TECHNISCHE UNIVERSITÄT MÜNCHEN

Department Chemie

Lehrstuhl für Biotechnologie

Analysis of different antibody domain mutations affecting the V_H/V_L interface, C_H1 folding and C_H2 stability

Christine John

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr. Bernd Reif
Prüfer der Dissertation:	1. Univ.-Prof. Dr. Johannes Buchner
	2. Univ.-Prof. Dr. Matthias Feige

Die Dissertation wurde am 27.10.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 05.12.2016 angenommen.

Contents

Contents

Summary

Antibodies are essential effectors of the adaptive immune system of higher eukaryotes. They are divided in different subclasses that cover multiple functionalities and places of action. With regard to the usage of antibodies for diagnostic and therapeutic applications, the improvement of antibodies by engineering represents a research field of great importance.

Within this study, different isolated antibody domains were investigated *in vitro.* The variable domains of the human 1HEZ IgM antibody were examined with regard to the highly conserved amino acids in the V_H/V_L interface, identified by Wang and coworkers within a covariation analysis [5]. The question of the potential importance of these residues for I) folding II) stability or III) V_H/V_L interaction was addressed by an alanine point mutation approach. It could be shown that V_H residues E46 and R38 are essential for folding. Moreover many of the residues affect the domain stability especially residue Y36 for V_L and residues E46 and W47 for V_H. In terms of V_H/V_L interaction, mutation of residues P44 in V_L and L45 and V37 in V_H led to severe impairment of the binding to the partner domain. In comparison to the data of Dr. Eva-Maria Herold, who investigated the same point mutations within a murine IgG antibody (MAK33), very similar results were obtained. This allows to draw a general conclusion about the observed effects and the associated relevance of the conserved residues. The generated data can be useful for the engineering of therapeutic antibodies, as it provides knowledge about the influence of the conserved V_H/V_L interface residues on folding, stability and V_H/V_L interaction.

The second project involved the stabilization of the isolated MAK33 C_H2 domain. It was shown that a total of 16 additional amino acids at the C-and N-terminus of C_H2 led to a strong increase of 21 °C for the T_m of thermal unfolding. In this study, truncation mutants with different numbers of C-and N-terminal extra amino acids were characterized under physiological and pH 2 conditions, with the aim to identify the residues that confer stability. The results show that a stabilizing effect can be achieved by addition of either N- or C-terminal extra amino acids, but the effect is slightly more pronounced with C-terminal residues. The number of amino acids was reduced down to GSA N-terminally and GSGS C-terminally. A similar stabilizing effect for all truncation mutants studies was observed. In this context, a further trimming to only one amino acid per terminus would

be of interest. This effect might be related to the consensus sequence GS or induced by a random C-or N-terminal addition of amino acids. An extension of these studies on other antibody domains could reveal whether this effect is of common nature or only associated to C_H2.

The third project aimed to identify key regions or amino acids responsible for the C_H1 domain to be unfolded in isolation. In contrast to all other immunoglobulin domains in isolation, C_H1 is unfolded [6]. Folding is induced by the presence of the C_L domain. In contrast to C_L and the other constant domains, C_H1 shows structural divergences including two helical regions that are present in the other Ig domains. Two approaches were applied to generate C_H1 mutants that are potentially able to fold in isolation by the transfer of structural features of C_L on C_H1. The first approach was based on a computational comparison of the MAK33 C_L and C_H1 structures, in order to identify divergent regions and amino acid interactions. As a result, two mutants with several amino acid exchanges between C_L and C_H1 were produced with the aim to generate a foldable C_H1 and an unfoldable C_L domain that requires the C_L wild type domain for folding. In the second approach, two additional C_H1 mutants were generated with the focus on the exchange of the two highly divergent helical regions and a proline residue. Unfortunately, within this study no foldable C_H1 domain was obtained. However, the mutants of the second approach exhibit an extreme acceleration of the folding event upon addition of C_L. They take only a few seconds for folding compared to the wild type C_H1 domain which requires about 50 minutes. This effect is assumed to be caused by proline residue 54.

Taken together, this thesis provides new insights in the role of conserved amino acids in the V_H and V_L interface as well as the stabilizing effect of N-terminal and C-terminal amino acid additions on C_H2, which might also be applicable to other antibody domains. Moreover, a key regulator for C_H1 folding was identified leading to an extraordinary acceleration of the folding reaction. The obtained data will be helpful for antibody engineering.

1. Introduction

1.1. Protein folding

Proteins represent the most abundant macromolecules in all forms of life. They fulfill manifold tasks involved in almost all biological processes. These include functions in the cell like DNA replication, catalysis, transport, storage, cell differentiation and growth as well as immunity. The immense variety of proteins is based on different combinations of the 20 naturally occuring amino acids in polypeptide chains, which are encoded in the gene sequence. After synthesis in the ribosome, the generated amino acid chains have to fold into a distinct three-dimensional structure in order to acquire functionality [7]. This specific conformational state obtained after the folding process is called the native state. In the crowded environment of a cell with proteins and macromolecules reaching a concentration of up to 300 – 400 mg/ml, correct protein folding can become quite challenging [8]. Anfinsen could show that *in vitro* many proteins are able to fold spontaneously governed by their amino acid sequence [9]. If protein folding happened by a random event, a 100 amino acid long protein would require more than one billion years to reach the one native conformation among 10^{30} possible conformations [10]. However, under physiological conditions, most of the proteins adopt a single, thermodynamically stable structure within a time scale of milliseconds to minutes [9, 11, 12]. These observations led to the postulation of “Levinthal’s paradox” which states that folding in one step is not sufficient to adopt the energetically favored native state of a protein. Instead, folding in a biological time frame has to follow a pathway with underlying conformational restrictions of partial structures, which allows a rapid formation of the most energetically stable structure [10, 13].

1.2. Driving forces of protein folding

Protein folding is accompanied by the formation of the so-called weak interactions. Weak interactions are composed of hydrophobic interactions, van der Waals interactions, peptide hydrogen bonds and peptide solvation. Despite the name, the total of all weak interactions is assumed to be the major force of protein folding [14]. Van der Waals interactions describe the attractive and repulsive forces between fixed or induced dipoles while hydrogen bonds are formed between two molecules that share a positively charged

H-atom that is covalently bound to a strongly electronegative atom. Peptide solvation involves the interaction of peptides and water based on an electrostatic character. Hydrogen bonds play a crucial role for peptide solvation since hydrogen bonds between water and the functional peptide groups N-O and C=O have to be disrupted to allow peptide hydrogen bond formation [15].

Hydrophobic free energy interactions are a major driving force of protein folding [15-17]. They are accompanied by a transfer of free energy which results from the repositioning of surface exposed nonpolar side chains to the hydrophobic core [16]. The burial of these residues in a non-aqueous solution is entropically favorable since hydrophobic side chains force the surrounding water molecules into an ordered state that is accompanied by an unfavorable decrease in entropy [17].

Additionally, there are further interactions that stabilize the native state. These include electrostatic interactions between charged side chains, which are crucial for α-helix and ß-sheet formation [15, 18]. But also hydrogen bonds strongly contribute to secondary structure formation. Some more specific pairwise interactions between side chains can be formed by salt bridges and cation-pi interactions. Moreover, covalent bonds like disulfide bonds between two cysteine residues are able to influence protein folding by restricting entropy [18].

The different types of interactions can be located to different regions within the native protein. While hydrogen bonds can be found in the hydrophobic core as well as on the protein surface, hydrophobic interactions are usually limited to the hydrophobic core and electrostatic interactions to the surface. Interactions can also be classified according to the distances between the interacting side chains. Here, van der Waals and hydrophobic interactions are short range interactions whereas Coulomb interactions and salt bridges can be assigned to long range interactions.

To sum up, within a native protein non polar side chains are predominantly directed towards the inner core where they contribute to hydrophobic free energy interactions. In contrast, polar and charged side chains are mainly surface-exposed and involved in electrostatic interactions governed by Coulomb's law, van der Waals and hydrogen bond interactions, including interactions with water molecules. Some polar residues involved in

secondary structure formation can also be found in the hydrophobic core where they are stabilized by hydrogen bonds. The favoring or disfavoring effect of charge-charge interactions energy depends strongly on the pH value and the associated total charge of the protein [15].

1.3. Protein folding mechanisms

The folding process is subject to great variations between different proteins. Hence, several models for protein folding were established which illustrate simplistic approaches to explain the existing experimental data. Three very prominent models represent the "framework", "hydrophobic collapse" and "nucleation condensation" model (Figure 1). The starting point of the "framework" model is an intermediate state that possesses a high degree of secondary structure [19]. The appearance of secondary structure induces the formation of tertiary structure. An alternative to the "framework" model is the "diffusion collision" model, which also starts with an intermediate rich in secondary structure. But here, the secondary structure is not completely formed as it is more unstable and prone to fluctuations. The interactions of secondary structure elements can lead to a further folding [20, 21]. This model provides a good explanation for fast folding proteins without a stable, detectable intermediate. In contrast, the "hydrophobic collapse" model is based on a collapse of hydrophobic regions that lead to a compaction of the unfolded protein. This enables the formation of native-like interactions that are able to induce the formation of tertiary structure [22, 23]. This model incorporates the presence of the molten globule state. The nucleation-condensation model includes contents of the first two mechanisms and fits very well to small, single domain proteins with a two-state folding behavior [24-26]. In this model, folding is initiated by a folding nucleus that accumulates secondary and tertiary structure in parallel. Thus instable short-range secondary structure interactions are stabilized by long-range interactions of the native-like tertiary structure.

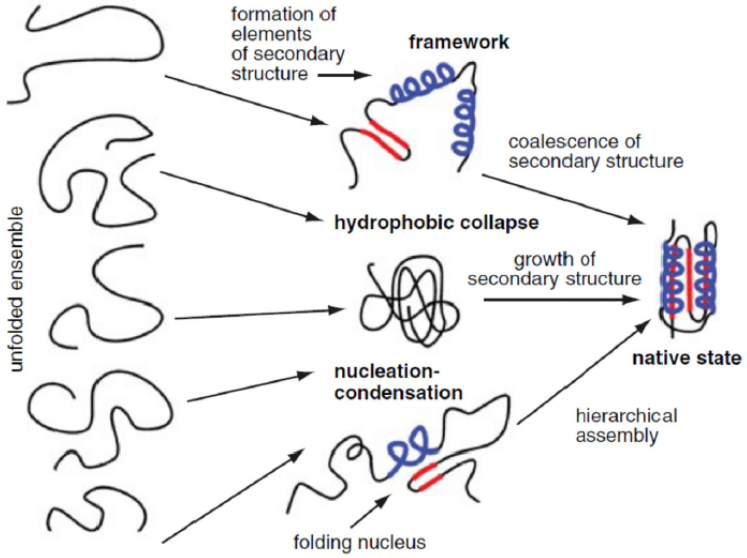


Figure 1: Prominent mechanisms for two-dimensional protein folding.
Three folding models with different principles of native state assembly are shown: framework, hydrophobic collapse and nucleation-condensation model. The entropy decreases with the increase of structural complexity (from left to right). The figure is taken from [13].

However, the complex process of protein folding with the involvement of multiple forces and energetic barriers requires a three-dimensional model. Hence, the folding funnel (Figure 2) was established to explain the cooperative character of folding with regard to the energy landscape and the enthalpic and entropic forces, that guide the folding process [27]. In this model, an unfolded protein moves downwards towards the native state which depicts the global energetic minimum of the funnel. The ragged energy surface provides several routes with different energetic intermediate conformational states. The landscape also includes some local minima traps in which proteins can get caught if they are not able to overcome certain energy barriers [28]. Although this model has some limitations, it accounts for the majority of the observed folding behaviors even unique ones [29].

The folding funnel also describes the effects observed for introduced point mutations. As long as the mutation does not affect the protein stability the same folding route is retained [30]. For some point mutations, though, it could be shown that they are able to reshape the folding landscape or redirect the folding pathway [30-32].

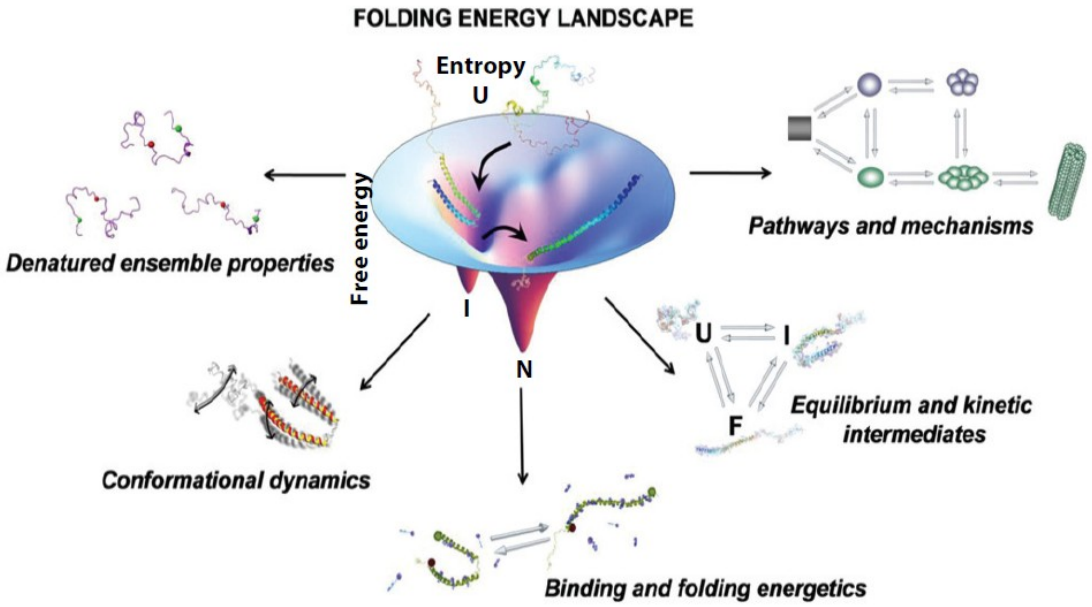


Figure 2: Three dimensional energy landscape of the folding funnel.
Protein folding starts with the unfolded state U and moves towards the native state N, across the rugged energy landscape of the funnel including kinetic traps and energy barriers with intermediates. The native state represents the favorable energetic minimum. The width of the funnel corresponds to the entropy, the depth to the free energy. Additionally, fields of importance for protein folding are schematically shown on the sides. The figure is taken from [33]

1.4. The difference between native and denatured state

All the introduced folding models illustrated the denatured state as a complete unstructured state and the native state as one distinct conformational state. Methods like NMR-spectroscopy and X-ray-crystallography with a high resolution down to 1 Å confirmed that the native state is usually assigned to one conformational state. However, recent studies point towards a residual, fluctuating structure of the native state [34-38]. Moreover, a new a class of proteins was detected, the intrinsically disordered proteins, which represent about 40% of the human proteome. Many of these proteins undergo conformational changes associated with an increase of structural elements upon ligand binding [39-41].

The denatured state is defined as a reversible, noncovalent and cooperative alteration from the native state [42]. It was shown that the denatured state does not only represent one conformational state but rather an ensemble of varying states [43]. So the denatured state, resulting from chemical and thermal denaturation can still differ in the remaining

structured segments [44]. Chaotropic agents like GdmCl and urea, for example, are more effective in interrupting noncovalent interactions than acidic or thermal denaturation [44]. Generally, denaturants induce the solvation of nonpolar residues in water and impair hydrophobic interactions [42, 45]. One also has to distinguish between denatured and unfolded state as not all denatured states can be assigned to the unfolded state. The unfolded state represents a state that shows as much similarity as possible with a random coil structure. In a random coil state, the rotation angles of all backbone bonds and side chains are independent of the angles of distant bonds leading to conformations with a comparable freedom of energy [46]. Restrictions are only set by steric repulsion between atoms with high proximity in the covalent structure and the excluded volume effect [46]. So only those denatured states involving a highly open and solvent-exposed conformation with little or no structure, which is usually the case under strong denaturing conditions, can be classified as unfolded [47].

What separates the folded from the unfolded state is the activation barrier. This has to be overcome to switch between the two states. The activation barrier which is governed by enthalpy and entropy is also called the Gibbs free energy of activation ($\Delta G = \Delta H - T\Delta S$; ΔH is the enthalpy and ΔS the entropy term) [48]. For some proteins, the activation barrier is so small that folding occurs without any barriers within a time scale of $10^6 – 10^4$ s^{-1} [49]. With regard to ΔG, the increasing order of the protein backbone decreases the entropy due to the accumulating restrictions of chain configurations. While enthalpy is associated with desolvation effects [50], the activation barrier and folding rates seem to correlate with the interaction distances. A majority in local contacts is associated with a fast folding reaction, the opposite effect is seen for long-range contacts [51].

1.5. Intermediates under extreme solvent conditions

Extreme solvent conditions provide a perfect environment to study intermediates since the native state is not stable under these conditions. Hence, intermediates that resemble naturally occurring intermediates on the folding pathway [52-54] can be trapped. At low pH or high temperatures, intermediates that correspond to energetic minima of the folding funnel energy landscape were detected. Among these intermediates the molten globule state, the alternatively folded state (AFS) and amyloid structures were identified [52-55].

The molten globule state depicts one of the best studied intermediates [55, 56]. It describes a partially folded protein with a high degree of native like secondary structure and little amounts of tertiary structure. The resulting weaker compaction affects the hydrodynamic radius, which is about 10-20 % larger than for the native state but smaller than for the unfolded state [40, 55]. Generally, the molten globule state possesses a very low stability and an uncooperative unfolding behavior.

Other trapped intermediates, like the formation of amyloid structures are connected with unfavorable effects in the cell. Fibril formation occurs under extreme solvent conditions by a nucleated self-assembly process [57]. This process can be separated in three phases. In the first, an oligomerization of non-native protein takes place. The second step is initiated by the excess of a critical concentration and involves an elongation phase, where either native or more non-native protein gets integrated. At last, a plateau is reached at which no further polymerization occurs [57-59]. The typical amyloid fibril comprises a cross-ß X-ray fiber diffraction pattern with a core structure of ß-sheets and strands perpendicular to the fiber axis [60]. Amyloid aggregates can be found associated with several severe degenerative disorders like Alzheimer's, Parkinson's and Huntington's disease [61]. In the context of antibodies, a disease causing light chain and heavy chain amyloidosis was observed [57].

Some proteins, especially antibodies, exhibit the formation of an alternatively folded state (AFS) under acidic conditions which is clearly differing from the native state and folding intermediates [62]. It is similar to the molten globule state with a high presence of secondary structure but usually possesses a defined quaternary structure. Additionally, compared to the molten globule state the AFS is very stable and shows a cooperative unfolding behavior [62]. The existence of the AFS was first discovered within the context of an IgG antibody [62]. Further studies confirmed the presence of the AFS also for the antibody Fab fragment [63] and single domains like C_H3 [64]. A heat-induced AFS state could be described for the Fc fragment [65].

1.6. Protein folding in the cell

Protein folding within the crowded space of the cytoplasm with protein concentrations between 300 – 400 mg/ml is a challenging task [66]. Hence, the cell established a complex

protein folding network including a stringent control mechanism that is able to initiate apoptosis in case of uncontrolled accumulation of misfolded and aggregated proteins. The so-called molecular chaperones interact with and assist newly synthesized proteins in their folding in order to prevent protein aggregation [67, 68]. Since the presence and up-regulation of most of these chaperones was first discovered with regard to the heat shock response, they are also called heat shock proteins or Hsps [69]. They are composed of five major classes that are named according to their molecular size: the Hsp100, Hsp 90, Hsp70 and Hsp60 families as well as the small heat shock proteins (sHsps). Many of these family members are constitutively expressed in the cell and can be up-regulated upon different types of stress including heat stress, oxidative stress and ER stress [4]. In higher eukaryotes a distinct cytosolic chaperone network can be found that involves the contribution of many Hsp members (Figure 3). Here, the client is transferred between different chaperones and cochaperones to provide a guided folding that can be adapted according to the individual needs of each of the multiple client proteins [70]. For most of the chaperones, except sHsps, the foldase function is associated with ATP-hydrolysis. This represents the driving force for the controlled binding and release of a huge variety of unfolded client proteins. Concerning sHsps, they form functional oligomers that can bind to unfolded proteins and prevent aggregation. However, for further folding assistance and release they require ATP-dependent chaperones [71]. The Hsp60/GroE family shows a different mechanism of protein folding based on substrate encapsulation [72]. Here, two ring-like structures of Hsp60/GroE form a cavity to which the unfolded client can bind, surrounded by a different physicochemical environment than provided in the cytoplasm. The most prominent and ubiquitous chaperones, though, are Hsp90 and Hsp70. Their close collaboration in client maturation can be complemented by several co-chaperones that participate in the formed complex and regulate the ATPase cycle or provide functions like peptidyl-prolyl isomerization and oxidoreductase activity [73, 74] (see Figure 3). Additionally, enzymes like the disulfide isomerase (PDIs), which catalyze the disulfide bond formation, ensure a correct protein folding [73, 75].

The hot spot of folding in the cell is the endoplasmic reticulum (ER) [76]. Here, folded and assembled proteins are transferred to the vesicular transport system via the golgi complex, while partially folded proteins are retained in the ER and misfolded proteins are

sentenced to proteasomal degradation in the cytosol [77, 78]. The ER associated degradation (ERAD) can be activated by different stress conditions that lead to an accumulation of misfolded or unfolded proteins: perturbations in calcium homeostasis or redox status, elevated secretory protein synthesis, expression of misfolded proteins, sugar/glucose deprivation, altered glycosylation or overloading of cholesterol [79-81].

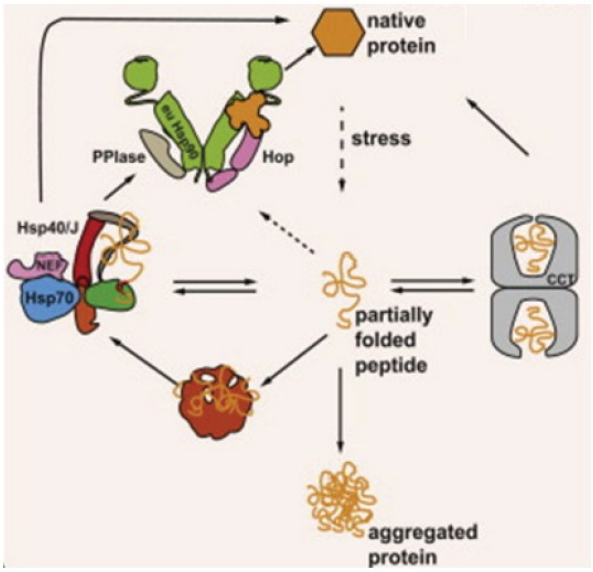


Figure 3: Overview of the molecular chaperone network in the cytosol of higher eukaryotes.
The chaperone network is composed of sHsps (red), the Hsp70 system (blue) with co-chaperones, the Hsp90 system (green) with co-chaperones and the Hsp60 family CCT (grey). Partially folded peptides (clients) can interact with any of the chaperones. In case of Hsp90, Hsp70 supplies the client via Hop. The transfer, binding and release of the clients within the chaperone network is regulated by co-chaperones. The figure is taken from [4] .

Additionally, upon detection of a misbalance in maintaining protein homeostasis, the unfolded protein response (UPR) is activated and initiates a complex signaling cascade that can up-regulate ER chaperones and in the worst case induce apoptosis, if all counteractions for restoring proteome homeostasis fail [82]. This protein quality control mechanism can be found in all kingdoms.

1.7. Immunoglobulins

Immunoglobulins are essential for the adaptive immune defense in higher eukaryotes. They belong to the immunoglobulin superfamily, a large group with members of diverse functionality that all share the immunoglobulin fold [83] (Figure 4). The physiological function of immunoglobulins is the protection of the organism from potential threats ranging from toxins to microbial pathogens [84]. Here, they are not only responsible for the removal of pathogens but also recruit molecular and cellular effector molecules that erase infections and tumor cells [85]. Immunoglobulins exist in two different forms, either as a membrane bound receptor complex that is involved in B lymphocyte development and activation or as a secreted effector molecule, called antibody [86]. They are complex

oligoproteins that are made up of two heavy chains (HCs) and two light chains (LCs) that assemble to a hetero-tetramer [84]. In higher vertebrates five different classes of antibodies exist: IgA, IgD, IgE, IgG and IgM which can be distinguished by their heavy chain constant regions (α,δ ,ε, ɣ and μ chains) that contain either three (IgG, IgA, IgD) or four heavy chain constant domains (IgE, IgM) [87]. Concerning the light chains there are only two types (λ and κ) that differ by distinct structural features and can be combined with any type of heavy chain. The antibody classes differ in their biological function, location in the organism and expression level. IgM antibodies represent the first humoral immune response upon antigen contact. They are secreted as pentamers or hexamers before somatic hypermutation occurs [88]. In this case avidity is assumed to compensate the low affinity [89]. Compared to monomeric antibodies, IgM antibodies possess an additional constant domain (C_H4) with a C-terminal tailpiece that guides polymerization [88]. However, owing to their size, they are not able to pass the extravascular space [90]. The IgG antibody is the most prominent antibody class with the longest serum half-life. In humans four IgG subclasses exist (IgG 1-4) that differ in structure, function and antigens [87]. Among them IgG1 is the most frequent subclass. Regarding a single B cell, only one light chain and one heavy chain allele is expressed in order to ensure specificity for only one antigen by the produced antibodies [91]. Since the gene loci encoding single antibody domains are limited, the organism developed several mechanisms to enhance the possible antibody combinations in order to be able to cope with the immense variety of antigens. These mechanisms mainly affect the variable domains, including somatic hypermutation within the hypervariable loops and combinatorial rearrangement between different gene segments (VDJ recombination of variable, diversity and joining gene segments). Together with the combinatorial association of heavy and light chains, a huge preimmune antibody repertoire of more than 10^{16} different antibodies is available [87].

Generally, hetero-tetrameric antibodies are composed of two heavy (V_H + C_H1-C_H3) and light chains (V_L + C_L). The whole antibody can be divided into two functionally diverse fragments: two antigen binding fragments (Fab) containing a heterodimer of the light chain and V_H and C_H1 of the heavy chain, as well as the crystallizable fragment (Fc) which consists of a heterodimer of C_H2 and C_H3 [92] (see Figure 4). The Fab fragment can be further shortened to the variable fragment (Fv) which is only composed of the variable domains that represent the antigen binding site. Each variable domain contains three

hypervariable loops (complementary determining regions/ CDRs) that interact with the antigen in its entirety. These six CDRs are subject of high variations in sequence and structure in order to provide a persistent high specificity and affinity for a multitude of antigens. They are embedded within a conserved scaffold, the so called framework region [93]. The Fc fragment fulfills different functions like activation of the complement system and the antibody-dependent cell-mediated-cytotoxicity (ADCC), which eradicates the Fab-bound antigens [94]. Additionally, the Fc part can regulate the serum half-life of an antibody by interacting with the neonatal Fc receptor which induces antibody recycling from endosomes after release of the bound antigen [84]. The second constant domain (C_H2) comprises a conserved glycosylation site on asparagine residue 297 [95]. The 30 possible glycosylation patterns for IgG can influence structure, binding and solubility as well as subcellular functions like transport, secretion, clearance and effector functions [87, 95].

The IgG antibody is covalently linked via two disulfide bridges in the hinge region and one between C_H1 and C_L of the Fab fragment. The hinge region allows a high flexibility between the Fab and Fc fragment in order to enhance the accessibility of the Fab fragment for antigen binding [96]. The topology of the antibody structure is highly conserved. The so called Ig fold consists of a greek key ß-barrel structure with seven ß-strands (a, b, c, d, e, f, g) for the constant domains (c-type) and nine ß-strands (a, b, c, c', c'' d, e, f, g) for the variable domains (v-type). They all share a conserved tryptophan residue in the hydrophobic core and a disulfide bridge between strands b and f (Figure 4). The conserved disulfide bond is important for the stabilization of the native structure [97]. This effect is based on the conformational restrictions that are caused by disulfide bonds. The destabilized, reduced C_L domain, for example, promotes the formation of misfolded off-pathway species on the way to the native [97].

Antibodies are of essential importance for diagnostic and therapeutic applications. They are used for the treatment of different diseases like cancer or autoimmune diseases [98] and recent antibody developments even achieved a crossing of the blood brain barrier [99, 100].

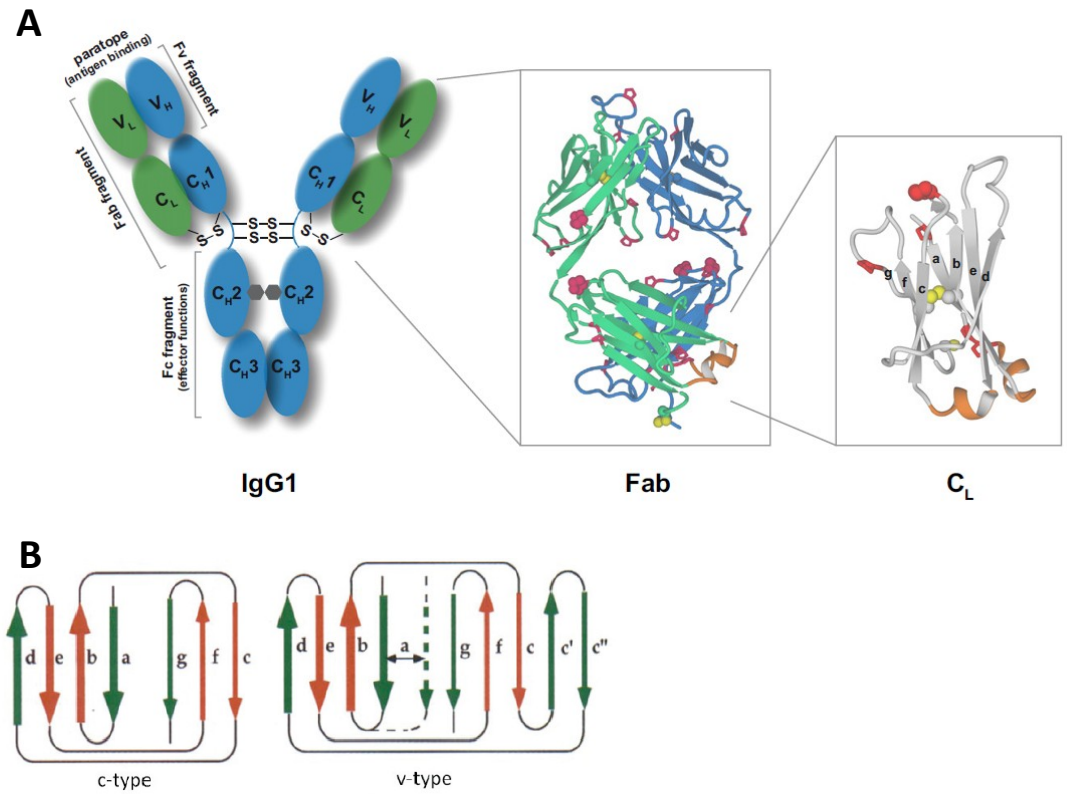


Figure 4: Schematic overview and structural composition of an IgG1 antibody.
(A) Left panel: schematic representation of an IgG1 antibody which consists of two heavy chains (blue) and two light chains (green) that together form two Fab fragments and one Fc fragment. The smallest functional unit depicts the Fv fragment containing the two variable domains that form the antigen binding site (paratope). The C_H2 domains are glycosylated (grey hexagons), the quaternary structure is stabilized by disulfide bridges (S-S) that connect the two heavy chains as well as the heavy and light chain. Middle panel: Structural representation of the Fab fragment, which is covalently linked by a disulfide bond. Additionally, each domain possesses the conserved intramolecular disulfide bond (yellow, CPK representation). Proline residues are highlighted in red, trans prolines are shown in stick representation, cis prolines in CPK representation. The small helices in C_L conserved within the constant domains except for C_H1 are highlighted in orange. Right panel: C_L domain with the seven Ig fold ß-strands (a-g), b and f are connected by the buried disulfide bond. Prolines and small helices are illustrated in the same color code as in the middle panel. (B) Schematic representation of the two conserved Ig fold types with the corresponding ß-strand nomenclature for the variable domains (v-type) and the constant domains (c-type). The figure is adapted from (A)[84]and (B)[101].

1.7.1 Folding of immunoglobulins

Antibody heavy and light chains are co-translationally translocated to the ER [102]. Here, correct folding is supported by the presence of corresponding chaperones and catalyzing enzymes such as PDIs and PPIases as well as oxidative conditions provided by a glutathione based redox system (GSH/GSSG). These oxidative conditions are necessary for the formation of the conserved inter- and intramolecular disulfide bridges [103]. Also post-translational modifications like the glycosylation at the conserved Asn-297

glycosylation site of IgGs is induced in the ER [104]. Single antibody domains initiate their folding by a nucleus formed by ß-strands b,c, e and f [84]. This nucleus is further stabilized by the intramolecular disulfide bridge between strands b and f [105-107]. A rate limiting factor for folding displays the peptidyl-prolyl isomerization of proline residues which is catalyzed by the PPIase cyclophilin B [6]. Antibody domains contain about 5- 10 % prolines, which are mostly located at turns connecting the ß-strands [108, 109].

An investigation of the V_L folding pathway revealed the presence of two folding intermediates within two possible pathways (Figure 5) [3].

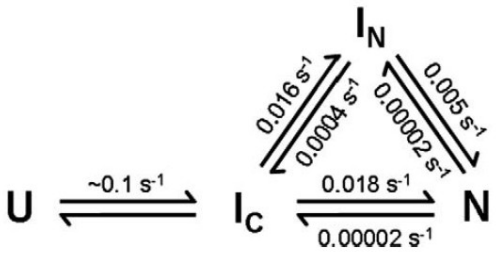


Figure 5: Model for the folding of the MAK33 V_L domain

From the unfolded state an initial intermediate (I_C) is formed, form which two different pathways diverge. The first pathway leads directly to the native state; the second involves a second intermediate (I_N). The figure is taken from [3]

.

The first intermediate (I_C) shows a molten globule-like structure with secondary structure and only small amounts of tertiary structure, from which the native state can be directly reached. An acceleration of this reaction can be achieved by addition of peptidyl-prolyl isomerases, which catalyze proline isomerization. For the alternative pathway, the first intermediate is followed by a second intermediate (I_N), which possesses a higher amount of tertiary structure. Moreover, the second intermediate resembles the native state of amyloid fibrils, associated with the disease light chain amyloidosis. Compared to the folding pathways of the constant domains C_L, C_H2 and C_H3, which belong to the Ig fold c-type, also multiple pathways were observed. However, the timescale of folding is slower for the V_L domain compared to C_L and C_H2, which fold within a few seconds. But due to the slow folding of V_L, the rate limiting effect of proline isomerization on folding is strikingly more pronounced for the C_L and C_H2 domains [97, 110-112].

The formation of the quaternary IgG structure usually starts with the assembly of the HC dimers to which the LC is covalently linked by a disulfide bridge between C_H1 and C_L [113]. Those HCs and LCs that do not find a partner are retained in the ER and sentenced to degradation. However, successful pairing results in B-cell expression of functional antigen

receptors. Upon binding of an antigen the differentiation into plasma cells, that produce huge amounts of antibodies, is initiated [114].

1.7.2 The variable domains

The high affinity and specificity of antibodies provides a high potential for therapeutic applications [115]. Within this context, the Fv fragment containing the variable domains with their antigen binding sites, the complementary determining regions (CDRs), is of special interest. Due to the weak interaction in the absence of the C_H1 and C_L domain, only scFv fragments with a covalent linkage between the variable domains can be used as therapeutic agents [116, 117]. ScFv fragments possess several advantages compared to whole antibodies including an easier expression and folding and a stronger tissue penetration and faster clearance due to their small size [118, 119]. The disadvantages of scFvs are a low thermal stability and higher aggregation tendency [118]. The stability of scFvs depends on the intrinsic stability of the V_H and the V_L domains as well as the dimer interface since V_H/V_L dissociation proceeds unfolding [120]. According to simulations, the disruption of the V_H/V_L interface is mainly caused by a loss of hydrophobic and aromatic interactions [120]. But also in terms of antigen binding the framework residues can play a crucial role. On the one hand, the association of V_L and V_H itself was shown to be crucial for antigen binding on the other hand interface residues are able to affect antigen binding by influencing the positioning of hypervariable loops [121-125]. A lot of research has already been done in the field of antibody engineering for further improvement of the stability and antigen binding to generate more efficient therapeutic antibodies. Studies with a focus on the variable domains depict a rather complex picture. Unfortunately, an increase in affinity often coincides with a decrease in stability and *vice versa* [118, 126-130]. Grafting studies show that the exchange of less stable frameworks with stable ones does not necessarily lead to a stabilization effect as the interactions between CDRs and framework also play an important role [123, 125, 131, 132]. In this context single amino acids can play a key role towards domain orientation and conformation [123, 125]. Hence, the design principles for the architecture of variable antibody domains needs to be elucidated in more detail, with regard to their high potential for therapeutic applications.

One approach for the identification of important residues within the variable domains is a bioinformatical one, more precisely a covariation analysis. This is based on a sequence

comparison of multiple variable domains from different species and subtypes with the aim of identifying conserved amino acid networks. The first studies in this context have been performed by Chothia and coworkers, identifying residues 98, 44 and 36 for V_L and 103, 47 and 37 for V_H [133]. The most recent study is from Wang and coworkers [5], based on a bigger data set for their analysis. In their study they included more than 2000 V-class sequences of human, mouse, cow, camel, llama, macaque and chicken with a bias towards human sequences (574 out of 2432). They could not only confirm most of the findings of Chothia and coworkers but were able to extent the number of conserved residues. One area of high conversation represents the V_H/V_L interaction interface, a second compromises the C_H1/V_H interface, which might be important for framework orientation. Within the V_H/V_L interface the residues with the highest covariation values were the following: Y36, Q37, P44, L46, F98 for V_L; V37, R38, G44, L45, E46, W47, W103 for V_H. All identified V_H residues except E46 and R38 are in direct contact with V_L. V_H W47 seems to be the central node based on the number and strength of its covariations with other interface residues, the same holds true for Y36 and P44 for the V_L domain. Some experimental studies already addressed the influence of the exchange of particular conserved amino acids, including some of the residues identified by Wang and coworkers, on the association of V_H and V_L via the stabilities of covalently linked scFv and Fab fragments [124, 134-138]. One investigation within the Fab format, tested the effect of V_L point mutations Y36A, P44A and F98A as well as V_H point mutations V37A, G44A, L45A and W103A, in terms of Fab expression and antigen binding [124]. They could show that most Fabs containing a point mutation, expressed worse than the wild type, especially L45A, W103A and V37A for V_H and P44A and F98A for V_L. Concerning antigen binding, a strong impairment of binding has been observed for V_H L45A and W103A and V_L F98A.

1.7.3 The C_H2 domain

The C_H2 domain is part of the heavy chain of antibodies; together with the C_H3 domain it forms the Fc fragment of IgG antibodies. One feature discriminates the C_H2 domain from all the other antibody domains: the conserved glycosylation site on the asparagine residue at position 297, which is buried between the two C_H2 domains in a whole molecule antibody (Figure 6) [139-141]. In isolation the C_H2 domain forms a monomer. Within the Fc fragment, the dimer formation is guided by the C_H3 domain [112]. The

glycosylation has a stabilizing effect on the C_H2 domain as it shields the inner hydrophobic residues from solvent exposure. The C_H2 in general, especially the unglycosylated form is quite unstable and aggregation-prone, it can even mediate IgG aggregation [111, 141-144]. In terms of functionality, C_H2 is involved in activation of effector functions and the complement system. Here, different interactions fulfill different functions, by interaction with the neonatal Fc receptor (FcRn) the serum half-life of antibodies is prolonged, by interaction with Fc gamma receptors (FcγRs) Fc-mediated effector functions are activated and the interaction with the C1q component of the complement system is important for pathogen removal [145-147].

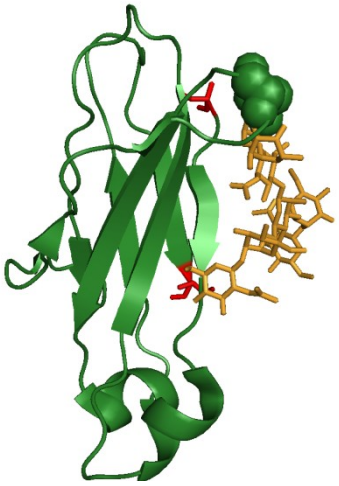


Figure 6: Crystal structure of the glycosylated C_H2 domain (pdb ID 3HKF)
Pymol cartoon representation of the glycosylated MAK33 C_H2 domain. The glycosylation is shown in orange and stick representation, the C_H2 glycosylation site (Asn 297) is illustrated as sphere. The last N- and C-terminal residue is highlighted in red and stick representation.

1.7.4 The C_H1 domain

The C_H1 domain belongs to the heavy chain and is part of the Fab fragment. It interacts with the C_L domain of the light chain, involving a covalent linkage via a disulfide bond. *In vivo* the C_H1 domain acts as a key regulator for the secretion of antibodies from the endoplasmic reticulum (ER) [6, 148]. To fulfill their functions antibodies have to fold into a defined quaternary structure [92, 149]. Hence, the secretion of misfolded antibodies or single chains has to be prevented by a quality control mechanism. This mechanism is based on the C_H1 domain [150], which is an intrinsically disordered protein in isolation, only upon binding of the C_L domain it folds into the typical Ig fold [6]. When entering the ER the unfolded C_H1 domain is recognized by an ER located Hsp70 chaperone called BiP, which interacts with C_H1 via the same interface residues as C_L [6, 151-153]. While light chains can be secreted, unassembled heavy chains are retained in the ER [148, 154, 155].

Domain swap experiments between light and heavy chain as well as diseases connected with C_H1 deletions and substitutions [6, 148, 156] could prove that the key domain for the interaction with BiP and the associated ER retention is the C_H1 domain. The interaction of BiP with C_H1 involves cycles of ATP-dependent binding and release [151, 152, 157]. As soon as C_L is present and interacts with C_H1 during a BiP release cycle, the formation of the intramolecular disulfide bridge is triggered. This is an essential event since C_L is not able to bind to the reduced C_H1 domain. After the release from BiP, the folding process by C_L induces the formation of the hydrophobic folding nucleus which is followed by complete folding of the C_H1 domain [6, 152] (see Figure 7).

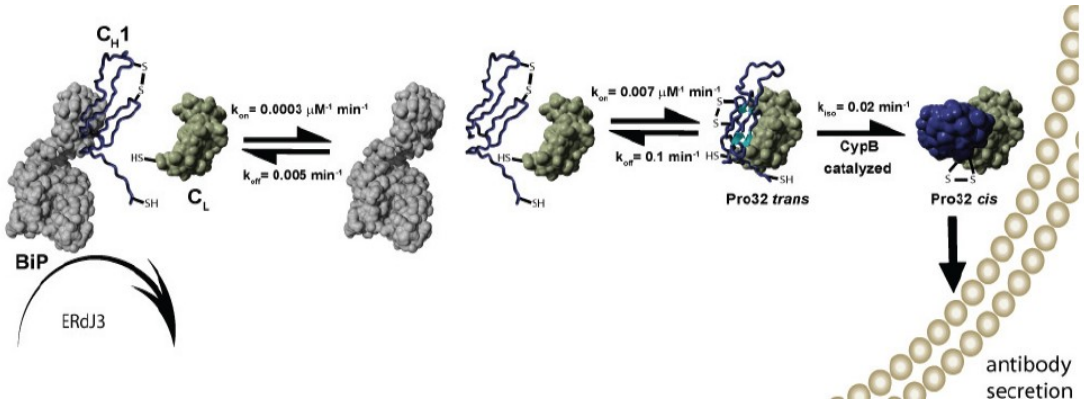


Figure 7: C_L induced folding of the C_H1 domain in the ER

Reduced C_H1 (dark blue) enters the ER as an intrinsically disordered protein and immediately binds to the ER-specific chaperone BiP. This interaction involves cycles of nucleotide dependent binding and release which allows the assembly of C_H1 with C_L (green) since C_L is not able to bind BIP-associated C_H1. Another prerequisite for C_L binding is the preceding oxidation of C_H1 which is probably promoted by additional components in the ER. As soon as C_L binds to C_H1 folding occurs via an intermediate, which is caused by the proline 32 cis-trans isomerization. This step can be accelerated by addition of cyclophilin B. After folding of C_H1 an intermolecular disulfide bridge is formed which covalently links both domains. The figure is taken from [6] and was modified by Dr. Natalia Sarmiento.

This involves the essential and rate limiting step of trans-cis isomerization of proline residue 32 [6]. Only in the cis configuration, the intermolecular disulfide bridge can be formed to complete the covalent assembly. When the intermolecular disulfide bridge between C_L and C_H1 is formed, BiP is no longer able to bind to C_H1.

1.8 Objective

Within this study four different isolated antibody domains were investigated *in vitro* with regard to three research topics:

1. The role of conserved amino acids in the V_L and V_H interface
2. Stabilization of the C_H2 domain by N- and C-terminal addition of amino acids
3. Investigation of potential key amino acids/regions responsible for the unfolded nature of the C_H1 domain in isolation

For the studies different well characterized antibodies were used as model systems. The murine MAK33 and human 1GAF IgG1 antibody as well as the human 1HEZ IgM antibody of which all contain a κ light chain [3, 6, 62, 88, 158, 159]. Among them the IgG1 antibody depicts the most frequently occurring antibody class. More importantly, for all antibodies crystal structures are available for the Fab fragment, Fc fragment or single domains (pdb entries: 1FH5, 3HKF; 1GAF; 1HEZ).

The first topic based on the covariation analysis of Wang and coworkers [5] in which they identified highly conserved amino acids in the V_H/V_L interface including five V_L residues (Y36, Q37, P44, L46, F98) and six V_H residues (R38, G44, L45, E46, W47, W103). The potential importance of these residues was assessed by an alanine mutation approach. A comprehensive characterization of the generated point mutations was performed in order to reveal their impact on folding, stability, V_H/V_L interaction and antigen binding.

The starting point of the second topic was the observation of a stabilizing effect of about 21 °C in the thermal unfolding for the C_H2 domain, achieved by the addition of 16 amino acids distributed at the N- and C-terminus. The aim of this study was the identification of the amino acids responsible for this stabilizing effect. Hence, different N- and C-terminal deletion and truncation mutants were generated, going down to either three N-terminal (GSA) or four C-terminal (GSGS) additional amino acids, and characterized with a focus on stability.

The third study focused on the C_H1 domain and its unfolded nature in the absence of the corresponding light chain interaction partner, the C_L domain. The physiological relevance of this behavior is an ER-associated quality control mechanism that ensures that only correctly assembled antibodies are secreted. Since C_H1 represents the only antibody domain which is unfolded in isolation, it is of importance to reveal the underlying structural constraints that cause this effect. Hence, two different approaches were

applied to generate C_H1 mutants that are potentially able to fold. The first approach based on a computational analysis which identified specific C_L and C_H1 related regions and amino acid networks. These data was used to produce C_H1 and C_L swap mutants that potentially possess the features of the *vice versa* domain. For the second approach, proline residue 54 within the BiP binding motif was mutated to alanine and helix 2 of the C_L domain was transferred to C_H1, followed by a characterization with regard to folding behavior.

2. Material and Methods

2.1. Materials

2.1.1. Devices

Autoclave Varioclav	EP-Z H+P
Biacore X100	GE Healthcare
Cell Disruption Apparatus Basic Z	Constant Systems
Centrifuges	
Avanti J25 and J26 XP	Beckman Coulter
Optima XL-A (equipped with FDS)	Beckman Coulter (Aviv)
Optima XL-I	Beckman Coulter
Rotina 46R	Hettich
Rotina 420R	Hettich
Universal 320R	Hettich
Tabletop centrifuge 5418	Eppendorf
Tabletop centrifuge Mikro R200	Hettich
Chromatography systems	
ÄKTA Prime	GE Healthcare
ÄKTA Purifier + REC 112	GE Healthcare
FPLC + REC 112	GE Amersham
Frac-900/950 fraction collectors	GE Healthcare
Superloops (various volumes)	GE Healthcare
Circular dichroism spectropolarimeters	
J710 (with PFD-350S Peltier device)	Jasco
J715 (with PTC 348 WI Peltier device)	Jasco
Thermoblocks	
Digital heat block	VWR
Eppendorf-Thermomixer	Eppendorf
TB1 Thermoblock	Biometra
Fluorescence spectrophotometers	
FluoroMax-4	Horiba Jobin Yvon
Gel documentation system	
Biodoc II	Biometra
ImageQuant 300	GE Healthcare
Image Scanner III	GE Healthcare
Gel electrophoresis and blotting devices	Hoefer
Homogeniser Ultra Turrax DIAX900	Heidolph
HPLC systems	Jasco

Ice maker	Zieger
Incubator	New Brunswick Scientific
Magnetic stirrer	
MR2000	Heidolph
MR3001	Heidolph
MR80	Heidolph
Ultraflex II MALDI ToF/ToF	Bruker Daltonics
Membrane vacuum pump	Sartorius
pH meter	WTW
Power amplifiers EPS 3500, 3501 and 1001	GE Healthcare
MARKII (refractometer)	Leica
Scales	
BP 121 S	Sartorius
BL 310	Sartorius
Thermal cycler Primus 25	MWG
GENios plate reader	Tecan
Thermal cycler MJ Mini 48 well	Biorad
Typhoon 9400	GE Healthcare
Ultra filtration cell 8050	Amicon
UV-Vis spectrophotometers	
Helios γ	Thermo Fisher
UltroSpec 1100 pro	Amersham Biosciences
Nanodrop	Peqlab
Vortex MS2	IKA

2.1.2. Chemicals

2-Mercaptoethanol	Sigma
5,5'-Dithiobis(2-nitrobenzoic acid) (DTNB)	Sigma
8-Anilino-1-naphthalenesulfonic acid (ANS)	Sigma
ABTS tablets	Roche
Acetic acid	Roth
Acrylamid/Bisacrylamide solution 38:2 (40% w:v)	Serva Electrophoresis
Agar Agar	Serva Electrophoresis
Agarose	Serva Electrophoresis
Ammonium persulphate (APS)	Roth

Blocking reagent	Roche
Bromphenol blue	Serva Electrophoresis
Coomassie Blue R	Serva Electrophoresis
Coomassie Brilliant Blue R-250	Serva Electrophoresis
Deoxynucleoside triphosphates (dNTPs)	Roche
Dimethyl sulfoxide (DMSO)	Sigma
Dithiothreitol (DTT)	Roth
EDC	GE Healthcare
EDTA	Merck
EMCH	Thermo scientific
Ethanol	Merck
Glutathione, oxidized (GSSG)	Sigma
Glutathione, reduced (GSH)	Sigma
Glycerol	Roth
Guanidinium chloride (GdmCl)	Sigma
HCl 32%	Merck
Imidazole	Sigma
Isopropyl β-d-1-thiogalaktopyranoside (IPTG)	Serva Electrophoresis
Kanamycin sulphate	Roth
KCl	Carl Roth
KH_2PO_4	Merck
L-Arginine	Sigma-Aldrich
LB medium	Serva Electrophoresis
NaCl	Merck
Na_2HPO_4*2 H_2O	Merck
NaH_2PO_4* H_2O	Merck
NHS	GE Healthcare
Protease inhibitor Mix G, HP	Serva Electrophoresis
Sodium dodecylsulphate (SDS)	Serva Electrophoresis
Stain G	Sigma
TCEP	Pierce
Tetraethylethylendiamin (TEMED)	Roth

Tris	Roth
Triton X-100	Merck
Tween-20	Merck
Urea	Merck

2.1.3. Consumables

Amicon Ultra-15 Centrifugal Filter Units	Millipore
Amicon Ultra-4 Centrifugal Filter Units	Millipore
Biacore Sensor Chip CM5	GE Healthcare
Cuvettes, plastic, 1 mL	Brand
Dialysis membranes Spectra/Por	Spectrum Laboratories
Membrane discs	Sartorius
PCR tubes	BioRad Laboratories
PE tubes, 15 and 50 mL	Greiner & Söhne
Petri dishes, PS, 94 mm	Greiner & Söhne
pH indicator	Merck
Reaction tubes, various volumes	Sarstedt
Sterile filter 0.2 µm	Zefa
Streptawell, streptavidin-coated microplate	Roche Diagnostics

2.1.4. Enzymes, Standards and Kits

SDS-PAGE Standard Low Range	Biorad
peqGold 1 kb ladder Orange G	Peqlab
Pfu DNA polymerase	Promega
Phusion HF DNA polymerase	NEB
Restriction enzymes	NEB and Promega
T4 ligase	Promega
Wizard Plus SV Minipreps DNA Purification System	Promega
Wizard SV Gel and PCR clean-up system	Promega

2.1.5. Biacore kits

Biacore Amine Coupling Kit	GE Healthcare
Biacore Maintenance Kit, type 2	GE Healthcare

2.1.6. Oligonucleotides

QC-PCR Name: Sequence (5`-3`)

1HEZ VL Y36A f: GCTATCTGAATTGGGCGCAGCAGAAACCG

1HEZ VL Y36A r: CGGTTTCTGCTGCGCCCAATTCAGATAGC

1HEZ VL Q37A f: CTGAATTGGTATGCGCAGAAACCG

1HEZ VL Q37A r: CGGTTTCTGCGCATACCAATTCAG

1HEZ VL P44A f: CCGGGTAAAGCAGCGAAACTGCTGA

1HEZ VL P44A r: TCAGCAGTTTCGCTGCTTTACCCGG

1HEZ VL L46A f: GTAAAGCAGCGAAAGCGCTGATTTATGC

1HEZ VL L46A r: GCATAAATCAGCGCTTTCGCTGCTTTAC

1HEZ VL F98A f: CACCGCGTACCGCGGGTCAGGGCAC

1HEZ VL F98A r: GTGCCCTGACCCGCGGTACGCGGTG

1HEZ VH V37A f: GGTATGCATTGGGCGCGTCAGGCACCG

1HEZ VH V37A r: CGGTGCCTGACGCGCCCAATGCATACC

1HEZ VH R38A f: GCATTGGGTTGCGCAGGCACCG

1HEZ VH R38A r: CGGTGCCTGCGCAACCCAATGC

1HEZ VH G44A f: GGCACCGGGTAAAGCGCTGGAATGGGTTG

1HEZ VH G44A r: CAACCCATTCCAGCGCTTTACCCGGTGCC

1HEZ VH L45A_2 f: GCGGAATGGGTTGCACTGATTAGCT

1HEZ VH L45A_2 r: ACCTTTACCCGGTGCCTGACGAAC

1HEZ VH E46A f: GTAAAGGTCTGGCGTGGGTTGCACTG

1HEZ VH E46A r: CAGTGCAACCCACGCCAGACCTTTAC

1HEZ VH W47A f: GTAAAGGTCTGGAAGCGGTTGCACTGATTAG

1HEZ VH W47A r: CTAATCAGTGCAACCGCTTCCAGACCTTTAC

1HEZ VH W103A f: CGAATGATTATGCGGGTCAGGGCAC

1HEZ VH W103A r: GTGCCCTGACCCGCATAATCATTCG

CDR / FW Name: Sequence (5`-3`)

1DH5_MAKVL_F98A_f2: GCGGGTGCAGGTACAAAACTGGAAC

1DH5_MAKVL_F98A_r2: GGTCGGAGGGGTGGTATAATGCTG

1DH5_MAKVL_L46A_f2: GCGCTGATTAAAGGTGCAAGCAGC

1DH5_MAKVL_L46A_r2: ACGCGGACTTTCATGGCTTTTCTGCT

MVH_CDR3DHU_R44A_f2: GCGCTGGAGTGGGTCGCAAC

MVH_CDR3DHU_R44A_r2: CTTTTCCGGAGTCTGGCGAACCC

ELISA and Biacore Name: Sequence (5`-3`)

VL_WT_Link_Nterm_F :ATGTGTGGGAGCGGGAGCGACATTCAGATGACCCAGAGCCC

VL_WT_Rev: TTATTATTTAATTTCAACTTTGGTGCCCTGACC

VL_WT_F: ATGGACATTCAGATGACCCAGAGCC

VL_WT_Link_C term_ R :
TTATTAACACCCGCTCCCGCTTTTAATTTCAACTTTGGTGCCCTGACCAAAG

VH_WT_Link_ N term_ F: ATGTGTGGGAGCGGGAGCGCACAGGTTCAGCTGGTTGAAAGC

VH_WT_R : TTATTAGCTGCTAACGGTAACCAGGGT

VH_WT_F : ATGGCACAGGTTCAGCTGGTTGAAAG

VH_WT_Link_C term_R : TTATTAACACCCGCTCCCGCTGCTGCTAACGGTAACCAGGGTG

VL FLAG cterm rev: TCATCCTTGTAATCTTTAATTTCAACTTTGGTGC

VL FLAG cterm fwd: CGACGATAAGTAATAAAAGCTTGCGGCC

VH FLAG cterm rev: ATCCTTGTAATCGCTGCTAACGGTAACCAG

VH FLAG cterm fwd: GACGACGATAAGTAATAAAAGCTTGCGGCC

VL_WT_N_rev: CGCACACATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTC

VL_WT_N_fwd: AGCGCGAGCGACATTCAGATGACCCAGAGCC

VH_WT_N_rev: CGCACACATGGTATATCTCCTTCT TAAAGT TAA ACAAAATTATTC

VH_WT_N_fwd: AGCGCGAGCGCACAGGTTCAGCTGGTTGAAAG

C_H1 Name Sequence (5`-3`)
CH1_c_SUMO_F: TTTTGGATCCGCCAGCACCAAAGCACCGAGC

CH1_c_SUMO_R: TTTTCTCGAGTTATTTGTCCACGCTTTTATCAACTTTGGTACTG

CH1_no tag _f: TTTTCCATGGCCAGCACCAAAGCACC

C_H2 Name Sequence (5`-3`)
NcoI-CH2-f: CATGCCATGGCGGATCTTGGATCCGCGTCATCTGTCTTCATCTTC

NcoI-CH2-f2: CATGCCATGGCGGATCTTGGATCCGCGTCATCTGTCTTCATCTTCCC

NcoI-CH2-f nat :CATGCCATGGGGTCATCTGTCTTC

CH2 ctr-f nat : CATGCCATGGGGTCATCTGTCTTCATCTTCCCCCC

CH2-r HindIII: GGGAAGCTTGGGTTATTACTGAAAATACAGGTTTT

CH2-r HindIII2: GGGAAGCTTGGGTTATTAGCTGCCACTACCTTTGGAGATG

NcoI-CH2-f : GGGCCATGGCGGATCTTGGATCC

CH2 ctr-r nat2: GGGAAGCTTGGGTTATTATTTGGAGATGGTTTTCTCGATGGG

NcoI-CH2-f-2: CATGCCATGGGATCCGCGTCATCTGTCTTCAT

All primers were purchased from Eurofins MWG Operon.

2.1.7. Bacterial strains and Plasmids

All constructs for expression in *Escherichia coli* were cloned into pET28-b (Novagen) with the exception of C_H1 camel, which was cloned into pET Sumo (Thermo Fisher).

Strain	Genotype	Origin
E. coli BL21 (DE3)	codon plus F– *ompT hsdS*(rB– mB–) *dcm*+ Tetr *gal endA* Hte (*argU proL* Camr)	Stratagene
E.coli Mach1	ΔrecA1398 endA1 tonA Φ80ΔlacM15 ΔlacX74 hsdR (r_{k-} m_{k+})	Invitrogen
E.coli Shuffle T7	F´ *lac*, *pro*, *lacI*^{Q} / *Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK λatt*::pNEB3-r1-*cDsbC* (Spec^R, *lacI*^{q}) *ΔtrxB rpsL150*(Str^R) *Δgor Δ(malF)3*	NEB

2.1.8. Media for *E. coli*

LB_0	medium:	20 g/l
	plates: Agar Agar	15 g/l
SOC Medium	0.5 % yeast extract (w/v) 2 % tryptone (w/v) 2.5 mM KCl 10mM $MgCl_2$ 10mM $MgSO_4$ 20 mM D-(+)-glucose	
Antibiotics (1000x stocks)	Kanamycin	35-50 µg/l (in ddH_2O)

All media were sterilized using an autoclave at 120°C for 20 min. Antibiotic stocks were passed through a sterile filter (0.22µm) and stored at -20°C.

2.1.9. Chromatography materials and columns

Superdex 75 Prep Grade	GE Healthcare
HisTrap Fast Flow	GE Healthcare
Q Sepharose Fast Flow	GE Healthcare
SP Sepharose Fast Flow	GE Healthcare

2.1.10. Buffers

2.1.10.1 Buffers for protein purification

Inclusion Body buffer	Tris / HCl pH 7.5	50 mM
	EDTA	5 mM
	NaCl	5 mM
Low salt buffer	Tris/HCl pH 7.8-8.0	25 mM
	EDTA	5 mM
	Urea	5 M
High salt buffer	Tris/HCl pH 7.8-8.0	25mM
	EDTA	5 mM
	Urea	5 M
	NaCl	1 M
IB dissolving buffer	Tris/HCl pH 7.8-8.0	25 mM
	EDTA	5 mM
	Urea	8 M
	ß-mercaptoethanol	10 mM

Refolding buffer V_L, C_L, C_H2, C_H1 (camel)	Tris/HCl pH 8.0	250 mM
	EDTA	5 mM
	L-Arginine	100 mM
	GSH	0.5 mM
	GSSG	1 mM
Refolding buffer C_H1	Tris/HCl pH 8.0	250 mM
	EDTA	5 mM
	GSSG	1 mM
Refolding buffer V_H	Tris/HCl pH 8.0	250 mM
	EDTA	5 mM
	L-Arginine	400 mM
	GSH	0.5 mM
	GSSG	1 mM
Phosphate-buffered saline (PBS)	NaCl	137 mM
	KCl	2.7 mM
	$Na_2HPO_4 \cdot H_2O$	10 mM
	KH_2PO_4	2.0 mM
	pH 7.4	
HisTrap FF binding / wash buffer	NaH_2PO_4	50 mM
	NaCl	150 mM
	GdmCl	3 M
	Imidazol	10 mM
	pH 8.0	
HisTrap FF elution buffer	NaH_2PO_4	50 mM
	NaCl	150 mM
	GdmCl	3 M
	Imidazol	300 mM
	pH 8.0	

2.1.10.2 Buffers for SDS Polyacrylamide Gelelectrophoresis

Fairbanks A (Coomassie staining)	2-Propanol	25% (v/v)
	Acetic acid	10 % (v/v)
	Coomassie Blue R	0.05 % (w/v)
Fairbanks D (destaining solution)	Acetic acid 10% (v/v)	
Laemmli sample buffer (5x)	Tris/HCl, pH 6.8	0.3 M
	SDS	10 % (w/v)
	Glycerol	50 % (v/v)
	2-Mercaptoethanol	5 % (v/v)
	Bromophenol blue	0.05 % (w/v)
SDS running buffer (10x)	Tris/HCl, pH 8	0.25 M
	Glycine	2 M
	SDS	1 % (w/v)
Separation gel buffer (2x)	Tris/HCl, pH8.8	1.5 M
	SDS	0.8 % (w/v)
Stacking gel buffer (4x)	Tris/HCl, pH 6.8	0.25 M
	SDS	0.4 % (w/v)

2.1.10.3 Buffers for Molecular Biology

TAE (50x)	Tris/Acetate pH 8.0	2 M
	EDTA	50 mM
Loading buffer (10x)	Glycerol	50 % (v/v)
	Tris/HCl pH 8.0	10 mM
	Orange G	0.2 % (w/v)
	Xylencyanol	0.2 (w/v)

2.1.10.4 Buffers for Biacore X100

HBS-P+ (10 x)	HEPES	0.1 M
	NaCl	1.5 M
	Surfactant P20	1.5 % (v/v)
	pH 7.4	

2.1.10.5 Buffers for ELISA

Reaction Mix I (MAK33)	human creatine kinase (biotinylated) 25 µg pre-dissolved in 2 ml H_20	
	blocking reagent	5 %
Reaction Mix II (1HEZ)	human IgG (biotinylated) 10 µl (1 mg/ml stock) diluted in 4.3 ml PBS	
	blocking reagent	5 %
Reaction Mix III (1HEZ)	human IgG Fc (biotinylated) 10 µl (1 mg/ml stock) diluted in 4.3 ml PBS	
	blocking reagent	5 %
Reaction Mix IV	Na_2HPO_4 pH 7.5	100 mM
	NaCl	50 mM
	EDTA	0.5 mM
	Tween-20	0.1 %
	Blocking Reagent	0.5 g/50 ml
Reaction Mix V	ABTS-buffer	0.835 g/50 ml
	ABTS	1 tablet

Blocking reagent for ELISA, ABTS tablets, ABTS buffer powder, as well as the Streptavidine-coated microwell plates were purchased from Roche Diagnostics.

2.1.10.6 Buffers for AFS formation

pH 2.0, NaCl	NaH_2PO_4	50 mM
	NaCl	100 mM
	pH 2.0 (o-phosphoric acid)	
pH 7.5, NaCl	NaH_2PO_4	50 mM
	NaCl	100 mM
	pH 2.0 (o-phosphoric acid)	

2.2. Software, Databases and Web-based Tools

Software

Adobe Illustrator CS5	Adobe Inc.
BioEdit	Ibis Biosciences
EndNote X6	Thomson Reuters
Epson TT8 Launch Silver Fast	Epson
ImageQuant	GE Healthcare
Microsoft Office 2010	Microsoft
MMass V2.4.0	Open Source
OriginPro 8.6	OriginLab Corp.
Pymol V1.3	Schrödinger
SedFit	Peter Schuck

Databases

PDB www.rcsb.org/pdb
PubMed www.ncbi.nlm.nih.gov/pubmed
UniProt www.uniprot.org

Web-based Tools

Denaturant Calculator	http://mcb.berkeley.edu/labs/krantz/tools/gdmcl.html
Kabat numbering	http://www.bioinf.org.uk/abs/
NCBI Blast	http://blast.ncbi.nlm.nih.gov/
Clustal W sequence alignment	http://www.ebi.ac.uk/Tools/msa/clustalw2/
NEBcutter V 2.0	http://tools.neb.com/NEBcutter2/
OligoAnalyzer 3.1	http://eu.idtdna.com/analyzer/applications/oligoanalyz
ProtParam	http://web.expasy.org/protparam/
Reverse complement	http://www.bioinformatics.org/sms/rev_comp.html

2.3. Molecular Biological Methods

All solutions applied were autoclaved or sterile filtered. If not stated otherwise all experiments were performed at room temperature.

2.3.1. *E. coli* cultivation

E.coli cultivation was performed on LB plates or in liquid medium in an incubator at 37°C and for liquid cultures additional shaking at 200 rpm. Antibiotics were added depending on the resistance gene on the plasmid transformed. Bacterial growth was monitored by measuring the optical density (OD) at 600 nm in cuvettes with 1 mm pathlength ($OD_{600} = 1$ corresponding to approximated 8×10^8 cells) via Vis spectroscopy. For cryo-stocks 800 µl of fresh overnight culture was mixed with 200 µl of 80 % glycerol, shock-frozen in liquid nitrogen and stored at -80°C.

2.3.2. DNA isolation and storage

For isolation of DNA plasmids from 5 ml *E.coli* cultures the Wizard® Plus SV Mini-Prep kit was used following the recommended protocol (Promega). PCR products and restricted fragments were purified using the Wizard® SV Gel and PCR Clean-Up System. The isolated DNA was stored in sterile nuclease free H_2O at -20 °C. For DNA sequencing, plasmid DNA samples with a concentration of 50 to 100 ng/µl, were sent to GATC Biotech AG for sequencing with the T7 forward primer.

2.3.3. Agarose gel electrophoresis

In order to isolate and determine the size of PCR products or digested fragments agarose gel electrophoresis was applied. With this method DNA fragments are separated by size when traversing a 1% agarose gel prepared in TAE buffer. By comparison with the bands of an additionally loaded 1 kb DNA ladder (Peqlab) the DNA size of the samples was determined. Gel runs were performed at a voltage of 120 mV, the addition of the DNA stain, Stain G (Serva Electrophoresis) to the gel mixture enabled a DNA detection under UV light.

2.3.4. Polymerase chain reaction (PCR)

PCR DNA amplification was used for quick change mutagenesis of existing DNA plasmids or selective DNA amplification for subsequent cloning. Primers with a melting temperature of 56- 60 °C were ordered from Eurofins MWG and controlled with OligoAnalyzer.

The PCR itself consists of three steps, denaturation of the double strand, annealing of the primers and elongation of the DNA strand by the added polymerase. Two polymerases with different elongations velocities (Pfu Polymerase with approximately two minutes/kb and the Phusion polymerase with 15-30 seconds/kb) were used. The following shows the applied cycler program pipetting scheme for the PCR reaction:

Cycler program:

Step 1: Denaturation 1 min; 95°C
Step 2: Repetition 30 x
    Denaturation: 30 sec; 95 °C
    Annealing: 30 sec; 53- 59 °C
    Amplification: 0.5 – 5 min; 72°C
Step 3: 10 min; 72° C

Protocol for a 50 µL reaction with Pfu/ Phusion Polymerase:
1 µL template DNA (50 ng/µL)
1 µL primer_f (10 pmol/µL)
1 µL primer_r (10 pmol/µL)
1 µL dNTPs (10 mM per dNTP)
5 µL 10x Pfu-Polymerase Puffer
1 µL Pfu-Polymerase
40 µL nuclease free ddH_2O

1 µL template (50 ng/µL)
2.5 µL primer_f (10 pmol/µL)
2.5 µL primer_r (10 pmol/µL)
1 µL dNTPs (10 mM per dNTP)
10 µL 5x Phusion HF Buffer
1.5 µL DMSO (3%)
0.5 µL Phusion HF DNA-Polymerase
31 µL Nuclease free ddH_2O

2.3.5. Cloning strategies

2.3.5.1. Quick Change PCR

In order to insert single point mutations in existing plasmids (usually pET28-b) forward and reverse primers carrying the target mutation were designed. The wild type plasmid served as the template DNA. To remove the template DNA after a successful PCR, DpnI (NEB) was added for 1h at 37°C, followed by a heat inactivation step for 20 minutes at 65°C. Then the PCR product was transformed into supercompetent Mach 1 *E. coli* cells (heat dependent transformation). Cells were plated onto LB plates containing the corresponding antibiotics and incubated overnight at 37°C. Single colonies were picked and incubated overnight at 37 °C in liquid LB with antibiotics. The day after, the plasmid DNA was isolated and sequenced.

2.3.5.2. Standard PCR

In case more than only 3 bases had to be added or deleted from the wild type plasmid a standard PCR was performed for subsequent cloning. Primers containing the desired 5`or 3`sequences were designed and a PCR was performed. The correct size of the PCR product was checked by agarose gel electrophoresis. After DNA purification insert and vector were digested with the appropriate restriction enzyme for 2 hours at 37 °C. Again

insert and vector were purified and ligated (1:2 ratio) in a total volume of 20 µl by adding the T4 DNA ligase from Promega together with the corresponding T4 ligase buffer. The ligation reaction was incubated overnight at 8°C in the fridge and then transformed into supercompetent Mach 1 *E. coli* cells.

2.4. Protein chemical methods

2.4.1. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS- PAGE was conducted according to the protocols of Laemmli [200]. The applied polyacrylamide gels had final concentrations of 5% (w/v) acrylamide/bisacrylamide 19:1 (40% w/v) for stacking and 15-18% (w/v) acrylamide/bisacrylamide 19:1 (40% w/v) for separation gels. Polymerization of the gel solution was induced by the addition of ammonium persulfate (APS, 10% w/v in H_2O) and tetramethylethylendiamin (TEMED). For sample preparation all samples were mixed with 5x Laemmli (+5% β-mercaptoethanol) sample buffer and incubated at 95°C for 5 min before loading on the gel. Gel electrophoresis was performed at a constant conductant currant of 35 mA/gel for a running time of 45-60 minutes. The gel staining was carried out with Coomassie according to a modified protocol of Fairbanks and co-workers [201].

2.4.2. Protein Expression and Purification

Purification of wild type C_L, C_H1, and V_L was performed as published by Feige and Simpson [119,191, 202], the purification of the newly generated different C_H1 mutants was carried out accordingly. Wild type V_H was purified referring to Feige et al. [192] with the exception of the additional use of a SP-sepharose for some mutants. The purification of the C_H2 constructs was performed as mentioned above with some slight deviations concerning the composition of the refolding buffer (see 2.1.10.1 buffers for purification).

2.4.2.1 Protein expression, harvest and cell disruption

All proteins were expressed in 5 L flasks with 2l LB medium containing antibiotic. Inoculation occurred with a 1:60 dilution of overnight culture and incubated at 37°C. At an OD_{600}= 0.6-0.8 protein purification was induced by addition of 1 mM IPTG (final concentration) and incubated overnight at 37°C. Cell harvest was carried out by centrifugation (7000 rpm, 10 min, JA 10). The cells were resuspended in IB buffer which

contained Protease inhibitor Mix G (or HP for His-Tag labelled proteins) and DNase I was added before cell disruption.

2.4.2.2. Preparation of inclusion bodies for purification

After cell disruption the lysate was incubated with 1% Triton X-100 at 4°C for at 2 hours to solubilize membrane proteins. Inclusion bodies were separated from the soluble parts of the lysate by centrifugation (20.000 rpm, 20 min, JA 25.50), washed twice with IB buffer and stored at – 20 °C. For purification inclusion bodies were dissolved for 2 hours in IB dissolving buffer with 10 mM β-mercaptoethanol at room temperature. After centrifugation (20.000 rpm, 25 min, JA 25.50) the supernatant with the solubilized protein was used for purification.

2.4.2.3. Affinity chromatography

The first step of purification for His_6-taged antibody domains like C_L and C_H1 (camel) was a 5 ml HisTrapFastFlow Ni-affinity (GE Healthcare) chromatography column. Here a chelate complex is formed between the matrix bound Ni ions and the His_6-tages proteins. The binding is reversible and can be disrupted by competitive displacement, which is achieved with an increasing imidazole concentration in the elution buffer. The protein was loaded with a flow rate of 2 ml/min and washed with 10 column volumes of HisTrap FF binding / wash buffer. For elution a gradient of 15 column volumes from 0 to 100 % of HisTrap FF elution buffer was applied and the fractions containing the target protein were pooled, the His_6-tag removed by enzyme induced cleavage over night at 4°C in HisTrap FF binding / wash buffer. The applied enzyme (thrombin, TEV protease, Ulp1) varied depending on the cleavage site attached to the target protein. After His_6-tag removal the protein was again loaded on the column to separate the cleaved tag from the protein since the protein without tag was then in the flow through. Ni- affinity chromatography was followed by protein refolding and size exclusion chromatography.

2.4.2.4. Ion exchange chromatography

For antibody domains without any tag the first step of purification was an ion exchange chromatography column. The principle of binding to this column material is reversible absorption of oppositely charged molecules to the ion exchange groups immobilized on the matrix. The net charge of a protein depends on its isoelectric point (pI) as well as the

pH of the buffer. In this work, mainly the anion exchange Q-sepharose FF column (GE Healthcare, Freiburg, Germany) was used with a flow rate of 4 ml/min. For some V_H mutants an additional SP-sepharose FF column (GE Healthcare) was utilized. After loading of the protein the column was washed with 3 column volumes of low salt buffer, followed by the elution of the protein via competitive displacement with high salt buffer containing 1 M NaCl. Ion exchange chromatography was followed by protein refolding and size exclusion chromatography.

2.4.2.5. Protein refolding

Refolding was performed in the refolding buffers as specified in 2.1.10.1 via overnight buffer exchange in dialysis tubes at 4 °C with an at least 20-fold excess of refolding buffer. For all C_H1 constructs concentrations < 100 µg/ml were applied.

2.4.2.6. Size exclusion chromatography

After refolding, proteins were concentrated to 10 mL and size exclusion chromatography with High Load Superdex 26/69 75 pg column (GE Healthcare) was performed. Here, the protein is not interacting with the column but has to traverse a matrix composite of dextran and agarose with different pore sizes, which leads to a separation by size. Protein purity was monitored by SDS-PAGE and the correct mass of the protein was confirmed by Matrix-assisted laser desorption ionization time-of-flight (MALDI-ToF) mass spectrometry. For storage the protein was shock-frozen with liquid nitrogen and stored at – 80 °C.

2.4.3. Protein labelling

Protein constructs with an additional N-terminal cysteine residue designed to be used for analytical ultracentrifugation interaction studies with fluorescence detection had to be labelled using maleimide Atto 488 (Atto-tech). For labelling the protein had to be incubated for 20 min at RT with 1 mM TCEP in order to reduce the free cysteine, followed by an 2 hour incubation with 2-3 fold excess of the fluorescence dye (dissolved in DMSO). After label quenching by adding 10 µl of a prepared cysteine solution (10 mM in H_20) free label was separated by HPLC using a Superdex 75 FPLC (Pharmacia Biotech). Labelling efficiency was determined using UV spectroscopy following the manufacturer's protocol.

2.5. Spectroscopy

2.5.1. UV-Vis spectroscopy

Protein concentrations can be determined by UV-Vis spectroscopy measuring the absorption at 280 nm. This is based on the absorbance behavior of chromophores with a system of conjugated double bonds as in aromatic amino acids (tryptophan, tyrosine, phenylalanine) between a wavelength of 230 – 280 nm. The exact numbers are the following:

Trp: λ_{max}(abs) = 280 nm, ε_{max}(abs) = 5600 $M^{-1}cm^{-1}$
Tyr: λ_{max}(abs) = 275 nm, ε_{max}(abs) = 1400 $M^{-1}cm^{-1}$
Phe: λ_{max}(abs) = 258 nm, ε_{max}(abs) = 200 $M^{-1}cm^{-1}$

So at a wavelength of 280 nm mainly tryptophan and tyrosine are responsible for the absorption. When applying the Beert-Lambert law, the protein concentration can be calculated from the determined absorption:

$$A = \varepsilon \, x \, c \, x \, d$$

Equation 1: Beer-Lambert law.

A = Absorbance, ε = molar extinction coefficient (M^{-1} cm^{-1}), c = molar protein concentration (M), d = path length (cm)

For the determination of the theoretical molar extinction coefficients the ProtParam tool was used. All UV spectra were recorded with a Helios γ (Thermo Fischer) or Ultrospec 1100 (Amersham Biosciences) UV/Vis spectrophotometer and baseline-corrected for buffer absorbance.

2.5.2. Circular Dichroism (CD) Spectroscopy

CD-spectroscopy measures the optical activity of chiral molecules in terms of absorption of left and right circularly polarized light. Concerning proteins, the left- and right-circularly polarized light in the far-UV region (260 – 170 nm) is absorbed by the periodically repeated binding angles of the secondary structure elements to a different extent. Secondary structure elements lead to a characteristic shape of the spectra, an α-helical structure has two minima at 222 nm and 208 nm, β-strands one minimum at 217 nm and random coil structures show a minimum at 200 nm. Altogether the CD-spectrum of a

native protein represents the sum of the signals for the different secondary structure elements, turns and loops. In the near-UV region (320 – 260 nm), though, the aromatic amino acids and to a small amount disulfide bonds generate the obtained CD signal. But it is the asymmetric environment of these aromatic amino acids that defines an individual, unpredictable fingerprint CD spectrum for each protein. So with the near-UV range the tertiary structure of a protein can be monitored but with less common characteristics than far-UV spectra. These spectra are a sensitive characteristic for the native state of a protein and mainly serve as an indicator of tertiary structure changes in comparison to a wildtype protein. The CD signal is measured as ellipticity (Θ) in degrees (Kelly et al, 2005).

$$\Delta A(\lambda) = A_L(\lambda) - A_R(\lambda) = [\varepsilon_L(\lambda) - \varepsilon_R(\lambda)]lc = \Delta\varepsilon lc$$

Equation 2: Absorption between left and right handed circularly polarized light.
A= absorption, λ= wavelength, R and L subscripts describe left- and right-handed circular polarized.

The measured ellipticity is further transformed to molar ellipticity θ_{MRW} which depicts the mean residue ellipticity, more precisely the ellipticity in relation to the average molecular weight of the amino acids of the protein, which is calculated as described in equation 3:

$$\theta_{MRW} = \frac{\theta \; x \; 100 \; xM}{d \; x \; c \; N_{AA}}$$

Equation 3: Determination of the mean residue ellipticity.
Θ_{MRW} = mean residue ellipticity (deg cm^2 $dmol^{-1}$), Θ= ellipticity (deg), M= molecular mass (g/mol), d= pathlength (cm), c = concentration (M), N_{aa}= number of amino acid residues.

In this work, far-UV and near-UV spectra were recorded to monitor the secondary and tertiary structure of proteins. To obtain the thermal stability of a protein by CD spectroscopy the sample was measured at a defined wavelength in the appropriate secondary structure far-UV region (210 – 215 nm), while heating up with 20 °C/h. With the obtained changes in CD signal a Boltzman fit was conducted to get a midpoint of the transition. The CD was also used to follow the real time folding of C_H1 by the C_L domain. The kinetics were conducted as described in Feige et al. 2009 at a constant wavelength of 205 nm, the midpoint was determined by exponential decay fit.

Parameters far-UV/near- UV:
Wavelength (nm): 260 - 195/ 320 - 250
Scan speed (nm/min): 20
Data pitch (nm): 0.2
Band width (nm): 1.0
Response (sec): 4
Temperature (°C): 20
Accumulations: 15
Concentration (µM): 10 / 50
Quartz cuvette pathlength (mm): 0.5 / 5

Parameters kinetics:
Wavelength (nm): 205
Data pitch (sec): 5
Band width (nm): 1.0
Response (sec): 4
Temperature (°C): 25
Accumulations: 1
Time (min): 180
Concentration (µM): 10 each
Quartz cuvette pathlength (mm): 1

Parameters temperature induced unfolding transition
Wavelength (nm): 210 - 215
Data pitch (nm): 0.2
Band width (nm): 1.0
Response (sec): 4
Temperature (°C): 20
Heating rate (°C/h): 20
Temperature (°C): 10 -95
Concentration (µM): 20
Quartz cuvette pathlength (mm): 0.5

2.5.3. Fluorescence spectroscopy

Excitation of an aromatic system generates fluorescence. It is caused when an electron shifts back from its excited state (S_1) to the ground state (S_0). A molecule in the S_1 returns to S_0 by different mechanisms, either driven by internal conversion, conversion to the triplet state (intersystem crossing) or emission of a photon (fluorescence). Within proteins, side chains of the amino acids phenylalanine (Phe), tyrosine (Tyr) and

tryptophan (Trp) exhibit fluorescence upon excitation. These amino acids show the following fluorescence properties at neutral pH:

Trp: λ_{max}(abs) = 280 nm, λ_{max}(fluor) = 353 nm

Tyr: λ_{max}(abs) = 275 nm, λ_{max}(fluor) = 304 nm

Phe: λ_{max}(abs) = 258 nm, λ_{max}(fluor) = 282 nm

Usually the fluorescence signal is dominated by the Trp side chains due to a stronger intensity of the emission compared to Phe and Tyr (205). Since the fluorescence emission and signal intensity depends strongly on the environment of the amino acids, instrinsic Trp fluorescence is suitable to monitor conformational changes of a protein. This principle is well established for antibody domains, containing a conserved Trp located in the inner, hydrophobic core, next to a conserved disulfide bridge which exerts a quenching effect on the Trp. Based on this, changes of the protein structure and loss of disulfide bridge linkage lead to pronounced changes in Trp fluorescence. In case of denaturant (GdmCl or Urea) induced protein unfolding by addition of increasing denaturant concentrations the Trp becomes solvent exposed. This displacement from the native, hydrophobic to a solvent exposed, polar environment leads to a change in Trp fluorescence visible by a shift of the maximum emission to a larger, lower energy wavelength (red shift) and an increase in signal intensity [160].

An additional application next to the determination of the intrinsic tryptophan fluorescence is the extrinsic fluorescence. Here, the hydrophobicity of a protein can be assigned by observing the fluorescence of the fluorescent dye 1-anilinonaphthalene-8-sulfonic acid (ANS, Sigma-Aldrich) which shows a strong binding to hydrophobic patches. Upon binding to exposed hydrophobic regions of a protein, ANS fluorescence increases. Hence, structural changes of proteins, from the native state with a mainly polar surface to the unfolded state with exposure of hydrophobic residues can be monitored [161].

Fluorescence measurements were carried out in a Spex FluoroMax IV fluorimeter (Horiba Jobin-Yvon). All experiments were performed at 20°C with a protein concentration of 1 µM in a 1-cm quartz cuvette. The excitation and emission slit widths were set to 2 nm and 3, respectively. Intrinsic tryptophan fluorescence spectra were recorded from 300 - 450 nm, the excitation wavelength was 280 nm and. For ANS measurements 50 -100 µM ANS

was added to the protein, the excitation wavelength was 380 nm and the spectra were recorded between 400 - 650 nm.

For GdmCl- or Urea- induced unfolding transitions a 8 M stock solution in PBS (pH 7.4 – 7.6) was prepared and the exact concentration determined refractrometrically. The samples were incubated over night at 20 °C in the corresponding increasing denaturant concentrations. Basic parameters were set as above and the emission signal was measured for 50 seconds at 355 nm, which depicts the wavelength with the largest signal difference between native and unfolded state due to the spectra red shift. Spectra were accumulated three times and buffer corrected, the obtained unfolding transitions were fitted according to a two-state unfolding model for dimeric proteins [162]:

$$y\,(D) = y_N^0 + m_N\,x[D] - \frac{(y_N^0 + m_N\,x[D]) - (y_U^0 + m_U\,x\,T)}{1 + \exp(-\frac{\Delta G_{stab} + m_c\,x\,[D]}{R\,x\,T})}$$

Equation 4: Two-state model with the assumption of linear dependency of emission from native and unfolded protein

y(D) = fluorescence or CD signal; y = y-intercepts; N = data of native protein; U = data of unfolded protein; m = slope of the best fit straight line; m_c = cooperativity; [D] = denaturant concentration; T = temperature in Kelvin (20 °C (RT~) 293.1 K); ΔG_{stab} = free enthalpy of denaturant induced unfolding; R = universal gas constant (8.314 J mol^{-1} K^{-1})

2.6. Quaternary structure analysis

2.6.1. Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) allows the real-time observation of the sedimentation of macromolecules in a centrifugal field. With the obtained sedimentation profile a hydrodynamic and thermodynamic characterization of a protein in solution, without disturbing interactions with a matrix or surface, can be performed. The sedimentation process itself is determined by three factors – the gravitational force, the buoyancy and the hydrodynamic friction. From the balance of these three forces, the Svedberg equation is derived [163]:

$$s = \frac{v}{\omega^2 r} = \frac{MD(1 - \bar{V}\rho)}{RT}$$

Equation 5: Svedberg equation
s = sedimentation coefficient (s), *v* = observed radial velocity (m/s), *ω* = angular velocity of the rotor (m/s^2), $\omega^2 r$ = centrifugal field, *M* = molar mass (g/mol), $\bar{V}$ = partial specific volume (cm^3/g), *ρ* = density of the solvent (g/cm^3), *D* = diffusion coefficient (m^2/s), *R* = gas constant (8.314472 J/K mol), *T* = absolute temperature (K)

The sedimentation coefficient s depicts the sedimentation velocity *v* in relation to the centrifugal field $\omega^2 r$. It is represented by Svedberg (S) units, which corresponds to 10^{-13} s. Differences in s values can be assigned to differences in the size or shape of a molecule, but usually higher s-values correspond to bigger molecules that sediment faster than small ones. AUC offers a wide range of applications depending on the optical system used for detection (absorbance, fluorescence or interference)[163]. In sedimentation velocity (SV) experiments, the sedimentation profile showing the movement of the molecule within a centrifugal force field towards the bottom of the centrifuge cell, can be used for a size distribution. Since some isolated antibody domains can form dimers, AUC represents a useful method to determine theses dimers as well as the formation of higher oligomers or aggregates. An additional application of SV runs is the measurement of protein interactions by the use of the fluorescence optics in combination with one interaction partner containing a fluorescence label. In order to determine a K_D a sample series with a constant protein concentration of the labeled protein and increasing concentrations of the unlabeled interaction partner is measured. Since only the labeled protein can be detected a shift towards higher s-values or the appearance of a second peak with higher s-value corresponds to the formation of a complex.

Another AUC approach for interaction studies are sedimentation equilibrium (SE) experiments. Here, the absorbance is monitored at three different speed settings, with a sample series of increasing concentrations with 1:1 ratio for both interaction partners. The measurements are performed after an 24 hour incubation step at each adjusted speed so that an equilibrium state between diffusion and sedimentation is reached. For an ideal non-interacting single component system, the equilibrium distribution is an exponential function of the buoyant mass of the molecule, which is described in the following equation:

$$C(r) = c_0 \, x \, e^{\frac{M(1-\overline{V}\rho)\omega^2(r^2-r_0^2)}{2RT}}$$

Equation 6: Equilibrium within an ideal non-interacting single component system.
c(r)= sample concentration at radial position r, c_0= sample concentration at reference radial distance r_0,V= partial specific volume, ω= angular velocity of the rotor, R= gas constant, T= absolute temperature

AUC was performed with a Beckman ProteomLab XL-A centrifuge (Beckman-Coulter). For SV experiments with absorbance optics, 400 µl of the protein samples and 450 µl of PBS were loaded into assembled cells with quartz windows of 12-mm pathlength and centrifuged at a speed of 42.000 rpm in an eight-hole Beckman AN50-Ti rotor. Sedimentation was monitored at 230 nm with a protein concentration related to an OD of 0.2 -0.8 for absorbance optics. In the case of fluorescence detection no buffer filled reference cell was necessary, detection was set to 490 nm with a signal gain of 50- 59 %. Data analysis was carried out with the program Sedfit using a non-model based continuous Svedberg distribution method (c(S)), with time (TI) and radial (RI) invariant noise on [164]. The obtained data was corrected to standard solvent conditions, density and viscosity of water at 20° C. The partial specific volume for the antibody domains was taken from literature [165]. For SE experiments 120 µl samples and 125 µl buffer were loaded into six channel epon centerpieces (12 mm window) and the equilibrium state was measured at three different velocities (25.000, 30.000 and 34.000 rpm). Single Scans were recorded at 280 nm, 250 nm and 230 nm, the best scan for each velocity with a good signal to noise ratio and an absorption range between 0.1 - 0.8 was chosen for further analysis. Data was globally fitted with the program Origin 8.6 to obtain the association constant with a heterogeneous interaction model (1:1 complex) based on the mass action law $C_{AB} = K_a c_A c_B$ (Equation 7). For the fit, scans of three different velocities and concentrations were implemented.

$$a(r) = C_A(r_0)\,\varepsilon_A\, d \exp\left[M_{b.A}\frac{\omega^2}{2RT}(r^2-r_0^2)\right] + C_B(r_0)\varepsilon_B\, d\, \exp\left[M_{b.B}\frac{\omega^2}{2RT}(r^2-r_0^2)\right] + K_a\, C_A(r_0)\, C_B(r_0)\,(\varepsilon_A + \varepsilon_B) d\, \exp\left[(M_{b.A} + M_{b.B})\frac{\omega^2}{2RT}(r^2-r_0^2)\right]$$

Equation 7: Model for heterogeneous interaction between two components.
c(A)/c(B)= sample concentration of interaction partner A and partner B, r_0= reference radius, $M_{b,A}$/$M_{b,B}$= buoyant masses of component A and B, ω= angular velocity of the rotor, R= gas constant, T= absolute temperature, d= optical path length, K_a= association constant

The K_D values can be calculated from the K_A values as described in equation 8.

$$K_D = \frac{1}{K_A}$$

Equation 8: K_D determination.
K_D = dissociation constant, K_A = association constant

2.6.2. Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) spectroscopy is a method for measuring real time protein interactions in a label free environment. For the generation of an SPR signal, a gold layer attached to a glass surface is irradiated with polarized light leading to an internal reflection at the interface between media of different refractive index. The reflected photons generate an electric field leading to the creation of plasmons within the gold layer extending to the medium as an evanescent wave on the opposite site of the interface. The velocity of the plasmons depends strongly on the medium composition, provoking changes in the angle of the incident light. The angle of the reflected light is detected and further converted to response units (RU) by the Biacore X100 data acquisition software. Within the experiment setup a protein (ligand) is immobilized via different coupling chemistry approaches on the glass/gold surface of a chip while the interaction partner (analyte) is passing under constant flow. When a protein is binding to the surface a change of the refractive index occurs leading to an angle shift of the reflected light. This change is linear to the amount of bound analyte.

Within this work a Biacore X100 (GE Helathcare, Uppsala, Sweden) was used with CM5 sensor chips following the manufacturers protocols and recommendations for maleimide immobilization and interaction experiment settings. Multi cycle kinetic measurements were performed with 10 µl/min at 15 – 20 °C, with HBSP+ buffer as running buffer and PBS as sample buffer. Regeneration was executed with 1.5 M NaCl, final RUs for immobilization varied (500 RU for V_L, 3000 RU for V_H), the analyte concentration range was individually adjusted depending on the K_D. Data was evaluated with the Biacore X100 evaluation software.

2.6.3. Enzyme-linked Immunosorbent Assay (ELISA)

Enzyme-linked Immunosorbent Assay (ELISA) was applied to assess the functionality of the isolated mutated V_H and V_L domains in terms of antigen binding. Within this work the ELISA established by Eva Herold in an 96- well format was used (see Figure 8 for the ELISA setup). For the ELISA the antigen is immobilized on a solid surface, the antibody or in this case the variable domains are added. If a complex is formed can be detected by addition of an enzyme- conjugated antibody, which specifically binds to the previous antibody. After some incubation with the substrate the enzyme activity is determined photometrical and conclusions can be drawn concerning the antigen/ antibody binding affinity. Microwell streptavidin-coated plates were applied to immobilize the appropriate biotinylated antigen, creatine kinase for MAK33 or F_c part of IgG for 1HEZ. The assay was performed according to the manufacturer´s recommendations. 10 µl a 1:1 V_H/V_L mixture (with the wild type domain containing an additional C-terminal FLAG-tag) was added to the biotinylated antigen solution (90µL) and incubated in the streptavidin wells for 45 min at 4 °C and 300 rpm shaking. After a washing step with water, the horseradish peroxidase-conjugated detection antibody or Protein L was added (1:15000 dilution) and incubated as mentioned before. The second washing step was followed by the addition of the substrate (synthetic peroxidase). Enzymatic activity leads to an instant visible color reaction of the substrate that can be monitored with a Tecan genios plate reader at an absorbance of 405 nm.

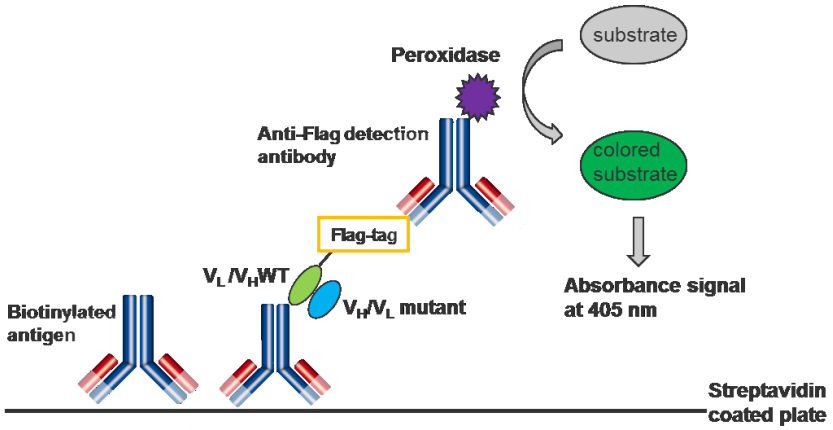


Figure 8 : ELISA set-up to test binding affinities of isolated domains.
The biotinylated antigen, human creatine kinase for MAK33 or the Fc part of an IgG antibody for 1HEZ, is coupled to a streptavidin coated 96 well plate; The wild type domain has a c-terminal FLAG-tag and is incubated with the corresponding mutant domain as interaction partner; For detection a horseradish peroxidase (HRP) coupled Anti-FLAG-tag antibody is added which leads to a color reaction of the substrate

3. Results and Discussion

3.1. Interaction of the variable domains

Within this study the role of the highly conserved amino acids in the V_H/V_L interface, identified by Wang and coworkers, is examined (see 1.7.2). Hence, the following alanine exchange point mutations of the conserved residues were generated and characterized within the context of a human IgM antibody (pdb ID: 1HEZ): R38A, G44A, L45A, E46A, W47A, W103A for V_H and Y36A, Q37A, P44A, L46A and F98A for V_L (Figure 9).

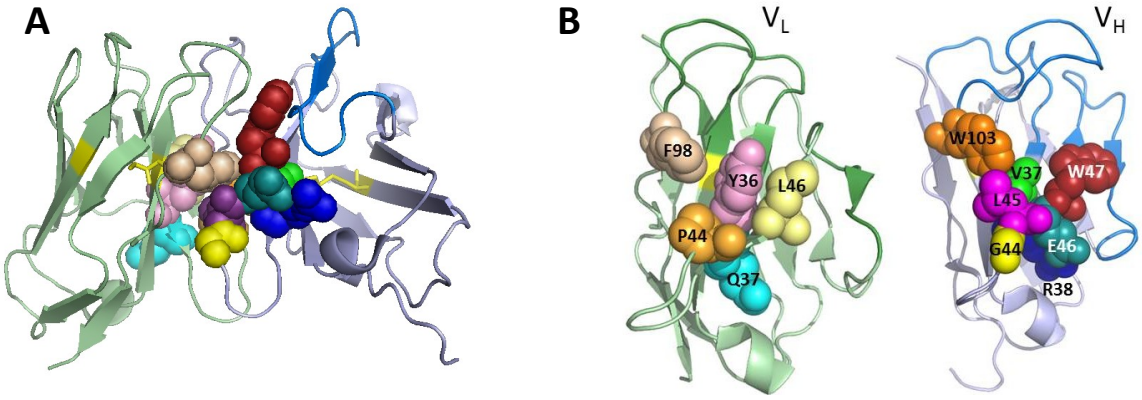


Figure 9 : Conserved residues within the V_H/V_L interface.
(A) The pymol cartoon representation of the 1HEZ Fv fragment shows the positions of the conserved interface amino acids. The V_L domain is depicted in light green with the three CDRs highlighted in dark green. V_H is shown in light blue and the CDRs are in dark blue. Conserved residues within the interface are illustrated as spheres and color-coded. The labeled residues were selected for an alanine-exchange mutational analysis. In (B) the top views of the interacting residues of the V_L (left) and V_H (right) domain of 1HEZ are shown. Six residues were selected for V_L (Y36, Q37, P44, L46, and F98) and seven for V_H (V37, R38, G44, L45, E46, W47, and W103).

This study is based on the assumption that these highly conserved amino acids must be of importance for the functionality of the variable domains. They might be crucial for structure formation, stability, V_H/V_L interaction or antigen binding. The crystal structure of the 1HEZ variable domains is shown in Figure 9 highlighting the corresponding residues included in the mutation study.

This work is continues the work of Dr. Eva-Maria Herold, where she performed a comparable analysis with a different antibody (pdb ID: 1FH5, κ light chain/IgG1) from mouse [166]. A sequence alignment of both Fv fragments is shown in Figure 10. The additional data of the 1HEZ antibody was generated in order to determine whether the observed effects can be transferred between different antibody subtypes of different

origin. Taken the two data sets together, it might be possible to draw a general conclusion about the influence of the conserved amino acids. The special approach of this analysis, in contrast to most other Fv studies, is the use of isolated variable domains for the characterization of mutants instead of Fab or scFv formats.

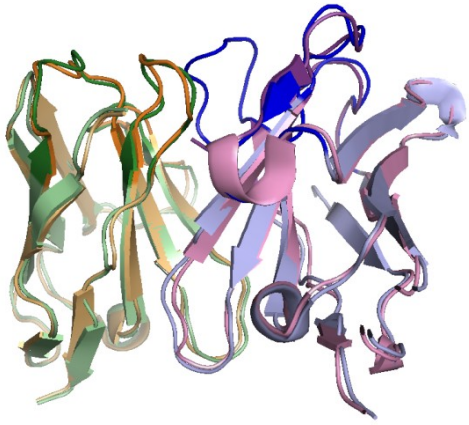


Figure 10 : Structure and sequence alignment of the 1HEZ Fv fragment with MAK33.

The 1HEZ V_L domain is depicted in light green with the three CDRs highlighted in dark green, the MAK33 (pdb ID 1FH5) V_L domain is represented in light orange with the CDRs highlighted in dark orange. 1HEZ V_H is shown in light blue and the CDRs in dark blue, MAK33 V_H is illustrated in light pink and the CDRs in dark pink. CDR H3 is missing for MAK33 as it was not included in the crystal structure. The Sequence alignment was performed with CLUSTALW, the structural alignment with pymol.

```
1HEZ VL: DIQMTQSPSSLSASVGDRVTITCRTSQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFS
MAK33 VL:DIVLTQSPATLSVTPGDSVSLSCRASQSISNNLHWYQQKSHESPRLLIKYASQSISGIPSRFS
         ** :****::**.: ** *:::**:*****. *:*****. ::*:***  **.  **:*****
GSGSGTDFTLTISSLQPEDFATYYCQQSYSTPRTFGQGTKVEIK-
GSGSGTDFTLSINSVETEDFGMYFCQQSNSWPLTFGAGTKLELKR
**********:*.*::.***. *:**** * * *** ***:*:*

1HEZ VH:  AQVQLVESGGGVVQPGRSLRLSCAASGFTFSGYGMHWVRQAPGKGLEWVALISYDESNKYYAD
MAK33 VH:  EVQGVESGGGLVKPGGSLKLSCAASGFTFSDYYMYWVRQTPEKRLEWVATISDGGSYTYYPD
           :** ******:*:** **:***********.* *:****:* * ***** ** . * .**.*

SVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKVKFYDP--TAPNDYWGQGTLVTVSS-
SVKGRFTISRDNAKNNLYLQMSSLKSEDTAMYYCARDKAYYGNYGDAMDYWGQGTSVTVSSA
************:**.*****.**::****:****: * *      . ******* *****
```

3.1.1. Structural effects of the variable domain point mutations

Before the actual measurements, some V_H point mutations already showed deviations from the wild type domain during the purification process. A very high aggregation propensity was observed for E46A, W47A and especially for R38A. Additionally, during gelfiltration chromatography, which separates proteins by size, the V_H mutants R38A, E46A and W47A, eluted with two peaks at 280 nm next to the aggregation peak. Both peaks were confirmed to be the corresponding to the V_H mutant by SDS PAGE and mass spectrometry. Size determination by AUC analysis, performed with a protein concentration of 5 - 10 µM monitored a 100 % dimer distribution for V_H wild type and all point mutations except for W47A (see Table 1). In this case, an equilibrium between monomer and dimer was observed for the pooled protein fractions of both gelfiltration

peaks of V_H mutant W47A. So surprisingly, no differences between the two elution peaks concerning V_H R38A, W47A and E46A were detected. But it cannot be excluded that the size distribution might shift with increasing or decreasing protein concentrations. Altogether the purification indicates structural changes for the point mutations R38A, E46A and W47A compared to the wild type. Concerning the V_L point mutations no noticeable problems during purification occurred, AUC showed a 100 % monomer for all mutants (Table 1).

Table 1 : Size distribution of the V_H and V_L point mutations

Domain	Mutation	% Monomer	% Dimer	Domain	Mutation	% Monomer	% Dimer
V_L	wild type	100	-	**V_H**	wild type	-	100
	Y36A	100	-		V37A	-	100
	Q37A	100	-		R38A-1	-	100
	P44A	100	-		R38A-2	-	100
	L46A	100	-		G44A	-	100
	F98A	100	-		L45A	-	100
					E46A -1	-	100
					E46A -2	-	100
					W47A-1	50	50
					W47A-2	50	50
					W103A	-	100

The size of the V_L and V_H point mutations was determined by AUC sedimentation velocity runs at a protein concentration of 5 - 10 µM with absorbance optics at 230 nm. Data evalution was perfomed with the program Sedfit (Peter Schuck) using a c(S) model.

The far-UV CD spectra of both, V_H and V_L point mutations indicate a ß-sheet structure with a minimum at 218 nm and a maximum at 200 nm (Table 4, Figure 11 A and B) [167]. For V_L, no differences of the point mutations compared to the wild type were visible. Concerning V_H, though, two point mutations, E46A- peak 1 and R38A, showed deviations from the wild type spectrum with higher signal intensity and a generally broader minimum range shifted towards ≈ 210 nm. In the case of E46A, for the two gelfiltration peaks different far-UV CD spectra were observed. While peak 2 exhibited a wild type like spectrum, peak 1 seemed to be misfolded, the same holds true for R38A which overlapped with the E46A-peak 1 spectrum. For R38A and W47A only one spectrum was shown since both peak spectra possessed the same shape. To sum up, no changes in secondary structure could be monitored for the V_L point mutations, while V_H point

mutations E46A-peak 1 and R38A cannot adopt a wild type like secondary structure. For further characterization R38A and E46A –peak 1 are excluded in this study due to their misfolded nature.

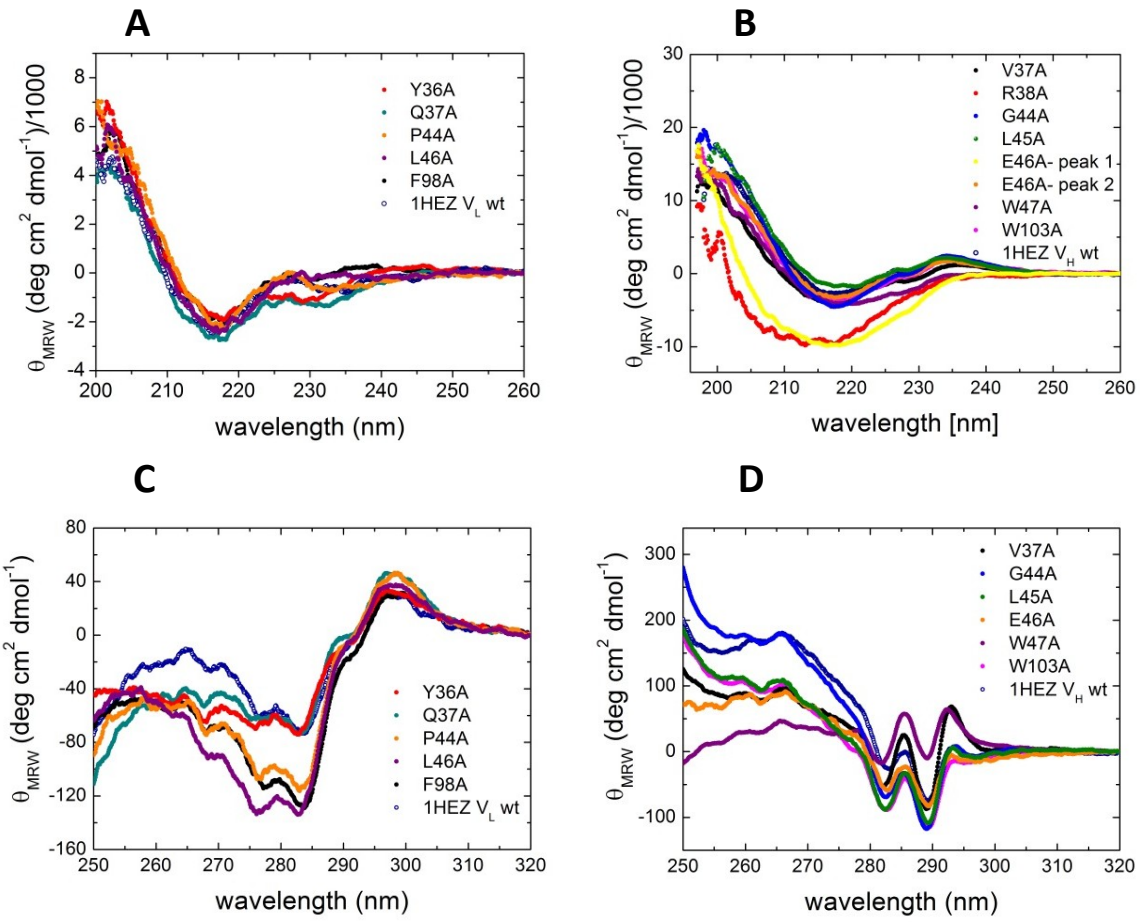


Figure 11 : Secondary and tertiary structure of the V_L and V_H point mutations.
In (A) and (B), far-UV-CD spectra of V_L mutants (A) and V_H mutants (B) are shown. For the spectra 16 accumulations each were recorded and buffer-corrected (PBS). All measurements were performed at a protein concentration of 20 µM (FUV-CD) and 50 µM (NUV-CD) in 0.5 mm (FUV) or 5 mm (NUV) quartz cuvette at 20°C.

The near-UV CD spectra can be applied to identify structual rearrangements around the environment of the aromatic amino acids of a protein. Regarding V_L, all spectra possessed maxima around 300 nm and 265 nm as well as two minima at 285 nm and 275 nm (Figure 11 C). Only the signal amplitude was subject of variations, while the wild type overlapped almost completely with Y36A and Q37A, the point mutations P44A, L46A and F98A had more pronounced minima. The V_H domain exhibited two minima around 290 nm and 280 nm as well as a maximum around 265 nm (Figure 11 D). As already observed for V_L, the overall spectra shape was the same with slight variations in the signal

amplitude, especially around 265 nm. The most striking difference was observed for W47A, with two maxima at 295 nm and 285 nm. Since the second tryptphan mutant (W103A) showed wild type like spectra, this effect might rather be related to the different oligomeric state of the W47A and the associated structural rearrangments than the lack of a CD signal contributing tryptophan residue. Altogether, for both variable domain point mutations only slight changes in the tertiary structure can be monitored.

3.1.2. Stability effects of the variable domain point mutations

For stability assessment thermal unfolding and GdmCl induced unfolding was performed (Table 2 ; Figure 12).

Table 2 : Stabilities of the V_L and V_H point mutations against thermal and chemical (GdmCl) denaturation.

Domain	Mutation	T_{melt} [°C]	Stability GdmCl $D_{1/2}$ [M]	Cooperativity [kJ mol^{-1} M^{-1}]
V_L	**wild type**	**56.5 ± 0.3**	**1.68 ± 0.02**	**12.1 ± 0.2**
	Y36A	34.4 ± 0.2	0.63 ± 0.02	15.2 ± 1.4
	Q37A	47.3 ± 0.2	1.20 ± 0.46	9.1 ± 0.8
	P44A	53.4 ± 0.2	1.54 ± 0.01	10.7 ± 0.9
	L46A	49.0 ± 0.2	1.06 ± 0.02	14.3 ± 1.4
	F98A	53.9 ± 0.2	1.66 ± 0.04	10.8 ± 1.8
V_H	**wild type**	**51.6 ± 0.3**	**0.91 ± 0.01**	**18.9 ± 3.9**
	V37A	42.3 ± 0.1	0.54 ± 0.04	16.8 ± 9.1
	G44A	50.5 ± 0.1	0.97 ± 0.05	19.4 ± 8.7
	L45A	43.0 ± 0.2	0.64 ± 0.02	25.4 ± 8.4
	E46A	41.1 ± 0.1	0.29 ± 0.03	27.6 ± 2.4
	W47A	37.8 ± 0.2	0.37 ± 0.03	15.1 ± 4.7
	W103A	50.0 ± 0.2	0.85 ± 0.03	24.1 ± 10.4

Although most of the GdmCl-induced unfolding transitions were not reversible, data evaluation according to a two-state equilibrium unfolding model was conducted to derive the midpoint of transitions ($D_{1/2}$), as well as the cooperativity parameter (m), for a qualitative comparison of the data. Thermal unfolding was measured by CD spectroscopy at a constant wavelength of 215 nm in a 0.5 mm quartz cuvette with a heating rate of 20 °C h^{-1}. The midpoint of transitions (T_{melt}) was determined by Boltzmann fit. The red color-code indicates a high deviation from the wild type domain, orange a mild deviation.

V_L wild type exhibited a melting temperature (T_{melt}) of 56.5 ± 0.3 °C and a GdmCl stability ($D_{1/2}$) of 1.68 ± 0.02 M. Concerning the point mutations, the picture was the same for both unfolding methods. P44A and F98A were only slightly less stable than the wild type (T_{melt}: 53.4 – 53.9 °C; $D_{1/2}$: 1.54 – 1.66 M) while Q37A and L46A showed a more

pronounced instabilty (T_{melt}: 47.3 – 49.0 °C; $D_{1/2}$: 1.06 – 1.20 M) and Y36A was about 50 % less stable than the wild type with a T_{melt} of 34.4 ± 0.2 and a $D_{1/2}$ of 0.63 ± 0.02 M. The cooperativity values all lie close to the range of the wild type domain.

For V_H, both methods also lead to an almost consistent result. The V_H wild type domain with a T_{melt} of 51.6 ± 0.3 °C and a GdmCl stability ($D_{1/2}$) of 0.91 ± 0.01 M is generally less stable than the V_L domain (Figure 12).

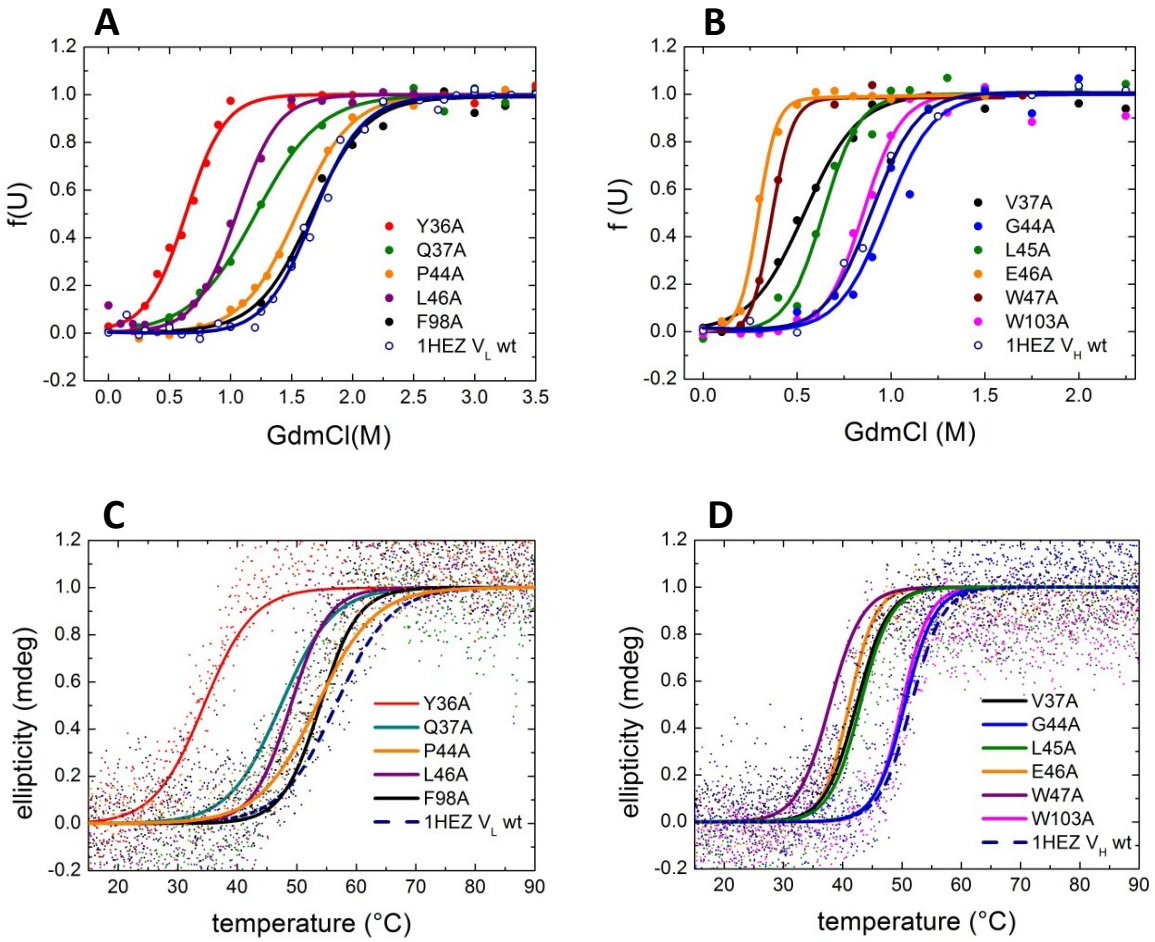


Figure 12 : Influence of the conserved residues on protein stability.
The stability of the V_L (A and C) and V_H (B and D) point mutations against thermal denaturation (C and D) was assessed by CD spectroscopy at a constant wavelength of 215 nm and a heating rate of 20 °C h^{-1}. Chemical denaturation (A and B) with GdmCl was determined by fluorescence spectroscopy, measuring the change of the intrinsic tryptophan fluorescence upon unfolding.

Comparing the stabilities of the V_H and V_L wild type domains with regard to differences between the two applied unfolding methods the observed deviation in chemical stability (≈50 % less stability for V_H) is clearly more pronounced than for thermal stability (≈10 % less stability for V_H). This observation might be due to the different unfolding principles of

the two methods but could also be associated with a dimer to monomer shift at low protein concentrations, since thermal unfolding is perfomed at a protein concentration of 20 µM and chemical unfolding at a very low protein concentration of 1 µM. Regarding the point mutations, for W103A and G44A the stabilty values lie within the wild type range (T_{melt}: 50.0 – 50.5 °C; $D_{1/2}$: 0.85 – 0.97 M), L45A and V37A are less stable (T_{melt}: 42.3 – 43.0 °C; $D_{1/2}$: 0.54 – 0.64 M) and E46A and W47A are the least stable point mutations (T_{melt}: 37.8 – 41.1 °C; $D_{1/2}$: 0.29 – 0.37 M). So for chemical unfolding E46A depicts the least stable mutant, for thermal unfolding this holds true for W47A. This observation fits very well to the proposed concentration-dependent monomer/dimer difference for both methods. Assuming a dimeric starting state for thermal unfolding, which leads to an overall stabilization of the V_H domain due to shielding of otherwise surface-exposed hydrophobic residues, W47A should exhibit a lower stabilty due to its partial monomeric nature. For chemical unfolding with concentrations underneath the concentrations applied for the AUC size determination, a potentially monomeric state for all point mutations might lead to a different stabilty distribution than observed for thermal unfolding, which is indeed the case. Of course, this assumption necessitates further data concerning the homodimer affinities and kinetics.

3.1.3. Influence of the conserved residues on V_H/V_L interaction

The interaction between the V_H and V_L domain was determined by SPR via amine coupling of either the V_H or V_L wild type domain. Equilibrium runs were not possible due to the presence of the V_H homodimer, which is not distinguishable by size from the V_H/V_L heterodimer. Sedimentation velocity runs with one interaction partner being fluorescently labeled via maleimide coupling at an N-terminally added cysteine residue only worked for the V_L wild type domain. With the labeled V_H wild type domain no binding could be measured. So either the GSGS linker with the additional cysteine residue, the labeling procedure or the label itself interferes with the V_L/V_H interaction. The determined SPR K_Ds for the wild type domains and point mutations are shown in Table 3.

Table 3 : Influence of the point mutations on the interaction between V_H and V_L.

	Mutation	Association K_D [µM] SPR	Domain	Mutation	Association K_D [µM] SPR	Association K_D [µM] AUC
V_L	wild type	**9.9**	**V_H**	wild type	**3.8**	**2.2**
	Y36A	**2.2**		V37A	**63.1**	**4.6**
	Q37A	**10.0**		G44A	**3.9**	**13.2**
	P44A	**110**		L45A	**100**	**>115**
	L46A	**11.1**		E46A	**6.5**	**2.8**
	F98A	**23.1**		W47A	**13.7**	**1.2**
	V_H wild type	**40.0**		W103A	**29.0**	**>70**

Interaction was assessed by SPR using a Biacore X100 or AUC by sedimentation velocity runs. The wild type domain was immobilized on a CM5 chip surface (final response units of 300 for V_L wild type and 1500 for V_H wild type). The obtained sensograms for the point mutations from multi cycle runs were fitted by using the Biacore X100 evaluation software in kinetics/affinity mode for K_D determination. For data evaluation of the AUC titration data, the integrated Svedberg values were plotted against the protein concentration. The K_D was assigned to the protein concentration corresponding to the half-maximum of the curve. Numbers highlighted in red correspond to a high deviation from the wild type values, orange indicates a mild deviation.

With an immobilized V_H wild type domain, the K_D for the V_H/V_L wild type heterodimer is 9.9 µM. The determined K_D for the V_H/V_H homodimer is rather high with 40.0 µM, indicating a favored binding to the V_L domain. The V_L point mutations Q37A and L46A lie within the wild type range, Y36A even shows a higher affinity with a slightly lower K_D of 2.2 µM. Less binding is observed for F98A with a K_D of 23.1 µM and the strongest impairment of the V_H/V_L interaction was measured for P44A with a K_D of 110 µM.

On the V_H side, with an immobilized V_L wild type domain, the K_D for the wild type heterodimer is 3.8 µM. This is about 2-fold lower than *vice versa*. Unfortunately, no errors are available for the obtained data. For the generation of standard deviations repetitions of the measurements have to be performed. Regarding the V_H point mutations, G44A and E46A show a wild type like interaction, while the tryptophan mutants W47A and W103A exhibit an impaired binding with a K_D of 13.7 and 29.0, respectively. The worst affinity was observed for V37A with a K_D of 63.1 µM and L45A with 100 µM.

For the V_H/V_L interaction data it has to be kept in mind that the determined values are no absolute but rather apparent K_D values. This is due to the homodimer nature of the V_H domain and the associated competition for binding between V_L and V_H, even though the heterodimer formation is favored according to the SPR results. For W47A, which depicts

the only V_H mutant with a monomeric fraction, the measured affinity is also not completely comparable to the other point mutations. In this case, it can be assumed that the homodimer affinity is significantly lower compared to the heterodimer, leading to an overall better K_D than expected for a completely dimeric W47A V_H domain.

The obtained results are in excellent agreement with the data of Dr. Eva-Maria Herold's characterization of the same alanine exchange mutants within the murine MAK33 IgG1 variable domains. Some variations can be observed but the effects of the mutations in terms of structure, stability and V_L/V_H interaction are the same. For both antibodies, R38A and E46A exhibited a very high aggregation propensity during purification and a misfolding behavior in the far-UV spectra. The monomeric MAK33 V_H domain is generally very unstable and therefore sensitive towards any mutation. Nevertheless, for both antibodies, W47A is the least stable V_H point mutation; on the V_L side this is Y36A. In terms of V_H/V_L interaction, P44A exhibits the highest impairment for V_L, the same is observed for L45A on the V_H side followed by V37A and W103A. Antigen binding cannot be compared due to the missing data for the 1HEZ variable domains.

3.1.4. Effects of the variable domain point mutations on antigen binding

The variable domains used in this study are derived from a human rheumatoid factor (RF) IgM antibody (pdb entry: 1HEZ) involved in the autoimmune disease rheumatoid arthritis [168, 169]. These autoantibodies can be found in the sera of patients with rheumatoid arthritis. They recognize different epitopes located at the Fc region of IgG antibodies, favoring to the C_H2/C_H3 interface region [170-172]. The highest affinity is observed for the IgG4 subclass, the lowest for IgG3, many cannot bind IgG3 at all. A crystal structure in complex with its antigen is not available. The existing crystal structures show the 1HEZ Fab fragment in complex with protein L and protein A [168, 169], where the binding is promoted by interaction with the framework regions of the variable domains. Protein L binds specifically to the V_L domain from κ light chains, while protein A binds to the human V_H3 subclass. The CDRs do not contribute at all to this mode of binding.

To determine the functionality of the variable domain point mutations regarding antigen binding, two different methods were used, ELISA and SPR (Figure 13). The applied ELISA approach with an immobilized biotinylated antigen (human IgG antibody) and detection via a FLAG-tagged wild type domain interaction partner was not successful. As can be

seen in Figure 13, a concentration dependent signal could be obtained for both combinations, V_H or V_L domain with an additional FLAG-tag. Unfortunately, for the control samples without untagged interaction partner the same signal increase was measured. This effect might either be caused by a different mode of antigen binding, meaning that the V_H as well as the V_L domain are able to bind to the antigen in the absence of the corresponding variable domain.

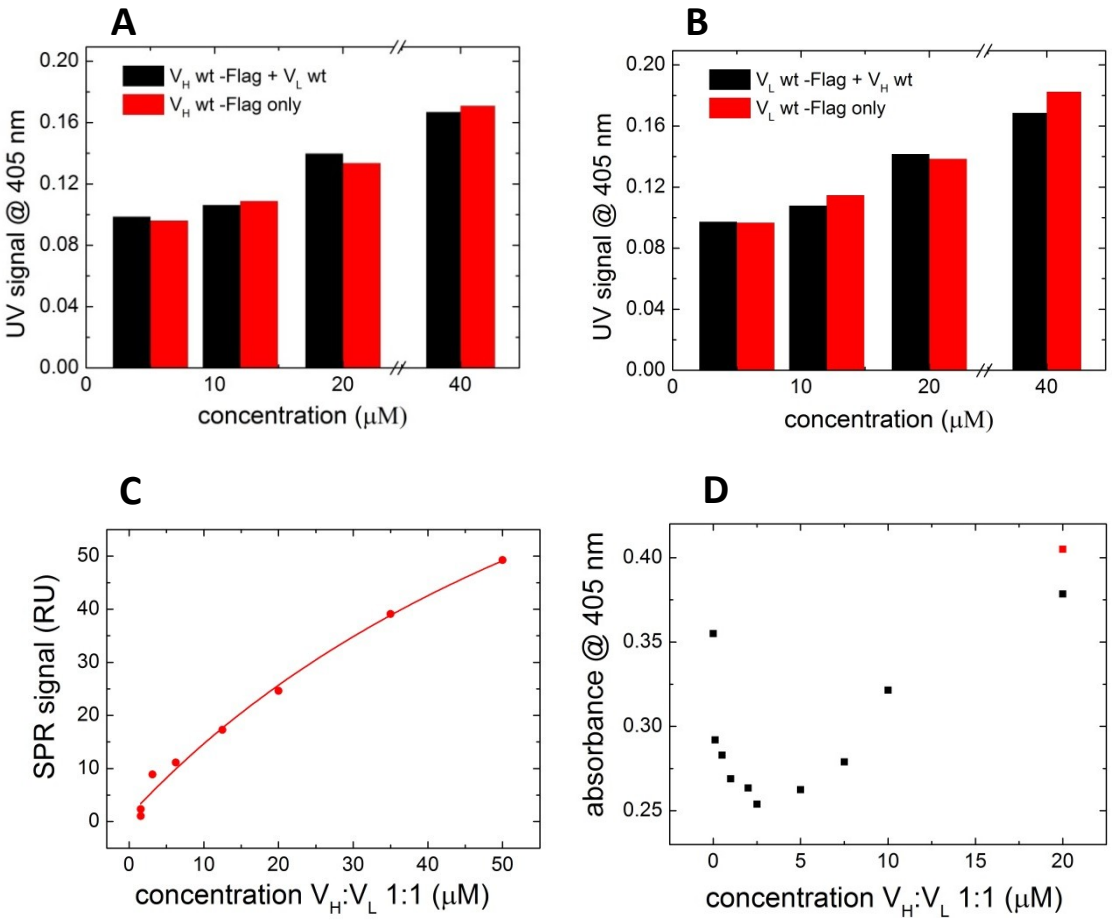


Figure 13 : Antigen binding of 1HEZ V_H and V_L wild type.

Antigen binding of the Fv fragment was assessed by ELISA (A ,B and D) and SPR (C). For ELISA either V_H (A) or V_L (B) contained a C-terminal FLAG-tag for detection; (D) contained tag-less V_H and V_L as well as human biotinylated IgG Fc fragments as antigen and HRP conjugated Protein L for detection; titrations of a 1:1 ratio were conducted. For SPR (C) biotinylated human IgG antibodies were immobilized on a Streptavidin Chip and titrations with 1:1 ratios of V_H and V_L wild type were measured by multicycle. For data evaluation the Biacore X100 evaluation software was applied.

A more likely explanation might be the unspecific binding of the FLAG-tag to the antigen, which could be tested with a control ELISA, where the FLAG-tag is added as a competitor. An additional ELISA approach with biotinylated human IgG Fc fragments as antigen and

HRP-conjugated Protein L for detection was tested. Unfortunately similar results were obtained, including a high signal for the PBS sample without protein. This indicates an unspecific binding of Protein L to the antigen, which should not be the case since the Fc fragment does not contain the Protein L binding partner V_L. So maybe the applied antigen contains Fab fragment impurities, which would explain the observed unspecific binding. The next approach for measuring antigen binding was SPR, using a Streptavidin Chip for immobilization of the biotinylated antigen. The result is shown in Figure 13 C. A concentration-dependent increase in the SPR signal was detected but due to insufficient regeneration conditions, the bound interaction partner could not be properly removed. This leads to an accumulation of the bound variable domains on the chip surface, diminishing possible antigen binding sites. This explains the observed SPR signal with an almost linear fitting curve and a highly overestimated K_D value. Further improvements of the regeneration conditions are necessary.

3.1.5. Additional point mutations within a different antibody context

Within the context of Dr. Eva-Maria Herold's studies on MAK33 (pdb ID 1FH5) IgG point mutations, CDR and framework grafting experiments with more stable variable domains (1DH5 V_L and 1DHU V_H) were performed. CDR grafting between two different V_H or V_L domains is a quite common approach for therapeutic applications [126]. The chosen variable domains were identified by Ewert and coworkers [131] as the most stable ones within a stability screening of isolated human variable domains (seven consensus sequences of the major human germline subclasses for each domain) retrieved from the human combinatorial antibody library (HuCAL). CDR sequences were chosen from the germline gene with highest usage in a family, only for V_H a solubility increasing long CDR3 was used to overcome the high aggregation tendency of this domain in isolation. This long CDR3 is supposed to partially shield the exposed hydrophobic surface which is usually covered by V_L [131]. The variable domains with the highest thermodynamic stability and expression yield were V_H3 and Vκ3.

The CDRs account for approximately 25% of the variable domains [173]. Even if they are primarily associated with antigen binding, the CDRs also comprise 25 % of the interface residues, mainly involving CDR1 and especially CDR3 residues [121, 173]. It could be shown that residues within the framework as well as interface contributing residues of

the CDRs can influence the interface and thus the V_H/V_L packing geometry [121, 122, 132, 173, 174]. An additional influence by hypervariable loop residues is exhibited by the so called proximate zone, which is situated at the base of the antigen binding site, comprising residues that do not actively participate in the interface [121].

In order to test the generality of the observed point mutation effects for the MAK33 antibody, the point mutation and grafting projects were combined. The V_L point mutations P44A which exhibited severe effects in V_H/V_L interaction and antigen binding as well as F98A and L46A, which behaved generally like the wild type with only small deviations in stability or V_H/V_L interaction and antigen binding, were transferred to the CDR grafting construct 1DH5_MAK. This contains the MAK33 framework and the 1DH5 CDRs. With this approach the CDR dependency of the observed effects can be examined. The structure and sequence alignment (Figure 14) of the MAK33 and 1DH5 V_L domains shows a high overlap. Regarding the CDRs, CDR2 and 3 own the same amino acid length, CDR1 differs by only one amino acid.

Figure 14 : Structure and sequence alignment between 1DH5 and MAK33 V_L.

A pymol cartoon representation of the MAK33 (pdb ID 1FH5) V_L domain is depicted in light green with the three CDRs highlighted in dark green, the 1DH5 V_L domain is represented in light orange with the CDRs highlighted in dark orange. The sequence alignment was performed with CLUSTALW, the CDRs are also color highlighted.

```
MAKVL           MDIVLTQSPATLSVTPGDSVSLSCRASQSISN-NLHWYQQKSHESPRLLIKYASQSISGI
1DH5            MDIVLTQSPATLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYGASSRATGV
                **************::**: .:********:*.  * *****. ::*****  **.  :*:

MAKVL           PSRFSGSGSGTDFTLSINSVETEDFGMYFCQQSNSWPLTFGAGTKLELKR
1DH5            PARFSGSGSGTDFTLTISSLEPEDFAVYYCQQHYTTPPTFGQGTKVEIKR
                *:**************:*.*:*.***.:*:***  : * *** ***:*:**
```

As already performed for the point mutations in the previous chapter, a characterization in terms of structure, stability, V_H/V_L interaction and antigen binding was conducted. Secondary structure determination by far-UV CD showed a typical ß-sheet structure with a minimum at 218 nm and maximum at 200 nm for all point mutations (Figure 15 A). Near-UV CD spectra of 1DH5_MAK V_L and the point mutations exhibited the same shape

with two minima at 285 nm and 275 nm (Figure 15 B). Only the signal intensity differed from 1DH5_MAK, while P44A deviated only slightly, F98A exhibited a 2-fold higher amplitude and L46A lies in between. Since the near-UV CD measures the environment of the aromatic amino acids and the F98A mutation lacks one aromatic amino acid this might be responsible for the signal change. For Dr. Eva Maria Herold's MAK33 V_L point mutations, similar data was obtained. The noticeable signal intensity increase for the near-UV CD spectra of the point mutations could only be observed for F98A in the MAK33 format.

Stability was assessed by thermal and chemical unfolding (Figure 15 C and D).

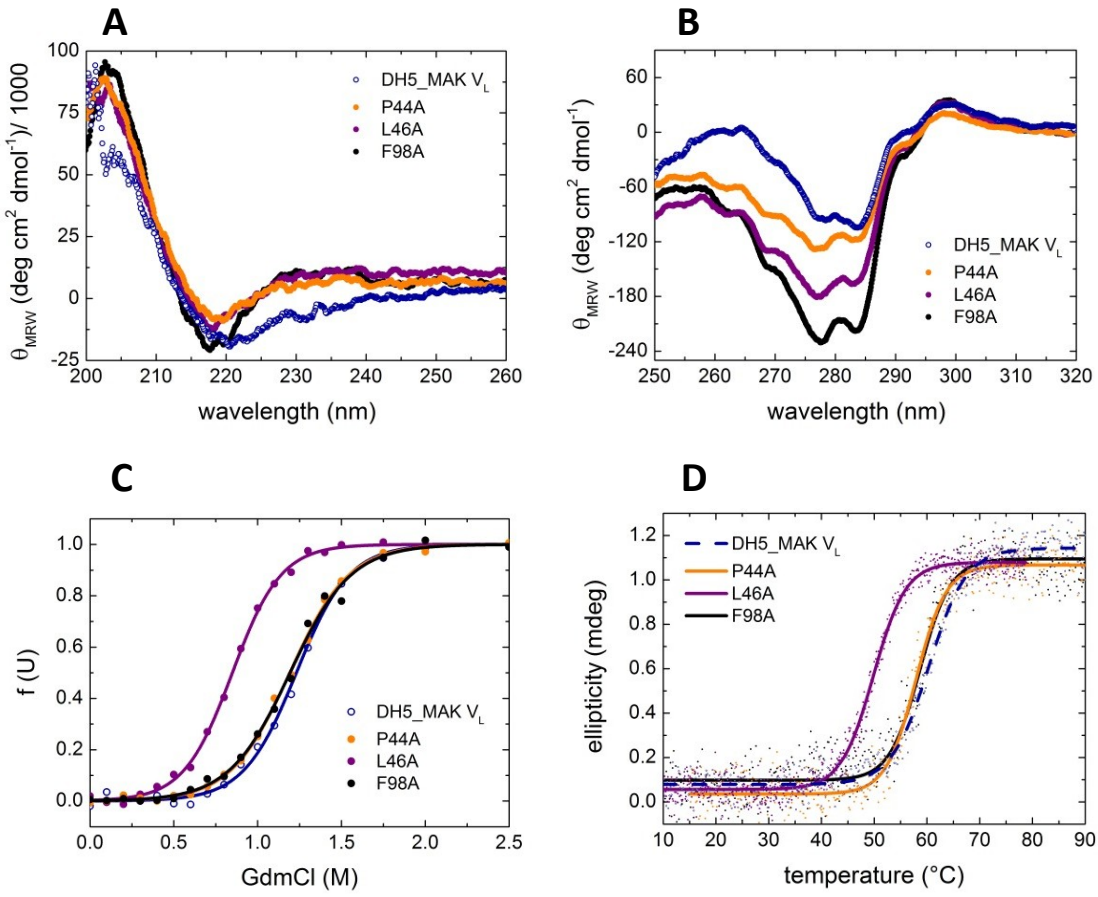


Figure 15 : Influence of the selected point mutations in MAK_1DH5 V_L on structure and stability.
In (A) far-UV CD spectra and in (B) near-UV CD spectra of the V_L mutants are shown. For the spectra 15 accumulations each were recorded and buffer-corrected (PBS). All CD measurements were performed at a protein concentration of 20 µM (FUV-CD) and 50 µM (NUV-CD) in 0.5 mm (FUV) or 5 mm (NUV) quartz cuvette at 20°C. The stability of the V_L point mutations against thermal denaturation (D) was assessed by CD spectroscopy at a constant wavelength of 215 nm. Chemical denaturation (C) with GdmCl was determined by fluorescence spectroscopy, measuring the change of the intrinsic tryptophan fluorescence upon unfolding at 20 °C.

While P44A and F98A behaved exactly like 1DH5_MAK, for L46A in both cases a decrease in stability was observed.

In comparison to the stability effects observed by Dr. Eva-Maria Herold for the point mutations in MAK33 V_L, L46A also exhibited a decrease in stability (Table 4). In the case of F98A and P44A, though, the data differed since they also showed a decrease in stability for MAK33 even if it was not as prominent as for L46A.

The V_H/V_L interaction was assessed by SPR with an immobilized V_H MAK33 domain via maleimide coupling (Table 4). So the following K_Ds obtained for the V_L point mutations and grafting mutations all account for the binding to MAK33 V_H wild type. Interestingly, the observed K_D for 1DH5_MAK is very high with 50 µM, in contrast to the *vice versa* grafting construct MAK_1DH5, containing the 1DH5 framework and MAK33 CDRs, with a K_D of 2.7 µM. Since the framework region is thought to be mainly responsible for the binding to V_H, an opposite result was expected for the grafting constructs. The K_D for MAK33 V_L wild type with 1.2 µM lies close to the MAK_1DH5 value while the 1DH5 wild type with a K_D of 80 µM is slightly above the 1DH5_MAK value. Since for 1DH5 no naturally occurring V_H counterpart exists, the weak binding to the V_H MAK33 domain is not surprising. A very similar distribution of binding impairments was observed in Dr. Eva Maria Herold's AUC equilibrium run experiments (Table 4), even if the maintained K_Ds are generally lower. This indicates a high importance of the CDRs for V_H/V_L interaction, which might be the result of conformational rearrangements in the frameworks dependent on the CDR sequence. Nevertheless for the point mutations within the 1DH5_MAK context a decreased binding was observed for F98A (K_D 156 µM) and especially for P44A (K_D 227 µM), L46A behaved like the wild type construct. These results correlate very well with the SPR data obtained for the same point mutations in MAK33 (Table 4). Only the corresponding AUC data for MAK33 F98A differed, exhibiting a wild type like binding.

Regarding the K_D deviation between AUC and SPR it has to be kept in mind that they are based on very different principles.

Due to fact that the MAK33 CDRs are exchanged within the 1DH5_MAK grafting construct, no binding to creatine kinase was observed for this domain and the corresponding point mutations.

Table 4 : Comparison of the characteristics of selected V_L point mutations within a CDR grafting and wild type context.

Domain	Mutation	Association K_D [µM] SPR	Association K_D [µM] AUC	T_{melt} [°C]	$D_{1/2}$ [M]	Cooperativity [kJ mol^{-1} M^{-1}]	Antigen binding app. K_D [µM]
1DH5_MAK	wild type	50	3.5	60.4 ± 0.2	1.2 ± 0.1	15.4 ± 2.0	NB
	P44A	227	ND	57.9 ± 0.2	1.2 ± 0.1	13.8 ± 1.1	NB
	F98A	156	ND	58.7 ± 0.2	1.2 ± 0.1	13.4 ± 1.5	NB
	L46A	44	ND	49.8 ± 0.1	0.8 ± 0.1	15.8 ± 0.7	NB
1DH5		85	*3.8*	*69.0 ± 0.6*	*2.4 ± 0.2*	*17.1 ± 6.1*	NB
MAK_1DH5		2.7	*0.3*	*55.5 ± 0.2*	*1.4 ± 0.2*	*15.6 ± 4.8*	*1.2 ± 0.3*
V_L MAK33	wild type	1.2	*0.2*	*52.3 ± 0.2*	*1.2 ± 0.7*	*17.1 ± 6.7*	*0.9 ± 0.1*
	P44A	47.8	*15.2*	*44.5 ± 0.5*	*1.0 ± 0.2*	*15.0 ± 3.2*	*10.5 ± 0.4*
	F98A	5.7	*0.6*	*48.1 ± 0.3*	*1.0 ± 0.2*	*19.0 ± 2.8*	*8.3 ± 0.5*
	L46A	0.5	*0.6*	*44.8 ± 0.4*	*0.7 ± 0.2*	*21.2 ± 3.5*	*4.6 ± 1.2*

K_D values for V_H/V_L interaction were determined by AUC sedimentation equilibrium experiments and SPR. For AUC measurements unmodified proteins were used, for SPR the wild type interaction partner was immobilized on a CM5 chip via maleimide coupling of an N-terminally added cysteine residue. Values in italic and grey mark the data adapted from Dr Eva-Maria Herold's studies [166] . Stabilities against the chemical and thermal denaturation are depicted by T_{melt} and $D_{1/2}$ values. GdmCl-induced unfolding data were evaluated according to a two-state equilibrium unfolding model to derive the midpoint of transitions ($D_{1/2}$), as well as the cooperativity parameter (m), for a qualitative comparison of the data. Thermal unfolding was measured by CD spectroscopy at a constant wavelength of 215 nm in a 0.5 mm quartz cuvette with a heating rate of 20 °C h^{-1}. The midpoint of transitions (T_{melt}) was determined by Boltzmann fit. Antigen binding of the Fv fragment was determined by ELISA with the antigen creatine kinase. Detection of the V_H and V_L wild type domains was possible via an introduced FLAG-tag at the C-termini. Samples were corrected for the signal of the single FLAG-tagged variable domains. Apparent K_D values were obtained by a Boltzmann fit. Experiments were performed at 20°C. NB indicates that no binding could be observed.

To sum up, the effects of the examined V_L point mutations within a CDR grafting context are highly comparable to the same point mutations in the MAK33 wild type context. The near-UV CD spectra for both V_L domains indicate small structural rearrangements for F98A, in the case of 1DH5_MAK also slightly for L46A and P44A. Concerning stability, the MAK33 V_L seems to be more sensitive towards the introduced point mutations than 1DH5_MAK. The overall picture is the same, though; L46A is the point mutation with the lowest stability. Concerning the V_H/V_L interaction, similar effects were observed, more or less pronounced in the two different backgrounds. F98A shows a lower affinity for the V_H domain while P44A results in a striking impairment of V_H binding. Interestingly, CDR

grafting had a severe effect on V_H/V_L binding in contrast to framework grafting. This indicates a greater contribution of the CDR regions than expected.

3.1.6. CDR grafting between MAK33 V_H and V_L

To further assess the influence of the CDRs on variable domain geometry and functionality, CDR swap experiments between the V_L and V_H domain of the MAK33 antibody were performed. This represents a very rough approach, considering the fact that the variable domains and their corresponding CDRs with distinct sequence features evolved over millions of years, exposed to the evolutionary pressure of adaption and optimization. Also, possible interactions of the CDRs with specific framework residues [123, 125, 134] as well as the different lengths of V_L and V_H CDRs, especially for CDR2 and CDR3 (see Figure 16) and the distinct canonical structures encoded by CDR sequence are not paid attention to within this approach [175].

```
MAK33 VL: DIVLTQSPATLSVTPGDSVSLSCRAS-QSISNNLHWYQQKSHESPRLLIKYASQSISG
MAK33 VH: EVQGVES-GGGLVKPGGSLKLSCAASGFTFSDYYMYWVRQTPEKRLEWVATISDGGSY
          ::  .:* .   *.**.*:.*** **  ::*:   :: ::: *.    :   *:. *

-----IPSRFSGS--GSGTDFTLSINSVETEDFGMYFCQQSNSWPLTFGAGTKLELKR--------
YYPDSVKGRFTISRDNAKNNLYLQMSSLKSEDTAMYYCARDKAYYGNYGDAMDYWGQGTSVTVSSA
     : .**: *  .: .:: *.:.*:::** .**:* :.:::  .:* . .   :
```

Figure 16: Sequence alignment of the MAK33 variable domains
A sequence alignment performed with CLUSTALW between the two variable domains of MAK33 (pdb ID 1FH5) is shown. The V_L CDRs are depicted in red the V_H CDRs in green, the consensus amino acids are highlighted in cyan.

Thus our swap experiments test to what extent the variable domains are able to tolerate severe CDR exchanges. The sequence alignment (Figure 16) between the V_H and V_L domains shows a modest degree of sequence identity, at least for the framework regions. Almost no overlap is visible for the CDRs, especially CDR2 and CDR3 differ greatly in terms of length. Apart from the CDRs, both variable domains exhibit a conserved ß-barrel framework stabilized by an internal disulfide bridge [101, 121, 173, 174].

A comparison of the mutant far-UV CD spectra with the corresponding MAK33 wild type domains exhibits severe secondary structure changes for both CDR grafting mutants (Figure 17 A, Table 5). For V_L with V_H CDRs, one minimum at about 210 nm is visible, for V_H with V_L CDRs two minima at 222 nm and 208 nm were detected. So both mutant domains are not unfolded (no minimum at 200 nm), V_L shows neither a typical ß-sheet

nor a helical structure but interestingly the CD spectrum of V_H is indicative of an α-helical structure, which is not the case for any immunoglobulin domain. This can be confirmed by the CDNN analysis of the far-UV spectra which assigns 70.9 % α-helical structure to V_H with V_L CDRs compared to 7.8 % for V_H wt, 6.7 % for V_L wt and 14.1 % for V_L with V_H CDRs (Table 5). Due to the low protein yield for V_H, near-UV CD spectra could only be measured for the V_L mutant (Figure 17 B). The spectrum of V_L with V_H CDRs is clearly different from that of the wild type, confirming structural rearrangements with high deviation from the wild type domains, as already observed for secondary structure.

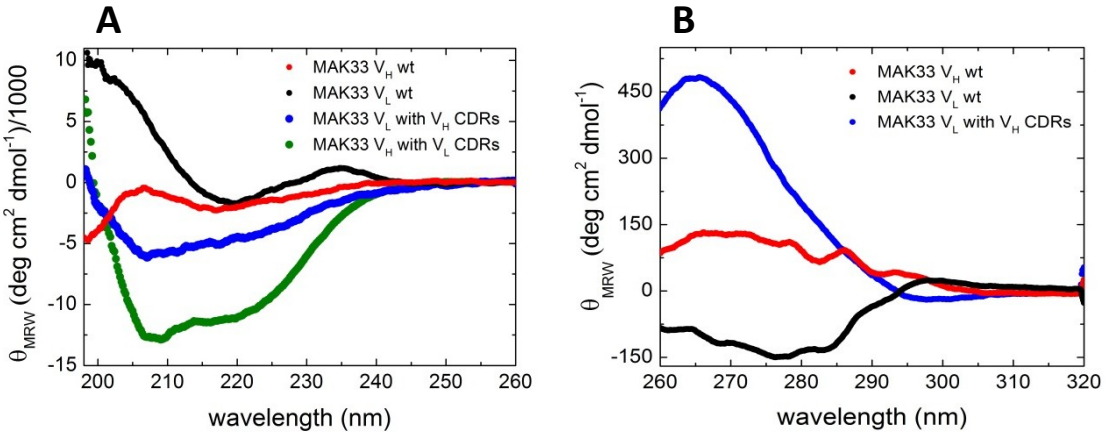


Figure 17: Structural characterization of the V_H /V_L CDR grafting constructs.
In (A) FUV-CD spectra and in (B) NUV-CD spectra of the V_L and V_H grafting constructs are shown. For the spectra 15 accumulations each were recorded and buffer-corrected (PBS). All CD measurements were performed at a protein concentration of 20 µM (FUV-CD) and 50 µM (NUV-CD) in 0.5 mm (FUV) or 5 mm (NUV) quartz cuvette at 20°C.

Table 5: Secondary structure composition of the V_H/V_L grafting constructs.

structure	V_H wt	V_L wt	V_H with V_L CDRs	VL with V_H CDRs
Helix %	7.8	6.7	70.9	14.1
Antiparallel %	47.0	39.3	1.0	23.2
Parallel %	6.6	5.4	3.6	5.3
ß-turn %	14.2	21.0	13.5	20.8
Random coil %	34.1	37.3	15.4	35.1
Error %	9.7	9.7	4.4	1.3

Secondary structure composition was determined with CDNN 2.1. This program assignes the fraction of each secondary structure element based on a buffer corrected far-UV spectrum ranging from 200-260 nm, comparing the uploaded data with 33 spectra of other proteins.

Further structural data was obtained by fluorescence spectroscopy (Figure 18). On the one hand, the intrinsic tryptophan fluorescence was monitored (Figure 18 A and B). The results show a typical red shift between the native and denatured state, indicating the

presence of a structure with a hydrophobic core involving the conserved tryptophan. Besides, V_H with the V_L CDRs exhibits a signal increase for the denatured sample. The opposite effect is measured for V_L with V_H CDRs, which again suggests a CDR-associated misfolding of the V_L mutant. On the other hand, the hydrophobicity of the surface was assessed by ANS binding (Figure 18 C). No binding accompanied by a signal decrease of ANS was visible, which suggests that surface-exposed hydrophobic regions are absent. This means that the hydrophobic core is present for both mutants. To a small extent for V_H wild type and V_L with V_H CDRs and strongly pronounced for V_H with V_L CDRs, though, a blue shift accompanied by a signal increase is visible [176]. This occurs when ANS interacts with the positively charged amino acids lysine or arginine [176].

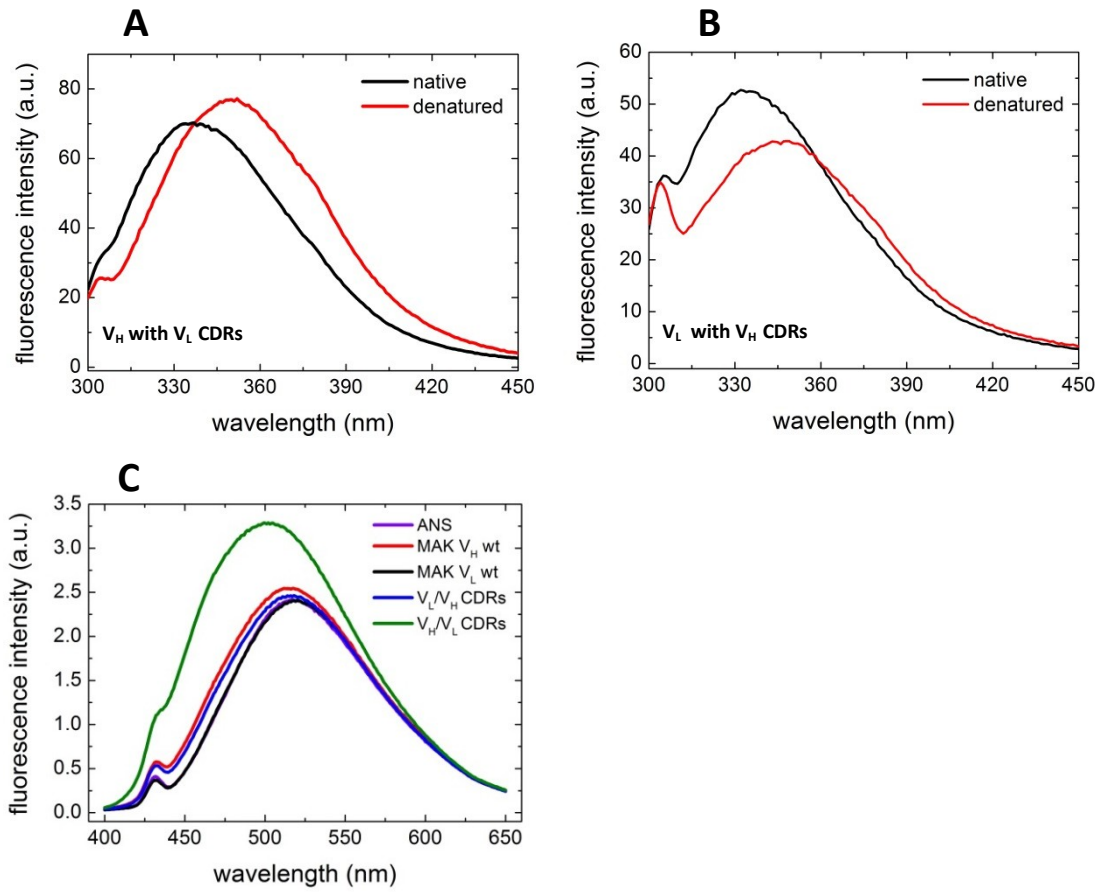


Figure 18: Fluorescence spectra and surface hydrophobicity of the V_H/V_L CDR grafting constructs.

Intrinsic tryptophan fluorescence was determined by fluorescence spectroscopy at an emission wavelength of 280 nm for (A) V_H with V_L CDRs and (B) V_L with V_H CDRs. The samples were measured under native and denaturing conditions (6M GdmCl). To analyze the change of surface hydrophobicity of the grafting constructs binding to 50 µM ANS was monitored at an excitation wavelength of 380 nm by fluorescence spectroscopy. Measurements were performed at a protein concentration of 1 µM at 20 °C.

Hence, these domains seem to possess surface-exposed positively charged amino acids that are able to interact with ANS. For the similar slight shift of MAK V_H wild type and V_L with V_H CDRs, the CDRs are probably responsible for this effect. Since the CDR exchanges decreased the number of the corresponding amino acids for V_H with V_L CDRs, the observed change is rather caused by framework misfolding than the introduced V_L CDR residues.

The size distribution by AUC sedimentation velocity runs shows an equilibrium between monomer and dimer for both grafting mutants. For the wild type domains, only the monomer exists. Hence, the observed misfolded state seems to favor the presence of a dimer.

To sum up, the CDR grafting between V_H and V_L leads to a high aggregation propensity during the purification process and severe structural rearrangements potentially corresponding to misfolded domains occur.

3.1.7. Discussion

Alanine exchange point mutants of the conserved amino acids in the V_H/V_L interface have been characterized in terms of structure, stability and V_H/V_L interaction within the context of a human IgM (1HEZ) antibody. The comparison with the previous work of Eva Herold, who performed the same alanine point mutation approach for a murine IgG antibody revealed a high correlation of the observed effects. For both antibodies, the V_H mutations R38A and E46A were highly unfavorable for reaching the immunoglobulin folding. In the covariation analysis a very high ϕ-value was assigned for these two residues. In the structure, a salt bridge is formed between them [5], R38 is buried behind E46 while E46 does not interact with other V_H interface residues, but might electrostatically affect V_L binding [5]. Camelids and cartilaginous fish possess naturally occurring heavy chain antibodies lacking the light chain [84, 177, 178]. Interestingly a sequence alignment of MAK33 V_H with the variable domain (V_{HH}) of the camelid HcAb (PDB entry 2XT1) and the variable domain of “monomeric” shark IgNAR (V_{NAR}) (PDB entry 2I24) shows a match for the residues E46 and R38 (Figure 19).

Concerning the shark IgNAR, these two residues are actually the only ones from the conserved network investigated in this study that can also be identified at corresponding

positions in the otherwise highly divergent shark sequence, which probably evolved from a cell surface receptor [177, 179].

```
VHH camel:  ------MAQVQLVESGGGLVQAGGSLRLSCAASGSFFMSNVMAWYRQAPGKARELIAAIR
VH mouse:   MTMITNSREVQGVESGGGLVKPGGSLKLSCAASGFTFSDYYMYWVRQTPEKRLEWVATIS
                     ** ********:.****:*******  * .  * * **:* *  * :*:*
GGDMSTVYDDSVKGRFTITRDDDKNILYLQMNDLKPEDTAMYYCK----------ASGSSWGQGTQVTVSS
DGGSYTYYPDSVKGRFTISRDNAKNNLYLQMSSLKSEDTAMYYCARDKAYYGNYGDAMDYWGQGTSVTVSSA
.*.  * * *********:**: ** *****..**.********            : . *****.*****


VHH Ignar: -ARVDQTPQRITKETG-------ESLTINCVVRDSRCVLSTGYWYRKPPGSRNE--ESIS
VH mouse:  MTMITNSREVQGVESGGGLVKPGGSLKLSCAASGFTFSDYYMYWVRQTPEKRLEWVATIS
             : : :: :    *:*        **.:.*.. .        ** *:.* .* *   :**
DGG----------
RYVETVNRGSKSFSLRINDLTVKDSGTYRCKPESRYGSYDAVCAALNDQYGGGTVVTVNAA
DGGSYTYYPDSVKGRFTISRDNAKNNLYLQMSSLKSEDTAMYYC---
ARDKAYYGNYGDAMDYWGQGTSVTVSSA
***           *:. : :...:.: *:..*. :*:. * *   :*  :* .  . * :  * **
***.:*
```

Figure 19: Sequence alignment of MAK33 V_H with V_{HH} from camel and shark.
The sequence alignment was performed with CLUSTALW, using pdb entries 2XT1 (camel), 2I24 (shark) and 1FH5 (mouse). Asterisks indicate sequence overlaps, the three CDRS from MAK33 are illustrated in green. The conserved amino acids in MAK33 are highlighted in yellow, the same holds true for the V_{HH} domains.

Both HcAbs show a high biophysical stability and their distinct structural patterns have by now been successfully applied to generate monomeric human V_H domains [180-182]. The increase in stability is associated with an increase of hydrophilic residues in the framework regions. This compensates for the shielding effect of the corresponding hydrophobic residues by interaction with the V_L domain. In the case of camelid V_{HH}s, a mutation of the hydrophobic V_H/V_L interface residues V37, G44, L45 and W47 in favor for hydrophilic ones was found (V37F/Y, G44E, L45R, W47G) (see Figure 20) [5, 178, 183]. These residues are among the highly conserved V_H/V_L interface residues investigated within this study.

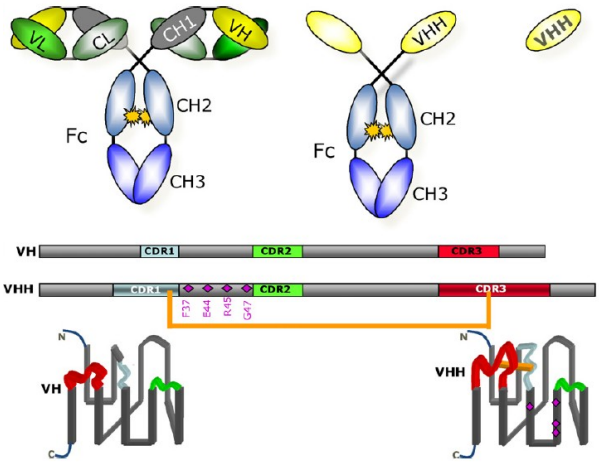


Figure 20 : Differences between V_H and V_{HH} domains from camel
On the left a conventional antibody, on the right a heavy chain antibody from camel is shown. Underneath the framework and CDR based differences between the V_H and V_{HH} domain are further highlighted, including amino acid exchanges in framework 2, increased CDR3 and CDR1 lengths and an additional disulfide bridge between CDR1 and CDR3. The figure is taken from [2].

Single alanine mutations of these hydrophobic amino acids showed a stability change only for W47A and V37A but in both cases it was a stability decrease. A comparison is difficult, though, as here only single amino acid exchanges towards a neutral amino acid were performed. Nevertheless, regarding a potential interaction between the four residues, it might be interesting to combine the four alanine point mutations.

The structural data of the two point mutation studies differs slightly. For V_L both identify a higher near-UV CD signal for F98A. In the case of 1HEZ, a higher near-UV CD signal was also observed for L46A and P44A. Since the monomeric MAK33 V_H domain is very unstable in isolation, almost any mutation has a disfavoring effect. This renders comparisons with the more stable dimeric 1HEZ V_H domain complicated. Additionally, for MAK33 no proper folded E46A mutant could be purified and characterized for comparison reasons. In terms of stability, almost all mutations, especially for V_H lead to a small decrease of stability. The outliers for each domain are V_L Y36A and V_H W47A for which the most striking stability decrease was observed. Residue W47 has already been in the focus of previous studies for improved solubility and stability but none of the mutations (W47L, W47R [184]) were favorable. In the study of Wang and coworkers W47 was identified as the central node based on the number and strength of its covariations with other interface residues [5]. In the context of 1HEZ, residue W47 also seems to be crucial for homodimer formation as W47A represented the only point mutation with a monomer fraction (50 % monomer, 50 % dimer). On the V_L side, Y36 and P44 were detected as the central node by Wang and coworkers. While Y36A was the outlier for V_L stability, the same holds true for P44A in terms of V_H/V_L interaction. Within this study the V_H/V_L interaction was determined with isolated variable domains. For MAK33 wild type, a K_D of 0.2 µM by AUC and 1.2 µM by SPR was measured by Dr. Eva-Maria Herold. Since AUC measurements with 1HEZ were not successful for both domains, only the K_D by SPR was compared, which was 9.9 µM when V_H was immobilized and 3.8 µM for an immobilized V_L wild type domain. Regarding the 1HEZ interaction data, it has to be kept in mind that all measured K_Ds are only apparent K_Ds, since absolute values cannot be determined due to the homodimer nature of the V_H domain. With regard to this homodimer formation it might be interesting to perform further experiments to

determine the K_D for this interaction. This has been tried within the SPR experiment resulting in a very high K_D of about 40 µM, which does not seem reasonable considering the AUC data with 100 % dimer down to a concentration of 5 µM. The different K_Ds for *vice versa* SPR coupling might be connected with the effect of sterical hindrance as both domains are randomly immobilized via lysine residues. So maybe for $V_{H,}$ less interaction sites are accessible which might explain the higher K_Ds for V_L. In comparison to the K_Ds of other antibody domains the V_H/V_L K_Ds measured within this study are comparable to the C_H1/C_L interaction. This was measured by C_L-induced change in intrinsic fluorescence of C_H, resulting in a K_D of 6.2 µM [6]. For the C_H3 dimer, a much higher affinity was observed with a $K_D < 10^{-4}$ M using SEC [185, 186]. Even if the interaction for V_H/V_L and C_H1/V_L is rather low, in the Fab format it strongly increases [187]. Concerning the effect of the point mutations on the V_H/V_L interaction, again the outliers coincide very well. Almost no binding was observed for V_H L45A, followed by V37A and W103A according to the AUC data of MAK33 and the SPR data for 1HEZ. The extent and order of affinities is subject of variations for the different methods but in all cases L45A shows almost no binding. Regarding the deviations between the AUC and SPR data, it has to be kept in mind that they are based on very different principles. While AUC measurements are performed in solution, SPR experiments necessitate one interaction partner to be immobilized on the chip surface. Next to the very harsh coupling conditions (acetate buffer at pH 4.5) which can affect the protein, effects like steric hindrance with diminished accessibility of the binding sites need to be considered. Also the different duration of the corresponding experiments might play a role with regard to the low stability of the V_H domain. On the V_L side, P44A resulted in the most striking impairment of binding to V_H. This can also be confirmed by the P44A mutation within the CDR grafting construct 1DH5_MAK. For the 1HEZ V_L domain, P44A is followed by F98A in terms of impaired binding behavior, the same holds true to different extents for 1DH5_MAK and the SPR data of MAK33. Concerning the effect of the point mutations on antigen binding, no data is available for 1HEZ due to the failed AUC and SPR experiments. This necessitates further optimization.

The additional point mutation data for the transfer of selected V_L point mutations (P44A, L46A, F98A) to the grafting construct 1DH5_MAK which contained the MAK33 framework and 1DH5_CDRs, were also in good agreement with the MAK33 data. With this approach the CDR dependence of the observed effects was examined. This is reasonable due to the

fact that 25 % of the interface residues are represented by CDR residues and several studies have proven that interface contributing residues of the CDRs can influence V_H/V_L association and thus the interface [121, 122, 132, 173, 174]. In Figure 21 the structure of the 1HEZ Fv fragment is shown, with regard to the location of the CDRs and their contribution to the V_H/V_L interface.

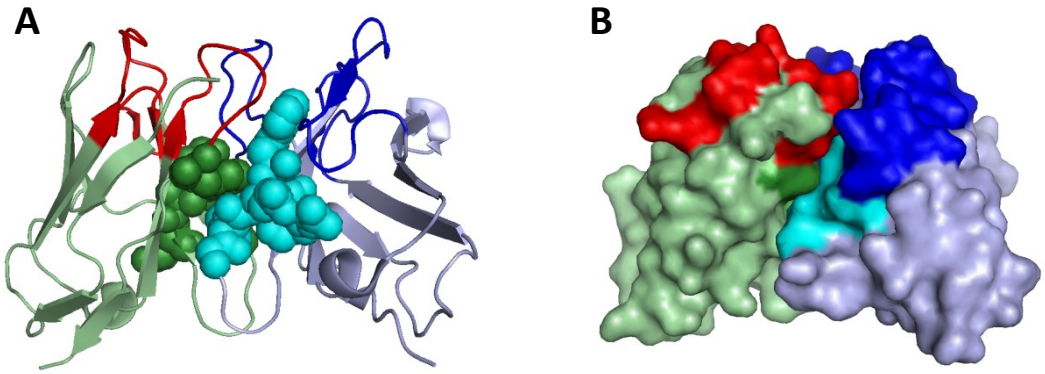


Figure 21: Structure of the MAK33 Fv fragment and the CDR contribution to the V_H/V_L interface.
(A) The pymol cartoon representation of the 1HEZ Fv fragment with the interface amino acids shown in sphere representation. The V_L domain is depicted in light green with the three CDRs highlighted in red and the interface residues in dark green. V_H is shown in light blue, the CDRs are in dark blue and the interface residues in cyan. (B) Pymol surface representation of the 1HEZ Fv fragment, containing the same color-code as applied for (A).

The obtained results can verify, at least for the V_L domain, that the observed point mutation effects are CDR-independent. Interestingly, the V_H/V_L interaction was severely decreased when exchanging the CDRs, the *vice versa* construct with exchanged framework exhibited an almost wild type like binding to V_H. This fits to the conclusion of a study from Honegger and coworkers, that examined the V_H domain stability, containing one specific framework scaffold within different CDR contexts, which states that an optimal framework is CDR-dependent [135].

In the context of the selected point mutations, it might be interesting to also transfer them to the MAK_1DH5 grafting construct. A similar approach for the V_H domain is also of interest especially since CDR H3 has shown to affect V_H/V_L orientation due to small position re-adjustments by the varying number of interface contributing residues [121].

Generally, the effect of the point mutations within a Fab fragment might be of interest, in order to exclude that structural rearrangements or the lack of additional interactions in absence of the adjacent antibody domains, C_L and C_H1, occur.

Within the course of investigating the impact of the CDRs on the variable domains, a CDR swap experiment between MAK33 V_H and V_L was conducted. This is a rather rough approach. But in this context it was interesting to see if this severe CDR exchange can be tolerated by the variable domains. The lengths of the exchanged CDRs differ greatly concerning CDR2 and CDR3. While CDR1 of the V_L domain (CDR-L1) is only amino acid longer than CDR-H1, the difference for CDR2 is the biggest one, with seven residues for CDR-L2 and fifteen for CDR-H2. But also CDR3-H3 with fourteen amino acids is five amino acids longer than CDR-L3. The constructs already showed a high aggregation propensity during the purification process resulting in very low yields. Structural characterization of the CDR grafting mutants indicated severe structural rearrangements compared to the wild type domains, including a helical structure for the V_H mutant and a misfolding behavior of the V_L mutant. Hence, CDRs execute a major effect on the variable domain integrity and grafting between V_H and V_L is not desirable as it implicates a loss of proper folding. This fits to studies of the isolated V_L domain, where they observed a change in stability associated with different CDR lengths and a complete exchange of the CDRs [188, 189]. In this context, it was proposed that the CDR loops might be important for the initial folding phases, since they bring distant regions of the domain in close proximity. Hence, the absence of distinct structural features of CDRs may promote the formation of a kinetically trapped intermediate [3]. This might explain the observed effects for the V_H/V_L CDR grafting constructs.

3.2. Stability improvement of the C_H2 domain

The studies of Dr. Eva-Maria Herold included the investigation of the alternatively folded state (AFS) of the C_H2 domain since this is the only antibody domain, for which the possibility to form this state had not been investigated [166]. It is a common feature of antibody domains that they form an AFS state at low pH [62-64, 108, 190]. The AFS is similar to the molten globule state with a high degree of secondary structure and a low degree of tertiary structure but differs in terms of stability. While the AFS possesses a high thermal and chemical stability, the molten globule state shows only a low stability [62, 191]. For her studies, Dr. Eva-Maria Herold included an unglycosylated and glycosylated C_H2 version expressed in *Sf9* cells. Unfortunately, the *Sf9* expression constructs had additional 16 amino acids distributed equally at the N-and C-terminus. The

N-terminal extra amino acids ADLGSA are the remainings of the secretion signal of the transfection vector pAcGP-67B. The additional C-terminal sequence GSGSENLYFQ, consists of a linker (GSGS) and leftovers (ENLYFQ) of the TEV protease cleavage site attached to the His_6-tag for purification. Alternative constructs with only one extra amino acid could not be purified successfully. The drawback of this C_H2 construct for expression in *Sf9* cells was the striking stability increase of ≈ 21 °C observed for thermal unfolding. This renders general comparisons with the less stable wild type domain rather complicated. In the end, Dr. Eva-Maria Herold characterized four different C_H2 constructs. The glycosylated and unglycosylated C_H2 with the mentioned additional amino acids from *Sf9* expression as well as the wildtype domain from *E.coli* expression and the control construct C_H2 ctr (Figure 22). The C_H2 ctr equals the amino acid sequence of the C_H2 from *Sf9* expression but based on its expression in *E.coli* it serves as a control, enabling a direct comparison with the C_H2 wild type domain.

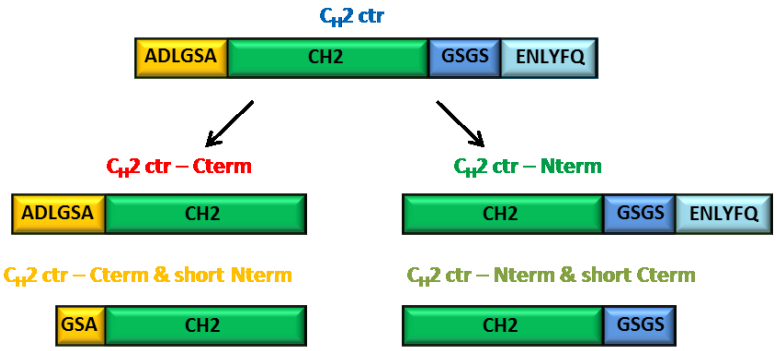


Figure 22 : Schematic overview of the four different C_H2 ctr stability constructs.
C_H2 (green) depicts the MAK33 wild type domain; the additional N- and C-terminal amino acids contained in C_H2 ctr are shown in detail. The same is applied for the four different C_H2 ctr N-teminal and C-terminal truncation constructs.

The following experiments focus on the high stability increase caused by the additional N- and C-terminal amino acids. As already mentioned, the C_H2 domain is rather unstable and aggregation prone in isolation, it can even provoke IgG aggregation. So a stability improvement is actually a desirable feature that can provide new application possibilities for single C_H2 domains in the context of therapeutic engineering [192]. But also an overall stability enhancement for the whole molecule antibody is of major interest. For revealing the key amino acids responsible for the observed stability improvement, four different constructs with increasing truncation of the additional amino acids, were expressed *in*

E.coli and characterized (see Figure 22). Due to the fact that the experiments executed with these constructs are not only interesting within the context of stability improvement of the C_H2 domain but also for the AFS studies of Eva Herold, all measurements were performed under physiological (50 mM sodium phosphate, 100 mM NaCl, pH 7.5) and AFS (50 mM sodium phosphate, 100 mM NaCl, pH 2.0) conditions and compared with the data of Eva Herold obtained for C_H2 wild type and C_H2 ctr.

In terms of stability improvement of the C_H2 domain different approaches have already been applied so far. One focused on the highly disordered N-terminal residues, where an N-terminal truncation of seven amino acids increased the thermal stability by 5 °C. This effect was accompanied by a low aggregation propensity and an elevated binding to FcRn [193, 194]. Another approach increased the thermal stability of the Fc part by 4.8 °C with the introduction of two point mutations within the N-glycosylation loop, leading to an additional interaction with the core fucose at the glycosylation site [195]. But also the introduction of an additional disulfide bridge could provoke a strong increase in thermal stability of 20 °C [196].

3.2.1 Structural characterization of the different C_H2 ctr mutants

In order to exclude structural changes associated with the shortened constructs secondary and tertiary structure was determined by CD spectroscopy and compared with C_H2 ctr and C_H2 wild type (Figure 23). The far-UV CD spectra of the four constructs at pH 7.5 and 2.0 (Figure 23 A and B) highly overlap with C_H2 ctr and C_H2 wild type. At pH 7.5 two minima at 230 nm and 210 nm as well as two maxima at 222 nm and 200 nm are visible, which represents a typical Far-UV spectrum for C_H2 in isolation [111]. The secondary structure undergoes severe changes at pH 2.0 showing one minimum at 210 nm and a maximum at 200 nm, indicating the presence of secondary structure elements since chemically denatured proteins usually show a random coil structure wit a minimum at 200 nm. Small deviations in signal intensity can be observed for the C-terminal truncations. Concerning the near-UV CD spectra a well-defined spectrum was determined for all C_H2 variants, indicating small tertiary structure rearrangments between C_H2 ctr and C_H2 wild type and also between the different truncation constructs (Figure 23 C and D). Although the N- and C-terminal pairs resemble each other quite well, they cannot be assigned to the wild type nor the C_H2 ctr spectrum, but seem to show a tertiary structure

somewhere in between. At pH 2.0, though, no signal is observed for almost all variants, indicating a complete loss of tertiary structure (Figure 23 C). The exception is C_H2 ctr–Cterm with a minimum at around 300 nm, assuming the presence of small amounts of tertiary structure.

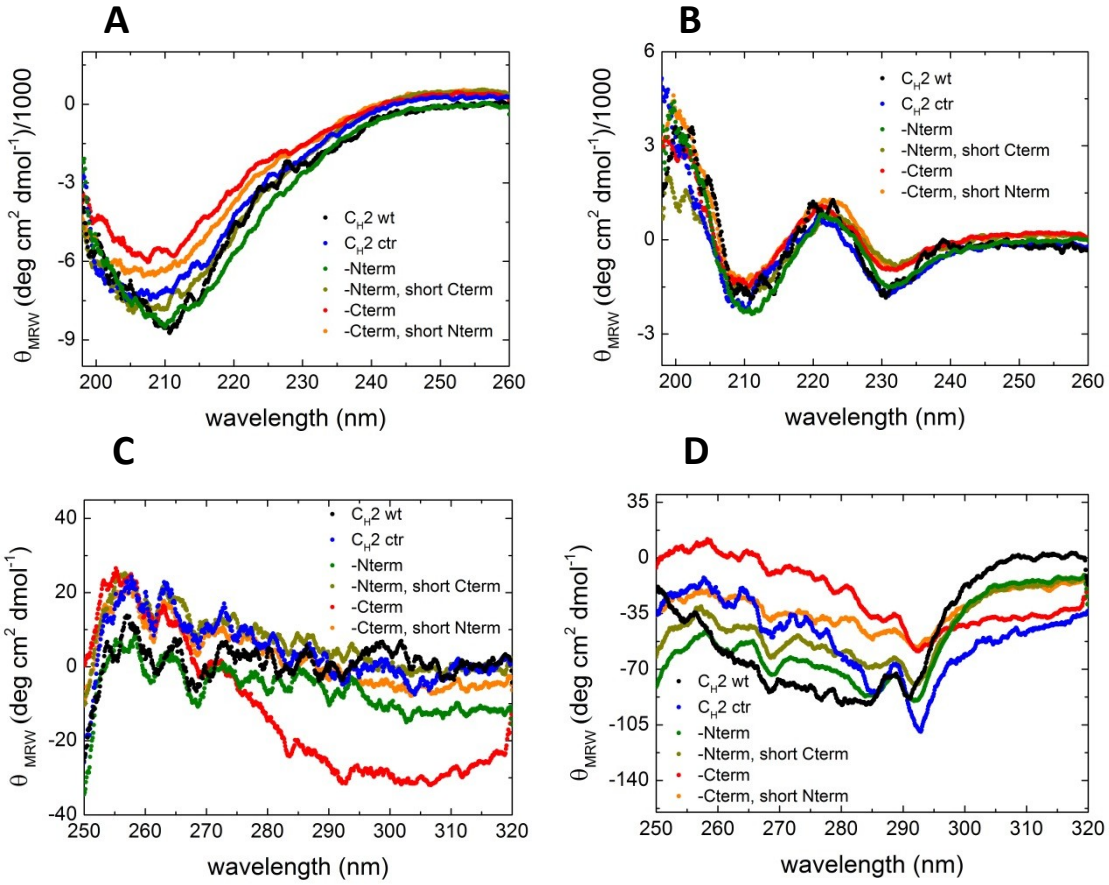


Figure 23 : Secondary and tertiary structure of the C_H2 ctr constructs under physiological and AFS conditions.

In (A) and (B), FUV-CD spectra of the C_H2 ctr constructs under AFS (A) and physiological (B) conditions are shown. (C) and (D) depict the NUV-CD spectra under AFS (C) and physiological (D) conditions. For the spectra 15 accumulations each were recorded and buffer-corrected. All measurements were performed at a protein concentration of 20 µM (FUV-CD) and 50 µM (NUV-CD) in 0.5 mm (FUV) or 5 mm (NUV) quartz cuvette at 20°C. The data for C_H2 wild type and C_H2 ctr were taken from Dr. Eva-Maria Herold's studies [166].

The hydrophobicity of the exposed surface was measured by ANS fluorescence. Under physiological conditions no change in ANS fluorescence was visible, corresponding to the absence of surface-exposed hydrophobic regions. Under AFS conditions, a strong increase in ANS fluorescence associated with the binding of ANS to exposed hydrophobic regions occurs (Figure 24). Thus under AFS conditions the secondary structure changes and the loss of tertiary structure is accompanied by a strong increase of surface-exposed

hydrophobic regions. Between C_H2 wild type and the different C_H2 ctr constructs no difference in the ANS spectra was observed (C_H2 wt and C_H2 ctr data not shown).

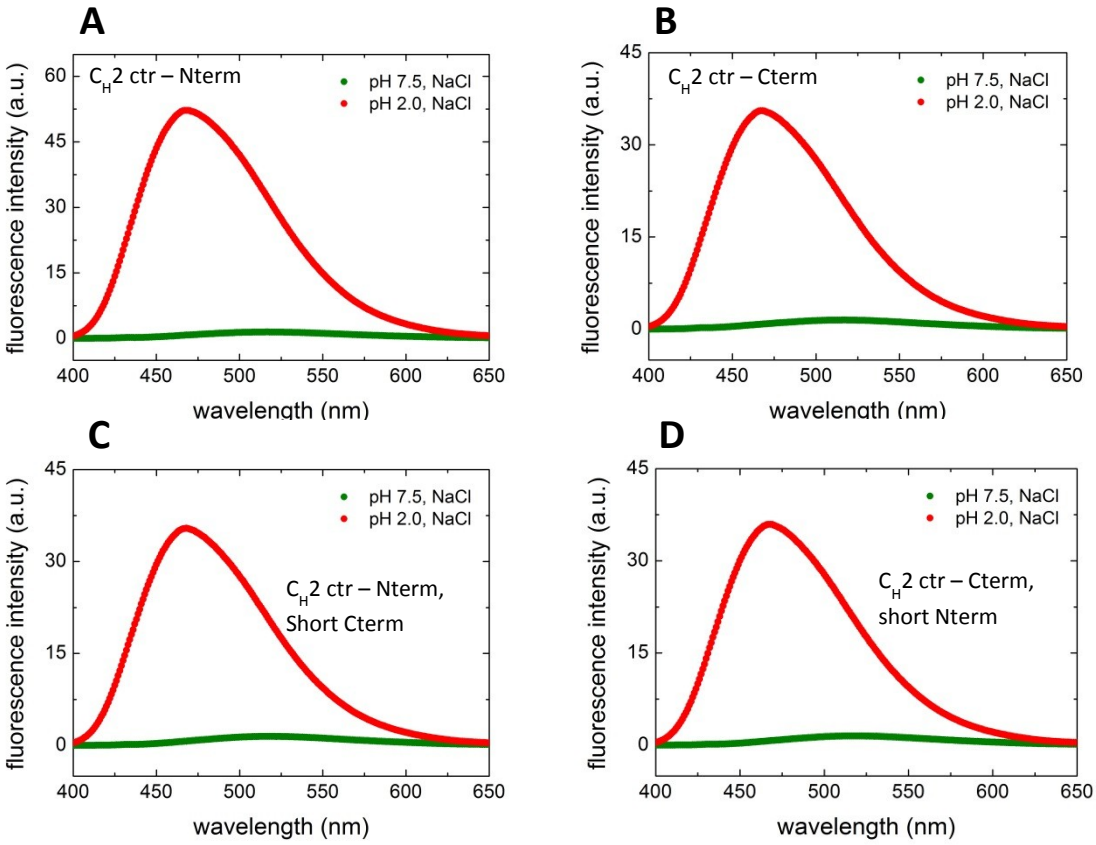


Figure 24 : Determination of the surface hydrophobicity by ANS binding.
To analyze the change of surface hydrophobicity of the C_H2 ctr constructs (10 µM protein), the complex formation in presence of 100 µM ANS under physiological (red) and AFS (green) conditions was monitored at an excitation wavelength of 380 nm by fluorescence spectroscopy. Measurements were performed at 20 °C.

3.2.2. The influence of the different C_H2 ctr constructs on stability

The determination of the thermal and chemical stability (Figure 25, Table 6) showed a very interesting picture. While the already observed stability difference of 21 °C and 2.9 M Urea between C_H2 wild type (T_m: 44.5 ± 0.2 °C; $D_1/_2$: 2.5 M ± 0.1) and C_H2 ctr (T_m: 65.6 ± 0.2 °C; $D_1/_2$: 5.4 M ± 0.4) is very prominent, it becomes more difficult turning to the C_H2 truncation constructs. Under physiological conditions, the constructs with removal of either the C-terminal or N-terminal amino acids including their further shortened versions (C_H2 ctr-Cterm & short Nterm; C_H2 ctr-Nterm & short Cterm) exhibited a very similar stability with T_ms between 57.0 – 59.3 °C and $D_1/_2$ values between 5.1 – 5.2 M for the –Nterm constructs and T_ms between 53.3 – 55.8 °C and a $D_1/_2$ value of 4.3 M for the –

Cterm constructs. So the C_H2 ctr-Nterm constructs are slightly more stable than the C_H2 ctr-Cterm constructs. Compared to the unstable wild type C_H2 and the very stable C_H2 ctr, the stabilities of the truncation constructs lie between the two with a bias towards C_H2 ctr according to the thermal unfolding data. For the chemical unfolding, the truncation construct values are very close to C_H2 ctr especially those for the C_H2 ctr–Nterm constructs. So surprisingly, the additional amino acids at both termini, especially at the C-terminus, contribute to the observed stability increase of C_H2 ctr. Furthermore, with high degree of overlap for the stability between C_H2 ctr–Cterm and C_H2 ctr–Cterm & short Nterm as well as C_H2 ctr-Nterm and C_H2 ctr–Nterm & short Cterm, a striking effect of stability improvement can be traced back to the amino acids GSA at the N-terminus and GSGS at the C-terminus.

Table 6 : Stability values of the C_H2 ctr constructs under physiological and AFS conditions.

Domain	T_{melt} [°C] pH 7.5	T_{melt} [°C] pH 2	$D_{1/2}$ [M] pH 7.5	$D_{1/2}$ [M] pH 2
C_H2 ctr	*65.6 ± 0.2*	*44.0 ± 0.5*	*5.4 ± 0.4*	*1.2 ± 0.6*
C_H2 –Nterm AS	59.3 ± 0.2	59.3 ± 0.3	5.2 ± 0.1	2.7 ± 0.1
C_H2 –Nterm AS, short Cterm	57.0 ± 0.2	25.9 ± 1.0	5.1 ± 0.1	0.6 ± 0.1
C_H2 –Cterm AS	55.8 ± 0.2	56.9 ± 1.8	4.3 ± 0.2	2.4 ± 0.4
C_H2 –Cterm AS, short Nterm	53.3 ± 0.2	75.9 ± 1.4	4.3 ± 0.1	3.7 ± 0.1
C_H2 wild type	*44.5 ± 0.2*	*41.2 ± 0.7*	*2.5 ± 0.1*	*1.6 ± 0.1*

Data evaluation of the GdmCl-induced unfolding transitions according to a two-state equilibrium unfolding model was conducted to derive the midpoint of transitions ($D_{1/2}$), as well as the cooperativity parameter (m), for a qualitative comparison of the data. Thermal unfolding was measured by CD spectroscopy at a constant wavelength of 215 nm in a 0.5 mm quartz cuvette with a heating rate of 20 °C h^{-1}. The midpoint of transitions (T_{melt}) was determined by Boltzmann fit. The data for C_H2 wild type and C_H2 ctr were adapted from Dr Eva-Maria Herold's studies [166].

Under AFS conditions, though, the stability data differ strongly (Table 6). The stability distribution is the same for thermal and chemical unfolding. C_H2 wild type and C_H2 ctr show, in contrast to the data under physiological conditions, a very similar stability with a T_m of 41.2 ± 0.7 and 44.0 ± 0.5 and $D_{1/2}$ values of 1.6 ± 0.1 and 1.2 ± 0.6, respectively. Interestingly, the C_H2 ctr truncation versions, -Cterm and –Nterm exhibit a higher stability with –Nterm being slightly more stable. With further reduction of the number of additional amino acids at the termini, stability changes in both directions occur. When the

C-terminal amino acids are further shortened to GSGS, the stability is drastically decreased to a T_m of 25.9 ± 0.6 and a $D_1/_2$ of 0.6 ± 0.1, which is by far the lowest stability observed. On the opposite site, when the N-terminal amino acids are further shortened to GSA an additional increase in stability is achieved with a T_m of 75.9 ± 0.6 and a $D_1/_2$ of ± 3.7 ± 0.1.

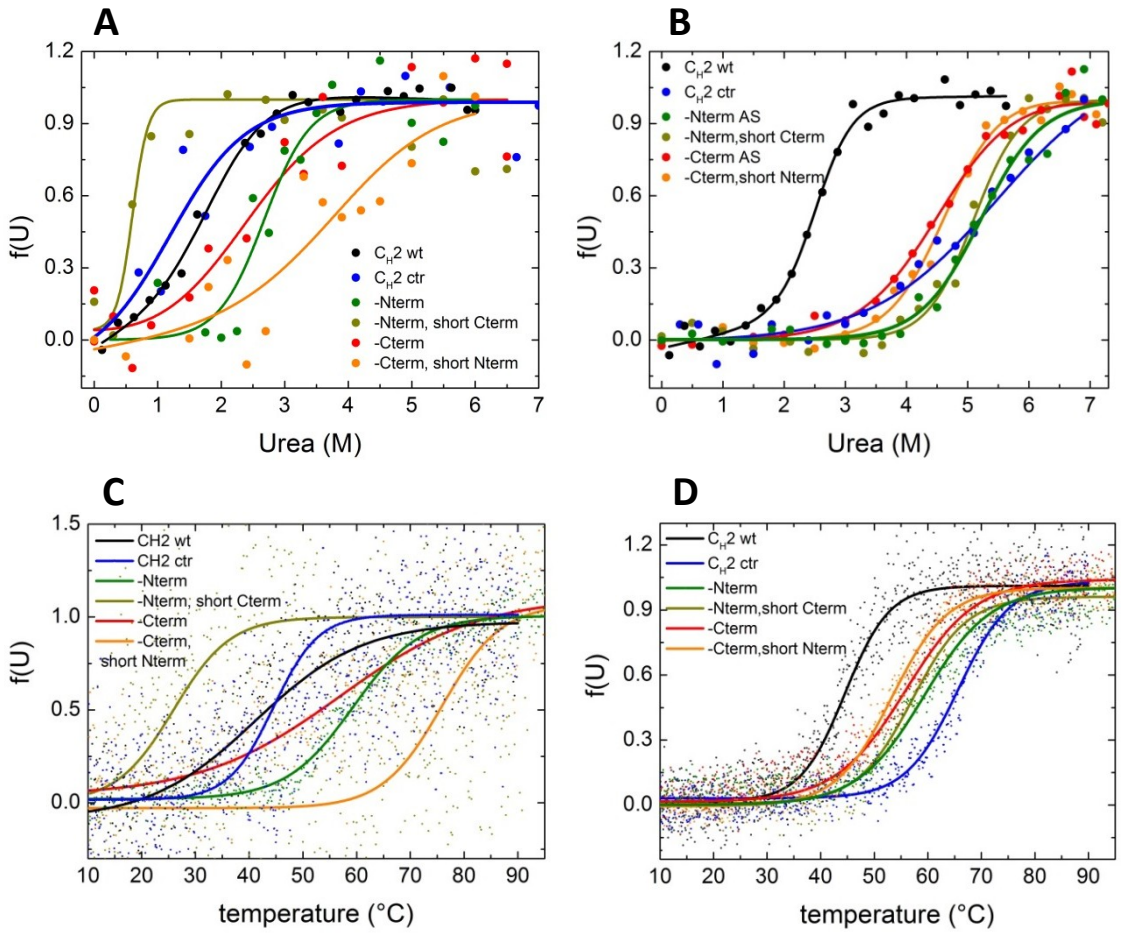


Figure 25: Influence of the additional N- and C-terminal amino acids on C_H2 stability under physiological and AFS conditions.

The stability of the C_H2 constructs under AFS (A and C) and physiological conditions (B and D) against thermal denaturation (C and D) was assessed by CD spectroscopy at a constant wavelength of 215 nm at pH 7.5 and 205 nm for pH 2.0 and a heating rate of 20 °C h^{-1}. Chemical denaturation (A and B) with GdmCl was determined by fluorescence spectroscopy, measuring the change of the intrinsic tryptophan fluorescence upon unfolding at an excitation wavelength of 280 nm. The data for C_H2 wild type and C_H2 ctr were taken from Dr. Eva-Maria Herold's studies [166].

The controversial stability data under AFS conditions fits quite well to the size distribution data measured by AUC (Figure 26, Table 6). All C_H2 variants form a monomer under physiological conditions. Under AFS conditions, C_H2 wild type shows an equilibrium between monomer and dimer. All other C_H2 ctr constructs show different distribution

patterns with a monomer and an oligomer fraction with a broad size distribution of 25 – 150 kDa, corresponding to dimers up to decamers (Figure 26).

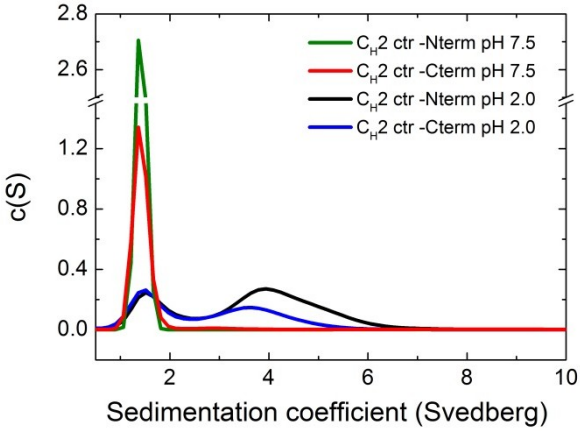


Figure 26: C(S) distribution of the AUC data under AFS and physiological conditions for two C_H2 ctr constructs.
The c(S) distribution of the sedimentation velocity run data for C_H2 ctr – Nterm and – Cterm is shown. Measurements were performed with a protein concentration of 5 - 10 µM and absorbance optics set to 230 nm. Data evaluation was performed with the program Sedfit (Peter Schuck) using a c(S) model with time and radially invariant noise on.

In general, the appearance of higher order oligomers under AFS conditions has already been observed for other antibody domains (V_H, V_L, C_H1) [64, 108]. But in these cases no monomer fraction has been measured. The observed amount of oligomers for C_H2 decreases in the following order of C_H2 ctr constructs, starting with an amount of 76 % and ending with 35 % oligomers: -Nterm, short Cterm (76 %); -Nterm (70 %); -Cterm, short Nterm (63 %); C_H2 ctr (47 %); -Cterm (35 %).

The observed increase in oligomer formation does not correlate with any other data obtained under AFS conditions. Even though the construct with the highest oligomer amount coincides with the lowest stability, the remaining oligomer data shows no correlation with stability. If only the C_H2 ctr AUC results are analyzed, an increase in oligomer formation correlates with the removal of the N-terminal extra amino acids. The amount of detected oligomers for the C_H2 –Nterm construct is 70 %, as soon as the C-terminus is additionally shortened, the oligomers even increase to 76 %. In this case, the very low stability might also contribute to the higher propensity of oligomer formation. By contrast, the removal of the C-terminal residues causes a reduction to 35 % oligomers compared to 47 % oligomers for C_H2 ctr. This observed decrease vanishes when the N-terminus is shortened, leading to an oligomer proportion of 63 %. Nevertheless when including the wild type domain, the conclusions from the C_H2 ctr variants cannot be transferred to the wild type data. Although it does not contain the N-terminal extra amino acids, C_H2 wild type does not form any higher oligomers.

Table 7: Size distribution of the C_H2 ctr truncation constructs.

Domain	Buffer conditions	Monomer [%]	25-150 kDa [%]
C_H2 wild type	pH 7.5, 100 mM NaCl	*100*	
	pH 2.0, 100 mM NaCl	Monomer-dimer equilibrium	
C_H2 ctr	pH 7.5, 100 mM NaCl	*100*	
	pH 2.0, 100 mM NaCl	*~ 53*	*~ 47*
C_H2 ctr - Nterm	pH 7.5, 100 mM NaCl	100	
	pH 2.0, 100 mM NaCl	~ 30	~ 70
C_H2 ctr – Nterm,	pH 7.5, 100 mM NaCl	100	
short Cterm	pH 2.0, 100 mM NaCl	~ 23	~ 76
C_H2 ctr – Cterm	pH 7.5, 100 mM NaCl	100	
	pH 2.0, 100 mM NaCl	~ 65	~ 35
C_H2 ctr – Cterm,	pH 7.5, 100 mM NaCl	100	
short Nterm	pH 2.0, 100 mM NaCl	~ 37	~ 63

The size of the different C_H2 ctr constructs was determined by AUC sedimentation velocity runs at a protein concentration of 5 - 10 µM. Data evaluation was performed with the program Sedfit (Peter Schuck) using a c(S) model. The data in grey and italic mark the data adapted from Dr. Eva-Maria Herold's studies[166].

3.2.3. Discussion

Altogether, under physiological conditions the C-terminal and N-terminal C_H2 ctr truncation pairs behave similar. The structural data indicate tertiary structure deviations between C_H2 wild type, C_H2 ctr and the two C_H2 truncation pairs (-Nterm; -Cterm), which might be responsible for the observed effects. The striking increase in stability can be traced back to the N-terminal residues GSA and the C-terminal residues GSGS, with a slightly higher contribution of the C-terminal extra amino acids. In order to identify the exact amino acids responsible for this stabillity effect, further amino acid reductions on both termini down to only one extra amino acid seem reasonable. As the residues GS depict the consensus of both termini one can speculate whether these are responsible for the stability improvement. In this regard it would be interesting if the higher stability of the C_H2 ctr–Nterm compared to C_H2 ctr-Cterm constructs can be associated with the GSGS pair instead of only one GS. For further insight, additional constructs have to be produced, including only GS at each termini as well as different numbers of GS repetitions at each terminus. Also a combination of the stabilizing amino acids on both termini might

be interesting in order to examine an additive effect. In a more general context, it might be interesting to investigate the effect of the identified stabilizing amino acids to other unstable antibody domains like V_H.

With regard to the whole antibody, it is of interest whether the monitored stability effects by addition of selected N- and C-terminal amino acids are artifically generated for the C_H2 domain in isolation or can also be found in nature. Hence, Figure 27 shows a schematic overview of the the natural amino acids adjacent to the C_H2 MAK33 domain. These depict the first amino acids of the N-terminal hinge region and the C-terminal C_H3 domain.

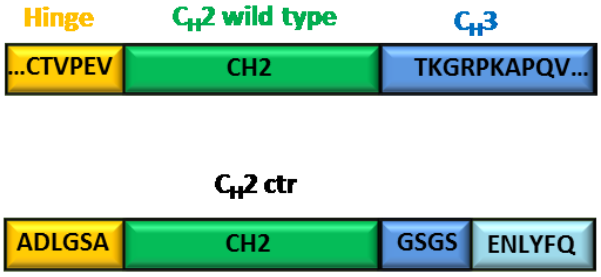


Figure 27: Schematic comparison of C_H2 ctr with the naturally occuring amino acids adjacent to C_H2 in the MAK33 heavy chain.

C_H2 (green) depicts the MAK33 wild type domain; Adjacent to the wild type domain the first natural amino acids present in the MAK33 heavy chain (the N-terminal hinge region and the C-terminal C_H3 domain) are shown in detail. The additional N- and C-terminal amino acids representing the C_H2 ctr construct are also depicted in detail.

The comparison with the natural amino acids exhibits no sequence consensus with the C_H2 ctr construct. With regard to C_H2 stability in the whole antibody, additional C_H2 constructs with additional N- and C-terminal natural amino acids might be of interest. With this approach it cannot only be evaluated whether the stability enhancement can be provoked by any amino acid added to the C- or N-terminus but also a general conclusion may be possible, concerning the stability of C_H2 within a completetly assembled antibody. Figure 28 shows the crystal structure of an IgG C_H2 dimer involving the first ten residues of the adjacent domains. Here, a human IgG1 structure was used for the illustration due to the fact, that there is no MAK33 crsytal structure of a whole antibody available, only of the Fab or Fc fragment.

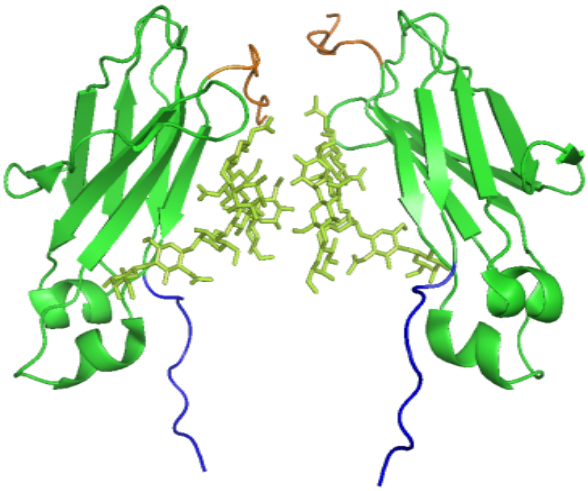


Figure 28: Structure of a glycosylated human IgG1 C_H2 dimer including the first residues of the adjacent domains.
Pymol cartoon representation of the glycosylated human 1HZH C_H2 dimer. The glycosylations are illustrated in lime green stick representation. The first ten amino acids from the adjacent domains involving the linker residues are also shown. The N-terminal residues of the hinge region are highlighted in orange, the C-terminal residues of the C_H3 domain in blue.

In studies with the MAK33 light chain, it could be shown that arginine residue 108 within the linker region between the V_L and C_L domain has a stabilizing effect on both domains [197]. For the V_L domain the exposure of the charged arginine residue increased solvation, which was accompanied by an overall stability increase. Concerning the C_L domain, the arginine residue interacts with the N-terminal loop residues, which is important for the domain integrity. In absence of the arginine residue, strutural fluctuations occur that lead to partial exposure of hydrophobic core residues. These stability effects associated with distinct linker residues, that connect the single antibody domains, might also be transferable to the C_H2 domain.

Under AFS conditions, all C_H2 ctr constructs as well as the wildtype exhibit a drastic change of secondary structure and an almost complete loss of tertiary structure, only C_H2 ctr – Cterm still shows a small amount of tertiary structure. These structural changes are accompanied by a strong increase of surface-exposed hydrophobic regions. Concerning the stability under AFS conditions, the wild type and C_H2 ctr exhibit similar stabilities while the removal of either the C-terminal or N-terminal extra amino acids leads to a stability improvement. So under AFS conditions extra amino acids on both termini are associated with a wild type like stability while additional amino acids on only one terminus enhances the stability. An additional reduction of the N-terminus to GSA results in a further stability increase of 19 °C. *Vice versa*, the shortening of the C-terminus to GSGS leads to a drastic loss of stability of 33 °C. This is rather unexpected, considering the fact that the sequential difference between these constructs and the wild type domain, causing T_{melt} values that differ from the wild type domain by ≈ 25 °C in the neagtive (-

Nterm, short Cterm) and positive direction (-Cterm, short Nterm), consists of GSA (N-terminal) or GSGS (C-terminal). In contrast to physiological conditions, the two C_H2 ctr truncation variant pairs do not behave the same. This implies a contribution of the N-terminal amino acids GSA in the stability increase and of the C-terminal amino acids GSGS in the stability decrease. While further addition of amino acids on both ends can significantly modify the stability towards a very similar stability value for C_H2 ctr – Nterm and C-term. The oligomerization state under AFS conditions is subject of great variations for all analysed samples. Other antibody domains like V_L and V_H form 100 % oligomers with a broad distribution in size under AFS conditions [108]. Interestingly, for C_H2, the wildtype domain under AFS conditions exhibits a monomer/dimer equlibrium and all the C_H2 ctr constructs possess a monomer and an oligomer fraction, ranging from 35 – 76 %. No correlation between surface hydrophobicity (ANS binding) and oligomerization could be observed, which has been the case for the Fab fragment under AFS conditions [108]. The same holds true for stability and oligomerisation even if the highest amount of oligomerisation corresponds to the construct with the least stability. The other constructs do not behave accordingly. For the C_H2 ctr –Cterm variant, which represents the only variant with little amounts of tertiary structure under AFS conditions, the lowest amount of oligomerization was observed. So the measuresd oligomerization differences might be associated with structural differences and in small amounts also with stability. It has to be kept in mind that the stability values from thermal and chemical unfolding are subject to a high degree of noise signal. Since most of the structural assignments vanish under AFS conditions, the difference to the completetly unfolded state is comparatively low.

3.3. C_H1 domain of immunoglobulins

3.3.1. The camelid C_H1 domain

In camelid and llama a completely different kind of antibody evolved, the so-called heavy chain antibody (HCAb), which comprises about 50 % of the IgGs found in the serum [198, 199]. The special feature of HCAbs is the absence of the light chain and the C_H1 domain of the heavy chain (see Figure 29).

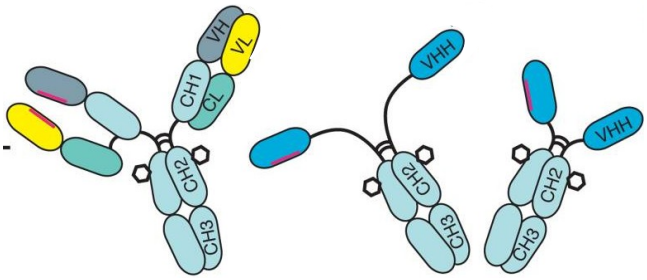


Figure 29: Structural composition of camelid heavy chain antibodies in comparison to conventional IgGs.
Left: a conventional IgG antibody; right: two variations of camelid heavy chain antibodies with different lengths of the hinge regions. The different domains contained in the corresponding antibody are labeled; the hexagon indicates a glycosylation site. The figure is taken from [1].

This unique form of IgGs is of special interest for therapeutic engineering in the context of the V_{HH} domains [2]. Single domain antibody domains have several advantages including an increased accessibility for epitopes, a more efficient penetration and fast clearance. Without the associated Fc fragment, sometimes undesired Fc-mediated effector functions can be abolished. In contrast to classic V_H domains, the V_{HH} is very stable in isolation [2]. This is due to some adaptations that evolved in the absence of the light chain, involving amino acid exchanges in the framework 2 region and for some V_{HH}s, an additional disulfide bridge between CDR1 and CDR3 and [200-203]. These turn the otherwise hydrophobic V_H/V_L interface region into a hydrophilic one. In order to overcome the increase in diversity enabled by different V_H/V_L epitope combinations, V_{HH} shows extended CDRs, especially for CDR3 with higher diversity beyond the limitations of the observed canonical CDR structures for V_H [175, 204]. Additional changes like an elongated hinge region which is associated with a genomic transposon insertion or an additional CDR, were also detected for some camel heavy chain IgG2a antibodies [198, 200].

Concerning the absent C_H1 domain, Nguyen and coworkers [205] determined the complete *camelus dromedarius* IgG ɣ2a heavy chain constant gene, derived from a liver genomic library and compared it to the mammalian ɣ2a heavy chain constant gene. On the genetic level different exons encode each Ig domain as well as the hinge and transmembrane region [206]. The sequences for the light and heavy chain lie on different chromosomes [207]. All the ɣ genes present in mammalians including the switch, the hinge region, C_H1, C_H2, C_H3, M1 and M2 exons could be identified in camel [205]. The sequence comparison of the camelid Cɣ2a gene with mammalian sequences showed the highest identity for bovine Cɣ1 and Cɣ2 (74 % and 70 %) followed by all human Cɣs (70 %)[205]. For the C_H1 domain only, a sequence identity of 70 – 74 % is observed for all mammalian Cɣs. So the silenced C_H1 domain exhibits fewer variations from the

mammalian genome than C_H2 and C_H3. Even though the C_H1 exon was intact, sequencing revealed point mutations in the adjacent splicing site, especially the G^{+1} to A^{+1} transversion seems to be of importance. This loss of consensus splicing signal leads to the removal of the C_H1 mRNA during RNA processing, resulting in the complete absence of the C_H1 domain on protein level in camel ɣ2a heavy chain immunoglobulins. Besides, there is a single G to C mutation of nucleotide 566 which leads to point mutation of one of the two conserved cysteine residues (Cys208Ser) in C_H1. But this is assumed to have succeeded the splice point mutations. In humans and mice heavy chain antibodies only exist associated with diseases. In this cases, partial or complete V_H or C_H1 deletions occur, which can also be found on the genetic level [208-210].

In context of the camel heavy chain antibodies there are two possible timeline scenarios for the C_H1 depletion. Either the heavy chain antibodies developed subsequently to the C_H1 removal. In the absence of the unfolded C_H1 domain, the associated quality control mechanism and ER retention are deactivated. Hence, the presence of the light chain with the C_L domain as the corresponding C_H1 interaction partner is no longer necessary. The second scenario involves a C_H1 removal after appearance of the heavy chain antibodies. This requires a C_H1 domain that is able to fold in the absence of the C_L domain. In this case the camel C_H1 domain would depict the only verified naturally occuring C_H1 domain which is able to fold without C_L interaction.

The question whether the camel C_H1 is able to fold in the absence of C_L is addressed in the following study where the silenced camel C_H1 domain is expressed in *E.coli* and characterized in terms of folding behavior. The only difference from the original sequence depicted the back mutation of the conserved cysteine (S208C).

The expression of soluble antibody domains in *E.coli* is not possible due to the reducing environment of the cytoplasm, instead they form insoluble inclusion bodies (IBs)[211]. First expression attempts with tag-less and His_6-tag C_H1 (codon optimized for *E.coli*) showed no protein expression in BL21 (DE3) cells. Hence, the next expression conditions were chosen within the context of solubility improvement, involving a change of the *E.coli* strain to engineered shuffle cells. These cells express DsbC, an oxidoreductase chaperone, which promotes disulfide bond formation in the cytoplasm [211]. An additional change concerned the expression vector. Here, pET SUMO was chosen, which adds an N-terminal

His_6- and a SUMO-tag to the protein of interest [212]. The 11kDa Sumo tag stabilizes associated proteins and facilitates their expression by an increased solubility. The applied pET SUMO vector includes a SUMO protease cleavage site, which allows the removal of the whole tag after successful purification. Under these modified conditions an expression of the camel C_H1 domain could be observed in the soluble and insoluble fraction (data not shown). Both were further used for purification. Inclusion bodies were dissolved and purified under denaturing conditions (4 M GdmCl) from the insoluble fraction; the C_H1 from the cytoplasm was purified under native conditions. The results of the characterization of the purified proteins are shown in Figure 30. Secondary structure was determined by far-UV CD from the two different purifications, with and without tag. Under denaturing conditions, though, only the tag-less C_H1 could be measured since SUMO cleavage was performed before refolding. The FUV spectra of all variants show a minimum at 208 nm and a maximum at 200 nm. This resembles rather the FUV spectrum of the unfolded murine MAK33 C_H1 than a typical folded Ig domain with ß-sheet structure (Figure 30). The only deviations between the C_H1 far-UV CD spectra from camel and mouse are the different signal intensity and a minimum shift to higher wavelengths for the camel C_H1. For the subsequent measurements only C_H1 without tag from the purification under denaturing conditions was applied. Thermal unfolding and fluorescence spectra under native, reduced and denatured conditions confirmed that the camel C_H1 possesses little amounts of structure since there are measurable differences (signal change for thermal unfolding; fluorescence red shift for the chemical denatured protein) detectable for C_H1 under native and denaturing conditions. Interestingly, no difference between C_H1 with and without SUMO tag was visible in the FUV, so a structural stabilization by the tag cannot be verified.

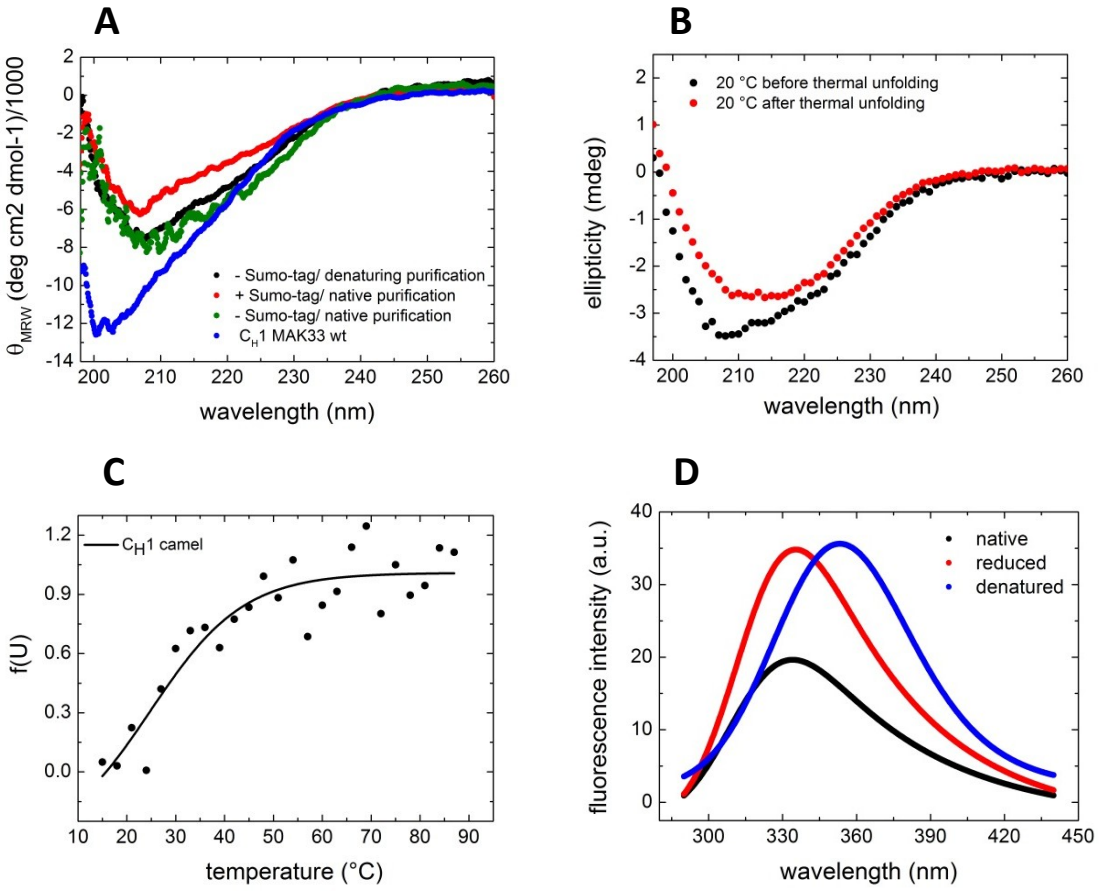


Figure 30: Structural and stability characterization of the camelid C_H1 domain

(A) and (B) show FUV-CD spectra of the camel C_H1 domain (A) from different purifications with and without SUMO-tag as well as a comparison (B) of the FUV spectra of tagless C_H1 before and after thermal unfolding. (C) depicts the thermal unfolding, fitted by Boltzmann fit. In (D) the instrinsic tryptophan fluorescence under native (PBS), reducing (PBS, 1mM TCEP) and denaturing (7 M GdmCl) conditions is shown.

To examine whether proper folding of the camel C_H1 domain can be induced by addition of C_L, kinetic experiments at a constant wavelength of 205 nm were performed (Figure 31). The addition of C_L from murine and human origin did not lead to a signal increase associated with C_H1 folding. This might be due to a high specifity of camel C_H1 to camel C_L. So in this context the result for the addition a camelid C_L domain would be interesting.

The oligomeric state of all purified versions was determined by AUC (Figure 32). The size distribution showed no monomer fraction; only higher oligomers or aggregates with a broad size variation, ranging from 70 kDa to 1MDa were monitored.

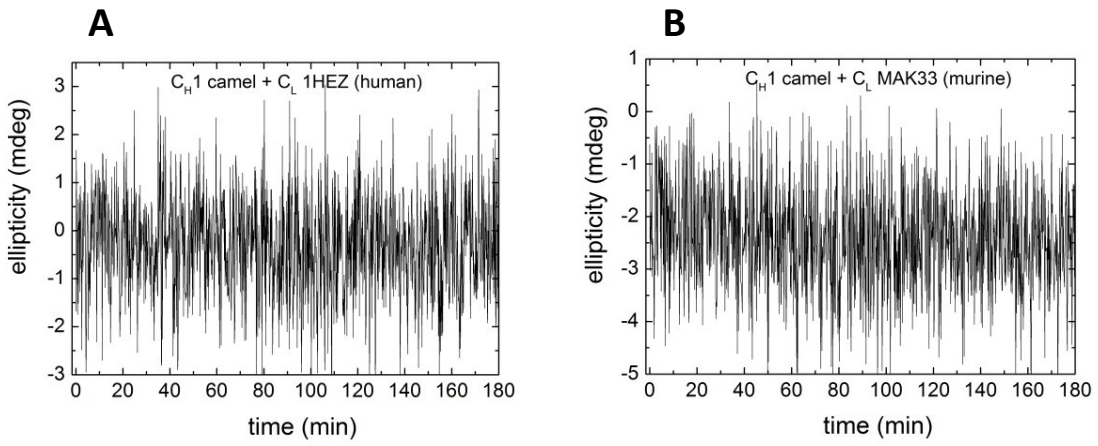


Figure 31: Kinetics of camel C_H1 upon addition of different C_L domains

Kinetic measurements by CD spectroscopy of camel C_H1 upon addtion of (A) human C_L (1HEZ) and (B) murine C_L (MAK33) at a constant wavelength of 205 nm. All measurements were performed at 25 °C and protein concentrations of 10 µM each in a 1 mm quartz cuvette.

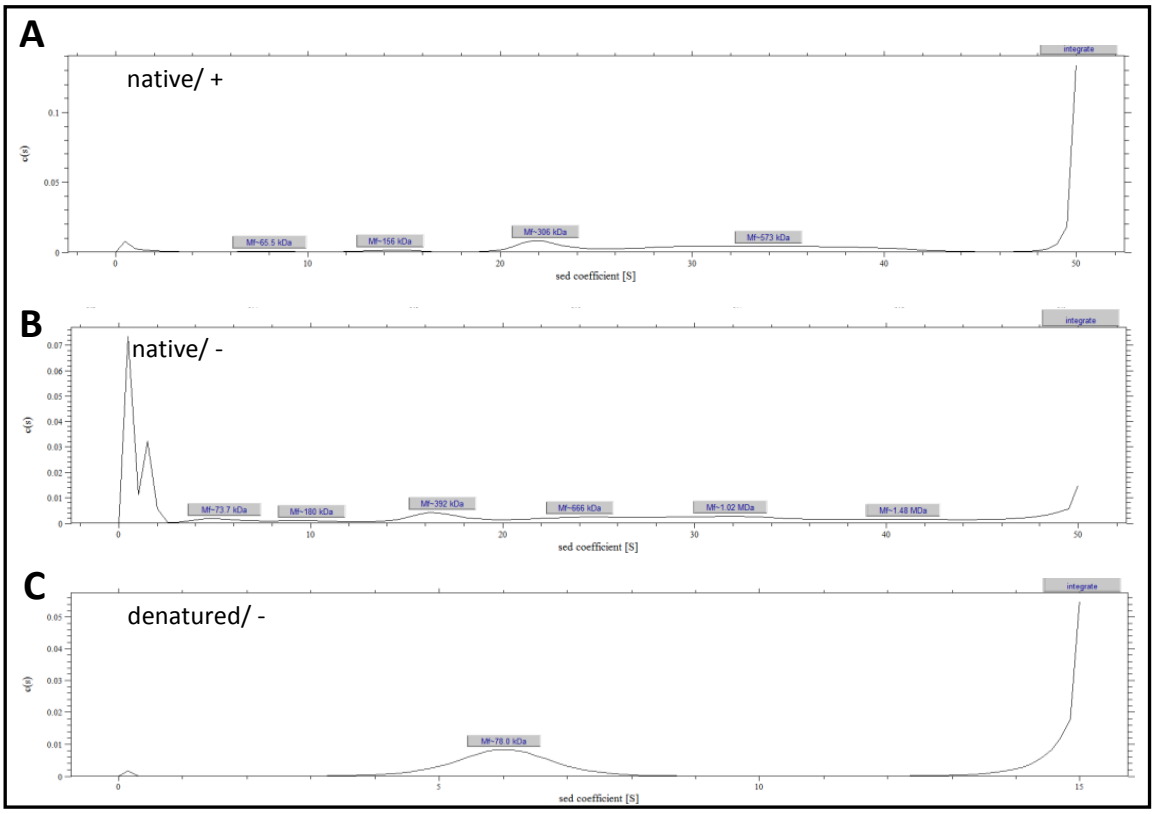


Figure 32: Size distribution of C_H1 camel from different purifications

Screenshots from AUC sedimentation velocity run evaluation with the program Sedfit (Peter Schuck) using a c(S) model. (A) and (B) show the data from C_H1 from native purification (A) with SUMO-tag and (B) without SUMO-tag as well as C_H1 from the denaturing purification without SUMO-tag(C).

The data obtained within this study indicates that C_H1 removal was indeed the prerequisite for the camel heavy chain antibody development. In isolation C_H1 from camel exhibits only slight features of a folded protein. In general, the results resemble the data obtained for intrinsically disordered murine (MAK33) C_H1 domain in isolation. With regard to the high sequence similarity (see Figure 33) this is not surprising.

```
CH1 camel: MASTKAPSVYPLTARSGDTPGSTVAFGCLVWGYIPEPVTVTWNSGAVSSGIHTFPSVLMS
CH1 MAK33: --MTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQS
             *..******:. *.  ..* *::**** **:***********::***:****:** *

LGLYSLSSLVTLPTSTSTGKTFICNVAHPASSTKVDKSVDK-
D-LYTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPR
  **:*** **:*:** .::*. **************.:
```

Figure 33: Sequence alignment of the murine (MAK33) and camelid C_H1 domain
The sequence alignment was performed with CLUSTALW, cyan highlights and asterisks indicate a sequence consensus.

But in contrast to the murine and human C_H1 domains, camel C_H1 exhibits a pronounced aggregation propensity with no presence of monomeric protein.

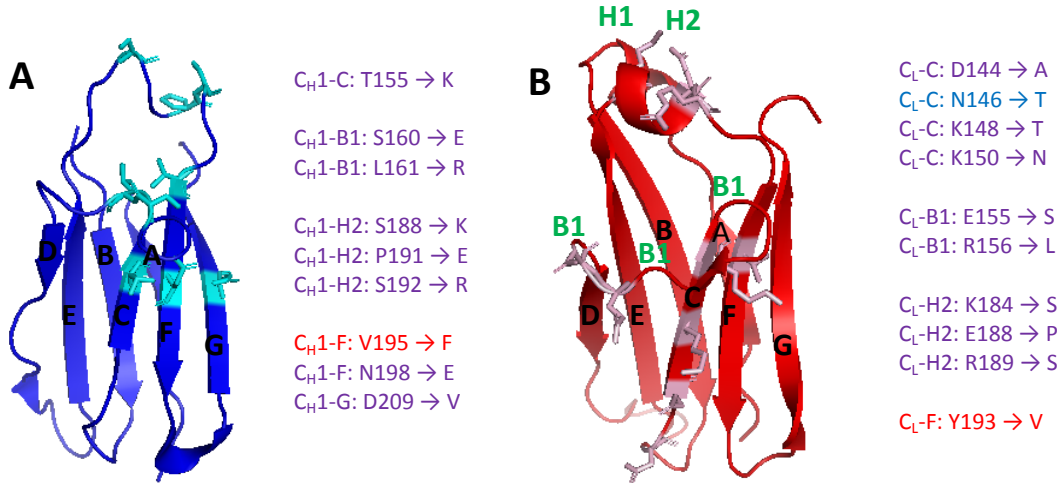
3.3.2. Investigation of potential key amino acid sequences involved in the C_H1 unfolding behavior in isolation

In the following study two different approaches were tested in order to generate a C_H1 domain that is able to fold in absence of the C_L domain. Here the focus was set on structural regions and amino acid patterns that are highly divergent not only between C_H1 and C_L but also between C_H1 and the other constant domains (C_H2, C_H3).

3.3.2.1 C_H1/ C_L swap experiments

The first approach for generating a foldable C_H1 domain is based on the bioinformatics analysis of George Stan and coworkers (University of Cincinnati). They performed coarse-grain simulations of folding and mechanical unfolding of the isolated C_H1 and C_L domains as well as simulations that probe the energy landscape using an implicit solvent model including atomic level details and electrostatic interactions. This comparison identified a high divergence between the two domains for the helical regions of C_L (H1 and H2) and a distinct region between strand C and D (B1) (see Figure 34) as well as in charged residues. Further structural comparisons with the remaining constant domains (C_H2, C_H3) resulted in a high consensus with the C_L domain concerning the identified regions (H1, H2, B1) that are different from C_H1. In the following *in vitro* study the obtained data was used to

generate a C_H1 (C_H1 GS) and a C_L mutant (C_L GS) that contain the calculated amino acid exchanges (Figure 34). So these mutants potentially possess the structural features of the opposite domain. The introduced amino acid exchanges (see Figure 34) in order to form H1, H2 and B1 in C_H1 include the following modifications: 1) addition of an aromatic contact between H2 and F, 2) increase of the number of charged-charged contacts, 3) formation of two salt-bridges inside helix H2, 4) formation of a salt-bridge between strand C and F, 5) increase of the number of H-bonds between strand C and F. The opposite effect was expected for the C_L domain.




```
CH1 MAK33: TTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNS--GSLSSGVHTFPAVLQS
CL  MAK33:  AAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDS
            :.*:*  :.*.*   *.. .:: *::::::*: :.*.*:   .. ..** .  :  :*

--DLYTLSSSVTVPSSTWP-SETVTCNVAHPASSTKVDKKIVPR-
KDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNE
  . *::**::*:... :   :: **:.:* :*:: : *.:  .
```

Figure 34: Amino acid exchanges for the generation of the mutants C_H1 GS and C_L GS
Pymol cartoon representation of MAK33 (A) C_H1 in blue with the mutated amino acids highlighted in cyan and (B) C_L in red with the mutated amino acids highlighted in light pink. Additionally the strand numberings (A, B, C, D, E, F, G) in black and regions that are structurally different to C_H1, containing helix 1 (H1), helix 2 (H2) and B1 (between strand C and D) in green are shown. A list with the amino acid exchanges and the corresponding strand or helix locations is depicted adjacent to the two domains. The particular exchanges are color-coded according to their involvement in structural contacts: aromatic contacts, charged-charged contacts, hydrogen bonds. Underneath, a sequence alignment between the two domains performed with CLUSTALW is shown, the mutation sites are again color-highlighted.

The far-UV CD spectra show the same spectra shape for C_H1 wild type, C_H1 GS and C_L GS with a minimum around 202 nm, while C_L exhibits a typical ß-sheet spectrum with a minimum at 218 nm and maximum at 200 nm (Figure 35 A). Near-UV CD spectra further confirm the unfolded nature of the mutants. Here, C_H1 and C_H1 GS result in a similar

spectrum shape with a minimum around 280 nm, only the amplitude is higher for C_H1 (Figure 35 B). C_L GS which is supposed to resemble C_H1 in terms of folding, has a shifted minimum at around 268 nm, for C_H1 GS almost no specific features are visible, indicating a loss of tertiary structure compared to C_H1 and C_L wild type. So concerning secondary structure, the mutants seem to possess the same intrinsically disordered state like C_H1 wild type. Similar results were obtained for the tertiary structure, while C_L GS seems to be folded differently than C_H1 and C_L wild type, C_H1 GS exhibits almost no tertiary structure.

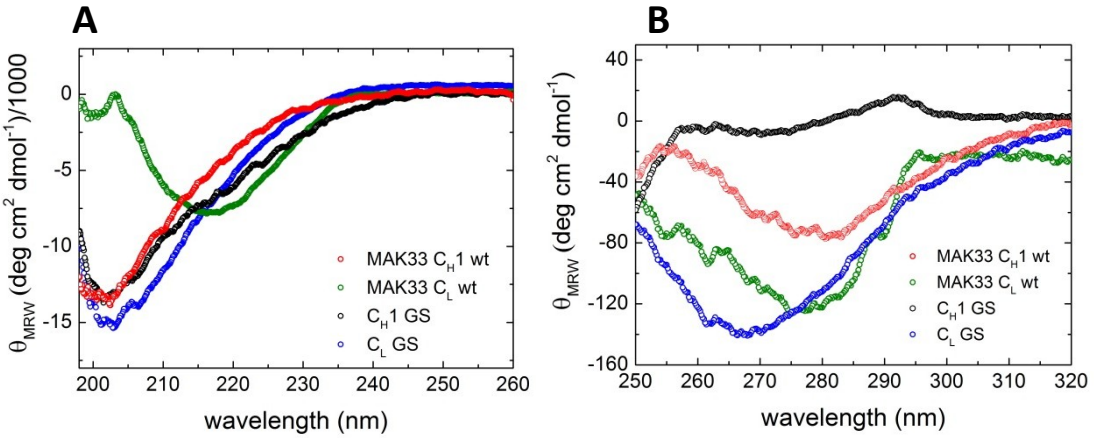


Figure 35: Structural characterization of the C_H1/C_L swap mutants
In (A) FUV-CD spectra and in (B) the NUV-CD spectra of the C_L and C_H1 swap mutants are shown. For the spectra 15 accumulations each were recorded and buffer-corrected. All measurements were performed at a protein concentration of 20 µM (FUV-CD) and 50 µM (NUV-CD) in 0.5 mm (FUV) or 5 mm (NUV) quartz cuvette at 20°C.

The size distribution by AUC detected 100 % monomeric protein for all wild type and mutant domains (Table 8). So the monitored structural deviations are not due to oligomer formation.

Additionally, correct folding of the mutant domains was assessed by intrinsic tryptophan fluorescence (Figure 36). For all domains a red shift from 340 to 355 nm between the native and denatured state was visible, indicating a change of the conserved tryptophan residue environment from hydrophobic to hydrophilic. Only for C_H1-GS, a small shift towards 350 nm is obtained for the native state. This fits to the near-UV CD data, which implied a severe loss of tertiary structure. In general, the difference in signal intensity between native and denatured state was subject of variations, for all the domains the native sample lies underneath the denatured one as expected. However, the signal

difference is more pronounced for C_H1 and C_H1 GS, while C_L GS shows only a small difference. This indicates that less tryptophan quenching is present in C_L GS. Additionally, a reduced sample under denaturing conditions was measured in order to confirm a correct disulfide bridge linkage. Since the conserved disulfide bridge exerts a quenching effect on the conserved tryptophan residue, a signal change should be visible under reducing conditions. This is the case for all wild type and mutant domains which demonstrates that the presence of the conserved disulfide bridge.

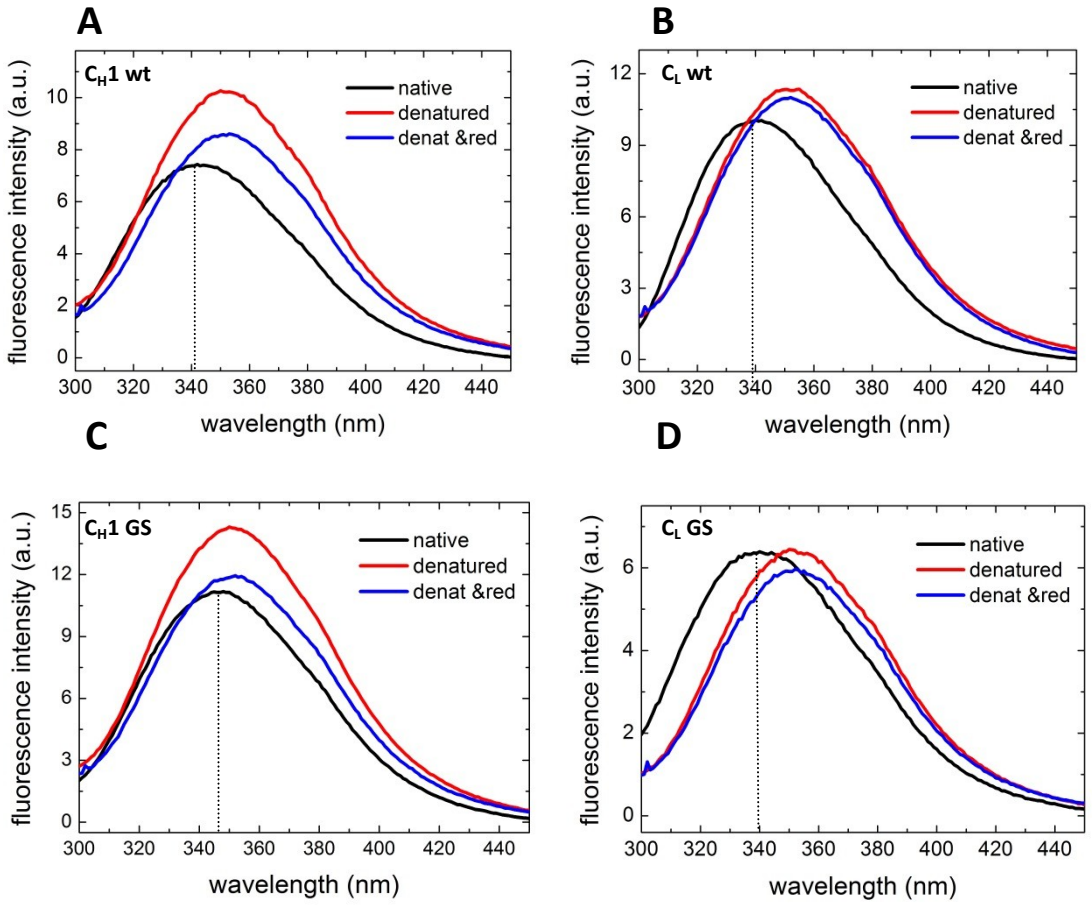


Figure 36: Intrinsic tryptophan fluorescence of the C_L/C_H1 swap mutants
Intrinsic tryptophan fluorescence was determined by fluorescence spectroscopy at an emission wavelength of 280 nm for (A) C_H1 wt, (B) C_L wt, (C) C_H1 GS and (D) C_L GS. The samples were measured under three different conditions: native, denatured (6M GdmCl) and denatured and reduced (6 M GdmCl, 1mM TCEP). Dotted lines indicate the peak maximum for the native sample.

Since the structural data pointed towards folding impairments of C_H1 GS and C_L GS, the hydrophobicity of the surface was investigated by ANS binding (Figure 37). Since no signal increase was monitored upon addition of protein (Figure 37 A), no hydrophobic residues seem to be surface-exposed. Interestingly, a blue shift was observed for C_H1 wild type,

C_H1 GS and most pronounced for C_L GS. This occurs when ANS is interacting with the positively charged amino acids lysine or arginine. In this context, it was interesting to see whether this interaction persists when the corresponding interaction partner is present (Figure 37 B). Since C_H1 GS is supposed to resemble the C_L domain and *vice versa*, different combinations were measured.

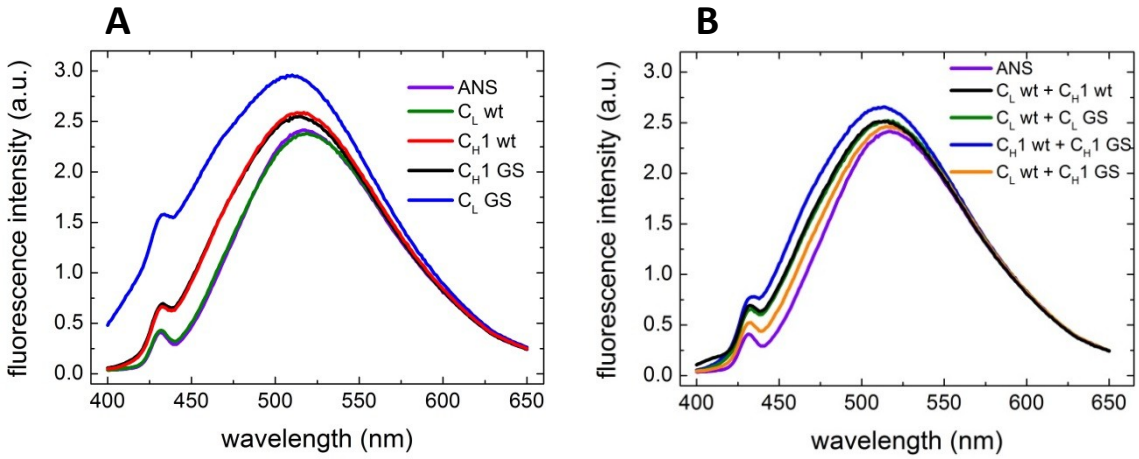


Figure 37: Surface hydrophobicity of the C_H1/C_L swap mutants
The presence of surface-exposed hydrophobic regions was determined by fluorescence spectroscopy via ANS binding. (A) shows the wild type domains and the C_H1/C_L swap mutants in comparison to ANS without protein. In (B) the mutants were mixed with C_L or C_H1 wt before ANS addition. Samples were measured with a protein concentration of 1uM for each protein and an ANS concentration of 50 µM in a 10 mm quartz cuvette at an emission wavelength of 380 nm at 20 °C.

Interestingly, when C_L wild type is added to C_L-GS or C_H1-GS the blue shift is less pronounced, which implies a potential charge-charge interaction between both mutants and the C_L domain. Stability of the constructs was assessed by chemical denaturation with GdmCl (Figure 38, Table 8)

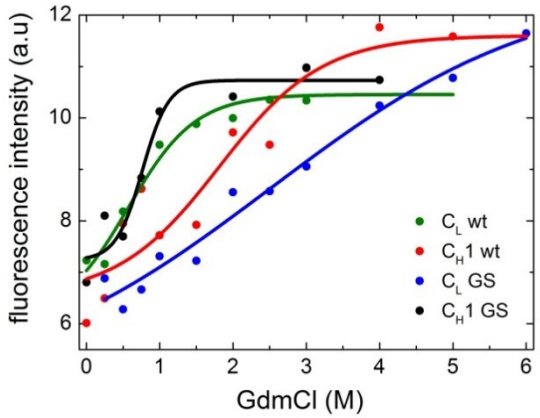


Figure 38: Stability of the C_H1/C_L swap mutants.
The stability of the C_H1/C_L swap mutants against chemical denaturation with GdmCl was determined by fluorescence spectroscopy, measuring the change of the intrinsic tryptophan fluorescence (emission at 280 nm) upon unfolding at 20 °C. The obtained data was Boltzmann fitted.

Table 8: Stability and size distribution of the C_H1 and C_L swap mutants

Domain	$D_{1/2}$ [M]	Monomer [%]
C_H1 wild type	1.8 ± 0.4	100
C_L wild type	0.76 ± 0.1	100
C_H1 GS	0.56 ± 0.3	100
C_L GS	2.5 ± 0.8	100

The size of the C_H1 and C_L swap mutants was determined by AUC sedimentation velocity runs at a protein concentration of 5 - 10 µM with absorbance optics at 230 nm. Data evalution was perfomed with the program Sedfit (Peter Schuck) using a c(S) model.

In order to analyze the ability of the mutants to fold or induce folding in the presence of the corresponding interaction partner, kinetic measurements by CD were executed (Figure 39). Again, combinations of the mutants with both wild type domains were tested as it was not clear after the obtained folding data, with which domain the mutants might interact. For the wild type kinetics of C_H1 and C_L (Figure 39 A), a signal increase at a constant wavelength of 205 nm is visible, reaching saturation after about 60 minutes. For all combinations with the mutant domains no change in the ellipticity signal was observed. So none of the mutants was able to induce C_H1 folding, nor is C_L capable to fold C_L GS or C_H1 GS. To exclude that folding occurred within the seconds between mixing and starting the measurement, far-UV CD spectra were measured after kinetics (Figure 39 F). These confirmed that folding was not induced in any combination containing a mutant domain. Since Feige and coworkers [6] mention that formation of the conserved disulfide bridge is essential for a successful kinetic measurement, this was investigated by SDS-PAGE for all C_H1 and C_L domains including the studied mutations, under oxidizing and reducing conditions (data shown for MAK33 C_H1 wt; Figure 40). The gel showed one band underneath the 14 kDa band for the sample with reducing agent and two bands for the sample without. Due to the more compact form in presence of the conserved disulfide bridge, oxidizied antibody domains migrate faster in an SDS-PAGE [213]. This means that about 80 % of the purified MAK33 C_H1 wild type domain possess the required disulfide bond. Similar or better results were obtained for the C_L domain and the two GS mutants (data not shown).

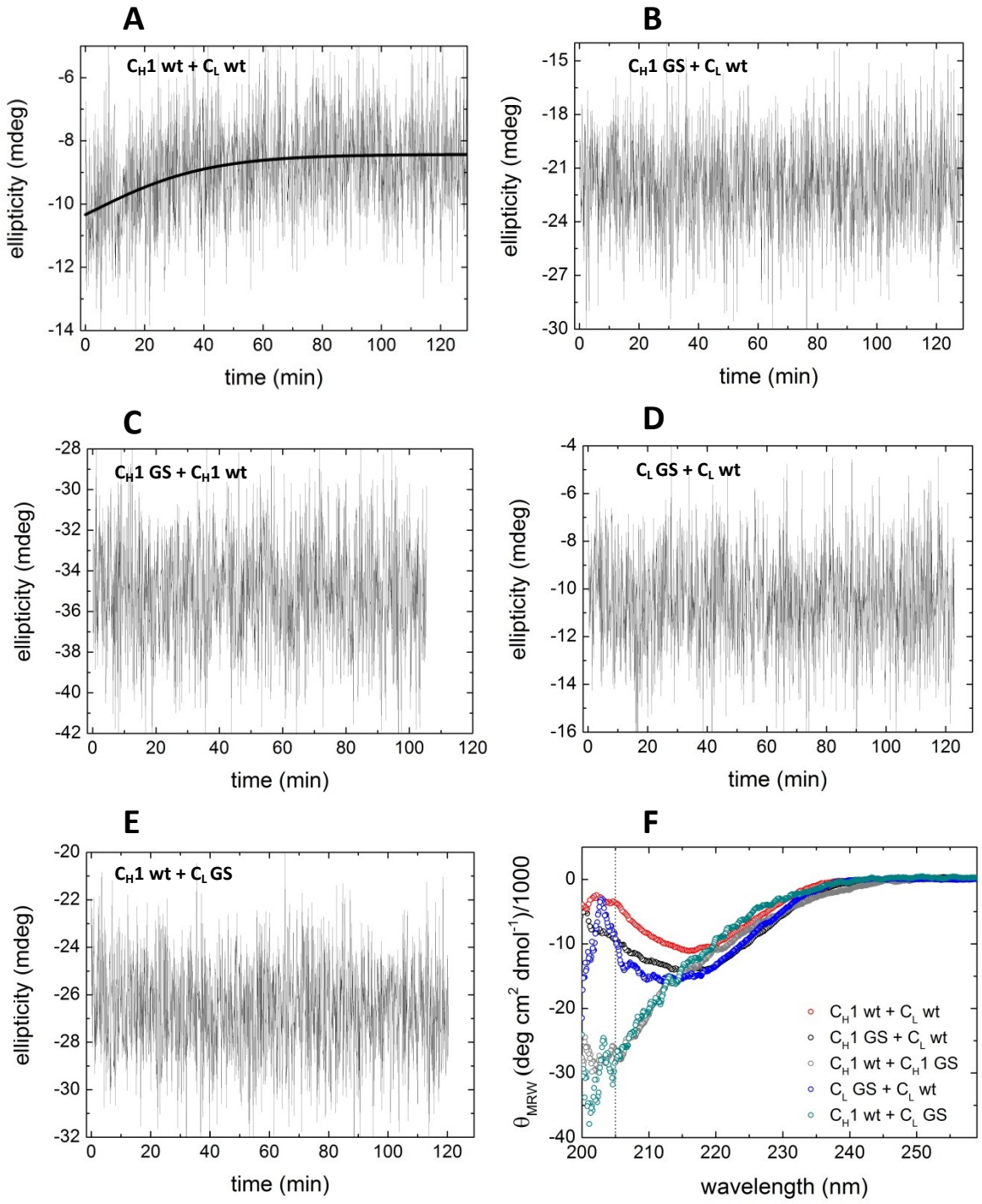


Figure 39: Kinetic measurements of the C_H1 MAK33 folding.

Folding of the C_H1 domain was monitored by CD spectroscopy at a constant wavelength of 205 nm and 25 °C for 120- 180 minutes. C_H1 alone was equilibrated for about 15 minutes at 25 °C, upon addition of C_L the kinetic was started. Next to wild type domains (A), different combinations of the C_L wt and C_H1 wt with the mutant C_H1 GS (B and C) and C_L GS (D and E) were tested. Directly after the kinetics a FUV spectrum with 3-5 accumulations was measured, which is shown in (F) including a line which labels the kinetics signal of 205 nm. The CD measurements were performed at a protein concentration of 10 µM each, in a 1 mm quartz cuvette.

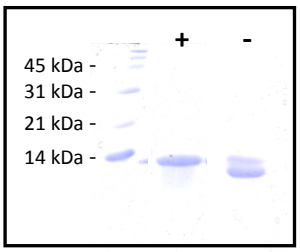
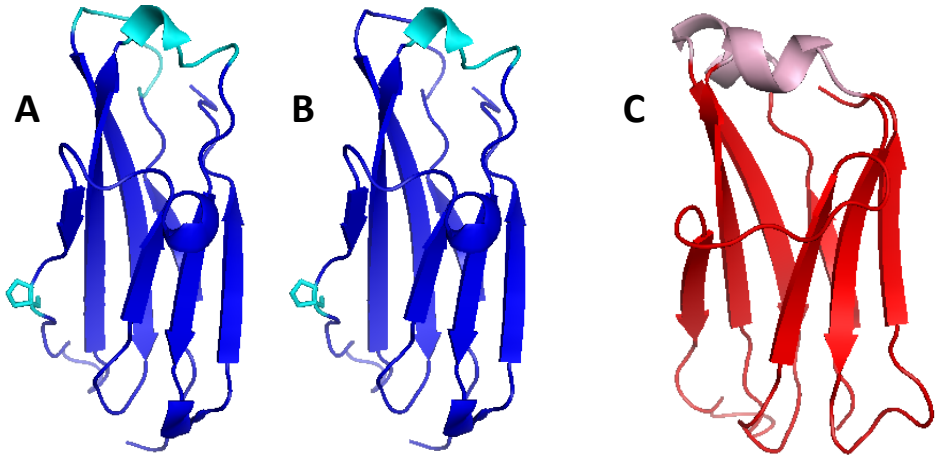


Figure 40: 18 % SDS gel of C_H1 MAK33 under oxidizing and reducing conditions.
The migration of C_H1 MAK33 samples with (+) and without (-) 1 mM ß-mercaptoethanol was compared. As a size standard, a low molecular weight marker (Biorad) was also applied on the gel. All SDS-PAGE samples contained 1x Lämmli buffer, separation was conducted at a voltage of 35 mA for 50 min.

To summarize, the obtained data for C_H1 GS and C_L GS hint towards a misfolded structure for both domains. C_H1 GS did not show any features of the C_L domain. C_L GS at least shared the intrinsically disordered secondary structure with C_H1 but this was not the case for the tertiary structure. Moreover, in the presence of C_L wild type C_L GS, which was supposed to behave like $C_H1$wild type, was not able to fold.

3.3.2.2 Selected C_H1 folding mutants

In the second approach, two C_H1 mutants (MF3 and MF4) were generated based on the ideas of Dr. Matthias Feige, focusing the two helical regions H1 and H2 of C_L (see Figure 41).



CH1 MF3: ASTKGPSVFPLAP~~SSKSTSGG~~**DEQLKSG**TAALGCLVKDYFPEPVTVSWNSGALTSGVHTF**P**AVLQ SSGLYSLSSVVTVP~~SSSLGT~~**KADYEKH**QTYICNVNHKPSNTKVDKKVEPKS

CH1 MF4: ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTF**P**AVLQSSGLYSL SSVPVTV~~SSSLGT~~**KSEYER**QTYICNVNHKPSNTKVDKKVEPKS

Figure 41: Mutation sites for the C_H1 folding experiments
Pymol cartoon representation of 1GAF (A and B) C_H1 in blue with the proline residue and the regions mutated for the (A) MF3 and (B) MF4 construct highlighted in cyan. The transferred regions of the C_L domain (C) in red are highlighted in light pink. Underneath the amino acid exchanges are shown for constructs MF3 and MF4 within the C_H1 sequence. Regions that were removed are highlighted in cyan as well as proline residue 54, the replacing amino acid sequences in red.

These helical regions correspond to disordered regions in C_H1, which contain a high portion of serine and glycine residues. These conserved helical regions have already previously been subject of folding experiments. Exchange of the helices of a human C_L domain with a modified version, which can be found in the highly stable constant domains of IgNAR increased the overall C_L stability as well as antibody secretion [177]. The exchange of the C_L helices with unstructured loop regions of the amyloid-prone $ß_2$-microglobulin induced a well-defined structure and the abolishment of amyloid formation [158]. During the process of the C_L domain folding, the two helices are one of the first assigned structural features which are already present in the intermediate state. They are suggested to fulfill various functions within the C_L domain: 1) spacer and orientation between strands A-B and E-F; 2) hydrogen acceptor and donor for adjacent loops and strands; 3) positioning of hydrophobic residues (like Y80 in H2) for hydrophobic core formation [158]. For the following *in vitro* study, a different IgG1/κ light chain antibody (pdb ID 1GAF) similar to MAK33 was used for the analysis, containing a human C_L and C_H1 domain. This human 1GAF C_H1 domain shows less complexity and less helical propensity compared to the murine MAK33 C_H1. The major difference, though, is the tryptophan residue in helix 2 only present in MAK33, which participates in the hydrophobic core. For mutant MF3, a complete helix transfer from C_L was conducted (see Figure 41). A very similar experiment with an exchange of the two helices has already been performed with MAK33 C_H1 (data not shown) and also *vice versa* with C_L [6] by Matthias Feige. This led to an unfolded C_L domain whereas the C_H1 domain remained unfolded. Concerning the C_H1 MF3 mutation, this experiment is not only performed within a different antibody context but also involves a proline to serine mutation within the HTFPAVL BIP binding motif [153] (highlighted in Figure 41) in order to erase the potential helix breaking effect of this proline. Generally, C_H1 contains a higher amount of prolines than the other constant domains including three cis prolines. The relevance of the three cis prolines in C_H1 MAK33 in terms of folding, has been investigated by Matthias Feige and coworkers (see Figure 42). They identified one essential cis proline at position 32 which is the prerequisite for the rate limiting folding step [6]. Mutation experiments showed that proline residues 34 and 74 were not necessary for folding of the C_H1 domain. However, the proline to alanine mutation of residue 74 in MAK33 C_H1 increased the heavy chain assembly and secretion [6]. In general, proline residues favor the formation of tight turns, which connect the

strands of Ig domains. They have also been suggested to decrease the aggregation propensity of Ig domains [214]. Hence, proline residues with potential effects on the folding behavior were also considered for the mutations in this study. Since proline 32 is essential for folding and 34 did not show any affect, only proline 74 was of interest (see Figure 42). However, in 1GAF C_H1, a glycine residue can be found at the corresponding position and moreover it is located in helix 2, which is anyways exchanged for the two mutants. Proline 74 present in MAK33 was among the mutations, though, introduced for the first C_H1 folding approach with the C_H1/C_L swaps. So for the two mutants, MF3 and MF4, proline 54 was chosen for a mutation to serine since it might be important due to its positioning within the BIP binding motif.

The second mutant, MF4, contains the exchange of helix 2 with a sequence very similar to MAK33 C_L helix 2 with the difference of a serine at position 2 instead of aspartate (KDEYER) (see Figure 41). Helix 2 often possesses an aromatic or hydrophobic amino acid which participates in the hydrophobic core [158, 177]. Adjacent to the hydrophobic contact, the formation of two salt bridges should be enabled by the introduced amino acids.

```
1GAF CH1:  ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSG
MAK33 CH1: --TTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQS D
             *. ***:****.* : :.. .:******.********:****:*:************ .
LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS
LYTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPR
**:*** ******:  ::*  *** * .*.******: *:
```

Figure 42: Investigated proline residues in C_H1.

Sequence alignment of the human (1GAF) and murine (MAK33) C_H1 performed by CLUSTALW. The two regions which correspond to the helical regions in C_L are highlighted in red. Cis prolines that were previously investigated by Feige and coworkers in MAK33 by alanine point mutations are highlighted in orange. The proline which was mutated within this study is illustrated in green.

The far-UV CD spectra show a high similarity for the C_H1 mutants MF3 and MF4 with a minimum around 200 nm (Figure 43 A). The C_H1 wild type spectrum exhibits less signal intensity with a minimum around 205 nm while C_L wild type shows a typical ß-sheet structure with a minimum at 218 nm and a maximum at 200 nm. Unfortunately, tertiary structure by near-UV CD could not be measured due to the low protein yield for C_H1 wild type. However, already on the secondary structure level a folding behavior similar to the C_L domain, which was the intention of the introduced mutations, can be excluded.

Investigation of the stability by thermal unfolding further proved the absence of folded regions (Figure 43 B). The normalized CD raw data of C_H1, C_H1 MF3 and C_H1 MF4 already shows a high signal to noise ratio, caused by a very low change in ellipticity signal. So in contrast to C_L only low amounts of secondary structure were present before thermal unfolding. Nevertheless C_L exhibits the highest stability with a T_m of 55.8 ± 0.1 °C compared to C_H1 wild type with a T_m of 37.2 ± 0.4 °C and the C_H1 mutant MF4 with a T_m of 32.9 ± 4.0 °C. (Figure 43, Table 9). Due to the linear fitting curve of C_H1 mutant MF3, no T_m value could be determined. This indicates that the introduced mutations seem to be disfavorable in terms of stability.

The AUC analysis showed that the wild type domains and the two mutants are completely monomeric (Table 9).

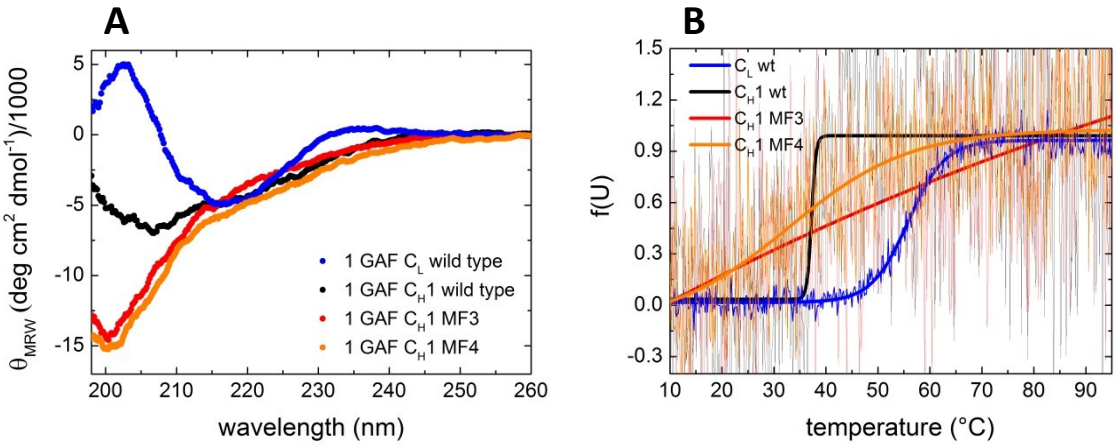


Figure 43: Secondary structure and stability of the C_H1 mutants MF3 and MF4.
In (A) FUV-CD spectra of the C_H1 mutants MF3 and MF4 in comparison to C_H1 and C_L wt are shown. For the spectra 15 accumulations each were recorded and buffer-corrected (PBS). In (B) the stability against thermal unfolding is depicted. All CD measurements were performed at a protein concentration of 20 µM (FUV-CD) in a 0.5 mm (FUV) quartz cuvette at 20°C. The heating rate for thermal unfolding was 20°C/h at a constant wavelength of 205 nm; the normalized data was Boltzmann fitted.

Table 9 : Stability values and size distribution of the C_H1 MF mutants

Domain	T_{melt}[°C]	Monomer [%]
C_H1 wt	37.2 ± 0.4	100
C_L wt	55.8 ± 0.1	100
C_H1 MF3	ND	100
C_H1 MF4	32.9 ± 4.0	100

The midpoint of transitions (T_{melt}) was determined by Boltzmann fit. ND stands for not determinable due to a poor fitting. The size of the C_H1 and C_L swap mutants was determined by AUC sedimentation velocity runs at a protein concentration of 5 - 10 µM. Data evaluation was performed with the program Sedfit (Peter Schuck) using a c(S) model.

The determination of the surface hydrophobocity by ANS binding revealed that no hydrophobic regions are surface-exposed for the wild type domains or the C_H1 mutants (Figure 44). Only a blue shift and signal increase was visible for C_H1 mutant MF4, indicating an interaction of ANS with arginine or lysine residues of MF4. A similar shift was observed for MF4 in presence of C_L (Figure 44 B). So no charge-charge interactions competing with ANS binding occur.

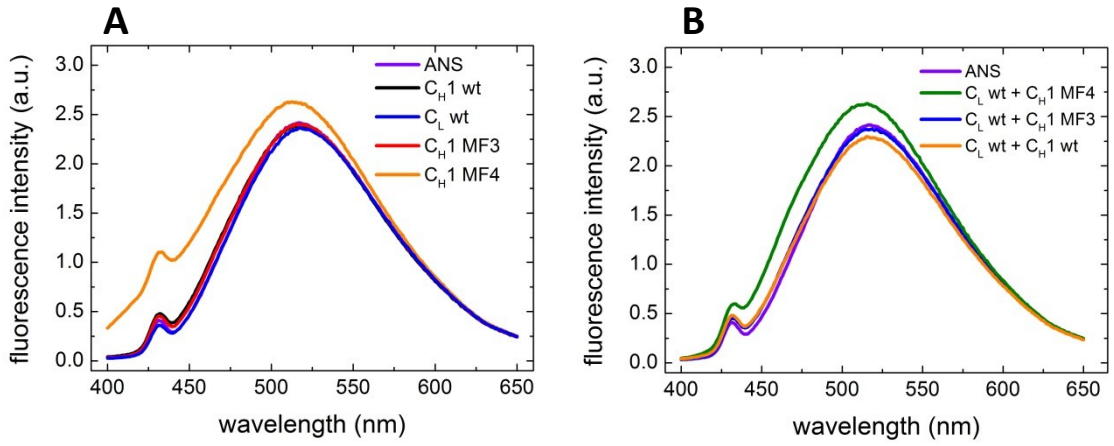


Figure 44: Surface hydrophobicity of the C_H1 mutants MF3 and MF4.

The presence of surface-exposed hydrophobic regions was determined by fluorescence spectroscopy via ANS binding. (A) shows the wild type domains and C_H1 mutants in comparison to ANS without protein. In (B) the mutants were mixed with C_L wt before ANS addition. Samples were measured with a protein concentration of 1uM for each protein and an ANS concentration of 50 µM in a 10 mm quartz cuvette at an emission wavelength of 380 nm at 20 °C.

The kinetics of C_H1 wild type folding upon addition of C_L wild type show a high signal increase which reaches saturation after about 50 minutes (Figure 45 A). However, all possible combinations of the C_H1 mutants with C_L or CH1 wild type results in no detectable change of signal (Figure 45 B-D). However, the far-UV CD spectra recorded subsequently to the kinetics showed a complete different picture. When substracting the far-UV CD spectrum of the single C_L domain the spectra of C_H1 wild type as well as the mutants MF3 and MF4 exhibit the spectrum of a folded C_H1 domain with a maximum at 217 nm (Figure 45 E)[6]. This means the complete folding of the C_H1 mutants must happen within the few seconds between addition of C_L wt and starting the measurement. Both mutants share the proline to serine mutation within the BIP binding motif and different helix 2 exchanges. Hence, the cause of this effect has to be further assessed. The same holds true for the folding rate constants of the mutant C_H1 domain. In this case,

stopped flow with a very small dead time between mixing and start of the measurement seems to be a more appropriate method than manual mixing.

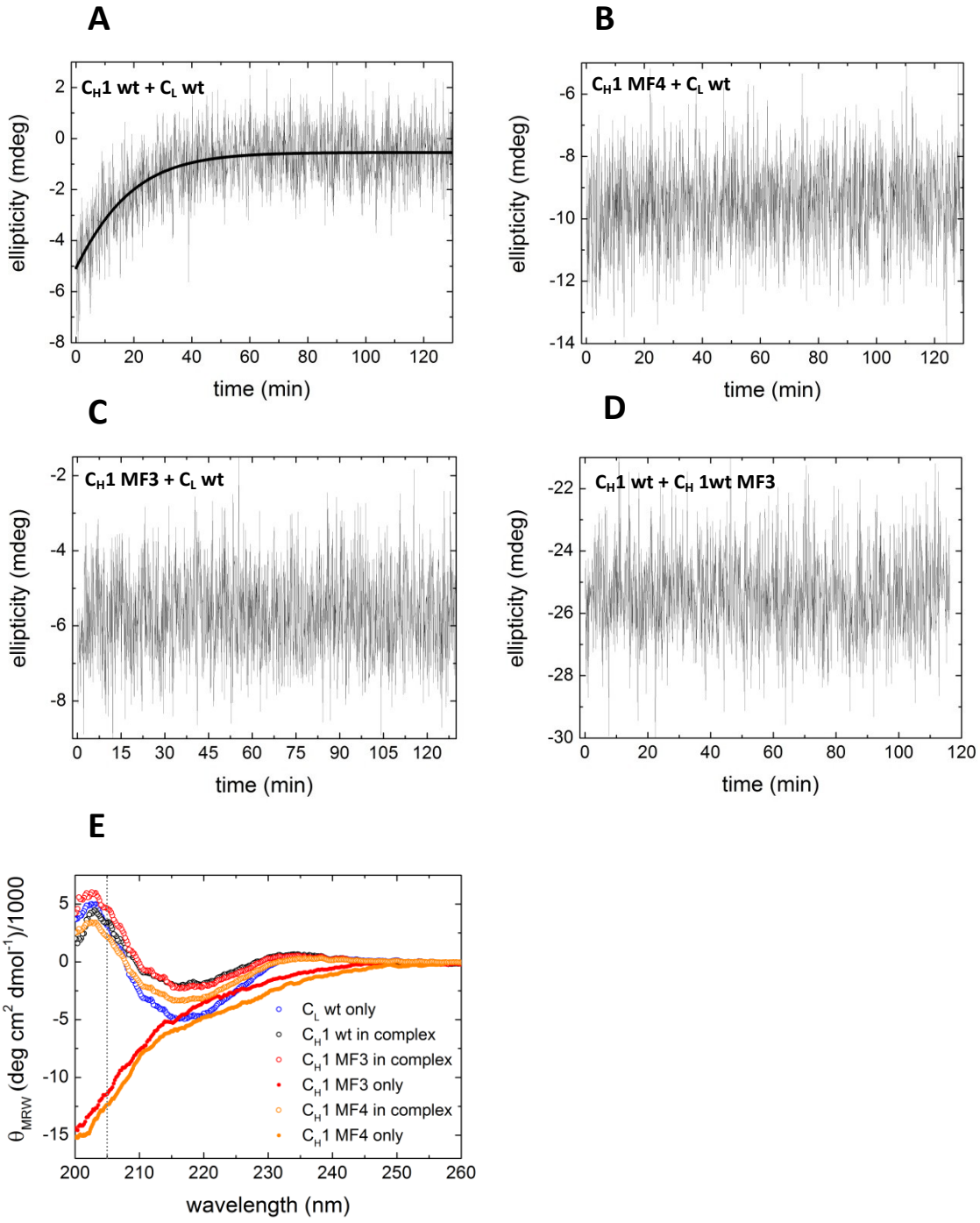


Figure 45: Kinetic measurements of C_H1 1GAF folding

Folding of the C_H1 domain was monitored by CD spectroscopy at a constant wavelength of 205 nm and 25 °C for 120- 180 minutes. C_H1 alone was equilibrated for about 15 minutes at 25 °C, upon addition of C_L the kinetic was started. Next to wild type domains (A), different combinations of the C_L wt and C_H1 wt with the C_H1 mutants C_H1 MF3 (C and D) and MF4 (B) were tested. Directly after the kinetics, a FUV spectrum with 5 accumulations was measured. The normalized and C_L subtracted data, in comparison to the single measurements of CL, C_H1 MF3 and MF4 is shown in (F) including a dotted line which labels the wavelength of used for kinetics. The CD measurements were performed at a protein concentration of 10 uM for each protein in a 1 mm quartz cuvette.

3.3.3. Discussion

Within this study none of the generated C_H1 mutants nor the C_H1 domain from camel HcAbs was able to fold in isolation or upon addition of C_L. Nevertheless some C_H1 mutants showed an improved folding behavior with an immense shift in the timescale of folding from about 50 minutes to less than 5 seconds.

Concerning C_H1 from camelid HcAbs which is not expressed due to a point mutation in the adjacent splicing site, it could clearly be shown that HcAbs must have evolved after the loss of the C_H1 domain. In isolation, camel C_H1 is unfolded, aggregation-prone and forms higher oligomers without any monomer fraction.

The two different mutational approaches of the C_H1 domain in order to generate a foldable C_H1 showed similar results in terms of folding behavior in isolation. Approach one is based on the comparative bioinformatics analysis of the C_H1 and C_L domain performed by George Stan and coworkers. This led to a list of amino acids associated with the structural features that are divergent between C_H1 and C_L. These calculated amino acid exchanges were introduced in C_L and C_H1 MAK33 (murine IgG1) in order to produce a foldable C_H1 domain (C_H1 GS) and an unfolded C_L domain (C_L GS). On the secondary structure level, spectra of an intrinsically disordered protein were observed for C_H1 wild type, C_H1 GS and C_L, which was desirable for C_L GS but not for C_H1 GS. When it came to tertiary structure, though, C_H1 GS did not resemble C_H1 wild type anymore but exhibited almost no tertiary structure instead. C_L GS showed a spectrum shape different from both wild type domains. Fluorescence spectroscopy verified the presence of the conserved disulfide bridge for all domains. The significant loss of tertiary structure for C_H1 GS was further confirmed by a small red shift of the native sample. Similar results were obtained for C_L GS, with a signal increase of the native sample indicating a less pronounced presence of the intrinsic tryptophan quenching effect. No surface hydrophobicity by ANS binding was detected which means that even the misfolded C_H1 GS and C_L GS mutants seem to possess the initial hydrophobic core. Interestingly, C_H1-GS, C_H1 wild type and especially C_L GS showed interaction with ANS mediated by arginine or lysine residues. This effect is reduced for C_H1 GS and C_L GS when C_L wild type is added but not for C_H1 wild type. This hints towards a charge-charge interaction between C_L and the two mutants. However this result is rather unexpected since for C_L GS many arginine and lysine residues

have been removed and the opposite was the case for C_H1 GS. Since this measurement was performed only once, the reproducibility of the obtained has to be shown. The kinetics revealed that only the combination of C_L wild type with C_H1 wild type is able to provoke a folding event accompanied by a CD signal increase, which takes about an hour until saturation is reached. All possible mutant combinations showed no change in signal. Hence, the mutants did not exhibit the expected behavior but seem to be somehow misfolded instead.

Approach two is based on the ideas of Matthias Feige focusing the exchange of the two helical regions H1 and H2 and the mutation of proline 54 within the BiP binding motif. Here, two C_H1 mutants, C_H1 MF3 and MF4, within a human IgG1 antibody (1GAF) context were generated with the aim of producing foldable C_H1 domains. Assessment of the secondary structure already showed spectra for C_H1 MF3 and MF4 as observed for C_L but spectra of an intrinsically disordered protein similar to C_H1 wild type. Thermal unfolding measurements with a high signal to noise ratio for C_H1 wild type and both C_H1 mutants further confirmed the presence of only very small amounts of secondary structure. As already observed for the mutants from the first approach, no surface-exposed hydrophobicity by ANS binding was detected for any of the domains. Thus also these C_H1 mutants seem to possess the initial hydrophobic folding core. C_H1 MF4, though, showed an interaction with ANS mediated by arginine or lysine residues. Since helix 2 of MF4 was exchanged by a sequence containing one lysine and one arginine this effect might be associated with these two residues. If this is the case, the two salt bridges that were supposed to be formed by the introduced amino acids were not established. Considering the fact that this was not observable for C_L wild type and C_H1 MF3 which both contain the 1GAF helix 2 including two lysines, this effect is limited to helix 2 used exclusively for MF4 showing a high sequence similarity to helix 2 from MAK33 C_L. For the measured kinetics, a signal was only obtained for the combination of C_L and C_H1 wild type. Folding of C_H1 wild type by C_L took about 50 minutes. This is slightly faster than the kinetic values obtained for MAK33 C_H1 and C_L. A comparison of the wild type kinetics from the two antibodies results in a four-times higher signal change for 1GAF. When comparing the MAK33 kinetics obtained within this study to the published MAK33 kinetics from Feige and coworkers [6], a similar signal deviation is visible. Interestingly, when combining C_H1 1GAF with C_L MAK33 and *vice versa,* a kinetic with a high change in signal is measured for the

first combination and no signal for the second (Figure 46). Thus, the low signal is probably connected to an impairment of MAK33 C_H1. Since it could be shown that about 80 % of the purified MAK33 C_H1 wild type domain possesses the required disulfide bond (Figure 40), the lack of the disulfide bridge can not explain the obtained results. Hence, the reason for the difference in the kinetics signal has to be further investigated.

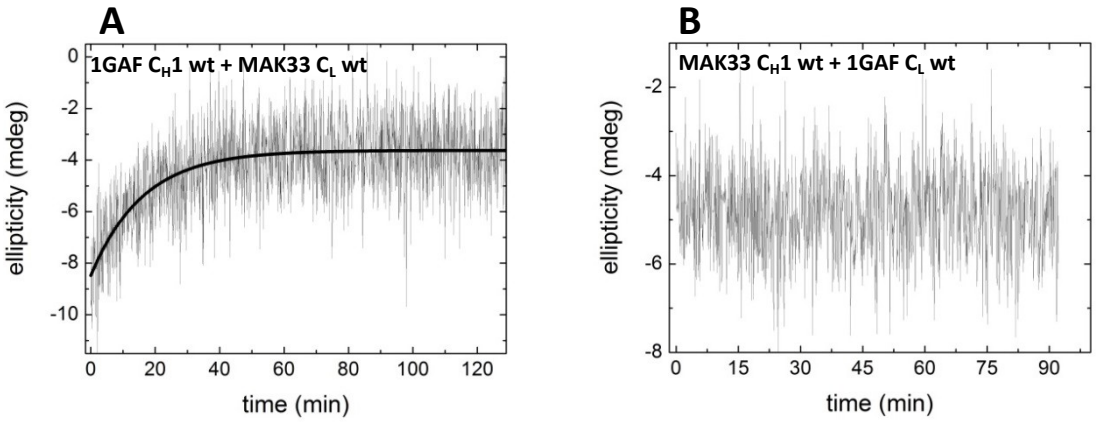


Figure 46: Kinetics of C_H1 folding with C_L domains of different origin.
Folding of the C_H1 domain was monitored by CD spectroscopy at a constant wavelength of 205 nm and 25 °C for 120- 180 minutes. C_H1 alone was equilibrated for about 15 minutes at 25 °C, upon addition of C_L the kinetic was started. C_H1 and C_L wild type domains from different origin were combined. (A) shows a 1GAF C_H1 with MAK33 C_L and (B) MAK33 C_H1 with 1GAF C_L. The CD measurements were performed at a protein concentration of 10 uM for each protein in a 1 mm quartz cuvette.

Interestingly, the far-UV spectra recorded after the kinetic measurements for 1GAF exhibited a completely folded spectrum for C_H1 wild type and the two mutants MF3 and MF4. This means folding of these mutants must have occurred within the dead time of the measurement in ≤ 5 seconds between addition of C_L and starting the measurement. So a dramatic shift in the timescale of C_H1 folding occurred compared to C_H1 wildtype which folds in about 50 minutes. A faster folding of the wild type domain has so far only been observed in the presence of the peptidyl-prolyl-cis-trans isomerase cyclophilin B, where C_H1 MAK33 folded in about 60 minutes instead of 120 minutes [6]. Now, it would be interesting to find the exact reason for this observed effect. The introduced mutations of MF3 and MF4 overlap concerning the proline to serine mutation and exchanges in helix 2 (see Figure 47). But the exchanges in helix 2 involved similar but different sequences. Considering the high impact of some proline residues of C_H1 in terms of folding [6] and the location of the related proline within the conserved BiP binding motif HTFPAVL, the obtained effect is assumed to be associated with the proline mutation. This can be easily

verified by the generation of a C_H1 mutant which contains either the proline mutation or an exchange of helix 2. Ideally, this is performed with MAK33 and 1GAF C_H1 to examine whether this effect is conserved among IgGs from different origin.

C_H1 MF3: ASTKGPSVFPLAP~~SSKSTSGG~~DEQLKSGTAALGCLVKDYFPEPVTV**S**WNSG**AL**TSGVHTF~~P~~**S**AVL QSSGLYSLSSVVTVP~~SSSLGT~~**KADYEKH**QT**Y**IC**N**VNHKPSNTKV**D**KKVEPKS
C_H1 MF4: ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTV**S**WNSG**AL**TSGVHTF~~P~~**S**AVLQSSGLYS LSSVPVTV~~SSSLGT~~**KSEYER**QT**Y**IC**N**VNHKPSNTKV**D**KKVEPKS

Figure 47: Comparison of the introduced sequence exchanges for the two investigated folded C_H1 domain approaches.
The introduced mutations for the generation of the 1GAF C_H1 MF3 and MF4 mutants are shown. Additionally, the corresponding amino acid positions that were subject of exchanges for the first approach are highlighted in green.

In case the proline residue of the BiP motif can be identified as the reason for this rapid folding, this might depict an additional quality control mechanism next to the unfolded nature of C_H1. In this scenario, the extreme deceleration of C_H1 folding by proline 54 might prevent uncontrolled folding events in the absence of BiP. Considering the wild type kinetics in which C_H1 folding requires about 50 minutes, this is no acceptable timescale for *in vivo* conditions in the cell, even if cyclophilin B is present. So it might be possible that upon interaction with BiP, conformational changes in C_H1 occur that enable a rapid folding in the range observed for the C_H1 MF mutants, as soon as C_L is present. Assuming these conformational changes are mainly induced by interactions of proline residue 54 with BIP, the introduced proline to serine mutation in the C_H1 mutants MF3 and MF4 might be able to imitate the conformational state that is otherwise only obtained in presence of BIP (see Figure 48). This would represent an explanation with regard to the physiological relevance of the fast folding event observed for the C_H1 MF mutants but of course this theory has to be validated by experimental data. In this context, it might also be interesting to express a Fab fragment containing the C_H1 domain with the proline mutation in mammalian cells. With this experiment, possible changes in the secretion of the Fab fragment, associated with the C_H1 mutation, can be investigated.

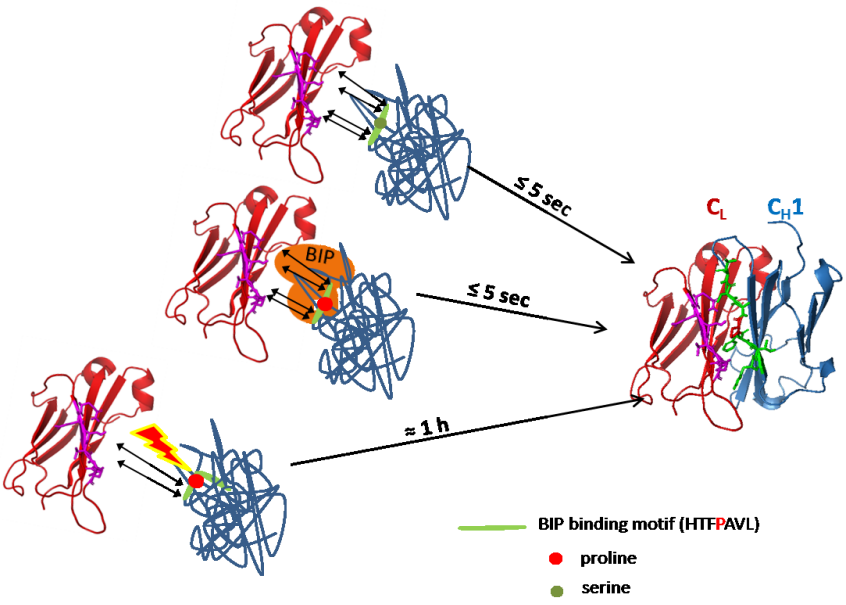


Figure 48: Potential involvement of the mutated proline residue 54 in C_H1 folding
Potential folding pathways of C_H1 by different modes of interaction with the folding partner C_L. Possible interaction sites (pink) of C_L (red) with the BIP binding motif HTFPAVL (green) including the corresponding proline residue in red are shown. Here, conformational changes in C_H1 occur upon binding of BIP that promotes a fast folding of C_H1 in presence of C_L. The proline to serine mutation imitates the conformational changes in the presence of BiP.

When comparing the two failed approaches for generation of a foldable C_H1 domain (Figure 47), the first one with a detailed bioinformatics analysis for the identification of structural deviations and the amino acid interaction networks associated with it, depicted the more sophisticated one. The second approach with the complete helix exchanges was rather rough. Nevertheless the highest overlap of both can be identified within the helix 2 sequence. So maybe both approaches do not tap their full potential as they might depend more than expected on the proline contribution for folding in isolation. In this regard more knowledge is required to dissect the role of the prolines within the C_H1 domain.

Conclusion

Within this study, different isolated antibody domains were investigated *in* vitro in the context of topics, relevant for understanding antibody structure with a view to allow engineering. The variable domains of the human 1HEZ IgM antibody were examined with regard to the highly conserved amino acids in the V_H/V_L interface, identified by Wang and coworkers within a covariation analysis [5]. The question of the potential importance of these residues for I) folding II) stability III) V_H/V_L interaction or IV) antigen binding was addressed by an alanine point mutation approach. It could be shown that V_H residues E46 and R38 are essential for folding. Residues Y36 for V_L and residues E46 and W47 for V_H strongly impaired the domain stability. In terms of V_H/V_L interaction, mutation of residues P44 in V_L and L45 and V37 in V_H led to severe impairment of the binding to the partner domain. By the transfer of some V_L point mutations to a MAK33 CDR grafting construct, it could be proven that the observed point mutation effects are not CDR-dependent. Together with the point mutation data of Dr. Eva-Maria Herold within the MAK33 variable domains, a general conclusion about the observed effects and the associated relevance of the conserved residues can be drawn.

The second project investigated the stabilizing effect of additional C- and N-terminal amino acids on the isolated MAK33 C_H2 domain. Hence, mutants with different numbers of C-and N-terminal extra amino acids were characterized under physiological conditions and pH 2. The results showed that a stabilizing effect can be achieved by addition of either N- or C-terminal extra amino acids, but the effect is slightly more pronounced with C-terminal residues. In this context, a further truncation to only one amino acid per terminus would be of interest. An extension of these studies on other antibody domains might reveal whether this effect is of common nature or only associated to C_H2. Moreover, it has to be examined whether the addition of the natural amino acids of the adjacent domains (hinge and C_H3 with a short linker in between) lead to a similar increase in stability. In case of the light chain domains, it could be shown that distinct linker residues are able to contribute to the domain stability. So for C_H2 a similar effect might be responsible for the observed stability increase.

The third project aimed to identify key regions or amino acids associated with the unfolded nature of the C_H1 domain in isolation. Two approaches were tested to generate

Conclusion

C_H1 mutants that are potentially able to fold in isolation, involving the transfer of structural divergent features of C_L on C_H1. The first approach was based on a computational comparison of the MAK33 C_L and C_H1 structures in order to identify divergent regions and amino acid interactions. As a result, two mutants with several amino acid exchanges between C_L and C_H1 were produced with the aim to generate a foldable C_H1 and an unfoldable C_L domain. In the second approach, two additional C_H1 mutants were generated with the focus on the exchange of two highly divergent helical regions and a proline residue. Unfortunately, within this study no foldable C_H1 domain was obtained. However, the mutants of the second approach show an extreme acceleration of the folding event upon addition of C_L. They fold within a few seconds while the wild type C_H1 domain requires about 50 minutes. This effect is probably caused by proline residue 54. But this has to be further investigated. With regard to some similarities of the targeted regions for the two different approaches, it might be reasonable to generate C_H1 mutants that combine both approaches.

Taken together, this thesis provides new insights in the role of conserved amino acids in the V_H and V_L interface as well as the stabilizing effect of N-terminal and C-terminal amino acid additions on C_H2. Moreover, a key regulator for C_H1 folding was identified leading to an extraordinary acceleration of the folding reaction. The obtained data will be helpful for antibody engineering.

Abbreviatons

ADCC	antibody-dependent cellular cytotoxicity
AFS	alternatively folded state
ANS	1-anilinonaphtalene-8-sulfonate
AUC	analytical ultracentrifugation
CD	circular dichroism
CDR	complementarity determining region
CDR-L1-3	CDR-1-3 of the light chain variable domain
CDR-H1-3	CDR-1-3 of the heavy chain variable domain
C_H1	first constant domain of the heavy chain
C_H2	second constant domain of the heavy chain
C_H3	third constant domain of the heavy chain
CM5	carboxymethylated dextran
Ctr	control
CypB	Cyclophilin B
DTT	di-thiothreitol
EDC	N,N'-1-ethyl-3-(3-dimethylaminopropyl)-carbodiimid
EDTA	ethylenedinitrilotetraacetic acid
ELISA	enzyme linked immunosorbent assay
EMCH	3,3´-N-(ε-maleimidocaproic acid) hydrazide, trifluoroacetic acid salt
ER	endoplasmic reticulum
Fab	fragment antigen binding
Fc	fragment crystallisable
FcRn	neonatal Fc receptor
FcγR	Fcγ receptor
Fd fragment	V_H+C_H1
Fv	fragment of variable domains
FW	framework
GdmCl	Guanidinium chloride
GSH	L-Glutathione reduced
GSSG	L-Glutathione oxidized

Abbreviations

HcAb	Heavy chain antibody
HRP	horse reddish peroxidase
HuCAL	human combinatorial library
Ig	immunoglobulin
IPTG	isopropyl-β-D-thiogalactopyranoside
LB	Lysogeny Broth
Θ MRW	mean residue ellipticity
NHS	*N*-Hydroxysuccinimid
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB ID	protein data bank identification
PDI	protein disulphide isomerase
PPIase	peptidyl-prolyl-*cis-trans* isomerase
RT	room temperature
scFv	single chain variable fragment
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
SE	sedimentation equilibrium
SEC	size exclusion chromatography
SOC	Super Optimal Broth with catabolite repression
SV	sedimentation velocity
TCEP	tris (2-carboxyethyl) phosphine hydrochloride
TEMED	*N,N,N′,N′*-Tetramethylethylendiamin
Tris	tris(hydroxymethyl)-aminomethan
V_H	variable domain of heavy chain
V_{HH}	variable domain of heavy chain of heavy chain antibodies
V_L	variable domain of light chain
wt	wild type

References

1. Wesolowski, J., et al., *Single domain antibodies: promising experimental and therapeutic tools in infection and immunity.* Med Microbiol Immunol, 2009. **198**(3): p. 157-74.
2. Muyldermans, S., et al., *Camelid immunoglobulins and nanobody technology.* Vet Immunol Immunopathol, 2009. **128**(1-3): p. 178-83.
3. Simpson, E.R., E.M. Herold, and J. Buchner, *The folding pathway of the antibody V(L) domain.* J Mol Biol, 2009. **392**(5): p. 1326-38.
4. Richter, K., M. Haslbeck, and J. Buchner, *The heat shock response: life on the verge of death.* Mol Cell, 2010. **40**(2): p. 253-66.
5. Wang, N., et al., *Conserved amino acid networks involved in antibody variable domain interactions.* Proteins, 2009. **76**(1): p. 99-114.
6. Feige, M.J., et al., *An unfolded CH1 domain controls the assembly and secretion of IgG antibodies.* Mol Cell, 2009. **34**(5): p. 569-79.
7. Ellis, R.J. and F.U. Hartl, *Principles of protein folding in the cellular environment.* Curr Opin Struct Biol, 1999. **9**(1): p. 102-10.
8. Ellis, R.J. and A.P. Minton, *Cell biology: join the crowd.* Nature, 2003. **425**(6953): p. 27-8.
9. Anfinsen, C.B., *Principles that govern the folding of protein chains.* Science, 1973. **181**(4096): p. 223-30.
10. Levinthal, C., *Are There Pathways for Protein Folding.* Journal De Chimie Physique Et De Physico-Chimie Biologique, 1968. **65**(1): p. 44-+.
11. Creighton, T.E., *Toward a better understanding of protein folding pathways.* Proc Natl Acad Sci U S A, 1988. **85**(14): p. 5082-6.
12. Baldwin, R.L., *Why is protein folding so fast?* Proc Natl Acad Sci U S A, 1996. **93**(7): p. 2627-8.
13. Morris, E.R. and M.S. Searle, *Overview of protein folding mechanisms: experimental and theoretical approaches to probing energy landscapes.* Curr Protoc Protein Sci, 2012. **Chapter 28**: p. Unit 28 2 1-22.
14. Tanford, C., *Protein denaturation. C. Theoretical models for the mechanism of denaturation.* Adv Protein Chem, 1970. **24**: p. 1-95.
15. Baldwin, R.L., *Energetics of protein folding.* J Mol Biol, 2007. **371**(2): p. 283-301.
16. Baldwin, R.L., *Weak Interactions in Protein Folding: Hydrophobic Free Energy, van der Waals Interactions, Peptide Hydrogen Bonds, and Peptide Solvation.* Protein Folding Handbook, 2005: p. 127-162.
17. Baldwin, R.L., *Properties of hydrophobic free energy found by gas-liquid transfer.* Proc Natl Acad Sci U S A, 2013. **110**(5): p. 1670-1673.
18. Dill, K.A., *Dominant forces in protein folding.* Biochemistry, 1990. **29**(31): p. 7133-55.
19. Kim, P.S. and R.L. Baldwin, *Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding.* Annu Rev Biochem, 1982. **51**: p. 459-89.
20. Karplus, M. and D.L. Weaver, *Protein-folding dynamics.* Nature, 1976. **260**(5550): p. 404-6.
21. Karplus, M. and D.L. Weaver, *Protein folding dynamics: the diffusion-collision model and experimental data.* Protein Sci, 1994. **3**(4): p. 650-68.
22. Pace, C.N., et al., *Contribution of hydrophobic interactions to protein stability.* J Mol Biol, 2011. **408**(3): p. 514-28.
23. Kellis, J.T., Jr., et al., *Contribution of hydrophobic interactions to protein stability.* Nature, 1988. **333**(6175): p. 784-6.
24. Fersht, A.R., *Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications.* Proc Natl Acad Sci U S A, 1995. **92**(24): p. 10869-73.
25. Otzen, D.E., et al., *Structure of the transition state for the folding/unfolding of the barley chymotrypsin inhibitor 2 and its implications for mechanisms of protein folding.* Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10422-5.

26. Fersht, A.R., *Nucleation mechanisms in protein folding.* Curr Opin Struct Biol, 1997. **7**(1): p. 3-9.
27. Neira, J.L. and A.R. Fersht, *Exploring the folding funnel of a polypeptide chain by biophysical studies on protein fragments.* J Mol Biol, 1999. **285**(3): p. 1309-33.
28. Dobson, C.M. and M. Karplus, *The fundamentals of protein folding: bringing together theory and experiment.* Curr Opin Struct Biol, 1999. **9**(1): p. 92-101.
29. Bryngelson, J.D., et al., *Funnels, pathways, and the energy landscape of protein folding: a synthesis.* Proteins, 1995. **21**(3): p. 167-95.
30. Onuchic, J.N. and P.G. Wolynes, *Theory of protein folding.* Curr Opin Struct Biol, 2004. **14**(1): p. 70-5.
31. Friel, C.T., G.S. Beddard, and S.E. Radford, *Switching two-state to three-state kinetics in the helical protein Im9 via the optimisation of stabilising non-native interactions by design.* J Mol Biol, 2004. **342**(1): p. 261-73.
32. Mogensen, J.E., et al., *Elimination of a misfolded folding intermediate by a single point mutation.* Biochemistry, 2004. **43**(12): p. 3357-67.
33. Ferreon, A.C. and A.A. Deniz, *Protein folding at single-molecule resolution.* Biochim Biophys Acta, 2011. **1814**(8): p. 1021-9.
34. Shortle, D., *The denatured state (the other half of the folding equation) and its role in protein stability.* Faseb Journal, 1996. **10**(1): p. 27-34.
35. Wong, K.B., et al., *Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding.* J Mol Biol, 2000. **296**(5): p. 1257-82.
36. Frauenfelder, H., G.A. Petsko, and D. Tsernoglou, *Temperature-dependent X-ray diffraction as a probe of protein structural dynamics.* Nature, 1979. **280**(5723): p. 558-63.
37. Eisenmesser, E.Z., et al., *Intrinsic dynamics of an enzyme underlies catalysis.* Nature, 2005. **438**(7064): p. 117-21.
38. Henzler-Wildman, K.A., et al., *Intrinsic motions along an enzymatic reaction trajectory.* Nature, 2007. **450**(7171): p. 838-44.
39. Dunker, A.K., et al., *Function and structure of inherently disordered proteins.* Curr Opin Struct Biol, 2008. **18**(6): p. 756-64.
40. Uversky, V.N., C.J. Oldfield, and A.K. Dunker, *Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling.* J Mol Recognit, 2005. **18**(5): p. 343-84.
41. Fink, A.L., *Natively unfolded proteins.* Curr Opin Struct Biol, 2005. **15**(1): p. 35-41.
42. Dill, K.A. and D. Shortle, *Denatured states of proteins.* Annu Rev Biochem, 1991. **60**: p. 795-825.
43. Bowler, B.E., *Residual structure in unfolded proteins.* Curr Opin Struct Biol, 2012. **22**(1): p. 4-13.
44. Matthews, C.R., *Pathways of protein folding.* Annu Rev Biochem, 1993. **62**: p. 653-83.
45. Lee, J.C. and S.N. Timasheff, *Partial specific volumes and interactions with solvent components of proteins in guanidine hydrochloride.* Biochemistry, 1974. **13**(2): p. 257-65.
46. Creighton, T.E., *The energetic ups and downs of protein folding.* Nat Struct Biol, 1994. **1**(3): p. 135-8.
47. Anfinsen, C.B., et al., *The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain.* Proc Natl Acad Sci U S A, 1961. **47**: p. 1309-14.
48. Fersht, A.R., *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, in *Freeman and Company*1999.
49. Dyer, R.B., *Ultrafast and downhill protein folding.* Curr Opin Struct Biol, 2007. **17**(1): p. 38-47.
50. Liu, Z. and H.S. Chan, *Desolvation is a likely origin of robust enthalpic barriers to protein folding.* J Mol Biol, 2005. **349**(4): p. 872-89.
51. Plaxco, K.W., K.T. Simons, and D. Baker, *Contact order, transition state placement and the refolding rates of single domain proteins.* J Mol Biol, 1998. **277**(4): p. 985-94.

52. Fink, A.L., *Compact intermediates states in protein folding.* Subcell Biochem, 1995. **24**: p. 27-53.
53. Fujiwara, K., et al., *Folding-unfolding equilibrium and kinetics of equine beta-lactoglobulin: equivalence between the equilibrium molten globule state and a burst-phase folding intermediate.* Biochemistry, 1999. **38**(14): p. 4455-63.
54. Arai, M. and K. Kuwajima, *Rapid formation of a molten globule intermediate in refolding of alpha-lactalbumin.* Fold Des, 1996. **1**(4): p. 275-87.
55. Arai, M. and K. Kuwajima, *Role of the molten globule state in protein folding.* Adv Protein Chem, 2000. **53**: p. 209-82.
56. Ptitsyn, O.B., *How the molten globule became.* Trends Biochem Sci, 1995. **20**(9): p. 376-9.
57. Blancas-Mejia, L.M. and M. Ramirez-Alvarado, *Systemic amyloidoses.* Annu Rev Biochem, 2013. **82**: p. 745-74.
58. Wilson, M.R., J.J. Yerbury, and S. Poon, *Potential roles of abundant extracellular chaperones in the control of amyloid formation and toxicity.* Mol Biosyst, 2008. **4**(1): p. 42-52.
59. Soto, C., *Protein misfolding and disease; protein refolding and therapy.* FEBS Lett, 2001. **498**(2-3): p. 204-7.
60. Dobson, C.M., *Principles of protein folding, misfolding and aggregation.* Semin Cell Dev Biol, 2004. **15**(1): p. 3-16.
61. Invernizzi, G., et al., *Protein aggregation: mechanisms and functional consequences.* Int J Biochem Cell Biol, 2012. **44**(9): p. 1541-54.
62. Buchner, J., et al., *Alternatively folded states of an immunoglobulin.* Biochemistry, 1991. **30**(28): p. 6922-9.
63. Lilie, H. and J. Buchner, *Domain interactions stabilize the alternatively folded state of an antibody Fab fragment.* FEBS Lett, 1995. **362**(1): p. 43-6.
64. Thies, M.J., et al., *The alternatively folded state of the antibody C(H)3 domain.* J Mol Biol, 2001. **309**(5): p. 1077-85.
65. Kanmert, D., et al., *Thermal induction of an alternatively folded state in human IgG-Fc.* Biochemistry, 2011. **50**(6): p. 981-8.
66. Zimmerman, S.B. and S.O. Trach, *Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli.* J Mol Biol, 1991. **222**(3): p. 599-620.
67. Seckler, R. and R. Jaenicke, *Protein folding and protein refolding.* Faseb Journal, 1992. **6**(8): p. 2545-52.
68. Walter, S. and J. Buchner, *Molecular chaperones--cellular machines for protein folding.* Angew Chem Int Ed Engl, 2002. **41**(7): p. 1098-113.
69. Lindquist, S. and E.A. Craig, *The heat-shock proteins.* Annu Rev Genet, 1988. **22**: p. 631-77.
70. Young, J.C., et al., *Pathways of chaperone-mediated protein folding in the cytosol.* Nat Rev Mol Cell Biol, 2004. **5**(10): p. 781-91.
71. Haslbeck, M., et al., *Some like it hot: the structure and function of small heat-shock proteins.* Nat Struct Mol Biol, 2005. **12**(10): p. 842-6.
72. Horwich, A.L. and W.A. Fenton, *Chaperonin-mediated protein folding: using a central cavity to kinetically assist polypeptide chain folding.* Q Rev Biophys, 2009. **42**(2): p. 83-116.
73. Frydman, J., et al., *Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones.* Nature, 1994. **370**(6485): p. 111-7.
74. Wandinger, S.K., K. Richter, and J. Buchner, *The Hsp90 chaperone machinery.* J Biol Chem, 2008. **283**(27): p. 18473-7.
75. Gupta, D. and N. Tuteja, *Chaperones and foldases in endoplasmic reticulum stress signaling in plants.* Plant Signal Behav, 2011. **6**(2): p. 232-6.
76. Zimmermann, R., et al., *Protein translocation across the ER membrane.* Biochim Biophys Acta, 2011. **1808**(3): p. 912-24.

References

77. Fra, A.M., et al., *Quality control of ER synthesized proteins: an exposed thiol group as a three-way switch mediating assembly, retention and degradation.* EMBO J, 1993. **12**(12): p. 4755-61.
78. Hellman, R., et al., *The in vivo association of BiP with newly synthesized proteins is dependent on the rate and stability of folding and not simply on the presence of sequences that can bind to BiP.* J Cell Biol, 1999. **144**(1): p. 21-30.
79. Kozutsumi, Y., et al., *The presence of malfolded proteins in the endoplasmic reticulum signals the induction of glucose-regulated proteins.* Nature, 1988. **332**(6163): p. 462-4.
80. Rutkowski, D.T. and R.S. Hegde, *Regulation of basal cellular physiology by the homeostatic unfolded protein response.* J Cell Biol, 2010. **189**(5): p. 783-94.
81. Woehlbier, U. and C. Hetz, *Modulating stress responses by the UPRosome: a matter of life and death.* Trends Biochem Sci, 2011. **36**(6): p. 329-37.
82. Su, Q., et al., *Modulation of the eukaryotic initiation factor 2 alpha-subunit kinase PERK by tyrosine phosphorylation.* J Biol Chem, 2008. **283**(1): p. 469-75.
83. Williams, A.F. and A.N. Barclay, *The immunoglobulin superfamily--domains for cell surface recognition.* Annu Rev Immunol, 1988. **6**: p. 381-405.
84. Feige, M.J. and J. Buchner, *Principles and engineering of antibody folding and assembly.* Biochim Biophys Acta, 2014. **1844**(11): p. 2024-2031.
85. Xu, Z., et al., *Immunoglobulin class-switch DNA recombination: induction, targeting and beyond.* Nat Rev Immunol, 2012. **12**(7): p. 517-31.
86. Casadevall, A. and A. Janda, *Immunoglobulin isotype influences affinity and specificity.* Proc Natl Acad Sci U S A, 2012. **109**(31): p. 12272-3.
87. Schroeder, H.W., Jr. and L. Cavacini, *Structure and function of immunoglobulins.* J Allergy Clin Immunol, 2010. **125**(2 Suppl 2): p. S41-52.
88. Muller, R., et al., *High-resolution structures of the IgM Fc domains reveal principles of its hexamer formation.* Proc Natl Acad Sci U S A, 2013. **110**(25): p. 10183-8.
89. Thomas, H.I. and P. Morgan-Capner, *The avidity of specific IgM detected in primary rubella and reinfection.* Epidemiol Infect, 1990. **104**(3): p. 489-97.
90. Boes, M., *Role of natural and immune IgM antibodies in immune responses.* Mol Immunol, 2000. **37**(18): p. 1141-9.
91. KMea, M., *Janeway's Immunobiology.* Garland Science, 2008. **7th edition**.
92. Huber, R., et al., *Crystallographic structure studies of an IgG molecule and an Fc fragment.* Nature, 1976. **264**(5585): p. 415-20.
93. Kabat, E.A., T.T. Wu, and H. Bilofsky, *Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites.* J Biol Chem, 1977. **252**(19): p. 6609-16.
94. Raghavan, M. and P.J. Bjorkman, *Fc receptors and their interactions with immunoglobulins.* Annu Rev Cell Dev Biol, 1996. **12**: p. 181-220.
95. Arnold, J.N., et al., *The impact of glycosylation on the biological function and structure of human immunoglobulins.* Annu Rev Immunol, 2007. **25**: p. 21-50.
96. Saphire, E.O., et al., *Contrasting IgG structures reveal extreme asymmetry and flexibility.* J Mol Biol, 2002. **319**(1): p. 9-18.
97. Feige, M.J., et al., *Influence of the internal disulfide bridge on the folding pathway of the CL antibody domain.* J Mol Biol, 2007. **365**(4): p. 1232-44.
98. Buss, N.A., et al., *Monoclonal antibody therapeutics: history and future.* Curr Opin Pharmacol, 2012. **12**(5): p. 615-22.
99. Chacko, A.M., et al., *Targeted delivery of antibody-based therapeutic and imaging agents to CNS tumors: crossing the blood-brain barrier divide.* Expert Opin Drug Deliv, 2013. **10**(7): p. 907-26.

100. Cheng, B., et al., *Inhibiting toxic aggregation of amyloidogenic proteins: a therapeutic strategy for protein misfolding diseases.* Biochim Biophys Acta, 2013. **1830**(10): p. 4860-71.
101. Bork, P., L. Holm, and C. Sander, *The immunoglobulin fold. Structural classification, sequence patterns and common core.* J Mol Biol, 1994. **242**(4): p. 309-20.
102. Wei, J. and L.M. Hendershot, *Protein folding and assembly in the endoplasmic reticulum.* EXS, 1996. **77**: p. 41-55.
103. Montero, D., et al., *Intracellular glutathione pools are heterogeneously concentrated.* Redox Biol, 2013. **1**: p. 508-13.
104. Helenius, A. and M. Aebi, *Roles of N-linked glycans in the endoplasmic reticulum.* Annu Rev Biochem, 2004. **73**: p. 1019-49.
105. Fowler, S.B. and J. Clarke, *Mapping the folding pathway of an immunoglobulin domain: structural detail from Phi value analysis and movement of the transition state.* Structure, 2001. **9**(5): p. 355-66.
106. Cota, E., et al., *The folding nucleus of a fibronectin type III domain is composed of core residues of the immunoglobulin-like fold.* J Mol Biol, 2001. **305**(5): p. 1185-94.
107. Hamill, S.J., A. Steward, and J. Clarke, *The folding of an immunoglobulin-like Greek key protein is defined by a common-core nucleus and regions constrained by topology.* J Mol Biol, 2000. **297**(1): p. 165-78.
108. Feige, M.J., et al., *Dissecting the alternatively folded state of the antibody Fab fragment.* J Mol Biol, 2010. **399**(5): p. 719-30.
109. Goto, Y. and K. Hamaguchi, *Unfolding and refolding of the reduced constant fragment of the immunoglobulin light chain. Kinetic role of the intrachain disulfide bond.* J Mol Biol, 1982. **156**(4): p. 911-26.
110. Goto, Y. and K. Hamaguchi, *Unfolding and refolding of the constant fragment of the immunoglobulin light chain.* J Mol Biol, 1982. **156**(4): p. 891-910.
111. Feige, M.J., S. Walter, and J. Buchner, *Folding mechanism of the CH2 antibody domain.* J Mol Biol, 2004. **344**(1): p. 107-18.
112. Thies, M.J., et al., *Folding and association of the antibody domain CH3: prolyl isomerization preceeds dimerization.* J Mol Biol, 1999. **293**(1): p. 67-79.
113. Baumal, R., et al., *The regulation of immunoglobulin synthesis and assembly.* Ann N Y Acad Sci, 1971. **190**: p. 235-49.
114. Hendershot, A.J., *Immunoglobulin assembly and secretion.* Molecular Biology of B cells, 2004: Academic Pr Inc.
115. Chames, P., et al., *Therapeutic antibodies: successes, limitations and hopes for the future.* Br J Pharmacol, 2009. **157**(2): p. 220-33.
116. Chen, L., et al., *Anti-hepatoma human single-chain Fv antibody and adriamycin conjugates with potent antitumor activity.* Int Immunopharmacol, 2014. **18**(1): p. 20-6.
117. Nicholson, I.C., et al., *Construction and characterisation of a functional CD19 specific single chain Fv fragment for immunotherapy of B lineage leukaemia and lymphoma.* Mol Immunol, 1997. **34**(16-17): p. 1157-65.
118. Worn, A. and A. Pluckthun, *Stability engineering of antibody single-chain Fv fragments.* J Mol Biol, 2001. **305**(5): p. 989-1010.
119. Harris, L.J., S.B. Larson, and A. McPherson, *Comparison of intact antibody structures and the implications for effector function.* Adv Immunol, 1999. **72**: p. 191-208.
120. Wang, T. and Y. Duan, *Probing the stability-limiting regions of an antibody single-chain variable fragment: a molecular dynamics simulation study.* Protein Eng Des Sel, 2011. **24**(9): p. 649-57.
121. Vargas-Madrazo, E. and E. Paz-Garcia, *An improved model of association for VH-VL immunoglobulin domains: asymmetries between VH and VL in the packing of some interface residues.* J Mol Recognit, 2003. **16**(3): p. 113-20.

References

122. Narayanan, A., B.D. Sellers, and M.P. Jacobson, *Energy-based analysis and prediction of the orientation between light- and heavy-chain antibody variable domains.* J Mol Biol, 2009. **388**(5): p. 941-53.
123. Nakanishi, T., et al., *Critical contribution of VH-VL interaction to reshaping of an antibody: the case of humanization of anti-lysozyme antibody, HyHEL-10.* Protein Sci, 2008. **17**(2): p. 261-70.
124. Chatellier, J., et al., *Functional mapping of conserved residues located at the VL and VH domain interface of a Fab.* J Mol Biol, 1996. **264**(1): p. 1-6.
125. Khalifa, M.B., et al., *Effects on interaction kinetics of mutations at the VH-VL interface of Fabs depend on the structural context.* J Mol Recognit, 2000. **13**(3): p. 127-39.
126. Ewert, S., A. Honegger, and A. Pluckthun, *Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering.* Methods, 2004. **34**(2): p. 184-99.
127. Honegger, A., *Engineering antibodies for stability and efficient folding.* Handb Exp Pharmacol, 2008(181): p. 47-68.
128. Kugler, M., et al., *Stabilization and humanization of a single-chain Fv antibody fragment specific for human lymphocyte antigen CD19 by designed point mutations and CDR-grafting onto a human framework.* Protein Eng Des Sel, 2009. **22**(3): p. 135-47.
129. Nieba, L., et al., *Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment.* Protein Eng, 1997. **10**(4): p. 435-44.
130. Weidenhaupt, M., et al., *Functional mapping of conserved, surface-exposed charges of antibody variable domains.* J Mol Recognit, 2002. **15**(2): p. 94-103.
131. Ewert, S., et al., *Biophysical properties of human antibody variable domains.* J Mol Biol, 2003. **325**(3): p. 531-53.
132. Abhinandan, K.R. and A.C. Martin, *Analysis and prediction of VH/VL packing in antibodies.* Protein Eng Des Sel, 2010. **23**(9): p. 689-97.
133. Chothia, C., I. Gelfand, and A. Kister, *Structural determinants in the sequences of immunoglobulin variable domain.* J Mol Biol, 1998. **278**(2): p. 457-79.
134. Honegger, A. and A. Pluckthun, *The influence of the buried glutamine or glutamate residue in position 6 on the structure of immunoglobulin variable domains.* J Mol Biol, 2001. **309**(3): p. 687-99.
135. Honegger, A., et al., *The influence of the framework core residues on the biophysical properties of immunoglobulin heavy chain variable domains.* Protein Eng Des Sel, 2009. **22**(3): p. 121-34.
136. Rothlisberger, D., A. Honegger, and A. Pluckthun, *Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability.* J Mol Biol, 2005. **347**(4): p. 773-89.
137. Hugo, N., et al., *Functional aspects of co-variant surface charges in an antibody fragment.* Protein Sci, 2002. **11**(11): p. 2697-705.
138. Tan, P.H., B.M. Sandmaier, and P.S. Stayton, *Contributions of a highly conserved VH/VL hydrogen bonding interaction to scFv folding stability and refolding efficiency.* Biophys J, 1998. **75**(3): p. 1473-82.
139. Dwek, R.A., A.C. Lellouch, and M.R. Wormald, *Glycobiology: 'the function of sugar in the IgG molecule'.* J Anat, 1995. **187 (Pt 2)**: p. 279-92.
140. Krapp, S., et al., *Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity.* J Mol Biol, 2003. **325**(5): p. 979-89.
141. Voynov, V., et al., *Dynamic fluctuations of protein-carbohydrate interactions promote protein aggregation.* PLoS One, 2009. **4**(12): p. e8425.
142. Wu, H., et al., *Competing aggregation pathways for monoclonal antibodies.* FEBS Lett, 2014. **588**(6): p. 936-41.

143. Chennamsetty, N., et al., *Aggregation-prone motifs in human immunoglobulin G.* J Mol Biol, 2009. **391**(2): p. 404-13.
144. Chennamsetty, N., et al., *Design of therapeutic proteins with enhanced stability.* Proc Natl Acad Sci U S A, 2009. **106**(29): p. 11937-42.
145. Jefferis, R., *Glycosylation as a strategy to improve antibody-based therapeutics.* Nat Rev Drug Discov, 2009. **8**(3): p. 226-34.
146. Kontermann, R.E., *Strategies for extended serum half-life of protein therapeutics.* Curr Opin Biotechnol, 2011. **22**(6): p. 868-76.
147. Fast, J.L., et al., *Physical instability of a therapeutic Fc fusion protein: domain contributions to conformational and colloidal stability.* Biochemistry, 2009. **48**(49): p. 11724-36.
148. Hendershot, L., et al., *Assembly and secretion of heavy chains that do not associate posttranslationally with immunoglobulin heavy chain-binding protein.* J Cell Biol, 1987. **104**(3): p. 761-7.
149. Porter, R.R., *Structural studies of immunoglobulins.* Science, 1973. **180**(4087): p. 713-6.
150. Kaloff, C.R. and I.G. Haas, *Coordination of immunoglobulin chain folding and immunoglobulin chain assembly is essential for the formation of functional IgG.* Immunity, 1995. **2**(6): p. 629-37.
151. Vanhove, M., Y.K. Usherwood, and L.M. Hendershot, *Unassembled Ig heavy chains do not cycle from BiP in vivo but require light chains to trigger their release.* Immunity, 2001. **15**(1): p. 105-14.
152. Lee, Y.K., et al., *BiP and immunoglobulin light chain cooperate to control the folding of heavy chain and ensure the fidelity of immunoglobulin assembly.* Mol Biol Cell, 1999. **10**(7): p. 2209-19.
153. Knarr, G., et al., *BiP binding sequences in antibodies.* J Biol Chem, 1995. **270**(46): p. 27589-94.
154. Melchers, F., *The secretion of a Bence-Jones type light chain from a mouse plasmacytoma.* Eur J Immunol, 1971. **1**(5): p. 330-5.
155. Melchers, F., *Biosynthesis, transport and secretion of immunoglobulin in plasma cells.* Histochem J, 1971. **3**(5): p. 389-97.
156. Wolfenstein-Todel, C., E. Mihaesco, and B. Frangione, *"Alpha chain disease" protein def: internal deletion of a human immunoglobulin A1 heavy chain.* Proc Natl Acad Sci U S A, 1974. **71**(3): p. 974-8.
157. Marcinowski, M., et al., *Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions.* Nat Struct Mol Biol, 2011. **18**(2): p. 150-8.
158. Feige, M.J., et al., *The structure of a folding intermediate provides insight into differences in immunoglobulin amyloidogenicity.* Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13373-8.
159. Patten, P.A., et al., *The immunological evolution of catalysis.* Science, 1996. **271**(5252): p. 1086-91.
160. Kayser, V., et al., *Tryptophan-tryptophan energy transfer and classification of tryptophan residues in proteins using a therapeutic monoclonal antibody as a model.* J Fluoresc, 2011. **21**(1): p. 275-88.
161. Easterbrook-Smith, S.B. and R.A. Dwek, *The use of ANS fluorescence as a probe for immunoglobulin flexibility.* FEBS Lett, 1980. **121**(2): p. 253-6.
162. Bolen, D.W. and M.M. Santoro, *Unfolding free energy changes determined by the linear extrapolation method. 2. Incorporation of delta G degrees N-U values in a thermodynamic cycle.* Biochemistry, 1988. **27**(21): p. 8069-74.
163. Patel, T.R., D.J. Winzor, and D.J. Scott, *Analytical ultracentrifugation: A versatile tool for the characterisation of macromolecular complexes in solution.* Methods, 2016. **95**: p. 55-61.
164. Schuck, P., *Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling.* Biophys J, 2000. **78**(3): p. 1606-19.

165. Brown, P.H., et al., *Density contrast sedimentation velocity for the determination of protein partial-specific volumes.* Plos One, 2011. **6**(10): p. e26221.
166. Herold, E.M., *Folding and association of antibody domains*, in *Chemie*2013, TU München: Biotechnologie.
167. Kelly, S.M., T.J. Jess, and N.C. Price, *How to study proteins by circular dichroism.* Biochim Biophys Acta, 2005. **1751**(2): p. 119-39.
168. Graille, M., et al., *Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5399-404.
169. Graille, M., et al., *Complex between Peptostreptococcus magnus protein L and a human antibody reveals structural convergence in the interaction modes of Fab binding proteins.* Structure, 2001. **9**(8): p. 679-87.
170. Artandi, S.E., et al., *Molecular analysis of IgM rheumatoid factor binding to chimeric IgG.* J Immunol, 1991. **146**(2): p. 603-10.
171. Bonagura, V.R., et al., *Mapping IgG epitopes bound by rheumatoid factors from immunized controls identifies disease-specific rheumatoid factors produced by patients with rheumatoid arthritis.* J Immunol, 1998. **160**(5): p. 2496-505.
172. Corper, A.L., et al., *Structure of human IgM rheumatoid factor Fab bound to its autoantigen IgG Fc reveals a novel topology of antibody-antigen interaction.* Nat Struct Biol, 1997. **4**(5): p. 374-81.
173. Chothia, C., et al., *Domain association in immunoglobulin molecules. The packing of variable domains.* J Mol Biol, 1985. **186**(3): p. 651-63.
174. Morea, V., et al., *Conformations of the third hypervariable region in the VH domain of immunoglobulins.* J Mol Biol, 1998. **275**(2): p. 269-94.
175. Decanniere, K., S. Muyldermans, and L. Wyns, *Canonical antigen-binding loop structures in immunoglobulins: more structures, more canonical classes?* J Mol Biol, 2000. **300**(1): p. 83-91.
176. Gasymov, O.K. and B.J. Glasgow, *ANS fluorescence: potential to augment the identification of the external binding sites of proteins.* Biochim Biophys Acta, 2007. **1774**(3): p. 403-11.
177. Feige, M.J., et al., *The structural analysis of shark IgNAR antibodies reveals evolutionary principles of immunoglobulins.* Proc Natl Acad Sci U S A, 2014. **111**(22): p. 8155-60.
178. Ewert, S., et al., *Biophysical properties of camelid V(HH) domains compared to those of human V(H)3 domains.* Biochemistry, 2002. **41**(11): p. 3628-36.
179. Streltsov, V.A., et al., *Structural evidence for evolution of shark Ig new antigen receptor variable domain antibodies from a cell-surface receptor.* Proc Natl Acad Sci U S A, 2004. **101**(34): p. 12444-9.
180. Kovalenko, O.V., et al., *Atypical antigen recognition mode of a shark immunoglobulin new antigen receptor (IgNAR) variable domain characterized by humanization and structural analysis.* J Biol Chem, 2013. **288**(24): p. 17408-19.
181. Muyldermans, S., *Nanobodies: natural single-domain antibodies.* Annu Rev Biochem, 2013. **82**: p. 775-97.
182. Davies, J. and L. Riechmann, *'Camelising' human antibody fragments: NMR studies on VH domains.* FEBS Lett, 1994. **339**(3): p. 285-90.
183. Harmsen, M.M. and H.J. De Haard, *Properties, production, and applications of camelid single-domain antibody fragments.* Appl Microbiol Biotechnol, 2007. **77**(1): p. 13-22.
184. Perchiacca, J.M., M. Bhattacharya, and P.M. Tessier, *Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions.* Proteins, 2011. **79**(9): p. 2637-47.
185. Isenman, D.E., D. Lancet, and I. Pecht, *Folding pathways of immunoglobulin domains. The folding kinetics of the Cgamma3 domain of human IgG1.* Biochemistry, 1979. **18**(15): p. 3327-36.

186. Bertz, M., J. Buchner, and M. Rief, *Mechanical stability of the antibody domain CH3 homodimer in different oxidation states.* J Am Chem Soc, 2013. **135**(40): p. 15085-91.
187. Horne, C., et al., *Noncovalent association of heavy and light chains of human immunoglobulins. III. Specific interactions between VH and VL.* J Immunol, 1982. **129**(2): p. 660-4.
188. Jung, S. and A. Pluckthun, *Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting.* Protein Eng, 1997. **10**(8): p. 959-66.
189. Helms, L.R. and R. Wetzel, *Destabilizing loop swaps in the CDRs of an immunoglobulin VL domain.* Protein Sci, 1995. **4**(10): p. 2073-81.
190. Vlasov, A.P., Z.I. Kravchuk, and S.P. Martsev, *[Non-native conformational states of immunoglobulins: thermodynamic and functional analysis of rabbit IgG].* Biokhimiia, 1996. **61**(2): p. 212-35.
191. Ptitsyn, O.B., *Molten globule and protein folding.* Adv Protein Chem, 1995. **47**: p. 83-229.
192. Dimitrov, D.S., *Engineered CH2 domains (nanoantibodies).* MAbs, 2009. **1**(1): p. 26-8.
193. Gong, R., et al., *Shortened engineered human antibody CH2 domains: increased stability and binding to the human neonatal Fc receptor.* J Biol Chem, 2011. **286**(31): p. 27288-93.
194. Gong, R., et al., *N-terminal truncation of an isolated human IgG1 CH2 domain significantly increases its stability and aggregation resistance.* Mol Pharm, 2013. **10**(7): p. 2642-52.
195. Chen, W., et al., *Stabilizing the CH2 Domain of an Antibody by Engineering in an Enhanced Aromatic Sequon.* ACS Chem Biol, 2016. **11**(7): p. 1852-61.
196. Gong, R., et al., *Engineered human antibody constant domains with increased stability.* J Biol Chem, 2009. **284**(21): p. 14203-10.
197. Nokwe, C.N., et al., *The Antibody Light-Chain Linker Is Important for Domain Stability and Amyloid Formation.* J Mol Biol, 2015. **427**(22): p. 3572-86.
198. Hamers-Casterman, C., et al., *Naturally occurring antibodies devoid of light chains.* Nature, 1993. **363**(6428): p. 446-8.
199. Nguyen, V.K., et al., *Heavy-chain antibodies in Camelidae; a case of evolutionary innovation.* Immunogenetics, 2002. **54**(1): p. 39-47.
200. Nguyen, V.K., et al., *Camel heavy-chain antibodies: diverse germline V(H)H and specific mechanisms enlarge the antigen-binding repertoire.* EMBO J, 2000. **19**(5): p. 921-30.
201. Nguyen, V.K., S. Muyldermans, and R. Hamers, *The specific variable domain of camel heavy-chain antibodies is encoded in the germline.* J Mol Biol, 1998. **275**(3): p. 413-8.
202. Muyldermans, S., et al., *Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains.* Protein Eng, 1994. **7**(9): p. 1129-35.
203. Conrath, K.E., et al., *Emergence and evolution of functional heavy-chain antibodies in Camelidae.* Dev Comp Immunol, 2003. **27**(2): p. 87-103.
204. Decanniere, K., et al., *A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops.* Structure, 1999. **7**(4): p. 361-70.
205. Nguyen, V.K., et al., *Loss of splice consensus signal is responsible for the removal of the entire C(H)1 domain of the functional camel IGG2A heavy-chain antibodies.* Mol Immunol, 1999. **36**(8): p. 515-24.
206. Sakano, H., et al., *Domains and the hinge region of an immunoglobulin heavy chain are encoded in separate DNA segments.* Nature, 1979. **277**(5698): p. 627-33.
207. Padlan, E.A., *Anatomy of the antibody molecule.* Mol Immunol, 1994. **31**(3): p. 169-217.
208. Cogne, M., J.L. Preud'homme, and P. Guglielmi, *Immunoglobulin gene alterations in human heavy chain diseases.* Res Immunol, 1989. **140**(5-6): p. 487-502.
209. Franklin, E.C., et al., *Heavy Chain Disease- a New Disorder of Serum Gamma-Globulins : Report of the First Case.* Am J Med, 1964. **37**: p. 332-50.
210. Khamlichi, A.A., et al., *Structure of abnormal heavy chains in human heavy-chain-deposition disease.* Eur J Biochem, 1995. **229**(1): p. 54-60.

References

211. Robinson, M.P., et al., *Efficient expression of full-length antibodies in the cytoplasm of engineered bacteria.* Nat Commun, 2015. **6**: p. 8072.
212. Kuo, D., M. Nie, and A.J. Courey, *SUMO as a solubility tag and in vivo cleavage of SUMO fusion proteins with Ulp1.* Methods Mol Biol, 2014. **1177**: p. 71-80.
213. Braakman, I., J. Helenius, and A. Helenius, *Manipulating disulfide bond formation and protein folding in the endoplasmic reticulum.* EMBO J, 1992. **11**(5): p. 1717-22.
214. Kiefhaber, T., *Protein folding kinetics.* Methods Mol Biol, 1995. **40**: p. 313-41.

Acknowledgement

Zu Beginn möchte ich mich bei meinem Betreuer Prof. Johannes Buchner bedanken, für die Möglichkeit am Biotechnologie Lehrstuhl, mit seinem zahlreichen Angebot an analytischen Methoden und Möglichkeiten, zu promovieren. Der nächste Dank gilt den Ideengebern einiger meiner Projekte, Eva Herold, George Stan und besonders Matthias Feige.

Weiterhin möchte ich mich bei den vielen derzeitigen und früheren Laborkollegen für eine schöne Zeit im Labor 6 bedanken: Eva Herold, Natalia Sarmiento, Matthias Rosam, Ruby Khan, Roger Müller, Christina Stutzer und Benedikt Weber.

Besonderer Dank geht an meine Vorgängerin Eva Herold, die mir immer mit Rat und Tat zur Seite stand.

Des Weiteren danke ich dem gesamten Lehrstuhl für die schöne, gemeinsame Zeit und unterhaltsame Mittagspausen. Danke für viele lustige Abende mit den Prosecco Mädels, Katrin, Eva und Betty sowie unterhaltsame Pausen mit den Kaffeemädels, Chrissy, Marina, Katrin, Patzi, Bine und Priyanka.

Ein großer Dank geht an das Büro 2, dafür dass ihr mich am Anfang, zu Zeiten räumlicher Knappheit, aufgenommen habt. Danke an meine Büromädels Marina und Katrin, dafür dass ihr in guten wie in schlechten Zeiten immer da wart (je nach Bedarf mit guten Ratschlägen, aufmunternden Worten oder einem Bier/Cuba). Liebe Marina, wenn Katrin und ich weg sind kümmere dich bitte um die Pflanzen und den Moritz ;)

Als letztes möchte ich meiner Familie für die Unterstützung während des Studiums und der Promotion danken.

Acknowledgement

List of publications

Nokwe CN, Hora M, Zacharias M, Yagi H, John C, Reif B, Goto Y, Buchner J.
The Antibody Light-Chain Linker Is Important for Domain Stability and Amyloid Formation.
J Mol Biol. 2015 Nov ; 427(22):3572-86

Janina Jamasbi, Remco T.A. Megens, Mariaelvy Bianchini, Kerstin Uhland, Götz Münch, Martin Ungerer, Shachar Sherman, Alexander Faussner, Richard Brandl, Christine John, Johannes Buchner, Christian Weber, Reinhard Lorenz, Natalie Elia, Wolfgang Siess
Cross-Linking GPVI-Fc by Anti-Fc Antibodies Potentiates Its Inhibition of Atherosclerotic Plaque- and Collagen-Induced Platelet Activation; JACC: basic to translational science, 2016 April

Allosteric Regulation Points Control the Conformational Dynamics of the Molecular Chaperone Hsp90.
Rehn A, Moroni E, Zierer BK, Tippel F, Morra G, John C, Richter K, Colombo G, Buchner J.
J Mol Biol. 2016 Sep

Declaration

I, Christine John, hereby declare that this thesis was prepared by me independently. I was using only the references and resources stated here. The work has so far not been submitted to any audit commission.

Hiermit erkläre ich, Christine John, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die Arbeit wurde bisher keiner Prüfungskommission vorgelegt.

_____

Christine John

München,