

# Structural characterization of membrane proteins by solid-state NMR spectroscopy

Dissertation  
zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultäten  
der Georg-August-Universität zu Göttingen

vorgelegt von

Karsten Seidel  
aus Bremerhaven

Göttingen 2008

D7

Referent: Prof. Dr. Tim Salditt

Korreferent: Prof. Dr. Christian Griesinger

Tag der mündlichen Prüfung: 19. Februar 2008

IN PUBLICA COMMODA



*Dans les champs de l'observation le hasard  
ne favorise que les esprits préparés.*

—Louis Pasteur, Université de Lille, Dec 7, 1854



## Abstract

In this work, nuclear magnetic resonance spectroscopy was employed to characterize the three-dimensional atomic structure of proteins and other biomolecules in the solid phase. A method for obtaining sequential resonance assignments from ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectroscopy on a uniformly labeled protein under magic angle spinning was developed. This facilitated a largely complete assignment of a nano-crystalline form of the 76 residue protein ubiquitin. Comparison to results from other sample preparations indicates that chemical shifts variations are most likely to occur in protein regions that exhibit an enhanced degree of molecular mobility. The use of measuring proton–proton distances via rare-spin encoding was described on L-tyrosine-ethylester (TEE), and on ubiquitin. On the example of TEE, it was shown that internuclear distance measurements can informatively be complemented by techniques that probe molecular motion. On ubiquitin, the calculation of a structural model from rare-spin encoded proton–proton distances was assessed using a homology model. Finally,  $\text{Ca}^{2+}$ -ATPase-bound phospholamban (PLN) was investigated in a functional, non-crystalline lipid environment. Chemical shifts in all domains of AFA-PLN in the complex were assigned. The results were combined with other biophysical and biochemical data to calculate a three-dimensional model of the complex, which provides a structural basis for understanding the molecular interactions. These experimental studies were complemented by a statistical analysis of chemical shift predictions for proteins in the solid state with semi-empirical algorithms that are trained on solution-state NMR and X-ray diffraction data. The findings validate for the first time the applicability of such computer programs for solid-state NMR data, and highlight possible areas of improvement.





## Kurzzusammenfassung

In dieser Arbeit wurde die kernmagnetische Resonanzspektroskopie zur Bestimmung der atomaren Raumstruktur von Proteinen und anderen Biomolekülen in fester Phase angewandt. Eine neue Methode zur sequenzspezifischen Resonanzzuordnung durch ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) Korrelationspektroskopie wurde entwickelt. Unter Verwendung dieser wurde eine nahezu vollständige Zuordnung der Resonanzen einer nano-kristallinen Form des 76-Reste Proteins Ubiquitin erreicht. Ein Vergleich mit Resultaten anderer Probenpräparationen weist auf einen Zusammenhang der Variation chemischer Verschiebungen mit dem Grad der molekularen Mobilität hin. Der Nutzen indirekt gemessener Proton-Proton Abstände wurde anhand von L-Tyrosin-Ethylester (TEE) und Ubiquitin beschrieben. Am Beispiel von TEE wurde gezeigt, dass die Verwendung internuklearer Abstände durch Techniken zur Messung molekulare Bewegungen sinnvoll ergänzt werden kann. Im Falle von Ubiquitin wurden Strukturrechnungen auf Grundlage der indirekt gemessenen Proton-Proton Abstände durchgeführt, wobei der Analyse eine Homologie zu bestehenden Modellen zu Grunde gelegt wurde. Abschließend wurde an  $\text{Ca}^{2+}$ -ATPase gebundenes Phospholamban (PLN) in einer funktionellen, nicht-kristallinen Lipid-Umgebung untersucht. Dabei wurden chemische Verschiebungen in allen Domänen von AFA-PLN zugeordnet. Aus der Kombination der Resultate mit anderen biophysikalischen und biochemischen Daten konnte ein Modell des Komplexes berechnet werden, das als Grundlage für weitere Analysen der intermolekularen Wechselwirkungen dienen kann. Die experimentellen Studien wurden durch eine statistische Analyse der Vorhersagegenauigkeit chemischer Verschiebungen durch teil-empirische Algorithmen, die mit Daten aus NMR in Lösungen und Röntgenstrukturanalysen trainiert wurden, ergänzt. Die Ergebnisse sichern zum ersten Mal umfassend die Verwendung dieser Programme für chemische Verschiebungen in Festkörpern ab, und erlauben Rückschlüsse auf Verbesserungsmöglichkeiten.



## Related Publications

Parts of the work presented in this thesis are based on contributions to the following publications. I wish to express my kindest regards to all co-authors for these fruitful collaborations, and to all persons acknowledged therein.

1. A. Lange, K. Seidel, L. Verdier, S. Luca, and M. Baldus. Analysis of proton–proton transfer dynamics in rotating solids and their use for 3D structure determination. *Journal of the American Chemical Society*, 125:12640 (2003). [160]
2. K. Seidel, A. Lange, S. Becker, C. E. Hughes, H. Heise, and M. Baldus. Protein solid-state NMR resonance assignments from ( $^{13}\text{C},^{13}\text{C}$ ) correlation spectroscopy. *Physical Chemistry Chemical Physics*, 6:5090 (2004). [253]
3. K. Seidel, M. Etzkorn, L. Sonnenberg, C. Griesinger, A. Sebald, and M. Baldus. Studying molecular 3D structure and dynamics by high-resolution solid-state NMR: Application to L-tyrosine-ethylester. *Journal of Physical Chemistry A*, 109:2436 (2005). [255]
4. A. Lange, S. Becker, K. Seidel, K. Giller, O. Pongs, and M. Baldus. A concept for rapid protein-structure determination by solid-state NMR spectroscopy. *Angewandte Chemie International Edition English*, 44:2089 (2005). [161]
5. H. Heise, K. Seidel, M. Etzkorn, S. Becker, and M. Baldus. 3D NMR spectroscopy for resonance assignment and structure elucidation of proteins under MAS: novel pulse schemes and sensitivity considerations. *Journal of Magnetic Resonance*, 173:64 (2005). [113]
6. K. Seidel, M. Etzkorn, H. Heise, S. Becker, and M. Baldus. High-resolution solid-state NMR studies on uniformly [ $^{13}\text{C},^{15}\text{N}$ ]-labeled ubiquitin. *ChemBioChem*, 6:1638 (2005). [254]

7. O. C. Andronesi, S. Becker, K. Seidel, H. Heise, H. S. Young, and M. Baldus. Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 127:12965 (2005). [9]
8. M. Etzkorn, S. Martell, O. C. Andronesi, K. Seidel, M. Engelhard, and M. Baldus. Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angewandte Chemie International Edition English*, 46:459 (2007). [81]
9. C. Ader, R. Schneider, K. Seidel, M. Etzkorn, and M. Baldus. Magic-angle-spinning NMR spectroscopy applied to small molecules and peptides in lipid bilayers. *Biochemical Society Transactions*, 35:991 (2007). [1]
10. K. Seidel, O. C. Andronesi, J. Krebs, C. Griesinger, H. S. Young, S. Becker, and M. Baldus. Structural characterization of  $\text{Ca}^{2+}$ -ATPase-bound phospholamban in lipid bilayers by solid-state NMR spectroscopy. *Biochemistry*, 47:4369 (2007). [256]
11. O. C. Andronesi, M. von Bergen, J. Biernat, K. Seidel, C. Griesinger, E. Mandelkow, and M. Baldus. Characterization of Alzheimer's-like paired helical filaments from the core domain of tau protein using solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 130:5922 (2008). [10]

# Acknowledgements

*I am very grateful to:*

Prof. Dr. Salditt of the University of Göttingen for supervising my doctoral thesis on behalf of the Department of Physics, and for his open mind on interdisciplinary research.

Prof. Dr. Griesinger of the Max Planck Institute for Biophysical Chemistry for continuous support of my work, and for providing excellent research conditions in the MPI-BPC's Department of NMR-based Structural Biology.

Dr. Marc Baldus for supervising my scientific work with true commitment, the interesting projects he entrusted me with, a sound balance of challenge and support, and for sustaining collaborations with excellent molecular biologists.

Gitta Angerstein for being a helping hand in countless lab situations. All former and current members of the Baldus group for valuable scientific discussions, in particular Dr. Adam Lange, Manuel Etzkorn, Dr. Henrike Heise, Dr. Ovidiu Andronesi, Dr. Colan Hughes, Dr. Sorin Luca, Christian Ader, Robert Schneider and Lars Sonnenberg for very cooperative team work and a personable atmosphere.

All members in MPI-BPC's Department for NMR-based Structural Biology and Laboratory of Cellular Dynamics for the productive neighborhood.

Dr. Stefan Becker, Kamila Budzyn, Karin Giller and Petra Widawka (MPI-BPC) for their biochemical expertise, in particular for supplying us with many high quality samples, such as AGG, ubiquitin, and phospholamban/SERCA, which are so valuable to our work.

Dr. Howard Young and Dr. Catherine Treiber from the University of Alberta, Canada, for support on phospholamban studies by sample preparation protocols and scientific discussions. Dr. J. Krebs for scientific discussions on PLN/SERCA interactions.

Prof. Dr. Angelika Sebald (U Bayreuth) for providing tyrosine-ethylester samples, for lots of helpful discussions on the TEE project, and for writing the publication with us.

Prof. Dr. Martin Engelhard (MPI Dortmund) and his group, especially Eric Schiffer and Svetlana Martell, for sensory rhodopsin / transducer samples, from which I gained much experience in membrane protein spectroscopy, and for the sample of Figure 1.4.

Prof. Dr. Robert Griffin (MIT) for giving me the opportunity to gain some first-hand experience in DNP, Vikram Bajaj and Dr. Patrick van der Wel for working with me on the project,

Melody Mak, Alexander Barnes, Marc Caporini and Marvin Bayro for additional training with spectrometers and sample preparation, and the entire Griffin group for their extraordinarily frank atmosphere. (Results will not be discussed in context of this work, but I greatly profited from all the experience gained by working at the 'Bitter'.)

Dr. Christian Boehme for assistance with computer resources at Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen.

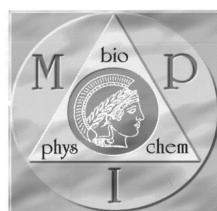
The Max Planck Society for continuous financial support, which allowed me to focus on my doctoral thesis work, the German Academic Exchange Service for personal funding during my research visit at MIT, EMBO for a summer course on biomolecular simulation, and EUROMAR '06 and ENC '07 for conference travel grants.

Furthermore, Prof. Dr. Stan Opella and Dr. Alex Nevzorov (UC San Diego) for raising my interest for biological solid-state NMR spectroscopy, Prof. Dr. Lu Sham (UCSD) for the quantum mechanics education, Prof. Dr. Montal (UCSD) for a lecture course that brought me to the field of structural biochemistry, Prof. Dr. Thomas Heimburg (U Göttingen) for his excellent courses on biophysics, and Eric and Isabelle Heinemeyer (U Göttingen) for stimulating discussions on graduate research in general.

Last but certainly not least, my parents Ingrid and Ralf for their constant support in any aim I have been pursuing, my brothers Jörn and Gunnar for being such refreshing characters, and Miriam for living with me through the highs and occasional lows that come along with scientific research.



Max-Planck-Gesellschaft  
zur Förderung  
der Wissenschaften e.V



Max-Planck-Institut für  
biophysikalische Chemie



Georg-August-Universität  
Göttingen

# Contents

<b>Abstract</b>	<b>vii</b>
<b>Kurzzusammenfassung</b>	<b>ix</b>
<b>Related Publications</b>	<b>xi</b>
<b>Acknowledgements</b>	<b>xiii</b>
<b>Contents</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Membrane Proteins . . . . .	1
1.2 Relation to Other Areas of Structural Biology . . . . .	4
1.3 Solid-State NMR Spectroscopy . . . . .	4
1.4 Scientific Challenges . . . . .	5
1.5 Outline of this Thesis . . . . .	6
<b>Principles, Technology and Methods</b>	<b>7</b>
<b>2 Concepts of NMR Spectroscopy</b>	<b>9</b>
2.1 Summary . . . . .	9
2.2 NMR Spectroscopy . . . . .	9
2.2.1 General Idea . . . . .	9
2.2.2 Fourier-Transform Spectroscopy . . . . .	10
2.2.3 Two-Dimensional Spectroscopy . . . . .	11
2.2.4 Signal Averaging . . . . .	11
2.3 Quantum-Mechanical Background . . . . .	12
2.3.1 Spin Dynamics . . . . .	12

2.3.2	Nuclear Spin Interactions . . . . .	14
2.4	Shift Referencing . . . . .	17
<b>3</b>	<b>Methods in SSNMR Spectroscopy</b>	<b>19</b>
3.1	Summary . . . . .	19
3.2	Decoupling of Nuclear Interactions . . . . .	19
3.2.1	Magic Angle Spinning . . . . .	19
3.2.2	RF Decoupling . . . . .	21
3.3	Selective Reintroduction of Nuclear Interactions . . . . .	22
3.3.1	Heteronuclear Polarization Transfer . . . . .	22
3.3.2	Homonuclear Polarization Transfer . . . . .	25
3.4	Spectral Editing by Isotope Labeling . . . . .	29
3.4.1	Uniform Sample Labeling . . . . .	30
3.4.2	Amino Acid Specific and Reverse Labeling . . . . .	31
<b>4</b>	<b>Molecular Modeling for SSNMR Structure Analysis</b>	<b>33</b>
4.1	Summary . . . . .	33
4.2	Introduction . . . . .	33
4.3	Target Function . . . . .	35
4.3.1	Empirical Force Field . . . . .	35
4.3.2	Common Approximations for Structure Calculation. . . . .	35
4.3.3	Experimental constraints . . . . .	37
4.4	Minimizing the Target Function . . . . .	38
4.4.1	Molecular Dynamics Simulated Annealing . . . . .	38
4.4.2	Conjugate Gradient Minimization . . . . .	41
4.4.3	Structure Assessment and Quality Control . . . . .	42
4.5	Structure Calculation Examples . . . . .	42
4.6	Molecular Docking . . . . .	44
4.6.1	<i>Ab Initio</i> Versus Experimental Data-Driven Docking . . . . .	44
4.6.2	High Ambiguity Data Driven Docking . . . . .	45
4.6.3	Protein Flexibility and Docking Schedule in HADDOCK . . . . .	46
	<b>New Techniques and Applications</b>	<b>49</b>
<b>5</b>	<b>On Solid-State NMR Shift Predictions</b>	<b>51</b>



5.1	Summary . . . . .	51
5.2	Introduction . . . . .	51
5.3	Methods . . . . .	54
5.4	Results and Discussion . . . . .	56
5.4.1	Overall Quality of Predictions . . . . .	56
5.4.2	Outliers . . . . .	58
5.4.3	Influence of Secondary Structure and Chemical Environment . . . . .	60
5.4.4	Residue-Specific Analysis . . . . .	61
5.4.5	Residue-Specific Effects from Secondary Structure and Environment . . . . .	62
5.4.6	Joining Forces . . . . .	63
5.5	Conclusions . . . . .	63
<b>6</b>	<b>Structure and Dynamics: l-Tyrosine-Ethylester</b>	<b>65</b>
6.1	Summary . . . . .	65
6.2	Introduction . . . . .	65
6.3	Material and Methods . . . . .	66
6.3.1	Sample Preparation . . . . .	66
6.3.2	NMR Experiments . . . . .	67
6.3.3	Quantum Mechanical Calculations . . . . .	68
6.3.4	Structure Calculation . . . . .	69
6.4	Results and Discussion . . . . .	70
6.4.1	Low-Temperature Analysis of TEE . . . . .	70
6.4.2	Solid-State NMR Analysis at Higher Temperature . . . . .	76
6.4.3	Molecular Dynamics As Seen by LG-CP . . . . .	79
6.4.4	Agreement of Model Structure Variations with Molecular Dynamics . . . . .	81
6.5	Conclusions . . . . .	82
<b>7</b>	<b>SSNMR Assignments from <math>^{13}\text{C}</math>-<math>^{13}\text{C}</math> spectroscopy</b>	<b>83</b>
7.1	Summary . . . . .	83
7.2	Introduction . . . . .	83
7.3	Numerical Analysis . . . . .	85
7.4	Spin Diffusion Under Weak Coupling Conditions . . . . .	87
7.5	Application to Uniformly Labeled Proteins . . . . .	89
7.6	Conclusion . . . . .	92

<b>8</b>	<b>Structural Characterization of Nanocrystalline Ubiquitin</b>	<b>93</b>
8.1	Summary . . . . .	93
8.2	Introduction . . . . .	93
8.3	Material and Methods . . . . .	95
8.3.1	Sample Preparation . . . . .	95
8.3.2	Solid-State NMR Experiments . . . . .	95
8.3.3	Data Analysis . . . . .	96
8.3.4	Structure Calculation . . . . .	96
8.4	Structural Characterization of UBI-P . . . . .	97
8.4.1	Resonance Assignments . . . . .	97
8.4.2	Secondary Structure Analysis . . . . .	102
8.4.3	Investigation of 3D Structure . . . . .	103
8.5	The Effect of Sample Preparation . . . . .	105
8.6	Conclusion . . . . .	110
<b>9</b>	<b>Ca<sup>2+</sup>-ATPase-bound Phospholamban in Lipid Bilayers</b>	<b>111</b>
9.1	Summary . . . . .	111
9.2	Introduction . . . . .	111
9.3	Material and Methods . . . . .	114
9.3.1	Sample Preparation . . . . .	114
9.3.2	Activity Assay . . . . .	114
9.3.3	Solid-State NMR Experiments and Analysis . . . . .	115
9.3.4	Modeling Using Flexible Protein-Protein Docking Simulation . . . . .	116
9.4	Results . . . . .	116
9.4.1	SsNMR Assignments and Secondary Structure of Mutant PLN . . . . .	116
9.4.2	Wild-Type PLN and PLN Mobility . . . . .	120
9.4.3	Docking Model . . . . .	122
9.5	Discussion . . . . .	124
9.5.1	Comparison to Previous Studies . . . . .	124
9.5.2	Implications for Complex Formation and Phosphorylation . . . . .	125
9.6	Conclusions . . . . .	126
<b>10</b>	<b>Summary and Outlook</b>	<b>127</b>

CONTENTS	xix
<b>Appendix</b>	<b>131</b>
<b>A Concepts of NMR Spectroscopy (Supplement)</b>	<b>133</b>
<b>B Molecular Modeling (Supplement)</b>	<b>137</b>
<b>C On SSNMR Chemical Shift Predictions (Supplement)</b>	<b>139</b>
<b>D l-Tyrosine-Ethylester (Supplement)</b>	<b>141</b>
<b>E SSNMR Assignments from <math>^{13}\text{C}</math>-<math>^{13}\text{C}</math> Spectroscopy (Supplement)</b>	<b>145</b>
<b>F Characterization of Nanocrystalline Ubiquitin (Supplement)</b>	<b>147</b>
<b>G <math>\text{Ca}^{2+}</math>-ATPase-bound PLN in Lipid Bilayers (Supplement)</b>	<b>151</b>
<b>H Primer on Amino Acids, Peptides and Proteins</b>	<b>161</b>
<b>I Pulse Programs</b>	<b>169</b>
<b>List of Symbols and Abbreviations</b>	<b>183</b>
<b>Bibliography</b>	<b>187</b>
<b>Lebenslauf</b>	<b>213</b>



# Chapter 1

## Introduction

*'... the opportunity cost incurred by the pharma industry from the absence of experimental structural information is immense.'*

—Peter Nollert and Lance Stewart, deCODE Genetics, Inc. [211]

### 1.1 Membrane Proteins

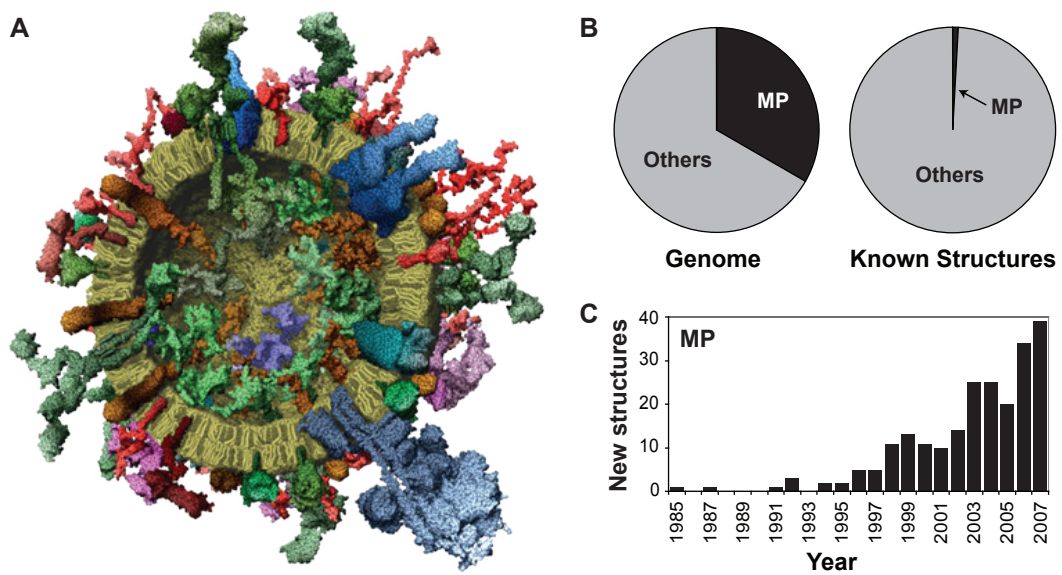
Life is based on the interaction of highly organized molecules that can fulfill specific functions. These macromolecules include, for example, DNA, RNA, Lipids, and Proteins. Proteins represent an extremely diverse class of organic macromolecules, contributing to intra- and extra-cellular structure, as well as basically all biological functions and interactions, such as signal transduction, transport, catalysis, immune defense or transcription [273]. Human DNA comprises about 25,000 protein-coding genes [128], which code for circa 100,000 proteins. Approximately one third of proteins coded for are membrane proteins [84] (Figure 1.1).

Knowledge of the three-dimensional (3D) structure and conformational dynamics of a protein can greatly facilitate understanding of biological function, prominent examples being the structural studies of Aquaporins and Potassium channels [66, 72]<sup>1</sup>. The first 3D structure of a protein, myoglobin, was solved in 1960 [144], while the first structure of a membrane protein was not solved until 25 years later [64]. Currently, the Protein Data Bank (PDB) [31] holds more than 47,000 experimentally determined protein structures with about one thousand different folds<sup>2</sup>, whereas, according to a recent survey [317], only 325 membrane proteins structures (of 144 unique proteins) are known (Figure 1.1B,C).

---

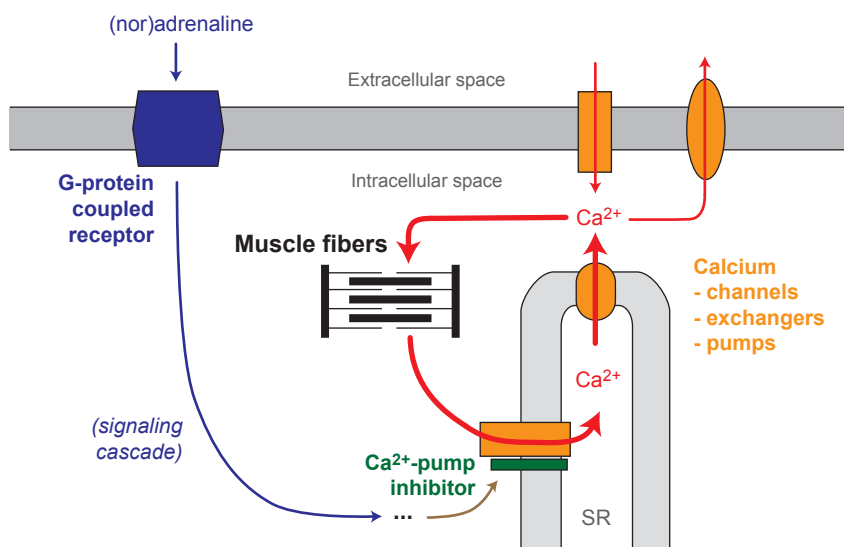
<sup>1</sup>Nobel price in chemistry, 2003

<sup>2</sup>See <http://www.pdb.org>; folds defined by SCOP



**Figure 1.1:** (A) Example of membrane proteins as found in a model of an average synaptic vesicle, sectioned in the middle. The model contains some macromolecular structures at near-atomic resolution, and homology models [280]. (B) Relative quantity of membrane proteins in the genome, and of membrane proteins with known structures with respect to protein models in the Protein Data Bank. (C) Histogramm of membrane protein structure determinations (unique structures) from data by the White lab [317] (data re-compiled). See ref. [38] for similar statistics of helix bundle membrane proteins.

This number is also in sharp contrast to the clinical relevance of membrane proteins. An example discussed in Chapter 9 of this thesis is the regulation of cardiac contractility, which is governed by a variety of membrane-integral ion pumps, exchangers and channels (Figure 1.2). Dysregulation of calcium cycling can cause heart disease, such as cardiac hypertrophy



**Figure 1.2:** Simplified scheme of cardiac  $\text{Ca}^{2+}$  flow (red) maintained by membrane-integral  $\text{Ca}^{2+}$  pumps, channels and exchangers (orange). The rate and intensity of heart strokes can be regulated through another membrane-integral protein, Phospholamban (green), based on an adrenaline-related signaling cascade that involves a  $\beta$ -adrenergic GPCR. See Chapter 9 for details.

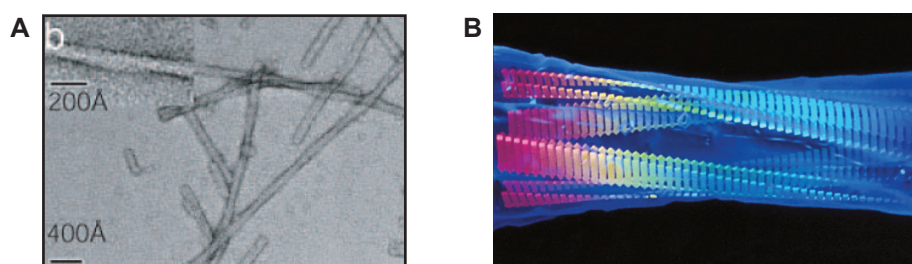
[187]. Understanding the molecular basis of interactions between the molecules involved may provide a ground for therapeutical interference with regulatory proteins in calcium cycling.

The poor accessibility of membrane proteins by structure determination methods is related to the amphipatic nature of these macromolecules: It can be extraordinarily difficult to grow large, well-diffracting crystals for high-resolution structural analysis by X-ray crystallography. Furthermore, the possibility of studying membrane proteins with NMR under fast-tumbling solution-state conditions is limited by the motional correlation time [49]. Here, the use of detergents that are employed to shield the hydrophobic interface of the protein increases the effective molecular weight, and may impose structural constraints on the molecule of interest. In this work, NMR methods developed for the study of solid-phase biomolecules (*solid-state NMR spectroscopy*) without long-range order will

be discussed, that do not require crystallization or motional averaging.

## 1.2 Relation to Other Areas of Structural Biology

Additional kinds of molecules that have chemical characteristics similar to membrane proteins may be studied using the very same methods and instruments as discussed in this work. Such systems include biomolecular aggregates in protein disorder diseases, like  $\alpha$ -synuclein (Parkinson's),  $A\beta$  and Tau (Alzheimer's), Amylin (type II diabetes) or PrP (prion encephalopathies, like Creutzfeldt–Jacob's and BSE) [257] which can form fibrils of  $\mu\text{m}$  length (Figure 1.3) . These fibrils commonly defy from crystallization under pathogenic



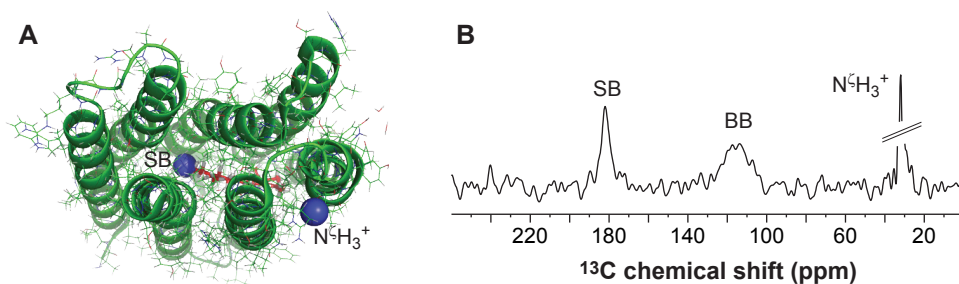
**Figure 1.3:** (A) Fibrils formed by the 140 amino acid protein  $\alpha$ -synuclein [111]. (B) Model derived from cryo-EM analysis (25 Å resolution) of fibrils grown from an SH3 domain, forming a hollow tube of 60 Å diameter [69, 136].

conditions. Likewise, other system of interest may be characterized, including peptidic and non-peptidic protein ligands that serve as templates or leads in drug discovery, see e.g. refs. [163, 185, 314].

## 1.3 Solid-State NMR Spectroscopy

*Solid-state NMR* (ssNMR) spectroscopy refers to the investigation of molecules by NMR spectroscopy under no or strongly reduced mobility. It is distinguished from solution-state NMR by the presence of generally strong anisotropic interactions [198]. These additional interactions lead to often poorly resolved spectra in particular for  $^1\text{H}$ , which form the basis of most solution-state NMR analyses [322]. Still, high-resolution conditions can be established for example by fast *magic angle* sample spinning and radio-frequency decoupling in high magnetic fields (Figure 1.4, and Chapter 3). Solid-state NMR spectroscopy can be applied to a large variety samples with low mobility, such as microcrystalline powders,





**Figure 1.4:** Example of a high-resolution solid-state NMR application. (A) Structure of Sensory Rhodopsin II (PDB ID: 1H68, from X-ray), lysine side-chain  $^{15}\text{N}$  shown as blue spheres (SB,  $\text{N}^{\epsilon}\text{H}_3^+$ ), one covalently bound to a retinal (red). (B) Chemical shifts of both  $^{15}\text{N}$  resonances in the presence of natural abundance backbone signal (BB).

gels, protein aggregates, or lipid bilayers, limited by spectral congestion and sensitivity, but not by demands on mobility or long-range order. Notably, ssNMR spectroscopy can provide information complementary to other techniques, such as solution-state NMR, X-ray, EM or FM methods, for example in the structural and dynamical characterization of molecular complexes with known or partially known structures of their constituents, see e.g. ref. [18], and Chapter 9.

## 1.4 Scientific Challenges

Correlation experiments involving carbon and nitrogen nuclei usually require samples enriched by the spin-1/2-isotopes  $^{13}\text{C}$  and  $^{15}\text{N}$ . In terms of sample preparation costs and measurement time, the exploitation of structural information from a few uniformly isotope enriched samples is preferred. In uniformly labeled samples, however, direct polarization transfer among carbons other than those in close proximity (bonded) is truncated by the interactions over short distances [115]. One objective of this work is to evaluate and improve methods for the structural characterization of uniformly  $^{13}\text{C}$  and  $^{15}\text{N}$  isotope-labeled polypeptides. As will be outlined below, this will involve (micro- and nano-)crystalline model peptides and proteins.

Another important aspect of this work is to demonstrate technical advancements and to gain insight of biological relevance by investigating a membrane protein in a functional environment. While a diversity of ssNMR methods have been developed for the determination of protein structure and dynamics, studies are often limited to favorable model systems, and applications to non-specifically labeled membrane proteins in non-

crystalline, lipid bilayers are still an exception. All studies shown here are accompanied by molecular modeling in order to assist the interpretation of results, and demonstrate the complementarity of the ssNMR data with results from other sources.

## 1.5 Outline of this Thesis

General concepts of nuclear magnetic resonance, solid-state NMR spectroscopy, and molecular modeling, as used in context of this work, are reviewed in Chapters 2–4. Before embarking on experimental studies, Chapter 5 evaluates the quality by which NMR chemical shifts of solid-phase polypeptides can be predicted from existing structural models using fast, semi-empirical computer programs. In Chapter 6, the determination of molecular structure by means of rare-spin encoded proton distance measurements, as introduced previously in refs. [80, 159, 160], is analyzed, with application to L-tyrosine-ethylester (TEE). The results are discussed in reference to molecular dynamics of TEE probed by  $^1\text{H}$ – $^{13}\text{C}$  polarization transfer. A prerequisite for structural studies on larger systems is identification of NMR resonances by sequential assignments. For this purpose, Chapter 7 presents an approach complementary to methods based on ( $^{15}\text{N}$ ,  $^{13}\text{C}$ )-transfer, improving sequential assignments by additionally using ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectroscopy. In Chapter 8, the previously discussed methods are then applied to fully labeled, nanocrystalline ubiquitin, a protein comprising 76 amino acids residues, and one of the first uniformly labeled proteins structurally characterized in the solid state. Aspects including sequential assignments, secondary structure, the use of proton-mediated rare-spin correlations and the influence of sample preparation on local structure are addressed. Chapter 9 concludes this work with an application to  $\text{Ca}^{2+}$ -ATPase bound phospholamban (PLN) in lipid bilayers. This study provides a secondary structure characterization from data directly obtained on PLN in the molecular complex, which could so far not be crystallized, and a consensus molecular model combining the solid-state NMR results with other biophysical and biochemical data.

# Principles, Technology and Methods



## Chapter 2

# Concepts of NMR Spectroscopy

## 2.1 Summary

Nuclear magnetic resonance (NMR) spectroscopy employs nuclear spin interactions that are sensitive to the local molecular environment. This chapter outlines fundamental concepts of NMR spectroscopy, followed by a summary of the quantum-mechanical treatment of spin systems.

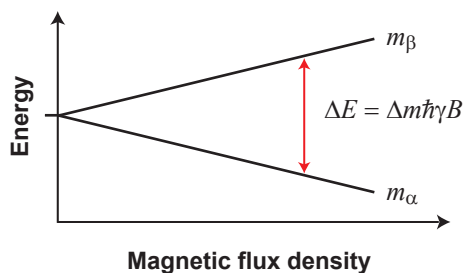
## 2.2 NMR Spectroscopy

### 2.2.1 General Idea

The fundamental idea of spectroscopy is to excite transitions between states with an energy difference  $\Delta E = \hbar\omega$ , and conclude from  $\omega$  to system properties. In magnetic resonance, the fundamental resonance condition is

$$\omega = -\gamma B, \tag{2.1}$$

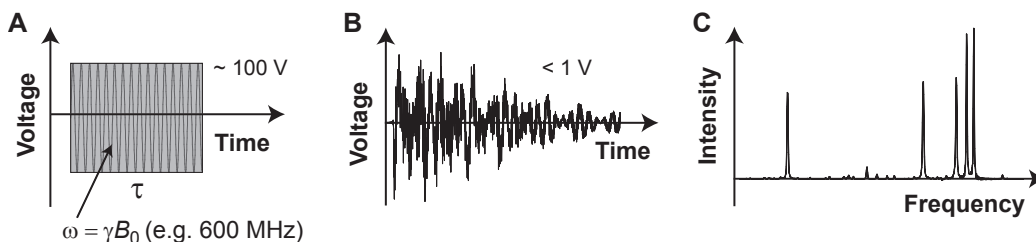
where,  $\gamma$  is the gyromagnetic ratio of a particle, and  $B$  is a magnetic field that splits up the otherwise degenerate energy levels (Figure 2.1). For nuclei in magnetic fields in the order of several Tesla,  $\omega$  takes values in the MHz (i.e., radio frequency, RF) range. The first successful NMR experiments on the resonant absorption of continuous RF radiation by bulk matter were performed in 1945/46 by Purcell, Torey and Pound on  $^1\text{H}$  of solid Paraffin [235], and by Bloch, Hansen and Packard on  $^1\text{H}$  of water [34]. The influence of the chemical environment on the resonance condition [68, 234] was originally probed using a field sweep of  $B$ .



**Figure 2.1:** Energy splitting  $\Delta E$  of spin eigenstates in a magnetic field  $B$ .

### 2.2.2 Fourier-Transform Spectroscopy

Attempts were undertaken to construct multichannel spectrometers that would allow simultaneous measurement of numerous points of a CW-spectrum, leading to an increase in performance proportional to the number of channels. The huge instrumental effort was circumvented by Ernst and Anderson, who introduced a spectrometer that recorded the linear response of magnetization (*free induction decay*, FID) in a sample subjected to a short RF pulse [77] (Figure 2.2). According to Fourier's theorem, a pulse can be



**Figure 2.2:** (A) Scheme of a radio-frequency (RF) pulse used to excite spin transitions for nuclei with a resonance frequency  $\omega$ . The magnitude may be one hundred volts or more. (B) Response of  $^{13}\text{C}$  magnetization after disturbing magnetization in thermal equilibrium with a  $90^\circ$  pulse, in the order of typically less than one volt. (C) Spectrum resulting from B after Fourier Transform.

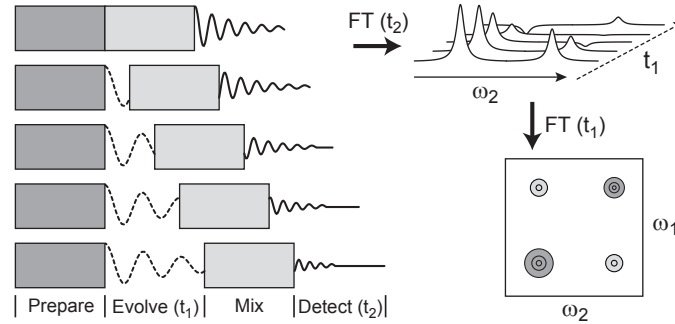
considered a multi frequency source, exciting a whole range of energy transitions at once. The frequency spectrum of the response can be obtained from Fourier transform

$$S(\omega) \propto \int_0^\infty s(t)e^{-i\omega t} dt \quad (2.2)$$

The practical realization of Fourier spectroscopy in the late-1960s was greatly facilitated by inexpensive computers that could perform fast Fourier transform (FFT) algorithms.

### 2.2.3 Two-Dimensional Spectroscopy

In 1971, Jeener proposed a two-pulse experiment in time-domain by using two independent precession periods during which coherences can evolve [135] (Figure 2.3). The indirect



**Figure 2.3:** General scheme for a 2D FT NMR experiment. With respect to a 1D experiment, a 2D experiment has an additional (indirect) period of evolution ( $t_1$ ), which is incremented while the rest of the experiment remains unchanged. The evolution during  $t_1$  modulates the recorded signal (FID,  $t_2$ ), which, after FT along  $t_1$ , yields an indirect spectral dimension  $\omega_1$ .

evolution is monitored through phase and amplitude of magnetization at the beginning of the detection time. The response signal  $s(t_1, t_2)$  of the system is then subjected to two-dimensional Fourier transformation,

$$S(\omega_1, \omega_2) \propto \int_0^\infty s(t_1, t_2) e^{-i\omega_1 t_1} e^{-i\omega_2 t_2} dt_1 dt_2, \quad (2.3)$$

for frequency analysis. Cross correlations of different resonance frequencies reveal interactions between different nuclei during mixing.

### 2.2.4 Signal Averaging

Summation of a repeated number of identical experiments leads to an increase of signal ( $S$ ) proportional to the number  $n$  of transients recorded, while noise ( $N$ ) increases proportional to  $\sqrt{n}$ , according to the law of large numbers, resulting in an increase of

$$S/N \propto \frac{n}{\sqrt{n}} = \sqrt{n}. \quad (2.4)$$

Depending on the sample, multidimensional ssNMR experiments may require several days of signal averaging, putting high demands on the stability of the NMR experiment, as regarding, for example,  $\mathbf{B}_0$  field stability, power sources, and the cooling system.

## 2.3 Quantum-Mechanical Background

The theoretical treatment of spin systems in NMR not only provides a basis for the description of NMR phenomena, but also allows for a numerical simulation of polarization transfer in small spin systems for the design and analysis of NMR experiments [17, 262, 271, 307].

### 2.3.1 Spin Dynamics

#### Density Operator

The state of a physical system is characterized by a density operator  $\rho$ . A mixed ensemble, composed of  $k$  pure ensembles with relative populations  $w_k$ , has the density operator  $\rho = \sum_k w_k |\psi^{(k)}\rangle\langle\psi^{(k)}|$  [245]. Measurement of an observable  $A$  yields  $\langle\langle A \rangle\rangle = \text{tr}(\rho A)$ . Time evolution of the density operator is governed by the Hamilton operator  $H$ , according to the Liouville–von Neumann equation

$$\frac{d}{dt}\rho(t) = \frac{1}{i\hbar}[H(t), \rho(t)]. \quad (2.5)$$

A suitable Hamiltonian for the formal solution can be obtained from Average Hamiltonian Theory (AHT) [315] (Section A.2.1). The initial density operator for a Hamiltonian with  $H/kT \ll 1$  in thermal equilibrium is  $\rho_0 \propto H$ .

#### Spin

The spin  $\mathbf{I} = (I_x, I_y, I_z)$  of a particle fulfills (and can be defined by) the commutator relation

$$[I_i, I_j] = i\hbar\epsilon_{ijk}I_k, \quad (2.6)$$

where  $\epsilon_{ijk}$  denotes the Levi-Civita-Tensor. The spin state is described by two quantum numbers  $\ell$  and  $m$  with  $I^2|\ell m\rangle = \hbar^2\ell(\ell + 1)|\ell m\rangle$  and  $I_z|\ell m\rangle = \hbar m|\ell m\rangle$ . Choice of the z-direction is arbitrary.

#### Rotations

Any rotation by an angle  $\phi$  about a unit vector  $\hat{\mathbf{n}}$  can be described by an operator

$$R(\phi, \hat{\mathbf{n}}) \equiv e^{-i\phi\frac{\hat{\mathbf{n}}\cdot\mathbf{J}}{\hbar}}, \quad (2.7)$$



with a *generator*  $G = \frac{\hat{\mathbf{n}} \cdot \mathbf{J}}{\hbar}$  containing the angular momentum operator  $\mathbf{J}$  [245]. For example,  $\phi = \omega t$ ,  $\hat{\mathbf{n}} = \hat{\mathbf{z}}$ , and  $\mathbf{J} = \mathbf{I}$  describe a rotation about the  $z$ -axis with angular frequency  $\omega$  in spin space.

### Rotating Frame

For a system with a Hamiltonian which is composed of a static and a time-dependent part,  $H(t) = H_0 + V(t)$ , one may define a *rotating frame*, where the time evolution of the density operator is determined by [198]

$$\frac{d}{dt}\rho' = \frac{1}{i\hbar}[V', \rho']. \quad (2.8)$$

Here,  $V' = R^\dagger V R$  and  $\rho' = R^\dagger \rho R$ , where the rotation operator is  $R = e^{-iH_0 t}$  with a generator with  $\mathbf{I}$  and  $H_0 = \omega I_z$ . In effect, the Hamiltonian is transformed into a *rotating frame*, rotating with an angular frequency  $\omega_0$  with respect to the *laboratory frame* [198].

### Tensors Rotations

The elements of a matrix  $D(R)$  representing a rotation operator  $R(\phi, \hat{\mathbf{n}})$  in the basis of  $I_z$  are given by

$$D_{m'm}^{(\ell)}(R) = \langle \ell m' | e^{-i\phi \hat{\mathbf{n}} \cdot \mathbf{J} / \hbar} | \ell m \rangle, \quad (2.9)$$

providing, in block-diagonal form, the  $(2\ell + 1)$ -dimensional *irreducible representation* of  $R$ . An arbitrary rotation between two coordinate systems in three-dimensional space can be described by *Euler rotations*  $R(\alpha, \beta, \gamma) \equiv R(\gamma, \hat{\mathbf{z}}')R(\beta, \hat{\mathbf{y}}')R(\alpha, \hat{\mathbf{z}})$ , which allow  $D_{m'm}^{(\ell)}$  to be conveniently expressed by *reduced Wigner elements* [245] (Section A.2.4). A tensor  $F$  can be decomposed into components  $F_m^{(\ell)}$  (or  $F_{\ell m}$ ) acting on the same subspaces as the rotation operator in irreducible representation (*irreducible spherical tensor*), transforming under rotations according to [245]

$$F_m'^{(\ell)} = D(R)F_m^{(\ell)}D^\dagger(R) = \sum_{m'=-\ell}^{\ell} F_{m'}^{(\ell)}D_{m'm}^{(\ell)}(R). \quad (2.10)$$

Calculations involving magic angle spinning (MAS, Chapter 3) experiments may require rotations between the *principle axis system* (PAS) of the tensor, the *molecular frame* (MOL), the *rotor-fixed axis system* (RAS) and the *laboratory frame* (LAB) [198].

### Structure of the NMR Hamiltonian

The Hamiltonian of a spin system can be written as the generalized scalar product of the spatial tensor  $A$  bearing the geometrical information and the spin tensor  $T$ ,

$$H = \sum_{\ell=0}^N A_{\ell} \cdot T_{\ell} = \sum_{\ell=0}^N \sum_{m=-\ell}^{\ell} (-1)^m A_{\ell m} T_{\ell, -m}, \quad (2.11)$$

where  $N$  denotes the maximum rank of the tensors,  $\ell$  the rank and  $m$  the component [198]. The Hamiltonian is invariant under rotation (rank 0). In the principle axis system, the components

$$A_{00}^{\text{PAS}} = -\frac{1}{\sqrt{3}} \text{tr} A, \quad A_{20}^{\text{PAS}} = \sqrt{\frac{3}{2}} \delta \quad \text{and} \quad A_{2\pm 2}^{\text{PAS}} = \frac{1}{2} \delta \eta \quad (2.12)$$

relate to the *isotropic average*  $a_{\text{iso}} = \frac{1}{3} \text{tr} A$ , *anisotropy*  $\delta = A_{zz} - a_{\text{iso}}$  and *asymmetry*  $\eta = (A_{yy} - A_{xx})/\delta$  the tensor  $A$  in cartesian coordinates [198]. Here,  $A_{xx}$ ,  $A_{yy}$  and  $A_{zz}$  denote the eigenvalues of  $A$ , by convention ordered by their magnitude of difference to  $a_{\text{iso}}$ . The form of the tensor in a new reference frame is then obtained from an Euler rotation as given in (2.10).

### 2.3.2 Nuclear Spin Interactions

The (relevant part of the) Hamiltonian of a system composed of  $N$  spin-1/2 particles in NMR is given by

$$H_{\text{NMR}} = \sum_{i=1}^N \left( H_Z^{(i)} + H_{\text{RF}}^{(i)} + H_{\text{CS}}^{(i)} \right) + \sum_{i=2}^N \sum_{j=1}^i \left( H_{\text{D}}^{(ij)} + H_{\text{J}}^{(ij)} \right), \quad (2.13)$$

with terms relating to the Zeeman (Z), radio-frequency (RF), chemical shielding (CS), *through-space* dipole–dipole (D) and *through-bond* spin–spin (J) interactions. Other interactions, such as the nuclear quadrupol coupling, hyperfine interaction and Knight shift [261] are not considered in context of this work.

#### **Zeeman Interaction**

The *Zeeman interaction* describes the coupling of nuclear spin  $\mathbf{I}$  to an external magnetic field  $\mathbf{B}_0$ . By convention, the  $z$ -axis of the system is taken along  $\mathbf{B}_0$ , therefore

$$H_Z = -\gamma I_z B_0 = \omega_0 I_z \quad (2.14)$$

with Larmor frequency  $\omega_0 = -\gamma B_0$ . For a nucleus with magnetic quantum numbers  $m = \pm 1/2$ , the nuclear Zeeman energy levels are separated by a magnitude  $\Delta E = \hbar\gamma B_0$  (Figure 2.1).

### Interaction with RF fields

Nuclear spins are manipulated by RF pulses generated by a current oscillating with frequency  $\omega_{\text{RF}}$  through a coil around the sample. The corresponding field  $\mathbf{B}_{\text{RF}}(t)$  leads to a term similar to the Zeeman Hamiltonian, differing by magnitude and time-dependence,

$$H_{\text{RF}}(t) = -\gamma \mathbf{I} \mathbf{B}_{\text{RF}}(t), \quad (2.15)$$

where  $|\mathbf{B}_{\text{RF}}| \ll B_0$ . Usually, the RF field is linearly polarized with a phase  $\phi$ . In a frame rotating with  $\omega_{\text{RF}}$  about  $\hat{z}$ , denoting  $\gamma|\mathbf{B}_{\text{RF}}| = \gamma B_1 = \omega_1$ , the RF Hamiltonian is

$$H'_{\text{RF}} = -\omega_1 (I_x \cos \phi + I_y \sin \phi). \quad (2.16)$$

The Larmor frequency of a spin in the rotating frame, i.e. the offset with respect to the carrier, is denoted by  $\Omega = \omega_0 - \omega_{\text{RF}}$  [78].

### Chemical Shielding

Chemical shielding is a modification to the local magnetic field of a nucleus, which is proportional to the external magnetic field  $\mathbf{B}_{\text{ext}} = \mathbf{B}_0 + \mathbf{B}_{\text{RF}}$ . It can be expressed by a shielding tensor  $\sigma$  that modifies the Zeeman interaction,

$$H_{\text{CS}} = \gamma \mathbf{I} \sigma \mathbf{B}_0. \quad (2.17)$$

In *secular* (i.e., *high field*) approximation [198], the Hamiltonian in the principle axis system of the shielding tensor is given by

$$H_{\text{CS}} = A_{00}^{\text{PAS}} T_{00} + A_{20}^{\text{PAS}} T_{20} = (\sigma_{\text{iso}} + \delta) I_z B_0. \quad (2.18)$$

with isotropic average  $\sigma_{\text{iso}}$  and anisotropy  $\delta$  as defined in Section 2.3.1. Notably,  $\sigma$  has anti-symmetric components  $A_{1m}$ , but the corresponding spin components  $T_{1m}$  do not commute with  $H_Z$ . Transformations of the secular chemical shielding Hamiltonian into other spatial reference frames affect only the anisotropic component  $A_{20}$ , see ref. [198] for an explicit expression of the component accounting for molecular, rotor and laboratory frame. In powder samples, the angular dependence of the anisotropic term leads to a characteristic *powder lineshape* (Figure 3.1C).

Distance	$d_{12}$ (rad · kHz)					
	H–H	H–C	H–N	C–C	C–N	N–N
1 Å	120	30	12	8	3	1.2
2 Å	15.0	3.8	1.5	0.9	0.4	0.2
3 Å	4.40	1.10	0.40	0.28	0.10	0.04

**Table 2.1:** Dipolar coupling constant  $d_{12}$  for selected nuclei.

## Dipolar Coupling

The Hamiltonian describing the magnetic dipole–dipole interaction (dipolar coupling) of two spins is of the form

$$H_D = \mathbf{I}_1 D \mathbf{I}_2 \quad (2.19)$$

where  $\mathbf{I}_1$  and  $\mathbf{I}_2$  are two nuclear spins, and  $D$  is the dipolar coupling tensor.  $D$  is traceless and axially symmetric, and can hence be averaged out by MAS (Chapter 3). In the principle axis system of dipolar coupling tensor, the Hamiltonian is

$$H_D = A_{20}^{\text{PAS}} T_{20} = d_{12} \cdot (\mathbf{I}_1 \mathbf{I}_2 - 3I_{1z} I_{2z}). \quad (2.20)$$

where  $d_{12}$  is the dipolar coupling constant,  $d_{12} = \frac{\mu_0 \hbar \gamma_1 \gamma_2}{4\pi r_{12}^3}$  (SI unit rad·Hz), with an internuclear distance  $r_{12}$ . The spin product may be expanded in terms of raising and lowering operators (*dipolar alphabet*, Appendix Section A.2.5). In secular (high-field) approximation, the Hamiltonian of the dipolar interaction for a *homonuclear* spin pair in an arbitrary frame of reference is

$$H'_D = A'_{20} T_{20} = d_{12} \cdot \frac{1 - 3 \cos^2 \theta}{2} \left( 2I_{1z} I_{2z} - \frac{1}{2} (I_1^+ I_2^- + I_1^- I_2^+) \right), \quad (2.21)$$

where  $\theta$  is the angle between by  $\mathbf{r}_{12}$  and  $\mathbf{B}_0$ . For a *heteronuclear* spin pair, the second term can be neglected since it does not commute with the heteronuclear Zeeman Hamiltonian, only  $I_{1z} I_{2z}$  is retained. The dipolar interaction is a dominant quantity in solid-state NMR. Values of  $d_{12}$  for common spin pairs are given in Table 2.1.

## J-Coupling

The *J-coupling* is an indirect interaction between nuclear spins, mediated by electron pairs involved in chemical bonding. The J-coupling Hamiltonian is of the form

$$H_J = \mathbf{I}_1 J \mathbf{I}_2, \quad (2.22)$$

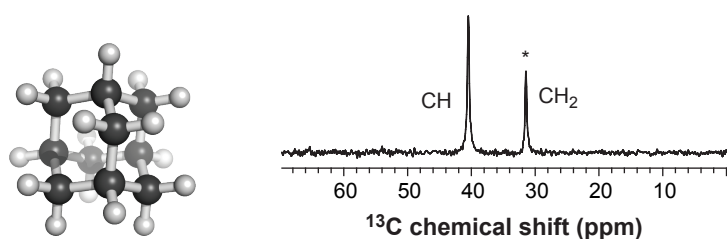
where  $\mathbf{I}_1$  and  $\mathbf{I}_2$  are two nuclear spins and  $J$  is the coupling tensor. By magnitude, the  $J$  is significantly smaller than the dipolar coupling,  $|J| \ll |D|$ . The anisotropic component of  $J$  can usually be neglected,  $J = 2\pi J_{\text{iso}}$ . The isotropic component of  $J$  (*scalar coupling*) can be used to mediate polarization transfer under both anisotropic (solid-state) and isotropic (solution-state) NMR conditions. Correlation experiments based on scalar transfer (Chapter 3) can, for example, be used to probe mobile components of otherwise 'solid' proteins (Chapter 9). Also,  $J$ -couplings may have to be taken into account in sensitive simulations of spin systems, such as for the derivation of long C–C distances in R<sup>2</sup>TR experiments of uniformly labeled compounds [266] (Chapter 6).

## 2.4 Shift Referencing

It is convenient to express the resonance frequency  $\omega$  (or  $\Omega$ ) by a *chemical shift*  $\delta(\omega, \omega_{\text{ref}})$  relative to a reference compound,

$$\delta(\omega, \omega_{\text{ref}}) \equiv \frac{\omega - \omega_{\text{ref}}}{\omega_{\text{ref}}} \cdot 10^6, \quad (2.23)$$

stating the result in *parts per million* (ppm) [108]. According to IUPAC, tetramethylsilane ((CH<sub>3</sub>)<sub>4</sub>Si, TMS) is the recommended standard reference for <sup>1</sup>H and <sup>13</sup>C chemical shifts, and nitromethane (CH<sub>3</sub>NO<sub>2</sub>) for <sup>15</sup>N, with the isotropic resonance lines of each spin species set to zero ppm [108]. Inconsistencies from using other reference compounds exist at a considerable degree [320, 329]. In this work, adamantane is used as external secondary standard with a chemical shift of 31.47 ppm (= 29.45<sup>TMS</sup> ppm + 2.01<sup>ΔDSS</sup> ppm) for the CH<sub>2</sub> resonance (Figure 2.4) to be compatible with internal DSS in D<sub>2</sub>O [202]. For <sup>15</sup>N,



**Figure 2.4:** Adamantane is a suitable external secondary standard for solid-state NMR shift referencing. Here, the CH<sub>2</sub> resonance (\*) is set to 31.47 ppm.

since liquid CH<sub>3</sub>NO<sub>2</sub> is not a convenient reference when low temperatures or temperature effects due to MAS are to be accounted for, it was practicable to use Ala–Gly–Gly as

secondary reference, as in ref. [183]. With a  $^{15}\text{NH}_3$  resonance at 38.9 ppm, this is in line with (liquid)  $\text{CH}_3\text{NO}_2$  or other secondary references, such as ammonium chloride [195].

## Chapter 3

# Methods in NMR Spectroscopy of Biological Solids

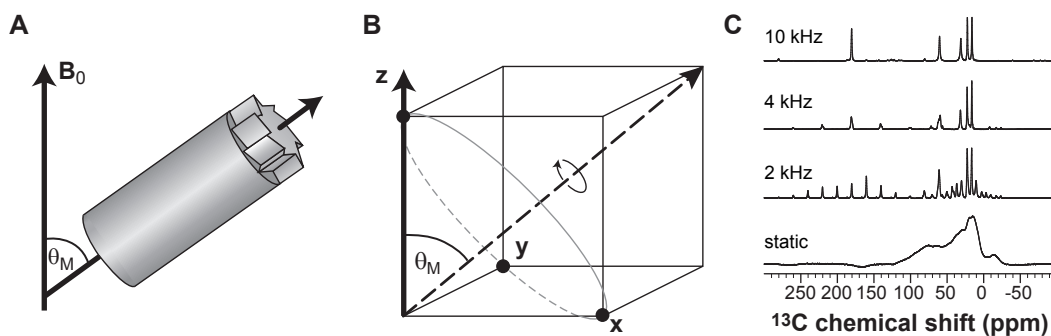
### 3.1 Summary

The efficient characterization of uniformly labeled biomolecules by NMR mandates tools for handling nuclear spins in a large number of different chemical sites. In the following, methods in solid-state NMR spectroscopy *used in context of this work* will be described that allow for (i) establishing conditions for high-resolution spectroscopy by *reducing* nuclear interactions, and (ii) measuring molecular characteristics by selectively *re-introducing* nuclear interactions. A selection of multidimensional correlation schemes that employ a combination of pulse blocks will be discussed, focusing on their application to biomolecular systems. A new concept for obtaining sequential resonance assignments from ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectroscopy will be described here briefly, and presented in greater detail in Chapter 7.

### 3.2 Decoupling of Nuclear Interactions

#### 3.2.1 Magic Angle Spinning

Mechanical sample rotation renders the space part of an NMR Hamiltonian time-dependent [4]. Sample spinning can be achieved by driving an air-bearing, impeller-equipped cylindrical sample container (rotor) by a tangential gas flow (Figure 3.1A). A rotation about the angle of the cubic space diagonal, the *magic angle* (MA) (Figure 3.1B), allows for a removal of rank 2 tensors, provided sufficiently fast spinning [5, 181]. The effect of magic angle spinning (MAS) on the spectrum of uniformly  $^{13}\text{C}$ -labeled valine is shown in Figure 3.1C. The time dependence introduced by rotating the sample by an angle  $\phi(t) = \omega_{\text{R}}t$  about the rotor axis results in a dependence of the Hamiltonian term  $A_{\ell m}^{\text{RAS}}$



**Figure 3.1:** Magic angle sample spinning. (A) Sample rotor with spinning axis aligned at the magic angle  $\theta_M$  with respect to the external magnetic field  $B_0$ . (B) Cubic symmetry at the magic angle  $\theta_M$ . (C) Dependence of MAS side band patterns on spinning frequency for uniformly  $^{13}\text{C}$ -labeled valine. Figure A adapted from ref. [265], Figure B adapted from ref. [166], spectra of Figure C courtesy of D. Nand.

from the rotor-fixed axis system (RAS)<sup>1</sup> (Section 2.3) of

$$A_{\ell m}^{\text{LAB}}(t) = \sum_{m'=-\ell}^{+\ell} A_{\ell m'}^{\text{RAS}} e^{-im\omega_R t} d_{m'm}^{(\ell)}(\theta_M) \quad (3.1)$$

in the laboratory frame (LAB), according to an Euler rotation  $R(\omega t, \theta_M, 0)$ , where  $\theta_M$  is the magic angle [198]. Hence, in addition to the isotropic resonance line, *side bands* appear at multiples of  $\omega_R$  and  $2\omega_R$  ( $\ell = 2$ ) away from the center band in the spectrum at moderate MAS frequencies, Figure 3.1.

The centrifugal force acting on the rotor walls, and stability requirements on the sample rotation, make use of high-performance materials, such as  $\text{ZrO}_2$  ceramics, mandatory. Permeability to electromagnetic irradiation at optical light microwave frequencies can be improved by using rotors made of sapphire crystals, for example.

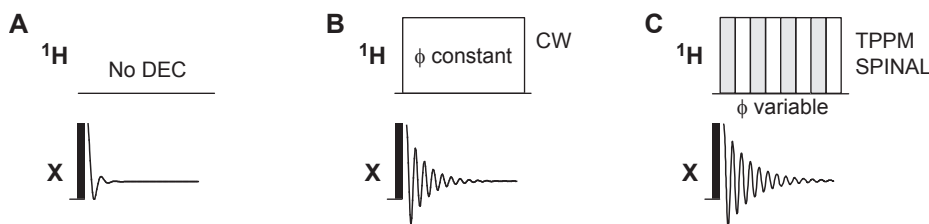
Applications presented in this thesis employ MAS at frequencies up to 13 kHz (4 mm rotor diameter), which is sufficient to decouple, for example, heteronuclear  $^{15}\text{N}$ - $^{13}\text{C}$  interactions and homonuclear  $^{13}\text{C}$ - $^{13}\text{C}$  interactions. Decoupling dipolar interactions of rare spins with  $^1\text{H}$ , and of homonuclear  $^1\text{H}$ - $^1\text{H}$  interactions in the absence of significant molecular mobility, additionally requires strong RF decoupling schemes.

<sup>1</sup> $A_{\ell m'}^{\text{RAS}}$  may be constructed from the principle axis system by Euler rotations to the molecular frame and rotor-fixed axis system



### 3.2.2 RF Decoupling

A substantial improvement of spectral resolution in biomolecular solids can be achieved by the decoupling of strong, heteronuclear dipolar interactions of  $^1\text{H}$  with other nuclei. Most simple,  $^1\text{H}$  spins are locked by continuous-wave (CW) irradiation (Figure 3.2B). In applications shown here, two-pulse phase modulation (TPPM) [28] or small phase incremental alternation (SPINAL) [88] was used, which perform superior to CW. Both sequences employ a continuous train of high-power RF pulses with alternating phases (Figure 3.2C), which are repeatedly increased and decreased for SPINAL.



**Figure 3.2:** High-power pulse schemes for heteronuclear decoupling. (A) No heteronuclear RF decoupling. (B) CW irradiation on the  $^1\text{H}$  channel employs a fixed RF phase, while (C) TPPM and SPINAL decoupling consist of pulses generating about  $165^\circ$  spin rotations with alternating phases. For example, TPPM-15 uses  $\pm 15^\circ$  phases, and SPINAL-64 cycles among  $\pm 10^\circ$ ,  $\pm 15^\circ$ , and  $\pm 20^\circ$  phases.

In the presence of motion on the  $\mu\text{s}$  time scale or faster, high-power RF decoupling is not required, as spatial averaging already facilitates a substantial reduction of dipolar interactions [9]. Instead, GARP decoupling [258] at low RF fields ( $\omega_1 \approx 10$  kHz) can be employed [9]. GARP (Globally optimized Alternating-phase Rectangular Pulse) decoupling is a computer-optimized pulse sequence comprised of variable-length pulses with alternating ( $0^\circ, 180^\circ$ ) phases.

In addition to dedicated heteronuclear decoupling schemes, symmetry-based RF sequences for homonuclear recoupling can have a decoupling effect on heteronuclear interactions, and high-resolution conditions can even be established completely without  $^1\text{H}$  decoupling [122]. Likewise, in the presence of  $\mu\text{s}$  time scale motion, we found that the symmetry-based TOBSY sequence, providing homonuclear  $^{13}\text{C}$  scalar recoupling, performs equally well without proton RF decoupling as with SPINAL or GARP [9].

Decoupling of *homonuclear*  $^1\text{H}$ - $^1\text{H}$  dipolar interactions can be achieved in the solid state for example by frequency-switched Lee-Goldburg (FSLG) [32, 169] decoupling, which

can be applied under cross-polarization conditions to a rare-spin (see below). Other approaches using numerically optimized, (quasi-)continuously modulated phase schemes, are given in ref. [244].

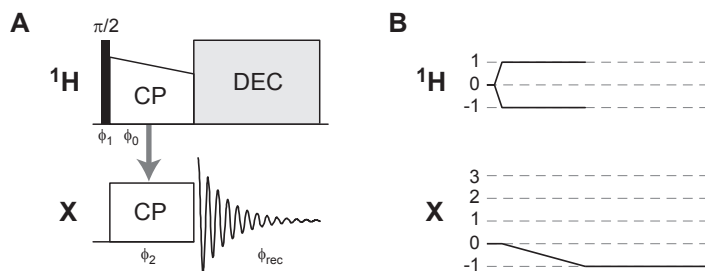
Furthermore, rendering a coupling ineffective may often also simply be achieved by refocusing of magnetization that is dephasing under an interaction, such as the refocusing of scalar couplings using an isolated  $180^\circ$  pulse [106], as employed below in INEPT-type experiments.

### 3.3 Selective Reintroduction of Nuclear Interactions

#### 3.3.1 Heteronuclear Polarization Transfer

##### Cross-Polarization

Cross-polarization (CP) establishes dipolar coupling-mediated polarization transfer by frequency matching in the doubly-rotating frame created by two transversal spin-lock RF fields (Figure 3.3) for the nuclei of interest [109, 231]. The RF fields must fulfill the



**Figure 3.3:** (A) Pulse sequence with  $^1\text{H}$ -X cross-polarization. Using a ramped RF field on one channel provides improved matching of the Hartmann-Hahn condition. Suitable pulse phases are e.g.  $\phi_1 = 90^\circ$ ,  $\phi_0 = \phi_2 = 0^\circ$ . (B) Corresponding coherence order diagram [36]. Coherence order  $-1$  is detected. The CP sequence example in Appendix Chapter I uses a two-step  $180^\circ$ -increment ( $^1\text{H}$ ) and four-step  $90^\circ$ -increment ( $^{13}\text{C}$ ) phase cycle.

*Hartmann-Hahn condition,*

$$\omega_{1,1} \pm \omega_{1,2} = n\omega_R, \quad n = \pm 1, \pm 2 \quad (3.2)$$

where  $\omega_{1,1}$  and  $\omega_{1,2}$  correspond to the strengths of the RF ( $B_1$ ) fields for the two spin species. CP can be employed for signal enhancement of low- $\gamma$  nuclei by polarization

transfer from high- $\gamma$  nuclei [231], such as  $^1\text{H}$ - $^{15}\text{N}$  or  $^1\text{H}$ - $^{13}\text{C}$ , and for spectral editing. CP is a central building block in many solid-state NMR experiments [21].

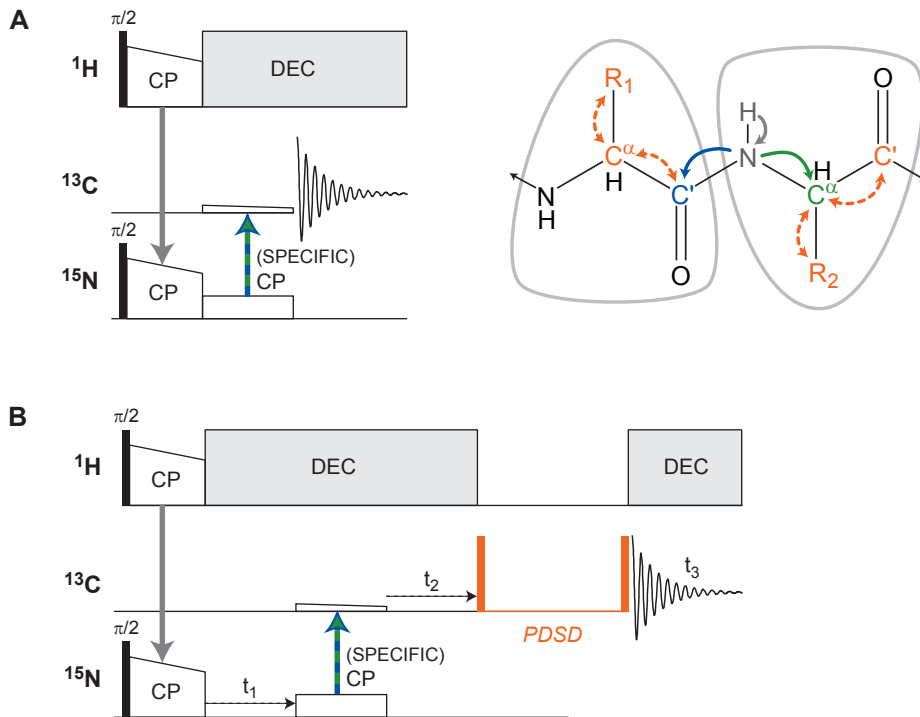
Suggested pulse and receiver phases, and phase cycle [36] for the CP experiment (and other pulse programs) are given in Appendix Chapter I.

### Double CP and SPECIFIC CP

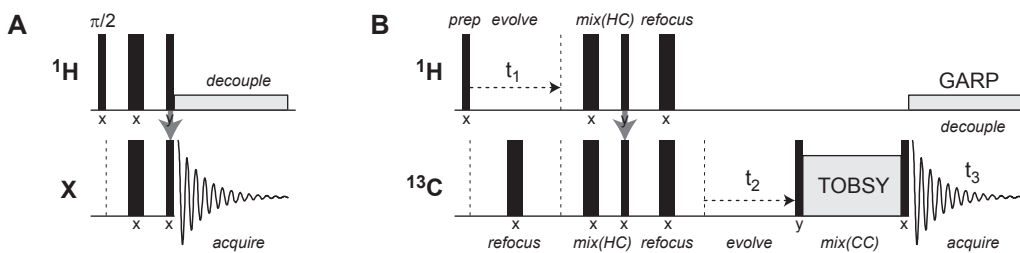
Spectral editing is a useful tool in NMR spectroscopy. Consecutive  $^1\text{H}$ - $^{15}\text{N}$  and  $^{15}\text{N}$ - $^{13}\text{C}$  transfers, for example, allow for exploring spin networks that are found along the backbone and some side chains of polypeptides. The corresponding RF scheme is termed *double CP* [246] (Figure 3.4). During  $^{15}\text{N}$ - $^{13}\text{C}$  CP, proton decoupling is required. In particular, RF fields can be chosen such that only N-C $^\alpha$  or N-C' transfer is established (SPECIFIC CP) [25], allowing to distinguish between intra-residue or inter-residue transfer for spectral resonance assignments in proteins.

### INEPT

INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) is a pulse sequence for heteronuclear polarization transfer based on  $J$ -coupling (Figure 3.5A) [203]. The transversal evolution during the preparation of desirable spin states on  $^1\text{H}$  is in the order of a few milliseconds, hence, transversal magnetization of  $^1\text{H}$  in completely rigid molecular domains will have decayed, in the absence of decoupling sequences, by the time polarization is to be transferred to the rare spin. In addition to polarization transfer and signal enhancement, this provides a mobility filter solid-state NMR, by which dynamics can be probed [9]. Multiple INEPT transfers can be used to obtain sequential resonance assignments from mobile protein domains, that otherwise elude from studies using dipolar based transfer mechanisms [9]. INEPT may be combined with  $J$ -coupling based homonuclear  $^{13}\text{C}$  mixing, such as TOBSY [23] under MAS, when residual anisotropic interactions are present (Figure 3.5B). Alternatively, INEPT may be used with homonuclear proton decoupling sequences where little or no mobility is present, for example when studying natural abundance organic solids, see ref. [243] for a review.



**Figure 3.4:** (A) HNC double cross-polarization signal enhancement and spectral editing. The second CP, from  $^{15}\text{N}$  to  $^{13}\text{C}$ , is performed under  $^1\text{H}$  decoupling, and can be rendered band-selective (SPECIFIC CP), for example to aliphatic resonances (i.e.,  $\text{C}^\alpha$ ) or carbonyls ( $\text{C}'$ ). (B) NCA and NCO transfers are often combined with homonuclear  $^{13}\text{C}$  mixing for an improved identification of spin systems. As an example, PDSM mixing is indicated, and the sequence is shown here as a 3D correlation experiment, as used for the assignment of ubiquitin (Chapter 8). Phases and phase cycles given in Appendix Chapter I.



**Figure 3.5:** (A) Pulse sequence for INEPT transfer, shown in context of signal acquisition on the X channel under  $^1\text{H}$  decoupling (e.g., GARP). (B) Refocused INEPT followed by TOBSY mixing, with two indirect evolution periods (on  $^1\text{H}$  and X). Phases for the  $\pi/2$ - and  $\pi$ -pulses are indicated by  $x$  ( $0^\circ$ ) and  $y$  ( $90^\circ$ ). Phase cycle given in Appendix Chapter I.

### 3.3.2 Homonuclear Polarization Transfer

#### Longitudinal $^{13}\text{C}$ mixing

A variety of methods for homonuclear polarization transfer involve an alignment of magnetization along  $\mathbf{B}_0$  by a  $90^\circ$  pulse from the transversal plane (*longitudinal mixing*) (Figure 3.6). Homonuclear  $^{13}\text{C}$  dipolar recoupling over distances up to several Å (Figure 3.6A) can be established by rotational resonance ( $\text{R}^2$ ), i.e. by exactly matching the chemical shift difference between two nuclei of interest with multiples  $n$  of the MAS rate,  $n \cdot \omega_R = |\Omega_1 - \Omega_2|$ , where  $\Omega_i$  are the resonance offsets from the carrier [6, 56, 154, 186, 236]. A detailed theoretical analysis for an isolated two-spin-1/2 system is presented in ref. [173]. Variations of the  $\text{R}^2$  concept allow for measuring exact internuclear distances in uniformly labeled compounds [239, 266]. For example, a weak RF field on the  $^{13}\text{C}$  channel may be used to shift the recoupling condition away from  $n \cdot \omega_R$  ( $\text{R}^2$  in the Tilted Rotating frame,  $\text{R}^2\text{TR}$ ) [281], avoiding complications with spinning side bands that would complicate the analysis of the experiment [266].

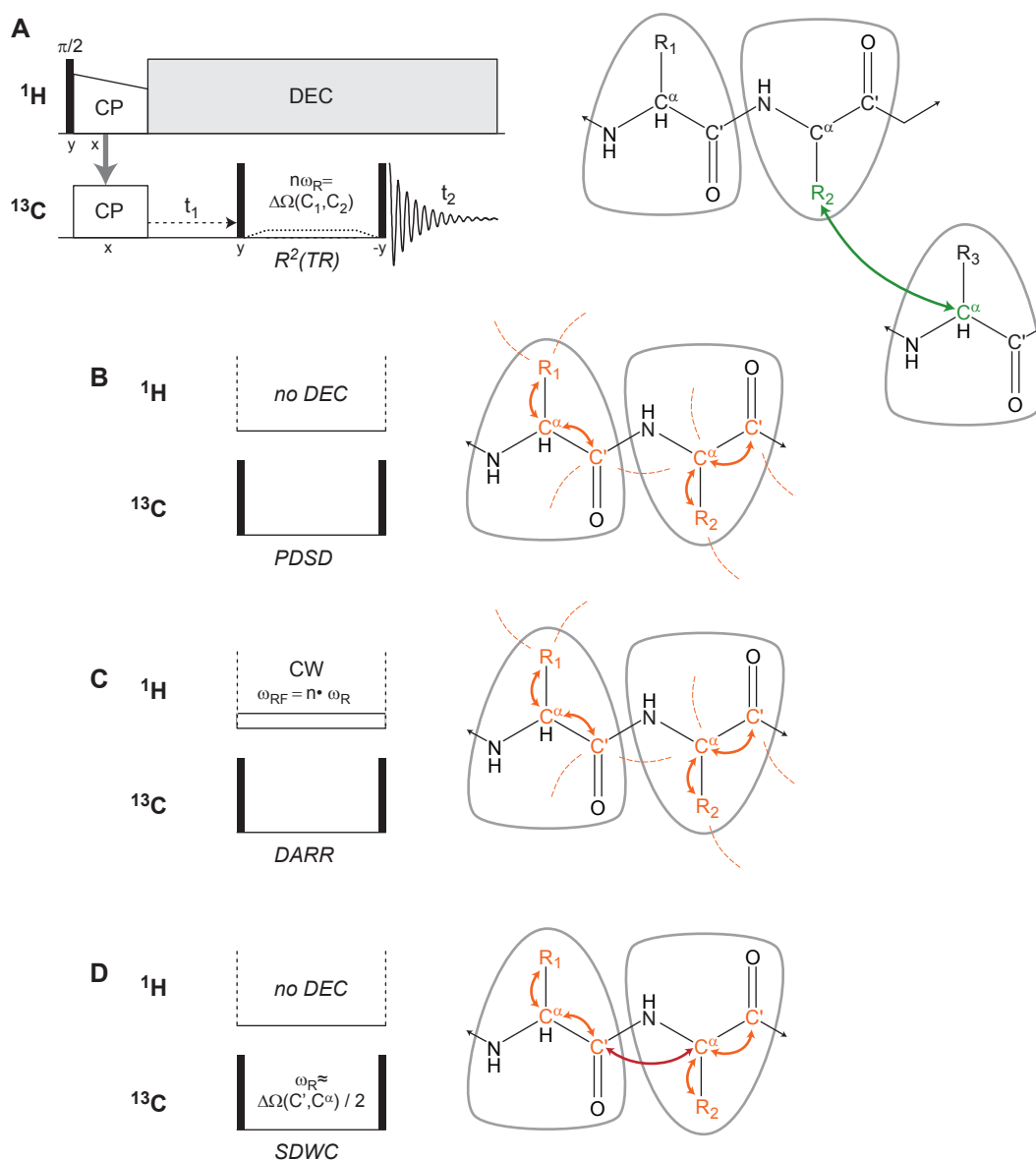
In the absence of  $^1\text{H}$  decoupling during longitudinal mixing, broadband  $^{13}\text{C}$ - $^{13}\text{C}$  proton-driven spin diffusion (PDS) is established [154, 275, 276]. In uniformly labeled compounds, however, efficient transfer is restricted to nuclei in close spatial proximity (Figure 3.6B). PDS is hence particularly useful in generating a large number of intra-residue ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlations (Chapters 6–9), as polarization is relayed among a homonuclear spin species, here  $^{13}\text{C}$ , but not along  $^{15}\text{N}$  of the peptide bond.<sup>2</sup>

The rate of PDS polarization transfer may be enhanced by increasing  $^1\text{H}$ - $^{13}\text{C}$  couplings, which broadens the spectral overlap between different nuclei (DARR) [282] (Figure 3.6C). For short and medium spin diffusion times (tens of ms) away from rotational resonance, no qualitative difference with respect to PDS was observed throughout this work; it may be favorable, however, using DARR to enhance transfer at high MAS rates, or for  $\text{C}'$ - $\text{C}^\alpha$  transfer. The effect of very long (hundreds of ms) PDS or DARR are so far not well understood. Recent simulations [98] and experiments (unpublished), however, indicate that sizable long-distance polarization transfer may be obtained and suffice to refine protein structures.

Furthermore, as discussed in greater detail in Chapter 7 and ref. [253], weak coupling conditions between the ( $\text{C}'$ ,  $\text{C}^\alpha$ ) moiety near the rotational resonance conditions, prefer-

---

<sup>2</sup>The transfer is created through space, but in terms of molecular topology, it is convenient to speak of bonds, e.g. one-bond transfer, two-bond transfer.

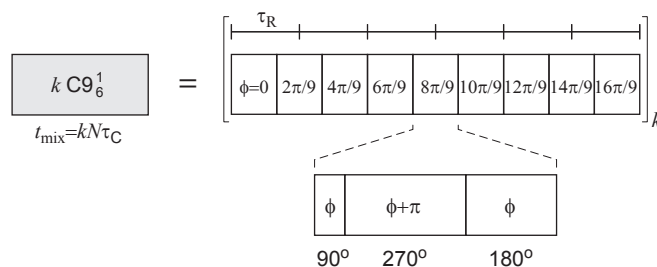


**Figure 3.6:** Selected schemes for longitudinal (0Q)  $^{13}\text{C}$  mixing. (A) General scheme for a 2D ( $^{13}\text{C}, ^{13}\text{C}$ ) correlation spectrum. Here,  $R^2$  (or  $R^2\text{TR}$ ) mixing is used, matching the shift difference of the nuclei of interest with the MAS rate under  $^1\text{H}$  decoupling. Long-distance  $^{13}\text{C}$ - $^{13}\text{C}$  polarization transfer for the selected spin pair is established. (B) Proton-Driven  $^{13}\text{C}$ - $^{13}\text{C}$  spin diffusion. Polarization transfer is largely confined to adjacent  $^{13}\text{C}$  nuclei, or to the closely connected  $^{13}\text{C}$  spin network within one residue, when spin diffusion delays of up to approximately 100 ms are employed. (C) PDSD transfer can be enhanced by (RF) rotary resonance on  $^1\text{H}$ . Again, transfer other than between adjacent  $^{13}\text{C}$  nuclei is marginal. (D) 'Clean' sequential correlations are obtained when the MAS rate is near the  $n = 2$  rotational resonance for the ( $\text{C}', \text{C}^\alpha$ ) moiety, and PDSD is active. Suggested phases indicated in (A). Phase cycle given in Appendix Chapter I.

ably ( $n = 2$ ), in combination with PDS may be used to generate sequential ( $^{13}\text{C}, ^{13}\text{C}$ ) correlations, in addition to intra-residue correlations (Figure 3.6D).

### Sequences Employing Symmetry Design Principles

A range of rotor-synchronized pulse sequences, such as REDOR [101], DRAMA [298], HORROR [208], C7 [170], POST-C7 [116], SPC5 [117] or TOBSY [23], can be constructed from symmetry rules that are known as  $C$ - or  $R$ -symmetry classes [172]. According to the selection rules of classes denoted by  $CN_n^\nu$  or  $RN_n^\nu$ , the first-order average Hamiltonian for an interaction vanishes,  $H_{\ell m \lambda \mu}^{(1)} = 0$ , when  $mn - \mu\nu = N$  or  $NZ_\lambda/2$ , respectively. For the  $R$  class, selection rules for the second order average Hamiltonian apply. The selection rules may be used to construct rotor-synchronized pulse sequences with desirable properties. As an example, the pulse sequence of a TOBSY (total through-bond correlation spectroscopy) experiment with POST- $C9_6^1$  element is illustrated in Figure 3.7<sup>3</sup>. It consists of a basic



**Figure 3.7:** Example of a homonuclear  $^{13}\text{C}$  mixing scheme that can be constructed from symmetry rules: TOBSY mixing (J-induced) using a POST-element with  $C9_6^1$  symmetry.

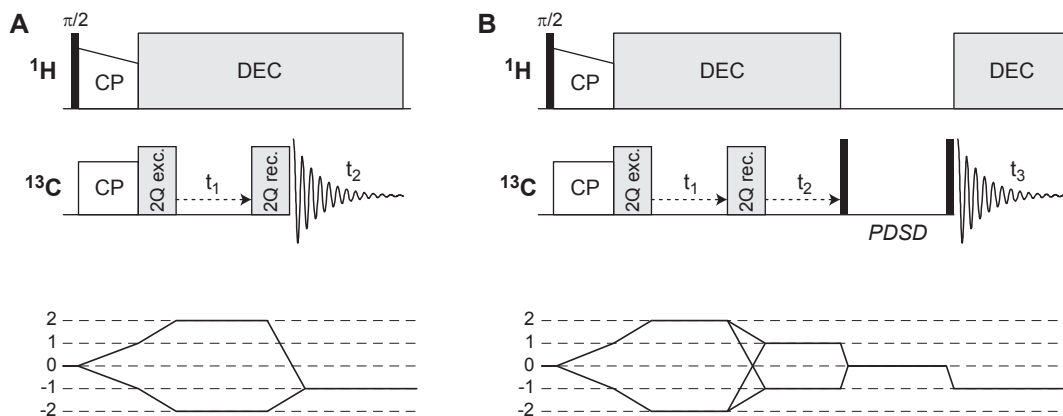
element, which may be a compound pulse, that is repeated  $N$  times during  $n$  rotor cycles, with a phase increment that is given by the selection rules. The width of the basic element, and the relation between number of elements and rotor cycles, determine a ratio between RF nutation frequency and rotor spinning frequency, hence the pulse sequence must be chosen in accordance with these. A detailed overview of both  $C$  and  $R$  sequences is given in ref. [172].

### Double-quantum recoupling

The generation of  $^{13}\text{C}$ - $^{13}\text{C}$  double-quantum (2Q) coherences can be useful for spectral simplification with respect to spin-diffusion type spectra. For example, a 2D ( $^{13}\text{C}, ^{13}\text{C}$ )

<sup>3</sup>The original TOBSY pulse sequence uses an  $R$ -element [23].

double-quantum/single-quantum correlation spectrum does not contain auto-correlations (a 'diagonal'). Furthermore, 2Q mixing based on strong dipolar couplings is limited to nuclei in close spatial proximity, i.e. bonded  $^{13}\text{C}$ , which may simplify spectral analysis, and it can be used to suppress natural abundance background signal. The pulse scheme for a multidimensional 2Q (or multi-quantum) / single quantum correlation spectrum no longer follows the standard scheme given in Section 2.2.3. Instead, excitation and reconversion of 2Q coherences frame the 2Q-evolution period (Figure 3.8A). The excited coherences need



**Figure 3.8:** (A) General scheme for a 2D 2Q/1Q correlation spectrum, with diagram indicating the coherence pathway that needs to be selected by the pulse and receiver phase cycle. In contrast to the standard 2D pulse scheme, mixing and evolution are no separate entities. (B) Three-dimensional homonuclear  $^{13}\text{C}$  correlation sequence combining 2Q-evolution with a longitudinal mixing scheme, such as PSD. Phases and phase cycles for selecting the indicated coherence pathways given in Appendix Chapter I, and for (B) also in ref. [113].

to be selected by an adequate phase cycle [296]. Pulse sequences that generate adequate 2Q Hamiltonian terms can, for example, be constructed from symmetry principles (Section 3.3.2).

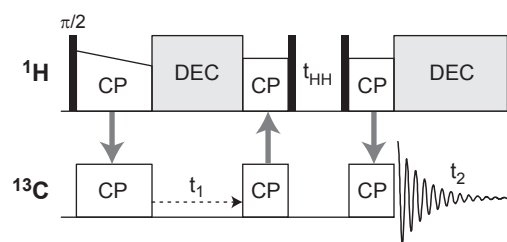
An example of a new homonuclear 3D correlation scheme, which we found to be particularly useful (Chapter 8), is the 3D 2Q/1Q/1Q  $^{13}\text{C}$  correlation scheme shown in Figure 3.8B. It provides the advantages and flexibility of 0Q mixing (e.g., PSD, SDWC, CHHC), but removes the cubic diagonal through 2Q evolution in  $t_1$  [113].

### Rare-Spin Detected Proton–Proton Mixing

While 2Q recoupling and PSD-type mixing schemes can be used to establish ( $^{13}\text{C},^{13}\text{C}$ ) correlations along the covalent structure of molecules, no true broadband  $^{13}\text{C}$  recoupling



sequences exist that allow to measure a large number of long-range internuclear distances in a uniformly labeled molecule with high accuracy, in a single experiment, such as done in solution with NOESY. Also,  $^1\text{H}$ - $^1\text{H}$  distances are most informative on the 3D fold of a protein [19, 20, 158]. A large number of correlations with sensitive distance-dependent buildup characteristics can be obtained by encoding homonuclear  $^1\text{H}$ - $^1\text{H}$  mixing in high resolution  $^{13}\text{C}$  or  $^{15}\text{N}$  dimensions [159, 160]. The 2D correlation scheme for CHHC mixing is shown in Figure 3.9. The use of CHHC correlations in structure determination will



**Figure 3.9:** CHHC sequence using longitudinal  $^1\text{H}$ - $^1\text{H}$  mixing for protons. The sequence may also be employed with other proton mixing schemes, or as NHHC correlation. Phases and phase cycle given in Appendix Chapter I.

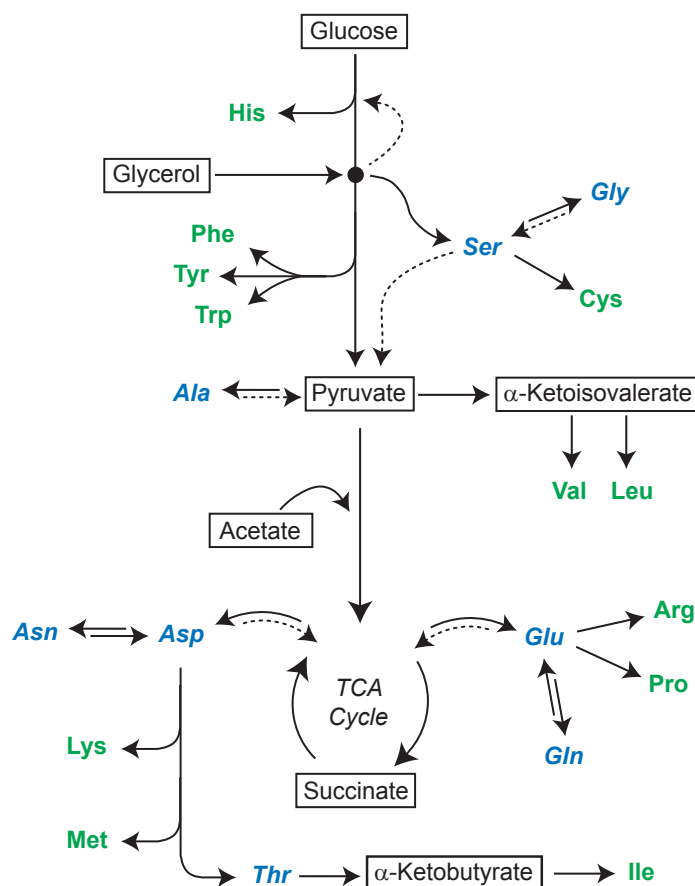
be assessed in Chapters 6 and 8. Similar to a CHHC correlation as shown in Figure 3.9, NHHC correlations can be obtained, starting with a  $^1\text{H}$ - $^{15}\text{N}$  CP. Combination with direct, intra-residue carbon transfer in higher-dimensional sequences, such as CHHCC and NHHCC can be used to additionally resolve chemical shifts [113].

### 3.4 Spectral Editing by Isotope Labeling

The low natural abundance of  $^{13}\text{C}$  and  $^{15}\text{N}$  and low accessibility of  $^1\text{H}$  spectra mandates isotope enrichment for most spectroscopic applications in solid-state NMR. In addition to signal enhancement, isotope labeling schemes can be used for editing NMR spectra, for example by generally reducing spectral crowding, enhancing signal from protein regions of interest, or influencing polarization transfer characteristics.

Site-specific isotope labeling, for example in solid-phase peptide synthesis, may be used to study local chemical sites of interest. While this is spectroscopically less demanding in terms of spectral congestion and resonance assignment, it implies the (often costly) preparation of multiple specifically labeled samples for the derivation of multiple structural parameters throughout an entire protein or a protein domain. In contrast, proteins studies

in this work performed using bacterial expression systems allow to introduce abundant isotope labels by means of amino acid metabolism [27] (Figure 3.10).



**Figure 3.10:** Schematic amino-acid biosynthesis pathways in *E. coli* [27]. Commonly used carbon sources are denoted by boxes. Amino acids that can be added to the expression medium without significant cross-labeling are shown in green, others in blue italics. For the sake of clarity, the scheme, adapted from ref. [176], is strongly simplified. Abbreviations: TCA, Tricarboxylic acid cycle. See Appendix Chapter H for amino acid abbreviations.

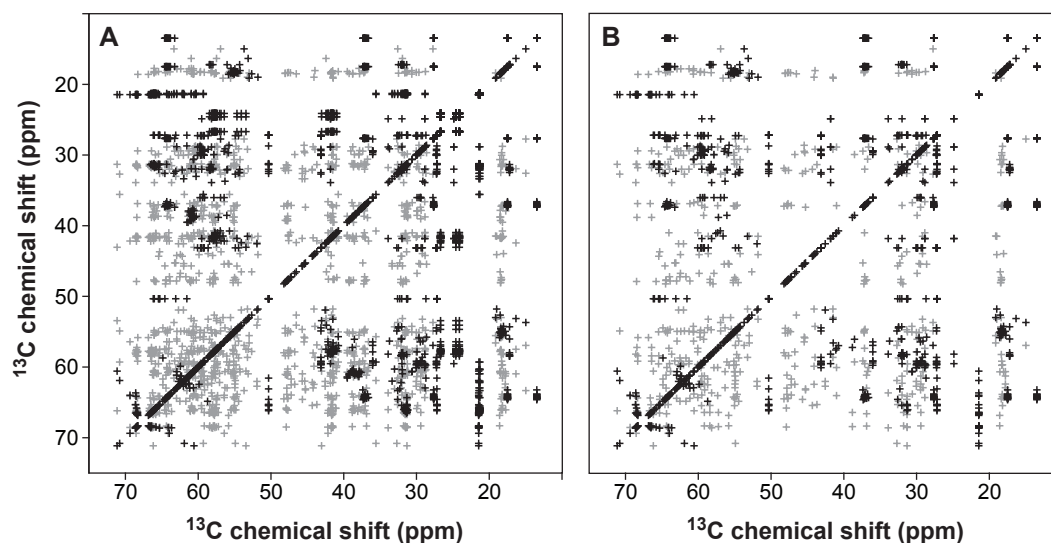
### 3.4.1 Uniform Sample Labeling

Using bacterial protein expression with amino acid synthesis pathways as summarized in Figure 3.10, several recent ssNMR studies [46, 328] employed  $^{13}\text{C}$  spin dilution with the labeling scheme by LeMaster and Kushlan [171], through (1,3)- $^{13}\text{C}$ - or 2- $^{13}\text{C}$ -glycerol as isotope source. In contrast, applications shown here use  $^{13}\text{C}_6$ -glucose that effect a (100%) uniform  $^{13}\text{C}$  labeling of amino acids. Additionally,  $^{15}\text{N}$  isotope labeling is achieved by

adding  $^{15}\text{N-NH}_4\text{Cl}$  to the expression medium. Several other common sources of  $^{13}\text{C}$  are denoted by boxes in Figure 3.10.

### 3.4.2 Amino Acid Specific and Reverse Labeling

Amino-acid type dependent sample labeling based on bacterial expression can be influenced by adding amino acids with natural  $^{13}\text{C}$  abundance (0.01%) to the isotope-labeled expression medium (*reverse labeling*), or by directly adding  $^{13}\text{C}$ -labeled amino acids to the natural isotope abundance expression medium (*amino acid specific labeling*). In both cases, it is essential to take into account cross-labeling by metabolic conversion of one amino acid into another, or through the activity of transaminase [27, 176]. A virtue of reverse labeling is the low cost of natural abundance amino acids that are added to the expression medium. As illustrated in Figure 3.11 and recently demonstrated experimentally in Chapter 9 and elsewhere [81, 111], solid-state NMR applications can significantly profit from reverse labeling schemes. Furthermore, amino acid specific labeling was used to identify resonances in a molecular complex that were difficult to identify in the fully and reverse-labeled samples (Chapter 9).



**Figure 3.11:** Effect of reverse labeling on spectral crowding, showing ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation patterns predicted from a crystal structure of Sensory Rhodopsin II (PDB ID: 1H68). Intra-residue correlations are shown in black, sequential correlations are given in gray. Backbone chemical shifts were calculated by SHIFTX [207] (Chapter 5), and side chain chemical shifts were retrieved from the BioMagResBank [300]. (A) Correlations to be expected for a uniformly  $^{13}\text{C}$  labeled sample. (B) Correlations remaining for the uniformly  $^{13}\text{C}$  labeled sample with LVFY reverse labeling (U\LVFY). Reduction of spectral crowding, likewise seen in ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) spectra, facilitated the spectral assignment of a large number of resonances in this seven transmembrane helix receptor [81].

## Chapter 4

# Molecular Modeling for SSNMR Structure Analysis

### 4.1 Summary

A brief overview of methods for molecular modeling is given, followed by a detailed description of biomolecular model building from experimental solid-state NMR data. Here, the focus lies on finding data-compliant structures by energy minimization through simulated annealing. Examples using different kinds of ssNMR data are discussed in reference to the Crystallography and NMR System (CNS) software suite, which was used for applications in Section 4.5 and Chapters 6–8. Furthermore, data-driven biomolecular docking using CNS within HADDOCK protocols, as in Chapter 9, is introduced.

### 4.2 Introduction

In nature, the *native* (biologically active and thermodynamically most stable [12]) conformation of a protein is completely determined by the amino acid sequence and the chemical environment [11]. In theory, the same information could be obtained by solving the time-dependent Schrödinger equation for the entire molecular system, but by all practical means, this is an impossible task [303]. Using a mean field approximation for the electronic interactions, *ab initio* calculations based on density functional theory (DFT) allow for calculating physical properties of small molecular systems, such as polarizability, or chemical shifts. Modern DFT algorithms [86, 251, 295] and computers can handle systems in the order of 100–1000 atoms, however, computational demands, accuracy and interactions with the chemical environment circumvent the *de novo* calculation of arbitrary molecular structures.

Generally, the covalent structure of proteins and other bio-molecules is well understood.

Separating nuclear and electronic Hamiltonians (Born-Oppenheimer approximation), motions of nuclei can be described classically, and the potential energy function that describes the electronic influence on motions of the nuclei can be approximated empirically [303]. This facilitates the calculation of *classical molecular dynamics* (MD) trajectories, that can yield the structure of small molecules and polypeptides. On the other hand, long-range interactions and entropic effects, that eventually lead to the native fold of a molecule, cannot yet efficiently be described in a quality that allows for the *de novo* calculation of large molecular structures from theory alone. Hence, a particular strength of MD currently is the refinement of molecular structures, and the detailed analysis of dynamics and conformational changes which are observed by experiments.

The conformational space that is accessible to a molecule according to theory can be restricted by experimental constraints. One usually distinguishes between *structure calculation* as the *de novo* calculation of a 3D structure from an unfolded starting structure, and *structure refinement* as the improvement of a given, well-defined fold. In addition to this, *molecular docking* describes the process of finding the structure of a complex formed by two or more molecules of known structure for the isolated constituents. The docking application in Chapter 9 demonstrates that borders between structure calculation, docking and refinement can be vanishingly small, and that a wide range of experimental information may be taken into account.

Once a large number of protein structures are known, protein structures can also be constructed from structural homology to other proteins. This approach is based on the observation that proteins with the same fold (structure homologs) exhibit a high degree of sequence identity (sequence homologs). It is a result of the evolutionary optimization of proteins, and as little as 30% sequence identity may suffice to build a correctly folded structure. On the other hand, the mutation of a few amino acids may completely alter a protein fold, and it is important to realize that homology modeling can only suggest protein folds that will then have to be verified by other means. In general, the concept of homology enters protein structure investigations in manifold ways. For example, screening intermolecular contacts by amino acid mutagenesis relies on the preservation of molecular folds, and scoring functions that predict the secondary structure of proteins based on chemical shifts (e.g., TALOS [58], PREDITOR [30]) or vice versa (Chapter 5) incorporate terms for sequence homology.

## 4.3 Target Function

A suitable target function for structure calculation and refinement is the overall energy of the molecule in its chemical environment,  $E_{\text{chem}}$ , combined with an energy term that increases the more a model violates a given set of experimental restraints,  $E_{\text{restr}}$ ,

$$E_{\text{target}} = E_{\text{chem}} + E_{\text{restr}} \quad (4.1)$$

The optimization problem is to find a molecular conformation for which  $E_{\text{target}}$  attains a global minimum.

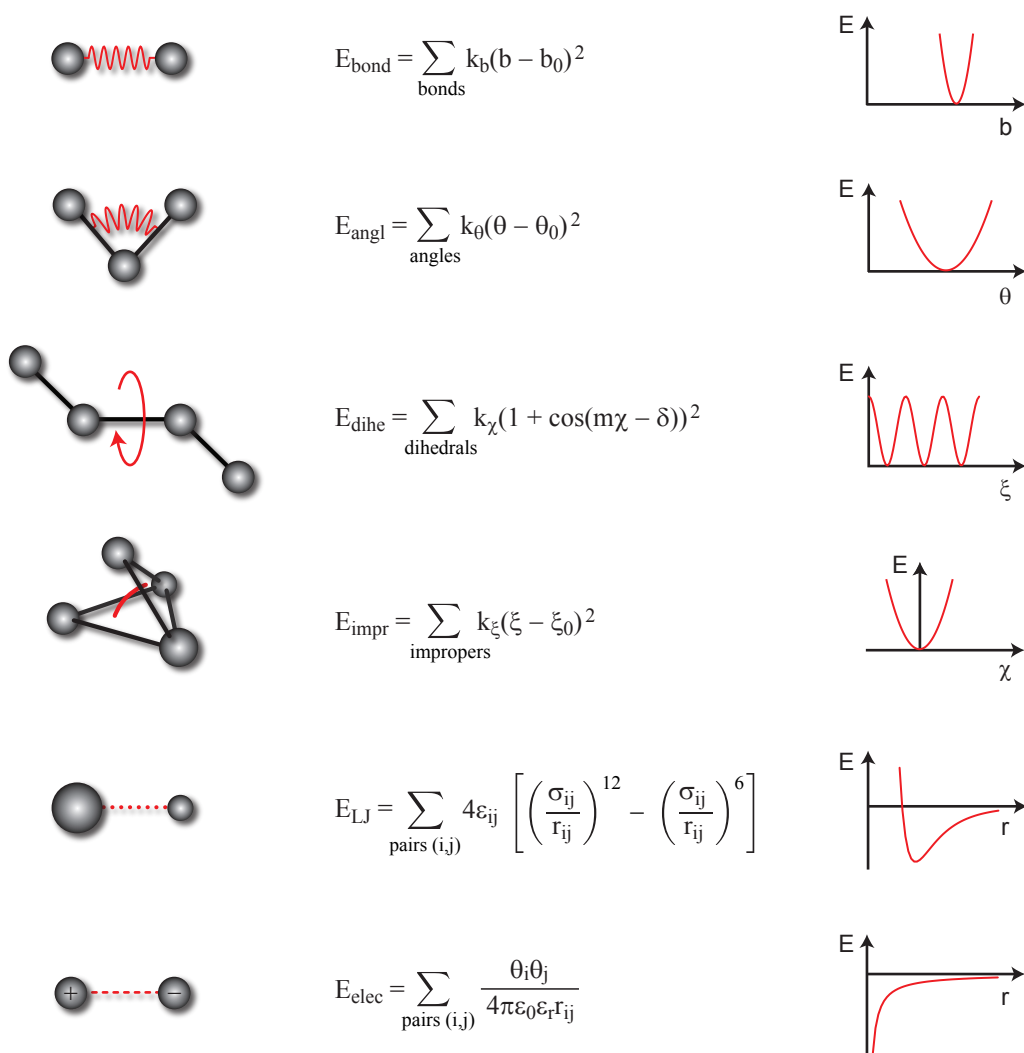
### 4.3.1 Empirical Force Field

$E_{\text{chem}}$  is given by the geometric relations between bonded atoms, such as bond lengths, angles and chirality, and non-bonded interactions, such as the electrostatic interactions, Pauli repulsion and van-den-Waals attraction. The contributions to  $E_{\text{chem}}$  as implemented in CNS are given in Figure 4.1. Regular angles (angl) relate to three nuclei, while *proper dihedrals* (dihe) and *improper dihedrals* (impr) relate to four nuclei. The term  $E_{\text{dihe}}$  is  $\frac{2\pi}{m}$ -periodic and is designed to avoid unfavorable geometries in periodic rotations about a bond. Improper dihedrals can be used to impose planarity constraints, e.g. on a phenyl ring. Pauli repulsion and van-den-Waals attraction are well represented by a Lennard-Jones (LJ) potential, but may be further approximated for data-driven structure calculation.

Parameters for the terms in Figure 4.1 (*force field*), such as force constants, bond lengths and angles, can in some cases be provided by theoretical analysis, but mostly need to be calibrated using experimental information. Parameters for bonded interactions sets used here for calculations by CNS have been obtained from a statistical analysis of proteins in the Protein Data Bank by Nilges and co-workers (ALLHDG force fields) [177, 178, 179]. Non-bonded interactions in HADDOCK docking simulations (*vide infra*) are obtained from the OPLS force field, which is derived from simulations on organic liquids and aqueous solutions of organic ions representative of subunits in the side chains and backbones of proteins [138].

### 4.3.2 Common Approximations for Structure Calculation.

The functional form of terms in  $E_{\text{chem}}$  (Figure 4.1) are chosen such that they are easy to compute. Still, within a cut-off of typically 8–10 Å, non-bonded interactions (LJ, elec) are



**Figure 4.1:** Empirical approximation to potential energies (*force field*) used in classical molecular dynamics simulations, as implemented in CNS. For experimental data-based structure calculation and refinement, the Lennard-Jones term and electrostatic term may be further approximated to provide a computational performance increase.



costly to evaluate. When experimental distances determine the overall fold, electrostatic interactions may be neglected, and the Lennard-Jones term may be replaced by a purely repulsive quartic potential,

$$E_{\text{elec}} = 0 \quad (4.2)$$

$$E_{\text{rep}}(r) = k_{\text{vwd}} (r_{\text{min}}^2 - r^2)^2, \quad (4.3)$$

where  $r_{\text{min}} = 0.8\sigma\sqrt[6]{2}$ , and  $\sigma$  is the distance at which the Lennard-Jones potential has a minimum [42]. For structure refinement, non-bonded interactions should be treated explicitly. Additionally, bond lengths are kept fix during structure calculations. As will be discussed below, this allows for a significant step size increase in the dynamic target function solver.

### 4.3.3 Experimental constraints

CNS provides various kinds of restraint potentials for  $E_{\text{restr}}$ . The work presented in this thesis makes use of square-well distance and dihedral angle constraints. Both are of the form

$$E_{\text{restr}}(q) = k \begin{cases} (q - q_{\ell})^2 & q < q_{\ell} \\ 0 & q_{\ell} < q < q_u \\ (q - q_u)^2 & q_u < q \end{cases} \quad (4.4)$$

where  $k$  is a force constant,  $q$  is either a distance or a dihedral angle, and  $[q_{\ell}, q_u]$  is an interval around an experimental restraint in which a violation of the restraint does not result in an energy penalty. It is also possible to define other restraining potentials.

#### Choice of distance restraint averaging

In N/CHHC spectra, ( $^{15}\text{N}, ^{13}\text{C}$ ) or ( $^{13}\text{C}, ^{13}\text{C}$ ) correlations may encode for multiple ( $^1\text{H}, ^1\text{H}$ ) pairs. In order to determine if a distance restraint is fulfilled, it becomes necessary to calculate an effective distance over all possible proton–proton-distances  $d_i$  that give rise to the observed correlation. In CNS, this is usually done in two ways,

$$d_{\text{SUM}}^{\text{eff}} = \left( \sum_{i=1}^N d_i^{-6} \right)^{-1/6} \quad \text{or} \quad d_{\text{R-6}}^{\text{eff}} = \left( \frac{1}{N} \sum_{i=1}^N d_i^{-6} \right)^{-1/6}. \quad (4.5)$$

The CNS terminology 'R-6' for the second option, may lead to some confusion, it differs from 'SUM' averaging by using the average ( $1/N$ ) instead of the sum alone before calcu-

lating the inverse root.<sup>1</sup> For methyl groups, the R-6 option should be used. Otherwise, the differences between SUM and R-6 is subtle, and is best illustrated with an example: If two distances are involved in the average, say 3.0 and 6.0 Å, the SUM averaging will produce an effective distance of 2.99 Å, while R-6 averaging will produce a distance of 3.36 Å. The fact that the effective distance in SUM averaging is always smaller than any single internuclear distance  $r_i$  is used in structure calculations employing *ambiguous distance restraints* [210] as implemented in the iterative calculations with ARIA [178], and in a more general way for implementing *ambiguous interaction restraints* in molecular docking with HADDOCK [70], see below.

## 4.4 Minimizing the Target Function

A search for molecular conformations that best fulfills the empirical and experimental constraints is achieved by minimization of  $E_{\text{target}}$ . Molecular Dynamics Simulated Annealing (MDSA) allows to find the global minimum, and Powell conjugate gradient minimization (CGM) is used for final optimization, efficiently locating local minima.

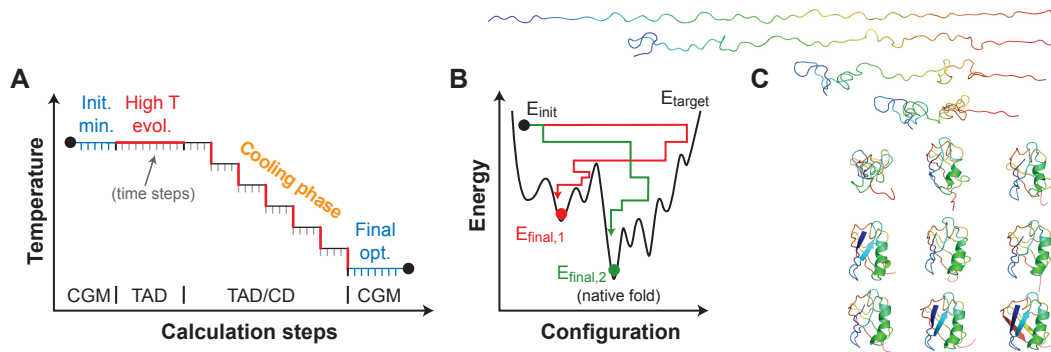
### 4.4.1 Molecular Dynamics Simulated Annealing

#### Simulated Annealing

Simulated annealing (SA) [151] has evolved as the method of choice for searching the conformation space and finding the global minimum of potential energy [103]. Kinetic energy determines which energy barriers can be overcome, hence the temperature schedule is important for success and efficiency of an SA scheme. As will be described below, SA here is performed by molecular dynamics simulation in contrast to, for example, Metropolis Monte Carlo. Typical CNS annealing schedules are described in Chapters 6 and 8. A schematic is given in Figure 4.2. When the starting structure is poor, performance of the SA schedule can be increased by initially performing high-temperature dynamics. Also, initial structures are usually subjected to a local minimization prior to SA, in order to remove high initial energy offsets due to e.g. significant bond stretching.

---

<sup>1</sup>A consistent nomenclature would be 'R-6 SUM' and 'R-6 AVE'.



**Figure 4.2:** Conformational search for the global energy minimum of  $E_{\text{target}}$  by simulated annealing molecular dynamics. (A) Structure calculation schedule with initial minimization, high-temperature search, simulated annealing, and final local optimization. (B) Schematic 1D representation of a protein energy landscape. Kinetic energy allows to overcome local minima. Calculations are repeated several times while scaling out kinetic energy via temperature coupling to find the global minimum. (C) Examples of folding intermediates during simulated annealing. Figure C obtained from ref. [41]. Abbreviations: TAD, Torsion Angle Dynamics; CD, Cartesian Dynamics; CGM, Conjugate Gradient Minimization.

### Molecular Dynamics Simulated Annealing

In Molecular Dynamics Simulated Annealing (MDSA), the molecular starting conformation is altered according to Newton's equation of motion while gradually scaling out kinetic energy via a temperature coupling term [29],

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = -\nabla_i E_{\text{target}} + \beta \left( \frac{T_0}{T} - 1 \right) \mathbf{v}_i, \quad (4.6)$$

where  $T$  is the temperature of the system,  $T_0$  is the bath temperature to which the system is coupled,  $\beta_i$  is the damping constant and  $\mathbf{v}_i$  is the velocity of each particle  $i$ . In practice, the temperature is reduced in discrete steps between a defined number of MD time steps (Figure 4.2A).

### Accelerated Calculations Using Holonomic Constraints

A significant performance increase in computational sampling of the conformation space of a macromolecule can be achieved by assuming fixed values for bond lengths or bond angles, when the covalent structure of a molecule is known. In particular, this applies to the determination of the overall fold of a protein. The conformational space of a typical protein is reduced by about one magnitude when only torsion angle rotations

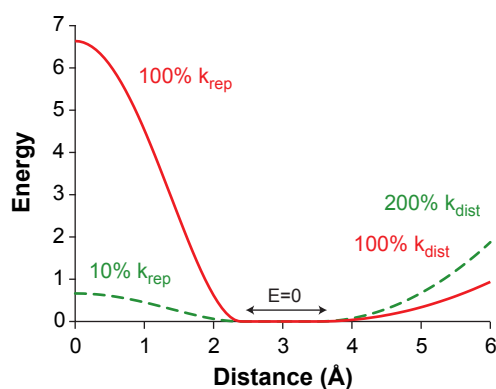
are considered (circa  $N/3$  torsion angles as compared to  $3N$  cartesian coordinates). The implementation of Torsion Angle Dynamics (TAD) in CNS [270] follows the algorithm by Bae and Haug [14, 15], which scales linearly with the number of atoms, as does Cartesian dynamics. It is slower than the implementation of the more recent algorithm by Jain [130] in DYANA [104], but sufficiently fast for applications shown herein.

In addition to the smaller number of calculations to be performed using TAD, the typical time step required to maintain a given level of energy conservation, and hence a proper functioning of the MD algorithm [2] can be significantly larger than in Cartesian dynamics, where bond stretching and bond angle variations are considered (CD: 1–2 fs). The maximum length of the time step that may be used depends on the temperature of the system. In DYANA, for example, time steps of about 100 fs, 30 fs and 7 fs at 1 K, 400 K and 10000 K were found to be reasonable for proper functioning of the molecular dynamics algorithm [103]. In CNS, typical time steps of 15 fs at 50000 K protein structure calculations in torsion angle space are used (20000 K for nucleic acids), with torsion angle force constants of about  $100 \text{ kcal}\cdot\text{mol}^{-1}\text{deg}^{-2}$  [270]. Note, however, that time steps between different programs and in particular with respect to 'classical' molecular dynamics are difficult to compare, because other parameters, such as the form of potential functions, are also of influence on the effective structural variation [103].

Models from simulated annealing in torsion angle space are usually refined with one or two rounds of Cartesian dynamics (CD) SA. Here, only bond stretching motion, in particular fast hydrogen oscillations, are removed using the SHAKE algorithm [241]. A typical time step in CNS for Cartesian SA is 5 fs, with annealing from 1000 to 300 K.

## Force Scaling

Folding of the polypeptide strand according to a set of long-range distance restraints is subject to sterical hindrance from (Pauli) repulsive forces among the atoms, represented by a quartic potential  $E_{\text{rep}}$  in structure calculation. In order to obtain an overall fold that fulfills the experimental restraints, forces may be scaled to smoothen the energy landscape and allow an easier passing of atoms. As an example, Figure 4.3 shows the downscaling of repulsive forces and upscaling of distance restraints during the initial high-temperature search of a typical CNS simulated annealing schedule. Values relate to the final stage of annealing. Similarly, dihedral angle constraints may be scaled.



**Figure 4.3:** Downscaling of repulsive force constants ( $k_{\text{rep}}$ ) and upscaling of distance restraints ( $k_{\text{dist}}$ ) during high-temperature search of a typical structure calculation schedule. In the given case, a distance constraint violated by more than 6 Å could enforce a full passage of two repulsive spheres (atoms) through each other.

### Computational Effort

Using MDSA as described above, the calculation of an ensemble of 100 structures of a small protein such as the 36-residue protein Kaliotoxin (*vide infra*) on one CPU can be performed within a few hours. Calculations of an ensemble of larger proteins, such as fibril segments of full-length  $\alpha$ -synuclein or docking of a large molecular complex, such as Phospholamban–SERCA (Chapter 9) with more than thousand residues, require several hours for each structure. Larger systems generally require longer simulated times than small systems in order to converge well [270], and an increased number of time steps improves the quality of structures in refinement [269]. Each structure can be calculated independently from the others, hence calculations of large structures are most effectively performed on computer clusters with fast nodes and small communication bandwidth instead of highly parallel machines.

#### 4.4.2 Conjugate Gradient Minimization

Final optimization of the molecular structure with respect to  $E_{\text{target}}$  is achieved by Powell Conjugate Gradient Minimization (CGM) [233]. CGM is an efficient optimizer for local minima, but it does not perform 'uphill' steps, hence, in order to overcome local energy minima, methods like simulated annealing must be employed. In contrast to other minimizers, which converge slowly near the minimum, the conjugate gradient algorithm produces steps that continually refine the direction towards the minimum. In CGM, the

new direction is computed by adding the gradient at the current point to the previous direction, scaled by a constant. The effect of the resulting minimization process is illustrated in Appendix Figure B.1.

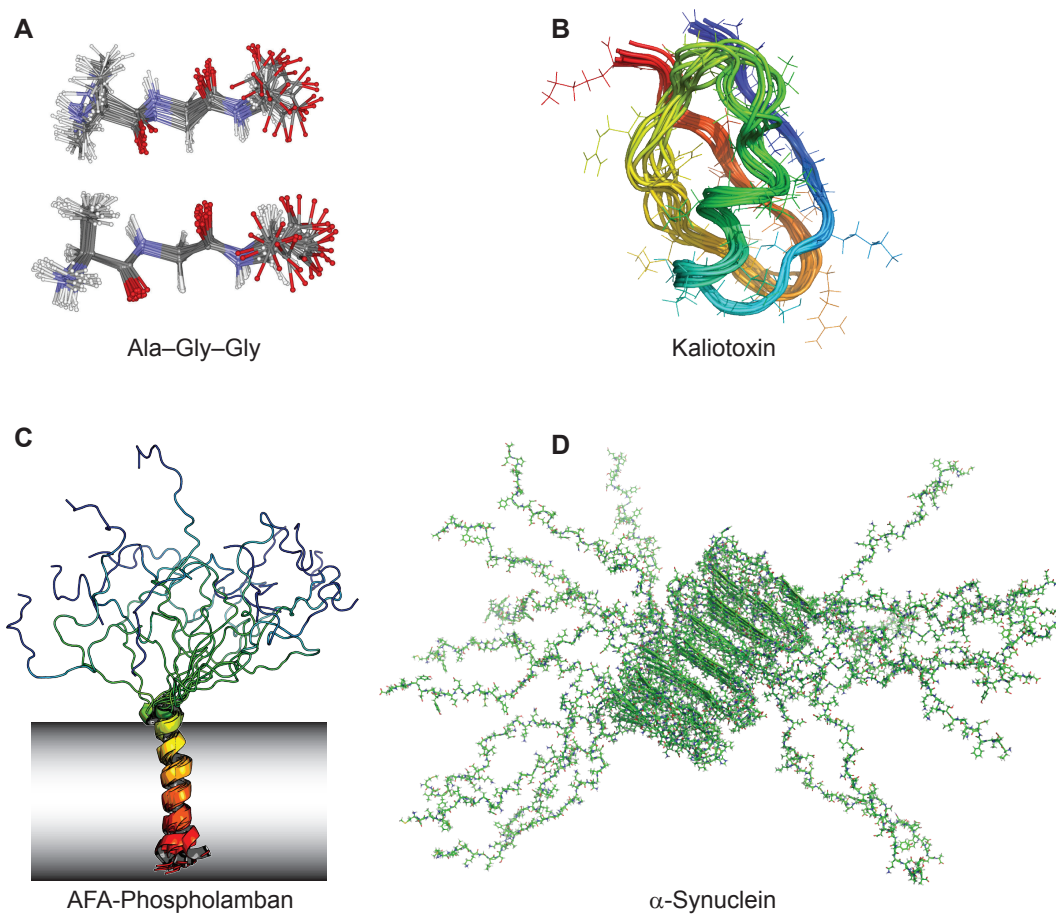
### 4.4.3 Structure Assessment and Quality Control

Repeated structure calculation result in a set of different molecular conformations, which vary depending on the quality of input data. The ensemble of structures that is ultimately chosen to represent the structure of the molecule typically comprises of 10–30 structures, reflecting a compromise between sufficient statistics and manageability in graphics and analysis programs [103]. Depending on the convergence of structures, i.e. the fraction of structures in agreement with the experimental data, an accordingly larger number of conformers needs to be computed. The distinction between acceptable and unacceptable conformers depends on the protein, the data set and the structure calculation algorithm used, it might be clear-cut or gradual. In any case, the chosen structures should not significantly violate the experimental data, and represent the range of possible conformations. [103]

The quality of a structural ensemble is commonly assessed in terms of (1) the residual empirical energy  $E_{\text{chem}}$ , (2) the number, size and distribution of restraint violations that contribute to  $E_{\text{restr}}$ , (3) atomic Root-Mean-Square Deviations within the ensemble (Appendix Chapter B), (4) the distribution of backbone torsion angles and other geometric parameters, using software like PROCHECK [165], WHAT IF [309] and WHAT\_CHECK [120] or MolMol [152], and (5) a description using 3D visualization software, such as MolMol, DeepView/Swiss-PdbViewer [100], VMD [124] or PyMOL [65].

## 4.5 Structure Calculation Examples

Four examples of biomolecular models based on solid-state NMR data from uniformly ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-labeled samples using CNS simulated annealing protocols are shown in Figure 4.4. The structure of Ala–Gly–Gly (AGG) was determined using chemical shift-derived backbone torsion angles for the central residue (Figure 4.4A, upper model), and refined with  $^1\text{H}$ – $^1\text{H}$  distances obtained from an NHHC correlation spectrum using a dual restraint classification (Figure 4.4A, lower model) [160]. Essentially the first structure of a fully uniformly labeled protein, the 38-residue toxin KTX (Figure 4.4B), was computed from unambiguously identified  $^1\text{H}$ – $^1\text{H}$  restraints in a CHHC correlation spectrum, in com-



**Figure 4.4:** Molecular models based on solid-state NMR data. (A–C) Structures determined unambiguously, as published in refs. [9, 160, 161]. (D) Possible structure of full-length  $\alpha$ -synuclein in line with solid-state NMR data. All structures were calculated with CNS simulated annealing protocols.

bination with chemical shift-derived torsion angle restraints [161] (PDB entry 1XSW). The calculation involved an iterative approach that allowed to identify one incorrectly assigned distance restraint. The structure of AFA-mutant Phospholamban (AFA-PLN) as seen by solid-state NMR spectroscopy in liposomes (Figure 4.4C) results from chemical shift-derived backbone torsion angles in the membrane-integral region, whereas the cytoplasmic domain is largely disordered, as determined by mobility-filtered correlation INEPT-type and NOESY experiments (a few detected NOESY cross peaks are included as distance restraints) [9]. Structures violating the lipid bilayer geometry were sorted out after structure calculation in vacuum. All structures presented in Figure 4.4A–C were determined unambiguously. In addition, a possible model for full-length  $\alpha$ -synuclein in line with solid-state NMR data is shown in Figure 4.4D.<sup>2</sup> Dihedral angle constraints and intra- and intermolecular distances were derived from a number of sources, including secondary structure identifications by ssNMR [111], X-ray diffraction,  $^1\text{H}/^2\text{D}$  exchange or protein digestion. Such models can be used to back-calculate intra- and intermolecular correlations, which may then be compared to, for example, NHHC and CHHC spectra of uniformly ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ) and mixed ( $^{13}\text{C}$  or  $^{15}\text{N}$ ) labeled samples.

## 4.6 Molecular Docking

### 4.6.1 *Ab Initio* Versus Experimental Data-Driven Docking

Molecular complexes can be constructed by *docking* its pre-structured binding partners. Different aspects may be emphasized in the docking procedure, depending on the size of the binding partners, computational method, or scientific scope of the work: Extensive surface sampling when the exact position of a small ligand is unknown, highly accurate binding energy calculation, compliance with experimental data, or extended flexibility including conformational changes of the receptor [37].

Docking of a small ligand to a large binding partner, like a protein receptor, without detailed experimental information on the binding interface, is best performed with algorithms that treat the molecular interface as a two-dimensional hypersurfaces [200, 204]. These are extremely fast, but allow for flexibility only in a small ligand or in the side chains of the receptor. In contrast, the approach described in the following allows for docking with arbitrary flexibility and ligand size. It is very similar to the structure calculations discussed above, and is driven by (possibly highly ambiguous) experimental data.

---

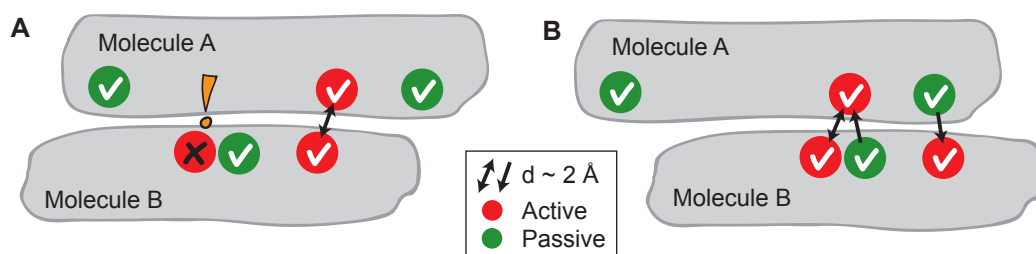
<sup>2</sup>Heise et al., EUROMAR 2006, York/UK



### 4.6.2 High Ambiguity Data Driven Docking

There is a wealth of experimental information on molecular interfaces, such as cross-linking, mutagenesis screening or chemical shift perturbation data, that is often not used in computational docking [301]. The HADDOCK approach ('High Ambiguity Driven biomolecular DOCKing based on biochemical and/or biophysical information') formalizes the way information from different sources is translated into (ambiguous) distance restraints, and provides CNS protocols for the calculation and evaluation of multi-molecular complex models [70].

First, any kind of constraint potential that is implemented in CNS, such as explicit distance and dihedral angle information, may be applied during HADDOCK docking. Second, ambiguous (i.e., not necessarily atom- or residue-specific) information is encoded in *Ambiguous Interaction Restraints* (AIRs) using *active* and *passive* residues. Active residues require an intermolecular contact to another active residue, or to a passive residue. Passive residues can satisfy intermolecular contact requirements from active residues, but do not need to be involved in any intermolecular contact themselves. A residue should be declared active when it was identified to form an interface contact. A residue should be declared passive when it is likely to form an interface contact, but no positive proof is available, for example, for surface neighbors to active residues. This relation is depicted in Figure 4.5. An AIR is defined as an intermolecular distance  $d_i$  with a maximum value



**Figure 4.5:** Illustration of the concept of active and passive residues in HADDOCK, see also text. (a) For the given molecular complex arrangement, an ambiguous interaction restraint defined for an active residue ( $\times$ ) in molecule B is not fulfilled by close spatial proximity to an active or passive binding residue of the binding partner, molecule A (" $\neq$ "). (b) Alternative arrangement of molecules A and B, where all AIRs are fulfilled.

of typically  $2 \text{ \AA}$  between all atoms of an active residue  $i$  of a protein  $A$  ( $m_i$ ) and any atom of both active and passive residues (together  $j$ ) of a protein  $B$  ( $n_j$ ). The effective distance

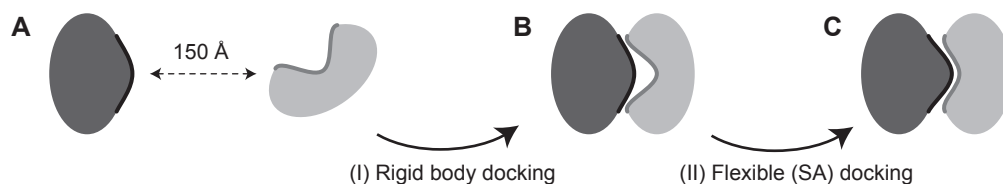
for each restraint is calculated using the equation

$$d_i^{\text{eff}} = \left( \sum_{m_i} \sum_j \sum_{n_j} d_{m_i n_j}^{-6} \right)^{-1/6} \quad (4.7)$$

Similar to the SUM method in (4.5), the restraint is fulfilled ( $d_i^{\text{eff}} < 2 \text{ \AA}$ ) as soon as any two atoms in the AIR set of the two proteins are in contact. Inversely, AIRs for residues in protein *B* are defined. The definition of passive residues ensures that residues at the interface which are not detected, for example when no change in binding can be identified upon mutation, are still able to satisfy the AIR restraint, i.e. provide a contact for active residues in the partner molecule.

### 4.6.3 Protein Flexibility and Docking Schedule in HADDOCK

Docking in HADDOCK is performed as a combination of rigid-body docking, simulated annealing molecular dynamics, and MD refinement, using an explicit spatial representation of all atoms. Hence, unlike other docking approaches, HADDOCK allows for an arbitrary degree of protein flexibility. The only limitation to flexibility is that there should be sufficient constraints by the empirical force field or by explicit constraints and AIRs to confine flexible elements to a realistic scale of change. For example, it is reasonable to assume for side chain flexibility and a certain degree of backbone flexibility at the binding interface (these residues are defined based on a set of rules, and comprise, among others, all active and passive residues). Since the empirical energy terms are approximated to achieve a sufficient computational speed-up, large conformational changes cannot be induced by the force field, but must result from experimental constraints. For example, in context with AFA-PLN–SERCA1a docking (Chapter 9), protein *B* (AFA-PLN) was allowed to be fully flexible at any stage besides of initial rigid-body docking, and was constrained by backbone dihedral angles restraints derived from solid-state NMR experiments. In a typical HADDOCK docking simulation, calculations start from the binding partners separated by 150 Å in random orientations, and are followed by several cycles of simulated annealing with different degrees of flexibility (Figure 4.6). Gentle water refinement may be performed using a shell of 8 Å TIP3P water [179]. Depending on the molecular context, rigid-body docking or water refinement may be omitted. A docking schedule is given in Table 4.1, referring to the docking of AFA-PLN with SERCA1a described in Chapter 9.



**Figure 4.6:** (A–C) Main stages in a docking simulation with HADDOCK. Details of the protocol are given in Table 4.1. Optionally, the complex may be refined in a shell of water, or with an external MD simulation software.

Stage	Dynamics	Temperature	Interface Flexibility	# Structures	CPU hours
I	Rigid-body	–	none	1000	20
II.a.i	Torsion angle	constant (high)	none	200	300
II.a.ii	Torsion angle	SA	none	200	
II.a.iii	Torsion angle	SA	side chain	200	
II.a.iv	Torsion angle	SA	side chain, backbone	200	
II.b.i	Cartesian	constant (high)	none	200	40
II.b.ii	Cartesian	SA	none	200	
II.b.iii	Cartesian	SA	side chain	200	
II.b.iv	Cartesian	SA	side chain, backbone	200	
III	Water refine	constant	interface sc & bb	20	40

**Table 4.1:** Docking schedule in HADDOCK. Abbreviations: bb, backbone; sc, side chain. CPU hours relate to docking AFA-PLN (54 residues) with SERCA1a (996 residues) on Intel Xeon 5160.



## **New Techniques and Applications**



## Chapter 5

# On the Prediction of NMR Chemical Shifts for Proteins in the Solid Phase

*Durch die Experiens wird sich leicht erweisen,  
welche Theorie mit der Wahrheit übereinkommt.*

—Leonhard Euler

### 5.1 Summary

A statistical analysis of nuclear chemical shift predictions of proteins in the solid state by rapid algorithms trained on and verified with solution-state NMR assignments is presented. The precision of predictions by three common computer programs evaluated for this purpose (SHIFTS, SHIFTX and SPARTA) was found to be close to chemical shift predictions of proteins in solution. Correlation coefficients depend on the kind of resonance (N, C', C $^{\alpha}$  and C $^{\beta}$ ) and on secondary structure ( $\beta$ -strand, random coil and  $\alpha$ -helix), but also on the molecular environment (membrane-integral or not). The findings establish a quantitative basis for using these rapid chemical shift prediction programs for solid-state NMR chemical shift analyses. On the other hand, prediction inaccuracies identified for certain resonance kind, residue type and molecular environment, point to possible areas of methodological improvement.

### 5.2 Introduction

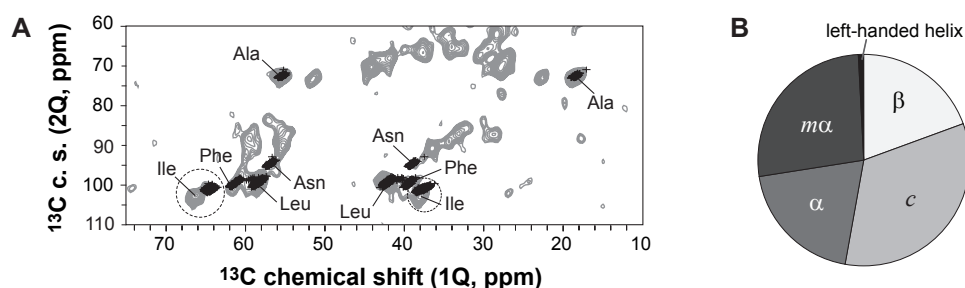
Physical properties of a system can be assessed by comparing experimental data to calculations from theory. A particularly sensitive probe to local molecular structure is the nuclear chemical shift. Backbone  $^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}$  chemical shifts depend on several factors, such

as hydrogen bonds, electric field effects, ring currents, and backbone and side-chain torsion angles. The influence of intra-residue backbone torsion angles  $\phi$  and  $\psi$  dominates the chemical shift of C', C $^\alpha$  and C $^\beta$  with respect to the residue-specific average (secondary chemical shift) [3, 63, 242, 268, 319], whereas  $^{15}\text{N}$  chemical shifts depend also on the conformation of the preceding residue, correlating with  $\psi_{i-1}$  in addition to  $\phi_i$ , where  $i$  is the residue number of  $^{15}\text{N}$  [63, 91, 168]. The contributions to  $^1\text{H}$  chemical shifts shall not be considered here.

The *ab initio* calculation of nuclear chemical shifts for entire proteins is generally precluded due to high computational demands. In recent years, however, a number of algorithms became available which allow for a rapid estimation of chemical shifts from existing structural data, based on a combination of approximative quantum-mechanical or empirical considerations [129, 199, 207, 259, 323]. Despite of limited precision, these shift predictions can be of great value in structural investigations by NMR spectroscopy. Under magic angle spinning or fast isotropic motion, for example, knowledge on the isotropic chemical shifts in a protein allows for an initial verification of correct overall fold and domain mobility by means of multidimensional correlation spectroscopy. In structural studies on systems with partially known components, such as membrane proteins with solvent-hosted globular domains, shift predictions can help to identify regions of significant structural difference. More recently, rapid chemical shift prediction algorithms have also been used for spectral fitting within structure calculation [48, 112, 194].

Today's chemical shift prediction algorithms have been trained with solution-state NMR assignments, and X-ray crystallographic or solution-state NMR structural data. The quality of predictions was evaluated in reference to soluble proteins. Due to the low availability of solid-state NMR resonance assignments, the validity of applying these routines in chemical shift predictions for protein in the solid phase could so far only be evaluated for three micro-crystalline proteins [112]. More recently, we observed a systematic deviation of predictions from measured chemical shifts in several membrane-integral proteins. For example, Figure 5.1A shows a comparison of predictions from an ensemble of structures for transmembrane helix residues N30–I48 of AFA-mutant phospholamban (AFA-PLN), reconstituted in liposomes [9], to an experimental  $^{13}\text{C}$  double-quantum/single-quantum correlation spectrum. While predictions for Ala, Asn, Leu and Phe residues center around the correlations of the isotropic chemical shifts, Ile predictions deviate by more than one resonance line width from the unambiguously identified cross peak of Ile residues. This observation raises the following question: *Can chemical shift predictions on solid-phase*





**Figure 5.1:** (A) Comparison of  $C^\alpha$  and  $C^\beta$  chemical shift predictions from an ensemble of AFA-PLN structures, obtained using SHIFTX, to a 2Q/1Q correlation spectrum of AFA-PLN reconstituted in liposomes for selected residues found within the membrane. (B) Distribution of residues in the data base according to different structure classes,  $\beta$ -strand ( $\beta$ ), random coil ( $c$ ),  $\alpha$ -helix ( $\alpha$ ,  $m\alpha$ ), and left-handed helix. In case of  $\alpha$ -helical conformations, a sufficiently large number of assignments was available to distinguish between residues found membrane-integral ( $m\alpha$ ) and not membrane-integral ( $\alpha$ ).

*proteins from software trained on solution-state NMR data be regarded a reliable tool for structural analysis?* If so, what are the error margins, and can predictions be improved by correcting for possible effects from the molecular environment? The need for a quantitative analysis of prediction results becomes even more imperative in the light of a growing number of applications where the inverse process, the prediction of backbone-torsion angles from chemical shifts by similar programs, e.g., TALOS [58] or PREDITOR [30], has already been used in structural investigations by solid-state NMR spectroscopy. Possibly, the analysis should not only be based on a larger number of assignments than in ref. [112], but also on a greater variety of sample preparations.

Here, the performance of three computer programs, SHIFTS (2002) [323], SHIFTX (2003) [207] and SPARTA (2007) [259], that allow for the prediction of  $^{15}\text{N}$  and  $^{13}\text{C}$  (isotropic) solid-state NMR chemical shifts, is evaluated. For this purpose, a database of more than 600 protein residues was compiled, for which solid-state NMR resonance assignments and high-resolution 3D structures are available. In particular, this includes several proteins with membrane-integral  $\alpha$ -helices ( $m\alpha$ ), see Figure 5.1B. All three considered programs require three-dimensional input coordinates for the polypeptide of interest. They predict backbone  $^{15}\text{N}$ ,  $^{13}\text{C}'$ ,  $^{13}\text{C}^\alpha$ , and side chain  $^{13}\text{C}^\beta$  chemical shifts. The quality of  $^1\text{H}$  predictions ( $^1\text{H}^\text{N}$ ,  $^1\text{H}^\alpha$  for all three programs, and side-chain  $^1\text{H}$  for SHIFTS and SHIFTX) will not be considered here, as the number of ssNMR assignments for these nuclei is too small.

SHIFTS [323] predicts  $^{15}\text{N}$  and  $^{13}\text{C}$  chemical shifts based on an additive model for conformational effects from a database of density functional theory calculations on a large number of peptides, complemented by empirical extensions for some conformational regions and residue types. SHIFTX [207] uses a combination of an empirical  $(\phi, \psi)$  torsion angle-dependent chemical shielding hypersurface with classical hydrogen bonding and secondary structure analysis. It provides relatively accurate and complete predictions [207, 259]. SPARTA [259] searches a database for matching triplets of adjacent residues to the query triplet in terms of  $(\phi, \psi, \chi^1)$  torsion angle and sequence similarity. The average secondary shifts for the 20 best-matching triplets are weighted according to empirically optimized parameters, and corrected for nearest neighbor, ring current and hydrogen bond effects. At time of the investigation, the PROSHIFT [199] web server was not available. The PROSHIFT prediction algorithm is based on neuronal networks, however, with an RMS deviation of 1.3 ppm for carbon and 3.6 ppm for nitrogen [199] of proteins in solution, its performance does not exceeded that of SHIFTX [207].

### 5.3 Methods

Solid-state NMR assignments were retrieved from the BioMagResBank (BMRB) [249], from published articles, or directly from the authors, see Table 5.1. In case of phospholamban and SRII, where assignments from mobility-selective scalar- and dipolar-based transfer schemes were available, only data for the rigid segments was used. Only shifts that can be predicted by all three programs were considered, hence,  $^{15}\text{N}$  of Pro, which is not predicted by SHIFTX, and all Cys resonances, which are not predicted by SHIFTS, were excluded. Also, amino- and carboxy-terminal residues were disregarded, as they exhibit unusual shifts, and are not predicted by SHIFTS and SPARTA. Twelve predictions of C' resonances by SHIFTS, of which eight were assigned in the solid state, were off by about a factor of two, and were excluded from the data base.<sup>1</sup> One further unusual shift,  $^{15}\text{N}$  in Crh Q3 (82.6 ppm), which was 36 ppm off the average prediction of all three programs, was also excluded. In proteins with resonance splitting (Crh, Ubiquitin), only the first resonance was used. No further shifts were discarded.

Chemical shifts predictions were re-referenced to yield zero offset for the average pre-

---

<sup>1</sup>These predictions showed no obvious correlation with residue type or residue number (such as SH3 S19 or Ubi I36), and input files for six out of seven affected proteins were approved structures from the Protein Data Bank.

Molecule	Kind	Asgn. Res.	SSNMR source, year (BMRB entry)	Prediction source (PDB code, method)
BPTI	globular	9	McDermott et al. [196], 2000	Wlodawer et al. [321] (6PTI, X-ray)
Crh	globular (dimer)	83	Böckmann et al. [44], 2003	Juy et al. [139] (1MU4_A, X-ray)
GB1	globular	54	Franks et al. [82], 2005	Franks et al. [83] (2GI9, X-ray)
KcsA (Kv1.3)	membrane integral	72	Lange et al. [162], 2006; Schneider et al., [248]	Zhou et al. [331] (1K4C, X-ray), variation [162]
KTX	globular	26	Lange et al. [161], 2005 (6341)	Gairi et al. [89] (2KTX_1, soln. NMR)
LH2	membrane integral	75	van Gammeren et al. [302], 2005 (6348)	Papiz et al. [219], (1NKZ.D/E, )
MPX	membrane bound	12	Fujiwara et al., 2004 [87] (6214)	Todokoro et al. [285] (2CZP_1, ssNMR)
PLN (AFA)	membrane integral	29	Andronesi et al. [9], 2005	Andronesi et al. [9] (ssNMR)
SH3	globular	53	Castellani et al. [47], 2003	Chevelkov et al. [54] (1U06, X-ray)
SRII	membrane integral	67	Etzkorn et al. [81], 2007	Gordel'iy et al. [93] (1H2S_A, X-ray)
TRX	globular	103	Marulanda et al. [191], 2005	Katti et al. [142], (2TRX, X-ray)
TTR	fibril	9	Jaroniec et al. [134], 2004	Jaroniec et al. [134] (1RVS_1, ssNMR)
UBI	globular	63	Seidel et al. [254], 2005	Cornilescu et al. [57], (1D3Z_1, soln. NMR)

**Table 5.1:** Proteins and polypeptides with solid-state NMR resonance assignments and high-resolution structure information. Column three shows the number of residues in the final data base with at least one assigned resonance.

dictions from all three programs for each protein individually [329], with a common offset for the three carbons sites. For this purpose, predictions that deviate by more than 3 ppm ( $^{13}\text{C}$ ) or 6 ppm ( $^{15}\text{N}$ ) from the experimental values were temporarily disregarded.

The final data base comprised of 512 N, 448 C', 648 C $^{\alpha}$  and 574 C $^{\beta}$  assignments. Overall, 655 residues had at least one solid-state NMR assignment ( $^{13}\text{C}$  or  $^{15}\text{N}$ ). The secondary structure for residues with at least one assignment divided into 20% beta-strand ( $\beta$ ), 33% coil ( $c$ ) and 47%  $\alpha$ -helix, with 20% contributed by not membrane-integral ( $\alpha$ ) and 27% membrane-integral ( $m\alpha$ ) residues, see Figure 5.1B. Five residues (1%) with assigned chemical shifts found in a left-handed helix will not be considered in the following.

## 5.4 Results and Discussion

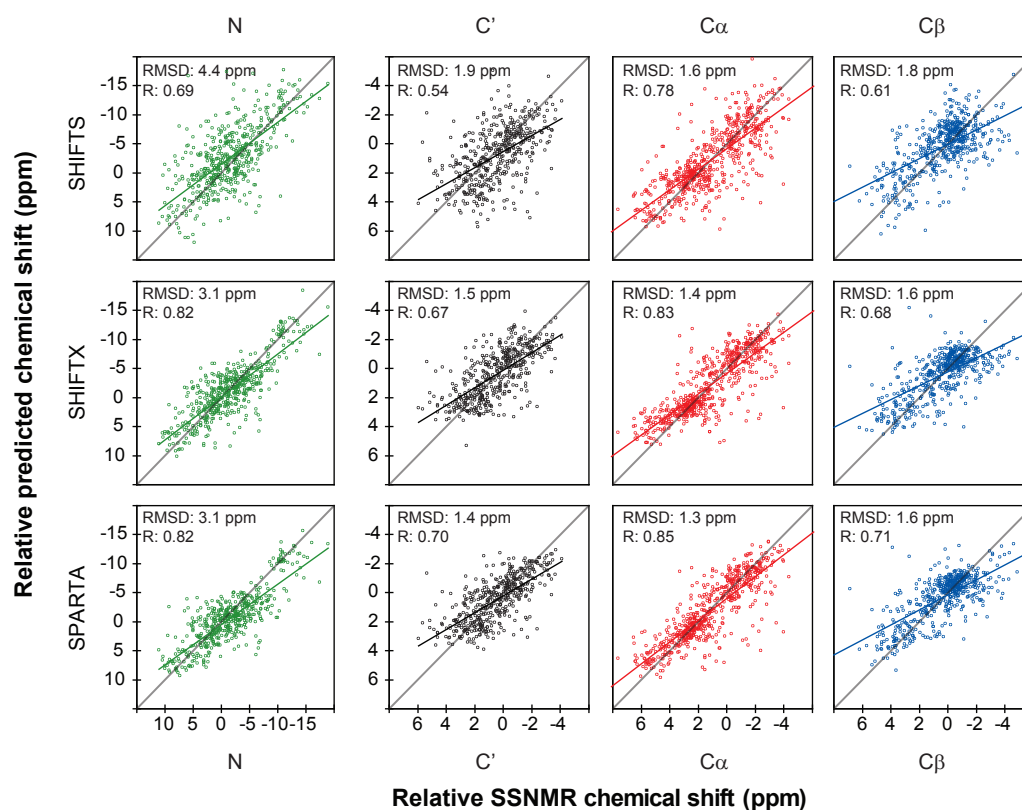
### 5.4.1 Overall Quality of Predictions

#### Reference Offsets

When the same  $^{15}\text{N}$  reference is used, the average over all  $^{15}\text{N}$  predictions is found for SHIFTS 1.1 ppm below, SHIFTX 0.3 ppm above, and SPARTA 0.9 ppm above the average of all three programs, hence, the shift offset of a prediction should be adjusted individually for each program used. Predictions for  $^{13}\text{C}$  can be adjusted to a common reference, with about 0.1 ppm variation among averages over all C', C $^{\alpha}$  and C $^{\beta}$  predictions of each program, and among the different programs.

#### Statistical Correlation

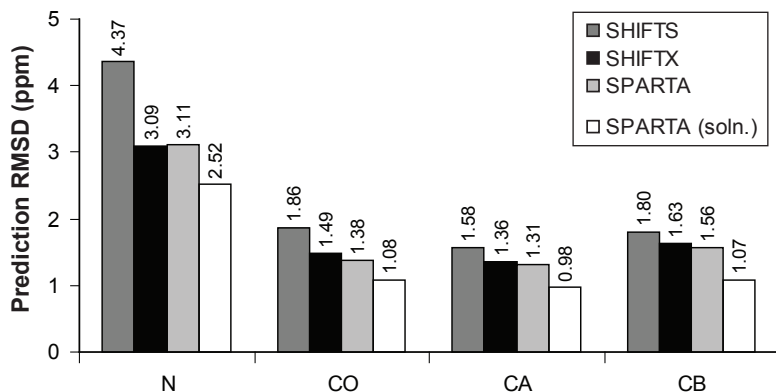
For all three programs, predicted and measured chemical shift are clearly correlated, as can be seen in Figure 5.2. In order to remove residue-specific offsets, the random coil value of the respective amino acid type was subtracted for each correlation (*relative shift*). All three programs show a slight tendency to under-estimate large shift deviations from the reference value, see for example correlations for C $^{\alpha}$  near 5–6 ppm in Figure 5.2. Chemical shifts for  $^{15}\text{N}$  deviate much more strongly from the random coil value than  $^{13}\text{C}$ , note the different scales for  $^{15}\text{N}$  and  $^{13}\text{C}$  in Figure 5.2. The (Pearson) correlation coefficient  $R$  from SHIFTX and SPARTA predictions for  $^{15}\text{N}$  is 0.82, and about 0.84 for C $^{\alpha}$ , see Figure 5.2 and Table 5.2.  $R$  is smaller for C' and C $^{\beta}$ . SHIFTS predictions correlate overall less good with the experimental values.



**Figure 5.2:** Comparison of chemical shifts predicted by SHIFTS, SHIFTX and SPARTA to values experimentally determined by ssNMR spectroscopy. For each predicted or measured shift, the random coil value of the respective amino acid type was subtracted (*relative shift*). Linear regression and the diagonal of perfect predictions are given. Note the different scale for  $^{15}\text{N}$ . The RMS prediction error is also discussed along Figure 5.3. The plot for SPARTA from comparing predictions to solution-state NMR shifts is given in Appendix Figure C.1 for reference.

### Mean Prediction Error

The RMS prediction error is about 3 ppm for  $^{15}\text{N}$  and around 1.5 ppm for  $^{13}\text{C}$  for SHIFTX and SPARTA, and larger for SHIFTS, see Figure 5.3. In terms of RMSD, SHIFTX per-



**Figure 5.3:** Comparison of root-mean square prediction errors for SHIFTS, SHIFTX and SPARTA with respect to ssNMR data. Values for SPARTA predictions of solution-state NMR chemical shifts are given as reference.

forms about 30% ( $^{15}\text{N}$ ) and 15% ( $^{13}\text{C}$ ) better than SHIFTS. SPARTA performs equally well as SHIFTX for  $^{15}\text{N}$ , and about 5% better for  $^{13}\text{C}$ . With respect to solubilized proteins, carbon predictions are about 1/3 less precise, as can be seen in Figure 5.3 from the comparison to SPARTA statistics of solution-state NMR test predictions. Notably, deviations of chemical shift predictions from experimentally determined values do not necessarily result from a methodological shortcoming, but may also indicate a difference between the prediction input model, in most cases a solution-state NMR or X-ray diffraction structure, and the molecular structure of the sample in the solid-state NMR experiment.

### 5.4.2 Outliers

Predictions were inspected regarding the probability of mis-estimating the experimentally determined shift beyond a certain threshold. For example, typical differences of average shifts between  $\beta$ -strands and  $\alpha$ -helices exceed 3 ppm for backbone carbons, with coil values found in between [312]. The fraction of  $^{13}\text{C}$  predictions differing from experimental values by more than 3 ppm is about 5% in SHIFTX and SPARTA, see Table 5.2. SHIFTS, in line with the increased prediction RMSD, yields a higher percentage of predictions above the threshold. For  $^{15}\text{N}$ , most (3/4) SHIFTS predictions deviate from the experimental values

Indicator	Program	$^{15}\text{N}$	C'	C $^{\alpha}$	C $^{\beta}$
Outliers (%)	SHIFTS	17	11	7	9
	SHIFTX	7	5	5	6
	SPARTA	7	4	5	6
for threshold (ppm)		6	3	3	3
Prediction RMSD (ppm) (including outliers)	SHIFTS	4.4	1.9	1.6	1.8
	SHIFTX	3.1	1.5	1.4	1.6
	SPARTA	3.1	1.4	1.3	1.6
Regression coefficient R (including outliers)	SHIFTS	0.69	0.54	0.78	0.61
	SHIFTX	0.82	0.67	0.83	0.68
	SPARTA	0.82	0.70	0.85	0.71

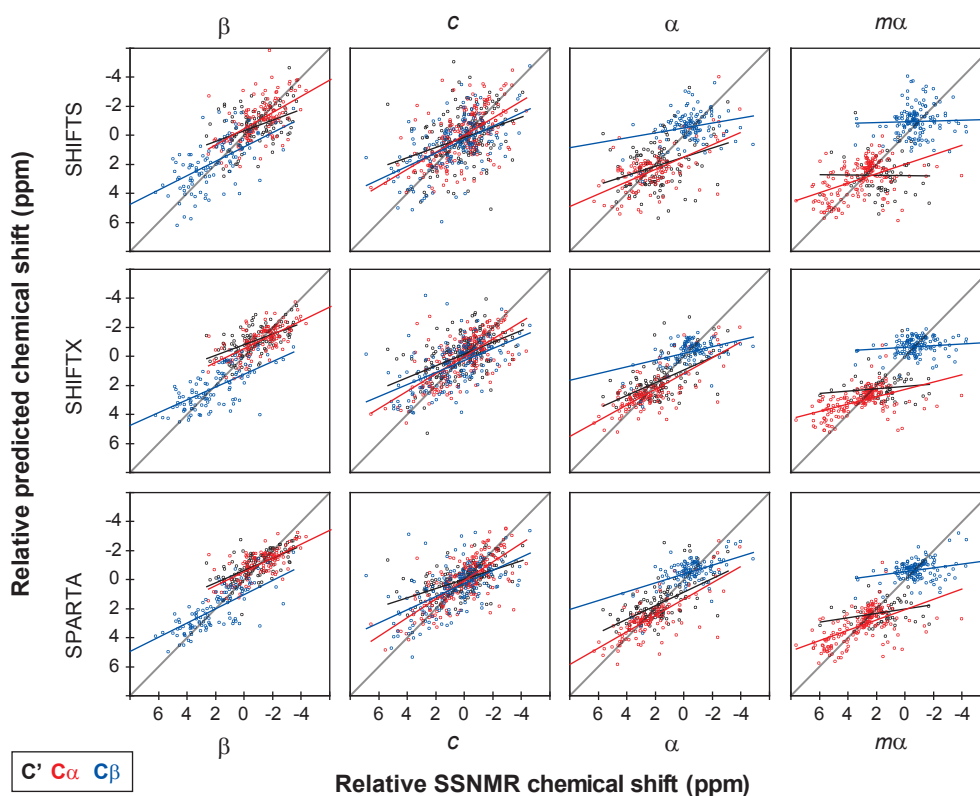
**Table 5.2:** Summary of indicators for the quality of solid-state NMR chemical shift predictions.

by more than 3 ppm, and only just a little more than half of SHIFTX and SPARTA predictions are within the threshold, owing to the larger RMS prediction error with respect to carbons. When the threshold for  $^{15}\text{N}$  is doubled, the number of outliers falls below 20% (SHIFTS) and below 10% (SHIFTX and SPARTA), see Table 5.2. Secondary structure-dependent chemical shift differences for  $^{15}\text{N}$  are in the same range [312], hence  $^{15}\text{N}$  shifts are only of small use in analyzing secondary structure, as in solution-state NMR. Furthermore, with a dropping number of outliers from SHIFTS to SHIFTX/SPARTA, predictions become more precise (Table 5.2). In other words, the higher probability of predicting a chemical shift within a given cutoff is not a trade-off for lower precision, encouraging the use of the more recent programs.

The incorrectness of a certain fraction of predictions is not a critical issue when a large number of predictions is being evaluated at the same time, allowing for statistical outliers, such as in context of general data assessment or structure fitting to spectra. In this case, quantitative information as presented above can be helpful in adjusting the weight of chemical shift contributions to structural fitting algorithms [48, 112, 194]. On the other hand, the correctness of individual predictions should not be relied on when the unambiguous identification of a single cross peak in a spectrum is essential. In these cases, reliable chemical shift identifications can only be obtained from sequential assignment by spectroscopic means.

### 5.4.3 Influence of Secondary Structure and Chemical Environment

Next, it was of interest to see what advantage shift predictions from 3D structure yielded, as compared to residue-specific, secondary-structure dependent average values reported, for example, by Wang and Jardetzky [312]. The distribution of  $C'$ ,  $C^\alpha$  and  $C^\beta$  resonances according to the four secondary structure classifications in the data base ( $\beta$ -strand ( $\beta$ ), random coil ( $c$ ),  $\alpha$ -helix ( $\alpha$ ) and membrane-integral  $\alpha$ -helix ( $m\alpha$ )) is shown in Figure 5.4. One finds that the correlation coefficient for  $C^\beta$  resonance predictions for an  $\alpha$ -helical



**Figure 5.4:** Evaluation of secondary structure-dependent  $^{13}\text{C}$  chemical shift predictions by SHIFTS, SHIFTX and SPARTA with respect to ssNMR measurements, for beta-strand ( $\beta$ ), random-coil ( $c$ ), alpha-helical ( $\alpha$ ) and membrane-integral alpha-helical ( $m\alpha$ ) residues. The random-coil value of the respective amino acid type was subtracted for each predicted or measured shift (*relative shift*).

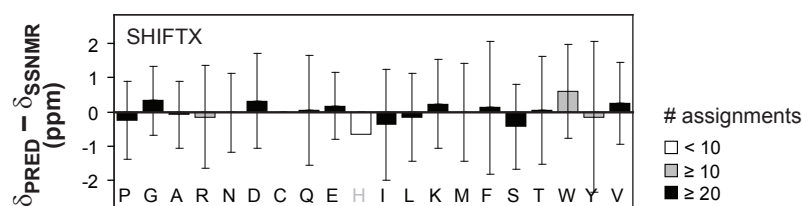
backbone ( $\alpha$ ) is lower than for extended  $\beta$ -strands. This results from a relatively small dispersion for  $C^\beta$  in helices, together with an RMSD for  $C^\beta$  that is not improved with respect to other secondary structure elements. A similar observation can be made for  $C'$



resonances. Also, correlation coefficients are smaller in  $m\alpha$  residues as compared to  $\alpha$  residues. Again, this results from a relatively low dispersion of values with respect to the prediction RMSD. On top of that, there is a trend to under-estimate large chemical shift offsets from the reference value, as can be seen from a larger number of correlation points above the diagonal for assigned shifts in the range of 5–6 ppm. This is in line with the observed misalignment of predicted and measured Ile  $C^\alpha$  shifts shown in Figure 5.1B. In order to reduce such effects, a more detailed identification of deviations is required, for example by looking at residue-specific statistics.

#### 5.4.4 Residue-Specific Analysis

The analysis can further be broken down to a residue-specific level. Residue-specific differences between predicted and observed chemical shifts are shown in Figure 5.5A for  $C^\alpha$  in SHIFTX. They are, by their magnitude, representative also for other resonances and programs, see e.g. Appendix Figure C.2. Results based on at least twenty  $C^\alpha$  assignments

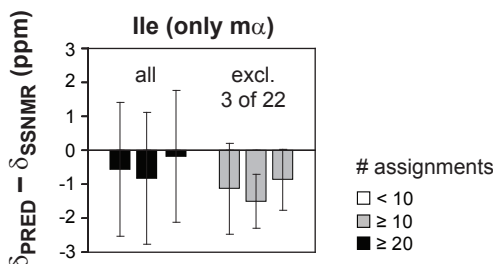


**Figure 5.5:** Residue-specific differences between predicted and observed  $C^\alpha$  chemical shifts for SHIFTX, averaging over all secondary structure elements. Black, gray and white bars indicate the availability of at least twenty, ten, or less than ten  $C^\alpha$  assignments, respectively.

are shown in black, results based on ten to twenty  $C^\alpha$  assignments in gray, others white. The average predictions scatter around mean values deviating by up to 0.8 ppm from the experimentally determined value. On the other hand, for all residues, the scatter is significantly larger than the average deviation. Particularly large deviations, such as for serine (S) in SHIFTS or tryptophan (W) in SPARTA are not found in all programs concurrently. The same is true for N,  $C'$  and  $C^\beta$  (data not shown). Deviations are generally smaller for residues with a large number of assignments, indicating that inaccuracies may average out with more data. The correction of residual inaccuracies like in Figure 5.5A could be used to improve the prediction quality.

### 5.4.5 Residue-Specific Effects from Secondary Structure and Environment

It would be beneficial for improving the prediction methods to understand if some inaccuracies can be traced back to specific groups of residues of the same kind, such as seen for membrane-integral Ile  $C^\alpha$ . A further discrimination of residue-specific shift predictions according to the four distinguished classes ( $\beta$ ,  $c$ ,  $\alpha$ ,  $m\alpha$ ), however, is difficult due to the low number of assignments available for the same residue type and structure class. In many cases, less than ten assignments were available, making a statistical assessment rather unreliable. Where at least twenty assignments per residue and structure class were present, deviations from measured values lie within the prediction error, provided the shift offset for  $^{15}\text{N}$  is adjusted for each program individually (data not shown). When at least ten, but less than twenty assignments are available, the only exception is Val  $C^\alpha(m\alpha)$  with an inaccuracy of  $-0.89$  ppm at an RMSD of  $0.63$  ppm for SHIFTX, and a similar trend observed with SHIFTS and SPARTA. This observation is based on ten assignments, and should be revised when data for other membrane-integral proteins will be available. Predictions for  $C^\alpha$  of Ile residues within membranes alone exhibit a trend



**Figure 5.6:** Inaccuracy of Ile  $C^\alpha$  predictions in membrane-integral  $\alpha$ -helices for (left to right) SHIFTS, SHIFTX and SPARTA. Results before and after exclusion of 3 outliers discussed in the text out of 22 residues are shown in black and gray, respectively.

to under-estimate the measured shift, which is most pronounced in SHIFTX, see Figure 5.6 (black bars). In average over 22 assigned residues, the SHIFTX prediction shows an offset of  $-0.83$  ppm with an RMSD of  $1.95$  ppm. This, however, is not representative of the Ile  $C^\alpha$  mismatch exemplified in Figure 5.1, wherein the scatter is clearly smaller than the mean offset. Interestingly, the prediction RMSD is more than three times larger than for Gly and Ala, where a large number of assignments is available, too. A detailed look into the data reveals that there are just three residues that do not follow the general trend: I26 in LH2 subunit  $\alpha$ , I16 in LH2 subunit  $\beta$  and I33 in AFA-PLN shift predictions

are off by 2.3, 3.2 and 4.8 ppm. In contrary, all other membrane-integral Ile C $\alpha$  shifts are off by about  $-2$  ppm, including LH2 and AFA-PLN predictions (SHIFTX). These prediction outliers are corroborated by SHIFTS and SPARTA. The input structures do not show differences to regular helices, hence these exceptions may point to interactions with other molecules, such as cofactors in LH2, or lipid head groups in case of AFA-PLN, provided the assignment is correct. Exclusion of these three residues from the data base finally leads to an average prediction inaccuracy of  $-1.50$  ppm compared to a relatively high precision of 0.80 ppm for the remaining 19 assignments, see Figure 5.6 (gray bars). The same trend can be observed for SHIFTS and SPARTA.

#### 5.4.6 Joining Forces

Last, one might ask the question if the performance of shift predictions can be improved by combined analysis using all three programs together. Indeed, averaging over SHIFTX and SPARTA predictions slightly improves the results: The number of outliers beyond a threshold of more than 6 ppm for N and 3 ppm for  $^{13}\text{C}$ , as described in Table 5.2, drops to (6, 4, 4, 5) %. The RMSD is slightly decreased for  $^{15}\text{N}$  and constant for  $^{13}\text{C}$  with (3.0, 1.4, 1.3, 1.6) ppm, while  $R$  remains approximately constant at (0.83, 0.71, 0.85, 0.71). In the end, the gain is rather marginal, but indicative for possible further improvements in prediction methodology.

### 5.5 Conclusions

Protein NMR chemical shift prediction computer programs trained with solution-state NMR chemical shifts and solution-state NMR or X-ray structural data are a valid tool for predicting the chemical shift measured by ssNMR spectroscopy. The RMS prediction error is found to be in the order of 3 ppm for  $^{15}\text{N}$  and 1.5 ppm for  $^{13}\text{C}$ , with a chance of 93% and 95% of finding the predicted shift within an interval of 6 and 3 ppm around the experimentally determined value for  $^{15}\text{N}$  and  $^{13}\text{C}$ , respectively. Including outliers, predictions are about 1/3 less precise than for proteins in solution. The highest precision is achieved for C $\alpha$ . Correlation coefficients are influenced by the kind of resonance, secondary structure type, and also by the molecular environment, with a tendency to under-estimate large chemical shift offsets from the residue-specific average. Given a sufficiently large number of assignments, none of the three considered computer programs shows prediction inaccuracies that exceed the prediction RMSD. Unusual deviations from experimental values

can occur in certain combinations of residue type and secondary structure or molecular environment, which should be reconsidered when more proteins will have been assigned. In case of Ile CA in membrane-integral,  $\alpha$ -helical residues, the deviations of SHIFTX predictions from measured chemical shifts become significant when three unusual outliers are excluded from the data base, and a clear trend of under-predicting this shift is then also visible in SHIFTS and SPARTA. These predictions may be manually corrected. Methodological improvement should be achievable. The use of SPARTA or SHIFTX instead of SHIFTS is encouraged, where SPARTA is found to be slightly more precise and reliable. On the other hand, when speed is a critical objective, such as applications within a structure calculation framework, SHIFTX provides optimum performance per time.

## Chapter 6

# Studying Molecular 3D Structure and Dynamics by High-Resolution Solid-State NMR: Application to L-Tyrosine-Ethylester

## 6.1 Summary

A unified approach to the study of 3D conformation and molecular dynamics using magic-angle-spinning solid-state NMR is demonstrated on a uniformly  $^{13}\text{C}$ -labeled sample of L-tyrosine-ethylester.

## 6.2 Introduction

Solid-state nuclear magnetic resonance (NMR) is an ideal tool to study molecular structure in systems that are insoluble or do not exhibit long-range structural order [40, 184, 297]. Techniques proposed to probe three-dimensional molecular structure in high spectral resolution include the combined measurements of individual internuclear distances and molecular torsion angles [73, 95, 172], the structural analysis based on the use of several multiply labeled samples [46, 61, 133, 134, 216, 228, 240], and the determination of 3D molecular structure from a single (uniformly isotope-labeled) sample [160, 161, 212].

In addition to the study of molecular structure, solid-state NMR also provides unique means to probe molecular dynamics on a wide time scale [218], including applications to probing chain order in lipids [250] and polymers [247], and the investigation of backbone and side-chain motion in (membrane) proteins [145, 174, 286]. Many applications have relied on studying anisotropic NMR interactions to monitor dynamics or have invoked relaxation measurements that are capable of probing molecular dynamics without inter-

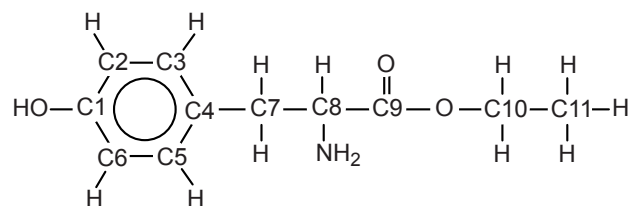
ference effects from the overall (isotropic) tumbling of the molecule in the solution state. Such studies, for example, can deliver order parameters to describe segmental motion in terms of amplitude fluctuations and may be used to study aromatic ring flips, side-chain motion in polypeptides, methyl group rotations, or structural phase transitions.

As has been known for a long time (see, for example, ref. [85] and references therein), and recently shown for the case of L-tyrosine-ethylester (TEE) [114], dynamic disorder affects many solid materials, including those made from small molecules. Small molecule crystalline solids such as TEE thus are ideal test systems to establish general concepts for monitoring 3D structure and dynamics using a single (uniformly) labeled sample. In this chapter, it is demonstrated that solid-state NMR methods tailored to 3D structure determination of uniformly labeled biomolecules under magic-angle spinning (MAS) [4] conditions in conjunction with techniques that probe local motion in a site-resolved manner can provide complementary insight into the details of molecular structure and dynamics in solid-phase systems. Such studies not only may aid the structural characterization of (bio)polymers and (membrane) proteins, but also may be used to study the relationship between free and receptor-bound conformations of a variety of pharmacophores of small molecular weight [185].

## 6.3 Material and Methods

### 6.3.1 Sample Preparation

All solid-state NMR experiments discussed in the following were conducted on a diluted, uniformly  $^{13}\text{C}$ -labeled sample of L-tyrosine-ethylester (TEE), see also refs. [114, 230] (Figure 6.1). 3D structural information was obtained using an NMR sample diluted at



**Figure 6.1:** Schematic representation of the L-tyrosine-ethylester (TEE) molecule according to nomenclature used in ref. [114].

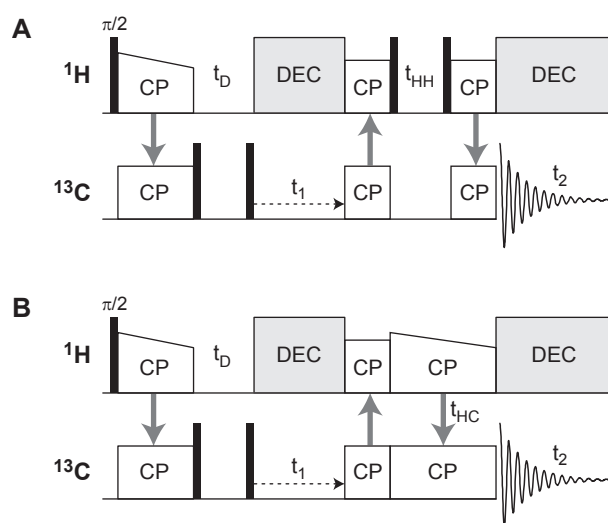
10% in unlabeled TEE ( $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$ ). For this purpose, 4.5 mg of  $\text{U}(^{13}\text{C})\text{-TEE}$  and

45 mg of TEE with  $^{13}\text{C}$  in natural abundance were dissolved in ethanol. After overnight evaporation of the solvent, the sample was ground and dried in a vacuum desiccator prior to transferring it into a standard 4 mm MAS rotor.

### 6.3.2 NMR Experiments

Two-dimensional ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) CHHC and CHC correlation experiments were carried out on a Bruker Avance 600 (Bruker Biospin, Germany) spectrometer at 14.1 T equipped with a standard 4-mm Bruker MAS probe, at a spinning frequency of 9.5 kHz. Rotational resonance in the tilted rotating frame ( $\text{R}^2\text{TR}$ ) and Lee-Goldburg cross polarization (LG-CP) buildup experiments were performed on an Avance 400 spectrometer at 9.4 T using a 4 mm rotor, at spinning frequencies described below. For evolution and detection periods, TPPM [28] proton decoupling was used.

Two-dimensional CHHC correlation spectroscopy was done according to methods described earlier [159, 160] using longitudinal (0Q)  $^1\text{H}$ - $^1\text{H}$  mixing in the laboratory frame, see also Figure 6.2A.  $^{13}\text{C}$ - $^1\text{H}$  transfer steps bracketing the ( $^1\text{H}$ ,  $^1\text{H}$ ) mixing unit were set to



**Figure 6.2:** Two-dimensional pulse schemes used for measuring ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) encoded (A)  $^1\text{H}$ - $^1\text{H}$  (CHHC) [159, 160] or (B)  $^1\text{H}$ - $^{13}\text{C}$  (CHC) [80, 157] polarization transfer. After an initial CP transfer and a z-filter to remove any residual  $^1\text{H}$  magnetization,  $^{13}\text{C}$  chemical shifts are recorded along the  $t_1$  evolution time. Following a short CP back-transfer step (175  $\mu\text{s}$ ), polarization transfer is monitored in two different ways, (A) for a variable  $^1\text{H}$ - $^1\text{H}$  mixing time  $t_{\text{HH}}$  with subsequent, short CP transfer or (B) with variable  $^1\text{H}$ - $^{13}\text{C}$  mixing time  $t_{\text{HC}}$ , and finally recorded in  $t_2$ . Phases and phase cycles given in Appendix Chapter I, including a phase cycle that allows to run the experiment without z-filter.

125  $\mu\text{s}$ . CHHC experiments were conducted on  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$  at 261 K for 75, 150, 250, and 500  $\mu\text{s}$   $^1\text{H}\text{-}^1\text{H}$  mixing times, with 312 points sampled in the indirect dimension, and 288, 128, 128, and 256 transients acquired per evolution period, respectively. At higher temperature (318 K), three 2D CHHC spectra with a  $^1\text{H}\text{-}^1\text{H}$  mixing time of 250  $\mu\text{s}$  and CHHC cross polarization (CP) contact times of 175  $\mu\text{s}$  were recorded over a time of 3.5 days and added together.

In Figure 6.2B, the pulse scheme for a 2D CHC correlation experiment is shown. After an initial cross-polarization (CP) [109] transfer and a  $z$ -filter to remove any residual  $^1\text{H}$  magnetization,  $^{13}\text{C}$  chemical shifts are recorded along the  $t_1$  evolution time. Note that in alternative phase cycles, the  $z$ -filter can be omitted [113] and Appendix Chapter I. Next, a short CP step ensures that only  $^1\text{H}$  magnetization of  $\text{C}(\text{H})_n$  groups is created. Establishing  $(^1\text{H},^{13}\text{C})$  CP for a variable mixing time  $t_{\text{HC}}$  allows for transfer of polarization to nearby  $^{13}\text{C}$  spins, that is read out along  $t_2$ . The second CP contact was established for 175  $\mu\text{s}$ , while the final CP mixing unit was set to 1.5 ms (see also refs [80, 157]). Note that the pulse scheme can be modified to include a homonuclear  $(^1\text{H},^1\text{H})$  decoupling unit during  $(^1\text{H},^{13}\text{C})$  CP [157].

Rotational resonance in the tilted rotating frame ( $\text{R}^2\text{TR}$ ) [59, 266, 281] was used to probe the C11–C2 distance in  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$ . Two-dimensional constant-time experiments were performed for mixing times of 3.1, 5.2, and 7.5 ms at a MAS frequency of 10790 Hz. The  $\text{R}^2\text{TR}$  condition for the chemical shift difference of 10280 Hz was established by an additional low-power pulse of 1.64 kHz, framed by short (35  $\mu\text{s}$ ) ramp-in and ramp-out pulses [59]. Another set of  $\text{R}^2\text{TR}$  spectra for shorter mixing times of 0.8 and 2.6 ms was recorded for the spin pair C4–C7, which served to calibrate parameters for the quantum mechanical simulation of the C11–C2 cross-peak buildup (see below).

Lee–Goldburg [169] cross-polarization [109] (LG-CP) transfers were set up using adamantane following established procedures [32]. LG-CP buildup measurements were performed at a MAS frequency of 10900 kHz in order to avoid rotational resonance [236] conditions. The applied field strength during the LG-CP was  $\omega_{^1\text{H},\text{eff}}/2\pi = 48$  kHz, and  $\omega_{^{13}\text{C},\text{eff}}/2\pi = 37$  kHz.

### 6.3.3 Quantum Mechanical Calculations

Quantum mechanical calculations using the C++ software library GAMMA [262] were carried out to simulate the cross-peak buildup for the C11–C2 spin pair in  $\text{R}^2\text{TR}$  experiments [266] and for the analysis of the LG-CP experiments. As noted elsewhere [160], a



quantum mechanical analysis does not always allow to adequately describe the details of proton–proton interactions in the solid state. In the R<sup>2</sup>TR experiments, the relevant spin system comprises the nuclear spins of interest (C11, C2), as well as all directly bonded <sup>13</sup>C nuclei (C10, C3, C1), and additionally all <sup>1</sup>H spins directly bonded to the aforementioned five carbon atoms [266]. Methylene and methyl proton groups were represented by single spins assuming a one-bond <sup>13</sup>C–<sup>1</sup>H distance of 1.09 Å. Test calculations in smaller spin systems (data not shown) reveal that proton–proton interactions within the C10 methylene group, that were not taken into account due to reasons of computational time, are of minor influence. Simulations under LG-CP conditions were conducted using RF fields as given by the experiment. In the case of nonuniaxial motion, the apparent C–H coupling under Lee-Goldburg irradiation depends on the motionally scaled dipolar coupling, and the details of the transfer profile are affected by a nonzero asymmetry parameter [119]. However, the polarization transfer dynamics are also influenced by multiple-spin effects and experimental aspects such as RF inhomogeneity and pulse imperfections. Hence, only an overall phenomenological scaling parameter  $S$  of the dipolar coupling  $D_{\text{eff}} = S \cdot D_{\text{CH}}$  with  $D_{\text{CH}} = \mu_0 \gamma_{\text{C}} \gamma_{\text{H}} \hbar / 4\pi r_{\text{CH}}^3$  (see also ref. [21]) was used to reproduce the experimentally detected buildup characteristics. 2-, 3-, and 4-spin systems were considered to simulate the polarization transfer dynamics of CH, CH<sub>2</sub>, and CH<sub>3</sub> groups, respectively. While one-bond <sup>13</sup>C–<sup>1</sup>H distances were set to 1.09 Å, <sup>1</sup>H–<sup>1</sup>H distances of 1.75 Å and 1.63 Å were assumed for CH<sub>2</sub> and CH<sub>3</sub> groups, respectively. For the simulation of the methyl group (C11), an additional  $D_{\text{CH}}$  scaling factor of  $S_{\text{CH}} = 1/3$  was introduced [206].

### 6.3.4 Structure Calculation

Model structures were calculated using CNS [42]. Internuclear distances obtained from CHHC and R<sup>2</sup>TR experiments were incorporated in the structure calculation by square-well distance restraints [160]. Ambiguous <sup>1</sup>H–<sup>1</sup>H or <sup>1</sup>H–<sup>13</sup>C distance restraints from methylene and methyl protons were accounted for by a sum averaging over all possible contacts. An elongated conformer of TEE was generated starting from the tyrosine amino acid parameter set in the PROTEIN-ALLHDG parameter files supplied with CNS version 1.1 [209]. In the generated starting structure, the distance between C11 and C1 amounts to 8.3 Å. A set of 50 conformers was calculated using a simulated annealing protocol consisting of three stages: (1) high-temperature annealing in torsion angle space, in 2000 time steps of 0.002 ps at 30000 K; (2) slow-cool annealing stage in torsion angle space, in 6000 steps of 0.002 ps, and temperature reduction from 30000 K to zero in steps of 250 K; (3)

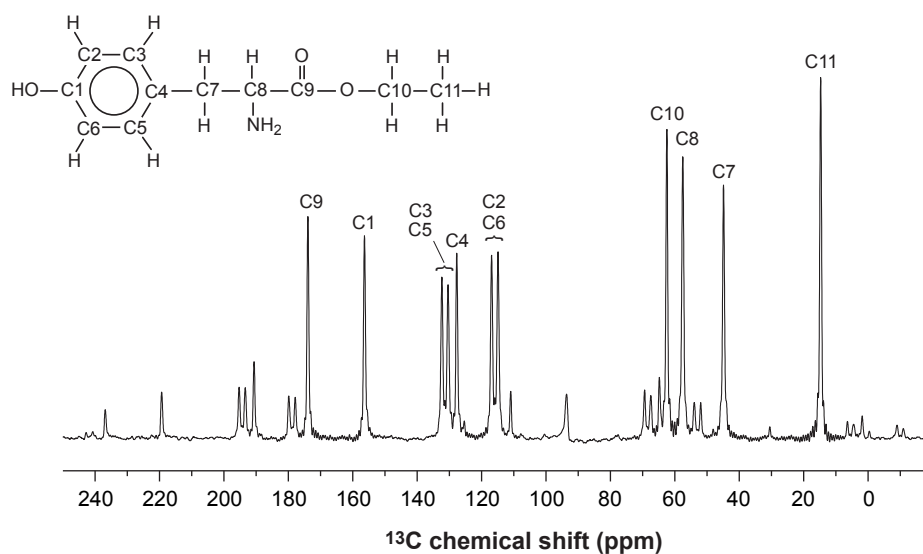
final conjugate gradient minimization in 20 cycles of 200 steps each. Distance restraints were invoked by force constants of  $150 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-2}$  during all three annealing stages. The resulting structures were sorted by total energy. Sets of 15 out of 50 structures were chosen to represent the molecular structure of TEE.

Detailed restraint classifications are given in Appendix Chapter D. Compared to our earlier model study on a tripeptide [160] where results of one 2D correlation experiment were used in a simplified, dual constraint classification, knowledge of the CHHC buildup characteristics here allows for a further refinement of the relationship between CHHC transfer efficiency and proton–proton distance. Three intervals were established corresponding to strong, medium, and weak transfer efficiencies. To reflect missing cross peaks, an additional interval was introduced. For cross-validation, the final set of structures was used to compute transfer CHHC profiles using the relaxation model introduced in ref. [160]. As shown in Appendix Chapter D, these transfer curves compare favorably to the experimental results.

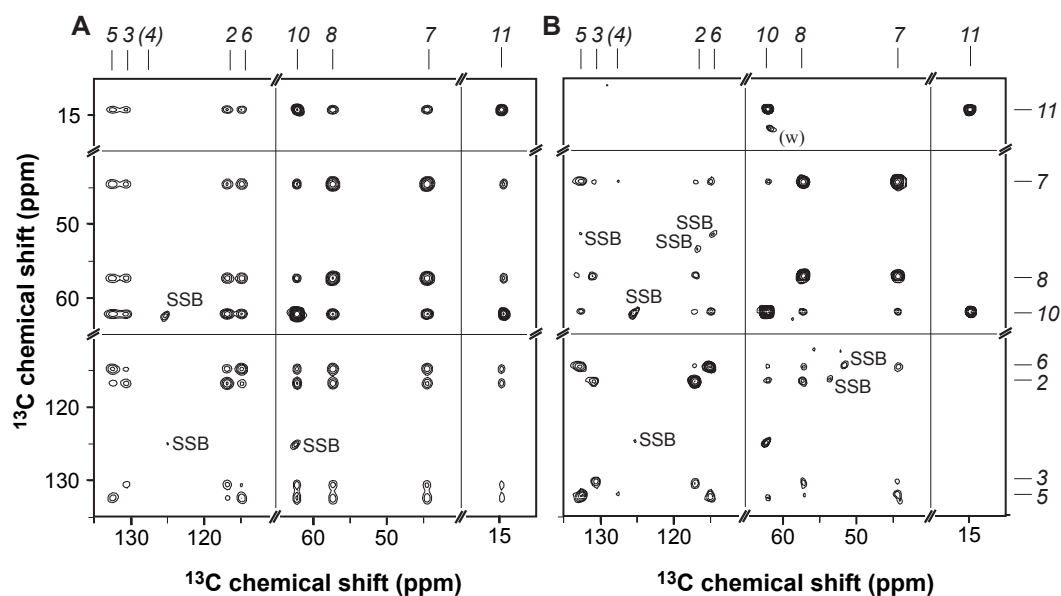
## 6.4 Results and Discussion

### 6.4.1 Low-Temperature Analysis of TEE

Earlier variable temperature  $^{13}\text{C}$  MAS NMR experiments [114], see also appendix Appendix Chapter D for a reproduction of these experiments, indicate that TEE exhibits phenyl ring  $\pi$ -flips with an activation energy of  $E_a = 50 \pm 12 \text{ kJ} \cdot \text{mol}^{-1}$  at a sufficiently small flip rate at temperatures  $T < 270 \text{ K}$ . A 1D CP spectrum of  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$  recorded at 261 K with 9.5 kHz MAS is shown in Figure 6.3. Resonances from C3 and C5 as well as C2 and C6 are clearly separated, though they cannot be unambiguously assigned at this stage. Accordingly, CHHC spectra were obtained at 261 K for a  $^1\text{H}\text{-}^1\text{H}$  mixing time of 250  $\mu\text{s}$ . These are depicted in Figure 6.4, comparing results obtained on  $\text{U}(^{13}\text{C})\text{-TEE}$  (A) and on  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$  (B). As expected for the CHHC concept, resonances of the nonprotonated C1, C4, and C9 carbon atoms are strongly suppressed in all CHHC correlation data discussed in the following. On the other hand, a variety of  $(^{13}\text{C}, ^{13}\text{C})$  cross peaks mediated by proton–proton interactions can be readily detected. While in Figure 6.4A correlations are seen among almost all  $^{13}\text{C}$  resonances, a distinct set of  $(^{13}\text{C}, ^{13}\text{C})$  correlations is visible in Figure 6.4B. Correspondingly, the data in Figure 6.4A could be further used to monitor intermolecular interactions [62, 299]; in the following, results obtained on  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$  will serve to study the structure and dynamic properties of the molecular moiety in solid

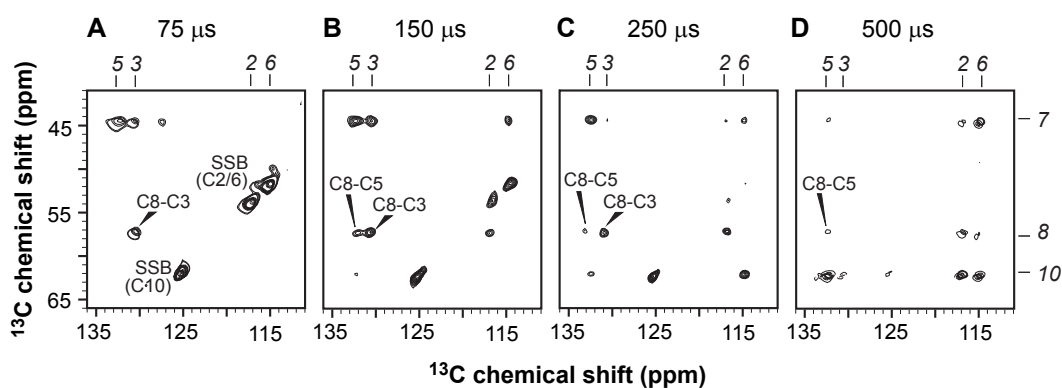


**Figure 6.3:** 1D <sup>1</sup>H-<sup>13</sup>C cross-polarization <sup>13</sup>C spectrum of U(<sup>13</sup>C)-TEE<sub>dil</sub> recorded at 261 K with 9.5 kHz magic angle spinning. Resonance labels refer to the eleven isotropic <sup>13</sup>C resonances; additionally, spinning side bands are present.



**Figure 6.4:** CHHC spectrum for a <sup>1</sup>H-<sup>1</sup>H mixing time of 250 μs of (A) U(<sup>13</sup>C)-TEE and (B) U(<sup>13</sup>C)-TEE<sub>dil</sub> obtained on a 600 MHz instrument. Spinning sidebands (SSB) are indicated.

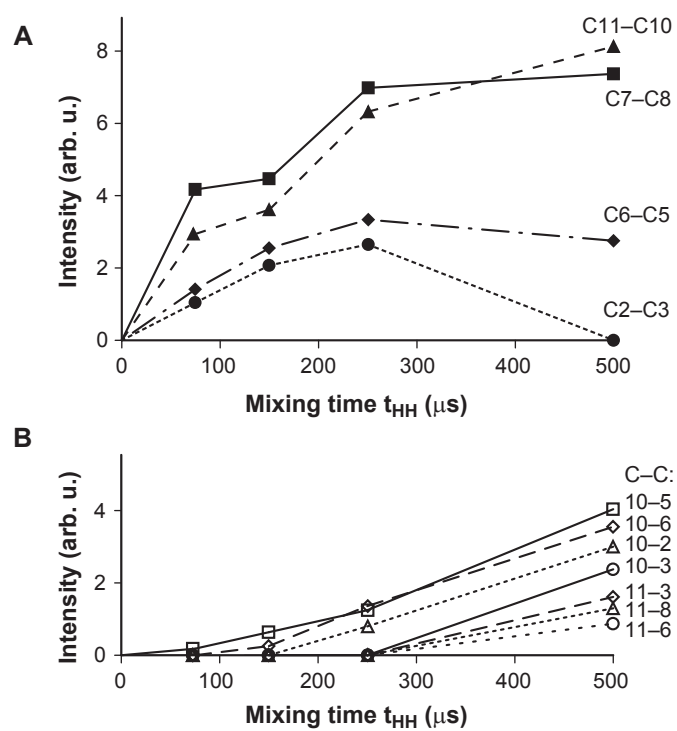
TEE. As discussed in further detail elsewhere [160], structural information can be obtained from CHHC 2D data if cross-peak intensities are evaluated in the initial rate regime and by analyzing transfer data obtained from several mixing times. For this reason, a series of CHHC experiments with mixing times of 75–500  $\mu\text{s}$  were conducted (see Material and Methods section). As an example, Figure 6.5 shows data involving C8-to-ring correlations highlighting the qualitative difference between correlations involving C3 and C5: cross peaks reflecting C8–C3 interactions build up faster than those originating from C8–C5 interactions, suggesting that the  $\text{H}_{\text{C8}}\text{--H}_{\text{C3}}$  distance must be significantly shorter than the  $\text{H}_{\text{C8}}\text{--H}_{\text{C5}}$  distance. In contrast to broadband ( $^{13}\text{C}, ^{13}\text{C}$ ) correlation experiments, indirect



**Figure 6.5:** Series of CHHC spectra with different mixing times, showing a spectral region with cross peaks between aliphatic and aromatic carbon resonances.

detection of  $^1\text{H}\text{--}^1\text{H}$  interactions hence allows for the spectral identification of each ring carbon. As discussed in further detail in ref. [114], total through-space ( $^{13}\text{C}, ^{13}\text{C}$ ) correlation spectroscopy [22, 24] does not provide a definitive answer as to whether TEE adopts an elongated or 'scorpion-like' structure in the polycrystalline state. For this aspect,  $^1\text{H}\text{--}^1\text{H}$  contacts mediating ( $^{13}\text{C}, ^{13}\text{C}$ ) correlations of C10 and C11 with ring carbons are of particular interest. As can be seen from Figure 6.4B, patterns correlating C10 with nuclei in the ring (C2,3,5,6) can be readily identified for short mixing times, consistent with proton–proton contacts over a length in the order of 3 Å. Since this data is obtained for the diluted sample  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$ , these interactions must reflect intramolecular interactions and thus provide information about the molecular conformation.

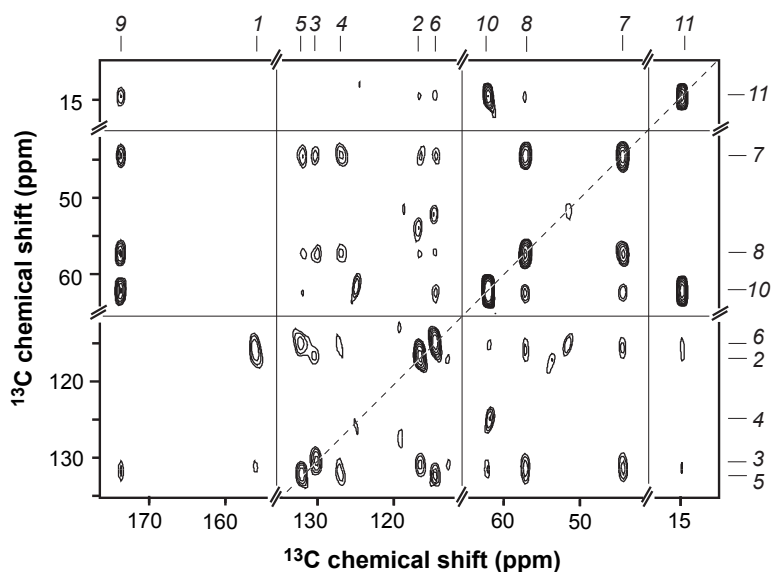
A series of buildup curves for one-bond and ethylester–phenyl-ring cross-peak intensities is displayed in Figure 6.6. The data were obtained by dividing the sum of the cross-peak volumes for each pair of spins by the sum of the corresponding diagonal peak



**Figure 6.6:** Intensity buildup of selected CHHC cross-peaks for (A) one-bond and (B) ester-tail-ring  $^1\text{H}$ - $^1\text{H}$  interactions.

volumes as a function of mixing time. Directly bonded  $^{13}\text{C}$  nuclei show the strongest  $^1\text{H}$ – $^1\text{H}$  polarization transfer (such as C7–C8), and buildup curves for these nuclei agree well with the CHHC transfer characteristics studied in other model compounds under comparable experimental conditions [160]. Interestingly, the buildup of the one-bond C10–C11 correlation shown in Figure 6.6B is slightly slower than for all other one-bond interactions considered in Figure 6.6A. This could be consistent with a reduction in  $(^1\text{H}, ^1\text{H})$  polarization transfer efficiency due to the methyl group rotation. Notably, earlier rotational resonance ( $\text{R}^2$ )  $^{13}\text{C}$  experiments concerning the C10–C11 pair of spins in  $\text{U}(^{13}\text{C})\text{-TEE}$  also indicate a seemingly reduced effective C10–C11 dipolar coupling constant [74]. According to Figure 6.6B, a variety of C10–phenyl-ring correlations can be observed at  $t_{\text{HH}} \geq 150 \mu\text{s}$ . On the other hand, C11-phenyl-ring correlations are in general weaker than those involving C10 and start to appear only for a mixing time of 500  $\mu\text{s}$ . As discussed in further detail in [160], cross-peak intensities detected for such long mixing times may reflect relay transfer involving more than two proton spins, making a direct structural analysis difficult. We also observe notable differences regarding  $(^1\text{H}, ^1\text{H})$  correlations among phenyl-ring carbon atoms. Nearest-neighbor interactions, i.e., C2–C3 and C5–C6, are visible for very short proton–proton mixing times, see Figure 6.4B. C3–C5 and C2–C6 correlations appear for  $t_{\text{HH}} \geq 150 \mu\text{s}$ . C3–C6 and C2–C5 interactions are not detected even for a mixing time of 500  $\mu\text{s}$ . The latter findings fully agree with the spectral assignments used in Figure 6.4 and Figure 6.5 that would require polarization transfer across the ring diagonal (i.e., 6 Å, Figure 6.1), which is strongly attenuated for the considered mixing times.

An independent means of probing molecular structure is brought about by measuring  $(^1\text{H}, ^{13}\text{C})$  correlations in two spectral dimensions. For this purpose, the N/CHHC polarization transfer concept [159] can be modified by removing the proton–proton mixing block and using the final CP polarization block as a variable  $(^1\text{H}, ^{13}\text{C})$  transfer unit [157], Figure 6.2. The corresponding NHC transfer scheme has recently been used for measuring intermolecular transfer in molecular mixtures [80]. Figure 6.7 shows that a mixing time  $t_{\text{HC}}$  of 1.5 ms delivers a multitude of through-space interactions in a CHC experiment applied to  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$ . In addition to CHHC correlations involving two protonated carbon atoms, correlations relating to C1, C4, and C9 resonances can now be observed. For the mixing time used, correlations of C1 with other nuclei are seen only within the phenyl ring, with more intense cross-peaks to C2 and C6 than for C3 and C5. On the other hand, a variety of three-bond correlations such as (C11,C9), (C10,C8), and (C4,C8) are detected. A cross peak between the resonances of non-protonated carbons C4 and C9,

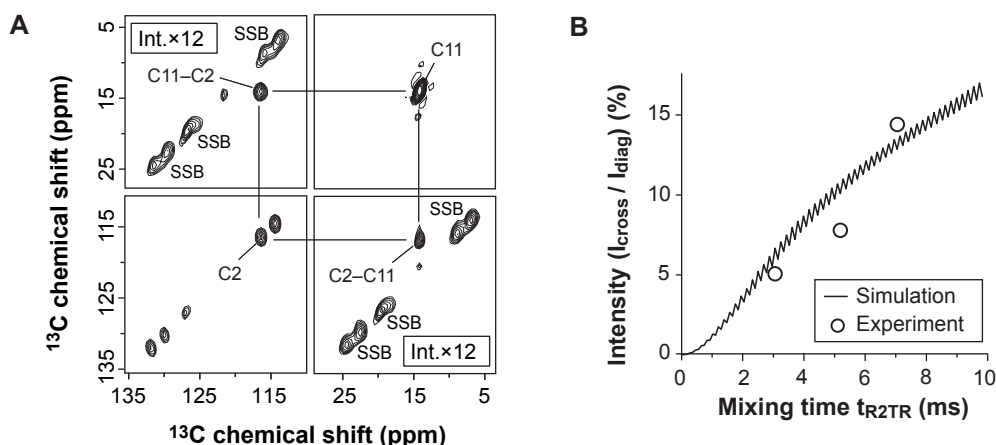


**Figure 6.7:** CHC spectrum obtained on  $U(^{13}\text{C})\text{-TEE}_{\text{dil}}$  at 261 K.  $t_{\text{HC}}$  was set to 1.5 ms.

also within a range of three bonds, is absent, as expected. Cross peaks relating the ester tail to the ring reveal qualitative agreement with the structural constraints derived from the CHHC correlations. As shown in [80], CHC-type transfer is, however, also influenced by the MAS rate. For this reason, a direct structural interpretation of the CHC correlation was not attempted here.

To further elucidate the structural topology involving the (C11,C2) pair of spins, 2D constant-time [266]  $\text{R}^2\text{TR}$  [281] experiments using mixing times of 3.1, 5.2, and 7.5 ms were conducted. In Figure 6.8A (mixing time 5.2 ms), cross-peaks can be readily identified without spectral overlap with the resonances of the pair of spins (C11, C2). The measured cross-peak buildup was further compared to a quantum mechanical calculation, with the C11–C2 bond distance amounting to 4.9 Å, Figure 6.8B (*results provided by L. Sonnenberg*). As discussed in more detail elsewhere [266], the accuracy of the distance determination, however, strongly depends on experimental conditions, in particular on the proton decoupling field strength during ( $^{13}\text{C},^{13}\text{C}$ ) mixing.

The distance restraints obtained from the analysis of our CHHC and  $\text{R}^2\text{TR}$  experiments were finally utilized in a standard 3D structure calculation within CNS. As shown in Figure 6.9, the use of CHHC distance restraints leads to a reduction of the available conformational space of TEE (Figure 6.9A) and firmly establishes the scorpion-like fold of the molecule in the crystalline state (Figure 6.9B). One  $\text{R}^2\text{TR}$  distance restraint (Figure



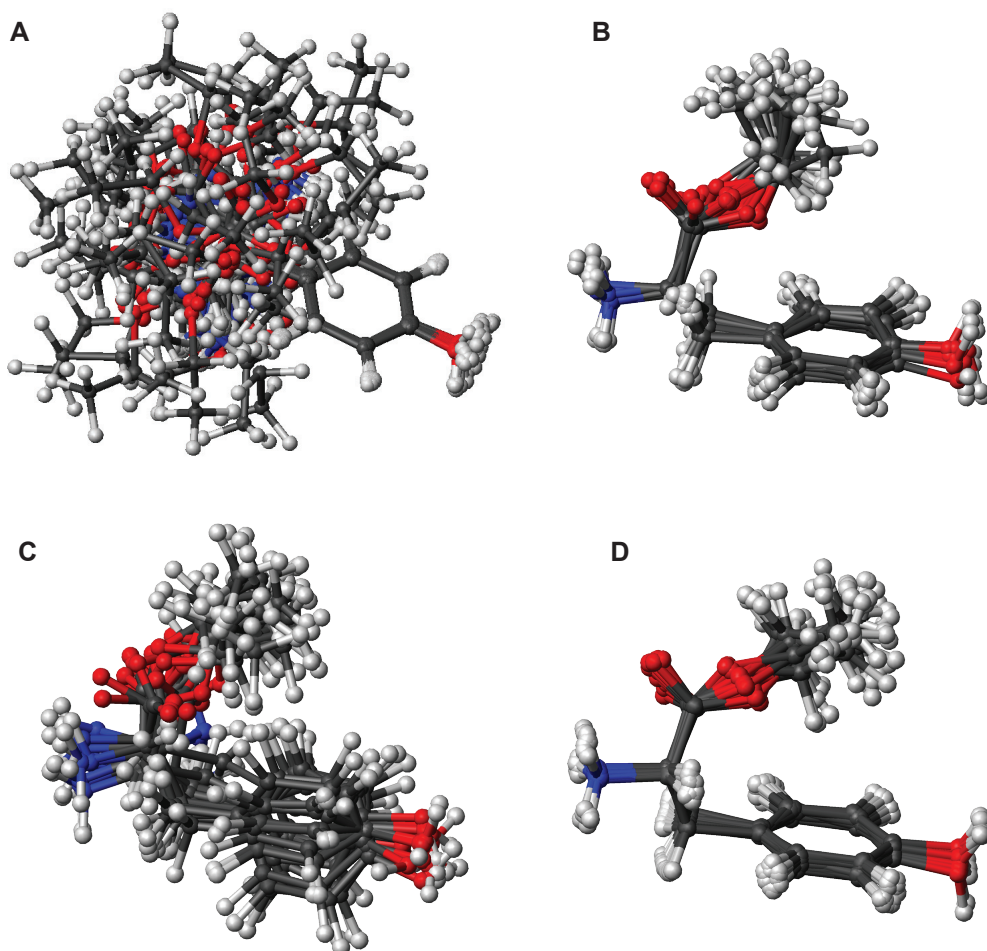
**Figure 6.8:**  $R^2\text{TR}$  experiments on  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$ . (A) Close-up of the spectral region of the C11–C2 diagonal and cross-peaks for a  $R^2\text{TR}$  experiment with a 5.2 ms mixing time. Cross peaks are separated by 510 Hz relative to the spinning sidebands (SSB). For the sake of clarity, intensities in the cross-peak region are scaled up by a factor of twelve. (B) Comparison of measured cross-peak intensities for three different mixing times (open circles) with simulated cross-peak buildup (solid line).

6.9C) alone is, on the other hand, insufficient to adequately describe the molecular structure of TEE at low temperatures, but it helps to further refine the CHHC-based structural model of TEE (Figure 6.9D).

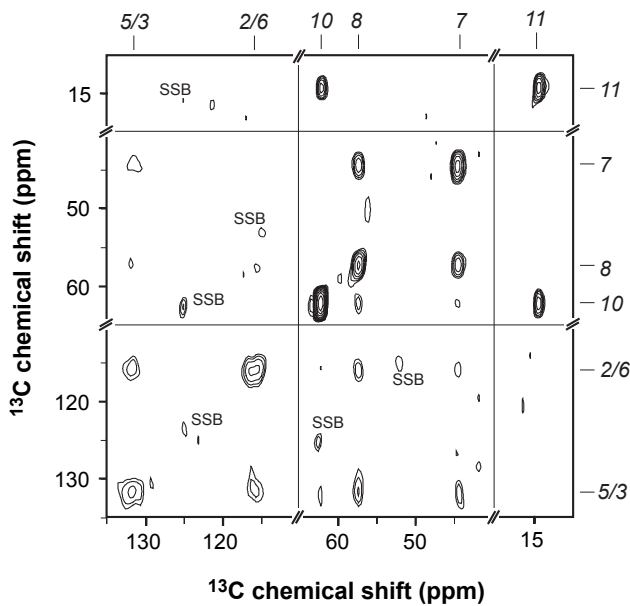
#### 6.4.2 Solid-State NMR Analysis at Higher Temperature

The phenyl ring in solid TEE undergoes  $\pi$ -flips at higher temperatures, see Appendix Chapter D; as a result of this, chemical-exchange broadening prevents the spectral separation of the C3/5 and  $^{13}\text{C}2/6$  resonances [114] at temperatures  $T > 318$  K. To investigate the influence of molecular dynamics upon  $^1\text{H}\text{-}^1\text{H}$  transfer, a CHHC spectrum of  $\text{U}(^{13}\text{C})\text{-TEE}_{\text{dil}}$  was recorded with a mixing time of 250  $\mu\text{s}$  at 318 K (Figure 6.10). Similar to Figure 6.4B, correlations of the phenyl-ring carbon resonances to C7, C8, and C10 resonances can be observed, albeit with an overall lower signal-to-noise ratio. Again, C11 shows cross peaks only with C10 for the given signal-to-noise ratio. For a more detailed analysis, we compare in Figure 6.11 the normalized CHHC cross-peak intensities at low (261 K) and high (318 K) temperatures. Although the overall signal intensity is decreased, the general transfer behavior is similar, suggesting that the overall structure (defined by nearest-neighbor proton–proton interactions) is preserved at higher temperature. Notably,

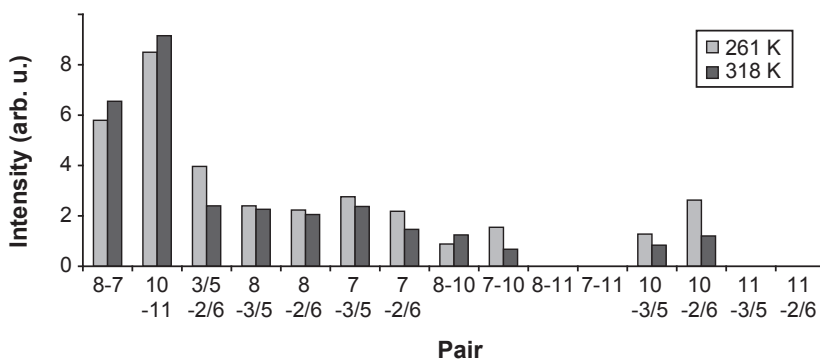




**Figure 6.9:** Three-dimensional structural models of TEE. (A) Complete ensemble of 50 structures calculated without additional experimental distance restraints. Structures do not differ significantly in total energy. (B) Ensemble of 15 out of 50 calculated structures representing TEE as seen by CHHC spectroscopy at low temperatures. (C) Ensembles with 20 out of 50 structures for TEE calculated using one C11–C2 distance restraint from R<sup>2</sup>TR alone. (D) Structure calculation using a combination of CHHC restraints at 261 K and one R<sup>2</sup>TR distance restraint.



**Figure 6.10:** CHHC spectrum for a  $^1\text{H}$ - $^1\text{H}$  mixing time of  $250 \mu\text{s}$  at a temperature of 318 K. Resonance lines from exchanging ring sites overlap.

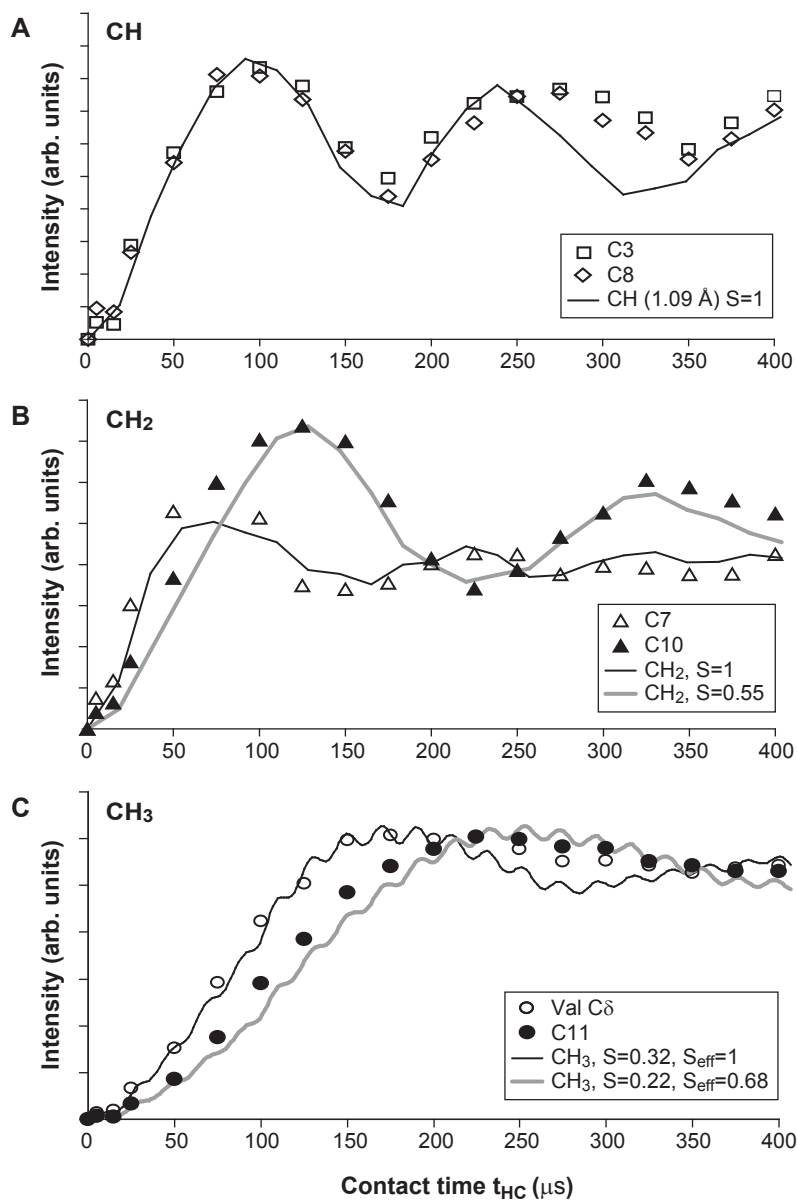


**Figure 6.11:** CHHC cross-peak amplitudes for the different TEE spin pairs for experiments conducted at two different temperatures. Intensities X/Y given at low temperatures reflect the sum of signal amplitudes measured for resonances X and Y. All cross-peak intensities were normalized to the total intensity of the 2D spectrum.

this conclusion would also be consistent with a ring flip that, as predicted in ref. [114], would occur on a time scale significantly larger than the inverse of the polarization transfer rate for proton–proton distances in the order of 3.5 Å.

### 6.4.3 Molecular Dynamics As Seen by LG-CP

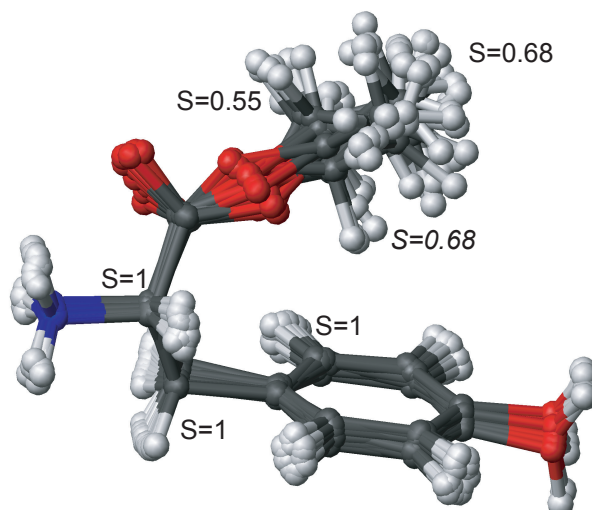
So far, the structural characterization of TEE was based on the detection of ( $^1\text{H}, ^1\text{H}$ ) contacts and ( $^{13}\text{C}, ^{13}\text{C}$ ) distance constraints that report on the 3D molecular arrangement. The underlying dipolar coupling interactions depend on the through-space distance but can be modulated by molecular motion. As an independent means to study molecular dynamics, solid-state NMR experiment that probe one-bond ( $^{13}\text{C}, ^{13}\text{C}$ ) and ( $^1\text{H}, ^{13}\text{C}$ ) dipolar interactions at variable temperatures may be employed. Since one-bond distances are well-known for both interactions, measurements of the corresponding effective dipolar coupling constants hence can reveal the effects of dynamic averaging. To suppress unwanted signal modulations due to ( $^1\text{H}, ^1\text{H}$ ) interactions during  $^1\text{H}$ – $^{13}\text{C}$  transfer, a variety of homonuclear decoupling schemes have been suggested [45, 99, 304]. In the following, the measurement of cross polarization  $^{13}\text{C}$  buildups under Lee-Goldburg [119, 169, 304] conditions will be employed. (*LG-CP data contributed by M. Etzkorn*) [255]. In Figure 6.12, results are shown for C8, C5 (A, CH), C7, C10 (B,  $\text{CH}_2$ ), and C11 (C,  $\text{CH}_3$ ) resonances as a function of the LG-CP buildup time. As discussed in Section 6.3.3, numerical simulations were conducted for systems of 2–4 spins with RF fields and resonance offsets as employed in the experiments. In Figure 6.12A, the one-bond C–H spin dynamics are well reproduced, assuming a dipolar order parameter  $S = 1$  for C8 and C5. The latter observation is in accordance with a general notion that  $\pi$ -ring flips in TEE at lower temperatures occur on a time scale larger than the (inverse) C–H dipolar interaction. At higher temperatures, the C–H coupling of ring carbon atoms can be further modulated [85] and a nonuniaxial motion must be considered [119]. By contrast, the CP buildup characteristics of the C10 resonance can only be reproduced using scaling factors  $S_{\text{CH}} = S_{\text{HH}} = 0.55$ . A similar scaling factor ( $S_{\text{eff}} = 0.68$ ) is found for the C11 spin if an overall methyl group rotation is taken into account. These results hence indicate an increase in molecular mobility for the ester tail, in agreement with results of broadband ( $^{13}\text{C}, ^{13}\text{C}$ ) 1Q correlation spectroscopy [114], and lead to a molecular order parameter for the C10–C11 dipolar coupling tensor of 0.68.



**Figure 6.12:** LG-CP buildups for (A) CH-type spin systems C8 and C3, (B) CH<sub>2</sub>-type spin systems C7 and C10, and (C) CH<sub>3</sub>-type spin system C11. Numerical simulations were conducted for systems of 2–4 spins and are included as solid lines, and are colored gray for an effective scaling factor  $S_{(eff)} < 1$ . For comparison to unscaled methyl characteristics, (C) additionally shows the buildup for a Val C $\delta$  methyl group.

#### 6.4.4 Agreement of Model Structure Variations with Molecular Dynamics

In agreement with ref. [114], the LG-CP order parameters observed for C10 and C11 are not affected by an increase in temperature from 261 to 318 K. In addition to the CHHC data at variable temperature, this suggests that the 3D structure derived at lower temperatures is largely preserved at higher temperatures and can be described by the order parameters shown in Figure 6.13.



**Figure 6.13:** Molecular model of polycrystalline TEE derived from the measurement of CHHC distance parameters and one ( $^{13}\text{C},^{13}\text{C}$ ) distance restraint. Order parameters have been obtained from LG-CP experiments, see text, and a result from ref. [114] is given in italics.

As discussed in ref. [114], broadband ( $^{13}\text{C},^{13}\text{C}$ ) 1Q correlation spectroscopy already leads the way to models of the dynamic disorder affecting the ester tail of solid TEE. These NMR data are, for example, consistent with rapid nutation in a  $25\text{--}30^\circ$  cone around the C10–C11 bond axis [114]. As discussed above, molecular mobility also affects the LG-CP buildup characteristics. Assuming an axially symmetric motion, the corresponding order parameters can be related to root-mean-square (RMS) angular fluctuations [180] via the equation

$$S = \frac{3 \cos^2 \theta - 1}{2} \quad . \quad (6.1)$$

For the  $^{13}\text{C}\text{--}^1\text{H}$  spin pairs, one finds RMS angular fluctuations of  $33^\circ$  (C10) and  $28^\circ$  (C11). Compared to the structural ensemble calculated from CHHC data, these parameters are well within the range of structural accuracy of the ester tail. Molecular mobility should

also affect results of ( $^{13}\text{C},^{13}\text{C}$ ) distance measurements as discussed above for  $\text{R}^2\text{TR}$  mixing, or as reported for rotational resonance ( $\text{R}^2$ ) experiments conducted on a selectively ( $^{13}\text{C}_{10},^{13}\text{C}_{11}$ )-labeled TEE sample [237]. The proposed cone motion would lead to fluctuations of the ( $^{13}\text{C},^{13}\text{C}$ ) ester-ring distance of 0.5–0.75 Å, which is close to the error margins reported for  $\text{R}^2$  and  $\text{R}^2\text{TR}$  experiments [237, 266].

## 6.5 Conclusions

In many biomolecular systems, information about both 3D structure and residual dynamics is of paramount importance for understanding biochemical functions such as protein folding, ligand binding, and protein–protein interactions. As is demonstrated here for TEE, constraints defining the 3D structure of a molecule are readily available using CHHC correlation experiments that indirectly probe nearest-neighbor  $^1\text{H}$ – $^1\text{H}$  interactions at high spectral resolution. These experiments confirm, in agreement with the findings of single-crystal X-ray diffraction, that TEE in the solid crystalline state adopts a scorpion-like structure. Assuming conservative  $^1\text{H}$ – $^1\text{H}$  distance constraints, the accuracy of the derived structure can be further improved by 2D  $\text{R}^2\text{TR}$  correlation experiments. While the location of the C10 atom relative to the phenyl ring is rather well-defined, C11–phenyl ring  $^1\text{H}$ – $^1\text{H}$  contacts are weak, leading to disorder in the resulting solid-state NMR model of TEE. An analysis of  $^{13}\text{C}$ – $^{13}\text{C}$  and  $^{13}\text{C}$ – $^1\text{H}$  one-bond interactions suggests that structural disorder is induced by molecular dynamics of the ester tail that is largely constant in the considered temperature range. Angular fluctuations deduced from the molecular order parameters are in qualitative agreement with the accuracy of the 3D structure.

As demonstrated here for TEE, the detection of through-space  $^1\text{H}$ – $^1\text{H}$  and  $^{13}\text{C}$ – $^{13}\text{C}$  interactions and one-bond dipolar couplings can provide an efficient means to elucidate molecular structure and dynamic disorder at atomic resolution. In particular, the combined application of both NMR techniques permits a clear distinction of structural or dynamic disorder in the solid state. Such information is, for example, relevant in the context of studying membrane insertion and penetration in a variety of biomolecules or may help characterizing ligand–receptor interactions in a non-crystalline environment.

## Chapter 7

# Protein Solid-State NMR Resonance Assignments from $^{13}\text{C}$ – $^{13}\text{C}$ Correlation Spectroscopy

## 7.1 Summary

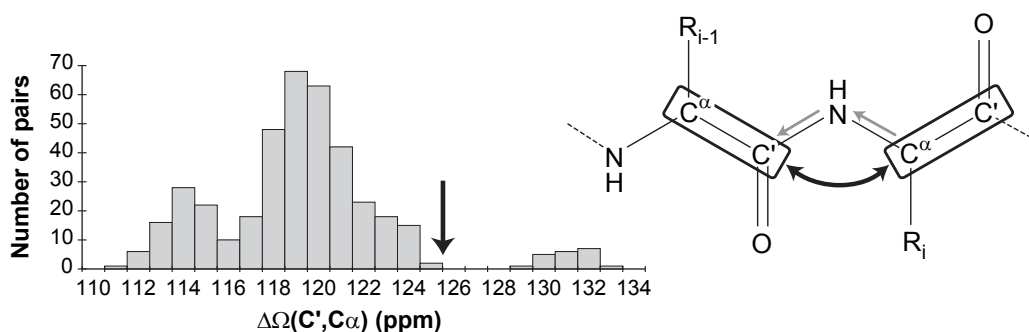
It is demonstrated that sequential resonance assignments can be obtained from ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectroscopy on a uniformly labeled protein under magic angle spinning. The experiment relies on weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) coupling conditions using a defined range of MAS rates, and can be employed at arbitrary magnetic field strength.

## 7.2 Introduction

For many studies of protein structure, near-complete sequential resonance assignments are a prerequisite and have now been obtained for several (membrane) proteins [44, 67, 75, 87, 126, 127, 222]. MAS-based spectral assignment methods have thus far relied on a collection of  $^{15}\text{N}$ – $^{13}\text{C}$  and  $^{13}\text{C}$ – $^{13}\text{C}$  one-bond dipolar or scalar polarization transfer schemes (see, for example, ref. [184]) to determine inter-residue connectivities in polypeptides. While side chain information may suffice to identify  $^{13}\text{C}$  resonances in small peptides [185], ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ) isotope labeling and a triple channel ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) NMR setup is mandatory in larger systems. In particular, several polarization transfer steps are necessary if sequential ( $\text{C}^\alpha$ ,  $\text{C}^\alpha$ ) [87] or side-chain–side-chain correlations are to be investigated in a single 2D experiment.

In principle, a reduction in the number of transfer units and hence an increase in efficiency of a sequential assignment protocol could be achieved if inter-residue side chain–side chain assignments are established using polarization transfer among local ( $\text{C}'$ ,  $\text{C}^\alpha$ ) moieties only.

For this purpose, polarization transfer must be mediated by a weak (inter-residue) coupling in the presence of two strong intra-residue  $\text{C}'$ - $\text{C}^\alpha$  interactions (Figure 7.1). At



**Figure 7.1:** Statistical analysis of ( $\text{C}'$ ,  $\text{C}^\alpha$ ) chemical shift differences using average chemical shifts compiled by the BioMagResBank data bank. Polarization transfer pathways to establish sequential ( $\text{C}'$ ,  $\text{C}^\alpha$ ) resonance assignments can invoke ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) polarization transfer (gray arrows) or can be mediated by ( $\text{C}'$ ,  $\text{C}^\alpha$ ) pairs (black arrow). Numerical simulations shown in Figure 7.2 were performed for the four-spin system highlighted by frames. The pointer at 125 ppm indicates the chemical shift difference (divided by two) equivalent to the MAS rates used for 2D correlation spectra shown in this chapter.

the same time, the mixing scheme of choice should be applicable at arbitrary magnetic field strength  $B_0$ . While the latter aspect suggests a transfer unit that is insensitive to isotropic or anisotropic chemical shielding interactions, broadband polarization transfer schemes have been shown to be unfavorable if weak couplings are to be studied in the presence of larger dipolar interactions [22, 115, 117, 146, 147]. Indeed, a series of quantum-mechanical (QM) model calculations assuming the four-spin topology given in Figure 7.1 reveals that broadband zero or double-quantum transfer schemes fail to generate sizable inter-residue polarization exchange.

Instead, the phenomenon originally termed *nuclear cross relaxation induced by specimen rotation* by Andrew and co-workers [6, 7] will be considered in the following. Here, the rate of polarization exchange in a dipolar coupled spin pair is enhanced if a multiple  $n$  of the MAS rate  $\omega_R$  (divided by the Larmor frequency  $\omega_0$ ) is equal to the chemical shift difference  $\Delta\Omega$ , i.e.  $\Delta\Omega = n\omega_R/\omega_0$ . Under proton decoupling conditions, this *rotational resonance* ( $\text{R}^2$ ) mechanism has been successfully used to determine structural constraints in a variety of pair-wise (see, for example, refs. [96, 164, 236]) and, more recently, uniformly labeled amino acids and polypeptides [26, 155, 239, 266, 318]. Unlike broadband mixing, the transfer dynamics are in this case explicitly dependent on isotropic resonance

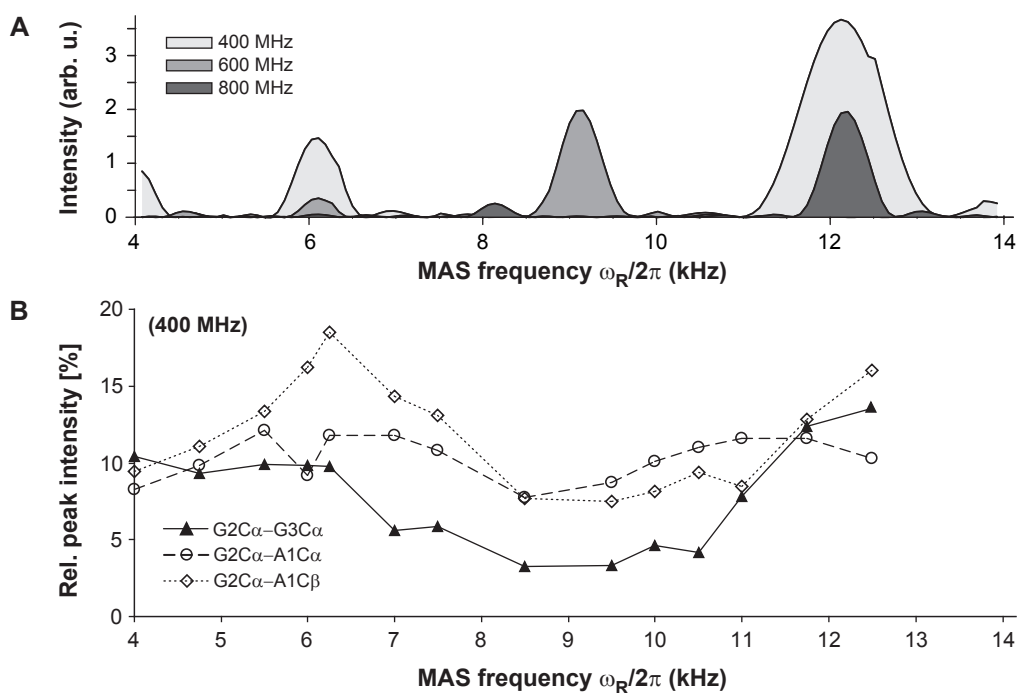


offsets and the magnitude and orientation of anisotropic dipolar and chemical shielding (CSA) interactions. While analytical solutions exist for the case of a dipolar-coupled two-spin system under exact  $R^2$  conditions [141, 173], a theoretical treatment of dipolar polarization transfer under off- $R^2$  conditions in a four-spin system is more complicated.

### 7.3 Numerical Analysis of ( $C'$ , $C^\alpha$ ) Polarization Transfer

Numerical calculations of a quantum mechanical (QM) four-spin ensemble are hence used to probe the signal intensity of the inter-residue transfer  $C'(i-1) \rightarrow C^\alpha(i)$  in the presence of two strongly coupled intra-residue ( $C'$ , $C^\alpha$ ) spin pairs. Simulations using the software package GAMMA [262] were conducted for a range of isotropic chemical shift differences  $\Delta\Omega(C',C^\alpha)$  (Figure 7.1) assuming standard dipolar and CSA [274, 316] interactions under variable molecular orientations.

As an example, the calculated maximum  $C^\alpha(i)$  signal intensity is plotted in Figure 7.2A as a function of the MAS rate for static magnetic fields  $\mathbf{B}_0$  corresponding to 400, 600 and 800 MHz  $^1\text{H}$  resonance frequency (*Simulation data provided by C. E. Hughes and M. Baldus.*). Assuming an initial density operator that reflects longitudinal magnetization on the  $C'(i-1)$  spin only, polarization transfer to  $C^\alpha(i)$  was monitored in the initial rate regime (0.8 ms mixing time) setting  $\Delta\Omega(i-1, i) \equiv \Delta\Omega(C'(i-1), C^\alpha(i)) = 122$  ppm. Intra-residue pairs with  $\Delta\Omega(i-1, i-1) = 111$  ppm and  $\Delta\Omega(i, i) = 112$  ppm were assumed. As visible in Figure 7.2A, inter-residue polarization transfer (including the complete four-spin system) is possible around the  $n = 2$  (400–800 MHz) and  $n = 1$  (400 MHz)  $R^2$  conditions for the considered MAS rates. Similar results are obtained if the magnitude or relative orientation of dipolar and CSA tensors is varied. Even in the degenerate case, i.e.  $\Delta\Omega(i-1, i) = \Delta\Omega(i, i) = \Delta\Omega(i-1, i-1)$ , sizable polarization transfer is possible. Comparative calculations assuming a ( $C'(i-1), C^\alpha(i), C'(i)$ ) three-spin topology show that the latter result is only found if two sequential ( $C', C^\alpha$ ) spin pairs are considered. According to the QM four-spin study, the polarization transfer profile shown in Figure 7.4A within the inter-residue ( $C', C^\alpha$ ) pair hence largely follows the general transfer characteristics of an isolated two-spin system. Normalized to the initial polarization, efficiencies of up to 20% for inter-residue polarization transfer are predicted from the simulation.

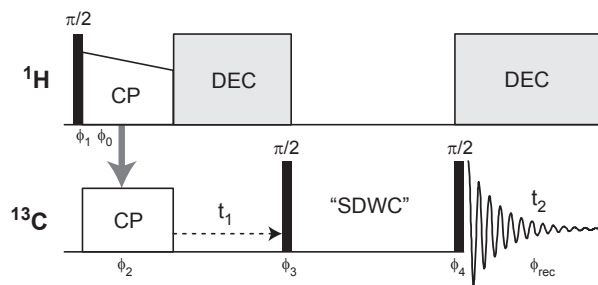


**Figure 7.2:** (A)  $\text{C}'(i-1) \rightarrow \text{C}^\alpha(i)$  polarization transfer as a function of MAS rate and magnetic field  $B_0$ . Isotropic chemical shift differences are described in the text. While anisotropic CSA interactions of  $\text{C}'$  and  $\text{C}^\alpha$  and the orientation of the  $\text{C}'$  CSA tensor were taken from the literature [274, 316], simulations were performed for different molecular orientations of the  $\text{C}^\alpha$  CSA tensor. (B) Inter-residue cross peaks, integrated and normalized against the  $\text{G2C}^\alpha$  diagonal peak intensity, obtained on  $\text{U-}(^{13}\text{C}, ^{15}\text{N})$ -labeled AGG for variable MAS rate. Data were recorded on a 400 MHz NMR instrument.

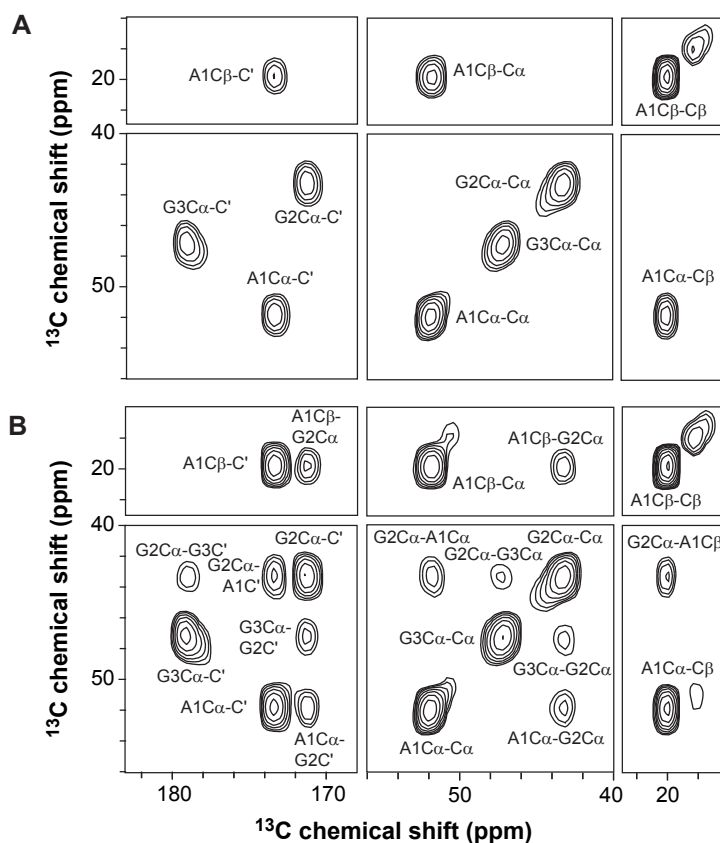
## 7.4 Spin Diffusion Under Weak Coupling Conditions

Setting the MAS rate at exact  $R^2$  conditions would result in severe line broadening effects and hence must be avoided. Furthermore, sequential resonance assignments should be possible at variable  $B_0$ . A band-selective ( $C'(i-1), C^\alpha(i)$ ) polarization transfer profile can be created by changing the MAS rate during mixing [305] or by incorporation of an additional radio frequency field on the  $^{13}\text{C}$  channel [50, 59, 67, 266, 281, 306]. While these methods permit the use of MAS rates away from  $R^2$  conditions, they also strongly modulate or remove the chemical shift dependence of the polarization scheme. Preserving the general chemical shift selectivity and, at the same time, broadening the transfer profile can be achieved by removing proton decoupling during mixing. The influence of the dipolar coupled proton reservoir has been described using first order perturbation theory [6, 154] and is, in the context of our experiment, commonly referred to as proton-driven spin diffusion (PDS) [35, 275, 276]. PDS mixing is possible for a variety of MAS rates and is characterized by an exponential rather than (damped) oscillatory (as in the two- or four-spin case) polarization transfer. On the other hand, the rate of polarization exchange depends on the zero-quantum line-shape function exhibiting a maximum at  $\Delta\Omega = n\omega_R/\omega_0$  in line with the theoretical four-spin simulations shown in Figure 7.2.

Spin Diffusion under 'Weak Coupling' conditions (SDWC) can hence be created if an MAS rate *close to but not exactly at the  $R^2$   $n = 2$  transfer condition* (i.e. around 125 ppm/2, Figure 7.1) is selected and PDS is active. As a result, polarization transfer is significantly slower than under exact  $R^2$  conditions Figure 7.2A, but the desired mixing characteristics are preserved. In addition to the  $C'(i-1) \rightarrow C^\alpha(i)$  transfer, proton-driven spin diffusion also allows for subsequent polarization transfer from  $C^\alpha(i)$  to side chain resonances of residue  $i$ . A standard two-dimensional PDS experiment (see Figure 7.3 and Chapter 3) hence should not only contain intra-residue correlations, but also inter-residue cross peaks such as ( $C'(i-1), C^\alpha(i)$ ), ( $C^\alpha(i-1), C^\alpha(i)$ ) or ( $C^\alpha(i-1), C^\beta(i)$ ). This effect is demonstrated in Figure 7.4 on the uniformly ( $^{13}\text{C}, ^{15}\text{N}$ ) labeled tri-peptide Ala-Gly-Gly for an MAS rate of 7 kHz on a 400 MHz instrument. For a mixing time of 10 ms (Figure 7.4A), intra-residue ( $^{13}\text{C}, ^{13}\text{C}$ ) correlations are detected, in agreement with the spectral assignments reported previously [183], and in line with the general PDS transfer profile. Increasing the mixing time to 50 ms (Figure 7.4B) leads, in addition, to the occurrence of sizable sequential ( $C', C^\alpha$ ), ( $C^\alpha, C^\alpha$ ) and ( $C^\alpha, C^\beta$ ) cross peaks. Inter-residue polarization transfer is enhanced if MAS rates close to the  $R^2$  conditions are selected, Figure 7.2B, in



**Figure 7.3:** Standard two-dimensional ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiment with longitudinal mixing, where for example PDSD or, under suitable conditions described in the text, Spin Diffusion under 'Weak Coupling' conditions can be active. Phases identical to the standard PDSD experiment, e.g.  $\phi_1 = y$ ,  $\phi_0 = \phi_2 = x$ ,  $\phi_3 = y$ ,  $\phi_4 = -y$ . Receiver phase and phase cycle given in Appendix Chapter I.



**Figure 7.4:** 2D ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiments on U-( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-labeled AGG at 7 kHz MAS for a mixing time of (A) 10 ms, and (B) 50 ms. During evolution and detection periods, TPPM decoupling [28] was used at 90 kHz.

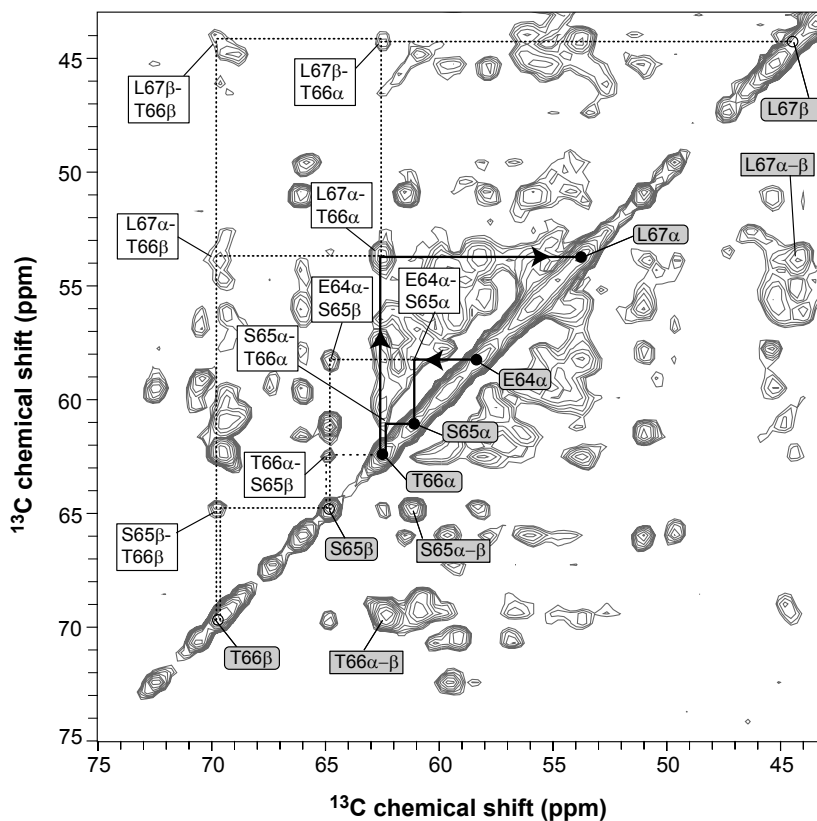
accordance with the QM model calculations. Line-broadening effects can be detected that are strongest at  $n = 1$  conditions, but range between 0–20% compared to the  $^{13}\text{C}$  line width far off  $R^2$  conditions if MAS rates close the  $n = 2$  condition are selected. Further experimental evidence that inter-residue polarization transfer indeed results from indirect ( $\text{C}^\alpha \leftrightarrow \text{C}' \leftrightarrow \text{C}^\alpha$ ) interactions comes from a 2D control experiment, in which  $\text{C}'$  resonances are decoupled by a weak on-resonance RF field during mixing (data not shown). In this case, both  $\text{C}'$  and  $\text{C}^\alpha$  peaks are strongly attenuated, indicative of a coupled spin state. Additional experiments using variable mixing times reveal that the maximum intensity of the sequential ( $\text{C}', \text{C}^\alpha$ ) and ( $\text{C}^\alpha, \text{C}^\alpha$ ) cross peaks is detected for mixing times in the order of 50 ms and 100 ms, respectively. For reasons of spectral resolution and dispersion, the latter interactions are, along with additional side-chain correlations, most informative for sequential ( $^{13}\text{C}, ^{13}\text{C}$ ) resonance assignments.

## 7.5 Application to Uniformly Labeled Proteins

To investigate the use of correlation spectroscopy with Spin Diffusion under ( $\text{C}', \text{C}^\alpha$ ) Weak Coupling conditions in larger systems and at higher magnetic fields, experiments on two uniformly labeled proteins were conducted.

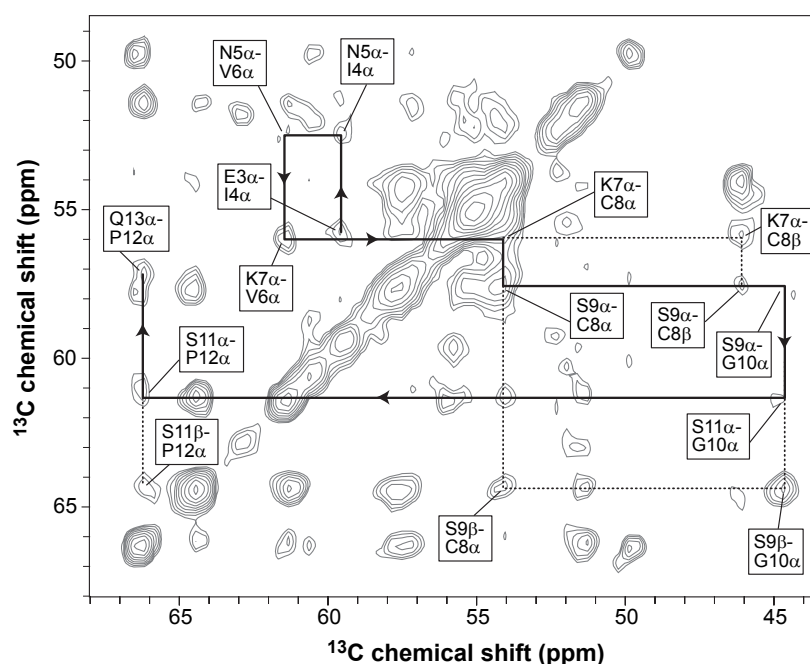
Results for the 76-residue protein ubiquitin, obtained on a 600 MHz wide bore instrument using an MAS rate of 9.375 kHz (i.e. 125 ppm/2, see Figure 7.1) and a mixing time of 100 ms are shown in Figure 7.5. Line broadening effects can be neglected, and a variety of inter-residue ( $\text{C}^\alpha, \text{C}^\alpha$ ) and ( $\text{C}^\alpha, \text{C}^\beta$ ) cross peaks are readily observed. These dependencies are highlighted in Figure 7.5 for the four amino-acid segment E64–L67. Inter- and intra-residue correlations are indicated on the left and right side of the diagonal, respectively. The backbone  $\text{C}^\alpha$ – $\text{C}^\alpha$  'walk' is facilitated by the occurrence of sequential ( $\text{C}^\alpha, \text{C}^\beta$ ) and ( $\text{C}^\beta, \text{C}^\beta$ ) correlations that can be used to cross validate the assignment procedure. Spectral assignments have been confirmed by additional NC correlation experiments as described in Chapter 8. Also, Figure 8.2 shows a spectrum for all aliphatic and carbonyl resonances obtained on the same sample using SDWC at a field corresponding to 800 MHz  $^1\text{H}$  resonance frequency.

The robustness of this approach is furthermore supported by results on a uniformly ( $^{13}\text{C}, ^{15}\text{N}$ ) labeled sample of the 38-residue protein Kalitoxin, shown in Figure 7.6 (*data provided by A. Lange*). The spectrum was recorded on an 800 MHz narrow-bore instrument. Again, the MAS rate was placed close to the  $n = 2$  condition (12.5 kHz) and a



**Figure 7.5:** ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiment conducted under weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) coupling conditions described in the text on U-( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-labeled ubiquitin. Sequential ( $\text{C}^\alpha$ ,  $\text{C}^\alpha$ ), ( $\text{C}^\alpha$ ,  $\text{C}^\beta$ ) or ( $\text{C}^\beta$ ,  $\text{C}^\beta$ ) correlations are indicated on the left of the diagonal. Intra-residue assignments are included in gray on the right side of the diagonal. During evolution and detection periods, TPPM decoupling [28] was used at 90 kHz. This sample of 8 mg was nano-crystallized by precipitation from polyethylene glycol [44, 190]

mixing time of 150 ms was employed. A variety of sequential correlations can be readily



**Figure 7.6:** 2D ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiment conducted at 800 MHz on U-( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-labeled Kaliotoxin, under weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) coupling conditions. A longitudinal mixing time of 150 ms was used. Temperature and MAS rate were set to  $-15^\circ\text{C}$  and 12.5 kHz, respectively. During evolution and detection periods, SPINAL64 [88] decoupling was used at 83 kHz. This sample of 6 mg was rehydrated before measurement.

identified. Cross validation using a series of NCACB and NCOCA experiments [161] reveals that these correlations again correspond to  $(i, i - 1)$  connectivities. As an example, the 'sequential walk' through ( $\text{C}^\alpha$ ,  $\text{C}^\alpha$ ) correlations of the eleven residue stretch E3–Q13 is highlighted in Figure 7.6.

In addition, solid-state NMR resonance assignments from ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectroscopy have meanwhile been obtained for a number of proteins. This includes membrane proteins reconstituted in lipid bilayers, see for example phospholamban in complex with  $\text{Ca}^{2+}$ -ATPase (Chapter 9), signaling rhodopsin II [81], and other native-like, non-crystalline environments. A ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectrum from employing SDWC on a sample of  $\alpha$ -synuclein fibrils is shown in Appendix Figure E.1 (*data provided by H. Heise*). Likewise, assignments have been obtained on Tau protein fibrils [8].

## 7.6 Conclusion

A set of experimental conditions that permit establishing inter-residue correlations under MAS conditions from a single ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiment has been described. For a given  $B_0$  field, the technique relies on establishing weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) coupling conditions using a defined range of MAS rates. Once sequential ( $\text{C}'$ ,  $\text{C}^\alpha$ ) transfer is achieved, additional correlations (mediated by proton-driven spin diffusion) between side chain moieties are possible. As confirmed by further test experiments, simply increasing the selected mixing time leads to the occurrence of additional interresidue correlations, including ( $i$ ,  $i + 2$ ) peaks. Such extensions can greatly facilitate the spectral assignment process and are only limited by  $T_1$  relaxation. Since no  $^{15}\text{N}$  evolution and mixing periods are required, the overall signal-to-noise ratio can be improved significantly. Experiments can be performed on uniformly labeled protein samples at various magnetic fields. In the presented applications, the MAS rate was selected based on statistical considerations. In the presence of ( $\text{C}'$ ,  $\text{C}^\alpha$ ) chemical shift variations, for example, due to the influence of secondary structure or hydrogen bonding, adjustments of MAS rate and mixing time may be necessary. Such conditions were successfully applied to the proteins considered here and reveal that weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) coupling conditions can be established for a variety of experimental setups. Compared to sequential resonance assignments obtained from chemical-shift selective ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) correlation spectroscopy [21], 2D ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) data recorded under weak coupling conditions contain both intra- and inter-residue correlations. However, both contributions can be easily separated by recording an additional 2D spectrum for short mixing times (Figure 7.4) or by using MAS rates  $\omega_R/\omega_0 \gg \Delta\Omega$ . The transfer scheme can be easily incorporated into other, multi-dimensional correlation schemes without loss in spectral resolution (for example, the 3D 2Q/1Q/1Q CCC correlation sequence shown in Chapter 3). In particular, these experiments may provide a useful reference for the analysis of proton-mediated ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiments [159, 160]. In addition to the applications discussed here, weak coupling conditions may also be employed to monitor correlations between  $\text{C}'$  spins and methyl or aromatic side-chain carbons.



## Chapter 8

# High-Resolution Solid-State NMR Studies on Uniformly ( $^{13}\text{C}$ , $^{15}\text{N}$ )-Labeled Ubiquitin

## 8.1 Summary

Understanding of the effects of intermolecular interactions, molecular dynamics, and sample preparation on high-resolution magic-angle spinning NMR data is currently limited. Using the example of a uniformly ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-labeled sample of ubiquitin, solid-state NMR methods tailored to the construction of 3D molecular structure are discussed, and the influence of solid-phase protein preparation on solid-state NMR spectra are studied. A comparative analysis of  $\text{C}'$ ,  $\text{C}^\alpha$ , and  $\text{C}^\beta$  resonance frequencies suggests that  $^{13}\text{C}$  chemical shift variations are most likely to occur in protein regions that exhibit an enhanced degree of molecular mobility. Results presented in this chapter can be refined by additional solid-state NMR techniques, and serve as a reference for ongoing efforts to characterize the structure and dynamics of (membrane) proteins, protein complexes, and other biomolecules by high-resolution solid-state NMR spectroscopy, such as presented in Chapter 9 and refs. [9, 81, 161].

## 8.2 Introduction

Significant progress has recently been made in the use of magic-angle spinning (MAS) [4] NMR for structural studies in large, noncrystalline complexes. Structural models have been obtained, for example, for polypeptides that form macromolecular assemblies [134, 140, 228] or bind to large membrane proteins [185]. If 3D molecular structures are available as a reference, MAS-based solid-state techniques can be used to gain insight into the details of protein functioning, as demonstrated by several groups in the case of rhodopsin [60, 97,

220]. In addition, polypeptides and globular proteins have been an important target area of solid-state NMR spectroscopy, including studies on protein dynamics [55, 90, 286], folding [44, 80, 110, 112], or ligand binding [327]. Whenever high-resolution X-ray or solution-state NMR 3D structures are available, globular proteins also represent a useful tool with which to develop and optimize solid-state NMR methodology and instrumentation. Such efforts have given rise to a handful of 3D structures ranging from two- and three-residue peptides to small proteins [46, 134, 161, 212, 240]. In particular, we have shown [160, 161, 255] (see also Chapter 6) that 3D molecular structures can be constructed through the use of a single, uniformly labeled sample.

Researchers have for a long time been interested in establishing a relationship between (supra)molecular structure and variations in solid-state NMR resonance frequencies, in contexts such as the polymorphism observed for compounds of low molecular weight [107]. In addition to variations in experimental conditions, such as hydration level or temperature [55], details of sample preparation have been shown to affect MAS NMR spectra of microcrystalline proteins [44, 190], protein fibrils formed by  $\beta$ -amyloid peptides [229], and  $\alpha$ -synuclein [111], and they can also play a significant role in the case of peptides and proteins reconstituted into lipid bilayers. While it is well known that hydration [94, 118, 182], crystallization [55, 126, 196, 221, 263], or the addition of cryoprotectants [131] can improve the spectral resolution of globular proteins, a detailed understanding of the relationship between solid-phase protein preparation and 3D molecular structure derived from MAS NMR is currently lacking. Such aspects have been thoroughly investigated by solution-state NMR in the case of ubiquitin, a 76-residue protein rich in secondary structure and involved in a variety of important cellular functions [192]. As well as the 3D structure [57] (PDB code: 1D3Z), folding intermediates [43] and, in particular, protein dynamics (see, for example, refs. [43, 227, 283]) have been examined in great detail.

Not surprisingly, ubiquitin has also become an attractive model system for MAS-based solid-state NMR studies, and yields MAS-NMR spectra that vary depending on the details of sample preparation [118, 126, 127, 159, 182, 190, 224, 272]. Here, first, sequential resonance assignments for uniformly labeled U-( $^{13}\text{C}$ ,  $^{15}\text{N}$ ) ubiquitin precipitated from polyethylene glycol (PEG) are derived. As demonstrated by M. Etzkorn in the publication related to this chapter [254], ( $^{15}\text{N}$ ,  $^{15}\text{N}$ ) correlations and indirectly detected proton-proton contacts can be used to monitor protein secondary structure (data not shown here). Insight gained from these experiments is used to refine the chemical shift-based secondary structure analysis. Results imply that the single-crystal structure of

ubiquitin and the 3D structure in PEG-microcrystals must be closely related, and so one can compare the experimental findings with results from MAS-NMR obtained on hydrated ubiquitin [252] and precipitants obtained by use of 2-methylpentane-2,4-diol (MPD) [126] in close reference to the X-ray structure and solution-state NMR data. It can be concluded that the influence of sample preparation on 3D protein structure determination in solids is most critical for molecular segments that exhibit a high degree of molecular mobility.

## 8.3 Material and Methods

### 8.3.1 Sample Preparation

Uniformly ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-labeled ubiquitin was purchased from VLI research (Malvern, PA) or expressed recombinantly in *E. coli* and purified by established procedures [167, 324]. After freeze-drying, two alternative routes of sample preparation were followed. As described in refs. [44, 190], ubiquitin (8 mg) was precipitated from polyethylene glycol (PEG) and transferred into a 4 mm MAS rotor. Additional experiments were conducted on U- ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-ubiquitin (6 mg) that was hydrated with  $\text{H}_2\text{O}$  (10 ml) after lyophilization in the rotor, see also [252]. For brevity, ubiquitin sample preparations relevant in the context of this study are described by UBI-P (PEG precipitation), UBI-H (rehydrated), and UBI-M (MPD precipitated). While experimental results are shown for UBI-P and UBI-H, resonance assignments for UBI-M were taken from refs. [126, 127].

### 8.3.2 Solid-State NMR Experiments

All NMR experiments were conducted with use of 4 mm triple-resonance ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) probe heads at a static magnetic field of 18.8 T and 14.1 T corresponding to 800 MHz and 600 MHz proton resonance frequencies (Bruker Biospin, Karlsruhe, Germany), respectively. Hartmann–Hahn [109] ( $^1\text{H}$ ,  $^{13}\text{C}$ ) and ( $^1\text{H}$ ,  $^{15}\text{N}$ ) cross polarization was established by use of ramped [201] radio frequency (RF) fields. For proton decoupling, TPPM [28] and SPINAL-64 [88] multiple-pulse schemes were used. Sequential resonance assignments were obtained from combining results of two- and three-dimensional ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) correlation experiments as described in refs. [21, 25] and Chapter 3, with results of ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiments performed under weak coupling conditions, see Chapters 3 and 6 [253]. Identifications of residue-specific spin systems were further assisted by a 3D ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ,  $^{13}\text{C}$ ) double-quantum/single-quantum/single-quantum (2Q/1Q/1Q) correlation spectrum (3D CCC) [113], see also Chapter 3. MAS rates between 9 and 12.5 kHz were employed at

sample temperatures between  $-14\text{ }^{\circ}\text{C}$  and  $-5\text{ }^{\circ}\text{C}$ . CHHC correlation experiments were conducted as described in refs. [159, 160], see also Chapters 3 and 6, with use of ( $^1\text{H}, ^1\text{H}$ ) mixing times between  $250\ \mu\text{s}$  and  $400\ \mu\text{s}$ , bracketed by short ( $^1\text{H}, ^{13}\text{C}$ ) CP transfers for contact times of 80–150 ms.

### 8.3.3 Data Analysis

Spectra were processed with QSINE window functions in all dimensions and analyzed with the aid of Sparky [92]. Resonance assignments were evaluated by using TALOS version 2003 [58], which predicts the backbone dihedral angles  $\phi$  and  $\psi$  on the basis of chemical shift and sequence homology with proteins of known structure and assigned (solution state) chemical shifts. Prior to use, the TALOS data base was re-compiled under exclusion of the solution-state NMR / X-ray diffraction ubiquitin data set. Results provided by M. Etzkorn from NN and NHHC experiments [79, 254] were used to resolve ambiguities for backbone dihedral angle predictions from TALOS using chemical-shift assignments for N, C',  $\text{C}^{\alpha}$ , and  $\text{C}^{\beta}$ , when pairs were not classified as reliable (i.e., 'good') in by the prediction algorithm. For cross validation, the crystal structure [308] (PDB ID: 1UBQ) was used.

### 8.3.4 Structure Calculation

Structure calculations were performed by use of a simulated annealing protocol in CNS [42] version 1.1 with the PROTEIN-ALLHDG [209] parameter file. From the combined TALOS and NN/NHHC analysis (*vide supra*), 45  $\phi$  angles and 47  $\psi$  angles were obtained for the simulation. Restraints were enforced by square-well potentials with no energy contribution for deviations within the predicted RMSD. Likewise, constraints for proton–proton distances were invoked with an allowed upper limit of  $3.5\ \text{\AA}$  and no lower bounds. Simulations started from an extended conformation generated from the amino acid sequence. The structure calculation protocol consisted of three stages: (1) high-temperature annealing in torsion-angle space, in 2000 time steps of 15 fs at 50 000 K, (2) slow-cool annealing in torsion-angle space, in 4000 steps of 15 fs, and temperature reduction from 50000 K to zero in steps of 250 K, and (3) final conjugate gradient minimization in 20 cycles of 100 steps each. Force constants were set to 300, 300, and  $150\ \text{kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$  for the distance restraints, and 100, 200, and  $400\ \text{kcal}\cdot\text{mol}^{-1}\cdot\text{rad}^{-2}$  for the backbone angle restraints during the three stages. Ambiguities in the assignments of methylene and methyl protons were accounted for by  $R^{-6}$  averaging over all possible contacts. A set of 100 structures was calculated, starting with different initial velocities. An ensemble of ten structures was

selected according to the lowest overall energy, and was aligned along the backbone atoms of residues M1 to V70 with MOLMOL 2K.2 [152].

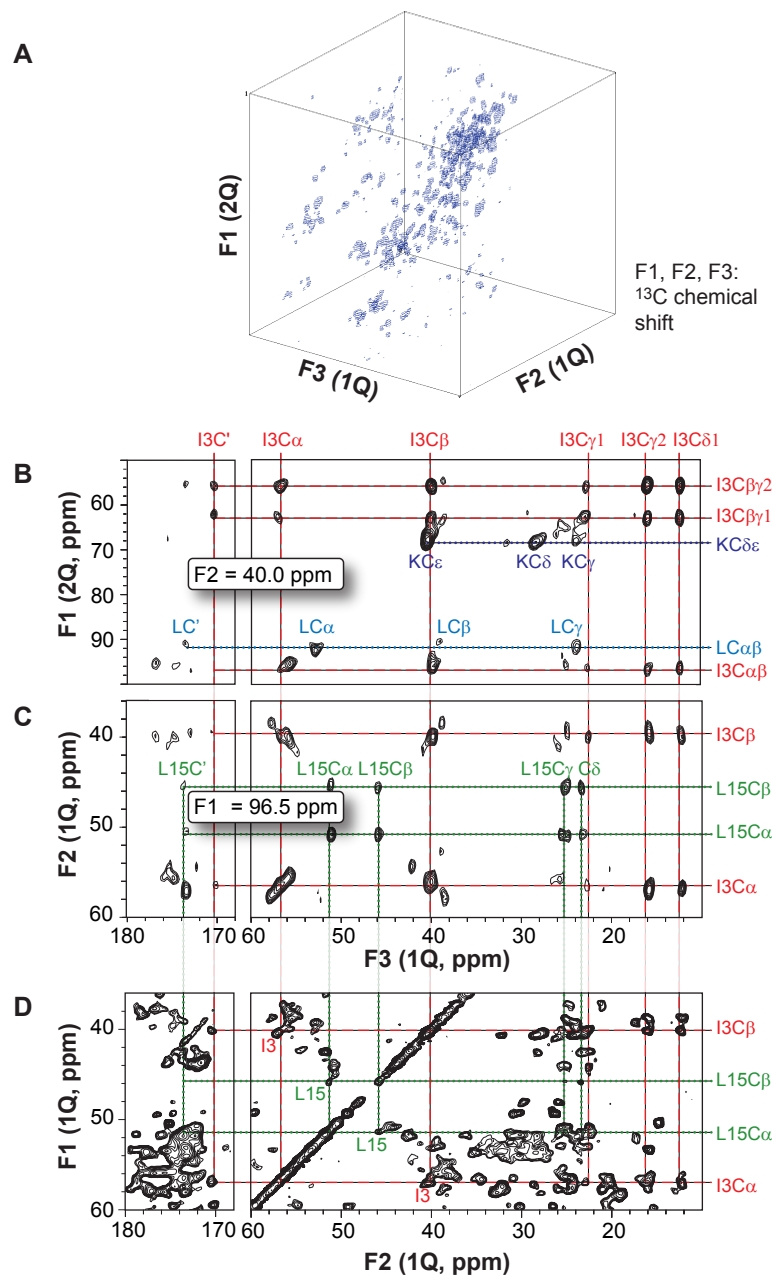
## 8.4 Structural Characterization of UBI-P

### 8.4.1 Resonance Assignments

A set of two- and three-dimensional ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) and ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiments was used to obtain sequential resonance assignments of UBI-P (precipitation with PEG),

#### Spin System Identification

The identification of residue-specific spin systems was facilitated using 2D ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectra with SD mixing times ranging from 5 to 50 ms or double-quantum mixing, 2D and 3D NCACX spectra with transfers established by NCA SPECIFIC CP [25] and PDSO ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) mixing, and the novel 3D 2Q/1Q/1Q CCC pulse sequence [113]. The 3D CCC proved to be particularly helpful in identifying a large number of C' resonances. As an example, Figure 8.1A displays the CCC spectrum of U( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-UBI-P. Excitation and reconversion of double-quantum coherences between directly bonded side chain carbon spins was achieved by  $R18_{10}^5$  [153] with a mixing time of 910  $\mu\text{s}$ . The second homonuclear mixing block consists of a spin-diffusion delay with a mixing time of 50 ms, which is long enough to enable intra-residue magnetization transfer along the full amino acid side chain. The spectrum was acquired with 108  $t_1$  and 90  $t_2$  increments, corresponding to  $t_{1,\text{max}} = 2.56$  ms and  $t_{2,\text{max}} = 3.73$  ms. With 16 scans per slice, the total experiment time was 90 h. In Figure 8.1B, the F1/F3 plane of the CCC spectrum at the resonance frequency of 40 ppm in F2 is displayed. This frequency corresponds to the  $C^\beta$  chemical shift of the amino acid I3. This slice exhibits cross-peaks between all three double-quantum coherences involving  $C^\beta$ , i.e.,  $C^\beta C^\alpha$ ,  $C^\beta C^{\gamma 1}$ , and  $C^\beta C^{\gamma 2}$ , in F1 and all resonances of the Ile residue, i.e., C',  $C^\alpha$ ,  $C^\beta$ ,  $C^{\gamma 1}$ ,  $C^{\gamma 2}$ , and  $C^{\delta 1}$  along F3, as indicated by the dashed lines. Furthermore, the spin system of one Lysine side chain at the  $C^\delta C^\epsilon$  double-quantum resonance and of one Leucine at the  $C^\alpha C^\beta$  double-quantum coherence can be seen (dotted line). In Figure 8.1C, the F2/F3 plane at 96.5 ppm, corresponding to the  $C^\alpha C^\beta$  double-quantum coherence of I3, is shown. The diagonal and back-transfer (anti-diagonal) peaks for  $C^\alpha$  and  $C^\beta$  of I3 span a square in the F1/F2-plane. Two sets of true 3D signals for all other spins belonging to the I3 chain are clearly visible at the  $C^\alpha$  as well as the  $C^\beta$  resonance in F2. Apart from the signals of I3, the L15 resonances related to the  $C^\alpha C^\beta$  double-quantum

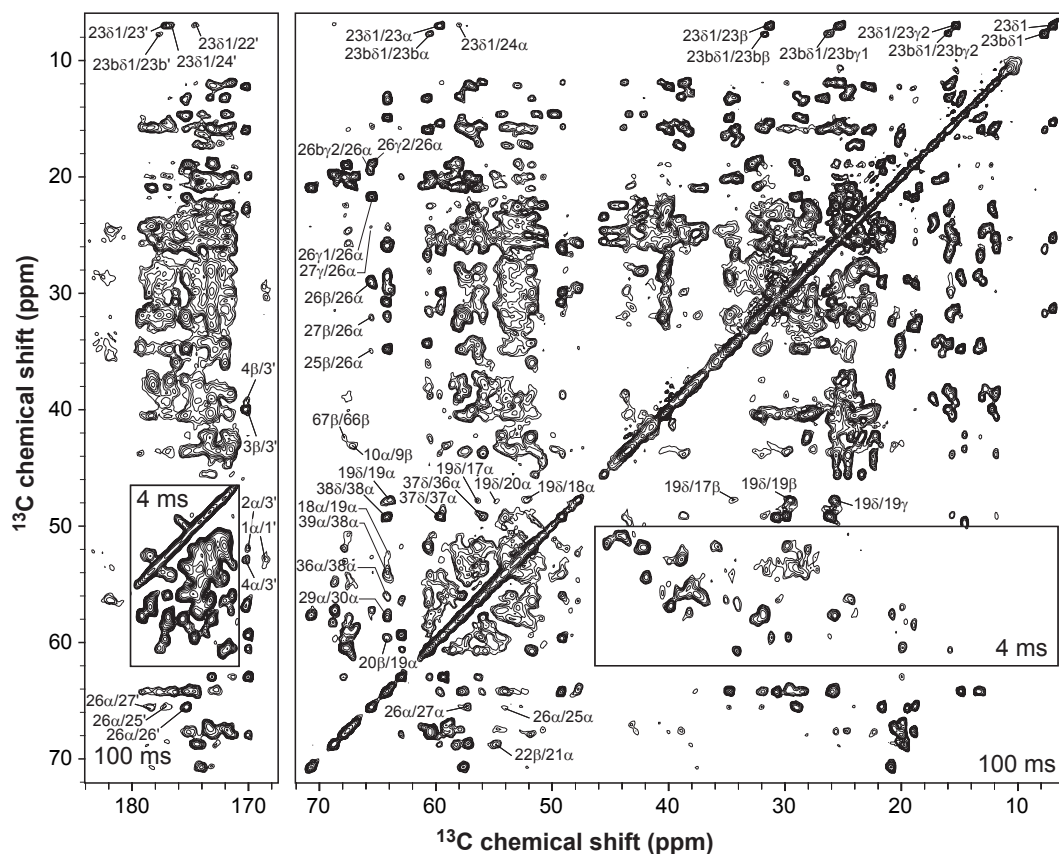


**Figure 8.1:** (A) 3D 2Q/1Q/1Q CCC spectrum of U( $^{13}\text{C},^{15}\text{N}$ )-UBI-P, obtained at 600 MHz proton frequency with 11 kHz MAS,  $R18_{10}^5$  2Q recoupling and 50 ms PDSD mixing. (B) F1/F3 plane at F2 = 40 ppm, corresponding to the  $\text{C}^\beta$  frequency of I3. (C) F2/F3 plane at F1 = 96.5 ppm, corresponding to the  $\text{C}^\alpha/\text{C}^\beta$  double-quantum frequency of I3. The I3 spin system is indicated with dashed lines, spin systems of other amino acids with dotted lines. (D) 2D ( $^{13}\text{C},^{13}\text{C}$ ) spectrum with 40 ms PDSD mixing, recorded with 11 kHz MAS at 600 MHz proton Larmor frequency.

coherence at 96.5 ppm (dotted line) are clearly visible. In Figure 8.1D, a 2D spin-diffusion spectrum obtained with a mixing time of 40 ms is displayed for comparison. In there, the resolution in the ( $C'$ ,  $C^\alpha$ ) cross-peak region is strongly limited due to spectral overlap, caused by rather poor spectral dispersion of  $C^\alpha$  as well as  $C'$  resonance frequencies. In contrast,  $C'$  signals of I3 and L15 in the 3D spectrum can clearly be identified. The same applies to the side chain resonances  $C^{\gamma 1,2}$  and  $C^{\delta 1}$ , which are now clearly separated from all other signals. It is straightforward to extend this CCC experiment for obtaining sequential ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) resonance assignments under 'weak coupling' conditions.

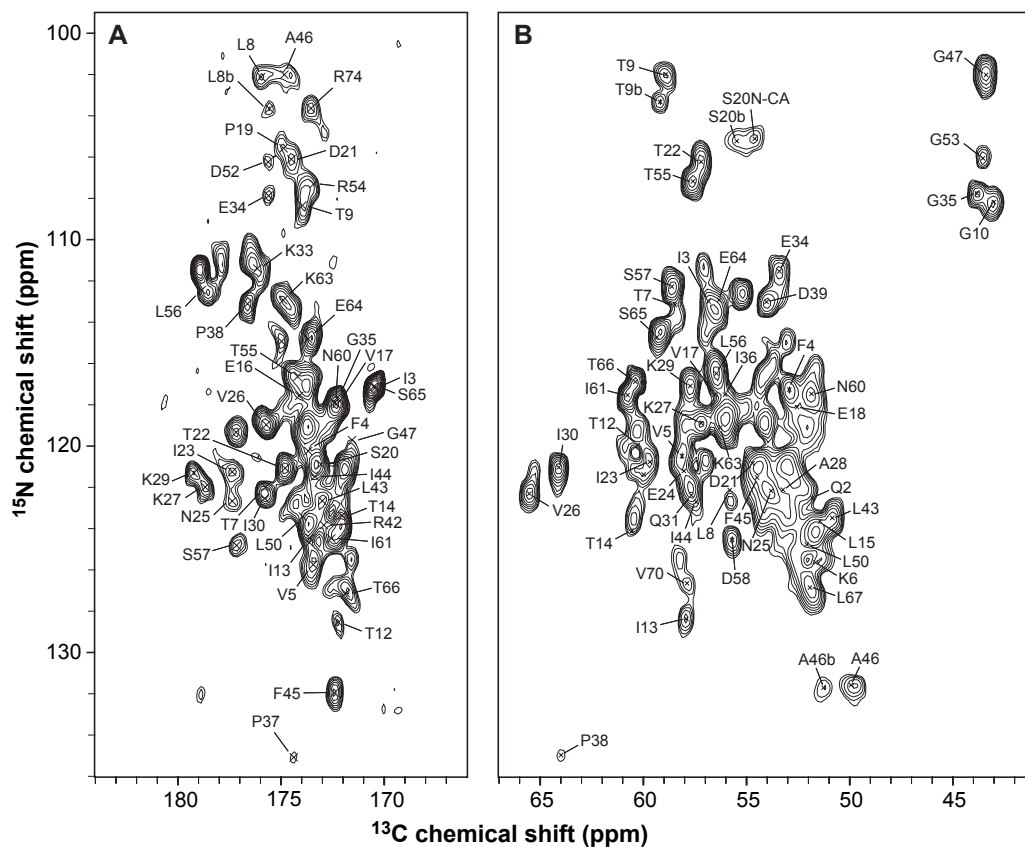
### Sequential Resonance Assignment

Sequence-specific assignments were obtained using NCO-type transfer within 2D and 3D NCOCX pulse sequences, and 2D ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlations from Spin Diffusion under Weak Coupling conditions (SDWC) [253], see also Chapter 7. As an example, Figure 8.2 shows the results of the two-dimensional ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) experiment on UBI-P for 100 ms SDWC mixing in comparison short mixing of only 4 ms. For short mixing times ( $t_{\text{CC}} = 4$  ms; insert), only intra-residue correlations are visible, whilst for a mixing time of 100 ms, inter-residue ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlations appear. These interactions are particularly easy to identify in the spectral region between 70 and 40 ppm, where intra-residue correlations are only possible for Ser, Thr, or Pro residues. In Figure 8.2, ( $i$ ,  $i \pm 1$ ) and ( $i$ ,  $i \pm 2$ ) correlations involving residues 1–4 and 17–27 are indicated as an example. To separate signal sets in an additional dimension, 2D and 3D NCOCX correlation experiments with PDSM mixing were conducted; alternatively, the 3D CCC sequence could also be employed here. Notably, no qualitative differences were detected in any of the spectra if the PDSM mixing unit was replaced by a DARR-type [282] irradiation scheme in which heteronuclear dipolar interactions involving protons and through-space proton–proton interactions are active. Figure 8.3 shows NCO and NCA-type spectra obtained on UBI-P from an 800 MHz instrument. Additional experiments using variable  $^1\text{H}$  decoupling strength reveal that the detected  $^{15}\text{N}$  line width (0.8 ppm) is sensitive to the  $^1\text{H}$  RF decoupling field strength in the 70–95 kHz range. The assignment of 3D NCOCX correlations profited substantially from a large number of  $C'$  chemical shift assignments obtained using the 3D CCC correlation sequence. Finally, analysis of the two- and three-dimensional ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) and ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) correlation spectra resulted in the *de novo* assignment of 86% of the 76 protein residues of ubiquitin. The assignments are given in Appendix Chapter F and reveal peak doubling for T9, I13, S20, I23, V26, P37, and A46. Two ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation



**Figure 8.2:** CC correlation spectrum obtained on U( $^{13}\text{C},^{15}\text{N}$ ) UBI-P for mixing times of 100 ms and 4 ms (inserts) under weak coupling conditions [253]. Spectra were recorded at 18.8 T  $B_0$  field, 12.5 kHz MAS rate, and a temperature of 261 K. Signals were acquired over 36 h, with a maximum  $t_1$  evolution time of 6.2 ms. Several  $(i, i \pm 1)$  and  $(i, i \pm 2)$  correlations are indicated as an example.





**Figure 8.3:** (A) NCO correlation spectrum under experimental conditions as in Figure 8.2 with SPINAL-64 proton decoupling at 93 kHz RF field strength. The duration of the experiment was 2.3 h, with a maximum  $t_1$  evolution time of 8 ms. Peak labels relate to the read-out residue, i.e. to residue  $i - 1$ , if the sequential transfer was initiated on residue  $i$ . (B) NCA correlation spectrum under the same conditions as in A.

patterns are seen for I23 and V26 in Figure 8.2, for example, and two NCA peaks are found for T9 and S20 in Figure 8.3. Such peak doubling has been observed previously in ( $^1\text{H}, ^{15}\text{N}$ ) correlation spectra obtained on perdeuterated UBI-P samples [224] and for the catabolite repression histidine-containing phosphocarrier protein Crh [44], and is indicative of the occurrence of multiple 3D conformations in the solid state. All sequential resonance assignments reported here are self-consistent and relate to the dominant peak intensities in the spectra.

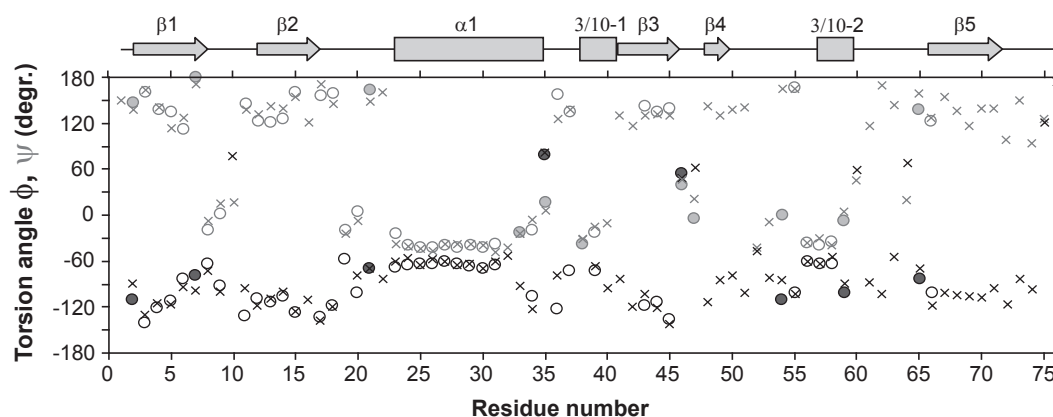
#### 8.4.2 Secondary Structure Analysis

As shown previously for selectively [242] or uniformly labeled [183] polypeptides, secondary chemical shifts provide a useful tool to describe protein secondary structure under MAS conditions. With the resonance assignments for UBI-P at hand, these parameters can be obtained readily. Helices and extended strands can be identified by manually comparing chemical shifts to average chemical shifts or shifts known for random coil protein segments. In order to provide quantitative information on the backbone torsion angles ( $\phi, \psi$ ), the computer program TALOS [58] was used here, which employs an advanced weighing function considering triplets of residues, and additionally includes protein sequence information.

In context of this study, it was demonstrated by M. Etzkorn that backbone topology can be refined using standard ( $^{15}\text{N}, ^{15}\text{N}$ ) or ( $^{15}\text{N}, ^{13}\text{C}$ ) correlation experiments [79, 254], as summarized here: (i) Dominating correlations in ( $^{15}\text{N}, ^{15}\text{N}$ ) proton driven spin diffusion experiments (NN) correspond to sequential transfer in right-handed helices and  $\beta$ -turn regions of the protein. Moreover, all correlations expected from  $\beta$ -sheet regions are in general weak or missing. The cross peak intensity is dependent on the MAS rate, with a zero-quantum (0Q) line-shape function ( $J^{0Q}(\omega_R)$ ) scaling with  $1/\omega_R^2$ . At a mixing time of 5 s and an MAS rate of 11 kHz, correlations should involve predominantly  $^{15}\text{N}$ - $^{15}\text{N}$  distances below 3 Å. (ii) Complementary to the NN case, ( $^{15}\text{N}, ^{13}\text{C}$ ) encoded ( $^1\text{H}, ^1\text{H}$ ) correlations (NHHC) are predominantly the result of  $\beta$ -sheet conformations. Longitudinal mixing in the order of 90  $\mu\text{s}$  reflects proton-proton interactions of a length scale of 2–3 Å. For glycine, strong intra-residue NHHC correlations are also possible for other secondary structure elements. In the case of ubiquitin, these mainly relate to  $\beta$ -turn regions of the protein. Additional correlations involving intra- and inter-residue  $\text{NH-HC}^\beta$  spin pairs may occur that encode additional information about the torsion angle  $\chi$  [322]. Finally, inter- $\beta$ -strand  $\text{NH-HC}^\alpha$  correlations relating to very short  $^1\text{H}$ - $^1\text{H}$  distances can occur. In two

spectral dimensions, these correlations cannot be assigned unequivocally for ubiquitin. As we have shown elsewhere, long-range intra-molecular NHHC distance constraint may be obtained using 3D spectroscopy to reduce spectral overlap (3D NHHCC) [113].

The results demonstrate that NN and NHHC correlations can provide a complementary means to characterize protein secondary structure if recorded under appropriate experimental conditions. Together with TALOS-derived torsion angle constraints (Figure 8.4, empty circles), these correlations can hence be used to increase the number of torsion-angle constraints. For UBI-P, torsion-angle predictions classified as *reliable* in TALOS are in



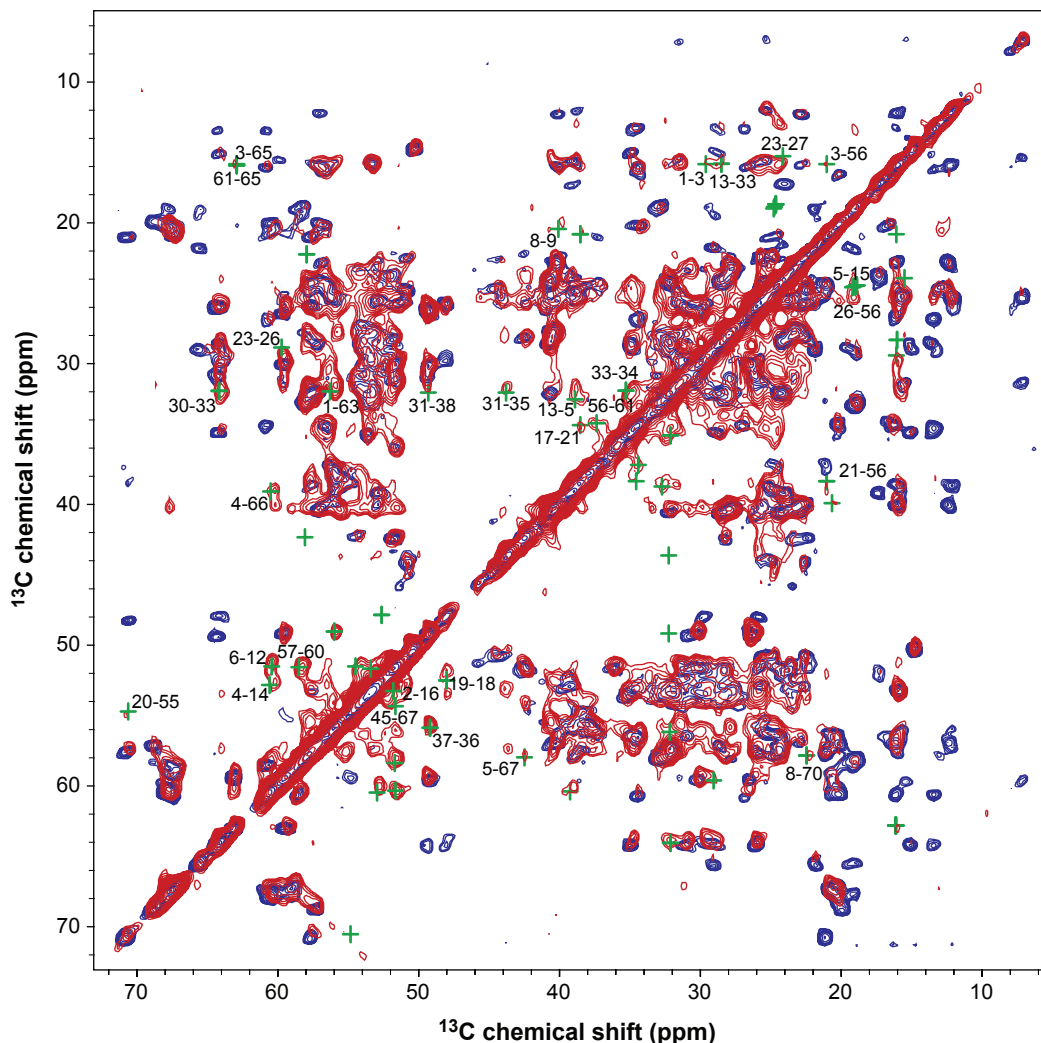
**Figure 8.4:** Comparison between dihedral angles as found in the crystal structure ( $\times$ ) with TALOS predictions ( $\circ$ ). Ambiguities in the TALOS analysis can be resolved with the aid of NN and NHHC spectra, resulting in a total of 19 additional dihedral angles (solid circles).

good agreement with the X-ray structure ( $\times$  symbols in Figure 8.4). The analysis of NN and NHHC spectra provides a total of 19 additional ( $\phi, \psi$ ) torsion-angle constraints (filled circles) that resolve ambiguities detected by TALOS and can be used in the context of a structure calculation. The corresponding residues are often found in loop regions of the protein, where TALOS predictions are known to be most unreliable. Only one ( $\psi_{54}$ ) of the total of 92 refined torsion angles differs significantly from values expected from the X-ray structure. These results confirm that the secondary structures of UBI-P microcrystals agree with data obtained by X-ray crystallography and solution-state NMR.

### 8.4.3 Investigation of 3D Structure

Figure 8.5 compares results of a ( $^{13}\text{C}, ^{13}\text{C}$ ) PDSM experiment (mixing time: 40 ms, MAS rate: 11 kHz, 600 MHz, blue) with CHHC [159, 160, 255] data (red) obtained for a

proton–proton mixing time of 250  $\mu\text{s}$ . There are notable differences between the two



**Figure 8.5:** Comparison between a ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) PDSD experiment (mixing time: 40 ms, MAS frequency 11 kHz, 600 MHz, blue) and a CHHC [159, 160, 255] spectrum (red) obtained for a proton–proton mixing time of 250  $\mu\text{s}$  on UBI-P. CHHC correlations that do not overlap with PDSD cross-peaks and are consistent with short, medium, or long-range (X–Y) contacts between residues X and Y as indicated by green symbols.

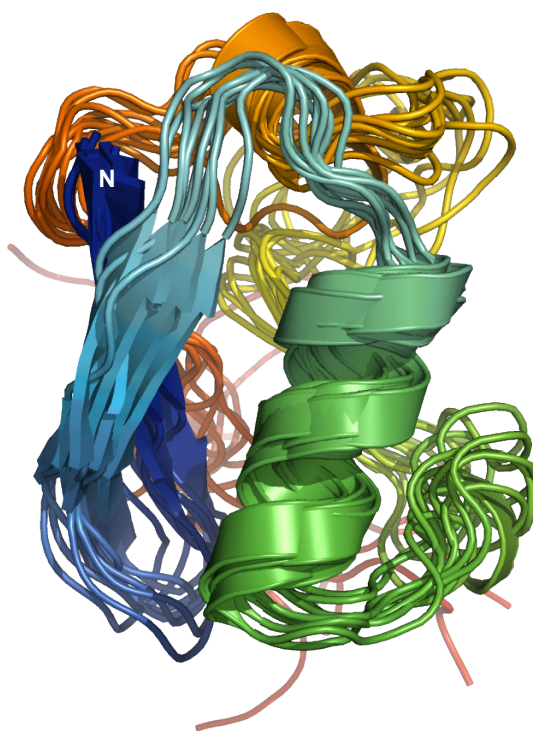
spectra, which are most pronounced in the  $\text{C}^\alpha$  region. Most of these CHHC correlations cannot be assigned unequivocally by 2D spectroscopy. However, by assuming a homology model, we can compare the CHHC data to predictions from the X-ray structure and resonance assignments obtained for UBI-P considering ( $^1\text{H}$ ,  $^1\text{H}$ ) distances up to 3.5  $\text{\AA}$ . According to ref. [160], such interactions should dominate the CHHC spectrum under the

experimental conditions considered here. Except for correlations involving methyl protons, which are attenuated due to the short CP time but appear for longer contact times, a variety of medium- and long-range contacts in the distance range under consideration can be identified. Correlations that do not overlap with intra-residue cross-peaks of the ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) spectrum are indicated by green symbols. For example,  $(i, i + 3)$  contacts such as 23–26 or 30–33 are observed for the central  $\alpha$ -helix. In addition, long-range contacts such as 4–66 or 5–67, connecting  $\beta$ -strand 1 ( $\beta_1$ ) and  $\beta$ -strand 5 ( $\beta_5$ ), are readily found. Likewise,  $\beta$ -strands 1 and 2 are connected by CHHC correlations 6–12 and 4–14. In total, 30 CHHC constraints were identified, relating to four sequential, nine medium-range, and 17 long-range contacts. The number of unequivocal CHHC distance restraints could be increased by use of three-dimensional correlation spectroscopy [113] or CHHC data with different mixing times (data not shown). Moreover, iterative methods demonstrated in ref. [161] or implemented in software routines such as CYANA [102] or ARIA [178] could be used to resolve ambiguities in the CHHC assignment process.

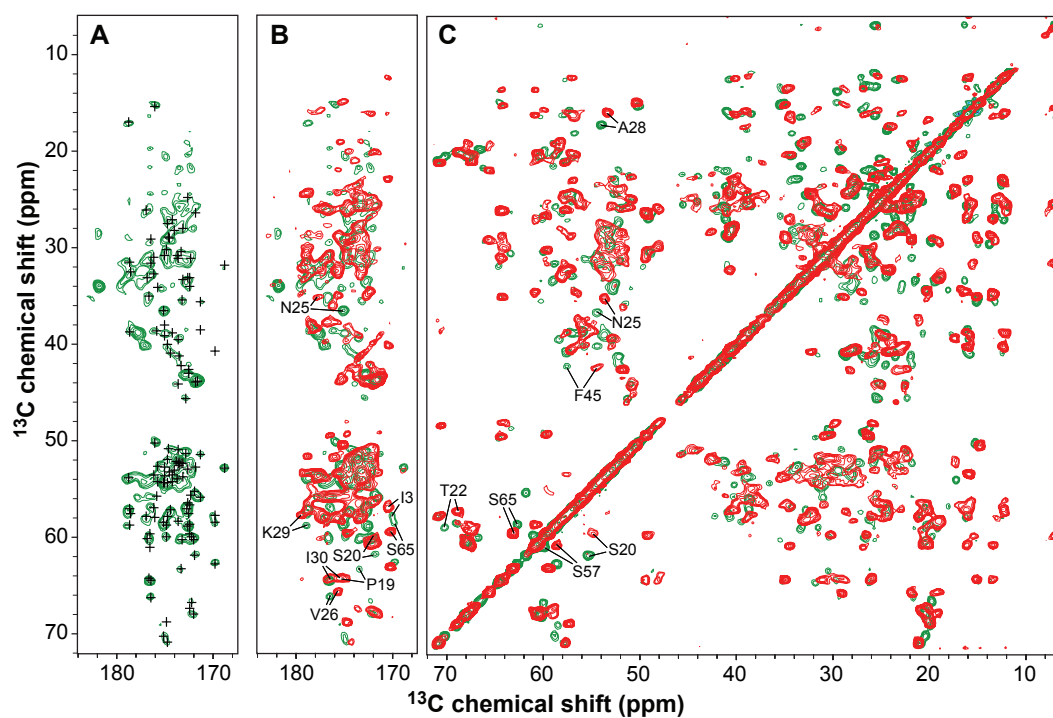
Next, it was investigated whether the CHHC constraints identified in Figure 8.5, together with torsion-angle constraints derived from a combined TALOS, NN, and NHHHC structure analysis, would be sufficient to construct a 3D molecular structure through the use of CNS. This approach hence does not deliver a *de novo* 3D structure of UBI-P microcrystals, but makes comparison of MAS-based NMR data sets to existing structures possible. The resulting ensemble of ten structures selected according to the lowest overall energy and aligned along the backbone atoms of residues M1 to V70 with the aid of MOLMOL is shown in Figure 8.6. The overall fold and the characteristic secondary structure elements are reproduced well, suggesting that the X-ray structure and the 3D structure of PEG-precipitated ubiquitin must be closely related. When comparing the effect of sample preparation on solid-phase protein structure one may hence utilize X-ray and solution-state NMR results as a reference.

## 8.5 The Effect of Sample Preparation

As previously demonstrated for Crh [44], comparison of MAS-NMR derived resonance assignments to structural data obtained from other resources provides a useful instrument with which to study solid-phase protein structure. First, ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) spectra of UBI-P (Figure 8.7B,C, red) were compared to results obtained with UBI-H (rehydrated, Figure 8.7B,C, green). While the detected  $^{13}\text{C}$  line width is comparable, chemical shift variations

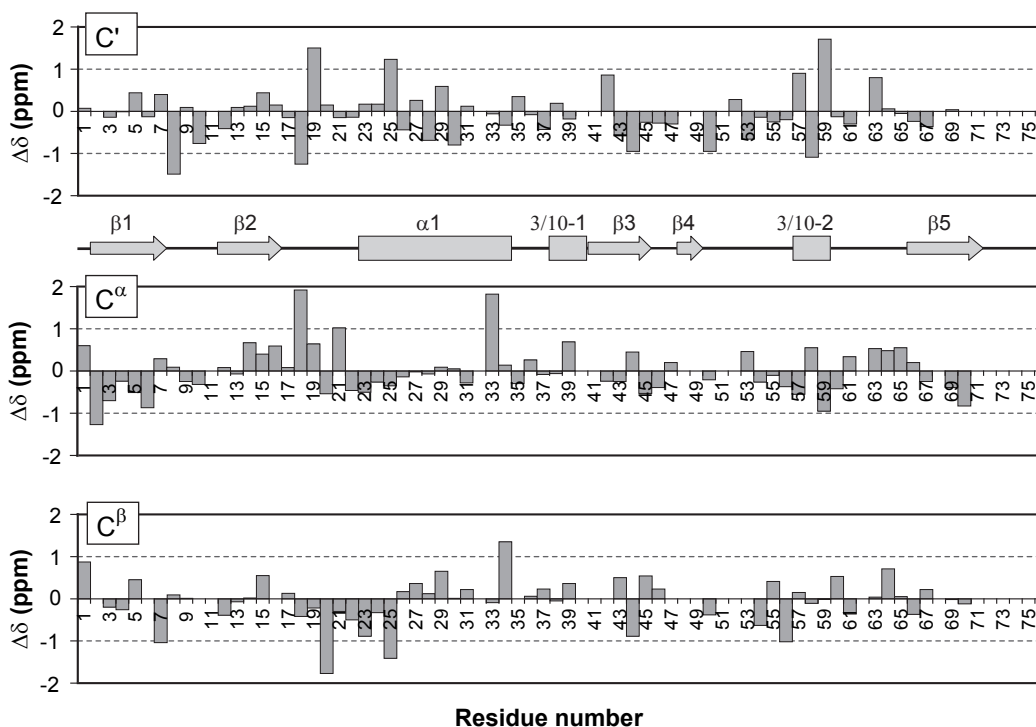


**Figure 8.6:** Ensemble of 10 UBI-P structures (backbone ribbon) calculated by use of CNS consistent with TALOS, NN, NHHc, and CHHC data. Structures were aligned along the backbone atoms of residues M1 to V70 with the aid of MOLMOL, the figure was prepared in PyMOL.



**Figure 8.7:** Comparison of ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) SD spectra for UBI-P (red) with experimental results obtained on UBI-H (green) with a SD mixing time of 40 ms SD and an MAS rate of 11 kHz at 600 MHz. Crosses in (A) mark predicted correlations from resonance assignments in UBI-M [126].

for both backbone and sidechain resonances are visible. Backbone  $C^\alpha$  and  $C^\beta$  resonances for S20, F45, and T22, for example, vary between 1 ppm and over 2 ppm, see (C). Similar variations are also seen for  $C'-C^\alpha$  and  $C'-C^\beta$  correlations, as visible in Figure 8.7B. Figure 8.7A compares correlations found for UBI-H with resonance assignments reported for UBI-M [126] (precipitated with MPD). Interestingly, the agreement between NMR spectra for those two preparations is far better than that between micro-crystals prepared from PEG and MPD. It can hence be concluded that hydrated ubiquitin adopts a solid-phase (supra)molecular structure that closely resembles that present in the case of MPD precipitants. Chemical-shift assignments of UBI-M have already been compared in ref. [126] with solution-state NMR data reported by Wand and co-workers [310], and these resonance assignments are in good agreement with values reported by Bax and co-workers [311]. In Figure 8.8, the latter set of values has been subtracted from chemical-shift assignments obtained for UBI-P. For the sake of clarity, chemical-shift changes are plotted only for  $C'$ ,



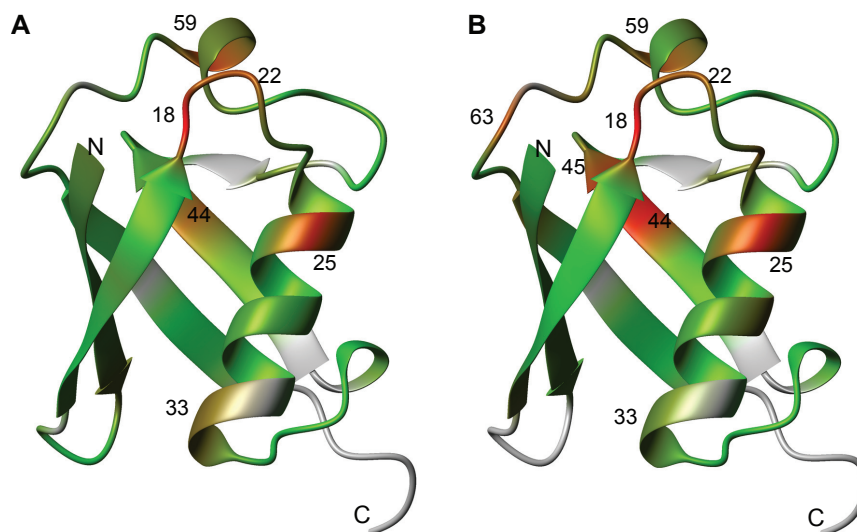
**Figure 8.8:** Difference between resonance assignments obtained for UBI-P and solution-state NMR shifts as reported by Wang et al. [311].

$C^\alpha$ , and  $C^\beta$ . For many residues, these variations are significantly larger than the natural  $^{13}\text{C}$  line width (ca. 0.5 ppm). At the N terminus, the strongest variations are observed for



residues 1 and 2 at the beginning of the first  $\beta$ -strand and, subsequently, within the loop connecting  $\beta$ -strand 2 and the central  $\alpha$ -helix. Interestingly, both regions exhibit a high degree of molecular motion when analyzed in weakly aligned, solubilized protein samples [227]. In the  $\alpha$ -helix  $\alpha 1$ , strong variations are found for residues 23 and 25, which have been shown to exhibit conformational exchange in solution [283, 284]. Such a mechanism [283] could also explain the large variations seen for E18. With the exception of residue 33, chemical shift variations are small for  $\alpha$ -helical residues 26–34, again in agreement with a recent solution-state NMR study on protein dynamics [227]. An accurate analysis for residues 40–50 is difficult, due to missing solid-state NMR assignments for residues 40, 41 and 49. Strong variations are again observed for the 3/10-helix 57–59, in agreement with ref. [227]. Finally, order parameters close to zero reported for the C terminus in the solution state are consistent with a highly flexible protein C-terminal segment in UBI-P, which gives rise to a strong reduction of ( $^1\text{H}$ ,  $^{13}\text{C}$ ) and ( $^1\text{H}$ ,  $^{15}\text{N}$ ) CP efficiency. These residues are missing in the ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) and ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) spectra for the experimental temperatures used in our study; this would be consistent with molecular mobility that gives rise to a reduction of through-space couplings and is only frozen out at lower temperatures.

In Figure 8.9A, the chemical shift changes shown in Figure 8.8 are displayed on the 3D (crystal) structure of ubiquitin. For simplicity, only the added norm of  $\text{C}'$ ,  $\text{C}^\alpha$ , and  $\text{C}^\beta$  chemical shift variations is plotted, ranging from small (green) to large (red) values. Unassigned residues are indicated in gray. With the exception of  $\alpha$ -helical residues known to undergo conformational exchange and the C terminus, the largest chemical variations are observed for the loop comprising residues 16–22 and the small 3/10-helix 57–59, which are found in close spatial proximity in the 3D structure. In contrast with changes at the end of  $\alpha$ -helix 1, these variations could hence speak in favor of structural differences between solution and solid-state conformation, due to increased molecular mobility. If this conclusion were correct, similar protein regions should be affected if protein microcrystals were prepared with use of MPD instead of PEG. Correspondingly, Figure 8.9B encodes chemical shift variations between UBI-P assigned in this study and resonance assignments on UBI-M as reported in ref. [126]. Indeed, many protein segments identified in Figure 8.9A also reveal strong chemical shift variations between UBI-P and UBI-M. In addition, substantial chemical shift changes are seen for F45 and K63. The reason for these variations can possibly be further elucidated by measuring site-resolved chemical shielding anisotropies,  $T_1$ ,  $T_2$ , and  $T_{1\rho}$  relaxation rates, or dipolar coupling order parameters under MAS conditions.



**Figure 8.9:** Chemical shift difference between (A) UBI-P and solution-state NMR [311], and (B) UBI-P and UBI-M [126], shown on the 3D (crystal) structure of ubiquitin. The added norm of  $C'$ ,  $C^\alpha$ , and  $C^\beta$  chemical shift variations is plotted, ranging from small (green, 0 ppm) to large (red, 3.6 ppm in (A), 4.7 ppm in (B)) values. Residues unassigned in at least one of the compared data sets are indicated in gray.

## 8.6 Conclusion

A series of correlation experiments was used to study the effect of sample preparation on a globular, solid-phase protein. The 3D molecular structure was studied by using a single  $U(^{13}C, ^{15}N)$ -labeled sample of ubiquitin, and chemical shift variations were evaluated in reference to solution-state NMR data. A comparative study of  $C'$ ,  $C^\alpha$ , and  $C^\beta$  resonance frequencies suggests that  $^{13}C$  chemical-shift variations are most likely to occur in protein regions that exhibit an enhanced degree of molecular mobility. Complementary information on molecular dynamics may be obtainable from a residue-specific analysis of  $^{13}C$  and  $^{15}N$  chemical-shielding anisotropies,  $^1H$  and  $^{15}N$  resonance frequencies,  $T_1$ ,  $T_2$ , and  $T_{1\rho}$  relaxation rates, or  $^{13}C$ - $^1H$  and  $^{13}C$ - $^{13}C$  dipolar coupling order parameters [255]. The design and application of these techniques in well characterized proteins such as ubiquitin provides a useful reference for future studies of molecular structure and dynamics in (membrane) proteins of unknown structure by high-resolution solid-state NMR.

## Chapter 9

# Structural Characterization of $\text{Ca}^{2+}$ -ATPase-bound Phospholamban in Lipid Bilayers

## 9.1 Summary

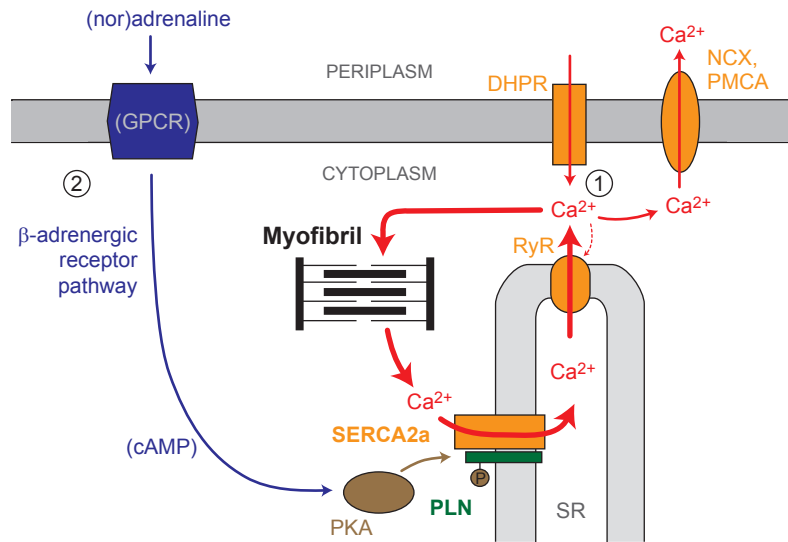
Phospholamban (PLN) regulates cardiac contractility via its modulation of Sarco(endo)plasmic Reticulum Calcium ATPase (SERCA) activity. Impairment of this regulatory process causes heart failure. While PLN and SERCA1a, an isoform from skeletal muscle, have been structurally characterized in great detail, direct information about the conformation of PLN in complex with SERCA has been limited.

Solid-state NMR (ssNMR) spectroscopy was used to deduce structural properties of both the  $\text{A}_{36}\text{F}_{41}\text{A}_{46}$ -mutant (AFA-PLN) and wild-type PLN (WT-PLN) when bound to SERCA1a after reconstitution in a functional lipid bilayer environment. Chemical shift assignments in all domains of AFA-PLN provide direct evidence for the presence of two terminal  $\alpha$ -helices connected by a linker region of reduced structural order that differs from previous findings on free PLN. SsNMR experiments on WT-PLN show no significant difference in binding compared to AFA-PLN, and do not support the coexistence of a significantly populated dynamic state of PLN after formation of the PLN/SERCA complex. Combination of spectroscopic data with biophysical and biochemical data using flexible protein–protein docking simulations provides a structural basis for understanding the interaction between PLN and SERCA1a.

## 9.2 Introduction

Sarco(endo)plasmic Reticulum  $\text{Ca}^{2+}$ -ATPase (SERCA) is the rate-determining enzyme for  $\text{Ca}^{2+}$  removal from muscle cytoplasm, a process that facilitates the relaxation of myo-

fibrils subsequent to contraction, see Figure 9.1 for a general scheme. Phospholamban (PLN) regulates cardiac contractility by modulating the activity of SERCA2a, the heart muscle SERCA isoform [260, 278]. Regulatory dysfunction of PLN on SERCA2a has been



**Figure 9.1:** Simplified 'transistor' scheme for interactions between two cardiac signaling pathways, following [187]. (1) Upon depolarization,  $\text{Ca}^{2+}$  enters the cytoplasm through channels in the plasma membrane, triggering the release of additional  $\text{Ca}^{2+}$  from storage in the sarcoplasmic reticulum (SR). The elevated  $\text{Ca}^{2+}$  concentration causes myofibril (i.e., muscle) contraction. Subsequent muscle relaxation is facilitated by removal of  $\text{Ca}^{2+}$  from the cytoplasm. Accounting for 70% of myoplasmic  $\text{Ca}^{2+}$  turnover, SERCA2a is the rate-determining enzyme. Unphosphorylated phospholamban (PLN) limits calcium cycling by reversely binding to and inhibiting SERCA2a. (2) The  $\beta$ -adrenergic signaling pathway targets PLN for phosphorylation by protein kinase A (PKA), resulting in loss of PLN inhibitory function. As a consequence, the rate of cardiac relaxation, and contractility on subsequent beats, is increased. PLN is dephosphorylated by a protein phosphatase. Abbreviations: DHPR, dihydropyridine receptors; PMCA, plasma-membrane  $\text{Ca}^{2+}$ -ATPases; NCX,  $\text{Na}^+/\text{Ca}^{2+}$  exchangers; RyR, Ryanodine receptors; GPCR, G-protein coupled receptors; cAMP, cyclic AMP.

identified as a cause of dilated cardiomyopathy, and alterations of PLN activity can have manifold influence on the progress of heart disease, making the PLN–SERCA complex an important target for therapeutic intervention [187].

SERCA is a membrane protein of  $\sim 110$  kDa molecular weight, existing in a number of highly conserved isoforms [225]. The skeletal muscle isoform SERCA1a (994 amino acid residues) has been structurally and functionally characterized in great detail, including the structure determination of several reaction intermediates [214, 215, 277, 289, 290, 291].

In particular, the structure of the  $\text{Ca}^{2+}$ -free  $\text{E}_2$  intermediate of SERCA1a, to which PLN can bind, is known at high resolution from X-ray crystallography [213, 215, 291].

PLN is a 52-residue protein, which structure has been studied by a variety of biophysical techniques. In line with other methods, NMR spectroscopy revealed a transmembrane helix and a mobile N-terminal domain with some differences in the secondary structure of monomeric PLN mutants [9, 156, 293, 325]. Wild-type PLN predominantly exists as a symmetric homopentamer with a pore-resembling core spanning the membrane [217, 264], where the cytoplasmic domains can associate with the lipid bilayer [294].

So far, detailed structural information obtained directly on PLN in complex with SERCA has been limited. The binding interface of PLN with SERCA1a or SERCA2a, regarding both the transmembrane and cytoplasmic domains of PLN, was described by mutagenesis screening of PLN [148, 149, 287] and SERCA [13, 288], as well as cross-linking studies [51, 52, 53, 132, 137, 292]. Magic angle spinning (MAS) ssNMR has been used to selectively probe internuclear distances in pairwise isotope-labeled PLN variants bound to SERCA1a [123], while for example EPR and FRET have been employed to study overall conformational changes upon complex formation [150, 205]. Current models for the structure of PLN when bound to SERCA1a ( $\text{E}_2$  state) were obtained from computational approaches using mutagenesis and cross-linking constraints for all domains of PLN [125, 292] or with a focus on the transmembrane region [52, 53]. In these studies, the initial template structure of PLN for molecular modeling was taken from investigations of free PLN [156, 232]. Solution-state NMR data obtained for AFA-PLN in the presence of SERCA were interpreted in terms of two or more transient conformations that are present during PLN-SERCA interaction, structurally different from results seen for free PLN in solution [326]. Additionally, the existence of a transient WT-PLN intermediate during complex formation with SERCA was hypothesized, based on a structural model of the pentamer [217]. However, these studies were conducted in micelles where complex formation and protein activity can be compromised [326] and did not provide direct spectroscopic information about PLN in complex with SERCA.

Here, high-resolution MAS ssNMR spectroscopy was applied to multiply ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-labeled variants of AFA- and WT-PLN in complex with SERCA1a in order to (i) directly infer the backbone binding mode of PLN in complex with SERCA in a functional bilayer environment, (ii) investigate the presence of any additional, dynamic PLN conformations in this preparation and (iii) generate a structural model that is compatible with biochemical data.

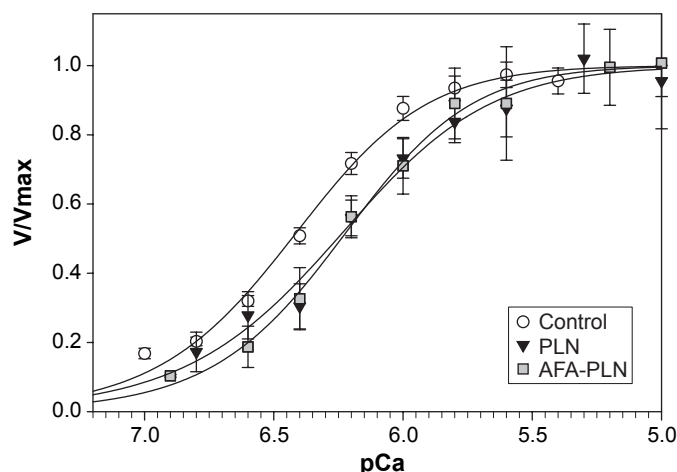
## 9.3 Material and Methods

### 9.3.1 Sample Preparation

Fully uniformly (U) (<sup>13</sup>C,<sup>15</sup>N) labeled recombinant AFA-PLN was expressed in *E. Coli* grown on U(<sup>13</sup>C)-glucose and (<sup>15</sup>N)-NH<sub>4</sub>Cl [71]. The primary sequence composed of 54 residues (PLN construct with initial Gly-Ser-) is given in Figure 9.3A. Skeletal muscle vesicles were prepared from rabbit hind leg [76], and the skeletal SR Ca<sup>2+</sup>-ATPase (SERCA1a) was purified by Reactive-Red Sepharose affinity chromatography [330]. Hydrated proteoliposomes were prepared with DOPC as described by Hughes et al. [123]. For the reconstitution, the molar ratio of SERCA1a/PLN/DOPC was 1:1.2:160. Typically, 22 mg of purified Ca<sup>2+</sup>-ATPase protein was reconstituted with 1.5 mg of labeled PLN or AFA-PLN. ATPase activity of the proteoliposomes was assayed as described below. Reversely IL and ILRKMYF labeled PLN samples were obtained from adding amino acids of natural <sup>13</sup>C and <sup>15</sup>N isotope abundance to the expression medium at concentrations of 100 mg/L for each amino acid. In contrast to Ile, Leu, Lys, Tyr and Phe, reverse labeling of Arg residues was not complete as judged by 2D (<sup>13</sup>C,<sup>13</sup>C) (Figure 9.3C) and 1D <sup>15</sup>N ssNMR (data not shown) experiments. Nevertheless, a significant reduction of the corresponding ssNMR signals facilitated the assignment process. In addition, amino-acid specific labeling was obtained by adding <sup>13</sup>C and/or <sup>15</sup>N labeled amino acids to the natural isotope abundance growth medium.

### 9.3.2 Activity Assay

*The following activity assay data was provided by H. Young, University of Alberta, Canada.* ATPase activity of the proteoliposomes was assayed by measuring the absorbance of NADH at 340 nm over time in a coupled-enzyme assay (Figure 9.2) as previously described [313]. The activity was measured in the range of 0.1 to 10 μM free Ca<sup>2+</sup>. Full pCa curves were measured for each co-reconstitution. The data for at least three co-reconstitutions were averaged. The activity of each co-reconstitution was matched to a control of Ca<sup>2+</sup>-ATPase reconstituted in the absence of PLN or AFA-PLN that had been prepared at the same time (Figure 9.2).



**Figure 9.2:** ATPase activity as a function of free  $\text{Ca}^{2+}$  concentration for  $\text{Ca}^{2+}$ -ATPase reconstituted in the absence (circles) and in the presence of recombinant wild-type PLN (triangles) or AFA-PLN (squares). The raw data were normalized to the maximal activity calculated based on fitting the data to the Hill equation ( $n = 5$  for Control;  $n = 4$  for wild-type PLN;  $n = 3$  for AFA PLN). The data are plotted as normalized ATPase activity versus pCa. Each data point is the mean  $\pm$  SEM.

### 9.3.3 Solid-State NMR Experiments and Analysis

Samples were packed in standard 4 mm  $\text{ZrO}_2$  rotors with top spacer and Kel-F caps. Experiments were carried out on Avance and Avance II NMR spectrometers (Bruker Biospin, Germany) equipped with triple-resonance ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) wide- or standard-bore probes, at magnetic fields corresponding to 400, 600 and 800 MHz  $^1\text{H}$  Larmor frequency. Magic angle spinning up to  $\omega_{\text{R}}/2\pi = 12.5$  kHz was performed using low-temperature VT gas chillers at temperatures ranging from 235 K to 275 K.

Residue identification and sequential resonance assignments were made from a homonuclear  $^{13}\text{C}$  double-quantum / single-quantum (2Q/1Q) correlation spectrum employing SPC5 [117] 2Q recoupling, and from 2D spectra with mixing units using Proton-Driven Spin Diffusion (PDS) [35, 275] or Spin Diffusion under Weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) Coupling Conditions (SDWC), see Chapter 7 [253]. Additionally, NCOX and NCACX correlation spectra [21], see also Chapter 3, were recorded with SPECIFIC CP [25] for NCO or NCA polarization transfer, followed by homonuclear  $^{13}\text{C}$  PDS or DARR [282] mixing. Mobile protein segments were probed for by a combination of through-bond ( $^1\text{H}$ ,  $^{13}\text{C}$ ) INEPT [203] and ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) TOBSY [23] mixing, suppressing signal from residues exhibiting little or no

mobility [9]. During dipolar and scalar-based correlation experiments, SPINAL-64 [88] or GARP [258] proton decoupling were applied at 85 kHz and 10 kHz, respectively. All but through-bond spectra were recorded with 5–7 days of signal averaging, the spectrum in Figure 9.3C was recorded over 9 days.

Signal processing was carried out in XWinNMR 3.5 or Topspin 2.0 (Bruker Biospin, Germany), employing zero-filling and exponential line broadening or square-sine-bell apodization functions. Two-dimensional spectral analysis was done in Sparky 3.110 [92]. Secondary chemical shifts  $\Delta\Delta\delta = (\delta_{CA}^{\text{exp}} - \delta_{CA}^{\text{av}}) - (\delta_{CB}^{\text{exp}} - \delta_{CB}^{\text{av}})$  [242] for each residue were calculated using average chemical shifts from the BioMagResBank [249].

### 9.3.4 Modeling Using Flexible Protein-Protein Docking Simulation

Docking models of PLN in complex with SERCA1a were calculated using the Crystallography and NMR System (CNS) [42] with protocols for High Ambiguity Driven biomolecular DOCKing (HADDOCK) [70] on a multi-node computer cluster. A linearly stretched, partially pre-folded docking template structure for PLN was generated using a standard simulated annealing protocol in CNS, and then docked to a structure for SERCA1a in the E2 state, Protein Data Bank (PDB) [31] entry 2AGV [213]. Both protein structures were initially separated by 150 Å in random orientations. Docking was driven by intermolecular distance constraints summarized in Table 9.1. For details, see Appendix Chapter G. Over the course of the docking simulation, PLN was constrained to a helical backbone conformation for the residues discussed below, and fully flexible elsewhere. Interface residues of SERCA were allowed to have flexible side chains or backbone during various stages of the docking. Other parts of SERCA remained rigid during the simulation.

## 9.4 Results

### 9.4.1 SsNMR Assignments and Secondary Structure of Mutant PLN

One- and two-dimensional (<sup>13</sup>C, <sup>13</sup>C) and (<sup>13</sup>C, <sup>15</sup>N) solid-state NMR correlation experiments were used to determine NMR chemical shifts of (monomeric) AFA-PLN in complex with SERCA1a. In order to reduce spectral overlap and ambiguities in the primary structure during assignments, studies were performed with fully and reverse-labeled AFA-PLN, as well as two samples with amino acid-selective labeling. Isotope labeling schemes are displayed in Figure 9.3A.



PLN	SERCA	Max. dist. (Å)	Link type	Reference
3	397', 400'	15	Denny-Jaffe	[132]
4C <sup>β</sup>	364'	3	Artificial	Potential binding groove [292]
12C <sup>β</sup>	358'	3	Artificial	Potential binding groove [292]
27C <sup>β</sup>	321'C <sup>β</sup>	5	SS	[292]
27	328'	10	EMCS	[51]
30	318'	10	BMH	[137]
30	321'	7	Cu-phe, but no SS	[292]
30	328'	15	KMUS	[51]
31	317'	10	BMH	[52]
49C <sup>β</sup>	89'C <sup>β</sup>	5	SS and Cu-phe	[53, 292]
50	89'	10	BMH	[53]
52	89'	7	bBBr	[53]

**Table 9.1:** Summary of cross-link restraints used for the docking between AFA-PLN and SERCA1a. Additional Ambiguous Interaction Restraints (AIR), as described for use in HADDOCK [70], were used to enforce interface contacts, such as residues with known loss-of-function mutations (see Appendix Chapter G). Of these residues, the following were defined as 'active', mandating an ambiguous interface contact: (SERCA1a) 89, 321, 802, 805, 809, and (AFA-PLN) 2, 4, 7, 9, 12, 27, 31, 34, 35, 38, 41, 42, 45, 48, 49, 52. Abbreviations: SS, disulphide bond; EMCS: N-( $\epsilon$ -maleimidocaproyloxy)succinimide ester; BMH, 1,6-bismaleimidohehexane; Cu-phe, Cu-phenantrolone; KMUS, N-( $\kappa$ -maleimidoundecanoyloxy)sulfosuccinimide ester; bBBr, dibromobimane.



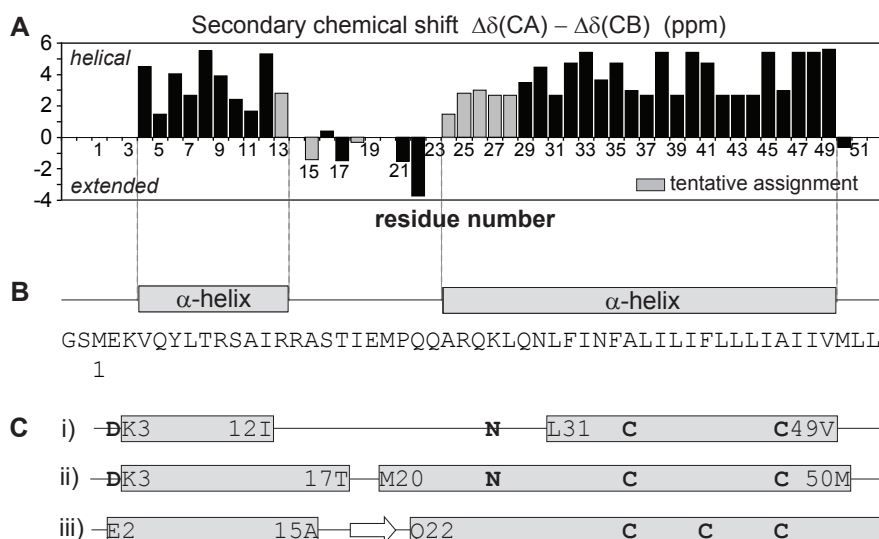
see Figure 9.3B. The spectrum of fully labeled AFA-PLN, however, is dominated by intra- and inter-residue correlations involving Ile and Leu. Using reverse-labeling, a significant reduction of spectral overlap (Figure 9.3C) and assignment ambiguities in the primary sequence (Figure 9.3A) was achieved. V4–Q5, S16–T17 and Q29–N30, for example, can now unambiguously be identified (Figure 9.3C). In addition, a sequential assignment for R13C' was obtained from a 1D SPECIFIC NCO-edited  $^{13}\text{C}$  spectrum for ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-Arg labeled AFA-PLN (Appendix Figure G.3). Spectroscopy of ( $^{13}\text{C}$ -Thr,  $^{15}\text{N}$ -Ile)-AFA-PLN assisted in resolving Thr resonances (Appendix Figure G.4). Assignments are in line with 2D NCACX and NCOCX correlation spectra for U( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-AFA-PLN (Figure G.2), however, due to lower resolution, dispersion and sensitivity in these spectra, the assignments were primarily obtained from the homonuclear carbon correlations.

Even without unambiguous sequence-specific residue identifications, many cross-peaks could be assigned to degenerate residue groups that occur exclusively within the same domain of AFA-PLN. This included LI/IL pairs between L37 and I44, which gave rise to numerous, strong sequential correlations (Figure 9.3B), or residues F32, F35 and F41. Met ( $\text{C}^\alpha, \text{C}^\beta$ ) chemical shift correlations corresponding to a  $\beta$ -strand [312] are not visible in the 2Q/1Q correlation spectrum, whereas signal was detected for random coil shifts [312], suggesting that there is no well-ordered  $\beta$ -strand around M20, [217].

A qualitative separation of monomeric PLN segments exhibiting helical, extended and coil backbone character was subsequently performed using secondary chemical-shift analysis. Backbone carbon resonances for V4, Q5, T8, S10–R13 as well as Q29, N30, N34, V49 and non-specifically assigned residues between F32 and I45 have helical character, whereas, for example, T17, Q22 and M50 show non-helical propensities. In addition, amino-acid specific resonances of Ala (weak), Glu, Met, Ile, Ser, and Leu were detected, corresponding to coil-like backbone structure (such as E2, A15, I18, E19), and Arg was found in different backbone configurations (Figure 9.3, Appendix Figure G.3). Due to signal overlap, Lys correlations indicative for a specific secondary structure could not be assigned.

Taken together, these findings suggest the existence of two helices comprising, at least, V4–R13 and Q29–V49, flanked by several residues of different secondary structure. Additional chemical shifts may be assigned to residues within helices which cannot be distinguished due to resonance overlap, such as Y6. Finally, taking into account earlier ssNMR distance measurements regarding A24C $^\alpha$ –Q26N [123] and the chemical shift of A24C $^\alpha$ , indicating a helical backbone in this region (including R25), one can deduce a non-helical

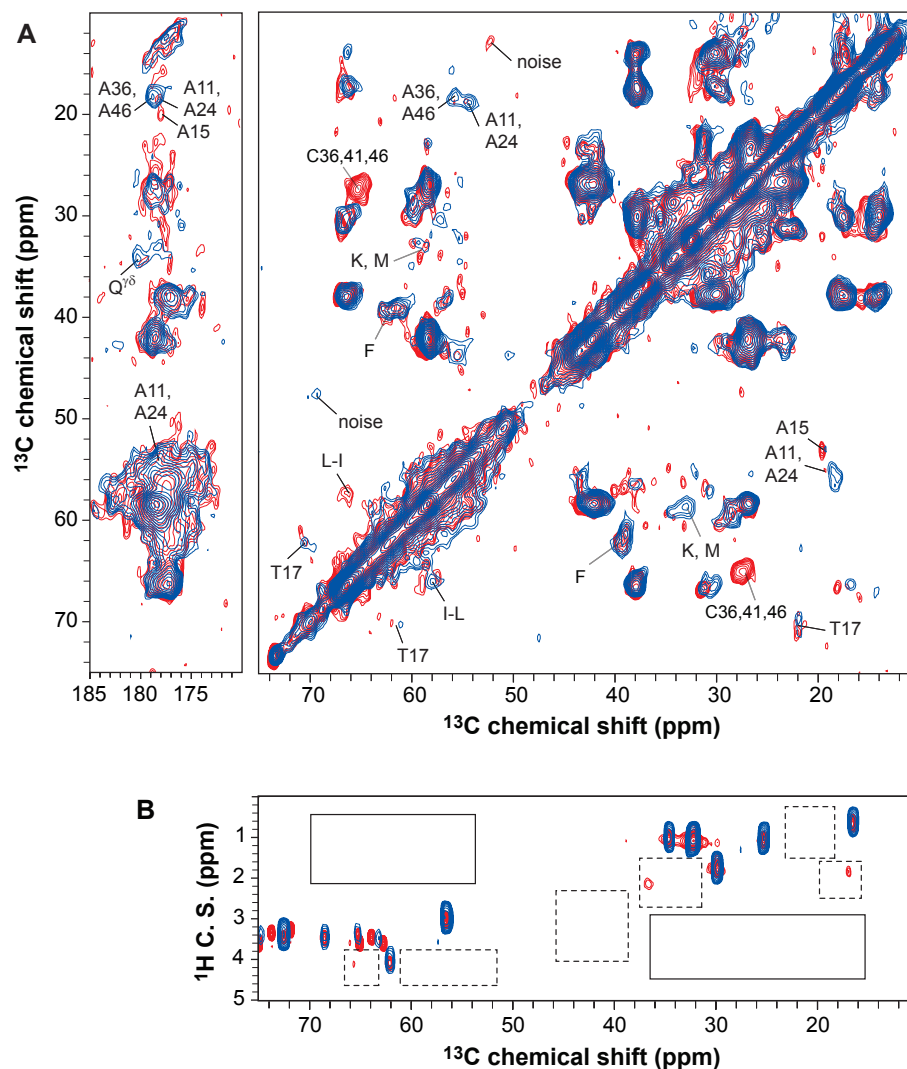
linker range from R14 to Q22. Secondary chemical shifts of P21 and Q22 exclude continuation of the C-terminal  $\alpha$ -helix in these two residues. A summary of the secondary chemical-shift analysis for AFA-PLN in complex with SERCA1a is displayed in Figure 9.4A. Secondary chemical shifts depending indirectly on previously published data for A24/Q26 [123], as well as R13, depending on R13C' only, are colored gray.



**Figure 9.4:** (A) Secondary structure analysis of PLN bound to SERCA by means of secondary chemical shifts from ssNMR, and (B) corresponding extent of a helical backbone, indicated by gray boxes. (C) Comparison to previous models, (i) AFA-PLN/SERCA model by Toyoshima et al. [292], (ii) solution-state NMR results by Lamberth et al. [156] as used for the complex model by Hutter et al. [125], and (iii) wild-type PLN in detergent micelles forming a pentamer [217], as discussed in context with rigid-body docking [53]. In another pentamer model (not shown), the N-terminal helix further includes S16 [294]. Mutants with respect to (A) are highlighted by bold letters.

### 9.4.2 Wild-Type PLN and PLN Mobility

In order to determine if structural changes previously seen between the AFA mutant and wild-type PLN alone are present in complex with SERCA, ssNMR experiments were conducted on a U(<sup>13</sup>C,<sup>15</sup>N)-WT-PLN/SERCA complex (Figure 9.5). Cysteine residues in WT-PLN gave rise to distinct ( $C^\alpha, C^\beta$ ) correlations not observed in AFA-PLN in a (<sup>13</sup>C,<sup>13</sup>C) correlation spectrum with 15 ms PDSM mixing (Figure 9.5A), corresponding to  $\alpha$ -helical backbone conformation of the non-oxidized form. Ala ( $C^\alpha, C^\beta$ ) cross peaks were



**Figure 9.5:** (A) Comparison of 2D  $(^{13}\text{C}, ^{13}\text{C})$  correlation spectra from 15 ms PSD mixing for AFA-PLN (blue) and WT-PLN (red). (B) Mobility-filtered  $(^1\text{H}, ^{13}\text{C})$  correlation spectrum from INEPT-TOBSY transfer of U-AFA-PLN/SERCA1a (blue) and INEPT transfer of U-WT-PLN/SERCA1a (red). Spectra were recorded with high-power SPINAL-64 or 10 kHz GARP decoupling in case of AFA- and WT-PLN, respectively, resulting in different signal intensities for various lipid sites.

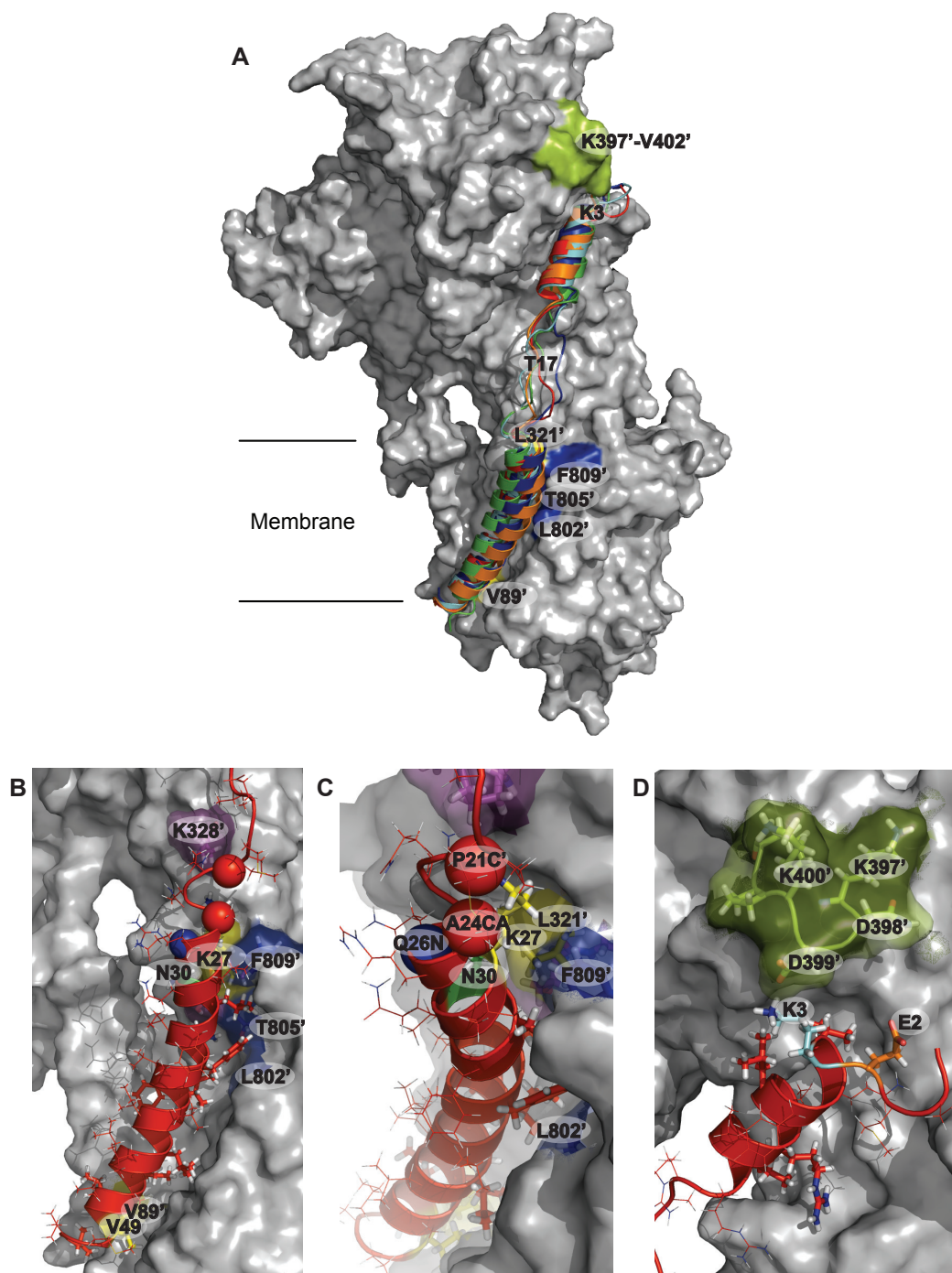
clearly diminished, owing to the absence of A36 and A46. While overall the sample yields a lower signal-to-noise ratio, other cross peaks appear identical at the given signal-to-noise level. In both spectra, recorded without freezing the samples, residues not buried in the membrane (e.g., M, E, Q, R) give rather weak signal, indicative of reduced structural disorder or increased mobility.

Protein segments exhibiting molecular motion in the micro- to nanosecond regime were probed by mobility-filtered INEPT or INEPT-TOBSY (<sup>1</sup>H,<sup>13</sup>C) correlation experiments (Figure 9.5B). In contrast to free PLN, these spectra were devoid of typical protein correlations after one-bond <sup>1</sup>H-<sup>13</sup>C and additional <sup>13</sup>C-<sup>13</sup>C transfer (dashed-framed and solid-framed boxes, respectively), as would be found for protein segments showing a high degree of mobility [9].

### 9.4.3 Docking Model

Next, docking calculations were performed to infer if the spectroscopic data are compatible with previous biochemical studies. The five models with best docking score are shown in Figure 9.6A. Close-up views for the best model, fulfilling all of the experimental constraints brought forward by ssNMR, cross-linking and mutagenesis, are given in Figure 9.6B–D.

Within the secondary structure boundaries given by ssNMR, the N-terminus of the cytoplasmic helix of AFA-PLN is found near the K397'–V402' loop of SERCA1a (Figure 9.6A,D) (*SERCA residues denoted with a prime*) while the C-terminus of the helix intimately connects to SERCA1a residues at the surface of transmembrane helices M4, M5, M6, M9 (Figure 9.6A–C). Residues 802', 805' and 809', which exhibit loss-of-function mutations [13], are in contact with AFA-PLN (Figure 9.6B,C). The distances between K27C<sup>β</sup> (AFA-PLN) and L321'C<sup>β</sup> (SERCA1a) as well as V49C<sup>β</sup> and V89'C<sup>β</sup> are compatible with the formation of disulphide cross-links of their respective cysteine mutants [52, 53, 292]. Furthermore, intra-molecular AFA-PLN distances described by Hughes and Middleton [123] for A24C<sup>α</sup>–Q26N and P21C'–A24C<sup>α</sup> (Table 9.1) are accounted for. Residues P21 to Q23 form a loop allowing for dihedral angles compatible with Q22 backbone torsion angle predictions from chemical shifts. The proximity of charged side chains E2 and K3 to K397'–V402', as deduced from charge-altering mutations leading to the loss of inhibitory function [287, 288], is fulfilled (Figure 9.6D).



**Figure 9.6:** Model of AFA-PLN in complex with SERCA1a. (A) Cluster of five conformations of AFA-PLN with lowest docking score (cartoon) and SERCA1a (PDB ID: 2AGV, gray surface). (B–D) Close-up views of the best scoring AFA-PLN conformer from (A), for the C-terminal (B,C) and N-terminal (D) regions. SERCA residues that can be cross-linked with PLN by disulphide bonds (V89' and L321') are displayed as yellow patches, other cross-linked residues magenta, residues 802', 805' and 809' are shown in blue, 397'–402' in green. Spheres indicate positions of specific isotope labels in an earlier ssNMR study [123]. Residues of AFA-PLN and SERCA related to loss-of-function mutants and residues studied by chemical cross-linking are displayed as sticks, other AFA-PLN residues are displayed as lines.

## 9.5 Discussion

### 9.5.1 Comparison to Previous Studies

Models of the PLN-SERCA complex have previously relied on structural data that was obtained for free PLN or deduced from biochemical data, such as mutagenesis and cross-linking. The corresponding secondary structures are given in Figure 9.4C and relate to (i) a model by Toyoshima and co-workers [292] based on two altered and computationally refined solution-state NMR structures for PLN [156, 232], and (ii) a model by Hutter et al. [125], which, during a docking simulation, enforced NOE-constraints in solution for a free PLN monomer [156]. As discussed in context of complex formation [53, 217], the secondary structure of WT-PLN pentamers seen by solution-state NMR spectroscopy in detergent micelles [217] is shown in Figure 9.4C (iii). Helix identifications from ssNMR spectroscopy are shown in Figure 9.4B, which points to differences in linker length in comparison to the previous studies.

In terms of the three-dimensional PLN/SERCA complex, the ssNMR-based model differs substantially from earlier results by Hutter and co-workers [125], in which the C-terminal domain is shifted towards the cytosol. Instead, the membrane-spanning helix ending close to the SR luminal side of the lipid bilayers, see Figure 9.6A–B, based on the recently described strong cross-link affinity of V49C to V89'C via a disulphide bridge [292] or Cu-Phe [53]. Elsewhere, the position of the C-terminal helix in the model presented here results from constraints derived from cross-links involving N30 and N27 [51, 52, 53], as well as the use of mutagenesis-derived Ambiguous Interaction Restraints (AIRs). It is in line with hydrogen-exchange measurements which indicate that L44 is buried within the membrane [123]. Furthermore, the secondary structure of AFA-PLN was directly obtained from the complex with SERCA1a, and not enforced to comply with the experimental NMR constraints on the structure of PLN alone [156], alleviating restraints on almost two full helix turns compared to free PLN.

The docking model described here shows similarity to the model proposed by Toyoshima and MacLennan [292]. For the C-terminus (Figure 9.6B), both models reflect the formation of a disulphide bond between V49C/V89'C double mutants and describe the absence of helical backbone structure for AFA-PLN residues L51 and L52. Unwinding of the helix at V49, however, is not in agreement with spectroscopic results and the resulting model shown here. Furthermore, the C-terminal helix extends beyond L31, including N30 and Q29 (Figures 9.4 and 9.6). Still, residue positions are in line with the K27C–L321'C



disulphide bond [292], and recently derived cross-linking constraints [51, 52, 53, 137]. In addition, the model described here is compatible with the finding that A24 and Q26 are part of the helix [123], and incorporates the P21C'–A24C $\alpha$  and A24C $\alpha$ –Q26N distances [123]. Despite a longer C-terminal helix, it was also possible to dock the cytoplasmic domain of AFA-PLN to residues in a pocket below the K397'–V402' SERCA1a loop. K3 is found within 15 Å of K397'/K400' [132], and proximity of E2 and K3 to K397'–V402' can facilitate interactions (Figure 9.6D). Residues exhibiting loss-of-function mutants are in contact with SERCA1a in both models.

Along with chemical cross-linking studies [51, 52, 53], rigid-body docking based on free PLN was used to align the transmembrane helix with the structure of SERCA1a, raising the question if PLN K3 could reach near SERCA K397'–V402' at all. Results illustrated here demonstrate that AFA-PLN computationally docked to the E<sub>2</sub> state of SERCA1a does satisfy the aforementioned interface constraints, when structural changes upon binding, facilitated by protein flexibility in the simulation, are taken into account. In addition, the SERCA structure in complex with PLN may differ from the docking template used (PDB ID: 2AGV). However, changes such as seen in recent crystal structures obtained using different or no inhibitors [215, 279] could be adapted for by small structural alterations of PLN in the model described here.

### 9.5.2 Implications for Complex Formation and Phosphorylation

The secondary structure of AFA-PLN in complex with SERCA1a (Figure 9.4B), implies a loss of secondary structure for several residues in the linker region compared to free AFA-PLN [156, 325], and to a lesser extent WT-PLN [217] (Figure 9.4C). In light of the structural model shown in Figure 9.6, this conformational flexibility is required for adaptation to the transmembrane and cytoplasmic SERCA binding sites. The existence of an extended linker already in the free WT form of PLN may be of advantage in complex formation. Indeed, the number of non-helical linker residues of PLN in the complex is rather similar to WT-PLN [217] (Figure 9.4C (iii)). However, this pentamer model cannot be docked to SERCA1a without violation of interface constraints or a structural rearrangement, as can be seen from manual docking using 3D visualization software (data not shown).

The importance of P21 to sustain unstructured characteristics in the segment preceding the C-terminal helix providing conformational flexibility has already been proposed from a structure determination of free PLN [156]. An increase of helix propensity by P21A

mutagenesis was found to reduce the ability of the N-terminal domain of PLN to efficiently inhibit SERCA1a [175]. In fact, P21 does not only show chemical shift characteristics of a statistical random coil protein backbone, but also bears an increase in ssNMR signal intensity when lowering the sample temperature (Figure 9.3B,C and Appendix Figure G.2), indicative of dynamic disorder in this protein segment. This notion further agrees with difficulties in assigning resonances from the linker segment in the 2D ssNMR data, due to unfavorable signal intensities or broadened resonance lines for these residues. Additional ssNMR experiments at elevated temperature could be used to further delineate the influence of structural and dynamic disorder and to potentially quantify the scale of motion for this segment. As noted earlier [188], this segment hosts the recognition sequence for phosphorylation which may be facilitated by local structural flexibility.

## 9.6 Conclusions

This study provides direct spectroscopic information on structural details of AFA- and WT-PLN in complex with SERCA1a in lipid bilayers. For AFA-PLN in complex with SERCA, a short N-terminal (V4–R13) and a long C-terminal (Q29–V49) helix intersected by a linker of reduced structural order from R14 to Q22 were identified. The C-terminal helix ends at V49. Spectra corroborate an earlier site-specific observation that the C-terminal helix comprises residues A24–Q26 [123], but additionally show that P21 and Q22 do not constitute another helix turn. 2D correlation spectra recorded on WT-PLN in complex with SERCA1a exhibit no significant difference with respect to AFA-PLN. Mobility-filtered (<sup>1</sup>H, <sup>13</sup>C) correlation experiments speak against the existence of a significantly populated PLN conformation with dynamics on the micro- to nanosecond timescale after formation of the complex. Combination of our spectroscopic results with biophysical and biochemical data using flexible protein–protein docking simulations leads to a structural model for AFA-PLN in complex with SERCA1a. While sharing similarities to the model by Toyoshima and co-workers, this model demonstrates that the extended length of the C-terminal  $\alpha$ -helix, obtained from experimental data, does not violate N-terminal binding near K397'–V402', and takes into account recent cross-linking results. Furthermore, structural differences seen in the transmembrane helix and the potential influence of protein flexibility around the phosphorylation sequence may provide an important starting point for a more detailed analysis of protein–protein recognition and the regulatory effect of protein kinase A.

## Chapter 10

### Summary and Outlook

The work presented here applied, evaluated and improved methods in NMR spectroscopy for the structural characterization of (membrane) proteins and other biomolecules in the solid phase. All experiments were performed under magic angle spinning. The studies employed uniform sample labeling schemes, which are demanding in terms of spectroscopy, but yield a large amount of experimental data for just one or a few samples.

**Chapter 5 – Shift Predictions.** A statistical analysis of nuclear chemical shift predictions of proteins in the solid state by rapid algorithms trained on and verified with solution-state NMR assignments was presented. The precision of predictions by three common computer programs (SHIFTS, SHIFTX and SPARTA) was found to be close to chemical shift predictions of proteins in solution. Correlation coefficients depend on the kind of resonance (N, C', C $\alpha$  and C $\beta$ ) and on secondary structure ( $\beta$ -strand, random coil and  $\alpha$ -helix), and additionally on the molecular environment (membrane-integral or not). It is suggested to manually correct Ile CA predictions in membrane-integral,  $\alpha$ -helical residues for all three programs, in particular for SHIFTX. The use of SPARTA or SHIFTX is encouraged due to their higher precision compared to SHIFTS.

The findings validated the use of rapid, semi-empirical chemical shift prediction algorithms in context of structural fitting [48, 112, 194], and provide a quantitative basis for their use in comparative modeling, two applications that may become more influential in structural investigations of membrane proteins in the light of increasing computational capacities and a larger data basis of known structures.

**Chapter 6 – TEE.** On a uniformly  $^{13}\text{C}$ -labeled sample of L-tyrosine-ethylester, it was demonstrated that the detection of through-space  $^1\text{H}$ - $^1\text{H}$  and  $^{13}\text{C}$ - $^{13}\text{C}$  interactions and

one-bond dipolar couplings can provide an efficient means to elucidate molecular structure and dynamic disorder at atomic resolution. Compared to an earlier model study on the tripeptide AGG [160], where results of one 2D correlation experiment were used in a simplified, dual constraint classification, the CHHC buildup characteristics were here further used to refine the structural modeling. In particular, the combined measurement of internuclear distances and one-bond dipolar couplings permits a clear distinction of structural or dynamic disorder in the solid state.

Such information is, for example, relevant in the context of characterizing structure and flexibility in ligand–receptor interactions in a non-crystalline environment.

**Chapter 7 – Sequential ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) Correlations.** For the study of larger systems, a new method was introduced that allows to obtain sequential resonance assignments from ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectroscopy on a uniformly labeled protein under magic angle spinning, delivering information complementary to established ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) correlation approaches. The experiment relies on weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) coupling conditions using a defined range of MAS rates, and can be employed at arbitrary magnetic field strength, as demonstrated on the tri-peptide AGG and the proteins ubiquitin and kalitoxin. The method was essential for resonance assignments discussed in Chapters 8 and 9, and found application in other studies, including membrane a seven-transmembrane helix receptor [81], and protein fibrils [111].

This experiment is a useful tool to study even larger membrane proteins or biomolecular assemblies, as it is robust, and can be implemented easily in higher-dimensional correlation sequences. On the other hand, the determination of resonance correlations under weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) coupling conditions may also be useful for the investigation of small molecules in the solid phase, without resorting to triple-resonance experiments, or for biomolecules which suffer from a lack of  $^{15}\text{N}$  resonances.

**Chapter 8 – Ubiquitin.** On example of a single sample of uniformly ( $^{13}\text{C}$ ,  $^{15}\text{N}$ )-labeled sample of ubiquitin (76 amino acid residues), solid-state NMR methods tailored to the construction of 3D molecular structure were discussed, and the influence of solid-phase protein preparation on solid-state NMR spectra was studied. The investigation was facilitated by a near-complete resonance assignment for a nano-crystalline sample preparation using precipitation from polyethylene-glycol buffer. This presented one of the first ssNMR resonance assignments for a uniformly labeled proteins of this size. A comparative analysis

of C', C<sup>α</sup>, and C<sup>β</sup> resonance frequencies suggested that <sup>13</sup>C chemical shift variations are most likely to occur in protein regions that exhibit an enhanced degree of molecular mobility.

It is suggested to improve the extraction of <sup>1</sup>H–<sup>1</sup>H distances for non-resolved <sup>13</sup>C resonances by employing iterative algorithms that allow for the incorporation of ambiguous distance information, such as ARIA [178]. Analytical experience gained from the application to ubiquitin was of direct use for the structure determination of KTX [161] and AFA-PLN [9]. The study provides a reference for future studies of molecular structure and dynamics in (membrane) proteins of unknown structure by high-resolution solid-state NMR spectroscopy.

**Chapter 9 – Ca<sup>2+</sup>-ATPase-Bound Phospholamban.** AFA- and WT-PLN reconstituted with SERCA1a in a functional lipid environment were characterized. Chemical shift assignments were obtained in all domains of AFA-PLN, relying mainly on sequential (<sup>13</sup>C, <sup>13</sup>C) correlation spectroscopy as discussed in Chapter 7. For AFA-PLN in complex with SERCA1a, this study provides *direct* spectroscopic evidence of a short N-terminal (V4–R13) and a long C-terminal (Q29–V49) helix intersected by a linker of reduced structural order ranging from R14 to Q22. Spectra support an earlier site-specific observation that the C-terminal helix comprises residues A24–Q26 [123], but additionally show that P21 and Q22 do not constitute another helix turn. Experiments on WT-PLN show no significant difference in binding compared to AFA-PLN, and do not support the coexistence of a significantly populated dynamic state of PLN on the μs to ns timescale after formation of the PLN/SERCA complex. Combining secondary structure information directly obtained for AFA-PLN bound to SERCA1a with other biophysical and biochemical data in a flexible protein–protein docking simulation with HADDOCK, a model of the complex was calculated. This model clearly demonstrated that the secondary structure of AFA-PLN is compatible with binding of the N-terminal domain near K397'–V402' of SERCA1a, despite the larger α-helix extent compared to the model by Toyoshima and co-workers, and takes into account recent cross-linking results.

Structural differences seen in the transmembrane helix and the potential influence of protein flexibility around the phosphorylation sequence may provide a starting point for a more detailed analysis of protein–protein recognition and the regulatory effect of protein kinase A. In this context, and in order to probe the exact location of interaction with the N-terminus of AFA-PLN that is suggested from cross-linking and mutagenesis data,

it would be desirable to also study spin-labeled SERCA1a. In addition, the structural model may serve as a template for molecular dynamics simulations of the complex and its interaction with protein kinase A.

# Appendix





## Appendix A

# Concepts of NMR Spectroscopy (Supplement)

### A.1 Bloch Equations

The Bloch equations [33] suffice to describe the interaction of bulk magnetization with a static magnetic field or a radio frequency (r.f.) pulse. In compact form, the Bloch equations read as

$$\frac{d\mathbf{m}(t)}{dt} = \gamma\mathbf{m}(t) \times \mathbf{B}(t) - R(\mathbf{m}(t) - \mathbf{m}_0), \quad (\text{A.1})$$

where  $\mathbf{m} = \langle\langle \mathbf{I} \rangle\rangle$ ,  $\mathbf{m}_0$  describes equilibrium magnetization,  $\mathbf{B}$  is a magnetic field and  $R$  the relaxation matrix

$$R = \begin{pmatrix} 1/T_2 & 0 & 0 \\ 0 & 1/T_2 & 0 \\ 0 & 0 & 1/T_1 \end{pmatrix} \quad (\text{A.2})$$

with coefficients for transversal ( $T_2$ ) and longitudinal ( $T_1$ ) relaxation. For  $\mathbf{B} = B_0\hat{z}$  and  $\omega_0 = \gamma B_0$ , the solution of (A.1) is given by

$$m_x(t) = m_x(0) \cos(\omega_0 t) e^{-t/T_2} \quad (\text{A.3})$$

$$m_y(t) = m_x(0) \sin(\omega_0 t) e^{-t/T_2} \quad (\text{A.4})$$

$$m_z(t) = m_0 + (m_z(0) - m_0) e^{-t/T_1} \quad (\text{A.5})$$

Transverse magnetization, associated to single quantum coherences in the density matrix, induces a current in the detection coil. The components are described in the complex plane,  $\mathbf{m}^\pm = m_x \pm im_y$ , and can be recorded by a single coil using quadrature detection [78].

### A.2 Theoretical Background

#### A.2.1 Average Hamiltonian Theory

The formal solution of the Liouville–von Neumann equation (here with  $\hbar = 1$ )

$$\frac{d}{dt}\rho(t) = -i[H(t), \rho(t)] \quad (\text{A.6})$$

may be written

$$\rho(t) = U(t)\rho(0)U^{-1}(t) \quad \text{with} \quad U(t) = T \exp(-i \int_0^t H(t') dt'), \quad (\text{A.7})$$

where the Dyson time-ordering operator  $T$  defines the evaluation of the exponential function in cases where the Hamiltonians at different times do not commute [78],

$$T = \left\{ \begin{array}{ll} H(t_1)H(t_2) & t_1 > t_2 \\ H(t_2)H(t_1) & t_1 < t_2 \end{array} \right\}. \quad (\text{A.8})$$

The idea of *Average Hamiltonian Theory* [105, 315] is to substitute the propagator

$$U(t) = T \exp \int_0^t H(t') dt', \quad (\text{A.9})$$

by a propagator  $U(t) = \exp \overline{H(t)} t$  with an average Hamiltonian  $\overline{H}$ , which has an identical effect on the time-evolution of the density matrix. Using the Magnus expansion, which is a continuous expression for the Baker-Campbell-Hausdorff equation

$$e^A \cdot e^B = e^{B+A} + \frac{1}{2}[A, B] + \frac{1}{12}(B[B, A] + [B, A]A) + \dots, \quad (\text{A.10})$$

the average Hamiltonian is given by a series

$$\overline{H(t)} = \overline{H^{(0)}}(t) + \overline{H^{(1)}}(t) + \overline{H^{(2)}}(t) + \dots \quad (\text{A.11})$$

with the first three terms given by

$$\overline{H^{(0)}}(t) = \frac{1}{t_c} \int_0^{t_c} H(t_1) dt_1 \quad (\text{A.12})$$

$$\overline{H^{(1)}}(t) = -\frac{1}{2t_c} i \int_0^{t_c} dt_2 \int_0^{t_2} dt_1 [H(t_1), H(t_2)] \quad (\text{A.13})$$

$$\begin{aligned} \overline{H^{(2)}}(t) = \frac{1}{6t_c} \frac{1}{\hbar^2} \int_0^{t_c} dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \{ [H(t_1), [H(t_2), H(t_3)]] \\ + [[H(t_1), H(t_2)], H(t_3)] \}. \end{aligned} \quad (\text{A.14})$$

## A.2.2 Initial Density Operator

At thermal equilibrium, the initial density operator  $\rho_0 \equiv \rho(0)$  at temperature  $T$  is given by a Boltzmann distribution

$$\rho_0 = \frac{e^{-H/kT}}{\text{tr}(e^{-H/kT})} \quad (\text{A.15})$$

with the Boltzmann constant  $k = k_B = 1.38 \times 10^{-23}$  J/K. At high temperature,  $|H|/kT \ll 1$ , for example  $|H|/kT \approx 2 \times 10^{-5}$  for  $^1\text{H}$  in a 14.1 T magnetic field at 0 °C. Hence, the exponential function can be approximated to first order by  $e^x = 1 + x$ ,

$$\rho_0 = \frac{e^{-H/kT}}{\text{tr} e^{-H/kT}} \approx \frac{1 - H/kT}{\text{tr}(1 - H/kT)} = 1 - \frac{H}{kT \cdot \text{tr} \mathbf{1}}. \quad (\text{A.16})$$

The largest term in (A.16) is the unity operator. It does not lead to an observable change of the state of the system and can therefore be neglected. For NMR experiments, the initial density operator may hence be given by  $\rho_0 \propto \frac{H}{kT}$  or  $\rho_0 \propto H$ .

### A.2.3 Coherences

The state  $|\psi\rangle$  may be expanded in terms of a complete orthonormal set  $\{|i\rangle\}_n$ ,

$$|\psi(t)\rangle = \sum_{i=1}^n c_i(t)|i\rangle. \quad (\text{A.17})$$

The density operator then reads as

$$\rho(t) = \sum_k w_k \sum_{ij} c_i^k(t) c_j^{*k}(t) |i\rangle\langle j| = \sum_{ij} \overline{c_i(t) c_j^*(t)} |i\rangle\langle j| \quad (\text{A.18})$$

where the bar denotes the ensemble average. In the set of eigenstates of the Hamiltonian  $H$  of the system described by  $\rho$ , the diagonal element

$$\rho_{rr} = \langle r|\rho(t)|r\rangle = \overline{|c_r(t)|^2} = P_r \quad (\text{A.19})$$

is equal to the probability of finding the spin system in the eigenstate  $|r\rangle$ .  $P_r$  is the population of the state  $|r\rangle$ . The off-diagonal element

$$\rho_{rs} = \langle r|\rho(t)|s\rangle = \overline{c_r(t) c_s^*(t)} \quad (\text{A.20})$$

indicates a *coherent superposition* of eigenstates  $c_r(t)|r\rangle + c_s(t)|s\rangle$  in the sense that the phase of the various members of the ensemble are correlated with respect to  $|r\rangle$  and  $|s\rangle$ . Such a coherent superposition is called *coherence* [78]. A matrix element represents  $p$ -quantum coherence, where  $p$  is the difference in magnetic quantum numbers (Section 2.3.1)  $\Delta m_{rs} = m_r - m_s$ . In NMR, single quantum coherence ( $p = \pm 1$ ) is related to transverse magnetization, and can be observed directly<sup>1</sup>. All other coherence terms can only be observed indirectly.

### A.2.4 Wigner Elements

The characterization of the rotation by Euler angles involved two rotations about  $\hat{\mathbf{z}}$ -axes, and can hence conveniently expressed using the eigenstates of  $\mathbf{I}_z$  as

$$D_{m'm}^{(j)}(\alpha, \beta, \gamma) = \langle jm'|e^{-i\alpha J_z'} e^{-i\beta J_y} e^{-i\gamma J_z}|jm\rangle \quad (\text{A.21})$$

$$= e^{-im'\alpha} \langle jm'|e^{-i\beta J_y}|jm\rangle e^{-im\gamma} \quad (\text{A.22})$$

$$= e^{-im'\alpha} d_{m'm}^{(j)}(\beta) e^{-im\gamma} \quad (\text{A.23})$$

where only the rotation about the  $y'$  axis remains to be calculated, and  $d_{m'm}^{(j)}(\beta) \equiv D_{m'm}^{(j)}(\beta, \hat{\mathbf{y}}')$  defines the *reduced Wigner elements*. An analytic expression for the reduced Wigner elements is given by Wigner's formula [245],

$$d_{m'm}^{(j)}(\beta) = \sum_k (-1)^{k-m+m'} \frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{(j+m-k)!(j-k-m')!(k-m+m')!} \quad (\text{A.24})$$

$$\times \cos\left(\frac{\beta}{2}\right)^{2j-2k+m-m'} \sin\left(\frac{\beta}{2}\right)^{2k-m+m'}. \quad (\text{A.25})$$

<sup>1</sup>Using *quadrature detection*, only one component can be detected. By convention, this is the coherence term with  $p = -1$ .

### A.2.5 Dipolar Alphabet

According to the correspondence principle, the spin-spin dipolar interaction Hamiltonian may be defined by

$$H_D = \frac{\mu_0}{4\pi} \hbar \frac{\gamma_1 \gamma_2}{r_{12}^3} \cdot (\mathbf{I}_1 \cdot \mathbf{I}_2 - 3(\mathbf{I}_1 \cdot \hat{\mathbf{r}}_{12})(\mathbf{I}_2 \cdot \hat{\mathbf{r}}_{12})). \quad (\text{A.26})$$

where  $\mathbf{I}_1$  and  $\mathbf{I}_2$  are two nuclear spins, and  $\hat{\mathbf{r}}$  the unit vector for the internuclear distance. Comparison to  $H_D = \mathbf{I}_1 D \mathbf{I}_2$  (Chapter 2) leads to an expression of the matrix representing the dipolar interaction,

$$D = d_{12} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad (\text{A.27})$$

where  $d_{12}$  is the dipolar coupling constant, defined as

$$d_{12} = \frac{\mu_0}{4\pi} \hbar \frac{\gamma_1 \gamma_2}{r_{12}^3}, \quad [d_{12}] = \text{rad} \cdot \text{Hz}. \quad (\text{A.28})$$

The dipolar tensor is traceless and axially symmetric, the only non-vanishing spatial tensor being  $A_{20} = -d_{12}\sqrt{6}$ . Together with the spin component  $T_{20} = -\frac{1}{\sqrt{6}}(\mathbf{I}_1 \mathbf{I}_2 - 3I_{1z} I_{2z})$ , the Hamiltonian for the dipolar interaction in the principle axis system of  $D$  is of the form

$$H_D = A_{20}^{\text{PAS}} T_{20} = d_{12} \cdot (\mathbf{I}_1 \mathbf{I}_2 - 3I_{1z} I_{2z}). \quad (\text{A.29})$$

Expanding the products, using spherical coordinates, raising- and lowering operators, the dipolar interaction Hamiltonian can be decomposed into the *dipolar alphabet*

$$\begin{aligned} H_D &= \frac{\mu_0}{4\pi} \hbar \frac{\gamma_1 \gamma_2}{|\mathbf{r}|^3} \cdot (A + B + C + D + E + F) \\ &\equiv d_{12} \cdot (A + B + C + D + E + F) \end{aligned} \quad (\text{A.30})$$

where

$$A = -I_{1z} I_{2z} \cdot (3 \cos^2 \theta - 1) \quad (\text{A.31})$$

$$B = (I_1^+ I_2^- + I_1^- I_2^+) \cdot \frac{3 \cos^2 \theta - 1}{4} \quad (\text{A.32})$$

$$C = -(I_1^+ I_{2z} + I_{1z} I_2^+) \cdot e^{-i\varphi} \cdot \frac{3 \sin \theta \cos \theta}{2} \quad (\text{A.33})$$

$$D = -(I_1^- I_{2z} + I_{1z} I_2^-) \cdot e^{i\varphi} \cdot \frac{3 \sin \theta \cos \theta}{2} \quad (\text{A.34})$$

$$E = -I_1^+ I_2^+ \cdot e^{-2i\varphi} \cdot \frac{3 \sin^2 \theta}{4} \quad (\text{A.35})$$

$$F = -I_1^- I_2^- \cdot e^{2i\varphi} \cdot \frac{3 \sin^2 \theta}{4}, \quad (\text{A.36})$$

and  $d_{12}$  is the dipolar coupling constant. Note that  $D = C^*$  and  $I_1^+ I_2^- + I_1^- I_2^+ = -2(I_{1z} I_{2z} - \mathbf{I}_1 \mathbf{I}_2)$ .

## Appendix B

### Molecular Modeling (Supplement)

#### B.1 Atomic Root-Mean-Square Deviations

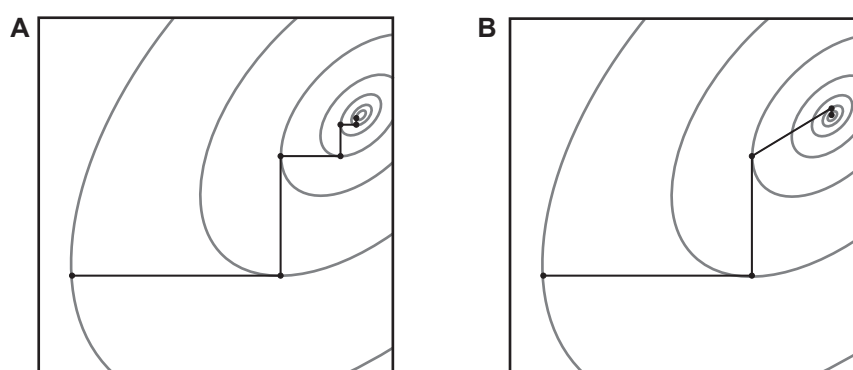
The standard measure used to quantify differences between 3D structures is the root-mean-square deviation (RMSD) for a given set of corresponding atoms [197]. For two sets of atoms  $\{\mathbf{r}_{1i}, \mathbf{r}_{2j}\}_{i,j=1,\dots,n}$  with  $\sum_i \mathbf{r}_{1i} = \sum_j \mathbf{r}_{2j} = 0$ , the RMSD is defined as the root mean square distance between the positions of corresponding atoms after optimal superposition of the two structures,

$$\text{RMSD} = \min \left\langle \frac{1}{n} \sum_{i=1}^n |\mathbf{r}_{1i} - R\mathbf{r}_{2i}|^2 \right\rangle^{1/2}, \quad (\text{B.1})$$

where  $R$  denotes a rotation matrix, and the minimum over all possible rotation matrices is taken [103]. The superposition (B.1) can be computed by molecular visualization software such as MOLMOL [152], or by using CNS scripts. The RMSD values may be calculated only for selected atoms, such as N, CO and CA, or for all heavy (i.e., non-hydrogen) atoms, excluding conformations that are not well-defined, such as chain termini and some loops.

##### B.1.1 Conjugate Gradient Minimization

In contrast to other methods, which miss the best path to the minimum, resulting in slow convergence, Powell's conjugate gradient algorithm produces steps that continually refine the direction towards the minimum. The algorithm is explained in detail e.g. in ref. [193]. The resulting minimization process is illustrated in Figure B.1B, in comparison to a less efficient algorithm.

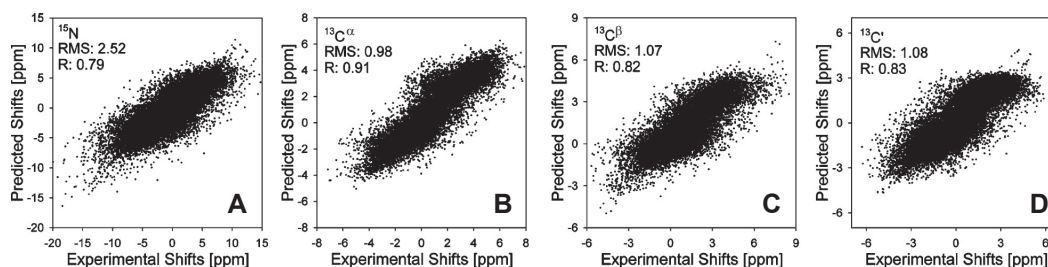


**Figure B.1:** (A) Minimization of a 2D potential function along standard base vectors using the taxi-cab method. In contrast, (B) Powell conjugate gradient minimization converges much better, here as of the third iteration. Figures adapted from [193].

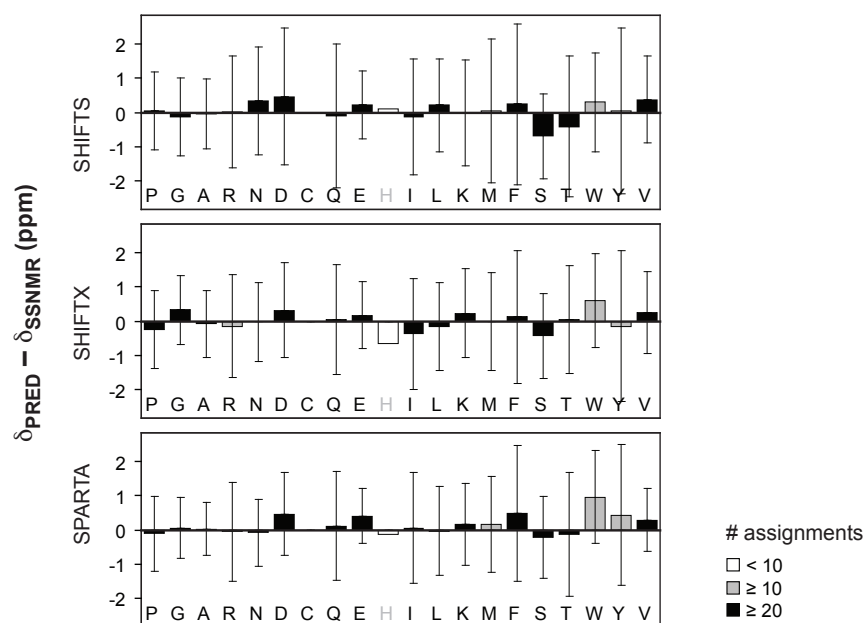
## Appendix C

### On SSNMR Chemical Shift Predictions (Supplement)

The plot for SPARTA from comparing predictions to solution-state NMR, as in Figure 5.2 for solid-state NMR shifts, is reproduced in Figure C.1 from the original publication [259], for reference. The results are, as described in the main text, just slightly better in terms of RMSD and  $R$  coefficient.



**Figure C.1:** SPARTA scatter plots comparing experimental and SPARTA-predicted secondary chemical shifts for backbone  $^{15}\text{N}$ ,  $\text{C}^{\alpha}$ ,  $\text{C}^{\beta}$ , and  $\text{C}'$  nuclei, from the original publication [259].



**Figure C.2:** Residue-specific differences between predicted and observed  $C^\alpha$  chemical shifts as shown in Figure 5.5 for SHIFTX, for all three considered programs. Averaging is done over all secondary structure elements. Results are representative also for other nuclei.



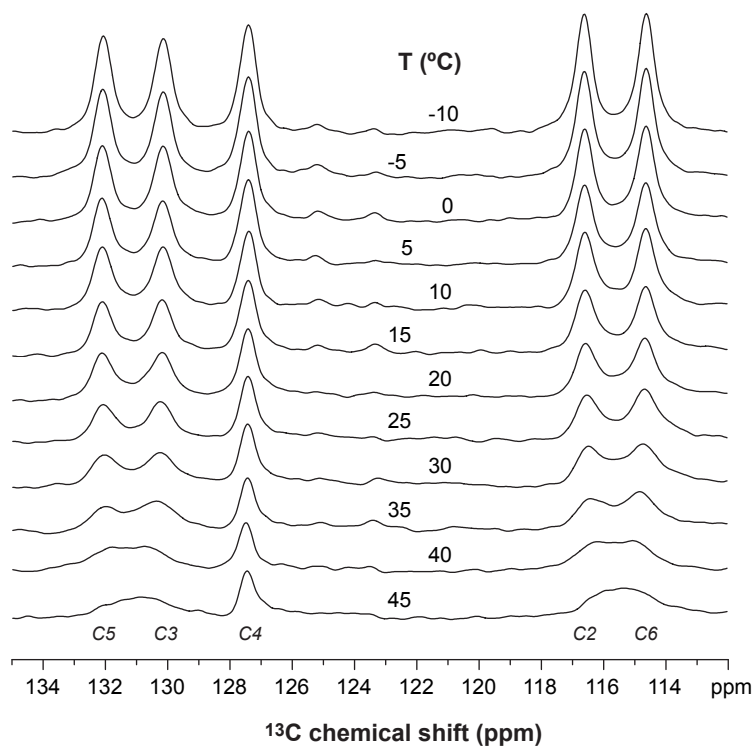
## Appendix D

### L-Tyrosine-Ethylester (Supplement)

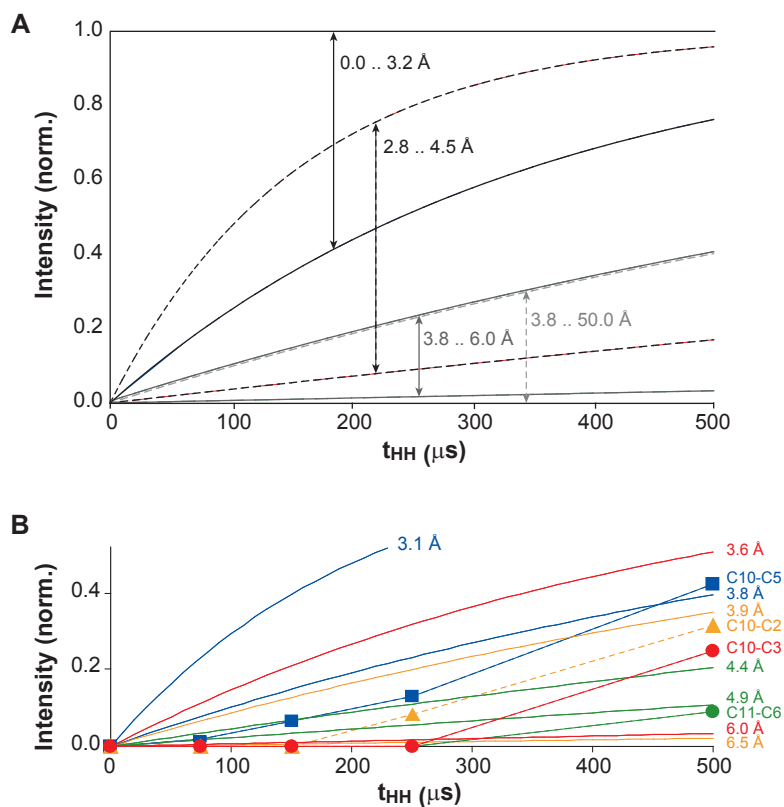
Distance constraints from measurements on U( $^{13}\text{C}$ )-TEE<sub>dil</sub> as used in the CNS structure calculation are given in Table D.1. H $x$  denotes a proton bound to carbon C $x$  ( $x = 2..11$ ), following the nomenclature given in Figure 6.1. Four classes of H–H distance restraints were set up from CHHC experiments by comparing the cross peak build-up data to theoretical curves, spin pairs with fixed distances known from the covalent structure of TEE, and results obtained in other model systems (Histidine, Ala-Gly-Gly) [160]. One C–C restraint was defined based on the R<sup>2</sup>TR measurement. Ambiguities from restraints involving methyl or methylene protons are denoted by a suffix (#), and were accounted for by the 'SUM' averaging method in CNS, which calculates an average distance from positions of all relevant protons. Restraints between protons with fixed distances, given for reference, were not used in the structure calculation.

Class	Value	Atom 1	Atom 2	Status
<b>Class 1</b>	$d_{\text{start}} = 2.5 \text{ \AA}$	H11#	H10#	used
	$d_{\text{min}} = 0.0 \text{ \AA}$	H7#	H5	used
	$d_{\text{max}} = 3.2 \text{ \AA}$	H7#	H8	used
		H2	H3	not used, fix, reference
		H8	H3	used
		H6	H5	used
		H7#	H3	used
<b>Class 2</b>	$d_{\text{start}} = 3.5 \text{ \AA}$	H10#	H5	used
	$d_{\text{min}} = 2.8 \text{ \AA}$	H3	H5	not used, fix, reference
	$d_{\text{max}} = 4.5 \text{ \AA}$	H6	H2	not used, fix, reference
		H8	H5	used
		H7#	H10#	used
		H8	H10#	used
		H10#	H6	used
		H7#	H6	used
		H8	H2	not used, build-up indicates relay transfer
	H7#	H2	not used, build-up indicates relay transfer	
<b>Class 3</b>	$d_{\text{start}} = 4.5 \text{ \AA}$	H11#	H8	used
	$d_{\text{min}} = 3.8 \text{ \AA}$	H11#	H6	used
	$d_{\text{max}} = 6.0 \text{ \AA}$	H10#	H3	used
		H11#	H7#	used
		H8	H6	used
		H10#	H2	used
		H11#	H3	used
<b>Class 4</b>	$d_{\text{start}} = 5.0 \text{ \AA}$	H11#	H5	used
	$d_{\text{min}} = 3.8 \text{ \AA}$	H11#	H2	used
	$d_{\text{max}} = 50.0 \text{ \AA}$	H6	H3	not used, fix, reference
		H2	H5	not used, fix, reference
<b>Class R<sup>2</sup>TR</b>	$d_{\text{start}} = 4.9 \text{ \AA}$	C11	C2	used
	$d_{\text{min}} = 3.0 \text{ \AA}$			
	$d_{\text{max}} = 5.2 \text{ \AA}$			

**Table D.1:** Distance restraints for TEE supplied to the CNS structure calculation.



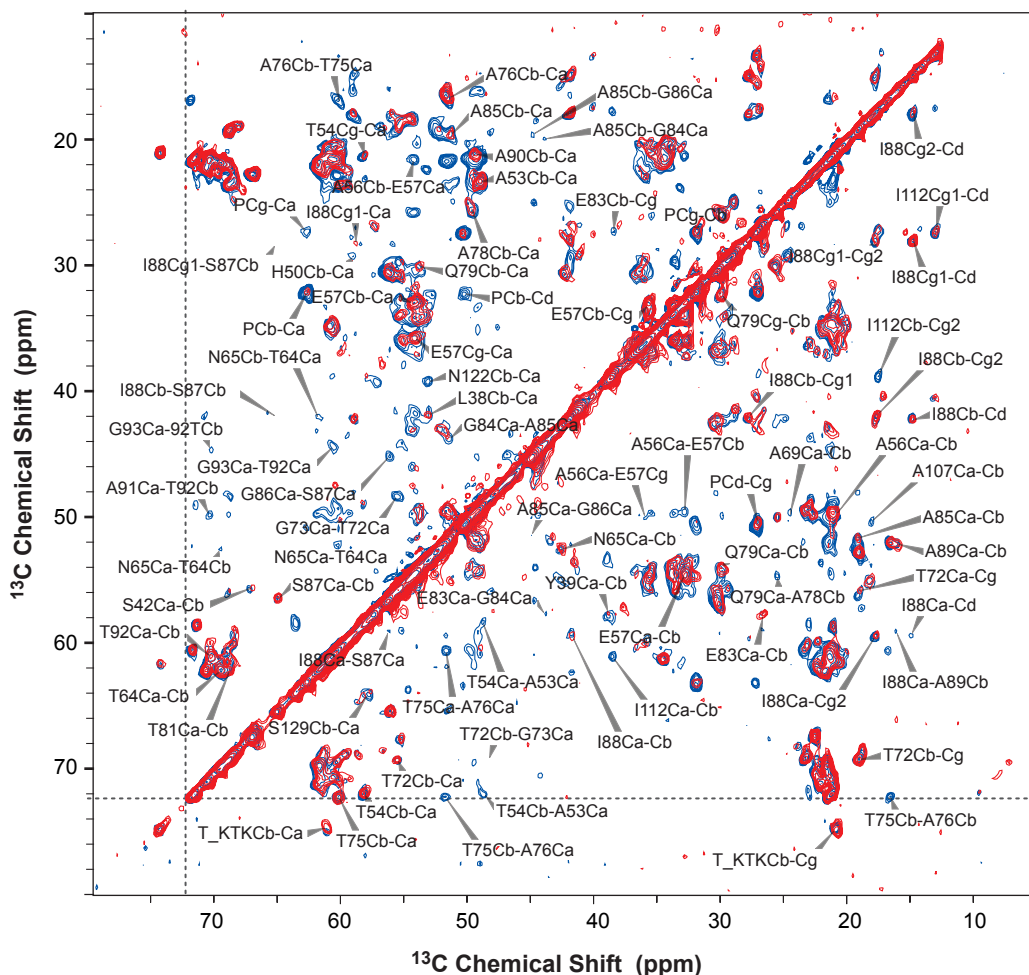
**Figure D.1:**  $^{13}\text{C}$  NMR spectra from  $^1\text{H}$ - $^{13}\text{C}$  cross-polarization of TEE, recorded with 9.5 kHz MAS at variable temperature, ranging from  $-10$  to  $+45^\circ\text{C}$ , reproducing results from [114]. At high temperatures, C2/C6 and C3/C5 rapidly interchange their chemical environment, giving rise to one common resonance line for each pair at 115 and 131 ppm average  $^{13}\text{C}$  chemical shift. With decreasing temperature, the ring flip occurs less often until at  $-5$  to  $-10^\circ\text{C}$ , the resonance lines are clearly separated. By contrast, the C4 (127 ppm) resonance line width is approximately constant over the given temperature range, indicating that C4 is not affected by the ring's  $\pi$ -flip.



**Figure D.2:** (A) Theoretical CHHC cross peak build-up curves assuming the theoretical model of ref. [160] and distance classifications of Table D.1. (B) Comparison of a selected set of experimental CHHC cross peak intensity build-ups to results from theory. The given distances reflect minimum and maximum proton–proton distances found in the structural ensemble shown in Figure 6.9. Transfer over distances larger than about 3.5 Å can be complicated by multi-spin effects [160], but in all cases, experimental data are found within or close to the dynamic range predicted by theory.

## **Appendix E**

Supplement on Protein Solid-State NMR Resonance  
Assignments from  $^{13}\text{C}$ – $^{13}\text{C}$  Correlation Spectroscopy



**Figure E.1:** 2D ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation experiment conducted on fibrils of U- $^{13}\text{C}$ ,  $^{15}\text{N}$  \ [K,V]-labeled  $\alpha$ -Synuclein in the A-form [111]. (red) CC SD spectrum with short, 20 ms mixing time, correlating intra-residue resonances, exclusively. (blue) CC SDWC correlation with 150 ms longitudinal mixing under weak ( $\text{C}'$ ,  $\text{C}^\alpha$ ) coupling conditions. In conjunction with complementary NCACX and NCOCX experiments, a number of protein residues can be assigned unambiguously. Figure kindly provided by H. Heise.

## Appendix F

### Supplement on High-Resolution Solid-State NMR Studies on Uniformly ( $^{13}\text{C}$ , $^{15}\text{N}$ )-Labeled Ubiquitin

#### F.1 Chemical Shift Assignments

Residue	N	CO	CA	CB	CG(1)	CG2	CD(1)	CD2	CE	CZ
M1		168.6	53.0	32.1	29.6					
Q2	122.6		51.8							
I3	113.1	170.3	56.9	40.0	22.8	16.0	12.3			
F4	117.4	173.3	53.0	39.2	137.6					
V5	120.3	173.3	58.1	32.7	20.8	18.8				
K6	125.9	175.0	51.6		22.8					
T7	113.4	175.3	58.8	67.6		19.1				
L8	122.0	175.3	55.7	40.1	25.5		22.4			
T9	102.3	173.6	59.1	67.1		20.6				
T9b	103.5		59.0	66.9		19.9				
G10	108.1	171.3	43.1							
K11										
T12	120.1	171.9	60.5	67.5		20.8				
I13	128.7	173.3	57.9	38.9	24.7	16.0	13.1			
I13b	128.8	172.2	58.0							
T14	124.5	171.9	60.6	67.7		20.1				
L15	123.6	173.1	51.2	45.6	25.3		24.6	23.4		
E16		174.0	53.4							
V17	117.6	172.0	56.5	34.5	20.2	16.6				
E18	118.2	172.9	52.6	27.9						
P19	133.5	174.8	64.1	29.7	25.8		48.0			
S20	105.6	172.8	54.9	59.6						
S20b	105.2		55.5							
D21	120.9	174.2	54.7	38.5						
T22	106.4	174.6	57.2	68.7		19.9				
I23	120.9	177.2	59.8	31.5	25.3	15.4	7.1			
I23b		177.7	60.5	31.7	26.0	16.0	7.7			
E24	121.3	176.8	58.0	25.9						





Residue	N	CO	CA	CB	CG(1)	CG2	CD(1)	CD2	CE	CZ
L69		173.3	51.5	42.4	25.7		24.5	22.4		
V70	126.7		58.0	32.8	19.2	18.7				
L71										
R72										
L73										
R74										
G75										
G76										

**Table F.1:** Chemical shift assignments for UBI-P. Residues marked with a *b* relate to doubled resonances.



## Appendix G

# Supplement on the Structural Characterization of Ca<sup>2+</sup>-ATPase-bound Phospholamban in Lipid Bilayers

## G.1 Chemical Shift Assignments

Chemical shift assignments are summarized in Table G.1.

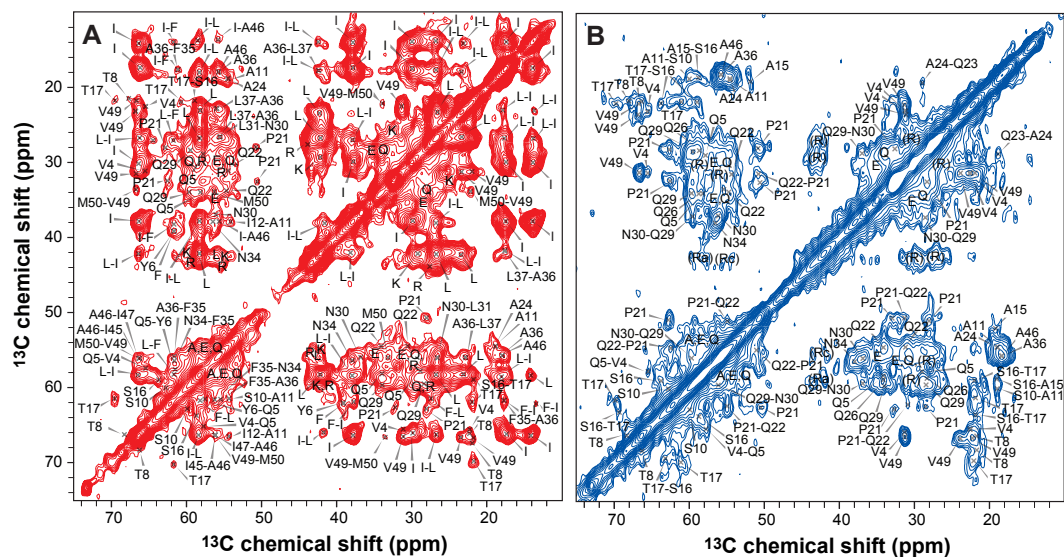
Residue	CO	CA	CB	CG(1)	CG2	CD(1)	CD2
G-1	170.9	43.5		-	-	-	-
S0				-	-	-	-
M1					-	-	-
E2					-		-
K3					-		-
V4	178.2	65.5	31.2	22.6	22.6	-	-
Q5		57.7	28.9		-		-
Y6		61.7	38.9				
L7	178.7	58.2	42.2	26.8	-	23.9	22.8
T8		66.3	68.3	-	21.7	-	-
R9		59.7	29.6	27.6	-	43.3	-
S10	176.6	60.0	62.8	-	-	-	-
A11		54.4	18.5	-	-	-	-
I12		66.3	38.0			14.3	-
R13	178.0	59.0	30.0	27.6	-	43.3	-
R14					-		-
A15		51.8	19.0	-	-	-	-
S16	174.9	58.9	63.7	-	-	-	-
T17	174.7	60.9	69.9	-	21.9	-	-
I18		60.9	38.2		17.8		-
E19					-		-
M20					-	-	-
P21	176.0	62.6	32.6	28.1	-	50.9	-
Q22		54.8	31.2	33.8	-		-
Q23					-		-

Residue	CO	CA	CB	CG(1)	CG2	CD(1)	CD2
A24		54.5	18.8	-	-	-	-
R25		59.0	30.0	27.6	-	43.3	-
Q26		58.6	28.3		-		-
K27		59.1	32.3		-		-
L28	178.7	58.2	42.2	26.8	-	23.9	22.8
Q29		59.3	28.5	34.5	-		-
N30	177.4	56.1	36.8		-	-	-
L31	178.7	58.2	42.2	26.8	-	23.9	22.8
F32	177.0	62.2	39.2				
I33	177.3	66.3	37.9	30.2	17.4	14.1	-
N34	177.4	56.4	37.9		-	-	-
F35	177.0	62.2	39.2				
A36	178.6	55.5	18.3	-	-	-	-
L37	178.7	58.2	42.2	26.8	-	23.9	22.8
I38	177.3	66.3	37.9	30.2	17.4	14.1	-
L39	178.7	58.2	42.2	26.8	-	23.9	22.8
I40	177.3	66.3	37.9	30.2	17.4	14.1	-
F41	177.0	62.2	39.2				
L42	178.7	58.2	42.2	26.8	-	23.9	22.8
L43	178.7	58.2	42.2	26.8	-	23.9	22.8
L44	178.7	58.2	42.2	26.8	-	23.9	22.8
I45	177.3	66.3	37.9	30.2	17.4	14.1	-
A46	178.6	55.5	18.3	-	-	-	-
I47	177.3	66.3	37.9	30.2	17.4	14.1	-
I48	177.3	66.3	37.9	30.2	17.4	14.1	-
V49	178.2	66.6	31.2	23.4	22.2	-	-
M50		56.1	33.6		-	-	-
L51					-		
L52					-		

**Table G.1:** Chemical shift assignments for AFA-PLN (with G<sub>-1</sub>, S<sub>0</sub>, A<sub>36</sub>, F<sub>41</sub>, A<sub>46</sub>) in complex with SERCA1a. The absence of an atom in a given residue is denoted with a dash. WT-PLN cysteine residues 36, 41 and 46 exhibit chemical shifts of C<sup>α</sup> = 65.1 ppm and C<sup>β</sup> = 27.3 ppm (Figure 9.5).

Key steps in the assignment process were: Unambiguous sequential assignments from the (<sup>13</sup>C,<sup>13</sup>C) spectrum of 150 ms SDWC mixing at 235 K (Figure 9.3C and Figure G.2B) of U(<sup>13</sup>C,<sup>15</sup>N \ ILRKMYF)-AFA-PLN/SERCA1a, involving V4C<sup>α</sup>-Q5C<sup>α</sup>, S10C<sup>α</sup>-A11C<sup>β</sup>, S16C<sup>α</sup>/C<sup>β</sup>-T17C<sup>γ2</sup>, P21C<sup>α</sup>-Q22C<sup>α</sup>, Q29C<sup>α</sup>-N30C<sup>β</sup>. Assignments for T8 and N34 follow readily, where the identification of Thr resonances was assisted by 1D spectroscopy of U-(<sup>13</sup>C-Thr,<sup>15</sup>N-Ile)-AFA-PLN/SERCA1a, see Figure G.4. Unambiguous sequential assignments from the (<sup>13</sup>C,<sup>13</sup>C) spectra of 150 ms SDWC mixing at 266 K (Figure 9.3B, Figure G.2A) or 243 K (Figure G.2C) of U(<sup>13</sup>C,<sup>15</sup>N)-AFA-PLN/SERCA1a, involving A11C<sup>α</sup>-I12C<sup>β</sup>/C<sup>δ1</sup>, V49C<sup>α</sup>-M50C<sup>β</sup>, and identification of overlapping resonances from segment-specific residue groups, as described in the main text (e.g., involving

Leu–Ile or Ile–Phe pairs). Unambiguous sequential assignment of R13C' (helical character;  $C^\alpha$  and  $C^\beta$  chemical shifts are not unambiguously assigned, and were taken from average shifts for Arg in a helix [312]), see also Figure G.3. Sequential assignments were facilitated by the identification of intra-residue correlations, see, for example, Figure G.2A,B.



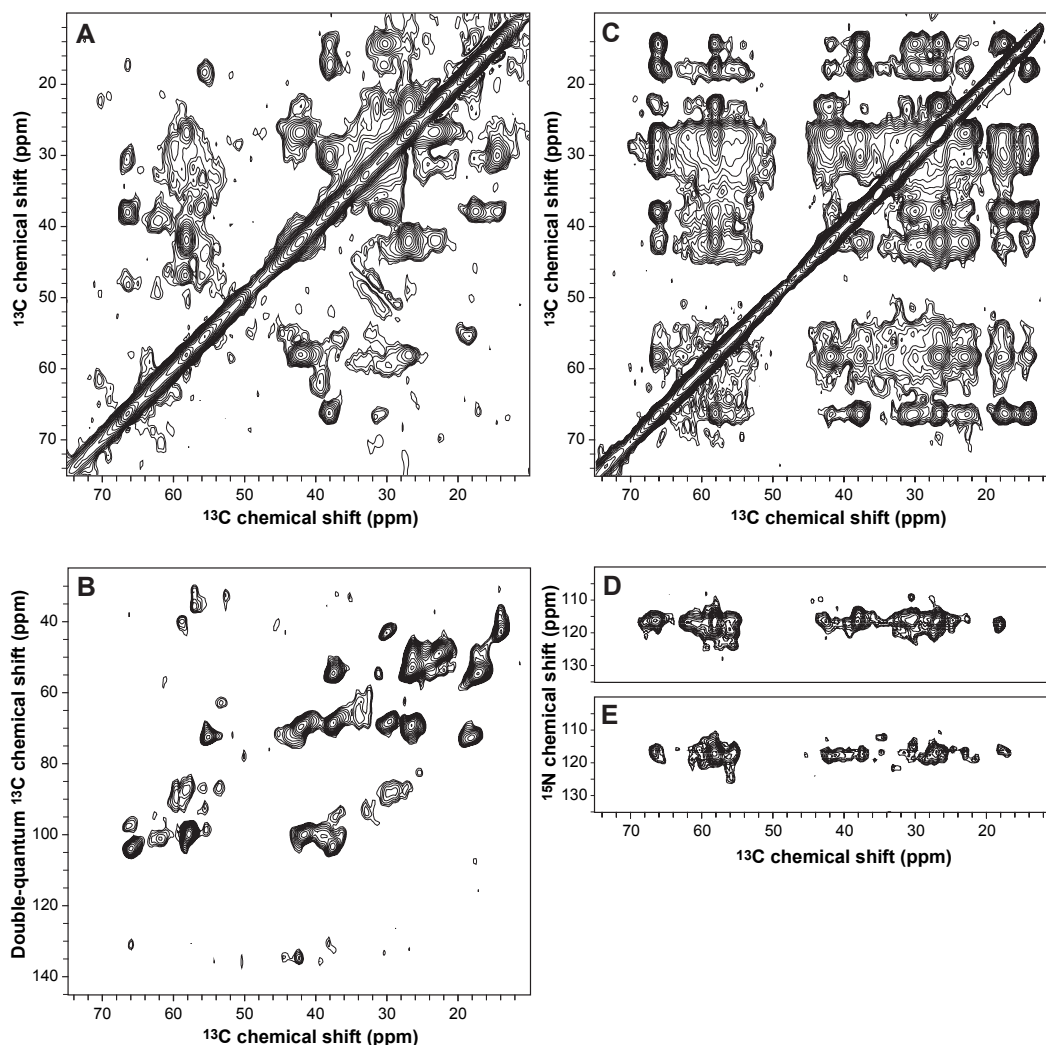
**Figure G.1:** Full spectra of aliphatic resonances from Figure 9.3, recorded at 18.1 T, MAS frequency of 12.5 kHz.

## G.2 Modeling Using Flexible Protein–Protein Docking Simulation

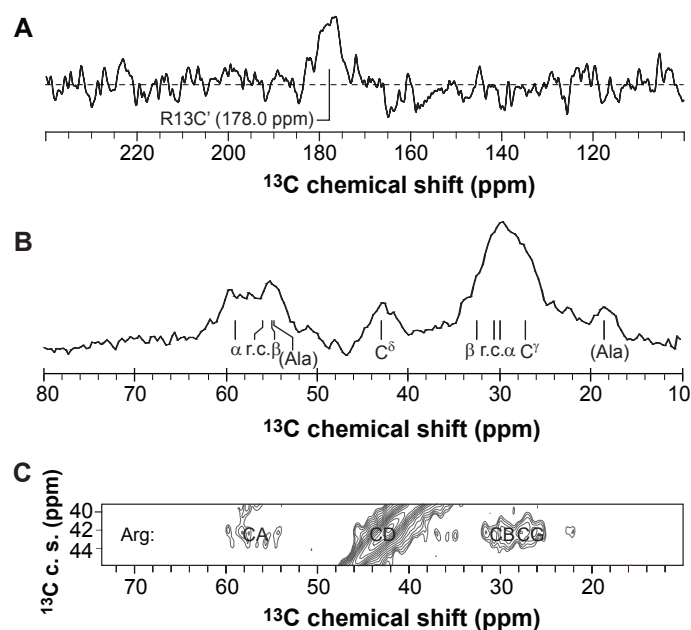
### G.2.1 Choice of Docking Constraints

Docking in HADDOCK is driven by unambiguous distance constraints derived from cross-linking experiments, and by Ambiguous Interaction Restraints (AIRs) reflecting mutagenesis screening of the complex interface.

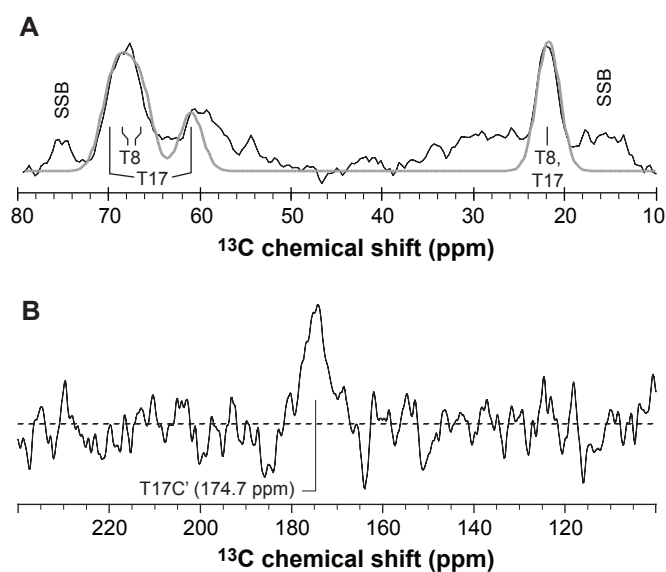
Unambiguous docking constraints are described in Table 9.1. HADDOCK uses ' $R^{-6}$  summation' over all selected atoms for ambiguous and unambiguous distance constraints, hence where no atom identifier is given, the distance constraint is fulfilled when any inter-atomic distance falls below the given maximum. Constraints derived from strong disulphide cross-linking were implemented using  $C^\beta$ – $C^\beta$  distances: Disulphide bridges generally allow for a  $C^\beta$ – $C^\beta$  distances up to 4.5 Å, and  $C^\alpha$ – $C^\alpha$  distances up to 6.5 Å [267]. Spatial proximity of residues derived from disulphide cross-linking of their respective Cys mutants is reflected by constraining the  $C^\beta$ – $C^\beta$  distance to a maximum of 5 Å, without an additional  $C^\alpha$ – $C^\alpha$  constraint, in order not to over-constrain the existing side chains. Artificial constraints were chosen to dock the cytoplasmic helix into the groove below K397'–V402', described earlier by Toyoshima and MacLennan [292], reflecting the cross-link constraint between K3 and K397' or K400' derived by James and co-workers [132], the relevance of a negative charge on either residue 398' or 399' [288], and the observation that E2 is an essential



**Figure G.2:** Supplementary spectra. (A) ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectrum of U-AFA-PLN/SERCA1a with 5 ms PDSM mixing recorded at 14.1 T magnetic field at a temperature of 243 K. The spectrum shows intra-residue correlations for through-space transfer over distances corresponding to one chemical  $^{13}\text{C}$ - $^{13}\text{C}$  bond. (B) 2Q/1Q ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectrum of U-AFA-PLN/SERCA1a, at 14.1 T magnetic field and a temperature of 243 K. The spectrum is showing intra-residue correlations for through-space transfer over distances corresponding to one chemical  $^{13}\text{C}$ - $^{13}\text{C}$  bond. (C) ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectrum of U-AFA-PLN/SERCA1a obtained with 150 ms SDWC mixing, at 14.1 T magnetic field and a temperature of 243 K. The spectrum shows essentially all possible intra-residue carbon correlations, and, additionally, correlations to sequential neighbors (mostly  $(i, i-1)$  and  $(i, i+1)$  correlations, where  $i$  denotes a residue number) [253]. (D,E) 2D ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) correlation spectra with SPECIFIC cross-polarization for (NCACX) intra-residue and (NCOCX)  $i \rightarrow i-1$  transfer.



**Figure G.3:** (A) 1D SPECIFIC NCO-edited  $^{13}\text{C}$  spectrum of U-[( $^{13}\text{C}$ , $^{15}\text{N}$ )-Arg]-AFA-PLN/SERCA1a, yielding the resonance of R13C', exclusively. (B) Double-quantum filtered  $^{13}\text{C}$  spectrum of U-[( $^{13}\text{C}$ , $^{15}\text{N}$ )-Arg]-AFA-PLN/SERCA1a. The presence of chemical shifts corresponding to helical and coil or extended backbone structure can readily be identified using statistical chemical shifts as published, for example by Wang and Jardetzky [312]. In addition, the spectrum yields signal at about 54 ppm and 18 ppm, which may result from amino acid residues converted from Arg during the bacterial expression process. (C) For comparison, residual intensity of Arg in reversely ( $\backslash$ ILRKM $\text{YF}$ )-labeled AFA-PLN/SERCA1a, from the spectrum shown in Figure 9.3. The overall signal intensity stemming from Arg in this spectrum reduced.

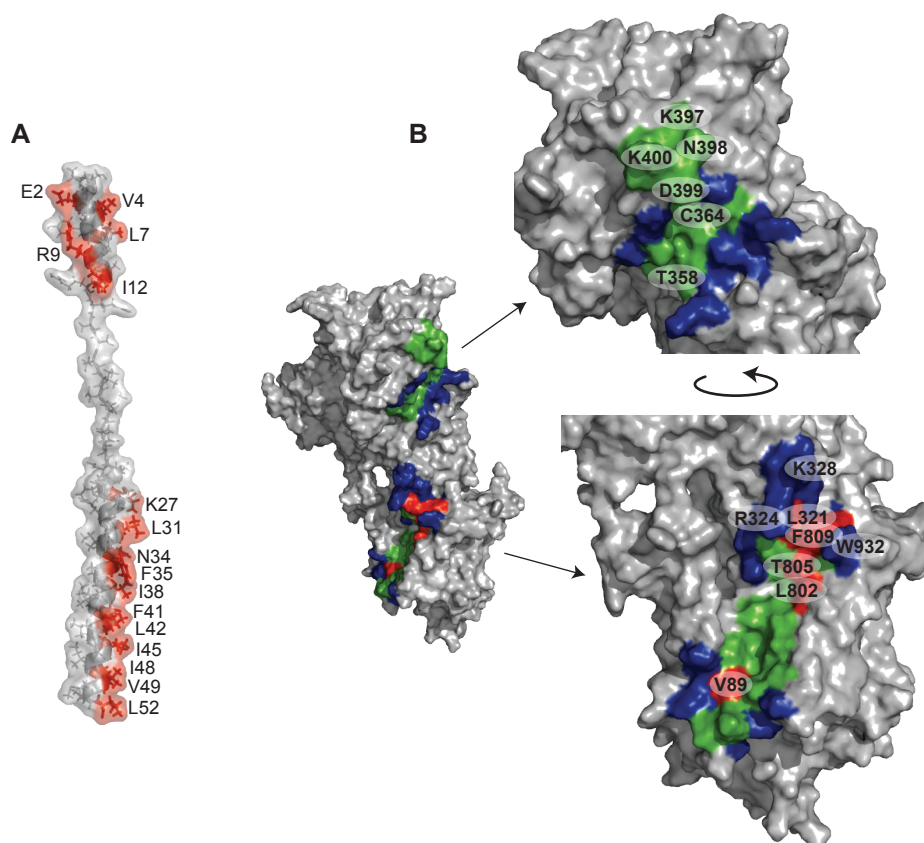


**Figure G.4:** (A) 1D Double-quantum filtered <sup>13</sup>C spectrum of U-(<sup>13</sup>C-Thr,<sup>15</sup>N-Ile)-AFA-PLN/SERCA1a, facilitating the identification of Thr residues with  $\alpha$ -helix-type chemical shifts (i.e., T8). T8 was difficult to identify in the 2Q/1Q correlation spectrum shown in Figure G.2B due to low signal-to-noise conditions for the Thr C $^{\alpha}$ -C $^{\beta}$  correlation, and in 1Q/1Q correlation spectra with zero-quantum mixing, see Figure 9.3B,C and Figure G.2A,C, due to the presence of strong diagonal signal. Gray lines show an approximation of the spectrum by Gaussian functions centered at the mean resonances value listed in Table G.1. (B) 1D SPECIFIC NCO-edited <sup>13</sup>C spectrum of U-(<sup>13</sup>C-Thr,<sup>15</sup>N-Ile)-AFA-PLN/SERCA1a, yielding the T17C' resonance, exclusively.



for the functional association of PLN with SERCA [287].

AIRs are generated from a list of *active* and *passive* residues. The terms *active*, *passive* and *interface residues* here are used in the sense of HADDOCK [70]. The terminology is not related to their biochemical role in complex formation. The surface of active residues of SERCA1a in Figure G.5 is shown in red. Active residues require an intermolecular contact to another active residue, or to a passive residue, *vide infra*. They are surface residues, selected according to HADDOCK based on biochemical or biophysical data, i.e. here, mutagenesis and disulphide cross-linking data. The



**Figure G.5:** Docking templates and HADDOCK interface residues. (A) PLN starting conformer generated by pre-folding the linear polypeptide chain in CNS as described in the Section G.2.3. Residues with cysteine mutants forming disulphide bonds with SERCA are show in yellow, other HADDOCK-‘active’ [70] residues in red. (B) SERCA structure (PDB ID: 2AGV), with ‘active’ residues shown in red and ‘passive’ residues colored green. The flexible interface comprises active and passive residues, and additionally, neighboring residues, including adjacent residues in the primary sequence, that are likely to exhibit structural changes upon complex formation (blue).

surface of passive residues of SERCA1a in Figure G.5 is colored green. Passive residues can satisfy intermolecular contact requirements from active residues, but do not need to be involved in any intermolecular contact themselves. For the transmembrane domain of SERCA1a, passive residues

were selected for those surface residues forming a stretch between L321' and V89', according to the finding that PLN forms a helix in this region. Additional residues beyond V89' (close to V49), were selected to satisfy potential contacts of the active C-terminal residue L52 of PLN. In the cytoplasmic domain of SERCA1a, passive residues are K397'-V402' and residues in the groove below, between T358' and C364'.

Two sets of AIRs were created to represent the transmembrane or cytoplasmic interactions independently. The complete lists of active and passive residues are as follows. Transmembrane: SERCA1a active: 89, 321, 802, 805, 809. SERCA1a passive: 84-86, 90, 93, 314, 317, 790, 793, 794, 797, 798, 801, 806, 953, 956, 957. AFA-PLN active: 27, 31, 34, 35, 38, 41, 42, 45, 48, 49, 52. AFA-PLN passive: none. Cytoplasmic: SERCA1a active: none. SERCA1a passive: 358, 360, 362, 364, 383, 384, 387, 397-402, 557, 559, 600, 602. PLN active: 2, 4, 7, 9, 12. PLN passive: none. Note that, according to ref. [292], a direct contact of loop residue L67 in SERCA with residues 21-30 in PLN, originally proposed from mutagenesis screening, was in the meantime identified to be of indirect nature in light of a high-resolution structural model of SERCA1a [291].

## G.2.2 Definition of Flexible Elements

In HADDOCK, flexible *interface* residues will be allowed to vary in their side chain or backbone conformation during different stages of the docking protocol. Interface residues shown in Figure G.5 are those colored other than gray. They comprise of all active and passive residues plus a number of neighboring residues that are likely to exhibit structural changes upon complex formation. Shown in blue in Figure G.5 are all such additional interface residues, which are sequence neighbors of active and passive residues (residues found below the surface not visible), or which have particularly exposed side chains close to the unambiguously identified binding regions (such as R324', K328', K605' and W932'). For AFA-PLN, full flexibility within the dihedral angle restraints was allowed during all docking stages, hence, no HADDOCK interface residues needed to be defined. The complete list of interface residues is as follows. SERCA1a: 83-94, 313-322, 324-328, 357-365, 382-388, 397-402, 556-560, 599-603, 605, 638, 789-799, 801-810, 932, 952-958.

## G.2.3 Docking Simulation Details

Docking models of PLN in complex with were calculated using the Crystallography and NMR System (CNS) [42] with protocols for High Ambiguity Driven biomolecular DOCKing (HADDOCK) [70], an approach allowing for flexible protein-protein docking. Covalent and non-bonded interactions were represented using the ALLHDG5.3 and OPLSX force fields, respectively [177, 179]. A docking template structure for AFA-PLN was generated using a standard simulated annealing protocol in CNS, with harmonic constraints on  $(\phi, \psi)$  backbone torsion angles of  $\phi = -57^\circ \pm 1^\circ$  and  $\psi = -48^\circ \pm 1^\circ$ , enforcing helical secondary structure for residues S0 to I12 and A24 to L52, and a minimum distance of 90 Å between K3C<sup>α</sup> and V49C<sup>α</sup>, leading to a linearly stretched, partially pre-folded conformer. This configuration of AFA-PLN was then docked to a structure for SERCA1a in the E<sub>2</sub> state determined by X-ray crystallography, PDB entry 2AGV [213]. Protons not defined by the X-ray diffraction model were added and energy refined by the docking software by means of a standard CNS protocol in HADDOCK. During the docking, harmonic constraints on  $(\phi, \psi)$  backbone torsion angles of  $\phi = -57^\circ \pm 10^\circ$  and  $\psi = -48^\circ \pm 10^\circ$  ( $\pm 15^\circ$  variation without energy penalty for the first and last residue of each helix) were used to enforce helical secondary structure

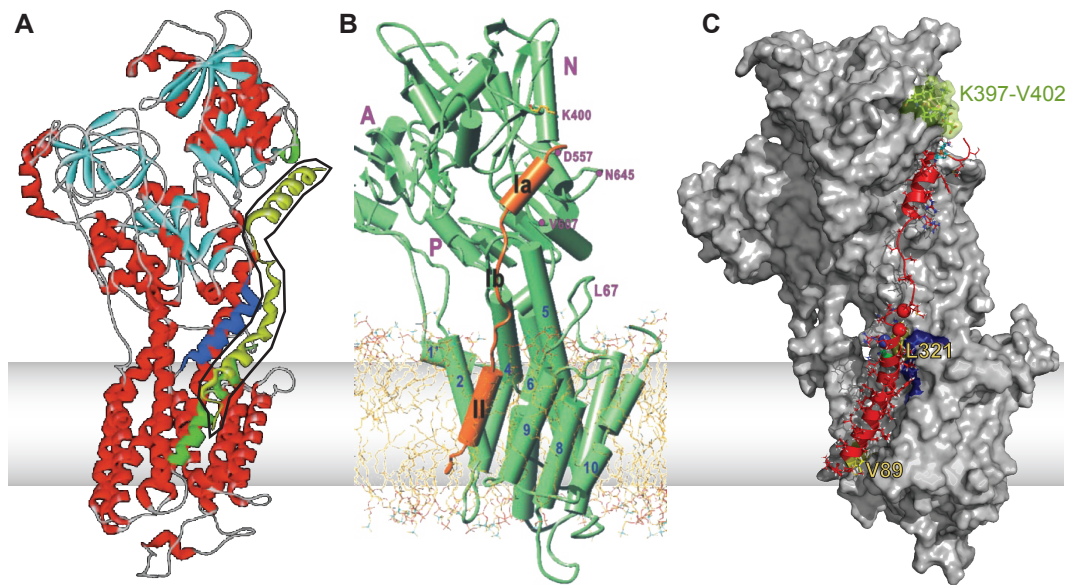
for residues constituting the two helices, V4–R13 and A24–V49 according to ssNMR (Figure 9.4). In order to check if the chemical shift assignments for Q22 (non-helical character) are in line with the  $P21C'$ – $A24C^\alpha$  distance measured by Hughes and Middleton [123], which was interpreted as an indication of an  $\alpha$ -helical backbone for P21–A24, backbone torsion angles were predicted for Q22 from chemical shifts using PREDITOR (<http://wishart.biology.ualberta.ca/predictor/>) [30]. The resulting  $(\phi, \psi)$  pair was used as a harmonic dihedral angle constraint.

Docking in HADDOCK was driven by intermolecular distance constraints derived from mutagenesis and cross-linking experiments reported in the literature, *vide supra*, as summarized in Table 9.1. The protein template structures were initially separated by 150 Å in random orientations. Over the course of the docking simulation, the backbone of PLN was constrained to a helical backbone conformation as described above, and left fully flexible within the limits of the force field and specific distance constraints elsewhere. The number of MD steps during first rigid body cooling stage, second cooling stage with flexible side-chains at interface and third cooling stage with fully flexible interface were increased by a factor of four (i.e. 2000, 2000, 4000, 4000). Otherwise, standard HADDOCK dynamics parameters were used with a constant dielectric of  $\epsilon = 10$  in the vacuum refinement. In the rigid-body docking stage, 1000 structures were calculated, out of which the 200 best-scoring models were thereafter subjected to flexible docking. Finally, refinement in a water shell was performed for the 20 best-scoring structures after the vacuum refinement of *iteration 1*. While this gives a more realistic picture for the extra-membrane domains, water refinement of the transmembrane domain is biophysically meaningless. Comparison of structural models before and after water refinement, however, showed that effects on the transmembrane helix are marginal, due to the relatively large number of constraints (dihedral angles from secondary chemical shift analysis, distances from mutagenesis and cross-linking). Hence, models presented in Figure 9.6 were taken from the last iteration in water.

The five best-scoring (weighted HADDOCK score) models from the last iteration were selected for presentation Figure 9.6. No additional clustering was performed to preserve the spread of unconstrained segments in AFA-PLN, which point to conformational flexibility in the complex. Other models showed excellent convergence towards the presented ensemble. For example, the model corresponding to the docking score median after *iteration 1* (vacuum refinement) showed no dihedral angle violation larger than  $5^\circ$ , no violation of an unambiguous distance constrain larger than 0.3 Å, and only one violation of an AIR in excess of 0.3 Å.

#### G.2.4 Model Comparison Supplement

For reference, structural models discussed in the main text (Section 9.5.1) are reproduced in Figure G.6 from the original publications.



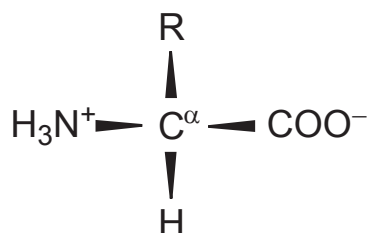
**Figure G.6:** Structural models discussed in the main text (Section 9.5.1), reproduced from the original publications, for comparison. (A) Model by Hutter et al. [125]. (B) Model by Toyoshima and co-workers [292]. (C) SsNMR-based docking model. For the sake of clarity, only one model from the ensemble shown in Figure 9.6 is reproduced, neglecting, however, structural variations that may be of relevance. The presentation and orientation of the models varies slightly, but structural aspects such as positioning of the C-terminus and helix extent can readily be identified. The gray bar approximately indicates the hydrophobic core of the membrane.

## Appendix H

# Primer on Amino Acids, Peptides and Proteins

## H.1 Amino Acids

Twenty common amino acids are the basic building blocks of proteins in living organisms [39]. Amino acids consist of a central carbon atom ( $C^\alpha$ , CA) with attached amino ( $NH_3^+$ ) and carboxyl ( $COO^-$ ) groups<sup>1</sup>, and a *side group* or *side chain* (R), by which the amino acids differ. All naturally expressed amino acids exist in L-form (chirality with counterclockwise order of priority), except from glycine, where L and R are degenerate. Figure H.1 depicts the basic structure of an amino acid. A list of all twenty common amino acids is given in Table H.1. A mnemonic device for linking



**Figure H.1:** Basic structure of an L-amino acid. R denotes a side group.

the one-letter code to the names of the amino acids can be found in [39]. The structure of each amino acid is given in Figure H.2.

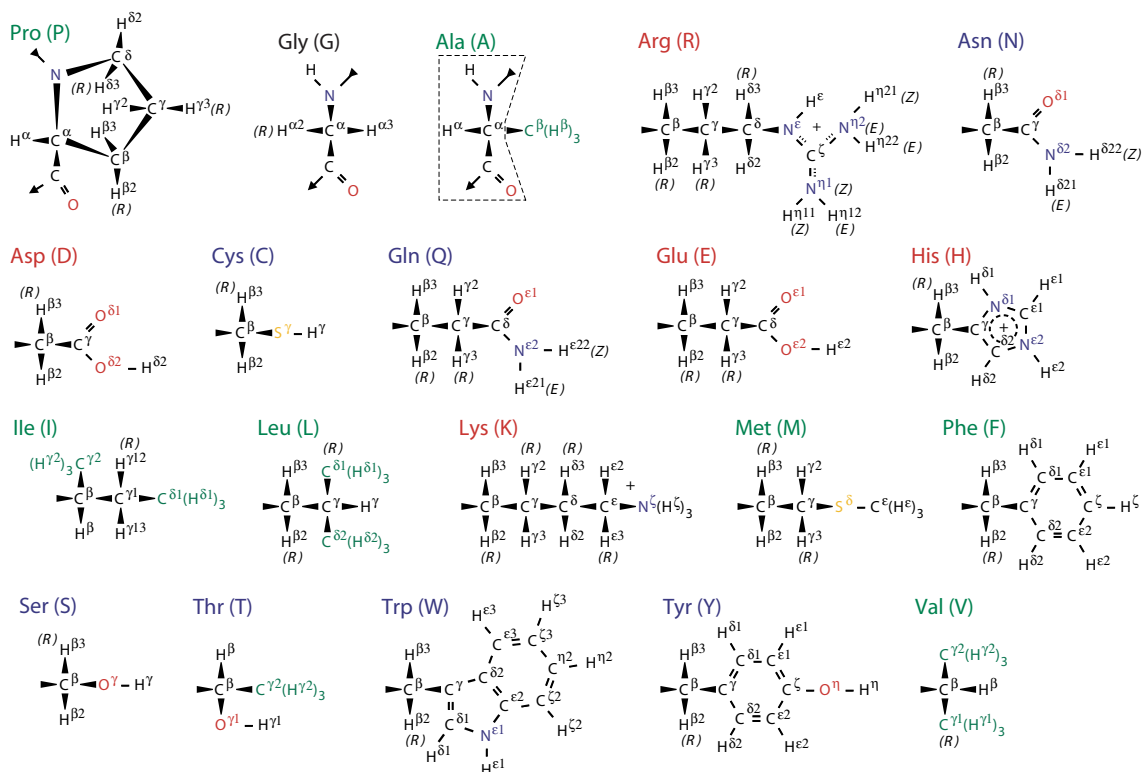
## H.2 Peptides

Amino acids can connect to each other through the formation of peptide bonds, see Figure H.3. The peptide bond is a partial double bond, delocalized over to the  $C_i-N_{i+1}$  bond, where  $i$  indicates a residue number. The peptide unit has a dipole moment due to this charge delocalization. Rotation around the peptide bond is restricted, and the dihedral angle given by O, C, N and H is fixed to approximately  $180^\circ$ , defining the *peptide plane*.

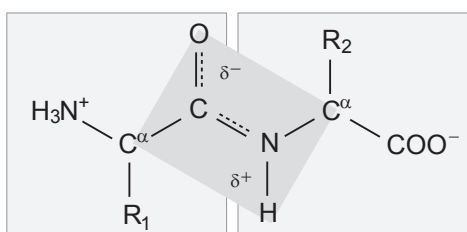
<sup>1</sup>IUPAC recommendations [189] suggest to denote the backbone carbonyl carbon by C or C'. For practical reasons, CO may be used here equivalently.

Amino Acid	Abbreviations		side chain properties
Glycine	Gly	G	unpolar, aliphatic
Alanine	Ala	A	
Valine	Val	V	
Leucine	Leu	L	
Isoleucine	Ile	I	
Methionine	Met	M	
Phenylalanine	Phe	F	aromatic
Tyrosine	Tyr	Y	
Tryptophane	Trp	W	
Serine	Ser	S	polar, uncharged
Proline	Pro	P	
Threonine	Thr	T	
Cysteine	Cys	C	
Asparagine	Asn	N	
Glutamine	Gln	Q	
Lysine	Lys	K	positively charged
Histidine	His	H	
Arginine	Arg	R	
Aspartate	Asp	D	negatively charged
Glutamate	Glu	E	

**Table H.1:** Amino acid names, abbreviations and side chain classification.



**Figure H.2:** The twenty common amino acids. Amino acid label colors differentiate between hydrophobic (green), polar (blue) and charged (red) amino acids. Glycine can have different affinities. Colors of atom identifiers point out nitrogen (blue), oxygen (red) and sulfur (orange), as well as methyl groups (green). Graphics reproduced from ref. [189], with changes.



**Figure H.3:** Schematic representation of two amino acids with side chains  $R_1$  and  $R_2$  connected by a peptide bond.

### H.3 Polypeptides (Primary Structure)

The order of amino acids in the polypeptide chain is called the *primary structure*, see e.g. Figure 9.3A.

### H.4 Secondary Structure and Structural Motifs

Local folding motifs are called the *secondary structure* of a polypeptide or protein. The two most common structural motifs in proteins are  $\alpha$ -helices and  $\beta$ -sheets, see below. The ubiquitin fold, for example, comprises a 12-residue  $\alpha$ -helix and a 5-stranded anti-parallel  $\beta$ -sheet. A number of folding motifs exist, which are themselves build from  $\alpha$ -helices and  $\beta$ -sheets, or in a similar fashion, such as the  $\beta$ -helix of Gramicidin A. In ubiquitin, one additionally finds two  $\beta$ -hairpins, consisting of a pair of anti-parallel  $\beta$ -strands joined by a small loop, and two left-handed or '3-10' helices. Furthermore, several residues form random *coils* that are usually not further classified. The textbook by Branden and Tooze [39] discusses various folding motifs and their biological relevance.

#### H.4.1 $\alpha$ -Helix

The  $\alpha$ -helix is one of the two major elements of secondary structure in proteins, described 1951 by Linus Pauling as a structure which would be stable and energetically favourable in proteins [223]. Complete verification was given by the first high-resolution protein structure, the structure of myoglobin, where basically all secondary structure is helical. The  $\alpha$ -helix has 3.6 residues per turn with hydrogen bonds between C'=O of residue  $n$  and NH of residue  $n + 4$ , Figure H.4A. Variations with 3 residues (3-10 helix) and 5 residues ( $\pi$ -helix) per turn are found, but are less common. All NH and CO groups are joined with hydrogen except at the ends. As a consequence, the ends of the  $\alpha$ -helix are polar and are often found at the surface of the proteins. The  $\alpha$ -helix has an overall dipole moment, which results from the peptide units dipole moments that are aligned in the same orientation along the helical axis. Side chains point away from the helix. As they are exposed to the environment, the amino acid sequence correlates with the possible locations of a helix with respect to an hydrophobic or hydrophilic environment.

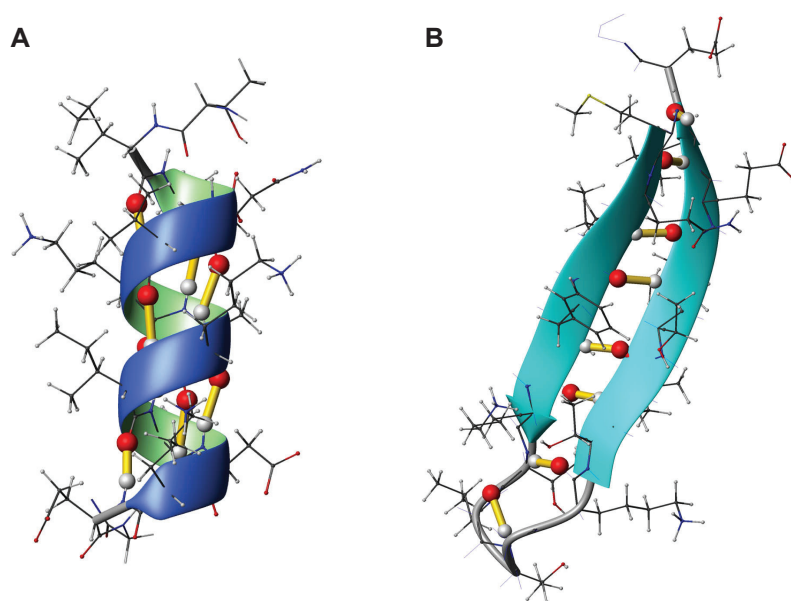
#### H.4.2 $\beta$ -sheet

The second major structural element in globular proteins is the  $\beta$ -sheet, formed by  $\beta$ -strands. These strands are aligned parallel or antiparallel to each other such that hydrogen bonds can form between C'=O groups of one strand and NH groups on an adjacent strand, Figure H.4B. Side chains point alternately to both directions of the  $\beta$ -sheet. Mixed  $\beta$ -sheets, where strands have parallel alignment to one neighboring strand and antiparallel alignment to the opposite strand are found in only about 20% of known sheet structures [39], whereas pure parallel or antiparallel alignment in a sheet is more common.

#### H.4.3 Ramachandran Plot

The backbone torsion angles  $\phi$  and  $\psi$  may be assessed in a Ramachandran plot [238]. Certain backbone torsion angle configurations are sterically not allowed, while others correspond to specific

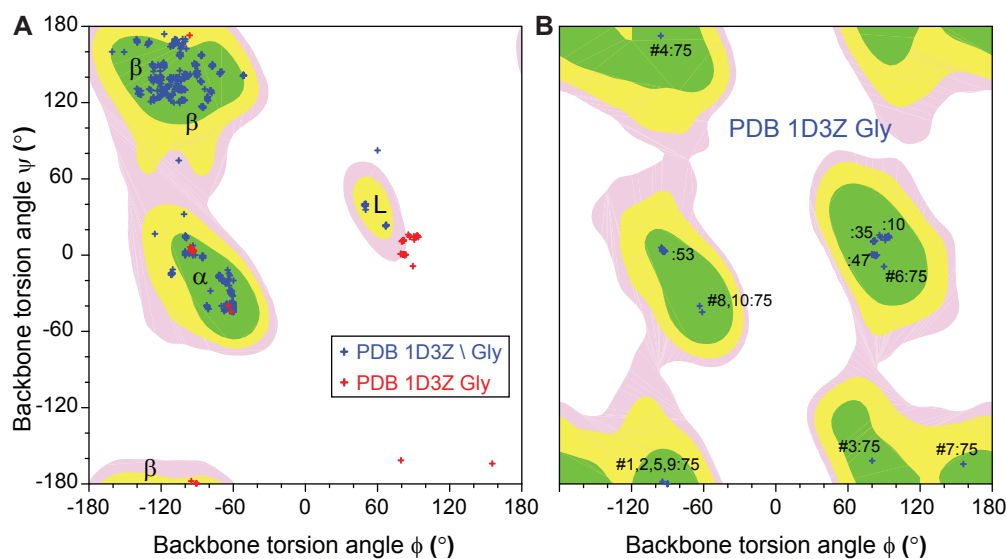




**Figure H.4:**  $\alpha$ -helix and  $\beta$ -sheet configurations, two common elements of secondary structure in proteins. Both are formed by hydrogen bonds between an amide proton (white) and a carbonyl oxygen (red), as indicated by yellow sticks. (A)  $\alpha$ -helix from ubiquitin (PDB ID 1D3Z) residues 23–34. Hydrogen bonds form between  $O_{C'}^{(n)}$  (red) and  $H_N^{(n+4)}$  (white). (B) Twisted antiparallel  $\beta$ -sheet formed of two strands of residues 1–18, comprising a hydrogen-bonded turn. This sheet is part of a larger  $\beta$ -sheet formed by five strands in ubiquitin.

backbone angles of secondary structure, as can be calculated. Colored areas in figure Figure H.5A indicate sterically allowed and energetically favorable manifolds for backbone torsion angles, derived from statistics on a large number of accurately determined protein structures [152]. Borders vary, depending on the set of structures used. Three plains that correspond approximately to the usual right-handed  $\alpha$ -helices,  $\beta$ -strands and left-handed  $\alpha$ -helices can be identified. Areas in between are sparsely occupied due to steric hindrance.

Glycine residues can adopt many different conformations, because a second proton on  $C^\alpha$  substitutes for the side chain, and grants more sterically allowed torsion angle combinations. Hence,  $(\phi, \psi)$  tuples cover regions in the Ramachandran plot, see Figure H.5B, which are not populated for other residue types.



**Figure H.5:** (A) Ramachandran plot for all residues of PDB entry 1D3Z (ubiquitin, 76 residues). Glycine residues are indicated by red marks. Background color areas indicate common regions of  $(\phi, \psi)$  angle pairs for helices and sheets. The three outliers at the region of left-handed  $\alpha$ -helices result from glycine residues, which can adapt a wider range of  $(\phi, \psi)$  angles because of a missing side chain. (B) Glycine residues only, for all ten conformers in the ensemble of PDB ID 1D3Z. '#' indicates a molecule number, ranging from 1 to 10, ':' indicates a residue number. Residue 76 has no  $\psi$  angle and is therefore not present.

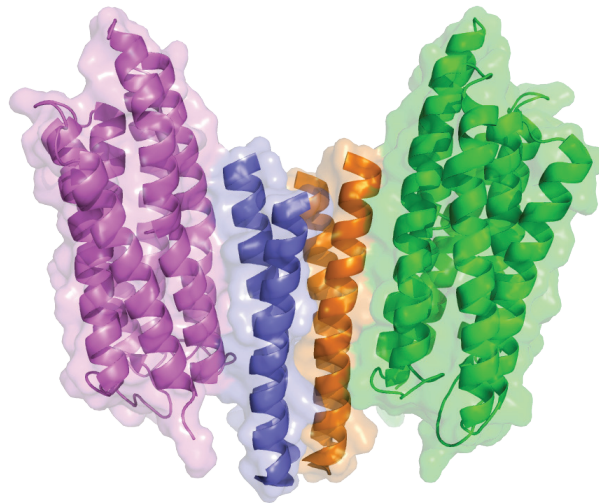
## H.5 Tertiary Structure

Secondary structure motifs combine to form domains. A domain is a polypeptide chain or part of it that can fold independently into a stable tertiary structure. The term tertiary structure is used both for the arrangement of motifs in domains, and for the way a single polypeptide chain folds into one or several domains. In ubiquitin, Figure 8.6, we find one domain, whereas SERCA1a (Figure 9.6, gray) has multiple domains. The tertiary structure of ubiquitin is characterized by the twist and arrangement of the five-stranded  $\beta$ -sheet, and the relative position of the  $\alpha$ -helix. In

more detail, a description of the tertiary structure also includes the position of smaller elements like the two 3-10-helices and loops.

## H.6 Quaternary Structure

Large proteins may comprise several polypeptide chains (subunits), the so-called quaternary structure. Subunits can have independent or cooperative function. Ubiquitin does not have a quaternary structure. Quaternary structure can for example be found in haemoglobin, the oxygen carrier in blood [226]. An illustration is given in Figure H.6 by the quaternary structure of the NpSRII/NpHtrII signaling rhodopsin–transducer complex (PDB entry 1H2S) [93].



**Figure H.6:** Tertiary and quaternary structure of the NpSRII/NpHtrII signaling rhodopsin–transducer complex (PDB entry 1H2S). The two signaling rhodopsin chains are colored magenta and green, the two transducer chains are given in blue and orange.



## Appendix I

### Pulse Programs

Example pulse programs used with Bruker spectrometers Avance 400 or Avance 600 running XWinNMR 3.5, or Avance-II 800 running Topspin 2.0, with double and triple resonance probes, running XWinNMR 3.5 or Topspin 2.0.

Phases:  $\phi_{31}$  (ph31) denotes the receiver phase. Unless a different modulo setting is given in parenthesis prior to a phase setting, phases 0, 1, 2 and 3 correspond to  $0^\circ$  ( $= x$ ),  $90^\circ$  ( $= y$ ),  $180^\circ$  ( $= -x$ ) and  $270^\circ$  ( $= -y$ ).

Phase cycling and the selection of coherence transfer pathways are described in great detail in ref. [36], and for example in refs. [16, 78, 121, 143, 296].

#### $\pi/2$ -pulse and acquisition (1D)

```
1 ze
2 d1

;90 f1
  1u p1:f1
  p1:f1 ph1

;acquisition (no f2 decoupling)
  10u
  gosc ph31
  lo to 2 times ns
  100m wr #0
exit

ph1 = 0 1 2 3
ph31= 2 3 0 1
```

$\pi/2$ -pulse with decoupling (1D)

```

1 ze
2 d1
  2u fq=0:f2

;90 f1
  1u pl1:f1
  p1:f1 ph1

;acquisition WITH f2 decoupling
  5u pl12:f2
  5u cpds2:f2
  gosc ph31
  1m do:f2

  lo to 2 times ns
  100m wr #0
exit

ph1 = 0 1 2 3
ph31= 0 1 2 3

```

## Cross-Polarization (CP)

```

1 ze
2 d1 do:f2
  2u fq=0:f2

;90 f2
  2u p12:f2
  p2:f2 ph1

;CP f2--f1
  1u p15:f1 p16:f2
  (p15 ph2):f1 (p15:spf0 p16 ph0):f2

;acquisition WITH decoupling
  5u pl12:f2
  5u cpds2:f2
  gosc ph31
  1m do:f2

  lo to 2 times ns
  100m wr #0
exit

ph0 = 0
ph1 = 1 1 1 1   3 3 3 3
ph2 = 0 1 2 3
ph31= 0 1 2 3   2 3 0 1

```

## Double-CP (1D)

```

1 ze
2 d1
  2u fq=0:f2
  2u fq=cnst20:f1

;90 f2
  2u pl2:f2
  (p2 ph1):f2

;CP f2--f3
  1u pl5:f3 pl6:f2
  (p15 ph2):f3 (p15:spf0 pl6 ph0):f2

;f2 decoupling for 2nd CP and Acquisition
  1u pl12:f2
  1u cpds2:f2

;CP f3--f1 with f2 decoupling
  1u pl15:f3 pl16:f1
  (p25 ph3):f3 (p25:spf1 pl16 ph4):f1

;acquisition
  5u fq=0:f1
  5u
  gosc ph31
  1m do:f2
  lo to 2 times ns
  100m wr #0
exit

ph0 = 0
ph1 = 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3
ph2 = 0
ph3 = 0 0 0 0 2 2 2 2
ph4 = 0 1 2 3
ph31= 0 1 2 3 2 3 0 1 2 3 0 1 0 1 2 3

```

## NCC SD (3D)

```

1 ze
2 d1

;90 f2
  1u pl2:f2
  p2:f2 ph1

;CP f2--f3
  1u pl5:f3 pl6:f2
  (p15 ph2):f3 (p15:spf0 pl6 ph0):f2

;t1 evolution f3
  1u pl12:f2
  1u cpds2:f2
  d0
  2u pl13:f2
  2u cw:f2

;CP f3--f1
  1u fq=cnst20:f1
  1u pl15:f3 pl16:f1
  (p25 ph3):f3 (p25:spf1 pl16 ph4):f1
  1u fq=0:f1

;t2 evolution f1
  1u pl12:f2
  1u cpds2:f2
  d10

;PDS
  1u pl1:f1
  2u do:f2
  p1:f1 ph5
  d11
  p1:f1 ph7

;Acquisition
  2u pl12:f2
  2u cpds2:f2
  gosc ph31
  1m do:f2
  lo to 2 times ns
  100m wr #0 if #0 zd
  1m id10
  1m ip4 ;increment phases prior to t2
  lo to 2 times td2

```

```

1m rd10
1m id0
1m ip2 ;increment phases prior to t1
lo to 2 times td1
exit

ph0 = 0 0 1 1 1 1 0 0
ph1 = 1 1 2 2 2 2 1 1 1 1 2 2 2 2 1 1
      3 3 0 0 0 0 3 3 3 3 0 0 0 0 3 3
ph2 = 0 0 1 1 1 1 0 0 2 2 3 3 3 3 2 2
ph3 = 0 0 1 1 1 1 0 0
ph4 = 0 0 1 1 3 3 2 2
ph5 = 1 1 2 2 2 2 1 1
ph7 = 0 1 3 0 1 2 2 3
ph31= 0 1 3 0 3 0 0 1 2 3 1 2 1 2 2 3
      2 3 1 2 1 2 2 3 0 1 3 0 3 0 0 1

```

## Refocused INEPT (1D)

```

;Echo delays: cnst2 = J(CH), cnst3 e.g. 6
"d4=1/(4*cnst2)"
"d3=1/(cnst3*cnst2)"

;compensation delays, p1>p2
"d5=p1-p2"
"d6=d5/2"

1 ze
2 d1

;90 1H
2u fq=0:f2
2u pl2:f2 pl1:f1
p2:f2 ph1

;Spinecho
d4
(d5 p2*2 ph2):f2 (p1*2 ph4):f1
d4

;Polarization Transfer
(d6 p2 ph3):f2 (p1 ph5):f1

;Spinecho
d3
(d5 p2*2 ph2):f2 (p1*2 ph6):f1
d3

;Acquisition
2u pl12:f2
2u cpds2:f2
gosc ph31
1m do:f2
lo to 2 times ns
100m wr #0
exit

ph1 = 0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph2 = 0 2
ph3 = 1 1 3 3
ph4 = 0 2
ph5 = 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph6 = 0 2 0 2 1 3 1 3
ph31= 0 0 2 2 1 1 3 3

```



## Refocused INEPT TOBSY (2D)

```

(d5 p2*2 ph2):f2 (p1*2 ph6):f1
d3

;INEPT echo delays
"d4=1/(4*cnst2)"
"d3=1/(cnst3*cnst2)"

;compensation delays
"d5=p1-p2"
"d6=d5/2"

;TOBSY mixing
;cnst31 = MAS rate
;p11 = 90 pulse for TOBSY

;P931 element:
;l5 = 9 is 3 rotor periods
;e.g. MAS=8kHz, Rf=48kHz, l5=144=6ms
"p11=1s/(24*cnst31)"

;P961 element:
;l5 = 9 is 6 rotor periods
;e.g. MAS=10kHz, Rf=30kHz, l5=90=6ms
;"p11=1s/(12*cnst31)"

1 ze
2 d1

;90 1H
2u fq=0:f2
2u fq=cnst20:f1
2u pl2:f2 pl1:f1
p2:f2 ph1

;t1 evolution
d0
(p1*2 ph0):f1
d0

;Spinecho
d4
(d5 p2*2 ph2):f2 (p1*2 ph4):f1
d4

;Polarization Transfer
(d6 p2 ph3):f2 (p1 ph5):f1

;Spinecho
d3

```

```

;TOBSY mixing
2u pl11:f1
p11:f1 ph10
3 p11*1:f1 ph7
p11*4:f1 ph8^
p11*3:f1 ph7^
lo to 3 times l5
p11:f1 ph11

;Acquisition
2u pl12:f2
2u cpds2:f2
2u fq=0:f1
10u
gosc ph31
1m do:f2
1m rpp7
1m rpp8
lo to 2 times ns
100m wr #0 if #0 zd
1m id0
1m ip1
lo to 2 times td1
exit

ph0 = 0
ph1 = 0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph2 = 0 2
ph3 = 1 1 3 3
ph4 = 0 2
ph5 = 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph6 = 0 2 0 2 1 3 1 3
ph7 = (16384) 0 1820 3641 5461
7282 9102 10923 12743 14566
ph8 = (16384) 8192 10012 11833 13653
15474 910 2730 4551 6371
ph10= 1 1 1 1 2 2 2 2 3 3 3 3 0 0 0 0
ph11= 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph31= 0 0 2 2 1 1 3 3

```

## CC Spin Diffusion (2D)

```

1 ze
2 d1 do:f2
  5u fq=0:f2

;90 f2
  1u pl2:f2
  p2:f2 ph1

;CP
  1u pl5:f1 pl6:f2
  (p15 ph2):f1 (p15:spf0 pl6 ph0):f2

;t1 evolution
  1u pl12:f2
  1u cpd2:f2
  d0
  2u do:f2

;(Proton-Driven) Spin Diffusion, d11
  1u pl1:f1
  p1:f1 ph3
  d11
  p1:f1 ph4

;Acquisition
  5u pl12:f2
  5u cpd2:f2
  gosc ph31
  1m do:f2
  lo to 2 times ns
  100m wr #0 if #0 zd
  1m id0
  1m ip2
  lo to 2 times td1
exit

ph0 = 0
ph1 = 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3
ph2 = 0 0 0 0 2 2 2 2
ph3 = 1
ph4 = 3 0 1 2
ph31= 0 1 2 3 2 3 0 1 2 3 0 1 0 1 2 3

```

R<sup>2</sup>TR constant time (2D)

```

1 ze
2 d1

;90 f2
  2u pl2:f2
  p2:f2 ph1

;CP
  2u pl5:f1 pl6:f2
  (p15 ph2):f1 (p15:sp0 ph0):f2

;Constant time delay
  2u pl1:f1 pl13:f2
  2u cpd2:f2
  p1:f1 ph3
  d24
  p1:f1 ph4

;t1 evolution
3 d0
  lo to 3 times l0

;R2TR mix
  p1:f1 ph5
  2u pl11:f1
  (p21:sp1 ph6):f1
; 2u pl11:f1
  p11:f1 ph7
  (p21:sp2 ph8):f1
  2u pl1:f1
  p1:f1 ph9
;Acquisition
  2u pl12:f2
  2u cpd2:f2
  10u:f1 ph30
  5u adc ph31
  aq
  1m do:f2
  rcyc=2
  100m wr #0 if #0 zd
  1m iu0
  1m ip4
  "d24=d24-d0"
  lo to 2 times td1
exit

```

```

ph0 = 0
ph1 = 1 1 1 1 3 3 3 3
ph2 = 0
ph3 = 1
ph4 = 3 0 1 2
ph5 = 1 2 3 0
ph6 = 2 3 0 1
ph7 = 2 3 0 1
ph8 = 2 3 0 1
ph9 = 3 0 1 2
ph30= 0
ph31= 0 1 2 3 2 3 0 1

```

## CHHC (2D)

```

1 ze
2 d1

;90 f2
  3u p12:f2
  p2:f2 ph1

;CP
  3u p15:f1 p16:f2
  (p16 ph2):f1 (p16:spf0 p16 ph0):f2

;Spinlock
  3u p11:f1
  p1:f1 ph3
  d2
  p1:f1 ph4

;t1 evolution
  3u p112:f2
  3u cpd2:f2
3 d0
  lo to 3 times l0
  3u do:f2

;CP to protons
  3u p15:f1 p16:f2
  (p15 ph5):f1 (p15:spf0 p16 ph6):f2

;Mixing protons
  3u p12:f2
  p2:f2 ph7
  d5
  p2:f2 ph8

;CP back to the carbons
  3u p15:f1 p16:f2
  (p15 ph10):f1 (p15:spf0 p16 ph9):f2

;Acquisition
  5u p112:f2
  5u cpd2:f2
  gosc ph31
  1m do:f2
  lo to 2 times ns
  100m wr #0 if #0 zd
  1m iu0

```

```

1m ip4
1m ip3
1m ip2
lo to 2 times td1
exit

ph0 = 0
ph1 = 1 1 1 1 1 1 1 1 1 1 1 1 1 1
      3 3 3 3 3 3 3 3 3 3 3 3 3 3
ph2 = 0
ph3 = 1
ph4 = 3 1
ph5 = 0
ph6 = 0 0 1 1 2 2 3 3
ph7 = 1 1 2 2 3 3 0 0
ph8 = 3 3 0 0 1 1 2 2 1 1 2 2 3 3 0 0
ph9 = 0 0 1 1 2 2 3 3
ph10= 0 0 1 1 2 2 3 3
ph31= 0 2 1 3 2 0 3 1 2 0 3 1 0 2 1 3
      2 0 3 1 0 2 1 3 0 2 1 3 2 0 3 1

```

## CHHC without $z$ -filter (2D)

```

1 ze
2 d1 do:f2

;90 f2
3u pl2:f2 pl5:f1
p2:f2 ph1

;CP H--C and t1 evolution
(p15 ph2):f1 (p15:sp0 ph0):f2
1u pl12:f2
d0 cpds2:f2
3u do:f2

;CHHC mixing
(p16 ph3):f1 (p16 ph4):f2
3u pl2:f2
p2:f2 ph5
d5
(p17 ph8):f1 (p17 ph7):f2
p2:f2 ph6
3u pl16:f2

;Acquisition
1u pl12:f2
1u cpds2:f2
10u
go=2 ph31
1m do:f2
100m wr #0 if #0 zd
1m id0
1m ip2
lo to 1 times td1
exit

ph0 = 0
ph1 = 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
      3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
ph2 = 0 0 2 2
ph3 = 0
ph4 = 1 3
ph5 = 0
ph6 = 2
ph7 = 1
ph8 = 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph30= 0
ph31= 0 2 2 0 1 3 3 1 2 0 0 2 3 1 1 3

```

2 0 0 2 3 1 1 3 0 2 2 0 1 3 3 1

## 2Q/1Q CC with SPC5 (2D)

```

;cnst31 MAS freq (Hz)
;l5 DQ exc/rec time:
; multiple of 80 = 4 rotor periods
;in0 rotor synchronize

;e.g. MAS=7.5kHz, in0=26.67us (tr/5),
; SW1=124ppm (600MHz), ipp7=819
;e.g. MAS=7.5kHz, in0=13.33us (tr/10),
; SW1=248ppm (600MHz), ipp7=410

;Change phase incr ip7
; according to spinning freq
;ip7 = 4k/n, e.g. 683 for n=6

"p11=1s/(cnst31*20)"
"in0=1s/(cnst31*5)" ;choose n
"d0=0u"

1 ze
2 d1
  1u fq=0:f2
  1m rpp6
  1m rpp7

;90 f2
  1u p12:f2
  p2:f2 ph1

;CP f2--f1
  1u p15:f1 p16:f2
  (p15 ph2):f1 (p15:spf0 p16 ph0):f2
  1u p113:f2
  1u cpds1:f2

;2Q excitation with SPC5
  1u p11:f1
  p1:f1 ph4
  1u p111:f1
3 p11:f1 ph6^ ipp7
  lo to 3 times l5

;t1 evolution (2Q) under cpds2 decoupling
  d0

;2Q reconversion with SPC5
4 p11:f1 ph7^

```

```

lo to 4 times 15
1u pl1:f1
p1:f1 ph5
7373 3277 3277 3277 3277 7373 7373 7373
819 4915 4915 4915 4915 819 819 819
2458 6554 6554 6554 6554 2458 2458 2458

;Acquisition
1u do:f2
1u pl12:f2
1u cpds2:f2
10u
gosc ph31
1m do:f2
1m ip5
1m ip7*2048
lo to 2 times ns
100m wr #0 if #0 zd
1m id0
1m ip7*819
1m ip2
1m ip4
1m ip6*1024
lo to 1 times td1
exit
ph31= 0 3 2 1 2 1 0 3

ph0 = 0
ph1 = 1 1 1 1 3 3 3 3
ph2 = (8) 0
ph4 = (8) 6
ph5 = 1
ph6 = (8192)
0 4096 4096 4096 4096 0 0 0
1638 5734 5734 5734 5734 1638 1638 1638
3277 7373 7373 7373 7373 3277 3277 3277
4915 819 819 819 819 4915 4915 4915
6554 2458 2458 2458 2458 6554 6554 6554
4096 0 0 0 0 4096 4096 4096
5734 1638 1638 1638 1638 5734 5734 5734
7373 3277 3277 3277 3277 7373 7373 7373
819 4915 4915 4915 4915 819 819 819
2458 6554 6554 6554 6554 2458 2458 2458

ph7 = (8192)
0 4096 4096 4096 4096 0 0 0
1638 5734 5734 5734 5734 1638 1638 1638
3277 7373 7373 7373 7373 3277 3277 3277
4915 819 819 819 819 4915 4915 4915
6554 2458 2458 2458 2458 6554 6554 6554
4096 0 0 0 0 4096 4096 4096
5734 1638 1638 1638 1638 5734 5734 5734

```

## 2Q/1Q CC with POST-C7 (2D)

```

;p111 POST-C7 power
;cnst31 spinning rate (Hz)
;l5 7 = 2 rotor periods
;l6 t1 incr (/p11)

"p11=(1s/cnst31)/28"

1 ze
2 d1

;Reset phase pointers
1m rpp6
1m rpp7
1m rpp8
1m rpp9
1u pl2:f2
p2:f2 ph1
1u pl5:f1 pl6:f2
(p15 ph2):f1 (p15:spf0 pl6 ph0):f2

;Start decoupling
1u pl13:f2
1u cw:f2

;90 on X
1u pl1:f1
p1:f1 ph4

;2QC exc with pc7
1u pl11:f1
3 p11*1:f1 ph6
p11*4:f1 ph7^ ipp8
p11*3:f1 ph6^ ipp9
lo to 3 times l5

;t1 2qc evolution
1u pl12:f2
1u cpds2:f2
d0
1u pl13:f2
1u cw:f2

;2QC rec with pc7
4 p11*1:f1 ph8
p11*4:f1 ph9^
p11*3:f1 ph8^
lo to 4 times l5

;90 on X
1u pl1:f1
p1:f1 ph5

;Acquisition
5u pl12:f2
5u cpds2:f2
gosc ph31
1m do:f2
1m ip8*16384
1m ip9*16384
lo to 2 times ns
100m wr #0 if #0 zd
id0

;TPPI
1m ip2
1m ip4
1m ip6*8192
1m ip7*8192

;Rotor synchronize
5 1m ip8*1170
1m ip9*1170
lo to 5 times l6
lo to 1 times td1
exit

ph0 = 0
ph1 = 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
      3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
ph2 = (8) 0
ph4 = (8) 6
ph5 = 1 1 1 1 2 2 2 2 3 3 3 3 0 0 0 0
ph6 = (65536) 0 9362 18725 28087
          37449 46811 56174
ph7 = (65536) 32768 42130 51493 60855
          4681 14043 23406
ph8 = (65536) 0 9362 18725 28087
          37449 46811 56174
ph9 = (65536) 32768 42130 51493 60855
          4681 14043 23406
ph31= 0 2 0 2 1 3 1 3 2 0 2 0 3 1 3 1
      2 0 2 0 3 1 3 1 0 2 0 2 1 3 1 3

```

2Q/1Q CC with R14<sub>2</sub><sup>6</sup> (2D)

```

;R14(6,2) = 180(77.14)/180(-77.14)
;R element 180 pulse

;cnst31 spinning speed
;p11      180 pulse at rf=3.5*MAS
; R element here 180 pulse
;l5 7 = 2 rotor periods

;R element 180
"p11=1s/(cnst31*7)"

;in0=tr/4 phase cycle rotor synchronized
; for 1/4 rotor period
"in0=1s/(cnst31*4)"\

1 ze
2 d1
  2u fq=0:f2
  2u fq=cnst20:f1

;90 f2
  2u pl2:f2
  p2:f2 ph1
;CP
  2u pl5:f1 pl6:f2
  (p15 spf0 p16 ph0):f2

;DQ excitation R14(6,2)
  2u pl12:f2
  2u cpds2:f2
  2u pl1:f1
  p1:f1 ph3
  2u pl11:f1
3 p11:f1 ph25
  p11:f1 ph26
  lo to 3 times l5

;t1 DQ evol
  d0

;DQ reconversion R14(6,2)
4 p11:f1 ph27
  p11:f1 ph28
  lo to 4 times l5
  2u pl1:f1
  p1:f1 ph4

;Acquisition
  2u pl12:f2
  2u cpds2:f2
  2u fq=0:f1
  10u
  gosc ph31
  1m do:f2
  lo to 2 times ns
  100m wr #0 if #0 zd
  1m id0
  1m ip2
  1m ip3
  1m ip25*2048
  1m ip26*2048

;Rotor synchronize reconversion block
  1m ip27*6144
  1m ip28*6144
  lo to 2 times td1
exit

;8-step phase cycle
;cycle only DQ reconversion and last 90
;with H spin temperature alternation

ph0 = 0
ph1 = 1 1 1 1 3 3 3 3
ph2 = (8) 0
ph3 = (8) 2
ph25 = (16384) 3511
ph26 = (16384) 12873
ph27 = (16384) 3511 7607 11703 15799
ph28 = (16384) 12873 585 4681 8777
ph4 = 0 1 2 3
ph31 = 0 3 2 1 2 1 0 3

```



## 2Q/1Q/1Q CCC with R18 and PDSB (3D)

```

;cnst31 MAS freq (Hz)
;d0 t1-domain (DQ), in0=dwell time,
; nd0=2 with TPPI
;d10 t2-domain (SQ), in10=dwell time,
; nd10=2 with TPPI

;AQorder:3-2-1
;18 R-elements equal 10 rotor periods
;R-element (90)0-(270)180
;l3=number of R-element pairs
;multiplier for ph13to16=ratio between
; in0 and rotor period times 32768

;90 pulse length for R18-5-5 recoupling
;for R18-10-5 use 5s/36*cnst31
"p11=4s/(36*cnst31)"

"d0=2u"
"d10=2u"
"l7=1"

1 ze
2 d1

90 f2
  2u p12:f2
  p2:f2 ph1

;CP
  2u p15:f1 p16:f2
  (p15 ph2):f1 (p15:spf0 p16 ph0):f2

;2Q excitation
  2u p113:f2
  2u cw:f2
  2u p11:f1
  p1:f1 ph3
  2u p111:f1
3 p11:f1 ph9
  p11*3:f1 ph10
  p11:f1 ph11
  p11*3:f1 ph12
  lo to 3 times l3

; t1-DQ-evolution
  2u p112:f2
  2u cpds2:f2
  d0

;2Q reconversion
  2u p113:f2
  2u cw:f2
4 p11:f1 ph13
  p11*3:f1 ph14
  p11:f1 ph15
  p11*3:f1 ph16
  lo to 4 times l3
  2u p11:f1
  p1:f1 ph5
  2u p112:f2
  2u cpds2:f2

;t2-SQ-evolution time
  d10

;(Proton-driven) spin-diffusion
  1u do:f2
  2u p11:f1
  p1:f1 ph7
  d11
  p1:f1 ph8

;Acquisition
  5u p112:f2
  5u cpds2:f2
  gosc ph31
  1m do:f2
  lo to 2 times ns
  100m wr #0 if #0 zd

;increment all 13C phases
; prior to t2 by 90 degr
; prior to t1 by 45degr/t1-increment
;rotor synchronize reconversion block
; in addition to TPPI
  1m id10
  1m ip2*2
  1m ip3*2
  1m ip9*16384
  1m ip10*16384
  1m ip11*16384
  1m ip12*16384
  1m ip5

```

```

1m ip13*16384
1m ip14*16384
1m ip15*16384
1m ip16*16384
lo to 2 times td2 ;t2-loop
1m rd10 ;reset t2
1m rp2 ;reset phase increments from t2
1m rp3
1m rp5
1m rp9
1m rp10
1m rp11
1m rp12
1m rp13
1m rp14
1m rp15
1m rp16
1m id0
7 1m ip2
1m ip3
10m ip9*8192
10m ip10*8192
10m ip11*8192
10m ip12*8192
10m ip13*8532
10m ip14*8532
10m ip15*8532
10m ip16*8532
lo to 7 times l7
10m iu7 ;increment loop counter
; for TPPI phase incrementation
lo to 1 times td1 ;t1-loop
exit

ph0 = 0 0 1 1 1 1 0 0 1 1 0 0 0 0 1 1
ph1 = 1 1 2 2 2 2 1 1 2 2 1 1 1 1 2 2
ph2 = (8) 0 2 6 0 6 0 0 2 2 4 4 6 4 6 2 4

ph3 = (8) 2 4 0 2 0 2 2 4 4 6 6 0 6 0 4 6

ph9 =(65536) 9102 25486 58254 9102
58254 9102 9102 25486
25486 41870 41870 58254
41870 58254 25486 41870
ph10=(65536) 41870 58254 25486 41870
25486 41870 41870 58254
58254 9102 9102 25486
9102 25486 58254 9102

ph11=(65536) 56433 7281 40049 56433
40049 56433 56433 7281
7281 23665 23665 40049
23665 40049 7281 23665
ph12=(65536) 23665 40049 7281 23665
7281 23665 23665 40049
40049 56433 56433 7281
56433 7281 40049 56433

ph13=(65536) 9102 9102 58254 58254
25486 25486 41870 41870
25486 25486 41870 41870
9102 9102 58254 58254
ph14=(65536) 41870 41870 25486 25486
58254 58254 9102 9102
58254 58254 9102 9102
41870 41870 25486 25486
ph15=(65536) 56433 56433 40049 40049
7281 7281 23665 23665
7281 7281 23665 23665
56433 56433 40049 40049
ph16=(65536) 23665 23665 7281 7281
40049 40049 56433 56433
40049 40049 56433 56433
23665 23665 7281 7281

ph5 = 0 0 3 3 1 1 2 2 1 1 2 2 0 0 3 3
ph7 = 0 0 1 1 1 1 0 0 1 1 0 0 0 0 1 1
ph8 = 0 0 1 1 2 2 1 1 3 3 2 2 3 3 0 0
ph31= 0 2 3 1 2 0 3 1 3 1 0 2 3 1 2 0

```

## List of Symbols, Acronyms, Abbreviations, and Terms

0Q	Zero quantum
1Q	Single quantum
2Q	Double quantum
2QF	Double-quantum filter
<i>A</i>	Spatial tensor
AFA-PLN	(C36A,C41F,C46A)-mutant phospholamban, usually a monomer
AGG	Ala–Gly–Gly tri-peptide
AHT	Average Hamiltonian theory
ARIA	Software for iterative structure calculation with ambiguous restraints
ATP	Adenosine triphosphate
<b><i>B</i></b>	Magnetic field (magnetic flux density)
<b><i>B</i><sub>0</sub></b>	External, static magnetic field
<b><i>B</i><sub>1</sub></b>	RF-generated magnetic field, strength in order of kHz
BMRB	BioMagResBank, repository for NMR data of Proteins, peptides and nucleic acids
CC	( <sup>13</sup> C, <sup>13</sup> C) correlation
CCC	( <sup>13</sup> C, <sup>13</sup> C, <sup>13</sup> C) correlation
CD	Molecular dynamics in Cartesian coordinates
CGM	Conjugate-gradient minimization
CHHC	( <sup>13</sup> C, <sup>13</sup> C)-encoded ( <sup>1</sup> H, <sup>1</sup> H) correlation sensitive to <sup>1</sup> H– <sup>1</sup> H distance
<i>CN<sub>n</sub><sup>ν</sup></i>	Pulse sequence class based on symmetry design principles
CNS	Crystallography and NMR System, simulation software
CP	Cross-polarization, a dipolar-coupling based heteronuclear transfer method
CS	Chemical shielding
CSA	Chemical shielding anisotropy
CW	Continuous wave (irradiation)
<i>D</i>	Dipolar coupling tensor (through-space coupling)
<i>D(R)</i>	Matrix representing a rotation operator
DARR	Dipolar-assisted rotational resonance
DEC	Decoupling
<i>δ</i>	Tensor anisotropy; chemical shift
$\Delta\delta$	Secondary chemical shift
$\Delta\Delta\delta$	Combined secondary chemical shifts of several nuclei (e.g., C <sup>α</sup> and C <sup>β</sup> )
DFT	Density functional theory
DNP	Dynamic nuclear polarization

DOPC	Dioleoyl-phosphocholine lipids
DRAMA	Dipolar recovery at the magic angle
DSS	Shift reference compound (2,2-Dimethyl-2-silapentane-5-sulfonic acid)
$E$	Energy
$E_{\text{angl}}$	Bond angle variation energy function
$E_{\text{bond}}$	Bond-stretching energy function
$E_{\text{chem}}$	Potential energy function representing chemical restraints
$E_{\text{dihe}}$	Dihedral energy function, for torsion angles
$E_{\text{elec}}$	Electrostatic potential energy function
$E_{\text{impr}}$	Improper dihedral energy function, e.g. for planarity constraints
$E_{\text{LJ}}$	Lennard–Jones attraction–repulsion energy function
$E_{\text{rep}}$	Repulsive energy function
$E_{\text{restr}}$	Restraining potential energy function for experimentally derived data
<i>E. coli</i>	<i>Escherichia coli</i> , a bacterium
EPR	Electron paramagnetic resonance
$\eta$	Tensor asymmetry
$F$	A general Tensor
FFT	Fast Fourier transform
FID	Free induction decay
FSLG	Frequency-switched Lee-Goldburg, a decoupling condition
FT	Fourier transform, usually FFT
$\gamma$	Gyromagnetic ratio
GARP	Globally optimized alternating-phase rectangular pulse, a decoupling scheme
GPCR	G-protein coupled receptor
$H$	Hamilton operator
$H_0$	Time-independent part of a Hamiltonian
HADDOCK	High ambiguity driven biomolecular docking based on biochemical and/or biophysical information, a docking software
$H_D$	Magnetic dipole–dipole coupling Hamiltonian
HETCOR	Heteronuclear correlation spectroscopy
HORROR	Homonuclear rotary resonance
$H_{\text{RF}}$	RF interaction Hamiltonian
$H_Z$	Zeeman interaction Hamiltonian
$I$	(Nuclear) spin
$I_x, I_y, I_z$	(Nuclear) spin components
INEPT	Insensitive nuclei enhanced by polarization transfer, a $J$ -based transfer
$J$	A general angular momentum operator
$J$	$J$ coupling tensor (through-bond coupling)
KTX	Kalioxin
LAB	Laboratory frame
LM	Longitudinal mixing
MA	Magic angle
MAS	Magic angle spinning
MD	Molecular dynamics

MDSA	Molecular dynamics simulated annealing
MolMol	Molecular visualization and analysis software
MPD	2-methylpentane-2,4-diol
NADH	Nicotinamide dehydrogenase
NCACX	Combined intra-residue N–C <sup>α</sup> and C <sup>α</sup> –C transfer or correlation.
NC	( <sup>15</sup> N, <sup>13</sup> C) correlation or transfer
NCC	Combined <sup>15</sup> N– <sup>13</sup> C and <sup>13</sup> C– <sup>13</sup> C transfer or correlation
NCO	Sequential N–C' transfer or correlation.
NCOCACX	Combined sequential N–C' and C'–C transfer or correlation.
NCX	Na <sup>+</sup> /Ca <sup>2+</sup> exchanger
NHHC	( <sup>15</sup> N, <sup>13</sup> C)-encoded ( <sup>1</sup> H, <sup>1</sup> H) correlation sensitive to <sup>1</sup> H– <sup>1</sup> H distance
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
$\omega$	Resonance frequency
$\omega_0$	Nuclear Larmor frequency of Zeeman interaction
$\omega_1$	Nuclear Larmor frequency of RF interaction
$\Omega$	Resonance frequency offset with respect to carrier frequency
$\omega_R$	Sample spinning frequency (unit: angular frequency)
PDB	Protein data bank
PDSF	Proton-driven spin diffusion
PAS	Principle axis system of a tensor
PEG	Polyethylene glycol
$\phi$	An angle, often phase angle, or backbone torsion angle between N and C <sup>α</sup>
PKA	Protein kinase A
PLN	Phospholamban, wild-type usually being a pentamer
PMCA	Plasma-membrane Ca <sup>2+</sup> -ATPase
POST	Composite pulse (90°) <sub>0°</sub> –(360°) <sub>180°</sub> –(270°) <sub>0°</sub> (subscript denoting phase)
ppm	Parts per million (here: relative unit of nuclear chemical shift)
PREDITOR	Software prediction backbone torsion angles from chemical shifts
$\psi$	Backbone torsion angle between C <sup>α</sup> and C', or state function
PyMOL	Molecular visualization and analysis software
$R$	Rotation operator
RF	Radio frequency
REDOR	Rotational echo double resonance
$\rho$	Density operator
RMSD	Root-mean-square deviation
$RN_n^\nu$	Pulse sequence class based on symmetry design principles
R <sup>2</sup>	Rotational resonance
R <sup>2</sup> TR	Rotational resonance in the tilted rotating frame
RyR	Ryanodine receptor
$S$	Scaling factor, order parameter
SA	Simulated annealing
SD	Spin diffusion

SDWC	Spin diffusion under weak coupling conditions
SERCA	Sarco(endo)plasmic reticulum $\text{Ca}^{2+}$ -ATPase
SHIFTS	Chemical shift prediction software
SHIFTX	Chemical shift prediction software
$\sigma$	Chemical shielding tensor
$\sigma_{\text{iso}}$	Isotropic average of (chemical shift) tensor
$S/N$	Signal-to-noise ratio
Sparky	NMR assignment software
SPARTA	Chemical shift prediction software
SPC5	Supercycled POST-C5
SPECIFIC-CP	Spectrally induced filtering in combination with cross-polarization
SR	Sarcoplasmic reticulum
ssNMR	Solid-state NMR
$t$	Time
$T$	A spin tensor
$T_1$	Spin-lattice relaxation time constant
$T_2$	Spin-spin relaxation time constant
TAD	Torsion-angle dynamics
TALOS	Software for predicting backbone torsion angles from chemical shifts
TCA	Citric or Tricarboxylic Acid cycle
TEE	Tyrosine-ethylester
$\theta$	An angle, usually between dipolar coupling vector and $\mathbf{B}_0$
$\theta_M$	Magic angle
TIP3P	Parameter set used for water simulation
TMS	Tetramethyl silane, a shift reference compound
TOBSY	Total through-bond correlation spectroscopy
Topspin	NMR spectrometer software
TPPM	Two-Pulse Phase Modulation
$U$	Quantum mechanical propagator
U(X)	Uniformly labeled with isotope X
U(X\A)	U(X) with reverse labeling of amino acid type A
UBI	Ubiquitin
UBI-H	Ubiquitin, rehydrated after lyophilization
UBI-M	Ubiquitin microcrystals from MPD-precipitation
UBI-P	Ubiquitin microcrystals from PEG-precipitation
$V$	Time-dependent part of a Hamiltonian
WT	Wild-type
$\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{n}$	Space vectors (denoted by bold print)
$\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}, \hat{\mathbf{n}}$	Unit space vectors (denoted by bold print and hat)
XWinNMR	NMR spectrometer software

## Bibliography

- [1] C. Ader, R. Schneider, K. Seidel, M. Etzkorn, and M. Baldus. Magic-angle-spinning NMR spectroscopy applied to small molecules and peptides in lipid bilayers. *Biochem. Soc. Trans.*, 35:991–995, 2007. DOI 10.1042/BST0350991.
- [2] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 1987.
- [3] I. Ando, H. Saito, R. Tabeta, A. Shoji, and T. Ozaki. Conformation-dependent carbon-13 NMR chemical shifts of poly(L-alanine) in the solid state: FPT INDO calculation of N-acetyl-N'-methyl-L-alanine amide as a model compound of poly(L-alanine). *Macromolecules*, 17: 457–461, 1984. DOI 10.1021/ma00133a036.
- [4] E. R. Andrew, A. Bradbury, and R. G. Eades. Nuclear Magnetic Resonance Spectra from a Crystal Rotated at High Speed. *Nature*, 182:1659, 1958. DOI 10.1038/1821659a0.
- [5] E. R. Andrew, A. Bradbury, and R. G. Eades. Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation. *Nature*, 183:1802–1803, 1959. DOI 10.1038/1831802a0.
- [6] E. R. Andrew, A. Bradbury, R. G. Eades, and V. T. Wynn. Nuclear Cross-Relaxation Induced by Specimen Rotation. *Phys. Lett.*, 4(2):99–100, 1963. DOI 10.1016/0031-9163(63)90123-9.
- [7] E. R. Andrew, S. Clough, L. F. Farnell, T. D. Gledhill, and I. Roberts. Resonant Rotational Broadening of Nuclear Magnetic Resonance Spectra. *Phys. Lett.*, 21(5):505–506, 1966. DOI 10.1016/0031-9163(66)91274-1.
- [8] O. C. Andronesi. *Solid-state NMR of (membrane) protein complexes: Novel methods and applications*. PhD thesis, Universität Göttingen, Germany, 2006. URL <http://resolver.sub.uni-goettingen.de/purl/?webdoc-759>.
- [9] O. C. Andronesi, S. Becker, K. Seidel, H. Heise, H. S. Young, and M. Baldus. Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. *J. Am. Chem. Soc.*, 127(37):12965–12974, Sep 2005. DOI 10.1021/ja0530164.
- [10] O. C. Andronesi, M. von Bergen, J. Biernat, K. Seidel, C. Griesinger, E. Mandelkow, and M. Baldus. Characterization of Alzheimer's-like paired helical filaments from the core domain of tau protein using solid-state NMR spectroscopy. *J. Am. Chem. Soc.*, 130(18):5922–5928, May 2008. DOI 10.1021/ja7100517.

- [11] C. B. Anfinsen. Principles that govern the folding of protein chains. *Science*, 181(96):223–230, Jul 1973.
- [12] C. B. Anfinsen, E. Haber, M. Sela, and F. H. White. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Ac. Sci. USA*, 47:1309–1314, Sep 1961.
- [13] M. Asahi, Y. Kimura, K. Kurzydowski, M. Tada, and D. H. MacLennan. Transmembrane helix M6 in sarco(endo)plasmic reticulum  $\text{Ca}^{2+}$ -ATPase forms a functional interaction site with phospholamban. Evidence for physical interactions at other sites. *J. Biol. Chem.*, 274(46):32855–32862, Nov 1999. URL <http://www.jbc.org/cgi/content/abstract/274/46/32855>.
- [14] D. S. Bae and E. J. Haug. A recursive formulation for constrained mechanical system dynamics: Part I. Open loop systems. *Mech. Struct. Mech.*, 15:359–382, 1987.
- [15] D. S. Bae and E. J. Haug. A recursive formulation for constrained mechanical system dynamics: Part II. Closed loop systems. *Mech. Struct. Mech.*, 15:481–506, 1988.
- [16] A. D. Bain. Coherence levels and coherence pathways in NMR. A simple way to design phase cycling procedures. *J. Magn. Reson.*, 56:418, 1984. DOI 10.1016/0022-2364(84)90305-6.
- [17] M. Bak, J. T. Rasmussen, and N. C. Nielsen. SIMPSON: A General Simulation Program for Solid-State NMR Spectroscopy. *J. Magn. Reson.*, 147:296–330, 2000. DOI 10.1006/jmre.2000.2179.
- [18] M. Baldus. Molecular interactions investigated by multi-dimensional solid-state NMR. *Curr. Op. Struct. Biol.*, 16(5):618–623, Oct 2006. DOI 10.1016/j.sbi.2006.08.003.
- [19] M. Baldus. Magnetic resonance in the solid state: applications to protein folding, amyloid fibrils and membrane proteins. *Eur. Biophys. J.*, 36(Suppl. 1):37–48, May 2007. DOI 10.1007/s00249-007-0174-y.
- [20] M. Baldus. ICMRBS founder’s medal 2006: biological solid-state NMR, methods and applications. *J. Biomol. NMR*, 39(1):73–86, Sep 2007. DOI 10.1007/s10858-007-9177-3.
- [21] M. Baldus. Correlation experiments for assignment and structure elucidation of immobilized polypeptides under magic angle spinning. *Progr. NMR Spectr.*, 41(1–2):1–47, Sep 2002. DOI 10.1016/S0079-6565(02)00007-9.
- [22] M. Baldus and B. H. Meier. Broadband polarization transfer under magic-angle spinning: Application to total through-space-correlation NMR spectroscopy. *J. Magn. Reson.*, 128(2):172–193, Oct 1997. DOI 10.1006/jmre.1997.1219.
- [23] M. Baldus and B. H. Meier. Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity. *J. Magn. Reson. A*, 121(1):65–69, Jul 1996. DOI 10.1006/jmra.1996.0137.



- [24] M. Baldus, M. Tomaselli, B. H. Meier, and R. R. Ernst. Broad-Band Polarization-Transfer Experiments for Rotating Solids. *Chem. Phys. Lett.*, 230(4–5):329–336, Dec 1994. DOI 10.1016/0009-2614(94)01209-1.
- [25] M. Baldus, A. T. Petkova, J. Herzfeld, and R. G. Griffin. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.*, 95(6):1197–1207, Dec 1998. DOI 10.1080/002689798166215.
- [26] M. Bechmann, X. Helluy, and A. Sebald. Selectivity of Double-Quantum Filtered Rotational-Resonance Experiments on Larger-than-Two-Spin Systems. In S. R. Kiihne and H. J. M. de Groot, editors, *Perspectives on Solid State NMR in Biology*, volume 1 of *Focus on Structural Biology*, pages 23–31. Kluwer Academic Publishers, 2001. ISBN 978-0-7923-7102-1.
- [27] D. A. Bender. *Amino Acid Metabolism*. John Wiley & Sons, 1985. ISBN 047106498X.
- [28] A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi, and R. G. Griffin. Heteronuclear Decoupling in Rotating Solids. *J. Chem. Phys.*, 103(16):6951–6958, Oct 1995. DOI 10.1063/1.470372.
- [29] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81:3684–3690, 1984. DOI 10.1063/1.448118.
- [30] M. V. Berjanskii, S. Neal, and D. S. Wishart. PREDITOR: a web server for predicting protein torsion angle restraints. *Nucl. Acids Res.*, 34(suppl 2):W63–69, July 1, 2006 2006. DOI 10.1093/nar/gkl341.
- [31] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The Protein Data Bank. *Nucl. Acids Res.*, 28(1):235–242, January 1, 2000 2000. URL <http://nar.oxfordjournals.org/cgi/content/abstract/28/1/235>.
- [32] A. Bielecki, A. C. Kolbert, and M. H. Levitt. Frequency-Switched Pulse Sequences – Homonuclear Decoupling and Dilute Spin NMR in Solids. *Chem. Phys. Lett.*, 155(4-5): 341–346, Mar 1989.
- [33] F. Bloch. Nuclear Induction. *Phys. Rev.*, 70(7–8):460–474, Oct 1946. DOI 10.1103/PhysRev.70.460.
- [34] F. Bloch, W. W. Hansen, and M. Packard. Nuclear Induction. *Phys. Rev.*, 69(3-4):127, Feb 1946. DOI 10.1103/PhysRev.69.127.
- [35] N. Bloembergen. On the Interaction of Nuclear Spins in a Crystalline Lattice. *Physica*, 15(3–4):386–426, 1949. DOI 10.1016/0031-8914(49)90114-7.
- [36] G. Bodenhausen, H. Kogler, and R. R. Ernst. Selection of Coherence-Transfer Pathways in NMR Pulse Experiments. *J. Magn. Reson.*, 58(3):370–388, 1984. DOI 10.1016/0022-2364(84)90142-2.

- [37] A. M. J. J. Bonvin. Flexible protein–protein docking. *Curr. Op. Struct. Biol.*, 16:194–200, 2006. DOI 10.1016/j.sbi.2006.02.002.
- [38] J. U. Bowie. Solving the membrane protein folding problem. *Nature*, 438:581–589, Dec 2005. DOI 10.1038/nature04395.
- [39] C. Branden and J. Tooze. *Introduction to Protein Structure*. Garland Publishing, 2nd edition, 1998.
- [40] S. P. Brown and H. W. Spiess. Advanced solid-state NMR methods for the elucidation of structure and dynamics of molecular, macromolecular, and supramolecular systems. *Chem. Rev.*, 101(12):4125–4156, Dec 2001. DOI 10.1021/cr990132e.
- [41] A. T. Brunger. *Molecular Dynamics: Applications in Molecular Biology*, chapter 5: Refinement of three-dimensional structures of proteins and nucleic acids, pages 137–178. Topics in Molecular and Structural Biology. Macmillan Press, 1991. ISBN 978-0333498866.
- [42] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst. D*, 54:905–921, Sep 1998. DOI 10.1107/S0907444998003254.
- [43] B. Brutscher, R. Bruschweiler, and R. R. Ernst. Backbone dynamics and structural characterization of the partially folded A state of ubiquitin by H-1, C-13, and N-15 nuclear magnetic resonance spectroscopy. *Biochemistry*, 36(42):13043–13053, Oct 1997. DOI 10.1021/bi971538t.
- [44] A. Böckmann, A. Lange, A. Galinier, S. Luca, N. Giraud, M. Juy, H. Heise, R. Montserret, F. Penin, and M. Baldus. Solid state NMR sequential resonance assignments and conformational analysis of the 2x10.4 kDa dimeric form of the Bacillus subtilis protein Crh. *J. Biomol. NMR*, 27(4):323–339, Dec 2003. DOI 10.1023/A:1025820611009.
- [45] P. Caravatti, L. Braunschweiler, and R. R. Ernst. Heteronuclear Correlation Spectroscopy in Rotating Solids. *Chem. Phys. Lett.*, 100(4):305–310, 1983. DOI 10.1016/0009-2614(83)80276-0.
- [46] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, and H. Oschkinat. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature*, 420(6911):98–102, Nov 2002. DOI 10.1038/nature01070.
- [47] F. Castellani, B.-J. van Rossum, A. Diehl, K. Rehbein, and H. Oschkinat. Determination of solid-state NMR structures of proteins by means of three-dimensional  $^{15}\text{N}$ - $^{13}\text{C}$ - $^{13}\text{C}$  dipolar correlation spectroscopy and chemical shift analysis. *Biochemistry*, 42(39):11476–11483, Oct 2003. DOI 10.1021/bi034903r.
- [48] A. Cavalli, X. Salvatella, C. M. Dobson, and M. Vendruscolo. Protein structure determination from NMR chemical shifts. *Proc. Natl. Ac. Sci. USA*, 104(23):9615–9620, Jun 2007. DOI 10.1073/pnas.0610313104.

- [49] J. Cavanagh, W. J. Fairbrother, A. G. I. Palmer, M. Rance, and N. J. Skelton. *Protein NMR Spectroscopy*. Elsevier, 2007.
- [50] J. C. C. Chan and R. Tycko. Broadband rotational resonance in solid state NMR spectroscopy. *J. Chem. Phys.*, 120(18):8349–8352, May 2004. DOI 10.1063/1.1737369.
- [51] Z. H. Chen, S. L. Stokes, W. J. Rice, and L. R. Jones. Spatial and dynamic interactions between phospholamban and the canine cardiac  $\text{Ca}^{2+}$  pump revealed with use of heterobifunctional cross-linking agents. *J. Biol. Chem.*, 278(48):48348–48356, Nov 2003. DOI 10.1074/jbc.M309545200.
- [52] Z. H. Chen, D. L. Stokes, and L. R. Jones. Role of Leucine 31 of Phospholamban in Structural and Functional Interactions with the  $\text{Ca}^{2+}$  Pump of Cardiac Sarcoplasmic Reticulum. *J. Biol. Chem.*, 280(11):10530–10539, Mar 2005. DOI 10.1074/jbc.M414007200.
- [53] Z. H. Chen, B. L. Akin, D. L. Stokes, and L. R. Jones. Cross-linking of C-terminal residues of phospholamban to the  $\text{Ca}^{2+}$  pump of cardiac sarcoplasmic reticulum to probe spatial and functional interactions within the transmembrane domain. *J. Biol. Chem.*, 281(20):14163–14172, May 2006. DOI 10.1074/jbc.M601338200.
- [54] V. Chevelkov, K. Faelber, A. Diehl, U. Heinemann, H. Oschkinat, and B. Reif. Detection of dynamic water molecules in a microcrystalline sample of the SH3 domain of  $\alpha$ -spectrin by MAS solid-state NMR. *J. Biomol. NMR*, 31(4):295–310, Apr 2005. DOI 10.1007/s10858-005-1718-z.
- [55] H. B. R. Cole and D. A. Torchia. An NMR-Study of the Backbone Dynamics of Staphylococcal Nuclease in the Crystalline State. *Chem. Phys.*, 158(2–3):271–281, Dec 1991. DOI 10.1016/0301-0104(91)87071-3.
- [56] M. G. Colombo, B. H. Meier, and R. R. Ernst. Rotor-driven spin diffusion in natural-abundance  $^{13}\text{C}$  spin systems. *Chem. Phys. Lett.*, 146:189, 1988. DOI 10.1016/0009-2614(88)87429-3.
- [57] G. Cornilescu, J. L. Marquardt, M. Ottiger, and A. Bax. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.*, 120(27):6836–6837, Jul 1998. DOI 10.1021/ja9812610.
- [58] G. Cornilescu, F. Delaglio, and A. Bax. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, 13(3):289–302, Mar 1999. DOI 10.1007/s10858-007-9166-6.
- [59] P. R. Costa, B. Q. Sun, and R. G. Griffin. Rotational resonance tickling: Accurate internuclear distance measurement in solids. *J. Am. Chem. Soc.*, 119(44):10821–10830, Nov 1997. DOI 10.1021/ja9812610.
- [60] A. F. L. Creemers, S. Kiihne, P. H. M. Bovee-Geurts, W. J. DeGrip, J. Lugtenburg, and H. J. M. de Groot. H-1 and C-13 MAS NMR evidence for pronounced ligand-protein interactions involving the ionone ring of the retinylidene chromophore in rhodopsin. *Proc. Natl. Ac. Sci. USA*, 99(14):9101–9106, Jul 2002. DOI 10.1073/pnas.112677599.

- [61] T. A. Cross and S. J. Opella. Solid-State NMR Structural Studies of Peptides and Proteins in Membranes. *Curr. Op. Struct. Biol.*, 4(4):574–581, Aug 1994. DOI 10.1073/pnas.112677599.
- [62] I. de Boer, J. Matysik, M. Amakawa, S. Yagai, H. Tamiaki, A. R. Holzwarth, and H. J. M. de Groot. MAS NMR structure of a microcrystalline Cd-bacteriochlorophyll d analogue. *J. Am. Chem. Soc.*, 125(44):13374–13375, Nov 2003. DOI 10.1021/ja0367492.
- [63] A. C. de Dios, J. G. Pearson, and E. Oldfield. Secondary and Tertiary Structural Effects on Protein Nmr Chemical-Shifts - an Abinitio Approach. *Science*, 260(5113):1491–1496, Jun 1993. DOI 10.1126/science.8502992.
- [64] J. Deisenhofer, O. Epp, K. Miki, R. Huber, and H. Michel. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3 Å resolution. *Nature*, 318:618–624, 1985. DOI 10.1038/318618a0.
- [65] DeLano Scientific. PyMOL. <http://pymol.sourceforge.net/>. K. S. acknowledges a free student licence for pre-compiled PyMOL.
- [66] B. M. Denker, B. L. Smith, F. P. Kuhajda, and P. Agre. Identification, purification, and partial characterization of a novel Mr 28,000 integral membrane protein from erythrocytes and renal tubules. *J. Biol. Chem.*, 263(30):15634–15642, Oct 1988. URL <http://www.jbc.org/cgi/content/abstract/263/30/15634>.
- [67] A. Detken, E. H. Hardy, M. Ernst, M. Kainosho, T. Kawakami, S. Aimoto, and B. H. Meier. Methods for sequential resonance assignment in solid, uniformly C-13, N-15 labelled peptides: Quantification and application to antamanide. *J. Biomol. NMR*, 20(3):203–221, Jul 2001. DOI 10.1023/A:1011212100630.
- [68] W. C. Dickinson. Dependence of the  $F^{19}$  Nuclear Resonance Position on Chemical Compound. *Phys. Rev.*, 77(5):736–737, Mar 1950. DOI 10.1103/PhysRev.77.736.2.
- [69] C. M. Dobson. Protein misfolding, evolution and disease. *Trends Biochem. Sci.*, 24(9):329–332, 1999. DOI 10.1016/S0968-0004(99)01445-0.
- [70] C. Dominguez, R. Boelens, and A. Bonvin. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.*, 125(7):1731–1737, 2003. DOI 10.1021/ja026939x. URL <http://www.nmr.chem.uu.nl/haddock/>.
- [71] J. L. Douglas, C. A. Trieber, M. Afara, and H. S. Young. Rapid, high-yield expression and purification of  $Ca^{2+}$ -ATPase regulatory proteins for high-resolution structural studies. *Prot. Expr. Purif.*, 40(1):118–125, Mar 2005. DOI 10.1016/j.pep.2004.11.015.
- [72] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. MacKinnon. The structure of the potassium channel: molecular basis of  $K^+$  conduction and selectivity. *Science*, 280(5360):69–77, Apr 1998. DOI 10.1126/science.280.5360.69.
- [73] S. Dusold and A. Sebald. Dipolar recoupling under magic-angle spinning conditions. volume 41 of *Annual Reports on NMR Spectroscopy*, pages 185–264. Academic Press, San Diego, 2000. ISBN 0066-4103.

- [74] S. Dusold and A. Sebald. Double-quantum filtration under rotational-resonance conditions: numerical simulations and experimental results. *J. Magn. Reson.*, 145(2):340–356, Aug 2000. DOI 10.1006/jmre.2000.2118.
- [75] T. A. Egorova-Zachernyuk, J. Hollander, N. Fraser, P. Gast, A. J. Hoff, R. Cogdell, H. J. de Groot, and M. Baldus. Heteronuclear 2D-correlations in a uniformly [ $^{13}\text{C}$ ,  $^{15}\text{N}$ ] labeled membrane-protein complex at ultra-high magnetic fields. *J. Biomol. NMR*, 19(3):243–253, Mar 2001. DOI 10.1023/A:1011235417465.
- [76] S. Eletr and G. Inesi. Phospholipid orientation in sarcoplasmic membranes: spin-label ESR and proton NMR studies. *Biochim. Biophys. Acta*, 282(1):174–179, Sep 1972.
- [77] R. R. Ernst and W. A. Anderson. Application of Fourier Transform Spectroscopy to Magnetic Resonance. *Rev. Sci. Instr.*, 37:93–102, 1966. DOI 10.1063/1.1719961.
- [78] R. R. Ernst, G. Bodenhausen, and A. Wokaun. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford University Press, 1st edition, 1987. ISBN 0198556470.
- [79] M. Etzkorn. Master’s thesis, Universität Göttingen, Department of Physics, 2004.
- [80] M. Etzkorn, A. Böckmann, A. Lange, and M. Baldus. Probing molecular interfaces using 2D magic-angle-spinning NMR on protein mixtures with different uniform labeling. *J. Am. Chem. Soc.*, 126(45):14746–14751, Nov 2004. DOI 10.1021/ja0479181.
- [81] M. Etzkorn, S. Martell, O. C. Andronesi, K. Seidel, M. Engelhard, and M. Baldus. Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.*, 46(3):459–462, 2007. DOI 10.1002/anie.200602139.
- [82] W. T. Franks, D. H. Zhou, B. J. Wylie, B. G. Money, D. T. Graesser, H. L. Frericks, G. Sahota, and C. M. Rienstra. Magic-angle spinning solid-state NMR spectroscopy of the beta 1 immunoglobulin binding domain of protein G (GB1): N-15 and C-13 chemical shift assignments and conformational analysis. *J. Am. Chem. Soc.*, 127(35):12291–12305, Sep 2005. DOI 10.1021/ja044497e.
- [83] W. T. Franks, B. J. Wylie, S. A. Stellfox, and C. M. Rienstra. Backbone conformational constraints in a microcrystalline  $\text{U-}^{15}\text{N}$ -labeled protein by 3D dipolar-shift solid-state NMR spectroscopy. *J. Am. Chem. Soc.*, 128(10):3154–3155, Mar 2006. DOI 10.1021/ja058292x.
- [84] C. M. Fraser, J. D. Gocayne, O. White, et al. The minimal gene complement of mycoplasma-genitalium. *Science*, 270:397–403, 1995. DOI 10.1126/science.270.5235.397.
- [85] M. H. Frey, J. A. Diverdi, and S. J. Opella. Dynamics of Phenylalanine in the Solid-State by NMR. *J. Am. Chem. Soc.*, 107(25):7311–7315, 1985. DOI 10.1021/ja00311a016.
- [86] M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al. *Gaussian 03, Revision C.02*. Gaussian, Inc., Wallingford, CT, 2004.

- [87] T. Fujiwara, Y. Todokoro, H. Yanagishita, M. Tawarayama, T. Kohno, K. Wakamatsu, and H. Akutsu. Signal assignments and chemical-shift structural analysis of uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled peptide, mastoparan-X, by multidimensional solid-state NMR under magic-angle spinning. *J. Biomol. NMR*, 28(4):311–325, Apr 2004. DOI 10.1023/B:JNMR.0000015377.17021.b0.
- [88] B. M. Fung, A. K. Khitrin, and K. Ermolaev. An improved broadband decoupling sequence for liquid crystals and solids. *J. Magn. Reson.*, 142(1):97–101, Jan 2000. DOI 10.1006/jmre.1999.1896.
- [89] M. Gairi, R. Romi, I. Fernandez, H. Rochat, M. F. Martin-Eauclaire, J. Van Rietschoten, M. Pons, and E. Giralt. 3D structure of kaliotoxin: Is residue 34 a key for channel selectivity? *J. Pept. Sci.*, 3(4):314–319, Jul-Aug 1997. DOI 10.1002/(SICI)1099-1387(199707)3:4<314::AID-PSC117>3.0.CO;2-E.
- [90] N. Giraud, A. Bockmann, A. Lesage, F. Penin, M. Blackledge, and L. Emsley. Site-specific backbone dynamics from a crystalline protein by solid-state NMR spectroscopy. *J. Am. Chem. Soc.*, 126(37):11422–11423, Sep 2004. DOI 10.1021/ja046578g.
- [91] J. Glushka, M. Lee, S. Coffin, and D. Cowburn. Nitrogen-15 chemical shifts of backbone amides in bovine pancreatic trypsin inhibitor and apamin. *J. Am. Chem. Soc.*, 111(20):7716–7722, 1989.
- [92] T. D. Goddard and D. G. Kneller. SPARKY 3. URL <http://www.cgl.ucsf.edu/home/sparky/>. University of California, San Francisco.
- [93] V. I. Gordeliy, J. Labahn, R. Moukhametzianov, R. Efremov, J. Granzin, R. Schlesinger, G. Buldt, T. Savopol, A. J. Scheidig, J. P. Klare, and M. Engelhard. Molecular basis of transmembrane signalling by sensory rhodopsin II–transducer complex. *Nature*, 419(6906):484–487, Oct 2002. DOI 10.1038/nature01109.
- [94] R. B. Gregory, M. Gangoda, R. K. Gilpin, and W. Su. The Influence of Hydration on the Conformation of Lysozyme Studied by Solid-State C-13-NMR Spectroscopy. *Biopolymers*, 33(4):513–519, Apr 1993. DOI 10.1002/bip.360331212.
- [95] R. G. Griffin. Dipolar recoupling in MAS spectra of biological solids. *Nature Struct. Biol.*, 5:508–512, Jul 1998. DOI 10.1038/749.
- [96] J. M. Griffiths and R. G. Griffin. Nuclear-Magnetic-Resonance Methods for Measuring Dipolar Couplings in Rotating Solids. *Analyt. Chim. Acta*, 283(3):1081–1101, Dec 1993.
- [97] G. Grobner, I. J. Burnett, C. Glaubitz, G. Choi, A. J. Mason, and A. Watts. Observations of light-induced structural changes of retinal within rhodopsin. *Nature*, 405(6788):810–813, Jun 2000. DOI 10.1038/35015604.
- [98] A. Grommek, B. H. Meier, and M. Ernst. Distance information from proton-driven spin diffusion under MAS. *Chem. Phys. Lett.*, 427:404–409, 2006.

- [99] J. D. Gross, D. E. Warschawski, and R. G. Griffin. Dipolar recoupling in MAS NMR: A probe for segmental order in lipid bilayers. *J. Am. Chem. Soc.*, 119(4):796–802, Jan 1997. DOI 10.1021/ja962951b.
- [100] N. Guex and M. Peitsch. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, 18:2714–2723, 1997. DOI 10.1002/elps.1150181505. URL <http://www.expasy.org/spdbv/>.
- [101] T. Gullion and J. Schaefer. Rotational-echo double-resonance NMR. *J. Magn. Reson.*, 81(1):196–200, Jan 1989. DOI 10.1016/0022-2364(89)90280-1.
- [102] P. Güntert. Automated NMR protein structure calculation. *Progr. NMR Spectr.*, 43(3–4):105–125, Dec 2003. DOI 10.1016/S0079-6565(03)00021-9.
- [103] P. Güntert. Structure Calculation of Biological Macromolecules from NMR Data. *Q. Rev. Biophys.*, 31:145–237, 1998.
- [104] P. Güntert, C. Mumenthaler, and K. Wüthrich. Torsion Angle Dynamics for NMR Structure Calculation with the New Program DYANA. *J. Mol. Biol.*, 273:283–298, 1997. DOI 10.1006/jmbi.1997.1284.
- [105] U. Haeberlen and J. S. Waugh. Coherent Averaging Effects in Magnetic Resonance. *Phys. Rev.*, 175:453–467, 1968. DOI 10.1103/PhysRev.175.453.
- [106] E. L. Hahn. Nuclear Induction Due to Free Larmor Precession. *Phys. Rev.*, 77(2):297–298, Jan 1950. DOI 10.1103/PhysRev.77.297.2.
- [107] R. K. Harris. Encyclopedia of Nuclear Magnetic Resonance, pages 3734–3740. John Wiley & Sons, 1996.
- [108] R. K. Harris, E. D. Becker, S. M. C. De Menezes, R. Goodfellow, and P. Granger. NMR nomenclature. Nuclear spin properties and conventions for chemical shifts (IUPAC recommendations 2001). *Pure Appl. Chem.*, 73:1795–1818, 2001. DOI 10.1351/pac200173111795.
- [109] S. R. Hartmann and E. L. Hahn. Nuclear Double Resonance in Rotating Frame. *Phys. Rev.*, 128(5):2042–2053, 1962. DOI 10.1103/PhysRev.128.2042.
- [110] R. H. Havlin and R. Tycko. Probing site-specific conformational distributions in protein folding with solid-state NMR. *Proc. Natl. Ac. Sci. USA*, 102(9):3284–3289, Mar 2005. DOI 10.1073/pnas.0406130102.
- [111] H. Heise, W. Hoyer, S. Becker, O. C. Andronesi, D. Riedel, and M. Baldus. Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Ac. Sci. USA*, 102(44):15871–15876, Nov 2005. DOI 10.1073/pnas.0506109102.
- [112] H. Heise, S. Luca, B. L. de Groot, H. Grubmüller, and M. Baldus. Probing conformational disorder in neurotensin by two-dimensional solid-state NMR and comparison to molecular dynamics simulations. *Biophys. J.*, 89(3):2113–2120, Sep 2005. DOI 10.1529/biophysj.105.059964.

- [113] H. Heise, K. Seidel, M. Etzkorn, S. Becker, and M. Baldus. 3D NMR spectroscopy for resonance assignment and structure elucidation of proteins under MAS: novel pulse schemes and sensitivity considerations. *J. Magn. Reson.*, 173(1):64–74, Mar 2005. DOI 10.1016/j.jmr.2004.11.020.
- [114] X. Helluy and A. Sebald. Structure and dynamic properties of solid L-tyrosine-ethylester as seen by C-13 MAS NMR. *J. Phys. Chem. B*, 107(14):3290–3296, Apr 2003. DOI 10.1021/jp0277210.
- [115] P. Hodgkinson and L. Emsley. The accuracy of distance measurements in solid-state NMR. *J. Magn. Reson.*, 139(1):46–59, Jul 1999. DOI 10.1006/jmre.1999.1759.
- [116] M. Hohwy, H. J. Jakobsen, M. Eden, M. H. Levitt, and N. C. Nielsen. Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence. *J. Chem. Phys.*, 108(7):2686–2694, Feb 1998. DOI 10.1063/1.475661.
- [117] M. Hohwy, C. M. Rienstra, C. P. Jaroniec, and R. G. Griffin. Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. *J. Chem. Phys.*, 110(16):7983–7992, Apr 1999.
- [118] M. Hong. Determination of multiple phi-torsion angles in proteins by selective and extensive C-13 labeling and two-dimensional solid-state NMR. *J. Magn. Reson.*, 139(2):389–401, Aug 1999. DOI 10.1006/jmre.1999.1805.
- [119] M. Hong, X. L. Yao, K. Jakes, and D. Huster. Investigation of molecular motions by Lee-Goldburg cross-polarization NMR Spectroscopy. *J. Phys. Chem. B*, 106(29):7355–7364, Jul 2002. DOI 10.1021/jp0156064.
- [120] R. W. W. Hooft, G. Vriend, C. Sander, and E. E. Abola. Errors in protein structures. *Nature*, 381:272, 1996.
- [121] P. Hore, J. Jones, and S. Wimperis. *NMR: The Toolkit*. Oxford Chemistry Primers, 92. Oxford University Press, 2000.
- [122] C. E. Hughes, S. Luca, and M. Baldus. Radio-frequency driven polarization transfer without heteronuclear decoupling in rotating solids. *Chem. Phys. Lett.*, 385(5-6):435–440, Feb 2004.
- [123] E. Hughes and D. A. Middleton. Solid-state NMR reveals structural changes in phospholamban accompanying the functional regulation of Ca<sup>2+</sup>-ATPase. *J. Biol. Chem.*, 278(23):20835–20842, Jun 2003.
- [124] W. Humphrey, A. Dalke, and K. Schulten. VMD – Visual Molecular Dynamics. *J. Mol. Graph.*, 14:33–38, 1996.
- [125] M. C. Hutter, J. Krebs, J. Meiler, C. Griesinger, E. Carafoli, and V. Helms. A structural model of the complex formed by phospholamban and the calcium pump of sarcoplasmic reticulum obtained by molecular mechanics. *ChemBioChem*, 3(12):1200–1208, Dec 2002.



- [126] T. I. Igumenova, A. E. McDermott, K. W. Zilm, R. W. Martin, E. K. Paulson, and A. J. Wand. Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.*, 126(21):6720–6727, Jun 2004. DOI 10.1021/ja030547o.
- [127] T. I. Igumenova, A. J. Wand, and A. E. McDermott. Assignment of the backbone resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.*, 126(16):5323–5331, Apr 2004. DOI 10.1021/ja030546w.
- [128] International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011):931–945, Oct 2004. DOI 10.1038/nature03001.
- [129] M. Iwadate, T. Asakura, and M. P. Williamson. C alpha and C beta carbon-13 chemical shifts in proteins from an empirical database. *J. Biomol. NMR*, 13(3):199–211, Mar 1999.
- [130] A. Jain, N. Vaidehi, and G. Rodriguez. A fast recursive algorithm for molecular dynamics simulation. *J. Comp. Phys.*, 106:258–268, 1993.
- [131] D. L. Jakeman, D. J. Mitchell, W. A. Shuttleworth, and J. N. S. Evans. Effects of sample preparation conditions on biomolecular solid-state NMR lineshapes. *J. Biomol. NMR*, 12(3):417–421, Oct 1998.
- [132] P. James, M. Inui, M. Tada, M. Chiesi, and E. Carafoli. Nature and Site of Phospholamban Regulation of the  $\text{Ca}^{2+}$  Pump of Sarcoplasmic-Reticulum. *Nature*, 342(6245):90–92, Nov 1989. DOI 10.1038/342090a0.
- [133] C. P. Jaroniec, C. E. MacPhee, N. S. Astrof, C. M. Dobson, and R. G. Griffin. Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril. *Proc. Natl. Ac. Sci. USA*, 99(26):16748–16753, Dec 2002. DOI 10.1073/pnas.252625999.
- [134] C. P. Jaroniec, C. E. MacPhee, V. S. Bajaj, M. T. McMahon, C. M. Dobson, and R. G. Griffin. High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl. Ac. Sci. USA*, 101(3):711–716, Jan 2004. DOI 10.1073/pnas.0304849101.
- [135] J. Jeener. Pulse Pair Techniques in High Resolution NMR. In *Ampere International Summer School II, Basko Polje, Yugoslavia*, volume Pulsed Magnetic and Optical Resonance, 1971.
- [136] J. L. Jiménez, J. I. Gujjarro, E. Orlova, J. Zurdo, C. M. Dobson, M. Sunde, and H. R. Saibil. Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *EMBO J.*, 18(4):815–821, Feb 1999. DOI 10.1093/emboj/18.4.815.
- [137] L. R. Jones, R. L. Cornea, and Z. Chen. Close Proximity between Residue 30 of Phospholamban and Cysteine 318 of the Cardiac  $\text{Ca}^{2+}$  Pump Revealed by Intermolecular Thiol Cross-linking. *J. Biol. Chem.*, 277(31):28319–28329, Jul 2002. DOI 10.1074/jbc.M204085200.
- [138] W. L. Jorgensen and J. Tirado-Rives. The OPLS Force Field for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.*, 110:1657–1666., 1988. DOI 10.1021/ja00214a001.

- [139] M. Juy, F. Penin, A. Favier, A. Galinier, R. Montserret, R. Haser, J. Deutscher, and A. Bockmann. Dimerization of Crh by reversible 3D domain swapping induces structural adjustments to its monomeric homologue Hpr. *J. Mol. Biol.*, 332(4):767–776, Sep 2003. DOI 10.1016/S0022-2836(03)00918-5.
- [140] R. A. Kammerer, D. Kostrewa, J. Zurdo, A. Detken, C. García-Echeverría, J. D. Green, S. A. Müller, B. H. Meier, F. K. Winkler, C. M. Dobson, and M. O. Steinmetz. Exploring amyloid formation by a de novo design. *Proc. Natl. Ac. Sci. USA*, 101(13):4435–4440, Mar 2004. DOI 10.1073/pnas.0306786101.
- [141] T. Karlsson and M. H. Levitt. Longitudinal rotational resonance echoes in solid state nuclear magnetic resonance: Investigation of zero quantum spin dynamics. *J. Chem. Phys.*, 109(13):5493–5507, Oct 1998. DOI 10.1063/1.477168.
- [142] S. K. Katti, D. M. Lemaster, and H. Eklund. Crystal-Structure of Thioredoxin from Escherichia-Coli at 1.68 Å Resolution. *J. Mol. Biol.*, 212(1):167–184, Mar 1990. DOI 10.1016/0022-2836(90)90313-B.
- [143] J. Keeler. *Understanding NMR Spectroscopy*. Wiley, 2005. ISBN 978-0470017876.
- [144] J. C. Kendrew, R. E. Dickerson, R. G. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, and V. C. Shore. Structure of Myoglobin. A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature*, 185:422, 1960.
- [145] M. A. Keniry, H. S. Gutowsky, and E. Oldfield. Surface dynamics of the integral membrane protein bacteriorhodopsin. *Nature*, 307(5949):383–386, 1984.
- [146] S. Kiihne, M. A. Mehta, J. A. Stringer, D. M. Gregory, J. C. Shiels, and G. P. Drobny. Distance measurements by dipolar recoupling two-dimensional solid-state NMR. *J. Phys. Chem. A*, 102(13):2274–2282, Mar 1998.
- [147] S. R. Kiihne, K. B. Geahigan, N. A. Oyler, H. Zebroski, M. A. Mehta, and G. P. Drobny. Distance measurements in multiply labeled crystalline cytidines by dipolar recoupling solid state NMR. *J. Phys. Chem. A*, 103(20):3890–3903, May 1999.
- [148] Y. Kimura, K. Kurzydowski, M. Tada, and D. H. MacLennan. Phospholamban inhibitory function is activated by depolymerization. *J. Biol. Chem.*, 272(24):15061–15064, Jun 1997.
- [149] Y. Kimura, M. Asahi, K. Kurzydowski, M. Tada, and D. H. MacLennan. Phospholamban domain Ib mutations influence functional interactions with the Ca<sup>2+</sup>-ATPase isoform of cardiac sarcoplasmic reticulum. *J. Biol. Chem.*, 273(23):14238–14241, Jun 1998. URL <http://www.jbc.org/cgi/content/abstract/273/23/14238>.
- [150] T. L. Kirby, C. B. Karim, and D. D. Thomas. Electron paramagnetic resonance reveals a large-scale conformational change in the cytoplasmic domain of phospholamban upon binding to the sarcoplasmic reticulum Ca-ATPase. *Biochemistry*, 43(19):5842–5852, May 2004. DOI 10.1021/bi035749b.

- [151] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983. DOI 10.1126/science.220.4598.671.
- [152] R. Koradi, M. Billeter, and K. Wuthrich. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.*, 14(1):51, Feb 1996. DOI 10.1016/0263-7855(96)00009-4.
- [153] P. E. Kristiansen, D. J. Mitchell, and J. N. S. Evans. Double-quantum dipolar recoupling at high magic-angle spinning rates. *J. Magn. Reson.*, 157(2):253–266, Aug 2002.
- [154] A. Kubo and C. A. McDowell. Spectral Spin Diffusion in Polycrystalline Solids under Magic-Angle Spinning. *J. Chem. Soc. Faraday Trans. I*, 84:3713–3730, 1988.
- [155] V. Ladizhansky and R. G. Griffin. Band-selective carbonyl to aliphatic side chain  $^{13}\text{C}$ – $^{13}\text{C}$  distance measurements in U- $^{13}\text{C}$ , $^{15}\text{N}$ -labeled solid peptides by magic angle spinning NMR. *J. Am. Chem. Soc.*, 126(3):948–958, Jan 2004. DOI 10.1021/ja037138c.
- [156] S. Lamberth, H. Schmid, M. Muenchbach, T. Vorherr, J. Krebs, E. Carafoli, and C. Griesinger. NMR solution structure of phospholamban. *Helv. Chim. Acta*, 83(9):2141–2152, 2000.
- [157] A. Lange. Master’s thesis, Universität Göttingen, Germany, 2002.
- [158] A. Lange. *Three-dimensional protein structure determination by high-resolution solid-state NMR spectroscopy*. PhD thesis, Universität Göttingen, Germany, 2006. URL <http://resolver.sub.uni-goettingen.de/purl/?webdoc-743>.
- [159] A. Lange, S. Luca, and M. Baldus. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.*, 124(33):9704–9705, Aug 2002.
- [160] A. Lange, K. Seidel, L. Verdier, S. Luca, and M. Baldus. Analysis of proton-proton transfer dynamics in rotating solids and their use for 3D structure determination. *J. Am. Chem. Soc.*, 125(41):12640–12648, Oct 2003. DOI 10.1021/ja034555g.
- [161] A. Lange, S. Becker, K. Seidel, K. Giller, O. Pongs, and M. Baldus. A concept for rapid protein-structure determination by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.*, 44(14):2089–2092, Mar 2005. DOI 10.1002/anie.200462516.
- [162] A. Lange, K. Giller, S. Hornig, M.-F. Martin-Eauclaire, O. Pongs, S. Becker, and M. Baldus. Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature*, 440(7086):959–962, Apr 2006. DOI 10.1038/nature04649.
- [163] A. Lange, T. Schupp, F. Petersen, T. Carlomagno, and M. Baldus. High-resolution solid-state NMR structure of an anticancer agent. *ChemMedChem*, 2(4):522–527, Apr 2007. DOI 10.1002/cmdc.200600299.
- [164] P. T. Lansbury, P. R. Costa, J. M. Griffiths, E. J. Simon, M. Auger, K. J. Halverson, D. A. Kocisko, Z. S. Hendsch, T. T. Ashburn, R. G. S. Spencer, B. Tidor, and R. G. Griffin. Structural Model for the Beta-Amyloid Fibril Based on Interstrand Alignment of an

- Antiparallel-Sheet Comprising a C-Terminal Peptide. *Nature Struct. Biol.*, 2(11):990–998, Nov 1995.
- [165] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and T. J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26:283–291, 1993.
- [166] D. D. Laws, H. M. L. Bitter, and A. Jerschow. Solid-state NMR spectroscopic methods in chemistry. *Angew. Chem. Int. Ed. Engl.*, 41:3096–3129, 2002. DOI 10.1002/1521-3773(20020902)41:17;3096::AID-ANIE3096j3.0.CO;2-X.
- [167] G. A. Lazar, J. R. Desjarlais, and T. M. Handel. De novo design of the hydrophobic core of ubiquitin. *Prot. Sci.*, 6(6):1167–1178, Jun 1997.
- [168] H. Le and E. Oldfield. Correlation between <sup>15</sup>N NMR chemical shifts in proteins and secondary structure. *J. Biomol. NMR*, 4(3):341–348, May 1994.
- [169] M. Lee and W. I. Goldberg. Nuclear-Magnetic-Resonance Line Narrowing by a Rotating Rf Field. *Phys. Rev.*, 140(4A):1261, 1965. DOI 10.1103/PhysRev.140.A1261.
- [170] Y. Lee, N. Kurur, M. Helmle, O. Johannessen, N. Nielsen, and M. Levitt. *Chem. Phys. Lett.*, 242:1995, 1995.
- [171] D. M. LeMaster and D. M. Kushlan. Dynamical mapping of E-coli thioredoxin via C-13 NMR relaxation analysis. *J. Am. Chem. Soc.*, 118(39):9255–9264, Oct 1996.
- [172] M. H. Levitt. Symmetry-Based Pulse Sequences in Magic-Angle Spinning Solid-State NMR. In D. M. Grant and R. K. Harris, editors, *Encyclopedia of Nuclear Magnetic Resonance: Supplementary Volume*, pages 165–196. John Wiley & Sons, 2002.
- [173] M. H. Levitt, D. P. Raleigh, F. Creuzet, and R. G. Griffin. Theory and Simulations of Homonuclear Spin Pair Systems in Rotating Solids. *J. Chem. Phys.*, 92(11):6347–6364, Jun 1990. DOI 10.1063/1.458314.
- [174] B. A. Lewis, G. S. Harbison, J. Herzfeld, and R. G. Griffin. NMR structural analysis of a membrane protein: bacteriorhodopsin peptide backbone orientation and motion. *Biochemistry*, 24(17):4671–4679, Aug 1985.
- [175] J. Li, C. B. Boschek, Y. Xiong, C. A. Sacksteder, T. C. Squier, and D. J. Bigelow. Essential role for Pro21 in phospholamban for optimal inhibition of the Ca-ATPase. *Biochemistry*, 44(49):16181–16191, Dec 2005. DOI 10.1021/bi051075o.
- [176] L. Lian and D. A. Middleton. Labelling approaches for protein structural studies by solution-state and solid-state NMR. *Progr. NMR Spectr.*, 39(3):171–190, Oct 2001. DOI 10.1016/S0079-6565(01)00034-6.
- [177] J. P. Linge and M. Nilges. Influence of non-bonded parameters on the quality of NMR structures: A new force field for NMR structure calculation. *J. Biomol. NMR*, 13(1):51–59, Jan 1999. DOI 10.1023/A:1008365802830.

- [178] J. P. Linge, M. Habeck, W. Rieping, and M. Nilges. ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics*, 19(2):315–316, Jan 2003.
- [179] J. P. Linge, M. A. Williams, C. Spronk, A. Bonvin, and M. Nilges. Refinement of protein structures in explicit solvent. *Prot. Struct. Funct. Gen.*, 50(3):496–506, Feb 2003.
- [180] G. Lipari and A. Szabo. Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.*, 104(17):4546–4559, 1982.
- [181] I. J. Lowe. Free Induction Decays of Rotating Solids. *Phys. Rev. Lett.*, 2:285–287, 1959. DOI 10.1103/PhysRevLett.2.285.
- [182] S. Luca and M. Baldus. Enhanced spectral resolution in immobilized peptides and proteins by combining chemical shift sum and difference spectroscopy. *J. Magn. Reson.*, 159(2):243–249, Dec 2002.
- [183] S. Luca, D. V. Filippov, J. H. van Boom, H. Oschkinat, H. J. M. de Groot, and M. Baldus. Secondary chemical shifts in immobilized peptides and proteins: A qualitative basis for structure refinement under Magic Angle Spinning. *J. Biomol. NMR*, 20(4):325–331, Aug 2001.
- [184] S. Luca, H. Heise, and M. Baldus. High-resolution solid-state NMR applied to polypeptides and membrane proteins. *Acc. Chem. Res.*, 36(11):858–865, Nov 2003. DOI 10.1021/ar020232y.
- [185] S. Luca, J. F. White, A. K. Sohal, D. V. Filippov, J. H. van Boom, R. Grisshammer, and M. Baldus. The conformation of neurotensin bound to its G protein-coupled receptor. *Proc. Natl. Ac. Sci. USA*, 100(19):10706–10711, Sep 2003. DOI 10.1073/pnas.1834523100.
- [186] W. E. J. R. Maas and W. S. Veeman. Natural abundance  $^{13}\text{C}$  spin diffusion enhanced by magic-angle spinning. *Chem. Phys. Lett.*, 149:170, 1988.
- [187] D. H. MacLennan and E. G. Kranias. Phospholamban: A crucial regulator of cardiac contractility. *Nature Rev. Mol. Cell Biol.*, 4(7):566–577, Jul 2003.
- [188] D. H. MacLennan, M. Abu-Abed, and C. Kang. Structure-function relationships in  $\text{Ca}^{2+}$  cycling proteins. *J. Mol. Cell. Cardiol.*, 34(8):897–918, Aug 2002.
- [189] J. L. Markley, A. Bax, Y. Arata, C. W. Hilbers, R. Kaptein, B. D. Sykes, P. E. Wright, and K. Wuthrich. Recommendations for the presentation of NMR structures of proteins and nucleic acids (IUPAC Recommendations 1998). *Pure Appl. Chem.*, 70(1):117–142, Jan 1998.
- [190] R. W. Martin and K. W. Zilm. Preparation of protein nanocrystals and their characterization by solid state NMR. *J. Magn. Reson.*, 165(1):162–174, Nov 2003.
- [191] D. Marulanda, M. L. Tasayco, M. Cataldi, V. Arriaran, and T. Polenova. Resonance Assignments and Secondary Structure Analysis of E. coli Thioredoxin by Magic Angle Spinning Solid-State NMR Spectroscopy. *J. Phys. Chem. B*, 109(38):18135–18145, September 29, 2005 2005.

- [192] J. Marx. Ubiquitin lives up to its name. *Science*, 297(5588):1792–1794, Sep 2002. DOI 10.1126/science.297.5588.1792.
- [193] J. H. Mathews and K. D. Fink. *Numerical Methods: Using Matlab*. Prentice-Hall, 4th edition, 2004. Taxi-cab and Powell’s conjugate gradient figures adapted from <http://math.fullerton.edu/mathews/n2003/PowellMethodBib.html>.
- [194] Y. Matsuki, H. Akutsu, and T. Fujiwara. Spectral fitting for signal assignment and structural analysis of uniformly  $^{13}\text{C}$ -labeled solid proteins by simulated annealing based on chemical shifts and spin dynamics. *J. Biomol. NMR*, 38(4):325–339, Aug 2007. DOI 10.1007/s10858-007-9170-x.
- [195] A. McDermott and G. Zhengtian. Carbon and Nitrogen Chemical Shifts: Applications to Solid State Proteins. In D. M. Grant and R. K. Harris, editors, *Encyclopedia of Nuclear Magnetic Resonance*, page 1137. John Wiley & Sons, 2007. DOI 10.1002/9780470034590.emrstm0020.
- [196] A. McDermott, T. Polenova, A. Bockmann, K. W. Zilm, E. K. Paulsen, R. W. Martin, and G. T. Montelione. Partial NMR assignments for uniformly (C-13, N-15)-enriched BPTI in the solid state. *J. Biomol. NMR*, 16(3):209–219, Mar 2000.
- [197] A. D. McLachlan. Gene duplication in the structural evolution of chymotrypsin. *J. Mol. Biol.*, 128:49–79, 1979. DOI 10.1016/0022-2836(79)90308-5.
- [198] M. Mehring. *Principles of High Resolution NMR in Solids*. Springer, 1983.
- [199] J. Meiler. PROSHIFT: protein chemical shift prediction using artificial neural networks. *J. Biomol. NMR*, 26(1):25–37, May 2003. DOI 10.1023/A:1023060720156.
- [200] E. Meng, B. Shoichet, and I. Kuntz. Automated docking with grid-based energy evaluation. *J. Comp. Chem.*, 13:505–524, 1992.
- [201] G. Metz, X. L. Wu, and S. O. Smith. Ramped-Amplitude Cross-Polarization in Magic-Angle-Spinning NMR. *J. Magn. Reson. A*, 110(2):219–227, Oct 1994.
- [202] C. R. Morcombe and K. W. Zilm. Chemical shift referencing in MAS solid state NMR. *J. Magn. Reson.*, 162(2):479–486, Jun 2003.
- [203] G. A. Morris and R. Freeman. Enhancement of Nuclear Magnetic-Resonance Signals by Polarization Transfer. *J. Am. Chem. Soc.*, 101(3):760–762, 1979.
- [204] G. M. Morris, D. S. Goodsell, R. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comp. Chem.*, 19:1639–1662, 1998. <http://autodock.scripps.edu/>.
- [205] B. Mueller, C. B. Karim, I. V. Negrashov, H. Kutchai, and D. D. Thomas. Direct detection of phospholamban and sarcoplasmic reticulum Ca-ATPase interaction in membranes using fluorescence resonance energy transfer. *Biochemistry*, 43(27):8754–8765, Jul 2004. DOI 10.1021/bi049732k.

- [206] M. Munowitz, W. P. Aue, and R. G. Griffin. Two-Dimensional Separation of Dipolar and Scaled Isotropic Chemical-Shift Interactions in Magic Angle NMR-Spectra. *J. Chem. Phys.*, 77(4):1686–1689, 1982. DOI 10.1063/1.444064.
- [207] S. Neal, A. M. Nip, H. Y. Zhang, and D. S. Wishart. Rapid and accurate calculation of protein H-1, C-13 and N-15 chemical shifts. *J. Biomol. NMR*, 26(3):215–240, Jul 2003.
- [208] N. C. Nielsen, H. Bildsoe, H. J. Jakobsen, and M. H. Levitt. Double-Quantum Homonuclear Rotary Resonance - Efficient Dipolar Recovery in Magic-Angle-Spinning Nuclear-Magnetic-Resonance. *J. Chem. Phys.*, 101(3):1805–1812, Aug 1994. DOI 10.1063/1.467759.
- [209] M. Nilges. Structure calculation from NMR data. *Curr. Op. Struct. Biol.*, 6(5):617–623, Oct 1996.
- [210] M. Nilges. Ambiguous distance data in the calculation of NMR structures. *Folding Design*, 2(4):S53–S57, 1997.
- [211] P. Nollert and L. Stewart. Climbing Mount Everest, the GPCR way. *Drug Disc. Today: Targets*, 3:2–4, 2004. DOI 10.1016/S1741-8372(04)02372-2.
- [212] K. Nomura, K. Takegoshi, T. Terao, K. Uchida, and M. Kainosho. Determination of the complete structure of a uniformly labeled molecule by rotational resonance solid-state NMR in the tilted rotating frame. *J. Am. Chem. Soc.*, 121(16):4064–4065, Apr 1999.
- [213] K. Obara, N. Miyashita, C. Xu, L. Toyoshima, Y. Sugita, G. Inesi, and C. Toyoshima. Structural role of countertransport revealed in Ca<sup>2+</sup> pump crystal structure in the absence of Ca<sup>2+</sup>. *Proc. Natl. Ac. Sci. USA*, 102(41):14489–14496, Oct 2005.
- [214] C. Olesen, T. L.-M. Sørensen, R. C. Nielsen, J. V. Møller, and P. Nissen. Dephosphorylation of the calcium pump coupled to counterion occlusion. *Science*, 306(5705):2251–2255, Dec 2004. DOI 10.1126/science.1106289.
- [215] C. Olesen, M. Picard, A. Lund Winther, C. Gyruup, J. P. Morth, C. Oxvig, J. V. Møller, and P. Nissen. The structural basis of calcium transport by the calcium pump. *Nature*, 450:1036–1042, Dec 2007. DOI 10.1038/nature06418.
- [216] S. J. Opella and F. M. Marassi. Structure determination of membrane proteins by NMR spectroscopy. *Chem. Rev.*, 104(8):3587–3606, Aug 2004. DOI 10.1021/cr0304121.
- [217] K. Oxenoid and J. J. Chou. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc. Natl. Ac. Sci. USA*, 102(31):10870–10875, Aug 2005. DOI 10.1073/pnas.0504920102.
- [218] A. G. Palmer, J. Williams, and A. McDermott. Nuclear magnetic resonance studies of biopolymer dynamics. *J. Phys. Chem.*, 100(31):13293–13310, Aug 1996.
- [219] M. Z. Papiz, S. M. Prince, T. Howard, R. J. Cogdell, and N. W. Isaacs. The structure and thermal motion of the B800-850 LH2 complex from *Rps. acidophila* at 2.0 Å resolution and 100 K: New structural features and functionally relevant motions. *J. Mol. Biol.*, 326(5):1523–1538, Mar 2003. DOI 10.1016/S0022-2836(03)00024-X.

- [220] A. B. Patel, E. Crocker, M. Eilers, A. Hirshfeld, M. Sheves, and S. O. Smith. Coupling of retinal isomerization to the activation of rhodopsin. *Proc. Natl. Ac. Sci. USA*, 101(27):10048–10053, Jul 2004.
- [221] J. Pauli, B. van Rossum, H. Forster, H. J. M. de Groot, and H. Oschkinat. Sample optimization and identification of signal patterns of amino acid side chains in 2D RFDR spectra of the alpha-spectrin SH3 domain. *J. Magn. Reson.*, 143(2):411–416, Apr 2000.
- [222] J. Pauli, M. Baldus, B. van Rossum, H. de Groot, and H. Oschkinat. Backbone and side-chain C-13 and N-15 signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 tesla. *ChemBioChem*, 2(4):272–281, Apr 2001.
- [223] L. Pauling, R. B. Corey, and H. R. Branson. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Ac. Sci. USA*, 37(4):205–211, Apr 1951.
- [224] E. K. Paulson, C. R. Morcombe, V. Gaponenko, B. Dancheck, R. A. Byrd, and K. W. Zilm. Sensitive high resolution inverse detection NMR spectroscopy of proteins in the solid state. *J. Am. Chem. Soc.*, 125(51):15831–15836, Dec 2003. DOI 10.1021/ja037315+.
- [225] M. Periasamy and A. Kalyanasundaram. SERCA pump isoforms: their role in calcium transport and disease. *Muscle & Nerve*, 35(4):430–442, Apr 2007. DOI 10.1002/mus.20745.
- [226] M. Perutz, M. Rossmann, A. Cullis, H. Muirhead, G. Will, and A. North. Structure of haemoglobin. A three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by X-ray analysis. *Nature*, 185:416–422, 1960.
- [227] W. Peti, J. Meiler, R. Bruschweiler, and C. Griesinger. Model-free analysis of protein backbone motion from residual dipolar couplings. *J. Am. Chem. Soc.*, 124(20):5822–5833, May 2002. DOI 10.1021/ja011883c.
- [228] A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, and R. Tycko. A structural model for Alzheimer’s beta-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Ac. Sci. USA*, 99(26):16742–16747, Dec 2002. DOI 10.1073/pnas.262663499.
- [229] A. T. Petkova, R. D. Leapman, Z. H. Guo, W. M. Yau, M. P. Mattson, and R. Tycko. Self-propagating, molecular-level polymorphism in Alzheimer’s beta-amyloid fibrils. *Science*, 307(5707):262–265, Jan 2005. DOI 10.1126/science.1105850.
- [230] A. F. Pieret, F. Durant, M. Griffe, and G. Germain. Crystal Structure of Tyrosin Ethyl Ester. *Acta Cryst. B*, 26(Dec):2117, 1970.
- [231] A. Pines, M. G. Gibby, and J. S. Waugh. Proton-Enhanced NMR of Dilute Spins in Solids. *J. Chem. Phys.*, 59(2):569–590, 1973. DOI 10.1063/1.1680061.
- [232] P. Pollesello, A. Annala, and M. Ovaska. Structure of the 1–36 amino-terminal fragment of human phospholamban by nuclear magnetic resonance and modeling of the phospholamban pentamer. *Biophys. J.*, 76(4):1784–1795, Apr 1999.



- [233] M. J. D. Powell. Restart Procedures for Conjugate Gradient Method. *Math. Progr.*, 12: 241–254, 1977.
- [234] W. G. Proctor and F. C. Yu. On the Magnetic Moments of  $\text{Mn}^{55}$ ,  $\text{Co}^{59}$ ,  $\text{Cl}^{37}$ ,  $\text{N}^{15}$ , and  $\text{N}^{14}$ . *Phys. Rev.*, 77(5):716–717, Mar 1950. DOI 10.1103/PhysRev.77.716.
- [235] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.*, 69(1-2):37–38, Jan 1946. DOI 10.1103/PhysRev.69.37.
- [236] D. P. Raleigh, M. H. Levitt, and R. G. Griffin. Rotational Resonance in Solid-State NMR. *Chem. Phys. Lett.*, 146(1-2):71–76, Apr 1988. DOI 10.1016/0009-2614(88)85051-6.
- [237] D. P. Raleigh, F. Creuzet, S. K. D. Gupta, M. H. Levitt, and R. G. Griffin. Measurement of Internuclear Distances in Polycrystalline Solids: Rotationally Enhanced Transfer of Nuclear-Spin Magnetization. *J. Am. Chem. Soc.*, 111(12):4502–4503, Jun 1989.
- [238] G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, 7:95–99, Jul 1963.
- [239] R. Ramachandran, V. Ladizhansky, V. S. Bajaj, and R. G. Griffin.  $^{13}\text{C}$ – $^{13}\text{C}$  rotational resonance width distance measurements in uniformly  $^{13}\text{C}$ -labeled peptides. *J. Am. Chem. Soc.*, 125(50):15623–15629, Dec 2003. DOI 10.1021/ja037761x.
- [240] C. M. Rienstra, L. Tucker-Kellogg, C. P. Jaroniec, M. Hohwy, B. Reif, M. T. McMahon, B. Tidor, T. Lozano-Pérez, and R. G. Griffin. De novo determination of peptide structure with solid-state magic-angle spinning NMR spectroscopy. *Proc. Natl. Ac. Sci. USA*, 99(16): 10260–10265, Aug 2002. DOI 10.1073/pnas.152346599.
- [241] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. Numerical-integration of cartesian equations of motion of a system with constraints—molecular dynamics of n-alkanes. *J. Comp. Phys.*, 23:327–341, 1977.
- [242] H. Saito. Conformation-Dependent C-13 Chemical-Shifts – a New Means of Conformational Characterization as Obtained by High-Resolution Solid-State C-13 NMR. *Magn. Res. Chem.*, 24(10):835–852, Oct 1986.
- [243] D. Sakellariou and L. Emsley. Through-bond experiments in solids. In D. M. Grant and R. K. Harris, editors, *Encyclopedia of Nuclear Magnetic Resonance*, volume 9, pages 196–211. John Wiley & Sons, 2002.
- [244] D. Sakellariou, A. Lesage, P. Hodgkinson, and L. Emsley. Homonuclear dipolar decoupling in solid-state NMR using continuous phase modulation. *Chem. Phys. Lett.*, 319:253, 2000. DOI 10.1016/S0009-2614(00)00127-5.
- [245] J. J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, revised edition, 1994.
- [246] J. Schaefer, R. A. McKay, and E. O. Stejskal. Double-Cross-Polarization NMR of Solids. *J. Magn. Reson.*, 34(2):443–447, 1979.

- [247] K. Schmidt-Rohr and H. W. Spiess. *Multidimensional Solid-State NMR and Polymers*. Academic Press, London, 1994.
- [248] R. Schneider, C. Ader, A. Lange, K. Giller, S. Hornig, O. Pongs, S. Becker, and M. Baldus. Solid-state NMR spectroscopy applied to a chimeric potassium channel in lipid bilayers. *J. Am. Chem. Soc.*, 130(23):7427–7435, Jun 2008. DOI 10.1021/ja800190c.
- [249] B. R. Seavey, E. A. Farr, W. M. Westler, and J. L. Markley. A relational database for sequence-specific protein NMR data. *J. Biomol. NMR*, 1(3):217–236, Sep 1991. DOI 10.1007/BF01875516.
- [250] J. Seelig and H. U. Gally. Investigation of Phosphatidylethanolamine Bilayers by Deuterium and P-31 Nuclear Magnetic-Resonance. *Biochemistry*, 15(24):5199–5204, 1976.
- [251] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. D. Payne. First-principles simulation: ideas, illustrations and the CASTEP code. *J. Phys.: Cond. Matt.*, 14:2717, 2002. DOI 10.1088/0953-8984/14/11/301.
- [252] K. Seidel. Structural Characterization of uniformly isotope-labeled biomolecules by solid-state nuclear magnetic resonance. Master’s thesis, Universität Göttingen, 2003.
- [253] K. Seidel, A. Lange, S. Becker, C. E. Hughes, H. Heise, and M. Baldus. Protein solid-state NMR resonance assignments from ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) correlation spectroscopy. *Phys. Chem. Chem. Phys.*, 6(22):5090–5093, 2004. DOI 10.1039/b411689e.
- [254] K. Seidel, M. Etzkorn, H. Heise, S. Becker, and M. Baldus. High-resolution solid-state NMR studies on uniformly [ $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled ubiquitin. *ChemBioChem*, 6(9):1638–1647, Sep 2005. DOI 10.1002/cbic.200500085.
- [255] K. Seidel, M. Etzkorn, L. Sonnenberg, C. Griesinger, A. Sebald, and M. Baldus. Studying molecular 3D structure and dynamics by high-resolution solid-state NMR: Application to L-tyrosine-ethylester. *J. Phys. Chem. A*, 109(11):2436–2442, Mar 2005. DOI 10.1021/jp045605m.
- [256] K. Seidel, O. C. Andronesi, J. Krebs, C. Griesinger, H. S. Young, S. Becker, and M. Baldus. Structural characterization of  $\text{Ca}^{2+}$ -ATPase-bound phospholamban in lipid bilayers by solid-state NMR spectroscopy. *Biochemistry*, 47(15):4369–4376, Apr 2008. DOI 10.1021/bi7024194.
- [257] D. J. Selkoe. Folding proteins in fatal ways. *Nature*, 426:900–904, 2003. DOI 10.1038/nature02264.
- [258] A. J. Shaka, P. B. Barker, and R. Freeman. Computer-Optimized Decoupling Scheme for Wideband Applications and Low-Level Operation. *J. Magn. Reson.*, 64(3):547–552, 1985.
- [259] Y. Shen and A. Bax. Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR*, 38(4):289–302, Aug 2007. DOI 10.1007/s10858-007-9166-6.

- [260] H. K. Simmerman and L. R. Jones. Phospholamban: protein structure, mechanism of action, and role in cardiac function. *Physiol. Rev.*, 78(4):921–947, Oct 1998.
- [261] C. P. Slichter. *Principles of Magnetic Resonance*. Springer, 3rd edition, 1989.
- [262] S. A. Smith, T. O. Levante, B. H. Meier, and R. R. Ernst. Computer-Simulations in Magnetic-Resonance - an Object-Oriented Programming Approach. *J. Magn. Reson. A*, 106(1):75–105, Jan 1994.
- [263] S. O. Smith, S. Farrjones, R. G. Griffin, and W. W. Bachovchin. Crystal Versus Solution Structures of Enzymes - NMR-Spectroscopy of a Crystalline Serine Protease. *Science*, 244(4907):961–964, May 1989.
- [264] S. O. Smith, T. Kawakami, W. Liu, M. Ziliox, and S. Aimoto. Helical structure of phospholamban in membrane bilayers. *J. Mol. Biol.*, 313(5):1139–1148, Nov 2001. DOI 10.1006/jmbi.2001.5101.
- [265] L. Sonnenberg. Abstandsmessungen an uniform isotopenmarkierten Biomolekülen mit Festkörper-NMR. Master's thesis, Universität Göttingen, 2003.
- [266] L. Sonnenberg, S. Luca, and M. Baldus. Multiple-spin analysis of chemical-shift-selective ( $^{13}\text{C}$ ,  $^{13}\text{C}$ ) transfer in uniformly labeled biomolecules. *J. Magn. Reson.*, 166(1):100–110, Jan 2004.
- [267] R. Sowdhamini, N. Srinivasan, B. Shoichet, D. Vonsanti, C. Ramakrishnan, and P. Balaram. Stereochemical Modeling of Disulfide Bridges – Criteria for Introduction into Proteins by Site-Directed Mutagenesis. *Prot. Eng.*, 3(2):95–103, Nov 1989.
- [268] S. Spera and A. Bax. Empirical Correlation between Protein Backbone Conformation and C-Alpha and C-Beta C-13 Nuclear-Magnetic-Resonance Chemical-Shifts. *J. Am. Chem. Soc.*, 113(14):5490–5492, Jul 1991.
- [269] C. Spronk, J. P. Linge, C. W. Hilbers, and G. W. Vuister. Improving the quality of protein structures derived by NMR spectroscopy. *J. Biomol. NMR*, 22(3):281–289, Mar 2002.
- [270] E. G. Stein, L. M. Rice, and A. T. Brünger. Torsion-Angle Molecular Dynamics as a New Efficient Tool for NMR Structure Calculation. *J. Magn. Reson.*, 124:154–164, 1997. DOI 10.1006/jmre.1996.1027.
- [271] F. Stickney de Bouregas and J. S. Waugh. ANTIOPE, A program for computer experiments on spin dynamics. *J. Magn. Reson.*, 96:280–289, 1992.
- [272] S. K. Straus, T. Bremi, and R. R. Ernst. Experiments and strategies for the assignment of fully C-13/N-15-labelled polypeptides by solid state NMR. *J. Biomol. NMR*, 12(1):39–50, Jul 1998.
- [273] L. Stryer, J. M. Berg, and J. L. Tymoczko. *Biochemie*. Spektrum Akademischer Verlag, 5th edition, 2007.

- [274] H. H. Sun, L. K. Sanders, and E. Oldfield. Carbon-13 NMR shielding in the twenty common amino acids: Comparisons with experimental results in proteins. *J. Am. Chem. Soc.*, 124(19):5486–5495, May 2002.
- [275] D. Suter and R. R. Ernst. Spectral Spin Diffusion in the Presence of an Extraneous Dipolar Reservoir. *Phys. Rev. B*, 25(9):6038–6041, 1982.
- [276] D. Suter and R. R. Ernst. Spin Diffusion in Resolved Solid-State NMR-Spectra. *Phys. Rev. B*, 32(9):5608–5627, 1985.
- [277] T. L.-M. Sørensen, J. V. Møller, and P. Nissen. Phosphoryl transfer and calcium ion occlusion in the calcium pump. *Science*, 304(5677):1672–1675, Jun 2004. DOI 10.1126/science.1099366.
- [278] M. Tada, M. A. Kirchberger, and A. M. Katz. Phosphorylation of a 22,000-dalton component of the cardiac sarcoplasmic reticulum by adenosine 3':5'-monophosphate-dependent protein kinase. *J. Biol. Chem.*, 250(7):2640–2647, Apr 1975.
- [279] M. Takahashi, Y. Kondou, and C. Toyoshima. Interdomain communication in calcium pump as revealed in the crystal structures with transmembrane inhibitors. *Proc. Natl. Ac. Sci. USA*, 104(14):5800–5805, Apr 2007. DOI 10.1073/pnas.0700979104.
- [280] Takamori, Jahn, S. others Takamori, M. Holt, K. Stenius, E. A. Lemke, M. Grønberg, D. Riedel, H. Urlaub, S. Schenck, et al. Molecular Anatomy of a Trafficking Organelle. *Cell*, 127:831–846, 2006. DOI 10.1016/j.cell.2006.10.030.
- [281] K. Takegoshi, K. Nomura, and T. Terao. Rotational Resonance in the Tilted Rotating-Frame. *Chem. Phys. Lett.*, 232(5-6):424–428, Jan 1995. DOI 10.1016/0009-2614(94)01399-G.
- [282] K. Takegoshi, S. Nakamura, and T. Terao.  $^{13}\text{C}$ - $^1\text{H}$  dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.*, 344(5-6):631–637, Aug 2001. DOI 10.1016/S0009-2614(01)00791-6.
- [283] N. Tjandra, S. E. Feller, R. W. Pastor, and A. Bax. Rotational diffusion anisotropy of human ubiquitin from N-15 NMR relaxation. *J. Am. Chem. Soc.*, 117(50):12562–12566, Dec 1995.
- [284] N. Tjandra, A. Szabo, and A. Bax. Protein backbone dynamics and N-15 chemical shift anisotropy from quantitative measurement of relaxation interference effects. *J. Am. Chem. Soc.*, 118(29):6986–6991, Jul 1996.
- [285] Y. Todokoro, I. Yumen, K. Fukushima, S.-W. Kang, J.-S. Park, T. Kohno, K. Wakamatsu, H. Akutsu, and T. Fujiwara. Structure of tightly membrane-bound mastoparan-X, a G-protein-activating peptide, determined by solid-state NMR. *Biophys. J.*, 91(4):1368–1379, Aug 2006. DOI 10.1529/biophysj.106.082735.
- [286] D. A. Torchia. Solid-State NMR-Studies of Protein Internal Dynamics. *Annual Review of Biophysics and Bioengineering*, 13:125–144, 1984.

- [287] T. Toyofuku, K. Kurzydowski, M. Tada, and D. H. MacLennan. Amino-Acids Glu(2) to Ile(18) in the Cytoplasmic Domain of Phospholamban Are Essential for Functional Association with the  $\text{Ca}^{2+}$ -ATPase of Sarcoplasmic-Reticulum. *J. Biol. Chem.*, 269(4):3088–3094, Jan 1994.
- [288] T. Toyofuku, K. Kurzydowski, M. Tada, and D. H. MacLennan. Amino-Acids Lys-Asp-Asp-Lys-Pro-Val(402) in the  $\text{Ca}^{2+}$ -ATPase of Cardiac Sarcoplasmic-Reticulum Are Critical for Functional Association with Phospholamban. *J. Biol. Chem.*, 269(37):22929–22932, Sep 1994.
- [289] C. Toyoshima and T. Mizutani. Crystal structure of the calcium pump with a bound ATP analogue. *Nature*, 430(6999):529–535, Jul 2004. DOI 10.1038/nature02680.
- [290] C. Toyoshima and H. Nomura. Structural changes in the calcium pump accompanying the dissociation of calcium. *Nature*, 418(6898):605–611, Aug 2002. DOI 10.1038/nature00944.
- [291] C. Toyoshima, M. Nakasako, H. Nomura, and H. Ogawa. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature*, 405(6787):647–655, Jun 2000. DOI 10.1038/35015017.
- [292] C. Toyoshima, M. Asahi, Y. Sugita, R. Khanna, T. Tsuda, and D. H. MacLennan. Modeling of the inhibitory interaction of phospholamban with the  $\text{Ca}^{2+}$  ATPase. *Proc. Natl. Ac. Sci. USA*, 100(2):467–472, Jan 2003.
- [293] N. J. Traaseth, J. J. Buffy, J. Zamoan, and G. Veglia. Structural dynamics and topology of phospholamban in oriented lipid bilayers using multidimensional solid-state NMR. *Biochemistry*, 45(46):13827–13834, Nov 2006. DOI 10.1021/bi0607610.
- [294] N. J. Traaseth, R. Verardi, K. D. Torgersen, C. B. Karim, D. D. Thomas, and G. Veglia. Spectroscopic validation of the pentameric structure of phospholamban. *Proc. Natl. Ac. Sci. USA*, 104(37):14676–14681, Sep 2007. DOI 10.1073/pnas.0701016104.
- [295] M. E. Tuckerman and M. Parrinello. Integrating the Car-Parrinello equations. I. Basic integration techniques. *J. Chem. Phys.*, 101(2):1302–1315, 1994. DOI 10.1063/1.467823.
- [296] D. L. Turner. Phase Cycling. In D. M. Grant and R. K. Harris, editors, *Encyclopedia of Nuclear Magnetic Resonance: Supplementary Volume*, pages 165–196. John Wiley & Sons. DOI 10.1002/9780470034590.emrstm0389.
- [297] R. Tycko. Solid-state NMR as a probe of amyloid fibril structure. *Curr. Op. Chem. Biol.*, 4(5):500–506, Oct 2000.
- [298] R. Tycko and Dabbagh. Measurement of nuclear magnetic dipole–dipole couplings in magic angle spinning NMR. *Chem. Phys. Lett.*, 173(5–6):461–465, Oct 1990. DOI 10.1016/0009-2614(90)87235-J.
- [299] R. Tycko and Y. Ishii. Constraints on supramolecular structure in amyloid fibrils from two-dimensional solid-state NMR spectroscopy with uniform isotopic labeling. *J. Am. Chem. Soc.*, 125(22):6606–6607, Jun 2003. DOI 10.1021/ja0342042.

- [300] E. L. Ulrich, H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. K. Wenger, H. Yao, and J. L. Markley. BioMagResBank. *Nucl. Acids Res.*, in press, Nov 2007. DOI 10.1093/nar/gkm957.
- [301] A. D. J. van Dijk, R. Boelens, and A. M. J. J. Bonvin. Data-driven docking for the study of biomolecular complexes. *FEBS J.*, 272(2):293–312, Jan 2005. DOI 10.1111/j.1742-4658.2004.04473.x.
- [302] A. J. van Gammeren, F. B. Hulsbergen, J. G. Hollander, and H. J. M. de Groot. Residual backbone and side-chain  $^{13}\text{C}$  and  $^{15}\text{N}$  resonance assignments of the intrinsic transmembrane light-harvesting 2 protein complex by solid-state Magic Angle Spinning NMR spectroscopy. *J. Biomol. NMR*, 31(4):279–293, Apr 2005. DOI 10.1007/s10858-005-1604-8.
- [303] W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hünenberger, M. A. Kastholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. A. van der Vegt, and H. B. Yu. Biomolecular Modeling: Goals, Problems, Perspectives. *Angew. Chem. Int. Ed. Engl.*, 45:4064–4092, 2006. DOI 10.1002/anie.200502655.
- [304] B. J. van Rossum, C. P. de Groot, V. Ladizhansky, S. Vega, and H. J. M. de Groot. A method for measuring heteronuclear (H-1-C-13) distances in high speed MAS NMR. *J. Am. Chem. Soc.*, 122(14):3465–3472, Apr 2000.
- [305] R. Verel, M. Baldus, M. Nijman, J. W. M. van Os, and B. H. Meier. Adiabatic homonuclear polarization transfer in magic-angle-spinning solid-state NMR. *Chem. Phys. Lett.*, 280(1-2): 31–39, Nov 1997.
- [306] R. Verel, M. Baldus, M. Ernst, and B. H. Meier. A homonuclear spin-pair filter for solid-state NMR based on adiabatic-passage techniques. *Chem. Phys. Lett.*, 287(3-4):421–428, May 1998.
- [307] M. Veshtort and R. G. Griffin. SPINEVOLUTION: a powerful tool for the simulation of solid and liquid state NMR experiments. *J. Magn. Reson.*, 178(2):248–282, Feb 2006. DOI 10.1016/j.jmr.2005.07.018.
- [308] S. Vijay-Kumar, C. E. Bugg, and W. J. Cook. Structure of Ubiquitin Refined at 1.8 Å Resolution. *J. Mol. Biol.*, 194(3):531–544, Apr 1987. DOI 10.1016/0022-2836(87)90679-6.
- [309] G. Vriend. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.*, 8: 52–56, 1990.
- [310] A. J. Wand, J. L. Urbauer, R. P. McEvoy, and R. Bieber. Internal Dynamics of Human Ubiquitin Revealed by  $^{13}\text{C}$ -Relaxation Studies of Randomly Fractionally Labeled Protein. *J. Biochem.*, 35:6116–6125, 1996. DOI 10.1021/bi9530144.
- [311] A. C. Wang, S. Grzesiek, R. Tschudin, P. J. Lodi, and A. Bax. Sequential Backbone Assignment of Isotopically Enriched Proteins in D<sub>2</sub>O by Deuterium-Decoupled HA(CA)N and HA(CACO)N. *J. Biomol. NMR*, 5(4):376–382, Jun 1995. DOI 10.1007/BF00182281.

- [312] Y. J. Wang and O. Jardetzky. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Prot. Sci.*, 11(4):852–861, Apr 2002. URL <http://www.proteinscience.org/cgi/content/abstract/11/4/852>.
- [313] G. B. Warren, P. A. Toon, N. J. Birdsall, A. G. Lee, and J. C. Metcalfe. Reconstitution of a Calcium-Pump Using Defined Membrane Components. *Proc. Natl. Ac. Sci. USA*, 71(3):622–626, 1974.
- [314] A. Watts. Solid-state NMR in drug design and discovery for membrane-embedded targets. *Nat. Rev. Drug. Discov.*, 4(7):555–568, Jul 2005. DOI 10.1038/nrd1773.
- [315] J. S. Waugh. Average Hailtonian Theory. In D. M. Grant and R. K. Harris, editors, *Encyclopedia of Nuclear Magnetic Resonance*. John Wiley & Sons, 2007. DOI 10.1002/9780470034590.emrstm0020.
- [316] Y. F. Wei, D. K. Lee, and A. Ramamoorthy. Solid-state  $^{13}\text{C}$  NMR chemical shift anisotropy tensors of polypeptides. *J. Am. Chem. Soc.*, 123(25):6118–6126, Jun 2001. DOI 10.1021/ja010145l.
- [317] S. H. White. The progress of membrane protein structure determination. *Prot. Sci.*, 13:948, 2004. See [http://blanco.biomol.uci.edu/Membrane\\_Proteins\\_xtal.html](http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html) for a recent list of membrane proteins of known structures.
- [318] P. T. F. Williamson, A. Verhoeven, M. Ernst, and B. H. Meier. Determination of internuclear distances in uniformly labeled molecules by rotational-resonance solid-state NMR. *J. Am. Chem. Soc.*, 125(9):2718–2722, Mar 2003. DOI 10.1021/ja028210u.
- [319] D. S. Wishart, B. D. Sykes, and F. M. Richards. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.*, 222(2):311–333, Nov 1991.
- [320] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes.  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, 6(2):135–140, Sep 1995.
- [321] A. Wlodawer, J. Nachman, G. L. Gilliland, W. Gallagher, and C. Woodward. Structure of form III crystals of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.*, 198(3):469–480, Dec 1987.
- [322] Wüthrich. *NMR of Proteins and Nucleic Acids*. Wiley Interscience, 1986. ISBN 978-0471828938.
- [323] X.-P. Xu and D. A. Case. Probing multiple effects on  $^{15}\text{N}$ ,  $^{13}\text{C}$  alpha,  $^{13}\text{C}$  beta, and  $^{13}\text{C}'$  chemical shifts in peptides using density functional theory. *Biopolymers*, 65(6):408–423, Dec 2002. DOI 10.1002/bip.10276.
- [324] J. X. You, R. E. Cohen, and C. M. Pickart. Construct for high-level expression and low misincorporation of lysine for arginine during expression of pET-encoded eukaryotic proteins in *Escherichia coli*. *Biotechniques*, 27(5):950–954, Nov 1999.

- [325] J. Zamoon, A. Mascioni, D. D. Thomas, and G. Veglia. NMR solution structure and topological orientation of monomeric phospholamban in dodecylphosphocholine micelles. *Biophys. J.*, 85(4):2589–2598, Oct 2003. URL <http://www.biophysj.org/cgi/content/abstract/85/4/2589>.
- [326] J. Zamoon, F. Nitu, C. Karim, D. D. Thomas, and G. Veglia. Mapping the interaction surface of a membrane protein: unveiling the conformational switch of phospholamban in calcium pump regulation. *Proc. Natl. Ac. Sci. USA*, 102(13):4747–4752, Mar 2005. DOI 10.1073/pnas.0406039102.
- [327] S. G. Zech, E. Olejniczak, P. Hajduk, J. Mack, and A. E. McDermot. Characterization of protein-ligand interactions by high-resolution solid-state NMR spectroscopy. *J. Am. Chem. Soc.*, 126(43):13948–13953, Nov 2004. DOI 10.1021/ja040086m.
- [328] S. G. Zech, A. J. Wand, and A. E. McDermott. Protein structure determination by high-resolution solid-state NMR spectroscopy: application to microcrystalline ubiquitin. *J. Am. Chem. Soc.*, 127(24):8618–8626, Jun 2005. DOI 10.1021/ja0503128.
- [329] H. Zhang, S. Neal, and D. S. Wishart. RefDB: a database of uniformly referenced protein chemical shifts. *J. Biomol. NMR*, 25(3):173–195, Mar 2003. DOI 10.1023/A:1022836027055.
- [330] P. Zhang, N. M. Green, C. Toyoshima, K. Yonekura, and D. L. Stokes. Transmembrane architecture of the calcium pump from sarcoplasmic reticulum by cryoelectron microscopy. *Biophys. J.*, 74(2):A328–A328, Feb 1998. Part 2.
- [331] Y. Zhou, J. H. Morais-Cabral, A. Kaufman, and R. MacKinnon. Chemistry of ion coordination and hydration revealed by a  $K^+$  channel-Fab complex at 2.0 Å resolution. *Nature*, 414(6859):43–48, Nov 2001. DOI 10.1038/35102009.



# Lebenslauf

Karsten Seidel

Geboren am 6. Dezember 1976 in Bremerhaven

Deutsche Staatsangehörigkeit

- Seit 2004            Wissenschaftlicher Mitarbeiter in der Arbeitsgruppe *Festkörper-NMR-Spektroskopie* von Dr. Baldus am Max-Planck-Institut für biophysikalische Chemie, Göttingen; Doktorarbeit
- 1997–2003        Studium der Physik mit Schwerpunkten Biophysik und physikalische Chemie an der Universität Göttingen; Diplomarbeit am Max-Planck-Institut für biophysikalische Chemie, AG Baldus, betreut durch Prof. Dr. Hofsäß und Prof. Dr. Griesinger; Abschluss mit Diplom 11/2003 (Vordiplom 10/1999)
- 1996–97           Zivildienst in einer Wohnstätte für Erwachsene mit körperlicher oder geistiger Behinderung
- 1983–96           Schulbesuch in Bremerhaven, Abitur am Schulzentrum Carl von Ossietzky

## Auslandsaufenthalte

- 2005                Projekt *DNP studies of a membrane protein* am Massachusetts Institute of Technology, Cambridge, USA, MIT/Harvard Center for Magnetic Resonance, bei Prof. Dr. Griffin (1 Semester)
- 2000–01            Studium der Physik mit Schwerpunkt Biophysik an der University of California, San Diego, USA (2 Semester)

## Wissenschaftliche Hilfskraft in der Lehre

- WS 2004/05        Integrierter Kurs Physik I (Prof. Dr. Salditt / Prof. Dr. Zippelius)
- WS 2001/02        Physik I / Theorie I (Prof. Dr. Samwer / Prof. Dr. Zippelius)

## Wissenschaftliche Hilfskraft in der Forschung

- WS 1999/2000    Institut für Psychologie der Universität Göttingen (Dr. Werner)

Mitglied der Deutschen Physikalischen Gesellschaft seit 1999