# Role of domain interactions in the aggregation of full-length immunoglobulin light chains

Enrico Rennella[a,b,c,1], Gareth J. Morgan[d,e,f,1,2], Jeffery W. Kelly[d,e,f,1], and Lewis E. Kay[a,b,c,g,1]

[a]Department of Molecular Genetics, The University of Toronto, Toronto, ON, Canada M5S 1A8; [b]Department of Biochemistry, The University of Toronto, Toronto, ON, Canada M5S 1A8; [c]Department of Chemistry, The University of Toronto, Toronto, ON, Canada M5S 1A8; [d]Department of Molecular Medicine, The Scripps Research Institute, La Jolla, CA 92037; [e]Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037; [f]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037; and [g]Program in Molecular Medicine, The Hospital for Sick Children, Toronto, ON, Canada M5G 1X8

Edited by David Baker, University of Washington, Seattle, WA, and approved November 30, 2018 (received for review October 11, 2018)

**Amyloid light-chain (LC) amyloidosis is a protein misfolding disease in which the aggregation of an overexpressed antibody LC from a clonal plasma cell leads to organ toxicity and patient death if left untreated. While the overall dimeric architecture of LC molecules is established, with each LC composed of variable ($V_L$) and constant ($C_L$) domains, the relative contributions of LC domain–domain interfaces and intrinsic domain stabilities to protection against LC aggregation are not well understood. To address these topics we have engineered a number of domain-destabilized LC mutants and used solution NMR spectroscopy to characterize their structural properties and intrinsic stabilities. Moreover, we used fluorescence spectroscopy to assay their aggregation propensities. Our results point to the importance of both dimerization strength and intrinsic monomer stability in stabilizing $V_L$ domains against aggregation. Notably, in all cases considered $V_L$ domains aggregate at least 10-fold faster than full-length LCs, establishing the important protective role of $C_L$ domains. A strong protective coupling is found between $V_L$–$V_L$ and $C_L$–$C_L$ dimer interfaces, with destabilization of one interface adversely affecting the stability of the other. Fibril formation is observed when either the $V_L$ or $C_L$ domain in the full-length protein is severely destabilized (i.e., where domain unfolding free energies are less than 2 kcal/mol). The important role of $C_L$ domains in preventing aggregation highlights the potential of the $C_L$–$C_L$ interface as a target for the development of drugs to stabilize the dimeric LC structure and hence prevent LC amyloidosis.**

light-chain amyloidosis | antibody light-chain domains | protein aggregation | solution NMR spectroscopy | proteinopathy

The misfolding and/or aggregation of a particular protein plays a causative role in ∼40 distinct human degenerative diseases, referred to as amyloid diseases (1–3). Some of these, including prion, Alzheimer's, Parkinson's, and Huntington's diseases, give rise to a central nervous system degeneration phenotype (4). Other diseases, including the peripheral amyloidoses, result from aggregation of proteins such as transthyretin (5, 6), antibody light chains (LCs) (7, 8), or islet amyloid polypeptides (9) that can be deposited as amyloid fibrils and other structures in the peripheral and autonomic nervous systems and/or in organs throughout the body such as the heart, kidneys, and digestive tract (1, 2). Whether accumulated amyloid affects tissue function, leading to its ultimate death, and/or whether smaller nonnative conformers cause proteotoxicity remains a debated topic. The most diagnosed systemic amyloid disease results from misfolding and/or aggregation of a different Ig LC in each Ig LC amyloidosis (AL) patient (7, 8). The Ig LC overexpressed and secreted by abnormally proliferating monoclonal plasma cells in AL patients typically misfolds and aggregates to form fibers in the heart, liver, kidneys, soft tissues, and nervous system (7). The mechanisms by which amyloidogenesis occurs and the processes leading to amyloid accumulation in organs are not well understood. X-ray structures of full-length (FL) Ig LCs show them to be dimeric, with each LC composed of N- and C-terminal variable ($V_L$) and constant ($C_L$) domains, respectively (10). Each LC monomer in the dimeric structure is arranged to form substantial $V_L$–$V_L$ and $C_L$–$C_L$ domain–domain interfaces, as shown schematically in Fig. 1, that bury 1,600 Å$^2$ and 2,400 Å$^2$ of protein surface area, respectively. Further stabilization of the dimer is provided by an inter-$C_L$ domain–$C_L$ domain disulfide bond and by a pair of intra-LC domain disulfides, one each in the $V_L$ and $C_L$ domains.

Much of the focus on understanding the mechanism of LC aggregation or amyloidogenesis has revolved around studies of $V_L$ domains, which were first identified, along with other LC fragments, as comprising amyloid fibrils (11–13). Isolated $V_L$ domains may arise from aberrant endoproteolysis of FL LCs, likely reflecting some degree of conformational dynamics within the FL molecule (14–16). Dimerization of these isolated domains is protective against aggregation (17, 18) and small molecules that stabilize the $V_L$–$V_L$ interface can reduce amyloid formation (19). However, in addition to fragments of $V_L$ and $C_L$ domains the FL LC protein is also found in amyloid deposits (20–23). There is not a consensus regarding whether FL LCs are able to independently form amyloid fibrils or, alternatively, whether an aberrant proteolysis event is required to release a fibril-competent fragment that is subsequently able to prime the aggregation of the FL species.

Notably, FL LCs appear to be much less aggregation-prone relative to the corresponding isolated $V_L$ domains in vitro (14, 21), suggesting that the FL dimeric LC structure is protective. The

## Significance

**The misfolding and/or aggregation of proteins can give rise to amyloid diseases, including immunoglobulin light-chain (LC) amyloidosis, leading to large proteinaceous deposits in organs throughout the body. A third of patients die within a year of diagnosis. The LC structure–proteotoxicity relationships and how protein aggregation causes degeneration are not understood. Using solution NMR we characterized the strengths of interactions between different domains in the dimeric 2-domain immunoglobulin LC structure to dissect their relative importance in protecting against aggregation. The constant domain ($C_L$) is particularly important in conferring protection, as are the stabilities of the individual domains. Our studies establish a threshold of stability, below which aggregation occurs rapidly, suggesting that stabilizing the $C_L$–$C_L$ domain interface could prevent aggregation.**

Author contributions: E.R., G.J.M., J.W.K., and L.E.K. designed research; E.R. and G.J.M. performed research; E.R., G.J.M., J.W.K., and L.E.K. analyzed data; and E.R., G.J.M., J.W.K., and L.E.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.



[1]To whom correspondence may be addressed. Email: rennella@pound.med.utoronto.ca, gjmorgan@bu.edu, jkelly@scripps.edu, or kay@pound.med.utoronto.ca.

[2]Present address: Department of Medicine, Boston University School of Medicine, Boston, MA 02118.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1817538116/-/DCSupplemental.

Published online December 31, 2018.

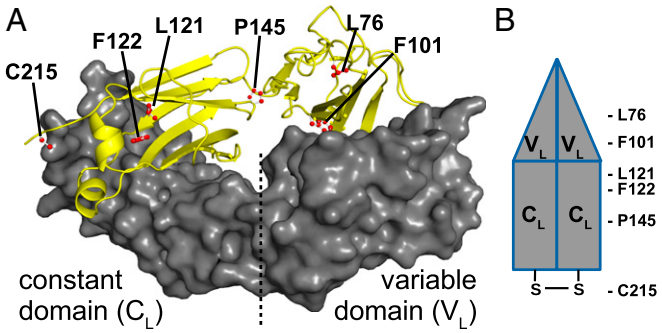


**Fig. 1.** Structure of the dimeric λ6 Ig LC. (*A*) One protomer comprising $V_L$ and $C_L$ domains is rendered in ribbon (yellow) and the other protomer is depicted as a surface representation (gray), highlighting the extensive $V_L$–$V_L$ and $C_L$–$C_L$ interfaces. The structure shown is a homology model built with the SWISS-MODEL server (https://swissmodel.expasy.org) using Protein Data Bank (PDB) ID code 1LIL (10) as a template. Sites (all side-chain heavy atoms) where mutations are introduced to evaluate the importance of domain interactions are indicated in red. (*B*) Schematic of the WT-dimeric LC structure used throughout the paper.

importance of folded multimeric structures in providing protection against aggregation-associated pathology is underscored by the recent development of a small molecule that binds to normally folded tetrameric transthyretin and slows transthyretin amyloid disease progression (24). Despite the clear link between stabilizing quaternary contacts and protection against aggregation, the structural elements that confer protection in most amyloidoses are not well understood. The clinical success realized by pharmacologic strategies that stabilize the native quaternary structure of an amyloidogenic protein suggests that a key target for AL will be stabilizing the FL LC molecule. As such, it is important to understand how each quaternary structural interface contributes to FL LC dimer stability and how mutations in key regions of the LC molecule might decrease quaternary structural stability, leading to conformations that are more aggregation-prone or conformations that enable endoproteolysis, facilitating LC fragment amyloidogenesis.

Solution NMR spectroscopy is a powerful tool for characterizing molecular interactions, including those between domains within the same molecule, in particular when the contacts are weak and thus not amenable to detailed, atomic-resolution analyses using other biophysical approaches (25, 26). Here we combine LC concentration-dependent NMR spin relaxation and chemical shift perturbation experiments, along with NMR-based hydrogen/deuterium exchange measurements, and biophysical studies of LC aggregation rates to quantify the relative contributions of different LC domain-domain interfaces to LC stability. We also employ this methodology to assess the role of primary amino acid sequence and intrinsic domain stabilities in providing protection against LC aggregation. We focus on two λ6 LC proteins (27, 28). The first is called JTO, originating from a patient with multiple myeloma (plasma cell proliferation with LC secretion but without LC amyloidogenesis). The second λ6 LC protein is WIL, derived from an AL patient. We employ the aforementioned methodology to study the individual $V_L$ and $C_L$ domains, subsequently extending the analyses to FL LCs. We show that the $V_L$ domain from JTO, JTO-$V_L$, has a dimerization dissociation constant that is three orders of magnitude smaller than that for WIL-$V_L$, while the intrinsic stability of the $V_L$ domain from JTO is ∼1.5 kcal/mol higher than the corresponding sequence from WIL, in agreement with previous findings (28). These results, together, provide a strong basis for understanding the decreased aggregation propensity of JTO-$V_L$ relative to WIL-$V_L$.

In studies on the FL LC we have used a mutagenesis approach to disrupt specific domain interfaces and, thus, assess the relative importance of specific domain–domain interactions both in preserving the dimeric nature of the LC molecule and, relatedly, in preventing LC aggregation. NMR studies show that the introduced mutations, in most cases, preserve secondary and tertiary structures of single domains but can destabilize interdomain contacts. Notably, mutations in the $C_L$ domain affect the stability of $V_L$–$V_L$ interactions, and vice versa, with certain mutant LCs retaining an overall dimeric structure yet lacking one of the $V_L$–$V_L$ or $C_L$–$C_L$ interfaces, while in other cases all interfaces are destabilized. The importance of both the intermolecular disulfide and the *cis* conformation of the peptidyl–prolyl amide bond at position 145 to the stability of the $C_L$ domain is clearly established by our studies. Characterization of the mutants in structural terms by NMR formed the basis for interpreting aggregation assays that allowed the assessment of intra- and interdomain contacts in protecting against amyloidogenesis. Remarkably, a monomeric LC (i.e., no $V_L$–$V_L$ or $C_L$–$C_L$ interactions) where the $V_L$–$C_L$ interface is significantly destabilized was protected against aggregation, while other mutants where the $V_L$ or the $C_L$ domain is destabilized are able to form fibrils, suggesting that the thermodynamic stability of the domains is a critical factor in preventing aggregation. Overall, our results highlight contributions to the stability of FL LC molecules from different intra- and interdomain contacts and, importantly, establish the significance of the $C_L$ domain in preventing aggregation, a first step in the targeted design of compounds that stabilize LC molecules.

## Results

Both WIL and JTO were originally described as V domains (28), with the pair of $V_L$ sequences 84% identical and 94% similar (*SI Appendix*, Fig. S1). To study the FL LCs we made constructs in which each V domain is fused to an identical λC3 constant domain (14). Fig. 1*A* highlights the dimeric λ6 LC structure that is essentially identical for WT JTO-FL and WIL-FL LCs examined in this study. Extensive interfaces between pairs of $V_L$ and $C_L$ domains are found; one of the two identical protomers is depicted in Fig. 1 as a yellow ribbon, while the other protomer is rendered as a gray CPK surface representation. The positions of a number of different mutations to be discussed below are indicated on the ribbon diagram depiction. For ease of illustration, the schematic shown in Fig. 1*B* will be used subsequently to depict the dimeric FL LC with all interfaces intact. As described in the Introduction, an important goal for understanding how the domain–domain interfaces contribute to FL LC stability and protection against aggregation is to evaluate the relative importance of each of the different domain–domain interactions ($V_L$–$V_L$, $C_L$–$C_L$, and intraprotomer $V_L$–$C_L$) to the integrity of the FL LC. Initially, we focus on individual $V_L$ or $C_L$ domains so as to provide a framework upon which to base our subsequent studies of the FL LC structure.

**JTO-$V_L$ Domain, but Not the WIL-$V_L$ Domain, Forms a Dimer.** Previous studies have established that monomeric and homodimeric $V_L$ domains exist in equilibrium (17, 29). Conditions resulting in a shift in the equilibrium from the homodimeric LC to the monomer and/or a decrease in the stability of the monomeric tertiary structure lead to an increasing propensity to form amyloid fibers (17, 30). Here we use a number of different NMR approaches to quantify the monomer–dimer equilibrium and the stability of each of the JTO-$V_L$ and WIL-$V_L$ domains to unfolding.

Initially a series of $^{1}H$-$^{15}N$ heteronuclear single quantum coherence (HSQC) spectra of JTO-$V_L$ (residues 1–111 of the FL protein) was recorded as a function of increasing protein concentration (Fig. 2*A*) showing a migration of peak positions toward those from dimeric JTO-FL (Fig. 2*A*, black), as illustrated for Q38 and Y50. In addition to shifts in peak positions, increased peak linewidths were also noted at the more dilute concentrations, despite the higher fraction of monomer, hence the expected decrease in peak linewidths in the absence of chemical exchange. Together this suggests that the JTO-$V_L$ domain exchanges between

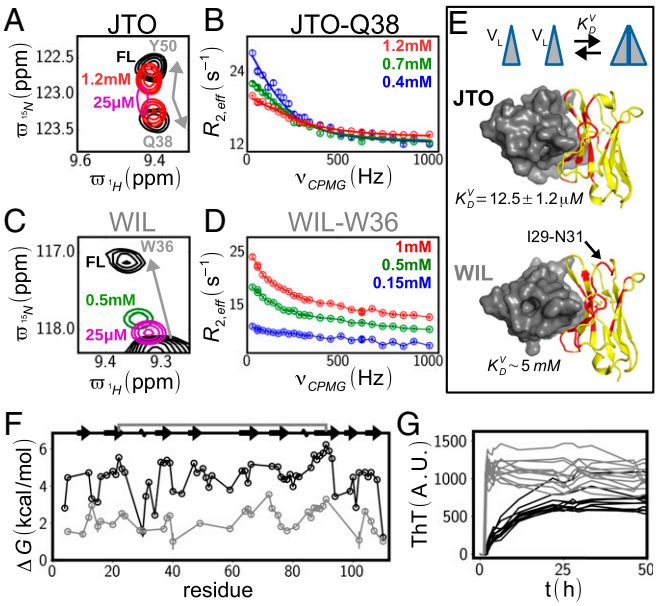


**Fig. 2.** Different stabilities for JTO-$V_L$ and WIL-$V_L$ dimers. (*A*) Superposition of selected regions of $^1H$,$^{15}N$ HSQC spectra of WT JTO-$V_L$ and WT JTO-FL LC (black, FL) recorded at 37 °C and 14.1 T highlighting Y50 and Q38. Data at two $V_L$ concentrations are shown, 0.025 mM (magenta) and 1.2 mM (red). Gray arrows indicate the trajectories of peak shifts for each probe. (*B*) Concentration-dependent $^{15}N$ CPMG profiles for Q38 of JTO-$V_L$ (circles) along with fits (solid curves) that include all residues with measurable profiles (maximum $\Delta R_{2,eff} > 10\ s^{-1}$ at one or more concentrations). (*C* and *D*) As in *A* and *B* but for WT WIL-$V_L$. (*E*) Kinetic model used to analyze the chemical shift and dispersion data, along with extracted dimer dissociation constants. Residues with $^{15}N$ dispersions such that $\Delta R_{2,eff} > 3\ s^{-1}$ are highlighted in red on the crystal structures of JTO-$V_L$ [PDB ID code 1CD0 (27)] and WIL-$V_L$ [PDB ID code 2CD0 (27)] dimers; one chain is rendered as a gray surface while the other as a yellow cartoon. $K_d$ values were estimated via global fits of both chemical shifts and CPMG data (JTO) or from protein concentration-dependent chemical shift changes alone (WIL; see *SI Appendix*). (*F*) Extracted free energy values, $\Delta G$, as a function of residue from hydrogen/deuterium exchange measurements for JTO-$V_L$ (black, 0.5 mM) and WIL-$V_L$ (gray, 0.8 mM). $\Delta G$ values include contributions from dimerization; inherent stabilities of the monomeric structures can be obtained on a per-residue basis by subtracting 1.4 and 0.1 kcal/mol from largest $\Delta G$ values for JTO and WIL, respectively (see *SI Appendix* for details). (*G*) Aggregation of JTO-$V_L$ (11 profiles in black) and WIL-$V_L$ (12 profiles in gray) as followed by ThT fluorescence.

monomeric and dimeric structures. This was confirmed by a concentration-dependent Carr–Purcell–Meiboom–Gill (CPMG) relaxation dispersion (31–33) study in which the $^{15}N$ transverse relaxation rates of individual backbone amides ($R_{2,eff}$) were quantified as a function of the frequency of application of chemical shift refocusing pulses ($\nu_{CPMG}$). In this experiment, the stochastic modulation of chemical shifts by conformational exchange leads to increased transverse relaxation rates, whose effects can be modulated through the successive application of pulses. As the number of pulses increases (larger $\nu_{CPMG}$) the effective chemical shift differences between the interconverting states become smaller, so that the dispersion profiles are decreased. Notably, the size of the dispersion profiles, $\Delta R_{2,eff}$, as defined by the difference between $R_{2,eff}$ values at the lowest and highest $\nu_{CPMG}$ values, increases with decreasing total protein concentrations (Fig. 2*B*), establishing an equilibrium favoring the dimeric structure at the concentrations explored. The data could be well fit to a two-state monomer–dimer equilibrium model with $K_D^V = 12.5 \pm 1.2\ \mu M$. A similar study of the WIL-$V_L$ domain was also performed, showing a linear progression of chemical shifts as a function of increasing concentration toward V domain peaks in the full length LC (Fig. 2*C*). Notably, and in contrast to what was observed for JTO-$V_L$, dispersion profile sizes increased with total protein concentration, indicating an equilibrium that is highly skewed toward the monomer in this case (Fig. 2*D*). Dispersion data for the WIL-$V_L$ domain could not be fit to a two-state model of exchange, suggesting a more complex set of dimeric interactions than for JTO. A value of $K_D^V \sim 5$ mM was estimated from concentration-dependent chemical shift changes that were fit for 13 well-resolved resonances (*SI Appendix*, Fig. S2). Interestingly, we have performed a similar analysis for a $V_L$ domain from a second LC AL protein, ALMC-$V_L$ (34), that also established a large $K_D^V \sim 5$ mM. Fig. 2*E* highlights the regions of the $V_L$ domains that show $\Delta R_{2,eff}$ values in excess of 3 $s^{-1}$ for JTO and WIL concentrations of 0.2 mM and 1 mM, respectively, corresponding to elements of structure that are sensitive to the conformational exchange process. These are in general agreement with the $V_L$–$V_L$ dimer interface determined by X-ray crystallography (27). In the case of WIL-$V_L$ an additional region that reports on exchange is observed, corresponding to residues I29–N31, which is consistent with a more complex association process than for JTO, as observed in the analysis of the $^{15}N$ CPMG dispersion data.

To evaluate relative stabilities of the domains, per-residue $\Delta G$ values were obtained for both $V_L$ constructs by carrying out hydrogen exchange experiments in which the intensities of correlations in $^1H$-$^{15}N$ HSQC spectra were monitored as a function of time after addition of $D_2O$ buffer (pD 5) to lyophilized protein (35, 36). The exchange rates, $k_{ex}$, so obtained were subsequently converted into $\Delta G$ values via the relation $\Delta G = -RTln(k_{ex}/k_{rc})$, using tabulated values for $k_{rc}$, the exchange rate of the corresponding completely unprotected residue (37). The analysis of exchange rates in terms of protection factors, as done here, is only justified under EX2 conditions (35). Intrinsic hydrogen exchange values are on the order of 0.1 $s^{-1}$ at pH 5 and 37 °C, an order of magnitude slower than the measured folding rate for JTO-$V_L$ (3 $s^{-1}$) under identical conditions that establishes the EX2 regime (35). $\Delta G$ values were uniformly higher for the JTO domain, reflecting, in part, that the protein is predominantly dimeric at the concentration used, while the WIL domain is largely a monomer. The contribution of the monomer–dimer equilibrium to the observed $\Delta G$ values can be accounted for (*SI Appendix*) and subtracted out so that opening/closing free energies, $\Delta G_{UN}$, are obtained on a per-residue basis from the hydrogen exchange data. The $\Delta G_{UN}$ values have similar patterns as for the $\Delta G$ profile in Fig. 2*F*, with the values for amides exchanging with the largest protection factors (largest $\Delta G_{UN}$) lowered by ~1.4 and 0.1 kcal/mol for JTO and WIL, respectively, at the 0.5 and 0.8 mM concentrations of JTO-$V_L$ and WIL-$V_L$ domains used in this study. The largest protection factors (largest $\Delta G_{UN}$) report directly on global stabilities of each protein, $\Delta G_{unf}$ (38), so that the monomeric JTO domain is ~1.5 kcal/mol more stable than its WIL counterpart. Thus, the stability of the JTO-$V_L$ domain is enhanced both by an increased propensity to form dimers and by an increased intrinsic stability of the folded, monomeric state. WIL-$V_L$ aggregates more readily than JTO-$V_L$ (Fig. 2*G*), consistent with its lower thermodynamic stability and lower propensity to form protective dimers.

**$C_L$ Domains Associate Weakly.** An extensive interface of 2,400 $Å^2$ is computed between adjacent $C_L$ domains from X-ray–derived structures of the FL λ-LCs (10), close to 1.5-fold larger than observed for their $V_L$ domain counterparts whose dimerization is known to be stabilizing and protective against aggregation (17, 29). It is of interest, therefore, to evaluate the relative stability of the isolated $C_L$–$C_L$ interaction focusing on (*i*) the importance of the interdomain disulfide bond connecting pairs of $C_L$ domains and on (*ii*) the significance of the *cis*-peptide bond at position 145. This was achieved by studying the isolated WT $C_L$ domain,

and a $C_L$ domain harboring C215S+P145A or C215S mutants (the isolated P145A mutant is difficult to obtain in pure form). Recall that the $C_L$ domains are identical for JTO and WIL LC constructs considered here. The $^{15}N$ CPMG relaxation dispersion profiles of C215S $C_L$ were linear, with $R_{2,eff}$ values decreasing over the 1,000 Hz $\nu_{CPMG}$ frequency range examined, consistent with an exchange process that is fast on the chemical shift timescale. Further, as with the WIL-$V_L$ domain, dispersion profiles increased with protein concentration, as would be expected for a monomer ($M$)–dimer ($D$) exchange process that is highly skewed toward $M$. All dispersion curves, along with chemical shift titration data recorded as a function of the concentration of C215S $C_L$, were analyzed together using a two-state exchange model, $2M \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} D$ (*SI Appendix*), and a value of $k_{-1} \sim 10{,}000\ s^{-1}$ was obtained (Fig. 3*A*, *Inset*). The fast exchange process results in linear changes to chemical shifts that depend on the relative populations of interconverting monomer and dimer (39), as illustrated for L139 in Fig. 3*B*. Here subspectra for the WT construct (blue), as well as for the C215S (green) and P145A+C215S (red) mutants, recorded in some cases at a number of protein concentrations, are superimposed, along with the amide correlation for L139 in the FL protein where the $C_L$ domains are fully dimeric (10) that is shown for reference (black, Fig. 3*B*). The concentration-dependent chemical shifts for each of the mutant domains can be quantified (Fig. 3*C*), as can the concentration-independent positions of the peaks from WT $C_L$ (dimer is disulfide-linked) that together allow extraction of $K_D^C$ values, with excellent agreement between fitted and observed chemical shifts obtained (Fig. 3*D*). Notably, from the slopes of the chemical shift vs. concentration profiles for the C215S and P145A+C215S $C_L$ domains (Fig. 3*C*, cf. red vs. green) it is clear that the dimerization constants for the two mutants are very different, as there are only very small changes in shifts in the case of the double mutant, which implies little association between monomers. Backbone amides from C215S $C_L$ that have substantial protein concentration-dependent chemical shift changes localize to a single β-strand surface of the constant domain and are color-coded in red on one of the two $C_L$ monomers comprising a $C_L$–$C_L$ dimer (Fig. 3*D*, *Inset*). As expected, the dimer interface regions show the largest chemical shift perturbations.

Our results clearly establish the importance of the interchain disulfide to dimer stabilization, as the equilibrium dissociation constant for the WT $C_L$ domain is 0.18 ($K_{D1}^C$, unimolecular reaction), while for the C215S mutant the dissociation constant for the bimolecular reaction ($K_{D2}^C$) is 11 mM (Fig. 3*E*). The effective molarity (40, 41) for the unimolecular reaction is given by $K_{D2}^C/K_{D1}^C$ = 63 mM (Fig. 3*E*), which is the concentration of the $C_L$ domain for which the uni- and bimolecular reactions proceed with the same forward rates. Our results also show that the double mutant P145A+C215S $C_L$, in which a *cis*-peptide bond is replaced by one that is *trans*, further destabilizes the $C_L$–$C_L$ dimer as $K_{D2}^C$ = 129 mM. Insight into why this is the case can be obtained by comparing backbone amide chemical shifts from spectra recorded of the single and double mutants at low concentration (25 μM) where both mutants are exclusively monomeric. Fig. 3*F* highlights those residues where shift differences exceed 0.1 ppm (magenta) and 0.4 ppm (red), with residue 145 drawn as a sphere. Regions that are affected by the P145A mutation extend to approximately half of the domain, including the dimer interface β-sheet, suggesting that subtle structural rearrangements are responsible for the lower affinity.

**Stabilities of Dimeric Interfaces Are Increased in the FL Protein.** Having characterized the oligomerization properties of isolated $V_L$ and $C_L$ domains from JTO and WIL, including domains harboring destabilizing mutations, we next focused on the FL

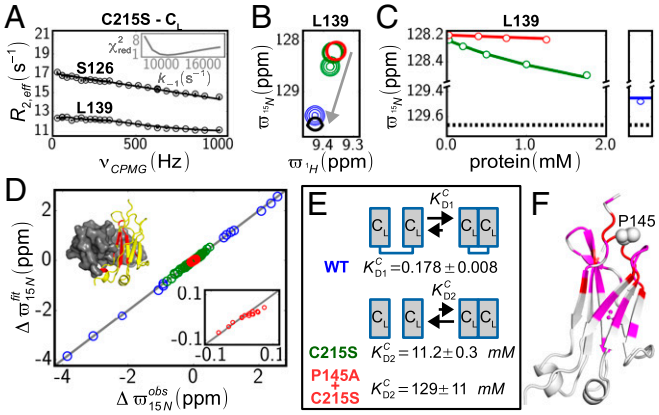


**Fig. 3.** $C_L$ domains associate weakly. (*A*) $^{15}N$ CPMG profiles of S126 and L139 of C215S $C_L$, 1.76 mM, 37 °C. The linear decrease in $R_{2,eff}$ vs. $\nu_{CPMG}$ is the hallmark of fast exchange on the NMR chemical shift timescale. The solid curves are fits of the data (circles); dispersion data were fit simultaneously with concentration-dependent chemical shift data (*B* and *C*) to a monomer–dimer exchange model, constraining $\Delta\varpi$ values to be the difference between shifts of the monomer as measured for C215S $C_L$ at 0.025 mM and the dimer as measured for WT WIL-FL. (*B*) Chemical shift titration of L139 as a function of protein concentration and mutation: red, P145A+C215S $C_L$, 0.025 mM; green, C215S $C_L$, 0.025 mM (single thick contour), 1.76 mM (multiple thin contours); blue, WT $C_L$; black, WT WIL-FL. (*C*) Concentration-dependent $^{15}N$ chemical shifts for L139 in P145A+C215S $C_L$ (red), C215S $C_L$ (green), and WT $C_L$ (blue; no concentration dependence due to C215 disulfide linkage); the $^{15}N$ chemical shift of L139 in WT WIL-FL is shown by dashed lines. (*D*) Comparison between observed and fitted chemical shift differences, $\Delta\varpi$ (inset for P145A+C215S $C_L$ only), for WT $C_L$ (blue), C215S $C_L$ (green) and P145A+C215S $C_L$ (red), where $\Delta\varpi = \varpi_i - \varpi_m$ and $\varpi_i$ and $\varpi_m$ are shifts for residue $i$ under the conditions of the experiment and for the corresponding residue in C215S $C_L$ (or P145A+C215S $C_L$; see *SI Appendix*) at 0.025 mM where the domain is monomeric, respectively. (*E*) Kinetic models and calculated equilibrium constants for the three $C_L$ domains studied. $K_d$ values were estimated via global fits of protein concentration-dependent chemical shift changes, assuming fast exchange (*SI Appendix*, Eq. **S2**). (*F*) Chemical shift differences between P145A+C215S and C215S $C_L$ domains at 0.025 mM (monomeric) superimposed on the structure of the $C_L$ domain [from PDB ID code 1LIL (10)] that has the same orientation as for the inset in *D*; Pro-145 is labeled and drawn with a balls and sticks representation. Residues shown in magenta and red have combined $\Delta\varpi_{HN}$ differences (*SI Appendix*) between P145A+C215S $C_L$ and C215S $C_L$ samples larger than 0.1 (magenta) and 0.4 ppm (red). All NMR data were recorded at 14.1 T.

protein in an effort to understand how domain–domain interactions in FL LCs might confer further stability. In particular, we used a strategy involving the introduction of specific mutations (*SI Appendix*, Fig. S3 and Table S1) that destabilize a particular interface in the dimer so as to quantify the importance of the remaining domain interactions in preserving the dimeric structure of the LC. We have obtained backbone resonance assignments for WT and mutant FL LC constructs so that perturbations made at a position of interest can be studied simultaneously in both $V_L$ and $C_L$ domains using large numbers of probes. In what follows we focus, initially, on the $V_L$ domain in the context of the FL LC, introducing the C215S or P145A+C215S substitutions that have been discussed previously, as well as F122D and F101D, because residues 122 and 101 are localized to the $C_L$–$C_L$ and $V_L$–$V_L$ interfaces, respectively. The latter two mutations introduce negative charges at the interfaces that, as we show below, are destabilizing.

Fig. 4*A* shows a superposition of regions of $^1H$-$^{15}N$ HSQC spectra recorded on F101D (orange) and F101D+C215S (magenta) JTO-FL proteins, focusing on peaks from Q38 and Y50 of the $V_L$ domains. By means of comparison the corresponding correlations for the WT dimeric JTO-$V_L$ domain are shown (gray, labeled D). Notably, the positions of the peaks in both

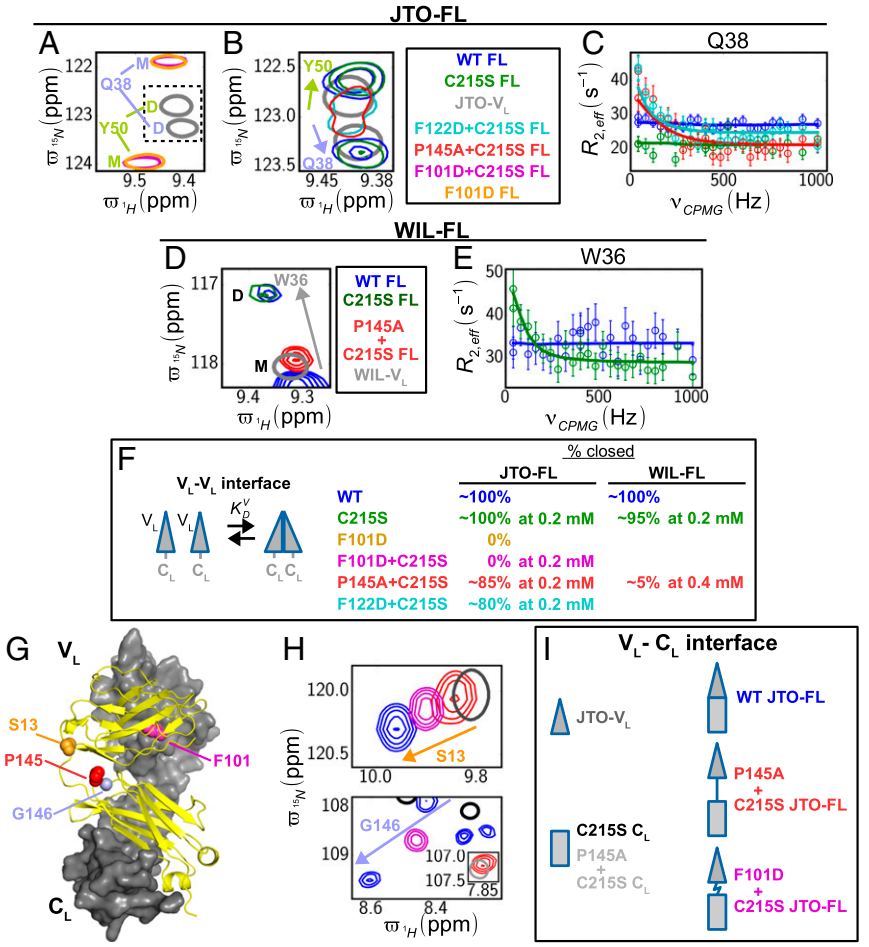


**Fig. 4.** The $V_L$ interface is affected by the stability of the $C_L$ interface. (*A* and *B*) Superposition of selected regions of $^1H$,$^{15}N$ HSQC spectra of WT and mutants of JTO-FL recorded at 37 °C and 14.1 T highlighting Y50 and Q38 that report on the $V_L$–$V_L$ interface, color-coded as shown to the right of *B*. The dashed rectangle in *A* indicates the region that is zoomed in *B*. Arrows in *B* point in the direction of chemical shift changes toward the dimer (WT FL). (*C*) The $^{15}N$ transverse relaxation optimized spectroscopy (TROSY)-CPMG dispersion profiles for Q38 from JTO-FL variants; only P145A+C215S (red) and F122D+C215S (teal) JTO-FL constructs show nonflat profiles that indicate a monomer–dimer equilibrium at the level of the $V_L$–$V_L$ interface, heavily skewed toward the dimer. Differences in plateau values of dispersion curves reflect the different protein concentrations used (*SI Appendix*). (*D* and *E*) As *A*–*C* but for WIL-FL. (*F*) Kinetic model and percentage of closed (WT-like) $V_L$–$V_L$ interfaces for variants of JTO-FL and WIL-FL. Fraction of closed interface for each mutant was calculated via a global fit of chemical shifts at a single concentration, assuming fast exchange. Concentrations are listed for all C215S FL mutants where the $C_L$ chains are not covalently linked (i.e., where formation of the closed structure is bimolecular). (*G*) Positions of key residues (S13 and G146) reporting on the integrity of the $V_L$–$C_L$ interface plotted on a homology model of the JTO-FL structure; one chain is rendered as a gray surface and the other as a cartoon; P145 and F101, which are mutated in the study, are also shown. (*H*) Superposition of small regions of $^1H$,$^{15}N$ HSQC spectra of WT JTO-FL focusing on S13 or G146. Peaks are color-coded as indicated in *I* that also shows either $V_L$ (S13) or $C_L$ (G146) domains that were used in the study, or LCs highlighting the nature of the $V_L$–$C_L$ interface (WT-like for WT JTO-FL, not formed for P145A+C215S JTO-FL or rapidly interconverting between WT and not formed for F101D+C215S JTO-FL, as established from the positions of S13 and G146 peaks). For simplicity only the monomeric forms of JTO-$V_L$ and FL LCs are drawn, although JTO-$V_L$, WT JTO-FL, and P145A+C215S JTO-FL are dimeric.

JTO-FL mutants have significantly changed relative to JTO-$V_L$ (or WT JTO-FL), by −1.5 and +1.4 ppm in the $^{15}N$ dimension for Q38 and Y50, respectively, consistent with the shift changes measured for the JTO-$V_L$ monomer–dimer transition using $^{15}N$ CPMG relaxation dispersion experiments, $\Delta\varpi_N = -1.3$ ppm (Q38), +1.3 ppm (Y50). Thus, the F101D mutation severely disrupts the $V_L$–$V_L$ interface and, as we show below, the double mutation F101D+C215S leads to a predominantly monomeric LC. In contrast, mutations to the $C_L$ domain (C125S, F122D+C215S, and P145A+C215S) have smaller effects on the $V_L$–$V_L$ interaction, as can be seen in Fig. 4*B*, where cross-peaks from Q38 and Y50 are located at positions very similar to those from the WT JTO-$V_L$ domain or WT JTO-FL, strongly suggesting that these proteins have predominantly native-like dimeric $V_L$ interfaces. Further insight into how the $V_L$ interface is affected by mutations that destabilize interactions between $C_L$ domains is obtained by recording $^{15}N$ CPMG dispersion profiles on the JTO-FL protein (Fig. 4*C*) because this class of experiment can be sensitive to interconverting populations as low as ∼1%. Notably, the WT FL LC shows flat profiles for $V_L$ domain residues (blue), yet nonflat dispersion curves are observed for a number of $V_L$ domain residues in both F122D+C215S (teal) and P145A+C215S (red) FL proteins that highlight a decreased stability of the $V_L$–$V_L$ interface when the dimeric $C_L$ interactions are broken.

Similar to JTO-FL, breaking the $C_L$ interface affects distal interactions between $V_L$ domains in the WIL-FL LC. The importance of the $C_L$ interface is made particularly clear by noting that (*i*) isolated WT WIL-$V_L$ is monomeric (Fig. 2 *C*–*E*) yet WT WIL-FL is largely dimeric with respect to $V_L$–$V_L$ interactions (Fig. 4*D*, compare blue vs. gray) and that (*ii*) the P145A+C215S double mutation largely eliminates the WIL-$V_L$–$V_L$ dimer interface (Fig. 4*D*, red vs. blue). Further, while the $V_L$–$V_L$ interface in the C215S WIL-FL mutant is largely present, based on the proximity of $V_L$ domain chemical shifts for both C215S WIL-FL and WT WIL-FL proteins

(Fig. 4*D*, green vs. blue), the nonflat dispersion profiles for some $V_L$ residues in C215S WIL-FL but not WT WIL-FL (Fig. 4*E*) suggest a destabilization of the $V_L$ domain interface in the context of the mutant protein.

In an effort to quantify further how the mutations affect the $V_L$–$V_L$ interface in the context of the FL protein, we compared the positions of cross-peaks of key reporter residues in each of the mutants relative to the corresponding positions in references with either fully dimeric or monomeric $V_L$ interfaces, as described above. In the case of the JTO-FL mutant proteins, WT JTO-FL (dimer) and F101D+C215S JTO-FL (0.25 mM, monomer) were used as references, while WT WIL-FL (dimer) and the isolated WIL-$V_L$ (0.25 mM, monomer) were used for WIL. Results from both JTO-FL and WIL-FL are summarized in Fig. 4*F* (*SI Appendix*, Fig. S4) where "closed" refers to the $V_L$–$V_L$ WT-like interface where the $V_L$ domains are dimeric. Notably, mutations in the $C_L$ domain can significantly affect the stability of the $V_L$–$V_L$ interface, shifting the closed–open equilibrium by as much as 20% from the WT fully closed conformation (at 0.2 mM LC concentration).

We next asked how weakening of $V_L$–$V_L$ or $C_L$–$C_L$ interactions would affect the $V_L$–$C_L$ interface that sequesters ∼500 Å$^2$ of surface area (10) (Fig. 4*G*). Fig. 4*H* shows spectra from S13 and G146 reporters that are proximal to the interface, located on the $V_L$ and $C_L$ domains, respectively. Interestingly a linear titration of cross-peak position as a function of mutation is noted that suggests, most simply, a rapid exchange between conformers where the $V_L$–$C_L$ interaction is either intact or broken. The data establish that the *cis*-peptide bond at position 145 is critical for stabilizing the interface, as S13 and G146 cross-peaks are at similar positions in spectra of the P145A+C215S JTO-FL construct and of either the WT JTO-$V_L$ domain (S13; compare gray and red contours) or the P145A+C215S $C_L$ domain (G146; compare gray and red, inset), with the isolated $V_L$ and $C_L$ domains, of course, lacking the interface. Simultaneous weakening of the $V_L$–$V_L$ and $C_L$–$C_L$ interfaces through the F101D and C215S substitutions, respectively, leads to an intermediate situation that reflects rapid exchange between conformers with and without the heterodomain interface with approximately equal populations (purple), as established by chemical shifts for S13 and G146 that are between those of the FL protein and of the individual domains.

As described above, the dimeric interface between isolated $C_L$ domains is weak and easily disrupted by removal of the interdomain disulfide bond. We were thus interested in establishing whether the presence of the $V_L$ domains in the context of the FL protein provides stability to the $C_L$–$C_L$ interface. Fig. 5*A* shows a superposition of selected regions from $^1$H-$^{15}$N HSQC spectra recorded on different mutants of JTO-FL that have been introduced in the previous sections, highlighting a subset of $C_L$ domain residues that report on the $C_L$–$C_L$ interface. Analysis of amide cross-peak positions for these residues and others like them that report on the relative populations of monomeric and dimeric $C_L$–$C_L$ interfaces provides a rapid means of assessing the stability of $C_L$-domain interactions. Notably, and unlike what was observed for the isolated C215S $C_L$ domain (gray), the C215S mutation in the JTO-FL background (green) only slightly decreases the extent of $C_L$–$C_L$ dimerization (200 μM protein concentration). This is readily noted by the similar positions of $C_L$ domain cross-peaks in the C215S JTO-FL and WT JTO-FL proteins, where in the latter case the $C_L$–$C_L$ interface is completely formed (compare green and blue in Fig. 5*A*), as shown in the X-ray structure of the FL LC (10). Thus, the presence of the $V_L$ interface in the full protein stabilizes interactions between the C domains. That this is the case can be further appreciated by the results obtained for the F101D FL protein where $V_L$–$V_L$ contacts are severely reduced by the presence of Asp residues at the $V_L$ dimer interface (Fig. 4). In this case the similar positions of $C_L$-domain cross-peaks for the F101D and C215S JTO-FL proteins suggests, qualitatively, that elimination of the $V_L$–$V_L$ interface through charge repulsion across the $V_L$ dimer (F101D mutant) destabilizes the $C_L$–$C_L$ interface to a similar degree as elimination of the interchain disulfide bond. A value of $K_{D1}^{C} = 0.12 \pm 0.03$ is calculated for the unimolecular dissociation of the C domains (still connected via a disulfide) in the context of the F101D JTO-FL protein that is in good agreement with

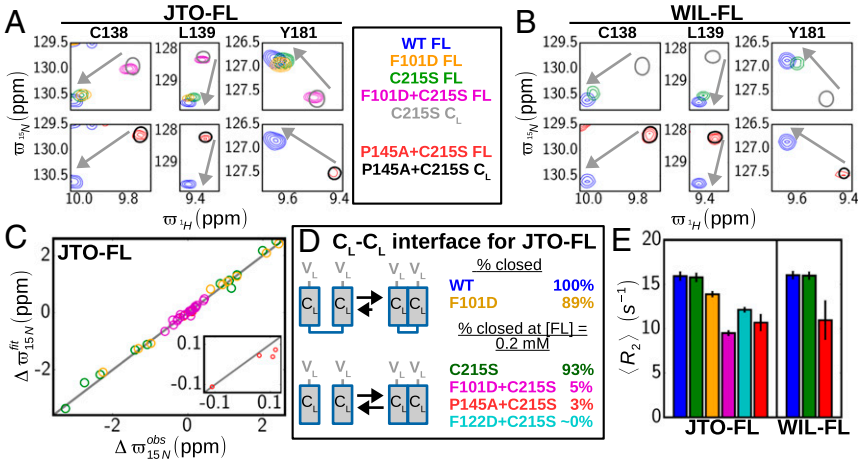


**Fig. 5.** A WT $V_L$ domain interface strengthens the $C_L$–$C_L$ interaction. (*A*) Superpositions of selected regions of $^1$H,$^{15}$N HSQC spectra of WT and mutants of JTO-FL recorded at 37 °C and 14.1 T that report on the $C_L$–$C_L$ interface and color-coded as indicated in the panel to the right. Gray arrows extend from peaks that derive from a monomeric $C_L$ interface (i.e., $C_L$–$C_L$ dimer interaction is not present; gray) to those from a WT-dimer interface (blue). (*B*) As *A* but for WIL-FL. (*C*) Correlation between observed and fitted $\Delta\varpi$ values (inset for P145A+C215S JTO-FL only) for variants of JTO-FL, where $\Delta\varpi = \varpi_i - \varpi_m$ and $\varpi_m$ is the shift for residue *i* from the isolated $C_L$ domain (either P145A+C215 or C215S, depending on whether the residue of interest derives from the P145A or P145 FL construct). (*D*) Kinetic models and percentages of closed (WT-like) $C_L$–$C_L$ interface for FL LCs. The disulfide bond connecting the $C_L$ domains in the C215 samples is indicated by a blue line. Concentrations of closed interfaces were calculated via a global fit of chemical shifts at a single protein concentration, assuming fast exchange. The concentration at which the percent closed $C_L$–$C_L$ interface values were estimated, based on chemical shifts, is indicated for the C215S samples where the LCs are not covalently linked. (*E*) Transverse relaxation rates of TROSY $^{15}$N magnetization components, averaged over structured regions of JTO-FL or WIL-FL with error bars indicating 1 SD. Rates have been adjusted to account for different protein concentrations used in the measurements (*SI Appendix*).

$K_{D1}^{C} = 0.18 \pm 0.01$ obtained from a similar analysis of peak positions for the isolated WT $C_L$ domain (Fig. 3*E*). Finally, when both the $V_L$ interface is eliminated and the disulfide reduced (F101D+C215S double mutant, purple) the interactions between the constant domains are also eliminated, as can be seen by the superposition of peaks from F101D+C215S FL and C215S $C_L$ constructs (compare purple and gray peaks). Recall that the C215S $C_L$ domain is monomeric ($K_d = 11$ mM; Fig. 3*E*). The F101D+C215S double mutant is thus monomeric. The importance of the $V_L$–$V_L$ interface in stabilizing the $C_L$–$C_L$ dimer is also made clear from similar analyses of the WIL-FL proteins establishing that, unlike for the isolated $C_L$ domain, the $C_L$–$C_L$ dimer is predominant in solution in the WT WIL-FL context (Fig. 5*B*, green vs. gray peaks). Our data thus clearly establish that both $V_L$–$V_L$ and $C_L$–$C_L$ interfaces cooperatively maintain the dimeric FL structure.

To quantify interactions between the domains more fully we have fit the relative positions of amide peaks from $C_L$ domain residues in each mutant of JTO-FL (Fig. 5*C*) to extract the extent to which the $C_L$–$C_L$ interface is formed in each case, as summarized in Fig. 5*D*. A value of 93% closed ($C_L$ dimer) was obtained for the C215S JTO-FL mutant, similar to that for the C215S WIL-FL protein. By means of comparison, a value of 3.5% closed is obtained for the isolated C215S $C_L$ domain based on a $K_D^C$ value of 11.2 mM (Fig. 3*E*) and assuming a total protein concentration of 0.2 mM, as was used in measurements of the FL protein, further emphasizing the role of the $V_L$–$V_L$ interface in providing stability to the $C_L$ domain interactions.

A similar picture is also obtained from measured $^{15}$N transverse relaxation rates, $R_2$, for the JTO-FL and WIL-FL constructs considered here. $R_2$ rates are sensitive both to molecular size and to internal dynamics that can change with structure (39, 42). Thus, a comparison of average $R_2$ values, $\langle R_2 \rangle$, in the FL constructs provides further insight into whether domain interfaces are properly formed. Fig. 5*E* plots $\langle R_2 \rangle$ from structured regions of the $C_L$ domains of the FL proteins. The similar $\langle R_2 \rangle$ rates for WT-FL (blue) and C215S FL (green) indicates, as expected based on chemical shift data, that both constructs are similarly dimeric, reinforcing the notion that the $V_L$–$V_L$ interface stabilizes interactions between $C_L$ domains, with a well-formed $C_L$ dimeric structure even in the absence of the disulfide. Notably, $\langle R_2 \rangle$ values for the P145A+C215S FL double mutant (red) decrease relative to the WT protein, consistent with a destabilized $C_L$–$C_L$ interface and in keeping with the nearly identical chemical shifts for constant domain amide probes in the P145A+C215S FL and $C_L$ constructs (Fig. 5 *A* and *B*, red vs. black). As expected, a lowered $\langle R_2 \rangle$ rate for $^{15}$N spins of the $C_L$ domains of the F122D+C215S FL protein is also measured, consistent with destabilization of the $C_L$–$C_L$ dimer interface. Notably, $\langle R_2 \rangle$ for the F101D+C215S FL construct, with destabilizing mutations in both of the $V_L$ (F101D) and $C_L$ (C215S) domains, is approximately half the value of the WT FL protein, consistent with the complete dissociation of the double mutant into monomers.

**Summary of Mutations at Dimeric Interfaces.** The combined chemical shift and relaxation data analysis that has been described allows definitive conclusions to be drawn about the formation of the $V_L$–$V_L$, $C_L$–$C_L$, and intramolecular $V_L$–$C_L$ interfaces for each of the mutants considered, as discussed above. It also establishes that destabilization of the $V_L$–$V_L$ interface has a negative effect on the $C_L$–$C_L$ interface, and vice versa, as summarized in Table 1. We have also included two additional mutants, L76G+C215S and L121G+C215S, in our analysis, not mentioned to this point (*SI Appendix*, Figs. S5–S9). The replacement of Leu at position 76 with a Gly severely destabilizes the $V_L$ domain so that it no longer folds (*SI Appendix*, Figs. S5 and S6). Cross-peaks derived from the $C_L$ domain of L76G+C215S JTO-FL superimpose with those from the isolated C215S $C_L$ domain (monomer), establishing that the $C_L$ domain of the double-mutant protein retains native structure but is not dimeric in the FL protein. Thus, L76G+C215S JTO-FL is monomeric at the millimolar concentrations used in the present study. The substitution of Gly for Leu at position 121 in the hydrophobic core of the isolated $C_L$ domain significantly destabilizes it, with approximately equal populations of both folded and unfolded $C_L$ domain observed in $^{1}$H,$^{15}$N HSQC spectra (*SI Appendix*, Fig. S7). Similar to the isolated L121G+C215S $C_L$ domain, the corresponding region in the FL protein is also in equilibrium between folded and unfolded conformers. Notably, however, the dimeric $V_L$ structure is preserved in this construct (*SI Appendix*, Figs. S8 and S9). We have cross-validated our NMR analyses described above by recording analytical ultracentrifugation (AUC) data on the JTO-FL and WIL-FL mutants (*SI Appendix*, Fig. S10). Very similar conclusions are reached regarding the influence of the mutations on the dimeric structure of the constructs, with the small differences reflecting the fact that the AUC data are recorded at significantly lower concentrations (10 μM) that shift the monomer–dimer equilibrium toward a monomeric state. Notably, the power of the NMR method lies in the fact that individual domains of the FL protein can be interrogated independently, unlike the case for other methods such as AUC where it is only possible to obtain a readout of whether the FL protein is monomeric or dimeric. Having characterized the structural properties of each mutant it is now possible to evaluate the relative importance of each of the domains and their interfaces in stabilizing the soluble LC form, as described subsequently. Free energy differences for the unfolding of $V_L$ ($\Delta G_V$) and $C_L$ ($\Delta G_C$) domains in the context of the FL protein, as measured by solvent hydrogen exchange NMR (35, 36), are reported in Fig. 6*A* (see *SI Appendix* for details and *SI Appendix*, Fig. S11) and Table 1, which also includes a summary of

**Table 1. Formation of interfaces and unfolding thermodynamics based on NMR data**

| Protein construct | $V_L$–$V_L$ | $C_L$–$C_L$ | $V_L$–$C_L$ | $\Delta G_V$ | $\Delta G_C$ |
|---|---|---|---|---|---|
| JTO-FL | | | | | |
| WT | ✓ | ✓ | ✓ | 7.9 ± 0.1 | 8.1 ± 0.1 |
| C215S | ✓ | ✓ | ✓ | 5.9 ± 0.1 | 5.5 ± 0.1 |
| L76G+C215S* | ✗ | ✗ | ✗ | <−2 | 2.9 ± 0.3 |
| F101D | ✗ | ✓ | ✓,✗ | 2.7 ± 0.1 | 5.4 ± 0.1 |
| F101D+C215S | ✗ | ✗ | ✓,✗ | 2.3 ± 0.1 | 4.1 ± 0.2 |
| L121G+C215S† | ✓ | ✓ | ✓ | 3.7 ± 0.2 | 0.0 ± 0.2 |
| F122D+C215S | ✓ | ✗ | ✓,✗ | 4.5 ± 0.1 | 3.2 ± 0.3 |
| P145A+C215S | ✓ | ✗ | ✗ | 3.9 ± 0.1 | <1.5 |
| WIL-FL | | | | | |
| WT | ✓ | ✓ | ✓ | 4.8 ± 0.1 | 7.3 ± 0.2 |
| C215S | ✓ | ✓ | ✓ | 2.7 ± 0.1 | 4.7 ± 0.2 |
| P145A+C215S | ✗ | ✗ | ✗ | n.a.‡ | n.a.‡ |

Details of NMR experiments are described in *SI Appendix*. Measurements were made at protein concentrations of 0.2 mM (monomer), with the exception of L76G+C215S JTO-FL (0.1 mM), WT JTO-FL (1.1 mM; note that ΔG values are concentration-independent as monomers are disulfide-linked), and C215S WIL-FL (0.4 mM). Check marks (crosses) indicate populations of closed interfaces higher than or equal to 80% (lower than 20%). In cases where the closed interface fractional population lies between 20% and 80% both check marks and crosses are indicated. ΔG values are given in kilocalories per mole. n.a., not available.

*$V_L$ domain is not able to fold so that the $V_L$–$V_L$ and $V_L$–$C_L$ interfaces are nonexistent while the $C_L$–$C_L$ interface is not present as well.

†Two sets of $C_L$ domain peaks are observed in NMR spectra, derived from molecules with folded and unfolded $C_L$ domains (*SI Appendix*, Fig. S7). For the folded population (~50%) both $C_L$–$C_L$ and $V_L$–$C_L$ interfaces are present, while they are absent in molecules containing unfolded $C_L$ domains.

‡Hydrogen exchange data not measured. (Note that it is expected $\Delta G_C <$ 1.5 as per the JTO construct.)

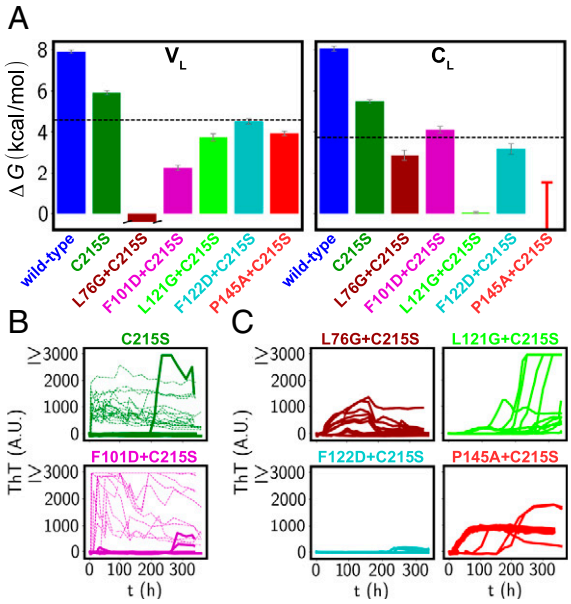


**Fig. 6.** Thermodynamic stabilities and aggregation propensities of LC variants. (*A*) ΔG values for unfolding of $V_L$ and $C_L$ domains in the context of the FL protein. In most cases the column heights and error bars indicate the estimated value and the SE, respectively, except for $V_L$ of L76G+C215S and for $C_L$ of P145A+C215S, where the only possible estimates are upper limits of −2 and +1.5 kcal/mol, respectively. Dashed horizontal lines indicate the $\Delta G_V$ value for WT JTO-$V_L$ and the $\Delta G_C$ value for C215S $C_L$. (*B*) ThT fluorescence-based kinetics of the aggregation of C215S JTO-FL (*Top*, 21 profiles in green thick solid lines) and F101D+C215S JTO-FL (*Bottom*, 18 profiles in purple thick solid lines) and the corresponding JTO-$V_L$ counterparts: WT (*Top*, 17 profiles in green thin dashed lines) and F101D (*Bottom*, 12 profiles in purple thin dashed lines). (*C*) ThT fluorescence kinetics of the aggregation of several JTO-FL mutants indicated at the top of each panel; 12 profiles are shown for each protein.

the interfaces that are formed. All of the constructs studied are listed in *SI Appendix*, Table S1.

**The $C_L$ Domain Is Protective Against Aggregation in FL Proteins.** The factors that influence whether a particular LC will form amyloid fibrils in vivo are not well known. As a first step to address this, we asked how different regions of an LC protect against aggregation in vitro by measuring aggregation of 10 μM LC variants in PBS at pH 7.4, 37 °C with continuous agitation in 96-well plates. We compared the kinetics of fibril formation of FL LC variants using thioflavin T (ThT) binding (43) and observed substantial variation between replicates, which we ascribe to the stochastic nature of the aggregation process under these conditions. A second method for quantifying aggregation using a differential filtration assay (44) was also performed and consistent results were obtained by both methods (*SI Appendix*). ThT binding data for JTO-FL variants is shown in Fig. 6 *B* and *C*, with similar behavior observed for WIL-FL variants (*SI Appendix*, Fig. S12). The morphology of the aggregated material was further inspected by electron microscopy, establishing the formation of fibrillar structures that are very similar for both WIL-FL and JTO-FL proteins (*SI Appendix*, Fig. S13). Both WIL-$V_L$ and JTO-$V_L$ aggregate in a few hours (Fig. 2*G*) but lose their aggregation propensity in the context of C215S FL, which includes the $C_L$ domain (Fig. 6*B*, *Top*). A similar scenario occurs when the F101D mutation, which destabilizes the $V_L$–$V_L$ interface (Fig. 4), is introduced into the $V_L$ domain. Notably, the isolated $V_L$ domain aggregates rapidly, with significantly reduced levels of aggregation for the corresponding C215S FL protein (Fig. 6*B*, *Bottom*). Taken together, these results provide strong evidence that the $C_L$ domains are protective against aggregation in FL proteins and further suggest that destabilization of the $C_L$ region might lead to increased aggregation.

We therefore looked at the aggregation of FL constructs with point mutations located in the $C_L$ domain (L121G+C215S, F122D+C215S, and P145A+C215S; Fig. 6*C*). Strikingly, fibril formation is observed on a timescale of 10–100 h for the two FL proteins with severely destabilized $C_L$ domains, L121G+C215S JTO/WIL-FL ($\Delta G_C = 0.05 \pm 0.04$ kcal/mol for JTO-FL mutant) and P145A+C215S JTO/WIL-FL ($\Delta G_C < 1.5$ kcal/mol for JTO-FL mutant). Aggregation is also observed for F122D+C215S WIL-FL (*SI Appendix*, Fig. S12) but not for F122D+C215S JTO-FL, which likely reflects the stronger $V_L$–$V_L$ interface for the JTO construct.

## Discussion

The most common systemic amyloid disease is LC amyloidosis that results from misfolding and/or aggregation of Ig monoclonal LCs (7, 8), leading to patient death within 1–10 y of onset in the absence of treatment. Although high-resolution X-ray structures of LCs are available (10) showing that they arrange as dimers with extensive $V_L$–$V_L$, $C_L$–$C_L$, and intramolecular $V_L$–$C_L$ interfaces (Fig. 1), sequestering surfaces of ∼1,600, 2,400, and 500 Å$^2$, respectively, the role of these interfaces in preventing aggregation and the importance of the intrinsic $V_L$ and $C_L$ domain stabilities in this regard are not well established. Herein we assess the relative importance of domain interactions in stabilizing the dimeric LC structure and, relatedly, in preventing LC aggregation using a mutagenesis strategy in which mutations are introduced that destabilize one or more interfaces. The resulting structural properties of the mutants are established by NMR spectroscopy, the effect of each mutation on domain stability is quantified by hydrogen exchange NMR experiments, and the mutants are assayed for aggregation by ThT fluorescence measurements.

A number of mutations were used in our study, including those that introduce charges in the $V_L$–$V_L$ (F101D) or $C_L$–$C_L$ (F122D) interfaces or cavities in the cores of $V_L$ (L76G) or $C_L$ (L121G) domains, as well as P145A and C215S substitutions that affect the $C_L$–$C_L$ interface. The P145A mutant, with a *trans* peptide bond at residue 145, acts as a surrogate for a LC folding intermediate, as the WT protein with a *cis*-prolyl peptide bond in its native state is formed via an on-pathway folding intermediate with a *trans* peptide bond at this position (45). Therefore, the P145A mutant may serve as a mimic of a misfolded LC in the plasma, in the absence of the necessary proteostasis machinery to properly refold the protein. The C215S mutation is biologically important as the interchain disulfide is reduced in 10–50% of FL LC proteins in the serum (46, 47).

We first evaluated the strengths of the $V_L$–$V_L$ interfaces for individual $V_L$ domains from JTO and WIL λ6 LC proteins. NMR spin relaxation experiments and chemical shift changes as a function of protein concentration establish that the $V_L$ dimer dissociation constant is three orders of magnitude lower for JTO-$V_L$ relative to WIL-$V_L$, while the intrinsic thermodynamic stability of the monomeric JTO-$V_L$ domain is ∼1.5 kcal/mol higher than for the corresponding WIL-$V_L$ monomer. Our results thus provide a rationale for the higher aggregation propensity of WIL-$V_L$ in comparison with JTO-$V_L$ (Fig. 2), in agreement with previous findings (28).

A similar NMR analysis of the isolated $C_L$ domains indicates an unexpectedly weak $C_L$–$C_L$ interface, with a unimolecular $K_d$ value of ∼0.18 for the WT domain (that contains the intermolecular C215S disulfide) and bimolecular $K_d$ values of 11 mM and 130 mM for C215S $C_L$ and P145A+C215S $C_L$, respectively (Fig. 3). These data establish that the intermolecular disulfide is required for the formation of the $C_L$–$C_L$ interface, as the covalent bond increases the effective molarity of the $C_L$

domains. Furthermore, the presence of a nonnative *trans*-prolyl bond at position 145 causes subtle structural changes in the adjacent β-sheet, as evidenced by chemical shift perturbations in NMR spectra, further decreasing the affinity of $C_L$ domain dimerization by at least one order of magnitude.

Our NMR studies of a variety of mutants of FL LCs (Figs. 4–6) show that the strengths of the $V_L$–$V_L$ and $C_L$–$C_L$ interfaces are interrelated. For example, C215S JTO-FL has a 100% closed, WT-like $V_L$–$V_L$ interface and a predominately closed $C_L$–$C_L$ interface (~93% closed), while at the same 200 μM protein concentration the isolated JTO-$V_L$ domain has a monomeric population of about 10% while the isolated C215S $C_L$ domain is largely monomeric. This effect is even more pronounced for C215S WIL-FL where both $V_L$–$V_L$ and $C_L$–$C_L$ interfaces are predominately closed (~95% and ~93%, respectively), despite the fact that the equilibria for the isolated domains are highly skewed toward the monomeric states. This cooperative effect is caused by the fact that when one of the two interfaces forms the second interface is generated through a unimolecular process that is favored by the increased effective molarity of the reactants. The net effect of the increased thermodynamic stability is a decrease in the unfolded state population for C215S FL in comparison with its isolated domains, as measured by solvent hydrogen exchange rates (Fig. 6*A*). Notably, free energy differences obtained by these measurements are in good agreement with calculated values based on measured affinities for $V_L$–$V_L$ and $C_L$–$C_L$ dimerization (*SI Appendix*), validating a proposed unfolding scheme that has been determined for FL Ig LCs (see *SI Appendix* for details).

As expected, introduction of mutations in either $V_L$ or $C_L$ domains or at $V_L$–$V_L$ or $C_L$–$C_L$ interfaces can have disruptive effects and we have exploited these to understand how domain–domain interactions and domain stability influences aggregation propensities of the LC Ig proteins (discussed below). Introduction of Asp residues not only breaks the interface within which they are located but also destabilizes the adjacent interface, due to the cooperative effect described above for the C215S FL protein. For example, the F101D mutation completely disrupts the $V_L$–$V_L$ interface but also leads to a predominantly opened $C_L$–$C_L$ interface (~5% closed) in F101D+C215S JTO-FL, as for the isolated C215S $C_L$ domain. Accordingly, the unfolding free energy for the $C_L$ domain in F101D+C215S JTO-FL, as calculated from hydrogen exchange, $\Delta G_C$, (4.1 ± 0.2 kcal/mol; Table 1) is almost the same as for the isolated C215S $C_L$ domain (3.7 ± 0.2 kcal/mol; Fig. 6*A*). A similar situation occurs for the double mutant F122D+C215S JTO-FL, with an abolished $C_L$–$C_L$ interface, where the strength of the $V_L$–$V_L$ interface is similar to that for the corresponding isolated JTO-$V_L$ domain. The cooperativity between $V_L$–$V_L$ and $C_L$–$C_L$ interfaces is abolished in P145A+C215S JTO-FL, which serves as a mimic of the LC folding intermediate. Like F122D+C215S JTO-FL, P145A+C215S JTO-FL no longer can form a $C_L$–$C_L$ interface, although the $V_L$–$V_L$ interface remains largely closed (~85%) and notably the $V_L$–$C_L$ interface is eliminated in the P145A+C215S double mutant as well. The effect of not forming a *cis* amide bond at position 145 is very destabilizing for the $C_L$ domain of JTO-FL, as evidenced by the very rapid exchange of amide protons with solvent. Finally, two mutations, L76G and L121G, were introduced in the hydrophobic core of $V_L$ and $C_L$ domains, respectively, to probe the effect of thermodynamic stability on aggregation without directly perturbing residues at the $V_L$–$V_L$ and $C_L$–$C_L$ interfaces. As expected, these mutants severely destabilize the domains in which they are introduced, leading to a completely unfolded $V_L$ domain in L76G+C215S JTO-FL and an approximately equal population of folded and unfolded $C_L$ domains in L121G+C215S JTO-FL. The presence/absence of domain interfaces for each of the examined mutants is summarized in Table 1.

Having characterized the structures and thermodynamics of a number of FL mutants we next carried out fibril formation assays (Fig. 6 and *SI Appendix*, Fig. S12) so as to identify the factors that determine aggregation risk in LC proteins. Aggregation rates of FL proteins are much slower than for isolated $V_L$ domains (compare Fig. 2*G* and Fig. 6), implying that the $C_L$ domains play a protective role against aggregation. This protective effect is not just the result of the cooperativity between the $C_L$–$C_L$ and $V_L$–$V_L$ interfaces, nor can it be explained exclusively in terms of the dimeric FL structure, since no aggregation is observed for monomeric F101D+C215S FL proteins. In this context it is noteworthy that the F101D mutation does not prevent fibril formation per se as F101D $V_L$ aggregates rapidly, as does the F101D+L121G+C215S JTO-FL triple mutant (*SI Appendix*, Fig. S12). Destabilization of $C_L$ domains partly abolishes protection against aggregation in FL proteins, as for the L121G+C215S, P145A+C215S and F101D+L121G+C215S variants, but fibril formation of the isolated $V_L$ domains is still at least one order of magnitude faster. Therefore, the presence of the $C_L$ domain slows down the aggregation cascade of FL proteins, even when the $C_L$ domain does not form native structure. The cause of this residual protection may reflect the fact that the $C_L$ domain is much less prone to form fibrils than the $V_L$ domain, as observed in our ThT fluorescence based assays (*SI Appendix*, Fig. S12), and/or that the unfolded $C_L$ domain may lead to off-pathway interactions that divert the protein from the fibril cascade.

Notably, fibril formation is observed in cases where either of the $V_L$ or $C_L$ domains in the FL protein is severely destabilized, with $\Delta G_V$ or $\Delta G_C$ values less than 2 kcal/mol (L76G+C215S, L121G+C215S, P145A+C215S, and F101D+L121G+C215S). Our data, however, do not allow us to distinguish whether unfolding of a single domain is sufficient for aggregation or whether ultimately both domains must lose their native structure to be accommodated in the fibril architecture. It is worth noting that while $\Delta G_i$ values less than 2 kcal/mol are predictors of the formation of fibrillar aggregates, is not possible to rank order aggregation propensity on the basis of thermodynamic stability alone. For example, the monomeric F101D+C215S JTO-FL ($\Delta G_V$ = 2.3 ± 0.1 kcal/mol) is less prone to form fibrils than WT JTO-$V_L$ ($\Delta G_V$ = 4.6 ± 0.1 kcal/mol), despite its lower thermodynamic stability.

Taken together our results demonstrate that differences in fibril formation between isolated $V_L$ domains and FL proteins cannot be ascribed only to the protective role of dimeric interfaces or only to the thermodynamic stability of $V_L$ domains. Instead, the $C_L$ domains play a critical role in protection. The fact that $C_L$ domains are more conserved than their $V_L$ counterparts may establish the $C_L$–$C_L$ interface as a preferred target for the development of drugs against AL. In fact, alignment with the Clustal Omega program (48) shows that the constant domains within the lambda family are 89% identical and 98% similar, while all human constant domains, including both kappa and lambda families, are 33% identical and 76% similar. A drug that tightly binds to the $C_L$–$C_L$ interface would be expected to (*i*) boost the population of closed states for both $V_L$–$V_L$ and $C_L$–$C_L$ interfaces, thus increasing both $\Delta G_V$ and $\Delta G_C$ so as to protect against aggregation, as well as to (*ii*) stabilize the native state of $C_L$ and hence reduce the risk of proteolysis within the $C_L$ domain, thus avoiding the release of aggregation-prone fragments containing the $V_L$ domain. It is clear that a rigorous understanding of the physical chemistry of FL immunoglobulin LC molecules will aid in the rational development of therapeutics to stabilize these proteins and hence prevent their aggregation.

## Materials and Methods

All NMR experiments were performed on a Bruker Avance III HD 14.1 T spectrometer equipped with a cryogenically cooled, x,y,z pulse-field gradient triple-resonance probe. All spectra were recorded at 37 °C using protein

concentrations that varied between 25 μM and 2 mM, depending on the application. Detailed information concerning preparation of samples for NMR, NMR experiments and data analysis, aggregation assays, electron microscopy, analytical ultracentrifugation, and a multistep model for unfolding of FL LCs with a C215S mutation is provided in *SI Appendix*.

**ACKNOWLEDGMENTS.** This work was supported by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada (to L.E.K.) and by NIH Grant DK46335 and the Skaggs Institute for Chemical Biology (to J.W.K.). L.E.K. holds a Canada Research Chair in Biochemistry.

1. Chiti F, Dobson CM (2017) Protein misfolding, amyloid formation, and human disease: A summary of progress over the last decade. *Annu Rev Biochem* 86:27–68.
2. Pinney JH, Hawkins PN (2012) Amyloidosis. *Ann Clin Biochem* 49:229–241.
3. Moreno-Gonzalez I, Soto C (2011) Misfolded protein aggregates: Mechanisms, structures and potential for disease transmission. *Semin Cell Dev Biol* 22:482–487.
4. Gadad BS, Britton GB, Rao KS (2011) Targeting oligomers in neurodegenerative disorders: Lessons from α-synuclein, tau, and amyloid-β peptide. *J Alzheimers Dis* 24: 223–232.
5. Mankad AK, Shah KB (2017) Transthyretin cardiac amyloidosis. *Curr Cardiol Rep* 19:97.
6. Dungu JN, Anderson LJ, Whelan CJ, Hawkins PN (2012) Cardiac transthyretin amyloidosis. *Heart* 98:1546–1554.
7. Milani P, Merlini G, Palladini G (2018) Light chain amyloidosis. *Mediterr J Hematol Infect Dis* 10:e2018022.
8. Falk RH, Alexander KM, Liao R, Dorbala S (2016) AL (light-chain) cardiac amyloidosis: A review of diagnosis and therapy. *J Am Coll Cardiol* 68:1323–1341.
9. Mukherjee A, Morales-Scheihing D, Butler PC, Soto C (2015) Type 2 diabetes as a protein misfolding disease. *Trends Mol Med* 21:439–449.
10. Huang DB, Ainsworth C, Solomon A, Schiffer M (1996) Pitfalls of molecular replacement: The structure determination of an immunoglobulin light-chain dimer. *Acta Crystallogr D Biol Crystallogr* 52:1058–1066.
11. Glenner GG, Terry W, Harada M, Isersky C, Page D (1971) Amyloid fibril proteins: Proof of homology with immunoglobulin light chains by sequence analyses. *Science* 172:1150–1151.
12. Glenner GG (1980) Amyloid deposits and amyloidosis. The beta-fibrilloses (first of two parts). *N Engl J Med* 302:1283–1292.
13. Glenner GG (1980) Amyloid deposits and amyloidosis: The beta-fibrilloses (second of two parts). *N Engl J Med* 302:1333–1343.
14. Morgan GJ, Kelly JW (2016) The kinetic stability of a full-length antibody light chain dimer determines whether endoproteolysis can release amyloidogenic variable domains. *J Mol Biol* 428:4280–4297.
15. Eulitz M, Linke RP (1993) The precursor molecule of a V lambda II-immunoglobulin light chain-derived amyloid fibril protein circulates precleaved. *Biochem Biophys Res Commun* 194:1427–1434.
16. Bellotti V, et al. (1990) Relevance of class, molecular weight and isoelectric point in predicting human light chain amyloidogenicity. *Br J Haematol* 74:65–69.
17. Brumshtein B, et al. (2014) Formation of amyloid fibers by monomeric light chain variable domains. *J Biol Chem* 289:27513–27525.
18. Baden EM, et al. (2008) Altered dimer interface decreases stability in an amyloidogenic protein. *J Biol Chem* 283:15853–15860.
19. Brumshtein B, et al. (2015) Inhibition by small-molecule ligands of formation of amyloid fibrils of an immunoglobulin light chain variable domain. *eLife* 4:e10935.
20. Enqvist S, Sletten K, Westermark P (2009) Fibril protein fragmentation pattern in systemic AL-amyloidosis. *J Pathol* 219:473–480.
21. Klimtchuk ES, et al. (2010) The critical role of the constant region in thermal stability and aggregation of amyloidogenic immunoglobulin light chain. *Biochemistry* 49: 9848–9857.
22. Olsen KE, Sletten K, Westermark P (1998) Fragments of the constant region of immunoglobulin light chains are constituents of AL-amyloid proteins. *Biochem Biophys Res Commun* 251:642–647.
23. Vrana JA, et al. (2009) Classification of amyloidosis by laser microdissection and mass spectrometry-based proteomic analysis in clinical biopsy specimens. *Blood* 114: 4957–4959.
24. Johnson SM, Connelly S, Fearns C, Powers ET, Kelly JW (2012) The transthyretin amyloidoses: From delineating the molecular mechanism of aggregation linked to pathology to a regulatory-agency-approved drug. *J Mol Biol* 421:185–203.
25. Bonvin AM, Boelens R, Kaptein R (2005) NMR analysis of protein interactions. *Curr Opin Chem Biol* 9:501–508.
26. Zuiderweg ER (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* 41:1–7.
27. Pokkuluri PR, Solomon A, Weiss DT, Stevens FJ, Schiffer M (1999) Tertiary structure of human lambda 6 light chains. *Amyloid* 6:165–171.
28. Wall J, et al. (1999) Thermodynamic instability of human lambda 6 light chains: Correlation with fibrillogenicity. *Biochemistry* 38:14101–14108.
29. Qin Z, Hu D, Zhu M, Fink AL (2007) Structural characterization of the partially folded intermediates of an immunoglobulin light chain leading to amyloid fibrillation and amorphous aggregation. *Biochemistry* 46:3521–3531.
30. Hurle MR, Helms LR, Li L, Chan W, Wetzel R (1994) A role for destabilizing amino acid replacements in light-chain amyloidosis. *Proc Natl Acad Sci USA* 91:5446–5450.
31. Korzhnev DM, Kay LE (2008) Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: An application to protein folding. *Acc Chem Res* 41:442–451.
32. Hansen DF, Vallurupalli P, Kay LE (2008) An improved 15N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J Phys Chem B* 112:5898–5904.
33. Palmer AG, 3rd, Kroenke CD, Loria JP (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol* 339:204–238.
34. Arendt BK, et al. (2008) Biologic and genetic characterization of the novel amyloidogenic lambda light chain-secreting human cell lines, ALMC-1 and ALMC-2. *Blood* 112:1931–1941.
35. Hvidt A, Nielsen SO (1966) Hydrogen exchange in proteins. *Adv Protein Chem* 21: 287–386.
36. Englander SW, Downer NW, Teitelbaum H (1972) Hydrogen exchange. *Annu Rev Biochem* 41:903–924.
37. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17:75–86.
38. Bai Y, Milne JS, Mayne L, Englander SW (1994) Protein stability parameters measured by hydrogen exchange. *Proteins* 20:4–14.
39. Cavanagh J, Fairbrother WJ, Palmer AG, III, Rance M, Skelton NJ (2007) *Protein NMR Spectroscopy* (Academic, New York), p 912.
40. Hunter CA, Anderson HL (2009) What is cooperativity? *Angew Chem Int Ed Engl* 48: 7488–7499.
41. Mazor MH, Wong CF, McCammon A, Deutch JM, Whitesidesd G (1990) Effective molarity in diffusion-controlled reactions. *J Phys Chem* 94:3807–3812.
42. Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry* 28:8972–8979.
43. Vassar PS, Culling CF (1959) Fluorescent stains, with special reference to amyloid and connective tissues. *Arch Pathol* 68:487–498.
44. Wanker EE, et al. (1999) Membrane filter assay for detection of amyloid-like polyglutamine-containing protein aggregates. *Methods Enzymol* 309:375–386.
45. Morgan GJ, Usher GA, Kelly JW (2017) Incomplete refolding of antibody light chains to non-native, protease-sensitive conformations leads to aggregation: A mechanism of amyloidogenesis in patients? *Biochemistry* 56:6597–6614.
46. Sölling K (1975) Free light chains of immunoglobulins in normal serum and urine determined by radioimmunoassay. *Scand J Clin Lab Invest* 35:407–412.
47. Sølling K (1980) Light chain polymerism in normal individuals in patients with severe proteinuria and in normals with inhibited tubular protein reabsorption by lysine. *Scand J Clin Lab Invest* 40:129–134.
48. Sievers F, Higgins DG (2014) Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol* 1079:105–116.