DISS. ETH NO. 20258

# Solid-State NMR Approaches for Protein Structure Determination

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

MATTHIAS HUBER

Master of Science ETH in Chemistry

Eidgenössische Technische Hochschule Zürich

born December 29, 1982

citizen of Schenkon LU

accepted on the recommendation of

Prof. Dr. Beat H. Meier, examiner

Prof. Dr. Roland Riek, co-examiner

Dr. Anja Böckmann, co-examiner

2012

*dedicated to my family*

*If you want to have good ideas you must have many ideas. Most of them will be wrong, and what you have to learn is which ones to throw away.*

*Linus Pauling*

# Contents

# Abstract

Proteins are biological macromolecules consisting of a chain of amino acid residues. Knowing their structure is a key step towards understanding their function. Since the three-dimensional structure of proteins can usually not be easily and reliably predicted from the amino acid sequence experimental methods for its determination are needed. The most common methods to determine protein structure are x-ray crystallography and solution NMR. Proteins that do not form crystals of good diffraction quality and that are at the same time not soluble are more difficult to study. Solid-state NMR does not suffer from those limitations and is thus an attractive alternative, particularly for the study of fibrils, membrane proteins or large protein complexes.

Structure calculations based on NMR data combine chemical information of the protein (amino acid sequence, bond lengths and bond angles) with experimentally measured restraints. For solid-state NMR the most common restraint classes are torsional angle restraints and distance restraints that are derived from correlation spectra, that have in the simplest case two dimensions. Since such 2D spectra for distance-restraints of proteins suffer from severe resonance overlap, we analyze in the first part of the thesis the potential of higher-dimensional (3D and 4D) proton-detected experiments, which show an increased number of unambiguous distance restraints for solid-state structure determination.

Sensitivity is a limiting factor in NMR measurements of proteins. Detection of protons instead of carbons or nitrogens offers a significant gain in sensitivity. In solid-state NMR narrow $^1$H line-widths can only be achieved if the strong homonuclear dipolar couplings are eliminated. This can most easily be done by partial deuteration. To maintain good resolution on $^{13}$C in such samples the J-coupling to deuterium needs to be decoupled. We established a simple and reversible way to expand a commercial 1.3 mm HCN MAS probe with a $^2$H channel with sufficient field strength for scalar deuterium decoupling (i.e. 2-3 kHz). A supplementary coil is placed at the outside of the stator and requires no significant modifications to the probe. The performance and the realizable gains in sensitivity and resolution are demonstrated using perdeuterated ubiquitin with selectively $^{13}$CHD$_2$-labeled methyl groups.

We then demonstrate on microcrystalline ubiquitin that a non-uniformly-sampled 4D $H^{methyl}$–$H^{methyl}$ spectrum can be be measured on just 2 mg of protein in 3 days with sufficient sensitivity to extract unambiguous distance restraints that permit successful structure calculation in combination with $H^N$–$H^N$ restraints from a 3D spectrum and TALOS+ torsional angle restraints.

In the second part of the thesis solid-state NMR methods are applied to amyloid fibrils. Amyloid fibrils play a role in various neurodegenerative diseases including Alzheimer's disease, the most common form of dementia. We present chemical shift assignments for fibrils formed by highly pure recombinant A$\beta$1-40 with the Osaka E22$\Delta$ mutation that is found in familial Alzheimer's disease. Amyloid $\beta$ 1-40 E22$\Delta$ fibrils are highly ordered and yield line widths in the order of 0.5 ppm. The fibrils consist of parallel in-register $\beta$-sheets and all regions of the peptide are rigid, including the N-terminal part which is believed to be flexible in A$\beta$ wt. Seeding experiments suggest that the same fold can be induced in A$\beta$ wt.

Finally we analyze structural features of fibrils and show how calculations of symmetric fibrils can be implemented in CYANA. In contrast to NMR based structure determination of globular proteins where peaks from spatial correlation spectra are mostly due to contacts between different parts of the primary sequence of one protein molecule, intermolecular interactions play an important role for fibrils. Amyloid fibrils have a characteristic cross-beta structure consisting of extended $\beta$-sheets perpendicular to the fibril axis that contain $\beta$-strands from different protein molecules. In the direction of the fibril-axis a network of hydrogen bonds connects different $\beta$-strands to the extended $\beta$-sheet. Laterally the fibril is stabilized by side-chain interactions. Both axial and lateral contacts can be intra- or intermolecular. A segment of the fibril is calculated that consists of several monomers. The features for the calculation of symmetric dimers have been extended to allow restraining several molecules in the transverse plane of the fibril into a symmetric arrangement as well as periodic repeats along the fibril axis.

# Zusammenfassung

Proteine sind biologische Makromoleküle, die aus einer Kette von Aminosäuren bestehen. Um die Funktion von Proteinen zu verstehen, ist es wichtig, ihre Struktur zu kennen. Weil man die dreidimensionale Struktur von Proteinen normalerweise nicht einfach und zuverlässig aus ihrer Aminosäurensequenz verhersagen kann, braucht man experimentelle Methoden um sie zu bestimmen. Die bekanntesten Methoden zur Strukturbestimmung sind die Kristallografie und die Lösungs NMR. Proteine, die keine Kristalle von guter Diffraktionsqualität bilden und gleichzeitig nicht löslich sind, sind schwieriger zu untersuchen. Festkörper NMR leidet nicht unter diesen Einschränkungen und ist deshalb eine attraktive Alternative, besonders um Fibrillen, Membranproteine oder grosse Proteinkomplexe zu untersuchen.

Strukturrechnungen, die auf NMR Daten basieren, kombinieren chemische Information über das Protein (Aminosäurensequenz, Bindungslängen und Bindungswinkel) mit experimentell bestimmten Daten. Mit Festkörper NMR werden in der Regel erlaubte Bereiche für Torsionswinkel bestimmt oder es werden Distanzobergrenzen zwischen zwei Atomen aus Korrelationsspektren extrahiert, die im einfachsten Fall zwei Dimensionen haben. Da die Signale in solchen 2D Spektren stark überlappen, untersuchen wir im ersten Teil dieser Arbeit die Möglichkeit, mit höher dimensionalen (3D und 4D) Protonen-detektieren Spektren die Anzahl eindeutig bestimmbarer Distanzobergrenzen für die Strukturbestimmung zu erhöhen.

Die Sensitivität ist ein limitierender Faktor in NMR Messungen an Proteinen. Die Detektion von Protonen statt Kohlenstoff- oder Stickstoffkernen ermöglicht eine signifikante Steigerung der Sensitivität. In der Festkörper NMR können scharfe $^1$H Linien nur erreicht werden, wenn man die starken homonuklearen dipolaren Kopplungen entfernt. Am einfachsten erreicht man dies durch teilweise Deuterierung. Um in der Kohlenstoffdimension eine gute Auflösung zu behalten, müssen in solchen Proben die skalaren Kopplungen zu Deuterium entkoppelt werden. Wir haben eine einfache und reversible Methode entwickelt, um einen kommerziellen 1.3 mm HCN MAS Probenkopf mit einem zusätzlichen Deuteriumkanal aufzurüsten, der 2-3 kHz Feldstärke für die skalare Entkopplung liefern kann. Dazu wird eine zusätzliche Spule ausserhalb des Stators installiert, wozu keine signifikante Modifikation des

Probenkopfes nötig ist. Die Funktion und die erzielbaren Gewinne bezüglich Empfindlichkeit und Auflösung demonstrieren wir auf einer perdeuterierten Probe von Ubiquitin mit selektiv $^{13}CHD_2$ markierten Methylgruppen. Wir zeigen dann auf einer mikrokristallinen Probe von Ubiquitin, dass ein nicht gleichmässig abgetastetes 4D H$^{methyl}$–H$^{methyl}$ Spektrum in drei Tagen mit nur 2 mg Protein gemessen werden kann und dass die Empfindlichkeit ausreicht um eindeutige Distanzobergrenzen zu extrahieren, mit denen es dann möglich ist, in Kombination mit Distanzdaten aus einem 3D H$^N$–H$^N$ Spektrum und mit TALOS+ Drehwinkeleinschränkungen, erfolgreich die Struktur zu bestimmen.

Im zweiten Teil dieser Arbeit werden Festkörper NMR Methoden auf Amyloid Fibrillen angewendet. Amyloid Fibrillen spielen in verschiedenen neurodegenerativen Krankheiten, wie z.B. der Alzheimer-Krankheit, der häufigsten Form der Demenz, eine Rolle. Wir zeigen die sequenz-spezifische Zuordnung der chemischen Verschiebungen für Fibrillen aus hoch reinem, rekombinant hergestelltem A$\beta$1-40 mit der Osaka E22$\Delta$ Mutation, die für eine familiäre Häufung von Alzheimer Erkrankungen verantwortlich ist. Amyloid-$\beta$ 1-40 E22$\Delta$ Fibrillen sind hochgeordnet und geben Linienbreiten in der Grössenordnung von 0.5 ppm. Die Fibrillen bestehen aus parallelen in-Register $\beta$-Faltblättern und alle Bereiche des Peptids sind unbeweglich, inklusive dem N-terminalen Teil, der im Wildtyp A$\beta$ als beweglich angenommen wird. Weiter zeigen wir, dass durch animpfen mit A$\beta$1-40 E22$\Delta$ Fibrillisationskeimen die gleiche Faltung auch in Wildtyp A$\beta$ induziert werden kann.

Schliesslich analysieren wir die Strukturelemente von Fibrillen und zeigen, wie Strukturrechnungen von symmetrischen Fibrillen in CYANA implementiert werden können. Im Gegensatz zur NMR-basierten Strukturbestimmung von globulären Proteinen, wo Signale in räumlichen Korrelationsspektren hauptsächlich durch Kontakte zwischen verschiedenen Teilen der Primärstruktur eines Proteinmoleküls verursacht werden, spielen intermolekulare Wechselwirkungen eine wichtige Rolle für Fibrillen. Amyloid Fibrillen haben eine charakteristische Kreuz-beta Struktur, die rechtwinklig zur Fibrillenachse aus ausgedehnten $\beta$-Faltblättern besteht, welche $\beta$-Stränge von verschiedenen Protein Molekülen enthalten. In Richtung der Fibrillenachse verbindet ein Netzwerk aus Wasserstoffbrücken die $\beta$-Stränge zu einem ausgedehnten $\beta$-Faltblatt. Lateral wird die Fibrille durch Seitenkettenwechselwirkungen stabilisiert. Sowohl axiale wie auch laterale Kontakte können intra- oder intermolekular sein. Berechnet wird ein Segment der Fibrille, das aus mehreren Monomeren besteht. Die Funktionen zur Strukturrechnung von symmetrischen Dimeren wurden ausgebaut um zu ermöglichen, dass mehrere Moleküle symmetrisch im Querschnitt der Fibrille angeordnet werden und sich gleichzeitig entlang der Fibrillenachse periodisch wiederholen.

# 1. General Introduction

## 1.1. Interactions in NMR

The Hamiltonian of a system of spin-$\frac{1}{2}$ nuclei is given by[1]:

$$\hat{\mathscr{H}} = \hat{\mathscr{H}}_Z + \hat{\mathscr{H}}_{RF} + \hat{\mathscr{H}}_{CS} + \hat{\mathscr{H}}_D + \hat{\mathscr{H}}_J. \tag{1.1}$$

The Zeeman Hamiltonian $\hat{\mathscr{H}}_Z$ describes the interaction of the spins with the static magnetic field, the rf-Hamiltonian $\hat{\mathscr{H}}_{RF}$ describes the interaction of the spins with an applied rf-field, $\hat{\mathscr{H}}_J$ is the isotropic J-coupling Hamiltonian and $\hat{\mathscr{H}}_{CS}$ the chemical-shift Hamiltonian which has the following form:

$$\hat{\mathscr{H}}_{CS} = -\sum_i \gamma_i \vec{\hat{I}}_i \, \underline{\sigma}_i \, \vec{B}_0. \tag{1.2}$$

Here $\gamma_i$ is the gyromagnetic ratio of spin i, $\vec{\hat{I}}_i$ the spin angular momentum operator and $\underline{\sigma}_i$ the chemical shift tensor. The observed chemical shift depends on the orientation of the chemical shift tensor with respect to the magnetic field.

The dipole Hamiltonian $\hat{\mathscr{H}}_D$ describes the through space interaction between two spins. It contains the orientation dependent dipole-dipole coupling constant $d_{i,j}$

$$d_{i,j} = \frac{\delta_D^{(i,j)}}{2} \frac{(1 - 3\cos^2 \theta_{ij})}{2} \tag{1.3}$$

where the dipolar constant $\delta_D^{(i,j)}$ is defined as

$$\delta_D^{(i,j)} = -2\frac{\mu_0}{4\pi} \frac{\gamma_i \gamma_j \hbar}{r_{ij}^3} \tag{1.4}$$

The dipole Hamiltonian has the form of

$$\hat{\mathscr{H}}_D = \sum_{ij,i<j} d_{i,j} \left( 2\hat{I}_{iz}\hat{I}_{jz} - \frac{1}{2}(\hat{I}_i^+\hat{I}_j^- + \hat{I}_i^-\hat{I}_j^+) \right). \tag{1.5}$$

in the homonuclear case and simplifies to

$$\hat{\mathscr{H}}_D = \sum_{ij,i<j} d_{i,j} 2\hat{I}_{iz}\hat{I}_{jz}. \tag{1.6}$$

in the heteronuclear case, since in the frame rotating with the two different Larmor frequencies the two terms $\hat{I}_i^+\hat{I}_j^-$ and $\hat{I}_i^-\hat{I}_j^+$ are time dependent and can be neglected as non-secular. In solution molecules undergo tumbling motion which averages the dipolar interaction to zero and the chemical shift to its isotropic value given by

$$\sigma_{\text{iso}} = \frac{1}{3} \left( \sigma_{xx,i} + \sigma_{yy,i} + \sigma_{zz,i} \right). \tag{1.7}$$

We thus obtain the Hamiltonian describing a system of spin-1/2 nuclei in solution:

$$\hat{\mathscr{H}} = \hat{\mathscr{H}}_Z + \hat{\mathscr{H}}_{\text{RF}} + \hat{\mathscr{H}}_{\text{CS}}^{\text{iso}} + \hat{\mathscr{H}}_J. \tag{1.8}$$

For molecules larger than about 40 kDa, the rotational correlation time $\tau_c$ approaches the order of tens of nanoseconds which accelerates $T_2$ relaxation and leads to broader lines. With relaxation optimized pulse-sequences (TROSY) and deuteration larger molecules of up to 1000 kDa can be studied[2, 3].

For solid samples, the orientation dependence of equations 1.2 and 1.4 has to be taken into account. For a powder in which all possible orientations of the molecule are present, the orientation dependence of chemical shift and dipolar coupling constants leads to a broad distribution of line positions. Static spectra of samples containing multiple spins are difficult to interpret since the powder patterns of the different spins overlap and lead to a broad featureless signal.

Magic-Angle-Spinning (MAS) has been introduced in the 1950s[4, 5] as a way to increase the resolution of spectra from powdered samples. When the sample is spun at an axis inclined by $\approx 54.7\,^\circ$ with respect to the magnetic field, the equation $cos^2(\theta_m) - 1 = 0$ is valid and the second rank tensors describing the anisotropy of the chemical shift and the dipolar coupling are averaged to zero in a first-order approximation. This approximation is only good if the

MAS frequency is considerably larger than the interaction. Today MAS rates of up to 80 kHz are possible which allow to suppress to first-order the anisotropy of the chemical shift and of the dipolar couplings between low $\gamma$ nuclei (e.g. $^{13}$C and $^{15}$N). Since the dipolar couplings between protons are stronger than between carbons and nitrogens, sharp $^1$H lines can only be obtained if the coupling network is diluted by partially deuterating the sample as is discussed in more detail in chapters 2 and 4.

The heteronuclear dipolar coupling between protons and carbons or between protons and nitrogens is also not sufficiently suppressed by MAS alone. To obtain narrow $^{13}$C or $^{15}$N lines, $^1$H decoupling is necessary. Protons are usually decoupled with multipulse sequences such as TPPM, SPINAL or XiX[6–8] with power levels around 100 kHz. Proton decoupling is typically the biggest contributor to the rf-load during an experiment. At fast MAS frequencies (>40 kHz) low-power decoupling schemes[9, 10] become an alternative.

## 1.2.  Isotope Labeling

NMR only allows observing nuclei with a nonzero nuclear spin. Spin-1/2 nuclei are most often used[11]. To study proteins the most convenient nuclei are $^1$H, $^{13}$C and $^{15}$N. Since $^{13}$C and $^{15}$N have a natural abundance of only 1.1 % and 0.37 % respectively[12] and since $^1$H line-widths in solid-state NMR are large due to the strong $^1$H–$^1$H dipolar coupling network, methods are needed to increase the abundance of $^{13}$C and $^{15}$N or to replace $^1$H by $^2$H. Isotope labeled proteins can be obtained relatively easily by recombinant expression in *E.coli* in a medium containing the desired isotope. Most commonly proteins are uniformly labeled with $^{13}$C and $^{15}$N by expression in M9 minimal medium which is prepared with $^{13}$C$_6$-glucose and $^{15}$N-ammonium chloride (M9 composition is given in Table 8.1 on page 113). For perdeuteration expression is done in a D$_2$O based medium. Since some hydrogen atoms are incorporated directly from glucose, perdeuterated glucose has to be used to achieve $^2$H labeling degrees above 80 %[13]. Sites of isotope labeling during recombinant expression can be tailored by analyzing the metabolic pathways and providing a partially isotope labeled compound which is then incorporated at specific positions of the protein[14]. Depending on whether a compound occurring early or late in the metabolic pathway is provided, fractional labeling is achieved for all amino acid residues or for a specific subset.

## 1.2.1. Glycerol Labeling

Alternate $^{12}$C-$^{13}$C labeling is possible by expression with [1,3-$^{13}$C-glycerol] or [2-$^{13}$C-glycerol] as carbon source. This method has been introduced by LeMaster et al.[15] and has been used by Castellani et al.[16] for the structure calculation of the $\alpha$-spectrin SH3 domain from solid-state NMR data. An advantage of the labeling scheme is that adjacent carbons are usually not both $^{13}$C labeled which suppresses the homonuclear $^1$J coupling and thereby increases resolution.

## 1.2.2. Glucose Labeling

Sparser $^{13}$C labeling can be achieved by using [1-$^{13}$C]- or [2-$^{13}$C]-glucose. The expected labeling patterns have been reported by Lundström et al.[17]. Recently this scheme has been used in solid-state NMR to probe supramolecular interactions[18]. The $^{13}$C labeling is dilute enough that dipolar truncation (see section 1.4) is much less significant than in uniformly $^{13}$C labeled samples.

## 1.2.3. Methyl Labeling
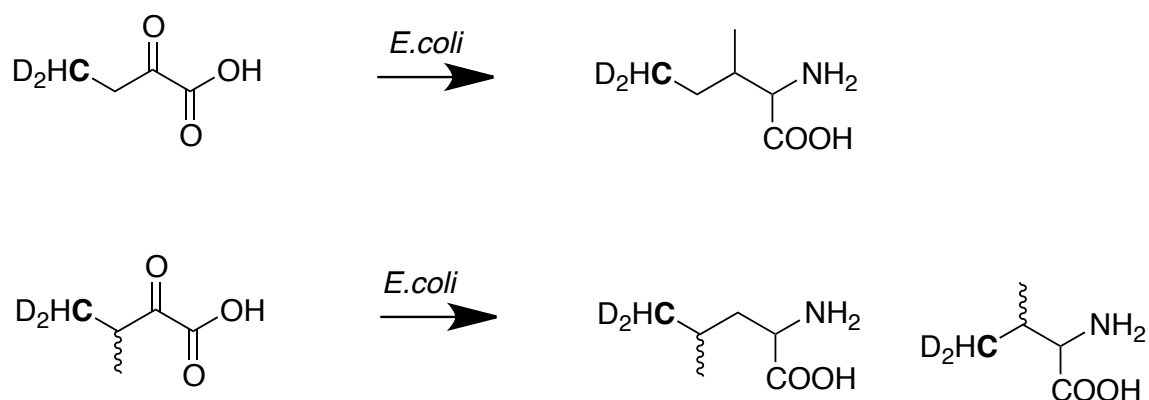


**Figure 1.1.** Supplying *E.coli* with $\alpha$-ketobutyrate one hour prior to induction leads to incorporation of a labeled methyl group into all isoleucines. $\alpha$-ketoisovalerate is converted to valine and leucine in the metabolic pathway for amino acid synthesis.

Methyl groups of isoleucine, leucine and valine can be selectively isotope labeled by adding appropriately isotope labeled $\alpha$-ketoisovalerate or, for isoleucine labeling, $\alpha$-ketobutyrate to

the medium[19]. According to the standard protocol[20] the ketoacid precursor (see Figure 1.1) is added to the medium one hour prior to induction. We have used this labeling scheme to prepare the samples for the collection of proton-detected distance restraints between methyl groups as described in chapter 4. If only one methyl group of valine or leucine should be labeled in a stereospecific fashion, an acetolactate precursor (see Figure 1.2) can be used, of which only one enatiomer is used in protein biosynthesis[21].
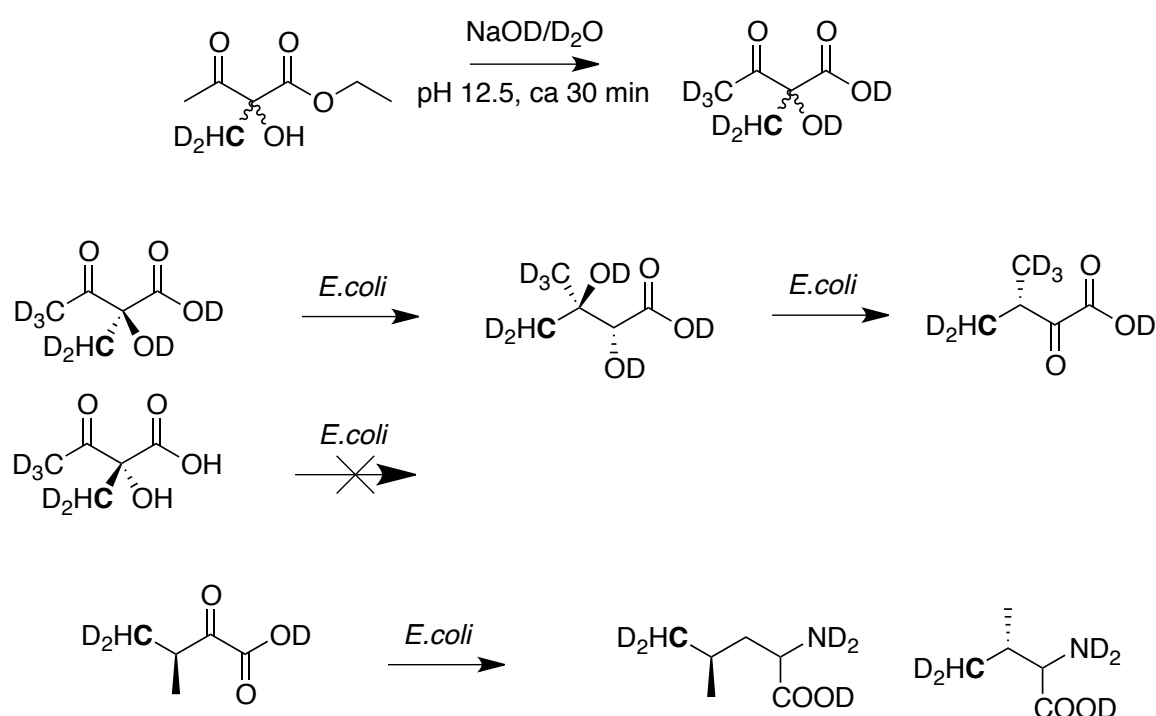
**Figure 1.2.** Only the 2S stereoisomer of the acetolactate obtained upon deprotection and ^{1}H/ [2]H exchange of ethyl oxobutanoate is the substrate of ketol-acid reductoisomerase. When acetolactate is supplied to the growth medium only the 2S stereoisomer is converted by keto-acid reductoisomerase and then dihydroxyacid dehydratase to $\alpha$-ketoisovalerate which is then incorporated into valine and leucine in a stereospecfic way. The 2R stereoisomer is not used for protein biosynthesis.

## 1.3. Recoupling

To transfer magnetization between nuclei one can use the J-coupling or the dipolar coupling. The dipolar coupling is considerably stronger than the J-coupling and therefore permits faster

transfers. In fully protonated samples J-based transfers such as INEPT[22] are difficult to implement since transverse magnetization decays fast. Since the dipolar couplings are averaged by MAS in a first-order approximation we need to recouple this interaction if we want to use it for magnetization transfer.

The most common heteronuclear polarization transfer in solid-state NMR is cross polarization (CP)[23]. Under MAS the resonance condition is

$$n_I \omega_{1I} + n_S \omega_{1S} + n_r \omega_r = 0 \tag{1.9}$$

were $\omega_{1I}$ and $\omega_{1S}$ are the effective nutation frequencies on the I and S spin, $\omega_r$ is the MAS frequency and $n_I$ and $n_S$ can have the values of $\pm 1$ and $n_r$ can be $\pm 1$ or $\pm 2$. At the resonance condition the magnetization oscillates between spin I and spin S and the frequency of this oscillation is given by the dipole-dipole coupling constant $d_{ij}$. Since the the dipole-dipole coupling constant is orientation dependent (see equation 1.3) for a powdered sample the superposition of the different transfer curves is observed. The maximum of the transient oscillation is 73% and for long mixing times a quasi equilibrium with 50% transfer is reached. 100% transfer can be reached by replacing the sudden experiment with an adiabatic version[24].

For homonuclear polarization transfer several methods are available. The first group of them reintroduces the dipolar coupling in the first-order average hamiltonian. Interference is created between the time dependence of the Hamiltonian due to sample rotation and due to rf-irradiation so that the dipolar coupling no longer averages to zero. One way to achieve this is to irradiate at the HORROR condition ($\omega_1 = \omega_r/2$). In the DREAM[25] experiment one sweeps adiabatically through this resonance condition to optimize transfer. Application of DREAM recoupling to a uniformly $^{13}$C-$^{13}$C labeled sample yields cross-peaks mostly for directly bound carbons. To observe transfer over longer distances other mechanisms are needed to avoid dipolar truncation as discussed in the following section.

## 1.4. Dipolar Truncation

Dipolar truncation describes the phenomenon that in the presence of a strong dipolar coupling polarization transfer to spins at a larger distances (corresponding to a smaller coupling) is strongly reduced[26]. We can understand this by looking at the evolution of the density operator of a dipolar-coupled three spin system:

The time evolution of the density operator under a Hamiltonian $\hat{\mathscr{H}}$ is described by the
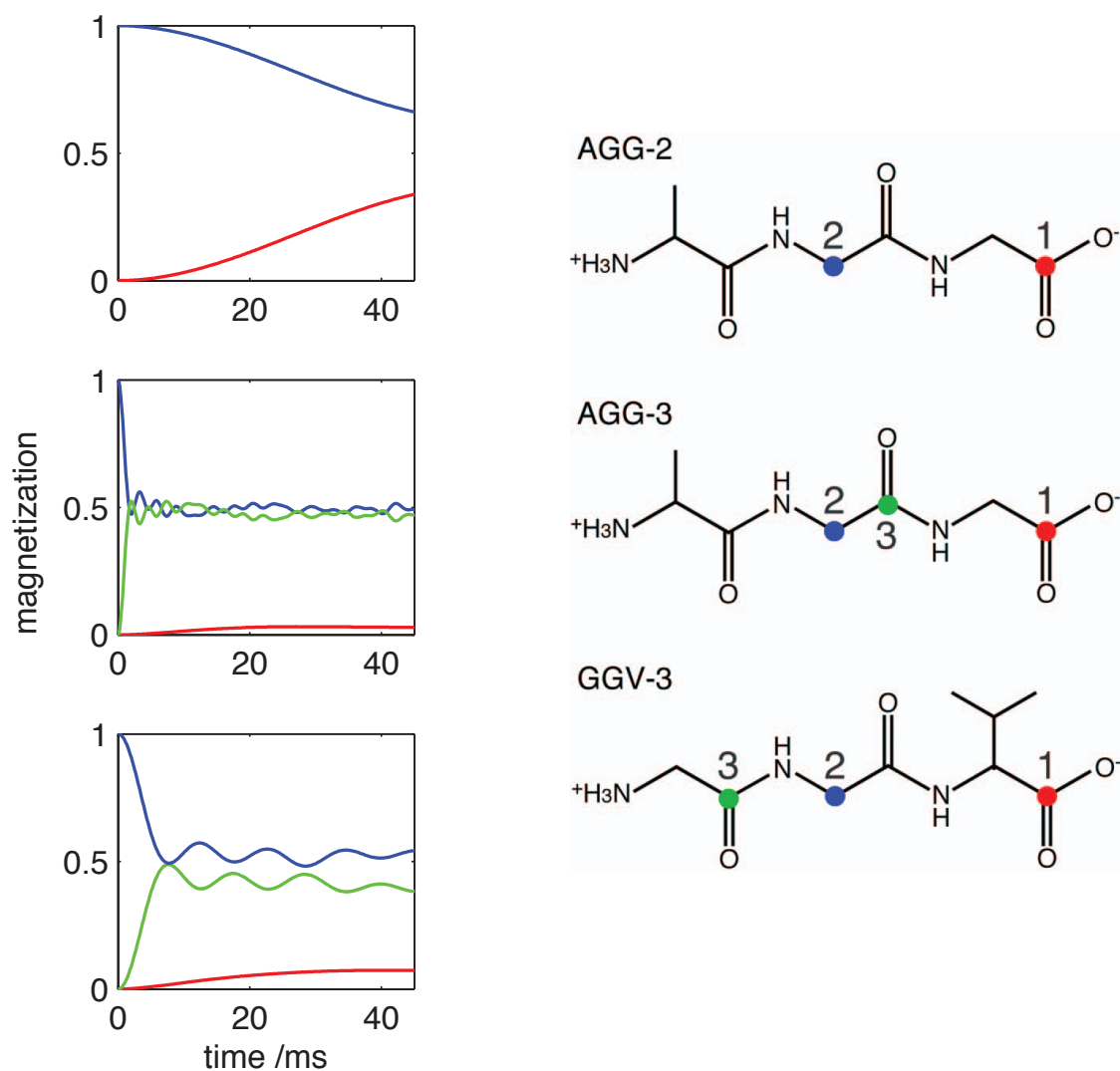
**Figure 1.3.** Simulated polarization transfer with SR4$_4^1$[27] ($\omega_1 = 30$ kHz, $\omega_{MAS}/(2\pi) = 30$ kHz) for the selectively $^{13}$C labeled tripeptides shown on the right (sequence indicated). While in the double $^{13}$C labeled AGG-2 the transfer across the weak dipolar coupling from spin 2 to 1 (66 Hz, 4.8 Å) is slow but takes place, this transfer is truncated in AGG-3 by the strong coupling between spin 2 and 3 (2150 Hz, 1.5 Å). In GGV-3 the 2-3 coupling is weaker (550 Hz, 2.4 Å) and dipolar truncation therefore a bit less severe. For details see [26].

Liouville von Neumann equation

$$\frac{d}{dt}\hat{\rho}(t) = -i\left[\hat{\mathscr{H}}, \hat{\rho}(t)\right] \tag{1.10}$$

with the solution

$$\hat{\rho}(t) = \exp\left(-i\hat{\mathscr{H}}t\right)\hat{\rho}(0)\exp\left(+i\hat{\mathscr{H}}t\right). \tag{1.11}$$

A number of recoupling sequences[28–30] create a zero-quantum first-order average Hamiltonian that has the form

$$\overline{\hat{\mathscr{H}}}_{ZQ}^{(1)} = \sum_{i,j=1;i<j}^{3} d_{ij}\left[\hat{I}_i^+\hat{I}_j^- + \hat{I}_i^-\hat{I}_j^+\right] \tag{1.12}$$

If we set the coupling between spins 2 and 3 ($d_{23}$) to zero we can analytically calculate the transfer of magnetization from spin 1 to spins 2 and 3. We calculate the density operator $\hat{\rho}(t)$ according to equation 1.11 by applying the first-order average Hamiltonian from equation 1.12 to the initial density operator $\hat{\rho}_0 = \hat{I}_{1z}$. The magnetization of spins 2 and 3 along the z-axis is then obtained as:

$$<\hat{I}_{iz}(t)> \quad = \quad Tr[\hat{I}_{iz}\rho(t)] \tag{1.13}$$

$$<\hat{I}_{2z}(t)> \quad = \quad \frac{d_{12}^2}{d_{12}^2 + d_{13}^2}sin^2(\omega t) \tag{1.14}$$

$$<\hat{I}_{3z}(t)> \quad = \quad \frac{d_{13}^2}{d_{12}^2 + d_{13}^2}sin^2(\omega t) \tag{1.15}$$

with $\omega = \sqrt{d_{12}^2 + d_{13}^2}$.

In the case of double-quantum recoupling that creates a first-order average Hamiltonian of the form

$$\overline{\hat{\mathscr{H}}}_{DQ}^{(1)} = \sum_{i,j=1;i<j}^{3} d_{ij}\left[\hat{I}_i^+\hat{I}_j^+ + \hat{I}_i^-\hat{I}_j^-\right] \tag{1.16}$$

we obtain analogous expression for $<\hat{I}_{2z}(t)>$ and $<\hat{I}_{3z}(t)>$ with a negative sign. Since the dipolar constant (see eq 1.4) is proportional $r^{-3}$, the ratio of transferred magnetization can be expressed as:

$$\frac{<\hat{I}_{2z}>}{<\hat{I}_{3z}>} = \left(\frac{r_{13}}{r_{12}}\right)^6 \tag{1.17}$$

If $r_{13} = 2r_{12}$ the intensity of the transfer to spin 3 is less than 2 % of the transfer to spin 2. Therefore to measure long-range distances in uniformly $^{13}$C labeled samples containing many

one-bond $^{13}$C-$^{13}$C contacts of 1.5 Å, a recoupling sequence has to be used that does not create pure ZQ or DQ Hamiltonian as described in equations 1.12 or 1.16. Second-order recoupling sequences, such as PAR[31] have been developed for this purpose. A numerical simulation of polarization transfer characteristics of two- and three spin systems that also takes into account magic-angle-spinning and chemical shift is shown in Figure 1.3.

# 1.5. Assignment

Before information on protein structure and dynamics can be extracted from spectra, the resonance frequencies of the observed nuclei have to be assigned site-specifically to the primary sequence of the protein. To do this, several spectra are recorded that allow to (i) identify spin systems (usually corresponding to amino acid residues), (ii) connect sequential spin-systems and (iii) to complete the (side chain) chemical shifts. The goal is to identify spin-system chains of sufficient length with an amino acid sequence that occurs only once in the protein so that these chains can be uniquely mapped to the primary sequence. For small proteins with narrow lines a sequential assignment can be obtained from a set of 2D spectra but to reduce ambiguities for more complicated cases usually several 3D spectra are recorded[32, 33].

## 1.5.1. Identifying spin systems

In solution NMR spin systems are identified by their peak in the 2D-[$^{1}$H;$^{15}$N] HSQC spectrum. Since the $^{1}$H line-width of fully protonated sample in the solid-state is too large to give a resolved 2D-[$^{1}$H;$^{15}$N] spectrum the standard solid-state NMR protocol relies only on heteronuclei and the 2D-NCA or 3D NCACB spectrum is used to identify the spin-systems. For deuterated proteins an assignment strategy based on the 2D-[$^{1}$H;$^{15}$N] spectrum as in solution NMR can be used[34].

## 1.5.2. Connecting spin systems

To connect sequential spin systems sets of 3D spectra are recorded that overlap in 2 dimensions. Two sequential residues can be connected along the backbone from NCACO, NCOCA and CANCO spectra (see Figure 1.4). Since the chemical shift dispersion on CO is limited, a second set of 3D spectra can be recorded where the CO shift is replaced by CB at the expense of an additional transfer step. The set of spectra then becomes NCACB, N(CO)CACB and
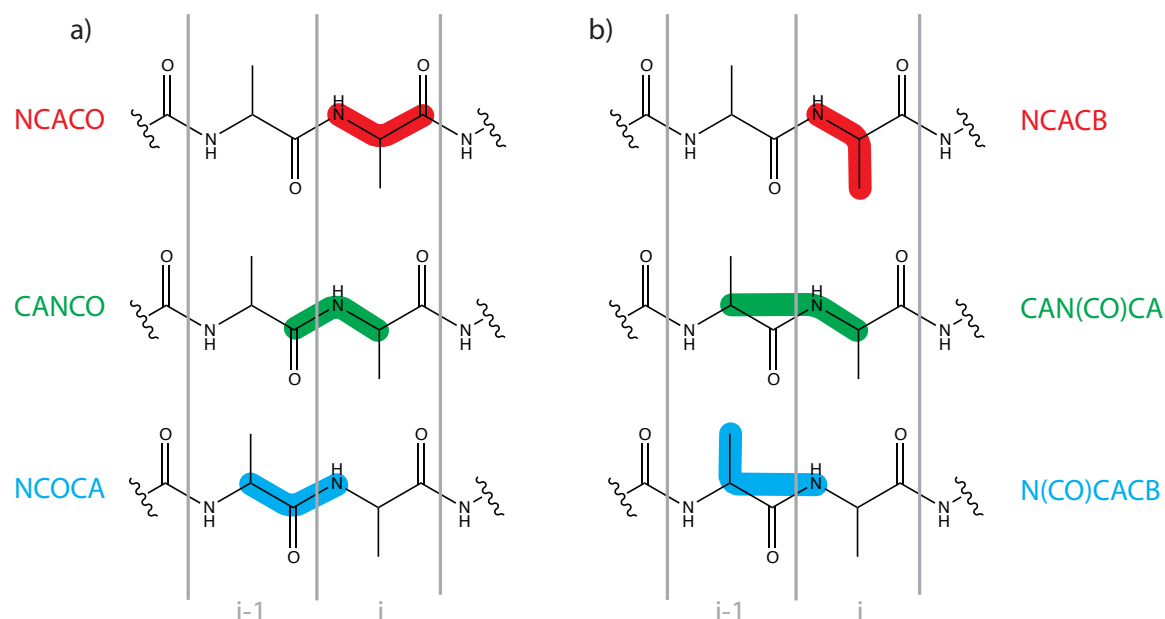
**Figure 1.4.** Observed connectivities in 3D assignment spectra. The most sensitive experiments (a) only involve two transfer steps and detect the C$\alpha$, C' and N resonances. With an additional transfer step the C$\beta$ shift can be obtained instead of C'. Since the observed chemical shift range of C$\beta$ is larger than that of C' peaks are more dispersed.

CAN(CO)CA (magnetization is transferred to but not detected on the resonances in parentheses).

## 1.5.3. Completing side chain assignments

To place chains of assigned residues on the primary sequence it is necessary to identify the types of the spin systems which is easiest when also the side chain atoms have been assigned. For this purpose one can record a 3D NCACX spectrum[35], where a spin-diffusion transfer is used as a last step which distributes the magnetization over all carbons in the spin-system. Alternatives are the 3D N(CA)CBCX spectrum or the 3D CCC that uses first a DREAM transfer to detect directly bound carbons (CA-CB) for instance and then spin-diffusion to find the remaining resonances. For structure calculation it is important to have assigned as many of the visible resonances as possible, ideally all, since missing assignments can result in incorrect distance restraints. Assignments of weak signals usually have to be completed from 2D [$^{13}$C;$^{13}$C] spectra.

# 1.6. Structure Calculation

Structure calculations based on NMR data combine chemical information of the protein (primary sequence, bond lengths and bond angles) with experimentally measured restraints. For solid-state NMR the most common restraint classes are torsional angle restraints and distance restraints.

Finding the structure of a protein corresponds to finding the minimum of a target function that is composed of physical energy function and pseudoenergy terms describing the agreement of the structure with experimental data. The degrees of freedom in this optimization are in principle the coordinates of all atoms. Finding the optimum of the target function is a so-called NP-complete problem. Finding the global minimum exactly is thus very time-consuming and typically heuristic algorithms are used to find an approximate solution. In simulated annealing[36] a molecular dynamics simulation is performed starting with very high kinetic energy that allows the system to escape from local minima which is then gradually minimized. In contrast to a molecular dynamics simulation that aims at obtaining a realistic trajectory, when the aim is only to optimize a structure to fit experimental restraints, a simplified force-field can be used. Bond-distances and bond-angles are known for proteins and their vibration is faster than the time-scale of NMR experiments. Introducing an internal coordinate system with only torsional angles variables allows to fix them in the calculation and thereby decreases the degrees of freedom and allows using longer integration time-steps. To keep the calculation fast and simple electrostatic interactions are usually neglected and the structure is optimized such that it fulfills experimental restraints and that steric conflicts are avoided. Algorithms for torsional angle dynamics are available that scale linearly with the size of the system[37].

**Torsional angle restraints** that define the secondary structure of the protein can be derived from the chemical shifts (obtained during assignment, see section 1.5) with software such as TALOS+[38] or DANGLE[39] that use the empirically found correlation between secondary chemical shift, amino-acid sequence and torsional angles.

**Distance restraints** are more difficult to obtain because spectra to collect them are normally more crowded and have a higher dynamic range than assignment spectra. The information containing long-range peaks are usually weaker than the trivial short range peaks[40, 41]. Signal overlap and dipolar truncation[26] (see section 1.4 and Figure 1.3) make it difficult to observe long-range peaks and once they are identified it remains challenging to resolve ambiguities in assigning them to a restraint between two atoms[40, 42].

true structure          without          with
                        restraint combination

**Figure 1.5.** The effect of restraint combination. Let's assume two restraints were extracted from two different peaks of the spectrum and one of them is correct, the other incorrect (for example due to a wrong shift assignment or because noise was peak-picked). To fulfill both of these restraints the calculated structure will be distorted. With restraint combination two restraints are combined into one ambiguous restraint. As long as one of the two restraints is correct no distortion occurs. (Figure adapted from [45])

A prerequisite for a structure calculation is that it is possible to create a coherent set of distance restraints that can be used to determine the initial fold of the protein. Based on that initial fold ambiguities in the assignment of peaks can be resolved and the structure can iteratively be refined. Typically several cycles of cross-peak assignment and structure calculation are required to obtain a highly resolved structure, and this process has been automated[43, 44]. Since the input to a structure calculation program inevitably contains some imperfections in the form of noise, artifacts and peak positions that have been measured with an error larger than the chemical shift tolerance, it is important to have an algorithm that filters this input and creates a set of restraints that can be used to derive an initial fold of the protein. Two components of such a filtering algorithm are:

**Network anchoring**, which only takes assignment possibilities into account if they can be integrated into a self-consistent set of distance restraints.

**Restraint combination** (see Figure 1.5), which combines distance-restraints from two or several peaks into a single upper distance limit. The probability of an incorrect restraint is reduced because the resulting restraint is only incorrect if both of the combined restraints are incorrect[43].

# Part I.

# Methods for Structure Determination

# 2. The need for 3D- and 4D spectrosopy

This chapter is based on the following publication:

Sebastian, Anja and Beat advised on how to design the manuscript and revised the text of the final version. I performed all structure calculations, prepared figures and wrote the initial manuscript.

## 2.1. Protein Structures using solid-state NMR

Determining the three dimensional structure of a protein is a key step towards understanding its function[46–49]. The most common methods to determine atomic-resolution protein structures are x-ray crystallography and solution NMR spectroscopy. Proteins that do not form crystals of good diffraction quality and that are at the same time not soluble are therefore underrepresented in the Protein Database (PDB)[50, 51] although solid-state NMR is now filling the gap, e.g. for fibrils[46, 52–55], membrane proteins[56] and large protein complexes[57, 58].

Structure calculation by NMR is based on the experimental measurement of geometry restraints, such as internuclear distances and dihedral angles. A prerequisite for the generation of such constraints is a list of the resonance frequencies of the different spins in the protein. This so-called sequence-specific resonance assignment can be established using multidimensional correlation spectra, which allow connecting the spins. The collection of distance constraints requires the measurement of cross-peaks in the nD spatial correlation spectra that correspond to n-tuples where the spins involved are located on amino-acid residues that are distant from each other by 5 or more residues, so called long-range correlations. Because the transfer is

mediated by the dipolar interaction, it is virtually impossible to detect these n-tuples without simultaneously measuring other tuples with i→i, i→(i+1) and i→(i+2) correlations which, in addition, are usually more intense. Therefore, spatial correlation spectra get crowded more quickly and the demand for higher-dimensional spectra is here more urgent.

The best way to measure long-range correlations involves measuring n-tuples where the involved spins all belong to either of two residues, e.g. a 2-tuple (iC$\alpha$, i+kC$\alpha$) or a 4-tuple (iC$\alpha$, iC$\beta$, i+kC$\alpha$, i+kC$\beta$) where i and i+k are the residue number with k>4. We immediately see from this consideration, that the 2-dimensional experiment is only sufficient to unequivocally identify the "source" spin and the "destination" spin for the long-range transfer when the respective 1D spectra are resolved, and the 4-dimensional spectrum when the 2D spectra are resolved. The next useful step would be a 6-dimensional experiment. The usage of odd-dimensional experiments will be discussed below.

Here, we want to describe the perspective for high-resolution structure calculations of larger proteins by exploiting 3D and 4D spatial correlation spectra to reduce signal overlap and assignment ambiguities in combination with sparse labeling that allows for efficient first-order recoupling. We will assess the practical requirements needed for such higher-dimensional spectroscopy, and discuss potentials and limitations of their use in protein structure calculation.

## 2.2. Resolution: nD spectra for structural restraints

Spatial correlation experiments typically contain an encoding, a mixing and a detection block [1]. The polarization is frequency-encoded at its origin, then transferred through space and finally detected at a destination nucleus. Mixing refers to the long-range transfer from spin i to i+k as described in the introduction. If we have the same N nuclei as sources and destinations of polarization (meaning a homonuclear transfer experiment), the components of a matrix contain all transfer efficiencies from which the distances between the nuclei are then derived (master equation approach)[40, 59]. Since the transfer efficiency between two spins is strongly distance-dependent, only the transfers between relatively close neighbors are detectable, typically in the range of up to 6 Å. Nevertheless, the number of resonances and therefore the spectral overlap is much higher than in assignment spectra. Assuming that spins in proteins have roughly m neighbors with which they interact in the chosen mixing time, our aim is thus to detect and resolve the m·N distance restraints and assign each of them to a

source and destination nucleus.

If the 1D spectrum of a protein is fully resolved, then the source and destination spins can unambiguously be identified in a 2D = 1D·1D spatial correlation spectrum. If not, there will be ambiguities in peak assignments. The overlap can be reduced by increasing the dimension of the subspectra that characterize the source and destination resonances with an additional frequency dimension from a directly bonded spin[1], including $^1$H–$^{15}$N, $^1$H–$^{13}$C, $^{13}$C–$^{13}$C and $^{13}$C–$^{15}$N correlations. The resonance used for the additional dimension should either already be isotope labeled or not have a negative impact (such as decreasing the intensity of long-range peaks) on the distance measurement. Adding a $^{15}$N or $^{13}$C label is unproblematic for $^1$H–$^1$H transfers, but changing from a methyl-$^{13}$C labeled sample to [U-$^{13}$C] has negative effects on $^{13}$C–$^{13}$C long-range transfer unless the long-range transfer is strictly band-selective. A 4D = 2D·2D would allow to unambiguously identify both source and destination spins as long as the 2D spectrum of the protein is resolved[60]. Note that we have jumped directly from 2 to 4 dimensions. A faster 3D experiment would allow to identify either the source or destination spin only. When the same spins act both as source and destination spins a symmetric spectrum is expected with each close spatial contact yielding two peaks representing the two directions of the transfer. If those two peaks can be identified and connected to each other the same information is obtained from a 3D spectrum as from the 4D spectrum.

## 2.3. Sensitivity: Proton detection and sparse labeling

Whereas 3D correlation spectra for resonance assignment purposes can be recorded with carbon-13 detection, 3D or 4D restraint spectra are often too insensitive to be recorded in a reasonable measurement time. A significant sensitivity gain can be achieved by substituting heteronuclear ($^{13}$C or $^{15}$N) detection by proton detection, if a narrow $^1$H line width can be obtained such that the [$^1$H;$^{15}$N] or [$^1$H;$^{13}$C] 2D spectrum is sufficiently resolved. Presently this can only be achieved if the dense network of $^1$H–$^1$H dipolar couplings is broken up by sparse $^1$H labeling against a perdeuterated background[61, 62]. This requirement demands expression of the protein in $D_2O$-based medium. Based on such a perdeuterated background, selected sites can then be protonated.

A way to produce partially protonated samples is the expression of uniformly [$^2$H$^{15}$N]- or [$^2$H$^{13}$C$^{15}$N]-labeled protein with subsequent partial back-exchange of deuterons at the backbone amide position with $^1$H spins during the purification and sample preparation steps. Best

results for 24 kHz MAS have been reported for back-exchange from a $H_2O$ / $D_2O$ mixture of 30% / 70%[63]. At 60 kHz MAS high resolution can also be achieved in samples back-exchanged with 100% $H_2O$[64]. While this amide-labeling scheme is biochemically relatively straightforward, it has the disadvantage that the labeled sites are limited to the protein backbone and exchangeable side-chain positions. These sites are useful to measure long-range contacts between $\beta$-strands, but distance restraints between helices are difficult to observe due to the long distances involved. To obtain distance restraints within the hydrophobic core of helical proteins it is thus beneficial to introduce proton labels on hydrophobic side-chain positions. A convenient way to achieve this is with Ile, Leu and Val methyl labelling[19]. This scheme has the advantage that it creates numerous contacts within the hydrophobic core of typical proteins. $^1H$–$^{13}C$ spin pairs have – in addition to making highly sensitive proton-detection possible – the advantage that they can be placed at more sites in the hydrophobic core of the protein when compared to $^{13}C$–$^{15}N$ spin pairs.

As a third option, a diverse spread of sparse $^1H$ labels can be created by expressing the protein in $D_2O$-based medium containing a few percent of $H_2O$. This prevents the formation of large $^1H$ dipolar coupling networks, which would lead to broad resonance lines. The probability of observing a particular long-range contact in these samples is reduced, as both the source and destination spins are only labeled to a small percentage. Asami et al.[65] recently showed that meaningful distance restraints can be measured on a sample produced with 5% $H_2O$ content in the expression medium.

## 2.4. Encoding distance information

Distance information in solid-state NMR is derived from measuring the dipolar interactions between spins by using rf-pulse sequences to recouple the dipolar couplings that are strongly attenuated by magic-angle spinning. Focussing on broadbanded pulse sequences, two categories can be distinguished. (i) First-order recoupling sequences directly reintroduce the dipolar coupling, resulting in high transfer efficiencies, but suffering from dipolar truncation[26], i.e. polarization transfer occurs only for the strongest dipolar coupling of each nucleus, and transfers over weaker dipolar couplings corresponding to longer distances are severely attenuated. (ii) Second-order recoupling sequences such as DARR[66] or PAR[31], which reintroduce the dipolar couplings in the second order of average Hamiltonian theory, resulting in lower transfer amplitudes, inversely proportional to the spinning frequency, but allowing

transfer over long distances even in the presence of strong couplings. First-order terms can promote almost complete polarization transfer in the context of adiabatic transfer schemes, e.g. DREAM[25]. Dipolar couplings not only depend on the distance, but also on the gyromagnetic ratio of the two spins between which polarization is transferred. Thus, transfers between protons yield best sensitivity of all stable isotopes, and can access the longest internuclear distances (we observed transfer for distances of up to about 6 Å - see Figure 4.7 on page 60 for details). For sensitivity reasons, it is therefore beneficial to apply a first-order recoupling scheme, e.g. DREAM, between protons.

## 2.5. Sampling schemes

Besides the sensitivity limitations discussed so far for 3D and 4D experiments, another issue is becoming central for high-dimensional NMR experiments, in particular for dimensionalities of four and above: It is not practical to fully sample the three indirect dimensions of a 4D spectrum uniformly since even a low number of time points per dimension leads overall to a very large number of required scans. When taking into account the need to record all complex components and the phase cycling to select the coherence transfer pathway, the recording of 4D spectra using uniform schemes requires prohibitively long data acquisition time, in the order of weeks to months.

This obstacle can be overcome by using non-uniform sampling schemes that record just a certain fraction of the time domain data. Biasing the selection of sampled points towards short times where the signal envelope is larger than at later time points increases S/N compared to random sampling[67]. Fourier Transformation (FT) of sparse time-domain data with missing points set to zero produces sampling artifacts described by a convolution of the hypothetical complete spectrum with the FT of the sampling scheme. Advanced reconstructions methods, like (Co)-MDD[68] and Forward Maximum Entropy Reconstruction[69] can be used to decrease these artifacts. It has been shown that these methods are of great value to remove the sampling noise caused by resonances with a high signal-to-noise value[68, 69]. The results obtained with Co-MDD processing of 4D spatial correlation spectra indicate that this processing method performs in the context of a structure determination even for weak diagonal planes at least as good as the linear FT[68]. Since adiabatic schemes generally allow to transfer polarization efficiently, the diagonal peaks (indicating no transfer) are often weaker than the cross signals when an adiabatic first-order DREAM scheme is used[25]. As discussed in the previ-

ous paragraph, this scheme is highly efficient and can be used as first-order recoupling scheme for diluted protons. Thus, processing of 4D spectra which contain two heteronuclear correlations devoid of strong diagonal peaks, and an adiabatic homonuclear transfer step resulting in very low diagonal signal intensity, can be processed in a simple, straightforward manner by setting all non-recorded points to zero (see section 4.2 and Figure 4.4 on page 58 for more details.)

## 2.6. The 3D HN(H)–H and 4D HC(H)–(H)CH experiments

In spectra for distance-restraints, we need dimensionalities of at least 3, better 4, to obtain sufficient resolution. For optimal sensitivity, we further want to use proton detection and efficient first-order recoupling schemes to record distances between protons. These schemes are chosen to be adiabatic for best sensitivity and for simple and artifact-free processing of the sparsely sampled spectra.

It thus remains to be decided what approach is used to correlate each proton out of the pair corresponding to a long-range correlation to their directly bound $^{15}$N or $^{13}$C nuclei in the same amino acid residue.

The 4D correlation spectra discussed below was obtained with the HX(H)–(H)XH pulse scheme of Figure 2.2. It allows peaks with distance information to be mapped to the 2D HC or HN spectrum of the source and destination sites. Optionally, we may consider 3D versions of these experiments. Whether we correlate one or both of the protons with its bonded heteronucleus (X), we in principle always work in the four-dimensional space defined by the two resonance frequencies HA and XA of the source spin pair and the two resonances frequencies HB and XB of the destination spin pair. A 3D spectrum reduces overlap for only either the source or destination of polarization, but since long-range transfers are symmetric, for one contact we observe two cross peaks in the 3D spectrum: HA–XBHB and HB–XAHA, in the 3D HX(H)–H. It is thus possible, under the condition that the cross peak is isolated in the $[^1H;^1H]$ 2D spectrum, to derive from these two transposed peaks the shifts of the corresponding 4D peak (Figure 2.1). The 3D correlation was implemented as the 3D HX(H)–H experiment illustrated in Figure 2.2a: $t_1$ encodes the proton frequency, and then the polarization is transferred to the attached heteronucleus, whose frequency is recorded in $t_2$. Polarization is transferred back to the proton, and then $^1H$–$^1H$ homonuclear recoupling is used to encode
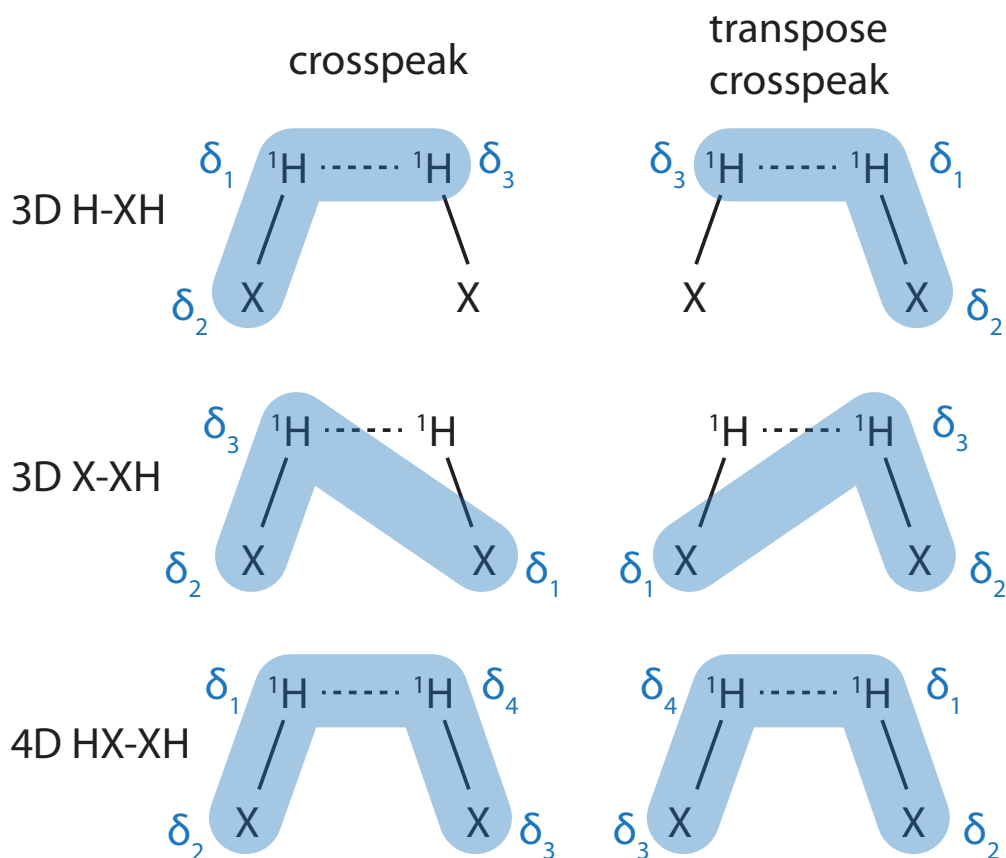
**Figure 2.1.** Combinations of nuclei correlated in 3D and 4D experiments for calculations of protein structures. For the three experiments 3D H–(H)XH, 3D (H)X(H)–(H)XH and 4D HX(H)–(H)XH, the correlated nuclei from two XH moieties are encircled and denoted $\delta_1$–$\delta_4$. X can stand for $^{15}$N or $^{13}$C.

the distance information between the proton spins. Detection is, as discussed above, done on protons. The one-bond heteronuclear transfers in the 3D or 4D experiments can be established using J-based (HSQC or HMQC) or dipolar based (CP) transfers. The evolution of J-based transfers requires long times. Since the coherence lifetime decreases at slower MAS[61] it can be beneficial to replace the J-based transfers by CPs when spinning slower than 55 kHz as in the examples below.

The experimental 3D and 4D spectra recorded on an ILV methyl protons containing sample displays $^1$H–$^1$H distance restraints between nearby methyl protons, which can be measured on the $^{13}$C–$^1$H frequency pairs to reduce overlap.
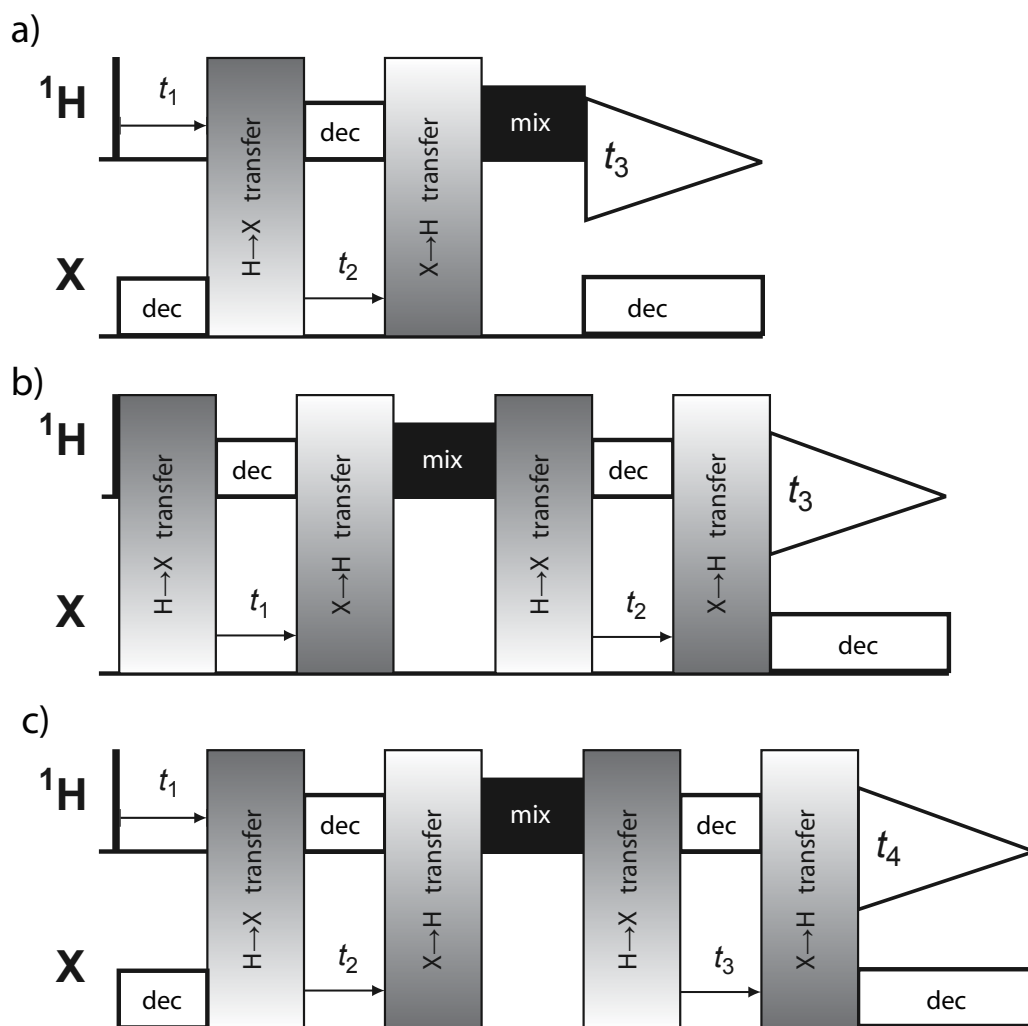
**Figure 2.2.** Schematic pulse sequences for 3D HX(H)–H (a), 3D (H)X(H)–(H)XH (b) and 4D HX(H)–(H)XH (c) experiments. X can stand for $^{15}$N or $^{13}$C. Heteronuclear transfers can be INEPTs or, for samples with short coherence life times, CPs as discussed in section 2.6. The pulse sequences used for the experimental spectra of ubiquitin are shown in chapter 4 as figures 4.1 and 4.6.

## 2.7. Structure Calculation

In the following, the advantage of higher-dimensional spectra is illustrated theoretically and experimentally using the example of ubiquitin, and also theoretically using the hypothetical case of Malate synthase G, an about ten times larger protein. The stability of a structure calculation relies strongly on a coherent set of distance restraints that is able to define the fold of the protein in the initial phase of the structure calculation program (e.g. cycle 1 in CYANA[43, 70]). Ambiguities in the assignment of peaks can subsequently be resolved and the structure can iteratively be refined[43, 44].

The two factors, resonance line width and spectral dimensionality, thus have a significant impact on the quality and resolution of calculated protein structures. To estimate the impact of these factors, we generated synthetic peak lists for 3D and 4D spectra of $^1$H–$^1$H contacts for two proteins, ubiquitin (76 aa) and Malate synthase G (MSG, 723 aa). Synthetic peaklists for 3D H–(H)CH, 3D (H)C(H)–(H)CH and 4D HC(H)–(H)CH experiments, as well as the corresponding amide spectra, were generated for all assigned $^1$H–$^1$H contacts up to a distance of 6 Å in the respective structures[71, 72].

For both proteins, the 3D HN(H)–H spectrum of the amide functionalities shows a better dispersion than the methyl HC(H)–H (Figure 2.3), due to the larger chemical shift dispersion of amide protons when compared to methyl protons. Furthermore, for both nitrogen and carbon, the average peak separation can be greatly increased with 4D experiments. This becomes particularly useful for larger proteins, such as MSG. Note that similar 4D approaches have been used for structure determination in solution-state NMR[72]. Whereas in the simulated 3D spectra of MSG, many signals overlap (16% and 26% overlap in HN(H)–H and HC(H)–H, respectively, for line widths of 30 Hz / 20 Hz / 20 Hz for $^1$H / $^{13}$C / $^{15}$N), the 4D spectra resolve the largest portion of the peaks (2.5% overlap for HC(H)–(H)CH) (Figure 2.3).

To estimate the influence of signal overlap, including the resulting missing and shifted peak positions on structure calculation, we used a sandbox calculation with the following rules to simulate the effect of peak overlap.

1. The probability that a peak is observed was estimated as follows: Peaks corresponding to inter-nuclear distances of up to 4 Å distance were given a 90% probability. For larger distances the probability decreases in proportion with the intensity ($1/r^3$) to be 0% at 6 Å. From this Monte-Carlo master list, signal overlap was calculated for 3D HC(H)–H, 3D (H)C(H)–(H)CH and 4D HC(H)–(H)CH spectra and line widths from 10 Hz to 150 Hz

(assuming a spectrometer operating at 850 MHz $^1$H Larmor frequency). The probability for a peak to be observed despite overlap was set to 100% for peaks separated by more than the line width from the spectrally nearest neighbor and this probability was linearly decreased to 50% for two peaks with identical position.

2. To each peak position, a random shift using a gauss distribution with a standard deviation of half the line width was added to simulate incorrect peak positions.

For each synthetic spectrum and line width, five peak lists with different random number seeds were generated and for each one a structure calculation was performed. The assignment tolerance in the CYANA calculation was set to the line width. In addition to the H$^{met}$ – H$^{met}$ peak lists, a synthetic peak list from a 3D HN(H)–H spectrum produced with the same algorithm and TALOS+ dihedral angle restraints were used for all calculations.

As expected, the RMSD between the calculated structures of ubiquitin and the correct solution increases with increasing line width (Figure 2.4). We further find that each type of input spectra has a certain tolerance up to which the correct three-dimensional structure is obtained. With the 3D HC(H)–H spectrum the correct structure is found for line widths up to 20 Hz. Both the 3D (H)C(H)–(H)CH and the combination with the 3D HC(H)–H spectra yield the correct protein structure solution at line widths up to 50 Hz. The 4D HC(H)–(H)CH experiment tolerates line widths of up to 100 Hz for ubiquitin and has the advantage to be fully resolved, allowing to unambiguously assign distance restraints directly. These calculations thus show the dramatic impact higher dimensionality has on resolving broad peaks and particularly emphasize the importance of four dimensions in spatial correlation spectra.

## 2.8. Experimental results for Ubiquitin

We measured a 3D HN(H)–H and a 4D HC(H)–(H)CH (Val/Leu) spectrum of ubiquitin to test the proposed method for structure calculation[73]. Here, we briefly summarize the results, focusing on the experimental choices made based on the framework presented in the previous sections. Perdeuterated $^{15}$N-labeled ubiquitin microcrystals were used with either 30% of all exchangeable hydrogens $^1$H-labeled or with one methyl group per valine and leucine residue labeled as $^{13}$CHD$_2$. In each case, 2 mg of protein was filled into a 1.3 mm rotor and NMR spectra were acquired on a 850 MHz magnet under magic-angle spinning of 55 kHz. Heteronuclear transfers were done by INEPT and the long-range $^1$H–$^1$H transfer by DREAM[25].
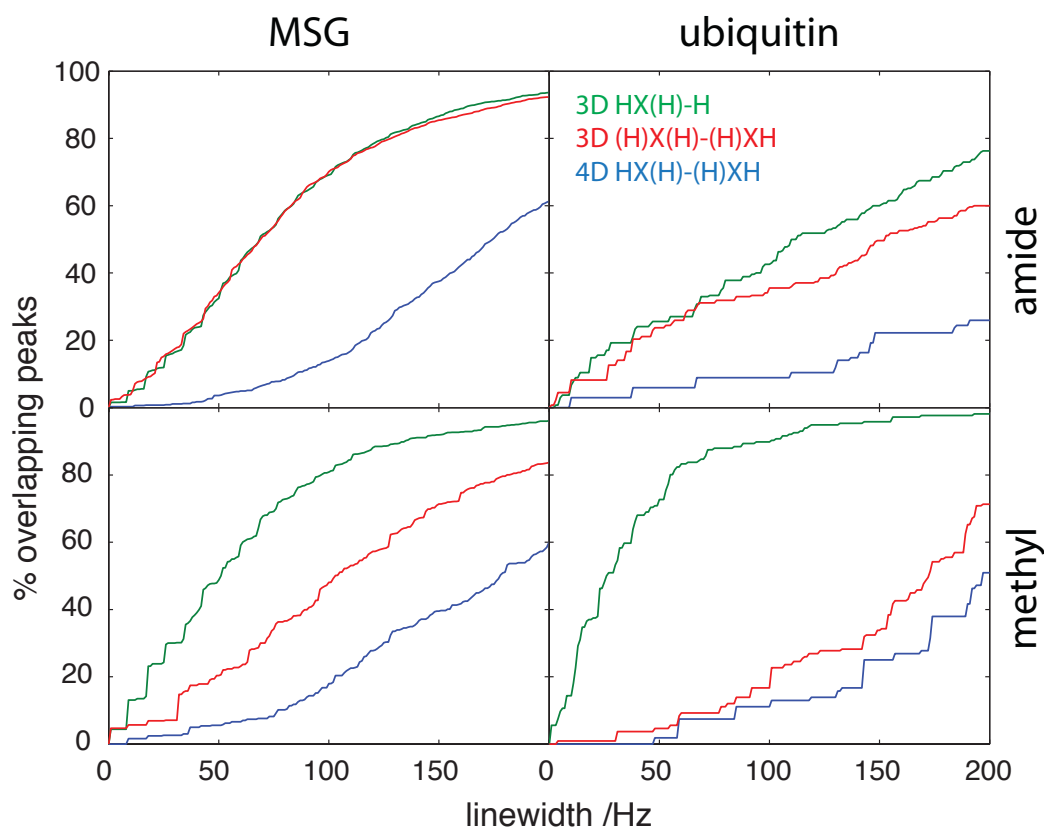
**Figure 2.3.** Fraction of overlapping peaks as a function of line width and spectrum type. Theoretical peak lists of MSG (723 aa) or ubiquitin (76 aa) were created for $^1$H–$^1$H distances closer than 6 Å. These peak lists thus contained 3422 $H^N$–$H^N$ peaks and 1022 $H^{met}$–$H^{met}$ peaks for MSG and 270 $H^N$–$H^N$ peaks and 216 $H^{met}$–$H^{met}$ peaks for ubiquitin. The separation is given in Hz for a $^1$H Larmor frequency of 850 MHz. Corresponding line widths for $^{15}$N and $^{13}$C were assumed 2/3 thereof. Note that while the degree of overlap in the 3D HN(H)–H spectrum is similar as in the 3D (H)N(H)–(H)NH, it is substantially higher in the HC(H)–H than in the 3D (H)C(H)–(H)CH.
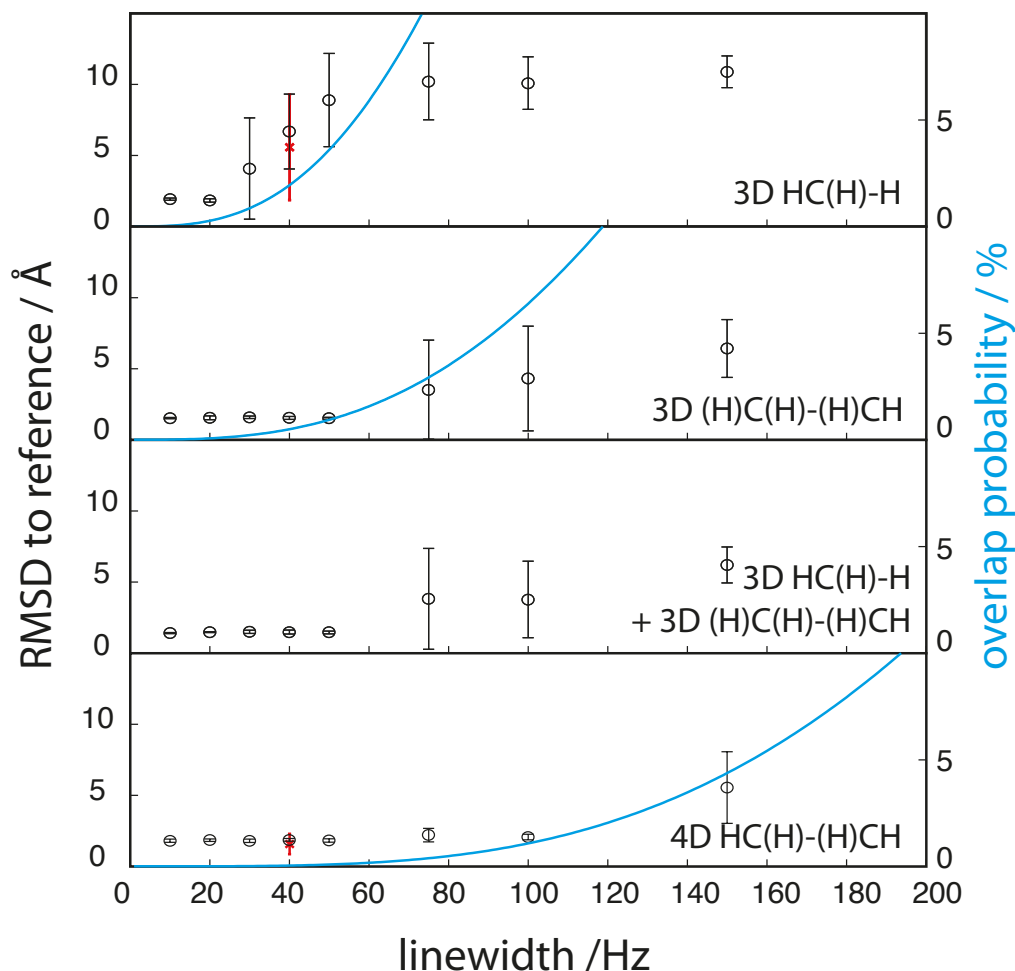
**Figure 2.4.** Result of structure calculations from synthetic peak lists for ILV $H^{met}$ –$H^{met}$ contacts < 6 Å in ubiquitin. The effects of line width and overlap were simulated by removing peaks and shifting peak positions as described in the main text. Circles correspond to the average RMSD to the reference structure (1UBQ) from five independent structure calculations. The error bars indicate the standard deviations. The results from experimental spectra are shown in red. The statistical probability that the peaks of two resonances with line widths $\Delta\omega$ in the populated shift range $\Omega$ overlap in a 1D spectrum is given by $\langle O \rangle = \Delta\omega/(\Omega + \Delta\omega)$ and for an nD spectrum by $\langle O \rangle = \prod_{i=1}^{n} (\Delta\omega_i/(\Omega_i + \Delta\omega_i))$. The probability of overlap between any of the mN peaks is $p = (1 - \langle O \rangle)^{m\cdot N}$ and plotted as a blue line. For the calculation the populated shift range $\Omega$ of Val/Leu methyl resonances was set to 10 ppm for $^{13}$C and 0.6 ppm for $^1$H, the conversion to Hz was done assuming a 20 T magnet.

The 3D H–NH spectrum was recorded in 24 h and the 14% sparse 4D HC–CH spectrum in 72 h. During the 72 h measurement the field drifted within a band of ($< 20$ Hz; $< \frac{1}{2}$ linewidth) and no correction was applied. In our experiment the field is relatively stable once the temperature of the probe reaches a steady state (typically after one night). If this is not the case and cannot be corrected by linear field compensation, active locking may be necessary.

Slices from the proton-detected 4D spectrum at 850 MHz (Figure 2.5) compared to a 2D [$^{13}$C;$^{13}$C]-DARR spectrum with 250 ms mixing time (Figure 2.5) illustrate the overlap problem encountered in low dimensional spectra. All cross peaks observed in the 4D spectrum can be unambiguously assigned to a source and destination methyl group, whereas the majority of the peaks in the 2D DARR are strongly overlapping.

The degree of overlap has direct consequences on the result of structure calculations made with different input spectra (Figure 2.7). Among these three calculations, only the restraint set obtained from the 4D HC(H)–(H)CH spectrum converged to the correct three-dimensional protein structure of ubiquitin. Already the substitution of the 4D spectrum by the lower-dimensional 3D HC(H)–H spectrum leads to a bundle of conformers that did not converge to a single backbone conformation. The structure calculation obtained using the 2D [$^{13}$C;$^{13}$C]-DARR spectrum (Figure 2.6) as the sole input did converge to an incorrect minimum with a RMS deviation from the X-ray structure of 7 Å, well outside the CYANA bundle with an (rmsd) of 1.0 Å (Figure 2.7). This behavior is a consequence of the high (spectral) ambiguity of the 2D spectrum (see section 8.1.7 on page 126 (appendix) for more details). The 2D [$^{13}$C;$^{13}$C]-DARR could be useful for structure refinement when combined with 3D- or 4D spectra that allow determining an approximate fold which could then be used to resolve ambiguities in the assignment of the 2D peaks.

## 2.9. Conclusions

The acquisition of 4D spectra that encode unambiguous spatial proton–proton correlations is feasible today if proton detection is combined with adiabatic first-order proton–proton polarization transfer and sparse sampling. This result is in line with the theoretical expectations and corresponds to measuring distance restraints between source and destination nuclei, which are both identified spectrally by a two-dimensional spectrum.

Unless significant further progress can be made in reducing the line width in the solid-state NMR spectra of proteins, 3D and 4D spectroscopy will thus be key elements for solid-state
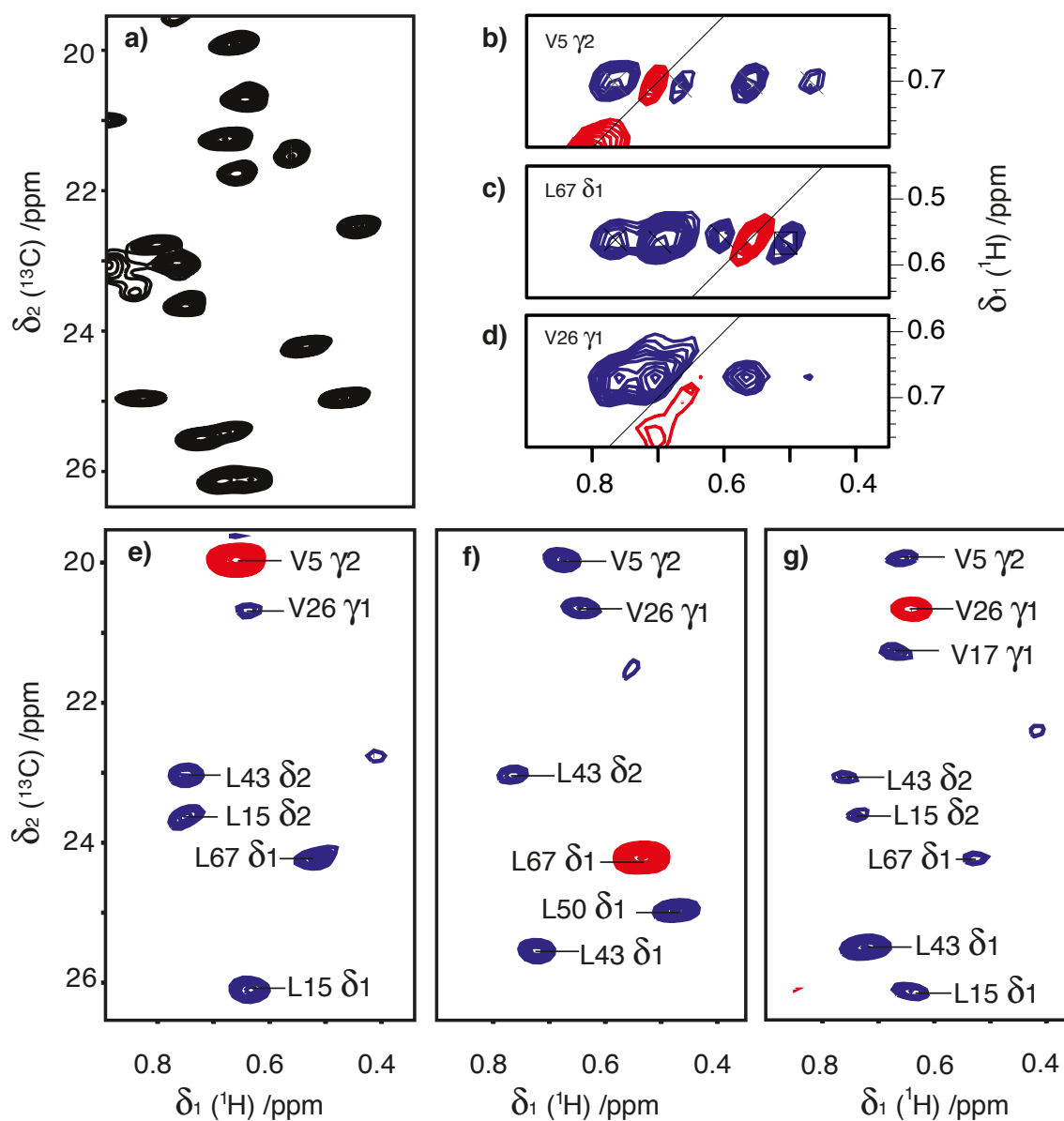
**Figure 2.5.** Proton detected solid-state NMR spectra of ubiquitin used for structure determinations. A $^{13}$C-detected 2D [$^{13}$C;$^{13}$C] DARR spectrum for comparison is shown in Figure 2.6. The proton-detected spectra utilize both the $^1$H and $^{13}$C chemical shift to identify methyl groups. (a) 2D [$^{13}$C;$^1$H]-HSQC spectrum of crystalline [U-$^2$H$^{15}$N; VL-$^{13}$CHD2] ubiquitin recorded at 55 kHz MAS. b–d) Strips extracted from the 3D HC(H)–H spectrum. e–g) 2D planes from the non-uniformly-sampled (NUS) 4D H$^{met}$–H$^{met}$ DREAM experiment. These 2D planes are extracted at the position of the diagonal peak ($\delta_1 = \delta_4$ and $\delta_2 = \delta_3$, red) from the 4D data set.
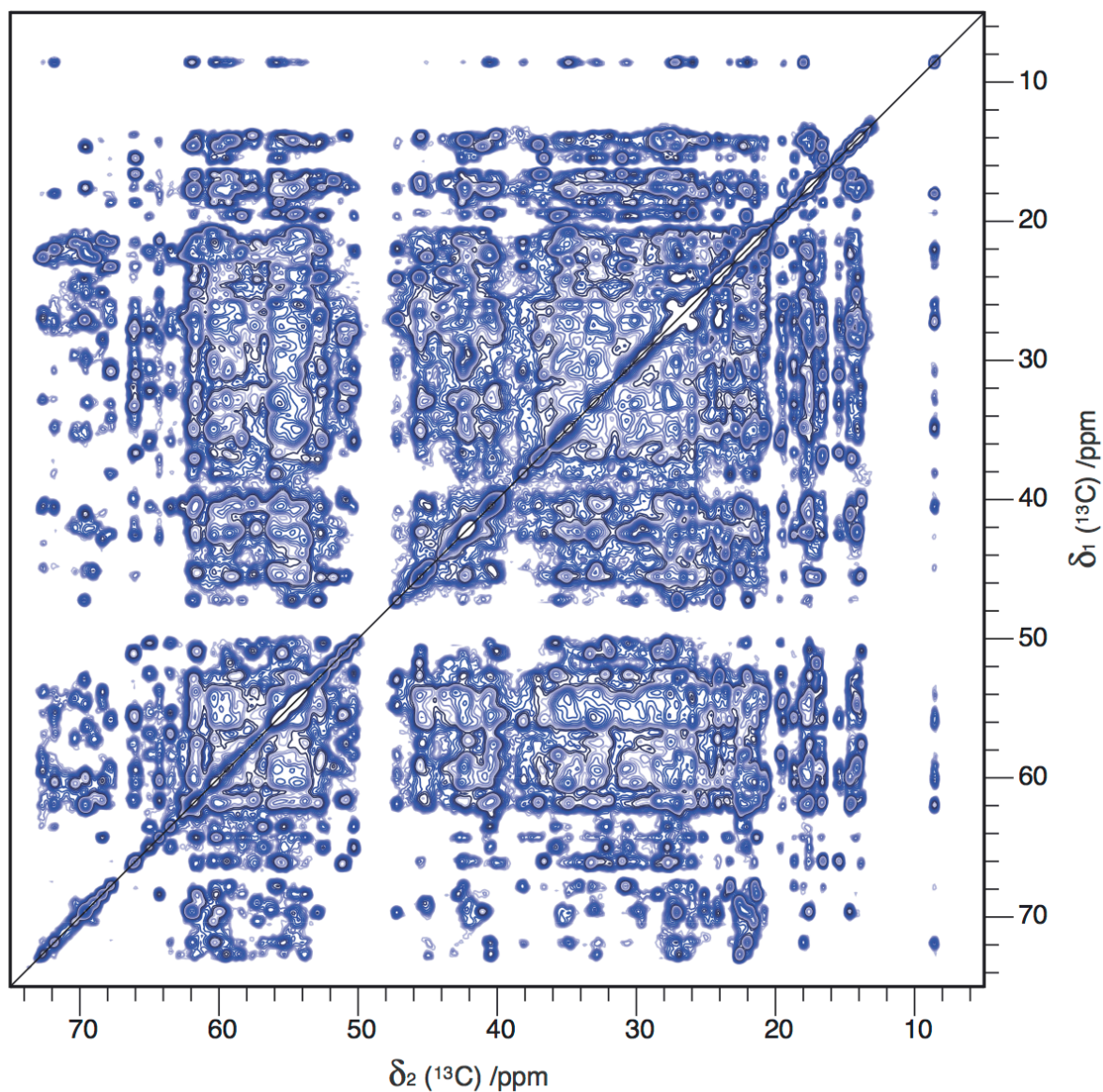
**Figure 2.6.** $^{13}C$-detected 2D [$^{13}C$;$^{13}C$] DARR spectrum with 250 ms mixing recorded at 14.1 T and 13 kHz MAS on [U-$^{13}C$] ubiquitin microcrystals. In addition to the long-range peaks many short- and medium range peaks are observed which increases peak overlap.

|  | 4D HC(H)-(H)CH<br>+ 3D HN(H)-H | 3D HC(H)-H<br>+ 3D HN(H)-H | 2D $^{13}$C-$^{13}$C DARR |
|---|---|---|---|
| bundle RMSD | 0.72 ± 0.12 Å | 3.72 ± 1.05 Å | 1.00 ± 0.40 Å |
| RMSD to reference | 1.57 Å | 5.57 Å | 7.01 Å |
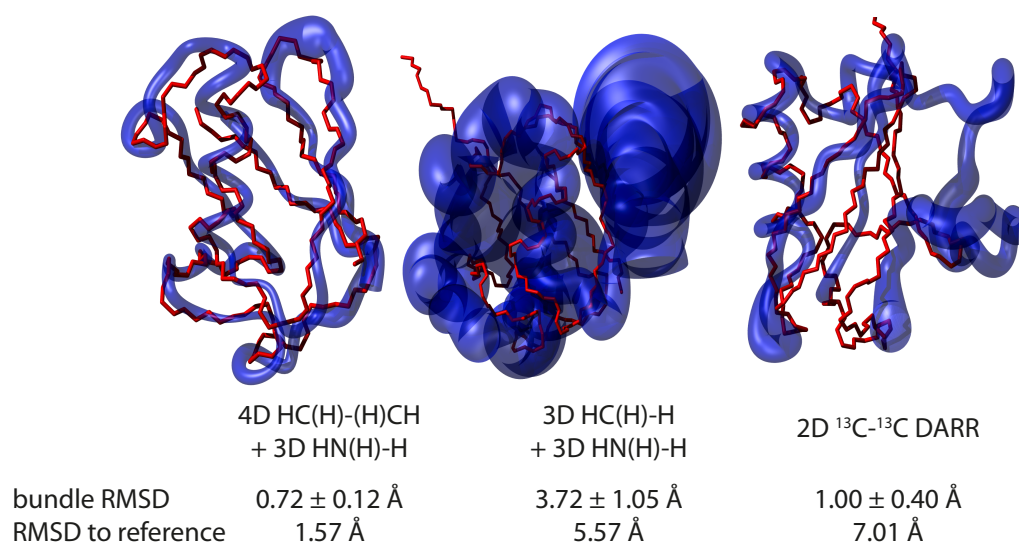
**Figure 2.7.** Structure bundles of ubiquitin in sausage representation (blue) obtained from the experimental solid-state NMR spectra indicated below each bundle. The structures are also aligned to the reference structure (1UBQ, red). In all calculations, TALOS+ dihedral restraints were used. The backbone RMSD of the bundle and the bias RMSD to reference are indicated. Only the leftmost calculation leads to the correct 3D reference structure.

NMR structure determination. Especially if current sensitivity limitations can be overcome by future developments, such as by DNP techniques or higher magnetic field strengths, 4D experiments will become increasingly important as an experimental method of choice for the determination of protein structures. Structure determination on considerably larger proteins will become possible in case that significant signal enhancement can be obtained.

# 3. A supplementary coil for $^2$H decoupling

This chapter is based on the following publication:

Oli designed and built the $^2$H coil and performed the numerical field calculations. Oli, Paul, René and I installed the coil, I measured all spectra and wrote the initial manuscript.

## 3.1. Introduction

While biological solid-state NMR is heavily based on the observation of $^{13}$C and $^{15}$N spins, proton detection is of increasing interest since the observation of sparsely placed protons in otherwise perdeuterated samples has been demonstrated to lead to highly resolved proton spectra. Such samples can be produced by expression in a mixture of $H_2O$ and $D_2O$ (and perdeuterated glucose)[65], by perdeuteration of the protein and crystallization/fibrillization from a partially protonated buffer[61, 62] or, for methyl labelling, through the use of $\alpha$-ketoacid precursors[19]. All of these approaches yields samples with $^{15}$N and/or $^{13}$C spins with directly bound $^2$H spins. As one of many examples $^{13}$CHD$_2$ labeling of methyl groups has been used to study protein structure[73] and dynamics[71, 74]. To obtain sensitive and highly resolved $^{13}$C–$^1$H correlation spectra $^{13}$C–$^2$H J-decoupling is essential, while the dipolar coupling is sufficiently reduced by the MAS. While some commercial triple resonance probes can be tuned to $^1$H -$^{13}$C -$^2$H , many probes for biomolecular NMR are restricted to $^1$H-$^{13}$C-$^{15}$N. In addition, many experiments also require a $^{15}$N dimension and therefore a quadruple resonance $^1$H-$^{13}$C-$^2$H-$^{15}$N probe.

Here, we present a way to install an auxiliary $^2$H coil in a commercial Bruker 1.3 mm HCN
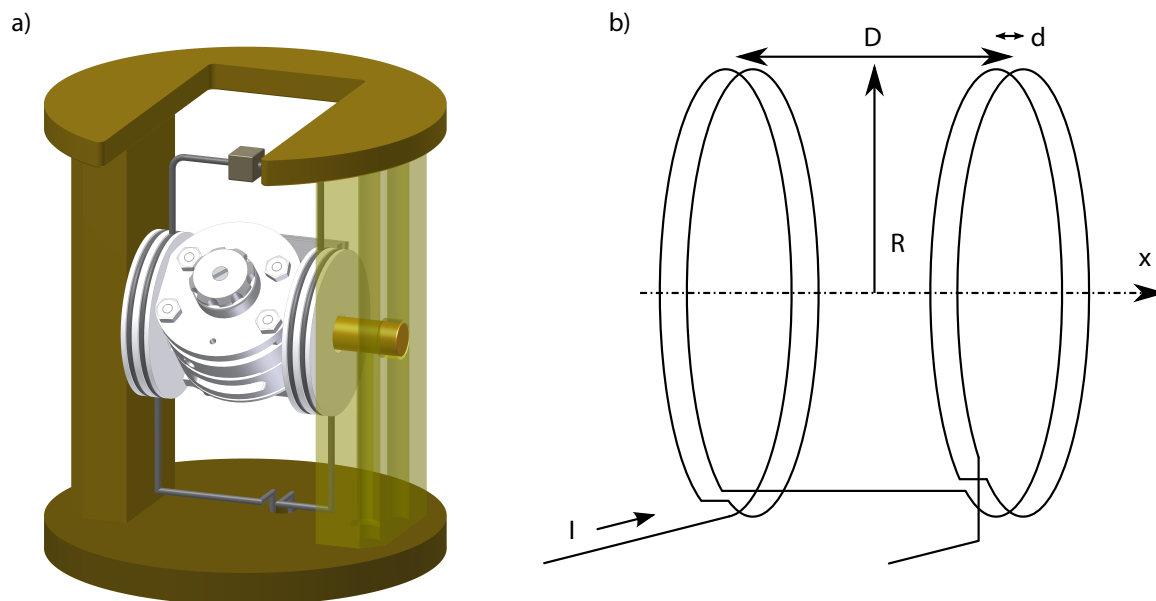
**Figure 3.1.** Geometry and installation of the auxiliary $^2$H coil. (a) Two loops of the coil are on each side of the MAS stator with a capacitor $C_{seg}$ above. This geometry ensures orthogonality of the fields generated by the two coils. The leads are guided through the bottom plate of the stator chamber, using the preexisting hole for optical spinning control and are connected to the tuning circuit in the lower part of the probe. (b) The $^2$H coil consists of two solenoids with radius $R = 12$ mm that are separated form each other by the width of the stator $D = 24$ mm. Each solenoid is two-turned with a spacing d = 2 mm.

probe that does not require significant modification of the existing hardware. We demonstrate on a microcrystalline [U-$^2$H;$^{15}$N, Ile-$^{13}$CHD$_2$] labeled sample of ubiquitin, that $^2$H decoupling with such an auxiliary $^2$H coil greatly enhances sensitivity and resolution of the [$^{13}$C;$^1$H]-HSQC spectrum.

## 3.2. Design Considerations

Different considerations guided our design of a $^2$H coil for J-decoupling purposes. (i) We aim for a supplementary coil which creates a B$_1$ field of an amplitude such that $\gamma$B$_1$ is of the order of several kHz. (ii) The inhomogeneity $\Delta B_1$ should be below 10% over the sample volume which can be easily handled by decoupling sequences[75]. (iii) The rf characteristics of the main coil should remain unchanged, i.e. the coupling of the coils must be minimal. (iv) All modifications to the probe should be simple and completely reversible. This is most easily achieved by placing the auxiliary coil outside the MAS stator.

In order to minimize the coupling between the two coils, we chose for a design where the field generated by the auxiliary coil and the one generated by the main coil, in our case a triply tuned solenoid coil, are perpendicular. Our design resembles a Helmholtz coil but without the condition that the distance between two sets of coils is equal to the coil's radius (see Figure 3.1). The field is parallel to the x-direction and perpendicular to the one produced by the main coil being at the magic angle in the yz-plane, the $B_1$ field of the $^2$H coil on the x-axis halfways between the two parts of the coils is given by[76]:

$$\vec{B}_x = \frac{\mu_0 \cdot I_{coil} \cdot n}{2} \cdot \frac{R^2}{\left( \left( D/2 \right)^2 + R^2 \right)^{3/2}} \cdot \vec{e}_x \tag{3.1}$$

Here, $I_{coil}$ is the current in the coil, $n$ the total number of loops, $D$ the distance between the two sets of loops and $R$ the radius of a loop. The maximum field is produced at $R = D/\sqrt{2}$. The inductance of the structure in Figure 3.1 is given by[77]:

$$L_1, L_2 = \left( \frac{n_{l,r}}{2} \right)^2 \cdot L_{loop} \tag{3.2}$$

where $n_{l,r}$ is the number of loops per side and

$$L_{loop} = R \cdot \mu_0 \cdot (ln\frac{4R}{a} - 2) \tag{3.3}$$

is the inductance of a single loop. The wire resistance associated with the coil is calculated by

$$R_s = l \cdot \sqrt{\frac{\omega}{2\sigma\mu_0}} \cdot \frac{1}{\pi a} \tag{3.4}$$

where $\sigma$ is the conductance and $a$ the diameter of the wire. The main stray capacitance is between the windings of the coil and is calculated by

$$C_{stray} = \frac{\pi \varepsilon \ell}{arcosh(d/a)} \tag{3.5}$$

where $l$ is the length of two parallel wires and $d$ the distance between the wires (see Fig 3.1b). Small stray capacitances also occur between conductors and ground but are difficult to specify in analytical form. These capacitances add on to the tuning capacitor $C_T$.
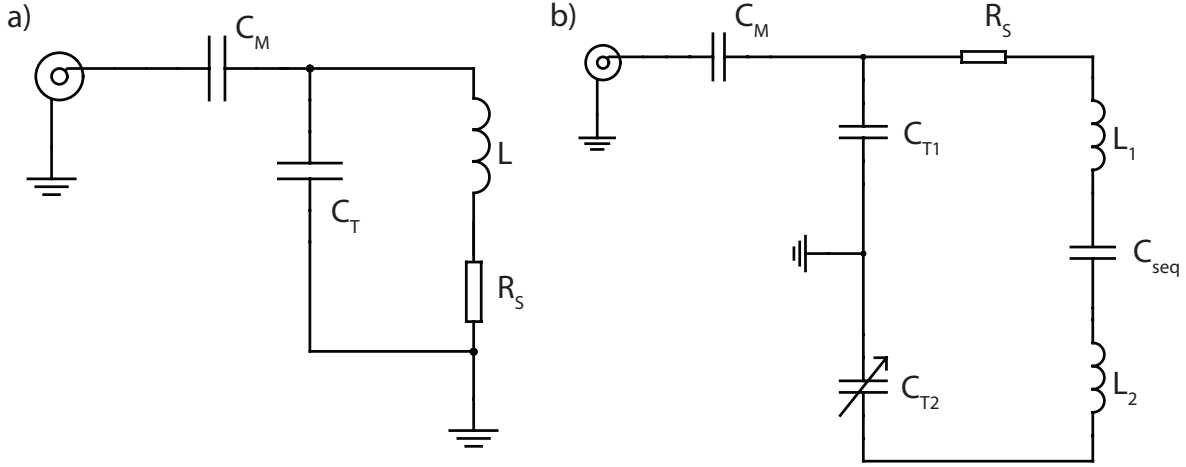
**Figure 3.2.** Schematic diagram of the $^2$H circuits. a) Basic tuning single resonance circuit b) Circuit used here. To reduce the overall inductance the coil is split into $L_1$ and $L_2$ by the segmentation capacitance $C_{seg}$. The tuning capacitance $C_T$ is split into $C_{T1}$ and $C_{T2}$ to achieve a balanced network. Since no tuning rods are installed $C_{T2}$ can only be adjusted while the probe is outside the magnet. Therefore, we chose the matching capacity $C_M$ slightly larger than optimal to increase the bandwidth. The capacitators used were $C_{T1}$ = 4.3 pF, $C_M$ = 1.6 pF, $C_{seg}$ = 10 pF and $C_{T2}$ tunable 1 .. 20 pF (set to $\approx$ 6 pF).

Using network theory the steady-state current in the coil of Fig 3.2b is calculated as[76]:

$$I_{coil} = \frac{2\sqrt{P \cdot R_i}}{\left(R_i + \left(i\omega C_T + \frac{1}{Z_{coil}+R_s}\right)^{-1} - \frac{1}{i\omega C_m}\right) \cdot \left(i\omega C_T + \frac{1}{Z_{coil}+R_s}\right) \cdot (Z_{coil}+R_s)} \tag{3.6}$$

where $R_i$ is the amplifier's internal impedance and $P$ the amplifier's output power.

Since the coil is unusually large, the resulting inductance of $L_1$, $L_2$ and the leads is roughly 600 nH. This would require a very low value for the tuning capacitance $C_T$ to fulfill the resonance condition $\omega_0 = 1/\sqrt{LC_T}$. With the stray capacitances between the loops acting as tuning capacitances $C_T$ would be even lower. Therefore the two coils are split with a capacitor $C_{seg}$ to lower the overall impedance of the coil. Furthermore, the tuning capacitance is split into $C_{T1}$ and $C_{T2}$. The resulting impedance ($Z_{coil}$) is much lower than the impedance of the coil alone and can be controlled by the choice of $C_{seg}$ and $C_{T2}$.

$$Z_{coil} = \left(j\omega L_1 + j\omega L_2 + \frac{1}{j\omega C_{seg}} + \frac{1}{j\omega C_{T2}}\right) \tag{3.7}$$

This balanced circuit (Figure 3.2b) has a number of advantages. The influence of stray capacitances to ground is cut in half, the voltages with respect to ground are only half as high

and the values of $C_{T1}$ and $C_{T2}$ can be chosen flexibly.

To avoid major modifications to the probe, no rods for tuning from the outside were installed. Therefore $C_{T2}$ can only be adjusted while the probe is outside the magnet and open and the matching capacity $C_M$ has a fixed value. By choosing the matching capacity $C_M$ somewhat larger than optimal we achieve a larger bandwidth while accepting a slight mismatch and reduction of the quality factor Q.

The component's values of the equivalent circuit (Figure 2) are as follows: $L_1$ and $L_2 \approx$ 300 nH each, $C_{T1} = 4.3$ pF, $C_{T2} = 1..20$ pF, $C_M = 1.6$ pF, $C_{seg} = 10$ pF, $R_s \approx 0.6 \ \Omega$ with $\sigma = 60 \cdot 10^6$ Sm$^{-1}$, $a = 1$ mm, $l \approx 0.5$ m and $R_i = 50 \ \Omega$. To create an rf-field of 2 kHz amplitude (corresponding to B$_1$=0.612 mT) we need, according to Eq. 3.1, a current $I$ of approximately 8 A. With Eq 3.6 and 3.7, we can calculate the required power at the amplifier to be $\approx 100$ W.

## 3.3. Results and Discussion

We tested the performance of the auxiliary coil by recording a nutation curve that permits to extract rf-field strength and homogeneity. We applied a power level of 125 W and pulse durations between 50 $\mu$s and 5 ms (Fig 3.3a), observing the D$_2$O resonance of a rotor filled with deuterated ubiquitin to be used later for the protein experiments. The first maximum of the nutation curve occurs at about 125 $\mu$s corresponding to a nutation frequency of 2 kHz. The B$_1$ field distribution in the sample volume is obtained from the Fourier Transformation of the nutation curve. Here an unusual bimodal distribution is observed (Fig 3.3c). We performed a finite elements simulation to gain insight into the spatial distribution of the B$_1$ field.

Fig 3.3b shows that for the central part of the MAS rotor the $^2$H field is partly shielded by the inner triple resonance coil, and we can thus ascribe the two components of the rf distribution to parts of the sample inside and outside the main coil. To confirm this experimentally, we measured a nutation curve of a rotor where the sample was restricted by PTFE spacers to the volume inside the main coil (center 2.5 mm of the 5.3 mm long cylindrical sample space in the rotor). Indeed only the low-field part of the distribution of B$_1$ fields was observed. It is important to note that this inhomogeneity of the $^2$H field does not deteriorate the decoupling performance in practice, because the part of the sample that experiences the higher $^2$H B$_1$ field is outside the main coil and is not detected in most experiments, not only because of the lower detection sensitivity but also because of inefficient excitation and polarization transfer for this part of the sample. To allow decoupling also during proton detection (which turns out to be
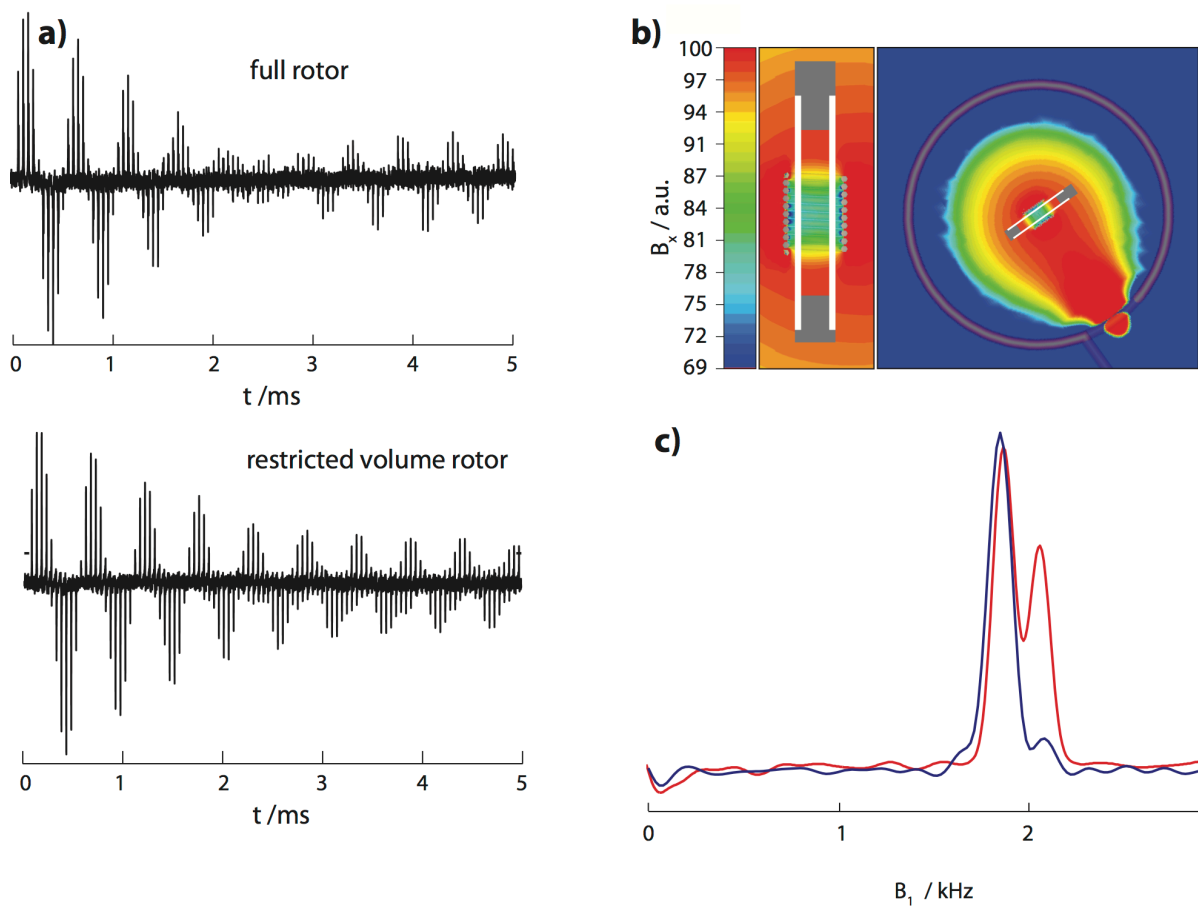
**Figure 3.3.** $B_1$ field strength and homogeneity induced in the sample by the external $^2$H coil. a) Stack plot of 1-pulse $^2$H experiments as a function of pulse length. b) Finite-Element simulated $B_1$ field distribution. On the right, a zoom on the 1.3 mm rotor inside the regular solenoid coil c) $B_1$ field strengths from the nutation experiment of the full rotor (red) and with the sample restricted to the volume inside the inner coil (blue). The 810°/90° ratio is 67%.

unnecessary in our experiments) we checked whether the rf on the $^2$H coil is picked up by the main coil, i.e. whether $^2$H decoupling creates noise in the $^1$H dimension. However, the noise level in the $^1$H FID with and without $^2$H decoupling is identical, indicating that $^2$H decoupling can be applied also during detection. We also investigated the possible impact of the additional coil on the performance of the main resonance circuit by testing rf fields and detection sensitivity on an alanine test sample before and after installation of the $^2$H coil. On the $^{13}$C channel, the power level for 100 kHz and the detection sensitivity remained unchanged. The $^1$H resonance frequency of the main coil decreased by about 5 MHz upon installation of the $^2$H coil, most likely due to parasitic capacitances. To allow tuning it to the Larmor frequency, the length of the transmission line was adjusted by sliding the telescope like inner conductor slightly together. The power requirement of the $^1$H channel increased by 2 dB and concurrently the detection sensitivity dropped by 15%, which is however significantly less than the gain due to the $^2$H decoupling with the installed $^2$H coil.

We then tested the decoupling performance by collecting 2D [$^{13}$C;$^1$H]-HSQC correlation spectra on a microcrystalline sample of the protein ubiquitin labeled with $^{13}$CHD$_2$ groups on Isoleucine-$\delta 1$ positions, and perdeuterated otherwise. We focus here on proton-detected spectra of these methyl groups. The one-bond scalar coupling between deuterium and carbon, $^1J_{CD}$ is approximately 20 Hz[78], and without decoupling the presence of two $^2$H nuclei gives rise to a quartet of lines with 1:2:2:1 intensity ratio, separated by the scalar coupling constant. The scalar coupling to the proton, $^2J_{HD}$, is much smaller, on the order of 2 Hz. We thus expect that scalar decoupling of $^2$H greatly reduces line widths of $^{13}$C, while the effect on a $^1$H dimension should be much smaller. Figure 3.4 shows experimental data of these methyl groups in the presence of a 2 kHz WALTZ decoupling field during the indirect $^{13}$C dimension (red) and without $^2$H decoupling. In all cases, the carbon-proton coupling was eliminated, either by a $\pi$ pulse applied to protons during $^{13}$C chemical-shift evolution, or by 2.5 kHz $^{13}$C WALTZ decoupling during $^1$H detection. The improvement in $^{13}$C line width provided by the $^2$H decoupling is shown by $^{13}$C FID data (Figure 3.4a) and frequency domain data (Figure 3.4b,c). Without decoupling $^{13}$C line-widths are 58$\pm$6 Hz (the $^1J_{CD}$ multiplet structure is not resolved), with decoupling they decrease to 21$\pm$3 Hz. The peak intensity increases by a factor of 2.8$\pm$0.5. This corresponds to a sensitivity gain of a factor 2.4 taking into account the 15% loss of $^1$H sensitivity upon installation of the $^2$H coil. This improvement indeed greatly accelerates data acquisition (by about a factor of 5.5 (i.e. 2.4$^2$) for a 2D spectrum, and a factor of 30 when methyl-methyl cross peaks are correlated, as in the context of structure determination[73]). We also tried applying $^2$H decoupling during $^1$H acquisition but did not
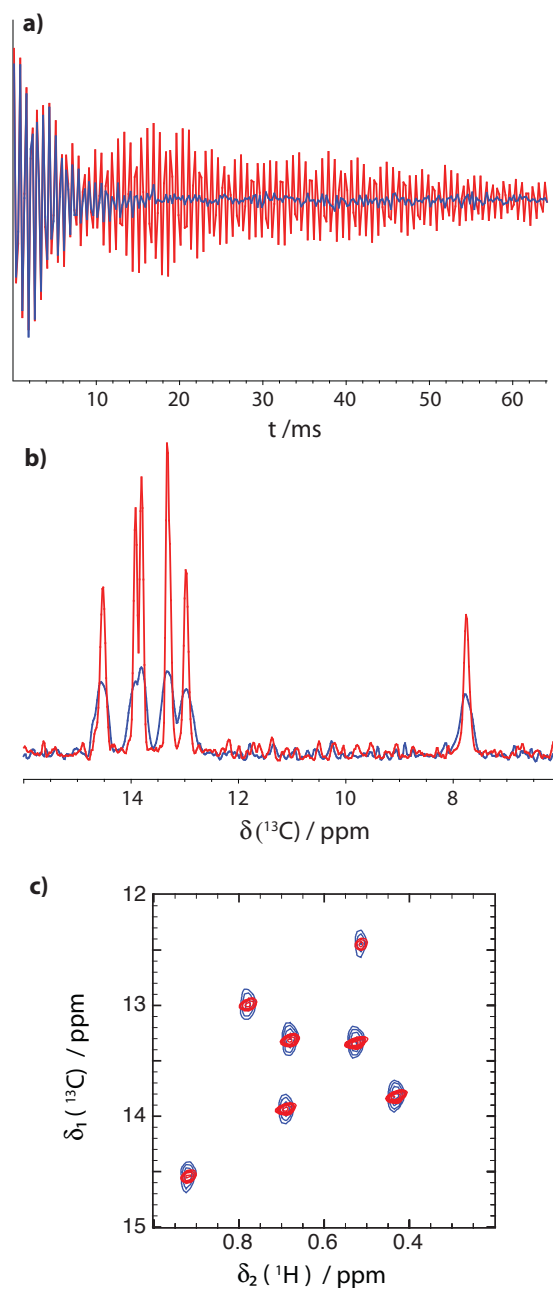
**Figure 3.4.** Effect of decoupling with the external $^2$H coil observed on [U-$^2$H;$^{15}$N, Ile-$^{13}$CHD$_2$] labeled ubiquitin micro-crystals at 55 kHz MAS. a) Indirect $^{13}$C FID extracted at the $^1$H frequency of Ile 3 from a proton-detected HSQC experiment with 2kHz $^2$H WALTZ decoupling (red) and without (blue). b) Projection of the resulting spectrum on the $^{13}$C axis. c) Overlay of the decoupled and not decoupled [$^{13}$C;$^1$H]-HSQC spectra. The peak at $\approx$ 8 ppm in (b) is aliased in this spectrum. The acquisition time for the 2D data set was 3 h. The decoupled peaks are slightly skewed due to correlated inhomogeneous broadening.

observe an improvement in line-width because other factors dominate.

In solution-state NMR the $^2$H circuit it typically used for field locking and decoupling. We used the auxiliary coil only for decoupling. Tests with a 3.2 mm probe and the main coil tuned to $^2$H have shown that it is possible to lock on the $D_2O$ line of microcrystalline protein samples but the sensitivity of the external coil is too low for this to be practical with a 1.3 mm rotor. Furthermore locking on the $D_2O$ resonance with fast magic-angle spinning (50kHz in this case) carries the risk of correcting for shifts that are actually due to thermal effects and not a change in the $B_0$ field.

## 3.4. Conclusions

We have shown that an auxiliary $^2$H coil outside of the MAS stator can be used for $^2$H decoupling in partially deuterated proteins, which greatly increases peak intensity and resolution. Deuterium-detection sensitivity is sufficient for easy calibration of the decoupling power level through a nutation experiment and direct $^2$H detection. The presented analytical and numerical calculations correctly predict the magnitude of the observed $B_1$ field as well as the form and origin of the inhomogeneity.

## 3.5. Experimental Methods

All NMR experiments were carried out on a Bruker Biospin AVANCE II+ spectrometer operating at 850 MHz $^1$H frequency using a 1.3 mm triple-resonance ($^1$H-$^{13}$C-$^{15}$N) probe (Bruker). The conductors (silver-plated copper wire, diameter 1 mm) of the auxiliary $^2$H coil were wound around PTFE disks, which are placed on the side of the stator. The tuning elements were placed below the stator chamber. The connecting coaxial cable was guided through a preexisting hole with the shield connected to the probe's frame at several points to achieve good grounding. Finite elements simulation of the $B_1$ field was performed with CST Microwave Studio 2010 using the time domain transient solver. The magnetic field was simulated at a fixed input current with minimum mesh size of 0.05 mm. [U-$^2$H$^{15}$N, Ile-$^{13}$CD$_2$H] ubiquitin was produced and crystallized as described in [73] with the exception that 4-$^{13}$CHD$_2$-3-d$_2$-$\alpha$-ketobutyrate (75 mg/L, Sigma) was added to the bacterial medium instead of $\alpha$-ketoisovalerate. [$^{13}$C;$^1$H] HSQC spectra shown in Figure 4 were with 512 $t_1$ increments and 16 scans, leading to an experimental time of 3 h for each spectrum. $B_0$ drift (maximally 20 Hz)

was corrected by applying a "first order phase correction" on the $^1$H free induction decay leading to a frequency shift in the spectrum. The angle was increased linearly between different $t_1$ traces.

# 4. Structure Calculation from 3D- and 4D $^1$H –$^1$H  DREAM spectra

This chapter is based on the following publication:

Matthias Huber, Sebastian Hiller, Paul Schanda, Matthias Ernst, Anja Böckmann, René Verel and Beat H. Meier, **2011**. A Proton-Detected 4D Solid-State NMR Experiment for Protein Structure Determination, *PhysChemPhys*, 12:915.

Sebastian helped to implement acquiring and processing non-uniformly sampled 4D spectra. Paul and René measured the 3D amide–amide spectrum. I prepared the samples, measured the methyl–methyl spectrum, performed structure calculations and wrote the initial manuscript.

## 4.1. Introduction

Owing to recent advances in instrumentation, methodology and sample-preparation techniques, solid-state NMR spectroscopy is providing unique insights into biological structures at atomic resolution. Three-dimensional structures of proteins and other biological macromolecules can now be determined that are difficult to characterize by x-ray diffraction or solution NMR[46, 52, 79, 80]. Still, the task remains demanding and new methodological developments are needed to make the structure determination more reliable. While it has been demonstrated that assignments are feasible for proteins with over 200 residues[32, 81]. The structure-determination step remains difficult, mainly because spectral overlap introduces ambiguities in restraint assignments during the process of structure generation. These ambiguities can in some cases be resolved[16, 40, 42, 46, 82–86], but remain a potential source of errors. Here, we present an experimental approach that combines the efficient measurement of long-range proton-proton distances in sparsely isotope-labeled samples with proton-detected 3D and 4D correlation spectroscopy. We demonstrate the method in the context of structure determination of the 76 amino-acid protein ubiquitin.

Sparsely distributed protons in an otherwise perdeuterated protein yield highly resolved $^1$H spectra[61–63, 87]. The corresponding samples are produced either by expressing perdeuterated protein followed by $^1$H back exchange at about 30% of the exchangeable sites[63] or by using suitable precursors to label exclusively a methyl group of alanine, isoleucine, valine or leucine during protein expression[19]. Under fast magic-angle spinning (55 kHz), coherences in such spin systems are sufficiently long lived to allow not only dipolar, but also scalar-coupling based polarization transfer[34, 88]. $^1$H-detection increases the sensitivity by a factor of 8 compared to $^{13}$C detection[89].

Similar deuteration schemes have been exploited in solution-state NMR to increase resolution and sensitivity[2, 90, 91], and to collect precise NOE restraints[91]. An additional advantage arises in the solid phase: in such samples, even the closest $^1$H neighbor often corresponds to a long-range contact and dipolar truncation[26] is not an issue allowing highly-efficient first-order dipolar-recoupling experiments to be used to obtain distance information.

## 4.2. Results and Discussion

Here we use the DREAM[25] mixing scheme (Figure 4.1) that provides particularly high polarization-transfer efficiencies due to its adiabatic nature. It can be seen from Figure 4.2a that the resolution in a 2D [$^{13}$C;$^1$H]-HSQC spectrum of crystalline [U-$^2$H$^{15}$N; VL-$^{13}$CHD$_2$] ubiquitin, is high enough to unambiguously identify the resonances of all 26 methyl groups and that the efficient INEPT transfer allows to record a 2D spectrum with high signal-to-noise ratio ($\approx$30) in 10 min. The abbreviation [U-$^2$H$^{15}$N; VL-$^{13}$CHD$_2$] refers to perdeuterated and $^{15}$N labeled protein with one and only one $^{13}$CHD$_2$ group on all Val and Leu residues with no stereospecificity. Connecting two such 2D-HSQC elements with a DREAM mixing time leads to the 4D experiment of Figure 4.1. Given the resolution in the 2D, all observed H$^{met}$–H$^{met}$ peaks in the corresponding 4D spectrum are resolved and can unambiguously be assigned.
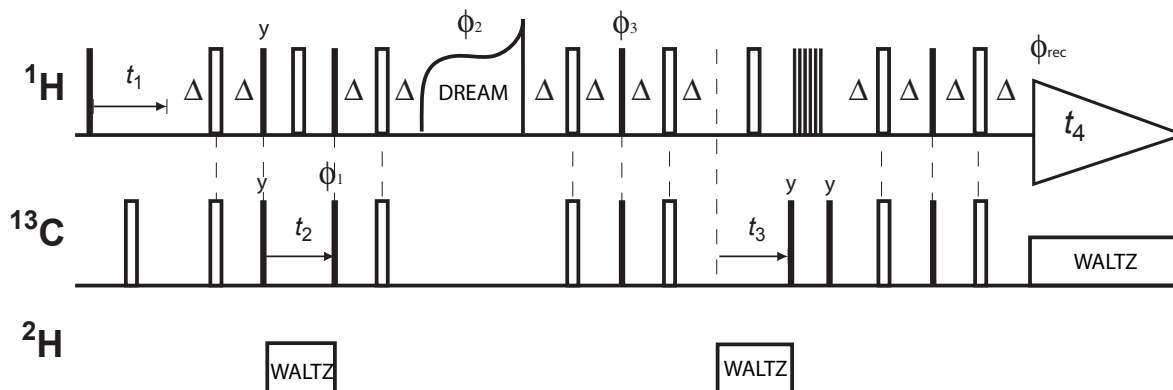


**Figure 4.1.** The 4D methyl-methyl DREAM experiment for [U- $^2$H; VL-$^{13}$CHD$_2$ ] labeled solid-state protein samples under MAS. Filled and open bars denote 90° and 180° pulses, respectively. The block of narrow bars denotes saturation pulses. The phase cycle is: $\phi_1$: [y, -y]; $\phi_2$: [x, x, x, x, -x, -x, -x, -x]; $\phi_3$: [x, x, -x, -x]; $\phi_{rec}$: [x, -x, -x, x], all unlabeled pulses were applied with x phase. See section 4.4.2 for experimental details. An optimized sequence is proposed in Figure 8.7 in the appendix.
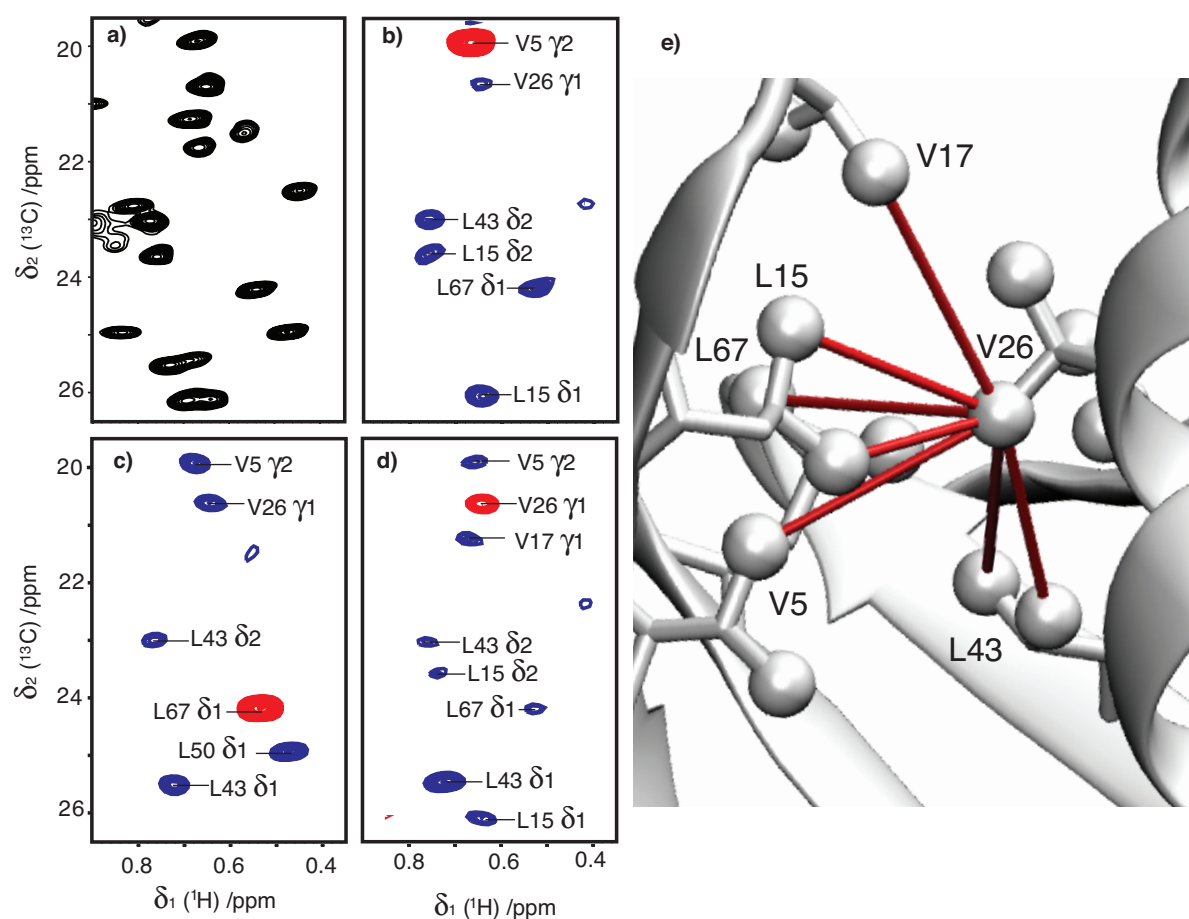
**Figure 4.2.** [$^{13}$C;$^1$H]-HSQC spectrum of crystalline [U-$^2$H$^{15}$N; VL- $^{13}$CHD$_2$] ubiquitin recorded 55 kHz MAS. The full spectrum is shown in Figure 4.3a. b-d) 2D planes from the non-uniformly-sampled (NUS) 4D H$^{met}$–H$^{met}$ DREAM experiment. Negative cross peaks are blue, positive diagonal peaks red. b) 2D plane of V5 C$\gamma$2 c) 2D plane of L67 C$\delta$1. d) 2D plane of V26 C$\gamma$1 e) $^1$H-$^1$H Structural restraints obtained from the 2D slice through V26 (Figure 2b) shown on the structure of Ubiquitin calculated here. Spheres indicate $^{13}$CHD$_2$ groups.
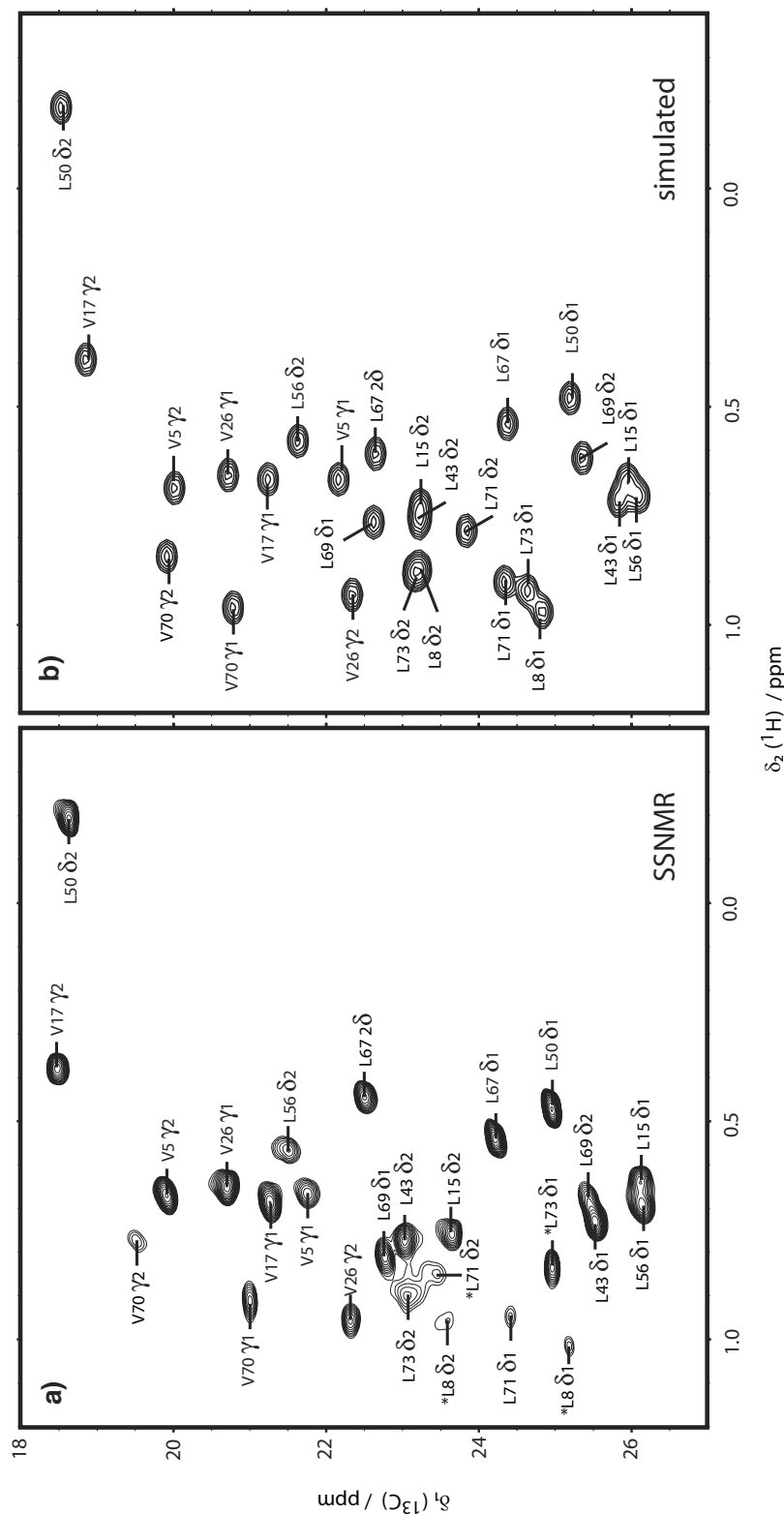
**Figure 4.3.** Assignment of methyl groups.  a) 2D [$^1$H;$^{13}$C]-HSQC spectrum of [U-$^2$H$^{15}$N; VL-$^{13}$CD$_2$H]-Ubiquitin crystallized from a solution containing 80% D$_2$O and 20% H$_2$O recorded at 55 kHz MAS and assignments based on the simulated spectrum (b) and confirmed by a 3D H$^{methyl}$C–H$^N$ DREAM experiment (Figure 4.12). Asterisks (*) indicates assignments based solely on the solution NMR shifts. b) Simulated $^1$H-$^{13}$C spectrum with stereospecific ssNMR $^{13}$C shifts (BMRB 7111) complemented with solution NMR (BMRB 5387) $^1$H and $^{13}$C shifts.

Fully sampling the three indirect dimensions of the 4D experiment would require a prohibitively long measurement time. Due to the high experimental sensitivity, the measurement time of the 4D $H^{met}$–$H^{met}$ DREAM is clearly sampling limited for ubiquitin and can thus be circumvented by non-uniform sampling (NUS) of the indirect time domains[92]. In our spectrum, 14% of the data points were randomly chosen and sampled in an experiment that took 72 h.

Fourier Transformation (FT) of the sparse time domain data with missing points set to zero produces artifacts described by the (unknown) signal convoluted with the FT of the sampling pattern. In the case of a random sampling pattern the artifacts although deterministic, can be appreciatively described as random noise. The apparent noise in the NUS spectrum is the sum of the thermal noise and the pseudo noise from sparse sampling artifacts. Non-linear processing methods[68, 93–97] have been introduced to minimize sampling artifacts. To investigate the ratio of sampling artefacts to thermal noise, we have created an artificial data set based on the experimentally observed peak intensities and line widths and deleted a given number of data-points. In our case more than 90% of the apparent noise is thermal noise (see Figure 4.4). The artifacts due to sparse sampling are low because the intensities of the various cross- and diagonal peaks are all within about one order of magnitude. Advanced reconstructions methods can only decrease sampling artifacts, not thermal noise. We have processed our data with (Co)-MDD[68] and Forward Maximum Entropy Reconstruction[69]. The results did not appear to be significantly better than direct FT of the sparse data. It has been shown that these methods are of great value if some signals are much stronger than thermal noise [68, 69].

Three 2D planes from the 4D methyl-methyl DREAM experiment are shown in Figures 4.2 b, c, and d. They characterize the source of the magnetization that arrives after DREAM mixing at the methyl groups of V5, L67 and V26, respectively. In total 49 distance restraints, all of them spectrally unambiguous, were extracted from the 4D data set.

The 49 $H^{met}$–$H^{met}$ distance restraints are structurally important since they yield restraints between methyl groups that are particularly abundant in the hydrophobic core of the protein. They were, however, not sufficient for obtaining the entire protein structure. We complemented them by TALOS+ predictions from the $^{13}C$ resonances and by measuring $H^N$–$H^N$ distance restraints in a perdeuterated sample where the $H^N$ (and other exchangeable) were backprotonated to 30%. While this could have been done with a similar 4D experiment (using $^{15}N$ instead of $^{13}C$), we recorded the slightly more sensitive 3D version of the experiment, a fully sampled 3D $H^N$–$H^N$ DREAM, in 24h (Figure 4.5, pulse sequence in Fig 4.6). 139 $H^N$–$H^N$ cross peaks could be identified in the 3D spectrum. DREAM transfer was observed
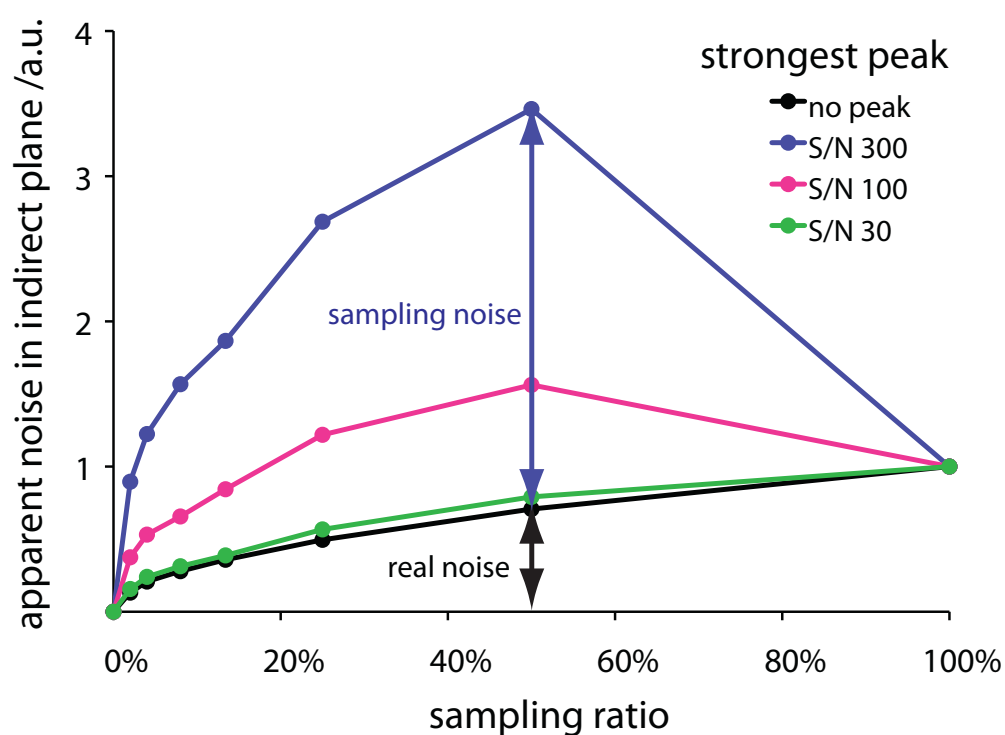
**Figure 4.4.** Noise intensities in indirect planes as a function of the sampling ratio for different signal-to-noise ratios of the strongest peak present in the spectrum.
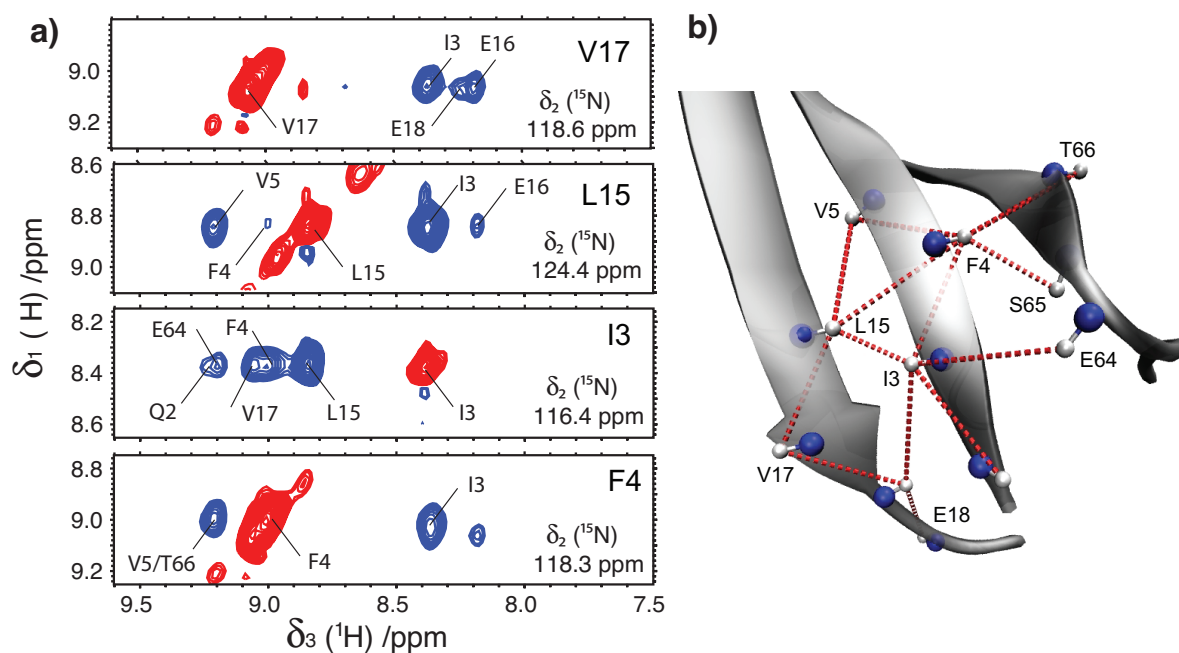
**Figure 4.5.** Spatial correlation of $H^N$–$H^N$ pairs with the 3D DREAM experiment. a) Four representative strips from a proton-detected 3D $^{15}$N-resolved [$^1$H;$^1$H]-DREAM experiment (Figure 4.6) of [U-$^2$H$^{15}$N; VL-$^{13}$CHD$_2$]-Ubiquitin. Negative and positive peaks are red and blue, respectively. b) Selected region of the ubiquitin structure with $H^N$–$H^N$ distance restraints derived from the observed cross peaks in a). Amide moieties are shown in ball-and-stick representation with their assignment given.
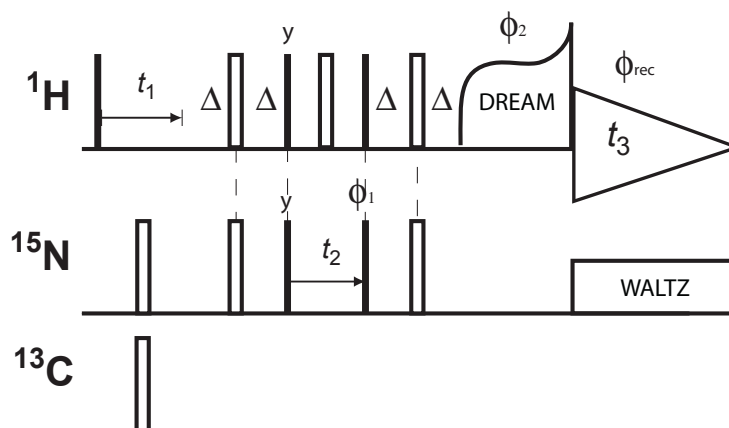


**Figure 4.6.** 3D $H^N$–$H^N$ DREAM experiment for [U-$^2$H$^{15}$N] labeled protein samples. Filled and open bars denote 90° and 180° pulses, respectively. The INEPT delay $\Delta$ is set to 2.1 ms. The tangential shape denotes a $^1$H–$^1$H DREAM transfer element with 4 ms length and amplitude modulation between 30% and 70% of the MAS frequency a phase angle of 87° as defined in eq. 4.1 in section 4.4.2. All unlabeled pulses are applied in x phase. The phase cycle is: $\phi_1$: [y, -y]; $\phi_2$: [x, x, -x, -x]; $\phi_{rec}$: [x, -x, x, -x].
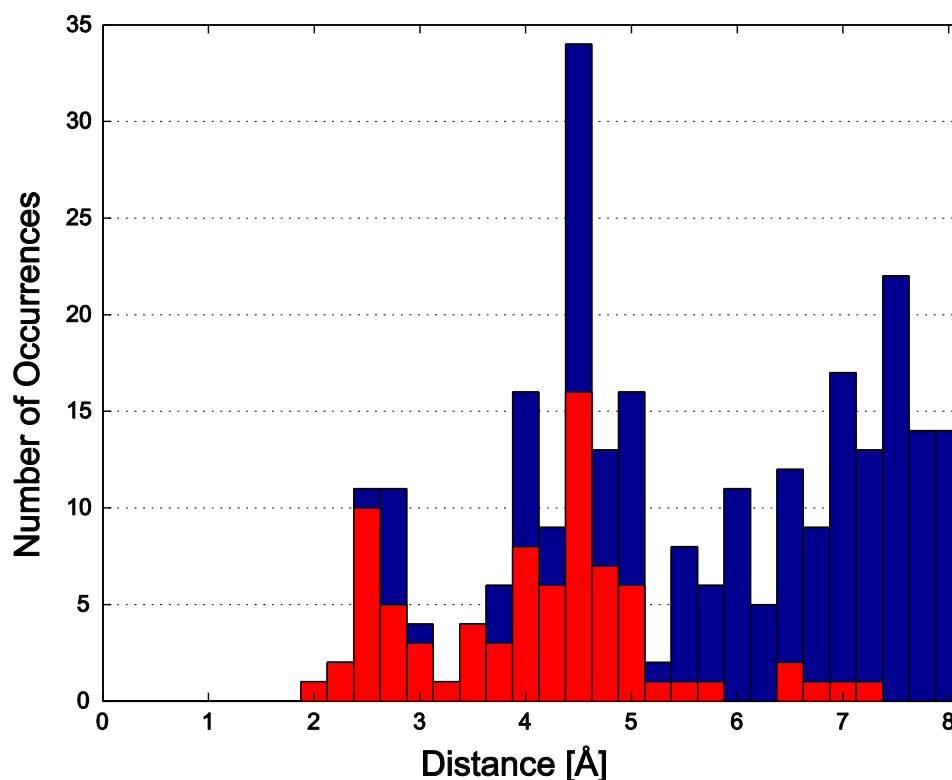
**Figure 4.7.** Histogram of possible (blue) and observed (red) $^1$H–$^1$H amide distances. In both cases the distances are extracted from the solution state NMR. The blue dataset corresponds to all potential $^1$H–$^1$H distances, whereas the red dataset is based on those contacts which were actually observed and assigned in the experimental 3D experiment. The red histogram qualitatively reproduces the shape of the blue plot but with a cutoff around 5-6 Å.

for $^1$H–$^1$H distances up to 6 Å for a transfer period of 8 ms. All cross peaks in both spectra had negative intensity, indicating a direct transfer mechanism rather than a relayed two-step transfer[98]. Figure 4.7 shows a histogram of expected and observed $^1$H–$^1$H amide contacts as a function of distance.

Structures were calculated using CYANA[70] 2.1 using the CANDID[43] automatic assignment routine for the H$^N$–H$^N$ peaks. For structure calculation the peaks of the 4D methyl–methyl spectrum were manually assigned. For the observed contacts between methyl protons (see Table 8.8 in the Appendix) distance restraints between the corresponding pseudoatoms were created. Upper distance limits were set to a fixed value without taking peak intensity into account. The length of the restraints was optimized as described in [52] (i.e. set to the lowest value that does significantly increase the CYANA target energy) and was set to 4.5 Å. This distance is relatively short and might force side chains into an unnatural conformation. Even
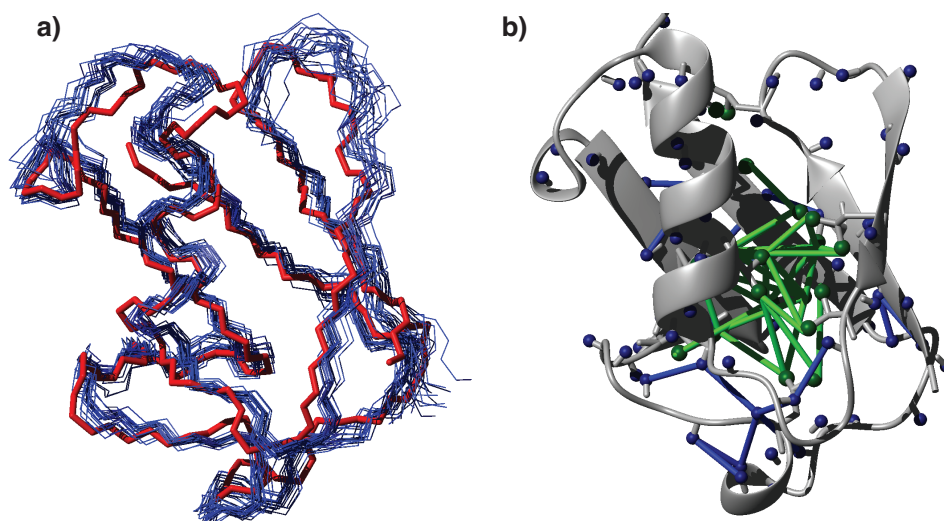
**Figure 4.8.** Structure determination of ubiquitin from proton-detected DREAM spectra. a) Bundle of the 20 lowest-energy conformers from 200 CYANA calculations using proton-proton distance restraints from a 3D $H^N$–$H^N$ DREAM and a 4D $H^{met}$–$H^{met}$ DREAM spectrum. (backbone RMSD 0.72 $\pm$ 0.12 Å; bias to x ray structure 1UBQ 1.57 Å, red overlay) . b) Extent of the constraint network shown on the lowest energy structure calculated here (PDB 2L3Z). labeled sites are shown as spheres. Constraints obtained from the 4D $H^{met}$–$H^{met}$ DREAM experiment are shown as green lines, constraints from the 3D $H^N$–$H^N$ dream as blue lines on the structure of ubiquitin.
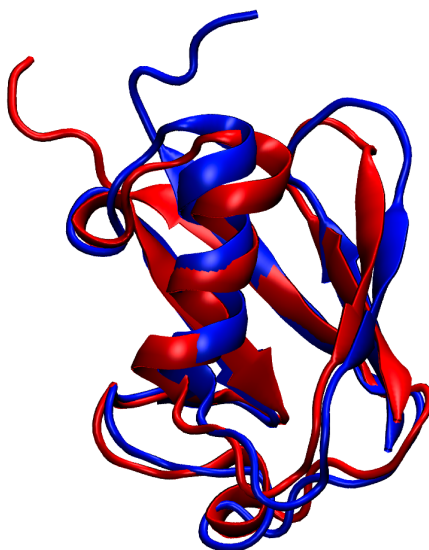


**Figure 4.9.** Cartoon representation of the structure calculated in this paper (blue, pdb: 2L3Z) and the x-ray structure (red, pdb: 1UBQ). Secondary structure elements are drawn as identified by the VMD viewer. The two structures were aligned by minimizing the backbone RMSD of the non-flexible residues 1-7 and 12-70.

| Cyana target function | average | 0.18 |
|---|---|---|
| | min | 0.10 |
| | max | 0.24 |
| violated distance constraints: | | 0 |
| violated angle constraints: | | 0 |
| violated van-der-waals contacts: | | 1 |
| | | (6 of 20 models, max 0.28 A) |
| *Procheck Ramachandran Analysis* | | |
| Residues in most favoured regions | 1244 | 94.2% |
| Residues in additional allowed regions | 75 | 5.7% |
| Residues in generously allowed regions | 1 | 0.1% |
| Residues in disallowed regions | 0 | 0.0% |

**Table 4.1.** Statistics of the structure calculation and the resulting bundle of conformers.

when the 4D spectrum was combined with 2D-[$^{13}$C;$^{13}$C] DARRs, structure calculations with a longer restraint distance for the 4D peaks decreased the quality the obtained structure (i.e. the RMSD of the obtained bundle and the RMSD to the reference structure increased).

The 139 3D H$^N$–H$^N$N peaks (out of which 72 were unambiguous, and 12 long-range) were compiled into an unassigned peak list and fed to the CANDID[43] routine to provide the UPLs. The distance calibration feature of CANDID was not used and restraints were uniformly set to 5.5 Å. In combination with TALOS+[38] dihedral-angle restraints derived from $^{13}$C chemical shifts, we obtained a bundle of conformers with a bias to the x-ray structure of 1.57 Å  (Figure 4.8a and Figure 4.9) and an RMSD of 0.72±0.12 Å. The final refinement in water was done using Opalp[99].  RMSDs were calculated based on non-flexible residues only (1-7 and 12-70).  Statistics of the structure calculation are given in Table 4.1.  The calculation proofed robust against variation of input parameters. The distance restraints used for the calculation are indicated in Figure 4.8b. Because two pairs of assignments (L43$\delta$2/L15$\delta$2 and L56$\delta$1/L15$\delta$1) were very likely but not completely certain, the corresponding restraints were, in an additional calculation, treated as ambiguous restraints in CYANA. The resulting structure was identical within the error margins.

Analysis of several variants of this structure calculation, using only subsets of the restraints (see Table 4.2) showed that the TALOS+ dihedral-angle restraints are necessary to obtain the secondary structure elements, the H$^N$–H$^N$ restraints define the arrangement of $\beta$-sheets, whereas the H$^{met}$–H$^{met}$ restraints ensured the correct global fold. H$^N$–H$^N$ restraints alone are neither sufficient to determine the structure to good precision (see Table 4.2)

**Table 4.2.** Structure calculation details for subsets of restraint classes.

| Restraints Used | | | RMSD of bundle /Å | Bias to xray structure /Å |
|---|---|---|---|---|
| 4D H$^{met}$–H$^{met}$ | 3D H$^{N}$–H$^{N}$ | TALOS dihedral angles | | |
| yes | yes | yes | 0.64±0.13 | 1.60 |
| yes | yes | no | 6.95±1.19 | 7.19 |
| yes | no | yes | 4.97±1.75 | 3.31 |
| no | yes | yes | 5.12±2.78 | 11.81 |
| yes | no | no | 7.77±1.12 | 6.94 |
| no | yes | no | 11.38±2.43 | 15.32 |
| no | no | yes | 11.01±1.46 | 11.77 |

## 4.3. Conclusions

It will be interesting to apply the method to larger proteins whose structure cannot yet be solved by solid-state NMR. Larger protein will lead to spectra with more signal overlap and decreased signal intensity, but roughly the same line width.

To obtain an estimate of the applicability of the present approach to larger proteins, we considered the three aspects resolution, sensitivity and assignability using the 82 kDa enzyme malate synthase G, for which solution shifts are available, as an example[100]. For larger proteins we expect, of course, spectra with more signal overlap and a signal intensity that decreases linearly with the molecular mass. We assume a constant line width irrespective of the size. We simulated the expected 2D [$^{13}$C;$^{1}$H]-correlation spectrum using the published chemical shifts[100] and the line widths observed in the present work (40 Hz for $^{1}$H, 30 Hz for $^{13}$C) (Figure 4.10) and predicted the expected signal overlap. From this calculation, we estimate that, in a 4D spectrum, the majority of the methyl peaks will be sufficiently separated from their respective nearest neighboring peak to allow an unambiguous assignment (Figure 4.11). When extrapolating from the present work, a loss of a factor 10 in signal per resonance is expected when going towards a 10-fold larger protein. However, the sensitivity can also be increased by a factor of four by employing stereospecific labeling of methyl groups[21]. The remaining factor of 2.5 can be compensated for by longer experiment times, and possibly the use of a larger MAS rotor. The third requirement is sequential assignments, here a combination of state-of-the-art solid-state NMR protocols[32, 81] and the adaptation of solution NMR strategies for methyl assignment can be employed[101].
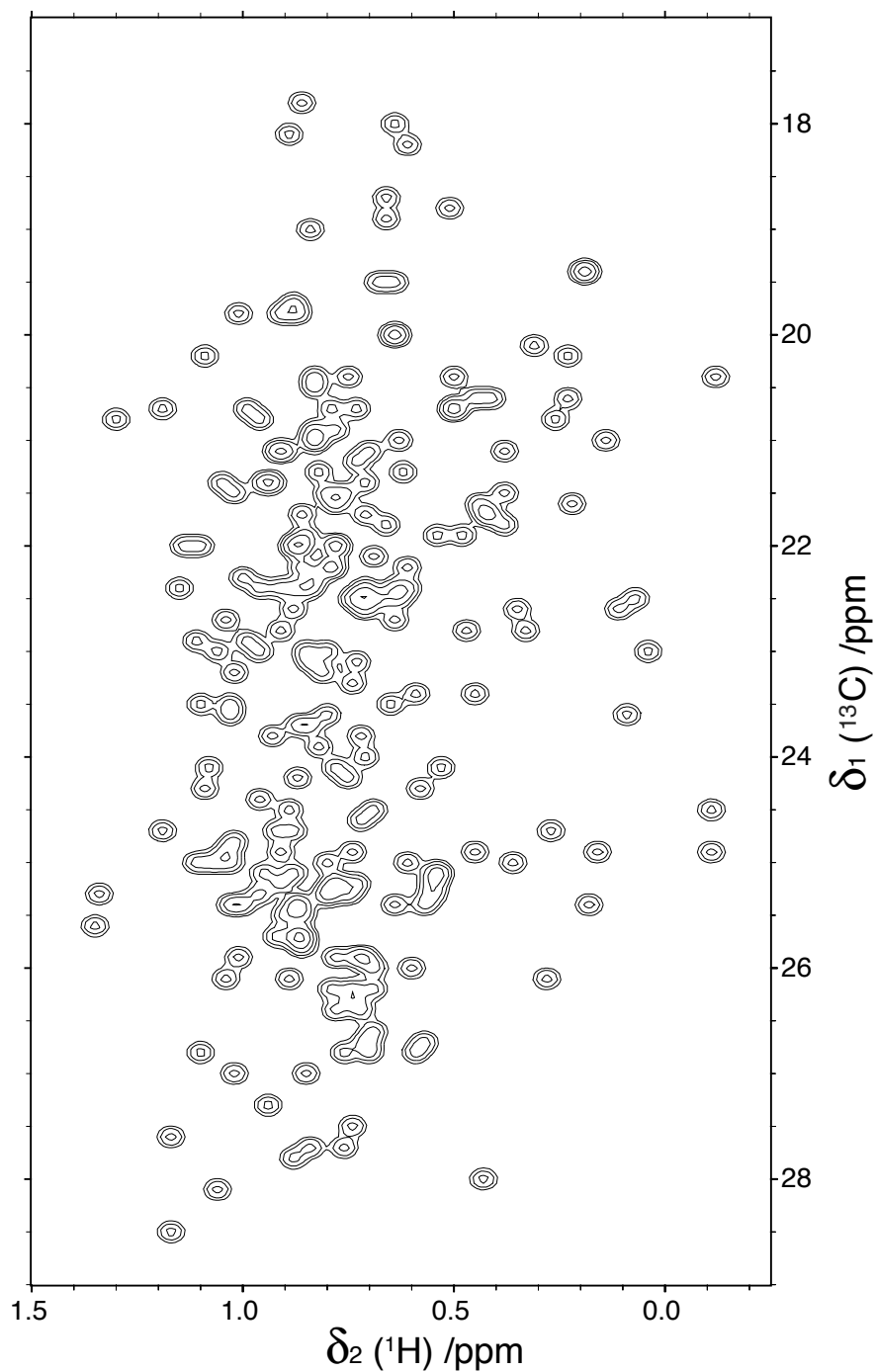
**Figure 4.10.** Simulated [$^{13}$C;$^1$H]-HSQC spectrum of [VL-$^{13}$CHD$_2$]- malate synthase G (shifts from ref[21]) with line width of 30 Hz ($^{13}$C) and 40 Hz ($^1$H) and the lowest contour level at 25% signal intensity.
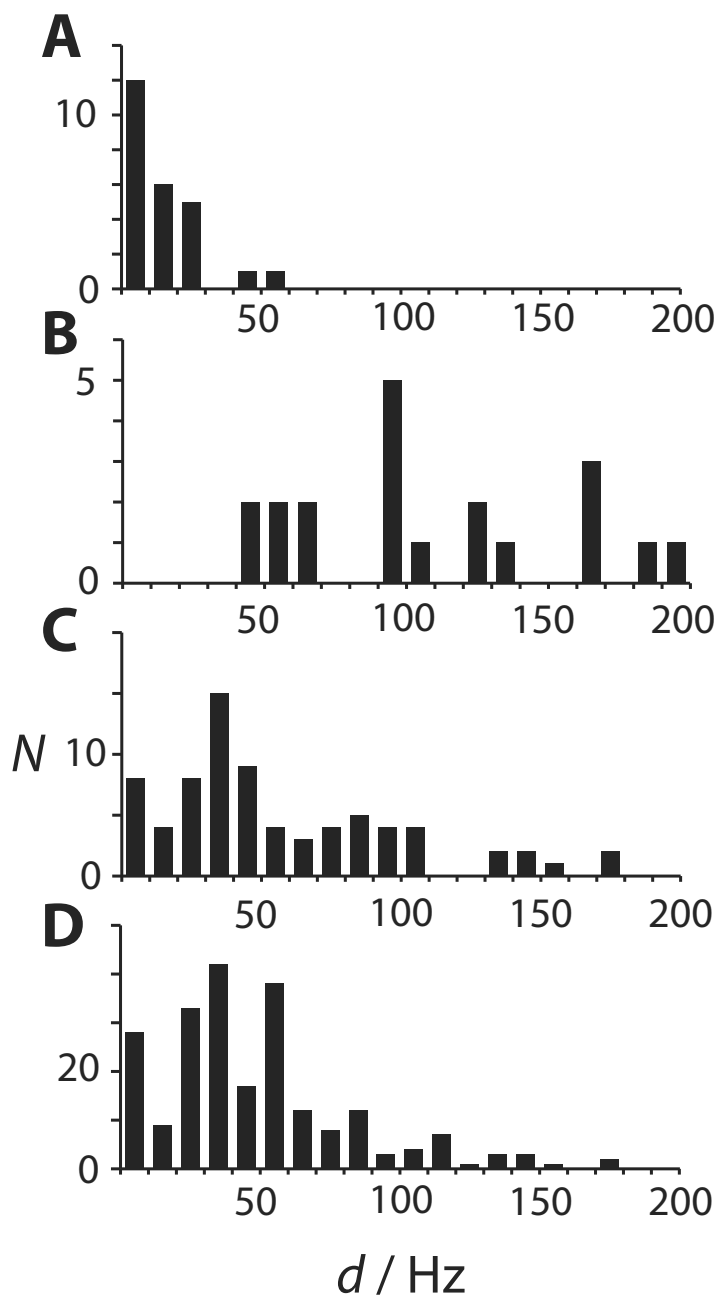
**Figure 4.11.** Histogram of the spectral distances of each methyl peak to its nearest neighbor for different proteins and types of NMR spectra. Spectral distances of less than the experimental line width in the current study are highlighted in red. A) Ubiquitin, $^1$H 1D. B) Ubiquitin (76aa) 2D [$^1$H;$^{13}$C]-HSQC, C) VDAC (283aa) 2D [$^1$H;$^{13}$C]-HSQC, D) MSG (723aa) 2D [$^1$H;$^{13}$C]-HSQC The effective separation in 2D has been calculated according to the following formula, scaling the separation in the $^{13}$C dimension with the experimentally observed difference in line width $\lambda$. (e.g.

$$d = \sqrt{(d_{1_H})^2 + \left(\frac{\lambda_{1_H}}{\lambda_{13_C}} d_{13_C}\right)^2} = \sqrt{(d_{1_H})^2 + (1.33 \cdot d_{13_C})^2})$$

In conclusion, using ubiquitin as model system we have shown that proton-detected [$^1$H;$^1$H] based 3D and 4D spectra of sparsely methyl- and amide-labeled samples contain sufficient information to determine de novo a protein structure in the solid state, when used in combination with dihedral-angle restraints derived from $^{13}$C secondary chemical shifts. The good dispersion in the spectra and the high sensitivity provided by the use of proton detection should make the protocol applicable to considerably larger proteins and we are presently exploring these opportunities.

## 4.4. Experimental Methods

### 4.4.1. Sample Preparation

[U-$^2$H$^{15}$N; VL-$^{13}$CD$^2$H] ubiquitin was produced by bacterial overexpression in D$_2$O-based M9 medium containing $^2$H-glucose (2g/L) and $^{15}$NH$_4$Cl (1g/L) with 3-$^{13}$CHD$_2$-$\alpha$-ketoisovalerate (125mg/L, CIL) added one hour before induction[19], and purified using standard protocols. Crystals were obtained using $^2$H$_{12}$-2-methyl-2,4-pentanediol (MPD) as described[102] (see also section 8.1.2 in the Appendix). The solutions used for crystallization of the methyl-labeled samples had an H$_2$O/D$_2$O ratio of 20%/80% (with the source of the protons being mainly the exchangeable protons of MPD), the ones for the sample used to collect amide restraints 30%/70%. About 4 mg of protein crystals were packed into a 1.3 mm Bruker rotor by ultracentrifugation[103].

### 4.4.2. NMR Spectroscopy

All NMR experiments were carried out on a Bruker Biospin AVANCE II+ spectrometer operating at 850 MHz $^1$H frequency using a 1.3 mm triple-resonance ($^1$H,$^{13}$C,$^{15}$N) probe (Bruker) upgraded with a home-built $^2$H coil positioned outside the stator. Chemical-shift calibration used sodium 3-(trimethylsilyl)-1-propanesulfonic acid (DSS) as an internal standard. The sample temperature was determined from the chemical shift of supernatant water[103] and adjusted to 20±5 °C. Spectra were processed with nmrpipe[104] or prosa[105] and analyzed with Sparky or NMRViewJ. Time points not recorded in the sparse 4D data set were set to zero before Fourier transformation. The MISSISSIPPI[106] scheme was used for water suppression.

The 4D methyl-methyl DREAM experiment for [U-$^2$H; VL–$^{13}$CHD2] used an INEPT delay

$\Delta$ = 1.65 ms, decoupling with WALTZ[107] with 5 kHz on $^{13}$C and 1.25 kHz on $^{2}$H and saturation pulses of 40 ms at 5kHz. The a [$^{1}$H;$^{1}$H] DREAM[25] transfer-period was 8 ms, with an tangential amplitude modulation between 16.5 and 38.5 kHz. Quadrature detection for the indirect dimensions $t_1$–$t_3$ was done with the States-TPPI method[108]. The sampling grid for the 14% random sampled had the dimension of 26 x 30 x 30 complex points in $t_1$-$t_3$ and the spectral width was 1.5 kHz for $^{1}$H and 2 kHz for $^{13}$C. For the 3D H$^{N}$-H$^{N}$ DREAM 74 x 64 complex indirect points ($t_1$ x $t_2$) were recorded with a spectral width of 6 kHz for $^{1}$H and 3.25 kHz for $^{15}$N. 8/4 scans were acquired for each hyper complex component with a recycle delay of 1.2 s and 1.1 s for the 4D and 3D DREAM respectively.

## Assignment of Methyl Resonances

Methyl [$^{1}$H,$^{13}$C]-HSQC peaks were assigned by comparison with known $^{13}$C solid-state chemical shifts[109] and solution-state shifts[110]. The good correlation between solid- and solution-state $^{1}$H methyl chemical shifts has been reported previously[62]. Remaining ambiguities were resolved with an additional H$^{met}$C$^{met}$-H$^{N}$ 3D spectrum (Figure 4.12).

## Details of the DREAM mixing period

The amplitude during the DREAM pulse is described by[25]:

$$\omega_{DREAM}(t) = \bar{\omega} + \delta_{est} \tan\left(\alpha\left(1 - \frac{2t}{\tau_{DREAM}}\right)\right) \tag{4.1}$$

for $0 \le t \le \tau_{DREAM}$ with the angular variation of the tangent defined as

$$\alpha = \arctan\left(\frac{\Delta}{d_{est}}\right) \tag{4.2}$$

The average field strength was set to the nominal HORROR condition of half the rotation frequency or $\bar{\omega}/(2\pi)$ = 27.5 kHz as calibrated by an independent nutation experiment. The initial and final offset from the central amplitude was set to $\Delta(2\pi)$ = 11 kHz and the angular variation of the tangent to $\alpha = 65°$. The total duration of the pulse, $\tau_{DREAM}$, was 8.0 ms.
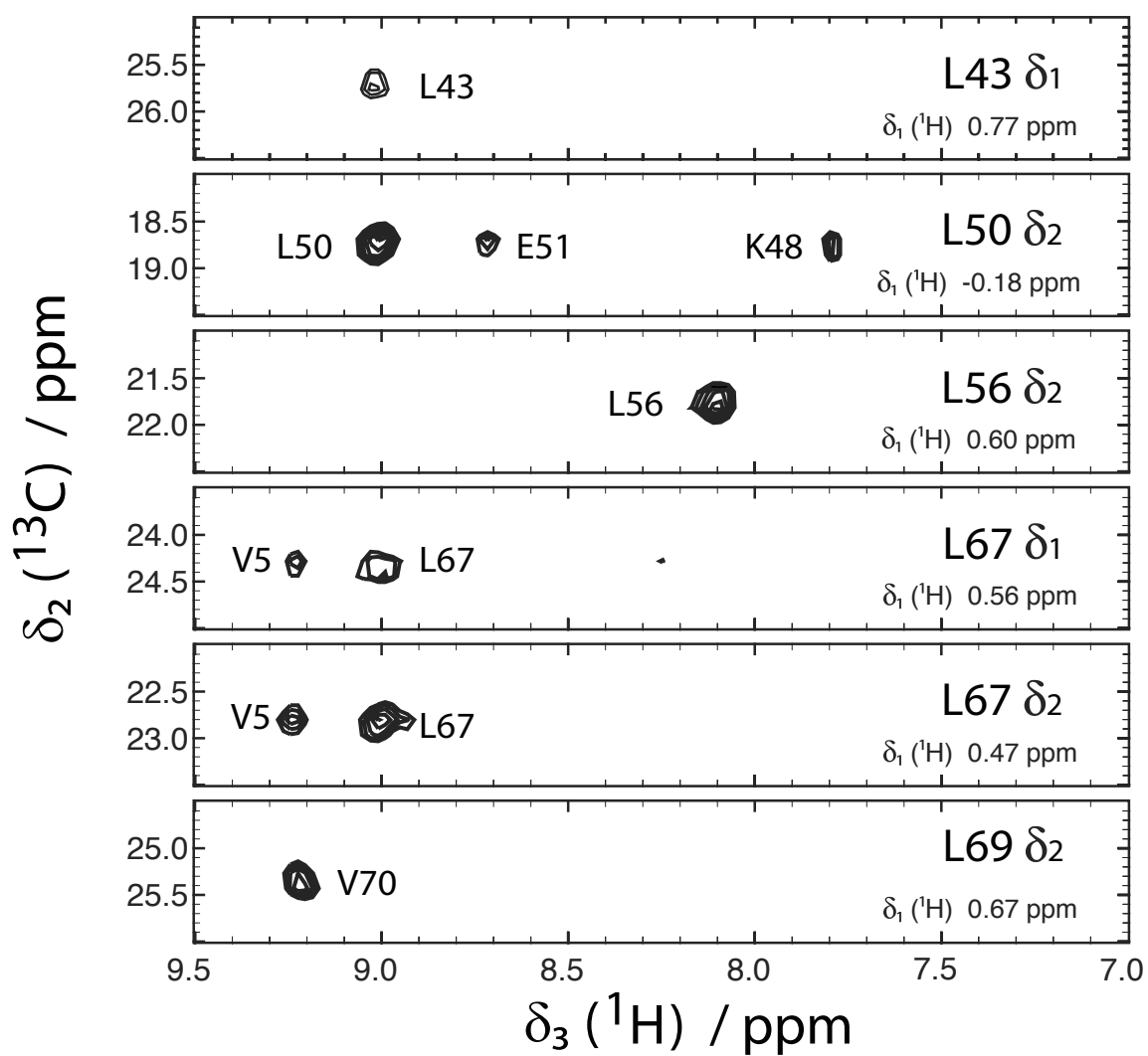
**Figure 4.12.** Spatial correlations of H$^{methyl}$–H$^N$ pairs with the 3D DREAM experiment. Representative strips are extracted for the methyl groups indicated and the most likely assignment possibility of the amide shift is given (an assigned [$^1$H;$^{15}$N]-HSQC spectrum is shown in Figure 4.13)
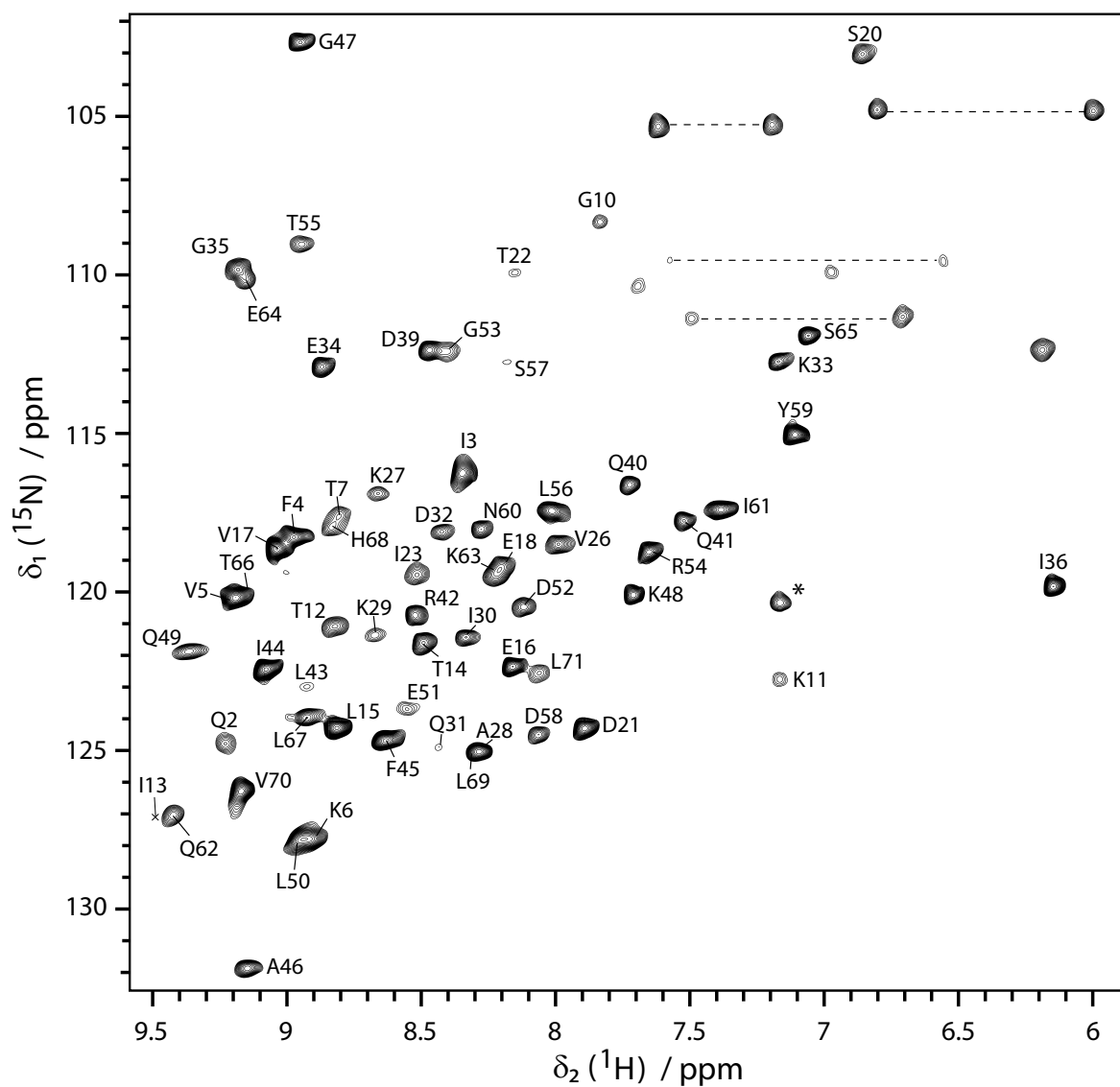
**Figure 4.13.** 2D [$^1$H;$^{15}$N]-HSQC spectrum at 55kHz MAS of [U-$^2$H$^{15}$N$^{13}$C]-Ubiquitin crystallized from a (60% (w/v) MPD-$^2$H$_{12}$ / 40% 50mM citrate pH 4.1 in D$_2$O) solution as described in ref[102]. The aqueous phase was adjusted such that 30% (70%) of the exchangeable sites in the protein were occupied with $^1$H($^2$H). Sequence-specific assignments are indicated and were derived from proton-detected 3D HNCO, HNCA and HN(CO)CA experiments as in ref[88]. Unlabeled peaks connected by a dashed horizontal line are side-chain resonances. A strip plot of the HNCA and HN(CO)CA spectra is shown in Figure 4.14

.
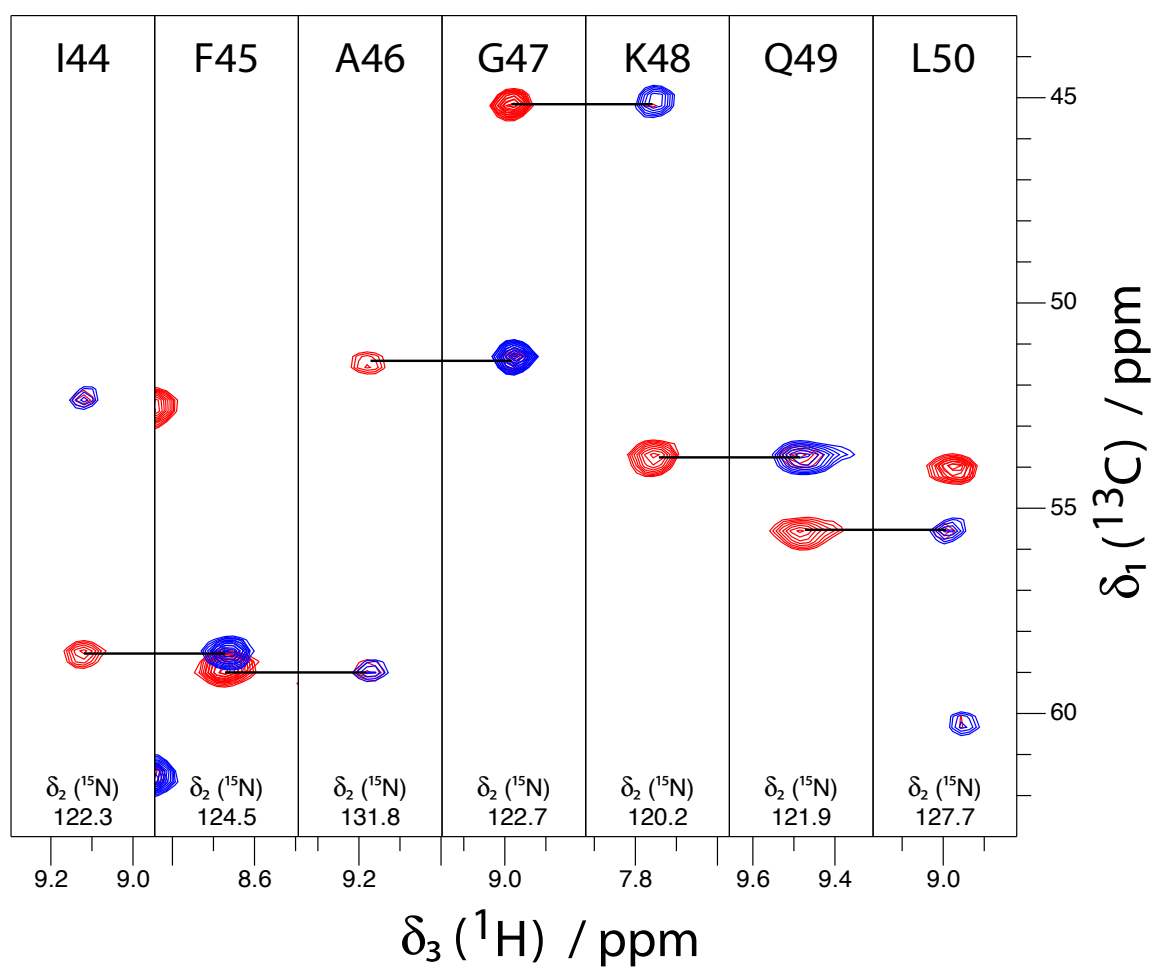
**Figure 4.14.** Backbone walk from the HNCA (red) and HN(CO)CA spectra of ubiquitin. Strips were extracted at the $^{15}$N and $^1$H frequency of residue specified on top. Black horizontal lines show identical C$\alpha$ shifts in the HNCA and HN(CO)CA strips of adjacent amino-acid residues.

# Part II.

# Application to Amyloid Fibrils

# 5. Osaka Mutant A$\beta$1-40 E22$\Delta$ fibrils

This chapter is based on a manuscript by Matthias Huber, Oxana Y. Ovchinnikova, Anne K. Schütz, Toni Vagt, Anja Böckmann, Rudi Glockshuber and Beat H. Meier.

Oxana prepared the samples and performed the seeding experiments. I measured 3D assignment spectra, derived the initial assignment, analyzed secondary chemical shifts and wrote the initial manuscript. Anne joined the project in early 2011 and is continuing the project with the aim of determining the structure and elucidating the binding-mode of congo-red. She measured the distance-restraint spectra and completed the assignment.

## 5.1. Introduction

The amyloid hypothesis states that aggregation of amyloid-$\beta$ (A$\beta$, a peptide of 39-43 residues length) and its accumulation in extracellular deposits in the brain are key events in Alzheimer's disease (AD), the most prevalent neurodegenerative disease in elderly people[111]. A$\beta$ is formed from proteolytic cleavage of the amyloid precursor protein (APP) by $\beta$- and $\gamma$-secretase. A$\beta$1-40 and A$\beta$1-42 are the main isoforms occurring in the human brain. A$\beta$1-42 is significantly less abundant than A$\beta$1-40[112] but is considered to be the pathologically relevant form due to its higher neurotoxicity and tendency to aggregate[113]. AD mostly occurs spontaneously with late onset. A small number of AD cases have a familial background, though, and are caused by mutations in the genes coding for APP or presenilin[114]. Recently the Osaka mutation E22$\Delta$ within the A$\beta$ sequence was found in Japanese AD patients[112]. A$\beta$1-40 E22$\Delta$ forms fibrils with a unique quaternary structure different from that of A$\beta$1-40 wt as seen by EM (see Figure 5.1). It has an increased tendency to form fibrillar bundles, a higher thioflavin T binding capacity and is significantly more toxic in rat primary neurons than A$\beta$1-40 wt[115], which might explain the early onset of AD caused by this mutation.

While A$\beta$ is generally known to form amyloid fibrils, determination its atomic-level structure remains a difficult task since the classical tools, X-ray crystallography and solution NMR
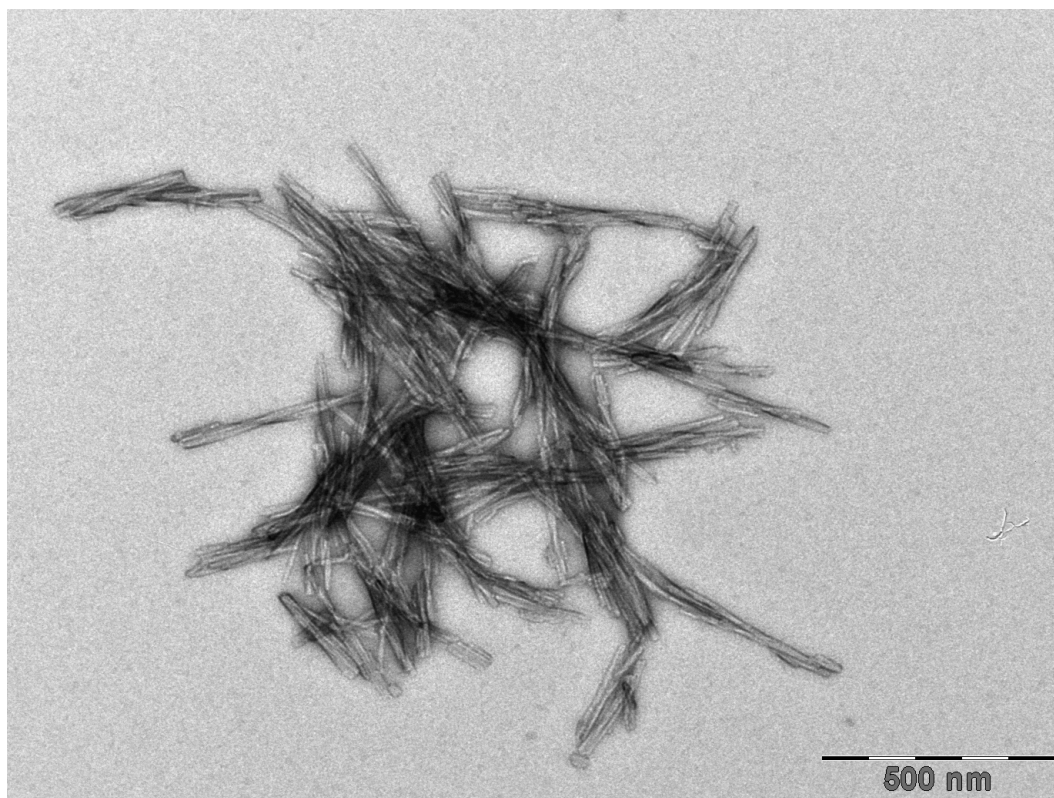
**Figure 5.1.** Negative-stain electron micrograph of fibrils formed from A$\beta$1-40 E22$\Delta$ .

spectroscopy, cannot provide sufficient information. Although numerous solid-state NMR studies on a multitude of polymorphs of A$\beta$ fibrils have been published to date (reviewed in[116, 117]) no atomic resolution structure has been obtained. A limiting step in the determination of the atomic resolution structure of A$\beta$ fibrils by solid-state NMR is the preparation of isotopically labeled A$\beta$ fibrils in sufficient quantity, and, even more critical, in sufficient quality and homogeneity to record high-resolution spectra. Most solid-state NMR studies on A$\beta$ have used selective isotope labeling which permits obtaining assignments and structural restraints from spectra of low resolution. In order to gain, using this method, information for all atoms, the preparation of a multitude of samples via chemical synthesis is required. Chemical synthesis typically introduces the problem of impurities and fragments that are difficult to remove and that may change the properties of A$\beta$[118]. In order to avoid these problems, we recombinantly expressed uniformly [$^{13}$C$^{15}$N]-isotope labeled fibrils, and present here the full sequential resonance assignment of A$\beta$1-40 E22$\Delta$ and compare the derived structural features to existing models. We have measured inter-monomer $\beta$-sheet contacts in fibrils of mixed [U-$^{13}$C] or [U-$^{15}$N] monomers and we show that the unique morphology of the A$\beta$1-40 E22$\Delta$

fibrils can also be induced in A$\beta$1-40 wt by seeding.

## 5.2. Results and Discussion

Fibrils of A$\beta$1-40 E22$\Delta$ yield 2D $^{13}$C-$^{13}$C and $^{15}$N-$^{13}$C spectra with narrow lines (line-width typically 0.5 ppm) reflecting a homogenous sample, as shown in Figure 5.2 and 5.3. Using 3D methods[32] (see also section 1.5) 95% of the backbone resonances could be sequentially assigned.
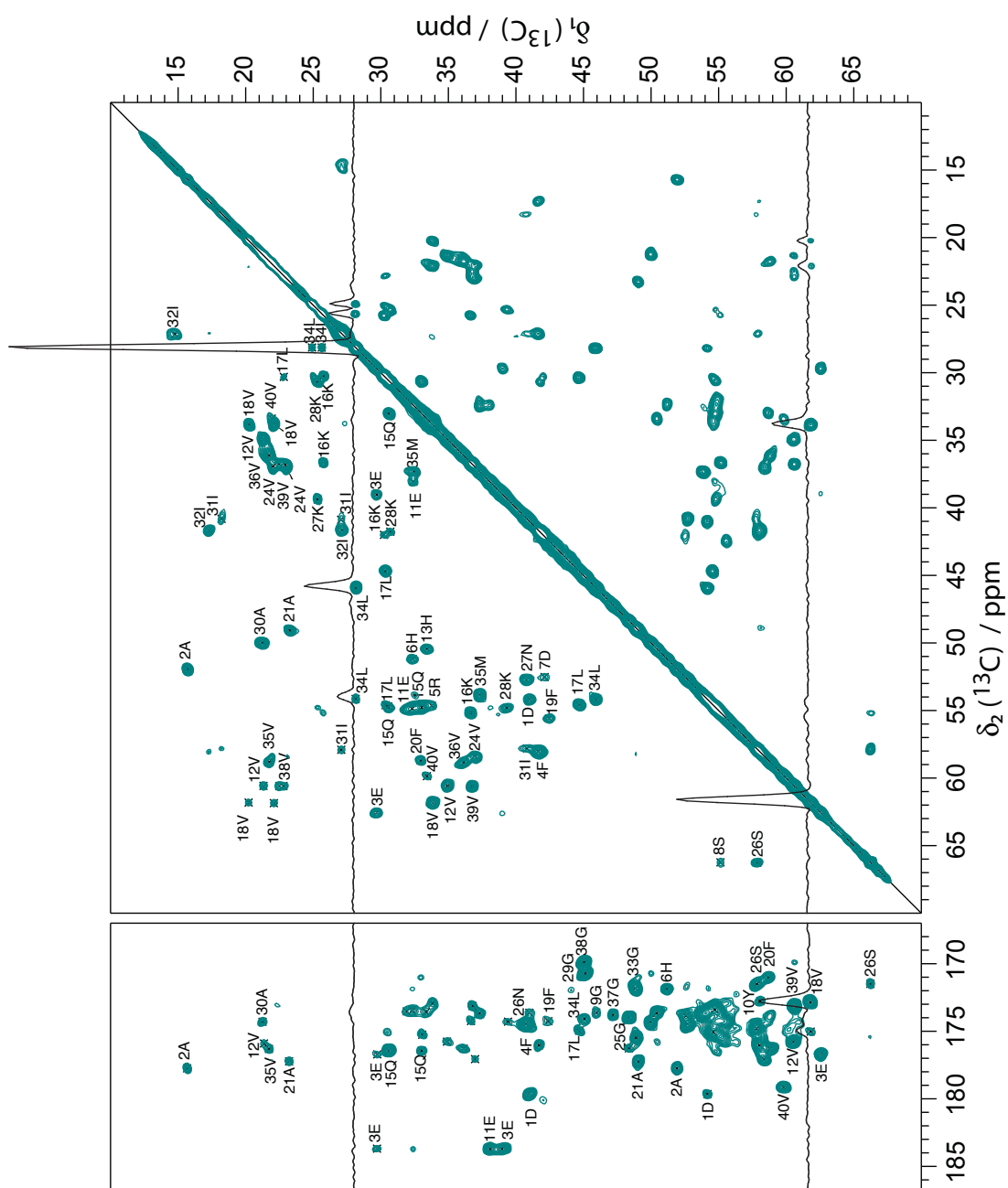
**Figure 5.2.** 2D $^{13}$C-$^{13}$C DARR spectrum of A$\beta$1-40 E22$\Delta$ fibrils with a mixing time of 15ms measured at a magnetic field of 20.0 T and magic-angle-spinning at a frequency of 19 kHz on [U-$^{13}$C] labeled fibrils. The lowest contour level plotted is at about 25 % of the intensity of cross-peaks corresponding to one bond $^{13}$C—$^{13}$C contacts.

**Figure 5.3.** NCA spectrum (20.0 T, 19 kHz MAS) of [U-¹³C¹⁵N] labeled Aβ1-40 E22Δ fibrils. The sequence specific assignments indicated were derived from a sequential backbone walk in 3D spectra as shown in Figure 2. Blue labels correspond to peaks from a second minor polymorph that is strongly decreased in the [U-¹³C] sample.
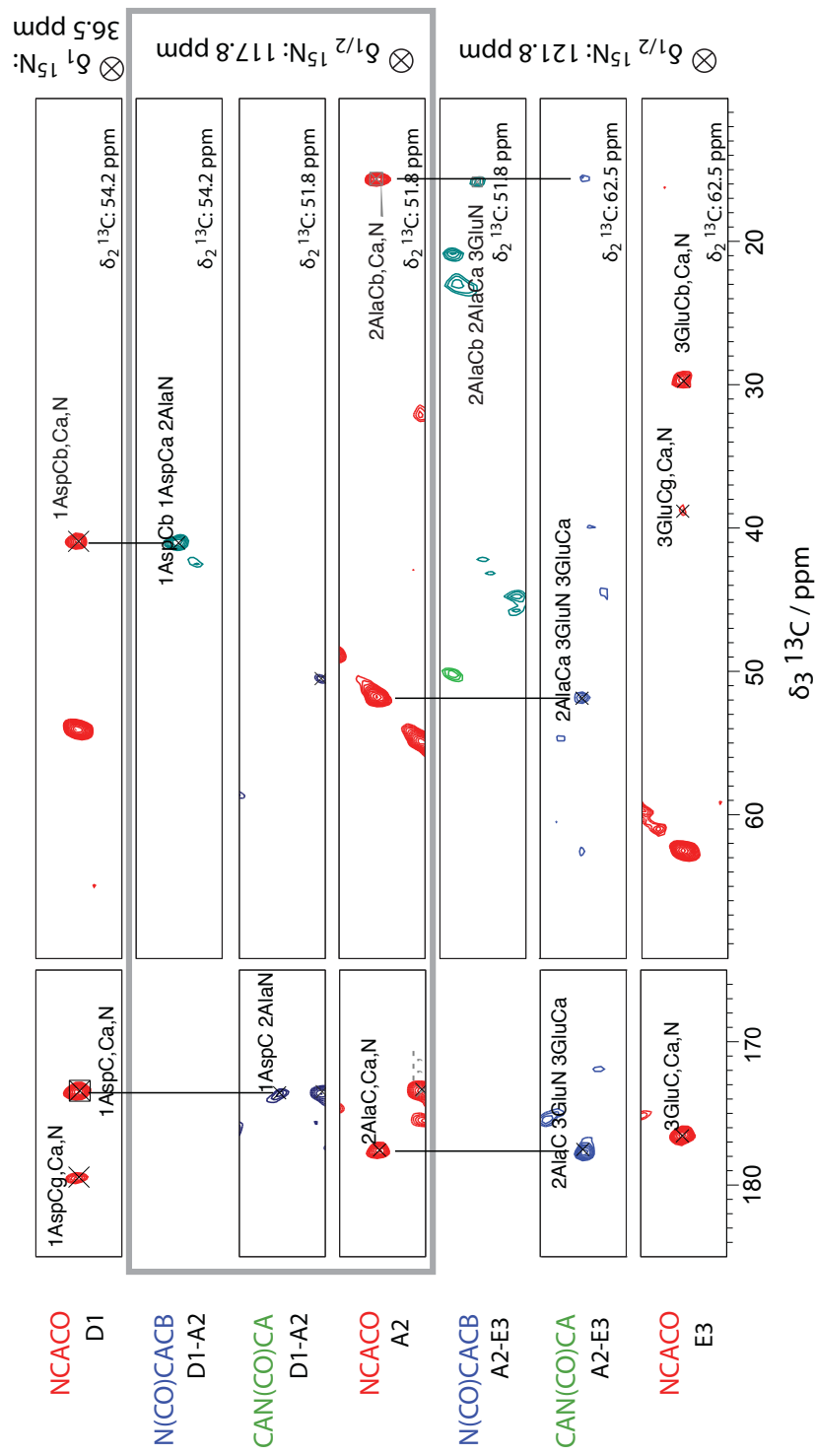
**Figure 5.4.** Strips extracted from the 3D NCACO, N(CO)CACB and CAN(CO)CA spectra of fibrillar Aβ1-40 E22Δ . Vertical lines were added to aid following the backbone walk starting at the N-terminus of the protein.

The remaining $^{13}$C resonances were assigned with a 2D [$^{13}$C;$^{13}$C] DARR spectrum with long mixing time (see Figure 5.5). Side chain $^{15}$N resonances were assigned from 2D [$^{15}$N;$^{13}$C] correlation spectra. Figure 5.6 shows the spectrum used for the assignment of histidine and arginine side chain nitrogens. The chemical shifts are shown in Table 5.1. From the chemical shifts information about the conformation of the peptide and also about the protonation state can be derived. The histidine N$\delta$ and N$\eta$ shifts that are around 175 ppm are characteristic of a doubly protonated and positively charged imidazolium ion [119]. The shift of N$\delta$ of His 13 is somewhat higher (192 ppm). If deprotonated $^{15}$N shifts around 245 ppm would be expected. The shift of His 13 N$\delta$ can be explained by assuming an H-bond to a carboxylate [119]. Dihedral angle restraints derived from secondary chemical shifts are discussed below.

All residues of the peptide are visible in spectra based on dipolar transfers, indicating that they are all rigid. In particular, also the N-terminal part of A$\beta$ that has been described as being flexible in wt-A$\beta$[120] yields strong, narrow signals. A representative stripplot of the sequential assignment is shown in Figure 5.4. For residues 11-14, 24-31 and 34, two or three sets of resonances are observed (see Figure 5.9 for details). The intensity of the second set of resonances with respect to the main resonances varies for the three different fibril preparations ([U-$^{13}$C$^{15}$N], [U-$^{13}$C] and mixed [U-$^{13}$C]/[U-$^{15}$N] ) studied between 60% and 20%. We were able to reduce the amount of polymorphic fibrils in the [U-$^{13}$C] rotor to 10% by removing the fibrils from the rotor, resuspending them in water and repacking the rotor (for an overlay of the DARR spectrum of the [U-$^{13}$C$^{15}$N] sample with the [U-$^{13}$C] sample after washing, where the polymorphic peaks are strongly decreased see Figure 5.7). This, combined with the fact that the resonances are sharp (ca 0.5 ppm linewidth) indicates that the fibril is well ordered and the packing of the monomers is highly symmetric. The second set of resonances must be from fibrils of a different morphology that have different sedimentation properties and cannot be due to structural inhomogeneity within the fibril.

The 400 ms DARR spectrum shown in Figure 5.5 could in principle also be used to extract distance restraints and not only to complete the assignment. The magnetization transfer within one residue and to sequentially connected residues is more efficient than the long range transfer (to residues separated by $>$ 4 amino acids in sequence). Therefore all of these peaks (that are useful for assignment) also occur and overlap with the long-range peaks. It is therefore difficult to include this spectrum in an automatic structure calculation. In regions of the spectrum without too much overlap long-range peaks can and should be manually assigned and included as manual distance restraints in structure calculations.

A possibility to reduce overlap due to many short range peaks is by decreasing the fraction

of $^{13}$C labeled sites, for example by 1-$^{13}$C or 2-$^{13}$C glucose labeling (see section 1.2.2). A PDSD spectrum of A$\beta$1-40 E22$\Delta$ fibrils prepared from an expression in 2-$^{13}$C glucose is shown in Figure 5.8. The line width is reduced because the one bond $^{13}$C–$^{13}$C J-couplings are not present in such a sample.

The secondary chemical shifts (Figure 5.9) and the secondary structure prediction with TALOS+[38] are consistent with most of the protein being in $\beta$-sheet conformation. Clear exceptions are residues Glu 3 and Gly 25, which show chemical shifts corresponding to turns. Gly 9, Phe 20, Asp 23, Gly 33 and Gly 37 have non $\beta$-sheet chemical shifts but the TALOS+ prediction of the secondary structure is inconclusive. The secondary chemical shifts of the minor polymorph of His 14 are indicative of a turn in contrast to His 14 of the major polymorph.
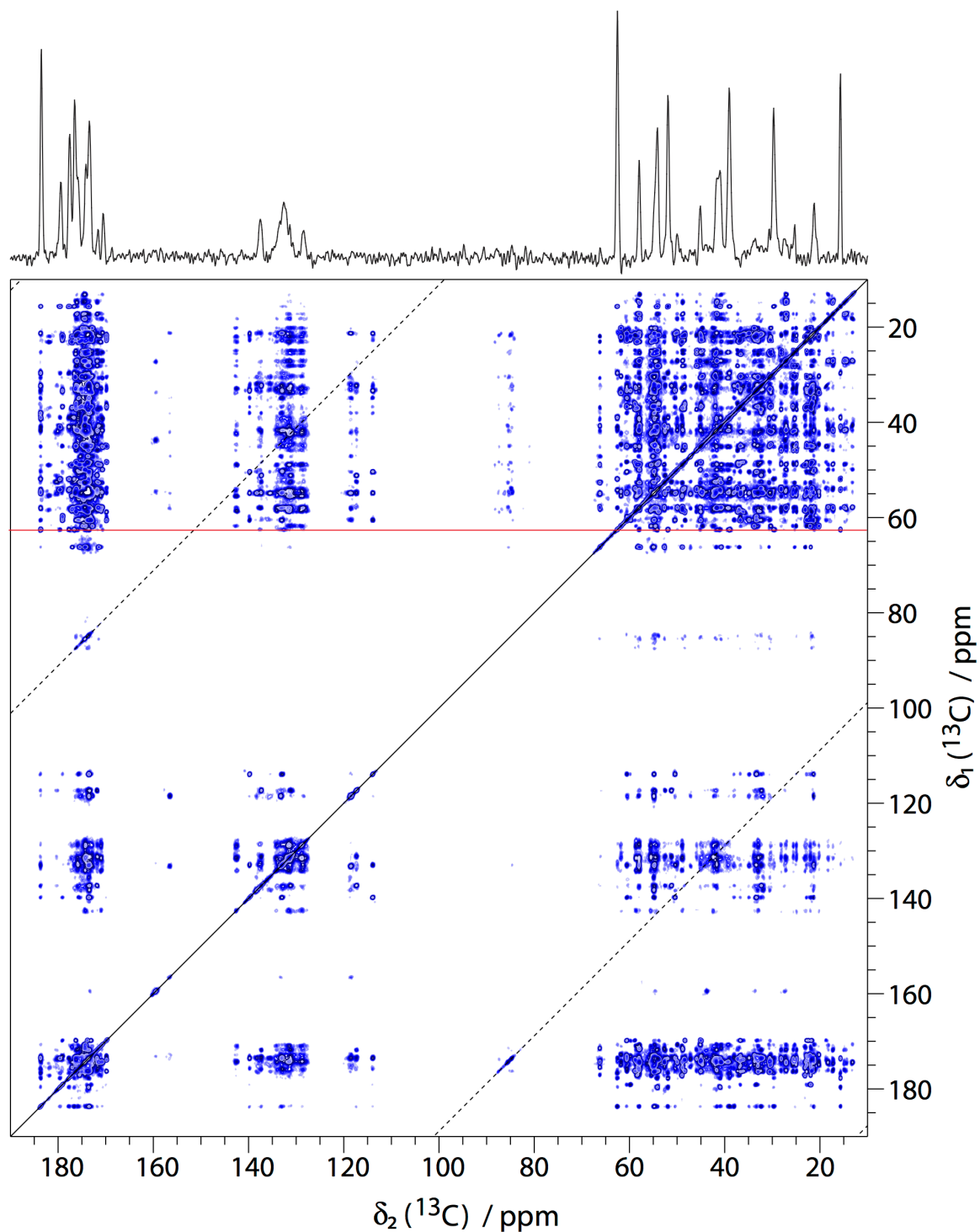
**Figure 5.5.** 2D [$^{13}$C;$^{13}$C] DARR spectrum of A$\beta$1-40 E22$\Delta$ fibrils with a mixing time of 400 ms measured at a magnetic field of 20.0 T and magic-angle-spinning at a frequency of 19 kHz on [U-$^{13}$C] labeled fibrils. The trace was extracted at the position of C$\alpha$ of glutamate 3 (indicated by the red line). The dashed-line indicates the position of spinning sidebands.
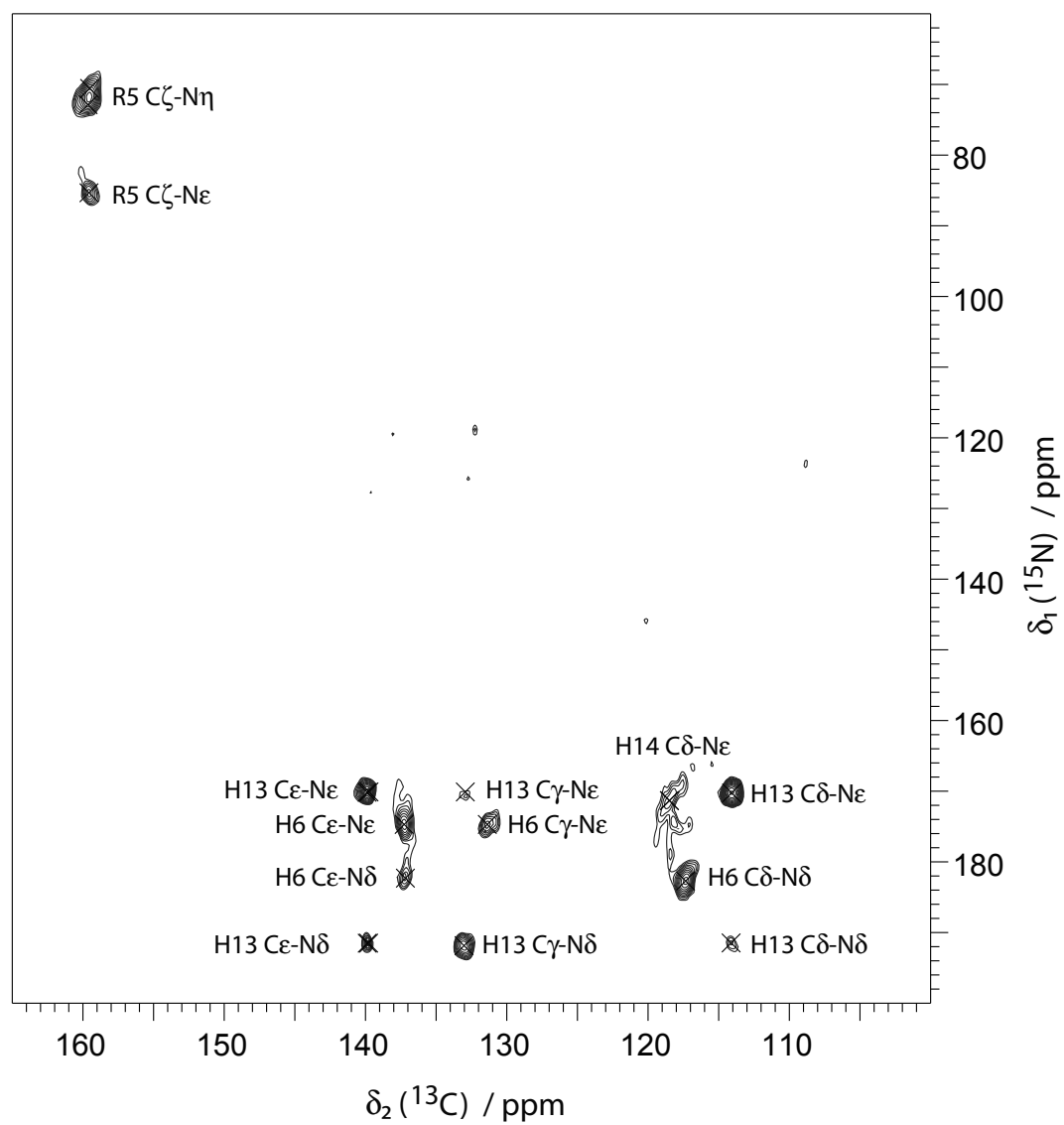
**Figure 5.6.** [$^{15}$N;$^{13}$C] correlation spectrum of arginine and histidine side chains.
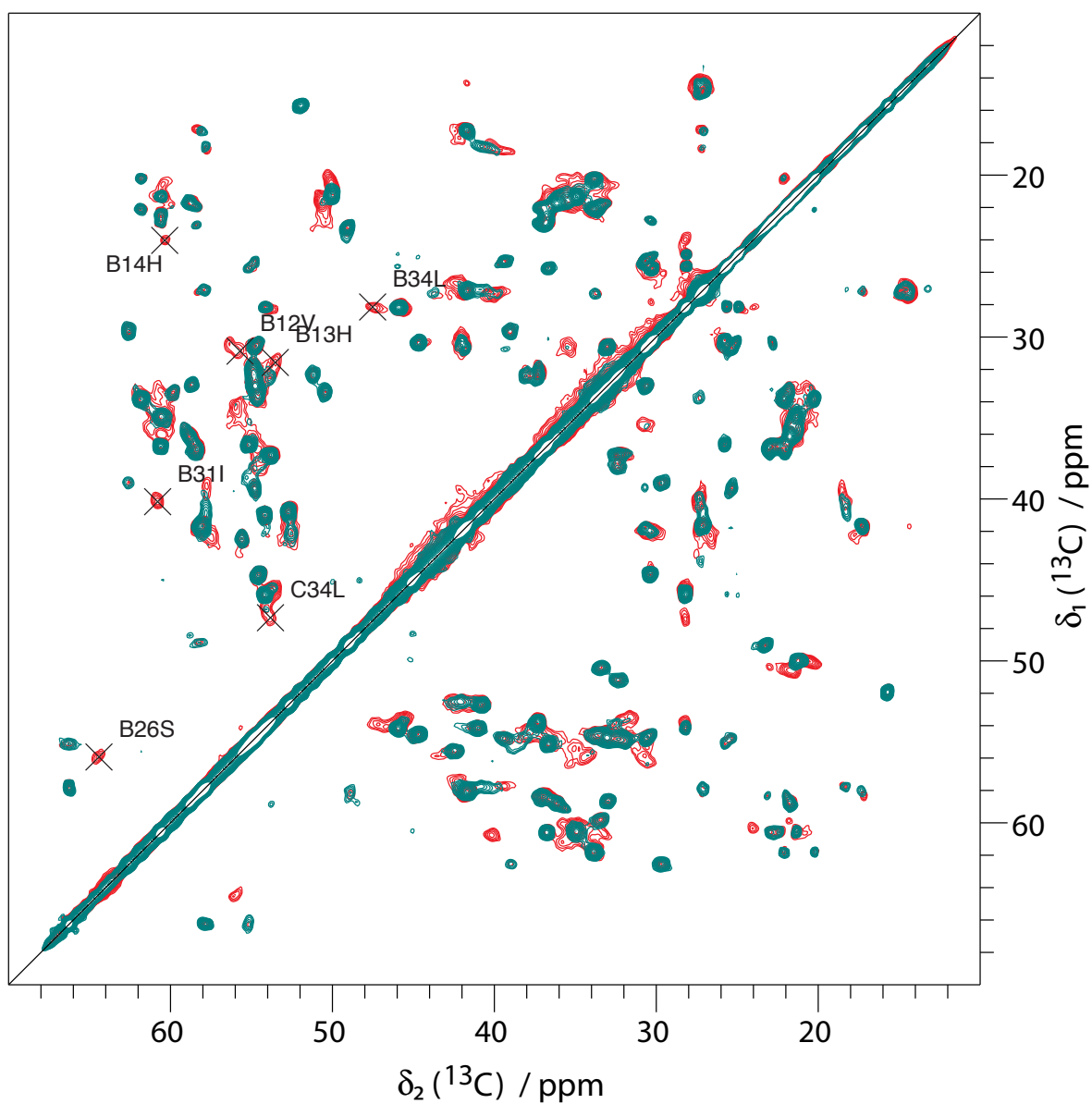
**Figure 5.7.** Overlay of 2D [$^{13}$C;$^{13}$C] DARR spectra (15 ms mixing, 19 kHz MAS, 20 T B$_0$) of Aβ1-40 E22Δ fibrils two independent preparations of the sample. Green: Sample 1, [U-$^{13}$C$^{15}$N]. Red: Sample 2, [U-$^{13}$C]. Polymorphic signals that are only observed in the sample 1 are labeled. The letters B and C in front of the residue names indicate polymorph B and C as shown in Figure 5.9
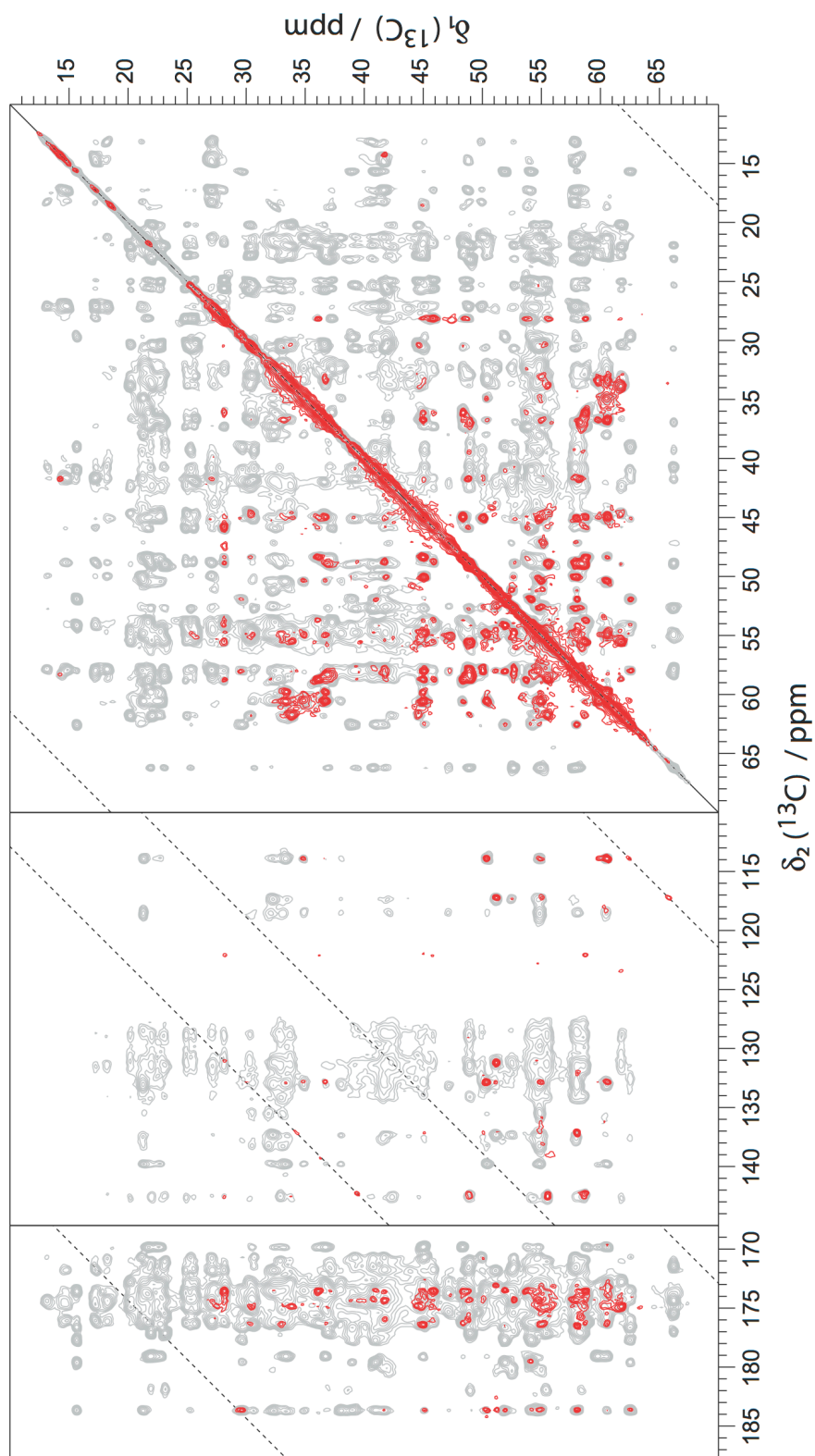
**Figure 5.8.** Comparison of 2D [$^{13}$C;$^{13}$C] spectra of Aβ1-40 E22Δ produced with $^{13}$C$_6$-glucose (grey, 400 ms DARR mixing, 19 kHz MAS) and 2-$^{13}$C glucose (red, 2 s PDSD mixing, 11 kHz MAS). Expression with 1-$^{13}$C glucose yielded a sample with a diluted network of $^{13}$C labeled sites. Magnetization transfer therefore happens mostly directly (is not relayed along chains of $^{13}$C atoms) and is therefore slower. The line-width is reduced since the $^{13}$C–$^{13}$C one-bond J-couplings are absent.

To determine the $\beta$-sheet contacts between A$\beta$1-40 E22$\Delta$ monomers in the fibril, we prepared fibrils from a 1:1 mixture of [U-$^{13}$C] and [U-$^{15}$N] monomers. Most peaks from the NCA and NCACB spectrum of [U-$^{13}$C$^{15}$N] are also present in the [$^{15}$N; $^{13}$C] PAIN[122] spectrum of the mixed fibril (Figure 5.10) indicating that nitrogen and carbon resonances of the same residue but from different monomers are in close spatial proximity as it is the case for parallel in-register $\beta$-sheets. Previously published solid-state NMR measurements on fibrils made from A$\beta$1-40 E22$\Delta$ peptide chemically synthesized to contain only one $^{13}$C label (either 1-$^{13}$C Val 18 or 1-$^{13}$C Val 36) are inconsistent with those labels forming chains with a spacing of 4.7 Å  as it would be expected for parallel in-register $\beta$-sheets[123]. We assume that the fibrils used for that study must be different than those used here. Cloe et al.[123] investigated fibrils that showed a decreased ThT fluorescence intensity compared to wt-fibrils, whereas we observed an increased ThT fluorescence intensity[115]. It is known for A$\beta$ wt that small differences in aggregation conditions or impurities can change the properties of the fibrils[124].

Comparison of the $\beta$-sheets regions of A$\beta$1-40 E22$\Delta$ with published models of A$\beta$ wt (Figure 5.11) shows that the secondary structure is clearly different. With the exception of the model proposed by Bertini et al. [131] the published models say that only the C-terminal part is rigid and forms $\beta$-sheets. Our secondary chemical shifts indicate that the site of the deletion mutation is in a turn. This in principle allows the peptide to adopt a similar fold as the wild-type A$\beta$and vice versa. A deletion within a $\beta$-strand would move all side chains of residues after the mutation from the outside to the inside of the fibril core and vice versa. In fact it is possible to seed fibrillization of A$\beta$ wt with mutant fibrils. For seeding experiments preformed fibrils (10% by mass) were added to the aggregation reaction of A$\beta$wt. The seeding capability of the preformed fibrils was enhanced by sonication. Six rounds of seeding were performed using A$\beta$1-40 E22$\Delta$ fibrils in the first cycle and the seeded A$\beta$wt fibrils of the previous cycle afterwards. A$\beta$1-40 E22$\Delta$ fibrils have a Thioflavin T fluorescence that is roughly four times stronger than that of A$\beta$wt fibrils. Seeded A$\beta$1-40 wt fibrils showed this increased ThT fluorescence and the morphology of A$\beta$1-40 E22$\Delta$ fibrils in negatively stained electron micrographs (Figure 5.12) after six seeding cycles.

| Residue | N | N$\delta$ | N$\eta$ | N$\zeta$ | C | C$\alpha$ | C$\beta$ | C$\gamma$ | C$\delta$ | C$\varepsilon$ | C$\zeta$ |
|---------|---|-----------|---------|----------|---|-----------|----------|-----------|-----------|----------------|----------|
| 1 Asp | 37.1 | - | - | - | 173.6 | 54.2 | 41.0 | 179.6 | - | - | - |
| 2 Ala | 117.8 | - | - | - | 177.7 | 51.9 | 15.7 | - | - | - | - |
| 3 Glu | 121.9 | - | - | - | 176.7 | 62.6 | 29.7 | 39.0 | 183.7 | - | - |
| 4 Phe | 114.6 | - | - | - | 176.0 | 58.0 | 41.7 | 137.8 | 133.3 | 131.2 | 128.6 |
| 5 Arg | 132.1 | - | 85.3 | - | 173.4 | 54.7 | 33.8 | 27.3 | 43.8 | - | 159.6 |
| 6 His | 125.8 | 174.6 | 182.4 | - | 171.9 | 51.2 | 32.3 | 131.3 | 117.4 | 137.4 | - |
| 7 Asp | 129.8 | - | - | - | 173.9 | 52.6 | 42.1 | 180.1 | - | - | - |
| 8 Ser | 116.5 | - | - | - | 175.4 | 55.1 | 66.2 | - | - | - | - |
| 9 Gly | 112.0 | - | - | - | 173.8 | 47.2 | - | - | - | - | - |
| 10 Tyr | - | - | - | - | 172.8 | 58.1 | 42.1 | 128.5 | 133.3 | 118.5 | 156.5 |
| 11 Glu | - | - | - | - | 173.4 | 54.9 | 32.4 | 38.1 | 183.7 | - | - |
| 12 Val | 126.1 | - | - | - | 175.8 | 60.6 | 34.9 | 21.3 | - | - | - |
| 13 His | 125.9 | 191.6 | 170.3 | - | 173.6 | 50.4 | 33.4 | 133.0 | 114.0 | 139.9 | - |
| 14 His | 118.1 | 179 | 170.8 | - | 173.5 | 55.0 | 31.9 | 135.6 | 118.6 | 133.2 | - |
| 15 Gln | 108.8 | - | 108.6 | - | 175.3 | 54.7 | 33.0 | 30.6 | 176.4 | - | - |
| 16 Lys | 120.0 | - | - | 33.7 | 174.2 | 55.2 | 36.7 | 25.7 | 30.2 | 42.0 | - |
| 17 Leu | 127.9 | - | - | - | 175.0 | 54.6 | 44.7 | 30.3 | 25.0, 22.8 | - | - |
| 18 Val | 123.4 | - | - | - | 172.9 | 61.8 | 33.8 | 22.1, 20.2 | - | - | - |
| 19 Phe | 129.4 | - | - | - | 174.3 | 55.6 | 42.4 | 142.8 | 130.6 | 131.5 | 128.9 |
| 20 Phe | 117.7 | - | - | - | 171.0 | 58.7 | 32.9 | 142.5 | 131.5 | 134.1 | 129.3 |
| 21 Ala | 114.9 | - | - | - | 177.2 | 49.1 | 23.2 | - | - | - | - |
| 23 Asp | 118.1 | - | - | - | 174.3 | 55.3 | 38.7 | 182.1 | - | - | - |
| 24 Val | 118.0 | - | - | - | 177.1 | 58.4 | 36.9 | 22.9, 22.0 | - | - | - |
| 25 Gly | 116.6 | - | - | - | 175.5 | 48.9 | - | - | - | - | - |
| 26 Ser | 113.8 | - | - | - | 171.5 | 57.8 | 66.2 | - | - | - | - |
| 27 Asn | 124.3 | 114.5 | - | - | 174.5 | 52.7 | 40.8 | 174.5 | - | - | - |
| 28 Lys | 119.8 | - | - | 34.1 | 174.3 | 54.8 | 39.4 | 25.3 | 30.7 | 41.8 | - |
| 29 Gly | 110.1 | - | - | - | 170.7 | 45.1 | - | - | - | - | - |
| 30 Ala | 131.0 | - | - | - | 174.3 | 50.0 | 21.2 | - | - | - | - |
| 31 Ile | 123.3 | - | - | - | 174.8 | 57.8 | 40.8 | 27.1, 18.2 | 13.4 | - | - |
| 32 Ile | 124.4 | - | - | - | 176.1 | 58.0 | 41.7 | 17.3, 27.1 | 14.6 | - | - |
| 33 Gly | 114.0 | - | - | - | 171.8 | 48.8 | - | - | - | - | - |
| 34 Leu | 125.3 | - | - | - | 173.6 | 54.2 | 45.9 | 28.1 | 24.9, 25.6 | - | - |
| 35 Met | 126.6 | - | - | - | 173.6 | 53.8 | 37.3 | 32.5 | - | 18.8 | - |
| 36 Val | 125.2 | - | - | - | 176.3 | 58.8 | 36.1 | 21.7 | - | - | - |
| 37 Gly | 117.0 | - | - | - | 174.0 | 48.4 | - | - | - | - | - |
| 38 Gly | 105.4 | - | - | - | 169.9 | 45.0 | - | - | - | - | - |
| 39 Val | 118.5 | - | - | - | 173.1 | 60.6 | 36.8 | 22.8, 22.5 | - | - | - |
| 40 Val | 127.8 | - | - | - | 179.2 | 59.8 | 33.4 | 21.9 | - | - | - |

**Table 5.1.** Chemical Shifts of A$\beta$1-40 E22$\Delta$ fibrils (major polymorph) in ppm
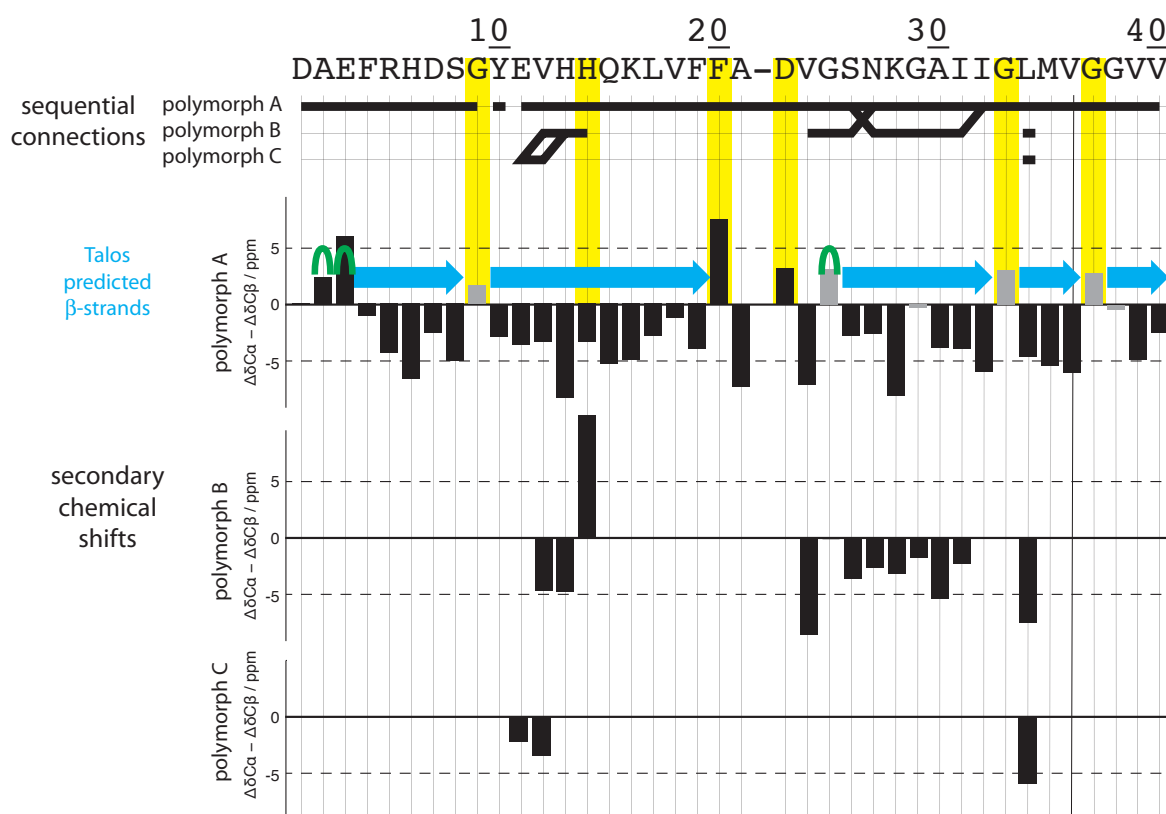
**Figure 5.9.** Observed sequential contacts in 3D spectra of labeled Aβ1-40 E22Δ fibrils and secondary chemical shifts[121]. For glycine $\Delta\delta C_\alpha$ is plotted and the corresponding residues bars are plotted in gray. Three negative secondary shifts in a row are indicative for β-sheet structure. Blue arrows are drawn for TALOS+ predicted β-sheet regions, green loops are drawn for residues predicted to not be involved in β-sheets and residues were TALOS+ gave out a warning are marked in yellow. The intensity of signal from polymorphs B and C varies for the different samples and is largest for the [U-$^{13}$C$^{15}$N] rotor (approx. 60% that of polymorph A) and lowest in the [U-$^{13}$C] rotor (approx. 10% of the intensity of polymorph A).
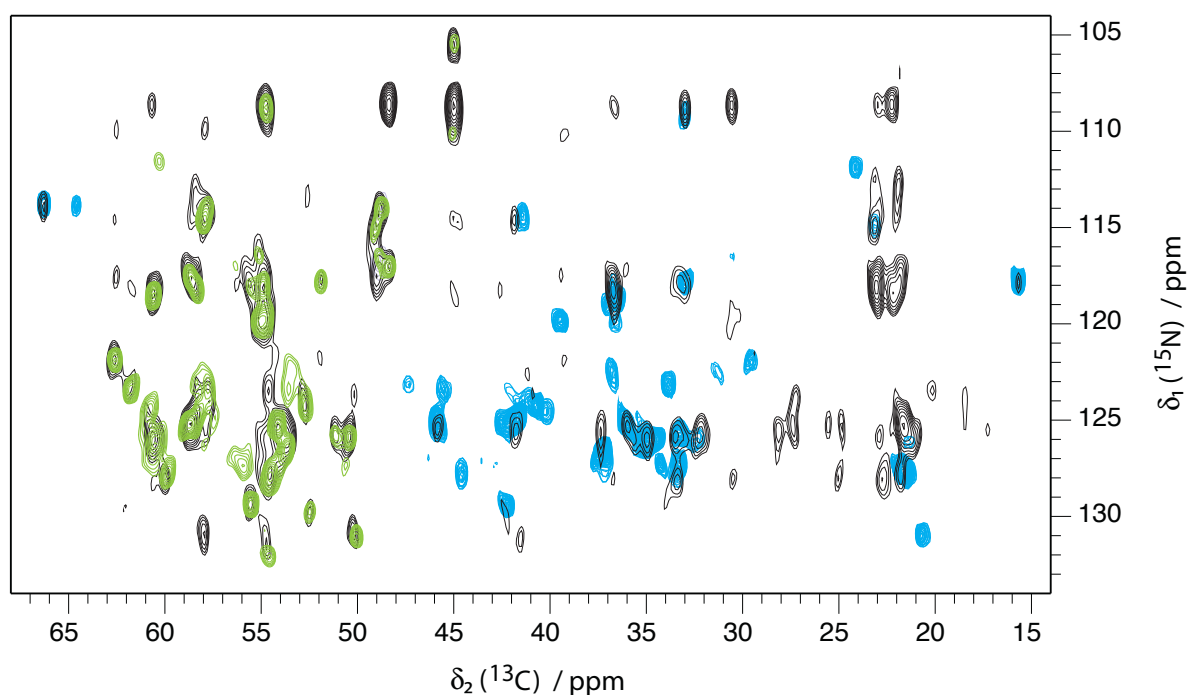
**Figure 5.10.** Overlay of long-range [$^{15}$N;$^{13}$C] correlation spectrum of A$\beta$1-40 E22$\Delta$ fibrils of a 1:1 mixture of [U-$^{13}$C] and [U-$^{15}$N] monomers (black) with an NCA (green) spectrum and the projection onto the NCB plane of the 3D NCACB spectrum (blue) of [U-$^{13}$C$^{15}$N] labeled fibrils. $^{15}$N–$^{13}$C contacts that can be observed both within a monomer (blue and green spectra) and between two different monomers (black spectrum) lead to identical peak positions and are indicative of parallel-in-register $\beta$-sheets.

**Location of β-strands**                                                    **Methodology**
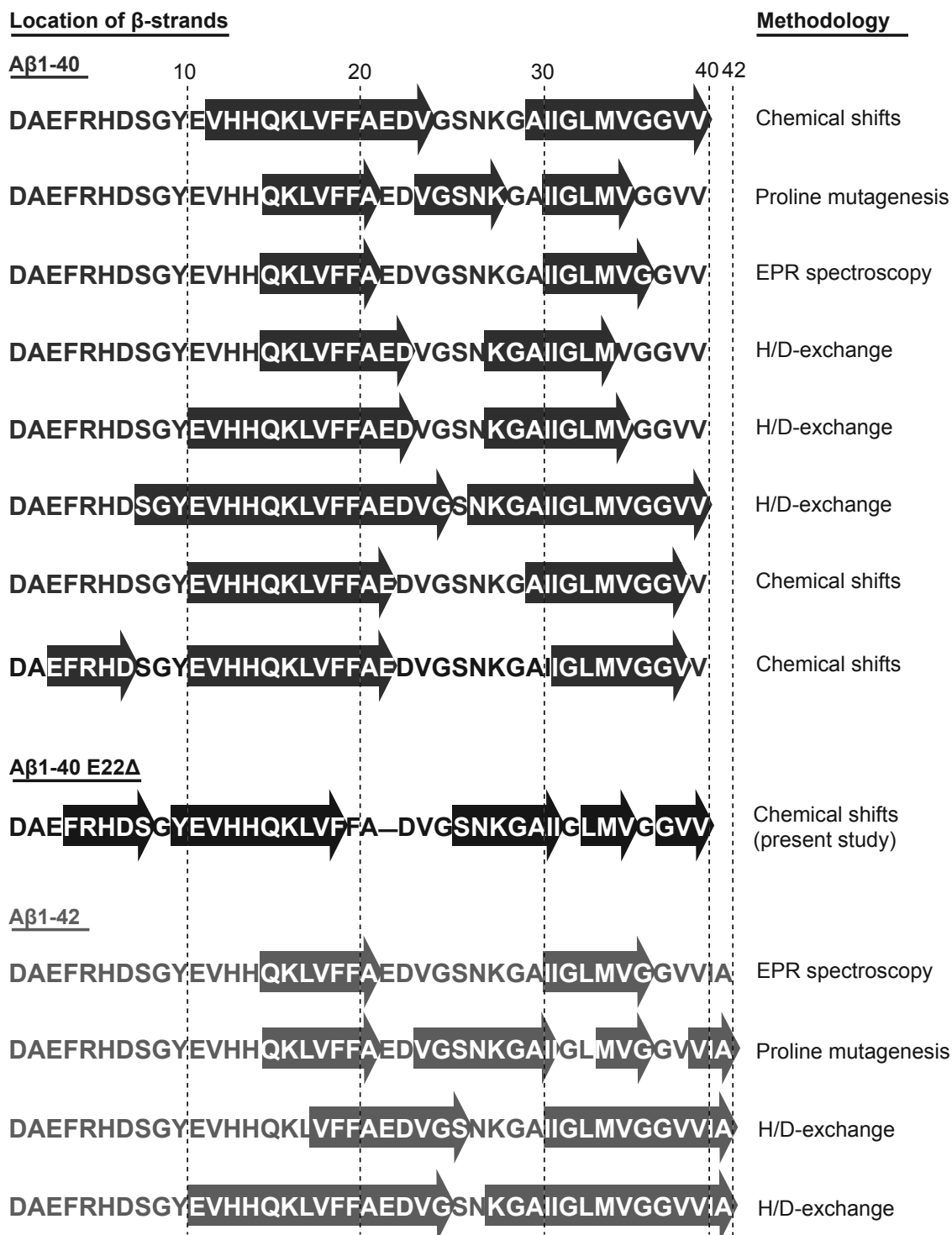


**Figure 5.11.** $\beta$-sheet regions in A$\beta$ wt compared to those derived for A$\beta$1-40 E22$\Delta$ in the present study. The structure and morphology of A$\beta$ wt fibrils is know to depend strongly on the details of the aggregation condition which leads to the differences between studies. Note however, that none of A$\beta$ wt measurements predicts the N-terminal part to be well structured and in a $\beta$-sheet region except for Bertinis model for A$\beta$1-40 wt and the present work on A$\beta$1-40 E22$\Delta$ . References for A$\beta$1-40 are: [53, 125–131], those for A$\beta$1-42 are: [126, 132–134].
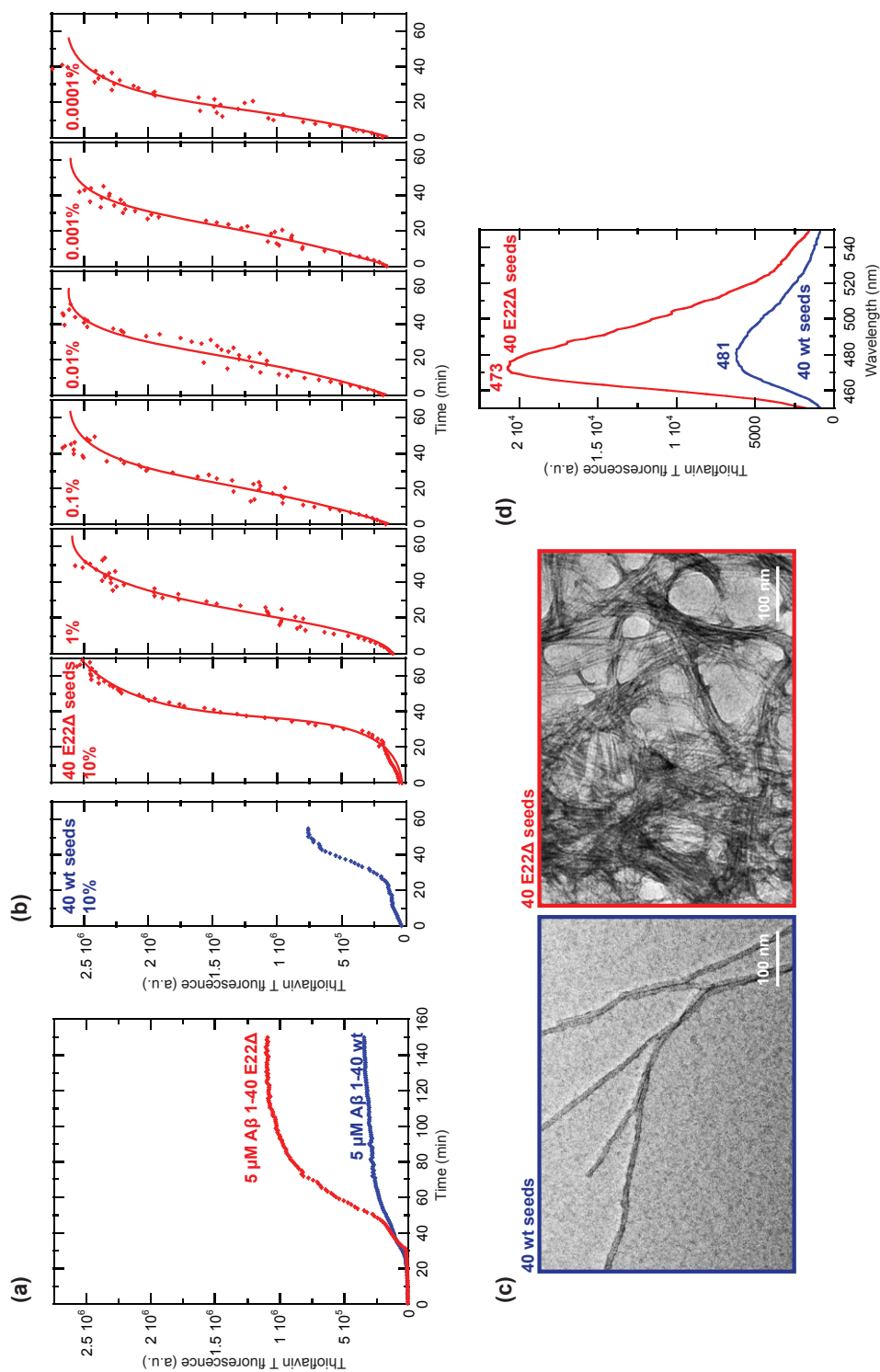
**Figure 5.12.** (a) Kinetics of in vitro fibril formation of Aβ1-40 wt (blue) and Aβ1-40 E22Δ (red), and (b) of Aβ1-40 wt 'seeds' (blue) and in presence of Aβ1-40 E22Δ 'seeds' (red). Mixed fibrils were sonicated to form seeds and used to seed more Aβ1-40 wt. (c) Negative-stain electron micrographs of Aβ1-40 wt fibrils formed in presence of its own 'seeds' (blue frame) and in presence of Aβ1-40 E22Δ 'seeds' after six seeding cycles (red frame). The seeded wt fibrils have the same morphology as Aβ1-40 E22Δ fibrils (see Figure. 5.1) (d) Emission spectra of thioflavin T ($\lambda$ex = 440 nm), bound to Aβ1-40 wt fibrils formed in presence of its own 'seeds' (blue) and in presence of Aβ1-40 E22Δ 'seeds' after six seeding cycles (red). (figure from Oxana Ovchinnikova, for details see [135])

## 5.3. Conclusions

We have described the sequential resonance assignment of A$\beta$1-40 E22$\Delta$ fibrils. Using 3D based assignment methods in combination with 2D [$^{13}$C;$^{13}$C] and 2D [$^{15}$N;$^{13}$C] spectra we achieved complete $^{13}$C and $^{15}$N assignment. In contrast to proposed structural models for A$\beta$-wt also the N-terminal residues are rigid and well structured. Intermonomer $^{15}$N-$^{13}$C contacts measured in mixed [$^{15}$N]/[$^{13}$C] labeled fibrils are consistent with a parallel in-register $\beta$-sheet arrangement. A$\beta$1-40 E22$\Delta$ cofibrillizes with A$\beta$1-40 wt and imposes its unique morphology on the wild-type peptide for several generations.

## 5.4. Experimental Methods

### 5.4.1. Preparation of [U-$^{13}$C,$^{15}$N]-A$\beta$1-40 E22$\Delta$ fibrils

The plasmid for expression in *E.coli* of A$\beta$1-40 E22$\Delta$ has been described previously [115]. The expression was performed in *E. coli* BL21 (DE3) under T7 promoter control in M9 minimal medium containing 100 $\mu$M CaCl$_2$, 1 $\mu$M FeCl$_3$, 10 $\mu$M ZnCl$_2$, 1 mM MgSO$_4$, $^{13}$C-Glucose (2g/L) and $^{15}$N-NH$_4$Cl (1g/L), and cells were induced at OD600 $\approx$ 1.0 for 10 hours at 37°C with 1 mM IPTG. A$\beta$1-40 E22$\Delta$ was produced as a fusion to the peptide sequence (NANP)$_{19}$ with N-terminal hexahistidine tag and subsequently cleaved of the fusion protein with TEV protease as described earlier[115]. The high purity and identity of the peptide was verified by MALDI-TOF mass spectrometry using sinapinic acid as matrix (see Figure 8.14 in the appendix). A$\beta$1-40 E22$\Delta$ was dissolved in 10 mM NaOH to concentration of 150-200 $\mu$M and subjected to ultracentrifugation for 1 h at 135,500 g and 4 °C. The stock solution was kept on ice and used for preparation of the fibrils within 24 h. The aggregation reaction was performed at 37 °C and 700 revolutions per minute (rpm) with 35-50 $\mu$M A$\beta$ (final concentration) in 10 mM H$_3$PO$_4$-NaOH pH 7.4 and 100 mM NaCl. The aggregation reaction was started by dilution of the A$\beta$ stock solution in 10 mM NaOH with an aggregation buffer mix resulting in pH 7.4 and the final concentrations indicated above.

## 5.4.2.  Preparation of fibrils from mixed [U-$^{13}$C] or [U-$^{15}$N] labeled monomers

The monomers were produced as described above with the exception that natural abundance glucose or ammonium chloride was used.  Stock solutions of the monomers were combined before aggregation.

## 5.4.3.  Sample preparation for NMR

The fibrils were centrifuged for 45 min at 29,400 g and 4°C, washed with MilliQ water and centrifuged again as described.  The pellet containing approximately 12 mg A$\beta$1-40 E22$\Delta$ fibrils was resuspended in ca 400 $\mu$l MilliQ water and packed into a 3.2 mm ZrO$_2$ rotor (Bruker) using ultracentrifugation and a home made filling device[103], in a SW41-TI swing-out rotor spinning at 100'000 g for 16 hours in an optima L90-K ultracentrifuge (Beckmann). The drive tip of the rotor was sealed with epoxy glue to prevent dehydration of the sample.

## 5.4.4.  NMR Spectroscopy

We used two sets of 3D experiments for backbone assignments, namely NCOCA, NCACO and CANCO as well as NCACB, N(CO)CACB and CAN(CO)CA as described previously[32] and completed the side-chain assignment with a 3D CCC and 2D DARR spectra as well as 2D NC spectra with carrier positions optimized to observe $^{15}$N–$^{13}$C correlations in histidine and arginine side-chains (Figure 5.6).  Spectra were recorded on a Bruker Avance II+ spectrometer operating at 850 MHz $^1$H Larmor frequency with magic-angle-spinning (MAS) at 19 kHz or on a Bruker Avance spectrometer operating at 600 MHz $^1$H Larmor frequency at MAS frequency of 13 kHz. Spectra were processed with Topspin (Bruker) and analyzed with CCPN[136]. Detailed experimental conditions are given in Table 5.2.

## 5.4.5.  Seeding experiments

The expression of A$\beta$1-40 wt and A$\beta$1-40 E22$\Delta$ was performed in E. coli BL21 (DE3) under T7 promoter control in TB medium as described before[115]. The high purity and identity of the peptides was verified by MALDI-TOF mass spectrometry using sinapinic acid as matrix (data not shown).  The stock solutions of A$\beta$1-40 wt and A$\beta$1-40 E22$\Delta$ were prepared by

| Experiment | NCACB | CANCoCX | NCOCX | CCC | NCACO | NcoCACB | NCA | DARR 15 ms |
|---|---|---|---|---|---|---|---|---|
| $^1$H Larmor frequency | 600 MHz | 850 MHz | 600 MHz | 600 MHz | 600 MHz | 600 MHz | 850 MHz | 850 MHz |
| Isotope Labeling | [U-$^{13}$C$^{15}$N] | [U-$^{13}$C$^{15}$N] | [U-$^{13}$C$^{15}$N] | [U-$^{13}$C$^{15}$N] | [U-$^{13}$C$^{15}$N] | [U-$^{13}$C$^{15}$N] | [U-$^{13}$C$^{15}$N] | [U-$^{13}$C] |
| MAS frequency [kHz] | 13 | 19 | 13 | 13 | 13 | 13 | 19 | 19 |
| Transfer 1 | HN-CP | HC-CP | HN-CP | HC-CP | HN-CP | HN-CP | HN-CP | HC-CP |
| field [kHz] -$^1$H | 71.1 | 72.8 | 81.6 | 130.4 | 70.7 | 74.1 | 62.3 | 75.8 |
| field [kHz] -X | 59.0 | 61.6 | 54.4 | 119.6 | 57.6 | 46.8 | 62.5 | 63.0 |
| shape | tangent | tangent | tangent | tangent | tangent | tangent | tangent | tangent |
| carrier [ppm] | - | CA | - | - | - | - | - | 100 |
| time [ms] | 1 | 0.4 | 1.2 | 0.5 | 1 | 1 | 1.2 | 1 |
| Transfer 2 | NC-CP | CN-CP | NC-CP | DREAM | NC-CP | NC-CP | NC-CP | DARR |
| field [kHz] -$^1$H | 94.4 | 93.3 | 109.7 | 97.7 | 79.4 | 83.1 | 88.1 | 18.7 |
| field [kHz] -$^{13}$C | 4.1 | 6.2 | 4.3 | 15.7 | 4.3 | 3.9 | 10.8 | - |
| field [kHz] - $^{15}$N | 17.7 | 13.0 | 16.7 | - | 18.6 | 16.6 | 24.0 | - |
| shape | tangent | tangent | tangent | tangent | tangent | tangent | tangent | - |
| carrier [ppm] | CA | CA | CO | 55 | CA | CO | CA | - |
| time [ms] | 6 | 5.5 | 4 | 4 | 5.5 | 7 | 6 | 15 |
| Transfer 3 | DREAM | NC-CP | DARR | DARR | DARR | MIRROR | - | - |
| field [kHz] -$^1$H | 94.4 | 83.1 | 12.9 | 12.8 | 13.4 | 17.9 | - | - |
| field [kHz] -$^{13}$C | 6.0 | 6.2 | - | - | - | - | - | - |
| field [kHz] - $^{15}$N | - | 13.4 | - | - | - | - | - | - |
| shape | tangent | tangent | - | - | - | - | - | - |
| carrier [ppm] | 59 | CO | - | - | - | - | - | - |
| time [ms] | 3 | 5.5 | 50 | 50 | 10 | 10 | - | - |
| Transfer 4 | - | MIRROR | - | - | - | DREAM | - | - |
| field [kHz] -$^1$H | - | 15.9 | - | - | - | 91.2 | - | - |
| field [kHz] -$^{13}$C | - | - | - | - | - | 6.0 | - | - |
| shape | - | - | - | - | - | tangent | - | - |
| carrier [ppm] | - | - | - | - | - | 59 | - | - |
| time [ms] | - | 30 | - | - | - | 3 | - | - |
| $t_1$ increments | 36 | 128 | 54 | 172 | 52 | 48 | 1024 | 2560 |
| sweep width ($t_1$) [kHz] | 3.35 | 11.7 | 4 | 20 | 3.65 | 4 | 40 | 100 |
| max. acq time ($t_1$) [ms] | 5.3 | 5.4 | 6.8 | 5.4 | 7.1 | 6 | 12.8 | 12.80 |
| $t_2$ increments | 90 | 64 | 96 | 172 | 108 | 100 | 1280 | 3072 |
| sweep width ($t_2$) [kHz] | 8.3 | 4.3 | 6.8 | 20 | 8.3 | 9 | 83.3 | 100 |
| max. acq time ($t_2$) [ms] | 5.4 | 7.3 | 7.1 | 5.4 | 6.5 | 5.5 | 15.4 | 15.36 |
| $t_3$ increments | 1280 | 1536 | 1280 | 1280 | 1280 | 1280 | - | - |
| sweep width ($t_3$) [kHz] | 50 | 100 | 50 | 50 | 50 | 100 | - | - |
| max. acq time ($t_3$) [ms] | 12.9 | 7.8 | 12.9 | 12.9 | 12.8 | 12.8 | - | - |
| $^1$H Spinal-64 Dec. [kHz] | 90 | 88 | 91 | 91 | 85 | 86 | 91 | 100 |
| interscan delay [s] | 1.8 | 2.6 | 2 | 1.9 | 1.4 | 1.6 | 2 | 2.5 |
| number of scans | 8 | 16 | 16 | 4 | 32 | 32 | 8 | 2 |
| measurem. time [h] | 13.1 | 95.5 | 47.7 | 64.6 | 71.4 | 69.6 | 4.6 | 3.6 |

**Table 5.2.** Experimental parameters for the NMR spectra. All field values in kHz calculated from reference fields (100 kHz or 62.5 kHz), the actual values may deviate significantly from calculated ones.

| Experiment | DARR 400ms | PAR 8ms | PARmix 8ms | gluc 2s PDSD | PAINmix | NC aro | NC Arg |
|---|---|---|---|---|---|---|---|
| $^1$H Larmor frequency | 850 MHz | 850 MHz | 850 MHz | 850 MHz | 850 MHz | 600 MHz | 600 MHz |
| Isotope Labeling | [U-$^{13}$C] | [U-$^{13}$C] | $^{13}$C/$^{15}$N mix | 2 $^{13}$C gluc | $^{13}$C/$^{15}$N mix | [U-$^{13}$C$^{15}$N] | [U-$^{13}$C$^{15}$N] |
| MAS frequency [kHz] | 19 | 19 | 19 | 19 | 19 | 13 | 13 |
| Transfer 1 | HC-CP | HC-CP | HC-CP | HC-CP | HN-CP | HN-CP | HN-CP |
| field [kHz] -$^1$H | 75.8 | 75.8 | 76.2 | 71.5 | 78.5 | 75.8 | 75.8 |
| field [kHz] -X | 63.0 | 63.0 | 64.1 | 61.6 | 50.0 | 63.6 | 63.6 |
| shape | tangent | tangent | tangent | tangent | tangent | tangent | tangent |
| carrier [ppm] | 100 | 100 | 100 | 100 | | | |
| time [ms] | 1 | 0.8 | 0.8 | 1 | 1.4 | 1.1 | 1.1 |
| Transfer 2 | DARR | PAR | PAR | PDSD | PAIN | NC | NC |
| field [kHz] -$^1$H | 18.7 | 63.7 | 63.7 | - | 67.5 | 100.0 | 100.0 |
| field [kHz] -$^{13}$C | - | 65.2 | 66.4 | - | 44.1 | 4.0 | 4.0 |
| field [kHz] -$^{15}$N | - | - | - | - | 42.1 | 16.7 | 16.7 |
| shape | - | - | - | - | tangent | tangent | tangent |
| carrier [ppm] | - | CA | CA | - | CA | 120ppm | 155ppm |
| time [ms] | 400 | 8 | 8 | 2000 | 6 | 7 | 4.6 |
| $t_1$ increments | 2048 | 2048 | 2048 | 2560 | 384 | 640 | 640 |
| sweep width ($t_1$) [kHz] | 100 | 100 | 100 | 100 | 25 | 40 | 40 |
| max. acq time ($t_1$) [ms] | 10.24 | 10.24 | 10.24 | 12.8 | 7.68 | 8 | 8 |
| $t_2$ increments | 3072 | 3072 | 3072 | 3584 | 2560 | 1536 | 1536 |
| sweep width ($t_2$) [kHz] | 100 | 100 | 100 | 100 | 50 | 50 | 50 |
| max. acq time ($t_2$) [ms] | 15.36 | 15.36 | 15.36 | 17.92 | 25.6 | 15.36 | 15.36 |
| $^1$H Spinal-64 Dec. [kHz] | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| interscan delay [s] | 2.3 | 2.5 | 2.5 | 1.3 | 2.2 | 2 | 2 |
| number of scans | 16 | 40 | 64 | 32 | 96 | 48 | 64 |
| measurem. time [h] | 24.7 | 57.4 | 91.8 | 75.5 | 22.8 | 17.2 | 23.0 |

**Table 5.3.** Experimental parameters for the NMR spectra. All field values in kHz calculated from reference fields (100 kHz or 62.5 kHz), the actual values may deviate significantly from calculated ones.

dissolving the peptides in 10 mM NaOH to concentration of 80-100 $\mu$M followed by ultra-centrifugation for 1 h at 135,500 g and 4 °C. The A$\beta$ concentration in the supernatant was determined via the absorbance of A$\beta$ at 280 nm in 10 mM NaOH (extinction coefficient at 280 nm: 1730 $M^{-1}$ $cm^{-1}$ (according to ref[118])). The stock solutions were kept on ice and used for preparation of the fibrils within 24 h. The fibrils of A$\beta$1-40 wt and A$\beta$1-40 E22$\Delta$ were formed in a volume of 150 $\mu$l in a black 96-well plate with a clear bottom (Greiner Bio-One GmbH, Germany) sealed with a VIEWseal foil to prevent evaporation (Greiner Bio-One, Germany) in a microtiter plate reader (BioTek instruments GmbH, Germany) at the following conditions: 5 $\mu$M A$\beta$, 10 mM $H_3PO_4$-NaOH pH 7.4, 100 mM NaCl, 35 $\mu$M thioflavin T, 37 °C, fast shaking. The aggregation reactions were started by dilution of the A$\beta$ stock solution in 10 mM NaOH with an aggregation buffer mix resulting in pH 7.4 and the final concentrations indicated above. The aggregation was monitored using the excitation filter 440/30 nm and the emission filter 485/20 nm. All aggregation curves represent the average of two experiments. The preformed fibrils (100 $\mu$l) were sonicated for 2 minutes in a cold water in a water bath based sonicator Sonorex Digital 10P (Bandelin Electronic GmbH, Germany) at the highest power settings. The seeding reactions were performed with 10% of freshly prepared and kept on ice 'seeds' and 90% of A$\beta$ monomer with the final A$\beta$ concentration of 5 $\mu$M. For aggregation of A$\beta$1-40 wt in presence of A$\beta$1-40 E22$\Delta$ sonicated fibrils, seeding reaction was repeated 6 times, each time with freshly formed fibrils, therefore diluting the initial A$\beta$1-40 E22$\Delta$ 'seeds' in the reaction mixture. Thioflavin T fluorescence spectra were recorded for A$\beta$1-40 wt fibrils formed in presence of A$\beta$1-40 wt 'seeds' and in presence of A$\beta$1-40 E22$\Delta$ 'seeds' (after the sixth cycle of aggregation) on a Quantamaster (QM-7/2003) fluorescence spectrometer (Photon Technology International) in the range of 450-550 nm (excitation at 440 nm). The same samples of fibrils were analyzed by electron microscopy. A$\beta$ fibrils were applied to carbon-coated copper grids (Quantifoil Micro Tools GmbH) for 1 min, washed with water and stained with 2% uranyl acetate for 30 s. Electron micrographs were recorded at 100 MeV on a FEI Morgagni 268 transmission electron microscope.

# 6. Structure Calculation of Symmetric Fibrils

This section describes work still in progress that is done in collaboration with Anne Schütz and Peter Güntert. Anne is recording and analyzing spectra for distance restraints and Peter implements fibril symmetry restraints in CYANA. All of us are performing structure calculations.

## 6.1. Introduction

Amyloid fibrils contain extended $\beta$-sheets perpendicular to the fibril axis that contain $\beta$-strands from different protein molecules[137]. In NMR based structure determination of globular proteins the peaks from spatial correlation spectra are mostly due to contacts between different parts of the primary sequence of one protein molecule. In contrast, intermolecular interactions play an important role for fibrils. If the fibrils yield spectra with only one narrow set of resonances the local environment of equivalent atoms in different molecules must be very similar. This is possible if the monomers are arranged in a symmetric fashion.

Fibrils can be quasi-one-dimensional systems where a single molecule is periodically repeated along the fibril axis[52] or several molecules can be in the transverse plane. In the direction of the fibril-axis a network of hydrogen bonds connects different $\beta$-strands to the extended $\beta$-sheet. Laterally the fibril is stabilized by side-chain interactions. Both axial and lateral contacts can be intra- or intermolecular.

In the following section we discuss how the different types of contacts can be observed and categorized and how the fibril structure can be calculated with CYANA[70]. We finally use this knowledge to calculate possible structures of A$\beta$1-40 E22$\Delta$ fibrils that are consistent we our experimental data.

## 6.2. Implementation in Cyana

The structure of a fibril consisting of symmetrically arranged monomers can be fully described by the structure of one monomer and the symmetry operation that allows to determine the positions of the neighbors. CYANA cannot treat symmetry with such operators since CYANA operates in torsional angle space. Instead a segment of the fibril is calculated that consists of several monomers. To enforce that all monomers have the same structure and are arranged in a symmetric way several pseudoenergy terms are added to the energy function that is optimized in the structure calculation.

Monomers are kept in the same conformation by adding an energy penalty if the backbone $\Phi$ and $\Psi$ angles of identical residues in different monomers are not equal. Additionally experimental distance restraints are always applied to all monomers and selected intramonomer distances are enforced to be equal in all monomers. Usually a small number of such symmetric distance restraints is sufficient to keep all monomers in symmetry, adding more slows down the calculation.

The organization of the monomers within the fibril is kept symmetric by applying all experimental restraints for interfaces between monomers to all equivalent interfaces and by restraining intermolecular distances to be the same across all equivalent interfaces.

To calculate the structure of a fibril with CYANA it is therefore necessary to know how many symmetric molecules have to be calculated and which interfaces between them should be equal. We discuss below how this information can be obtained.

## 6.3. Axial stacking

For cross-$\beta$-sheet structures the axial interface is defined by the hydrogen bonds between the $\beta$-strand layers. Determining the axial stacking (see Figure 6.1a) therefore can be reduced to determining whether the $\beta$-sheet is parallel or antiparallel and what the register-shift is. Intermolecular contacts between $\beta$-strands can most easily be identified from [$^{15}$N;$^{13}$C] correlation spectra (PAIN or NHHC) of fibrils produced from a mixture of only $^{15}$N and only $^{13}$C labeled monomers. Nitrogen atoms are easy to isotope label and mostly occur on the protein backbone. They are well suited to probe the $\beta$-sheet register since they are involved in the relevant hydrogen bonds.

There are two classes of $\beta$-strand–loop–$\beta$-strand motifs that can occur in fibrils. As a first possibility, the two $\beta$-strands can form an antiparallel $\beta$-sheet with hydrogen-bonds between
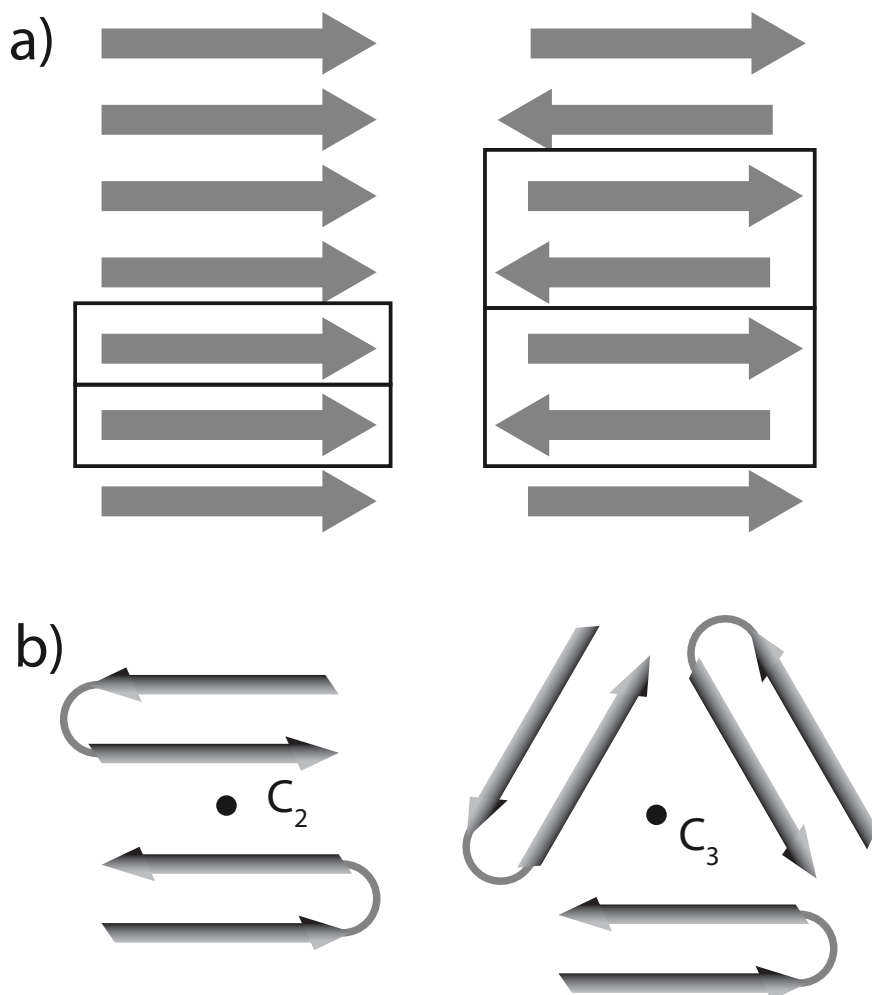
**Figure 6.1.** Intermolecular (quaternary) arrangement of monomers in amyloid fibrils. In cross-$\beta$-structures monomers are periodically repeated along the fibril axis. A network of hydrogen bonds connects different $\beta$-strands to the extended $\beta$-sheet. **a)** The $\beta$-strands can be stacked antiparallel or parallel either in-register or with a register offset (not shown). **b)** If several monomers are present in the transverse cross-section but give a single set of peaks they must be symmetrically arranged, such that equivalent atoms in different monomers have the same local environment. This criterium is fulfilled if a symmetry operation such as a rotation about a $C_n$ axis converts the coordinates of one monomer into that of another. Note: In cross-beta fibrils $\beta$-strands in the same transverse plane are connected by $\beta$ arcs [138] and the contacts between them are via the side chains. Backbone hydrogen bonds only occur along the fibril axis.

the two strands. The loop is then called $\beta$-turn and the motif a $\beta$-hairpin. In that case the repeating unit along the fibril axis (black box) in Figure 6.1a) would contain two $\beta$-strands of the same monomer in antiparallel arrangement. The second possibility is that the two $\beta$-strands interact with their side-chains and each of the $\beta$-strands is in a different parallel $\beta$-sheet. In this case the loop is called $\beta$-arc and the motif is called a $\beta$-arch [138] (see Figure 6.2). Fibrils consisting of parallel-in-register $\beta$-strands such as A$\beta$1-40 E22$\Delta$ fibrils (see Figure 5.10) can therefore only have $\beta$-arches but no $\beta$-hairpins.
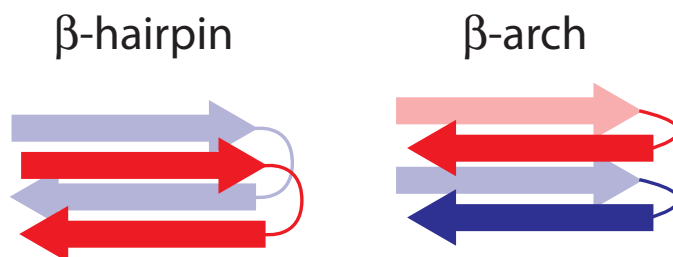


**Figure 6.2.** Schematic drawing of two hairpins and two $\beta$-arches. In a hairpin the two $\beta$-strands of a monomer fold back on each other and form a hydrogen bonded anti-parallel $\beta$-sheet. In a $\beta$-arch the $\beta$-strands from the same monomer interact with via their side-chains and are in two different parallel-in-register $\beta$-sheets.

## 6.4. Lateral contacts

Identification of lateral contacts (see Figure 6.1b) is challenging because they are side chain contacts. Side chains make both axial and lateral contacts and only a subset of them (namely methyl groups, see section 1.2.3) can be selectively isotope labeled. Analysis of spectra from diluted samples (prepared by co-fibrillization of labeled and unlabeled monomers) allows identifying intermolecular $^{13}$C–$^{13}$C peaks. The intensity of intra- and intermolecular peaks ($I_{inter}$ and $I_{intra}$) in the diluted fibril, compared that those in the uniformly labeled fibril ($I_0$) is given by:

$$I_{intra} = I_0 \cdot p_{labeled} \tag{6.1}$$
$$I_{inter} = I_0 \cdot p_{labeled}^2 \tag{6.2}$$

where $p_{labeled}$ stands for the fraction of $^{13}$C labeled monomers. For these equations it was assumed that only one interaction between two atoms contributes to the signal, i.e. that the
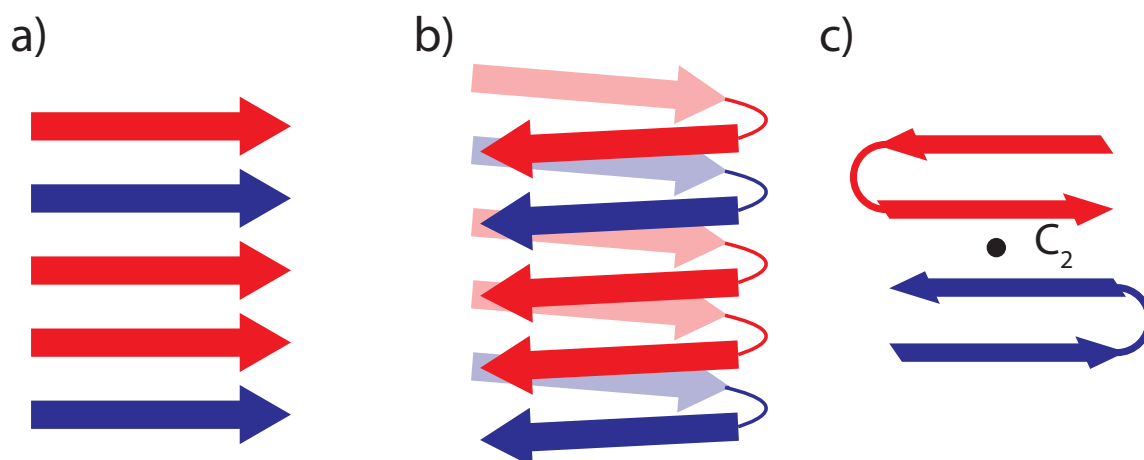
**Figure 6.3.** Schematic representation of fibrils consisting of differently isotope labelled monomers. Intermolecular contacts can be identified by measuring heteronuclear contacts in mixed [$^{13}$C]/[$^{15}$N] labeled samples or by comparing the peak intensity between diluted samples that in fully labelled samples. Intermolecular contacts are expected along the direction of the fibril ($\beta$-sheet register contacts, (a)), or transverse to the fibril direction. Transverse intermolecular contacts can be due to staggered side-chain contacts (b) or to due to several monomers in the transverse cross-section (c).

peak is either intra or intermolecular but not both. Intermolecular peaks are attenuated more strongly than intramolecular peaks because they are only observed when both monomers are isotope labeled.

Observed intermolecular $^{13}$C–$^{13}$C contacts can be axial or lateral, classifying them as lateral is only possible if the axial contacts are known and cannot explain the contact. Intermolecular lateral contacts can occur if the fibril contains more than one monomer in the transverse cross-section or if the $\beta$-sheets are staggered (see Figure 6.3). A staggered fibril refers to a fibril containing monomers in a $\beta$-strand-turn-$\beta$-strand, where the side chains of $\beta$-strand 1 are in contact with those of $\beta$-strand 2 of a different monomer[139]. Lührs et al. [133] proposed a staggered model for A$\beta$1-42 wt based on the observation that a mixture of D23K and K28D mutants forms fibrils whereas the two mutants do not from fibrils by themselves.

## 6.5. Fibril Width

While for axial contacts it is sufficient to know the interfaces between adjacent monomers and the number of them is assumed to be infinite, it is necessary to know the lateral interfaces and the number of monomers per transverse layer to determine the structure of a fibril. NMR

is well suited to gain insight into the local structure. The number of molecules per transverse fibril layer has to be measured with complementary methods that provide information on a larger length-scale. The most common way is by mass-per length measurements from scanning transmission electron microscopy (STEM)[140, 141].

## 6.6.  Structure calculation of A$\beta$1-40 E22$\Delta$ fibrils

The manually identified distance restraints for A$\beta$1-40 E22$\Delta$ fibrils are shown in Figure 6.4. Peaks from the [U-$^{13}$C] fibrils were compared to those in the mixed [U-$^{13}$C][U-$^{15}$N] (i.e. 50% diluted) fibrils (see Figure 6.5). Whenever possible, restraints were classified as intra- or intermolecular based on the observed intensity ratio between the sample (eq. 6.1 and 6.2). The dilution analysis in the C-terminal part of the peptide is inconclusive because there are differences between the [U-$^{13}$C] sample and the mixed [U-$^{13}$C][U-$^{15}$N] (i.e. 50% diluted) sample. Some intraresidue peaks of residues 28-31 in 15 ms DARR spectrum shift by more than 0.5 ppm and the intensity of peaks belonging to residues 35, 38 and 39 vary by more than 30%. These residues could thus not reliably be included in the dilution analysis.

Since for A$\beta$1-40 E22$\Delta$ fibrils mass-per-length data is not yet available we calculated structures for two and three molecules per transverse layer that are symmetrically placed. In addition to the manually identified unambiguous distance restraints, unassigned peak lists were automatically converted to ambiguous distance restraints by the CANDID algorithm[43]. Based on the secondary chemical shifts (see Figure 5.9) TALOS+[38] dihedral angle restraints were derived and hydrogen-bond restraints for a parallel-in register $\beta$-sheet were created from the predicted $\beta$-sheet regions. The results of the calculations are shown in Figures 6.6 and 6.7. Both the intra- and intermonomer contacts are very similar in both structure models, illustrating the need for complementary data.

The calculated structures dimer and trimer structures (Figures 6.6 and 6.7) do no explain why the contacts between residues 13/15 and the C-terminal residues 37-40 are intermolecular (see interaction matrix in Figure 6.4). This intermolecular contact could be explained by a staggered arrangement of the $\beta$-strands in the fibrils (see Figure 6.3) or because of domain-swapping. Two domain-swapped structure models are shown in Figure 6.8.
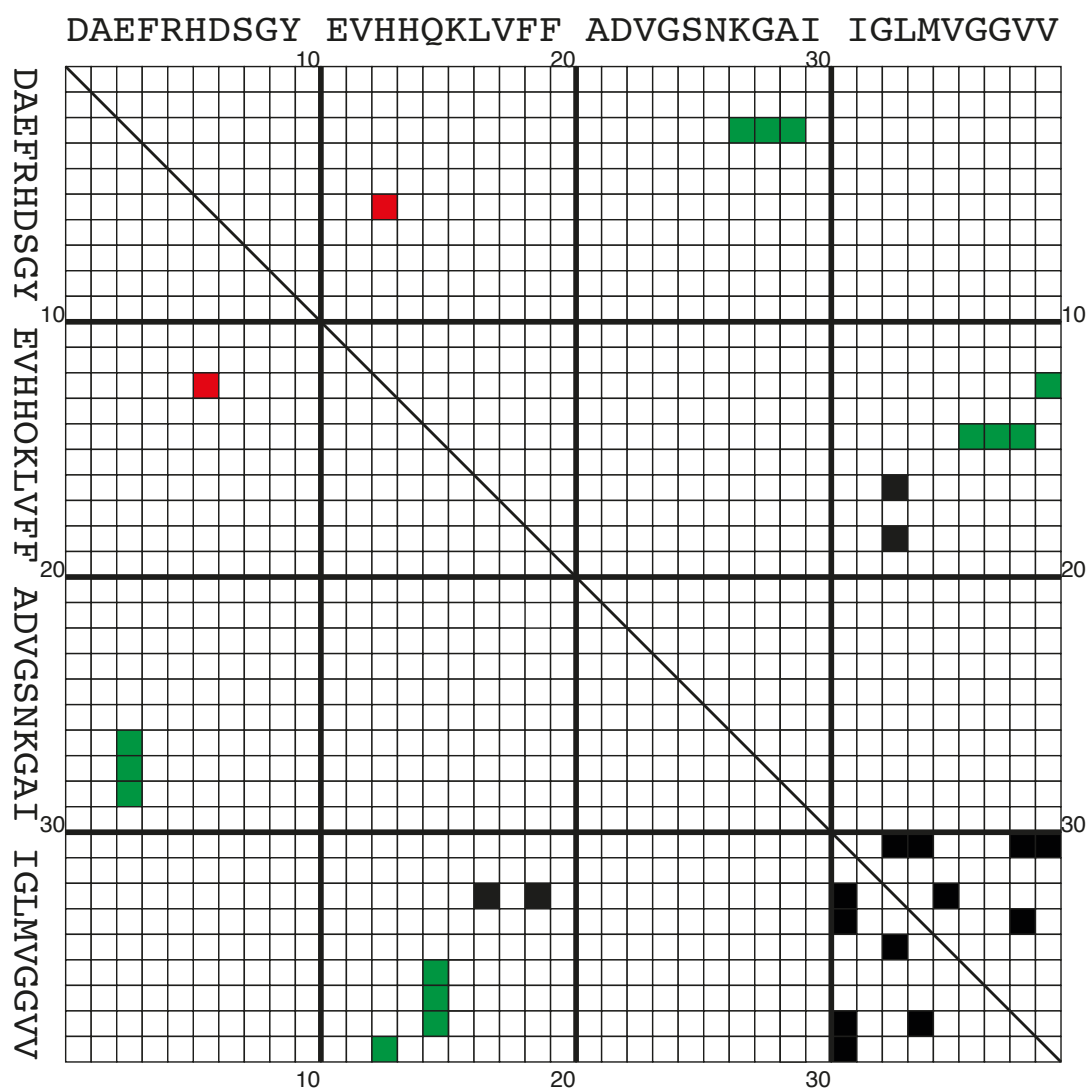
**Figure 6.4.** Interaction matrix of manually identified unambiguous contacts in Aβ 1-40 E22Δ fibrils. Classification into intra- and intermolecular restraints was done based on equations 6.1 and 6.2. Residues that had a different peak position ($> 0.5$ ppm) or peak intensity ($> 30\%$ deviation) in the 15 ms DARR spectrum of the [U-$^{13}$C] and mixed [U-$^{13}$C][U-$^{15}$N] sample were not classified as intra or intermolecular.
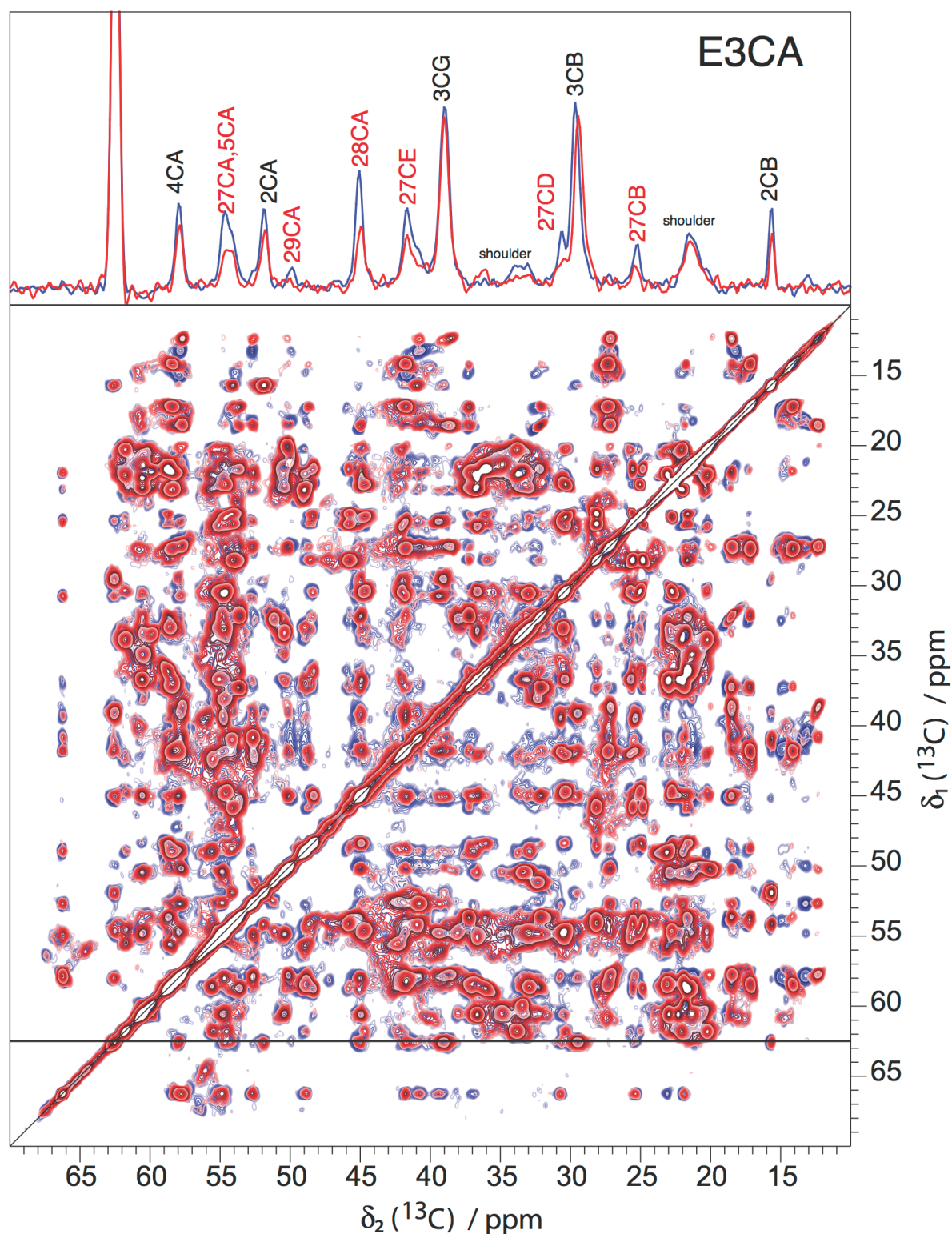
**Figure 6.5.** Superposition of an 8 ms PAR spectrum recorded at 20 T and 19 kHz MAS on a fully $^{13}$C labeled sample (blue) and on a 50% $^{13}$C labeled sample (red). The trace was extracted at the position of glutamate 3 C$\alpha$. The traces are scaled such that the intensity of the diagonal peak is equal. Cross-peaks corresponding to intermolecular contacts are expected to be attenuated by 50% compared to intramolecular contacts. Such peaks are labeled in red in the trace.
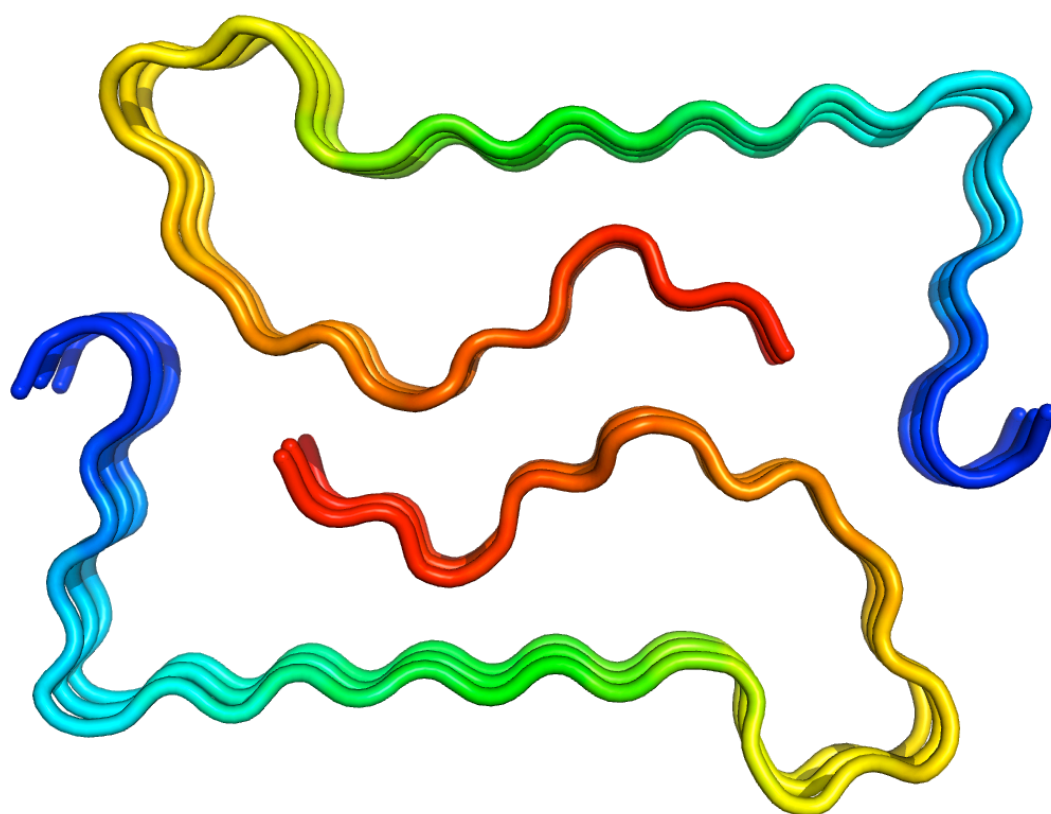
**Figure 6.6.** Result of a structure calculation with two protofibrils, imposed $C_2$ symmetry and restraints ambiguous between intra- and intermolecular. Only the backbone is plotted color coded in blue for the N-terminus to red for the C-terminus.
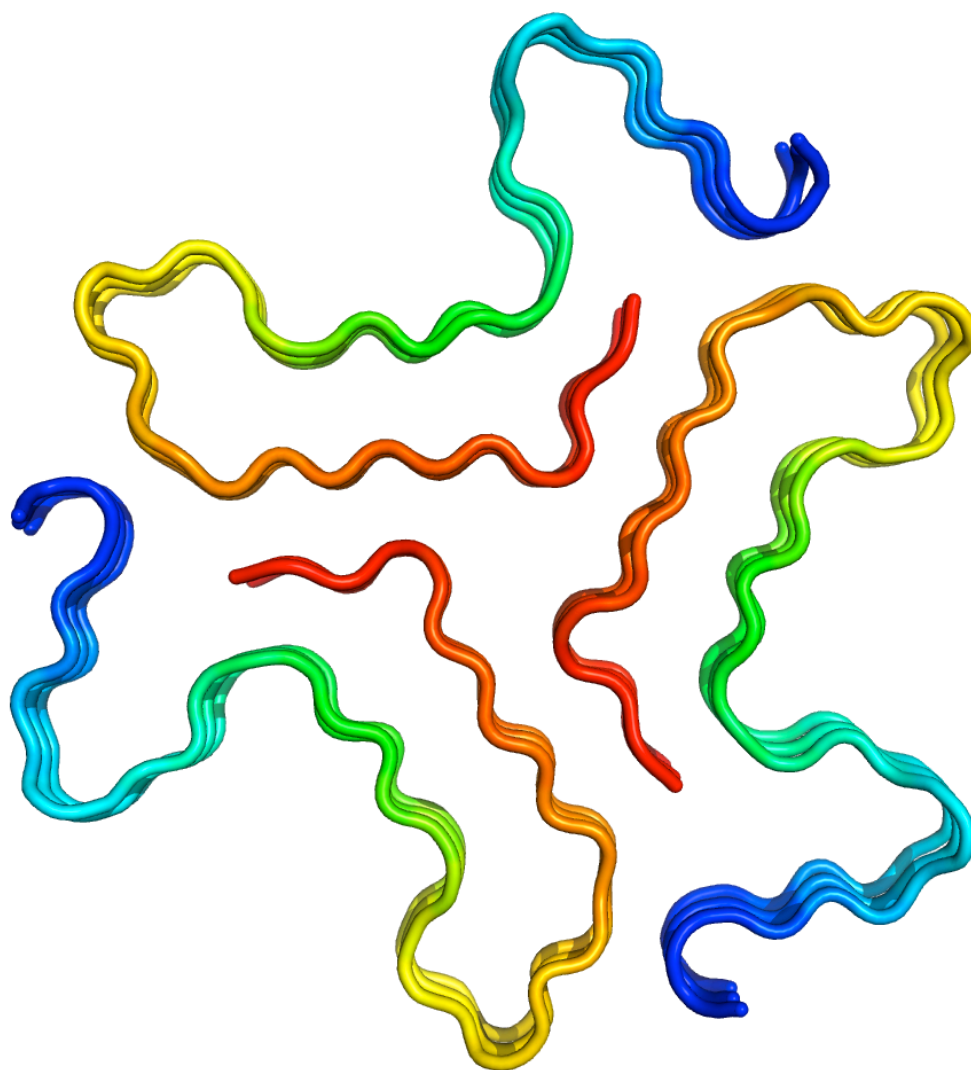
**Figure 6.7.** Result of a structure calculation with three protofibrils and imposed $C_3$ symmetry.
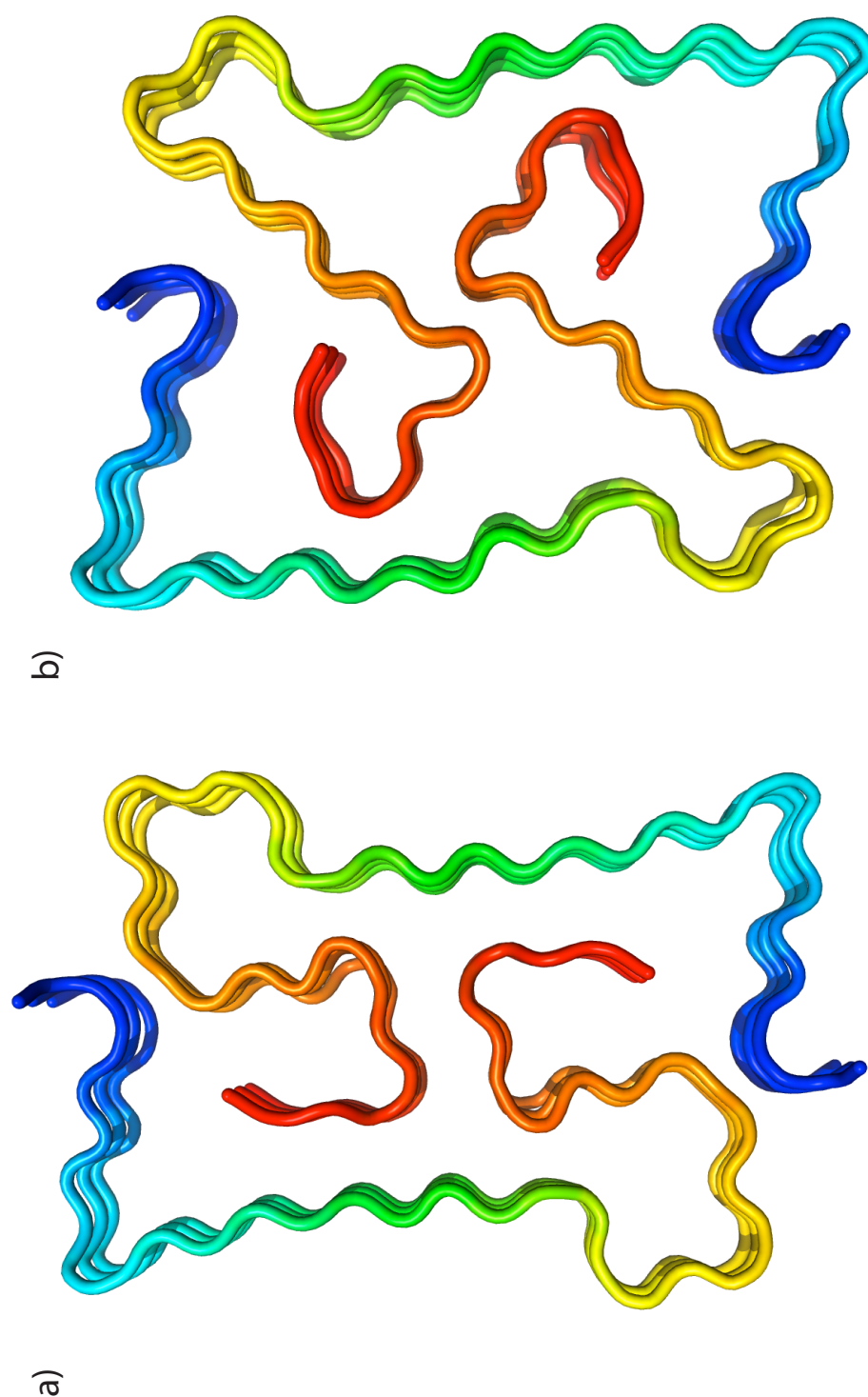
**Figure 6.8.** Results of structure calculations with two protofibrils and imposed C₂ symmetry. The contact between residues 36 and 15/17 is restrained to be intermolecular. In calculation a) the interaction 33-19 is restrained as intramolecular whereas in b) it restrained intermolecularly.

## 6.7. Verification of the calculated structures

Good structural models should explain the observed distance restraints and shifts and also be biologically plausible. Hydrophobic residues should be buried in the interior of the fibril and internal charge residues should be compensated (see Figure 6.9). We are able to calculate models that fulfill the experimental restraints and look plausible but currently the experimental restraints are insufficient to find a unique solution.

## 6.8. Conclusions

Fibrils with more than one molecule per transverse slice can be calculated in CYANA by adding more symmetry restraints and defining both the axial and lateral interfaces between monomers. Currently we do not have enough data to conclusively determine the structure of our A$\beta$1-40 E22$\Delta$ fibrils. Unambiguous distance restraints have been identified in the spectra and are summarized in the interaction matrix in Figure 6.4.

Because the fully labeled and diluted sample had different peak positions and intensities in the C-terminus, we do not know for these contacts whether they are intra- or intermolecular. A further complication to the experimental confirmation of lateral contacts in the C-terminus of A$\beta$1-40 E22$\Delta$ is that some of the proposed lateral contacts (see Figure 6.9) are between identical residues on different protofibrils (Val-35 for example). Such peaks are difficult to confirm as lateral because they can also be explained as axial register contacts. Even when contacts have been identified as lateral it remains unknown whether they are observed because of a staggered arrangement of the $\beta$-sheets (Figure 6.3b) or because of domain swapping (Figure 6.9).

The last missing elements for the determination of the structure of A$\beta$1-40 E22$\Delta$ fibrils are mass-per-length data to determine the number of monomers in the transverse cross-section of the fibril and spectra from a more highly diluted sample to distinguish intra- and intermolecular contacts within the C-terminus, which might help to decide whether A$\beta$1-40 E22$\Delta$ fibrils are staggered or laterally domain-swapped. The other necessary experimental data (secondary structure, $\beta$-sheet register and high quality restraint-spectra) as well as the computational tools for the calculation of the structure are available and have been tested for the calculation of the structural models shown in this chapter.
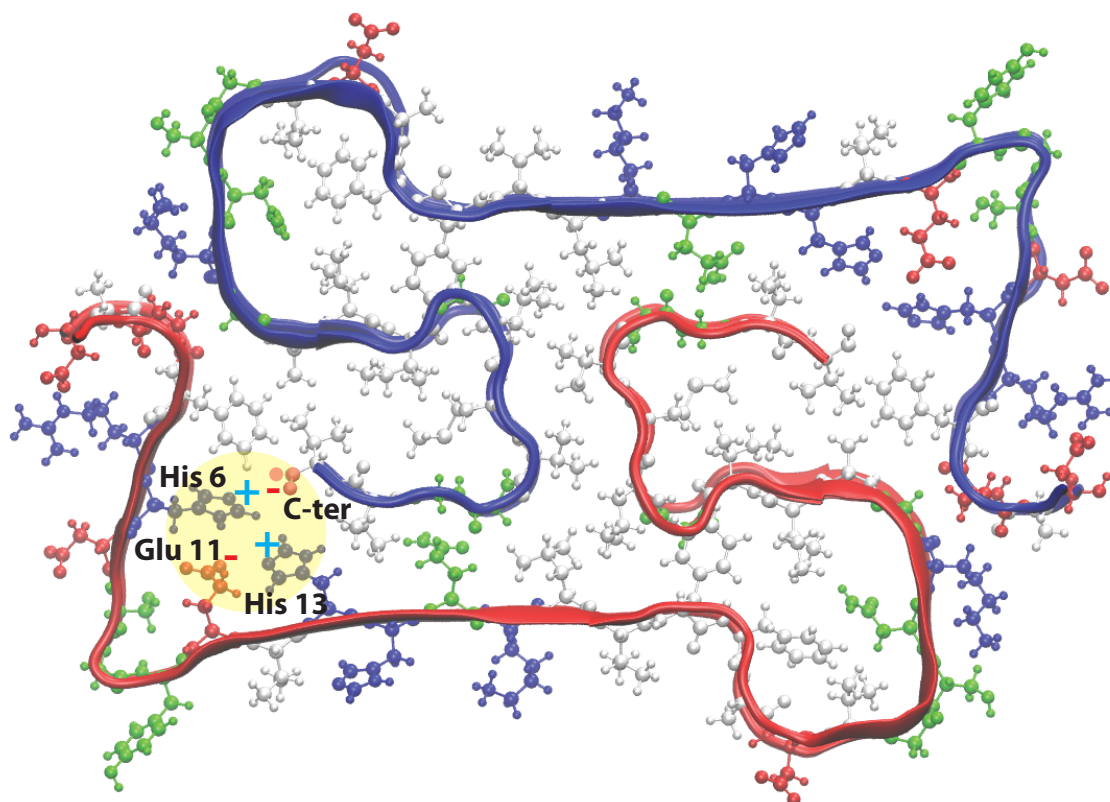
**Figure 6.9.** Structure model of Aβ1-40 E22Δ fibrils as in Figure 6.8 color-coded by amino acid type (white–hydrophobic, green–polar, red–negative, blue–positive). The two positive charges His 6 and His 13 are compensated by the two negative charges of Glu 11 and the C-terminal carboxylate of Val 40. (Charge compensation is also possible in other models which also are consistent with current experimental restraints).

# 7. Conclusions and Outlook

In this thesis solid-state NMR methods were developed and applied to the structural characterization of proteins. In chapter 2 the challenges of traditional solid-state NMR methods for structure determination were analyzed and it was shown that 3D and 4D proton-detected spectra can be used to increase resolution and sensitivity for measuring distance restraints. We developed hardware (a supplementary $^2$H coil, chapter 3) and a protocol (chapter 4) for structure determination of proteins in the solid-state and tested them on microcrystalline ubiquitin. Finally, solid-state NMR was applied to study fibrils of amyloid-$\beta$ (chapter 5), a peptide involved in Alzheimer's disease and a way to calculate fibril structures with several monomers in the transverse plane has been implemented in CYANA (chapter 6).

## Proton Detected Restraint Spectra

Ubiquitin sparsely $^1$H labeled against a deuterated background yields spectra of high resolution and sensitivity. Since the shortest contacts between protons in the methyl labeled sample often correspond to structurally interesting long-range peaks efficient first-order recoupling techniques were used to detect distance restraints. The highly sensitive $^1$H detection allowed recording a 4D HC(H-H)CH spectrum from which the restraints needed for successful structure calculation could be directly and unambiguously extracted. Attempts to increase the size of proteins amendable to solid-state NMR by applying the protocol used for ubiquitin to microcrystalline HET-s (1-227) (spectra shown in the appendix 8.1.8) have not yet succeeded due to problems with maintaining proper hydration levels in the 1.3 mm rotors and poor efficiency of stereospecific valine and leucine labeling. As these problems are not directly related to the protocol for structure determination from $^1$H-$^1$H contacts in sparsely labeled sample but rather a consequence of a particular choice of rotor and sample preparation, we expect that our protocol for structure calculation can be extended to proteins significantly larger than ubiquitin.

## Increasing Resolution in $^{13}$C–$^{13}$C spectra

An alternative method to increase resolution compared to 2D-[$^{13}$C;$^{13}$C] spectra from uniformly $^{13}$C labeled samples (if sparse $^1$H labeling is not possible or not advantageous) is sparse $^{13}$C labeling. Proteins expressed in media containing partially $^{13}$C labeled carbon sources (see sections 1.2.1 and 1.2.2) yield significantly narrower $^{13}$C lines as the homonuclear $^1$J couplings are mostly removed. We have prepared samples of ubiqutin, HET-s (218-289) and A$\beta$1-40 E22$\Delta$ in medium containing 1-$^{13}$C or 2-$^{13}$C glucose. The spectra are well resolved (see Figure 5.8) and long-range peaks have been identified. Incorporation of these restraints into structure calculation is under way.

## Amyloid $\beta$

We established a preparation of A$\beta$1-40 E22$\Delta$ fibrils that reproducibly yields highly resolved spectra. The sequential assignment of all residues was obtained, which represents the first step towards a high-resolution study of this peptide. $\beta$-strands were predicted from chemical shifts for residue 4 to 19 and 26-40 with possible turns/kincks at the glycines 9, 33 and 37. Peaks in the $^{15}$N–$^{13}$C correlations spectrum of mixed $^{13}$C / $^{15}$N fibrils are characteristic for parallel-in-register arrangement of the $\beta$-sheets. From dilution analysis it is apparent that some of the identified distance restraints must be between different monomers, indicating more than one monomer per layer. Since there is only one major set of peaks, all monomers must have the same structure and be symmetrically arranged. To determine the correct structure of the fibril solid-state NMR data will have to be combined with complementary techniques, such as a mass per length measurement.

Structure calculation for fibrils differs from that of globular proteins. Since intermolecular contacts (both longitudinally and possible laterally) play an important role in fibrils, recording and analyzing restraint spectra has to split up into work on differently isotope labeled samples. Finally some parameters such as the symmetry of the fibril have to be manually set for the calculation but then also decrease the available degrees of freedom.

## Solid-state NMR for structure determination of proteins

Using solid-state NMR for the determination of protein structures involves several steps:

1. producing and purifying the isotope labeled protein

2. bringing the protein into a suitable form (sample condition) for NMR experiments

3. assigning chemical shifts

4. recording restraint spectra

5. analyzing the restraint spectra

6. calculating a structure

7. verifying the structure

After spending about five years doing these steps successfully and sometimes also unsuccessfully on different proteins I take myself the right to use some space here to write down my personal opinion on how I think these different aspects should best be done and combined:

It is important to remember that each of these steps depends on the results of the previous steps. (i.e. once a sample condition has been chosen and assignment- as well as restraint measurements have started, changing the sample condition is not possible without repeating all spectroscopic work.) The better one aspect is solved, the easier the next step becomes. I believe that the best overall outcome can be achieved by giving attention to all aspects. While doing one step very badly has a good chance of rendering the entire work of structure determination impossible, optimizing one step to perfection will not solve the problem. Once a certain minimum level is achieved for each step the outcome can be improved at different places, i.e. with a medium quality sample it will be more difficult to record good spectra and take more time to analyze them. So either the sample preparation can be improved or the method to analyze spectra of lower quality can be improved.

A relatively simple way to produce protein samples and record restraint spectra is measuring 2D [$^{13}$C;$^{13}$C] DARR spectra of uniformly $^{13}$C labeled protein, yielding spectra as shown in Figure 2.5 on page 39. The problem is that peak overlap and restraint ambiguity is very high, which makes manual analysis very difficult. Automatic structure calculation from such a spectrum will usually find a solution. To verify if the solution is correct is quite difficult, though. In fact most structures derived from 2D $^{13}$C–$^{13}$C spectra of [U-$^{13}$C]-labeled proteins have been reported for model systems of known structure for which verification of the result can be simplified by checking whether the found structure matches the known structure. From the DARR spectrum of Ubiquitin recorded on the 600 MHz magnet (shown in Figure 2.6) I was unable to automatically calculate the correct structure but I did get it from a CHHC spectrum measured at 850 MHz (Figure 8.8 in the Appendix). Verifying that all long-range

restraints derived from the spectrum are correct and explaining why some peaks were ignored in the structure calculation is very difficult though.

I believe that sparse labeling and higher dimensional spectra as shown in chapter 4 provide a better balance of the work necessary for preparing a sample and recording spectra on the one hand and analyzing the spectra and verifying the calculated spectrum on the other hand.

Currently 3D- and 4D spectra for distance restraints are difficult to measure with sufficient sensitivity without proton-detection and it is not possible to get $^1$H line-widths and coherence lifetimes as seen for ubiquitin, SH3 or other microcrystalline (model) proteins for all other sample such as fibrils[142]. For samples not suitable for sparse $^1$H labeling it might be an option to use sparse $^{13}$C labeling (see section 1.2.1 and 1.2.2) to increase resolution by eliminating homonuclear one-bond $^{13}$C–$^{13}$C J-couplings and decreasing the effect of dipolar truncation (see section 1.4).

Since the different steps needed for structure determination need to be done sequentially and each of them should receive its due attention it is important to know when to move from one step to the next:

Before starting assignment and measuring restraint spectra, it should be established that (i) sufficient quantities of protein can be produced in high purity and with different isotope labeling patterns, that (ii) an NMR sample with sufficient resolution can reproducibly be prepared and (iii) that the sample is stable for the time it takes to record the different spectra.

Similarly before spending too much time analyzing restraint spectra, calculating structures and verifying them one should be convinced that the shift assignment is reasonably complete and correct and that the restraint spectra have the necessary quality to extract a coherent set of distance restraints.

The tricky part, for which experience or examples as presented in this thesis become important, is knowing what *sufficient*, *reasonable* or *necessary* mean.

# 8. Appendix

## 8.1. Part I: Methods

### 8.1.1. Expression and purification of ubiquitin

An overnight pre-culture is used to inoculate the desired medium. Cells are grown at 37 °C and 100-120 rpm until OD600 reaches 0.8 in LB or 0.6 in M9, respectively. The culture is induced and grown for 3-4 h in LB or until no further growth occurs in M9 (M9 composition is given in Table 8.1), respectively (For a SDS-PAGE gel of cell harvested at different induction times see Figure 8.1). Cells are centrifuged for 10 minutes at 6000 g and resuspended in lysis buffer for the Microfluidzer.

Perchloric acid is added to the lysis supernatant to a concentration of 3.5% w/v and vigorously shaken for a short time. After an incubation of 10 min at 40 °C precipitated proteins are pelleted at 4 °C and 20 min at 48000 g. The supernatant is neutralized using saturated Tris buffer (1.5-2.5x volume of perchloric acid, check pH with pH-strip!). The solution is exchanged with ubiquitin binding buffer by gel filtration using a HiPrep 26/10 desalting column. The protein is boiled at 85 °C for 10 min and the precipitate pelleted at 4 °C and 48000 g.

The supernatant is loaded on a 6 ml Resource S ion exchange column, washed with 5 column volumes of binding buffer and eluted with a gradient of 0-100% elution buffer over 20 column volumes (the chromatogram is shown as Fig 8.3). The fractions containing ubiquitin are dialyzed against MilliQ water overnight at 4°C and then lyophilized.

### 8.1.2. Preparation of Ubiquitin microcrystals

Ubiquitin is dissolved in 20 mM ammonium acetate buffer at pH 4.3 containing 0.05% sodium azide. The final ubiquitin concentration is 25 mg/ml ($M_{ubiquitin}$ = 8.56 kDa). The protein solution is filter sterilized and centrifuged for 5 min at 15000 rpm (4°C).

The precipitation buffer consists of 2-methylpentanediol-2,4 and citrate buffer at pH 4.1

**Table 8.1.** Components of M9 medium (for 1 liter, modified from [143]). For $^{15}$N or $^{13}$C labeling replace ammonium chloride, glucose or glycerol with isotope labeled compounds.

| Ingredient | Amount |
|---|---|
| $Na_2HPO_4$ | 6.8 g |
| $KH_2PO_4$ | 3.0 g |
| NaCl | 0.5 g |
| $NH_4Cl$ | 1 ml |
| water (to 1.0 l) | |
| $MgSO_4$ (1M) | 1 ml |
| $ZnCl_2$ (10 mM) | 1 ml |
| $FeCl_3$ (1 mM) | 1 ml |
| $CaCl_2$ (100 mM) | 1 ml |
| MEM Vitamin Mix | 10 ml |
| Glucose or Glycerol | 2.5 g |
| antibiotic | |
| Kanamycin (30 mg/ml) | 1 ml |
| or | |
| Ampicillin (150 mg/ml) | 1 ml |

**Table 8.2.** Buffers for Ion Exchange chromatography

| | |
|---|---|
| Ubiquitin Binding Buffer | 20mM $NH_4$ Acetate at pH 4.5 |
| Ubiquitin Elution Buffer | 250 mM $NH_4$ Acetate at pH 7.6 |
| Saturated Tris Buffer | Heating $H_2O$ overnight containing more Tris than is soluble ($> 3$ M) |

in a 3:2 volume ratio, containing 0.05% sodium azide. The citrate buffer can be prepared from citric acid and adjusted carefully with 1M NaOH to the desired pH of 4.1. The final concentration of citric acid in the precipitation buffer is 20 mM.

In each crystallization reservoir 500 $\mu$l of precipitation buffer is added. 37 $\mu$l of protein solution and 10 $\mu$l of precipitation buffer are mixed in the crystallization well. The crystallization reservoirs are sealed with adhesive tape and stored at 4°C. After about two weeks the wells should contain crystals (for a photo of the crystals see Figure 8.4 which can be harvested and centrifuged into MAS rotors. Rotor caps are glued to prevent water evaporation.

Buffers for crystallization can be prepared in advance and stored for several months at 4°C. The exact composition of ammonium acetate buffer and of the precipitation buffer is listed in Table 8.3 and 8.4. The crystallization of 10 mg purified and lyophilized [U-$^{13}$C$^{15}$N]-labeled ubiquitin is described in detail in Table 8.5. To fill a 3.2mm Bruker rotor 20-25 mg of Ubiquitin need to be crystallized.

**Table 8.3.** Composition of the Ammonium Acetate Buffer

| I. 40 ml, 0.1 M ammonium acetate, pH 4.3 (stock solution) |
|---|
| 1.  36 ml $H_2O$ |
| 2.  add 229 $\mu$l acetic acid (98.5% purity) |
| 3.  adjust with $NH_3$ to pH 4.3 |
| 4.  add 400 $\mu$l 5% $NaN_3$ |
| 5.  fill with $H_2O$ to final volume |
| 6.  filter sterilize the solution and store at 4°C until usage |
| **II. 50 ml, 20 mM ammonium acetate** |
| 1.  10 ml 1 M stock solution |
| 2.  40 ml $H_2O$ |

**Table 8.4.** Composition of the Precipitation Buffer

| I. 40 ml, 50 mM citrate buffer, pH 4.1 (stock solution) |
|---|
| 1.  385 mg citric acid [M=192.43 g/mol] |
| 2.  add 35 ml $H_2O$ |
| 3.  adjust with 1M NaoH to pH 4.1 |
| 4.  add 400 $\mu$l 5% $NaN_3$ |
| 5.  fill with $H_2O$ to final volume |
| 6.  filter sterilize the solution and store at 4°C until usage |
| **II. 25 ml Precipitation Buffer** |
| 1.  10 ml 50 mM citrate buffer, pH 4.1 |
| 2.  15 ml MPD (mix well) |

**Table 8.5.** Protocol for the Crystallization of Ubiquitin

| Crystallization of 10 mg uniformly $^{13}$C , $^{15}$N labeled ubiquitin |
|---|
| 1.  10 mg ubiquitin (purified, lyophilized), [M=8560 g/mol] |
| 2.  add 400 $\mu$l 20 mM ammonium acetate, pH 4.3 |
| 3.  filter sterilize and centrifuge for 5 min at 15000 rpm, 4°C |
| 4.  fill crystallization reservoir with 500 $\mu$l of precipitation buffer |
| 5.  mix 37 $\mu$l of protein solution and 10 $\mu$l of precipitation buffer in the bridge reservoir |
| 6.  crystallize for 10-14 days at 4°C |
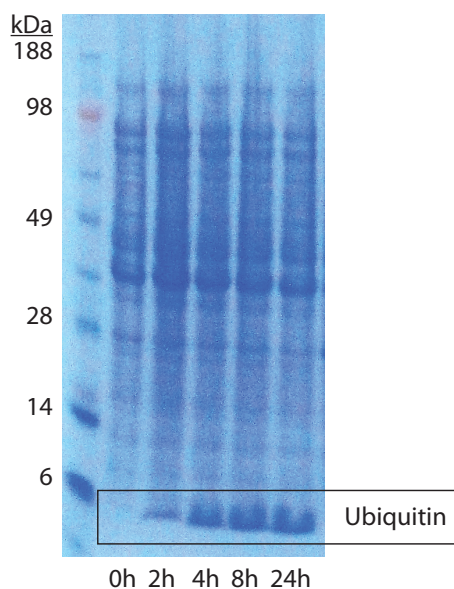
**Figure 8.1.** SDS-PAGE gel of BL21(DE3) cells transformed with the Ubiquitin plasmid growing in M9 medium in $D_2O$ before and at various time points after induction with 1 mM IPTG. For optimal yield of ubiquitin expression needs to be at least 4h but can also be longer (i.e. overnight) without lowering yield due to protein degradation.
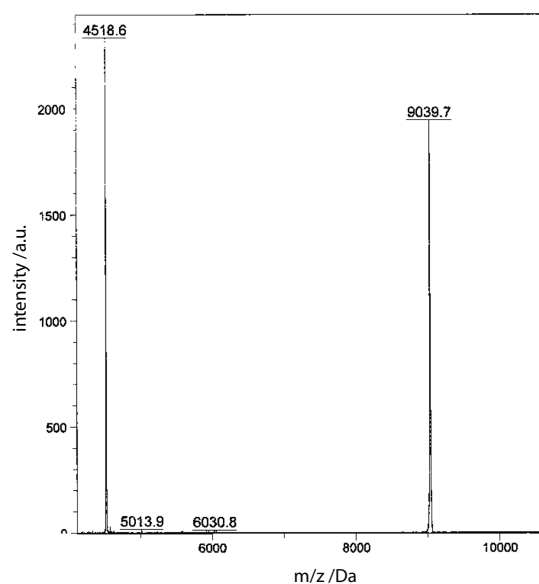


**Figure 8.2.** MALDI-TOF mass spectrum of [U-$^{13}$C$^{15}$N]-Ubiquitin.
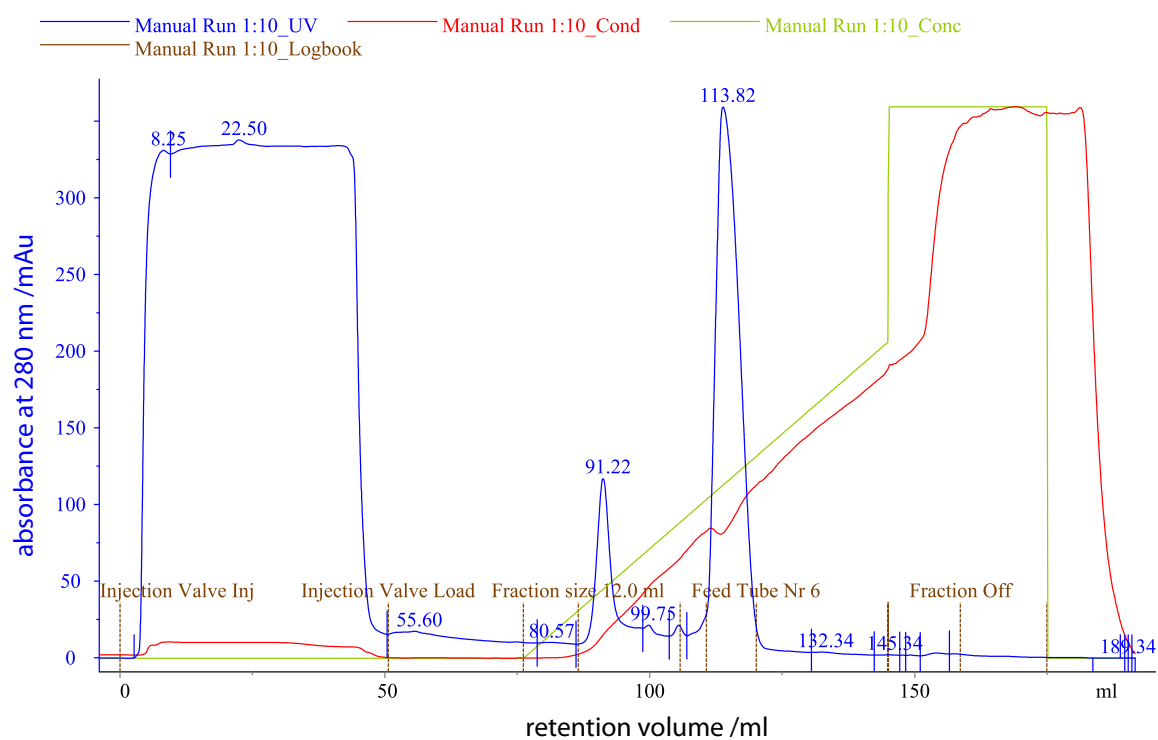
**Figure 8.3.** Chromatogram of the purification of Ubiquitin on a 6m Resource S column on an Aekta Prime system with conditions as given in section 8.1.1. Ubiquitin elutes at a concentration of about 40% of buffer B. The peak with a maximum of UV absorption at 113.82 ml corresponds to ubiquitin as verified by SDS-PAGE and MS (see Figure 8.2).
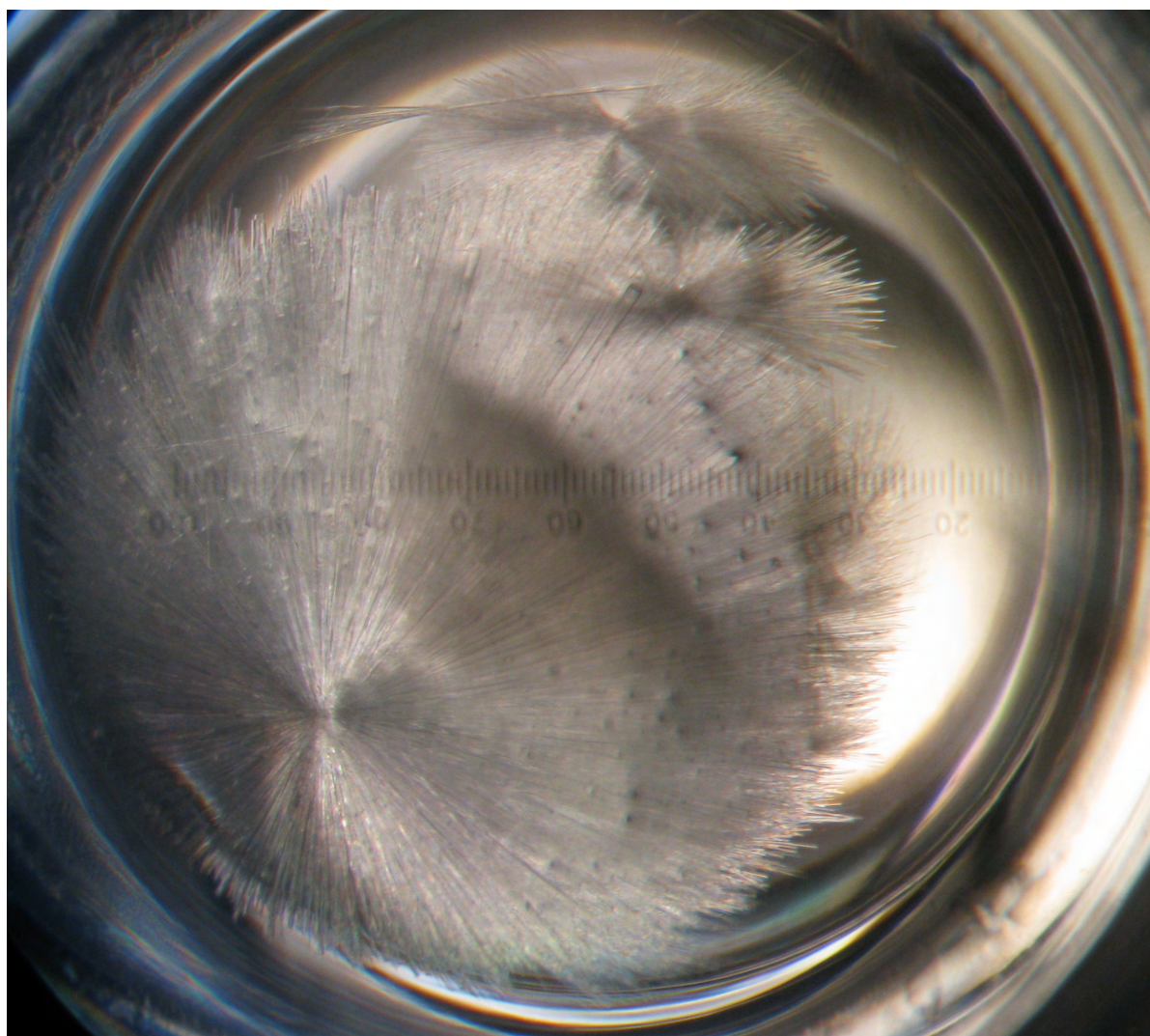
**Figure 8.4.** Ubiquitin crystals formed in a sitting drop crystallization setup from a solution containing MPD. The diameter of the drop is approximately 2 mm.

## 8.1.3. Chemical Shifts and Peaklist of the $^1$H-$^1$H DREAM spectra of Ubiquitin

Chemical shifts for the amide and ILV methyl resonances are given in Table 8.6. The peak list from the 3D amide–amide spectrum (see Table 8.7) was automatically assigned and converted to distance restraints during the structure calculation. Upper distance limits were set to 5.5 Å. For the observed contacts between methyl protons (see Table 8.8 in the Appendix) distance restraints between the corresponding pseudoatoms were created and set to 4.5 Å.

**Table 8.6.** Chemical Shifts of amide and ILV methyl groups of Ubiquitin.

| Residue | H | N | H | C | Residue | H | N | H | C |
|---|---|---|---|---|---|---|---|---|---|
| 2 Gln | 9.24 | 124.7 | | | 43 Leu | 8.97 | 123.0 | 0.73 | 25.5 |
| 3 Ile | 8.37 | 116.2 | 0.50 | 13.3 | | | | 0.77 | 23.0 |
| 4 Phe | 9.00 | 118.2 | | | 44 Ile | 9.10 | 122.4 | 0.65 | 13.3 |
| 5 Val | 9.22 | 120.1 | 0.67 | 19.9 | 45 Phe | 8.66 | 124.6 | | |
| | | | 0.67 | 21.8 | 46 Ala | 9.17 | 131.8 | | |
| 6 Lys | 8.96 | 127.8 | | | 47 Gly | 8.97 | 102.6 | | |
| 7 Thr | 8.84 | 117.6 | | | 48 Lys | 7.73 | 120.0 | | |
| 8 Leu | | | 0.96 | 25.2 | 49 Gln | 9.39 | 121.8 | | |
| | | | 1.01 | 23.6 | 50 Leu | 8.96 | 127.8 | -0.19 | 18.6 |
| 10 Gly | 7.87 | 108.3 | | | | | | 0.47 | 25.0 |
| 11 Lys | 7.19 | 122.5 | | | 51 Glu | 8.58 | 123.7 | | |
| 12 Thr | 8.84 | 121.0 | | | 52 Asp | 8.14 | 120.4 | | |
| 13 Ile | | | 0.66 | 13.9 | 53 Gly | 8.42 | 112.4 | | |
| 14 Thr | 8.52 | 121.6 | | | 54 Arg | 7.68 | 118.7 | | |
| 15 Leu | 8.84 | 124.2 | 0.64 | 26.1 | 55 Thr | 8.97 | 109.0 | | |
| | | | 0.8 | 23.7 | 56 Leu | 8.04 | 117.4 | 0.70 | 26.1 |
| 16 Glu | 8.18 | 122.3 | | | | | | 0.57 | 21.5 |
| 17 Val | 9.06 | 118.6 | 0.69 | 21.3 | 57 Ser | 8.21 | 112.7 | | |
| | | | 0.38 | 18.5 | 58 Asp | 8.09 | 124.5 | | |
| 18 Glu | 8.24 | 119.3 | | | 59 Tyr | 7.13 | 115.0 | | |
| 20 Ser | 6.88 | 103.0 | | | 60 Asn | 8.30 | 118.0 | | |
| 21 Asp | 7.92 | 124.2 | | | 61 Ile | 7.41 | 117.4 | 0.40 | 13.8 |
| 23 Ile | 8.55 | 119.4 | | | 62 Gln | 9.45 | 127.0 | | |
| 23 Ile | | | 0.49 | 7.8 | 63 Lys | 8.24 | 119.3 | | |
| 26 Val | 8.02 | 118.5 | 0.65 | 20.7 | 64 Glu | 9.20 | 109.9 | | |
| | | | 0.96 | 22.3 | 65 Ser | 7.09 | 111.9 | | |
| 27 Lys | 8.69 | 116.8 | | | 66 Thr | 9.22 | 120.1 | | |
| 29 Lys | 8.71 | 121.4 | | | 67 Leu | 8.95 | 123.9 | 0.45 | 22.5 |
| 30 Ile | 8.36 | 121.4 | 0.89 | 14.6 | | | | 0.54 | 24.2 |
| 31 Gln | 8.47 | 124.9 | | | 68 His | 8.83 | 117.6 | | |
| 32 Asp | 8.44 | 118.1 | | | 69 Leu | 8.31 | 125.0 | 0.67 | 22.8 |
| 33 Lys | 7.20 | 112.7 | | | | | | 0.81 | 25.4 |
| 34 Glu | 8.90 | 112.9 | | | 70 Val | 9.19 | 126.2 | 0.91 | 21.0 |
| 35 Gly | 9.20 | 109.8 | | | | | | 0.77 | 19.5 |
| 36 Ile | 6.17 | 119.8 | 0.75 | 13.0 | 71 Leu | 8.08 | 122.5 | 0.85 | 24.4 |
| 39 Asp | 8.50 | 112.3 | | | | | | 0.95 | 23.5 |
| 40 Gln | 7.75 | 116.6 | | | 73 Leu | | | 0.83 | 25.0 |
| 41 Gln | 7.55 | 117.7 | | | | | | 0.90 | 23.1 |
| 42 Arg | 8.55 | 120.7 | | | | | | | |

**Table 8.7.** Peaklist of the 3D H-HN DREAM spectrum with the most likely assignments. For structure calculation the unassigned peaklist was used.

| H | H(N) | N | H | HN |
|---|------|---|---|----|
| 9.25 | 9.25 | 124.8 | Q2 | Q2 |
| 8.39 | 8.37 | 116.2 | I3 | I3 |
| 8.85 | 8.35 | 116.2 | L15 | I3 |
| 9.20 | 8.38 | 116.2 | Q2 | I3 |
| 8.99 | 8.36 | 116.2 | F4 | I3 |
| 9.05 | 8.36 | 116.2 | V17 | I3 |
| 9.46 | 8.37 | 116.3 | Q62 | I3 |
| 9.01 | 9.00 | 118.2 | F4 | F4 |
| 8.37 | 9.01 | 118.2 | I3 | F4 |
| 9.21 | 8.99 | 118.3 | | F4 |
| 8.80 | 9.00 | 118.3 | L15 | F4 |
| 7.09 | 9.01 | 118.3 | S65 | F4 |
| 8.94 | 9.22 | 120.2 | K6 | V5 |
| 8.84 | 9.22 | 120.2 | L15 | V5 |
| 8.94 | 8.93 | 127.8 | K6 | K6 |
| 8.32 | 8.95 | 127.8 | L69 | K6 |
| 9.22 | 8.93 | 127.8 | V5 | K6 |
| 8.82 | 8.96 | 127.8 | T7 | K6 |
| 8.85 | 8.84 | 117.6 | T7 | T7 |
| 7.20 | 8.84 | 117.6 | K11 | T7 |
| 7.87 | 7.86 | 108.3 | G10 | G10 |
| 7.18 | 7.88 | 108.3 | K11 | G10 |
| 7.88 | 7.18 | 122.5 | G10 | K11 |
| 8.83 | 7.19 | 122.5 | T12 | K11 |
| 7.20 | 7.20 | 122.4 | K11 | K11 |
| 8.85 | 8.85 | 121.0 | T12 | T12 |
| 8.52 | 8.52 | 121.5 | T14 | T14 |
| 8.85 | 8.84 | 124.3 | L15 | L15 |
| 8.37 | 8.85 | 124.2 | I3 | L15 |
| 9.21 | 8.84 | 124.2 | V5 | L15 |
| 9.00 | 8.82 | 124.2 | F4 | L15 |
| 8.18 | 8.84 | 124.2 | E16 | L15 |
| 8.21 | 8.20 | 122.3 | E16 | E16 |
| 8.83 | 8.17 | 122.3 | L15 | E16 |
| 9.07 | 8.18 | 122.4 | V17 | E16 |
| 7.93 | 9.08 | 118.6 | V17 | V17 |
| 7.93 | 9.06 | 118.6 | V17 | V17 |
| 8.19 | 9.06 | 118.6 | E16 | V17 |
| 8.69 | 9.06 | 118.6 | | V17 |
| 7.69 | 9.06 | 118.6 | | V17 |
| 9.06 | 8.24 | 119.2 | V17 | E18 |
| 7.92 | 8.24 | 119.3 | D21 | E18 |
| 6.89 | 6.89 | 103.1 | S20 | S20 |
| 7.93 | 6.87 | 103.1 | D21 | S20 |
| 7.93 | 7.91 | 124.3 | D21 | D21 |
| 8.05 | 7.91 | 124.2 | L56 | D21 |
| 8.24 | 7.92 | 124.3 | E18 | D21 |
| 6.89 | 7.93 | 124.2 | S20 | D21 |
| 8.55 | 8.55 | 119.5 | I23 | I23 |
| 7.68 | 8.55 | 119.5 | R54 | I23 |
| 8.03 | 8.55 | 119.5 | L56 | I23 |
| 8.03 | 8.02 | 118.6 | V26 | V26 |
| 8.70 | 8.00 | 118.5 | V26 | V26 |
| 8.70 | 8.00 | 116.9 | K27 | V26 |
| 8.01 | 8.71 | 116.9 | V26 | K27 |
| 8.36 | 8.72 | 116.9 | I30 | K27 |
| 8.70 | 8.71 | 121.3 | K29 | K29 |
| 8.36 | 8.72 | 121.4 | I30 | K29 |
| 8.37 | 8.36 | 121.4 | I30 | I30 |
| 8.70 | 8.35 | 121.4 | K29 | I30 |
| 8.48 | 8.47 | 124.9 | Q31 | Q31 |
| 8.45 | 8.44 | 118.1 | D32 | D32 |
| 7.19 | 8.46 | 118.0 | K33 | D32 |
| 7.21 | 7.20 | 112.8 | K33 | K33 |
| 8.90 | 7.20 | 112.7 | E34 | K33 |
| 9.20 | 7.20 | 112.7 | G35 | K33 |
| 8.43 | 7.19 | 112.7 | D32 | K33 |
| 8.46 | 7.19 | 112.7 | Q31 | K33 |
| 8.91 | 8.90 | 112.9 | E34 | E34 |
| 7.21 | 8.88 | 112.9 | K33 | E34 |
| 6.18 | 8.90 | 112.9 | I36 | E34 |
| 8.45 | 8.90 | 113.0 | Q31 | E34 |
| 9.19 | 8.89 | 112.9 | G35 | E34 |
| 8.38 | 8.89 | 112.9 | I30 | G35 |
| 9.23 | 9.23 | 109.9 | G35 | G35 |
| 6.18 | 9.22 | 109.9 | I36 | G35 |
| 7.20 | 9.21 | 109.8 | K33 | G35 |
| 8.89 | 9.22 | 109.7 | E34 | G35 |
| 6.19 | 6.18 | 119.8 | I36 | I36 |
| 9.21 | 6.16 | 119.8 | G35 | I36 |
| 8.90 | 6.17 | 119.8 | E34 | I36 |
| 8.51 | 8.50 | 112.3 | D39 | D39 |
| 7.75 | 8.51 | 112.4 | Q40 | D39 |
| 7.54 | 8.49 | 112.3 | Q41 | D39 |
| 7.00 | 8.50 | 112.4 | | D39 |
| 7.76 | 7.75 | 116.6 | Q40 | Q40 |
| 8.51 | 7.74 | 116.6 | D39 | Q40 |
| 7.55 | 7.55 | 117.6 | Q41 | Q41 |
| 8.51 | 7.55 | 117.7 | D39 | Q41 |
| 8.56 | 8.55 | 120.8 | R42 | R42 |
| 9.20 | 8.54 | 120.7 | V70 | R42 |
| 8.96 | 8.96 | 123.0 | L43 | L43 |
| 9.09 | 9.00 | 122.9 | I44 | L43 |
| 9.11 | 9.11 | 122.5 | I44 | I44 |
| 8.67 | 9.11 | 122.5 | F45 | I44 |
| 9.20 | 9.09 | 122.5 | A46 | I44 |
| 8.32 | 9.11 | 122.5 | L69 | I44 |
| 8.97 | 9.11 | 122.5 | L43 | I44 |
| 8.85 | 9.11 | 122.5 | H68 | I44 |
| 8.67 | 8.66 | 124.7 | F45 | F45 |
| 8.97 | 8.65 | 124.7 | G47 | F45 |
| 9.10 | 8.65 | 124.7 | I44 | F45 |
| 9.18 | 8.66 | 124.7 | A46 | F45 |
| 7.74 | 8.66 | 124.6 | K48 | F45 |
| 9.40 | 8.66 | 124.6 | Q49 | F45 |
| 9.18 | 9.18 | 131.9 | A46 | A46 |
| 8.95 | 9.19 | 131.8 | G47 | A46 |
| 8.66 | 9.17 | 131.8 | F45 | A46 |
| 8.97 | 8.97 | 102.7 | G47 | G47 |
| 8.65 | 8.97 | 102.6 | F45 | G47 |
| 7.73 | 8.98 | 102.7 | K48 | G47 |
| 9.18 | 8.95 | 102.6 | A46 | G47 |
| 7.75 | 7.75 | 120.1 | K48 | K48 |
| 8.66 | 7.72 | 120.1 | F45 | K48 |
| 8.97 | 7.72 | 120.1 | G47 | K48 |
| 9.18 | 7.74 | 120.1 | A46 | K48 |
| 9.39 | 9.39 | 121.9 | Q49 | Q49 |
| 8.98 | 8.98 | 127.8 | L50 | L50 |
| 8.65 | 8.97 | 127.7 | F45 | L50 |
| 9.11 | 8.98 | 127.8 | I44 | L50 |
| 9.39 | 8.97 | 127.8 | A46 | L50 |
| 8.58 | 8.58 | 123.7 | E51 | E51 |
| 8.15 | 8.15 | 120.5 | D52 | D52 |
| 8.43 | 8.43 | 112.3 | G53 | G53 |
| 7.69 | 7.67 | 118.8 | R54 | R54 |
| 8.55 | 7.67 | 118.7 | I23 | R54 |
| 8.97 | 7.68 | 118.7 | T55 | R54 |
| 8.98 | 8.98 | 109.0 | T55 | T55 |
| 8.06 | 8.05 | 117.4 | T55 | T55 |
| 7.92 | 8.05 | 117.4 | D21 | T55 |
| 8.55 | 8.05 | 117.4 | I23 | L56 |
| 8.97 | 8.04 | 117.4 | T55 | L56 |
| 7.14 | 8.05 | 117.4 | Y59 | L56 |
| 8.21 | 8.21 | 112.7 | S57 | S57 |
| 8.09 | 8.09 | 124.4 | D58 | D58 |
| 7.14 | 8.09 | 124.5 | Y59 | D58 |
| 8.98 | 8.07 | 124.6 | T55 | D58 |
| 7.15 | 7.14 | 115.0 | Y59 | Y59 |
| 8.30 | 7.12 | 115.0 | N60 | Y59 |
| 8.10 | 7.12 | 115.0 | D58 | Y59 |
| 7.42 | 7.12 | 115.0 | I61 | Y59 |
| 8.98 | 7.12 | 115.0 | T55 | Y59 |
| 8.03 | 7.11 | 115.0 | L56 | Y59 |
| 8.31 | 8.29 | 117.9 | N60 | N60 |
| 7.41 | 8.31 | 117.9 | I61 | N60 |
| 7.14 | 8.32 | 117.9 | Y59 | N60 |
| 7.42 | 7.40 | 117.4 | I61 | I61 |
| 7.13 | 7.41 | 117.4 | Y59 | I61 |
| 8.30 | 7.39 | 117.3 | N60 | I61 |
| 8.02 | 7.40 | 117.3 | L56 | I61 |
| 7.22 | 7.41 | 117.4 | | I61 |
| 9.45 | 9.45 | 127.0 | Q62 | Q62 |
| 9.21 | 9.46 | 127.0 | T66 | Q62 |
| 8.38 | 9.19 | 110.1 | I3 | E64 |
| 7.08 | 9.20 | 110.2 | S65 | E64 |
| 9.19 | 9.17 | 110.2 | E64 | E64 |
| 8.99 | 9.20 | 110.0 | F4 | E64 |
| 8.25 | 9.19 | 110.2 | K63 | E64 |
| 7.10 | 7.09 | 111.9 | S65 | S65 |
| 9.00 | 7.08 | 111.9 | F4 | S65 |
| 9.20 | 7.09 | 111.9 | E64 | S65 |
| 9.46 | 9.19 | 120.2 | Q62 | T66 |
| 8.96 | 8.96 | 124.0 | L67 | L67 |
| 8.88 | 8.87 | 117.9 | H68 | H68 |
| 9.12 | 8.85 | 118.0 | I44 | H68 |
| 8.35 | 8.84 | 117.9 | L69 | H68 |
| 9.10 | 8.31 | 125.1 | I44 | L69 |
| 9.20 | 8.31 | 125.0 | V70 | L69 |
| 8.94 | 8.31 | 125.0 | K6 | L69 |
| 8.85 | 8.33 | 125.0 | H68 | L69 |
| 9.21 | 9.20 | 126.3 | V70 | V70 |
| 9.09 | 9.20 | 126.4 | I44 | V70 |
| 8.55 | 9.20 | 126.3 | R42 | V70 |
| 8.32 | 9.19 | 126.3 | L69 | V70 |
| 8.97 | 9.21 | 126.4 | L43 | V70 |
| 8.08 | 9.19 | 126.3 | L71 | V70 |
| 8.08 | 8.08 | 122.5 | L71 | L71 |
| 6.59 | 6.59 | 109.4 | V70 | L71 |
| 7.60 | 7.60 | 109.5 | | |
| 6.73 | 6.72 | 111.2 | | |
| 7.52 | 7.52 | 111.3 | | |
| 7.66 | 7.66 | 105.3 | | |
| 7.23 | 7.23 | 105.3 | | |
| 6.84 | 6.84 | 104.8 | | |
| 6.03 | 6.03 | 104.8 | | |
| 6.22 | 6.21 | 112.3 | | |
| 8.33 | 8.32 | 125.0 | | |
| 9.23 | 9.23 | 120.2 | | |
| 8.25 | 8.24 | 119.4 | | |
| 7.71 | 7.71 | 110.3 | | |
| 7.24 | 7.23 | 109.9 | | |
| 7.01 | 7.01 | 109.9 | | |
| 8.50 | 6.99 | 109.9 | D39 | |
| 7.42 | 7.22 | 105.3 | I61 | |
| 9.00 | 9.23 | 120.2 | F4 | |
| 7.71 | 7.21 | 105.2 | | |

**Table 8.8.** Observed contacts in the 4D HC-CH DREAM spectrum of Ubiquitin

| 5  | VAL | H$\gamma$1 | - | 43 | LEU | H$\delta$1 |
|----|-----|-----|---|----|-----|-----|
| 5  | VAL | H$\gamma$1 | - | 43 | LEU | H$\delta$2 |
| 5  | VAL | H$\gamma$1 | - | 69 | LEU | H$\delta$2 |
| 5  | VAL | H$\gamma$2 | - | 15 | LEU | H$\delta$1 |
| 5  | VAL | H$\gamma$2 | - | 26 | VAL | H$\gamma$1 |
| 5  | VAL | H$\gamma$2 | - | 43 | LEU | H$\delta$2 |
| 5  | VAL | H$\gamma$2 | - | 67 | LEU | H$\delta$1 |
| 5  | VAL | H$\gamma$2 | - | 15 | LEU | H$\delta$1 |
| 15 | LEU | H$\delta$1 | - | 17 | VAL | H$\gamma$2 |
| 15 | LEU | H$\delta$1 | - | 26 | VAL | H$\gamma$1 |
| 5  | VAL | H$\gamma$2 | - | 15 | LEU | H$\delta$2 |
| 15 | LEU | H$\delta$2 | - | 26 | VAL | H$\gamma$1 |
| 15 | LEU | H$\delta$1 | - | 17 | VAL | H$\gamma$1 |
| 17 | VAL | H$\gamma$1 | - | 26 | VAL | H$\gamma$1 |
| 17 | VAL | H$\gamma$1 | - | 26 | VAL | H$\gamma$2 |
| 17 | VAL | H$\gamma$2 | - | 56 | LEU | H$\delta$1 |
| 5  | VAL | H$\gamma$2 | - | 26 | VAL | H$\gamma$1 |
| 15 | LEU | H$\delta$1 | - | 26 | VAL | H$\gamma$1 |
| 26 | VAL | H$\gamma$1 | - | 43 | LEU | H$\delta$1 |
| 26 | VAL | H$\gamma$1 | - | 67 | LEU | H$\delta$1 |
| 26 | VAL | H$\gamma$2 | - | 56 | LEU | H$\delta$1 |
| 26 | VAL | H$\gamma$2 | - | 56 | LEU | H$\delta$2 |
| 26 | VAL | H$\gamma$1 | - | 43 | LEU | H$\delta$1 |
| 43 | LEU | H$\delta$1 | - | 50 | LEU | H$\delta$1 |
| 43 | LEU | H$\delta$1 | - | 67 | LEU | H$\delta$1 |
| 5  | VAL | H$\gamma$1 | - | 43 | LEU | H$\delta$2 |
| 5  | VAL | H$\gamma$2 | - | 43 | LEU | H$\delta$2 |
| 26 | VAL | H$\gamma$1 | - | 43 | LEU | H$\delta$2 |
| 43 | LEU | H$\delta$2 | - | 67 | LEU | H$\delta$1 |
| 50 | LEU | H$\delta$1 | - | 67 | LEU | H$\delta$1 |
| 50 | LEU | H$\delta$1 | - | 67 | LEU | H$\delta$2 |
| 50 | LEU | H$\delta$2 | - | 67 | LEU | H$\delta$2 |
| 26 | VAL | H$\gamma$2 | - | 56 | LEU | H$\delta$1 |
| 26 | VAL | H$\gamma$2 | - | 56 | LEU | H$\delta$2 |
| 50 | LEU | H$\delta$1 | - | 56 | LEU | H$\delta$2 |
| 56 | LEU | H$\delta$2 | - | 67 | LEU | H$\delta$1 |
| 5  | VAL | H$\gamma$2 | - | 67 | LEU | H$\delta$1 |
| 26 | VAL | H$\gamma$1 | - | 67 | LEU | H$\delta$1 |
| 43 | LEU | H$\delta$1 | - | 67 | LEU | H$\delta$1 |
| 43 | LEU | H$\delta$2 | - | 67 | LEU | H$\delta$1 |
| 50 | LEU | H$\delta$1 | - | 67 | LEU | H$\delta$1 |
| 5  | VAL | H$\gamma$2 | - | 67 | LEU | H$\delta$2 |
| 26 | VAL | H$\gamma$1 | - | 67 | LEU | H$\delta$2 |
| 43 | LEU | H$\delta$1 | - | 67 | LEU | H$\delta$2 |
| 43 | LEU | H$\delta$2 | - | 67 | LEU | H$\delta$2 |
| 50 | LEU | H$\delta$1 | - | 67 | LEU | H$\delta$2 |
| 5  | VAL | H$\gamma$1 | - | 69 | LEU | H$\delta$2 |
| 69 | LEU | H$\delta$1 | - | 71 | LEU | H$\delta$1 |

## 8.1.4. Correlation of peak intensity and distance in the 4D $^1$H-$^1$H DREAM
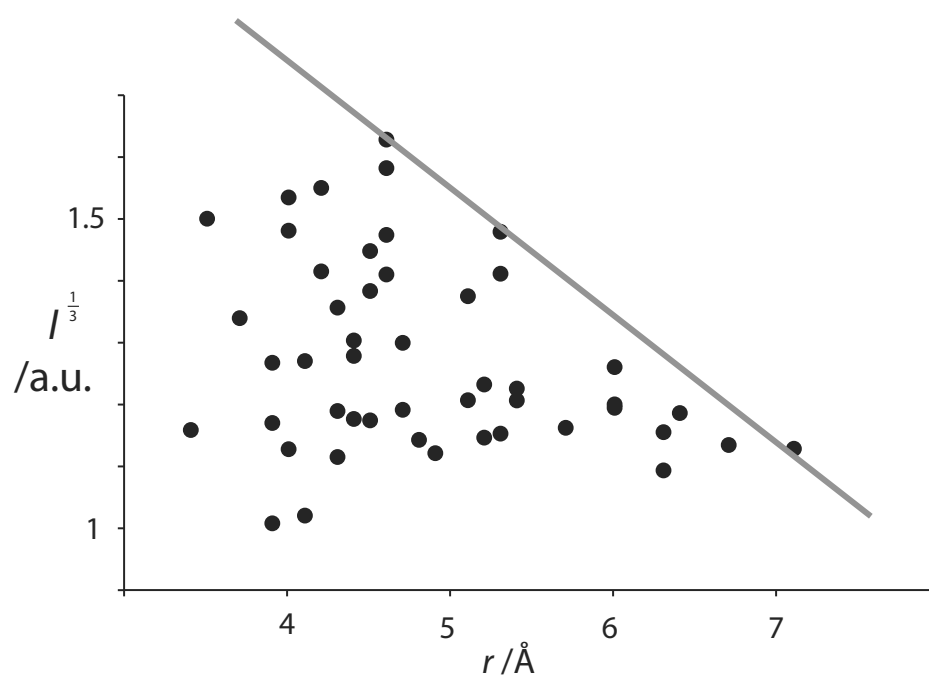


**Figure 8.5.** Observed cross-peak intensities in a 4D DREAM of [U-$^2$H$^{15}$N; VL-$^{13}$CD$_2$H] ubiquitin crystals versus the corresponding $^{13}$C–$^{13}$C distance in the xray-structure (pdb 1UBQ). The grey line indicates a possible calibration curve for the conversion of peak intensities into upper limits for distance constraints. For the structure calculation the lengths of upper-distance restraints were not calibrated based on intensity and set to 4.5 Å between the hydrogen pseudoatoms instead.

## 8.1.5. T$_2$ measurements of [U-$^2$H$^{15}$N; VL-$^{13}$CHD$_2$]-Ubiquitin
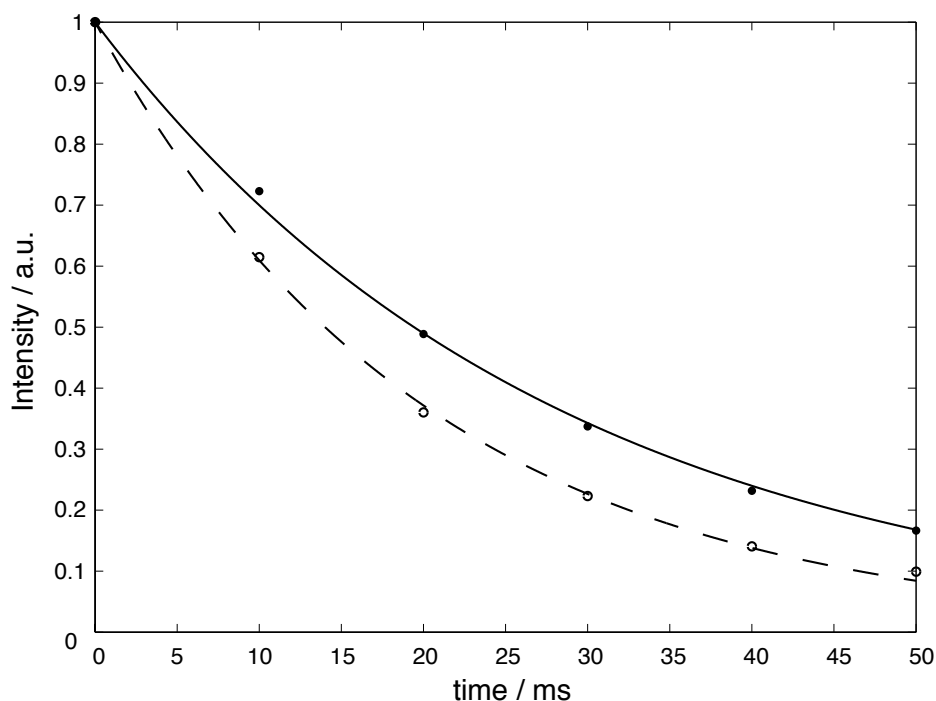


**Figure 8.6.** T$_2$ relaxation curves of integrated $^1$H (black) and $^{13}$C (dashed) methyl region in crystalline [U-$^2$H$^{15}$N; VL-$^{13}$CD$_2$H] ubiquitin under the conditions used for the DREAM experiments measured by a spin-echo pulse sequence. The solid lines are exponential fits with a relaxation time of 28 ms and 20 ms respectively. This corresponds a natural line width of 11 Hz and 16 Hz (FWHH). The experimental line width is 40 Hz in $^1$H and 30 Hz in $^{13}$C.

## 8.1.6. Pulse sequence for 4D $^1$H-$^1$H DREAM

A significant reduction of the length of the sequence is possible if water-suppression is not needed or can be achieved without the MISSISSIPPI [106] scheme, which would allow to use simple INEPTs instead of refocussed INEPTs for the magnetization transfers for the t$_3$-evolution. Pulse sequences without water suppression were used for measurements on HETs(1-227) crystallized in 100% D$_2$ and perdeuterated PEG (see section 8.1.8).

An alternative for solvent suppression is the insertion of a spin-lock purge pulse into an IN-EPT element[144]. After the last $\pi/2$ pulse on protons (see Figure 8.7) we have heteronuclear antiphase magnetization $2I_{-y}S_z$ and (unwanted) in phase magnetization $I_{-y}$ of the solvent. The heteronuclear J-coupling acts on the signal during $2\Delta$ converting the antiphase magneti-

zation into in-phase $I_x$ magnetization that is then spin-locked along the x-axis. The solvent signal remains aligned along the -y axis and is dephased by the rf-inhomogeneity of the spin-lock pulse. Implementing this idea under MAS might not be as straight-forward as in solution NMR since resonance conditions between spin lock and MAS that occur at ratios $\kappa = \omega_1/\omega r$ of $\kappa = 1/4$, $1/3$, $1/2$, $1$, $3/2$ and $2$ lead to a fast loss of spin-locked magnetization[145] and therefore have to be avoided.

The length of the sequence can be further reduced by incorporating the first INEPT transfer into the $t_1$ evolution period[146]. The semi-constant-time evolution is done such that for $t_1 = 0$ the length of the evolution corresponds to $2 \cdot \Delta$ where the INEPT delay is $\Delta = \frac{1}{4 \cdot J_{IS}}$ and for $t_1 = t_1^{max}$ the evolution times becomes $t_1^{max}$ (assuming $t_1^{max} > 2\Delta$). The delays $t_{1a}$, $t_{1b}$ and $t_{1c}$ (see Figure 8.7) have to be chosen such that

$$t_1 \quad = \quad t_{1a} + t_{1b} - t_{1c} \tag{8.1}$$

$$2\Delta \quad = \quad t_{1a} - t_{1b} + t_{1c} \tag{8.2}$$

For $t_1 = 0$ these equations can be fulfilled with $t_{1a}(0) = t_{1c}(0) = \Delta$ and $t_{1b}(0) = 0$. For arbitrary $t_1$ times we find:

$$t_{1a} \quad = \quad \frac{t_1}{2} + \Delta \tag{8.3}$$

$$t_{1b} \quad = \quad \frac{t_1}{2} \cdot \left(1 - \frac{2\Delta}{t_1^{max}}\right) \tag{8.4}$$

$$t_{1c} \quad = \quad \Delta \cdot \left(1 - \frac{t_1}{t_1^{max}}\right) \tag{8.5}$$
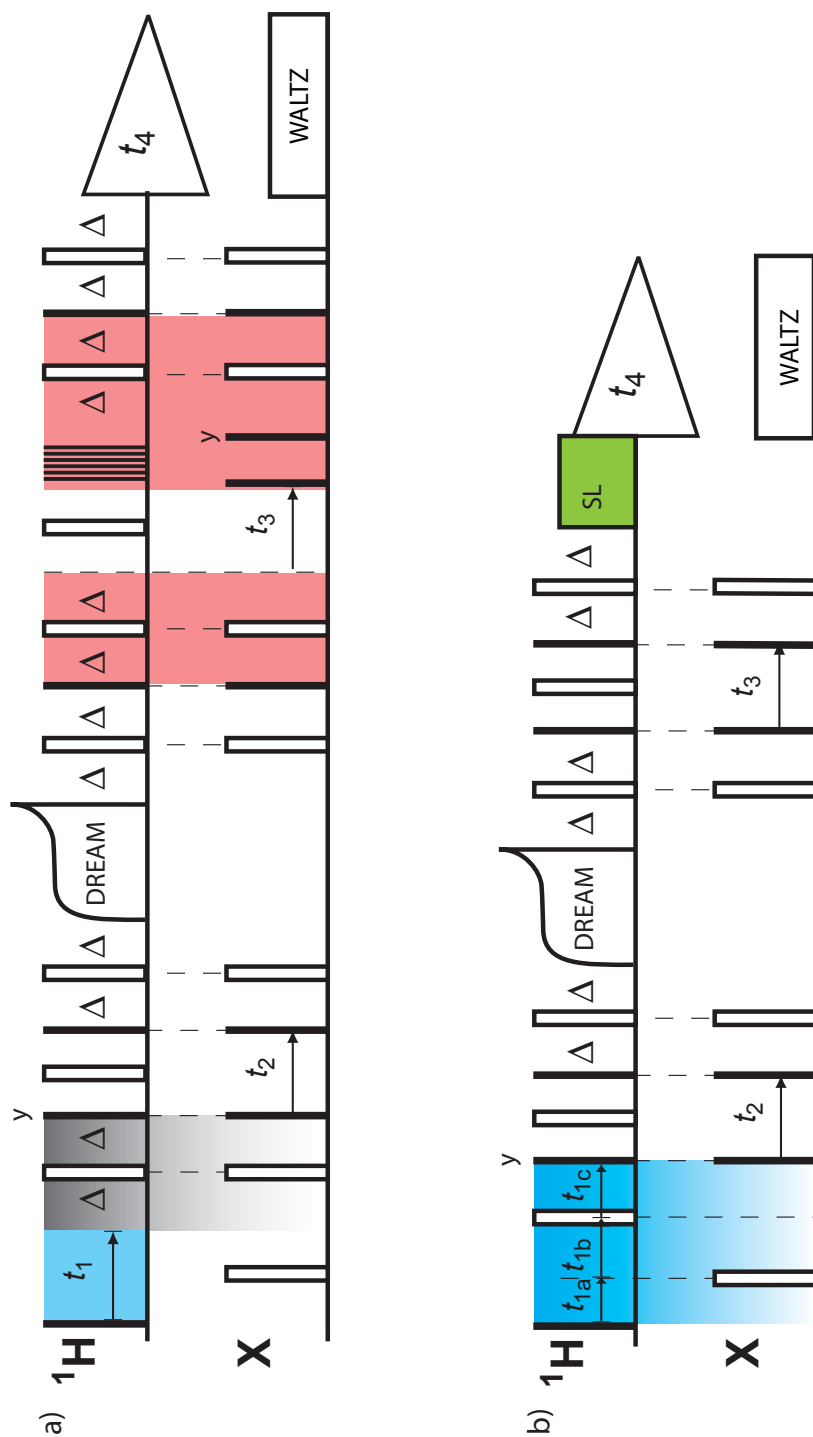
**Figure 8.7.** Pulse sequence for the 4D methyl-methyl DREAM. The sequence used for the measurement on ubiquitin (a) can be shortened by introducing a semi-constant-time $t_1$ evolution[146] and by replacing the proton saturation pulses that require an refocused INEPT by a spin-lock purge pulse[144] which can be integrated into the INEPT building block. Filled and open bars denote 90° and 180° pulses, respectively. All unlabeled pulses were applied with x phase. The INEPT delay $\Delta = \frac{1}{4 \cdot J_{IS}}$ was set to 1.65 ms for C-H transfer, $t_{1a}$, $t_{1b}$ and $t_{1c}$ are defined in section 8.1.6. For optimal $^{13}$C resolution of CHD$_2$ labeled methyl groups, $^2$H decoupling has be applied during the $^{13}$C chemical shift evolution. If necessary a auxiliary $^2$H coil can be installed for this purpose (see chapter 3).

## 8.1.7. Structure Calculation from 2D CHHC spectrum of Ubiquitin

Automatic structure calculation from 2D [$^{13}$C;$^{13}$C] spectra of [U-$^{13}$C] labeled proteins is difficult due high ambiguity in the cross-peak assignment and peak overlap that prevents identification of some weak peaks as discussed in chapter 2. To have a good comparison with structure calculations based on $^{1}$H–$^{1}$H restraints as described in chapter 4 we recorded a CHHC spectrum on the same spectrometer (850 MHz) and for the same time (3 days) as we did for 4D $^{1}$H–$^{1}$H methyl–methyl DREAM spectrum. Differences were that the 4D DREAM was recorded on $\approx$ 2 mg of perdeuterated and selectively ILV methyl $^{1}$H labeled ubiquitin in a 1.3 mm rotor at 55 kHz MAS whereas the 2D CHHC was measured on $\approx$ 20 mg of [U-$^{13}$C] labeled ubiquitin in a 3.2 mm rotor at 19 kHz MAS. In both experiments the long-range magnetization transfer took place between protons. Measuring $^{1}$H–$^{1}$H contacts is advantageous compared to $^{13}$C–$^{13}$C contacts to derive meaningful distance restraints, since even for a short distance cutoff such as 3.5 Å 40% of expected contacts are long- or medium range while for $^{13}$C–$^{13}$C contacts 95% correspond to intraresidue one- or two bond contacts[42]. The CHHC spectrum shown in Figure 8.8 was automatically peakpicked and the peak list was used for an automatic structure calculation in CYANA[43, 70] in combination with TALOS+[38] torsional angle restraints. Peaks were converted to ambiguous distance restraints by applying a tolerance window of 0.3 ppm in both dimensions and using network anchoring to reduce assignment possibilities. Upper distance limits were set to 4.5 Å and applied between the protons directly bound to the carbons that were frequency encoded or detected. Initially 905 distance restraints were derived from the 2245 picked peaks, with 147 of those restraints being long-range. On average restraints had 9 assignment possibilities. To reduce the impact of imperfections in the peak- and shift lists, restraint combination (see section 1.6 for details) was used in the first cycles of the structure calculation protocol. Structure calculation runs with different random seeds yielded cycle 1 bundles with an RMSD within the bundle between 1.8 and 3.4 Å. In the course of the calculation the bundle RMSD dropped to 0.3 Å. The RMSD of the final structure bundle is a bad indicator for the quality of the structure as it only displays the structural variability under the assumption that ambiguities were correctly resolved using intermediary structures. A bundle created from the best structures of independent calculations is a better indicator. Such a bundle of the best structures from different calculations is shown in Figure 8.9. The RMSD within the bundle of 1.9 Å is very similar to the RMSD between the bundle and the reference structure (2.1 Å).
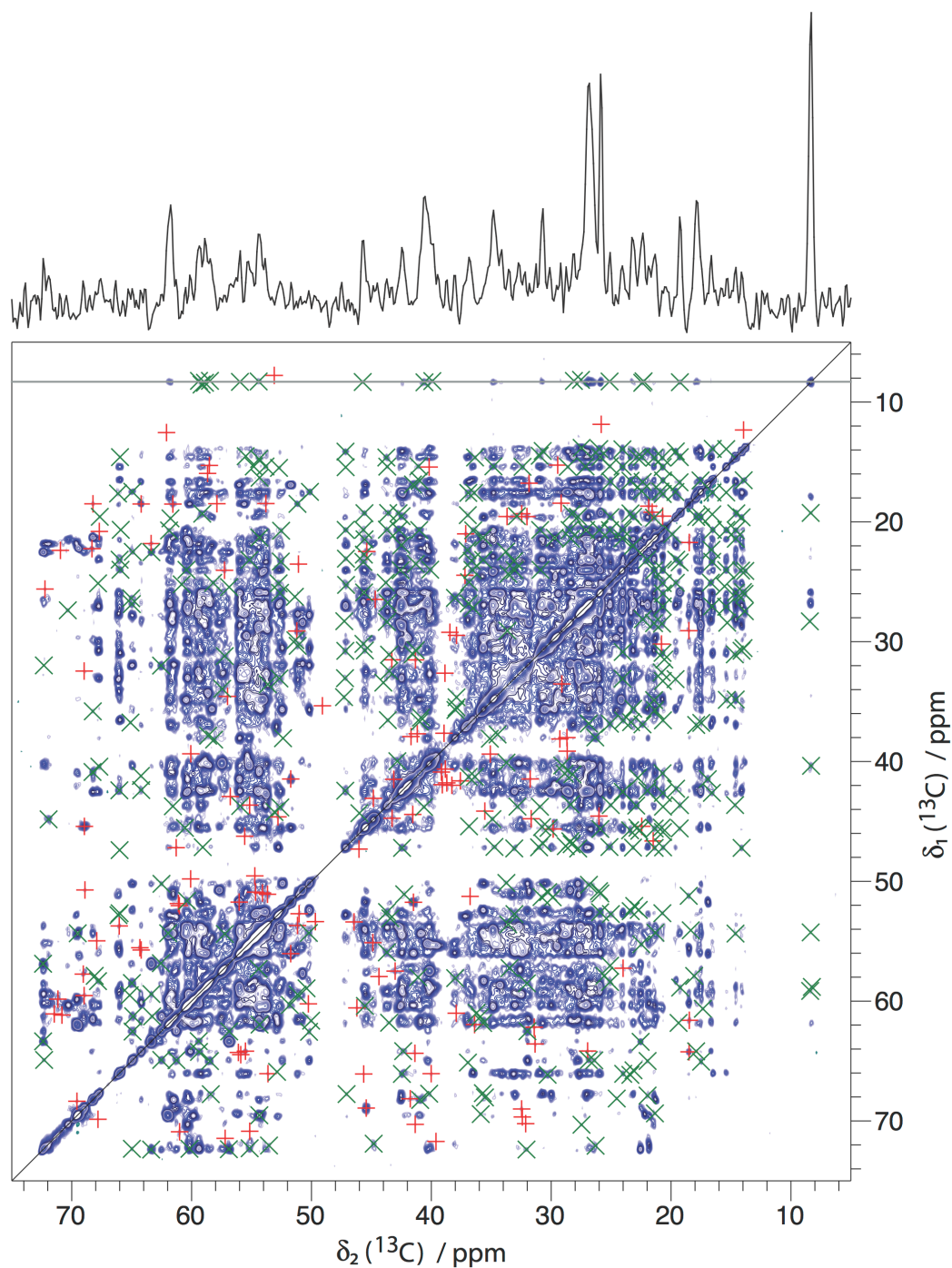
**Figure 8.8.** 2D $^{13}$C-$^{13}$C CHHC spectrum of [U-$^{13}$C] labeled ubiquitin micro crystals. The spectrum was recorded for 3 days at 20.0 T and with a mixing time of 600 $\mu$s. The trace shown on top was extracted at the position of Ile 23 C$\delta_1$. A total of 2245 peaks were automatically picked and used for structure calculation. The 456 peaks identified as long range (separation in sequence $> 4$ resides) are shown as green 'x' and the 162 peaks that could not be explained and were neglected in the final cycle of the CYANA calculation are shown as red '+'.
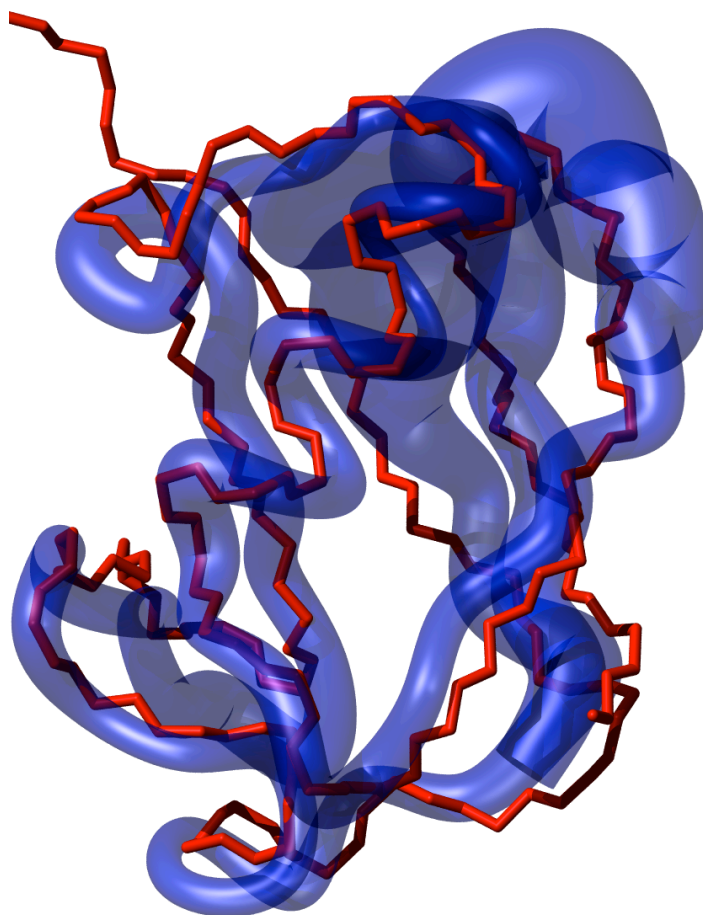
**Figure 8.9.** Structure bundle of ubiquitin in sausage representation (blue) obtained from 10 independent structure calculations from the CHHC spectrum shown in Figure 8.8. The structures are aligned to the reference structure (1UBQ, red). TALOS+ dihedral restraints were used in addition to distance restraints automatically derived from the peak list. The backbone RMSD of the bundle of the best structures from 10 independent structure calculations is 1.9±1.1 Å and the average bias RMSD to reference structure is 2.1 Å.

## 8.1.8. Proton-detected spectra of ILV methyl labeled HET-s (1-227)

Structure determination from proton-detected $^1$H-$^1$H DREAM spectra as described in chapter 4 for ubiquitin should extend the size of proteins amendable to structure calculation from solid-state NMR spectra. We measured HSQC and $^1$H-$^1$H DREAM spectra of HET-s (1-227), a larger model system for which the structure is also known from x-ray diffraction. Since HET-s (1-227) is mostly $\alpha$-helical, we only expect 26 H$^N$–H$^N$ longrange ($> 5$ residue separation) contacts to be observable for a distance cutoff of 6 Å. More than half of those involve the short $\beta$-sheet stretch of residues 27-32. Therefore measuring methyl–methyl contacts is expected to be important for determining the structure. Figure 8.10 shows contacts of less than 6 Å between methyl groups of isoleucine, valine and leucine. Figure 8.11 shows the HSQC of ILV labeled HET-s (1-227) produced from $\alpha$-ketoacid precursors leading that give random labeling of one methyl group of valine and leucine and for comparison the HSQC spectrum from a sample produced with a precursor for stereospecific methyl labeling[21]. Figure 8.12 shows slices from a 3D HCH and a 3D HCC $^1$H–$^1$H DREAM spectrum that could be used to drive distance restraints
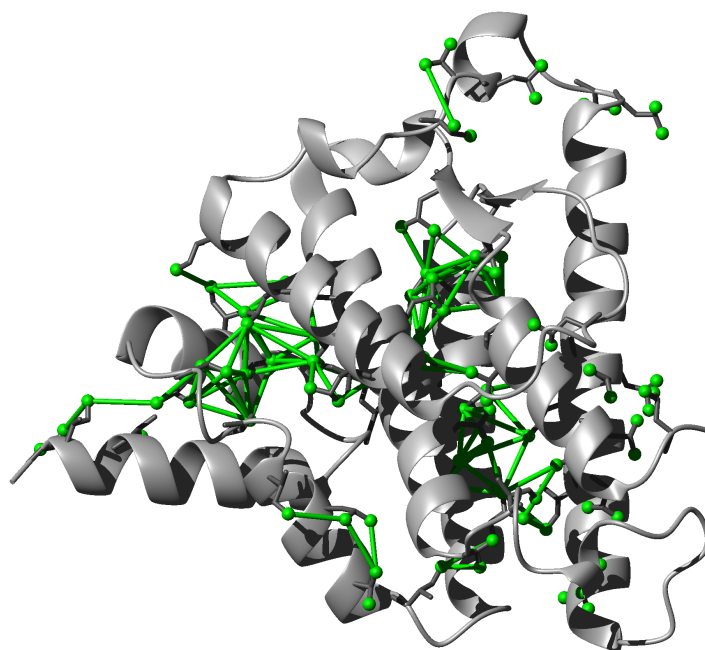


**Figure 8.10.** X-ray structure of HETs(1-227) with ILV methyl groups shown as green spheres. Contacts between methyl groups of less than 6 Å that are expected to be visible in $^1$H-$^1$H DREAM spectra are shown as green lines.
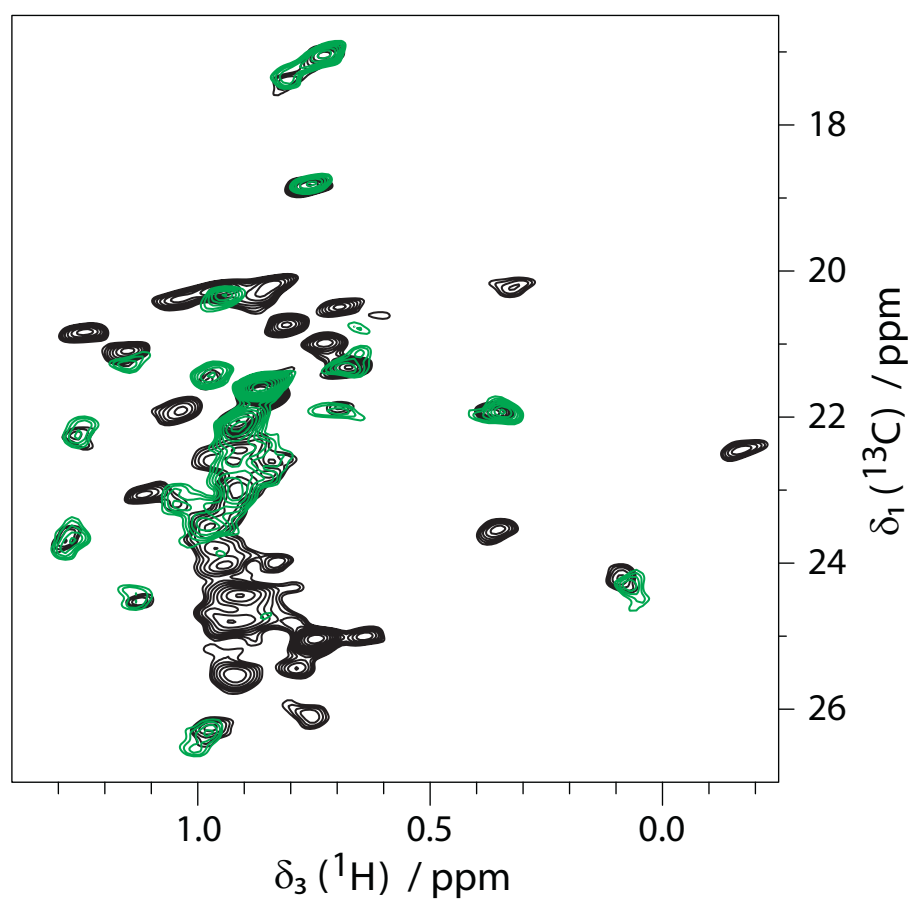
**Figure 8.11.** 2D [$^1$H;$^{13}$C]-HSQC spectrum of [U-$^2$H$^{15}$N$^{13}$C,ILV $^{13}$CHD$_2$] HET-s(1-227) crystallized in [U-$^2$H] PEG-4000 and D$_2$O. In the black spectrum one methyl group is labeled as $^{13}$CHD$_2$ and one unlabeled ($^{12}$CD$_3$) in a random way. The spectrum in green is from a sample with stereospecific methyl labeling were for valine and leucine only the $\gamma_2$ or $\delta_2$ methyl groups are labeled.
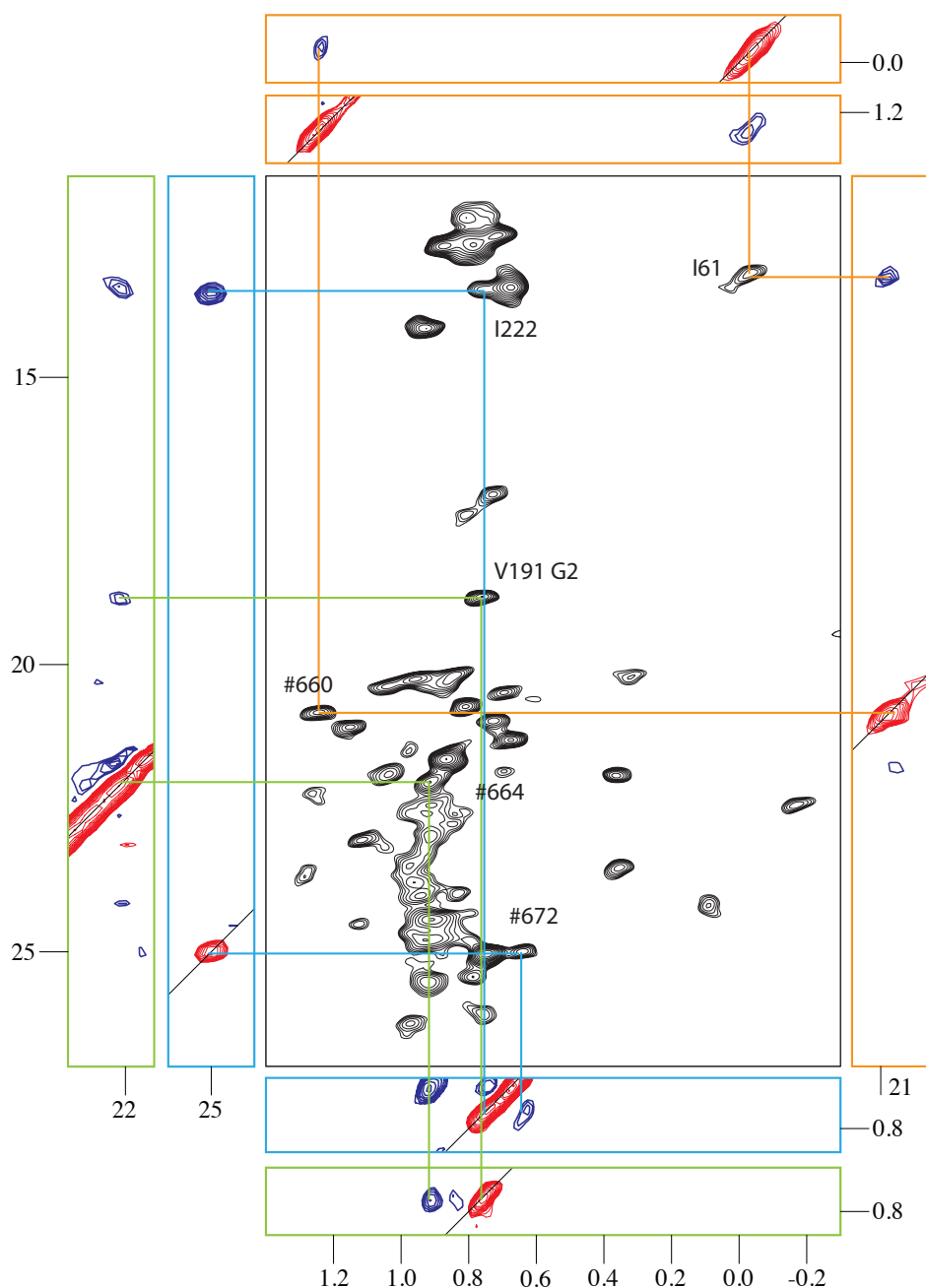
**Figure 8.12.** Strips from 3D H*methyl*–H*methyl* DREAM spectra of ILV-labelled HETs(1-227). Red peaks are positive diagonal peaks and blue peaks are negative DREAM cross-peaks. **Vertical strips** are $^{13}$C–$^{13}$C planes from the 3D CCH-DREAM spectrum. The $^1$H frequency is that of the peak in the HSQC which has a line to the diagonal peak. **Horizontal strips** are $^1$H–$^1$H planes from the 3D HCH-DREAM spectrum . The $^{13}$C frequency is that of the peak in the HSQC which has a line to the diagonal peak. Peaks in strips of the same color (green, blue and orange border) were used to derive distance restraints between the peaks indicated in the 2D HSQC spectrum.

## 8.2. Part II: Application to A$\beta$

### 8.2.1. A$\beta$1-40 E22$\Delta$

**Fusion construct for A$\beta$1-40 E22$\Delta$ expression**

The fusion construct for expression under the T7 promoter/lac operator in *E. coli* BL21 (DE3) has the sequence:

(H)$_6$ *GS* (NANP)$_{19}$ *RS ENLYFQ*

DAEFRHDSG YEVHHQKLVF FADVGSNKGAI IGLMVGGVV

It consists of a His-tag, linker, solubility tag (NANP)$_{19}$, linker, TEV recognition site and the sequence of A$\beta$1-40 E22$\Delta$ . Although the solubility tag is removed and the peptide purified by HPLC traces of the full lenght fusion protein have been detected in INEPT spectra of A$\beta$ fibrils (see Figure 8.13).
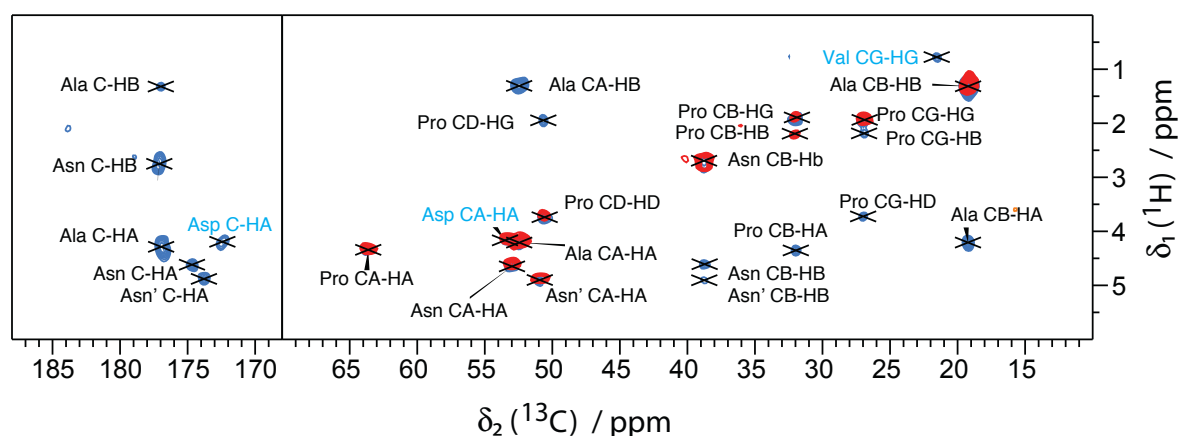


**Figure 8.13.** 2D $^1$H-$^{13}$C INEPT spectrum (red) and 2D $^1$H-$^{13}$C INEPT-CCtobsy spectrum (blue) of A$\beta$1-40 E22$\Delta$ fibrils. Amino-acid type specific assignments were done by comparison with predicted random coil-values [147]. The peaks labeled in black can be explained by signals from the NANP solubility tag in the fusion protein (see section 8.2.1 that is repeated 19 times. Only the peaks with blue labels are from the A$\beta$ peptide.

**Purification of A$\beta$1-40 E22$\Delta$**

The purification of A$\beta$1-40 E22$\Delta$ has been described by Oxana Ovchinnikova et al. in [115] and is repeated here for completeness:

The cells were harvested and resuspended in 20 mM $H_3PO_4$–NaOH, 6 M guanidinium chloride, and 0.3 M NaCl (pH 8.0; buffer A), and the suspension was ultracentrifuged after incubation at 4°C for 90 min. The supernatant was diluted with buffer A containing 500 mM imidazole–HCl to a concentration of 10 mM imidazole and loaded on a $Ni^{2+}$-NTA agarose column. After a wash with buffer A containing 10 mM imidazole–HCl, the fusion proteins were eluted with buffer A containing 250 mM imidazole–HCl and further purified via reversed-phase high-performance liquid chromatography at 80°C in aqueous acetonitrile (ACN) containing 0.1% trifluoroacetic acid on a Zorbax SB300 C8 column 9.4 x 250 mm, 5 $\mu$m semi-preparative column (Agilent). The fusion proteins were eluted isocratically at ACN concentrations of 28–30.5% and lyophilized. Cleavage of the fusion proteins (100 $\mu$M) with 7.5 $\mu$M tobacco etch virus protease was performed in 10 mM Tris–HCl (pH 8.0), 0.5 mM ethylenediaminetetraacetic acid, and 1 mM DTT (buffer B) for 1 h at room temperature, followed by incubation at 4 °C overnight. The cleaved A$\beta$ peptides precipitated during the cleavage reactions. They were pelleted by centrifugation (4500g, 20 min, 4 °C), dissolved in 8 M guanidinium chloride–HCl (pH 2.0) and afterwards diluted with water to 4-5 M GdHCl (peptide concentration ca 0.5 mM). They were purified via reversed-phase high-performance liquid chromatography, on a ZORBAX column as described above. Sample volumes of 3-4 ml were loaded with 10% ACN, 0.1% TFA and a flow rate of 4 ml/min. A$\beta$1-40 E22$\Delta$ was eluted at 28% ACN (the chromatogram is shown in Figure 8.15). The eluted peptide was aliquoted in Protein LoBind Eppendorf tubes (Vaudaux-Eppendorf), lyophilized, and stored at -80 °C.
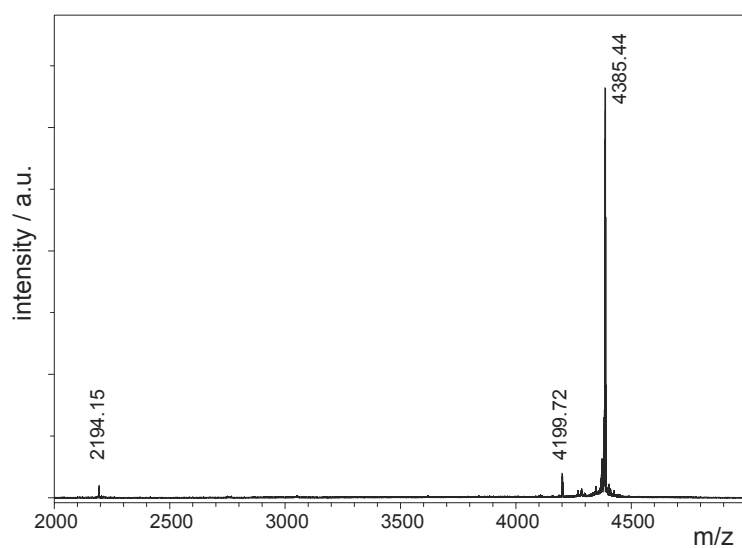
**Figure 8.14.** MALDI-TOF mass spectrum of [U-$^{13}$C] A$\beta$1-40 E22Δ using sinapinic acid as matrix. The expected mass of the isotope labelled peptide is 4389 Da.
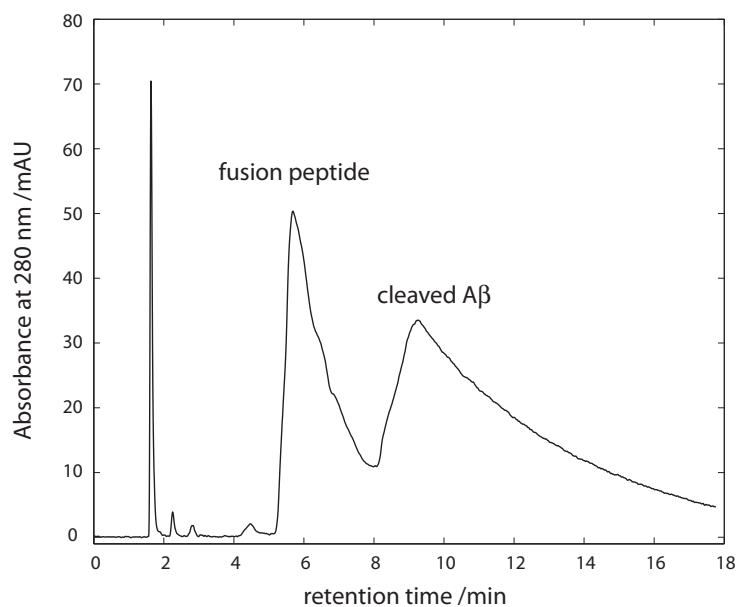


**Figure 8.15.** Chromatogram of the purification of A$\beta$1-40 E22Δ peptide after TEV cleavage using reverse phase HPLC on a Zorbax SB300 C8 column.
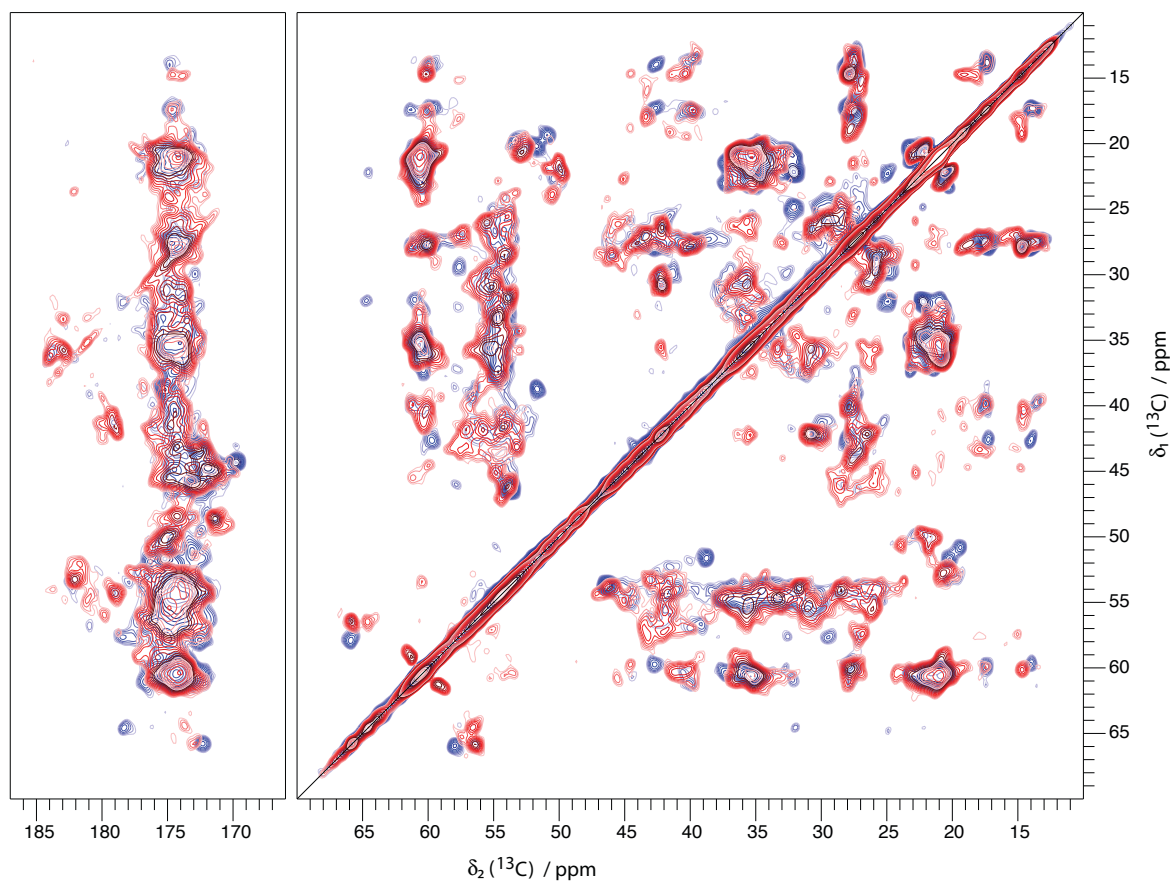
## 8.2.2. A$\beta$1-42 wt



**Figure 8.16.** 50ms PDSD spectra of A$\beta$1-42 wt recorded at 600 MHz. The spectrum in red is from the first batch, fibrillized at a concentration of 1.5 mg/ml within 20 h (fast-fibrillized) and the spectrum in blue is from a fibrillization with a concentration of 0.2mg/ml within 8 weeks (slow-fibrillized). The fibrillization for both samples was performed in 10 mM Trip pH 7.4, 300 mM NaCl and 0.5 mM EDTA at 37 $^\circ C$ and with slow agitation. There are some strong differences in the peak positions, notably a serine with an $\alpha$-helical shift in the fast-fibrillized and a valine with downfield C$\alpha$ shift in the slow-fibrillized sample.
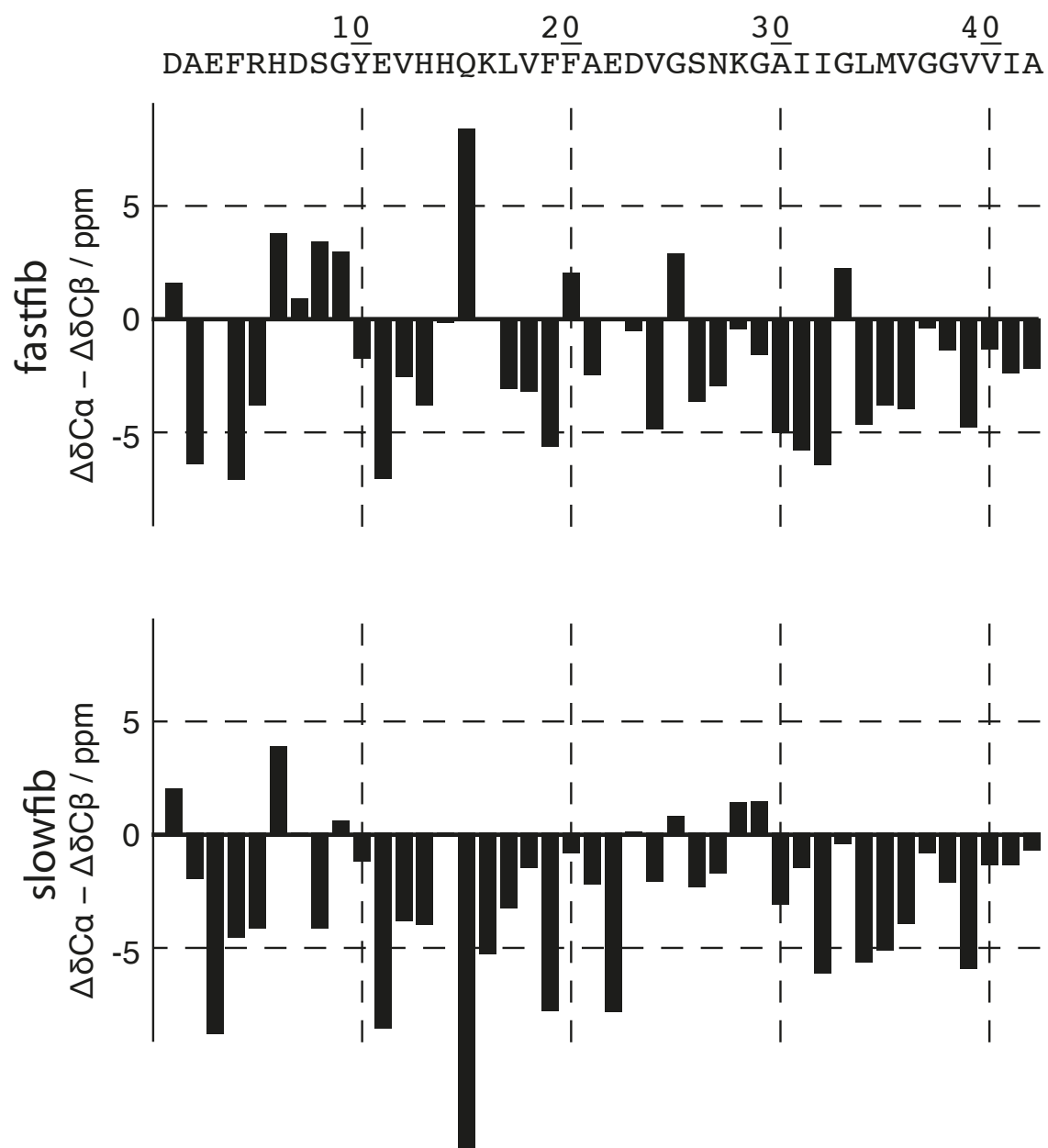
**Figure 8.17.** Secondary chemical shifts[121] of the two Aβ1-42 sample preparations. Three negative secondary shifts in a row are indicative for β-sheet structure. Chemical shift assignments were done by Alice Soragni and are very preliminary. The sample quality and quantity was insufficient for a reliable and complete sequential assignment.

# Bibliography

[1] Ernst, R. R., G. Bodenhausen, and A. Wokaun, 1987. Principles of nuclear magnetic resonance in one and two dimensions. International series of monographs on chemistry. Oxford University Press.

[2] Pervushin, K., R. Riek, G. Wider, and K. Wuthrich, 1997. Attenuated T-2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Nat. Acad. Sci.* 94:12366.

[3] Fiaux, J., E. B. Bertelsen, A. L. Horwich, and K. Wüthrich, 2002. NMR analysis of a 900K GroEL GroES complex. *Nature* 418:207–211.

[4] Lowe, I. J., 1959. Free Induction Decays of Rotating Solids. *Phys. Rev. Lett.* 2:285.

[5] Andrew, E. R., A. Bradbury, and R. G. Eades, 1958. Nuclear Magnetic Resonance Spectra from a Crystal rotated at High Speed. *Nature* 182:1659.

[6] Bennett, A., C. Rienstra, and M. Auger, 1995. Heteronuclear decoupling in rotating solids. *The Journal of Chemical Physics* 103:6951.

[7] Fung, B. M., A. K. Khitrin, and K. Ermolaev, 2000. An improved broadband decoupling sequence for liquid crystals and solids. *J Magn Reson* 142:97–101.

[8] Detken, A., E. Hardy, and M. Ernst, 2002. Simple and efficient decoupling in magic-angle spinning solid-state NMR: the XiX scheme. *Chem. Phys. Lett.* 356:298.

[9] Ernst, M., A. Samoson, and B. H. Meier, 2003. Low-power XiX decoupling in MAS NMR experiments. *J. Magn. Reson.* 163:332.

[10] Kotecha, M., N. P. Wickramasinghe, and Y. Ishii, 2008. Efficient low-power heteronuclear decoupling in13C high-resolution solid-state NMR under fast magic angle spinning. *Magn. Reson. Chem.* 45:S221.

[11] Breukels, V., A. Konijnenberg, S. M. Nabuurs, J. F. Doreleijers, N. V. Kovalevskaya, and G. W. Vuister, 2001. Overview on the Use of NMR to Examine Protein Structure: Current Protocols in Protein Science. John Wiley & Sons, Inc., Hoboken, NJ, USA.

[12] De Laeter, J., J. Böhlke, P. De Bièvre, H. Hidaka, H. Peiser, K. Rosman, and P. Taylor, 2003. Atomic weights of the elements: review 2000. *Pure Appl. Chem* 75:683–800.

[13] Leiting, B., F. Marsilio, and J. F. O'Connell, 1998. Predictable deuteration of recombinant proteins expressed in Escherichia coli. *Anal. Biochem.* 265:351–355.

[14] Hoogstraten, C. G., and J. E. Johnson, 2008. Metabolic labeling: Taking advantage of bacterial pathways to prepare spectroscopically useful isotope patterns in proteins and nucleic acids. *Concepts Magn. Reson. A* 32A:34.

[15] LeMaster, D., and D. Kushlan, 1996. Dynamical mapping of E. coli thioredoxin via 13C NMR relaxation analysis. *J. Am. Chem. Soc.* 118:9255.

[16] Castellani, F., B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, and H. Oschkinat, 2002. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420:98.

[17] Lundström, P., K. Teilum, T. Carstensen, I. Bezsonova, S. Wiesner, D. F. Hansen, T. L. Religa, M. Akke, and L. E. Kay, 2007. Fractional 13C enrichment of isolated carbons using [1-$^{13}$C]- or [2-$^{13}$C]-glucose facilitates the accurate measurement of dynamics at backbone C$\alpha$ and side-chain methyl positions in proteins. *J. Biomol. NMR* 38:199–212.

[18] Loquet, A., K. Giller, S. Becker, and A. Lange, 2010. Supramolecular interactions probed by 13C-13C solid-state NMR spectroscopy. *J. Am. Chem. Soc* 132:15164–15166.

[19] Goto, N. K., K. H. Gardner, G. A. Mueller, R. C. Willis, and L. E. Kay, 1999. A robust and cost-effective method for the production of Val, Leu, Ile ($\delta$ 1) methyl-protonated N-15-, C-13-, H-2-labeled proteins. *J. Biomol. NMR* 13:369.

[20] Tugarinov, V., V. Kanelis, and L. E. Kay, 2006. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat. Protocols* 1:749.

[21] Gans, P., O. Hamelin, R. Sounier, I. Ayala, M. A. Dura, C. D. Amero, M. Noirclerc-Savoye, B. Franzetti, M. J. Plevin, and J. Boisbouvier, 2010. Stereospecific Isotopic Labeling of Methyl Groups for NMR Spectroscopic Studies of High-Molecular-Weight Proteins. *Angew. Chem. Int. Ed.* 49:1958.

[22] Morris, G., and R. Freeman, 1979. Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* 101:760.

[23] Hartmann, S., and E. Hahn, 1962. Nuclear double resonance in the rotating frame. *Phys. Rev.* 128:2042.

[24] Hediger, S., and B. Meier, 1995. Adiabatic passage Hartmann-Hahn cross polarization in NMR under magic angle sample spinning. *Chem. Phys. Lett.* 240:449.

[25] Verel, R., M. Ernst, and B. H. Meier, 2001. Adiabatic dipolar recoupling in solid-state NMR: The DREAM scheme. *J. Magn. Reson.* 150:81.

[26] Bayro, M. J., M. Huber, R. Ramachandran, T. C. Davenport, B. H. Meier, M. Ernst, and R. G. Griffin, 2009. Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J. Chem. Phys.* 130:114506.

[27] Brinkmann, A., J. Schmedt auf der Günne, and M. H. Levitt, 2002. Homonuclear Zero-Quantum Recoupling in Fast Magic-Angle Spinning Nuclear Magnetic Resonance. *J. Magn. Reson.* 156:79–96.

[28] Bennett, A. E., C. M. Rienstra, J. M. Griffiths, W. Zhen, P. T. Lansbury, and R. G. Griffin, 1998. Homonuclear radio frequency-driven recoupling in rotating solids. *J. Chem. Phys.* 108:9463.

[29] Ishii, Y., 2001. $^{13}$C–$^{13}$C dipolar recoupling under very fast magic angle spinning in solid-state nuclear magnetic resonance: Applications to distance measurements, spectral assignments, and high-throughput secondary-structure determination. *J. Chem. Phys.* 114:8473.

[30] Raleigh, D., M. Levitt, and R. Griffin, 1988. Rotational resonance in solid state NMR. *Chem. Phys. Lett.* 146:71–76.

[31] De Paëpe, G., J. R. Lewandowski, A. Loquet, A. Böckmann, and R. G. Griffin, 2008. Proton assisted recoupling and protein structure determination. *J. Chem. Phys.* 129:245101.

[32] Schuetz, A., C. Wasmer, B. Habenstein, R. Verel, J. Greenwald, R. Riek, A. Bockmann, 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.

[33] Habenstein, B., C. Wasmer, L. Bousset, Y. Sourigues, A. Schuetz, A. Loquet, B. H. Meier, R. Melki, and A. Böckmann, 2011. Extensive de novo solid-state NMR assignments of the 33 kDa C-terminal domain of the Ure2 prion. *J. Biomol. NMR* 51:235.

[34] Linser, R., U. Fink, and B. Reif, 2008. Proton-detected scalar coupling based assignment strategies in MAS solid-state NMR spectroscopy applied to perdeuterated proteins. *J. Magn. Reson.* 193:89.

[35] Pauli, J., M. Baldus, B. van Rossum, H. de Groot, and H. Oschkinat, 2001. Backbone and Side–Chain $^{13}$C and $^{15}$N Signal Assignments of the $\alpha$–Spectrin SH3 Domain by Magic Angle Spinning Solid–State NMR at 17.6 Tesla. *ChemBioChem* 2:272–281.

[36] Kirkpatrick, S., C. Gelatt, and M. Vecchi, 1983. Optimization by simulated annealing. *Science* 220:671.

[37] Jain, A., N. Vaidehi, and G. Rodriguez, 1993. A Fast Recursive Algorithm for Molecular-Dynamics Simulation. *J. Comput. Phys.* 106:258–268.

[38] Shen, Y., F. Delaglio, G. Cornilescu, and A. Bax, 2009. TALOS plus : a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* 44:213.

[39] Cheung, M.-S., M. L. Maguire, T. J. Stevens, and R. W. Broadhurst, 2010. DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J. Magn. Reson.* 202:223.

[40] Manolikas, T., T. Herrmann, and B. H. Meier, 2008. Protein structure determination from C-13 spin-diffusion solid-state NMR spectroscopy. *J. Am. Chem. Soc.* 130:3959.

[41] Baldus, M., 2002. Correlation experiments for assignment and structure elucidation of immobilized polypeptides under magic angle spinning. *Prog. NMR Spectroscopy* 41:1.

[42] Loquet, A., B. Bardiaux, C. Gardiennet, C. Blanchet, M. Baldus, M. Nilges, T. Malliavin, and A. Bockmann, 2008. 3D structure determination of the Crh protein from highly ambiguous solid-state NMR restraints. *J. Am. Chem. Soc.* 130:3579.

[43] Herrmann, T., P. Güntert, and K. Wüthrich, 2002. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* 319:209.

[44] Rieping, W., M. Habeck, B. Bardiaux, A. Bernard, T. E. Malliavin, and M. Nilges, 2007. ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* 23:381.

[45] Güntert, P., 2011. Protein NMR Spectroscopy: Practical Techniques and Applications: Practical Techniques and Applications. Wiley.

[46] 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.

[47] Hiller, S., R. G. Garces, T. J. Malia, V. Y. Orekhov, M. Colombini, and G. Wagner, 2008. Solution Structure of the Integral Human Membrane Protein VDAC-1 in Detergent Micelles. *Science* 321:1206–1210.

[48] Riek, R., S. Hornemann, G. Wider, M. Billeter, R. Glockshuber, and K. Wüthrich, 1996. NMR structure of the mouse prion protein domain PrP(121-231). *Nature* 382:180–182.

[49] Nielsen, P. R., D. Nietlispach, H. R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A. G. Murzin, N. V. Murzina, and E. D. Laue, 2002. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* 416:103–107.

[50] Berman, H., K. Henrick, and H. Nakamura, 2003. Announcing the worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* 10:980.

[51] Böckmann, A., 2008. 3D protein structures by solid-state NMR spectroscopy: Ready for high resolution. *Angew. Chem. Int. Ed.* 47:6110.

[52] Van Melckebeke, H., C. Wasmer, A. Lange, E. Ab, A. Loquet, A. Bockmann, and B. H. Meier, 2010. Atomic-Resolution Three-Dimensional Structure of HET-s(218-289) Amyloid Fibrils by Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 39:13765.

[53] 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 beta -amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Nat. Acad. Sci.* 99:16742.

[54] Debelouchina, G. T., G. W. Platt, M. J. Bayro, S. E. Radford, and R. G. Griffin, 2010. Intermolecular Alignment in $\beta$(2)-Microglobulin Amyloid Fibrils. *J. Am. Chem. Soc.* 132:17077.

[55] Schütz, A. K., A. Soragni, S. Hornemann, A. Aguzzi, M. Ernst, A. Böckmann, and B. H. Meier, 2011. The Amyloid-Congo Red Interface at Atomic Resolution. *Angew. Chem. Int. Ed.* 50:5956.

[56] Cady, S. D., K. Schmidt-Rohr, J. Wang, C. S. Soto, W. F. Degrado, and M. Hong, 2010. Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* 463:689.

[57] Jehle, S., P. Rajagopal, B. Bardiaux, S. Markovic, R. Kühne, J. R. Stout, V. A. Higman, R. E. Klevit, B.-J. van Rossum, and H. Oschkinat, 2010. Solid-state NMR and SAXS studies provide a structural basis for the activation of alphaB-crystallin oligomers. *Nat. Struct. Mol. Biol.* 17:1037.

[58] Mainz, A., S. Jehle, B. J. van Rossum, H. Oschkinat, and B. Reif, 2009. Large protein complexes with extreme rotational correlation times investigated in solution by magic-angle-spinning NMR spectroscopy. *J. Am. Chem. Soc.* 131:15968.

[59] Jeener, J., B. H. Meier, P. Bachmann, and R. R. Ernst, 1979. Investigation of Exchange Processes by 2-Dimensional Nmr-Spectroscopy. *J. Chem. Phys.* 71:4546.

[60] Sørensen, O., 1990. Aspects and prospects of multidimensional time-domain spectroscopy. *J. Magn. Reson.* 89:210.

[61] Chevelkov, V., K. Rehbein, A. Diehl, and B. Reif, 2006. Ultrahigh resolution in proton solid-state NMR spectroscopy at high levels of deuteration. *Angew. Chem. Int. Ed.* 45:3878.

[62] Agarwal, V., A. Diehl, N. Skrynnikov, and B. Reif, 2006. High resolution H-1 detected H-1,C-13 correlation spectra in MAS solid-state NMR using deuterated proteins with selective H-1,H-2 isotopic labeling of methyl groups. *J. Am. Chem. Soc.* 128:12620.

[63] Akbey, U., S. Lange, W. T. Franks, R. Linser, K. Rehbein, A. Diehl, B. J. van Rossum, B. Reif, and H. Oschkinat, 2010. Optimum levels of exchangeable protons in perdeuterated proteins for proton detection in MAS solid-state NMR spectroscopy. *J. Biomol. NMR* 46:67.

[64] Lewandowski, J. R., J.-N. Dumez, U. 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. *J. Phys. Chem. Lett.* 2:2205–2211.

[65] Asami, S., P. Schmieder, and B. Reif, 2010. High resolution 1H-detected solid-state NMR spectroscopy of protein aliphatic resonances: access to tertiary structure information. *J. Am. Chem. Soc.* 132:15133.

[66] Takegoshi, K., S. Nakamura, and T. Terao, 2001. C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* 344:631–637.

[67] Rovnyak, D., J. Hoch, A. Stern, and G. Wagner, 2004. Resolution and sensitivity of high field nuclear magnetic resonance spectroscopy. *J. Biomol. NMR* 30:1–10.

[68] Hiller, S., I. Ibraghimov, G. Wagner, and V. Y. Orekhov, 2009. Coupled Decomposition of Four-Dimensional NOESY Spectra. *J. Am. Chem. Soc.* 131:12970.

[69] Hyberts, S. G., D. P. Frueh, H. Arthanari, and G. Wagner, 2009. FM reconstruction of non-uniformly sampled protein NMR data at higher dimensions and optimization by distillation. *J. Biomol. NMR* 45:283.

[70] Güntert, P., C. Mumenthaler, and K. Wüthrich, 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* 273:283.

[71] Schanda, P., M. Huber, J. Boisbouvier, B. H. Meier, and M. Ernst, 2011. Solid-state NMR measurements of asymmetric dipolar couplings provide insight into protein side chain motion. *Angew. Chem. Int. Ed.* 50:11005.

[72] Tugarinov, V., W.-Y. Choy, V. Y. Orekhov, and L. E. Kay, 2005. Solution NMR-derived global fold of a monomeric 82-kDa enzyme. *Proc. Nat. Acad. Sci.* 102:622.

[73] Huber, M., S. Hiller, P. Schanda, M. Ernst, A. Böckmann, R. Verel, and B. H. Meier, 2011. A Proton-Detected 4D Solid-State NMR Experiment for Protein Structure Determination. *ChemPhysChem* 12:915.

[74] Reif, B., Y. Xue, V. Agarwal, M. S. Pavlova, M. Hologne, A. Diehl, Y. E. Ryabov, and N. R. Skrynnikov, 2006. Protein side-chain dynamics observed by solution- and solid-state NMR: Comparative analysis of methyl H-2 relaxation data. *J. Am. Chem. Soc.* 128:12354.

[75] Shaka, A. J., and J. Keeler, 1987. Broadband spin decoupling in isotropic liquids. *Prog. NMR Spectrosc.* 19:47.

[76] Mispelter, J., M. Lupu, and A. Briguet, 2006. NMR probeheads for biophysical and biomedical experiments : theoretical principles and practical guidelines. Imperial College Press, London.

[77] Elliott, R. S., I. Antennas, and P. Society., 1993. Electromagnetics : history, theory, and applications. IEEE Press series on electromagnetic waves. IEEE Press, Piscataway, NJ.

[78] Alei Jr, M., and W. Wageman, 1978. Gas phase C chemical shifts and coupling constants in the deuteromethanes. *J. Chem. Phys.* 68:783.

[79] Wasmer, C., A. Schütz, A. Loquet, C. Buhtz, J. Greenwald, R. Riek, A. Böckmann, and B. Meier, 2009. The Molecular Organization of the Fungal Prion HET-s in Its Amyloid Form. *J. Mol. Biol.* 394:119.

[80] Jehle, S., B. van Rossum, J. R. Stout, S. M. Noguchi, K. Falber, K. Rehbein, H. Oschkinat, R. E. Klevit, and P. Rajagopal, 2009. alphaB-crystallin: a hybrid solid-state/solution-state NMR investigation reveals structural aspects of the heterogeneous oligomer. *J. Mol. Biol.* 385:1481.

[81] Sperling, L. J., D. A. Berthold, T. L. Sasser, V. Jeisy-Scott, and C. M. Rienstra, 2010. Assignment strategies for large proteins by magic-angle spinning NMR: the 21-kDa disulfide-bond-forming enzyme DsbA. *J. Mol. Biol.* 399:268.

[82] Lange, A., S. Becker, K. Seidel, K. Giller, O. Pongs, and M. Baldus, 2005. A concept for rapid protein-structure determination by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed.* 44:2089.

[83] Zech, S. G., A. J. Wand, and A. E. Mcdermott, 2005. Protein structure determination by high-resolution solid-state NMR spectroscopy: application to microcrystalline ubiquitin. *J. Am. Chem. Soc.* 127:8618.

[84] Li, S., Y. Zhang, and M. Hong, 2010. 3D 13C–13C–13C correlation NMR for de novo distance determination of solid proteins and application to a human $\alpha$-defensin. *J. Magn. Reson.* 202:203.

[85] Zhou, D. H., G. Shah, M. Cormos, C. Mullen, D. Sandoz, and C. M. Rienstra, 2007. Proton-detected solid-state NMR Spectroscopy of fully protonated proteins at 40 kHz magic-angle spinning. *J. Am. Chem. Soc.* 129:11791.

[86] Korukottu, J., R. Schneider, V. Vijayan, A. Lange, O. Pongs, S. Becker, M. Baldus, and M. Zweckstetter, 2008. High-resolution 3D structure determination of kaliotoxin by solid-state NMR spectroscopy. *PLoS ONE* 3:e2359.

[87] Paulson, E. K., C. R. Morcombe, V. Gaponenko, B. Dancheck, R. A. Byrd, and K. W. Zilm, 2003. Sensitive high resolution inverse detection NMR spectroscopy of proteins in the solid state. *J. Am. Chem. Soc.* 125:15831.

[88] Schanda, P., M. Huber, R. Verel, M. Ernst, and B. H. Meier, 2009. Direct Detection of (3h)J(NC ') Hydrogen-Bond Scalar Couplings in Proteins by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* 48:9322.

[89] Abragam, A., 1961. The Principles of Nuclear Magnetism. Clarendon Press, Oxford.

[90] Ollerenshaw, J. E., V. Tugarinov, and L. E. Kay, 2003. Methyl TROSY: explanation and experimental verification. *Magn. Reson. Chem.* 41:843.

[91] Sounier, R., L. Blanchard, Z. Wu, and J. Boisbouvier, 2007. High-accuracy distance measurement between remote methyls in specifically protonated proteins. *J. Am. Chem. Soc.* 129:472.

[92] Hoch, J. C., A. S. Stern, D. L. Donoho, and I. M. Johnstone, 1990. Maximum-Entropy Reconstruction of Complex (Phase-Sensitive) Spectra. *J. Magn. Reson.* 86:236.

[93] Chylla, R. A., and J. L. Markley, 1995. Theory and application of the maximum likelihood principle to NMR parameter estimation of multidimensional NMR data. *J. Biomol. NMR* 5:245.

[94] Schmieder, P., A. S. Stern, G. Wagner, and J. C. Hoch, 1997. Quantification of maximum-entropy spectrum reconstructions. *J. Magn. Reson.* 125:332.

[95] Hyberts, S. G., G. J. Heffron, N. G. Tarragona, K. Solanky, K. A. Edmonds, H. Luithardt, J. Fejzo, M. Chorev, H. Aktas, K. Colson, K. H. Falchuk, J. A. Halperin, and G. Wagner, 2007. Ultrahigh-resolution (1)H-(13)C HSQC spectra of metabolite mixtures using nonlinear sampling and forward maximum entropy reconstruction. *J. Am. Chem. Soc.* 129:5108.

[96] Tugarinov, V., L. E. Kay, I. Ibraghimov, and V. Y. Orekhov, 2005. High-resolution four-dimensional 1H-13C NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. *J. Am. Chem. Soc.* 127:2767.

[97] Orekhov, V. Y., I. V. Ibraghimov, and M. Billeter, 2001. MUNIN: a new approach to multi-dimensional NMR spectra interpretation. *J. Biomol. NMR* 20:49.

[98] Sun, B. Q., P. R. Costa, D. Kocisko, P. T. Lansbury, and R. G. Griffin, 1995. Internuclear Distance Measurements in Solid-State Nuclear-Magnetic-Resonance - Dipolar Recoupling Via Rotor Synchronized Spin Locking. *J. Chem. Phys.* 102:702.

[99] Luginbühl, P., P. Güntert, M. Billeter, and K. Wüthrich, 1996. The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR* 8:136.

[100] Tugarinov, V., and L. Kay, 2003. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J. Am. Chem. Soc* 125:13868.

[101] Otten, R., B. Chu, K. D. Krewulak, H. J. Vogel, and F. A. Mulder, 2010. Comprehensive and cost-effective NMR spectroscopy of methyl groups in large proteins. *J. Am. Chem. Soc.* 132:2952.

[102] Igumenova, T. I., A. E. McDermott, K. W. Zilm, R. W. Martin, E. K. Paulson, and A. J. Wand, 2004. Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* 126:6720.

[103] 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. *J. Biomol. NMR* 45:319.

[104] Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, 1995. Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes. *J. Biomol. NMR* 6:277.

[105] Güntert, P., V. Dötsch, G. Wider, and K. Wüthrich, 1992. Processing of Multidimensional Nmr Data with the New Software Prosa. *J. Biomol. NMR* 2:619.

[106] Zhou, D. H., and C. M. Rienstra, 2008. High-performance solvent suppression for proton detected solid-state NMR. *J. Magn. Reson.* 192:167.

[107] Shaka, A., J. Keeler, T. Frenkiel, and R. Freeman, 1983. An improved sequence for broadband decoupling: WALTZ-16. *J. Magn. Reson.* 52:335.

[108] Marion, D., M. Ikura, R. Tschudin, and A. Bax, 1989. Rapid Recording of 2d Nmr-Spectra without Phase Cycling - Application to the Study of Hydrogen-Exchange in Proteins. *J. Magn. Reson.* 85:393.

[109] Schubert, M., T. Manolikas, M. Rogowski, and B. H. Meier, 2006. Solid-state NMR spectroscopy of 10% 13C labeled ubiquitin: spectral simplification and stereospecific assignment of isopropyl groups. *J Biomol NMR* 35:167.

[110] Flynn, P. F., M. J. Milton, C. R. Babu, and A. J. Wand, 2002. A simple and effective NMR cell for studies of encapsulated proteins dissolved in low viscosity solvents. *J. Biomol. NMR* 23:311.

[111] Hardy, J., and D. J. Selkoe, 2002. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* 297:353.

[112] 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 beta variant favoring oligomerization in Alzheimer's-type dementia. *Ann. Neurol.* 63:377.

[113] Irie, K., K. Murakami, Y. Masuda, A. Morimoto, H. Ohigashi, R. Ohashi, K. Takegoshi, M. Nagao, T. Shimizu, and T. Shirasawa, 2005. Structure of beta-amyloid fibrils and its relevance to their neurotoxicity: implications for the pathogenesis of Alzheimer's disease. *J. Biosci. Bioeng.* 99:437.

[114] Levy-Lahad, E., W. Wasco, P. Poorkaj, D. M. Romano, J. Oshima, W. H. Pettingell, C. E. Yu, P. D. Jondro, S. D. Schmidt, and K. Wang, 1995. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* 269:973.

[115] 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 $\beta$ Fibrils and Modulates A$\beta$ Neurotoxicity. *J. Mol. Biol.* 408:780.

[116] Tycko, R., 2006. Molecular structure of amyloid fibrils: insights from solid-state NMR. *Q. Rev. Biophys.* 39:1.

[117] Tycko, R., 2011. Solid-State NMR Studies of Amyloid Fibril Structure. *Annu. Rev. Phys. Chem.* 62:279.

[118] Finder, V. H., I. Vodopivec, R. M. Nitsch, and R. Glockshuber, 2010. The recombinant amyloid-beta peptide Abeta1-42 aggregates faster and is more neurotoxic than synthetic Abeta1-42. *J. Mol. Biol.* 396:9.

[119] Bachovchin, W. W., 2001. Review: Contributions of NMR spectroscopy to the study of hydrogen bonds in serine protease active sites. *Mag. Res. Chem.* 39:S199–S213.

[120] Fändrich, M., M. Schmidt, and N. Grigorieff, 2011. Recent progress in understanding Alzheimer's $\beta$-amyloid structures. *Trends Biochem. Sci.* 36:338–345.

[121] Wishart, D. S., and B. D. Sykes, 1994. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J. Biomol. NMR* 4:171.

[122] Lewandowski, J. R., G. De Paëpe, and R. G. Griffin, 2007. Proton assisted insensitive nuclei cross polarization. *J. Am. Chem. Soc.* 129:728.

[123] Cloe, A. L., J. P. R. O. Orgel, J. R. Sachleben, R. Tycko, and S. C. Meredith, 2011. The Japanese mutant A$\beta$ ($\Delta$E22-A$\beta$(1-39)) forms fibrils instantaneously, with low-thioflavin T fluorescence: seeding of wild-type A$\beta$(1-40) into atypical fibrils by $\Delta$E22-A$\beta$(1-39). *Biochemistry* 50:2026.

[124] Zagorski, M., J. Yang, H. Shao, K. Ma, H. Zeng, and A. Hong, 1999. Methodological and chemical factors affecting amyloid beta peptide amyloidogenicity. *Method Enzymol* 309:189.

[125] Williams, A., S. Shivaprasad, and R. WetzelL, 2006. Alanine Scanning Mutagenesis of A$\beta$(1-40) Amyloid Fibril Stability. *J. Mol. Biol.* 357:1283–1294.

[126] Torok, M., 2002. Structural and Dynamic Features of Alzheimer's Abeta Peptide in Amyloid Fibrils Studied by Site-directed Spin Labeling. *J. Biol. Chem.* 277:40810–40815.

[127] Whittemore, N. A., R. Mishra, I. Kheterpal, A. D. Williams, R. Wetzel, and E. H. Serpersu, 2005. Hydrogen–Deuterium (H/D) Exchange Mapping of A$\beta$ 1-40Amyloid Fibril Secondary Structure Using Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* 44:4434–4441.

[128] Kheterpal, I., M. Chen, K. D. Cook, and R. Wetzel, 2006. Structural Differences in A$\beta$ Amyloid Protofibrils and Fibrils Mapped by Hydrogen Exchange – Mass Spectrometry with On-line Proteolytic Fragmentation. *J. Mol. Biol.* 361:785–795.

[129] Olofsson, A., M. Lindhagen-Persson, A. E. Sauer-Eriksson, and A. Öhman, 2007. Amide solvent protection analysis demonstrates that amyloid-$\beta$(1–40) and amyloid-$\beta$(1–42) form different fibrillar structures under identical conditions. *Biochem. J.* 404:63.

[130] Paravastu, A., R. Leapman, W. Yau, and R. Tycko, 2008. Molecular structural basis for polymorphism in Alzheimer's beta-amyloid fibrils. *Proc. Natl. Acad. Sci.* 105:18349–18354.

[131] Bertini, I., L. Gonnelli, C. Luchinat, J. Mao, and A. Nesi, 2011. A New Structural Model of A$\beta$40 Fibrils. *J. Am. Chem. Soc.* 113:16013.

[132] Morimoto, A., K. Irie, K. Murakami, Y. Masuda, H. Ohigashi, M. Nagao, H. Fukuda, T. Shimizu, and T. Shirasawa, 2004. Analysis of the Secondary Structure of $\beta$-Amyloid (1-42) Fibrils by Systematic Proline Replacement. *J. Biol. Chem.* 279:52781–52788.

[133] Lührs, T., C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Doeli, D. Schubert, and R. Riek, 2005. 3D structure of Alzheimer's amyloid-beta(1-42) fibrils. *Proc. Natl. Acad. Sci.* 102:17342–17347.

[134] Olofsson, A., A. Sauer-Eriksson, and A. Öhman, 2006. The solvent protection of Alzheimer amyloid-$\beta$-(1–42) fibrils as determined by solution NMR spectroscopy. *J. Biol. Chem.* 281:477.

[135] Ovchinnikova, O. Y., 2011. A$\beta$ peptide in familial Alzheimers's disease: from aggregation kinetics to fibril structure. Ph.D. thesis, ETH Zurich.

[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* 59:687.

[137] Nelson, R., M. R. Sawaya, M. Balbirnie, A. Ø. Madsen, C. Riekel, R. Grothe, and D. Eisenberg, 2005. Structure of the cross-$\beta$-spine of amyloid-like fibrils. *Nature* 435:773–778.

[138] Hennetin, J., B. Jullian, A. C. Steven, and A. V. Kajava, 2006. Standard conformations of $\beta$-arches in $\beta$-solenoid proteins. *J. Mol. Biol.* 358:1094–1105.

[139] Petkova, A. T., W.-M. Yau, and R. Tycko, 2006. Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry* 45:498–512.

[140] Thomas, D., P. Schultz, A. C. Steven, and J. S. Wall, 1994. Mass analysis of biological macromolecular complexes by STEM. *Biol. Cell* 80:181–192.

[141] Müller, S. A., and A. Engel, 2001. Structure and mass analysis by scanning transmission electron microscopy. *Micron* 32:21–31.

[142] Linser, R., M. Dasari, M. Hiller, V. Higman, U. Fink, J.-M. Lopez Del Amo, S. Markovic, L. Handel, B. Kessler, P. Schmieder, D. Oesterhelt, H. Oschkinat, and B. Reif, 2011. Proton-detected solid-state NMR spectroscopy of fibrillar and membrane proteins. *Angew. Chem. Int. Ed.* 50:4508.

[143] Marley, J., M. Lu, and C. Bracken, 2001. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* 20:71–75.

[144] Messerle, B., G. Wider, G. Otting, C. Weber, and K. Wüthrich, 1989. Solvent Suppression Using a Spin Lock in 2D and 3D NMR-Spectroscopy with $H_2O$ Solutions. *J. Magn. Reson.* 85:608–613.

[145] Lange, A., I. Scholz, T. Manolikas, M. Ernst, and B. H. Meier, 2010. Low-power cross polarization in fast magic-angle spinning NMR experiments. *Chem. Phys. Lett.* 468:100–105.

[146] Grzesiek, S., and A. Bax, 1993. Amino acid type determination in the sequential assignment procedure of uniformly 13 C/15 N-enriched proteins. *J.Biomol. NMR* 3:185–204.

[147] Tamiola, K., B. Acar, and F. A. A. Mulder, 2010. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc* 132:18000–18003.

# Acknowledgments

This thesis would not have been possible without the support of a number of people.

First, I would like to thank my supervisor, **Beat Meier**, for giving me the opportunity to work in his research group, for giving me the freedom to pursue my own interests and for his support and advice when I needed it. I thank **Roland Riek** and **Anja Böckmann** not only for co-examining my thesis but also for the fruitful discussions.

The research presented in this thesis was done in collaboration with great researchers that invested a lot of their time and shared their knowledge. I thank **Paul Schanda**, **Sebastian Hiller**, **René Verel** and **Oliver With** for the collaboration on the proton-detection project as well as **Oxana Ovchinnikova** and **Anne Schütz** for the collaboration on A$\beta$. At various times during my thesis I collaborated with **Peter Güntert**, **Matthias Ernst**, **Giorgia Zandomeneghi**, **Andrea Cavalli**, **Inés Garcia Rubio**, **Jason Greenwald**, **Kathrin Székely**, **Theofanis Manolikas**, **Ingo Scholz** and **Alice Soragni** and I would like to thank them.

Then there are people like **Thomas Westfeld**, **Hélène van Melckebeke**, **Christian Wasmer** and **Jacco van Beek** with whom I never had a project but who nevertheless gave me help and advice and therefore deserve a big thank you.

Experimental work got a lot easier with the support of our technical staff. **Andreas Hunkeler** and **Alexander Däpp** often made seemingly impossible things possible. I would also like to thank **Riccardo Cadalbert**, **Diego Sánchez** and **Fabienne Arn** for their work in the chemistry- and biolab.

I'm grateful to my office mates **Marcin Krajewski**, **Stephanie Köneke**, **Patrick Léger** and then later **Francesco Ravotti** and **Emilie Testori** for the good atmosphere we had.

I thank **Andreas Grommek**, **Michael Batel**, **Rosmarie Joss**, **Anders Nielsen** and the rest of the Meier group including our regular guests from Lyon, **Birgit Habenstein** and **Nina Luckgei**, for the many scientific and not so scientific discussions we had in the coffee room and at lunch.

Special thanks go to **Julia Gath** who always asked me about my progress in writing this thesis and thereby helped me stay focussed.

Finally, I would like to thank my family and friends for their support.

# Curriculum Vitae

| | |
|---|---|
| **Name:** | Matthias Huber |
| **Date of birth:** | 29. December 1982 |
| **Place of birth:** | Altdorf UR |
| **Nationality:** | Swiss |

## Education

| | |
|---|---|
| December 2006 – Present | **Ph.D. studies** in the group of Prof. Beat Meier at the laboratory of physical chemistry. *ETH Zürich* (Zürich, Switzerland) |
| 2005 – 2006 | **Master studies** in Chemistry at *ETH Zürich* |
| 2002 – 2005 | **Bachelor studies** in Chemistry at *ETH Zürich* |
| October 2004 – March 2005 | **Erasmus Exchange** at *Ecole Polytechnique* (Palaiseau, France) |
| 1996 – 2002 | **Abitur** at the *Kantonale Mittelschule Uri* (Altdorf) |

# Work Experience

| | |
|---|---|
| December 2006 – Present | **ETH Zürich** (Zürich, Switzerland) <u>Ph.D. thesis</u> in the solid-state NMR group of Prof. Dr. Beat H. Meier |
| December 2006– Present | **ETH Zürich** (Zürich, Switzerland) <u>Teaching assistant</u> accompanying the following lectures and practical courses: physical chemistry (thermodynamics), magnetic resonance, practical course in physical chemistry and chemistry |
| April 2006 – July 2006 | **Massachusetts Institute of Technology** (Cambridge, MA, USA) <u>Master Thesis</u> in the group of Prof. Dr. Robert G. Griffin<br><br>Title: $^{13}C$–$^{13}C$ *Distance Measurements with Rotational Resonance Width* |

# Publications

10. **M. Huber**, A. Böckmann, S. Hiller, B.H. Meier, 2012, *Perspective: 4D solid-state NMR for protein structure determination,* PCCP, doi:10.1039/C2CP23872A

9. **M. Huber**, O. With, P. Schanda, R. Verel, M. Ernst, B.H. Meier, 2012, *A supplementary coil for 2H decoupling with commercial HCN MAS probes,* J. Magn. Reson. 214:76-80

8. P. Schanda, **M. Huber**, J. Boisbouvier, B.H. Meier, M. Ernst, 2011, *Solid-State NMR Measurements of Asymmetric Dipolar Couplings Provide Insight Into Protein Side-Chain Motion,* Angew. Chem. Int. Ed. 50:11005-11009

7. **M. Huber**, S. Hiller, P. Schanda, M. Ernst, A. Böckmann, R. Verel, B.H. Meier, 2011, *A Proton-Detected 4D Solid-State NMR Experiment for Protein Structure Determination,* ChemPhysChem, 12:915-918

6. P. Schanda, **M. Huber**, R. Verel, M. Ernst, B.H. Meier, 2009, *Direct Detection of* $^{3h}J_{NC'}$ *Hydrogen-Bond Scalar Couplings in Proteins by Solid-State NMR*, Angew. Chem. Int. Ed. 48:9322-9325

5. A.B. Barnes, L. Andreas, **M. Huber**, R. Ramachandran, P.C.A. van der Wel, M. Veshtort, R.G. Griffin, M.A. Mehta, 2009, *High-Resolution Solid-State NMR Structure of Alanyl-Prolyl-Glycine,* J. Magn. Reson. 200:95-100

4. M.J. Bayro, **M. Huber**, R. Ramachandran, T. C. Davenport, B.H. Meier, M. Ernst, R.G. Griffin, 2009, *Dipolar truncation in magic-angle spinning solid-state NMR recoupling experiments,* J. Chem. Phys. 130:114506

3. B. Zagrovic, Z. Gattin, J. Lau, **M. Huber**, W. F. van Gunsteren, 2008, *Validation of the GROMOS 53A6 force field for the study of beta-peptides in solution,* Eur. Biophys. J. 37:903-912

2.  I. Scholz, **M. Huber**, T. Manolikas, B. H. Meier and M. Ernst., 2008, *MIRROR recoupling and its application to spin diffusion under fast magic-angle spinning.* Chem. Phys. Lett. 460:278-283

1.  **M. Huber**, W.J. Stark, S. Loher, M. Maciejewski, F. Krumeich, A. Baiker, 2005, *Flame synthesis of calcium carbonate nanoparticles*, Chem. Commun. 5:648-650

*Front Cover*:  $^{13}$C–$^{13}$C DARR spectrum of A$\beta$1-40 E22$\Delta$ fibrils
see figure 5.8 on page 83 for full caption

*Back Cover*:  Ubiquitin structure calculated from 3D- and 4D $^{1}$H–$^{1}$H DREAM spectra
see figure 2.7 on page 41 for full caption