The antibody light chain linker regulates domain orientation and amyloidogenicity

Benedikt Weber[1], Manuel Hora[1], Pamina Kazman[1], Christoph Göbl[1,2], Carlo Camilloni[3], Bernd Reif[1], Johannes Buchner[1]*

Running title: Linker regulation of antibody light chain

[1]*Center for Integrated Protein Science Munich at the Department Chemie, Technische Universität München, Lichtenbergstr. 4, 85748 Garching, Germany*

[2]*Helmholtz Zentrum München, Institute of Structural Biology, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany*

[3]*Dipartimento di Bioscienze, Università degli studi di Milano, 20133 Milano, Italy*

* To whom correspondence should be addressed: johannes.buchner@tum.de

phone: +498928913340; fax: +498928913345

Keywords: antibody folding, protein stability, light chain linker, amyloid, intramolecular interactions

Abstract

The antibody light chain (LC) consists of two domains and is essential for antigen binding in mature immunoglobulins. The two domains are connected by a highly conserved linker which comprises the structurally important Arg108 residue. In antibody light chain (AL) amyloidosis, a severe protein amyloid disease, the LC and its N-terminal variable domain (V_L) convert to fibrils deposited in the tissues causing organ failure. Understanding the factors shaping the architecture of the LC, is important for basic science, biotechnology and for deciphering the principles that lead to fibril formation. In this study, we examined the structure and properties of LC variants with a mutated or extended linker. We show that under destabilizing conditions, the linker modulates the amyloidogenicity of the LC. The fibril formation propensity of LC linker variants and their susceptibility to proteolysis directly correlate implying an interplay between the two LC domains. Using NMR and RDC-based simulations, we found that the linker residue Arg108 is a key factor regulating the relative orientation of the V_L and C_L domains, keeping them in a bent and dense, but still flexible conformation. Thus, inter-domain contacts and the relative orientation of V_L and C_L to each other are of major importance for maintaining the structural integrity of the full-length LC.

Introduction

The antibody light chain (LC) consists of a variable (V_L) and a constant (C_L) domain, connected by a short unstructured linker. Both, the V_L and C_L domain, share the highly conserved immunoglobulin (Ig) fold [1]. The antigen binding regions (CDRs) of the V_L domain are extremely variable due to V(D)J rearrangement and subsequent somatic hypermutation [2,3]. To buffer adverse effects of the sequence variability, the LC must exhibit an intrinsic high tolerance against misfolding by destabilizing residues [4]. However, in some cases, compensation and cellular protein quality control fail and in consequence certain mutations in the light chain result in systemic antibody light chain (AL) amyloidosis, a fatal protein misfolding disease [5]. Here, aberrant plasma cells produce high quantities of altered immunoglobulin light chains, which are secreted into the blood stream. These free serum light chains then undergo amyloid fibril formation [6]. Amyloid fibrils and their precursors accumulate in various tissues, ultimately leading to organ failure and death [7]. In the past, primarily N-terminal LC fragments, i.e. the V_L domain with parts of the truncated C_L, were found in fibril samples extracted from human tissues [8–12]. However, more recent data provide evidence for the presence of both, the full-length LC and the V_L in amyloid deposits, along with other proteins like apolipoprotein E [13–15]. In this context, an important role of protein dynamics, transient unfolding and proteolytic cleavage of LCs for fibril formation and the pathogenesis of AL patients has been suggested [11,16–19]. Interestingly, some of the mentioned factors are also associated with other amyloidogenic precursor proteins such as TTR [20]. In previous studies, we identified the linker residue R108 to be important for the conformational integrity and stability of both, the isolated V_L and C_L domains [21]. For V_L, R108 stabilizes the domain by increasing its solubility due to a solvent-exposed charged C-terminus. The isolated C_L benefits from intramolecular interactions of R108 with N-terminal loop residue, which protect the hydrophobic core of the domain. This raises the question how the LC modulates the properties of its individual domains (V_L and C_L) and how the LC linker is involved in AL amyloidosis. Klimtchuk et al. found a destabilizing effect of an amyloidogenic V_L domain on its $C_L\lambda$ in the full-length LC context, which initiated LC aggregation compared to its germline counterpart [15]. In contrast, Blancas-Mejía and coworkers showed that a $C_L\kappa$ does not significantly impact the stability of a full-length LC, but modulates its amyloidogenic properties and fibril morphology [22]. They hypothesized that rather structural features than the LC stability are likely to be responsible for amyloid formation. Of notice, recent studies concluded that thermodynamic stability can only be considered as one of several factors rendering LC proteins amyloidogenic [23,24]. Thus, the role of the linker and the individual domains in the LC remains vague. To address the influence of a kappa C_L domain and the LC linker on LC fibril formation, we analyzed V_L, C_L and LC variants with a focus on the linker including R108 and the communication between V_L

and C_L (Fig. 1). We characterized all variants with respect to their fold, conformational stability and fibril formation propensity to gain deeper insight in the importance of the LC linker and the presence of the C_L domain. We find that the linker directly modulates LC properties, particularly the relative orientation of the V_L and C_L domains and its fibril formation propensity.

Results

How the C_L domain and the LC linker affect the amyloid fibril formation propensity of amyloidogenic V_L variants in the context of the LC is still enigmatic. To address this question, we varied the linker connecting the two individual domains of a kappa LC. This approach allows us to assess the communication between the V_L and C_L domains. Since the presence of the LC linker residue R108 is important for both, V_L and C_L integrity [21], we substituted R108 by either alanine or glutamate to test the impact of this residue in the context of the LC. In addition, we created variants in which we introduced an additional glycine (LC link+G) or a glycine-alanine spacer (LC link+GAGAG) between R108 and A109 to allow a higher degree of LC linker flexibility. It was not possible to completely delete the linker as this variant aggregated quantitatively during refolding and purification. An overview of the constructs used in this study is given in figure 1.

Linker mutations do not alter the structure of the V_L, LC and C_L domains

To characterize the secondary, tertiary and quaternary structures of the V_L, LC and C_L variants, we performed CD spectroscopy and SEC-MALS experiments (Fig. 2). Far-UV CD measurements revealed a characteristic β-strand-rich structure for all variants indicated by a minimum at 218 nm (Fig. 2A). The recorded CD spectra match perfectly with the respective wild type spectra. Near-UV CD experiments were used to assess the tertiary structure (Fig. 2B). All spectra showed similar ellipticity characterized by a minimum at ~ 275 nm. We observed only minor changes between the variants, indicating a similar tertiary structure. Thus, the introduced mutations did not alter the conserved Ig fold. SEC-MALS, which allows determining absolute molar masses for isolated protein species, was used to assess the oligomeric state of the variants (Fig. 2C). All V_L, LC and C_L variants were monomeric (Table S1). For V_L wt, we found a minor fraction of larger species, most likely dimers, as a shoulder of the main peak. Due to the low concentration of this species, we could not determine its molecular mass. Thus, the quaternary structure of the LC is not affected by mutating R108 or extending the LC linker. Taken together, our data show that the introduced mutations do not alter the native structure of the variants.

The LC linker stabilizes the individual domains and the full-length LC

Next, we determined the conformational stability of the V_L, LC and C_L variants by thermal unfolding transitions (Fig. 3 and Table 1). V_L wt and V_L R108E are the most stable V_L variants exhibiting melting temperatures (T_ms) of 52 °C (Fig. 3 left). V_L R108A (T_m 51 °C) revealed a

slightly decreased conformational stability. The R108 deletion variant V_L ΔR108 (T_m 48 °C) showed the lowest T_m among the V_L variants, which is in agreement with the literature [21]. C_L wt and C_L R108A (both T_ms 54 °C) are the most stable proteins in this study (Fig. 3 right). The R108E (T_m 52 °C) mutation in C_L decreased the stability by ~ 2 °C, which is more pronounced than the impact of the R108A mutation in V_L (decrease of ~ 1 °C). Similar to the effect on the V_L variant, the deletion of R108 in C_L (T_m 50 °C) results in a ~ 4 °C lower conformational stability. Thus, for both individual domains, the R108 deletion variants, V_L ΔR108 and C_L ΔR108, reveal the highest decrease in conformational stability. For the full-length LC variants, we found similar trends (Fig. 3 middle). Again, LC wt (T_m 51 °C) showed the highest stability among the LC variants. LC R108A (T_m 50 °C) as well as the extension variants LC link+G (T_m 49 °C) and LC link+GAGAG (T_m 49 °C) are slightly less stable than the wild type, but more stable than LC R108E (T_m 48 °C). Interestingly, this suggests that the dislocation of R108 relative to the C_L domain by introducing a glycine or a glycine-alanine spacer has the same impact as the substitution of R108 by alanine. It was not possible to purify a LC linker deletion variant (LC Δlink) due to quantitative protein aggregation during refolding. Thus, mutating the important linker residue R108 to alanine has no significant impact on protein stability, neither for the individual domains V_L and C_L, nor for the full-length LC. Interestingly, the substitution of R108 by glutamate resulted in a T_m decrease of ~ 2 to 3 °C for both, LC R108E (T_m 48 °C) and the C_L R108E variant (T_m 52 °C). In contrast, V_L R108E (T_m 52 °C) is completely unaffected by this mutation. Changes in the length of the LC linker (LC link+G, LC link+GAGAG) negatively affected the stability of the LC. R108 substitutions by alanine or glutamate (LC R108A, LC R108E), however, did not uniformly impact the conformational stability of the LC. Interestingly, our data show that the stability of the LC is generally lower than the stability of the individual V_L and C_L domains. This is true for the wild type proteins as well as for the variants. For LC R108E, the decrease in conformational stability can be attributed to the individual C_L domain, whereas the slightly decreased stability of V_L R108A does not affect the LC R108A variant.

C_L preserves LC integrity and prevents amyloid fibril formation

To monitor fibril formation of the variants, we utilized the extrinsic fluorescent dye thioflavin T (ThT), as a probe for amyloid structures (Fig. 4A). Samples showing a positive ThT signal were further analyzed by transmission electron microscopy (TEM) (Fig. 4B). In all cases, except for LC wt + SDS, we observed fibrillar structures, proving the reliability of the ThT assay. Noteworthy, for LC wt + SDS, the ThT signal was substantially lower than for fibril-positive samples. For comparison of the kinetic profiles, we determined the time point when 50 % of the maximal ThT signal (t_{50}) was reached (Tables 2 and 3). After 15 days of

incubation under gentle shaking in PBS pH 7.4 buffer at 37°C, all V_L variants formed fibrils albeit with different kinetics. Both, V_L R108A (t_{50} 6.99 d) and V_L R108E (t_{50} 6.85 d) exhibited the fastest fibril formation kinetics. The R108 deletion variant V_L ΔR108 (t_{50} 7.52 d) formed fibrils only slightly slower than the R108 substitution variants V_L R108A and V_L R108E. Fibril formation of V_L wt (t_{50} 9.58 d) was ~ 2.5 d slower than that of the R108 substitution and the deletion variants. All V_L variants revealed partially altered fibril morphology and different degrees of fragmentation as observed by TEM analysis (Fig. 4B). Particularly, TEM micrographs of V_L ΔR108 fibrils showed strongly fragmented fibrillar clusters. In contrast, we observed no fibril formation for the C_L variants, including C_L wt, C_L R108A and C_L R108E (Fig. 4A) Even the destabilized variant C_L ΔR108 was resistant to amyloid formation under the given conditions.

To test how the presence of the C_L domain affects mutations shown to enhance V_L amyloidogenicity, we performed the ThT assay with LCs carrying the respective mutations (R108A, R108E). Surprisingly, none of the LC variants showed any ThT binding or fibrils in the TEM analysis when incubated in PBS pH 7.4 for up to 20 days (Fig. 4A). In contrast to V_L wt, LC wt is resilient against fibril formation under the experimental conditions. Although the mutation R108E decreases the T_m of the LC, it did not render the LC amyloidogenic. Furthermore, LC variants with an extended LC linker, LC link+G and LC link+GAGAG, did not form fibrils. To enable fibril formation of the LC and reveal differences between the variants, we performed further fibril formation kinetics under slightly destabilizing conditions using 0.5 mM SDS (submicellar concentration) added to PBS pH 7.4, as Yamamoto et al. had shown that 0.5 mM SDS is able to induce fibril formation in β-stand-rich proteins like β2-m which shares the same domain architecture with antibody domains [1,25]. By this approach, we were able to monitor amyloid fibrils using ThT fluorescence (Fig. 4A) and TEM (Fig. 4B). All LC variants, except LC wt, readily formed fibrils, however with different t_{50} (Table 3). LC link+GAGAG exhibited the fastest formation (t_{50} 0.96 d), closely followed by LC R108E (t_{50} 1.36 d). LC link+G (t_{50} 2.86) and LC R108A (t_{50} 3.54 d) revealed a lower fibril formation propensity. Our data indicate, that C_L wt is under non-destabilizing conditions intrinsically resistant against fibril formation, which is in agreement with the literature [26]. Moreover, destabilizing mutations in the linker do not have a negative impact on the integrity of the isolated C_L. In the full-length LC context, the C_L domain retains this resilience and expands it to V_L via the LC linker even in the presence of unfavorable destabilizing mutations. However, as shown by using destabilizing conditions, the LC variants revealed mutation-dependent fibril formation propensities suggesting a distinct impact of the linker.

Proteolysis susceptibility correlates with fibril formation propensity

Recent studies showed that the susceptibility of the LCs to proteolysis is modulated by their kinetic stabilities and the presence of non-native conformations [17,19]. For another amyloidogenic protein, TTR, it was shown that a specific mutation increases proteolytic cleavage and fibril formation propensity [20]. Therefore, we tested the structural integrity of the LC variants by limited proteolysis using proteinase K (Fig. 5A and S1). Interestingly, LC wt was most resistant against proteolytic cleavage, followed by LC R108A and LC link+G. LC R108E was substantially more sensitive to proteolysis than LC wt. For LC link+GAGAG, almost no residual full-length protein was observed after 10 min of incubation with proteinase K. Interestingly, the order of proteolysis susceptibility and the fibril formation propensity was identical for all LC variants. For the amyloidogenic LC variants, we found a linear correlation between both parameters with a Pearson R value of 0.979 (Fig. 5B). This strong correlation and the fact that we did not find a correlation with conformational stability (Fig. 3), indicate that a distinct structural feature induced by the introduced linker mutations renders the LC variants amyloidogenic.

R108 maintains orientational rigidity of the two LC domains

In order to reach a structural understanding of our results, we conducted NMR experiments. All variants produced spectra of folded proteins. Despite some minor chemical shift changes, all spectra were similar enough to confirm, that the variants adopt the characteristic Ig fold (Fig. S2 and S3). The largest chemical shift differences between V_L wt and LC wt NMR spectra were observed for T80 and G84 (Fig. 6A and B). These residues also displayed considerable chemical shift changes in the mutants LC R108A and LC R108E. Comparison of C_L wt and LC wt showed a prominent shift difference for T172. In a previous study, we found two distinct conformations for C_L wt, which resulted in two sets of NMR resonances [21]. In that study, we showed that the R108 deletion variant C_L ΔR108 caused a shift to the weakly populated C_L wt conformer. In LC wt, T172 yields a significant chemical shift difference in comparison to the C_L wt major peak. Still, the shift was much closer to the major population peak than to the minor population. This is supported by the LC wt N138 peak being very close to the main conformation of C_L wt. Similar observations were made for LC R108A. While we could observe two weak signals for N138 in this case, the position of the T172 peak indicates a conformation closer to the major state of C_L wt. In contrast, LC R108E exhibits chemical shifts for N138 and T172, which are in good agreement with the weakly populated conformation of C_L wt. In summary, the conformations of LC wt and LC R108A are more similar to the major population, albeit with some differences, while the R108E mutation

caused a similar conformational change as the C_L deletion ΔR108. Due to their close structural proximity, R108 likely interacts with T172 or D170 in LC wt, providing stabilizing contacts (Fig. 6C). However, the latter residue (D170) was not assigned in the NMR spectra. In the case of LC R108A, these favorable interactions would be absent. This would result in an increased flexibility of the D170 loop and a relaxation of the rigid orientation between both domains. In LC R108E, the attractive interactions might be replaced by repelling forces, thereby pushing the two individual domains apart.

The linker residue R108 keeps the LC in a closed conformation

In order to confirm our hypotheses about different domain orientations affected by the linker residues, we measured residual dipolar couplings (RDCs). These NMR parameters report on relative orientations of nuclear spins to an alignment medium [27]. We determined 67 RDCs for LC wt and 63 RDCs for LC R108A (Fig. S4 and Table S2). Our experiments yield different RDC values for LC wt and LC R108A. Since chemical shift differences between these variants were rather small, we consider it unlikely that the different RDCs result from changes in the domain structures. Instead, the RDCs likely report on different relative orientations of C_L and V_L with respect to each other.

To obtain a structural interpretation of the effect of the R108A mutation on the dynamics of the two LC domains we made use of Metadynamic Metainference (M&M) simulations and the theta-method for the interpretation of the experimentally derived RDCs. Two M&M simulations (cf. Methods section) were performed to compare the conformational space accessible for LC wt and LC R108A. The two resulting conformational ensembles (Fig. 7C) show that the single domains maintain a stable fold (in both cases we have average RMSDs of ~ 1.4 and ~ 1.8 Å for the backbone with respect to the crystal structure of V_L and C_L, respectively), while showing marked differences in the global dynamics. In Fig. 7A, the free energy surfaces are shown as a function of the relative orientation of the two domains defined as the two angles shown in the cartoon (phi and psi). Remarkably, while in LC wt and LC R108A the two domains can fully rotate around the linker (phi angle), the R108A mutation allows the two domains to adopt more open orientations to each other (psi angle), however with a preference for a relatively close orientation (psi 1.5 - 2.5 rad). For the phi angle, both systems prefer angles between -2 and 0 rad, as well as around 1 rad for LC wt. Of notice, the X-ray structure falls into the low free-energy region, adopting psi and phi angles of ~ 1.7 and -1, respectively (Fig. 7A, yellow square). From a structural perspective, R108 is able to form a salt-bridge with D170 (absent in the crystal structure) that is 75 % populated in the ensemble, while the R108A substitution increases the average distance between the

residues A108 and D170 from 5 Å to 8 Å. Consequently, a contact between these two residues is only 20 % populated in LC R108A. In contrast, the salt-bridge K103 - D165 formed in the crystal structure is only weakly populated in the solution ensembles (2 % and 8 % for LC wt and LC R108A, respectively).

The analysis of the contact probabilities shows a remarkable similarity between the contacts within the two domains (V_L and C_L) for both, LC wt and LC R108A, confirming the common interactions maintaining the Ig fold (Fig. 7B). In the case of LC R108A, a number of inter-domain contacts are lost, which is in agreement with the more open conformations accessible upon mutating R108 to A108. Noteworthy, the overall amount of inter-domain contacts in both LC variants is relatively low, highlighting the importance of the LC linker to covalently connect V_L and C_L.

Discussion

Since the individual V_L domain and full-length LC, have been found deposited in amyloid plaques, it is important to determine the amyloidogenic properties of V_L, LC and C_L in comparison [8,10,13–15]. Here, we analyzed the impact of the C_L domain on the V_L domain in the LC framework and specifically focused on the importance of the LC linker in this context.

Based on previous work on the isolated V_L and C_L domains, the presence of R108 in the linker between the two domains was suspected to be of crucial importance for the full-length LC [21]. Substituting the linker residue R108 by alanine, however, did not substantially affect the stability of V_L or LC. Since the positively charged side chain of R108 does not undergo distinct intramolecular interactions in V_L as judged from the crystal structure (PDB 1FH5, L chain), it was expected that its replacement by alanine does not affect the domain structure. Moreover, potentially missing electrostatic interactions due to the lack of R108 in LC R108A might also be compensated by K107, resulting in similar structural properties of the V_L variants as determined for V_L wt and V_L R108A. For LC R108A, however, we conclude that the decrease of polar interactions between R108 and its potential interaction partners Y140, D165, D170 and T172 in the C_L domain of the LC results in a less rigid orientation of V_L and C_L to each other. Our RDC-based simulations revealed that there are fewer transient inter-domain contacts between C_L and V_L for LC R108A than for LC wt. As a result, the structure of LC R108A exhibits increased flexibility, shifting towards opened conformations compared to LC wt. In general, the overall amount of inter-domain contacts in LC wt and LC R108A is relatively low and transient, highlighting the importance of the LC linker to covalently connect V_L and C_L. Substituting R108 by a negatively charged residue, here glutamate, reduced the stabilities of C_L and LC, but not that of V_L. In contrast to V_L, there are likely intramolecular interactions of R108 with the C_L domain, i.e. polar interactions with Y140, D165, D170 and T172, which anchor the exposed “bc” and “de” loops of C_L to the LC linker, thereby keeping V_L and C_L in close proximity. Interestingly, the deletion of R108 or its substitution by glutamate (C_L ΔR108, C_L R108E) result in a more pronounced decrease in stability than the substitution of R108 by alanine (C_L R108A). Thus, structurally necessary polar interactions of R108 can be maintained by the backbone carbonyl oxygen of A108 as implied by the RDC-based simulations. The introduction of a negative charge (E108) or deletion of residue 108 (ΔR108), however, leads to a repulsion of interaction partners. The same is true for the corresponding full-length LC variants as indicated by a similar decrease in T_m and NMR chemical shifts observed for N138 and T172 in LC R108E, which are similar to those of the destabilized variant C_L ΔR108 [21]. In contrast, for LC R108A and LC wt we found the chemical shifts for these residues to be close to the major and more stable C_L conformer.

Hence, a conformational shift of the C_L domain and an open orientation of V_L and C_L upon introducing E108 are likely. Our NMR and RDC data indicate that this is not the case for the full-length LC as there are no substantial chemical shift changes or loss of signals for the core residues of any LC variant studied. Surprisingly, the conformational stability of the LC is generally lower than the average of the stabilities of its individual domains suggesting a negative influence of the observed transient inter-domain contacts and the connection via the linker. Linker extensions by one or five non-polar residues (LC link+G, LC link+GAGAG) resulted in a further decrease in stability consistent with the idea that in these variants R108 is sterically separated from its potential interactions partners (Y140, D165, D170 and T172), similar to the missing R108 in the variant LC R108A.

Although some mutations destabilized and locally altered the conformation of the LC as well as the relative orientation of its domains, no LC variant tested formed fibrils under physiological conditions (PBS pH 7.4) even after 20 days of incubation. Only under destabilizing conditions, the LC variants, except LC wt, formed amyloid fibrils This supports the view that the conformational stability does not necessarily correlate with the fibril formation propensity of a given protein [22,23,28,29]. In contrast, all V_L variants studied readily fibrillated when incubated in PBS pH 7,4: V_L R108A, V_L R108E and V_L ΔR108 revealed comparable fibril formation kinetics, although only V_L ΔR108 exhibits a decreased conformational stability. This emphasizes the importance of R108 in particular for the structural integrity, not necessarily for the stability, of the V_L domain. In that context, Nokwe et al. found the deletion of R108 in the isolated V_L to result in a partially solvent-exposed hydrophobic core leading to a higher fibril formation propensity [21]. Its presence in V_L seems to inhibit structural transitions that lead to fibril formation. It could well be that R108 helps stabilizing a specific folding intermediate as V_L can pursue two distinct folding pathways, each comprising an intermediate, which folds to the native state [30]. It was suggested that specific folding intermediates shift towards amyloid formation [30–32]. Our finding of a C_L domain protecting V_L from fibrillation, in the context of a LCκ is consistent with results on other LC proteins [15,22].. Recent studies suggested that aspects other than the conformational stability of the LC have to be factored in, such as kinetically controlled fibril reactions or specific structural features [22,24,29,33] as well as proteolysis [19]. While LC wt did not form fibrils even when incubated under destabilizing conditions, all LC linker variants exhibited fibril propensity, albeit to different extents. By correlating fibril formation propensity with susceptibility to proteolysis with respect to the relative domain orientation, we were able to determine a direct correlation between this structural feature and amyloid formation, independent from the thermodynamic stability of the LC variants. Thus, for the antibody LC, the orientation of the two constituent domains plays a key role in amyloid fibril formation. A

compact and stable orientation assured by R108 is favored and maintains amyloid-resistance.

C_L did not form any fibrils, neither the wild type protein nor the comparably instable C_L ΔR108 variant, which is in agreement with the literature [26]. A major difference between the two domains is the presence of helical elements in C_L, which are missing in the V_L domain. These helices rapidly initialize folding of the C_L domain by orientating the β-sheets correctly [26,30,34]. It was suggested that specific residues and sequences within V_L modulate its intrinsically amyloidogenic [35], which might not be the case for C_L. These features seem to make C_L intrinsically more resistant against misfolding and amyloid formation than V_L. However, it was hypothesized that amyloid fibril formation is a generic, condition-dependent structural form of all proteins [36]. Supporting this, it had been shown that also C_Lκ is capable of forming amyloid fibrils *in vitro* under distinct experimental conditions [37].

As C_L protects V_L from fibril formation even in the presence of amyloidogenic linker mutations, it seems to modulate the folding pathway of V_L under mild conditions. This might be achieved by stabilizing the local fold of the C-terminal segment of V_L, which forms β-strands in amyloid fibrils derived from V_L domains as demonstrated by MAS solid-state NMR [38,39]. The fact that the full-length LC does not form fibrils under physiological conditions suggests that an unrestrained C-terminus is a key requirement for the transition from natively folded monomers to amyloid fibrils. Alternatively, as a looser domain orientation promotes fibril formation, a correctly oriented C_L domain might inhibit this pathway by blocking interaction sites in V_L required for the fibril emergence. Thus, the LC is optimized for preventing off-pathway reactions of the V_L domain while maintaining its conformational flexibility required for antigen binding and the interaction with the heavy chain.

Materials & Methods

Oligonucleotides were from MWG Biotech (Ebersberg, Germany). All chemicals were purchased from Merck (Darmstadt, Germany) or Sigma (St. Louis, USA). Unless otherwise stated all measurements were carried out in PBS buffer (10 mM Na_2HPO_4 x 2 H_2O; 1.8 mM KH_2PO_4; 2.7 mM KCl; 137 mM NaCl) pH 7.4 at 25 °C.

Cloning, mutagenesis, expression and purification

MAK33 V_L wt (D1 - R108) and ΔR108 variants, as wells as MAK33 C_L wt (R108 - E214), C_L ΔR108 and MAK33 LC wt (D1-E214) encoding plasmids were previously described [21,23,40–42]. MAK33 V_L variants (R108A, R108E), MAK33 C_L variants (R108A, R108E), as well the MAK33 LC variants (R108A, R108E, link+G, link+GAGAG), were generated by QuikChange mutagenesis PCR (Agilent, Santa Clara, CA, USA) according to the manufacturers protocol. MAK33 V_L wt, C_L wt and LC wt served as PCR templates for generation of the respective single amino acid substitution variants. All variants were expressed and purified as previously described [30,40]. For NMR experiments, expression of ^{15}N-labelled proteins was performed in M9 minimal medium containing $^{15}NH_4Cl$ (Cambridge Isotope Laboratories Inc., Andover, MA, USA) as the only nitrogen source. Briefly, the plasmids were transformed into *E. coli* BL21 (DE3)-star cells. Protein expression at 37 °C was induced with 1 mM IPTG at an OD_{600} of 0.6 – 0.8. Cells were harvested after overnight protein expression and inclusion bodies were prepared as previously described [43]. The pellet was solubilized and unfolded in 25 mM Tris-HCl (pH 8.0), 5 mM ethylenediaminetetraacetic acid (EDTA), 8 M urea and 2 mM β-mercaptoethanol at room temperature for at least 2 h. The solubilized protein was injected on a Q-Sepharose column equilibrated in 25 mM Tris-HCl (pH 8.0), 5 mM EDTA and 5 M urea. The proteins were all eluted in the flow-through fraction and subsequently refolded by dialysis against 250 mM Tris-HCl (pH 8.0), 100 mM L-Arg, 5 mM EDTA, 1 mM oxidized glutathione and 0.5 mM reduced glutathione at 4 °C overnight. To remove aggregates and remaining impurities, the proteins were purified using a Superdex 75 16/60 column (GE Healthcare, Uppsala, Sweden) equilibrated in PBS buffer. The recovery and purity of intact protein was verified by SDS-PAGE.

CD measurements

CD measurements were performed using a Jasco J-715 spectropolarimeter (Jasco, Grossumstadt, Germany) equipped with a Peltier element. Far-UV CD spectra were

measured using 10 µM protein in a 0.1-cm-pathlength cuvette between 260 and 200 nm wavelengths. Spectra were accumulated 16 times, buffer-corrected and normalized for mean residue molar ellipticity. Near-UV CD spectra were recorded using 50 µM protein in a 0.2-cm-pathlength quartz cuvette between 320 and 260 nm. Spectra were accumulated 16 times, buffer-corrected and normalized for mean residue molar ellipticity. Thermal transitions were recorded using 10 µM protein in a 0.1-cm-pathlength quartz cuvette at 205 nm with a heating rate of 30 °C/h.

Thioflavin T assay

Thioflavin T (ThT) assays were performed in triplicates in black 96 well microplates (#437112, Nunc, ThermoFisher Scientific, Roskilde, Denmark). Samples with and without SDS were measured with a Tecan Infinite 200 PRO M Nano or a Tecan Genios plate reader (Tecan Group Ltd., Männedorf, Switzerland) at 440 and 480 nm excitation and emission wavelengths, respectively [44]. To start the assay with defined samples and prevent seed contamination, monomer isolation was performed prior to the ThT assay by loading the proteins on a Superdex 75 10/300 GL (GE Healthcare, Uppsala, Sweden) and collecting the monomer peak fraction. Assays were conducted using 20 µM monomeric protein, 10 µM ThT in PBS buffer pH 7.4 with 0.05 % NaN_3 and 0.5 mM SDS for indicated experiments with a final volume of 200 µl per well. Microplates were covered with a Crystal Clear PP sealing foil (HJ-Bioanalytik GmbH, Erkelenz, Germany). Each well was measured every 30 min for 20 days. Between the measurements, the 96 well microplates were incubated at 37°C under continuous orbital shaking at 225 rpm within the plate reader. For data analysis, means of the replicates were calculated and individually normalized for better comparison.

Transmission Electron Microscopy (TEM)

For TEM micrographs, 10 µl samples were taken from the ThT assay wells and placed on a 200-mesh activated copper grid and incubated for 1 min. The samples were washed twice with 10 µl H_2O and negatively stained with 8 µl of a 1.5% (w/v) uranyl acetate solution for 1 min. Excess solutions on the grid were removed with a filter paper. Micrographs were recorded with a JEOL JEM-1400 Plus transmission electron microscope (JEOL Germany GmbH, Freising, Germany) at 120 kV.

SEC equipped with multi angle light scattering (SEC-MALS)

A Superdex 75 10/300 GL column (GE Healthcare, Uppsala, Sweden) connected to a Shimadzu HPLC system (Shimadzu, Munich, Germany) equipped with a Dawn Heleos II multi-angle light scattering detector (Wyatt Technology, Dernbach, Germany), a RID-20A detector and a SPD-20A detector (Shimdazu, Munich, Germany) were employed for determination of the absolute mass of all variants. The column was equilibrated for at least 24 h to obtain stable light scattering baseline signals before data collection. Determination of inter-detector delay volumes, band broadening correction and light-scattering detector normalization were performed using a standard 1 mg/ml BSA solution run, according to manufacturer's protocol. 50 µl of 50 µM protein samples were loaded on the column. All experiments were performed at room temperature at a flow rate of 0.45 ml/min in PBS buffer containing 0.05 % NaN_3. The molecular weight was determined using the ASTRA 5.3.4.20 software (Wyatt Technology, Dernbach, Germany) according to the manufacturer's protocol.

Limited Proteolysis

Limited proteolysis using proteinase K was performed in a protein:protease ratio of 10:1 (m/m). LC stocks were adjusted to a concentration of 43 µM and diluted with the assay buffer (50 mM Tris, 5 mM $CaCl_2$, pH 7.5) to a final concentration of 21 µM (0.5 mg/ml). Proteinase K stocks of 1 mg/ml were diluted with the same buffer to a final concentration of 0.05 mg/ml. Proteolysis was stopped at the indicated time points using a final concentration of 2 mM PMSF. To prevent residual protease activity, stopped samples were stored on ice, mixed with Laemmli buffer and analyzed by SDS-PAGE after the last sample (60 min) was taken. For quantification of the full-length LC spots, NIH ImageJ [45] was utilized.

NMR spectroscopy

All NMR spectra were conducted using uniformly ^{15}N-labeled protein. Experiments were recorded at a 600 MHz Bruker Avance III spectrometer (Bruker Biospin, Karlsruhe) equipped with a triple-resonance cryoprobe. $^{1}H,^{15}N$-HSQCs were acquired with a 50 µM protein solution in a buffer containing 20 mM sodium phosphate, 50 mM NaCl and 10 % (v/v) D_2O at pH 6.5 and a temperature of 298 K. Backbone resonance assignments were transferred from previous assignments [21].

For measurement of residual dipolar couplings (RDCs), we employed 320 µM protein solutions in a buffer containing 10 mM NaH_2PO_4, 50 mM NaCl, 0.5 mM EDTA and 10 % D_2O

at pH 6.5. A first IPAP-type $^{1}H,^{15}N$-HSQC experiment was conducted to determine isotropic J couplings, then Pf1 phage (ASLA Biotech, Riga) was added to a final concentration of ~ 8 mg/ml, producing a D_2O splitting of 6.0 Hz for LC wt and 11.9 Hz for LC R108A. Experiments were recorded and processed using TopSpin 3.2 (Bruker BioSpin) and analyzed with CcpNmr analysis 2.4 [46]. Chemical shift changes were calculated using the following formula:

$$\Delta\delta^{\mathrm{res}} = \sqrt{(\Delta\delta^{^{1}\mathrm{H}})^2 + \frac{1}{25}(\Delta\delta^{^{15}\mathrm{N}})^2}$$

PDB structures were analyzed using UCSF chimera 1.11.2 [47].

RDC-based ensembles

RDCs based ensembles were modelled for LC wt and LC R108A variants making use of Metadynamic Metainference (M&M) [48] and the theta-method [49]. M&M enables introducing experimental information into Molecular Dynamics (MD) simulations optimally balancing their information content by means of Bayesian Inference. This allows taking into account at the same time force field inaccuracies as well as the possible sources of error affecting experimental data like random noise, systematic errors etc., leading, at least in principle, to structural ensembles that do not depend on the initial choice of the MD force field and that are in agreement with some given experimental evidence [50].

In M&M multiple molecular dynamics (MD) simulations (replicas) are run in parallel coupled together by a potential that, by using Bayesian inference, optimally balance the information content of the experimental data with that of the MD force field [51] while at the same time the overall sampling of the conformational space is enhanced by Metadynamics [52].

MD simulations were performed for LC wt and LC R108A using GROMACS 2016 [53], PLUMED 2 [54] and the ISDB PLUMED module [55]. The two systems were prepared from the PDB 1FH5 (chain L) [56] applying the following mutations using SCWRL4 [57] (E17D and Y87F in both cases and R108A for the second system). The systems were solvated with ~ 25000 TIP4P 2005 [58] water molecules and 2 or 3 Na^+ ions in a dodecahedron box of ~ 765 nm^3. The force field employed was the Amber03W [59]. After an energy minimization followed by a NVT and NPT thermalization at 298 K and 1 atm the systems were simulated in the NPT ensemble (using velocity rescaling [60] and Parrinello-Rahman [61]) for 16 ns and 16 conformations were extracted as the starting point for the 16 replicas employed in the M&M simulations.

M&M simulations with RDCs were run as described previously [50]. A complete setup to reproduce the simulations, including starting conformations, topologies, GROMACS parameters file and PLUMED input files, can be downloaded from GitHub (https://github.com/carlocamilloni/papers-data). Shortly, a preliminary optimization of the relative orientation of the 16 replicas with respect to the Z-axis was obtained by maximizing the correlation between calculated and measured RDCs. The Metainference potential was then turned on making use of a Gaussian noise per data point [62] and a scaling factor sampled from a prior uniform distribution [50]. The sampling of the conformational space was boost by Metadynamics using Parallel Bias and sharing the bias among the replicas [48,63] using a BIASFACTOR of 20 and an initial energy deposition rate of 2.5 kJ/mol/ps. In particular because the aim is that of sampling the inter-domain dynamics, we selected the following collective variables (CVs) as defined in PLUMED: the 1) ALPHABETA and 2) DIHCOR defined over the psi angles of the linker residues; 3) the DISTANCE between the center of masses of the two domains defined using the Cα atoms; 4) the TORSION angle between the major axis of the two domains; and 5) the DHENERGY (Debye-Hückel energy) between the two domains calculated using the changed side-chains [64]. The width of the Metadynamics bias for the above CVs was set to 0.1, 0.1, 0.05 nm, 0.1, and 0.2 kJ/mol, respectively. For both systems each replica was evolved for 200 ns, for a total nominal time of 3.2 μs of simulation per system studied.

The collected sampling was eventually reweighted taking into account the final bias in the quasi-static regime [65] so to obtain a statistical ensemble of conformations with each conformer associated to a specific statistical weight.

Acknowledgements

This work was supported by a grant from the DFG to J.B. B.R. acknowledges support from the Helmholtz-Gemeinschaft, the Center for Integrated Protein Science Munich (CIPS-M) and the Deutsche Forschungsgemeinschaft (SFB-1035, project B07).

Figures

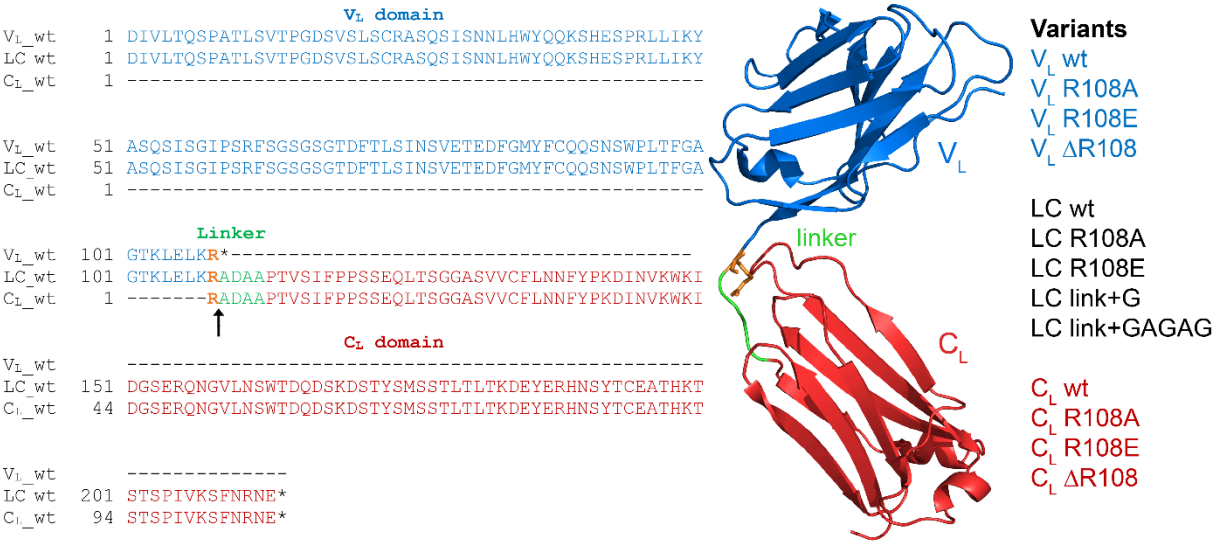

Figure 1. Sequence alignment and crystal structure. Alignment of the amino acid sequences of V_L wt, LC wt and C_L wt and the crystal structure of the MAK33 light chain (PDB: 1FH5) indicating the V_L domain (blue), the light chain linker (green), the C_L domain (red) and the mutated R108 (orange). In this study R108 was mutated to alanine, glutamate or deleted. Additional linker residues (LC link+G, LC link+GAGAG) were introduced between R108 and A109 (arrow) to assess the importance of the LC linker.

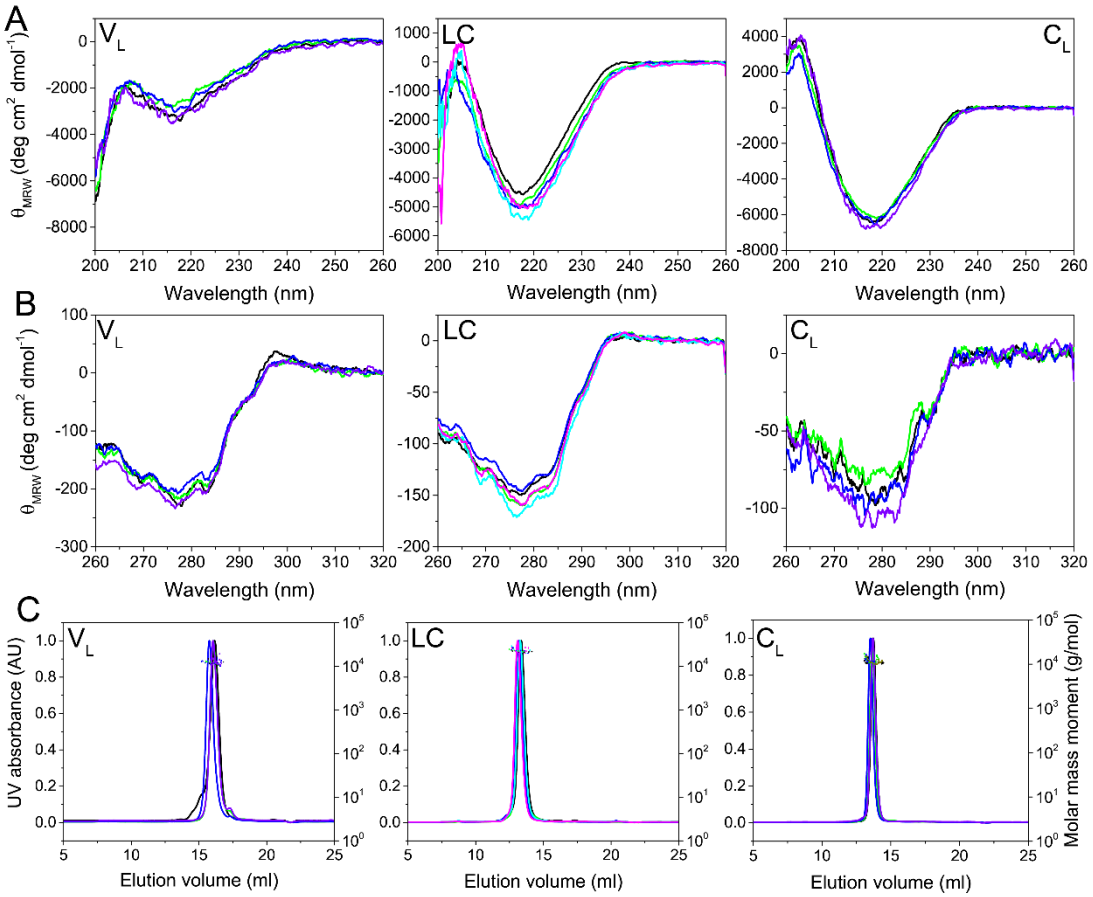

Figure 2. Structural analysis of V_L, LC and C_L variants. A) Far-UV CD spectra of V_L (left), LC (middle) and C_L (right) variants. **B)** Near-UV CD spectra of V_L (left), LC (middle) and C_L (right) variants. **C)** SEC-MALS chromatograms of V_L (left), LC (middle) and C_L (right) variants. Molecular mass moments are indicated by dots. All variants are predominantly monomeric. Colors indicate a mutation in the respective variants, wt (black), R108A (green), R108E (blue), ΔR108 (purple), link+G (turquoise), link+GAGAG (magenta).

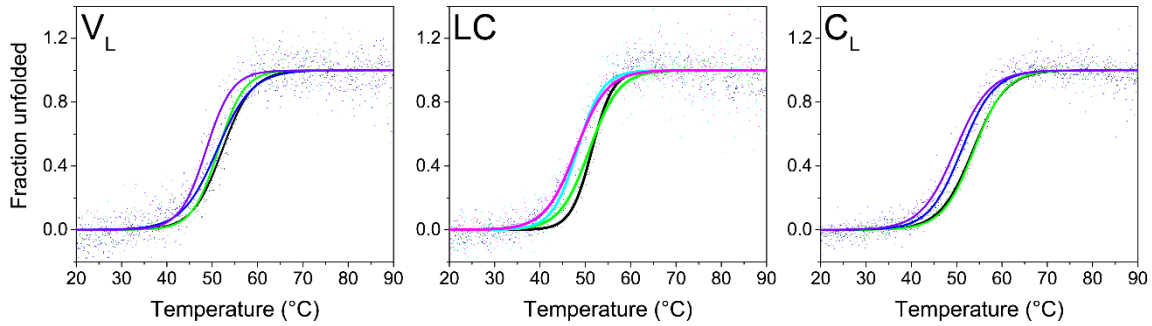

Figure 3. Thermal unfolding transitions monitored by CD spectroscopy. For V_L variants (left), only V_L ΔR108 reveals a decreased melting temperature (T_m). T_m values with errors are given in Table 1. Colors indicate a mutation in the respective variants, wt (black), R108A (green), R108E (blue), ΔR108 (purple), link+G (turquoise), link+GAGAG (magenta).

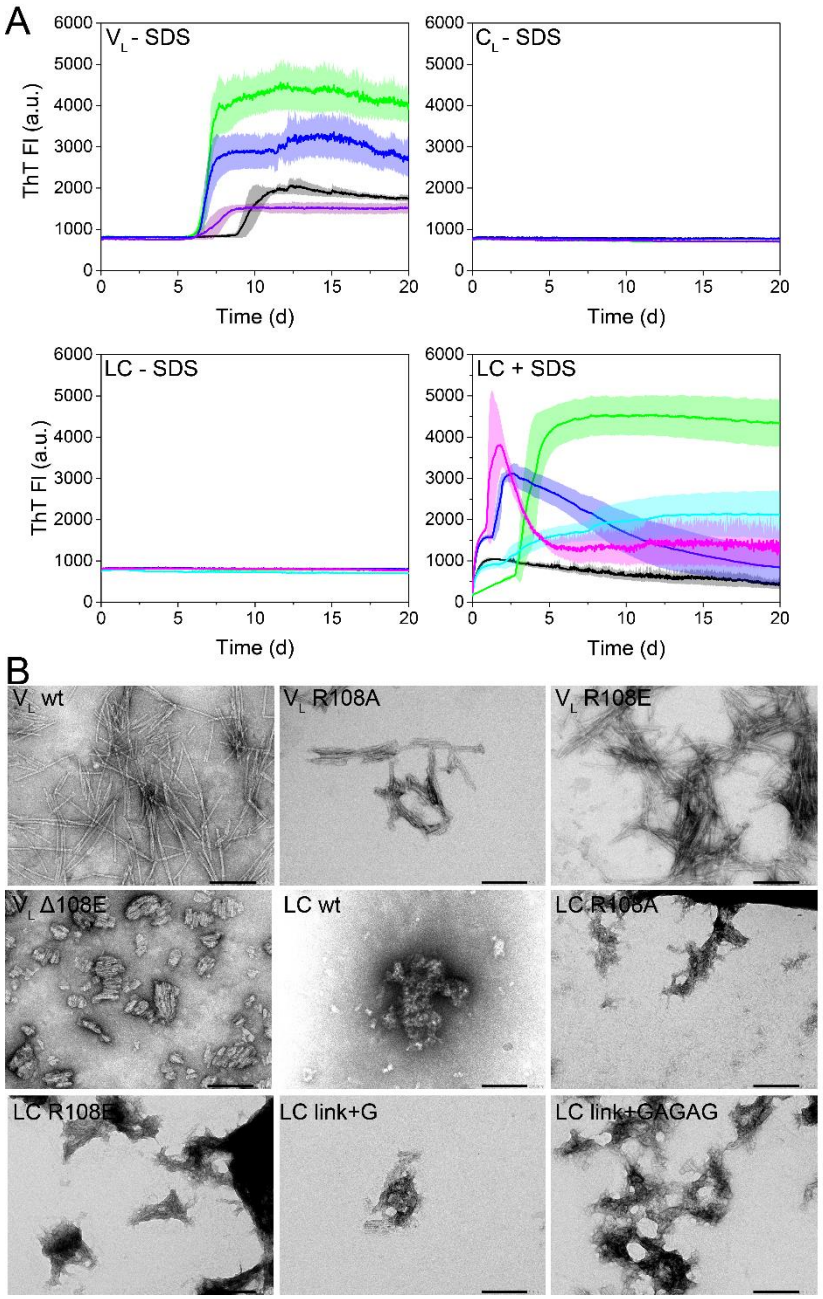

Figure 4. Fibril formation propensity. A) Fibril formation kinetics of V_L, LC and C_L variants monitored by ThT fluorescence. V_L, LC and C_L variants were incubated in PBS pH 7.4. Additionally, LC variants were incubated in the same buffer with 0.5 mM SDS added. For all measurements, 20 µM monomeric protein samples were incubated under shaking with 10 µM ThT and 0.05% NaN_3 for 20 days at 37°C. t_{50} values are given in Table 2. Errors represent the SD of technical duplicates. Colors indicate a mutation in the respective variants, wt (black), R108A (green), R108E (blue), ΔR108 (purple), link+G (turquoise), link+GAGAG (magenta). **B)** TEM images of fibril samples. LC wt revealed amorphous aggregates, but no fibrils. The scale bar represents 200 nm.

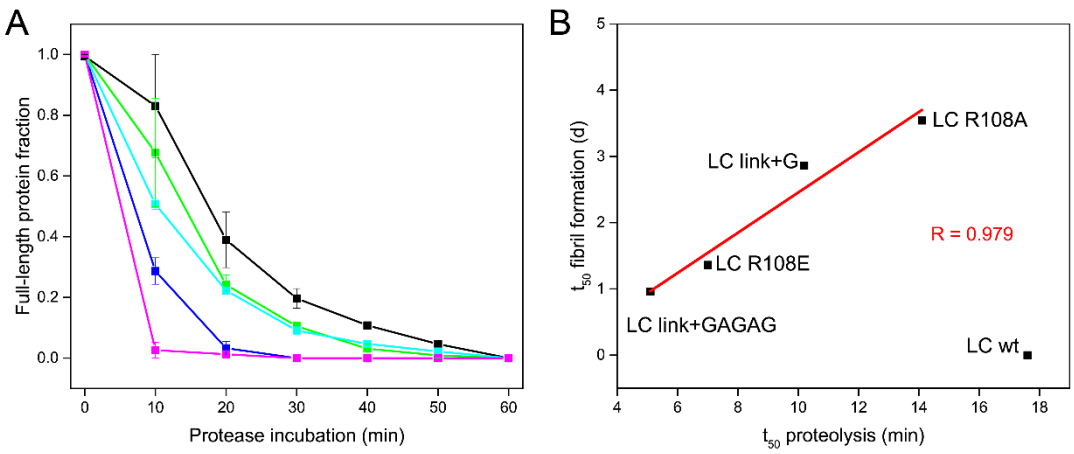

Figure 5. Limited proteolysis and correlation with fibril formation kinetics. A) Quantification of limited proteolysis using LC variants incubated with proteinase K (LC:protease ratio 10:1) for the indicated time points. Colors indicate a mutation in the respective variants, wt (black), R108A (green), R108E (blue), link+G (turquoise), link+GAGAG (magenta). Error represents the SD of technical duplicates. B) Correlation of t_{50} limited proteolysis with t_{50} fibril formation (+SDS). For all samples forming fibrils (Fig. 4), a linear fit with a Pearson R value of 0.979 is applicable. Of notice, LC wt did not form fibrils and is therefore not included in the fit.

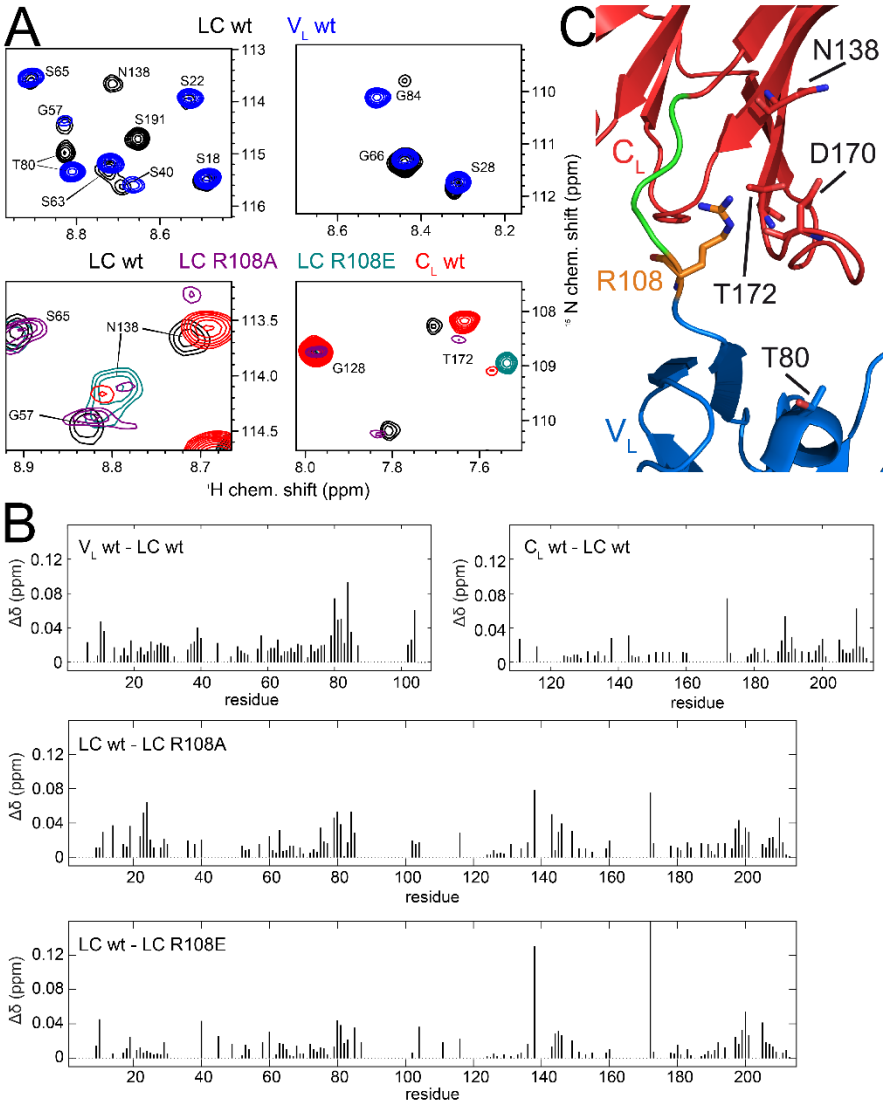

Figure 6. NMR spectra and chemical shifts. A) Top: Chemical shift differences between LC wt (black) and V_L wt (blue). Bottom: Chemical shifts differences between LC wt (black), LC R108A (purple), LC R108E (green) and C_L wt (red). We have observed previously that C_L wt undergoes slow chemical exchange between two conformations. In the spectra, this exchange is shown for residues N138 and T172, with a major and a minor population. LC wt N138 and T172 displayed similar chemical shifts as the major population of C_L wt. For N138 of LC R108A, two weak peaks were visible, however they were close to the noise. LC R108A T172 indicated this variant to adopt a conformation similar to the major population of C_L wt. In contrast, LC R108E produced chemical shifts close to those of the minor population of C_L wt. **B)** Chemical shift differences between V_L, C_L and LC constructs. For C_L wt, the major population for residues N138 and T172 were used to calculate chemical shift differences. **C)** Structural representation of the LC domain interface. In addition to the LC linker residue Arg108 (orange, other linker residues in green), important residues in V_L (T80) and C_L (N138, D170, T172) are highlighted.

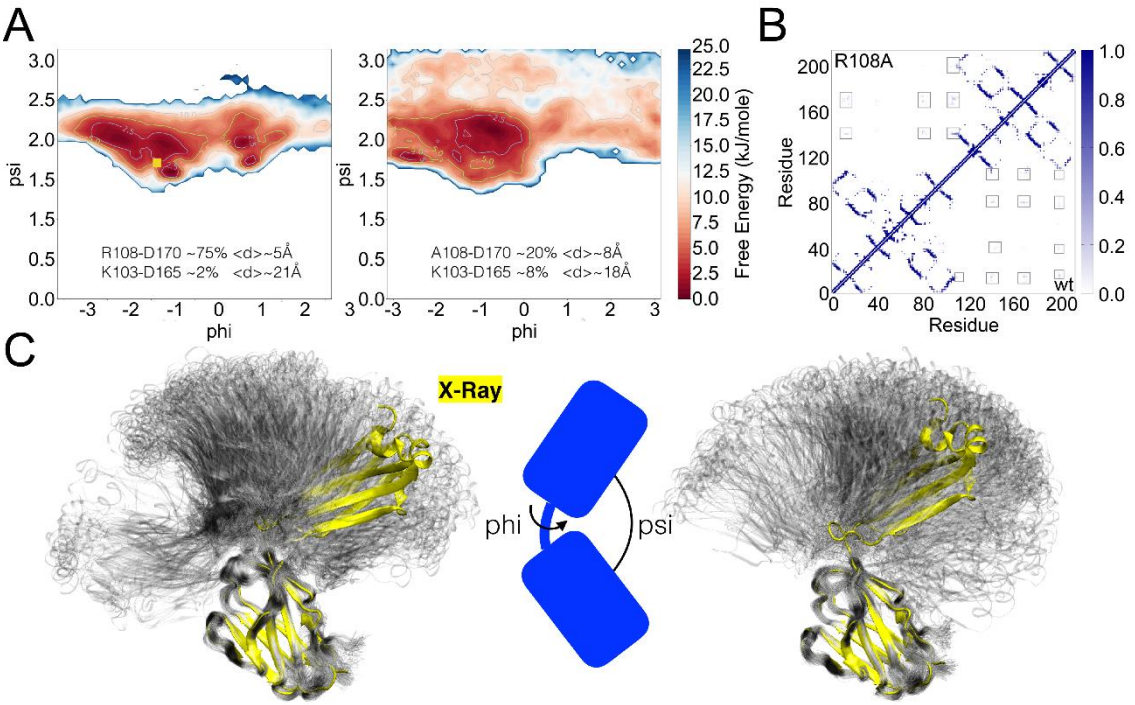

Figure 7. RDC-based ensembles. A) Free energy surfaces as a function of the domains relative orientation for LC wt (left) and LC R108A (right). The populations and average distance of two salt-bridges is also reported. In particular the salt-bridge R108-D170 (not formed in the crystal) is strongly populated in the wild type and abolished by the mutation. The K103-D165 salt-bridge that is formed in the crystal is instead only marginally populated in both ensembles. **B)** Probability contact maps for LC wt (bottom right) and LC R108A (top left). The overall effect of the R108A mutation is that of decreasing the populations of all contacts between the V_L and C_L domains. Highlighted are clusters of contacts formed between the two domains with a population > 5%. **C)** Comparison of the conformational ensembles for LC wt (left) and LC R108A (right) based on RDCs. The ensembles (grey) are represented below and compared with the crystal structure (PDB: 1FH5) in yellow. On average the effect of the R108A mutation is that of increasing the accessible conformational space towards more open conformations.

Tables

Table 1

Variant	Melting temperature (°C)		
	V_L	**LC**	**C_L**
wt	52	51	54
R108A	51	50	54
R108E	52	48	52
ΔR108	48	N/A	50
link+G	N/A	50	N/A
link+GAGAG	N/A	49	N/A

Table 1. Conformational stability. Melting temperatures (T_m) of all V_L, LC and C_L variants as determined by CD thermal transitions in °C (Fig. 3). Of notice, the SD of the T_ms, SE of the T_ms and SE of the Boltzmann fit for all variants are < 1 °C, < 0.3 °C and < 0.6 °C, respectively.

Table 2

Variant	t_{50} (days)		
	V_L	**LC**	**C_L**
wt	9.58 ± 0.47	no fibrils	no fibrils
R108A	6.99 ± 0.08	no fibrils	no fibrils
R108E	6.85 ± 0.20	no fibrils	no fibrils
ΔR108	7.52 ± 0.45	N/A	no fibrils
link+G	N/A	no fibrils	N/A
link+GAGAG	N/A	no fibrils	N/A

Table 2. Fibril formation kinetics. t_{50} values represent the time point when 50 % of the total fibril formation was reached. Data was taken from ThT fibril kinetics (Fig. 4A). Errors represent the SD of technical duplicates. N/A: construct does not exist.

Table 3

Variant	t_{50} (days)
	LC
wt	no fibrils
R108A	3.54 ± 0.48
R108E	1.36 ± 0.54
ΔR108	N/A
link+G	2.86 ± 0.01
link+GAGAG	0.96 ± 0.27

Table 3. Fibril formation kinetics for LCs with 0.5 mM SDS. t_{50} values represent the time point when 50 % of the total fibril formation was reached. Data was taken from ThT fibril kinetics (Fig. 4A). Errors represent the SD of technical duplicates. N/A: construct does not exist.

References

[1] M.J. Feige, L.M. Hendershot, J. Buchner, How antibodies fold., Trends Biochem. Sci. 35 (2010) 189–98. doi:10.1016/j.tibs.2009.11.005.

[2] D. French, R. Laskov, M. Scharff, The role of somatic hypermutation in the generation of antibody diversity, Science (80-.). 244 (1989) 1152–1157. doi:10.1126/science.2658060.

[3] H.W. Schroeder, L. Cavacini, Structure and function of immunoglobulins., J. Allergy Clin. Immunol. 125 (2010) S41-52. doi:10.1016/j.jaci.2009.09.046.

[4] L. Ellgaard, a. Helenius, Quality control in the endoplasmic reticulum, Nat. Rev. Mol. Cell Biol. 4 (2003) 181–191. doi:10.1038/nrm1052.

[5] E.M. Baden, L.A. Sikkink, M. Ramirez-Alvarado, Light chain amyloidosis - current findings and future prospects., Curr. Protein Pept. Sci. 10 (2009) 500–8. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3898330&tool=pmcentrez&rendertype=abstract (accessed May 20, 2015).

[6] M.B. Pepys, Amyloidosis, Annu. Rev. Med. 57 (2006) 223–241. doi:10.1146/annurev.med.57.121304.131243.

[7] H.P. McWilliams-Koeppen, J.S. Foster, N. Hackenbrack, M. Ramirez-Alvarado, D. Donohoe, A. Williams, S. Macy, C. Wooliver, D. Wortham, J. Morrell-Falvey, C.M. Foster, S.J. Kennel, J.S. Wall, Light chain amyloid fibrils cause metabolic dysfunction in human cardiomyocytes, PLoS One. 10 (2015). doi:10.1371/journal.pone.0137716.

[8] G.G. Glenner, J. Harbaugh, J.I. Ohms, M. Harada, P. Cuatrecasas, An amyloid protein: The amino-terminal variable fragment of an immunoglobulin light chain, Biochem. Biophys. Res. Commun. 41 (1970) 1287–1289. doi:10.1016/0006-291X(70)90227-5.

[9] G.G. Glenner, D. Ein, E.D. Eanes, H.A. Bladen, W. Terry, D.L. Page, Creation of “amyloid” fibrils from Bence Jones proteins in vitro., Science. 174 (1971) 712–4. doi:10.1126/science.174.4010.712.

[10] J.N. Buxbaum, J. V. Chuba, G.C. Hellman, A. Solomon, G.R. Gallo, Monoclonal immunoglobulin deposition disease: Light chain and light and heavy chain deposition diseases and their relation to light chain amyloidosis. Clinical features, immunopathology, and molecular analysis, Ann. Intern. Med. 112 (1990) 455–464. doi:10.7326/0003-4819-76-3-112-6-455.

[11] J. Buxbaum, Aberrant immunoglobulin synthesis in light chain amyloidosis. Free light chain and light chain fragment production by human bone marrow cells in short-term tissue culture, J. Clin. Invest. 78 (1986) 798–806. doi:10.1172/JCI112643.

[12] V. Bellotti, P. Mangione, G. Merlini, Review: Immunoglobulin light chain amyloidosis - The archetype of structural and pathogenic variability, J. Struct. Biol. 130 (2000) 280–289. doi:10.1006/jsbi.2000.4248.

[13] S. Dasari, J.D. Theis, J.A. Vrana, O.M. Meureta, P.S. Quint, P. Muppa, R.M. Zenka, R.C. Tschumper, D.F. Jelinek, J.I. Davila, V. Sarangi, P.J. Kurtin, A. Dogan, Proteomic detection of immunoglobulin light chain variable region peptides from amyloidosis

patient biopsies, J. Proteome Res. 14 (2015) 1957–1967. doi:10.1021/acs.jproteome.5b00015.

[14] F. Lavatelli, D.H. Perlman, B. Spencer, T. Prokaeva, M.E. McComb, R. Théberge, L.H. Connors, V. Bellotti, D.C. Seldin, G. Merlini, M. Skinner, C.E. Costello, Amyloidogenic and Associated Proteins in Systemic Amyloidosis Proteome of Adipose Tissue, Mol. Cell. Proteomics. 7 (2008) 1570–1583. doi:10.1074/mcp.M700545-MCP200.

[15] E.S. Klimtchuk, O. Gursky, R.S. Patel, K.L. Laporte, L.H. Connors, M. Skinner, D.C. Seldin, The Critical Role of the Constant Region in Thermal Stability and Aggregation of Amyloidogenic Immunoglobulin Light Chain, Biochemistry. 49 (2010) 9848–9857. doi:10.1021/bi101351c.

[16] L. Oberti, P. Rognoni, A. Barbiroli, F. Lavatelli, R. Russo, M. Maritan, G. Palladini, M. Bolognesi, G. Merlini, S. Ricagno, Concurrent structural and biophysical traits link with immunoglobulin light chains amyloid propensity, Sci. Rep. 7 (2017) 16809. doi:10.1038/s41598-017-16953-7.

[17] G.J. Morgan, G.A. Usher, J.W. Kelly, Incomplete Refolding of Antibody Light Chains to Non-Native, Protease-Sensitive Conformations Leads to Aggregation: A Mechanism of Amyloidogenesis in Patients?, Biochemistry. 56 (2017) 6597–6614. doi:10.1021/acs.biochem.7b00579.

[18] R.P. Linke, F.W. Tischendorf, D. Zucker-Franklin, E.C. Franklin, The Formation of Amyloid-Like Fibrils in Vitro from Bence Jones Proteins of the VλI Subclass, J Immunol. 111 (1973) 24–26. http://www.jimmunol.org/content/111/1/24.short.

[19] G.J. Morgan, J.W. Kelly, The Kinetic Stability of a Full-Length Antibody Light Chain Dimer Determines whether Endoproteolysis Can Release Amyloidogenic Variable Domains, J. Mol. Biol. 428 (2016) 4280–4297. doi:10.1016/j.jmb.2016.08.021.

[20] P.P. Mangione, R. Porcari, J.D. Gillmore, P. Pucci, M. Monti, M. Porcari, S. Giorgetti, L. Marchese, S. Raimondi, L.C. Serpell, W. Chen, A. Relini, J. Marcoux, I.R. Clatworthy, G.W. Taylor, G.A. Tennent, C. V. Robinson, P.N. Hawkins, M. Stoppini, S.P. Wood, M.B. Pepys, V. Bellotti, Proteolytic cleavage of Ser52Pro variant transthyretin triggers its amyloid fibrillogenesis, Proc. Natl. Acad. Sci. 111 (2014) 1539–1544. doi:10.1073/pnas.1317488111.

[21] C.N. Nokwe, M. Hora, M. Zacharias, H. Yagi, C. John, B. Reif, Y. Goto, J. Buchner, The Antibody Light-Chain Linker Is Important for Domain Stability and Amyloid Formation, J. Mol. Biol. 427 (2015) 3572–3586. doi:10.1016/j.jmb.2015.09.012.

[22] L.M. Blancas-Mejía, T.J. Horn, M. Marin-Argany, M. Auton, A. Tischer, M. Ramirez-Alvarado, Thermodynamic and fibril formation studies of full length immunoglobulin light chain AL-09 and its germline protein using scan rate dependent thermal unfolding, Biophys. Chem. 207 (2015) 13–20. doi:10.1016/j.bpc.2015.07.005.

[23] C.N. Nokwe, M. Hora, M. Zacharias, H. Yagi, J. Peschek, B. Reif, Y. Goto, J. Buchner, A Stable Mutant Predisposes Antibody Domains to Amyloid Formation through Specific Non-Native Interactions, J. Mol. Biol. 428 (2016) 1315–1332. doi:10.1016/j.jmb.2016.01.015.

[24] L.M. Blancas-Mejía, A. Tischer, J.R. Thompson, J. Tai, L. Wang, M. Auton, M. Ramirez-Alvarado, Kinetic control in protein folding for light chain amyloidosis and the differential effects of somatic mutations., J. Mol. Biol. 426 (2014) 347–61. doi:10.1016/j.jmb.2013.10.016.

[25] S. Yamamoto, K. Hasegawa, I. Yamaguchi, S. Tsutsumi, J. Kardos, Y. Goto, F. Gejyo,

H. Naiki, Low Concentrations of Sodium Dodecyl Sulfate Induce the Extension of β 2 - Microglobulin-Related Amyloid Fibrils at a Neutral pH †, Biochemistry. 43 (2004) 11075–11082. doi:10.1021/bi049262u.

[26] M.J. Feige, S. Groscurth, M. Marcinowski, Z.T. Yew, V. Truffault, E. Paci, H. Kessler, J. Buchner, The structure of a folding intermediate provides insight into differences in immunoglobulin amyloidogenicity, Proc. Natl. Acad. Sci. 105 (2008) 13373–13378. doi:10.1073/pnas.0802809105.

[27] K. Chen, N. Tjandra, The use of residual dipolar coupling in studying proteins by NMR, Top. Curr. Chem. 326 (2012) 47–67. doi:10.1007/128_2011_215.

[28] C. Camilloni, B.M. Sala, P. Sormanni, R. Porcari, A. Corazza, M. De Rosa, S. Zanini, A. Barbiroli, G. Esposito, M. Bolognesi, V. Bellotti, M. Vendruscolo, S. Ricagno, Rational design of mutations that change the aggregation rate of a protein while maintaining its native structure and stability, Sci. Rep. 6 (2016). doi:10.1038/srep25559.

[29] M. Marin-Argany, J. Güell-Bosch, L.M. Blancas-Mejía, S. Villegas, M. Ramirez-Alvarado, Mutations can cause light chains to be too stable or too unstable to form amyloid fibrils., Protein Sci. 24 (2015) 1829–40. doi:10.1002/pro.2790.

[30] E.R. Simpson, E.M. Herold, J. Buchner, The Folding Pathway of the Antibody VL Domain, J. Mol. Biol. 392 (2009) 1326–1338. doi:10.1016/j.jmb.2009.07.075.

[31] J.W. Kelly, The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways, Curr. Opin. Struct. Biol. 8 (1998) 101–106. doi:10.1016/S0959-440X(98)80016-X.

[32] V.N. Uversky, A.L. Fink, Conformational constraints for amyloid fibrillation: The importance of being unfolded, Biochim. Biophys. Acta - Proteins Proteomics. 1698 (2004) 131–153. doi:10.1016/j.bbapap.2003.12.008.

[33] L. Saelices, L.M. Johnson, W.Y. Liang, M.R. Sawaya, D. Cascio, P. Ruchala, J. Whitelegge, L. Jiang, R. Riek, D.S. Eisenberg, Uncovering the mechanism of aggregation of human transthyretin, J. Biol. Chem. 290 (2015) 28932–28943. doi:10.1074/jbc.M115.659912.

[34] Z. Qin, D. Hu, M. Zhu, A.L. Fink, Structural characterization of the partially folded intermediates of an immunoglobulin light chain leading to amyloid fibrillation and amorphous aggregation, Biochemistry. 46 (2007) 3521–3531. doi:10.1021/bi061716v.

[35] M.R. Hurle, L.R. Helms, L. Li, W. Chan, R. Wetzel, A role for destabilizing amino acid replacements in light-chain amyloidosis., Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 5446–5450. doi:10.1073/pnas.91.12.5446.

[36] C.M. Dobson, Protein misfolding, evolution and disease, Trends Biochem. Sci. 24 (1999) 329–332. doi:10.1016/S0968-0004(99)01445-0.

[37] K. Yamamoto, H. Yagi, Y.H. Lee, J. Kardos, Y. Hagihara, H. Naiki, Y. Goto, The amyloid fibrils of the constant domain of immunoglobulin light chain, FEBS Lett. 584 (2010) 3348–3353. doi:10.1016/j.febslet.2010.06.019.

[38] M. Hora, R. Sarkar, V. Morris, K. Xue, E. Prade, E. Harding, J. Buchner, B. Reif, MAK33 antibody light chain amyloid fibrils are similar to oligomeric precursors, PLoS One. 12 (2017). doi:10.1371/journal.pone.0181799.

[39] D.W. Piehl, L.M. Blancas-Mejía, J.S. Wall, S.J. Kennel, M. Ramirez-Alvarado, C.M.

Rienstra, Immunoglobulin Light Chains Form an Extensive and Highly Ordered Fibril Involving the N- and C-Termini, ACS Omega. 2 (2017) 712–720. doi:10.1021/acsomega.6b00494.

[40] C.N. Nokwe, M. Zacharias, H. Yagi, M. Hora, B. Reif, Y. Goto, J. Buchner, A Residue-specific Shift in Stability and Amyloidogenicity of Antibody Variable Domains, J. Biol. Chem. 289 (2014) 26829–26846. doi:10.1074/jbc.M114.582247.

[41] J. Buchner, R. Rudolph, H. Lilie, Intradomain disulfide bonds impede formation of the alternatively folded state of antibody chains, J. Mol. Biol. 318 (2002) 829–836. doi:10.1016/S0022-2836(02)00171-7.

[42] M.J. Feige, E.R. Simpson, E.M. Herold, A. Bepperling, K. Heger, J. Buchner, Dissecting the Alternatively Folded State of the Antibody Fab Fragment, J. Mol. Biol. 399 (2010) 719–730. doi:10.1016/j.jmb.2010.04.032.

[43] M.J.. Thies, F. Pirkl, Chromatographic purification of the CH2 domain of the monoclonal antibody MAK33, J. Chromatogr. B Biomed. Sci. Appl. 737 (2000) 63–69. doi:10.1016/S0378-4347(99)00377-1.

[44] K. Gade Malmos, L.M. Blancas-Mejia, B. Weber, J. Buchner, M. Ramirez-Alvarado, H. Naiki, D. Otzen, ThT 101: a primer on the use of thioflavin T to investigate amyloid formation, Amyloid. 24 (2017) 1–16. doi:10.1080/13506129.2017.1304905.

[45] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nat. Methods. (2012). doi:10.1038/nmeth.2089.

[46] W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, E.D. Laue, The CCPN data model for NMR spectroscopy: Development of a software pipeline, Proteins Struct. Funct. Genet. 59 (2005) 687–696. doi:10.1002/prot.20449.

[47] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera, J. Comput. Chem. 25 (2004) 1605–12. doi:10.1002/jcc.20084.

[48] M. Bonomi, C. Camilloni, M. Vendruscolo, Metadynamic metainference: Enhanced sampling of the metainference ensemble using metadynamics, Sci. Rep. 6 (2016). doi:10.1038/srep31232.

[49] C. Camilloni, M. Vendruscolo, A Tensor-Free Method for the Structural and Dynamical Refinement of Proteins using Residual Dipolar Couplings, J. Phys. Chem. B. 119 (2015) 653–661. doi:10.1021/jp5021824.

[50] T. Löhr, A. Jussupow, C. Camilloni, Metadynamic metainference: Convergence towards force field independent structural ensembles of a disordered peptide, J. Chem. Phys. 146 (2017). doi:10.1063/1.4981211.

[51] S. Tavaré, D.J. Balding, R.C. Griffiths, P. Donnelly, Inferring coalescence times from DNA sequence data, Genetics. 145 (1997) 505–518. doi:10.1126/sciadv.1501177.

[52] A. Laio, M. Parrinello, Escaping free-energy minima, Proc. Natl. Acad. Sci. 99 (2002) 12562–12566. doi:10.1073/pnas.202427399.

[53] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1–2 (2015) 19–25. doi:10.1016/j.softx.2015.06.001.

[54] G.A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, PLUMED 2: New feathers for an old bird, Comput. Phys. Commun. 185 (2014) 604–613. doi:10.1016/j.cpc.2013.09.018.

[55] M. Bonomi, C. Camilloni, Integrative structural and dynamical biology with PLUMED-ISDB, Bioinformatics. 33 (2017) 3999–4000. doi:10.1093/bioinformatics/btx529.

[56] J.G. Augustine, A. de la Calle, G. Knarr, J. Buchner, C.A. Frederick, The Crystal Structure of the Fab Fragment of the Monoclonal Antibody MAK33, J. Biol. Chem. 276 (2001) 3287–3294. doi:10.1074/jbc.M005221200.

[57] G.G. Krivov, M. V. Shapovalov, R.L. Dunbrack, Improved prediction of protein side-chain conformations with SCWRL4, Proteins Struct. Funct. Bioinforma. 77 (2009) 778–795. doi:10.1002/prot.22488.

[58] J.L.F. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, J. Chem. Phys. 123 (2005). doi:10.1063/1.2121687.

[59] R.B. Best, J. Mittal, Protein simulations with an optimized water model: Cooperative helix formation and temperature-induced unfolded state collapse, J. Phys. Chem. B. 114 (2010) 14916–14923. doi:10.1021/jp108618d.

[60] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007). doi:10.1063/1.2408420.

[61] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190. doi:10.1063/1.328693.

[62] M. Bonomi, C. Camilloni, A. Cavalli, M. Vendruscolo, Metainference: A Bayesian inference method for heterogeneous systems, Sci. Adv. 2 (2016). doi:10.1126/sciadv.1501177.

[63] J. Pfaendtner, M. Bonomi, Efficient Sampling of High-Dimensional Free-Energy Landscapes with Parallel Bias Metadynamics, J. Chem. Theory Comput. 11 (2015) 5062–5067. doi:10.1021/acs.jctc.5b00846.

[64] T.N. Do, P. Carloni, G. Varani, G. Bussi, RNA/peptide binding driven by electrostatics - Insight from bidirectional pulling simulations, J. Chem. Theory Comput. 9 (2013) 1720–1730. doi:10.1021/ct3009914.

[65] D. Branduardi, G. Bussi, M. Parrinello, Metadynamics with adaptive gaussians, J. Chem. Theory Comput. 8 (2012) 2247–2254. doi:10.1021/ct3002464.

Supplemental Material

Figure S1

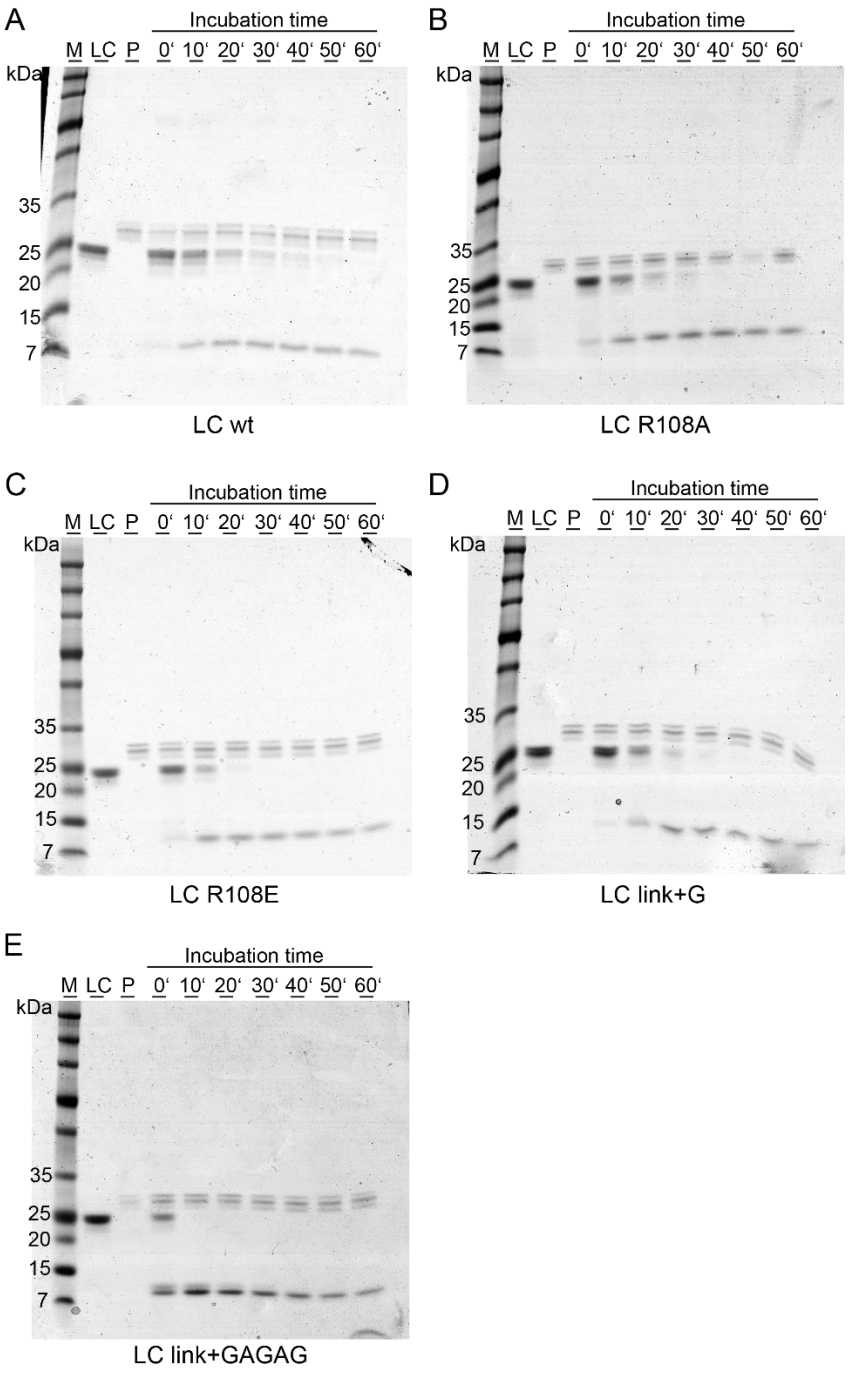

Figure S1. SDS-PAGE of limited proteolysis samples. LC variants were incubated with proteinase K (protein:proteinase K ratio 1000:1) for indicated time points and subsequently analyzed by SDS-PAGE. LC wt (A), LC R108A (B), LC R108E (C), LC link+G (D) and LC link+GAGAG (E). Abbreviations: Marker (M), protein only (LC), proteinase K only (P), incubation time in minutes (0’ to 60’).

Figure S2

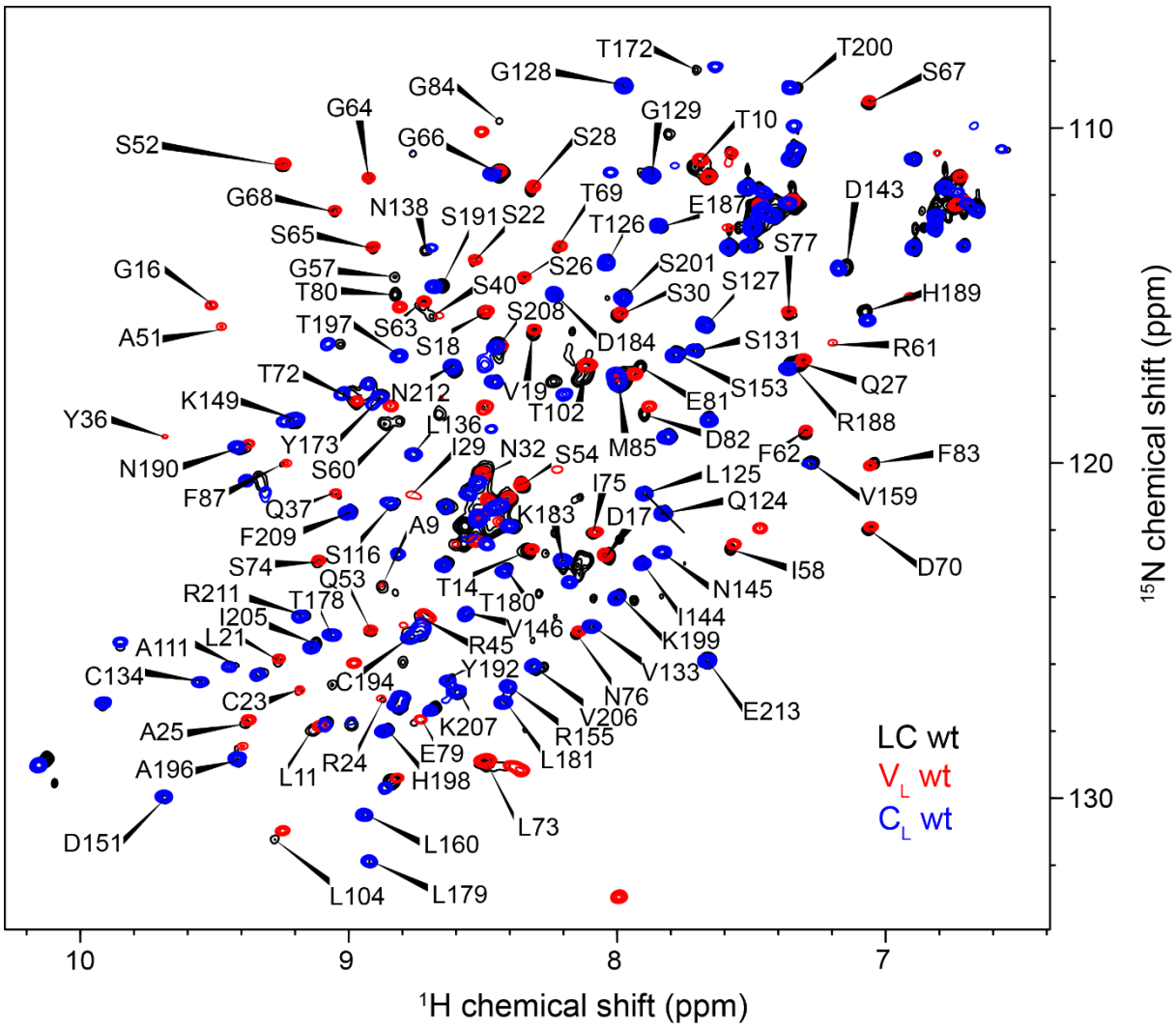

Figure S2. $^{1}H,^{15}N$-HSQC of LC wt (black), V_L wt (red) and C_L wt (blue).

Figure S3

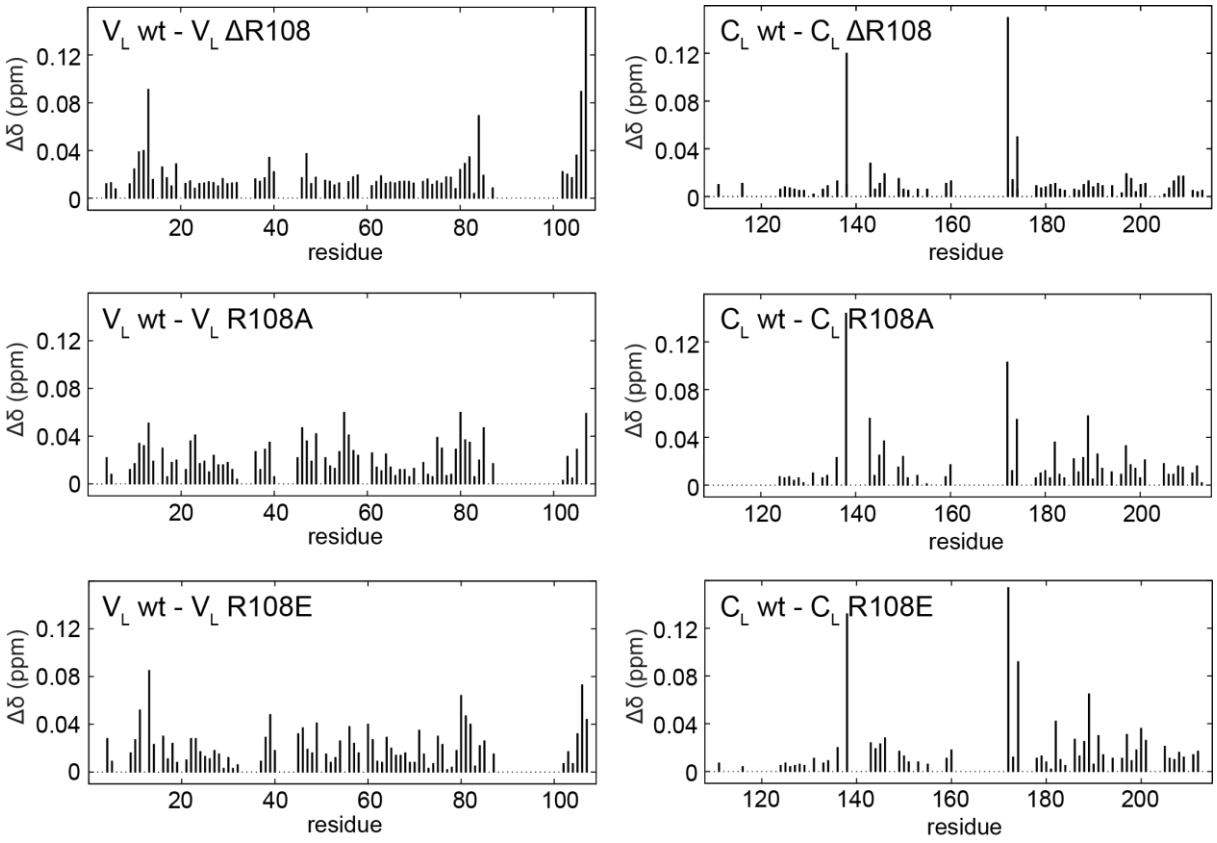

Figure S3. Chemical shift differences between various V_L (left) and C_L (right) variants.

Figure S4

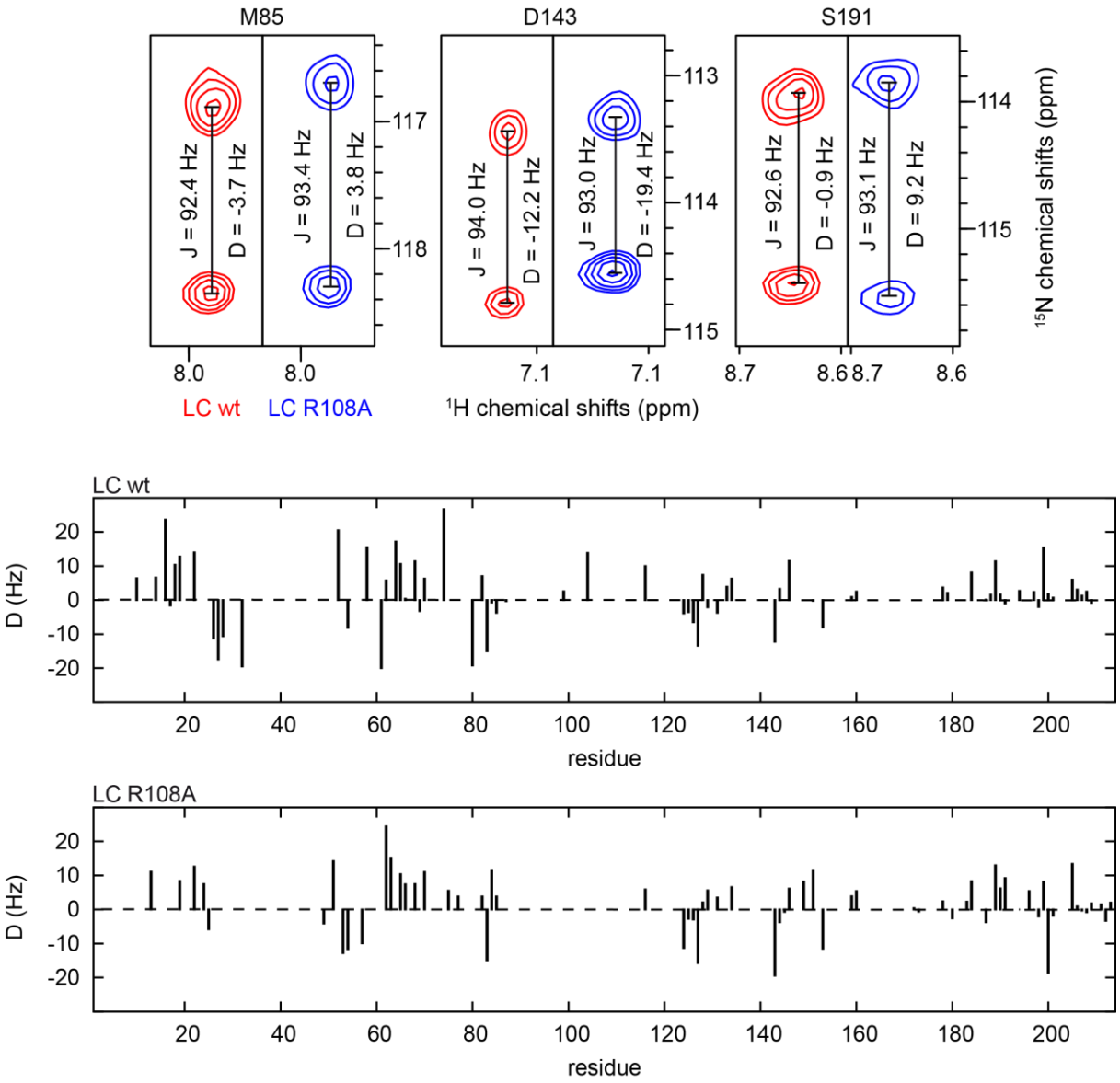

Figure S4. RDC spectra for LC wt and LC R108A. The size of the doublet splitting of the cross peak is dependent on the sum of J (scalar coupling, ~ 92 Hz for $^{1}J_{HN}$) and D (dipolar coupling, dependent on orientation). Subtraction of ^{15}N chemical shift differences measured with and without alignment medium yields the RDC values. RDCs were measured for residues along the whole sequence, yielding information about the relative orientation of the V_L and C_L domain.

Table S1

Variant	Molecular mass (kDa)		
	V_L	LC	C_L
wt	11.7	20.9	11.2
R108A	11.9	22.9	11.5
R108E	12.6	23.4	12.6
ΔR108	11.5	N/A	10.6
link+G	N/A	21.2	N/A
link+GAGAG	N/A	21.3	N/A

Table S1. Molecular masses. Molecular masses of the main species of the size-exclusion chromatograms (Fig. 2C) as determined by SEC-MALS. The calculated molecular masses are 11.8 kDa for V_L wt, 23.4 kDa for LC wt and 11.8 kDa for C_L wt. All variants are monomeric.

Table S2

	LC wt	LC R108A		LC wt	LC R108A		LC wt	LC R108A
residue	**D (Hz)**		**residue**	**D (Hz)**		**residue**	**D (Hz)**	
10	6.4		**70**	6.3	11.0	**159**	0.9	3.9
13		11.1	**74**	26.7		**160**	2.5	5.4
14	6.6		**75**		5.5	**172**		0.4
16	23.6		**77**		3.8	**173**		-0.6
17	-1.6		**80**	-19.2		**178**	3.7	2.4
18	10.4		**82**	7.0	3.8	**179**	2.1	
19	12.8	8.4	**83**	-15.0	-14.9	**180**		-2.6
22	14.0	12.6	**84**	-0.7	11.6	**183**		2.3
24		7.5	**85**	-3.7	3.8	**184**	8.1	8.3
25		-5.8	**87**	-0.3		**187**	0.1	-3.7
26	-11.2		**99**	2.6		**188**	1.6	
27	-17.4		**104**	13.9		**189**	11.4	13.0
28	-10.6		**116**	10.0	5.9	**190**	1.7	6.2
32	-19.5		**124**	-3.9	-11.3	**191**	-0.9	9.2
49		-4.1	**125**	-3.5	-2.7	**194**	2.7	
51		14.2	**126**	-6.5	-2.9	**196**		5.4
52	20.5		**127**	-13.4	-15.7	**197**	2.4	
53		-12.8	**128**	7.4	2.1	**198**	-2.0	-2.0
54	-8.1	-11.6	**129**	-2.1	5.6	**199**	15.4	8.1
57		-9.9	**131**	-3.7	3.5	**200**	1.8	-18.6
58	15.5		**133**	3.9		**201**	0.7	-1.8
61	-20.0		**134**	6.3	6.6	**205**	6.0	13.4
62	5.8	24.4	**143**	-12.2	-19.4	**206**	3.1	0.9
63		15.2	**144**	3.3	-3.7	**207**	1.3	-0.3
64	17.2		**145**		-0.7	**208**	2.5	-0.8
65	10.6	10.4	**146**	11.5	6.1	**209**	-0.8	1.8
66	0.4	7.4	**149**		8.2	**211**		1.5
68	11.4	7.5	**151**	-0.2	11.6	**212**		-3.3
69	-3.2		**153**	-8.0	-11.5	**213**		2.0

Table S2. RDC values for LC wt and LC R108A for the respective residue.