# Structural investigation of CA150.WW2 amyloid fibrils by MAS NMR spectroscopy

Dissertation zur Erlangung des akademischen Grades der
Doktorin der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von

**Johanna Becker**
aus Frankfurt am Main

April 2008

1. Gutachter Prof. Dr. Hartmut Oschkinat
2. Gutachter Prof. Dr. Udo Heinemann

Disputation am 08.08.2008

# Contents


**1 Introduction** **1**

1.1 Amyloid diseases . . . 2
1.2 A structural view on amyloid fibrils . . . 3
1.2.1 Short amyloidogenic peptides . . . 5
1.2.2 Alzheimer's A$\beta$-peptide . . . 7
1.2.3 Prions . . . 9
1.2.4 Other amyloid systems: $\alpha$-synuclein, $\beta_2$-microglobulin, and amylin . . . 11
1.2.5 Artificial amyloid systems . . . 12
1.2.6 A general amyloid fold? . . . 13
1.3 CA150.WW2 – a WW-domain which can form amyloid fibrils . . . 13

**2 Theoretical Background** **15**

2.1 Principles of NMR spectroscopy . . . 15
2.2 MAS NMR spectroscopy . . . 20
2.3 Structure determination . . . 23
2.3.1 Assignment of solid state NMR spectra . . . 23
2.3.2 Structural restraints and structure calculation . . . 24

**3 Experiments and Methods** **26**

3.1 Protein preparation . . . 26
3.1.1 Expression protocol . . . 27
3.1.2 Purification protocol . . . 27
3.2 CA150.WW2 fibrillation and solid state NMR sample preparation . . . 28
3.3 NMR spectroscopy . . . 30
3.3.1 Solution state NMR experiments . . . 30
3.3.2 MAS NMR experiments . . . 30
3.4 Data analysis and structure calculation . . . 34

**4 Results** **36**

4.1 The CA150.WW2 Y19F variant and MAS NMR sample preparation . . . 36
4.2 Initial characterisation of CA150.WW2 fibrils by MAS NMR . . . 38
4.3 Assignment of CA150.WW2 amyloid fibrils employing new methods . . . 43
4.3.1 A novel $^{13}$C-assignment strategy . . . 43
4.3.2 $^{13}$C-assignment of CA150.WW2 amyloid fibrils . . . 44
4.3.3 Towards amide nitrogen and amide proton assignment . . . 56
4.3.4 3D-HNC-TEDOR – a new experimental scheme . . . 57
4.4 Assignment of the natively folded CA150.WW2 domain in solution . . . 60
4.5 Experimental constraints on the fibrillar structure . . . 60
4.5.1 Secondary structure prediction – dihedral angle constraints . . . 60

4.5.2 Long range distance constraints from MAS NMR spectra . . . . . 63
4.5.3 Constraints from other methods: Optical diffraction and mutagenesis . . 66
4.6 The fold of the monomeric unit in CA150.WW2 fibrils . . . . . 68
4.7 CA150.WW2 structure determination . . . . . 69
4.7.1 Quaternary structure constraints employing peptide mixtures . . . . . 69
4.7.2 Structure calculation on CA150.WW2 amyloid fibrils . . . . . 76

**5 Discussion** **82**
5.1 MAS NMR as a tool to study amyloid fibrils – a strategy . . . . . 82
5.2 CA150.WW2 and related structures – a $\beta$-solenoid as a common amyloid fold? . 85
5.3 Outlook . . . . . 88
5.4 Publications . . . . . 90

**6 Summary** **91**

**7 Zusammenfassung** **94**

**References** **97**

**Acknowledgement – Danksagung** **108**

**Curriculum vitae** **110**

# 1 Introduction

Amyloid deposits are associated with more than 20 known syndromes [Gazit 05, Bellotti 07], including prion diseases, Alzheimer's, and Parkinson's disease. Despite the prevalence of some of these disorders they are yet incurable. At the same time, the role of the amyloid deposits in disease and disease progression is poorly understood. It is known that these deposits and especially their prefibrillar aggregates are in most cases toxic to cells. However, the mechanism by which cells are damaged is still under discussion, focusing e.g. on the cytotoxicity of the prefibrillar aggregates or the ability of amyloid fibrils to perforate biological membranes [Stefani 07]. The deposits consist of fibrils which are mainly built from proteins characteristic for the different disorders. *In vivo* the fibrillar core is surrounded by additional non-fibrillar constituents such as the serum amyloid P component [Sipe 00]. However, the amyloidogenic protein alone is usually sufficient to form amyloid fibrils *in vitro*. They are often regarded as a result of protein misfolding. Nevertheless, functional fibril formation is also observed, e.g. for the prion protein HET-s of *Podospora anserina* involved in heterokaryon incompatibility [Saupe 00], for the CsgA curlin subunit observed in biofilms of *Escherichia* and *Salmonella* species [Chapman 02], and even in humans where the M$\alpha$ fragment of the protein Pmel17 forms amyloid fibrils which are involved in the chemical synthesis of melanin [Fowler 06].

Detailed structural information, ideally at atomic resolution, is needed for a better understanding of the impact of amyloid fibrils on cells and tissues. However, amyloid fibrils are non-crystalline and insoluble. Thus, solution NMR (**n**uclear **m**agnetic **r**esonance) spectroscopy and X-ray crystallography, generally used for structure determination in biological systems, are of limited use. On the contrary, MAS (**m**agic **a**ngle **s**pinning) NMR spectroscopy is in principle well suited for such studies. Recent technical developments, especially with respect to hardware and pulse sequence design [Baldus 02], enable the assignment of immobile proteins of intermediate size and the determination of structural constraints in these systems. Therefore, several amyloid fibrils are studied using MAS NMR, e.g. A$\beta$ [Tycko 06], HET-s [Ritter 05], and $\alpha$-synuclein [Heise 05b, Kloepper 07].

In this work, the fibrillar form of the second WW-domain of the human transcriptional activator CA150 (CA150.WW2) was studied by MAS NMR. An assignment strategy was developed to assign the spectra which showed relatively broad lines [Becker 08]. The identification of long

range distance constraints in combination with the results from alanine scanning and electron microscopy led to a structural model for the monomeric units [Ferguson 06]. Dedicated labelling schemes were designed for the investigation of the quaternary arrangement of the monomeric units in the fibrils. Finally, the structures of three slightly different geometries were calculated.

## 1.1 Amyloid diseases

Syndromes associated with amyloid fibril deposition are referred to as amyloid diseases. The onset of these diseases is generally age-related and can be influenced by mutations in the gene coding for the protein that constitutes the main fibril component. Here, a brief overview on the different classes of amyloid disorders and the involved proteins is given and several reviews can be found in the literature [Westermark 05a, Uversky 04].

In systemic amyloidosis extracellular amyloid deposits are observed in several tissues and organs in parallel, generally with the exception of the brain. The amyloid fibril protein is secreted into the plasma and transported to the deposition site. The amount of deposits can be very large (10-15 kg) which severely disturbs the function of the affected organs. Examples are **a**myloidoses derived from the immunoglobulin **l**ight chain (AL) or **t**rans**t**hy**r**etin (ATTR). The occurrence of extracellular deposits of the islet amyloid polypeptide (amylin) in the pancreas causes type II diabetes and is an example for an organ-specific amyloidosis. In the wide spread Alzheimer's disease extracellular deposits of the Amyloid-$\beta$ (A$\beta$) peptide are found in the brain.

Diseases with intracellular amyloid deposits are not always referred to as amyloidoses but nevertheless represent an important class of amyloid disorders. They include Parkinson's disease, where neuronal inclusions of $\alpha$-synuclein, the Lewy bodies, are observed. Another example is Huntington's disease, which is an inherited disorder caused by an abnormally long polyglutamine expansion of the protein huntingtin deposited as fibrils in the neurons. Another important group are the prion diseases such as the human Creutzfeldt-Jacob disease or the **b**ovine **s**pongiform **e**ncephalopathy (BSE) associated with the deposition of the prion protein in the brain. These diseases can occur sporadically but also through contact with infected tissue. The feeding of cattle with animal by-products has lead to the BSE epidemic in Great Britain in the 1990s.

Most of these diseases are age-related and thus relevant in populations with increasing lifespan. To develop therapeutic concepts it is necessary to understand the principles behind protein aggregation, amyloid formation, what causes the onset of these diseases, the protein structure in the amyloid deposits, and the way they interact with other cellular compounds. Only few structural data has been collected on the different amyloid systems up to now, and this work aims to contribute by structural investigations at atomic resolution.

## 1.2 A structural view on amyloid fibrils

The term amyloid is often used in different ways. In the context of diseases it was defined by the Nomenclature Committee of the International Society of Amyloidosis as "extracellular depositions of protein fibrils with characteristic appearance in electron microscope, typical X-ray diffraction pattern, and affinity for Congo Red with concomitant green birefringence" [Westermark 05b]. Here, the definition of Fändrich [Fandrich 07] is used. It is based on structural considerations and describes amyloid fibrils as fibrillar protein aggregates with a cross-$\beta$ conformation, normally revealed by X-ray or electron diffraction.

The core of amyloid fibrils is built from disease specific polypeptide chains, so-called primary components which are surrounded by diverse secondary components including apolipoprotein E, amyloid P component, heparan sulfate proteoglycans and some lipids [Sipe 00, Fandrich 07]. While the amyloidogenic peptide is characteristic for a certain disease, often the same secondary components are found. The proteins forming the amyloid core present a diverse group with respect to size, amino acid composition, and native structure. However, kinetic experiments and studies by electron microscopy, X-ray crystallography and NMR spectroscopy as described in the next paragraphs showed that the biophysical and structural properties of amyloid fibrils are very similar, indicating that these aggregates share a common overall structure.

Amyloid fibrils form with distinct kinetics. Initially, nucleation with a characteristic lag phase occurs, but seeding of the protein solution with preformed fibrils removes this lag phase [Dobson 03]. Before formation of mature fibrils, prefibrillar aggregates, sometimes termed amorphous aggregates, occur which can have a spherical shape. Interestingly, specific antibodies can be designed which bind to the prefibrillar aggregates of different proteins but not to the mature fibrils, showing that prefibrillar aggregates and fibrils differ substantially from

each other [Kayed 03]. It has been demonstrated that the prefibrillar aggregates are generally more toxic to cells than the mature amyloid fibrils [Stefani 07]. This led to the speculation that amyloid cytotoxicity is based on the structural properties of the prefibrillar aggregates [Stefani 03]. Only little data is available on how the structure of the prefibrillar aggregates compares with the arrangement of the polypeptide chain in mature amyloid fibrils. A study on A$\beta$(1-40), however, indicates that the fold of the toxic prefibrillar aggregates and the mature amyloid fibrils are similar [Chimon 07]. The mature fibrils are the end point of the aggregation process and thus are accessible for structural studies.

Due to the non-crystalline insoluble nature of amyloid fibrils a variety of methods has been applied to elucidate their structure. The most widely applied technique is electron microscopy which typically shows straight, unbranched amyloid fibrils which are 0.1-1.6 $\mu$m in length and have a diameter of 60-120 Å. They are built from protofibrils with a diameter of 25-35 Å [Sipe 00, Stefani 03]. Similar proteins may form amyloid fibrils with different morphologies exhibiting a twist of the fibril or lateral association forming an untwisted fibril. Oriented amyloid fibrils show a typical so-called cross-$\beta$ X-ray diffraction pattern with a meridional 4.6-4.8 Å spacing, which results from the main chain distance within $\beta$-sheets running parallel to the fibril axis [Serpell 99]. The variable equatorial 8-12 Å spacing corresponds to the distance between two $\beta$-sheets [Stefani 03, Fandrich 07]. Interestingly, amyloid fibrils show a different amid I' band in FTIR spectra in comparison with native $\beta$-sheet-rich globular proteins [Zandomeneghi 04]. The underlaying common structural properties of amyloid fibrils lead to specific reactions with staining dyes (Congo Red [Puchtler 65], Thioflavine T [LeVine 93]) and antibodies [O'Nuallain 02].

Proton-deuterium exchange solution NMR spectroscopy can be used to investigate the accessibility of amide groups in amyloid fibrils yielding information on the position of the $\beta$-strands within the polypeptide chain. Using solid-state NMR the structural properties of amyloid fibrils can be investigated directly. From the resonance assignment information on the amyloid fold can be obtained by secondary structure prediction. Ideally, long range contacts are detected enabling the determination of the overall fold of the polypeptide chain in the amyloid conformation. If a sufficient number of such constraints is available, a structure can be calculated which has been shown for crystalline proteins [Castellani 02, Zech 05]. Recently, this technique led to the first high resolution structure of an amyloid fibril, the prion protein

HET-s(218-289) fibril from *Podospora anserina* [Wasmer 08]. A different approach to obtain atomic resolution structural information can be applied to small fragments of amyloidogenic proteins. In this case, sometimes single crystals can be grown which resemble amyloid fibrils and enable structure determination by X-ray crystallography [Nelson 05, Sawaya 07].

The structural information obtained to date on several amyloid systems indicates a fold of the polypeptide chain which can be described by a $\beta$-arch [Fandrich 07]. Two $\beta$-strands are separated by a loop region and incorporated in different $\beta$-sheets that interact via their side chains. This basic structural motif can be extended by additional $\beta$-sheets. A related structural fold is the $\beta$-helix or $\beta$-solenoid [Kajava 06] where $\beta$-strands are arranged in a helical manner to form two to four $\beta$-sheets that are in contact with each other via their side chains. This structure is e.g. found in the extraordinary stable tail spike proteins [Steinbacher 94] and was also suggested to be the structural core of amyloid fibrils, whereby each coil is formed by a single polypeptide chain [Wetzel 02]. The structure determined on HET-s(218-289) amyloid fibrils also resembles a $\beta$-helix with two $\beta$-helical coils per monomeric unit [Wasmer 08].

A general picture of the amyloid structure is still missing and to date only few systems have been studied at high resolution. A recent review on solid state NMR spectroscopy of amyloid fibrils is given by Heise [Heise 08]. Open questions are, how amyloid fibrils built from different polypeptide chains or under different conditions compare, and what the structural basis of the different morphologies is. So far, detailed structural information could only be obtained on a relatively small number of amyloid systems. An overview on structural studies on amyloid fibrils is given in the following paragraphs.

### 1.2.1 Short amyloidogenic peptides

Short amyloid-forming fragments of 5-12 residues could be derived from many amyloidogenic polypeptide chains. In their amyloid form they often display a very similar morphology to the full length fibrils and are sometimes also cytotoxic [Gazit 05]. They were used as model systems, e.g. in positional scanning mutagenesis [de la Paz 04] to investigate the sequence determinants of amyloid formation. Peptides shorter than 13 residues were also the first systems for which high resolution structural information was obtained [Jaroniec 04, Makin 05, Nelson 05, Sawaya 07].

In 2004, Jaroniec et al. published the MAS NMR structure of the monomeric unit in amyloid

fibrils formed by the 11-residue fragment YTIAALLSPYS from transthyretin (TTR, residues 105-115) [Jaroniec 04]. TTR in its native form is a homotetramer which is involved in the transport of thyroxine and retinol in the plasma. However, in the human body it can cause senile systemic amyloidosis [Westermark 90], and some variants are associated with familial amyloid polyneuropathy [Saraiva 01]. The structure was calculated based on nitrogen-carbon distances, chemical shift derived and experimentally determined backbone dihedral angle constraints. The obtained conformation in TTR(105-115) fibrils is that of an extended $\beta$-strand. However, the overall fibril architecture was not determined. The side chains of the central region of the peptide are oriented perpendicular to the amide plane and thus mediate the contact to the neighbouring $\beta$-sheets.

Short amyloidogenic peptides were also used to grow fibril-like crystals for structure determination by X-ray crystallography. This was demonstrated by Makin et al. for the designed 12-residue peptide KFFEAAAKKFFE [Makin 05]. In the group of Eisenberg peptides (four to seven residues) derived from different amyloidogenic proteins (e.g. the yeast protein Sup35 or the A$\beta$-peptide) were crystallised [Nelson 05, Sawaya 07]. For 13 peptides amyloid fibril-like diffracting crystals could be grown. Their structures were solved and revealed a so-called steric zipper motif as main structural feature, Figure 1. It is built by a pair of $\beta$-sheets. The side chains of the different $\beta$-sheets that face each other interdigitate to form a dry interface. The $\beta$-sheets are displaced with respect to each other along the crystal symmetry axis by half the

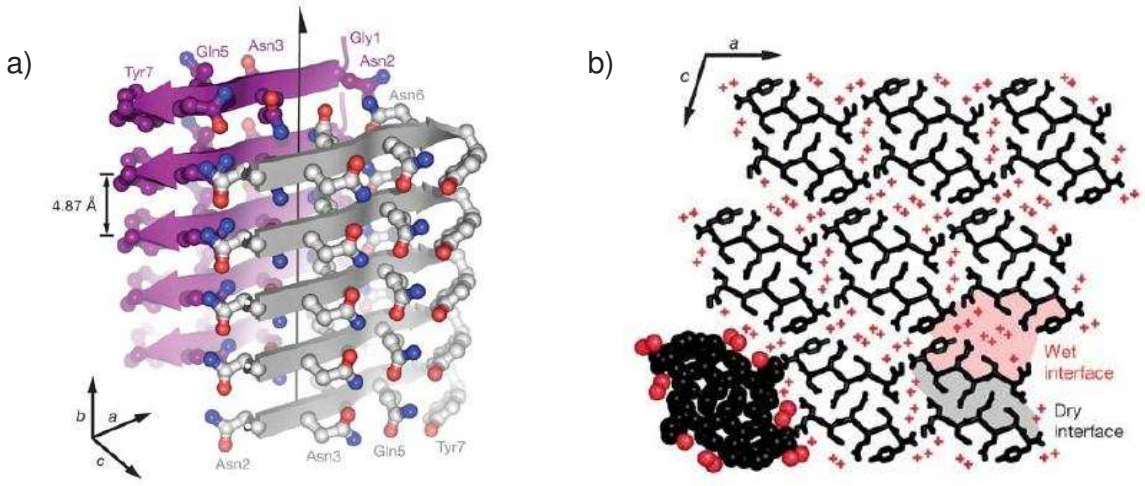



Figure 1: Structure of the GNNQQNY-peptide in its crystal form, a) side view on two $\beta$-sheets forming the dry interface, b) cross section perpendicular to the fibril axis showing the arrangement of the dry and the wet interfaces in the crystal and the tight side chain interdigitation in the dry interface. From Nelson et al. [Nelson 05].

strand-strand separation. Therefore, each side chain is in contact with the side chains from two strands of the mating sheet. In the various crystals different arrangements of the $\beta$-strands are observed. They can either form parallel or antiparallel sheets, which pack against each other via the same or the different face, and are oriented parallel or antiparallel with respect to one another. Combination of these possibilities results in a total of eight arrangements. Five of them were so far observed experimentally. Interestingly, the studied peptides comprise a variety of different sequences ranging from highly polar to highly apolar and from small to large amino acids. It is suggested that the interdigitation of side chains of similar size is crucial for the observed steric zipper arrangements [Sawaya 07]. In summary, structural studies on short amyloidogenic peptides indicate that amyloid fibrils are built from $\beta$-sheets which interact via a steric zipper mechanism.

### 1.2.2 Alzheimer's A$\beta$-peptide

One of the most intensively studied amyloidogenic peptides is Alzheimer's A$\beta$ peptide. It is the major component of the deposits found in the brain of Alzheimer's disease patients. The peptide is generated from the membrane-integrated $\beta$-amyloid precursor protein by proteolytic cleavage [Selkoe 94]. Several isoforms of the peptide are observed *in vivo* ranging in length from 37 to 43 amino acids. The major species in the brain is A$\beta$(1-40) [Mori 92], however in the deposits A$\beta$(1-42) is prevalent [Iwatsubo 94]. It has also been shown that A$\beta$(1-42) (in comparison with A$\beta$(1-40)) has an increased propensity to form amyloid fibrils *in vitro* [Jarrett 93]. In addition, many shorter fragments form amyloid fibrils *in vitro* [Burdick 92].

A couple of solid state NMR studies were carried out on A$\beta$(1-40) fibrils, which in contrast to A$\beta$(1-42) can efficiently be synthesised. In comparison to other solid biological samples, such as 2D crystals of membrane proteins [Hiller 05], they show very broad lines. Assignment was facilitated by selective labelling. The observed chemical shifts indicate two $\beta$-strand regions ranging from residues $G_9$ to $A_{21}$ and $A_{30}$ to $V_{36}$. A salt-bridge was detected connecting residues $D_{23}$ and $K_{28}$ [Petkova 02]. Multiple quantum experiments showed that the $\beta$-sheets in the fibril are parallel with an in-register arrangement of the $\beta$-strands [Antzutkin 00]. Using $^{13}$C-$^{13}$C correlations, contacts between different $\beta$-sheets were observed. The strand-strand distance in the $\beta$-sheets is 4.8 Å. The displacement along the fibril axis of the $\beta$-sheets with respect to each other was probed by dilution experiments and revealed a shift of $\frac{3}{2}$ times

the strand-strand spacing which was referred to as a STAG(2) arrangement. However, due to steric considerations in this study only odd (and not even) numbered multiples of half the strand-strand separation were taken into account. The direction of this shift could not be determined. In combination with the data from electron microscopy a structural model was constructed [Petkova 06]. The basic building block consits of two monomeric units each forming two $\beta$-strands. Thus two pairs of $\beta$-sheets constitute the protofilament, whereby the $\beta$-sheets formed by strand one of the monomeric unit face towards the outside of the protofilament while the two $\beta$-sheets formed by strand two interact with each other, Figure 2a).

Small prefibrillar aggregates of A$\beta$(1-40) that occur as an intermediate during fibril formation and are highly toxic to cells were investigated by solid state NMR $^{13}$C-$^{13}$C correlations after freeze-trapping these intermediates [Chimon 05, Chimon 07]. The spectra are nearly identical to those recorded on mature fibrils indicating that the structural arrangement in both forms is very similar.

Only few solid state NMR data exists on A$\beta$(1-42) fibrils. Homo- and heteronulcear recoupling experiments showed that the $\beta$-sheets are parallel and in-register as observed for A$\beta$(1-40) [Antzutkin 02]. Using additional data from other methods Lührs et al. have constructed a structural model, Figure 2b) [Lührs 05]. Proton-deuterium exchange experiments revealed two $\beta$-strand regions ranging from residues $V_{18}$ to $S_{26}$ and $I_{31}$ to $A_{42}$. The intersheet packing was investigated by mixing of several variants, including combinations that recover wild-type fibril formation as monitored by electron microscopy. The resulting model

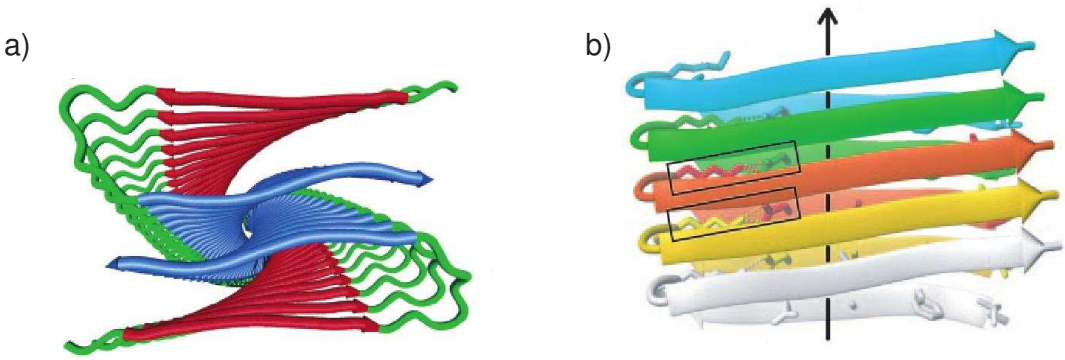



Figure 2: a) Structural model of A$\beta$(1-40) from Petkova et al. [Petkova 06], the basic building block contains four $\beta$-strands formed by two monomeric units, strand one and two of the monomeric units are given in red and blue, respectively. b) Structural model of A$\beta$(1-42) from Lührs et al. [Lührs 05], the basic building block contains two $\beta$-strands formed by a single peptide, different monomers are given in different colours, the salt-bridge investigated by mixing of several variants is indicated by black boxes.

contained only two $\beta$-sheets. As in A$\beta$(1-40) the side chain contacts are intermolecular but the register of the interaction between the two $\beta$-strands is shifted by 10 Å, thus showing a different side chain packing in the core of the structure. A displacement along the fibril axis of the $\beta$-sheets with respect to each other by approximately the strand-strand distance was observed.

In addition to these studies, structural data was also obtained on shorter A$\beta$-peptides. The A$\beta$(10-35) peptide forms fibrils with an in-register parallel $\beta$-sheet arrangement, similar to the structure of the full length A$\beta$-peptides [Antzutkin 02]. A$\beta$(10-40) lacking the N-terminus which is unstructured in the fibrils [Kheterpal 01] was used to show that fibrils grown under different conditions with respect to pH, agitation, and seeding display different morphologies. MAS NMR $^{13}$C-$^{13}$C correlations indicated that different morphologies are related to different registers of the strand-strand interaction [Paravastu 06]. Interestingly, amyloid fibrils formed by A$\beta$(11-25) consist of just a single $\beta$-strand and contain an antiparallel $\beta$-sheet which was revealed by MAS NMR spectroscopy [Petkova 04]. These investigations, in combination with studies that show a strong dependence of fibril morphology on experimental conditions, suggest that there might not be a unique structure of A$\beta$-fibrils.

### 1.2.3 Prions

Prions are **pr**oteinaceous **i**nfectious particles (-on by analogy to viri**on**) [Dobson 05] that cause the transmissible spongiform encephalopathies which include the Creutzfeldt-Jakob disease and Kuru in humans as well as scrapie and BSE in animals. All these disorders are associated with amyloid formation of the prion protein and neuronal cell death [Harris 06]. Functional prion formation is found in some fungi, and so far most structural studies on prion amyloid fibrils was done on these functional prions (see below). In addition, many prion proteins have been characterised in their soluble form [Lysek 05, Gossert 05, Calzolai 05].

Ure2p is a prion protein of *Saccharomyces cerevisiae* forming functional amyloid fibrils that cause a certain metabolic phenotype of this yeast [Wickner 94]. The prion domain of this 354-residue protein consists of the N-terminal 89 amino acids. Ure2p(10-39) is the most highly conserved segment in this domain and was used in a MAS NMR study [Chan 05]. The full prion domain Ure2p(1-89) was also investigated with this technique [Baxa 07]. These studies revealed that amyloid fibrils formed from the N-terminal domain contain in-register

parallel $\beta$-sheets. The domain is rich in asparagine and glutamine residues. Therefore, it was suggested that in amyloid fibrils of Ure2p(1-89) these residues form polar zippers connecting the monomeric units. Two structural models with a strand-turn-strand conformation were postulated [Chan 05].

HET-s is the prion protein of the yeast *Podospora anserina* and is involved in heterokaryon incompatibility, a mechanism that controls vegetative cell fusion [Saupe 00]. To date, the only known high resolution amyloid fibril structure was obtained by MAS NMR spectroscopy on fibrils formed by the prion domain of HET-s consisting of the C-terminal residues 218-289 [Wasmer 08]. In an earlier study, the $\beta$-strand regions were derived from proton-deuterium exchange experiments and solid state MAS NMR chemical shifts [Ritter 05]. In contrast to other MAS NMR studies of amyloid fibrils, the observed lines are very narrow. Interestingly, in CP-MAS spectra only 43 of the 78 residues were detected [Siemer 06b]. Most of the missing residues were, however, observed using proton detection and solution NMR techniques, indicating a high mobility in some parts of the protein [Siemer 06a]. In addition, the observed chemical shifts suggest a random coil conformation of these residues which comprise the N and C-terminus of the protein and residues in the loop regions. Recently, the HET-s(218-289) structure at atomic resolution was presented [Wasmer 08]. The monomeric unit forms eight short $\beta$-strands. Six of them are arranged in a manner similar to a left-handed $\beta$-helix or

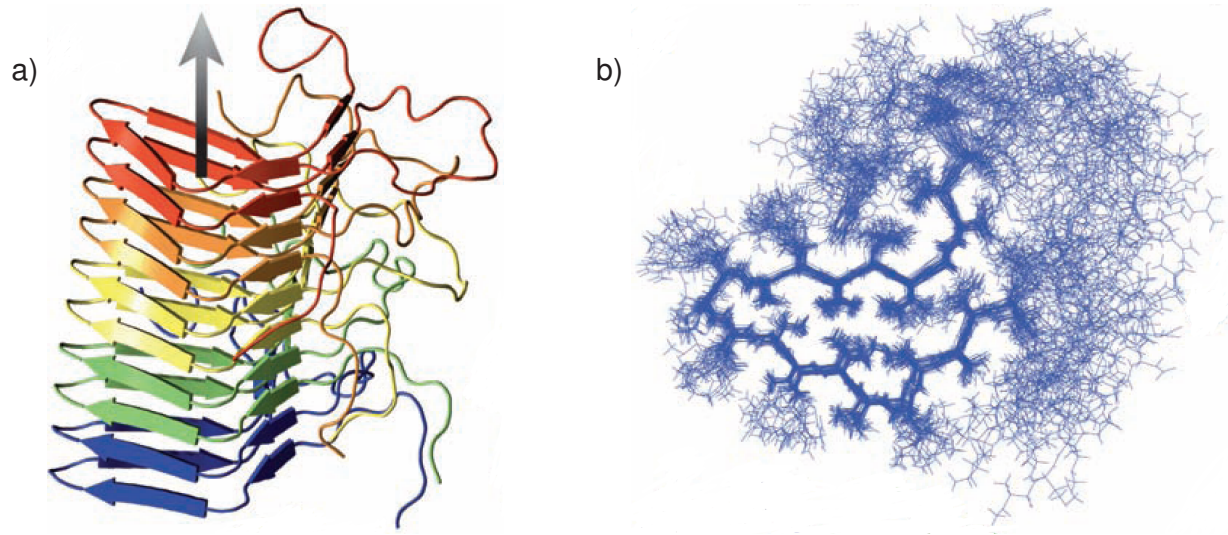



Figure 3: Representations of the high resolution MAS NMR structure of HET-s(218-289), a) backbone fold of five monomeric units, the single monomeric units are shown in different colours, b) structural ensemble of a monomeric unit. From Wasmer et al. [Wasmer 08].

$\beta$-solenoid with two coils per monomeric unit, Figure 3. Thereby, the $\beta$-strands incorporated in the same $\beta$-sheet are pseudo repeats of each other. Between strands three and four a long loop region, not incorporated in the fibril core, is observed.

### 1.2.4 Other amyloid systems: $\alpha$-synuclein, $\beta_2$-microglobulin, and amylin

$\alpha$-Synuclein is a natively unfolded 140-residue protein. In its amyloid form it is the main component of the Lewy bodies, the intracellular inclusions in dopaminergic neurons found in Parkinson's disease patients [Cookson 05]. EPR (**e**lectron **p**aramagnetic **r**esonance) studies indicate a central core region with parallel in-register $\beta$-sheets [Der-Sarkissian 03] for $\alpha$-synuclein fibrils. In MAS NMR spectra remarkably narrow lines are observed [Heise 05a, Kloepper 07]. In a study by Heise et al. [Heise 05a] two polymorphs were detected. Both have a disordered N-terminus, an immobile core region, and a highly flexible C-terminus. 48 and 32 residues could be assigned in the core regions of the two polymorphs, respectively. Secondary chemical shift analysis shows $\beta$-strand conformation for most assigned residues in the core region ranging from $L_{38}$ to $V_{95}$. MAS NMR measurements at low temperature (-40°C) by Kloepper et al. [Kloepper 07] revealed an increased sensitivity, probably due to a more efficient cross polarisation transfer. In the region between $T_{22}$ and $S_{129}$, 44 residues were assigned. Particularly strong signals were observed between residues $T_{64}$-$K_{80}$ and $A_{85}$-$G_{93}$. TALOS analysis revealed five $\beta$-strands within these regions. Interestingly, in the samples used in the course of this study only one polymorph was observed.

$\beta_2$-microglobulin is a part of the class I major histocompatibility complex (MHC) and forms amyloid fibrils in dialysis-related amyloidosis [Naiki 97]. By digesting $\beta_2$-microglobulin the 22-residue K3-peptide is obtained which forms amyloid fibrils *in vitro* [Kozhukh 02]. MAS NMR spectra recorded on these fibrils show relatively broad lines. Main chain signal assignment revealed two $\beta$-strands separated by a loop region. In addition, DARR correlations between side chains of the two strands were observed. In combination with dihedral angles from TALOS analysis they were used to determine a three dimensional structure. Based on inter molecular contacts, two structural models were derived. They show a displacement of the two $\beta$-sheets with respect to each other along the main fibril axis by half the strand-strand separation up or down, respectively. Structure calculations revealed that only the so-called STAG(+1) arrangement was consistent with the experimental data [Iwata 06], Figure 4.

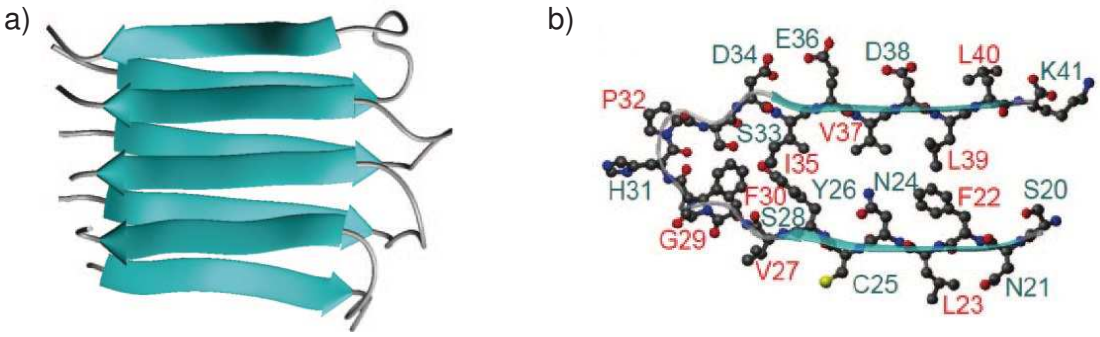



Figure 4: Structure of the K3-peptide form $\beta_2$-microglobulin a) backbone representation of the amyloid fibril and b) a monomeric unit. From Iwata et al. [Iwata 06].

In the pancreas of 90% of the patients suffering from type two diabetes, depositions of the 37-residue islet amyloid peptide, also known as amylin, are found. It forms amyloid fibrils *in vitro* which were studied by MAS NMR spectroscopy [Luca 07]. Relative broad lines, comparable to the A$\beta$(1-40) peptide ([Petkova 02]), are observed. The assignment was facilitated by selective labelling. From a TALOS analysis the presence of two $\beta$-strands ranging from $A_8$ to $V_{17}$ and $S_{28}$ to $Y_{37}$ was concluded. In addition, an inter strand side chain/side chain contact (either $D_{14}/I_{26}$ or $F_{15}/I_{27}$) was observed. Mass-per-length measurments indicated that two monomeric units per strand-strand distance occur along the fibril axis. Using this information and a $C_2$-rotational symmetry around the fibril axis, two possible structural models for the protofilament were suggested that are very similar to the A$\beta$(1-40) model developed by Petkova et al. [Petkova 06].

### 1.2.5 Artificial amyloid systems

Some proteins and peptides form amyloid fibrils *in vitro* although they are not correlated to diseases or functional amyloid formation, for example PI3-SH3 (from the bovine phosphatidyl-inositol-3'-kinase) [Ventura 04]. By inserting the short amyloidogenic peptide sequences, STVIIE or KLVFFA, into the $\alpha$-spectrin SH3 domain, amyloid formation can be triggered [Esteras-Chopo 05]. Nevertheless, the modified domains were as stable as the wild type domain. Proteolytic cleavage experiments showed that the amyloid core in these fibrils was built by the inserted peptide stretches including some adjacent residues. The major part of the protein was not incorporated. This study supports the "amyloid stretch hypothesis" stating that the self-assembly of a polypeptide chain to an amyloid fibril can be triggered by a short highly amyloidogenic stretch of amino acids. Also some polyaminoacids such as poly-E (74 kDa),

poly-T (7.6 kDa), and poly-K (77.3 kDa) form amyloid fibrils stimulating discussions that the amyloid fold might be a general property of the protein backbone [Fändrich 02].

### 1.2.6 A general amyloid fold?

In summary, the structural studies conducted on different amyloid systems within the size range of 20 to 140 residues reveal at least two $\beta$-strands separated by a loop region as the amyloid core. The $\beta$-strands are incorporated in parallel in-register $\beta$-sheets that interact with each other. So far, antiparallel $\beta$-sheets were only observed for very short polypeptide chains. In combination with the cross-$\beta$ diffraction pattern generally detected for amyloid fibrils, these observations indicate that there is a general amyloid fold. However, the sequence determinants of amyloid formation and how and if the different sequences can be accommodated in a single structural model are still a matter of intense debate pending on the determination of high resolution structures. Open questions are how the amyloid fold compares to the $\beta$-helix, the $\beta$-solenoid, and the steric zipper motifs. More precisely, the side chain packing and the shift of the $\beta$-sheets along the fibril axis with respect to each other need to be elucidated for a range of amyloid systems. Furthermore, it is unclear how many $\beta$-sheets or groups of interacting $\beta$-sheets constitute a protofilament and whether this varies among the different systems.

## 1.3 CA150.WW2 – a WW-domain which can form amyloid fibrils

The WW-domain is a small ubiquitous protein module characterised by two highly conserved tryptophans (donating its name) and one strictly conserved proline residue. It is one of the simplest natural $\beta$-sheet structures and folds extremely fast [Ferguson 01] into a three-stranded antiparallel $\beta$-sheet, binds usually proline-rich ligands and is involved in protein-protein interactions. WW-domains are found in a wide range of structural, regulatory, and signalling proteins [Sudol 96].

The second WW-domain of the human transcriptional activator CA150 (CA150.WW2, from **co**a**c**tivator protein with apparent molecular mass of **150** kDa) [Sune 97] is in sequence identical to the murine formin binding protein 28 (FBP28) [Chan 96] and forms amyloid fibrils under physiological conditions [Ferguson 03]. The small size (37 residues) of CA150.WW2 and the wealth of kinetic and mutagenesis data obtained in the laboratory of Alan R. Fersht

Figure 5: Backbone representation of the solution NMR structure of CA150.WW2 [Macias 00] showing the three-stranded antiparallel $\beta$-sheet.

and Neil Ferguson at the Medical Research Council (MRC) in Cambridge, UK, make it a good model system for structural studies of amyloid fibrils. Its native fold was determined by Macias et al. [Macias 00] and is shown in Figure 5.

Interestingly, increasing evidence indicates a role of CA150 in Huntington's disease, a dominantly inherited neurodegenerative disorder caused by the N-terminal expansion of a glutamine repeat in huntingtin [Macdonald 93]. The N-terminal region of huntingtin forms intra cellular inclusions in neurons [DiFiglia 97]. These deposits were shown to be amyloid-like [McGowan 00] by Congo Red staining. A yeast two-hybrid interactor screen showed interactions of CA150 with huntingtin involving the WW-domains of CA150 [Faber 98]. In the brain of Huntington's disease patients CA150 expression was found to be increased. In addition, CA150 colocalises with huntingtin [Holbert 01] and an animal study indicated that CA150 disfunction might be an important factor that promotes the striatal cell death in Huntington's disease [Arango 06].

Here, CA150.WW2 amyloid fibrils were chosen for structural studies by MAS NMR spectroscopy to widen the range of amyloid systems for which structural information is available and to provide a high resolution structure. Thereby, MAS NMR techniques were developed to deal with heterogeneous samples and to pave the way for structural analysis of amyloid systems with increased line widths in MAS NMR spectra.

# 2 Theoretical Background

## 2.1 Principles of NMR spectroscopy

In this section a brief summary on the theoretical background of the NMR experiments used in this work is given. Complete presentations can be found e.g. in the reviews by Laws et al. [Laws 02] and Baldus [Baldus 02] or in the textbooks by Levitt [Levitt 01] and Duer [Duer 04].

NMR spectroscopy can be used to probe the chemical environment and the dynamic properties of atomic nuclei with a nuclear spin quantum number $I > 0$. It is thus a very valuable tool when investigating chemical or biological systems on a molecular level. The technique uses the interaction of the intrinsic magnetic moments of these nuclei, the so-called nuclear spins, with an external magnetic field. This results in a splitting of nuclear energy levels. Hence the nuclei can interact with electromagnetic waves in the radio frequency range.

Nuclei with a nuclear spin quantum number $I > \frac{1}{2}$, so-called quadrupolar nuclei, have an electric quadrupolar moment. The resulting quadrupolar interaction is very strong in comparison to the interaction of the magnetic moments with the external magnetic field and thus complicates the spectra of these nuclei. Therefore, in biological NMR only nuclei with $I = \frac{1}{2}$ are used. The following theoretical considerations and experimental schemes are restricted to these non-quadrupolar nuclei, Table 1. In addition, the so-called secular approximation will be used throughout this work. It requires a large external magnetic field. Then the interaction of the nuclear spins with this external field is strong and all other spin interactions can be treated as perturbations. In the secular approximation only the zeroth order perturbations are considered.

The nuclear spin is described by the magnetic moment operator $\hat{\mathbf{I}} = (\hat{I}_x, \hat{I}_y, \hat{I}_z)$. In a macroscopic sample, a large ensemble of spins has to be considered. The state of a spin ensemble is described by the density operator $\hat{\rho}(t)$. The ensemble-averaged matrix elements are given

| Isotope | Natural abundance | Gyromagnetic ratio in $\frac{1}{sT} * 10^6$ |
|---|---|---|
| $^{1}$H | 100 % | 267.522 |
| $^{13}$C | 1.1 % | 67.283 |
| $^{15}$N | 0.37 % | -27.126 |
| $^{31}$P | 100 % | 108.394 |

Table 1: $I = \frac{1}{2}$ nuclei commonly used in NMR spectroscopy.

by: $\hat{\rho}_{ij}(t) = \overline{c_i^*(t)c_j(t)}$. The diagonal elements of this matrix are the eigenstates of the spin system, whereas the off-diagonal elements are termed coherences. These coherences are referred to as single quantum coherence, double quantum coherence, and so on, depending on their distance from the diagonal. Due to the NMR selection rules only single quantum coherences are observables. From the density matrix it is possible to understand and predict the outcome of NMR experiments.

In the presence of a static magnetic field in the z-direction ($\mathbf{B} = (0, 0, B_0)$) the density matrix in thermal equilibrium is proportional to the z-component of the nuclear magnetic moment operator: $\hat{\rho} \sim \hat{I}_z$. The time dependence of the density matrix is given by the Liouville-von-Neumann equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) = -i(\hat{H}(t)\hat{\rho}(t) - \hat{\rho}(t)\hat{H}(t)),$$

where $\hat{H}(t)$ is the total energy operator of the system, the so-called Hamiltonian. Introducing the propagator, $\hat{U}(t)$, a general solution of this differential equation is given by $\hat{\rho}(t) = \hat{U}(t)\hat{\rho}(0)\hat{U}^{-1}(t)$. For a time-independent Hamiltonian an expression for $\hat{U}(t)$ can be found easily: $\hat{U}(t) = \exp(-i\hat{H}t)$.

In more complicated situations, however, approximations need to be used. Sometimes it is for example possible to define a Hamiltonian that is piecewise constant in time. A very powerful method to deal with such situations is the Average Hamiltonian theory which can be applied when the Hamiltonian is cyclic. The obtained average Hamiltonian then describes the state of the spin system after a complete cycle but not in between.

Based on the properties of classical magnetic moments the Hamiltonian can be directly written down. The main interaction Hamiltonian, the so-called Zeeman Hamiltonian, $\hat{H}_Z$, describes the interaction of the nuclear spins with the static magnetic field:

$$\hat{H}_Z = \gamma\hbar\hat{I}_z = \omega_0\hat{I}_z,$$

where $\gamma$ is the gyromagnetic ratio of the respective nuclei and $\hbar$ the reduced Planck's constant. Electromagnetic waves of the Larmor frequency, $\omega_0 = \gamma\hbar$, are absorbed by the nuclei – this process is termed **n**uclear **m**agnetic **r**esonance. NMR spectra of certain nuclei can therefore be recorded by applying a radio frequency pulse at the Larmor frequency to a sample in the magnetic field. This pulse changes the density matrix and can lead

to detectable single quantum coherences. In other words, the longitudinal equilibrium magnetisation along the magnetic field is changed and transverse magnetisation in a plane perpendicular to the magnetic field is created which precesses around the z-axis with the Larmor frequency. This induces voltages in the receiver coil. Before digitisation the received signal is mixed with the radio frequency which was initially used to irradiate the sample (the carrier frequency). Thus only deviations from this carrier frequency are sampled. The resulting time-dependent spectrum is termed free induction decay (FID) and can be Fourier-transformed into a frequency-dependent spectrum.

The Hamiltonian describing the interaction of the spins with a radiofrequency field of the amplitude $2B_1$, the frequency $\omega_{\mathrm{rf}}$, and the phase $\varphi$ is given by:

$$\hat{H}_{\mathrm{rf}} = \gamma B_1(\hat{I}_{\mathrm{x}} \cos(\varphi + \omega_{\mathrm{rf}} t) + i\hat{I}_{\mathrm{y}} \sin(\varphi + \omega_{\mathrm{rf}} t)).$$

Due to the internal referencing of the received radiofrequency, only the modulation by the offset frequency $\omega_{\mathrm{off}} = \omega_0 - \omega_{\mathrm{rf}}$ is visible in the FID.

In addition to the external interactions, the spins are influenced by their electrostatic environment and by their neighbouring spins. The total Hamiltonian under the secular approximation reads: $\hat{H}(t) = \hat{H}_Z + \hat{H}_{\mathrm{rf}} + \hat{H}_{\mathrm{int}}$, where the internal Hamiltonian, $\hat{H}_{\mathrm{int}}$, contains several contributions: $\hat{H}_{\mathrm{int}} = \hat{H}_{CS} + \hat{H}_D + \hat{H}_J$, which are described in the following paragraphs.

**Chemical shift:** An important interaction occurs between the spins and their electrostatic environment. Variations in the electron density around each nucleus, e.g. due to chemical bonds, lead to a shielding or deshielding of the external magnetic field. This results in a slightly different local field at the place of the nucleus and thus in a different resonance frequency. The effect is called chemical shielding and leads to a dispersion of resonance frequencies from nuclei in different environments. This chemical shift is often characteristic for the chemical entity in which a nucleus is incorporated. The chemical shielding is orientation dependent which is described by the chemical shift tensor, $\sigma$, that is diagonal in the principle axis frame with the diagonal elements $\sigma_{\mathrm{xx}}$, $\sigma_{\mathrm{yy}}$, and $\sigma_{\mathrm{zz}}$. The chemical shift Hamiltonian $\hat{H}_{CS}$ is given by:

$$\hat{H}_{CS} = \mathbf{B}\sigma\hat{\mathbf{I}} = \omega_0 \sigma_{\mathrm{zz}} \hat{I}_{\mathrm{z}},$$

where $\sigma_{zz}$ in the laboratory frame reads:

$$\sigma_{zz}^{LAB} = \sigma_{iso} + \frac{\delta_{CS}}{2\omega_0}(3\cos^2\beta - 1 + \eta_{CS}\sin^2\beta\cos 2\alpha), \tag{1}$$

with the isotropic chemical shift $\sigma_{iso} = \frac{1}{3}(\sigma_{xx} + \sigma_{yy} + \sigma_{zz})$, the anisotropy parameter $\delta_{CS} = \omega_0(\sigma_{zz} - \sigma_{iso})$, the asymmetry parameter $\eta_{CS} = \frac{\sigma_{yy}-\sigma_{xx}}{\sigma_{zz}-\sigma_{iso}}$, and the Euler angles $\beta$ and $\alpha$, connecting the principle axis frame with the laboratory frame.

**Dipolar coupling:** The interaction of two spins with each other due to their magnetic moment is called dipolar coupling. It is distance dependent and the corresponding dipolar Hamiltonian, $\hat{H}_D$, for each spin pair reads:

$$\hat{H}_D = -D_{ij}\frac{3\cos^2\vartheta - 1}{2}(2\hat{I}_{zi}\hat{I}_{zj} - \hat{I}_{xi}\hat{I}_{xj} - \hat{I}_{yi}\hat{I}_{yj}), \tag{2}$$

where $\vartheta$ is the angle between the internuclear vector and the external magnetic field. The dipolar coupling constant, $D_{ij}$, depends on the distance, $r_{ij}$, of the respective nuclei: $D_{ij} = \frac{\mu_0\hbar\gamma_i\gamma_j}{4\pi r_{ij}^3}$. In the heteronuclear case, due to the large differences in the resonance frequency of different isotopes the Hamiltonian simplifies to:

$$\hat{H}_D = -D_{ij}\frac{3\cos^2\vartheta - 1}{2}(2\hat{I}_{zi}\hat{I}_{zj}).$$

**Scalar coupling:** Nuclei which are connected with each other by chemical bonds interact through their shared electrons. Neglecting the very week anisotropy of this scalar or *J*-coupling and assuming a considerable chemical shift difference between the respective nuclei, the Hamiltonian for a spin pair, $\hat{H}_J$, reads:

$$\hat{H}_J = 2\pi J_{ij}\hat{I}_i\hat{I}_j,$$

where $J_{ij}$ is the scalar coupling constant.

**Relaxation:** A spin system which has been disturbed, e.g. by a radio frequency pulse, will evolve back to the equilibrium state. This process is termed relaxation. Two different types of relaxation can be distinguished. The longitudinal, spin-lattice, or $T_1$-relaxation leads to

a build-up of equilibrium magnetisation along the z-axis. It is caused by the interaction of the nuclear spins with small fluctuating magnetic fields that result from the motion and the vibrations of the molecules. The transverse, spin-spin, or $T_2$-relaxation describes the decrease of the non-equilibrium transverse magnetisation. It is the main source for the decay of the FID over time. Therefore, the line widths in NMR spectra depend crucially on $T_2$, i.e. the longer $T_2$ the narrower the line. The $T_2$-relaxation arises from the loss of coherence among the spins, precessing around the z-axis with the Larmor frequency, due to slight variations in the magnetic field.

**Phase cycling:** Due to the small nuclear energy level separation NMR is an insensitive method. To overcome this, generally many transients are recorded and added to increase the signal-to-noise ratio. Before each repetition the spin systems need to return to equilibrium. The relaxation time $T_1$, therefore, crucially affects the total time needed to record a spectrum. When several transients are recorded, the signal averaging procedure can be used to remove artefacts that occur due to imperfections in the experimental setup. Thereby radio frequency pulses of different phases in subsequent transients are applied. This technique is referred to as phase cycling. It can also be used to conveniently select desired coherence pathways.

**Multidimensional spectroscopy:** The spectral resolution can be increased by two or higher dimensional spectroscopy. At the same time nuclei that interact with each other via the described NMR interactions can be identified. Such spectra are recorded by the introduction of indirect evolution periods. They are incremented and for each increment a direct FID is recorded to sample the indirect FID. The multidimensional spectum is then obtained by Fourier-transformation of all time dimensions. This technique is especially valuable in protein NMR to obtain sufficient resolution. It is also essential when assigning spectra and to obtain structural constraints (see section 2.3).

**Isotopes:** Protons are popular in NMR spectroscopy due to their high natural abundance in combination with a large gyromagnetic ratio which results in strong signals. When investigating proteins often additional nuclei have to be used to increase the spectral resolution. The frequently employed $^{13}$C and $^{15}$N-nuclei, however, have a low natural abundance and a smaller gyromagnetic ratio which often makes isotope enrichment mandatory. In addition, in solid state

NMR spectroscopy the usage of the $^{1}$H-nucleus is very difficult due to its large anisotropic interactions, see below. Therefore, with this technique mainly the $^{13}$C and the $^{15}$N-nuclei are used. In addition, often selective labelling schemes are applied to reduce the spectral overlap and to simplify the assignment.

## 2.2 MAS NMR spectroscopy

In the solid state, the various internal interactions of the nuclear spins lead to broad lines with complicated lineshapes. Hence, solid state NMR spectra are often difficult to interpret. The main source of the observed broadening are the dipolar interaction and the chemical shift anisotropy. However, from equation 2 follows that the influence of the dipolar coupling vanishes when the internuclear vector of two spins with the external magnetic field adopts the angle $\vartheta = \arccos\frac{1}{\sqrt{3}} = 54.74°$. This holds also for the time-averaged dipolar tensor under fast rotation of the sample, where $\vartheta$ is then the angle between the rotation axis and the external field. Therefore, fast rotation at the "magic angle" $54.74°$, the so-called **m**agic **a**ngle **s**pinning (MAS), removes the dipolar interactions. However, a prerequisite for complete removal of the dipolar couplings is that the spinning frequency is much larger than the dipolar coupling constant $D$. Due to the dependence of $D$ on the gyromagnetic ratios of the involved nuclei this is easier to achieve for low $\gamma$-nuclei. MAS results also in an averaging of the chemical shift tensor. Hence, the asymmetry parameter $\eta_{CS}$ in equation 1 vanishes and the anisotropic part of the chemical shift interaction depends on $\cos^2\vartheta - 1$ in analogy to the dipolar coupling tensor. Thus the chemical shift anisotropy can also be removed by MAS and only the isotropic contribution of the chemical shielding interaction remains.

Using standard MAS rotors (3-4 mm outer diameter) spinning frequencies of 10-20 kHz are routinely achieved. They are sufficient to largely remove the chemical shift anisotropy for most $I = \frac{1}{2}$ nuclei and the dipolar couplings between low-$\gamma$ nuclei such as $^{13}$C or $^{15}$N. Incomplete averaging of the anisotropic NMR interactions due to low MAS frequencies results in additional signals on both sides of the main signal. They are referred to as spinning side bands and occur at multiples of the spinning frequency with respect to the centre signal.

Due to the strong proton-proton dipolar coupling MAS is generally not sufficient to record high resolution $^{1}$H-NMR spectra on samples with a high proton density. The influence of this coupling can be removed or attenuated by deuteration or special pulse schemes.

Nevertheless, decoupling of the homonuclear proton-proton interaction is experimentally challenging and therefore low-$\gamma$ nuclei are usually detected in MAS NMR studies. To increase the signal-to-noise ratio in $^{13}$C or $^{15}$N-detected experiments cross polarisation from the protons is used to create excess magnetisation of these nuclei. In addition, the time between two transients can be decreased in such experiments because the proton $T_1$-relaxation is generally faster than the $^{13}$C or $^{15}$N-relaxation. Furthermore, in MAS NMR spectroscopy a number of experimental schemes are used to selectively reintroduce anisotropic interactions. In the following paragraphs the most important MAS NMR techniques are described. They present the basic building blocks of the experiments used in this work, see section 3.3.

**Cross Polarisation (CP):** The dipolar coupling between two types of nuclei, here I and S, can be used to transfer magnetisation between them by applying radio frequency irradiation on both channels so that the Hartmann-Hahn condition is fullfilled: $\omega_{1\mathrm{I}} = \omega_{1\mathrm{S}}$ [Hartmann 62], where $\omega_{1\mathrm{I}}$ and $\omega_{1\mathrm{S}}$ are the nutation frequencies of the I and S spins, respectively. This leads to cross polarisation which distributes the magnetisation equally between the two types of nuclei. In a situation when I has transverse magnetisation and S the equilibrium longitudinal magnetisation, for example after a $\frac{\pi}{2}$-pulse on I, transfer of transverse magnetisation from S to I occurs. For $\gamma_{\mathrm{I}}$ larger than $\gamma_{\mathrm{S}}$ the created magnetisation of the S spins is larger than their equilibrium magnetisation. Thus, the cross polarisation scheme is routinely used to create $^{13}$C or $^{15}$N-magnetisation in biological samples. In addition, the technique can be used to transfer magnetisation between $^{13}$C and $^{15}$N-nuclei in heteronuclear correlations.

Under MAS cross polarisation occurs at the so-called side band match conditions: $\omega_{1\mathrm{I}} - \omega_{1\mathrm{S}} = \pm\omega_r, \pm 2\omega_r$ [Wu 93]. These match conditions are generally quite narrow, but in practice they are broadend by the application of a ramped contact pulse [Metz 94] on one of the channels.

**Decoupling:** MAS is generally not sufficient to remove the proton homo- and heteronuclear dipolar interactions, especially in biological material which contains many protons forming a strongly coupled spin network. During the detection of low-$\gamma$ nuclei the heteronuclear coupling to the proton network can be removed by the application of high power radio frequency irradiation on the proton channel. This results in a fast interconversion between the two proton spin states. If this rate of conversion is much larger than the heteronuclear

dipolar coupling, the low-$\gamma$ nucleus experiences a vanishing dipolar coupling. The high power decoupling performance can be improved by combining the continuous irradiation with phase modulation. Several such phase modulation schemes have been developed, e.g. **t**wo-**p**ulse **p**hase-**m**odulation (TPPM) [Bennett 95], **s**mall **p**hase **i**ncremental **al**ternation with **64** steps (SPINAL64) [Fung 00], or the XiX-scheme [Detken 02]. The performance of high power proton decoupling schemes is restricted by the amount of power the probe and the sample can take without damage. Special care has to be taken when investigating protein samples because due to their high water content severe sample heating is often observed.

Proton homonuclear decoupling schemes are mandatory to achieve reasonable resolution in heteronuclear correlations involving protons when investigating fully protonated samples. A number of experimental homonuclear decoupling techniques have been developed, e.g. the phase-modulated [Vinogradov 99] or frequency-switched [Bielecki 89] Lee-Goldberg schemes. However, it is often difficult to achieve a sufficient decoupling performance. In addition, scaling of the proton frequency axis reduces the spectral resolution. Another approach to remove the strong proton-proton couplings is the reduction of the proton density by partial deuteration. Partial deuteration, e.g. of non-exchangeable sites in proteins, can also simplify the situation with respect to the heteronuclear decoupling performance. In this case MAS can be sufficient to remove the heteronuclear dipolar couplings. For such samples sometimes low power decoupling schemes are used to remove the *J*-couplings.

**Spin diffusion:** Dipolar coupled spins exchange longitudinal magnetisation which is therefore distributed among the coupled spins in a spin system. This process is termed spin diffusion. Under MAS it is generally attenuated and does not occur among low $\gamma$-nuclei as their dipolar couplings are averaged to zero. However, strong couplings to a proton spin network broaden the signals of low $\gamma$-nuclei and enable the exchange of magnetisation. This process is referred to as **p**roton-**d**riven **s**pin-**d**iffusion (PDSD) [Szeverenyi 82] and can be accelerated by irradiating the protons at the rotary resonance conditions (one or two times the MAS frequency), a scheme referred to as **d**ipolar-**a**ssisted **r**otational-**r**esonance (DARR) [Takegoshi 01]. The efficiency of the magnetisation transfer depends on the distance of the involved nuclei and can be used to correlate nuclei which are close in space (e.g. up to 7-8 Å for $^{13}$C).

**Recoupling:** The dipolar couplings that are removed by MAS can be reintroduced by certain pulse schemes. Thereby a train of radio frequency pulses is used which inverts the magnetisation every half rotor period. This results in an incomplete averaging of the dipolar couplings and drives magnetisation transfer between the coupled spins. This can either be done in a homo- or a heteronuclear fashion and is referred to as **r**adio-**f**requency **d**riven **r**ecoupling (RFDR) [Bennett 92] or (**t**ransferred) **r**otational-**e**cho **do**uble **r**esonance (TEDOR/REDOR) [Hing 92, Gullion 89], respectively.

## 2.3 Structure determination

NMR spectroscopy is ideally suited to collect structural information at atomic resolution on complex systems, such as proteins, polymers, or inorganic materials. In the following section the assignment and structure determination process of proteins will be described. A prerequisite for studying proteins by NMR spectroscopy is a preparation that guarantees that the protein is maintained in the conformation of interest throughout the series of experiments. Examples are natively folded proteins in solution, protein crystals, and proteins in micelles or proteoliposomes.

### 2.3.1 Assignment of solid state NMR spectra

To obtain site specific information from NMR spectra, the observed isotropic chemical shifts (resonances) have to be assigned to the nuclei in the protein. In solid state NMR spectroscopy complete assignment of small proteins has become possible only during the last decade, and is still not routine. Examples are microcrystalline proteins of the $\alpha$-spectrin SH3 domain [Pauli 01], ubiquitin [Igumenova 04b], or the catabolite repression histidine-containing phosphocarrier protein (Crh) from *Bacillus subtilis* [Böckmann 03], and the non-crystalline amyloid fibrils of the HET-s prion protein [Siemer 06b]. Some larger membrane protein systems could be partially assigned, such as sensory rhodopsin II from *Natronomonas pharaonis* (NpSRII) in proteoliposomes [Etzkorn 07] and the outer membrane protein G (OmpG) from *Escherichia coli* in 2D crystalls [Hiller 08].

The established sequential assignment procedures are based on $^{13}$C-$^{13}$C and $^{13}$C-$^{15}$N correlations. A state of the art set of experiments has been assembled by e.g. Siemer et al. [Siemer 06b]. It comprises NCA and NCO experiments which were extended by a

DREAM scheme yielding N(CA)CB, N(CO)CA and N(CA)CO experiments which can also be run in a three dimensional fashion. This set of experiments is completed by DREAM $^{13}$C-$^{13}$C correlations and $CA_i(NCO)CA_{i-1}$ experiments. In ideal cases, these schemes can be extended to record four dimensional spectra, which has been demonstrated by Franks et al. [Franks 07] who recorded a CANCOCX experiment on GB1 (B1 domain of the G protein).

However, for such experimental schemes a sufficient signal-to-noise is mandatory. Therefore, also special labelling schemes such as reverse (NpSRII [Etzkorn 07]) or selective (OmpG [Hiller 08]) labelling have been applied to reduce the spectral overlap. In addition, the efficiency of the NC-transfer is often poor, because it depends on the intrinsic mobility of the sample, is very sensitive to the probe architecture, and weakened by high magnetic fields. As an alternative assignment strategy, based on direct $^{13}$C-$^{13}$C transfer, an approach was suggested by Seidel et al. [Seidel 04] to identify neighbouring residues using C$\alpha$/C$\alpha$ cross peaks.

### 2.3.2 Structural restraints and structure calculation

In many cases, first structural information can be quickly obtained from the dependence of the chemical shifts on the protein backbone conformation. This correlation has been analysed empirically and can provide dihedral angle restraits, e.g. by using the TALOS software [Cornilescu 99]. However, these restraints are predictions and have to be used with care. In addition, distance restraints can be collected from NMR experiments employing distance dependent magnetisation transfer. In protein MAS NMR spectroscopy, spin diffusion is often used to obtain correlations between nuclei distant in sequence. PDSD [Szeverenyi 82] or DARR [Takegoshi 01] experiments are used to transfer magnetisation between low-$\gamma$ nuclei. Spin diffusion among protons is probed, using nitrogen and carbon atoms for detection, in CHHC or NHHC experiments [Lange 02]. In addition, experiments which recouple the dipolar interaction, such as TEDOR [Hing 92] and REDOR [Gullion 89], can yield distance constraints [Antzutkin 02]. Using these techniques in the last years a number of protein structures were determined by MAS NMR, e.g. microcrystalline SH3 [Castellani 02] and ubiquitin [Zech 05] or the scorpion toxin kalitoxin in its free state and bound to a chimeric potassium channel (KcsA-Kv1.3) [Lange 06].

To calculate a protein structure from experimental constraints many software tools are available. They can be based on force fields, e.g. XPLOR-NIH [Schwieters 03]. A forced

heatup-cooldown annealing cycle is used to calculated a protein structure that is in accordance with the experimental constraints. Several runs are performed to cover the energy landscape. An ensemble of structures that are all in agreement with the experimental data is generated and should represent the real structure. Other software uses a distance geometry approach, where a distance matrix is employed to generate protein structures that are in agreement with the experimental data, e.g. Diana [Braun 87].

Regardless of the computational method, a well defined structural ensemble is only obtained if a large number of experimental constraints well distributed over the molecule is available. Thereby distance restraints between residues distant in sequence (so-called long-range constraints) are essential to obtain the correct overall fold.

# 3 Experiments and Methods

## 3.1 Protein preparation

The samples of CA150.WW2 studied in this work contained various isotope labelling patterns, see Table 2. They were expressed and purified either by Neil Ferguson at the MRC in Cambridge, UK (*na-CA150.WW, u-CA150.WW2, 1,3-CA150.WW2, 2-CA150.WW2, d-CA150.WW2, s-CA150.WW2*), or by Anne Diehl and Martina Leidert at the FMP (*2-CA150.WW2, la-CA150.WW2, v-CA150.WW2, ta-CA150.WW2, n-CA150.WW2*) following the same protocol. The *u-CA150.WW2* sample used for solution NMR studies was prepared by myself.

| Preparation | Isotopically labelled compounds |
|---|---|
| *na-CA150.WW2* | none |
| *u-CA150.WW2* | u-$^{13}$C-glucose, $^{15}NH_4Cl$ |
| *1,3-CA150.WW2* | 1,3-$^{13}$C-glycerol, $^{15}NH_4Cl$ |
| *2-CA150.WW2* | 2-$^{13}$C-glycerol, $^{15}NH_4Cl$ |
| *d-CA150.WW2* | u-$^{13}$C,$^{2}H_6$-glucose, $^{15}NH_4Cl$, $D_2O$ |
| *s-CA150.WW2* | u-$^{13}$C,$^{15}$N-Ala, u-$^{13}$C,$^{15}$N-Gly, 2,3-$^{13}$C,$^{15}$N-Phe, 2,3-$^{13}$C,$^{15}$N-Tyr |
| *la-CA150.WW2* | 3,4-$^{13}$C-Leu, 1-$^{13}$C-Ala |
| *v-CA150.WW2* | 2-$^{13}$C-glycerol, natural abundant Phe, Leu, Lys, Tyr |
| *ta-CA150.WW2* | u-$^{13}$C-Thr, 1-$^{13}$C-Ala |
| *n-CA150.WW2* | 100% $^{12}$C-glucose, $^{15}NH_4Cl$ |

Table 2: List of CA150.WW2 preparations with the isotopically labelled material used during recombinant expression in *Escherichia coli*.

CA150.WW2 Y19F was expressed in the *Escherichia coli* strains BL21(DE3) or Rosetta2 (both from Novagen) using a pGAT2 expression vector [Peranen 96]. The vector was generated by PCR (polymerase chain reaction) from the wild type construct (GenBank accession code U40749, [Macias 00]), by Neil Ferguson, MRC, Cambridge, UK. This introduced an N-terminal HIS (6*histidine) - GST (glutathione-S-transferase) extension which was subsequently removed by proteolytic cleavage with thrombin yielding the desired domain including a short N-terminal extension: GSM-CA150.WW2 Y19F. In the following, CA150.WW2 will always refer to this 40-residue construct (4.77 kDa): GSM GATAVSEWTE YKTADGKTFY YNNRTLESTW EKPQELK. The sequence numbering used in this work ranges from -2 to 0 for the GSM extension and from 1 to 37 for the WW-domain.

Here, the protocol used by myself is described. It relies on the wealth of experience gained

by Neil Ferguson, Anne Diehl, and Martina Leidert on this system. The protocol varied slightly for the different preparations, see below.

### 3.1.1 Expression protocol

The vector was transformed into the *Escherichia coli* strain Rosetta2. A clone which showed expression of CA150.WW2 in rich as well as in minimal media was chosen for an overnight starter culture in rich medium at 37°C. All media contained the antibiotics chloramphinicol (3.4 $\mu$g/mL) and carbenicillin (60 $\mu$g/mL). In some of the other preparations, the BL21(DE3) strain was used which showed similar levels of expression. Generally, expression was better when freshly transformed cells were used. 2 L of rich medium were seeded to yield an optical density of $\approx 0.1$ at 600 nm. At 37°C the cultures were grown to an optical density of $\approx 0.7$, then the cells were condensed by centrifugation at 750 g for 20 minutes. The pellets were resuspended in 1 L M9-medium containing 2 g $^{13}$C-glucose and 1 g $^{15}NH_4Cl$. This change from rich to minimal medium [Marley 01] reduces the consumption of isotopically enriched reagents and has been done in most preparations, but is not strictly necessary. Proteins with a different isotope labelling pattern, Table 2, were obtained using other isotopically labelled compounds in concentrations of 2 g/L glycerol and 0.2 g/L labelled or unlabelled amino acids. In preparations with isotopically labelled amino acids, the remaining standard amino acids were added to the medium in natural abundance to suppress the bacterial metabolism of the amino acids. For the preparation of the deuterated sample, the expression was done in $D_2O$-medium. All subsequent steps were carried out in buffers containing $H_2O$, thus leaving deuterium atoms only at non-exchangeable sites in the protein. After one hour in minimal medium the culture was induced with IPTG (isopropyl $\beta$-D-1-thiogalactopyranoside). Expression was done over night at 25°C or in some other cases at 30°C. Cells were harvested by centrifugation (5000 g) at 8°C, washed with 150 mM NaCl solution, and stored at -80°C when purification could not be done immediately.

### 3.1.2 Purification protocol

The cell pellet was suspended in a buffer (phosphate, HEPES, or Tris/HCl) at pH 8.0 containing EDTE-free protease inhibitors, benzonase, and 1 mM $MgCl_2$. The cells were ruptured in a homogenisation machine. In other cases, french press or sonification was used. After

centrifugation (58500 g) the lysate was purified using a Ni-affinity column in a phosphate buffer system at pH 8.0. In other preparations HEPES or Tris/HCl buffers were used. The protein containing fractions were pooled and concentrated. GST was cleaved by incubation with thrombin (100 units/20 mg fusion protein) at 8°C. After filtration the digest solution was applied to a gel filtration column (Superdex 75, GE Healthcare) using a 200 mM $(NH_4)HCO_2$ solution. The CA150.WW2 containing fractions were pooled and lyophilisated to yield 2.3 mg of protein. Generally, the preparations yielded between 2 and 10 mg CA150.WW2 per litre minimal medium. The identity of the purified protein and the isotope labelling were confirmed by MALDI (**m**atrix-**a**ssisted **l**aser **d**esorption/**i**onization) mass spectrometry. However, the gel filtration step was not completely reproducible between the two laboratories involved despite the usage of the same column from the same vendor. While the preparations at the MRC in Cambridge yielded nearly salt free CA150.WW2, at the FMP some other compounds eluted together with the protein. Solution NMR analysis showed that these non-proteinatious components contained large amounts of Tris, when a Tris/HCl buffer was used for Ni-affinity chromatography. Using the phosphate buffer system, large amounts of imidazole were eluted together with CA150.WW2.

## 3.2 CA150.WW2 fibrillation and solid state NMR sample preparation

The fibrillation procedure of CA150.WW2 amyloid fibrils was established by Neil Ferguson at the MRC in Cambridge [Ferguson 03]. An overview on the prepared solid state samples is given in Table 3. They were prepared in Cambridge and I was involved in the preparation of vl-CA150.WW2 and nta-CA150.WW2 during a visit there.

| Solid state NMR sample | Protein preparation | MAS rotor |
|---|---|---|
| u-CA150.WW2 | *u-CA150.WW2* | 3.2 mm |
| 1,3-CA150.WW2 | *1,3-CA150.WW2* | 4.0 mm |
| 2-CA150.WW2 | *2-CA150.WW2* | 4.0 mm |
| d-CA150.WW2 | *d-CA150.WW2* | 3.2 mm |
| s-CA150.WW2 | *s-CA150.WW2* | 3.2 mm |
| 2dil-CA150.WW2 | 20% *2-CA150.WW2* and 80% *na-CA150.WW2* | 3.2 mm |
| vl-CA150.WW2 | 40% *v-CA150.WW2* and 60% *la-CA150.WW2* | 3.2 mm |
| nta-CA150.WW2 | 50% *n-CA150.WW2* and 50% *ta-CA150.WW2* | 3.2 mm |

Table 3: Solid state NMR samples used in this work, the isotope labelling of the protein preparations is given in Table 2.

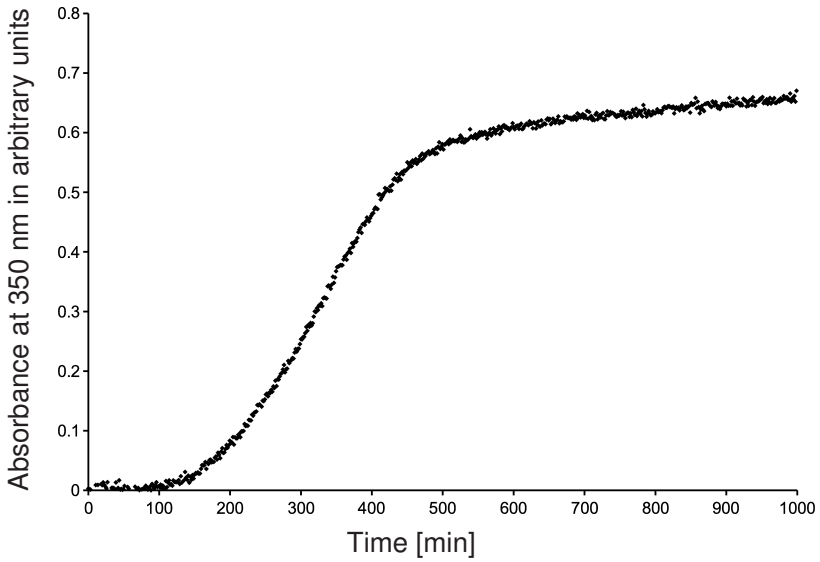



Figure 6: Growth curve of d-CA150.WW2 amyloid fibrils.

In order to grow amyloid fibrils lyophilisated protein was dissolved on ice over one hour in 10 mM sodium phosphate buffer pH 7.0 with 0.02% w/v $NaN_3$. Then the solution was treated briefly in an ultrasonic bath to ensure complete dissolution. The protein concentration was determined from the absorbance at 280 nm. This value was corrected for the scattering contribution which was estimated by extrapolaiting the absorbance between 320 and 350 nm to 280 nm. The protein solutions were then diluted to a concentration of 100 $\mu$M. The solutions were incubated in sealed cuvettes, 3.5 ml each, at 37°C and continuously agitated by a rapidly moving stirring bar. To avoid seeding, new stirring bars and cuvettes were used for each preparation. Fibril growth was monitored by the absorbance at 350 nm. Figure 6 shows an example growth curve, recorded during the preparation of d-CA150.WW2. An initial lag time, typical for the nucleation process in amyloid formation, was followed by rapid fibril growth until the absorbance reached a plateau after eight hours. The fibril suspensions were kept at these conditions for a further six days to obtain mature amyloid fibrils. They were harvested by centrifugation, resuspended in little buffer, and stored at -80°C.

At the MRC in Cambridge, Thioflavin T staining of CA150.WW2 amyloid fibrils was investigated with an Amincon-Bowman SLM series 2 luminescence spectrometer (SLM Instruments, Urbana, IL). Samples were excited at 440 nm and their emission was recorded at 460 to 650 nm. Preformed fibrils were added to a 65 $\mu$M thioflavin T solution, both in 20 mM sodium phosphate buffer at pH 7.0. A strong increase in the emission intensity at 482 nm was characteristic for amyloid formation [LeVine 93]. Low-dose uranyl acetate cryoelectron

micrographs and optical diffraction pattern were recorded on a FEI G2 F20 field emission gun electron microscope at 200 keV by John Berriman at the SBC in New York.

To fill the MAS rotors, the samples were thawed and concentrated by centrifugation. This was done by gradually increasing the centrifugal force on the samples from 5000 to 60000 g over 5-6 steps. Before each increase of the spinning frequency, the supernatant was removed. For each fibril preparation around 15 mg of protein was centrifuged into MAS rotors with an outer diameter of 4.0 or 3.2 mm, see Table 3. The larger rotors contained spacers which insured that the sample was positioned in the rotor centre. Before and between NMR measurements the samples were stored in their MAS rotors at 4°C. Fibrils which were produced from protein preparations containing large amounts of salt (vl-CA150.WW2 and nt-CA150.WW2) were washed with the fibrillation buffer before the final concentration steps.

## 3.3 NMR spectroscopy

### 3.3.1 Solution state NMR experiments

*u-CA150.WW2* was dissolved in 600 $\mu$l 20 mM sodium phosphate buffer, pH 7.0, containing 10% $D_2O$ and 0.02% w/v $NaN_3$ to yield a 0.8 mM solution. On a Bruker DRX 600 MHz spectrometer equipped with a cryo-probe the following experiments were recorded at 285 K: 2D-HSQC, 3D-HN(CO)CACB, 3C-HNCACB, 3D-HNCO, 3D-HN(CO)CC-TOCSY. Pulsed field gradients were employed for water suppression and the experimental schemes by Sattler et al. [Sattler 99] were used.

### 3.3.2 MAS NMR experiments

A large number of different MAS NMR experiments were recorded on the different samples at magnetic fields ranging from 400 to 900 MHz. Each experimental scheme is briefly described here, base on the building blocks discussed in section 2. The employed pulse sequences are shown in Figure 7, with the exception of the 2D-TEDOR and 3D-HNC-TEDOR experiments which are given in Figure 18 on page 58. If not stated otherwise, all spectra were recorded using ramped cross polarisation from protons ($\sim$ 60 kHz) to $^{13}$C or $^{15}$N during 750-2000 $\mu$s; high power proton TPPM [Bennett 95] or SPINAL64 [Fung 00] decoupling at 80-85 kHz during recoupling pulse trains, direct, and indirect evolution; a nominal temperature of 285 K; and a

recycle delay of 3 s. The applied phase cycling schemes are given in the caption of Figure 7. Table 4 lists the recorded MAS NMR experiments.

**PDSD/DARR:** After CP-excitation and $^{13}$C-evolution, longitudinal $^{13}$C-magnetisation is created by a $\frac{\pi}{2}$-pulse. During $t_{mix}$ PDSD [Szeverenyi 82] or DARR mixing [Takegoshi 01] follows before another $^{13}$C $\frac{\pi}{2}$-pulse and detection. DARR irradiation was done at the $\omega_{DARR} = \omega_r$ condition.

**RFDR:** $^{13}$C-magnetisation is excited directly by a $^{13}$C $\frac{\pi}{2}$-pulse. Bracketed by a pair of $^{13}$C $\frac{\pi}{2}$-pulses, a radio-frequency driven recoupling (RFDR) [Bennett 92] pulse train is used during $t_{mix}$ to recouple the $^{13}$C homonuclear dipolar couplings. Then $^{13}$C-magnetisation is detected. A recycle delay of 5 s was used. In this work the sequence was only applied to the deuterated sample (d-CA150.WW2) at 20 kHz MAS. Therefore, only low power proton decoupling was applied during evolution and recoupling periods using the WALTZ16 scheme [Shaka 83] at 5 kHz.

**NCA/NCO and NCACX/NCOCX:** After CP-excitation of $^{15}$N-magnetisation and $^{15}$N-evolution the magnetisation is transferred to carbon by a selective cross polarisation step. This was established either by SPECIFIC CP [Baldus 98] or an adiabatic-passage CP [Baldus 96]. Depending on the frequency offset of the $^{13}$C-irradiation this results in selective magnetisation transfer to the CO or the C$\alpha$-sites. NCA or NCO spectra are obtained by detecting this $^{13}$C-magnetisation. NCACX or NCOCX spectra are recorded by inserting a PDSD step before detection.

**NHC:** Nitrogen and carbon atoms can be correlated via protons which are close to both nuclei. Here, the scheme introduced by Etzkorn et al. [Etzkorn 04] was used. $^{15}$N-magnetisation was generated by CP followed by $^{15}$N-evolution. After a z-filter, a short CP-step is used to transfer magnetisation to the protons directly connected to nitrogen sites. A final CP step is applied to transfer magnetisation from these protons to nearby $^{13}$C-sites.

**HC/HN:** The HC or HN experiments are started with a proton evolution period which is followed by a CP-transfer from the protons to the $^{13}$C or $^{15}$N-nuclei, respectively. These

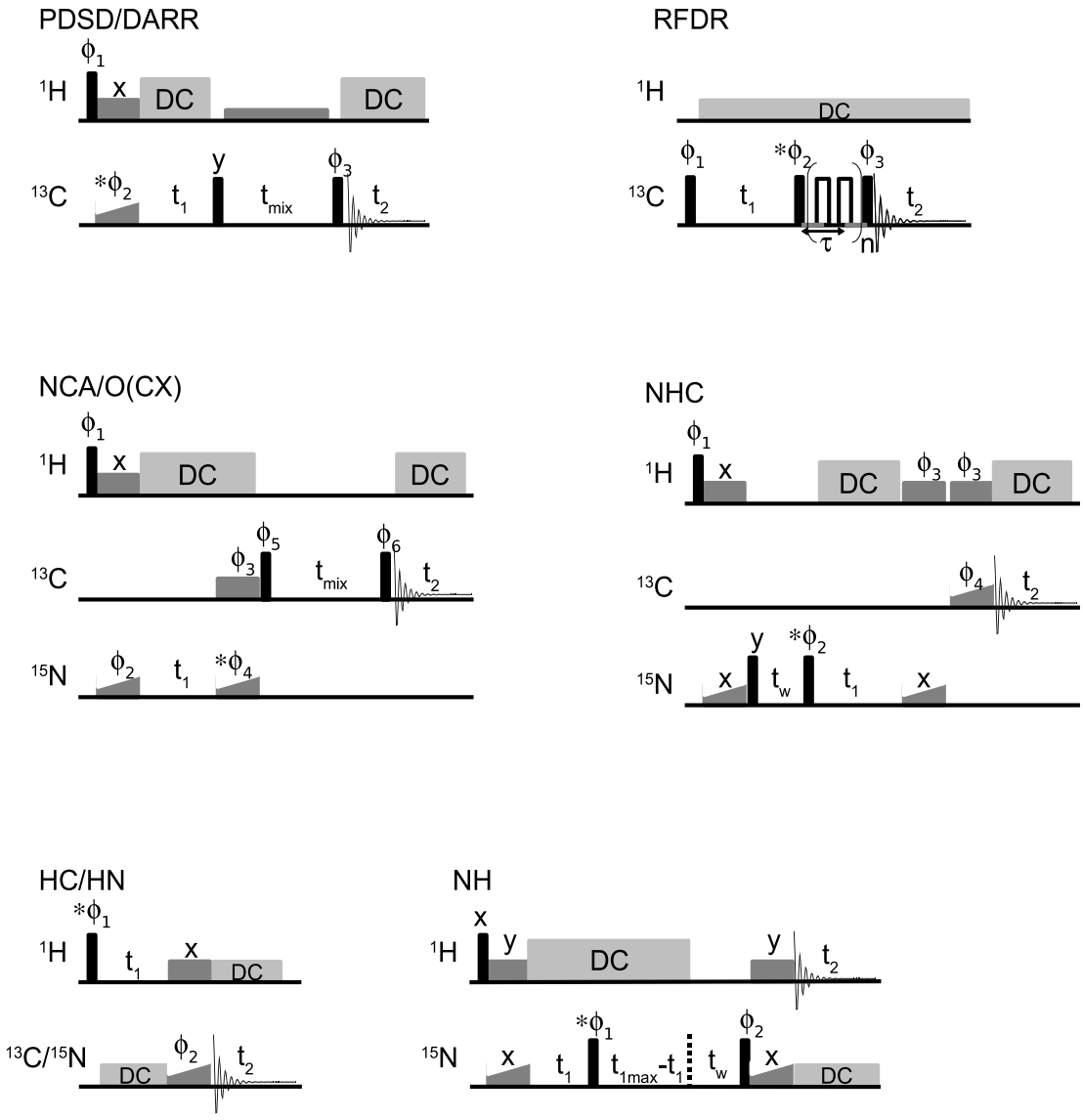



Figure 7: Pulse sequences used in this work with the exception of the TEDOR pulse sequences which are given in Figure 18 on page 58. $t_1$ and $t_2$ are the direct or indirect evolution times, filled bares represent $\frac{\pi}{2}$-pulses, open bars $\pi$-pulses. Decoupling times are indicated by DC-boxes, all other boxes represent CP-irradiation. $\tau$ is the length of a rotor period, $t_{mix}$ is a mixing time and $t_w$ is a water dephasing period. The phase that is incremented during TPPI (**T**ime-**P**roportional **P**hase **I**ncrementation) [Marion 83] indirect detection is indicated by an asterisk. PDSD/DARR: Proton irradiation during $t_{mix}$ was only applied for DARR experiments, $\phi_1 = y\overline{y}$, $\phi_2 = xxxx\ xxxx\ \overline{xxxx}\ \overline{xxxx}$, $\phi_3 = yy\overline{xx}\ \overline{yy}xx$, $\phi_{rec} = \overline{x}x\overline{y}y\ x\overline{x}y\overline{y}\ \overline{x}x\overline{y}y\ x\overline{x}\overline{y}y$; RFDR: 5 s recycle delay, $\phi_1 = \overline{y}y$, $\phi_2 = xxxx\ xxxx\ \overline{xxxx}\ \overline{xxxx}$, $\phi_3 = \overline{yy}yy\ xx\overline{xx}$, $\phi_{rec} = \overline{x}xx\overline{x}\ \overline{y}yy\overline{y}\ x\overline{x}\overline{x}x\ y\overline{y}\overline{y}y$, $\pi$-pulses $y\overline{x}y\overline{x}\ \overline{x}y\overline{x}y$; NCA/O(CX): Detection directly after the $^{13}$C CP-pulse yield NCA/O spectra, NCA/OCX spectra are recorded using the full sequence, $\phi_1 = y\overline{y}$, $\phi_2 = xx\overline{xx}$, $\phi_3 = xxxx\ \overline{xxxx}$, $\phi_4^{NCA/O} = x\overline{x}x\overline{x}\ \overline{x}x\overline{x}x$, $\phi_{rec}^{NCA/O} = xx\overline{xx}$, $\phi_4 = x\overline{x}x\overline{x}$, $\phi_5 = y\overline{y}$, $\phi_6 = y\overline{y}y\overline{y}\ \overline{x}x\overline{x}x\ \overline{y}y\overline{y}y; x\overline{x}x\overline{x}$, $\phi_{rec} = \overline{xx}xx\ yy\overline{yy}\ xx\overline{xx}\ \overline{yy}yy$; NHC: $\phi_1 = yyyy\ yyyy\ yyyy\ yyyy\ \overline{yyyy}\ \overline{yyyy}\ \overline{yyyy}\ \overline{yyyy}$, $\phi_2 = \overline{y}y$, $\phi_3 = xxyy\ \overline{xxyy}$, $\phi_4 = xxyy\ \overline{xxyy}\ \overline{xxyy}\ xxyy$, $\phi_{rec} = x\overline{x}y\overline{y}\ \overline{x}x\overline{y}y\ \overline{x}x\overline{y}y\ x\overline{x}y\overline{y}\ \overline{x}x\overline{y}y\ x\overline{x}y\overline{y}\ x\overline{x}y\overline{y}\ \overline{x}x\overline{y}y$; HC/HN: $\phi_1 = y\overline{y}$, $\phi_2 = xx\overline{xx}\ yy\overline{yy}$, $\phi_{rec} = x\overline{xx}x\ y\overline{yy}y$; NH: $\phi_1 = yy\overline{yy}$, $\phi_2 = \overline{y}y$, $\phi_{rec} = y\overline{yy}y$.

| **2D $^{13}$C-$^{13}$C correlations** | | | | | |
|---|---|---|---|---|---|
| Experimental scheme | Sample | $t_{mix}$ | MAS-frequency | $^{1}$H-frequency | Result section |
| PDSD | u-CA150.WW2* | 20 ms | 10.5 kHz | 900 MHz | 4.2 |
| | u-CA150.WW2* | 100 ms | 10.5 kHz | 900 MHz | 4.3.2 |
| | 1,3-CA150.WW2 | 20 ms | 10.5 kHz | 900 MHz | 4.2 |
| | 1,3-CA150.WW2 | 100 ms | 12.0 kHz | 700 MHz | 4.3.2 |
| | 1,3-CA150.WW2* | 100 ms | 11.0 kHz | 900 MHz | 4.3.2 |
| | 1,3-CA150.WW2 | 200 ms | 10.5 kHz | 900 MHz | 4.3.2 |
| | 2-CA150.WW2 | 20 ms | 10.5 kHz | 900 MHz | 4.2 |
| | 2-CA150.WW2 | 100 ms | 12.0 kHz | 700 MHz | 4.3.2 |
| | 2-CA150.WW2 | 100 ms | 10.5 kHz | 900 MHz | 4.3.2 |
| | 2-CA150.WW2 | 200 ms | 10.5 kHz | 900 MHz | 4.3.2 |
| | 2-CA150.WW2 | 400 ms | 10.5 kHz | 900 MHz | 4.5.2 |
| | d-CA150.WW2 | 20 ms | 10.5 kHz | 900 MHz | 4.2 |
| | s-CA150.WW2 | 20 ms | 10.5 kHz | 900 MHz | 4.3.2 |
| | 2dil-CA150.WW2 | 400 ms | 10.5 kHz | 900 MHz | 4.7.1 |
| | vl-CA150.WW2 | 400 ms | 10.5 kHz | 900 MHz | 4.7.1 |
| DARR | 1,3-CA150.WW2 | 400 ms | 10.5 kHz | 900 MHz | 4.3.2 |
| | 1,3-CA150.WW2 | 700 ms | 12.0 kHz | 700 MHz | 4.5.2 |
| | 2-CA150.WW2 | 700 ms | 12.0 kHz | 700 MHz | 4.3.2 |
| | d-CA150.WW2 | 400 ms | 10.5 kHz | 900 MHz | 4.3.2 |
| | d-CA150.WW2 | 700 ms | 10.5 kHz | 900 MHz | 4.3.2 |
| RFDR | d-CA150.WW2 | 1.2 ms | 20.0 kHz | 900 MHz | 4.2 |
| | d-CA150.WW2 | 4.8 ms | 20.0 kHz | 900 MHz | 4.3.2 |
| **2D $^{15}$N $^{13}$C-correlations** | | | | | |
| Experimental scheme | Sample | $t_{mix}$ | MAS-frequency | $^{1}$H-frequency | Result section |
| NCA | 1,3-CA150.WW2 | - | 10.0 kHz | 400 MHz | 4.2 |
| NCO | 1,3-CA150.WW2 | - | 10.0 kHz | 400 MHz | 4.2 |
| NCACX | 1,3-CA150.WW2 | 100 ms | 10.0 kHz | 400 MHz | 4.2 |
| NCOCX | 1,3-CA150.WW2 | 100 ms | 10.0 kHz | 400 MHz | 4.2 |
| TEDOR | 2-CA150.WW2 | 1.6 ms | 10.0 kHz | 400 MHz | 4.2 |
| | 2-CA150.WW2 | 12.8 ms | 10.0 kHz | 400 MHz | 4.5.2 |
| | d-CA150.WW2 | 0.8 ms | 10.0 kHz | 900 MHz | 4.2 |
| **Other experiments** | | | | | |
| Experimental scheme | Sample | $t_{CP/mix}$ | MAS-frequency | $^{1}$H-frequency | Result section |
| HC | d-CA150.WW2 | 1.0 ms | 20.0 kHz | 600 MHz | 4.3.3 |
| HN | d-CA150.WW2 | 2.0 ms | 20.0 kHz | 900 MHz | 4.3.3 |
| NH-Zilm | d-CA150.WW2 | 1.0 ms | 20.0 kHz | 900 MHz | 4.3.3 |
| HNC-TEDOR | d-CA150.WW2 | 1.2 ms | 20.0 kHz | 600 MHz | 4.3.4 |
| NHC | nta-CA150.WW2 | 2.0 ms | 10.0 kHz | 400 MHz | 4.7.1 |

Table 4: Overview on the solid state NMR experiments described in this work, * indicates spectra that were recorded by Jeremy Flinders.

experiments were recorded on the deuterated sample (d-CA150.WW2) at 20 kHz MAS, and thus no high power proton decoupling was used. The WALTZ16 scheme [Shaka 83] at 5 kHz was applied to the $^{13}$C, the $^{15}$N, or the $^{1}$H-channel during indirect and direct evolution.

**$^{1}$H-detected NH:** This experiment relies on 20 kHz MAS and deuteration for proton-proton decoupling during direct detection and employed the scheme by Paulson et al. [Paulson 03] for water suppression. $^{15}$N-magnetisation is created by cross polarisation and then during $^{15}$N-evolution high power proton decoupling is applied. After evolution, longitudinal $^{15}$N-magnetisation is created but decoupling continues, to run the experiment in a constant time fashion. Then high power proton decoupling is switched off to dephase the water magnetisation during $t_w$, typically 4 ms. Finally, $^{15}$N-magnetisation is transferred back to the transverse plane and a cross polarisation step creates $^{1}$H-magnetisation. Low power decoupling is applied on the $^{15}$N-channel during proton detection.

**2D-NC-TEDOR:** The TEDOR experiment follows the schemes published by Michal et al. [Michal 97] and Jaroniec et al. [Jaroniec 00]. $^{13}$C-magnetisation is transferred to $^{15}$N by a REDOR pulse train [Gullion 89] during $t_{mix}$ followed by a pair of $\frac{\pi}{2}$-pulses on the $^{13}$C and the $^{15}$N-channels. After $^{15}$N-evolution the magnetisation is transferred back to $^{13}$C for detection by a second pair of $\frac{\pi}{2}$-pulses and another REDOR pulse train. The symmetric arrangement of the pulse trains refocuses the $^{13}$C homonuclear dipolar couplings which leads to an efficient REDOR transfer when applied to multiply labelled samples.

**3D-HNC-TEDOR:** In the course of this work, a three-dimensional HNC-TEDOR experiment, including a proton dimension, was assembled. The experiment is described in detail in section 4.3.4 on page 57.

## 3.4 Data analysis and structure calculation

The $^{1}$H-chemical shift in the solution NMR spectra was referenced externally to DSS (2,2-dimethyl-2-silapentane-5-sulfonic acid) and was then used for indirect referencing of the $^{13}$C and $^{15}$N-dimensions [Wishart 95]. $^{13}$C-referencing of the solid state MAS NMR spectra was done externally using the upfield signal in the adamantane spectrum at 31.5 ppm which

corresponds to a referencing to DSS [Morcombe 03]. From that $^{15}$N was referenced indirectly [Wishart 95]. $^{1}$H-referencing was done using the water signal at 4.8 ppm.

NMR spectra were acquired and processed with XWINNMR 2.6, TOPSPIN 1.3, or TOPSPIN 2.0 (all Bruker, Karlsruhe, Germany). The assignment was done with CCPN analysis [Vranken 05], and figures showing NMR spectra were produced with Sparky 3.1 (T. D. Goddard, D. G. Keller, University of California, San Francisco, USA). Chemical shift analysis was done with TALOS [Cornilescu 99].

Fibril structures were calculated using Xplor-NIH 2.16 [Schwieters 03]. To calculate an oligomeric fibril, translational symmetry was enforced between the monomeric units. For each structure calculation 200 runs were done and 1000 cycles were used for annealing.

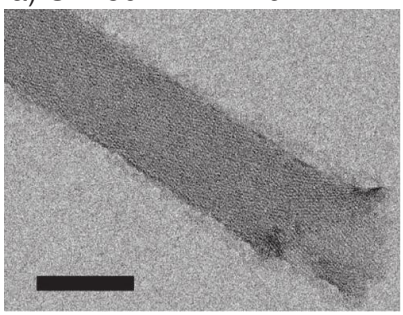
# 4 Results

In this work, structure determination of CA150.WW2 Y19F amyloid fibrils by MAS NMR is described. Fibrils formed by the CA150.WW2 Y19F variant were selected for the structural studies based on their morphology, fibrillation kinetics, and initial NMR experiments, section 4.1. In the following, CA150.WW2 refers to this variant if not stated otherwise. By recombinant expression in *Escherichia coli*, a set of differently isotopically labelled samples was prepared. The MAS NMR spectra showed intermediate line widths compared to other amyloid fibrils, section 4.2. Due to the low resolution in $^{13}$C-$^{15}$N correlations the assignment of the spectra using established procedures [Pauli 01, Siemer 06a] was hindered. Instead, an alternative strategy was developed making use of the isotope labelling pattern in glycerol-labelled samples to establish a sequential assignment using $^{13}$C-$^{13}$C correlations, section 4.3. Furthermore, $^{13}$C-$^{13}$C correlations of a triply labelled sample ($^{2}$H,$^{13}$C,$^{15}$N-labelled, $^{1}$H back-exchanged) were employed and showed improved line widths. This sample was also used in a new NMR experiment (HNC-TEDOR) for the assignment of 35% of the amide nitrogen and amide proton sites. The backbone assignment of the natively folded CA150.WW2 in solution was determined for comparison of the WW with the amyloid fold, section 4.4. In spectra of CA150.WW2 amyloid fibrils, long-range distance constraints were obtained from $^{13}$C-$^{13}$C correlations with long mixing times. A salt-bridge was detected by a two dimensional TEDOR experiment, section 4.5. These constraints were combined with data from TALOS dihedral angle predictions, electron microscopy, and alanine scanning to determine the fold of the monomeric unit in CA150.WW2 amyloid fibrils, section 4.6. Quaternary contacts were probed using mixtures of differently isotopically labelled samples. Finally, three CA150.WW2 amyloid fibril structures with slightly different geometries were calculated, section 4.7.

## 4.1 The CA150.WW2 Y19F variant and MAS NMR sample preparation

The preparation of conformationally homogeneous samples is crucial for protein structure determination. This is of particular importance when investigating amyloid fibrils, because they often show various morphologies in electron micrographs and are inherently non-crystalline. Initial MAS NMR studies on fibrils formed by the wild type CA150.WW2, yielded spectra with broad lines indicating such sample heterogeneities. Different morphologies and structural

a) CA150.WW2 Y19F



b) CA150.WW2 wild type

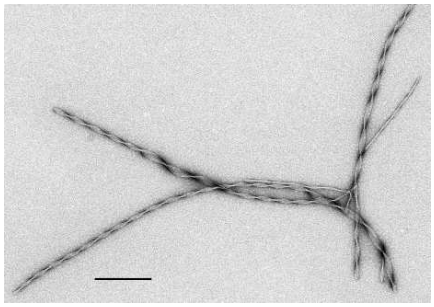


Figure 8: Low-dose uranyl acetate electron micrographs of negatively stained CA150.WW2 a) Y19F and b) wild type fibrils, scale bars are 100 nm and 200 nm, respectively.

homogeneity can be triggered by small sequence variations like point mutations or truncation of terminal residues. Therefore, a number of CA150.WW2 variants (alanine and glycine mutants, and tyrosine or tryptophan to phenylalanine mutants [Petrovich 06]) were investigated by Neil Ferguson and Henning Tidow at the MRC in Cambridge, UK, concerning their suitability for MAS NMR studies. The morphologies of the fibrils of these variants grown at 37°C and pH 7.0 were monitored by electron microscopy by John Berriman at the New York Structural Biology Center. Especially large and homogeneous fibrils were formed by the Y19F variant, Figure 8a). They build large tubular associations which comprise about hundred protofilaments [Ferguson 06]. In contrast, in case of wild type CA150.WW2, 3-18 protofilaments associate laterally to form a twisted ribbon [Ferguson 03], see Figure 8b). The large regular structure of the Y19F fibrils indicates that the single protofilaments are in a highly ordered environment. In addition, their fibrillation kinetic was found to be extremely reproducible (oral communication Neil Ferguson). The cross-$\beta$ optical diffraction pattern of oriented Y19F tubes, see Figure 22 on page 66, and their staining reaction with thioflavin T showed that they are indeed amyloid fibrils. Most importantly, MAS NMR spectra of the CA150.WW2 Y19F fibrils showed signals with an improved line width compared to those of the wild type fibrils. As a result of these initial studies, the Y19F variant was chosen for the structural investigations by MAS NMR.

For these studies, isotopically labelled CA150.WW2 was prepared by recombinant expression as a GST fusion protein in *Escherichia coli* (section 3.1.1). During purification (section 3.1.2) GST was cleaved leaving a small N-terminal extension (GSM). Several differently labelled samples were prepared. A uniformly $^{13}$C,$^{15}$N-labelled sample (u-CA150.WW2) was used to assess the suitability of the CA150.WW2 fibril preparation for MAS NMR studies. To simplify the assignment procedure and facilitate the collection of long range distance restraints,

two samples, 1,3-CA150.WW2 and 2-CA150.WW2, were prepared from bacteria grown on $^{15}NH_4Cl$ and 1,3-$^{13}$C-glycerol or 2-$^{13}$C-glycerol as the only nitrogen and carbon sources, respectively. In addition, a $^{2}$H,$^{13}$C,$^{15}$N-labelled, $^{1}$H back-exchanged sample (d-CA150.WW2) was prepared to facilitate proton MAS NMR spectroscopy and to ease proton-carbon decoupling.

To address specific questions with respect to the assignment, structural homogeneity (section 4.3), and the quaternary structural arrangement (section 4.7.1), specifically labelled fibrils and fibrils from mixtures of differently labelled proteins were prepared. s-CA150.WW2 was **s**electively labelled by feeding the bacteria u-$^{13}$C,$^{15}$N-alanine, u-$^{13}$C,$^{15}$N-glycine, 2,3-$^{13}$C, $^{15}$N-phenylalanine, and 2,3-$^{13}$C,$^{15}$N-tyrosine. Another sample, 2dil-CA150.WW2, was prepared by **dil**uting the fibrillation solution of **2**-$^{13}$C-glycerol-labelled CA150.WW2 with unlabelled CA150.WW2 in a 1:4 ratio. vl-CA150.WW2 is also a mixture (2:3) of two differently labelled peptides: a mainly 2,3-$^{13}$C-**v**aline-labelled sample (made by reverse labelling of a 2-$^{13}$C-glycerol sample with unlabelled phenylalanine, leucine, lysine, and tyrosine) was mixed with selectively 1-$^{13}$C-alanine, 3,4-$^{13}$C-**l**eucine labelled CA150.WW2. Finally the mixed sample nta-CA150.WW2 was prepared from $^{15}$**N**-CA150.WW2 and u-$^{13}$C-**t**hreonine, 1-$^{13}$C-**a**lanine selectively labelled CA150.WW2 in a 1:1 ratio. An overview on all employed MAS NMR samples is given in Table 3 on page 28.

## 4.2 Initial characterisation of CA150.WW2 fibrils by MAS NMR

The suitability and the reproducibility of the CA150.WW2 amyloid sample preparation was investigated by MAS NMR $^{13}$C-$^{13}$C correlations. These spectra can be recorded using different excitation schemes which vary with respect to their dependence on the mobility in the sample. CP-transfer from the protons to the carbons is only efficient for rigid parts of the sample, whereas extremely mobile residues are detectable employing solution NMR methods such as INEPT, which was shown for HET-s fibrils [Siemer 06a]. Direct excitation of the $^{13}$C-nuclei is not sensitive to the mobility within the sample, but suffers from the low gyromagnetic ratio of the $^{13}$C-nucleus and its long $T_1$-relaxation times.

The CP-transfer scheme is most efficient for the immobile residues in the amyloid core region. It was used to record 20 ms $^{13}$C-$^{13}$C PDSD spectra on the samples u-CA150.WW2, 1,3-CA150.WW2, and 2-CA150.WW2 at 10.5 kHz MAS, see Figure 9a-c). At this short mixing

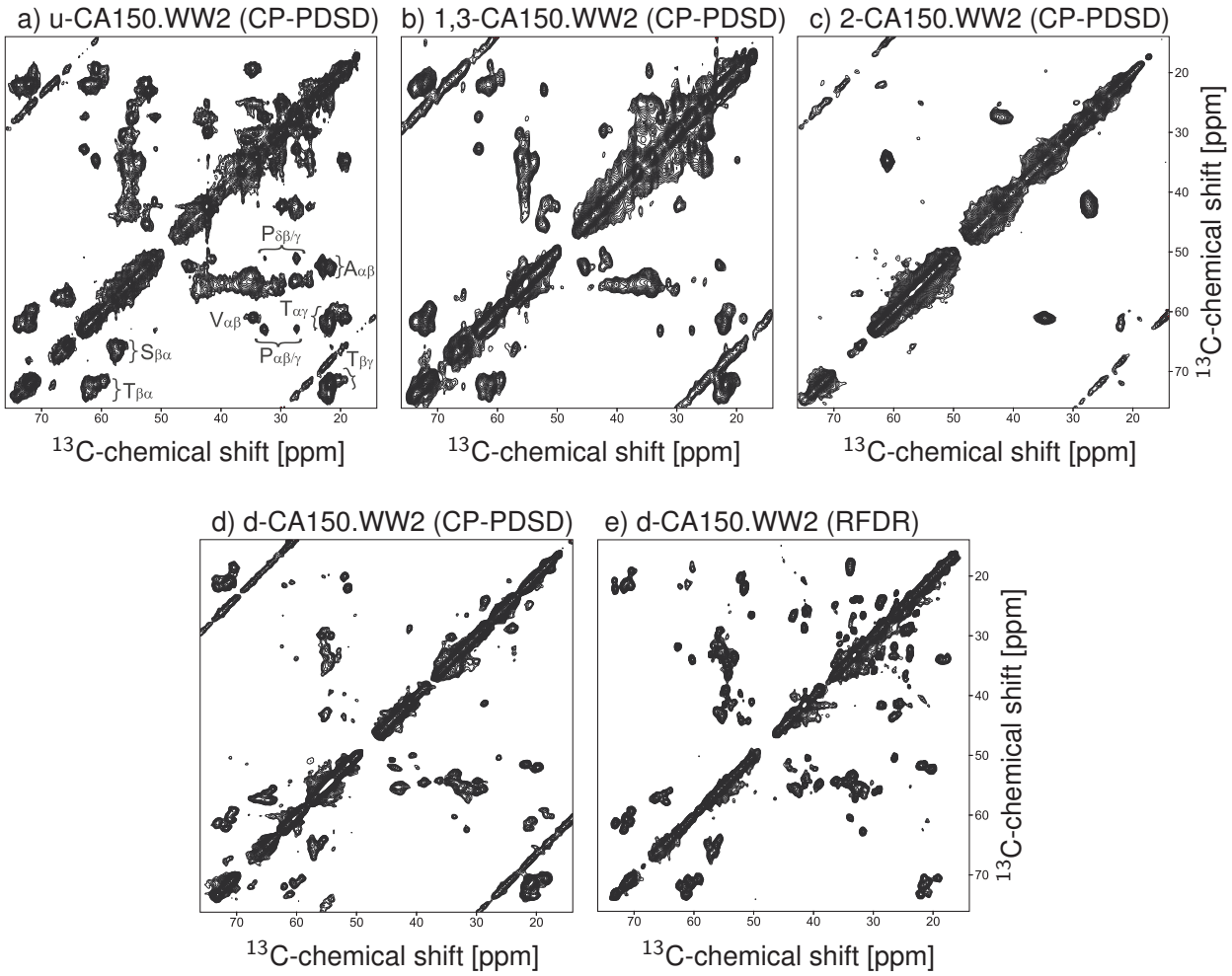



Figure 9: $^{13}$C-$^{13}$C correlation spectra from differently labelled samples. a-d) CP-PDSD experiments with a mixing times of 20 ms at 10.5 kHz MAS. e) 1.2 ms RFDR spectrum with direct excitation at 20 kHz MAS.

time only correlations between signals from sites very close in space, i.e. intra residue, are expected. The spectrum of u-CA150.WW2 shows a large number of such cross peaks. Many of them strongly overlap, see region between 45-20 and 50-60 ppm in Figure 9a), despite the small size of the protein. The observed line width is relatively large compared to crystalline proteins such as SH3 [Pauli 01] or ubiquitin [Igumenova 04a]. With respect to other amyloid fibrils (see section 1.2), however, the line width is in an intermediate range. The cross peak patterns of the threonine, serine, valine, alanine, and proline residues can be identified directly on the basis of their characteristic chemical shifts, see Figure 9a). The spectrum did not show signs of peak doubling for these residues, but the sequential assignment procedure revealed signal splitting for a threonine and a serine residue, see section 4.3. In general, and taking the cross peak pattern into account, the spectra recorded on 1,3-CA150.WW2

and 2-CA150.WW2 amyloid fibrils compare well with the spectrum of u-CA150.WW2 fibrils, showing good reproducibility of sample preparation. This, together with the observed line widths, indicates that the sample preparation is suitable for structural studies.

CP and PDSD transfers both depend on the strong proton-carbon dipolar coupling present in biological samples. However, the deuterated sample d-CA150.WW2 is only protonated at the exchangeable sites. Due to this dilution of the proton spins the proton-carbon dipolar coupling is reduced by MAS to a larger extent which makes CP and PDSD-transfers less efficient. Figure 9d) shows a 20 ms CP-PDSD spectrum recorded on d-CA150.WW2 at 10.5 kHz MAS. Off-diagonal cross peaks are observed despite the low abundance of protons in this sample. They appear mainly from carbon atoms in the vicinity of a proton, such as the C$\alpha$ and C$\beta$-sites near the amide proton or side chain atoms close to an exchangeable site as for example in serine, threonine, or lysine. At higher spinning frequencies, e.g. 20 kHz, no off-diagonal signals occur in d-CA150.WW2 PDSD spectra. The cross peak positions in spectra of d-CA150.WW2 vary from those in the other samples due to the isotope effect [Hansen 00] that causes a decrease in the carbon chemical shift depending on the number of $^{2}$H-nuclei at the respective carbon site.

Interestingly, in comparison with the protonated samples the line width of most cross peaks is significantly reduced in the spectra recorded on d-CA150.WW2. This might be due to either the low amount of protons in the sample which eases decoupling or differences in the dynamic behaviour in this sample. A different structural arrangement was excluded on the basis of the chemical shifts which, taking the isotope effect into account, are very similar for the protonated and the deuterated samples. Another difference to the spectra of the protonated samples is the occurrence of signals from an additional serine residue. In the protonated samples these signals might be broadened beyond detection, possibly due to dynamic properties that interfere with the proton-carbon decoupling.

Signals from mobile residues suffer from a low CP-transfer efficiency. In order to detect them, a direct excitation spectrum was recorded on d-CA150.WW2, Figure 9e). 20 kHz MAS was used which to a large extent reduces the proton-carbon couplings and sufficient decoupling can be achieved by a low power proton decoupling scheme (5 kHz WALTZ16). A 1.2 ms RFDR pulse train was used for $^{13}$C-$^{13}$C mixing. The spectrum shows narrow lines and some signals exhibit line widths even smaller than in the CP-PDSD spectrum of d-CA150.WW2,

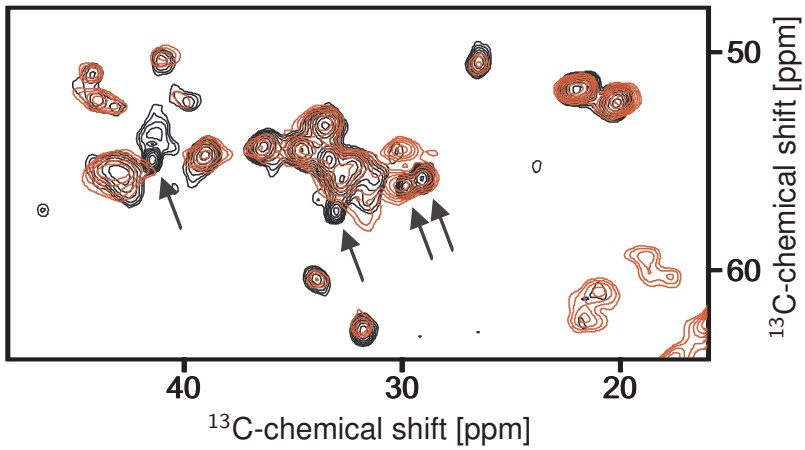



Figure 10: Overlay of the direct excitation RFDR spectrum (black) with the CP-PDSD spectum (red), both recorded on d-CA150.WW2. The black arrows indicate cross peaks which show remarkably narrow lines in the direct excitation spectrum and are not visible or broadened in the CP spectrum.

Figure 9d,e). In addition, some strong signals appear which are weaker or not visible in the CP-PDSD spectrum, arrows in Figure 10. These cross peaks presumably originate from residues which are relatively mobile and are therefore hardly excited by the cross polarisation scheme. Furthermore, in CP-PDSD spectra cross peaks from mobile sites suffer from a slow proton-driven spin-diffusion mixing step in the proton-diluted sample.

$^{13}$C-$^{15}$N correlation spectra which can be used to correlate the atoms along the protein backbone and are crucial experiments in most MAS NMR assignment procedures, see section 2.3, were recorded on the different CA150.WW2 samples. Figure 11 shows CP-$^{13}$C-$^{15}$N correlations of the 1,3-CA150.WW2 sample. Despite the reduced isotope labelling in this sample the spectral resolution is relatively poor in the NCA and the NCO spectra due to the intermediate line width observed in the $^{13}$C and the $^{15}$N-dimensions. In addition, NCACX and NCOCX spectra with a PDSD mixing time of 100 ms were recorded, Figure 11. These techniques did not yield efficient magnetisation transfer apart from the threonine C$\beta$-sites in the NCOCX spectrum due to the enhancement of magnetisation transfer by rotational resonance. Similar results were observed for the other samples and thus CP-$^{15}$N-$^{13}$C correlation experiments were not used further in the course of this work.

Another experimental scheme that establishes correlations between nitrogen and carbon sites is the TEDOR experiment. Figure 12 shows two examples for TEDOR-$^{15}$N-$^{13}$C correlations. In the spectrum recorded on 2-CA150.WW2, see Figure 12a), severe signal overlap similar

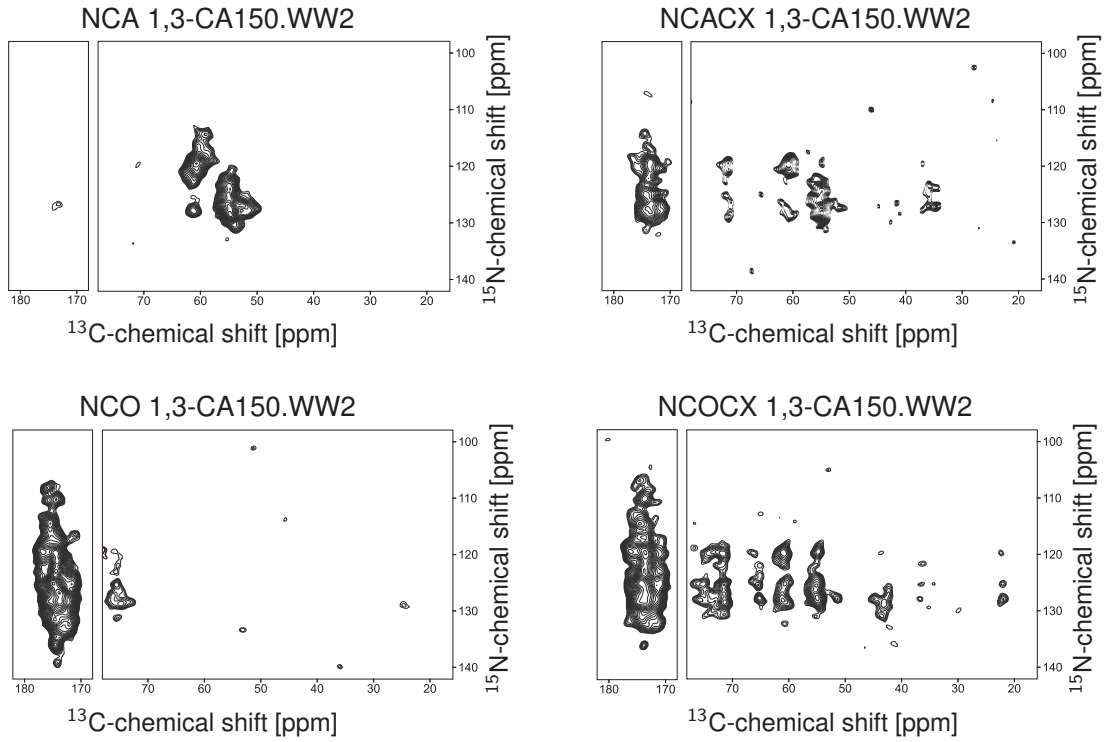



Figure 11: Carbonyl and aliphatic regions of CP-$^{15}$N-$^{13}$C correlation spectra recorded on 1,3-CA150.WW2. The PDSD mixing time in NCACX and NCOCX spectra was 100 ms.

to the CP-$^{15}$N-$^{13}$C spectra of 1,3-CA150.WW2, Figure 11, occurs. In contrast, despite the uniform $^{13}$C-labelling, the signals in the d-CA150.WW2 spectrum, see Figure 12b), are better resolved, however, it is still crowded. As correlations of each nitrogen atom to the neighbouring C$\alpha$ and CO-sites are observed at the same time, further magnetisation transfer to side chain carbon atoms, by either a longer TEDOR mixing or a $^{13}$C-$^{13}$C correlation step, leads to crowded spectra. Therefore, dispersion of the signals into an additional third dimension would be very useful and is demonstrated in this work using an amide proton dimension, see section 4.3.4.

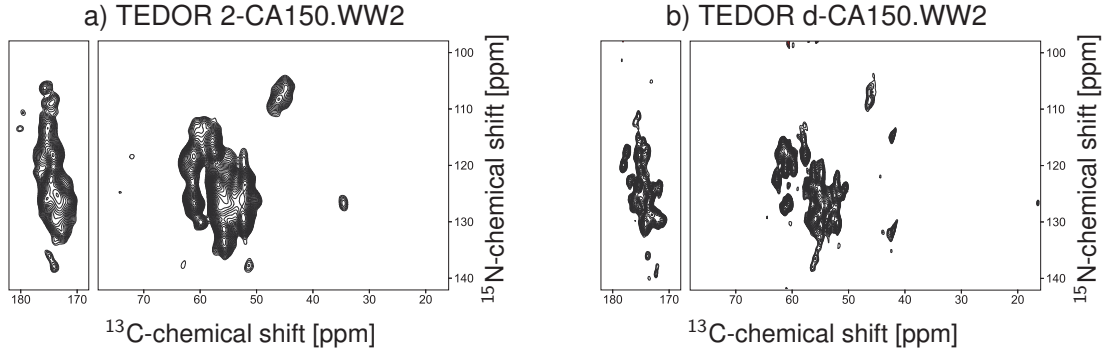



Figure 12: Carbonyl and aliphatic regions of TEDOR-$^{15}$N-$^{13}$C correlation spectra recorded on a) 2-CA150.WW2 and b) d-CA150.WW2 with a TEDOR mixing time of 1.6 and 0.8 ms, respectively.

## 4.3 Assignment of CA150.WW2 amyloid fibrils employing new methods

### 4.3.1 A novel $^{13}$C-assignment strategy

The low spectral resolution in $^{15}$N-$^{13}$C correlations on CA150.WW2 amyloid fibrils hindered the usage of standard MAS NMR assignment procedures along the protein backbone, see section 2.3. In the following, an assignment strategy is outlined that relies on the labelling pattern observed in spectra of glycerol-labelled samples, [Lemaster 82]. Cross peaks between neighbouring C$\alpha$ or CO-sites occur in the spectra from these samples only for distinct sequential pairs depending on the sequence. In Figure 13 the labelling pattern at the C$\alpha$ and the CO-sites is given for the 40-residue CA150.WW2 construct.

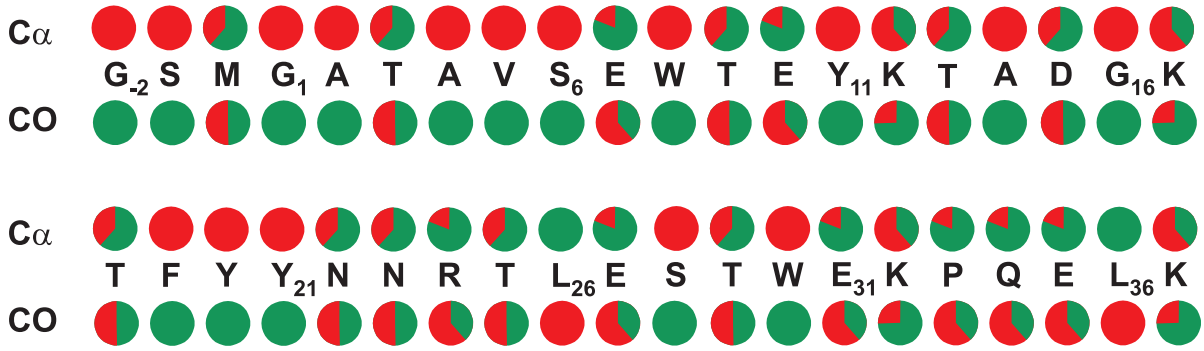



Figure 13: The degree of C$\alpha$ and CO-labelling of the CA150.WW2 sequence for 1,3 and 2-CA150.WW2 samples in green and red, respectively.

In PDSD spectra with a mixing time $\geq$100 ms C$\alpha$/C$\alpha$ cross peaks from neighbouring residues are observed. In a uniformly labelled sample thus 39 such cross peaks are expected leading to severe spectral overlap. In contrast, the same spectrum recorded on a 2-$^{13}$C-glycerol labelled sample shows only cross peaks between neighbouring amino acids in case of isotope labelling at their C$\alpha$-sites. In 2-CA150.WW2 this is for example the case for $A_4$/$V_5$ but not for $S_6$/$E_7$. In spectra with a good signal-to-noise ratio weak signals may also be detectable for residue pairs with partial labelling such as $K_{12}$/$T_{13}$. In spectra of 2-CA150.WW2 and 1,3-CA150.WW2, eight and nine strong C$\alpha$/C$\alpha$ cross peaks are expected, respectively. This reduced number of signals in the C$\alpha$-region of the spectra greatly enhances the resolution and at the same time simplifies the assignment due to the restricted number of assignment possibilities which complement each other in the two glycerol-labelled samples. In a similar way, other contacts such as C$\alpha$/C$\beta$, C$\alpha$/proline-C$\delta$, or C$\alpha$/CO can be used in PDSD spectra and other correlations. An example would be TEDOR spectra in which an amide nitrogen only

shows correlations to the sequential C$\alpha$ and CO-sites for pairs of amino acids with labelling at both sites.

### 4.3.2 $^{13}$C-assignment of CA150.WW2 amyloid fibrils

The assignment process of the MAS NMR spectra of CA150.WW2 amyloid fibrils was started with the identification of the observed signal sets. PDSD and RFDR spectra with short mixing times, as shown in Figure 9, contain only cross peaks between atoms that belong to the same residue and are ideally suited for this purpose. Each type of amino acid has a typical cross peak pattern that is further characterised by the specific labelling pattern in spectra of the glycerol-labelled samples [Castellani 02]. In this way, most signal sets could directly be assigned to a certain type of residue. The improved line widths observed in spectra of the deuterated sample helped in some cases to distinguish spin systems of the same type. Due to the deuterium isotope effect [Hansen 00], separate chemical shift tables for the deuterated and the protonated samples were compiled during the assignment procedure. Sequential connectivities between spin systems that were found in spectra from either the protonated or the deuterated samples were used in the interpretation of all spectra.

PDSD spectra with mixing times $\geq$100 ms were used to detect correlations between atoms of different spin systems. The C$\alpha$/C$\alpha$ correlations observed in the two glycerol-labelled samples were of special importance as described in the preceding paragraph. C$\alpha$/C$\alpha$ cross peaks may, however, also occur between residues located in adjacent $\beta$-strands. Hence, it is necessary to distinguish such long range contacts from sequential correlations. The shorter the mixing time at which inter residue cross peaks appear, the more likely they are sequential in nature. Cross peaks that exclusively appear in spectra with long mixing times (300-400 ms) were only treated as sequential when this was supported by the occurrence of several cross peaks between the respective residues. A large number of sequential contacts was observed in a 400 ms DARR spectrum of d-CA150.WW2. They appeared only at long mixing times because of the decreased transfer efficiency in the deuterated sample. The strongest of these cross peaks result from residues which have exchangeable sites in their side chains such as serine, lysine, or threonine.

In the next paragraphs, the sequential assignment procedure for each residue is described and the most important sequential cross peaks are discussed. Most of these cross peaks

are indicated in Figure 14 showing a) a 100 ms PDSD spectrum of 2-CA150.WW2 and b) a 400 ms DARR spectrum of d-CA150.WW2. In addition, Figure 14c) shows a 4.8 ms RFDR spectrum of the deuterated sample with annotated assignments. In total 35 (out of 40) C$\alpha$, 30 (out of 37) C$\beta$, and 25 (out of 40) CO-resonances were assigned. The results are compiled in Table 5. Some chemical shifts were only determined on the deuterated sample. The corresponding chemical shift values for the protonated sample were calculated from these: C$\alpha$ (glycine) = +0.4 ppm, C$\alpha$ (all other amino acids) = +0.3 ppm, C$\beta$ ($CD_2$) = +0.5 ppm [Hansen 00]. To identify the possible resonance frequencies for each atom in the protein, the restricted set of protein chemical shifts provided by the BMRB was used: www.bmrb.wisc.edu/ref_info/statsel.htm. Here, the assignment procedure is described on a residue type basis.

**Valine ($V_5$):** The CA150.WW2 sequence contains only one valine which was easily identified by its characteristic cross peak pattern. Due to the absence of isoleucine in the CA150.WW2 sequence the valine is the only type of residue for which cross peaks are expected around 30-35/10-23 ppm in both dimensions. They correspond to the C$\beta$/C$\gamma_{1/2}$ signals. In case of $V_5$, two isolated cross peaks are observed at 35.4/18.8 and 35.4/19.7 ppm in the spectra from the uniformly labelled samples. The identity of these signals is further confirmed by the cross peaks involving C$\beta$ and C$\gamma_{1/2}$ in one dimension and the C$\alpha$-site at 60.9 ppm in the other. In addition, as a very characteristic feature, a strong C$\alpha$/C$\beta$ cross peak occurs in spectra of 2-CA150.WW2 recorded at short mixing times, see Figure 9 on page 39. Similarly, 1,3-CA150.WW2 spectra show CO/C$\gamma_{1/2}$ cross peaks at 175.2/18.8 and 175.2/19.7 ppm.

**Proline ($P_{33}$):** Proline is the only residue for which cross peaks are expected around 50/60 ppm due to the C$\alpha$/C$\delta$ correlation. Such cross peaks are indeed observed in spectra of u-CA150.WW2, d-CA150.WW2, and 2-CA150.WW2 at 62.9/50.9 ppm. Cross peaks involving these two resonances were identified at the C$\beta$ and C$\gamma$-chemical shifts of 32.6 and 27.3 ppm, respectively, in spectra of u-CA150.WW2 where they are well isolated from other signals.

**Alanine ($A_2$, $A_4$, $A_{14}$):** The pattern of the alanine residues are easily found due to their unique C$\alpha$/C$\beta$ cross peaks around 51-53/21-23 ppm, appearing in spectra of uniformly labelled samples. Two of the alanines have almost the same C$\alpha$ and C$\beta$ shifts (52.1/

| | H | N | CO | C$\alpha$ | C$\beta$ | C$\gamma/\delta/\varepsilon/\zeta$ | d-CO | d-C$\alpha$ | d-C$\beta$ | d-C$\gamma/\delta/\varepsilon/\zeta$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $G_{-2}$ | | | (169.9) | (43.6) | | | 169.9 | 43.2 | | |
| $S_{-1}$ | | | (173.6) | (55.9) | (64.5) | | 173.6 | 55.6 | 64.0 | |
| $M_0$ | | | (174.5) | 54.9 | 35.4 | 31.1/-/17.3 | 174.5 | 54.5 | 34.4 | 30.2/-/16.6 |
| $G_1$ | 7.92 | 108.6 | 171.3 | 46.3 | | | 171.1 | 45.9 | | |
| $A_2$ | 8.66 | 124.5 | 177.0 | 52.1 | 22.8 | | 177.0 | 51.5 | 22.0 | |
| $T_3$ | 9.40 | 117.6 | 172.7 | 62.3 | 73.7 | 21.8 | 173.3 | 61.2 | 73.2 | 21.8 |
| $T'_3$ | | | | | 72.2 | | | | 72.4 | 18.9 |
| $A_4$ | 9.90 | 131.0 | 175.9 | 52.1 | 22.7 | | 175.3 | 51.5 | 21.8 | |
| $V_5$ | 9.54 | 126.7 | 175.2 | 60.9 | 34.5 | 18.8/19.7 | 174.8 | 60.2 | 33.7 | 17.6/18.7 |
| $S_6$ | 9.28 | 123.8 | 173.2 | 56.8 | 65.6 | | 173.4 | 56.2 | 65.0 | |
| $E_7$ | | | | 56.1 | | | | | | |
| $W_8$ | | | | | (30.4) | | | | 29.9 | |
| $T_9$ | | | 173.2 | 61.0 | 71.9 | 22.1 | 173.4 | 60.3 | 71.2 | 21.0 |
| $E_{10}$ | | | | 55.1 | 34.5 | 36.5/183.6 | | | 33.3 | 35.6 |
| $Y_{11}$ | | | | | | | | | | |
| $K_{12}$ | | | (175.0) | 54.6 | 37.0 | 25.5/29.7/42.2 | 175.0 | 54.2 | 36.3 | 24.6/28.9/41.2 |
| $T_{13}$ | 9.24 | 119.2 | 173.5 | 59.3 | 71.3 | 20 | 172.8 | 59.2 | 70.5 | 18.8 |
| $A_{14}$ | 8.00 | 125.8 | 175.9 | 52.5 | 21.6 | | 176.5 | 52.1 | 20.0 | |
| $D_{15}$ | | | | 55.2 | 39.8 | 180.9 | | 54.6 | 38.9 | 180.3 |
| $G_{16}$ | 8.15 | 108.6 | 174.6 | 44.9 | | | 174.2 | 44.7 | | |
| $K_{17}$ | 7.92 | 125.1 | (176.5) | 55.2 | 33.9 | 25.1/29.7/42.1 | 176.5 | 55.1 | 32.2 | 23.8/28.4/41.3 |
| $T_{18}$ | 9.01 | 122.2 | 171.9 | 62.3 | 72.2 | 21.8 | 171.7 | 62.1 | 71.7 | 21.6 |
| $F_{19}$ | 9.64 | 129.9 | (172.2) | (56.2) | (43.1) | | 172.2 | 55.9 | 42.6 | 136.6 |
| $Y_{20}$ | | | | | | | | | | |
| $Y_{21}$ | | | | | | | | | | |
| $N_{22}$ | | | | | | | | | | |
| $N_{23}$ | | | | 52.3 | | | | | | |
| $R_{24}$ | | | | 54.9 | 34.5 | 27.3/43.7/159.6 | | | 33.3 | 26.3/43.3 |
| $T_{25}$ | | | 173.2 | 60.9 | 71.8 | 21.8 | 173.4 | 61.0 | 70.9 | 21.0 |
| $L_{26}$ | | | | 55.1 | 42.2 | 27.5/21.7/22.6 | | 53.7 | 41.5 | 25.9 |
| $E_{27}$ | | | (174.0) | 55.0 | 34.3 | 36.6/183.6 | 174.0 | 54.5 | 33.2 | 35.5 |
| $S_{28}$ | 8.84 | 115.9 | 173.5 | 57.5 | 67.0 | | 173.5 | 57.5 | 66.3 | |
| $S'_{28}$ | | | | 57.0 | 65.5 | | 173.9 | 57.1 | 64.7 | |
| $T_{29}$ | | | 173.2 | 61.1 | 71.9 | 21.7 | 173.4 | 60.6 | 71.2 | 20.4 |
| $W_{30}$ | | | | | (30.4) | | | | 29.9 | |
| $E_{31}$ | | | | 56.2 | | | | 55.9 | | |
| $K_{32}$ | 9.21 | 130.3 | | 53.7 | 34.5 | 25.0/29.8/42.2 | | 53.0 | 33.3 | 23.8/28.7/41.4 |
| $P_{33}$ | | | 176.6 | 62.9 | 32.6 | 27.3/50.9 | 176.7 | 62.5 | 31.7 | 26.4/50.4 |
| $Q_{34}$ | | | (175.9) | 56.1 | 29.9 | 34.0/180.4 | 175.9 | 55.6 | 29.0 | 33.3/180.4 |
| $E_{35}$ | | | (176.1) | 56.4 | 30.6 | 36.3/183.8 | 176.1 | 56.0 | 29.7 | 35.7/184.0 |
| $L_{36}$ | | | | (55.2) | 42.2 | 26.9/23.6/25.1 | | 54.9 | 41.4 | 26.1/24.1/22.7 |
| $K_{37}$ | | | 181.1 | 57.6 | 33.9 | | 181.2 | 57.3 | 33.0 | 23.8 |

Table 5: Solid state MAS NMR assignments in ppm of fibrillar CA150.WW2, the last four columns contain the carbon chemical shifts obtained on d-CA150.WW2. Some chemical shifts were only determined for d-CA150.WW2 and the corresponding resonances for the protonated samples (in brackets) were calculated by correcting for the deuterium isotope effect [Hansen 00], see main text.

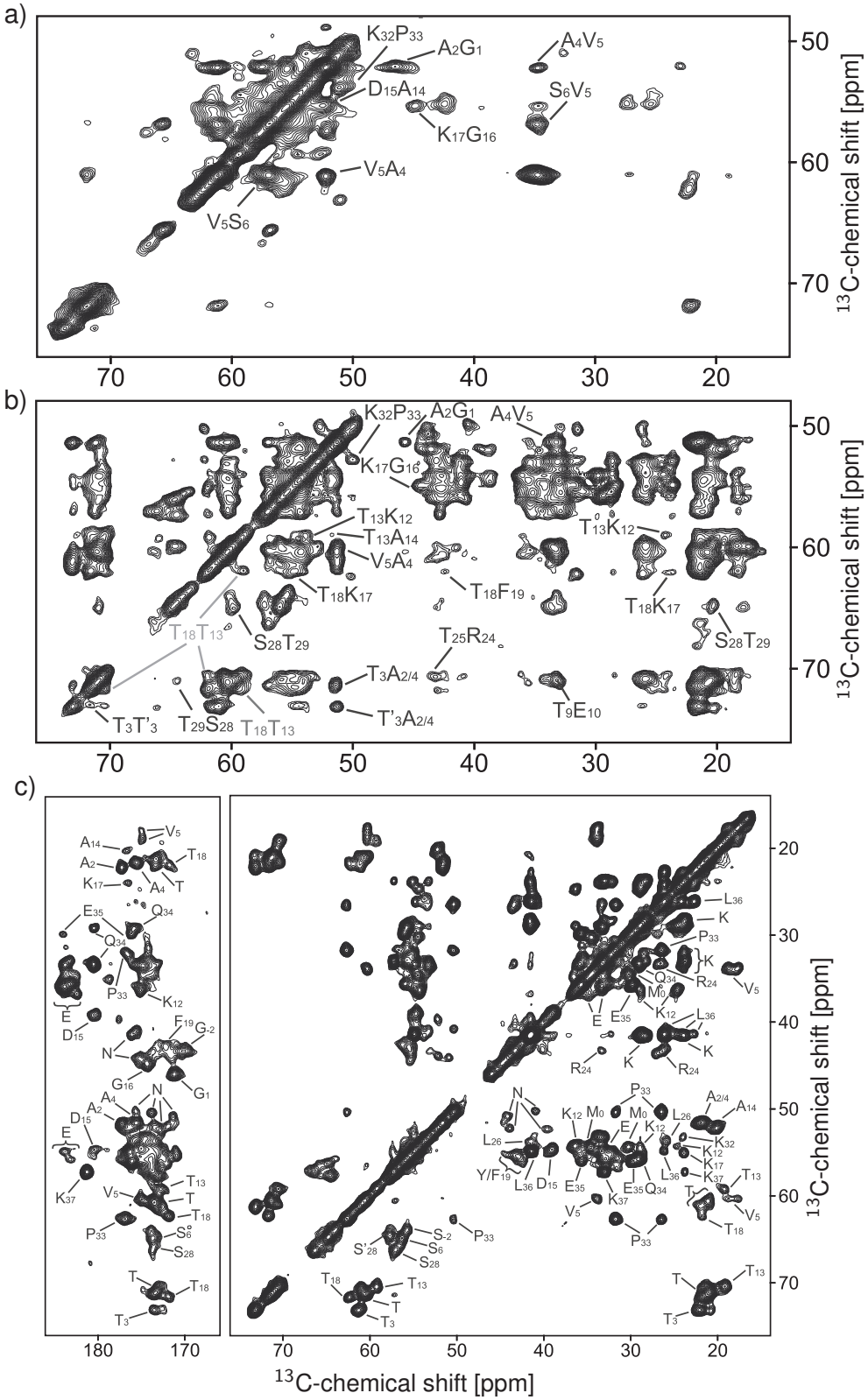



Figure 14: a) 100 ms PDSD spectrum of 2-CA150.WW2 and b) 400 ms DARR spectrum of d-CA150.WW2 with annotated sequential correlations. In b) the signal-pattern of a $T_{13}/T_{18}$ long range contact is indicated in grey, c) 4.8 ms RFDR spectrum of d-CA150.WW2 with $^{13}$C-assignments.

22.75 ppm) but can be distinguished well by their carbonyl resonances at 177 and 175.9 ppm in u-CA150.WW2, d-CA150.WW2, and 1,3-CA150.WW2 spectra. The third alanine (C$\alpha$/C$\beta$ at 52.5/21.6 ppm) shows weaker and often broadened cross peaks. This can for example be seen for the CO/C$\beta$ alanine cross peak at 176.5/20 ppm in Figure 14c).

For one of the two alanines with overlapping C$\alpha$/C$\beta$ signals, a large number of cross peaks to $V_5$ was observed. Similarly, cross peaks to a glycine residue occur, see Figure 14a). These two residues were assigned to $A_4$ and $A_2$, because as $A_2$ is the only alanine next to a glycine. The correct assignment of the corresponding alanine CO-resonances was established from a cross peak in the 100 ms PDSD spectrum of u-CA150.WW2 between the alanine CO-resonance at 175.9 ppm and the methyl groups of $V_5$. At the C$\alpha$-chemical shift of $A_2$ and $A_4$ also cross peaks to threonine C$\alpha$ and C$\beta$-resonances could be detected, e.g. in the 100 ms PDSD spectrum of u-CA150.WW2, in accordance with the triplet sequence $A_2T_3A_4$.

The remaining alanine is then $A_{14}$ which is confirmed by a C$\alpha$/C$\alpha$ cross peak at 52.5/55.2 ppm in F2/F1 in the 100 ms PDSD spectrum of 2-CA150.WW2, Figure 14a). The resonance at 55.2 ppm can neither come from $G_1$, $T_3$, $V_5$, nor $T_{13}$ and thus the cross peak was assigned to $A_{14}$C$\alpha$/$D_{15}$C$\alpha$. Additionally, some weak cross peaks from $A_{14}$ to a threonine residue were observed in spectra of the deuterated sample, e.g. at 52.1/59.2 ppm in F2/F1 in Figure 14b), in agreement with $A_{14}$ being next to a threonine residue.

**Aspartic acid ($D_{15}$):** The aspartic acid carbonyl side chain resonance occurs at a characteristic frequency around 179 ppm. In this region, two cross peaks were detected in spectra of u-CA150.WW2 at 180.9/55.2 and 180.9/39.8 ppm in F2/F1 and were assigned to $D_{15}$CO/C$\alpha$ and $D_{15}$CO/C$\beta$, respectively. A corresponding C$\alpha$/C$\beta$ cross peak was detected in spectra of u-CA150.WW2 and also appeared in 1,3-CA150.WW2 spectra. The identity of the residue was confirmed by cross peaks to $A_{14}$ (see above). All cross peaks of the $D_{15}$ signal pattern showed relatively broad lines, comparable to the line widths observed for $A_{14}$.

**Glycine ($G_{-2}$, $G_1$, $G_{16}$):** Glycine is a very small residue and gives rise to a single cross peak in $^{13}$C-$^{13}$C correlations. This peak is only observed in spectra of uniformly $^{13}$C-labelled samples. The C$\alpha$-chemical shifts can be identified in NCA or TEDOR spectra since they appear at a unique carbon chemical shift range (43-47 ppm). Figure 12 on page 42 shows a TEDOR spectrum of 2-CA150.WW2 where two glycine signals appear at C$\alpha$ chemical shifts of

46.3 and 44.9 ppm. Using this information, CO/C$\alpha$ cross peaks for two glycines were identified in 20 ms PDSD spectra of u-CA150.WW2 at 171.3/46.3 and 174.6/44.9 ppm.

$G_1$ was identified by the $G_1C\alpha/A_2C\alpha$ cross peak at 46.3/52.1 ppm in F1/F2 in spectra of 2-CA150.WW2, Figure 14a). In addition, at the position of the $G_1C\alpha$-resonance two symmetric pairs of cross peaks (45.9/34.4 and 45.9/30.2 ppm) are observed in the d-CA150.WW2 400 ms PDSD spectrum – outside the plotted region in Figure 14b). These signals were attributed to correlations to the C$\beta$ and C$\gamma$-sites of the neighbouring residue $M_0$.

In spectra of 2-CA150.WW2, see Figure 14a), a cross peak involving a glycine C$\alpha$-chemical shift is observed at 44.9/55.2 ppm in F2/F1. This is consistent with a contact to a lysine residue (see below), hence the cross peak was assigned to $G_{16}C\alpha/K_{17}C\beta$. At 44.7/ 38.9 ppm weak correlations are observed in a 400 ms DARR spectrum of d-CA150.WW2. Based on the assignment of $D_{15}$ (see above), the symmetric cross peaks were assigned to $G_{16}C\alpha/D_{15}C\beta$ confirming the assignment of the $G_{16}/K_{17}$ cross peak. Interestingly, the $G_{16}$CO/ C$\alpha$ cross peak in F2/F1 is relatively well defined in the spectra of d-CA150.WW2 (174.2/ 44.7 ppm, Figure 14c)) but has a very broad, inhomogeneous line in spectra of u-CA150.WW2.

The first residue in the sequence is $G_{-2}$ which is not seen in the standard NCA or TEDOR experiments. However, in d-CA150.WW2 spectra at 169.9/43.6 ppm in F2/F1 a cross peak occurs. At the CO-chemical shift of 169.9 ppm no other cross peaks occur indicating that it is part of a glycine signal pattern, Figure 14c). In addition, the intensity of the signal is strongest in direct carbon excitation spectra, indicating some mobility of this residue. Thus the signal was assigned to the very first amino acid of the sequence, $G_{-2}$.

**Methionine ($M_0$):** On the diagonal of all $^{13}C$-$^{13}C$ spectra a signal at 17.3 ppm occurs. This resonance was assigned to the methyl group of the methionine residue due to the absence of isoleucine in the CA150.WW2 sequence and the knowledge of the chemical shifts of the alanine and valine methyl groups. In addition, no cross peaks from this signal to resonances around 60 or 70 ppm characteristic for a threonine residue were observed. Correlations from the resonance at 17.3 ppm to the remaining methionine sites occurred only at long mixing times due to the sulphur atom interupting the methionine side chain. Most sites in the methionine residue are predominantly labelled in the 1,3-CA150.WW2 sample. In the 400 ms DARR spectrum of 1,3-CA150.WW2 cross peaks occurred at 17.3/35.4 and 17.3/31.1 ppm and were assigned to $M_0C\varepsilon/C\beta$ and $M_0C\varepsilon/C\gamma$, respectively. The $M_0C\alpha$-resonance was

identified from the $M_0C\alpha/C\gamma$ cross peak at 54.9/35.4 ppm in spectra of 1,3-CA150.WW2. The identity of the residue was also confirmed by correlations involving the C$\alpha$-atom of $G_1$, see above. No cross peaks to $S_{-1}$ were observed.

**Serine ($S_{-1}$, $S_6$, $S_{28}$):** The serine residues can be identified in the spectra of the uniformly labelled sample by their characteristic C$\alpha$/C$\beta$ cross peaks around 56-61/62-66 ppm which are isolated from the other signals in the spectra. Three such cross peaks are observed in most CA150.WW2 spectra. However, spectra of the deuterated sample show four serine C$\alpha$/C$\beta$ cross peaks.

In the 200 ms PDSD spectrum of 1,3-CA150.WW2 cross peaks between two of the serine C$\beta$-resonances to a signal at 71.9 ppm were observed characteristic for a threonine C$\beta$-resonance. However, in the CA150.WW2 sequence there is only one serine ($S_{28}$) next to a threonine ($T_{29}$). Therefore, both serine signal sets that show correlations with this threonine C$\beta$-resonance were assigned to $S_{28}$. The observed peak doubling of $S_{28}$ is most pronounced for the $S_{28}$C$\beta$-resonance. The other serine residue visible in spectra of all samples shows many cross peaks to $V_5$, e.g. $V_5$C$\alpha$/S$\alpha$ at 60.9/56.8 ppm in Figure 14a), and was assigned to $S_6$.

At the unique serine C$\beta$-chemical shifts of $S_6$, $S_{28}$, and S'$_{28}$, further correlations were observed in spectra of most samples which were assigned to contacts involving the neighbouring $E_7$ and $E_{27}$-residues. The respective serine-C$\beta$/glutamic acid-C$\alpha$ cross peaks at 65.6/56.1 ($S_6/E_7$) and 67.0/55.0 ppm ($S_{28}/E_{27}$) were assigned in the 400 ms DARR spectrum of 1,3-CA150.WW2 lacking valine C$\beta$-signals.

The serine residue only visible in spectra of d-CA150.WW2 was assigned to $S_{-1}$. No correlations to other residues were observed. This indicates some mobility at the N-terminus.

**Lysine ($K_{12}$, $K_{17}$, $K_{32}$, $K_{37}$):** The lysine residue has a large cross peak pattern which partially overlaps with the signals from other residues. The lysine C$\beta$/C$\gamma$, C$\beta$/C$\delta$, and C$\gamma$/C$\delta$ cross peaks are observed in spectra of 1,3-CA150.WW2 in the range from 25-35 ppm as the dominant feature. In a 20 ms PDSD spectrum of d-CA150.WW2, see Figure 9d) on page 39, the lysine C$\gamma$/C$\delta$ cross peaks present one of the few signals outside the C$\alpha$/C$\beta$-region and thus can be assigned unambiguously in this spectrum. These cross peaks show that the chemical shifts of the lysine C$\gamma$-resonances are between 24 and 26 ppm. In the region

52-60/24-26 ppm in F2/F1, where thus the lysine C$\alpha$/C$\gamma$ cross peaks are expected, three signals are observed in the 100 ms PDSD spectrum of d-CA150.WW2. In the 4.8 ms direct excitation spectrum a fourth signal appears in this region, see Figure 14c). These four cross peaks were tentatively assigned to lysine and the spectra were further investigated for sequential correlations supporting this assignment.

The residue with C$\alpha$/C$\gamma$ cross peaks at 55.2/25.1 ppm was assigned to $K_{17}$ due to contacts to $G_{16}$ (see above). Correlations involving resonances of this lysine signal set and a threonine C$\alpha$-resonance at 55.1/62.1 and 23.8/62.1 ppm in F2/F1 were observed in the 400 ms DARR spectrum of d-CA150.WW2, Figure 14b), and assigned to $K_{17}$/$T_{18}$.

One of the lysine C$\alpha$/C$\gamma$ cross peaks has a relatively isolated C$\alpha$-chemical shift of 53.7 ppm. A correlation to the $P_{33}$C$\delta$-site at 53.7/50.9 ppm in F1/F2 was observed in spectra of 2-CA150.WW2, Figure 14a), and assigned to $K_{32}$/$P_{33}$. In addition, in a 400 ms DARR spectrum of 1,3-CA150.WW2 cross peaks from the $K_{32}$C$\alpha$-resonance to another C$\alpha$-site were observed at 53.7/56.2 ppm. These correlations were assigned to $E_{31}$/$K_{32}$.

In most spectra the lysine residue with the largest C$\gamma$-chemical shift (25.5 ppm) has an increased line width and only in the 400 ms DARR spectrum of d-CA150.WW2 inter residue cross peaks to a threonine were observed, e.g. lysine-C$\alpha$/threonine-C$\alpha$ at 54.2/59.2 ppm in F2/F1, Figure 14b). This correlation was assigned to $K_{12}$/$T_{13}$, because the only other sequential lysine-threonine pair in the sequence ($K_{17}$/$T_{18}$) was already identified. However, no correlations from $K_{12}$ to $Y_{11}$ were unambiguously identified due to spectral overlap, further obscured by the line broadening observed for $K_{12}$ and the tyrosine residues (see below).

The forth C$\alpha$/C$\gamma$ cross peak at 57.3/23.8 ppm in F1/F2 occurs only in the direct excitation spectrum of d-CA150.WW2. No sequential correlations involving the C$\alpha$-chemical shift were observed. The CO-resonance of this residue occurs at a very high chemical shift value (181.1 ppm). This is in accordance with an assignment to $K_{37}$ which has a free C-terminus. The increased intensity of this signal set in direct carbon excitation spectra, see Figure 10 on page 41, indicates mobility at the C-terminus.

**Arginine ($R_{24}$):** Arginine residues can be identified by the chemical shift of the carbon atom in their guanidinium group which appears around 160 ppm. However, signals from this group are hardly visible in most CA150.WW2 spectra. Such a correlation was only observed in a 700 ms DARR spectrum of 2-CA150.WW2 at 159.6/43.7 ppm in F2/F1 and was assigned to

$R_{24}C\zeta/C\delta$. The signals of the other side chain carbons were then identified by correlations to the C$\delta$-site.

A cross peak was observed at 43.3/70.9 ppm in F1/F2 in the 400 ms DARR spectrum of d-CA150.WW2, see Figure 14b). This contact was assigned to $R_{24}/T_{25}$ based on the characteristic threonine C$\beta$-chemical shift. A cross peak at 54.9/52.3 ppm in F1/F2 in the 100 ms PDSD spectrum of 1,3-CA150.WW2 was assigned to $R_{24}C\alpha/N_{23}C\alpha$.

**Threonine ($T_3$, $T_9$, $T_{13}$, $T_{18}$, $T_{25}$, $T_{29}$):** Threonines are easily identified by their C$\alpha$/C$\beta$ (around 60/71 ppm), C$\alpha$/C$\gamma$ (around 60/25 ppm), and C$\beta$/C$\gamma$ (around 71/25 ppm) cross peaks. However, even the relatively high resolution in d-CA150.WW2 spectra is not sufficient to resolve the signals from all six threonine residues. Three isolated C$\alpha$/C$\beta$-cross peaks and a larger peak conglomerate were observed. Many sequential correlations involving threonine C$\alpha$ and C$\beta$-sites appeared in the 400 ms DARR spectrum of d-CA150.WW2, see Figure 14b).

Relatively weak cross peaks from a threonine residue (C$\alpha$ = 59.2 ppm, C$\beta$ = 70.5 ppm) to $K_{12}$ and $A_{14}$ occurred in this spectrum (see above). The corresponding threonine signal set belongs to one of the more isolated ones and was assigned to $T_{13}$. In most spectra, the C$\gamma$-signal of this residue shows broad lines.

The only other sequential contacts between alanine and threonine residues are expected for the $A_2T_3A_4$-triplet. Surprisingly, two different threonine C$\beta$-resonances show strong correlations to the degenerate alanine C$\alpha$-chemical shifts of $A_2$ and $A_4$. In the carbonyl region, where $A_2$ and $A_4$ have very different chemical shifts, both alanine residues show correlations to these two threonine-C$\beta$ resonances. The cross peaks at the $A_2$CO-resonance are stronger in accordance with the shorter distance to the $T_3$C$\beta$-atom compared to the $A_4$CO-site. In addition, cross peaks between the two $T_3$C$\beta$-resonances are observed at 72.4/73.2 ppm in Figure 14b), indicating chemical exchange. It was concluded that $T_3$, in addition to $S_{28}$, shows peak doubling.

$T_{18}$ was identified via its correlations to $K_{17}$ (see above) and is the third threonine which shows an isolated C$\alpha$/C$\beta$ cross peak. From the C$\alpha$-site a weak correlation at 62.1/42.6 ppm in F2/F1 was observed and assigned to $T_{18}C\alpha/F_{19}C\beta$.

The remaining three threonines have very similar chemical shifts. Correlations at 71.1/ 33.3 ppm in F1/F2 appeared in d-CA150.WW2 spectra, Figure 14b), and in 1,3-CA150.WW2

spectra. Based on the labelling pattern of the 1,3-CA150.WW2 sample this cross peak was assigned to $T_9C\beta/E_{10}C\beta$. No cross peaks between threonine signals and resonances around 30 ppm occurred which would be expected for the $W_8/T_9$ and $T_{29}/W_{30}$ correlations. $T_{29}$ and $T_{25}$ were identified by contacts to $S_{28}$ (see above) and $R_{24}$ (see above), respectively. A possible $T_{25}/L_{26}$ cross peak was only observed in a 700 ms DARR spectrum of 2-CA150.WW2 but due to the long mixing time this signal could also arise from a long range contact.

**Leucine ($L_{26}$, $L_{36}$):** The leucine residue can be identified by its strong $C\beta/C\gamma$ cross peak around 42/27 ppm in 2-CA150.WW2 spectra with short mixing times. Symmetric cross peaks from one such correlation are observed in the 20 ms PDSD spectrum of 2-CA150.WW2, see Figure 9c) on page 39. In addition, leucine shows a characteristic isopropyl-group signature in spectra of uniformly labelled samples. In a 100 ms PDSD spectrum of u-CA150.WW2 such a pattern at 26.9/25.1/23.6 ppm is observed showing very sharp lines. In the same spectrum, cross peaks appear at 55.1/22.6 and 55.1/21.7 ppm. These signals are assigned to leucine-$C\alpha/C\delta_{1/2}$ as all other residues with chemical shifts around 22 ppm, such as lysine, valine, or alanine were already assigned.

The isopropyl-group pattern with sharp lines observed in the 100 ms PDSD spectrum of u-CA150.WW2 shows correlations to $C\beta$ and $C\alpha$-signals only in direct carbon excitation d-CA150.WW2 spectra. In addition, no inter residue cross peaks are observed involving these resonances. The residue thus belongs, as $K_{37}$, to the small group of residues which give rise to only weak signals when using cross polarisation and are more intense in direct carbon excitation spectra. Another two residues, $Q_{34}$ and $E_{35}$ (see below), behave similarly, see Figure 10 on page 41. The identity of these four residues indicates that they belong to the C-terminus, QELK, which seems to exhibit increased mobility. Therefore, this leucine signal set was assigned to $L_{36}$.

The leucine-$C\alpha/C\delta_1$ and $C\alpha/C\delta_2$ cross peaks at 55.1/22.6 and 55.1/21.7 ppm, respectively, were thus assigned to $L_{26}$. They are well excited in CP-spectra and therefore the strong leucine $C\beta/C\gamma$ cross peak in 2-CA150.WW2 spectra is also assigned to this residue. In a 400 ms DARR spectrum of 1,3-CA150.WW2 a $L_{26}C\delta_{1/2}$/glutamic acid correlation pattern occurs and was assigned to $L_{26}/E_{27}$.

**Glutamine and glutamic acid ($E_7$, $E_{10}$, $E_{27}$, $E_{31}$, $Q_{34}$, $E_{35}$):** Glutamic acid and glutamine residues may be identified by their cross peak pattern involving side chain carbonyl resonances which show a downfield shift with respect to the backbone carbonyl sites. The cross peaks of different glutamine and glutamic acid residues in the aliphatic and the carbonyl region overlap strongly in CA150.WW2 spectra. Only two signal sets are isolated from the others. They both belong to the group of residues with stronger signals in direct carbon excitation spectra, Figure 10 on page 41. In addition, the side chain carbonyl resonance of one of these residues has an upfield chemical shift (180.4 ppm) compared to the others (around 184 ppm) which is characteristic for glutamine. From the C-terminal sequence, QELK, these two residues were assigned to $Q_{34}$ and $E_{35}$. All other assignments of glutamic acids were done using cross peaks to assigned neighbouring amino acids as described in the respective paragraphs.

**Asparagine ($N_{22}$, $N_{23}$):** The C$\alpha$/C$\beta$ cross peaks of the asparagines were identified by their chemical shifts around 52/42 ppm in F2/F1 in Figure 14c). Furthermore, strong C$\alpha$/C$\beta$ cross peaks appear in 1,3-CA150.WW2 spectra but not in spectra from 2-CA150.WW2. Two pairs of symmetric C$\alpha$/C$\beta$ cross peaks for the two asparagines are expected, however, several C$\alpha$/C$\beta$ cross peaks were observed. A $N_{23}$C$\alpha$-resonance was assigned due to cross peaks involving $R_{24}$C$\alpha$ (see above). On the basis of this assignment, in spectra of 1,3-CA150.WW2 the strongest asparagine C$\alpha$/C$\beta$-cross peaks at 51.1/41.5 and 52.3/45.2 ppm were tentatively assigned to $N_{22}$ and $N_{23}$, respectively. In spectra of d-CA150.WW2, four to five asparagine C$\alpha$/C$\beta$ cross peaks appear on each side of the diagonal. They are equally strong, and the assignment of $N_{22}$ and $N_{23}$ could not be safely transferred to spectra of d-CA150.WW2.

**Tyrosine, phenylalanine and tryptophan ($W_8$, $Y_{11}$, $F_{19}$, $Y_{20}$, $Y_{21}$, $W_{30}$):** The aromatic residues can be identified by cross peaks in the aromatic region of the $^{13}$C-$^{13}$C correlations at 100-160 ppm. However, in most spectra broadened lines appear in this region. Only the d-CA150.WW2 sample yields spectra that show well defined peaks. At a PDSD mixing time of 20 ms mainly C$\beta$/C$\gamma$ cross peaks are observed at 100-160/25-45 ppm, see Figure 15a). Two cross peaks appear at chemical shifts around 110 ppm in F2 and 29.9 ppm in F1, characteristic for the tryptophan C$\gamma$ and C$\beta$-sites. At longer mixing times, no sequential cross peaks were identified for the tryptophan signal sets, probably due to spectral overlap in spectra

of d-CA150.WW2. Nevertheless, both tryptophan C$\beta$-sites have the same chemical shift. $F_{19}$ was assigned on the basis of a cross peak at a characteristic chemical shift of 136.6/42.6 ppm in F2/F1, see Figure 15a). This is in accordance with the observed $T_{18}$/$F_{19}$ sequential contact, see above.

The C$\gamma$-signals of the three tyrosines are expected at chemical shifts around 130 ppm in F2. However, an excess of cross peaks occurs in this region, see Figure 15a). The same situation is encountered for the C$\alpha$/C$\beta$-cross peaks in spectra of d-CA150.WW2 and u-CA150.WW2. To investigate these residues in more detail, the sample s-CA150.WW2 was prepared which was selectively labelled with 2,3-$^{13}$C,$^{15}$N-phenylalanine, 2,3-$^{13}$C,$^{15}$N-tyrosine, u-$^{13}$C,$^{15}$N-alanine and u-$^{13}$C,$^{15}$N-glycine. A 20 ms PDSD spectrum of this sample is shown in Figure 15b). A large C$\alpha$/C$\beta$ cross peak conglomerate was observed for the phenylalanine and tyrosine residues. Clearly, more than the four expected C$\alpha$/C$\beta$ cross peaks occur which indicates the presence of multiple signals for each tyrosine residue. Thus the individual tyrosine signals could not be assigned. The aromatic C$\alpha$/C$\beta$ cross peak conglomerate is somewhat less broadened in spectra of the deuterated samples, see Figure 14c), and chemical shift estimates for the tyrosine residues were obtained from the centre of this peak at 56.3/43.9 ppm in F2/F1 (values are corrected for the isotope effect).

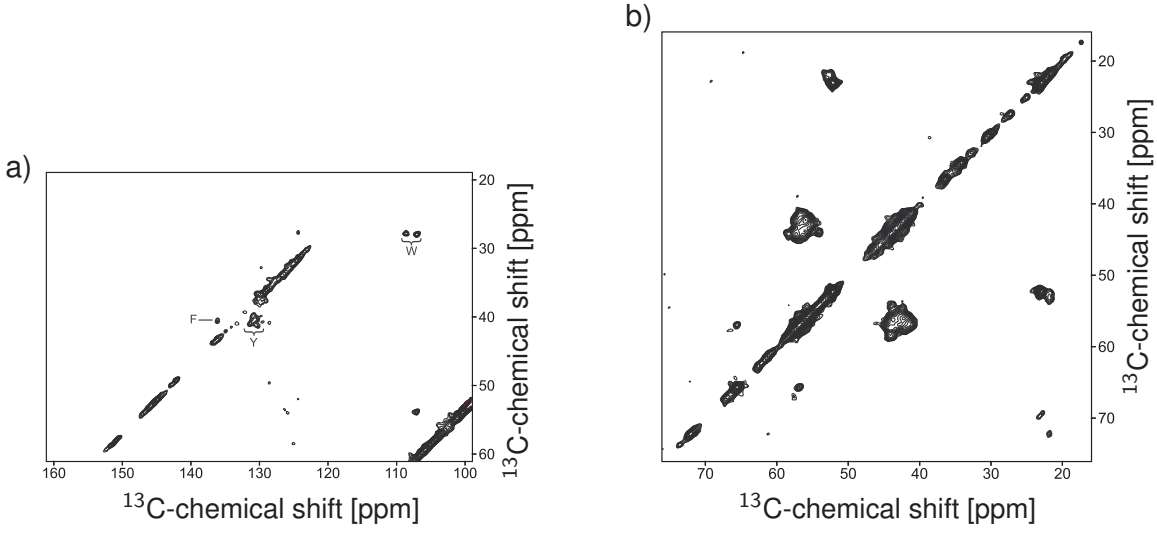



Figure 15: a) Aromatic region of the 20 ms PDSD spectrum of d-CA150.WW2, the C$\alpha$/C$\gamma$ cross peaks for the different types of aromatic amino acids are indicated, b) 20 ms PDSD spectrum of s-CA150.WW2 showing the C$\alpha$/C$\beta$-cross peak conglomerate of the phenylalanine and tyrosine residues.

### 4.3.3 Towards amide nitrogen and amide proton assignment

$^{13}$C-$^{15}$N correlations recorded on CA150.WW2 amyloid fibrils show spectral overlap which is most severe in the carbonyl/amide nitrogen region, see section 4.2. To increase the spectral resolution it is desireable to introduce an amide proton dimension. However, proton MAS NMR is challenging due to the large dipolar homonuclear interactions. To some extend this can be overcome by the application of homonuclear decoupling sequences during proton evolution. Well known examples are the frequency-switched and phase-modulated Lee-Goldberg decoupling schemes, see section 2.2. The observed proton line width can also be improved by reducing the number of protons in the sample by deuteration.

The d-CA150.WW2 sample which is protonated only at the exchangeable sites was used to record spectra with a $^{1}$H-dimension at 20 kHz MAS. At this MAS-frequency the proton-proton dipolar couplings are only partially removed, but the situation is strongly improved compared to a fully protonated sample. Figure 16 shows a $^{1}$H-$^{13}$C correlation. The signal intensity in the proton aliphatic region results from incomplete deuteration of the aliphatic sites, which is most severe for the methyl groups (0-2 ppm). This phenomenon is typical for biosynthetically prepared proteins. In the amide proton chemical shift region (7-11 ppm) correlations to the C$\alpha$-atoms are visible (50-65 ppm). In addition, sharp signals appear at a water proton chemical shift of 4.8 ppm. They occur either due to relayed transfer or because of chemical exchange with strongly bound water during the cross polarisation step. From the carbon chemical shifts of these cross peaks it can be concluded that mainly lysines are in contact with tightly bound

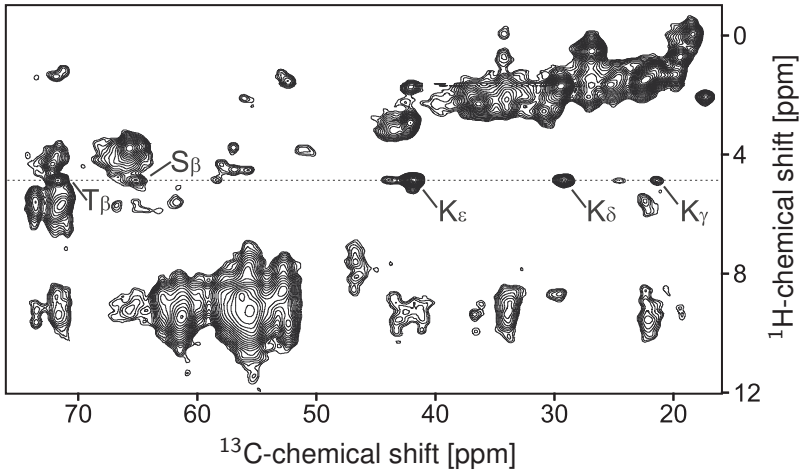



Figure 16: $^{1}$H-$^{13}$C correlation of d-CA150.WW2, the dashed line indicates the water proton chemical shift at 4.8 ppm, residues in contact with water are indicated.

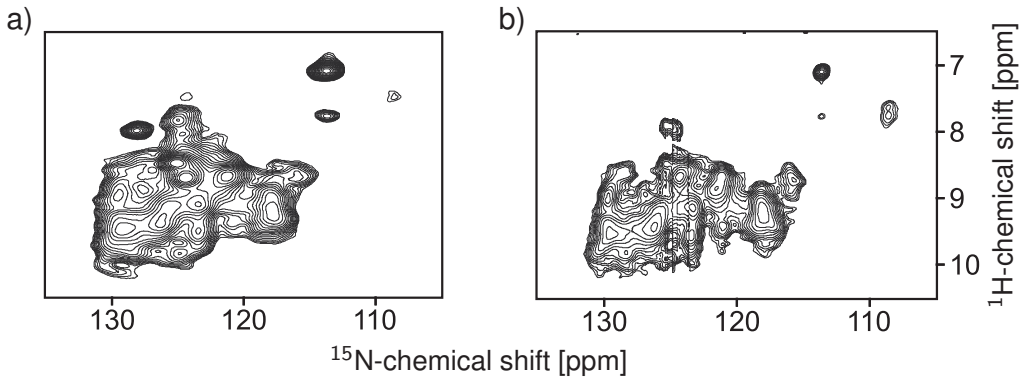



Figure 17: $^{1}$H-$^{15}$N correlations of d-CA150.WW2, a) $^{15}$N or b) $^{1}$H-detected, respectively.

water molecules but also at least one serine and a threonine residue, see Figure 16.

A $^{1}$H-$^{15}$N correlation is shown in Figure 17a). Strong signal overlap occurs, but nevertheless some resolution is apparent from the contours. A very similar spectrum, Figure 17b), is obtained using direct proton detection in combination with water suppression following the scheme by Paulson et al. [Paulson 03], although not the same high level of water suppression as reported in the original publication was observed.

In summary, spectra with a proton dimension recorded on d-CA150.WW2 at 20 kHz MAS show some resolution that used as a third dimension might be suitable to increase the spectral dispersion in $^{13}$C-$^{15}$N correlations, see the following section 4.3.4.

### 4.3.4 3D-HNC-TEDOR – a new experimental scheme

The resolution observed in $^{13}$C-$^{15}$N spectra was not sufficient to assign the amide nitrogen signals of CA150.WW2 amyloid fibrils on the basis of these spectra alone, see section 4.2. To overcome this problem, it would be useful to record $^{13}$C-$^{15}$N correlations with a proton dimension to increase the spectral resolution. In the previous section 4.3.3 it was shown that the proton resolution in spectra of d-CA150.WW2 at 20 kHz MAS might be sufficient to yield resolved cross peaks in such a three-dimensional spectrum.

A three-dimensional experimental scheme was assembled by combining the TEDOR sequence with a proton evolution. In two-dimensional TEDOR experiments, Figure 18a), the nitrogen evolution period is surrounded by two symmetric REDOR pulse trains. This refocuses the homonuclear $^{13}$C-dipolar couplings which otherwise would lead to a dephasing of the $^{13}$C-magnetisation in multiply labelled samples. For the three-dimensional experiment it is, however, desirable to start with a proton evolution period followed by a transfer of

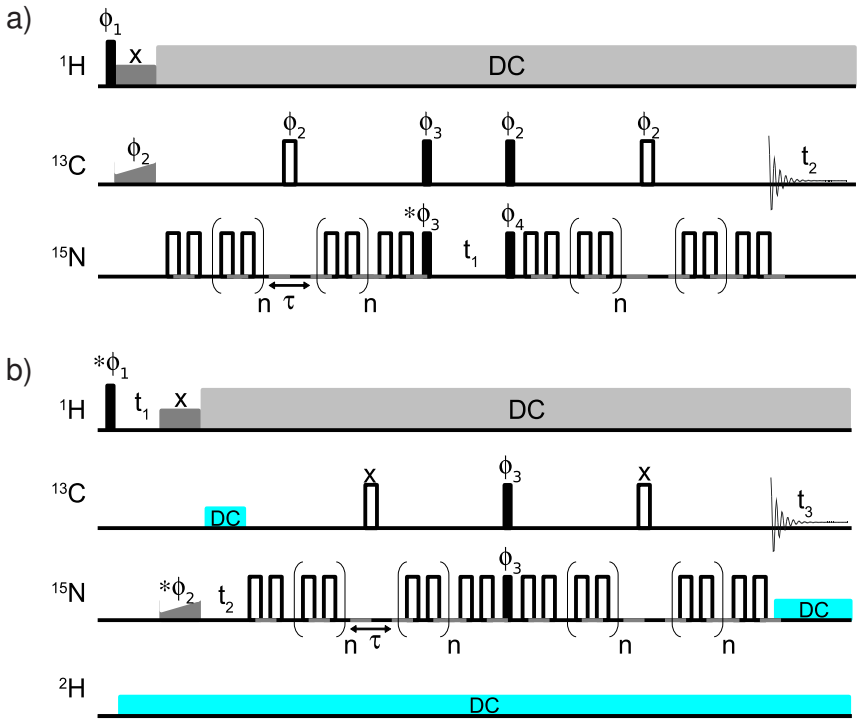



Figure 18: TEDOR pulse sequences. $t_1$, $t_2$, and $t_3$ are the direct or indirect evolution times, filled bares represent $\frac{\pi}{2}$-pulses, open bars $\pi$-pulses. Decoupling periods are indicated by DC-boxes, all other boxes represent CP-irradiation. $\tau$ is the length of a rotor period. The phase that is incremented during indirect detection (TPPI [Marion 83]) is indicated by an asterisk. a) 2D-NC-TEDOR experiment following the experimental schemes published by Michal et al. [Michal 97] and Jaroniec at al. [Jaroniec 00], see section 3.3.2. Phase cycling: $\phi_1 = y\bar{y}$, $\phi_2 = \overline{xxyy}\ xxyy$, $\phi_3 = xxyy\ \overline{xxyy}$, $\phi_4 = \overline{xxyy}\ xxyy\ xxyy\ \overline{xxyy}$, $\phi_{\text{rec}} = x\bar{x}y\bar{y}\ \bar{x}x\bar{y}y\ \bar{x}x\bar{y}y\ x\bar{x}y\bar{y}$, $\pi$-pulses *xyxy yxyx* $\overline{xyxy}$ $\overline{yxyx}$. b) 3D-HNC-TEDOR experiment, blue rectangles indicate decoupling units that might further improve the observed line widths, see main text. Phase cycling: $\phi_1 = y\bar{y}$, $\phi_2 = \overline{xxyy}\ xxyy$, $\phi_3 = xxyy\ \overline{xxyy}$, $\phi_{\text{rec}} = x\bar{x}y\bar{y}\ \bar{x}x\bar{y}y$, $\pi$-pulses *xyxy yxyx* $\overline{xyxy}$ $\overline{yxyx}$.

the magnetisation to the amide nitrogens. At the same time, carbon should be employed for detection for several reasons: in the $^{13}$C-dimension a large spectral window has to be recorded; the sensitivity of the $^{13}$C-nucleus is much better compared to the $^{15}$N-nucleus; proton detection is difficult due to the lack of probes with gradient coils.

The assembled three-dimensional HNC-TEDOR experiment is shown in Figure 18b). It starts with a proton evolution period that is followed by a short CP-transfer to nitrogen. Then nitrogen evolution takes place during $t_2$. A subsequent TEDOR-transfer is used to create $^{13}$C-magnetisation for detection. Thereby, the symmetry of the pulse train is kept as in the two-dimensional TEDOR experiment, to refocuses the homonuclear $^{13}$C-dipolar couplings.

The TEDOR mixing time (1.2 ms) was chosen such, that only a one-bond carbon-nitrogen magnetisation transfer occurred. In summary, the experiment yields spectra with $H_iN_iC\alpha_i$ and $H_iN_iCO_{i-1}$ cross peaks.

Figure 19 shows subspectra of the 3D-HNC-TEDOR experiment. The $^{13}C$-$^{15}N$ subspectra in the C$\alpha$ and the CO-region can be used to correlate the C$\alpha$-atoms from each residue with the CO-atom from the preceding residue, Figure 19a). The obtained proton resolution is shown in a $^{1}H$-$^{15}N$ subspectrum, Figure 19b). Using and at the same time confirming the carbon assignment, 14 amide nitrogen and proton resonances were assigned from the three-dimensional HNC-TEDOR spectrum. Another 15 amide nitrogen and proton spin systems were identified but could not be assigned unambiguously due to the overlap of the corresponding CO and C$\alpha$-resonances. It can thus be concluded that most of the expected HNC$\alpha$ and HNCO cross peaks are visible in the spectrum. The application of longer TEDOR mixing times to additionally obtain correlations to the C$\beta$-atoms was not successful due to the low signal-to-noise ratio of experiments employing long TEDOR pulse trains.

The main source of line broadening in the 3D-HNC-TEDOR spectrum are the residual proton heteronuclear and homonuclear couplings. Other contributions are the deuterium-carbon and the nitrogen-carbon scalar couplings ($N_i/C\alpha_i$ = 11 Hz, $N_i/C\alpha_{i-1}$ = -7 Hz, $N_i/CO_{i-1}$ = 15 Hz). To further improve the resolution in 3D-HNC-TEDOR spectra it might thus be desirable to apply carbon-deuterium decoupling and carbon-nitrogen decoupling schemes, as indicated

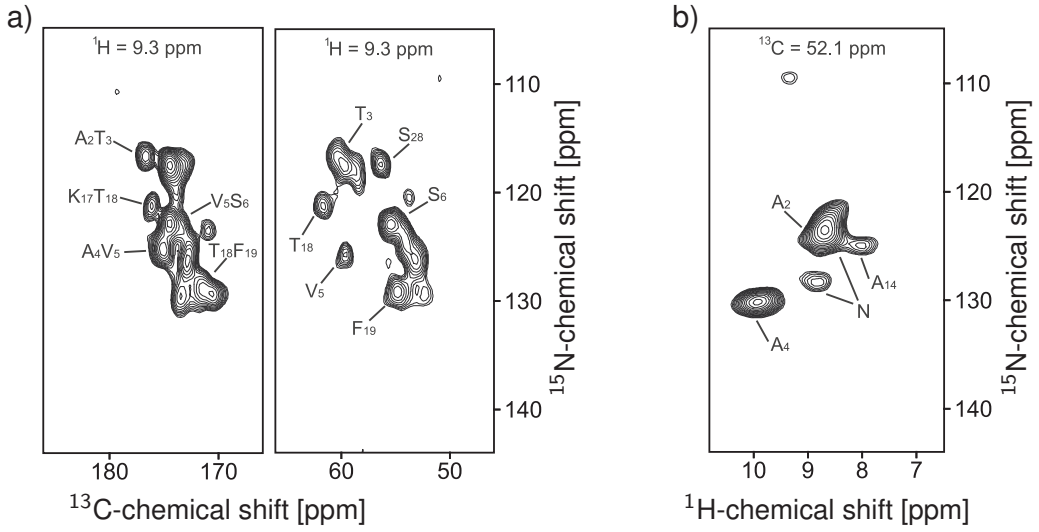



Figure 19: a) CO and C$\alpha$-regions of a $^{13}C$-$^{15}N$ subspectrum of the three-dimensional HNC-TEDOR spectrum at a proton chemical shift of 9.3 ppm, b) $^{1}H$-$^{15}N$ subspectrum of the same experiment at a carbon chemical shift of 52.1 ppm. Assigned cross peaks are indicated.

by the blue rectangles in Figure 18. In addition, it should be possible to increase the resolution in the proton dimension by homonuclear decoupling schemes (frequency-switched or phase-modulated Lee-Goldberg).

## 4.4 Assignment of the natively folded CA150.WW2 domain in solution

The chemical shifts of CA150.WW2 amyloid fibrils can be used to obtain information on the type and the position of secondary structure elements. It is known that amyloid fibrils have a high $\beta$-sheet content. This is also the case for the native WW-fold of CA150.WW2. Large structural differences between the amyloid and the native WW-fold should, nevertheless, result in different chemical shifts of the peptide in its two forms. To compare them, the solution state NMR spectra of native CA150.WW2 were assigned. For this study *u-CA150.WW2* was dissolved in the fibrillation buffer to a concentration of 0.8 mM. A standard set of triple resonance spectra (see section 3.3) was recorded and yielded assignments for most of the residues, see Table 6. No HSQC signals were detected for residues $S_{-1}$, $D_{15}$, $T_{18}$, $F_{19}$, and $E_{27}$, either due to signal broadening or interference with an imidazole artifact which hindered the detection of week signals with proton chemical shifts around 7.01 and 7.95 ppm. However, analysis of triple resonance spectra led in most cases to the assignment of carbon sites for these residues. $T_{18}$ is the only residue for which no atom could be assigned.

## 4.5 Experimental constraints on the fibrillar structure

### 4.5.1 Secondary structure prediction – dihedral angle constraints

The chemical shift of the main chain atoms is sensitive to the backbone conformation. It can be used to predict secondary structure elements for example with the program TALOS [Cornilescu 99]. The software predicts the backbone conformation by comparing up to five chemical shifts (HA, N, C$\alpha$, C$\beta$, and CO) of the respective residue and its two neighbours with the chemical shifts of similar triplets from proteins with known structure. From the ten results with the best chemical shift agreement, the backbone dihedral angles are compared. If all or most of them are similar, the analysed residue is likely to be in the same range.

The MAS NMR chemical shifts of CA150.WW2, see Table 5, were used as input for the program TALOS, excluding the entries for residues $T_3$ and $S_{28}$ which show peak doubling. The

| residue | N | H | CO | C$\alpha$ | C$\beta$ |
|---|---|---|---|---|---|
| $G_{-2}$ | | | | | |
| $S_{-1}$ | | | 174.63 | 58.20 | 63.61 |
| $M_0$ | 122.39 | 8.47 | 176.60 | 55.33 | 32.10 |
| $G_1$ | 110.19 | 8.25 | 173.60 | 44.86 | |
| $A_2$ | 123.82 | 8.06 | 178.01 | 52.32 | 19.38 |
| $T_3$ | 113.64 | 8.01 | 174.05 | 61.52 | 69.64 |
| $A_4$ | 126.74 | 8.15 | 177.51 | 52.18 | 19.39 |
| $V_5$ | 119.46 | 8.01 | 176.08 | 62.02 | 32.58 |
| $S_6$ | 118.97 | 8.21 | 174.88 | 57.99 | 63.85 |
| $E_7$ | 124.29 | 8.61 | 175.38 | 57.20 | 29.35 |
| $W_8$ | 118.31 | 8.04 | 176.53 | 55.98 | 31.95 |
| $T_9$ | 119.06 | 9.24 | 171.99 | 61.68 | 71.54 |
| $E_{10}$ | 126.29 | 8.30 | 173.90 | 54.87 | 31.62 |
| $Y_{11}$ | 125.71 | 8.34 | 173.20 | 56.52 | 41.67 |
| $K_{12}$ | 117.98 | 8.10 | 178.07 | 53.93 | 35.26 |
| $T_{13}$ | 113.45 | 8.92 | 177.45 | | |
| $A_{14}$ | 124.58 | 8.99 | | 54.58 | 18.27 |
| $D_{15}$ | | | 176.30 | 52.70 | 39.45 |
| $G_{16}$ | 107.37 | 7.81 | 174.36 | 45.75 | |
| $K_{17}$ | 121.09 | 7.63 | | 55.53 | 33.50 |
| $T_{18}$ | | | | | |
| $F_{19}$ | | | 170.63 | 54.96 | 39.55 |
| $Y_{20}$ | 117.16 | 8.64 | 174.25 | 56.22 | 40.64 |
| $Y_{21}$ | 124.07 | 9.12 | 173.19 | 55.79 | 43.04 |
| $N_{22}$ | 129.75 | 8.10 | 174.45 | 51.23 | 38.38 |
| $N_{23}$ | 121.52 | 8.22 | 174.52 | 54.85 | 38.66 |
| $R_{24}$ | 119.43 | 8.31 | 177.33 | 57.33 | 30.57 |
| $T_{25}$ | 109.33 | 7.74 | 175.77 | 61.47 | 69.26 |
| $L_{26}$ | 116.52 | 7.61 | | | |
| $E_{27}$ | | | 174.74 | 56.00 | 31.46 |
| $S_{28}$ | 115.86 | 8.31 | 173.70 | 56.63 | 66.12 |
| $T_{29}$ | 117.08 | 9.12 | 174.43 | 59.10 | 69.61 |
| $W_{30}$ | 125.21 | 8.49 | 175.81 | 57.40 | 29.35 |
| $E_{31}$ | 121.70 | 8.08 | 174.51 | 54.52 | 30.58 |
| $K_{32}$ | 127.05 | 8.22 | | 54.49 | 32.37 |
| $P_{33}$ | | | 176.91 | 61.85 | 31.70 |
| $Q_{34}$ | 122.62 | 8.35 | 177.24 | 57.61 | 28.32 |
| $E_{35}$ | 117.21 | 8.92 | 176.27 | 57.49 | 29.30 |
| $L_{36}$ | 120.00 | 7.33 | 175.72 | 54.00 | 41.86 |
| $K_{37}$ | 126.67 | 7.41 | | 58.04 | 33.01 |

Table 6: Solution state NMR assignments in ppm of the natively folded CA150.WW2 domain.

dihedral angle predictions were accepted when $\geq$9 out of the 10 similar triplets had dihedral angles in the same region of the Ramachandran Plot. In addition, only results from residues with more than one assigned site were regarded as highly reliable, filled symbols in Figure 20. Chemical shift estimates for the tyrosine and asparagine residues, which were obtained as described in the respective paragraphs in section 4.3.2, were used as additional input in a second TALOS run. These results and predictions for residues which have only one assigned chemical shift are given with open symbols in Figure 20. Two long stretches rich in $\beta$-sheet conformation, ranging from residues $A_4$ to $T_{13}$ and $K_{17}$ to $K_{32}$, respectively, are observed. No TALOS predictions could be obtained for the N and C-termini and the region $A_{14}$-$G_{16}$. This indicates that the fibrillar fold contains two long $\beta$-strands separated by a loop region.

Figure 20 shows TALOS predictions also for CA150.WW2 in its native fold. They were obtained using the chemical shifts from the solution NMR assignment, see Table 6. Similarly to fibrillar CA150.WW2, the majority of the residues are in the $\beta$-strand conformation as expected from the known WW-fold. The predictions compare well with the structure determined

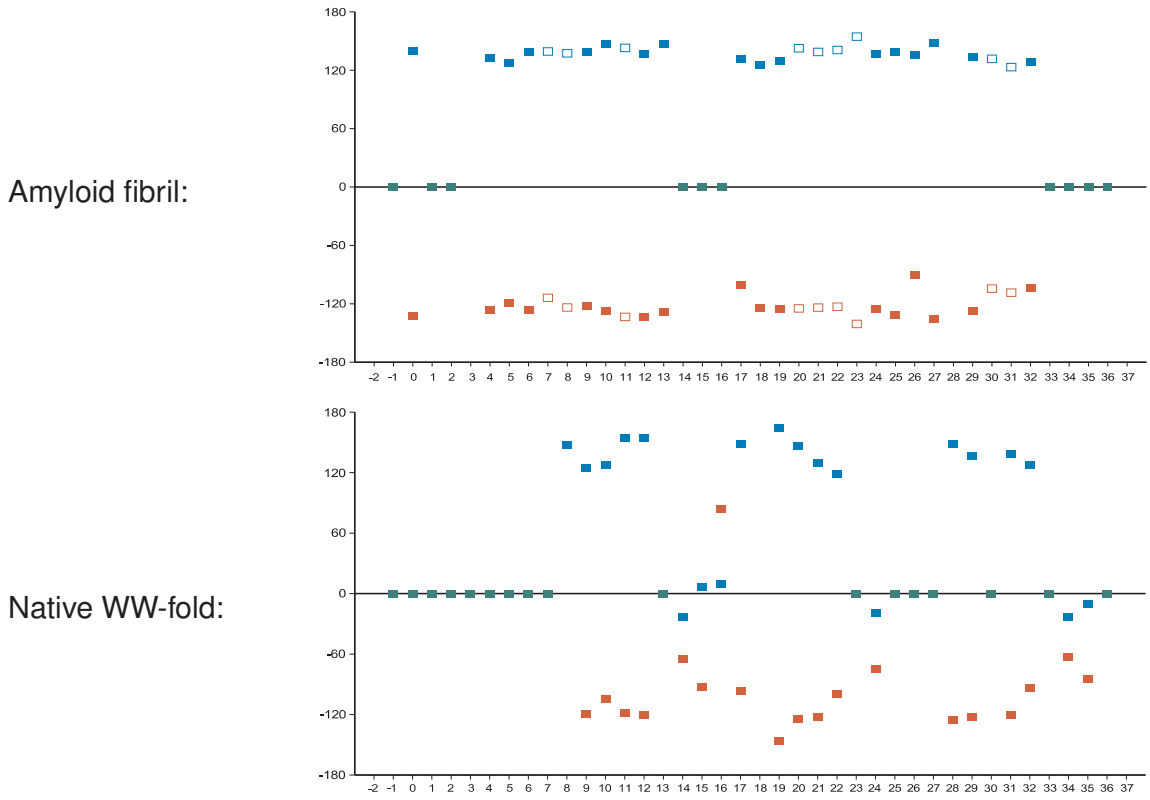



Figure 20: TALOS secondary structure prediction for CA150.WW2 in the amyloid and the WW-fold. Filled symbols indicated reliable predictions, open symbols indicate less reliable predictions, see main text. Blue: $\psi$-dihedral angles, red: $\varphi$-dihedral angles, green: inconsistent TALOS predictions.

by Macias et al. [Macias 00] for the wild type sequence (with GSM-tag), which has an unstructured N-terminus and three antiparallel $\beta$-strands ranging from $W_8$-$T_{13}$, $K_{17}$-$N_{23}$, and $L_{26}$-$E_{31}$, respectively. The position of the first loop, $A_{14}$-$G_{16}$, coincides with the loop region in the amyloid fibrils. However, in the fibrillar fold the N-terminus is not unstructured and the residues that form the second loop region in the native fold are part of the second $\beta$-strand in the fibril. This indicates that the amyloid fibril fold significantly differs from the WW-fold. In addition, the overall $\beta$-strand content of the WW-fold is lower than in the fibrillar fold.

### 4.5.2 Long range distance constraints from MAS NMR spectra

Using the resonance assignment of CA150.WW2 amyloid fibrils, homo- and heteronuclear correlation spectra were investigated for non-sequential cross peaks. They are observed between residues close in space but distant in sequence. Generally, they occur at longer mixing times compared to e.g. C$\alpha$/C$\alpha$ sequential correlations due to the larger inter atomic distances. Cross peak patterns between six non-sequential residue pairs were detected comprising 29 unambiguous distance constraints, see Table 8 in section 4.7. Especially valuable for this procedure was the usage of spectra recorded on the glycerol labelled samples, 1,3-CA150.WW2 and 2-CA150.WW2, due to the reduced spectral overlap.

Figure 21a+b) shows $^{13}$C-$^{13}$C correlations with long mixing times of these samples. The 20 ms PDSD spectrum of 2-CA150.WW2, Figure 9c) on page 39, contains in its aliphatic region only two strong cross peaks, $V_5$C$\alpha$/C$\beta$ (60.9/34.5 ppm) and $L_{26}$C$\beta$/C$\gamma$ (42.2/27.5 ppm). In addition, weaker lysine C$\alpha$/C$\gamma$ signals are observed. At longer mixing times, cross peaks occur that correlate the valine and leucine signals with each other via a complete cross peak pattern, see Figure 21a). The inter residue cross peaks involving the valine C$\beta$-resonance at 34.5 ppm are stronger and appear at shorter mixing times than the ones from the valine C$\alpha$-resonance at 60.9 ppm, indicating that $V_5$ and $L_{26}$ are in contact with each other via their side chains. Around 22/19 ppm correlations between the methyl groups of both residues could be observed in a 400 ms DARR spectrum of 1,3-CA150.WW2, see Figure 21b). This spectrum also shows cross peaks that match $V_5$C$\gamma_1$/$L_{26}$C$\alpha$ (18.8/55.1 ppm) and $V_5$C$\gamma_2$/$L_{26}$C$\alpha$ (19.7/55.1 ppm) contacts. However, the $L_{26}$C$\alpha$ chemical shift is not unique and contributions to these signals from other residues cannot be excluded. Based on the other observed long range contacts (see below) these cross peaks might also be assigned to $V_5$C$\gamma_1$/$R_{24}$C$\alpha$ and

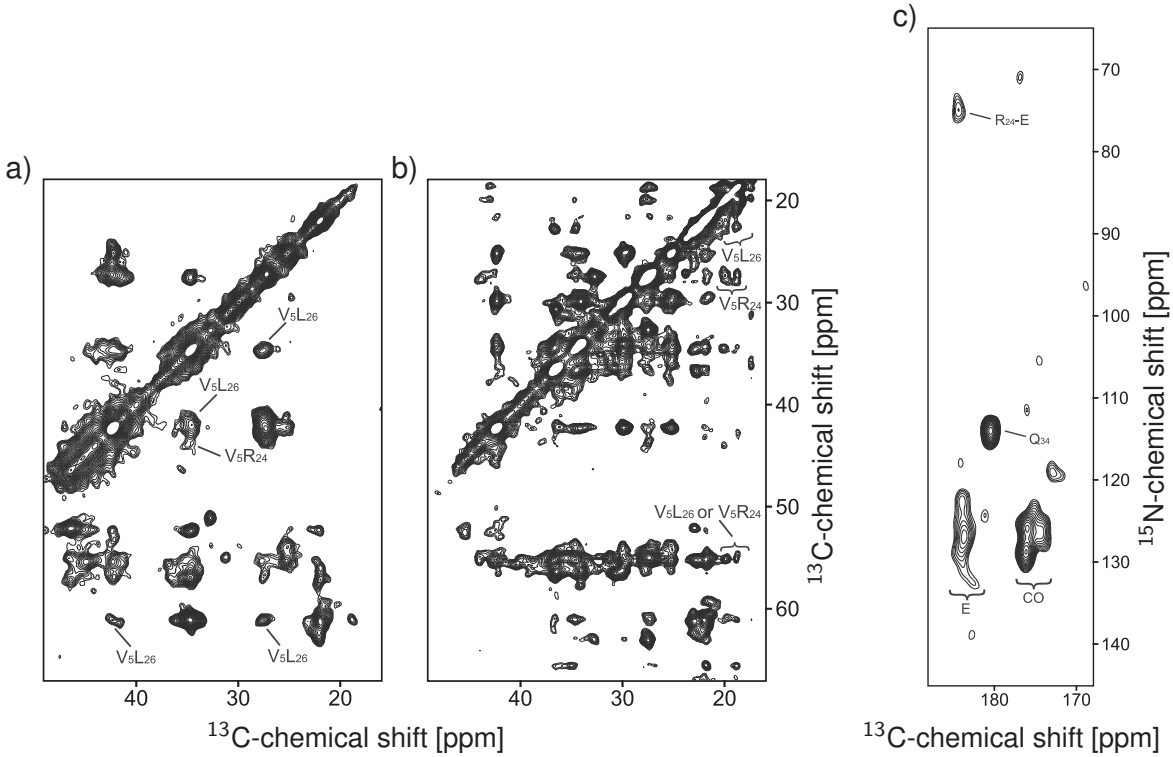



Figure 21: Spectra showing long range correlations a) 400 ms PDSD spectrum of 2-CA150.WW2, b) 400 ms DARR spectrum of 1,3-CA150.WW2, c) 12.8 ms TEDOR spectrum of 2-CA150.WW2.

$V_5C\gamma_2/R_{24}C\alpha$.

In the 400 ms PDSD spectrum of 2-CA150.WW2, see Figure 21a), a correlation involving $V_5C\beta$ is observed at 34.5/43.7 ppm. Taking the labelling pattern into account this cross peak can only result from a $V_5C\beta/R_{24}C\delta$ contact. This is further supported by the observation of $V_5C\gamma_1/R_{24}C\gamma$ (18.8/27.3 ppm) and $V_5C\gamma_2/R_{24}C\gamma$ (19.7/27.3 ppm) cross peaks in $^{13}C$-$^{13}C$ correlations of 1,3-CA150.WW2, see Figure 21b).

At a relatively short mixing time (100 ms) a cross peak at 52.2/57.4 ppm is observed in spectra of 2-CA150.WW2. The chemical shift of 52.2 ppm is characteristic for the $A_2C\alpha$ and $A_4C\alpha$-resonances. However, their sequential neighbours could be unambiguously identified and have $C\alpha$-chemical shifts very different from 57.4 ppm. The only resonances that match this value are the $S_{28}C\alpha$ and $S'_{28}C\alpha$-resonances. This is in accordance with the observation of correlations from the $S_{28}C\beta$ and $S'_{28}C\beta$-signals to the alanine $C\alpha$-resonance in a 700 ms DARR spectrum of d-CA150.WW2. Thus $S_{28}$ and $S'_{28}$ are in contact with either $A_2$ or $A_4$. In addition, the $S_{28}C\alpha$-resonance shows a correlation to the unique $T_3C\beta$ chemical shift at 73.7 ppm in the 700 ms DARR spectrum of 2-CA150.WW2.

In the 400 ms DARR spectrum recorded on d-CA150.WW2 only very few long range cross peaks are observed due to the slow DARR magnetisation transfer in this sample. Nevertheless, a complete C$\alpha$/C$\beta$ cross peak pattern between $T_{13}$ and $T_{18}$ was observed in this spectrum, grey labels in Figure 14b) on page 47. The magnetisation transfer between these two residues is efficient due to the protons in the threonine hydroxyl-groups.

In $^{13}$C-$^{15}$N TEDOR spectra with long mixing times inter residue cross peaks can occur. This is especially interesting for resonances with isolated $^{15}$N-chemical shifts such as lysine or arginine side chains because the arising cross peaks do not overlap with other signals in the crowded nitrogen amide region. Figure 21c) shows a TEDOR spectrum of 2-CA150.WW2 with a long mixing time of 12.8 ms. Comparison to spectra recorded with short mixing times, see Figure 12 on page 42, shows that the carbonyl/amide cross peaks (170-180/120-135 ppm) are already strongly relaxed. Further signals appear at a carbon chemical shift of 181-186 ppm characteristic for the C$\delta$-site of the glutamic acid residues. From the observed nitrogen chemical shifts of these cross peaks at 120-130 ppm, it is concluded that they arise from correlations to amide nitrogens. However, the resolution is too low to assign these peaks unambiguously. One glutamic acid shows an additional correlations to a nitrogen resonance at 75 ppm which is characteristic for the asparagine guanidinium group. Thus a contact between the side chain of the only asparagine, $R_{24}$, and the acidic group of one of the glutamic acids is observed. Due to the charged character of the head groups of these amino acids formation of a salt-bridge or a hydrogen bond is likely.

Interestingly, in the TEDOR spectrum a strong cross peak is observed at a carbon chemical shift of 180.5 ppm. It was assigned to $Q_{34}$C$\delta$/N$\epsilon$ although the cross peak is very week in spectra with short mixing times characteristic for such a one bond transfer (Figure 12 on page 42). This can be explained by the observed mobility of $Q_{34}$ which scales the dipolar interaction.

In summary, non-trivial contacts were observed between residues $T_{13}$/$T_{18}$, $S_{28}$/$T_3$, $S_{28}$/$A_{2or4}$, $V_5$/$L_{26}$, $V_5$/$R_{24}$, and $R_{24}$/$E_?$. Most of these interactions connect the two $\beta$-strands identified by TALOS analysis, section 4.5.1, indicating that the two $\beta$-strands interact with each other via their side chains as observed for other amyloid fibrils, such as A$\beta$(1-40) [Petkova 06] or the K3-peptide from $\beta_2$-microglobulin [Iwata 06].

### 4.5.3 Constraints from other methods: Optical diffraction and mutagenesis

Amyloid fibrils exhibit a typical so-called cross-$\beta$ diffraction pattern which was also found for CA150.WW2 fibrils. Figure 22 shows the optical diffraction pattern of oriented CA150.WW2 tubes which was recorded by John Berriman at the New York SBC. All reflexes occur in a duplicate manner due to the tubular structure of fibrillar CA150.WW2, see Figure 8a) on page 37. The reflexes at the farthest distance (wide-angle) from the centre of the diffraction pattern correspond to a 4.73 Å spacing and are characteristic for the strand-strand separation in $\beta$-sheets. The small-angle reflexes close to the centre of the diffraction pattern arise from a 30 Å distance perpendicular to the 4.73 Å spacing. Another small-angle reflex of lower intensity corresponds to a 10 Å distance. The two small-angle reflexes describe the dimensions of the protofilaments. Their diameter is thus 30 Å and the distance between the $\beta$-sheets 10 Å.

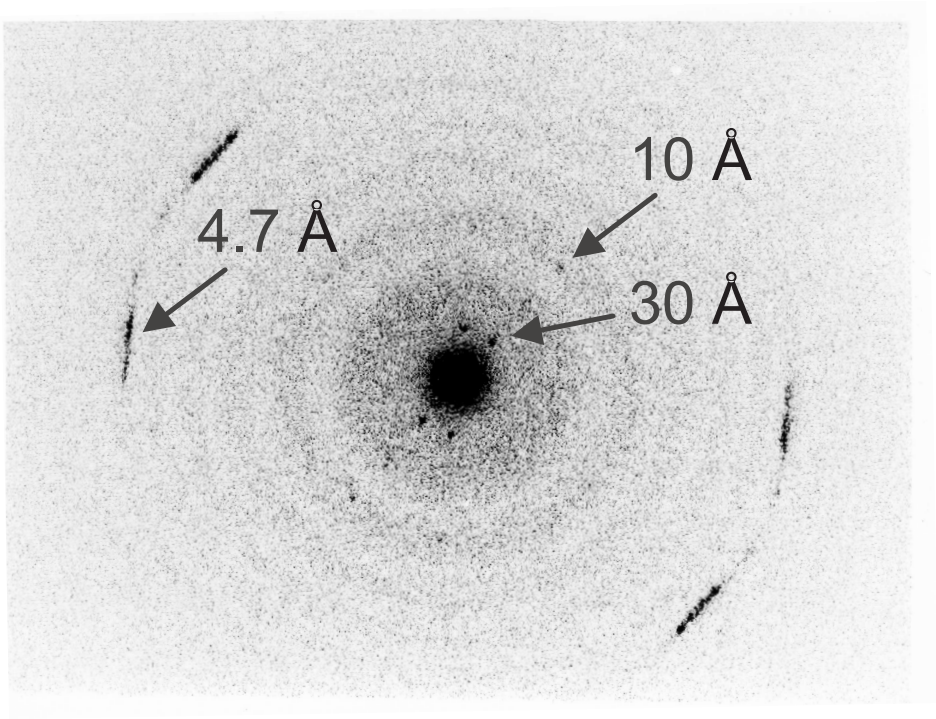



Figure 22: Optical diffraction pattern of an oriented CA150.WW2 tube, arrows indicate reflexes, see main text.

Parallel to the NMR analysis, a mutational study of wild type CA150.WW2 was carried out by Neil Ferguson and Henning Tidow at the MRC in Cambridge. Here, their results are summarised. Nearly all residues between $T_3$ and $L_{36}$ were mutated to alanine or glycine (in case of alanine) and their ability to form amyloid fibrils was monitored, by a thioflavin T

assay, see section 3.2. In addition, the fibrillation kinetics of the mutants was determined by measuring the absorbance at 350 nm.

Only few of the mutants were not investigated due to low expression levels, formation of inclusion bodies, or because they were not soluble anymore after lyophilisation: $W_8A$, $N_{22}A$, $T_{25}A$, $E_{31}A$, $K_{32}A$, and $P_{33}A$. Some of these residues are crucial for stabilising the native WW-fold ($W_8$ and $P_{33}$) and thus replacing them leads to unfolded protein which is quickly degraded by proteases or deposited in inclusion bodies. Interestingly, some of the investigated alanine variants did not form amyloid fibrils at all: $V_5A$, $Y_{11}A$, $Y_{20}A$, $R_{24}A$, $L_{26}A$, $T_{29}A$, and $W_{30}A$. For all other mutants the apparent half-time of the fibril growth curves was determined. Three types of variants were distinguished with respect to their kinetic properties in relation to the wild type. An increased rate of fibril formation was observed for the variants $E_{10}A$, $T_{13}A$, $A_{14}G$, $D_{15}A$, $G_{16}A$, $K_{17}A$, and $Y_{21}A$. A slower rate was found for $T_3A$, $E_7A$, and $Q_{34}A$. The mutants of all other residues showed fibrillation kinetics similar to the wild type, see Table 7.

| Variant | $\tau_{\frac{1}{2}}$ [min] | Variant | $\tau_{\frac{1}{2}}$ [min] |
|---|---|---|---|
| WT | $300 \pm 10$ | $Y_{20}A$ | No fibril |
| $T_3A$ | $620 \pm 30$ | $Y_{21}A$ | $30 \pm 2$ |
| $A_4G$ | $330 \pm 60$ | $N_{22}A$ | No data |
| $V_5A$ | No fibril | $N_{23}A$ | $260 \pm 8$ |
| $S_6A$ | $320 \pm 40$ | $R_{24}A$ | No fibril |
| $E_7A$ | $1600 \pm 200$ | $T_{25}A$ | No data |
| $W_8A$ | No data | $L_{26}A$ | No fibril |
| $T_9A$ | $340 \pm 40$ | $E_{27}A$ | $180 \pm 10$ |
| $E_{10}A$ | $170 \pm 15$ | $S_{28}A$ | $230 \pm 15$ |
| $Y_{11}A$ | No fibril | $T_{29}A$ | No fibril |
| $K_{12}A$ | $270 \pm 40$ | $W_{30}A$ | No fibril |
| $T_{13}A$ | $120 \pm 10$ | $E_{31}A$ | No data |
| $A_{14}G$ | $40 \pm 5$ | $K_{32}A$ | No data |
| $D_{15}A$ | $70 \pm 2$ | $P_{33}A$ | No data |
| $G_{16}A$ | $140 \pm 10$ | $Q_{34}A$ | $630 \pm 40$ |
| $K_{17}A$ | $40 \pm 5$ | $E_{35}A$ | $190 \pm 20$ |
| $T_{18}A$ | $310 \pm 40$ | $L_{36}A$ | $240 \pm 80$ |
| $Y_{19}A$ | $255 \pm 5$ | | |

Table 7: Half-time of the fibril formation process ($\tau_{\frac{1}{2}}$) for several mutants.

## 4.6 The fold of the monomeric unit in CA150.WW2 fibrils

Prior to structure calculation, the fold of the monomeric units in the CA150.WW2 amyloid fibrils was determined using MAS NMR chemical shifts, some long range distance constraints, electron microscopy, and alanine scanning data. The basic secondary structure elements, two long $\beta$-strands separated by a loop region, were identified by TALOS analysis. Both $\beta$-strands are in contact with each other via their side chains which has been shown by MAS NMR distance constraints between residues $T_3$/$S_{28}$, $V_5$/$R_{24}$, $V_5$/$L_{26}$, and $T_{13}$/$T_{18}$. In $\beta$-strand one ($A_4$-$T_{13}$) the odd numbered residues ($T_3$, $V_5$, $E_7$, $T_9$, $Y_{11}$, $T_{13}$) and in $\beta$-strand two ($K_{17}$-$K_{32}$) the even numbered residues ($T_{18}$, $Y_{20}$, $N_{22}$, $R_{24}$, $L_{26}$, $S_{28}$) are oriented towards the interior. A sketch of the described model is shown in Figure 23. The observed distance constraints are indicated by black arrows.

In the region $T_9$, $Y_{11}$, and $Y_{20}$, $N_{22}$ no such long range distance constraints could be detected. In addition, for many of these residues ($Y_{11}$, $Y_{20}$, $N_{22}$) very heterogeneous line shapes are observed in the MAS NMR spectra, see section 4.3.2. This indicates that the structure is less ordered in this region.

The model is in line with the results from the mutational analysis, see section 4.5.3. Residues which cannot be replaced by alanine without removing the ability of CA150.WW2

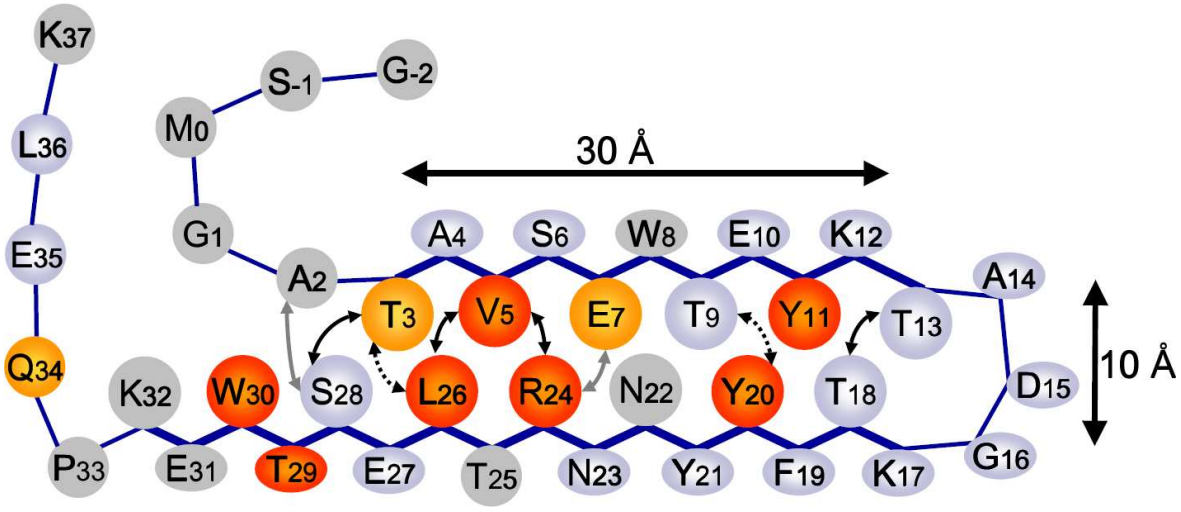



Figure 23: Sketch of the monomeric unit of CA150.WW2 fibrils; alanine scanning results are indicated by the colour of the residues: no fibril formation (red), very slow fibril formation (orange), wild-type or fast fibril formation (blue), no data (grey); long range contacts are indicated by arrows, unambiguous (black), unambiguous based on the model (grey), ambiguous (dashed); residues that show a $\beta$-strand conformation in TALOS analysis are connected with a thick line; the distances determined by electron microscopy are indicated.

to form amyloid fibrils at physiological conditions are coloured in red in Figure 23. In addition, alanine mutants which form fibrils very slowly are given in orange. Most of these residues are situated in the interface between the two strands: $T_3$, $V_5$, and $E_7$ in $\beta$-strand one and $Y_{20}$, $R_{24}$, and $L_{26}$ in $\beta$-strand two. Interestingly, the residues in the ordered part of the amyloid core region have side chains of similar size, except for $R_{24}$. This indicates that the strand-strand interaction might be steric in nature as observed by Eisenberg et al. [Nelson 05, Sawaya 07] in the steric zipper structures of fibril-like crystals.

Based on this structural model for the monomeric unit the observed long range contacts $S_{28}/A_{2or4}$ and $R_{24}/E_?$, see section 4.5.2, were assigned to $S_{28}/A_2$ and $R_{24}/E_7$ (grey arrows in Figure 23). From the model further long range contacts are expected. Cross peaks that could result from $T_3/L_{26}$ and $T_9/Y_{20}$ contacts were observed in the MAS NMR spectra, but could not be unambiguously assigned due to spectral overlap (dashed arrows in Figure 23). Thus they were not included in the structure calculation (see section 4.7.2).

The sheet-sheet distance is 10 Å as determined by electron microscopy indicating that the sheets are tightly packed against each other. The diameter of the fibrils is 30 Å which corresponds to a $\beta$-strand of around ten residues in good agreement with the length of $\beta$-strand one ($A_4$-$T_{13}$). TALOS analysis indicates that $\beta$-strand two is longer ($K_{17}$-$K_{32}$), but based on the electron microscopy data, the structural core of the fibrils seems to be shorter, comprising, approximately residues $K_{17}$-$E_{27}$. Combining the TALOS and electron microscopy data, the structural core of the amyloid fibrils comprises at least residues $A_4$-$E_{27}$. This coincides with the occurrence of peak doubling for residues $T_3$ and $S_{28}$. Interestingly, alanine mutation of $T_{29}$ and $W_{30}$ hinder fibril formation, and it may be speculated that these residues play a role in protecting the fibrillar core.

Based on this preliminary structural model, it is difficult to understand the observed decelerating effect of the $Q_{34}$A mutation on the fibrillation kinetics. Further experiments are necessary to investigate whether $Q_{34}$ is really essential for fibril formation.

## 4.7 CA150.WW2 structure determination

### 4.7.1 Quaternary structure constraints employing peptide mixtures

For structure calculation of CA150.WW2 amyloid fibrils the arrangement of the monomeric units in the fibrils need to be determined. From the cross-$\beta$ amyloid diffraction pattern, see

section 4.5.3, it is known that CA150.WW2 amyloid fibrils contain $\beta$-sheets and that the single $\beta$-strands are oriented perpendicular to the fibril axis. According to the model of the monomeric unit of CA150.WW2, the fibrils consist of two $\beta$-strands incorporated in $\beta$-sheets which interact via their side chains. In general, it is possible that the pair of $\beta$-sheets further associates to form thicker protofilaments, as in the case of A$\beta$(1-40) [Petkova 06], or that the long $\beta$-strands are interrupted as observed for HET-s [Lange 08].

Different terminologies for two $\beta$-strands separated by a short loop region are found in the literature. Here, the definitions by Hennetin et al. [Hennetin 06] are used. Two $\beta$-strands that are connected by hydrogen bonds forming an antiparallel $\beta$-sheet are referred to as $\beta$-hairpin and the loop region is a $\beta$-turn, see Figure 24a). Another possibility is, that the $\beta$-strands interact via their side chains. This is called $\beta$-arch and the loop region $\beta$-arc, see Figure 24b).

Several arrangements of the monomeric strand-turn-strand motif into two interacting $\beta$-sheets are conceivable, see Figure 25. The perspective in this figure was chosen such, that the hydrogen bonds are in the plane of the drawing, and the side chains of the residues are perpendicular to this plane. In most structural models of amyloid fibrils which are based on experimental constraints, parallel $\beta$-sheets are found. Often two different $\beta$-sheets are observed, each composed of one of the $\beta$-strands from the monomeric unit, Figure 25a). Thus each $\beta$-sheet contains just a single type of $\beta$-strand (uni-$\beta$-sheet). In a further architecture the two $\beta$-strands of the peptide would for example alternate within one parallel $\beta$-sheet

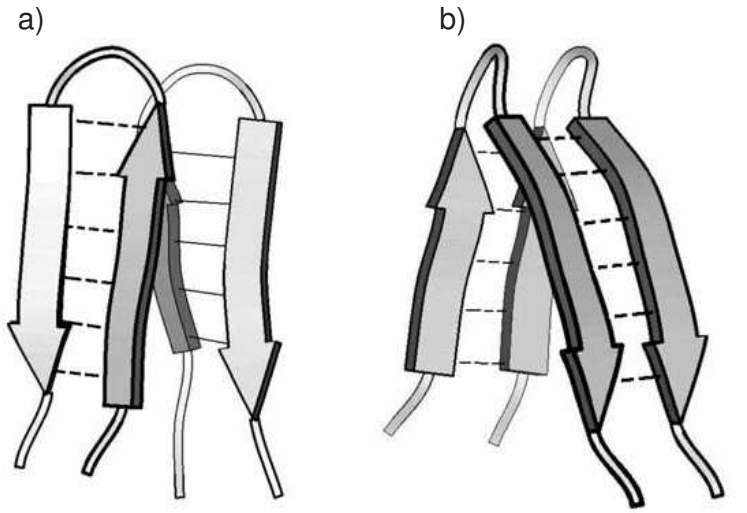



Figure 24: Schematic diagrams of a) two $\beta$-hairpins where the two strand are connected by a $\beta$-turn and b) two $\beta$-archs with a $\beta$-arc loop. The black lines indicate hydrogen bonds. Reproduced from Hennetin et al. [Hennetin 06].

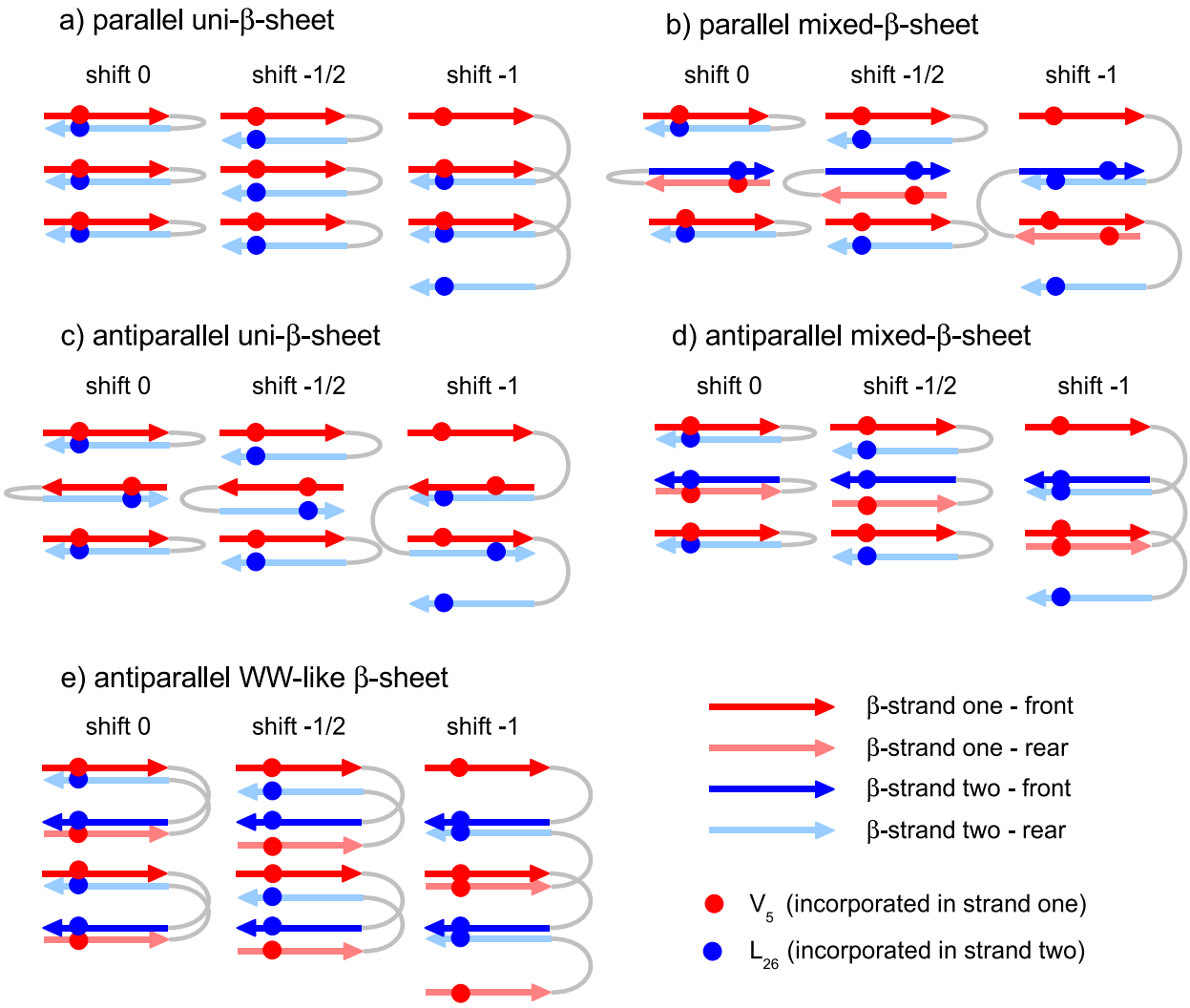



Figure 25: Different architectures for the polymeric arrangement of two $\beta$-strands into two interacting $\beta$-sheets. For the shift $\pm\frac{1}{2}$ and shift $\pm1$ arrangements only the shift in the negative direction is shown.

(mixed-$\beta$-sheet), Figure 25b). In addition to these cases, antiparallel uni and mixed-$\beta$-sheets occur, Figure 25c) and d), respectively. All these architectures contain a strand-turn-strand motif in a $\beta$-arch arrangement. Another possibility is shown in Figure 25e). In this case $\beta$-strands one and two occur within one $\beta$-sheet in an alternating manner similar to the mixed-$\beta$-sheets, but belong to the same polypeptide chain. This resembles the typical $\beta$-hairpin motif that occurs in the native WW-fold and which is frequently observed in globular proteins. Due to the relatively short loop region in fibrillar CA150.WW2 this arrangement would result in antiparallel $\beta$-sheets.

An additional degree of freedom in all discussed architectures is the shift of the two $\beta$-sheets with respect to each other along the main fibril axis. The backbone of a $\beta$-strand

might be directly on top of the corresponding $\beta$-strand from the other $\beta$-sheet (shift 0). It could be between two such $\beta$-strands (shift $\frac{1}{2}$), or shifted by a complete strand-strand distance (shift 1). Larger displacements are unlikely due to the restricted loop length. Thus in total 15 different arrangements should be considered, Figure 25. Depending on the direction of the shift, two forms of the shifted arrangements (shift $\pm\frac{1}{2}$ and shift $\pm1$) can be distinguished.

In the literature, the term STAG can be found to describe the displacement of $\beta$-sheets in amyloid fibrils with respect to each other [Petkova 06, Iwata 06]. This definition takes only odd multiples of half the strand-strand spacing into account. Thus a STAG 1 arrangement corresponds to shift $\frac{1}{2}$ in the nomenclature used in this work. The STAG terminology was not used because in this nomenclature no description for the shift 0 or the shift 1 arrangements exists.

To distinguish the different conceivable architectures experimentally, it would be useful to determine whether the observed long range contacts between the two $\beta$-strands are intra or inter molecular. Inter molecular contacts are suppressed in samples diluted with unlabelled material. Ideally, in such sample every labelled monomeric unit would only have unlabelled neighbours. Then the cross peaks in NMR spectra would arise only from intra molecular contacts. A sample from a peptide mixture close to such an infinite dilution would, however, yield spectra of very poor signal-to-noise ratio. Therefore, in practice often a 1:2-5 ratio of labelled and unlabelled material is used. Inter molecular contacts can be probed using samples from a mixture of two peptides which are labelled in $\beta$-strand one and $\beta$-strand two, respectively.

In case of CA150.WW2 amyloid fibrils, the correlation $V_5/L_{26}$ which shows strong and relatively isolated cross peaks in spectra recorded on 2-CA150.WW2 was chosen to probe intra and inter molecular contacts in the fibrils. Therefore, two specially labelled samples were made. The sample, 2dil-CA150.WW2, was prepared from a 1:4 mixture of 2-$^{13}$C-glycerol labelled and unlabelled peptide. With this labelling scheme cross peaks arising from inter molecular contacts are less probable and thus intra molecular contacts are probed. A second sample was designed to detect inter molecular contacts without intra molecular contributions: vl-CA150.WW2 is a mixture of two differently labelled peptides, 40% *v-CA150.WW2* and 60% *la-CA150.WW2*. *v-CA150.WW2* contains 2,3-$^{13}$C-valine from bacteria grown on 2-$^{13}$C-glycerol. The only $^{13}$C-labelled compounds in *la-CA150.WW2* were

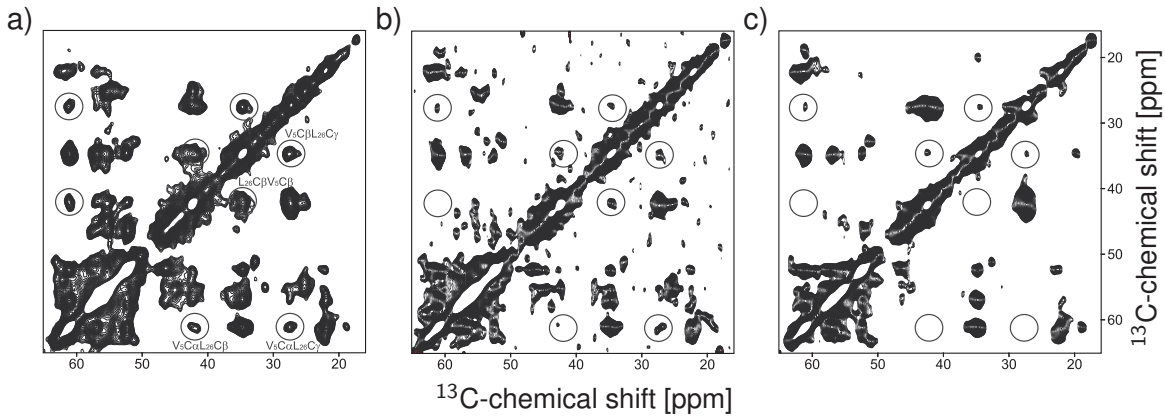



Figure 26: 400 ms PDSD spectra of a) 2-CA150.WW2, b) 2dil-CA150.WW2, and c) vl-CA150.WW2. The $V_5/L_{26}$ cross peak pattern is highlighted.

3,4-$^{13}$C-leucine and 1-$^{13}$C-alanine. To avoid leucine labelling in *v-CA150.WW2*, this amino acid was added in natural abundance to the growth medium. In addition, spectral overlap was reduced by also feeding unlabelled lysine, phenylalanine and tyrosine.

The complete $V_5/L_{26}$ cross peak pattern appears in the 400 ms PDSD spectrum of 2-CA150.WW2, see Figure 26a). The strongest inter residue $V_5/L_{26}$ signals are the $V_5C\beta/L_{26}C\beta$ and $V_5C\beta/L_{26}C\gamma$ cross peaks, whereas the $V_5C\alpha/L_{26}C\beta$ signal is the weakest and occurs only at relatively long mixing times. The same experiment was recorded on 2dil-CA150.WW2, Figure 26b), and vl-CA150.WW2, Figure 26c). Both spectra were plotted such that the intensity for the valine $C\alpha/C\beta$ cross peak is the same as in Figure 26a) to compensate for the differences in the amount of labelled material in the samples. With the exception of the $V_5C\alpha/L_{26}C\beta$ signal, the $V_5/L_{26}$ cross peak pattern is clearly observed in the spectra of 2dil-CA150.WW2 and vl-CA150.WW2. Therefore, it was concluded that the $V_5/L_{26}$ cross peaks contain intra and inter molecular contributions.

The intensity of the $V_5/L_{26}$ cross peaks in the spectrum of 2dil-CA150.WW2 is around 40% compared to the fully labelled material normalised to the $V_5C\alpha/C\beta$ cross peak. However, the intensity of this cross peak in the spectrum of 2dil-CA150.WW2 arises not purely from intra molecular contributions as 20% of the peptide is labelled. In comparsion to fully labelled material, the observed cross peaks in spectra of 2dil-CA150.WW2 account for 100% of the intra and 20% of the inter molecular contributions. Thus less than 40% of the $V_5C\alpha/C\beta$ cross peak intensity in the spectrum of 2-CA150.WW2 should arise from intra molecular contacts. Only inter molecular contacts contribute to the $V_5/L_{26}$ signals in the spectrum of vl-CA150.WW2. Based on the degree of valine labelling the observed cross peak intensity

(around 25% compared to the fully labelled material) corresponds to 60% of the inter molecular contacts. Thus around 42% of the cross peak intensity in the fully labelled material should arise from inter molecular contacts. Clearly, the cross peak intensity observed in both spectra is weaker than expected (42% and 40% should add up to more than 100%, as 20% of the inter molecular contacts contribute to both spectra). This might be due spin diffusion which leads to a loss of magnetisation in the spin-diluted samples. Thus a quantitative analysis of the spectra is ambiguous. Nevertheless, it can be concluded that $V_5$/$L_{26}$ cross peaks arise from intra and inter molecular contacts which contribute similary to the observed intensity. Based on this results, the different architectures shown in Figure 25 are discussed in the following paragraphs, $V_5$ and $L_{26}$ are indicated in red and blue, respectively.

The parallel uni-$\beta$-sheet arrangement, Figure 25a), with shift 0 seems to contain mainly intra molecular $V_5$/$L_{26}$ contacts. However, each monomer has two nearest neighbours which are only 4.7 Å away. Thus, the two inter molecular contacts together could lead to a similar cross peak intensity as the intra molecular correlation. Hence, this arrangement would be in agreement with the data. In the shift $\frac{1}{2}$ arrangement only one of the two neighbouring molecules is expected to contribute to the $V_5$/$L_{26}$ cross peaks. The inter and intra molecular distances would be of very similar size. The shift $\frac{1}{2}$ displacement would therefore also account for the observed experimental data. The distances that would occur in the shift 1 architecture are very similar to the ones for shift 0. However, the shortest and one of the longer distances would be inter molecular. Only one of the longer distances would be intra molecular which is not in agreement with the experimentally observed similar size of the inter and intra molecular contributions.

In the cases of the parallel mixed-$\beta$-sheet, Figure 25b), and the antiparallel uni-$\beta$-sheet, Figure 25c), architectures, the inter molecular $V_5$ and $L_{26}$ distances are $\approx$ 15-25 Å and should not give rise to inter molecular cross peaks. Thus these architectures can be excluded for CA150.WW2 amyloid fibrils. The antiparallel mixed-$\beta$-sheet architecture, Figure 25d), is in a way very similar to the parallel uni-$\beta$-sheet architecture, Figure 25a). However, the inter molecular distances in e.g. the shift 0 arrangement occur between neighbouring strands of the same $\beta$-sheet. This cannot explain why in the spectra of vl-CA150.WW2 the $V_5C\beta$/$L_{26}C\beta$ and $V_5C\beta$/$L_{26}C\gamma$ cross peaks are observed whereas the $V_5C\alpha$/$L_{26}C\beta$ cross peaks are below the detection level. Thus CA150.WW2 amyloid fibrils do not contain antiparallel mixed-$\beta$-sheets. Using the same argument for the intra molecular contacts, monitored in the spectrum of

2dil-CA150.WW2, the antiparallel WW-like $\beta$-sheet architecture is also not in agreement with the observed experimental data.

In summary, CA150.WW2 amyloid fibrils form with a parallel uni-$\beta$-sheet architecture with a shift 0, shift +$\frac{1}{2}$, or shift -$\frac{1}{2}$ displacement of the $\beta$-sheets.

Concentrating on this parallel fibril architecture, it is considered, that the $\beta$-sheets can form with different hydrogen bonding registers, while satisfying the intra and inter molecular $V_5$/$L_{26}$ contacts. The amino group of $V_5$ can then be hydrogen bonded to the CO-sites of $A_2$, $A_4$, or $S_6$ from the neighbouring molecule. Here, the three possible registers are referred to as in-register, +2, or -2 register, Figure 27a). Hydrogen bonding of the $V_5$ amino group to the CO-site of $T_3$, $V_5$, or $E_7$ from the neighbouring monomer would result in different orientations of the side chains in neighbouring monomeric units, and was thus excluded.

To investigate the register, a sample, nta-CA150.WW2, was prepared from a 1:1 mixture of a uniformly $^{15}$N-labelled (*n-CA150.WW2*) peptide and u-$^{13}$C-threonine, 1-$^{13}$C-alanine labelled (*ta-CA150.WW2*) peptide. $^{13}$C-$^{15}$N correlation spectra of such a sample yield only inter molecular cross peaks. Neglecting residue $A_{14}$ which is in a loop region and therefore should not contribute to inter molecular $^{13}$C-$^{15}$N correlations, the $^{13}$C-dimension of spectra recorded on this sample shows isolated CO-resonances for residues $A_2$ and $A_4$. Therefore, for the in-register arrangement $A_4$CO/$V_5$N (175.9/126.7 ppm) and $A_2$CO/$T_3$N (177.0/117.6 ppm) cross

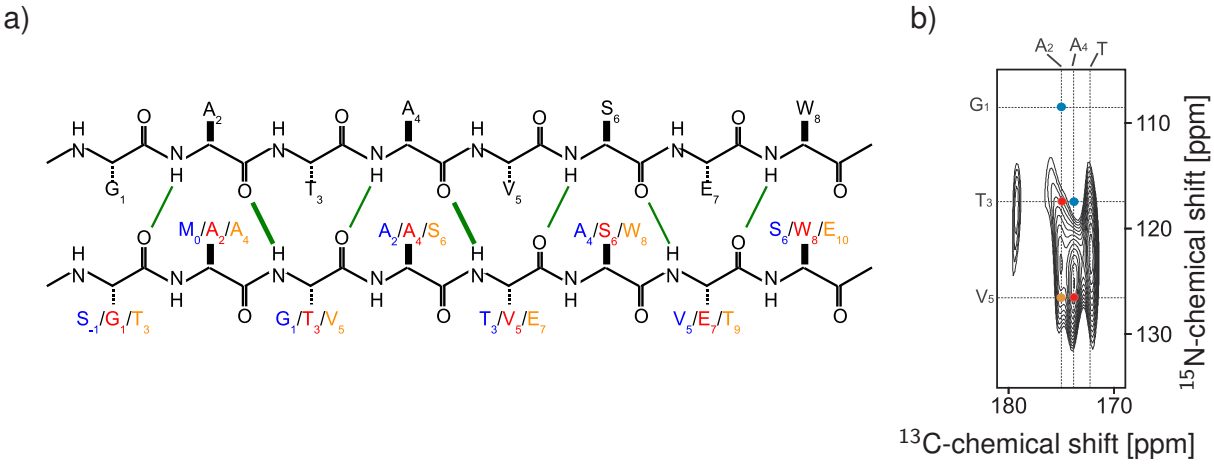



Figure 27: a) -2, in, and +2 registers for the parallel fibril architecture in blue, red, and orange, respectively, hydrogen bonds are drawn in green, contacts monitored by the NHC-experiment on the nta-CA150.WW2 sample are indicated with a bold line, b) NHC spectrum of nta-CA150.WW2, amide nitrogen and carbonyl resonances potentially involved in inter molecular hydrogen bonding are indicated by dashed lines, cross peaks expected for the different registers are marked in the respective colour.

peaks are expected. In case of the -2 register the crosspeaks $A_4CO/T_3N$ (175.9/117.6 ppm) and $A_2CO/G_1N$ (177.0/108.6 ppm) would occur. $A_4CO/E_7N$ (175.9/? ppm) and $A_2CO/V_5N$ (177.0/126.7 ppm) cross peaks would be indicative for the +2 register.

To probe these contacts a NHC spectrum with a short NH (300 $\mu$s) and a long HC (2000 $\mu$s) cross polarisation transfer was recorded, Figure 27b). It shows inter molecular correlations between nitrogen amide and carbonyl sites. The measurement lasted five days, but the signal-to-noise ratio was still poor. Thus the spectrum is badly resolved especially in the indirect dimension. No signal intensity occurs at 175.9/117.6 and 177.0/108.6 ppm which would correspond to the -2 register (labelled in blue). However, on the basis of this spectrum it is difficult to distinguish between the in-register (labelled in red) and the +2-register (labelled in orange) as the signal at the $A_2CO$-resonance is badly defined and the $E_7N$-chemical shift is unknown. Due to the poor signal-to-noise ratio, analysis of the threonine carbonyl/amide nitrogen cross peaks was not possible. The cross peak at 179/118 ppm in F2/F1 was attributed to noise as no signal intensity was observed at 179 ppm in a one-dimensional carbon spectrum of this sample.

In summary, studying three different samples (2dil-CA150.WW2, vl-CA150.WW2, and nta-CA150.WW2) narrowed down the different possibilities for the quaternary structure of CA150.WW2 amyloid fibrils. They consist of two parallel in-register or +2 register uni-$\beta$-sheets, which interact via their side chains. The displacement of the $\beta$-sheets along the fibril axis with respect to each other is either shift 0, shift -$\frac{1}{2}$, or shift +$\frac{1}{2}$.

### 4.7.2 Structure calculation on CA150.WW2 amyloid fibrils

Based on the different possibilities for the quaternary structure of CA150.WW2 amyloid fibrils described in the previous section, the experimental constraints were converted to distance constraints for structure calculation. From the assigned long range cross peaks two distance constraint classes were defined, based on the classification by Castellani et al. [Castellani 02]. Class one (2.5 - 5.5 Å) are cross peaks that occur at PDSD mixing times up to 200 ms. Class (2.5 - 6.5 Å) are cross peaks that appear only at longer mixing times, see Table 8. For the REDOR cross peak $R_{24}/E_7$ it is not known which of the nitrogen atoms of the guanidinium group is involved. Thus the contact was converted into a class two distance constraint between $R_{24}C\zeta$ and $E_7C\delta$. In different calculations, the long range constraints were either treated purely

| Class one (2.5 - 5.5 Å) | Class two (2.5 -6.5 Å) |
|---|---|
| $V_5C\alpha$ - $L_{26}C\gamma$ | $V_5C\alpha$ - $L_{26}C\beta$ |
| $V_5C\beta$ - $L_{26}C\beta$ | $V_5C\alpha$ - $L_{26}C\delta_1$ |
| $V_5C\beta$ - $L_{26}C\gamma$ | $V_5C\alpha$ - $L_{26}C\delta_2$ |
| $V_5C\gamma_1$ - $L_{26}C\alpha$ or $R_{24}C\alpha$ | $V_5C\alpha$ - $R_{24}C\delta$ |
| $V_5C\gamma_2$ - $L_{26}C\alpha$ or $R_{24}C\alpha$ | $T_{13}C\alpha$ - $T_{18}C$ |
| $V_5C\gamma_1$ - $L_{26}C\delta_1$ | $T_{13}C\alpha$ - $T_{18}C\beta$ |
| $V_5C\gamma_1$ - $L_{26}C\delta_2$ | $T_{13}C\alpha$ - $T_{18}C\gamma$ |
| $V_5C\gamma_2$ - $L_{26}C\delta_1$ | $T_{13}C\beta$ - $T_{18}C$ |
| $V_5C\gamma_2$ - $L_{26}C\delta_2$ | $T_{13}C\beta$ - $T_{18}C\beta$ |
| $V_5C\beta$ - $R_{24}C\delta$ | $T_{13}C\gamma$ - $T_{18}C$ |
| $V_5C\gamma_1$ - $R_{24}C\gamma$ | $T_{13}C\gamma$ - $T_{18}C\gamma$ |
| $V_5C\gamma_2$ - $R_{24}C\gamma$ | $T_3C\beta$ - $S_{28}C\alpha$ |
| $T_{13}C\alpha$ - $T_{18}C\alpha$ | $A_2C\alpha$ - $S_{28}C\beta$ |
| $T_{13}C\gamma$ - $T_{18}C\alpha$ | $E_7C\delta$ - $R_{24}C\zeta$ |
| $A_2C\alpha$ - $S_{28}C\alpha$ | |

Table 8: Distance constraints used in the structure calculations

intra molecular or inter and intra molecular at the same time. Thus fibril structures resembling the shift 0 (only intra molecular contacts), the shift $+\frac{1}{2}$ (additionally inter molecular contacts to the following molecule), or the shift $-\frac{1}{2}$ (additionally inter molecular contacts to the preceding molecule) arrangements, were obtained.

The TALOS software provided a total of 50 dihedral angle constraints which were used with generous error margins of 30 degrees. Translational symmetry was imposed to form a fibrillar fragment comprising six polypeptide chains. The single monomers were kept together by sets of inter molecular hydrogen bonds involving residues $A_4$-$T_{13}$ and $K_{17}$-$E_{27}$. Two different registers, corresponding to an in-register and a +2 register arrangement were used in different calculations. However, for the +2 register architecture it was not possible to satisfy all constraints. Thus this register probably does not present the correct CA150.WW2 amyloid fold. So far such a structure has not been described in the literature for any amyloid fibril.

Using the in-register arrangement of the hydrogen bonds, three different structures were calculated corresponding to the shift 0, $+\frac{1}{2}$, and $-\frac{1}{2}$ arrangements. In each case 200 structures were generated. For the shift 0 and the shift $+\frac{1}{2}$ arrangements 45-50% (96 for shift 0 and 91 for shift $+\frac{1}{2}$) structures without constraint violations were obtained. In contrast, this was only the case for 10% of the calculated structures in the shift $-\frac{1}{2}$ arrangement. For each of the three architectures an average structure was calculated and the 15 structures closest to this average were chosen for the structural ensemble.

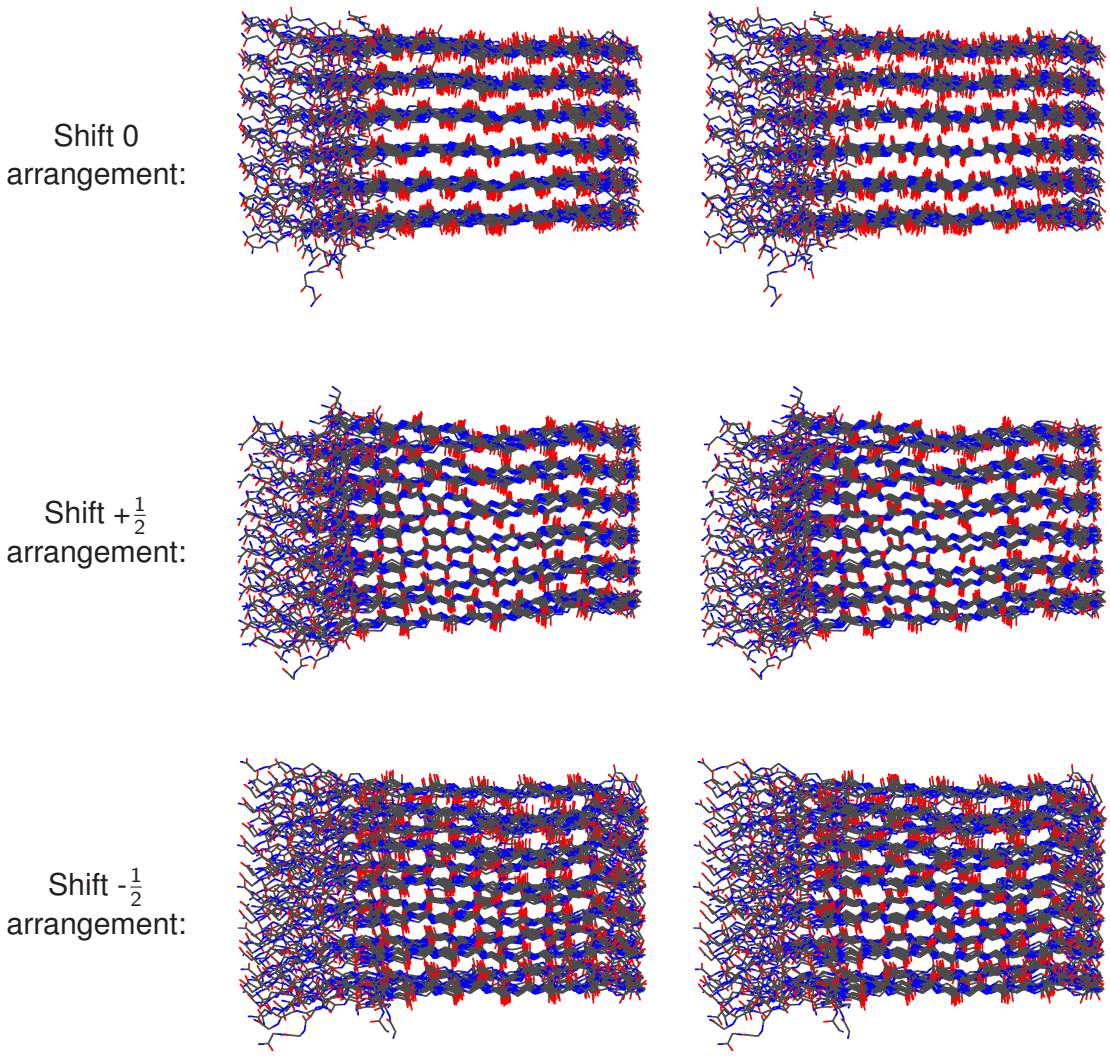



Figure 28: Stereo view of the fibril structure ensembles, residues $G_1$-$K_{32}$ of the 15 structures closest to the average are shown, the $\beta$-sheet that is formed by $\beta$-strand two is on top.

Figure 28 shows a side view of the different ensembles. In the shift 0 arrangement the backbone atoms of both $\beta$-strands of a monomeric unit are in one plane. In contrast, in the shift $+\frac{1}{2}$ and shift $-\frac{1}{2}$ structural ensembles the $\beta$-strands are out-of-plane. The smallest backbone RMSD (**r**oot **m**ean **s**quare **d**eviation) was observed for the shift $+\frac{1}{2}$ arrangement (0.53 Å for residues $A_4$-$T_{13}$ and $K_{17}$-$E_{27}$). The RMSD (0.68 Å) of the shift 0 ensemble is slightly higher, whereas the shift $-\frac{1}{2}$ arrangement has the largest RMSD (0.96 Å).

Interestingly, the variation in the loop region among the different structures within one ensemble is different for the three architectures. In the shift $+\frac{1}{2}$ structures the dihedral angles of $A_{14}$, $D_{15}$, and $G_{16}$ cluster around one or two values. In contrast, for the structures in the shift $-\frac{1}{2}$ and shift 0 ensembles, these residues show a larger distribution of dihedral angles. In the shift $+\frac{1}{2}$ structures the conformation of $A_{14}$ and $G_{16}$ is in or close to the $\beta$-sheet region

of the Ramachandran Plot, whereas $D_{15}$ has dihedral angles close to the $\alpha$-helical and loop regions.

A lower probability for the occurrence of structures with a loop conformation that enables a shift -$\frac{1}{2}$ displacement during structure calculation, might present a reason why only very few structures were obtained for this architecture. It may be speculated that an entropic contribution, plays a role during fibril formation and thus makes a shift -$\frac{1}{2}$ arrangement of the monomeric units in CA150.WW2 fibrils eventually less likely. Similarly, during the structure determination of the K3-peptide (from $\beta_2$-microglobulin) fibrils only a shift +$\frac{1}{2}$ displacement could successfully be calculated [Iwata 06].

In Figure 29 different views along the fibril axis are shown for the shift +$\frac{1}{2}$ arrangement. The ensemble of the six calculated monomeric units, Figure 29a), exhibits a slight twist for each monomer with respect to its neighbour. This twist is an intrinsic property of $\beta$-sheets and was also observed in the other two structural ensemble and other amyloid fibril models [Petkova 06, Correia 06]. The ensemble for a single monomeric unit is shown in Figure 29b) for the case of the shift +$\frac{1}{2}$ arrangement. The side chain packing in the core of the structure is very dense and leaves hardly space for solvent molecules. Different side chain orientations are observed for the aromatic residues, due to the lack of constraints. These residues show heterogeneous line shapes in MAS NMR spectra, leading to the speculation that different side chain orientations might be present in the fibrils. A very similar situation is encountered in the structural ensembles of the other two architectures. A single monomeric unit for the shift +$\frac{1}{2}$ architecture is shown in Figure 29c) illustrating the intra molecular distance constraints used during structure calculation.

The different calculated architectures were analysed to determine whether they can be distinguished by the $V_5$/$L_{26}$ inter residue distances. For each of the three structural ensembles $V_5$/$L_{26}$ inter residue distances were calculated. Here, the result for the $V_5C\beta$/$L_{26}C\beta$ contact is given, see Table 9. In all three structural ensembles, relatively short distances were observed for the intra molecular contacts enforced by the distance constraints during structure calculation. A long and a short contact is observed for the inter molecular distances to the two neighbouring monomeric units in all three architectures. Interestingly, the majority of the shift 0 ensemble structures (12 out of 15), showed a side chain orientation resembling the shift -$\frac{1}{2}$ structural ensemble.

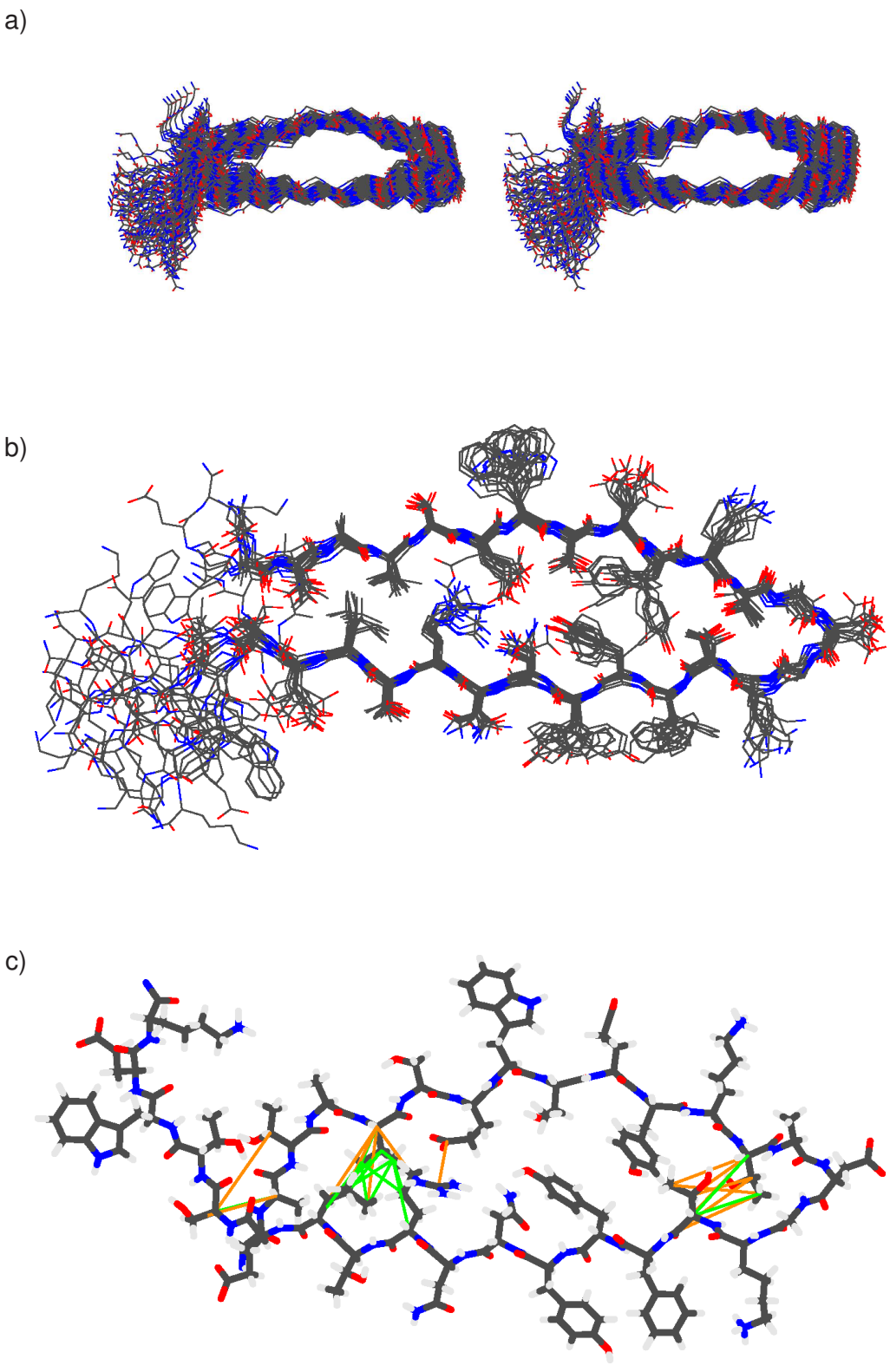



Figure 29: Different views along the fibril axis for the shift $+\frac{1}{2}$ arrangement, residues $G_1$-$K_{32}$ are shown, a) stereo view of the hexamer ensemble, b) ensemble of a monomer, c) a single monomer, the class one and two long range distance constraints are indicated, in green and orange, respectively.

The distances measured in the different arrangements are remarkably similar. Therefore, here no attempt was made to further distinguish between the three arrangements based on the spectra recorded on 2dil-CA150.WW2 and vl-CA150.WW2. In addition, the observed line widths in spectra of CA150.WW2 fibrils indicate some structural heterogeneity and it cannot be excluded that in CA150.WW2 amyloid fibrils slightly different architectures are present in parallel.

| $V_5C\beta/L_{26}C\beta$ contact | shift 0 | shift $+\frac{1}{2}$ | shift $-\frac{1}{2}$ |
|---|---|---|---|
| intra molecular | $5.1 \pm 0.5$ | $5.1 \pm 0.6$ | $5.3 \pm 0.7$ |
| inter molecular to the preceding monomer | $6.1 \pm 2.0$ | $8.5 \pm 0.8$ | $5.1 \pm 1.0$ |
| inter molecular to the following monomer | $8.1 \pm 1.6$ | $5.4 \pm 0.4$ | $9.1 \pm 1.0$ |

Table 9: $V_5C\beta/L_{26}C\beta$ distances in Å in the structural ensembles.

# 5 Discussion

## 5.1 MAS NMR as a tool to study amyloid fibrils – a strategy

Amyloid fibrils are challenging objects for structural investigations because they are insoluble and non-crystalline in nature which hinders studies by X-ray crystallography and solution NMR spectroscopy. MAS NMR spectroscopy is in principle well suited to study these systems. However, assignment of MAS NMR spectra recorded on amyloid fibrils and collection of constraints on the amyloid structure is so far not routine. In this work, CA150.WW2 amyloid fibrils were investigated. Using a novel strategy, carbon and partial amide proton and amide nitrogen assignments were obtained. Subsequently dihedral angle prediction and distance constraint determination was possible. The NMR data agreed well with the results from electron microscopy and alanine scanning. The combined data was used to derive the fold of the monomeric unit in CA150.WW2 fibrils. Special isotopically labelled samples were used to probe the quaternary structure of the amyloid fibrils. The distance constraints derived from the NMR data were used for structure calculation of CA150.WW2 amyloid fibrils. Here, based on the experience gained in this work and recent achievements from other groups, a strategy for structural investigations of amyloid fibrils by MAS NMR is outlined.

To obtain structural information from MAS NMR spectra, a sufficient amount of $^{13}$C and $^{15}$N-signals need to be assigned. Using uniformly $^{13}$C and $^{15}$N-labelled samples this can be done by tracing the protein backbone in $^{13}$C-$^{15}$N correlation spectra. These methods have been developed on crystalline proteins [Pauli 01]. They were successfully applied to the rigid parts in HET-s [Siemer 06b] and $\alpha$-synuclein [Heise 05a, Kloepper 07] fibrils which both have line widths comparable to crystalline proteins. However, the spectra of many amyloid systems show considerably broader lines which lead to strong signal overlap. In the study of amyloid fibrils formed by the 22-residue K3-peptide of $\beta_2$-microtubulin [Iwata 06] this was tolerable due to the small size of the monomeric unit. In other cases though, assignment strategies tracing the protein backbone were not applied. An example are amyloid fibrils formed by the A$\beta$-peptide which exhibit broad lines in MAS NMR spectra. The assignment problem in MAS NMR studies of these systems was overcome by sequence-specific labelling using peptide synthesis [Petkova 02].

The line widths observed in spectra of CA150.WW2 amyloid fibrils also hindered the usage

of assignment strategies employing NCA/NCACX and NCO/NCOCX spectra. Therefore, a novel assignment strategy was used in this work to obtain carbon and partial amide proton and amide nitrogen assignments. Thereby, four differently labelled samples were used: a uniformly $^{13}$C,$^{15}$N-labelled, a 1,3-$^{13}$C-glycerol,$^{15}$N-labelled, a 2-$^{13}$C-glycerol,$^{15}$N-labelled, and a uniformly $^{2}$H,$^{13}$C,$^{15}$N-labelled, $^{1}$H back-exchanged sample. $^{13}$C-$^{13}$C PDSD correlations were recorded on these samples. In spectra with short mixing times, cross peaks occurred from intra residue correlations. The glycerol-labelled samples yielded spectra with a specific labelling pattern for each type of residue which has been described in the literature [Lemaster 82, Castellani 02]. The spectra recorded on the deuterated sample showed improved line widths for most cross peaks (see below). This sample was also used to record spectra with direct carbon excitation which are more sensitive for signals from residues with increased mobility. Indeed, a few additional cross peaks from the C and N-termini were observed in such spectra.

In $^{13}$C-$^{13}$C spectra recorded with PDSD mixing times that exceed 50 ms, C$\alpha$/C$\alpha$ correlations involving neighbouring amino acids are observed. In principle, they can be used for sequential assignment. However, even for small proteins, the C$\alpha$/C$\alpha$ region in PDSD spectra of uniformly labelled samples suffers from spectral crowding. This is markedly reduced in spectra recorded on the glycerol-labelled samples. Thus the assignment possibilities are narrowed down and sequential assignments can be obtained based on the characteristic glycerol-labelling pattern of each residue. In a similar way, other contacts between neighbouring residues can be used, e.g. CO/C$\alpha$ cross peaks in TEDOR spectra. Further sequential correlations can be obtained from $^{13}$C-$^{13}$C spectra with long mixing times (e.g. 400 ms DARR) of a deuterated $^{1}$H back-exchanged sample. They occur mainly from sites near a proton nucleus.

Using the described strategy, a $^{13}$C-assignment of CA150.WW2 fibrils was obtained. The strategy should be suitable to assign other amyloid systems which exhibit line widths that hinder conventional procedures. One might argue that the described strategy is only applicable to systems with monomeric units of a size up to 40-50 residues. However, larger amyloidogenic sequences often yield MAS NMR spectra with a similarly small number of cross peaks due to the presence of highly flexible parts which do not give rise to signals in MAS NMR spectra [Siemer 06a].

The origin of the improved line widths in spectra of the deuterated sample (around 220 Hz and 190 Hz in the case of protonated and deuterated fibrils, respectively) might be differences

in the structural or dynamic properties compared to the protonated samples. Another reason might be insufficient proton-carbon decoupling during the experiments recorded on the protonated samples. This view is supported by a study on nanocrystalline ubiquitin which yields very well resolved MAS NMR spectra indicative for good decoupling. Consequently, the spectral resolution in $^{13}$C-spectra of ubiquitin does not improve upon deuteration [Morcombe 04]. To explain the improvement of the line width in CA150.WW2 amyloid fibrils upon deuteration, it may be speculated that CA150.WW2 amyloid fibrils exhibit mobility on a timescale that interferes with the heteronuclear proton decoupling. This hypothesis could be tested experimentally by comparing spectra recorded at different temperatures, or by applying very high decoupling powers, e.g. using a very small rotor to prevent sample heating. The experience gained in this work indicates, that deuteration might serve as a general approach to improve the resolution in amyloid fibrils with increased line widths.

Deuteration also leads to smaller and fewer homonuclear proton dipolar interactions. This can be exploited in $^{1}$H-$^{13}$C-$^{15}$N correlations. In this work, a three-dimensional HNC-TEDOR experiment was assembled. The analysis of the spectrum yielded partial amide nitrogen and amide proton assignments. The resolution in the proton dimension in this spectrum might be further improved by homonuclear proton decoupling using for example the frequency-switched Lee-Goldberg scheme [Bielecki 89]. Another approach would be, to further reduce the number of protons in the sample. This was recently demonstrated by Chevelkov et al. [Chevelkov 06]. They used microcrystalline SH3 with a very high degree of deuteration (90% of all exchangeable sites and nearly 100% of all other sites) to record well resolved proton detected $^{1}$H-$^{15}$N MAS NMR correlations. However, the sensitivity of such experiments is relatively low due to the small amount of protons in such samples.

Secondary structure in amyloid fibrils can be probed by TALOS dihedral angle prediction. Based on the established assignment, it is also possible to detect and assign long range contacts in MAS NMR spectra. For this purpose, PDSD and DARR spectra with long mixing times were used in this work on CA150.WW2 amyloid fibrils. In the work on HET-s amyloid fibrils additionally proton spin diffusion experiments (CHHC and NHHC) were employed to obtain long range distance constraints [Wasmer 08]. These spectra are less crowded but at the same time the techniques are less sensitive.

Due to the polymeric nature of amyloid fibrils it is necessary to use fibrils grown from mixtures of differently labelled peptides to probe the overall fibril architecture. In this work and many other studies on amyloid fibrils, spectra were recorded on fibrils grown from mixtures of labelled and unlabelled peptides to distinguish intra from inter molecular contacts. In addition, spectra from fibrils grown from a mixture of $^{13}$C and $^{15}$N-labelled peptides can reveal the register of the $\beta$-sheets.

Amyloid structure calculation can be done by imposing translational symmetry along the fibril axis. Due to the high symmetry in the polymeric arrangement often a relatively small number of long range distance constraints is sufficient to determine a relatively well-defined amyloid fibril structure (e.g. 134 total and 103 long range distance constraints in HET-s [Wasmer 08]).

In summary, using the described strategy, a collection of 5-6 differently labelled samples, including glycerol-labelling, deuteration, and spin-dilution, should for many amyloid systems be sufficient to assign the MAS NMR spectra, to determine some long range distance constraints, to probe the quaternary structure, and finally to obtain an overall fold. The procedure is robust and thus applicable to systems that show intermediate line widths. The described assignment strategy might also present a way to tackle non-amyloid systems of intermediate line widths.

## 5.2 CA150.WW2 and related structures – a $\beta$-solenoid as a common amyloid fold?

In this work, based on MAS NMR spectroscopy, alanine scanning, and electron microscopy, the fold of the monomeric unit and the architecture of CA150.WW2 amyloid fibrils was determined. The core of CA150.WW2 amyloid fibrils was found to consist of two parallel in-register $\beta$-sheets, each composed of one of the $\beta$-strands of the monomeric unit (uni-$\beta$-strand). To build these $\beta$-sheets, the monomeric unit folds into a so-called $\beta$-arch consisting of two $\beta$-strands which are separated by a short $\beta$-arc loop region. The two $\beta$-sheets interact with each other via their side chains forming a dense structural core that has elements of the steric zipper arrangement observed in fibril-like crystals of short peptides [Nelson 05, Sawaya 07]. The C-terminus ($Q_{34}$-$K_{37}$) is not involved in the structural core and shows increased mobility.

Similar structures were determined or are suggested as structural models for other amyloid systems. Pairs of uni-$\beta$-strand, in-register, parallel $\beta$-sheets that interact via their side chains were described for fibrils formed by the A$\beta$(1-40) peptide [Petkova 02, Petkova 06], the A$\beta$(1-42) peptide [Lührs 05], amylin [Luca 07], and the K3-peptide of $\beta_2$-microglobulin. For A$\beta$(1-40) and amylin it was further suggested that two pairs of $\beta$-sheets interact with each other to produce a twice as large structural core. More complex structures were observed for longer amyloidogenic sequences such as the prion domain of the HET-s prion protein or $\alpha$-synuclein. HET-s fibrils are composed of three interacting parallel $\beta$-sheets. Each of the $\beta$-sheets consits of two alternating $\beta$-stands that are pseudo repeats of each other [Wasmer 08]. Secondary structure analysis of $\alpha$-synuclein fibrils revealed five $\beta$-strand regions in the monomeric unit [Kloepper 07], however, the fold of the monomeric unit and the fibril architecture are still unknown.

For the different amyloid systems various displacements of the $\beta$-sheets along the main fibril axis relative to each other, in this work referred to as shift, are observed. Here, an overview of these findings together with the results on CA150.WW2 fibrils is given. On CA150.WW2 fibrils the shift was determined to a precision of the strand-strand distance and can either be shift 0, shift -$\frac{1}{2}$, or shift +$\frac{1}{2}$. The structure calculated from MAS NMR data on the K3-peptide of $\beta_2$-microglobulin showed a STAG(+1) arrangement [Iwata 06] which corresponds to the shift +$\frac{1}{2}$ displacement in the nomenclature used in the present work. The STAG(-1) arrangement (equivalent to shift -$\frac{1}{2}$) in K3-peptide fibrils was excluded during the structure calculation process. Similarly, the corresponding shift -$\frac{1}{2}$ arrangement in CA150.WW2 fibril calculations yielded the fewest structures without restraint violations. However, in the study of the K3-peptide fibrils a STAG(0) (shift 0) arrangement was not considered on the basis of spectra from spin diluted samples (labelled and unlabelled 1:2).

Similarly, in the work on A$\beta$(1-40) fibrils by Petkova et al. [Petkova 06] only odd multiples of half the interstrand spacing (shift $\frac{1}{2}$, shift $\frac{3}{2}$) were taken into account based on the assumption that this leads to a favourable side chain packing between two $\beta$-sheets. From spin-dilution experiments a STAG($\pm$2) arrangement was concluded (corresponding to a shift $\pm\frac{3}{2}$ displacement). In contrast, a shift +1 arrangement was determined for A$\beta$(1-42), based on the fibril growth properties of mixtures of several variants and on inhibitor binding experiments [Lührs 05]. To accommodate the relatively short loop region between the two

$\beta$-strands in A$\beta$(1-42) fibrils, the suggested structural model contains slightly tilted $\beta$-strands.

The well resolved structure of HET-s amyloid fibrils [Wasmer 08] shows a displacement of the $\beta$-sheets along the fibrils axis corresponding to a shift in the positive direction. Finally, in X-ray structures of fibril-like crystals of short peptides, the monomeric units from different interacting $\beta$-sheets are displaced by half a strand-strand distance with respect to each other [Nelson 05, Sawaya 07].

Parallel $\beta$-sheets that interact through their side chains are not a unique feature of amyloid fibrils. A well known example for the occurrence of such $\beta$-sheets are the parallel $\beta$-helices or $\beta$-solenoids [Jenkins 01, Kajava 06]. They consist of two to four parallel $\beta$-sheets. The $\beta$-strands of the different $\beta$-sheets are connected by short loops. The majority of the $\beta$-helices are composed of three $\beta$-sheets. They can be right or left-handed, e.g. the phage P22 tailspike protein from *Salmonella typhimurium* [Steinbacher 94] and UDP-*N*-acetylglucosamine acyltransferase from *Escherichia coli* [Raetz 95], respectively. The arrangement of the $\beta$-sheets in these proteins corresponds to a positive or negative shift in case of the left or right-handed helices, respectively. The side chains in one coil of these structures are directed towards each other. An interdigitation as observed in the fibril-like crystals investigated by Eisenberg et al. [Nelson 05, Sawaya 07] does not occur as three $\beta$-sheets build the interface in these examples. However, right-handed parallel $\beta$-helices with an L-shaped cross sections (e.g. pektate lyase C from *Erwinia chrysanthemi* [Yoder 93]) or even double-stranded $\beta$-helices, also referred to as $\beta$-roles (e.g. alkaline protease from *Pseudomonas aeroginosa* [Baumann 93]), have been observed. In these examples an interdigitation for some of the side chains occurs. CA150.WW2 amyloid fibrils can be regarded as a parallel $\beta$-solenoid with two $\beta$-sheets and only one loop. In a similar way, other amyloid fibrils can be described as parallel $\beta$-solenoids and the well resolved HET-s structure [Wasmer 08] greatly resembles a triple-stranded, left-handed, parallel $\beta$-helix.

In most parallel $\beta$-helix structures the formation of amino acid stacks inside the $\beta$-helix core which can be hydrophobic (valine, leucine, isoleucine), aromatic (tyrosine, phenylalanine) or hydrogen-bonded (asparagine, glutamine) are observed. This is also reflected in the primary sequences of these proteins. Such interactions are easily established in parallel in-register amyloid fibrils where each monomeric unit forms a single coil identical to its neighbouring coils. Thus, these stacking interactions might be a driving force for the relatively high structural order

in amyloid fibrils. This is supported by the observation that in the functional HET-s fibrils, where the monomeric unit forms two coils, the single coils are pseudo-repeats of each other.

In summary, CA150.WW2 and other amyloid fibrils described in the literature have many structural properties in common supporting the view that there is a general amyloid fold which resembles the parallel $\beta$-helices. This is in accordance with the observation that many parallel $\beta$-helix proteins as well as most amyloid fibrils are very stable with respect to high temperatures and denaturing agents. So far, it is unknown whether the amyloid fold is related to the left or right-handed parallel $\beta$-helices or whether both forms occur. Studies on the HET-s and the K3-peptide amyloid fibrils revealed a shift in the positive direction, corresponding to the left-handed $\beta$-helix. Intriguingly, the CA150.WW2 structure with the best backbone RMSD also has a positive shift (shift $+\frac{1}{2}$).

## 5.3 Outlook

In this work, CA150.WW2 amyloid fibrils were studied intensively by MAS NMR to determine the fold of their monomeric unit and their fibril architecture. Some uncertainty remained with respect to the relative shift of the $\beta$-sheets along the fibril axis. The calculated fibril architectures shift 0, shift $+\frac{1}{2}$, and shift $-\frac{1}{2}$ are relatively similar. Therefore, it might be possible that the shift is not the same throughout a fibril sample, varying e.g. within a single protofilament or among the different fibrils.

To distinguish the different shifts experimentally, it is necessary to measure inter nuclear distances to a very high precision. MAS NMR experiments involving several spins, however, are very difficult to interpret in a quantitative way. Therefore, it would be desirable to prepare samples which contain just a spin pair – ideally further spin-diluted with unlabelled material. Using such a sample it would be possible to measure dipolar build-up curves from a series of one dimensional spectra. A similar experiment was conducted by Balbach et al. [Balbach 02] to confirm the parallel in-register architecture of A$\beta$(1-40) amyloid fibrils.

Another approach to distinguish the different shifts and to improve the structural quality would be to measure the dihedral angles in the loop region directly. Such an experiment was demonstrated by Rienstra et al. [Rienstra 02] on a uniformly $^{13}$C,$^{15}$N-labelled tripeptide. It is, however, experimentally challenging to do this experiment on CA150.WW2 amyloid fibrils due the line broadening observed for the residues in the loop region.

In this work, experimental determination of the hydrogen bonding register was done using inter molecular $^{13}$C-$^{15}$N contacts in fibrils grown from a mixture of $^{13}$C and $^{15}$N-labelled peptides. However, the obtained signal-to-noise ratio in NHC experiments from this sample was very poor. To improve this, it would be useful to reduce the loss of magnetisation due to proton spin diffusion. This could be done by introducing deuterons in one of the peptides, e.g. a $^{2}$H,$^{15}$N-labelled, $^{1}$H back-exchanged peptide.

Further improvement of the CA150.WW2 amyloid fibril structure is expected from experiments elucidating the role of the residues $T_{29}$ and $W_{30}$. They are important because their replacement by alanine eliminates the ability of CA150.WW2 to form amyloid fibrils. In addition, TALOS predictions indicate that $\beta$-strand two might continue until $K_{32}$. However, no distance constraits could be determined for this part of the monomeric unit due to the heavy signal overlap of the involved residue ($T_{29}W_{30}E_{31}K_{32}$) with other parts of the sequence. Selective labelling using peptide synthesis might be a way to overcome this problem. In such a preparation it would also be useful to label additional residues, e.g. $A_2$ and $S_{28}$, to probe putative long range contacts between $T_{29}$-$K_{32}$ and the structural core.

The determination of amyloid fibril structures has been slow in the past and only few structural information is available on these systems. This work in conjunction with other reports, e.g. on A$\beta$ [Petkova 06], the K3-peptide from $\beta_2$-microglobulin [Iwata 06], and the HET-s prion protein [Wasmer 08], has shown that MAS NMR is a valuable tool to investigate these non-crystalline and non-soluble materials. Many amyloid fibrils show relatively broad lines in MAS NMR spectra and the parallel in-register architecture makes it difficult to distinguish inter from intra molecular contacts. In the presented work, these problems were overcome using specially labelled samples (including glycerol-labelling and sample mixture) and deuteration. Based on the developed techniques it should be possible to determine more and better resolved amyloid structures in the near future in order to develop a general understanding of the amyloid fibril fold.

The role of mature amyloid fibrils in diseases is under debate and special interest lies in the understanding of the structural properties of the oligomeric amyloid precursors which, in many cases, are more toxic to cells than the mature amyloid fibrils [Stefani 07]. To develop preparations of these precursors suitable for MAS NMR studies and to compare their structural properties to mature amyloid fibrils will be of interest in future amyloid work. First progress

has been made by Chimon et al. who studied prefibrillar aggregates of A$\beta$(1-40) by MAS NMR and observed chemical shifts quite similar to those from the mature amyloid fibrils [Chimon 05, Chimon 07]. Techniques applicable to samples with broadened lines and a low signal-to-noise ratio, as developed in the present work, may also be suitable to study prefibrillar aggregates in detail. Using the obtained structural information, it will eventually become possible to investigate the interaction of amyloid fibrils or their precursors with cellular compounds such as membranes or other proteins to further develop the understanding of amyloid diseases.

## 5.4 Publications

The work described here has so far resulted in two publications:

Ferguson et al. [Ferguson 06] describes alanine scanning, electron microscopy, and MAS NMR long range distance constraint determination on CA150.WW2 amyloid fibrils. The results of these studies were combined to yield a structural model of CA150.WW2 amyloid fibrils.

Becker et al. (accepted) [Becker 08] presents the assignment of CA150.WW2 amyloid fibrils and the assignment strategy, a TALOS analysis, and the 3D-HNC-TEDOR experiment.

# 6 Summary

Amyloid fibrils are the major protein components of deposits found in patients suffering from amyloid diseases such as Alzheimer's or Parkinson's disease. The role of the amyloid deposits in these disorders is under debate, ranging from suggestions that they are merely byproducts that occur during the course of the disease to the hypothesis that they are the main toxic species. It has been discovered that *in vitro* oligomeric precursors of the fibrils are often more toxic to cells than mature amyloid fibrils raising the suspicion that mature fibrils could serve as a sink for the more dangerous intermediates.

To improve understanding of the function of amyloid deposits in disease, it is necessary to study their interaction with cellular compounds at a molecular level. Thus, high resolution structural information on these systems is needed. Amyloid fibrils are generally non-crystalline and insoluble which largely hinders studies by X-ray crystallography and solution NMR, both being well established techniques in structural biology. Solid state NMR spectroscopy is in principle well suited to study fibrillar aggregates at atomic resolution. However, so far no routine procedure for structure determination of immobile proteins by solid state NMR spectroscopy has been established.

Nevertheless, much progress has been made during the last decade and the first solid state NMR structure was solved in 2002 using a microcrystalline protein. A range of amyloid systems is under study in several solid state NMR laboratories. Based on NMR and other experimental evidence a number of structural models have been determined on different amyloid systems and recently the first high resolution structure of an amyloid fibril (the HET-s prion) was presented. This system is characterised by remarkably narrow lines in solid state NMR spectra in contrast to many other amyloid samples. A more general approach for tackling amyloid samples with increased line width is lacking. Therefore, the total number of amyloid systems studied at atomic resolution so far is very limited and further information is necessary to develop a general understanding of amyloid structure, formation of amyloid deposits, and amyloid toxicity.

In this work, the structural properties of amyloid fibrils formed by the second WW-domain of CA150, a human transcriptional activator, were studied. The system was chosen for structural investigations because a wealth of information on the fibrillation properties of CA150.WW2 and several of its variants was available from a collaboration with Neil Ferguson and Alan R.

Fersht from the MRC in Cambridge, UK. In preliminary studies, CA150.WW2 Y19F amyloid fibrils showed MAS NMR spectra with considerably narrower lines compared to wild type CA150.WW2 fibrils. Consequently, this variant was chosen for the structural studies. The observed line width are, nevertheless, relatively broad compared to crystalline proteins. Therefore, standard solid state NMR assignment procedures using two or three-dimensional NCA/NCACX and NCO/NCOCX experiments were not suitable for assigning the spectra recorded on CA150.WW2 amyloid fibrils.

A novel assignment strategy was devised. It is based on the amino acid labelling pattern introduced by using 1,3-$^{13}$C-glycerol or 2-$^{13}$C-glycerol as the only carbon source during recombinant protein expression. The observed labelling pattern facilitates the assignment of the signal sets to a certain type of residue. In addition, sequential cross peaks, such as C$\alpha$/C$\alpha$ signals, are only observed when both sites are labelled which is only the case for certain residue combinations in the glycerol-labelled samples. This reduces the spectral overlap as well as the number of assignment possibilities. Therefore, two-dimensional $^{13}$C-$^{13}$C correlations can be used for the assignment. Additionally, a u-$^{2}$H,$^{13}$C,$^{15}$N-labelled sample was investigated which carried protons at all exchangeable sites. Interestingly, the deuteration led to spectra with significantly reduced line width. Using spectra of this and the two glycerol-labelled samples, the $^{13}$C-assignment of CA150.WW2 amyloid fibrils was determined. For the assignment of proton and amide nitrogen signals a three-dimensional HNC-TEDOR experiment was designed. It is expected that the described strategy might be useful in assigning other amyloid or non-amyloid systems exhibiting broad lines.

From $^{13}$C-$^{13}$C and $^{13}$C-$^{15}$N experiments recorded with long mixing times, 29 distance constaints were extracted. Additonal structural constraints were obtained by TALOS dihedral angle prediction, electron microscopy, and alanine scanning. The data was combined to derive a structural model of the monomeric unit in CA150.WW2 amyloid fibrils. It consists of two long $\beta$-strands connected by a short loop. The two $\beta$-strands are incorporated into two different $\beta$-sheets which interact with each other via their side chains. Direct excitation solid state NMR spectra showed that the C-terminal residues exhibit increased mobility.

The overall fibril architecture was determined using amyloid fibrils grown from mixtures of specially labelled peptides. This enabled analysis of the general arrangement of the two $\beta$-strands in the amyloid fibrils. The shift along the main fibril axis of the two $\beta$-sheets with respect to each other also was investigated. The shift nomenclature used here refers to the

strand-strand distance. CA150.WW2 fibrils are formed by parallel, in-register, uni-stranded $\beta$-sheets (i.e. only one strand of the monomeric unit is incorporated into each $\beta$-sheet) with a shift between $+\frac{1}{2}$ and $-\frac{1}{2}$. Structures were calculated for the shift $+\frac{1}{2}$, shift 0, and shift $-\frac{1}{2}$ architectures. Considering only the defined regions, the lowest backbone RMSD was observed for the shift $+\frac{1}{2}$ architecture but the RMSD for the other two ensembles was also $< 1.0$. However, for the shift $-\frac{1}{2}$ arrangement, only 10% of the calculated structures were free of restraint violations. The interface between the two $\beta$-sheets in the three structures is very similar. Therefore, it was not possible to distinguish which of these represents best the fibril architecture. In all cases, the interface contains interdigitating side chains comparable to the steric zipper motif observed by X-ray crystallography on fibril-like crystals of short peptides.

The $\beta$-arch motif of the monomeric unit of CA150.WW2 amyloid fibrils is distinct from the $\beta$-hairpin observed in antiparallel $\beta$-sheets of globular proteins. However, it occurs in a number of other amyloid systems, e.g. in amyloid fibrils formed by the A$\beta$ peptide or the K3-peptide from $\beta_2$-microglobulin. A similar arrangement was found in the high resolution solid state NMR structure of the HET-s prion fibril. There two $\beta$-arch motifs occur and alternate to form the protofilament whereby one of the $\beta$-sheets is interupted by a short bend region. The determined structures of CA150.WW2, A$\beta$, K3-peptide, and HET-s amyloid fibrils all contain two or three $\beta$-sheets that interact with each other in a manner similar to the two and three-stranded parallel $\beta$-helices. Thus, this work supports the assumption that there is a general amyloid folding motif which resembles the $\beta$-helical (more generally $\beta$-solenoid) structures.

# 7 Zusammenfassung

Amyloidfibrillen sind die Hauptbestandteile der Ablagerungen, die im Verlauf von Amyloiderkrankungen wie z.B. der Alzheimer- oder der Parkinson-Krankheit entstehen. Die Rolle, die diese Ablagerungen im jeweiligen Krankheitsverlauf spielen, wird kontrovers diskutiert. Die Vorschläge reichen dabei von der Annahme, dass es sich bei den Ablagerungen um Nebenprodukte anderer schädlicher Prozesse handelt zur Hypothese, dass sie ursächlich für das Auftreten von Symptomen wie dem Absterben von Zellen sind. Inzwischen weiß man, dass die oligomeren Vorstufen der Fibrillen Zellen *in vitro* meist stärker schädigen als die voll ausgebildeten Amyloidfibrillen, so dass ihr Auftreten einen Weg darstellen könnte, die toxischen Vorstufen unschädlich zu machen.

Um die Bedeutung der Amyloidablagerungen besser zu verstehen, ist es notwendig ihre Wechselwirkung mit anderen Zellbestandteilen auf molekularer Ebene zu untersuchen. Dies erfordert die Kenntnis der dreidimensionalen Fibrillenstruktur in hoher Auflösung. In der Regel sind die Fibrillen jedoch unlöslich und nicht kristallin, so dass die etablierten strukturbiologischen Methoden der Röntgenkristallographie und der Lösungs-NMR-Spektroskopie nur von sehr begrenztem Nutzen sind. Die Festkörper-NMR-Spektroskopie ist dagegen prinzipiell sehr gut geeignet, um Strukturdaten mit atomarer Auflösung an fibrillären Aggregaten zu gewinnen. Bisher wurde aber noch keine allgemeingültige Methode zur Strukturbestimmung von immobilisierten Proteinen mit der Festkörper-NMR-Spektroskopie etabliert.

Auf diesem Gebiet wurden im letzten Jahrzehnt jedoch große Fortschritte erzielt und 2002 konnte die erste Festkörper-NMR-Struktur eines kristallinen Proteins gelöst werden. In einer Reihe von Festkörper-NMR-Laboratorien werden zur Zeit Amyloidsysteme untersucht. Basierend auf den Ergebnissen dieser Studien wurden, oft in Kombination mit anderen Methoden, einige Amyloidstrukturmodelle bestimmt und kürzlich wurde die erste hochaufgelöste Struktur (von Fibrillen der HET-s Prionen) vorgestellt. Im Gegensatz zu den meisten anderen Amyloidpräparationen zeichnen sich HET-s Fibrillen durch besonders schmale Linien in Festkörper-NMR-Spektren aus. Zur Untersuchung von Amyloidfibrillen mit verbreiterten Linien gibt es bisher jedoch keinen allgemeingültigen Ansatz, so dass nur von einer geringen Anzahl von Amyloidsystemen hochaufgelöste Strukturinformationen vorliegen. Für das allgemeine Verständnis von Amyloidstrukturen, der Ausbildung von Amyloidablagerungen und ihrer Zelltoxizität ist es deshalb notwendig, Strukturuntersuchungen an einer größeren Anzahl von Amyloidsystemen durchzuführen.

In dieser Arbeit wurden die strukturellen Eigenschaften von Amyloidfibrillen, die von der zweiten WW-Domäne des humanen Transkriptionsaktivators CA150 gebildet werden, untersucht. Dieses System wurde für die Strukturuntersuchungen ausgewählt, da aufgrund einer Kooperation mit Neil Ferguson und Alan R. Fersht vom MRC in Cambridge in Großbritannien im großen Umfang Daten zu den Fibrillierungseigenschaften von CA150.WW2 und vielen Mutanten zur Verfügung standen. Aus Vorstudien war bekannt, dass Fibrillen der Y19F-Mutante schmaleren Linien in NMR-Spektren zeigen als Fibrillen des Wildtyps. Daher wurde die Mutante für die strukturellen Untersuchungen ausgewählt. Im Vergleich zu mikrokristallinen Proteinpräparationen wiesen jedoch auch die Festkörper-NMR-Spektren der Mutante relativ breite Linien auf, so dass etablierte Methoden zur Zuordnung, die auf zwei- oder dreidimensionalen NCA/NCACX- und NCO/NCOCX-Spektren basieren, nicht erfolgreich angewandt werden konnten.

Daher wurde eine neuartige Zuordnungsstrategie entwickelt. Diese basiert auf dem Markierungsmuster, das bei der rekombinanten Proteinsynthese entsteht, wenn 1,3-$^{13}$C oder 2-$^{13}$C-Glycerol als einzige Kohlenstoffquelle zur Verfügung stehen. Diese Markierungsmuster erleichtern die Zuordnung einzelner Signalsätze zu einem bestimmten Aminosäurerest. Außerdem werden sequenzielle Kreuzsignale, wie z.B. C$\alpha$/C$\alpha$-Kontakte, nur beobachtet, wenn beide Aminosäuren markiert sind. Für die Glycerol-markierten Proben trifft dies nur für bestimmte sequenzielle Paare zu, so dass sich die Signalüberlagerung reduziert, und $^{13}$C-$^{13}$C-Korrelationen für die Zuordnung genutzt werden können. Zusätzlich wurde eine $^{2}$H,$^{13}$C,$^{15}$N-markierte Probe untersucht, bei der die austauschbaren Gruppen protoniert vorlagen. Interessanterweise führte die Deuterierung zu Spektren mit schmaleren Linien. Durch die Auswertung von Spektren der deuterierten sowie der beiden Glycerol-markierten Proben konnte die $^{13}$C-Zuordnung ermittelt werden. Für die Zuordnung der Protonen- und Amidstickstoffsignale wurde ein dreidimensionales HNC-TEDOR-Experiment konzipiert. Die beschriebene Strategie sollte sich auch für die Zuordnung weiterer Amyloidfibrillen sowie anderer Systeme mit verbreiterten Linien eignet.

Die sequentielle Zuordnung der NMR-Signale der CA150.WW2 Amyloidfibrillen ermöglichte die Bestimmung von 29 Abstandsbedingungen aus $^{13}$C-$^{13}$C- und $^{13}$C-$^{15}$N-Spektren mit langer Mischzeit. Zusätzliche Strukturdaten wurden mittels TALOS-Diederwinkelvorhersage, Elektronenmikroskopie und Alaninmutagenese bestimmt. Ausgehend von diesen Daten wurde ein Strukturmodell der monomeren Einheit in CA150.WW2 Amyloidfibrillen erstellt. Es besteht aus zwei langen $\beta$-Strängen, die über eine kurze Schleife verbunden sind. Die einzelnen $\beta$-Stränge sind Teil von zwei unterschiedlichen $\beta$-Faltblättern, die über ihre Seitenketten

miteinander in Kontakt stehen. NMR Spektren mit direkter Anregung zeigten, dass das C-terminale Ende der monomeren Einheit eine erhöhte Mobilität aufweist.

Die Gesamtarchitektur der Amyloidfibrillen wurde mithilfe von Proben aus Mischungen unterschiedlich markierter Peptide bestimmt. Dabei wurde die Anordnung der einzelnen $\beta$-Stränge und die gegenseitige Verschiebung der beiden $\beta$-Faltblätter entlang der Fibrillenhauptachse untersucht. Die hier verwendete Nomenklatur für diese Verschiebung (Shift) bezieht sich auf den Abstand zweier $\beta$-Stränge. CA150.WW2-Fibrillen bestehen aus parallelen, $\beta$-Faltblättern mit Wasserstoffbrücken zwischen immer denselben (uni-stranded) $\beta$-Strängen in äquivalenten Positionen (in-register), die gegeneinander um bis zu einen halben Interstrangabstand in die eine oder die andere Richtung verschoben sind, d.h. zwischen Shift $+\frac{1}{2}$ und Shift $-\frac{1}{2}$. Für die Shift $+\frac{1}{2}$-, Shift 0- und den Shift $-\frac{1}{2}$-Anordnungen wurden Strukturrechnungen durchgeführt. Berücksichtigt man nur die strukturell gut definierten Regionen, weist die Shift $+\frac{1}{2}$-Anordnung den niedrigsten RMSD auf. Auch der Wert für die beiden anderen Architekturen lag bei weniger als 1.0. Für die Shift $-\frac{1}{2}$-Anordnung zeigten allerdings nur 10% der berechneten Strukturen keine Verletzungen der Abstandsbedingungen. Die Interaktionsfläche zwischen den beiden $\beta$-Faltblättern war in allen drei Strukturen sehr ähnlich, so dass nicht unterschieden werden konnte, welche dieser Möglichkeiten der wirklichen Amyloidfibrillenstruktur am nächsten kommt. In allen Strukturen erfolgte die Wechselwirkung zwischen den $\beta$-Faltblättern durch eine Verzahnung der Seitenketten, welche dem sogenannten sterischen Reißverschluss, der in Röntgenstrukturanalysen an fibrillenartigen Kristallen von kleinen Peptiden gefunden wurde, ähnelt.

Das $\beta$-Bogen-Motiv der monomeren Einheit in CA150.WW2 Amyloidfibrillen unterscheidet sich deutlich von der $\beta$-Haarnadel, die in antiparallelen $\beta$-Faltblättern globulärer Proteine gefunden wird. Es kommt jedoch häufig in Amyloidfibrillen, etwa den Fibrillen aus dem A$\beta$-Peptid oder dem K3-Peptid aus $\beta_2$-Microglobulin, vor. Eine ähnliche Anordnung zeigt auch die hochaufgelöste Festkörper-NMR-Struktur der HET-s Prionenfibrille. Sie enthält zwei $\beta$-Bogen-Motive, die abwechselnd aufeinander folgen und das Protofilament bilden, wobei eines der $\beta$-Faltblätter eine Biegung aufweist. Die Strukturen von CA150.WW2, A$\beta$, K3-Peptid und HET-s Amyloidfibrillen enthalten jeweils zwei oder drei $\beta$-Faltblätter, die miteinander auf eine Art und Weise in Wechselwirkung stehen, die den zwei- bzw. dreisträngigen parallelen $\beta$-Helizes ähnelt. Die Ergebnisse dieser Arbeit unterstützen daher die Annahme, dass es ein allgemeines Faltungsmotiv für Amyloidfibrillen gibt, welches als $\beta$-helikal (oder allgemeiner als $\beta$-solenoid) beschrieben werden kann.

# References


[Antzutkin 00] O. N. Antzutkin, J. J. Balbach, R. D. Leapman, N.W. Rizzo, J.Reed & R. Tycko. *Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of $\beta$-sheets in Alzheimer's $\beta$-amyloid fibrils*. PNAS, vol. 97, pages 13045–13050, 2000.

[Antzutkin 02] O. N. Antzutkin, R. D. Leapman, J. J. Balbach & R. Tycko. *Supramolecular structural constraints on Alzheimer's beta-amyloid fibrils from electron microscopy and solid-state nuclear magnetic resonance*. Biochem., vol. 41, no. 51, pages 15436–15450, 2002.

[Arango 06] M. Arango, S. Holbert, D. Zala, E. Brouillet, J. Pearson, E. Régulier, A. K. Thakur, P. Aebischer, R. Wetzel, N. Déglon & C. Néri. *CA150 Expression Delays Striatal Cell Death in Overespression and Knock.In Conditions for Mutant Huntingtin Neurotoxicity*. The Journal of Neuroscience, vol. 26, pages 4649–4659, 2006.

[Balbach 02] J. J. Balbach, A. T. Petkova, N. A. Oyler, O. N. Antzutkin, D. J. Gordon, S. C. Meredith & R. Tycko. *Supramolecular Structure in Full-Length Alzheimer's $\beta$-Amyloid Fibrils: Evidence for a Parallel $\beta$-Sheet Organization from Solid-State Nuclear Magnetic Resonance*. Biophysical Journal, vol. 83, pages 1205–1216, 2002.

[Baldus 96] M. Baldus, D.G. Geurts, S. Hediger & B. H. Meier. *Efficient 15N-13C Polarization Transfer by Adiabatic-Passage Hartmann-Hahn Cross Polarization*. J. Magn. Reson. Series A, vol. 118, pages 140–144, 1996.

[Baldus 98] M. Baldus, A. t. Petkova, J. Herzfeld & R. G. Griffin. *Cross polarization in the tited frame: assignment and spectral simplification in heteronuclear spin systems*. Molecular Physics, vol. 95, pages 1197–1207, 1998.

[Baldus 02] M. Baldus. *Correlation experiments for assignment and structure elucidation of immobilized polypeptides under magic angle spinning*. Prog. Nucl. Magn. Res. Spec., vol. 41, pages 1–47, 2002.

[Baumann 93] U. Baumann, S. Wu, K. M. Flaherty & D. B. Mckay. *3-dimensional structure of the alkaline protease of pseudomonas-aeruginosa - a 2-domain protein with a calcium-binding parallel-beta roll motif*. Embo J., vol. 12, no. 9, pages 3357–3364, 1993.

[Baxa 07] U. Baxa, R. B. Wickner, A. C. Steven, D. E. Anderson, L. N. Marekov, W. M. Yau & R. Tycko. *Characterization of beta-sheet structure in Ure2p(1-89) yeast prion fibrils by solid-state nuclear magnetic resonance*. Biochem., vol. 46, no. 45, pages 13149–13162, 2007.

[Becker 08] J. Becker, N. Ferguson, J. Flinders, B.-J. van Rossum, A. R. Fersht & H. Oschkinat. *A sequential assignment procedure for proteins that have intermediate line widths in MAS NMR spectra: Amyloid fibrils of human CA150.WW2*. ChemBioChem, 2008.

[Bellotti 07] V. Bellotti, M. Nuvolone, S. Giorgetti, L. Obici, G. Palladini, P. Russo, F. Lavatelli, V. Perfetti & G. Merlini. *The workings of the amyloid diseases*. Annals Medicine, vol. 39, no. 3, pages 200–207, 2007.

[Bennett 92] A. E. Bennett, J. H. Ok, R. G. Griffin & S. Vega. *Chemical-shift correlation spectroscopy in rotating solids - radio frequency-driven dipolar recoupling and longitudinal exchange*. J. Chem. Phys., vol. 96, no. 11, pages 8624–8627, 1992.

[Bennett 95] A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi & R. G. Griffin. *Heteronuclear decoupling in rotating solids*. J. Chem. Phys., vol. 103, no. 16, pages 6951–6958, 1995.

[Bielecki 89] A. Bielecki, A. C. Kolbert & M. H. Levitt. *Frequency-Switched Pulse Sequences: Homonuclear Decoupling and Dilute Spin NMR in Solids*. Chem. Phys. Lett., vol. 155, pages 341–346, 1989.

[Braun 87] W. Braun. *Distance geometry and related methods for protein-structure determination from nmr data*. Quarterly Rev. Biophys., vol. 19, no. 3-4, pages 115–157, 1987.

[Burdick 92] D. Burdick, B. Soreghan, M. Kwon, J. Kosmoski, M. Knauer, A. Henschen, J. Yates, C. Cotman & C. Glabe. *Assembly and aggregation properties of synthetic alzheimers a4/beta amyloid peptide analogs*. J. Biological Chem., vol. 267, no. 1, pages 546–554, 1992.

[Böckmann 03] A. Böckmann, A. Lange, A. Galinier, S. Luca, N. Giraud, M. Juy, H. Heise, R. Montserret, F. Penin & M. Baldus. *Solid state NMR sequential resonance assignments and conformational analysis of the 2 x 10.4 kDa dimeric form of the Bacillus subtilis protein Crh*. Journal of Biomolecular NMR, vol. 27, pages 323–339, 2003.

[Calzolai 05] L. Calzolai, D. A. Lysek, D. R. Perez, P. Guntert & K. Wuthrich. *Prion protein NMR structures of chickens, turtles, and frogs*. Proc. National Acad. Sciences United States Am., vol. 102, no. 3, pages 651–655, 2005.

[Castellani 02] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein & H. Oschkinat. *Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy*. Nature, vol. 420, no. 6911, pages 98–102, 2002.

[Chan 96] D. C. Chan, M. T. Bedford & P. Leder. *Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains*. EMBO J., vol. 15, pages 1045–1054, 1996.

[Chan 05] J. C. C. Chan, N. A. Oyler, W.-M. Yau & R. Tycko. *Parallel $\beta$-Sheets and Polar Zippers in Amyloid Fibrils Formed by Residues 10-39 of the Yeast Prion Protein Ure2p*. Biochemistry, vol. 44, pages 10669–10680, 2005.

[Chapman 02] M. R. Chapman, L. S. Robinson, J. S. Pinkner, R. Roth, J. Heuser, M. Hammar, S. Normark & S. J. Hultgren. *Role of Escherichia Coli Curli Operons in Directing Amyloid Fibre Formation*. Science, vol. 295, pages 851–855, 2002.

[Chevelkov 06] V. Chevelkov, K. Rehbein, A. Diehl & B. Reif. *Ultrahigh resolution in proton solid-state NMR spectroscopy at high levels of deuteration*. Angewandte Chemie – international Edition, vol. 45, no. 23, pages 3878–3881, 2006.

[Chimon 05] S. Chimon & Y. Ishii. *Capturing intermediate structures of Alzheimer's beta-amyloid, A beta(1-40), by solid-state NMR spectroscopy*. J. Am. Chem. Soc., vol. 127, no. 39, pages 13472–13473, 2005.

[Chimon 07] S. Chimon, M. A. Shaibat, C. R. Jones, D. C. Calero, B. Aizezi & Y. Ishii. *Evidence of fibril-like beta-sheet structures in a neurotoxic amyloid intermediate of Alzheimer's beta-amyloid*. Nature Struct. Mol. Biol., vol. 14, no. 12, pages 1157–1164, 2007.

[Cookson 05] M. R. Cookson. *The biochemistry of Parkinson's disease*. Ann. Rev. Biochem., vol. 74, pages 29–52, 2005.

[Cornilescu 99] G. Cornilescu, F. Delaglio & A. Bax. *Protein backbone angle restraints from searching a database for chemical shift ans sequence homology*. Journal of Biomolecular NMR, vol. 13, pages 289–302, 1999.

[Correia 06] B. E. Correia, N. Loureiro-Ferreira, J. R. Rodrigues & R. M. M. Brito. *A structural model of an amyloid protofilament of Transthyretin*. Protein Science, vol. 15, pages 28–32, 2006.

[de la Paz 04] M. Lopez de la Paz & L. Serrano. *Sequence determinates of amyloid fibril formation*. PNAS, vol. 101, pages 87–92, 2004.

[Der-Sarkissian 03] A. Der-Sarkissian, C. C. Jao, J. Chen & R. Langen. *Structural organization of alpha-synuclein fibrils studied by site-directed spin labeling*. J. Biological Chem., vol. 278, no. 39, pages 37530–37535, 2003.

[Detken 02] A. Detken, E. H. Hardy, M. Ernst & B. H. Meier. *Simple and efficient decoupling in magic-angle spinning solid-state NMR: the XiX scheme*. Chem. Phys. Lett., vol. 356, no. 3-4, pages 298–304, 2002.

[DiFiglia 97] M. DiFiglia, E. Sapp, K. O. Chase, S. W. Davies, G. P. Bates, J. P. Vonsattel & N. Aronin. *Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain*. Science, vol. 277, no. 5334, pages 1990–1993, 1997.

[Dobson 03] C. M. Dobson. *Protein folding and misfolding*. Nature, vol. 426, pages 884–890, 2003.

[Dobson 05] C. M. Dobson. *Prying into prions*. Nature, vol. 435, pages 747–749, 2005.

[Duer 04] M. J. Duer. *Solid-State NMR spectroscopy*. Blackwell Publishing, 2004.

[Esteras-Chopo 05] A. Esteras-Chopo, L. Serrano & M. L. de la Paz. *The amyloid stretch hypothesis: Recruiting proteins toward the dark side*. Proc. National Acad. Sciences United States Am., vol. 102, no. 46, pages 16672–16677, 2005.

[Etzkorn 04] M. Etzkorn, A. Bockmann, A. Lange & M. Baldus. *Probing molecular interfaces using 2D magic-angle-spinning NMR on protein mixtures with different uniform labeling*. J. Am. Chem. Soc., vol. 126, no. 45, pages 14746–14751, 2004.

[Etzkorn 07] M. Etzkorn, S. Martell, O. C. Andronesi, K. Seidel, M. Engelhard & M. Baldus. *Secondary Structure, Dynamics, and Topology of a Seven-Helix Receptor in Native Membranes, Studies by Solid-State NMR Spectroscopy*. Angew. Chem. Int. Ed., vol. 46, pages 459–462, 2007.

[Faber 98] P. W. Faber, G. T. Barnes, J. Srinidhi, J. Chen, J. F. Gusella & M. E. MacDonalds. *Huntingtin interacts with a family of WW domain proteins*. Human Molecular Genetics, vol. 7, pages 1463–1474, 1998.

[Fandrich 07] M. Fandrich. *On the structural definition of amyloid fibrils and other polypeptide aggregates*. Cellular Mol. Life Sciences, vol. 64, no. 16, pages 2066–2078, 2007.

[Ferguson 01] N. Ferguson, C. M. Johnson, M. Macias, H. Oschkinat & A. Fersht. *Ultrafast folding of WW domains without structured aromatic clusters in the denatured state*. Proc. National Acad. Sciences United States Am., vol. 98, no. 23, pages 13002–13007, 2001.

[Ferguson 03] N. Ferguson, J. Berriman, M. Petrovich, T. D. Sharpe, J. T. Finch & A. R. Fersht. *Rapid amyloid fiber formation from the fast-folding WW domain FBP28*. PNAS, vol. 100, pages 9814–9819, 2003.

[Ferguson 06] N. Ferguson, J. Becker, H. Tidow, S. Tremmel, T. D. Sharpe, G. Krause, J. Flinders, M. Petrovich, J. Berriman, H. Oschkinat & A. R. Fersht. *General structural motifs of amyloid protofilaments*. Proc. National Acad. Sciences United States Am., vol. 103, no. 44, pages 16248–16253, 2006.

[Fowler 06] D. M. Fowler, A. V. Koulov, C. Alory-Jost, M. S. Marks, W. E. Balch & J. W. Kelly. *Functional amyloid formation within mammalian tissue*. Plos Biol., vol. 4, no. 1, pages 100–107, 2006.

[Franks 07] W. T. Franks, K. D. Kloepper, B. J. Wylie & C. M. Rienstra. *Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins*. J. Biomolecular Nmr, vol. 39, no. 2, pages 107–131, 2007.

[Fung 00] B. M. Fung, A. K. Khitrin & K. Ermolaev. *An Improved Broadband Decoupling Sequence for Liquid Crystals and Solids*. J. Magn. Resn., vol. 142, pages 97–101, 2000.

[Fändrich 02] M. Fändrich & C. M. Dobson. *The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation*. EMBO J., vol. 21, pages 5682–5690, 2002.

[Gazit 05] E. Gazit. *Mechanisms of amyloid fibril self-assembly and inhibition (Model short peptides as e key research tool*. FEBS Journal, vol. 272, pages 5971–5978, 2005.

[Gossert 05] A. D. Gossert, S. Bonjour, D. A. Lysek, F. Fiorito & K. Wuthrich. *Prion protein NMR structures of elk and of mouse/elk hybrids*. Proc. National Acad. Sciences United States Am., vol. 102, no. 3, pages 646–650, 2005.

[Gullion 89] T. Gullion & J. Schaefer. *Rotational-echo double-resonance nmr*. J. Magnetic Resonance, vol. 81, no. 1, pages 196–200, 1989.

[Hansen 00] P. E. Hansen. *Isotope effects on chemical shifts of proteins and peptides*. Magnetic Resonance In Chem., vol. 38, no. 1, pages 1–10, 2000.

[Harris 06] D. A. Harris & H. L. True. *New insights into prion structure and toxicity*. Neuron, vol. 50, no. 3, pages 353–357, 2006.

[Hartmann 62] S. R. Hartmann & E. L. Hahn. *Nuclear double resonance in rotating frame*. Phys. Rev., vol. 128, no. 5, page 2042, 1962.

[Heise 05a] H. Heise, W. Hoyer, S. Becker, O. C. Andronesi, D. Riedel & M. Baldus. *Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR*. Proc. National Acad. Sciences United States Am., vol. 102, no. 44, pages 15871–15876, 2005.

[Heise 05b] H. Heise, K. Seidel, M. Etzkorn, S. Becker & M. Baldus. *3D NMR spectoscopy for resonance assignment and structure elucidation of proteins under MAS: novel pulse schemes and sensitivity considerations*. J. Magn. Resn., vol. 173, pages 64–74, 2005.

[Heise 08] H. Heise. *Solid-state NMR spectroscopy of amyloid proteins*. Chembiochem, vol. 9, no. 2, pages 179–189, 2008.

[Hennetin 06] J. Hennetin, B. Jullian, A. C. Steven & A. V. Kajava. *Standard Conformations of $\beta$-Arches in $\beta$-Solenoid Proteins*. J. Mol. Biol., vol. 358, pages 1094–1105, 2006.

[Hiller 05] M. Hiller, L. Krabben, K. R. Vinothkumar, F. Castellani, B.-J. van Rossum, W. Kühlbrand & H. Oschkinat. *Solid-State Magic-Angle-Spinning NMR of Outer-Membrane Protein G from Escherichia coli*. ChemBioChem, vol. 6, pages 1679–1684, 2005.

[Hiller 08] M. Hiller, V. A. Higman, S. Jehle, B. J. van Rossum, W. Kuhlbrandt & H. Oschkinat. *[2,3-C-13]-labeling of aromatic residues – getting a head start in the magic-angle-spinning NMR assignment of membrane proteins*. J. Am. Chem. Soc., vol. 130, no. 2, pages 408–+, 2008.

[Hing 92] A. W. Hing, S. Vega & J. Schaefer. *Transferred-echo double-resonance nmr*. J. Magnetic Resonance, vol. 96, no. 1, pages 205–209, 1992.

[Holbert 01] S. Holbert, I. Denghien, T. Kiechle, A. Rosenblatt, C. Weelington, M. R. Hayden, R. L. Margolis, C. A. Ross, J. Dausset, R. J. Ferrante & C. Néri. *The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: Neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis*. PNAS, vol. 98, pages 1811–1816, 2001.

[Igumenova 04a] T. I. Igumenova, A. E. McDermott, K. W. Zilm, R. W. Martin, E. K. Paulson & A. J. Wand. *Assignment of Carbon NMR Resonances for Microcrystalline Ubiquitin*. J. Am. Chem. Soc., vol. 126, pages 6720–6727, 2004.

[Igumenova 04b] T. I. Igumenova, A. J. Wand & A. E. McDermott. *Assignment of the Backbone Resonances for Microcrystalline Ubiquitin*. J. Am. Chem. Soc., vol. 126, pages 5323–5331, 2004.

[Iwata 06] K. Iwata, T. Fujiwara, Y. Matsuki, S. Takahashi, H. Naiki & Y. Goto. *3D structure of amyloid protofilaments of $\beta_2$-microglobulin fragment probed by solid state NMR*. PNAS, vol. 103, pages 18119–18124, 2006.

[Iwatsubo 94] T. Iwatsubo, A. Odaka, N. Suzuki, H. Mizusawa, N. Nukina & Y. Ihara. *Visualization of A-Beta-42(43) and A-Beta-40 in senile plaques with end-specific A-Beta monoclonals - evidence that the initially deposited species is A-Beta-42(43)*. Neuron, vol. 13, pages 45–53, 1994.

[Jaroniec 00] C. P. Jaroniec, B. A. Tounge, C. M. Rienstra, J. Herzfeld & R. G. Griffin. *Recoupling of Heteronuclear Dipolar Interactions with Rotational-Echo Double-Resonance at High Magic-Angle Spinning Frequencies*. J. Magn. Reson., vol. 146, pages 132–139, 2000.

[Jaroniec 04] C. P. Jaroniec, C. E. MacPhee, V. S. Bajaj, M. T. McMahon, C. M. Dobson & R. G. Griffin. *High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy*. PNAS, vol. 101, pages 711–716, 2004.

[Jarrett 93] J. T. Jarrett, E. P. Berger & P. T. Lansbury. *The carboxy terminus of the beta-amyloid protein is critical for the seeding of amyloid formation - implications for the pathogenesis of alzheimers-disease*. Biochem., vol. 32, no. 18, pages 4693–4697, 1993.

[Jenkins 01] J. Jenkins & R. Pickersgill. *The architecture of parallel beta-helices and related folds*. Progress In Biophys. & Mol. Biol., vol. 77, no. 2, pages 111–175, 2001.

[Kajava 06] A. V. Kajava & A. C. Steven. *beta-rolls, beta-helices, and other beta-solenoid proteins*. Fibrous Proteins: Amyloids, Prions Beta Proteins, Advances in Protein Chemistry, vol. 73, pages 55–+, 2006.

[Kayed 03] R. Kayed, E. Head, J. L. Thompson, T. M. McIntire, S. C. Milton, C. W. Cotman & C. G. Glabe. *Common Structure of Soluble Amyloid Oligomers Implies Common Mechanism of Pathogenesis*. Science, vol. 300, pages 486–489, 2003.

[Kheterpal 01] I. Kheterpal, A. Williams, C. Murphy, B. Bledsoe & R. Wetzel. *Structural features of the A beta amyloid fibril elucidated by limited proteolysis*. Biochem., vol. 40, no. 39, pages 11757–11767, 2001.

[Kloepper 07] K. D. Kloepper, D. H. Zhou, Y. Li, K. A. Winter, J. M. George & C. M. Rienstra. *Temperature-dependent sensitivity enhancement of solid-state NMR spectra of alpha-synuclein fibrils*. J. Biomolecular Nmr, vol. 39, no. 3, pages 197–211, 2007.

[Kozhukh 02] G. V. Kozhukh, Y. Hagihara, T. Kawakami, K. Hasegawa, H. Naiki & Y. Goto. *Investigation of a peptide responsible for amyloid fibril formation of beta(2)-microglobulin by Achromobacter protease I*. J. Biological Chem., vol. 277, no. 2, pages 1310–1315, 2002.

[Lange 02] A. Lange, S. Luca & M. Baldus. *Structural Constraints from Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids*. J. Am. Chem. Soc., vol. 124, pages 9704–9705, 2002.

[Lange 06] A. Lange, K. Giller, S. Hornig, M.-F. Martin-Eauclaire, O. Pongs, S. Becker & M. Baldus. *Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR*. Nature, vol. 440, pages 959–962, 2006.

[Lange 08] A. Lange, H. van Melkebeke, C. Wasmer, R. Verel & B. H. Meier. *Amyloids and Prions as seen by NMR*. Annual Meeting of (CA) NMR-Life, Königstein, 2008.

[Laws 02] D.D. Laws, H.-M. L. Bitter & A. Jerschow. *Solid-State NMR Spectroscopic Methods in Chemistry*. Angew. Chem. Int. Ed., vol. 41, pages 3096–3129, 2002.

[Lemaster 82] D. M. Lemaster & J. E. Cronan. *Biosynthetic production of c-13-labeled amino-acids with site-specific enrichment*. J. Biological Chem., vol. 257, no. 3, pages 1224–1230, 1982.

[LeVine 93] H. LeVine. *Thioflavine T interaction with synthetic Alzheimer's disease $\beta$-amyloid peptides: Detection of aggregation in solution*. Protein Science, vol. 2, pages 404–410, 1993.

[Levitt 01] M. H. Levitt. *Spin Dynamics*. Wiley Chichester, 2001.

[Luca 07] S. Luca, W. M. Yau, R. Leapman & R. Tycko. *Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR*. Biochem., vol. 46, no. 47, pages 13505–13522, 2007.

[Lysek 05] D. A. Lysek, C. Schorn, L. G. Nivon, V. Esteve-Moya, B. Christen, L. Calzolai, C. von Schroetter, F. Fiorito, T. Herrmann, P. Guntert & K. Wuthrich. *Prion protein NMR structures of cats, dogs, pigs, and sheep*. Proc. National Acad. Sciences United States Am., vol. 102, no. 3, pages 640–645, 2005.

[Lührs 05] T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Dobeli, D. Schubert & R. Riek. *3D structure of Alzheimer's amyloid-$\beta$(1-42) fibrils*. PNAS, vol. 102, pages 17342–17347, 2005.

[Macdonald 93] M. E. Macdonald, C. M. Ambrose, M. P. Duyao, R. H. Myers, C. Lin, L. Srinidhi, G. Barnes, S. A. Taylor, M. James, N. Groot, H. Macfarlane, B. Jenkins, M. A. Anderson, N. S. Wexler, J. F. Gusella, G. P. Bates, S. Baxendale, H. Hummerich, S. Kirby, M. North, S. Youngman, R. Mott, G. Zehetner, Z. Sedlacek, A. Poustka, A. M. Frischauf, H. Lehrach, A. J. Buckler, D. Church, L. Doucettestamm, M. C. Odonovan, L. Ribaramirez, M. Shah, V. P. Stanton, S. A. Strobel, K. M. Draths, J. L. Wales, P. Dervan, D. E. Housman, M. Altherr, R. Shiang, L. Thompson, T. Fielder, J. J. Wasmuth, D. Tagle, J. Valdes, L. Elmer, M. Allard, L. Castilla, M. Swaroop, K. Blanchard, F. S. Collins, R. Snell, T. Holloway, K. Gillespie, N. Datson, D. Shaw & P. S. Harper. *A novel gene containing a trinucleotide repeat that is expanded and unstable on huntingtons-disease chromosomes*. Cell, vol. 72, no. 6, pages 971–983, 1993.

[Macias 00] M. J. Macias, V. Gervais, C. Civera & H. Oschkinat. *Structural analysis of WW domains and design of a WW prototype*. nature structural biology, vol. 7, pages 375–379, 2000.

[Makin 05] O. S. Makin & L. C. Serpell. *Structures for amyloid fibrils*. FEBS J, vol. 272, pages 5950–5961, 2005.

[Marion 83] D. Marion & K. Wuthrich. *Application of phase sensitive two-dimensional correlated spectroscopy (cosy) for measurements of h-1-h-1 spin-spin coupling-constants in proteins*. Biochemical Biophys. Research Comm., vol. 113, no. 3, pages 967–974, 1983.

[Marley 01] J. Marley, M. Lu & C. Bracken. *A method for efficient isotopic labeling of recombinant proteins*. J. Biomolecular Nmr, vol. 20, no. 1, pages 71–75, 2001.

[McGowan 00] D. P. McGowan, W. van Roon-Mom, H. Holloway, G. P. Bates, L. Mangiarini, G. J. S. Cooper, R. L. M. Faull & R. G. Snell. *Amyloid-like inclusions in Huntington's disease*. Neuroscience, vol. 100, no. 4, pages 677–680, 2000.

[Metz 94] G. Metz, X. Wu & S. O. Smith. *Ramped-Amplitude Cross Polarization in Magic-Angle-Spinning NMR*. J. Magn. Resn. A, vol. 110, pages 219–227, 1994.

[Michal 97] C. A. Michal & L. W. Jelinski. *REDOR 3D: Heteronuclear distance measurements in uniformly labeled and natural abundance solids*. J. Am. Chem. Soc., vol. 119, no. 38, pages 9059–9060, 1997.

[Morcombe 03] C. R. Morcombe & K. W. Zilm. *Chemical shift referencing in MAS solid state NMR*. J. Magn. Reson., vol. 162, pages 479–486, 2003.

[Morcombe 04] C. R. Morcombe, V. Gaponenko, R. A. Byrd & K. W. Zilm. *$^{13}C$ CPMAS Spectroscopy of Deuterated Proteins: CP Dynamics, Line Shapes, and $T_1$ Relaxation*. J. Am. Chem. Soc., vol. 127, pages 397–404, 2004.

[Mori 92] H. Mori, K. Takio, M. Ogawara & D. Selkoe. *Mass-spectrometry of purified amyloid-beta protein in alzheimers-disease*. J. Biological Chem., vol. 267, no. 24, pages 17082–17086, 1992.

[Naiki 97] H. Naiki, N. Hashimoto, S. Suzuki, H. Kimura, K. Nakakuki & F. Gejyo. *Establishment of a kinetic model of dialysis-related amyloid fibril extension in vitro*. Amyloid-international J. Experimental Clinical Investigation, vol. 4, no. 4, pages 223–232, 1997.

[Nelson 05] R. Nelson, M. R. Sawaya, M Balbirnie, A. O. Madsen, C. Riekel, R. Grothe & D. Eisenberg. *Structure of the cross-$\beta$ spine of amyloid-like fibrils*. Nature, vol. 435, pages 773–778, 2005.

[O'Nuallain 02] B. O'Nuallain & R. Wetzel. *Conformational Abs recognizing a generic amyloid fibril epitope*. Proc. National Acad. Sciences United States Am., vol. 99, no. 3, pages 1485–1490, 2002.

[Paravastu 06] A. K. Paravastu, A. T. Petkova & R. Tycko. *Polymorphic Fibril Formation by Residue 10-40 of the ALzheimer's $\beta$-Amyloid Peptide*. Biophys. J., 2006.

[Pauli 01] J. Pauli, M. Baldus, B. van Rossum, H. de Groot & H. Oschkinat. *Backbone and Side-Chain 13C and 15N Signal Assignment of the $\alpha$-Spectrin SH3 Domain by Magic Angle Spinning Soild-State NMR at 17.6 Tesla*. ChemBioChem, vol. 2, pages 272–281, 2001.

[Paulson 03] E. K. Paulson, C. R. Morcombe, V. Gaponenko, B. Dancheck, R. A. Byrd & K. W. Zilm. *High-Sensitivity Observation of Dipolar Exchange and NOEs between Exchangable Protons in Proteins by 3D Solid-State NMR Spectroscopy*. J. Am. Chem. Soc., vol. 125, pages 14222–14223, 2003.

[Peranen 96] J. Peranen, M. Rikkonen, M. Hyvonen & L. Kaariainen. *T7 vectors with a modified T7lac promoter for expression of proteins in Escherichia coli*. Analytical Biochem., vol. 236, no. 2, pages 371–373, 1996.

[Petkova 02] A. T. Petkova, Y. Ishii, J. J. Babach, O. N. Antzutkin, R. D. Leapman, F. Delaglio & R. Tycko. *A structural model for Alzheimer's $\beta$-amyloid fibrils based on experimental constraints from solid state NMR*. PNAS, vol. 99, pages 16742–16747, 2002.

[Petkova 04] A. T. Petkova, G. Buntkowsky, F. Dyda, R. D. Leapman, W.-M. Yau & R. Tycko. *Solid State NMR Reveals a PH-dependent Antiparallel $\beta$-Sheet Registry in Fibrils Formed by a $\beta$-Amyloid Peptide*. J. Mo. Biol., vol. 335, pages 247–260, 2004.

[Petkova 06] A. T. Petkova, W.-M. Yau & R. Tycko. *Experimental Constraints on Quarternary Structure in Alzheimer's $\beta$-Amyloid Fibrils*. Biochemistry, vol. 45, pages 498–512, 2006.

[Petrovich 06] M. Petrovich, A. L. Johnsson, N. Ferguson, V. Daggett & A. R. Fersht. *$\Phi$-Analysis at the Experimental Limits: Mechanism of $\beta$-Hairpin Formation*. J. Mol. Biol., vol. 360, pages 865–881, 2006.

[Puchtler 65] H. Puchtler & F. Sweat. *Congo Red as a stain for fluorescence microscopy of amyloid*. J. Histochem. Cytochem., vol. 13, pages 693–694, 1965.

[Raetz 95] C. R. H. Raetz & S. L. Roderick. *A left-handed parallel beta-helix in the structure of udp-n-acetylglucosamine acyltransferase*. Science, vol. 270, no. 5238, pages 997–1000, 1995.

[Rienstra 02] C. M. Rienstra, M. Hohwy, L. J. Mueller, C. P. Jaroniec, B. Reif & R. G. Griffin. *Determination of Multiple Torsion-Angle Constraints in U-$^{13}$C-$^{15}$N-Labeled Peptides: 3D $^{1}$H-$^{15}$N-$^{13}$C-$^{1}$H Dipolar Chemical Shift NMR Spectriscopy in Rotating Solids*. J. Am. Chem. Soc., vol. 124, pages 11908–11922, 2002.

[Ritter 05] C. Ritter, M.-L-Maddelein, A. B. Siemer, T. Lührs, M. Ernst, B. H. Meier, S. J. Saupe & R. Riek. *Correlation of structural elements and infectivity of the HET-s prion*. Nature, vol. 435, pages 844–848, 2005.

[Saraiva 01] M. J. M. Saraiva. *Transthyretin mutations in hyperthyroxinemia and amyloid diseases*. Human Mutation, vol. 17, no. 6, pages 493–503, 2001.

[Sattler 99] M. Sattler, J. Schleucher & C. Griesinger. *Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients*. Progress In Nuclear Magnetic Resonance Spectr., vol. 34, no. 2, pages 93–158, 1999.

[Saupe 00] S. J. Saupe. *Molecular Genetics of Heterokaryon Incompatibility in Filamentous Ascomycetes*. Microbiol. Mol. Biol. Rev., vol. 64, pages 489–502, 2000.

[Sawaya 07] M. R. Sawaya, S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. O. Madsen, C. Riekel & D. Eisenberg. *Atomic structures of amyloid cross-$\beta$ spines reveal varied steric zippers*. Nature, vol. 447, pages 453–457, 2007.

[Schwieters 03] C. D. Schwieters, J. J. Kuszewski, N. Tjandra & G. M. Clore. *The Xplor-NIH NMR molecular structure determination package*. J. Magnetic Resonance, vol. 160, no. 1, pages 65–73, 2003.

[Seidel 04] K. Seidel, A. Lange, S. Becker, C. E. Hughes, H. Heise & M. Baldus. *Protein solid-state NMR resonance assignment from(13C,13C) correlation spectroscopy*. Phys. Chem. Chem. Phys., vol. 6, pages 5090–5093, 2004.

[Selkoe 94] D. J. Selkoe. *Normal and abnormal biology of the beta-amyloid precursor protein*. Ann. Rev. Neuroscience, vol. 17, pages 489–517, 1994.

[Serpell 99] L. C. Serpell, P. E. Fraser & M. Sunde. *X-ray fiber diffraction of amyloid fibrils*. Amyloid, Prions, Other Protein Aggregates, vol. 309, pages 526–536, 1999.

[Shaka 83] A. J. Shaka, J. Keeler, T. Frenkiel & R. Freeman. *An Improved Sequence for Broadband Decoupling: WALTZ-16*. J. Magn. Resn., vol. 52, pages 335–338, 1983.

[Siemer 06a] A. B. Siemer, A. A. Arnold, C. Ritter, T. Westfeld, M. Ernst, R. Riek & B. H. Meier. *Observation of Highly Flexible Residues in Amyloid Fibrils of the HET-s Prion*. J. AM. Chem. Soc., vol. 128, pages 13224–13228, 2006.

[Siemer 06b] A. B. Siemer, C. Ritter, M. O. Steinmetz, M. Ernst, R. Riek & B. H. Meier. *$^{13}$C, $^{15}$N Resonance assignment of parts of the HET-s prion protein in its amyloid form*. J. Biomol. NMR, vol. 34, pages 75–87, 2006.

[Sipe 00] J. D. Sipe & A S Cohen. *Review: History of the Amyloid Fibril*. J. Struc. Biol., vol. 130, pages 88–89, 2000.

[Stefani 03] M. Stefani & C. M. Dobsen. *Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution*. J. Mol. Med., vol. 81, pages 678–699, 2003.

[Stefani 07] M. Stefani. *Generic cell dysfunction in Neurodegenerative disorders: Role of surfaces in early protein misfolding, aggregation, and aggregate cytotoxicity*. Neuroscientist, vol. 13, no. 5, pages 519–531, 2007.

[Steinbacher 94] S. Steinbacher, R. Seckler, S. Miller, B. Steipe, R. Huber & P. Reinemer. *Crystal-structure of p22 tailspike protein - interdigitated subunits in a thermostable trimer*. Science, vol. 265, no. 5170, pages 383–386, 1994.

[Sudol 96] M. Sudol. *The WW domain binds polyprolines and is involved in human diseases*. Experimental Mol. Medicine, vol. 28, no. 2, pages 65–69, 1996.

[Sune 97] C. Sune, T. Hayashi, Y. Liu, W. S. Lane, R. A. Young & M. A. GarciaBlanco. *CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription*. Mol. Cellular Biol., vol. 17, no. 10, pages 6029–6039, 1997.

[Szeverenyi 82] N. M. Szeverenyi, M. J. Sullivan & G. E. Maciel. *Observation of spin exchange by two-dimensional fourier-transform c-13 cross polarization-magic-angle spinning*. J. Magnetic Resonance, vol. 47, no. 3, pages 462–475, 1982.

[Takegoshi 01] K. Takegoshi, S. Nakamura & T. Terao. *$^{13}C$-$^{1}H$ dipolar assisted rotational resonance in magic-angle spinning NMR*. Chem. Phys. Lett., vol. 344, pages 631–637, 2001.

[Tycko 06] R. Tycko. *Molecular structure of amyloid fibrils: insights from solid-state NMR*. Quarterly Reviews of Biophysics, vol. 39, pages 1–55, 2006.

[Uversky 04] V. N. Uversky & A. L. Fink. *Conformational constraints for amyloid fibrillation: the importance of being unfolded*. Biochimica Et Biophysica Acta-proteins Proteomics, vol. 1698, no. 2, pages 131–153, 2004.

[Ventura 04] S. Ventura, J. Zurdo, S. Narayanan, M. Parreno, R. Mangues, B. Reif, F. Chiti, E. Giannoni, C. M. Dobson, F. X. Aviles & L. Serrano. *Short amino acid stretches can mediate amyloid formation in globular proteins: The Src homology 3 (SH3) case*. PNAS, vol. 101, pages 7158–7263, 2004.

[Vinogradov 99] E. Vinogradov, P. K. Madhu & S. Vega. *High-resolution protein solid-state NMR spectroscopy by phase-modulated Lee-Goldberg experiment*. Chem. Phys. Lett., vol. 314, pages 443–450, 1999.

[Vranken 05] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, P. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides & E. D. Laue. *The CCPN data model for NMR spectroscopy: Development of a software pipeline*. Proteins-structure Function Bioinformatics, vol. 59, no. 4, pages 687–696, 2005.

[Wasmer 08] C. Wasmer, A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek & B. H. Meier. *Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core*. Science, vol. 319, no. 5869, pages 1523–1526, 2008.

[Westermark 90] P. Westermark, K. Sletten, B. Johansson & G. G. Cornwell. *Fibril in senile systemic amyloidosis is derived from normal transthyretin*. Proc. National Acad. Sciences United States Am., vol. 87, no. 7, pages 2843–2845, 1990.

[Westermark 05a] P. Westermark. *Aspects on human amyloid forms and their fibril polypeptides*. FEBS J., vol. 272, pages 5942–5949, 2005.

[Westermark 05b] P. Westermark, M. D. Benson, J. N. Buxbaum, A. S. Cohen, B. Frangione, S. I. Ikeda, C. L. Masters, G. Merlini, M. J. Saraiva & J. D. Sipe. *Amyloid: Toward terminology clarification - Report from the Nomenclature Committee of the International Society of Amyloidosis*. Amyloid-journal Protein Folding Disorders, vol. 12, no. 1, pages 1–4, 2005.

[Wetzel 02] R. Wetzel. *Ideas of Order for amyloid Fibril Structure*. Structure, vol. 10, pages 1031–1036, 2002.

[Wickner 94] R. B. Wickner. *[ure3] as an altered ure2 protein - evidence for a prion analog in saccharomyces-cerevisiae*. Science, vol. 264, no. 5158, pages 566–569, 1994.

[Wishart 95] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley & B. D. Sykes. *1H, 13C and 15N chemical shift referencing in biomolecular NMR*. J. Biomol. NMR, vol. 6, pages 135–140, 1995.

[Wu 93] X. L. Wu & K. W. Zilm. *Cross-polarization with high-speed magic-angle-spinning*. J. Magnetic Resonance Series A, vol. 104, no. 2, pages 154–165, 1993.

[Yoder 93] M. D. Yoder, N. T. Keen & F. Jurnak. *New domain motif - the structure of pectate lyase-c, a secreted plant virulence factor*. Science, vol. 260, no. 5113, pages 1503–1507, 1993.

[Zandomeneghi 04] G. Zandomeneghi, M. R. H. Krebs, M. G. McCammon & M. Fändrich. *FTIR reveals structural differences between native $\beta$-sheet proteins and amyloid fibrils*. Protein Science, vol. 13, pages 3314–3321, 2004.

[Zech 05] S. G. Zech, A. J. Wand & A. E. McDermott. *Protein Structure Determination by High-Resolution Solid-State NMR Spectroscopy: Application to Microcrystalline Ubiquitin*. J. Am. Chem. Soc., vol. 127, pages 8618–8626, 2005.

# Acknowledgement – Danksagung

This work would not have been possible without the help of many people. Some of them were directly involved in the work with the CA150.WW2 amyloid project, others taught me solid and solution state NMR, sample preparation, introduced me to the world of structural biology, and supported me during my everyday PhD-student's life.

Diese Arbeit habe ich am Leibnizinstitut für Molekulare Pharmakologie (FMP) in Berlin-Buch in der Abteilung Strukturbiologie unter der Leitung von **Hartmut Oschkinat** durchgeführt. Bei ihm möchte ich mich für die Möglichkeit, die CA150.WW2 Amyloidfibrillen zu untersuchen und für sein großes Interesse an dem Projekt bedanken. Dazu gehörte vor allem auch die uneingeschränkte Möglichkeit, auf die Ressourcen der Gruppe zurückgreifen zu können.

I want to thank the main collaborators of the project, **Alan R. Fersht** and **Neil Ferguson** from the MRC in Cambridge, UK. Their intense work with WW-domains built the basis of the project and without the close collaboration throughout the project it would not have been possible to do this research. I specially thank Neil who prepared most of the solid state NMR samples. Together with **Henning Tidow** he studied the fibrillation kinetics of a vast amount of CA150.WW2 variants. In addition, he really pushed the whole project forward, showed huge enthusiasm and taught me a lot, especially during my visit at the MRC. I thank **John Berriman** from the SBC in New York for the electron microscopy work. **Jeremy Flinders** did a lot of initial solid state NMR work on the CA150.WW2 amyloid system at the FMP. Due to his work the project was well established when I started on it. I very much thank **Elisabeth Dowler, Michael Lisurek, Stefan Jehle, Janet Zapke, Barth-Jan van Rossum, Vivien Lange, and Carolyn Vargas** for proof-reading the manuscript.

Bei **Anne Diehl** und **Martina Leidert** möchte ich mich sehr herzlich für die unermütliche Präparation zahlreicher NMR-Proben bedanken. Ich möchte mich auch bei allen anderen Kollegen für ihre Unterstützung vor allem bei meinem kurzen Ausflug in die Herstellung rekombinanter Proteine bedanken, besonders bei **Vivien Lange**, **Janina Hahn**, **Kristina Rehbein** und **Lilo Handel**.

**Gerd Krause** danke ich für die Moleküldynamikrechnungen der CA150.WW2 Amyloid-fibrille, sowie **Ronald Kühne** und **Stefan Jehle** für die Hilfe bei der Strukturrechnung mit X-plor. Ich danke auch **Christoph Brockmann** auf dessen Skript die Strukturrechnungen

basierten. Desweiteren gilt mein Dank **Frank Eisenmenger**, der sich stets darum gekümmert hat, dass ich Hard- und Software zur Verfügung hatte.

Bei **Barth-Jan van Rossum** und **Stefan Jehle** möchte ich mich für ihre Unterstützung in allen Fragen der Festkörper-NMR-Spektroskopie und für NMR-Erste Hilfe bedanken. **Uwe Lange** (von Bruker) verdanke ich viele praktische Tipps und Einblicke in die Funktionsweise der Magneten und Konsolen. Außerdem möchte ich mich bei **Peter Schmieder** für die schnellen und unkomplizierten Lösungs-NMR-Messungen und vor allem auch für seine Unterstützung in allen NMR-Fragen bedanken. Mein besonderer Dank gilt auch **Kay Saalwächter**, der mir während meiner Diplomarbeit die Grundlagen der Festkörper-NMR-Spektroskopie beigebracht hat.

Ganz besonders möchte ich mich bei denjenigen bedanken, die dazu beigetragen haben, dass ich immer gerne ans FMP gekommen bin und die sich stets für die Hochs und Tiefs meines Alltags interessiert haben: **Vivien Lange**, **Britta Kunert**, **Janet Zapke** und **Vicky Davies**. Bei Vicky möchte ich mich auch für ihren wissenschaftlichen Rat und ihre Hilfe bei CCPN bedanken. Mein besonderer Dank gilt Vivien für ihre Freundschaft. Bedanken möchte ich mich auch bei **Andrea Steuer** für ihre Unterstützung bei allen organisatorischen Fragen und Problemen. Für die Organisation diverser Sportveranstaltungen und ihre ansteckende Sportbegeisterung möchte ich mich bei **Kristin Pankow** bedanken. Ebenso gilt mein Dank **Juliane Walter** und **Janina Hahn** für die interessante und spannende Zusammenarbeit als Doktorandenvertreterinnen. Ich möchte mich auch bei meinen Bürokollegen **Anna Schrey**, **Michael Lisurek** und **Matthias Hiller** für die schöne Atmosphäre bedanken, letzterem auch für die fachliche Zusammenarbeit. Mein Dank gilt auch meinen ersten Bürokollegen bzw. -nachbarn **Michael Soukenik** und **Urs Wiedemann** für ihre Unterstützung in meiner Anfangszeit am FMP.

Dem **SFB 449** verdanke ich die Finanzierung meiner Arbeit und der **Forschergruppe 475** die Möglichkeit, meine Arbeit auf dem internationalen Symposium der DFG Forschergruppe 475 “Folding and Stability of $\beta$-sheets“ vorzustellen.

Besonderer Dank geht an alle, die mich außerhalb des Institutes unterstützt haben, vor allem an meine Familie.

# Curriculum vitae

Due to data privacy part of the Curriculum vitae cannot be found in the online version of this thesis.

## Veröffentlichungen – Publications

Becker J, Dunmur DA, Dunn CJ, Luckhurst GR, Marchant-Lane SE, Timimi BA, *On the investigation of field-induced director dynamics: A novel ESR experiment*, Molecular Crystals and Liquid Crystals 394, 45-61, 2003.

Becker J, Becher F, Hucke O, Labahn A, Koslowski T, *Theory and simulation of vibrational coupling in deuterated proteins: Toward a new structural probe?*, Journal of Physical Chemsitry B 107, 12878-12883, 2003.

Becker J, Comotti A, Simonutti R, Sozzani P, Saalwächter K, *Molecular motion of isolated alkanes in nanocrystals*, Journal of Physical Chemsitry B 109, 23285-23294, 2005.

Ferguson N, Becker J, Tidow H, Tremmel S, Sharp TD, Krause G, Flinders J, Petrovich M, Berriman J, Oschkinat H, Fersht AR, *General structural motifs of amyloid protofilaments*, PNAS 103, 16248-16253, 2006.

Becker J, Furgoson N, Flinders J, van Rossum B-J, Fersht AR, Oschkinat H, *A sequential assignment procedure for proteins that have intermediate line width in MAS NMR spectra: Amyloid fibrils of human CA150.WW2*, ChemBioChem, accepted.

## Tagungsbeiträge – Conference Contributions

Bunsentagung 2003 in Kiel, Poster P12: *Theory and Simulation of vibrational coupling in deuterated proteins: Towards a new structural probe?*

Jahrestagung der Fachgruppe Magnetresonance in Aachen 2004, Poster SOL9: *Local dynamics of single sites in n-alkyl inclusion compounds by means of solid state NMR.*

ICMRBS 2006 in Göttingen, Poster P262: *Sample Preparation for Membrane Proteins by Solid State NMR.*

Alpine Conference on Solid State NMR in Chamonix 2007, Poster P12: *Structural Study of CA150.WW2 in its Amyloid Form by MAS NMR.*

International Symposium of the DFG research unit 475 in Golm 2007, Talk: *Structural investigations on amyloid fibrils formed by the CA150.WW2 domain.*