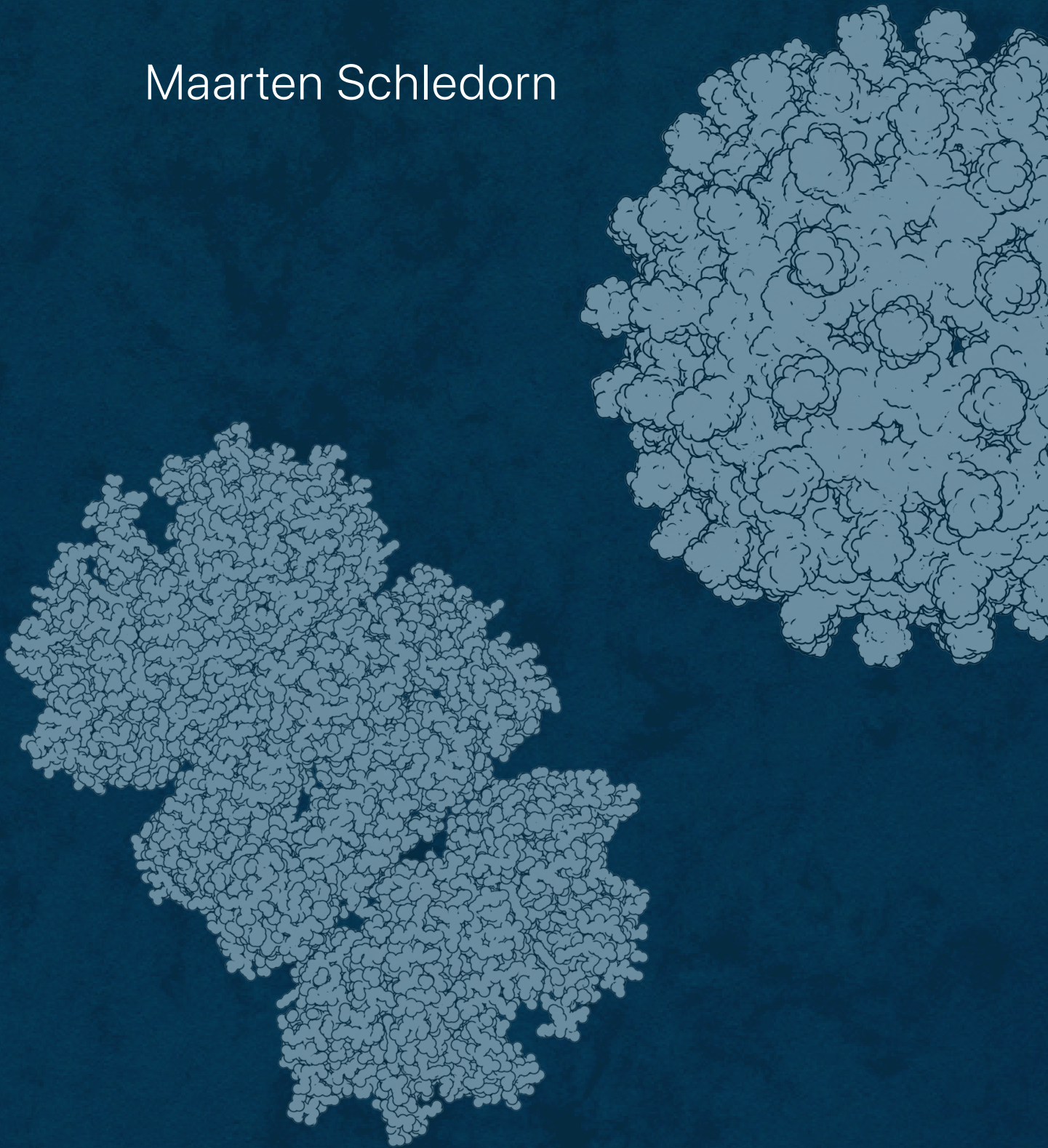# Solid-State NMR Studies of Large Protein Assemblies

Maarten Schledorn

DISS. ETH NO. 26430

Solid-State NMR Studies of
Large Protein Assemblies

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

MAARTEN SCHLEDORN

MSc Medical Biotechnology, KTH Royal Institute of Technology

born on 23.02.1988

Citizen of the Kingdom of the Netherlands

accepted on the recommendation of

Prof. Dr. Beat H. Meier
Prof. Dr. Roland Riek
Dr. Anja Böckmann

2019

Contents

Appendix   143

Abbreviations

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
Aβ	amyloid beta
AD	Alzheimer's disease
ADP	adenosine diphosphate
AHT	average Hamiltonian theory
AMP-PCP	adenosine methylene triphosphate
ATP	adenosine triphosphate
BMRB	biological magnetic resonance bank
CAM	capsid assembly modulator
CcpNmr	collaborative computing project for NMR (*cf.* NMR)
CD	circular dichroism
CFPS	cell-free protein synthesis
Cp	HBV core protein (*cf.* HBV)
CP	cross polarisation
CSA	chemical-shift anisotropy
CSP	chemical-shift perturbation
DARR	dipolar assisted rotational resonance
DNA	deoxyribonucleic acid
DNP	dynamic nuclear polarisation
DREAM	dipolar recoupling enhanced by amplitude modulation
DSS	dimethyl-silapentane-sulfonate
DSSP	define secondary structure of proteins
EM	electron microscopy
EPR	electron paramagnetic resonance
FID	free induction decay

fMLF	formyl-methionine-leucine-phenylalanine
FWHM	full width at half maximum
GB1	immunoglobulin-binding protein G, domain B1
HBs S/M/L	small/medium/large HBV envelope protein (*cf.* HBV)
HBV	hepatitis B virus
IR	infrared radiation
MAS	magic-angle spinning
MPL	mass-per-length
NA	natural abundance
NMR	nuclear magnetic resonance
PAIN	proton-assisted insensitive nuclei
PAR	proton-assisted recoupling
PCNA	proliferating cell nuclear antigen
PDB	protein data bank
PDSD	proton-driven spin diffusion
ppm	parts per million
rf	radio frequency
RNA	ribonucleic acid
RPO	RNA polymerase (*cf.* RNA)
SANS	small-angle neutron scattering
SAXS	small-angle X-ray scattering
SNR	signal-to-noise ratio
SST	somatostatin
STEM	scanning transmission electron microscopy
STRIDE	structural identification
SVP	sub-viral particle
TALOS	torsion angle likelihood obtained from shift and sequence similarity
TEDOR	transferred echo double resonance
TEM	transmission electron microscopy
TMS	tetramethylsilane
UL	uniformly labelled (commonly referring to ^{13}C and ^{15}N)
WGE	wheat germ extract
XFEL	X-ray free-electron laser
$x \pm y$	if not otherwise specified, y indicates a single standard deviation

Abstract

Amongst other applications, solid-state nuclear magnetic resonance (NMR) spectroscopy can provide atomic-resolution data for the determination of protein structures and dynamics. This is true in particular for protein assemblies that cannot be crystallised or are too large for solution-state NMR. While understanding the dynamics of a protein is of great importance for a full appreciation of its molecular machinery, this thesis has a focus on the structural aspects of a number of large protein assemblies.

The first study, presented in **Chapter 2**, builds on previously determined chemical shifts of a mutant form of amyloid-β, which is compared to a brain-seeded form of the wild-type. This comparison suggests that the determined mutant fold can also be adopted by the wild-type, with small conformational adaptations which accommodate the E22 deletion in the Osaka mutant. In addition, this chapter illustrates how other mutants could conform to this model. The stabilisation of the N-terminal part of the protein via an intermolecular salt bridge to K28 may represent a common structural motif for the mutants that are related to early-onset Alzheimer's disease. This feature may connect to the observed increased toxicity of the mutant forms compared to wild-type Amyloid-β1-40, where the salt bridge involving K28 is intramolecular instead of intermolecular.

A continuation of amyloid studies follows in **Chapter 3**, but for a different kind of amyloid: in recent years the idea that a number of peptide hormones and neuropeptides are transiently stored in aggregated form has accumulated support. These reversible, functional amyloids are believed to be packed into dense-core vesicles, which function as temporary depots of messenger peptides in secretory cells. Somatostatin (SST) is such a peptide hormone that occurs physiologically both aggregated and as a soluble monomer. The structure of human SST-14 in the context of a fibril was determined to atomic resolution using magic-angle spinning (MAS) solid-state NMR spectroscopy. In addition to scanning transmission electron microscopy data, the complete backbone resonance assignment is presented in this chapter. Subsequently, dipolar-based experiments that provide spectrally unambiguous long-range distance restraints are combined with a prediction of secondary-structure ele-

ments by secondary chemical-shift calculations and dihedral-angle restraints. The collective data culminate in the molecular structure presented in this chapter.

In **Chapter 4**, both ^{13}C- and ^{1}H-detected experiments are presented. Both approaches are compared in general, and more specifically in the context of several proteins related to the hepatitis B virus (HBV). HBV is a small enveloped DNA virus whose genomic information encodes only a few genes: the envelope proteins S, M and L (collectively known as hepatitis B surface antigen/HBsAg), the core protein (Cp), the polymerase (P), and the X protein (HBx). This chapter presents structural studies of the envelope protein S and the core protein Cp in its full-length (including C-terminal domain (CTD)) and reduced (without CTD) forms.

Proton-detection is applied to probe interactions between protein and nucleic acids (ATP analogues and the deoxyribonucleotides of DNA) in combination with phosphorus-detected experiments in **Chapter 5**. Protein-nucleic acid interactions play important roles not only in energy-providing reactions such as ATP hydrolysis, but also in reading, extending, packaging or repairing genomes. While they can often be analysed in detail with X-ray crystallography, complementary methods are necessary to visualise these interactions in complexes which are not crystalline. This chapter describes how solid-state NMR can detect and classify protein-nucleic acid interactions via site-specific ^{1}H- and ^{31}P-detected spectroscopy. The sensitivity of ^{1}H chemical-shift values for non-covalent interactions involved in these molecular recognition processes is exploited to directly probe the chemical bonding state, a characteristic that cannot be directly obtained from an X-ray structure. Despite its rather challenging size, the method is applied to study interactions in the 669 kDa dodecameric DnaB helicase in complex with ADP:AlF$_4^-$:DNA.

Finally, **Chapter 6** investigates proton-detection in solid-state NMR rather from a more methodological point of view in the context of MAS and resolution. Spectral resolution is key to unleash the structural and dynamic information contained in NMR spectra. The advent of ever faster MAS, today exceeding 100 kHz, is what enabled proton detection in solid-state NMR. In this respect, it is valuable to evaluate the benefit of a continued investment in faster spinning. To address this question, MAS up to 150 kHz is used to investigate a protein complex of archaeal RNA polymerase subunits 4 and 7. Using a rotor with an outer diameter of 0.5 mm and a sample content of approximately 170 µg, the total linewidth of Rpo4/7 improves by a factor of 1.23 ± 0.05 by going from 100 to 150 kHz, and signal intensity increases by a factor 1.48 ± 0.13 in the same MAS range. With some further considerations demonstrated in this chapter, the conclusion is that continued investment in faster MAS is indeed meaningful.

Zusammenfassung

Neben anderen Anwendungen beitet die Festkörper-Kernresonanzspektroskopie (NMR) die Möglichkeit die Struktur und Dynamik von Proteinen auf atomarer Ebene zu analysieren. Diese sind insbesondere für Proteinzusammensetzungen relevant, die sich nicht kristallisieren lassen oder die für Lösungs-NMR zu groß sind. Obwohl die Kenntnis der Proteindynamik für ein fundamentelles Verständnis molekularer Mechanismen wichtig ist, liegt der Fokus dieser Arbeit auf der Aufklärung von strukturellen Aspekten großer Proteinkomplexe.

Kapitel 2 befasst sich mit dem Vergleich eines Amyloid-β Mutanten mit seinem Wildtyp, dessen Fibrillen gekeimt wurden mit Hilfe klinischer Fibrillen-Proben von an Alzheimer erkrankten Patienten. Die Grundlage hierzu liefern die zuvor bestimmten chemischen Verschiebungen des Mutanten. Dieser Vergleich deutet darauf hin, dass die Faltung des Mutanten mit kleinen Anpassungen der Konformation auch von der Wildform angenommen werden kann. Zudem zeigt dieses Kapitel, wie fünf weitere der früh einsetzenden Alzheimer-Krankheit zugeschriebene Mutanten mit diesem Modell übereinstimmen könnten. Die Stabilisierung des N-Terminus des Proteins mittels einer intermolekularen elektrostatischen Bindung zu K28 könnte ein generelles strukturelles Motiv für alle Mutanten darstellen. Diese Beobachtung steht möglicherweise im Zusammenhang mit der erhöhten Toxizität der Mutanten im Vergleich zur Wildform, wo die elektrostatische Bindung zu K28 intramolekular anstatt intermolekular vorliegt.

Eine weitere Amyloidstudie folgt in **Kapitel 3**, jedoch für eine andere Art von Amyloiden: in den letzten Jahren wurde die Hypothese, dass Peptidhormone und Neuropeptide vorübergehend als Amyloiden gespeichert werden, immer weiter unterstützt. Diese reversibele funktionelle Amyloiden werden gemäss den neuesten Erkenntnissen in sogenannte Dichter-Kern Vesikel zusammengepackt, die als Speicher in Sekretionszellen dienen. Somatostatin (SST) ist ein Beispiel solcher Peptidhormone, das sowohl in aggregierter als auch in löslicher Form physiologisch vorkommt. Die Proteinstruktur von menschlichem SST-14 im Faltungskontext eines Amyloiden wurde mittels Festkörper-NMR mit Drehung im magischen Winkel (magic-angle spinning; MAS) mit atomarer Auflösung bestimmt. Zusätzlich zu

rastertransmissions-elektronenmikroskopischen Daten wird die vollständige Rückgratsresonanzzuordnung in diesem Kapitel präsentiert. Anschließend werden auf Dipolarkopplung bezogene Experimente zur Bestimmung von spektral eindeutigen Distanzbeschränkungen mit einer Sekundärstrukturvoraussage kombiniert, welche auf chemischen Verschiebungen zweiter Ordnung und Torsionswinkelbeschränkungen basiert. Sämtliche Daten führen zur in diesem Kapitel vorgelegten Proteinstruktur.

In **Kapitel 4** werden sowohl ^{13}C- als auch ^{1}H-detektierte Experimente präsentiert. Beide Detektionsweisen werden miteinander verglichen, sowohl allgemein als auch im Zusammenhang mit Proteinen, welche mit dem Hepatitis-B-Virus (HBV) verwandt sind. HBV ist ein sphärisches behülltes Virus, dessen Genom nur wenige Gene enthält: die HBV-Hüllproteine S (Klein), M (Medium), und L (Groß) welche zusammengefasst als das HBV-Ober-flächenantigen bezeichnet werden, das Kernprotein (Cp), die Polymerase (P), und das X-Protein (HBx). Dieses Kapitel präsentiert Strukturanalysen in Bezug auf das Hüllprotein S und das Kernprotein mit (Cp183) und ohne (Cp149) C-terminalen Bereich.

Protonendetektion wird in Kombination mit ^{31}P-detektierten Experimenten zur Erforschung der Wechselwirkung zwischen Proteinen und Nukleinsäuren (ATP-Analoga und Desoxyribonukleotiden der DNA) angewandt. Protein-Nukleinsäurewechselwirkungen sind nicht nur in energiefreisetzenden Reaktionen wie zum Beispiel der ATP-Hydrolyse sondern auch beim Auslesen, Verlängern, Zusammenpacken oder Reparieren des Genoms von großer Bedeutung. Obwohl sie oft auch mittels Röntgenkristallographie analysiert werden können, sind zusätzliche Analysemethoden für nicht-kristalline Anordnungen unverzichtbar. In **Kapitel 5** wird beschrieben, wie Festkörper-NMR Protein-Nukleinsäurewechsel-wirkungen mittels ortsaufgelöster ^{1}H- und ^{31}P-detektierten Spektren feststellen und einordnen kann. Die Sensitivität der chemischen Verschiebung der Protonen aufgrund der nicht-kovalenten Wechselwirkungen, die mit diesen molekularen Erkennungsprozessen verbunden sind, wird zur Erforschung des chemischen Bindungszustands ausgenutzt; eine Information die nicht direkt aus Röntgenkristallographiedaten ermittelt werden kann. Trotz seiner herausfordernden Größe, wird die Methode auf die dodekamerische DnaB Helikase (669 kDa) in einer Komplexverbindung zu ADP:AlF$_4^-$:DNA angewandt.

Zum Abschluss dieser Dissertation wird in **Kapitel 6** die ^{1}H-Detektion mittels Festkörper-NMR aus einer eher methodologischer Sicht in Bezug auf MAS und Auflösung untersucht. Die spektrale Auflösung ist von größter Bedeutung, um die strukturellen und dynamischen Inhalte des NMR-Spektrums zugänglich zu machen. Die Protonendetektion in Festkörper-NMR wurde durch das Aufkommen immer höherer MAS-Frequenzen (heutzu-

tage über 100 kHz) ermöglicht. In diesem Zusammenhang ist es sinnvoll sich zu überlegen, welche Vorteile weiterführende Untersuchungen immer höherer MAS Frequenzen mit sich bringen. Diese Frage wird in diesem Kapitel angegangen, indem auf ein Proteinkomplex von archäale RNA Polymerase-Untereinheiten 4 und 7 MAS-Frequenzen von bis zu 150 kHz angewendet werden. Hierfür wird ein Rotor mit einem Aussendurchmesser von 0.5 mm und einem Proteininhalt von etwa 170 µg verwendet. Bei Drehfrequenzen von 100 und 150 kHz findet man, dass sich die totale Linienbreite und Signalintensität um einen Faktor 1.23 ± 0.05, respektive 1.48 ± 0.13 verbessern. Aufgrund dieser Resultate und weiteren Überlegungen, die in diesem Kapitel angestellt werden, wird die Schlussfolgerung gezogen, dass eine weitere Entwicklung und Erforschung höherer MAS-Frequenzen sinnvoll und lohnend ist.

1. General introduction

From archaea to eukaryotes, all living cells in nature contain proteins. In fact, proteins make up the majority of a cell's dry mass, corresponding to around 2 billion proteins in a single human cell [1]. Almost all functions of a cell are executed by proteins and their wide array of functions includes import and export of small molecules into and out of the cell, signal reception and intracellular relay, transport, replication, stability and so forth. To a large extent, our understanding of the molecular mechanisms behind a protein's function are based on three-dimensional structural models detailing the atomic coordinates of a protein in its native fold.

The structure of a protein is defined on four levels (Figure 1.1). Modern peptide sequencing can efficiently identify the **primary** protein structure: the sequence of amino-acid residues in the polypeptide chain [2]. In addition, an estimate of a protein's **secondary** structure can be made in a timespan in the order of a few hours with the use of infrared (IR) or Raman spectroscopy for instance, or circular dichroism (CD). Such methods are particularly useful to monitor conformational changes in the general fold of the protein as a result of temperature changes, mutations, or binding reactions [3, 4]. However, they do not give residue-specific information. True understanding of the **tertiary** and **quaternary** structure of a protein, as well as its function, requires information at atomic resolution. Since the advent of structural biology, X-ray crystallography and solution-state NMR spectroscopy have been robust protein structure determination methods to provide such data.

Since those beginnings, the 'palette' of structural biology methods has expanded. In particular, some proteins may be difficult to crystallise (as is often the case for membrane proteins) and/or too large for solution-state NMR, when a decreased molecular tumbling rate reduces T_2 relaxation times and results in line broadening. In such cases, magic-angle spinning (MAS) solid-state NMR can offer a solution. Indeed, solid-state NMR spectroscopy has proven its merit in the field, notably for structure determination of amyloid fibrils [6–11]

Primary

```
1     MFEARLVQGS ILKKVLEALK DLINEACWDI SSSGVNLQSM DSSHVSLVQL
51    TLRSEGFDTY RCDRNLAMGV NLTSMSKILK CAGNEDIITL RAEDNADTLA
101   LVFEAPNQEK VSDYEMKLMD LDVEQLGIPE QEYSCVVKMP SGEFARICRD
151   LSHIGDAVVI SCAKDGVKFS ASGELGNGNI KLSQTSNVDK EEEAVTIEMN
201   EPVQLTFALR YLNFFTKATP LSSTVTLSMS ADVPLVVEYK IADMGHLKYY
251   LAPKIEDEEG S
```
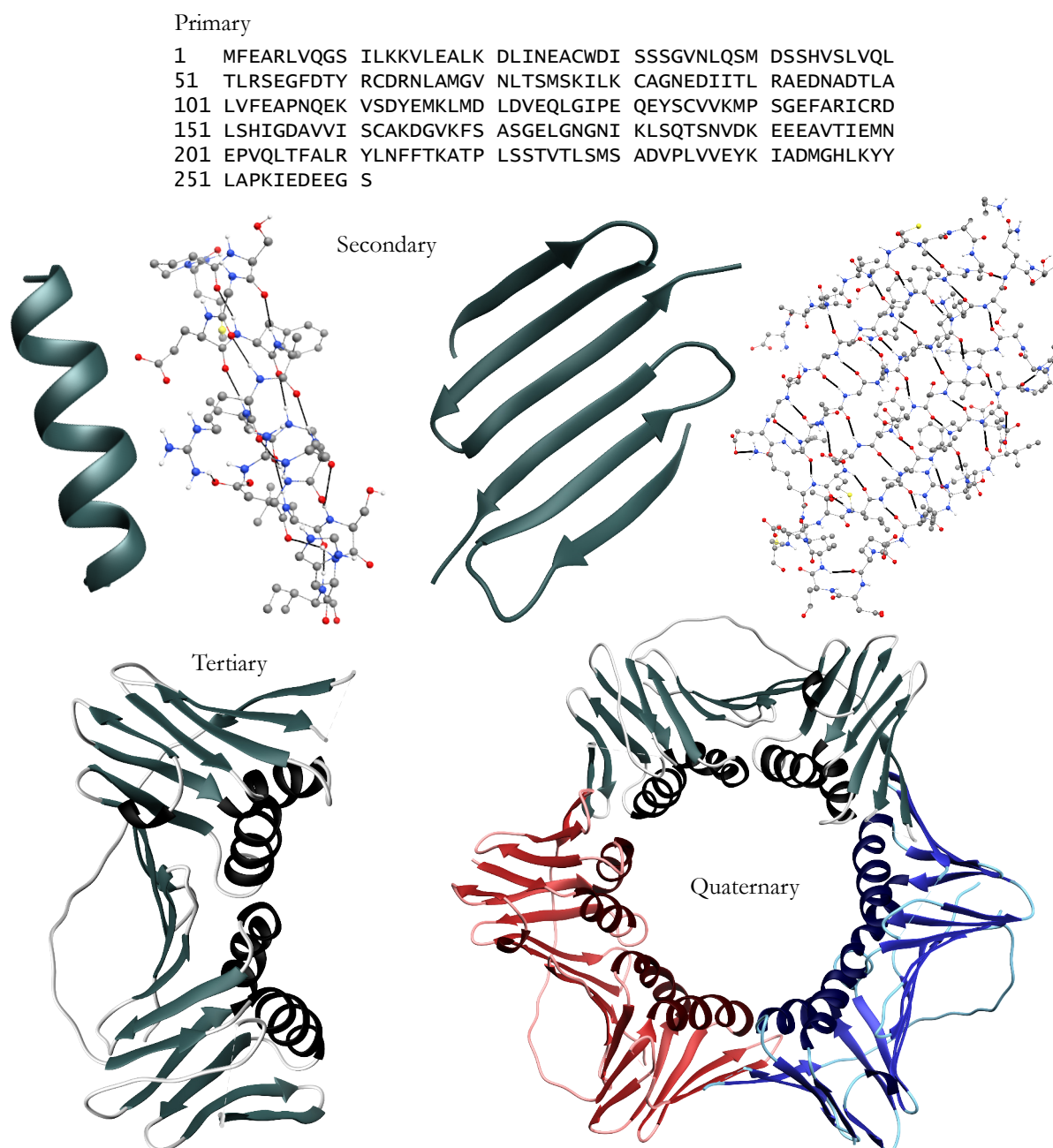

Figure 1.1. *From primary to quaternary protein structure.* The amino-acid sequence of *Proliferating cell nuclear antigen* (PCNA) counts 261 residues in its polypeptide chain, which form the **primary** structure. Various sections of the sequence assemble into local **secondary** structure elements, here exemplified by an α-helix on the left and a β-sheet on the right. Individual β-strands in the sheet are represented by arrows in the ribbon representation. The ball-and-stick models of both secondary structure elements show hydrogen bonds as black lines, illustrating typical bonding-patterns for each. In the **tertiary** structure, all these structure elements are packed together into one or several domains that may represent the mature and functional protein. In some cases, the functional protein consists of several polypeptide chains arranged into a **quaternary** structure. The two additional chains in the quaternary structure of PCNA are coloured in shades of red and blue. (Data used for this figure were obtained from PDB entry 1AXC [5])
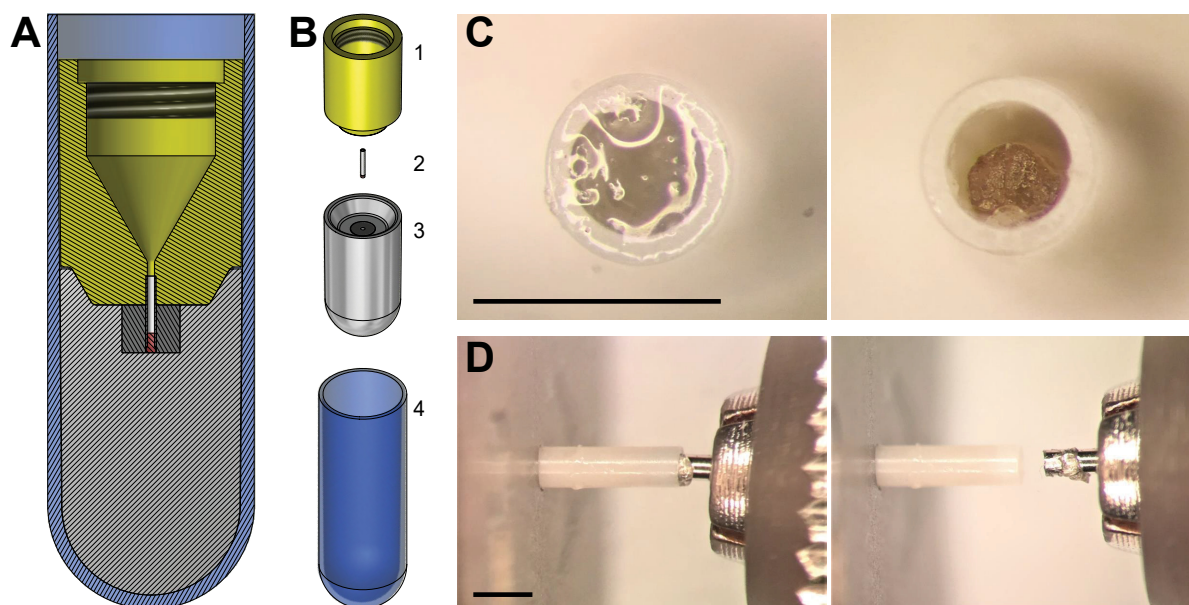
Figure 1.2. *Filling process of a 0.7 mm solid-state NMR rotor.* A schematic representation of the custom designed filling tool is given in **(a)**. The tool is used for protein sedimentation in a swing-out rotor of an ultracentrifuge. The separate components shown in **(b)** represent **1:** a composite polymer funnel which fits **2:** the 0.7 mm rotor, and seals tightly around it under pressure of ultracentrifugation. **3:** A metal rotor holder is designed to withstand the high centrifugal forces during sedimentation, while **4:** a sleeve holds the components together and fits exactly in the swing-out bucket of the ultracentrifuge. The technical drawings are by courtesy of Andreas Hunkeler. **(c)** Typical examples of resulting protein sediments after ultracentrifugation. The consistency of the protein sediment can be probed, as shown in **(d)**, with a tool that serves to free up enough space inside the rotor to fit the drive cap that closes the rotor. Scale bars represent 1 mm.

as also represented in part of this thesis (Chapters 2 and 3). An advantage of solid-state NMR over X-ray crystallography is that a highly concentrated protein solution can be centrifuged directly into the NMR sample container (a rotor), to form a gel-like protein sediment, as illustrated in Figure 1.2. Besides crystallography and NMR, following a recent 'resolution revolution' [12], (cryo-)electron microscopy (EM) forms an important addition to high-resolution structure methods and has already produced many valuable results (examples of which include [13–17]).

Further contributions to structure elucidation may for instance come from dynamic nuclear polarisation (DNP) [18, 19], small-angle neutron or X-ray scattering (SANS/SAXS) [20], mass spectrometry [21], or in-cell NMR [22]. In the future, continued technical developments may even allow for protein structure determination by single-molecule X-ray diffraction with the use of highly brilliant X-ray free-electron lasers (XFELs). In conclusion, it seems

Name	AA count	MW$_{monomer}$ / kDa	Units	MW$_{assymbly}$ / kDa	Chapter
GB1	56	6.15	1	6	
Ubiquitin	76	8.56	1	9	
Rpo4/7	294	33.52	1	34	6
PCNA	261	28.77	3	86	1
DnaB	488	55.73	12	669	5
HET-s(218-289) *	72	7.83	500	3915	6
HBV Cp149	149	16.85	240	4044	4
HBV Cp183	183	21.06	240	5054	4
Amyloid-β1-40 *	40	4.32	2000	8640	2
Somatostatin-14 *	14	1.64	6000	9840	3

* Assuming a fibril of 1000 layers, which corresponds to roughly 0.5 µm (*cf.* Figure 3.2)

Table 1.1. *Molecular weights of selected proteins and protein assemblies.* AA count gives the number of amino-acid residues in a single chain. The molecular weight (MW) is given for the monomer and the protein assymbly respectively.

almost inevitable that the future of structural biology will be characterised by collaborations and a methodologically hybrid approach to structure determination [11, 23, 24].

1.1. Large protein assemblies

This thesis presents various aspects of solid-state NMR, applied to a range of biological systems. What most of them have in common is that they are "large protein assemblies". This is, of course, a relative term, but to place it in context, Table 1.1 lists the molecular weights of the protein assemblies discussed in this thesis. These are compared to the system presented in Figure 1.1 (PCNA), and two model systems commonly used in structural biology in general, and NMR in particular: ubiquitin [25–28], and GB1 [29–31].

As explained in Figure 1.1, several polypeptide chains can combine into the quaternary structure of a final functional protein. PCNA was chosen to illustrate this, because it also contains relatively equal amounts of secondary-structure elements: coils, α-helices, and β-strands. Yet, it also relates functionally to one of the "large assemblies" discussed in this thesis: the DnaB helicase. Both PCNA and the DnaB helicase are essential in DNA replication, a process that is further discussed and illustrated in Chapter 5. The DnaB helicase forms a dodecameric quaternary structure. With a peptide chain of 488 residues, that brings its

quaternary assembly to 669 kDa; a molecular weight that is two orders of magnitude larger than ubiquitin and GB1.

Another class of proteins that forms large supramolecular assemblies is represented by viral capsids. Capsids are protein shells that protect the viral genome and deliver it to the host cell, amongst other functions, and in many cases show icosahedral symmetry [32]. As shown in Table 1.1, two variations of the hepatitis B virus core protein (HBV Cp149 and 183) assemble into protein capsids of 4 and 5 MDa respectively. As discussed in Chapter 4, there is in fact another large supramolecular assembly that forms around the viral genome: the capsid envelope. This is not listed in the table, since the envelope is not a purely polypeptide-based structure. Instead, it consists of many HBV surface antigen proteins embedded in a lipid membrane. While the envelope is a large viral construct, it is consequently not strictly a protein assembly.

A final category that appears several times in the lower half of Table 1.1 is that of amyloids. Amyloid fibrils are protein aggregates formed by macromolecular assemblies of multiple copies of the same polypeptide chain. While the number of copies in the above-described capsid assemblies is often precisely defined for a particular virus capsid, the number of copies in an amyloid fibril depends on its length. Although it can be influenced by growth conditions [33], the length of a fibril is not a strongly regulated characteristic. In contrast, monomer stacking along the so-called fibrillar axis is strictly repetitive, leading to unbranched elongated fibrils. β-strands run perpendicularly to the fibrillar axis and the fibril structure is stabilised by hydrogen bonds between the strands. This common feature of amyloids is known as the cross-β motif [34, 35]. Further details concerning amyloids are discussed in Chapters 2 and 3.

The protein assemblies discussed in this thesis have a molecular weight near, or in the megadalton range. However, it is important to realise that some of these large assemblies consist of many repeats of relatively short polypeptide chains. Although the repeating unit is not always entirely symmetric (*e.g.* the subunits forming HBV capsids), such repetition reduces the complexity of NMR spectra. Nevertheless, also in terms of the number of amino-acid residues in the single polypeptide chain, DnaB is almost an order of magnitude larger than the model system GB1. In conclusion, while it is primarily the assemblies that are large, some of the protein monomers are large too, and the term appropriately describes the work presented in this thesis.

1.2. Interactions in solid-state NMR

The spectroscopic observations made in this thesis are based on a set of well-described interactions, extensively discussed in [36–43]. Summarising these interactions, one may separate them to be either between a nuclear spin k and a magnetic field

$$\hat{\mathscr{H}}^{(k,B)} = \hat{\vec{\mathrm{I}}}_k \underline{A}^{(k,B)} \vec{B} \,, \tag{1.1}$$

or between two spins k and l

$$\hat{\mathscr{H}}^{(k,l)} = \hat{\vec{\mathrm{I}}}_k \underline{A}^{(k,l)} \hat{\vec{\mathrm{I}}}_l \,. \tag{1.2}$$

Here, tensor $\underline{A}$ describes the strength and angular dependency of the interaction; $\vec{B}$ is the magnetic field vector; and $\hat{\vec{\mathrm{I}}}$ contains the x, y and z spin operators. This thesis only describes interactions for nuclei with spin quantum number $\frac{1}{2}$ (which excludes quadrupolar coupling), and the spin operators can consequently be represented by the Pauli matrices

$$\hat{\mathrm{I}}_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{\mathrm{I}}_y = \frac{i}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{\mathrm{I}}_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.3}$$

All Hamiltonians in this chapter are given in units of angular frequency, where $\hat{\mathscr{H}} = \frac{1}{\hbar}\hat{\mathrm{H}}$.

1.2.1. The Zeeman Hamiltonian

By far the largest interaction in NMR is the interaction of a nuclear spin with a static magnetic field $\vec{B}_0$. This interaction depends only on the strength of the magnetic field and the gyromagnetic ratio γ_k of a nuclear isotope k (concerning the isotopes discussed in this thesis, for $k = \left[{}^1\mathrm{H}, {}^{13}\mathrm{C}, {}^{15}\mathrm{N}, {}^{31}\mathrm{P}\right]$, $\gamma_k = [267.522, 67.283, -27.116, 108.291] \cdot 10^6 \ \mathrm{rad} \cdot \mathrm{s}^{-1} \cdot \mathrm{T}^{-1}$). In such a static field, the energy difference between the two energy levels available to a single nuclear spin-half $\hat{\vec{\mathrm{I}}}_k$ is given by the Larmor frequency:

$$\omega_0 = -\gamma_k \vec{B}_0 \,. \tag{1.4}$$

By convention in NMR, $\vec{B}_0$ is assumed to be along the z-axis such that $\vec{B}_0 = (0,0,\mathrm{B}_0)$. Then, the Zeeman hamiltonian is given by the Larmor frequency and the z-component of the spin operator $\hat{\vec{\mathrm{I}}}$ (Equation 1.3):

$$\hat{\mathscr{H}}_Z = \omega_0 \hat{\mathrm{I}}_z \,. \tag{1.5}$$

1.2.2. Chemical shift

In the above, the described interaction was for a 'bare' nucleus in a magnetic field. However, a nucleus is usually surrounded by a cloud of electrons. This oscillating probability density of charge will give rise to a small induced magnetic field that adds to the interaction of the static magnetic field B_0 with the nucleus. As a result, the resonance frequency of that particular nucleus (Equation 1.4) will shift slightly, depending on its electronic environment. The effective magnetic field is the sum of B_0 and the induced local field. The latter is described by the product of B_0 and the chemical-shift tensor $\underline{\sigma}$:

$$\vec{B}_{\text{ind}} = \underline{\sigma}\vec{B}_0 = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix} . \tag{1.6}$$

The locally induced magnetic field at a nucleus is the first example of information that NMR provides about the structure of a molecule, and appears in the chemical-shift Hamiltonian as:

$$\hat{\mathcal{H}}_{\text{CS}} = \omega_0 \left(\sigma_{xz}\hat{I}_x + \sigma_{yz}\hat{I}_y + \sigma_{zz}\hat{I}_z \right) . \tag{1.7}$$

For high-field magnets, the Zeeman interaction generally is orders of magnitude larger than the chemical-shift interaction. In this case it is useful to define an axis system that rotates around the z-axis at the Larmor frequency ω_0. In this 'rotating frame', terms of Equation 1.7 that are proportional to $\hat{I}_x$ and $\hat{I}_y$ will also rotate around the z-axis with ω_0, and average to 0. In this high-field or *secular* approximation, the chemical-shift Hamiltonian simplifies to:

$$\hat{\mathcal{H}}_{\text{CS}} = \omega_0 \sigma_{zz}\hat{I}_z . \tag{1.8}$$

While NMR spectra principally involve resonance frequencies in Hertz, by convention, their axes report a chemical shift δ in parts-per-million (ppm), relative to a reference compound, and independent of the applied magnetic field B_0. This is defined as:

$$\delta\left[\text{ppm}\right] = 10^6 \cdot \frac{\nu_{\text{obs}} - \nu_{\text{ref}}}{\nu_{\text{ref}}} . \tag{1.9}$$

Here, ν_{obs} is the observed resonance frequency, ν_{ref} is the resonance frequency of the reference compound, and δ is the reported chemical shift. Clearly, in this definition the chemical shift δ of the reference compound will be zero.

1.2.3. Radio-frequency pulses

The equilibrium magnetisation is described by a net magnetic moment of the nuclear spins in the sample, aligned along the z-axis of the static magnetic field. At equilibrium, this magnetisation is 'in the shadow' of the external field, and to observe the nuclear spins, the net magnetic moment must be tilted into the x/y-plane. In this orientation, the magnetic moment will precess around the z-axis at the Larmor frequency (Equation 1.4). To achieve such a tilt, a radio-frequency (rf) pulse is applied with the use of a small coil around the sample, which generates a field commonly referred to as the B_1 field. Since B_1 is a pulse-generated field it is time-dependent. Like the rotating frame defined for the chemical-shift interaction described above, an interaction frame can be defined that rotates around the z-axis with ω_{rf}, the frequency of the B_1 field. In this interaction frame, the B_1 field is time-independent, and in analogy to the Larmor frequency, a nutation frequency ω_1 can be defined for a spin k as:

$$\omega_1 = -\gamma_k B_1 \,. \tag{1.10}$$

A transformation to the interaction frame also removes time-dependencies of the x- and y-components of magnetic moment precessing in the x/y-plane. The radio-frequency Hamiltonian is then given by:

$$\begin{aligned}\hat{\mathscr{H}}_{RF} &= \omega_1 \hat{\vec{I}} \\ &= \omega_1 (\hat{I}_x \cos\psi + \hat{I}_y \sin\psi) \,. \end{aligned} \tag{1.11}$$

where ψ contains the phase of the rf-pulse.

Assuming an ideal on-resonance pulse ($\omega_{rf} = \omega_0$) with phase $\psi = 0$, the resulting B_1 field generates a rotation of the magnetisation vector around the interaction-frame x-axis with the nutation frequency ω_1. The time τ and the nutation frequency ω_1 determine the flip angle ϑ. With the right τ, a flip angle $\vartheta = 90°$ will bring the full magnetisation in the x/y-plane and generate maximum signal intensity.

1.2.4. Dipolar coupling

The above-described interactions are all of the general form of Equation 1.1, concerning interactions between a nuclear spin and a magnetic field. This section and the section below will instead discuss interactions between spins. The one described here is an interaction

between the magnetic dipole moments of two spins that interact through space. By contrast, in the section below, the described interaction is mediated by the chemical bond. Both are further examples (besides the chemical shift) of information that NMR can give about the structure of the molecule.

The dipolar coupling Hamiltonian between two spins k and l is often defined in terms of the "dipolar alphabet"

$$\hat{\mathscr{H}}_{\mathrm{D}} = \frac{\mu_0}{4\pi} \frac{\gamma_k \gamma_l \hbar}{r_{kl}^3} (A + B + C + D + E + F) \, , \tag{1.12}$$

where μ_0 is the vacuum permeability, γ_k and γ_l are the gyromagnetic ratios of the two nuclei k and l involved in the interaction, $\hbar$ is the reduced Planck constant and r_{kl} is the internuclear distance between spins k and l. The single terms A to F are given by

$$A = 2\hat{\mathrm{I}}_{kz}\hat{\mathrm{I}}_{lz}\frac{(3\cos^2\vartheta - 1)}{2}$$

$$B = -\frac{1}{2}(\hat{\mathrm{I}}_k^+\hat{\mathrm{I}}_l^- + \hat{\mathrm{I}}_k^-\hat{\mathrm{I}}_l^+)\frac{(3\cos^2\vartheta - 1)}{2}$$

$$C = (\hat{\mathrm{I}}_k^+\hat{\mathrm{I}}_{lz} + \hat{\mathrm{I}}_{kz}\hat{\mathrm{I}}_l^+)\frac{3\sin\vartheta\cos\vartheta e^{-i\varphi}}{2}$$

$$D = (\hat{\mathrm{I}}_k^-\hat{\mathrm{I}}_{lz} + \hat{\mathrm{I}}_{kz}\hat{\mathrm{I}}_l^-)\frac{3\sin\vartheta\cos\vartheta e^{i\varphi}}{2}$$

$$E = \frac{1}{2}\hat{\mathrm{I}}_k^+\hat{\mathrm{I}}_l^+\frac{3\sin^2\vartheta e^{-2i\varphi}}{2}$$

$$F = \frac{1}{2}\hat{\mathrm{I}}_k^-\hat{\mathrm{I}}_l^-\frac{3\sin^2\vartheta e^{2i\varphi}}{2} \, ,$$

where $\hat{\mathrm{I}}^+ = \hat{\mathrm{I}}_x + i\hat{\mathrm{I}}_y$ and $\hat{\mathrm{I}}^- = \hat{\mathrm{I}}_x - i\hat{\mathrm{I}}_y$; and the polar angles ϑ and φ describe the orientation of the internuclear vector in the laboratory frame. The terms C, D, E, and F contain transverse spin-operator components (regarding $\hat{\mathrm{I}}_x$ or $\hat{\mathrm{I}}_z$) that are time-dependent in an interaction frame that rotates around the z-axis. These terms can be neglected in the secular approximation, and for two spins k and l of the same isotope, the homonuclear dipolar coupling Hamiltonian reduces to:

$$\hat{\mathscr{H}}_{\mathrm{D}}^{\mathrm{homo}} = \frac{\mu_0}{4\pi} \frac{\gamma_k \gamma_l \hbar}{r_{kl}^3} \frac{(3\cos^2\vartheta - 1)}{2} \left(2\hat{\mathrm{I}}_{kz}\hat{\mathrm{I}}_{lz} - \frac{1}{2}(\hat{\mathrm{I}}_k^+\hat{\mathrm{I}}_l^- + \hat{\mathrm{I}}_k^-\hat{\mathrm{I}}_l^+)\right) \, . \tag{1.13}$$

For two isotopes that are not the same (*e.g.* ^{1}H and ^{13}C), one uses separate rotating frames for both nuclei, since they have different Larmor frequencies. In this case, the transverse spin-operator components of the so-called "flip-flop" term $\left(\hat{I}_k^+ \hat{I}_l^- + \hat{I}_k^- \hat{I}_l^+\right)$ do not cancel each other and the B-term of the dipolar alphabet becomes non-secular like the terms C-F. For the heteronuclear dipolar coupling Hamiltonian in the secular approximation, this leaves:

$$\hat{\mathcal{H}}_D^{\text{hetero}} = \frac{\mu_0}{4\pi} \frac{\gamma_k \gamma_l \hbar}{r_{kl}^3} \frac{(3\cos^2 \vartheta - 1)}{2} 2\hat{I}_{kz}\hat{S}_{lz} \, . \tag{1.14}$$

The two symbols $\hat{I}$ and $\hat{S}$ underline the spin operators to concern different isotopes.

1.2.5. J coupling

As mentioned above, the J-coupling is a spin-spin interaction that is mediated by the electrons in the covalent bond. Analogous to Equation 1.2, the J-coupling Hamiltonian can be described as:

$$\hat{\mathcal{H}}_J = 2\pi \hat{\vec{I}}_k \underline{J}^{(k,l)} \hat{\vec{I}}_l \, . \tag{1.15}$$

The J-coupling contains anisotropic contributions described by the tensor $\underline{J}^{(k,l)}$. Most of the time, the anisotropic contributions are small and hard to distinguish from the dipolar coupling. For this reason, often only the isotropic part of $\underline{J}^{(k,l)}$ is considered. In contrast to the dipolar coupling, the isotropic value of the J-coupling is not zero. Again, a homonuclear and heteronuclear case can be separately defined in secular approximation as:

$$\hat{\mathcal{H}}_J^{\text{homo}} = 2\pi J_{\text{iso}}^{(k,l)} \hat{\vec{I}}_k \cdot \hat{\vec{I}}_l \tag{1.16}$$

$$\hat{\mathcal{H}}_J^{\text{hetero}} = 2\pi J_{\text{iso}}^{(k,l)} \hat{I}_{kz}\hat{S}_{lz} \, . \tag{1.17}$$

The full Hamiltonian

All interactions described above combine linearly into the full Hamiltonian of a spin-$\frac{1}{2}$ system as

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_Z + \hat{\mathcal{H}}_{CS} + \hat{\mathcal{H}}_{RF} + \hat{\mathcal{H}}_D + \hat{\mathcal{H}}_J \, . \tag{1.18}$$

1.3. Magic-angle spinning

Molecules in a solution tumble rapidly around their centre of gravity. This reduces anisotropic NMR interactions (*e.g.* the chemical shift and dipolar coupling) to their isotropic values. In a solid, every molecular orientation has an equal chance to be present and these orientations do not average. Spectroscopically, this is observed as 'powder patterns' in the spectrum, in which separate nuclear spins produce a manyfold of overlapping resonances. To reduce anisotropic chemical shift and dipolar interactions to their isotropic values, the sample can be mechanically rotated in the static magnetic field. This is called magic-angle spinning (MAS) [44–46]. An example of the effect of MAS on solid-state NMR spectra is given in Figure 1.3.
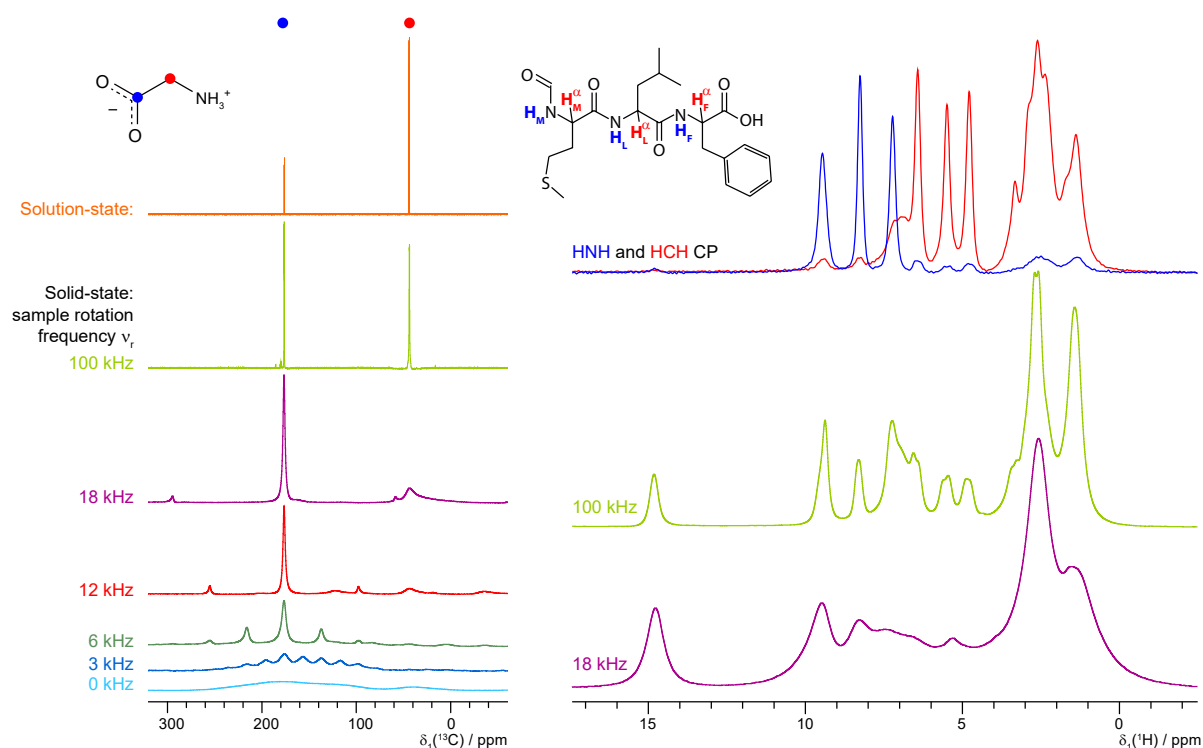
Figure 1.3. *MAS applied to* ^{13}*C-detected glycine and* ^{1}*H-detected fMLF.* On the **left**, solid-state NMR ^{13}C spectra at various MAS frequencies are compared to a solution-state spectrum of the single amino acid glycine (in the solid-state spectra, glycine is chemically stabilised as glycine ethylester). The carbonyl carbon and its resonance are indicated with a blue dot, and the Cα carbon and resonance are marked with a red dot. On the **right**, the two highest MAS frequencies of the ^{13}C spectra (18 and 100 kHz) are shown for ^{1}H spectra of the tripeptide fMLF. Double-CP 1D spectra through either ^{15}N (in blue) or ^{13}C (in red) facilitate resolution of H^{N} and H$^{\alpha}$ protons.

The chemical-shift anisotropy (CSA) as well as the dipolar coupling are second rank tensors. Their angular dependence is given by $(3\cos^2\vartheta - 1)/2$, where ϑ is the angle between the applied magnetic field and the principal axis of the interaction. During MAS, the sample is rotated at an angle $\vartheta_m = \arccos\frac{1}{\sqrt{3}}$. At this magic angle ϑ_m of about 54.7° with respect to the static magnetic field B_0, the angular dependence of the second rank tensors reduces to zero ($3\cos^2\vartheta_m - 1 = 0$). As a result, MAS at a frequency that is significantly larger than the interaction reduces the interaction to its isotropic value to first-order approximation. Further details concerning the orders of this approximation are discussed in Section 1.3.2 below.

From the dipolar coupling Hamiltonian (Equation 1.12), it is straightforward to see that the interaction becomes stronger for spins with larger gyromagnetic ratios γ. Consequently, a ^{1}H-^{1}H dipole coupling is much stronger than that between for example two carbons, or a ^{13}C-^{1}H spin pair. In the spectra on the left of Figure 1.3, the Cα carbon has more protons in its vicinity than the carbonyl carbon. As a consequency, it needs faster MAS to be resolved, as the spectra clearly show. A further illustration of this point is given in the right half of the figure, where primarily ^{1}H-^{1}H couplings strongly broaden the signals, and even a sample rotation frequency of 100 kHz is not enough to resolve all resonances. What is not directly apparent from the spectra is that not only the resolution is affected by the anisotropic interactions, but also the signal-to-noise (SNR) ratio. While the spectra on the left were recorded in the order of a few minutes for the solid-state spectra, and a few seconds for the solution-state spectrum, the two proton spectra on the right (at 18 and 100 kHz) recorded for 1.5 h each.

1.3.1. Spherical-tensor notation

As discussed above, some interactions in solid-state NMR contain both isotropic and anisotropic contributions. The isotropic contributions (*e.g.* to the chemical shift or J-coupling Hamiltonians) are invariant under rotation, and indeed all NMR experiments as well as MAS introduce rotations of the spatial and/or the spin part of the Hamiltonians. With this in mind, in contrast to a definition in a Cartesian coordinate system, it can be useful to define Hamiltonians in a symmetry-adapted spherical notation. The irreducible spherical-tensor

notation [47] provides such a symmetry-adapted basis, in which the Hamiltonian is defined as a sum over all relevant interactions i, the rank of the tensor ℓ, and its components m

$$\hat{\mathscr{H}} = \sum_i \sum_{\ell=0}^{2} \sum_{m=-\ell}^{\ell} (-1)^m A_{\ell,m}^{(i)} \hat{\mathscr{T}}_{\ell,-m}^{(i)} . \tag{1.19}$$

In this definition, the spatial parts A of the Hamiltonian are separated from the spin-tensor operators $\hat{\mathscr{T}}$. In an NMR experiment, each of these becomes time-dependent for different reasons: MAS of the sample generates a spatial time-dependence, while the B_1 field that originates from a pulsed and time-dependent rf-field modulates the spin part $\hat{\mathscr{T}}$ of the Hamiltonian.

1.3.2. Time-dependent Hamiltonians

MAS and rf-irradiation during the NMR experiment affect the Hamiltonian as described in Equation 1.19, such that the time-dependent Hamiltonian is given by:

$$\hat{\mathscr{H}}(t) = \sum_i \sum_{\ell=0}^{2} \sum_{m=-\ell}^{\ell} (-1)^m A_{\ell,m}^{(i)}(t) \hat{\mathscr{T}}_{\ell,-m}^{(i)}(t) . \tag{1.20}$$

Besides MAS, random molecular motion also modulates the spatial part of the Hamiltonian, but in a non-coherent way. This causes relaxation of longitudinal and transverse magnetisation (*e.g.* T_1, T_2, $T_{1\rho}$). Relaxation is not further discussed here in the context of time-dependent Hamiltonians, but is for example well described in [48].

For MAS around an axis at an angle $\vartheta_m = \arccos\frac{1}{\sqrt{3}}$ to the direction of B_0, anisotropic interactions are averaged and the Hamiltonian in the secular approximation reduces to:

$$\hat{\mathscr{H}}(t) = \sum_i \sum_{\ell=0}^{2} A_{\ell,0}^{(i)}(t) \hat{\mathscr{T}}_{\ell,0}^{(i)} . \tag{1.21}$$

In the special case where the time-dependence of the Hamiltonian is periodic with a cycle time τ_c such that a given $\hat{\mathscr{H}}(t)$ will reappear after an *integer* number of cycle times, and the observation of the Hamiltonian is synchronised to an integer multiple of that cycle time, Average-Hamiltonian Theory (AHT) [47, 49] facilitates the calculation of a time-independent average Hamiltonian $\overline{\hat{\mathscr{H}}}$.

To establish a relation between $\hat{\mathscr{H}}(t)$ and $\overline{\hat{\mathscr{H}}}$, usually the Magnus expansion [50] is employed to obtain an average Hamiltonian in different orders of approximation

$$\overline{\hat{\mathscr{H}}} = \overline{\hat{\mathscr{H}}}^{(1)} + \overline{\hat{\mathscr{H}}}^{(2)} + \overline{\hat{\mathscr{H}}}^{(3)} + \ldots \tag{1.22}$$

where the first three contributions to this series expansion are

$$\overline{\hat{\mathscr{H}}}^{(1)} = \frac{1}{\tau_c} \int_0^{\tau_c} dt_1 \hat{\mathscr{H}}(t_1) \tag{1.23}$$

$$\overline{\hat{\mathscr{H}}}^{(2)} = \frac{-i}{2\tau_c} \int_0^{\tau_c} dt_2 \int_0^{t_2} dt_1 \left[\hat{\mathscr{H}}(t_2), \hat{\mathscr{H}}(t_1) \right] \tag{1.24}$$

$$\overline{\hat{\mathscr{H}}}^{(3)} = \frac{-1}{6\tau_c} \int_0^{\tau_c} dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \left[\hat{\mathscr{H}}(t_3), \left[\hat{\mathscr{H}}(t_2), \hat{\mathscr{H}}(t_1) \right] \right]$$
$$+ \left[\left[\hat{\mathscr{H}}(t_3), \hat{\mathscr{H}}(t_2) \right], \hat{\mathscr{H}}(t_1) \right]. \tag{1.25}$$

Applying this for example to the chemical-shift Hamiltonian leads to a first order contribution of $\overline{\hat{\mathscr{H}}}_{CS}^{(1)} = -\omega_0 \sigma_{iso} \hat{I}_z$, while all higher-order terms are zero. This means that the first-order average Hamiltonian is in fact equal to the full average Hamiltonian, if the experimental observation is synchronous to $n\tau_c$. With $\overline{\hat{\mathscr{H}}}_{CS} = \overline{\hat{\mathscr{H}}}_{CS}^{(1)}$, the chemical-shift Hamiltonian forms an example of an interaction that is *heterogeneous* with respect to MAS.

By contrast, applying AHT to the dipolar coupling Hamiltonian under MAS at ϑ_m leads to a vanishing first-order term, but a non-vanishing second-oder term for multiple dipolar couplings (for example, for three spins k, ℓ, and m that are not equivalent, $\hat{T}_{2,0}^{k,\ell}$ and $\hat{T}_{2,0}^{k,m}$ do not commute). As a result, higher-order average Hamiltonian contributions remain and the interaction is *homogeneous* with respect to MAS. Such remaining higher-order terms typically contribute to the spectrum in the from of a broadening of the observed signals.

1.3.3. Spectral proton linewidth

With respect to the broadening mentioned above and as a prediscussion for Chapter 6, it is interesting to summarise the various contributions to the width Δ of a proton signal at full-width half max (FWHM) in a solid-state MAS NMR spectrum.

First off, the ^{1}H linewidth contributions can be separated in coherent, incoherent and inhomogeneous contributions [28, 43, 51], in which the combination of the coherent and

incoherent contributions is jointly referred to as the homogeneous contributions. This relation can be summarised as the following equation:

$$\Delta^{tot}(\nu_r) = \Delta^{coh}(\nu_r) + \Delta^{incoh}(\nu_r) + \Delta^{inhomo} = \Delta^{homo}(\nu_r) + \Delta^{inhomo}. \tag{1.26}$$

Coherent contributions arise from homonuclear interactions, which include the anisotropic (Equation 1.7) and isotropic (Equation 1.8) chemical shift as well as homonuclear proton dipolar couplings (Equation 1.13). Inserting these terms in the average Hamiltonian as described in Equation 1.22 leaves a first-order term that consists of the isotropic chemical shift only. Higher-order terms include cross-terms between anisotropic and isotropic chemical shifts as well as multiple terms for homonuclear dipolar couplings. These second- and third-order terms scale with the inverse of the MAS frequency (or the frequency squared for the third-order term) and vanish in the limit of infinitely fast spinning. However, at finite MAS frequencies they give rise to coherent line broadening in the spectrum [28, 41, 51].

Incoherent contributions are due to molecular motions. These cause stochastic processes such as Redfield-type T_2 relaxation or chemical exchange, which form a more fundamental limit to the maximum resolution that can be achieved in a solid-state NMR spectrum. Nevertheless, for slow molecular motions (which are the relevant motions for line broadening), R_2 rates decrease for increasing MAS frequencies. The homogeneous broadening, that is defined as the sum of the coherent and incoherent contributions (Equation 1.26), can be experimentally quantified with spin-echo decay curves [52].

Finally, inhomogeneous effects are related to sample inhomogeneities such as disorder and crystal defects leading to chemical-shift distributions, or B_0 inhomogeneities and contributions from an imperfect probe shimming [28, 51]. While homogeneous interactions scale down with the MAS frequency, the inhomogeneous contributions remain uninfluenced by MAS [28]. Some of the material presented in this thesis implies that at 150 kHz MAS, the homogeneous and inhomogeneous linewidth contributions in large protein assemblies are approximately of the same size. This will be discussed further in Chapter 6.

2. Alternative salt bridge formation in Amyloid-β

The work presented in this chapter was part of a collaboration with Anja Böckmann, and builds on work from Schütz *et al.* [53] in which the 3D structure of the Osaka mutant form (E22Δ) of Amyloid-β1-40 was determined.

This chapter is based on the following publication:

[54] Schledorn, M., Meier, B. H., & Böckmann, A. (2015). Alternative salt bridge formation in Aβ − a hallmark of early-onset Alzheimer's disease?. *Frontiers in Molecular Biosciences*, 2, 14.

Structure calculations [MS]; Analysis of results [MS, BHM, AB]; Writing of the original manuscript [AB]; Discussion and editing of final text [MS, AB, BHM]; Study design and supervision [BHM, AB].

2.1. Introduction

Amyloid-β (Aβ) in its different conformations and aggregations states is a central player in the amyloid-cascade hypothesis for Alzheimer's disease (AD) [55]. Brain-derived or synthetic aggregates were shown to propagate between cells when injected into transgenic mice [56, 57]. It is not established whether the Aβ fibrils or smaller oligomers, or both, are the toxic species [58–60]. The existence of mutants that lead to early-onset AD [61] offers an opportunity to characterise the conformational space available to the Aβ peptide in fibrillar or oligomeric state. Amyloids in general, and Aβ in particular, are known for their ability to form a number of polymorphs [62], and it has been proposed that they are at the origin of different phenotypes of the disease, reminiscent of the appearance of strains in prion disease [63–65]. Different mutant forms of Aβ have been identified involving a single-residue deletion or substitution at, or adjacent to, residue E22, e.g. the Flemish A21G [66], Arctic E22G [67, 68], Dutch E22Q [69, 70], Italian E22K [71, 72] and Iowa D23N [73] mutants, as well as the Osaka deletion mutation E22Δ [74, 75]. These mutants cause early-onset AD, and they display different toxicity profiles in *in vitro* and *in vivo* studies compared to wild-type Aβ1-40 and Aβ1-42.

Knowledge of the atomic-resolution 3D structure is central for the understanding of the molecular basis underlying the amyloid diseases, and solid-state NMR is a powerful method to determine and characterise structures of amyloid fibrils at atomic resolution [6] and to map the conformational space available to these proteins. In addition to the Osaka mutant structure [53], several wild-type Aβ fibril polymorphs have been characterised [76–81], and a model for the Iowa mutant was presented in 2015 [82].

Despite a large body of literature, there is presently not enough high-resolution structural data available to establish a detailed structure-toxicity relationship for amyloids in general, and Aβ in particular [83]. However, virtually complete chemical-shift information for all 40 residues has recently become available for two wild-type forms and a mutant of Aβ1-40. This information is used here to suggest that, on the basis of a comparison of NMR chemical-shift values between the Osaka mutant [53, 84] and one of the wild-type polymorphs [80], the wild-type peptide does indeed have the ability to assume the fold established for Aβ1-40 E22Δ [53], albeit with a modified in-out pattern of the amino-acid residues in the loop comprising residues 20-31. Furthermore, it is illustrated below how other early-onset Aβ mutants can in principle form a similar fold. Hence, we speculate that the key feature of the mutant folds is the formation of an *inter*molecular salt bridge, attaching the N-terminal residues

to the fibril core, as opposed to the wild-type protein where the situation is substantially different, as residue K28 was experimentally shown to be involved in an *intra*molecular salt bridge [80]. As a consequence, the N-terminus is less tightly attached, which impacts properties of formed fibrils, and potentially those of prefibrillar states [85, 86].

2.2. Results and discussion

2.2.1. The wild-type protein can adopt a fold similar to that of the Osaka mutant

Chemical-shift comparisons between proteins represent a sensitive means to assess differences and similarities in 3D structures, as shift differences point to conformational differences for the observed residue. Comparison of wild-type Aβ chemical shifts described by Lu *et al.* [80] to those of the Aβ1-40 E22Δ mutant [53] (Figure 2.1A) reveals that the two forms display almost coinciding shifts for large parts of the protein, with noticeable exceptions in two regions where the differences surpass the 1 ppm ($\delta_{(^{13}C)}$) limit, which we consider to be a significant change.

   To illustrate the location of these changes, we plotted the absolute value of the mean chemical-shift differences from Figure 2.1A, displayed in Figure 2.1B, on the Aβ E22Δ mutant structure in Figure 2.1C: residues with similar shifts are colored in blue, others in red. Blue residues cover the entire hydrophobic core, as well as the largest parts of β-sheets 1 and 2. Importantly, the two regions displaying larger chemical-shift differences are distant in sequence, but spatially contiguous if plotted on the Aβ1-40 E22Δ mutant structure. They concern the loop as well as the N-terminus attached to the loop via the E3-K28 salt bridge. For comparison, the other wild-type polymorph with complete assignments [78] shows substantial chemical-shift differences spanning the entire sequence, with only few and possibly fortuitous exceptions (see Appendix Figure A.1).

2.2.2. Secondary structure comparison between WT Aβ and the Osaka mutant

Secondary chemical shifts (the difference to random-coil shifts [87, 88]) inform on protein secondary structures. The secondary shift differences ΔδCα-ΔδCβ can be evaluated according to a pattern of three consecutive negative values corresponding to a β-strand, four
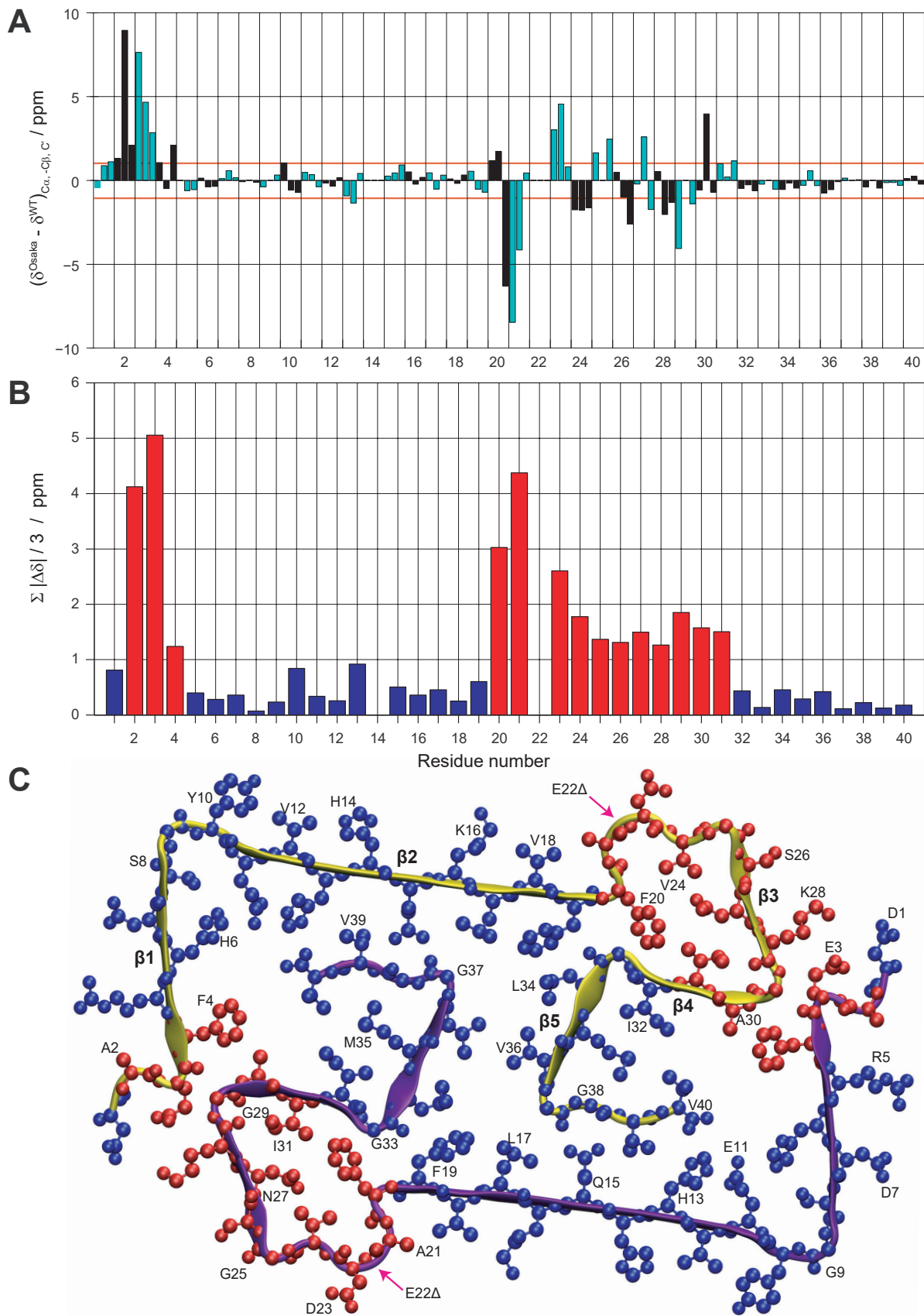
Figure 2.1. (a) Difference of the chemical shifts Cα, -Cβ and C', between the (continued on next page)
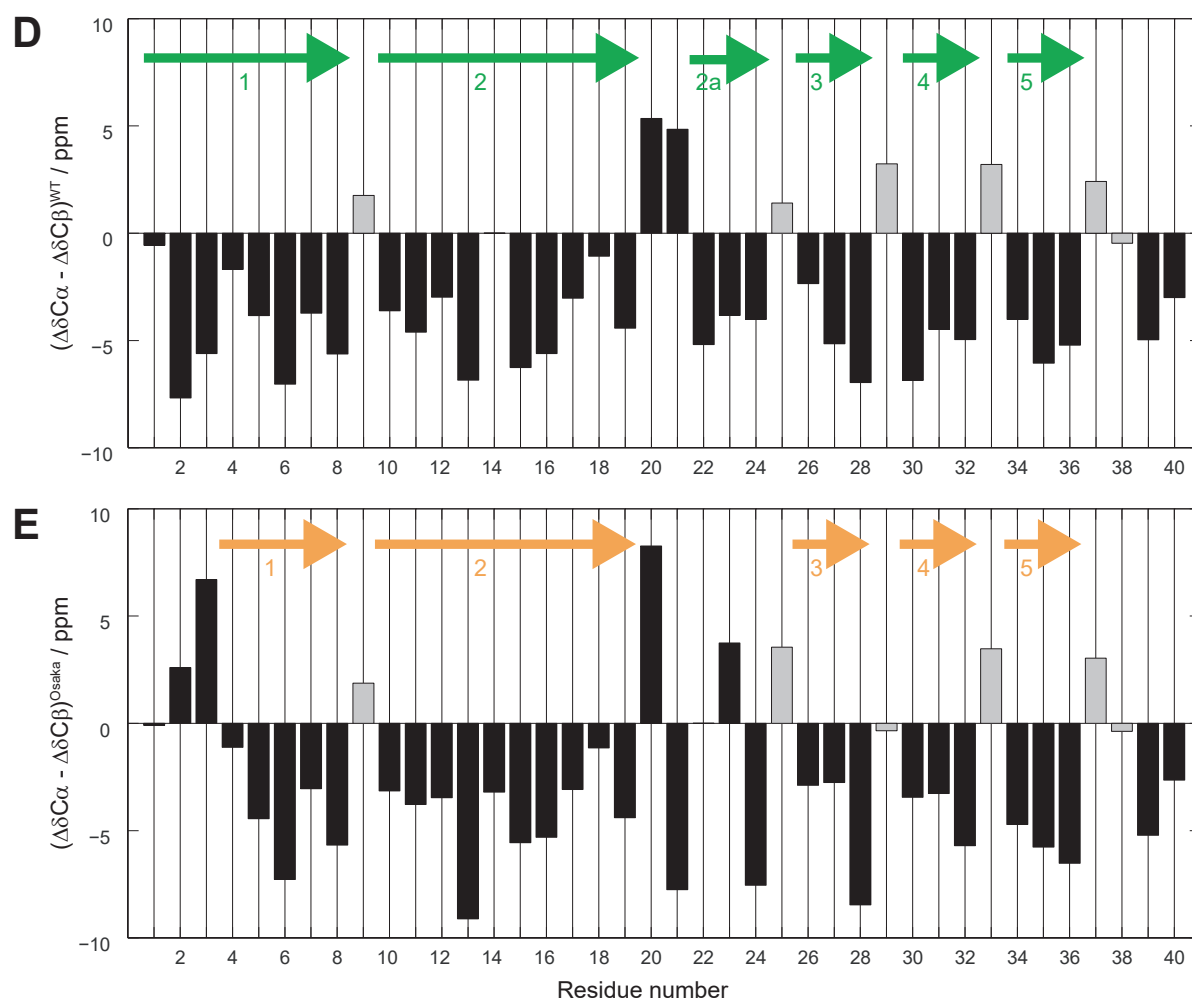
Figure 2.1. (continued from previous page) Osaka mutant [53] and wild-type Aβ1-40, Patient I of Lu *et al.* [80]. Cβ is given as negative values to highlight differences in secondary structure propensity such that all chemical-shift differences point in the same direction: where the Osaka mutant shows less β-strand propensity, all three values are positive, and *vice versa*. The chemical shifts were re-referenced to yield the smallest possible differences. Odd residues are shown in cyan, and even ones in black. The red line indicates a limit of 1 ppm, that we consider to be a significant change. **(b)** Average absolute values of the mean chemical shift differences between wild-type Aβ1-40 [80] and the Osaka mutant [53], calculated as $\left(\left| \delta C\alpha^{WT} - \delta C\alpha^{Osaka} \right| + \left| \delta C\beta^{WT} - \delta C\beta^{Osaka} \right| + \left| \delta C'^{WT} - \delta C'^{Osaka} \right| \right) \div 3$. Deviations ≥1 ppm are plotted in red, and <1 ppm in blue. **(c)** Representation of chemical-shift differences from (b) plotted on the Aβ1-40 E22Δ structure (PDB ID 2MVX), with the same colour coding as the bar chart. **(d-e)** Differences between Cα and Cβ secondary chemical shifts for (d) wild-type Aβ1-40 of Patient I from [80], and (e) for the Osaka mutant [53]. Arrows indicate locations of β-sheets according to the rule that three negative values in a row are indicative for this secondary structure element, with the exception of glycines (marked in grey).

	1	10	20	30
Sequence	DAEFRHDSG	YEVHHQKLVF	FAEDVGSNKG	AIIGLMVGGVV
β-Strands	1----	2---------	2a- 3--	4-- 5--
Orientation of first residue	In	Out	Out In	Out In
	Residues 5-19 and 32-40		**Residues 1-4 and 20-31**	
Distance restraints	Inter/Intra-molecularly and spectrally unambiguous restraints		None	
TALOS+ angles	BMRB 25289 (E22Δ)		BMRB 19009 (WT)	

Table 2.1. Information used as input for WT model building. Distance restraints from the Osaka mutant were used between the residues which show small chemical-shift differences (blue in Figure 2.1C). A list is given in Appendix Table A.1. TALOS+ angles were calculated from the given chemical shift depositions and the restraints for CYANA [89] calculations had a range over the top 10 database matches ± 15°. Residues without dihedral angle restraints were: all glycines; residues D1 and V40 being the terminal residues; and residues A21 and V24 which had a warning in the TALOS+ prediction. Although F20 also has a warning in the TALOS+ prediction, the restraint was still used, since secondary chemical shifts in both the WT [80] and the Osaka mutant [53] convincingly indicate that the β-strand that starts from Y10 ends at F19. F20 thus should discontinue the strand and have the same orientation as F19. β-strands were established according to Figure 3 and Figure S8 in Schütz *et al.* [53]. Strand 2a does not exist in the Osaka mutant, but is suggested in the WT form by the chemical shifts [80] (Figure 2.1D). The orientation of the sidechains is defined by the hydrogen bonding pattern between the different layers of the fibril and the resultant orientation of the first residue in each strand is given as inward or outward with respect to the core of the dimer.

positive values in a row indicating α-helical secondary structure, and the remainder turns or loops. $\Delta\delta C\alpha$-$\Delta\delta C\beta$ of Aβ1-40 E22Δ is plotted in Figures 2.1D-E and the resulting pattern of β-sheets is found to be quite similar, with a couple of noticeable differences: for residues A2 and G3, where the mutant protein structure presents a turn, the WT protein chemical shifts indicate a straight β-strand, extending the strand which in Aβ1-40 E22Δ spans residues F4 to S8. In the WT peptide, the turn comprising residues F20-V24 in the mutant form is reduced to residues F20-A21, and an additional β-strand E22-V24 is present. We note from Figure 2.1 that the β-sheet S26-K28 is conserved, but that the individual chemical shifts (Figure 2.1A) differ significantly (>1 ppm). Residues F4-F19 and residues I30-V40 feature very similar shifts for the WT and the Osaka mutant fibril.

2.2.3. A speculative structure for a WT Aβ1-40 polymorph

We have noted that the monomer conformation for the residues mentioned above is very similar for the WT from the brain-seeded sample and the Osaka mutation despite the fact that the polymorph of the mutant form is described by a two-fold symmetry [53] and the WT polymorph by a three-fold symmetry [80]. In the following, we construct a two-fold model of a monomer in a WT fibril using the information provided in Table 2.1. The model mainly relies on the data coming from the mutant form, complemented by information from the brain seeded WT form for the residues showing chemical shift differences between the two forms. Details are given in the Materials and Methods, Section 2.4. The corresponding calculation did not result in any restraint violations, and the atomic-detail model is displayed in Appendix Figure A.2A. Figure 2.2B represents the region of interest, presented as a schematic drawing in order to emphasize its nature as a speculative model. When compared to the Osaka mutant structure shown in Figure 2.2A, the two β-sheets E22-V24 and S26-K28 in this model are turned around (inside/outside interchanged) to accommodate the K28-D23 salt bridge inside the loop (see Materials and methods, Section 2.4). The attachment of β-sheet 1 (Figure 2.1C) is now only mediated by the H6-E11-H13-V40 salt bridge network [53], as well as F4-A30 hydrophobic interactions, and lacks the strong electrostatic stabilisation by E3-K28. This would explain the dynamic nature of the N-terminal part of roughly 10 residues, which has systematically been observed in the WT forms. These residues were shown to display a certain flexibility in most polymorphs described (allowing for H/D exchange or giving rise to weak or missing signals in solid-state NMR spectra), and could only be sequentially assigned for Aβ1-40 E22Δ and three wild-type models [78, 80, 81]. Even in these, no unambiguous distance restraints could be established between the N-terminal segment and the fibril core. Only the Aβ1-40 E22Δ mutant distinguishes itself by a rigid N-terminal segment, structurally unambiguously defined by the intermolecular E3-K28 salt bridge. We speculate that this stabilising salt bridge is a direct consequence of the E22Δ mutation, and Section 2.2.4 below describes how it might also occur in other early-onset AD-related Aβ variants, where mutations are found in the loop around residues A21-D23.

2.2.4. Structural models for other mutant forms

A common feature shared by all the discussed early-onset mutants is that their amino-acid composition is such that the presence of negative charge in the fibril interior can be avoided, which in turn allows an outside orientation of Lys28 and a rigidification of
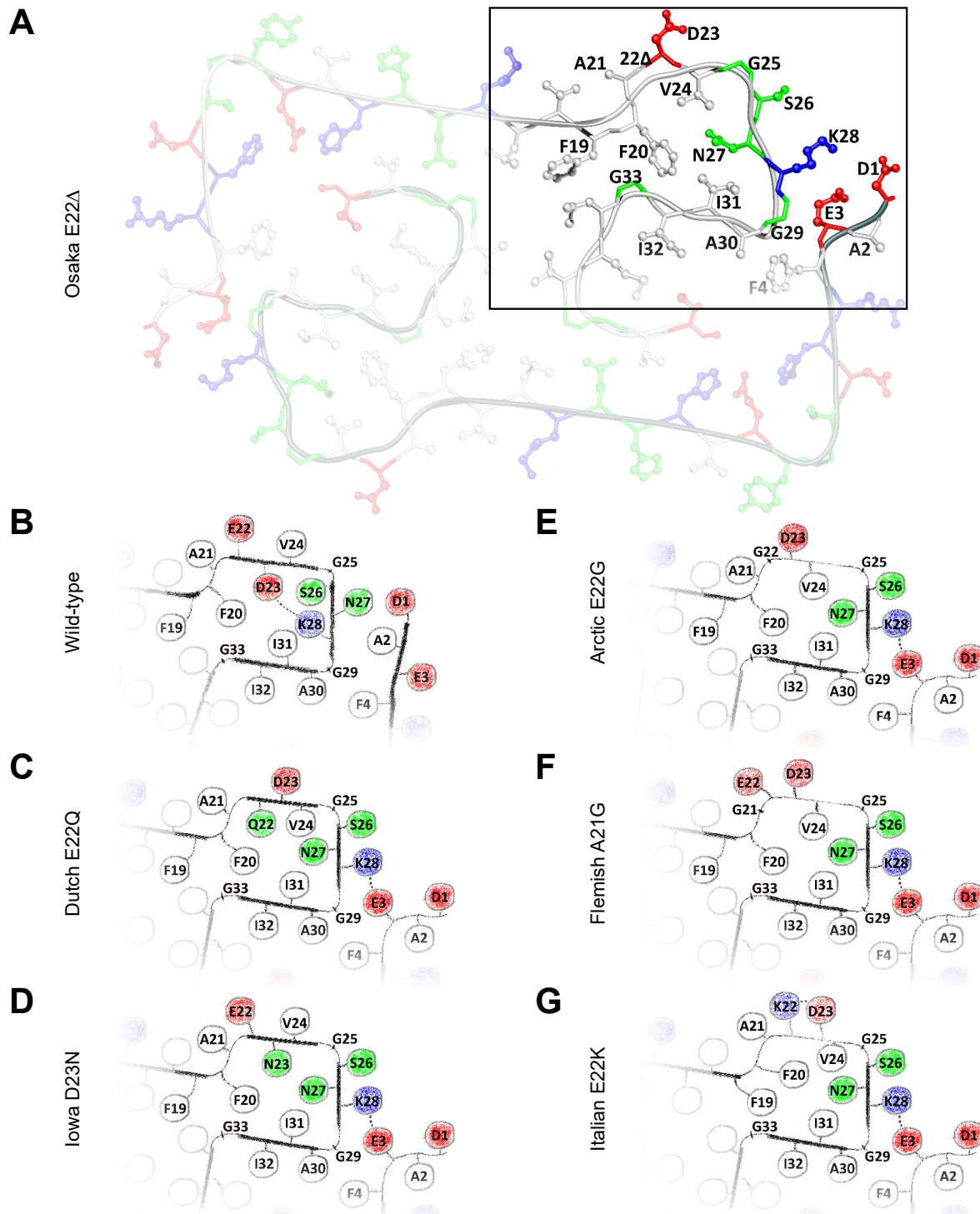
Figure 2.2. (a) The structure of the Osaka mutant E22Δ from Schütz *et al.* [53]. **(b)** Speculative model of
WT Aβ as described in the main text, and hypothetical models of **(c)** Dutch mutant E22Q; **(d)** Iowa mutant
D23N; **(e)** Arctic mutant E22G; **(f)** Flemisch mutant A21G; and **(g)** Italian mutant E22K. Models (c-g) are
built based on the Osaka mutant structure and additional considerations as described in the main text. All
these mutants favour, in this model, an outside orientation of K28 leading to an intermolecular salt bridge
to E3 of the opposing monomer in the fibril layer. Hydrophobic residues are coloured white, glycines and
polar residues are green, and charged residues are coloured blue when positive and red when negative.

the N-terminal β-strand by its attachment to the fibril core via the E3-K28 salt bridge. Hypothetical models for the different mutant forms are shown in Appendix Figure A.2 (for details of the calculations, see Section 2.4), and drawings of the regions of interest are shown in Figures 2.2C-G. The Dutch, Iowa, and Arctic mutants all result in the replacement of a negative charge of the wild-type peptide from either D23 or E22. The replacing residues in these mutants are polar or hydrophobic, which consequently easily accommodates a charge-free fibril interior in the loop encompassing residues 20-33. In the Flemish mutant, the A21G mutation introduces additional flexibility such that both E22 and D23 can have an outward orientation, and in the Italian mutant, replacement of E22 by a positive charge (K22) allows to stabilise D23 outside by a salt bridge with K22. The calculations carried out for the mutant forms Appendix Figure A.2 show that the above discussed hypothetical conformations can be realised without steric hindrance.

2.3. Conclusion

In summary, while the presented data cannot ascertain that K28 must be pointing outside in the early-onset Aβ mutants discussed here, our calculations show that there are no restraint violations, static clashes, or unbalanced charges inside the loop of residues 20-33 when K28 points outside. Consequently, the residue can engage in an energetically favourable salt bridge with the N-terminus of the protein. The Osaka polymorph fold is thus in principle accessible to the other mutant forms, and the possibility to assume this fold could be of importance to early onset. In any case, the models presented here provide presently testable hypotheses in mutational studies allowing to experimentally confirm the central role of K28 in the increased toxicity of mutant forms [66, 68–74]. More chemical-shift information and high-resolution structures of different polymorphs of WT and mutant fibrils of both Aβ1-40 and 1-42, as well as, hopefully, Aβ oligomers [90], will be needed to reliably establish a clear relationship between structure and biological activity of the protein.

	Wild-type	Dutch E22Q	Iowa D23N
Target function	$(9 \pm 4) \cdot 10^{-3}$	$(4.1 \pm 0.2) \cdot 10^{-2}$	$(2 \pm 3) \cdot 10^{-4}$
Distance constraints			
Mean violations (Å)	$(3.0 \pm 1.0) \cdot 10^{-4}$	$(2.0 \pm 0.1) \cdot 10^{-4}$	$< 10^{-5}$
Max. violations (Å)	$(9 \pm 4) \cdot 10^{-3}$	$< 10^{-5}$	$< 10^{-5}$
Dihedral angle constraints			
Mean violations (°)	$(1.8 \pm 0.7) \cdot 10^{-2}$	$(2.1 \pm 0.1) \cdot 10^{-2}$	$(1 \pm 1) \cdot 10^{-3}$
Max. violations (°)	$(1.4 \pm 0.5) \cdot 10^{-1}$	$(2.2 \pm 0.2) \cdot 10^{-1}$	$(1 \pm 1) \cdot 10^{-2}$
Average pairwise RMSD			
Backbone (Å)	0.60 ± 0.17	0.39 ± 0.13	1.2 ± 0.4
Heavy atom (Å)	0.87 ± 0.11	0.74 ± 0.09	1.5 ± 0.4

	Arctic E22G	Flemish A21G	Italian E22K
Target function	$(5 \pm 2) \cdot 10^{-5}$	$(3.4 \pm 1.1) \cdot 10^{-5}$	$(4 \pm 3) \cdot 10^{-4}$
Distance constraints			
Mean violations (Å)	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$
Max. violations (Å)	$< 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$
Dihedral angle constraints			
Mean violations (°)	$< 10^{-5}$	$< 10^{-5}$	$(2 \pm 1) \cdot 10^{-3}$
Max. violations (°)	$< 10^{-5}$	$< 10^{-5}$	$(2 \pm 1) \cdot 10^{-2}$
Average pairwise RMSD			
Backbone (Å)	1.0 ± 0.3	0.66 ± 0.11	1.1 ± 0.3
Heavy atom (Å)	1.3 ± 0.3	0.92 ± 0.08	1.4 ± 0.3

Table 2.2. CYANA [89] statistics for the structural models presented in Figure 2.2 and Appendix Figure A.2.

2.4. Materials and methods

2.4.1. Calculation of the model for the WT Aβ dimer

The calculation was performed as described for the "manual calculations" in [53]. For residues 2-4, and 20-31, TALOS+ angles [91] were calculated from the deposited chemical shifts [80]. Unambiguous distance restraints, as determined in Schütz *et al.* [53] for the mutant form, were used for the residues which show concomitant chemical shifts in both proteins (blue residues in Figure 2.1C). The list of distance restraints used is therefore a subset of the list given in [53], namely the ones between two residues from the range 1, 5-19, and 32-40, and is listed in Appendix Table A.1. The orientations of the β-strands were defined for β-sheets 1, 2, 4, and 5 in the same manner as in the mutant form (for the strand numbering, see Figure 2.1D-E). This input was not sufficient to distinguish between the possible orientations of the two β-strands spanning residues E22-V24 (strand 2a) and S26-K28 (strand 3) in the wild-type protein (Figure 2.1D) on the basis of restraint violations. The following additional considerations were used as constraints for the structure calculations: *(i)* The assigned sidechain Cγ/δ chemical shifts in the brain-seeded WT polymorph for D23 and E22 respectively show chemical-shift values typical of *charged* COO⁻ sidechain moieties [80]. *(ii)* The β-sheet structure of E22-V24 evidenced by the secondary chemical shifts (Figure 2.1D) observed for the brain-seeded polymorph implies that either E22 or D23 point to the inside of the fibril. *(iii)* The thus introduced E22 or D23 negative charge inside the fibril should be compensated by a positive charge in order to yield a stable fibril. Here, K28 is the only possible partner. From these three points, we can infer that S26-K28 shows an in-out-in pattern, allowing for charge compensation. This leaves the two options of an out-in-out or in-out-in orientation for E22-V24, but steric clashes between V24 and S26 (in an in-out-in orientation of E22-V24) strongly suggest an out-in-out pattern, with D23 pointing inside. Indeed, the D23-K28 salt bridge has been described in most WT models [80, 92].

2.4.2. Structural models for the mutants

For all mutant form models, dihedral angle restraints from Schütz *et al.* [53] were used and distance restraints from the same source were entered into CYANA [89] calculations, except for residues 20-30 and 2-4. Hydrogen bonding patterns were adopted from the Osaka mutant for the Arctic, Flemish, and Italian mutant. Like the WT model, the Dutch and Iowa mutant calculations included an additional β-sheet 2a for residues 22-24 (see Figure 2.1D).

Its orientation followed from two arguments: *(i)* we avoid uncompensated charges inside the core of the fibril, and *(ii)* we avoid steric clashes between residues V24 and S26. As a result, the Iowa mutant has an out-in-out pattern (Figure 2.2D and Appendix Figure A.2C), as also predicted in a recent model for the Iowa mutant [82]. While the Dutch mutant assumes an in-out-in pattern for residues Q22-V24 (Figure 2.2C and Appendix Figure A.2B), the Arctic mutant shows an outside-pointing D23 (Figure 2.2E and Appendix Figure A.2D), and in the Flemish mutant both E22 and D23 are oriented outwards and β-strand "2a" (E22-V24) is not present (Figure 2.2F and Appendix Figure A.2E). Finally, in the Italian mutant (Figure 2.2G and Appendix Figure A.2F), the formation of a K22-D23 salt bridge is forced by an additional restraint to the CYANA [89] calculation, in which electrostatic interactions are otherwise not considered. Without restraint violations, this residue pair can only be accommodated outside the loop due to limited space on the inside. The structural statistics for all mutant calculations as well as the wild-type are given in Table 2.2.

3. Molecular structure of *human* somatostatin-14 in the amyloid form

This work has been a collaboration with the group of Roland Riek, where Nadezhda Nespovitaya prepared fibril samples and performed EM.

The chapter is a preparation for the following manuscript:

Schledorn, M., Ravotti, F., Nespovitaya, N., Böckmann, A., Riek, R., & Meier, B.H. *(To be submitted)* Molecular structure of *human* somatostatin-14 in the amyloid form.

Protein expression, purification, and fibril formation [NN]; TEM and STEM [NN]; NMR spectroscopy [MS, FR]; Sequential resonance assignment [MS, FR]; MPL, TALOS, SCS, and structure calculations [MS]; Writing of the original manuscript [MS]; Study design and supervision [BHM, RR, AB]

3.1. Introduction

Mid-previous century, the hypothesis was established that the hypothalamus releases molecular entities which control hormone release from the anterior pituitary. In an undoubtedly exciting time where the identification of such species called for the collection of hypothalamic extract from about 500,000 sheep from abattoirs across France [93], it was the team around Roger Guillemin that identified the cyclic tetradecapeptide somatostatin(SST)-14 to inhibit growth hormone with convincing evidence [94] after McCann and co-workers had found early indications for the same inhibitive function [95].

The evolutionary highly conserved (Figure 3.1) peptide hormone that is of cyclic nature due to an intra-molecular disulphide bond between 3Cys and 14Cys [96, 97] was shown to have a plethora of functions after its initial discovery as a growth hormone-inhibiting hormone. It is a common feature for all of these functions to be inhibitive with respect to the function or expression level of other hormones, including well-known examples such as glucagon, insulin, norepinephrine, gastrin, neurotensin and renin [98]. Coinciding with a broad spectrum of physiological functions, SST-14 is present throughout the body and excreted in various tissue types. While native SST has a short half-life (1-3 min), several analogues have been developed for clinical use as anti-tumour and anti-inflammatory drugs, as well as treatment of diabetes and obesity [99].

The amino-acid sequence of SST-14 is `AGCKNFFWKTFTSC` and particularly the `NFFWK` motif of just five amino-acid residues is unique in the entire human genome for only somatostatin-14 and its related paralog cortistatin-17. The adjacent phenylalanine rings may *a priori* be an indicative motif for amyloid fibril formation through stacking of aromatic rings (often referred to as π-π interaction [100]), as postulated in [101]. Indeed, spontaneous fibril formation of SST-14 has been described [102], and interpreted as a possible storage mechanism in the form of dense-core secretory granules, common to a multitude of peptide hormones, while the formation of SST-14 amyloid fibrils *in vivo* as a reversible storage mechanism has been shown in rat hypothalamic tissue [103].

Besides the description as a functional amyloid, SST-14 has been linked to Aβ in the context of Alzheimer's disease (AD) [104–106], and the general concept of cross-talk between functional and disease-related amyloids, as would be the case for SST-14 and Aβ, has been suggested before [107, 108]. As described in Chapter 2, Aβ fibrils are linked to senile plaques in AD-affected brain tissue, but recent evidence suggests that soluble oligomeric pre-fibrillar forms of Aβ are the toxic species in AD development, rather than the matured fibrils [58, 109–
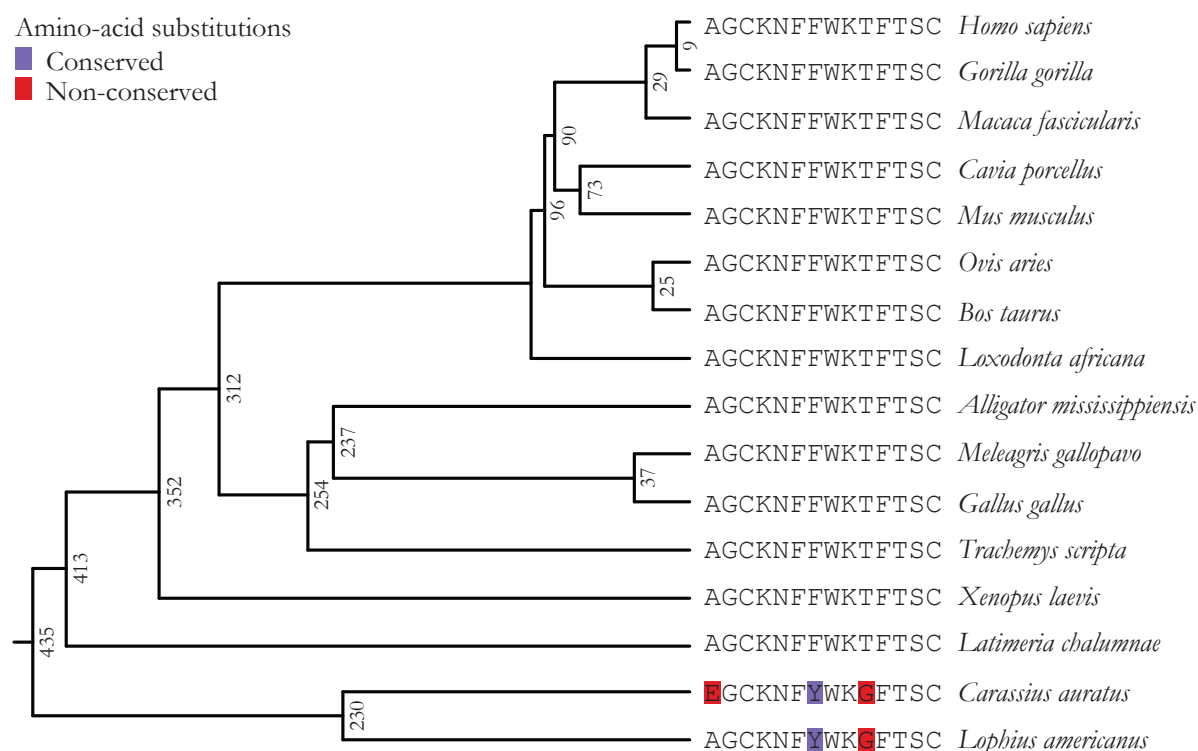
Figure 3.1. *Phylogenetic tree of the amino-acid sequence of somatostatin-14.* Multiple sequence alignment was performed using the CLUSTALW algorithm, based on a BLOSUM weight matrix (for amino-acid sequences) giving a penalty of 10.0 for gap opening and 0.05 for gap extension. Hierarchical clustering into the rooted tree was performed by an Unweighted Pair-Group Method with Arithmetic Mean (UPGMA) where branch length represents evolutionary distance, indicated as an estimated number in mya (million years ago) for the Most Recent Common Ancestor. The bottom two entries illustrate the development of a separate variant in some fish, where preproSST I is cleaved to SST-14 and preproSST II to SST-28. In other fish, and in most terrestrial vertebrae however, one preproSST is cleaved into both products and for all entries besides the bottom two, the 14 C-terminal amino acids of preproSST are a 100 % conserved from vertebrate reptiles and fish to mammals.

111]. Consequently, for the continued efforts towards therapeutic treatment of Aβ-mediated toxicity, it is of paramount importance to identify and characterise molecules that bind Aβ oligomers. Recent *interactome* analysis by mass spectrometry with Aβ monomers and oligomers as "bait" for interactions with human frontal lobe brain extract [112] identified SST-14 to selectively interact with Aβ oligomers. The interaction extends the fibril-forming lag phase, thereby inhibiting Aβ aggregation. Notably, it does so if, and only if at least one of the two binding partners Aβ and SST-14 is in pre-aggregated oligomeric form. As a basis for the further understanding of structure and function of SST-14, this chapter presents the molecular structure of an SST-14 monomer in the context of an amyloid fibril.

3.2. Results and discussion

3.2.1. Fibril formation

SST-14 occurs *in vivo* both as a soluble monomer and as fibrils, as for example shown in rat hypothalamic tissue [103]. To date (Oct 2019), the Protein Data Bank [113] contains one entry with a solution-state NMR structure of the SST-14 monomer [103]. Logically, since it is a solution-state structure, it is not in amyloid form. Yet, this form is important for storage in dense-core secretory granules prior to synaptic release [102], as well as for the interaction with Aβ for which it has been shown that at least one of the binding partners needs to be in non-monomeric state for the two to bind [106]. Therefore, this chapter investigates the SST-14 monomer in the fibrillar context and chemical shifts were recorded on fibrillated peptides (Figure 3.2). Experimental details, including fibril preparation of SST-14 in the presence of heparin, are described in Section 3.4.

3.2.2. Mass-per-length analysis

Besides the transmission electron micrographs for global characterisation of the fibril preparations (Figure 3.2), STEM data (Figure 3.3A) allow for an estimation of the mass-per-length (MPL) of the fibrils, and consequently of the number of monomers per fibril layer. It is not uncommon for fibrils to comprise more than one monomer per layer (*e.g.* Aβ [8, 9, 53, 78–80]), or even less than one monomer per layer (*i.e.* one monomer occupies several fibril layers), *e.g.* HET-s [6, 7] or PrP^{Sc} [114]. The data can be somewhat ambiguous, since *protofilaments* (the smallest fibril units) may bundle together (see also Figure 3.2B) to form larger fibrils, and heparin can be incorporated into these suprafibrillar assemblies.

A total of 66 data points from the STEM micrographs are summarised in Figure 3.3B. Three populations can be distinguished, with respective MPLs of 41 ± 3, 67 ± 4, and 87 ± 3 kDa·nm^{-1}, stemming from the aforementioned bundling of protofilaments. However, none of these can represent the MPL of the protofilament: considering the existence of a population of 67 ± 4 kDa·nm^{-1}, roughly in between the other two values, the population of 41 ± 3 kDa·nm^{-1} cannot possibly represent the protofilament. Instead, it should be at least about half that value, *i.e.* ~21 kDa·nm^{-1}. A separate population at this value may be absent in the data because it bundles together with at least one other protofilament. It seems unlikely that the protofilament would be even smaller, *e.g.* at an MPL around 10-11 kDa·nm^{-1}, considering the absence of populations around 55 and 77 kDa·nm^{-1}.
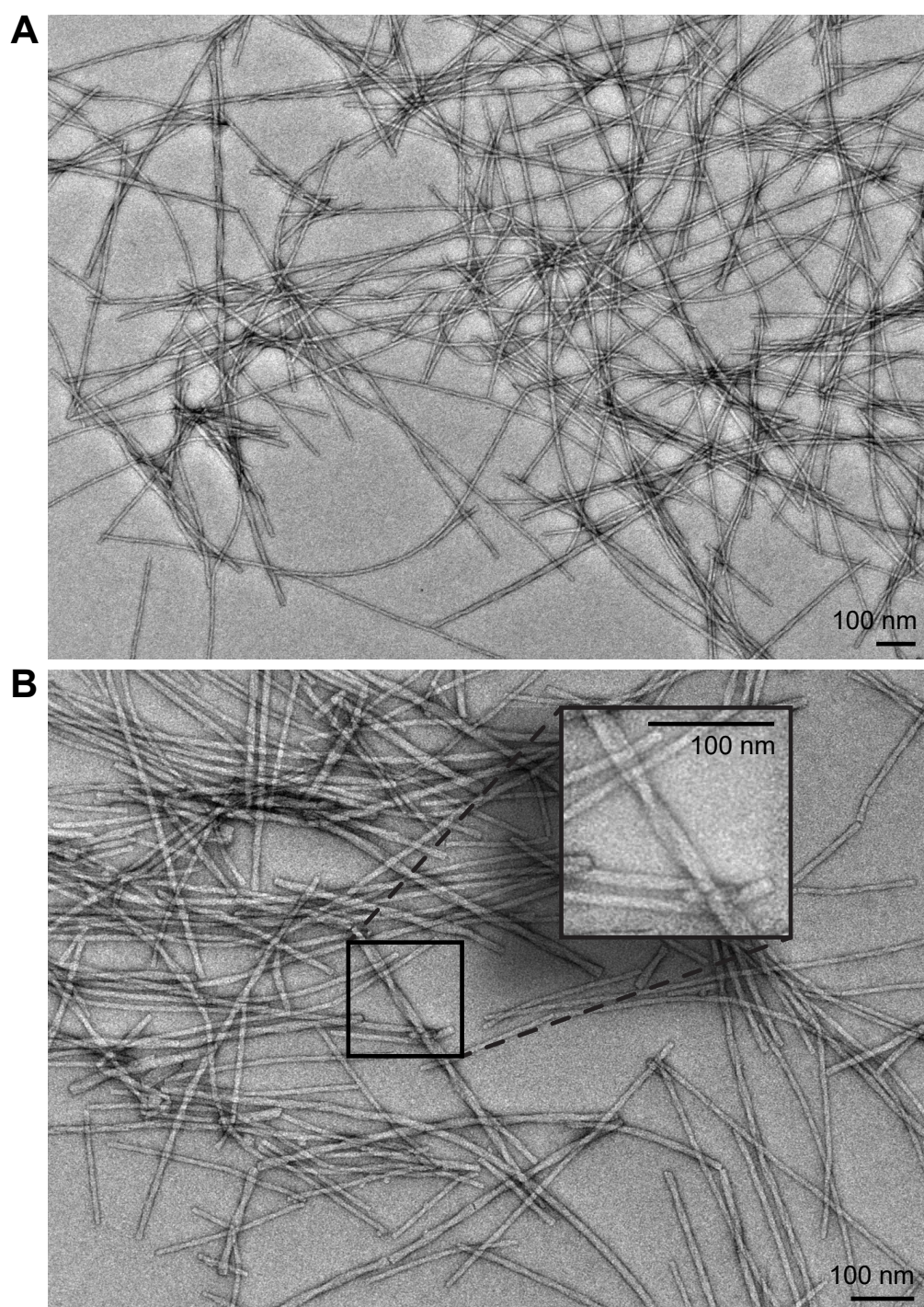
Figure 3.2. *Negative-stain bright-field TEM of diluted SST-14 fibrils.* Electron micrograph (**b**) is roughly 1.6 times more magnified than (**a**), and particularly in (**b**) there are several examples visible (see also the zoomed-in panel) of fibrils twisting around their fibrillar axis, exhibiting a fibril morphology distinct from some of the fibrils *e.g.* towards the bottom-right of figure (b). Note that the twist period is constant, *i.e.* the fibril length per twist, does not change within a single fibril.
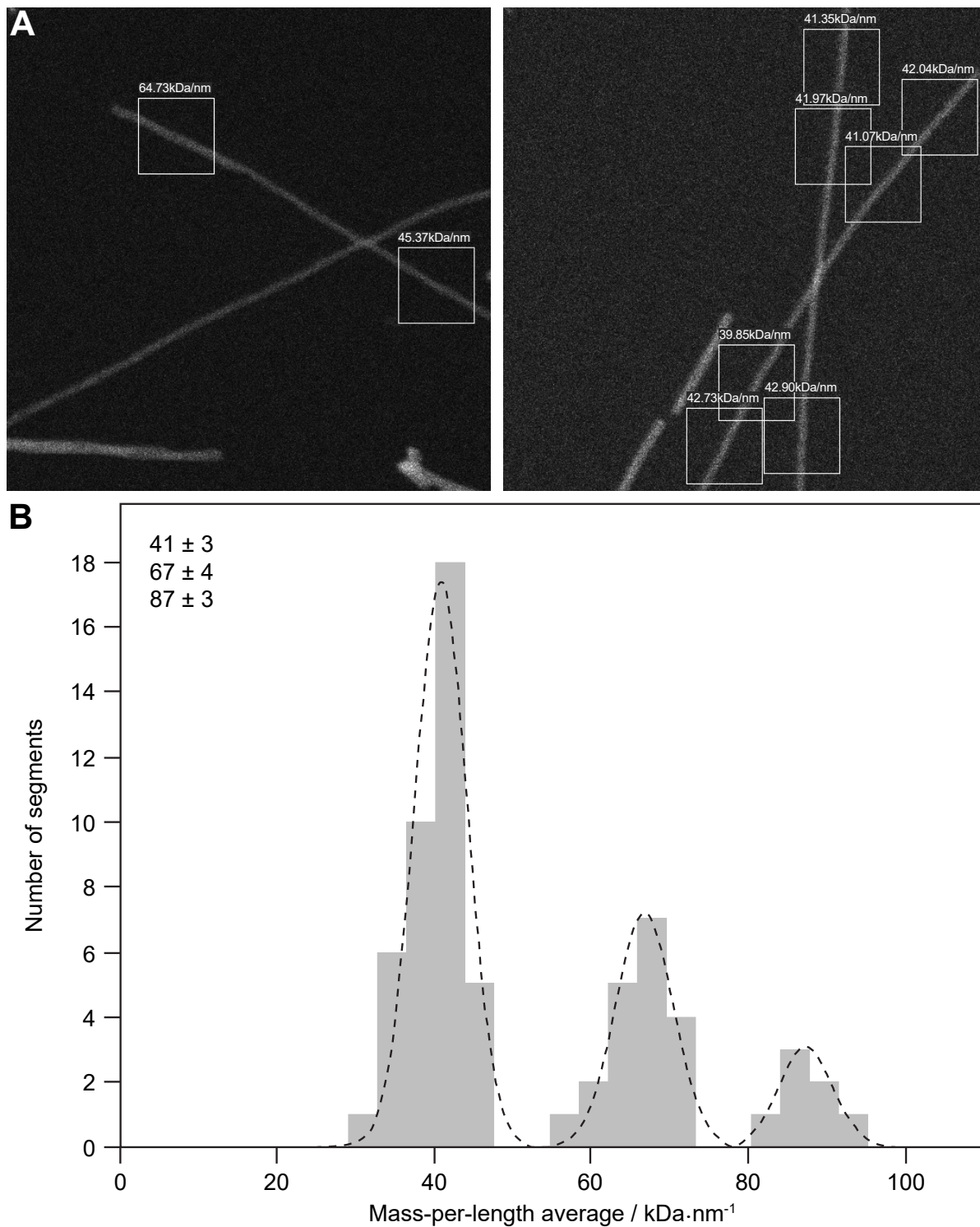
Figure 3.3. *STEM and mass-per-length data from fibrillised SST-14.* **(a)** STEM electron micrographs of the fibrils, from which a total of 66 measurements were drawn (two examples with nine measurements shown here). **(b)** Three MPL classes can be identified from the combined 66 measurements, with average MPL values of 41 ± 3, 67 ± 4, and 87 ± 3 kDa·nm^{-1}.

Consequently, the presumed protofilament MPL is one half of 41 ± 3 kDa·nm^{-1} (the higher values may be comparatively more obscured by incorporated heparin). The resulting MPL of 20.5 ± 1.7 kDa·nm^{-1} is converted with a molecular weight of 1.64 kDa [115], and an *inter*-β-strand distance of 0.48 nm [35, 116], to an expected number of monomers per fibril layer of 6.0 ± 0.5.

3.2.3. Sequential resonance assignment

A number of two- and three-dimensional spectra were recorded on uniformly ^{13}C-^{15}N labelled (diluted, see Section 3.4) somatostatin-14 amyloid fibrils in a 3.2 mm rotor. Figure 3.4 gives a general impression of the well-resolved spectra, with a median ^{13}C FWHM linewidth of about 150 Hz (Figure 3.4B), which corresponds to 0.7 or 1.0 ppm at the static magnetic fields of 20.0 and 14.1 T respectively, that were used for data acquisition. Such resolution compares well to the resolved spectra of other amyloid fibrils, as for example HET-s [117], Aβ [84], or α-synuclein [10].

As a basis for the sequential assignment of SST-14 fibrils, three 3D spectra were used, namely an NCACB, NCOCX, and CANCO, besides 2D NCA and 50 ms CC DARR spectra. Particularly the interpretation of the 2D NCA spectrum (Figure 3.4A) should be relatively straightforward given the small number of amino-acid residues in the sequence: the spectrum should contain exactly 14 NCα cross-peaks. However, a quick inspection of the 2D, in combination with *e.g.* the serine/threonine-region of the 50 ms DARR spectrum, suggests that more than one molecular conformation is present in the sample. Upon further analysis, two separate molecular conformations (labelled in black and red in Figure 3.4) could be identified and sequentially assigned. Figure 3.5 shows part of the assignment for both conformations.

With the use of 3D NCACX and CCC spectra, as well as a 400 ms DARR 2D ^{13}C-^{13}C correlation spectrum, most of the sidechain chemical shifts could be assigned. However, the aromatic resonances of the three phenylalanine residues present in the primary structure showed to be difficult to assign unambiguously, due to overlapping peaks. This was especially due to the two molecular conformations, that have distinct chemical shifts for example for the threonine residues, but less so in the C-terminal cysteine for instance, or the aromatic sidechain resonances. In addition, dynamics (*e.g.* ring flips) may play a role in the absence of some resonances in the spectra. Figure 3.6 shows the assignment graphs for
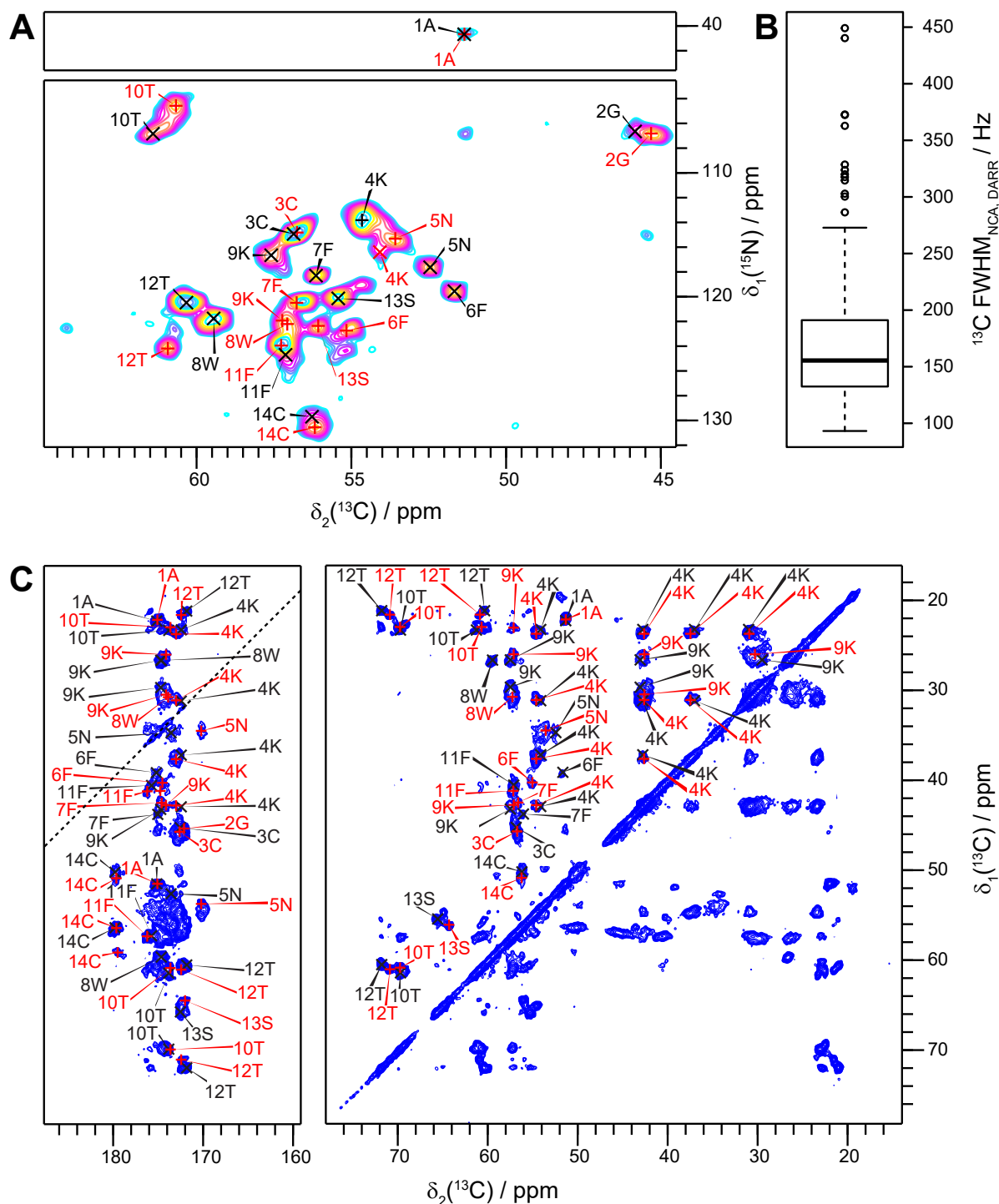
Figure 3.4. *2D solid-state NMR spectra of SST-14 fibrils.* **(a)** An NCA spectrum for which typically each amino acid in the sequence produces one cross-peak. This spectrum shows doubling for many peaks (*e.g.* Phe6 and 7, Lys9, and Thr10 and 12), indicating polymorphism (or oligomeric organisation per fibril layer). The backbones of two systems were separately assigned and are consistently represented with black and red labels. **(b)** FWHM linewidth of shown NCA and aliphatic DARR peaks in the direct ^{13}C dimension. **(c)** 2D ^{13}C-^{13}C DARR spectrum with a 50 ms mixing time where peak doubling is again indicated by assignment to two separate molecules represented by red and black labels. A third peak is observed for the Ser13 Cα-Cβ cross-peak.
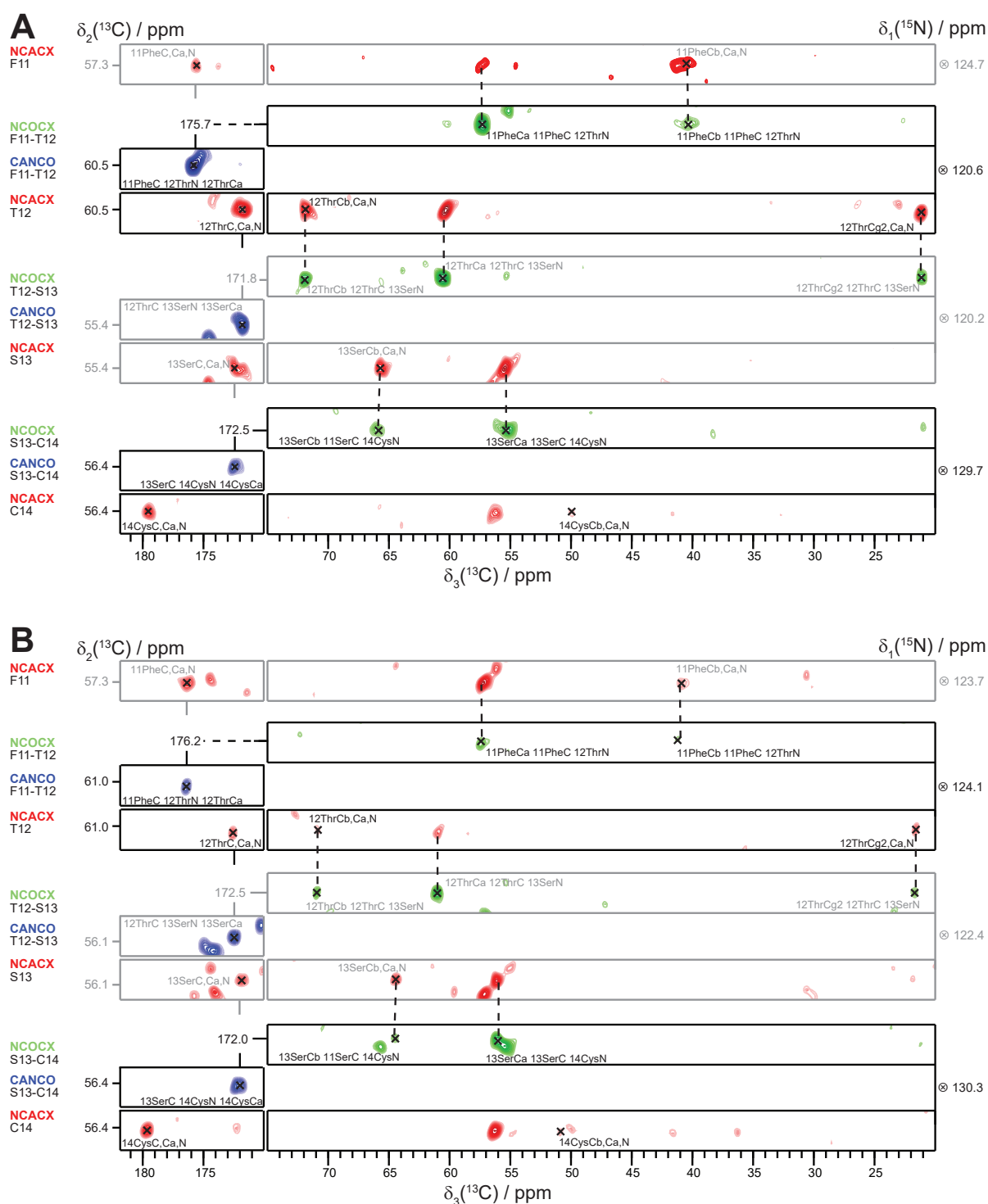
Figure 3.5. *Sequential backbone-assignment of SST-14 from Phe11 to Cys14.* Black labels in Figure 3.4 and 3.6A correspond to the molecular conformation assigned in **(a)**, while the assignment of the other conformation (in red, in Figure 3.4 and 3.6) is illustrated in the strip plot in **(b)**.

both conformations, with 100 % of the backbone Cα, C', and N resonances assigned, 95 % of the Cβ and Cγ resonances assigned, and 37 % of the remaining ^{13}C resonances assigned.

3.2.4. Polymorphism

The structure of fibrils may vary, both at the molecular and the supramolecular level [118], despite having the same primary peptide sequence. Such polymorphism is a common feature of amyloids [10, 33, 119], and can present itself in microscopic images, *e.g.* as protofilaments combining to form ribbons next to, or twists around each other [120], while it may also be apparent from peak doubling in solid-state NMR experiments [77, 118, 120]. Other techniques that can give insight into polymorphism include crystallography, that can distinguish between various segments of the peptide chain as an origin for polymorphism on the one hand, contrasted to identical segments of the chain packing in different ways on the other [121].

The number of polymorphs that a particular protein or peptide adopts, depends on the primary sequence as well as the conditions under which the fibrils form, and commonly varies between a few up to a couple dozen [35]. Specifically *in vitro*, fibril growth conditions such as pH, temperature, buffer composition and peptide concentration can be controlled to favour one predominant morphology [33], especially when using small fragments of one parent fibril ("seeds") for the nucleation of daughter fibrils. Nevertheless, producing 50 mg of a single fibril polymorph, possibly multiple times for different isotope labelling schemes, remains to be a difficult task.

In this context it is hardly a surprise to observe multiple molecular conformations in the SST-14 spectra, as described in Section 3.2.3. In the spectra, two separate but complete sequential connections can be made from N- to C-terminus. Looking closely, three sets of peaks can be distinguished for Ser13, which may indicate the presence of a third conformation variation toward the C-terminus of the peptide, but none of the other residues exhibit a third resonance.

When comparing peaks for some of the resonances in both molecular conformations (Figure 3.7), the system that is labelled in black in Figures 3.4 and 3.6 seems to consistently have a somewhat higher intensity than the system labelled in red. This notion is also apparent from the 2D spectral slices from 3D spectra in Figure 3.5. If the two spectrally distinct molecular conformations would be the result of for example a heterodimer, one would expect both sets of peaks to be equally intense. Instead, in the context of the common
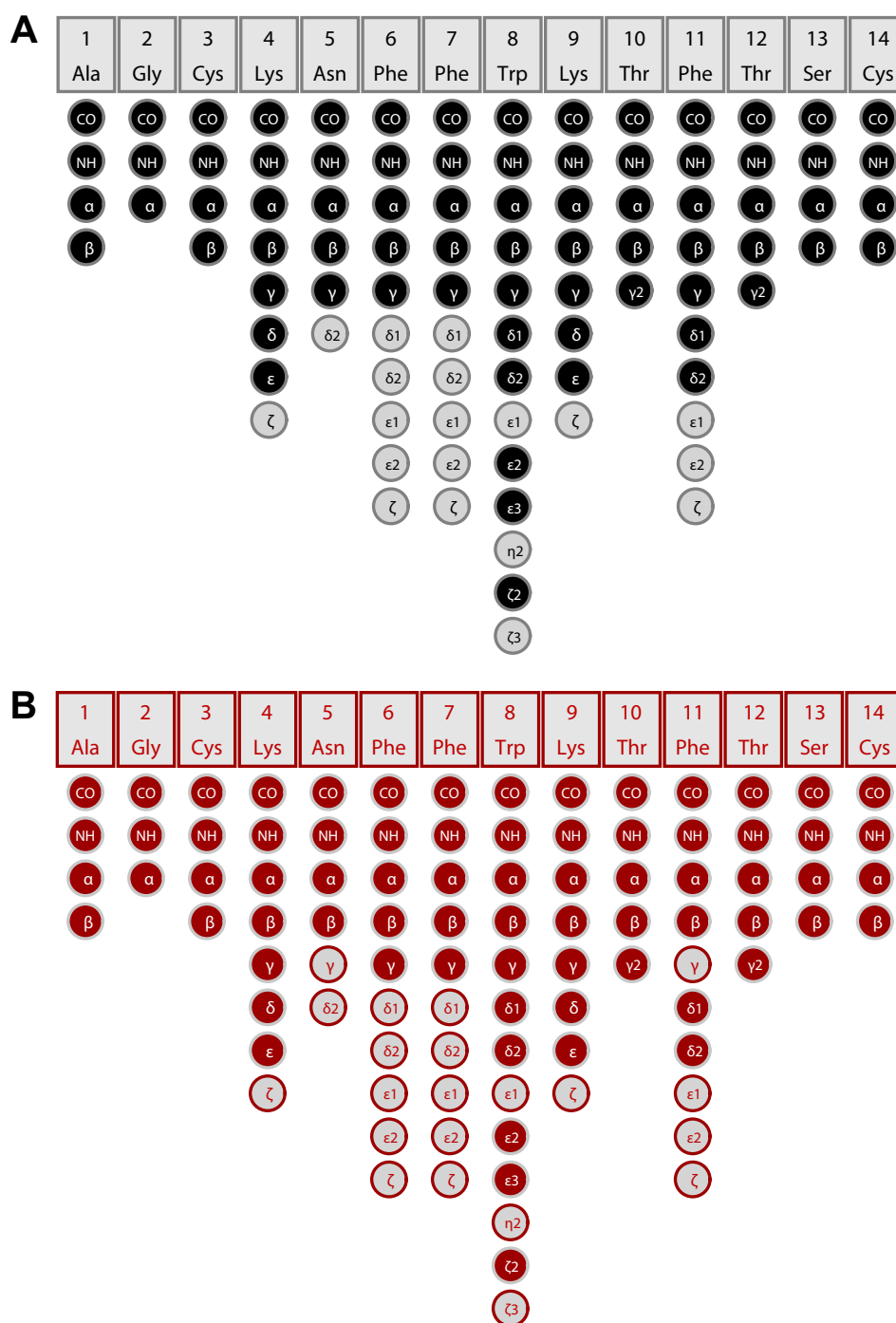
Figure 3.6. *Assignment graphs for both molecular conformations in the spectra.* Assignment graph **(a)** corresponds to the black labels in Figure 3.4 and to the sequential assignment strip plot in Figure 3.5A; **(b)** corresponds to Figure 3.4 (red labels) and 3.5B.
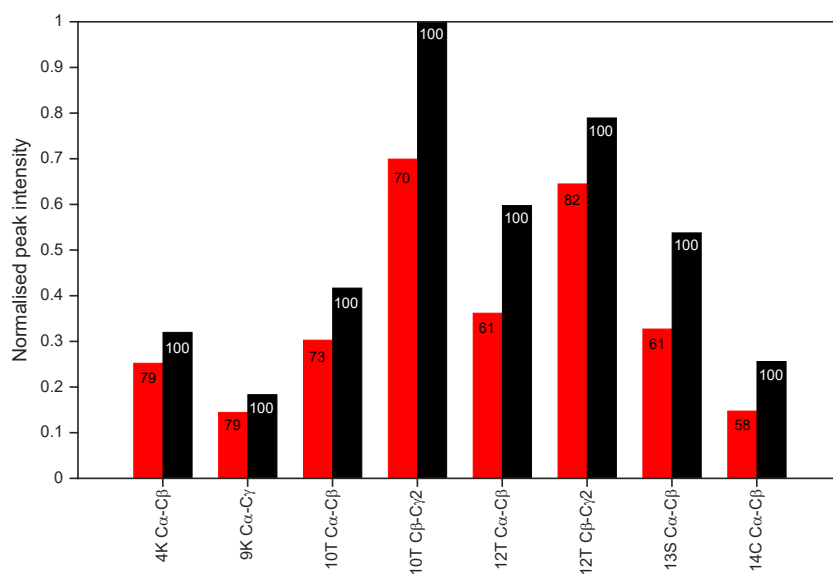
Figure 3.7. *Normalised peak intensities from a 50 ms DARR.* All peak intensities are normalised to the most intense peak (the 10Thr Cβ-Cγ2 cross-peak). Relative intensities for the red peaks compared to the black peaks *within* each pair of two are given as a percentage in each bar, to a mean value of 70 ± 9 % for the red peaks.
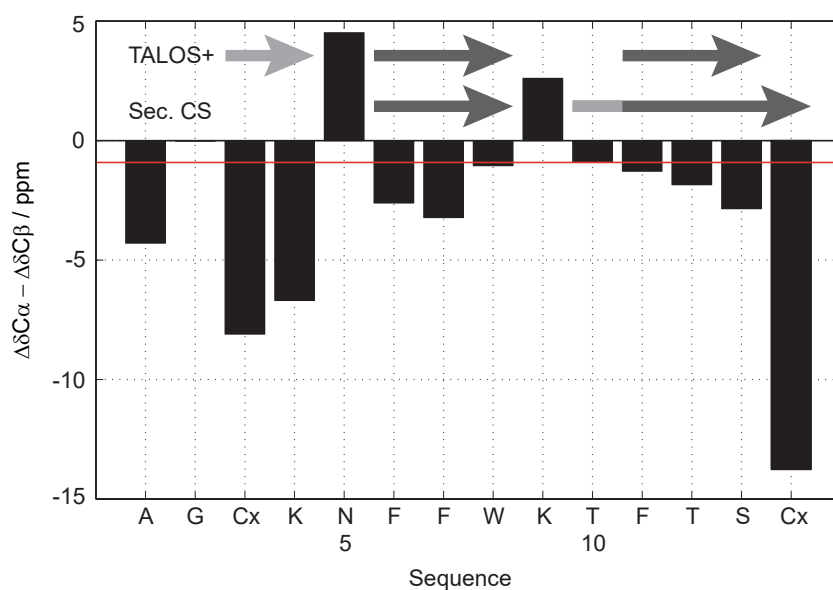
Figure 3.8. *Secondary chemical shifts (SCS) and TALOS+ dihedral-angle predictions.* Three or more negative secondary chemical shift values in a row are indicative of β-strand propensity, and arrows are accordingly drawn above the bars, both for the secondary shifts shown here, and for TALOS+ [91] secondary structure predictions. (Dihedral-angle predictions are given in Table 3.1.) The light grey β-strand predictions are borderline cases because of the shortness of the strand (for the top-left strand), or because of the not-so-negative SCS (for residue 10). The 'x' on the two cysteines indicates that random-coil values of the oxidised form were used [122], and the red line indicates a value of -1 ppm.

incidence of polymorphism in amyloids and *in vitro* fibril preparations, it seems more likely to interpret the two conformations that have distinct but very similar chemical shifts as polymorphs. While it remains important to keep the two systems apart, for instance in the collection of distance restraints, the rest of the presented data will focus on the major, most intense polymorph, *i.e.* the system labelled in black in Figures 3.4, 3.6 and 3.7, for which the sequential assignment is illustrated in Figure 3.5A.

3.2.5. Distance restraints

For the collection of distance restraints, NMR spectra were recorded on the sample that was also used for sequential assignment, namely uniformly ^{13}C,^{15}N-labelled (UL) SST-14 fibrils, diluted with peptide fibrils at natural abundance (NA), in a UL:NA ratio of 1:3 (see Section 3.4). The performed experiments include a 400 ms DARR ^{13}C-^{13}C correlation 2D, as well as 6.5 ms PAR and 300 µs CHHC 2D spectra. Due to the sample dilution, all observed cross-correlations are assumed to be *intra*molecular.

Besides these experiments, a 2D PAIN spectrum was recorded on a mixed-labelled sample, where half of the peptides are ^{15}N-labelled, and half are ^{13}C-labelled. As a result, in a ^{15}N-^{13}C 2D correlation spectrum, only *inter*residual contacts will appear, since none of the peptides (save the 1 % NA ^{13}C) have both ^{15}N- and ^{13}C-labelled nuclei in their sequence.

3.2.5.1. In-register parallel stacking

Amyloids are characterised by a cross-β-sheet structure that was first described in 1935 on the basis of X-ray diffraction of poached egg-white films [123]. In such a structure, many β-strands (in the order of thousands) lie on top of each other, forming long β-sheets packed together at a distance of roughly 10 Å, in the direction of the fibrillar axis [35]. The tight packing requires careful interdigitation of sidechains, known as a 'steric zipper' [34], categorised into eight different classes. The orientation and placement of strands within the sheet, and the sheets with respect to each other, determines the class of the zipper.

First, the direction of each subsequent β-sheet (determined by the N- and C-terminal ends of the polypeptide chain) can be the same or opposite, leading to 'parallel' or 'anti-parallel' β-strands. Second, each parallel strand (*i.e.* every strand in a parallel sheet, or every second strand in an anti-parallel sheet) can lie exactly on top of the strand below, leading to an in-register stacking [35]. Although less common, each parallel strand can also be shifted with respect to the strand below, known as an out-of-register β-sheet [124].

Residue	Φ_{min}	Φ_{max}	Ψ_{min}	Ψ_{max}
3 Cys	-164	-103	142	181
4 Lys	-166	-117	148	176
6 Phe	-134	-88	106	168
7 Phe	-167	-85	113	178
8 Trp	-193	-74	86	192
10 Thr	-190	-62	121	188
11 Phe	-138	-66	111	148
12 Thr	-150	-90	104	148
13 Ser	-145	-79	104	136

Table 3.1. *Dihedral-angle restraints.* All angles are given in degrees and the range for both Φ and Ψ angles is given at two standard deviations around the value predicted by TALOS+ [91]. For detailed output, see Appendix Table B.1

Figure 3.9 shows an overlay of the mixed-labelled PAIN spectrum described above, with a uniformly (1:3 diluted) labelled NCA 2D. All intraresidual cross-peaks present in the NCA are also present in the PAIN spectrum. The high number of intraresidual cross-peaks in the PAIN spectrum, that must at the same time be intermolecular due to the labelling, are clearly indicative of an in-register parallel stacking of the SST-14 amyloid fibrils (see also Figure 3.10A)

3.2.5.2. Hydrogen bonding and dihedral angles

The cross-β-sheet structure of the amyloid, as described above, is built from individual β-strands that are defined in the molecular structure calculation by intermolecular hydrogen-bonding distance limits between the O or N and the H atom involved in the bond. The location of these bonds can be predicted from the primary structure by algorithms such as GOR4 [125] (Appendix Table B.2). More advanced algorithms like DSSP [126] or STRIDE [127] use the atomic coordinates of the molecule to predict secondary structure. Although the latter is often used in molecular viewing software if a PDB file does not contain any secondary structure information (*e.g.* as implemented in PyMOL [128], Chimera [129], and VMD [130]), it is not helpful before calculation of a molecular structure has taken place.

Instead, the chemical-shift assignment provides data that can be used to predict secondary structure in comparison to deposited chemical shifts of known structures (*e.g.* TALOS+ [91]) or in comparison to deposited random-coil shifts (Figure 3.8). The difference between the observed chemical shift and the random-coil shift for the same residue type (known as the secondary chemical shift, SCS, or $\Delta\delta$), is sensitive to the environment, for example an α-helix
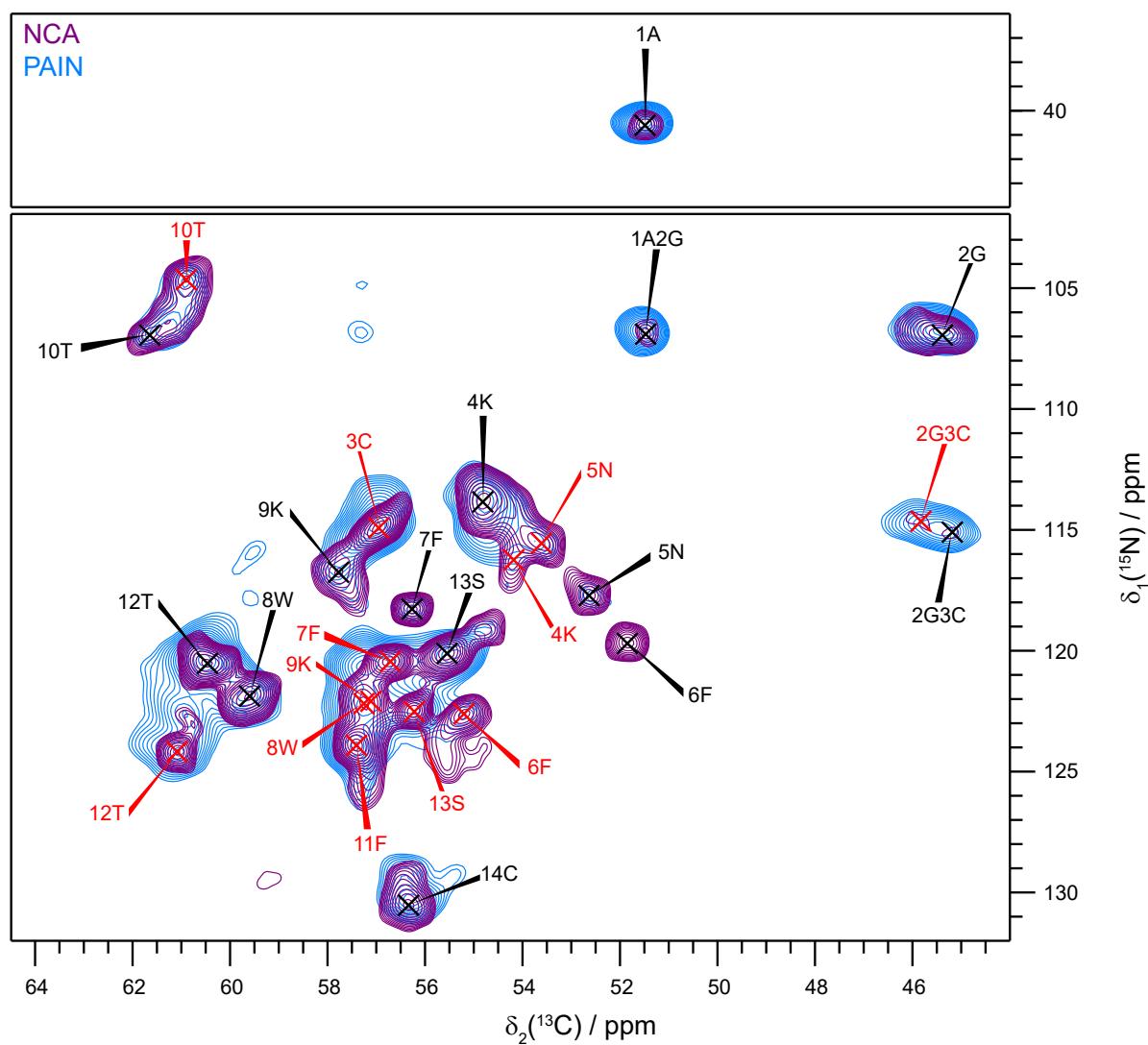
Figure 3.9. *Overlay of a PAIN spectrum of the mixed $^{13}C,^{15}N$-labelled sample (magenta) and the NCA spectrum of the uniformly $^{13}C,^{15}N$-labelled sample (blue).* Most of the intra-residual $^{13}C\alpha$-^{15}N cross peaks in the NCA also appear in the PAIN spectrum, which is characteristic of an in-register parallel fibril stacking. The red and black labels correspond to the colour-coding as it was used in Figures 3.4-3.7.

or a β-strand. In particular, the Cα and Cβ chemical shifts have been observed to shift to a higher and lower ppm value respectively, if located in an α-helix, compared to a random-coil shift, while the exact opposite is the case for a β-strand [87]. In solids, usually $\Delta\delta C\alpha - \Delta\delta C\beta$ is calculated, which has the advantage that it is independent of referencing. For glycines, often the $\Delta\delta C\alpha$ is calculated (for lack of a Cβ shift), but Gly2 in SST-14 (Figure 3.8) has a $\Delta\delta C\alpha$ of -0.012 ppm, which is not visible in the plot.

The restraints used in the calculation are for hydrogen bond upper- and lower-distance limits corresponding to two β-strands, from residues 6-8 and 10-13, following the prediction by the secondary chemical shifts, except for the C-terminal cysteine. The latter has a highly negative value in Figure 3.8, stemming from a Cβ chemical shift that is much higher (~ 51 ppm) than observed in a random coil (~ 38 ppm for the oxidised form [122]). However, this strong downfield shift may well be the result of the deshielding negative charge of the C-terminal carboxyl group and have less to do with a β-strand propensity.

The β-strand predicted by TALOS+ for residues 3 and 4 (see Figure 3.8) is not included in these restraints, but the dihedral-angle restraints for these residues are taken into the calculation (Table 3.1). In this table, the terminal residues have no predictions, because to make a prediction, TALOS+ requires chemical shifts of neighbouring residues on each side of the concerned residue. In addition, values for residues 2, 5, and 9 were marked with a warning in the TALOS+ predictions, which means that there is no consensus in the database for these particular triplets of residues and their observed chemical shifts. They were consequently not taken into account for the calculation.

3.2.5.3. NMR distance restraints

Experiments based on second-order recoupling in the mixing time give cross-correlations that are less sensitive to dipolar truncation. Such correlations are useful for extraction of structure calculation restraints over longer distances. Particularly informative are restraints for atom pairs that are close in space, but remote in the primary sequence.

The spectra discussed in the beginning of this section (p. 57) provided distance restraints for the structure calculation as presented in Table 3.2. The total number of spectrally unambiguous intramolecular distance restraints is 17. This number, for a 14-residue polypeptide, can be compared to 26 manual unambiguous intramolecular restraints for the calculation of Aβ E22Δ with 39 residues (see Appendix Table A.1)[53], or 19 manual unambiguous intramolecular restraints for β-endorphin with 31 residues (Table 5.1 in [131]). The

	ω1			ω2		UPL / Å	d / Å
DARR 400 ms	3	CYSS	CB	1	ALA	CB	7	7.0
	3	CYSS	C	5	ASN	CB	7	6.6
	4	LYS	CD	14	CYSS	CB	7	4.4
	4	LYS	CB	14	CYSS	CB	7	6.0
	4	LYS	CB	12	THR	CG2	7	5.5
	10	THR	CB	8	TRP	CB	7	6.7
	10	THR	CB	12	THR	CG2	7	4.9
	11	PHE	C	13	SER	CB	7	5.7
	14	CYSS	C	12	THR	CB	7	6.7
PAR 6 ms	4	LYS	CG	12	THR	CB	6	4.8
	6	PHE	CA	12	THR	CG2	6	6.0
	12	THR	CB	4	LYS	CG	6	4.8
CHHC	4	LYS	HA	2	GLY	QA	6.5	6.0
	6	PHE	QB	10	THR	QB	6.5	4.6
	6	PHE	QB	12	THR	QB	6.5	6.3
	12	THR	HB	10	THR	HB	6.5	6.0
	12	THR	QG2	14	CYSS	HA	6.5	6.5

Table 3.2. *Intramolecular spectrally unambiguous distance restraints.* Only restraints at a distance of at least $i \pm 2$ are given. UPL gives the upper distance limit in ångström, and d gives the actual distance in the calculated model (see also Figure 3.10C).

400 ms DARR, PAR, and CHHC spectra can be found in the appendix (Figures B.1 and B.2). The criterium for 'relevance' of distance restraints for a calculation depends, of course, on the structure that is calculated. For a structure that is rich in α-helices, restraints below $i \pm 4$ could be considered less relevant since many of those will merely confirm the helical structure. In the SST-14 monomer, the well-documented disulphide bond [96, 97] forms a strong determinant for the structure (see also Materials and Methods Section 3.4.2), and is confirmed by high Cβ chemical shifts for Cys3 and Cys14. Important information that we want to extract from the NMR distance restraints is the orientation of sidechains, especially with respect to pointing into or out of the fibril core. Many of the observed restraints (Figure 3.10C) indeed confirm a network of interresidual contacts between those sidechain atoms that are on the inside of the cyclic molecule.

3.2.6. Structure calculations

The calculation of the monomer structure of SST-14 in amyloid form is based on a five-layer fibril context (Figure 3.10A), performed with CYANA [89]. In the calculation, the amino-acid

residues in the layers are numbered 1-14, 101-114, 201-214, 301-314 and 401-414, linked by pseudo-atoms that are removed at the end of the calculation. In the resulting pdb file, only the residues 201-214 are of interest for the correct monomer structure from the SST-14 fibril, and presented figures and statistics in this chapter concern only that layer.

Primary input for the calculation was *i)* the polypeptide sequence (p. 46), *ii)* a 3Cys-14Cys disulphide bond [96, 97], and *iii)* two β-strands for residues 6-8 and 10-13 (Section 3.2.5.2), based on secondary chemical shifts, ambiguously defined as upper- and lower-distance limits to either the fibril layer above, or below the current layer. In addition, also based on the chemical shifts, *iv)* dihedral-angle predictions from TALOS+ [91] were included as restraints in the calculation (Table 3.1). The upper and lower limits for each angle were defined as two standard deviations from the predicted mean. Finally, *v)* symmetry restraints with respect to atomic distances and dihedral angles (generated by CYANA [89]) ascertained symmetry of the monomers. With these restraints, a structure calculation of 300 structures produced a bundle of the ten best conformations with a backbone RMSD to mean of 1.57 ± 0.39 Å.

To improve this structure, the unambiguous distance restraints from the discussed 2D spectra (DARR, CHHC, PAR; Table 3.2) were included as upper-distance limits varying between 6 and 7 Å. With these restraints included, a total of 300 structures (each containing five layers) was calculated with CYANA [89], in which the ten best structures had an average backbone RMDS to mean of 0.69 ± 0.34 Å and a mean target function value of 4.00 ± 0.65 Å^2. The target function is defined such that it is zero if and only if all experimental distance restraints and torsion angle restraints are fulfilled, and all non-bonded atom pairs satisfy steric non-overlap conditions [89]. Consequently, the molecular conformation that satisfies the given restraints best will have the lowest target function value. We consider the mean target function value of 4 Å^2 for a five-layer fibril a good result of the calculation, in combination with an average backbone RMDS to mean of 0.69 ± 0.34 Å and a heavy atom RMSD to mean of 1.15 ± 0.23 Å. More details concerning the resulting structure are discussed below.

3.2.7. Atomic-resolution structure

The monomer structure of SST-14 in amyloid form is presented in Figure 3.11. Each monomer has two β-strands spanning residues 6Phe-8Trp and 10Thr-13Ser. The restraints used for the calculation are discussed above, and in the resulting structure only very few of those restraints are violated. An upper-distance limit of 6 Å between 6PheCα and 12ThrCγ_2 is violated by 0.11 Å in four of the ten best structures, although not in the presented monomer
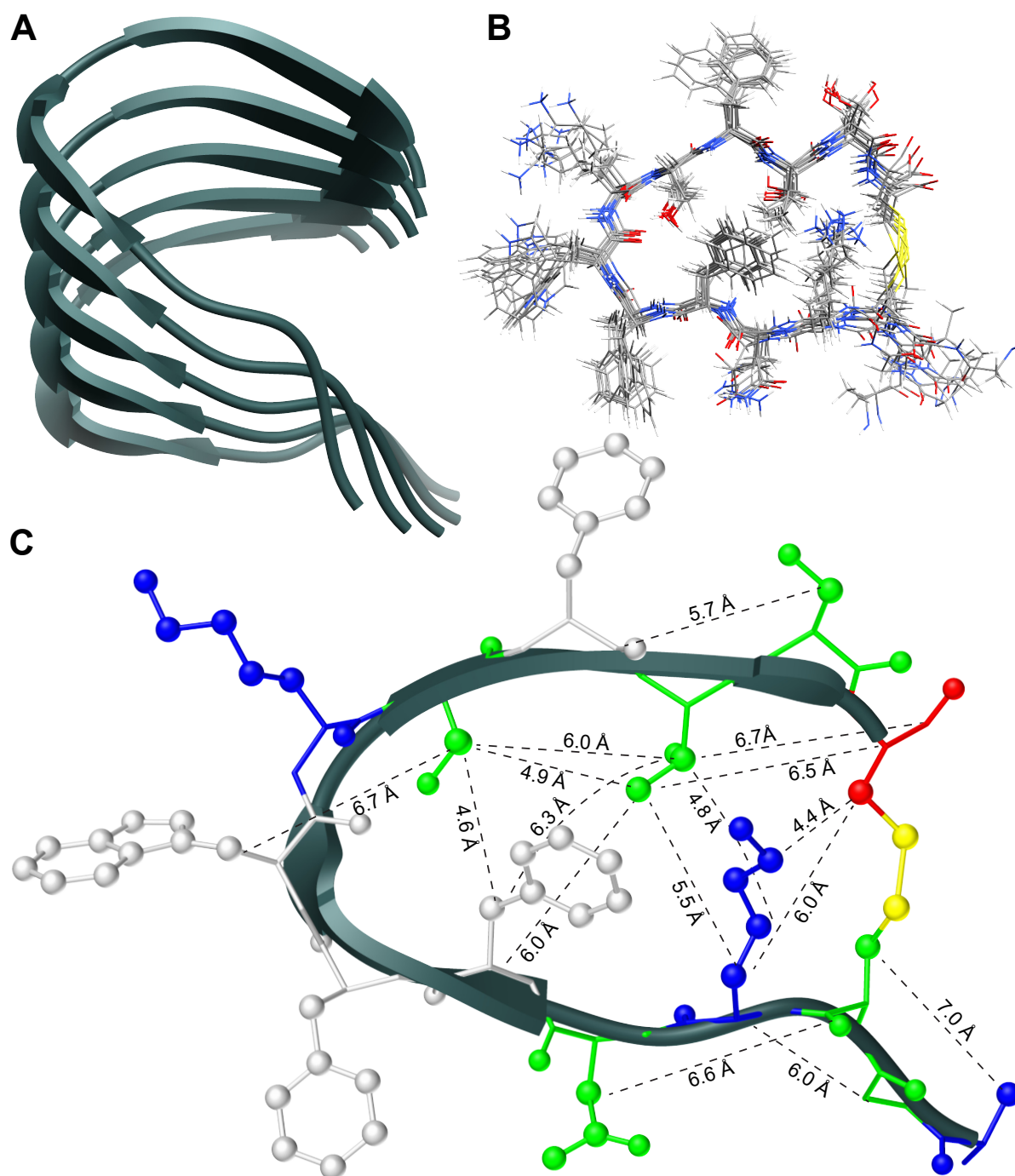
Figure 3.10. *Structure calculation of the SST-14 monomer in the context of a fibril.* **(a)** The calculated and discussed monomer is flanked by two peptide molecules on each side. This five-layer fibril serves to give the monomer the right context in terms of intermolecular hydrogen bonding and β-sheet formation, as it is shown here in an in-register parallel stacking. **(b)** Top view of the ten best structures from the CYANA [89] calculation for the central layer (the third from either side) of the fibril. **(c)** Distance restraints used for the structure calculation were set to upper limits of 6-7 Å, as given in Table 3.2. Displayed here are the actual distances in the calculated model, which are all below the upper limits defined for the calculation. (Hydrogen atoms not shown; *cf.* Figure 3.11A, also for an explanation of colours used in (c).)

A
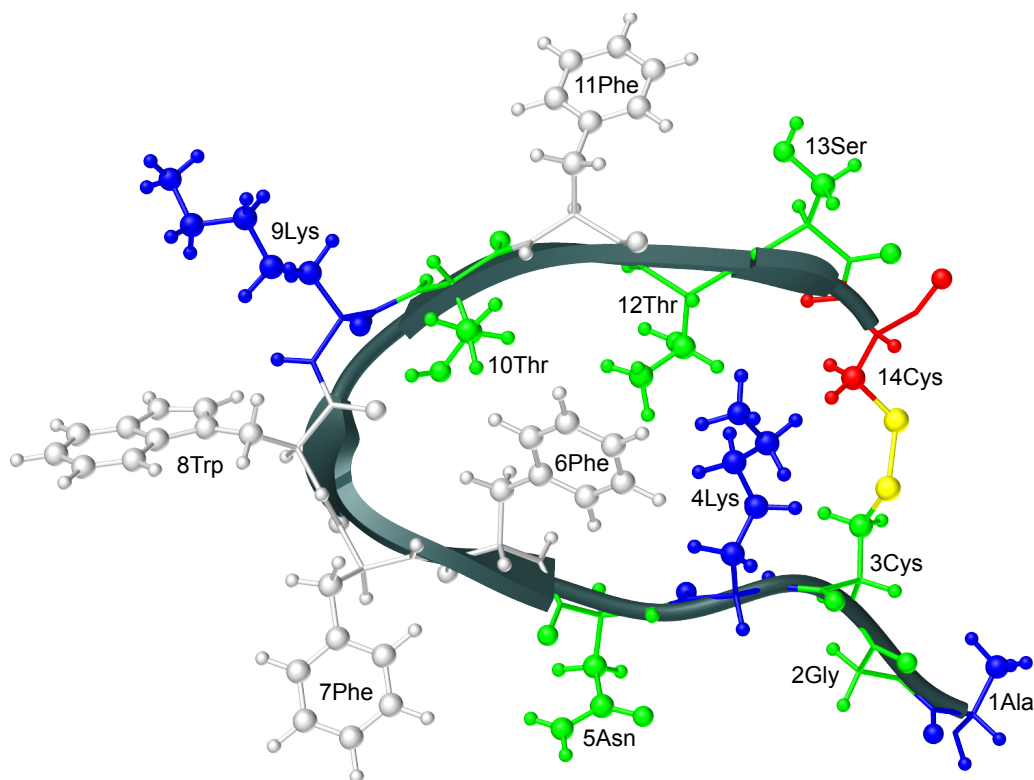
B
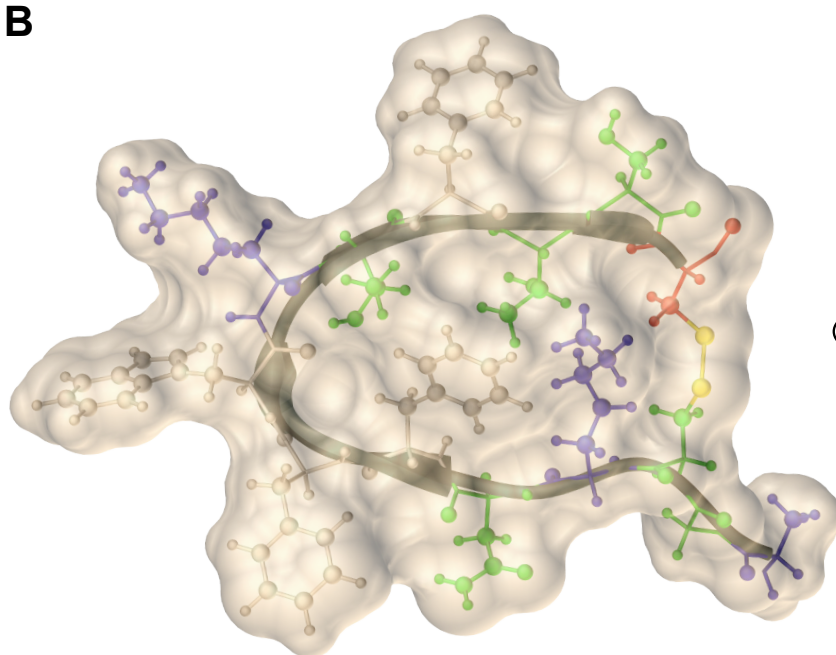
C
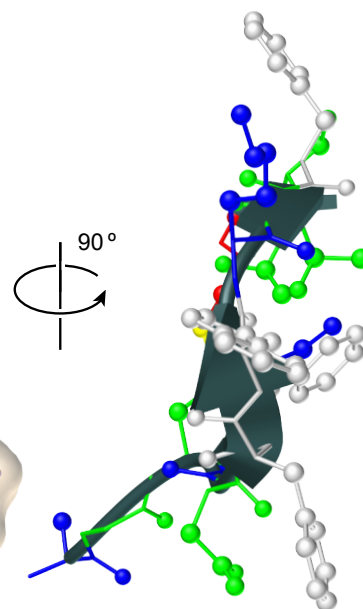
Figure 3.11. *Final structure of SST-14 monomer in amyloid form.* **(a)** The monomer structure with all atoms displayed. Hydrophobic residues are coloured in white, polar residues and glycines in green, and charged residues in blue when positive, and red when negative. The disulphide bond is in yellow. **(b)** A surface trace of the molecule (performed with Chimera [129]) shows the complete fill of the fibril core. **(c)** A side view of the monomer structure illustrates the relatively planar conformation of the molecule, except for the N-terminal 1Ala and 2Gly that exhibit some conformational variation, as can also be seen in Figure 3.10B.

with the lowest target function value of all. Another violation that is present in four of the ten best conformations but not in the presented monomer, is an intermolecular overlap of the van der Waals radii of 9LysO and 9LysHα of the layer below by 0.16 Å. Finally, another van der Waals violation in four conformations in the bundle, including the best one presented in Figure 3.11, concerns 6PheN and 6PheHδ_1 at an overlap of 0.09 Å.

The Ramachandran plot of the presented monomer (Appendix Figure B.3) indicates that 84 % of the dihedral angles is in the most favoured regions, while 12 % is in the additionally allowed regions. The dihedral angles for 2Gly are not restrained by TALOS+ [91] predictions (predictions were not included based on a warning by the programme; see Appendix Table B.1) and the N-terminus of the molecule exhibits some conformational variation, as can be seen in the bundle in Figure 3.10B. Some of those conformations, including the best one, contain dihedral angles for 2Gly that are outside the favoured regions. Interestingly, some other conformations position the N-terminus such that both 3Cys and 2Gly are in the right-handed α-helical region. For *in vitro* fibril formation of SST-14, it has been observed that the SST-14 peptide passes from a random-coil conformation to an amyloid through a helix-rich conformation, based on far UV-CD observations during aggregation [103, 132]. The turn at the N-terminus of the peptide in some of the calculated conformations may reflect a remnant of this amyloid-forming mechanism in SST-14. Additional angle and distance restraints would be necessary to more unambiguously define the orientation of the two N-terminal residues.

In the context of a fibril, two things are important for the monomer structure. First, considering the stacking of monomers on top of each other in the fibril, the orientation of the amino-acid sidechains should be relatively planar (although a staggered conformation would form an exception to this rule). This prevents steric clashes with neighbouring molecules above and below in the fibril. Figure 3.11C shows that the molecule is indeed rather flat, with the exception of the two N-terminal residues. They point out of the plane a bit, but do not cause van der Waals violations in the fibril calculation, considering that all N-termini are pointing slightly out of the plane (Figure 3.10A).

Second, it would seem counter-intuitive if the inside of the cyclic molecule would contain "empty" space. The surface trace of the monomer (Figure 3.11B) clearly shows that the core of the molecule is occupied by the inwards-pointing sidechain atoms. In this respect it is somewhat unusual for molecular structures in general, but not surprising for this SST-14 monomer, to see that a number of hydrophobic residues (*i.e.* 7Phe, 8Trp, and 11Phe) are solvent-exposed: considering the limited space on the inside of the loop and the 'bulky'

nature of the aromatic rings, it is to be expected that some of the hydrophobic sidechains point outward. In addition, earlier in this chapter (p. 47) it has been discussed that SST-14 and Aβ interact, and indeed 8Trp was shown to be crucial for this interaction in mutation studies [106, 112]. This makes it highly unlikely for at least 8Trp to be oriented to the inside of the molecule.

In general review of the fold of this relatively small peptide, the two β-strands are more or less facing each other. This facilitates the cross-β structure that is characteristic for a generic amyloid conformation [34], as also observed in HET-s [7], Aβ [8, 9, 53], and α-synuclein [10, 11, 133, 134], for example. Noteworthy in this structure is the 4Lys sidechain pointing to the inside of the cyclic monomer. A compensation for this positive charge is found in the negative charge of 14Cys at the C-terminus of the sequence. Indeed, the presented structure can be calculated with an upper distance restraint of 4 Å on the Nζ of 4Lys and the backbone O of 14Cys, combined with a 3 Å restraint on 4LysHζ (ambiguously on one of the three Hζ's) and 14CysO. Such a calculation produces a very similar structure without significantly elevating the target function value (*i.e.* without major restraint violations). However, the presented structure does not include such a restraint, as it would currently only be based on the assumption that such a 4Lys-14Cys salt-bridge is formed, and bias the molecular conformation to this end. Instead, a ^{1}H detected ^{15}N-^{1}H 2D NMR spectrum of the lysine sidechain resonances (see Figure 5.6C on p. 106) could possibly provide chemical-shift based evidence for this case, as discussed in the context of DNA-nucleotide interactions in Chapter 5.

As a somewhat speculative conclusion to this section, it is important to return to the mass-per-length data (discussed in Section 3.2.2) which indicate SST-14 to form hexameric layers in the fibril. Taking into account the adjacent hydrophobic residues 7Phe and 8Trp to the outside of the molecule, it is not unreasonable to assume that this could form the core of the hexamer in an SST-14 fibril, such that, although oriented to the outside of the molecule, the residue sidechains do not point to the outside of the protofibril. Figure 3.12 gives an indication of what the hexamer could look like, presented as a "hand-drawn sketch" to underline the speculative nature of the illustrated conformation. Solving the complete molecular fibril structure of SST-14 is outside the scope of this chapter, but could be approached with precise inter- and intramolecular restraints from a dilution analysis, as has been previously applied to fibrils of *e.g.* HET-s [7] and Aβ [9, 53].
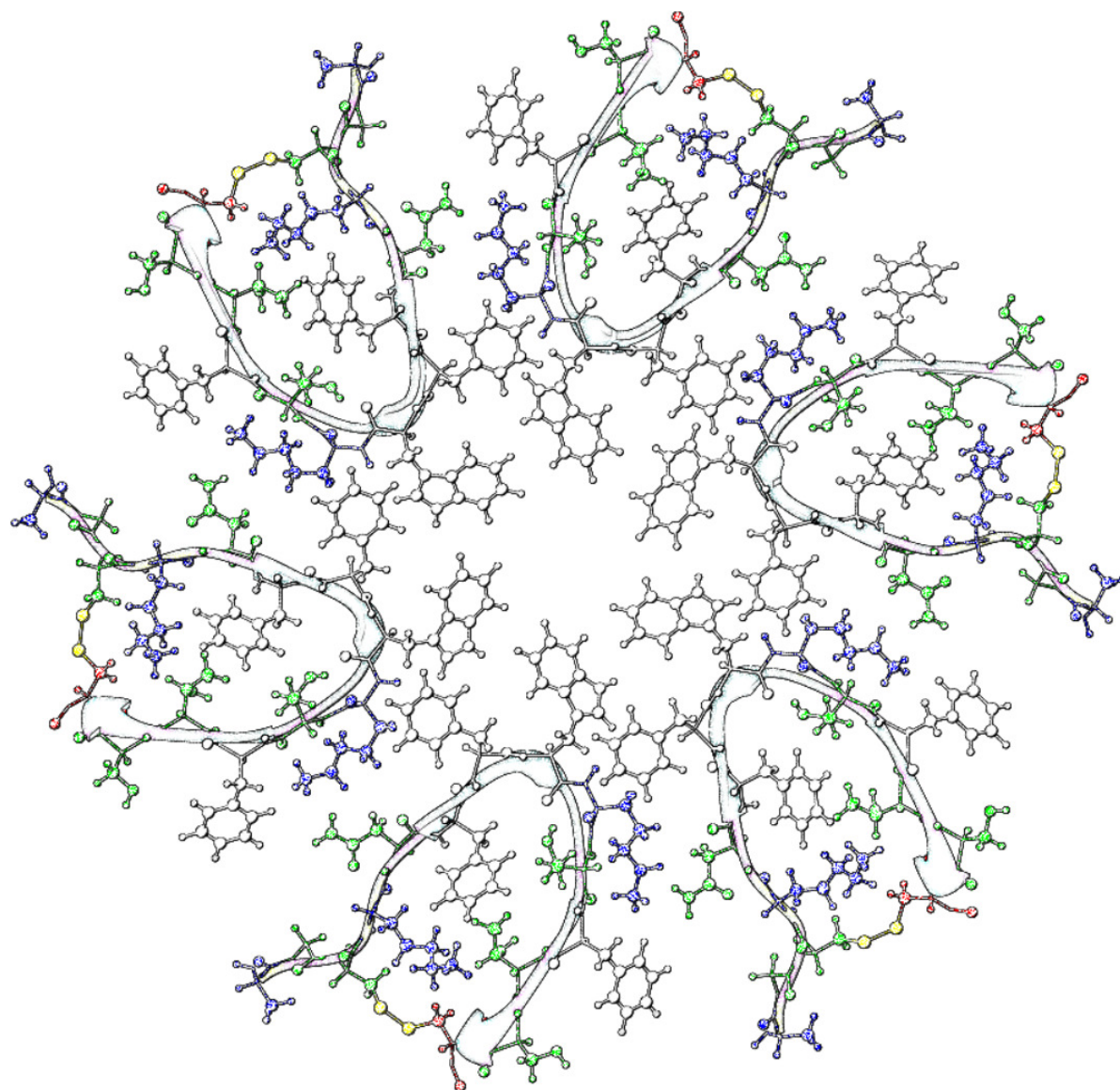
Figure 3.12. *Sketch of what a hexameric SST-14 fibril layer could look like.* Following the mass-per-length data (Section 3.2.2) and the calculated monomer structure (Figure 3.11), one layer of the SST-14 protofibril could have each of the six molecular cores in addition to a core that is on the outside of the cyclic molecules, but on the inside of the protofibril, with a high concentration of hydrophobic amino-acid residues forming the core of the protofibril.

3.3. Conclusion

This chapter discussed the functional amyloid somatostatin-14, assumed to be stored *in vivo* in secretory granules [102], prior to synaptic release as a monomer peptide hormone functioning as an inhibitive agent for growth hormone, insulin, neurotensin and a range of other hormones [98]. As also discussed above, SST-14 has been shown to inhibit Aβ fibril formation by extending its lag phase in recent literature [106, 112].

The data presented above have allowed us to sequentially assign the chemical shifts of the SST-14 monomer in a diluted fibril sample. The assignment procedure revealed the presence of two molecular conformations in the spectra, assumed to be indicative of polymorphism, as it often occurs in amyloids [33, 35]. The assignment and subsequent peak-intensity analysis identified one of the polymorphs for further structural characterisation. This was possible with the use of TEM and STEM images, in particular to determine the mass-per-length of the fibril, and with distance restraint 2D NMR spectra (DARR, CHHC, and PAR). Employing TALOS+ [91] and CYANA [89], predictions and calculations were made.

A calculation of 300 five-layered structures produced a monomer structure bundle of ten conformations with the lowest target function values (4.00 ± 0.65 Å^2) and an average backbone RMDS to mean of 0.69 ± 0.34 Å. This structure contains two β-strands from 6Phy to 8Trp, and from 10Thr to 13Ser. The sidechain atoms of 8Trp are located outside the cyclic core of the peptide hormone molecule, in agreement with its crucial role in aforementioned interaction with Aβ as determined by mutation studies [106, 112]. In addition, the outward orientation of two aromatic sidechains may form the hydrophobic core of a possible hexameric protofibril layer, in accordance with the mass-per-length data. Further studies could confirm the SST-14 fibril structure, as an extension to the experimentally determined three-dimensional monomer structure presented at atomic resolution.

3.4. Materials and methods

3.4.1. Expression and purification of recombinant protein

A detailed description of the protocal is given in [21]. In summary, SST-14 was expressed in *E. coli* One Shot® BL21 Star™ (DE3) strain (Invitrogen, Carlsbad, California, USA). 1 L of M9 medium containing 3 g of $^{13}C_6$-D-glucose and 1 g of $^{15}NH_4Cl$ was inoculated with a starting culture to grow cells at 37 °C. Recombinant gene expression was induced by adding IPTG to 1 mM when the culture reached an OD^{600} of 1.4, and expression was subsequently performed at 20 °C overnight (16 h). The cells were harvested at 10,000 g (Avanti J-26 XP, Beckman Coulter, Inc.) at +4 °C for 30 min and resuspended in lysate buffer. To this buffer, PMSF and lysozyme were added at a final concentration of 1 mM and 0.5 mg/mL, respectively. Cells were disrupted by passing the suspension several times through a 110S microfluidizer (Microfluidics, Newton, Massachusetts, USA). The solution was centrifuged at 25,000 rpm (rotor 45 Ti, Optima L-90 K Ultracentrifuge, Beckman Coulter, Inc.), +6 °C for 1 h, after which the precipitate was resuspended in 25 mL of extraction buffer (50 mM Tris-HCl, 8 M urea, 150 mM NaCl, 10 mM imidazole, 1 mM DTT, pH 7.5). Extraction of the fusion protein was carried out for 16 h at room temperature. The extract was centrifuged for 1 h at +22 °C, 40,000 rpm (rotor 45 Ti, Optima L-90K Ultracentrifuge, Beckman Coulter, Inc.). Purification of the supernatant followed on a HisTrapFF (fast flow) 5 mL column (GE Healthcare). The fusion protein was eluted by elution buffer 1 (EB1) (50 mM Tris-HCl, 7 M urea, 150 mM NaCl, 1 mM DTT, 300 mM imidazole, pH 7.5). The solution was then desalted on a HiPrep 26/10 column (GE Healthcare) and EB1 was exchanged to a cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, pH 7.5). Next, the fusion protein was treated for 2 h with TEV-protease in a molar ratio of 1:100 at room temperature, after which the solution was centrifuged at 4000 rpm (Biofuge primo R, Heraeus) for 15 min at +22 °C. This was followed again by affinity chromatography on a HisTrapFF 5 mL column. Acetonitrile was added to the peptide solution up to 5 % (vol.) prior to reversed-phase (RP) chromatography. Then the solution was acidified to pH 2.0 by the addition of TFA (CF_3COOH) to a final concentration of 0.3 % (vol.). The HisTrapFF column was recovered by washing with EB2 (50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 300 mM imidazole, pH 7.0) and equilibrated again in the lysis buffer with 20 mM imidazole. This loading-elution procedure was repeated 3 times.

3.4.2. Oxidation of SST-14

To make sure that the NMR sample contained cyclic SST-14 with the disulphide bond formed as described and assumed in the work presented in this chapter, lyophilised SST-14 was dissolved at 7 μM concentration in a 100 mM ammonium acetate solution pH 6.8. A filtered solution of 50 mM potassium ferricyanide ($K_3[Fe(CN)_6]$) was added dropwise to the peptide solution under nitrogen flow and constant stirring, until the colour of the solution turned to slight yellow. The pH was maintained at 6.8. At the end of the reaction, the pH was decreased to 2.0 by addition of TFA (CF_3COOH). Acetonitrile was added up to 5 % (vol.). The final purification step to obtain oxidised SST-14 was RP-HPLC (see [21] for the RP-HPLC protocol). The presence and purity of the recombinant protein at different stages was monitored by ESI-MS (Q-TOF Ultima API, Micromass, Manchester, UK). Fractions containing cyclic SST-14 were merged and lyophilised.

3.4.3. Fibril formation

Formation of SST-14 fibrils (and of other amyloid fibrils) is described in the Supplementary Information of [102]. Initially, lyophilised SST-14 was dissolved in 0.5 mL of 5 % D-Mannitol (pH 5.5) and 0.01 % sodium azide at a concentration of 2 mg/mL in 1.5 mL Eppendorf tubes. The Eppendorf tubes were placed into an EchoTherm model RT11 rotating mixture (Torrey Pines Scientific) with a speed corresponding to 50 rpm inside a 37 °C incubator. Afterwards, the content was dissolved in 0.5 mL of 5 % D-Mannitol (pH 5.5) and 0.01 % sodium azide at a concentration of 2 mg/mL in presence of 400 μM LMW heparin (5 kDa heparin from CalBioChem) in 1.5 mL Eppendorf tubes and incubated again in the EchoTherm at 50 rpm and 37 °C. Fibril formation was monitored by EM, CD and time-resolved Thio T binding studies (as described in further detail in [102]).

3.4.4. TEM imaging

SST-14 fibril samples were diluted in buffer and adsorbed to thin carbon films for 60 s. These films lie across a thick fenestrated carbon layer on top of a 200 by 200 mesh per inch gold-plated copper grid (a STEM grid). The grids were subsequently blotted, rinsed with quartz double-distilled water, and stained with 2 % (w/v) uranyl acetate. For imaging of the samples, a CM10 transmission electron microscope was used, operative at 80 kV (Philips, Eindhoven, The Netherlands). Final images were recorded with a 2000 by 2000 pixel charge-coupled

device camera (Veleta; Olympus soft imaging solutions GmbH, Münster, Germany) at a nominal magnification of 60 000x (Figure 3.2A) and 100 000x (Figure 3.2B).

3.4.5. Solid-state NMR spectroscopy

All spectra were recorded on Bruker Avance III 600 and 850 MHz spectrometers, operating at static magnetic fields of 14.1 and 20.0 T respectively, using magic-angle spinning 3.2 mm wide-bore triple-resonance probes. Acquired spectra were processed with Topspin 3.2 (Bruker Biospin), using a shifted squared sine-bell (QSINE) apodisation function with SSB = 2.5 to 3, and zero-filling to the next power of two in all dimensions. ^{13}C data in the direct dimension were treated with an automatic baseline correction. Following previously described assignment procedures [135], the sequential assignment of diluted SST-14 fibrils was carried out in CcpNmr Analysis 2.4.2 [136, 137]. Detailed experimental parameters are given in Table 3.3.

3.4.6. Structure calculation

Structure calculation were performed with CYANA [89] version 3.97. A total of 300 structures were calculated with 35,000 steps per calculation, of which the 10 structures with the lowest target function values were selected for the structural bundle. All restraints that were used for the calculation are discussed in Section 3.2.5 of this thesis.

Experiment	NCACB	NCACX	NCOCX	CANCO
Sample	UL:NA 1:3 diluted	UL:NA 1:3 diluted	UL:NA 1:3 diluted	UL:NA 1:3 diluted
Spectrometer B_0 field / T	20.0	14.1	14.1	14.1
Probe	3.2 mm Bruker	3.2 mm Bruker	3.2 mm Bruker	3.2 mm Bruker
MAS / kHz	17	12	12	12
Measurement time / h	89	96	95	102
Number of scans	16	8	8	8
Interscan delay / s	2.7	2.6	2.6	2.8
Transfer 1	HN-CP	HN-CP	HN-CP	HC-CP
Time / ms	0.9	0.8	0.8	0.4
Field / kHz	$26(^{15}N)/50(^1H)$	$45(^{15}N)/56(^1H)$	$45(^{15}N)/56(^1H)$	$35(^{13}C)/56(^1H)$
Shape	Tangent(^{1}H)	Tangent(^{1}H)	Tangent(^{1}H)	Tangent(^{1}H)
Carrier ^{13}C/ ppm	-	-	-	60
Transfer 2	NCA-CP	NCA-CP	NCO-CP	CAN-CP
Time / ms	5	5.5	5	4
Field / kHz	$13(^{15}N)/3(^{13}C)/88(^1H)$	$14(^{15}N)/3(^{13}C)/90(^1H)$	$14(^{15}N)/3(^{13}C)/77(^1H)$	$14(^{15}N)/3(^{13}C)/84(^1H)$
Shape	Tangent(^{13}C)	Tangent(^{13}C)	Tangent(^{13}C)	Tangent(^{13}C)
Carrier ^{13}C / ppm	58	56	176	176
Transfer 3	DREAM	DARR	DARR	NCO-CP
Time / ms	2	35	25	6
Field / kHz	$7(^{13}C)/88(^1H)$	$12(^1H)$	$12(^1H)$	$14(^{15}N)/3(^{13}C)/77(^1H)$
Shape	Tangent(^{13}C)	-	-	Tangent(^{13}C)
Carrier ^{13}C / ppm	58	-	-	176
All ^{1}H,^{15}N carriers / ppm	4.8 / 120	4.8 / 120	4.8 / 120	4.8 / 120
t1 Increments	72	144	144	112
Sweep width / kHz	5	8	8	8
Carrier / ppm	120	120	120	60
Acquisition time / ms	8	9	9	7
td proc	512	512	512	256
window function	qsine 2.5	qsine 2.5	qsine 2.5	qsine 2.5
t2 Increments	96	112	112	144
Sweep width / kHz	8	8	8	8
Carrier / ppm	58	56	176	120
Acquisition time / ms	6	7	7	9
td proc	512	256	256	512
window function	qsine 2.5	qsine 2.5	qsine 2.5	qsine 2.5
t3 Increments	2048	1216	1216	1216
Sweep width / kHz	100	50	50	50
Carrier / ppm	58	100	100	176
Acquisition time / ms	10	12	12	12
td proc	8192	4096	4096	4096
window function	qsine 2.5	qsine 2.5	qsine 2.5	qsine 2.5
Decoupling	SPINAL64	SPINAL64	SPINAL64	SPINAL64
Field / kHz	90	90	90	90

Table 3.3. Conditions for solid-state NMR experiments of uniformly ^{13}C,^{15}N-labelled somatostatin-14 amyloid fibrils.

Experiment	CCC	NCO	NCA	PAR
Sample	UL:NA 1:3 diluted	UL:NA 1:3 diluted	UL:NA 1:3 diluted	UL:NA 1:3 diluted
Spectrometer B_0 field / T	14.1	14.1	20.0	20.0
Probe	3.2 mm Bruker	3.2 mm Bruker	3.2 mm Bruker	3.2 mm Bruker
MAS / kHz	10	17	17	17
Measurement time / h	175	15	9	161
Number of scans	8	32	24	224
Interscan delay / s	2.2	2.8	2.5	2.5
Transfer 1	HC-CP	HN-CP	HN-CP	HC-CP
Time / ms	0.4	0.8	0.9	0.6
Field / kHz	$36(^{13}C)/55(^1H)$	$45(^{15}N)/56(^1H)$	$26(^{15}N)/50(^1H)$	$50(^{13}C)/80(^1H)$
Shape	Tangent(^{1}H)	Tangent(^{1}H)	Tangent(^{1}H)	Tangent(^{1}H)
Carrier ^{13}C/ ppm	100	-	-	100
Transfer 2	DREAM	NCO-CP	NCA-CP	PAR
Time / ms	2.5	6.5	6.0	6.0
Field / kHz	$5(^{13}C)/87(^1H)$	$14(^{15}N)/3(^{13}C)/77(^1H)$	$13(^{15}N)/3(^{13}C)/88(^1H)$	$69(^{13}C)/69(^1H)$
Shape	Tangent(^{13}C)	Tangent(^{13}C)	Tangent(^{13}C)	-
Carrier ^{13}C / ppm	50	176	58	-
Transfer 3	DARR	-	-	-
Time / ms	80	-	-	-
Field / kHz	$10(^1H)$	-	-	-
Shape	-	-	-	-
Carrier ^{13}C / ppm	-	-	-	-
All ^{1}H,^{15}N carriers / ppm	4.8 / -	4.8 / 120	4.8 / 120	4.8 / -
t_1 Increments	192	608	512	1024
Sweep width / kHz	16	20	20	50
Carrier / ppm	50	120	120	100
Acquisition time / ms	6	15	13	10
TD proc	512	2048	2048	8192
Window function	qsine 2	qsine 3	qsine 4	qsine 2.5
t_2 Increments	176	1280	2790	2560
Sweep width / kHz	16	50	100	100
Carrier / ppm	45	176	58	100
Acquisition time / ms	6	13	14	13
TD proc	512	4096	8192	8192
Window function	qsine 2	qsine 3	qsine 4	qsine 2.5
t_3 Increments	1280	-	-	-
Sweep width / kHz	50	-	-	-
Carrier / ppm	100	-	-	-
Acquisition time / ms	13	-	-	-
TD proc	2048	-	-	-
Window function	qsine 2	-	-	-
Decoupling	SPINAL64	SPINAL64	SPINAL64	SPINAL64
Field / kHz	90	90	90	90

Table 3.3 (Continued) Conditions for solid-state NMR experiments of uniformly ^{13}C,^{15}N-labelled somatostatin-14 amyloid fibrils.

Experiment	DARR 50 ms	DARR 400 ms	CHHC	PAIN
Sample	UL:NA 1:3 diluted	UL:NA 1:3 diluted	UL:NA 1:3 diluted	^{13}C:^{15}N 1:1 mixed
Spectrometer B_0 field / T	20.0	20.0	14.1	20.0
Probe	3.2 mm Bruker	3.2 mm Bruker	3.2 mm Bruker	3.2 mm Bruker
MAS / kHz	17	15	12	17
Measurement time / h	15	184	150	160
Number of scans	8	80	288	352
Interscan delay / s	2.5	2.8	2.2	3.0
Transfer 1	HC-CP	HC-CP	HC-CP	HN-CP
Time / ms	0.6	0.6	0.6	0.4
Field / kHz	40(^{13}C)/65(^{1}H)	39(^{13}C)/65(^{1}H)	43(^{13}C)/54(^{1}H)	27(^{15}N)/45(^{1}H)
Shape	Tangent(^{1}H)	Tangent(^{1}H)	Tangent(^{1}H)	Tangent(^{1}H)
Carrier ^{13}C/ ppm	41	41	55	-
Transfer 2	DARR	DARR	HH	PAIN-CP
Time / ms	50	400	0.35	6
Field / kHz	17(^{1}H)	15(^{1}H)	-	57(^{15}N)/61(^{13}C)/60(^{1}H)
Shape	-	-	-	-
Carrier ^{13}C / ppm	-	-	-	41
Transfer 3	-	-	CH-CP	-
Time / ms	-	-	0.6	-
Field / kHz	-	-	43(^{13}C)/54(^{1}H)	-
Shape	-	-	Tangent(^{1}H)	-
Carrier ^{13}C / ppm	-	-	55	-
Transfer 4	-	-	HC-CP	-
Time / ms	-	-	0.6	-
Field / kHz	-	-	43(^{13}C)/54(^{1}H)	-
Shape	-	-	Tangent(^{1}H)	-
Carrier ^{13}C / ppm	-	-	55	-
All ^{1}H,^{15}N carriers / ppm	4.8 / -	4.8 / -	4.8 / -	4.8 / 118
t_1 Increments	2560	2560	2048	512
Sweep width / kHz	100	100	100	20
Carrier / ppm	41	41	55	118
Acquisition time / ms	13	13	10	13
TD proc	8192	8192	4096	2048
Window function	qsine 3.0	qsine 2.0	qsine 2.0	qsine 2.0
t_2 Increments	3072	3072	2560	2048
Sweep width / kHz	100	100	100	100
Carrier / ppm	100	100	55	41
Acquisition time / ms	15	15	13	10
TD proc	8192	8192	4096	4096
Window function	qsine 3.0	qsine 2.0	qsine 2.0	qsine 2.0
Decoupling	SPINAL64	SPINAL64	SPINAL64	SPINAL64
Field / kHz	90	90	90	90

Table 3.3 (Continued) Conditions for solid-state NMR experiments of uniformly ^{13}C,^{15}N-labelled somatostatin-14 amyloid fibrils.

4. Envelope and capsids of the hepatitis B virus

This work has been performed in collaboration with the group of Anja Böckmann at the IBCP in Lyon, where proteins were recombinantly expressed in *E. coli* as well as by cell-free protein synthesis, in particular by Shishan Wang, Guillaume David, and Marie-Laure Fogeron.

This chapter contains parts of the following publications:

[138] David, G.[#], Fogeron, M.L.[#], Schledorn, M.[#], Montserret, R., Haselmann, U., Penzel, S., Badillo, A., Lecoq, L., André, P., Nassal, M., Bartenschlager, R., Meier, B.H. & Böckmann, A. (2018). Structural Studies of Self-Assembled Subviral Particles: Combining Cell-Free Expression with 110 kHz MAS NMR Spectroscopy. *Angewandte Chemie International Edition*, 57(17).

[139] Lecoq, L.[#], Schledorn, M.[#], Wang, S., Penzel, S., Malär, A.A., Callon, M., Nassal, M., Meier, B.H., & Böckmann, A. (2019). 100 kHz MAS proton-detected NMR spectroscopy of hepatitis B virus capsids. *Frontiers in Molecular Biosciences*, 6(58).

[140] Wang, S., Fogeron, M.L., Schledorn, M., Dujardin, M., Penzel, S., Burdette, D., Berke, J.M., Nassal, M., Lecoq, L., Meier, B.H., & Böckmann, A. (2019). Combining cell-free protein synthesis and NMR into a tool to study capsid assembly modulation. *Frontiers in Molecular Biosciences*, 6(67).

# These authors contributed equally.
Protein expression and purification [GD, MLF, SW, LL]; NMR rotor filling [MS, LL, SP]; NMR spectroscopy [MS, LL, GD, AAM, MC, SP]; Sequential assignment [LL]; Data analysis [MS, LL, GD, SW, AB, BHM]; Study design and supervision [AB, BHM].

4.1. Introduction

Chronic hepatitis B virus (HBV) infection remains a global concern for public health despite the existence of effective vaccines [141]. This chapter presents solid-state NMR studies of two proteins which form the structural basis of the capsid and envelope enclosing the viral DNA (Figure 4.1). Infection is detected using one of these two proteins (a surface antigen) as a serological marker, and based on measurements of this marker, estimated prevalence has increased from about 248 million affected individuals in 2010 [142] to about 356 million in 2015 [141]; a 44 % increase compared to 6 % global population growth in the same period. The importance of fighting HBV is self-evident, and indeed both proteins are potential drug targets for treatment of infection [138, 143].

The HBV surface antigen (HBsAg) is the collective name for three similar proteins embedded in the viral envelope. The smallest variant, HBs S (where S is for 'small') is the topic of Section 4.4, in this case of *anatine* origin. An extension at the N-terminus of the protein with the preS2 domain is denoted HBs M (for medium), while an extension with both the preS2 and preS1 domains forms HBs L (for large). Besides the surface antigen, the viral genome encodes only three other proteins: the polymerase P, the X protein HBx, and the HBV core protein Cp. The Sections 4.2 and 4.3 describe work with *human* Cp in various forms.

More in particular, Cp is a 183-residue protein with two domains: the N-terminal assembly domain that forms the capsid shell built of 240 Cp molecules. It self-assembles into this supramolecular structure even in the absence of the C-terminal domain (CTD, residues 150-183) that is responsible for RNA packaging, amongst other functions [145]. In HBV infected cells, Cp packs pregenomic RNA into the capsid upon assembly [146], as well as a copy of the viral polymerase [147]. The RNA inside the capsid is subsequently transcribed to double-stranded relaxed circular DNA; such 'mature capsids' are ready for envelopment. The basic building block of these capsids is the 20 kDa core protein that is mainly α-helical, and forms a stable dimer through a four-helix bundle which protrudes from the capsid surface as a 'spike', as can be seen in the insert in Figure 4.2B.

The colouring in this figure indicates four groups of symmetrically non-equivalent subunits, each consisting of otherwise chemically identical Cp molecules. Sixty dimers of subunits A and B, and sixty more of subunits C and D, assemble into an icosahedral capsid with T = 4 symmetry, where 12 fivefold vertices are exclusively composed of subunit A (in dark blue), and 30 quasi-sixfold vertices where the other subunits meet (in light blue, red, and yellow). The asymmetry of the subunits can be detected by NMR, as is shown for capsids
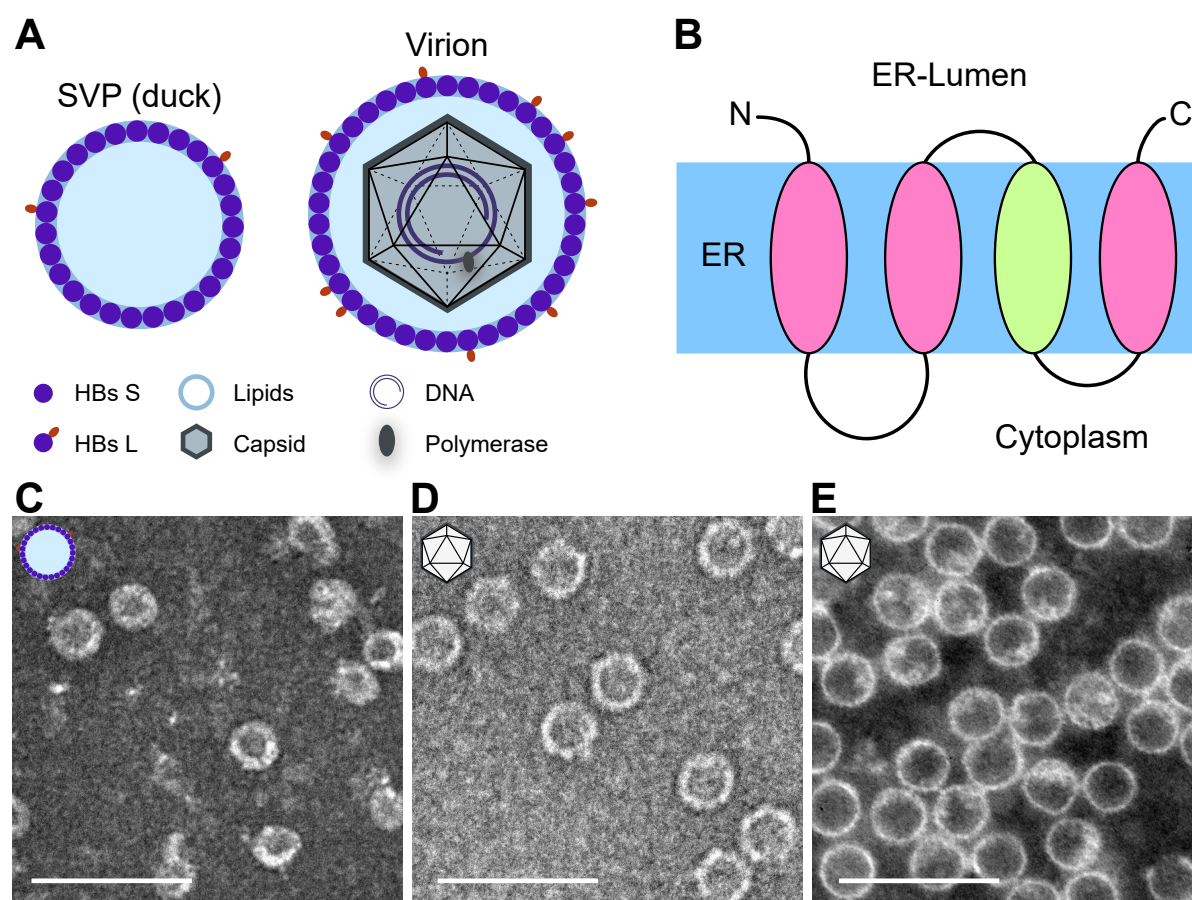
Figure 4.1. *Hepatitis B virus envelope and capsids.* **(a)** Schematic representation of the HBV virion with a protein-lipid based envelope surrounding a capsid with typical icosahedral symmetry (see also Figure 4.2B). The capsid consists of the core protein Cp and encloses partially double-stranded DNA and the viral polymerase. This is contrasted to the sub-viral particle (SVP), an 'empty envelope' that consists of a phospholipid membrane with the embedded proteins HBs S (predominantly) and HBs L (sparsely). The representation in (a) is for a duck SVP, while human HBV SVPs also contain the HBs M protein. **(b)** Membrane topology model of the HBs S protein, with the third helix that is only predicted for the HBs S protein (in light green). **(c-e)** Negative stain electron micrographs of, from left to right, duck HBs S SVPs (^{2}H,^{13}C,^{15}N-labelled), human Cp183 capsids (^{13}C,^{15}N-labelled), and human Cp149 capsids (^{13}C,^{15}N-labelled). The three images are scaled to the same magnification to illustrate the somewhat smaller size of SVPs compared to the virion. Scale bars are 100 nm. This figure consists of combined and edited components from figures in [138, 140, 144].

consisting of the assembly domain Cp149 only (Figure 4.3B and Section 4.2.3, also discussed in [148]). Detailed descriptions of materials and methods are given in Appendix Section C, p. 155, as well as in [138–140, 149].

4.2. HBV Cp149

From the perspective of the previous two chapters on Aβ (Chapter 2) and Somatostatin-14 (Chapter 3), this section of the current chapter links the work on HBV capsids to the development of faster MAS and ^{1}H detection, as discussed in the Introduction (Section 1.3, p. 27). It describes sequential assignment of fully protonated Cp149 using proton-detected solid-state NMR experiments at 100 kHz MAS, and makes a comparison to carbon-detected experiments in 3.2 mm rotors at 17.5 kHz MAS. In addition, a protonated preparation (*pCp149*) and a deuterated and back-exchanged preparation (*dCp149*) are evaluated with respect to each other, while linewidth, sensitivity, and back-exchange are discussed.

4.2.1. Sequential assignments of the amide proton resonances

For amide proton assignments, the pCp149 capsid sample was used. Backbone atoms were assigned using four 3D spectra: an hCANH showing intra-residue connections; hCONH and hCAcoNH spectra showing inter-residue connections, and an hncaCBcaNH connecting the Cβ (and Cα) to the NH (see Appendix Table C.2). Representative extracts of the spectra are shown in Figure 4.2A. The four spectra allowed for assignment of 90 % of the H^N resonances of the protein for residues 1-139 (the ten residues from the NTD-CTD linker are excluded from the statistic analysis as they are not visible in the NMR spectra). Using only the proton-detected spectra shown in Figure 4.2, 83 % of N (including prolines), 91 % of Cα, 23 % of Cβ and 80 % of C' could be assigned. Peak assignments are summarised as the mean chemical shift from the combined 3D experiments, and back-predicted onto the 2D hNH spectrum in Figure 4.2B.

4.2.2. Comparison to ^{13}C-detection-based resonance assignments and SNR

The ^{15}N and ^{13}C backbone and sidechain resonances of pCp149 capsids were assigned in previous work, using ^{13}C-detected solid-state NMR at 17.5 kHz MAS in a 3.2 mm (thin-wall) rotor [144]. There, the assignment for residues 1-139 was complete at 97 % for N, C', Cα and Cβ atoms, and 76 % of sidechain heteronuclei. C-terminal residues 140 to 149 are excluded from the statistics as no cross peaks were detected in either ^{13}C- or ^{1}H-detected experiments. Except for this region, only 4 residues are missing from the ^{13}C-detection-based assignment
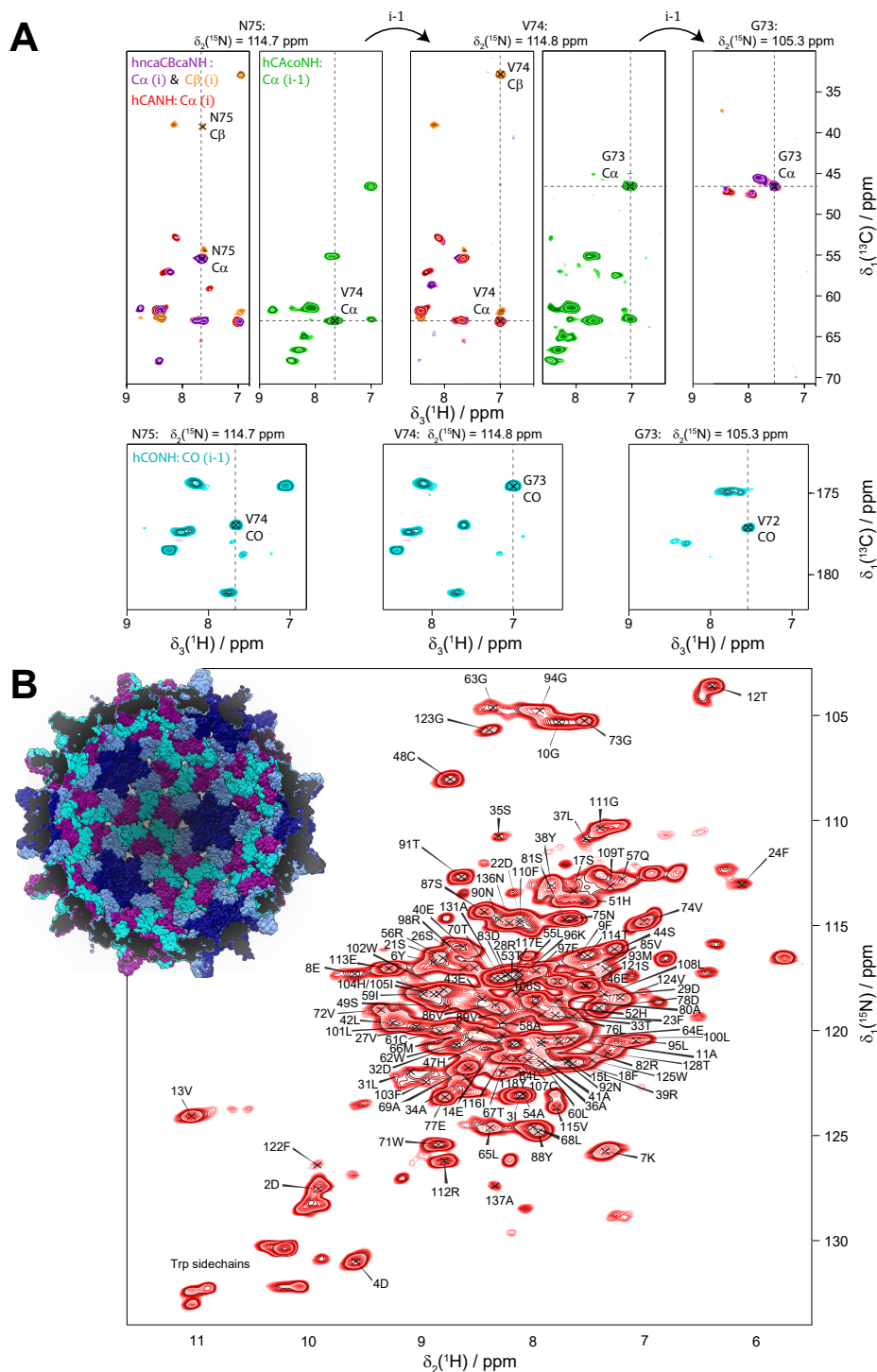
Figure 4.2. *Sequential resonance assignment of Cp149.* **(a)** 3D planes of pCp149 capsid spectra showing a sequential walk for residues N75, V74 and G73. Four spectra were used: hncaCBcaNH in purple (positive peaks: Cα) and orange (negative peaks: Cβ); hCANH in red (Cα); hCAcoNH in green (Cα[i-1]); and hCONH in cyan (CO[i-1]). **(b)** Assigned hNH spectrum of pCp149 capsids. The insert is an illustration of the four subunits of the capsid structure (in shades of blue), based on PDB entry 1QGT [150].

of HBV capsids. Compared to the ^{13}C-detection based procedure, in the ^{1}H-detected spectra an additional 11 residues could not be assigned due to low intensity or missing signals (L16, L19, L30, Q99, V120, S121, R127, P129, Y132, P138, I139, see Appendix Figure C.2). Thus, the ^{13}C-^{15}N backbone assignment completeness using ^{1}H-detection is slightly lower when compared to ^{13}C-detection NMR. Of course, the latter completely lacks information on ^{1}H shifts, due to the strong proton line broadening at 17.5 kHz.

Experimental parameters used for proton and carbon-detection spectra are compared in Figure 4.3A. Measurement times are roughly the same between the two approaches for this protein (despite the difference in sample amount of about a factor of one hundred), with about 8-12 h used to record a good 2D fingerprint spectrum (C-C DARR or NCA for carbon-detection, and hNH for proton-detection), and about 2 weeks to record the set of 3D experiments used for the assignment. The reason for the lower assignment completeness in ^{1}H-detection lies mainly in the signal-to-noise ratio (SNR) per square root time of the proton-detected spectra that is lower by a factor of about 2 (Figure 4.3). This brings the divided intensity of some residues with peak doubling or quadrupling due to asymmetry (*vide infra*) below the detection limit. These remained therefore unassigned. In detail, the SNR per square root time calculated on 2D experiments shows a gain by a factor of 1.8 between hNH and DARR-CC spectra, and of 2.4 between hNH and NCA spectra, based on the intensities of 14 representative residues (Figure 4.3F). Corresponding 3D spectra show an SNR per square root time increase by a factor of 2.3 between the hCONH and the CANCO spectra, and almost a factor of 3 between the hCAcoNH and the NCOCX spectra for the [Cα]$_{i-1}$ correlation signals (Figure 4.3F). With respect to the latter, one should note that alternative schemes exist for the hCAcoNH experiment, where J-coupling based transfers [151] could improve sensitivity under certain conditions. 3D NCACX spectra yield almost twice the SNR per square root time compared to hCANH spectra on the [Cα]$_i$ correlation signal, and it additionally provides the carbon sidechain shifts. Finally, the main difference is observed for the experiment leading to the Cβ assignment, where the NCACB shows better sensitivity than the hncaCBcaNH. Globally, the SNR per square root time is therefore still higher within similar experimental time when using 3.2 mm and carbon detection versus 0.7 mm and proton detection by a factor of about 2 for both 2D and 3D spectra. Yet, especially when available sample amounts prove limiting, the gain in mass sensitivity clearly outweighs the loss in SNR for this protein. Indeed, with a total mass here of 0.56 ± 0.04 mg in a 0.7 mm rotor versus 55 ± 3 mg in a thin-wall 3.2 mm rotor (mass ratio ~100, Appendix Table C.1). Even when considering a loss of roughly a factor of two in SNR, the overall gain in sensitivity remains ~50 in favour of the

[1]H-detection approach using 0.7 mm rotors. In addition, we compared the SNR for dCp149 capsids at 60 kHz in a 1.3 mm rotor (Figure 4.3F). We found a similar sensitivity for the 3D hCANH, even if the 2D hNH reveals a better SNR compared to pCp149 capsids at 100 kHz. With a total mass of 4.3 ± 0.1 mg in the 1.3 mm rotor (mass ratio ~8 compared to 0.7 mm), there is therefore no clear advantage of using 1.3 mm rotors rather than 0.7 mm rotors in terms of sensitivity for this protein.

One may wonder if the expected sensitivity effects could be predicted. Data and predictions related to sensitivity have been described in a recent publication [152] where a gain of a factor 1.5 was predicted, using proton (0.7 mm rotor) rather than carbon detection (3.2 mm rotor) for a 3D spectrum of microcrystalline SH3. As also detailed there, some parameters can be easily compared between the two approaches, as for example the rotor volume (roughly a factor 1/100, *vide supra*), the increased efficiency of the smaller coil which we estimated to scale with the inverse of the rotor diameter, namely $3.2/0.7 \approx 4.6$ (and $1.3/0.7 \approx 1.8$) [153, 154]. A factor of 8 is gained by detecting protons instead of carbons. A comparison of these parameters results indeed in a factor around 2.5 in favour of carbon detection (for a simple 1D spectrum detected starting with proton equilibrium polarisation). In principle, other factors need to be taken into account. For instance, the ratio of the product of the linewidth in all dimensions of the experiment (in Hz), which is approximately one, since the proton linewidth is often comparable with the carbon linewidth. Next, the efficiency of the probe circuits, the noise figure of the preamplifier, and efficiencies of the different polarisation transfers used during the 2D and 3D experiments play a major role. Because these values are highly dependent on the protein, temperature, spectrometer and/or probe parameters, it is difficult to make good predictions for the relative SNR in proton-detected (0.7 mm) and carbon-detected (3.2 mm) experiments. Nevertheless, for full rotors and to a reasonable approximation, the relative SNR in both cases is roughly equal, within a factor of two or three.

Nonetheless, [1]H-detection has the crucial advantage to provide amide-proton resonances. These are not only sensitive probes for conformational differences (Figure 4.3B), but also for non-covalent interactions such as hydrogen bonds or ring-current effects (Chapter 5). We conclude that in cases where enough sample is available, the carbon- and proton-detected approaches are truly complementary. Obviously, when sample amount is limiting, proton detection above 100 kHz is a must, and clearly shows competitive sensitivity.

A

	¹³C-detection	¹H-detection
MAS	17.5 kHz	100-110 kHz
Rotor size	3.2 mm	0.7 mm
Protein quantity	~30 mg	< 0.5 mg
Labeling	¹³C-¹⁵N	¹³C-¹⁵N & ²H-¹³C-¹⁵N
Temperature	~4 °C	~22 °C
Fingerprint experiment	2D DARR CC or NCA	2D hNH
Fingerprint exp. time	~12 h	~12 h
3D experiments	NCACX, NCOCX, CANCO, NCACB	hCANH, hncaCBcaNH, hCONH, hCAcoNH
Total experimental time	~15 days	~15 days
Backbone assignment	97 %	85 %

B

A137 N-Cα @ 17.5 kHz   A137 N-H @ 100 kHz

C

D

E

F

¹³C-detection

Protonated Cp149 @ 17.5kHz

Spectrum	Exp time / h	SNR	SNR / √h
2D NCA	8	28.2	10.0
2D DARR	15	28.7	7.4
3D NCACB	60	27.9	3.6
3D CANCO	92	13.0	1.4
3D NCACX	72	11.2	1.3
3D NCOCX	141	10.4	0.9

¹H-detection

Protonated Cp149 @ 100 kHz

Spectrum	Exp time / h	SNR	SNR / √h
2D hNH	12	14.5	4.2
3D hCANH	22	3.8	0.8
3D hCONH	66	4.8	0.6
3D hCAcoNH	130	3.4	0.3
3D hncaCBcaNH	142	2.5	0.2

Deuterated Cp149 @ 110 kHz

Spectrum	Exp time / h	SNR	SNR / √h
2D hNH	12	17.3	5.0
3D hCANH	23	5.4	1.1

Deuterated Cp149 @ 60 kHz

Spectrum	Exp time / h	SNR	SNR / √h
2D hNH	11.2	26.3	7.8
3D hCANH	23	3.8	0.8

Figure 4.3. *Proton- and carbon-detection.* **(a)** Comparison of acquisition parameters and assignments of Cp149 capsids (residues 1-139) using ¹³C-detected experiments recorded in a 3.2 mm rotor (thin-wall) at 17.5 kHz MAS, and ¹H-detected experiments recorded in a 0.7 mm rotor at 100 kHz MAS. The backbone assignment percentages include N, Cα and C'. **(b)** Extract of hNH (in red) and NCA (in blue) spectra showing peak splitting for residue A137, which displays four resonances corresponding to the four subunits of the pCp149 capsid (Figure 4.2B). Both 2D spectra were recorded in ~8 h. The NCA extract is adapted from previous data [144]. **(c)** Relative intensities of the four signals of A137 in the hNH and NCA spectra. **(d)** Signal to noise ratio of the four signals of A137 in the hNH and NCA spectra extracted by dividing the peak's height by the noise rmsd and by the square root of the experimental time in hours. A factor of 1.0625 was applied to the NCA spectrum to take into account the field difference (850 MHz for ¹H-detection *versus* 800 MHz for ¹³C-detection). **(e)** H, N, Cα and Cβ CSPs between the four subunit peaks of A137, with CSP = $(\gamma_x/\gamma_H)|\Delta\delta_x|$ (x, nucleus of interest; Δδ chemical shift difference between the peak of interest and the reference peak, here number 1). **(f)** SNR divided by the square root of the experimental time in hours for ¹³C-detected and ¹H-detected experiments. SNR was calculated based on the intensities of 14 representative and isolated peaks (D2, D4, G10, T12, V13, D22, E46, C48, W71, D78, T91, C107, G111 and I116) and compared to the mean noise of each individual spectrum as estimated in CcpNmr [136]. For the DARR, the probe was set to double-resonance mode (¹⁵N-insert was removed). All ¹³C-detected spectra were recorded on the same thin-wall 3.2 mm rotor as described in [144]. All ¹H-detected spectra of pCp149 and dCp149 were recorded on the same rotors. The four rotors were full (Appendix Table C.1). A factor of 1.0625 was applied to carbon-detected spectra and a factor of 1.214 to proton-detected spectra at 60 kHz to take into account the field difference (850 MHz for ¹H-detection at 100-110 kHz versus 800 MHz for ¹³C-detection and 700 MHz for ¹H-detection at 60 kHz). For 3D spectra with multiple correlation peaks, the intensities of the following resonances were taken into account for the SNR calculation: H-N-Cα for the hncaCBcaNH, N-Cα-Cβ for the NCACX (except for glycines: N-Cα-C'), and N-[C'-Cα]i-1 for the NCOCX. For the deuterated samples, C107 and I116 were excluded from the calculation as they were not back-exchanged (Appendix Figure C.1).

4.2.3. Proton chemical shifts are more sensitive to detect capsid subunits

Previous work by Lecoq *et al.* [148] demonstrated peak splitting in carbon-detected spectra due to the presence of the A, B, C, and D subunits in the icosahedral HBV capsid (see the four colours in the capsid illustration in Figure 4.2B and reference [155]). This behaviour was observed for residues A11-T12, L16-D22, T33-S35, L108-F110, V115, S121-W125, R127-T128, P130-Y132, and N136-P140, representing a total of 28 residues out of 139 visible ones, *i.e.* about 20 % of the protein. In proton-detected spectra, NMR peak splitting was detected for roughly the same residues, including A11-T12, D22, A34-S35, T109, V115, F122-V124, T128, A131, and N136-A137. This represents a total of 14 residues out of 128 possibly visible NH correlations (prolines were removed), *i.e.* about 11 % of the protein. The congruency between the two approaches shows once more that the high sensitivity of NMR to detect asymmetric features in molecular assemblies applies for proton chemical shifts as well. The remaining 9 % of peaks for which one would also expect a splitting in the proton spectra are too close to the noise and could therefore not be detected.

For residues for which chemical shifts are different for the four protein subunits, the proton's high sensitivity to the different chemical environments for these subunits led to larger chemical shift differences between them. This is illustrated for residue A137 in Figure 4.3B, which shows extracts of the hNH and NCA spectra of pCp149 capsids, with the signals assigned to the four Cp subunits. The respective intensities of the four correlation peaks are displayed in Figure 4.3C, and show roughly the same distribution in both spectra. Figure 4.3D shows that each of the four A137 signals has a higher SNR in the NCA than in the hNH when comparing two experiments with a similar experimental time ($\sim$ 8 h). The chemical shift perturbations (CSPs) between the peaks of the four different subunits reveal that, while Cα remains almost unaffected, HN, N and Cβ nuclei are the most sensitive for A137, with ^{1}H showing CSPs up to nearly 1 ppm (Figure 4.3E). This emphasises the complementarity of the three types of nuclei to measure CSPs as indicator in interaction studies.

4.2.4. ^{1}H and ^{2}H: incomplete back-exchange *versus* proton linewidths

The choice between deuterated and protonated proteins for NMR studies using proton detection has been discussed in recent literature [156–159]. It has been suggested that MAS frequencies above 100 kHz present an important step towards the use of fully-protonated proteins, since resolution starts to become high enough to resolve most resonances in 100-200 amino-acid residue proteins. This allows to bypass the more complex deuterated sample preparation and, even more importantly, back-exchange of amide protons, which can be difficult without a denaturation-renaturation step if they are not solvent accessible. Full protonation also represents an advantage to ease access to sidechain resonances. A systematic linewidth comparison of protonated and deuterated ubiquitin at 126 kHz MAS is given in [28], and predicts significant differences in linewidths (> 80 Hz) for this MAS frequency.

While we used pCp149 capsids for assignments, we wanted to assess which advantages, if any, can be obtained by the use of dCp149. The hNH spectra of the pCp149 and dCp149 capsids are shown in Figures 4.4A and B. A 1D extract of C48 HN is shown as inset to illustrate the observed difference in proton linewidth on a single resonance. Figure 4.4C shows the FIDs of the 2D spectra in the direct acquisition dimension, where the signal decay in the proton dimension can be seen. The signal decay is clearly faster in the protonated sample than in the deuterated one. To quantify this, we measured bulk T_2' relaxation times, which resulted in 2.5 ± 0.1 ms for the protonated sample and 11.6 ± 0.2 ms for the deuterated sample (Appendix Figure C.3). This shows that the homogeneous contribution to the linewidth is smaller in dCp149 than in pCp149 by a factor of 4.6, as is expected from dilution of the strongly coupled proton network with deuterium.

We also measured a 2D hNH of dCp149 at 60 kHz in a 1.3 mm rotor for comparison, which is shown in Appendix Figure C.4A. The proton linewidths of the three samples were measured for the assigned residues in the 3D hCANH spectra at 60 kHz and 100 kHz for dCp149, and 110 kHz for pCp149. The resulting linewidths are shown in Figures 4.4D-E and Appendix Figures C.4B-C. The mean proton linewidth is 168 ± 7 Hz for pCp149 at 100 kHz, 113 ± 4 Hz for dCp149 at 110 kHz, and 128 ± 6 Hz for dCp149 at 60 kHz. This corresponds to a difference of a factor 1.48 ± 0.08 between the linewidths of pCp149 and dCp149 at 100-110 kHz. The difference between dCp149 in 0.7 mm rotor at 110 kHz and dCp149 in 1.3 mm rotor at 60 kHz is a factor 1.14 ± 0.07. The improvement of 15 Hz between
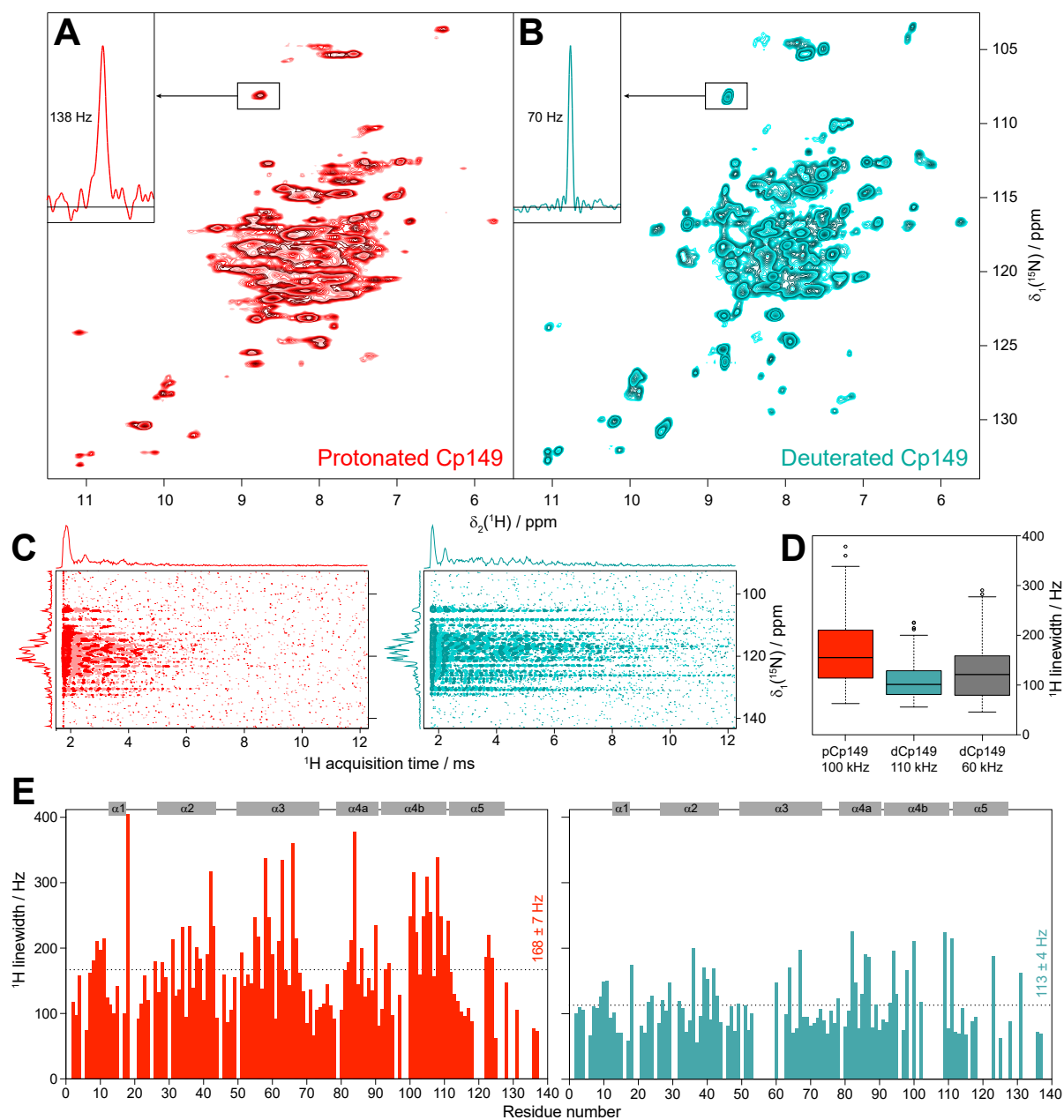
Figure 4.4. *Protonation and deuteration.* hNH spectra of **(a)** pCp149, and **(b)** dCp149 capsids. The proton line of C48 is shown in the inserts. Both spectra were recorded with identical acquisition parameters (100 kHz MAS, 40 scans, 12 h experimental time, VTU 273 K, and 25 ms acquisition in the ^{15}N dimension) and processing parameters (without apodisation function, cut at 12.9 ms acquisition in the ^{1}H dimension and zero-filled to 4096 and 1024 points in ^{1}H and ^{15}N dimensions, respectively). **(c)** FIDs from hNH spectra processed in the time domain using the xf1 command in TopSpin for Cp149 protonated (red) and deuterated capsids (cyan). **(d)** Dispersion of total proton linewidths measured on all assigned residues in 3D hCANH spectra of pCp149 at 100 kHz (red) and dCp149 capsids at 110 kHz (cyan) and 60 kHz (grey). Median values are indicated as a black line within each box and outliers as circles, defined as exceeding 1.5 times the interquartile range above the third or below the first quartile. **(e)** Total proton linewidths for assigned residues observed in the 3D hCANH spectra of pCp149 (red) and dCp149 capsids (cyan) using parabolic fit in CcpNmr [136]. Both 3D spectra were run with a similar experimental time of about 22 h and were processed identically (without apodisation function, cut at 12.9 ms acquisition in ^{1}H dimension and zero-filled to 2048 points in ^{1}H dimension and 128 points in ^{13}C and ^{15}N dimensions). The mean linewidth for both samples is indicated as a dotted line and secondary structure elements are illustrated at the top of each graph.

proton linewidths measured on dCp149 at 110 kHz compared to 60 kHz is close to values obtained for deuterated ubiquitin (improvement of 13 Hz in [28]) and for deuterated GB1 (improvement of only 6 Hz in [156]).

Despite the better resolution of dCp149 *versus* pCp149 at MAS ≥ 100 kHz, it should be mentioned that 15 % of the assigned residues are completely missing from the 2D and 3D spectra of the deuterated sample due to incomplete back-exchange (Appendix Figures C.1A and C.4C). In particular, mainly two regions are affected, namely A54 to T62 and L101 to L108, both located at the base of the spike where the bottoms of helices α3 and α4 interact in the dimer. These residues are buried and thus protected from the solvent, as shown in Appendix Figure C.1B. Only the sidechains of I59, W62 and I116 are accessible, while the H^N remain inaccessible. In HBV capsids, the advantage of deuteration obtained in linewidths must therefore be weighed against the disadvantage of incomplete back exchange, which obscures information on residues at the centre of the capsid spike.

4.3. HBV Cp183

Below, the HBV core protein Cp remains the topic of interest, but in its full-length form. In addition, rather than recombinant expression in *E. coli*, cell-free protein synthesis (CFPS) is exploited here, although protein from both methods is compared. As mentioned in the introduction to this chapter (p. 76), the HBV capsid and envelope that protect the viral genome and allow for cell entry and fusion are crucial targets for treatment of infection. Existing treatment aims at suppression of virus replication [160–162] but does not eradicate viral persistence [143]. Instead, capsid assembly modulators (CAMs) can induce disassembly of nucleocapsids, affecting multiple steps of HBV replication and reducing the existing pool of circular viral DNA [163, 164].

The advantage of CFPS, being an open system, is that activity of CAMs can be monitored during capsid assembly and subsequently studied by NMR spectroscopy, if CFPS is a suitable method for the latter. The presented data show that the this is indeed the case, and the addition of CAMs to the CFPS system produces results comparable to those from preformed Cp183 capsids to which CAMs were added afterwards.

4.3.1. Cell-free synthesised capsids analysed by NMR spectroscopy

Conformational details can be revealed by NMR in so-called fingerprint spectra, which show the typical signature of the protein preparation either in two or three dimensions. Structural variations can be sensitively identified by comparing spectra recorded under different conditions, and analysing the differences in the observed chemical shifts, *i.e.* the NMR frequencies [165]. An opportunity of the combination of CFPS and NMR is the fact that only the synthesised protein, which is the sole isotopically labelled protein, will be observed in the spectra. The use of a simple sucrose gradient concentration step thus might not produce perfectly pure protein; still, only the protein of interest will produce signal in the spectra. A possible drawback might lie in a loss of signal-to-noise ratio (SNR) in the spectra, since the NMR sample container (rotor) also might contain residual contaminating proteins (Appendix Figure C.5A). It is thus important to establish whether protein samples prepared by CFPS are compatible with the recording of 2D and in particular 3D spectra in a reasonable amount of time (see Appendix Table C.4).

The hNH 2D correlation spectrum recorded in 16 h on the protonated cell-free Cp183 displays a highly similar spectrum to the one recorded on the capsids purified from *E. coli* (Figure 4.5A and Appendix Figure C.7) in 10 h. The NMR signal amplitude of the sample from CFPS is about 35 % of the spectra obtained on the preparation from purified *E. coli* protein recorded under the same experimental conditions. As both rotors were full with protein sediment, this means that the contaminating unlabelled proteins from the WGE fill almost 2/3rd of the rotor. A 3D hCANH spectrum was recorded on the sample in 4 days and 15 h, and an overlay of all 3D NH planes onto the 2D NH plane shows that most signals in the 2D hNH spectrum are also observed in the 3D (Figure 4.5B).

The 2D spectrum recorded on the deuterated sample is shown in Figure 4.5C-D. The SNR is very favourable in this sample, since the deuterated protein surprisingly showed better purity (Appendix Figure C.5B). The spectrum reveals narrower lines than the spectrum from the protonated sample, as also observed in model systems [28] and, in particular, also in capsid preparations purified from *E. coli* (as also shown for Cp149, see Figure 4.4): 140 Hz on average for the protonated versus 100 Hz for the deuterated sample, as measured on six isolated resonances. The SNR and proton linewidths for the four samples are summarised in Figures 4.5E and F, respectively. It reveals that CFPS samples show a greater variability in sample amounts than the well-established *E. coli* samples, which in this case may stem
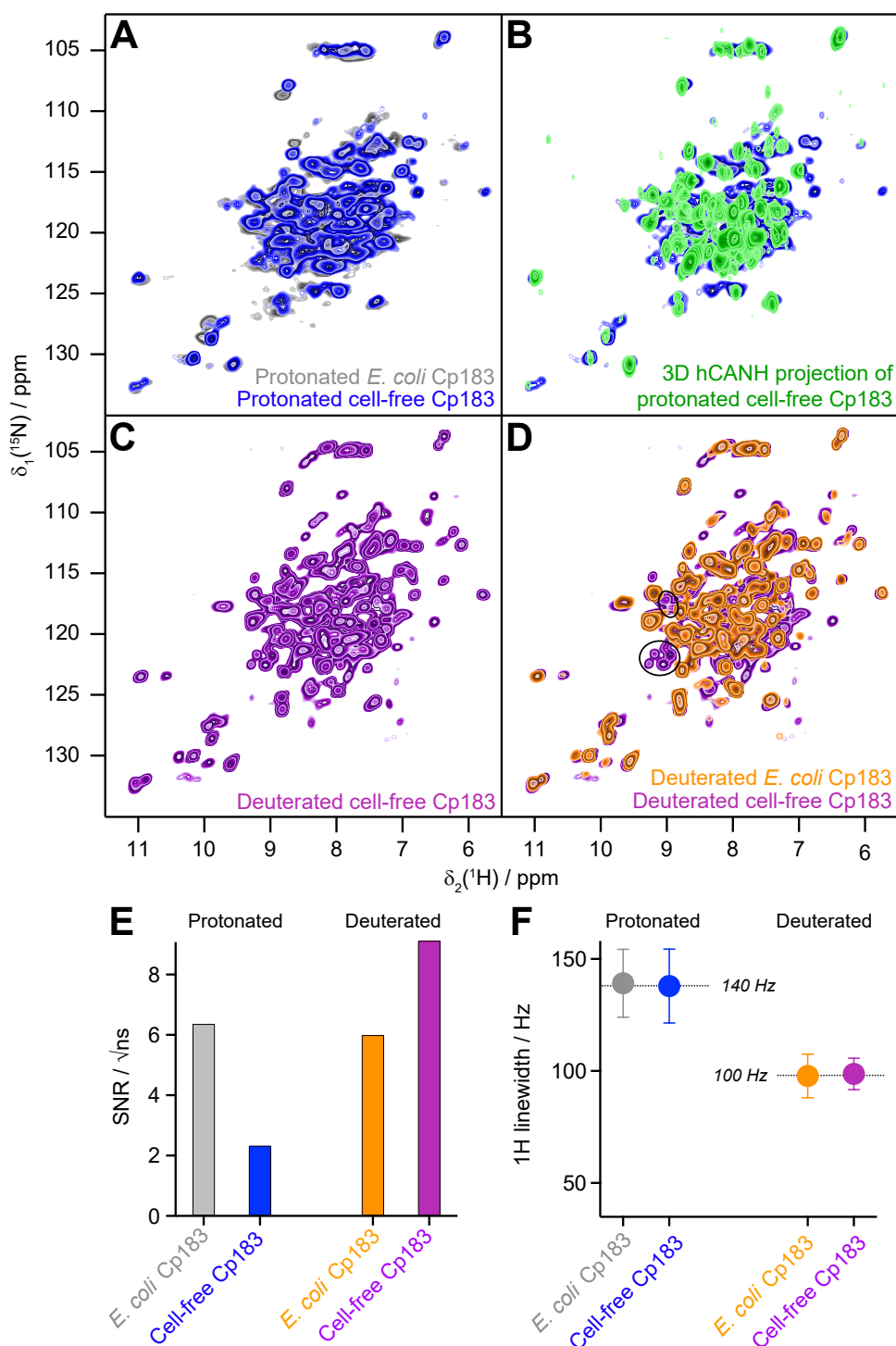
Figure 4.5. *Comparisons of NMR spectra between the capsids from CFPS and capsids purified from E. coli.*
(a) Overlay of the 2D hHN spectra of the protonated Cp183 capsids from CFPS (in blue) and purified from *E. coli* (in grey). **(b)** Overlay of 2D planes from the 3D hCANH spectrum recorded on protonated Cp183 CFPS capsids. **(c)** 2D hNH spectrum of the deuterated Cp183 CFPS capsids. **(d)** Overlay of the 2D hNH spectra of the deuterated capsids from CFPS (in purple) and purified from *E. coli* (in orange), with resonances not observed in the *E. coli* sample highlighted as black circles. Spectra are shown individually in Appendix Figure C.7. **(e)** Comparison of signal-to-noise ratios of the different samples. **(f)** Comparison of proton linewidths in the different samples. The mean values for the two protonated and the two deuterated samples are indicated.

from the concentration of contaminating proteins rather than the available amount of labelled Cp183. Further experience is needed to evaluate parameters allowing reproducible sample preparation using CFPS. The proton linewidths are virtually similar between the two protonated and two deuterated samples, indicating that production by CFPS or *E. coli* expression does not make a difference with respect to linewidth and therefore conformational homogeneity.

Importantly, several peaks are present in the cell-free synthesised dCp183 which could not be observed in the deuterated sample purified from *E. coli*, as emphasised in Figure 4.5D. The origin of this observation lies in the incomplete back-exchange in *E. coli* produced samples (see also Section 4.2.4 and Appendix Figure C.1). Indeed, when deuterated protein is expressed in *E. coli*, synthesis takes place in D_2O, and exchange of deuterons to protons is achieved during the subsequent purification steps, carried out in H_2O. Still, solvent-inaccessible deuterons can remain in the protein over long periods of time, and often denaturation/renaturation of the protein is applied to complete proton exchange. However, this step can be very difficult for more complex proteins, and the present experiment highlights this interesting feature of CFPS, where the protein is synthesised from the beginning in H_2O, and deuteration is achieved not via metabolism, but by addition of deuterated amino acids to the cell-free reaction. This results in fully protonated amide (and exchangeable sidechain) protons in the synthesised protein, which is essential for the recording of NMR spectra showing resonances for all amino acids. Of course, this substantial benefit should be weighed against relatively modest protein yields and a consequently rather costly sample.

4.3.2. Capsids synthesised in the presence of antiviral compounds

CFPS proceeds in an open system, and a variety of substances can be added to the reaction mixture. We added different capsid assembly modulators to the reaction, in order to analyse whether this produces comparable phenotypes to those observed on capsids purified from *E. coli*. Appendix Figure C.6A shows the Coomassie blue stained gels of the cell-free solutions without compounds, in the presence of DMSO used for solubilisation of the antiviral, and in presence of AT-130, JNJ-623 (CAM-N), and JNJ-890 (CAM-A). The corresponding Western blots are shown in Figure C.6B. None of the compounds inhibited protein synthesis. We analysed the total cell-free solutions, without any concentration or purification, under the
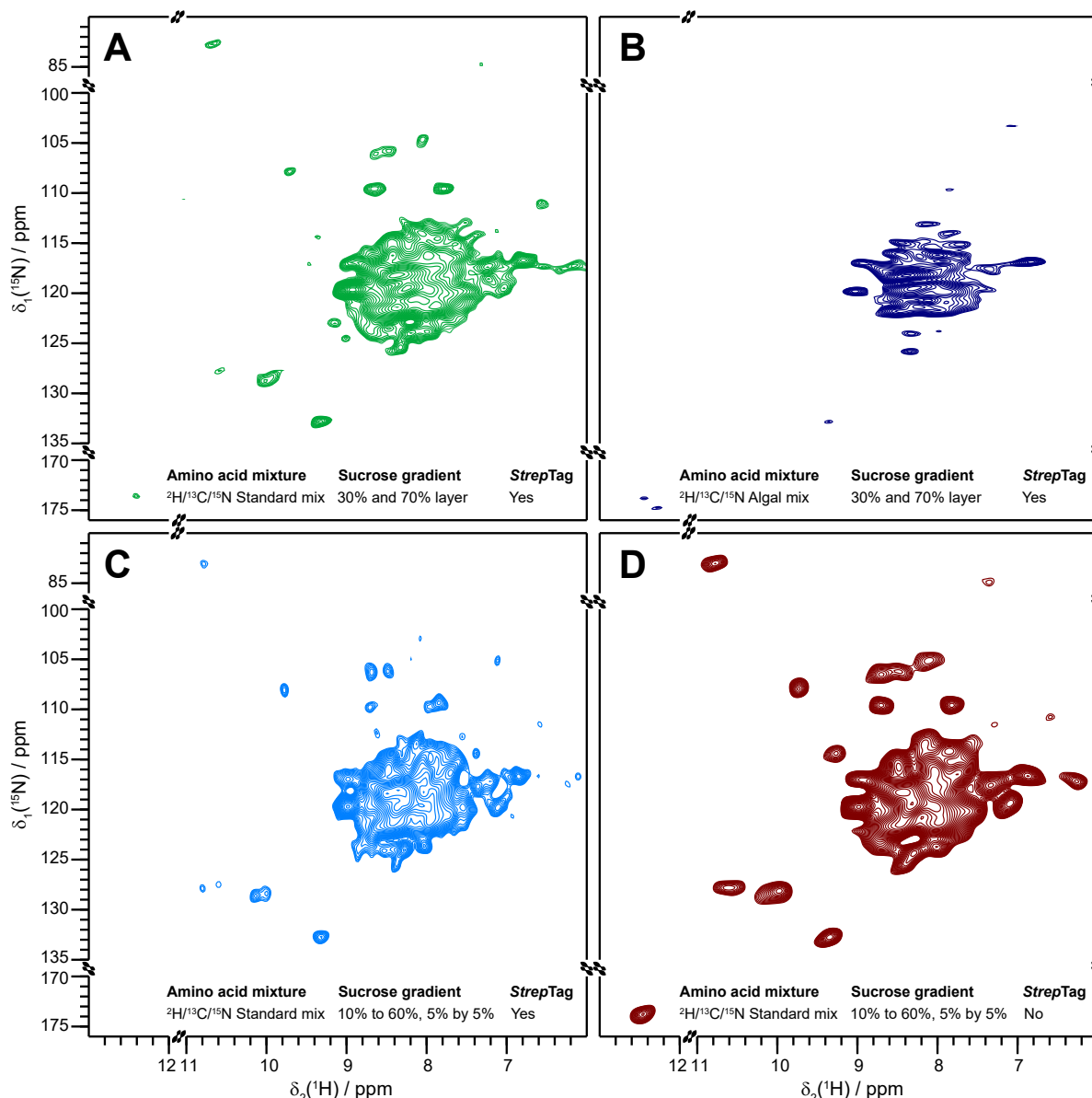
Figure 4.6. *Comparison of spectra recorded on different HBs S samples.* Main parameters that changed are written in the bottom-left corner. Samples a-c contained a *Strep*-tag II, which was removed in sample d as it was not used for purification. For samples a and b, a sucrose cushion was used in order to concentrate the protein in a small fraction, yielding volumes compatible with the filling tool with a 0.5 mL funnel. Better separation and thus sample purity was achieved using a layered gradient from 10 to 60 %, and a selection of most concentrated sections. This resulted in higher protein volumes, for which a new filling tool with a larger funnel (1.0 mL) was designed. For sample b, a triply-labelled algal amino-acid mix was used, which resulted in a lower expression yield, since it lacks certain amino acids. Even if supplementation with those amino acids did restore the yield to a large extent in small-scale tests, amino-acid mixes prepared from single amino acids were used for further samples. Experiment times were similar for all samples (20 h, 16 h, 11 h, and 19 h for samples a through d), but yielding a different SNR of 4, 3, 3, and 8 for isolated peaks in the spectra (again for samples a through d respectively). Spectra a-c were recorded at 100 kHz MAS, while d was recorded at 110 kHz. For other experimental parameters, see Appendix Table C.5.

electron microscope, and compared the observed capsids with the ones obtained from addition of compounds to capsids purified from *E. coli*. One can see in the micrographs (Figure C.6) that the resulting objects closely resemble those obtained by addition to preformed capsids: DMSO and CAM-Ns produced no visible effect, whereas CAM-As showed the typical disruption of capsids also reported in the literature [163, 166]. Notably, the presence of AT-130 led to poorer contrast in the EM micrographs of both preparations.

4.4. HBV envelope protein S

The final section of this chapter switches from the soluble HBV core protein to the membrane-embedded surface antigen HBs S. There is a topology model available for HBs S (Figure 4.1B) but, unlike for HBV Cp, there is no high-resolution structure determined yet. Although some work with *human* HBs S is ongoing (Appendix Figure C.10), the data presented below concern HBs S from duck HBV, which is similar to that of human HBV, but lacks its cysteine-rich antigenic loop, thus reducing the number of amino-acid residues from 226 to 167. As illustrated in Figure 4.1A, HBs S can form empty envelopes. Characteristic of HBV-infected cells is that they produce large amounts of such subviral particles (SVPs) mainly composed of HBs S, embedded in a membrane of phospholipids that are derived from the host [167]. Like in the previous section concerning Cp183, the HBs S SVPs were produced in a cell-free system. Besides structural characterisation of this serological marker for HBV infection, CFPS would allow for native-like interaction studies between HBs S and potential drug compounds.

4.4.1. HBs S CFPS

NMR samples of the protein were prepared using different protocols. Since affinity purification resulted in particle dissociation, it was not used. Instead, a sucrose gradient was used for purification and all samples had in common that the protein-containing fractions from the sucrose gradient were diluted twice and directly sedimented into the NMR rotor using dedicated filling tools [168]. Variation of parameters included sucrose gradients, amino-acid mixtures used for expression, rotor-filling procedures and constructs. We evaluated the sample quality using 2D ^{1}H-^{15}N 100-110 kHz MAS solid-state NMR spectra recorded on 0.7 mm NMR rotors, which are shown in Figure 4.6. It can be seen there that spectrum quality depended on sample preparation; with the use of an algal amino-acid mix resulting

in poor expression and signal-to-noise ratio (SNR), and with protein efficiently filled using modified filling tools (Figure 4.6A *vs.* B-D). Whereas first trials were made using productions containing 2 mL of WGE, the use of 0.5 mL proved sufficient in the end to fill a rotor. While here home-made WGE was used [169] resulting in costs of € 650 per triply labelled sample, the cost remains in the range of standard NMR samples also with commercial extracts (approx. total of € 1000 per sample). A more detailed discussion of the CFPS preparation of HBs S is given in [138] and the PhD thesis of Guillaume David [149].

4.4.2. HBs S solid-state NMR

Figure 4.7 shows the NMR spectrum recorded on the sample of Figure 4.6D where the unused *Strep*-tag II (for solid-state NMR sample preparation the protein is exclusively purified by sucrose gradient centrifugation), was removed from the expression plasmid. It shows the highest sensitivity (SNR = 7.6 ± 0.6, as measured on isolated signals) compared to all other preparations (Figure 4.6), and the 2D spectrum recorded in 19 h shows many distinct signals. The spectrum is processed with a 60° shifted sine-bell function (SSB = 3 in Bruker Topspin), and plotted just above the noise level. Appendix Figure C.8 shows the spectrum processed with other apodisation functions, in order to illustrate the artificial resolution-enhancement that can be achieved with apodisation. Linewidths for different signals (on a spectrum processed without resolution enhancement using a pure cosine wave, SSB = 2), span 128-192 Hz for the four resolved peaks marked by crosses in Figure 4.7.

Measured on series of 1D spectra, bulk relaxation times are 3.9 ± 0.6 ms for T_2', and 18 ± 3 ms for $T_1\rho$ (see Appendix Table C.5). Linewidths of DHBs S are larger than linewidths of microcrystalline proteins like ubiquitin [27], SH3 [170], or GB1 [171]. However, the linewidths are comparable to those observed for deuterated H^N back-exchanged and fully protonated model membrane proteins, *e.g.* proteorhodopsin [172] or OmpG [170], which show around 190 and 100 Hz proton linewidths respectively. 1D traces along different nitrogen dimensions are shown in Figure 4.7 to illustrate the signals in the proton dimension and their SNR. The signals in the spectrum are distributed over a spectral width that can be expected for a fully α-helical protein, which by nature shows a smaller chemical-shift dispersion than proteins containing β-strands or both α-helical and β-strand conformation. Indeed, the spectral region of the backbone amide correlations in the DHBs S spectrum strikingly coincides with NH correlations of a mainly α-helical protein, proteorhodopsin [173], and lacks the larger chemical-shift distribution of the β-barrel OmpXs [174], as shown in Ap-
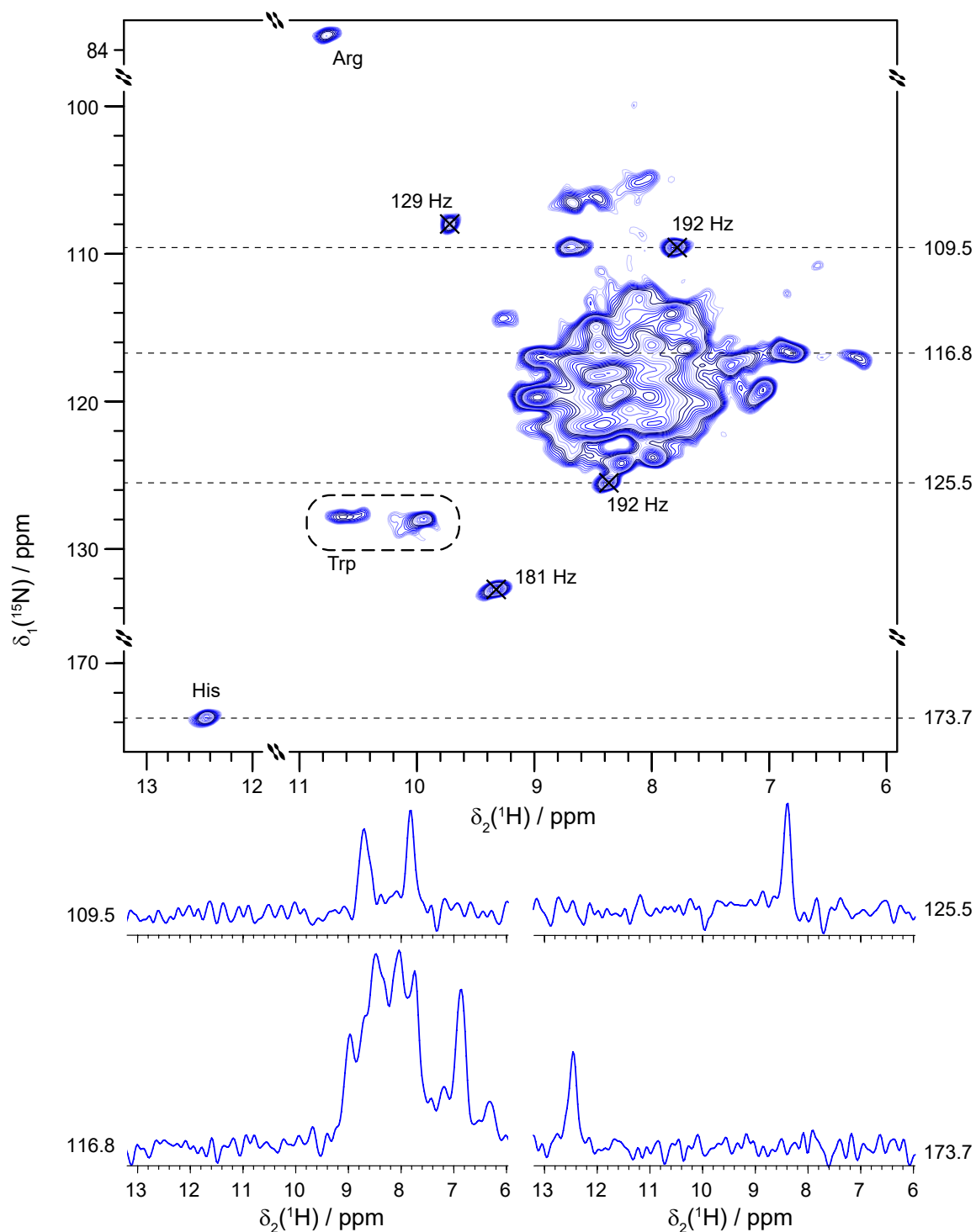
Figure 4.7. *2D ^{1}H-^{15}N spectrum of DHBs WT at 110 kHz MAS and a proton Larmor frequency of 850 MHz.* The 0.7 mm (outer diameter) rotor was filled using the 40 % and 45 % fractions of a discontinuous sucrose gradient which were directly sedimented into the rotor using a dedicated filling tool [168]. Linewidths here have been measured on spectra processed with pure cosine waves as apodisation function (SSB = 2 in Bruker Topspin), and are given for a selection of four signals (marked with a cross) likely to present single resonances. The 2D spectrum has been processed using a 60° shifted sine bell function (SSB = 3 in Bruker Topspin), and is plotted just above the noise level. 1D traces show extracts of the 2D plot to illustrate the signal-to-noise ratio. All values on the left and right edges of the figure give ^{15}N chemical shifts in ppm.

pendix Figure C.9. Besides the backbone NH correlations, one sidechain signal is observed
for His60 (the only His in the protein, located in the cytosolic loop) and three or four for Trp
(out of 7), with dispersed chemical shifts due to sequence-specific environments. The signal
in the upper left corner probably stems from an Arg sidechain (three in total), which shows
an outlier $H\varepsilon$ chemical shift. This might be indicative for interaction, presumably with the
membrane, since all three arginines are predicted to be located at the membrane interface,
at the beginning or the end of the cytosolic loop. One can notice a certain distribution of
linewidths and intensities of the peaks in the spectrum, which is likely to reflect variable
structure or order and/or dynamics in different parts of the protein, *e.g.* in transmembrane
helices or turns. Further studies of HBV envelope protein S would be necessary to uncover
these features.

4.5. Conclusion

The HBV Cp and HBs S proteins were brought to the fore from different perspectives.
Section 4.2 presented sequential resonance assignment of fully-protonated Cp149, using
100 kHz MAS solid-state NMR spectroscopy. Sequential assignment was obtained from a
set of four spectra, recorded within 15 days. This yielded 85 % backbone assignment, while
the previously obtained ^{13}C,^{15}N backbone assignment from ^{13}C-detected experiments was
about 10 % more complete. Besides being more complete, ^{13}C-detected spectroscopy al-
lowed for sidechain assignment within similar experimental time. However, the ^{1}H-detected
experiments yielded, obviously, H^{N} chemical shifts, besides N, Cα, Cβ, and C' chemical
shifts. The proton shifts can provide powerful information with respect to binding modes
through chemical-shift perturbations (see also Chapter 5), or through peak splitting from
structural asymmetry, as was shown here (Figure 4.3B) and elsewhere [148]. For interaction
studies of HBV nucleocapsids with drug molecules, such data may prove to be crucial for
understanding the drug compound's mechanism of action.

   The presented data in both Section 4.2 and 4.3 unequivocally show that spectral resolution
at 100 kHz MAS (and at 60 kHz) is better in deuterated and back-exchanged samples than
in fully protonated ones. Nevertheless, the near-complete assignment of Cp149 stemmed
from fully protonated protein spectroscopy that was evidently resolved enough for this
purpose. In addition, H^{N} resonances in parts of the protein with low solvent-accessibility
will be lost when working with deuterated material, due to incomplete back-exchange. In

this respect, HBV capsids are an important subject of study for treatment of infection, but also as a suitable model system for NMR studies of a more methodological character. The latter also shows from the reappearance of Cp149 in Chapter 6.

Sections 4.3 and 4.4 presented NMR data for capsids and SVPs that were successfully produced by CFPS. The 2D and 3D fingerprint spectra of Cp183 clearly illustrate that SNR and resolution are sufficient for more advanced NMR experiments and the investigation of conformational changes upon binding of *e.g.* CAMs. CFPS would allow for such interactions to take place directly on protein synthesis from the ribosome, and could include other interactions with for instance lipids, chaperones, or enzymes of relevance *in vivo*. In this respect, the presented data (including Appendix Figure C.6) show that the interaction of CAMs of three different classes interfere with capsid assembly in the same way for CFPS-self-assembled Cp183 and pre-formed capsids purified from expression *E. coli*. The effect of these CAMs can consequently be studied in CFPS-produced (NMR) samples in future research.

The same is true for the membrane protein HBs S from duck and human HBV, although continued structure studies would benefit from improvement in SNR in particular, and to a lesser extent from further improved resolution. Currently, the spectrum of duck HBs S confirms the topology model (Figures 4.1B and Appendix Figure C.9) suggesting a primarily α-helical structure. The compatibility of CFPS with HBs S would allow for selective labelling strategies that would be relatively easy to apply considering the open system, as was previously shown for a membrane protein from the hepatitis C virus [175]. Implementation of such labelling should allow for further structural characterisation of HBs S, and consequently for improved understanding of the viral envelope, hopefully contributing to alleviating the global burden of HBV viral infection.

5. Nucleotide binding modes in *Hp*DnaB helicase

This chapter is based on the following publication:

[176] Wiegand, T., Schledorn, M., Malär, A. A., Cadalbert, R., Däpp, A., Terradot, L., Meier, B. H., & Böckmann, A. (2019) Nucleotide binding modes in a motor protein revealed by ^{31}P- and ^{1}H-detected MAS solid-state NMR. *ChemBioChem*, 21(3).

Protein expression and purification [RC]; ^{13}C-and ^{31}P-detected NMR spectroscopy [TW]; ^{1}H-detected NMR spectroscopy and design of ^{1}H assignment strategy [MS]; Resonance assignment [MS, TW]; Relaxation experiments [TW, AAM]; Further data analysis [MS, TW, AAM]; Writing of the original manuscript [TW]; Discussion and editing of final text [MS, TW, AAM, BHM, AB]; Study design and supervision [BHM, AB].

5.1. Introduction

In this chapter, data is presented for the dodecameric DnaB helicase from *Helicobacter pylori* (*Hp*DnaB) with a molecular weight of 12·56 kDa [177]. *Hp*DnaB is involved in unwinding dsDNA to prepare a ssDNA template for DNA replication. A schematic representation of this process is given in Figure 5.1. With *Hp*DnaB as a model, this chapter establishes approaches to identify protein-nucleic-acid interactions in large proteins by solid-state NMR.

Nucleotide-protein interactions play a central role in many biological functions, for example post-translational modifications, metabolism, or signalling. This chapter has a focus on two classes of interactions in particular: *i)* those that involve energy-providing reactions, where ATP is hydrolysed to yield energy to motor domains driving reactions [178, 179]; and *ii)* those that involve interactions with RNA or DNA. Binding of nucleotides, such as ATP and deoxyribonucleotides of DNA, occurs via non-covalent interactions including hydrogen bonds, electrostatic, and van-der-Waals interactions [180, 181] (the latter include interactions between aromatic rings [182]). These interactions have typically been studied by high-resolution X-ray crystallography [181, 183, 184]. However, many of the above-described scenarios involve protein complexes which are difficult to crystallise, and if they do crystallise they may diffract to insufficient resolution for a clear identification of said interactions. Therefore, solid-state NMR forms a welcome alternative to provide information about the molecular structure and dynamics of large biomolecular complexes, with samples that are simply sedimented into an NMR rotor (see p. 19). In previous studies, solid-state NMR has been used to identify residues at protein-RNA interfaces in smaller proteins [185–188].

Two approaches are particularly promising to probe protein-nucleotide interactions: phosphorus- (^{31}P) and proton- (^{1}H) detected spectroscopy. Distances between ^{31}P spins of DNA and ^{15}N spins of a protein have been measured using TEDOR experiments [185]. Intermolecular information can also be obtained from ^{31}P-detected, heteronuclear correlation experiments probing the spatial proximity of nucleotide ^{31}P and protein ^{15}N or ^{13}C nuclei [185, 189]. Proton chemical-shift values are highly sensitive to hydrogen bonding [190–193] as shown in theoretical [192, 193], but also in experimental studies [190, 191, 194]. Empirical correlations between the ^{1}H chemical-shift values of amide protons [195] and the strength of the hydrogen bond (characterised by the hydrogen bond distance) have been established for biological systems [193, 195–198]. Still, one has to keep in mind that proton chemical shifts can also be influenced by anisotropic neighbour effects, ring current
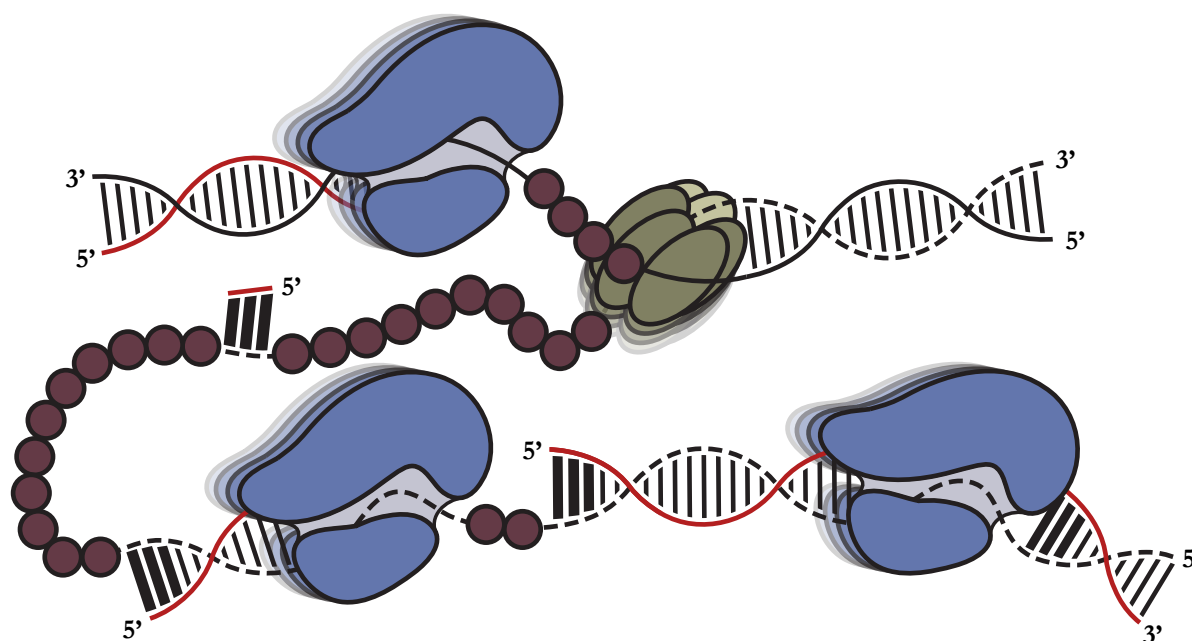
Figure 5.1. *Schematic representation of DNA replication.* The **top-right** of the figure shows the double-stranded parent strand. This double-stranded helix is forced apart by the DnaB helicase (in green), and the leading daughter strand continues outside of the helicase, to the **top-left** of the figure. While it is single-stranded, ssDNA-binding proteins (brown circles) stabilise and protect it, until the polymerase (in blue) completes the daughter strand with complementary nucleotides (in red). On the other hand, the lagging daughter strand leads through the core of the DnaB helicase down to the **bottom** of the figure. Associated with the helicase is a primase (not explicitly drawn) that synthesises short primers as starting points for the polymerase to add new nucleotides to the nascent double-stranded daughter. Since the polymerase can do this only in one direction, the process is continuous on the leading strand, but fragmentary on the lagging strand.

effects [199–201] and the secondary structure [202]. This underlines the importance of a combination with evidence from ^{31}P correlations delivering direct geometric information.

DnaB helicases are motor proteins that interact with both ATP and DNA during their functional cycle [178]. ATP and a Mg^{2+} cofactor are bound by the Walker A and B motifs (Figure 5.2) as well as the arginine finger (R-finger) connecting two adjacent subunits of the oligomeric assembly [203]. Instead, DNA binds in the central space of the hexameric proteins to so-called DNA binding loops. It has been revealed from crystal structures [181, 183, 204] that major coordination partners are Lys and Arg sidechains. Besides DNA ribose or base moieties [181], DNA phosphate groups are important recognition motifs through hydrogen bonds or salt bridges [205–208]. Using ADP:AlF$_4^-$ as an ATP analogue that mimics the hydrolysis transition state, it was previously established that this state preorganises the

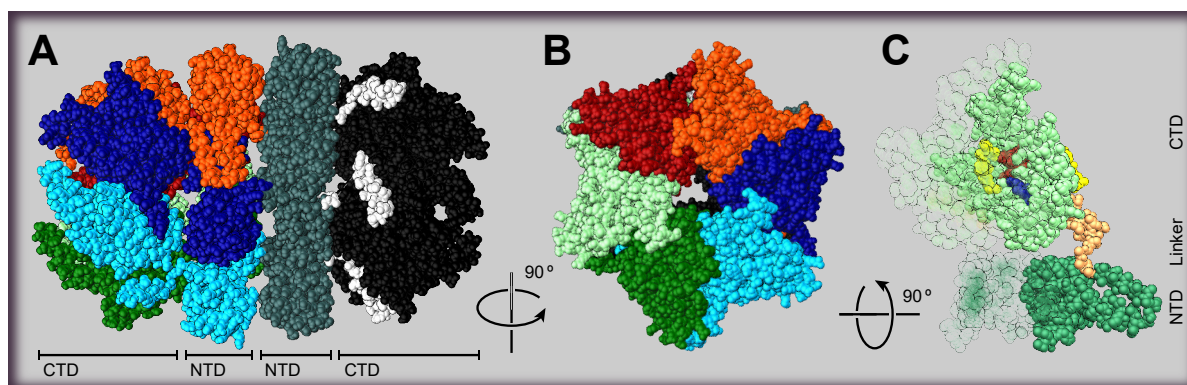**Figure 5.2.** *Structural model of HpDnaB.* **(a)** The left half of the dodecamer is coloured by subunit, while the right half is coloured by domain. This highlights the NTD (1M-145G) in grey, a linker (146S-176V) in white, and the CTD (177T-488T) in black. **(b)** Six subunits in a top view of the dodecamer. **(c)** One subunit is illustrated as opaque van der Waals spheres, with a second subunit in front of it as transparent spheres. The opaque subunit shows the phosphate-binding loop 203A-210T in red (Walker A motif), and the Mg^{2+} cofactor coordinating Walker B motif in blue (309I-316Q). The R-finger (445N-451G) binds the γ-phosphate and is shown twice, in yellow. The R-finger that is closest to the Walker A and B motifs belongs to the transparent neighbouring subunit, while the "own" R-finger is almost hidden behind the opaque subunit, to the far right of the CTD. The structure representations are based on PDB ID 4ZC0 [177] in which one of the CTDs is not resolved, but added here for the purpose of illustrating the full dodecamer.

helicase for binding single-stranded DNA to the C-terminal domain (CTD) in *Hp*DnaB [209]. Upon DNA binding, a large fraction of the protein covering the DNA binding loops rigidifies, including 357R and 373K, potentially involved in DNA binding [209]. Using ^{13}C-detected experiments, 70 % of the N-terminal domain (NTD) and approximately 60 % of the CTD (311 residues) of DnaB (^{15}N, ^{13}Cα and ^{13}Cβ, BMRB accession number 27879) were formerly sequentially assigned [210]. Also, previous work by Wiegand *et al.* [209] noted that the NTD is not observed in the DnaB:ADP:AlF$_4^-$:DNA sample, presumably due to dynamics, which reduces the number of observed amino-acid residues to the CTD.

For the ^{1}H resonances of the protein (~45 % of the CTD, see BMRB accession number 27879 and Appendix Table D.1), the assignment and the assignment strategy are described below, together with ^{31}P-^{15}N/^{13}C heteronuclear correlation spectra. This chapter demonstrates how these data can be used to reveal nucleotide-protein interactions, and to determine binding modes, in particular for DNA coordination by DnaB via the phosphate groups or base edges. It also highlights key residues involved in ATP and DNA binding, located in the Walker A motif and the DNA binding loops. Finally, a comparison is made to data for DnaB from *Bacillus stearothermophilus* (*Bst*DnaB) (as described in [184]), for which a crystal structure is available of the GDP:AlF$_4^-$:DNA-bound state.
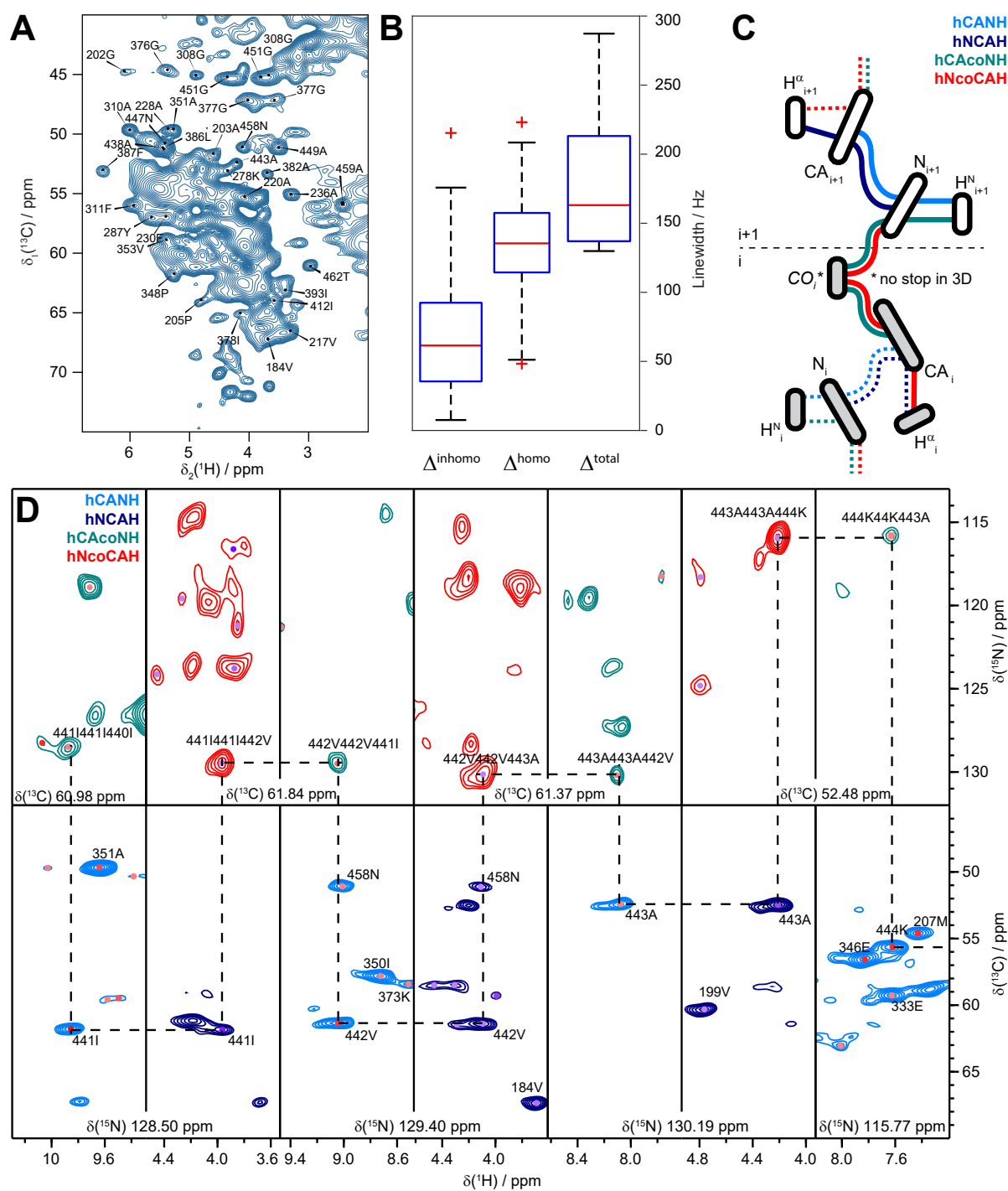
Figure 5.3. *Sequential assignments of the proton resonances.* **(a)** 2D hCH spectrum with assignments of isolated resonances. **(b)** Boxplots for the ^{1}H inhomogeneous, homogeneous, and total linewidth respectively, as described in Section 5.2.2. **(c)** Assignment strategy using both H^N and H$^\alpha$ protons. The dashed lines highlight connectivities at the same ^{13}C resonance frequency. **(d)** Representative 2D planes of 3D spectra used for sequential resonance assignment (*"backbone walk"*).

5.2. Results and discussion

5.2.1. Assigning fully protonated DnaB in complex with ADP:AlF$_4^-$ and DNA

H^N-and H$^\alpha$-detected two-dimensional hNH and hCH spectra of DnaB:ADP:AlF$_4^-$:DNA recorded at an MAS frequency of 110 kHz show a significant number of resolved signals, and are shown in Figure 5.3A and 5.6A. For resonance assignments, 3D hCANH, hNCAH, hCA-coNH and hNcoCAH spectra were recorded. The assignment strategy is shown in Figure 5.3C. It allows to walk along the protein backbone and delivers the H^N, H$^\alpha$, ^{13}Cα and ^{15}N chemical shifts. Each correlation between a pair of nuclei appears in two independent experiments, as already proven central in ^{13}C-detected experiments [135]. Figure 5.3D shows 2D planes of the 3D spectra illustrating the assignment strategy at the example of residues 441I to 444K (for a second example see Appendix Figure D.1). Around 45 % of the CTD H^N and H$^\alpha$ resonances (142 and 139 resonances, respectively) could be assigned, most of them in a sequential manner (100 and 118 correlations, respectively, were sequentially assigned in the *inter*residual hCAcoNH and hNcoCAH spectra, as sketched in Figure 5.3C). Further ^{1}H resonances were assigned by making use of ^{15}N and ^{13}Cα assignments obtained by ^{13}C-detected experiments. These allowed to find unambiguous ^{15}N/^{13}Cα/^{13}CO spin systems in ^{1}H-detected 3D experiments, to which the H^N and/or H$^\alpha$ resonance then could be added. The largest fraction of the NTD remains invisible in ^{1}H- and ^{13}C-detected spectra [209].

The assigned H^{N}and H$^\alpha$ chemical shifts are shown on a structural model of the CTD of the DnaB helicase [177] in Figure 5.4A/B, colour-coded with the corresponding ^{1}H chemical-shift values. These values are plotted in Figure 5.4C. Note that the H^N chemical shifts show a quite large spectral dispersion (*i.e.* ∼5 ppm). This is attributed to the high sensitivity of proton shifts to non-covalent interactions. The data show shielded resonances associated with ring-current effects (*e.g.* 314Y in the Walker B motif and 459A, 461F and 463R possibly located in a loop above the ADP base plane) and deshielded resonances due to hydrogen bond formation (many of them located in β-strands, see Figures 5.4C and 5.4 and Appendix Figure D.5). Assignments of outlier resonances are shown in Figure 5.3A and 5.6A . For both ^{1}H species, residues in β-strands were nearly completely assigned, while many residues in α-helices remained unassigned. This is a consequence of the larger chemical-shift dispersion of β-strand residues, but also of the usually broader homogeneous ^{1}H lines of residues in α-helices due to a denser proton network [26].
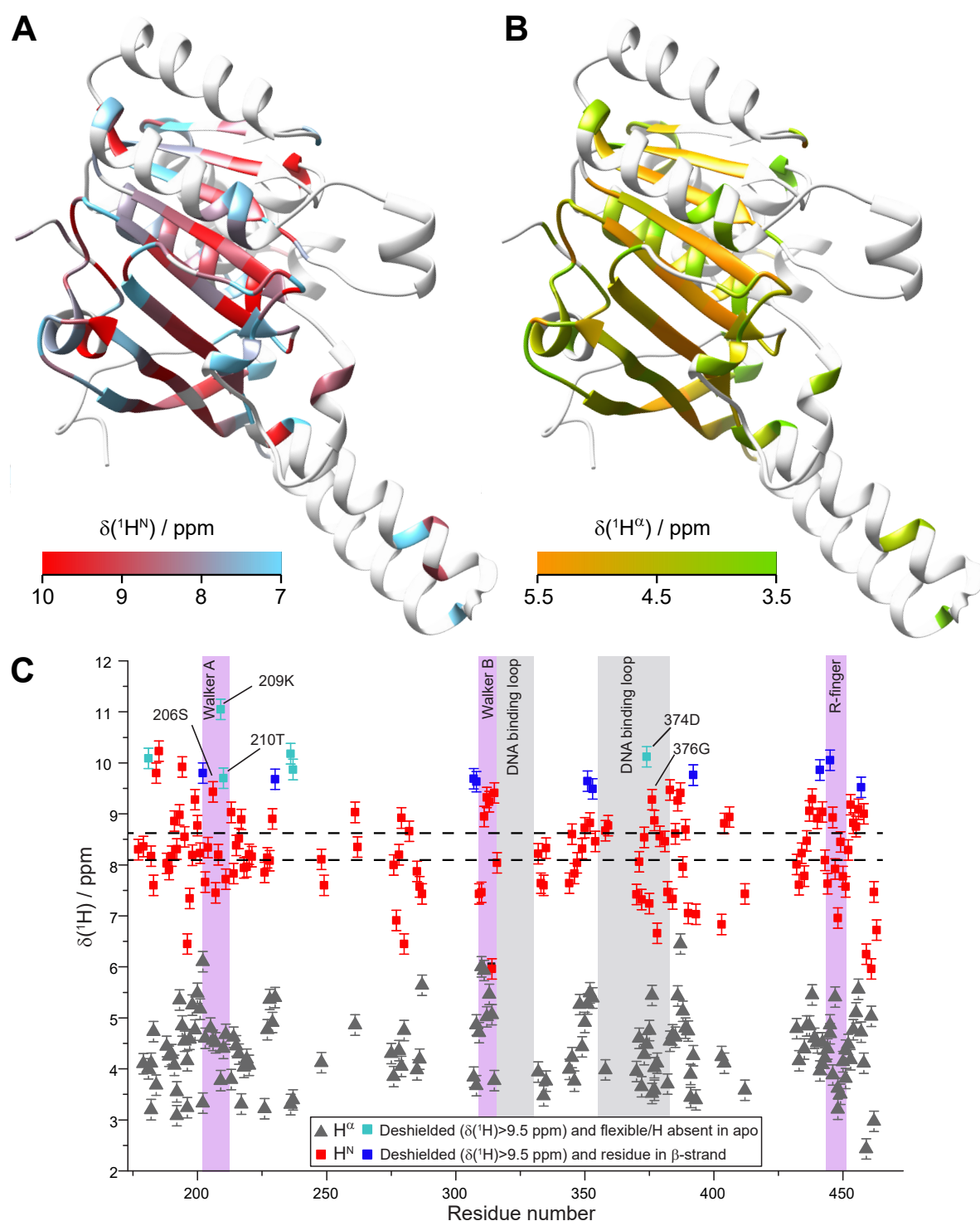
Figure 5.4. *Proton assignment with respect to the molecular structure.* **(a)** Assigned H^N resonances plotted on a DnaB model, based on the *Aa*DnaB:ADP complex (PDB ID 4NMN) using a colour gradient from 7 to 10 ppm, as shown in the legend. **(b)** Assigned H$^\alpha$ resonances, plotted on the same model, using a colour gradient from 3.5 to 5.5 ppm. **(c)** Site-specific H^N and H$^\alpha$ chemical shift values. The dashed horizontal lines represent the average H^N shifts in α-helices (8.1 ppm) and β-strands (8.6 ppm). The error bars are estimated to 0.1 ppm.
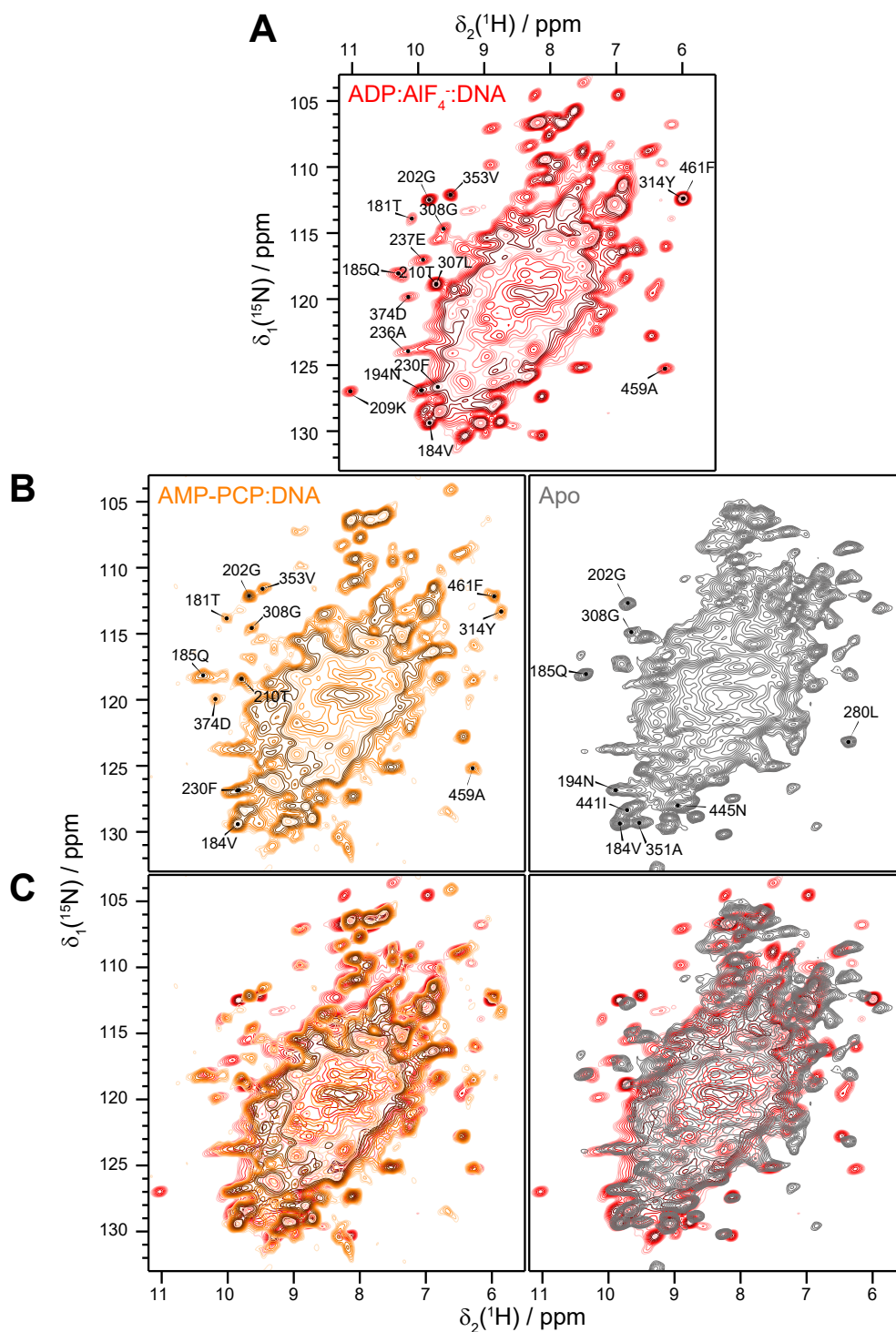
Figure 5.5. *Deshielded and shielded H^N resonances differ between the apo and the DNA-bound state.*
The DnaB:ADP:AlF$_4^-$:DNA complex **(a)** is compared to the pre-hydrolytic state mimicking DnaB:AMP-PCP:DNA complex **(b, left)** and the apo state of the protein **(b, right)**. For easier comparison, **(c)** shows the spectra from (b), overlaid with the one from (a). Characteristic deshielded and shielded isolated peaks discussed in Section 5.2.3 are marked with their assignment.

5.2.2. Proton linewidths of DnaB

The ease of ^{1}H resonance assignment strongly depends on the observed total proton linewidths (Δ^{tot}) that are a sum of homogeneous (Δ^{homo}) and inhomogeneous (Δ^{inhomo}) contributions (see Section 1.3.3, p. 30 and Appendix Figures D.2-D.4). Figure 5.3B shows the contributions to the ^{1}H linewidths determined for isolated peaks in the 2D hNH spectrum (see Appendix Figure D.3). The mean total linewidth of $\Delta^{\text{tot}} = 181 \pm 10$ Hz contains a homogeneous broadening of $\Delta^{\text{homo}} = 125 \pm 6$ Hz on average, which is comparable to other protonated systems [28, 156]. The average inhomogeneous contribution is $\Delta^{\text{inhomo}} = 61 \pm 11$ Hz. From these values, it is clear that ^{1}H linewidths of DnaB are dominated by homogeneous broadening effects. This topic is further explored in Chapter 6.

5.2.3. Probing protein-nucleotide interactions

In order to complement the information from the assigned chemical shifts (Section 5.2.1), a number of *Hp*DnaB complexes were investigated (Figure 5.5). These served to compare the DNA unbound ('apo') and bound states, where the latter was observed in the ATP prehydrolytic state (using AMP-PCP as an ATP analogue together with ssDNA), and the ATP hydrolysis transition state (using ADP:AlF$_4^-$ and ssDNA). Polarisation transfer was measured in correlation spectra for a range of resonances, including backbone ^{15}N, sidechain ^{15}N, ^{31}P in both nucleotides, ^{1}H , and finally aliphatic and carbonyl ^{13}C , the latter in particular for lysine and arginine sidechains (Figure 5.6).

For example, Figure 5.6B illustrates direct polarisation transfer between spins in spatial proximity (< 8-9 Å) for a ^{31}P-detected 2D NHHP correlation experiment (for a CHHP spectrum see Appendix Figure D.6). In contrast to the previously described ^{31}P,^{15}N TEDOR experiment used for identifying protein-RNA contacts in a smaller test system [185], the polarisation in NHHP experiments is mediated between close-by ^{15}N and ^{31}P nuclei via H-H spin diffusion, which allows to extract medium-to-long-range structural restraints [211] and is thus an alternative to TEDOR experiments. Sensitivity in such experiments is becoming an issue if the ^{31}P spin concentration is small as in the investigated system. The spectra reveal correlations between phosphate groups and backbone amides, as well as arginine and lysine sidechain nitrogen atoms. The spectrum clearly distinguishes different ^{31}P shifts for ADP and DNA [209]. As a matter of fact, assignments in the 2D spectrum remain ambiguous for such a large protein, but could be resolved considering the primary amino-acid sequence and the motifs to which nucleotides were predicted to bind [177]
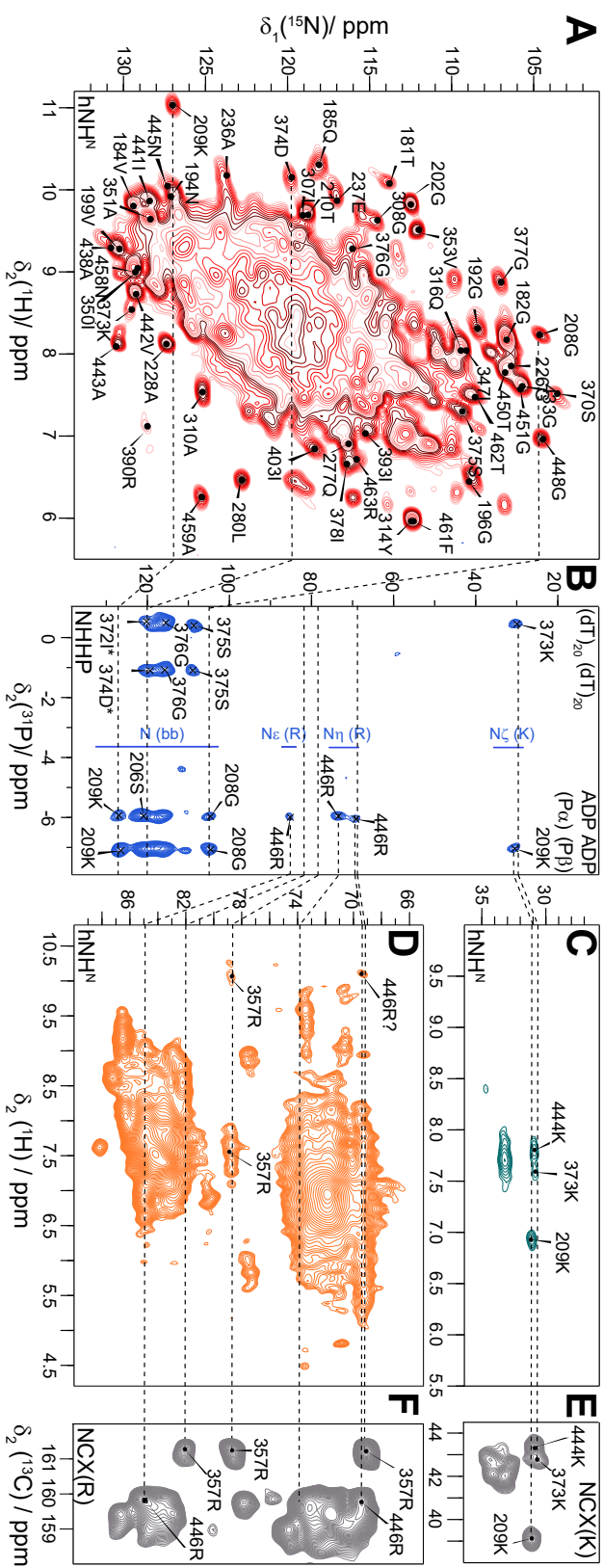
Figure 5.6. *Solid-state NMR spectra to probe protein-nucleotide interactions.* **(a)** 2D hNH spectrum with assignment of isolated resonances. **(b)** NHHP spectrum with assigned resonances. **(c)** Lysine sidechain 2D hNH spectrum with assignments. **(d)** Arginine sidechain 2D hNH spectrum with assignments. **(e)** Lysine sidechain 2D NC spectrum with assignments. **(f)** Arginine sidechain 2D NC spectrum with assignments. Subfigures b, e, and f are adapted from [209]. Dashed lines are guidance for the eye for the correlations discussed throughout Section 5.2.3.
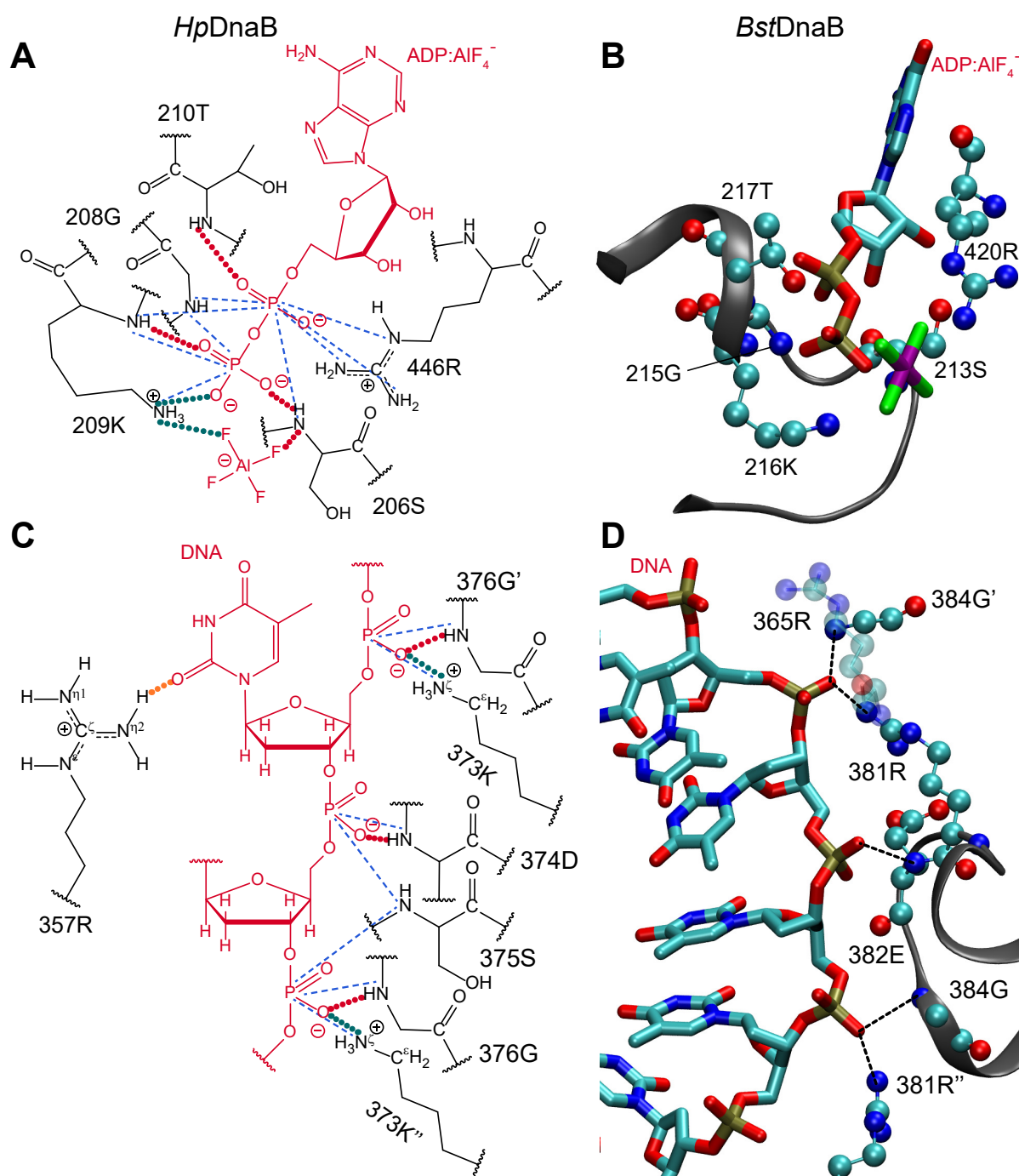
Figure 5.7. *Protein-nucleotide interactions in DnaB derived from solid-state NMR.* **(a)** Schematic drawing of *Hp*DnaB contacts to ADP:AlF$_4^-$ (in red) deduced from NMR data. Dashed lines represent spatial correlations observed in NMR spectra, coloured corresponding to the colour coding of spectra in Figure 5.6. Dotted lines represent hydrogen bonds derived from chemical shifts. **(b)** Protein-GDP:AlF$_4^-$ contacts in *Bst*DnaB as determined from the crystal structure (PDB ID 4ESV [184]). **(c)** Schematic drawing similar to (a), showing contacts between *Hp*DnaB and DNA (in red). ' and " indicate amino-acid residues from adjacent subunits. **(d)** Protein-DNA contacts in *Bst*DnaB as determined from the crystal structure (PDB ID 4ESV [184]).

(*i.e.* residues 203A-210T, 445N-451G and residues 371D-382A [212], respectively), and taking also further NMR spectroscopic information into account (*e.g.* CSPs and dynamic changes upon nucleotide binding, see reference [209]).

5.2.3.1. High-frequency H^N shifts in the Walker A motif

The residue 209K is located in the Walker A motif (Figure 5.2C). A clear 209K Nζ-ADP Pβ cross signal can be observed in the NHHP spectrum (Figure 5.6B), which favourably positions the sidechain to interact with the β-phosphate and/or with the AlF$_4^-$. And indeed, a sidechain correlation peak involving ^{15}N is observed for 209K in both the hNH and NC spectrum (Figure 5.6C,E); observation of such cross signals is often related to involvement of the NζH$_3^+$ group in salt bridge formation [213]. Interestingly, the 209K H^N chemical shift represents the most deshielded ^{1}H resonance (11.0 ppm) for the DnaB:ADP:AlF$_4^-$:DNA complex. Also, an (ambiguous) signal is observed at the 209K H^N shift in the NHHP spectrum. 209K H^N is thus with high confidence involved in a strong hydrogen bond which is established only upon nucleotide-binding, since the 209K HN resonance is not observed in the apo form. Such a conclusion can be drawn from the hNH spectra shown in Figure 5.5 which compare the DnaB:ADP:AlF$_4^-$:DNA state with the pre-hydrolytic DnaB:AMP-PCP:DNA and the apo state. This agrees with previous findings that 209K rigidifies upon ADP:AlF$_4^-$ binding and remains flexible in the apo form. Interestingly, this is the case in the DnaB:AMP-PCP:DNA-bound state as well, where one would also expect that 209K binds to the nucleotide (see Appendix Table D.2 and Figure 5.5 for the 2D hNH spectra) [209].

5.2.3.2. Lysine and arginine sidechains interact with ATP and DNA

The NHHP spectrum indicates spatial proximity between further 209K-neighbouring residues and the nucleotide, among which 210T is most likely involved in hydrogen bonding (H^N shift of 9.7 ppm). 206S is also relatively deshielded (9.4 ppm) for DnaB:ADP:AlF$_4^-$:DNA. The equivalent residue to 206S has been identified in other NTPases, whose structures locate this residue in the Walker A motif near the fluorine atom of the γ-phosphate mimic, identifying its key role in stabilising the γ-phosphate during ATP hydrolysis [214]. Besides 209K, correlation signals for Nη1/η2 and Nε sidechain resonances are observed in the NHHP/CHHP spectra for 446R, showing that it is in close spatial proximity to Pα of ADP (Figure 5.6B, and Appendix Figure D.6) . 446R is located in the R-finger which connects two adjacent monomers, and plays a central role in ATP binding and hydrolysis [178, 214–216]. While in

the hNH spectrum (Figure 5.6A), the 446R resonances cannot be unambiguously assigned, a deshielded ^{1}H resonance (10.1 ppm) is particularly observed for Nη1/η2 at the corresponding ^{15}N shift, which could support assignments to this residue based on the expected hydrogen-bond formation with the ADP phosphate group. To further analyse nucleotide binding, spectra were recorded on two additional protein samples: the apo protein (no nucleotides bound) and DnaB:AMP-PCP:DNA using a pre-hydrolytic ATP-mimic [209].

Further interactions can be identified in the arginine sidechain region of the hNH spectrum in Figure 5.6D. The isolated 357R ^{15}Nη2 chemical-shift (Figure 5.6F and Appendix Figure D.8) can be clearly identified. Two correlation peaks to ^{1}H are detected at this frequency, one around 10.2 ppm, and the other around 7.6 ppm (Appendix Figure D.9). While the latter is a typical amide proton chemical-shift value, the former is deshielded, indicating hydrogen-bonding interactions, but no correlations to the DNA phosphates are observed in NHHP (Figure 5.6B) and only very weak ones are seen in the CHHP spectra, in contrast to 446R of the R-finger (Appendix Figure D.6). Arginine residues can show three different sidechain-DNA binding modes [181, 183]: either to the DNA phosphate group; to the base edges or the ribose via hydrogen bonding or electrostatic interactions; or to the DNA base plane via electrostatic cation-π interactions [217–219]. The absence of NHHP correlations, combined with the deshielded proton shift, positions 357R in coordination with the DNA base edge. As only one H^N is deshielded, only one NηH$_2$ group is involved in a hydrogen bond to one of the thymidine oxygens (see Figure 5.7C for a schematic drawing). The relatively broad resonances (predominantly in the ^{1}H dimension) might be a consequence of structural disorder, *e.g.* a small structural inequivalence of the DnaB monomers in the oligomeric protein-DNA complex, leading to a chemical-shift distribution. These observations show that NMR allows us to distinguish between hydrogen bonds to the DNA phosphate groups and the DNA base edge. Arginine sidechain-thymidine base edge interactions are rather rarely observed and much less frequent than for example arginine-guanidine pairs [181, 183].

The DNA chemical-shift region of the NHHP spectrum in Figure 5.6B reveals two ^{31}P chemical shifts for two ^{31}P DNA phosphate resonances. The binding of two DNA nucleotides per DnaB monomer is a common feature of SF4 helicases [184, 209] and is reflected in solid-state NMR by two ^{31}P resonances with different ^{31}P chemical-shift values. The 150 ms ^{31}P-^{31}P 2D DARR spectrum in Figure 5.8 shows the phosphorus resonances for the ADP phosphate groups α and β as well as two DNA resonances. Strong cross-peaks between the two ADP resonances indicate the close spatial proximity of the two ADP phosphate groups, while

for the two DNA resonances such cross-peaks are less intense. Despite being less intense, the presence of these cross-peaks still indicates that the two ^{31}P DNA resonances can be assigned to two adjacent but structurally distinct phosphate groups of DNA.

However, returning to the top-left of Figure 5.6B, only one of the phosphorus DNA resonances (the one around 0.3 ppm) shows a correlation to the lysine sidechain of 373K (see Appendix Figure D.7 for the assignment of this residue). The Nζ-Hζ sidechain resonance of this residue is also observed in the hNH spectrum (7.6 ppm, Figure 5.6C) supporting its involvement in a salt bridge with the DNA phosphate. For apo DnaB, no lysine sidechain correlations are detected, in line with our previous findings [209]. Concomitant to this, the neighbouring residue, 374D shows a particularly deshielded resonance of 10.1 ppm, indicating that it is involved in a hydrogen bond as well, possibly with the DNA (see Figures 5.5 and 5.7 and Appendix Table D.2).

5.2.3.3. A comparison to *Bst*DnaB:GDP:AlF$_4^-$:DNA

The findings from the ^{31}P intermolecular correlation experiments and ^{1}H chemical shifts are summarised in Figure 5.7A, and are compared to the interactions available from the *Bst*DnaB:GDP:AlF$_4^-$:DNA crystal structure (Figure 5.7B). One can see that the contacts with the nucleotide, as defined by NMR, are similar in *Hp*DnaB to those revealed by the X-ray structure of *Bst*DnaB. For example, the close spatial proximity of the sidechain of *Hp*DnaB 209K (in the Walker A motif) to the nucleotide is equivalent for *Bst*DnaB 216K. Despite these similarities, NMR also reveals small differences in the binding mode. This can for instance be observed for the spatial proximity of the R-finger in DnaB:ADP:AlF$_4^-$:DNA (446R) exclusively to Pα as deduced from the NHHP and CHHP spectra (Figure 5.6B and Appendix Figure D.6), which distinguishes it from *Bst*DnaB:GDP:AlF$_4^-$:DNA, where 420R is close to both GDP phosphate groups (Figure 5.7B) [184]. Note that there is an important difference between the *Bst*DnaB structure and our NMR data for *Hp*DnaB regarding the occupancy of ATP-binding sites: while for *Bst*DnaB only five out of six nucleotide binding domains are occupied, full occupancy was derived from the NMR data for *Hp*DnaB [209], indicating further small structural differences between these proteins.

Figure 5.7C schematically summarises the obtained information on the DNA-binding mode of DnaB, and compares it to the DNA-binding modes from equivalent residues in the *Bst*DnaB:GDP:AlF$_4^-$:DNA structure [184]. 373K contacts the DNA phosphate backbone via a salt bridge. Among the nearby residues, 374D and 376G are potentially involved in binding

as well, based on their deshielded H^N resonances. 357R possibly contacts the DNA base edge via hydrogen bonding. Figure 5.7D shows that equivalent contacts can be seen in the crystal structure of *Bst*DnaB:GDP:AlF$_4^-$:DNA [184]), where the sidechain of 381R (corresponding to 373K in *Hp*DnaB) is in spatial proximity to the phosphate group of DNA, while the amide nitrogen atoms of 382E and 384G (374D and 376G, respectively, in *Hp*DnaB) may contact the phosphate group of DNA via hydrogen bonds. However, 365R in *Bst*DnaB (equivalent to *Hp*DnaB 357R) is close to the DNA phosphate groups in the crystal structure, indicating the existence of small differences in DNA coordination in *Bst*DnaB compared to *Hp*DnaB.
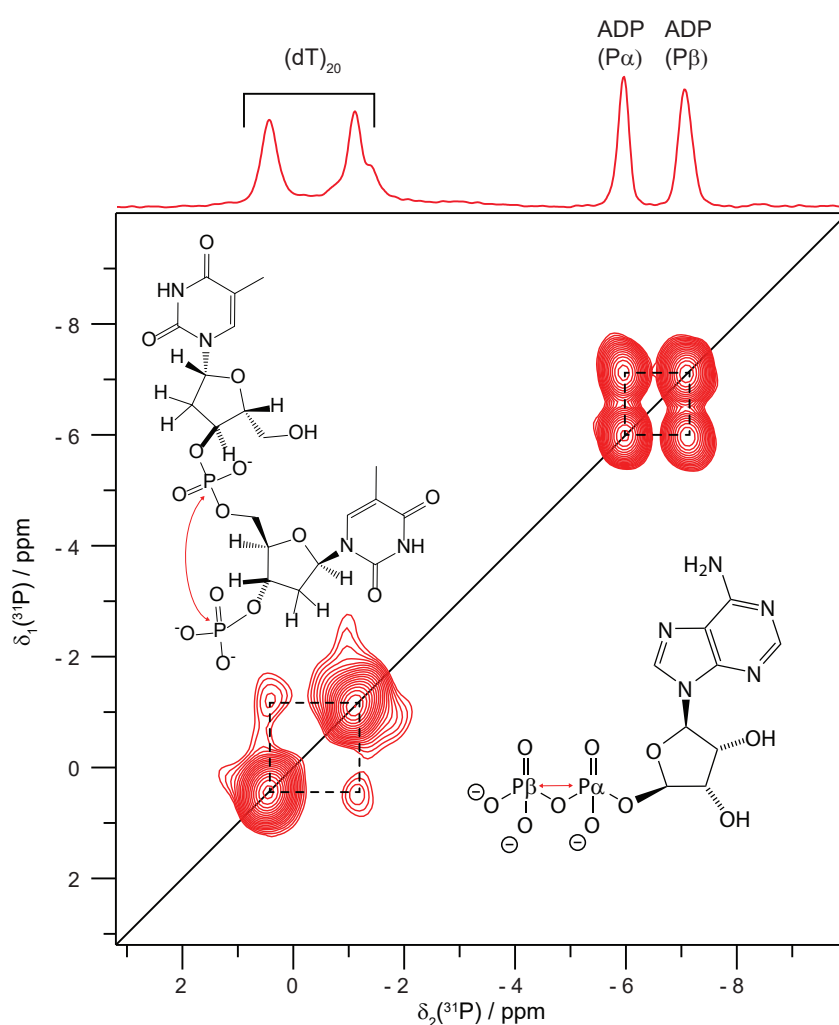

Figure 5.8. *Two DNA nucleotides bind per DnaB monomer and are in spatial proximity.* ^{31}P-^{31}P 150 ms DARR spectrum of DnaB:ADP:AlF$_4^-$:DNA, showing that the two bound DNA nucleotides are in close spatial proximity. On top of the 2D spectrum, a 1D ^{31}P CP spectrum is shown (taken from reference [209]).

5.3. Conclusion

The beginning of this chapter illustrated that a large part of the ^{1}H resonances in the do-decameric motor protein *Hp*DnaB can be assigned by making use of the well-dispersed H^N and H$^\alpha$ frequencies in a combined 3D assignment approach. The strategy exploits two intraresidual 3D correlations (hNCAH & hCANH), and two interresidual (relayed) 3D correlations (hNcoCAH & hCAcoNH). In addition, the chapter has briefly discussed ^{1}H linewidths, concluding that homogeneous broadening clearly is the dominant contribution to DnaB proton linewidths at this magic-angle spinning frequency. ^{1}H linewidths have also been discussed in the context of HBV Cp149, Cp183, and HBs S in Chapter 4, relating rather to differences between deuterated (and back-exchanged) samples compared to fully proto-nated ones. In this respect, it is interesting to mention that all the ^{1}H-detected work in this chapter has been performed on fully protonated samples. This demonstrates that the MAS frequencies used here (between 100 and 110 kHz) sufficiently reduce dipolar H-H broadening for a system of the size of DnaB. The topic will return for further discussion in the context of fast(er) magic-angle spinning in Chapter 6.

The main part of this chapter described in detail how nucleotide binding modes of DnaB with ATP analogues and DNA can be characterised to identify key residues and residue types, such as arginine and lysine, for protein-nucleotide interactions. This is achieved by the combination of direct 2D correlation spectra (including ^{31}P-detected ones) and ^{1}H as-signed spectra of various sample preparations. The findings compare well with data on a homologous protein, for which a crystal structure is available, which validates the presented approach. The procedure described here consequently allows to detect non-covalent inter-actions in molecular recognition processes involving nucleotides in further non-crystalline protein assemblies, be it in the context of nucleic acid synthesis, transcription, repair, or packaging (*e.g.* in viral capsids; see Chapter 4). With the presented methods, it is even possible to investigate changes in protein-DNA contacts *during* functional cycles, *e.g.* in DNA replication, revealing the intricacies of dynamic binding and release in such transitory states.

5.4. Materials and methods

5.4.1. Sample preparation

AMP-PCP and ADP was purchased from Sigma-Aldrich and $(dT)_{20}$ from Microsynth.

5.4.1.1. Expression and purification of *Hp*DnaB

The protein was cloned into the vector pACYC-duet1 (making use of the forward primer 5'-agtcatatggatcatttaaagcatttgcag-3' containing an *Nde*I restriction site and reverse primer 5'-atactcgagttcaagttgtaactatatcataatcc-3' containing an *Xho*I site) [177], and expressed in the *E. coli* strain BL21 Star (DE3) (One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*, Invitrogen™). The overexpression was performed in M9 minimal medium [220] using ^{13}C-enriched glucose 2 g·L^{-1} (Cambridge Isotope Laboratories, Inc. CLM-1396-PK) and ^{15}N-enriched ammonium chloride 2 g·L^{-1} (Sigma-Aldrich® 299251) as sole carbon and nitrogen sources. The cells were lysed by a microfluidisation process. ^{13}C-^{15}N labelled *Hp*DnaB was purified by heparin-agarose affinity chromatography using a 5 mL HiTrap Heparin HP column (GE Healthcare Life Sciences) followed by anion exchange chromatography using a 5 mL HiTrap Q HP column (GE Healthcare Life Sciences). The purified protein was concentrated up to 30 mg·mL^{-1} by centrifugation in buffer A (2.5 mM sodium phosphate, pH 7.5, 130 mM NaCl). For more details see reference [221].

5.4.1.2. Preparation of *Hp*DnaB:nucleotide:DNA complexes

The *Hp*DnaB:nucleotide complexes were prepared by incubating DnaB with 5 mM of nucleotide and 5 mM $MgCl_2·6H_2O$ for 2 h at 4 °C. In the case of $ADP:AlF_4^-$, an NH_4AlF_4 solution was prepared by incubating 1 M $AlCl_3$ solution with a 6-fold excess of 1M NH_4F solution (compared to $AlCl_3$) for 5 min in H_2O. 1 mM of $(dT)_{20}$ was added to the complexes and reacted for 30 min at RT. The protein solution was sedimented in the MAS-NMR rotor (16 h at 4 °C at 210'000 g).

5.4.2. Solid-state NMR experiments

^{1}H-detected solid-state NMR spectra were acquired at 20.0 T static magnetic field strength using a 0.7 mm Bruker probe operated at an MAS frequency of 105-110 kHz. ^{13}C-detected spectra were recorded at the same magnetic field, but in a 3.2 mm Bruker "Efree" probe [222].

The 2D and 3D spectra were processed with the software TOPSPIN (version 3.5, Bruker Biospin) with a shifted (2.5 to 3.5) squared cosine apodisation function and automated baseline correction in the indirect and direct dimensions. ^{1}H-detected spectra were cut in the F1-dimension after 1k points. ^{31}P-detected experiments were acquired at 11.74 T in a Bruker 3.2 mm probe (using home-built ^{31}P/^{13}C and ^{31}P/^{15}N inserts) at an MAS frequency of 17.0 kHz. ^{31}P-detected experiments were processed with an exponential line broadening of 10 Hz in the two dimensions. The sample temperature was set to 278 K [168]. All spectra were analysed with the software CcpNmr [136, 137, 223] and referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). All experimental details are provided in Appendix Table D.3.

6. Magic-angle spinning ever faster

This work has been in collaboration with the group of Ago Samoson in Tallinn, Estonia, responsible for building the 0.5 mm probe head. The chapter contains a small part of:

[224] Torosyan, A.#, Wiegand, T.#, Schledorn, M.#, Klose, D., Günthert, P., Böckmann, A., & Meier, B. H. (2019). Including protons in solid-state NMR resonance assignment and secondary structure analysis: The example of RNA polymerase II subunits Rpo4/7. *Frontiers in Molecular Biosciences*, 6, 100.

And is a preparation for the following manuscript:

Schledorn M.#, Malär, A.A.#, Torosyan, A.#, Oss, A., Org, M.L., Penzel, S., Klose, D., Wang, S., Lecoq, L., Cadalbert, R., Samoson, A., Böckmann, A., & Meier, B. H. *(To be submitted)* Protein NMR spectroscopy at 150 kHz magic-angle spinning continues to improve resolution and mass sensitivity.

# These authors contributed equally. For the work in the manuscript: Probe head design and engineering [AO, MLO, AS]; Protein expression and purification [AT, DK, SW, LL, RC]; NMR spectroscopy [MS, AAM]; Data analysis [MS, AAM]; Writing of the original manuscript [MS]; Discussion and editing of the final text [MS, AAM, AT, AB, BHM]; Study design and supervision [AB, AS, BM].

6.1. Introduction

The three-dimensional structure of a protein or protein complex is an essential first step to understanding its function, and, ultimately, to interfering with it, for example in the context of drug development. Solid-state NMR plays an important role in de novo 3D structure determination of particular protein-classes, *e.g.* amyloids [6, 8, 9, 11, 53]. The method allows to study conformational details with particularly high sensitivity [148, 225, 226], and can for example identify drug-binding sites or characterise dynamic features [176, 227–232].

Almost invariably, the information content of the NMR spectrum increases with improved spectral resolution. For biomolecular NMR in solids, such resolution requires magic-angle spinning (MAS) [44, 45]. Until very recently, most applications of solid-state NMR have been based on carbon-13 detection, because proton resonances were not resolved enough, which was principally caused by less-than-perfect averaging of ^{1}H dipole interactions by MAS. The development of faster magic-angle spinning has recently enabled ^{1}H-detected spectroscopy. Indeed, with MAS frequencies approaching and exceeding 100 kHz, the linewidth in deuterated proteins (fully back-protonated on exchangeable sites), as well as in fully protonated proteins, is reduced to a degree which makes proton-detected solid-state NMR an attractive and accessible alternative to carbon-13 detection [26, 27, 138, 139, 158, 170, 233–237].

Proton detection has the important advantage of reducing the required sample amount for NMR spectroscopy at comparable sensitivity by roughly a factor 100 [139, 152], comparing a 3.2 mm rotor (standard for ^{13}C-detected experiments on proteins) with 50 mg of sample to a 0.7 mm rotor with about 0.5 mg for ^{1}H-detection.

Increased spinning frequencies with current technology will require a decrease of the rotor diameter d (see *e.g.* the review in [238]). The detection with optimised smaller coils further increases the sensitivity per unit volume (the mass sensitivity) by a factor d approximately, further reducing sample needs and enabling structural and dynamic studies of proteins available in small quantities. However, the reduced sample amount leads a decrease of the total signal with roughly d^{-2}. To keep the measurement times for larger proteins reasonable, this signal loss needs to be partially countered by further intensity gains, namely through a combination of a decrease in linewidths and a gain in transfer efficiencies exploiting the longer T_2' and $T_1\rho$ relaxation times which are central in scalar- and dipolar-coupling based transfers in solid-state NMR pulse sequences. The proton linewidth can be described by three different contributions: $\Delta^{tot}(\nu_r) = \Delta^{coh}(\nu_r) + \Delta^{incoh}(\nu_r) + \Delta^{inhomo}$. The coherent

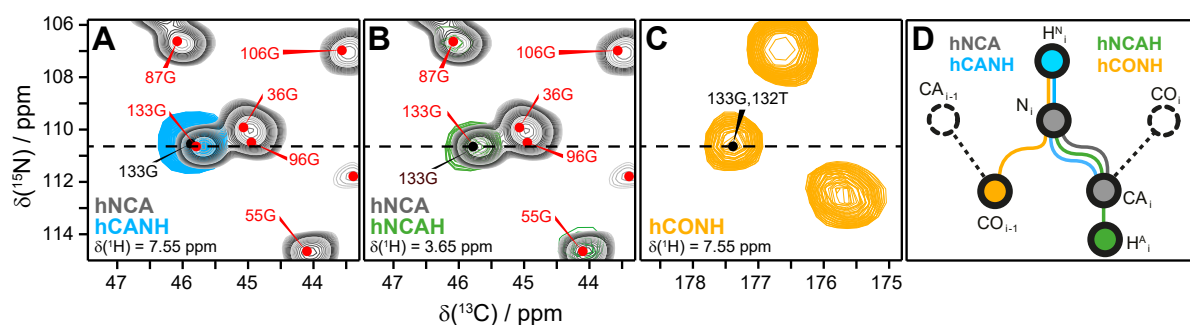**Figure 6.1.** *Assigment strategy for Rpo4/7 H^N and H^α resonances.* **(a)** ^{13}C-detected 2D hNCA spectrum (grey) and 2D plane of a ^{1}H-detected 3D hCANH spectrum (cyan) at $\delta_3(^1$H$) = 7.6$ ppm showing an example of the H^N resonance assignment of 133Gly, using known ^{13}C and ^{15}N chemical shifts. **(b)** The same 2D hNCA spectrum (grey) as in (a), with a 2D plane of a 3D hNCAH spectrum (blue) at $\delta_3(^1$H$) = 3.7$ ppm showing an example of the H^α resonance assignment of 133Gly. **(c)** 2D plane of an hCONH spectrum at $\delta_3(^1$H$) = 7.6$ ppm, to verify the H^N resonance assigned in (a), with the known C' resonance of the preceding spin system (CO$_{i-1}$). **(d)** Schematic representation of the resonance assignment of H^N and H^α atoms, as exemplified in (a-c).

contribution is often found to be dominant for highly ordered samples [28, 239]. Then the decrease in proton linewidth is, for the fast-spinning limit, linear with spinning frequency (and therefore with *d*) leading to a roughly linear improvement in signal amplitude in the spectrum. Improvements in polarisation-transfer efficiency have an even larger potential to enhance the sensitivity at faster MAS. It has been demonstrated that the polarisation-transfer efficiency at 90 kHz MAS is between 0.22 and 0.9 for each transfer step, leading, for more complex experiments, to low overall efficiencies below 0.10 (*e.g.* 0.03 for an hcocaCBcacoNH 3D experiment [27]) Therefore, there is considerable potential for improvement and higher-dimensional experiments can entail a correspondingly larger benefit.

This chapter presents data of a protein complex of 34 kDa recorded at 150 kHz MAS. This represents another step forward in MAS frequency for proteins (from 126 kHz [28]; for a peptide 140 kHz was reported [240]), and expands the scope from model proteins and peptides to the protein complex Rpo4/Rpo7 with 294 amino-acid residues, of which the 187 amino-acid residues of Rpo7 are uniformly ^{13}C-^{15}N-labelled. The data demonstrate that the expected improvement in spectral resolution when going from 100 kHz to 150 kHz can be realised and that the overall sensitivity is comparable to standard experiments at 110 kHz.
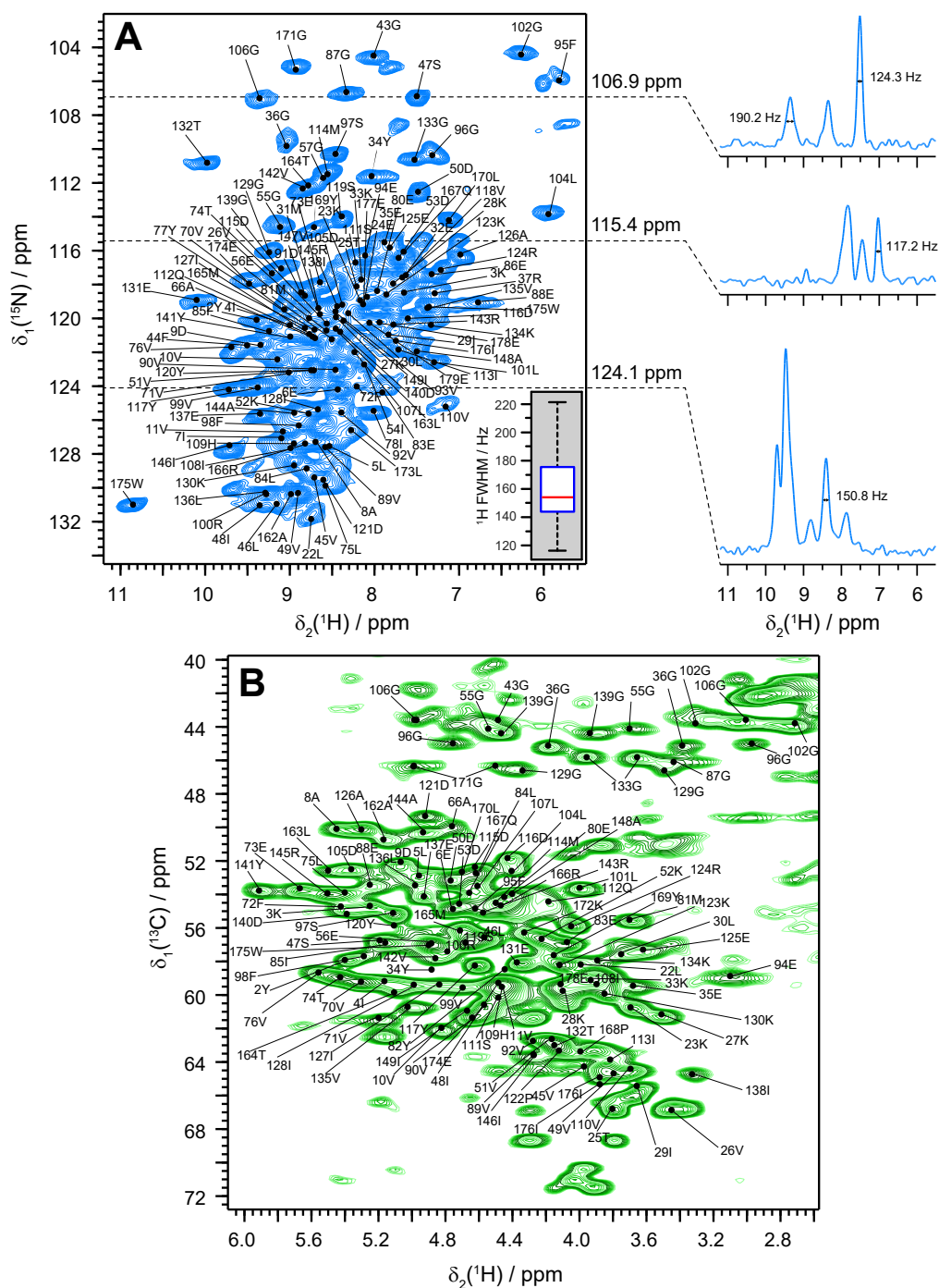
Figure 6.2. **(a)** 2D hNH correlation spectrum of fully protonated Rpo4/7 at 110 kHz MAS. The spectrum includes labels for the ^{15}N-^{1}H peaks as predicted from the manually created shift list. On the right side of the figure, 1D traces for the ^{1}H dimension are presented at the corresponding ^{15}N frequencies. The ^{1}H linewidth characteristics of the full population of marked cross-peaks are summarised in the boxplot in the bottom right, indicating the maximum, 3rd quartile, mean, 1st quartile and minimum value of ^{1}H FWHM linewidth in Hz with a mean value of 157 ± 5 Hz. **(F)** 2D hCH correlation spectrum of fully protonated Rpo4/7 at 110 kHz MAS with peaks labelled as in (a).

6.2. **Results and discussion**

To assign the amide H^N and aliphatic H^α protons of fully protonated UL Rpo7 in complex with Rpo4 (at NA), we used proton-detected spectroscopy at 110 kHz MAS frequency. The assignment strategy for these protons is illustrated in Figure 6.1. The assignment was done using three 3D spectra, namely hCANH, hNCAH and hCONH [27, 170], and taking advantage of the ^{13}C and ^{15}N peak assignment described in [224]. Details of the experiments are given in Appendix Table E.1. The ^{13}C and ^{15}N assignment, as illustrated in small parts of a ^{13}C-detected NCA spectrum, was used to assign H^N resonances in an hCANH spectrum (Figure 6.1A) and to assign H^α resonances in an hNCAH spectrum (Figure 6.1B). To confirm the assignment of amide protons, an additional hCONH spectrum was used to verify the CO chemical shift of the previous spin system (Figure 6.1C). hNH and hCH fingerprint spectra are shown in Figure 6.2, with the assigned resonances back-predicted onto the 2D spectra. In total, 97 % of the amide protons and 93 % of the H^α protons for which $C\alpha$ and N assignments exist could be assigned. The mean value and standard deviation of the ^1H linewidths of the fully protonated sample in the hNH spectrum are 157 $\pm$ 5 Hz for the marked peaks in Figure 6.2A. On the right side of the spectrum, 1D traces of the ^1H dimension are shown at the corresponding ^{15}N frequencies with linewidths of selected peaks.

6.2.1. **Solid-state 2D and 3D NMR spectroscopy at 150 kHz MAS**

Using a 0.5 mm probe triple-resonance MAS probe, we recorded a set of proton-detected spectra (hNH, hCH, hCANH) from the protein complex of subunits 4 and 7 of the archaeal RNA polymerase, in which only subunit 7 is uniformly $^{13}C/^{15}N$-labelled. The sample preparation and resonance assignment is reported in [224]. The hNH 2D spectrum, given in Figure 6.3A, predominantly shows backbone-nitrogen to amide-proton (H^N) correlations, that are then additionally correlated to the backbone $C\alpha$ resonances in a set of planes from an hCANH 3D experiment in Figure 6.3B. Resolved Hα-Cβ and Hβ-Cα cross-peaks are visible in the hCH Figure 6.3C, besides further aliphatic side-chain cross-peaks. The spectra were recorded within 12, 14 and 45 hours for the three experiments (Table 6.2), comparable to the 17, 20, and 83 h used for the same experiments at 110 kHz MAS in a Bruker 0.7 mm probe (Table E.1). We therefore conclude, that the total sensitivity is roughly the same despite the reduction of sample volume by a factor of 2 (0.59 μL *vs.* 0.31 μL) (*vide infra*).
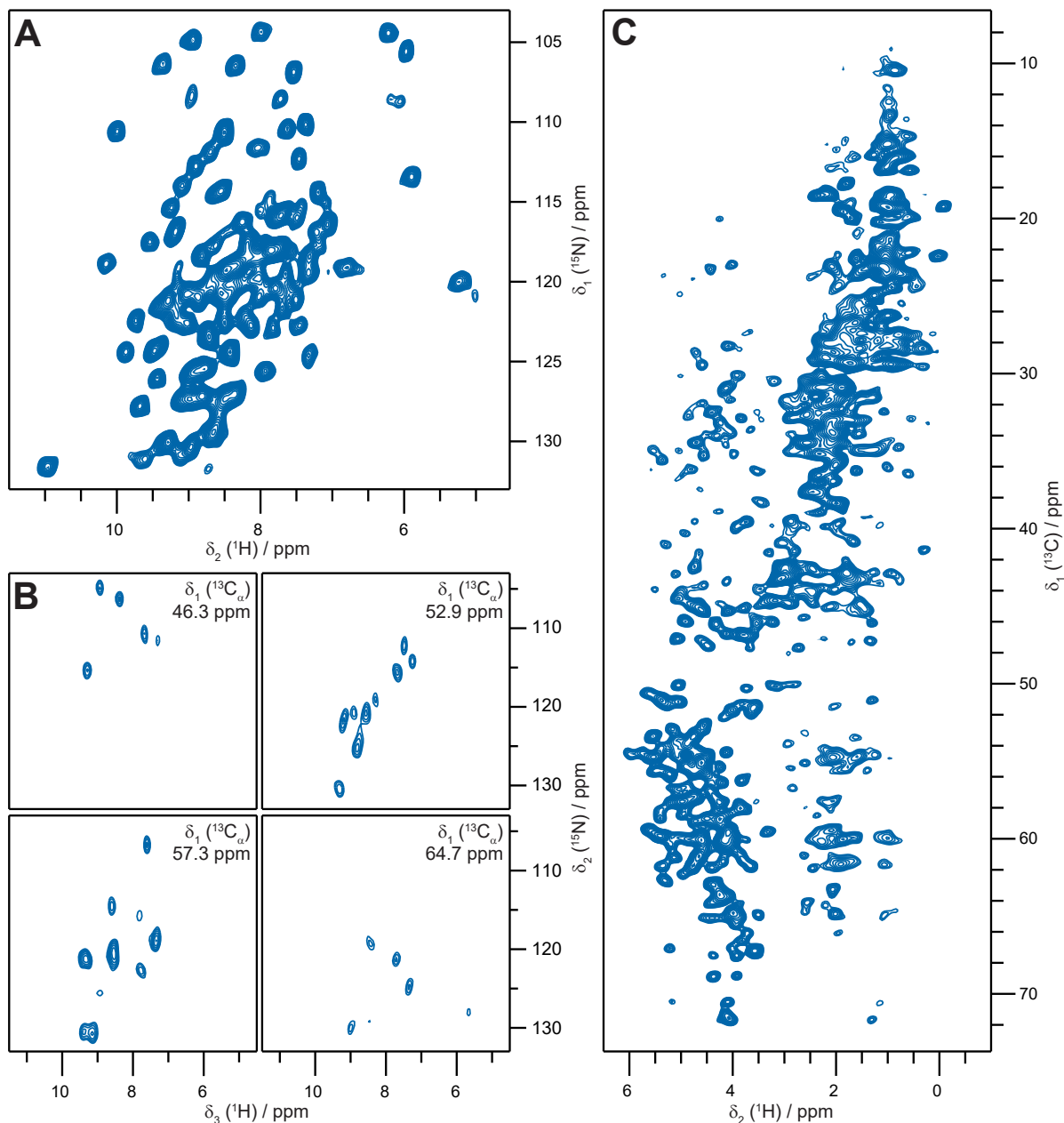
Figure 6.3. *2D and 3D heteronuclear correlation spectra of the Rpo4/7 complex recorded at 150 kHz MAS.*
The blue contours in **(a)** represent an hNH correlation spectrum processed with a QSINE 4 apodisation
function and cut off just above the noise level (exp. time 20 h). Right below, **(b)** shows representative
planes of an hCANH (*cf.* Appendix Figure E.1 for further details) (exp. time 45 h). While Cα-Hα cross-
peaks are resolved in the ^{15}N-dimension in this 3D, the hCH 2D correlation spectrum (exp. time 14 h) in
(c) illustrates further cross-peaks to Hβ, Cβ and other aliphatic carbons (aromatic region of the spectrum
not shown).

An automatic peak-pick by CcpNmr Analysis [136, 137] for the hNH spectrum of Figure 6.3A finds 83 resolved peaks (limited by signal overlap), while the same procedure resolves 101 peaks in the hCANH. A handful of peaks appear in the 2D while they are missing in the 3D (see Appendix Figure E.1A), some of which are side-chain cross-peaks that must be absent in the 3D. We ascribe other missing peaks to losses during the polarisation-transfer periods caused by dynamical processes. Notably, all except one of these missing peaks come from amino-acid residues located in turns or on the very edge of a β-strand or α-helix secondary structure element of the pr otein, as depicted in Appendix Figure E.1B.

   The spectra in Figure 6.3 form a fingerprint of the protein, as they display all its backbone ^{1}H, ^{13}C and ^{15}N resonances. The spectra illustrate the improved overall resolution obtained at 150 kHz MAS. The observed difference in resolution (some spectral regions are more overlapped than others) is influenced by the chemical-shift dispersion as well as by residue-specific differences in the linewidths, as we will demonstrate further below. Experimental parameters can be found in the Materials and Methods section, Table 6.2.

6.2.2. Linewidth and sensitivity at faster MAS

The spinning-frequency dependence of the hNH spectrum is illustrated by comparing spectra recorded at 100 and 150 kHz in Figure 6.4. The spectra were obtained in a back-to-back measurement with identical acquisition and processing parameters, except for the spin-lock fields matching the two different MAS frequencies. From the stack plots in Figure 6.4A and B it becomes immediately appreciable that the total linewidth reduces while peak amplitudes concomitantly increase at faster spinning. Contour plots of the 2D spectra (Figure 6.4C), with contours in red for 100 kHz and in blue for 150 kHz MAS, allow for a residue-specific analysis of this improvement, using the chemical-shift assignment as deposited in BMRB entry 27959 [224]. With respect to sensitivity, for the peaks for which assignments are indicated in Figure 6.4C, we observe an average increase in signal intensity by a factor 1.48 ± 0.13 when MAS increases from 100 to 150 kHz (Figure 6.5). This improvement can be due to both line-narrowing and higher transfer efficiencies. With respect to resolution, for the peaks labelled in the traces of Figure 6.4D, we observe a mean linewidth decrease by 36 ± 7 Hz (from 193 ± 6 to 157 ± 3 Hz), which corresponds to a reduction in linewidth by a factor of 1.23 ± 0.05 indicating that much of the gain is obtained from the reduction of the linewidth.
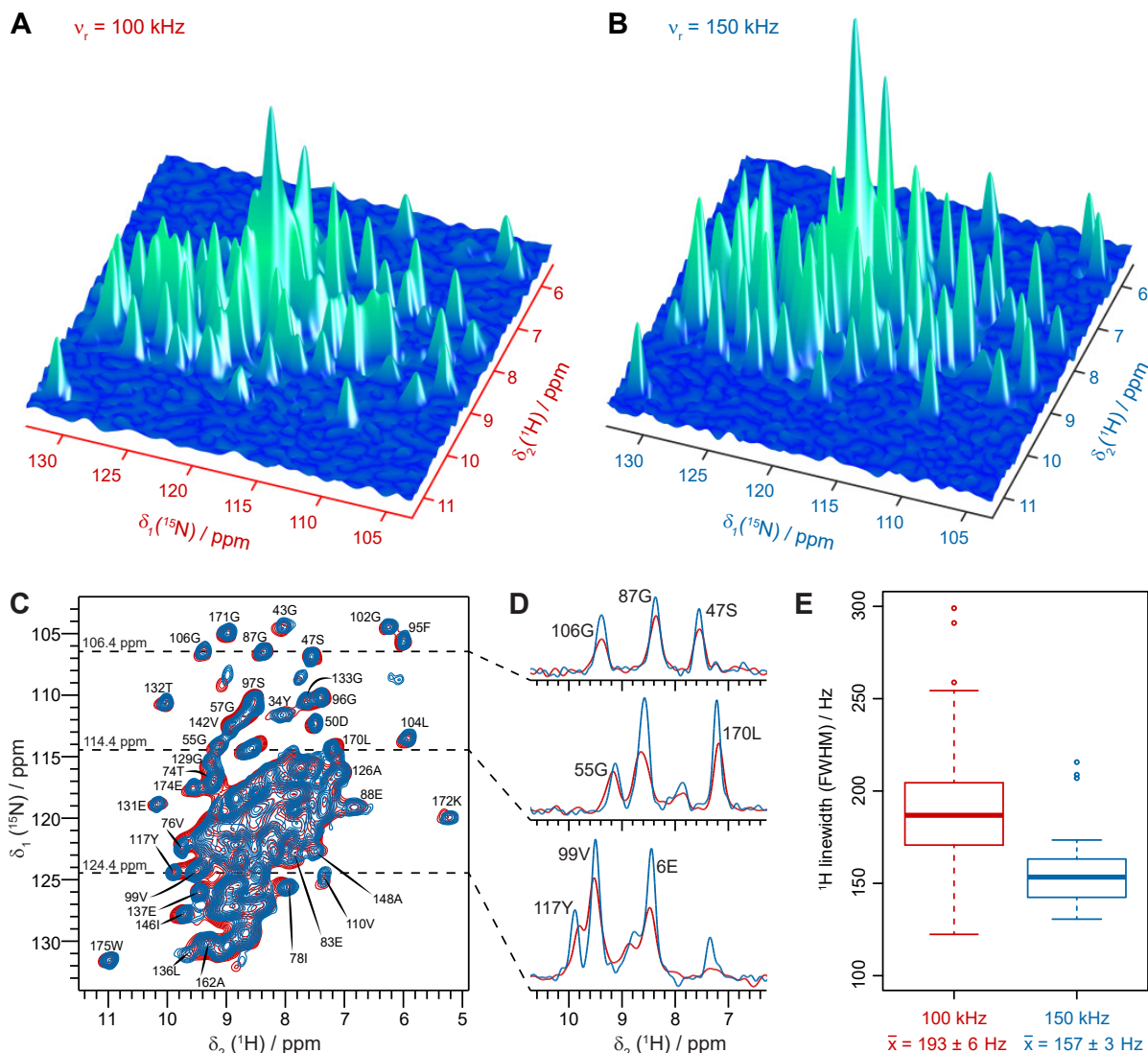
Figure 6.4. Comparison of resolution and signal intensity for an hNH correlation spectrum at 100 and 150 kHz MAS. Panels **(a)** and **(b)** show 3D stack plots of a 2D correlation experiments. An experiment at 100 kHz MAS **(a)** is contrasted to a measurement with identical acquisition and processing parameters at 150 kHz **(b)**. Both spectra are plotted at the same absolute intensity scale. The two spectra are overlaid in panel **(c)**, where some of the peaks are shown as a trace in panel **(d)**, with 100 kHz in red and 150 kHz in black. The corresponding linewidths of the peaks in panel D are summarised as boxplots in panel **(e)** to a mean value of 193 ± 6 and 157 ± 3 Hz respectively for 100 and 150 kHz MAS. The peaks that are indicated with an assignment label in panel C were used for site-specific relaxation analysis (Figure 6.8).
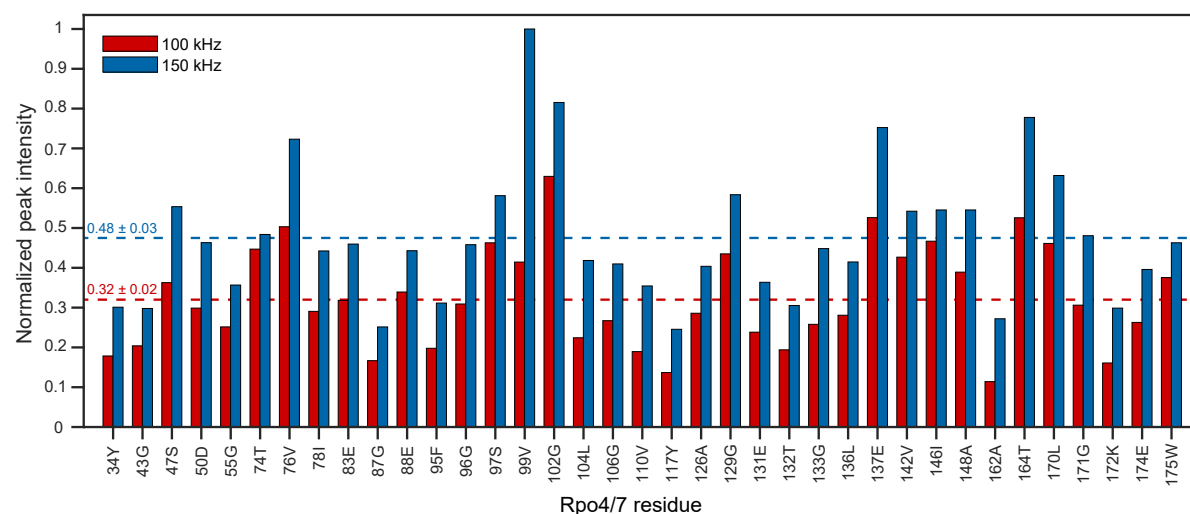
Figure 6.5. *Site-specific signal intensity for Rpo4/7 in a 0.5 mm rotor at 100 and 150 kHz MAS.* All peak intensities are normalised to the most intensive peak (99V at 150 kHz), with mean normalised peak intensities indicated as dashed lines at 0.32 ± 0.02 (100 kHz) and 0.48 ± 0.03 (150 kHz), corresponding to an intensity gain of a factor 1.48 ± 0.13 for 150 compared to 100 kHz MAS.

The 100 kHz experiments described above could have been done in a 0.7 mm probe instead of 0.5 mm increasing the sample volume by a factor of 2. Therefore, we experimentally compared full 0.7 and 0.5 mm rotors in their respective probes at 100 and 150 kHz MAS respectively. We observe the SNR in the 0.5 mm rotor at 150 kHz to be reduced by a factor 0.85, compared to a 0.7 mm rotor at 100 kHz, for a well-resolved isolated peak in a 2D hNH spectrum (see Appendix Table E.2). This implies that the largest part of the expected signal-to-noise loss due to the reduced sample amount in a 0.5 mm rotor, is counterbalanced by a gain due to the smaller coil, reduced linewidths, and improved polarisation-transfer efficiency, although the latter is not expected to be predominant for the hNH experiment.

In addition, the information content of the spectrum increases with peaks that only become visible at faster MAS where they sharpen up. Examples of this effect are observed in Figure 6.4C (*e.g.* at δ_1 = 109 and δ_2 = 6 ppm) and for meldonium (*vide infra*) (*e.g.* at δ_1 = 2.5 ppm).

6.2.3. Quantifying bulk coherence lifetimes at increasing MAS frequencies

Besides a direct comparison of the Rpo 4/7 spectra between 100 and 150 kHz MAS, we investigated bulk R_2' relaxation-rate constants as a function of MAS frequency in four different fully-protonated proteins. We measured homogeneous linewidths $\Delta^{\mathrm{homo}}(\nu_r)$ by Hahn-echo experiments. The bulk values for $\Delta^{\mathrm{homo}}(\nu_r)$ in Rpo4/7 are given in Figure 6.6A in a MAS-frequency range from 90-150 kHz, according to $\Delta^{\mathrm{homo}}(\nu_r) = c^{(1)} + c^{(2)}/\nu_r + c^{(3)}/\nu_r^2$. As expected for the fast spinning regime, the linewidth decreases linearly with the inverse of the MAS frequency $\Delta^{\mathrm{homo}}(\nu_r) = c^{(2)}/\nu_r$, with $c^{(2)} = 14 \pm 1$ kHz^2 for Rpo4/7 (the error is given as two standard deviations, to roughly correspond to the 95 % confidence intervals drawn in Figure 6.6). The value of this linear decrease seems to be rather universal for fully protonated proteins (Table 6.1). For HETs (218-289), a fungal prion in its β-sheet rich amyloid form [6, 7], we obtain $c^{(2)} = 15 \pm 1$ kHz^2 (Figure 6.6B); for Cp149 capsids of the Hepatitis B Virus (see Chapter 4 and Figure 6.6C) we find $c^{(2)} = 13 \pm 1$ kHz^2 ; and for ubiquitin [25] (Figure 6.6D) we find $c^{(2)} = 14 \pm 1$ kHz^2.

For comparison to the above described protein systems, we investigated the organic drug molecule meldonium (Figure 6.6E-F) (spectra and assignment are given in Figure 6.7A). In contrast to the investigated protein samples, meldonium shows a very different correlation between homogeneous linewidth and MAS frequency, with linear fit components of $c^{(2)}_{\mathrm{NH}} = 13 \pm 5$ kHz^2, $c^{(2)}_{\mathrm{CH3}} = 12 \pm 1$ kHz^2 and $c^{(2)}_{\mathrm{CH2}} = 47 \pm 4$ kHz^2, as well as quadratic contributions of $c^{(3)}_{\mathrm{NH}} = 4.0 \pm 0.6$ kHz^3, $c^{(3)}_{\mathrm{CH3}} = 1.1 \pm 0.2$ kHz^3 and $c^{(3)}_{\mathrm{CH2}} = 0.7 \pm 0.5$ kHz^3. This indicates significantly stronger proton-proton couplings. Although the spectral lines are thus substantially broader than the protein resonances at all spinning frequencies, their width also decreases considerably with a higher spinning frequency (Figure 6.7A).

The homogeneous linewidth at 150 kHz for the three assigned proton resonances in the meldonium spectrum is 340 ± 20 Hz for the CH$_2$ group; 121 ± 5 Hz for the CH$_3$; and 250 ± 20 Hz for the NH group (Figure 6.7B). In comparison, the fitted total FWHM (Figure 6.7C) for each of these components respectively, is 660, 270, and 470 Hz indicating that also in this case, the homogeneous broadening still forms a substantial contribution to the linewidth. We note that in particular the CH$_2$ group is very broad and of relatively low amplitude compared to the 9:4 abundance of methyl to methylene protons, even at 150 kHz MAS. To verify that these broader lines are indeed caused by the dipolar-coupling network we predicted coherent line widths from the crystal structure and the observed chemical
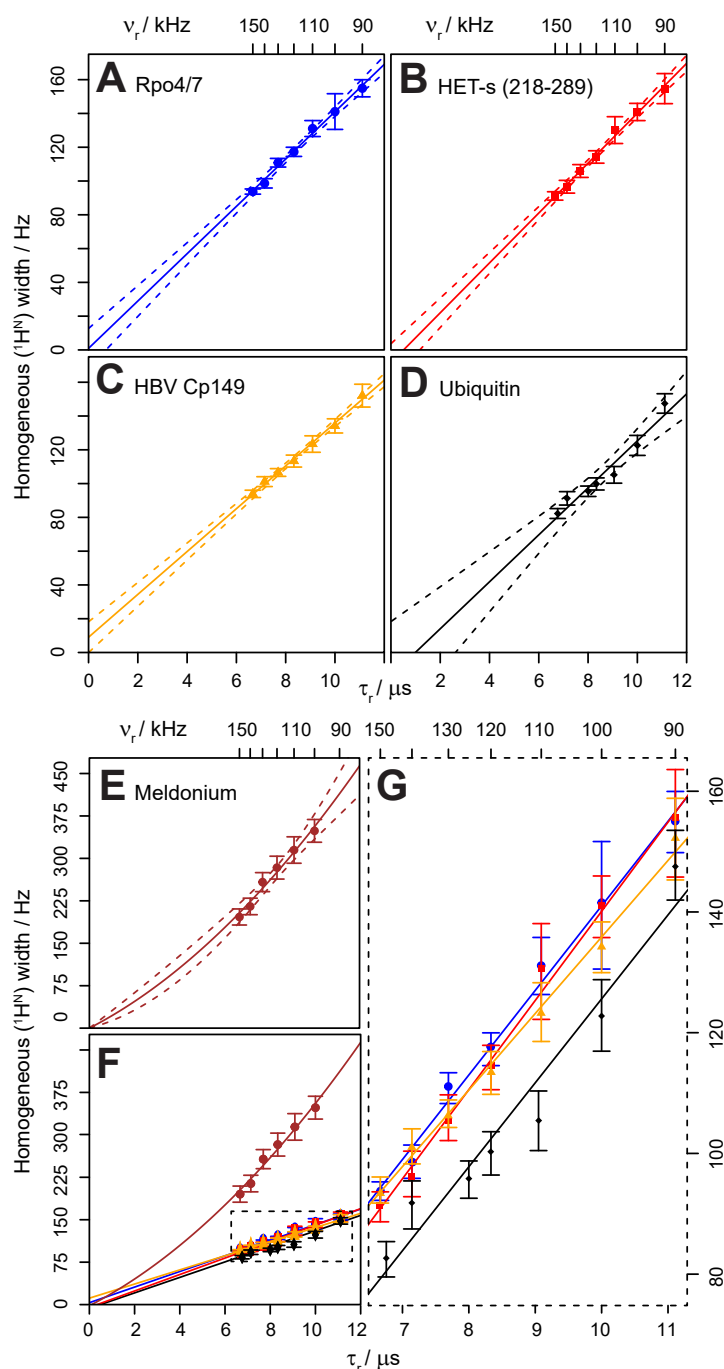
Figure 6.6. *Bulk homogeneous linewidths in a MAS range of 90 to 150 kHz are highly comparable for various protein samples but not for meldonium.* The fully protonated ^{13}C-^{15}N labelled samples of Rpo4/7 (panel **(a)**) (as presented in the main text), HETs (218-289) (panel **(b)**), and *human* HBV Cp149 (panel **(c)**) all can be fit linearly at a slope of 13-15 kHz2 while intercepting the origin within a confidence interval of 90 % indicated by dashed lines above and below each fit. Ubiquitin (panel **(d)**) shows a similar trend, although a small quadratic contribution is observed. In contrast, the homogeneous width for meldonium (panel **(e)**) decreases much faster for shorter rotor periods and a significant quadratic component is observed. At the same time, the general homogeneous broadening is much stronger than for protein samples, which can be explained by proton-proton coupling in the molecular structure of the small compound (see also Figure 6.7). Panel **(f)** shows all samples in relation to each other, with a blow-up in panel **(g)** to distinguish the various protein samples that are very similar in slope and intercept.
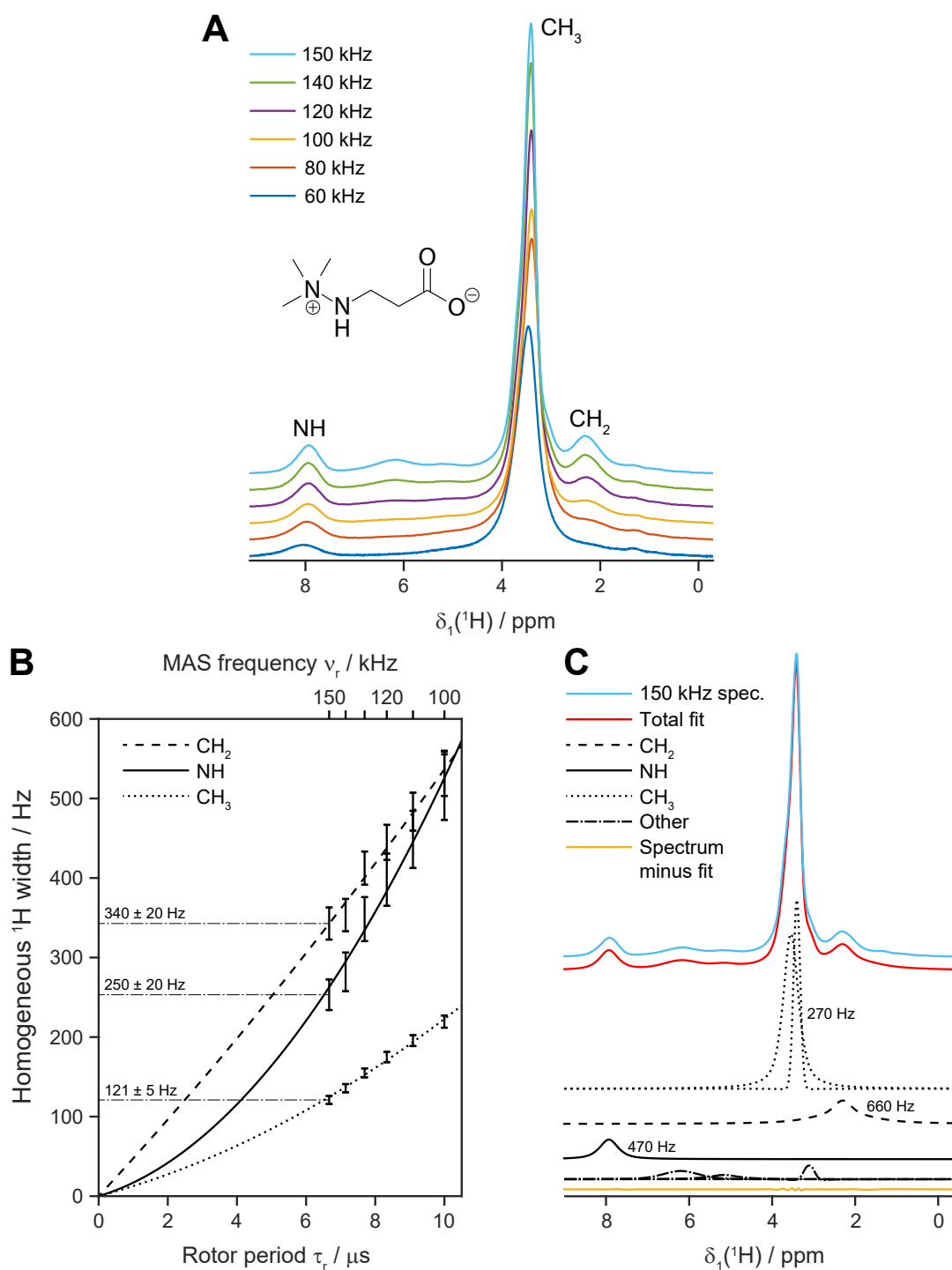
Figure 6.7. *Resolution improvement in meldonium at 60-150 kHz MAS.* **(a)** 1D proton spectra at 60 up to 150 kHz MAS, with assignment of the NH and triple CH3 groups, as well as one of the CH2 groups shown in the molecular structure of meldonium. The second CH2 group is suspected to be hidden around 3 ppm, under the right shoulder of the methyl peak. **(b)** Experimental homogeneous linewidths for the three assigned groups, with widths at 150 kHz (rotor period of 6.7 μs) of 690, 510 and 240 Hz, for CH2, NH, and CH3 respectively. **(c)** FWHM fit of the three assigned groups in the 150 kHz 1D spectrum, performed with DMFIT [241].

	$\Delta^{\mathrm{homo}}(\nu_{\mathrm{r}}) \; = \; c^{(1)} + c^{(2)}/\nu_{\mathrm{r}} + c^{(3)}/\nu_{\mathrm{r}}^2$		
	$c^{(1)}$ / Hz	$c^{(2)}$ / kHz^2	$c^{(3)}$ / kHz^3
Rpo4/7	1 ± 10	14 ± 1	0
HET-s (218-289)	-8 ± 9	15 ± 1	0
HBV Cp149	9 ± 10	13 ± 1	0
Ubiquitin	-13 ± 25	14 ± 3	0
Rpo4/7	0	14 ± 1	0.0 ± 0.1
HET-s (218-289)	0	13 ± 1	0.1 ± 0.1
HBV Cp149	0	15 ± 1	-0.1 ± 0.1
Ubiquitin	0	10 ± 1	0.2 ± 0.1
Meldonium NH	0	13 ± 5	4.0 ± 0.6
Meldonium CH$_3$	0	12 ± 1	1.1 ± 0.2
Meldonium CH$_2$	0	47 ± 4	0.7 ± 0.5

Table 6.1. *Parameters describing the spinning-frequency dependence of the homogeneous linewidth in a range of MAS frequencies from 90 to 150 kHz (Figure 6.6).* Values at zero (without error bars) were fixed in the respective evaluation. Errors are given as 2σ, roughly equivalent to the 95 % confidence intervals drawn as dashed lines in Figure 6.6.

shifts using a second-moment approach [239]. We obtain of 228 Hz for the CH$_3$, 398 Hz for the NH and 415 Hz for the CH$_2$ groups, which shows again that homogeneous broadening is the dominant contribution. The meldonium spectrum, expected to be representative for many drugs, thus strongly benefits from MAS frequencies beyond those presently realised.

6.2.4. Site-specific homogeneous linewidths improve for all residues, with variable inhomogeneous contributions

Site-specific homogeneous linewidths ($\Delta^{\mathrm{homo}} = R_2'/\pi$) were extracted from baseline-separated peaks of Rpo4/7 at MAS frequencies of 90, 120 and 150 kHz (Figure 6.8A). Consistently, the R_2' rate constants and consequently the homogeneous linewidths decrease with increasing spinning frequency as expected from the bulk behaviour (Figure 6.6A). Δ^{homo} for the three data sets at 90, 120, and 150 kHz, reaches a median value of 140 ± 10, 116 ± 5, and 90 ± 3 Hz respectively. Interpolating these values for a mean homogeneous linewidth at 100 kHz MAS gives 132 ± 6 Hz. The corresponding improvement factor from 100 to 150 kHz (132 ± 6 Hz / 90 ± 3 Hz = 1.47 ± 0.09) corresponds well to the expected improvement based
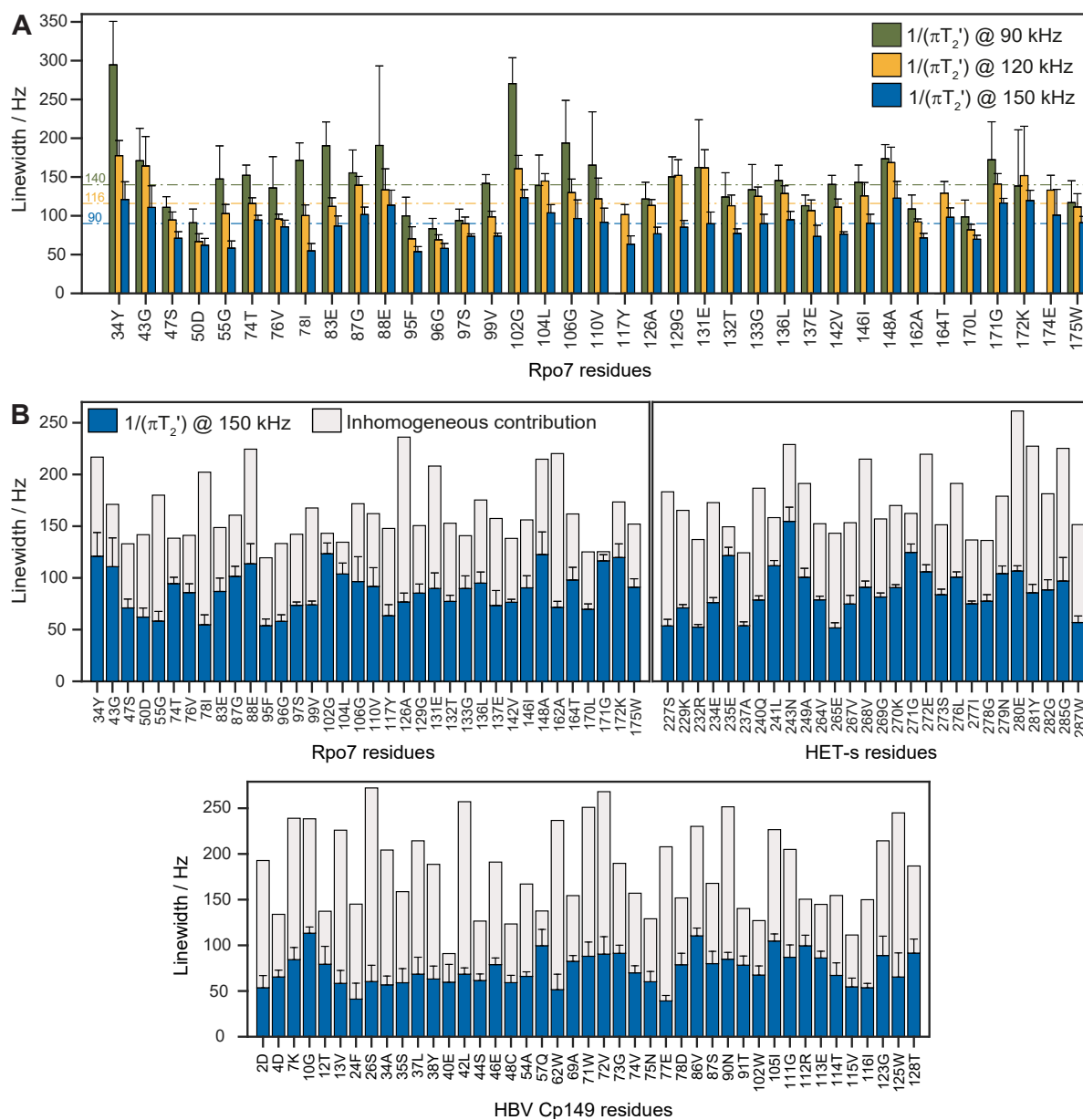
Figure 6.8. *Protein site-specific T$_2$' relaxation analyses for varying MAS frequencies.* A: Relaxation data were recorded at three MAS frequencies in eight 2D spectra for each, with various delays. Peaks that were resolved in all 24 spectra (see Figure 6.3C) were used for site-specific analysis, of which the result is expressed as $(\pi \cdot T_2')^{-1}$ in Hz, representing the homogeneous contribution. The three instances where a 90 kHz histogram bin is missing are troubled by peak overlap in the 2D spectrum and were consequently excluded from this figure. The shown errorbars correspond to 2σ, where σ indicates the experimental error, and are symmetric to the negative side although not shown in this direction for graphic clarity. Median values for each of the three data sets are given as a dashed line, with the corresponding value in Hz toward the far left of the panel. B: Homogeneous and inhomogeneous linewidth contributions for Rpo4/7, HETs (218-289), and HBV Cp149 residues at 150 kHz MAS. Mean inhomogeneous contributions to the total linewidth are 45 %, 51 % and 58 % respectively.

on a linear correlation between the ratio of the spinning frequencies (150 kHz / 100 kHz = 1.5) and indicates that the homogeneous broadening by the dipole interaction is the dominant contribution to the linewidth.

The inhomogeneous contribution $\Delta^{\mathrm{inhomo}} = \Delta^{\mathrm{tot}} - \Delta^{\mathrm{homo}}$ for Rpo4/7 (data at 150 kHz MAS) is given on a per-residue basis in Figure 6.8B, with a mean value of 76 ± 6 Hz, compared to a mean homogeneous contribution of 87 ± 3 Hz. Figure 6.8B also gives site-specific inhomogeneous contributions in HETs (218-289) and HBV Cp149. All three data sets in Figure 6.8B show a varying size of inhomogeneous contributions across residues. These data suggest that for most residues, the largest contribution to the linewidth still comes from coherent contributions, but the inhomogeneous contribution is of comparable magnitude, and often larger for the HBV capsid. As the HBV capsid has four different molecules in the unit cell (see Figure 4.2B, p. 79), unresolved splitting (see Figure 4.3B, p. 82) of the four expected signals might contribute to the inhomogeneous linewidth.

6.3. Conclusion

This chapter presented protein NMR spectra at 150 kHz MAS. While frictional heating was significant (30 °C without external cooling), the cooling system allowed to keep the sample temperature at approximately 10 °C. Spectra of the Rpo4/7 protein complex were recorded within experimental times that are the same as for 110 kHz experiments.

Homogeneous linewidths were shown to improve linearly with the inverse spinning frequency in a range between 90 and 150 kHz MAS. For an MAS increase by a factor 1.5 (from 100 to 150 kHz), we find a narrowing of the total linewidth from 193 ± 6 to 157 ± 3 Hz, improving by a factor 1.23 ± 0.05. In the same range, the homogeneous linewidth improves by a factor 1.47 ± 0.09, corresponding well with said linear relation between MAS frequency and homogeneous linewidth.

The data clearly demonstrate that protein NMR spectroscopy on fully-protonated protein samples of small and intermediate size (*i.e.* up to several hundred amino-acid residues) can greatly benefit from MAS spinning frequencies beyond the presently often applied 110 kHz. The linewidth as a function of the spinning frequency shows a very similar behaviour for all four proteins investigated. Proton networks of strongly coupled spins with near-degenerate resonance frequencies, like the drug meldonium investigated here, still exhibit broader lines, even at 150 kHz. However, their linewidth improvement with increased MAS is larger than

what is observed for proteins. Such small-molecule (drug) compounds can consequently take particular advantage of faster spinning.

6.4. Materials and methods

6.4.1. Sample preparation

The protein complex Rpo4/7* consists of two subunits, Rpo4 (107 amino acids, 12.3 kDa) and Rpo7 (187 amino acids, 21.2 kDa) of the archaeal RNA polymerase from the thermophile *Methanocaldococcus jannaschii* [242]. Each of the two subunits was recombinantly expressed and purified as described previously and, in vitro, they form a stable complex of 33.5 kDa during a urea-based refolding dialysis [243]. The asterisk (*) denotes isotope labeling of Rpo7 with ^{15}N and ^{13}C in M9-minimal medium, while subunit Rpo4 contains these isotopes at natural abundance. The M9-minimal medium [220] consisted of 6.8 g Na_2HPO_4, 3 g KH_2PO_4, 0.5 g NaCl, 1 mL each of 1 M $MgSO_4$, 10 mM $ZnCl_2$, 1 mM $FeCl_3$, 100 mM $CaCl_2$ per 1 L, supplemented with 10 mL MEM vitamin solution (100x), 1 g ^{15}N H_4Cl and 2.5 g ^{13}C-glucose. Purification from inclusion bodies was carried out as described previously, as was expression and purification of Rpo4 (Klose et al 2012).

Filling of a 0.5 mm (outer diameter) rotor with a sample volume of 0.31 μL was carried out by first performing sedimentation of an approximately 30 mg/mL protein solution into a 0.7 mm (outer diameter) rotor with an inner volume of 0.59 μL by 10 h ultracentrifugation at 35,000 RPM (210,000 g) and at 4 °C. To ensure that the protein was sedimented after ultracentrifugation, the UV/vis absorption at 280 nm was measured and compared to the one measured before ultracentrifugation, showing that the protein concentration in the supernatant was reduced by at least two orders of magnitude. Probing the consistency of the sediment after filling is another factor that is taken into account in the process: the amount of resistance on the sample removal tool that is used to make space for the drive cap is indicative of the viscosity of the protein sediment. A custom-made funnel was adjusted to fit the open end of the 0.5 mm rotor. Subsequently, the content of the 0.7 mm rotor was transferred via the funnel into the 0.5 mm rotor, by pushing the sediment down using the blunt end of a 0.4 mm drill. The mass of the protein sediment inside the 0.5 mm rotor was determined to be 0.17 ± 0.02 mg, in which the Rpo4 subunit (*i.e.* 36 % of the protein complex) is at natural abundance, and up to half of a typical sediment may be water [168]. A dedicated filling tool for direct sedimentation into 0.5 mm rotors is presently under construction.

6.4.2. NMR spectroscopy

For all NMR experiments we used a triple-resonance 0.5 mm/310 nL 150 kHz MAS probe, optimised for proton detection. The experiments were carried out on a Bruker Biospin AVANCE III 850 MHz spectrometer. The 150 kHz magic-angle sample rotation was stable within a 10 Hz standard deviation over a period of days, using 3.6 bar bearing pressure and 5.0 bar drive pressure. Spinning was tested up to a stable maximum of 153 kHz. Cooling was performed with an additional gas flow from a cooling unit (BCU II, Bruker) with a gas flow of 360 L/hr at maximum cooling capacity. The resulting sample temperature, extrapolated from the water chemical shift [168], was estimated to 10 °C (see Appendix Figure E.2). The probe was shimmed using DSS (0.3 M) dissolved in D2O (99.96 %) to a proton linewidth of 0.03 ppm on the trimethylsilyl peak at an MAS frequency of 11 kHz and the magic angle was initially set on the same sample and adjusted directly on Rpo4/7 [244].

All experiments used MISSISSIPPI solvent suppression [245], WALTZ 64 decoupling on ^{13}C and ^{15}N during acquisition [246, 247], and frequency-swept low-power TPPM decoupling [248] on ^{1}H during indirect time increments in 2D and 3D experiments. Further experimental details can be found below, in Table 6.2.

The spectra were processed in TopSpin 3.5 (Bruker) and analysed in CcpNmr Analysis 2.4.2 [136, 137].

6.4.3. Computational details

For the estimation of the coherent contributions to the meldonium ^{1}H line widths, we used the codes and the second moment-based simulation method presented and described in detail in [239]. The proton coordinates were extracted from the meldonium crystal structure deposited on the Cambridge Structural Database (CSD ID: 1822460 [249]). Proton positions within one CH_3 group have been replaced by averaged positions, accounting for the fast CH_3 rotation. The simulation run over 150 atoms in a supercell generated in VESTA [250]. Chemical shifts for CH_3 and CH_2 were taken from ATB3.0 [251] computed according to [252]. For the amide shift we took the experimental value.

6.4.4. Relaxation analysis

For site-specific investigation of homogeneous linewidth (for definitions see [28]), an additional eight hNH proton T_2' relaxation experiments were recorded, each at 90, 120 and 150 kHz MAS.

Bulk R_2' relaxation rate constants were extracted by exporting the peak areas of the bulk HN signal as function of the variable delay time into Matlab (The MathWorks 2016, version 9.1.0.441655) and fitting the peak intensities with a monoexponential decay (see Appendix Figures E.3-E.7). Site-specific relaxation-rate constants were extracted from the 2D hNH relaxation experiments, using the INFOS software [253], using as an input the peak positions determined with CcpNMR Analysis [136, 137]. Experimental errors have been calculated using a bootstrapping approach with 500 iterations and error propagation. The shown error bars correspond to 2σ, where σ indicates the experimental standard deviation. Site-specific total linewidths were extracted from spectra processed with a squared sine-bell shift of 2.5 in both dimensions, with the fitting routine included in CcpNMR [136, 137].

Sample	Rpo4[NA] / Rpo7[^{13}C-^{15}N labelled] 0.51 mm			
Spectrometer	AVANCE III 850 MHz (20.0 T)			
Experiment	**hNH 2D**	**hNH 2D**	**hCH 2D**	**hCANH 3D**
MAS frequency / kHz	100	150	150	150
t_1 increments	374	374	388	84
Sweep width (t_1) / ppm	180	180	180	50
Acquisition time (t_1) / ms	12	12	9	4
t_2 increments	2048	2048	2048	28
Sweep width (t_2) / ppm	47	47	46.7	33
Acquisition time (t_2) / ms	26	26	26	5
t_3 increments	-	-	-	2048
Sweep width (t_3) / ppm	-	-	-	46.7
Acquisition time (t_3) / ms	-	-	-	26
Proton decoupling (swfTPPM) / kHz	10	10	10	10
Nitrogen decoupling (WALTZ64) / kHz **	5	5	-	5
Carbon decoupling (WALTZ64) / kHz **	-	-	5	5
Water suppression (120 ms MISS.) / kHz	20	20	20	20
Interscan delay / s	1.2	1.2	1.2	1.2
Number of scans	88	88	96	32
Measurement time / hh:mm	12:13	12:13	13:58	45:28
Transfer I	**HN-CP**	**HN-CP**	**HC-CP**	**HC-CP**
^{1}H /X field / kHz	69/26	114/27	132/30	94/49
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
Carrier ^{13}C / ppm	-	-	56	56
Time / ms	1.0	1.0	0.7	0.5
Transfer II	**NH-CP**	**NH-CP**	**CH-CP**	**CN-CP**
^{1}H /^{13}C/^{15}N field / kHz	69/-/26	114/-/27	132/30/-	-/68/78
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{13}C
Carrier ^{13}C / ppm	-	-	56	56
Time / ms	1.0	1.0	1.0	9.0
Transfer III	-	-	-	**NH-CP**
^{1}H/^{15}N field / kHz	-	-	-	116/26
Shape	-	-	-	Tangent ^{1}H
Time / ms	-	-	-	1.3

* Carriers for ^{1}H and ^{15}N were 4.8 and 117.5 ppm, respectively, for all experiments
** In the 3D, π-pulses were applied on ^{13}C during t_1 and on ^{15}N during t_2

Table 6.2. *Experimental parameters for the presented 2D and 3D measurements.* All spectra were recorded at the same spectrometer, with the same probe head, on the same Rpo4/7 sample, in one continuous measurement session.

7. Conclusion

This thesis presented a number of solid-state NMR studies of large protein assemblies. In those studies, not only did the structure of those proteins play an important role, but also methodological aspects like protonated compared to deuterated proteins, or proton-detected compared to carbon-detected experiments, and fast(er) magic-angle spinning.

   Chapter 2 discussed various familial mutant forms of amyloid-β 1-40. Many mutations in these sequences cluster around residues A21-D23 and the logical question arising from this observation is if a common structural basis exists, which explains the phenotype of early-onset Alzheimer's disease observed for each mutant. The conclusions in this chapter are based on a second important observation, namely that the atomic-resolution structure determinations of WT Aβ 1-40 and of the Osaka mutant E22Δ exhibit strikingly similar chemical shifts for a large portion of the chains, with the exception of the loop that contains the mutation. With the restraints from the Osaka mutant for the unchanged portion of the molecule, and conformational freedom for the mutant loop, the calculations in this chapter demonstrate that the Arctic, Dutch, Flemish, Italian, and Iowa mutant can each adopt the same fold without restraint violations, static clashes, or unbalanced charges inside this loop *if* the K28 sidechain reorients toward the outside of the molecule. This position is different from the WT and allows for the formation of an energetically favourable salt bridge that stabilises the N-terminal end of the second monomer in the same fibril layer. Although the presented data are insufficient to draw a conclusion towards this end, the resulting hypothesis postulates a central role for K28 in the increased toxicity of the discussed mutant forms of Aβ 1-40.

   In contrast to the hypothetical structure calculations for Aβ mutants, **Chapter 3** presented the complete data for structure determination of the somatostatin-14 peptide hormone in its amyloid form. The peptide hormone is stored in secretory granules in the form of amyloid fibrils, before synaptic release as active monomers. The chapter introduced SST-14 fibrils in

EM images in the context of fibril morphology, and EM mass-per-length data that suggest six monomers per fibril layer. The chapter continued to demonstrate complete backbone resonance assignment of two SST-14 fibril polymorphs, one of which was found to be present in the NMR spectra as slightly more intense signals. This presumed major polymorph was then further characterised in terms of dihedral angles and secondary structure, based on secondary chemical shifts and TALOS angles. This was combined with chemical shifts for two cysteines that are indicative of the oxidised form of these residues, which confirms the formation of a disulphide bond as it is also reported in literature, and molecular distance restrains from DARR, PAR, and CHHC spectra. The joined information allowed for the molecular structure calculation of an SST-14 monomer in amyloid form. The resulting fold is characteristic of the cross-β motif commonly observed in amyloids. Two notes were made on somewhat remarkable features of the calculated structure. First, a positively charged lysine sidechain was found to be on the inside of the molecule, and postulated to be charge-compensated by the negative charge of the C-terminus. This assumption could be tested by proton-detected experiments on the Hζ protons. Second, two adjacent aromatic residues were found to have their sidechains pointing to the outside of the monomer. While hydrophobic residues are commonly observed to be buried in the core of a folded protein, the final portion of the chapter suggested that the hexameric conformation of SST-14 in each fibril layer could explain how these residues are jointly forming a hydrophobic core of the fibril, despite, or in fact because of being oriented to the outside of each monomer. This assumption was discussed to be supported at least in part by mutation studies described in the literature, in which one of the two aromatic residues (W8) was found to be crucial for interactions between Aβ and SST-14.

From the exclusively ^{13}C-detected data concerning SST-14, **Chapter 4** proceeded with a presentation of both ^{13}C- and ^{1}H-detected data of human hepatitis B viral capsids and the duck HBV viral envelope. The first part of that chapter in fact compared ^{13}C- and ^{1}H-detection in terms of information content, resolution, and sensitivity. Both methods were found to be relatively comparable in terms of SNR, with the important difference that ^{13}C-detected data require roughly a factor 100 more labelled protein sample. The backbone assignment from ^{13}C-detected experiments was slightly more complete, but did evidently not include proton resonance assignments. The assignment with ^{1}H-detected experiments was completed with spectra from a fully protonated HBV Cp149 sample. In addition, the chapter presented a comparison to deuterated and back-exchanged capsids. The data clearly showed that deuteration significantly reduces the linewidth, even at 100 kHz MAS,

but at the cost of missing signals for solvent-inaccessible H^N sites that remain shielded from back-exchange in the interior of the protein. This conclusion was further supported by data from full-length HBV Cp183 in fully protonated as well as deuterated preparations. In contrast to the Cp149 capsids, Cp183 capsids were prepared by cell-free protein synthesis. The data demonstrated that CFPS is a feasible method for the production of NMR samples. In addition, for deuterated preparations, CFPS was shown to have all exchangeable sites protonated since the complete process is carried out in H_2O and deuteration is achieved by adding deuterated amino acids to the CFPS system. The section of this chapter that discussed Cp183 also demonstrated interaction studies with capsid assembly modulators. The presented data led to the conclusion that CAMs interact with capsids upon assembly in CFPS in the same way as they do with pre-assembled capsids from expression in *E. coli.* This finding forms a basis for future studies of CFPS-produced proteins in which interactions with *e.g.* lipids, enzymes, or chaperones can be probed directly upon protein synthesis from the ribosome. The final portion of the chapter presented another study from a CFPS-produced sample, namely the HBs S surface antigen of the viral envelope, primarily in duck, although preliminary data was included on human HBs S. In contrast to Cp149 and Cp183, the structure of HBs S is not known and the presented data confirmed an existing topology model that suggests a predominantly α-helical conformation. Further studies with HBs S would be necessary to determine an atomic-resolution structure, and it was suggested that such work could include selective labelling, taking advantage of the fact that such a strategy is relatively straightforward to implement in CFPS.

As an application of ^{1}H-detection, **Chapter 5** presented the investigation of binding modes of the *Hp*DnaB helicase to ATP analogues and deoxyribonucleotides of single-stranded DNA. This work, as all structural investigations in NMR, started with assignment. The ^{13}C and ^{15}N resonances of a large portion of the 488 residue monomer had been assigned in previous work, and this chapter started with sequential assignment of H^N and H^α resonances. This was achieved with two intraresidual 3D correlation spectra in opposite directions to detect H^N or H^α respectively, and two relayed 3D correlation spectra from one spin system to the adjacent residue. All spectra were acquired from fully protonated *Hp*DnaB. With this assignment at hand, the chapter proceeded with 2D spectra correlating nitrogen and/or carbon to protons and phosphorus. The results allowed for the residue-specific, and in some cases atom-specific identification of binding partners between *Hp*DnaB and ATP analogues as well as DNA. The data confirm an important role for arginine and lysine sidechains in the coordination of these interaction partners. In addition, the findings

were contrasted to reported data of the homologous protein *Bst*DnaB and were found to compare well. In summary, the chapter presented a method to analyse non-covalent interactions between proteins and dynamic binding partners in general, and specifically between (deoxyribo)nucleotides and motor proteins in transitory states of binding and release during their functional cycle.

The final part of this thesis shifted from protein structure analysis to a focus on magic-angle spinning. **Chapter 6** discussed data up to 150 kHz MAS, for various protein systems. The development of faster MAS is technically challenging and one may rightfully ask if the benefits justify the investment. It is exactly this question that the chapter addressed and the simple answer is: yes, the benefits of faster MAS are worth facing the challenges. In more detail, the chapter started briefly discussing proton assignment of the archaeal RNA polymerase complex of subunits 4 and 7. Rpo4/7 was then used to quantify gains in resolution and sensitivity in 2D correlation spectra at 100 and 150 kHz MAS. The total linewidth of Rpo4/7 improved by a factor of 1.23 ± 0.04 by going from 100 to 150 kHz, and signal intensity increased by a factor 1.48 ± 0.13. The presented data also offered a comparison of several other protein systems and the drug compound meldonium, and their bulk homogeneous linewidths in a range from 90 to 150 kHz MAS. These results showed that all protein systems improve at a strikingly similar linear rate of about 14 Hz decreased homogeneous linewidth per µs decrease in MAS rotor period. By contrast, meldonium showed to be much broader, but also showed to improve at a much higher rate. In conclusion, both for protein and small molecule solid-state NMR, the continued development of faster MAS remains meaningful and beneficial for high-resolution spectroscopy.

The work in this thesis has been performed in a time where fast MAS has transformed solid-state NMR. At the start of the research presented in this thesis (2015), the solid-state NMR group at ETH still consisted of many members who only sporadically or never recorded ^{1}H-detected spectra. By contrast, at the time when this research is brought to conclusion (2019), a number of group members never or only seldomly record ^{13}C-detected experiments. This is neither to say that the transformation did not start in the preceding decade or decades, nor that it has finished. In fact, the statement is not even made to give the impression that carbon detection is a passed station. On the contrary, the results in this thesis confirm that both proton- and carbon-detection have their merits. The choice for either of the two will depend on the research question at hand, and given constraints with regards to for instance the available sample amount, probe heads, and static field strength. Ultimately, it is likely that both methods can complement each other in many cases.
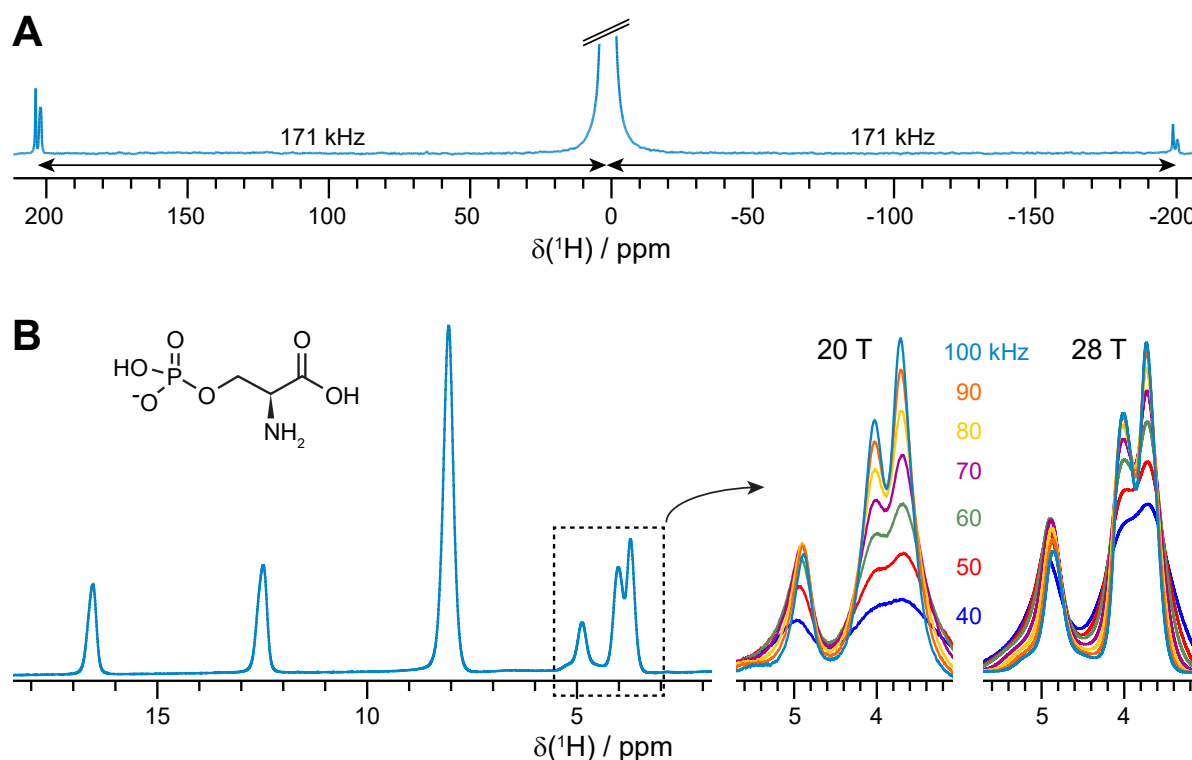
A

B

Figure 7.1. *Continued developments in solid-state NMR.* This figure illustrates two ongoing efforts in solid-state NMR development: the adamantane ^{1}H 1D spectrum in **(a)** has its central line cut off at less than 1 % of its full signal intensity. This is because the focus of this figure is not on the spectrum itself, but rather on the weak MAS sidebands. The spinning sidebands in this spectrum show at 171 kHz, the MAS frequency of the rotor at the time of recording this spectrum. The second development, illustrated in **(b)**, is that of higher static magnetic fields. Spectra are shown for ortho-phospho-L-serine, at various MAS frequencies and at fields of 20 and 28 T (corresponding to ^{1}H Larmor frequencies of 850 MHz and 1.2 GHz, respectively).

Technical developments and improvements in all fields of structural biology continue to shape the research that can be conceived and performed. To illustrate where developments in solid-state NMR may be headed, this conclusive section includes a figure with two examples of developments that were not explicitly presented in the thesis, but were ongoing while this work was written (Figure 7.1). First, while the maximum MAS frequency reported in this thesis is 150 kHz, a new MAS probe from the team around Prof. A. Samoson has arrived in our laboratory and a rotation frequency up to 171 kHz could be reached. Second, Bruker has succeeded in producing a 1.2 GHz magnet, to which our group was invited to collect some preliminary data. A 1D spectrum of ortho-phospho-L-serine at this field (28 T) is shown in Figure 7.1B.

The combination of these two developments is a powerful duo to advance solid-state NMR. The smaller rotor sizes that are associated with faster MAS have a consequence for the amount of labelled sample that can be observed in an NMR experiment. While faster MAS and smaller coil sizes compensate for the reduced sample amount to some extent, a significantly increased B_0 field strength will push sensitivity. This is of particular relevance when considering that the expected MAS frequency that is necessary to resolve *e.g.* the strongly coupled protons in an α-helix is above 200 kHz. This means that the current rotor volume for 150 kHz MAS (around 300 nL) is likely to decrease even further to reach that speed. It is consequently the combination of both faster MAS and higher fields that can shape the near future of solid-state NMR.

In the bigger picture of structural biology, these developments are of importance too. Throughout the thesis, it has been mentioned several times that the future of structural biology will most likely be characterised by collaborations and a hybrid approach to structure determination. In addition, the study of protein dynamics will probably play an increasingly important role. With the help of fast MAS and strong magnetic fields, the high sensitivity of solid-state NMR for even the smallest changes in molecular conformation can continue to contribute to the combined efforts of X-ray crystallography, cryo-EM, MD, EPR, solution-state NMR, and further methods for determination of protein structure and dynamics.

Appendix

A. Supplementary information to alternative salt bridge formation in Aβ

A.1. Figures

The figures in this appendix were also presented in the supporting documentation to [54], the publication presented in Chapter 2.

Figure A.1. (a) Difference of the chemical shifts Cα, -Cβ and C', between the Osaka mutant [53] and wild-type Aβ1-40 as described by Bertini *et al.* in [78]. Cβ is given as negative values to highlight differences in secondary structure propensity such that all chemical-shift differences point in the same direction: where the Osaka mutant shows less β-strand propensity, all three values (Cα, -Cβ and C') are positive, and *vice versa*. The chemical shifts were re-referenced to yield the smallest possible differences. Odd residues are shown in cyan, and even ones in black. **(b)** Average absolute values of the mean chemical shift differences between wild-type Aβ1-40 [78] and the Osaka mutant [53], calculated as $\left(\left| \delta C\alpha^{WT} - \delta C\alpha^{Osaka} \right| + \left| \delta C\beta^{WT} - \delta C\beta^{Osaka} \right| + \left| \delta C'^{WT} - \delta C'^{Osaka} \right| \right) \div 3$. Deviations ≥1 ppm are plotted in red, and <1 ppm in blue (figure continued on next page).

C

D
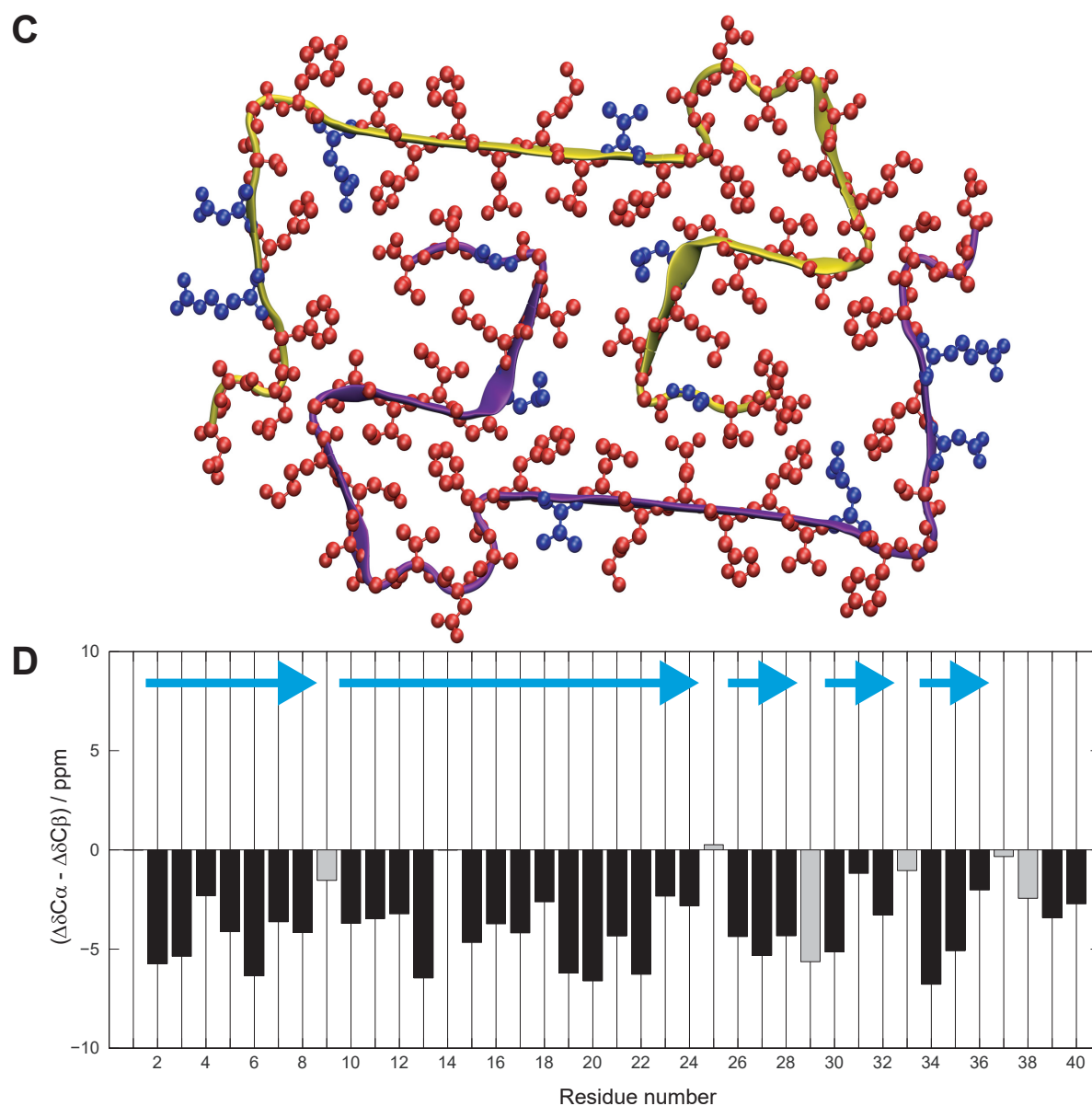
Figure A.1. (continued from previous page) **(c)** Representation of chemical-shift differences from (b) plotted on the Aβ1-40 E22Δ structure (PDB ID 2MVX), with the same colour coding as the bar chart. **(d)** Differences between Cα and Cβ secondary chemical shifts for the polymorph described by Bertini *et al.*[78], suggesting five β-strands, assuming that each glycine (marked as a grey bar) interrupts the strand, even for negative secondary Cα chemical shifts [6]. Indeed, G9 seems to interrupt the first long stretch, as no long-range restraints were detected for these residues and the structural model displays a manually added dynamic first β-sheet including residues 1-8 [78]. Glycine 25 clearly points to a turn, and, considering the resulting structural model [78], Gly29 must represent a second turn. However, unambiguous distance restraints have been detected neither for the N-terminal strand, nor for residues 22-30, for which the structure remains undefined.

A Wild-type

B Dutch E22Q

C Iowa D23N
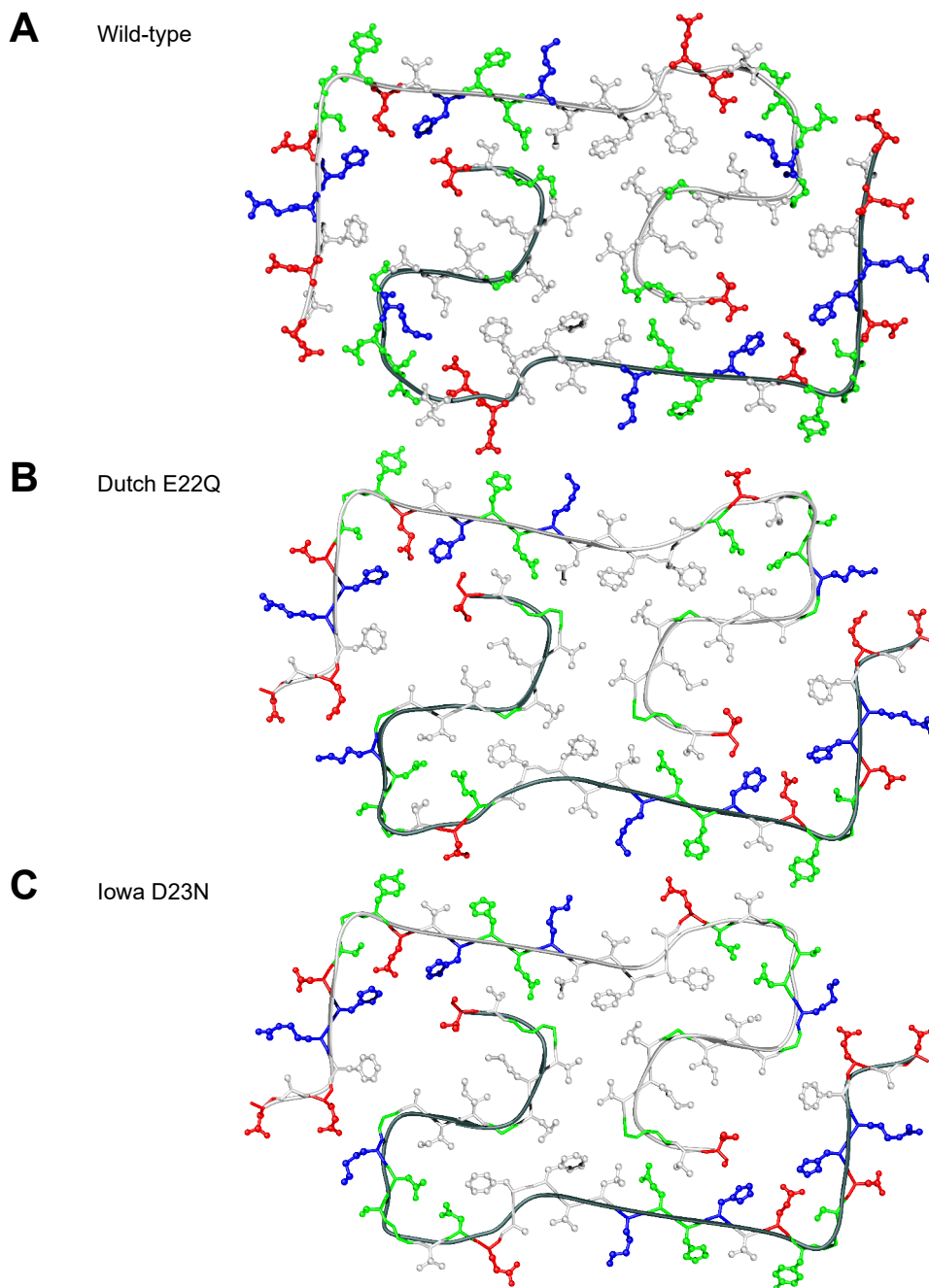
Figure A.2. Atomic-detail models from CYANA structure calculations (continued on next page)

D   Arctic E22G

E   Flemish A21G

F   Italian E22K
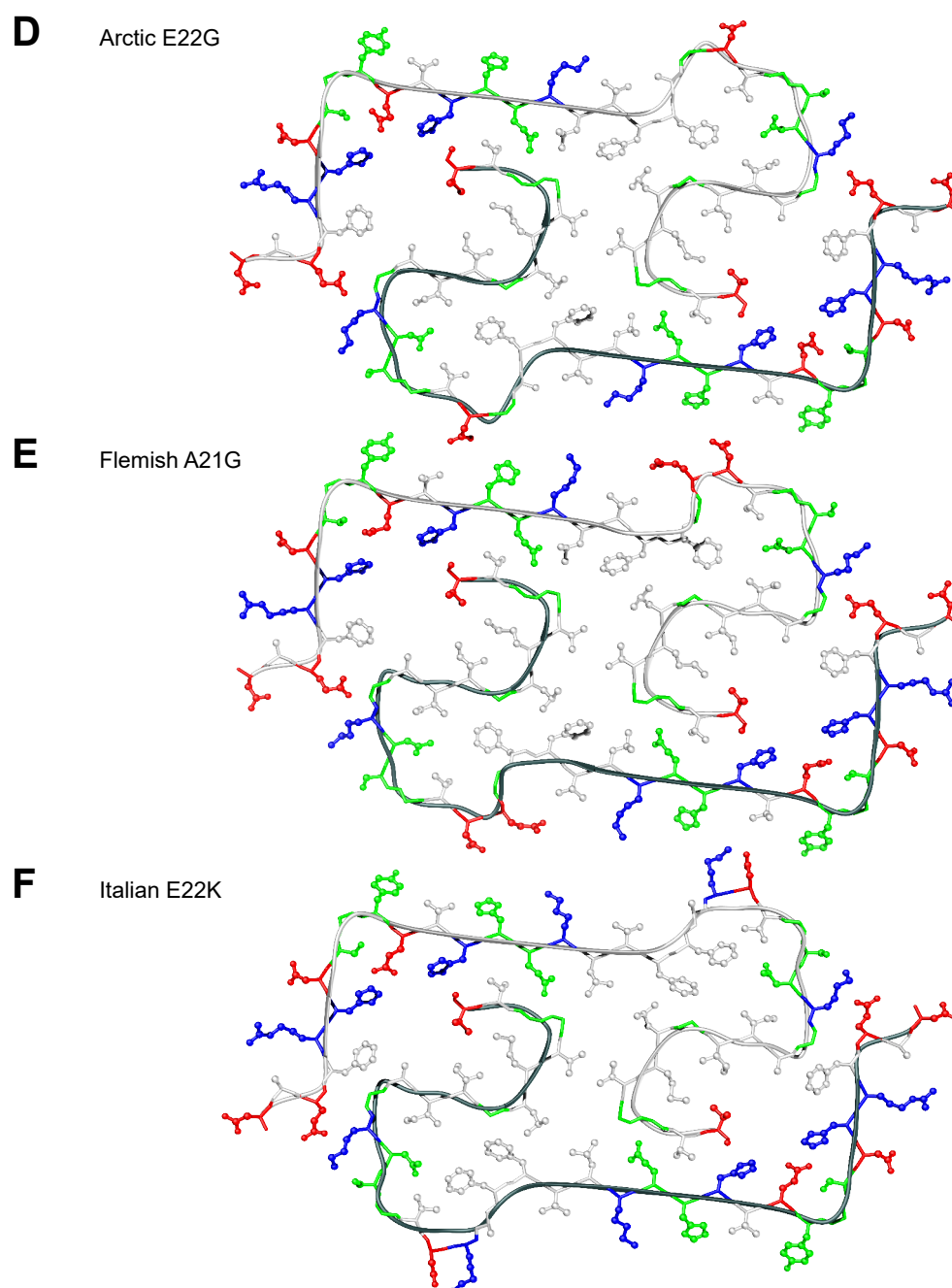
Figure A.2. (continued from previous page) Atomic-detail models that resulted from CYANA structure calculations, forming the basis for the hand-drawn pictures in Figure 2.2, where the drawing by hand is to emphasise the hypothetical nature of the models. Hydrophobic residues are coloured white, glycines and polar residues are green, and charged residues are coloured blue when positive and red when negative. The statistics for the structure calculations is given in Table 2.2.

A.2. Tables

The table in this appendix was also presented in the supporting documentation to [54], the publication presented in Chapter 2.

DARR	ω_1	ω_2
~~E3 - K28~~	~~E3 CA~~	~~K28 CG~~
	~~K28 CG~~	~~E3 CA~~
	~~E3 CA~~	~~K28 CD~~
~~E3 - G29~~	~~G29 C~~	~~E3 CA~~
~~E3 - A30~~	~~E3 CA~~	~~A30 CA~~
	~~A30 CA~~	~~E3 CA~~
H13 - V40	H13 CE1	V40 CA
	V40 CA	H13 CE1
	V40 C	H13 CA
	V40 C	H13 CD2
	V40 C	H13 CE1
	V40 CA	H13 CD2
H6- E11	H6 CD2	E11 CG
	E11 CG	H6 CE1
F19 - L34	F19 CB	L34 CG
	L34 CG	F19 CA
	L34 CG	F19 CZ
~~F20 - I31~~	~~F20 CG~~	~~I31 CG2~~
	~~F20 CE1~~	~~I31 CG2~~
I32 - M35	I32 CG2	M35 CB
	M35 CB	I32 CG2
M35- G38	M35 CA	G38 C
	G38 C	M35 CA
	G38 C	M35 CE

PAR	ω_1	ω_2
~~E3 - K28~~	~~E3 CA~~	~~K28 CG~~
	~~K28 CG~~	~~E3 CA~~
~~E3 - A30~~	~~E3 CA~~	~~A30 CA~~
	~~A30 CA~~	~~E3 CA~~
H6- H13	H6 CE1	H13 CE1
	H13 CE1	H6 CE1
F19 - L34	L34 CG	F19 CA
	L34 CG	F19 CB
I32 - M35	I32 CG2	M35 CB
	M35 CB	I32 CG2
	I32 CD1	M35 CB
	M35 CB	I32 CD1

CHHC	ω_1	ω_2
~~E3 - K28~~	~~E3 HA~~	~~K28 QG~~
H6- H13	H6 HE1	H13 HE1
	H13 HE1	H6 HE1
F19 - L34	F19 QB	L34 HG
I32 - M35	I32 QG2	M35 QB
	M35 QB	I32 QG2
	M35 QB	I32 QD1

PAIN	ω_1	ω_2
~~E3 - G29~~	~~E3 CA~~	~~G29 N~~
H13 - V40	V40 C	H13 ND1
	V40 CA	H13 ND1
Q15- G37	G37 CA	Q15 NE2
Q15- G38	G38 C	Q15 NE2

Table A.1. Unambiguous restraints extracted from Aβ1-40 E22Δ. The basic table (without strikeouts) is reproduced from [53] (with permission). Restraints that are struck through were not used here or the CYANA manual calculations of WT and all mutants (besides E22Δ). A red background indicates *inter*molecular restraints, while *intra*molecular restraints have a grey background. For details on the restraint collection and calculations, see [53].

B. Supplementary information to molecular structure of SST-14

B.1. Figures

Figure B.1. *PAR and CHHC distance restraint spectra.* Unambiguous assignments are indicated for cross-peaks that are at least $i \pm 2$.

DARR

3CC' 5NCβ

11FC' 13SCβ

14CC' 12TCβ

10TCβ 12TCγ2

4KCβ 12TCγ2

3CCβ 1ACβ

14CCβ 4KCδ

10TCβ 8WCβ

14CCβ 4KCβ

$\delta_1(^{13}C)$ / ppm

$\delta_2(^{13}C)$ / ppm

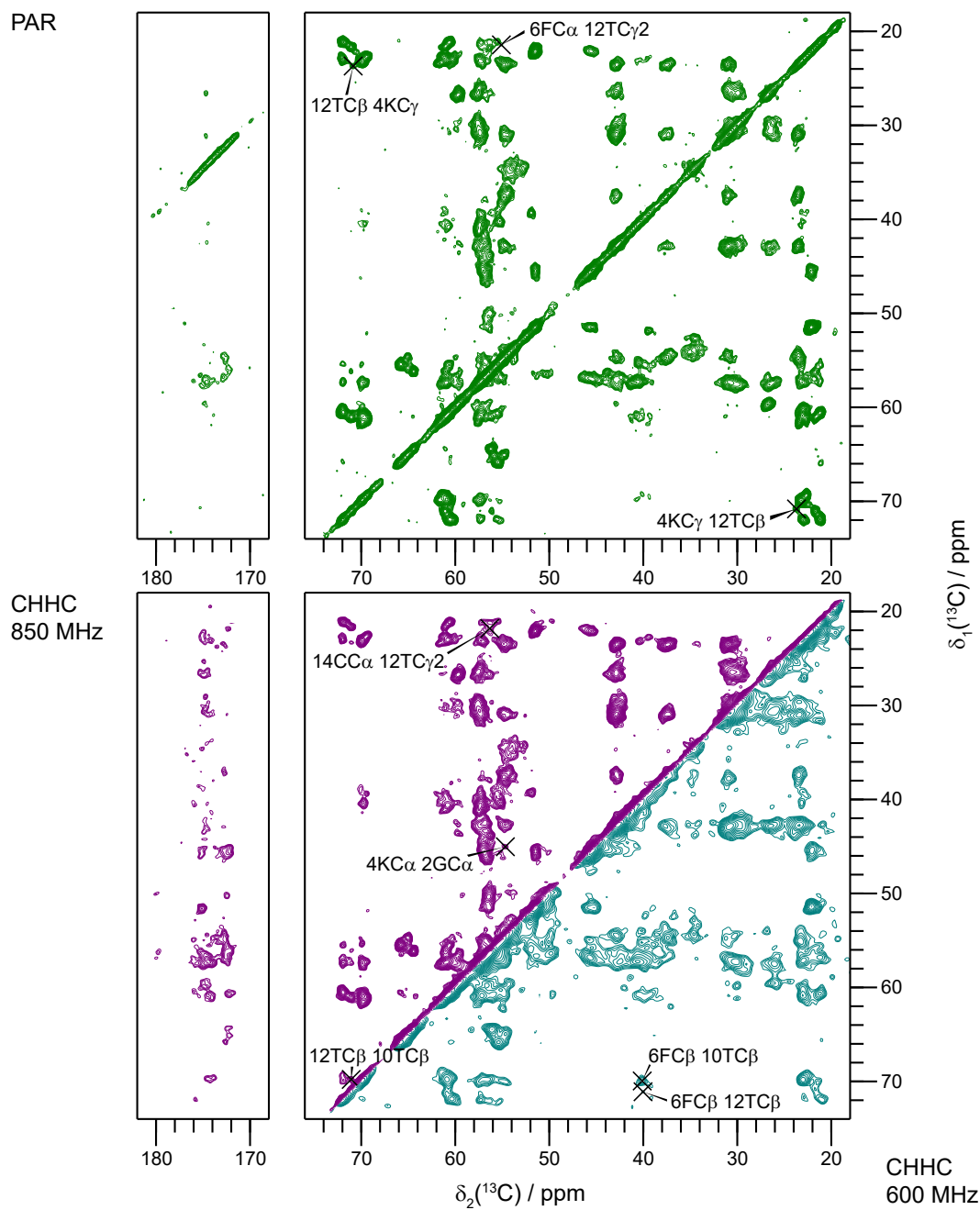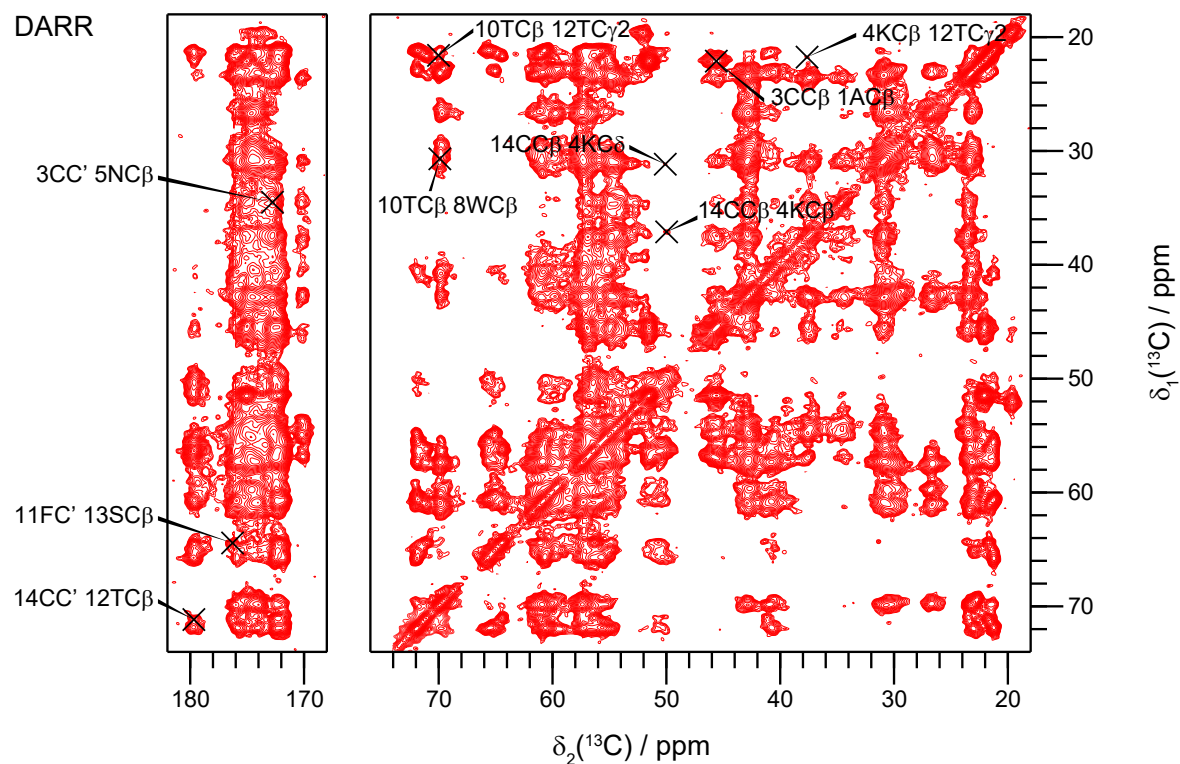**Figure B.2.** *400 ms DARR distance restraint spectrum.* Unambiguous assignments are indicated for cross-peaks that are at least $i \pm 2$.
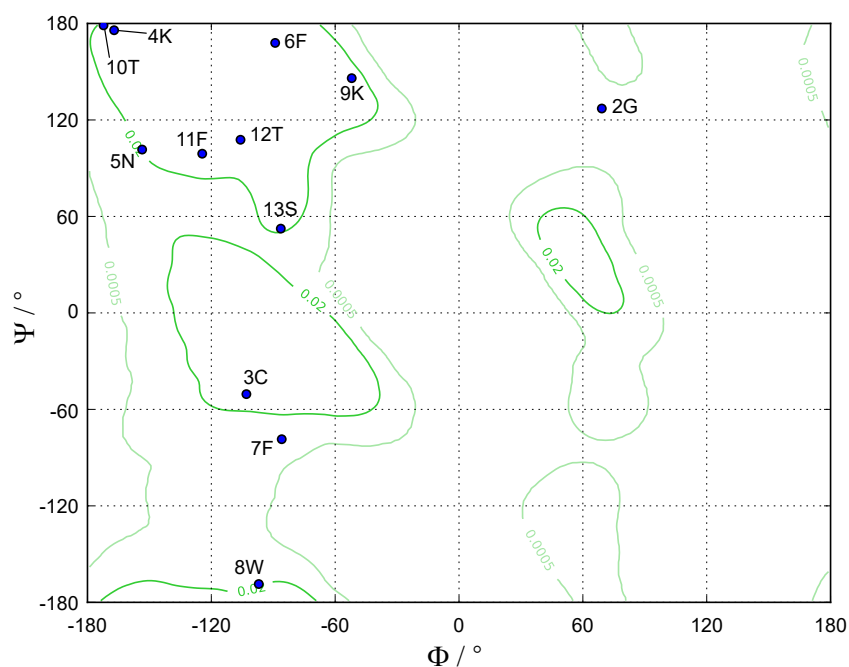
Ψ / °

Φ / °

Figure B.3. *Ramachandran plot of the SST-14 fibril monomer.*

B.2. Tables

	Phi / °	σ_{Phi} / °	Psi / °	σ_{Psi} / °	DIST	S^2	Count	CS_count	Class
1A	9999	0	9999	0	0	0	0	7	None
2G	-159.662	52.706	163.296	20.444	49.184	0.904	10	11	Warn
3C	-133.329	15.322	161.256	9.801	57.419	0.916	10	11	Good
4K	-141.759	12.150	161.992	6.902	49.678	0.921	10	12	Good
5N	-152.657	50.357	152.299	38.270	53.319	0.911	9	12	Warn
6F	-111.297	11.468	137.043	15.700	39.840	0.893	10	12	Good
7F	-125.994	20.520	145.177	16.220	26.889	0.867	10	12	Good
8W	-133.780	29.781	138.956	26.371	43.169	0.847	10	12	Good
9K	53.712	6.597	44.283	7.582	83.708	0.817	10	12	Warn
10T	-125.984	31.835	154.480	16.915	43.125	0.814	10	12	Good
11F	-102.262	17.986	129.286	9.358	33.318	0.803	10	12	Good
12T	-120.176	14.853	126.300	11.009	28.093	0.794	10	12	Good
13S	-112.271	16.599	120.211	8.095	50.771	0.770	10	12	Good
14C	9999	0	9999	0	0	0	0	8	None

DIST is the TALOS+ database matching score.

S^2 is the Wishart RCI chemical shift order parameter [254].

Count is the number of database triplets used to form the torsion angle predictions.

CS_count is the count of exp. chem. shifts of the target residue and its two neighbours.

Class is the classification of the prediction result:

  None: no torsion prediction was made.

  Good: majority consensus in database matches; prediction is likely to be good.

  Warn: no consensus in database matches, do not use prediction.

  Dyn: RCI S2 value indicates that residue has dynamic conformation.

Table B.1. *Talos output of dihedral angle predictions for SST-14.*

#	Residue	H/E/C
1	A	C
2	G	C
3	C	C
4	K	C
5	N	C
6	F	E
7	F	E
8	W	E
9	K	E
10	T	E
11	F	E
12	T	E
13	S	E
14	C	C

Table B.2. *GOR4 [125] secondary structure prediction based on the primary sequence of SST-14.* The predicted secondary structure elements are abbreviated according to the DSSP standard: H for α-helix, E for extended β-strand, and C for coil.

C. Supplementary information to HBV envelope and capsids

C.1. Materials and methods

The details concerning sample preparation and biochemical work with the HBV capsids and envelope were predominantly performed in and by the group of Anja Böckmann in Lyon, while most NMR spectroscopy was performed in the solid-state NMR group of Beat Meier in Zürich. The presented experimental details were also described in [138–140] on which Chapter 4 is based.

C.1.1. Cp149

Sample preparation

Uniformly ^{13}C-^{15}N labelled fully-protonated pCp149 capsid samples were prepared as previously described in [144]. Preparation of the uniformly ^{2}H-^{13}C-^{15}N labelled dCp149 resulted in a slightly modified expression protocol. Cells from 10 mL inoculated LB medium were collected by centrifugation (4,000 g, 10 min, 20 °C) and resuspended in 100 mL of D$_2$O-based M9 minimal medium containing 1 g/L of ^{15}NH$_4$Cl and 2 g/L of deuterated ^{13}C-glucose as sole nitrogen and carbon sources, respectively. The deuterated culture was incubated overnight at 37 °C, and then transferred into 900 mL of fresh D$_2$O-based M9 medium. When the culture reached an OD$_{600}$ of 2.0, protein expression was induced by adding 1 mM IPTG and cells were grown overnight at 25 °C. Protonated and deuterated capsids were purified as previously described in [144], allowing the accessible labile deuterons in the deuterated sample to back-exchange to protons. For fast-MAS NMR measurements, capsids were dialysed overnight at 4 °C in the solid-state-NMR buffer (50 mM TRIS pH 7.5, 5 mM DTT) and ~0.6 mg (Appendix Table C.1) of capsids were filled into 0.7 mm rotors using home-made filling tools [168] by centrifugation (200,000 g, 14h, 4 °C). A minute amount of saturated to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) solution was added to the protein sediment before closing the rotor for chemical shift referencing.

NMR spectroscopy and data processing

The carbon-detected spectra used for comparison are described in [144]. Shortly, they were acquired on a wide-bore 800 MHz Bruker Avance II spectrometer equipped with a 3.2 mm triple-resonance MAS probe at 17.5 kHz MAS and a sample temperature of 4 °C, as

determined from the relationship T(°C)$= 455 - 90 \cdot \delta_{H_2O}$ described in [255], where the water chemical shift (δ_{H_2O}) corresponds to the supernatant water signal [168].

The [1]H-detected spectra at 100-110 kHz were acquired on a wide-bore 850 MHz Bruker Avance III spectrometer equipped with a 0.7 mm triple-resonance MAS probe and referenced DSS. The magic angle was set using a 0.7 mm rotor with glycine ethyl ester, optimising intensity and J-coupling-based splitting of the CO resonance [244]. The MAS frequency was set to 100 kHz and the VT gas temperature to 273 K using a nitrogen gas flow of 400 L/h, corresponding to a sample temperature of 22 °C, extrapolated from the water chemical shift in a [1]H 1D [168, 255] as detailed above. On the uniformly [13]C-^{15}N labelled fully-protonated sample, a set of four three-dimensional (3D) spectra (hCANH, hCONH, hCAcoNH, hncaCB-caNH) [27] and one two-dimensional (2D) fingerprint spectrum (hNH) were recorded. For comparison, 2D hNH and hCANH spectra with virtually identical acquisition parameters (except for the MAS frequency which was set to 110 kHz in the hCANH) were recorded on dCp149 capsids. Acquisition parameters are given in Appendix Table C.2.

The 2D hNH and 3D hCANH spectra [27, 170] [1]H-detected spectra at 60 kHz on dCp149 were acquired on a narrow-bore 700 MHz Bruker Avance II spectrometer equipped with a 1.3 mm triple-resonance MAS probe and referenced to DSS. The MAS frequency was set to 60 kHz and the VT gas temperature to 248 K using a nitrogen gas flow of 1400 L/h, corresponding to a sample temperature of 16 °C, extrapolated from the water chemical shift in a [1]H 1D [168, 255] as detailed above. Acquisition parameters are given in Appendix Table C.3.

All spectra were processed using TopSpin 4.0.3 (Bruker Biospin) by zero filling to no more than double the number of acquired points. Spectra were apodised in the direct and indirect dimensions with a shifted sine-bell window function (SSB = 3), except for the determination of linewidths, where no apodization was applied. Spectral analyses and resonance assignments were performed using CcpNmr Analysis 2.4.2 [136, 137]. Peak positions, linewidths and peak intensities were fitted using the parabolic fit function integrated in CcpNmr. The standard deviations of the average linewidths were calculated by the square root of the difference between the individual linewidths minus their mean value, divided by the number of lines minus one.

For comparison of sensitivites of spectra from 3.2 mm (thin-wall), 1.3 mm and 0.7 mm rotors, we calculated the total mass by weighting the empty and full rotors (Appendix Table C.1).

The T_2' bulk relaxation times were measured using a Hahn-echo inserted after an hNH dipolar-coupling based polarization transfer sequence. The resulting proton-detected bulk amide signal has been recorded for different variable delay points. The peak area of each signal was extracted using TopSpin and then exported to MATLAB (MATLAB 2016R, The MathWorks Inc., Natick, MA 2016), where the relaxation series was fitted with a mono-exponential decay function. The fit error was derived using a bootstrapping approach.

Assignment deposition

H^N chemical shifts of the fully-protonated ^{13}C-^{15}N-labelled Cp149 capsids at 100 kHz MAS were deposited in the Biological Magnetic Resonance Data Bank (BMRB) under accession number 27845. ^{15}N, $^{13}C'$, $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts, for which some resonances slightly differ from ^{13}C-detection based assignments (BMRB 27317 [144]), were deposited as well.

C.1.2. Cp183

Plasmids

The genes corresponding either to the full-length core protein (Cp183) or to its truncated form Cp149 were cloned into the pEU-E01-MCS vector (CellFree Sciences, Japan) for WGE-CF expression. The plasmids were amplified in DH5α bacteria, and purified using a Nu-cleoBond Xtra Maxi kit (Macherey-Nagel, France). An additional purification step was performed with a phenol/chloroform extraction to ensure the purity of the plasmid according to the recommendations of CellFree Sciences (Yokohama, Japan).

mRNA transcription

Transcription was performed according to [256] in 1.5 mL Eppendorf tubes using 100 μg/mL plasmid, 2.5 mM NTP mix (Promega), 1U/μL SP6 RNA Polymerase (CellFree Sciences) and 1U/μL RNase inhibitor (CellFree Sciences) in transcription buffer (CellFree Sciences) containing 80 mM Hepes-KOH pH 7.6, 16 mM magnesium acetate, 10 mM DTT and 2 mM spermidine. After incubation for 6 h at 37 °C, mRNA was used directly for translation.

Wheat germ cell-free protein synthesis

Non-treated durum wheat seeds (Sud Céréales, France) were used to prepare home-made WGE as described in [169], according to the protocol of Takai and colleagues [256] with

minor modifications. Translation was performed using the bilayer method as described in [169, 256] for small scale expression tests in the presence of compounds, or using the dialysis mode as described in [138] for larger scale production followed by isolation on a sucrose density gradient. For the bilayer method, the bottom layer (20 µL) corresponding to the translation mixture contains per well 10 µL of mRNA, 10 µL of WGE, 40 ng/µL of creatine kinase and 6 mM of amino-acid mix (0.3 mM per amino acid, average concentration). The upper layer (200 µL) corresponding to the feeding buffer contains SUB-AMIX NA (CellFree Sciences; 30 mM Hepes-KOH pH 7.6, 100 mM potassium acetate, 2.7 mM magnesium acetate, 16 mM creatine phosphate, 0.4 mM spermidine, 1.2 mM ATP, 0.25 mM GTP and 4 mM DTT), and 6 mM of amino-acid mix (0.3 mM per amino acid, average concentration). For Cp183 expression in the presence of different compounds, 10 nmol of antiviral (dissolved in DMSO at a concentration of 10 mM) was added into 200 µL feeding buffer and translation was performed at 22 °C for 16 h.

For large-scale production, dialysis cassettes with a volume of either 500 µL or 3 mL, depending on the production scale, and a MWCO of 10 kDa were used. The translation mixture contained $\frac{1}{2}$ by volume of feeding buffer, $\frac{1}{3}$ of mRNA, $\frac{1}{6}$ of WGE, 40 ng/µL of creatine kinase, 0.3 mM of amino-acid mix. The feeding buffer (either 20 mL or 124 mL for a 500 µL or a 3-mL dialysis cassette, respectively) contains SUB-AMIX NA (CellFree Sciences) as described above, supplemented with 0.3 mM of amino-acid mix. The dialysis cassette containing the translation mix was soaked in the feeding buffer, and incubated for 16 h under shaking at 60 rpm, 22 °C. A mix containing all twenty isotopically labelled amino acids (Cambridge Isotope Laboratory) was used for the production of ^{13}C-^{15}N-Cp183 for NMR studies in a 3 mL-translation reaction experiment.

Isolation of the capsids on a sucrose density gradient

The total cell-free reaction mixture (CFS) was treated with 25,000 units/mL of benzonase for 30 min at room temperature before centrifugation at 20,000 g, 4 °C for 30 min. The supernatant (SN) was loaded onto a discontinuous sucrose gradient with layers of 10, 20, 30, 40, 50, and 60 % sucrose (w/v), each with a volume of 350 µL for a production in a 500 µL cassette. For the production of a ^{13}C-^{15}N-Cp183 sample in a 3-mL dialysis cassette, the supernatant (SN) was split into two fractions and loaded onto two sucrose gradients with layers of 10, 20, 30, 40, 50, and 60 % sucrose (w/v), each with a volume of 1.5 mL. The gradients were centrifuged at 200,000 g, 4 °C for 12 h. After centrifugation, the different

sucrose fractions were harvested and analysed by SDS-PAGE and Western blotting, as well as by electron microscopy after negative staining as described below.

Capsids from *E. coli*

Cp183 capsids used as reference for negative stain EM with CAMs were obtained from BL21*-CodonPlus (DE3) cells using plasmid pRSF-T7-HBc183opt. Expression and purification were done as previously reported [144, 257]. In brief, protein was expressed overnight after induction with 1 mM IPTG at 20 °C, and cell lysate was separated with a 10-60 % sucrose gradient. Cp183 capsids were precipitated after the sucrose gradient by 40 % saturated ammonium sulfate, and resuspended in final buffer (50 mM Tris pH 7.5, 5 mM DTT, 1 mM EDTA, 5 % sucrose). The interaction between preformed Cp183 capsids and compounds was performed at a monomer:compound molar ratio of 1:4, at 37 °C for 2 h.

Rotor filling and NMR data acquisition

Four different Cp183 NMR samples were prepared: two from cell-free protein synthesis, one synthesised using $^{13}C/^{15}N$, and the other one $^{2}H/^{13}C/^{15}N$ amino acids, resulting in a protonated sample, and a deuterated, but 100 % protonated on exchanging protons, as synthesis is carried out in H_2O; and for reference two samples from *E. coli* expression, one deuterated and back exchanged on exchangeable sites, and one protonated [144, 257]. NMR samples were filled into 0.7 mm rotors as sediment obtained by ultracentrifugation directly into the rotor [168] at 200,000 g for approximately 16 h at 4 °C, yielding approximately 0.5 mg of sediment. As an internal chemical-shift reference, about 30 µL of saturated (0.3 M) 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was added to 250 µL of protein solution before sedimentation.

On each of the samples a two-dimensional (2D) fingerprint hNH spectrum was recorded. On the protonated, uniformly ^{13}C-^{15}N labelled cell-free produced sample, an hCANH 3D [27] was recorded in addition. All spectra were acquired on a wide-bore 850 MHz Bruker Avance III spectrometer with a 0.7 mm triple-resonance MAS probe (Bruker Biospin) operated at 100 kHz MAS. Magic angle and shim for this probe were set using a 0.7 mm rotor with glycine ethylester by optimising the intensity and J-coupling based splitting of the CO resonance [244]. The sample was cooled with a BCU (Bruker Cooling Unit) gas flow of 400 L/h with a VT (Variable Temperature) set to 272 K, corresponding to a sample temperature of

approximately 22 °C, extrapolated from the water chemical shift in a ^{1}H 1D [168, 255]. Detailed acquisition parameters can be found in Appendix Table C.4.

NMR data processing

TopSpin 4.0.3 (Bruker Biospin) was used for the data acquisition and processing. 2D hNH spectra were processed with 1024 points in the ^{1}H dimension (corresponding to 12.9 ms of acquisition time) and zero filling was applied to respectively 4096 points in ^{1}H and 1024 points in ^{15}N dimension. The 3D hCANH was processed with zero filling to respectively 2048 points in ^{1}H, 128 points in ^{15}N and 256 points in ^{13}C dimensions. All spectra were apodised with a shifted sine-bell window function using SSB = 3.5 in TopSpin. Linear prediction to twice the recorded number of points was applied in the ^{15}N dimension for 2D hNH spectra of the protonated capsids produced by CFPS, and the deuterated *E. coli* capsids, in order to reach a similar number of points as acquired for the other samples. Spectral analyses were performed using CcpNmr Analysis 2.4.2. The proton linewidths were obtained using the parabolic fit function integrated on CcpNmr on six isolated peaks in the hNH spectra. The errors given represent the standard deviations between the six values. Signal-to-noise ratios were calculated on the bulk signals from 1D hNH spectra recorded and processed with similar parameters and divided by the square root of the number of scans.

SDS-PAGE & Western blotting analysis

The expression of Cp183 was assessed by 15 % Coomassie blue stained SDS-PAGE and Western blotting as described in [258]. A polyclonal rabbit antiserum *'a-c149'* against the N-terminal domain of the HBV core protein was used to detect both Cp149 and Cp183 on blots.

Negative staining electron micrographs

Samples for electron microscopy were negatively stained as described in [144]. Briefly, 5 μL of each fraction were loaded on a carbon-coated grid (EMS Microscopy) and incubated for 2 min at room temperature. Remaining liquid was drained using Whatman paper. Grids were negatively stained on a 50 μL drop of 2 % phosphotungstic acid (pH 7) for 2 min at room temperature and observed with a JEM-1400 transmission electron microscope operating at 100 kV.

C.1.3. HBs S

Plasmids

cDNA of DHBs S (DHBV Strain QCA34 ; UniProt accession number Q66405/ENA accession number X60213) was amplified by PCR and cloned into pEU-E01-MCS vector (CellFree Sciences, Japan). A Strep-Tag II has been fused at the C-terminal end of the protein (WSH-PQFEK). The resulting plasmids were transformed into *Escherichia coli* TOP10 competent cells (Life Technologies). DNA was prepared using a NucleoBond Xtra Maxi kit (Macherey-Nagel, France). Plasmids were further purified by a phenol/chloroform extraction, according to CellFree Sciences (Yokohama, Japan) recommendations.

Wheat germ cell-free expression

Home-made wheat germ extracts were prepared using non-treated durum wheat seeds (Semences du Sud, France) according to Takai and colleagues [256], with uncoupled transcription and translation. Transcription was performed using 100 ng/µL plasmid, 2.5 mM NTP mix (Promega, France), 1U/µL RNase inhibitor (CellFreeSciences, Japan) and 1U/µL SP6 RNA polymerase (CellFreeSciences, Japan) in transcription buffer (CellFreeSciences, Japan). After 6 h incubation at 37 °C, the mRNA was used directly for translation. For small scale expression tests, translation was performed using the bilayer method as described in [256] and [169, 258], whereas a dialysis method [259] was applied for larger scale production in order to have reasonable volume for sucrose gradient analysis. The translation mix was transferred in a dialysis cassette (Slide-A-Lyzer, Thermofisher, USA) with a 10 kDa cut-off. It contained mRNA, wheat germ extract, 6 mM amino-acid mix, 4 µg/mL creatine kinase, 0.2 % MNG-3 (w/v) and a feeding buffer, constituted by all the substrates needed for protein synthesis: 30 mM HEPES-KOH pH 7.8, 4 mM DTT, 0.25 mM GTP, 1.2 mM ATP, 0.4 mM spermidine, 16 mM creatine phosphate, 2.7 mM magnesium acetate and 100 mM potassium acetate. For the translation step, the translation mix was dialysed against 40 volumes of feeding buffer for 16 h at 22 °C under shaking.

SDS-PAGE and Western blotting analysis

All experiments were assessed using 15 % polyacrylamide SDS-PAGE gels as described in [258]. Samples were resuspended in a loading buffer containing 62.5 mM Tris-HCl pH 6.8, 10 % glycerol (v/v), 2 % SDS (w/v), 5 % β-mercaptoethanol (v/v) and 0.01 % bromophenol

blue (w/v) and incubated at room temperature for 15 min before loading. Western blotting analysis was carried out by protein transfer onto a nitrocellulose membrane using an iBlot2® gel transfer device. The nitrocellulose membrane has then been blocked with 5 % non-fat milk powder in PBS-T buffer, which contains 12 mM sodium phosphate pH 7.4, 137 mM NaCl, 2.7 mM KCl, 0.05 % Tween® 20 (v/v). The membrane was then incubated with a StrepMAB-Classic mouse primary antibody (IBA Lifesciences, Germany) for ^{1}H at RT and further incubated with an anti-mouse IgG HRP conjugated secondary antibody (Promega, France) for ^{1}H at RT. Epitope-containing bands were observed using a Amersham ECL Prime Western Blotting Detection Reagent kit (GE Healthcare, France).

Analysis on discontinuous sucrose density gradient

For analysis, the soluble fraction of DHBs S was layered onto a discontinuous sucrose gradient made of 350 µL layers of 10, 20, 30, 40, 50, and 60 % of sucrose. For preparation of NMR samples, the gradient was either made of one 6 mL layer of 30 % of sucrose and one 3 mL layer of 70 % of sucrose, or 750 µL layers of 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, and 60 % of sucrose. The gradients were centrifuged at 200 000 g for 12 h at 4 °C in a SW 60 or SW 41 Ti rotor, for 3.2 mL and 14 mL centrifugal tubes respectively. Each fraction has then been harvested and analysed by SDS-PAGE and Western blotting, as described above.

Affinity chromatography purification

The protein was expressed in the presence of 0.2 % MNG-3 using the dialysis method in a 0.5-mL dialysis cassette, as described above. The CFS has been incubated with 2.5 U/µL benzonase and 0.25 % (w/v) DDM on a rolling wheel for 30 min at RT. Afterwards, CFS have been centrifuged at 20 000 g, 4 °C for 30 min. The supernatant obtained has been loaded on a 200 µL Strep-Tactin® Superflow® gravity flow column (IBA Lifesciences, Germany). Purification was performed as described by the manufacturer, with all buffers containing 0.1 % (w/v) DDM. The protein of interest was eluted in 100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM D-desthiobiotin and 0.1 % DDM. Final protein concentration was determined at 280 nm using a NanoDrop spectrometer.

Transmission electron microscopy

For negative staining experiments presented in Figure 4.1, 5 µL of sample were incubated on 300-mesh Cu-Formvar grids for 2 minutes. The grids were then layered on top of 50 µL

of 2 % phosphotungstic acid (w/v) pH 7 and incubated at room temperature for 2 minutes. Grids were finally stored at room temperature until observation. Samples were analysed using a Jeol JEM 1400 transmission electron microscope (Jeol, Tokyo, Japan) with a built-in Orius 600 camera (Gatan, USA).

Solid-state NMR experiments

NMR spectra were acquired on a Bruker AVANCE III wide-bore spectrometer operating at 850 MHz [1]H Larmor frequency using a 0.7 mm probe. The details of NMR experiments are given in the Appendix Table C.5. The sample temperature was set to 292 K using DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) as an internal standard [168]. The 2D spectra were processed using TOPSPIN 3.5 with a shifted sine-bell window function as indicated and analysed using CcpNMR software [136, 137].

C.2. Figures

The figures presented below were taken from main texts and appendices of [138–140] on which Chapter 4 is based, with the exception of Appendix Figure C.10.
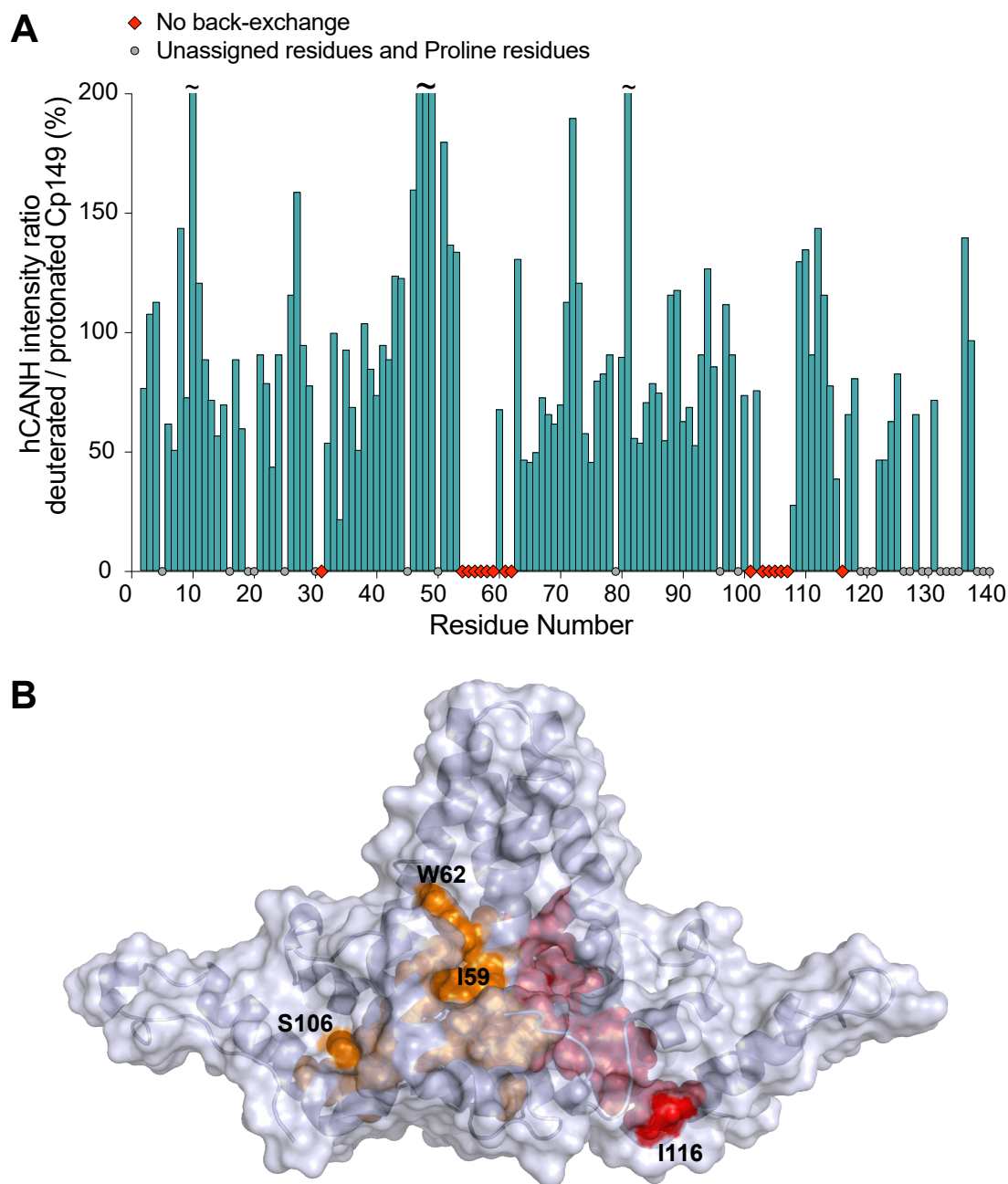
Figure C.1. *Back-exchange in deuterated Cp149.* **(a)** Intensity ratio of NMR signals in the 3D hCANH spectra of back-exchanged deuterated Cp149 capsids versus protonated Cp149 capsids. Unassigned residues and prolines are shown as grey circles and residues without signal in the deuterated sample are shown as red diamonds. The noise was adjusted to the same level for both spectra in CcpNmr [136] before fitting the peak intensities and there was no apodisation applied to either spectrum. **(b)** Residues whose H^N could not be back-exchanged in the deuterated capsids are shown in orange for chain C and in red for chain D on the X-ray structure 1QGT [150].

```
           10          20          30          40          50
MDIDPYKEFG ATVELLSFLP SDFFPSVRDL LDTASALYRE ALESPEHCSP
           60          70          80          90          100
HHTALRQAIL CWGELMTLAT WVGVNLEDPA SRDLVVSYVN TNMGLKFRQL
           110         120         130         140         149
LWFHISCLTF GRETVIEYLV SFGVWIRTPP AYRPPNAPIL STLPETTVV
```

█ Assigned in ^{13}C- and ^{1}H-detected experiments

█ Assigned in ^{13}C-detected experiments only

Figure C.2. *Cp149 sequence showing the residues assigned in both ^{13}C- and ^{1}H-detected experiments (light blue) and in ^{13}C-detected experiments only (dark blue).* Residues with at least 2 assigned atoms are coloured, including prolines in ^{1}H-detected experiments for which carbon resonances could be assigned from the hCONH and hCAcoNH 3D spectra. Unassigned residues are not coloured.
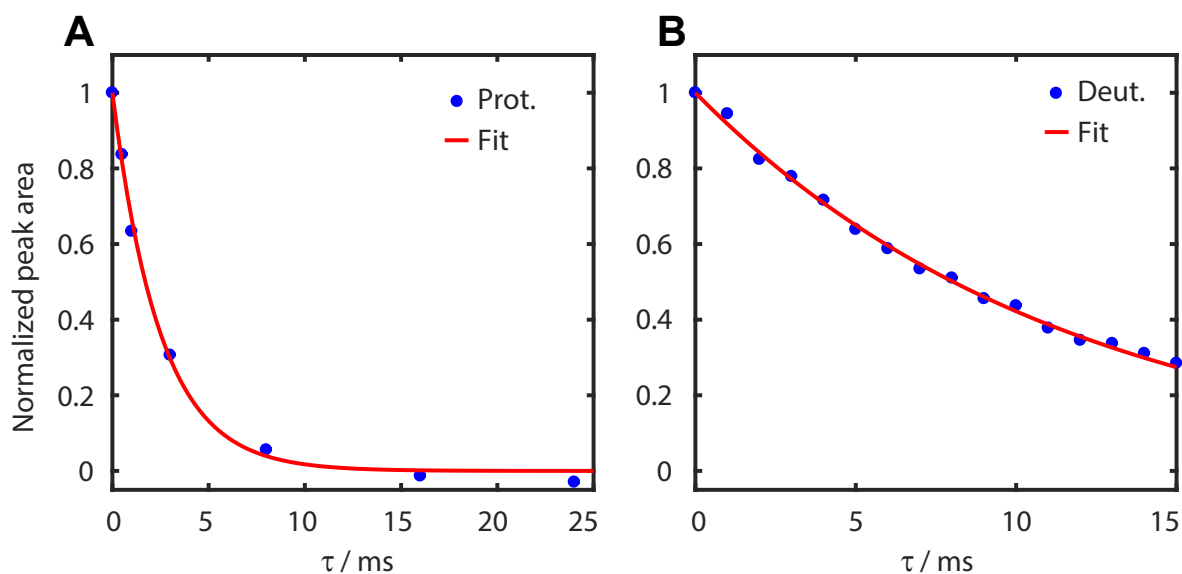

Figure C.3. *Bulk T_2' decays.* Decay **(a)** for the pCp149 capsids, fitted to a value of $T_2' = 2.5 \pm 0.1$ ms, and **(b)** for dCp149 capsids, fitted to a value of 11.6 ± 0.2 ms. Note that these values give the homogeneous, but not the full contribution to the signal decay for the two samples, as depicted in the main text in Figure 4.4C/D.
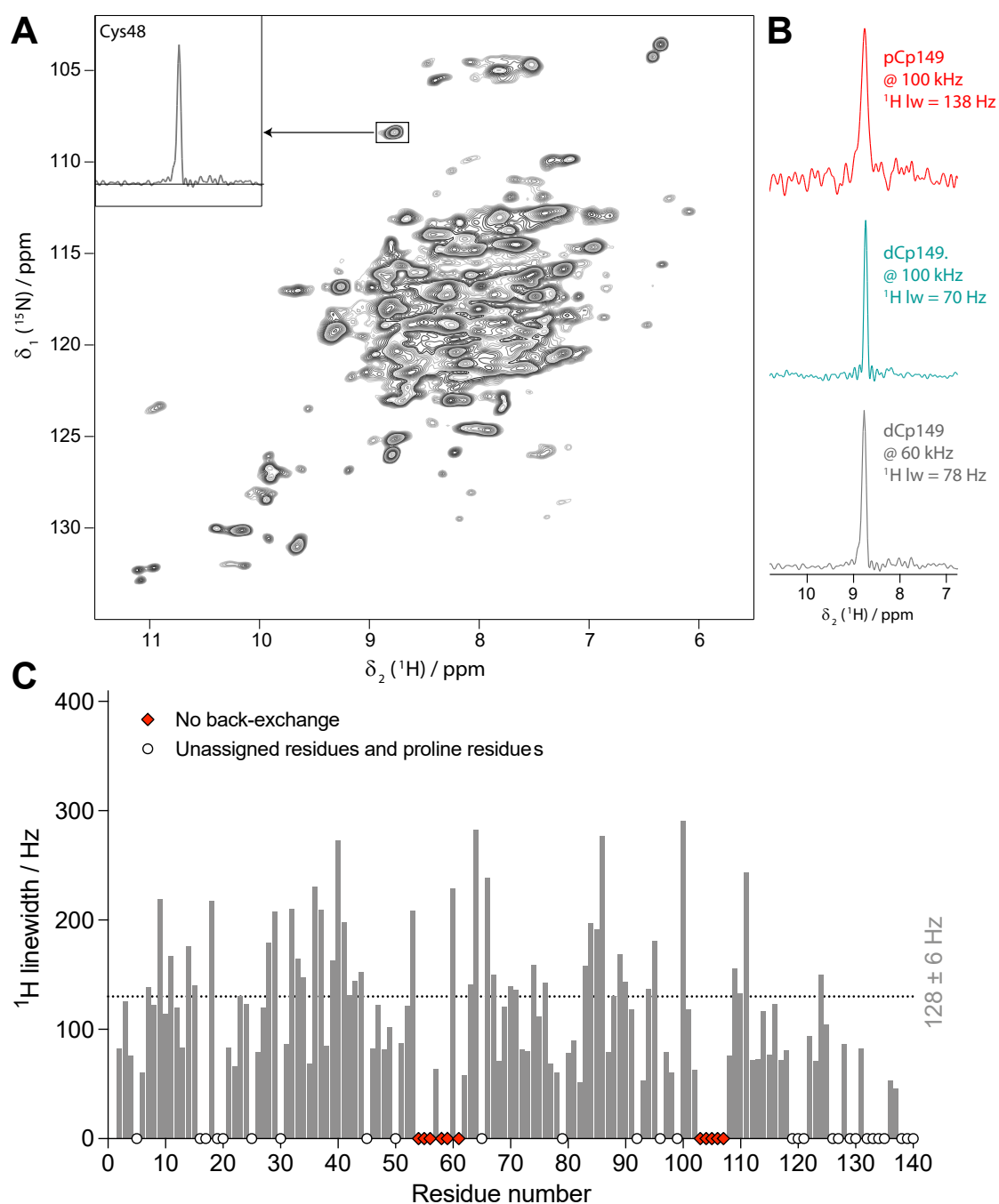
Figure C.4. **(a)** hNH spectrum recorded in a 1.3 mm rotor of dCp149 capsids, with the proton linewidth of the isolated C48 residue shown in the insert. The spectrum was recorded at 60 kHz MAS at a 16.4 T field with 80 scans, with similar acquisition and processing parameters as data recorded at 100 kHz MAS (11.2 h experimental time, 25 ms acquisition in the ^{15}N dimension, processed without apodisation function, cut at 12.9 ms acquisition in ^{1}H dimension and zero-filled to 4096 points (^{1}H) and 1024 points (^{15}N)). **(b)** Comparison of C48 ^{1}H linewidths for pCp149 at 100 kHz MAS (in red), dCp149 at 110 kHz (in cyan) and at 60 kHz MAS (in grey). **(c)** Total proton linewidths for assigned residues observed in the 3D hCANH spectra of dCp149 capsids at 60 kHz MAS using a parabolic fit in CcpNmr [136]. The 3D hCANH spectrum was run with a similar experimental time (22 h) as the hCANH spectra recorded at 100-110 kHz, and was processed identically. The mean linewidth is indicated as a dotted line. Unassigned and proline residues are indicated as white circles, and residues without back-exchange (invisible in the 3D) as red diamonds.
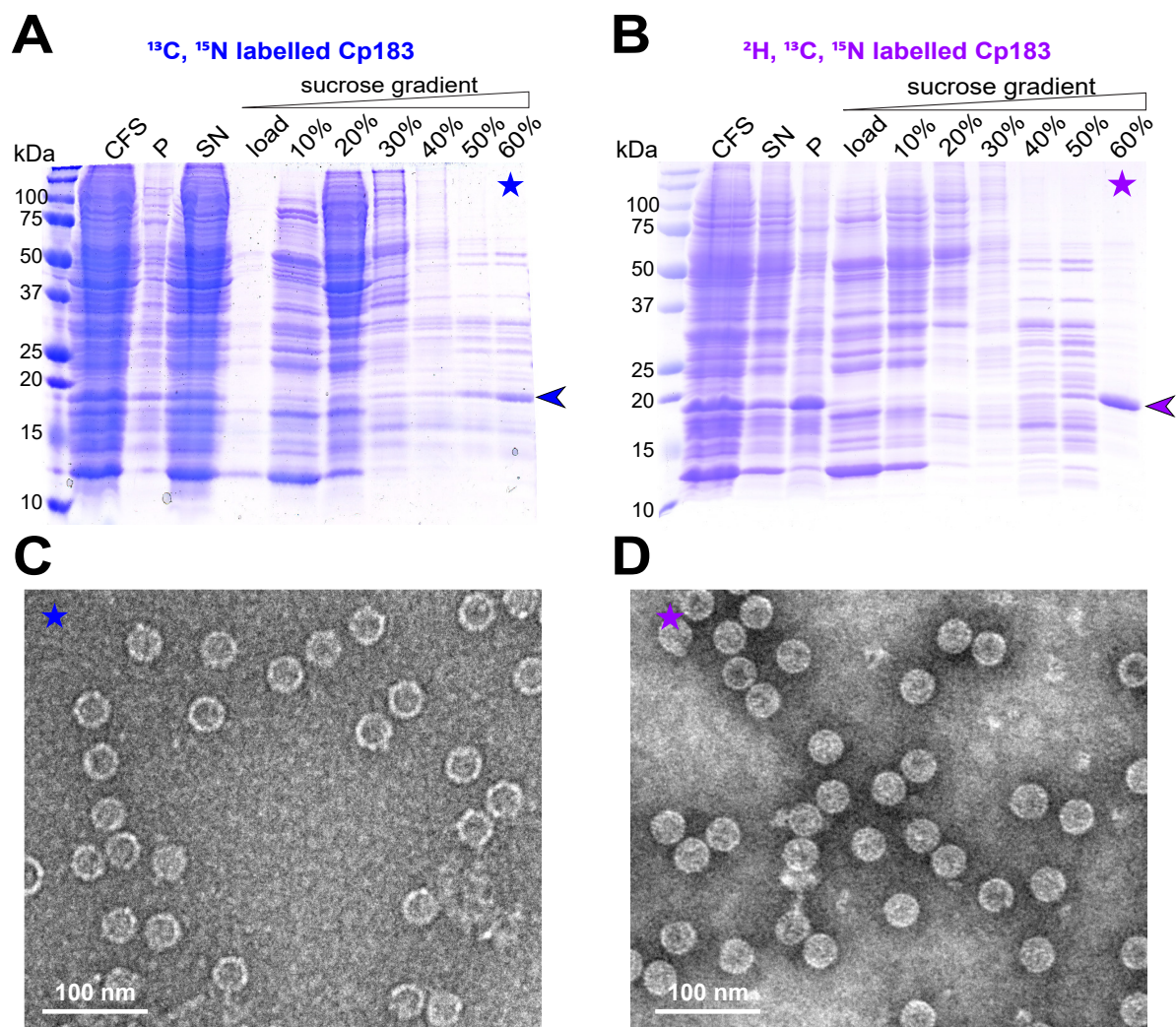
Figure C.5. *Analysis of CFPS purification of Cp183.* Sucrose gradient isolation on 15 % SDS-PAGE gels of **(a)** ^{13}C, ^{15}N and **(b)** ^{2}H, ^{13}C, ^{15}N isotopically labelled dCp183. CFS: total cell-free reaction mixture; P and SN: pellet and supernatant obtained after centrifugation of the CFS at 20,000 g, 4 °C for 30 min; 0-60 %: fractions from the sucrose gradient. Negative staining electron micrographs display **(c)** the ^{13}C, ^{15}N labelled and **(d)** ^{2}H, ^{13}C, ^{15}N labelled capsids from the 60 % sucrose fractions. Scale bar is 100 nm.
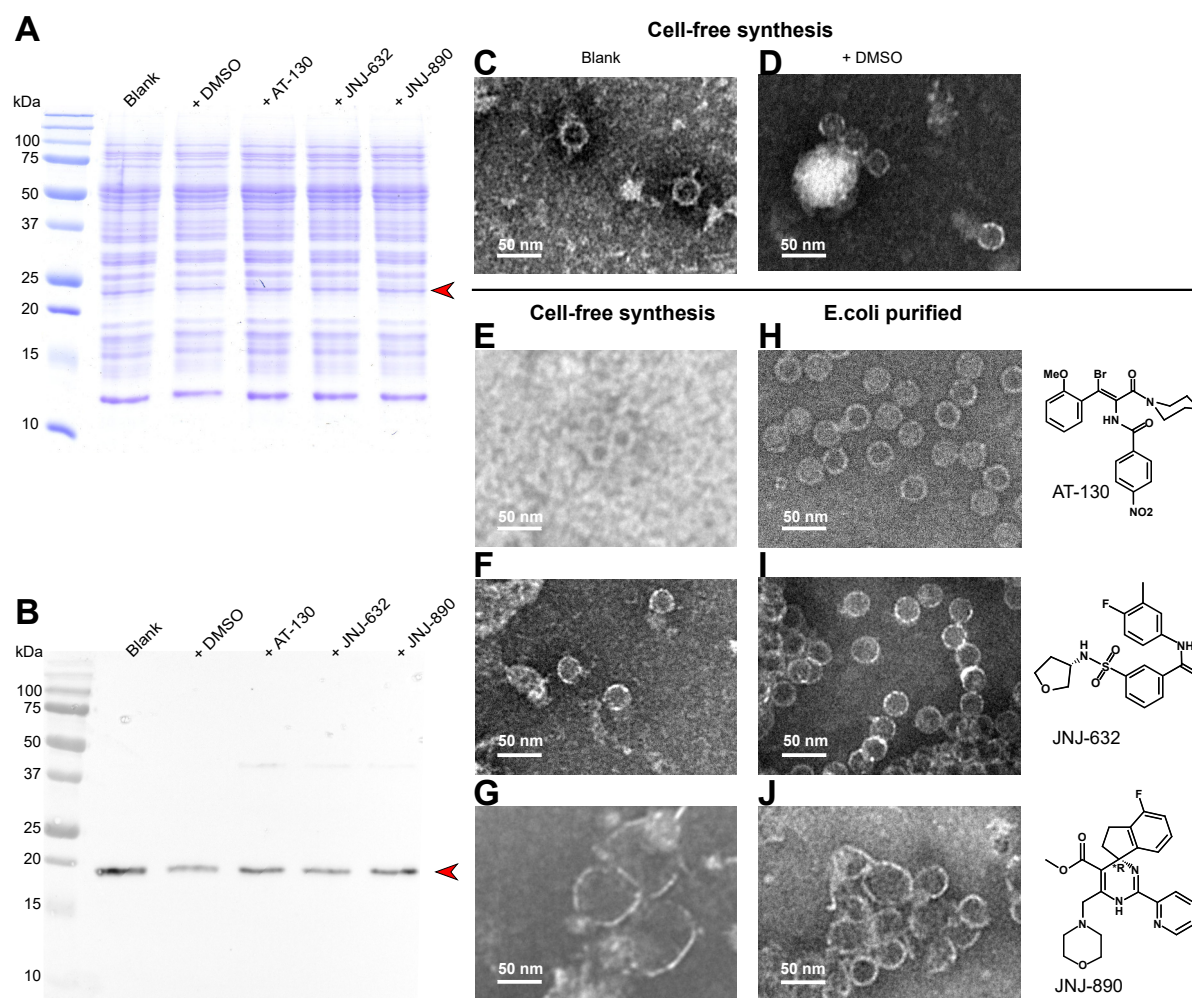
Figure C.6. *CFPS of Cp183 in presence of different compounds.* **(a)** 15 % SDS-PAGE and **(b)** Western blot analysis on Cp183 protein produced in 220 μL CFPS system with 10 nmol drugs dissolved in 1 μL DMSO. The lanes of two controls, indicated as blank and +DMSO, and productions with compounds are highlighted on the top. Electron microscopy (negative staining) analysis of cell-free products is shown for **(c)** the blank, **(d)** +DMSO, **(e)** +AT-130, **(f)** +JNJ-632, **(g)** +JNJ-890, as well as for capsids from *E. coli* incubated with **(h)** AT-130, **(i)** JNJ-632, **(j)** JNJ-890. Scale bar is 50 nm.

Figure C.7. *Individual spectra of Cp183.* For reference with respect to the overlays as presented in Figure 4.5.

Figure C.8. *Comparison of processing protocols with different signal apodisation.* **(a)** The SSB parameters used by the Bruker Topspin spectrometer software are given. SSB is used to calculate the apodisation function $f(t)$, according to $f(t) = \sin\left(\left(\pi - \frac{\pi}{\text{SSB}}\right) \cdot \f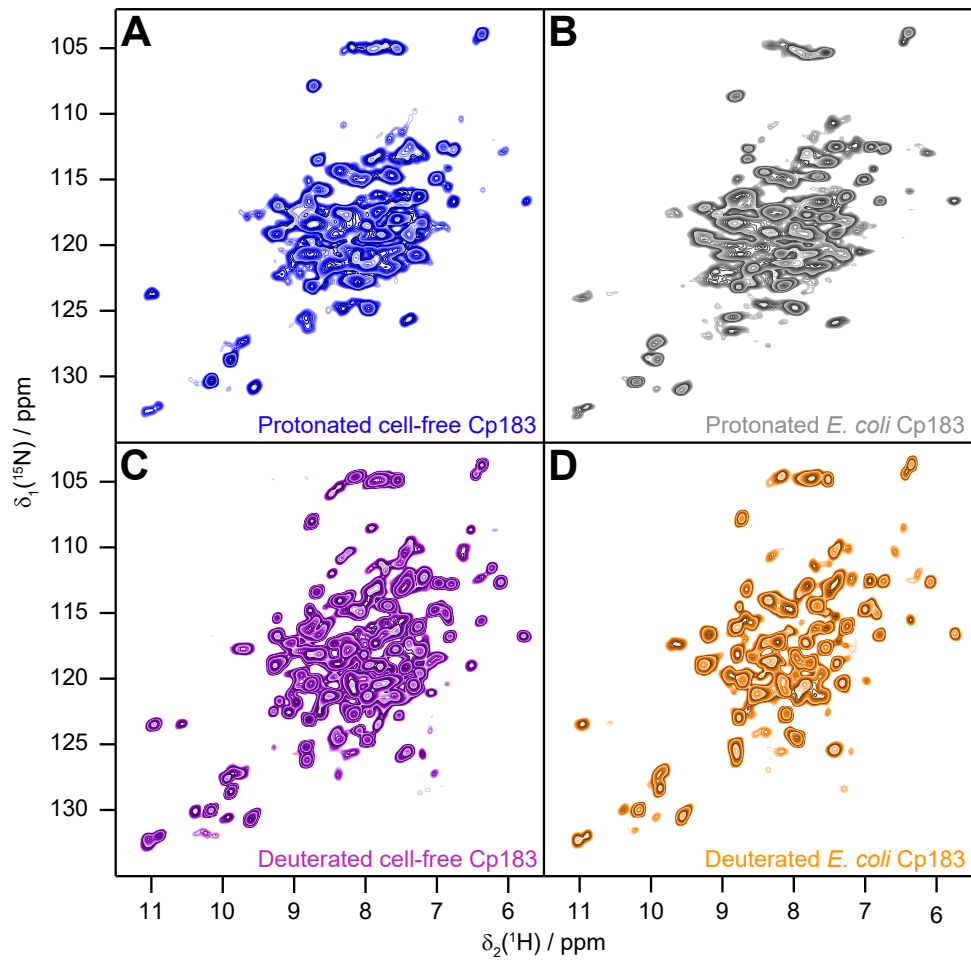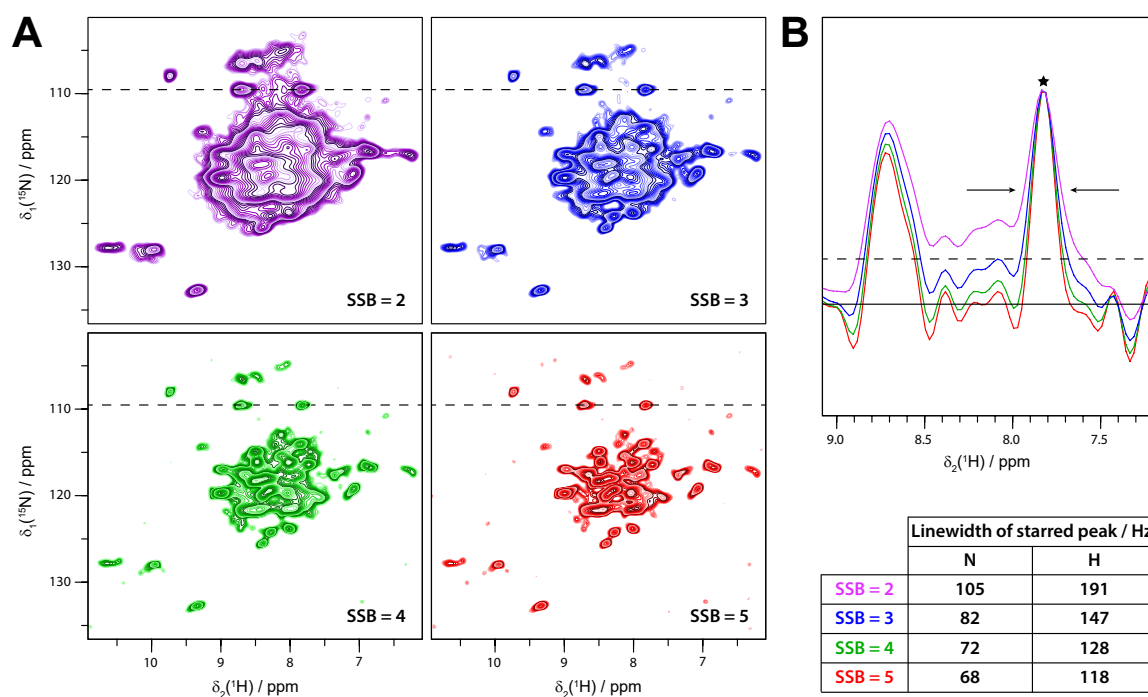rac{t}{\text{AQ}} + \frac{\pi}{\text{SSB}}\right)$, with AQ the acquisition time of the FID. A value of SSB = 2 corresponds to a pure cosine wave (90° shift). SSB = 3, 4, 5 correspond to 60°, 45°, and 36° shifts respectively. In the literature, sine functions shifted by 60° are often used [172, 233, 234], but shifts down to 30° (SSB = 6) have been reported for proton-detected spectra [260, 261], resulting in spectra with artificially higher resolution. In the main text, all linewidths are given for spectra apodised with a pure cosine function. **(b)** Influence of the apodisation function on the linewidth. 1D extracts at the dotted lines in the 2D spectra (109 ppm) are shown. Besides narrowing lines, strong apodisation also suppresses weak signals, which on the one hand gives the spectrum a visually cleaner look but on the other hand may lead to loss of information. The dotted line in the 1D extract indicates the base level plotted in the 2D spectra; if spectra are plotted at higher levels, they only show peak summits which result in additional cleaning of spectra. The table lists the different linewidths measured for the peak labelled with a star.
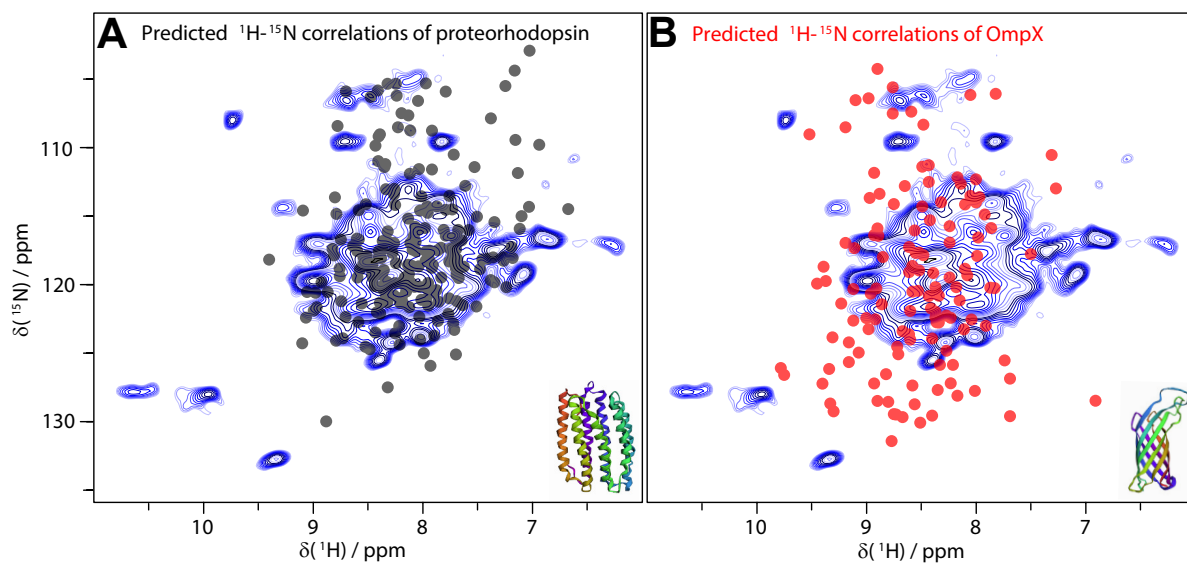
Figure C.9. *Comparison of the spectrum in Figure 4.6D and 4.7 with ^1H-^{15}N predictions from the BMRB.*
(a) Prediction for an entirely α-helical protein (proteorhodopsin, BMRB entry 17327 [173]) and for **(b)** an entirely β-sheet protein (OmpX, BMRB entry 18797 [174]). The good coincidence of the DHBs S spectrum with the typical α-helical resonances of proteorhodopsin supports the mainly α-helical fold from the circular dichroism analysis (reported in more detail in [138]).
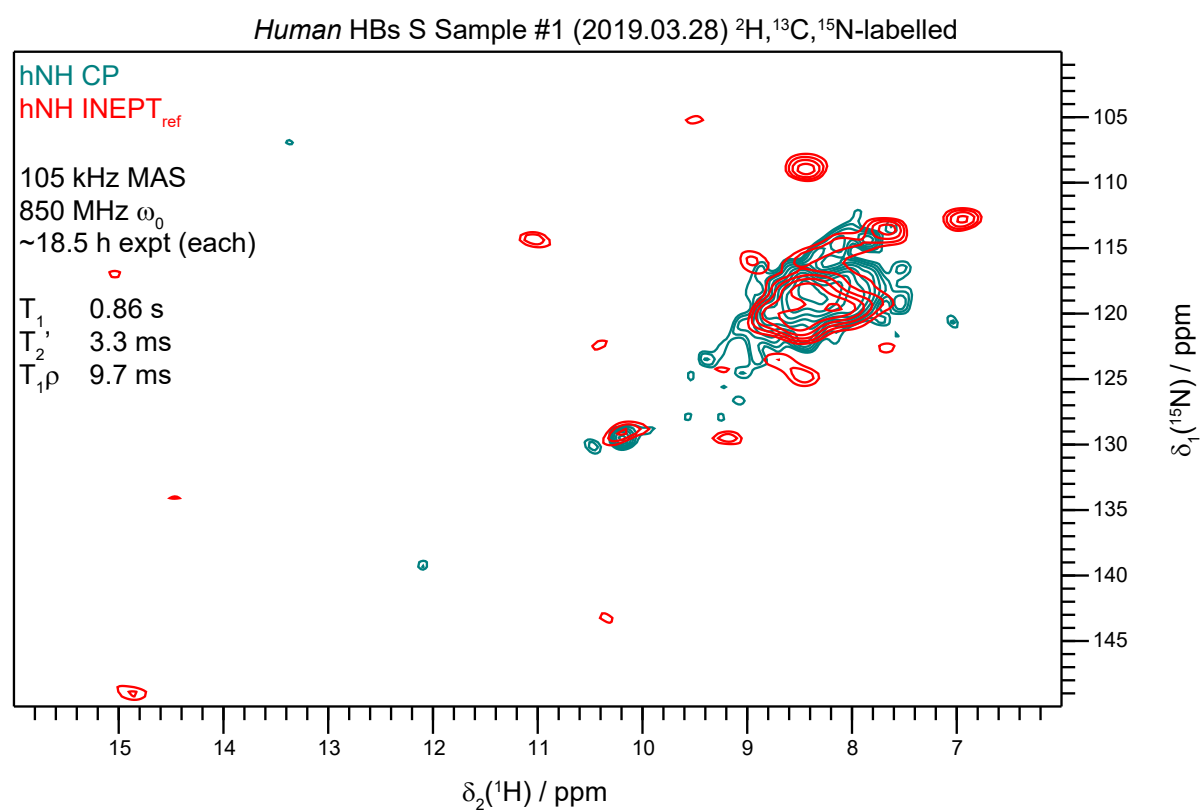
Figure C.10. *Preliminary spectra of human HBs S.* Two ^{15}N-^{1}H correlation spectra are shown, where in one (teal), magnetisation transfer is based on dipolar coupling, while in the other (orange) it is based on scalar coupling.

C.3. Tables

The tables presented below were taken from main texts and appendices of [138–140] on which Chapter 4 is based.

	Active volume / μL	Empty weight / mg	Full weight / mg	Total sample weight / mg	Mass ratio to 0.7 mm protonated
0.7 mm, protonated	0.37	6.10 ± 0.04	6.66 ± 0.01	0.56 ± 0.04	1
0.7 mm, deuterated	0.37	6.12 ± 0.01	6.67 ± 0.01	0.55 ± 0.01	1
1.3 mm, deuterated	1.45	35.6 ± 0.1	39.9 ± 0.1	4.3 ± 0.1	7.8
3.2 mm thin-wall, protonated	33	394 ± 2	449 ± 3	55 ± 3	98.2

**Table C.1.** *Rotors volumes and weights of Cp149 samples.* The active volume of each rotor was calculated using the inner rotor area based on the inner diameter as indicated by Bruker, multiplied by the coil length. For the 3.2 mm rotor weight, an average was calculated on 3 empty rotors and 4 full rotors filled with similar core protein samples which give a standard deviation of 3 mg. For the weight of 0.7 mm and 1.3 mm rotors, the original rotor samples were used. The mass ratios with respect to the protonated 0.7 mm rotors are given in the right column. All samples were filled using overnight ultracentrifugation at 200 000 g, and residual protein was observed in the funnel for all samples, indicating that the rotors were full. Note that the actual protein quantity required to fill each rotor is about half of the total sample weight indicated, as the sedimented capsids contains solvent. Typical starting amounts of capsids used to fill a rotor are in our hands ~30 mg for a thin-wall 3.2 mm rotor, ~2 mg for a 1.3 mm rotor and ~0.4 mg for a 0.7 mm rotor.

	pCp149					dCp149	
Experiment	CP hNH 2D	hCANH 3D	hCONH 3D	hCAcoNH 3D	hncaCBcaNH 3D	CP hNH 2D	hCANH 3D
MAS frequency / kHz	100	100	100	100	100	100	110
Field / T	20.0	20.0	20.0	20.0	20.0	20.0	20.0
t1 increments	776	64	42	82	112	776	64
Sweep width (t1) / ppm	180	30	16	30	55	180	30
Acquisition time (t1) / ms	25	5	6.1	6.4	4.8	25	5
t2 increments	1024	38	38	48	64	1024	43
Sweep width (t2) / ppm	46.7	35	35	35	35	46.7	40
Acquisition time (t2) / ms	13	6.3	6.3	8.0	10.6	13	6.2
t3 increments	-	1024	1024	1024	1024	-	1024
Sweep width (t3) / ppm	-	46.7	46.7	46.7	46.7	-	46.7
Acquisition time (t3) / ms	-	12.9	12.9	12.9	12.9	-	12.9
^{1}H dec. (swfTPPM) / kHz	10	10	10	10	10	10	10
^{15}N dec. (WALTZ64) / kHz	5	5	5	5	5	5	5
^{13}C dec (WALTZ64) / kHz	-	5	5	5	5	-	5
180° ref. pulse t1 / channel	^{13}C	^{15}N	^{15}N	^{15}N	^{15}N	-	^{15}N
180° ref. pulse t2 / channel	-	^{13}C	^{13}C	^{13}C	^{13}C	-	^{13}C
Water sup. (120 ms) / kHz	20	20	20	20	20	20	20
Interscan delay / s	1.2	2.0	1.0	1.3	1.3	1.2	1.1
Number of scans	40	16	128	80	48	40	24
Experiment time	12 h	23 h	2 days 18 h	5 days 10 h	5 days 22 h	12 h	23 h
Transfer I	HN (dip.c.)	HC (dip.c.)	HC (dip.c.)	HC (dip.c.)	HN (dip.c.)	HN (dip.c.)	HC (dip.c.)
^{1}H field / kHz	78	72	80	77	79	69	76
X field / kHz	17	16	15	15	15	15	16.5
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
Carrier ^{13}C / ppm (*)	-	56	174	56	-	-	51
Time / ms	1.0	1.2	0.8	1.0	1.0	2.5	4.4
Transfer II	NH (dip.c.)	CN (dip.c.)	CN (dip.c.)	CC (DREAM)	NCA (sc.c.)	NH (dip.c.)	CN (dip.c.)
^{1}H field / kHz	78	-	-	-	-	69	-
^{13}C field / kHz	-	71	66	49	-	-	63
^{15}N field / kHz	17	25	35	-	-	15	36
Shape	Tangent ^{1}H	Tangent ^{13}C	Tangent ^{13}C	Tangent ^{13}C	-	Tangent ^{1}H	Tangent ^{13}C
Carrier ^{13}C / ppm	-	56	174	130	42	-	51
Time / ms	1.0	10.0	20.0	9.0	22	2.5	18.0
Transfer III		NH (dip.c.)	NH (dip.c.)	CN (dip.c.)	CACB (sc.c.)		NH (dip.c.)
^{1}H field / kHz		74	84	-	-		74
^{13}C field / kHz		-	-	60	-		-
^{15}N field / kHz		15	15	36	-		16
Shape		Tangent ^{1}H	Tangent ^{1}H	Tangent ^{13}C	-		Tangent ^{1}H
Carrier ^{13}C / ppm		-	-	174	42		-
Time / ms		1.0	1.6	12.0	7		2.5
Transfer IV				NH (dip.c.)	CBCA (sc.c.)		
^{1}H field / kHz				77	-		
^{15}N field / kHz				15	-		
Shape				Tangent ^{1}H	-		
Carrier ^{13}C / ppm				-	42		
Time / ms				1.2	7		
Transfer V					CN (sc.c.)		
Carrier ^{13}C / ppm					42		
Time / ms					22		
Transfer VI					NH (dip.c.)		
^{1}H field / kHz					79		
^{15}N field / kHz					15		
Shape					Tangent ^{1}H		
Time / ms					1.0		

(*) Carriers for ^{1}H and ^{15}N were 4.8 and 117.5 ppm respectively, for all experiments.

Table C.2. *Cp149 experimental parameters for NMR spectroscopy ≥ 100 kHz MAS.* (dip.c.) is written for dipolar-coupling based transfers and (sc.c.) for scalar-coupling based transfers of magnetisation.

	dCp149	
Experiment	**CP hNH 2D**	**hCANH 3D**
MAS frequency / kHz	60	60
Field / T	16.4	16.4
t1 increments	248	58
Sweep width (t1) / ppm	70	30
Acquisition time (t1) / ms	25	5.4
t2 increments	1800	40
Sweep width (t2) / ppm	99.2	40
Acquisition time (t2) / ms	13	7
t3 increments	-	1800
Sweep width (t3) / ppm	-	99.2
Acquisition time (t3) / ms	-	13
^{1}H dec. (WALTZ16) / kHz	10	10
^{15}N dec. (WALTZ16) / kHz	5	9
^{13}C dec. (WALTZ16) / kHz	10	10
180° ref. pulse t1 / channel	^{13}C	^{15}N
10 kHz WALTZ16 t2 / channel	-	^{13}C
Water sup. (100 ms) / kHz	15	15
Interscan delay / s	2	1
Number of scans	80	32
Experiment time	11.2 h	23 h
Transfer I	**HN (dipolar)**	**HC (dipolar)**
^{1}H field / kHz	100	85
X field / kHz	40	50
Shape	Ramp ^{1}H	Ramp ^{1}H
Carrier ^{13}C / ppm (*)	-	54
Time / ms	0.6	5
Transfer II	**NH (dipolar)**	**CN (dipolar)**
^{1}H field / kHz	100	-
^{13}C field / kHz	-	34
^{15}N field / kHz	43.8	26
Shape	Ramp ^{1}H	Tangent ^{15}N
Carrier ^{13}C / ppm	-	54
Time / ms	0.6	7
Transfer III		**NH (dipolar)**
^{1}H field / kHz		95
^{13}C field / kHz		-
^{15}N field / kHz		44
Shape		Ramp ^{1}H
Carrier ^{13}C / ppm		-
Time / ms		0.6

(*) Carriers for ^{1}H and ^{15}N were 4.8 and 120 ppm respectively.

Table C.3. *Cp149 experimental parameters for NMR spectroscopy at 60 kHz MAS.* Decoupling is abbreviated as dec.

Sample production	*E. coli*	*E. coli*	Cell-free	Cell-free	
Sample labeling	^{2}H ^{13}C ^{15}N	^{1}H ^{13}C ^{15}N	^{2}H ^{13}C ^{15}N	^{1}H ^{13}C ^{15}N	
Experiment	**hNH 2D**	**hNH 2D**	**hNH 2D**	**hNH 2D**	**hCANH 3D**
MAS frequency / kHz	100	100	100	100	100
Field / T	20	20	20	20	20
t1 increments	284	774	620	330	104
Sweep width (t1) / ppm	180	180	180	180	36
Acquisition time (t1) / ms	9.2	25	20	11	6.8
t2 increments	5550	5550	2048	2048	44
Sweep width (t2) / ppm	47	47	47	40	34
Acquisition time (t2) / ms	70	70	26	30	7.5
t3 increments	-	-	-	-	2048
Sweep width (t3) / ppm	-	-	-	-	47
Acquisition time (t3) / ms	-	-	-	-	26
Proton decoupling (swfTPPM) / kHz	11	12	10	10	10
Nitrogen decoupling (WALTZ64) / kHz	6	6	5	5	5
Carbon decoupling (WALTZ64) / kHz	-	-	-	-	5
Water suppression (120 ms MISS.) / kHz	22	24	20	20	20
Interscan delay / s	2.0	1.2	2.0	1.2	1.2
Number of scans	32	32	64	128	64
Measurement time / dd:hh:mm	00:15:34	00:09:50	00:16:41	00:15:59	04:14:32
Carrier ^{1}H / ppm	4.8	4.8	4.8	4.8	4.8
Carrier ^{15}N / ppm	117.5	117.5	117.5	117.5	117.5
Transfer I	**HN CP**	**HN CP**	**HN CP**	**HN CP**	**HC CP**
^{1}H field / kHz	72	77	78	77	82
X field / kHz	17	16	15	15	15
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
Carrier ^{13}C / ppm	-	-	-	-	56
Time / ms	1.0	1.0	1.8	1.6	0.6
Transfer II	**NH CP**	**NH CP**	**NH CP**	**NH CP**	**CN CP**
^{1}H field / kHz	72	77	78	77	-
^{13}C field / kHz	-	-	-	-	63
^{15}N field / kHz	17	16	15	15	35
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{13}C
Carrier ^{13}C / ppm	-	-	-	-	56
Time / ms	1.0	1.0	2.4	1.6	20.0
Transfer III	**-**	**-**	**-**	**-**	**NH CP**
^{1}H field / kHz	-	-	-	-	78
^{15}N field / kHz	-	-	-	-	15
Shape	-	-	-	-	Tangent ^{1}H
Time / ms	-	-	-	-	H.4

Table C.4. Cp183 experimental parameters for NMR spectroscopy

Sample	#1 (A)	#2 (B)	#3 (C)	#4 (D)
Relaxation parameters bulk protein				
T_1 / s	0.77 ± 0.04	0.70 ± 0.09	0.94 ± 0.02	0.83 ± 0.01
T_2' / ms	5.3 ± 1.3	4.0 ± 3.0	3.3 ± 0.7	3.9 ± 0.6
$(\pi \cdot T_2')^{-1}$ / Hz	60 ± 15	80 ± 40	100 ± 30	82 ± 13
$T_1\rho$ / ms	14 ± 2	9 ± 2	11 ± 2	18 ± 3
Experiment	**CP hNH 2D**	**CP hNH 2D**	**CP hNH 2D**	**CP hNH 2D**
MAS frequency / kHz	100	100	100	110
Field / T	20	20	20	20
t1 increments	256	512	256	420
Sweep width (t1) / ppm	120	120	180	180
Acquisition time (t1) / ms	12	25	8	14
t2 increments	5550	5550	5550	5550
Sweep width (t2) / ppm	46.7	46.7	46.7	46.7
Acquisition time (t2) / ms	70	70	70	70
Proton decoupling (swfTPPM) / kHz	10	10	10	10
Nitrogen decoupling (WALTZ64) / kHz	5	5	5	5
Water suppression (120 ms MISS.) / kHz	20	20	20	20
Interscan delay / s	1.1	1.0	1.2	1.2
Number of scans	208	94	112	112
Measurement time / hh:mm	19:36	15:51	11:09	18:19
SNR ("bulk") in a reference 1D (128 scans)	13	10	12	32
SNR of isolated peaks in 2D	4.2 ± 0.3	3.4 ± 0.4	3.2 ± 0.3	7.6 ± 0.6
Mean ^{1}H linewidth of 2D isolated peaks / Hz	270 ± 70	280 ± 80	150 ± 40	170 ± 40
Mean ^{15}N linewidth of 2D isolated peaks / Hz	130 ± 50	70 ± 10	140 ± 30	90 ± 10
Transfer I	**HN-CP**	**HN-CP**	**HN-CP**	**HN-CP**
^{1}H field / kHz	74	78	82	91
^{15}N field / kHz	17	15	15	15
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
^{15}N carrier / ppm	117.5	117.5	117.5	117.5
Time / ms	1.0	1.0	1.0	1.2
Transfer II	**NH-CP**	**NH-CP**	**NH-CP**	**NH-CP**
^{1}H field / kHz	74	78	82	91
^{15}N field / kHz	17	15	15	15
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
^{1}H carrier / ppm	4.8	4.8	4.8	4.8
Time / ms	1.0	1.5	1.8	1.2

Table C.5. *HBs S experimental parameters for NMR spectroscopy.*

D. Supplementary information to nucleotide binding modes in *Hp*DnaB

D.1. Figures

The figures in this appendix were also presented in the supporting documentation to [176], the publication presented in Chapter 5.
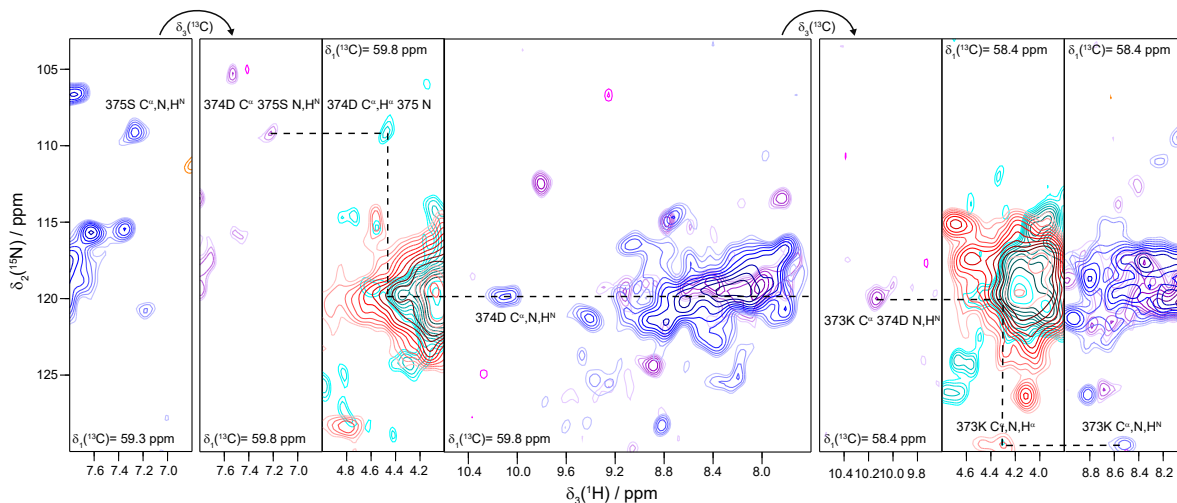
Figure D.1. *Assignment ambiguities in* ^{1}*H-detected spectra.* Example for a sequential walk along residues 375S-373K using the assignment scheme presented in Figure 5.3. Assignment ambiguities are resolved by additionally using the information from ^{13}C-detected assignment experiments [209].
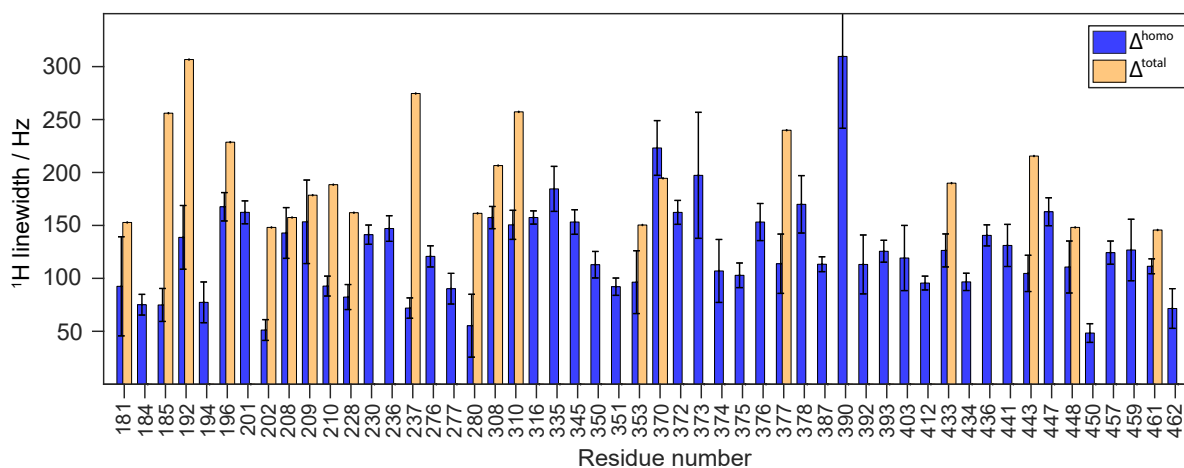
Figure D.2. ^{1}*H linewidths in solid-state NMR spectra of DnaB are dominated by homogeneous contributions.* Comparison of site-specific homogeneous contributions (blue bars) *versus* total experimental linewidths for a set of isolated resonances (orange bars).
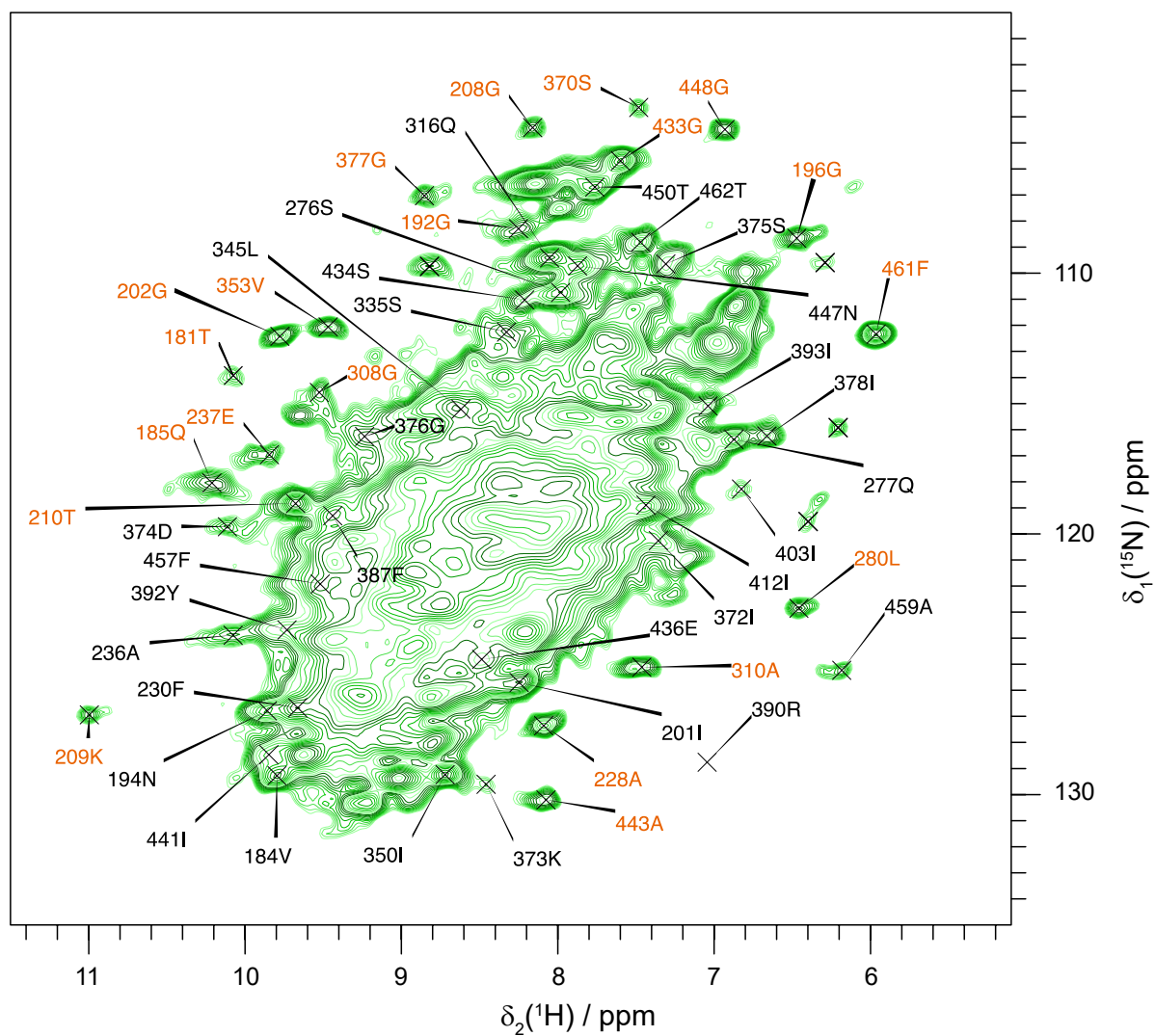
Figure D.3. *Reference 2D hNH spectrum and assignment used for the T_2' analysis.* Orange peaks indicate those peaks for which a total linewidth could be extracted.
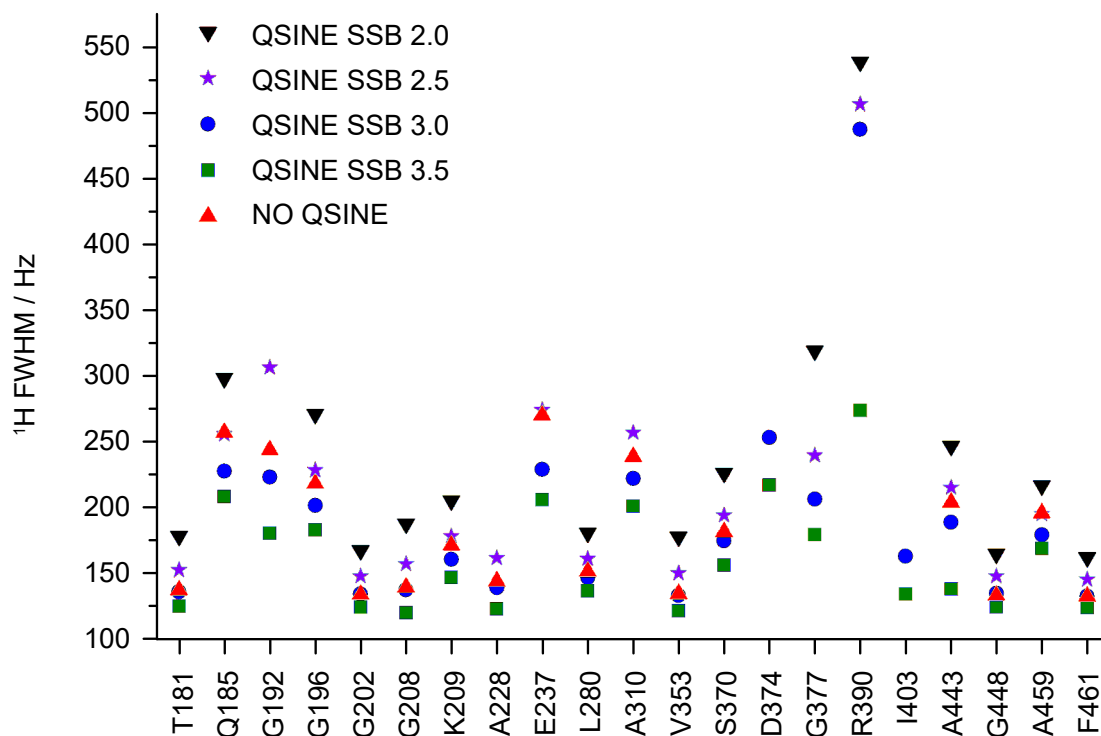
Figure D.4. *Signal processing influences ^1H linewidths extracted from 2D spectra.* ^{1}H FWHM for a selection of isolated peaks taken from the hNH spectrum shown in Figure 5.6A. For the determination of the inhomogeneous contribution to the linewidth a QSINE window function (SSB 2.5) was used which is for most residues very close to the value without window function (NO QSINE). Note that some residues have a larger (nonphysical) FWHM than 575 Hz most likely due to peak overlap and are thus not shown here. The spectrum in Figure 5.3A is processed with a QSINE (SSB 3.0) window function.
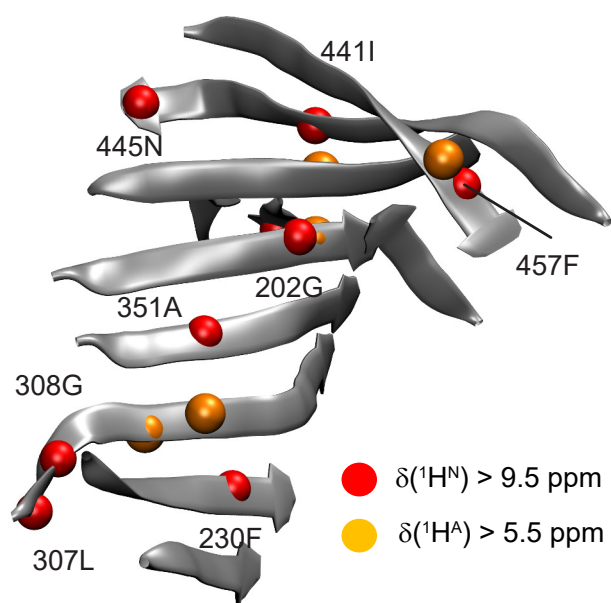
Figure D.5. *Deshielded H^N and H$^\alpha$ resonances form a band through the central β-barrel.* Residues with deshielded ^{1}H resonances (see legend) plotted on the central half-β barrel (taken from a DnaB model [177]). The residues form a network through this structural element.
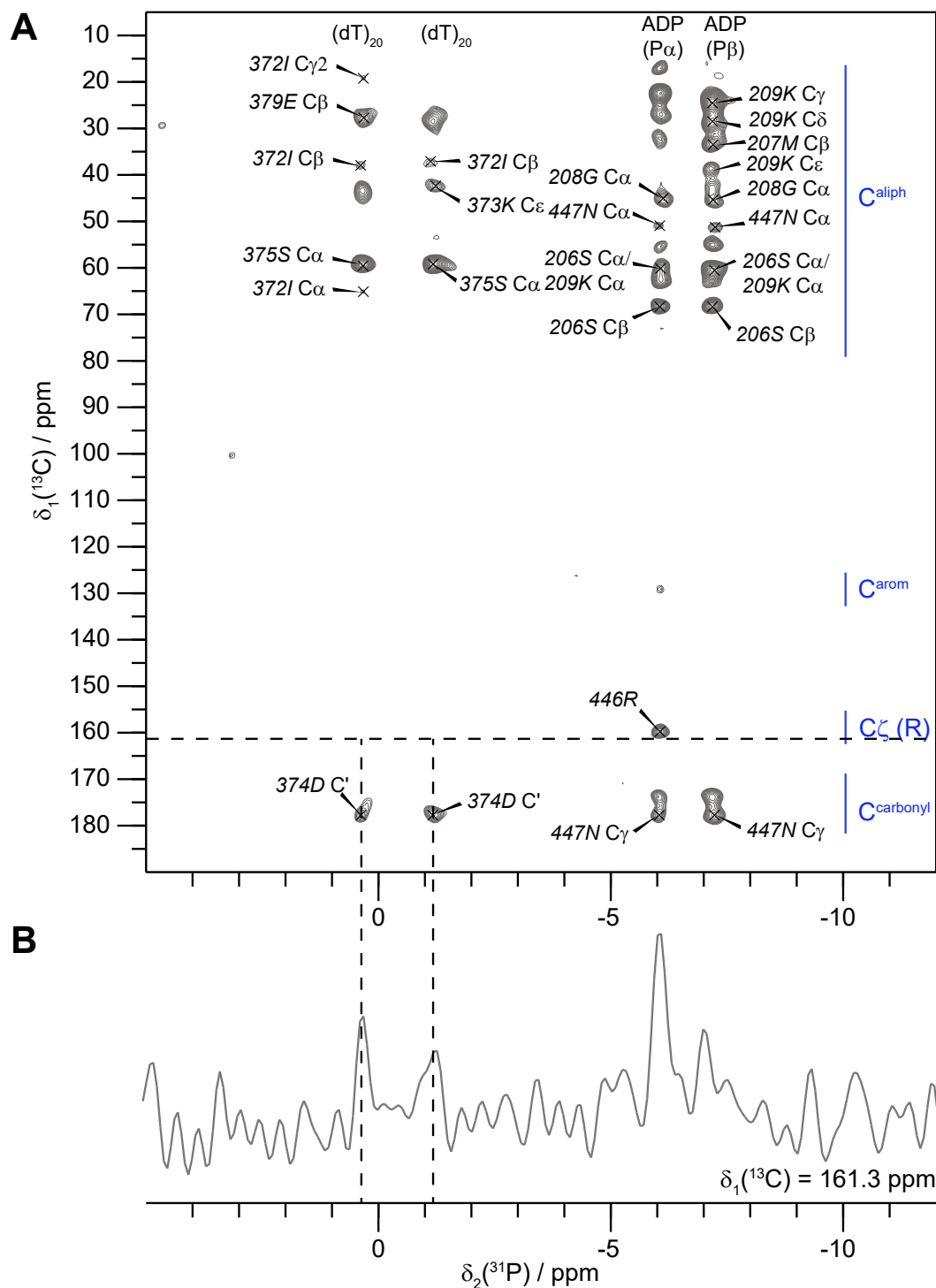
Figure D.6. *CHHP to probe protein-nucleotide contacts.* **(a)** CHHP spectrum of DnaB:ADP:AlF$_4^-$:ssDNA. Ambiguous assignments are shown in italic.**(b)** Slice along the direct dimension of the CHHP 2D spectrum at $\delta(^{13}C)$ = 161.3 ppm (357R Cζ resonance). A very weak DNA-arginine correlation appears which is significantly weaker than the one observed for 446R indicating that 357R is not in very close spatial proximity to the DNA phosphate groups.
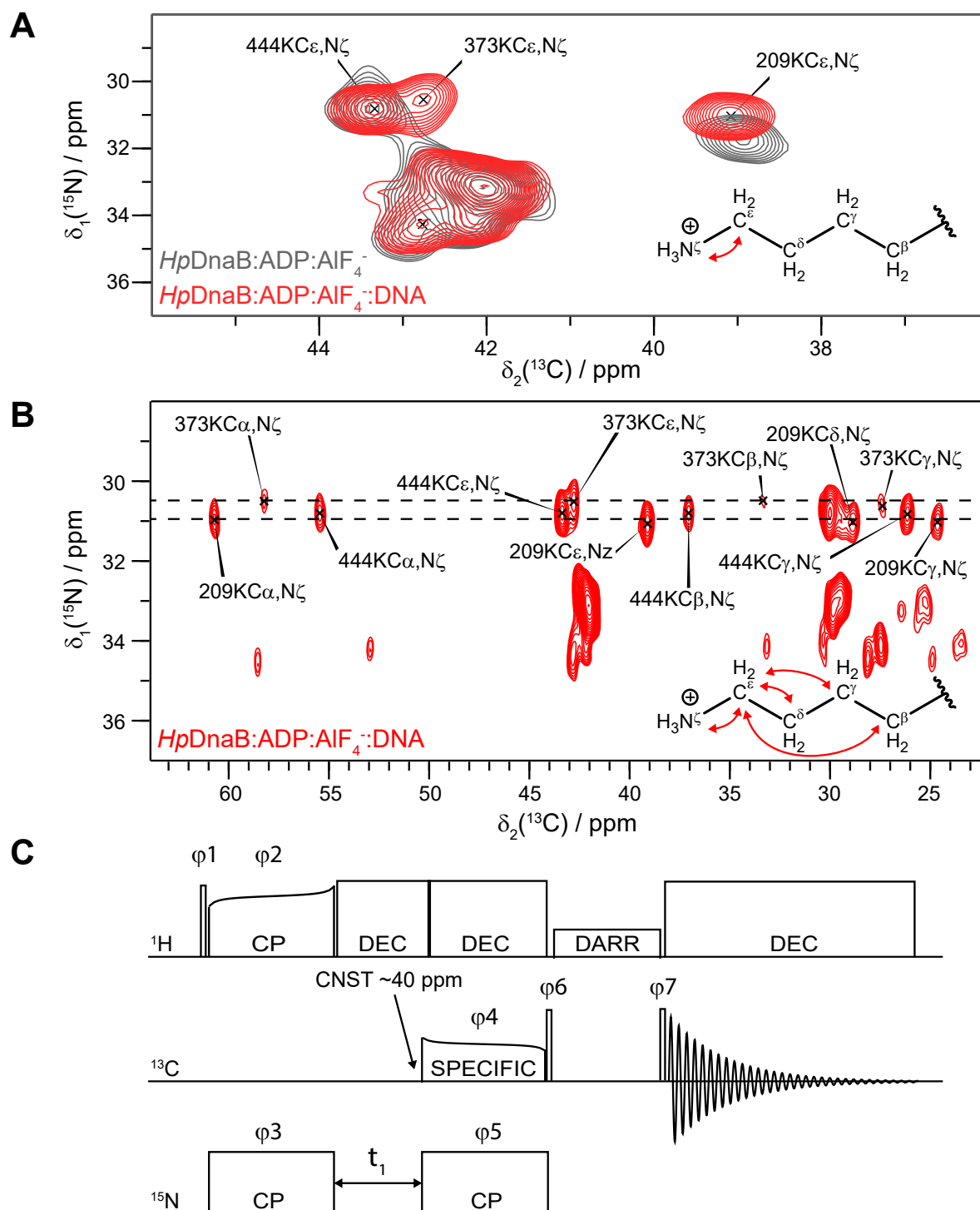
Figure D.7. *Lysine 373K rigidifies upon DNA binding and coordinates to the DNA.* **(a)** ^{15}N-^{13}C correlation spectra for the lysine sidechain region of DnaB:ADP:AlF$_4^-$ and DnaB:ADP:AlF$_4^-$:ssDNA (the latter spectrum is from [209]). **(b)** N-CC correlation spectrum (for the pulse sequence see **(c)**) allowing to unambiguously assign 373K. The following phases were used:

ϕ1= y −y, ϕ2 = x, ϕ3 = x, ϕ4 = x x x x −x −x −x −x y y y y −y −y −y −y,
ϕ5= x x −x −x, ϕ6= y y y y −y −y −y −y −x −x −x −x x x x x,
ϕ7= −y −y −y −y y y y y x x x x −x −x −x −x,
ϕrec= x −x −x x −x x x −x y −y −y y −y y y −y.
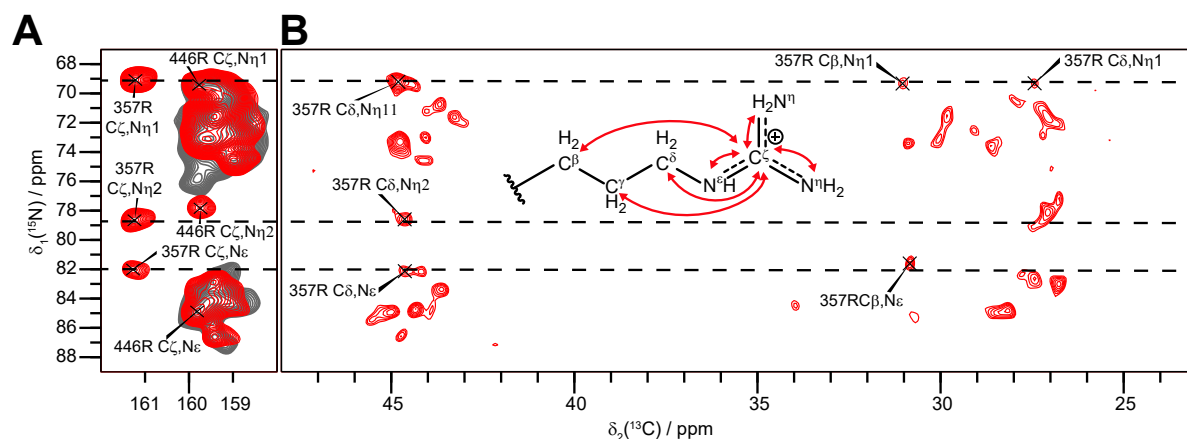
Figure D.8. *Arginine 357R rigidifies upon DNA binding and is not located in the NBD.* **(a)** ^{15}N-^{13}C correlation spectra for the arginine sidechain region of DnaB:ADP:AlF$_4^-$ (shown in grey) and DnaB:ADP:AlF$_4^-$:ssDNA (red). The red spectrum is taken from [209]. **(b)** N-CC correlation spectrum allowing to unambiguously assign 357R (for the pulse sequence see Appendix Figure D.7C with the SPECIFIC CP step optimised for the arginine sidechain resonances).
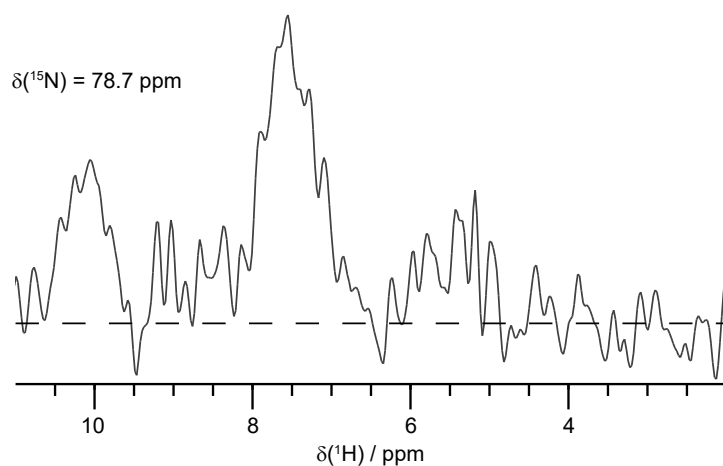
Figure D.9. *Arginine 357R is involved in hydrogen bonding via its sidechain.* 1D trace along the direct dimension of a 2D hNH spectrum of DnaB:ADP:AlF$_4^-$:ssDNA recorded at 110 kHz.

D.2. Tables

The tables in this appendix were also presented in the supporting documentation to [176], the publication presented in Chapter 5.

Residue		$\delta(H^N)$ / ppm	$\delta(H^\alpha)$ / ppm	Residue		$\delta(H^N)$ / ppm	$\delta(H^\alpha)$ / ppm
177	Thr	8.3	-	261	Asp	9	4.9
179	Ile	8.4	4.1	262	Asp	8.4	-
180	Pro	-	-	275	Leu	-	4.3
181	Thr	10.1	4	276	Ser	8	3.9
182	Gly	8.2	3.2, 4.1	277	Gln	6.9	-
183	Phe	7.6	4.7	278	Lys	8.2	4.4
184	Val	9.8	3.7	279	Lys	8.9	4.1
185	Gln	10.2	-	280	Leu	6.5	4.8
188	Asn	8	4.4	282	Phe	8.7	-
189	Tyr	7.9	4.3	285	Lys	7.9	4
190	Thr	8.2	4.3	286	Ser	7.6	4.2
191	Ser	8.9	4.1	287	Tyr	7.4	5.6
192	Gly	8.3	3.1, 3.6	307	Leu	9.7	3.8
193	Phe	9	5.4	308	Gly	9.6	3.7, 4.9
194	Asn	9.9	4.8	309	Ile	7.4	4.7
195	Lys	8.6	4.5	310	Ala	7.5	6
196	Gly	6.5	3.2, 4.2	311	Phe	9	5.9
197	Ser	7.3	4.6	312	Ile	9.3	5
198	Leu	8.2	5.3	313	Asp	9.2	5.5
199	Val	9.3	4.8	314	Tyr	6	5.1
200	Ile	8.8	5.5	315	Leu	9.4	3.8
201	Ile	8.2	5.2	316	Gln	8	-
202	Gly	9.8	6.1, 3.3	332	Ala	8.2	3.9
203	Ala	7.7	4.6	333	Glu	7.6	-
204	Arg	8.3	-	334	Ile	7.6	3.5
205	Pro	-	4.8	335	Ser	8.3	3.8
206	Ser	9.4	4.6	344	Glu	7.6	4
207	Met	7.5	4.5	345	Leu	8.6	4.2
208	Gly	8.2	-	346	Glu	7.8	3.8
209	Lys	11	3.8	347	Ile	8	-
210	Thr	9.7	4.4	348	Pro	-	5.3
211	Ser	7.7	4.7	349	Ile	8.3	4.4
213	Met	9	3.8	350	Ile	8.7	4.9
214	Met	7.8	4.6	351	Ala	9.6	5.3
215	Asn	8.4	4.5	352	Leu	8.8	5.5
249	Ser	7.6	-	353	Val	9.5	5.4

Table D.1. *Summary of H^N and H^α chemical-shift values of DnaB:ADP:AlF$_4^-$:DNA.* The values are graphically summarised in Figure 5.4C. Given values have a standard deviation of ± 0.1 ppm.

Residue		$\delta(H^N)$ / ppm	$\delta(H^\alpha)$ / ppm	Residue		$\delta(H^N)$ / ppm	$\delta(H^\alpha)$ / ppm
354	Gln	8.5	-	438	Ala	9.3	5.5
358	Ser	8.7	4	439	Glu	-	4.6
359	Leu	8.8	-	440	Ile	8.9	4.5
370	Ser	7.4	4	441	Ile	9.9	4
371	Asp	8.1	4.6	442	Val	9	4.1
372	Ile	7.3	3.7	443	Ala	8.1	4.2
373	Lys	8.5	4.3	444	Lys	7.6	4.4
374	Asp	10.1	4.5	445	Asn	10.1	4.7
375	Ser	7.2	4.8	446	Arg	8.9	3.9
376	Gly	9.3	3.5, 5.4	447	Asn	7.9	5.4
377	Gly	8.9	3.6, 4.0	448	Gly	7	3.2, 3.6
378	Ile	6.7	4.1	449	Ala	8.5	3.5
379	Glu	8.6	-	450	Thr	7.8	4.1
381	Asp	8.5	-	451	Gly	7.6	3.8, 4.4
382	Ala	7.5	3.7	452	Thr	8.3	4.5
383	Asp	9.5	4.5	453	Val	9.2	4.2
384	Ile	7.3	4.7	454	Tyr	8.8	4.8
385	Val	8.6	4.6	455	Thr	8.8	5.1
386	Leu	9.3	5.4	456	Arg	9.1	5.6
387	Phe	9.4	6.5	457	Phe	9.5	4.7
388	Leu	8	5.1	458	Asn	9	4.1
389	Tyr	8.7	4.8	459	Ala	6.3	2.4
390	Arg	7.1	4.7	461	Phe	6	5
391	Gly	-	3.4, 3.9	462	Thr	7.5	3
392	Tyr	9.8	4.3	463	Arg	6.7	-
393	Ile	7	3.4				
403	Ile	6.8	4.2				
404	Asp	8.8	4.1				
406	Leu	8.9	-				
412	Ile	7.4	3.6				
432	Asn	8	4.8				
433	Gly	7.6	4.2				
434	Ser	8.2	-				
435	Ile	7.8	4.4				
436	Glu	8.5	4.8				
437	Glu	9.1	4.8				

Table D.1 (Continued) *Summary of H^N and H^α chemical-shift values of DnaB:ADP:AlF$_4^-$:DNA.* The values are graphically summarised in Figure 5.4C. Given values have a standard deviation of ± 0.1 ppm.

Residue	$\delta(H^N)$ / ppm DnaB:ADP:AlF$_4^-$:DNA	$\delta(H^N)$ / ppm DnaB:AMP-PCP:DNA	$\delta(H^N)$ / ppm Apo DnaB
209K	11.1	flexible	flexible
185Q	10.2	10.4	10.4
236A	10.2	H absent	flexible
374D	10.1	10.2	flexible
181T	10.1	10.0	flexible
445N	10.1	9.5	9.0
194N	9.9	9.9	9.8
237E	9.9	9.7	H absent
441I	9.9	9.8	9.8
184V	9.8	9.9	9.8
202G	9.8	9.7	9.7
392Y	9.8	9.8	9.0
210T	9.7	9.8	H absent
307L	9.7	9.7	flexible
230F	9.7	9.8	flexible
351A	9.6	9.6	9.6
308G	9.6	9.6	9.7
457F	9.5	9.5	9.0
353V	9.5	9.4	H absent
383D	9.5	9.2	7.6
206S	9.4	flexible	flexible

Table D.2. *Summary of deshielded H^N resonances.* Residues with H^N chemical shifts > 9.4 ppm are given. Residues highlighted in **green** are part of the central β-barrel. Resonances classified as *flexible* are not visible in the spectra, most likely due to dynamics. Resonances classified as *H absent* are visible in ^{13}C-detected spectra, but remain unassigned in the ^{1}H dimension. Chemical-shift values are given with a standard deviation of ± 0.1 ppm.

Experiment	NHHP	CHHP
MAS frequency / kHz	17	17
Field / T	11.7	11.7
Transfer I	**HN-CP**	**HC-CP**
^{1}H field / kHz	45	60
X field / kHz	29.6	42.0
Shape	Tangent ^{1}H	Tangent ^{1}H
Time / ms	0.4	0.5
Transfer II	**NH-CP**	**CH-CP**
^{1}H field / kHz	45	60
^{15}N field / kHz	29.6	42.0
Shape	Tangent ^{1}H	Tangent ^{1}H
Time / ms	0.4	0.5
Transfer III	**H-H SD**	**H-H SD**
Time / ms	0.2	0.2
Transfer IV	HP-CP	HP-CP
^{1}H field / kHz	60	60
^{31}P field / kHz	44.7	41.2
Shape	Tangent ^{1}H	Tangent ^{1}H
Time / ms	1.5	1.2
t2 increments	3072	3072
Sweep width (t2) / kHz	125	125
Acquisition time (t2) / ms	12.3	12.3
t1 increments	56	96
Sweep width (t1) / kHz	7.1	25
Acquisition time (t1)/ ms	3.9	1.9
^{1}H Spinal64 decoupling power / kHz	90	90
Interscan delay / s	2	2
Number of scans	9984	5632
Measurement time / h	311	300

Table D.3. *Experimental solid-state NMR parameters for DnaB helicase.*

Residue sidechain	Lysine		
Experiment	**NCX**		**NCC**
Complex	**ADP:AlF$_4^-$:ssDNA**	**ADP:AlF$_4^-$**	**ADP:AlF$_4^-$:ssDNA**
MAS frequency / kHz	17	17	17
Field / T	20	20	20
Transfer I	**HN-CP**	**HN-CP**	**HN-CP**
^{1}H field / kHz	60	60	60
X field / kHz	42.4	43.6	43.4
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
Time / ms	1.2	1.6	1.4
Transfer II	**NC-CP**	**NC-CP**	**NC-CP**
^{13}C field / kHz	6	6	6
^{15}N field / kHz	20.4	21.5	21.5
Shape	Tangent ^{13}C	Tangent ^{13}C	Tangent ^{13}C
Carrier / ppm	42	42	42
Time / ms	7	7	7
Transfer III	-	-	**DARR**
^{1}H field /kHz	-	-	17.0
Tims / ms	-	-	125
t1 increments	1792	1792	1792
Sweep width (t1) / kHz	66.7	66.7	66.7
Acquisition time (t1) / ms	13.4	13.4	13.4
t2 increments	3072	3072	3072
Sweep width (t2) / kHz	100	100	100
Acquisition time (t2) / ms	15.4	15.4	15.4
^{1}H SPINAL 64 dec. power / kHz	90	90	90
Interscan delay / s	2.7	2.7	3
Number of scans	16	16	32
Measurement time / h	22	22	48

Table D.3 (Continued) *Experimental solid-state NMR parameters for DnaB helicase.*

| Residue sidechain | Arginine | | |
| Experiment | NCX | | NCC |
Complex	ADP:AlF$_4^-$	ADP:AlF$_4^-$:ssDNA	ADP:AlF$_4^-$:ssDNA
MAS frequency / kHz	17	17	17
Field / T	20	20	20
Transfer I	**HN-CP**	**HN-CP**	**HN-CP**
^{1}H field / kHz	60	60	60
X field / kHz	40	40	43.4
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
Time / ms	0.7	0.7	0.7
Transfer II	**NC-CP**	**NC-CP**	**NC-CP**
^{13}C field / kHz	6	6	6
^{15}N field / kHz	9.4	9.4	20.7
Shape	Tangent ^{13}C	Tangent ^{13}C	Tangent ^{13}C
Carrier / ppm	159.3	159.3	159.3
Time / ms	3	3	3.5
Transfer III	-	-	**DARR**
^{1}H field /kHz	-	-	17.0
Tims / ms	-	-	100
t1 increments	1536	1536	1536
Sweep width (t1) / kHz	66.7	66.7	66.7
Acquisition time (t1) / ms	11.5	11.5	11.5
T2 increments	3072	3072	3072
Sweep width (t2) / kHz	100	100	100
Acquisition time (t2) / ms	15.4	15.4	15.4
^{1}H SPINAL64 dec. power / kHz	90	90	90
Interscan delay / s	2.5	2.5	3
Number of scans	24	16	32
Measurement time / h	27	18	41

Table D.3 (Continued) *Experimental solid-state NMR parameters for DnaB helicase.*

Complex	Apo	AMP-PCP:ssDNA	ADP:AlF$_4^-$:ssDNA	
Experiment	hNH	hNH	hNH	hCH
MAS frequency / kHz	108	105	110	110
Field / T	20	20	20	20
Transfer I	**HN-CP**	**HN-CP**	**HN-CP**	**HC-CP**
^{1}H field / kHz	135	135.0	85	89
X field / kHz	36	34.6	15	15
Shape	Tang. ^{1}H	Tang. ^{1}H	Tang. ^{1}H	Tang. ^{1}H
^{15}N carrier / ppm	117.5	117.5	78	56
Time / ms	0.6	0.6	1.0	0.9
Transfer II	**NH-CP**	**NH-CP**	**NH-CP**	**CH-CP**
^{1}H field / kHz	135	135.0	85	89
^{15}N field / kHz	36	34.6	-	15
^{13}C field / kHz	-	-	15	-
Shape	Tang. ^{1}H	Tang. ^{1}H	Tang. ^{1}H	Tang. ^{1}H
^{1}H carrier / ppm	4.8	4.8	4.8	4.8
Time / ms	1.0	1.5	1.0	0.7
t1 increments	620	512	512	488
Sweep width (t1) / ppm	180	180	180	190
Acquisition time (t1) / ms	20.0	16.5	16.5	6.0
t2 increments	1024	5550	1024	1024
Sweep width (t2) / ppm	46.7	43.2	40.0	46.7
Acquisition time (t2) / ms	12.9	69.9	15.0	12.9
^{1}H swfTPPM decoupling / kHz	10	10	10	10
^{15}N WALTZ64 decoupling / kHz	5.0	5.0	5.0	-
^{13}C WALTZ64 decoupling / kHz	-	-	-	5.0
Interscan delay / s	1.2	1.0	1.2	0.9
Number of scans	64	184	658	120
Measurement time / h	15	27	128	18

Table D.3 (Continued) *Experimental solid-state NMR parameters for DnaB helicase.*

Complex	(ADP:AlF$_4^-$:ssDNA)			
Experiment	hCANH	hCAcoNH	hNCAH	hNcoCAH
MAS frequency / kHz	110	110	110	110
Field / T	20	20	20	20
Transfer I	**HC-CP**	**HC-CP**	**HN-CP**	**HN-CP**
^{1}H/X field / kHz	89/15	89/15	85/15	85/15
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
^{15}N carrier / ppm	56	56	117.5	117.5
Time / ms	0.9	0.7	1.0	1.0
Transfer II	**CN-CP**	**CC DREAM**	**NC-CP**	**NC-CP**
^{13}C/^{15}N field / kHz	70/35	50/-	69/35	71/36
Shape	Tangent ^{13}C	Tangent ^{13}C	Tangent ^{13}C	Tangent ^{13}C
Carrier / ppm	117.5	145	56	176
Time / ms	15.0	7.0	16.0	18.0
Transfer III	**NH-CP**	**CN-CP**	**CH-CP**	**CC DREAM**
^{1}H/^{13}C/^{15}N field / kHz	89/-/15	-/67/35	89/15/-	-/51/-
Shape	Tangent ^{1}H	Tangent ^{13}C	Tangent ^{1}H	Tangent ^{13}C
Carrier / ppm	4.8	117.5	4.8	175
Time / ms	1.2	16.0	0.7	9.5
Transfer IV	**-**	**NH-CP**	**-**	**CH-CP**
^{1}H/X field / kHz	-/-	85/15	-/-	89/15
Shape	-	Tangent ^{1}H	-	Tangent ^{1}H
Carrier / ppm	-	4.8	-	4.8
Time / ms	-	1.0	-	0.7
t1 increments	86	104	48	42
Sweep width (t1) / ppm	30	30	32	32
Acquisition time (t1) / ms	6.7	8.1	8.7	7.6
t2 increments	48	48	120	120
Sweep width (t2) / ppm	32	32	40	40
Acquisition time (t2) / ms	8.7	8.7	7.0	7.0
t3 increments	1024	1024	1024	1024
Sweep width (t3) / ppm	46.7	46.7	46.7	46.7
Acquisition time (t3) / ms	12.9	12.9	12.9	12.9
^{1}H swfTPPM decoupling / kHz	10	10	10	10
^{15}N WALTZ64 decoupling / kHz	5	5	5	5
^{13}C WALTZ64 decoupling / kHz	5	5	5	5
MISSISSIPPI wat.supp. / kHz	20	20	20	20
Interscan delay / s	1.0	1.0	1.0	1.0
Number of scans	64	120	96	112
Measurement time / h	89	198	181	187

Table D.3 (Continued) *Experimental solid-state NMR parameters for DnaB helicase.*

Experiment	hNH T_2' (^{1}H)
Complex	ADP:AlF$_4^-$:DNA
MAS frequency / kHz	110
Field / T	20
Transfer I	**HN-CP**
^{1}H field / kHz	81
^{15}N field / kHz	27
Shape	Tangent ^{1}H
Time / ms	1.1
Transfer II	**NH-CP**
^{1}H field / kHz	76
^{15}N field / kHz	27
Shape	Tangent ^{1}H
T_2'(^{1}H) Measurement	Hahn-Echo block
Relaxation delays / ms	0.001, 0.250, 0.500, 0.750, 1, 1.25, 1.5, 1.75
^{1}H carrier / ppm	4.7
^{15}N carrier / ppm	117.5
t1 increments	2048
Sweep width (t1) / kHz	34.0
Acquisition time (t1) / ms	30.0
t2 increments	512
Sweep width (t2) / kHz	14.6
Acquisition time (t2) / ms	17.5
Water Suppression	MISSISIPPI
^{1}H field / kHz	15
Time / ms	120
^{1}H swfTPPM decoupling power / kHz	10
^{1}H swfTPPM decoupling pulse length / μs	46
^{15}N WALTZ64 decoupling power / kHz	15
^{15}N WALTZ64 decoupling pulse length / μs	1.58
Interscan delay / s	1
Number of scans	64
Measurement time / h	86

Table D.3 (Continued) *Experimental solid-state NMR parameters for DnaB helicase.*

E. Supplementary information to 150 kHz MAS NMR spectroscopy

E.1. Figures
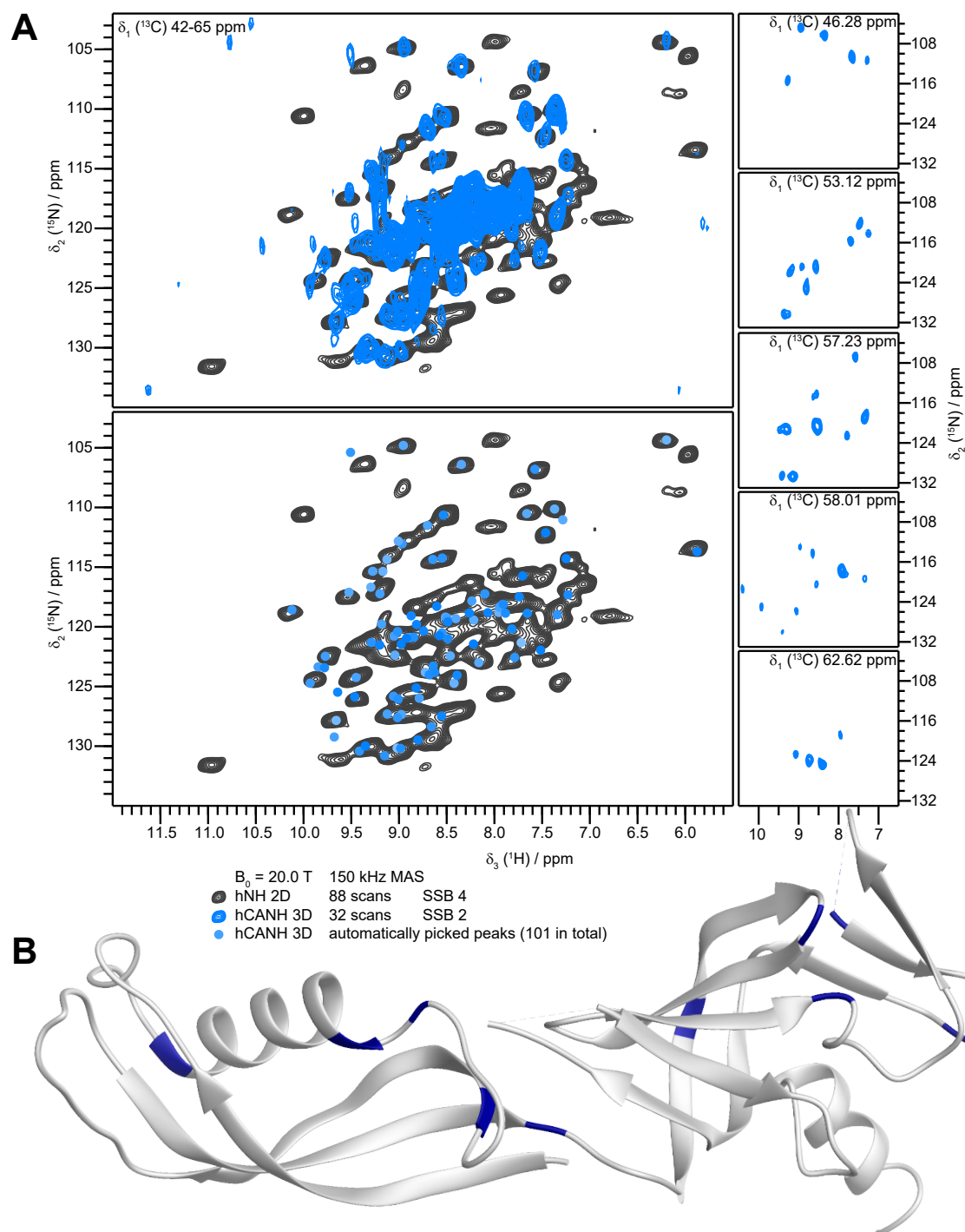
Figure E.1. *Projection and back-prediction of a 3D hCANH onto a 2D hNH spectrum.* In **(a)**, representative planes of the 3D are given to the right of the figure, while on the top left, the full $^{13}C\alpha$ dimension (42-65 ppm) is projected onto the 2D. On the bottom left, the 101 automatically picked peaks (with CcpNmr Analysis 2.4.2 [136, 137]) from the 3D are marked onto the 2D. In **(b)**, the 2D peaks for which the resonances are missing in the 3D spectrum are depicted on the crystal structure (PDB ID 1GO3) in blue.
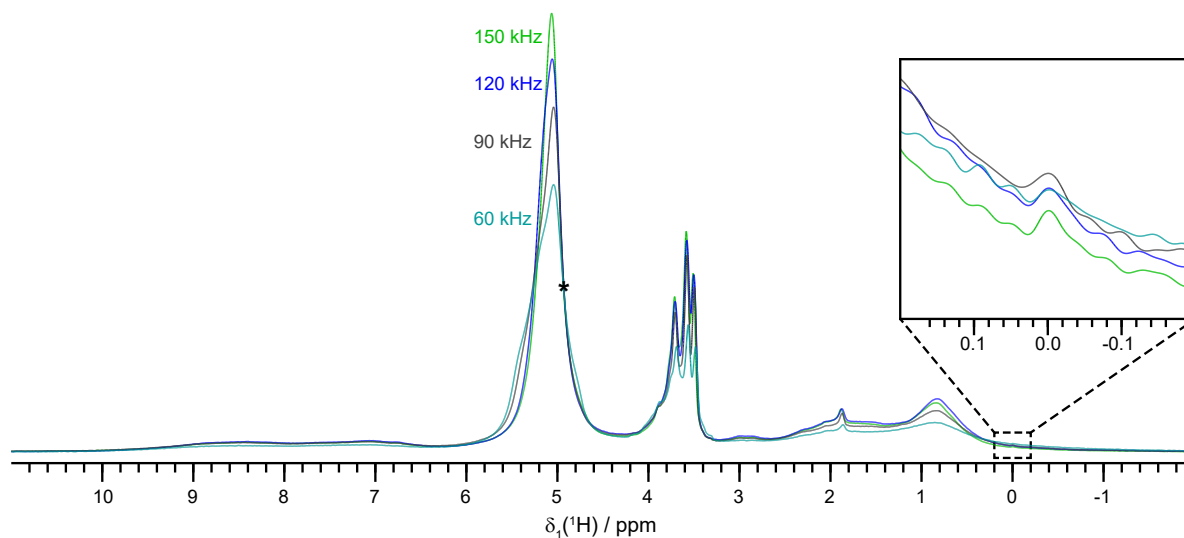
Figure E.2. *Temperature control at various MAS frequencies.* This figure shows 1D proton spectra of Rpo4/7 at 60-150 kHz MAS, referenced to DSS (signal shown in zoom-in panel). Although a resolved peak for the free-water pool described in Böckmann et al. [168] is not seen in these spectra, it would be expected to appear on the right flank of the water peak (indicated with an asterisk). For this reason, cooling was adjusted such that this point was overlaid at all MAS frequencies, assuming that the temperature in this case is approximately 10 °C.
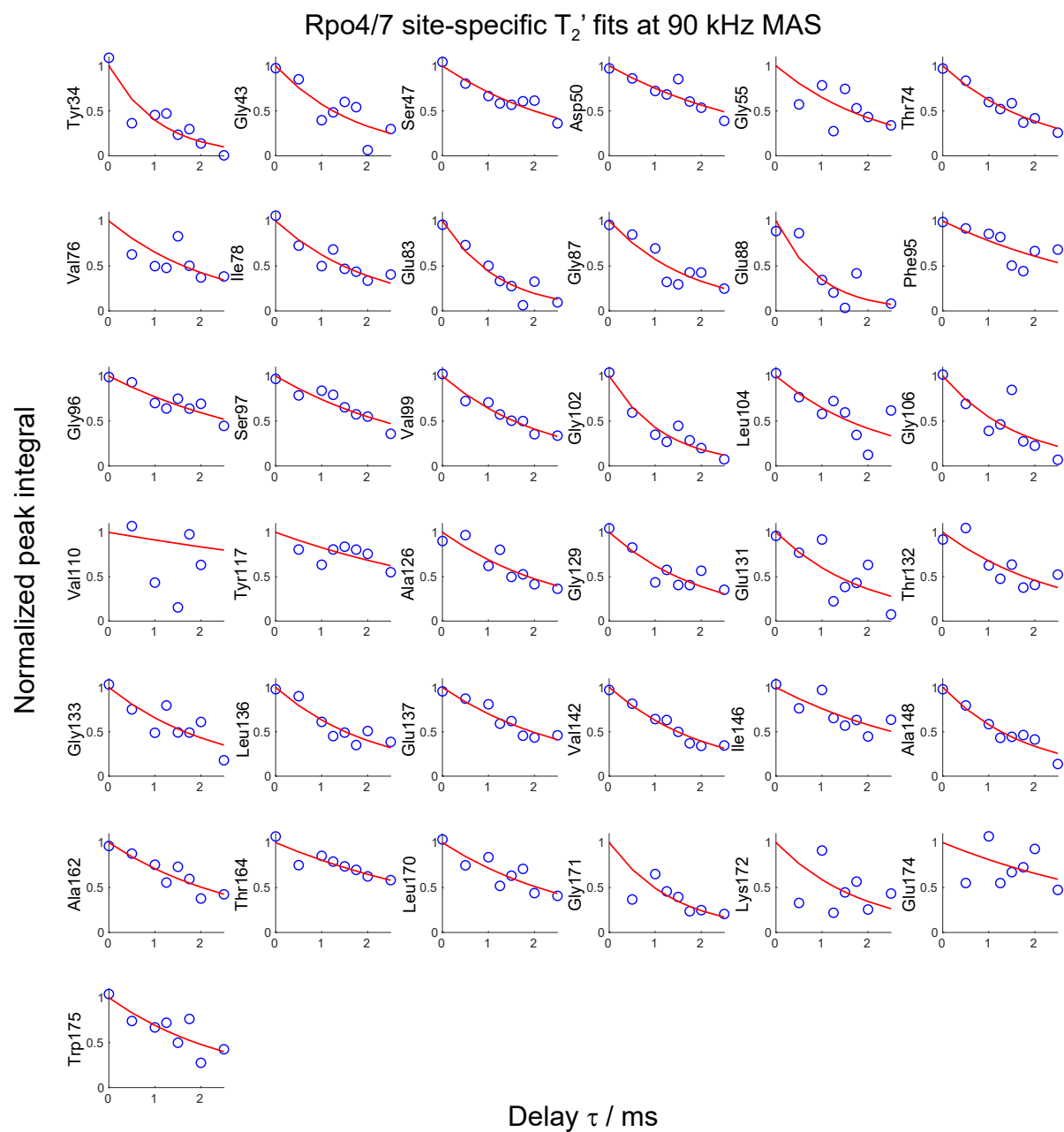
Figure E.3. *Fitting curves for site-specific T$_2$' relaxation rates in Rpo4/7.* Each curve represents one cross peak in an NH 2D correlation spectrum recorded at 90 kHz MAS. The points in each curve represent integrals of that cross peak in eight spectra with different delays τ, given in ms.
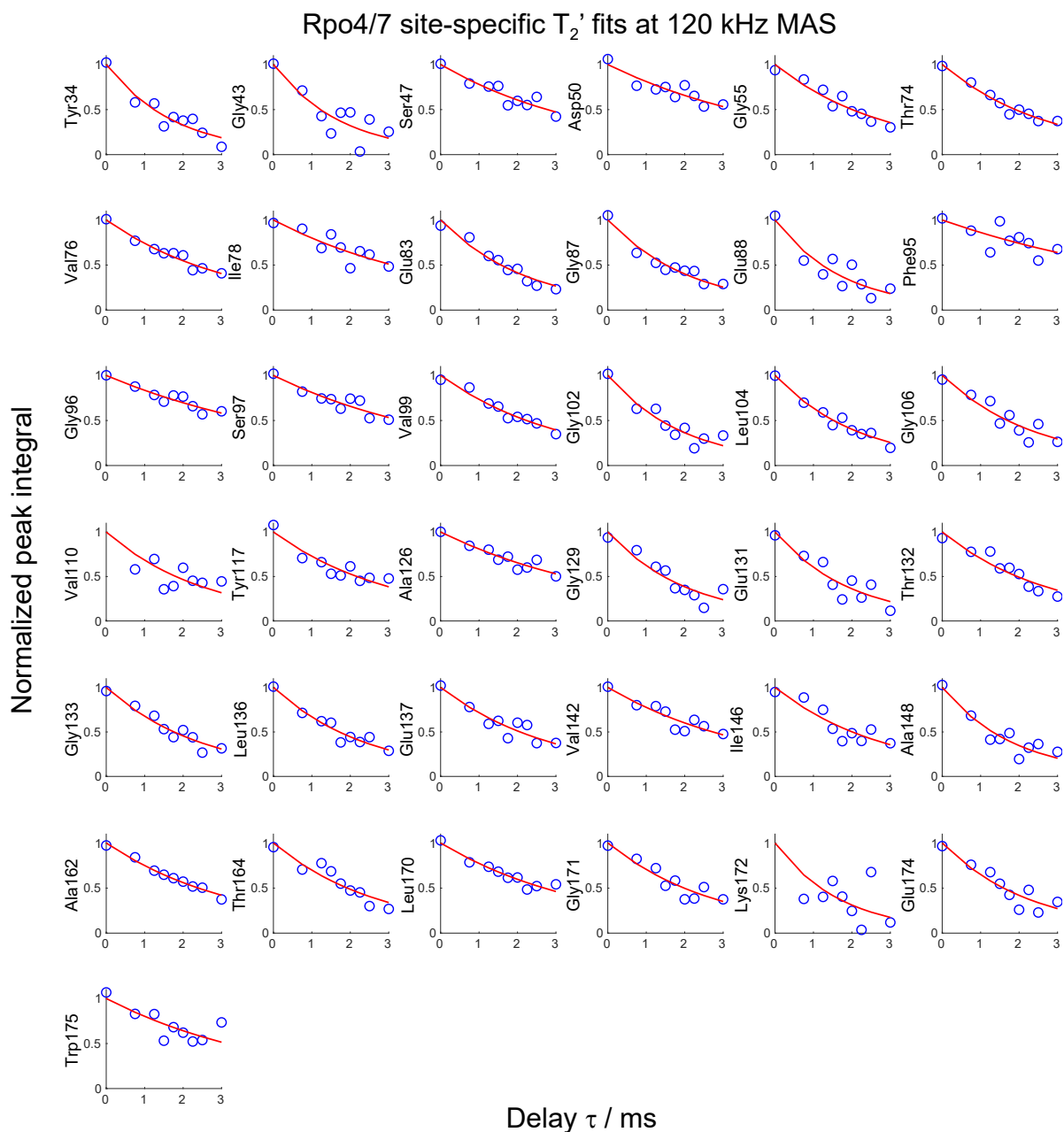
Figure E.4. *Fitting curves for site-specific T₂' relaxation rates in Rpo4/7.* Each curve represents one cross peak in an NH 2D correlation spectrum recorded at 120 kHz MAS. The points in each curve represent integrals of that cross peak in eight spectra with different delays τ, given in ms.
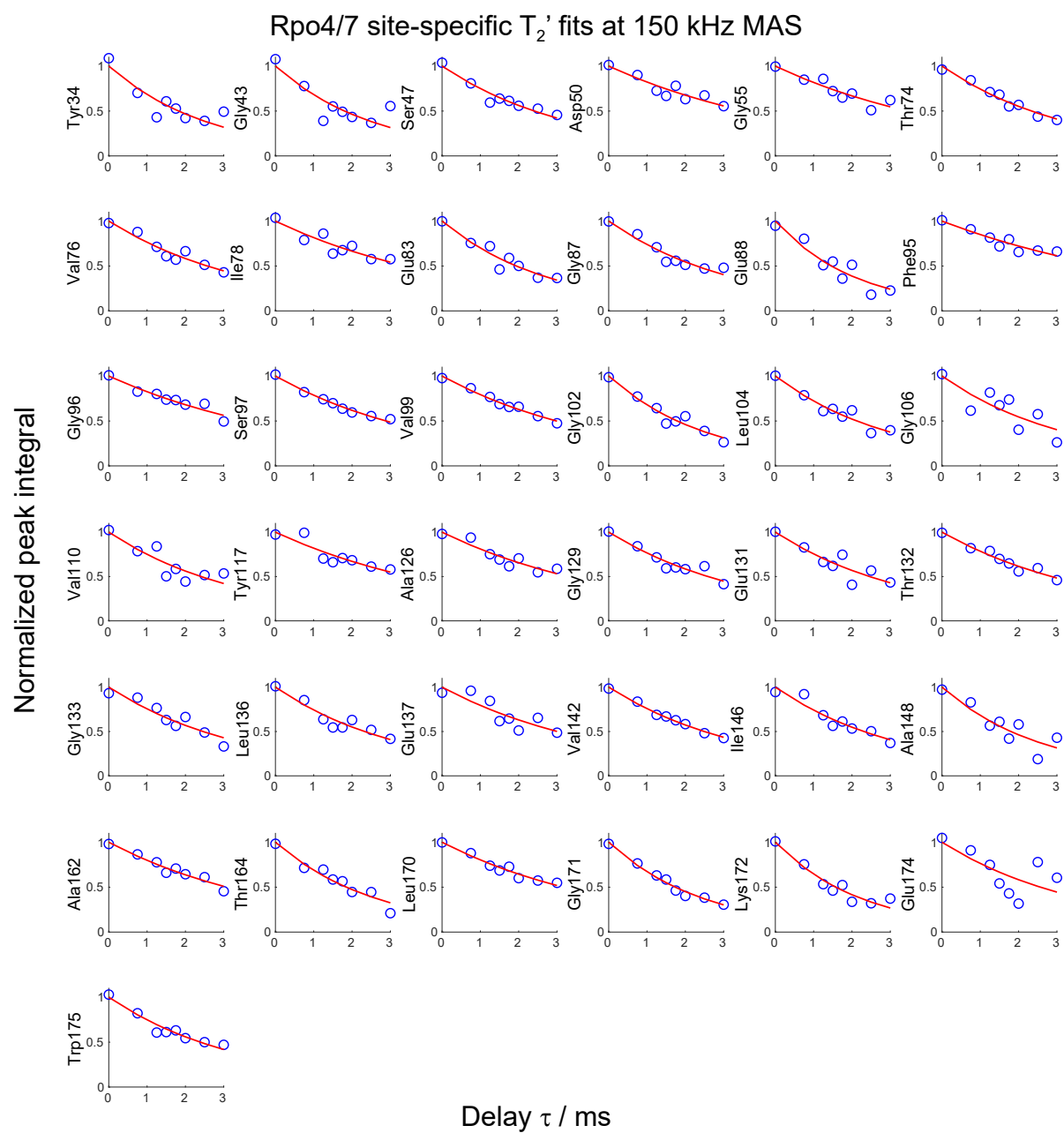
Figure E.5. *Fitting curves for site-specific T_2' relaxation rates in Rpo4/7.* Each curve represents one cross peak in an NH 2D correlation spectrum recorded at 150 kHz MAS. The points in each curve represent integrals of that cross peak in eight spectra with different delays τ, given in ms.
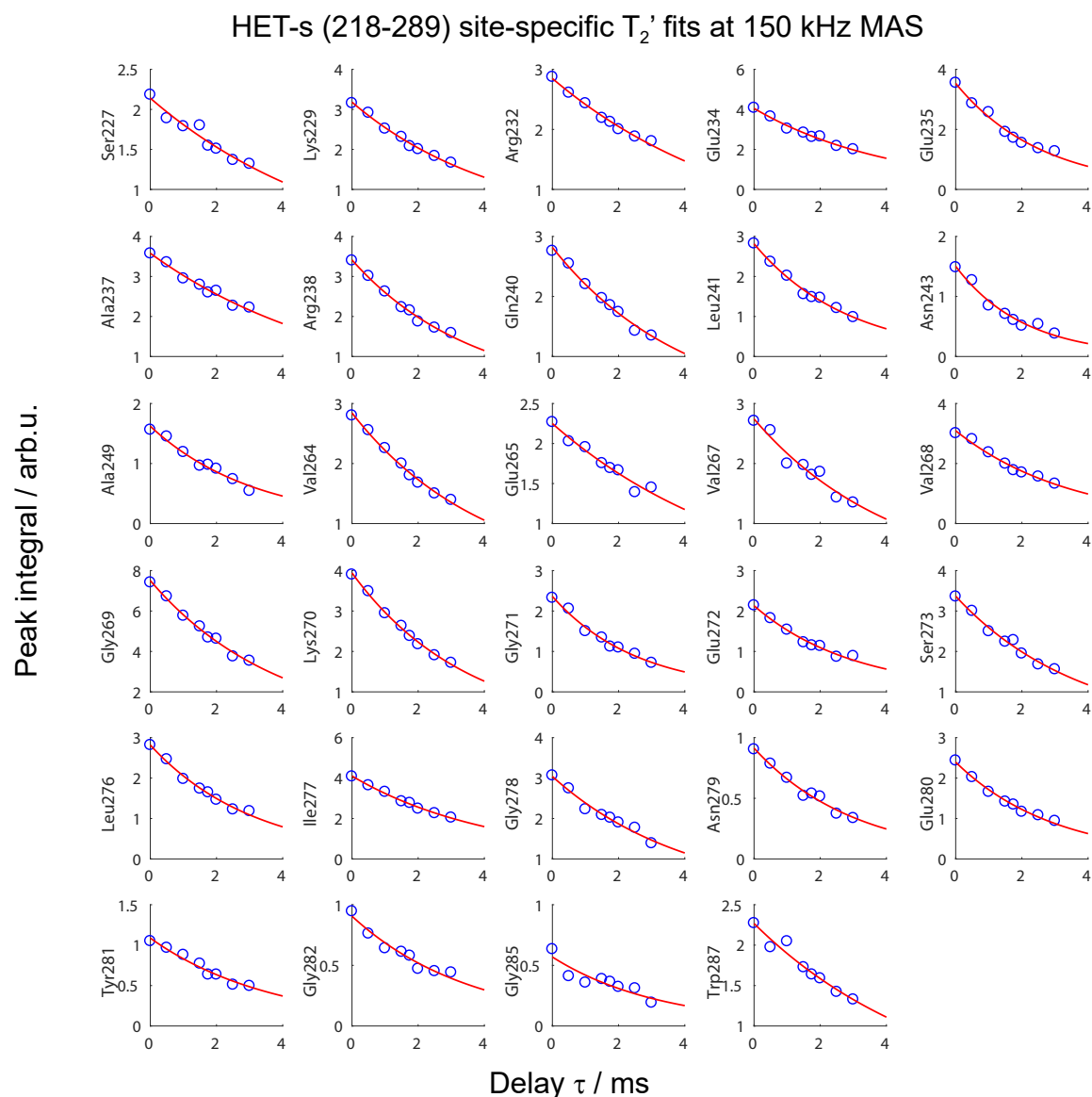
Figure E.6. *Fitting curves for site-specific T_2' relaxation rates in HET-s (218-289).* Each curve represents one cross peak in an NH 2D correlation spectrum recorded at 150 kHz MAS. The points in each curve represent integrals of that cross peak in eight spectra with different delays τ, given in ms.
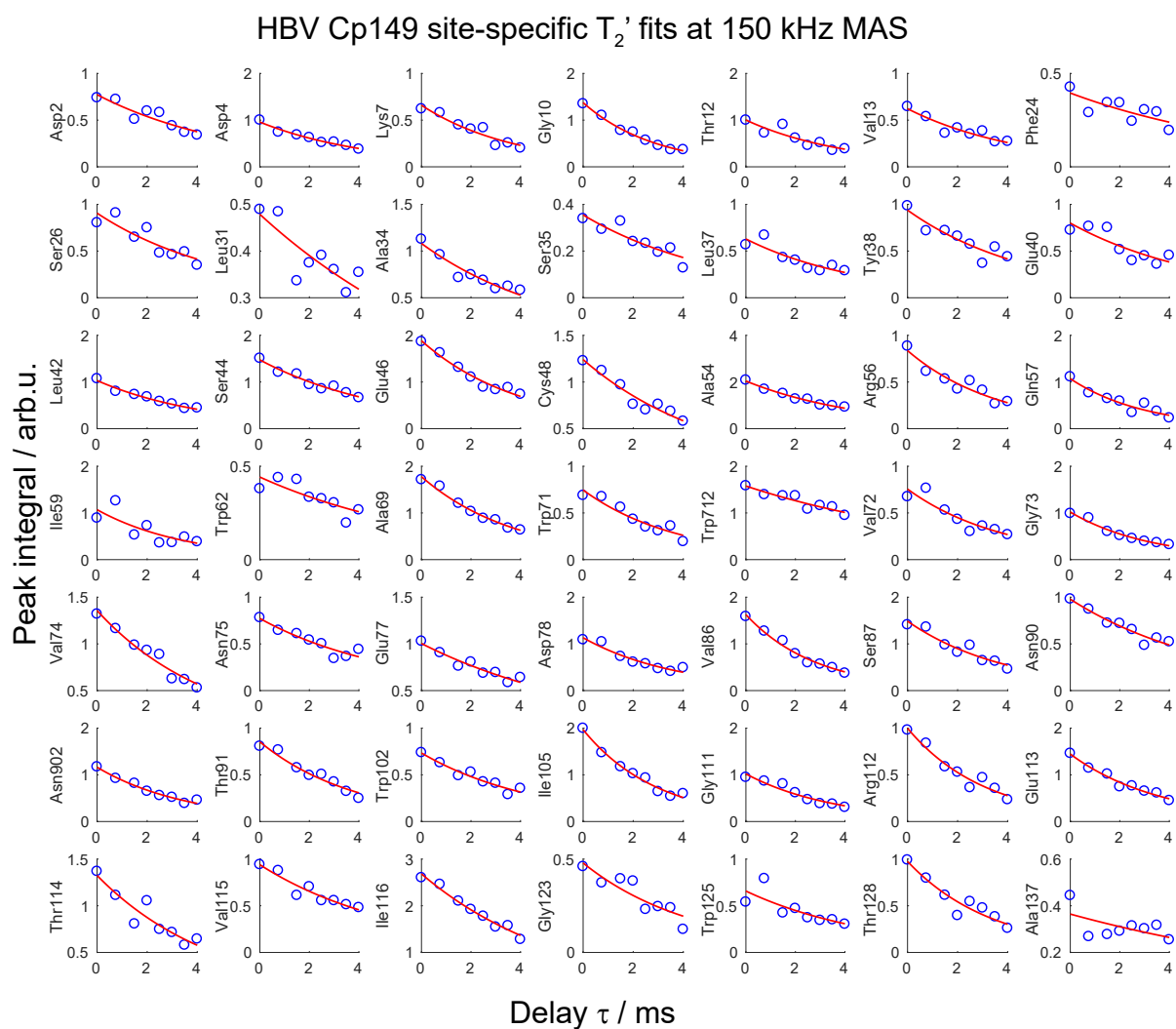
Figure E.7. *Fitting curves for site-specific T_2' relaxation rates in HBV Cp149.* Each curve represents one cross peak in an NH 2D correlation spectrum recorded at 150 kHz MAS. The points in each curve represent integrals of that cross peak in eight spectra with different delays τ, given in ms.

E.2. Tables

| Experiment	hNH	hCH	hCANH	hNCAH	hCONH
MAS frequency / kHz	110	110	110	110	110
Field / T	20.0	20.0	20.0	20.0	20.0
t_1 increments	788	1624	142	32	58
Sweep width (t_1) / ppm	180	200	55	32	20
Acquisition time (t_1) / ms	25	19	6	5.8	6.8
t_2 increments	5550	2048	32	74	38
Sweep width (t_2) / ppm	47	47	32	30	34
Acquisition time (t_2) / ms	70	26	5.8	5.8	6.5
t_3 increments	-	-	2048	3072	3072
Sweep width (t_3) / kHz	-	-	47	47	47
Acquisition time (t_3) / ms	-	-	26	38.7	38.7
Proton decoupling (swfTPPM) / kHz	10	10	10	10	10
Nitrogen decoupling (WALTZ64) / kHz	5	5	5	5	5
Water suppression (120 ms MISS) / kHz	20	20	20	20	20
Interscan delay / s	1.0	1.2	1.0	1.1	1.0
Number of scans	64	32	56	48	88
Measurement time / hh:mm	16:48	19:50	82:58	40:00	64:00
Transfer I	**HN-CP**	**HC-CP**	**HC-CP**	**HN-CP**	**HC-CP**
^{1}H field / kHz	89	89	89	89	89
X field / kHz	15	15	15	15	15
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
Carrier / ppm	117.5	56	56	117.5	175
Time / ms	1.6	0.9	0.6	1.5	1.3
Transfer II	**NH-CP**	**CH-CP**	**CN-CP**	**NC-CP**	**CN-CP**
^{1}H field / kHz	89	86	-	-	-
^{13}C field / kHz	-	15	71	69	69
^{15}N field / kHz	15	-	35	35	35
Shape	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{13}C	Tangent ^{13}C	Tangent ^{13}C
Carrier / ppm	4.8	4.8	117.5	56	117.5
Time / ms	1.0	1.0	16.5	17.0	16.0
Transfer III	**-**	**-**	**NH-CP**	**CH-CP**	**NH-CP**
^{1}H field / kHz	-	-	89	85	89
^{13}C field / kHz	-	-	-	15	-
^{15}N field / kHz	-	-	15	-	15
Shape	-	-	Tangent ^{1}H	Tangent ^{1}H	Tangent ^{1}H
Carrier / ppm	-	-	4.8	4.8	4.8
Time / ms	-	-	1.05	0.55	1.25

Table E.1. *Overview of experimental parameters of the 110 kHz MAS solid-state NMR experiments for H^N and H^α assignment of Rpo4/7.* Experiments were performed on a 0.7 mm rotor. Carrier frequencies during transfers are given for the channel to which magnetisation is transferred.

	0.7 mm 100 kHz	0.5 mm 150 kHz
Acquisition time t in hours	3.57	13.7
Mean noise $\overline{n}$	$18.8 \cdot 10^5$	$5.51 \cdot 10^5$
Standard error of mean noise $s_{\overline{n}}$	$0.37 \cdot 10^5$	$0.13 \cdot 10^5$
$\partial \overline{SNR}/\partial \overline{n}$ (*)	$-1.56 \cdot 10^{-12}$	$-12.0 \cdot 10^{-12}$
CCPN scaling factor F	$4.88 \cdot 10^{-4}$	$39.1 \cdot 10^{-4}$
$\overline{n} \cdot F$	$9.18 \cdot 10^2$	$21.5 \cdot 10^2$
Mean peak height $\overline{p}$	$9.58 \cdot 10^3$	$29.1 \cdot 10^3$
$s_{\overline{p}}$	$1.40 \cdot 10^3$	$4.09 \cdot 10^3$
$\partial \overline{SNR}/\partial \overline{p}$ (*)	$5.76 \cdot 10^{-4}$	$1.26 \cdot 10^{-4}$
$\overline{SNR}$	10	14
$\overline{SNR}/\sqrt{t} \cdot h$	5.52	3.66
$s_{\overline{SNR}/\sqrt{t} \cdot h}$	0.81	0.51
Normalised to largest rotor	**1.00 ± 0.15**	**0.66 ± 0.09**
Cys48 peak height	$1.55 \cdot 10^5$	$6.06 \cdot 10^5$
$\overline{SNR}$	17	28
$\overline{SNR}/\sqrt{t} \cdot h$	8.94	7.61
$s_{\overline{SNR}/\sqrt{t} \cdot h}$	$5.81 \cdot 10^{-8}$	$15.5 \cdot 10^{-8}$
Normalised to largest rotor	**1.00**	**0.85**

(*) Components of $s_{\overline{SNR}} = \sqrt{\left(\frac{\partial SNR}{\partial n} \cdot s_{\overline{n}}\right)^2 + \left(\frac{\partial SNR}{\partial p} \cdot s_{\overline{p}}\right)^2}$

Table E.2. *Signal to noise comparison in a 0.7 mm and a 0.5 mm probe.* Experiments were performed on full 0.7 mm and 0.5 mm rotors with fully protonated HBc149, operated at 100 and 150 kHz MAS respectively.

Bibliography

[1] Milo, R., 2013. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* 35:1050–1055. http://doi.wiley.com/10.1002/bies.201300066.

[2] Steen, H., and M. Mann, 2004. The abc's (and xyz's) of peptide sequencing. *Nature Reviews Molecular Cell Biology* 5:699–711. http://www.nature.com/articles/nrm1468.

[3] Greenfield, N. J., 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* 1:2876–2890. http://www.nature.com/articles/nprot.2006.202.

[4] Pelton, J. T., and L. R. McLean, 2000. Spectroscopic Methods for Analysis of Protein Secondary Structure. *Analytical Biochemistry* 277:167–176. https://www.sciencedirect.com/science/article/pii/S0003269799943208.

[5] Gulbis, J. M., Z. Kelman, J. Hurwitz, M. O'Donnell, and J. Kuriyan, 1996. Structure of the C-Terminal Region of p21WAF1/CIP1 Complexed with Human PCNA. *Cell* 87:297–306. https://www.sciencedirect.com/science/article/pii/S0092867400813471{#}FIG1.

[6] Wasmer, C., A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek, and B. H. Meier, 2008. Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 319:1523–1526.

[7] Van Melckebeke, H., C. Wasmer, A. Lange, E. AB, A. Loquet, A. Böckmann, and B. H. Meier, 2010. Atomic-Resolution Three-Dimensional Structure of HET-s(218-289) Amyloid Fibrils by Solid-State NMR Spectroscopy. *Journal of the American Chemical Society* 132:13765–13775. https://pubs.acs.org/doi/10.1021/ja104213j.

[8] Colvin, M. T., R. Silvers, Q. Z. Ni, T. V. Can, I. Sergeyev, M. Rosay, K. J. Donovan, B. Michael, J. Wall, S. Linse, and R. G. Griffin, 2016. Atomic Resolution Structure of Monomorphic Aβ 42 Amyloid Fibrils. *Journal of the American Chemical Society* 138:9663–9674. http://pubs.acs.org/doi/10.1021/jacs.6b05129.

[9] Wälti, M. A., F. Ravotti, H. Arai, C. G. Glabe, J. S. Wall, A. Böckmann, P. Güntert, B. H. Meier, and R. Riek, 2016. Atomic-resolution structure of a disease-relevant

Aβ(1-42) amyloid fibril. *Proceedings of the National Academy of Sciences* 113:E4976–84. http://www.ncbi.nlm.nih.gov/pubmed/27469165.

[10] Verasdonck, J., L. Bousset, J. Gath, R. Melki, A. Böckmann, and B. H. Meier, 2016. Further exploration of the conformational space of α-synuclein fibrils: solid-state NMR assignment of a high-pH polymorph. *Biomolecular NMR Assignments* 10:5–12. http://link.springer.com/10.1007/s12104-015-9628-9.

[11] Tuttle, M. D., G. Comellas, A. J. Nieuwkoop, D. J. Covell, D. A. Berthold, K. D. Kloepper, J. M. Courtney, J. K. Kim, A. M. Barclay, A. Kendall, W. Wan, G. Stubbs, C. D. Schwieters, V. M. Lee, J. M. George, and C. M. Rienstra, 2016. Solid-state NMR structure of a pathogenic fibril of full-length human alpha-synuclein. *Nat Struct Mol Biol* 23:409–415.

[12] Kühlbrandt, W., 2014. The resolution revolution. *Science* 343:1443–4. http://www.ncbi.nlm.nih.gov/pubmed/24675944.

[13] Kondo, Y., J. Ognjenović, S. Banerjee, D. Karandur, A. Merk, K. Kulhanek, K. Wong, J. P. Roose, S. Subramaniam, and J. Kuriyan, 2019. Cryo-EM structure of a dimeric B-Raf:14-3-3 complex reveals asymmetry in the active sites of B-Raf kinases. *Science* 366:109–115. http://www.ncbi.nlm.nih.gov/pubmed/31604311.

[14] Laverty, D., R. Desai, T. Uchański, S. Masiulis, W. J. Stec, T. Malinauskas, J. Zivanov, E. Pardon, J. Steyaert, K. W. Miller, and A. R. Aricescu, 2019. Cryo-EM structure of the human α1β3γ2 GABAA receptor in a lipid bilayer. *Nature* 565:516–520. http://www.nature.com/articles/s41586-018-0833-4.

[15] Sugita, Y., H. Matsunami, Y. Kawaoka, T. Noda, and M. Wolf, 2018. Cryo-EM structure of the Ebola virus nucleoprotein-RNA complex at 3.6 Å resolution. *Nature* 563:137–140. http://www.nature.com/articles/s41586-018-0630-0.

[16] Ruan, J., S. Xia, X. Liu, J. Lieberman, and H. Wu, 2018. Cryo-EM structure of the gasdermin A3 membrane pore. *Nature* 557:62–67. http://www.nature.com/articles/s41586-018-0058-6.

[17] Sirohi, D., Z. Chen, L. Sun, T. Klose, T. C. Pierson, M. G. Rossmann, and R. J. Kuhn, 2016. The 3.8 Å resolution cryo-EM structure of Zika virus. *Science* 352:467–70. http://www.ncbi.nlm.nih.gov/pubmed/27033547.

[18] Costello, W. N., Y. Xiao, and K. K. Frederick, 2019. DNP-Assisted NMR Investigation of Proteins at Endogenous Levels in Cellular Milieu. *Methods in enzymology* 615:373–406.

[19] Xiao, Y., W. Costello, C. Madrid, and K. Frederick, 2018. Combining DNP NMR with Segmental and Specific Labeling to Study the Quaternary Structures of Yeast Prion Protein Strains. *Biophysical Journal* 114:220a.

[20] Zhao, C., and D. Shukla, 2018. SAXS-guided enhanced unbiased sampling for structure determination of proteins and complexes. *Scientific Reports* 8:17748. `http://www.nature.com/articles/s41598-018-36090-z`.

[21] Nespovitaya, N., K. Barylyuk, C. Eichmann, R. Zenobi, and R. Riek, 2014. The production of recombinant ^{15}N, ^{13}C-labelled somatostatin 14 for NMR spectroscopy. *Protein Expression and Purification* 99:78–86. `https://www.sciencedirect.com/science/article/pii/S1046592814000680`.

[22] Luchinat, E., and L. Banci, 2017. In-cell NMR: a topical review. *IUCrJ* 4:108–118. `http://www.ncbi.nlm.nih.gov/pubmed/28250949`.

[23] Russell, R. W., M. P. Fritz, J. Kraus, C. M. Quinn, T. Polenova, and A. M. Gronenborn, 2019. Accuracy and precision of protein structures determined by magic angle spinning NMR spectroscopy: for some 'with a little help from a friend'. *Journal of Biomolecular NMR* 73:333–346. `http://link.springer.com/10.1007/s10858-019-00233-9`.

[24] Cuniasse, P., P. Tavares, E. V. Orlova, and S. Zinn-Justin, 2017. Structures of biomolecular complexes by combination of NMR and cryoEM methods. *Current Opinion in Structural Biology* 43:104–113. `https://www.sciencedirect.com/science/article/pii/S0959440X16302366`.

[25] Vijay-Kumar, S., C. E. Bugg, and W. J. Cook, 1987. Structure of ubiquitin refined at 1.8 Å resolution. *Journal of Molecular Biology* 194:531–544. `https://www.sciencedirect.com/science/article/pii/0022283687906796?via{%}3Dihub`.

[26] Agarwal, V., S. Penzel, K. Szekely, R. Cadalbert, E. Testori, A. Oss, J. Past, A. Samoson, M. Ernst, A. Böckmann, and B. H. Meier, 2014. De Novo 3D Structure Determination from Sub-milligram Protein Samples by Solid-State 100 kHz MAS NMR Spectroscopy. *Angewandte Chemie International Edition* 53:12253–12256. `http://doi.wiley.com/10.1002/anie.201405730`.

[27] Penzel, S., A. A. Smith, V. Agarwal, A. Hunkeler, M.-L. Org, A. Samoson, A. Böckmann, M. Ernst, and B. H. Meier, 2015. Protein resonance assignment at MAS frequencies approaching 100 kHz: a quantitative comparison of J-coupling and dipolar-coupling-based transfer methods. *Journal of Biomolecular NMR* 63:165–186. `http://link.springer.com/10.1007/s10858-015-9975-y`.

[28] Penzel, S., A. Oss, M.-L. Org, A. Samoson, A. Böckmann, M. Ernst, and B. H. Meier, 2019. Spinning faster: protein NMR at MAS frequencies up to 126 kHz. *Journal of Biomolecular NMR* 73:19–29. `http://link.springer.com/10.1007/s10858-018-0219-9`.

[29] Clore, G. M., A. M. Gronenborn, N. Essig, A. Achari, M. Whitlow, P. Wingfield, and G. Clore, 1991. Structures of larger proteins in solution: three- and four-dimensional

heteronuclear NMR spectroscopy. *Science* 252:1390–9. http://www.ncbi.nlm.nih.gov/pubmed/2047852.

[30] Gallagher, T., P. Alexander, P. Bryan, and G. L. Gilliland, 1994. Two Crystal Structures of the B1 Immunoglobulin-Binding Domain of Streptococcal Protein G and Comparison with NMR. *Biochemistry* 33:4721–4729. https://pubs.acs.org/doi/abs/10.1021/bi00181a032.

[31] Kirsten Frank, M., F. Dyda, A. Dobrodumov, and A. M. Gronenborn, 2002. Core mutations switch monomeric protein GB1 into an intertwined tetramer. *Nature Structural Biology* 9:877–885. http://www.nature.com/doifinder/10.1038/nsb854.

[32] Vernizzi, G., and M. Olvera de la Cruz, 2007. Faceting ionic shells into icosahedra via electrostatics. *Proceedings of the National Academy of Sciences of the United States of America* 104:18382–6. http://www.ncbi.nlm.nih.gov/pubmed/18003933.

[33] Petkova, A. T., R. D. Leapman, Z. Guo, W.-M. Yau, M. P. Mattson, and R. Tycko, 2005. Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science* 307:262–5. http://www.ncbi.nlm.nih.gov/pubmed/15653506.

[34] Nelson, R., M. R. Sawaya, M. Balbirnie, A. Ø. Madsen, C. Riekel, R. Grothe, and D. Eisenberg, 2005. Structure of the cross-β spine of amyloid-like fibrils. *Nature* 435:773–778. http://www.nature.com/articles/nature03680.

[35] Riek, R., and D. S. Eisenberg, 2016. The activities of amyloids from a structural perspective. *Nature* 539:227–235. http://www.nature.com/articles/nature20416.

[36] Meier, B. H., and M. Ernst, 2017. Magnetic resonance. Notes of the Physical Chemistry IV lecture at ETH Zurich.

[37] Meier, B. H., and M. Ernst, 2018. High-resolution solid-state NMR: principles and applications. Notes of the Advanced Magnetic Resonance lecture at ETH Zurich.

[38] Keeler, J., 2011. Understanding NMR spectroscopy. John Wiley & Sons.

[39] Levitt, M. H., 2008. Spin dynamics: basics of nuclear magnetic resonance. John Wiley & Sons, 2nd edition.

[40] Duer, M. J., 2008. Introduction to solid-state NMR spectroscopy. Blackwell Publishing, 2nd edition.

[41] Schmidt-Rohr, K., and H. Spiess, 1994. Multidimensional solid-state NMR and polymers. Academic Press.

[42] Ernst, R., G. Bodenhausen, and A. Wokaun, 1990. Principles of magnetic resonance in one and two dimensions. Oxford University Press.

[43] Abragam, A., 1983. Principles of nuclear magnetism. Oxford University Press.

[44] Andrew, E. R., A. Bradbury, and R. Eades, 1958. Nuclear Magnetic Resonance Spectra from a Crystal rotated at High Speed. *Nature* 182:1659–1659. http://www.nature.com/articles/1821659a0.

[45] Lowe, I. J., 1959. Free Induction Decays of Rotating Solids. *Physical Review Letters* 2:285–287. https://link.aps.org/doi/10.1103/PhysRevLett.2.285.

[46] Maricq, M., and J. S. Waugh, 1979. NMR in rotating solids. *The Journal of Chemical Physics* 70:3300–3316. http://aip.scitation.org/doi/10.1063/1.437915.

[47] Mehring, M., 1983. Principles of High Resolution NMR in Solids. Springer, Berlin, 2nd edition.

[48] Kowalewski, J., and L. Mäler, 2006. Nuclear spin relaxation in liquids: theory, experiments, and applications. CRC Press, 1st edition.

[49] Haeberlen, U., and J. S. Waugh, 1968. Coherent Averaging Effects in Magnetic Resonance. *Physical Review* 175:453–467. https://link.aps.org/doi/10.1103/PhysRev.175.453.

[50] Magnus, W., 1954. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics* 7:649–673. http://doi.wiley.com/10.1002/cpa.3160070404.

[51] Sternberg, U., R. Witter, I. Kuprov, J. M. Lamley, A. Oss, J. R. Lewandowski, and A. Samoson, 2018. [1]H line width dependence on MAS speed in solid state NMR - Comparison of experiment and simulation. *Journal of Magnetic Resonance* 291:32–39. https://www.sciencedirect.com/science/article/pii/S1090780718301071.

[52] Paëpe, G. D., N. Giraud, A. Lesage, P. Hodgkinson, A. Böckmann, and L. Emsley, 2003. Transverse Dephasing Optimized Solid-State NMR Spectroscopy. *Journal of the American Chemical Society* 125:13938–13939. https://pubs.acs.org/doi/abs/10.1021/ja037213j.

[53] Schütz, A. K., T. Vagt, M. Huber, O. Y. Ovchinnikova, R. Cadalbert, J. Wall, P. Güntert, A. Böckmann, R. Glockshuber, and B. H. Meier, 2015. Atomic Resolution Three Dimensional Structure of Amyloid β Fibrils Bearing the Osaka Mutation. *Angewandte Chemie International Edition* 54:331–335.

[54] Schledorn, M., B. H. Meier, and A. Böckmann, 2015. Alternative salt bridge formation in Aβ – a hallmark of early-onset Alzheimer's disease? *Frontiers in molecular biosciences* 2:14. http://www.ncbi.nlm.nih.gov/pubmed/25988181.

[55] Hardy, J. A., and G. A. Higgins, 1992. Alzheimer's disease: the amyloid cascade hypothesis. *Science* 256:184–5. http://www.ncbi.nlm.nih.gov/pubmed/1566067.

[56] Kane, M. D., W. J. Lipinski, M. J. Callahan, F. Bian, R. A. Durham, R. D. Schwarz, A. E. Roher, and L. C. Walker, 2000. Evidence for seeding of beta -amyloid by intracerebral infusion of Alzheimer brain extracts in beta -amyloid precursor protein-transgenic mice. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 20:3606–11. http://www.ncbi.nlm.nih.gov/pubmed/10804202.

[57] Eisele, Y. S., U. Obermüller, G. Heilbronner, F. Baumann, S. A. Kaeser, H. Wolburg, L. C. Walker, M. Staufenbiel, M. Heikenwalder, and M. Jucker, 2010. Peripherally applied Abeta-containing inoculates induce cerebral beta-amyloidosis. *Science* 330:980–2. http://www.ncbi.nlm.nih.gov/pubmed/20966215.

[58] Benilova, I., E. Karran, and B. D. Strooper, 2012. The toxic Aβ oligomer and Alzheimer's disease: an emperor in need of clothes. *Nature neuroscience* 15:349–357.

[59] Cohen, S. I. A., S. Linse, L. M. Luheshi, E. Hellstrand, D. A. White, L. Rajah, D. E. Otzen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, 2013. Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proceedings of the National Academy of Sciences* 110:9758–63. http://www.ncbi.nlm.nih.gov/pubmed/23703910.

[60] Matsuzaki, K., 2014. How Do Membranes Initiate Alzheimer's Disease? Formation of Toxic Amyloid Fibrils by the Amyloid β-Protein on Ganglioside Clusters. *Accounts of Chemical Research* 47:2397–2404. http://pubs.acs.org/doi/10.1021/ar500127z.

[61] Wu, L., P. Rosa-Neto, G.-Y. R. Hsiung, A. D. Sadovnick, M. Masellis, S. E. Black, J. Jia, and S. Gauthier, 2012. Early-Onset Familial Alzheimer's Disease (EO-FAD). *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques* 39:436–445. https://www.cambridge.org/core/product/identifier/S0317167100013949/type/journal{_}article.

[62] Meier, B. H., and A. Böckmann, 2015. The structure of fibrils from 'misfolded' proteins. *Current Opinion in Structural Biology* 30:43–49. https://www.sciencedirect.com/science/article/pii/S0959440X14001596.

[63] Meyer-Luehmann, M., J. Coomaraswamy, T. Bolmont, S. Kaeser, C. Schaefer, E. Kilger, A. Neuenschwander, D. Abramowski, P. Frey, A. L. Jaton, J.-M. Vigouret, P. Paganetti, D. M. Walsh, P. M. Mathews, J. Ghiso, M. Staufenbiel, L. C. Walker, and M. Jucker, 2006. Exogenous induction of cerebral beta-amyloidogenesis is governed by agent and host. *Science* 313:1781–4. http://www.ncbi.nlm.nih.gov/pubmed/16990547.

[64] Colby, D. W., and S. B. Prusiner, 2011. Prions. *Cold Spring Harbor perspectives in biology* 3:a006833. http://www.ncbi.nlm.nih.gov/pubmed/21421910.

[65] Stöhr, J., J. C. Watts, Z. L. Mensinger, A. Oehler, S. K. Grillo, S. J. DeArmond, S. B. Prusiner, and K. Giles, 2012. Purified and synthetic Alzheimer's amyloid beta (Aβ) prions. *Proceedings of the National Academy of Sciences* 109:11025–30. http://www.ncbi.nlm.nih.gov/pubmed/22711819.

[66] Hendriks, L., C. van Duijn, P. Cras, M. Cruts, W. van Hul, F. van Harskamp, A. Warren, M. G. McInnis, S. E. Antonarakis, J.-J. Martin, C. van Broeckhoven, and A. Hofman, 1992. Presenile dementia and cerebral haemorrhage linked to a mutation at codon 692 of the β-amyloid precursor protein gene. *Nature Genetics* 1:218–221. http://hdl.handle.net/1765/5803.

[67] Kamino, K., H. T. Orr, H. Payami, E. M. Wijsman, M. E. Alonso, S. M. Pulst, L. Anderson, S. O'dahl, E. Nemens, and J. a. White, 1992. Linkage and mutational analysis of familial Alzheimer disease kindreds for the APP gene region. *American journal of human genetics* 51:998–1014.

[68] Nilsberth, C., a. Westlind-Danielsson, C. B. Eckman, M. M. Condron, K. Axelman, C. Forsell, C. Stenh, J. Luthman, D. B. Teplow, S. G. Younkin, J. Näslund, and L. Lannfelt, 2001. The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation. *Nature neuroscience* 4:887–893.

[69] Levy, E., M. D. Carman, I. J. Fernandez-Madrid, M. D. Power, I. Lieberburg, S. G. van Duinen, G. T. Bots, W. Luyendijk, and B. Frangione, 1990. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. *Science* 248:1124–1126. http://www.jstor.org/stable/2874859.

[70] Van Broeckhoven, C., J. Haan, E. Bakker, J. a. Hardy, W. Van Hul, A. Wehnert, M. Vegter-Van der Vlis, and R. a. Roos, 1990. Amyloid beta protein precursor gene and hereditary cerebral hemorrhage with amyloidosis (Dutch). *Science* 248:1120–1122.

[71] Tagliavini, F., G. Rossi, A. Padovani, M. Magoni, G. Andora, M. Sgarzi, A. Bizzi, M. Savoiardo, F. Carella, and M. Morbin, 1999. A New APP Mutation Related to Hereditary Cerebral Haemorrhage. *Alzheimers Reports* 2:S28.

[72] Bugiani, O., G. Giaccone, G. Rossi, M. Mangieri, R. Capobianco, M. Morbin, G. Mazzoleni, C. Cupidi, G. Marcon, A. Giovagnoli, A. Bizzi, G. Di Fede, G. Puoti, F. Carella, A. Salmaggi, A. Romorini, G. M. Patruno, M. Magoni, A. Padovani, and F. Tagliavini, 2010. Hereditary Cerebral Hemorrhage With Amyloidosis Associated With the E693K Mutation of APP. *Archives of neurology* 67:987–995.

[73] Grabowski, T. J., H. S. Cho, J. P. G. Vonsattel, G. William Rebeck, and S. M. Greenberg, 2001. Novel amyloid precursor protein mutation in an Iowa family with dementia and severe cerebral amyloid angiopathy. *Annals of Neurology* 49:697–705.

[74] Tomiyama, T., T. Nagata, H. Shimada, R. Teraoka, A. Fukushima, H. Kanemitsu, H. Takuma, R. Kuwano, M. Imagawa, S. Ataka, Y. Wada, E. Yoshioka, T. Nishizaki, Y. Watanabe, and H. Mori, 2008. A new amyloid β variant favoring oligomerization in Alzheimer's-type dementia. *Annals of Neurology* 63:377–387.

[75] Ovchinnikova, O. Y., V. H. Finder, I. Vodopivec, R. M. Nitsch, and R. Glockshuber, 2011. The Osaka FAD mutation E22Δ leads to the formation of a previously unknown type of amyloid β fibrils and modulates Aβ neurotoxicity. *Journal of Molecular Biology* 408:780–791. http://dx.doi.org/10.1016/j.jmb.2011.02.049.

[76] Petkova, A. T., Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, and R. Tycko, 2002. A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proceedings of the National Academy of Sciences* 99:16742–16747.

[77] Paravastu, A. K., R. D. Leapman, W.-M. Yau, and R. Tycko, 2008. Molecular structural basis for polymorphism in Alzheimer's β-amyloid fibrils. *Proceedings of the National Academy of Sciences* 105:18349–18354.

[78] Bertini, I., L. Gonnelli, C. Luchinat, J. Mao, and A. Nesi, 2011. A new structural model of Aβ40 fibrils. *Journal of the American Chemical Society* 133:16013–22. http://www.ncbi.nlm.nih.gov/pubmed/21882806.

[79] Lopez del Amo, J. M., M. Schmidt, U. Fink, M. Dasari, M. Fändrich, and B. Reif, 2012. An Asymmetric Dimer as the Basic Subunit in Alzheimer's Disease Amyloid β Fibrils. *Angewandte Chemie International Edition* 51:6136–6139. http://doi.wiley.com/10.1002/anie.201200965.

[80] Lu, J.-X., W. Qiang, W.-M. Yau, C. D. Schwieters, S. C. Meredith, and R. Tycko, 2013. Molecular structure of β-amyloid fibrils in Alzheimer's disease brain tissue. *Cell* 154:1257–68. http://www.ncbi.nlm.nih.gov/pubmed/24034249.

[81] Niu, Z., W. Zhao, Z. Zhang, F. Xiao, X. Tang, and J. Yang, 2014. The Molecular Structure of Alzheimer β-Amyloid Fibrils Formed in the Presence of Phospholipid Vesicles. *Angewandte Chemie International Edition* 53:9294–9297.

[82] Sgourakis, N. G., W.-m. Yau, and W. Qiang, 2015. Modeling an In-Register, Parallel "Iowa" Aβ Fibril Structure Using Solid-State NMR Data from Labeled Samples with Rosetta. *Structure/Folding and Design* 23:216–227. http://dx.doi.org/10.1016/j.str.2014.10.022.

[83] Tiller, K. E., and P. M. Tessier, 2013. Lifting the veil on amyloid drug design. *eLife* 2. https://elifesciences.org/articles/01089.

[84] Huber, M., O. Y. Ovchinnikova, A. K. Schütz, R. Glockshuber, B. H. Meier, and A. Böckmann, 2014. Solid-state NMR sequential assignment of Osaka-mutant amyloid-beta (Aβ1-40 E22Δ) fibrils. *Biomolecular NMR Assignments* 1–8.

[85] Tarus, B., J. E. Straub, and D. Thirumalai, 2006. Dynamics of Asp23-Lys28 Salt-Bridge Formation in Aβ10-35 Monomers. *Journal of the American Chemical Society* 128:16159–16168.

[86] Reddy, G., J. E. Straub, and D. Thirumalai, 2009. Influence of preformed Asp23-Lys28 salt bridge on the conformational fluctuations of monomers and dimers of Aβ peptides with implications for rates of fibril formation. *Journal of Physical Chemistry B* 113:1162–1172.

[87] Spera, S., and A. Bax, 1991. Empirical correlation between protein backbone conformation and C alpha and C beta ^{13}C nuclear magnetic resonance chemical shifts. *Journal of the American Chemical Society* 113:5490–5492. http://pubs.acs.org/doi/abs/10.1021/ja00014a071.

[88] Wishart, D., and B. Sykes, 1994. The ^{13}C Chemical-Shift Index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *Journal of Biomolecular NMR* 4:171–180. http://link.springer.com/10.1007/BF00175245.

[89] Güntert, P., 2004. Automated NMR Structure Calculation With CYANA. *In* Protein NMR Techniques, Humana Press, New Jersey, 353–378. http://link.springer.com/10.1385/1-59259-809-9:353.

[90] Stroud, J. C., C. Liu, P. K. Teng, and D. Eisenberg, 2012. Toxic fibrillar oligomers of amyloid-β have cross-β structure. *Proceedings of the National Academy of Sciences* 109:7717–22. http://www.ncbi.nlm.nih.gov/pubmed/22547798.

[91] Shen, Y., F. Delaglio, G. Cornilescu, and A. Bax, 2009. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of Biomolecular NMR* 44:213–223.

[92] Paravastu, A. K., A. T. Petkova, and R. Tycko, 2006. Polymorphic Fibril Formation by Residues 10-40 of the Alzheimer's β-Amyloid Peptide. *Biophysical Journal* 90:4618–4629. https://www.sciencedirect.com/science/article/pii/S0006349506726364.

[93] Meites, J., 1977. The 1977 Nobel Prize in physiology or medicine. *Science* 198:594–6. http://www.ncbi.nlm.nih.gov/pubmed/17619283.

[94] Brazeau, P., W. Vale, R. Burgus, N. Ling, M. Butcher, J. Rivier, and R. Guillemin, 1973. Hypothalamic polypeptide that inhibits the secretion of immunoreactive pituitary growth hormone. *Science* 179:77–79.

[95] Krulich, L., A. Dhariwal, and S. McCann, 1968. Stimulatory and inhibitory effects of purified hypothalamic extracts on growth hormone release from rat pituitary in vitro. *Endocrinology* 83:783–790.

[96] Hernández, B., C. Carelli, Y.-M. Coïc, J. De Coninck, and M. Ghomi, 2009. Vibrational Analysis of Amino Acids and Short Peptides in Aqueous Media. V. The Effect of the Disulfide Bridge on the Structural Features of the Peptide Hormone Somatostatin-14. *The Journal of Physical Chemistry B* 113:12796–12803. http://pubs.acs.org/doi/abs/10.1021/jp904737v.

[97] Yazdanparast, R., P. Andrews, D. L. Smith, and J. E. Dixon, 1986. A new approach for detection and assignment of disulfide bonds in peptides. *Analytical Biochemistry* 153:348–353.

[98] Patel, Y., J. Epelbaum, D. Rubinow, C. Davis, R. Post, V. Schusdziarra, P. Maton, and R. Arakaki, 1992. Somatostatin, volume 4. Springer Science, Berlin, 1st edition.

[99] Rai, U., T. R. Thrimawithana, C. Valery, and S. A. Young, 2015. Therapeutic uses of somatostatin and its analogues: Current view and potential applications. *Pharmacology & Therapeutics* 152:98–110.

[100] Martinez, C. R., and B. L. Iverson, 2012. Rethinking the term "pi-stacking". *Chemical Science* 3:2191. http://xlink.rsc.org/?DOI=c2sc20045g.

[101] Makin, O. S., E. Atkins, P. Sikorski, J. Johansson, and L. C. Serpell, 2005. Molecular basis for amyloid fibril formation and stability. *Proceedings of the National Academy of Sciences* 102:315–320.

[102] Maji, S. K., M. H. Perrin, M. R. Sawaya, S. Jessberger, K. Vadodaria, R. A. Rissman, P. S. Singru, K. P. R. Nilsson, R. Simon, D. Schubert, D. Eisenberg, J. Rivier, P. Sawchenko, W. Vale, and R. Riek, 2009. Functional Amyloids As Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science* 325:328–332. http://science.sciencemag.org/content/325/5938/328.abstract.

[103] Anoop, A., S. Ranganathan, B. D. Dhaked, N. N. Jha, S. Pratihar, S. Ghosh, S. Sahay, S. Kumar, S. Das, M. Kombrabail, K. Agarwal, R. S. Jacob, P. Singru, P. Bhaumik, R. Padinhateeri, A. Kumar, and S. K. Maji, 2014. Elucidating the role of disulfide bond on amyloid formation and fibril reversibility of somatostatin-14: Relevance to its storage and secretion. *Journal of Biological Chemistry* 289:16884–16903.

[104] Davies, P., R. Katzman, and R. D. Terry, 1980. Reduced somatostatin-like immunoreactivity in cerebral cortex from cases of Alzheimer disease and Alzheimer senile dementa. *Nature* 288:279–280. http://www.nature.com/articles/288279a0.

[105] Gahete, M. D., A. Rubio, M. Durán-Prado, J. Avila, R. M. Luque, and J. P. Castaño, 2010. Expression of Somatostatin, Cortistatin, and Their Receptors, as well as Dopamine Receptors, but not of Neprilysin, are Reduced in the Temporal Lobe of Alzheimer's Disease Patients. *Journal of Alzheimer's Disease* 20:465–475. http://www.medra.org/servlet/aliasResolver?alias=iospress{&}doi=10.3233/JAD-2010-1385.

[106] Solarski, M., H. Wang, H. Wille, and G. Schmitt-Ulms, 2018. Somatostatin in Alzheimer's disease: A new Role for an Old Player. *Prion* 12:1–8. https://www.tandfonline.com/doi/full/10.1080/19336896.2017.1405207.

[107] Chen, S. G., V. Stribinskis, M. J. Rane, D. R. Demuth, E. Gozal, A. M. Roberts, R. Jagadapillai, R. Liu, K. Choe, B. Shivakumar, F. Son, S. Jin, R. Kerber, A. Adame, E. Masliah, and R. P. Friedland, 2016. Exposure to the Functional Bacterial Amyloid Protein Curli Enhances Alpha-Synuclein Aggregation in Aged Fischer 344 Rats and Caenorhabditis elegans. *Scientific Reports* 6:34477. http://www.nature.com/articles/srep34477.

[108] Hervás, R., L. Li, A. Majumdar, M. d. C. Fernández-Ramírez, J. R. Unruh, B. D. Slaughter, A. Galera-Prat, E. Santana, M. Suzuki, Y. Nagai, M. Bruix, S. Casas-Tintó, M. Menéndez, D. V. Laurents, K. Si, and M. Carrión-Vázquez, 2016. Molecular Basis of Orb2 Amyloidogenesis and Blockade of Memory Consolidation. *PLOS Biology* 14:e1002361. https://dx.plos.org/10.1371/journal.pbio.1002361.

[109] Haass, C., and D. J. Selkoe, 2007. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide. *Nature Reviews Molecular Cell Biology* 8:101–112. http://www.nature.com/articles/nrm2101.

[110] Mroczko, B., M. Groblewska, A. Litman-Zawadzka, J. Kornhuber, and P. Lewczuk, 2018. Amyloid β oligomers (AβOs) in Alzheimer's disease. *Journal of Neural Transmission* 125:177–191. http://link.springer.com/10.1007/s00702-017-1820-x.

[111] Shea, D., C.-C. Hsu, T. M. Bi, N. Paranjapye, M. C. Childers, J. Cochran, C. P. Tomberlin, L. Wang, D. Paris, J. Zonderman, G. Varani, C. D. Link, M. Mullan, and V. Daggett, 2019. α-Sheet secondary structure in amyloid β-peptide drives aggregation and toxicity in Alzheimer's disease. *Proceedings of the National Academy of Sciences* 116:8895–8900. http://www.ncbi.nlm.nih.gov/pubmed/31004062.

[112] Wang, H., L. D. Muiznieks, P. Ghosh, D. Williams, M. Solarski, A. Fang, A. Ruiz-Riquelme, R. Pomès, J. C. Watts, A. Chakrabartty, H. Wille, S. Sharpe, and G. Schmitt-Ulms, 2017. Somatostatin binds to the human amyloid β peptide and favors the formation of distinct oligomers. *eLife* 6. https://elifesciences.org/articles/28401.

[113] Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, 2000. The Protein Data Bank. *Nucleic Acids Research* 28:235–242. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/28.1.235.

[114] Wille, H., and J. Requena, 2018. The Structure of PrPSc Prions. *Pathogens* 7:20. http://www.mdpi.com/2076-0817/7/1/20.

[115] Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch, 2016. Protein identification and analysis tools on the ExPASy server. *In* J. M. Walker, editor, The proteomics protocols handbook, Springer Science & Business Media, chapter 52, 571–607.

[116] Seuring, C., J. Verasdonck, P. Ringler, R. Cadalbert, H. Stahlberg, A. Böckmann, B. H. Meier, and R. Riek, 2017. Amyloid Fibril Polymorphism: Almost Identical on the Atomic Level, Mesoscopically Very Different. *The Journal of Physical Chemistry B* 121:1783–1792. https://pubs.acs.org/doi/10.1021/acs.jpcb.6b10624.

[117] Siemer, A. B., C. Ritter, M. Ernst, R. Riek, and B. H. Meier, 2005. High-Resolution Solid-State NMR Spectroscopy of the Prion Protein HET-s in Its Amyloid Conformation. *Angewandte Chemie International Edition* 44:2441–2444. http://doi.wiley.com/10.1002/anie.200462952.

[118] Tycko, R., 2014. Physical and structural basis for polymorphism in amyloid fibrils. *Protein Science* 23:1528–1539. http://doi.wiley.com/10.1002/pro.2544.

[119] Gath, J., L. Bousset, B. Habenstein, R. Melki, B. H. Meier, and A. Böckmann, 2014. Yet another polymorph of α-synuclein: solid-state sequential assignments. *Biomolecular NMR Assignments* 8:395–404. http://link.springer.com/10.1007/s12104-013-9526-y.

[120] Bousset, L., L. Pieri, G. Ruiz-Arlandis, J. Gath, P. H. Jensen, B. Habenstein, K. Madiona, V. Olieric, A. Böckmann, B. H. Meier, and R. Melki, 2013. Structural and functional characterization of two alpha-synuclein strains. *Nature Communications* 4:2575. http://www.nature.com/articles/ncomms3575.

[121] Wiltzius, J. J. W., M. Landau, R. Nelson, M. R. Sawaya, M. I. Apostol, L. Goldschmidt, A. B. Soriaga, D. Cascio, K. Rajashankar, and D. Eisenberg, 2009. Molecular mechanisms for protein-encoded inheritance. *Nature Structural & Molecular Biology* 16:973–978. http://www.nature.com/articles/nsmb.1643.

[122] Wang, Y., and O. Jardetzky, 2002. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Science* 11:852–861.

[123] Astbury, W. T., S. Dickinson, and K. Bailey, 1935. The X-ray interpretation of denaturation and the structure of the seed globulins. *The Biochemical journal* 29:2351–2360. http://www.ncbi.nlm.nih.gov/pubmed/16745914.

[124] Soriaga, A. B., S. Sangwan, R. Macdonald, M. R. Sawaya, and D. Eisenberg, 2016. Crystal Structures of IAPP Amyloidogenic Segments Reveal a Novel Packing Motif of Out-of-Register Beta Sheets. *The Journal of Physical Chemistry B* 120:5810–5816. https://pubs.acs.org/doi/10.1021/acs.jpcb.5b09981.

[125] Garnier, J., J.-F. Gibrat, and B. Robson, 1996. GOR method for predicting protein secondary structure from amino acid sequence. *Methods in Enzymology* 266:540–553. https://www.sciencedirect.com/science/article/pii/S0076687996660340.

[126] Kabsch, W., and C. Sander, 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577–2637. http://doi.wiley.com/10.1002/bip.360221211.

[127] Frishman, D., and P. Argos, 1995. Knowledge-based protein secondary structure assignment. *Proteins: Structure, Function, and Genetics* 23:566–579. http://doi.wiley.com/10.1002/prot.340230412.

[128] Schrödinger LLC., 2015. The PyMOL Molecular Graphics System, Version 1.8.

[129] Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, 2004. UCSF Chimera-a visualization system for exploratory research and analysis. *Journal of computational chemistry* 25:1605–1612.

[130] Humphrey, W., A. Dalke, and K. Schulten, 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 14:33–38. https://www.sciencedirect.com/science/article/pii/0263785596000185.

[131] Verasdonck, J., 2017. Solid-State NMR Observations of Amyloid Fibrils. Ph.D. thesis, ETH Zurich. https://doi.org/10.3929/ethz-b-000161480.

[132] Anoop, A., S. Ranganathan, B. Das Dhaked, R. S. Jacob, A. Kumar, R. Padinhateeri, and S. Kumar Maji, 2013. Understanding the mechanism of somatostatin-14 amyloid formation in vitro. *Biophysical Journal* 104:50a. https://linkinghub.elsevier.com/retrieve/pii/S0006349512015615.

[133] Guerrero-Ferreira, R., N. M. Taylor, D. Mona, P. Ringler, M. E. Lauer, R. Riek, M. Britschgi, and H. Stahlberg, 2018. Cryo-EM structure of alpha-synuclein fibrils. *eLife* 7:1–18.

[134] Guerrero-Ferreira, R., N. M. Taylor, A.-A. Arteni, P. Kumari, D. Mona, P. Ringler, M. Britschgi, M. E. Lauer, J. Verasdock, R. Riek, R. Melki, B. H. Meier, A. Böckmann,

L. Bousset, and H. Stahlberg, 2019. Two new polymorphic structures of alpha-synuclein solved by cryo-electron microscopy. *bioRxiv* 1–35. https://www.biorxiv.org/content/10.1101/654582v1.

[135] Schuetz, A., C. Wasmer, B. Habenstein, R. Verel, J. Greenwald, R. Riek, A. Böckmann, and B. H. Meier, 2010. Protocols for the Sequential Solid-State NMR Spectroscopic Assignment of a Uniformly Labeled 25 kDa Protein: HET-s(1-227). *ChemBioChem* 11:1543–1551. http://doi.wiley.com/10.1002/cbic.201000124.

[136] Vranken, W. F., W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, and E. D. Laue, 2005. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Structure, Function, and Bioinformatics* 59:687–696.

[137] Stevens, T. J., R. H. Fogh, W. Boucher, V. A. Higman, F. Eisenmenger, B. Bardiaux, B.-J. van Rossum, H. Oschkinat, and E. D. Laue, 2011. A software framework for analysing solid-state MAS NMR data. *Journal of biomolecular NMR* 51:437–447.

[138] David, G., M.-L. Fogeron, M. Schledorn, R. Montserret, U. Haselmann, S. Penzel, A. Badillo, L. Lecoq, P. André, M. Nassal, R. Bartenschlager, B. H. Meier, and A. Böckmann, 2018. Structural studies of self-assembled subviral particles: Combining cell-free expression with 110 kHz MAS NMR spectroscopy. *Angewandte Chemie International Edition* 57:4787–4791. http://doi.wiley.com/10.1002/anie.201712091.

[139] Lecoq, L., M. Schledorn, S. Wang, S. Smith-Penzel, A. A. Malär, M. Callon, M. Nassal, B. H. Meier, and A. Böckmann, 2019. 100 kHz MAS Proton-Detected NMR Spectroscopy of Hepatitis B Virus Capsids. *Frontiers in Molecular Biosciences* 6:58. https://www.frontiersin.org/article/10.3389/fmolb.2019.00058/full.

[140] Wang, S., M.-L. Fogeron, M. Schledorn, M. Dujardin, S. Penzel, D. Burdette, J. M. Berke, M. Nassal, L. Lecoq, B. H. Meier, and A. Böckmann, 2019. Combining Cell-Free Protein Synthesis and NMR Into a Tool to Study Capsid Assembly Modulation. *Frontiers in Molecular Biosciences* 6:67. https://www.frontiersin.org/article/10.3389/fmolb.2019.00067/full.

[141] Vos, T., C. Allen, M. Arora, ..., M. Zhou, S. Zodpey, L. J. Zuhlke, and C. J. L. Murray, 2016. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* 388:1545–1602. https://www.sciencedirect.com/science/article/pii/S0140673616316786.

[142] Schweitzer, A., J. Horn, R. T. Mikolajczyk, G. Krause, and J. J. Ott, 2015. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. *The Lancet* 386:1546–1555. https://www.sciencedirect.com/science/article/pii/S014067361561412X.

[143] Yang, L., and M. Lu, 2018. Small Molecule Inhibitors of Hepatitis B Virus Nucleocapsid Assembly: A New Approach to Treat Chronic HBV Infection. *Current Medicinal Chemistry* 25:802–813. https://doi.org/10.2174/0929867324666170704121800.

[144] Lecoq, L., S. Wang, T. Wiegand, S. Bressanelli, M. Nassal, B. H. Meier, and A. Böckmann, 2018. Solid-state ^{13}C-^{15}N NMR resonance assignment of hepatitis B virus core protein. *Biomolecular NMR Assignments* 12:205–214. http://link.springer.com/10.1007/s12104-018-9810-y.

[145] Birnbaum, F., and M. Nassal, 1990. Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. *Journal of virology* 64:3319–30. http://www.ncbi.nlm.nih.gov/pubmed/2191149.

[146] Nassal, M., 1992. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *Journal of virology* 66:4107–16. http://www.ncbi.nlm.nih.gov/pubmed/1602535.

[147] Bartenschlager, R., M. Junker-Niepmann, and H. Schaller, 1990. The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. *Journal of virology* 64:5324–32. http://www.ncbi.nlm.nih.gov/pubmed/2214019.

[148] Lecoq, L., S. Wang, T. Wiegand, S. Bressanelli, M. Nassal, B. H. Meier, and A. Böckmann, 2018. Localizing Conformational Hinges by NMR: Where Do Hepatitis B Virus Core Proteins Adapt for Capsid Assembly? *ChemPhysChem* 19:1336–1340. http://doi.wiley.com/10.1002/cphc.201800211.

[149] David, G., 2019. Towards structural studies of Hepadnavirus subviral particles using wheat germ cell-free expression and solid-state NMR. Ph.D. thesis, Université de Lyon. http://www.theses.fr/2019LYSE1336.

[150] Wynne, S., R. Crowther, and A. Leslie, 1999. The Crystal Structure of the Human Hepatitis B Virus Capsid. *Molecular Cell* 3:771–780. https://www.sciencedirect.com/science/article/pii/S1097276501800095.

[151] Barbet-Massin, E., A. J. Pell, K. Jaudzems, W. T. Franks, J. S. Retel, S. Kotelovica, I. Akopjana, K. Tars, L. Emsley, H. Oschkinat, A. Lesage, and G. Pintacuda, 2013. Out-and-back ^{13}C-^{13}C scalar transfers in protein resonance assignment by proton-detected solid-state NMR under ultra-fast MAS. *Journal of Biomolecular NMR* 56:379–386. http://link.springer.com/10.1007/s10858-013-9757-3.

[152] Mandala, V. S., and M. Hong, 2019. High-sensitivity protein solid-state NMR spectroscopy. *Current Opinion in Structural Biology* https://www.sciencedirect.com/science/article/pii/S0959440X18300587.

[153] Webb, A. G., 1997. Radiofrequency microcoils in magnetic resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy* 31:1–42. https://linkinghub.elsevier.com/retrieve/pii/S0079656597000046.

[154] Samoson, A., T. Tuherm, J. Past, A. Reinhold, I. Heinmaa, T. Anupõld, M. E. Smith, and K. J. Pike, 2010. Fast Magic-Angle Spinning: Implications. *In* Encyclopedia of Magnetic Resonance, John Wiley & Sons, Ltd, Chichester, UK. http://doi.wiley.com/10.1002/9780470034590.emrstm1017.

[155] Crowther, R., N. Kiselev, B. Böttcher, J. Berriman, G. Borisova, V. Ose, and P. Pumpens, 1994. Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. *Cell* 77:943–950. https://www.sciencedirect.com/science/article/pii/0092867494901422.

[156] Cala-De Paepe, D., J. Stanek, K. Jaudzems, K. Tars, L. B. Andreas, and G. Pintacuda, 2017. Is protein deuteration beneficial for proton detected solid-state NMR at and above 100 kHz magic-angle spinning? *Solid State Nuclear Magnetic Resonance* 87:126–136. https://www.sciencedirect.com/science/article/pii/S0926204017300462.

[157] Linser, R., 2017. Solid-state NMR spectroscopic trends for supramolecular assemblies and protein aggregates. *Solid State Nuclear Magnetic Resonance* 87:45–53. https://www.sciencedirect.com/science/article/pii/S092620401730070X.

[158] Schubeis, T., T. Le Marchand, L. B. Andreas, and G. Pintacuda, 2018. ^{1}H magic-angle spinning NMR evolves as a powerful new tool for membrane proteins. *Journal of Magnetic Resonance* 287:140–152. https://www.sciencedirect.com/science/article/pii/S1090780717302884.

[159] Xue, K., R. Sarkar, C. Motz, S. Asami, D. C. R. Camargo, V. Decker, S. Wegner, Z. Tosner, and B. Reif, 2017. Limits of Resolution and Sensitivity of Proton Detected MAS Solid-State NMR Experiments at 111 kHz in Deuterated and Protonated Proteins. *Scientific Reports* 7:7444. http://www.nature.com/articles/s41598-017-07253-1.

[160] Kang, L., J. Pan, J. Wu, J. Hu, Q. Sun, and J. Tang, 2015. Anti-HBV Drugs: Progress, Unmet Needs, and New Hope. *Viruses* 7:4960–4977. http://www.mdpi.com/1999-4915/7/9/2854.

[161] Clark, D. N., and J. Hu, 2015. Hepatitis B virus reverse transcriptase - Target of current antiviral therapy and future drug development. *Antiviral Research* 123:132–137. https://www.sciencedirect.com/science/article/pii/S0166354215002223.

[162] Deres, K., C. H. Schröder, A. Paessens, S. Goldmann, H. J. Hacker, O. Weber, T. Krämer, U. Niewöhner, U. Pleiss, J. Stoltefuss, E. Graef, D. Koletzki, R. N. A. Masantschek, A. Reimann, R. Jaeger, R. Groß, B. Beckermann, K.-H. Schlemmer, D. Haebich, and

H. Rübsamen-Waigmann, 2003. Inhibition of Hepatitis B Virus Replication by Drug-Induced Depletion of Nucleocapsids. *Science* 299:893–896. https://science.sciencemag.org/content/299/5608/893.

[163] Lahlali, T., J. M. Berke, K. Vergauwen, A. Foca, K. Vandyck, F. Pauwels, F. Zoulim, and D. Durantel, 2018. Novel Potent Capsid Assembly Modulators Regulate Multiple Steps of the Hepatitis B Virus Life Cycle. *Antimicrobial agents and chemotherapy* 62:1–15. http://www.ncbi.nlm.nih.gov/pubmed/30012770.

[164] Nijampatnam, B., and D. C. Liotta, 2019. Recent advances in the development of HBV capsid assembly modulators. *Current Opinion in Chemical Biology* 50:73–79. https://www.sciencedirect.com/science/article/pii/S1367593118301741{#}!

[165] Williamson, M. P., 2013. Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy* 73:1–16. https://www.sciencedirect.com/science/article/pii/S0079656513000265.

[166] Berke, J. M., P. Dehertogh, K. Vergauwen, E. V. Damme, W. Mostmans, K. Vandyck, and F. Pauwels, 2017. Capsid Assembly Modulators Have a Dual Mechanism of Action in Primary Human Hepatocytes Infected with Hepatitis B Virus. *Antimicrobial Agents and Chemotherapy* 61:1–14. http://www.ncbi.nlm.nih.gov/pubmed/28584155.

[167] Seeger, C., F. Zoulim, and W. Mason, 2007. Hepadnaviruses. *Fields virology* 2:2977–3029.

[168] Böckmann, A., C. Gardiennet, R. Verel, A. Hunkeler, A. Loquet, G. Pintacuda, L. Emsley, B. H. Meier, and A. Lesage, 2009. Characterization of different water pools in solid-state NMR protein samples. *Journal of Biomolecular NMR* 45:319–327. http://link.springer.com/10.1007/s10858-009-9374-3.

[169] Fogeron, M.-L., A. Badillo, F. Penin, and A. Böckmann, 2017. Wheat Germ Cell-Free Overexpression for the Production of Membrane Proteins. *In* Methods in Molecular Biology, Humana Press, chapter 5, 91–108. http://link.springer.com/10.1007/978-1-4939-7151-0{_}5.

[170] Barbet-Massin, E., A. J. Pell, J. S. Retel, L. B. Andreas, K. Jaudzems, W. T. Franks, A. J. Nieuwkoop, M. Hiller, V. Higman, P. Guerry, A. Bertarello, M. J. Knight, M. Felletti, T. Le Marchand, S. Kotelovica, I. Akopjana, K. Tars, M. Stoppini, V. Bellotti, M. Bolognesi, S. Ricagno, J. J. Chou, R. G. Griffin, H. Oschkinat, A. Lesage, L. Emsley, T. Herrmann, and G. Pintacuda, 2014. Rapid proton-detected NMR assignment for proteins with fast magic angle spinning. *J Am Chem Soc* 136:12489–12497.

[171] Wang, S., S. Parthasarathy, Y. Xiao, Y. Nishiyama, F. Long, I. Matsuda, Y. Endo, T. Nemoto, K. Yamauchi, T. Asakura, M. Takeda, T. Terauchi, M. Kainosho, and Y. Ishii,

2015. Nano-mole scale sequential signal assignment by [1]H-detected protein solid-state NMR. *Chemical Communications* 51:15055–15058. http://xlink.rsc.org/?DOI=C5CC04618A.

[172] Lalli, D., M. N. Idso, L. B. Andreas, S. Hussain, N. Baxter, S. Han, B. F. Chmelka, and G. Pintacuda, 2017. Proton-Based Structural Analysis of a Heptahelical Transmembrane Protein in Lipid Bilayers. *Journal of the American Chemical Society* 139:13006–13012. http://pubs.acs.org/doi/10.1021/jacs.7b05269.

[173] Reckel, S., D. Gottstein, J. Stehle, F. Löhr, M.-K. Verhoefen, M. Takeda, R. Silvers, M. Kainosho, C. Glaubitz, J. Wachtveitl, F. Bernhard, H. Schwalbe, P. Güntert, and V. Dötsch, 2011. Solution NMR Structure of Proteorhodopsin. *Angewandte Chemie International Edition* 50:11942–11946. http://doi.wiley.com/10.1002/anie.201105648.

[174] Hagn, F., M. Etzkorn, T. Raschle, and G. Wagner, 2013. Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *Journal of the American Chemical Society* 135:1919–1925. https://pubs.acs.org/doi/10.1021/ja310901f.

[175] Fogeron, M.-L., V. Jirasko, S. Penzel, D. Paul, R. Montserret, C. Danis, D. Lacabanne, A. Badillo, J. Gouttenoire, D. Moradpour, R. Bartenschlager, F. Penin, B. H. Meier, and A. Böckmann, 2016. Cell-free expression, purification, and membrane reconstitution for NMR studies of the nonstructural protein 4B from hepatitis C virus. *Journal of Biomolecular NMR* 65:87–98. http://link.springer.com/10.1007/s10858-016-0040-2.

[176] Wiegand, T., M. Schledorn, A. A. Malär, R. Cadalbert, A. Däpp, L. Terradot, B. H. Meier, and A. Böckmann, 2019. Nucleotide binding modes in a motor protein revealed by [31]P- and ^{1}H-detected MAS solid-state NMR. *ChemBioChem* 21:324–330. https://onlinelibrary.wiley.com/doi/abs/10.1002/cbic.201900439.

[177] Bazin, A., M. V. Cherrier, I. Gutsche, J. Timmins, and L. Terradot, 2015. Structure and primase-mediated activation of a bacterial dodecameric replicative helicase. *Nucleic Acids Research* 43:8564–8576.

[178] Spies, M., 2012. DNA helicases and DNA motor proteins. Springer Science & Business Media, New York, NY.

[179] Levin, M. K., and S. S. Patel, 2004. Molecular Motors. Wiley-VCH Verlag GmbH & Co, KGaA, Weinheim, FRG, Weinheim.

[180] Kollman, P. A., 1977. Noncovalent interactions. *Accounts of Chemical Research* 10:365–371. http://pubs.acs.org/doi/abs/10.1021/ar50118a003.

[181] Lejeune, D., N. Delsaux, B. Charloteaux, A. Thomas, and R. Brasseur, 2005. Protein-nucleic acid recognition: Statistical analysis of atomic interactions and influence of DNA structure. *Proteins: Structure, Function, and Bioinformatics* 61:258–271. http://doi.wiley.com/10.1002/prot.20607.

[182] Grimme, S., 2008. Do special noncovalent π-π stacking interactions really exist? *Angewandte Chemie International Edition* 47:3430–3434. http://doi.wiley.com/10.1002/anie.200705157.

[183] Luscombe, N. M., R. A. Laskowski, and J. M. Thornton, 2001. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Research* 29:2860–2874. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/29.13.2860.

[184] Itsathitphaisarn, O., R. A. Wing, W. K. Eliason, J. Wang, and T. A. Steitz, 2012. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* 151:267–277. http://dx.doi.org/10.1016/j.cell.2012.09.014.

[185] Jehle, S., M. Falb, J. P. Kirkpatrick, H. Oschkinat, B.-J. van Rossum, G. Althoff, and T. Carlomagno, 2010. Intermolecular protein-RNA interactions revealed by 2D ^{31}P-^{15}N magic angle spinning solid-state NMR spectroscopy. *Journal of the American Chemical Society* 132:3842–3846. https://pubs.acs.org/doi/10.1021/ja909723f.

[186] Asami, S., M. Rakwalska-Bange, T. Carlomagno, and B. Reif, 2013. Protein-RNA Interfaces Probed by ^{1}H-Detected MAS Solid-State NMR Spectroscopy. *Angewandte Chemie International Edition* 52:2345–2349. http://doi.wiley.com/10.1002/anie.201208024.

[187] Huang, W., G. Varani, and G. P. Drobny, 2011. Interactions of protein side chains with RNA defined with REDOR solid state NMR. *Journal of Biomolecular NMR* 51:347–356. http://link.springer.com/10.1007/s10858-011-9573-6.

[188] Huang, W., G. Varani, and G. P. Drobny, 2010. ^{13}C/^{15}N-^{19}F intermolecular REDOR NMR study of the interaction of TAR RNA with Tat peptides. *Journal of the American Chemical Society* 132:17643–17645. https://pubs.acs.org/doi/10.1021/ja1051439.

[189] Morag, O., G. Abramov, and A. Goldbourt, 2014. Complete Chemical Shift Assignment of the ssDNA in the Filamentous Bacteriophage fd Reports on Its Conformation and on Its Interface with the Capsid Shell. *Journal of the American Chemical Society* 136:2292–2301. http://pubs.acs.org/doi/10.1021/ja412178n.

[190] Jeffrey, G. A., Y. Yeon, and IUCr, 1986. The correlation between hydrogen-bond lengths and proton chemical shifts in crystals. *Acta Crystallographica Section B Structural Science* 42:410–413. http://scripts.iucr.org/cgi-bin/paper?S0108768186098038.

[191] Berglund, B., and R. W. Vaughan, 1980. Correlations between proton chemical shift tensors, deuterium quadrupole couplings, and bond distances for hydrogen bonds in solids. *The Journal of Chemical Physics* 73:2037–2043. http://aip.scitation.org/doi/10.1063/1.440423.

[192] Barfield, M., 2002. Structural Dependencies of Interresidue Scalar Coupling h3JNC' and Donor [1]H Chemical Shifts in the Hydrogen Bonding Regions of Proteins. *Journal of the American Chemical Society* 124:4158–4168. https://pubs.acs.org/doi/abs/10.1021/ja012674v.

[193] Parker, L. L., A. R. Houk, and J. H. Jensen, 2006. Cooperative Hydrogen Bonding Effects Are Key Determinants of Backbone Amide Proton Chemical Shifts in Proteins. *Journal of the American Chemical Society* 128:9863–9872. https://pubs.acs.org/doi/abs/10.1021/ja0617901.

[194] Harris, R. K., R. E. Wasylishen, and M. J. Duer, 2009. NMR crystallography. Wiley.

[195] Wagner, G., A. Pardi, and K. Wuethrich, 1983. Hydrogen bond length and proton NMR chemical shifts in proteins. *Journal of the American Chemical Society* 105:5948–5949. http://pubs.acs.org/doi/abs/10.1021/ja00356a056.

[196] Asakura, T., K. Taoka, M. Demura, and M. Williamson, 1995. The relationship between amide proton chemical shifts and secondary structure in proteins. *Journal of Biomolecular NMR* 6:227–236. http://link.springer.com/10.1007/BF00197804.

[197] Zhou, N. E., B. Y. Zhu, B. D. Sykes, and R. S. Hodges, 1992. Relationship between amide proton chemical shifts and hydrogen bonding in amphipathic .alpha.-helical peptides. *Journal of the American Chemical Society* 114:4320–4326. http://pubs.acs.org/doi/abs/10.1021/ja00037a042.

[198] Cordier, F., and S. Grzesiek, 1999. Direct Observation of Hydrogen Bonds in Proteins by Interresidue 3hJNC' Scalar Couplings. *Journal of the American Chemical Society* 121:1601–1602. https://pubs.acs.org/doi/10.1021/ja983945d.

[199] Perkins, S. J., and K. Wüthrich, 1979. Ring current effects in the conformation dependent NMR chemical shifts of aliphatic protons in the basic pancreatic trypsin inhibitor. *Biochimica et Biophysica Acta (BBA) - Protein Structure* 576:409–423. https://www.sciencedirect.com/science/article/pii/0005279579904161.

[200] Sitkoff, D., and D. A. Case, 1997. Density Functional Calculations of Proton Chemical Shifts in Model Peptides. *Journal of the American Chemical Society* 119:12262–12273. https://pubs.acs.org/doi/abs/10.1021/ja9721430.

[201] Case, D., 1995. Calibration of ring-current effects in proteins and nucleic acids. *Journal of Biomolecular NMR* 6:341–346. http://link.springer.com/10.1007/BF00197633.

[202] Wishart, D., B. Sykes, and F. Richards, 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of Molecular Biology* 222:311–333. https://www.sciencedirect.com/science/article/pii/002228369190214Q.

[203] Walker, J., M. Saraste, M. Runswick, and N. Gay, 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO Journal* 1:945–951. http://doi.wiley.com/10.1002/j.1460-2075.1982.tb01276.x.

[204] Wilson, K. A., J. L. Kellie, and S. D. Wetmore, 2014. DNA-protein π-interactions in nature: abundance, structure, composition and strength of contacts between aromatic amino acids and DNA nucleobases or deoxyribose sugar. *Nucleic Acids Research* 42:6726–6741. http://academic.oup.com/nar/article/42/10/6726/2435280/DNAprotein-πinteractions-in-nature-abundance.

[205] Mackenzie, H. W., and D. F. Hansen, 2017. A ^{13}C-detected ^{15}N double-quantum NMR experiment to probe arginine side-chain guanidinium ^{15}Nη chemical shifts. *Journal of Biomolecular NMR* 69:123–132. http://link.springer.com/10.1007/s10858-017-0137-2.

[206] Esadze, A., C. Chen, L. Zandarashvili, S. Roy, B. M. Pettitt, and J. Iwahara, 2016. Changes in conformational dynamics of basic side chains upon protein-DNA association. *Nucleic Acids Research* 44:6961–6970. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw531.

[207] Raman, B., C. Guarnaccia, K. Nadassy, S. Zakhariev, A. Pintar, F. Zanuttin, D. Frigyes, C. Acatrinei, A. Vindigni, G. Pongor, and S. Pongor, 2001. Nomega-arginine dimethylation modulates the interaction between a Gly/Arg-rich peptide from human nucleolin and nucleic acids. *Nucleic Acids Research* 29:3377–3384. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/29.16.3377.

[208] Frigyes, D., F. Alber, S. Pongor, and P. Carloni, 2001. Arginine-phosphate salt bridges in protein-DNA complexes: a Car-Parrinello study. *Journal of Molecular Structure: THEOCHEM* 574:39–45. https://www.sciencedirect.com/science/article/pii/S0166128001003682.

[209] Wiegand, T., R. Cadalbert, D. Lacabanne, J. Timmins, L. Terradot, A. Böckmann, and B. H. Meier, 2019. The conformational changes coupling ATP hydrolysis and translocation in a bacterial DnaB helicase. *Nature Communications* 10:31. http://www.nature.com/articles/s41467-018-07968-3.

[210] Wiegand, T., R. Cadalbert, C. von Schroetter, F. H.-T. Allain, and B. H. Meier, 2018. Segmental isotope labelling and solid-state NMR of a 12 x 59 kDa motor protein:

identification of structural variability. *Journal of Biomolecular NMR* 71:237–245. http://link.springer.com/10.1007/s10858-018-0196-z.

[211] Lange, A., S. Luca, and M. Baldus, 2002. Structural Constraints from Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids. *Journal of the American Chemical Society* 124:9704–9705. https://pubs.acs.org/doi/abs/10.1021/ja026691b.

[212] Stelter, M., I. Gutsche, U. Kapp, A. Bazin, G. Bajic, G. Goret, M. Jamin, J. Timmins, and L. Terradot, 2012. Architecture of a Dodecameric Bacterial Replicative Helicase. *Structure* 20:554–564. https://www.sciencedirect.com/science/article/pii/S0969212612000512.

[213] Sepuru, K. M., J. Iwahara, and K. Rajarathnam, 2018. Direct detection of lysine side chain NH3+ in protein-heparin complexes using NMR spectroscopy. *The Analyst* 143:635–638. http://xlink.rsc.org/?DOI=C7AN01406F.

[214] Shalaeva, D. N., D. A. Cherepanov, M. Y. Galperin, A. V. Golovin, and A. Y. Mulkidjanian, 2018. Evolution of cation binding in the active sites of P-loop nucleoside triphosphatases in relation to the basic catalytic mechanism. *eLife* 7:1–35. https://elifesciences.org/articles/37373.

[215] Jin, Y., R. W. Molt, and G. M. Blackburn, 2017. Metal Fluorides: Tools for Structural and Computational Analysis of Phosphoryl Transfer Enzymes. *In* H. Jessen, editor, Phosphate Labeling in Chemical Biology, Springer International Publishing, 35–65. http://link.springer.com/10.1007/978-3-319-60357-5{_}2.

[216] Ogura, T., S. W. Whiteheart, and A. J. Wilkinson, 2004. Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. *Journal of Structural Biology* 146:106–112. https://www.sciencedirect.com/science/article/pii/S1047847703002843.

[217] Gallivan, J. P., and D. A. Dougherty, 1999. Cation-π interactions in structural biology. *Proceedings of the National Academy of Sciences* 96:9459–9464. https://www.pnas.org/content/96/17/9459.short.

[218] Burley, S. K., and G. A. Petsko, 1985. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* 229:23–8. http://www.ncbi.nlm.nih.gov/pubmed/3892686.

[219] Ma, J., and D. A. Dougherty, 1997. The Cation-π Interaction. *Chemical Reviews* 97:1303–1324. https://pubs.acs.org/doi/abs/10.1021/cr9603744.

[220] Studier, F. W., 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 41:207–234. https://www.sciencedirect.com/science/article/pii/S1046592805000264.

[221] Gardiennet, C., A. K. Schütz, A. Hunkeler, B. Kunert, L. Terradot, A. Böckmann, and B. H. Meier, 2012. A sedimented sample of a 59 kDa dodecameric helicase yields high-resolution solid-state NMR spectra. *Angewandte Chemie International Edition* 51:7855–7858. http://doi.wiley.com/10.1002/anie.201200779.

[222] Gor'kov, P. L., R. Witter, E. Y. Chekmenev, F. Nozirov, R. Fu, and W. W. Brey, 2007. Low-E probe for ^{19}F-^{1}H NMR of dilute biological solids. *Journal of Magnetic Resonance* 189:182–189. https://www.sciencedirect.com/science/article/pii/S1090780707002765.

[223] Fogh, R., J. Ionides, E. Ulrich, W. Boucher, W. Vranken, J. P. Linge, M. Habeck, W. Rieping, T. Bhat, J. Westbrook, K. Henrick, G. Gilliland, H. Berman, J. Thornton, M. Nilges, J. Markley, and E. Laue, 2002. The CCPN project: an interim report on a data model for the NMR community. *Nature Structural Biology* 9:416–418. http://www.nature.com/doifinder/10.1038/nsb0602-416.

[224] Torosyan, A., T. Wiegand, M. Schledorn, D. Klose, P. Günthert, A. Böckmann, and B. H. Meier, 2019. Including protons in solid-state NMR resonance assignment and secondary structure analysis: The example of RNA polymerase II subunits Rpo4/7. *Frontiers in Molecular Biosciences* 6:1–8. https://www.frontiersin.org/articles/10.3389/fmolb.2019.00100/abstract.

[225] Cady, S. D., T. V. Mishanina, and M. Hong, 2009. Structure of Amantadine-Bound M2 Transmembrane Peptide of Influenza A in Lipid Bilayers from Magic-Angle-Spinning Solid-State NMR: The Role of Ser31 in Amantadine Binding. *Journal of Molecular Biology* 385:1127–1141. https://www.sciencedirect.com/science/article/pii/S0022283608014538.

[226] Spadaccini, R., H. Kaur, J. Becker-Baldus, and C. Glaubitz, 2018. The effect of drug binding on specific sites in transmembrane helices 4 and 6 of the ABC exporter MsbA studied by DNP-enhanced solid-state NMR. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1860:833–840. https://www.sciencedirect.com/science/article/pii/S0005273617303358.

[227] Elkins, M. R., J. K. Williams, M. D. Gelenter, P. Dai, B. Kwon, I. V. Sergeyev, B. L. Pentelute, and M. Hong, 2017. Cholesterol-binding site of the influenza M2 protein in lipid bilayers from solid-state NMR. *Proceedings of the National Academy of Sciences* 114:12946–12951. https://www.pnas.org/content/114/49/12946?collection={&}utm{_}source=TrendMD{&}utm{_}medium=cpc{&}utm{_}campaign=Proc{_}Natl{_}Acad{_}Sci{_}U{_}S{_}A{_}TrendMD{_}1.

[228] Henzler-Wildman, K., and D. Kern, 2007. Dynamic personalities of proteins. *Nature* 450:964–972. http://www.nature.com/doifinder/10.1038/nature06522.

[229] Lacabanne, D., C. Orelle, L. Lecoq, B. Kunert, C. Chuilon, T. Wiegand, S. Ravaud, J.-M. Jault, B. H. Meier, and A. Böckmann, 2019. Flexible-to-rigid transition is central for substrate transport in the ABC transporter BmrA from Bacillus subtilis. *Communications Biology* 2:149. http://www.nature.com/articles/s42003-019-0390-x.

[230] Quinn, C. M., M. Wang, M. P. Fritz, B. Runge, J. Ahn, C. Xu, J. R. Perilla, A. M. Gronenborn, and T. Polenova, 2018. Dynamic regulation of HIV-1 capsid interaction with the restriction factor TRIM5α identified by magic-angle spinning NMR and molecular dynamics simulations. *Proceedings of the National Academy of Sciences* 115:11519–11524. http://www.ncbi.nlm.nih.gov/pubmed/30333189.

[231] Schütz, A. K., S. Hornemann, M. A. Wälti, L. Greuter, C. Tiberi, R. Cadalbert, M. Gantner, R. Riek, P. Hammarström, K. P. R. Nilsson, A. Böckmann, A. Aguzzi, and B. H. Meier, 2018. Binding of Polythiophenes to Amyloids: Structural Mapping of the Pharmacophore. *ACS Chemical Neuroscience* 9:475–481. http://pubs.acs.org/doi/10.1021/acschemneuro.7b00397.

[232] Vasa, S. K., H. Singh, P. Rovó, and R. Linser, 2018. Dynamics and Interactions of a 29 kDa Human Enzyme Studied by Solid-State NMR. *The Journal of Physical Chemistry Letters* 9:1307–1311. http://pubs.acs.org/doi/10.1021/acs.jpclett.8b00110.

[233] Andreas, L. B., K. Jaudzems, J. Stanek, D. Lalli, A. Bertarello, T. Le Marchand, D. Cala-De Paepe, S. Kotelovica, I. Akopjana, B. Knott, S. Wegner, F. Engelke, A. Lesage, L. Emsley, K. Tars, T. Herrmann, and G. Pintacuda, 2016. Structure of fully protonated proteins by proton-detected magic-angle spinning NMR. *Proceedings of the National Academy of Sciences* 113:9187–92. http://www.ncbi.nlm.nih.gov/pubmed/27489348.

[234] Stanek, J., L. B. Andreas, K. Jaudzems, D. Cala, D. Lalli, A. Bertarello, T. Schubeis, I. Akopjana, S. Kotelovica, K. Tars, A. Pica, S. Leone, D. Picone, Z.-Q. Xu, N. E. Dixon, D. Martinez, M. Berbon, N. El Mammeri, A. Noubhani, S. Saupe, B. Habenstein, A. Loquet, and G. Pintacuda, 2016. NMR Spectroscopic Assignment of Backbone and Side-Chain Protons in Fully Protonated Proteins: Microcrystals, Sedimented Assemblies, and Amyloid Fibrils. *Angewandte Chemie International Edition* 55:15504–15509. http://doi.wiley.com/10.1002/anie.201607084.

[235] Knight, M. J., A. L. Webber, A. J. Pell, P. Guerry, E. Barbet-Massin, I. Bertini, I. C. Felli, L. Gonnelli, R. Pierattelli, L. Emsley, A. Lesage, T. Herrmann, and G. Pintacuda, 2011. Fast resonance assignment and fold determination of human superoxide dismutase by high-resolution proton-detected solid-state MAS NMR spectroscopy. *Angew Chem Int Ed Engl* 50:11697–11701.

[236] Lewandowski, J. R., J.-N. Dumez, Ü. Akbey, S. Lange, L. Emsley, and H. Oschkinat, 2011. Enhanced Resolution and Coherence Lifetimes in the Solid-State NMR Spectroscopy of Perdeuterated Proteins under Ultrafast Magic-Angle Spinning. *The Journal*

of Physical Chemistry Letters 2:2205–2211. http://pubs.acs.org/doi/10.1021/jz200844n.

[237] Reif, B., 2018. Proton-Detection in Biological MAS Solid-State NMR Spectroscopy, Springer International Publishing, chapter 45, 879–910. https://doi.org/10.1007/978-3-319-28388-3{_}69.

[238] Böckmann, A., M. Ernst, and B. H. Meier, 2015. Spinning proteins, the faster, the better? *Journal of Magnetic Resonance* 253:71–79.

[239] Malär, A. A., S. Smith-Penzel, G.-M. Camenisch, T. Wiegand, A. Samoson, A. Böckmann, M. Ernst, and B. H. Meier, 2019. Quantifying proton NMR coherent linewidth in proteins under fast MAS conditions: a second moment approach. *Physical Chemistry Chemical Physics* 21:18850–18865. http://xlink.rsc.org/?DOI=C9CP03414E.

[240] Lin, Y.-L., Y.-S. Cheng, C.-I. Ho, Z.-H. Guo, S.-J. Huang, M.-L. Org, A. Oss, A. Samoson, and J. C. C. Chan, 2018. Preparation of fibril nuclei of beta-amyloid peptides in reverse micelles. *Chemical Communications* 54:10459–10462. http://xlink.rsc.org/?DOI=C8CC05882B.

[241] Massiot, D., F. Fayon, M. Capron, I. King, S. Le Calvé, B. Alonso, J.-O. Durand, B. Bujoli, Z. Gan, and G. Hoatson, 2002. Modelling one- and two-dimensional solid-state NMR spectra. *Magnetic Resonance in Chemistry* 40:70–76. http://doi.wiley.com/10.1002/mrc.984.

[242] Ouhammouch, M., F. Werner, R. O. J. Weinzierl, and E. P. Geiduschek, 2004. A fully recombinant system for activator-dependent archaeal transcription. *The Journal of biological chemistry* 279:51719–21. http://www.ncbi.nlm.nih.gov/pubmed/15485836.

[243] Klose, D., J. P. Klare, D. Grohmann, C. W. M. Kay, F. Werner, and H.-J. Steinhoff, 2012. Simulation vs. Reality: A Comparison of In Silico Distance Predictions with DEER and FRET Measurements. *PLoS ONE* 7:e39492. http://dx.plos.org/10.1371/journal.pone.0039492.

[244] Penzel, S., A. A. Smith, M. Ernst, and B. H. Meier, 2018. Setting the magic angle for fast magic-angle spinning probes. *Journal of Magnetic Resonance* 293:115–122. https://www.sciencedirect.com/science/article/pii/S1090780718301575.

[245] Zhou, D. H., and C. M. Rienstra, 2008. High-performance solvent suppression for proton detected solid-state NMR. *Journal of Magnetic Resonance* 192:167–172. https://www.sciencedirect.com/science/article/pii/S1090780708000438.

[246] Shaka, A., J. Keeler, T. Frenkiel, and R. Freeman, 1983. An improved sequence for broadband decoupling: WALTZ-16. *Journal of Magnetic Resonance (1969)* 52:335–338. https://www.sciencedirect.com/science/article/pii/002223648390207X.

[247] Zhou, Z., R. Kümmerle, X. Qiu, D. Redwine, R. Cong, A. Taha, D. Baugh, and B. Win-niford, 2007. A new decoupling method for accurate quantification of polyethylene copolymer composition and triad sequence distribution with 13C NMR. *Journal of Magnetic Resonance* 187:225–233. https://www.sciencedirect.com/science/article/pii/S1090780707001504.

[248] Thakur, R. S., N. D. Kurur, and P. Madhu, 2006. Swept-frequency two-pulse phase mod-ulation for heteronuclear dipolar decoupling in solid-state NMR. *Chemical Physics Letters* 426:459–463. https://www.sciencedirect.com/science/article/pii/S000926140600830X.

[249] Nazarenko, A. Y., 2019. Construction of Meldonium (3-(1,1,1-Trimethylhydrazin-1-ium-2-yl)propanoate) Crystals: An X-ray View. *Journal of Chemical Crystallography* 1–7. http://link.springer.com/10.1007/s10870-019-00769-6.

[250] Momma, K., and F. Izumi, 2011. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *Journal of Applied Crystallography* 44:1272–1276. http://scripts.iucr.org/cgi-bin/paper?S0021889811038970.

[251] Malde, A. K., L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink, and A. E. Mark, 2011. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *Journal of Chemical Theory and Computation* 7:4026–4037. https://pubs.acs.org/doi/10.1021/ct200196m.

[252] Abraham, R. J., and M. Mobli, 2008. Modelling ^{1}H NMR Spectra of Organic Compounds : Theory, Applications and NMR Prediction Software. John Wiley & Sons.

[253] Smith, A. A., 2017. INFOS: spectrum fitting software for NMR analysis. *Jour-nal of Biomolecular NMR* 67:77–94. http://link.springer.com/10.1007/s10858-016-0085-2.

[254] Berjanskii, M. V., and D. S. Wishart, 2005. A Simple Method To Predict Protein Flex-ibility Using Secondary Chemical Shifts. *Journal of the American Chemical Society* 127:14970–14971. https://pubs.acs.org/doi/abs/10.1021/ja054842f.

[255] Gottlieb, H. E., V. Kotlyar, and A. Nudelman, 1997. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *Journal of Organic Chemistry* 62:7512–7515. https://pubs.acs.org/doi/abs/10.1021/jo971176v.

[256] Takai, K., T. Sawasaki, and Y. Endo, 2010. Practical cell-free protein synthesis system using purified wheat embryos. *Nature Protocols* 5:227–238. http://www.nature.com/articles/nprot.2009.207.

[257] Heger-Stevic, J., P. Kolb, A. Walker, and M. Nassal, 2018. Displaying Whole-Chain Proteins on Hepatitis B Virus Capsid-Like Particles. *In* Methods in Molecular Biology, Humana Press, chapter 33, 503–531. http://link.springer.com/10.1007/978-1-4939-7808-3{_}33.

[258] Fogeron, M.-L., A. Badillo, V. Jirasko, J. Gouttenoire, D. Paul, L. Lancien, D. Moradpour, R. Bartenschlager, B. H. Meier, F. Penin, and A. Böckmann, 2015. Wheat germ cell-free expression: Two detergents with a low critical micelle concentration allow for production of soluble HCV membrane proteins. *Protein Expression and Purification* 105:39–46. https://www.sciencedirect.com/science/article/pii/S1046592814002289.

[259] Makino, S.-i., E. T. Beebe, J. L. Markley, and B. G. Fox, 2014. Cell-Free Protein Synthesis for Functional and Structural Studies. *In* Methods in Molecular Biology, Humana Press, chapter 11, 161–178. http://link.springer.com/10.1007/978-1-62703-691-7{_}11.

[260] Guo, C., G. Hou, X. Lu, B. O'Hare, J. Struppe, and T. Polenova, 2014. Fast magic angle spinning NMR with heteronucleus detection for resonance assignments and structural characterization of fully protonated proteins. *Journal of Biomolecular NMR* 60:219–229. http://link.springer.com/10.1007/s10858-014-9870-y.

[261] Struppe, J., C. M. Quinn, M. Lu, M. Wang, G. Hou, X. Lu, J. Kraus, L. B. Andreas, J. Stanek, D. Lalli, A. Lesage, G. Pintacuda, W. Maas, A. M. Gronenborn, and T. Polenova, 2017. Expanding the horizons for structural analysis of fully protonated protein assemblies by NMR spectroscopy at MAS frequencies above 100 kHz. *Solid State Nuclear Magnetic Resonance* 87:117–125. https://www.sciencedirect.com/science/article/pii/S0926204017300486.

Curriculum vitae

Personal data

Name	Maarten Schledorn
Date of Birth	February 23rd, 1988
Place of Birth	Enter
Citizen of	Kingdom of the Netherlands

Education

2015 – 2019   **Doctoral studies**
Laboratory of Physical Chemistry, ETH Zürich, Switzerland
Group of Prof. Dr. Beat H. Meier, *Solid-State Nuclear Magnetic Resonance*
Thesis: "Solid-state NMR studies of large protein assemblies"

2014 – 2015   **Research internship**
Laboratory of Physical Chemistry, ETH Zürich, Switzerland
Group of Prof. Dr. Beat H. Meier, *Solid-State Nuclear Magnetic Resonance*

2012 – 2014   **MSc Medical biotechnology**
KTH Royal Institute of Technology, Sweden
Thesis: "Bacterial transmembrane glycosyltransferase *XCGT*"

2007 – 2012   **BSc Interdisciplinary degree - Profile: biochemistry**
University College Maastricht, The Netherlands
Thesis: "Biochemical reduction of dimensionality in blood coagulation"

2007 – 2012   **BA Violin**
Maastricht Conservatory, The Netherlands
Thesis: "Dyslexia in the musician's teaching practice"

2000 – 2006   **High school**
Pius X College Almelo, The Netherlands

Professional experience

2015 – 2019   **Researcher - Solid-state NMR**
Department of Chemistry and Applied Biosciences, ETH Zürich, Switzerland

2015 – 2019   **Teaching Assistant - Physical chemistry**
Department of Chemistry and Applied Biosciences, ETH Zürich, Switzerland

2013 – 2014   **Researcher - X-ray crystallography**
Medical Biochemistry and Biophysics Dept., Karolinska Institutet, Sweden

2011 – 2012   **Intern Violin Teacher**
Eindhoven Arts Centre, Eindhoven, Netherlands

2008 – 2012   **Historical Tour Guide**
VVV - The Dutch National Tourist Agency, Maastricht, The Netherlands

Scientific career

Conferences: *EUROMAR / ISMAR*, Berlin, Germany, 2019
ICMRBS, Dublin, Ireland, 2018
ISMAR / Rocky Mountain conference on EPR, Québec City, Canada, 2017
Conference on Protein Structure and Function, Tällberg, Sweden, 2016
Ampère Biological Solid-State NMR School, Palma de Mallorca, Spain, 2016
Conference on Protein Structure and Function, Tällberg, Sweden, 2014

Awards: *Poster award*, EUROMAR 2019, Berlin, Germany
Student travel stipend, ISMAR 2017, Québec City, Canada
Poster award, ISMAR 2017, Québec City, Canada
Best oral presentation of a young scientist, CPSF 2016, Tällberg, Sweden

Presentations: *Speedy-spinning spectroscopy* AMP Events 2019, Zürich, Switzerland
Kernspinresonantie in vaste stof Radboud visit 2018, Zürich, Switzerland
Magnete, Moleküle und der magische Winkel, TOL 2017, Zürich, Switzerland
Mol. struc. of SST-14 in the amyloid form CPSF 2016, Tallberg, Sweden

Languages

Dutch:	native		
Swedish:	fluent	English:	fluent
German:	good	Schwiizertüütsch:	intermediate
French:	intermediate	Italian:	basic

Publications

1. Torosyan, A.[#], Wiegand, T.[#], Schledorn, M.[#], Klose, D., Günthert, P., Böckmann, A., & Meier, B. H. (2019). Including protons in solid-state NMR resonance assignment and secondary structure analysis: The example of RNA polymerase II subunits Rpo4/7. *Frontiers in Molecular Biosciences*, 6, 100.

2. Lecoq, L.[#], Schledorn, M.[#], Wang, S., Penzel, S., Malär, A.A., Callon, M., ... & Böckmann, A. (2019). 100 kHz MAS proton-detected NMR spectroscopy of hepatitis B virus capsids. *Frontiers in molecular biosciences*, 6(58).

3. Wiegand, T., Schledorn, M., Malär, A. A., Cadalbert, R., Däpp, A., Terradot, L., Meier, B. H., & Böckmann, A. (2019) Nucleotide binding modes in a motor protein revealed by ^{31}P- and ^{1}H-detected MAS solid-state NMR. *ChemBioChem*, Volume 20.

4. Wang, S., Fogeron, M.L., Schledorn, M., Dujardin, M., Penzel, S., Burdette, D., ... & Böckmann, A. (2019). Combining cell-free protein synthesis and NMR into a tool to study capsid assembly modulation. *Frontiers in molecular biosciences*, 6(67).

5. David, G.[#], Fogeron, M.L.[#], Schledorn, M.[#], Montserret, R., Haselmann, U., Penzel, S., ... & Bartenschlager, R. (2018). Structural Studies of Self-Assembled Subviral Particles: Combining Cell-Free Expression with 110 kHz MAS NMR Spectroscopy. *Angewandte Chemie International Edition*, 57(17).

6. Schledorn, M., Meier, B. H., & Böckmann, A. (2015). Alternative salt bridge formation in Aβ − a hallmark of early-onset Alzheimer's disease?. *Frontiers in molecular biosciences*, 2, 14.

\# These authors contributed equally

Manuscripts in preparation

1. Schledorn M.[#], Malär, A.A.[#], Torosyan, A.[#], Oss, A., Org, M.L., Penzel, S., Klose, D., Wang, S., Lecoq, L., Cadalbert, R., Samoson, A., Böckmann, A., & Meier, B. H. Protein NMR spectroscopy at 150 kHz magic-angle spinning continues to improve resolution and mass sensitivity.

2. Schledorn, M., Ravotti, F., Nespovitaya, N., Böckmann, A., Riek, R., & Meier, B.H. Molecular structure of *human* somatostatin-14 in the amyloid form.

   # These authors contributed equally

Acknowledgements

This work sits humbly on the shoulders of giants and I would like to explicitly thank those whom I have been in contact with during my time at ETH .

The opportunity to be a cog in the ETH science machine is exciting and rare, and I have **Prof. Dr. Beat H. Meier** to thank for that. You have offered me the chance of an internship in your group, followed by your trust in me to start this thesis, and you allowed me to develop my knowledge and skills with exceptional freedom. The intellectual stimulus and the scientific education that you have given me are invaluable.

A genuine and indefatigable care for the well-being of the group is what I would like to thank **Prof. Dr. Matthias Ernst** for most. In fact, when I first joined the group I was a little intimidated by your radiating wisdom, especially on Fridays at 7.45 am. I am happy to have gotten to know you as very social, helpful and funny, and appreciate all the advice and explanations you have given me.

Already after a few weeks into my start here, I had a first meeting with **Dr. Anja Böckmann**, when she was visiting from Lyon. Many meetings and visits have followed since, and I would like to thank you for all your advice, and for your open way of communication. With a little bit of an outsider-perspective, you have often given refreshing input and support to the projects that I have worked on.

I would also like to thank **Prof. Dr. Roland Riek** for being so kind to be my co-examiner. In addition, I would like to thank **Prof. Dr. Christina Divne** and **Dr. Rosaria Gandini** for that initial spark. The field of structural biology would not have the same glow to me without your contagious enthusiasm for it.

The ETH Laboratory of Physical Chemistry has a large number of members of staff who carry the institute with their priceless support. Technical members of staff whom I would like to thank in particular for sharing their extensive knowledge and experience are **Riccardo Cadalbert**, **Alexander Däpp**, **Andreas Hunkeler**, and **Daniel Zindel**. Equally cherished are the administrative members of staff who keep the institute rolling: first and foremost, I would like to thank the irreplaceable **Gabriele Roder**. Your tireless efforts "in the background" leave an indelible imprint on the progress of our entire group. In addition, I would like to thank **Kristina Comiotto** and **Claudia Hilti** for their work and friendly chats.

Of course, the daily life during my time in the solid-state NMR laboratory has been defined by the group members who joined me on parts of the journey. **Francesco Ravotti, Kathrin Székely**, and **Thomas Wiegand** have been crucial for keeping me afloat in the beginnings of the ride, with countless hours of patient and friendly explanations and instructions. In addition, I am thankful for all the jokes, the gallons of coffee, the endless discussions that were equally impassioned when they concerned crucial scientific details as when they concerned ludicrous nonsense, and all the food and drinks (Jägerbombe anyone?) that I shared with **Johannes Zehnder, Marco E. Weber, Fabian Jähnig, Alons Lends, Joeri Verasdonck, Emilie Testori, Anahit Torosyan, Aaron Himmler, Kong Ooi Tan, Matías Chávez, Johannes Hellwagner, Sara Pfister, Susanne Smith-Penzel, Sebastian Schnoz, Morgane Callon, Johannes Wittmann, Alex Malär, Thomas Bauer, Nils-Alexander Lakomek, Giorgia Zandomeneghi, Denis Lacabanne, Vlastimil Jirásko,** and **Andy Smith-Penzel**. All of you are worth your weight in gold!

Also outside of ETH there have been wonderful travel companions on my NMR trip. From Anja's group in Lyon, I would like to thank **Lauriane Lecoq**, **Guillaume David**, **Marie-Laure Fogeron**, and **Shishan Wang**, not in the least for their scientific contribution to this thesis, but also for the very pleasant and plentiful social exchanges. Other adorable people from NMR labs who make science so much fun are **Arndt Wallmann**, **Max Zinke**, **Kitty Hendriks**, **Claudia Bohg**, and **Whitney Costello**.

This thesis would have been unimaginable in silence. In a colossal "Thank you for the music", I would like to include **Marco Tovani, Johannes Schlegel, Martin Kley, Elisabeth Büchler, Jacques Trösch, Laura Piveteau, Marco Cristuzzi, Ruben Melik, Antoine Schneider** and **Grégoire May**, as well as the complete **Accente e Accenti**, the wonderfully shapeshifting

PSI·HCI·HIT ensemble (that is you, Renato, Emma, Ulrich, Nicola and Markus... we need a better name), and finally, for their immaculate Bach-interpretations that have been on infinite repeat in the past months, the eminent **Viktoria Mullova** and **Hilary Hahn**.

Papa en mama, ik weet niet eens waar ik zou moeten beginnen met bedanken. Jullie zijn er, altijd en onvoorwaardelijk, en daarmee scheppen jullie een 'thuis' dat de kabouters moeiteloos kunnen overdragen, naar waar dan ook in de wereld. Daar ben ik jullie enorm dankbaar voor. I am equally thankful for all the love and support from all the other members of my (extended) family: **Jaap, Janneke, Sam, Anna, Ulf, Klara, Sandra,** and **Jonas**, who even joined our lab for a week for his final high-school science project. And here, in the very end, for being my $|\alpha\rangle$ and my $|\beta\rangle$, I thank you **Magnus**, den sötaste gröten på jorden :)