



Technische Universität München

TUM School of Natural Sciences

Molecular mechanisms of antibody light chain amyloidogenicity

Georg Josef Rottenaicher

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzende: Prof. Dr. Cathleen Zeymer

Prüfer der Dissertation: 1. Prof. Dr. Johannes Buchner
2. Prof. Dr. Bernd Reif

Die Dissertation wurde am 02.09.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Natural Sciences am 29.11.2022 angenommen.

Für meine Eltern

Content

ABBREVIATIONS

SUMMARY	1
ZUSAMMENFASSUNG	3
1. INTRODUCTION	5
1.1 Protein folding	5
1.1.1 General aspects of protein folding	5
1.1.2 Protein folding in the cell	8
1.2 Protein misfolding and associated diseases	13
1.2.1 Misfolding and aggregation	13
1.2.2 The amyloid state of proteins	15
1.2.3 Amyloidoses	17
1.3 Immunoglobulins	19
1.3.1 Structure of antibodies and the immunoglobulin fold	19
1.3.2 Emergence of sequence variability	22
1.4 Systemic light chain (AL) amyloidosis	23
1.4.1 Disease characteristics	23
1.4.2 Diagnosis and treatment	25
1.5 Objectives of the thesis	27
2. MATERIAL AND METHODS	28
2.1 Chemicals	28
2.2 Consumables	30
2.3 Buffers and solutions	31
2.3.1 Buffers for protein purification	31
2.3.2 Buffers for SDS-PAGE	32
2.3.3 Buffers and solutions for molecular biology	32
2.3.4 Buffers for protein biochemistry	33
2.4 Enzymes, Standards, and Kits	34
2.5 Bacterial strains	34
2.6 Plasmids and constructs	35
2.7 Oligonucleotides	36
2.8 Devices and instruments	37
2.9 Software, databases, and web-based tools	40

2.10 Molecular biology methods	41
2.10.1 Amplification and modification of DNA (PCR and site-directed mutagenesis)	41
2.10.2 Agarose gel electrophoresis	43
2.10.3 Preparation of chemically competent bacterial cells	43
2.10.4 Transformation of bacteria	43
2.10.5 Isolation of DNA	43
2.11 Protein expression and purification	44
2.11.1 Protein expression as insoluble inclusion bodies	44
2.11.2 Purification by refolding from inclusion bodies	44
2.11.3 Native protein expression	45
2.11.4 Native protein purification	46
2.12.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)	47
2.12.2 Liquid chromatography and mass spectrometry (LC-MS)	47
2.13 Spectroscopy	48
2.13.1 UV/VIS spectroscopy	48
2.13.2 Circular dichroism spectroscopy (CD)	49
2.13.3 Fluorescence spectroscopy	50
2.13.3.1 Intrinsic tryptophan fluorescence	50
2.13.3.2 Chemical unfolding transitions	51
2.13.3.3 Acrylamide quenching	52
2.13.3.4 Red edge excitation shift (REES)	52
2.13.3.5 Anilinonaphtalene sulfonate (ANS) binding	53
2.13.3.6 Thioflavin T (ThT) binding kinetics	53
2.14 Analytical ultracentrifugation	53
2.15 Limited Proteolysis	54
2.16 Differential Scanning Calorimetry (DSC)	55
2.17 Transmission electron microscopy (TEM)	55
2.18 Hydrogen/deuterium exchange coupled to mass spectrometry (HDX-MS)	56
2.19 Molecular dynamics (MD) simulations	56
3. RESULTS AND DISCUSSION	58
3.1 Amyloid formation of the FOR005 variable domain (V_L) is mediated by patient-specific mutations in hypervariable CDR loops	58
3.1.1 Sequence differences between FOR005-PT and FOR005-GL	58
3.1.2 FOR005-PT and FOR005-GL differ in their biophysical properties and amyloid propensity	60
3.1.3 The effect of point mutations on stability and aggregation of the V _L domain	65
3.1.4 Conformational dynamics are linked to decreased protein stability and amyloid formation	69
3.1.5 Energetically unfavorable conformations of the CDR loops negatively impact the stability of the V _L	73
3.1.6 Discussion	76
3.2 Disease onset in the case of FOR005 is linked to a non-conservative mutation in the constant light chain (C_L) domain	80

3.2.1 The sequence of FOR005 contains a rare C _L mutation	80
3.2.2 Structure and stability of the C _L are altered by the mutation V136G	81
3.2.3 The mutation V136G causes partial unfolding of the C _L	85
3.2.4 The full-length LC and the C _L domain do not form amyloid fibrils <i>in vitro</i>	90
3.2.5 Proteolysis of LCs is mediated by an unstable C _L domain	91
3.2.6 The C _L mutation impairs homo-dimerization	95
3.2.7 Discussion	98
4. CONCLUSIONS AND PERSPECTIVES	102
5. REFERENCES	105
6. PUBLICATIONS AND PRESENTATIONS	126
7. ACKNOWLEDGEMENTS	127
8. DECLARATION	128

Abbreviations

°C	Degree Celsius
Å	Angstrom
ADP	Adenosine diphosphate
AL	Amyloid light chain
ATP	Adenosine triphosphate
BiP	Immunoglobulin Binding Protein
CD	Circular dichroism
cDNA	Complementary desoxyribonucleic acid
CDR	Complementarity determining region
CDRH3	CDR3 of the heavy chain
C _H	Heavy chain constant domain
C _L	Light chain constant domain
C _m	Denaturant concentration of 50 % unfolding
cm	Centimeter
CNS	Central nervous system
COC	Cycloolefin-copolymer
<i>E. coli</i>	Escherichia coli
EM	Electron microscopy
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ERQC	ER quality control
ESI	Electrospray ionization
Erdj	ER-resident J-domain containing protein
F _{ab}	Antigen-binding fragment
F _c	Crystallizable fragment
FLC	Free light chain
FR	Framework regions
FUV	Far ultraviolet
fwd	Forward
GL	Germline
HC	Heavy chain
HDX	Hydrogen deuterium exchange
HMR	Hydrogen mass repartitioning
HPLC	High pressure liquid chromatography
Hsp	Heat shock protein

IMGT	International Immunogenetics information system
Ig	Immunoglobulin
IPTG	Isopropyl thiogalactopyranoside
kDa	Kilodalton
kJ	Kilojoule
L	Liter
LB _{Kan}	Luria-Bertani medium with kanamycin
LC	Light chain
M	Mol per liter
MD	Molecular dynamics
mg	Milligram
mL	Milliliter
μl	Microliter
MM	Multiple myeloma
mM	Millimol per liter
μM	Micromol per liter
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MWCO	Molecular weight cut-off
m/z	Mass over charge
nm	Nanometer
NMR	Nuclear magnetic resonance
NUV	Near ultraviolet
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data base
PE	Polyethylene
PP	Polypropylene
PS	Polystyrol
PT	Patient
rev	Reverse
RMSD	Root mean square deviations
RMSF	Root mean square fluctuations
rpm	Rotations per minute
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
t ₅₀	Time of 50 % completed fibril formation

ThT	Thioflavin T
T _m	Melting temperature
TOF	Time of flight
UPLC	Ultra-high pressure liquid chromatography
US	Umbrella sampling
UV/VIS	Ultraviolet/Visible
V _H	Heavy chain variable domain
V _L	Light chain variable domain
v/v	volume per volume
w/v	weight per volume
WHAM	Weighted histogram analysis method

Amino acids (three letter and single letter code)

Amino acid	Three letter code	Single letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valin	Val	V

Summary

Light chain (AL) amyloidosis is the most common form of systemic amyloidosis and represents a debilitating disease characterized by the deposition of insoluble amyloid fibrils in the extracellular space of various organs. Most commonly, the heart and kidneys are affected although also liver, gastro-intestinal tract, and nervous system can be involved. The underlying cause is most often a plasma cell dyscrasia (e.g., multiple myeloma) leading to excessive secretion of free antibody light chains (FLCs) by aberrant plasma cell clones. In a fraction of patients, these secreted FLCs contain destabilizing mutations causing them to misfold and aggregate into highly ordered fibrils that irreversibly damage patient organs upon deposition.

In the scope of this thesis, the underlying molecular mechanisms that govern the misfolding and amyloid formation of antibody LCs were investigated. To this end, the patient case FOR005 was examined. The cDNA sequence of the FOR005 LC was obtained at the University Clinic Heidelberg from a patient with severe cardiac amyloidosis and in a following study by Annamalai et al., it was shown that the amyloid deposits from patient tissue only contain the LC variable domain (V_L). Therefore, the first part of the FOR005 project focuses only on the V_L domain, which contains five mutations compared to the closest corresponding germline V_L sequence. The specific effects of these point mutations on the features of the V_L domain were examined by a variety of biochemical and biophysical approaches and it was found that two mutations in particular are responsible for the destabilization and fibrillar aggregation of the FOR005 V_L . These two mutations – namely G49R and G94A – are located in the hypervariable CDR2 and CDR3 loop, respectively, and induce energetically unfavorable backbone conformations in these loops. As a result, conformational dynamics in conserved, surrounding framework regions are enhanced leading to a strong destabilization of the V_L domain, which coincides with its amyloid propensity.

In the second part of the project, the full-length LC of FOR005 was investigated. The patient sequence does not only contain five mutations in the V_L domain but also a non-conservative substitution in the constant domain (C_L). The effects of this disease-related C_L mutation were characterized in detail. Mutations in the C_L domain are very rare and, therefore, their influence in the context of AL amyloidosis is poorly understood. The results show that the V136G substitution has a strong impact on the properties of the C_L and the LC, respectively, as it leads to partial unfolding of the C_L and a significant decrease in thermodynamic stability both for the isolated C_L and the entire LC. Furthermore, LC homo-dimerization is impaired by both the C_L and V_L mutations since the constituent LC dimerization interface is disrupted. The resulting LC monomers are far more susceptible to proteolytic cleavage than dimeric LCs and are, therefore, the species presumably responsible for release of the amyloidogenic V_L domain.

The enhanced proteolytic susceptibility of LC monomers is caused by a disruption of native V_L-C_L domain interactions due to the unfavorable point mutations. Thus, the findings enable the definition of a molecular mechanism to describe the disease onset in the case FOR005. Conclusively, the results obtained during this thesis reveal novel insights into the molecular underpinnings of AL amyloidosis and can, thereby, contribute to the development of new therapeutic options and diagnostic tools in the future.

Zusammenfassung

Leichtkettenamyloidose (AL) ist die häufigste systemische Form der Amyloiderkrankungen und eine schwere Krankheit, bei der sich unlösliche Amyloidfibrillen im extrazellulären Raum verschiedener Organe ablagern. Meistens sind Herz und Nieren betroffen, jedoch können auch Leber, Gastrointestinaltrakt oder das Nervensystem involviert sein. Zu Grunde liegt häufig eine Plasmazell-Dyskrasie (z.B. Multiples Myelom), was eine exzessive Sekretion von freien Antikörper-Leichtketten (FLC) durch entartete Plasmazellen zur Folge hat. In einem Teil der Patienten enthalten diese sekretierten Leichtketten (LCs) destabilisierende Mutationen, was zur Fehlfaltung und Aggregation in Form geordneter Amyloidfibrillen führt. Diese Fibrillen verursachen irreversible Organschäden.

Im Rahmen dieser Dissertation wurden die molekularen Mechanismen untersucht, welche die Fehlfaltung und Amyloidbildung der Leichtketten steuern. Dazu wurde der Patientenfall FOR005 analysiert. Die cDNA-Sequenz der FOR005 LC einer Patientin mit schwerer, kardialer Amyloidose wurde am Universitätsklinikum Heidelberg identifiziert und die darauf aufbauende Studie von Annamalai et al. zeigte, dass die Ablagerungen im Patientengewebe nur aus der V_L -Domäne bestehen. Daher lag der Fokus im ersten Teil des FOR005-Projekts nur auf der V_L -Domäne, welche fünf Mutationen im Vergleich zur entsprechenden Keimbahnsequenz enthält. Die Effekte dieser Punktmutationen auf die Eigenschaften der V_L wurden mit verschiedenen biochemischen und biophysikalischen Methoden ermittelt, mit dem Ergebnis, dass insbesondere zwei spezifische Mutationen für die Destabilisierung und Aggregation der V_L -Domäne verantwortlich sind. Diese beiden Mutationen – genauer G49R und G94A – befinden sich in den hypervariablen CDR2- bzw. CDR3-Schleifen der Domäne und verursachen energetisch ungünstige Konformationen in diesen Bereichen. In der Folge weisen die umliegenden, konservierten Framework-Regionen eine erhöhte strukturelle Dynamik auf, was zu einer starken Destabilisierung und einhergehender Anfälligkeit für Fibrillenbildung führt.

Im zweiten Teil des Projekts wurde die komplette LC von FOR005 untersucht. Die Patientensequenz enthält nicht nur V_L -Mutationen, sondern auch eine nicht-konservative Substitution in der konstanten Domäne (C_L). Die Effekte dieser krankheitsassoziierten C_L Mutation wurden im Detail charakterisiert. Mutationen der C_L sind sehr selten, weshalb über deren Einfluss im Kontext der AL Amyloidose wenig bekannt ist. Wie die Ergebnisse zeigen hat die V136G-Mutation einen starken Effekt sowohl auf die Eigenschaften der C_L -Domäne als auch der kompletten LC. Die Mutation bedingt die partielle Entfaltung der C_L , wodurch die thermodynamische Stabilität sowohl der isolierten C_L als auch der LC signifikant erniedrigt wird. Des Weiteren wird durch die V_L - und C_L -Mutationen die LC-Homodimerisierung beeinträchtigt, da diese das Dimerisierungs-Interface zerstören. Die resultierenden LC-

Monomere sind weitaus anfälliger für Proteolyse als LC-Dimere, weshalb die Monomere die Spezies sind, die mutmaßlich für die Freisetzung der amyloidogenen V_L-Domäne verantwortlich sind. Die erhöhte proteolytische Anfälligkeit der Monomere wird durch die mutationsbedingte Störung der nativen V_L-C_L-Domäneninteraktionen verursacht. Somit ermöglichen diese Erkenntnisse die Definition eines molekularen Mechanismus der Krankheitsentstehung im Fall FOR005. Zusammenfassend eröffnen die Ergebnisse dieser Arbeit neue Einsichten in die molekularen Grundlagen der systemischen Leichtkettenamyloidose und können dadurch zur zukünftigen Entwicklung neuer Therapieoptionen und diagnostischer Methoden beitragen.

1. Introduction

1.1 Protein folding

1.1.1 General aspects of protein folding

Proteins are among the most central molecular building blocks of life and fulfill important functions in virtually every aspect of biology. They partake in a large variety of biochemical processes such as molecular transport, signal transduction, cell development, enzyme catalysis, cell-cell communication, immune responses, cellular energetics, and they are also crucial constituents of the cytoskeleton. In general, proteins are linear biopolymers consisting of 20 different proteinogenic amino acids, which are covalently linked by peptide bonds, hence the name polypeptides. The specific amino acid sequence of a protein is called primary structure and it is encoded by the corresponding DNA segment, or gene, which is transcribed into mRNA that is further recognized and translated into the amino acid sequence by the ribosome (Voet et al., 2016). To function properly, proteins need to adopt a complex three-dimensional structure, which is subdivided into four layers. The primary structure, as mentioned above, represents the amino acid sequence of the protein. Secondary structure refers to local structural elements which are built up through short-range hydrogen bonds between the peptide bonds of amino acids that are in close proximity to each other in the sequence (Chothia, 1984; Creighton, 1993). The different amino acids have varying tendencies for the formation of certain secondary structure elements such as α -helices, β -strands, and turns (Malkov et al., 2008). Through combination of secondary structure elements recurring structural motifs can be created which are often associated with a specific function, as for instance helix-turn-helix motifs in DNA-binding proteins (Harrison and Aggarwal, 1990). The tertiary structure of a protein is determined by the relative orientation and long-range interactions between segments with defined secondary structure. Primary, secondary, and tertiary structure define the three-dimensional conformation of a single polypeptide chain (Chothia, 1984; Creighton, 1993). However, multiple polypeptides can be associated with each other and the arrangement of these multiple chains is described as quaternary structure (Klotz et al., 1970; Marsh and Teichmann, 2015). Protein structure is primarily defined by the non-covalent interactions of the peptide backbone and amino acid side chains which engage in hydrogen bonding, electrostatic interactions between charged residues, and van-der-Waal's or hydrophobic interactions (Chothia, 1984; Lazaridis et al., 1995; Pace, 1995; Dobson et al., 1998; Dobson, 2003; Dill et al., 2008). Additionally, the covalent connection of side chains is also possible, for instance in the form of disulfide bonds between cysteine residues which can increase a protein's stability (Go and Jones, 2008). Disulfide bridges are a type of post-translational modification which play an important role in protein functionality since they are

able to modulate structure and activity (Kikuchi et al., 1986; Bulaj, 2005). Among the most common post-translational modifications are phosphorylation and glycosylation of side chains (Walsh et al., 2005). The main determinant of protein functionality, however, is its unique three-dimensional structure, the so-called native state. It represents the thermodynamically most stable conformation of the polypeptide chain with the lowest degree of Gibbs free energy (Dobson, 2003). In this native state, hydrophobic residues are typically buried inside the protein core to reduce energetically unfavorable interactions with the surrounding polar solvent, i.e., water molecules, whereas polar, hydrophilic side chains are arranged on the outside of the protein (Kauzmann, 1959; Chen and Stites, 2001; Dill et al., 2008). The transition of a linear polypeptide chain, as it emerges from the ribosome, into a complex three-dimensional structure is described as protein folding (Dobson, 2003).

The mechanisms and principles of protein folding have puzzled scientists for decades and are still subject to ongoing research today. In the early days of protein folding research, Levinthal's paradox was formulated: if a linear polypeptide would fold into its native structure by a random search mechanism, i.e., by sampling every possible conformation until the minimum in Gibbs free energy is reached, it would take an astronomical amount of time to reach the native state. This is in vast contrast to the actual time frame in which proteins normally fold into their native structure which ranges from milliseconds to minutes. Therefore, protein folding cannot occur randomly but must follow a predefined pathway instead (Levinthal, 1968). This pathway is determined by the amino acid sequence of the protein, as stated by Anfinsen (Anfinsen, 1973). Mechanistically, folding pathways can be perceived by different models. In the hydrophobic collapse model, the hydrophobic side chains build up the protein core in a first, rapid step due to repulsive interactions with the surrounding polar water molecules (hydrophobic effect). In a slower second step, the conformations of the residual, more hydrophilic segments on the protein surface are established. (Kauzmann, 1959; Dill, 1985; Gutin et al., 1995). The nucleation-condensation model also describes a two-phased mechanism: in the first step a folding nucleus consisting of a small number of secondary structure elements is built up by a few key residues. Once this nucleus is formed, the residual structure condenses around it in a faster, second step (Wetlaufer, 1973; Udgaonkar and Baldwin, 1988; Fersht, 1997). Further, the diffusion-collision model involves the formation of so-called microdomains, i.e., small segments with defined structure. These microdomains randomly fluctuate until they collide and engage in native interactions that establish the correct tertiary structure (Karplus and Weaver, 1979, 1994).

With ongoing progress in the field, our view of protein folding was refined by the emergence of the folding funnel model (Dill and Chan, 1997). Rather than following a specific pathway, the unfolded polypeptide chain performs a stochastic search along a free energy landscape that

is strongly biased towards the native state, i.e., the global minimum of the free energy surface, hence the funnel description (Onuchic et al., 1997). A schematic representation of the folding funnel is depicted in Figure 1 (Hartl and Hayer-Hartl, 2009). The coordinates spanning up the landscape are defined as Gibbs free energy and entropy with each point of the surface representing a specific conformational state of the protein (Dobson et al., 1998). Thus, as the protein moves along the landscape, down towards the global minimum, its thermodynamic stability increases due to the decline in free energy (Dobson, 2003). However, the native-like interactions that are established within the protein along the way, stabilize the structure at the cost of entropic loss which means that protein folding can be viewed as a trade-off between a gain in thermodynamic stability and loss in conformational entropy (Dobson et al., 1998). Yet, the formation of the hydrophobic protein core involves the displacement of water molecules, which in turn increases the entropy of the surrounding solvent thus partially compensating the negative entropy change of the protein folding reaction (Makhatadze and Privalov, 1996; Baldwin and Rose, 2013).

The free energy landscape is not a smooth surface, but rather rugged with multiple local minima that represent folding intermediates (Dill and Chan, 1997; Onuchic et al., 1997; Dobson, 2003). These intermediates can be seen as transition states of a chemical reaction and they can already resemble the native conformation to some extent. The molten globule state, for instance, is characterized by the presence of unstable segments with compact, native-like secondary structure while the tertiary structure is not yet established. Thus, the molten globule refers to a structurally heterogeneous ensemble of folding intermediates (Ptitsyn et al., 1990; Onuchic et al., 1997; Dobson et al., 1998). Furthermore, proteins can adopt alternatively folded states depending on the conditions, e.g., low pH value. This means the protein assumes a folded, stable conformation with well-defined tertiary contacts, yet it differs from the native structure that is adopted under physiological conditions (Buchner et al., 1991). In addition, the native state itself does not necessarily correspond to a single conformation at the lowest point of the folding funnel. The global minimum can also be a rugged region at the bottom of the folding funnel meaning that multiple native structures exist in parallel, which are equally stable but slightly different regarding their structure. Thereby, proteins can undergo conformational rearrangements in the native state and are equipped with a certain degree of structural flexibility which is often a requirement for fulfilling their biological function (Karplus and Kuriyan, 2005; Shehu et al., 2007).

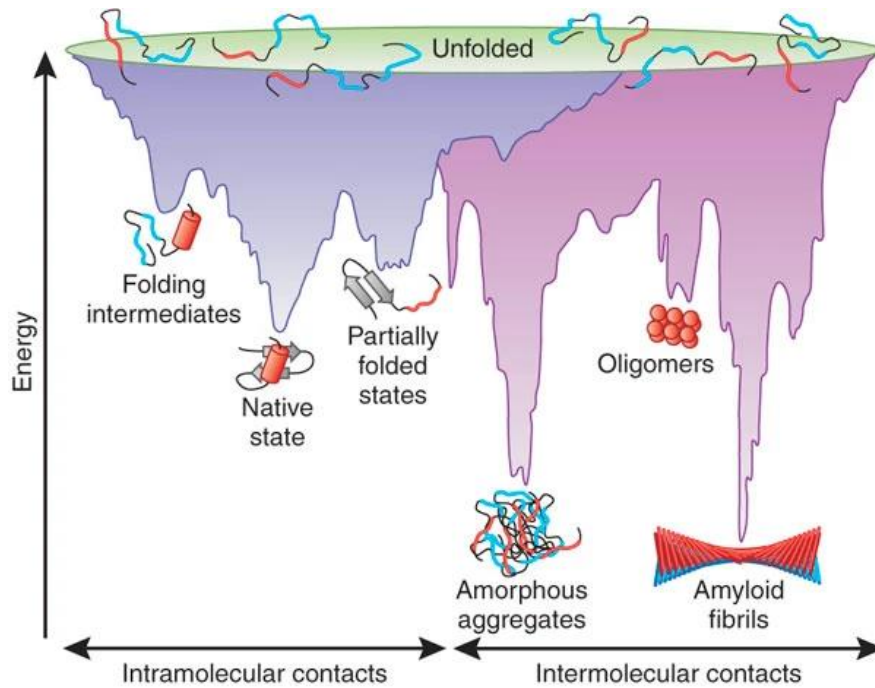


Figure 1: Schematic representation of the energy landscape of protein folding (Folding funnel). In *de novo* folding, the protein is funneled towards the native state along the surface by establishing native intramolecular contacts. Through intermolecular contact formation, aggregation from partially folded intermediates can occur. The folding landscape is a function of the Gibbs energy and the conformational entropy. Figure adapted with permission from Hartl and Hayer-Hartl, 2009.

1.1.2 Protein folding in the cell

In general, the folding of a protein is determined by its amino acid sequence which shapes the intramolecular interactions and the free energy landscape (Anfinsen, 1973). In living cells, however, there are numerous other factors that also affect this process such as molecular crowding (Ellis and Minton, 2006). Therefore, cells possess a complex protein homeostasis network to ensure correct folding and protein functionality in which molecular chaperones play a crucial role (Hartl and Hayer-Hartl, 2009). Molecular chaperones are proteins that assist in the folding of other proteins (so-called clients or substrates) either by binding to the unfolded protein and assuring transition into the native state, or by preventing the misfolding and aggregation of (partially) unfolded proteins (Balchin et al., 2020). Due to their vital function in cellular protein homeostasis, chaperones are often up-regulated during cell stress, for instance during heat stress, which is why many of them are called heat shock proteins (Hsp) (Richter et al., 2010). There are several different classes and networks of molecular chaperones that are often connected and that cooperate in client protein folding (Balchin et al., 2016).

The Hsp90 system is one of the most important chaperone networks involved in the folding of multiple client proteins that fulfill important cellular functions including kinases, E3 ligases, transcription factors, and hormone receptors (Pratt and Dittmar, 1998; Taipale et al., 2012). The chaperone Hsp90 is a 90 kDa protein consisting of three domains: an N-terminal domain

responsible for nucleotide binding, a middle domain, and a C-terminal domain that facilitates homo-dimerization. This chaperone system and its function are conserved from bacteria to all eukaryotes in which several isoforms exist (Chen et al., 2006). On one hand, there are different isoforms of Hsp90 in the respective cellular compartments, on the other hand, there is a constitutively expressed form (Hsp90 β) and an induced form (Hsp90 α) that is expressed in the cytosol during cell stress (Biebl and Buchner, 2019). Hsp90 is a nucleotide-dependent chaperone that undergoes an ATPase cycle with different conformational states when it assists folding of a client protein. In the so-called open state, the homo-dimer exhibits a V-shaped conformation and both ATP and unfolded client proteins can be bound. Through several subsequent structural rearrangements, Hsp90 adopts a closed conformation in which the two monomers are in close contact to each other. Upon client folding and ATP hydrolysis, the complex disassembles, thus releasing the properly folded client protein, ADP, and phosphate. Hsp90 returns into its open state and another client can be bound (Hessling et al., 2009; Balchin et al., 2016). The structure and functional cycle of Hsp90 are shown in Figure 2 (Balchin et al., 2016). Further, the activity and client interactions of Hsp90 are often modulated by co-chaperones that bind in a specific manner. While bacteria possess no Hsp90 co-chaperones, eukaryotes have developed a multitude of interaction partners that affect Hsp90-dependent client folding, such as Hop, Aha1, Sgt1, p23, Cdc37, CHIP, etc. (Schopf et al., 2017). These co-chaperones can inhibit or accelerate distinct steps of the ATPase cycle, bind to specific conformational states of Hsp90, or act in a client-specific manner, e.g., Cdc37 specifically modulates kinase maturation (Roe et al., 2004). Arguably, Hsp90 participates in a multitude of cellular processes and is a key player in protein homeostasis. Since a high protein folding load is also a known trait of cancer cells, the possibility of addressing Hsp90 as a therapeutic target has been subject to intensive research. Furthermore, Hsp90 is also associated with various neurodegenerative diseases due to its vital role in the folding and regulation of client proteins. However, its exact function in this context is unclear which so far prevents targeting Hsp90 in these conditions (Whitesell and Lindquist, 2005; Biebl and Buchner, 2019).

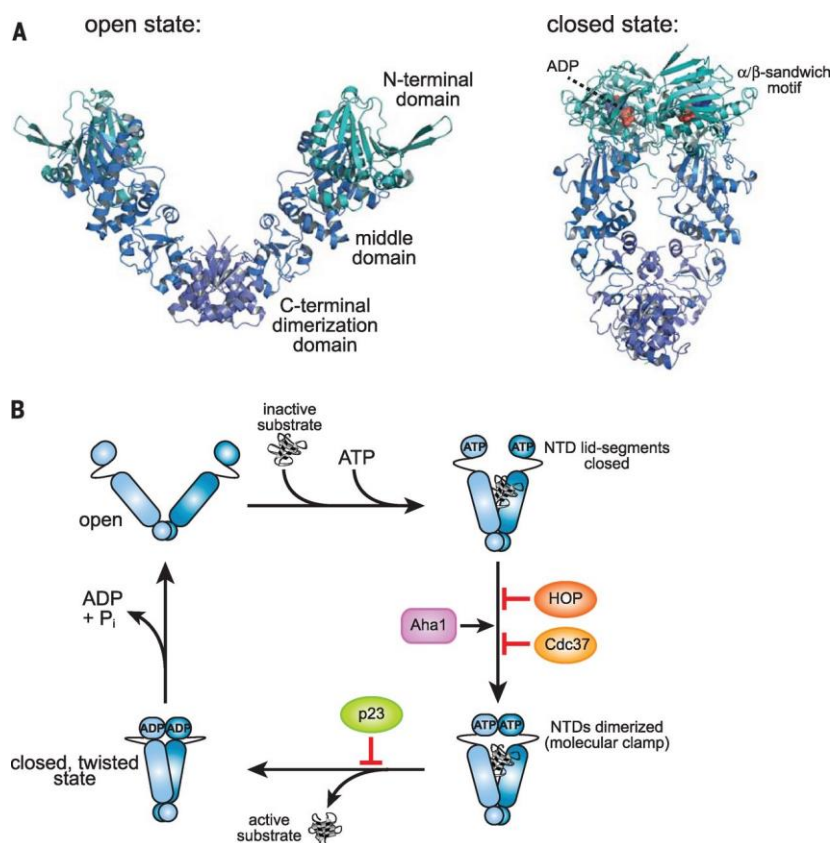


Figure 2: Structure and conformational cycle of Hsp90. A) Hsp90 is a homo-dimer with each protomer consisting of an N-terminal nucleotide binding domain, a middle domain, and a C-terminal domain responsible for dimerization. B) Hsp90 undergoes a series of structural rearrangements during its ATPase cycle. Co-chaperones can bind during different steps of the cycle and thereby modulate Hsp90's activity. Figure adapted from Balchin et al., 2016. Reprinted with permission from AAAS.

Another crucial chaperone network involved in cellular protein homeostasis is the Hsp40-Hsp70 system which is also conserved from prokaryotes to eukaryotes (Balchin et al., 2016). Structurally, Hsp70 consists of a substrate-binding domain (SBD), a nucleotide-binding domain (NBD), and a flexible lid. Similar to Hsp90, Hsp70 binds to various unfolded or partially folded proteins and aids in obtaining their native three-dimensional structure (Rosenzweig et al., 2019). Further, Hsp70 also has numerous binding partners that regulate its function and activity, for instance its nucleotide exchange factors (NEF) Hsp110/Grp170, Sil1, and Bag domain proteins (Bracher and Verghese, 2015). One of the most important interaction partners of Hsp70, however, is Hsp40. This molecular chaperone belongs to the family of J-domain containing proteins (JDPs) and is a key factor of protein homeostasis (Rosenzweig et al., 2019). Hsp40, of which several classes and isoforms exist in eukaryotes, recognizes and binds unfolded or misfolded substrate proteins and transfers them to Hsp70 which is nucleotide-dependent and passes through a conformational cycle while performing its physiological tasks as shown in Figure 3 (Kampinga and Craig, 2010; Kampinga et al., 2019; Rosenzweig et al., 2019). When ATP is bound, Hsp70 adopts an open, low-affinity-state. A non-native substrate protein is transferred to Hsp70 by Hsp40 causing ATP-hydrolysis and rearrangement into the

ADP-bound high-affinity-state. Upon release of ADP from the complex, a nucleotide exchange factor (NEF) binds to Hsp70 and facilitates rebinding of ATP and release of the natively folded client (Rosenzweig et al., 2019). Further, Hsp70 is not only able to participate in the maturation of nascent proteins, but also to recognize misfolded proteins and protein aggregates. Misfolded proteins and aggregates are bound, unfolded, and are then either refolded or directed towards cellular protein degradation (Balchin et al., 2016; Mogk et al., 2018). Therefore, Hsp40 and Hsp70 together with Hsp110 can act as a disaggregase, i.e., a protein complex that catalyzes disassembly of potentially harmful protein aggregates (Gao et al., 2015; Mogk et al., 2018). The Hsp70 system and the Hsp90 system are not completely isolated networks, but rather they can cooperate in cellular protein folding. Therefore, certain bridging factors exist, i.e., proteins that connect the two systems and hand over clients from one chaperone network to the other (Genest et al., 2019). The Hsp70-Hsp90 organizing protein (Hop) is the best-known bridging component so far, facilitating client transfer from the Hsp70 system to Hsp90 (Johnson et al., 1998; Schmid et al., 2012). However, additional proteins with similar function (e.g., NudC) have been identified recently (Biebl et al., 2022).

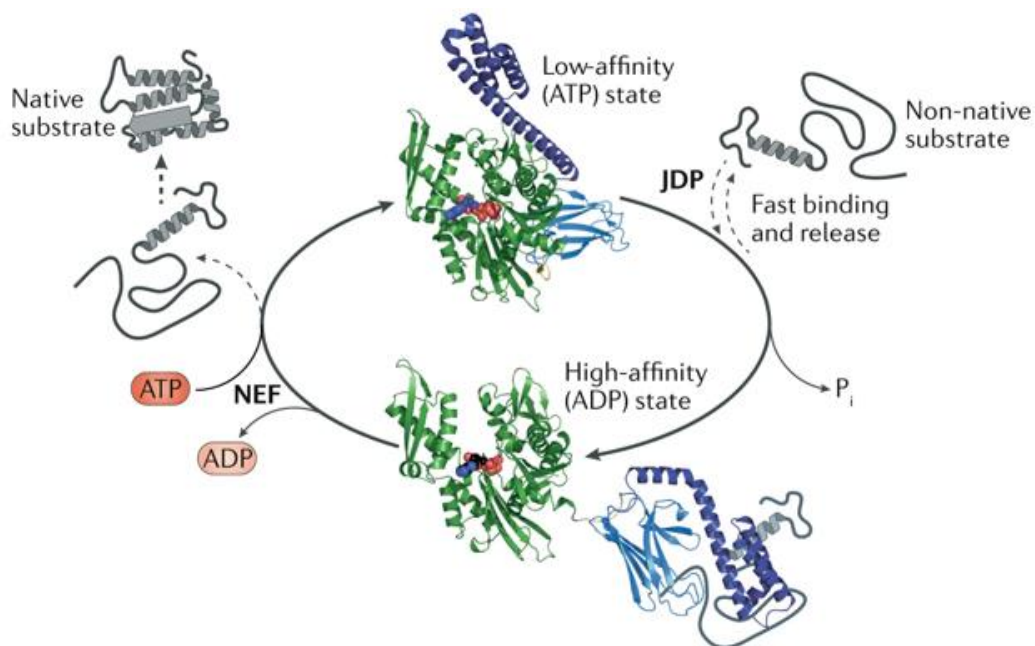


Figure 3: The Hsp70 conformational cycle. The nucleotide binding domain (NBD) is shown in green, the substrate binding domain (SBD) in light blue, and the lid domain in dark blue. In the ATP-bound state, Hsp70 receives unfolded substrate proteins from Hsp40 (JDP). Upon ATP-hydrolysis and release of ortho-phosphate (P_i), structural rearrangements take place. Substrate release and nucleotide exchange are facilitated by a nucleotide exchange factor (NEF). Figure adapted and modified with permission from Rosenzweig et al., 2019.

Small heat shock proteins (sHsp) constitute another class of chaperones with relatively small molecular weights that are conserved in all organisms. All sHsps are ATP-independent and share common structural characteristics, i.e., a disordered N-terminal region (NTR), a

conserved α -crystallin domain (ACD), and a C-terminal domain harboring a conserved IXI/V sequence motif (Haslbeck et al., 2019). Primarily, sHsps suppress the aggregation of substrate proteins and function as so-called holdases, i.e., they bind to misfolded substrate proteins and hold them in an unfolded state while inhibiting aggregate formation (Richter et al., 2010). Further, this ubiquitous class of chaperones preferentially forms dynamic, higher order homo-oligomers or hetero-oligomers of varying size which is also a mode of functional regulation (Haslbeck and Vierling, 2015). Often smaller oligomers show higher chaperone activity than large oligomer assemblies and the oligomerization can be regulated by post-translational modifications, e.g., phosphorylation, or by variations in pH or temperature, respectively (Fleckenstein et al., 2015; Mühlhofer et al., 2021). The sHsp network is also connected to other chaperone systems such as the Hsp70 system via bridging factors (e.g., Bag3) although the interplay of sHsps with other chaperone systems is only poorly understood so far (Rauch et al., 2017).

Like a plethora of processes, the division into cellular compartments regulates protein folding in eukaryotic cells. Cellular proteins are synthesized and folded in the cytoplasm while translation, folding, and modification of secretory proteins occurs in the endoplasmic reticulum (ER) (Aviram and Schuldiner, 2017). This intricate membrane network contains an oxidizing environment that allows the formation of disulfide bridges in proteins as opposed to the reducing environment of the cytoplasm in which free cysteine residues are common (Bulaj, 2005). Nascent secretory proteins that are translated on cytoplasmic ribosomes are bound by the signal recognition particle (SRP), transported to the ER membrane, and translocated into the ER lumen via the Sec61 translocon machinery. ER-resident chaperones such as BiP (Hsp70 in the ER) and ER DnaJ proteins (ERdj; Hsp40 family of the ER) among others facilitate the folding of secretory substrate proteins upon their translocation (Fenech et al., 2020). Subsequently, proteins can be modified by ER-resident enzymes prior to secretion, e.g., by glycosyl transferases or by protein disulfide isomerases (PDI) which are able to rearrange disulfide bridges (Kozlov et al., 2010; Cherepanova et al., 2016).

Furthermore, the ER contains a complex protein quality control mechanism (ERQC) which ensures that only properly folded proteins are being secreted into the extracellular milieu (Wiseman et al., 2022). The proteins calnexin and calreticulin act as sensors that detect misfolded substrate proteins and direct them to a pathway of glycosylation and de-glycosylation events until it is probed by the two folding sensors again (Braakman and Bulleid, 2011). A substrate protein can undergo this cycle a number of times until it either adapts its correct fold and is being secreted or it is directed towards the ER-associated degradation pathway (ERAD) (Vembar and Brodsky, 2008). Thus, the ERQC system is responsible for assuring that only correctly folded and functional proteins leave the cell.

1.2 Protein misfolding and associated diseases

1.2.1 Misfolding and aggregation

As mentioned above, protein folding can be viewed as a sort of trade-off between the gain in thermodynamic stability, i.e., a decrease in free energy due to native interactions, and an entropic loss due to the decreasing conformational entropy (Dobson et al., 1998). Therefore, the native tertiary structure of proteins is often only marginally more stable compared to its denatured state (Dill et al., 2008). In this context, protein folding can be viewed as an equilibrium chemical reaction in which the equilibrium is shifted towards the native state of the protein under physiological conditions



where U is the unfolded state and N is the native state. The thermodynamic description is given by

$$\Delta G_f = \Delta H_f - T\Delta S_f \quad (\text{Equation. 2})$$

where ΔG_f is the Gibbs free energy of the folding reaction, ΔH_f is the change in enthalpy, ΔS_f is the change in entropy, and T is the absolute temperature. The change in enthalpy ΔH is negative since the stabilized native state of the protein is lower in enthalpy than the unfolded state. The entropic change ΔS is also negative due to the emergence of the three-dimensional protein structure that exhibits a far higher degree of order than the unfolded polypeptide chain. Therefore, if the temperature T is sufficiently high, the entropic term compensates the enthalpic term so that the Gibbs free energy ΔG becomes positive. Thus, the reaction is endergonic and the folding equilibrium is shifted towards the unfolded state. Other than elevated temperature, extrinsic factors such as pH value, ionic strength, and the presence of denaturing agents (e.g., urea, guanidinium chloride) affect the equilibrium and the stability of the folded state (Santoro and Bolen, 1988). Intrinsically, mutations in the amino acid sequence influence the folding equilibrium and stability of the native state. The primary sequence determines the interactions on which the native fold depends and, therefore, substitutions in the amino acid sequence can disrupt crucial interactions. Thus, the stability of the native state is decreased and the folding equilibrium can be shifted towards the unfolded state (Dobson et al., 1998; Dobson, 2003). Figure 4 represents a simplified schematic overview of the thermodynamics of protein unfolding and misfolding.

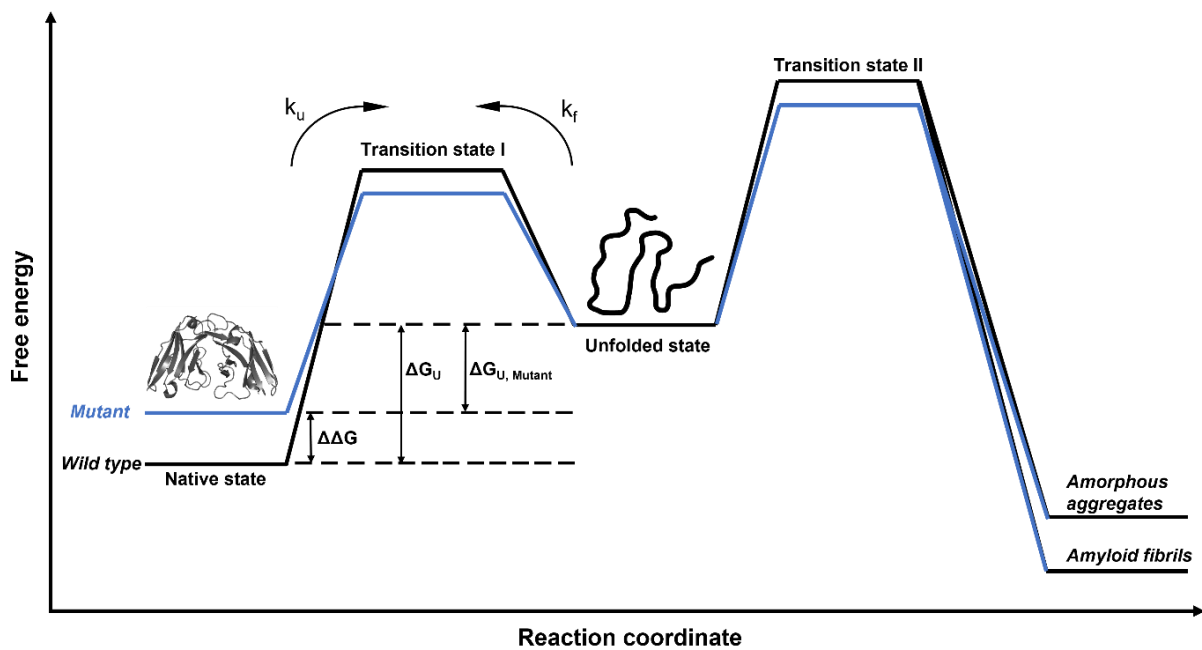


Figure 4: Schematic free energy diagram of protein unfolding and aggregation. The difference in free energy between the native and unfolded state determines the thermodynamic protein stability, whereas the difference between native and transition state I is termed kinetic protein stability. Mutations (“Mutant”; labeled in blue) destabilize the protein both thermodynamically and kinetically, thus influencing the rate constant of unfolding (k_u). Further, mutations can lower the energy barrier for the transition from the unfolded state into protein aggregates.

Further, mutations decrease the kinetic energy, that is the difference in Gibbs energy between the native and transition state, which means that partially folded or misfolded states are more readily accessible. These intermediate states (on- or off-pathway) can be kinetically trapped, i.e., the activation energy barrier is too high to reverse back to the native state meaning that the partially folded intermediate state stays significantly populated. This paves the way for protein aggregation due to the emergence of hydrophobic sequence patches on the protein surface which would normally be buried inside the protein core in the native state. These hydrophobic patches then engage in intermolecular contacts leading to non-native oligomerization and the subsequent formation of aggregates (Dobson, 2001; Chiti et al., 2002; Chiti et al., 2003). An overview of the different conformational states that proteins can assume is given in Figure 5 (Dobson, 2003). Protein aggregation is implicated in numerous diseases although one needs to discriminate between amorphous aggregates and so-called amyloid fibrils which represent a special kind of protein aggregate that will be discussed in more detail in the following chapter (Dobson, 2003).

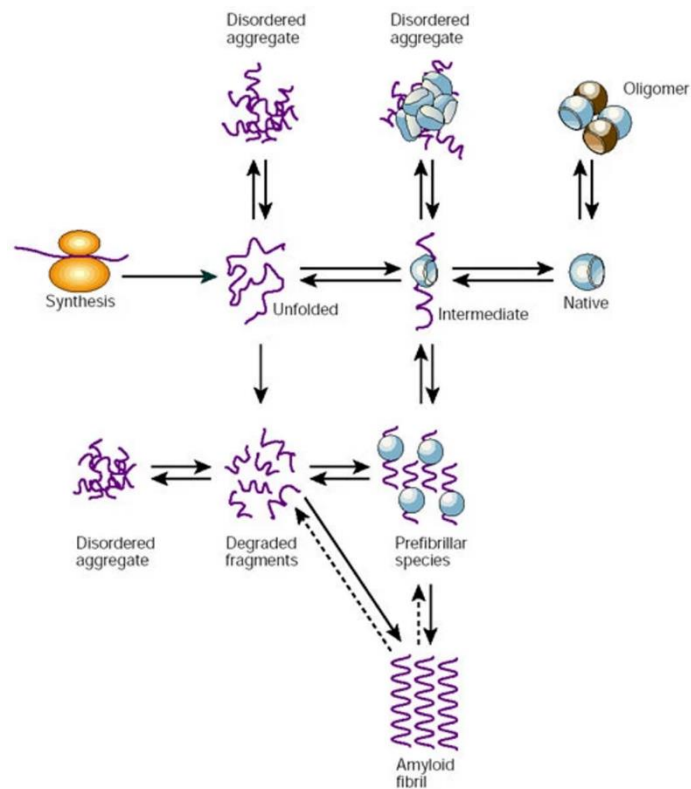


Figure 5: Overview of the conformational states of proteins. Upon ribosomal synthesis, the protein has to fold into its native structure via intermediate states. These can be kinetically trapped thus enabling protein aggregation which can occur either in amorphous form or as highly-ordered amyloid fibrils. Adapted and modified with permission from Dobson, 2003.

1.2.2 The amyloid state of proteins

As mentioned above, amyloid fibrils represent a kind of protein aggregate with a certain characteristic feature: the aggregates are not amorphous but rather exhibit a highly ordered structural organization. The precursor protein (i.e., the soluble monomer that aggregates) misfolds and the monomers stack upon each other to form protofilaments. These protofilaments pack together to form the insoluble amyloid fibrils which typically exhibit a twisted conformation and are approximately 6-12 nm in diameter. Amyloid fibrils can be identified by emission of a typical green birefringence under cross polarization when stained with Congo red. Further, they show a characteristic x-ray diffraction pattern indicating a spacing between the stacked monomers of around 5 Å and a distance of approximately 10 Å between two opposing protofilaments (Riek and Eisenberg, 2016). The fibril precursor protein does not possess its original, native structure but rather undergoes a major structural reorganization during the fibrillation process (Lecoq et al., 2019). Thereby, the typical cross-β motif found in amyloid fibrils is built up which consists of stacked parallel β-strands oriented perpendicularly to the fibril axis. This cross-β architecture allows formation of an extensive network of hydrogen bonds that stabilize the fibrillar structure (Astbury et al., 1935; Sunde et

al., 1997; Sunde and Blake, 1997; Fitzpatrick et al., 2013). Further stabilization is conferred by side chain interactions between individual protofilaments which is known as steric zipper. Due to the highly ordered structure and the large interaction network, amyloid fibrils are extremely stable and rigid (Jimenez et al., 2002; Schmidt et al., 2015). The structural organization of amyloid fibrils is shown in Figure 6.

The kinetics of amyloid formation are described by a nucleated polymerization mechanism characterized by sigmoidal growth behavior in which an initial lag phase is followed by a rapid, exponential growth phase, and a final plateau phase (Knowles et al., 2014). During the lag phase, the process of primary nucleation takes place including unfolding and oligomerization of soluble precursor monomers which undergo structural reorganization (Kazman et al., 2021). In the subsequent growth phase, fibril elongation occurs, i.e., soluble monomers attach to the end of nuclei or already existing fibrils, thereby adopting the cross- β conformation. Thus, nuclei and fibrils act as structural templates for added monomers. In the plateau phase, the pool of soluble monomers is depleted and fibril elongation stops, i.e., monomer attachment and dissociation from the fibrils are in equilibrium resulting in an apparent stop of fibril growth (Knowles et al., 2014). Apart from primary nucleation, the mechanism of secondary nucleation describes the branched addition of monomers to fibrils rather than simple linear elongation. In this process, the surface of the existing fibril acts as a heterogeneous catalyst facilitating the structural conversion and attachment of soluble precursor proteins to the fibril (Cohen et al., 2012; Cohen et al., 2013). Further, the fragmentation of already existing fibrils also represents a mechanism of secondary nucleation since the fracture of a fibril results in the emergence of two novel fibril ends or elongation sites, respectively (Knowles et al., 2014). The mechanisms of fibril formation are shown in Figure 7.

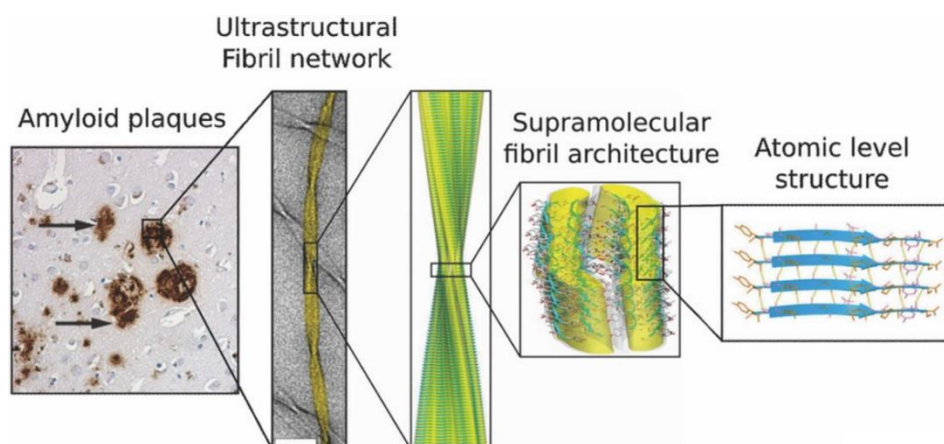


Figure 6: Amyloid plaques and the structure of amyloid fibrils at atomic level resolution. The depicted fibrils consist of a fragment of the amyloidogenic protein transthyretin. The reconstructed section of the cryo-EM image shows the twisted protofilaments with the stacked β -sheets running perpendicularly to the fibril axis. Figure adapted with permission from Fitzpatrick et al., 2013 and Knowles and Mezzenga, 2016.

Amyloid fibrils are most known as a disease hallmark and there is an entire family of human conditions, termed amyloidoses, that involve amyloid fibrils (Chiti and Dobson, 2017). However, it has been shown that the ability to form amyloid fibers is not only restricted to disease-associated proteins. Instead, depending on the surrounding conditions almost any protein or peptide can engage in the amyloid pathway *in vitro* (Fändrich and Dobson, 2002). Therefore, it is hypothesized that amyloid formation is a generic feature of proteins and peptides rather than a state that is necessarily connected to disease (Knowles et al., 2014). Accordingly, it has been shown that the amyloid state can be thermodynamically more stable than the soluble native state of a protein (Baldwin et al., 2011). This is supported by the observation of functional amyloid, i.e., amyloid fibrils that fulfil a beneficial physiological function instead of causing damage. These types of fibrils have for instance been found in microbial biofilms and are subject to extensive research because the advantageous properties of amyloid fibrils (e.g., high stability) could enable novel applications (Fowler et al., 2007; Knowles and Mezzenga, 2016).

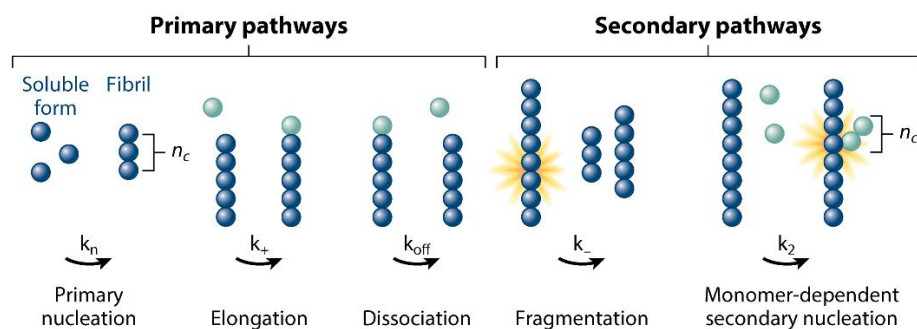


Figure 7: Schematic representation of the primary and secondary mechanisms of amyloid formation. Primary pathways are the initial nucleation from (partially) misfolded monomers and the subsequent fibril elongation, which is in equilibrium with the reverse process of monomer dissociation from the fibril end. Secondary pathways describe the fragmentation of fibrils, which releases two additional elongation sites, i.e., the two new fibril ends. Secondary nucleation occurs on the surface of already existing fibrils leading to branched aggregation form. Adapted and modified with permission from Chiti and Dobson, 2017.

1.2.3 Amyloidoses

As already established, protein misfolding and aggregation are hallmarks of numerous human diseases. Conditions that involve the formation of highly ordered amyloid fibrils and subsequent pathological deposition as so-called amyloid plaques are collectively termed amyloidoses. Today, around 40 different human amyloidoses are known and they can be classified by their individual precursor proteins. A few of these conditions are summarized in Table 1. The majority of the precursor proteins are secreted from the cells and, therefore, the

amyloid deposits occur most often in the extracellular space. Only a small number of cytosolic proteins engage in amyloid formation. Furthermore, these illnesses can occur either systemically or in a localized manner, i.e., restricted to a certain organ or part of the body. Among the localized amyloidoses, the ones that affect the brain and nervous system, termed neurodegenerative conditions, are of special interest due to their detrimental nature (Chiti and Dobson, 2017).

The most widely known amyloidoses are Alzheimer's disease (AD) and Parkinson's disease (PD) which both belong to the aforementioned neurodegenerative diseases. In AD, the β -amyloid precursor protein (APP), a ubiquitously expressed transmembrane protein, is proteolytically cleaved by a secretase, thereby releasing the 40- or 42-residue containing amyloid- β peptide ($A\beta$) which subsequently engages in fibrillar aggregation. The resulting amyloid plaques deposit in the brain of the patient where they cause irreversible damage to neural cells and lead to dementia and ultimately death. However, not only mature amyloid fibrils but also soluble oligomers of the $A\beta$ peptide can affect synaptic function and the metabolism of neuronal cells, thus contributing to the disease-progression of AD (Haass and Selkoe, 2007; Querfurth and LaFerla, 2010). In PD, the protein α -synuclein acts as precursor and eventually aggregates forming so-called Lewy bodies. The intrinsically disordered α -synuclein protein consists of 140 amino acids and its normal function in neuronal cells is not fully understood. The Lewy body plaques are deposited in the central nervous system causing severe harm to motor neurons which leads to the characteristic tremors (Poewe et al., 2017). Amyloid diseases are becoming increasingly prevalent in an ageing society and are starting to pose a major public health issue (Olshansky et al., 2005; Querfurth and LaFerla, 2010). Further, protein aggregation, either amorphous or amyloid, is associated with other diseases (e.g., cancer), too, and not only with typical protein misfolding diseases (Chiti and Dobson, 2017). Despite intensive efforts by academia and pharmaceutical industry, amyloid diseases are still hard to tackle and the possibilities for therapeutic intervention are scarce (Sacchettini and Kelly, 2002; Iadanza et al., 2018). One type that is particularly challenging due to its inherent precursor heterogeneity is system light chain (AL) amyloidosis, which will be covered in more detail in chapter 1.4.

Table 1: Overview of a selection of human amyloid diseases.

Disease	Aggregating peptide or protein	Precursor structure	Affected organ(s)
<i>Neurodegenerative diseases</i>			
Alzheimer's disease	Amyloid- β peptide (A β)	Intrinsically disordered	Brain
Parkinson's disease	α -synuclein	Intrinsically disordered	Brain and CNS
Huntington's disease	Huntingtin (HTT)	Intrinsically disordered	Brain and CNS
Amyotrophic lateral sclerosis (ALS)	Superoxide dismutase (SOD)	β -sheet	Brain and CNS
Familial amyloidotic polyneuropathy	Transthyretin mutants	β -sheet	Brain and CNS
Creutzfeld-Jakob disease	Prion protein (PrP)	Intrinsically disordered and α -helical	Brain
<i>Systemic, non-neuropathic amyloidoses</i>			
Light chain (AL) amyloidosis	Fragments of Ig light chains	β -sheet	Heart, kidney, liver, soft tissue, gastrointestinal tract
Heavy chain (AH) amyloidosis	Fragments of Ig heavy chains	β -sheet	Kidneys, heart, liver
Senile systemic amyloidosis	Transthyretin wild type	β -sheet	Heart, gastrointestinal tract, liver, kidney
Serum amyloid A (AA) amyloidosis	Serum Amyloid A1 protein (SAA)	α -helical	Spleen, kidney, gastrointestinal tract
Dialysis-related amyloidosis	β_2 -microglobulin	β -sheet	Bones, joints, osteoarticular surfaces
<i>Localized, non-neuropathic amyloidoses</i>			
Type-II diabetes	Islet amyloid polypeptide (IAPP)	Intrinsically disordered	Pancreas
Injection-localized amyloidosis	Insulin	α -helical	Injection site

1.3 Immunoglobulins

1.3.1 Structure of antibodies and the immunoglobulin fold

The human immune system can be divided into two parts: the innate system, that confers an unspecific, first line of defense, and the adaptive system, which is responsible for tackling threats (e.g., bacteria, viruses, cancer cells) in a more selective and targeted manner. One of the most important components of the adaptive immune response are immunoglobulins, also known as antibodies. They play major roles as B cell surface receptors in the cellular immune

response and as secreted effector molecules in the humoral response. Antibodies can bind antigens (e.g., bacterial toxins or surface proteins) with high selectivity and specificity in a process called opsonization which leads to the recruitment of immune cells that eventually eradicate the extrinsic threat (Murphy and Weaver, 2016). Immunoglobulins exhibit a characteristic Y-shaped molecular structure and are built up as hetero-tetrameric glycoproteins consisting of two identical heavy chains (HC) which are connected to two light chains (LC) via disulfide bonds. The HCs have an N-terminal variable domain (V_H) followed by three constant domains (C_{H1-3}) although there are different types of antibodies that can contain more constant domains. The LC consists of an N-terminal variable domain (V_L) and one constant domain (C_L) (Feige et al., 2010). Further, all antibody domains contain a highly conserved disulfide bridge which is important for domain stability (Goto et al., 1979; Bork et al., 1994). The variable domains are crucial for antigen binding which is why they contain three segments that exhibit a very high sequence variability, so-called complementarity-determining regions (CDRs) surrounded by conserved framework regions (FRs). Furthermore, antibodies can be divided into an antigen-binding fragment (F_{ab}), comprising the LC and V_H - C_{H1} , and constant, so-called crystallizable fragment (F_c) consisting of the residual C_{H2} and C_{H3} domains. While the F_{ab} contains the paratope (i.e., the six CDRs), the F_c fulfils effector functions and can bind to surface receptors for immune cell recruitment (Chiu et al., 2019; Hussain et al., 2021).

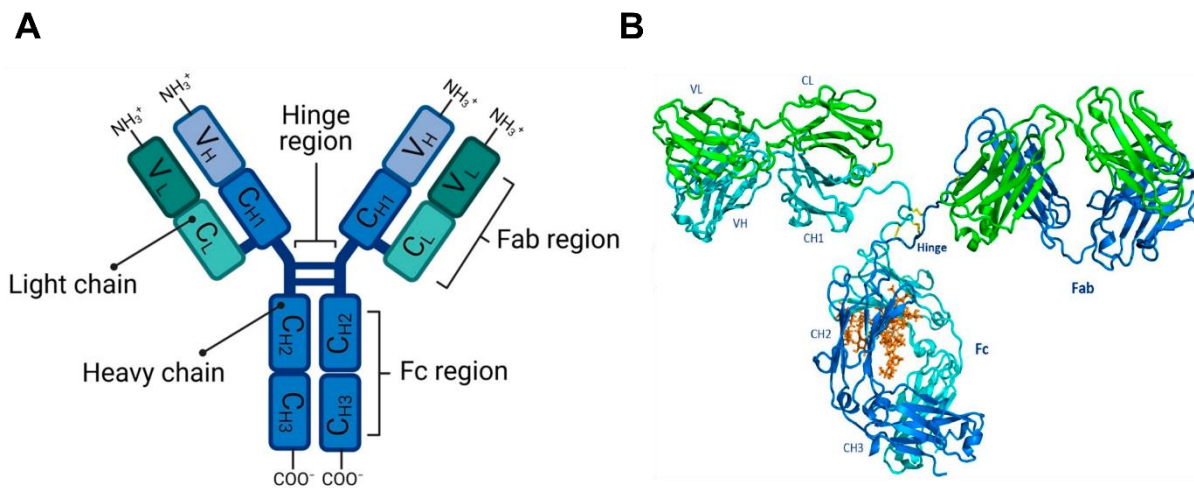


Figure 8: Structure of an IgG antibody. A) Schematic representation of the domain arrangement of antibodies. Adapted from Hussain et al., 2021. B) Crystal structure (PDB: 1IGT) of a murine IgG antibody. The LCs are shown in green, the HCs in blue and cyan, glycosyl moieties in orange, and disulfide bridges in yellow. Adapted and modified from Chiu et al., 2019.

There are different classes of antibodies exhibiting slightly different structural features and functions, with human antibodies being divided into the isotypes IgA, IgD, IgE, IgG, and IgM (Schroeder Jr and Cavacini, 2010). The Ig class is determined by the HC of the antibody, which can affect the structure since IgE and IgM, for instance, contain a fourth constant domain (C_{H4}). Further, IgM adopts a different quaternary structure by assembling into pentamers and

hexamers via a short C-terminal peptide extension called tailpiece and with the help of the J chain protein (Pasalich et al., 2017; Li et al., 2020). Similarly, IgA forms dimers, whereas the other classes are only present as monomers. The Ig classes differ functionally, i.e., IgA is predominantly present on mucosal surfaces and in secretory fluids, IgE participates in the defense against multicellular parasites but is also the antibody associated with allergies. IgG represents the most prevalent antibody class and occurs mainly in the blood stream, while IgM serves as a first line of defense in blood and mucosal tissue. IgD is found both on cell surfaces and as secreted antibody in the blood, however, only very little is known about the function of IgD so far (Schroeder Jr and Cavacini, 2010). In addition to the HC, also the LCs are categorized into different subtypes namely lambda (λ) and kappa (κ) which can further be divided into subfamilies, i.e., λ I, λ II, λ III, etc. The two LC subtypes are expressed in different amounts, i.e., in humans the physiological κ/λ -ratio is approximately 2/1, whereas other species show different ratios (Dorshkind and Rawlings, 2018; Absmeier et al., 2022).

Immunoglobulin domains share a highly conserved β -sandwich fold containing two antiparallel β -sheets. Constant domains exhibit a c-type Ig fold comprising seven β -strands (abcdefg) whereas variable domain belong to the v-type Ig fold which is built up of nine β -strands (abcc'c''defg) (Bork et al., 1994; Chatterjee et al., 2021). The individual strands are connected by short protruding loops (including the CDRs) which often contain proline residues. These proline residues play an important role in Ig domain architecture and proline cis-trans-isomerization is a crucial step in antibody folding, catalyzed by a class of chaperones called peptidyl-prolyl-cis-trans-isomerases (PPIs) (Feige et al., 2010). Notably, the CDR loops do not exhibit a random disordered structure, but rather they can be grouped into several so-called canonical classes with restricted loop conformations with the HC CDR3 being the only exception (Chothia et al., 1989; Al-Lazikani et al., 1997). While the Ig fold is generally highly conserved throughout all species, differences regarding overall antibody architecture can occur. For example, sharks and camelids are known to possess heavy chain only antibodies, i.e., antibodies that lack a light chain (Feige et al., 2014; Los Rios et al., 2015). Further, cattle have a subset of antibodies containing V_H domains with ultralong CDR3 segments. These bovine ultralong CDR3 are between 50 to 70 amino acids long, whereas regular human CDR3 loops normally comprise only 8 to 16 residues. Regarding their structure, ultralong CDR loops form a long stalk protruding from the V_H and a so-called knob mini domain sitting on top whose local conformation is largely determined by disulfide bridges. Such unusual antibody formats can offer novel properties in terms of structural stability and binding capacity which is particularly interesting for the numerous biotechnological and pharmaceutical applications of antibodies (Wang et al., 2013; Svilenov et al., 2021).

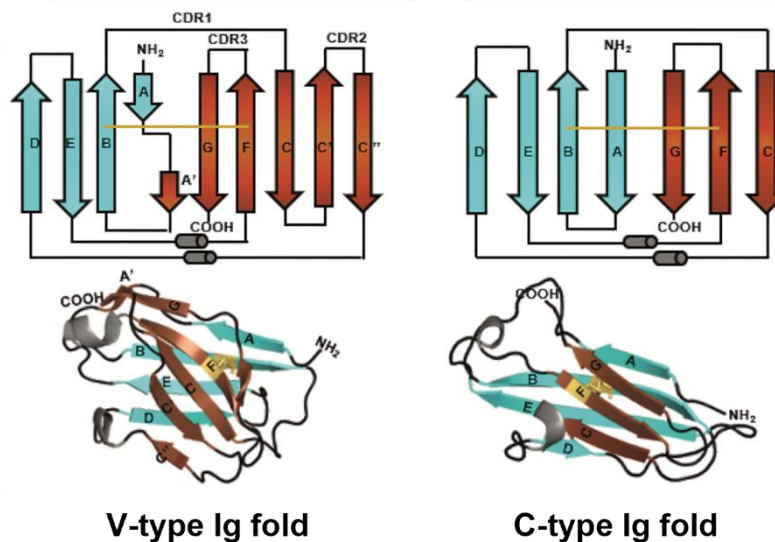


Figure 9: Ig domain fold topology of V-type and C-type domains. V-type Ig domains consist of nine β -strands, although strand A can be split up as shown here. C-type domains consist of seven β -strands. The conserved disulfide bridge is depicted in yellow and short helical segments are shown as grey cylinders. Adapted and modified with permission from Chatterjee et al., 2020.

1.3.2 Emergence of sequence variability

As stated above, antibody variable domains contain three hypervariable CDR loops with a high degree of sequence diversity that is needed to recognize the vast number of antigens our immune system is confronted with. This sequence variability is the result of gene rearrangement events, a mechanism termed VDJ recombination, and of somatic hypermutation (SHM) which both occur during the maturation of antibody-producing cells (Bernard et al., 1978; Tonegawa, 1983). In the course of an infection for instance, naïve B cells in the bone marrow express antibodies as receptors on their cell surface and can thereby bind to a pathogen-derived antigen. Each B cell expresses a different antibody on its surface and via clonal selection, the antibody with the highest affinity is identified. This particular B cell then develops into a plasma cell and secretes large amounts of that specific antibody to eradicate the infecting agent (Murphy and Weaver, 2016). In order to generate a highly variable antibody repertoire in B cells, the antibody encoding gene loci consist of numerous segments which can be divided into Variable (V), Diversity (D), Joining (J), and Constant (C) gene segments. The HC locus on chromosome 14 consists of approximately 40-50 V segments, 20-30 D segments, 6 J segments, and the different C segments corresponding to the aforementioned heavy chain isotypes. In the course of affinity maturation, one segment of each type is randomly connected by RAG recombinases to form the exon coding for the complete heavy chain (Teng and Schatz, 2015; Dorshkind and Rawlings, 2018). Furthermore, specific enzymes can introduce additional nucleotides which alter the amino acid sequence of the resulting antibody, thereby conferring an even higher degree of sequence diversity. This process is termed somatic hypermutation

(SHM) and it occurs predominantly in the joining region where V, D, and J segments are connected (Motea and Berdis, 2010; Pavri and Nussenzweig, 2011). Notably, this region encodes the CDR3 loop of the antibody, thus explaining the high variability of this particular region (Tonegawa, 1983; Alt et al., 1987). Sequence diversification for antibody LCs is achieved by the same mechanisms although there are some minor differences (Collins and Watson, 2018). The LC gene loci on chromosomes 2 (κ) and 22 (λ) do not contain D segments but only V, J, and C segments with one C segment on the κ -locus and four C segments on the λ -locus. Thus, four different isotypes of λ constant domains exist. Regarding the number of V and J segments on the LC gene loci, there are around 40 different functional V gene segments and five different J segments on both the κ -locus and the λ -locus, respectively (Collins and Watson, 2018). The λ gene locus encodes V_L s that can be grouped into ten different subfamilies by structural and sequence similarity, whereas the V_L s emerging from the κ -locus can be grouped in up to seven subfamilies (Dorshkind and Rawlings, 2018). As the structure of the different gene loci already suggests, V_H domains exhibit a higher sequence diversity than V_L domains and accordingly the CDR3 loop of the HC – the most diverse region of the antibody – is the part that is predominantly involved in antigen binding (Al-Lazikani et al., 1997).

1.4 Systemic light chain (AL) amyloidosis

1.4.1 Disease characteristics

Systemic light chain amyloidosis is a debilitating protein misfolding disease characterized by the extracellular deposition of amyloid fibrils made up of antibody LCs. The underlying cause is most commonly a plasma cell dyscrasia which is why AL amyloidosis often occurs as a co-morbidity of multiple myeloma (MM) (Zhang et al., 2017). The aberrant, monoclonal plasma cells produce large amounts of a certain antibody LC which is predominantly not associated with a corresponding HC but is rather secreted into the blood stream as free light chains (FLCs) (Morgan and Wall, 2020). It has been shown that these pathogenic plasma cells are not naïve but have undergone antigen-driven selection and somatic hypermutation before they home in the bone marrow (Perfetti et al., 1998). The secretion of FLCs can either occur in the monomeric form or as covalently linked LC homodimers that are connected via a C-terminal disulfide bridge. These covalent LC homodimers are called Bence-Jones proteins and can be detected in large amounts in patient urine (Kaplan et al., 2009; Kaplan et al., 2011). Secretion of vast amounts of FLCs into the blood stream is also a hallmark of MM and other diseases, so-called monoclonal gammopathies. Yet, only in a small fraction of these patients amyloid deposits form and they develop AL amyloidosis (Merlini and Bellotti, 2003). Therefore, the LC primary sequence and the presence of unfavorable point mutations play an important role in AL (Absmeier et al., 2022). Upon secretion from plasma cells, the LCs can misfold due to

destabilizing mutations and assemble into amyloid fibrils, which subsequently deposit in the extracellular space of various organs. Most commonly the heart and kidneys are affected although deposits in liver, gastrointestinal tract, and nervous system can also occur (Merlini et al., 2018). However, in the majority of cases it is not the complete LC that undergoes amyloid aggregation but rather fragments thereof, most commonly the isolated V_L domain, although deposition of full-length LCs and C_L domains has also been reported (Glenner et al., 1970; Solomon et al., 1998; Lavatelli et al., 2008; Blancas-Mejía and Ramirez-Alvarado, 2013). In the target organ, the amyloid fibrils cause severe damage leading to dysfunction and eventually organ failure (Brenner et al., 2004; Marin-Argany et al., 2016). However, studies show that not only the mature fibrils play a role in the pathogenicity, but also smaller oligomeric fibril precursors have been identified as cytotoxic agents that can affect organ function (Reixach et al., 2004; Maritan et al., 2020). This phenomenon is also known for other amyloid diseases, e.g., for AD, where soluble oligomers are thought to play an equally important role due to their cytotoxicity (Haass and Selkoe, 2007). The amyloid plaques often also contain other proteins that are co-deposited such as serum amyloid P (SAP), ApoE4, or the extracellular chaperone Clusterin (Zhang et al., 2017). Both SAP and ApoE4 are associated with amyloid diseases and are known as pro-amyloidogenic factors (Saunders et al., 1993; Tennent et al., 1995). The co-deposited chaperone Clusterin, however, counteracts amyloid formation by contributing to the clearance of fibrils (Greene et al., 2011; Wyatt et al., 2011). Additionally, non-proteinaceous components such as glycosaminoglycans (GAGs), for instance, can be associated with the fibrils as well and have been shown to affect fibril formation (Blancas-Mejía et al., 2015a; Zhang et al., 2017). These accompanying factors raise the question how proteins of the blood plasma, extracellular chaperones, and components of extracellular matrices are involved in the fibril formation process (Zhang et al., 2017; Gottwald and Röcken, 2021).

AL amyloidosis is set apart from other amyloid diseases by the vast sequence variability of the precursor protein. Due to gene rearrangement and somatic mutations, antibody LCs and especially V_L domains display a highly diverse set of sequences and mutations that is needed for antigen recognition. However, some of these mutations can also be the cause of LC destabilization, misfolding, and aggregation. The enormous sequence variability is, therefore, problematic in the disease context since every patient is unique and it is not possible to predict which mutations will drive an LC towards misfolding and amyloid formation and which mutations have no effect in this regard. Typically, amyloidogenic LCs from patients contain roughly between four and fifteen mutations with the majority occurring in the variable CDRs rather than in conserved framework regions (Blancas-Mejía and Ramirez-Alvarado, 2013; Absmeier et al., 2022). From a biophysical point of view, these different mutations can have quite diverse effects on the structure and the properties of the LC. The amino acid substitutions

can disrupt native interactions which can either decrease the thermodynamic stability or the kinetic stability, thereby enabling the enhanced population of partially folded or misfolded states (Hurle et al., 1994; Raffin et al., 1999; Wall et al., 1999; Kim et al., 2000; Khurana et al., 2001; Qin et al., 2007; Feige et al., 2008). Further, protein backbone dynamics have been shown to play a key role in LC fold stability and mutations can affect the structural dynamics drastically (Valdés-García et al., 2017; Maya-Martinez et al., 2019; Lewkowicz and Gursky, 2022). In line with the structural flexibility that mutations can cause, an LC or V_L domain has to undergo major conformational rearrangements and reorganizations to assume the cross- β of amyloid fibrils (Lecoq et al., 2019). These structural changes that pave the way for the subsequent molecular cascade leading to amyloid formation take place early in the lag phase. It has been shown, that these structural conversions can lead to enhanced surface hydrophobicity and the formation of non-native dimers that already contain the parallel β -sheet motif later found in the fibrils (Kazman et al., 2021). In addition to affecting dynamics, mutations can also disrupt or destabilize the LC dimerization interface, which is responsible for association with the HC or with another LC in Bence-Jones proteins. The role of quaternary structure in the context of AL amyloidosis is well known and studies demonstrated the protective effect of dimer association which is why mutations that hamper this dimer association can have deleterious effects (Baden et al., 2008a; Brumshtein et al., 2014; Wolwertz et al., 2016b). Apart from somatic mutations, the underlying germline gene can also have an impact as the distribution of occurring LC subtypes in AL patients is approximately $\lambda/\kappa = 3:1$. In healthy individuals the ratio of κ to λ -LCs is approximately 2:1 (Absmeier et al., 2022). Further, there are certain germline genes that are primarily associated with the disease, as they are found with higher prevalence in patients, for instance, the $\lambda 6a$ and $\lambda 3r$ germline segments (Solomon et al., 1982; Perfetti et al., 2002). Strikingly, a connection between certain amyloidogenic germlines and specific organ involvement has also been reported. For example, the $\lambda 6a$ germline is associated with renal involvement although the reason behind this pathogenic connection is yet unknown (Merlini, 2017).

1.4.2 Diagnosis and treatment

AL amyloidosis is the most common form of systemic amyloid diseases with an incidence of approximately 9-12 newly diagnosed patients per million people each year. The disease is lethal if left untreated and the median survival rate without effective treatment is one to two years (Duhamel et al., 2017; Zhang et al., 2017). The condition occurs primarily in older individuals since the average age of diagnosed patients is 65 years and roughly 10 % of patients are under 50 (Desport et al., 2012). The symptoms of AL amyloidosis are very diverse and include restrictive cardiomyopathy (heart involvement), impaired renal function and proteinuria (renal involvement), macroglossia, dyspnoea, fatigue, and asthenia among others.

These unspecific symptoms are the cause of delayed diagnosis and misdiagnosis which is problematic since an early diagnosis and therapeutic intervention are crucial for AL patients, especially considering that the irreversible organ damage caused by amyloid plaques often creates the need for organ transplantation (Desport et al., 2012; Merlini, 2017). Diagnosis usually involves tissue biopsy and Congo Red staining of abdominal fat and the affected organ, followed by amyloid-typing by mass spectrometry or scintigraphy (Merlini, 2017). For cardiac AL amyloidosis, N-terminal pro-brain natriuretic peptide (NT-proBNP) can be used as a sensitive biomarker for myocardial dysfunction and as a useful prognostic tool (Palladini et al., 2003).

The standard approaches for treatment of AL are myeloma-oriented chemotherapy and autologous stem cell transplantation (ASCT) to eradicate the malignant plasma cell clones and thereby prevent further FLC secretion (D'Souza et al., 2015; Gertz, 2018). For this purpose, high dose dexamethasone or combinatorial approaches with dexamethasone-melphalan and bortezomib-cyclophosphamide-dexamethasone, respectively, are commonly applied (Palladini et al., 2004; Palladini et al., 2015; Zhang et al., 2017). However, these treatments only aim at stopping further secretion of FLCs from plasma cells and cannot target already existing amyloid plaques. In addition, chemotherapy is often accompanied by severe side effects for patients which further necessitates the development of additional therapeutic options (Absmeier et al., 2022). Therefore, the novel proteasome-inhibitors carfilzomib (Cohen et al., 2016) and ixazomib (Santhorawala et al., 2017) are being investigated, as well as CD38-targeting antibodies daratumumab and elotuzumab (Iqbal et al., 2019; Hassan et al., 2021). For the removal of amyloid plaques, studies with fibril-targeting antibodies 11-1F4 (Edwards et al., 2017) and NEOD001 (Gertz et al., 2016) are being conducted. Another compound that could potentially be administered for the clearance of fibrillar deposits is epigallocatechin-3-gallate (EGCG) which is a plant-derived polyphenol found in green tea (Andrich et al., 2017; Hora et al., 2017). A completely novel approach is currently undertaken by the development of molecules that stabilize precursor proteins by binding to the LC homo-dimers, thus preventing them from misfolding and subsequent aggregation. Coumarin-based and hydantoin-based compounds can be used in such a way for instance (Morgan et al., 2019; Yan et al., 2020; Yan et al., 2021). A similar strategy has been successfully applied with the drug tafamidis in transthyretin amyloidosis (ATTR). Here, the native transthyretin homo-tetramer is bound by the pharmacological chaperone tafamidis leading to structural stabilization thus preventing the dissociation of potentially aggregation-prone monomers (Bulawa et al., 2012). Nonetheless, AL amyloidosis is still a challenging disease and a more detailed understanding of the underlying disease mechanisms is key for the development of new therapeutic approaches.

1.5 Objectives of the thesis

The scope of this scientific thesis was to investigate the underlying molecular mechanisms of systemic light chain (AL) amyloidosis. Therefore, the patient-derived LC FOR005 was examined and characterized *in vitro* by an array of biophysical and biochemical methods. This sequence was obtained from a patient at the Amyloidosis Center of the University Hospital Heidelberg and was provided by our collaborators Dr. Rolf Köhler, Prof. Dr. Ute Hegenbart, and Prof. Dr. Stefan Schönland. The FOR005 project is divided into two parts: in the first part, the isolated V_L domain of FOR005 was characterized to obtain insight into the role of amyloidogenic point mutations and their influence on the biophysical properties of the domain. The second part aimed at a deeper understanding of the mechanisms of disease onset by investigating the FOR005 full-length LC and characterizing the effects of a C_L domain mutation.

The first part is based on the findings of Annamalai et al. who had reported that the amyloid deposits derived from patient tissue contained only the V_L domain of FOR005. By determination of the closest related germline sequence, five point mutations were identified in the V_L domain. The impact of these individual point mutations on the biophysical properties of the domain were explored in the context of misfolding and amyloid formation, respectively. Therefore, the V_L domain constructs were produced recombinantly, purified, and subsequently tested by a variety of methods concerning protein stability, aggregation, and conformational dynamics.

In the second part of the project the focus was set on the full-length LC of FOR005 and a point mutation that was found in the C_L domain. Substitutions in this normally constant part of the antibody LC are rare and accordingly the effects of such mutations in the context of AL amyloidosis are enigmatic. Combining recombinant LC constructs and biophysical characterization methods provided insights into the role of V_L-C_L domain interactions, LC quaternary structure, and proteolytic cleavage of LCs in the context of FOR005. The results of this study allowed the deduction of a stepwise mechanistic model for the disease onset in patient FOR005 that can be integrated in a broader cell-biological context thus providing insights into the mechanisms that govern onset and progression of AL amyloidosis.

2. Material and Methods

2.1 Chemicals

Chemical	Supplier
2-Mercaptoethanol	Merck (Darmstadt, Germany)
8-Anilino-naphthalene-1-sulfonate (ANS)	Sigma (St. Louis, USA)
Acetic acid	Roth (Karlsruhe, Germany)
Acetonitrile	Merck (Darmstadt, Germany)
Acrylamide	SERVA (Heidelberg, Germany)
Acrylamide/Bisacrylamide (40 % w/v)	SERVA (Heidelberg, Germany)
Agar agar	SERVA (Heidelberg, Germany)
Agarose	SERVA (Heidelberg, Germany)
Ammonium chloride (NH ₄ Cl)	Merck (Darmstadt, Germany)
Ammonium persulfate (APS)	Roth (Karlsruhe, Germany)
Ampicillin sodium salt	Roth (Karlsruhe, Germany)
Bromphenol blue	SERVA (Heidelberg, Germany)
Calcium chloride (CaCl ₂)	Merck (Darmstadt, Germany)
Coomassie Blue R	SERVA (Heidelberg, Germany)
Coomassie Brilliant Blue R-250	SERVA (Heidelberg, Germany)
Desoxynucleoside triphosphates (dNTPs)	NEB (Frankfurt, Germany)
Deuterium oxide (D ₂ O)	Merck (Darmstadt, Germany)
Dimethyl sulfoxide (DMSO)	NEB (Frankfurt, Germany)
Dithiothreitol (DTT)	Roth (Karlsruhe, Germany)
Ethylenediaminetetraacetic acid (EDTA)	Merck (Darmstadt, Germany)
Ethanol	Merck (Darmstadt, Germany)
Formic acid	Merck (Darmstadt, Germany)
Glutathione, oxidized (GSSG)	Merck (Darmstadt, Germany)
Glutathione, reduced (GSH)	Merck (Darmstadt, Germany)
Glycerol	Roth (Karlsruhe, Germany)
Glycine	Roth (Karlsruhe, Germany)
Guanidinium chloride (GdmCl)	Merck (Darmstadt, Germany)
Hydrochloric acid (HCl)	Merck (Darmstadt, Germany)
Imidazole	Sigma (St. Louis, USA)
Isopropanol	Merck (Darmstadt, Germany)

Chemical	Supplier
Isopropyl-β-d-1-thiogalactopyranoside (IPTG)	SERVA (Heidelberg, Germany)
Kanamycin sulfate	Roth (Karlsruhe, Germany)
L-Arginine	Merck (Darmstadt, Germany)
Luria-Bertani (LB) medium	Merck (Darmstadt, Germany)
Magnesium chloride (MgCl ₂)	Merck (Darmstadt, Germany)
Manganese (II) chloride (MnCl ₂)	Merck (Darmstadt, Germany)
N,N,N',N'-Tetramethyl ethylenediamine (TEMED)	Roth (Karlsruhe, Germany)
NP 40 (Igepal)	Sigma (St. Louis, USA)
Phenylmethylsulfonyl fluoride (PMSF)	Roth (Karlsruhe, Germany)
Phosphoric acid (H ₃ PO ₄)	Roth (Karlsruhe, Germany)
Potassium chloride (KCl)	Roth (Karlsruhe, Germany)
Potassium phosphate, monobasic (KH ₂ PO ₄)	Merck (Darmstadt, Germany)
Protease Inhibitor Mix HP	Roth (Karlsruhe, Germany)
Sodium acetate, anhydrous (NaOAc)	Merck (Darmstadt, Germany)
Sodium azide (NaN ₃)	Merck (Darmstadt, Germany)
Sodium chloride (NaCl)	Merck (Darmstadt, Germany)
Sodium dodecylsulfate (SDS)	SERVA (Heidelberg, Germany)
Sodium hydroxide (NaOH)	Merck (Darmstadt, Germany)
Sodium phosphate, dibasic (Na ₂ HPO ₄ x 2 H ₂ O)	Merck (Darmstadt, Germany)
Sodium phosphate, monobasic (NaH ₂ PO ₄ x 2 H ₂ O)	Merck (Darmstadt, Germany)
Stain G	Merck (Darmstadt, Germany)
Thioflavin T (ThT)	Sigma (St. Louis, USA)
Tris(2-carboxyethyl) phosphine hydrochloride (TCEP)	Merck (Darmstadt, Germany)
Tris(hydroxymethyl)aminomethane (Tris)	Sigma (St. Louis, USA)
Triton X-100	Merck (Darmstadt, Germany)
Tween-20	Merck (Darmstadt, Germany)
Uranyl acetate	Science Services (Munich, Germany)
Urea	Merck (Darmstadt, Germany)

2.2 Consumables

Consumable	Supplier
200-mesh copper grids	Merck (Darmstadt, Germany)
Amicon Centrifugal filter units (0.5/4/15 ml)	Millipore (Darmstadt, Germany)
Crystal Clear PP sealing foil	HJ-Bioanalytik GmbH (Erkelenz, Germany)
Cuvettes, plastic 1 ml	Brand (Darmstadt, Germany)
Dialysis membranes Spectra/Por (MWCO: 6-8 kDa)	Spectrum Laboratories (Houston, USA)
Membrane discs	Sartorius (Göttingen, Germany)
Membrane filter 0.22 µm	Millipore (Darmstadt, Germany)
Microwellplate 96 well, black, PP	Greiner Bio-One (Frickenhausen, Germany)
Microwellplate 96 well, UV-Star, COC	Greiner Bio-One (Frickenhausen, Germany)
Millex-GS 0.22 µm syringe filter units	Millipore (Darmstadt, Germany)
Millex-GV 0.22 µm syringe filter units	Millipore (Darmstadt, Germany)
PCR tubes	Bio-Rad (Munich, Germany)
PE tubes (15/50 ml)	Greiner Bio-One (Frickenhausen, Germany)
Petri dishes, PS, 94 mm	Greiner Bio-One (Frickenhausen, Germany)
Reaction tubes (0.5/1.5/2 ml)	Sarstedt (Nümbrecht, Germany)
SERVAGel TG PRiME 4-20 % pre-cast gel	SERVA (Heidelberg, Germany)
Sealing tape for 96 well plate	Thermo Fisher (Waltham, USA)

2.3 Buffers and solutions

2.3.1 Buffers for protein purification

Buffer	Composition
Dialysis A buffer	50 mM Tris/HCl 5 M Urea pH 8.0
Dialysis B buffer	50 mM Tris/HCl 3 M Urea pH 8.5
HisTrap buffer A	20 mM Na ₂ HPO ₄ 20 mM NaH ₂ PO ₄ 300 mM NaCl 10 mM Imidazole pH 7.5
HisTrap buffer B	20 mM Na ₂ HPO ₄ / NaH ₂ PO ₄ 300 mM NaCl 300 mM Imidazole pH 7.5
10x Phosphate buffered saline (PBS)	100 mM Na ₂ HPO ₄ x 2 H ₂ O 18 mM KH ₂ PO ₄ 27 mM KCl 1.37 M NaCl
Q-Sepharose buffer A	50 mM Tris/HCl 5 M Urea pH 8.0
Q-Sepharose buffer B	50 mM Tris/HCl 5 M Urea 1 M NaCl pH 8.0
Refolding buffer	50 mM Tris/HCl 100 mM L-Arginine 0.5 mM GSH 1 mM GSSG pH 8.0

Buffer	Composition
2x Ulp1 storage buffer	50 mM Tris/HCl 500 mM NaCl 1 mM DTT pH 8.0

2.3.2 Buffers for SDS-PAGE

Buffer	Composition
Fairbanks A (staining)	25 % (v/v) Isopropanol 10 % (v/v) Acetic acid 0.05 % (w/v) Coomassie Blue R
Fairbanks D (staining)	10 % (v/v) Acetic acid
5x Laemmli buffer, reducing	0.3 M Tris/HCl, pH 6.8 10 % (w/v) SDS 50 % (v/v) Glycerol 5 % (v/v) 2-Mercaptoethanol 0.05 % (w/v) Bromphenol blue
5x Laemmli buffer, non-reducing	0.3 M Tris/HCl, pH 6.8 10 % (w/v) SDS 50 % (v/v) Glycerol 0.05 % (w/v) Bromphenol blue
10x SDS running buffer	0.25 M Tris/HCl, pH 8.0 2 M Glycine 1 % (w/v) SDS
4x Separation gel buffer	1.5 M Tris/HCl, pH 8.8 0.8 % (w/v) SDS
2x Stacking gel buffer	0.25 M Tris/HCl, pH 6.8 0.4 % (w/v) SDS

2.3.3 Buffers and solutions for molecular biology

Solution	Composition
1000x Antibiotic stock (in ddH ₂ O)	50 mg/ml for Ampicillin 35 mg/ml for Kanamycin sterile filtered 0.22 µm

Solution	Composition
1000x IPTG stock (in ddH ₂ O)	1 M IPTG sterile filtered 0.22 µm
Luria Bertani (LB) medium for bacteria	20 g/L LB-medium (SERVA) in ddH ₂ O
Solution A for competent bacterial cells	13 ml 3 M NaOAc, pH 5.5 100 ml 1 M CaCl ₂ 25 ml 2.8 M MnCl ₂ 862 ml H ₂ O
Solution B for competent bacterial cells	69 ml Glycerol (87 %) 331 ml Solution A
50x TAE buffer for agarose gels	2 M Tris/Acetate 50 mM EDTA pH 8.0

2.3.4 Buffers for protein biochemistry

Buffer	Composition
HDX Quenching buffer	200 mM Na ₂ HPO ₄ 200 mM NaH ₂ PO ₄ 250 mM TCEP 3 M GdmCl pH 2.2
Limited proteolysis buffer	100 mM Tris/HCl 100 mM NaCl 10 mM CaCl ₂ pH 7.8
1x Phosphate buffered saline (PBS)	10 mM Na ₂ HPO ₄ 1.8 mM KH ₂ PO ₄ 2.7 mM KCl 137 mM NaCl pH 7.4

2.4 Enzymes, Standards, and Kits

Product	Supplier
Bsal	NEB (Frankfurt, Germany)
CutSmart buffer	NEB (Frankfurt, Germany)
DNA ladder, 1 kb	NEB (Frankfurt, Germany)
DNAseI	Roche (Mannheim, Germany)
DpnI	NEB (Frankfurt, Germany)
Phusion HF DNA Polymerase	NEB (Frankfurt, Germany)
5x Phusion HF buffer	NEB (Frankfurt, Germany)
5x Phusion High GC buffer	NEB (Frankfurt, Germany)
Prestained protein standard	NEB (Frankfurt, Germany)
Protein test mixture 6	SERVA (Heidelberg, Germany)
Q5 HF DNA Polymerase	NEB (Frankfurt, Germany)
5x Q5 High GC enhancer	NEB (Frankfurt, Germany)
5x Q5 Reaction buffer	NEB (Frankfurt, Germany)
T4 DNA Ligase	NEB (Frankfurt, Germany)
10x T4 DNA Ligase buffer	NEB (Frankfurt, Germany)
T4 Polynucleotide kinase (PNK)	NEB (Frankfurt, Germany)
Wizard Plus SV Miniprep DNA purification system	Promega (Madison, USA)
Wizard Plus SV Gel and PCR Clean-up system	Promega (Madison, USA)
XbaI	NEB (Frankfurt, Germany)

2.5 Bacterial strains

Strain	Genotype	Supplier
E. coli BL21-CodonPlus (DE3)-RIL	<i>F</i> – <i>ompT hsdS</i> (<i>rB</i> – <i>mB</i> –) <i>dcm</i> + <i>Tetr gal</i> λ (<i>DE3</i>) <i>endA Hte [argU</i> <i>ileY leuW Camr</i>	Stratagene (LaJolla, USA)
E. coli SHuffle T7 Express B strain (C3029)	<i>fhuA2 lacZ::T7 gene1</i> [<i>lon</i>] <i>ompT</i> <i>ahpC gal</i> λ <i>att::pNEB3-r1-</i> <i>cDsbC</i> (<i>Spec</i> ^R , <i>lacI</i> ^q) Δ <i>trxB</i> <i>sulA11 R(mcr-73::miniTn10--</i> <i>Tet</i> ^S <i>)2 [dcm] R(zgb-210::Tn10 --</i> <i>Tet</i> ^S <i>) endA1</i> Δ <i>gor</i> Δ (<i>mcrC-</i> <i>mrr</i>)114:: <i>IS10</i>	NEB (Ipswich, USA)

Strain	Genotype	Supplier
E. coli SHuffle T7 Express K12 strain (C3026)	<i>F' lac, pro, lacI^q / Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)Pvull phoR ahpC* galE (or U) galK λatt::pNEB3-r1-cDsbC (Spec^R, lacI^q) ΔtrxB rpsL150(Str^R) Δgor Δ(malF)3</i>	NEB (Ipswich, USA)
E. coli XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F proAB lacIqZΔM15 Tn10 (Tetr)]</i>	Stratagene (LaJolla, USA)

2.6 Plasmids and constructs

Construct	Vector	Origin
FOR005 PT (V _L domain)	pET28b(+)	GeneArt (Regensburg, Germany)
FOR005 PT S31Y	pET28b(+)	This work
FOR005 PT F48Y	pET28b(+)	This work
FOR005 PT R49G	pET28b(+)	This work
FOR005 PT S51N	pET28b(+)	This work
FOR005 PT A94G	pET28b(+)	This work
FOR005 PT R49G/A94G	pET28b(+)	This work
FOR005 GL (V _L domain)	pET28b(+)	GeneArt (Regensburg, Germany)
FOR005 GL Y31S	pET28b(+)	Benedikt Weber
FOR005 GL Y48F	pET28b(+)	Benedikt Weber
FOR005 GL G49R	pET28b(+)	Benedikt Weber
FOR005 GL N51S	pET28b(+)	Benedikt Weber
FOR005 GL G94A	pET28b(+)	Benedikt Weber
FOR005 GL Y31S/G94A	pET28b(+)	This work
FOR005 GL Y48F/G94A	pET28b(+)	This work
FOR005 GL G49R/G94A	pET28b(+)	This work
FOR005 GL N51S/G94A	pET28b(+)	This work
FOR005 PT+linker	pET28b(+)	GeneArt (Regensburg, Germany)
FOR005 P-LC	pET28b(+)	GeneArt (Regensburg, Germany)
FOR005 P-LC	pE-SUMO	This work
FOR005 P-LC G136V	pET28b(+)	This work
FOR005 P-LC C214S	pET28b(+)	This work

Construct	Vector	Origin
FOR005 P-LC G136V/C214S	pET28b(+)	This work
FOR005 P-LC	pET-SUMO	GeneArt (Regensburg, Germany)
FOR005 G-LC	pET28b(+)	This work
FOR005 G-LC V136G	pET28b(+)	GeneArt (Regensburg, Germany)
FOR005 G-LC C214S	pET28b(+)	This work
FOR005 G-LC V136G/C214S	pET28b(+)	This work
FOR005 P-CL (C _L domain incl. linker)	pET28b(+)	GeneArt (Regensburg, Germany)
FOR005 G-CL (C _L domain G136V)	pET28b(+)	This work
FOR005 P-CLΔlinker	pET28b(+)	GeneArt (Regensburg, Germany)
Ulp-1 (Sumo protease)	pET28b(+)	Christina Nickels

2.7 Oligonucleotides

Template(s)	Name	Sequence (5' to 3')
FOR005 PT	S31Y fwd	gcgtagctattatgcaagctggt
	S31Y rev	aggctatcacccctgacag
	F48Y fwd	tctggtatttatcgtaaataatcgctccgagcgg
	F48Y rev	accggtgctgaccggt
	R49G fwd	ggttattttggcaaatctaatcgctccgagcgg
	R49G rev	agaaccggtgctgaccc
	S51N fwd	tttcgtaaaaacaatcgctccgagcgg
	S51N rev	ataaccagaaccggtgcc
	A94G fwd	aacacctgatggtgccgaactatcgcgac
	A94G rev	gtcgcgatagttcgggcaaccatcagggtgt
FOR005 GL G94A	Y31S/G94A fwd	gcgtagctattctgcaagctggt
	Y31S/G94A rev	aggctatcacccctgacag
	Y48F/G94A fwd	cgattgtttaccaaaaataaccagaaccggtgcctga
	Y48F/G94A rev	tcaggcaccggttctggtatttttgtaaaaacaatcg
	G49R/G94A fwd	ctcggacgattgttttacgataaataaccagaaccggtg
	G49R/G94A rev	caccggttctggtatttatcgtaaaaacaatcgctccgag
	N51S/G94A fwd	cgctcggacgattgctttaccataaataaccagaaccg
	N51S/G94A rev	cggttctggtatttatggtaaaagcaatcgctccgagcg
FOR005 P-LC	PLC_pESumo_Bsal fwd	gaggctctaggttagcagcgaactgac

Template(s)	Name	Sequence (5' to 3')
FOR005 P-LC	PLC_pESumo_Xbal rev	gctctagattagctacattcgggtcgggtgc
FOR005 P-LC	C214S fwd	ctcgaattcttagctactttcgggtcgggtgcaacgg
FOR005 G-LC	C214S rev	ccgttgaccgaccgaaagtagctaagaattcgag
FOR005 P-LC	G136V fwd	aagcaaccctggtgtgtctgattag
FOR005 G-LC V136G FOR005 P-CL	G136V rev	tattgctgcagttcttc

2.8 Devices and instruments

Device/instrument	Supplier
<u>Analytical Ultracentrifugation</u>	
XL-I with absorbance and interference optics	Beckman Coulter (Krefeld, Germany)
Optima AUC with absorbance and interference optics	Beckman Coulter (Krefeld, Germany)
<u>Centrifuges</u>	
Avanti J25	Beckman Coulter (Krefeld, Germany)
Avanti J26 XP	Beckman Coulter (Krefeld, Germany)
Optima MAX-E preparative ultracentrifuge	Beckman Coulter (Krefeld, Germany)
Rotina 420 R	Hettich (Tuttlingen, Germany)
Rotina 46 R	Hettich (Tuttlingen, Germany)
Tabletop centrifuge 5418	Eppendorf (Hamburg, Germany)
Tabletop centrifuge MIKRO 200R	Hettich (Tuttlingen, Germany)
Tabletop centrifuge UNIVERSAL 320 R	Hettich (Tuttlingen, Germany)
<u>Centrifugation rotors</u>	
JA-10	Beckman Coulter (Krefeld, Germany)
JA-25.50	Beckman Coulter (Krefeld, Germany)
JLA-16.250	Beckman Coulter (Krefeld, Germany)
TLA-45 (for prep. Ultracentrifuge)	Beckman Coulter (Krefeld, Germany)
<u>Chromatography columns</u>	
HiLoad Superdex 75 pg 16/60	GE Healthcare (Freiburg, Germany)
HiLoad Superdex 75 pg 26/60	GE Healthcare (Freiburg, Germany)
HisTrap Fast Flow (FF), 5 ml	GE Healthcare (Freiburg, Germany)
PD10 desalting column	GE Healthcare (Freiburg, Germany)

Device/instrument	Supplier
<u>Chromatography columns</u> Q-Sepharose Fast Flow Superdex 75 10/300 GL	GE Healthcare (Freiburg, Germany) GE Healthcare (Freiburg, Germany)
<u>Chromatography systems</u> ÄKTA Explorer ÄKTA FPLC ÄKTA Prime ÄKTA Pure ÄKTA Purifier Frac-900/950 fraction collectors Shimadzu Prominence HPLC system Superloops (0.5/2/10/50/150 ml)	GE Healthcare (Freiburg, Germany) GE Healthcare (Freiburg, Germany) GE Healthcare (Freiburg, Germany) GE Healthcare (Freiburg, Germany) GE Healthcare (Freiburg, Germany) GE Healthcare (Freiburg, Germany) Shimadzu (Munich, Germany) GE Healthcare (Freiburg, Germany)
<u>Differential scanning calorimeter</u> MicroCal PEAQ-DSC	Malvern Panalytical (Kassel, Germany)
<u>Fluorescence spectrophotometers</u> FluoroMax-4 FP-8500	Horiba Jobin Yvon (Edison, USA) Jasco (Großumstadt, Germany)
<u>Gel electrophoresis equipment</u> Hoefer Mighty Small II dual gel caster Pharmacia EPS 301/3500 power supplies	GE Healthcare (Freiburg, Germany) GE Healthcare (Freiburg, Germany)
<u>H/DX-MS equipment</u> ACQUITY M-Class UPLC ACQUITY UPLC BEH C18 column Enzymate BEH Pepsin column HTS PAL Leap robot Synapt G2-S ESI-TOF mass spectrometer	Waters Corp. (Milford, USA) Waters Corp. (Milford, USA) Waters Corp. (Milford, USA) Leap Technologies (Carborro, USA) Waters Corp. (Milford, USA)
<u>Mass spectrometers</u> LCQ Fleet ESI-ion trap Synapt XS ESI-TOF	Thermo Fisher (Waltham, USA) Waters Corp. (Milford, USA)
<u>Microplate readers</u> GENios Infinite M Nano	Tecan (Männedorf, Schweiz) Tecan (Männedorf, Schweiz)
<u>Microscopes</u> JEM-1400 Plus Transmission electron microscope	JEOL (Freising, Germany)

Device/instrument	Supplier
<u>Scales</u> BL 310 BP 121 S	Sartorius (Göttingen, Germany) Sartorius (Göttingen, Germany)
<u>Spectropolarimeters</u> Chirascan Plus J-1500	Applied Photophysics (Leatherhead, UK) Jasco (Großumstadt, Germany)
<u>Thermal cyclers</u> MJ Mini 48 well Primus 25	Bio-Rad (Munich, Germany) MWG (Ebersberg, Germany)
<u>Temperature-controlled incubator</u> Digital heat block Eppendorf-Thermomixer TB1 Thermoblock	VWR (Darmstadt, Germany) Eppendorf (Hamburg, Germany) Biometra (Göttingen, Germany)
<u>UV/VIS spectrophotometers</u> Nanodrop ND-1000 Nanodrop 2000 UltroSpec 1100 pro	Peqlab (Erlangen, Germany) Peqlab (Erlangen, Germany) Amersham Biosciences (Amersham, UK)
<u>Additional equipment</u> Autoclave Varioclav EP-Z Cell disruption apparatus Basic Z Homogenizer SilentCrusher M Icemaker Image Scanner III Incubator MARK II Abbé refractometer Membrane vacuum pump MR80 magnetic stirrer MR2000 magnetic stirrer MR3001 magnetic stirrer pH meter 538 MultiCal Vortex MS2 Water bath F6-K	H+P (Oberschleißheim, Germany) Constant Systems (Warwick, USA) Heidolph (Schwabach, Germany) Ziegler (Isernhagen, Germany) GE Healthcare (Freiburg, Germany) New Brunswick Scientific (Nürtingen, Germany) Leica (Wetzlar, Germany) Sartorius (Göttingen, Germany) Heidolph (Schwabach, Germany) Heidolph (Schwabach, Germany) Heidolph (Schwabach, Germany) WTW (Weilheim, Germany) IKA (Staufen, Germany) Haake (Karlsruhe, Germany)

2.9 Software, databases, and web-based tools

Name	Supplier/ URL
abYsis	http://www.abysis.org/abysis/
ALBase	https://wwwapp.bumc.bu.edu/BEDAC_ALBase/
AmylPred2	http://thalis.biol.uoa.gr/AMYL_PRED2/
BestSel	http://bestsel.elte.hu/index.php
Chromeleon	Thermo Fisher (Waltham, USA)
Citavi 6	https://www.citavi.com/de
Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/
Dichroweb	http://dichroweb.cryst.bbk.ac.uk/html/home.shtml
DynamX 3.0	Waters Corp. (Milford, USA)
ExPASy ProtParam	https://web.expasy.org/protparam/
Google Scholar	https://scholar.google.de/
ImageJ	NIH (Bethesda, USA)
IMGT V-QUEST	https://www.imgt.org/
Inkscape	https://inkscape.org/de/
MagTran	Amgen Inc. (Thousand Oaks, USA)
MetAmyl	http://metamyl.genouest.org/
NCBI IgBLAST	https://www.ncbi.nlm.nih.gov/igblast/
NEBaseChanger	https://nebasechanger.neb.com/
NEBioCalculator	https://nebiocalculator.neb.com/#!/ligation
NEB Tm calculator	https://tmcalculator.neb.com/#!/main
Origin	OriginLab Corp. (Northampton, USA)
Protein Data Bank (PDB)	https://www2.rcsb.org/
Protein Lynx Global Server PLGS 3.0.3	Waters Corp. (Milford, USA)
PubMed	https://pubmed.ncbi.nlm.nih.gov/
PyMOL	Schrödinger LLC (New York, USA)
Sedfit	Peter Schuck
Serial Cloner 2.5	Franck Perez
SWISS-MODEL	https://swissmodel.expasy.org
UniProt	https://www.uniprot.org/
VBase2	www.vbase2.org
LabScan 6.0	GE Healthcare (Freiburg, Germany)
Xcalibur	Thermo Fisher (Waltham, USA)
ZipperDB	https://services.mbi.ucla.edu/zipperdb/

2.10 Molecular biology methods

2.10.1 Amplification and modification of DNA (PCR and site-directed mutagenesis)

Polymerase chain reaction (PCR) was used for the amplification and site-directed mutagenesis of DNA, which allows the targeted introduction of point mutations into a protein of interest. Primers for site-directed mutagenesis were designed either with NEBaseChanger or with Agilent QuickChange and then purchased from Eurofins (Ebersberg, Germany). For the amplification, either Phusion DNA polymerase or Q5 DNA polymerase and the following protocols were used.

PCR pipetting scheme	
Template DNA (plasmid)	1 μ l (= 1-5 ng DNA)
5x Q5 Reaction buffer/ 5x Phusion HF or GC buffer	10 μ l
dNTPs (10 mM)	1 μ l
Forward primer (10 μ M)	2.5 μ l
Reverse primer (10 μ M)	2.5 μ l
5x Q5 High GC enhancer (optional)	(10 μ l)
DMSO (optional)	(1.5 μ l)
Q5 DNA polymerase/ Phusion DNA polymerase	0.5 μ l
Nuclease-free water	to 50 μ l

Thermocycler protocol for PCR			
Step	Temperature	Time	
Initial denaturation	98 °C	3-4 min	
Denaturation	98 °C	30 sec	25-35 cycles
Annealing	50–72 °C (determined with NEB Tm Calculator)	30 sec	
Elongation	72 °C	1 min/kb	
Final elongation	72 °C	5 min	
Hold	4-10 °C	∞	

Afterwards, the successful amplification of DNA was checked by agarose gel electrophoresis (described below). In order to digest remaining methylated template DNA and to circularize the modified PCR product, a so-called KLD reaction was set up.

KLD reaction component	Volume
PCR product	3.5 µl
Nuclease-free water	4 µl
T4 DNA ligase buffer	1 µl
T4 DNA ligase	0.5 µl
DpnI	0.5 µl
T4 Polynucleotide kinase (PNK)	0.5 µl
Σ	10 µl

The KLD reaction was carried out for one to two hours at room temperature. Afterwards, the complete reaction volume (10 µl) was transformed into 100 µl *E. coli* XL1 blue (see below) and plated on LB agar plates supplemented with an antibiotic (e.g., kanamycin). DNA isolation from bacterial clones and sequence verification by Sanger sequencing were performed as described below (section 2.10.5).

For subcloning into the pE-Sumo vector, the insert PLC3706 was amplified by PCR as described above. The entire PCR product was purified from an agarose gel using the Wizard Plus SV gel clean up kit (Promega). The subsequent restriction digest was performed overnight at 37 °C and gentle shaking using the following set up:

Vector digestion		Insert digestion	
pE-Sumo	2 µg plasmid DNA	Insert DNA	2 µg
CutSmart buffer	10 µl	CutSmart buffer	10 µl
Bsal restriction enzyme	3 µl	XbaI	3 µl
Nuclease-free water	fill up to 100 µl	Bsal	3 µl
		Nuclease-free water	fill up to 100 µl
Σ	100 µl	Σ	100 µl

The digested insert and vector DNA were purified via spin columns (Promega Wizard Plus SV gel clean up kit) and ligated overnight at room temperature using T4 DNA Ligase (NEB) following the manufacturer's protocol. Five reactions were set up, each with 50 ng of digested vector and varying vector:insert ratios (1:1/ 1:3/ 1:5/ 1:7). The amounts of digested insert DNA were calculated using NEBioCalculator. After the overnight ligation, the complete reaction volumes were transformed into *E. coli* XL1 blue for plasmid preparation and sequence control.

2.10.2 Agarose gel electrophoresis

Agarose gel electrophoresis was performed to isolate and analyze DNA fragments and PCR products. Therefore, 1 % agarose gels (w/v) were prepared in TAE buffer and supplemented with 0.003 % (v/v) Stain G (SERVA, Heidelberg) which allows visualization of DNA under UV light. DNA samples were mixed with 6x purple loading dye before loading onto the gel and FastLoad 1 kb DNA ladder (SERVA, Heidelberg) was used as a size standard. The gels were run for 20-30 minutes at 120 mV in TAE buffer.

2.10.3 Preparation of chemically competent bacterial cells

Chemically competent bacteria were prepared according to Sambrook et al. (Sambrook et al., 1989). Briefly, *E. coli* were grown in LB₀ medium under continuous shaking at 37 °C until an OD₆₀₀ of 0.5-1.0 was reached. The bacterial suspension was then supplemented with 2 ml of 1 M MgCl₂, incubated under shaking at 37 °C for another 10 min, and then cooled down on ice. Bacteria were harvested by centrifugation (5 minutes/ 5000 rpm) and then resuspended in cold solution A and incubated on ice for one hour. Subsequently, bacteria were harvested again by centrifugation and resuspended in 2 ml cold solution B. Aliquots were prepared, snap-frozen in liquid nitrogen, and stored at -80 °C.

2.10.4 Transformation of bacteria

Approximately 100 µl of frozen, chemically competent *E. coli* were thawed on ice and then incubated with DNA (50-100 ng) for 15-20 minutes on ice. A heat shock was performed for 45 seconds at 42 °C before the cells were cooled down on ice again for 5 minutes. Afterwards, 900 µl of LB₀ medium were added and the bacteria were incubated at 37 °C under shaking for one hour. The bacteria were centrifuged for 5 minutes at 5000 rpm, resuspended in 200 µl LB₀ medium, and plated on agar plates containing an appropriate antibiotic. The plates were incubated overnight at 37 °C.

2.10.5 Isolation of DNA

For isolation and analysis of DNA, single clones of transformed bacteria were picked from the agar plates and incubated overnight at 37 °C in 4 ml LB medium supplemented with antibiotic. Plasmid DNA was isolated using the Wizard Plus SV DNA purification kit (Promega, USA) following the manufacturer's protocol. Concentration and purity of isolated DNA was checked by absorbance at 260 nm using the Nanodrop ND-1000 (PepLab, Germany). Afterwards, 15 - 20 µl of isolated plasmid DNA were sent to Eurofins Genomics (Ebersberg, Germany) for Sanger sequencing with either T7 or T7 terminator primer.

2.11 Protein expression and purification

In order to obtain recombinant proteins, two general strategies were applied: either the protein was over expressed in *E. coli* BL21 (DE3) as insoluble inclusion bodies, or the protein was expressed in its native state using *E. coli* SHuffle.

2.11.1 Protein expression as insoluble inclusion bodies

For inclusion body expression, first, 50 ml pre-cultures of transformed *E. coli* BL21 (DE3) in LB_{Kan} were grown overnight at 37 °C under continuous shaking. The next day flasks with 2 L LB_{Kan} were inoculated with the entire pre-cultures. Bacteria were grown at 37 °C under shaking until an OD₆₀₀ of 0.7-1.0 was reached. Protein expression was started by addition of 1 mM IPTG and was then carried out overnight at 37 °C under shaking. The next day, bacteria were harvested by centrifugation for approximately 20 minutes at 7000-8000 rpm and 4-8 °C. The bacterial pellet was resuspended in 100-200 ml PBS pH 7.4 supplemented with protease inhibitor mix HP and DNaseI. Afterwards, the bacteria were lysed under 1.8 kbar pressure using a cell disruptor and the resulting lysate was centrifuged for 45-60 minutes at 10000 rpm at 4-8 °C to obtain the insoluble inclusion bodies as pellet. Inclusion bodies were frozen in liquid nitrogen and stored at -20 °C until purification.

2.11.2 Purification by refolding from inclusion bodies

The refolding of antibody domains from inclusion bodies was performed by two slightly different approaches.

V_L domains and LCs

Inclusion bodies of V_Ls and LCs were solubilized in 50 ml PBS containing 8 M urea and 0.1 % 2-mercaptoethanol at room temperature for 2-4 hours under stirring and afterwards centrifuged for 30-60 minutes at 20000 rpm and 4-8 °C to clear insoluble cell debris. The supernatant containing the dissolved, unfolded protein was dialysed overnight against a 20-fold excess of dialysis buffer A at approximately 10 °C under gentle stirring to get rid of the excess 2-mercaptoethanol. After dialysis, the protein solution was loaded onto a Q-sepharose column and purified by anion exchange chromatography (bulk elution with Q-sepharose buffer B). The fractions containing the protein of interest were identified by SDS-PAGE (see below) and pooled. Then, protein concentration was determined by absorbance at 280 nm using a NanoDrop device. Subsequently, the protein was diluted to approximately 0.5 mg/ml with the residual Q-sepharose buffer A and dialysed against a 20-fold excess of dialysis buffer B overnight at approximately 10 °C under gentle stirring. The next day, dialysis buffer B was

removed and diluted 1:2 with PBS pH 7.4 to obtain a dialysis containing 1 M urea, which was used for further dialysis for approximately 6-8 hours. Then a final dialysis step against PBS pH 7.4 was carried out at 10 °C overnight. Subsequently, the refolded protein was concentrated either using Amicon centrifugal filters or nitrogen-driven stirred concentrator cells before loading onto a Superdex75 column for size-exclusion chromatography. Afterwards, protein-containing fractions were identified by SDS-PAGE and the identity and purity of the protein was checked by SDS-PAGE and mass spectrometry (see below).

C_L domains and dimeric P-LC (Bence-Jones protein)

For C_L domains the inclusion bodies were dissolved as described above. However, instead of dialysis against dialysis buffer A, anion exchange chromatography with Q-sepharose was carried out directly. After this first purification step, refolding was performed by overnight dialysis against a 20-fold excess of refolding buffer at approximately 10 °C. In case of the P-LC Bence-Jones dimer, the refolding buffer contained 2 mM oxidized glutathione. Subsequently, dialysis against PBS pH 7.4, size-exclusion chromatography, and protein quality control were carried out as described above.

2.11.3 Native protein expression

Antibody domains

Native protein expression and purification was carried out using His-Sumo-tagged antibody domains in *E. coli* SHuffle B or K12 strains. Pre-cultures of 50 ml LB_{Kan} were inoculated with transformed bacteria and grown overnight at 37 °C under shaking. The next day, 2 L flasks of LB_{Kan} were inoculated with the complete pre-cultures and grown 37 °C to an OD₆₀₀ of 0.6-0.8, before protein expression was induced with 1 mM IPTG. Then, the bacteria were shifted to 20-25 °C and protein expression was carried out overnight under continuous shaking. The bacteria were subsequently harvested by centrifugation at 7000-8000 rpm and 4-8 °C for 20 minutes. The cell pellet was resuspended in approximately 40-50 ml PBS pH 7.4 supplemented with protease inhibitor mix HP, frozen in liquid nitrogen, and stored at -20 °C until purification.

Sumo-protease (Ulp1)

The sumo-specific protease Ulp1 was expressed natively in *E. coli* BL21 (DE3) similarly to the antibody domains described above. However, for Ulp1 the expression in BL21 cells was carried out for 4-5 hours at 37 °C before bacteria were harvested and stored.

2.11.4 Native protein purification

Antibody domains

The frozen *E. coli* SHuffle cells were thawed and lysed using a cell disruptor (pressure: 1.8 kbar) equilibrated with HisTrap A buffer and PMSF was added to the lysate to a final concentration of approximately 1 mM. The lysate was cleared by centrifugation at 20000 rpm and 4-8 °C for 45-60 minutes and the supernatant was loaded onto a HisTrap FF column (5 ml, equilibrated in HisTrap buffer A). Then the column was washed with 5-8 column volumes of buffer A before a gradient elution (0-100 % B over 70 ml) with buffer B was performed. The fractions were analyzed by SDS-PAGE, protein-containing elution fractions were pooled, and digested with Ulp1 (1-2 aliquots) overnight at 4-8 °C while performing dialysis against a 20-fold excess of PBS pH 7.4 under gentle stirring. The next day, the digested protein solution was loaded again onto the HisTrap column, a washing step with 5 column volumes buffer A, and gradient elution were performed. Fractions were again analyzed by SDS-PAGE and the protein-containing fractions (flow-through) were pooled and dialysed against a 20-fold excess of PBS pH 7.4 overnight at 4-8 °C under stirring. The following day, the protein was concentrated and size-exclusion chromatography with a Superdex75 column equilibrated in PBS pH 7.4 was performed as a polishing step.

Sumo protease (Ulp1)

For Ulp1, the cell disruption and the first purification step on a HisTrap FF column were carried out as described above. After analysis via SDS-PAGE, the elution fractions containing Ulp1 were pooled and concentrated to a volume of approximately 5-6 ml using an amicon centrifugal filter. Then, a buffer exchange into 2x Ulp1 storage buffer was performed using a PD10 desalting column. The PD10 column was equilibrated with 2x Ulp1 storage buffer, 2-3 ml of concentrated protein were loaded, and subsequently eluted with 3-4 ml of the buffer. This procedure was repeated, so that the entire purified protein was now in 2x Ulp1 storage buffer. The protein solution was diluted 1:1 with glycerol and supplemented with 1 % (v/v) of NP40. Aliquots of 200 - 500 µl were prepared, snap-frozen in liquid nitrogen, and stored at -80 °C.

2.12 Protein analytics

2.12.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was carried out to estimate protein size and purity of purification fractions. The gels for protein purification were self-made according to the following protocol:

	15 % separation gel	18 % separation gel	5 % stacking gel
Acrylamide/bisacrylamide (40 % w/v)	3.75 ml	4.5 ml	0.625 ml
2x stacking gel buffer	-	-	2.5 ml
4x separation gel buffer	2.5 ml	2.5 ml	-
10 % APS	100 μ l	100 μ l	100 μ l
TEMED	10 μ l	10 μ l	10 μ l
ddH ₂ O	3.75 ml	3 ml	1.875 ml
Σ	10 ml	10 ml	5 ml

Glass plates and ceramic plates with spacers were mounted in the dual gel caster and the freshly prepared separation gel solution was poured into the space between the plates. A thin layer of isopropanol was poured on top and the solution was allowed to polymerize. After polymerization was completed, the isopropanol was discarded, the freshly prepared stacking gel solution was poured in, and a comb was put in to generate sample wells. The protein samples were mixed 4:1 with 5x Laemmli buffer (\pm 2-mercaptoethanol depending on whether disulfide bridges should be reduced or oxidized), heated for 2-3 minutes at 95 °C and loaded into the sample wells. The gel chamber was filled with SDS running buffer and a constant current of 20-30 mA was applied for 40-60 minutes depending on the gel. As a size standard SERVA protein test mixture 6 or NEB Color Prestained Protein Standard (broad range) were used. The gels were stained with Fairbanks A for 5-10 minutes including a short heating step (30 seconds microwave). Afterwards, the gels were destained with Fairbanks D until clear bands were visible.

2.12.2 Liquid chromatography and mass spectrometry (LC-MS)

LC-MS was used for determining the exact molecular weight and purity of recombinant proteins. Full-length measurements were carried out on two mass spectrometers with electrospray ionization.

LCQ Fleet ESI-Ion trap

The LCQ Fleet has a heated electrospray ion source coupled to an ion trap mass analyzer with a standard m/z-range between 300 and 2000. Prior to ionization, samples (c = 0.1 mg/ml; injection volume = 1-5 µl) were applied to a Dionex Ultimate 3000 HPLC system (Thermo Fisher) with heated autosampler (15 °C) and a heated column compartment (T = 30 °C constant). For full-length measurements of proteins, either a desalting column or a monolithic column running in ddH₂O (+ 1 % formic acid) were used and an elution gradient from 5-100% acetonitrile was applied. After the measurement, data was evaluated with Xcalibur software (Thermo Fisher) and m/z-spectra were deconvoluted using MagTran (Amgen) to calculate the molecular weight.

Synapt XS ESI-TOF

The mass spectrometer is equipped with an electrospray ion source and a time-of-flight mass analyzer with an m/z-range of 400-4000. Samples (c = 0.1 mg/ml; injection volume = 2 µl) were applied to a C4 column attached to an ACQUITY HPLC system (Waters Corp.). An elution gradient of 0-95 % acetonitrile was used and MassLynx software (Waters Corp.) was used for data analysis.

2.13 Spectroscopy

2.13.1 UV/VIS spectroscopy

UV/VIS spectroscopy is based on the absorbance of light in the ultraviolet and visible range by molecules with excitable π-electrons, so called chromophores. In the case of DNA, the aromatic ring systems of the heterocyclic purin and pyrimidin moieties of the nucleobases are mainly responsible for the absorption of light (Cantor and Schimmel, 1980). In proteins, the π-electrons of the peptide bond and aromatic amino acid side chains (tryptophan, tyrosine, phenylalanine) absorb UV light. Absorbance of the peptide bond occurs between 190-220 nm, whereas the aromatic side chains absorb light mostly between 260-295 nm with tryptophan showing a strong absorption maximum at 280 nm (Stoscheck, 1990). These characteristics allow determining the concentration of DNA and protein solutions by absorption at 260 nm and 280 nm, respectively, according to the Lambert-Beer law

$$A(\lambda) = \varepsilon \cdot c \cdot d \quad (\text{Equation. 3})$$

where A is the absorbance at a specific wavelength λ, ε is the calculated molar extinction coefficient (M⁻¹cm⁻¹), c is the concentration (M), and d is the pathlength (cm).

Absorbance measurements were carried out on a NanoDrop spectrophotometer with baseline correction. ExPASy ProtParam was used to calculate molar extinction coefficients.

2.13.2 Circular dichroism spectroscopy (CD)

Light can be circularly polarized in either clockwise or counterclockwise manner by passing linearly polarized light through a modulator quartz crystal that is exposed to an alternating electrical field. The resulting light will switch between clockwise and counterclockwise circular polarization in the same rate at which the electrical field in the quartz crystal changes. Chiral chromophores or chromophores in an asymmetric environment absorb these two kinds of circularly polarized light to different extents (Kelly and Price, 2000). This phenomenon is called circular dichroism (CD) and it can be used for the structural characterization of proteins since the α -carbon atom in amino acids (other than glycine) is a chiral center (Woody, 1995; Kelly et al., 2005). Additionally, non-chiral chromophores that are surrounded by an asymmetric chemical environment, such as aromatic amino acids embedded in a protein core, will also absorb the two kinds of circularly polarized light to different extents. Therefore, CD can report on the secondary and tertiary structure of proteins. Signals in the far UV region (FUV; 180-250 nm) arise from $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions of the peptide bonds and give information about secondary structure elements, whereas signals in the near UV region (NUV; 250-320 nm) come mostly from aromatic amino acids and contribute to a so-called tertiary structure fingerprint (Kelly and Price, 2000). Further, CD spectroscopy can be used to determine the fold stability and melting temperatures of proteins e.g., by gradually heating the sample using a peltier element and following the signal at specific wavelengths (Woody, 1995; Kelly et al., 2005).

CD experiments were carried out on a J-1500 spectropolarimeter (JASCO, Großumstadt, Germany) or on a Chirascan Plus spectropolarimeter (Applied Photophysics, Leatherhead, UK) both equipped with a peltier element and water bath for temperature control. For FUV and thermal transition experiments, 300 μ l of protein solution with a concentration of 0.1 mg/ml in PBS were used in a 1 mm quartz cuvette (Hellma, Germany). NUV CD spectra were obtained in a 2 mm quartz cuvette filled with 450-500 μ l of protein solution at a concentration of 0.5 mg/ml. CD spectra were recorded at 20 °C with FUV spectra ranging from 200-260 nm and NUV spectra ranging from 250-320 nm. Thermal transitions were recorded at a wavelength of 205 nm in a temperature range between 20-90 °C with an applied heating rate of 1 °C per minute.

For data analysis, first the buffer spectra were subtracted from the protein CD spectra and then the mean residue weight ellipticity was calculated according to equation 4

$$\Theta_{MRW} = \frac{\Theta \cdot 100 \cdot M}{d \cdot c \cdot N_{aa}} \quad (\text{Equation. 4})$$

where Θ_{MRW} is the mean residue weight ellipticity ($\text{deg cm}^2 \text{dmol}^{-1}$), Θ is the measured ellipticity (mdeg), M is the molecular mass (g/mol), d is the pathlength (cm), c is the concentration (M), and N_{aa} is the number of amino acids in the protein. To assess the relative secondary structure contents of the obtained protein spectra, web-based tools such as Dichroweb and BeStSel can be used (Micsonai et al., 2022; Miles et al., 2022). For thermal unfolding analysis, the measured ellipticity at 205 nm was plotted against the respective temperature and the data was fitted applying a Boltzmann fit equation

$$y = \frac{A_1 - A_2}{1 + \exp((x - x_0)/dx)} + A_2 \quad (\text{Equation. 5})$$

The transition midpoint of the fit equation represents the apparent melting temperature T_m of the protein.

2.13.3 Fluorescence spectroscopy

2.13.3.1 Intrinsic tryptophan fluorescence

When a molecule with delocalized π -electrons absorbs light of a certain wavelength, the energy of the light can be transferred to the molecule so that it reaches a so-called excited state (S_1). When the electrons descend from the excited state back to their original energy level (S_0), the energy is emitted again as light of a greater wavelength, so-called fluorescent light (Jablonski, 1933). In proteins, the aromatic amino acids tryptophan, tyrosine, and phenylalanine represent natural fluorophores with tryptophan being the most important one since the effect is most pronounced for this amino acid. Tryptophan has an absorbance maximum at 280 nm with a local maximum (peak shoulder) at around 295 nm and it shows an emission maximum at around 350 nm (Lakowicz, 2006). Because tryptophan fluorescence is strongly dependent on the chemical surroundings of the indole ring of the side chain, it can be utilized to monitor protein folding or unfolding. Antibody light chains contain several highly conserved tryptophan residues that are mostly buried in the hydrophobic protein core when the protein is properly folded. This makes them ideal reporter fluorophores for the investigation of antibody folding and stability.

2.13.3.2 Chemical unfolding transitions

Antibody LCs contain a highly conserved tryptophan residue at position 35 of the V_L that is buried in the hydrophobic protein core in the native state and whose fluorescence emission is quenched by the adjacent disulfide bridge. When the protein is unfolded by a denaturant such as urea, tryptophan fluorescence exhibits a red shift and intensity increase due to the change of chemical environment (non-polar to polar environment) and the lost quenching effect of the disulfide bridge (Goto et al., 1979; Goto and Hamaguchi, 1982; Feige et al., 2010). By incubating a protein with increasing concentrations of urea and determining the change of tryptophan fluorescence as a function of denaturant concentration it is possible to derive the thermodynamic stability of a protein, that is the free enthalpy of the unfolding process.

Antibody domains at a concentration of 1 μM in PBS were incubated with increasing concentrations of urea at room temperature overnight in triplicates. The samples were then transferred into the wells of a 96-well UV-Star plate (Greiner Bio-One) and fluorescence spectra of the samples were recorded at 25 °C in a Tecan Infinite M Nano plate reader. The excitation wavelength was 280 nm and emission spectra were obtained between 300-400 nm. For data analysis, the spectrum of the native protein (0 M urea) was subtracted from the spectrum of the completely unfolded protein (6 M urea) and the maximum wavelength of the resulting differential spectrum was chosen for further data analysis since at this wavelength the largest difference between the folded and unfolded state of the protein are observed. Subsequently, the fluorescence intensities at this specific wavelength are averaged and plotted against the respective urea concentration and a sigmoidal fit function according to a two-state model of protein folding (Pace, 1986; Santoro and Bolen, 1988) was applied

$$y = y_N^0 + m_N \cdot [D] - \frac{y_N^0 + m_N \cdot [D] - (y_U^0 + m_U \cdot [D])}{1 + \exp\left\{-\frac{\Delta G_u^0 + m \cdot [D]}{R \cdot T}\right\}} \quad (\text{Equation. 6})$$

Where y_N^0 and m_N^0 determine the baseline equation for the native protei state, y_U^0 and m_U^0 determine the linear equation for the unfolded state, $[D]$ is the denaturant concentration (in this case urea), R is the gas constant, T is the absolute temperature, m is the cooperativity, and ΔG_u^0 is the Gibbs free energy of unfolding. The parameters y_N and y_U were kept at a constant value of 1 for the fit. Further a Boltzmann fit function (see section 2.13.2) was applied to determine the transition midpoint concentration C_m which represents the chemical stability.

Tryptophan fluorescence spectroscopy was also used to assess whether the antibody domains unfold reversibly or irreversibly. Protein at a concentration of 1 μM in PBS was incubated with 0.6 M urea at room temperature overnight and the same protein at a concentration 10 μM in PBS was unfolded with 6 M urea for two hours at room temperature. The unfolded protein

solution was then diluted 1:9 with PBS to obtain a protein solution containing 1 μM refolded protein and 0.6 M urea. If the tryptophan fluorescence spectra of both samples are identical the unfolding of the protein is considered to be reversible (Morgan and Kelly, 2016). The fluorescence spectra of the native and the refolded proteins at 0.6 M urea were obtained at 25 $^{\circ}\text{C}$ on a FluoroMax-4 spectrofluorometer (Horiba Jobin Yvon) to determine the reversibility of protein unfolding.

2.13.3.3 Acrylamide quenching

Fluorescence quenching describes the energy transfer of excited fluorophores onto another molecule (quencher) which results in a decrease in measured fluorescence intensity. In proteins, tryptophan fluorescence quenching (e.g., by acrylamide or iodide) can be used to obtain information about conformational differences and dynamics in proteins because the quenching efficiency depends on how accessible the tryptophan fluorophores are to the quenching agent (Eftink and Ghiron, 1981; Lakowicz, 2006). The antibody domains were incubated at a concentration of 15 μM in PBS pH 7.4 with increasing concentrations of acrylamide (0-1 M in steps of 0.1 M) overnight in the dark at room temperature. The samples were then transferred into a 96-well UV-Star plate (Greiner Bio-One) and fluorescence spectra were recorded in triplicates on a Tecan Infinite M Nano plate reader at 25 $^{\circ}\text{C}$. The excitation wavelength was 292 nm and emission spectra were obtained between 315 and 380 nm. The emission maximum wavelength of the native state (without acrylamide) was determined and the quenching rates F_0/F were calculated, where F_0 is the fluorescence intensity of the native state (without quenching agent) and F is the intensity at a specific acrylamide concentration. Subsequently, F_0/F was plotted against the quenching agent concentration and a linear fit function was applied to the data points. The slope of the linear fit represents the Stern-Volmer constant k_{sv} of quenching.

2.13.3.4 Red edge excitation shift (REES)

In REES experiments a 3D excitation-emission scan is performed which can ultimately give information on protein folding and dynamics by reporting on changes of the local chemical environments of intrinsic or extrinsic fluorophores (Lakowicz, 2006; Catici et al., 2016). Buffer controls and samples containing 20 μM protein in PBS pH 7.4 were equilibrated at 25 $^{\circ}\text{C}$ and 37 $^{\circ}\text{C}$ for approximately three hours before the measurements. Denatured samples also contained 6 M urea. The 3D scans were then recorded in duplicates at 25 $^{\circ}\text{C}$ and 37 $^{\circ}\text{C}$ on a JASCO FP-8500 spectrofluorimeter equipped with a peltier element for temperature control. Excitation wavelengths ranged from 280-300 nm and emission spectra were recorded between 315-400 nm. For data analysis, the duplicate spectra were buffer subtracted, averaged, and the center of spectral mass (CSM) was calculated according to equation (7)

$$\text{CSM} = \frac{\sum(f_i * \lambda_{em})}{\sum(f_i)} \quad (\text{Equ. 7})$$

where f_i is the fluorescence intensity and λ_{em} represents the emission wavelength. The CSM values were then plotted against the respective excitation wavelengths. The CSM starting value gives information about fluorophore solvent exposure, whereas the difference between the first and last CSM value correlates with the dynamics of the fluorophore environment.

2.13.3.5 Anilinonaphthalene sulfonate (ANS) binding

The fluorescent dye ANS is able to bind hydrophobic segments in proteins upon which its fluorescence spectrum exhibits a blue-shift and an intensity increase (Stryer, 1965). Therefore, this dye is often used in protein folding studies (Yu and Strobel, 1996). Buffer controls and samples with 10 μM protein were incubated with 100 μM ANS in PBS pH 7.4 for approximately 2 hours at room temperature in the dark. Then, the samples were transferred into a UV-Star 96-well plate and ANS emission spectra were recorded on a Tecan Infinite M Nano plate reader at 25 °C. The excitation wavelength was 380 nm and emission spectra were obtained between 400-600 nm.

2.13.3.6 Thioflavin T (ThT) binding kinetics

The benzothiazole-based dye thioflavin T (ThT) binds to the cross β structures of amyloid fibrils which leads to an increase in its fluorescence intensity and thereby enables monitoring the kinetics of amyloid formation (Gade Malmos et al., 2017). For ThT assays, aggregates were first removed from the thawed protein aliquots by preparative ultracentrifugation for 3-4 hours (40000-45000 rpm/4-8 °C) using an Optima MAX-E ultracentrifuge with a TLA-45 rotor (Beckman-Coulter). Subsequently, samples containing 15 μM protein, 7.5 μM ThT, 0.05 % (v/v) sodium azide were prepared. The buffer was either PBS pH 7.4, PBS pH 6.4, or PBS pH 7.4 including 0.5 mM SDS. The samples (200 μl) were transferred into a black 96-well PP plate which was covered with an adhesive, transparent sealing foil. The measurements were carried out in triplicates over a period of two weeks in a Tecan GENios plate reader at a constant temperature of 37 °C. Continuous orbital shaking with high intensity was applied to accelerate the fibril formation process. The excitation wavelength was 440 nm and fluorescence emission was measured at 480 nm.

2.14 Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) enables the investigation of protein oligomerization and complex formation (Laue, 2001). The sedimentation velocity (SV) experiments were carried out on either an Optima or ProteomLab XL-I device (Beckman-Coulter) and protein concentrations for AUC experiments were adjusted according to the absorbance value at 280

nm. The samples (350 μ l) were transferred into assembled cells with quartz windows and 12 mm path length charcoal filled epon double sector center pieces. Measurements were performed at 20 °C and 42000 rpm using an AN50-ti rotor (Beckman-Coulter). Absorbance was measured at 280 nm or 295 nm for higher protein concentrations and data analysis was performed using SEDFIT (Schuck, 2000; Brown and Schuck, 2006). For the analyses, a continuous c(S) distribution model was applied and fitting parameters were as follows:

resolution:	200
S _{min} :	0
S _{max} :	10-15
frictional ratio:	1.20
confidence interval:	0.95
spatial resolution:	0.73
buffer density:	1.0056
buffer viscosity:	1.0020

AUC measurements were performed with kind help of Ramona Absmeier.

2.15 Limited Proteolysis

In limited proteolysis experiments, the protein of interest is incubated with a protease and the kinetics of proteolytic cleavage or degradation are followed by SDS-PAGE and quantified densitometrically. These experiments give information about protein stability, dynamics, and general susceptibility to cleavage by proteases (Fontana et al., 1999). The substrate protein was used at a concentration of 0.3 mg/ml diluted in limited proteolysis buffer (100 mM Tris, 100 mM NaCl, 10 mM CaCl₂, pH 7.8). The protease trypsin was applied at a final concentration of 0.02 mg/ml for V_L domains (substrate/enzyme ratio = 15/1) and 0.015 mg/ml for LCs (substrate/enzyme ratio = 20/1), whereas proteinase K was used at a final concentration of 0.002 mg/ml for V_{LS} (substrate/enzyme ratio = 150 /1) and 0.0012 mg/ml for LCs (substrate/enzyme ratio = 250/1). The experiments were carried out at room temperature over 90 minutes and at defined time points samples were taken from the reaction and cleavage was stopped by diluting the sample in a solution containing 5x Laemmli buffer and 2 mM PMSF. The samples were subsequently analyzed on an SDS gel (SERVA TG Prime 4-20 %) for 80-90 minutes applying a constant current of 25 mA. Afterwards, gels were stained with Quick Coomassie Stain (SERVA) and gel band intensity was quantified using NIH ImageJ.

2.16 Differential Scanning Calorimetry (DSC)

DSC is a direct method to assess the stability of proteins by thermal unfolding (Bruylants et al., 2005). The calorimeter contains two capillary chambers one of which is loaded with the protein sample and the other one with the sample buffer. Both chambers are kept at the same temperature while a defined heating gradient is applied, however, due to the unfolding enthalpy of the proteins, the two chambers need to be supplied with different amounts of energy or electrical power, respectively, to maintain the same temperature. This difference in power is monitored and used to calculate the change in heat capacity (C_p) depending on the temperature from which the melting temperature (T_m) and unfolding enthalpy (ΔH) can be determined (Bruylants et al., 2005).

The measurements were performed in duplicates on a MicroCal PEAQ DSC (Malvern Panalytical) with protein samples of 1 mg/ml in PBS pH 7.4. Thermograms between 20-90 °C were obtained with a scan rate of 1 °C/min and a pre-scan thermostat of 5 minutes was applied. Before samples were measured, the device was equilibrated with water and PBS pH 7.4 to ensure baseline stability and reproducibility. After every sample measurement, the chambers were cleaned with 10 % Decon90 (v/v), rinsed with water, and a buffer measurement was performed to ensure that the chambers were clean. Data analysis was carried out using the PEAQ DSC software (Malvern Panalytical). The buffer baseline was subtracted with a progressive baseline fit and a non-two state model was applied to fit the thermogram data.

2.17 Transmission electron microscopy (TEM)

TEM measurements were performed to check for amyloid fibrils or amorphous protein aggregates after a completed ThT assay. To prepare TEM samples, 10 µl of sample solution were taken from the microplate wells of the ThT assay and were applied to a 200-mesh activated copper grid. After incubation for 1 minute, the grids were washed with 2x 10 µl ddH₂O and subsequently stained with 8 µl uranyl acetate solution (1.5 % w/v) for 1 minute. All excess solutions were removed from the grid with a filter paper. TEM micrographs were recorded on a JEOL JEM-1400 Plus transmission electron microscope at 120 kV (JEOL, Freising, Germany). TEM measurements were performed with kind help of Ramona Absmeier and Dr. Carsten Peters.

2.18 Hydrogen/deuterium exchange coupled to mass spectrometry (HDX-MS)

HDX-MS is a powerful technique to study the structure, dynamics, and interactions of proteins (Masson et al., 2017). Depending on their accessibility, the amide protons of the peptide backbone can readily exchange with protons of the surrounding water. The exchange rate of amide protons rises with conformational flexibility and with structural disorder (Englander, 2006; Mazur et al., 2017). In HDX experiments, the protein of interest is incubated in a buffer containing deuterium oxide (D_2O), so the exchange of amide protons with deuterium changes the mass of the protein or a peptide, respectively. This peptide mass change is detected via high resolution mass spectrometry and allows the quantitative assessment of protein dynamics. The HDX-MS experiments were carried out on a fully automated system comprising a Leap robot (HTS PAL, Leap Technologies, USA), a Waters ACQUITY M-Class UPLC (Waters, USA), an HDX manager (Waters, USA), and a Synapt G2 mass spectrometer (Water, USA). Protein samples with a concentration of 30 μM were diluted 1:20 in PBS buffer pH 7.4 containing D_2 and incubated for 10 seconds, 1 minute, 10 minutes, 30 minutes, and 2 hours. The exchange reaction was stopped by diluting the protein 1:1 in quenching buffer (200 mM $Na_2HPO_4 \times 2 H_2O$, 200 mM $NaH_2PO_4 \times 2 H_2O$, 250 mM TCEP, 3 M GdmCl, pH 2.2) at 1 °C. The labeled protein was then proteolytically digested online using an immobilized Waters Enzymate BEH Pepsin column (2.1 x 30 mm) at 20 °C. Subsequently, peptide separation was carried out on a Waters ACQUITY UPLC BEH C18 column (1.7 μm , 1.0 x 100 mm) by applying a water/acetonitrile gradient with both eluents containing 0.1 % formic acid (v/v). This step was performed at 0 °C to avoid back exchange. Eluting peptides were subjected to electrospray ionization and the time-of-flight (TOF) mass analyzer. An additional peptide separation by drift time was performed prior to fragmentation and mass detection. For data analysis, the Waters Protein Lynx Global Server PLGs (version 3.0.3) and the DynamX software package (version 3.0) were used.

2.19 Molecular dynamics (MD) simulations

All MD simulations were performed by Dr. Manuel Hitzenberger and Prof. Dr. Martin Zacharias from the chair of theoretical biophysics at TUM.

The MD simulations were performed and analyzed using the Amber18 simulation package (Case et al., 2018). For analysis of the V_L domain, simulations were started from the FOR005-PT V_L variant for which a crystal structure is available (PDB: 5L6Q) and on the *in silico* generated variants with the R49G und A94G substitutions, the FOR005-GL (wild type sequence), the GL G49R/G94A and the GL N51S/G94A sequence variants. All proteins were

solvated in TIP3P water in a periodic octahedral box with a minimum distance of 10 Å between protein atoms and the box boundary (Jorgensen et al., 1983). The ff14SB force field (Maier et al., 2015) was applied and Na⁺ and Cl⁻ ions were added to neutralize the system reaching an ion concentration of 0.15 M. Energy minimization of each system was carried out using the sander module of Amber18 (2500 minimization cycles). The system temperature was increased in steps of 100 K (50 ps per step) to a final temperature of 310 K with the solute non-hydrogen atoms harmonically restraint to the start structure, keeping all bonds involving hydrogen atoms at optimal length. Stepwise removal of the harmonic restraints was performed in four additional steps. For the subsequent production simulations, hydrogen mass repartitioning (HMR) was employed allowing a time step of 4 fs (instead of 2 fs used during heating and equilibration). Unrestrained production simulations were extended to 1 μs for each system and the coordinates were saved every 8 ps. Using the cpptraj module of Amber18, root mean square deviation (RMSD), root mean square fluctuations (RMSF), and analysis of dihedral angle distributions were performed.

To analyze the influence of the C_L mutation, MD simulations were performed starting from the P-CL (G136) domain (residues 110-212 of the P-LC structure) and the G-CL (V136) structure. The simulations were carried out exactly as described above for the V_L domains, except that unrestrained production simulations were extended to 1.6 μs. Additionally, umbrella sampling (US) simulations were performed using the center of mass distance between backbone atoms of residues 135-137 (including either the V136 or G136 residue) and residues 120-121 which belong to the N-terminal β-strand of the C_L domain. The reference distance in the quadratic restraining potential (force constant 2.5 kcal mol⁻¹Å⁻²) varied from 4.5 – 9.0 Å in steps of 0.5 Å. Sampling in each US interval was performed using 5 ns equilibration followed by 30 ns data gathering. The associated potential of mean force (PMF), i.e., the free energy change along the reaction coordinate, was calculated using the weighted-histogram analysis method (WHAM).

3. Results and Discussion

3.1 Amyloid formation of the FOR005 variable domain (V_L) is mediated by patient-specific mutations in hypervariable CDR loops

3.1.1 Sequence differences between FOR005-PT and FOR005-GL

The FOR005 project was based on a collaboration of our lab with research groups at the University Clinic Heidelberg and the University Ulm. Our collaborators derived patient material from a 50-year-old, female patient with heart failure due to cardiac light chain amyloidosis and reported the cDNA sequence and crystal structure (PDB: 5L6Q) of the amyloid forming V_L domain (FOR005-PT) in 2017 (Annamalai et al., 2017). We then determined the corresponding germline V_L sequence (FOR005-GL) with the highest possible sequence identity by using IgBLAST, IMGT, and abYsis (Brochet et al., 2008; Lefranc et al., 2009; Ye et al., 2013; Swindells et al., 2017). Genetically, the FOR005 sequence belongs to the λ3I light chain subfamily (gene segments: IGLV3-19/IGLJ2). Of note, the related germline V segment λ3r has been reported to be frequently associated with AL amyloidosis (Perfetti et al., 2002; Villalba et al., 2015). By sequence comparison, five point mutations were identified in the patient-derived V_L domain (Y31S, Y48F, G49R, N51S, G94A) with respect to the germline sequence (Fig. 10). For definition of the CDRs, the Kabat and Chothia numbering systems were applied (Wu and Kabat, 1970; Chothia and Lesk, 1987). Four of the mutations are located in hypervariable CDRs (Y31S in CDR1, G49R and N51S in CDR2, G94A in CDR3) and only one mutation is in a conserved framework region (Y48F in FR2).

Residue conservation analysis performed with ConSurf showed the lowest degree of conservation (ConSurf score = 1) for all four CDR mutations and an average conservation degree (ConSurf score = 5) for the framework residue F48 (Fig. 11) (Ashkenazy et al., 2010). The web-based prediction tools AmylPred2, MetAmyl, and ZipperDB (Goldschmidt et al., 2010; Emily et al., 2013; Tsohis et al., 2013) were used to assess aggregation-prone regions and individual mutational effects in the V_L domain. Yet, these predictions did not suggest significant alterations of the amyloid aggregation or steric zipper propensity of the patient V_L sequence in comparison to its corresponding germline sequence (Fig. 10A). To visualize structural effects of the point mutations in more detail, a homology model of FOR005-GL based on the template structure 5BV7 was created using the SWISS-MODEL web server (Waterhouse et al., 2018). The template V_L and the FOR005-PT V_L domain share 98.2 % sequence identity. By aligning the homology model with the crystal structure of FOR005-PT we observed that the overall structure was conserved, although the conformations of the CDR2 and CDR3 loops are slightly

changed (Fig. 10B). However, a homology model is only an approximate prediction of the actual native protein structure, which needs to be considered regarding data interpretation.

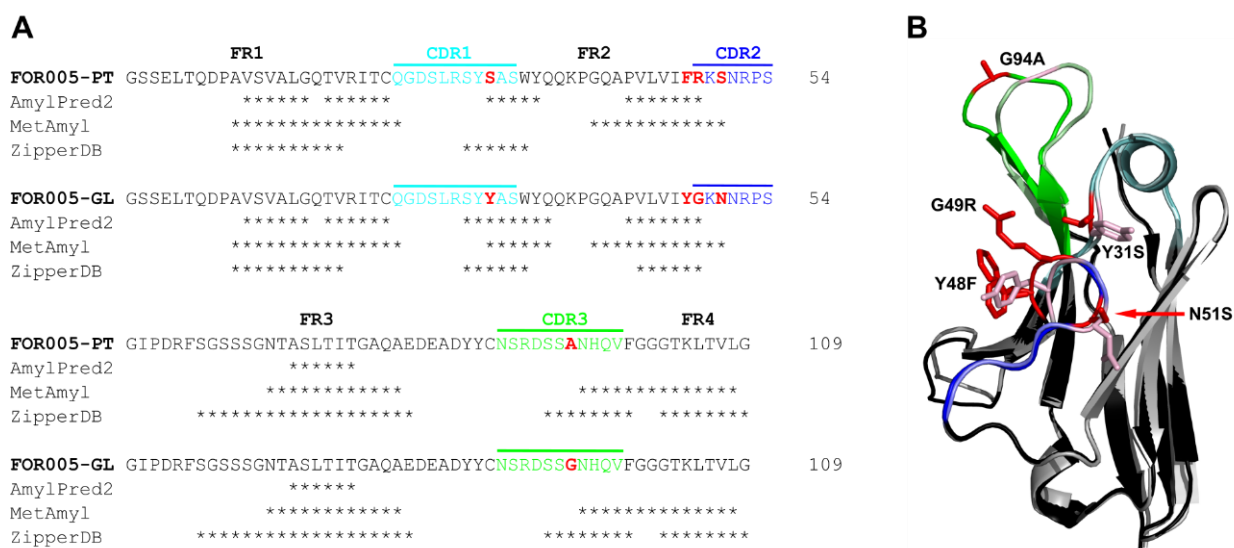


Figure 10: Sequence alignment, aggregation prediction, and structural comparison of FOR005-PT and FOR005-GL. A) Sequence alignment of patient and germline V_L with the five point mutations shown in red and the variable CDR loops colored in cyan (CDR1), blue (CDR2), and green (CDR3). Predicted amyloidogenic regions (AmylPred2, MetAmyl, ZipperDB) are indicated by asterisk. According to the predictions, the substitutions do not introduce new amyloid driving segments. The sequence numbering as derived from Annamalai et al. starts with the first serine residue, Ser1. The N-terminal glycine is numbered as Gly0 in our sequence. B) Structural alignment of FOR005-PT shown in black (PDB: 5L6Q) with the homology model derived for FOR005-GL depicted in grey. The homology model was created using the SWISS-MODEL server and the template structure 5BV7. CDRs are colored according to the sequence alignment in A). Mutations are shown in red on the patient V_L and light red on the germline V_L domain with side chains depicted as sticks. For F48 on the patient structure two rotamers are shown. Adapted from Rottenaicher et al., 2021.

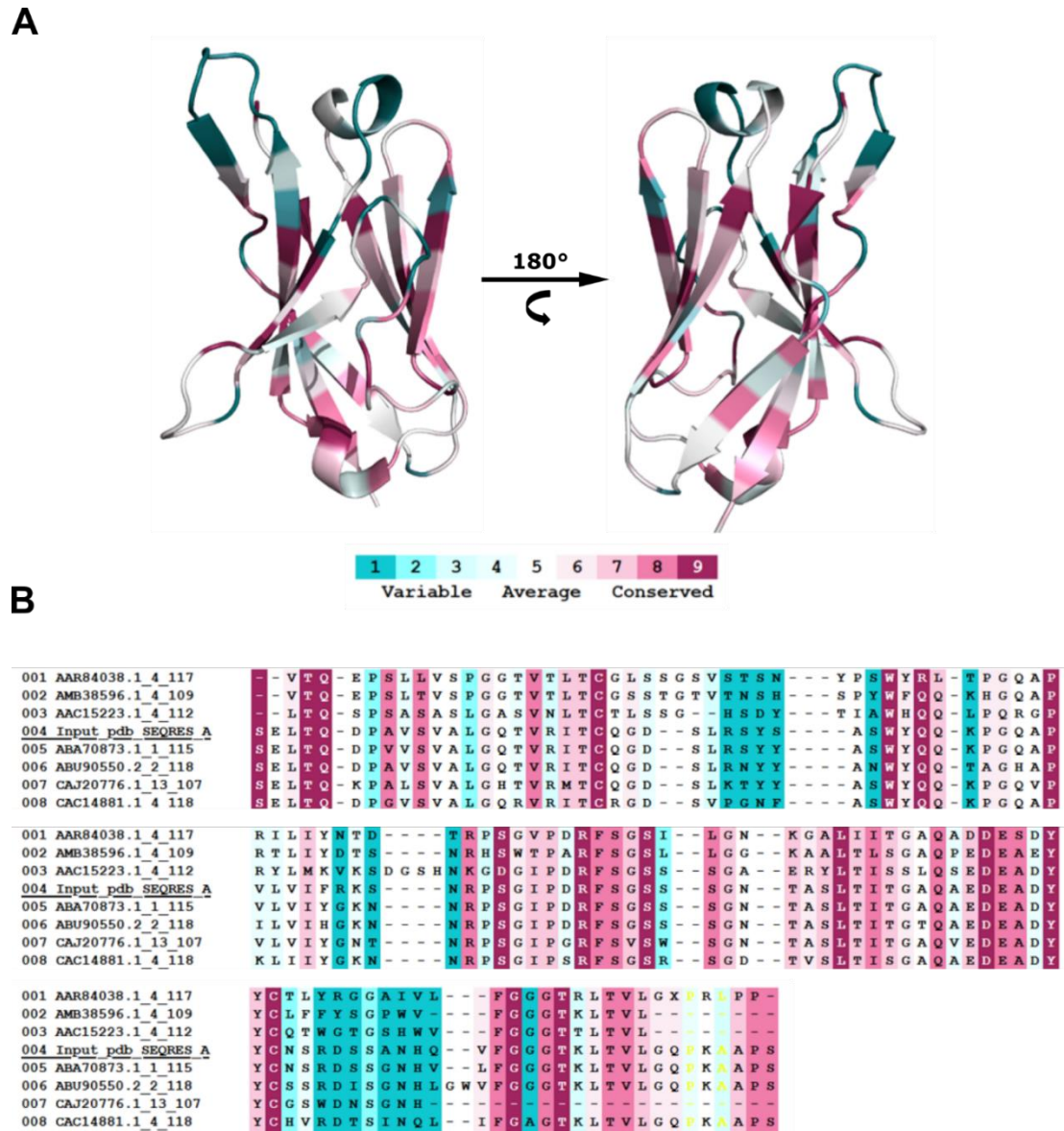


Figure 11: Residue conservation in FOR005-PT analyzed with ConSurf. A) Color-coded conservation scores derived from ConSurf were plotted onto the crystal structure of FOR005-PT. Highly variable residues with a conservation score of 1 are shown in dark cyan, highly conserved positions with a score of 9 are shown in purple. B) Outtake of the ConSurf multiple sequence alignment comprising eight sequences including the sequence of FOR005-PT (“004_Input_pdb_SEQRES_A”). Adapted from Rottenaicher et al., 2021.

3.1.2 FOR005-PT and FOR005-GL differ in their biophysical properties and amyloid propensity

The V_L domains FOR005-PT and FOR005-GL were recombinantly produced in *E. coli* as insoluble inclusion bodies. After refolding, purification, and quality control by mass spectrometry and SDS-PAGE, the biochemical and biophysical characterization was carried out. For secondary structure analysis, far-UV CD spectra of the purified proteins were obtained. In the far-UV region both V_L domains exhibit the typical β -sheet-rich immunoglobulin fold as indicated by the minimum at around 218 nm (Woody, 1995) although FOR005-GL

shows slightly higher secondary structure content as can be seen by the increased signal around 200 nm (Fig. 12D). The tertiary structure was probed by near-UV CD spectroscopy which yields a protein-specific structural “fingerprint” (Kelly et al., 2005). The two V_L domains exhibit highly similar spectra indicating that the overall topology is hardly altered by the mutations and the structural differences seemingly affect mainly local secondary structural elements (Fig. 12A). Additionally, analytical ultracentrifugation (AUC) was performed to assess the quaternary structure. Sedimentation coefficients of 1.52 S and 1.59 S, respectively, show that both the patient and germline V_L domains are monomeric in solution (Fig. 12B). To gain further insight into the effects of the mutations on the folding pathway, the reversibility of urea-induced unfolding was assessed by fluorescence spectroscopy. Neither of the V_L domains can be completely refolded into their native structure within 24 hours at room temperature as shown by the differences of the fluorescence spectra of refolded and native proteins (Fig. 12C). Nonetheless, the germline V_L exhibits a higher degree of unfolding reversibility than the patient variant. Thus, one might speculate that FOR005-GL could refold entirely into its native structure under different experimental conditions, e.g., longer refolding time (Fig. 12C). Further, refolding after thermal denaturation was analyzed by far-UV CD spectroscopy. Similarly to the fluorescence data, incomplete refolding was observed for both domains, although FOR005-GL shows slightly higher ability to regain its native conformation after denaturation (Fig. 12D).

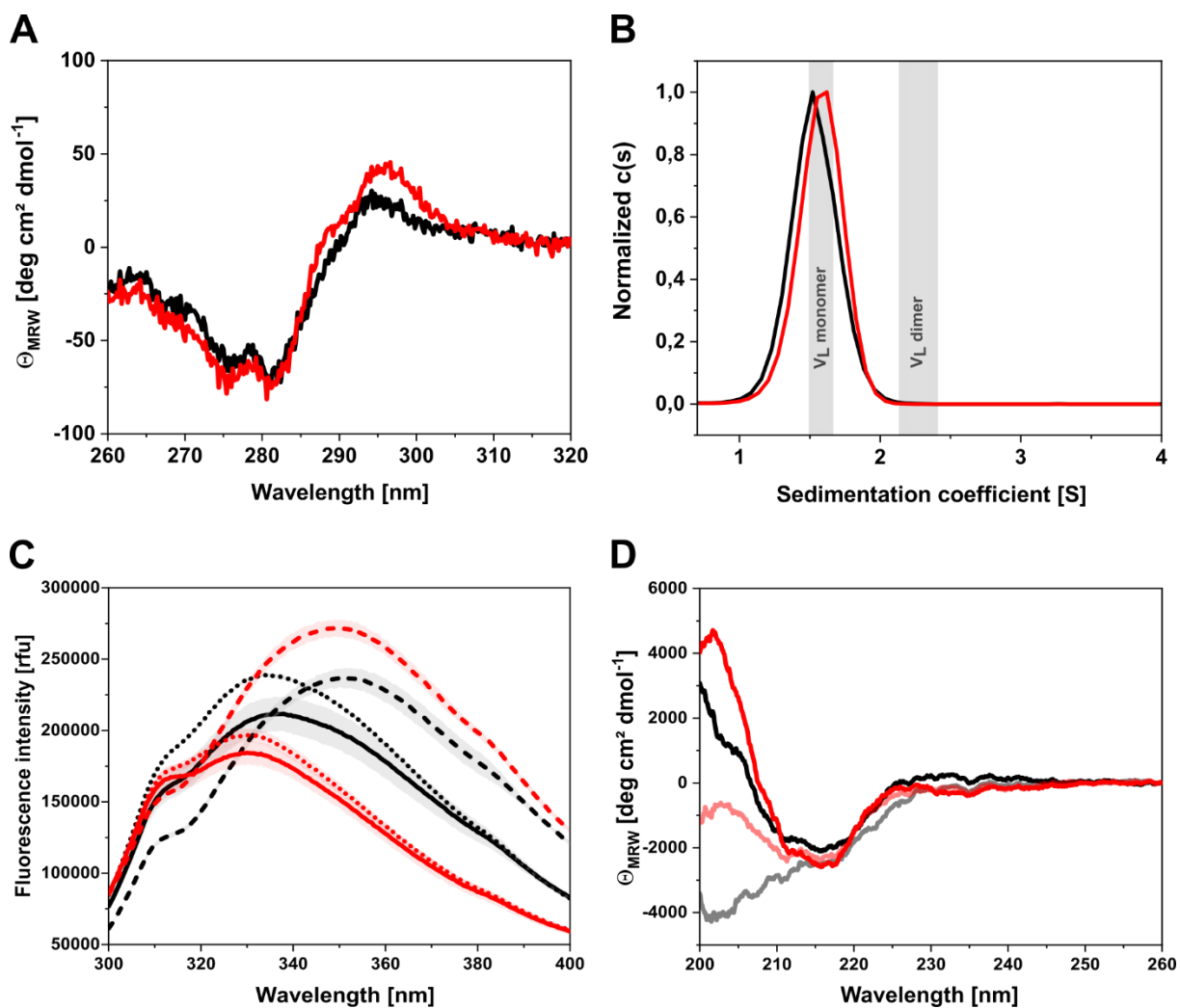


Figure 12: Biophysical characterization of FOR005-PT and FOR005-GL. A) NUV CD spectra of FOR005-PT (black) and -GL (red) at 20 °C. B) AUC of PT (black) and GL (red) shows S-values of 1.5 - 1.6 Svedberg corresponding to a V_L monomer with an estimated molecular mass of around 12 kDa (according to SEDFIT). C) Fluorescence spectra of folded (straight line), completely unfolded (dashed line), and refolded (pointed line) FOR005-PT (black) and FOR005-GL (red) were recorded in triplicates at a protein concentration of 1 μ M using a 1 cm fluorescence quartz cuvette. The native V_L domains were equilibrated with 0.6 M urea for 24 hours at room temperature; the unfolded proteins were equilibrated with 6 M urea. For refolding, the native V_L domains were first unfolded for two hours at room temperature using 6 M urea and subsequently diluted to 0.6 M urea using PBS. The refolded samples were then equilibrated for 24 hours at room temperature prior to measurement. The difference of the spectra shows the irreversibility of chemical unfolding for FOR005 variants. D) FUV CD spectra of native FOR005-PT and FOR005-GL at 20 °C are shown in black and red, respectively. The samples were heated up to 90 °C and subsequently cooled down back to 20 °C. FUV spectra were again recorded after overnight equilibration at 20 °C and are shown in grey for FOR005-PT and in light red for FOR005-GL. Adapted from Rottenaicher et al., 2021.

To test fibril formation propensities of the two V_L domains, the proteins were incubated in PBS (pH 7.4) under shaking at 37 °C and fibril formation was monitored using the amyloid-binding, thiazol-based fluorescent dye Thioflavin T (ThT) (Gade Malmos et al., 2017). As the ThT binding kinetics demonstrate, the patient V_L domain starts to form amyloid fibrils *in vitro* after approximately four days, whereas the corresponding germline protein does not engage in

amyloid aggregation (Fig. 13A, Table 2). The presence of fibrils in the samples was further confirmed by transmission electron microscopy (TEM). Amyloid fibrils were only observed on the TEM micrographs of the patient V_L sample and not in the germline control (Fig. 13B).

Thermodynamic stability is a crucial factor in protein aggregation and in AL amyloidosis (Blancas-Mejía and Ramirez-Alvarado, 2013). Therefore, the domain stabilities were investigated by chemical and thermal denaturation experiments. Unfolding transitions in the presence of increasing urea concentrations were obtained by fluorescence spectroscopy to determine the chemical domain stability and unfolding cooperativity (Fig. 13C, Table 2). FOR005-PT showed a midpoint of unfolding at a urea concentration (C_m) of 1.90 M, whereas the transition midpoint of the germline domain is at 4.28 M urea (Fig. 13C, Table 2). A two-state fit model was used to calculate unfolding free energies (ΔG_{un}) from the data, although this is in principle only applicable if unfolding is completely reversible. Since this is not the case under the conditions used (Fig. 12C), the results do not represent true ΔG_{un} values but are rather apparent unfolding free energies (ΔG_{un}^{app}) (Table 2). The calculated values of 11.51 kJ·mol⁻¹ for FOR005-PT and 19.53 kJ·mol⁻¹ for FOR005-GL imply a significant stability difference. In thermal denaturation experiments, the patient protein exhibited a melting temperature (T_m) of 43.5 °C, whereas the germline protein showed a melting temperature of 56.3 °C (Fig. 13D). The T_m values correspond to the temperatures at which 50 % of the protein is unfolded. However, since thermal unfolding of both FOR005-PT and FOR005-GL is also irreversible as demonstrated by CD spectroscopy (Fig. 12D), the obtained transition midpoints accordingly represent apparent melting temperatures (Table 2). Taken together, the data show that the patient V_L domain has a significantly decreased thermodynamic stability compared to its corresponding germline counterpart (Fig. 13, Table 2).

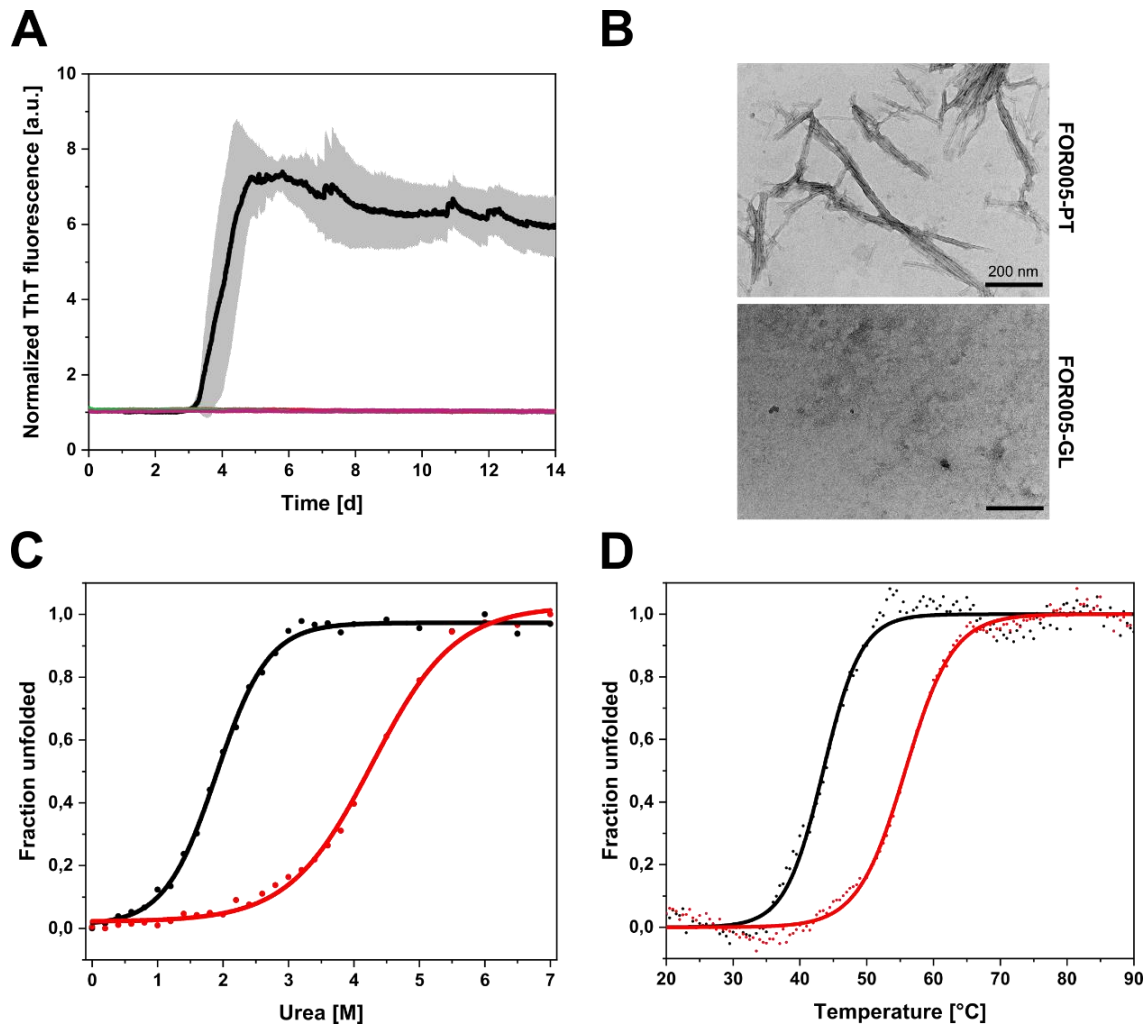


Figure 13: Fibril formation propensity and stability of the patient V_L domain. A) Thioflavin T binding kinetics of FOR005-PT (black) and FOR005-GL (red) obtained at 37 °C and pH 7.4 under continuous shaking. The increase in fluorescence shows that only the patient V_L domain is engaging in amyloid fibril formation after approximately three days. All kinetic curves were normalized to a fluorescence start value of 1. The grey area represents the standard deviation of triplicates. B) TEM micrographs of samples from finished ThT assays were recorded after negative stain with uranyl acetate. The amyloid fibers of FOR005-PT can be seen in the upper panel, the scale bar represents 200 nm. C) For chemical unfolding transitions 1 μ M protein was equilibrated with increasing concentrations of urea over night at room temperature. Fluorescence spectra ($\lambda_{ex} = 280$ nm/ $\lambda_{em} = 300 - 400$ nm) were recorded at 25 °C in a 96 well plate. The transition of FOR005-PT is shown with black dots, the data for FOR005-GL is shown as red dots. The black and red sigmoidal lines represent the individual fit functions. D) Thermal unfolding transitions of FOR005-PT and -GL were obtained by recording the CD signal at 205 nm while applying a temperature gradient from 20 – 90 °C with a heating rate of 1 °C/min. Adapted and modified from Rottenaicher et al., 2021.

3.1.3 The effect of point mutations on stability and aggregation of the V_L domain

The initial results of FOR005-PT and FOR005-GL show that the patient-specific mutations affect conformational stability and fibril formation, however, predicting or rationalizing which of the mutations are responsible for amyloidogenesis was impossible. Therefore, we replaced them individually in the germline sequence by the respective patient residues (Y31S, Y48F, G49R, N51S, and G94A) to decipher the specific effects of each of the five mutations. Basic structural characterization by CD spectroscopy and AUC analysis revealed that all point mutants possess the conserved β -sheet structure of the Ig fold and were monomeric in solution (Fig. 14A/C). Moreover, the amino acid substitutions have only minor effects on the global tertiary structure as similar near-UV CD spectra suggest (Fig. 14B).

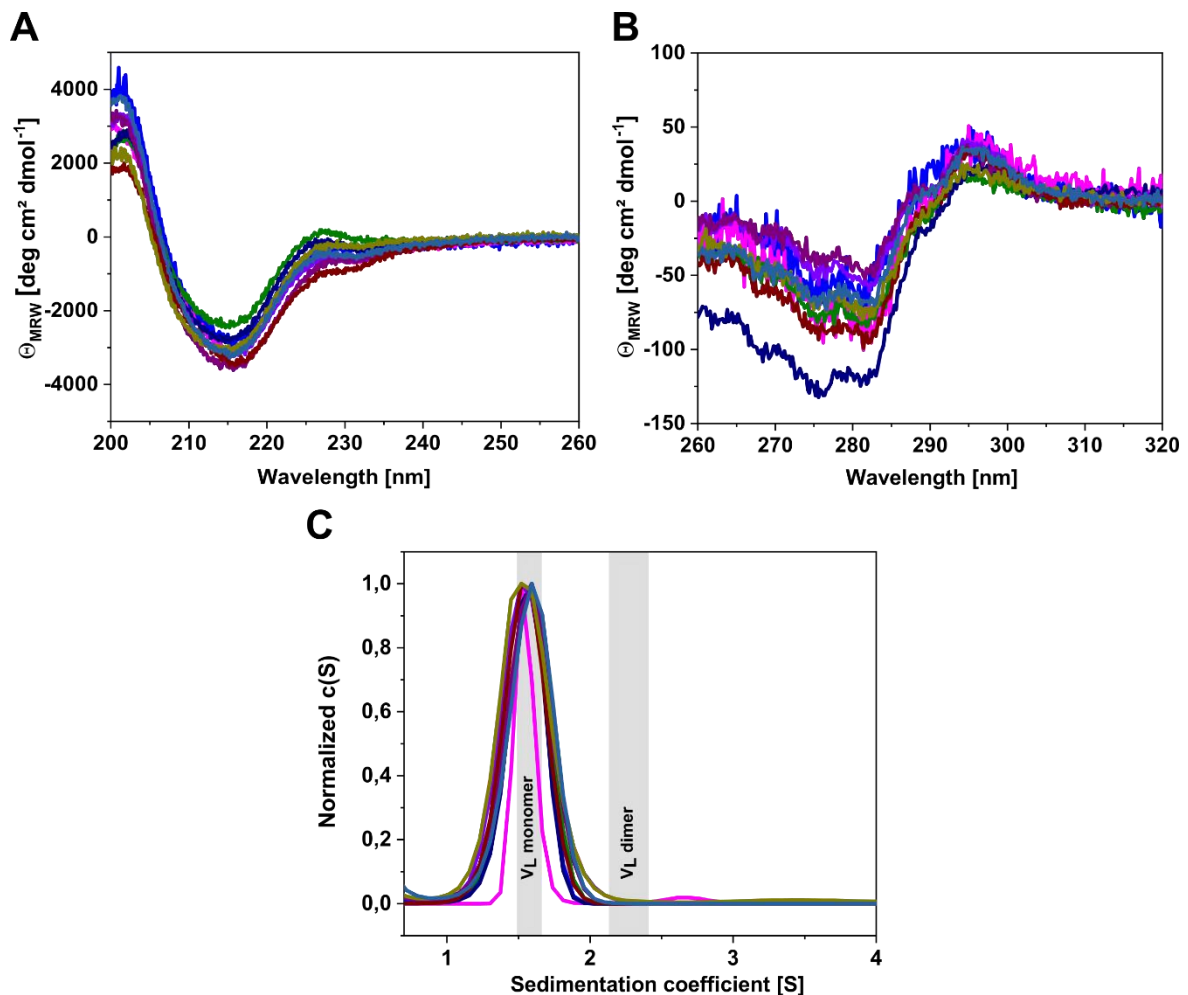


Figure 14: Structural characterization of FOR005-GL single and double mutants. A) FUV CD spectra of FOR005-GL mutants Y31S (blue), Y48F (pink), G49R (green), N51S (dark blue), G94A (violet), Y31S/G94A (purple), Y48F/G94A (brown), G49R/G94A (dark yellow), and N51S/G94A (pale blue). All variants share the same secondary structure. B) NUV CD spectra of FOR005-GL mutants (color code as in A) indicate a highly similar tertiary structure for all mutants. C) AUC of FOR005-GL single and double mutants shows that all proteins are monomeric in solution. Grey bars indicate the expected sedimentation regions of monomeric V_Ls (1.5 - 1.6 S) and V_L dimers (2.2 - 2.4 S), respectively. Adapted and modified from Rottenaicher et al., 2021.

Thermal unfolding experiments of the germline V_L domain mutants showed the most pronounced decrease in stability for the G94A mutant with a T_m value of 50.5 °C followed by the G49R mutant with a transition temperature of 52.2 °C (Fig. 15A, Table 2). The remaining mutants Y31S, Y48F, and N51S exhibited slightly decreased thermal stabilities with T_m values of 56.1 °C, 55.4 °C, and 54.4 °C, respectively (Fig. 15A, Table 2). Chemical unfolding experiments with urea revealed the strongest decrease in stability for the G49R mutant with a C_m value of 2.91 M, whereas the G94A variant unfolded at a concentration of 3.25 M urea. Further, both the G49R and G94A variant show comparable ΔG_{un}^{app} values of 16.49 kJ·mol⁻¹ and 16.18 kJ·mol⁻¹, respectively (Fig. 15B, Table 2). As seen for thermal unfolding, the transition midpoints of the Y31S (4.2 M) and Y48F (4.15 M) mutants lie only slightly below that of the germline reference, while the N51S variant unfolded at 3.44 M urea (Fig. 15B, Table 2).

The substitutions G49R and G94A showed the largest destabilizing effect on the germline V_L domain of all five mutations. These results were somewhat unexpected since G94A is a small, conservative mutation located in the hypervariable CDR3 loop. Therefore, double mutants were created by individually combining G94A with the remaining four mutations resulting in the double mutants Y31S/G94A, Y48F/G94A, G49R/G94A, and N51S/G94A. Of these the double mutants G49R/G94A and N51S/G94A exerted the largest effect on thermal stability with melting temperatures of 47.0 °C and 48.6 °C, respectively (Fig 15A, Table 2). Accordingly, in chemical unfolding experiments the mutations G49R/G94A and N51S/G94A were significantly affected with C_m values of 2.09 M and 2.56 M urea, and ΔG_{un}^{app} values of 14.14 kJ·mol⁻¹ and 15.45 kJ·mol⁻¹, respectively (Fig. 15B, Table 2).

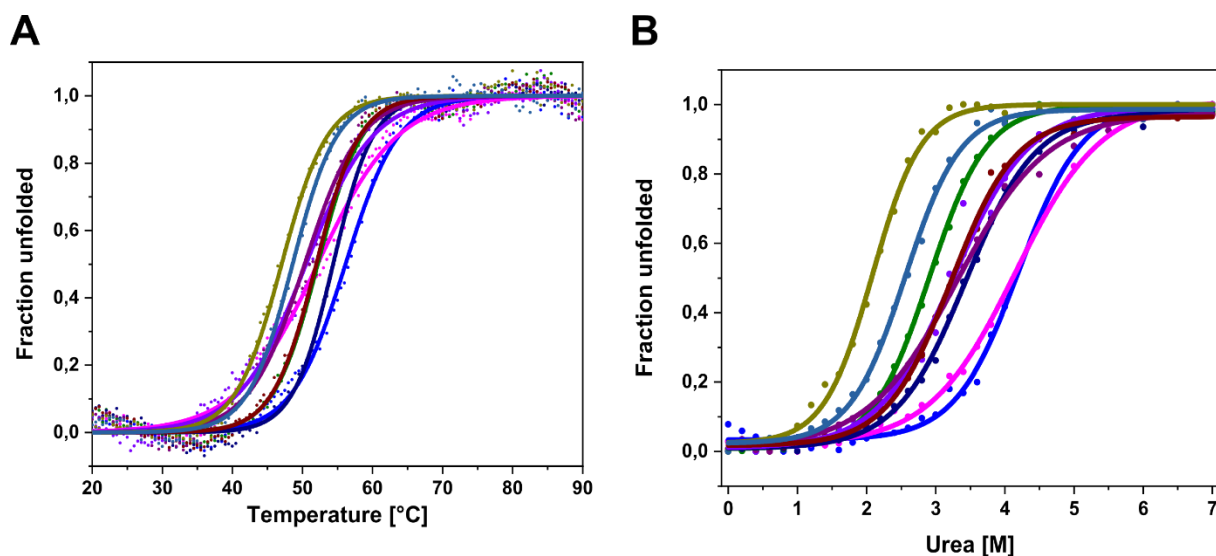


Figure 15: Stability of FOR005 V_L single and double mutants. Coloring: Y31S blue, Y48F pink, G49R green, N51S dark blue, G94A violet, Y31S/G94A purple, Y48F/G94A brown, G49R/G94A dark yellow, and N51S/G94A pale blue. A) Thermal transitions between 20 and 90 °C of FOR005-GL mutants recorded by CD at 205 nm with an applied heating rate of 1 °C/min. B) Urea induced unfolding transitions were recorded in triplicates at 25 °C using 1 μM protein. Samples were equilibrated over night at room temperature, excitation wavelength was 280 nm and emission was measured from 300 – 400 nm. Fluorescence intensities of the triplicates were averaged, normalized, and fitted with Origin. Adapted and modified from Rottenaicher et al., 2021.

To assess the effects of all mutations on fibril formation, ThT binding kinetics and TEM micrographs were obtained (Fig. 16). The data revealed that G49R/G94A is the only mutant that forms amyloid fibrils *in vitro* at pH 7.4 and 37 °C (Fig. 16A). The N51S/G94A mutant, however, did not form fibrils within two weeks at pH 7.4, despite exhibiting a significantly decreased thermodynamic stability similar to G49R/G94A. Previous studies have reported that thermodynamic destabilization is not necessarily the only critical factor in the amyloid formation pathway. Additionally, protein dynamics and the population of non-native folding intermediate states are also of importance (Khurana et al., 2001; Nokwe et al., 2016; Oberti et al., 2017). Therefore, additional fibril formation assays were performed at pH 6.4, since acidification can decrease stability and favor the population of alternatively folded intermediate states (Fig. 16B) (Buchner et al., 1991; Creighton, 1993; Feige et al., 2010). As expected, the decreased pH value led to accelerated fibril formation in ThT assays. FOR005-PT and G49R/G94A formed amyloid fibers faster than at physiological pH, but also for N51S/G94A fibril formation was observed after approximately 10 days under the altered conditions (Fig 16B). Again, the presence of amyloid fibrils in the samples was confirmed by TEM micrographs after the ThT assay (Fig. 16C). It seems that amyloid aggregation of FOR005-PT relies on a mechanism in which domain destabilization is an important, yet not the only decisive factor.

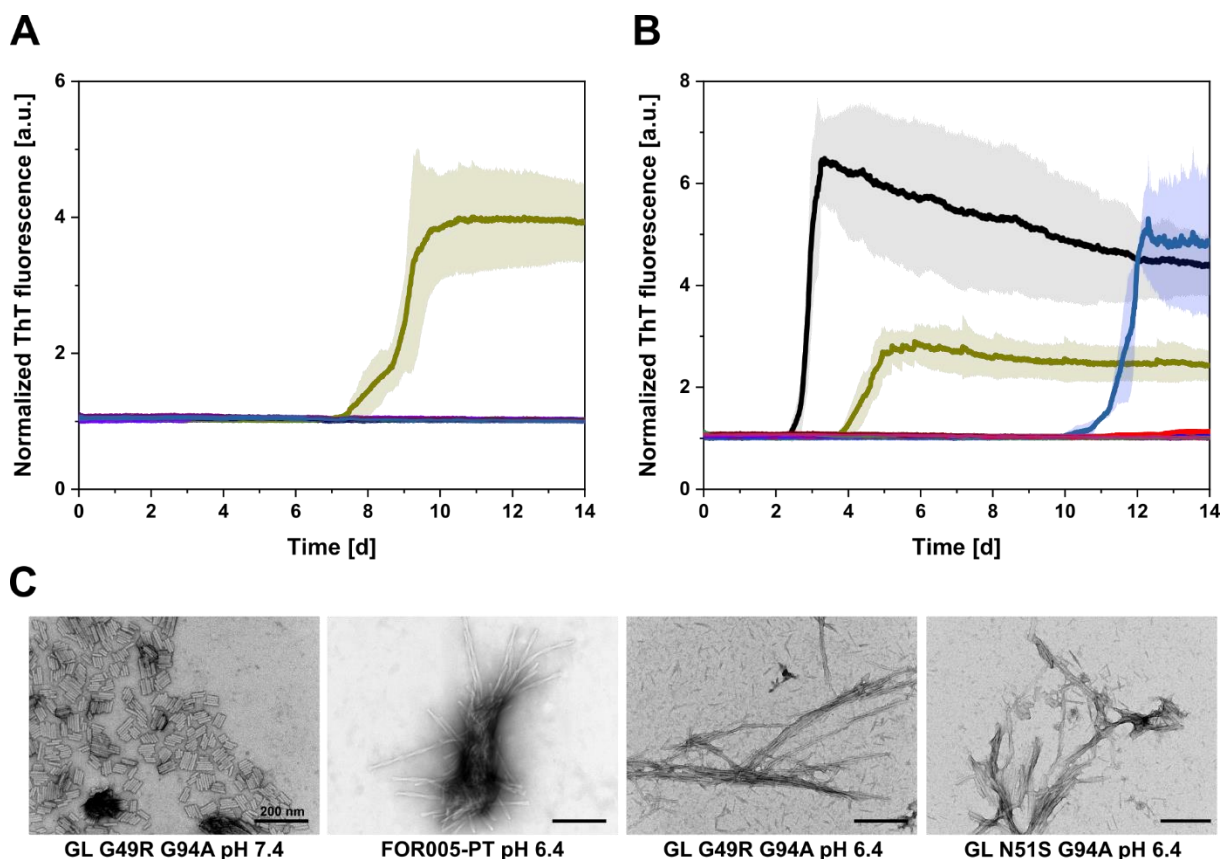


Figure 16: The effects of point mutations on fibril formation propensity. A) Fibril formation kinetics at 37 °C, pH 7.4, and continuous shaking show that G49R/G94A (dark yellow) is the only one of the nine investigated mutants that forms amyloid fibrils in vitro. All ThT kinetics were normalized to a fluorescence start value of 1. B) At pH 6.4 fibril formation of FOR005-PT (black) and G49R/G94A (dark yellow) is accelerated and also amyloid aggregation of N51S/G94A (pale blue) can be observed after approximately ten days. C) TEM micrographs of all FOR005 variants were obtained after two weeks of incubation at 37 °C, pH 7.4 or 6.4, and continuous shaking. Samples were stained using uranyl acetate. The four panels show the only samples that exhibited an increase in ThT fluorescence signal in the ThT assays depicted in A) and B). The scale bar represents 200 nm. TEM micrographs were recorded by Ramona Absmeier. Adapted from Rottenaicher et al., 2021.

In conclusion, the results for the single and double mutations revealed that only two out of the five mutations are mainly responsible for the significant loss in thermodynamic stability and the gain in amyloid formation propensity. Surprisingly, the conservative G94A mutation in the exposed, hypervariable CDR3 loop plays a vital role regarding the biophysical properties of the patient V_L domain.

Table 2: Stability parameters, unfolding cooperativity and fibril formation midpoints of FOR005 constructs.

V _L domain	T _m	C _m	ΔG_{un}^{app}	m	t ₅₀ pH 7.4	t ₅₀ pH 6.4
	°C	M urea	kJ mol ⁻¹	kJ mol ⁻¹ M ⁻¹	d	d
FOR005-PT	43.5 ± 0.14	1.90 ± 0.02	11.51 ± 0.63	5.99 ± 0.33	3.8	2.8
FOR005-GL	56.3 ± 0.11	4.28 ± 0.04	19.53 ± 1.20	4.33 ± 0.31	-	-
Y31S	56.1 ± 0.12	4.20 ± 0.03	21.42 ± 2.00	5.12 ± 0.47	-	-
Y48F	55.4 ± 0.13	4.15 ± 0.03	19.37 ± 1.43	4.56 ± 0.32	-	-
G49R	52.2 ± 0.10	2.91 ± 0.01	16.49 ± 0.36	5.55 ± 0.13	-	-
N51S	54.4 ± 0.10	3.44 ± 0.02	17.08 ± 0.79	5.01 ± 0.27	-	-
G94A	50.5 ± 0.17	3.25 ± 0.05	16.18 ± 3.23	4.95 ± 0.99	-	-
Y31S/G94A	50.3 ± 0.15	3.31 ± 0.03	16.48 ± 1.86	5.17 ± 0.58	-	-
Y48F/G94A	51.9 ± 0.10	3.22 ± 0.02	17.11 ± 0.81	5.32 ± 0.28	-	-
G49R/G94A	47.0 ± 0.09	2.09 ± 0.02	14.14 ± 0.84	6.69 ± 0.41	8.9	4.4
N51S/G94A	48.6 ± 0.08	2.56 ± 0.02	15.45 ± 0.81	5.98 ± 0.33	-	11.5

T_m values were obtained by recording the CD signal at 205 nm between 20 to 90 °C at a heating rate of 1 °C/min. Chemical unfolding transitions were obtained by fluorescence spectroscopy using 1 μM V_L domain with increasing concentrations of urea. Since both thermal and chemical unfolding are irreversible, the stability parameters T_m and C_m represent apparent values. Transition midpoints were derived from a Boltzmann fit, standard deviation represents goodness of fit. Chemical unfolding data was subjected to a two-state unfolding fit model to determine cooperativity and ΔG_{un}^{app} values. The t₅₀ values represent the time point at which fibril formation is 50 % completed.

3.1.4 Conformational dynamics are linked to decreased protein stability and amyloid formation

The causal link between aggregation propensity and conformational dynamics, as well as cellular toxicity of prefibrillar species has previously been shown by various studies (Oberti et al., 2017; Le Marchand et al., 2018; Kazman et al., 2020; Maritan et al., 2020). Thus, the structural flexibility and dynamics of the investigated V_L domains were probed by limited proteolysis with model proteases and hydrogen/deuterium exchange mass spectrometry (H/DX-MS).

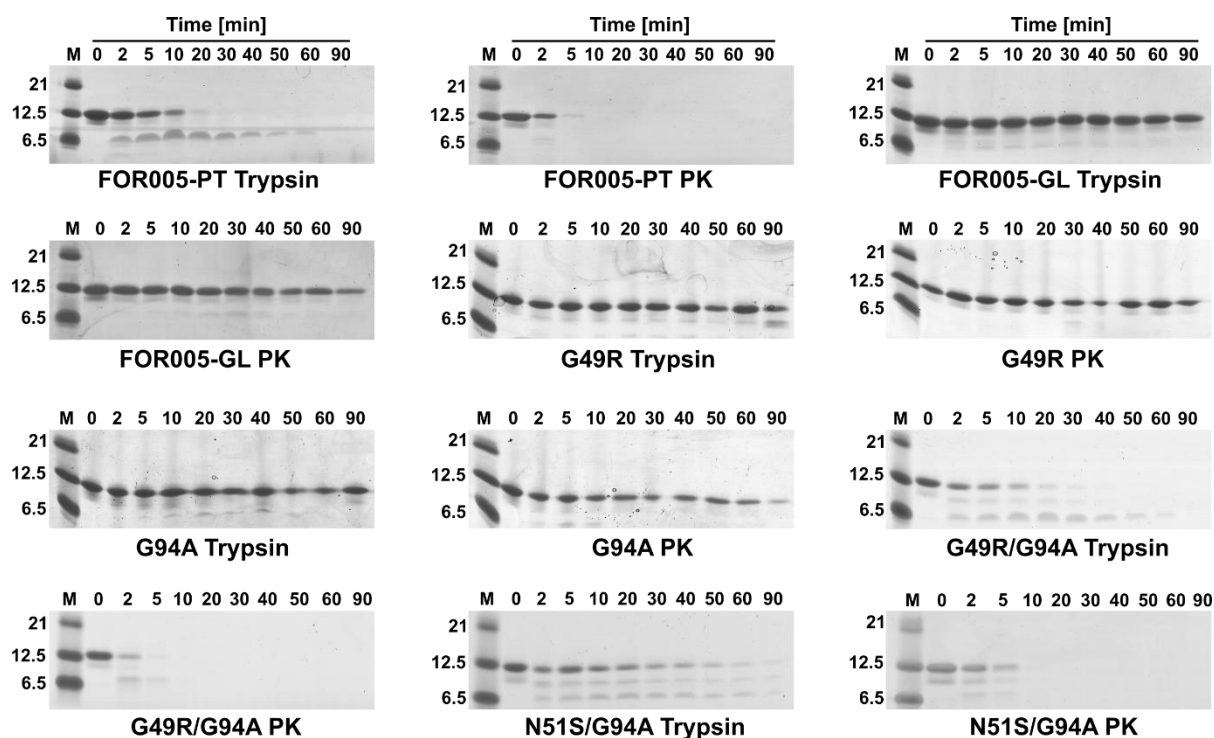


Figure 17: SDS-PAGE gels of limited proteolysis experiments of FOR005 V_L domains with trypsin and proteinase K. The experiments were performed at room temperature. Only one gel per proteolysis reaction is shown. SERVA protein test mixture 6 was used as marker (M). For digestion with trypsin a substrate/enzyme ratio of 15/1 (w/w) was used, for proteinase K reactions a ratio of 150/1 (w/w) was used. Adapted from Rottenaicher et al., 2021.

Structural flexibility can be assessed by limited proteolysis experiments since proteolytic degradation is increased by enhanced protein dynamics and partial unfolding (Fontana et al., 1999). The experiments were carried out with the proteases trypsin or proteinase K. We found that the patient-derived V_L domain was degraded much faster than its germline counterpart suggesting a highly dynamic behavior (Figs. 17, 18). In addition, the single point mutants G49R and G94A exhibited slow degradation kinetics similar to FOR005-GL. However, G94A is degraded faster and to a greater extent than G49R and FOR005-GL. The double mutant N51S/G94A is also cleaved more readily than the germline V_L and the observed single mutants, yet not as fast as the patient-derived V_L domain. Notably, the fibril-forming double mutant G49R/G94A is degraded even more readily than the patient V_L domain FOR005-PT (Figs. 17, 18).

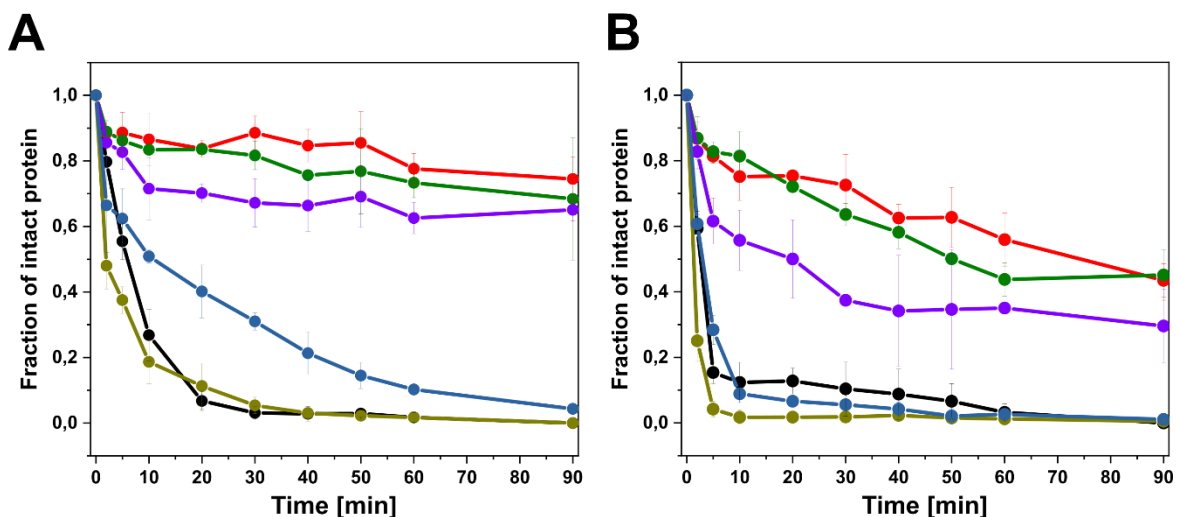


Figure 18: Quantification of limited proteolysis experiments of FOR005 V_L domains. A) Limited proteolysis with trypsin (substrate/enzyme ratio = 15/1). B) Limited proteolysis with proteinase K (substrate/enzyme ratio = 150/1). The quantification of SDS-PAGE gel bands was performed with NIH ImageJ. Color coding: FOR005-PT black, GL red, G49R green, G94A violet, G49R/G94A dark yellow, and N51S/G94A pale blue. Increased susceptibility to proteolytic degradation implies enhanced structural dynamics. Adapted and modified from Rottenaicher et al., 2021.

To gain detailed insights into the conformational dynamics of the V_L domains, H/DX-MS was applied. This method is based on the exchange of backbone amide hydrogen atoms with the surrounding solvent in flexible protein regions. From this, peptide-resolved information about the dynamics can be derived after pepsin cleavage and LC-MS/MS analysis (Englander, 2006). The fractional deuterium uptake was analyzed for variants FOR005-PT, FOR005-GL, G49R, G94A, as well as the double mutants G49R/G94A and N51S/G94A (Fig. 19A). By dividing uptake ratios of the investigated mutants by the uptake ratios of the germline V_L domain, the fold change in fractional uptake was calculated (Fig. 19B). Thus, the stable FOR005-GL serves as a reference point to evaluate how certain mutations affect conformational dynamics. A value below 1 indicates that the germline exhibits higher deuterium uptake, whereas a value above 1 shows increased deuterium uptake for the observed mutant. Conformational dynamics are especially enhanced for residues 12 – 20, residues 65 – 85, and residues 97 – 105 in the case of FOR005-PT and for the double mutant G49R/G94A (Fig. 19). For the double mutant, however, a slightly lower fold change of uptake was observed compared to the actual patient V_L that contains all five substitutions. Interestingly, the stretch from residues 50 – 60 which includes the CDR2 loop is more dynamic in the germline V_L domain and in G49R/G94A than in FOR005-PT. The double mutant N51S/G94A also exhibits increased dynamic behavior, especially for residues 80 – 105, whereas the single mutants G49R and G94A, in comparison, do not impose a strong effect on the conformational flexibility. Notably, G94A affects overall dynamics slightly more than G49R (Fig. 19).

Table 3: Results of H/DX-MS data acquisition.

Protein	Number of peptides	Sequence coverage (%)	Redundancy
FOR005-PT	41	96.4	4.24
FOR005-GL	59	92.7	6.18
G49R	48	100.0	5.15
G94A	49	98.2	5.18
G49R/G94A	39	96.4	4.06
N51S/G94A	42	96.4	4.43

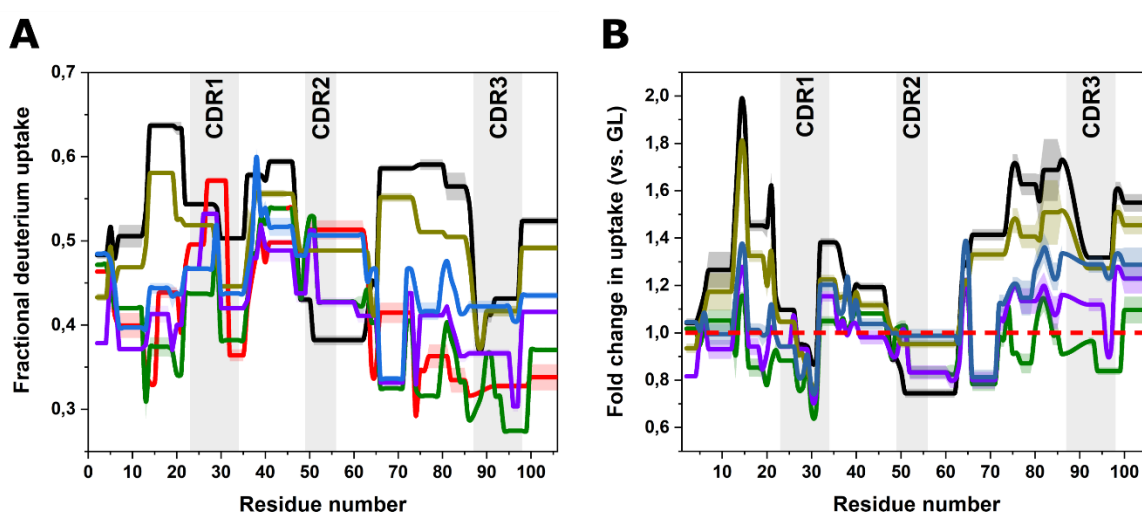


Figure 19: H/DX-MS of FOR005 V_L domains. A) Absolute fractional deuterium uptake after two hours incubation with D₂O detected by ESI-TOF/TOF mass spectrometry gives peptide-resolved information on protein backbone dynamics. B) The fold change in fractional uptake (2 hours) compared to the germline V_L was calculated by dividing the uptake values of the respective mutants by the uptake values of FOR005-GL. The dashed red line at a value of 1 represents the germline V_L. Coloring: FOR005-PT black, GL red, G49R green, G94A violet, G49R/G94A dark yellow, and N51S/G94A pale blue. B) is adapted from Rottenaicher et al., 2021.

For better visualization, the change in fractional uptake was plotted onto the crystal structure of FOR005-PT (PDB: 5L6Q) (Fig. 20). Structurally, the most affected regions correspond to the β -strands A2 and B and the small loop connecting them (residues 12 – 20). Further, for the β -strands E and F including the small interspaced helical segment (residues 65 – 85), and the C-terminal β -strands G1 and G2 an increase in dynamics can be observed (Fig. 20). In summary, the results display a connection between conformational dynamics and amyloid aggregation. Remarkably, the strongest increase in dynamics is observable in conserved framework regions rather than the CDR segments where the point mutations are located. These results suggest that the mutation could induce small changes in CDR loop conformations, which potentially propagate through the domain architecture. Thereby,

backbone dynamics in framework regions are enhanced which lowers the thermodynamic stability and enhances aggregation propensity.

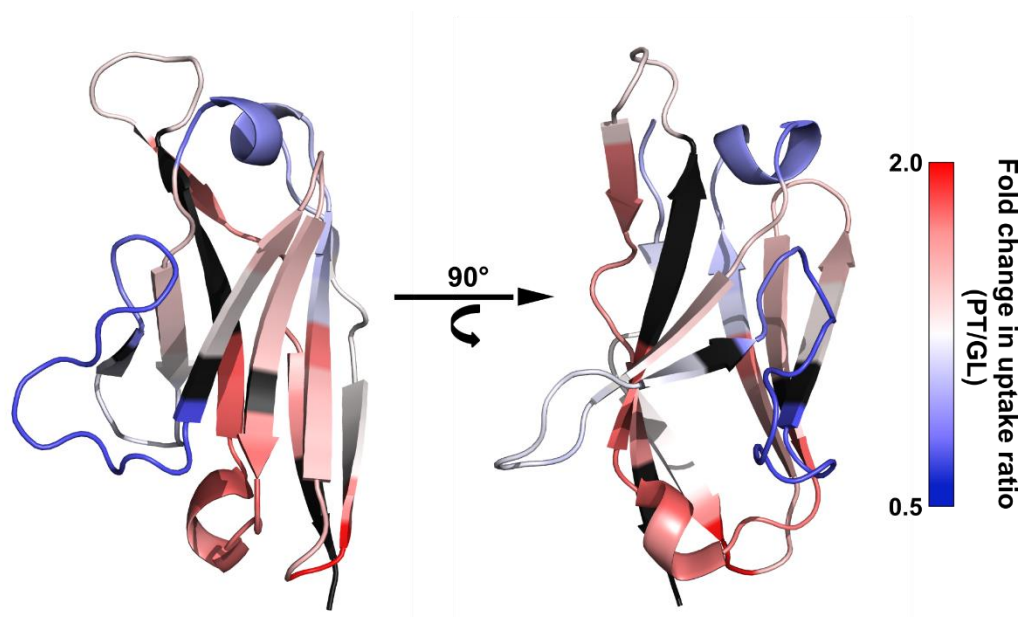


Figure 20: Fold change in uptake of FOR005-PT plotted onto the crystal structure of the patient V_L domain. Red segments show strongly enhanced dynamics in the patient V_L domain, blue segments are more dynamic in the germline protein. For residues colored in black no fold change data could be obtained. Adapted and modified from Rottenaicher et al., 2021.

3.1.5 Energetically unfavorable conformations of the CDR loops negatively impact the stability of the V_L

For additional information on the flexibility of the patient V_L domain, molecular dynamics (MD) simulations were performed by Dr. Manuel Hitzenberger and Prof. Dr. Martin Zacharias from the Chair of Theoretical Biophysics at TUM. The simulations were carried out in explicit solvent on FOR005-PT and a PT variant containing the R49G and A94G double substitution (PT R49G/A94G) (Fig. 21), as well as FOR005-GL, GL G49R/G94A, and GL N51S/G94A (Fig. 22). The variants exhibited only small deviations from the start structure and were stable during the 1 μs MD simulations (Figs. 21, 22). Root-mean-square fluctuations (RMSF) were compared and the lowest fluctuations were observed for FOR005-GL, whereas GL N51S/G94A showed slightly enhanced fluctuations. Significantly increased fluctuations, especially around residue 49 and 94, were found in the case of the GL G49R/G94A variant (Fig. 22). Further, FOR005-PT exhibited larger conformational fluctuations on the MD time scale than the variant PT R49G/A94G (Fig. 21).

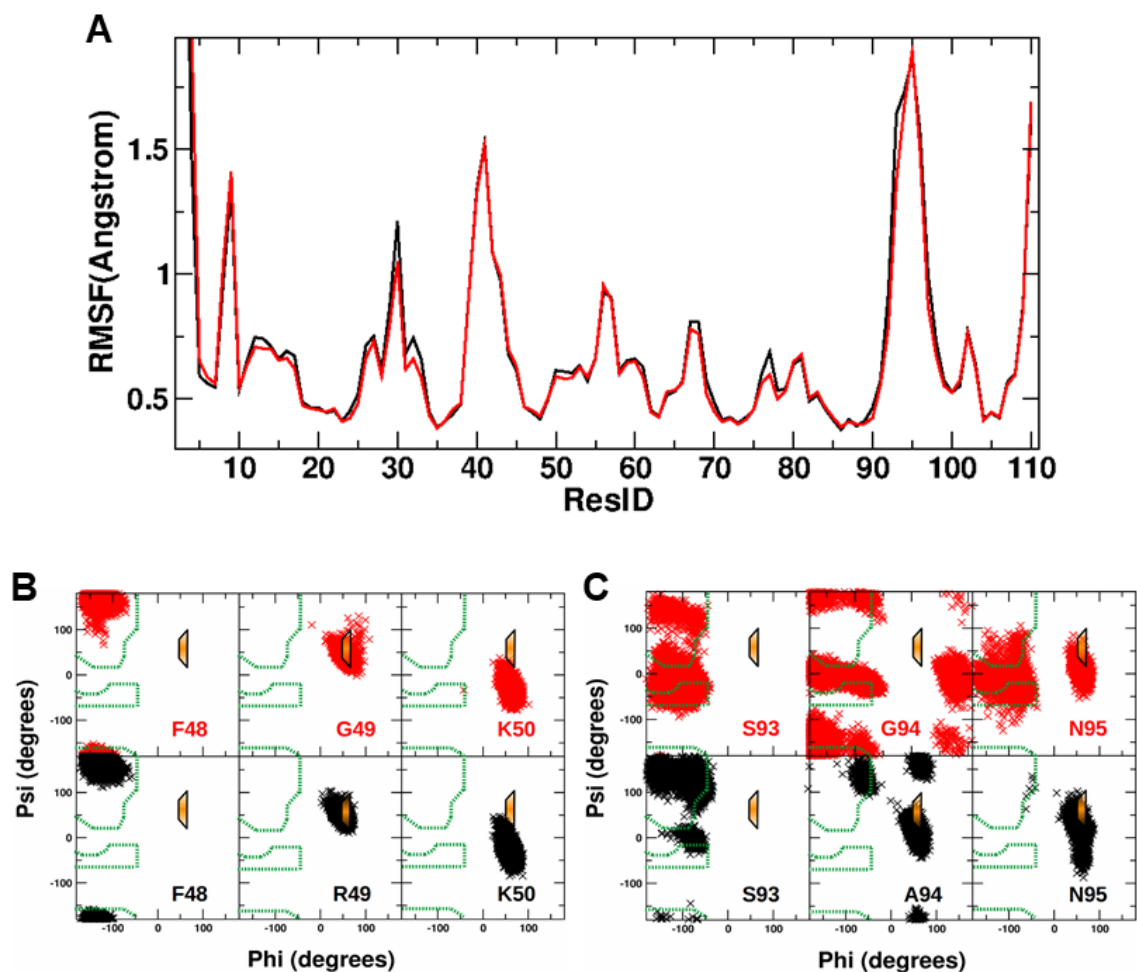


Figure 21: MD simulations of FOR005-PT V_L and the R49G and A94G double substitution. A) Root-mean-square fluctuations (RMSF) observed in MD simulations (1 μ s, at 310 K) along the residue sequence for the FOR005-PT V_L variant (black line) and the R49G and A94G double substitution (red line). B) Sampled backbone dihedral angles phi and psi plotted as Ramachandran plots for residues 48-50 observed during 1 μ s simulations of FOR005-PT V_L (black dots) and the R49G and A94G variant (red dots). Favorable regions for non-glycine residues are indicated by a green dashed boundary in the Ramachandran plots and a regime favorable for glycine but less for other amino acids is indicated in orange with a blue boundary. C) Same as B) but Ramachandran plots for residues 93 - 95 observed during 1 μ s simulations of FOR005-PT V_L (black dots) and the R49G and A94G variant (red dots). The figure was kindly provided by Prof. Dr. Martin Zacharias. Adapted from Rottenbacher et al., 2021.

Furthermore, the peptide backbone dihedral angles in the loop regions near residues 49 and 94 were analyzed. It was found that residues 49 and 50, as well as residues 94 and 95 (but not residues 51 or 52) sample left-handed helical regimes in the Ramachandran plots (Fig. 22) which is sterically favorable for glycine. Hence, the presence of R at position 49 and A at position 94 creates a steric strain on these loop structures. Glycine residues, however, can relax this strain and would thus be favorable for the loop conformation. The same sampling of sterically unfavorable peptide backbone states was also observed in the MD simulations of FOR005-PT (Fig. 21). These unfavorable loop backbone conformations are also found in the crystal structure of the patient protein. Hence, in the structure of the patient V_L, the loop

residues are being forced into an energetically unfavorable backbone conformation. This forced strain can be released by the germline substitutions R49G and A94G because the glycine residues at positions 49 and 94 are more compatible with the required loop structure. The backbone sampling of neighboring residues is also affected by these substitutions (Fig. 21, Fig. 22). According to these findings, the enforced backbone conformation of the CDR loops – mediated by the mutations at position 49 and 94 - could be the reason for the much lower stability of FOR005-PT and of the variants with non-glycine residues at positions 49 and 94. The N51S substitution, however, had seemingly no significant effect since for residues 51 and 52 approximately the same sampling of favorable backbone states was observed (Fig 22).

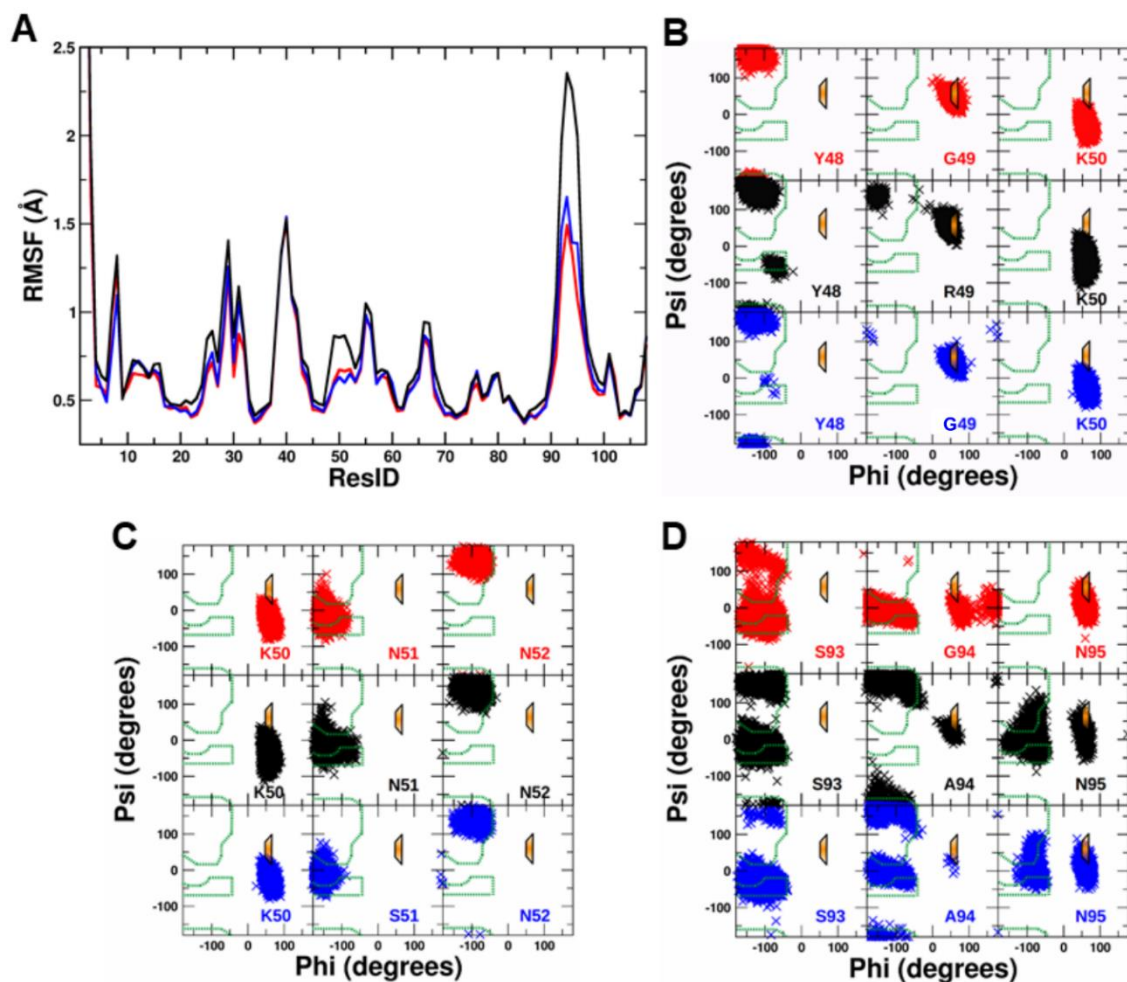


Figure 22: MD simulations show energetically unfavorable backbone conformations in CDR2 and CDR3. A) Root-mean-square fluctuations (RMSF) observed in MD simulations (1 μ s, at 310 K) along the residue sequence for the FOR005-GL V_L variant (red line), the FOR005-GL G49R/G94A (black line), and the GL N51S/G94A (blue line) substitutions. B) Sampled backbone dihedral angles phi and psi plotted as Ramachandran plots for residues 48-50 (same color code as in A). Favorable regions for non-Gly residues are indicated by a green dashed boundary in the Ramachandran plots and a regime favorable for Gly but less for other amino acids is indicated in orange with a blue boundary. C) Same as in B) but for residues 50-52. D) Same as in B) but for residues 93-95. The figure was kindly provided by Prof. Dr. Martin Zacharias. Adapted from Rottenaicher et al., 2021.

3.1.6 Discussion

The vast sequence variability of the precursor LC renders systemic AL amyloidosis a highly complex protein misfolding disease. Thus, it is challenging to acquire a detailed mechanistic understanding of the LC amyloid aggregation process. Moreover, there are numerous other factors that affect fibril formation, such as proteolytic processing of the precursor LCs (Merlini, 2017; Blancas-Mejia et al., 2018).

It is still unclear whether proteolytic processing of LCs is a prerequisite or a consequence of fibril formation and this topic is still subject to debate in the field (Enqvist et al., 2009; Morgan and Kelly, 2016; Rennella et al., 2019a). Studies have shown that the C_L domain can protect LCs from amyloid aggregation *in vitro* (Klimtchuk et al., 2010; Weber et al., 2018; Kazman et al., 2020). In the case of FOR005, amyloid deposits in the patient's tissue consist only of the V_L domain (Annamalai et al., 2017). In line with this finding, only the patient-derived V_L domain but not the corresponding germline V_L engaged in the amyloid pathway *in vitro*. The role of the FOR005 C_L domain will be discussed in detail in chapter 3.2.

FOR005 is an interesting case, because the V_L domain has four CDR mutations and only one framework mutation compared to its germline counterpart. Notably, the location of the CDRs depends on the choice of the domain numbering system. The Kabat, Chothia, and IMGT systems are the ones most commonly used (Wu and Kabat, 1970; Chothia and Lesk, 1987; Lefranc et al., 2009). In the IMGT numbering scheme, the FOR005 mutation N51S would be considered a framework mutation rather than a CDR2 mutation. In the Kabat and Chothia classification systems, however, this residue belongs to CDR2 which is in line with the ConSurf residue conservation analysis. The result of germline sequence identification for a given V_L domain depends on the applied method/ database. Here, abYsis, IgBLAST, and IMGT were used to identify a V_L domain with highest possible amino acid sequence identity (Brochet et al., 2008; Lefranc et al., 2009; Ye et al., 2013; Swindells et al., 2017).

Framework mutations are known determinants of LC amyloid aggregation (Hurle et al., 1994; Baden et al., 2008b; Hernández-Santoyo et al., 2010; González-Andrade et al., 2013; Del Pozo-Yauner et al., 2014; Nokwe et al., 2016; Luna-Martínez et al., 2017; Weber et al., 2018; Rennella et al., 2019b; Kazman et al., 2020). The only framework mutation in FOR005-PT – Y48F – has already been investigated in a previous study. Using the V_L domain REI, it has been shown that this mutation has little to no influence on domain stability and aggregation propensity (Hurle et al., 1994). Therefore, the hypothesis was established that only the CDR mutations are important in the case of FOR005, which was then confirmed experimentally. The CDR loops of antibodies fulfill more functions than just antigen recognition. They also play a critical part in antibody domain architecture and V_H/V_L domain association. Their strong

influence on the folding pathway, stability, and conformation has been demonstrated by both experimental and computational studies on V_L domains (Helms and Wetzel, 1995; Kuroda et al., 2009; Gaćiarz and Ruddock, 2017; Herold et al., 2017; van der Kant et al., 2019). Regarding the role of CDR mutations in LC amyloidogenicity, it has been shown that deletion of a proline residue in the CDR3 loop of an amyloidogenic V_L domain resulted in enhanced stability and delayed fibril formation (Blancas-Mejía et al., 2014). Furthermore, it was reported that the kinetic stability of V_L domains and LCs from AL and multiple myeloma (MM) patients was affected by non-conservative mutations – including some in the CDR3 loops (Blancas-Mejía et al., 2015b). Moreover, CDR1 has been shown to represent a hotspot for aggregation and that a CDR3-based peptide can drive amyloid fibril formation due to enhanced steric zipper propensity (Brumshtein et al., 2018; Ruiz-Zamora et al., 2019). Nonetheless, the exact role of CDR residues in the context of the disease is still enigmatic (Poshusta et al., 2009).

The misfolding and amyloid aggregation of LCs are mediated by a number of factors: thermodynamic and kinetic stability, structural dynamics, partial unfolding, LC dimerization, and local changes to the native fold (Baden et al., 2008a; Baden et al., 2008b; Feige et al., 2008; DiCostanzo et al., 2012; Oberti et al., 2017; Blancas-Mejia et al., 2018). Decreased thermodynamic and kinetic stability have been viewed as the major driving forces in the misfolding and aggregation pathway of LCs (Marin-Argany et al., 2015; Nokwe et al., 2016; Kazman et al., 2020; Maritan et al., 2020). Regarding FOR005, the synergy between thermodynamic destabilization and altered conformational dynamics seemingly governs the pathway of the soluble V_L monomer towards amyloid aggregation. Based on the fibril formation kinetics and thermodynamic data of the FOR005 double mutants it can be deduced that CDR mutations do not drive the V_L towards amyloid formation only through destabilization, since the severely destabilized N51S/G94A mutant only forms fibrils after prolonged incubation at lower pH. It is the additional increase in conformational dynamics that further determines the shift towards amyloid aggregation. This increase in dynamics is initiated by the two critical CDR mutations G49R and G94A. Remarkably, it is not the hypervariable loop segments containing the mutations that experience the strongest increase in dynamics, but rather the conserved framework regions and the protein core are affected. According to the MD simulations, the loop residues 49, 50, 94, and 95 sample energetically unfavorable backbone conformations for non-glycine residues which negatively impacts the stability. Considering that glycine residues are structurally preferable at positions with certain ϕ/ψ angles (Creighton, 1993; Hurle et al., 1994), it appears plausible that the energetically unfavorable conformation of the loops affects the structural interplay of CDRs and framework. The resulting framework dynamics due to these strained loop structures are likely the cause for the lower stability of variants with non-glycine residues at the aforementioned positions.

These findings seem counter-intuitive at first sight. When considering the primary function of CDRs, one of their basic traits is the sequence diversity that results in conformational flexibility. This enables the loops to structurally adapt to antigens and interact with them. It is known, however, that out of the six CDRs of an antibody F_{ab} , five adopt only a limited number of backbone conformations. These conformations are called canonical classes and the heavy chain CDR3 (CDRH3) represents the only exception as it is not structurally restrained (Chothia and Lesk, 1987; Chothia et al., 1989; Al-Lazikani et al., 1997; Kuroda et al., 2009; Dondelinger et al., 2018). Therefore, a mutation-induced deviation from the canonical CDR class resulting in unfavorable loop conformations that put structural strain on the framework appears as a plausible mechanism. This deviation is visualized by alignment of the crystal structure of FOR005-PT with the structures of similar, non-amyloidogenic V_L domains (Fig. 23). Similarly, a canonical class alteration has been reported for the CDR1 loop of some amyloidogenic LCs of the $\lambda 6$ subfamily (del Pozo Yauner et al., 2006). Nonetheless, a drastic alteration of protein structure and stability by conservative mutations in exposed loops was unexpected (Chang et al., 2014). Yet, in FOR005 the CDR mutations, especially G94A, strongly impact stability and conformation of the V_L domain by inducing unfavorable CDR loop conformations. The resulting changes seem to propagate through the entire protein leading to increased flexibility, which causes the enhanced population of partially unfolded, aggregation-competent states (Khurana et al., 2001; Feige et al., 2008). Thus, the altered stability and amyloidogenicity of the V_L domain can be attributed to an altered interplay of hypervariable loops and conserved framework regions (Helms and Wetzel, 1995; Blancas-Mejía et al., 2014). Conclusively, in the FOR005 case the onset of fibril formation is directly and mechanistically correlated to the substitution of two distinct amino acids in CDR loops. Surprisingly, one of these two decisive mutations is the small, conservative G94A in the surface-exposed, hypervariable CDR3 loop.

Recently, Pradhan and co-workers have demonstrated by NMR spectroscopy that the R49 residue in FOR005-PT has an important function in stabilizing the structure of the fibril core (Pradhan et al., 2020). Therefore, it can be argued that different mutations in amyloid-forming LCs can serve different purposes and can thus be categorized. For instance, the G94A mutation induces conformational changes in the CDR3 loop, thereby causing a structural deviation from the canonical CDR class, which enhances framework dynamics and destabilizes the domain. This concept is illustrated by a structural alignment of FOR005-PT with three highly similar, non-amyloidogenic V_L domains derived from the PDB (Fig. 23). However, residue A94 is unimportant in the final core structure of FOR005-PT fibril, thus its only effect is the destabilization of the precursor V_L domain. On the other hand, the CDR2 mutation G49R favors amyloidogenesis by both altering CDR2 loop conformation and by stabilizing the fibril core through a side chain interaction (Pradhan et al., 2020). Yet, since G94A mediates a larger increase in conformational dynamics than G49R and since the CDR2

mutation N51S is also capable of inducing fibril formation, the primary role of G49R appears to lie in stabilizing the final product of the aggregation pathway – the core of the amyloid fiber. However, the two point mutations do not induce fibril formation individually *in vitro* but need to be combined in order to act synergistically. In conclusion, the decisive, amyloid-driving mutations of an LC can act in different ways as exemplified by G49R and G94A in FOR005: either the mutation establishes a specific side chain interaction inside the fibril core structure, or it destabilizes the soluble precursor protein in a specific manner. In the case of FOR005, the enhanced framework dynamics lead to rearrangements ultimately forming up the fibril core, thus, rendering the relationship of the mutations and fibril formation topologically indirect.

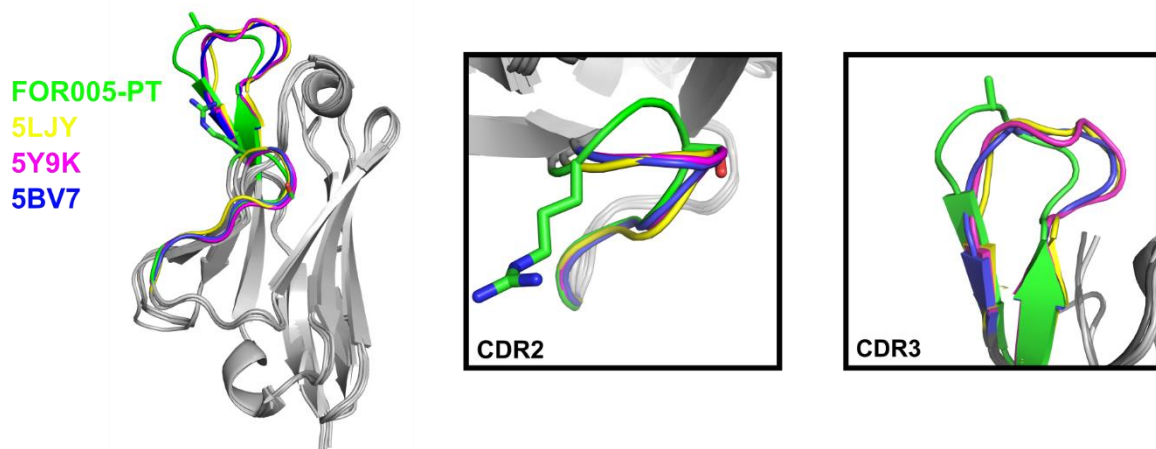


Figure 23: Structural alignment of FOR005-PT with three highly similar, non-amyloidogenic V_L domains shows the changed CDR loop conformation representing a deviation from canonical CDR classes. The PDB was searched for primary sequences with high sequence similarity to FOR005-PT and the crystal structures of the V_L monomers were aligned using PyMOL. The CDR2 and CDR3 regions are colored as follows: FOR005-PT = green, 5LJY = yellow, 5Y9K = magenta, 5BV7 = blue. The conformational differences between the CDR loops represent canonical class deviations and could provide a possible explanation for the observed destabilizing effects of CDR mutations. Adapted from Rottenaicher et al., 2021.

3.2 Disease onset in the case of FOR005 is linked to a non-conservative mutation in the constant light chain (C_L) domain

3.2.1 The sequence of FOR005 contains a rare C_L mutation

As stated above, the cDNA sequence coding for the FOR005 LC was derived from the patient and reported in 2017. The amyloid deposits found in the patient tissue, however, consisted of the V_L domain only (Annamalai et al., 2017). Using IMGT, abYsis, and IgBLAST the most closely related germline LC sequence (G-LC) was identified and it was found that the patient LC (P-LC) contains six mutations in total (Fig. 24) (Brochet et al., 2008; Lefranc et al., 2009; Ye et al., 2013; Swindells et al., 2017). Five of the six mutations are found in the V_L domain and four of those V_L mutations lie in the hypervariable complementarity determining regions (CDRs) as shown in Chapter 3.1. However, FOR005 also carries a point mutation (V136G) in the C_L domain - which is quite rare. Conservation analysis of the G-LC using ConSurf showed the valine at position 136 to be highly conserved and buried which implies an integral structural role (Fig. 25) (Ashkenazy et al., 2010). Further, we applied homology modelling using SWISS-MODEL (template structure: 5BV7) and observed that position 136 is located in the middle of the C_L β-strand b which is part of the highly conserved β-sandwich topology of Ig domains (Kiefer et al., 2009). Conclusively, a non-conservative substitution that alters side chain properties at this position is likely to have a strong effect on the biophysical and biochemical characteristics of the C_L domain and, therefore, on the entire LC.

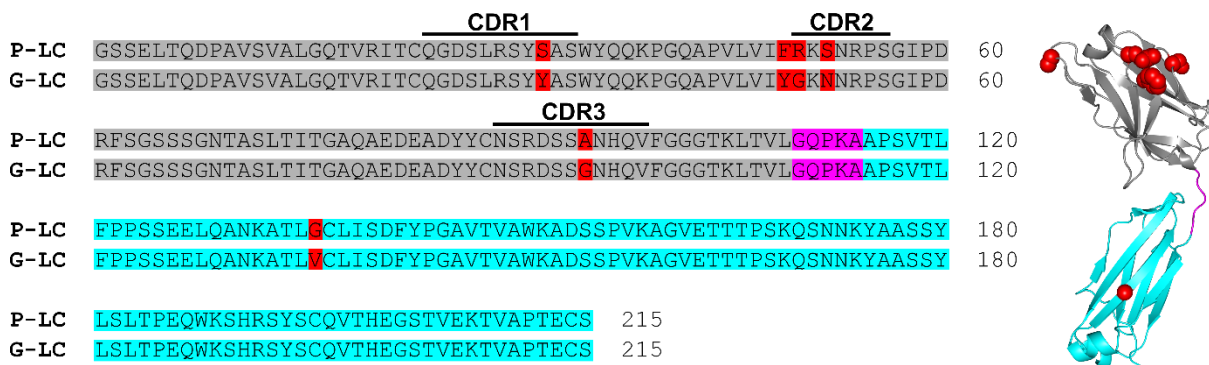
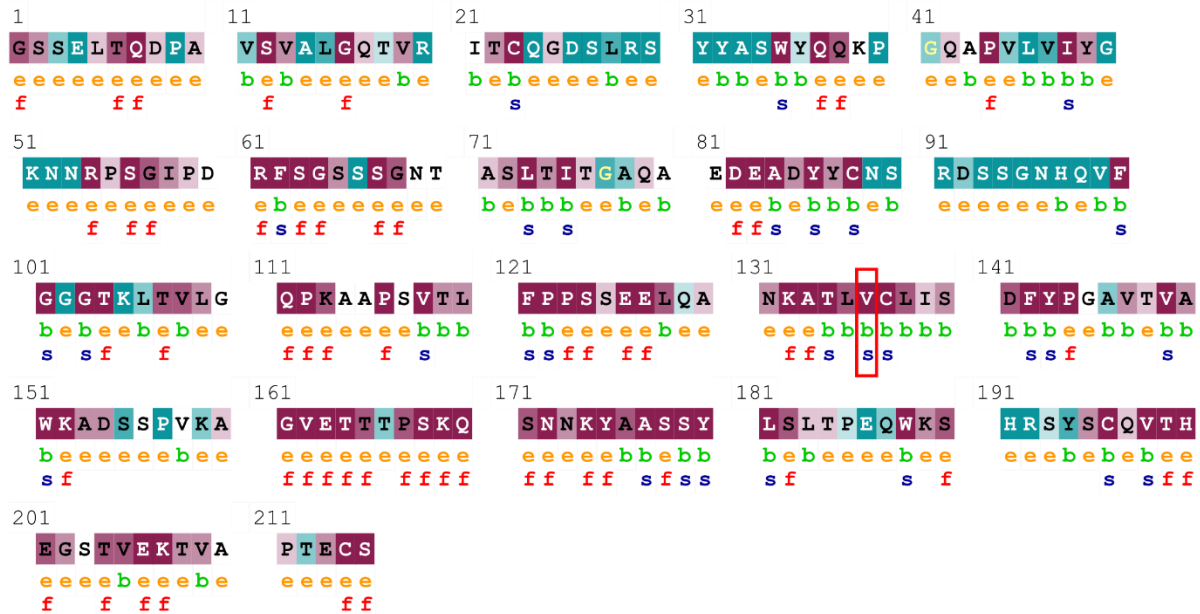


Figure 24: Sequence alignment and homology model of full-length FOR005. The V_L domain is colored in gray, the linker region in magenta, the C_L domain in cyan. Mutations are indicated in red. The homology model was created with SWISS-MODEL using the template structure 5BV7. The C-terminal cysteine at position 214 has been mutated to a serine in all constructs to allow purification of monomeric LCs.



The conservation scale:



- e** - An exposed residue according to the neural-network algorithm.
- b** - A buried residue according to the neural-network algorithm.
- f** - A predicted functional residue (highly conserved and exposed).
- s** - A predicted structural residue (highly conserved and buried).
- x** - Insufficient data - the calculation for this site was performed on less than 10% of the sequences.

Figure 25: Consurf analysis of G-LC indicates integral role of V136. The red box marks the valine residue at position 136 that is mutated to glycine in P-LC. As indicated by the conservation scale, residues colored in cyan have low conservation and residues in white have an average conservation degree. Purple residues are highly conserved. According to the Consurf algorithm, valine 136 is highly conserved and buried inside the protein core of G-LC.

3.2.2 Structure and stability of the C_L are altered by the mutation V136G

To decipher the role of the constant domain mutation, the recombinant patient C_L (P-CL), germline C_L (G-CL), and the respective patient and germline LCs (P-LC, G-LC) were investigated. Additionally, the chimeric mutants P-LC G136V (= patient V_L + germline C_L) and G-LC V136G (= germline V_L + patient C_L) were created to obtain more detailed insights into the interplay between the two LC domains. In all of these constructs, the C-terminal cysteine residue at position 214 - normally involved in covalent linkage of LC and heavy chain (HC) or LC homo-dimers via disulfide bond (Murphy and Weaver, 2016) – was mutated to serine to prevent the covalent dimerization. All proteins were recombinantly produced as inclusion bodies in *E. coli*, refolded, and purified. To assess the secondary structure, far UV circular dichroism spectroscopy (FUV CD) was performed (Fig. 26A). All constructs exhibit spectra indicative of a β -sheet structure, typical for immunoglobulin domains (Bork et al., 1994).

However, the secondary structure content varies between the different proteins as can be seen by the differences in molar ellipticity around 200 nm. Especially the P-CL (black curve) exhibits a decreased signal, which points towards a larger portion of disordered segments or partial unfolding. The dashed lines represent spectra of refolded proteins after thermal unfolding. The G-CL (red curve) and G-LC (blue curve) regain their native structure to a large extent after unfolding but not completely. However, it should be noted that thermal unfolding is often accompanied by aggregation which could explain the signal discrepancy between the spectra of the native and refolded germline variants (Benjwal et al., 2006). The remaining proteins, which contain patient-specific mutations in either of the two domains, stay largely unfolded after thermal denaturation and refolding, indicating irreversible unfolding (Fig. 26A). Near UV (NUV) CD spectra represent a protein-specific tertiary structure fingerprint. Since all NUV CD spectra look similar, it can be concluded that the overall structural topology does not differ strongly between the different LCs and C_Ls (Fig. 26B). Therefore, the main differences observed by CD seemingly lie in local secondary structure elements. To investigate protein stability, thermal denaturation curves between 20 and 90 °C followed by CD at 205 nm were recorded (Fig 26C, Table 4). The transition midpoint (T_m) represents the melting temperature at which 50 % of the protein is unfolded. As expected, the patient C_L, V_L, and LC exhibit considerably lower thermal stability compared to their respective germline counterparts with melting point differences (ΔT_m) ranging between approx. 13 - 16 °C (Fig. 26C, Table 4). As for the two chimeric LCs, the P-LC G136V variant is stabilized by approx. 6 °C compared to the P-LC, whereas G-LC V136G is destabilized by roughly 6 °C in comparison to G-LC (Fig. 26C, Table 4). Thus, the chimeric mutants exhibit intermediate T_m values lying in between the stabilities of P-LC and G-LC, which correlates well with the secondary structure content of the LCs (Fig. 26, Table 4). Of note, the derived melting temperatures represent apparent T_m values since the protein unfolding is not fully reversible.

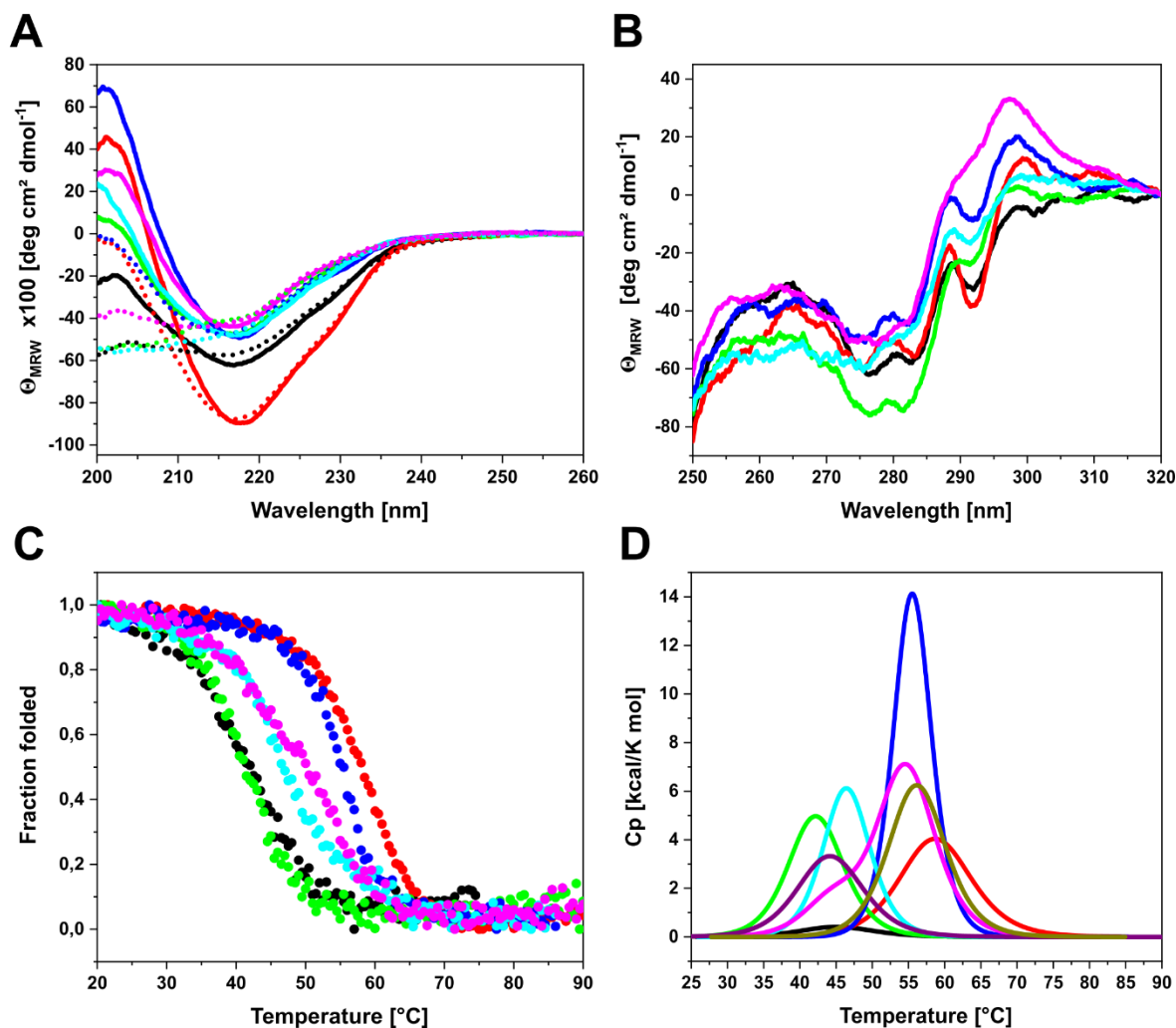


Figure 26: Structural and thermodynamic stability of LCs and CL domains. A) FUV-CD spectra of the native (solid lines) and refolded (dotted lines) proteins. Color code: P-CL in black, G-CL in red, P-LC in green, G-LC in blue, P-LC G136V in cyan, G-LC V136G in magenta. B) NUV-CD spectra of the native proteins representing tertiary structure fingerprints. C) Normalized thermal unfolding transitions followed by CD at 205 nm. The transition midpoint T_m indicates the melting temperature at which 50 % of the protein is unfolded. D) Thermograms obtained by DSC. For comparison, the P-VL (purple) and G-VL (olive) were also included. Unfolding temperatures and enthalpies were determined using a non-two-state fit model.

To characterize the thermodynamic stability of the variants in more detail, differential scanning calorimetry (DSC) was performed (Fig. 26D). This method allows determination of the apparent T_m and of the unfolding enthalpy either calorimetrically (ΔH_{cal}) or via the van't Hoff relation (ΔH_{VH}) by applying a non-two-state fit model (Bruylants et al., 2005). The melting temperatures obtained for the constructs by DSC match the values derived from the thermal unfolding experiments monitored by CD spectroscopy (Fig. 26, Table 4). Generally, all proteins tested unfold in a single transition except for G-LC V136G, which shows separate unfolding peaks for the two constitutive domains (Fig. 26D). For P-LC and G-LC it seems plausible that the T_m values of the individual domains are too close to each other to be resolved by DSC. In the case of P-LC G136V, however, the two domains should exhibit quite different melting points in DSC

experiments, based on the results obtained for the individual domains. Yet, this LC still unfolds in a single transition, which indicates a cooperative interplay between the two domains. Presumably, this cooperative unfolding is not observable in the G-LC V136G due to the loss of V_L - C_L interactions which are disrupted in the mutant. Furthermore, in DSC experiments ΔH_{cal} and ΔH_{VH} should be identical if the protein exhibits an ideal two-state unfolding transition (Johnson, 2013). Discrepancies between the two enthalpy values point towards the existence of intermediate states along the unfolding pathway that divert the unfolding reaction from a simple two-state mechanism (Boone et al., 2013). To test this, the $\Delta H_{cal}/\Delta H_{VH}$ ratios for all samples were determined (Fig. 27). Notably, the ratios of all germline proteins (G-CL, G-VL, G-LC) tend to be closer to 1 than the ratios of the proteins which carry patient mutations (P-CL, P-VL, P-LC, P-LC G136V, G-LC V136G), implying that the mutations in both the patient V_L and patient C_L favor the population of partially unfolded states along the folding trajectory (Fig. 27). Such folding intermediates are known to play a key role in the amyloid pathway of LCs and V_L s (Khurana et al., 2001).

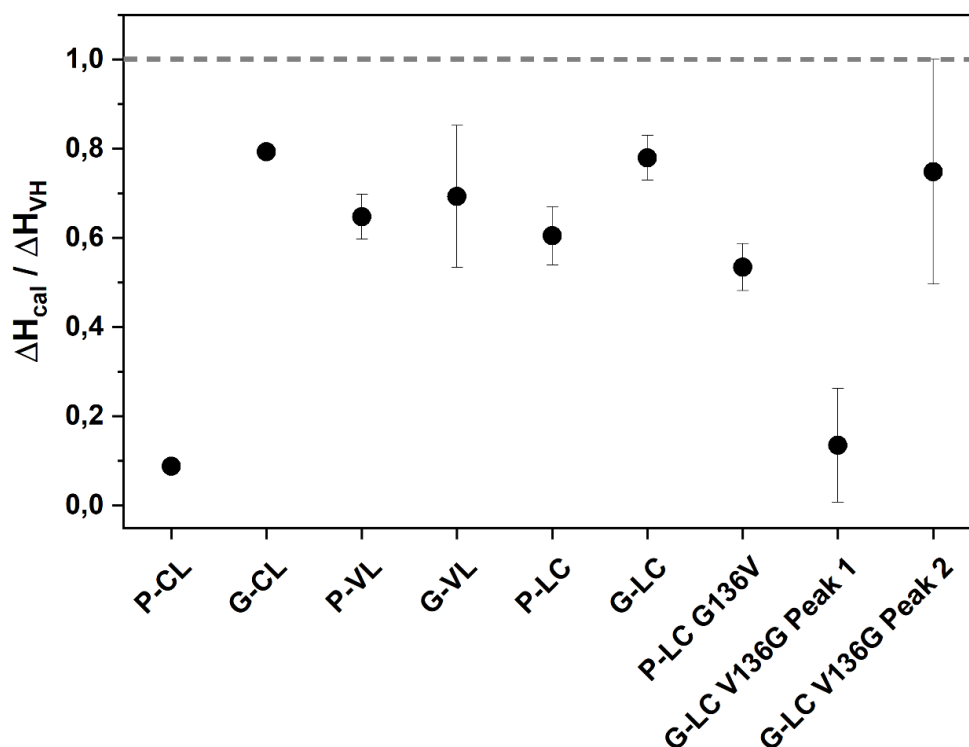


Figure 27: $\Delta H_{cal} / \Delta H_{VH}$ ratios imply the population of partially unfolded states. As indicated, the only protein that exhibited two distinct transitions was G-LC V136G. All other proteins either unfolded in a single, cooperative transition or the two transition temperatures of the individual domains were too close to each other to be resolved by DSC. Theoretically, the fraction of $\Delta H_{cal} / \Delta H_{VH}$ is 1 for an ideal two-state unfolding event (indicated by the dashed grey line). The stronger a value deviates from 1, the more likely partially folded states are being populated among the folding trajectories.

Table 4: Thermodynamic stability parameters determined by thermal unfolding using CD spectroscopy and DSC.

Protein	T _m by CD	T _m by DSC	ΔH _{cal}	ΔH _{VH}
	°C	°C	kcal/mol	kcal/mol
P-CL	41.04 ± 0.31	44.55 ± 0.95	5.19 ± 0.60	58.95 ± 0.64
G-CL	57.78 ± 0.16	58.80 ± 0.01	52.80 ± 0.71	66.60 ± 0.42
P-VL	43.50 ± 0.14	44.15 ± 0.18	42.50 ± 2.97	65.65 ± 0.49
G-VL	56.30 ± 0.11	56.22 ± 0.01	58.95 ± 9.69	85.75 ± 5.73
P-LC	41.38 ± 0.41	42.34 ± 0.11	48.25 ± 3.32	79.95 ± 3.18
G-LC	55.28 ± 0.21	55.58 ± 0.04	97.00 ± 3.54	124.5 ± 3.54
P-LC G136V	47.46 ± 0.19	46.57 ± 0.12	49.90 ± 4.81	93.35 ± 0.21
G-LC V136G Peak 1	49.35 ± 0.23	44.44 ± 0.29	10.41 ± 8.62	84.55 ± 15.91
G-LC V136G Peak 2		54.92 ± 0.28	62.40 ± 16.69	84.45 ± 6.15

All experiments were carried out in duplicates. CD data was analyzed applying a Boltzmann fit, for DSC data a non-two-state fitting model was used.

3.2.3 The mutation V136G causes partial unfolding of the C_L

To gain further insight into the folding and conformational properties of the LCs and especially the mutated C_L domain, fluorescence spectroscopy, including red edge excitation shift (REES), ANS binding, and acrylamide quenching experiments (Figs. 28, 29; Table 5) were used. In tryptophan fluorescence spectra (excitation wavelength = 295 nm), a red shift in the emission maxima and increased fluorescence intensities for constructs carrying patient mutations were observed (Fig. 28A/B). These spectral features hint at partial unfolding, especially regarding the striking difference between the isolated G-CL and P-CL domain spectra (Fig. 28B). This is further confirmed by stronger binding of the dye ANS to P-CL as compared to G-CL indicating higher exposure of hydrophobic patches on the protein surface which suggests partial unfolding (Fig. 28C).

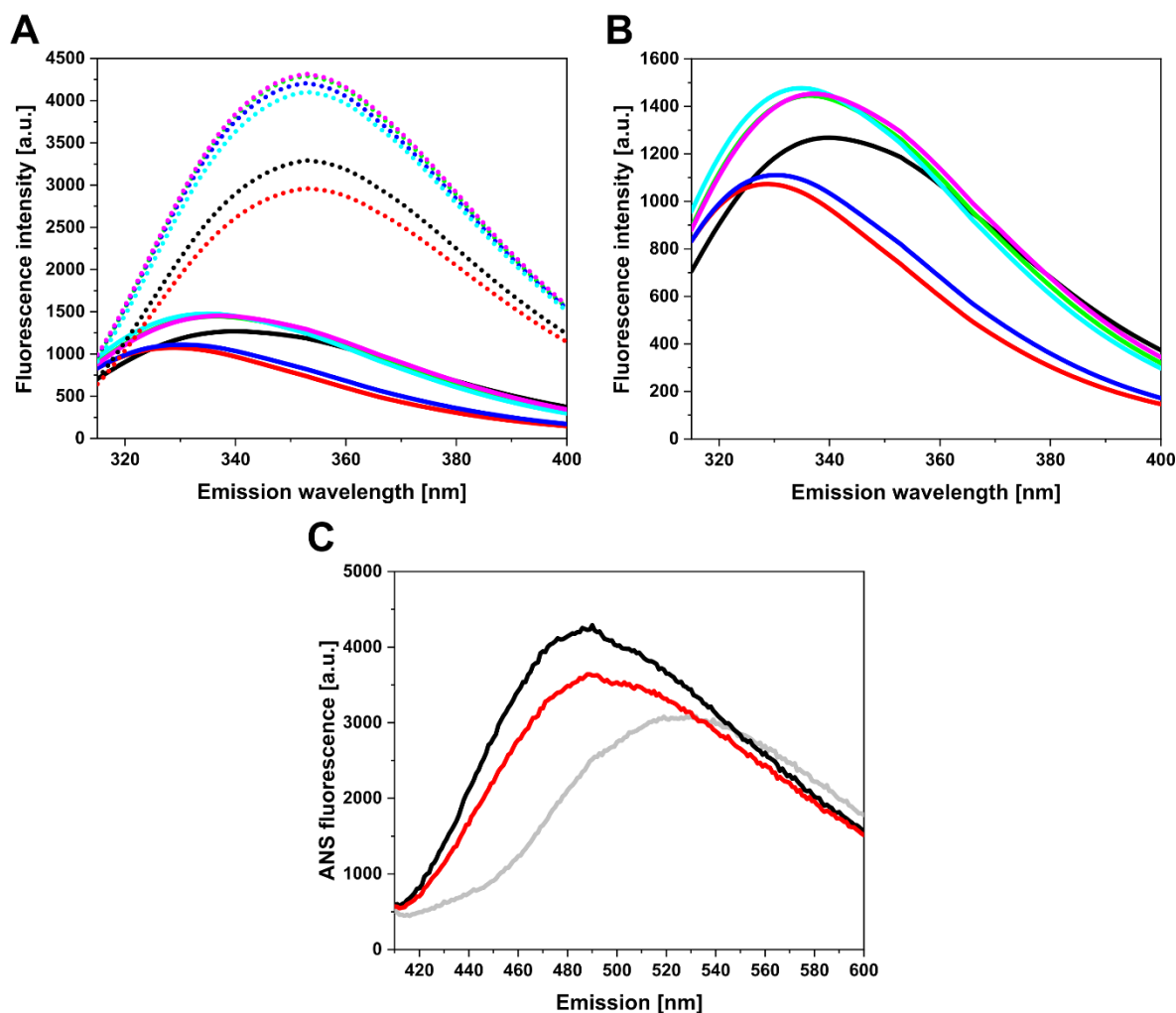


Figure 28: Tryptophan fluorescence and ANS binding show that P-CL is partially unfolded. A) Fluorescence spectra of proteins denatured with 6 M urea (dotted lines) and native proteins (solid lines) in PBS pH 7.4. B) Spectra of the native proteins. Fluorescence spectra were recorded between 315-400 nm with an excitation wavelength of 295 nm at 25 °C. C) ANS binding of P-CL (black) and G-CL (red) with 10 μ M protein and 100 μ M ANS at 25 °C. The wavelength of excitation was 380 nm. The PBS control is shown in grey. Coloring: P-CL in black, G-CL in red, P-LC in green, G-LC in blue, P-LC G136V in cyan, G-LC V136G in magenta.

REES experiments of P-CL and G-CL at 25 °C and 37 °C showed increased center of spectral mass values (CSM) for P-CL which indicates higher solvent exposure of the tryptophan fluorophores and, therefore, also a higher degree of partial unfolding (Fig. 29). Further, the effect of elevated temperature (25 vs. 37 °C) is more pronounced in P-CL than in G-CL (Fig. 29A). Regarding the LCs, the G-LC shows the lowest CSM value, as expected (Fig. 29B/C). P-LC and the two chimeric mutants all exhibit very similar REES profiles at 25 °C, although the difference between starting and end value ($\Delta\text{CSM} = \text{CSM}_{300} - \text{CSM}_{280}$) varies between the curves indicating alterations in conformational dynamics (Catici et al., 2016). In addition, the shift towards higher CSM values is more pronounced at 37 °C for P-LC and G-LC V136G, which demonstrates the beneficial effect of valine 136 on the stability and dynamics of the C_L domain (Fig. 29B/C, Table 5).

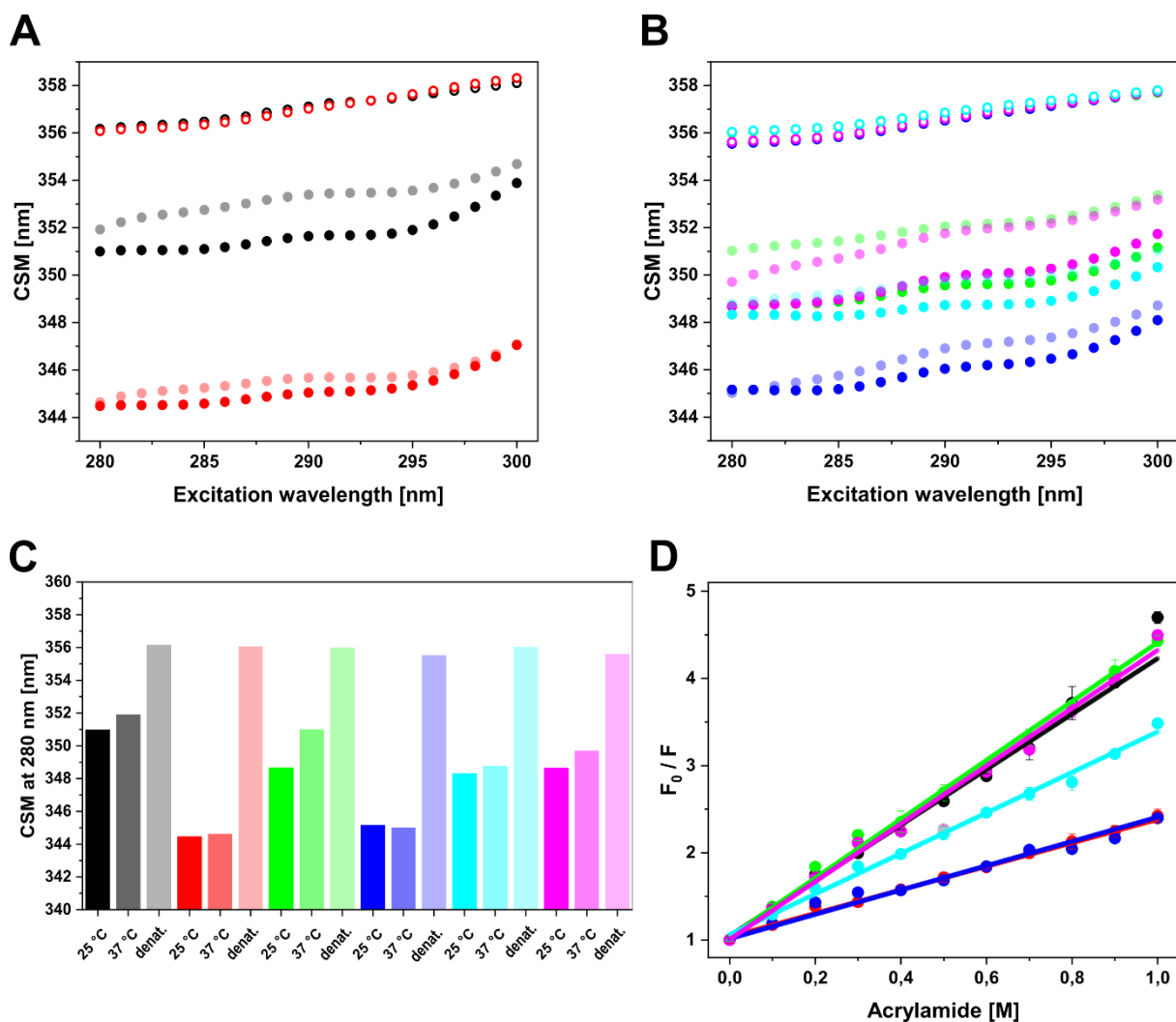


Figure 29: REES fluorescence and acrylamide quenching give insights into folding and conformational flexibility. A) REES data of P-CL (black) and G-CL (red) at 25 °C (solid), 37 °C (light colored), and in the presence of 6 M urea (circled). The CSM value at 280 nm indicates solvent exposure of the tryptophan fluorophores. B) REES data of P-LC (green), G-LC (blue), P-LC G136V (cyan), and G-LC V136G (magenta). Data at 25 °C is shown as solid dots, data at 37 °C as light dots, and data with 6 M urea as circled dots. C) Bar graph the CSM values at 280 nm for all constructs under the three conditions tested. The data in presence of 6 M urea is termed “denat.” for denatured. Coloring as in A and B. D) Acrylamide quenching plots showing F_0/F as a function of acrylamide concentration, where F_0 is the native fluorescence. Coloring is the same as in A, B, and C. The slope of the linear fit represents the Stern-Volmer constant k_{sv} (Table 5). For G-LC V136G (magenta), the data point at 0.5 M was excluded from the fit.

Acrylamide quenching can be used to investigate the accessibility of fluorophores and can thus provide insights into domain architecture and dynamics (Lakowicz, 2006). The Stern-Volmer constant (k_{sv}) is the slope of the linear fitting curves and represents the degree to which tryptophan fluorescence is quenched (Fig. 29D, Table 5). G-CL and G-LC expectedly show the lowest degree of quenching which correlates to their conformational stability. The proteins P-CL, P-LC, and G-LC V136G (i.e., all constructs with a mutated, destabilized C_L domain) exhibit very similar quenching curves and k_{sv} values (Fig. 29D, Table 5). The value for P-LC

G136V lies in between. Therefore, the increased quenching compared to the two germline constructs can be attributed to the effects of the patient-specific mutations in the V_L domain. In conclusion, the fluorescence data point toward partial unfolding of the P-CL and further demonstrate the negative effects of the V136G mutation on the FOR005 patient C_L .

Table 5: CSM values determined by REES fluorescence spectroscopy and Stern-Volmer constants (k_{sv}) derived from acrylamide quenching experiments.

Protein	CSM ₂₈₀ at 25 °C	CSM ₂₈₀ at 37 °C	k_{sv} at 25 °C
	nm	nm	L/mol
P-CL	350.00	351.93	3.18 ± 0.12
G-CL	344.47	344.64	1.35 ± 0.03
P-LC	348.69	351.03	3.38 ± 0.07
G-LC	345.16	345.02	1.39 ± 0.05
P-LC G136V	348.32	348.79	2.33 ± 0.06
G-LC V136G	348.67	349.71	3.32 ± 0.09

The error of the k_{sv} values indicates goodness of fit.

In addition, Molecular Dynamics (MD) simulations on the P-CL and G-CL variants were performed by Prof. Dr. Martin Zacharias from the Chair of Theoretical Biophysics at TUM. Starting from the folded domains, no unfolding or any other major conformational change was observed on the time scale of the MD simulations (500 ns). Both P-CL and G-CL resulted in similar root-mean-square deviation (RMSD) from the start structure and root-mean-square fluctuations (RMSF) (Fig. 30). Residue 136 is located next to a cysteine that is involved in the single disulfide bond in the C_L domain and forms backbone hydrogen bonds to the N-terminal β -strand.

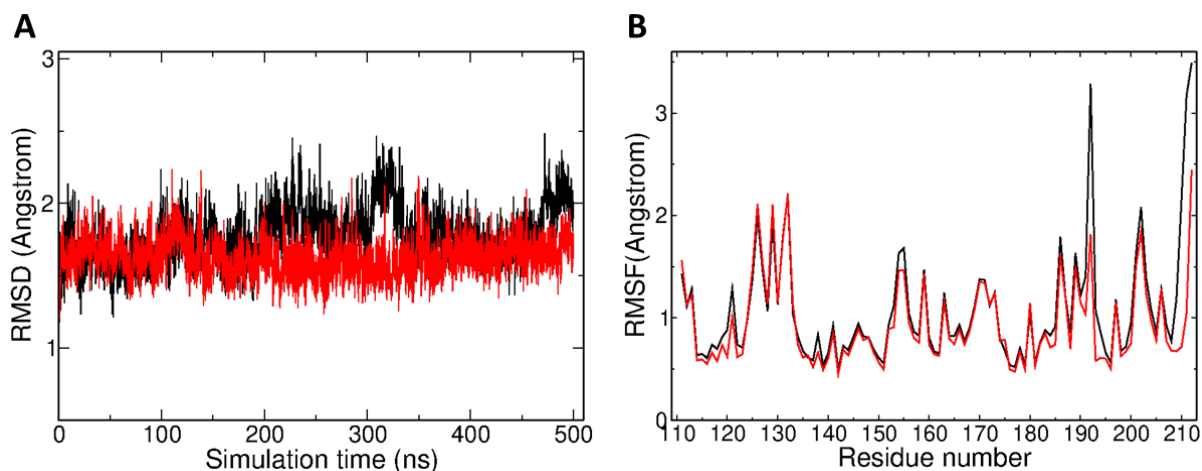


Figure 30: Molecular dynamics simulation of P-CL and G-CL. A) Root mean square deviation (RMSD of non-hydrogen atoms) from the start structure during 500 ns MD simulation time (Coloring: P-CL in black, G-CL in red). B) Root mean square fluctuations of P-CL and G-CL residues with respect to mean structure recorded during MD-simulations (same coloring as in A). Figure kindly provided by Prof. Dr. Martin Zacharias.

To investigate if the greater conformational flexibility of a G136 may affect the interaction with the N-terminal β -strand, which in turn may affect the stability, free energy simulations were performed to partially dissociate the N-terminal β -strand and disrupt the backbone hydrogen bonds with the residues 135-137 (Fig. 31). Indeed, the calculated free energy penalty to locally disrupt the β -sheet segment is lower for the P-CL variant by ~ 1 kcal/mol. This indicates that the V136G mutation in the C_L domain destabilizes the β -sheet segment near the N-terminus, which contributes to a lower stability of the P-CL vs. G-CL.

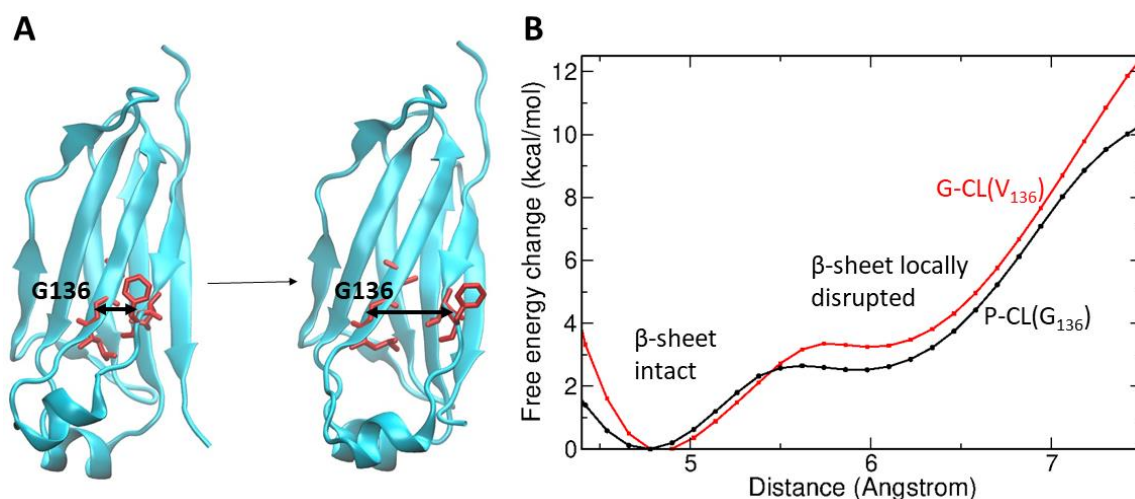


Figure 31: Free energy simulations on P-CL and G-CL. A) Umbrella sampling free energy simulations on partial dissociation of the N-terminal β -strand near residue 136 were performed on the P-CL and G-CL (cartoon illustration) variants. The center of mass (com) distance between backbone atoms of residues 135-137 and residues 120-121 (red sticks) served as reaction coordinate (double arrow). B) Calculated free energy changes (potential of mean force) along the reaction coordinate for P-CL (black line) and G-CL (red line). Figure kindly provided by Prof. Dr. Martin Zacharias.

3.2.4 The full-length LC and the C_L domain do not form amyloid fibrils *in vitro*

As mentioned above, in the majority of cases the fibrillar deposits contain mainly the V_L domain. There are only few cases where full-length LCs or C_L domains undergo amyloid formation (Blancas-Mejía and Ramirez-Alvarado, 2013; Absmeier et al., 2022). The fibrils isolated from FOR005 patient tissue contained only the V_L domain (Annamalai et al., 2017). Here, it was tested whether the P-LC or the G-LC form amyloid fibrils *in vitro* at a physiological pH of 7.4. Further, fibril formation of the germline or patient C_L domains and LCs under destabilizing conditions was investigated. Therefore, Thioflavin T (ThT) binding assays were carried out under continuous shaking at pH 7.4 or pH 6.4 or in the presence of 0.5 mM SDS which can accelerate fibril formation (Pertinhez et al., 2002). However, no fibril formation was observed for any of the constructs under the conditions tested despite their low thermodynamic stabilities (Fig. 32A). Transmission electron microscopy (TEM) of the samples revealed that the proteins form amorphous aggregates instead of fibrils (Fig. 32B). Thus, the presence of the C_L domain and its interactions with the V_L domain alter the folding pathway and seemingly prevent the LCs from accessing the fibrillary state. Although the patient C_L domain exhibits significantly lower stability, it still exerts a protective function, which favors amorphous aggregation over fibril assembly.

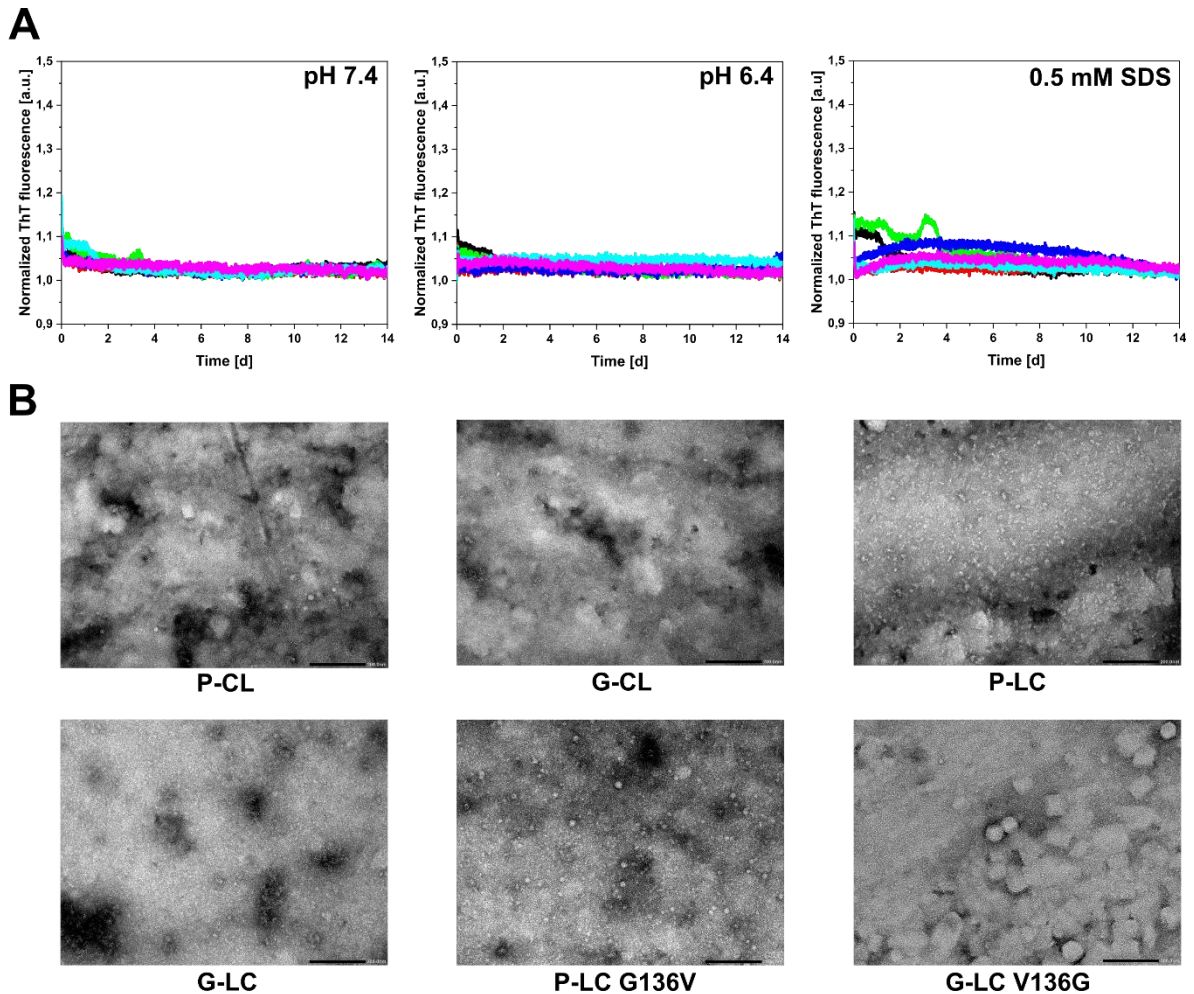


Figure 32: ThT binding kinetics and TEM micrographs demonstrate that the CLs and all LCs are resistant to amyloid formation *in vitro*. A) ThT kinetics were obtained as triplicates over two weeks under continuous shaking in a microplate reader at 37 °C and either pH 7.4 (left panel), pH 6.4 (middle panel), or in the presence of 0.5 mM SDS (right panel). B) TEM micrographs of finished ThT assays at pH 7.4 show the presence of amorphous aggregates and confirm the absence of amyloid fibrils. Coloring: P-CL in black, G-CL in red, P-LC in green, G-LC in blue, P-LC G136V in cyan, G-LC V136G in magenta.

3.2.5 Proteolysis of LCs is mediated by an unstable CL domain

The finding that the FOR005 P-LC does not form fibrils *in vitro* whereas the patient V_L domain does, implies that the LC needs to be cleaved in order to set the destabilized V_L domain free. In general, the proposed proteolytic cleavage as a prerequisite for fibril formation is still subject to debate in AL amyloidosis (Absmeier et al., 2022). For FOR005, limited proteolysis was applied to address this question. The LC proteins were incubated with the model proteases trypsin or proteinase K and the degradation kinetics were followed by SDS-PAGE (Figs. 33, 34). The experiments show that the G-LC is cleaved far more slowly compared to constructs carrying patient mutations in either of the two domains (Fig. 33): the P-LC was cleaved completely within a few minutes, whereas the two chimeric mutants P-LC G136V and G-LC V136G showed slightly increased proteolytic stability although they are also cleaved

completely within 90 minutes. Furthermore, proteolysis of the P-CL and G-CL were also tested since limited proteolysis can also give information about conformational dynamics. Notably, very rapid and complete degradation of the P-CL compared to the G-CL was observed, which is indicative of high structural flexibility, thus further supporting the hypothesis that the isolated P-CL is partially unfolded. Of note, the experiment with G-CL and proteinase K yielded a rather unusual time course with a large, rapid drop in the intact protein fraction down to approximately 40 % rather than a continuous exponential decrease. However, no complete degradation of G-CL by proteinase K was observed (Figs. 33, 34).

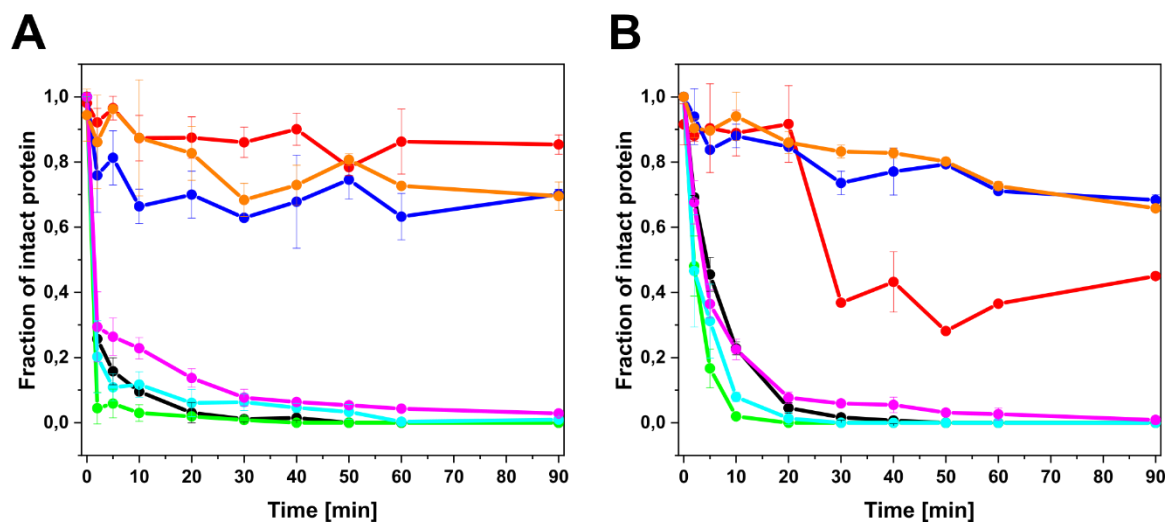


Figure 33: Limited proteolysis experiments with trypsin and proteinase K reveal impact of the C_L mutation on LC cleavage. A) Limited proteolysis at room temperature with trypsin (substrate/protease = 20/1). Coloring: P-CL in black, G-CL in red, P-LC in green, G-LC in blue, P-LC G136V in cyan, G-LC V136G in magenta, covalent P-LC dimer (Bence-Jones protein) in orange. B) Limited proteolysis with proteinase K at room temperature (substrate/protease = 250/1). Coloring as in A.

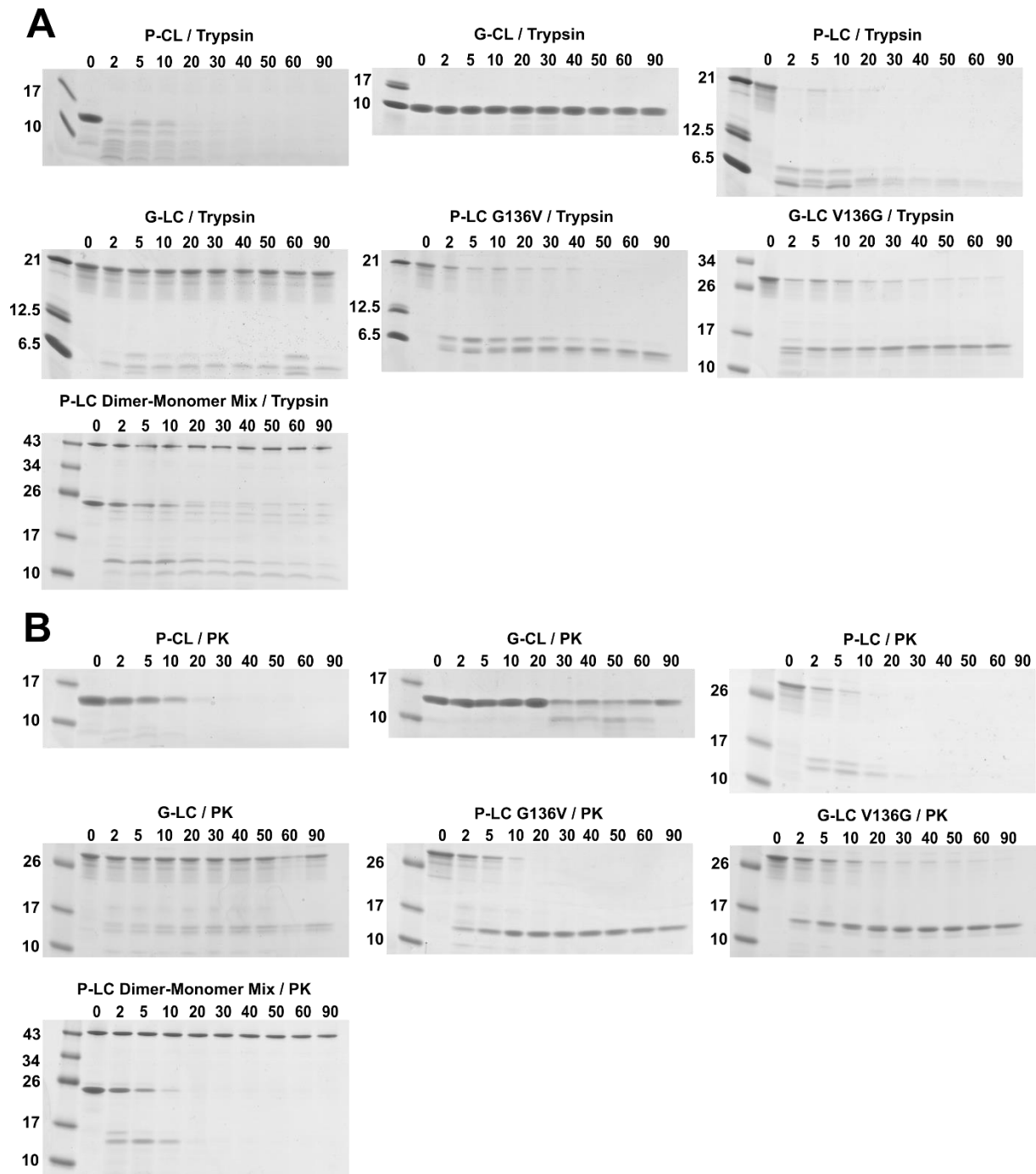


Figure 34: SDS-PAGE gels of limited proteolysis experiments with trypsin and proteinase K. All experiments were performed at room temperature. A) For trypsin, a 20-fold substrate excess was used. B) For experiments with proteinase K, a 250-fold substrate excess was used. Samples were run on SERVA TG Prime 4-20 % SDS gels. Limited proteolysis experiments were carried out in duplicates, however, only one gel is shown per investigated protein.

The proteolysis data clearly show the mutation-dependent differences in susceptibility for the monomeric LCs. However, *in vivo* LCs are also secreted from plasma cells as covalent, disulfide-linked dimers, so-called Bence-Jones proteins (Kaplan et al., 2009; Kaplan et al., 2011). Therefore, an additional construct was created in which the C-terminal cysteine residue of the P-LC was re-introduced to allow dimer formation during *in vitro* refolding. By refolding

from inclusion bodies, a mixture of P-LC dimers and monomers was obtained (roughly a 50/50 distribution) which allowed determining their susceptibility for degradation in the same experiment. The limited proteolysis experiments with the P-LC dimer-monomer mixture demonstrated that the covalent P-LC dimer is far more resistant to proteolytic cleavage than the monomeric P-LC (Figs. 33, 34). Therefore, covalent dimerization exerts a protective function implying that *in vivo* not the Bence-Jones protein but primarily the monomeric LC of FOR005 is responsible for the proteolytic release of the amyloidogenic V_L domain. In this context, characterization of the P-LC dimer-monomer mixture by CD spectroscopy also revealed changes in structure, folding, and stability as a result of the covalent dimerization via the C-terminal disulfide bridge (Fig. 35). FUV and NUV CD spectra of the P-LC dimer-monomer mix show differences in comparison to the completely monomeric P-LC (Fig. 35A/B). Further, the mixture containing dimeric P-LC exhibits a slightly higher T_m value with 44.24 °C and a higher degree of unfolding reversibility (Fig. 35A/C). In conclusion, both the V_L and C_L mutations favor proteolytic cleavage of the monomeric patient LC that - in the case of FOR005 - appears to be a prerequisite for amyloid aggregation. Strikingly, the covalent P-LC dimer is protected from proteolytic cleavage *in vitro*, which potentially relies on small structural changes due to the covalent dimerization.

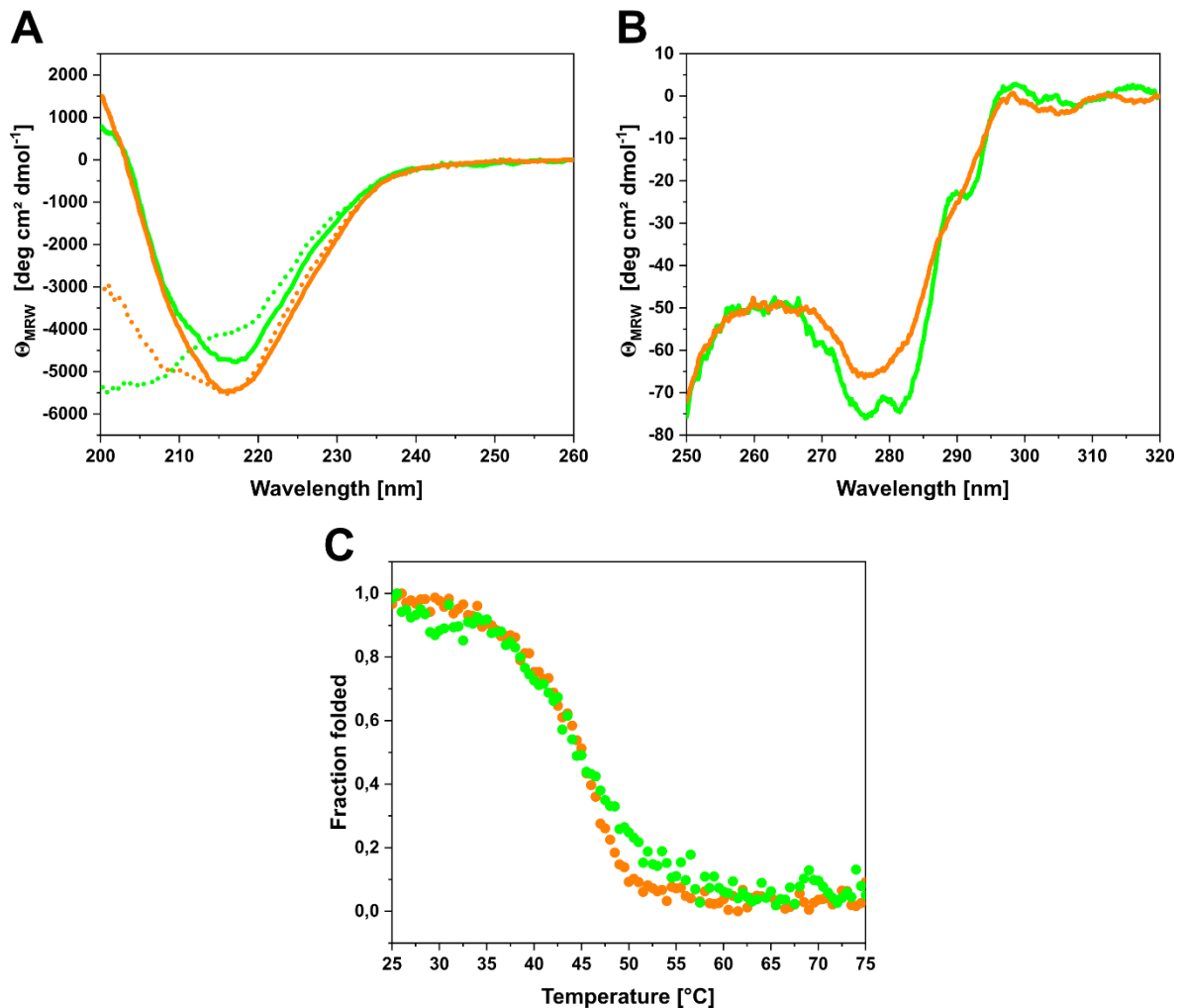


Figure 35: CD spectroscopy of the P-LC dimer/monomer mixture reveals dimerization-induced differences in structure and stability. A) FUV-CD spectrum of the native (solid orange line) and refolded (dotted orange line) P-LC dimer/monomer mixture. Respective spectra of the regular P-LC monomer are shown in green. FUV spectra were recorded at 0.1 mg/ml. The different spectra indicate altered secondary structure and partially reversible unfolding of P-LC due to covalent dimerization. B) NUV-CD spectrum of the P-LC monomer/dimer mix (orange) and the P-LC monomer (green). NUV spectra were recorded at 0.5 mg/ml. C) Thermal denaturation of P-LC monomer/dimer mix followed by CD spectroscopy at 205 nm. Thermal unfolding was recorded at 0.1 mg/ml protein concentration between 20-75 °C. The melting temperature T_m was determined to be 44.24 ± 0.28 °C by Boltzmann fit. Therefore, the dimer/monomer mix exhibits slightly higher apparent thermal stability than the completely monomeric P-LC ($T_m = 41.38 \pm 0.41$ °C).

3.2.6 The C_L mutation impairs homo-dimerization

Apart from stability, partial unfolding, and proteolytic cleavage, the LC quaternary structure also plays a key role in the onset of AL amyloidosis (Absmeier et al., 2022). As mentioned above, free LCs can either be secreted by plasma cells as covalently linked homo-dimers, so called Bence-Jones proteins, or as monomeric LCs (Kaplan et al., 2009; Kaplan et al., 2011). Considering the complete structure of the template F_{ab} fragment 5BV7 that was used for

homology modeling of the LC structures, positioning of the Val136 residue in the HC-LC association interface is observable (Fig. 36).

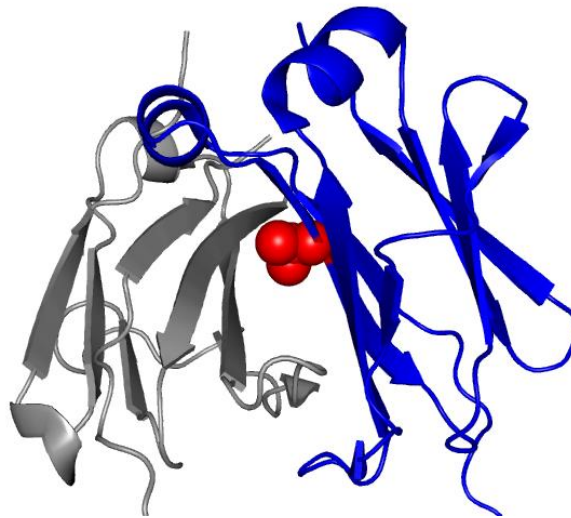


Fig 36: The crystal structure of F_{ab} fragment 5BV7 shows that Val136 is located at the HC-LC association interface. The C_{H1} is colored in grey, the C_L is colored in blue, and the valine residue at position 136 is depicted as red spheres. The location of the residue suggests that the patient mutation V136G affects LC homo-dimerization.

Therefore, the dimerization propensity of FOR005 constructs was investigated by analytical ultracentrifugation (AUC) at varying protein concentrations (Fig. 37). In all proteins analyzed by AUC, the C-terminal cysteine residue was mutated to serine, so that no intermolecular disulfide formation could occur. Thus, only non-covalent association was monitored reflecting the interactions of the dimer interfaces. Both isolated C_L domains and V_L domains stay monomeric regardless of the concentration used (Fig. 37). Of the four full-length LCs, P-LC shows the lowest dimerization propensity with an apparent K_D of 62.92 ± 0.07 μM (Fig. 38, green curves). In contrast, G-LC preferentially forms non-covalent dimers already at low concentrations with a significant shift towards dimer formation with increasing concentrations. The apparent K_D of G-LC is 6.09 ± 1.08 μM (Fig. 38, dark blue). The chimeric mutant P-LC G136V also exhibits a concentration-dependent monomer/dimer equilibrium with a K_D of 12.97 ± 2.28 μM, thus the propensity of homo-dimerization is lower than in G-LC. Regarding the variant G-LC V136G, only dimerization at relatively high protein concentrations was observed, similar to the behavior of P-LC, resulting in an apparent dimerization K_D of 53.67 ± 0.01 μM. These results suggest that the V136G mutation disrupts the dimer interface and thereby prevents homo-dimerization. However, the data also imply that both V_L and C_L domain contribute to the dimer interface since dimerization of the P-LC G136V variant is not fully restored when compared to G-LC and G-LC V136G shows higher dimerization propensity than P-LC. Thus, the patient-specific mutations in the V_L domain also diminish the dimerization

propensity, although the impact of the C_L domain on quaternary structure is considerably more pronounced.

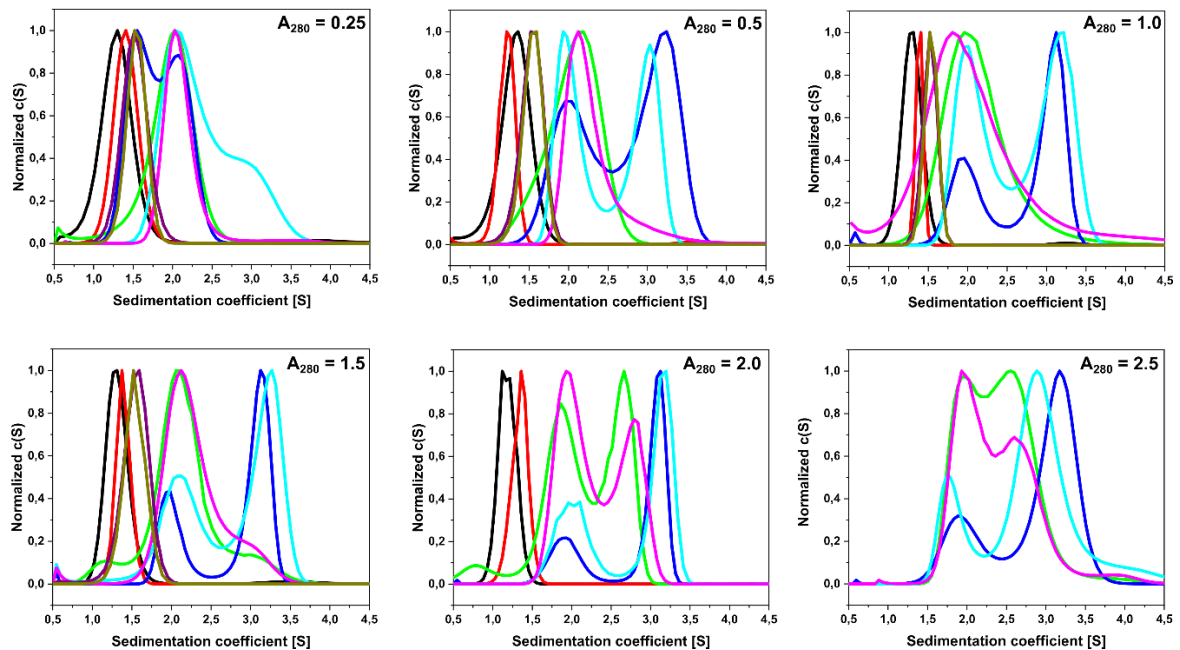


Figure 37: AUC shows the differences in dimerization propensity between the investigated variants. AUC runs at A₂₈₀ = 0.25/ 0.5/ 1.0/ 1.5/ 2.0/ 2.5. In the last run (A₂₈₀ = 2.5) the single V_L and C_L domains were not investigated. (Coloring: P-CL in black, G-CL in red, P-LC in green, G-LC in blue, P-LC G136V in cyan, G-LC V136G in magenta, P-VL in purple, G-VL in olive). For data analysis Sedfit was used.

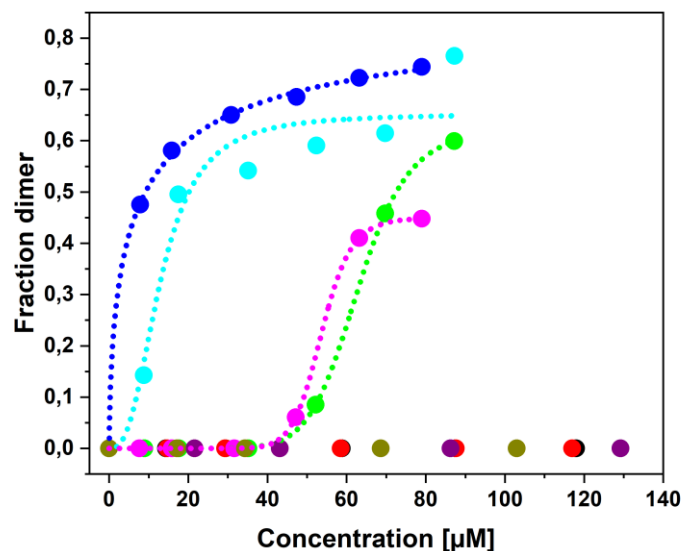


Figure 38: The C_L mutation leads to diminished LC homo-dimer affinity. The fraction of non-covalent dimers depends on the protein concentration. Coloring: P-CL black, G-CL red, P-VL purple, G-VL olive, P-LC green, G-LC blue, P-LC G136V cyan, and G-LC V136G magenta. Data was fitted using the Hill equation to determine the apparent K_D values of the non-covalent LC dimerization.

3.2.7 Discussion

AL amyloidosis is a challenging disease as each patient has a unique precursor LC sequence and point mutations. Accordingly, there are still many open questions about the mechanisms that underlie AL onset and progression. The large sequence diversity of precursor LCs also complicates the development of novel diagnostic and therapeutic approaches (Absmeier et al., 2022). Further, the amyloid deposits in patients can contain either entire LCs or different fragments thereof, i.e., the V_L or the C_L (Blancas-Mejía and Ramirez-Alvarado, 2013). This raises the questions when proteolytic cleavage takes place and which factors determine the occurrence of such endoproteolysis events (Absmeier et al., 2022). In the case of FOR005, the patient's amyloid fibrils are made up only by the V_L domain which carries five mutations (Chapter 3.1). FOR005, however, also carries a mutation in the conserved C_L domain, which is rare since it is usually not subject to somatic hypermutation during B cell maturation (Murphy and Weaver, 2016). Therefore, the influence of C_L mutations and the C_L in general on the properties of the LC and its involvement in disease onset is still poorly understood.

The ConSurf web tool identified the valine at position 136 of the LC to be highly conserved and to be a buried residue which means it is most likely an integral residue for the protein's structure and stability (Ashkenazy et al., 2010). Thus, it was expected that mutation of the valine to glycine, as it occurs in the patient's LC, could have a pronounced effect on the properties of the LC. Indeed, the structure and the stability of the LCs containing the mutation and of the isolated P-CL were found to be strongly affected. The CD spectra show reduced secondary structure content for variants containing the V136G substitution and decreased thermodynamic stability. Further, unfolding reversibility after thermal unfolding is impaired by the presence of the C_L mutation. Comparative MD free energy simulations on the P-CL and G-CL domains indicate a larger free energy penalty of disrupting the β -sheet interaction of residues around G136 with the N-terminal β -strand compared to the G-CL with a V136. However, the calculated magnitude of the effect may not be sufficient to explain the significantly lower stability of the P-CL vs. G-CL. It might be possible that the G136 mutation in P-CL allows formation of an increased number of stable unfolded or misfolded conformations that overall shift the equilibrium towards unfolded states. Such states cannot be sampled on the time scale of the present MD simulations. Despite the structural alterations and the resulting destabilization, the mutant C_L still exerts a protective function regarding amyloid fibril formation. In the ThT assays, none of the LC constructs formed fibrils even under destabilizing conditions (lowered pH, addition of SDS) while the patient V_L readily forms fibrils *in vitro* (Chapter 3.1). This phenomenon has been reported for other cases before (Weber et al., 2018; Kazman et al., 2020), yet, there are also cases where full-length LCs readily aggregate into fibrils (Klafki et al., 1992; Lavatelli et al., 2008; Andrich et al., 2017). In the case of FOR005,

however, all LC variants and both C_Ls formed only amorphous aggregates as visualized by TEM. Therefore, the presence of the C_L domain diverts the LC from the amyloid pathway in favor of amorphous aggregation as these two pathways of protein aggregation can be viewed as competing processes (Adachi et al., 2018).

This protective effect of the C_L domain likely relies on interdomain contacts between C_L and V_L which are crucial regarding the biophysical properties of LCs (Klimtchuk et al., 2010). Rennella et al. have shown that V_L-C_L interactions can have a strong influence on the unfolding and aggregation pathway of an LC and that even a destabilized and partially unfolded C_L can still protect the LC from amyloid formation, or at least slow down the process considerably (Rennella et al., 2019a). A previous study on the LC of the murine antibody MAK33 further demonstrated that amyloidogenic mutations in the V_L domain can shift V_L-C_L interactions from native to non-native thereby destabilizing the entire LC and abolishing the C_L's protective function (Weber et al., 2020). Here, the effects of V_L-C_L interactions were also observed and it is shown that partial unfolding of the mutated C_L domain abrogates inter-domain contacts and their cooperative behavior as seen by the DSC experiments. However, even when the domain interactions are lost, the LCs do not enter the fibril pathway which could be explained by the C_L's low intrinsic amyloid propensity which outweighs that of the V_L (Rennella et al., 2019a).

The influence of interdomain contacts also becomes apparent when proteolysis is considered. The general importance of proteolytic cleavage of LCs in AL amyloidosis is still subject to debate but increasing evidence suggests a case- and sequence-dependent phenomenon (Absmeier et al., 2022). The FOR005 P-LC monomer and the two chimeric mutants are readily cleaved by model proteases. The sizes of the resulting fragments correspond roughly to the two constituent domains of the LC, implying cleavage in the linker region. Therefore, it looks as though inter-domain contacts also govern the accessibility of the LC linker for endoproteases since proteolysis is observed as soon as one of the two domains carries destabilizing substitutions. Only G-LC, in which neither of the domain is mutated, is resistant to proteolytic cleavage. In conclusion, no matter how intrinsically stable the V_L or C_L is, it cannot prevent cleavage of the full-length LC if the other domain is sufficiently destabilized by mutations. Another explanation involves the population of partially folded intermediates since Morgan and coworkers demonstrated that partially folded LCs are far more prone to proteolytic cleavage than natively folded ones and that transient unfolding and incomplete refolding of LCs can promote endoproteolysis (Morgan et al., 2017). The DSC data shows that P-LC and both chimeric mutants are less likely to exhibit two-state unfolding than G-LC meaning they more readily populate folding intermediates. Thus, if one domain carries unfavorable mutations, the population of partially folded intermediates is enhanced, kinetic stability is decreased, and proteolysis can occur more easily. Conclusively, the mutations of the V_L and

C_L strongly favor the proteolytic processing of FOR005 P-LC giving rise to its amyloidogenic V_L domain.

Proteolytic cleavage apparently plays a major role in the disease pathway of FOR005. However, it occurs only for the monomeric LC whereas the covalently linked P-LC dimer is protected from proteolysis. Thus, a direct link between quaternary structure and proteolytic susceptibility exists. The *in vitro* data suggests that in FOR005, proteolytic cleavage has to occur to allow amyloid formation of the V_L and that disulfide-bridged P-LC homo-dimers (Bence Jones dimers) are not susceptible to proteolysis. In consequence, it can be hypothesized that impaired LC dimer association in the endoplasmic reticulum of the plasma cell is a critical step in the disease pathway since it likely leads to enhanced secretion of P-LC monomers, which are seemingly the decisive species. Accordingly, the mutations – especially V136G – were found to affect the homo-dimer association of the FOR005 LC to a large extent. While the single, isolated V_L and C_L domains are not able to dimerize at the concentrations investigated, the full-length LCs exhibited concentration-dependent monomer-dimer equilibria. This demonstrates that LC dimerization depends on both the V_L-V_L interface and the C_L-C_L interface which are interdependent and can act cooperatively (Rennella et al., 2019a). The dimerization propensities of the P-LC ($K_D = 62.92 \mu\text{M}$) and the G-LC V136G ($K_D = 53.67 \mu\text{M}$) were considerably lower than those of the G-LC ($K_D = 6.09 \mu\text{M}$) and the P-LC G136V ($K_D = 12.97 \mu\text{M}$) which shows the importance of the natively folded C_L domain for proper LC dimer association. Previous studies had demonstrated the importance of homo-dimerization in AL amyloidosis and it is known to protect LCs and V_Ls from amyloid formation (Brumshtein et al., 2014; Wolwertz et al., 2016b, 2016a; Nawata et al., 2017). Therefore, the V136G mutation in the P-LC paves the way for amyloid formation by strongly diminishing the dimerization propensity and, thereby, abrogating an important protective feature of the LC.

The preferential cleavage of LC monomers by proteases as compared to LC homo-dimers has already been implied in a previous study (Morgan and Kelly, 2016). Furthermore, secretion of monomeric LCs from plasma cells even in healthy individuals has also been reported (Kaplan et al., 2009; Kaplan et al., 2011). To integrate our findings for FOR005 into a general, stepwise disease mechanism, the following scenario is proposed: (I) the mutation in the C_L domain together with V_L mutations disrupt the LC dimer interface, most likely leading to the secretion of LC monomers from plasma cells; (II) these LC monomers are prone to endoproteolysis since the mutations in V_L and C_L weaken the stabilizing inter-domain interactions and domain cooperativity; (III) through proteolytic processing the highly dynamic and aggregation-prone V_L is set free and is now able to form amyloid fibrils since protection by the C_L is abolished. The stepwise disease mechanism is summarized in a pathway model in Figure 39.

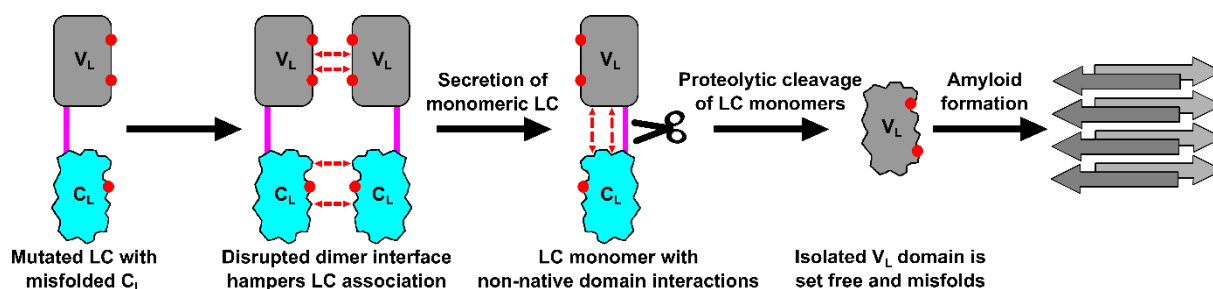


Figure 39: Pathway scheme representing the mechanistic steps in the disease progression of FOR005. The V_L domain is shown in grey, the C_L in cyan, and the linker connecting the two domains is drawn in magenta. The point mutations are indicated as red dots on the LC domains, the dashed red arrows represent non-native interactions or disrupted native interactions, respectively.

However, these proposed mechanistic steps also give rise to new, open questions regarding the disease pathway. Secretion of LC monomers from plasma cells is a known phenomenon, yet our data imply partial unfolding of the C_L . It is unknown how and why such a partially unfolded LC can escape ER quality control (ERQC) and subsequent degradation (Absmeier et al., 2022). In this context, pharmacological targeting of the ER protein homeostasis network has previously been shown to compensate the failed cellular retention of partially folded LCs in AL amyloidosis (Rius et al., 2021). Furthermore, the question arises which protease(s) is/are involved in cleavage of the monomeric LCs in the blood stream and why the destabilized V_L is not degraded or subjected to renal clearance afterwards. Accordingly, interactions with plasma proteins and components of extracellular matrices likely play a key role in the onset of AL amyloidosis (Hawkins, 1988; Wyatt et al., 2013). Additionally, our findings shed light on another problem: in the case of FOR005, the decisive species seems to be the LC monomer since it is highly susceptible to proteolytic cleavage due to the loss of the protective stabilization by LC homo-dimerization. Thus, the recently employed therapeutic approach to stabilize dimeric LCs by small molecules – while elegant and effective for dimers - could potentially be rendered ineffective in a case like FOR005 where monomeric LCs likely need to be targeted instead (Kaplan et al., 2009; Kaplan et al., 2011; Morgan et al., 2019; Yan et al., 2020; Yan et al., 2021). In the future, these and other questions in the context of the onset and progression of AL amyloidosis need to be addressed to gain further mechanistic insight required for progress in AL diagnosis and therapeutics.

4. Conclusions and Perspectives

Light chain (AL) amyloidosis is a debilitating disease in which mutant antibody LCs, secreted by aberrant plasma cell clones, misfold and form insoluble fibrils which can be deposited in various organs (Merlini et al., 2018). The aim of this thesis was to acquire new insights into the mechanisms underlying systemic AL amyloidosis by detailed analysis of the patient-derived sequence FOR005. The FOR005 project was divided into two parts: the first study investigated the effects of individual, patient-specific point mutations of the amyloid-forming V_L domain. The second part focused on a mutation in the constant domain and its impact on the V_L - C_L domain interplay and consequences for the proteolytic release of the amyloidogenic V_L domain by LC cleavage.

The first part of the project revolved around the V_L domain of FOR005 because the amyloid fibrils derived from patient tissue were entirely built up of this fragment (Annamalai et al., 2017). The link between patient-specific mutations in the antibody V_L domain and the disease are widely established (Absmeier et al., 2022), but the causative molecular mechanisms by which these mutations induce misfolding and amyloid aggregation are still poorly understood. Further, it is not possible to predict the effect of a specific mutation just from primary sequence information. Therefore, the patient V_L domain was compared to its non-amyloidogenic germline counterpart and it was found that two out of the five mutations – namely G49R and G94A - destabilize the protein and induce amyloid fibril formation. Unexpectedly, the decisive mutations are found in the highly variable complementarity determining regions (CDRs) of the patient V_L domain, yet they show a remarkable impact on the dynamics of conserved framework regions including the protein core. The underlying cause of this effect is a deviation from the canonical CDR structures in the CDR2 and CDR3 loops. This deviating backbone conformation and the resulting structural strain are induced by the acquired, patient-specific substitutions. Thus, the mutations destabilize the V_L domain by increasing the framework dynamics, enabling their conformational rearrangement into the fibril core which shows an unexpected influence on CDR-framework interactions. However, the amyloid-driving mutation G49R is also involved in a specific side chain interaction within the fibril core structure representing another mechanism by which mutations promote fibrillar aggregation. Therefore, FOR005 represents an example in which different mutations pave the way towards amyloid formation by either destabilizing the precursor LC or by stabilizing the final fibril core structure or even both. Furthermore, the concept that thermodynamic stability is an important, yet not the only crucial molecular determinant in the LC fibril formation pathway is further confirmed by these findings. Additionally, the study provides mechanistic information on the limitations of CDR flexibility, and how mutations in the hypervariable CDRs affect the biophysical properties of a V_L domain and its integrity.

While the results obtained for FOR005 provide novel insights regarding the role of V_L mutations, it will be important in the future to investigate further patient sequences and characterize the effects that specific substitutions can have on the features of an LC or V_L domain. Ultimately, the detailed analysis of patient mutations will contribute to a deep understanding of the disease mechanisms which is crucial for the development of novel drugs and therapeutic options. Further, the information of such patient sequence analyses can be used to develop new diagnostic tools for AL amyloidosis which is of great importance considering the problems that arise from false and late diagnoses. For instance, sequencing of FLCs from the blood of MM patients would allow identification of known amyloid-driving point mutations and, therefore, enable risk assessment for development of AL amyloidosis.

In the second part of the FOR005 project, the focus was set on the full-length LC and specifically on the C_L domain which also contains a point mutation. The mechanistic role of C_L mutations in the context of AL amyloidosis is still poorly understood because such mutations are quite rare. During this thesis work, the patient and germline C_L domains, full-length patient and germline LCs, as well as chimeric LC variants (i.e., patient V_L with germline C_L and vice versa) were investigated. The biochemical and biophysical analyses revealed the influence of the C_L mutation on the amyloid pathway of FOR005: (i) the mutation leads to impaired folding and stability of the patient C_L by weakening secondary structure elements of the domain; (ii) the LC dimer interface is disrupted and LC homo-dimerization is prevented, thus abolishing a protective mechanism; (iii) the C_L mutation promotes proteolytic cleavage of the monomeric LC by disrupting native V_L - C_L domain interactions, thus releasing the isolated, amyloidogenic V_L domain while dimeric LCs (Bence-Jones dimers) are not susceptible to proteolysis. The inability of full-length LCs to form amyloid fibrils even in the presence of a destabilized C_L domain and the enhanced proteolytic cleavage of LC monomers point towards a mechanism in which proteolysis precedes amyloid formation since the C_L domain exerts a protective function.

The findings of the FOR005 C_L domain study contribute to the understanding of the causative mechanisms and steps governing the onset of AL amyloidosis. However, the results also raise new, important questions that will spark future research efforts. The role of transient V_L - C_L domain interactions regarding the conformation and proteolytic susceptibility of amyloidogenic LCs can, for instance, be studied by solution-state NMR spectroscopy. Such high-resolution structural data can provide novel insights into antibody LC architecture in general and can contribute to the development of compounds that stabilize LCs by binding the native V_L - C_L interface. Further, the role of the ER protein homeostasis network and ERQC components can be investigated both by *in vitro* and *in vivo* methods to elucidate the mechanism by which monomeric, partially misfolded LCs evade physiological quality control steps in order to be

secreted from plasma cells. This will pave the way for a connection of biophysics of amyloid formation with cell biology, thus enabling new insights into the biogenesis of secretory protein and providing a possible new target for pharmacological intervention. In addition, the elucidation of the role of proteolysis in AL amyloidosis will be an important topic for future research. Concerning the FOR005 project, there is now strong evidence supporting an initial proteolytic cleavage step that is followed by aggregation of an isolated V_L. Nonetheless, the situation is case-dependent and possibly different for other patient sequences (Absmeier et al., 2022). Obtaining a more comprehensive picture about the role of proteases in the disease pathway of AL amyloidosis will be critical in the future. Identification and pharmacological targeting of these proteases, however, are challenging tasks.

In summary, this thesis represents an in-depth analysis of the disease pathway of the patient sequence FOR005 starting from the full-length LC that is affected by a C_L mutation that impairs LC dimer association, thereby enhancing secretion of monomeric LCs which are susceptible to proteolysis. This leads to the release of an isolated V_L domain that contains destabilizing CDR mutations which alter CDR loop conformations, thermodynamic stability, and structural dynamics, thereby inducing the amyloid aggregation pathway. The results of this thesis can contribute to a better understanding of underlying disease mechanisms and accelerate the development of novel therapeutic options and diagnostic tools.

5. References

- Absmeier, R.M., Rottenaicher, G.J., Svilenov, H.L., Kazman, P., and Buchner, J. (2022). Antibodies gone bad - the molecular mechanism of light chain amyloidosis. *The FEBS Journal*, DOI: 10.1111/febs.16390
- Adachi, M., Noji, M., So, M., Sasahara, K., Kardos, J., Naiki, H., and Goto, Y. (2018). Aggregation-phase diagrams of β 2-microglobulin reveal temperature and salt effects on competitive formation of amyloids versus amorphous aggregates. *Journal of Biological Chemistry* 293: 14775–14785.
- Al-Lazikani, B., Lesk, A.M., and Chothia, C. (1997). Standard conformations for the canonical structures of immunoglobulins. *Journal of Molecular Biology* 273: 927–948.
- Alt, F.W., Blackwell, T.K., and Yancopoulos, G.D. (1987). Development of the Primary Antibody Repertoire. *Science* 238: 1079–1087.
- Andrich, K., Hegenbart, U., Kimmich, C., Kedia, N., Bergen, H.R., Schönland, S., Wanker, E., and Bieschke, J. (2017). Aggregation of Full-length Immunoglobulin Light Chains from Systemic Light Chain Amyloidosis (AL) Patients Is Remodeled by Epigallocatechin-3-gallate. *Journal of Biological Chemistry* 292: 2328–2344.
- Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science* 181: 223–230.
- Annamalai, K., Liberta, F., Vielberg, M.-T., Close, W., Lilie, H., Gührs, K.-H., Schierhorn, A., Koehler, R., Schmidt, A., and Haupt, C., et al. (2017). Common Fibril Structures Imply Systemically Conserved Protein Misfolding Pathways In Vivo. *Angew. Chem.* 129: 7618–7622.
- Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 38: W529-33.
- Astbury, W.T., Dickinson, S., and Bailey, K. (1935). The X-ray interpretation of denaturation and the structure of the seed globulins. *Biochemical Journal* 29: 2351.
- Aviram, N. and Schuldiner, M. (2017). Targeting and translocation of proteins to the endoplasmic reticulum at a glance. *J Cell Sci* 130: 4079–4085.
- Baden, E.M., Owen, B.A.L., Peterson, F.C., Volkman, B.F., Ramirez-Alvarado, M., and Thompson, J.R. (2008a). Altered dimer interface decreases stability in an amyloidogenic protein. *Journal of Biological Chemistry* 283: 15853–15860.
- Baden, E.M., Randles, E.G., Aboagye, A.K., Thompson, J.R., and Ramirez-Alvarado, M. (2008b). Structural insights into the role of mutations in amyloidogenesis. *Journal of Biological Chemistry* 283: 30950–30956.

- Balchin, D., Hayer-Hartl, M., and Hartl, F.U. (2016). In vivo aspects of protein folding and quality control. *Science* 353: aac4354.
- Balchin, D., Hayer-Hartl, M., and Hartl, F.U. (2020). Recent advances in understanding catalysis of protein folding by molecular chaperones. *FEBS Letters* 594: 2770–2781.
- Baldwin, A.J., Knowles, T.P.J., Tartaglia, G.G., Fitzpatrick, A.W., Devlin, G.L., Shammas, S.L., Waudby, C.A., Mossuto, M.F., Meehan, S., and Gras, S.L. (2011). Metastability of native proteins and the phenomenon of amyloid formation. *Journal of the American Chemical Society* 133: 14160–14163.
- Baldwin, R.L. and Rose, G.D. (2013). Molten globules, entropy-driven conformational change and protein folding. *Current Opinion in Structural Biology* 23: 4–10.
- Benjwal, S., Verma, S., Röhm, K.-H., and Gursky, O. (2006). Monitoring protein aggregation during thermal unfolding in circular dichroism experiments. *Protein Science* 15: 635–639.
- Bernard, O., Hozumi, N., and Tonegawa, S. (1978). Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell* 15: 1133–1144.
- Biebl, M.M. and Buchner, J. (2019). Structure, Function, and Regulation of the Hsp90 Machinery. *Cold Spring Harbor Perspectives in Biology* 11.
- Biebl, M.M., Delhommel, F., Faust, O., Zak, K.M., Agam, G., Guo, X., Mühlhofer, M., Dahiya, V., Hillebrand, D., and Popowicz, G.M., et al. (2022). NudC guides client transfer between the Hsp40/70 and Hsp90 chaperone systems. *Molecular Cell* 82: 555-569.e7.
- Blancas-Mejía, L.M., Misra, P., Dick, C.J., Cooper, S.A., Redhage, K.R., Bergman, M.R., Jordan, T.L., Maar, K., and Ramirez-Alvarado, M. (2018). Immunoglobulin light chain amyloid aggregation. *Chemical Communications* 54: 10664–10674.
- Blancas-Mejía, L.M., Hammernik, J., Marin-Argany, M., and Ramirez-Alvarado, M. (2015a). Differential Effects on Light Chain Amyloid Formation Depend on Mutations and Type of Glycosaminoglycans. *Journal of Biological Chemistry* 290: 4953–4965.
- Blancas-Mejía, L.M., Horn, T.J., Marin-Argany, M., Auton, M., Tischer, A., and Ramirez-Alvarado, M. (2015b). Thermodynamic and fibril formation studies of full length immunoglobulin light chain AL-09 and its germline protein using scan rate dependent thermal unfolding. *Biophysical Chemistry* 207: 13–20.
- Blancas-Mejía, L.M. and Ramirez-Alvarado, M. (2013). Systemic amyloidoses. *Annual Review of Biochemistry* 82: 745–774.
- Blancas-Mejía, L.M., Tischer, A., Thompson, J.R., Tai, J., Wang, L., Auton, M., and Ramirez-Alvarado, M. (2014). Kinetic control in protein folding for light chain amyloidosis and the differential effects of somatic mutations. *Journal of Molecular Biology* 426: 347–361.
- Boone, C.D., Gill, S., Tu, C., Silverman, D.N., and McKenna, R. (2013). Structural, catalytic and stabilizing consequences of aromatic cluster variants in human carbonic anhydrase II. *Archives of Biochemistry and Biophysics* 539: 31–37.

- Bork, P., Holm, L., and Sander, C. (1994). The immunoglobulin fold. Structural classification, sequence patterns and common core. *Journal of Molecular Biology* 242: 309–320.
- Braakman, I. and Bulleid, N.J. (2011). Protein folding and modification in the mammalian endoplasmic reticulum. *Annual Review of Biochemistry* 80: 71–99.
- Bracher, A. and Verghese, J. (2015). GrpE, Hsp110/Grp170, HspBP1/Sil1 and BAG Domain Proteins: Nucleotide Exchange Factors for Hsp70 Molecular Chaperones. In: Blatch, G.L., Edkins, A.L. (Eds.). *The Networking of Chaperones by Co-chaperones: Control of Cellular Protein Homeostasis*. Springer International Publishing, Cham, pp. 1–33.
- Brenner, D.A., Jain, M., Pimentel, D.R., Wang, B., Connors, L.H., Skinner, M., Apstein, C.S., and Liao, R. (2004). Human amyloidogenic light chains directly impair cardiomyocyte function through an increase in cellular oxidant stress. *Circulation Research* 94: 1008–1010.
- Brochet, X., Lefranc, M.-P., and Giudicelli, V. (2008). IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res* 36: W503-W508.
- Brown, P.H. and Schuck, P. (2006). Macromolecular Size-and-Shape Distributions by Sedimentation Velocity Analytical Ultracentrifugation. *Biophysical Journal* 90: 4651–4661.
- Brumshtein, B., Esswein, S.R., Landau, M., Ryan, C.M., Whitelegge, J.P., Phillips, M.L., Cascio, D., Sawaya, M.R., and Eisenberg, D.S. (2014). Formation of amyloid fibers by monomeric light chain variable domains. *Journal of Biological Chemistry* 289: 27513–27525.
- Brumshtein, B., Esswein, S.R., Sawaya, M.R., Rosenberg, G., Ly, A.T., Landau, M., and Eisenberg, D.S. (2018). Identification of two principal amyloid-driving segments in variable domains of Ig light chains in systemic light-chain amyloidosis. *Journal of Biological Chemistry* 293: 19659–19671.
- Bruylants, G., Wouters, J., and Michaux, C. (2005). Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design. *Current Medicinal Chemistry* 12: 2011–2020.
- Buchner, J., Renner, M., Lilie, H., Hinz, H.J., Jaenicke, R., Kiefhaber, T., and Rudolph, R. (1991). Alternatively folded states of an immunoglobulin. *Biochemistry* 30: 6922–6929.
- Bulaj, G. (2005). Formation of disulfide bonds in proteins and peptides. *Biotechnology Advances* 23: 87–92.
- Bulawa, C.E., Connelly, S., DeVit, M., Wang, L., Weigel, C., Fleming, J.A., Packman, J., Powers, E.T., Wiseman, R.L., and Foss, T.R. (2012). Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *Proceedings of the National Academy of Sciences* 109: 9629–9634.

- Cantor, C.R. and Schimmel, P.R. (1980). *Biophysical chemistry: Part II: Techniques for the study of biological structure and function*. Macmillan.
- Case, D.A., Ben-Shalom, I.Y., Brozell, S.R., Cerutti, D.S., Cheatham, T.E., III, Cruzeiro, V.W.D., Darden, T.A., Duke, R.E., Ghoreishi, D., and Gilson, M.K., et al. (2018). *Amber 18*. University of California, San Francisco.
- Catici, D.A.M., Amos, H.E., Yang, Y., van den Elsen, J.M.H., and Pudney, C.R. (2016). The red edge excitation shift phenomenon can be used to unmask protein structural ensembles: implications for NEMO-ubiquitin interactions. *The FEBS Journal* 283: 2272–2284.
- Chang, H.-J., Jian, J.-W., Hsu, H.-J., Lee, Y.-C., Chen, H.-S., You, J.-J., Hou, S.-C., Shao, C.-Y., Chen, Y.-J., and Chiu, K.-P., et al. (2014). Loop-Sequence Features and Stability Determinants in Antibody Variable Domains by High-Throughput Experiments. *Structure* 22: 9–21.
- Chatterjee, S., Basak, A.J., Nair, A.V., Duraivelan, K., and Samanta, D. (2021). Immunoglobulin-fold containing bacterial adhesins: molecular and structural perspectives in host tissue colonization and infection. *FEMS Microbiol Lett* 368: fnaa220.
- Chen, B., Zhong, D., and Monteiro, A. (2006). Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. *BMC Genomics* 7: 1–19.
- Chen, J. and Stites, W.E. (2001). Packing Is a Key Selection Factor in the Evolution of Protein Hydrophobic Cores. *Biochemistry* 40: 15280–15289.
- Cherepanova, N., Shrimal, S., and Gilmore, R. (2016). N-linked glycosylation and homeostasis of the endoplasmic reticulum. *Current Opinion in Cell Biology* 41: 57–65.
- Chiti, F. and Dobson, C.M. (2017). Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annual Review of Biochemistry* 86: 27–68.
- Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C.M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 424: 805–808.
- Chiti, F., Taddei, N., Baroni, F., Capanni, C., Stefani, M., Ramponi, G., and Dobson, C.M. (2002). Kinetic partitioning of protein folding and aggregation. *Nature Structural Biology* 9: 137–143.
- Chiu, M.L., Goulet, D.R., Teplyakov, A., and Gilliland, G.L. (2019). Antibody structure and function: the basis for engineering therapeutics. *Antibodies* 8: 55.
- Chothia, C. (1984). Principles That Determine The Structure Of Proteins. *Annual Review of Biochemistry* 53: 537–572.
- Chothia, C. and Lesk, A.M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *Journal of Molecular Biology* 196: 901–917.

- Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, D., and Tulip, W.R., et al. (1989). Conformations of immunoglobulin hypervariable regions. *Nature* 342: 877–883.
- Cohen, A.D., Landau, H., Scott, E.C., Liedtke, M., Kaufman, J.L., Rosenzweig, M., Gasparetto, C., Vesole, D.H., Sanchorawala, V., and Lentzsch, S. (2016). Safety and efficacy of carfilzomib (CFZ) in previously-treated systemic light-chain (AL) amyloidosis. *Blood* 128: 645.
- Cohen, S.I.A., Linse, S., Luheshi, L.M., Hellstrand, E., White, D.A., Rajah, L., Otzen, D.E., Vendruscolo, M., Dobson, C.M., and Knowles, T.P.J. (2013). Proliferation of amyloid- β 42 aggregates occurs through a secondary nucleation mechanism. *Proceedings of the National Academy of Sciences* 110: 9758–9763.
- Cohen, S.I.A., Vendruscolo, M., Dobson, C.M., and Knowles, T.P.J. (2012). From macroscopic measurements to microscopic mechanisms of protein aggregation. *Journal of Molecular Biology* 421: 160–171.
- Collins, A.M. and Watson, C.T. (2018). Immunoglobulin light chain gene rearrangements, receptor editing and the development of a self-tolerant antibody repertoire. *Frontiers in Immunology* 9: 2249.
- Creighton, T.E. (1993). *Proteins: structures and molecular properties*. Macmillan.
- del Pozo Yauner, L., Ortiz, E., and Becerril, B. (2006). The CDR1 of the human lambdaVI light chains adopts a new canonical structure. *Proteins* 62: 122–129.
- Del Pozo-Yauner, L., Wall, J.S., González Andrade, M., Sánchez-López, R., Rodríguez-Ambríz, S.L., Pérez Carreón, J.I., Ochoa-Leyva, A., and Fernández-Velasco, D.A. (2014). The N-terminal strand modulates immunoglobulin light chain fibrillogenesis. *Biochemical and Biophysical Research Communications* 443: 495–499.
- Desport, E., Bridoux, F., Sirac, C., Delbes, S., Bender, S., Fernandez, B., Quellard, N., Lacombe, C., Goujon, J.-M., and Lavergne, D. (2012). AL amyloidosis. *Orphanet Journal of Rare Diseases* 7: 1–13.
- DiCostanzo, A.C., Thompson, J.R., Peterson, F.C., Volkman, B.F., and Ramirez-Alvarado, M. (2012). Tyrosine residues mediate fibril formation in a dynamic light chain dimer interface. *Journal of Biological Chemistry* 287: 27997–28006.
- Dill, K.A. (1985). Theory for the folding and stability of globular proteins. *Biochemistry* 24: 1501–1509.
- Dill, K.A. and Chan, H.S. (1997). From Levinthal to pathways to funnels. *Nature Structural Biology* 4: 10–19.
- Dill, K.A., Ozkan, S.B., Shell, M.S., and Weikl, T.R. (2008). The protein folding problem. *Annual Review of Biophysics* 37: 289–316.

- Dobson, C.M. (2001). The structural basis of protein folding and its links with human disease. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356: 133–145.
- Dobson, C.M. (2003). Protein folding and misfolding. *Nature* 426: 884–890.
- Dobson, C.M., Šali, A., and Karplus, M. (1998). Protein Folding: A Perspective from Theory and Experiment. *Angewandte Chemie International Edition* 37: 868–893.
- Dondelinger, M., Filée, P., Sauvage, E., Quinting, B., Muyldermans, S., Galleni, M., and Vandevenne, M.S. (2018). Understanding the Significance and Implications of Antibody Numbering and Antigen-Binding Surface/Residue Definition. *Frontiers in Immunology* 9: 2278.
- Dorshkind, K. and Rawlings, D.J. (2018). B-cell development. In: *Hematology*. Elsevier, pp. 210–220.
- D'Souza, A., Dispenzieri, A., Wirk, B., Zhang, M.-J., Huang, J., Gertz, M.A., Kyle, R.A., Kumar, S., Comenzo, R.L., and Gale, R.P. (2015). Improved outcomes after autologous hematopoietic cell transplantation for light chain amyloidosis: a Center for International Blood and Marrow Transplant Research Study. *Journal of Clinical Oncology* 33: 3741.
- Duhamel, S., Mohty, D., Magne, J., Lavergne, D., Bordessoule, D., Aboyans, V., and Jaccard, A. (2017). Incidence and prevalence of light chain amyloidosis: a population-based study. *Blood* 130: 5577.
- Edwards, C.V., Gould, J., Langer, A.L., Mapara, M.Y., Radhakrishnan, J., Maurer, M.S., Raza, S., Mears, J.G., Leng, S., and Wall, J.S. (2017). Final analysis of the phase 1a/b study of chimeric fibril-reactive monoclonal antibody 11-1F4 in patients with relapsed or refractory AL amyloidosis. *Blood* 130: 509.
- Eftink, M.R. and Ghiron, C.A. (1981). Fluorescence quenching studies with proteins. *Analytical Biochemistry* 114: 199–227.
- Ellis, R.J. and Minton, A.P. (2006). Protein aggregation in crowded environments. *Biological Chemistry* 387: 485–497.
- Emily, M., Talvas, A., and Delamarche, C. (2013). MetAmyl: a METa-predictor for AMYLoid proteins. *PloS one* 8: e79722.
- Englander, S.W. (2006). Hydrogen exchange and mass spectrometry: A historical perspective. *Journal of the American Society for Mass Spectrometry* 17: 1481–1489.
- Enqvist, S., Sletten, K., and Westermark, P. (2009). Fibril protein fragmentation pattern in systemic AL-amyloidosis. *The Journal of Pathology* 219: 473–480.
- Fändrich, M. and Dobson, C.M. (2002). The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *The EMBO Journal* 21: 5682–5690.
- Feige, M.J., Gräwert, M.A., Marcinowski, M., Hennig, J., Behnke, J., Ausländer, D., Herold, E.M., Peschek, J., Castro, C.D., and Flajnik, M. (2014). The structural analysis of shark

- IgNAR antibodies reveals evolutionary principles of immunoglobulins. *Proceedings of the National Academy of Sciences* 111: 8155–8160.
- Feige, M.J., Groscurth, S., Marcinowski, M., Yew, Z.T., Truffault, V., Paci, E., Kessler, H., and Buchner, J. (2008). The structure of a folding intermediate provides insight into differences in immunoglobulin amyloidogenicity. *Proceedings of the National Academy of Sciences* 105: 13373–13378.
- Feige, M.J., Hendershot, L.M., and Buchner, J. (2010). How antibodies fold. *Trends in Biochemical Sciences* 35: 189–198.
- Fenech, E.J., Ben-Dor, S., and Schuldiner, M. (2020). Double the Fun, Double the Trouble: Paralogs and Homologs Functioning in the Endoplasmic Reticulum. *Annual Review of Biochemistry* 89: 637–666.
- Fersht, A.R. (1997). Nucleation mechanisms in protein folding. *Current Opinion in Structural Biology* 7: 3–9.
- Fitzpatrick, A.W.P., Debelouchina, G.T., Bayro, M.J., Clare, D.K., Caporini, M.A., Bajaj, V.S., Jaroniec, C.P., Wang, L., Ladizhansky, V., and Müller, S.A. (2013). Atomic structure and hierarchical assembly of a cross- β amyloid fibril. *Proceedings of the National Academy of Sciences* 110: 5468–5473.
- Fleckenstein, T., Kastenmüller, A., Stein, M.L., Peters, C., Daake, M., Krause, M., Weinfurter, D., Haslbeck, M., Weinkauf, S., and Groll, M., et al. (2015). The Chaperone Activity of the Developmental Small Heat Shock Protein Sip1 Is Regulated by pH-Dependent Conformational Changes. *Molecular Cell* 58: 1067–1078.
- Fontana, A., Laureto, P.P. de, Filippis, V. de, Scaramella, E., and Zambonin, M. (1999). Limited Proteolysis in the Study of Protein Conformation. In: Sterchi, E.E., Stöcker, W. (Eds.). *Proteolytic Enzymes: Tools and Targets*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 253–280.
- Fowler, D.M., Koulov, A.V., Balch, W.E., and Kelly, J.W. (2007). Functional amyloid—from bacteria to humans. *Trends in Biochemical Sciences* 32: 217–224.
- Gąciarz, A. and Ruddock, L.W. (2017). Complementarity determining regions and frameworks contribute to the disulfide bond independent folding of intrinsically stable scFv. *PLoS one* 12: e0189964.
- Gade Malmos, K., Blancas-Mejia, L.M., Weber, B., Buchner, J., Ramirez-Alvarado, M., Naiki, H., and Otzen, D. (2017). ThT 101: a primer on the use of thioflavin T to investigate amyloid formation. *Amyloid* 24: 1–16.
- Gao, X., Carroni, M., Nussbaum-Krammer, C., Mogk, A., Nillegoda, N.B., Szlachcic, A., Guilbride, D.L., Saibil, H.R., Mayer, M.P., and Bukau, B. (2015). Human Hsp70 Disaggregase Reverses Parkinson's-Linked α -Synuclein Amyloid Fibrils. *Molecular Cell* 59: 781–793.

- Genest, O., Wickner, S., and Doyle, S.M. (2019). Hsp90 and Hsp70 chaperones: Collaborators in protein remodeling. *Journal of Biological Chemistry* 294: 2109–2120.
- Gertz, M.A. (2018). Immunoglobulin light chain amyloidosis: 2018 Update on diagnosis, prognosis, and treatment. *American Journal of Hematology* 93: 1169–1180.
- Gertz, M.A., Landau, H., Comenzo, R.L., Seldin, D., Weiss, B., Zonder, J., Merlini, G., Schönland, S., Walling, J., and Kinney, G.G. (2016). First-in-human phase I/II study of NEOD001 in patients with light chain amyloidosis and persistent organ dysfunction. *Journal of Clinical Oncology* 34: 1097.
- Glennner, G.G., Harbaugh, J., Ohms, J.I., Harada, M., and Cuatrecasas, P. (1970). An amyloid protein: the amino-terminal variable fragment of an immunoglobulin light chain. *Biochemical and Biophysical Research Communications* 41: 1287–1289.
- Go, Y.-M. and Jones, D.P. (2008). Redox compartmentalization in eukaryotic cells. *Biochimica et Biophysica Acta* 1780: 1273–1290.
- Goldschmidt, L., Teng, P.K., Riek, R., and Eisenberg, D. (2010). Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proceedings of the National Academy of Sciences* 107: 3487.
- González-Andrade, M., Becerril-Luján, B., Sánchez-López, R., Ceceña-Álvarez, H., Pérez-Carreón, J.I., Ortiz, E., Fernández-Velasco, D.A., and Del Pozo-Yauner, L. (2013). Mutational and genetic determinants of λ 6 light chain amyloidogenesis. *The FEBS Journal* 280: 6173–6183.
- Goto, Y., Azuma, T., and Hamaguchi, K. (1979). Refolding of the immunoglobulin light chain. *J Biochem* 85: 1427–1438.
- Goto, Y. and Hamaguchi, K. (1982). Unfolding and refolding of the constant fragment of the immunoglobulin light chain. *Journal of Molecular Biology* 156: 891–910.
- Gottwald, J. and Röcken, C. (2021). The amyloid proteome: A systematic review and proposal of a protein classification system. *Critical Reviews in Biochemistry and Molecular Biology* 56: 526–542.
- Greene, M.J., Sam, F., Hoo, P.T.S., Patel, R.S., Seldin, D.C., and Connors, L.H. (2011). Evidence for a functional role of the molecular chaperone clusterin in amyloidotic cardiomyopathy. *The American Journal of Pathology* 178: 61–68.
- Gutin, A.M., Abkevich, V.I., and Shakhnovich, E.I. (1995). Is Burst Hydrophobic Collapse Necessary for Protein Folding? *Biochemistry* 34: 3066–3076.
- Haass, C. and Selkoe, D.J. (2007). Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β -peptide. *Nature Reviews Molecular Cell Biology* 8: 101–112.
- Harrison, S.C. and Aggarwal, A.K. (1990). DNA recognition by proteins with the helix-turn-helix motif. *Annual Review of Biochemistry* 59: 933–969.

- Hartl, F.U. and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. *Nature Structural Biology* 16: 574–581.
- Haslbeck, M. and Vierling, E. (2015). A first line of stress defense: small heat shock proteins and their function in protein homeostasis. *Journal of Molecular Biology* 427: 1537–1548.
- Haslbeck, M., Weinkauf, S., and Buchner, J. (2019). Small heat shock proteins: Simplicity meets complexity. *Journal of Biological Chemistry* 294: 2121–2132.
- Hassan, H., Anwer, F., Javaid, A., and Hashmi, H. (2021). Progress in research: Daratumumab improves treatment outcomes of patients with AL amyloidosis. *Critical Reviews in Oncology/Hematology* 165: 103435.
- Hawkins, P.N. (1988). Amyloidosis. *Blood Reviews* 2: 270–280.
- Helms, L.R. and Wetzel, R. (1995). Destabilizing loop swaps in the CDRs of an immunoglobulin VL domain. *Protein Science* 4: 2073–2081.
- Hernández-Santoyo, A., del Pozo Yauner, L., Fuentes-Silva, D., Ortiz, E., Rudiño-Piñera, E., Sánchez-López, R., Horjales, E., Becerril, B., and Rodríguez-Romero, A. (2010). A single mutation at the sheet switch region results in conformational changes favoring lambda6 light-chain fibrillogenesis. *Journal of Molecular Biology* 396: 280–292.
- Herold, E.M., John, C., Weber, B., Kremser, S., Eras, J., Berner, C., Deubler, S., Zacharias, M., and Buchner, J. (2017). Determinants of the assembly and function of antibody variable domains. *Scientific Reports* 7: 12276.
- Hessling, M., Richter, K., and Buchner, J. (2009). Dissection of the ATP-induced conformational cycle of the molecular chaperone Hsp90. *Nature Structural Biology* 16: 287–293.
- Hora, M., Carballo-Pacheco, M., Weber, B., Morris, V.K., Wittkopf, A., Buchner, J., Strodel, B., and Reif, B. (2017). Epigallocatechin-3-gallate preferentially induces aggregation of amyloidogenic immunoglobulin light chains. *Scientific Reports* 7: 1–12.
- Hurle, M.R., Helms, L.R., Li, L., Chan, W., and Wetzel, R. (1994). A role for destabilizing amino acid replacements in light-chain amyloidosis. *Proceedings of the National Academy of Sciences* 91: 5446–5450.
- Hussain, A.F., Grimm, A., Sheng, W., Zhang, C., Al-Rawe, M., Bräutigam, K., Abu Mraheil, M., Zeppernick, F., and Meinhold-Heerlein, I. (2021). Toward homogenous antibody drug conjugates using enzyme-based conjugation approaches. *Pharmaceuticals* 14: 343.
- Iadanza, M.G., Jackson, M.P., Hewitt, E.W., Ranson, N.A., and Radford, S.E. (2018). A new era for understanding amyloid structures and disease. *Nature Reviews Molecular Cell Biology* 19: 755–773.
- Iqbal, S.M., Stecklein, K., Sarow, J., Krabak, M., Hillengass, J., and McCarthy, P. (2019). Elotuzumab in combination with lenalidomide and dexamethasone for treatment-resistant

- immunoglobulin light chain amyloidosis with multiple myeloma. *Clinical Lymphoma Myeloma and Leukemia* 19: e33-e36.
- Jablonski, A. (1933). Efficiency of anti-Stokes fluorescence in dyes. *Nature* 131: 839–840.
- Jimenez, J.L., Nettleton, E.J., Bouchard, M., Robinson, C.V., Dobson, C.M., and Saibil, H.R. (2002). The protofilament structure of insulin amyloid fibrils. *Proceedings of the National Academy of Sciences* 99: 9196–9201.
- Johnson, B.D., Schumacher, R.J., Ross, E.D., and Toft, D.O. (1998). Hop Modulates hsp70/hsp90 Interactions in Protein Folding *. *Journal of Biological Chemistry* 273: 3679–3686.
- Johnson, C.M. (2013). Differential scanning calorimetry as a tool for protein folding and stability. *Archives of Biochemistry and Biophysics* 531: 100–109.
- Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* 79: 926–935.
- Kampinga, H.H., Andreasson, C., Barducci, A., Cheetham, M.E., Cyr, D., Emanuelsson, C., Genevaux, P., Gestwicki, J.E., Goloubinoff, P., and Huerta-Cepas, J., et al. (2019). Function, evolution, and structure of J-domain proteins. *Cell Stress and Chaperones* 24: 7–15.
- Kampinga, H.H. and Craig, E.A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology* 11: 579–592.
- Kaplan, B., Livneh, A., and Sela, B.-A. (2011). Immunoglobulin free light chain dimers in human diseases. *The Scientific World Journal* 11: 726–735.
- Kaplan, B., Ramirez-Alvarado, M., Sikkink, L., Golderman, S., Dispenzieri, A., Livneh, A., and Gallo, G. (2009). Free light chains in plasma of patients with light chain amyloidosis and non-amyloid light chain deposition disease. High proportion and heterogeneity of disulfide-linked monoclonal free light chains as pathogenic features of amyloid disease. *British Journal of Haematology* 144: 705–715.
- Karplus, M. and Kuriyan, J. (2005). Molecular dynamics and protein function. *Proceedings of the National Academy of Sciences* 102: 6679–6685.
- Karplus, M. and Weaver, D.L. (1979). Diffusion–collision model for protein folding. *Biopolymers* 18: 1421–1437.
- Karplus, M. and Weaver, D.L. (1994). Protein folding dynamics: The diffusion-collision model and experimental data. *Protein Science* 3: 650–668.
- Kauzmann, W. (1959). Some Factors in the Interpretation of Protein Denaturation. In: Anfinsen, C.B., Anson, M.L., Bailey, K., Edsall, J.T. (Eds.). *Advances in Protein Chemistry*. Academic Press, pp. 1–63.

- Kazman, P., Absmeier, R.M., Engelhardt, H., and Buchner, J. (2021). Dissection of the amyloid formation pathway in AL amyloidosis. *Nature Communications* 12: 1–10.
- Kazman, P., Vielberg, M.-T., Pulido Cendales, M.D., Hunziger, L., Weber, B., Hegenbart, U., Zacharias, M., Köhler, R., Schönland, S., and Groll, M., et al. (2020). Fatal amyloid formation in a patient's antibody light chain is caused by a single point mutation. *eLife* 9.
- Kelly, S.M., Jess, T.J., and Price, N.C. (2005). How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1751: 119–139.
- Kelly, S.M. and Price, N.C. (2000). The use of circular dichroism in the investigation of protein structure and function. *Current Protein & Peptide Science* 1: 349–384.
- Khurana, R., Gillespie, J.R., Talapatra, A., Minert, L.J., Ionescu-Zanetti, C., Millett, I., and Fink, A.L. (2001). Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates. *Biochemistry* 40: 3525–3535.
- Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* 37: D387-92.
- Kikuchi, H., Goto, Y., and Hamaguchi, K. (1986). Reduction of the buried intrachain disulfide bond of the constant fragment of the immunoglobulin light chain: global unfolding under physiological conditions. *Biochemistry* 25: 2009–2013.
- Kim, Y.-s., Wall, J.S., Meyer, J., Murphy, C., Randolph, T.W., Manning, M.C., Solomon, A., and Carpenter, J.F. (2000). Thermodynamic modulation of light chain amyloid fibril formation. *Journal of Biological Chemistry* 275: 1570–1574.
- Klafki, H.W., Kratzin, H.D., Pick, A.I., Eckart, K., Karas, M., and Hilschmann, N. (1992). Complete amino acid sequence determinations demonstrate identity of the urinary Bence Jones protein (BJP-DIA) and the amyloid fibril protein (AL-DIA) in a case of AL-amyloidosis. *Biochemistry* 31: 3265–3272.
- Klimtchuk, E.S., Gursky, O., Patel, R.S., Laporte, K.L., Connors, L.H., Skinner, M., and Seldin, D.C. (2010). The critical role of the constant region in thermal stability and aggregation of amyloidogenic immunoglobulin light chain. *Biochemistry* 49: 9848–9857.
- Klotz, I.M., Langebman, N.R., and Dahnall, D.W. (1970). Quaternary Structure of Proteins. *Annual Review of Biochemistry* 39: 25–62.
- Knowles, T.P.J. and Mezzenga, R. (2016). Amyloid Fibrils as Building Blocks for Natural and Artificial Functional Materials. *Adv. Mater.* 28: 6546–6561.
- Knowles, T.P.J., Vendruscolo, M., and Dobson, C.M. (2014). The amyloid state and its association with protein misfolding diseases. *Nature Reviews Molecular Cell Biology* 15: 384–396.
- Kozlov, G., Määttänen, P., Thomas, D.Y., and Gehring, K. (2010). A structural overview of the PDI family of proteins. *The FEBS Journal* 277: 3924–3936.

- Kuroda, D., Shirai, H., Kobori, M., and Nakamura, H. (2009). Systematic classification of CDR-L3 in antibodies: implications of the light chain subtypes and the VL–VH interface. *Proteins: Structure, Function, and Bioinformatics* 75: 139–146.
- Lakowicz, J.R. (2006). *Principles of Fluorescence Spectroscopy*. Springer US, Boston, MA.
- Laue, T. (2001). Biophysical studies by ultracentrifugation. *Current Opinion in Structural Biology* 11: 579–583.
- Lavatelli, F., Perlman, D.H., Spencer, B., Prokaeva, T., McComb, M.E., Théberge, R., Connors, L.H., Bellotti, V., Seldin, D.C., and Merlini, G., et al. (2008). Amyloidogenic and associated proteins in systemic amyloidosis proteome of adipose tissue. *Molecular & Cellular Proteomics* 7: 1570–1583.
- Lazaridis, T., Archontis, G., and Karplus, M. (1995). Enthalpic contribution to protein stability: insights from atom-based calculations and statistical mechanics. *Advances in Protein Chemistry* 47: 231–306.
- Le Marchand, T., Rosa, M. de, Salvi, N., Sala, B.M., Andreas, L.B., Barbet-Massin, E., Sormanni, P., Barbiroli, A., Porcari, R., and Sousa Mota, C., et al. (2018). Conformational dynamics in crystals reveal the molecular bases for D76N beta-2 microglobulin aggregation propensity. *Nature Communications* 9: 1658.
- Lecoq, L., Wiegand, T., Rodriguez-Alvarez, F.J., Cadalbert, R., Herrera, G.A., Del Pozo-Yauner, L., Meier, B.H., and Böckmann, A. (2019). A Substantial Structural Conversion of the Native Monomer Leads to in-Register Parallel Amyloid Fibril Formation in Light-Chain Amyloidosis. *Chembiochem* 20: 1027–1031.
- Lefranc, M.-P., Giudicelli, V., Ginestoux, C., Jabado-Michaloud, J., Folch, G., Bellahcene, F., Wu, Y., Gemrot, E., Brochet, X., and Lane, J., et al. (2009). IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res* 37: D1006-12.
- Levinthal, C. (1968). Are there pathways for protein folding? *Journal de Chimie Physique* 65: 44–45.
- Lewkowicz, E. and Gursky, O. (2022). Dynamic protein structures in normal function and pathologic misfolding in systemic amyloidosis. *Biophysical Chemistry* 280: 106699.
- Li, Y., Wang, G., Li, N., Wang, Y., Zhu, Q., Chu, H., Wu, W., Tan, Y., Yu, F., and Su, X.-D., et al. (2020). Structural insights into immunoglobulin M. *Science* 367: 1014–1017.
- Los Rios, M. de, Criscitiello, M.F., and Smider, V.V. (2015). Structural and genetic diversity in antibody repertoires from diverse species. *Current Opinion in Structural Biology* 33: 27–41.
- Luna-Martínez, O.D., Hernández-Santoyo, A., Villalba-Velázquez, M.I., Sánchez-Alcalá, R., Fernández-Velasco, D.A., and Becerril, B. (2017). Stabilizing an amyloidogenic λ 6 light chain variable domain. *The FEBS Journal* 284: 3702–3717.

- Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* 11: 3696–3713.
- Makhatadze, G.I. and Privalov, P.L. (1996). On the entropy of protein folding. *Protein Science* 5: 507–510.
- Malkov, S.N., Živković, M.V., Beljanski, M.V., Hall, M.B., and Zarić, S.D. (2008). A reexamination of the propensities of amino acids towards a particular secondary structure: classification of amino acids based on their chemical structure. *Journal of Molecular Modeling* 14: 769–775.
- Marin-Argany, M., Güell-Bosch, J., Blancas-Mejía, L.M., Villegas, S., and Ramirez-Alvarado, M. (2015). Mutations can cause light chains to be too stable or too unstable to form amyloid fibrils. *Protein Science* 24: 1829–1840.
- Marin-Argany, M., Lin, Y., Misra, P., Williams, A., Wall, J.S., Howell, K.G., Elsbernd, L.R., McClure, M., and Ramirez-Alvarado, M. (2016). Cell damage in light chain amyloidosis: fibril internalization, toxicity and cell-mediated seeding. *Journal of Biological Chemistry* 291: 19813–19825.
- Maritan, M., Romeo, M., Oberti, L., Sormanni, P., Tasaki, M., Russo, R., Ambrosetti, A., Motta, P., Rognoni, P., and Mazzini, G., et al. (2020). Inherent Biophysical Properties Modulate the Toxicity of Soluble Amyloidogenic Light Chains. *Journal of Molecular Biology* 432: 845–860.
- Marsh, J.A. and Teichmann, S.A. (2015). Structure, Dynamics, Assembly, and Evolution of Protein Complexes. *Annual Review of Biochemistry* 84: 551–575.
- Masson, G.R., Jenkins, M.L., and Burke, J.E. (2017). An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opinion on Drug Discovery* 12: 981–994.
- Maya-Martinez, R., French-Pacheco, L., Valdés-García, G., Pastor, N., and Amero, C. (2019). Different dynamics in 6aJL2 proteins associated with AL amyloidosis, a conformational disease. *International Journal of Molecular Sciences* 20: 4078.
- Mazur, S.J., Gallagher, E.S., Debnath, S., Durell, S.R., Anderson, K.W., Miller Jenkins, L.M., Appella, E., and Hudgens, J.W. (2017). Conformational Changes in Active and Inactive States of Human PP2C α Characterized by Hydrogen/Deuterium Exchange-Mass Spectrometry. *Biochemistry* 56: 2676–2689.
- Merlini, G. (2017). AL amyloidosis: from molecular mechanisms to targeted therapies. *Hematology - American Society of Hematology - Education Program 2017*: 1–12.
- Merlini, G. and Bellotti, V. (2003). Molecular mechanisms of amyloidosis. *New England Journal of Medicine* 349: 583–596.

- Merlini, G., Dispenzieri, A., Santhorawala, V., Schönland, S.O., Palladini, G., Hawkins, P.N., and Gertz, M.A. (2018). Systemic immunoglobulin light chain amyloidosis. *Nature Reviews Disease Primers* 4: 38.
- Micsonai, A., Moussong, É., Wien, F., Boros, E., Vadász, H., Murvai, N., Lee, Y.-H., Molnár, T., Réfrégiers, M., and Goto, Y., et al. (2022). BeStSel: webserver for secondary structure and fold prediction for protein CD spectroscopy. *Nucleic Acids Res* 50: W90-W98.
- Miles, A.J., Ramalli, S.G., and Wallace, B.A. (2022). DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. *Protein Science* 31: 37–46.
- Mogk, A., Bukau, B., and Kampinga, H.H. (2018). Cellular Handling of Protein Aggregates by Disaggregation Machines. *Molecular Cell* 69: 214–226.
- Morgan, G.J. and Kelly, J.W. (2016). The Kinetic Stability of a Full-Length Antibody Light Chain Dimer Determines whether Endoproteolysis Can Release Amyloidogenic Variable Domains. *Journal of Molecular Biology* 428: 4280–4297.
- Morgan, G.J., Usher, G.A., and Kelly, J.W. (2017). Incomplete Refolding of Antibody Light Chains to Non-Native, Protease-Sensitive Conformations Leads to Aggregation: A Mechanism of Amyloidogenesis in Patients? *Biochemistry* 56: 6597–6614.
- Morgan, G.J. and Wall, J.S. (2020). The Process of Amyloid Formation due to Monoclonal Immunoglobulins. *Hematology/Oncology Clinics of North America* 34: 1041–1054.
- Morgan, G.J., Yan, N.L., Mortenson, D.E., Rennella, E., Blundon, J.M., Gwin, R.M., Lin, C.-Y., Stanfield, R.L., Brown, S.J., and Rosen, H., et al. (2019). Stabilization of amyloidogenic immunoglobulin light chains by small molecules. *Proceedings of the National Academy of Sciences* 116: 8360–8369.
- Motea, E.A. and Berdis, A.J. (2010). Terminal deoxynucleotidyl transferase: The story of a misguided DNA polymerase. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1804: 1151–1166.
- Mühlhofer, M., Peters, C., Kriehuber, T., Kreuzeder, M., Kazman, P., Rodina, N., Reif, B., Haslbeck, M., Weinkauff, S., and Buchner, J. (2021). Phosphorylation activates the yeast small heat shock protein Hsp26 by weakening domain contacts in the oligomer ensemble. *Nature Communications* 12: 6697.
- Murphy, K. and Weaver, C. (2016). *Janeway's immunobiology*. Garland science.
- Nawata, M., Tsutsumi, H., Kobayashi, Y., Unzai, S., Mine, S., Nakamura, T., Uegaki, K., Kamikubo, H., Kataoka, M., and Hamada, D. (2017). Heat-induced native dimerization prevents amyloid formation by variable domain from immunoglobulin light-chain REI. *The FEBS Journal* 284: 3114–3127.

- Nokwe, C.N., Hora, M., Zacharias, M., Yagi, H., Peschek, J., Reif, B., Goto, Y., and Buchner, J. (2016). A Stable Mutant Predisposes Antibody Domains to Amyloid Formation through Specific Non-Native Interactions. *Journal of Molecular Biology* 428: 1315–1332.
- Oberti, L., Rognoni, P., Barbiroli, A., Lavatelli, F., Russo, R., Maritan, M., Palladini, G., Bolognesi, M., Merlini, G., and Ricagno, S. (2017). Concurrent structural and biophysical traits link with immunoglobulin light chains amyloid propensity. *Scientific Reports* 7: 16809.
- Olshansky, S.J., Passaro, D.J., Hershov, R.C., Layden, J., Carnes, B.A., Brody, J., Hayflick, L., Butler, R.N., Allison, D.B., and Ludwig, D.S. (2005). A Potential Decline in Life Expectancy in the United States in the 21st Century. *New England Journal of Medicine* 352: 1138–1145.
- Onuchic, J.N., Luthey-Schulten, Z., and Wolynes, P.G. (1997). Theory of protein folding: the energy landscape perspective. *Annual Review of Physical Chemistry* 48: 545–600.
- Pace, C.N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* 131: 266–280.
- Pace, C.N. (1995). Evaluating contribution of hydrogen bonding and hydrophobic bonding to protein folding. *Methods in Enzymology* 259: 538–554.
- Palladini, G., Campana, C., Klersy, C., Balduini, A., Vadacca, G., Perfetti, V., Perlini, S., Obici, L., Ascari, E., and d'Eril, G.M., et al. (2003). Serum N-Terminal Pro-Brain Natriuretic Peptide Is a Sensitive Marker of Myocardial Dysfunction in AL Amyloidosis. *Circulation* 107: 2440–2445.
- Palladini, G., Perfetti, V., Obici, L., Caccialanza, R., Semino, A., Adami, F., Cavallero, G., Rustichelli, R., Virga, G., and Merlini, G. (2004). Association of melphalan and high-dose dexamethasone is effective and well tolerated in patients with AL (primary) amyloidosis who are ineligible for stem cell transplantation. *Blood* 103: 2936–2938.
- Palladini, G., Sachchithanantham, S., Milani, P., Gillmore, J., Foli, A., Lachmann, H., Basset, M., Hawkins, P., Merlini, G., and Wechalekar, A.D. (2015). A European collaborative study of cyclophosphamide, bortezomib, and dexamethasone in upfront treatment of systemic AL amyloidosis. *Blood* 126: 612–615.
- Pasalic, D., Weber, B., Giannone, C., Anelli, T., Müller, R., Fagioli, C., Felkl, M., John, C., Mossuto, M.F., and Becker, C.F.W., et al. (2017). A peptide extension dictates IgM assembly. *Proceedings of the National Academy of Sciences* 114: E8575-E8584.
- Pavri, R. and Nussenzweig, M.C. (2011). Chapter 1 - AID Targeting in Antibody Diversity. In: Alt, F.W., Austen, K.F., Honj, T., Melchers, F., Uhr, J.W., Unanue, E.R. (Eds.). *Advances in Immunology*. Academic Press, pp. 1–26.
- Perfetti, V., Casarini, S., Palladini, G., Vignarelli, M.C., Klersy, C., Diegoli, M., Ascari, E., and Merlini, G. (2002). Analysis of V(λ)-J(λ) expression in plasma cells from

- primary (AL) amyloidosis and normal bone marrow identifies 3r (lambdall) as a new amyloid-associated germline gene segment. *Blood* 100: 948–953.
- Perfetti, V., Ubbiali, P., Vignarelli, M.C., Diegoli, M., Fasani, R., Stoppini, M., Lisa, A., Mangione, P., Obici, L., and Arbustini, E., et al. (1998). Evidence That Amyloidogenic Light Chains Undergo Antigen-Driven Selection. *Blood* 91: 2948–2954.
- Pertinhez, T.A., Bouchard, M., Smith, R.A.G., Dobson, C.M., and Smith, L.J. (2002). Stimulation and inhibition of fibril formation by a peptide in the presence of different concentrations of SDS. *FEBS Letters* 529: 193–197.
- Poewe, W., Seppi, K., Tanner, C.M., Halliday, G.M., Brundin, P., Volkman, J., Schrag, A.-E., and Lang, A.E. (2017). Parkinson disease. *Nature Reviews Disease Primers* 3: 17013.
- Poshusta, T.L., Sikkink, L.A., Leung, N., Clark, R.J., Dispenzieri, A., and Ramirez-Alvarado, M. (2009). Mutations in specific structural regions of immunoglobulin light chains are associated with free light chain levels in patients with AL amyloidosis. *PloS one* 4: e5169.
- Pradhan, T., Annamalai, K., Sarkar, R., Huhn, S., Hegenbart, U., Schönland, S., Fändrich, M., and Reif, B. (2020). Seeded fibrils of the germline variant of human λ -III immunoglobulin light chain FOR005 have a similar core as patient fibrils with reduced stability. *Journal of Biological Chemistry* 295: 18474–18484.
- Pratt, W.B. and Dittmar, K.D. (1998). Studies with purified chaperones advance the understanding of the mechanism of glucocorticoid receptor–hsp90 heterocomplex assembly. *Trends in Endocrinology & Metabolism* 9: 244–252.
- Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zerovnik, E., and Razgulyaev, O.I. (1990). Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Letters* 262: 20–24.
- Qin, Z., Hu, D., Zhu, M., and Fink, A.L. (2007). Structural characterization of the partially folded intermediates of an immunoglobulin light chain leading to amyloid fibrillation and amorphous aggregation. *Biochemistry* 46: 3521–3531.
- Querfurth, H.W. and LaFerla, F.M. (2010). Alzheimer's Disease. *New England Journal of Medicine* 362: 329–344.
- Raffen, R., Dieckman, L.J., Szpunar, M., Wunschl, C., Pokkuluri, P.R., Dave, P., Stevens, P.W., Cai, X., Schiffer, M., and Stevens, F.J. (1999). Physicochemical consequences of amino acid variations that contribute to fibril formation by immunoglobulin light chains. *Protein Science* 8: 509–517.
- Rauch, J.N., Tse, E., Freilich, R., Mok, S.-A., Makley, L.N., Southworth, D.R., and Gestwicki, J.E. (2017). BAG3 Is a Modular, Scaffolding Protein that physically Links Heat Shock Protein 70 (Hsp70) to the Small Heat Shock Proteins. *Journal of Molecular Biology* 429: 128–141.

- Reixach, N., Deechongkit, S., Jiang, X., Kelly, J.W., and Buxbaum, J.N. (2004). Tissue damage in the amyloidoses: Transthyretin monomers and nonnative oligomers are the major cytotoxic species in tissue culture. *Proceedings of the National Academy of Sciences* 101: 2817–2822.
- Rennella, E., Morgan, G.J., Kelly, J.W., and Kay, L.E. (2019a). Role of domain interactions in the aggregation of full-length immunoglobulin light chains. *Proceedings of the National Academy of Sciences* 116: 854–863.
- Rennella, E., Morgan, G.J., Yan, N., Kelly, J.W., and Kay, L.E. (2019b). The Role of Protein Thermodynamics and Primary Structure in Fibrillogenesis of Variable Domains from Immunoglobulin Light Chains. *Journal of the American Chemical Society* 141: 13562–13571.
- Richter, K., Haslbeck, M., and Buchner, J. (2010). The heat shock response: life on the verge of death. *Molecular Cell* 40: 253–266.
- Riek, R. and Eisenberg, D.S. (2016). The activities of amyloids from a structural perspective. *Nature* 539: 227–235.
- Rius, B., Mesgarzadeh, J.S., Romine, I.C., Paxman, R.J., Kelly, J.W., and Wiseman, R.L. (2021). Pharmacologic targeting of plasma cell endoplasmic reticulum proteostasis to reduce amyloidogenic light chain secretion. *Blood Advances* 5: 1037–1049.
- Roe, S.M., Ali, M.M.U., Meyer, P., Vaughan, C.K., Panaretou, B., Piper, P.W., Prodromou, C., and Pearl, L.H. (2004). The Mechanism of Hsp90 Regulation by the Protein Kinase-Specific Cochaperone p50(cdc37). *Cell* 116: 87–98.
- Rosenzweig, R., Nillegoda, N.B., Mayer, M.P., and Bukau, B. (2019). The Hsp70 chaperone network. *Nature Reviews Molecular Cell Biology* 20: 665–680.
- Rottenaicher, G.J., Weber, B., Rührnöfl, F., Kazman, P., Absmeier, R.M., Hitzenberger, M., Zacharias, M., and Buchner, J. (2021). Molecular mechanism of amyloidogenic mutations in hypervariable regions of antibody light chains. *Journal of Biological Chemistry* 296: 100334.
- Ruiz-Zamora, R.A., Guillaumé, S., Al-Hilaly, Y.K., Al-Garawi, Z., Rodríguez-Alvarez, F.J., Zavala-Padilla, G., Pérez-Carreón, J.I., Rodríguez-Ambriz, S.L., Herrera, G.A., and Becerril-Luján, B., et al. (2019). The CDR1 and Other Regions of Immunoglobulin Light Chains are Hot Spots for Amyloid Aggregation. *Scientific Reports* 9: 3123.
- Sacchettini, J.C. and Kelly, J.W. (2002). Therapeutic strategies for human amyloid diseases. *Nature Reviews Drug Discovery* 1: 267–275.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Cold spring harbor laboratory press.
- Santhorawala, V., Palladini, G., Kukreti, V., Zonder, J.A., Cohen, A.D., Seldin, D.C., Dispenzieri, A., Jaccard, A., Schönland, S.O., and Berg, D. (2017). A phase 1/2 study of

- the oral proteasome inhibitor ixazomib in relapsed or refractory AL amyloidosis. *Blood* 130: 597–605.
- Santoro, M.M. and Bolen, D.W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27: 8063–8068.
- Saunders, A.M., Strittmatter, W.J., Schmechel, D., George-Hyslop, P.S., Pericak-Vance, M.A., Joo, S.H., Rosi, B.L., Gusella, J.F., Crapper-MacLachlan, and Alberts, M.J. (1993). Association of apolipoprotein E allele ϵ 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43: 1467.
- Schmid, A.B., Lagleder, S., Gräwert, M.A., Röhl, A., Hagn, F., Wandinger, S.K., Cox, M.B., Demmer, O., Richter, K., and Groll, M., et al. (2012). The architecture of functional modules in the Hsp90 co-chaperone Sti1/Hop. *The EMBO Journal* 31: 1506–1517.
- Schmidt, M., Rohou, A., Lasker, K., Yadav, J.K., Schiene-Fischer, C., Fändrich, M., and Grigorieff, N. (2015). Peptide dimer structure in an A β (1–42) fibril visualized with cryo-EM. *Proceedings of the National Academy of Sciences* 112: 11858–11863.
- Schopf, F.H., Biebl, M.M., and Buchner, J. (2017). The HSP90 chaperone machinery. *Nature Reviews Molecular Cell Biology* 18: 345–360.
- Schroeder Jr, H.W. and Cavacini, L. (2010). Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology* 125: S41-S52.
- Schuck, P. (2000). Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophysical Journal* 78: 1606–1619.
- Shehu, A., Kaviraki, L.E., and Clementi, C. (2007). On the Characterization of Protein Native State Ensembles. *Biophysical Journal* 92: 1503–1511.
- Solomon, A., Frangione, B., and Franklin, E.C. (1982). Bence Jones proteins and light chains of immunoglobulins. Preferential association of the V lambda VI subgroup of human light chains with amyloidosis AL (lambda). *The Journal of Clinical Investigation* 70: 453–460.
- Solomon, A., Weiss, D.T., Murphy, C.L., Hrcic, R., Wall, J.S., and Schell, M. (1998). Light chain-associated amyloid deposits comprised of a novel kappa constant domain. *Proceedings of the National Academy of Sciences* 95: 9547–9551.
- Stoscheck, C.M. (1990). [6] Quantitation of protein. In: Deutscher, M.P. (Ed.). *Methods in Enzymology : Guide to Protein Purification*. Academic Press, pp. 50–68.
- Stryer, L. (1965). The interaction of a naphthalene dye with apomyoglobin and apohemoglobin: A fluorescent probe of non-polar binding sites. *Journal of Molecular Biology* 13: 482–495.
- Sunde, M. and Blake, C. (1997). The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Advances in Protein Chemistry* 50: 123–159.

- Sunde, M., Serpell, L.C., Bartlam, M., Fraser, P.E., Pepys, M.B., and Blake, C.C.F. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* 273: 729–739.
- Svilenov, H.L., Sacherl, J., Protzer, U., Zacharias, M., and Buchner, J. (2021). Mechanistic principles of an ultra-long bovine CDR reveal strategies for antibody design. *Nature Communications* 12: 1–13.
- Swindells, M.B., Porter, C.T., Couch, M., Hurst, J., Abhinandan, K.R., Nielsen, J.H., Macindoe, G., Hetherington, J., and Martin, A.C.R. (2017). abYsis: Integrated Antibody Sequence and Structure-Management, Analysis, and Prediction. *Journal of Molecular Biology* 429: 356–364.
- Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K.D., Karras, G.I., and Lindquist, S. (2012). Quantitative Analysis of Hsp90-Client Interactions Reveals Principles of Substrate Recognition. *Cell* 150: 987–1001.
- Teng, G. and Schatz, D.G. (2015). Chapter One - Regulation and Evolution of the RAG Recombinase. In: Murre, C. (Ed.). *Advances in Immunology : Molecular Mechanisms that Orchestrate the Assembly of Antigen Receptor Loci*. Academic Press, pp. 1–39.
- Tennent, G.A., Lovat, L.B., and Pepys, M. (1995). Serum amyloid P component prevents proteolysis of the amyloid fibrils of Alzheimer disease and systemic amyloidosis. *Proceedings of the National Academy of Sciences* 92: 4299–4303.
- Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature* 302: 575–581.
- Tsolis, A.C., Papandreou, N.C., Iconomidou, V.A., and Hamodrakas, S.J. (2013). A consensus method for the prediction of 'aggregation-prone' peptides in globular proteins. *PloS one* 8: e54175.
- Udgaonkar, J.B. and Baldwin, R.L. (1988). NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature* 335: 694–699.
- Valdés-García, G., Millan-Pacheco, C., and Pastor, N. (2017). Convergent mechanisms favor fast amyloid formation in two lambda 6a Ig light chain mutants. *Biopolymers* 107: e23027.
- van der Kant, R., Bauer, J., Karow-Zwick, A.R., Kube, S., Garidel, P., Blech, M., Rousseau, F., and Schymkowitz, J. (2019). Adaption of human antibody λ and κ light chain architectures to CDR repertoires. *Protein Engineering, Design & Selection* 32: 109–127.
- Vembar, S.S. and Brodsky, J.L. (2008). One step at a time: endoplasmic reticulum-associated degradation. *Nature Reviews Molecular Cell Biology* 9: 944–957.
- Villalba, M.I., Canul-Tec, J.C., Luna-Martínez, O.D., Sánchez-Alcalá, R., Olamendi-Portugal, T., Rudiño-Piñera, E., Rojas, S., Sánchez-López, R., Fernández-Velasco, D.A., and Becerril, B. (2015). Site-directed mutagenesis reveals regions implicated in the stability and fiber formation of human λ 3r light chains. *Journal of Biological Chemistry*: 2577–2592.

- Voet, D., Voet, J.G., and Pratt, C.W. (2016). *Fundamentals of biochemistry: life at the molecular level*. John Wiley & Sons.
- Wall, J., Schell, M., Murphy, C., Hrcic, R., Stevens, F.J., and Solomon, A. (1999). Thermodynamic instability of human lambda 6 light chains: correlation with fibrillogenicity. *Biochemistry* 38: 14101–14108.
- Walsh, C.T., Garneau-Tsodikova, S., and Gatto Jr., G.J. (2005). *Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications*. *Angewandte Chemie International Edition* 44: 7342–7372.
- Wang, F., Ekiert, D.C., Ahmad, I., Yu, W., Zhang, Y., Bazirgan, O., Torkamani, A., Raudsepp, T., Mwangi, W., and Criscitiello, M.F. (2013). Reshaping antibody diversity. *Cell* 153: 1379–1393.
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., Beer, T.A.P. de, Rempfer, C., and Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46: W296-W303.
- Weber, B., Hora, M., Kazman, P., Göbl, C., Camilloni, C., Reif, B., and Buchner, J. (2018). The Antibody Light-Chain Linker Regulates Domain Orientation and Amyloidogenicity. *Journal of Molecular Biology* 430: 4925–4940.
- Weber, B., Hora, M., Kazman, P., Pradhan, T., Rührnößl, F., Reif, B., and Buchner, J. (2020). Domain Interactions Determine the Amyloidogenicity of Antibody Light Chain Mutants. *Journal of Molecular Biology* 432: 6187–6199.
- Wetlaufer, D.B. (1973). Nucleation, Rapid Folding, and Globular Intrachain Regions in Proteins. *Proceedings of the National Academy of Sciences* 70: 697–701.
- Whitesell, L. and Lindquist, S.L. (2005). HSP90 and the chaperoning of cancer. *Nature Reviews Cancer* 5: 761–772.
- Wiseman, R.L., Mesgarzadeh, J.S., and Hendershot, L.M. (2022). Reshaping endoplasmic reticulum quality control through the unfolded protein response. *Molecular Cell* 82: 1477–1491.
- Wolwertz, M.L., Nguyen, P.T., Quittot, N., and Bourgault, S. (2016a). Probing the role of lambda 6 immunoglobulin light chain dimerization in amyloid formation. *Biochimica et Biophysica Acta* 1864: 409–418.
- Wolwertz, M.L., Nguyen, P.T., Quittot, N., and Bourgault, S. (2016b). Probing the role of lambda 6 immunoglobulin light chain dimerization in amyloid formation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1864: 409–418.
- Woody, R.W. (1995). [4] Circular dichroism. In: *Biochemical Spectroscopy*. Elsevier, pp. 34–71.

- Wu, T.T. and Kabat, E.A. (1970). 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: 211–250.
- Wyatt, A.R., Yerbury, J.J., Berghofer, P., Greguric, I., Katsifis, A., Dobson, C.M., and Wilson, M.R. (2011). Clusterin facilitates in vivo clearance of extracellular misfolded proteins. *Cellular and Molecular Life Sciences* 68: 3919–3931.
- Wyatt, A.R., Yerbury, J.J., Ecroyd, H., and Wilson, M.R. (2013). Extracellular chaperones and proteostasis. *Annual Review of Biochemistry* 82: 295–322.
- Yan, N.L., Santos-Martins, D., Nair, R., Chu, A., Wilson, I.A., Johnson, K.A., Forli, S., Morgan, G.J., Petrassi, H.M., and Kelly, J.W. (2021). Discovery of Potent Coumarin-Based Kinetic Stabilizers of Amyloidogenic Immunoglobulin Light Chains Using Structure-Based Design. *Journal of Medicinal Chemistry* 64: 6273–6299.
- Yan, N.L., Santos-Martins, D., Rennella, E., Sanchez, B.B., Chen, J.S., Kay, L.E., Wilson, I.A., Morgan, G.J., Forli, S., and Kelly, J.W. (2020). Structural basis for the stabilization of amyloidogenic immunoglobulin light chains by hydantoins. *Bioorganic & Medicinal Chemistry Letters* 30: 127356.
- Ye, J., Ma, N., Madden, T.L., and Ostell, J.M. (2013). IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res* 41: W34-40.
- Yu, X.C. and Strobel, H.W. (1996). Interactions of 8-anilino-1-naphthalenesulfonic acid (ANS) and cytochrome P450 2B1: role of ANS as an effector as well as a reporter group. *Molecular and Cellular Biochemistry* 162: 89–95.
- Zhang, C., Huang, X., and Li, J. (2017). Light chain amyloidosis: Where are the light chains from and how they play their pathogenic role? *Blood Reviews* 31: 261–270.

6. Publications and Presentations

Publications

Rottenaicher, G.J., Weber, B., Rührnößl, F., Kazman, P., Absmeier, R.M., Hitzemberger, M., Zacharias, M., and Buchner, J. (2021). Molecular mechanism of amyloidogenic mutations in hypervariable regions of antibody light chains. *Journal of Biological Chemistry*, 296: 100334, doi: 10.1016/j.jbc.2021.100334

Schürmann, J., Gottwald, J., Rottenaicher, G., Tholey, A., and Röcken, C. (2021). MALDI mass spectrometry imaging unravels organ and amyloid-type specific peptide signatures in pulmonary and gastrointestinal amyloidosis. *Proteomics Clinical Applications*, 6: e2000079, doi: 10.1002/prca.202000079

Absmeier, R.M., Rottenaicher, G.J., Svilenov, H.L., Kazman, P., and Buchner, J. (2022). Antibodies gone bad – the molecular mechanism of light chain amyloidosis. *FEBS Journal*, doi: 10.1111/febs.16390 (Review article)

Conference attendances with presentations

Poster: Investigations on the molecular pathogenesis of systemic light chain (AL) amyloidosis. *4th Meeting on Biophysics of Amyloid Formation, Ulm, Germany, February 2020*

Talk: Insights into the mechanism of amyloid formation in systemic light chain amyloidosis. *6th Meeting on Biophysics of Amyloid Formation, Ulm, Germany, February 2022*

7. Acknowledgements

An erster Stelle möchte ich mich bei meinem Doktorvater Johannes Buchner bedanken: für die Möglichkeit meine Doktorarbeit am Lehrstuhl für Biotechnologie anzufertigen, für die Hilfe und das Vertrauen, das er mir entgegengebracht hat und für die nötige Freiheit, um mein Promotionsprojekt selbst mitzugestalten und meine eigenen Ideen umzusetzen.

Außerdem bedanke ich mich bei allen Mitgliedern der Forschergruppe FOR2969. Es war eine tolle Erfahrung mit Wissenschaftlerinnen und Wissenschaftlern aus unterschiedlichen Fachbereichen zusammenzuarbeiten, sich auszutauschen, dabei Neues zu lernen und Fragestellungen aus einem anderen Blickwinkel zu betrachten. Des Weiteren gilt mein Dank den Kooperationspartnern an der TUM: Teji, Olga und Bernd Reif vom Bayerischen NMR-Zentrum sowie Manuel Hitzenberger und Martin Zacharias vom Lehrstuhl für Theoretische Biophysik für die gute Zusammenarbeit. Auch meinen Praktikantinnen und Praktikanten möchte ich herzlich danken für die Unterstützung im Labor und dafür, dass sie mir oft eine neue Sichtweise auf manche Dinge vermittelt haben.

Ganz besonders möchte ich mich bei meinen aktuellen und ehemaligen Kollegen am Lehrstuhl für Biotechnologie bedanken, insbesondere bei Tuan, Laura, Ramona, Flo, Chris, Maxi, Sonja, Annika, Benji, Pamina, Bene und Mareike. Die letzten Jahre mit euch allen waren eine sehr schöne und lehrreiche Zeit. Seien es die interessanten Fachdiskussionen oder die gemütlichen Abende bei einem kühlen Bier: Dank euch wird mir meine Promotion immer als eine tolle Zeit in Erinnerung bleiben.

Natürlich möchte ich mich auch bei meiner ganzen Familie, bei meiner Schwester Monika, meiner Nichte Julia und meinen Neffen Ben und Alex bedanken. Doch der wohl größte Dank, den ich hier loswerden möchte, geht an meine Eltern, denen diese Doktorarbeit gewidmet ist. Ihr habt mich immer bedingungslos unterstützt und ohne euch hätte ich es niemals bis hierher geschafft!

Zu guter Letzt noch ein großes Dankeschön an Steffi: wir haben in den vergangenen Jahren sehr viel zusammen erlebt und geschafft und ohne dich hätte ich das Studium und die Promotion wohl kaum gepackt. Du warst immer für mich da und hattest immer ein offenes Ohr für meine Probleme und Sorgen. Ich bin sehr froh dich an meiner Seite zu haben. Vielen Dank für Alles!

8. Declaration

I, Georg Josef Rottenaicher, hereby declare that this thesis was prepared by me independently and using only references and resources stated here. The work has not been submitted to any audit commission, yet. Parts of this work were published in scientific journals.

Hiermit erkläre ich, Georg Josef Rottenaicher, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die Arbeit wurde bisher noch keiner Prüfungsbehörde vorgelegt. Teile dieser Arbeit wurden in wissenschaftlichen Journalen veröffentlicht.

Georg J. Rottenaicher