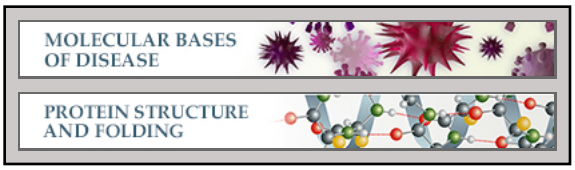
Molecular Bases of Disease:

Formation of Amyloid Fibers by Monomeric Light-chain Variable Domains



Boris Brumshtein, Shannon R. Esswein, Meytal Landau, Christopher M. Ryan, Julian P. Whitelegge, Martin L. Phillips, Duilio Cascio, Michael R. Sawaya and David S. Eisenberg

J. Biol. Chem. *published online August 19, 2014*

Access the most updated version of this article at doi: 10.1074/jbc.M114.585638

JBC Papers in Press. Published on August 19, 2014 as Manuscript M114.585638
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M114.585638

Formation of Amyloid Fibers by Monomeric Light-chain Variable Domains*

Boris Brumshtein[1], Shannon R. Esswein[1**], Meytal Landau[1,2], Christopher M. Ryan[3], Julian P. Whitelegge[3], Martin L. Phillips[4], Duilio Cascio[1], Michael R. Sawaya[1] & David S. Eisenberg[1].**

[1]Department of Biological Chemistry and Department of Chemistry and Biochemistry, University of California Los Angeles, Howard Hughes Medical Institute, UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA.

[**]These authors contributed equally to the work.

[2]Present address: Faculty of Biology, Technion-Israel Institute of Technology, Haifa 3200000, Israel.

[3]The Pasarow Mass Spectrometry Laboratory, NPI-Semel Institute, David Geffen School of Medicine, University of California Los Angeles, CA 90025, USA.

[4]Department of Chemistry and Biochemistry, University of California Los Angeles, UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA.

*Running title: Pathway from Ig light-chains to amyloid fibers

Keywords: amyloid, protein aggregation, antibody, multiple myeloma, X-ray crystallography, light-chain amyloidosis, systemic amyloidosis, Bence-Jones proteins, light-chain variable domains.

Background
Amyloid fibers are protein aggregates associated with numerous pathologies.
Results
Mcg light-chain variable domains form amyloid fibers through monomers.
Conclusion
This light-chain variable domain monomer is the fundamental unit required to form amyloid fibers.
Significance
Understanding the molecular mechanism of Mcg light-chain amyloid fiber formation has implications for treating systemic amyloidosis.

ABSTRACT

Systemic light-chain amyloidosis is a lethal disease characterized by excess immunoglobulin light-chains and light-chain fragments composed of variable domains, which aggregate into amyloid fibers. These fibers accumulate and damage organs. Some light-chains induce formation of amyloid fibers while others do not, making it unclear what distinguishes amyloid formers from non-formers. One mechanism by which sequence variation may reduce propensity to form amyloid fibers is by shifting the equilibrium toward an amyloid-resistant quaternary structure. Here we identify the monomeric form of the Mcg immunoglobulin light-chain variable domain as the quaternary unit required for amyloid fiber assembly. Dimers of Mcg variable domains remain stable and soluble, yet become prone to assemble into amyloid fibers upon disassociation into monomers.

Bence-Jones proteins were first reported over a century and a half ago as disease-associated substances and later shown to be immunoglobulin light-chains involved in systemic light-chain amyloidosis, amyloid formation frequently associated with multiple myeloma (1–5). Light-chains (LC) are classified by amino acid sequence as either of two types, κ or λ, and consist of two domains, the constant (C_L) and the variable (V_L) connected by a joining (J) segment (6). In this disease, immunoglobulin light-chains are over-expressed, increasing the ratio of light to heavy chains. This disturbance of the balance between light and heavy chains results in the formation of LC homo-dimers termed Bence-Jones proteins.

Amyloid fibers are insoluble protein deposits, often associated with disease, which are induced by particular peptide sequences and exhibit common structural and biochemical characteristics. Specifically, all amyloid fibers contain a β-sheet rich spine, which serves as a scaffold for fiber extension and gives rise to a distinctive cross-β X-ray diffraction pattern (7–9). The core of amyloid fibers consists of a pair of β-sheets, whose side chains interdigitate forming a stable spine termed a steric zipper (10). Formation of a steric zipper requires the presence of a short peptide segment with an amino acid sequence capable of forming a self-complimentary propagating structure upon exposure to solvent (11). In systemic LC amyloidosis, both the full-length LCs (V_L-J-C_L) and their V_Ls are found to be constituents of amyloid fibers, yet the ubiquitous presence of V_Ls indicates that this domain may be essential for their assembly (12–17). The peptide segments within V_Ls responsible for forming the steric zipper remain unknown. Due to somatic recombination of immunoglobulins, the LC amino acid sequence is different in every patient; however, the disease has a common progression and pathology leading to the secretion of excess monoclonal LCs through the renal tract (18). Despite renal clearance, full-length LCs or LC fragments that contain the variable domain still form amyloid deposits in patients' tissues, thereby resulting in organ failure (19).

Although excess LCs and V_Ls form amyloid fibers, most of their known molecular structures are homo-dimers with the two variable domains resembling a light-chain and a heavy-chain variable domain in a full antibody (20–22). The V_L homo-dimer resembles the structure of the region of antibodies termed the antigen-binding fragment (Fab), which is composed of both light-chains and heavy-chains (HC), each with constant (C_H) and variable (V_H) immunoglobulin domains. In the Fab, a disulfide bond covalently links the C-termini of the LC and HC, and non-covalent dimer interfaces form hydrophobic cavities between the C_H and C_L and between the V_H and V_L. Similarly to the Fab, a disulfide bond links full-length LC homo-dimers at the C-termini, and the two V_Ls form a non-covalent dimer. The dimer interfaces of both the V_H-V_L hetero-dimer of Fab and pathologic V_L-V_L homo-dimer are lined with apolar residues enclosing a cavity capable of accommodating hydrophobic molecules (23, 24).

The pathway by which variable domains form amyloid fibers remains uncertain; however, there are several hypotheses. Early models of V_L amyloid assembly proposed that continuous semi-crystalline fibers are formed by V_L dimers which preserve their native or a native-like conformation with a conserved hydrophobic dimer interface (25). We term these dimers *canonical dimers* and this model the *canonical dimer model*. This model suggests that canonical V_L dimers, which resemble the structure of antigen-binding variable domains in a full antibody, stack to form non-covalent filamentous polymers with the width of a single V_L dimer. The dimer interface of the two V_L domains would need to form a continuous pair of β-sheets to give rise to the observed cross-β X-ray diffraction pattern. Then, several of these filaments pack into stable amyloid fibers through intermolecular interactions between strands G-F-C-C' of the Greek fold (26). This model proposes that semi-crystalline amyloid fibers contain globular V_Ls in a conformation similar to that observed in crystals (PDB 1REI) (27). Since the canonical dimer model assumes the presence of globular V_Ls with a conserved three-dimensional structure, it also assumes that amyloid formation is reversible.

Though most V_Ls crystallize as canonical dimers, several V_Ls, both κ and λ types, display conformations in which the orientation of domains relative to each other is altered. The most pronounced change includes a rotation up to 180^O around an axis perpendicular to the plane of the dimer interface, while the hydrophobic environment of the cavity between the V_Ls is conserved (28, 29).

These structural alterations suggest the possibility of alternative amyloid formation pathways through which non-canonical dimers, lacking the stability of canonical dimers, assemble amyloid fibers in either the globular or the partially unfolded state. If the globular state is maintained in the fiber, the non-canonical dimers would need to stack as continuous β-sheets, similar to the model for amyloid formation by canonical dimers. Alternatively, the V_Ls of a non-canonical dimer may partially unfold without disassociation in order to form steric zippers, while the hydrophobic interface between domains is maintained. Based on this model, which we term the *non-canonical dimer model*, destabilizing mutations would induce partial unfolding to an amyloid-prone tertiary conformation, whereas stabilizing mutations would inhibit amyloidosis (30–32).

In addition to forming canonical and non-canonical dimers, V_Ls also exist as monomers in solution. Experiments conducted in denaturing conditions indicate that reducing the thermodynamic stability of the monomeric state promotes amyloid fiber formation (33–35). Specifically, the introduction of mutations that induce dimer disassociation or promote monomer unfolding increases the propensity to form amyloid fibers. These findings gave rise yet another model for V_L amyloid assembly: partially unfolded monomers polymerize into amyloid fibers, whereas dimers protect against amyloidosis. We term this the *monomer model*.

Although many V_Ls have been examined and different models for formation of their amyloid fibers have been proposed, it remains unclear which V_L quaternary state promotes amyloid fiber assembly and whether fibers contain unfolded, partially folded, or globular V_Ls (36). Such ambiguity limits our ability to identify peptide segments that form the amyloid spine. If amyloid fibers contain globular dimers with a preserved dimer interface, amyloidogenic segments can map only to solvent accessible areas of V_Ls. However, if unfolded V_L monomers compose amyloid fibers, then amyloidogenic segments can map anywhere within the entire V_L sequence.

We seek to identify the immunoglobulin V_L quaternary state that is the precursor to formation of amyloid fibers and to identify which of the three models best describes the amyloid-forming process (Figure 1). Since the three-dimensional structure of canonical V_L dimers mimics the conformation of physiological Fab antigen-binding domains, which are abundant in human blood as a part of antibodies and are not pathological, the canonical dimers formed by excess V_Ls and full-length LCs may not form amyloid fibers prior to a structural rearrangement. The existence of various non-canonical V_L dimers and monomers suggests a transition between alternative conformations through a monomeric V_L – the quaternary state that may be prone to form amyloid. Here, we investigate whether V_L monomers alone are able to form fibers and thus clarify which of the models explains the molecular pathway of V_L amyloid formation.

EXPERIMENTAL PROCEDURES

Preparation of Mcg mutant proteins. The peptide sequence for the Mcg protein was obtained from the PDB entry 3MCG. The DNA sequence was synthesized by Genscript and cloned into the pET26b+ expression vector. The N-termini of the proteins contained a 6xHis tag followed by a TEV protease cleavage site. Mcg-Y38E-F99A-F101E and Mcg-A45C-F101C were prepared by means of site-directed mutagenesis PCR. Plasmids with the appropriate sequences were cloned into BL21(DE3)Gold (Agilent Technologies) E. Coli expression cells.

TB media were inoculated with expression cultures and allowed to reach $O.D._{600nm} \sim 1.0$. At this density, production of proteins was induced with 1 mM IPTG and performed overnight at $20^{O}C$. Cells were harvested, dissolved in a 50 mM Tris pH8, 100 mM NaCl buffer, sonicated, and a inclusion bodies pellet was separated from the supernatant by centrifugation. V_Ls were found in inclusion bodies of the pellet. Inclusion bodies were washed by dissolving the pellet in 50 mM Tris pH 8, 100 mM NaCl buffer, centrifuged, and separated from the wash supernatant. To solubilize V_Ls from inclusion bodies we used a solution of 7 M urea, 50 mM Tris pH 8, 100 mM NaCl, 0.05% Tween20. V_Ls were purified by His-tag capture on Ni-NTA columns and eluted with 7 M urea, 50 mM Tris pH 8, 100 mM NaCl, 0.05% Tween20, 0.5 M Imidazole. Following Ni-NTA capture, the proteins were dialyzed against a solution of 50 mM Tris pH 8, 100 mM NaCl, 10% Glycerol, 0.1 M L-Arginine, 0.05% Tween20. TEV protease digestion was performed by addition of 1% w/w of TEV protease to the total protein mass and addition of reduced/oxidized glutathione to 3 mM/

0.3 mM respectively. TEV protease digestion was conducted overnight at room temperature and monitored by SDS-PAGE. Urea was added to a final concentration of 7 M and undigested protein, cleaved His-tags, and His-tagged TEV protease were separated from V_Ls by means of Ni-NTA columns. The flow through was collected and contained V_Ls denatured in urea buffer. Proteins were dialyzed against 50 mM Tris pH 8, 100 mM NaCl, 10% Glycerol, 0.1 M L-Arginine, 0.05% Tween20. To assist with proper refolding of the Mcg-A45C-F101C mutant, DTT was added to a final concentration of 5 mM in the dialysis buffer. After refolding, the proteins were concentrated in Centricon devices with a 10 kDa cutoff and washed three times with 25 mM Tris pH 8, 50 mM NaCl.

Amyloid fiber formation assays. ThT assays were performed by diluting the protein with 50 mM Acetic acid/NaAcetate pH 4 buffer, 150 mM NaCl, and 0.05 mM ThT to a final concentration of 0.5 mg/ml. In reducing conditions, the proteins were first treated with 1 mM DTT for 5 minutes and then diluted to 0.5 mg/ml with 50 mM Acetic acid/NaAcetate pH 4 buffer, 150 mM NaCl, 0.05 mM ThT, and 1 mM DTT. Formation of amyloid fibers was performed at 37 ^{O}C with constant shaking in black NUNC micro plates with Teflon balls with radii of 0.125" as stirrers. Fluorescent measurements of ThT were acquired with a Varioskan spectrophotometer at 444/482 nm excitation/emission wavelengths (37).

Electron microscopy. Following the completion of ThT assays, samples were collected, diluted with water to 10% v/v, and applied onto electron microscopy copper grids. The grids were stained with 2% uranyl acetate and images were collected by means of a Tecnai T12 electron microscope at 120 kV with a Gatan CCD camera.

Crystal structure determination. Crystallization trials were set up in hanging drop plates with a Mosquito micro-crystallization robot. Crystals were cryo-protected with 25% glycerol and data were collected with Rigaku HTC or at APS, Argonne, IL, 24-ID-C and 24-ID-E. Data were processed with Denzo and scaled with Scalepack or with XDS package. Phases were resolved by means of molecular replacement with Phaser, using a variable domain fragment from chain A of PDB 3MCG as the initial model. Protein models were refined with Refmac5. Data collection and refinement statistics are summarized in Table 3. Graphics were rendered with Pymol (http://www.pymol.org/). PDB accession codes and X-ray data collection and refinement statistics are in Table 1.

Native state mass spectrometry, QTOF. For native mass spectrometry, protein samples were dialyzed into 50 mM ammonium bicarbonate immediately prior to analysis under native conditions. The sample solution was infused by a syringe pump (20 µL/minute) to the electrospray source of a quadrupole time-of-flight mass spectrometer operating in positive ion mode (Agilent Technologies 6550 with dual AJS source). The instrument was operated in the 'extended mass range' mode at 2 GHz for scanning to 10,000 m/z. Quadrupole transmission was set to 1,000 m/z and various parameters were adjusted to lower the energy delivered to the analyte, including the sheath gas temperature of 125 ^{O}C and flow of 3 L/minute with a fragmentor voltage of 200 V. Mass spectra were analyzed with MassHunter Qualitative Analysis software (Agilent Technologies).

Analytical ultracentrifugation. Sedimentation equilibrium runs were performed on the native Mcg, the covalent Mcg dimer, and the V_L monomer at 4 ^{O}C in a Beckman Optima XL-A analytical ultracentrifuge using absorption optics at 280 nm. Samples were at concentrations of 0.094 mg/ml, 0.28 mg/ml, and 0.44 mg/ml. Additional concentrations of 0.38 mg/ml and 1.5 mg/ml were also run for the native Mcg. A twelve mm path length six-sector cell was used for all samples, except for the sample at 1.5 mg/ml, for which a three mm pathlength double sector cell was used. Samples were in 50 mM NaCl, 25 mM Tris pH 8. Sedimentation equilibrium profiles were measured at speeds of 14,000 rpm, 17,000 rpm, and 22,000 rpm. The data were initially fitted with a nonlinear least-squares exponential fit for a single ideal species using the Beckman Origin-based software (Version 3.01). The Beckman global analysis software (the "multifit" option of the above mentioned software) was then used to analyze multiple scans, corresponding to different input concentrations and speeds, simultaneously for various models. Partial specific volumes of 0.709 (native Mcg) and 0.707 (V_L monomer and covalent Mcg dimer) were used. These were calculated from the amino acid composition and corrected to 4 ^{O}C (38, 39).

Chemical denaturation analysis. Stability of the proteins with and without 1 mM DTT was assessed by titrating the samples with urea and monitoring

tryptophan fluorescence (40). The excitation wavelength was set at 295 nm and the emission spectra were recorded from 325 to 400 nm. The maximal emission change along the titration occurred at 355 nm. The experiment was performed with a Molecular Devices SpetraMax fluorometer and quartz cuvette with a 1 cm path length. The concentration of protein in all samples was 0.1 mg/ml.

Fluorescence data points indicating cooperative denaturation through logarithmic analysis were fit to the sigmoidal equation (Y=Bottom + (Top-Bottom)/(1+10^((LogEC50-X)*HillSlope)) using GraphPad Prism version 6 for Windows (GraphPad Software, La Jolla, California USA, www.graphpad.com). The urea concentration corresponding to denaturation of half of the protein (IC50) was used to assess stability of the proteins.

RESULTS

Design of a covalently linked V_L dimer and an induced V_L monomer

The pathologic Mcg identified by Edmundson *et al.* is a λ LC homo-dimer that was isolated from a patient and served as a model in many previous experiments, making it a particularly informative variant for studying the amyloidogenic behavior of V_Ls in our research (21, 41). The crystal structure of Mcg V_Ls appears as a homo-dimer (PDB 3MCG) with an overall conformation similar to the canonical V_H-V_L hetero-dimer of Fab (42, 43). The structural integrity of the Fab is mimicked in Mcg since the dimer interface between the V_Ls, which contains strands G-F-C-C' of the immunoglobulin Greek key fold, is preserved (26).

To verify whether pure V_L monomer or a dimer form amyloid fibers and to examine in physiological conditions findings previously found in denaturing conditions (33, 44, 45), we generated distinct, soluble, quaternary forms of the model V_L domains: the native Mcg dimer consisting of two individual V_Ls bound by non-covalent interactions, a covalently linked Mcg dimer with disulfide bonds welding the two individual V_Ls as a dimer, and an induced V_L monomer. In conditions near physiological, we examined which of the three states of Mcg can assemble into amyloid fibers.

To produce the covalent Mcg dimer, residues appropriately located for site-directed mutagenesis to cysteine were identified based on the crystal structure of the native Mcg V_L dimer (Figure 2, Table 1). The Mcg crystal structure shows proximity of residues A45 and F101 at the interface of the identical V_Ls of the dimer. Therefore, we created a covalent Mcg dimer by mutagenesis of both of these residues to cysteine (Mcg-A45C-F101C). When purified, a non-reducing SDS-PAGE shows a molecular weight of ~30 kDa as expected for a Mcg dimer. In addition, the crystal structure of Mcg-A45C-F101C clearly shows the presence of the two designed disulfide bonds between the adjacent V_Ls (Figure 2). Upon reduction of the disulfide bonds, the SDS-PAGE shows a molecular weight that matches that of the native Mcg, ~15 kDa.

To induce a soluble V_L monomer, residues at the hydrophobic interface between the two V_L domains of the native Mcg were mutated to decrease inter-domain van der Waals forces and increase solubility. These residues were identified based on the sequence and structural similarity between the native Mcg and its analogous Fab (Figure 3) and according to their effect on dimerization constants as noted by Stevens *et al.* (14, 46). The crystal structure of the most similar full-length antigen-binding fragment (Fab) (PDB 3TNM) was determined by means of amino acid sequence alignment of the native Mcg with proteins deposited in the Protein Data Bank. Comparison of the three-dimensional V_L models indicates that not only do the native Mcg and the corresponding Fab overlay closely, but apolar residues at the hydrophobic interface overlay as well. Upon the alignment of crystal structures, Mcg residues Y38, F99, and F101 were identified among several conserved key residues that form the hydrophobic interface of the V_L dimer. The residues Y38 and F101 were also noted previously by Stevens *et al.* to affect disassociation constants of V_L dimers in κ variable domains (46). The triple mutant of Y38E, F99A, and F101E removed hydrophobic contacts and introduced two additional charges per monomer, resulting in a soluble V_L monomer (Mcg-Y38E-F99A-F101E).

Quaternary states of the proteins in solution

The distribution of quaternary states of the native Mcg and two mutants in solution was verified by two methods: QTOF native mass spectrometry and analytical ultracentrifugation. Native state mass spectrometry identified molecular masses for the three proteins (Figure 4). The V_Ls of the native Mcg were found to be both dimeric and monomeric, 23.1

kDa and 11.5 kDa respectively, indicating that equilibrium exists between the two quaternary states. The molecular weight of the covalent Mcg dimer was 23.1 kDa and the V_L monomer was 11.4 kDa. These results verified the intended effects of the designed mutations on the Mcg quaternary states in solution and confirmed that the native Mcg exists in equilibrium between a V_L monomer and dimer.

Analytical ultracentrifugation data for the covalent Mcg dimer and V_L monomer gave concentration independent molecular weights consistent with a dimer and monomer respectively. The native Mcg, however, gave molecular weights that were both speed dependent (indicating molecular weight heterogeneity) and concentration dependent (indicating association). A group analysis of twelve files (four concentrations at three speeds) of Mcg found the best fit to be a monomer-dimer equilibrium with a disassociation constant (K_d) of 0.2±0.01 mM (Figure 5).

Stability of the native Mcg, covalent Mcg dimer, and V_L monomer

The stability of the three proteins was examined by titrating with urea and monitoring fluorescence. Logarithmic analysis of the fluorescence data indicates that all the proteins proceed through cooperative unfolding at lower concentrations of urea followed by non-cooperative unfolding at higher concentrations (Figure 6). The urea concentration corresponding to denaturation of half of the protein (IC50) is highest for the covalent Mcg dimer, followed by the native Mcg, and then the induced V_L monomer, indicating that the monomer is the least stable. In the presence of DTT, the IC50 of each protein decreases, indicating that the proteins are less stable under reducing conditions. This shows that in addition to reducing the disulfide bond that links the Mcg covalent dimer, the intra-chain disulfide bond between residues 22 and 90 of each variable domain is also reduced despite its location in the core of the domain.

Effect of quaternary structure on amyloid fiber formation

The ability of the native Mcg, covalent Mcg dimer, and V_L monomer to form fibers was monitored by an increase in thioflavin T (ThT) fluorescence and by analysis of transmission electron micrographs for the presence of fibers after the ThT assays. When amyloid fibers formed, catalyzed by an acidic environment and temperature of 37 ^{O}C, the ThT assay indicated an increase in fluorescence and electron micrographs confirmed the presence of fibers (Figure 7). Although the increase in ThT fluorescence correlates with formation of amyloid fibers, the experiments indicated a variation in intensity and lag times. This variation probably occurs due to differences in nucleation heterogeneity, and therefore, we qualitatively examined ThT fluorescence.

The native Mcg and V_L monomer readily formed fibers in oxidizing and reducing conditions. However, the covalent Mcg dimer did not form fibers in oxidizing conditions. Upon reduction of the disulfides with dithiothreitol (DTT), the dimer formed amyloid fibers resembling those of the native Mcg and the V_L monomer (Figure 7). Confinement of the covalent Mcg to its canonical dimer conformation abolishes its amyloidogenic property, whereas both the native Mcg, existing in equilibrium between dimers and monomers, and the pure V_L monomer are able to form amyloid fibers.

Structures of designed V_L domains

The structures of the covalent Mcg dimer and of the V_L monomer were determined by means of X-ray crystallography, and their conformations remained similar to the native Mcg dimer. The covalent Mcg dimer showed two disulfide bonds between residues 45 and 101 of adjacent V_Ls, which remained in the same orientation relative to each other as in the native Mcg (Figure 2). Despite mutations that induced the V_L monomer in aqueous solution, it also crystallized in a conformation similar to the native Mcg. However, its crystal structure displayed a phenomenon new to immunoglobulin variable domains: G-strands were swapped among crystallographic dimers, yet the overall immunoglobulin fold remained preserved (Figure 2, Figure 8).

DISCUSSION

Identification of pathways of assembly from globular proteins into amyloid fibers is essential for understanding amyloid associated diseases. Until now, it has not been clear whether a monomer, dimer, or some other multimeric state of immunoglobulin LCs is the precursor to assembly of amyloid fibers. Three models for the assembly pathway are presented in Figure 1. Our experiments

show that the covalent Mcg dimer is unable to form fibers due to disulfide bonds that bind pairs of V_Ls and prevent transformation into monomers or a non-canonical dimer. Yet upon reduction of the disulfide bonds, the dimer regains the ability to form fibers as readily as the native Mcg and the V_L monomer. Therefore, this demonstrates that a canonical dimer of Mcg is protective against amyloid formation while a V_L monomer is prone (29, 32, 33). In other words, the *monomer model* is the most consistent with our data.

Differences in amino acid sequence between various V_Ls are likely to affect the stability of the dimer and the monomer, which can account for why some V_Ls are more likely than others to disassociate into monomers and form amyloid fibers (31). In the vast number of possible immunoglobulin variants, each differs from others by only several residues within each of the λ or κ families, and these differences may affect V_L dimer-monomer disassociation constants (31, 47). Unstable dimers are more likely to transform between alternative conformations, which protect against amyloid formation to various degrees according to their thermodynamic stability. This transformation is likely to occur by disassociation of the dimer into amyloid-prone monomers, which then can partially unfold and expose a segment that forms the steric zipper spine of the fiber.

Our results are consistent with previous findings that destabilization of V_L dimers and unfolding of V_L monomers are consecutive yet separate processes (29, 32, 44). Experiments with κ1 O18/O8, AL-09, and SMA V_Ls show that disassociation of dimers and unfolding of monomers promotes amyloid assembly, and in denaturing conditions, mutations that destabilize V_L monomers induce formation of amyloid fibers. Similarly, our chemical denaturation experiments suggest the monomer is the quaternary state that is more prone to unfold, while dimer is protective. Since formation of amyloid fibers depends on a soluble yet unstable monomer, both a stable dimer and a stable monomer would reduce propensity to form amyloid fibers. However, reduction of amyloid formation occurs by two different mechanisms: stable dimers reduce the concentration of soluble monomers, and stable monomers are less likely to unfold and form amyloid fibers. This is similar to pathways by which other amyloidogenic proteins form amyloid fibers; for example, transthyretin disassociates into monomers and lysozyme must transition through a partially unfolded state to form amyloid fibers (33, 48–51).

Many variants of V_Ls have been found in patients with systemic light-chain amyloidosis, and these variants only differ by a few amino acids (14). Although our experiments were performed with only one variant and a single conformation was imposed on the covalently linked dimer, the fact that a pure, soluble monomer forms amyloid fibers indicates that the monomeric state of the V_L is the precursor to amyloid fibers. Due to the similarity in sequence, structure and biochemical behavior of all light-chain variable domains, we infer that our finding for Mcg may well be applied to other light-chains and explain the differences in propensity to form amyloid: a stable dimer is less likely to disassociate into monomers and therefore, less likely to form amyloid.

While V_L dimers must disassociate into monomers in order to form fibers, structures of full-length LCs indicate that this conformational change can occur regardless of a connection to a constant domain (42). In full-length LCs, V_Ls are connected to C_Ls through a joining (J) segment, which provides the domains with the flexibility to acquire conformations independent of each other. Such an architecture of the protein does not impose a particular orientation on V_Ls relative to connected C_Ls, as affirmed by the existence of several crystal structures (42, 52, 53). Therefore, a dimer of V_Ls maintains the ability to disassociate into amyloid prone V_L monomers, while still covalently attached to C_Ls. This flexibility explains why amyloid fibers of immunoglobulin can consist of full-length LCs, just their V_Ls, or both.

The crystal structure of our induced V_L monomer domain displays an overall conformation similar to the dimer formed by the native Mcg, yet it also shows a phenomenon new to immunoglobulin variable domains where strands are swapped among molecules of the asymmetric unit (Figure 8). Swapping between segments of protein domains is known to be concentration dependent and was previously observed in other proteins, such as diphtheria toxin, human prion protein, cystatin C, β_2-microglobulin, T7 endonuclease, and helicase (54–58). In closed swapping, protein segments are exchanged between a discrete number of molecules, whereas in runaway swapping, the exchange occurs between an unlimited number, resulting in a

continuous polymer (55). In the crystal structure of V_L monomers, closed swapping of G-strands occurs between four V_L dimers, with both domains of each dimer being swapped. In addition, canonical dimer interfaces link the pairs of swapped dimers in a non-crystallographic octamer with distorted 422 symmetry. Although swapping between domains of V_L monomers in solution may play a role in linking molecules in a fiber, the closed swapping in the crystal structure does not resemble a steric zipper spine that would account for the cross-β X-ray diffraction pattern. Though the domain-swapping cannot be dismissed as unrelated, its existence in immunoglobulins and relation to their amyloidogenic property is yet unclear.

A practical insight emerging from both past studies and the present study is the possible binding of molecules to the hydrophobic cavity between the variable domains in order to stabilize the dimer (59). Structures of V_L dimers show that a variety of ligands bind to the hydrophobic cavity between the domains (24). Crystallographic models of Mcg indicate that the protein maintains its structure as a canonical dimer while bound to a ligand. The binding of molecules yields an energetically favorable complex, and the equilibrium between different conformations might shift towards a stable dimer rather than an amyloid prone monomer. Such complexes may therefore inhibit the transition of LC or V_L dimers to amyloid fibers, resulting in alleviated symptoms in systemic amyloidosis patients.

REFERENCES

1. Glenner, G. G., Harbaugh, J., Ohma, J. I., Harada, M., and Cuatrecasas, P. (1970) An amyloid protein: the amino-terminal variable fragment of an immunoglobulin light chain. *Biochem. Biophys. Res. Commun.* **41**, 1287–9
2. Glenner, G. G. (1973) Immunoglobulin and amyloid fibril proteins. *Br J Haematol* **24**, 533–7
3. Glenner, G. G., Ein, D., Eanes, E. D., Bladen, H. A., Terry, W., and Page, D. L. (1971) Creation of "amyloid" fibrils from Bence Jones proteins in vitro. *Science* **174**, 712–4
4. Dahlin, D. C., and Dockerty, M. B. (1950) Amyloid and myeloma. *Am. J. Pathol.* **26**, 581–593
5. Jones, H. B. (1848) On a New Substance Occurring in the Urine of a Patient with Mollities Ossium. *Philos. Trans. R. Soc. Lond.* **138**, 55–62
6. Marchalonis, J. J., and Schluter, S. F. (1989) Evolution of variable and constant domains and joining segments of rearranging immunoglobulins. *Faseb J* **3**, 2469–79
7. Sipe, J. D., and Cohen, A. S. (2000) Review: history of the amyloid fibril. *J Struct Biol* **130**, 88–98
8. Geddes, A. J., Parker, K. D., Atkins, E. D., and Beighton, E. (1968) "Cross-beta" conformation in proteins. *J Mol Biol* **32**, 343–58
9. Hobbs, J. R. (1973) An ABC of amyloid. *Proc R Soc Med* **66**, 705–10
10. Eisenberg, D., and Jucker, M. (2012) The amyloid state of proteins in human diseases. *Cell* **148**, 1188–203
11. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. Ø., Riekel, C., Grothe, R., and Eisenberg, D. (2005) Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **435**, 773–8
12. Buxbaum, J. (1992) Mechanisms of disease: monoclonal immunoglobulin deposition. Amyloidosis, light chain deposition disease, and light and heavy chain deposition disease. *Hematol Oncol Clin North Am* **6**, 323–46
13. Falk, R. H., Comenzo, R. L., and Skinner, M. (1997) The systemic amyloidoses. *N Engl J Med* **337**, 898–909
14. Bodi, K., Prokaeva, T., Spencer, B., Eberhard, M., Connors, L. H., and Seldin, D. C. (2009) AL-Base: a visual platform analysis tool for the study of amyloidogenic immunoglobulin light chain sequences. *Amyloid* **16**, 1–8
15. Olsen, K. E., Sletten, K., and Westermark, P. (1998) Fragments of the constant region of immunoglobulin light chains are constituents of AL-amyloid proteins. *Biochem. Biophys. Res. Commun.* **251**, 642–7

16. Lavatelli, F., Perlman, D. H., Spencer, B., Prokaeva, T., McComb, M. E., Théberge, R., Connors, L. H., Bellotti, V., Seldin, D. C., Merlini, G., Skinner, M., and Costello, C. E. (2008) Amyloidogenic and associated proteins in systemic amyloidosis proteome of adipose tissue. *Mol. Cell. Proteomics MCP* **7**, 1570–1583
17. Vrana, J. A., Gamez, J. D., Madden, B. J., Theis, J. D., Bergen, H. R., 3rd, and Dogan, A. (2009) Classification of amyloidosis by laser microdissection and mass spectrometry-based proteomic analysis in clinical biopsy specimens. *Blood* **114**, 4957–4959
18. Sakano, H., Hüppi, K., Heinrich, G., and Tonegawa, S. (1979) Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* **280**, 288–294
19. Buxbaum, J. (1986) 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**, 798–806
20. Colman, P. M., Schramm, H. J., and Guss, J. M. (1977) Crystal and molecular structure of the dimer of variable domains of the Bence-Jones protein ROY. *J Mol Biol* **116**, 73–9
21. Edmundson, A. B., Wood, M. K., Schiffer, M., Hardman, K. D., and Ainsworth, C. F. (1969) A crystallographic investigation of the Mcg myeloma protein. ANL-7635. *ANL*, 283–285
22. Firca, J. R., Ely, K. R., Kremser, P., Westholm, F. A., Dorrington, K. J., and Edmundson, A. B. (1978) Interconversion of conformational isomers of light chains in the Mcg immunoglobulins. *Biochemistry (Mosc.)* **17**, 148–158
23. Edmundson, A. B., Harris, D. L., Fan, Z. C., Guddat, L. W., Schley, B. T., Hanson, B. L., Tribbick, G., and Geysen, H. M. (1993) Principles and pitfalls in designing site-directed peptide ligands. *Proteins* **16**, 246–267
24. Edmundson, A. B., Ely, K. R., and Herron, J. N. (1984) A search for site-filling ligands in the Mcg Bence-Jones dimer: crystal binding studies of fluorescent compounds. *Mol. Immunol.* **21**, 561–576
25. Stevens, F. J., Myatt, E. A., Chang, C., Westholm, F. A., Eulitz, M., Weiss, D. T., Murphy, C., Solomon, A., and Schiffer, M. (1995) A Molecular Model for Self-Assembly of Amyloid Fibrils: Immunoglobulin Light Chains. *Biochemistry (Mosc.)* **34**, 10697–10702
26. Richardson, J. S. (1977) beta-Sheet topology and the relatedness of proteins. *Nature* **268**, 495–500
27. Epp, O., Lattman, E. E., Schiffer, M., Huber, R., and Palm, W. (1975) The molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI refined at 2.0-A resolution. *Biochemistry (Mosc.)* **14**, 4943–4952
28. Hernandez-Santoyo, A., del Pozo Yauner, L., Fuentes-Silva, D., Ortiz, E., Rudino-Pinera, E., Sanchez-Lopez, R., Horjales, E., Becerril, B., and Rodriguez-Romero, A. (2010) A single mutation at the sheet switch region results in conformational changes favoring lambda6 light-chain fibrillogenesis. *J Mol Biol* **396**, 280–92
29. Peterson, F. C., Baden, E. M., Owen, B. A. L., Volkman, B. F., and Ramirez-Alvarado, M. (2010) A single mutation promotes amyloidogenicity through a highly promiscuous dimer interface. *Struct. Lond. Engl. 1993* **18**, 563–570
30. Edmundson, A. B., Ely, K. R., Abola, E. E., Schiffer, M., Panagiotopoulos, N., and Deutsch, H. F. (1976) Conformational isomerism, rotational allomerism, and divergent evolution in immunoglobulin light chains. *Fed. Proc.* **35**, 2119–2123
31. Hurle, M. R., Helms, L. R., Li, L., Chan, W., and Wetzel, R. (1994) A role for destabilizing amino acid replacements in light-chain amyloidosis. *Proc Natl Acad Sci U A* **91**, 5446–50
32. Baden, E. M., Owen, B. A., Peterson, F. C., Volkman, B. F., Ramirez-Alvarado, M., and Thompson, J. R. (2008) Altered dimer interface decreases stability in an amyloidogenic protein. *J Biol Chem* **283**, 15853–60
33. Qin, Z., Hu, D., Zhu, M., and Fink, A. L. (2007) Structural Characterization of the Partially Folded Intermediates of an Immunoglobulin Light Chain Leading to Amyloid Fibrillation and Amorphous Aggregation. *Biochemistry (Mosc.)* **46**, 3521–3531
34. Bernier, G. M., and Putnam, F. W. (1963) Monomer-dimer forms of Bence Jones proteins. *Nature* **200**, 223–225

35. Kishida, F., Azuma, T., and Hamaguchi, K. (1975) A type kappa Bence Jones protein containing a cysteinyl residue in the variable region. *J. Biochem. (Tokyo)* **77**, 481–491
36. Blancas-Mejía, L. M., and Ramirez-Alvarado, M. (2013) Systemic amyloidoses. *Annu. Rev. Biochem.* **82**, 745–774
37. Wall, J., Murphy, C. L., and Solomon, A. (1999) In vitro immunoglobulin light chain fibrillogenesis. *Methods Enzymol.* **309**, 204–217
38. Cohn, E. J., and Edsall, J. T. (1943) Density and apparent specific volume of proteins. In: Proteins, Amino Acids and Peptides as Ions and Dipolar Ions. *Reinhold Publishing Corporation, New York.*, 370–381
39. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-Aided Interpretation of Analytical Sedimentation Data for Proteins. In: Analytical Ultracentrifugation in Biochemistry and Polymer Science. *Cambridge, Great Britain, The Royal Society of Chemistry*, 90–125
40. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280
41. Hanson, B. L., Bunick, G. J., Harp, J. M., and Edmundson, A. B. (2002) Mcg in 2030: new techniques for atomic position determination of immune complexes. *J. Mol. Recognit. JMR* **15**, 297–305
42. Ely, K. R., Herron, J. N., Harker, M., and Edmundson, A. B. (1989) Three-dimensional structure of a light chain dimer crystallized in water. Conformational flexibility of a molecule in two crystal forms. *J. Mol. Biol.* **210**, 601–615
43. Bourne, P. C., Ramsland, P. A., Shan, L., Fan, Z. C., DeWitt, C. R., Shultz, B. B., Terzyan, S. S., Moomaw, C. R., Slaughter, C. A., Guddat, L. W., and Edmundson, A. B. (2002) Three-dimensional structure of an immunoglobulin light-chain dimer with amyloidogenic properties. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 815–23
44. Khurana, R., Gillespie, J. R., Talapatra, A., Minert, L. J., Ionescu-Zanetti, C., Millett, I., and Fink, A. L. (2001) Partially Folded Intermediates as Critical Precursors of Light Chain Amyloid Fibrils and Amorphous Aggregates. *Biochemistry (Mosc.)* **40**, 3525–3535
45. Baden, E. M., Sikkink, L. A., and Ramirez-Alvarado, M. (2009) Light chain amyloidosis - current findings and future prospects. *Curr Protein Pept Sci* **10**, 500–8
46. Stevens, F. J., Westholm, F. A., Solomon, A., and Schiffer, M. (1980) Self-association of human immunoglobulin kappa I light chains: role of the third hypervariable region. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 1144–1148
47. Lefranc, M. (1999) IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res.* **27**, 209–212
48. Colon, W., and Kelly, J. W. (1992) Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry (Mosc.)* **31**, 8654–8660
49. Lai, Z., Colón, W., and Kelly, J. W. (1996) The acid-mediated denaturation pathway of transthyretin yields a conformational intermediate that can self-assemble into amyloid. *Biochemistry (Mosc.)* **35**, 6470–6482
50. Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C. F., and Pepys, M. B. (1997) Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **385**, 787–793
51. Fink, A. L. (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **3**, R9–R23
52. Terzyan, S. S., Bourne, C. R., Ramsland, P. A., Bourne, P. C., and Edmundson, A. B. (2003) Comparison of the three-dimensional structures of a human Bence-Jones dimer crystallized on Earth and aboard US Space Shuttle Mission STS-95. *J Mol Recognit* **16**, 83–90
53. Makino, D. L., Henschen-Edman, A. H., Larson, S. B., and McPherson, A. (2007) Bence Jones KWR protein structures determined by X-ray crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 780–792
54. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Domain swapping: entangling alliances between proteins. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 3127–3131
55. Bennett, M. J., Sawaya, M. R., and Eisenberg, D. (2006) Deposition diseases and 3D domain swapping. *Structure* **14**, 811–24

56. Liu, C., Sawaya, M. R., and Eisenberg, D. (2011) β_2-microglobulin forms three-dimensional domain-swapped amyloid fibrils with disulfide linkages. *Nat. Struct. Mol. Biol.* **18**, 49–55
57. Guo, Z., and Eisenberg, D. (2006) Runaway domain swapping in amyloid-like fibrils of T7 endonuclease I. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 8042–8047
58. Sawaya, M. R., Guo, S., Tabor, S., Richardson, C. C., and Ellenberger, T. (1999) Crystal Structure of the Helicase Domain from the Replicative Helicase-Primase of Bacteriophage T7. *Cell* **99**, 167–177
59. Miroy, G. J., Lai, Z., Lashuel, H. A., Peterson, S. A., Strang, C., and Kelly, J. W. (1996) Inhibiting transthyretin amyloid fibril formation via protein stabilization. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15051–15056

Acknowledgements-We thank the Amyloidosis Foundation (http://www.amyloidosis.org/) for a fellowship to B.B., NSF MCB-0958111 and NIH AG-029430 for supporting this research, the UCLA-DOE crystallization facility for setting up crystallization trials, the beamline scientists at 24-ID-C NE-CAT, APS, Argonne (IL), USA for facilitating X-ray data collection experiments, and Dr. L. Salwinski for his valuable comments on the manuscript.

FOOTNOTES

B.B. and S.E. contributed equally to the work described here. B.B. conceived the hypothesis and developed the experimental plan and methods. B.B. and S.E. performed the experiments. M.L. identified amino acids that induce the monomeric conformation of the protein. C.R. and J.W. performed QTOF experiments. M.P. performed analytical ultracentrifugation experiments. M.S. and D.C. assisted in crystallographic data collection and model building. D.S.E. supervised the research. B.B., S.E., and D.S.E. wrote the paper, incorporating suggestions from all authors.

Three-dimensional protein models and structure factors are available in the Protein Data Bank with accession codes as indicated in Table 1. There are no competing financial interests involved in this work. Correspondence and requests for materials should be addressed to D.S.E. (david@mbi.ucla.edu).

FIGURE LEGENDS

Table 1. X-ray data collection and refinement statistics.

Figure 1. Hypothetical pathways for the conversion of immunoglobulin light-chain variable domain (V_L) dimers to amyloid fibers. The left-hand pathway depicts the *canonical dimer* observed in several crystal structures, which then stacks into an amyloid fiber. The right-hand pathway depicts one of several *non-canonical dimers*, also seen in crystal structures. In these dimers, the quaternary structure of the two light-chain variable domains differs from the canonical dimer, but each domain retains its tertiary conformation, and the dimer stacks into an amyloid fiber. The central pathway depicts the disassociation of the dimer to monomers, which then partially unfold and form an amyloid fiber. In this paper, we assess which of these three models best describes the quaternary state most prone to forming amyloid fibers.

Figure 2. Crystal structures of the native Mcg (PDB 4UNU), the covalent Mcg dimer (PDB 4UNV), and the induced Mcg V_L monomer (PDB 4UNT). **A.** Native Mcg. **B.** Mcg-A45C-F101C covalent dimer. Two disulfide bonds (shown in yellow) between residues 45 and 101 covalently link adjacent V_L domains of the dimer. **C.** The induced monomer Mcg-Y38E-F99A-F101E with swapped G-strands. **D.** SDS-PAGE of the three samples in oxidizing (-) and reducing (+) conditions, where the reductant is DTT. Markers on both sides of the gel indicate molecular weight in kDa. Notice that in reducing conditions, the molecular weight of the "Dimer" corresponds to the weight of the original Mcg and induced monomer.

Figure 3. Similarity of the Mcg V_L homo-dimer to the V_L-V_H variable domain hetero-dimer of a physiological Fab (PDB 3TNM). **A.** The crystal structure of the Fab (PDB 3TNM) reveals a hydrophobic cavity in between the variable domains of its light and heavy chains. The cavity between the V_L-V_H is lined with side chains of hydrophobic residues. V_L is shown in blue and V_H in green. V_L denotes a variable domain of a light chain. V_H – variable domain of a heavy chain. C_L – constant domain of a light chain. C_H – constant domain of a heavy chain. **B.** Overlay of Fab with Mcg (yellow). Despite differences in residues, the overall structures have identical orientations of domains. **C.** Crystal structure of the Mcg V_L-V_L dimer. **D.** Sequence alignment of 3TNM and Mcg. Yellow indicates residues that form the hydrophobic interface between variable domains. Notice the similarity of the interface-forming residues of the Mcg V_L-V_L homo-dimer and the V_L-V_H hetero-dimer.

Figure 4. Deconvoluted protein masses from native state mass spectrometry measurements. **A.** Detected masses for Mcg. The two major peaks show that the protein is found as both a monomer (11557.6 Da) and a dimer (23113.19 Da) in solution. **B.** Detected mass for the covalent Mcg dimer (Mcg-A45C-F101C, 23150.53 Da). **C.** Detected mass for the induced monomer (Mcg-Y38E-F99A-F101E, 11427.35 Da). Thus, mass spectrometry shows that Mcg is in equilibrium between monomers and dimers.

Figure 5. Analytical ultracentrifugation data for the native Mcg, covalent dimer, and induced V_L monomer. **A.** Fit of three models of the distribution of quaternary states in solution to the native Mcg ultracentrifugation data. The blue curve shows the fit of a pure dimer model, the red curve shows the fit of a pure monomer model, and the green curve shows the fit of a dimer-monomer equilibrium model, which fits best to the data. The fit is shown for the native Mcg at 0.4 mg/ml and a speed of 22,000 rpm. **B.** Analytical ultracentrifugation data for the three samples. The top row shows data for the native Mcg, the middle row for the covalent dimer, and the bottom row for the induced V_L monomer. The left column shows data for proteins at 0.4 mg/ml, the middle column for 0.3 mg/ml, and the right for 0.1 mg/ml. O.D. is a measurement of absorption at 280 nm along the radius of the quartz cell from the center of the analytical ultracentrifugation rotor. Measurements were taken only after samples reached equilibrium and the distribution of protein along the cells did not change. Three measurements were made for each analyte at speeds of 22,000, 17,000 and 14,000 rpm. Models of a dimer-monomer equilibrium, pure dimer,

and pure monomer were fit to data of the native Mcg, covalent dimer, and induced monomer. The curves of the fit are not clearly visible because of the almost perfect alignment with the experimental data. Residuals show the corresponding error of the fit model to the data.

Figure 6. Chemical denaturation of the native Mcg, covalent dimer, and induced V_L monomer. Cooperative denaturation was detected for the three protein constructs at low concentrations of urea, followed by non-cooperative denaturation. The concentrations of urea corresponding to denaturation of half of the protein indicate that the covalent Mcg dimer is the most stable quaternary state and the monomer is the least stable.

Figure 7. Thioflavin T fiber formation assays and electron micrographs of the native Mcg, covalent Mcg dimer, and induced V_L monomer. **A and B.** Averaged fluorescence readings for each protein based on three repeated Thioflavin T assays in oxidized and reduced conditions. Error bars are not shown for clarity. **C and D.** Electron micrographs. The only sample that does not form amyloid fibers within five days is the covalent Mcg dimer. However, when the covalent cysteine linkages are reduced by DTT, it forms amyloid fibers as readily as the native Mcg and the induced V_L monomer. **E.** Raw ThT fluorescence data for the three replicated experiments, which are shown averaged in columns A and B.

Figure 8. The asymmetric unit of domain-swapped V_L monomers. Although the Mcg-Y38E-F99A-F101E mutant is monomeric in solution, it resembles the structure of a swapped dimer in the crystal. Note that G-strands of every peptide chain are swapped into a domain in an adjacent dimer.

Table 1

	Mcg PDB 4UNU	Mcg-A45C-F101C (covalent dimer) PDB 4UNV	Mcg-Y38E-F99A-F101E (monomer) PDB 4UNT
Data collection			
Space group	P 1 2_1 1	C 2 2 2_1	C 1 2 1
Cell dimensions			
a, *b*, *c* (Å)	41.3, 34.1, 63.3	37.4, 60.2, 79.4	103.1, 90.3, 99.2
α, β, γ (°)	90.0, 104.9, 90.0	90, 90, 90	90, 118.8, 90
Resolution (Å)	40-0.95(0.97)*	30-1.6(1.64)	65-2.7(2.77)
R_{sym} or R_{merge}	0.07(0.7)	0.06(0.4)	0.09(0.3)
$I/\sigma I$	21(1.9)	39(4.2)	13(4.3)
Completeness (%)	88(34)	98(86)	98(99)
Redundancy	12(3.2)	7(3.6)	4(3.8)
Refinement			
Resolution (Å)	40-0.95	30-1.6	64-2.7
No. reflections	89567	11350	20681
$R_{work/}$ R_{free}	0.11/0.13	0.20/0.26	0.25/0.29
No. atoms			
Protein	1682	794	6255
Ligand/ion	41	10	65
Water	373	113	56
B-factors			
Protein	9.7	24.1	31.3
Ligand/ion	21.9	38.8	54.1
Water	28.8	35.4	18.9
R.m.s deviations			
Bond lengths (Å)	0.026	0.018	0.010
Bond angles (°)	2.291	1.962	1.448
Crystallization conditions			
	2 M NaCl, 2 M $(NH4)_2SO_4$	0.1 M CH_3COONa pH 4.6, 2 M $(NH4)_2SO_4$	0.5 M LiCl, 1.65 M $(NH4)_2SO_4$

Each structure was derived from a single crystal.
*Highest resolution shell is shown in parenthesis.

Figure 1

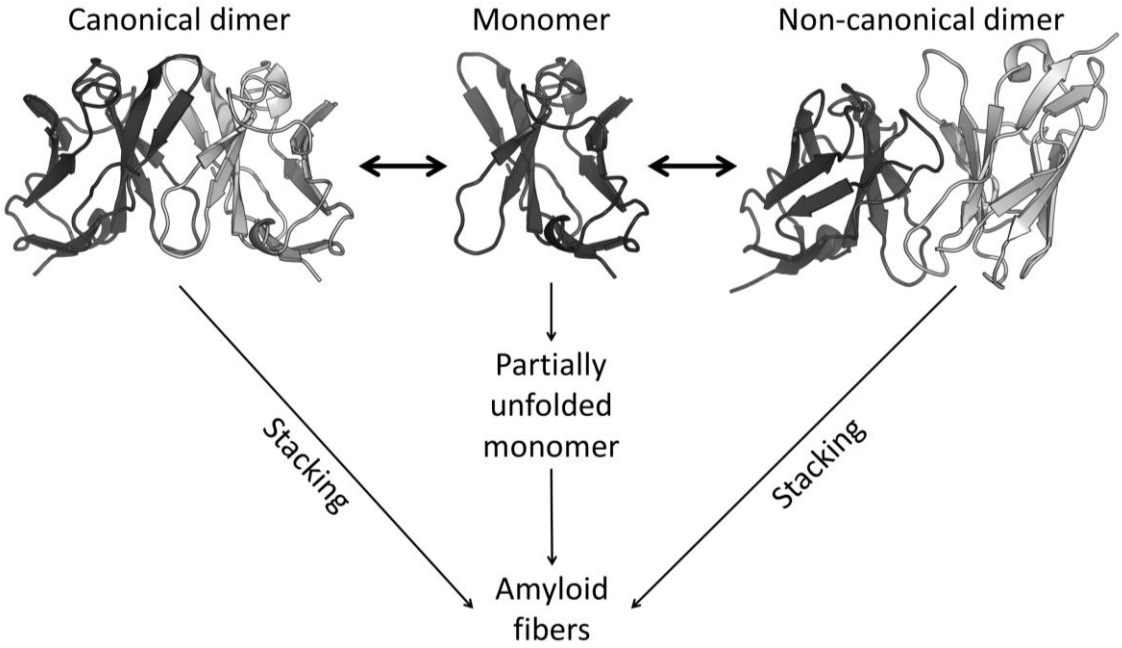
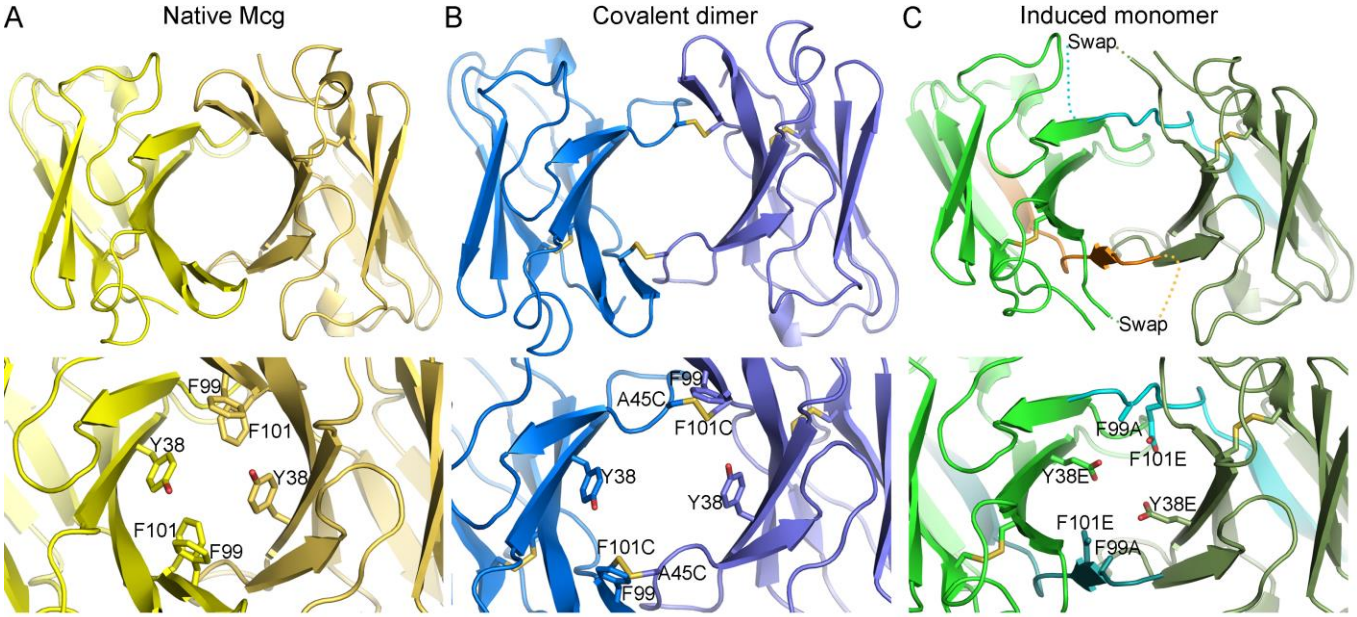


Figure 2



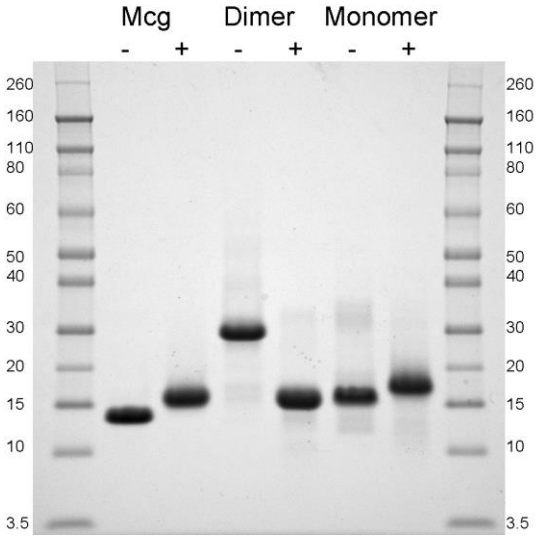
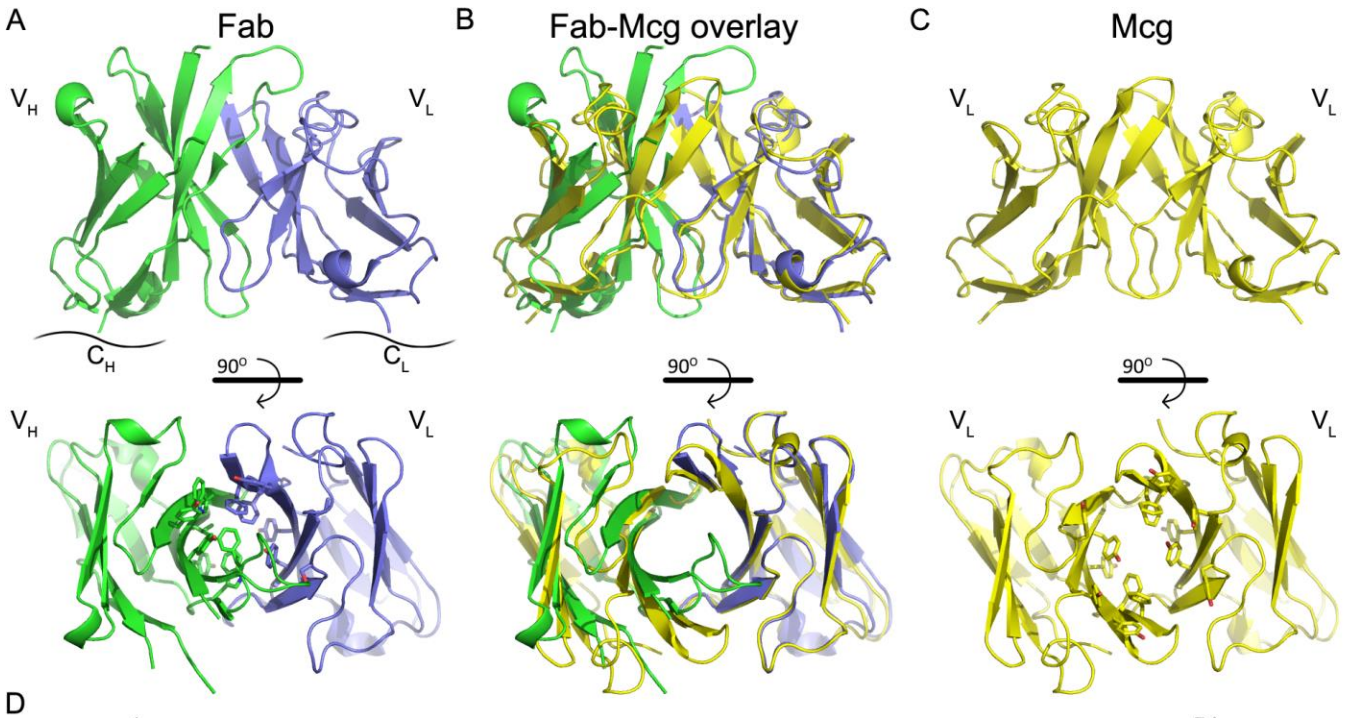


Figure 3



D

```
                1                                                   54
Mcg             QSALTQPPSASGSLGQSVTISCTGTS-SDVGGYNYVSWYQQHAGKAPKVIIYEVN------
3TNM V_L        QSVLTQPPSASGSPGQSVTISCTGTS-SDVGGYNYVSWYQHHPGKAPKLIISEVN------
3TNM V_H        QVQLQESGPGLVKPSQTLSLSCTVSGGSSSSGAHYWSWIRQYPGKGLEWIGYIHYSGNTYY

               55                                                  110
Mcg             KRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGS----DNF-VFGTGTKVTVL
3TNM V_L        NRPSGVPDRFSGSKSGNTASLTVSGLQAEDEAEYYCSSYTDI----HNF-VFGGGTKLTVL
3TNM V_H        NPSLKSRITISQHTSENQFSLKLNSVTVADTAVYYCARGTRLRTLRNAFDIWGQGTMVTVS
```

Figure 4

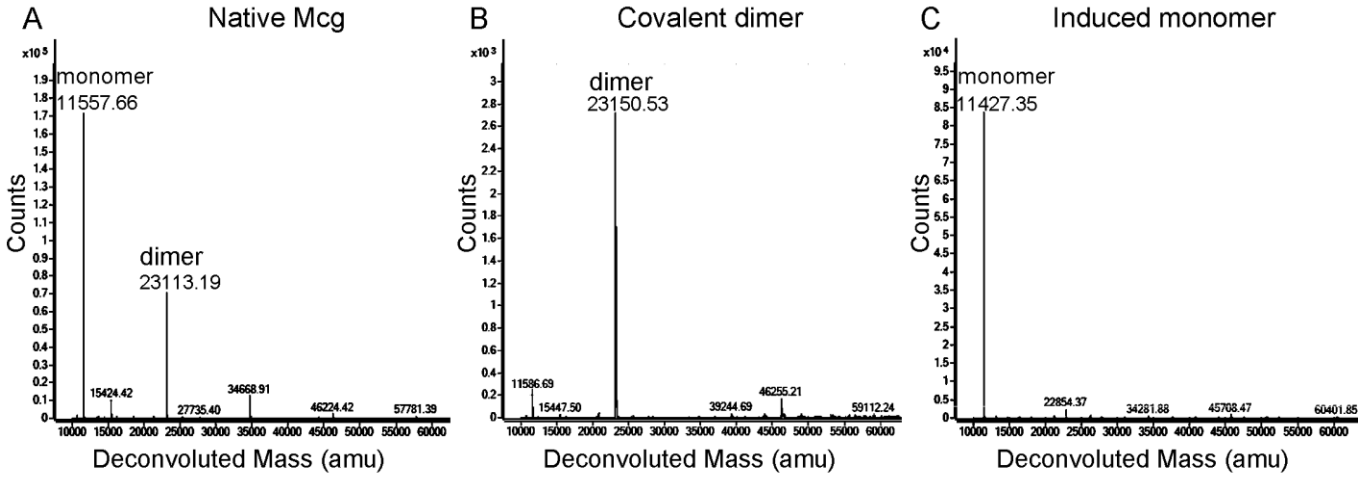
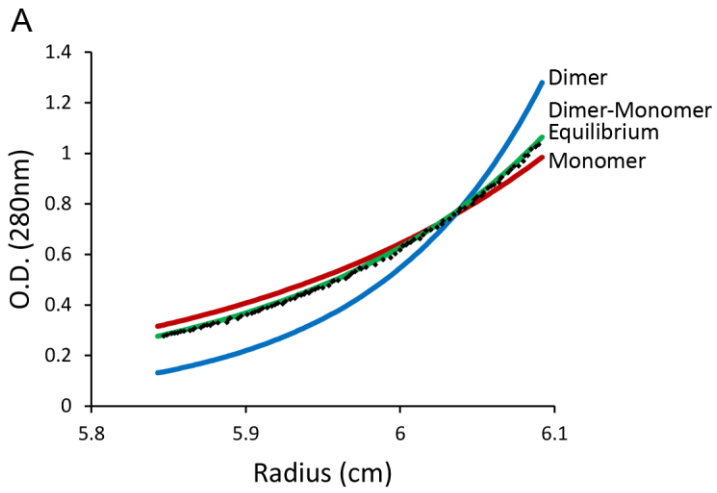


Figure 5



B

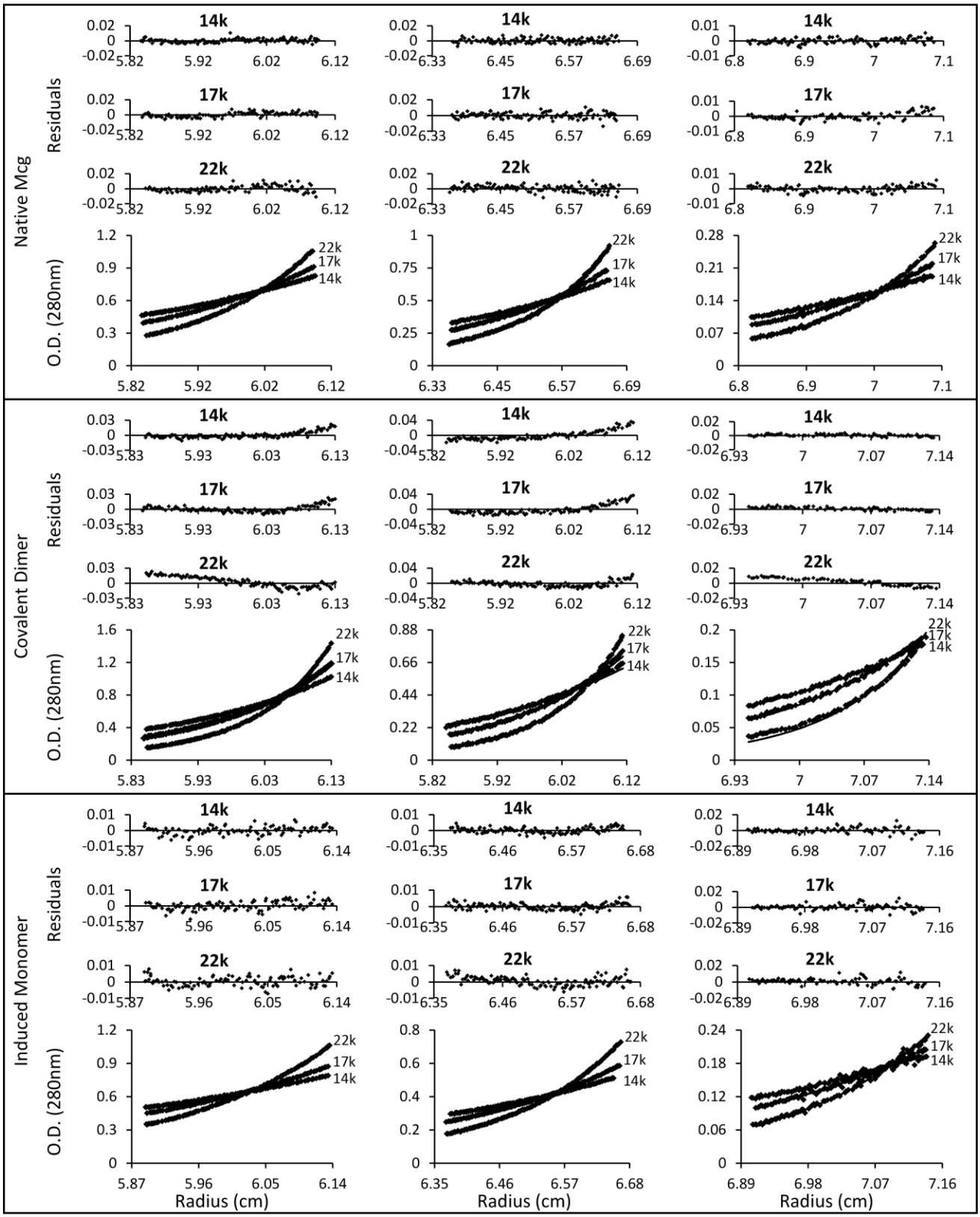
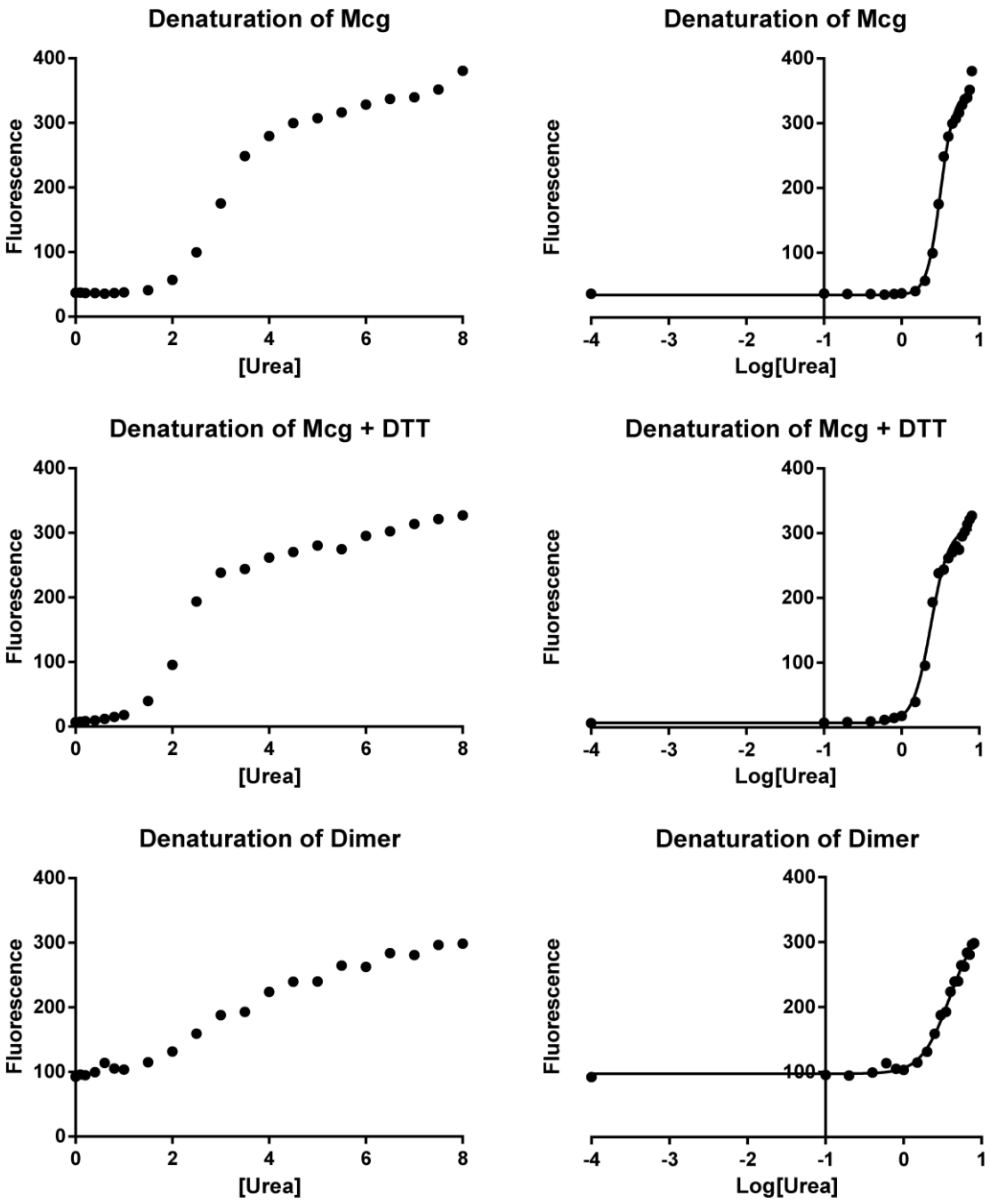
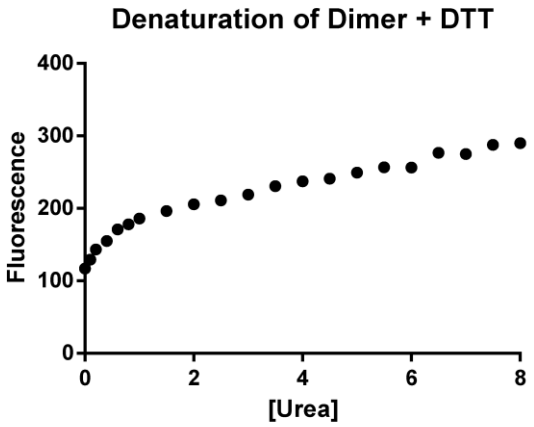
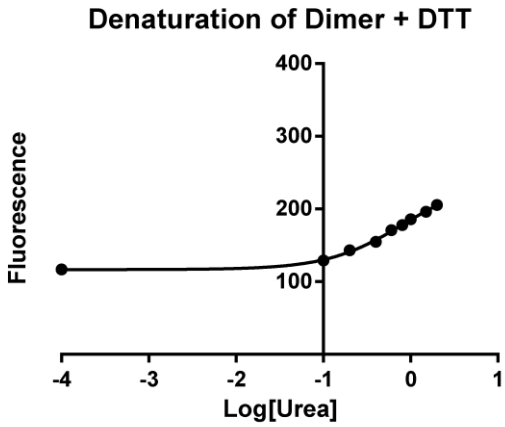
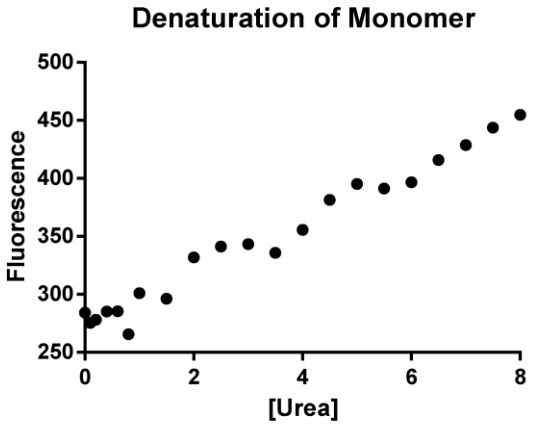


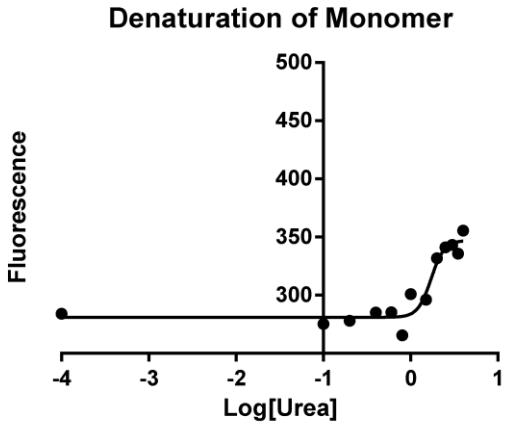
Figure 6

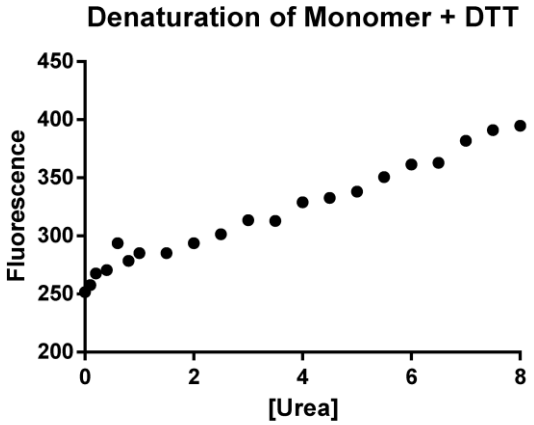


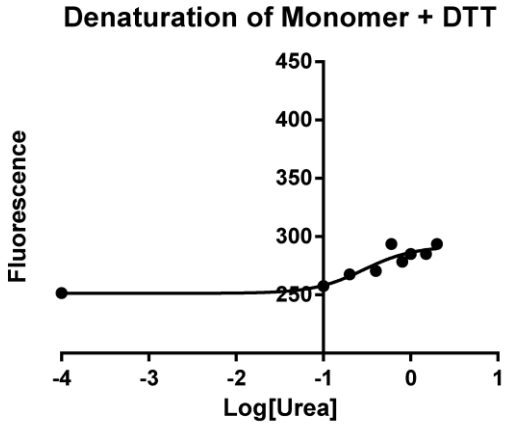












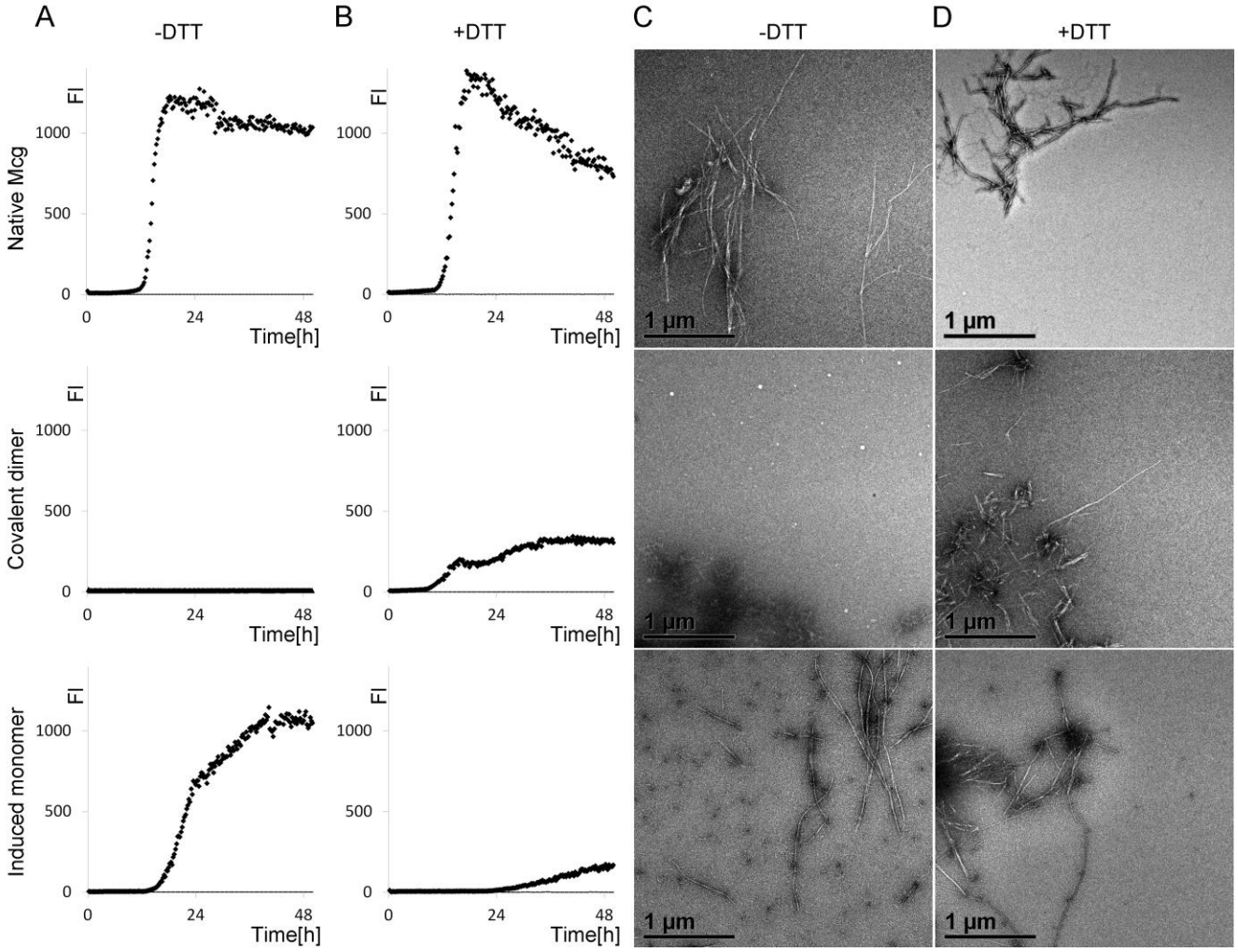
Protein	Fit*, r^2	IC50: Half of the protein is denatured (95% confidence) [Urea, M]	Data points used for IC50 calculation [Urea, M]**
Mcg	Sigmoidal, 0.99	3.2 (3.0-3.3)	0-8
Mcg + 1 mM DTT	Sigmoidal, 0.99	2.4 (2.2-2.6)	0-8
Dimer	Sigmoidal, 0.99	3.9 (3.4-4.5)	0-8
Dimer + 1 mM DTT	Sigmoidal, 0.97	0.8 (0.4-1.4)	0-2
Monomer	Sigmoidal, 0.92	1.7 (1.6-2.1)	0-4

Monomer + 1 mM DTT	Sigmoidal, 0.86	0.3 (0.1-1.1)	0-2

* Sigmoidal fit to Y=Bottom + (Top-Bottom)/(1+10^((LogEC50-X)*HillSlope))

** Data points at 0 M Urea were assigned a value of 0.0001 M Urea for inclusion in logarithmic analysis.

Figure 7



E

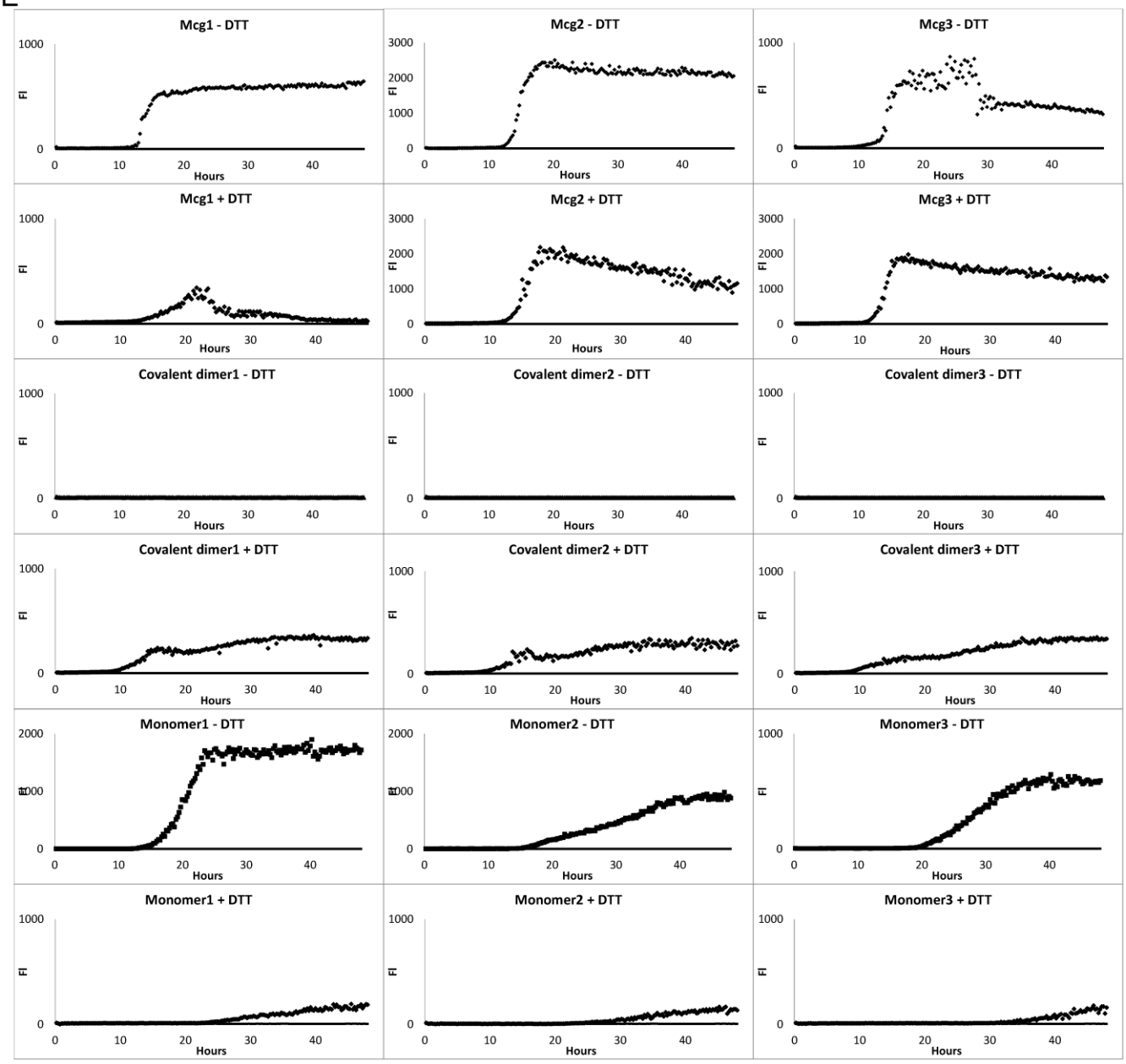


Figure 8