Technische Universität München

Fakultät für Chemie

Festkörper-NMR-Spektroskopie

Solid- and Solution-State Nuclear Magnetic Resonance Spectroscopic Analysis of Amyloidogenic Immunoglobulin Light Chains and their Interactions with Epigallocatechin Gallate

Manuel Christian Hora

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:	Prof. Dr. Sevil Weinkauf
Prüfer:	
	1. Prof. Dr. Bernd Reif
	2. Prof. Dr. Johannes Buchner
	3. Prof. Dr. Marina Ramirez-Alvarado

Die Dissertation wurde am 22.6.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 7.9.2016 angenommen.

Table of Contents

Acknowledgements

Bernd, I really enjoyed the time in your group. Your contagious enthusiasm was only sometimes exhausting, but most of all motivating. Whenever I thought, I had nice results and could take a break, you incited me to strive for more. I am grateful for your patient explanations and for honest and familiar conversations. You encouraged me to think outside the box and offered the freedom to come up with my own ideas about the project. It was also great, that I could attend so many conferences. Thank you for the opportunity to grow.

Thank you, Johannes. When I applied at TUM for a place in the Bachelor's programme, I had my interview with you. Now, nine years later, I'm glad that you will also be an examiner of my PhD thesis.

Marina, although we met only once before in Heidelberg, you immediatelly agreed to be examiner of my thesis. Thank you for your advice, the great time in Uppsala and most of all, thank you that you came the long way to Munich for my defense.

Teji, I liked working together with you. It was great to see your progress during the Master thesis. I wish you all the best for your project and I am happy, that you continue the work on AL.

Thanks also to Cardine. You shared many plasmids and gave practical advice on the handling of MAK. It was great to participate in your project on the different mutants and I am grateful for the collaboration.

Thank you to Martin and Birgit in Jülich and Benedikt here at TUM for your contributions to my project. I appreciate your efforts.

Thank you, Severin, for the work during your lab rotation. It was good to see you again after several years. While I supervised you, I actually also learned from you.

Special thanks to Riddhiman. You spent a lot of time setting up experiments with me and spent even more time explaining NMR to me. Together, we kept our server running. Most important, you were a pleasant neighbour in the office!

I also want to thank Kai and Zdenek for help with both NMR theory and practice. Kai, thank you for the invitation to celebrate Chinese New Year!

Rainer, Gerd, Sam, Ralf, thank you for quick help, whenever the spectrometers did not work as expected. Thanks also for all the chats about NMR, computers and weird science.

Thank you to Christoph (G.) for SAXS, RDCs and lots of advice. Thanks also to Christoph (H.) for the great time in Italy!

Marie-Theres, thank you for the crystallisation experiments. The MAK V_L didn't want to form crystals, but I enjoyed working with you.

Very special thanks to Elke, Diana, Matthias, Markus and Carina. It was a great time with you. Elke, I could always trust you and rely on you. Thank you so much. Diana, your vivid style brought fun and excitement to our lab. Matthias, I will miss our discussions about politics and the future. Think positive! Markus, thank you for bringing people together. Carina, you did so much to establish good working conditions here. Also thanks to Agathe, the green friend on our desk.

Thank you, Antje, for all the purifications. It made my work much easier. Asita, thank you for the care and the cakes.

I had great colleagues here in our office, in the lab or for common lunch. Thank you, Vanessa, Zheng, Maristella, Eugen, Maria, Justin, Alex, Thomas and Carlo.

Special thanks to the whole EM group. Carsten, Christoph, Marianne, you helped me a lot, whenever I forgot, how to handle an electron microscope. Also, when there was trouble in our lab, I could always take refuge at your offices.

Sevil, thank you for all of our discussions, not only about science, but about life. Thank you for your trust and your honesty, no matter, on which topic.

Many thanks to the GERAMY colleagues. Karthik, Marcus, Young-In, Erich, Martin, Kathrin, Jan, after our first meeting in Berlin, every autumn I looked forward to see you again.

Special thanks to Ute and Stefan. I remember that day, when you emailed us about the transplantation. This was the moment, when I fully realised, that I didn't just work on an arbitrary scientific question, but on an actual disease that affected real people.

I wish to express my deep gratitude to my family and my friends, for all the support, the interest and also the distractions and the opportunities to relax after work.

Abbreviations

AFS	alternatively folded state
AL amyloidosis	antibody light chain amyloidosis
BMRB	Biological Magnetic Resonance Data Bank
BSH-CP	band selective homonuclear cross polarisation
CD spectroscopy	circular dichroism spectroscopy
CDR	complementarity determining region
C_L domain	constant domain of the immunoglobulin light chain
CP	cross polarisation
CSI	chemical shifts index
cw	continuous wave
DARR	dipolar assisted rotational recoupling
DREAM	dipolar recoupling enhanced by amplitude modulation
EGCG	epigallocatechin gallate
FID	free induction decay
FROSTY	freezing rotational diffusion of protein solutions at low temperature and high viscosity
HSQC	heteronuclear single quantum coherence experiment
K_D	dissociation constant
LC	light chain
MAS	magic angle spinning
MD	molecular dynamics
MW	molecular weight
MWCO	molecular weight cut off
NMR spectroscopy	nuclear magnetic resonance spectroscopy
NUS	nonuniform sampling
PAR	proton assisted recoupling
PDB	protein data bank
PDSD	proton driven spin diffusion
RMSF	root mean square fluctuation
UV	ultraviolet
V_L domain	variable domain of the immunoglobulin light chain
S/N ratio	signal-to-noise ratio
SPECIFIC-CP	spectrally induced filtering in combination with cross polarisation
SPINAL-64 decoupling	small phase incremental alternation decoupling with 64 steps
T_m	melting temperature
TEM	transmission electron microscopy
ThT	thioflavin T
WT	wild type

1 Abstract

1.1 Abstract (English)

Antibody light chain amyloidosis (AL amyloidosis) is a rare systemic disease, caused by amyloid aggregation of secreted antibody light chain (LC) variable domains (V_L). The hallmark of this disease is the broad variety of different V_L sequences, which can form the amyloid deposits. This variety leads to the question, which features render a V_L sequence amyloidogenic. Furthermore, the plethora of amyloidogenic sequences is a serious challenge for the development of drug candidates.

In this thesis, the MAK33 V_L domain was analysed as a model system for misfolding of LC V_L domains. We used NMR spectroscopy for the research on MAK33 V_L, as this method is able to characterise each of the physiologically and pathologically relevant states of the protein, ranging from the native structure via oligomers to the amyloid fibril form. To this end, we combined solution-state and MAS solid-state NMR spectroscopy.

We studied the effects of several amyloidogenic point mutations. These comprise the highly destabilising I2E mutation, by which the N-terminus is detached from the core of the domain, as well as the S20N and D70N substitutions. The latter ones are of special interest, since the thermodynamic stability is not altered by the mutations. Instead, we identified H-bonding networks on the surface of the protein structure, which are changed by the mutations. This leads to further structural changes in the hydrophobic core. Finally, we studied the role of the linker residue R108 for both the V_L domain and the constant (C_L) domain. While the effects of the deletion of R108 on MAK33 V_L are only minor, the C_L domain adopts an alternative conformation upon removal of R108. This conformation is also accessible by the C_L WT, but has not been identified before due to its small population. In all of these cases, our NMR spectroscopic analyses were highly complementary to thermodynamic measurements and molecular dynamics (MD) simulations.

Next, we examined epigallocatechin gallate (EGCG) as a candidate for therapy of AL amyloidosis. The green tea polyphenol EGCG is known to interfere with the amyloid formation of many other proteins. We combined solution-state NMR spectroscopy, site-directed mutagenesis and MD simulations to identify two binding sites of EGCG to MAK33 V_L. Interaction with one of these binding sites causes precipitation of the protein. With MAS solid-state NMR spectrocopy, we confirmed the non-amyloid nature of these aggregates. Precipitation kinetics revealed, that EGCG interacts prefentially with more amyloidogenic V_L variants. This feature makes it a highly interesting drug candidate. Sequence analyses showed, that the relevant EGCG binding site is conserved among V_L domains, enabling the broad use against different pathogenic LC sequences.

Finally, we report the first assignment experiments of an antibody V_L domain amyloid fibril. With MAS solid-state NMR spectroscopy, we identified 36 rigid residues, which constitute the core of the fibril. The rigid residues are located in the C-terminal part of the sequence. With respect to the native structure, they correspond mostly to the framework regions, whereas the CDR3 (complementarity determining region) residues are completely missing in the structured part of the fibril. Of special interest was the analysis of an oligomeric state of the V_L domain. The solid-state NMR spectra of these oligomers are highly similar to those of the fibrils. This supports the theory of oligomeric intermediates as on-pathway precursors of the amyloid fibrils. Complementary solution-state NMR experiments revealed the presence of many flexible residues. The flexible regions were different for the oligomers and the fibrils. The assignment of the flexible regions in the fibrils indicated an overlap with the rigid residues in the fibril core. Based on these results, we suggest two models: The multiple resonance sets might be caused by polymorphism or by partial dissociation of monomers from the fibril.

In summary, our results provide new insights into the structure and dynamics of the native, the oligomeric and the fibrillar state of V_L domains. This expands our understanding of AL amyloidosis and facilitates the development of therapeutic strategies. With EGCG, we presented an intriguing small molecule, which can discriminate V_L sequences with respect to their amyloidogenicity.

1.2 Abstract (Deutsch)

Antikörper-Leichtketten-Amyloidose (AL Amyloidose) ist eine seltene systemische Erkrankung, bei der sich variable Domänen (V_L Domänen) von sekretierten Antikörper-Leichtketten als Amyloidfibrillen ablagern. Eine Besonderheit der Krankheit ist die große Anzahl an verschiedenen V_L Sequenzen, die zu dieser Krankheit führen können. Es ist bislang nicht geklärt, durch welche Eigenschaften eine V_L Sequenz anfällig für die Fibrillenbildung wird. Die Vielzahl an in Frage kommenden Sequenzen stellt außerdem eine große Herausforderung für die Entwicklung von Medikamenten dar.

Im Rahmen dieser Dissertation wurde die V_L Domäne von MAK33 als Modellsystem für die Fehlfaltung von V_L Domänen untersucht. Eine Kombination aus Lösungs- und Festkörper-NMR Spektroskopie ermöglichte es uns, alle relevanten Faltungszustände zu betrachten: die native Struktur, Oligomere und die Fibrillen.

Wir konnten verschiedene Punktmutationen charakterisieren, durch die MAK33 V_L leichter Fibrillen bilden kann. Bei der instabilen I2E Mutation löst sich der N-Terminus leicht von der Domäne ab. Die S20N und D70N Mutanten hingegen haben die gleiche thermodynamische Stabilität wie der Wildtyp. Durch diese Mutationen verändern sich die Wasserstoffbrücken auf der Proteinoberfläche. Somit kommt es auch zu Veränderungen im hydrophoben Kern der Domäne. Wir untersuchten des Weiteren die Rolle von R108 im Linker für die Stabilität der V_L und der konstanten (C_L) Domäne. Die Deletion ΔR108 hatte nur geringe Auswirkungen auf die V_L Domäne. Die C_L Domäne hingegen nimmt durch die Deletion einen anderen Zustand an. Obwohl dieser Zustand in geringem Maße auch vom Wildtyp mit R108 populiert wird, war er bislang unbekannt. Unsere NMR spektroskopischen Experimente führten bei allen diesen Mutanten zu Erkenntnissen, die sich gut mit thermodynamischen Messungen und Molekulardynamik-Simulationen ergänzt haben.

Weiterhin haben wir untersucht, inwiefern der Naturstoff Epigallocatechin Gallat (EGCG) als Wirkstoff gegen AL Amyloidose geeignet ist. Von diesem Polyphenol ist bekannt, dass es die Fibrillenbildung von vielen anderen Proteinen verhindert. Mit einer Kombination aus Lösungs-NMR Spektroskopie, zielgerichteter Mutagenese und Molekulardynamik-Simulationen konnten wir zwei unterschiedliche Bindestellen von EGCG an MAK33 V_L identifizieren. Die Bindung von EGCG an eine dieser Stellen führt zur Präzipitation. Mittels Festkörper-NMR Spektroskopie konnten wir nachweisen, dass diese Präzipitate unstrukturiert sind. Weitergehende kinetische Messungen ergaben, dass EGCG bevorzugt amyloidogene V_L Varianten präzipitiert. Diese Eigenschaft ist besonders interresant für eine therapeutische Anwendung von EGCG. Sequenzvergleiche zeigten schließlich, dass die relevante Bindestelle in vielen V_L Domänen konserviert ist. Damit ist EGCG prinzipiell für den Einsatz gegen viele verschiedene pathogene Sequenzen von Leichtketten geeignet.

Als weiteres Ergebnis präsentieren wir hier die erste Zuordnung von NMR Spektren von V_L Fibrillen. Mittels Festkörper-NMR Spektroskopie konnten wir 36 Reste identifizieren, die zum strukturierten Kern der Fibrille gehören. Alle zugeordneten Reste befinden sich im C-terminalen Bereich der V_L Sequenz. Im Bezug auf die native Struktur entsprechen die Reste überwiegend dem konservierten Rahmen der Immunoglobu-

lindomäne. Die Reste aus CDR3 (complementarity determining region) fehlen hingegen vollständig in den strukturierten Bereichen der Fibrille. Besonders interessant war die Untersuchung der V_L Oligomere. Die Festkörper-NMR Spektren der Oligomere waren außerordentlich ähnlich zu denen der Fibrillen. Das unterstützt die Hypothese, dass diese Oligomere ein Intermediat auf dem Weg zur Fibrillenbildung darstellen. Bei ergänzenden Lösungs-NMR spektroskopischen Experimenten entdeckten wir viele flexible Reste, die sich in den Oligomeren und den Fibrillen deutlich unterscheiden. Die Zuordnung der dynamischen Reste in den Fibrillen ergab einen Überlapp mit den rigiden Resten im Kern der Fibrille. Auf Basis dieser Ergebnisse schlagen wir zwei Modelle für die Fibrillenstruktur vor: Die mehrfachen Signalsätze können eine Folge von Polymorphismus sein oder endständige Monomere können teilweise von der Fibrille dissoziieren.

Zusammenfassend haben wir neue Erkenntnisse über die Strukturen und Dynamiken von dem nativen Zustand ebenso wie den Oligomern und den Fibrillen der V_L Sequenzen erlangt. Dies erweitert unser Verständnis von AL Amyloidose und erleichtert die zukünftige Entwicklung neuer Therapien. Beispielhaft dafür präsentieren wir EGCG als ein kleines Molekül, dass in der Lage ist, V_L Sequenzen anhand ihrer Toxizität zu unterscheiden.

2 Introduction

2.1 Protein Structure and Amyloid Fibrils

In 1972, Christian Anfinsen was awarded the Nobel Prize in chemistry for his finding that 3D protein structure is determined by amino acid sequence [1]. This fundamental principle in protein folding led to practical applications like sequence based structure predictions. However, the sequence cannot be the only factor determining the structure, as some proteins can adopt multiple different structures. Interconversion between different conformations can be triggered e.g. by ligand binding. The actual native structure is not necessarily the most stable one by thermodynamics. Other states might just be not accessible due to kinetic barriers. A special type of protein structure is the amyloid state. Previously known from protein-misfolding diseases, it was found that destabilising conditions can cause myoglobin to form amyloid fibrils [2]. Systematic analysis revealed the amyloid state as mostly independent of side-chain interactions, instead relying on favourable backbone interactions [3]. Hence, the amyloid state could be accessible for all protein sequences and usually just prevented by kinetic barriers. Once a fibril nucleus has formed, extension proceeds faster, so the presence of a template reduces the required free energy for transiton [4].

The amyloid state is characterised by the cross-β fold. The β-sheets are oriented along the fibril axis and consist of strands contributed from many monomers. Amyloid fibrils formed from several different proteins produced very similar and characteristic X-ray diffraction patterns, indicating a common core structure [5]. The highly repetitive structure could be continued at both ends by recruiting additional monomers. An amyloid fibril structure at atomic resolution is displayed in Fig. 1 [6]. Several of these protofilaments could assemble to form mature amyloid fibrils with helical superstructure. All these features distinguish amyloid aggregates from unstructured protein aggregates.

2.2 Diseases Related to Amyloid Fibrils

While amyloid fibrils are also interesting from a point of view of protein folding, the driving force behind most research on amyloids comes from their involvement in severe diseases. Many different proteins are known to form amyloid deposits related to diseases. These so-called amyloidoses comprise widespread and well known diseases like Alzheimer's disease, Parkinson's disease, Creutzfeldt-Jakob disease or type II diabetes. Detailed lists of amyloidoses, together with the proteins forming the fibrils, can be found in several reviews [7, 8].
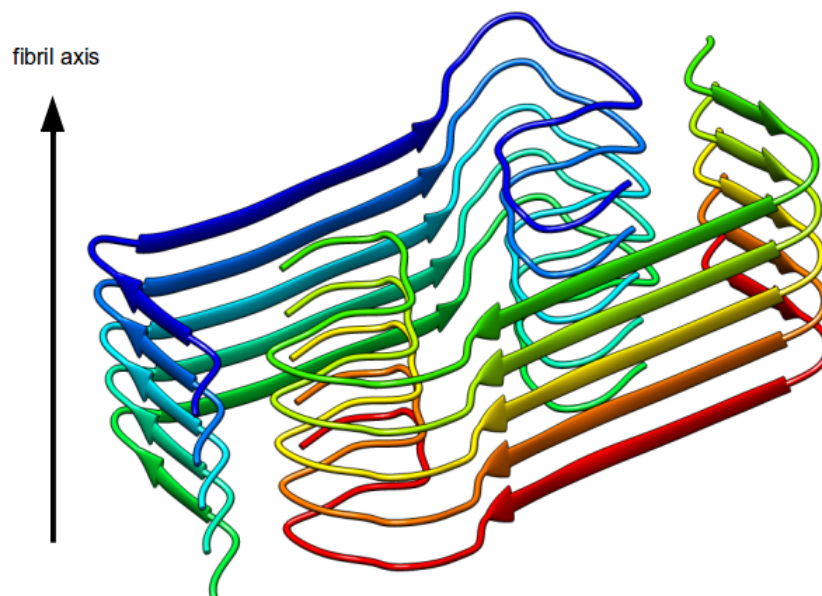
Figure 1: Structure of Amyloid-β-Fibrils with the Osaka Mutation
Amyloid-β is an archetypical protein for amyloid fibril formation. Each layer of the fibrils consists of two monomers. The fibril could be extended in both directions along the fibril axis. The image is based on PDB code 2MVX [6].

What is common to these diseases is ongoing discussion about their cause. The early assumption of fibril deposits causing the disease was questioned by findings that oligomeric intermediates are cytotoxic and fibrils might be inert aggregates. This hypothesis is supported e.g. by the aggressive arctic mutation of amyloid precursor protein, which leads to elevated levels of protofibrils, but does not increase the final amyloid load [9]. Similar results exist for familial amyloidotic polyneuropathy, where the protein transthyretin forms cytotoxic nonfibrillar aggregates prior to amyloid formation. The native tetramer and mature fibrils, in contrast are not neurotoxic [10]. The suggested common mechanism for this toxicity is attachment of intermediates to biological membranes, followed by pore formation [11].

Most of the research efforts so far naturally focussed on the most common amyloidoses. Alzheimer's disease, Parkinson's disease and type II diabetes are all well known examples for localised amyloidosis, as the deposits are found only in specific tissues or organs, e.g. nerves or pancreatic β-cells. In contrast, there are systemic amyloidoses, where fibril deposits are found in multiple organs. Systemic amyloidoses are very rare diseases and thus in general unknown to the broader public. The most common among these conditions is immunoglobulin light chain amyloidosis, typically abbreviated as AL amyloidosis for amyloidogenic light chains. The incidence is estimated at 9 patients per million inhabitants per year, according to a recent review [12], making AL amyloidosis an orphan disease.

2.3 Immunoglobulin Light Chain Amyloidosis

2.3.1 Pathological Background of Immunoglobulin Light Chain Amyloidosis

AL amyloidosis starts with a B-cell dyscrasia, which is ususally non-symptomatic on its own. The proliferating B-cells continue to produce antibodies, with a sequence unique for every clone of B-cells (see Fig. 2a). Despite sophisticated mechanisms of the cell for quality control of antibody assembly [13], in addition to whole antibodies, also antibody light chains (LC) are secreted into the circulatory system. Due to historical reasons, these LCs are also referred to as Bence-Jones protein [14]. The amount of free LCs in the serum is an important biomarker for AL amyloidosis, especially to monitor efficacy of treatment [15]. However, not all secreted LCs form amyloid fibrils.
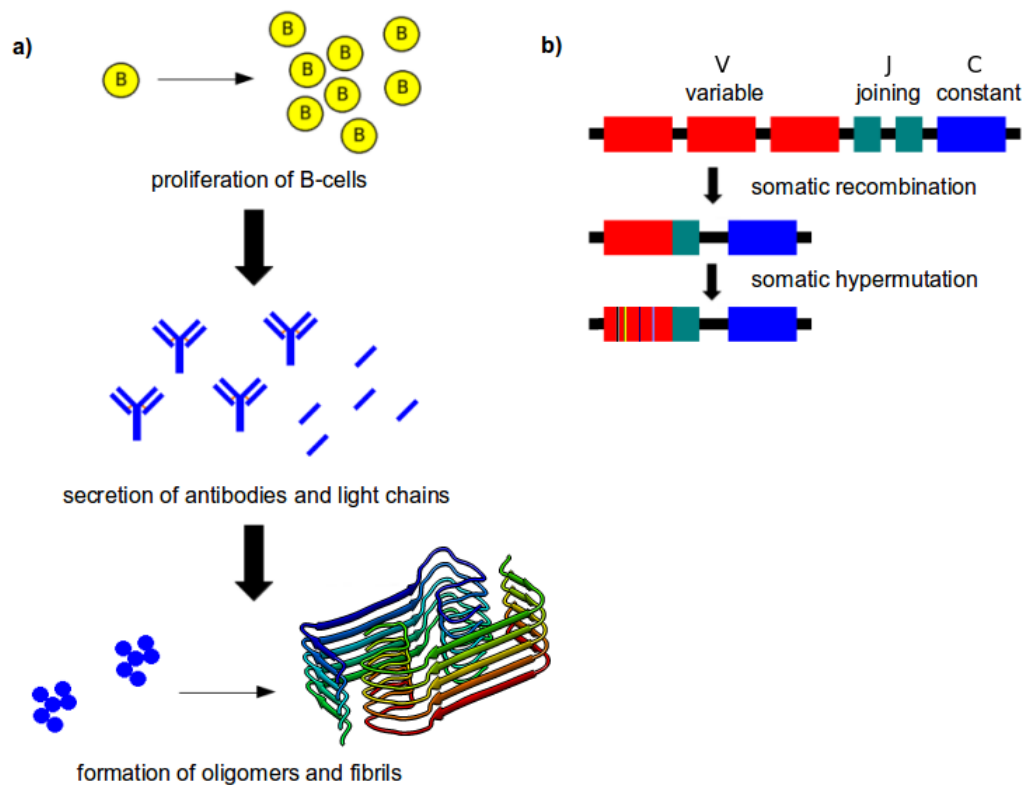
Figure 2: Development of AL Amyloidosis
     a) Origin of AL fibrils
     b) Formation of functional light chain genes during B-cell maturation

While it is of crucial importance for the adaptive immune response to generate a broad diversity of antibody sequences, this comes at the cost of accidentally producing also unstable proteins. The processes, which lead to genetic variability of LC sequences, are briefly displayed in Fig. 2b. The genetic information for immunoglobulin LCs is

encoded in V (variable), J (joining) and C (constant) segments. The copy numbers for these segments are given in Table 1 [16]. The V segment codes for the first 95-101 amino acids of the LC, the J segment for up to 13 remaining residues in the variable domain. The C segment finally contains information for the whole constant domain.

	V	J	C
λ	30	4	4
κ	40	5	1

Table 1: Number of gene segments for antibody light chains
V (variable), J (joining) and C (constant) gene segment copies in humans [16].

During somatic recombination, one copy from each segment type is randomly selected and combined to a functional LC gene. Between V and J, some base pairs might be added or removed during this process, resulting in increased sequence variety in the region corresponding to complementarity determining region 3 (CDR3). The CDRs are the loops, which are responsible for antigen binding. Later, in secundary diversification of the antibody repertoire, the sequence is subjected to random mutations, mostly in the regions encoding for the three CDRs. This mechanism is referred to as somatic hypermutation. Overall, the genetic variability generated by somatic recombination and hypermutation for both light and heavy chain leads to a repertoire of 10^{11} antibodies in humans [16].

Considering the random nature of inserted mutations, it is natural that some of the resulting proteins are unstable and likely to form oligomeric complexes and finally amyloid fibrils. These aggregates deposit in the whole circulatory system. Among the strongly affected organs are typically heart, kidneys, liver and nerves [12, 17]. Most patients die from cardiac amyloidosis [18] and mortality still remains high, with 4-years overall-survival of 42% [19].

As explained above for amyloidoses in general, also for AL amyloidosis specifically it is still a topic of ongoing research, whether the oligomeric intermediates or the mature fibrils are cause of the disease. Recent reports showed oxidative stress in cardiomyocytes caused by amyloidogenic LCs, while non-amyloidogenic sequences do not have such an effect [20]. This effect is caused by monomers, dimers and small oligomers, but not by fibrils [21]. A *Caenorhabditis elegans* model was used successfully to discriminate cardiotoxic and other LC sequences [22]. Going more into detail, LCs are internalised via clathrin-mediated endocytosis [23] and transported into lysosomes [24]. Damage of lysosomes and breakdown of autophagy lead to cell death [25].

On the other hand, also fibrils were found to be toxic against cardiomyocytes [26]. In addition to toxicity, it is also conceivable that the sheer load of amyloid deposits in the heart mechanically impairs cardiac function. This is supported by increased heart wall thickness observed in many AL amyloidosis patients. Thus, according to current knowledge, both precursors and mature fibrils might be harmful to patients. Hence, therapy should aim at reducing amounts of both species.

2.3.2 Antibody Structure and Immunoglobulin Fold

In order to discuss important characteristics of AL amyloidosis on a molecular level, it is helpful to look at the general structure of antibodies and specifically of LCs. There are five different isotypes of antibodies, called IgA, IgD, IgG, IgE and IgM. All of them share the same LC sequences and differ only be the constant domains of the heavy chain. Fig. 3 shows the architecture of an IgG antibody. The LC consists of a variable and a constant domain, V_L and C_L and the heavy chain of one variable and three constant domains, V_H and C_H1, C_H2 and C_H3. LCs are connected to heavy chains by a disulfide bond between the constant domains. Another two disulfide bonds bind both heavy chains together, these are located in the hinge region between C_H1 and C_H2.
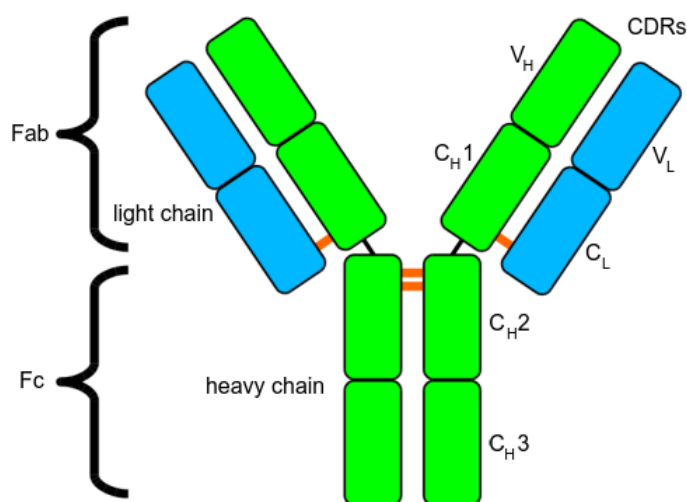
Figure 3: Architecture of Immunoglobulin G
> IgG antibodies are composed of two identical LCs and two identical heavy chains. The chains are connected via intermolecular disulfide bonds (shown in orange). The variable domains V_L and V_H form antigen binding sites and the constant domains in the Fc part are important for effector functions.

All domains of an antibody are similar. They typically comprise approximately 110 residues. Characteristic is the tertiary structure, composed of two β-sheets connected by an intramolecular disulfide bond. This fold is a common motif seen in many proteins, and is called immunoglobulin fold due to its prominence in antibodies.

In Fig. 4, the V_L domain of the antibody MAK33 is shown. MAK33 is a murine antibody of subclass κ/IgG1, targeted against muscle-type creatine kinase. The structure of the Fab fragment was solved with x-ray cristallography [27]. Since this V_L domain was used for most of the experiments presented in this thesis, it will be introduced here in more detail. The structure shows the two β-sheets ABED and C'CFG. In addition, between strands C' and D there is a loop region, for consistency referred to as C".
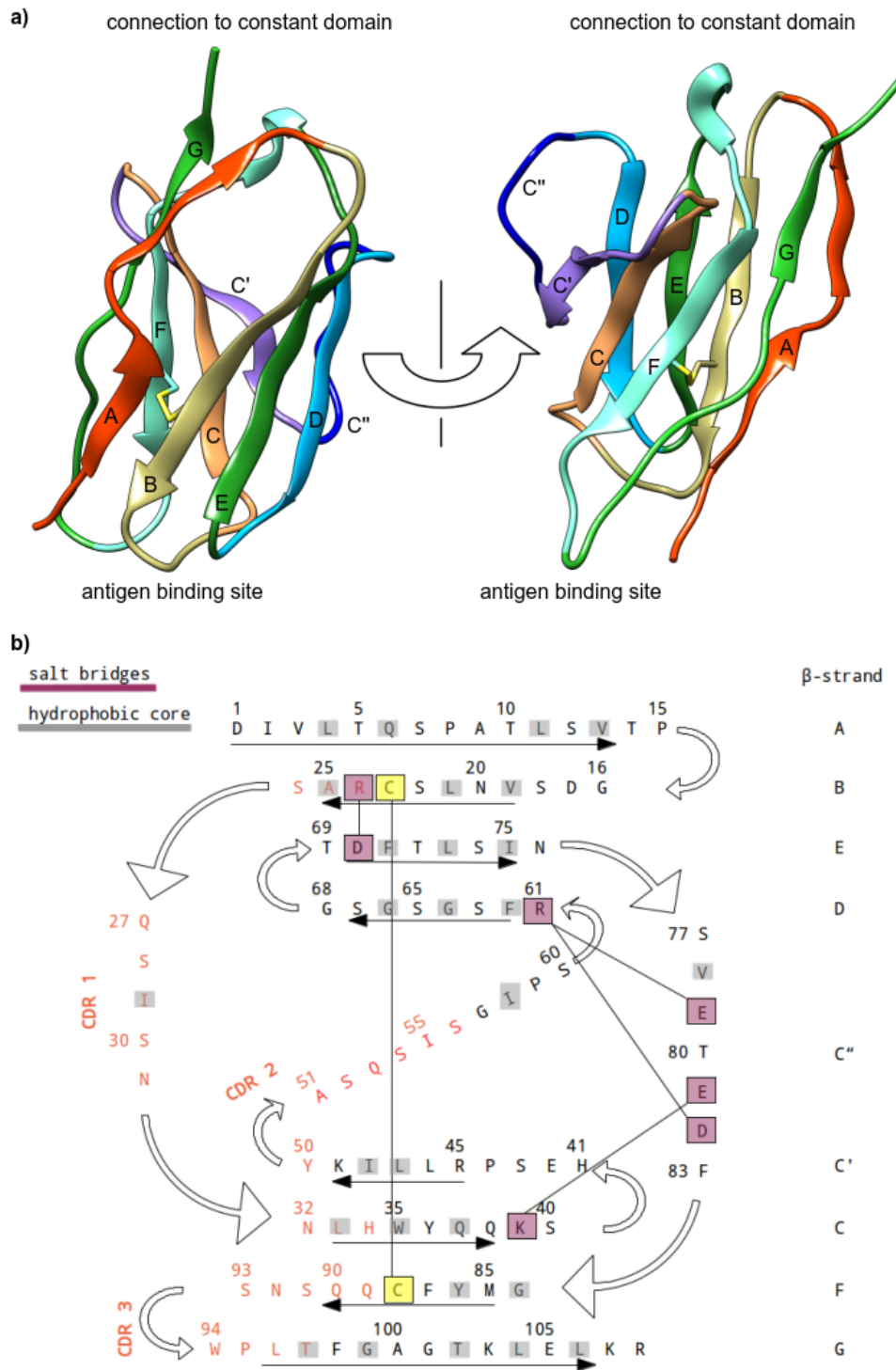
Figure 4: Structure of MAK33 V_L

a) PDB structure of MAK33 V_L [27]. In the right image, the structure is turned by 180°. β-sheet C'CFG comprises the interface for binding to the V_H domain in an intact Fab fragment.

b) Topology of MAK33 V_L

17

The sequence contains five proline residues, at positions 8, 15, 44, 59 and 95. These prolines are highly conserved in V_L domains. P15 and P95 are in cis-conformation in the native state and isomerisation is limiting for the folding rate [27, 28, 29, 30]. Once it is folded, the structure is stabilised by the conserved disulfide bond between C23 and C88 and also several salt bridges, most of which are also conserved in V_L domains. While the framework of the immunoglobulin domain is more conserved, the CDRs are highly variable in order to be able to bind different antigenes. The CDRs are located between strands B and C (CDR1), C' and C'' (CDR2) and F and G (CDR3). As can be seen in Fig. 4a), these loops are close to each other in the structure, enabling all of them to bind to one antigene.

In addition to the antigene binding site, V_L domains also have a dimerisation interface to V_H domains, as mentioned above. Due to the high similarity in all immunoglobulin domains, LCs in absence of heavy chains tend to form homodimers, albeit the K_D may vary over several orders of magnitude, depending on the specific sequence [31].

The MAK33 V_L domain is especially interesting for studying amyloid fibrils, since it can adopt an alternatively folded state (AFS). This state was first described for the whole antibody at acidic pH [32]. Each domain of the Fab part is able to form such an AFS [33]. MAK33 V_L is oligomeric in its AFS. The molecular weight of the complex varies between 50 kDa and 1 MDa, with the majority consisting of ca. 30 monomers. Far-UV CD spectra indicate mostly β-sheet structure, whereas there is hardly any near-UV CD signal. Exposed hydrophobic surface increases and the state is stable against heating and denaturants. These characteristics already resemble typical features of amyloid fibrils [34, 35].

Indeed, the MAK33 V_L domain can form fibrils at similar conditions as required for the AFS. While pH 2 is sufficient to transform to the state of unstructured oligomers, transition to fibrils also needs agitation and elevated temperatures of 37 °C [29]. Although these conditions seem rather artificial, they can be put into physiological context. Considering that fibrils *in vivo* form in lysosomes [24], where pH is around 4.5-5, pH 2 is just a step further to compensate for absence of other cellular components like membranes, which could facilitate fibril formation [36]. Agitation mimics the heart beat, as most LC fibril deposits are found in the heart and 37 °C is human body core temperature.

The advantage of the MAK33 V_L domain as model to study AL amyloidosis is the abundance of literature on biochemical and biophysical characteristics [27, 29, 32, 33, 37]. In addition, the MAK33 V_L domain is special, since the AFS can be stabilised at acidic pH, without proceeding to amyloid fibrils. In general, such oligomeric states, which might be the precursors for fibril formation, are very short lived and hence difficult to characterise [38]. Since these olgiomers come more and more into focus of research due to their aforementioned cytotoxicity and to understand fibril formation mechanistically, the MAK33 sequence offers a unique possibility to trap this state and study it in detail.

Considering that MAK33 is a murine antibody, this should not limit this study, as sequence identity is also rather low among human LC sequences. Actually, the similarity between this murine sequence and human sequences is comparable to that between different human LCs. Fig. 5 shows a sequence alignment of a diverse set of V_L domains, from human and mouse and κ and λ. Prolines or the cysteines for the disulfide bond are

highly conserved. The β-strands display moderate conservation, also between different gene families. As expected, there are many differences in the CDRs [39].
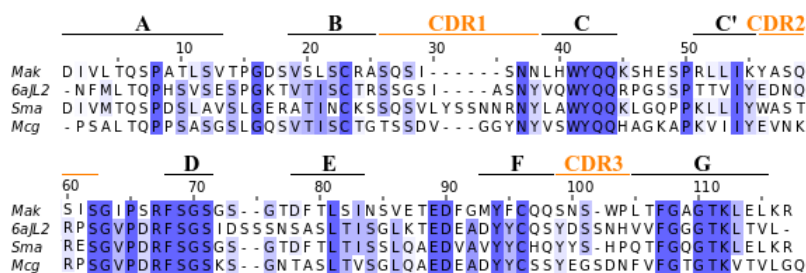

Figure 5: V$_L$ Sequence Alignments

MAK33 is a murine κ sequence [27]. 6aJL2 is a predicted human germline sequence for the highly amyloidogenic germline λ-6 [40]. Sma is a κ-4 sequence obtained from an AL patient [41, 42]. Mcg was obtained from a patient suffering from multiple myeloma and λ-AL amyloidosis [43, 44]. β-strands and CDRs are indicated above the sequences. Despite different origins (murine and human, κ and λ), all sequences share similar regions. The CDRs vary both in sequence and, especially for CDR1, in length.

2.3.3 Important Factors for Amyloidogenicity

Due to this variety of sequences, it is especially difficult to identify destabilising elements or other factors, which render the protein prone to aggregation. Still, some common characteristics could be identified. In most amyloid deposits found in patients, only the V$_L$ domain is completely present. The C$_L$ domain is typically cleaved off and found rarely in the aggregates [45, 46]. λ-type sequences are overrepresented in AL sequences, compared to gene usage in healthy humans [47]. Specific germlines, e.g. λ-6, are frequently associated with AL amyloidosis [48]. Germlines also influence, in which organs amyloids are found [49]. This tropism may be a result of varying membrane compositions in different tissues, as membranes can facilitate fibril formation [36]. Moreover, thermodynamic stability is especially important for fibril formation. Obviously, low stability can favour fibril formation, but recent results indicated, that stability can also be too low [50, 51]. This is in agreement with observations, that fibrils form from partially unfolded intermediates, but not from completely unstructured states [52, 53]. Population of such partially unstructured states can be increased by applying acidic conditions or addition of denaturants [52, 54, 55].

Many point mutations have been reported to enhance fibril formation. These substitions frequently destabilise the native state [56, 57, 58, 59]. Mechanistically more interesting are those mutations, which affect amyloidogenicity without changing stability. E.g. Tyr-to-Phe mutations could alter the dimerisation behaviour [60]. Such mutations can lead to non-canonical dimer interfaces, which enhance fibril formation (Fig. 6) [61]. This is in contrast to the canonical V$_L$ dimer, which protects against conversion to amyloids [62, 63].
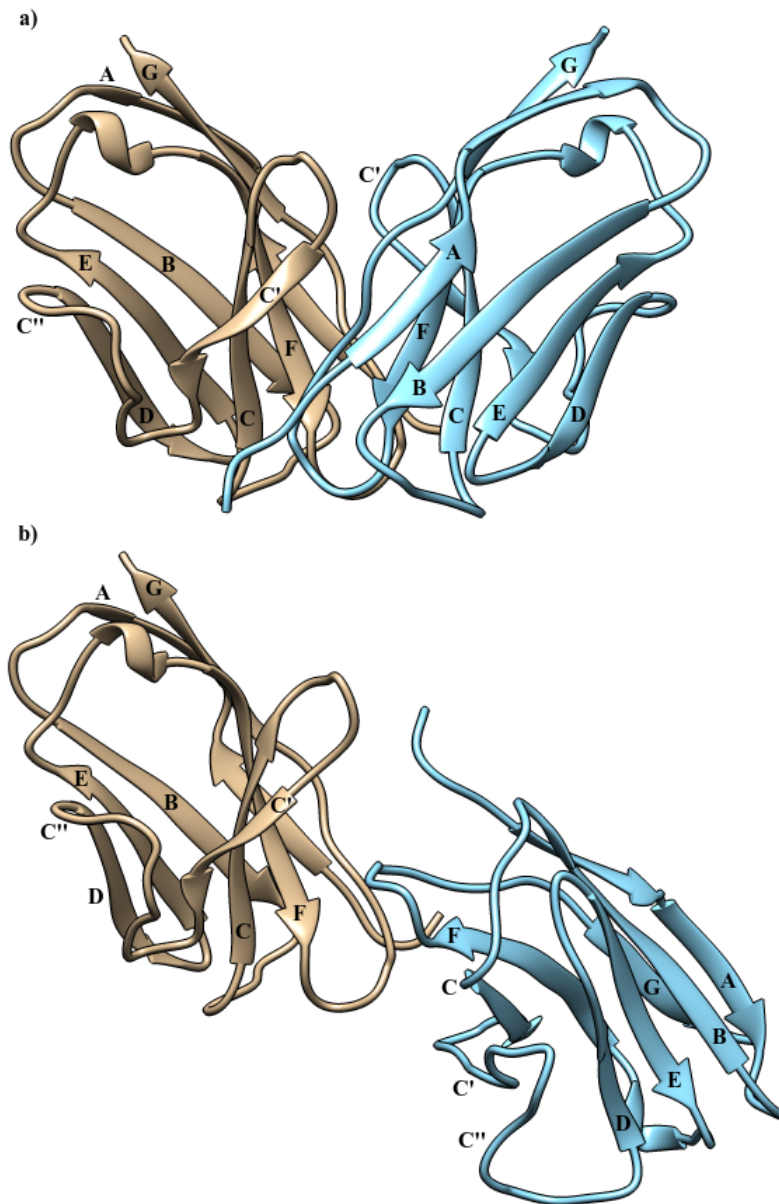
Figure 6: V$_L$ Dimer Interfaces

The κI O18/O8 V$_L$ domain and its mutant AL-09 can dimerise in different orientations.
a) Canonical dimer interface, formed by κI O18/O8 V$_L$ (PDB: 2Q20). The orientation of the monomers corresponds to the heterodimer between V$_L$ and V$_H$ in a native antibody.
b) Non-canonical dimer interface, formed by AL-09 V$_L$ (PDB: 2Q1E). One monomer is rotated by 90 °C. This interface enhances fibril formation [61].

Another factor to be considered is the structure of the edge-strands in the immunoglobulin fold. β-sheet proteins are naturally prone to aggregate by association of β-strands from different monomers. In many proteins, such behaviour is prevented by negative design, e.g. kinks in the strands at the edge of a β-sheet or introduction of charged residues [64]. In this respect, the conserved proline in the N-terminal strand protects against amyloid aggregation [65].

Amyloidogenic LCs are frequently subject to post-translational modifications. In addition to the cleavage of the constant domain, disulfide-linked dimerisation and glycosylation have been observed [66]. Introduction of N-linked glycosylation sites by mutations is a common feature of many studied LC deposits [67].

Despite much effort by the scientific community to identify characteristics, which render an LC prone to amyloid aggregation, no unique feature was found yet. It is more likely, that many different factors contribute to amyloidogenicity. This diversity both of sequences and of causes for instability is a hallmark of AL amyloidosis, which also presents obstacles for targeted therapies.

2.3.4 Therapies for Immunoglobulin Light Chain Amyloidosis

Within the last decades, overall-survival of AL amyloidosis patients improved steadily. Still, mortality remains high, with a four-years overall-survival of 42% [19]. Early diagnosis is crucial for successful therapies [18, 68]. However, as AL amyloidosis is a rare disease and difficult to distinguish from other cardiac diseases, proper diagnosis often requires a specialist. Typical methods for identification of AL amyloidosis are echocardiography [69], cardiac magnetic resonance [70] and congo red stain of biopsies [71, 72]. Efficacy of treatment can be evaluated by using biomarkers like serum free LCs [73, 74, 75], troponin T [76] or natriuretic peptides [77].

Current treatment typically involves chemotherapy with high-dose melphalan, in combination with hematopoietic stem cell transplantation [78, 79]. Bortezomib is also a promising therapeutic agent [80, 81]. All these therapeutic options target the underlying plasma cell dyscrasia. However, they may be poorly tolerated, causing severe side-effects [82]. In addition, these therapies do not affect already secreted serum free LCs or amyloid deposits.

Direct targeting of the amyloid deposits or the serum free LCs is difficult, due to the heterogeneity of LC sequences. Recently, a very promising and elegant approach used small molecules against serum amyloid P component (SAP) [83]. SAP is a protein, which deposits together with all amyloidogenic proteins. As the aggregates are stabilised by SAP, this protein is a universal target for therapy of amyloidoses [84]. Depletion of SAP by therapeutic antibodies [85] or the small molecule (R)-1-[6-[(R)-2-carboxy-pyrrolidin-1-yl]-6-oxo-hexanoyl]pyrrolidine-2-carboxylic acid (CPHPC) [86] destabilises amyloid aggregates and promotes clearance *in vivo*. Another promising therapeutic approach involves upregulation of the unfolded protein response [87]. Both strategies, depletion of SAP and upregulation of unfolded protein response, make use of indirect effects to prevent aggregation of immunoglobulin LC. This renders therapy independent from individual LC sequences and makes it applicable to a broad range of patients.

It is obvious, that drugs for AL amyloidosis must rely on universal features of the amyloidogenic LC in order to be useful for patients without knowledge about the specific LC sequence. The small molecules methylene blue and sulfasalazine also make use of such a feature, specifically the monomer-dimer equilibrium of V_L domains. As amyloid formation starts from the monomers [62, 63], these compounds shift the equilibrium towards the amyloid resistant dimer [88].

2.3.5 Epigallocatechin Gallate as Drug Candidate for Immunoglobulin Light Chain Amyloidosis

Epigallocatechin gallate (EGCG) has been studied in context of many types of amyloidosis. It is likely, that EGCG also makes use of universal features of amyloid deposits, as it was shown to interact with aggregates from a diverse group of proteins and peptides. The list of amyloidogenic proteins interacting with EGCG comprises α synuclein [89], amyloid-β [90], huntingtin [91], islet amyloid precursor protein [92], transthyretin [93], tau [94], human serum albumin [95], semen derived enhancer of viral infection (SEVI) [96] and chicken cystatin [97].
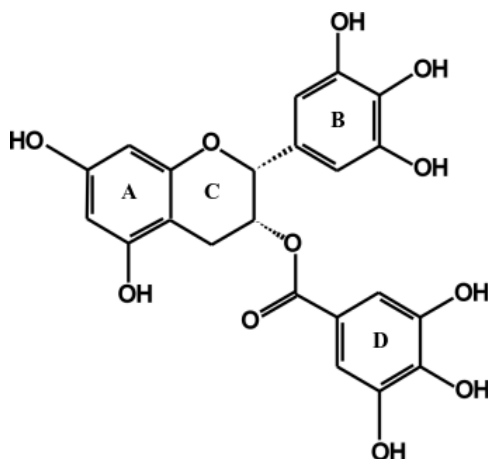
Figure 7: Epigallocateching Gallate (EGCG)

Chemically, EGCG belongs to the group of polyphenols (Fig. 7). It is a secundary metabolite in plants and found in large amounts in green tea, which is made from *Camellia sinensis*. In addition to its activity with regards to amyloid proteins, it is also assumed to have plenty of other beneficial effects on health [98]. EGCG is reported to protect against cancer, improve cardiovascular health, enhance weight loss and protect against damage caused by UV light [99]. Most of these effects are attributed to its antioxidant properties [100]. At the same time, EGCG also exerts prooxidant properties, which promote apoptosis and might act anticarcinogenic [101]. Other assumed biological activities include prevention of protein glycation in the eye lens [102] or antibacterial properties mediated by inhibition of fatty acid synthesis and damage to the cell wall [103].

Albeit only few of these effects are confirmed by clinical trials, EGCG has become a popular dietary supplement. It is not only consumed in pills, but even added to some dentrifices [104]. Such widespread consumption of EGCG makes it an intriguing candidate for therapeutic research. Not only could it be used at a high dose for treatment of a specific disease, but it might already exert beneficial actions due to daily consumption, before pathologic effects become obvious. Since AL amyloidosis develops slowly and early diagnosis is still a major problem [105], it would be especially interesting to learn about protective activity with regards to AL amyloidosis even before diagnosis. Pathologic increase of clonal serum free LCs could be measured already more than ten years before diagnosis of AL amyloidosis [75]. If EGCG protects against AL amyloidosis, routine ingestion of green tea could interfere with development of the disease many years before a targeted therapy could start.

First anecdotal evidence for activity of EGCG with respect to AL amyloidosis came from Werner Hunstein [106]. Hunstein, himself a hematologist in Heidelberg, was diagnosed with AL amyloidosis in 2004, when he was 75 years old. Chemotherapy stopped the progression of the disorder, but had to be interrupted in 2006 due to severe side effects. At this time, Hunstein started to drink 1.5 to 2 L green tea per day, encouraged by findings from Erich Wanker on *in vitro* interaction of EGCG with α-synuclein and amyloid-β [107]. Without further chemotherapy, intraventricular cardiac septum width, a biomarker for AL amyloidosis, improved. Hunstein continued to drink green tea and did not receive other treatment for AL amyloidosis, until he passed away in 2012, at the age of 83 years. Werner Hunstein's personal narration can be read online at http://www.hunstein-egcg.de/tven.html [108].

In the meantime, improvement of cardiac biomarkers upon comsumption of EGCG has been reported for a series of patients [109]. Two phase II clinical trials are ongoing (TAME-AL [110], EpiCardiAL [111]).

On a molecular level, some insights can be gained from EGCG interactions with other amyloid proteins. Unstructured monomers of α-synuclein or amyloid-β are redirected into stable, non-toxic and non-amyloid aggregates [107]. Mature fibrils formed from these proteins can also be dissolved into non-toxic oligomers by addition of EGCG [112].

In vivo experiments with *C. elegans* showed protective effects of EGCG concerning toxicity of amyloidogenic LCs [22]. *In vitro* experiments with the λ-type germline sequence 6aJL2 proved specific binding of EGCG and inhibition of fibril formation at substochiometric amounts [113]. In combination, all these data make EGCG a promising candidate to interfere with AL amyloidosis, which could react with amyloidogenic LCs in general, irrespective of the specific primary sequence.

2.4 Nuclear Magnetic Resonance Spectroscopy in Structural Biology

Mechanistic insights into misfolding of amyloid proteins as well as ways to interfere with such events require structural information about the respective proteins. Nuclear magnetic resonance (NMR) spectroscopy is a well suited method for such endeavours. The combination of solid- and solution-state NMR spectroscopy is unique in its ability to elucidate structures of both aggregated and soluble complexes. In addition, NMR spectroscopy yields information on dynamics and can be used to identify specific interactions. This makes NMR spectroscopy our method of choice to elucidate AL amyloidosis and related activity of EGCG on a molecular level.

NMR spectroscopy has evolved to a highly sophisticated method. Many research groups are commited to further improvement of the underlaying theories and new or better applications. As this study was about application of NMR spectroscopy, the next sections will emphasise obervable parameters and their meanings. In particular, the two most obvious parameters obtained from NMR experiments will be discussed: chemical shift and linewidth.

2.4.1 Chemical Shift

To start at the very basics, NMR resonance frequency is directly related to the energy gap between the ground state and the excited state of the respective nuclear spin:

$$\nu = \frac{\Delta E}{h}$$

with ν being the NMR resonance frequency, ΔE being the difference between the energy levels and h being Planck's contant [114]. Because ν depends on the external magnetic field B_0, it is converted into the field independent chemical shift:

$$\delta_{ppm} = 10^6 \times \frac{\nu - \nu_{reference}}{\nu_{reference}}$$

with δ_{ppm} being the chemical shift in parts per million [114].

Chemical shifts correspond to the local magnetic field experienced by a specific nucleus, relative to the average external field B_0. Such local differences are caused by shielding or deshielding factors from the chemical environment. A simple example for deshielding of a nucleus is a hydrogen atom bound to an electronegative atom, which reduces electron density at the hydrogen. Thus, the hydrogen is less shielded and experiences a stronger magnetic field.

In proteins, the ^{1}H chemical shift typically results from a combination of electric fields due to charges, aromatic ring currents and magnetic anisotropy around bonds [115]. Partial or formal electric charges contribute considerably to the chemical shifts of atoms in different functional groups. With regards to deviations from average chemical shifts in proteins, however, their contributions are limited. This is a consequence of opposite charges attracting each other and thus frequently canceling out each other. In addition,

the electric field is attenuated by the dielectric constant, which is large in water [115]. Ring currents are more important for deviations from average chemical shifts [115]. This makes ring currents a sensitive tool for identification of ligand binding sites, as the majority of drugs contains at least one aromatic ring system [116]. Anisotropic contributions are especially important for backbone amide ^{1}H, which are close to the anisotropic amide and carbonyl groups.

This brings up two further topics: Chemical shift anisotropy and secondary chemical shifts. The chemical shift observed in solution-state NMR spectroscopy is the isotropic chemical shift. It is the average of the principal values of the anisotropic chemical shift [117]:

$$\delta^{iso} = \frac{1}{3} \times (\delta_{xx} + \delta_{yy} + \delta_{zz})$$

In solution, the chemical shift anisotropy cannot be observed due to molecular tumbling. However, the anisotropy is still present and is for example crucial for transverse relaxation-optimised spectroscopy (TROSY) [118]. In the solid state, chemical shift anisotropy is removed by magic angle spinning (MAS) [119, 120]. This method removes all contributions from anisotropic second rank tensors in the interaction hamiltonians [121]. Still, if the chemical shift anisotropy is measured with special experiments, it is a useful parameter for structure calculation [122].

The isotropic chemical shift still contains information on the anisotropy of amide and carbonyl groups. These depend on the orientation of the protein backbone elements and consequently contain information on the secundary structure [123]. Hence, chemical shifts can be used to predict backbone torsion angles, e.g. with TALOS+ [124] and protein structures, e.g. with CS-ROSETTA [125].

For such analyses, chemical shifts need to be assigned to the respective atoms in the protein. In solution, assignment is mostly accomplished using well established 3D heteronuclear correlation experiments. These correlate amide ^{1}H and amide ^{15}N with a ^{13}C dimension. The latter is a selection of ^{13}Cα, ^{13}Cβ or ^{13}CO of either one or two consecutive amino acids [126]. The ^{1}H dimension is used for detection due to high sensitivity, ^{15}N is used as additional dimension to increase dispersion and ^{13}C dimensions are used to establish sequential contacts. Residue types can then be identified using average chemical shifts, especially the characteristic side chain chemical shift of ^{13}Cβ [127]. Once backbone resonances are assigned, side chain assignment can be completed using TOCSY-based experiments, which correlate amide ^{1}H and ^{15}N with all ^{1}H or ^{13}C side chain resonances, respectively [126]. In solid-state NMR spectroscopy, up to date there is no standard method for resonance assignment.

As a final remark to this section, the precise chemical shift of a resonance is not always easy to determine. Spectral overlap often hampers determination of the chemical shift. In order to resolve this problem, scientists exploit multidimensional experiments, use higher magnetic fields or design experiments, which improve spectral resolution. The relation between resolution and determination of chemical shift becomes obvious with so-called "pure shift" NMR experiments [128]. These remove all multiplets otherwise resulting from scalar coupling, thereby dramatically reducing spectral overlap.

2.4.2 Linewidth

This leads to the second parameter, which can be obtained directly from each NMR spectrum: Linewidths of the resonances, which is intrinsically linked to spectral resolution (Fig. 8). The lineshape of an NMR resonance is the real part or absorptive part of a complex Lorentzian:

$$A(\lambda, \omega, \omega^0) = Re\{L(\lambda, \omega, \omega^0)\} = \frac{\lambda}{\lambda^2 + (\omega - \omega^0)}$$

with λ being the coherence decay rate constant, ω being the frequency and ω^0 being the resonance frequency [117].
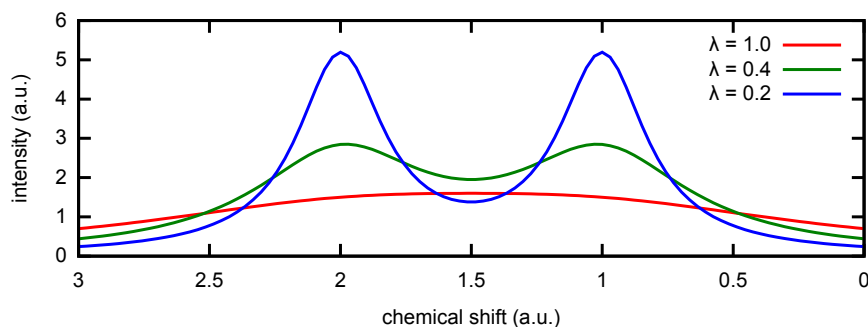
Figure 8: Separation of Resonances at Different Linewidths
The sum of two Lorentzians is plotted at different λ values. In all cases, the chemical shifts of the two resonances are 1 and 2, respectively. At $\lambda = 1$, the resonances cannot be resolved at all. With lower λ values, separate resonances become visible, although baseline separation is still not reached at $\lambda = 0.2$. As the integral of the Lorentzians is constant, smaller linewdiths also result in higher signal intensities.

The linewidth depends on λ, which is defined as the inverse of the transverse relaxation time T_2. This relaxation time constant itself is governed by local fluctuations in the magnetic field, with important contributions of the spectral density function at J(0), J(ω^0) and J($2\omega^0$):

$$T_2^{-1} = \frac{3}{20}b^2 \left\{3J(0) + 5J(\omega^0) + 2J(2\omega^0)\right\}$$

with b being the dipole-dipole coupling constant and J being the spectral density of the dipole-dipole coupling [117]. Longer tumbling correlation times (corresponding to larger molecules in solution) result in a narrower spectral density function, which leads to shorter T_2 times (Fig. 9).

In solution, molecular tumbling is the main source of local fluctuations, hence the correlation between molecular weight and T_2 time. In solids, rotational diffusion is at least restricted, if not absent. Thus, local dynamics within the molecule are the cause of fluctuations in the magnetic field and determine the spectral density function. This, in turn, means, that relaxation and thus, also linewidth, is already a reporter on local dynamics.
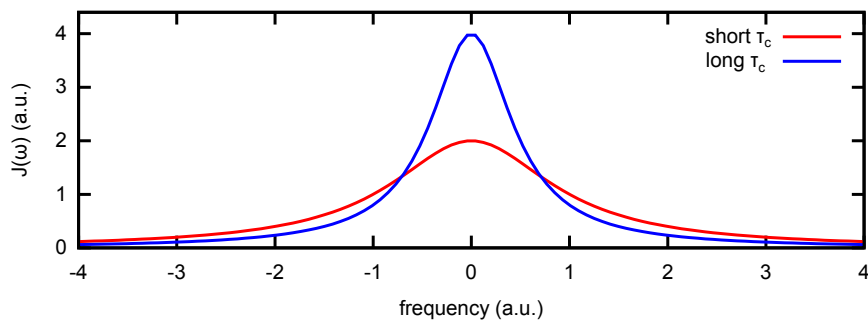
Figure 9: Spectral Density Functions at Long and Short Tumbling Correlation Times
With longer tumbling correlation time, the spectral density function becomes narrower, resulting in larger contributions at J(0).

T_2 determines the optimal linewidth, as explained above. In order to achieve this linewidth, long acquisition times are necessary. However, experimental restraints can limit the acquisition time. The main limitations for long evolution periods are sensitivity and sample heating due to high-power decoupling during acquisition. Concerning the relation between acquisition time and sensitivity, it helps to visualise these parameters in a free induction decay (FID) (Fig. 10). At short acquisition times, the signal is high, but longer times improve resolution. However, the picture is not complete yet, as longer acquisition times result in smaller linewidths, which also increases intensity, as was shown before (Fig. 8). Corrected for the line narrowing, the sensitivity as a function of acquisition time is

$$S/N \propto \frac{T_2}{\sqrt{t}} \left(1 - e^{-\frac{t}{T_2}} \right)$$

The function has a rather broad maximum at 1.26 T_2 times (Fig. 11) [129].
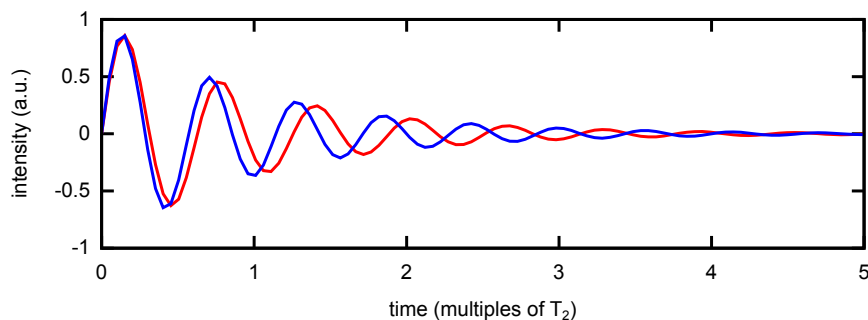
Figure 10: Free Induction Decay of Two Spins
At short acquisition times, the signal is high, which contributes to a high signal-to-noise ratio (S/N). However, a phase difference between both spins evolves only at long acquisition times. Thus, long acquisition times are necessary to resolve the resonance frequencies.

Sensitivity decreases only slowly, if acquisition times are extended beyond this maximum. Hence, in the direct dimension, it is good practice to use long acquisition times
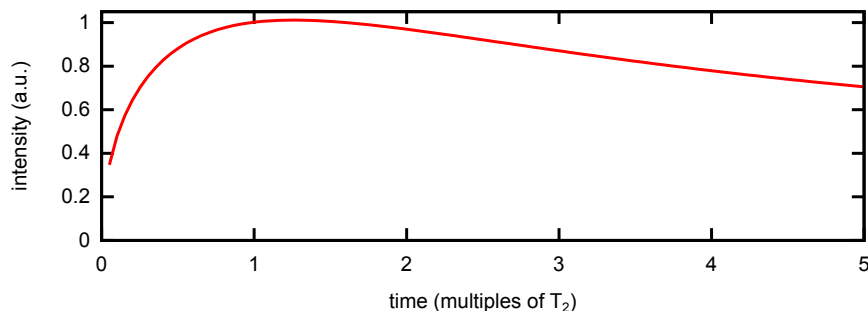
Figure 11: Sensitivity of 1D NMR Experiments as a Function of Acquisition Time
With increasing acquisition times, sensitivity increases due to enhanced resolution. At the same time, transversal relaxation reduces the sensitivity. The optimum is at ca. 1.26 T_2 [129].

and aim for optimal resolution. In the indirect dimensions, however, shorter acquisition times might be reasonable in case of insensitive experiments or bad sample quality. This results from the additional measurement time required for long indirect evolution periods. While smaller linewidths increase the signal intensity, increasing the number of repetitions of the experiment often yields a larger gain. Thus, the objective in most experiments is to obtain enough signal so that all resonances can be observed and then go for optimal resolution.

The trade-off between sensitivity and resolution can be circumvented by use of nonuniform sampling (NUS). This method changes the actual sampling scheme, measuring only a fraction of the typically measured points in the indirect evolution periods (Fig. 12) [130]. Many samples at short evolution periods benefit from high S/N, while few samples after long acquisition times allow for higher spectral resolution. Iterative Fourier transforms, thresholding of spectra and inverse Fourier transforms are applied to reconstruct the missing data points in the FID [131, 132]. The saved experimental time due to less increments in the indirect dimension can be used to measure more transients for signal averaging, thereby improving sensitivity [133, 134].

Linewidth, however, is not only limited by evolution time and T_2. Several interactions of the spins can also cause line broadening. Couplings between spins produce splitting of resonances. If this splitting is not resolved, it just appears as line broadening (Fig. 13). In solution state, the most important interaction leading to such splittings is scalar coupling. In order to prevent evolution of heteronuclear scalar coupling during acquisition, π-pulses are used in indirect dimensions to refocus coupling evolution. This refocussing is a practical application of the spin echo experiment [135]. Heteronuclear coupling evolution during the direct acquisition period is prevented by heteronuclear composite pulse decoupling [136]. Homonuclear scalar coupling evolution can be avoided by use of constant-time evolution in indirect dimensions [137]. In direct dimensions, homonuclear scalar coupling evolution could be avoided by decoupling and stroboscopic detection [138].

In the solid state, dipolar couplings evolve in addition to scalar couplings. Dipolar couplings are much stronger than scalar couplings. Typical coupling values are given in
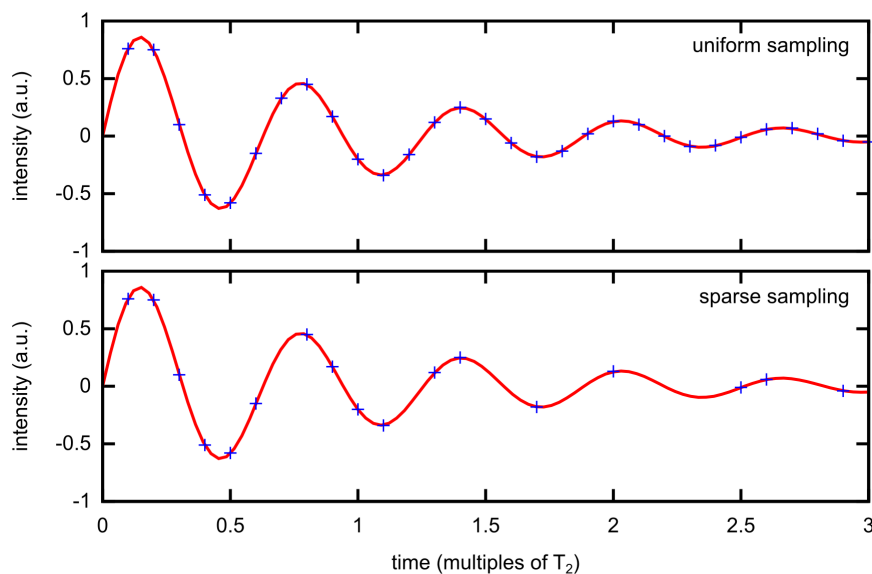
Figure 12: Sensitivity Enhancement by Sparse Sampling in Indirect Dimensions
In the upper panel, the FID is sampled uniformly. In the lower panel, only a fraction of the uniform grid is sampled, with higher sampling density at short evolution times. Thus, high sensitivity of short acquisition periods can be combined with resolution of long acquisition times, if the uniform FID can be reconstructed.
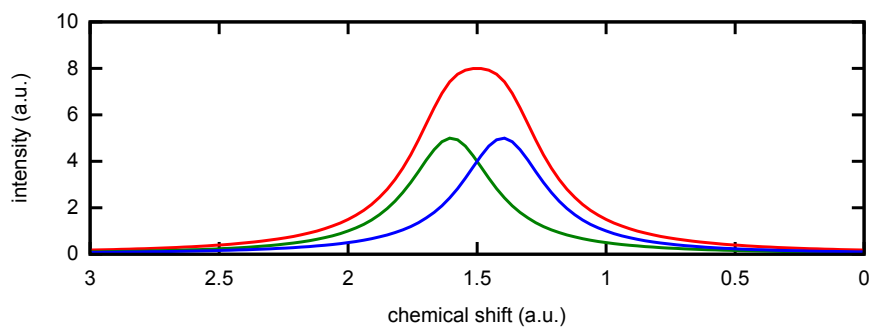
Figure 13: Resonance Broadening due to Unresolved Splitting
If the resolution is not sufficient to separate resonances resulting from splitting due to coupling, the observed sum of the resonances appears as one resonance. The observed linewidth is broader than the actual linewidth of the individual signals.

table 2 [139]. With such large couplings, ^{1}H linewidths would exceed the spectral width of typcial experiments and resonances would be broadened beyond detection. Since the dipolar coupling is an anisotropic interaction, it can be partially removed by magic angle spinning. As magic angle spinning needs to be considerably faster than the interaction it should remove, dipolar couplings are not averaged completely. Up to the present, the fastest spinning frequency achieved is approximately 110 kHz [140, 141]. Increasing the spinning frequency from 50 kHz to 100 kHz reduced ^{1}H linewidth by a factor of two [141].

Before such ultrafast MAS rotors for frequencies of 60 kHz or more were developped, biogical solid-state NMR spectroscopy relied heavily on detection of heteronuclei. In general, ^{1}H detection yields better sensitivity due to the larger gyromagnetic ratio. However, ^{1}H detection was not popular with standard samples at MAS rates of up to 20 kHz due to the strong dipolar interactions of protons. As ultrafast spinning was not available, isotope dilution was applied to reduce dipolar interactions. The idea is to avoid the presence of several ^{1}H atoms next to each other in order to remove the strongest dipolar couplings. Deuteration made detection of residual protons feasable [142, 143]. Protonation of the exchangeable amide hydrogens enabled measurement of ^{1}H-^{1}H distances [144]. Methyl protons become accessible by selective protonation [145] and other aliphatic protons by reduced adjoining protonation (RAP) [146]. Despite the smaller number of detectable protons, the actual S/N ratio improved with these approaches, due to dramatic reductions in linewidths. Deuteration for enhancing ^{1}H linewidth was succesfully applied for structure determinations [147, 148].

^{1}H	-	^{1}H	CH_3 group	60 kHz
^{13}C	-	^{1}H	directly bonded	23 kHz
^{15}N	-	^{1}H	directly bonded	11 kHz
^{13}C	-	^{13}C	directly bonded	3 kHz
^{13}C	-	^{15}N	directly bonded	1 kHz

Table 2: Dipolar Coupling Values
Coupling values are taken from reference [139].

Similar to the deuteration approaches, also ^{13}C linewidths could be reduced by spin dilution. Selective labeling with 2-^{13}C$_1$ glycerol or 1,3-^{13}C$_2$ glycerol as nutrients avoids labeling of adjacent carbon atoms. Thus, scalar couplings can be removed [149, 150]. Labeling with 1-^{13}C$_1$ glucose or 2-^{13}C$_1$ glucose as carbon sources, which also relies on removal of scalar couplings, reduces linewidths by a factor of two [151]. Selective labeling with glycerol as nutrient was used to solve the first MAS solid-state NMR structure [152], as well as several others [148, 153]. These structure determinations exploited another benefit of selective labeling: Not only were scalar couplings removed, but also strong dipolar couplings of close neighbours. Thus, isotope dilution can circumvent dipolar truncation, where interesting long-range dipolar transfers are not observed due to much stronger short-range transfers.

Besides T_2 time and coupling interactions, linewidth can also be affected by exchange dynamics. This can occur for one molecule undergoing conformational changes or for

ligand binding experiments. As explained above, such changes are likely to change chemical shifts. If the exchange rate is similar to the difference in resonance frequencies of both states, signals are broadened [154]. This arises from one spin experiencing different chemical environments within the experiment. In some cases, such line broadening can be prevented by changing the dynamics of the interaction, e.g. by changing temperature, or measuring at a different field strength. Although exchange broadening hampers analysis of NMR data, it can also be used to learn about the dynamics and the different states. This can be accomplished by use relaxation dispersion experiments, based on the Carr-Purcell-Meiboom-Gill sequence [155, 156, 157].

The last two sections explained the meaning, potential uses and some experimental strategies concerning chemical shift and linewidth. These concepts outlined directions, how to gain information on biological systems by means of NMR spectroscopy.

2.5 Scope of this Study

NMR spectroscopy can be used for research of both soluble proteins and insoluble aggregates. Therefore, it is the method of choice for structural analysis of amyloid proteins, which are found in both states in biologically relevant contexts. The goal of this thesis is to expand the knowledge on AL amyloidosis in several fields:

Analysis of the native state of the MAK33 V_L domain and structural characterisation of amyloidogenic mutants

Many different mutants or experimental conditions have been reported to enhance amyloidogenicity. The explanations for this are manifold, e.g. destabilisation of the native state, change of intrinsic dynamics or sampling of alternative conformations. We aim to characterise the effects of specific point mutations on the native V_L domain structure. Thus, we want to rationalise their effects on thermodynamic parameters and misfolding propensity.

Characterisation of EGCG as a small molecule compound targeted at AL amyloidosis

The goal is to understand the interaction between EGCG and V_L domains. Objectives are the identification of the binding site, the mode of action and the specifity. We will conduct these analyses with the model system MAK33 V_L. Subsequently, we will study, whether the results can be transferred also to other amyloidogenic V_L sequences.

Structural characterisation of the fibril and oligomer states of MAK33 V_L

Up to date, there is hardly any structural data available on V_L fibrils. However, such information is essential to elucidate the mechanism, by which native V_L domains transform to amyloid deposits. We aim to study MAK33 V_L fibrils by MAS solid-state NMR spectroscopy. This will enable the identification of the hydrophobic core residues. Complementary NMR spectroscopic experiments based on scalar couplings can be used to identify flexible parts in the fibrils. In addition, we will expand these analyses to the MAK33 V_L oligomers, which might be an intermediate on the pathway to amyloid fibrils. This could not only provide information on the actual fibril structure, but also on the process of their formation.

3 Material and Methods

3.1 Materials

3.1.1 Chemicals

Chemicals were supplied by Carl Roth (Karlsruhe, Germany), SERVA Electrophoresis (Heidelberg, Germany), VWR International (Ismaning, Germany), Roche Diagnostics (Unterhaching, Germany) and Sigma-Aldrich Chemie (Taufkirchen, Germany), unless stated otherwise. Epigallocatechin gallate (Sunphenon EGCG) was a kind gift by Stefan Schönland, supplied by Taiyo (Yokkaichi, Japan). $^{15}NH_4Cl$, U-^{13}C D-glucose and D_2O were purchased from Cambridge Isotope Laboratories (Tewksbury, USA) and Eurisotop (Saint-Aubin, France). 2-$^{13}C_1$ glycerol, 1,3-$^{13}C_2$ glycerol and $NaH^{13}CO_3$ were supplied by Sigma-Aldrich Chemie (Taufkirchen, Germany).

3.1.2 Equipment

Äkta Purifier and Äkta Prime systems for liquid chromatography were purchased from GE Healthcare (Munich, Germany). The used columns comprise HiLoad 16/10 Q Sepharose High Performance, HiLoad 16/600 Superdex 75 prep grade and HiLoad 26/600 Superdex 75 prep grade, all produced by GE Healthcare.

NMR spectrometers and associated equipment were manufactured by Bruker BioSpin (Rheinstetten, Germany). The rotor filling tool for the ultracentrifuge was provided by Giotto Biotech (Florence, Italy) [158].

3.1.3 Software

NMR experiments were recorded using TopSpin 2.1 or 3.2 (Bruker BioSpin, Rheinstetten, Germany). Uniformly sampled NMR spectra were processed using TopSpin 3.2. NUS spectra were processed both with TopSpin 3.2 (including the NUS plugin) or with hmsIST [132] in combination with NMRPipe [159] und GNU Parallel [160]. Spectra were analysed using CcpNmr analysis 2.4.0 [161].

PDB structures were analysed using UCSF Chimera [162]. Protein sequence alignments were calculated in Jalview [163] using Clustal W [164]. Amino acid frequency plots were created with WebLogo 2.8.2, using the frequency plot option [165]. Physical formulae were fitted with Matlab 8.4 R2014b (MathWorks, Ismaning, Germany). Data were plotted with Gnuplot 4.4 [166].

3.1.4 Bacterial Strains and Plasmids

The pET28a(+) vector and *Escherichia coli* BL21 (DE3) cells were supplied by Novagen (Merck, Darmstadt, Germany). *E. coli* XL1-Blue cells were purchased from Stratagene (Agilent Technologies Deutschland GmbH, Waldbronn, Germany).

3.1.5 Bacterial Growth Media

Lysogeny Broth (LB) was used for expression of unlabeled proteins, minipreps or precultures. LB medium was autoclaved prior to use. M9 medium was supplemented with $^{15}NH_4Cl$ for ^{15}N labeling and u-^{13}C D-glucose for ^{13}C isotope labeling, respectively. Otherwise, unlabeled NH_4Cl and D-glucose were added to the media. In case of selective labeling with glycerol, 2 g/l glucose were substituted with 2 g/l of 2-$^{13}C_1$ glycerol and 2 g/l $NaH^{13}CO_3$ or with 2 g/l of 1,3-$^{13}C_2$ glycerol and 2 g/l $NaH^{12}CO_3$. Isotopes, biotin and thiamin-HCl were prepared freshly and sterile filtered. All other M9 components were autoclaved prior to use.

LB medium	
10 g	tryptone
10 g	NaCl
5 g	yeast extract
ad 1 l	water

10x M9 salts	
60 g	Na_2HPO_4
30 g	KH_2PO_4
5 g	NaCl
ad 1 l	water

M9 medium	
100 ml	10x M9 salts
10 ml	100x trace elements
1 ml	1M $MgSO_4$
0.3 ml	1M $CaCl_2$
10 ml	biotin (0.1 mg/ml)
1 ml	thiamin-HCl (1 mg/ml)
10 ml	glucose (20% w/v)
5 ml	NH_4Cl (10% w/v)
ad 1 l	water

100x M9 trace elements	
5g	EDTA
0.83 g	$FeCl_3 \cdot 6\ H_2O$
84 mg	$ZnCl_2$
13 mg	$CuCl_2 \cdot 2\ H_2O$
10 mg	$CoCl_2 \cdot 6\ H_2O$
10 mg	H_3BO_3
1.6 mg	$MnCl_2 \cdot 6\ H_2O$
ad 1 l	water

3.2 Methods

3.2.1 Molecular Biology

Plasmid Purification

Plasmids were purified from *E. coli* XL1 LB overnight cultures using the Wizard Plus SV Miniprep kit (Promega, Mannheim, Germany). The DNA was sequenced by GATC Biotech (Konstanz, Germany) and stored at -20 °C. Pure plasmid concentration was determined by measuring the UV absorbance at 260 nm.

Site Directed Mutagenesis

Mutagenesis primers for point mutations or deletions of single amino acids were designed by adjusting the DNA sequence to the aimed mutation and considering codon bias [167]. Upstream and downstream of the mutation site, 20 nucleotides each were added to complete the primer. The mutagenesis protocol was adapted from the QuikChange site-directed mutagenesis kit from Stratagene (Santa Clara, USA). Polymerase chain reactions (PCR) were set up with the following mixture:

	PCR reaction
2.5 μl	Pfu Ultra Buffer (10x)
200 ng	template
0.5 μl	0.5 μM forward primer
0.5 μl	0.5 μM reverse primer
0.5 μl	10 mM dNTP
0.5 μl	Pfu Ultra DNA polymerase
ad 25μl	water

PCR comprised 20 cycles of:

1 min	94 °C
1 min	56 °C
2 min / kb	68 °C

After PCR, the template was digested by adding 1 μl Dpn I and incubating for 2h at 37 °C. 5 μl of the product were used for transformations.

Transformation

A 100 μl aliquot of cells was thawed on ice. Around 20 ng of the plasmid DNA were added to the cells and mixed gently. The mixture was incubated for 30 min on ice, heat shocked for 40 s at 42 °C and immediately incubated on ice for 5 min. 1 ml of LB was added to the cells, followed by incubation for 1 h at 37 °C with 1000 rpm. Cells were pelleted, supernatant was discarded and cells resuspended in the remaining liquid. This was spread on LB agar plates containing kanamycin (50 μg/ml). Cells were allowed to grow overnight at 37 °C.

3.2.2 Recombinant Gene Expression and Protein Purification

Precultures and Main Cultures

All V_L and C_L variants were cloned into pET28a(+) containing a gene for kanamycin resistance. Precultures of 100 ml LB_{Kan} or unlabeled $M9_{Kan}$ were inoculated from LB_{Kan} plates and grown overnight at 37 °C und 120 rpm. Kanamycin concentration was 50 μg/ml for each medium type. For unlabeled protein, *E. coli* cultures were grown in LB_{Kan} main cultures. For ^{15}N labeled protein or ^{13}C,^{15}N labeled protein, $M9_{Kan}$ with ^{15}NH$_4$Cl and unlabeled glucose or ^{15}NH$_4$Cl and ^{13}C glucose were used, respectively. For selectively glycerol labeled cultures, glycerol and sodium bicarbonate were added as carbon sources instead of glucose, as described above.

Cells from the preculture were transferred to the main culture in a ratio to set the start OD_{600} to 0.1. If the main culture was labeled, cells were pelleted before the transfer and resuspended in the new medium to avoid isotope scrambling. Main cultures were grown at 37 °C und 120 rpm in 2 l baffled flasks per 0.5 l of culture. At OD_{600} 0.6-0.8, recombinant expression was induced by adding IPTG to a final concentration of 1 mM. Cells were harvested after over night expression at 37 °C. If protein purification did not start immediatelly, cell pellets were washed in cell lysis buffer (vide infra) and stored at -20 °C.

Purification of Recombinant MAK33 V_L and C_L Domains

For protein purification, the following buffers were used:

cell lysis buffer	
50 mM	Tris
10 mM	NaCl
10 mM	EDTA

refolding buffer (pH 8)	
250 mM	Tris
5 mM	EDTA
100 mM	L-arginine

inclusion body dissolving buffer (pH 8)	
50 mM	Tris
5 mM	EDTA
8 M	urea
1%	β-mercaptoethanol

gel filtration buffer (pH 6.5)	
20 mM	sodium phosphate
50 mM	NaCl

anion exchange chromatography buffer (pH 8)	
25 mM	Tris
5 mM	EDTA
5 M	urea

The purification of MAK33 V_L and C_L was published previously [29, 33]. The following description refers to 1 l of culture. For other culture volumes, all further volumes scale

linearly. Cell pellets were resuspended in 50 ml cell resuspension buffer. One tablet of EDTA-free protease inhibitor (Roche, Penzberg, Germany) was added. Cells were subjected to ultrasonication for lysis. Afterwards, DNase was added to a final concentration of 10 μg/ml. Inclusion bodies were pelleted at 24,000 g and separated from the soluble fraction. The inclusion bodies could safely be stored at -20 °C at this point.

Inclusion bodies were dissolved in 30 ml inclusion body dissolving buffer and incubated for 2 h at 25 °C with agitation to ensure solubilisation. Subsequently, insoluble components were separated by centrifugation at 20,000 g and the inclusion bodies were purified by anion exchange chromatography, using a 20 ml Q Sepharose column. Inclusion bodies eluted in the flowthrough fraction, while impurities bound to the column. In order to remove impurities, the column was thouroughly cleaned with 2 column volumes of 1 M NaOH and 2 column volumes of 2 M NaCl. The flowthrough (approximately 80 ml) was diluted to a final volume of ca. 150 ml and dialysed over night at 4 °C against 1 l of refolding buffer. The MWCO for the dialysis tubes was 3.5 kDa.

The refolded protein was gently concentrated to ca. 15 ml and then purified by size exclusion chromatography, using a 120 ml Superdex 75 column. Although this exceeds the suggested loading volume for the column, the resulting fractions were found to be pure and further concentration only increased the amounts of aggregates. Pure fractions were pooled. The total yield was determined using a molar extinction coefficient of ϵ_{280} = 15,595 M^{-1}cm^{-1} and the MW of 11.8 kDa for MAK33 V$_L$ WT. For other constructs, these values were adjusted according to the results from ProtParam [168]. All further experiments were conducted using the gel filtration buffer, unless stated otherwise. For long-time storage, the protein was aliquoted, frozen in liquid nitrogen and stored at -80 °C. For short-time storage, most MAK33 V$_L$ and C$_L$ variants could be stored at 4 °C. Solutions of the unstable variants, e.g. MAK33 V$_L$ I2E, were always prepared from fresh or frozen aliquots.

3.2.3 Protein Analysis

Polyacryl Amide Gel Electrophoresis

Tricine sodium dodecyl sulphate polyacryl amide gel electrophoresis (SDS-PAGE) was used to assess purity or check for degradation. Tricine SDS-PAGE is a variant of SDS-PAGE optimised for small MW proteins. The procedure was carried out as published [169].

Mass Spectrometry

Matrix Assisted LASER Desorption Ionisation Time of Flight Mass Spectrometry (MALDI-TOF MS) was used to ensure the correct MW of purified proteins. The analysis was carried out by an in-house MS routine using an Ultraflex2 MALDI-TOF/TOF mass spectrometer (Bruker, Bremen, Germany). The α-cyano-4-hydroxycinnamic acid (HCCA) matrix was dissolved in 0.1% tricarboxylic acid (TCA), 80% acetonitrile and 20% water. Proteins were spotted in the matrix solution and analysed after the evaporation of the solvent.

Melting Curves

Circular dichroism spectroscopy (CD spectroscopy) was used to determine melting temperatures (T_m) of proteins. Measurements were performed using a J715 spectropolarimeter (Jasco, Gross-Umstadt, Germany) with a Peltier unit (PTC-348WI) and a thermostat (Julabo725). Melting curves were recorded at a wavelength of 212 nm and 20 °C/h at 10 μM protein concentration. Curves were fitted with Matlab version R2014b (TheMathWorks, Natick, USA) using the equation:

$$[\theta] = [\theta]_{min} + \frac{[\theta]_{max} - [\theta]_{min}}{1 + exp((T_m - T)/c)}$$

with $[\theta]$ being the ellipticity, T_m being the melting temperature and c being the cooperativity index.

Transmission Electron Microscopy

Negative stain transmission electron microscopy (TEM) was used to assess the presence and morphology of amyoid fibrils. Copper grids with 300 meshes coated with formvar/carbon film (Electron Microscopy Sciences, Hatfield, USA) were glow-discharged in argon atmosphere for 30 s at 3 mA. 5 μl of a 50 μM protein sample were incubated for 60 s on the grid. After removing the protein solution, the grid was washed with water. 5 μl uranyl acetate solution (2% w/v) were applied on the grid for staining and removed after 30 s.

Photographs were taken on a Jeol JEM 100CX transmission electron microscope (Jeol, Tokio, Japan) operating at 100 kV. Images were recorded on Kodak SO163 films, which were developed with Eukobrom developer (2.1 l distilled water, 350 ml Eukobrom) for 5 minutes. Photographs were scanned at a resolution of 1,000 dpi.

Thioflavin T Assays (conducted by Benedikt Weber)

Thioflavin T (ThT) assays were used to monitor fibril growth [170]. The ThT assays for this project were conducted by Benedikt Weber from the research group of Johannes Buchner (chair for biotechnology, chemistry department, Technische Universität München). ThT assays were performed in triplicates in 96 medium binding well microplates (Greiner Bio-One, Frickenhausen, Germany) with a FP-8500 fluorescence spectrometer (Jasco, Gross-Umstadt, Germany) equipped with an FMP-825 fluorescence microplate reader (Jasco, Gross-Umstadt, Germany). Assays were conducted with 50 μM MAK33 V_L I2E protein solution, 10 μM ThT solution and EGCG from 0 to 250 μM (0 to 5 equivalents) in PBS buffer pH 7.4 with 0.05% NaN_3 with a final reaction volume of 250 μl. ThT fluorescence was recorded at 440 / 482 nm excitation respectively emission wavelengths. Between the fluorescence measurements, 96 well microplates were incubated at 37 °C under continuous orbital shaking (350 rpm) using a PHMP thermoshaker (Grant Instruments, Cambridge, UK).

3.2.4 Amyloid Fibril Preparation

De novo amyloid fibrils from MAK33 V_L S20N were formed by thoroughly exchanging the buffer to a solution, which contained 25 mM o-phosphoric acid, 25 mM acetic acid and 50 mM NaCl. This buffer was adjusted to pH 2. In order to inhibit bacterial growth, 0.05% NaN_3 were supplemented. The protein concentration was adjusted to 50 μM. The batch was transferred into a glass vial to prevent adsorption to the wall or templating effects. Upon incubation at 37 °C and agitation at 120 rpm, white precipitate was typically visible within one day. Fibrils were observed within 1 week of incubation, as confirmed by TEM. To prevent agglomeration of fibrils, agitation of the samples continued until the fibrils were packed into MAS solid-state NMR rotors.

In order to increase homogeneity and reproducibility, fibril samples for solid-state NMR spectroscopy were grown after several generations of seeding. The procedure is the same as for *de novo* fibril formation, but 1 to 2% of seeds were added. These seeds were generated by taking an aliquot from a previous generation and subjecting it to ultrasonication for 10 minutes. Seeds for a new generation were taken at earliest after 7 days of incubation, if TEM analysis confirmed the presence of fibrils. All solid-state NMR samples used in this project were derived from the 7th generation or later of continuous seeding.

3.2.5 Magic Angle Spinning Solid-State Nuclear Magnetic Resonance Spectroscopy

Sample Preparation for Magic Angle Spinning Solid-State Nuclear Magnetic Resonance Spectroscopy

An OptimaL-100 XP ultracentrifuge (Beckman Coulter, Krefeld, Germany) equipped with an SW 32 Ti swinging bucket rotor and a rotor filling device (Giotto Biotech, Florence, Italy) [158] were used to pack the protein aggregate into an MAS rotor. The rotation frequency of the centrifuge was set to 28,000 rpm. Samples were packed into 3.2 mm thin wall ZrO_2 rotors with vespel caps (CortecNet, Voisins Le Bretonneux, France) employing house-made teflon spacers (height approximately 1 mm).

General Setup of Experiments in Magic Angle Spinning Solid-State Nuclear Magnetic Resonance Spectroscopy

MAS solid-state NMR experiments were recorded on Avance II or Avance III spectrometers operating at ^{1}H Larmor frequencies of 750, 500 and 400 MHz. The spectrometers are equipped with triple resonance (^{1}H/^{13}C/^{15}N) MAS probes for 3.2 mm rotors. All measurements were conducted at a nominal temperature of 0 °C (MAS then increases temperature slightly above freezing temperature). Experiments started either with ^{1}H,^{13}C cross polarisation (CP) or ^{1}H,^{15}N CP. In both cases, the CP was used to select for rigid parts of the structure and reduce longitudinal relaxation times. Contact times were approximately 1 ms for ^{1}H,^{13}C CP or 1-2 ms for ^{1}H,^{15}N CP, respectively. In both cases, a ramp from 100% to 70% power was used on ^{1}H and constant power

irradiation on the heteronucleus. During acquisition periods, ^{1}H were decoupled with 70-80 kHz SPINAL-64 decoupling [171]. During mixing steps, continuous wave (cw) decoupling of 70-100 kHz was applied, if appropriate.

^{13}C,^{13}C-Correlation Experiments

^{13}C,^{13}C-correlations were acquired either as proton driven spin diffusion (PDSD) [172] or dipolar assisted rotational recoupling (DARR) [173]. Mixing times were set to 50-70 ms for intraresidue correlations. MAS frequency was set to 10 kHz, 11 kHz or 17 kHz. This frequency was adjusted to avoid rotational side bands in the signal region.

^{13}C,^{15}N-Correlation Experiments

^{13}C,^{15}N-correlations started with a ^{1}H,^{15}N CP, followed by an indirect evolution period on ^{15}N. Subsequently, coherence was transferred to ^{13}C by use of SPECIFIC-CP [174]. The contact time for SPECIFIC-CP was 4-6 ms, with a ramp from 100% to 90% power on ^{15}N and constant power irradiation on ^{13}Cα or ^{13}CO, respectively.

NCACX and NCOCX Assignment Experiments

NCACX and NCOCX experiments were used for obtaining sequential assignments [175]. The experiments combine the SPECIFIC-CP based ^{13}C,^{15}N-correlation and the PDSD or DARR based ^{13}C,^{13}C-correlation explained above. Both experiments were recorded as 2D (without ^{13}Cα/^{13}CO evolution) for enhanced sensitivity or as 3D for increased dispersion. MAS frequency was set to either 10 or 17 kHz.

NCOCA and CANCO Experiments

NCOCA and CANCO [176] were used as complementary approach for assignment, together with NCACX. The CO-CA transfer was achieved by use of bandselective homonuclear CP (BSH-CP)[177]. The BSH-CP was implemented with a 4.4 ms ramp from 100% to 70% power on resonance on ^{13}Cα. CANCO used two SPECIFIC-CP transfers. MAS frequency was set to 17 kHz.

NCACB and N(CO)CACB Experiments

NCACB and N(CO)CACB were another set of complementary assignment experiments, making use of the good dispersion of ^{13}Cβ [176]. CO-CA transfer was achieved using the BSH-CP [177] and CA-CB mixing was accomplished with DREAM [178]. The DREAM transfer was implemented with a 5 ms tangent shaped pulse with an amplitude spread from 80% to 100%. The MAS frequency was set to 17 kHz.

Nonuniform Sampling

NUS was used in order to maximise sensitivity of the experiments [134]. For NUS experiments, sparsity was set to 50%. The experiments used exponentially decaying

sampling densities, with decay constants between 3 ms (biased sampling for higher signal enhancement) up to 9 ms (matched sampling with sampling decay rate equal to the ^{15}N T_2 time).

NUS spectra were processed either with the compressed sensing function of the Top-Spin 3.2 NUS plugin or with hmsIST [132] in combination with NMRPipe [159].

3.2.6 Solution-State Nuclear Magnetic Resonance Spectroscopy

Sample Preparation for Solution-State Nuclear Magnetic Resonance Spectroscopy

For assignment experiments, protein samples were concentrated up to 1 mM. For other experiments, the concentration was adjusted to 50 μM. All samples contained 10% D$_2$O for locking. For assignment experiments, Shigemi NMR tubes (Shigemi, Tokyo, Japan) were used with at least 250 μl of sample, for other experiments, standard NMR tubes (Norell, Landisville, USA) were filled with ca. 500 μl of sample.

For titrations, a stock solution of 20 mM or 50 mM EGCG in water was prepared. To ensure complete solubilisation, the stock solution was sonified for 5 minutes. The protein solution was kept in the NMR tubes and EGCG was added directly to the NMR tube. Samples were mixed gently by inversion. Immediate measurement after mixing was crucial, as the reaction with EGCG is time dependent.

Experiments in Solution-State Nuclear Magentic Resonance Spectroscopy

Solution-state NMR experiments were recorded on Avance III spectrometers operating at ^{1}H Larmor frequencies of 900, 600 and 500 MHz. The spectrometers are equipped with cryogenic (except for one of the 600 MHz spectrometers) triple resonance gradient probes (^{1}H/^{13}C/^{15}N). All measurements were conducted at 25 °C (experiments at lower temperatures displayed less resonances, whereas at higher temperatures, the protein denatured too fast).

All experiments contained a watergate pulse sequence for solvent suppression [179]. 2D ^{1}H,^{15}N-correlations were recorded as heteronuclear single quantum coherence (HSQC) experiments with ^{13}C decoupling. Assignment experiments comprised sensitivity enhanced HNCA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO and HCC(CO)NH [126]. For the assignment of dynamic regions of the fibrils, an HN(CA)NNH experiment [180] was acquired as complementary approach to the aforementioned experiments. Due to the T_1 of 1 s for MAK33 V_L, the relaxation delay of all 2D or 3D experiments was set to 1.3 s.

3.2.7 Analysis of Nuclear Magnetic Resonance Spectroscopic Data

Resonance assignments were conducted in CcpNmr Analysis 2.4.0 [161]. Chemical shift changes were calculated using the following equation:

$$\Delta \delta^{res} = \sqrt{(\delta^{1H}) + \frac{1}{25} \times (\delta^{15N})}$$

Secondary chemical shifts were both analysed with TALOS+ [124] or calculated as

$$CSI = (\delta_{observed}^{C\alpha} - \delta_{random\ coil}^{C\alpha}) - (\delta_{observed}^{C\beta} - \delta_{random\ coil}^{C\beta})$$

with CSI being the chemical shifts index.

3.2.8 Docking and Molecular Dynamics Simulations (conducted by Martin Carballo Pacheco)

Docking experiments were used for identification of the EGCG binding site to MAK33 V_L. Molecular dynamics (MD) simulations were conducted to study interactions and assess binding properties. These experiments were conducted by Martin Carballo Pacheco from the research group of Birgit Strodel (computational biochemistry group, Forschungszentrum Jülich).

Docking simulations of EGCG binding to MAK33 V_L were performed using both AutoDock Vina 1.1.2 [181] and HADDOCK 2.1 [182]. The simulations were started from the MAK33 V_L structure of PDB code 1FH5 [27] (residues 2-109 of the LC, which from now are called 1-108) with the mutations E17D, S20N and Y87F. With both docking programs semi-flexible docking was carried out, allowing EGCG to be flexible while the protein is considered rigid. The docking simulation with AutoDock Vina 1.1.2 was performed with an exhaustiveness factor of 100 and the entire surface of the protein as a possible target for EGCG binding. The docking simulation with HADDOCK 2.1 was performed with I58 and P59 as predefined contacts for EGCG binding, as predicted from the NMR experiments. Therefore, the simulation based on docking with AutoDock Vina is referred to as unguided simulation, while the simulation with HADDOCK is referred to as NMR-guided simulation.

In order to test the stability of the obtained complexes and for allowing MAK33 V_L to adjust to the EGCG bound to it, MD simulations were performed using Gromacs 4.6.4 [183]. The systems were modeled with the Amber99sb-ildn force field [184] and the Tip4p-Ew water model [185]. EGCG was parametrised using acpype [186], with bonded and Lennard-Jones parameters taken from the general Amber force field (GAFF) [187] and RESP charges calculated with antechamber [188] from a HF/6-31G* single point calculation after the geometry was optimised at the B3LYP/6-31G* level. Before starting the MD simulations, we performed steepest descent energy minimisations. The systems were then equilibrated with a 1 ns NVT simulation followed by a 1 ns NPT simulation. Afterwards, 500 ns production runs were performed. In all MD simulations a cutoff of 1.0 nm was used for short-range van der Waals and electrostatic interactions, while long-range electrostatic interactions were calculated using the particle mesh Ewald method [189]. Periodic boundary conditions and a 2 fs timestep were used. The temperature was kept constant at 310 K with the Nos-Hoover thermostat [190] and the pressure was kept constant at 1.0 bar using the Parrinello-Rahman barostat [191]. For the analysis, the first 20 ns of each simulation were discarded, considering it as equilibration. The trajectories were analysed using Gromacs tools [183] and MDAnalysis 0.8.1 [192]. Contacts between EGCG and the protein were considered when the distance between any atom of the two molecules was less than 0.3 nm.

4 Characterisation of the Natively Folded MAK33 V_L and C_L Domain and the Effects of Point Mutations

4.1 Resonance Assignment of MAK33 V_L

The murine κ-type MAK33 V_L was employed as a model system to study fibril formation in AL amyloidosis. The first focus was on the WT V_L domain and the S20N mutant. The motivation for this mutation was a report on glycosylation and its role in amyloidogenicity [66]. Specifically, in a patient sequence, residue 20 was suggested to be an N-linked glycosylation site. The aim of this project was, to determine whether glycosylation is necessary to enhance fibril formation, or if the introduction of an asparagine is sufficient.

This work was conducted in close collaboration with Cardine Nokwe from the research group of Johannes Buchner (Technische Universität München). As a result of this research project, the S20N mutant was identified as facilitating fibril formation, independent of glycosylation (vide infra) [193].

We assigned most backbone NMR resonances of MAK33 V_L WT and S20N by use of standard 3D sequential assignment experiments (Fig. 14). Although HNCACB and CBCA(CO)NH were crucial for identifying residue types and resolving ambiguities, it is worth noting that ^{13}CO resonances were also very helpful for completing the assignment. The primary sequence of MAK33 V_L contains 19 serines and 8 threonines, which have many overlapping ^{13}Cα and ^{13}Cβ resonances. Here, the surprisingly good dispersion of ^{13}CO resonances was essential to establish sequential connectivities.

Not all residues could be assigned. Different sample conditions were tested in order to complete the assignment. This revealed a strong effect of protein concentration on the chemical shifts (Fig. 15). Such effects were found for both WT and S20N mutant as well as for the other mutants studied later. Systematic analysis of concentration dependent effects on different residues revealed several hot spot regions with large chemical shift changes (Fig. 16). These were located close to the gaps in the assignment. These gaps comprise the first N-terminal residues as well as many residues in the region 40 - 50 and 90 - 100. As concentration dependent effects indicate interaction between monomers, these data were analysed further with respect to dimerisation. Therefore, a crystal structure of a homologous V_L dimer was used as template [194]. Mapping of MAK33 V_L on this dimer and then highlighting unassigned residues showed a clear correlation (Fig. 17). All non-proline residues, which could not be assigned, are located in the dimer interface.
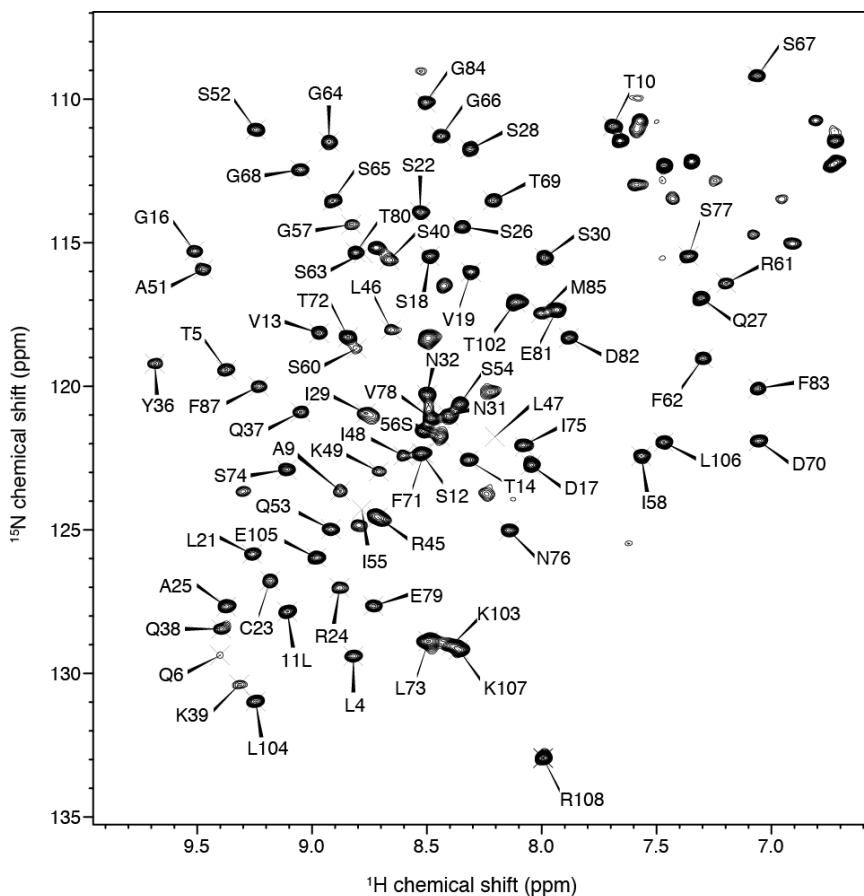
Figure 14: ^{1}H,^{15}N **HSQC of MAK33 V$_L$ WT**
With few exceptions, all observable backbone resonances were assigned sequentially.
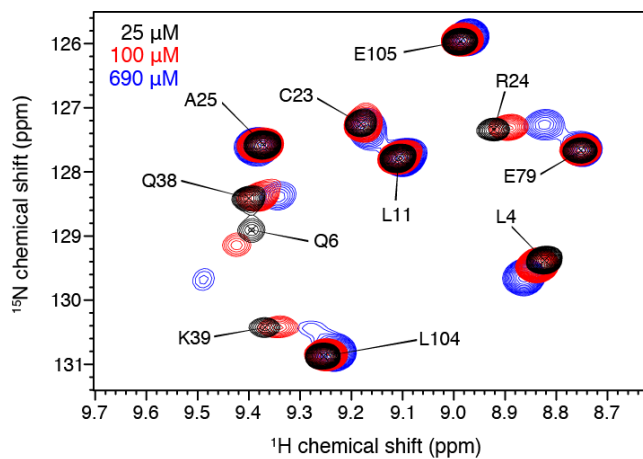
Figure 15: Effects of Protein Concentration on MAK33 V$_L$ S20N
Some resonances displayed large chemical shift changes in the ^{1}H,^{15}N HSQC, e.g. Q6, R24 or K39. At least small changes were seen for all resonances.
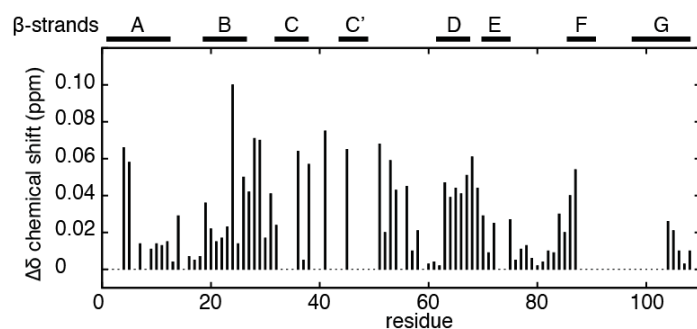
Figure 16: Concentration Dependent Chemical Shift Changes of MAK33 V_L S20N
Changes in backbone chemical shifts are plotted for MAK33 V_L S20N at 25 μM and 690 μM. Gaps represent residues, for which assignment is missing at least at one of the concentrations.
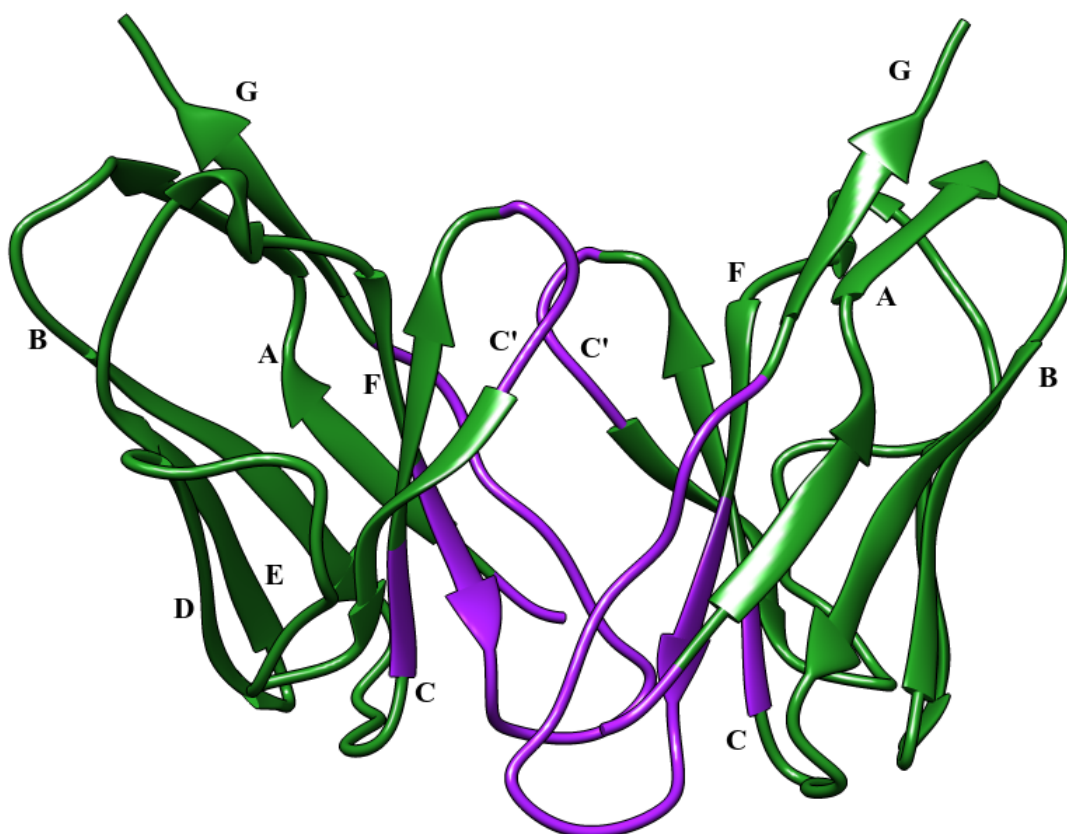
Figure 17: Unassigned Residues are Located in the Dimer Interface of MAK33 V_L
A dimer model was constructed. Those residues, which could not be assigned by NMR spectroscopy, are coloured in purple. These residues are located in the dimer interface. The dimer structure was created by mapping the monomers on a crystal structure of a V_L dimer (PDB ID: 1LGV) [194].

45

This, first of all, indicates that MAK33 V_L does form a dimer, which was not known before. Second, this explains, why the NMR resonance assignment could not be completed. Within the applied range of concentrations, the protein exists in an equilibrium between monomer and dimer, as indicated by the chemical shift changes. This causes chemical exchange broadening, as described in the introduction of this thesis. Briefly, nuclei in the dimer interface sample environments, which cause different resonance frequencies, within the time frame of one experiment. This causes averaging and thus, broadening of the resonances. Hence, while some residues might still be observable in sensitive ^{1}H,^{15}N HSQC experiments, they vanish in the noise in the less sensitive backbone assignment experiments. More dynamic residues even cannot be seen in the HSQC experiment.
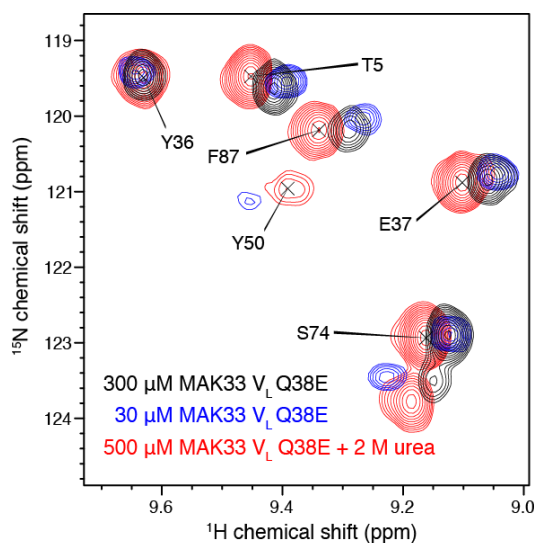
Figure 18: Dimerisation can be Modulated by Addition of Urea
NMR spectra of the MAK33 V_L Q38E mutant are shown at different concentrations in presence and absence of urea. Without denaturants, the resonance of Y50, a residue in the dimer interface, could only be observed at low concentrations. At 300 μM, it disappeared due to chemical exchange broadening. In presence of 2 M urea, it reappeared, as the urea reduced the fraction of dimeric protein. The spectra at 300 μM and 500 μM were acquired with analogous setup and are displayed with identical contour levels.

In order to extend the resonance assignment, the equilibrium needs to be shifted. For practical reasons, this cannot be accomplished by changing the protein concentrations. At lower concentrations, where the protein exists mostly monomeric, the sensitivity of NMR experiments is too low for assignment experiments. At higher concentrations, where the dimer should be the dominant species, the protein precipitates quickly. Further strategies to affect the monomer-dimer equilibrium could be site directed mutagenesis or change of environmental conditions. In the following, we combined both approaches. We introduced the Q38E mutation, based on literature information on conserved residues

in the V_L dimer interface [31]. Further conceivable mutations were not undertaken, as it is not known, which residues precisely affect dimerisation. In addition, mutagenesis always causes chemical shift changes, especially for residues close to the mutation site. Therefore, so far unassigned resonances might experience so large chemical shift changes, that they cannot be tracked back from a mutant to the WT.

In addition, we modified the environment in order to increase the fraction of monomers. This was accomplished by adding urea to the buffer. Urea is a commonly used denaturant. Hence, the concentration must bet set considerably lower than the concentration neccessary for an unfolding transition. Initial screenings revealed a urea concentration of 2 M as safe, i.e. the protein was long enough stable to conduct assignment experiments. Fig. 18 displays, that the addition of urea indeed increased signal intensities of residues located in the dimer interface.

In combination, the Q38E mutation and the addition of urea enabled us to extend the backbone resonance assignment. Although many residues within the dimer interface still remained unassigned, many adjacent residues could be added to the assignment. Overall, we could successfully assign 83 of the 103 non-proline backbone resonances (Fig. 19).
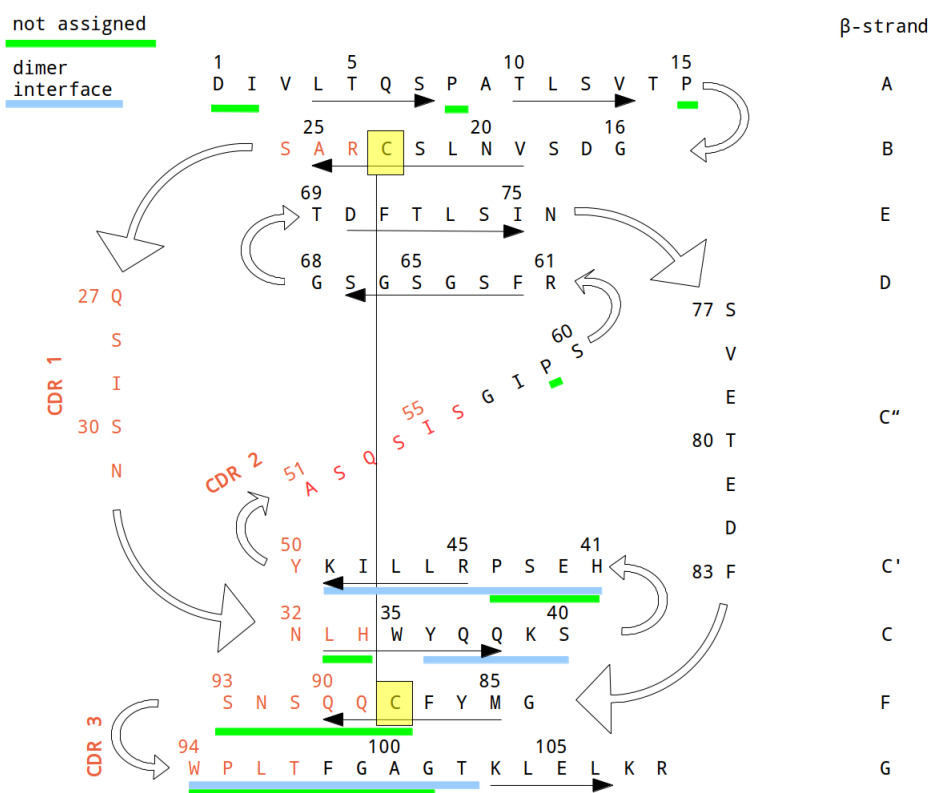

Figure 19: Assigned Residues in MAK33 V_L S20N
The majority of the unassigned residues is located in the dimer interface.

47

4.2 Analysis of Amyloidogenic Point Mutants of MAK33 V_L

We used the obtained resonance assignments to analyse the effects of several point mutations, which are known to be associated with AL amyloidosis. These mutants comprise MAK33 V_L I2E, MAK33 V_L S20N and MAK33 V_L D70N. As described in the previous section, the S20N mutant was used to analyse the effects of an asparagine substitution without glycosylation. The same goal was pursued with the D70N mutation. Again, the position of the mutation site was selected according to literature information on patient AL sequences [66, 195]. The I2E mutation, finally, originated from sequence comparisons between MAK33 V_L and the highly similar V_L domain of the OPG2 antibody with the PDB ID 1OPG [196]. The latter one is identical to MAK33 V_L except for five residues, with the I2E substitution being one of the differences. For all three mutations, i.e. I2E, S20N and D70N, biochemical and biophysical analysis in Johannes Buchner's research group revealed, that the mutations enhance fibril formation [197, 193]. Using NMR spectroscopic analysis, our goal was to rationalise these observations on a structural level.

4.2.1 The I2E Mutation in MAK33 V_L

At first, the I2E mutation shall be discussed here. In a ^{1}H,^{15}N HSQC experiments, many resonances shifted in comparison to MAK33 V_L (Fig. 20). Still, the assignment could be transferred to the I2E mutant. Analysis of the chemical shift changes revealed hot spots in CDR1 between β-strands B and C and also in strands D and E. Mapping of the chemical shift changes on the protein structure showed effects even on many residues, which were distant from the mutation site (Fig. 21).

Residue 2 sticks into a hydrophobic pocket, formed by I4, A9, I29 and F71. All these residues were strongly affected by the mutation. The native isoleucine at position 2 could bind there. The charged glutamate, instead is repelled from the pocket. Without this interaction, the N-terminus dissociated from the hydrophobic core. This is in agreement with MD simulations of the WT and the I2E mutant [197].

The perturbations of the hydrophobic core had large effects on the whole structure, as monitored by the widespread changes in chemical shifts. Sequentially close to the I29 in the hydrophobic pocket are C23 and R24. Both residues are involved in crucial interactions. R24 forms a conserved salt bridge with D70, which, together with the hydrophobic pocket residue F71, explains the large effects of the I2E mutation on residues 67 to 71. C23, on the other hand, forms a highly conserved disulfide bridge with C88. As also C23 displayed considerable chemical shift changes, the rearrangements in the hydrophobic core could be mediated by C23 and R24 to further distant residues.

Concordantly, thermodynamic stability of the mutant was drastically reduced. T_m was reduced by 10 °C [197]. Such destabilisation enabled formation of amyloid fibrils at native like conditions, i.e. agitation at 37 °C. In contrast, the WT protein required pH 2 in addition, in order to induce fibril formation. Overall, the I2E mutation supports the model that destabilising mutants facilitate amyloid aggregation [56].
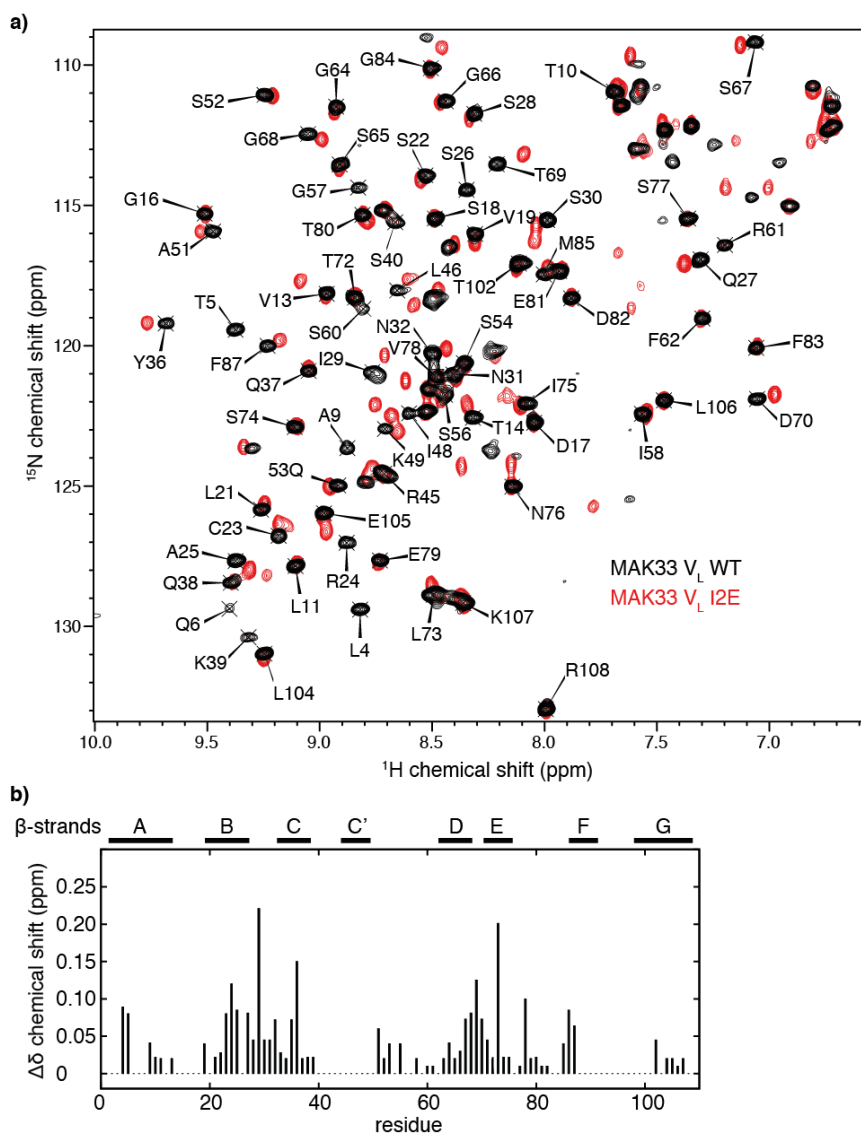
Figure 20: Analysis of Chemical Shift Changes due to I2E Mutation in MAK33 V_L
a) ^{1}H,^{15}N spectra of MAK33 V_L WT and IE2 at 50 μM concentration.
b) Differences in chemical shifts. Gaps indicate resonances, which could not be assigned unambiguously in both proteins or where resonances overlap. Many resonances displayed large chemical shift changes.
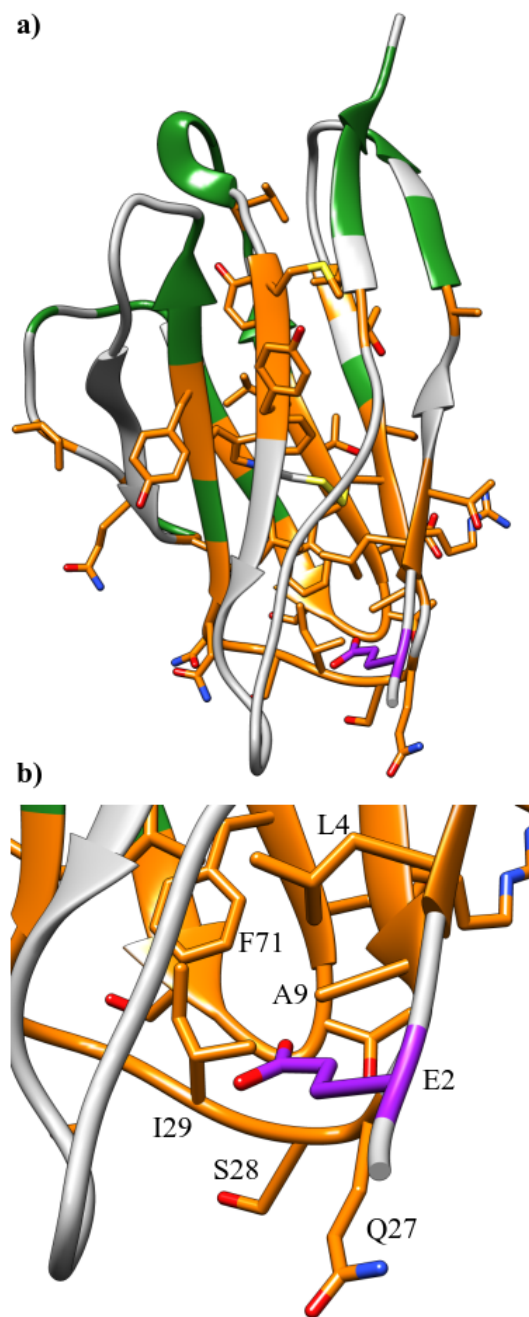
Figure 21: Structural Effects of the I2E Mutation on MAK33 V_L

    a) All resiudes, for which the chemical shift changes were larger than 0.02 ppm, are highlighted in orange. Unassigned residues are shown in grey. Residue E2 (mutated *in silico*) is displayed in purple. Residues all over the structure were affected by the mutation.

    b) Close up view on the hydrophobic pocket, in which residue 2 binds. A charged residue causes the N-terminus to dissociate from the hydrophobic core, thereby causing structural rearrangements throughout the whole protein.

4.2.2 The S20N and D70N Mutations in MAK33 V_L

The S20N and D70N mutations are different from the I2E substitution, since these mutants do not affect thermodynamic stability [193]. Nevertheless, they enhance fibril formation. Addition of 0.5 mM SDS or sonication is sufficient to enable formation of amyloid fibrils, in contrast to the WT, which requires pH 2. As not only the thermodynamic stability, but also folding and unfolding rates of the S20N and D70N mutants are equal to those of the WT sequence, the reason for increased amyloidogenicity must be related to structural changes.

Analysis of the ^{1}H,^{15}N HSQC experiments of the mutants revealed moderate chemical shift changes, in comparison to the WT (Fig. 22). Most resonances could be tracked easily. In comparison to the I2E mutation, the observed changes were smaller. For S20N, the strong chemical shift changes clustered at several spots in the sequence (Fig. 23). As expected, resonances around the mutation site were strongly affected. Additionally, we found hot spot regions from residue 63 to 66 and also from 72 to 74.
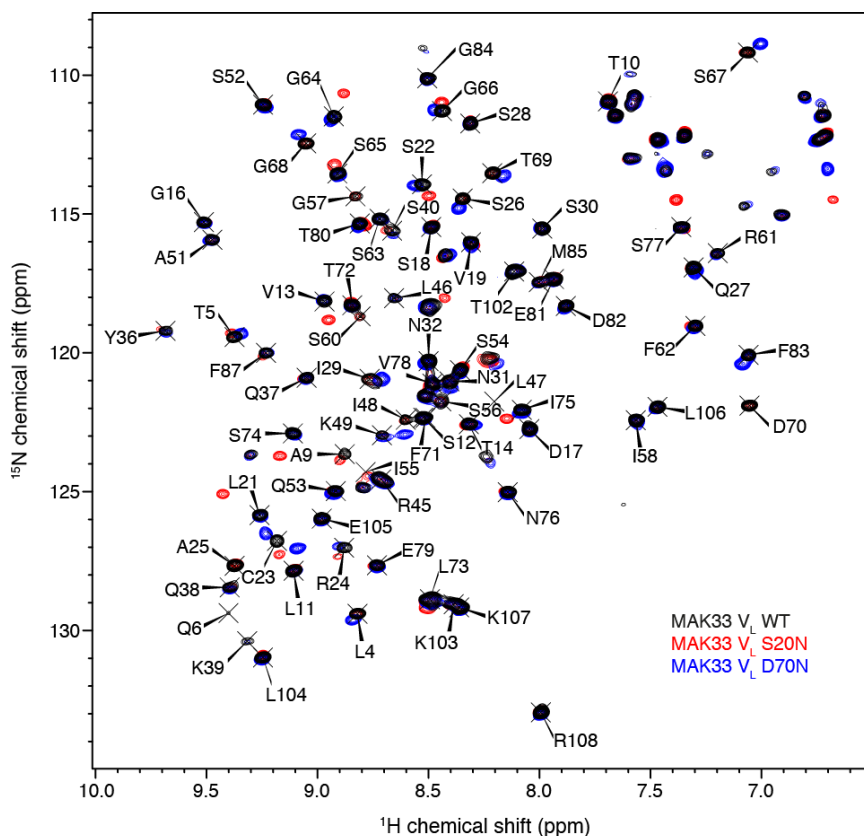
Figure 22: NMR Spectroscopic Analysis of the MAK33 V_L S20N and D70N Mutants
^{1}H,^{15}N spectra of MAK33 V_L WT, S20N and D70N at 50 μM concentration.
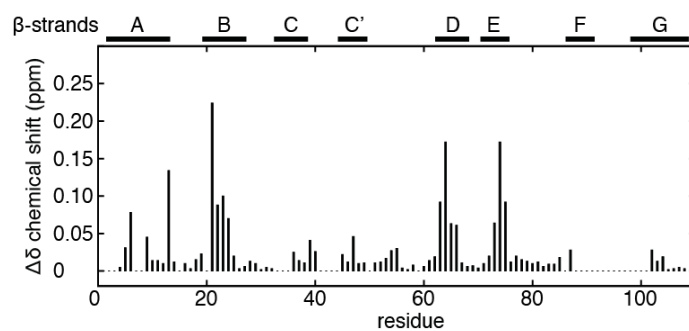
Figure 23: Chemical Shift Changes in the MAK33 V$_L$ S20N Mutant
The mutation caused localised chemical shift changes.

Figure 24: Structural Effects of the S20N Mutation
All resiudes, for which the chemical shift changes were equal to or larger than 0.02 ppm, are highlighted in orange. Unassigned residues are shown in grey. Residue N20 (mutated *in silico*) is displayed in purple. Many affected residues are part of a H-bonding network on the protein surface.

An analysis of the structure revealed, that these residues are spacially close. At the protein surface, there is a large network of serines and threonines, which interact via H-bonds. Both the native S20 and the mutated N20 could contribute to this network. MD simulations revealed, that N20 interacted simultaneously with several other residues in this network, while the native serine had only one interaction partner at a time [193]. Importantly, the effects were not restricted to the surface, but spread again to the hydrophobic core, similar to the I2E mutation. C23 with its disulfide bond to C88 was again affected, which is reasonable, since it is very close to N20. Also, D70 displayed strong shift changes, which links the S20N mutation with the D70N mutation. Actually, D70 respectivly N70 are part of the same interaction network at the protein surface. Hence, these two mutations could be expected to behave similar.

The differences between MAK33 V_L D70N and the WT were indeed comparable to those between S20N and the WT, albeit less pronounced (Fig. 25). Among the affected residues were residues 22 to 30 and of course those around the mutated N70. Looking at the structure, the pattern still looks different from S20N. For D70N, the effects were more localised (Fig. 26). The resonance of R24 shifted not as much as expected, considering that D70 forms a structurally important conserved salt bridge with R24. Possibly, the effects are more pronounced on the side chain of R24, while the backbone might undergo only minor rearrangements. Regarding the H-bonding network mentioned above, the concerned residues were partially affected here. Mostly, the residues close to the mutation site displayed chemical shift changes. Hence, N70 might play a smaller role in the solvent exposed H-bonding network. Conceivably, the effects were minor, since the penetration into the hydrophobic core was weaker. C23 might play a key role here with its disulfide bridge to C88. While N20 is not only spatially, but also sequentially close to C23, N70 exerts only indirect effects via the broken salt bridge with R24. Weaker effects on C23 might remove an important effect for mediating long range structural rearrangements.
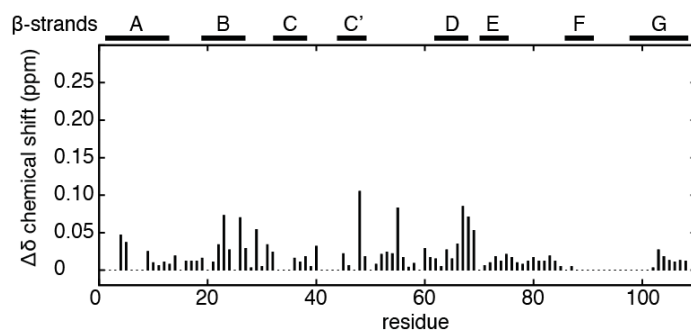
Figure 25: Chemical Shift Changes in the MAK33 V_L D70N Mutant
Those residues, which were affected by the D70N mutation, were also affected by the S20N mutation.

Despite the overall smaller effects of the D70N substitution, amyloidogenicity was similarly enhanced [193]. Hence, it might be worth to compare, which residues were affected by both mutations. The hydrophobic core residues affected in both cases were mainly located in β-strands A and B. Strikingly, both mutations caused changes for
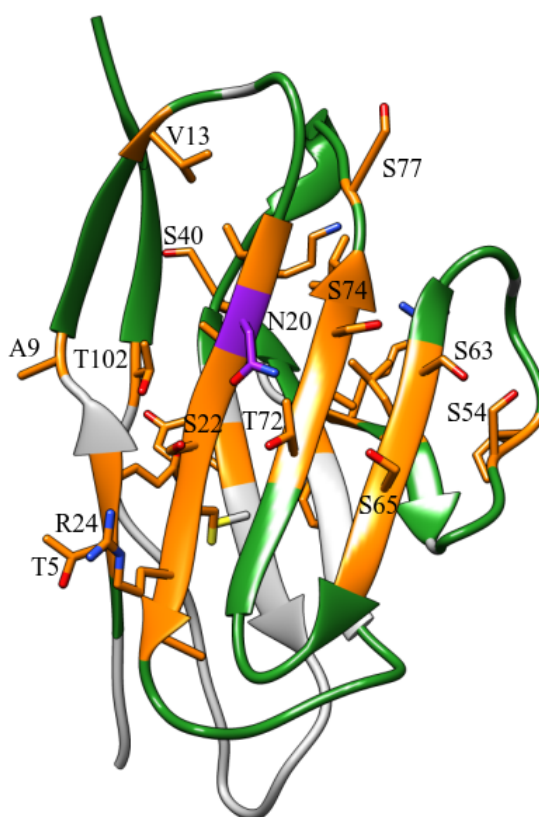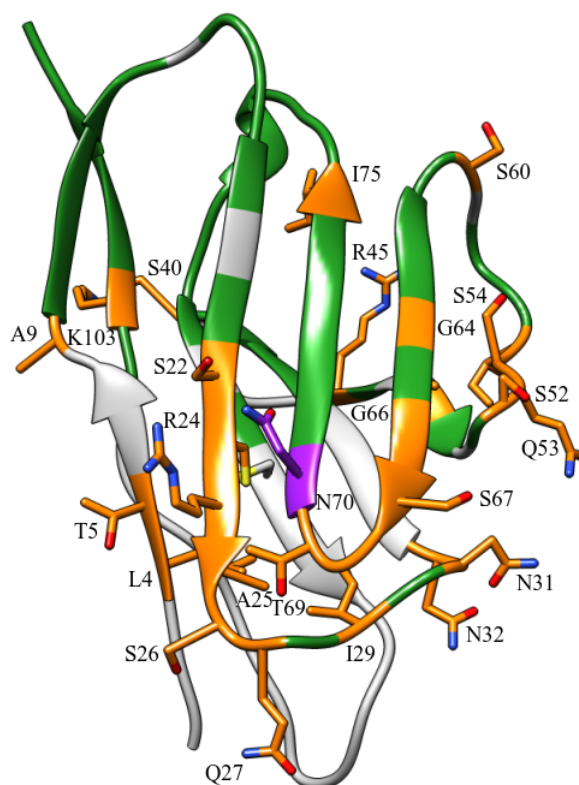
Figure 26: Structural Effects of the D70N Mutation
All resiudes, for which the chemical shift changes were equal to or larger than 0.02 ppm, are highlighted in orange. Unassigned residues are shown in grey. Residue D70 (mutated *in silico*) is displayed in purple. Most affected residues are located at the antigene binding site (which is the bottom part in the shown orientation).

L6 and the binding site for the N-terminus to the core of the domain. This situation might thus be similar to the effects of the I2E mutation. While the current data are not conclusive enough for a final statement, dissociation of the N-terminus from the core might be a crucial step for fibril formation. This could be related to the edge-strand hypothesis, where terminal strands of β-sheet proteins are determined as important for liability towards fibril formation [64]. Specifically, the N-terminus was already identified in other amyloidogenic light chain sequences as key factor for amyloid aggregation [65].

4.3 The Linker Residue R108

The last section described effects of point mutations within the V_L domain. Now, the topic will be extended towards the C_L domain. Since this domain is frequently subject to truncation in AL amyloidosis, it is likely to modulate amyloid formation [45, 46]. Naturally, the linker between both domains lends itself for first analyses. The last residue of MAK33 V_L is R108. On a protein level, this arginine is actually within the linker, while the attribution to the V_L domain is rooted in the gene level. A brief summary of the respective genetics was given in the introduction of this thesis.

In further cooperation with Cardine Nokwe and Johannes Buchner, we studied the effects of R108 deletion mutants on both V_L (as C-terminal truncation) and C_L (as N-terminal truncation). NMR spectra of the truncated V_L ΔR108 looked mostly similar to that of the WT domain (Fig. 27). Strong changes were only observed for the C-terminal residues. K107 shifted a lot, as it became the new terminal residue and thus negatively charged at the carboxyl group. Other residues exhibited only small changes, so the structure is not affected considerably.

Complementary analyses, however, revealed reduced stability [198]. The melting temperature dropped from 52.0 °C to 47.7 °C and ΔG_u from -19.4 kJ/mol to -17.8 kJ/mol, compared to the WT. Moreover, the deletion mutant could fibrillise in presence of 0.5 mM SDS with ultrasonication, while the WT did not aggregate under these conditions.

Such behaviour cannot be explained by structural effects, as the chemical shifts remained mainly the same. One might consider effects of the additional charge due to the native R108. In fact, two basic residues are located next to each other, with K107 as predecessor. Within the four previous residues, there are K103 and E105. Thus, positive charges in the terminal region are not fully compensated. However, as all of the aforementioned residues are solvent exposed, this is unlikely to compromise stability. Actually, this solvent exposition might be the key for understanding the effects of the mutation. In the WT protein, the additional charged residue might increase solubility. Still, the results should not be over-intellectualised, as the effects both on thermal and on chemical stability were rather small.

The picture was different for the C_L domain. NMR resonance assignments for the domain were transferred from published assignments [199]. In the WT sequence, which includes R108 at the N-terminus, conformational exchange in the slow regime was observed for a subset of resonances (Fig. 28). Fast and intermediate exchange have been described in previous sections of this thesis (see introduction and resonance assignment of MAK33 V_L). In slow exchange, the exchange rate is slower than the chemical shift difference in frequency units. The smallest distances between corresponding resonances were approximately 30 Hz. Therefore, the exchange rate observed here needs to be in the order of 30 s^{-1}. The ratio of the resonance integrals corresponds to the relative population of the sampled states. Accordingly, we found a major population at 76 $\pm$ 4% and a minor state populated by 24 $\pm$ 4% of the protein molecules.

In the deletion mutant ΔR108, such slow exchange was not observed any more. However, the observed resonances corresponded to the minor state of the conformations of the WT. Hence, the N-terminal truncation stabilised the previously unfavoured state.
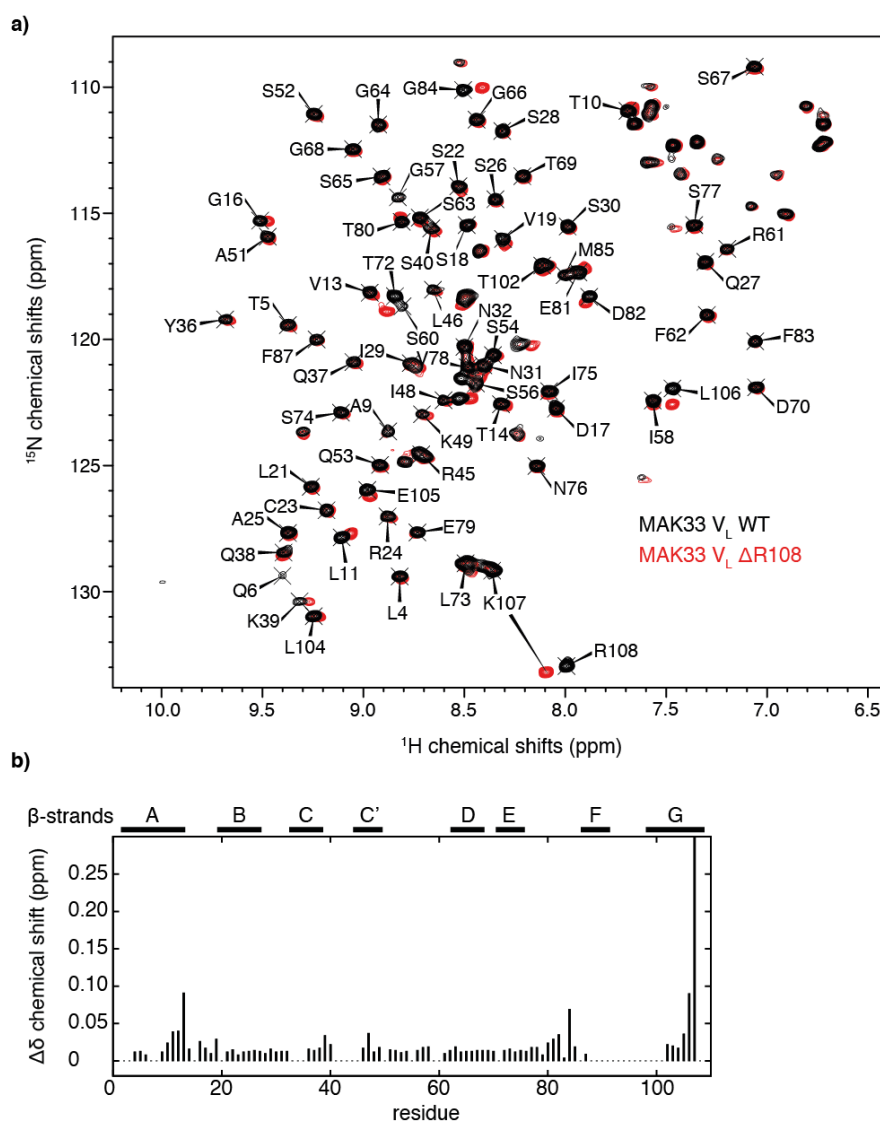
Figure 27: Chemical Shift Changes in the MAK33 V$_L$ $\triangle$R108 Mutant
a) ^{1}H,^{15}N spectra of MAK33 V$_L$ WT and $\triangle$R108 at 50 μM concentration.
b) Differences in chemical shifts. Gaps indicate resonances, which could not be assigned unambiguously in both proteins or where resonances overlap.
Chemical shift changes were mostly small, except for the linker residues close to the deletion site.
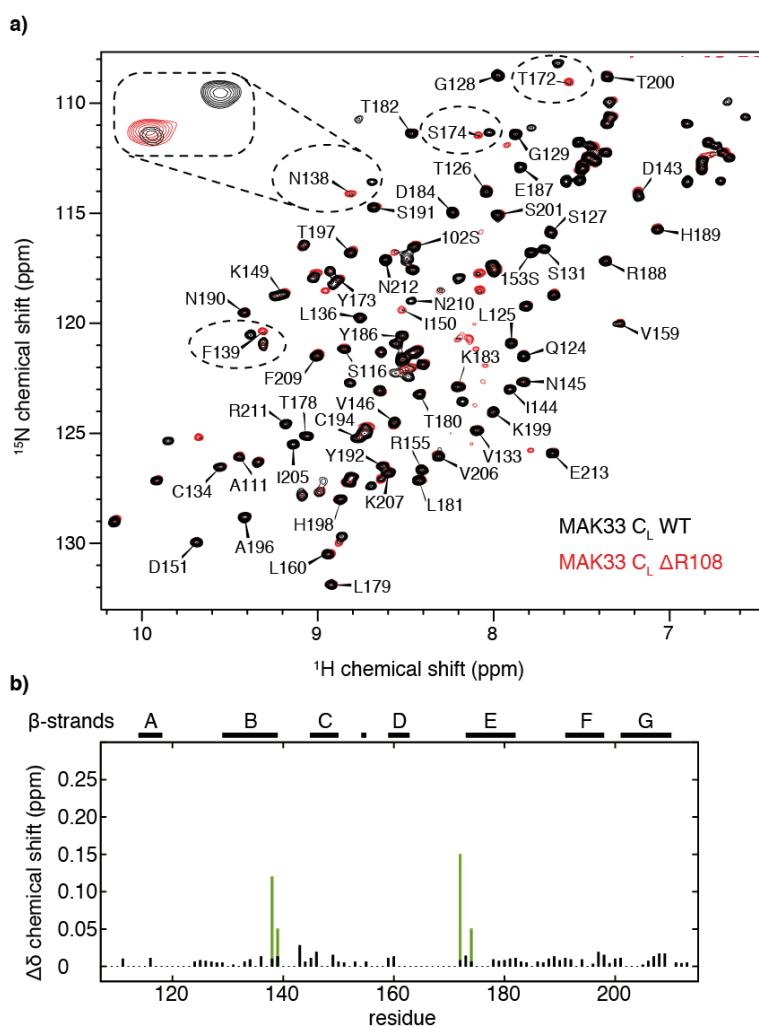
Figure 28: Chemical Shift Changes in the MAK33 C$_L$ ΔR108 Mutant
a) ^{1}H,^{15}N spectra of MAK33 C$_L$ WT and ΔR108 at 50 μM concentration. Dashed circles indicate resonances, which displayed slow exchange in the WT. For N138, the resonances are shown close-up.
b) Differences in chemical shifts. Gaps indicate resonances, which could not be assigned unambiguously in both proteins or where resonances overlap. Green bars indicate resonances, for which slow exchange was observed in the WT, but not in the mutant.

Obviously, those resonances, for which slow exchange was observed, were affected by the mutation, as the deletion prevented the exchange. Other resonances displayed only small chemical shift differences and thus were not relevant for the effects of the mutation. This revealed, that the minor conformation of the C_L WT was very similar to the ΔR108 mutant. Those residues, which exhibited slow exchange, are structurally close to the N-terminus, as expected (Fig. 29). Most likely, the N-terminal linker to the V_L domain is flexible in the WT protein. It might bind to the C_L domain. Otherwise, the linker might dissociate from the core, corresponding to the minor population in the NMR spectra. In absence of attracting interactions of R108, the mutant populates only the conformation with the free linker, thereby shifting the equilibrium.
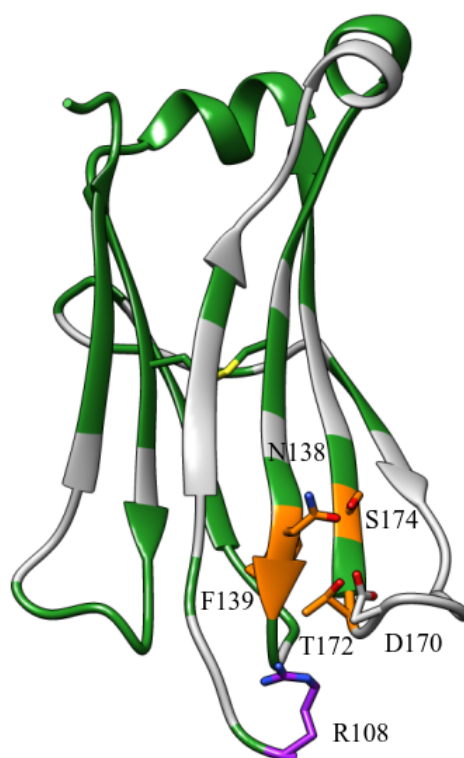
Figure 29: Residues Displaying Slow Exchange in MAK33 C$_L$ WT, but not in the $\triangle$R108 mutant
All resiudes, for which slow exchange was observed, are highlighted in orange. Unassigned residues are shown in grey. Residue R108 is displayed in purple. Interactions between R108 and the unassigned D170 bind the N-terminus to the core of the domain.

Complementary experiments conducted by Cardine Nokwe and Martin Zacharias supported this picture [198]. MD simulations revealed interactions between R108 and D170, explaining the binding of the N-terminus to the core of the domain. The exchange between two distinct conformations could not be observed by MD simulations, as the

simulations are restricted to analysis of fast events. Still, the mutant exhibited larger fluctuations in the N-terminus, indicating a dissociated state. Concordantly, ΔG_u was considerably smaller for the mutant, -7.5 kJ/mol compared to -16.8 kJ/mol for the WT. This reflects the loss of interaction between N-terminus and the domain core, which was present in the major population of the WT.

Combining the observations on the deletion of R108 in V_L respectively C_L, R108 seems to be more important for the constant domain. Here, specific interactions could be identified and the truncation shifted the equilibrium between two states to a state, where only one conformation could be observed. Also, the effects on stability were larger for the C_L domain, although the V_L domain was also destabilised enough to enhance formation of amyloid fibrils. While R108 genetically is part of th V_L domain, it affects both domains and likely plays an important role in domain interactions. The results are most likely general, as R108 and the interacting D170 are conserved in κ-type LCs.

4.4 Conclusions on Factors Enhancing Fibril Formation

We presented data on the natively folded MAK33 V_L, where assignment gaps and concentration dependent chemical shifts already revealed dynamic regions. Subsequently, different point mutations were studied, which increase the tendency to form fibrils. Finally, the linker between V_L and C_L was analysed. The effects of the individual mutations could be rationalised on a structural level and with complementary biophysical information. Now, the goal is to generalise the insights gained from specific point mutations.

While it is easy to simply explain fast fibril formation of the MAK33 V_L I2E mutant with the reduced thermal stability, this information is not new. Considering, that the S20N and D70N mutations did not alter thermodynamic stability, the picture obviously remains incomplete, if only global physical parameters are observed. Instead, we would like to emphasise the role of specific regions in the V_L domain. Therefore, we compared the observed chemical shift changes due to the different V_L mutations as well as due to different concentrations (Fig. 30). While the precise residues, which displayed chemical shift changes, varied, a pattern appears. All mutations affected β-strands B and D. As was shown in the previous sections, in each of these cases, hydrophobic core residues were affected, also in these particular strands. So, these β-strands might be promising for further analysis. It is noteworthy, that β-strand B contains residues C23 and R24. Both residues have been affected by all of the amyloidogenic mutations presented here. As already mentioned in the above sections, both residues are highly conserved and important for domain integrity, due the disulfide bond respectively the salt bridge to D70. This matches with results from other amyloid forming proteins, were cystine bonds were shown to modulate fibril formation. In islet amyloid precursor protein (IAPP), reduction of the internal disulfide bond alters the pathway for fibril formation [200, 201]. This has alse been observed for larger proteins, e.g. β-2-microglobulin. In this case, reduction prevents the amyloid aggregation [202, 203]. Superoxide dismutase forms fibrils only, if the cystine bond is reduced [204]. While the cysteins were not reduced in any of our studied mutants, the recurring chemical shift changes at C23 suggest, that the disulfide bond is important for fibril formation. C88, the other cystein involved in this bond, could not be assigned, however the residues just before it also displayed considerable changes in the I2M mutant.

To a lesser extent, also the very N-terminal residues as well as the E-strand appear to be affected in several cases. The first N-terminal residues of the V_L domain were also identified in other LC sequences as important for protection against aggregation [65]. The relevance of the N-terminus in the native state might be related to its role in the fibril form. There, the N-terminus acts as epitope for an AL fibril specific antibody, which does not bind to the native form [205].

Considering the observed chemical shift changes upon increase of concentration, it is now established that V_L dimers do not form fibrils [62, 63]. Although it is unclear, whether this is due to protection of interfaces or due to dynamics, NMR based dynamics measurements with the Len LC revealed millisecond dynamics, which are linked to fibril formation [206]. In detail, Len V_L forms fibrils at pH 2, similar to the MAK33 V_L WT.
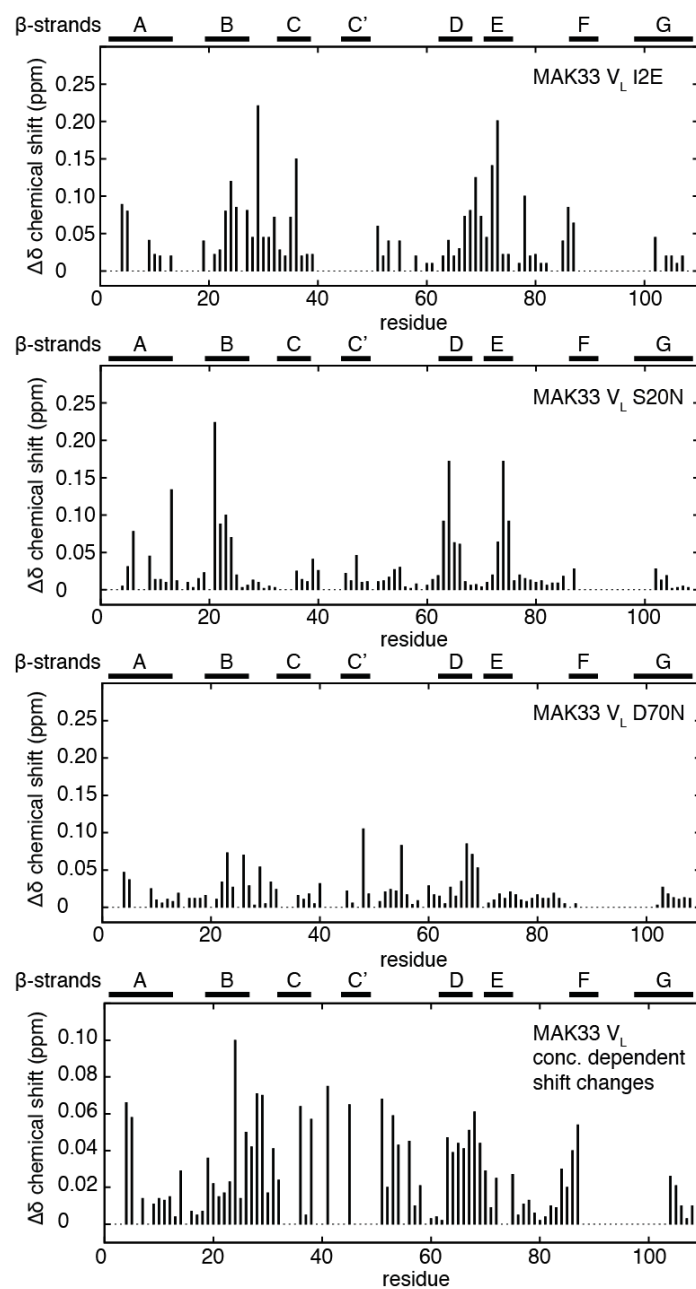
Figure 30: Overview of Chemical Shift Changes Observed for Different V$_L$ Variants and Conditions

β-strands B and D were affected by all studied perturbations of the native systems. The N-terminus and β-strand E are also affected in several cases.

61

At this pH, three regions exhibited elevated dynamics, compared to neutral pH. These comprise the N-terminal residues up to residue 9, i.e. β-strand A, residues 42 to 57, i.e. strand C' and CDR2 and residues 89 to 101, i.e. CDR3. These regions are close to the dimer interface. Len compares well with MAK33, as these regions also were difficult to assign in MAK33 as well, or could not be assigned at all. It is striking, that again the N-terminal residues behave special. In MAK33, dimerisation caused also chemical shift changes distant from the dimer interface. The picture is incomplete here, precisely due to the concentration dependent exchange broadening. However, if we connect these information on concentration dependent effects with partially similar affected regions due to the studied mutations, it is likely that these dimer-related dynamics are actually meaningful for fibril formation.

The deletion mutant V_L ΔR108 needs special consideration. As shown above, the effects of the mutation were only localised close to the deletion site. Hence, the above explanations cannot be applied here. However, considering that V_L ΔR108 is less stable than the WT, it is tempting to speculate about dissociation of the C-terminus from the domain. Then, the situation might be quite similar to the I2E mutant, where the N-terminus easily dissociates. It is again a terminal strand, which governs the amyloid aggregation. For β-2-microglobulin, the deletion of the six N-terminal residues enabled fibril formation at native like conditions [207]. Here, modifications at both the N-terminus or the C-terminus might increase amyloidogenicity. Both is in agreement with the protecitve role of edge strands to prevent aggregation of β-sheet proteins [64].

If we expand the system by including the C_L domain, it is important, whether the truncation of the C_L domain happens prior to fibril formation, or afterwards. Findings of full-length LCs in aggregates indicated the latter one [46]. However, the question is still not solved [208]. If the cleavage occurs before the aggregation, then partial cleavage of the constant domain might disrupt interactions of R108. This, in turn, could then link the effects of R108 on both V_L and C_L with the propensity to form amyloids.

As concluding remark to this chapter, we did not identify that one characteristic, which renders all LCs prone to aggregation, since there is not only one such feature. We could, however, rationalise the effects of several mutations enhacing fibril formation. Our explanations match with literature observations, paving the way to a holistic understanding of the variety of LC sequences. So far, we did not reach the point, were we can predict from a sequence, whether an LC is likely to form fibrils. But step by step the scientific community gathers information on hot spot regions, which are linked to AL amyloidosis.

5 Interactions between Soluble MAK33 V_L and Epigallocatechin Gallate

Within the last chapter, we described several mutants of MAK33 V_L, which display different degrees of amyloidogenicity. These can be used now as established system for analysis of interactions with EGCG. As was described in the introduction, EGCG has been shown to prevent toxic effects of AL LCs *in vivo* in *C.elegans* [22]. *In vitro*, fibril formation of the 6aJL2 sequence was inhibited by EGCG [113] and for another sequence the aggregation pathway was modified [209]. EGCG is thus a promising candidate to deal with a variety of different AL sequences. We aim now at elucidating the mechanism, by which EGCG prevents amyloid formation.

5.1 EGCG Induced Precipitation is Uncoupled from Chemical Shift Changes

The first step for such an analysis was an NMR titration with EGCG as ligand. The MAK33 V_L S20N mutant was employed for this experiment, as it displayed already considerable tendency to form amyloid deposits, but at the same time, it was stable enough for convenient handling. The latter fact facilitated reproducibility, as the onset of fibril formation was easy to control.

Addition of a 10-fold molar excess of EGCG to MAK33 V_L S20N resulted only in moderate chemical shift changes (Fig. 31). Still, the effects were well reproducible and also mostly similar for the WT protein (data not shown). Considerable changes were observed for K39 and I58, among others. Stepwise addition of EGCG to the protein solution caused steadily increasing chemical shift changes (Fig. 32a).

At the same time, addition of EGCG induced visible precipitation. Precipitation of only EGCG without protein could be excluded, as EGCG was well soluble in the used buffer. For EGCG alone, concentrations up to 50 mM EGCG resulted in an increase of NMR signals proportional to the concentration, ensuring complete solubilisation of EGCG monomers (data not shown). In contrast, addition of EGCG to the protein led to a decrease in protein signal intensities (Fig. 32b). This indicated coaggregation of EGCG and protein. Without agitation or other perturbations, precipitation occurred slowly. The process of precipitation continued after the titration was finished. As the series of experiments comprised 11 measurement points, with a duration of approximately 45 minutes per point, the whole titration took ca. 8 hours. Upon over night incubation, the fraction of precipitated protein increased by a factor of 2. At the same time, chemical shift perturbations decreased slightly. This probably resulted from changed ratios of
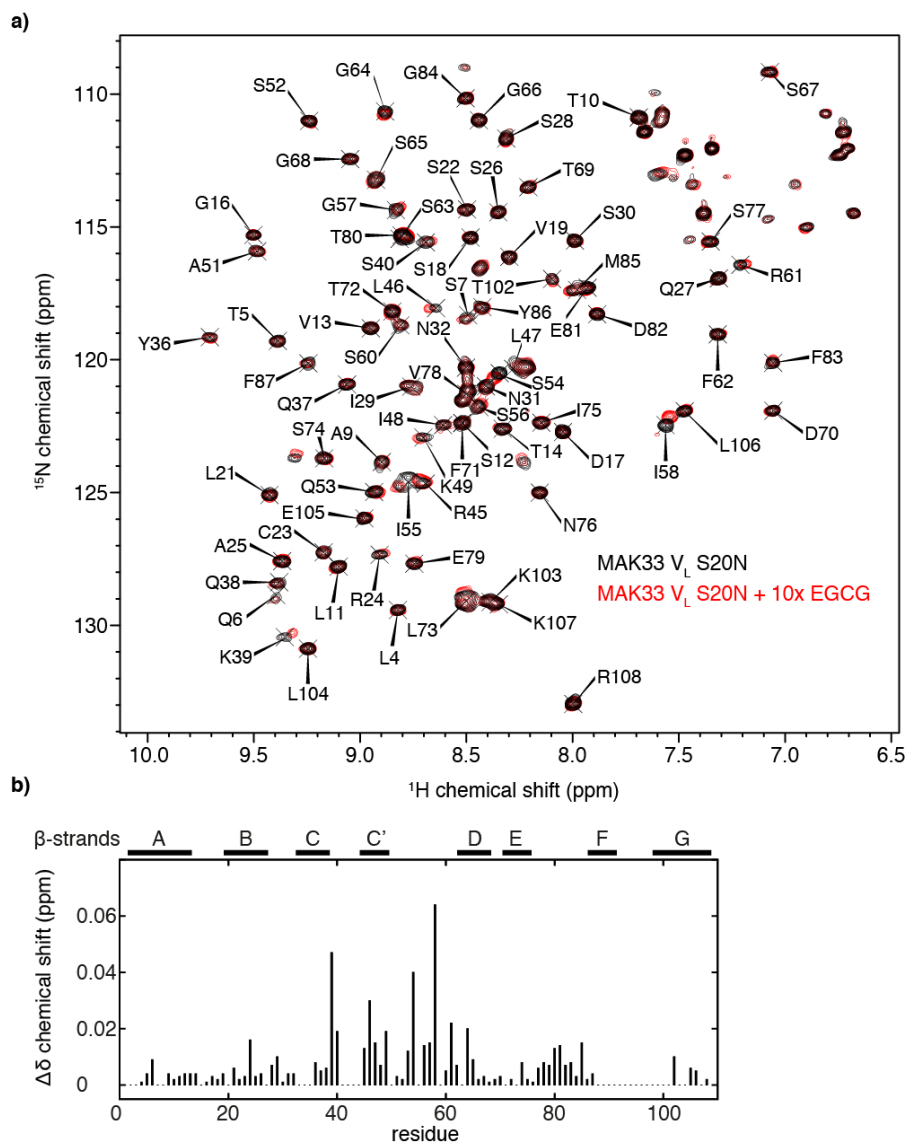
Figure 31: Chemical Shift Changes Induced by EGCG
a) ^{1}H,^{15}N HSQCs of 50 μM MAK33 V$_L$ S20N in presence and absence of 500 μM EGCG.
b) Differences in amide backbone chemical shifts. Gaps indicate resonances, which could not be assigned unambiguously in both proteins or where resonances overlap. Chemical shift changes were small, with the most prominent changes at K39 and I58.

EGCG and protein in solution due to precipiation losses. In this case, the ratio of precipitated EGCG to protein would be larger than 1. An important lesson for further experiments was, that time is critical for the titrations, due to slow reactions, which continuously change chemical shifts and concentrations of soluble components.
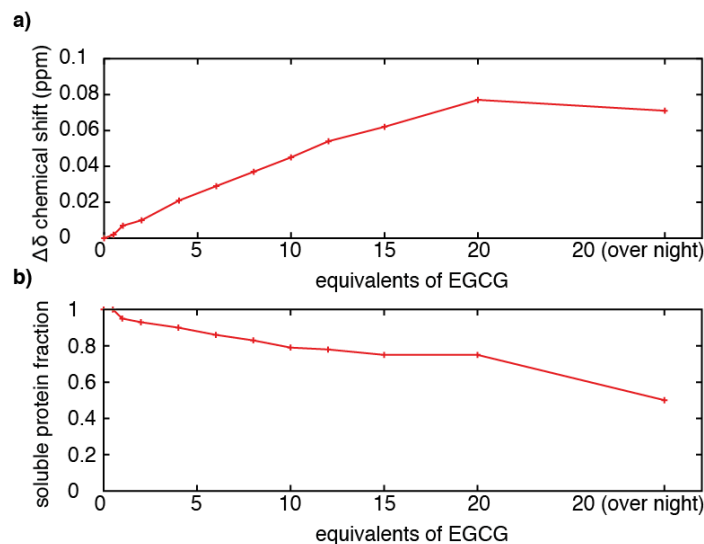
Figure 32: Concentration Dependent Effects of EGCG
EGCG was added stepwise to MAK33 V_L S20N.
a) Build-up of chemical shift perturbations at I58.
b) Reduction of soluble monomers, assessed by 1D NMR signal intensities.

Next, the chemical shift perturbations were examined from a structural point of view. All affected residues are located in the same region of the protein (Fig. 33). They form a potential binding pocket for EGCG. At a first glance, it was surprising, that I58 displayed the largest effect upon addition of EGCG. A likely explanation is the next residue in the primary sequence, P59. EGCG is well known in literature for binding to proline residues [210, 211, 212]. As we used ^{1}H,^{15}N HSQC experiments to assess the effects of EGCG binding, prolines are not represented in such experiments. This is a result of the lacking amide hydrogen atom necessary for a 1-bond ^{1}H,^{15}N-transfer.

In order to confirm the relevance of P59 for binding of EGCG, we mutated it to alanine. As P59 is located in a flexible loop region, the mutation was unlikely to affect the overall structure. Concordantly, thermal transitions with CD spectroscopy were well comparable to the WT (Fig. 34). Melting temperatures were almost identical, with 50.7 $\pm$ 0.2 °C for the WT and 50.0 $\pm$ 0.2 °C for the P59A mutant. ^{1}H,^{15}N HSQC spectra were also highly similar (data not shown), proving that the mutation did not have a large effect on the structure.
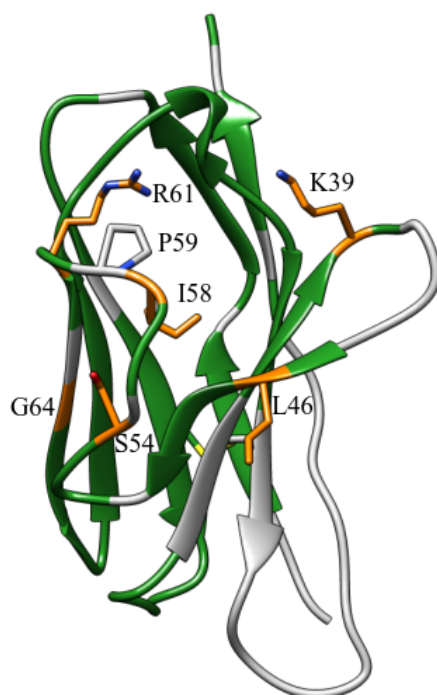
Figure 33: Structural Representation of Chemical Shift Changes upon Addition of EGCG
All residues with chemical shift changes equal to or larger than 0.02 ppm are highlighted in orange. Unassigned residues are shown in grey. The affected residues cluster in a presumed binding site for EGCG. P59 is located within this pocket, next to the strongly affected I58.
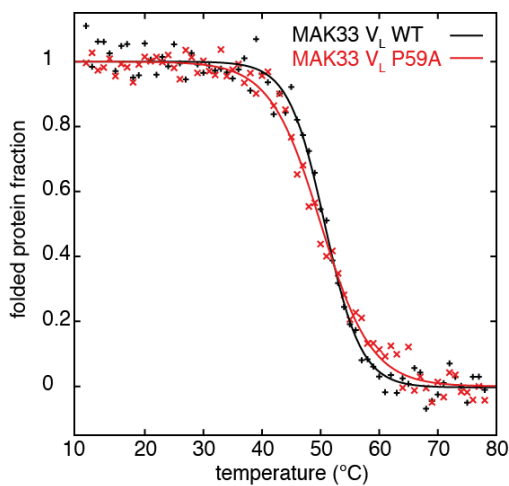
Figure 34: Thermal Transitions observed by Circular Dichroism Spectroscopy
The P59A mutation did not reduce thermal stability of MAK33 V_L.
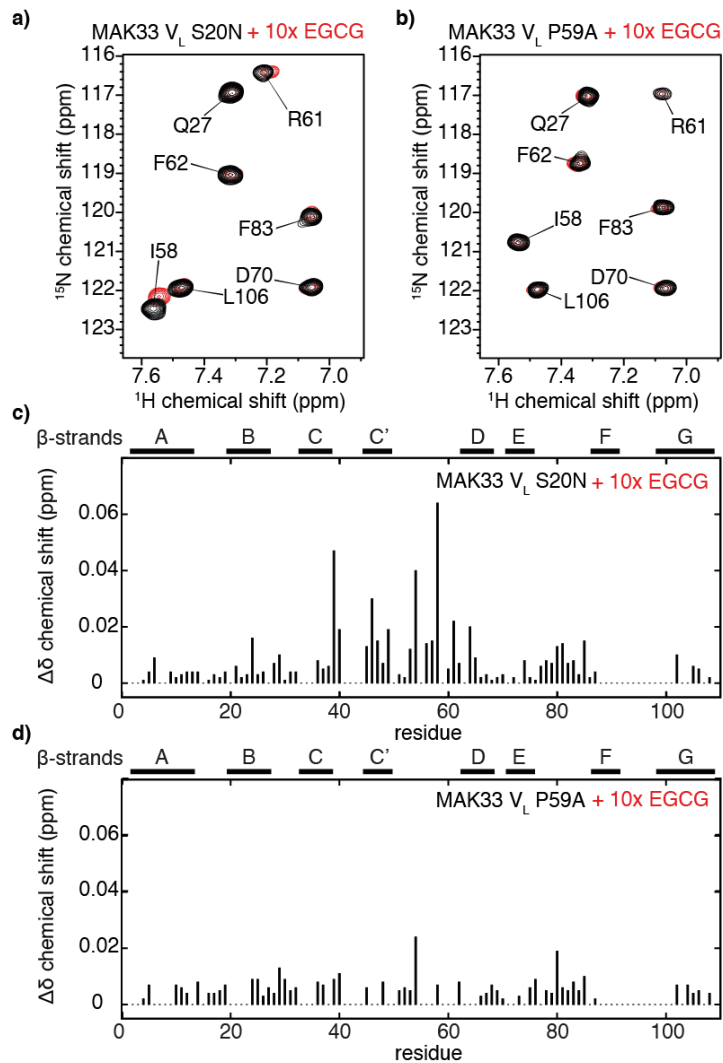
Figure 35: The P59A Mutation Prevents Chemical Shift Changes Induced by EGCG
 a, c) EGCG induced chemical shift changes on MAK33 V_L S20N.
 b, d) EGCG induced chemical shift changes on MAK33 V_L P59A.

Having confirmed, that the P59A mutant is suitable to test interactions between EGCG and P59, we repeated the NMR titration. In contrast to the WT or the S20N variant, here EGCG had only negligible effects on chemical shifts (Fig. 35). Especially I58 and R61 did not shift at all. This confirmed the assumed binding of EGCG to P59. However, the EGCG-induced precipitation still took place, despite the absence of chemical shift changes (Fig. 36). Contrary to the expectations, the mutant precipitated faster than the WT protein. This means, that binding of EGCG to P59 is not relevant for precipitation. Hence, EGCG needs to interact with V_L via another mode, in addition to the binding to P59.
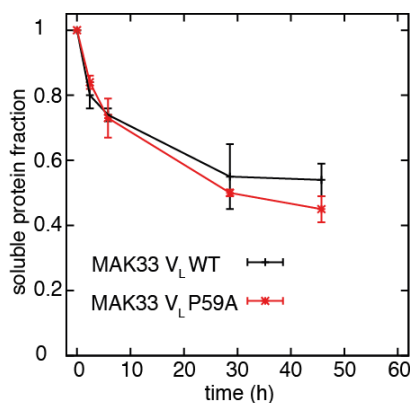
Figure 36: EGCG-Induced Precipitation of the WT and the P59A Mutant
Precipitation was analysed by measuring protein signal intensities in ^{1}H 1D NMR experiments. Error bars represent the standard deviation of triplicate measurements. Although the P59A mutant prevented EGCG-related chemical shift changes, it is still precipitated by EGCG.

5.2 EGCG can Bind to MAK33 V_L via Two Distinct Binding Sites

Further insights were gained from docking experiments and molecular dynamics simulations. These experiments were conducted by Martin Carballo Pacheco from the research group of Birgit Strodel (Forschungszentrum Jülich).

For the first experiments, the docking of EGCG to MAK33 V_L was restrained by NMR chemical shift changes. Based on our previous results, contacts between EGCG and I58 and P59 were predefined. According to the simulation, EGCG bound to the protein, mainly due to hydrophilic interactions with serine residues (Fig. 37a). In the subsequent MD simulation, EGCG fluctuated between different orientations and eventually dissociated. Concordantly, this binding mode was considered as weak interaction only. Binding of EGCG affected dynamics in β-strand F and CDR3, otherwise the protein was not changed (Fig. 37c). This mode of interaction between EGCG and V_L is probably the cause of the chemical shift changes observed by NMR.

Since our experimental data suggested a second interaction mode, the docking experiment was repeated without NMR restraints. This time, EGCG bound to a completely different position (Fig. 37b). Strikingly, also the second binding site involved a proline residue as interaction partner, P44. In addition to H-bonds with the hydroxyl groups of EGCG, this complex was stabilised by π-π-stacking interactions with aromatic amino acids. During the course of an MD trajectory, EGCG remained stably bound to this site, whereas in the previous simulation it dissociated from the P59 site. Hence, this new binding site can be considered to have a higher affinity. Fluctuations of the amino acid positions upon binding of EGCG revealed now two regions of interest (Fig. 37d). Similarly to the P59 site, binding to P44 reduced the fluctuations within β-strand F and CDR3. This can be easily rationalised by the stacking interactions with F87 and F98. It was surprising to find a second affected region, namely the CDR2 loop between β-strand C' and D. This region contains also P59. Structurally, it is adjacent to strand C' with P44. While the actually binding C' strand is not affected by EGCG, it displayed intrinsically large RMSFs. Interactions with EGCG might then perturb interactions with the adjacent loop and thus change the loop's dynamics.

In our previous NMR spectroscopic analysis, we could not observe interactions between EGCG and this P44 binding site. This results from the aforementioned exchange broadening due to homodimerisation of the V_L domain. Thus, by conventional solution-state NMR spectroscopy, this binding site cannot be confirmed. Instead, we focussed on analysis of the precipitates. Transmission electron microscopy showed clearly, that EGCG prevents amyloid fibril formation of V_L domains (Fig. 38). While the protein without EGCG formed long, well structured amyloid aggregates, the sample with EGCG precipitated as compact aggregates without well defined structure.

We collected such precipitates and analysed them by use of MAS solid-state NMR spectroscopy (Fig. 39a, b). The WT protein precipitate with EGCG produced spectra with resolved resonances in the ^{13}C,^{13}C-correlation. Linewidths were rather broad, between 250 and 300 Hz. For homogeneous fibrils, typical ^{13}C linewidths are approximately 100 Hz [6]. This indicates already the mostly unstructured nature of the coprecipitates. As the experiment relied on dipolar transfers, only rigid residues produced observable
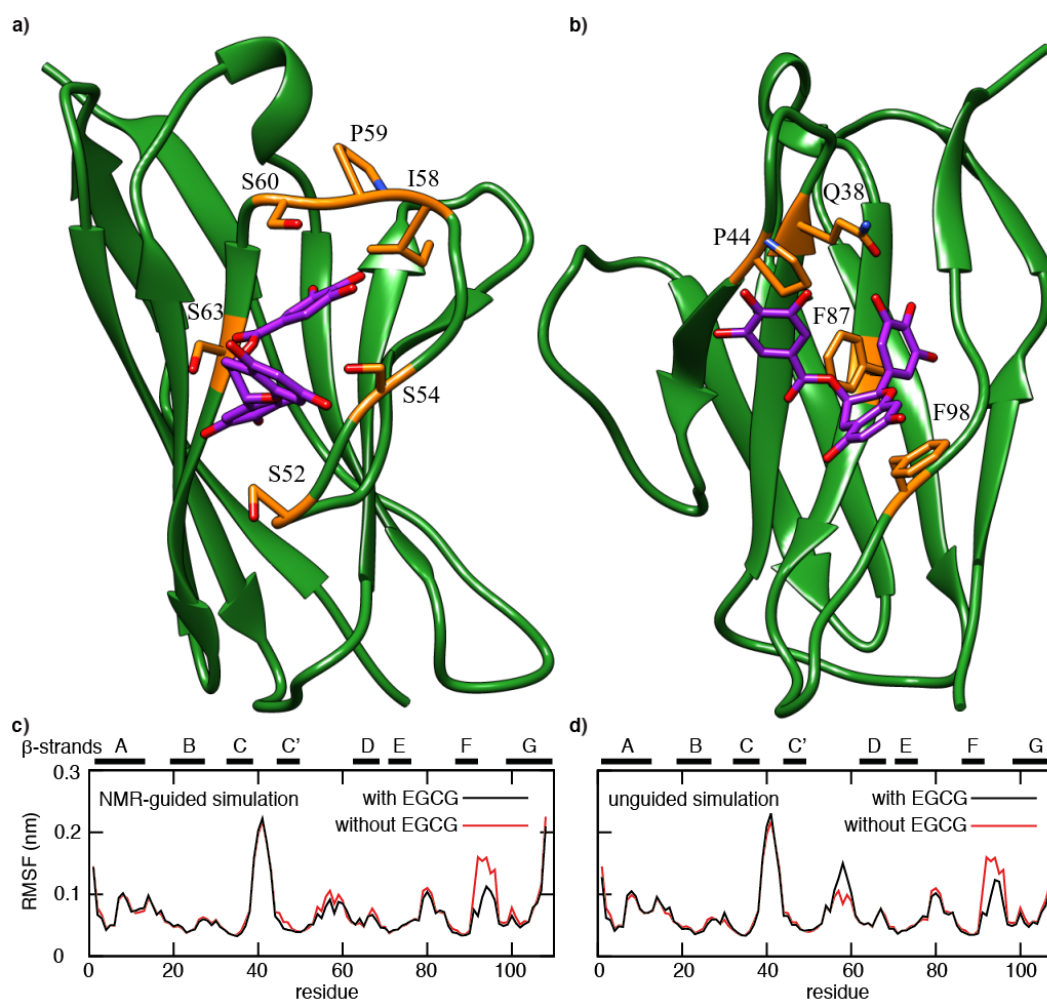
Figure 37: Docking of EGCG to MAK33 V$_L$

    a, c) NMR-guided docking and root mean square fluctuations (RMSF) during MD simulation. For this docking, contacts between EGCG and I58 as well as P59 were pre-defined. EGCG is bound mainly via H-bonds to different serine residues. In presence of EGCG, the dynamics within β-strand F and the subsequent CDR3 were reduced. Otherwise, the binding did not cause changes in dynamics.

    b, d) Unguided docking and RMSF during MD simulation. No experimental restraints were used. EGCG was stabilised by several favourable interactions. These include π-π stacking interactions between EGCG and F87 as well as with F98. The H-bond with Q38 was stable during the course of the simulation. In addition to a reduction of dynamics in β-strand F and CDR3, EGCG also caused increased dynamics around P59.

70

signals. For dynamic residues, dipolar couplings average out. Consequently, those spin systems, which were visible, represent the small structured part of the aggregate. It is reasonable to assume, that these residues were rigidified by interaction with EGCG.
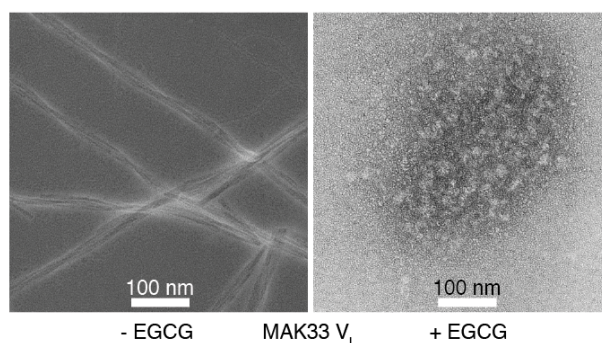
Figure 38: MAK33 V$_L$ I2E in Fibril State or as Unstructured Precipitate
   The I2E mutant was employed, as it can form fibrils at native-like conditions. Except for presence or absence of 10-fold molar excess of EGCG, both samples were treated identically, i.e. incubation at 37 °C with agitation.

Albeit sequential resonance assignment was not feasible with such a resolution, the visible spin systems could be assigned to amino acid types. The observed amino acids include arginine (only side chain), glutamate, isoleucine, leucine, lysine, proline, serine, threonine and tyrosine (only side chain). The proline Cβ and Cγ resonances had a chemical shift difference of 5 ppm. This is characteristic for the trans conformation [213]. Out of the five prolines in the MAK33 V$_L$ sequence, the cis prolines P8 and P95 could thus be excluded, as it was unlikely that the proline ring pucker changed upon binding of EGCG. This left three remaining prolines as conceivable interaction sites: P15, P44 and P59.

Since the P59A mutant was still precipitated by EGCG and the docking experiments suggested interactions with P44, we assumed the observed proline spin system to be the latter one. By making use of the complex structure from the docking experiment, we tentatively assigned the remaining resonances. Therefore, for each amino acid type, we assigned the spin system to that residue, which was closest to EGCG in the structure. With this approach, four spin systems could be assigned: Y36, E42, S43 and R45 (Fig. 39c). For leucine, lysine and threonine, several assignments were conceivable. Only for isoleucine, no close candidate was found. However, the isoleucine spin system consisted of very weak resonances, hence this residue might already be more dynamic or further distant from the binding site. Due to sequential proximity to other assigned residues, I48 might be the correct assignment. The good agreement between observed spin systems and residues close to the binding site is not a prove, but qualifies this interaction site as reasonable to assume. If we assumed other prolines than P44 as binding site, no such good sequence coverage would be obtained (data not shown).
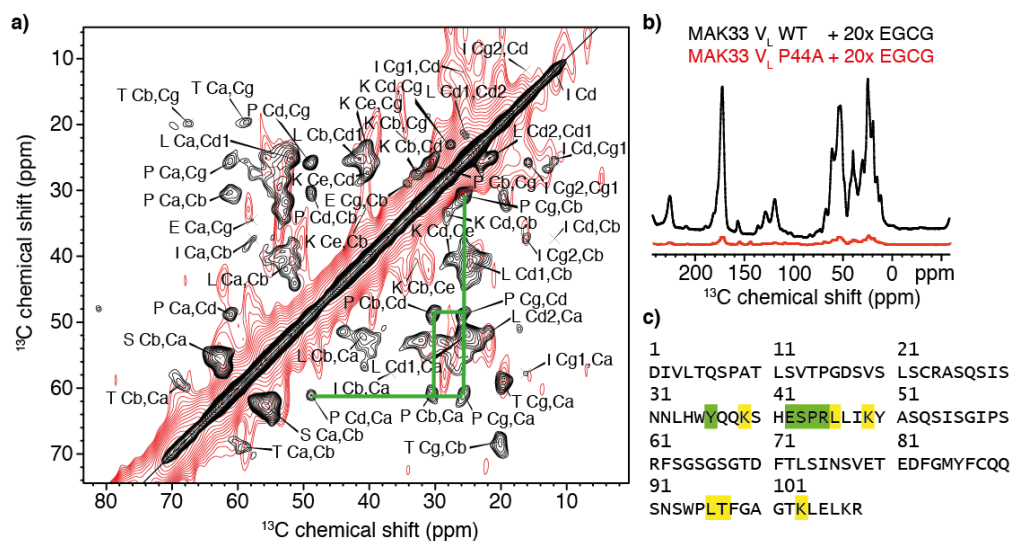
Figure 39: EGCG-Induced Precipitates of MAK33 V$_L$ WT and P44A Analysed by MAS Solid-State NMR Spectroscopy

a) ^{13}C,^{13}C-correlations of the EGCG-induced precipitates with the WT V$_L$ (black) and the P44A mutant (red). The proline spin system in the spectrum of the WT is marked with green lines in the bottom right part of the symmetric spectrum. The experiments were conducted using 50 ms PDSD mixing.

b) Comparison of signal intensities of ^{13}C-1D CP experiments for coprecipitates with both mutants. The intensity of the P44A mutant's spectrum is approximately 15 times less than that of the WT.

c) Mapping of identified spin systems on the amino acid sequence. Green residues were assigned unambiguously, while yellow marks indicate ambiguous assignments. Residues were assigned according to the distance from EGCG in the complex structure from the unguided docking experiment.

In order to further test the assignment of P44, we mutated it to alanine. The consequences were striking. The protein still precipitated, but to a lesser extent. The NMR signal intensity dropped by a factor of ca. 15, corrected for slightly different initial amounts of substance (Fig. 39b). In a ^{13}C,^{13}C-correlation experiment, no spin systems could be observed for the coprecipitate of EGCG and MAK33 V$_L$ P44A, despite three days of measurement time. The drastic effects of the mutation confirmed P44 as crucial for the interaction with EGCG and supplied experimental evidence for the binding mode suggested by the unguided docking experiments.

5.3 Precipitation Kinetics of Different Mutants with EGCG

After we identified the second mode of interaction between EGCG and V_L proteins, we continued to study the precipitation behaviour. In order to learn about the kinetics of amyloid fibril formation and EGCG induced-precipitation, we conducted Thioflavin T (ThT) assays (Fig. 40). ThT is a dye that binds selectively to amyloid fibrils [170]. In the bound state, the fluorescence emission spectrum changes. This makes ThT a common probe for detecting and quantifying amyloid deposits.

Without EGCG, the I2E mutant started immediatelly to form fibrils. The protein is unstable enough to skip the otherwise typical lag phase for amyloid formation. After approximately 10 days, fibril formation was completed and the amount of fibrils remained stable.
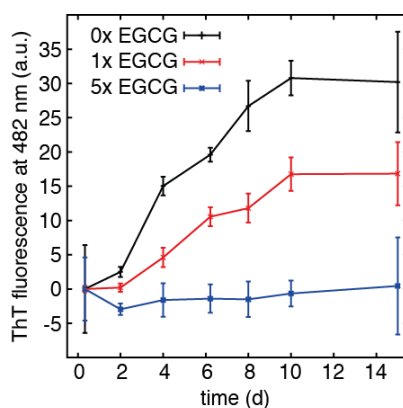
Figure 40: Fibril Formation Kinetics In Presence and Absence of EGCG
Fibrils were formed with MAK33 V_L I2E by agitation at 37 °C. Increasing amounts of EGCG reduced the total amount of amyloid fibrils.

Addition of equimolar amounts of EGCG was sufficient to considerably reduce the amount of fibrils. However, the observation should be treated with caution, as EGCG changed the emission spectrum of ThT also in absence of protein (data not shown). Since we could confirm the unstructured nature of EGCG-induced precipitates already with TEM and by MAS solid-state NMR linewidths, the data were considered to be at least qualitatively reliable. Quantitative interpretations will be avoided due to the interaction between EGCG and ThT. Still, it can be seen that EGCG did not delay the onset of fibril formation. 10-fold molar excess of EGCG prevented fibril formation completely, in agreement with the TEM data (Fig. 38).

Next, we studied the precipitation induced by EGCG (Fig. 41). All of these experiments displayed fast kinetics, which produced considerable effects within two to three days. After comparison with the ThT kinetics, we assume the NMR-based precipitation kinetics here to reflect mostly EGCG-induced aggregation, while fibril formation can be considered mostly negligible within this time scale.

Similar to the ThT kinetics, aggregation was studied as a function of the ligand concentration. In the control experiment, in which water was added instead of EGCG
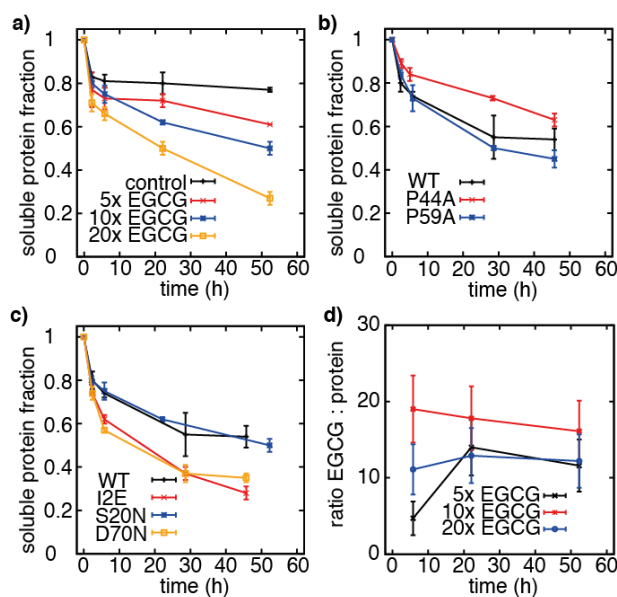
Figure 41: Kinetics of EGCG-Induced Precipitation
    a) Precipitation of MAK33 V_L S20N as a function of EGCG concentration.
    b) Comparison of the precipitation rates of the proline mutants with altered EGCG binding properties. Experiments were conducted with 10-fold excess of EGCG
    c) Precipitation of V_L variants with varying degrees of amyloidogenicity upon addition of 10-fold excess of EGCG.
    d) Stochiometric analysis of the composition of coprecipitates from the S20N mutant with EGCG.
    For all experiments, quantification was achieved by analysis of solution NMR ^{1}H-1D signal intensities. Between the NMR measurements, all samples were incubated at 25 °C with agitation.

solution, the observed protein signals displayed a quick initial decrease (Fig. 41a). This reflected unfolding due to mixing of the protein solution, concordantly with a small foam layer on top of the solution after mixing. This mixing-related signal loss was comparable for all further experiments.

Increasing EGCG concentrations resulted in faster aggregation. A concentration dependent mode of action of EGCG matches the expectations, as aggregation reactions become faster with increasing concentrations.

In another experiment, the proline mutants were compared to the WT sequence with respect to precipitation (Fig. 41b). The WT and the P59A variant have been compared already before (Fig. 36). Including the P44A mutant into this picture still yields consistent data. In agreement with the MAS solid-state NMR analysis, the P44A mutant displayed the slowest precipitation kinetics. The effect of the mutation is, however, not as pronounced as the respective signal intensities obtained from the precipitates by solid-state NMR spectroscopy. While there the amounts of aggregates differed by a factor of 15, the observed effect here is considerably smaller. It might be, that the P44A mutant is

still effectively precipitated by EGCG. However, the aggregates might be less stable and thus less visible by dipolar-based MAS solid-state NMR experiments. Then the presence of a proline at position 44 would still be important for reliable long-term removal of V_L proteins from the soluble fraction. Thus, the reservoir, from which fibrils are formed, could be depleted.

We repeated this experiment with the other V_L mutants studied so far. EGCG-induced aggregation was studied with the I2E, the S20N and the D70N mutants. To briefly summarise the important characteristics about these variants, they differ with respect to amyloidogenicity. Whereas the WT sequence requires pH 2, 37 °C and agitation for fibril formation, the S20N or D70N mutants did not need the acidic pH. Instead, ultrasonication in combination with elevated temperatures and agitation was sufficient [193]. The I2E variant needs only 37 °C and agitation, due to reduced thermal stability [197]. With this prior knowledge, it was expectable that the I2E mutant precipitated faster in presence of EGCG (Fig. 41c). The solid-state NMR spectroscopic experiments confirmed already, that the aggregates are mostly unstructured. It is reasonable that the unstable I2E mutant could easily transition into such an unfolded state.

The S20N mutant behaved identical to the WT V_L sequence. However, the D70N variant surprisingly aggregated equally fast as the I2E mutant. This was especially striking, as the D70N V_L sequence had the same thermodynamic stability as the WT. It is unlikely, that the faster precipitation resulted from specific structural changes induced by the D70N mutation, as positions 44 and 70 are on opposite sides. It seems, as if EGCG can discriminate amyloidogenic species, at least to some extent. While this feature does not apply to all amyloigenic variants, as was shown with the S20N variant, it is for sure an intriguing property for a drug candidate.

Finally, we evaluated the stochiometric composition of the coprecipitates (Fig. 41d). The ratio of EGCG to protein varied a lot, also as a function of the initial stochiometry in solution. Between 5 and 20 EGCG molecules were found to aggregate with one protein. The large fluctuations support the model of unstructured aggregates. Once the protein has unfolded, EGCG molecules could just accidentally associate with previously protected binding sites.

5.4 Conclusions on Interactions between EGCG and Amyloidogenic V_L Domains

In summary, we characterised two different modes of interaction between EGCG and MAK33 V_L. Both binding sites involve prolines as crucial interaction partners. While we focussed on a κ-type LC, similar studies were conducted already for the λ V_L domain 6aJL2 [113]. The results are complementary. While the binding to P59 was only observed for the MAK33 sequence, the P44 binding site was also visible in 6aJL2. There, the respective residues were visible by solution-state NMR spectroscopy, as dimerisation-related exchange broadening was not a concern for that sequence. The residues involved in binding of EGCG were similar regarding the sequences. This prompted us to align the respective sequences and analyse the conservation of the important residues. To obtain a more complete picture, we included also the 1AL1 sequence for this analysis. This V_L sequence originated from an AL amyloidosis patient. *In vitro* studies showed, that EGCG also modulated the aggregation behaviour of this sequence [209].
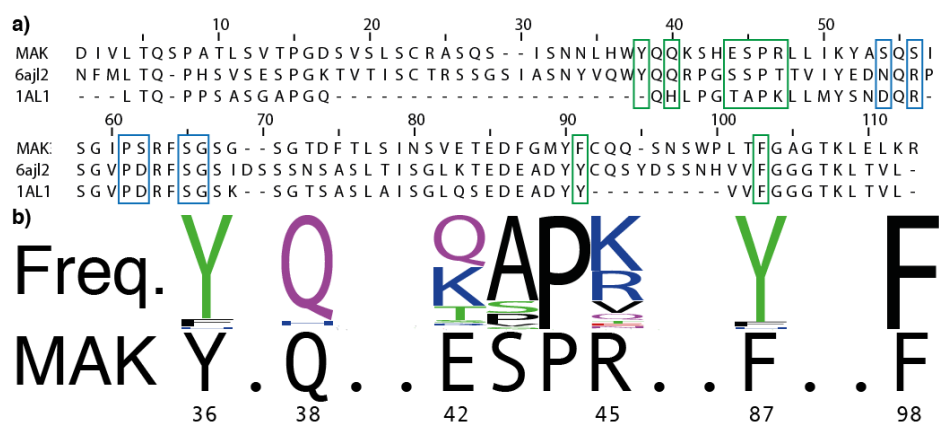
Figure 42: Conservation of EGCG Binding Elements in V_L Sequences

a) Sequence alignments of MAK33 V_L, 6aJL2 and 1AL1. All sequences bind to EGCG. Important residues for binding to the P44 site are highlighted in green, whereas interacting residues at the P59 site are marked in blue. The 1AL1 sequence information was kindly provided by Kathrin Andrich.

b) Conservation of those amino acid positions in V_L domains, which are important for binding of EGCG to the P59 site. Residues were considered as important, if they were identified by MAS solid-state NMR spectroscopy of the precipitates, or if the contact probability according to MD simulations was at least 80%. The conservation was calculated using amino acid frequencies from more than 11,000 human LC sequences, which were deposited in abYsis 2.3.3 (http://www.bioinf.org.uk/abysis/).

According to the sequence alignments, both the P44 site and the P59 site are conserved in the analysed sequences (Fig. 42a). With respect to the P59 binding site, S52, S54 and S60 (amino acid numbering of MAK33 V_L) are only present in the MAK33 sequence, while P59, S63 and G64 were conserved in all three sequences. More importantly, the P44 site, which is crucial for precipitation, displayed considerable conservation. This was not

only the case for the three sequences, for which experimental evidence for interaction with EGCG is available, but also for a broad range of more then 11,000 human V$_L$ sequences (Fig. 42b). Specifically, in addition to P44, all aromatic residues are conserved, i.e. Y36, F78 and F98. Q38 is highly conserved, only in 1AL1 it is replaced by the also hydrophilic histidine. Position 45 (arginine in MAK33 V$_L$) is a basic residue in most V$_L$ sequences. Positions 42 and 43 are more diverse in the comparison of the 11,000 sequences, but since this is also the case in the alignment with 6aJL2 and 1AL1, these residues seem to be less crucial for binding of EGCG. Based on these information, we consider EGCG to be able to bind to many different V$_L$ sequences.

For its use as a drug candidate, mechanistic understanding of EGCG's mode of action is important. Based on our experimental results, we suggest a model for the interaction with amyloidogenic V$_L$ proteins. In absence of EGCG, the protein domains can form amyloid fibrils via partially unfolded intermediates [52, 53, 54, 55]. If EGCG is added, it can interact both with the P59 site and the P44 site. The P59 binding site has the weaker affinity, as indicated by our MD simulations. Although this binding site is the cause of the observed chemical shift changes, it is not relevant for precipitation.
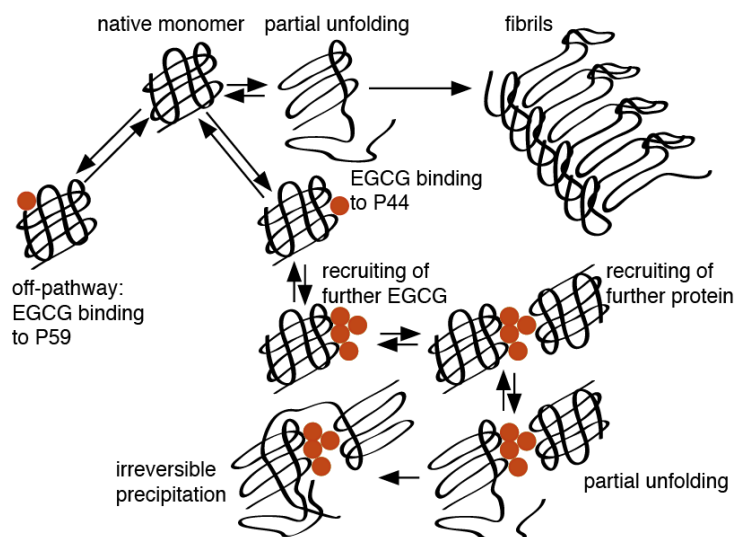
Figure 43: Mechanism for the Interaction between EGCG and Amyloidogenic V$_L$ sequences
Without interaction with EGCG, V$_L$ domains transform via intermediate states to mature amyloid fibrils. The native EGCG monomer can bind EGCG at the P59 site. However, as this interaction does not lead to precipitation, it is not considered further. Alternatively, EGCG can bind at the P44 site. Since EGCG has intrinsic tendencies to form homo-oligomers, it can act as molecular glue for several V$_L$ proteins. As amyloidogenic V$_L$ sequences tend to adopt partially unfolded states, the increased local protein concentration then favours unstructured aggregation.

Binding to the P44 site, in contrast, initiates unstructured aggregation. EGCG often forms homo-oligomers [214]. Thus, several proteins, which are already associated with EGCG, can form complexes mediated by EGCG oligomerisation. Such a model is in agreement with reported 1:1 stochiometries for initial binding of EGCG to V$_L$

proteins [113]. Only after the initial binding events, further EGCG molecules would be recruited by its self-association property. Later on, protein molecules within these complexes might unfold, similar to the unfolding events prior to amyloid formation. During such unfolding events, cryptic binding sites for EGCG might be exposed, explaining the fluctuations observed in the stochiometry of the final precipitates. In addition, such unfolding events could expose hydrophobic patches. As the local protein concentration is increased due to EGCG-mediated complexes, this paves the way for unstructured aggregation. A similar mechanism for EGCG-induced precipitation was suggested for interactions between EGCG and β-casein [215].

The key feature here is the partial unfolding event. As this is vital for amyloid formation, this allows EGCG to discriminate amyloidogenic species from other LCs. Such a selectivity of EGCG for reaction with amyloidogenic mutants was recently also reported for TGFBIp, which is related to lattice corneal dystrophy [216]. Despite EGCG's reputation of binding promiscuously to proteins, it might be the right drug candidate to deal with the diverse AL sequences. The recognition elements in the V$_L$ proteins are conserved, while amyloidogenic species are preferred targets for the precipitation. Another protective feature for stable LCs is the location of the binding site. As it is found in the dimer interface, intact antibodies cannot be precipitated by EGCG. Also V$_L$ homodimers, which are less prone to amyloid formation, are protected against unspecific interaction with EGCG [62, 63].

Altogether, the *in vitro* characterisation of EGCG as small molecule compound to treat AL amyloidosis yielded promising results. First *in vivo* indications for its effectivity are already available [22, 109]. The pharmaceutical use for AL amyloidosis patients is currently adressed by the clinical trials EpiCardiAL and TAME-AL.

6 Analysis of the Fibrillar State of MAK33 V_L

The previous parts of this thesis focussed on the native state of MAK33 V_L. We adressed several mutations, which enhanced fibril formation. In the context of EGCG as therapeutical agent, the pathway to fibril formation via intermediates was briefly discussed. The topic of this chapter is the structure of the fibrils themselves. All these topics are closely related. While the native state obviously can have characteristics, which enhance fibril formation, understanding of the fibril state is crucial as well for elucidating, how and why some LC sequences form amyloid aggregates. Knowledge about the fibril structure could enable rational design of strategies to interfere with the formation itself or for destabilisation of already present fibrils to assist clearing. Furthermore, characterisation of the fibrils and intermediates allows to tackle the question, which species is most toxic to the patients.

Up to date, there is hardly any information available on the fibril structure. While fibril features have been discussed on the basis of low-resolution negative stain TEM images, there are no studies available that characterised LC V_L fibrils at high resolution.

As was discussed in the introduction, the first step for MAS solid-state NMR spectroscopic analysis is resonance assignment. In contrast to solution-state NMR spectroscopy, where assignment experiments are well established, solid-state NMR spectroscopy requires thorough optimisation of the sample preparation and often many complementary assignment experiments. This is a consequence of the lower sensitivity due to ^{13}C-detection. In addition, fibrils typically yield larger linewidths than crystalline samples, due to heterogeneity.

6.1 Preparation of Fibril Samples of MAK33 V_L S20N

In consequence, we carefully optimised the sample preparation. Solid-state NMR analysis sets three obectives to this preparation: 1) The sensitivity must be high. This is achieved by maximising the amount of substance within the MAS rotor. 2) Resolution must be good. Therefore, it is crucial to avoid mixing of polymorphs. In addition, the fibrils should be kept at conditions, were they are stable. Otherwise, the samples might degrade and become heterogeneous. As was discussed in the introduction, resolution is also linked to the sensitivity, as polymorphism spreads the signal over a larger area, thereby reducing the absolute signal intensity. 3) The preparation must be well reproducible. This is a special problem of amyloid samples. Due to the aforementioned polymorphism, it seems like hardly any solid-state NMR spectrum of one amyloid protein is identical for different laboratories (a comparison of different chemical shifts for Aβ fibrils is given in [217], SI). In some cases, even identical preparation conditions within the same laboratory yielded different polymorphs of amyloid protein fibrils. In order to

achieve these three goals, we used seeding of amyloid fibrils. As was explained briefly in the introduction, seeding can be used to imprint the structure of one polymorph on a fresh preparation. At least seven generations of continuous seeding were used to prepare our MAK33 S20N V_L fibril samples used for MAS solid-state NMR analysis. Spectra were well reproducible, then. For all spectra shown later on, chemical shifts of different sample preparations could be well superimposed.

We tried to increase the amount of substance of protein in the rotor by lyophilising the fibrils [218]. The rationale is to reduce the amount of water, thereby increasing the available space for protein. Since freeze-drying can impair the structure, it is advisable to rehydrate samples afterwards, only with less water than it was previously present. A water content of 36 to 50% yielded good signal enhancements at high resolution for α-synuclein fibrils [218].
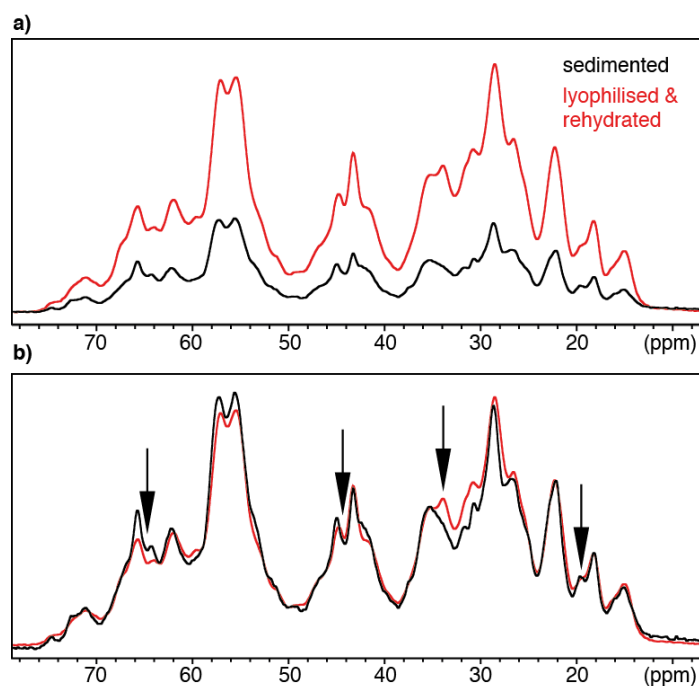
Figure 44: Comparison of Sedimented and Rehydrated MAK33 S20N V_L Fibrils
a) ^{13}C-1D CP spectra at same scale. The lyophilised and rehydrated sample had approximately 2.7x more signal to noise.
b) The same spectra are shown scaled to similar intensities to ease comparison of the resolution. In detail, the spectra displayed some differences, as highlighted by the arrows. The sedimented fibrils produced a spectrum with smaller linewidths.

We prepared a sample from 21 mg lyophilised fibrils, rehydrated with 32 μg acidic water. The water was acidified in order to stabilise the MAK33 V_L S20N fibrils, which were formed at pH 2. In order to guarantee equal distribution of the water, the sample was rehydrated stepwise. A small amount of approximetely 5 mg dry fibrils was filled into the rotor with an ultracentrifuge, then water was added on top. Subsequently, this

procedure was repeated, until the rotor was completely filled. Thereby, the amount of water could be reduced a lot. While only 14 mg of sedimented fibrils fitted into a 3.2 mm thin wall rotor with one teflon spacer, 21 mg (dry mass) of rehydrated protein could be filled into the same volume.

The signal improved even further. The lyophilisation and rehydration increased the signal to noise by a factor of approximately 2.7 (Fig. 44). However, the spectrum changed, compared to a conventionally packed sample. While ^{13}C linewidths were similar for both samples, small differences could be observed in the resonance chemical shifts. The picture was worse, when we compared ^{13}C,^{15}N-correlations (Fig. 45). ^{15}N resonances are typicalls more sensitive than ^{13}C resonances with respect to sample quality. The resolution in the ^{15}N dimension was worse than that of the sedimented fibril sample. In addition, some resonances shifted.
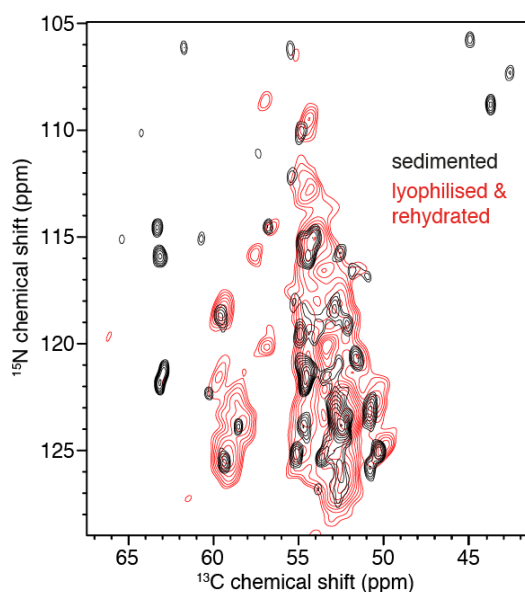
Figure 45: ^{13}C,^{15}N-Correlation of Rehydrated MAK33 S20N V$_L$ Fibrils
The N(CA)CX of sedimented fibrils is superimposed with the NCA of the lyophilised and rehydrated fibrils. Acquisition times in the ^{15}N dimension were 14 ms for the sedimented sample and 20 ms for the rehydrated sample.

Albeit the improvement of the signal-to-noise ratio was still ca. 1.5 in the NCA experiment, we decided not to use lyophilised samples for further analyses. The increased sensitivity does not compensate for the loss of resolution. In addition, the change of chemical shifts raises the question, whether the rehydrated sample contained still intact fibrils. While lyophilising and rehydration may work well for other amyloid fibril forming proteins, it caused more drawbacks than benefits in this case.

6.2 Experimental Strategies for Dipolar-Based Resonance Assignment

All resonance assignment experiments for the MAK33 V_L S20N fibrils were conducted using 3.2 mm rotors. Hence, the experiments relied on ^{13}C detection, as ^{1}H detection is demanding at spinning frequencies slower than 20 kHz. For ^{13}C detected assignment experiments, several strategies have been demonstrated. An excellent description of many MAS solid-state NMR assignment experiments was published by Schütz et al. [176]. As many of these experiments yield complementary information, we applied several different experiments. The assignment experiments used here comprise:
- NCACX & NCOCX (both as 2D and as 3D)
- NCOCA & CANCO (3D)
- NCACB (2D and 3D) & N(CO)CACB (2D)

For further description of the results and advantages of the different approaches, the NMR spectroscopic characteristics of MAK33 V_L S20N fibrils shall be briefly described. The ^{13}C T_2 relaxation times are comparable to those of Aβ fibrils. The ^{15}N T_2 times are far better than observed for other amyloid samples. According to spin echo experiments, the average ^{15}N T_2 time is 9 ms. For a fibril sample, this indicates a high level of structural homogeneity.

6.2.1 2D and 3D Experiments for Assignment

Typically, assignment relies heavily on 3D experiments. The additional dimension greatly enhances spectral dispersion. For NCACX, NCACB or N(CO)CACB type experiments, this requires good dispersion of the ^{13}Cα resonances. Except for the isoleucines and threonines, most ^{13}Cα resonances in the MAK33 V_L S20N fibrils displayed only little dispersion (Fig. 46). Most ^{13}Cα resonances are spread between 51 and 55 ppm.
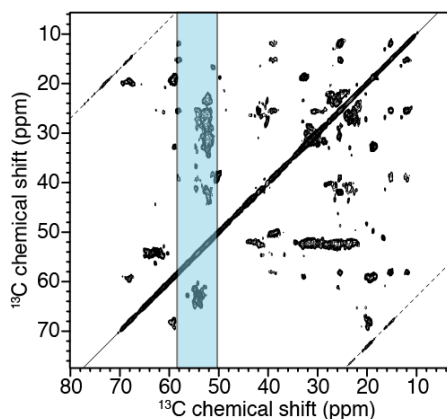
Figure 46: Dispersion of ^{13}Cα Resonances
   In a PDSD experiment, the ^{13}Cα resonances (highlighted in blue) showed only little dispersion.

At a spectrometer, which operates at 750 MHz ^{1}H Larmor frequency, an indirect acquisition time of 7.8 ms in the ^{13}Cα dimension allows for an acquisition-limited linewidth of 130 Hz. In practice, resolution was already sample-limited with this setup, with apodised linewidths of approximately 200 Hz, corresponding to 1 ppm. Considering now the small dispersion, 3D NCACX type experiments have only little benefits compared to N(CA)CX 2D experiments.

The 2D versions of these experiments, on the other side, had much higher signal-to-noise ratios. In detail, we acquired a 3D NCACX with 94 points in the ^{13}Cα dimension, corresponding to 7.8 ms evolution time. The S/N ratio of the 3D experiments is approximately by a factor of 1.7 higher than that of the first N(CA)CX 2D plane of the same experiment (Fig. 47). Corrected for the longer measurement time for 94 points in the indirect dimension, the 2D had approximately 3.4 times higher sensitivity than the 3D experiment. This results from the measurement time spent on indirect points at long acquisition times, were the S/N ratio is low. In the 2D version, the same time can be spent on acquiring more transients in order to improve the S/N ratio.
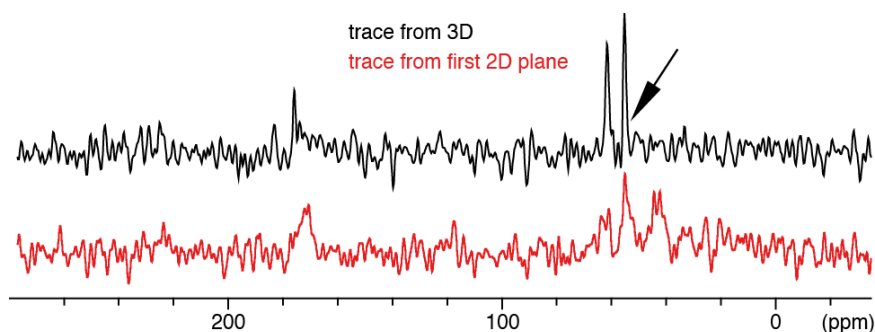
Figure 47: S/N Ratio of 2D and 3D experiments
The shown 1D traces originate from a 3D NCACX and from the first 2D plane of the same NCACX. S/N ratios were compared for the resonance marked with an arrow. The 3D experiment had approximately 1.7 times higher sensitivity. At the same time, the required measurement time increased by a factor of 94 for 7.8 ms evolution time on ^{13}Cα.

While 3D experiments benefit from better dispersion, 2D experiments are a serious alternative for resonance assignment, given the characteristics of our fibril sample. The small dispersion of ^{13}Cα in this case and the small dispersion of ^{13}CO in general reduces the advantage of 3D experiments. At the same time, the improved sensitivity of 2D experiments can make more resonances visible. We emphasise, however, that this approach relies on the good resolution in the ^{15}N dimension. To make best use of the long T_2 times on ^{15}N, 2D experiments were run with approximately 15 ms evolution time. This resulted in excellent resolution. With an acquisition-limited resolution of 70 Hz, we could achieve actual apodised linewidths of 70 to 90 Hz, or 0.9 to 1.1 ppm. Thus, most resonances could be resolved already in the 2D experiments (Fig. 48). Compared with the 3D version with shorter ^{15}N acquisition times of 6.6 ms, the S/N ratio was still higher by a factor of approximetely 2.8. This allowed for assignment of many resonances, which

were invisible in the 3D experiment. 3D NCACX type experiments were still measured, in order to determine precise $^{13}C\alpha$ chemical shifts, which were partially obscured due to overlap. The considerations regarding 2D and 3D experiments are also valid for the NCOCX, as dispersion of ^{13}CO is typically smaller than that of $^{13}C\alpha$.
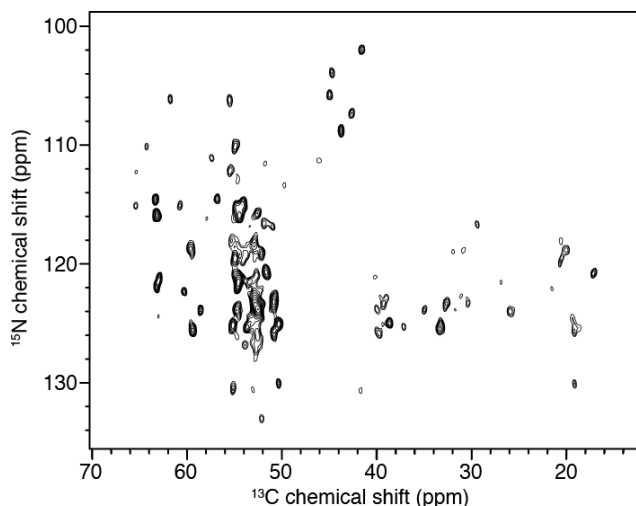
Figure 48: Spectral Resolution and Dispersion in 2D N(CA)CX Experiments.
The spectrum was acquired with 14.4 ms indirect evolution time on ^{15}N. Most resonances can be resolved in the 2D experiment, even if not all of them are baseline-separated. Several weak resonances can be seen, which vanish in the 3D version of this experiment.

6.2.2 Pulse Sequences

In the following section, the advantages of different pulse sequences will be briefly discussed. While general features of these experiments are already described elsewhere [176], we will focus specifically on the use on MAK33 V_L S20N fibrils. NCACX and NCOCX were very helpful, as they yield chemical shift information for all side chain atoms. This made them especially useful in the 2D experiments, when sequential connectivities could be established not only based on poorly dispersed $^{13}C\alpha$ or ^{13}CO resonances. Side chain resonances beyond $^{13}C\beta$ turned out to be precious for unambiguous assignments. This was especially the case for e.g. isoleucines or valines. In general, however, NCACX and NCOCX were most useful for amino acids with long aliphatic side chains, since aromatic side chains were typically invisible even after e.g. 70 ms of DARR mixing. As a side note, the use of either PDSD or DARR had only small effects, according to our observations. At higher MAS frequencies, i.e. 17 kHz, 50 ms DARR performed slightly better than 50 ms PDSD, although the difference was mostly negligible.

The NCACB and N(CO)CACB experiments generally contain similar information as NCACX and NCOCX. The motivation for NCACB type experiments is the higher sensitivity due to the DREAM transfer, instead of the rather inefficient PDSD or DARR

mixing. In general, we could confirm the better performance of NCACB compared to NCACX (Fig. 49). The first one resulted in sensitivity gains of up to factor of 2. However, the precise positioning of the carrier frequency was critical for the performance of DREAM [219]. A carrier in the center of the $^{13}C\alpha$ region produced the best results, as it enabled transfers also to the many downfield serine and threonine $^{13}C\beta$ resonances. Still, several resonances, which were visible in the NCACX, did not appear in the NCACB, since the DREAM could not cover the whole spectral region of interest with good performance. As many $^{13}C\alpha$ resonances vanished in the NCACB, 3D NCACB experiments were helpful to obtain $^{13}C\alpha$ chemical shifts.

The N(CO)CACB, in contrast, suffered a lot from the additional transfer step. With our setup and sample, the N(CO)CACB performed worse than the NCOCX. Hence, while NCACB and NCACX can be complementary, we advice against the use of N(CO)CACB experiments.
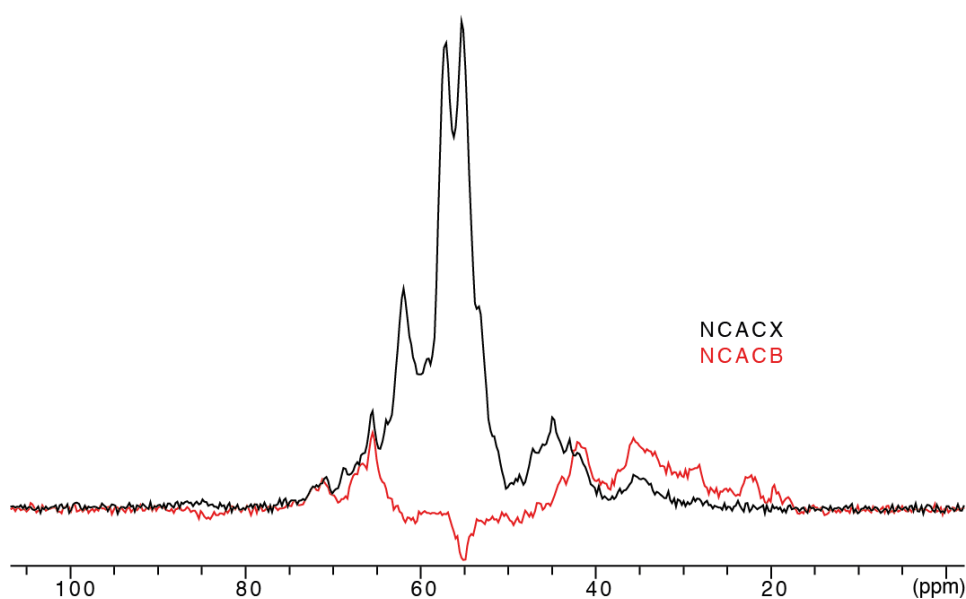
Figure 49: Sensitivity of NCACX and NCACB
Due to the DREAM double quantum transfer, the $^{13}C\alpha$ resonances have opposite signs in the NCACB. The residual $^{13}C\alpha$ signal depended strongly on the carrier frequency and mixing time for the DREAM step. At ca. 50 ppm, $^{13}C\alpha$ and $^{13}C\beta$ signals canceled out each other in the 1D. For signals between 40 and 30 ppm, the NCACB performed better than the NCACX by approximately a factor of 2.

The NCACO, NCOCA and CANCO experiments, finally, rely completely on backbone resonances. NCACO and NCOCA use the BSH-CP for $^{13}C\alpha$-^{13}CO mixing. Again, the motivation is to avoid the unefficient PDSD or DARR mixing. According to literature, BSH-CP can achieve a transfer efficiency of 33% and perform better than PDSD-based NCOCX by 100% [177]. We cannot guarantee, that our implementation of BSH-CP achieved the maximim possible efficiency. However, the transfer from $^{13}C\alpha$ to ^{13}CO had the same efficiency for both BSH-CP and PDSD (50 or 100 ms mixing time). Hence,

we used NCACX instead of NCACO. The performance was better in the opposite direction. NCOCA had a higher sensitivity than NCOCX with our setup and sample. The NCOCA had approximately 30% more S/N. The CANCO provides completely complementary information, which could not be obtained with any other of the aforementioned experiments. It is thus valuable also in combination with NCACX and NCOCX, as it can be used to doublecheck assignments. For spin systems with weak intensities, we could assign resonances by use of NCACX, NCOCA and CANCO. While the sensitivities were good, this approach is limited by resolution, since both $^{13}C\alpha$ and ^{13}CO had little dispersion. Hence, the experiments were suited to extend existing assignments towards more dynamic residues with weaker signals.

6.2.3 Nonuniform Sampling

NUS was presented already in the introduction as a method to enhance sensitivity [134]. The motivation for NUS is to sample more indirect points at short evolution times, where the S/N ratio is high. At long indirect acquisition times, only few points should be sampled to improve the resolution. We used NUS for several 2D and 3D assignment experiments. Especially the long T_2 time of ^{15}N made NUS attractive.

The sampling schedule for NUS was determined by a decaying exponential function. We tried matched NUS with a sparsity of 50%, as described in the literature [134]. Matched sampling refers to a decay rate of the sampling density function equal to the transversal relaxation time constant. We tested this settings with the actual ^{15}N T_2 being 9 ms and a total evolution time of 15 ms (Fig. 50).
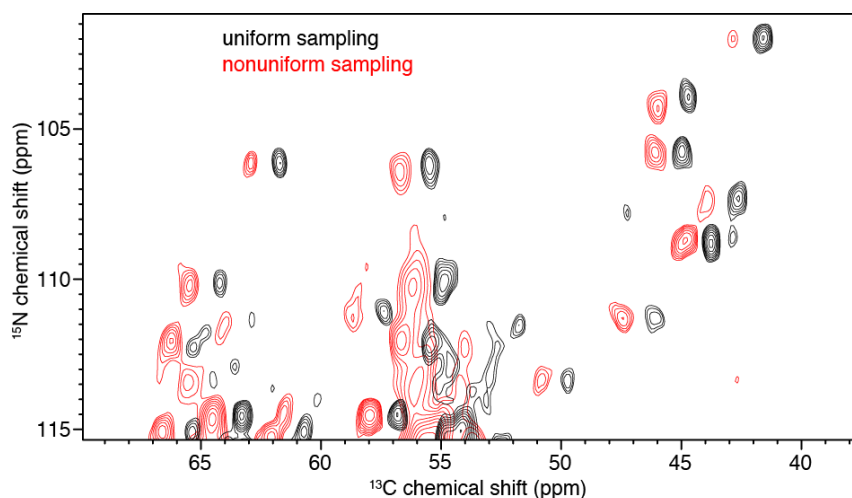
Figure 50: Performance of Nonuniform Sampling
An NCACX was acquired with uniform sampling and with 50% matched NUS. The indirect acquisition time was 15 ms. The NUS spectrum was processed with the compressed sensing option of the TopSpin NUS plugin. To enable comparison of the peak shapes, the NUS spectrum was shifted in the ^{13}C dimension.

We observed no artifacts due to NUS. The quality of the reconstruction differed for different resonances. This might be caused by different relaxation times. In average, the ^{15}N linewidths increased by $18 \pm 22\%$. Effects on the sensitivity varied even stronger for different resonances. We observed improvements of the S/N ratio by a factor of 3.3 ± 1.4. The huge gains in sensitivity convinced us to use NUS for further assignment experiments. However, the results show that relative signal intensities were not conserved by NUS. While NUS was shown to produce spectra, in which relative intensities were maintained, results with our setup were not suited for quantification. Hence, it is adviseable to record uniformly sampled experiments as reference. While we did not quantify the effects on NUS for 3D experiments, we expect even larger sensitivity enhancements there.

6.2.4 Selective Labeling

In ^{13}C,^{13}C-correlations, the linewidth is limited by evolution of scalar coupling, as described in the introduction. Selective labeling with 2-^{13}C$_1$ glycerol or 1,3-^{13}C$_2$ glycerol can circumvent this problem. With selective glycerol labeling, adjacent carbon atoms are rarely both labeled at the same time [150].
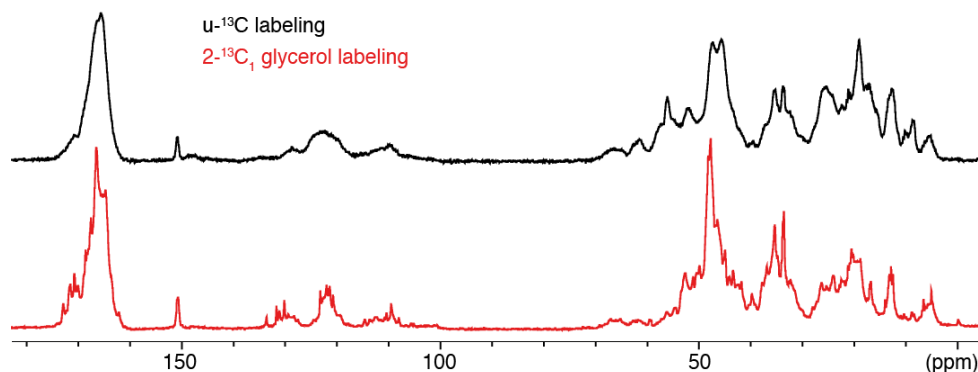
Figure 51: Increase in Resolution due to Selective Labeling
u-^{13}C,^{15}N and 2-^{13}C$_1$ glycerol,^{15}N labeled MAK33 V$_L$ S20N fibrils are compared. An acquisition time of 45 ms was used. No apodisation was applied.

The resolution of the ^{13}C dimension improved considerably by 2-^{13}C$_1$ glycerol labeling (Fig. 51). Especially in the aromatic region (110 to 130 ppm), the smaller linewidths can be appreciated already in the 1D experiment. While there were no isolated signals in the ^{13}C 1D of the uniformly labeled sample, we measured methyl linewidths of 45 Hz for the 2-^{13}C$_1$ glycerol labeled sample.

The drawback of selective labeling with glycerol as ^{13}C source was a reduction of the S/N ratio. This was unexpected and not described in literature. With selectively labeled glycerol as nutrient and sodium bicarbonate as additional carbon source, roughly 50% of the carbon atoms should be labeled. However, the S/N ratio dropped approximately by a factor of 8. This was reproducible for several preparations of both 2-^{13}C$_1$ glycerol or 1,3-^{13}C$_2$ glycerol labeled samples. Solution-state NMR spectroscopic analysis of the

two glycerol types confirmed the correct labeling patterns. Labeled respectively unlabeled sodium bicarbonate was added to prevent isotope scrambling, as described in the literature [220]. In theory, one would expect an increase in sensitivity, as the intensities should increase upon removal of scalar coupling. It remains to be seen, whether this loss of sensitivity will be observed also for other systems.

Nevertheless, we could obtain new information from the selectively labeled samples. Only with such samples, we could observe correlations between aromatic and aliphatic resonances (Fig. 52). In an overlay with spectra from a uniformly labeled sample, this enabled the identification of aromatic residues. Furthermore, selective labeling was used to confirm the assignment of residue types, as the labeling results in characteristic patters for different amino acids [150].
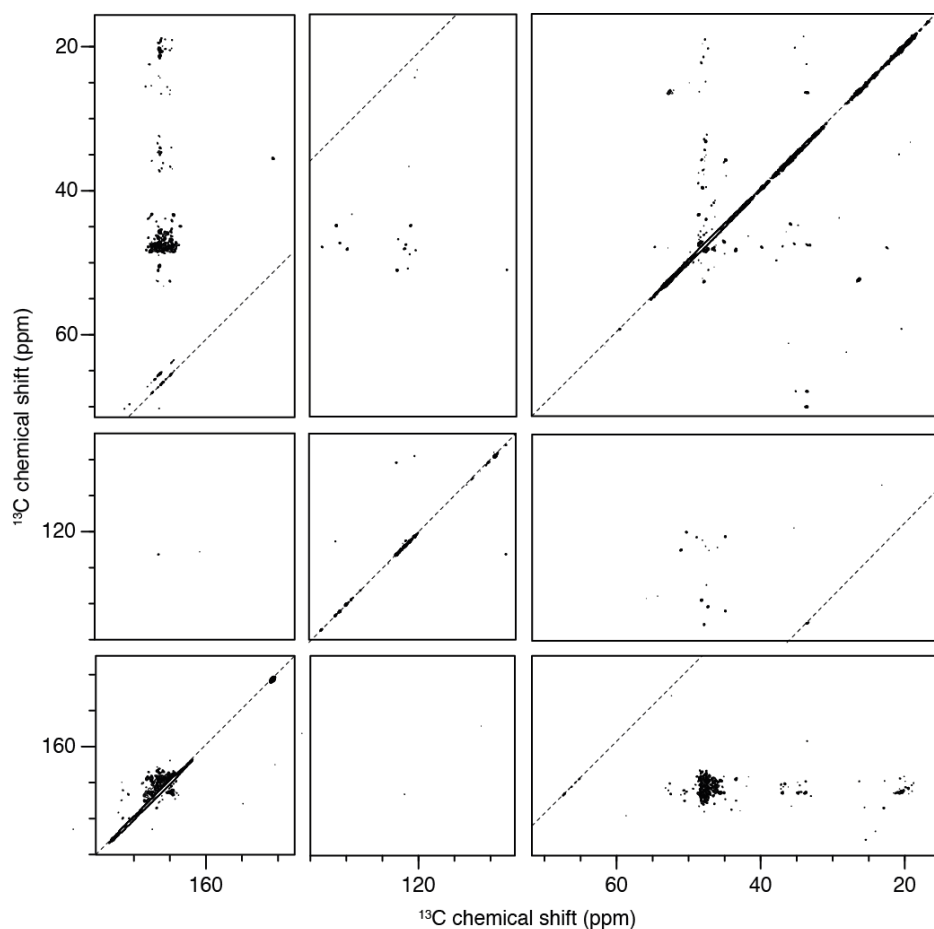
Figure 52: Aromatic Resonances in a 2-^{13}C$_1$ Glycerol Labeled Sample
Cross peaks between aromatic and aliphatic resonances could only be observed with selective labeling. In 2-^{13}C$_1$ glycerol labeled samples, mostly ^{13}Cα and ^{13}Cγ are labeled, which allowed for more efficient transfer into the aromatic ring system. The small number of aliphatic cross peaks results from the mostly unlabeled ^{13}Cβ atoms.

6.3 Assignment of Structured Regions within MAK33 V_L S20N Fibrils

The strategies described in the last section enabled the assignment of 36 rigid residues in MAK33 V_L S20N fibrils. A detailed list with all experiments used for the assignment is given in the supplementary information. We estimate to see approximately 15 to 20 further spin systems, which could not be assigned. Most of the isolated resonances in the ^{13}C,^{13}C-correlation are assigned (Fig. 53). Strong unassigned resonances there indicate the presence of one further isoleucine and two or three threonines, for which no sequential connectivities were found. In the ^{15}N,^{13}C-correlation (Fig. 54), the good dispersion of the ^{15}N resonances enabled better separation of the resonances.

All experiments used for the assignment relied on dipolar transfers. As the effective dipolar couplings are stronger for well structured residues, there is a correlation between signal intensity and rigidity. We could assign most of the strong resonances, which are most likely to form the structurally well defined hydrophobic core of the fibril. The vast majority of assigned resonances is an almost continuous sequence from residue 60 to 87 (Fig. 55).

For several amino acids with long aliphatic side chains, the complete spin systems were visible, e.g. L73, I75 or V78. The ^{13}Cζ resonance of an arginine could also be observed, probably of R61. Both observations indicate well structured side chains. In contrast, the side chains of aromatic amino acids produced hardly any cross peaks in the uniformly ^{13}C labeled sample. Hence, the aromatic rings might be rotating or trapped in different orientations, thus producing only weak signals.

Most of the assigned chemical shifts matched the expectations for the respective amino acid types. Noteworthy is the ^{15}N shift of N76, which is 132.7 ppm. However, this could be validated by several complementary assignment experiments, including the CANCO. The ^{13}CO shift of D82 was marked as unlikely by the CcpNmr analysis software. Our observed value was 170.6 ppm, which we consider as still acceptable, regarding the BMRB value of 176.4 ± 3.7 ppm. The situation was similar for G84 ^{13}Cα (observed: 41.6 ppm, BMRB: 45.4 ± 2.4 ppm). Y86 ^{13}Cβ had the strongest deviation from the random coil carbon chemical shifts (observed: 32.3 ppm, BMRB: 39.4 ± 3.4 ppm). As this resonance could only be observed in the NCACX experiment, we cannot exclude an error here. However, the ring currents of the tyrosine itself or another residue interacting via π-π stacking interactions could also justify larger deviations from the random coil chemical shifts. The other resonances were considered as within the tolerances by the CcpNmr analysis quality check.

Figure 53: ^{13}C,^{13}C-Correlation of MAK33 V$_L$ S20N Fibrils
A PDSD with 50 ms mixing time was used to obtain intraresidue correlations.

Figure 54: ^{15}N,^{13}C-Correlation of MAK33 V$_L$ S20N Fibrils
An N(CA)CX with 50 ms mixing time was used to obtain intraresidue correlations.
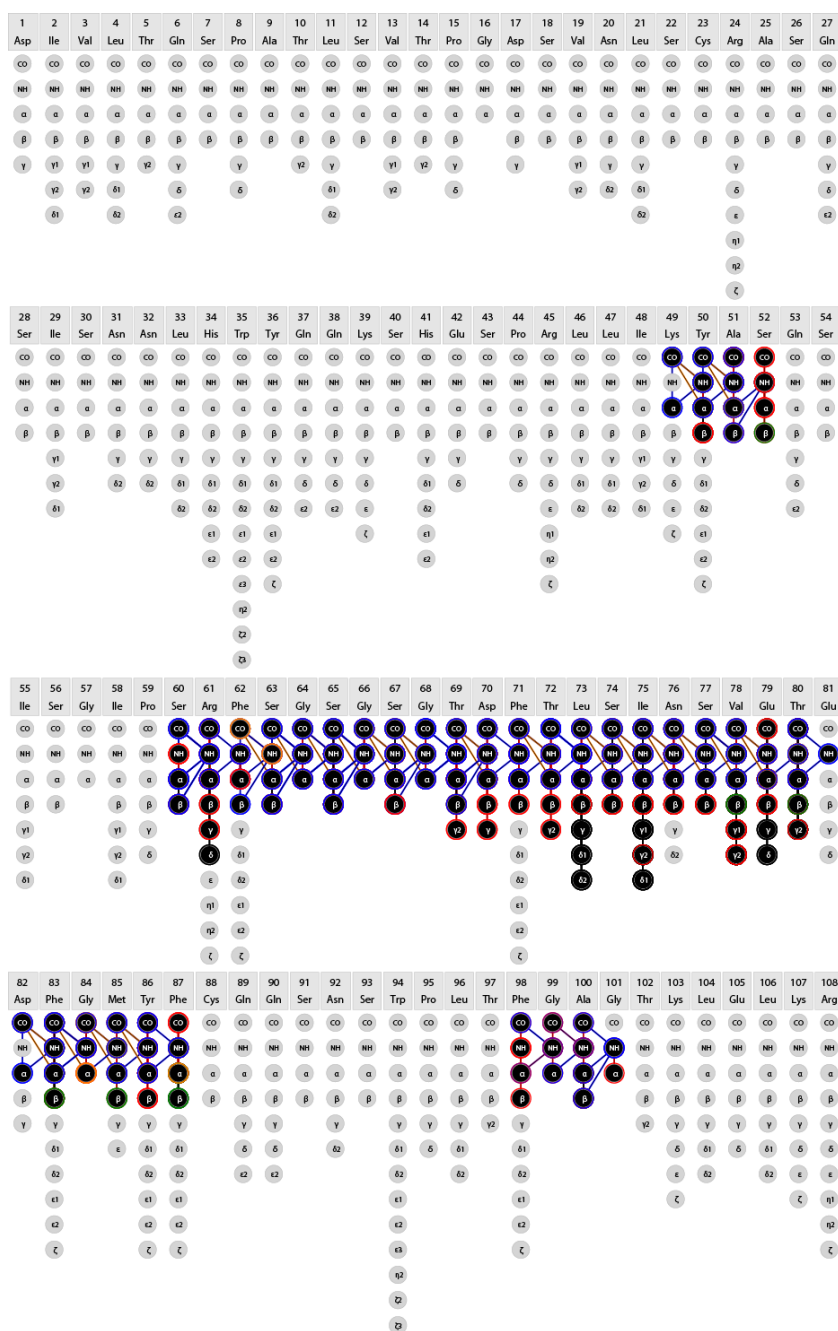
Figure 55: Assigned Atoms by Use of MAS Solid-State NMR Spectroscopy
The colours refer to different types of experiments. Black indicates ^{13}C,^{13}C-correlations. Red indicates NCACX type experiments. Blue indicates NCOCX type experiments and the NCOCA. Green indicates NCACB type experiments. Purple indicates the N(CO)CACB. Orange indicates the CANCO.

According to the secondary structure prediction by TALOS+, most of the residues were either in β-strand conformation or random coil (Fig. 56). Based on these predictions, we suggest the following pattern of β-strands:

```
 50    60   65   70   75   80   85      100
  YAS...SRFSGSGSGTDFTLSINSVETEDFGMYF...FGAG
```

with underlined residues indicating β-strands. The predicted strand at residues F83 and G84 seems problematic. The confidence values by TALOS+ were rather low in this region. Also, the $^{13}C\alpha$ shift of G84 was unusual, as mentioned before. Finally, glycine is unlikely to be found in β-strands.
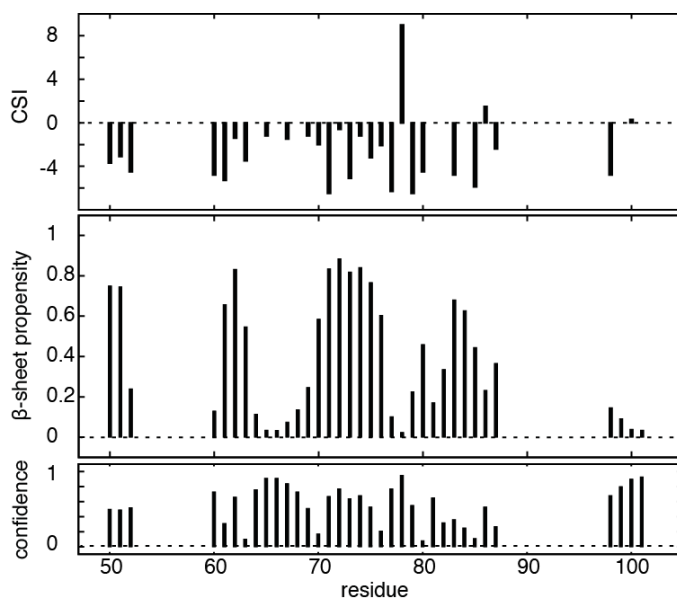
Figure 56: Secondary Structure Prediction for Rigid Residues
The chemical shifts index (CSI) is the difference of the distance between $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts for observed chemical shifts and random coil chemical shifts. Negative CSI values are indications for β-strands.
The β-strand propensities were calculated with TALOS+ [124] based on chemical shifts. For regions with low β-sheet probability, TALOS+ predicted random coil structure, but in no case α-helices.

With respect to the other predicted β-strands, the situation was clearer. The confidence of the prediction was rather high, especially for the loops between the strands. Hence, it is probable, that we not only found the location of the strands, but also got a precise estimate for their boundaries.

6.4 Analysis of Rigid Regions within MAK33 V$_L$ S20N Oligomers

An intriguing feature of the MAK33 V$_L$ domain is its ability to form stable oligomers, which are not amyloid fibrils [33]. This alternatively folded state (AFS) forms at pH 2. Under quiescent conditions, the oligomers are stable at this pH. Only with agitation and elevated temperatures, fibril formation can be observed. While there is no prove yet, the oligomers are likely to be precursors of the mature amyloid fibrils. As discussed in the introduction, such intermediate states are of special interest, as they might be the most toxic species. Using FROSTY NMR [221], such soluble oligomers can be studied with conventional MAS solid-state NMR spectroscopy methods.

The sample of MAK33 V$_L$ WT oligomers was provided by the group of Johannes Buchner and the ^{13}C,^{13}C-correlation was measured by Elke Prade. The experimental conditions for these oligomers were slightly different from those for the MAK33 V$_L$ S20N fibrils. First of all, the fibrils were formed from a mutant and the oligomers from the WT. However, since all assigned residues where located in the C-terminal half of the V$_L$ sequence, the S20N mutation is unlikely to cause major changes. In addition, the NMR experiment with the oligomers was acquired at 277 K instead of 273 K. This could cause little changes in chemical shifts, but not more. With this in mind, the similarity between the spectra of the oligomers and the fibrils is astonishing (Fig. 57).

Figure 57: Comparison Between Fibrils and Oligomers
The fibrils of the S20N mutant are compared with the oligomers of the WT. The additional proline spin system in the oligomers is highlighted with dashed green lines.

Many resonances were identical in both experiments. The resolution was worse, but this was expected. Oligomers are likely to be less structured than fibrils, resulting in broader lines and less efficient dipolar transfers. Still, there was one spin system, which produced more intense signals in the oligomers in comparison to the fibrils. This spin system corresponded to a proline. Some of its resonances could be observed in the fibrils, but only with very weak intensity. Only in the oligomers, the whole spin system was visible. Using the chemical shifts from the oligomers, we could not find this proline in an N(CA)CX experiment with the fibrils. We need to be careful with the discussion of spin systems, which cannot be assigned sequentially. However, given these observations, the structure of this proline might be a distinguishing feature between the otherwise similar oligomers and fibrils.

With Cβ and Cγ chemical shifts of 30.3 ppm respectively 25.5 ppm, the proline was in trans conformation [213]. The chemical shifts were identical in the fibrils, only much weaker. Thus, a change of the proline ring pucker between oligomers and fibrils can be excluded, while such a change during the transition from the native state to oligomers is conceivable.

The quality of the oligomer spectrum was not sufficient for further analysis. However, it is obvious that the oligomers are already very similar to the fibrils. This supports the concept of the soluble oligomers as precursors of the amyloid deposits. Differences arise probably only at the level of quaternary structure, resulting in the distinct superstructures.

6.5 Analysis of Dynamic Regions within MAK33 V_L S20N Fibrils and Oligomers

We have identified several rigid residues within the oligomers and the fibrils. As most parts of the sequence could not be observed by dipolar-based MAS solid-state NMR spectroscopic experiments, it is obvious that many residues must be flexible. If the local dynamics are fast enough to compensate for the slow molecular tumbling of oligomers and fibrils, such residues can be observed by conventional solution-state NMR spectroscopy.

Thus, we measured ^{1}H,^{15}N HSQC experiments with both the oligomers and the amyloid fibrils (Fig. 58). There were some similarities, but the spectra differed much more than the respective dipolar-based ^{13}C,^{13}C-correlations. It is tempting to speculate, that the dynamic regions make the main difference between oligomers and fibrils.
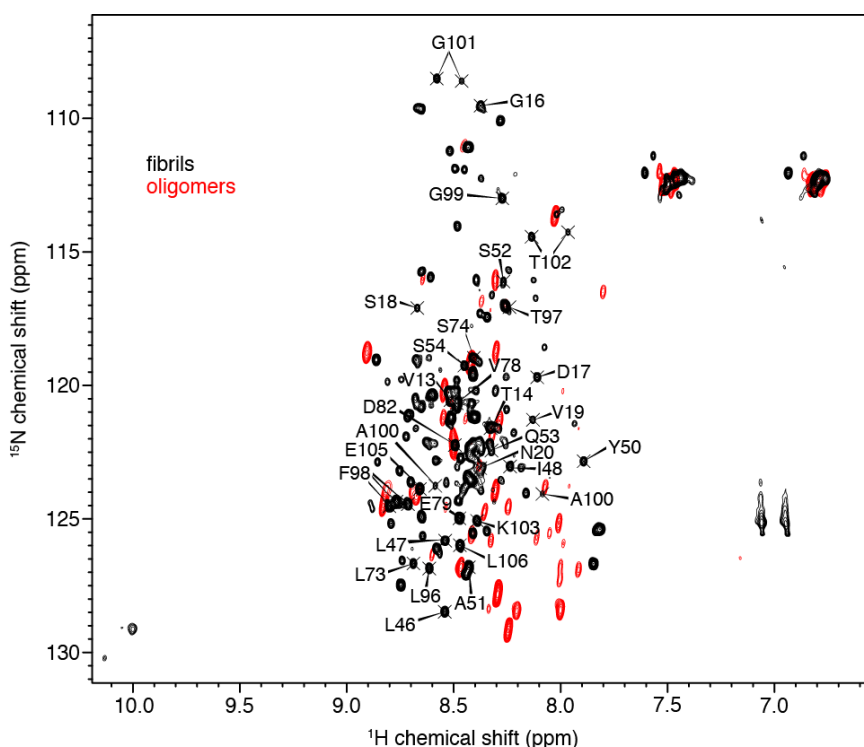
Figure 58: Dynamic Regions in MAK33 V_L S20N Fibrils and Oligomers
Both spectra were acquired as conventional ^{1}H,^{15}N HSQCs by solution-state NMR spectroscopy.

The small dispersion in the ^{15}N dimension was remarkable. Whereas the amide ^{1}H resonances of the native state where spread between 7 and 9.5 ppm, here the dispersion ranged mostly from 8 to 9 ppm. According to literature, amide ^{1}H resonances of unstructured polypeptides are between 8 and 8.4 ppm [222]. Hence, the flexible parts of both the oligomers and the fibrils can be considered mostly unstructured.

95

The large number of resonances in both spectra was surprising. As the quality of the spectra was good, despite the high molecular weights, we acquired experiments for sequential assignment of the fibrils. To enhance sensitivity, the experiments were acquired with 50% NUS. The spectra revealed multiple sets of resonances (Fig. 59a). Such doubling was observed for several different residues, where sequential walks enabled unambiguous assignment. In order to exclude degradation of the protein as a cause for the multiple resonance sets, we analysed the fibrils with SDS PAGE (Fig. 59b). There were no bands at lower MW, compared to the intact protein. The difference between oligomer and fibril HSQC experiments is another argument against degradation artifacts. Thus, the resonances reflected conformational heterogeneity of the flexible regions.
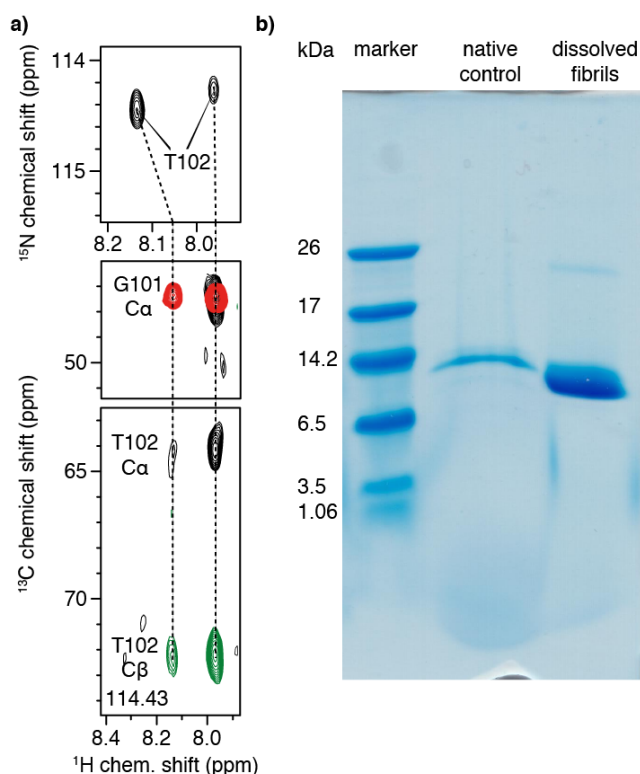
Figure 59: Multiple Sets of Resonances in the Flexible Regions of MAK33 S20N V$_L$ Fibrils
a) In the ^{1}H,^{15}N HSQC, two resonances could be observed for T102. The corresponding ^{13}C resonances were almost identical. The spectra in the lower panel are HNCACB (black and green for positive and negative contours) and CBCACONH (red).
b) The multiple resonances did not result from degradation. Dissolved fibrils displayed the same MW as the native control (11.8 kDa).

As a consequence, sequential assignment was difficult, due to the degeneracy of ^{13}C chemical shifts. In order to circumvent this problem, we acquired an HN(CA)NNH [180]. This experiment allows for sequential walks without the use of ^{13}C resonances (Fig. 60). It works, however, best for highly dynamic intrinsically unfolded proteins [223]. The fact,

that we could assign several resonances with this experiment, emphasised the fast dynamics of the flexible regions. We used the HN(CA)NNH to establish sequential connectivities and the conventional 3D assignment experiments for obtaining $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts, which allowed us to identify the amino acid types. The combination of these two approaches enabled the assignments shown in Fig. 58.
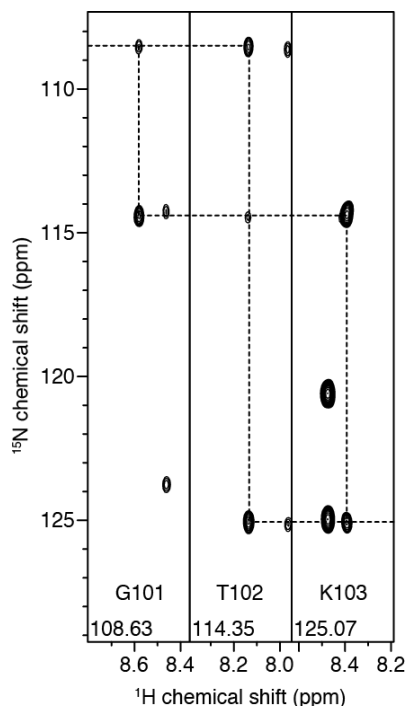
Figure 60: Sequential Walk with HN(CA)NNH
In the ideal case, the amide ^{15}N chemical shifts of residues i-1, i and i+1 can be seen in the strip corresponding to residue i. This requires fast dynamics of the respective residues.

Noteworthy was also the tryptophan side chain resonance at 10 ppm in the ^{1}H dimension of the fibril ^{1}H,^{15}N HSQC (Fig. 58). Two or three resonances could be seen in the fibril spectrum, while none were visible in the oligomer spectrum. The sequence contains two tryptophans, W35 and W94. Without assignment, it cannot be said, whether the visible resonances originated from the same tryptophan or different ones. Due to the bulky side chain, tryptophan is easily trapped in different conformations, resulting in several resonances. Considering, that no residues were sequentially assigned between residue 21 and 45, it is unlikely that the W35 side chain can be observed. In contrast, W94 is right adjacent to the assigned P95 (only $^{13}C\alpha$ and $^{13}C\beta$) and the sequentially assigned L96. Hence, we assume to observe W94. According to the ^{1}H,^{15}N HSQC, it occurs in two or three different conformations.

In summary, we could assign 38 residues sequentially and observed up to three different sets of resonances per residue. While more resonances could be observed, assignment

was not feasible due to resonance doubling and missing sequential connectivities. The results are visualised in the following, with <u>assigned residues underlined</u> and residues with multiple resonances in red:

```
1         11         21         31         41         51
DIVLTQSPAT LSVTPGDSVN LSCRASQSIS NNLHWYQQKS HESPRLLIKY ASQSISGIPS
61        71         81         91         101
RFSGSGSGTD FTLSINSVET EDFGMYFCQQ SNSWPLTFGA GTKLELKR
```

We also analysed secondary structure propensities. As expected, the vast majority of residues observed with solution-state NMR spectroscopy is in random coil conformation (Fig. 61). TALOS+ predicted β-strands only for residues 45 to 47 and for L106. Possibly, there is a structured β-strand prior to residue 45. This strand could comprise those residues, which were visible by MAS solid-state NMR experiments, but could not be assigned sequentially.
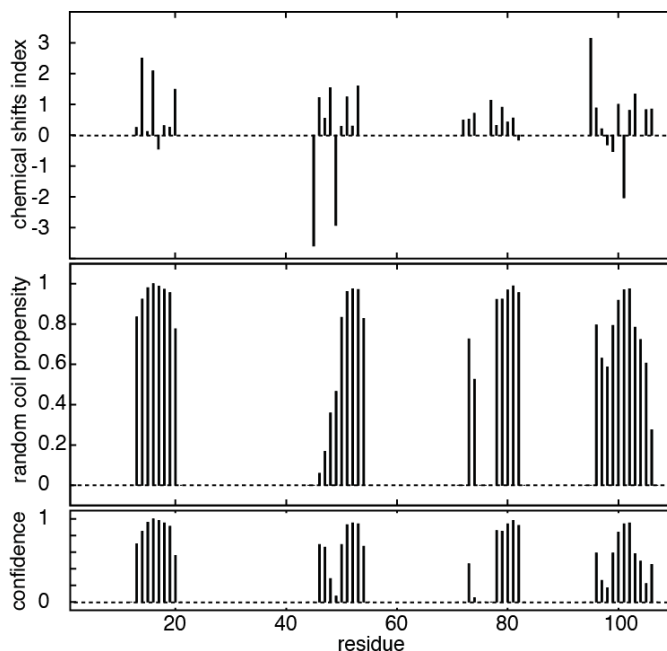
Figure 61: Dynamic Regions are Mostly Unstructured
The chemical shifts index is the difference of the distance between ^{13}Cα and ^{13}Cβ chemical shifts for observed chemical shifts and random coil chemical shifts. If no series of either positive or negative CSI values is observed for several subsequent residues, this indicates random coil structure.
The random coil propensity was calculated with TALOS+.

6.6 Conclusions on Structural Features of MAK33 V$_L$ S20N Fibrils and Oligomers

We presented the first characterisation of antibody LC V$_L$ domain fibrils by NMR spectroscopy. In addition to an analysis of the rigid parts of the fibrils, we also studied a non-fibrillar oligomeric state and investigated both states by solution-state NMR spectroscopy.

6.6.1 Discussion of Specific Residues in the Oligomers and Fibrils

Proline in the Rigid Parts of the Oligomers

Some residues and resonances deserve further discussion. In the solid-state NMR ^{13}C,^{13}C-correlation of the oligomers, we observed a proline spin system (Fig. 57). While some of its resonances were also visible in the fibril sample, the intensity was much less there. The different intensities indicate a change in structural homogeneity or rigidity. The rigid proline is probably close to other assigned residues. Of the five prolines in the MAK33 V$_L$ S20N sequence, P59 is directly adjacent to the sequentially assigned residues. P59 is especially likely, since also one isoleucine could not be assigned so far and residue 58 is an isoleucine:

```
1          11         21         31         41         51
DIVLTQSPAT LSVTPGDSVN LSCRASQSIS NNLHWYQQKS HESPRLLIKY ASQSISGIPS
61         71         81         91         101
RFSGSGSGTD FTLSINSVET EDFGMYFCQQ SNSWPLTFGA GTKLELKR
```

with unassigned isoleucines and prolines highlighted in red and sequentially assigned residues underlined. If oligomers and fibrils differ in the loop including P58, this could mean a displacement of the short β-strand containing Y50 and A51. However, as these residues are already weak in the fibrils and cannot be observed clearly in the oligomers, this is just speculative.

One of the sparse information in literature on the structure of LC V$_L$ fibrils is a special conformation of P8 [205]. This highly conserved proline is in the cis conformation in the native state. mAb 11-1F4 is an antibody, which binds to LC V$_L$ fibrils, but not to the native state. Epitope mapping revealed a β-turn involving P8 as binding site of this antibody. The authors considered both a cis or a trans conformation of P8 as possible in the fibril state. We observed a trans proline, but consider it unlikely to be P8, since we assigned no other rigid residues in the N-terminal half of the V$_L$ domain. Still, the epitope of mAb 11-1F4 suggests further structured regions at the N-terminus.

Tryptophan in the Flexible Parts of the Fibrils

In the solution-state NMR spectra of the fibrils, we observed tryptophan side chain signals, which were absent in the oligomer state. As explained above, we assume this to be W94, possibly in different polymorphs. In a previous characterisation of the

oligomer state of MAK33 V$_L$, Feige et al. observed a slight blueshift of the tryptophan fluorescence, compared to the native state signal [33]. This indicates a hydrophobic environment. As W35 is already buried in the hydrophobic core in the native structure, the shift is probably caused by W94, which is completely solvent exposed in the native state. Upon oligomer formation, W94 becomes buried, explaining the blueshift. Concordantly, we did not observe the tryptophan signals by solution-state NMR spectroscopy in the oligomer state. In the fibril state, tryptophan, propably W94, becomes visible by solution-state NMR. Hence, we suggest a conversion from exposed W94 in the native state via W94 with restricted motions in the oligomers to flexible W94 in the fibrils.

Relevance of the pH and Compensation of Charges

Both the oligomers and the fibrils were characterised at pH 2. We tried to shift the pH back to neutral after formation of the oligomer respectively fibril states. However, this resulted in loss of spectral quality and probably dissolution of the fibrils (data not shown). Hence, the acidic pH stabilises the examined states. Such a behaviour was already reported for acid-induced β2-microglobulin fibrils [224]. We checked the assigned carboxyl atoms of acidic amino acids with respect to their protonation state. We got no indications for charged acidic amino acids from the chemical shifts (data not shown). In contrast, we observed the $^{13}C\zeta$ of an arginine, which is for sure charged at pH 2. With selectively glycerol labeled samples, we confirmed the characteristic labeling pattern of arginine and observed cross peaks to the aliphatic resonances, which were not visible with uniformly labeled samples. The resonance belongs to R61. The clear $^{13}C\zeta$ signal in the $^{13}C,^{13}C$-correlation indicates a structured side chain. As there are no negatively charged residues at pH 2, the compensation of the arginine charge in the fibril core is a problem. We suggest either binding to a Cl$^-$ ion from the buffer or cation-π interactions for stabilisation of the arginine within the core of the fibril. Seven aromatic amino acids were identified by MAS solid-state NMR spectroscopy: Y49, F62, F71, F83, Y86, F87 and F98. The quadrupole moments of these aromatic systems could shield the arginine charge. However, the aromatic resonances themselves were only poorly visible. Thus, we could not validate an interaction partner experimentally.

6.6.2 Heterogeneity and Structure of the Oligomers and Fibrils

MAK33 V$_L$ is special, as it can form stable oligomers [33]. This enabled the characterisation of this usually only transiently populated species by solid- and solution-state NMR spectroscopy. The MAK33 V$_L$ oligomers turned out to be very similar to the fibrils with respect to the rigid parts. Such a similarity in rigid parts has already been reported for Amyloid-β oligomers and protofibrils and might be common for amyloidogenic proteins [225]. In contrast, there are considerable differences in the flexible parts. Most likely, the transformation from oligomers to fibrils proceeds via rearrangements of the quaternary structure, which enable the continuous addition of monomers at the edges. Changes of the tertiary structure, however, seem to play a minor role in this process. Overall, our data support the hypothesis of the oligomers being on-pathway precursors of the fibrils.

Regarding the fibril structure, we assigned 36 residues by MAS solid-state NMR spectroscopy and 38 residues by solution-state NMR spectroscopy. In both cases, several more unassigned spin systems were visible. In these cases, weak signal intensities hampered further assignments. Still, not all of the 108 residues yielded signals. This is likely due to heterogeneity, which causes many or broad signals with low intensities. Both static heterogeneity, i.e. polymorphism, or slow sampling of different conformations would result in invisible residues.

Of special interest is the fact, that we observed some residues both by solid- and solution-state NMR spectroscopy. The following overview summarises the assignments:

```
1         11         21         31         41         51
DIVLTQSPAT LSVTPGDSVN LSCRASQSIS NNLHWYQQKS HESPRLLIKY ASQSISGIPS
61         71         81         91         101
 RFSGSGSGTD FTLSINSVET EDFGMYFCQQ SNSWPLTFGA GTKLELKR
```

with underlined residues being visible by solid-state NMR spectroscopy and red residues being visible by solution-state NMR spectroscopy. Obviously, one amino acid cannot be rigid and flexible at the same time. In order to be visible by MAS solid-state NMR spectroscopy, the effective tumbling correlation time τ_c of a residue needs to be longer than the MAS rotor period (100 μs at 10 kHz MAS frequency). Otherwise, efficient averaging of anisotropic interactions cannot be achieved [221]. Solution-state NMR spectroscopy, in contrast, is suited for proteins with τ_c on the order of nanoseconds [226]. Hence, we observed different states of the respective residues, if they are visible with both methods. As explained in the results section, degradation could be excluded by SDS-PAGE, which showed the correct MW. In addition, we can exclude incomplete fibril formation and leftover monomers in solution. When we pelleted the fibrils for 10 minutes at 21,000 g, no monomers where left in the supernatant, according to the UV absorption at 280 nm (data not shown). Thus, the different states need to be of high MW.

Multiple resonances could mean static heterogeneity or dynamic exchange between several states. An example for static heterogeneity is a heterodimer as subunit of the fibril core [227]. Another possibility are different polymorphs of fibrils (Fig. 62). In both

cases, a residue can be rigid in one polymorph and flexible in the other. This model could also explain the multiple resonances observed by solution-state NMR spectroscopy. They would correspond to the different polymorphs, which could also interact with either the same or another polymorph, resulting in distinct chemical environments. In addition, this hypothesis of distinct polymorphs can explain the rather weak signals. The rigid residues could not be assigned completely, despite large efforts, due to poor signal intensity. With two polymorphs, the amount of substance for each polymorph would be smaller than expected, producing less intense signals.
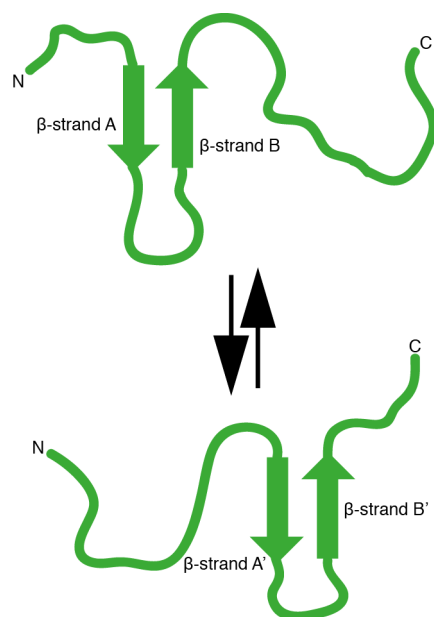
Figure 62: Coexistence of Two Polymorphs
According to this model, the V$_L$ monomer exists in two distinct polymorphs in the fibril state. The rigid residues are different in both forms. The polymorphs might form pure fibrils, consisting of only one polymorph, or mixed fibrils. Exchange between these polymorphs might be very slow.

Another hypothesis to explain the data is partial dissociation at the fibril ends. There is already experimental evidence for fibril disassembly in literature [228, 229]. This requires a dynamic equilibrium between fibril and monomer state. However, if complete monomers would dissociate, we should observe all residues by solution-state NMR spectroscopy. Also, localised polymorphism would be difficult to explain for free unstructured monomers. Finally, dissociated monomers would not precipitate with the fibrils.

Instead, we suggest the dissociation of individual strands in contrast to full monomers (Fig. 63a, b). Our MAS solid-state NMR spectroscopic data indicate a structured core of the fibril, located in the C-terminal half. Besides, there is an unstructured region. Such unstructured regions are typically assumed to be flexible. However, given the high local protein concentrations, they might as well form unstructured aggregates. Due to their heterogeneity, they would be invisible by experiments based on dipolar
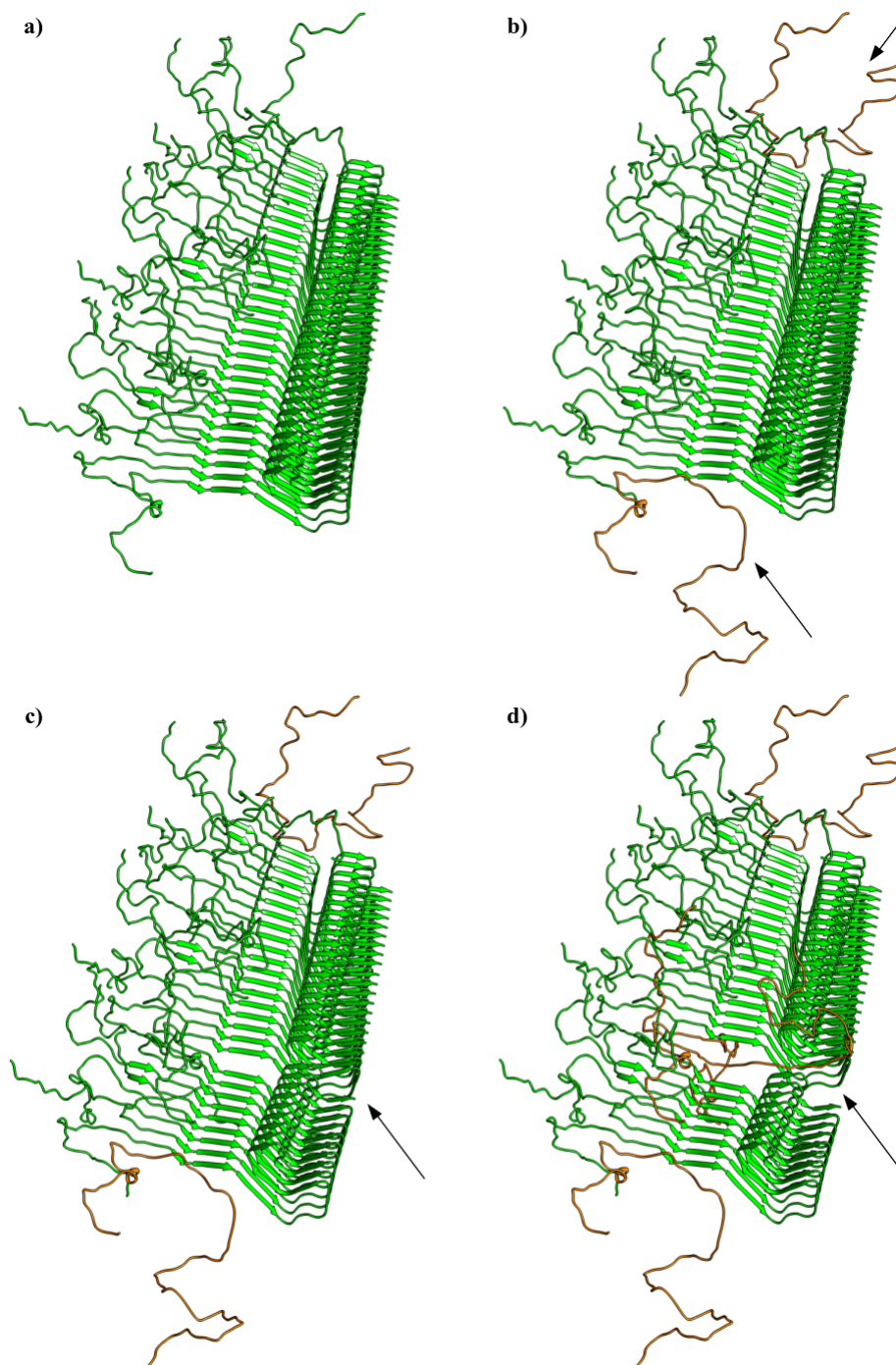
Figure 63: Partial Dissociation of β-strands at the Fibril Ends
Starting from a fully structured fibril core **(a)**, the edge strands (highlighted in orange) could dissociate and form unstructured random coil elements **(b)**. The always unstructured aggregated regions outside of the fibril core keep the monomers together. Ruptures of the structured core could increase the number of flexible tails **(c, d)**. A modified fibril structure of α-synuclein (PDB ID: 2N0A, [153]) was used for visualisation of the concept.

couplings. As unstructured aggregation is often irreversible, this would prevent the release of monomers from the fibril ends. The structured β-strands at the fibril ends, however, could dissociate. Then, these tails might form regions of dynamic random coil structure, which are still attached to the fibril via the unstructured aggregated part. Hence, this model can explain, why several regions can be observed at the same time by experiments based on dipolar coupling and those based on J-couplings.

This model would indicate rather low stability of the fibrils, if the core dissociates so easily. This is in agreement with the loss of structure upon change from acidic to neutral pH. Moreover, low stability might cause small-scale fluctuations, which result in less efficient transfers via dipolar couplings. This would explain, why the rigid parts of the fibrils were so difficult to assign.

A problem of this hypothesis would be the unexpected high signal intensity in the solution-state NMR spectra. Integration of the resonances and comparison to samples of known concentrations indicated a fraction of 10 - 20% of the monomers to produce signals in solution-state NMR spectroscopy (data not shown). According to our model of loose monomers at the fibril ends, the fibrils then could consist only of 5 to 10 monomers, on average. However, we observed fibrils with a length of several 100 μm by TEM. Therefore, this model can only be valid, if also monomers in the middle of a fibril could become loose. Considering again the low stability of the fibrils, shear forces might rupture the rigid core (Fig. 63c). Unstructured aggregated N-terminal regions would keep the fibril together, while the rigid part would get additional ends, where the strands could dissociate (Fig. 63d).

The actual structures of the oligomers and fibrils remain unknown so far, as we did not measure distance restraints. Therefore, we cannot say yet, which hypothesis is correct, polymorphism or partial dissociation. However, our data allowed first structural insights into V_L amyloid fibril structure. We identified structured regions and β-strands. Goals for future research projects would be to elucidate the topology and finally atomic structure of the oligomers and fibrils. This would pave the way for an understanding of the fibril formation process. Another open question is, how the fibrils of different V_L sequences compare. Despite the large sequence variety, we expect a conserved structure for the V_L fibrils, as the rigid residues are found mostly in the more conserved framework of the domain. According to our assignment, CDR3 is precisely excluded from the rigid parts of the fibril structure:

```
1         11        21        31        41        51
DIVLTQSPAT LSVTPGDSVN LSCRASQSIS NNLHWYQQKS HESPRLLIKY ASQSISGIPS
61        71        81        91        101
RFSGSGSGTD FTLSINSVET EDFGMYFCQQ SNSWPLTFGA GTKLELKR
```

with red indicating the assigned rigid parts and underlined residues indicating the CDRs. Only a short stretch in CDR2 is visible by solid-state NMR spectroscopy. This supports the hypothesis, that the framework is more important for amyloid formation than the CDRs [230, 231]. The insertion of two residues at positions 68 and 69 in the highly amyloidogenic λ-6 germline is also located right in the structured fibril core, according to

our assignments [230, 231]. While such findings were so far only discussed with respect to the native structure, we hope that insights into the fibril structure will provide new explanations for the effects of amyloidogenic sequence variations. The question, which V_L sequences are prone to amyloid aggregation, cannot be answered completely without knowledge about the fibril structure.

7 References

[1] Christian B Anfinsen. Some observations on the basic principles of design in protein molecules. *Comparative biochemistry and physiology*, 4(2-4):229–240, 1962.

[2] Marcus Fändrich, Matthew A Fletcher, and Christopher M Dobson. Amyloid fibrils from muscle myoglobin. *Nature*, 410(6825):165–166, 2001.

[3] Marcus Fändrich and Christopher M Dobson. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *The EMBO journal*, 21(21):5682–5690, 2002.

[4] Paolo Arosio, Tuomas PJ Knowles, and Sara Linse. On the lag phase in amyloid fibril formation. *Physical Chemistry Chemical Physics*, 17(12):7606–7618, 2015.

[5] Margaret Sunde, Louise C Serpell, Mark Bartlam, Paul E Fraser, Mark B Pepys, and Colin CF Blake. Common core structure of amyloid fibrils by synchrotron x-ray diffraction. *Journal of molecular biology*, 273(3):729–739, 1997.

[6] Anne K Schütz, Toni Vagt, Matthias Huber, Oxana Y Ovchinnikova, Riccardo Cadalbert, Joseph Wall, Peter Güntert, Anja Böckmann, Rudi Glockshuber, and Beat H Meier. Atomic-Resolution Three-Dimensional Structure of Amyloid β Fibrils Bearing the Osaka Mutation. *Angewandte Chemie International Edition*, 54(1):331–335, 2015.

[7] Giampaolo Merlini and Vittorio Bellotti. Molecular mechanisms of amyloidosis. *New England Journal of Medicine*, 349(6):583–596, 2003.

[8] Fabrizio Chiti and Christopher M Dobson. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.*, 75:333–366, 2006.

[9] Camilla Nilsberth, Anita Westlind-Danielsson, Christopher B Eckman, Margaret M Condron, Karin Axelman, Charlotte Forsell, Charlotte Stenh, Johan Luthman, David B Teplow, Steven G Younkin, et al. The'Arctic'APP mutation (E693G) causes Alzheimer's disease by enhanced Aβ protofibril formation. *Nature neuroscience*, 4(9):887–893, 2001.

[10] Mónica Mendes Sousa, Isabel Cardoso, Rui Fernandes, António Guimaraes, and Maria Joao Saraiva. Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy: Evidence for toxicity of nonfibrillar aggregates. *The American journal of pathology*, 159(6):1993–2000, 2001.

[11] Samuel A Kotler, Patrick Walsh, Jeffrey R Brender, and Ayyalusamy Ramamoorthy. Differences between amyloid-β aggregation in solution and on the membrane: insights into elucidation of the mechanistic details of Alzheimer's disease. *Chemical Society Reviews*, 43(19):6692–6700, 2014.

[12] Estelle Desport, Frank Bridoux, Christophe Sirac, Sébastien Delbes, Sébastien Bender, Béatrice Fernandez, Nathalie Quellard, Corinne Lacombe, Jean-Michel Goujon, David Lavergne, Julie Abraham, Guy Touchard, Jean-Paul Fermand, and Arnaud Jaccard. AL Amyloidosis. *Orphanet journal of rare diseases*, 7(1):1–13, 2012.

[13] Matthias J Feige, Sandra Groscurth, Moritz Marcinowski, Yuichiro Shimizu, Horst Kessler, Linda M Hendershot, and Johannes Buchner. An unfolded C_H1 domain controls the assembly and secretion of IgG antibodies. *Molecular cell*, 34(5):569–579, 2009.

[14] William D Terry, David L Page, Shigeru Kimura, Takashi Isobe, Elliott F Osserman, and George G Glenner. Structural identity of Bence Jones and amyloid fibril proteins in a patient with plasma cell dyscrasia and amyloidosis. *Journal of Clinical Investigation*, 52(5):1276, 1973.

[15] Giovanni Palladini, Francesca Lavatelli, Paola Russo, Stefano Perlini, Vittorio Perfetti, Tiziana Bosoni, Laura Obici, Arthur R Bradwell, GianVico Melzi D'Eril, Roberto Fogari, et al. Circulating amyloidogenic free light chains and serum N-terminal natriuretic peptide type B decrease simultaneously in association with improvement of survival in AL. *Blood*, 107(10):3854–3858, 2006.

[16] Kenneth P Murphy, Paul Travers, Mark Walport, and Charles Janeway. *Immunologie*. Spektrum Akad. Verlag, 2009.

[17] Morie A Gertz, Ray Comenzo, Rodney H Falk, Jean Paul Fermand, Bouke P Hazenberg, Philip N Hawkins, Giampaolo Merlini, Philippe Moreau, Pierre Ronco, Vaishali Sanchorawala, et al. Definition of organ involvement and treatment response in immunoglobulin light chain amyloidosis (AL): a consensus opinion from the 10th International Symposium on Amyloid and Amyloidosis. *American journal of hematology*, 79(4):319–328, 2005.

[18] Giampaolo Merlini, David C Seldin, and Morie A Gertz. Amyloidosis: pathogenesis and new therapeutic options. *Journal of Clinical Oncology*, 29(14):1924–1933, 2011.

[19] Shaji K Kumar, Morie A Gertz, Martha Q Lacy, David Dingli, Suzanne R Hayman, Francis K Buadi, Kristen Short-Detweiler, Steven R Zeldenrust, Nelson Leung, Philip R Greipp, et al. Recent improvements in survival in primary systemic amyloidosis and the importance of an early mortality risk score. In *Mayo Clinic Proceedings*, volume 86, pages 12–18. Elsevier, 2011.

[20] Daniel A Brenner, Mohit Jain, David R Pimentel, Bo Wang, Lawreen H Connors, Martha Skinner, Carl S Apstein, and Ronglih Liao. Human amyloidogenic light chains directly impair cardiomyocyte function through an increase in cellular oxidant stress. *Circulation research*, 94(8):1008–1010, 2004.

[21] Laura A Sikkink and Marina Ramirez-Alvarado. Cytotoxicity of amyloidogenic immunoglobulin light chains in cell culture. *Cell death & disease*, 1(11):e98, 2010.

[22] Luisa Diomede, Paola Rognoni, Francesca Lavatelli, Margherita Romeo, Elena del Favero, Laura Cantù, Elena Ghibaudi, Andrea di Fonzo, Alessandro Corbelli, Fabio Fiordaliso, et al. A *Caenorhabditis elegans*–based assay recognizes immunoglobulin light chains causing heart amyloidosis. *Blood*, 123(23):3543–3552, 2014.

[23] Jiamin Teng, William J Russell, Xin Gu, James Cardelli, M Lamar Jones, and Guillermo A Herrera. Different types of glomerulopathic light chains interact with mesangial cells using a common receptor but exhibit different intracellular trafficking patterns. *Laboratory investigation*, 84(4):440–451, 2004.

[24] Jiamin Teng, Elba A Turbat-Herrera, and Guillermo A Herrera. An animal model of glomerular light-chain-associated amyloidogenesis depicts the crucial role of lysosomes. *Kidney international*, 86(4):738–746, 2014.

[25] Jian Guan, Shikha Mishra, Yiling Qiu, Jianru Shi, Kyle Trudeau, Guy Las, Marc Liesa, Orian S Shirihai, Lawreen H Connors, David C Seldin, et al. Lysosomal dysfunction and impaired autophagy underlie the pathogenesis of amyloidogenic light chain-mediated cardiotoxicity. *EMBO molecular medicine*, 6(11):1493–1507, 2014.

[26] Helen P McWilliams-Koeppen, James S Foster, Nicole Hackenbrack, Marina Ramirez-Alvarado, Dallas Donohoe, Angela Williams, Sallie Macy, Craig Wooliver, Dale Wortham, Jennifer Morrell-Falvey, et al. Light Chain Amyloid Fibrils Cause Metabolic Dysfunction in Human Cardiomyocytes. *PloS one*, 10(9):e0137716, 2015.

[27] John G Augustine, Agustin de la Calle, Gerhard Knarr, Johannes Buchner, and Christin A Frederick. The Crystal Structure of the Fab Fragment of the Monoclonal Antibody MAK33. *Journal of Biological Chemistry*, 276(5):3287–3294, 2001.

[28] Marcus Jäger and Andreas Plückthun. The rate-limiting steps for the folding of an antibody scFv fragment. *FEBS letters*, 418(1):106–110, 1997.

[29] Emma R Simpson, Eva M Herold, and Johannes Buchner. The folding pathway of the antibody V_L domain. *Journal of molecular biology*, 392(5):1326–1338, 2009.

[30] Robert Huber and Wolfgang Steigemann. Two cis-prolines in the Bence-Jones protein Rei and the cis-pro-bend. *FEBS letters*, 48(2):235–237, 1974.

[31] Batia Kaplan, Avi Livneh, and Ben-Ami Sela. Immunoglobulin Free Light Chain Dimers in Human Diseases. *The Scientific World Journal*, 11:726–735, 2011.

[32] Johannes Buchner, Michael Renner, Hauke Lilie, Hans Juergen Hinz, Rainer Jaenicke, Thomas Kiefhaber, and Rainer Rudolph. Alternatively folded states of an immunoglobulin. *Biochemistry*, 30(28):6922–6929, 1991.

[33] Matthias Johannes Feige, Emma Rhiannon Simpson, Eva Maria Herold, Alexander Bepperling, Klaus Heger, and Johannes Buchner. Dissecting the alternatively folded state of the antibody fab fragment. *Journal of molecular biology*, 399(5):719–730, 2010.

[34] Bertrand Morel, Lorena Varela, and Francisco Conejero-Lara. The thermodynamic stability of amyloid fibrils studied by differential scanning calorimetry. *The Journal of Physical Chemistry B*, 114(11):4010–4019, 2010.

[35] Weronika Surmacz-Chwedoruk, Iwona Malka, Łukasz Bożycki, Hanna Nieznańska, and Wojciech Dzwolak. On the heat stability of amyloid-based biological activity: insights from thermal degradation of insulin fibrils. *PloS one*, 9(1):e86320, 2014.

[36] Xiaoyun Meng, Anthony L Fink, and Vladimir N Uversky. The Effect of Membranes on the *in vitro* Fibrillation of an Amyloidogenic Light-Chain Variable-Domain SMA. *Journal of molecular biology*, 381(4):989–999, 2008.

[37] Hauke Lilie, Johannes Buchner, Kurt Lang, and Rainer Rudolph. Prolyl isomerases catalyze antibody folding in vitro. *Protein Science*, 2(9):1490–1496, 1993.

[38] Anika Gladytz, Evgeny Lugovoy, Ales Charvat, Tilman Häupl, Katrin R Siefermann, and Bernd Abel. Intermediates caught in the act: tracing insulin amyloid fibril formation in time by combined optical spectroscopy, light scattering, mass spectrometry and microscopy. *Physical Chemistry Chemical Physics*, 17(2):918–927, 2015.

[39] Tai Te Wu and Elvin A Kabat. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *The Journal of experimental medicine*, 132(2):211–250, 1970.

[40] Luis del Pozo Yauner, Ernesto Ortiz, Rosalba Sánchez, Rosana Sánchez-López, Leopoldo Güereca, Charles L Murphy, Amy Allen, Jonathan S Wall, D Alejandro Fernández-Velasco, Alan Solomon, et al. Influence of the germline sequence on the thermodynamic stability and fibrillogenicity of human lambda 6 light chains. *Proteins: Structure, Function, and Bioinformatics*, 72(2):684–692, 2008.

[41] Mordechai Pras, Maxwell Schubert, Dorothea Zucker-Franklin, Abraham Rimon, and Edward C Franklin. The Characterization of Soluble Amyloid Prepared in Water. *Journal of Clinical Investigation*, 47(4):924, 1968.

[42] Priscilla Wilkins Stevens, Rosemarie Raffen, Deborah K Hanson, Ya-Li Deng, Maria Berrios-Hammond, Florence A Westholm, Marianne Schiffer, Fred J Stevens, Charles Murphy, Alan Solomon, et al. Recombinant immunoglobulin variable domains generated from synthetic genes provide a system for in vitro characterization of light-chain amyloid proteins. *Protein Science*, 4(3):421–432, 1995.

[43] Allen B Edmundson, Mical K Wood, Marianne Schiffer, Karl D Hardman, Clinton F Ainsworth, Kathryn R Ely, and Harold F Deutsch. A Crystallographic Investigation of a Human IgG Immunoglobulin. *Journal of Biological Chemistry*, 245(10):2763–2764, 1970.

[44] Marianne Schiffer, Rowland L Girling, Kathryn R Ely, and Allen B Edmundson. Structure of a λ-type Bence-Jones protein at 3.5-Å resolution. *Biochemistry*, 12(23):4620–4631, 1973.

[45] Alan Solomon, Deborah T Weiss, Charles L Murphy, Rudi Hrncic, Jonathan S Wall, and Maria Schell. Light chain-associated amyloid deposits comprised of a novel κ constant domain. *Proceedings of the National Academy of Sciences*, 95(16):9547–9551, 1998.

[46] Karen Ege Olsen, Knut Sletten, and Per Westermark. Fragments of the Constant Region of Immunoglobulin Light Chains Are Constituents of AL-Amyloid Proteins. *Biochemical and biophysical research communications*, 251(2):642–647, 1998.

[47] Roshini S Abraham, Susan M Geyer, Tammy L Price-Troska, Cristine Allmer, Robert A Kyle, Morie A Gertz, and Rafael Fonseca. Immunoglobulin light chain variable (V) region genes influence clinical presentation and outcome in light chain–associated amyloidosis (AL). *Blood*, 101(10):3801–3807, 2003.

[48] Jonathan Wall, Maria Schell, Charles Murphy, Rudi Hrncic, Fred J Stevens, and Alan Solomon. Thermodynamic Instability of Human $\lambda 6$ Light Chains: Correlation with Fibrillogenicity. *Biochemistry*, 38(42):14101–14108, 1999.

[49] Raymond L Comenzo, Yana Zhang, Carmen Martinez, Keren Osman, and Guillermo A Herrera. The tropism of organ involvement in primary systemic amyloidosis: contributions of Ig V_L germ line gene use and clonal plasma cell burden. *Blood*, 98(3):714–720, 2001.

[50] Tanya L Poshusta, Nagaaki Katoh, Morie A Gertz, Angela Dispenzieri, and Marina Ramirez-Alvarado. Thermal Stability Threshold for Amyloid Formation in Light Chain Amyloidosis. *International journal of molecular sciences*, 14(11):22604–22617, 2013.

[51] Marta Marin-Argany, Jofre Güell-Bosch, Luis M Blancas-Mejía, Sandra Villegas, and Marina Ramirez-Alvarado. Mutations can cause light chains to be too stable or too unstable to form amyloid fibrils. *Protein Science*, 24(11):1829–1840, 2015.

[52] Ritu Khurana, Joel R Gillespie, Anupam Talapatra, Lauren J Minert, Cristian Ionescu-Zanetti, Ian Millett, and Anthony L Fink. Partially Folded Intermediates as Critical Precursors of Light Chain Amyloid Fibrils and Amorphous Aggregates. *Biochemistry*, 40(12):3525–3535, 2001.

[53] Luis M Blancas-Mejia, Luis A Tellez, Luis del Pozo-Yauner, Baltazar Becerril, Jose M Sanchez-Ruiz, and D Alejandro Fernandez-Velasco. Thermodynamic and Kinetic Characterization of a Germ Line Human $\lambda 6$ Light-Chain Protein: the Relation between Unfolding and Fibrillogenesis. *Journal of molecular biology*, 386(4):1153–1166, 2009.

[54] Sergey P Martsev, Anatoly P Dubnovitsky, Alexander P Vlasov, Masaru Hoshino, Kazuhiro Hasegawa, Hironobu Naiki, and Yuji Goto. Amyloid Fibril Formation of the Mouse V_L Domain at Acidic pH. *Biochemistry*, 41(10):3389–3395, 2002.

[55] Zhijie Qin, Dongmei Hu, Min Zhu, and Anthony L Fink. Structural Characterization of the Partially Folded Intermediates of an Immunoglobulin Light Chain Leading to Amyloid Fibrillation and Amorphous Aggregation. *Biochemistry*, 46(11):3521–3531, 2007.

[56] Elizabeth M Baden, Edward G Randles, Awo K Aboagye, James R Thompson, and Marina Ramirez-Alvarado. Structural Insights into the role of Mutations in Amyloidogenesis. *Journal of Biological Chemistry*, 283(45):30950–30956, 2008.

[57] Martín González-Andrade, Baltazar Becerril-Luján, Rosana Sánchez-López, Héctor Ceceña-Álvarez, Julio I Pérez-Carreón, Ernesto Ortiz, D Alejandro Fernández-Velasco, and Luis Pozo-Yauner. Mutational and genetic determinants of $\lambda 6$ light chain amyloidogenesis. *FEBS Journal*, 280(23):6173–6183, 2013.

[58] Yuta Kobayashi, Hirotaka Tsutsumi, Tetsuyuki Abe, Kyohei Ikeda, Yuki Tashiro, Satoru Unzai, Hironari Kamikubo, Mikio Kataoka, Hidekazu Hiroaki, and Daizo Hamada. Decreased Amyloidogenicity Caused by Mutational Modulation of Surface Properties of the Immunoglobulin Light Chain BRE Variable Domain. *Biochemistry*, 53(31):5162–5173, 2014.

[59] Manikanthan Bhavaraju and Ulrich HE Hansmann. Effect of single point mutations in a form of systemic amyloidosis. *Protein Science*, 24(9):1451–1462, 2015.

[60] Ara Celi DiCostanzo, James R Thompson, Francis C Peterson, Brian F Volkman, and Marina Ramirez-Alvarado. Tyrosine Residues Mediate Fibril Formation in a Dynamic Light Chain Dimer Interface. *Journal of Biological Chemistry*, 287(33):27997–28006, 2012.

[61] Elizabeth M Baden, Barbara AL Owen, Francis C Peterson, Brian F Volkman, Marina Ramirez-Alvarado, and James R Thompson. Altered Dimer Interface Decreases Stability in an Amyloidogenic Protein. *Journal of Biological Chemistry*, 283(23):15853–15860, 2008.

[62] Boris Brumshtein, Shannon R Esswein, Meytal Landau, Christopher M Ryan, Julian P Whitelegge, Martin L Phillips, Duilio Cascio, Michael R Sawaya, and David S Eisenberg. Formation of Amyloid Fibers by Monomeric Light Chain Variable Domains. *Journal of Biological Chemistry*, 289(40):27513–27525, 2014.

[63] Mathieu Laporte Wolwertz, Phuong Trang Nguyen, Noé Quittot, and Steve Bourgault. Probing the role of $\lambda 6$ immunoglobulin light chain dimerization in amyloid formation. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 2016.

[64] Jane S Richardson and David C Richardson. Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proceedings of the National Academy of Sciences*, 99(5):2754–2759, 2002.

[65] Luis del Pozo-Yauner, Jonathan S Wall, Martín González Andrade, Rosana Sánchez-López, Sandra L Rodríguez-Ambriz, Julio I Pérez Carreón, Adrián Ochoa-Leyva, and D Alejandro Fernández-Velasco. The N-terminal strand modulates immunoglobulin light chain fibrillogenesis. *Biochemical and biophysical research communications*, 443(2):495–499, 2014.

[66] Lawreen H Connors, Yan Jiang, Marianna Budnik, Roger Théberge, Tatiana Prokaeva, Kip L Bodi, David C Seldin, Catherine E Costello, and Martha Skinner. Heterogeneity in Primary Structure, Post-Translational Modifications, and Germline Gene Usage of Nine Full-Length Amyloidogenic κ1 Immunoglobulin Light Chains. *Biochemistry*, 46(49):14259–14271, 2007.

[67] Fred J Stevens. Four structural risk factors identify most fibril-forming kappa light chains. *Amyloid*, 7(3):200–211, 2000.

[68] Morie A Gertz. Immunoglobulin light chain amyloidosis: 2012 update on diagnosis, prognosis, and treatment. *American Journal of Hematology*, 87(2):183–189, 2012.

[69] Francesco Cappelli, Samuele Baldasseroni, Franco Bergesio, Stefano Perlini, Francesco Salinaro, Luigi Padeletti, Paola Attana, Alessandro Paoletti Perini, Alberto Moggi Pignone, Elisa Grifoni, et al. Echocardiographic and Biohumoral Characteristics in Patients with AL and TTR Amyloidosis at Diagnosis. *Clinical cardiology*, 38(2):69–75, 2015.

[70] Michal Fikrle, Tomas Palecek, Martin Masek, Petr Kuchynka, Jan Straub, Ivan Spicka, Romana Rysava, and Ales Linhart. The diagnostic performance of cardiac magnetic resonance in detection of myocardial involvement in al amyloidosis. *Clinical physiology and functional imaging*, 2014.

[71] Anupama Bhat, Carlo Selmi, Stanley M Naguwa, Gurtej S Cheema, and M Eric Gershwin. Currents Concepts on the Immunopathology of Amyloidosis. *Clinical reviews in allergy & immunology*, 38(2-3):97–106, 2010.

[72] Barbara Kieninger, Magdalena Eriksson, Reinhard Kandolf, Philipp A Schnabel, Stefan Schönland, Arnt V Kristen, Ute Hegenbart, Peter Lohse, and Christoph Röcken. Amyloid in endomyocardial biopsies. *Virchows Archiv*, 456(5):523–532, 2010.

[73] Tilmann Bochtler, Ute Hegenbart, Christiane Heiss, Axel Benner, Friedrich Cremer, Martin Volkmann, Jochen Ludwig, Jolanta B Perz, Anthony D Ho, Hartmut Goldschmidt, et al. Evaluation of the serum-free light chain test in untreated patients with AL amyloidosis. *Haematologica*, 93(3):459–462, 2008.

[74] Shaji K Kumar, Angela Dispenzieri, Martha Q Lacy, Suzanne R Hayman, Francis K Buadi, Steven R Zeldenrust, Tow Tan, Shirshendu Sinha, Nelson Leung, Robert A Kyle, et al. Changes in serum-free light chain rather than intact monoclonal immunoglobulin levels predicts outcome following therapy in primary amyloidosis. *American journal of hematology*, 86(3):251–255, 2011.

[75] Brendan M Weiss, Joseph Hebreo, Daniel V Cordaro, Mark J Roschewski, Thomas P Baker, Kevin C Abbott, and Stephen W Olson. Increased Serum Free Light Chains Precede the Presentation of Immunoglobulin Light Chain Amyloidosis. *Journal of Clinical Oncology*, 32(25):2699–2704, 2014.

[76] Arnt V Kristen, Evangelos Giannitsis, Stephanie Lehrke, Ute Hegenbart, Matthias Konstandin, David Lindenmaier, Corina Merkle, Stefan Hardt, Philipp A Schnabel, Christoph Röcken, et al. Assessment of disease severity and outcome in patients with systemic light-chain amyloidosis by the high-sensitivity troponin T assay. *Blood*, 116(14):2455–2461, 2010.

[77] Giovanni Palladini, Andrea Foli, Paolo Milani, Paola Russo, Riccardo Albertini, Francesca Lavatelli, Laura Obici, Stefano Perlini, Remigio Moratti, and Giampaolo Merlini. Best use of cardiac biomarkers in patients with AL amyloidosis and renal failure. *American journal of hematology*, 87(5):465–471, 2012.

[78] Stefan O Schönland, Peter Dreger, Theo De Witte, and Ute Hegenbart. Current status of hematopoietic cell transplantation in the treatment of systemic amyloid light-chain amyloidosis. *Bone marrow transplantation*, 47(7):895–905, 2012.

[79] Saulius Girnius, David C Seldin, Hans K Meier-Ewert, J Mark Sloan, Karen Quillen, Frederick L Ruberg, John L Berk, Gheorghe Doros, and Vaishali Sanchorawala. Safety and efficacy of high-dose melphalan and auto-SCT in patients with AL amyloidosis and cardiac involvement. *Bone marrow transplantation*, 49(3), 2014.

[80] Donna E Reece, Ute Hegenbart, Vaishali Sanchorawala, Giampaolo Merlini, Giovanni Palladini, Joan Bladé, Jean-Paul Fermand, Hani Hassoun, Leonard Heffner, Vishal Kukreti, et al. Long-term follow-up from a phase 1/2 study of single-agent bortezomib in relapsed systemic AL amyloidosis. *Blood*, 124(16):2498–2506, 2014.

[81] Robert F Cornell, Xiabao Zhong, Carlos Arce-Lara, Ehab Atallah, Linda Blust, William R Drobyski, Timothy S Fenske, Marcelo C Pasquini, J Douglas Rizzo, Wael Saber, et al. Bortezomib-based induction for transplant ineligible AL amyloidosis and feasibility of later transplantation. *Bone marrow transplantation*, 50(7):914–917, 2015.

[82] Donna E Reece, Vaishali Sanchorawala, Ute Hegenbart, Giampaolo Merlini, Giovanni Palladini, Jean-Paul Fermand, Robert A Vescio, Xiangyang Liu, Yusri A Elsayed, Andrew Cakana, et al. Weekly and twice-weekly bortezomib in patients with systemic AL amyloidosis: results of a phase 1 dose-escalation study. *Blood*, 114(8):1489–1497, 2009.

[83] Duncan B Richards, Louise M Cookson, Alienor C Berges, Sharon V Barton, Thirusha Lane, James M Ritter, Marianna Fontana, James C Moon, Massimo Pinzani, Julian D Gillmore, et al. Therapeutic clearance of amyloid by antibodies to serum amyloid P component. *New England Journal of Medicine*, 373(12):1106–1114, 2015.

[84] Mark B Pepys, J Herbert, Winston L Hutchinson, Glenys A Tennent, HJ Lachmann, J Ruth Gallimore, LB Lovat, T Bartfai, A Alanine, C Hertel, et al. Targeted pharmacological depletion of serum amyloid P component for treatment of human amyloidosis. *Nature*, 417(6886):254–259, 2002.

[85] Karl Bodin, Stephan Ellmerich, Melvyn C Kahan, Glenys A Tennent, Andrzej Loesch, Janet A Gilbertson, Winston L Hutchinson, Palma P Mangione, J Ruth Gallimore, David J Millar, et al. Antibodies to human serum amyloid P component eliminate visceral amyloid deposits. *Nature*, 468(7320):93–97, 2010.

[86] Raya Al-Shawi, Glenys A Tennent, David J Millar, Angela Richard-Londt, Sebastian Brandner, David J Werring, J Paul Simons, and Mark B Pepys. Pharmacological removal of serum amyloid P component from intracerebral plaques and cerebrovascular Aβ amyloid deposits in vivo. *Open biology*, 6(2):150202, 2016.

[87] Christina B Cooley, Lisa M Ryno, Lars Plate, Gareth J Morgan, John D Hulleman, Jeffery W Kelly, and R Luke Wiseman. Unfolded protein response activation reduces secretion and extracellular aggregation of amyloidogenic immunoglobulin light chain. *Proceedings of the National Academy of Sciences*, 111(36):13046–13051, 2014.

[88] Boris Brumshtein, Shannon R Esswein, Lukasz Salwinski, Martin L Phillips, Alan T Ly, Duilio Cascio, Michael R Sawaya, and David S Eisenberg. Inhibition by small-molecule ligands of formation of amyloid fibrils of an immunoglobulin light chain variable domain. *eLife*, page e10935, 2015.

[89] Nikolai Lorenzen, Søren B Nielsen, Yuichi Yoshimura, Brian S Vad, Camilla Bertel Andersen, Cristine Betzer, Jørn D Kaspersen, Gunna Christiansen, Jan S Pedersen, Poul Henning Jensen, et al. How epigallocatechin gallate can inhibit α-synuclein oligomer toxicity in vitro. *Journal of Biological Chemistry*, 289(31):21299–21310, 2014.

[90] Juan Miguel Lopez del Amo, Uwe Fink, Muralidhar Dasari, Gerlinde Grelle, Erich E Wanker, Jan Bieschke, and Bernd Reif. Structural Properties of EGCG-Induced, Nontoxic Alzheimer's Disease Aβ Oligomers. *Journal of molecular biology*, 421(4):517–524, 2012.

[91] Dagmar E Ehrnhoefer, Martin Duennwald, Phoebe Markovic, Jennifer L Wacker, Sabine Engemann, Margaret Roark, Justin Legleiter, J Lawrence Marsh, Leslie M Thompson, Susan Lindquist, Paul Muchowski, and Erich Wanker. Green tea (-)-epigallocatechin-gallate modulates early events in huntingtin misfolding and reduces toxicity in Huntington's disease models. *Human molecular genetics*, 15(18):2743–2751, 2006.

[92] Qianqian Wang, Jingjing Guo, Pingzu Jiao, Huanxiang Liu, and Xiaojun Yao. Exploring the Influence of EGCG on the β-Sheet-Rich Oligomers of Human Islet Amyloid Polypeptide (hIAPP 1–37) and Identifying Its Possible Binding Sites from Molecular Dynamics Simulation. *PloS one*, 9(4):e94796, 2014.

[93] Masanori Miyata, Takashi Sato, Miyuki Kugimiya, Misato Sho, Teruya Nakamura, Shinji Ikemizu, Mami Chirifu, Mineyuki Mizuguchi, Yuko Nabeshima, Yoshiaki Suwa, et al. The Crystal Structure of the Green Tea Polyphenol (-)-Epigallocatechin Gallate- Transthyretin Complex Reveals a Novel Binding Site Distinct from the Thyroxine Binding Site. *Biochemistry*, 49(29):6104–6114, 2010.

[94] Heike J Wobst, Apurwa Sharma, Marc I Diamond, Erich E Wanker, and Jan Bieschke. The green tea polyphenol (-)-epigallocatechin gallate prevents the aggregation of tau protein into toxic oligomers at substoichiometric ratios. *FEBS letters*, 589(1):77–83, 2015.

[95] Susmita Bhattacharya, Nitin K Pandey, Anushree Roy, and Swagata Dasgupta. Effect of (-)-epigallocatechin gallate on the fibrillation of human serum albumin. *International journal of biological macromolecules*, 70:312–319, 2014.

[96] Laura M Castellano, Rebecca M Hammond, Veronica M Holmes, Drew Weissman, and James Shorter. Epigallocatechin-3-gallate rapidly remodels PAP85-120, SEM1 (45-107), and SEM2 (49-107) seminal amyloid fibrils. *Biology open*, 4(9):1206–1212, 2015.

[97] Na Wang, Jianwei He, Alan K Chang, Yu Wang, Linan Xu, Xiaoying Chong, Xian Lu, Yonghui Sun, Xichun Xia, Hui Li, et al. (-)-Epigallocatechin-3-gallate Inhibits Fibrillogenesis of Chicken Cystatin. *Journal of agricultural and food chemistry*, 63(5):1347–1351, 2015.

[98] Stéphane Quideau, Denis Deffieux, Celine Douat-Casassus, and Laurent Pouységu. Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. *Angewandte Chemie International Edition*, 50(3):586–621, 2011.

[99] Dale G Nagle, Daneel Ferreira, and Yu-Dong Zhou. Epigallocatechin-3-Gallate (EGCG): Chemical and Biomedical Perspectives. *Phytochemistry*, 67(17):1849–1855, 2006.

[100] Nathan R Perron and Julia L Brumaghim. A Review of the Antioxidant Mechanisms of Polyphenol Compounds Related to Iron Binding. *Cell biochemistry and biophysics*, 53(2):75–100, 2009.

[101] Sonish Azam, Naghma Hadi, Nizam Uddin Khan, and Sheikh M Hadi. Prooxidant property of green tea polyphenols epicatechin and epigallocatechin-3-gallate: implications for anticancer properties. *Toxicology in vitro*, 18(5):555–561, 2004.

[102] Yingdong Zhu, Yantao Zhao, Pei Wang, Mohamed Ahmedna, Chi-Tang Ho, and Shengmin Sang. Tea Flavanols Block Advanced Glycation of Lens Crystallins Induced by Dehydroascorbic Acid. *Chemical research in toxicology*, 28(1):135–143, 2014.

[103] Wanda C Reygaert. The antimicrobial possibilities of green tea. *Front Microbiol*, 5:434, 2014.

[104] Jong-Hwa Jang, Yong-Duk Park, Hyo-Kwang Ahn, Seung-Jin Kim, Joo-Young Lee, Eun-Cheol Kim, Yeon-Soo Chang, Yun-Jung Song, and Ha-Jeong Kwon. Analysis of Green Tea Compounds and Their Stability in Dentifrices of Different pH Levels. *Chemical and Pharmaceutical Bulletin*, 62(4):328–335, 2014.

[105] Angela Dispenzieri, Morie A Gertz, and Francis Buadi. What do I need to know about immunoglobulin light chain (AL) amyloidosis? *Blood reviews*, 26(4):137–154, 2012.

[106] Werner Hunstein. Epigallocathechin-3-gallate in AL amyloidosis: a new therapeutic option? *Blood*, 110(6):2216–2216, 2007.

[107] Dagmar E Ehrnhoefer, Jan Bieschke, Annett Boeddrich, Martin Herbst, Laura Masino, Rudi Lurz, Sabine Engemann, Annalisa Pastore, and Erich E Wanker. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nature structural & molecular biology*, 15(6):558–566, 2008.

[108] Werner Hunstein. Details of my illness. *accessed 30. Mar 2016*, http://www.hunstein-egcg.de/tven.html, 2010.

[109] Derliz Mereles, Sebastian J Buss, Stefan E Hardt, Werner Hunstein, and Hugo A Katus. Effects of the main green tea polyphenol epigallocatechin-3-gallate on cardiac involvement in patients with AL amyloidosis. *Clinical research in cardiology*, 99(8):483–490, 2010.

[110] Stefan Schönland. TAME-AL. *ClinicalTrials.gov Identifier: NCT02015312*, 2013.

[111] Giampaolo Merlini. EpiCardiAL. *ClinicalTrials.gov Identifier: NCTNCT01511263*, 2012.

[112] Jan Bieschke, Jenny Russ, Ralf P Friedrich, Dagmar E Ehrnhoefer, Heike Wobst, Katja Neugebauer, and Erich E Wanker. EGCG remodels mature α-synuclein and amyloid-β fibrils and reduces cellular toxicity. *Proceedings of the National Academy of Sciences*, 107(17):7710–7715, 2010.

[113] Angel E Pelaez-Aguilar, Lina Rivillas-Acevedo, Leidys French-Pacheco, Gilberto Valdes-Garcia, Roberto Maya-Martinez, Nina Pastor, and Carlos Amero. Inhibition of Light Chain 6aJL2-R24G Amyloid Fiber Formation Associated with Light Chain Amyloidosis. *Biochemistry*, 54(32):4978–4986, 2015.

[114] James Keeler. *Understanding NMR spectroscopy*. John Wiley & Sons, 2011.

[115] Mike P Williamson. Using chemical shift perturbation to characterise ligand binding. *Progress in nuclear magnetic resonance spectroscopy*, 73:1–16, 2013.

[116] Mark A McCoy and Daniel F Wyss. Spatial Localization of Ligand Binding Sites from Electron Current Density Surfaces Calculated from NMR Chemical Shift Perturbations. *Journal of the American Chemical Society*, 124(39):11758–11763, 2002.

[117] Malcolm H Levitt. *Spin Dynamics: Basics of Nuclear Magnetic Resonance, 2nd Edition.* John Wiley & Sons, 2008.

[118] Konstantin Pervushin, Roland Riek, Gerhard Wider, and Kurt Wüthrich. Attenuated T_2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences*, 94(23):12366–12371, 1997.

[119] Edward R Andrew, A Bradbury, and Robert G Eades. Nuclear Magnetic Resonance Spectra From A Crystal Rotated At High Speed. *Letters to Nature*, 1958.

[120] Irving J Lowe. Free Induction Decays of Rotating Solids. *Physical Review Letters*, 2(7):285, 1959.

[121] Melinda J Duer. *Solid state NMR spectroscopy: principles and applications.* John Wiley & Sons, 2008.

[122] Rebecca S Lipsitz and Nico Tjandra. Carbonyl CSA Restraints from Solution NMR for Protein Structure Refinement. *Journal of the American Chemical Society*, 123(44):11065–11066, 2001.

[123] Silvia Spera and Ad Bax. Empirical Correlation between Protein Backbone Conformation and Cα and Cβ ^{13}C Nuclear Magnetic Resonance Chemical Shifts. *Journal of the American Chemical Society*, 113(14):5490–5492, 1991.

[124] Yang Shen, Frank Delaglio, Gabriel Cornilescu, and Ad Bax. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR*, 44(4):213–223, 2009.

[125] Yang Shen, Oliver Lange, Frank Delaglio, Paolo Rossi, James M Aramini, Gaohua Liu, Alexander Eletsky, Yibing Wu, Kiran K Singarapu, Alexander Lemak, et al. Consistent blind protein structure generation from NMR chemical shift data. *Proceedings of the National Academy of Sciences*, 105(12):4685–4690, 2008.

[126] Michael Sattler, Jürgen Schleucher, and Christian Griesinger. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution. *Progress in nuclear magnetic resonance spectroscopy*, 34:93–158, 1999.

[127] Eldon L Ulrich, Hideo Akutsu, Jurgen F Doreleijers, Yoko Harano, Yannis E Ioannidis, Jundong Lin, Miron Livny, Steve Mading, Dimitri Maziuk, Zachary Miller, et al. BioMagResBank. *Nucleic acids research*, 36(suppl 1):D402–D408, 2008.

[128] Juan A Aguilar, Stephen Faulkner, Mathias Nilsson, and Gareth A Morris. Pure Shift ^{1}H NMR: A Resolution Of The Resolution Problem? *Angewandte Chemie International Edition*, 49(23):3901–3903, 2010.

[129] Gerald B Matson. Signal Integration and the Signal-to-Noise Ratio in Pulsed NMR Relaxation Measurements. *Journal of Magnetic Resonance (1969)*, 25(3):477–480, 1977.

[130] Jennifer CJ Barna and Ernest D Laue. Conventional and Exponential Sampling for 2D NMR Experiments with Application to a 2D NMR Spectrum of a Protein. *Journal of Magnetic Resonance (1969)*, 75(2):384–389, 1987.

[131] Iddo Drori. Fast Minimization by Iterative Thresholding for Multidimensional NMR Spectroscopy. *EURASIP Journal on Advances in Signal Processing*, 2007(1):1–10, 2007.

[132] Sven G Hyberts, Alexander G Milbradt, Andreas B Wagner, Haribabu Arthanari, and Gerhard Wagner. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled with Multidimensional Poisson Gap Scheduling. *Journal of biomolecular NMR*, 52(4):315–327, 2012.

[133] Sven G Hyberts, Scott A Robson, and Gerhard Wagner. Exploring Signal-to-Noise Ratio and Sensitivity in Non-Uniformly Sampled Multi-Dimensional NMR Spectra. *Journal of biomolecular NMR*, 55(2):167–178, 2013.

[134] Melissa R Palmer, Christopher L Suiter, Geneive E Henry, James Rovnyak, Jeffrey C Hoch, Tatyana Polenova, and David Rovnyak. Sensitivity of Nonuniform Sampling NMR. *The Journal of Physical Chemistry B*, 119(22):6502–6515, 2015.

[135] Erwin L Hahn. Spin Echoes. *Physical review*, 80(4):580, 1950.

[136] Malcolm H Levitt and Ray Freeman. Composite Pulse Decoupling. *Journal of Magnetic Resonance (1969)*, 43(3):502–507, 1981.

[137] Robert Powers, Angela M Gronenborn, G Marius Clore, and Ad Bax. Three-Dimensional Triple-Resonance NMR of $^{13}C/^{15}N$-Enriched Proteins Using Constant-Time Evolution. *Journal of Magnetic Resonance (1969)*, 94(1):209–213, 1991.

[138] Veniamin Chevelkov, Zhongjing Chen, Wolfgang Bermel, and Bernd Reif. Resolution enhancement in MAS solid-state NMR by application of ^{13}C homonuclear scalar decoupling during acquisition. *Journal of Magnetic Resonance*, 172(1):56–62, 2005.

[139] Emeline Barbet-Massin and Guido Pintacuda. Biomolecular Solid-State NMR/Basics. *NMR of Biomolecules: Towards Mechanistic Systems Biology*, pages 345–364, 2012.

[140] Yue Qi Ye, Michal Malon, Charlotte Martineau, Francis Taulelle, and Yusuke Nishiyama. Rapid measurement of multidimensional ^{1}H solid-state NMR spectra at ultra-fast MAS frequencies. *Journal of Magnetic Resonance*, 239:75–80, 2014.

[141] Vipin Agarwal, Susanne Penzel, Kathrin Szekely, Riccardo Cadalbert, Emilie Testori, Andres Oss, Jaan Past, Ago Samoson, Matthias Ernst, Anja Böckmann, et al. De Novo 3D Structure Determination from Sub-milligram Protein Samples by Solid-State 100 kHz MAS NMR Spectroscopy. *Angewandte Chemie International Edition*, 53(45):12253–12256, 2014.

[142] Ann E McDermott, FJ Creuzet, AC Kolbert, and Robert G Griffin. High-Resolution Magic-Angle-Spinning NMR Spectra of Protons in Deuterated Solids. *Journal of Magnetic Resonance (1969)*, 98(2):408–413, 1992.

[143] Ling Zheng, Kenneth W Fishbein, Robert G Griffin, and Judith Herzfeld. Two-Dimensional Solid-State Proton NMR and Proton Exchange. *Journal of the American Chemical Society*, 115(14):6254–6261, 1993.

[144] Bernd Reif, Christopher P Jaroniec, Chad M Rienstra, Morten Hohwy, and RG Griffin. ^{1}H-^{1}H MAS Correlation Spectroscopy and Distance Measurements in a Deuterated Peptide. *Journal of Magnetic Resonance*, 151(2):320–327, 2001.

[145] Vipin Agarwal, Anne Diehl, Nikolai Skrynnikov, and Bernd Reif. High Resolution ^{1}H detected ^{1}H, ^{13}C Correlation Spectra in MAS Solid-State NMR using Deuterated Proteins with Selective ^{1}H, ^{2}H Isotopic Labeling of Methyl Groups. *Journal of the American Chemical Society*, 128(39):12620–12621, 2006.

[146] Sam Asami, Peter Schmieder, and Bernd Reif. High Resolution ^{1}H-Detected Solid-State NMR Spectroscopy of Protein Aliphatic Resonances: Access to Tertiary Structure Information. *Journal of the American Chemical Society*, 132(43):15133–15135, 2010.

[147] Rasmus Linser, Benjamin Bardiaux, Loren B Andreas, Sven G Hyberts, Vanessa K Morris, Guido Pintacuda, Margaret Sunde, Ann H Kwan, and Gerhard Wagner. Solid-State NMR Structure Determination from Diagonal-Compensated, Sparsely Nonuniform-Sampled 4D Proton-Proton Restraints. *Journal of the American Chemical Society*, 136(31):11002–11010, 2014.

[148] Chaowei Shi, Pascal Fricke, Lin Lin, Veniamin Chevelkov, Melanie Wegstroth, Karin Giller, Stefan Becker, Martin Thanbichler, and Adam Lange. Atomic-resolution structure of cytoskeletal bactofilin by solid-state NMR. *Science advances*, 1(11):e1501087, 2015.

[149] Mei Hong and Karen Jakes. Selective and extensive ^{13}C labeling of a membrane protein for solid-state NMR investigations. *Journal of biomolecular NMR*, 14(1):71–74, 1999.

[150] Victoria A Higman, Jeremy Flinders, Matthias Hiller, Stefan Jehle, Stefan Markovic, Sebastian Fiedler, Barth-Jan van Rossum, and Hartmut Oschkinat. Assigning large proteins in the solid state: a MAS NMR resonance assignment strategy using selectively and extensively ^{13}C-labelled proteins. *Journal of biomolecular NMR*, 44(4):245–260, 2009.

[151] Antoine Loquet, Guohua Lv, Karin Giller, Stefan Becker, and Adam Lange. ^{13}C Spin Dilution for Simplified and Complete Solid-State NMR Resonance Assignment of Insoluble Biological Assemblies. *Journal of the American Chemical Society*, 133(13):4722–4725, 2011.

[152] Federica Castellani, Barth van Rossum, Annette Diehl, Mario Schubert, Kristina Rehbein, and Hartmut Oschkinat. Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature*, 420(6911):98–102, 2002.

[153] Marcus D Tuttle, Gemma Comellas, Andrew J Nieuwkoop, Dustin J Covell, Deborah A Berthold, Kathryn D Kloepper, Joseph M Courtney, Jae K Kim, Alexander M Barclay, Amy Kendall, et al. Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. *Nature Structural & Molecular Biology*, 2016.

[154] Jerome Kaplan. Exchange Broadening in Nuclear Magnetic Resonance. *The Journal of Chemical Physics*, 28(2):278–282, 1958.

[155] Herman Y Carr and Edward M Purcell. Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Physical review*, 94(3):630, 1954.

[156] Saul Meiboom and David Gill. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Review of scientific instruments*, 29(8):688–691, 1958.

[157] Arthur G Palmer, Michael J Grey, and Chunyu Wang. Solution NMR Spin Relaxation Methods for Characterizing Chemical Exchange in High-Molecular-Weight Systems. *Methods in enzymology*, 394:430–465, 2005.

[158] Ivano Bertini, Frank Engelke, Leonardo Gonnelli, Benno Knott, Claudio Luchinat, David Osen, and Enrico Ravera. On the use of ultracentrifugal devices for sedimented solute NMR. *Journal of biomolecular NMR*, 54(2):123–127, 2012.

[159] Frank Delaglio, Stephan Grzesiek, Geerten W Vuister, Guang Zhu, John Pfeifer, and AD Bax. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR*, 6(3):277–293, 1995.

[160] Ole Tange. GNU Parallel: The Command-Line Power Tool. *The USENIX Magazine*, 36(1):42–47, 2011.

[161] Wim F Vranken, Wayne Boucher, Tim J Stevens, Rasmus H Fogh, Anne Pajon, Miguel Llinas, Eldon L Ulrich, John L Markley, John Ionides, and Ernest D Laue. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins: Structure, Function, and Bioinformatics*, 59(4):687–696, 2005.

[162] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. UCSF Chimera - a Visualization System for Exploratory Research and Analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.

[163] Andrew M Waterhouse, James B Procter, David MA Martin, Michèle Clamp, and Geoffrey J Barton. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9):1189–1191, 2009.

[164] Mark A Larkin, Gordon Blackshields, NP Brown, R Chenna, Paul A McGettigan, Hamish McWilliam, Franck Valentin, Iain M Wallace, Andreas Wilm, Rodrigo Lopez, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21):2947–2948, 2007.

[165] Gavin E Crooks, Gary Hon, John-Marc Chandonia, and Steven E Brenner. WebLogo: a Sequence Logo Generator. *Genome research*, 14(6):1188–1190, 2004.

[166] Thomas Williams, Colin Kelley, HB Broker, EA Merritt, J Campbell, R Cunningham, D Denholm, G Elber, R Fearick, C Grammes, et al. Gnuplot 4.4: an interactive plotting program. *Official gnuplot documentation, http://sourceforge. net/projects/gnuplot*, 2010.

[167] Stanley Maloy. Codon usage in E. coli. *accessed 13. Apr 2016*, http://www.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/in-vitro-genetics/codon-usage.html, 2002.

[168] Elisabeth Gasteiger, Christine Hoogland, Alexandre Gattiker, S'everine Duvaud, Marc R Wilkins, Ron D Appel, and Amos Bairoch. *Protein Identification and Analysis Tools on the ExPASy Server*. Springer, 2005.

[169] Hermann Schägger. Tricine - SDS-PAGE. *Nature Protocols*, 1:16–22, 2006.

[170] Minna Groenning, Mathias Norrman, James M Flink, Marco van de Weert, Jens T Bukrinsky, Gerd Schluckebier, and Sven Frokjaer. Binding mode of Thioflavin T in insulin amyloid fibrils. *Journal of structural biology*, 159(3):483–497, 2007.

[171] Bing M Fung, Anatoly K Khitrin, and Konstantin Ermolaev. An improved broadband decoupling sequence for liquid crystals and solids. *Journal of Magnetic Resonance*, 142(1):97–101, 2000.

[172] Andreas Grommek, Beat H Meier, and Matthias Ernst. Distance information from proton-driven spin diffusion under MAS. *Chemical physics letters*, 427(4):404–409, 2006.

[173] Kiyonori Takegoshi, Shinji Nakamura, and Takehiko Terao. ^{13}C-^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters*, 344(5):631–637, 2001.

[174] Marc Baldus, Aneta T Petkova, Judith Herzfeld, and Robert G Griffin. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Molecular Physics*, 95(6):1197–1207, 1998.

[175] Jutta Pauli, Marc Baldus, Barth van Rossum, Huub de Groot, and Hartmut Oschkinat. Backbone and Side-Chain ^{13}C and ^{15}N Signal Assignments of the α-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem*, 2(4):272–281, 2001.

[176] Anne Schuetz, Christian Wasmer, Birgit Habenstein, René Verel, Jason Greenwald, Roland Riek, Anja Böckmann, and Beat H Meier. Protocols for the Sequential Solid-State NMR Spectroscopic Assignment of a Uniformly Labeled 25 kDa Protein: HET-s (1-227). *ChemBioChem*, 11(11):1543–1551, 2010.

[177] Veniamin Chevelkov, Chaowei Shi, Hannes Klaus Fasshuber, Stefan Becker, and Adam Lange. Efficient band-selective homonuclear CO-CA cross-polarization in protonated proteins. *Journal of biomolecular NMR*, 56(4):303–311, 2013.

[178] René Verel, Marc Baldus, Matthias Ernst, and Beat H Meier. A homonuclear spin-pair filter for solid-state NMR based on adiabatic-passage techniques. *Chemical physics letters*, 287(3):421–428, 1998.

[179] Maili Liu, Xi-an Mao, Chaohui Ye, He Huang, Jeremy K Nicholson, and John C Lindon. Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. *Journal of Magnetic Resonance*, 132(1):125–129, 1998.

[180] Rüdiger Weisemann, Heinz Rüterjans, and Wolfgang Bermel. 3D Triple-resonance NMR techniques for the sequential assignment of NH and ^{15}N resonances in ^{15}N-and ^{13}C-labelled proteins. *Journal of biomolecular NMR*, 3(1):113–120, 1993.

[181] Oleg Trott and Arthur J Olson. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2):455–461, 2010.

[182] Sjoerd J De Vries, Marc van Dijk, and Alexandre MJJ Bonvin. The HADDOCK web server for data-driven biomolecular docking. *Nature protocols*, 5(5):883–897, 2010.

[183] Berk Hess, Carsten Kutzner, David Van Der Spoel, and Erik Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of chemical theory and computation*, 4(3):435–447, 2008.

[184] Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L Klepeis, Ron O Dror, and David E Shaw. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8):1950–1958, 2010.

[185] Hans W Horn, William C Swope, Jed W Pitera, Jeffry D Madura, Thomas J Dick, Greg L Hura, and Teresa Head-Gordon. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of chemical physics*, 120(20):9665–9678, 2004.

[186] Alan W Sousa da Silva and Wim F Vranken. ACPYPE-AnteChamber PYthon Parser interfacE. *BMC research notes*, 5(1):367, 2012.

[187] Junmei Wang, Romain M Wolf, James W Caldwell, Peter A Kollman, and David A Case. Development and Testing of a General Amber Force Field. *Journal of computational chemistry*, 25(9):1157–1174, 2004.

[188] Junmei Wang, Wei Wang, Peter A Kollman, and David A Case. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of molecular graphics and modelling*, 25(2):247–260, 2006.

[189] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An $N \cdot \log (N)$ method for Ewald sums in large systems. *The Journal of chemical physics*, 98(12):10089–10092, 1993.

[190] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of chemical physics*, 81(1):511–519, 1984.

[191] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

[192] Naveen Michaud-Agrawal, Elizabeth J Denning, Thomas B Woolf, and Oliver Beckstein. MDAnalysis: a Toolkit for the Analysis of Molecular Dynamics Simulations. *Journal of computational chemistry*, 32(10):2319–2327, 2011.

[193] Cardine N Nokwe, Manuel Hora, Martin Zacharias, Hisashi Yagi, Jirka Peschek, Bernd Reif, Yuji Goto, and Johannes Buchner. A Stable Mutant Predisposes Antibody Domains to Amyloid Formation through Specific Non-Native Interactions. *Journal of Molecular Biology*, 2016.

[194] Simon S Terzyan, Christina R Bourne, Paul A Ramsland, Philip C Bourne, and Allen B Edmundson. Comparison of the three-dimensional structures of a human Bence-Jones dimer crystallized on Earth and aboard US Space Shuttle Mission STS-95. *Journal of Molecular Recognition*, 16(2):83–90, 2003.

[195] Tanya L Poshusta, Laura A Sikkink, Nelson Leung, Raynell J Clark, Angela Dispenzieri, and Marina Ramirez-Alvarado. Mutations in Specific Structural Regions of Immunoglobulin Light Chains are Associated with Free Light Chain Levels in Patients with AL Amyloidosis. *PLoS One*, 4(4):e5169, 2009.

[196] Ramadurgam Kodandapani, B Veerapandian, Thomas J Kunicki, and Kathryn R Ely. Crystal Structure of the OPG2 Fab. *Journal of Biological Chemistry*, 270(5):2268–2273, 1995.

[197] Cardine N Nokwe, Martin Zacharias, Hisashi Yagi, Manuel Hora, Bernd Reif, Yuji Goto, and Johannes Buchner. A Residue-specific Shift in Stability and Amyloidogenicity of Antibody Variable Domains. *Journal of Biological Chemistry*, 289(39):26829–26846, 2014.

[198] Cardine N Nokwe, Manuel Hora, Martin Zacharias, Hisashi Yagi, Christine John, Bernd Reif, Yuji Goto, and Johannes Buchner. The Antibody Light-Chain Linker Is Important for Domain Stability and Amyloid Formation. *Journal of Molecular Biology*, 427(22):3572–3586, 2015.

[199] Matthias J Feige, Sandra Groscurth, Moritz Marcinowski, Vincent Truffault, Emanuele Paci, Horst Kessler, Johannes Buchner, et al. The structure of a folding intermediate provides insight into differences in immunoglobulin amyloidogenicity. *Proceedings of the National Academy of Sciences*, 105(36):13373–13378, 2008.

[200] Bon W Koo and Andrew D Miranker. Contribution of the intrinsic disulfide to the assembly mechanism of islet amyloid. *Protein Science*, 14(1):231–239, 2005.

[201] Alexandre I Ilitchev, Maxwell J Giammona, Thanh D Do, Amy G Wong, Steven K Buratto, Joan-Emma Shea, Daniel P Raleigh, and Michael T Bowers. Human Islet Amyloid Polypeptide N-Terminus Fragment Self-Assembly: Effect of Conserved Disulfide Bond on Aggregation Propensity. *Journal of The American Society for Mass Spectrometry*, pages 1–9, 2016.

[202] Yiwen Chen and Nikolay V Dokholyan. A Single Disulfide Bond Differentiates Aggregation Pathways of β2-Microglobulin. *Journal of molecular biology*, 354(2):473–482, 2005.

[203] Kaori Yamamoto, Hisashi Yagi, Daisaku Ozawa, Kenji Sasahara, Hironobu Naiki, and Yuji Goto. Thiol Compounds Inhibit the Formation of Amyloid Fibrils by β 2-Microglobulin at Neutral pH. *Journal of molecular biology*, 376(1):258–268, 2008.

[204] Madhuri Chattopadhyay, Ekeoma Nwadibia, Cynthia D Strong, Edith Butler Gralla, Joan Selverstone Valentine, and Julian P Whitelegge. The Disulfide Bond, but Not Zinc or Dimerization, Controls Initiation and Seeded Growth in Amyotrophic Lateral Sclerosis-linked Cu, Zn Superoxide Dismutase (SOD1) Fibrillation. *Journal of Biological Chemistry*, 290(51):30624–30636, 2015.

[205] Brian O'Nuallain, Amy Allen, Stephen J Kennel, Deborah T Weiss, Alan Solomon, and Jonathan S Wall. Localization of a Conformational Epitope Common to Non-Native and Fibrillar Immunoglobulin Light Chains. *Biochemistry*, 46(5):1240–1247, 2007.

[206] Sujoy Mukherjee, Simon P Pondaven, and Christopher P Jaroniec. Conformational Flexibility of a Human Immunoglobulin Light Chain Variable Domain by Relaxation Dispersion Nuclear Magnetic Resonance Spectroscopy: Implications for Protein Misfolding and Mmyloid Assembly. *Biochemistry*, 50(26):5845–5857, 2011.

[207] Genarro Esposito, R Michelutti, Giuliana Verdone, Paolo Viglino, Helena Hernandez, Carol V Robinson, Angela Amoresano, Fabrizio Dal Piaz, Maria Monti, Pietro Pucci, et al. Removal of the N-terminal hexapeptide from human β 2-microglobulin facilitates protein aggregation and fibril formation. *Protein Science*, 9(05):831–845, 2000.

[208] Marina Ramirez-Alvarado. Amyloid formation in light chain amyloidosis. *Current topics in medicinal chemistry*, 12(22):2523, 2012.

[209] Kathrin Andrich and Jan Bieschke. The Effect of (-)-Epigallo-catechin-(3)-Gallate on Amyloidogenic Proteins Suggests a Common Mechanism. In *Natural Compounds as Therapeutic Agents for Amyloidogenic Diseases*, pages 139–161. Springer, 2015.

[210] Ann E Hagerman and Larry G Butler. The Specificity of Proanthocyanidin-Protein Interactions. *Journal of Biological Chemistry*, 256(9):4494–4497, 1981.

[211] Adrian J Charlton, Edwin Haslam, and Michael P Williamson. Multiple Conformations of the Proline-Rich Protein/Epigallocatechin Gallate Complex Determined by Time-Averaged Nuclear Overhauser Effects. *Journal of the American Chemical Society*, 124(33):9899–9905, 2002.

[212] Francis Canon, Franck Pate, Véronique Cheynier, Pascale Sarni-Manchado, Alexandre Giuliani, Javier Pérez, Dominique Durand, Joaquim Li, and Bernard Cabane. Aggregation of the Salivary Proline-Rich Protein IB5 in the Presence of the Tannin EgCG. *Langmuir*, 29(6):1926–1937, 2013.

[213] Mario Schubert, Dirk Labudde, Hartmut Oschkinat, and Peter Schmieder. A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ^{13}C chemical shift statistics. *Journal of biomolecular NMR*, 24(2):149–154, 2002.

[214] Karol Wroblewski, Ranjith Muhandiram, Avi Chakrabartty, and Anders Bennick. The molecular interaction of human salivary histatins with polyphenolic compounds. *European Journal of Biochemistry*, 268(16):4384–4397, 2001.

[215] Elisabeth Jöbstl, John O'Connell, J Patrick A Fairclough, and Mike P Williamson. Molecular Model for Astringency Produced by Polyphenol/Protein Interactions. *Biomacromolecules*, 5(3):942–949, 2004.

[216] Marcel Stenvang, Gunna Christiansen, and Daniel E Otzen. Epigallocatechin Gallate Remodels Fibrils of Lattice Corneal Dystrophy Protein, Facilitating Proteolytic Degradation and Preventing Formation of Membrane-Permeabilizing Species. *Biochemistry*, 55(16):2344–2357, 2016.

[217] Elke Prade, Heiko J Bittner, Riddhiman Sarkar, Juan Miguel Lopez del Amo, Gerhard Althoff-Ospelt, Gerd Multhaup, Peter W Hildebrand, and Bernd Reif. Structural Mechanism of the Interaction of Alzheimer Disease $A\beta$ Fibrils with the Non-steroidal Anti-inflammatory Drug (NSAID) Sulindac Sulfide. *Journal of Biological Chemistry*, 290(48):28737–28745, 2015.

[218] Marcus D Tuttle, Joseph M Courtney, Alexander M Barclay, and Chad M Rienstra. Preparation of Amyloid Fibrils for Magic-Angle Spinning Solid-State NMR Spectroscopy. *Protein Amyloid Aggregation: Methods and Protocols*, pages 173–183, 2016.

[219] Thomas Westfeld, René Verel, Matthias Ernst, Anja Böckmann, and Beat H Meier. Properties of the DREAM scheme and its optimization for application to proteins. *Journal of biomolecular NMR*, 53(2):103–112, 2012.

[220] Federica Castellani. *Structure Determination of Immobilized Proteins by Solid-State NMR Spectroscopy*. PhD thesis, Freie Universität Berlin, 2003.

[221] Andi Mainz, Stefan Jehle, Barth J van Rossum, Hartmut Oschkinat, and Bernd Reif. Large Protein Complexes with Extreme Rotational Correlation Times Investigated in Solution by Magic-Angle-Spinning NMR Spectroscopy. *Journal of the American Chemical Society*, 131(44):15968–15969, 2009.

[222] David S Wishart, Colin G Bigam, Arne Holm, Robert S Hodges, and Brian D Sykes. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of biomolecular NMR*, 5(1):67–81, 1995.

[223] Juan Lopez, Puneet Ahuja, Melanie Gerard, Jean-Michel Wieruszeski, and Guy Lippens. A new strategy for sequential assignment of intrinsically unstructured proteins based on ^{15}N single isotope labelling. *Journal of Magnetic Resonance*, 236:1–6, 2013.

[224] Itaru Yamaguchi, Kazuhiro Hasegawa, Naoki Takahashi, Fumitake Gejyo, and Hironobu Naiki. Apolipoprotein E Inhibits the Depolymerization of β2-Microglobulin-Related Amyloid Fibrils at a Neutral pH. *Biochemistry*, 40(29):8499–8507, 2001.

[225] Holger A Scheidt, Isabel Morgado, and Daniel Huster. Solid-state NMR reveals a close structural relationship between amyloid-β protofibrils and oligomers. *Journal of Biological Chemistry*, 287(27):22822–22826, 2012.

[226] T Szyperski, P Luginbühl, G Otting, P Güntert, and K Wüthrich. Protein dynamics studied by rotating frame ^{15}N spin relaxation times. *Journal of biomolecular NMR*, 3(2):151–164, 1993.

[227] Juan Miguel Lopez del Amo, Matthias Schmidt, Uwe Fink, Muralidar Dasari, Marcus Fändrich, and Bernd Reif. An Asymmetric Dimer as the Basic Subunit in Alzheimers Disease Amyloid β Fibrils. *Angewandte Chemie International Edition*, 51(25):6136–6139, 2012.

[228] József Kardos, András Micsonai, Henriett Pál-Gábor, Éva Petrik, László Gráf, János Kovács, Young-Ho Lee, Hironobu Naiki, and Yuji Goto. Reversible Heat-Induced Dissociation of β2-Microglobulin Amyloid Fibrils. *Biochemistry*, 50(15):3211–3220, 2011.

[229] Kristoffer Brännström, Anders Öhman, Malin Lindhagen-Persson, and Anders Olofsson. Ca^{2+} enhances Aβ polymerization rate and fibrillar stability in a dynamic manner. *Biochemical Journal*, 450(1):189–197, 2013.

[230] Francis E Dwulet, Kandice Strako, and Merrill D Benson. Amino Acid Sequence of a λ VI Primary (AL) Amyloid Protein (WLT). *Scandinavian journal of immunology*, 22(6):653–660, 1985.

[231] Vittorio Bellotti, Palma Mangione, and Giampaolo Merlini. Review: immunoglobulin light chain amyloidosisthe archetype of structural and pathogenic variability. *Journal of structural biology*, 130(2):280–289, 2000.

Supplementary Information

Sequence Information for the MAK33 Light Chain

Primary sequence of the MAK33 LC:

V_L:
DIVLTQSPAT LSVTPGDSVS LSCRASQSIS NNLHWYQQKS HESPRLLIKY ASQSISGIPS
RFSGSGSGTD FTLSINSVET EDFGMYFCQQ SNSWPLTFGA GTKLELKR

C_L:
ADAAPTVSIF PPSSEQLTSG GASVVCFLNN FYPKDINVKW KIDGSERQNG VLNSWTDQDS
KDSTYSMSST LTLTKDEYER HNSYTCEATH KTSTSPIVKS FNRNE

Gene sequences (optimised for expression in *E. coli*):

V_L:
ATG GAT ATT GTG CTA ACT CAG TCT CCA GCC ACC CTG TCT GTG ACT CCA GGA GAT AGC
GTC AGT CTT TCC TGC AGG GCC AGC CAA AGT ATT AGC AAC AAC CTA CAC TGG TAT CAA
CAA AAA TCA CAT GAG TCT CCA AGG CTT CTC ATC AAA TAT GCT TCC CAG TCC ATC TCT
GGG ATC CCC TCC AGG TTC AGT GGC AGT GGA TCA GGG ACA GAT TTC ACT CTC AGT ATC
AAC AGT GTG GAG ACT GAA GAT TTT GGA ATG TAT TTC TGT CAA CAG AGT AAC AGC TGG
CCT CTC ACG TTC GGT GCT GGG ACC AAG CTG GAG CTG AAA AGA TAA

C_L:
ATG GCT GAT GCT GCA CCA ACT GTA TCC ATC TTC CCA CCA TCC AGT GAG CAG TTA ACA
TCT GGA GGT GCC TCA GTC GTG TGC TTC TTG AAC AAC TTC TAC CCC AAA GAC ATC AAT
GTC AAG TGG AAG ATT GAT GGC AGT GAA CGA CAA AAT GGC GTC CTG AAC AGT TGG ACT
GAT CAG GAC AGC AAA GAC AGC ACC TAC AGC ATG AGC AGC ACC CTC ACG TTG ACC AAG
GAC GAG TAT GAA CGA CAT AAC AGC TAT ACC TGT GAG GCC ACT CAC AAG ACA TCA ACT
TCA CCC ATT GTC AAG AGC TTC AAC AGG AAT GAG

Experiments for the Assignment of Rigid Residues in MAK33 V_L S20N Fibrils

All MAS solid-state NMR experiments used for assignment of MAK33 V_L S20N fibrils are listed here. B_0 refers to the external magnetic field in units of MHz of ^{1}H Larmor frequency. NUS experiments were run with matched sampling on ^{15}N (time constant of the NUS density function is equal to the T_2 time of 9 ms). For ^{13}C, the NUS density function was set to 3 ms time constant. NS is the number of scans, F1, F2 and F3 the different dimensions. TD is the time domain, given in measured points and aq. is the acquisition time, given in ms.

type	B_0	sampling	labeling	NS	F1 TD	F1 aq.	F2 TD	F2 aq.	F3 TD	F3 aq.
PDSD	750	uniform	uniform	64	1024	11	2048	12		
N(CA)CX	750	uniform	uniform	288	132	15	3072	18		
N(CA)CX	750	50% NUS	uniform	2048	140	15	3072	18		
NCACX	750	uniform	uniform	24	60	7	94	8	2048	12
NCACX	750	50% NUS	uniform	32	70	8	84	7	1536	14
N(CO)CX	750	uniform	uniform	320	100	11	3072	18		
N(CO)CX	750	50% NUS	uniform	2048	140	15	3072	18		
NCOCX	750	uniform	uniform	32	70	8	60	9	3072	18
NCOCA	750	50% NUS	uniform	32	106	9	56	8	1536	14
CANCO	750	50% NUS	uniform	32	84	7	70	8	1634	14
N(CA)CB	750	uniform	uniform	512	110	12	1536	9		
NCACB	750	50% NUS	uniform	32	80	9	110	5	1024	6
N(COCA)CB	750	uniform	uniform	1088	90	10	1536	9		
PDSD	750	uniform	1,3 glyc.	32	650	6	2048	18		
PDSD	750	uniform	1,3 glyc.	48	1536	20	2048	22		
PDSD	500	uniform	1,3 glyc.	16	900	18	2048	27		
PDSD	400	uniform	2 glyc.	64	850	21	1024	24		
NCA	400	uniform	2 glyc.	1024	80	20	1536	19		

Curriculum Vitae

Manuel Hora

Date and Place of Birth: 22.2.1988 in München

Research Activities:

2012 - 2016	Dissertation in the group of Prof. Bernd Reif at the Technische Universität München Topic: Solid- and Solution-State Nuclear Magnetic Resonance Spectroscopic Analysis of Amyloidogenic Immunoglobulin Light Chains and their Interactions with Epigallocatechin Gallate	
2012	Master thesis in the group of Prof. Sevil Weinkauf at the Technische Universität München Topic: Structural Investigation of Human αA-crystallin by Transmission Electron Microscopy and Homology Modeling	
2011	Member of the team "TUM Nanu", supervised by Prof. Hendrik Dietz at the Technische Universität München; Participation in the Biomod 2011 Competition, organised by Harvard University, USA Topic: Design of a DNA Origami Device for Measuring Effects of DNA Binding Molecules on DNA Structure	

Education:

2010 - 2012	Master of Science in Biochemistry at the Technische Universität München	
2007 - 2010	Bachelor of Science in Biochemistry at the Technische Universität München	
2007	Allgemeine Hochschulreife (higher education entrance qualification) at the Apian-Gymnasium Ingolstadt	

Publications and Presentations

Parts of this thesis have been published in the following scientific articles:

Cardine N Nokwe, Martin Zacharias, Hisashi Yagi, <u>Manuel Hora</u>, Bernd Reif, Yuji Goto, Johannes Buchner:
A Residue-specific Shift in Stability and Amyloidogenicity of Antibody Variable Domains, *J. Biol. Chem.* (2014), doi: 10.1074/jbc.M114.582247

Cardine N Nokwe, <u>Manuel Hora</u>, Martin Zacharias, Hisashi Yagi, Christine John, Bernd Reif, Yuji Goto, Johannes Buchner:
The Antibody Light-Chain Linker Is Important for Domain Stability and Amyloid Formation, *J. Mol. Biol.* (2015), doi:10.1016/j.jmb.2015.09.012

Cardine N Nokwe, <u>Manuel Hora</u>, Martin Zacharias, Hisashi Yagi, Jirka Peschek, Bernd Reif, Yuji Goto, Johannes Buchner:
A Stable Mutant Predisposes Antibody Domains to Amyloid Formation through Specific Non-Native Interactions, *J. Mol. Biol.* (2016), doi:10.1016/j.jmb.2016.01.015

The following oral presentation was given:

Solid- and Solution-State NMR Studies on Antibody Light Chain Amyloid Formation and Interactions with Epigallocatechin Gallate, *Computational Aspects - Biomolecular NMR Gordon Research Conference, Lucca (Barga), Italy, 10th of June, 2015*

The following manuscripts are in preparation:

<u>Manuel Hora</u>, Martin Carballo Pacheco, Benedikt Weber, Antje Wittkopf, Johannes Buchner, Birgit Strodel, Bernd Reif:
Epigallocatechin-3-gallate Interacts via Two Distinct Binding Modes with Amyloidogenic Immunoglobulin Light Chains, *in preparation*

Andreas Schmidt, <u>Manuel Hora</u>, William Close, Karthikeyan Navalpur Annamalai, Matthias Schmidt, Bernd Reif, Nikolaus Grigorieff, Marcus Fändrich:
Cryo-EM Reveals the Polymorphism of Light Chain-Derived Peptide Fibrils, *in preparation*

Declaration

I declare that I wrote this thesis independently. I was using only the references and resources stated here. The work has so far not been submitted to any audit commission. Parts of this work have been published in scientific journals.

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die Arbeit wurde bisher keiner Prüfungskommission vorgelegt. Ein Teil dieser Arbeit wurde bereits in wissenschaftlichen Journalen veröffentlicht.

. . . .