The structure of a folding intermediate provides insight into differences in immunoglobulin amyloidogenicity

Matthias J. Feige*†, Sandra Groscurth*†, Moritz Marcinowski*, Zu Thur Yew‡, Vincent Truffault*§, Emanuele Paci‡, Horst Kessler*, and Johannes Buchner*¶

*Center for Integrated Protein Science Munich and Department Chemie, Technische Universität München, Lichtenbergstrasse 4, 85747 Garching, Germany; and ‡Astbury Centre for Structural Molecular Biology, University of Leeds, LS2 9JT Leeds, United Kingdom

Edited by Alan R. Fersht, University of Cambridge, Cambridge, United Kingdom, and approved July 2, 2008 (received for review March 20, 2008)

Folding intermediates play a key role in defining protein folding and assembly pathways as well as those of misfolding and aggregation. Yet, due to their transient nature, they are poorly accessible to high-resolution techniques. Here, we made use of the intrinsically slow folding reaction of an antibody domain to characterize its major folding intermediate in detail. Furthermore, by a single point mutation we were able to trap the intermediate in equilibrium and characterize it at atomic resolution. The intermediate exhibits the basic β-barrel topology, yet some strands are distorted. Surprisingly, two short strand-connecting helices conserved in constant antibody domains assume their completely native structure already in the intermediate, thus providing a scaffold for adjacent strands. By transplanting these helical elements into β_2-microglobulin, a highly homologous member of the same superfamily, we drastically reduced its amyloidogenicity. Thus, minor structural differences in an intermediate can shape the folding landscape decisively to favor either folding or misfolding.

amyloids | NMR | protein folding | antibodies | molecular dynamics

In the current view, almost all proteins are believed to populate partially folded species, so-called folding intermediates, along their pathways to the native state (1–6). The characteristics of folding intermediates are critical in determining whether a protein is able to fold robustly or has the tendency to misfold (6, 7). A detailed structural characterization of folding intermediates is thus key for the understanding of protein folding in general. Because they are transient, however, only very few folding intermediates have so far been described in atomic detail (8–13). In general, experimental data argue for near-native topology with incompletely folded or partially misfolded structural elements such as side-chain interactions (8–13). In this context it is of particular importance that partially folded states have recently been associated with a large variety of pathologies (14). In the case of amyloid diseases—e.g., transthyretin (TTR) familial amyloid polyneuropathy (15), light chain amyloidosis (16), or dialysis-related amyloidosis (17)—it is thought that folding intermediates are key precursors for the formation of amyloid fibrils (18). Amyloids are long-lived dissociation- and degradation-resistant structures. They are made up of β-strands that are arranged into sheets lying perpendicular to the long fiber axis and possess a core cross-β structure (19). Despite the large variety of native folds shown by amyloidogenic proteins, these structural features seem to be a recurring motif in amyloids suggesting a common assembly mechanism (20). A particularly well studied example in this respect is the MHCI component β_2-microglobulin (β_2m), a member of the widespread immunoglobulin (Ig) superfamily (21, 22). β_2m has been shown to form amyloid fibrils if partially unfolded—e.g., at acidic pH (23), where the reaction is thought to be initiated by the population of a partially folded intermediate (23, 24). When such intermediates are populated for long periods, they are particularly susceptible to misfolding and misassembly reactions as has been shown for β_2m. Its productive folding from a native-like intermediate to the native state is limited by an intrinsically slow *trans*-to-*cis* peptidyl-prolyl isomerization reaction (24, 25). Experiments in which the critical proline residue was held in a *trans* state confirmed that this intermediate is a major determinant in amyloid formation (24, 26). In this regard, several studies showed that the most probable amyloidogenic precursor already possesses a large part of the native β-sheet topology with only the outer strands and loop regions being distorted (24, 25, 27).

Bearing in mind that intermediates are a rather general aspect of a protein folding reaction and that most polypeptides are in principle susceptible to amyloid formation (28), the question arises of how proteins avoid aggregation in the majority of cases. To address this issue we set out to study the folding pathway of the constant domain of the antibody light chain (C_L) with high structural resolution. The C_L domain is a particularly instructive model system because it also belongs to the Ig superfamily and, like β_2m, forms a β-sandwich composed of seven strands stabilized by a single disulfide bond between strands B and F (29, 30). The *cis* proline residue associated with the amyloidogenic potential of β_2m is conserved in the C_L domain (29). Furthermore, the overall folding mechanisms of the two proteins are highly similar (24, 30), each populating an intermediate state en route to the native state. Nevertheless, the C_L domain has never been directly associated with amyloidogenic diseases even if present at much higher concentrations than β_2m in the blood (31). By the structural characterization of its major folding intermediate, we show how the C_L antibody domain might avoid such harmful misfolding reactions.

Results

The Major Kinetic Folding Intermediate of C_L is Highly Structured. The C_L domain folds via an obligatory intermediate on two parallel pathways to its native state, the slower one being limited by the isomerization of the Y34–P35 bond to the native *cis* conformation (30, 32). This bond is predominantly *trans* in the unfolded state. As a consequence, only ≈10% of the molecules are able to fold to the native state within a few seconds (30, 32), and ≈90% of the molecules have to undergo the intrinsically slow isomerization reaction before complete folding to the native state (30, 32). At 2°C

Author contributions: M.J.F., S.G., M.M., Z.T.Y., E.P., H.K., and J.B. designed research; M.J.F., S.G., M.M., Z.T.Y., and V.T. performed research; M.J.F., S.G., M.M., Z.T.Y., V.T., E.P., H.K., and J.B. analyzed data; and M.J.F., S.G., Z.T.Y., E.P., H.K., and J.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

†M.J.F. and S.G. contributed equally to this work.

§Present address: Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany.

¶To whom correspondence should be addressed. E-mail: johannes.buchner@ch.tum.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0802809105/DCSupplemental.

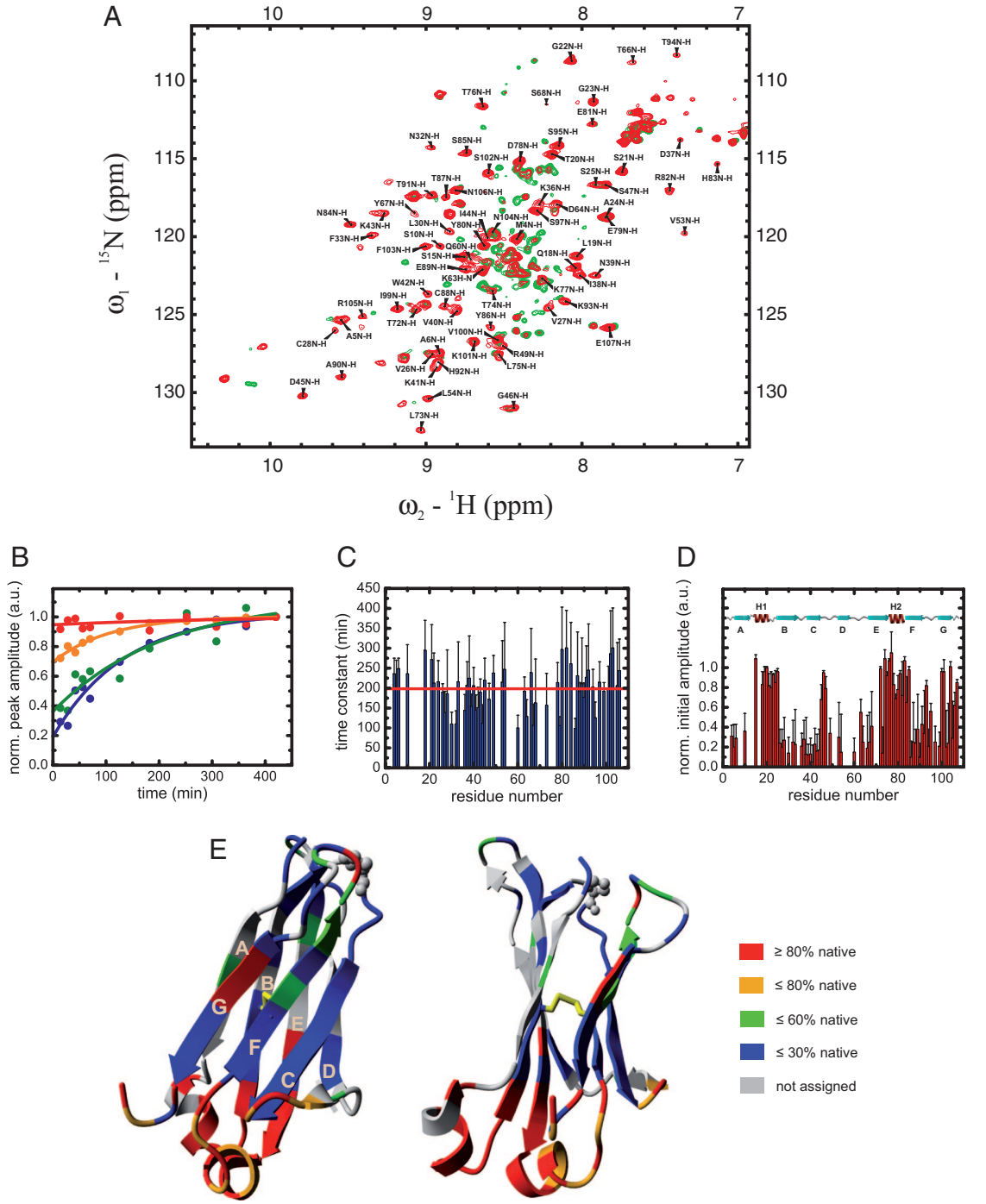

Fig. 1. Structural characterization of the major C_L folding intermediate by NMR spectroscopy. (*A*) ^{15}N-^{1}H HSQC spectrum of the native C_L domain including the backbone resonance assignment is shown in red and the one of the intermediate in green. The spectrum of the intermediate was derived from the first HSQC spectrum measured during refolding after 14 min and corrected for the 10% C_L molecules that refold within the dead-time of the experiment. (*B*) The change in intensity over the time for each peak was fitted by a single exponential function and extrapolated to time 0 (blue, K36; green, E89; orange, E79; red, L19 selected as representative residues). The time constants for every assigned residue are shown in *C*. The red line denotes their mean value. Initial amplitudes are shown in *D*. (*E*) Initial HSQC amplitudes mapped on the crystal structure of the native C_L domain (PDB entry 1FH5). The colors represent the initial HSQC amplitudes relative to the native HSQC amplitudes. All spectra were recorded at 2°C. Protein concentrations of 0.5–1.0 mM in PBS buffer with a GdmCl concentration of 0.2 M were used.

this reaction takes several hours to complete [see supporting information (SI) Fig. S1], allowing the major kinetic intermediate to be populated for a significant amount of time. CD spectra of the intermediate argue for a partially formed β-sheet framework and the absence of defined asymmetric environment around the aromatic amino acids (see Fig. S1).

To structurally characterize the intermediate state as well as the folding process on a residue level, >70% of the C_L domain backbone was assigned by standard NMR techniques (Fig. 1*A*), and real-time ^{15}N-^{1}H HSQC spectra were measured during refolding from the chemically denatured state. The first spectrum recorded after 14 min reflects almost exclusively the kinetic intermediate and

had only to be corrected for 10% of the C_L molecules possessing the correct Y34–P35 isomerization state (Fig. 1*A*; see *Materials and Methods* for details). Because the chemical shifts of the amide protons strongly depend on their molecular environment, overlaying the HSQC spectra of the intermediate and the native state reveals similarities and changes in their environment during the folding process (Fig. 1*A*). The HSQC spectra of the native C_L domain and the folding intermediate are superimposable for some residues but non-superimposable for others where significant differences in the chemical shifts are observed (Fig. 1*A*). To obtain more insights into the structural properties of the intermediate, the change in the peak intensity at the native chemical shift position was followed over time for each assigned residue. In every case the change in peak intensity could be well described by a single exponential function, if not already showing a native-like intensity after the dead-time of the experiment (Fig. 1*B*). As can be seen in Fig. 1*C*, the time constants of the folding of the individual residues show stochastic behavior around a mean value of $\tau = 199$ min at 2°C without any significant systematic deviations for any part of the protein. In contrast, initial HSQC amplitudes in the folding intermediate show interesting patterns: almost native initial amplitudes are found in correspondence of the two short helices connecting strand A and B as well as E and F and adjacent β-sheet termini suggesting that they are already in a native environment in the intermediate whereas low initial amplitudes are observed for some of the β-strands, in particular strands C and D, which suggests a lack of native structure (Fig. 1*D*). In Fig. 1*E* regions of high or low initial amplitudes are mapped on the crystal structure of C_L revealing that the two helices and their local environment are highly structured in the intermediate.

An Intermediate Structure at Equilibrium Trapped by a Single Point Mutation. The initial HSQC amplitudes only provide hints on the structural properties of the intermediate. Equilibrium spectroscopic data could provide information more directly related to structure. Therefore, we tried to trap the intermediate at equilibrium by exploiting the isomerization reaction separating the intermediate from the native state. We hypothesized that mutating the P35 residue against another amino acid that preferentially adopts a *trans* peptide bond (33), such as Ala (C_L^{P35A}), might "trap" the kinetic intermediate making it populated at equilibrium. Indeed, far-UV and near-UV CD spectra of C_L^{P35A} were found to be very similar to the respective spectra of the kinetic intermediate (data not shown). To determine the stability of the mutant in comparison to the wild type (C_L^{wt}), denaturant-induced unfolding transitions were performed. The unfolding of both proteins, C_L^{wt} and C_L^{P35A}, was a two-state process because there was concurrent loss of secondary structure (monitored by far-UV CD-spectroscopy) and tertiary structure (monitored by the change in the intrinsic tryptophan fluorescence) (Fig. 2*A*). Consistent with a partially folded species being the major equilibrium state, the P35A mutation led to a stability reduction of the C_L domain by >50% from $\Delta G_{unfolding} = 13.4 \pm 0.9$ kJ·mol^{-1} for the wild-type protein to $\Delta G_{unfolding} = 6.1 \pm 0.5$ kJ·mol^{-1} for the mutant (Fig. 2*A*). The cooperativity parameter of the transition decreased by 20% from $m_{eq} = 15.9 \pm 0.9$ kJ·mol^{-1}·M^{-1} for C_L^{wt} to $m_{eq} = 12.8 \pm 0.8$ kJ·mol^{-1}·M^{-1} for C_L^{P35A}. If the kinetic intermediate is trapped by the Pro35Ala mutation, a simplified folding mechanism devoid of the kinetic phases leading to the native state is expected for C_L^{P35A} as compared to C_L^{wt}. To assess the folding mechanism of C_L^{P35A}, a chevron plot was determined. The mutant only showed one major folding/unfolding phase in the chevron plot (Fig. 2*B*) giving rise to one folding and one unfolding microscopic rate constant ($k_f = 4.4 \pm 0.1$ s^{-1}/$k_u = 0.30 \pm 0.02$ s^{-1} at 20°C and 0 M GdmCl). In contrast to this, two separate folding phases had been reported for the major folding pathway of C_L^{wt} (30). From a global fit of the overall folding mechanism, the rate constants for each individual folding process could be determined for C_L^{wt} (30). It folds to its intermediate state with a very similar rate

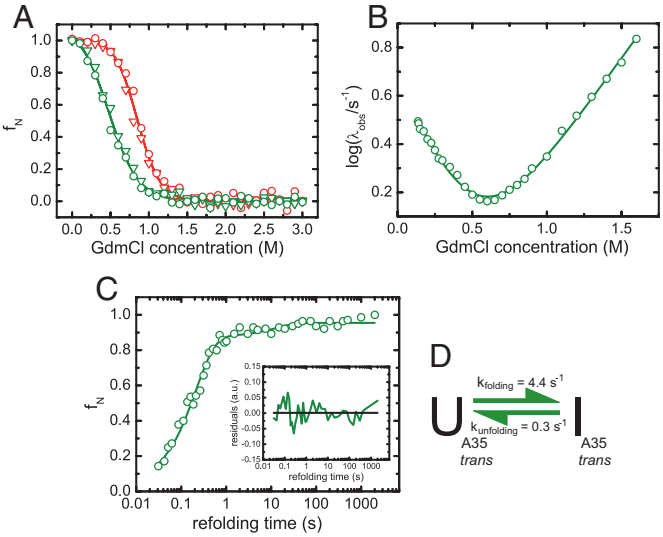

Fig. 2. Stability and folding mechanism of the C_L^{P35A} mutant. (*A*) Equilibrium unfolding transitions of C_L^{P35A} (green) and C_L^{wt} (red) determined by the intrinsic tryptophan fluorescence excited at 280 nm and detected at 360 nm (circles) as well as the far-UV CD-signal at 218 nm (inverted triangles). The data were fit to a two-state unfolding model. (*B*) Chevron plot for C_L^{P35A} determined by stopped-flow fluorescence spectroscopy. It could be described by a two-state folding reaction. (*C*) Formation of the final refolding species of C_L^{P35A} was followed by interrupted refolding experiments. The data were fit by a double exponential function (residuals shown as *Inset*) with 93% of the molecules folding via the fast pathway. (*D*) The folding mechanism of C_L^{P35A} can be described by a simple two-state model neglecting the 7% slow folding species. Transitions were measured at 10 μM protein concentration, and all kinetic experiments were performed at a final protein concentration of 2 μM. All measurements were carried out at 20°C in PBS.

constant as C_L^{P35A} folds to its final state (30). From the derived rate constants, the stability of the C_L^{wt} folding intermediate could be calculated to 11 ± 2 kJ·mol^{-1}, which is slightly higher than the stability of C_L^{P35A} ($\Delta G_{unfolding} = 6.1 \pm 0.5$ kJ·mol^{-1}). $\Delta G_{unfolding}$ for the intermediate of C_L^{wt}, however, is only an indirect estimate. Additionally, C_L^{P35A} unfolds faster than the C_L^{wt} intermediate (30) pointing towards a certain kinetic destabilization of the intermediate structure by the P35A exchange itself.

Notably, for C_L^{P35A} $\Delta G_{kinetic} = 6.6 \pm 0.3$ kJ·mol^{-1}, the stability calculated from the rate constants obtained from the chevron plot, is in very good agreement with the value from the equilibrium unfolding experiment ($\Delta G_{unfolding} = 6.1 \pm 0.5$ kJ·mol^{-1}) confirming a two-state kinetic process. This was corroborated by interrupted refolding experiments (Fig. 2*C*), which show that 93% of the molecules fold to their final state with a rate constant of 4.8 ± 0.4 s^{-1} and only 7% with a rate constant of 0.08 ± 0.01 s^{-1}, which can likely be attributed to the isomerization of one of the four remaining Proline residues from a nonnative *cis* to the native *trans* state. All data argue for the C_L^{P35A} mutant populating the kinetic intermediate in equilibrium, which is trapped by the *trans* state of the bond preceding A35. The data are summarized in the folding model shown in Fig. 2*D*.

To confirm this, the structure of C_L^{P35A} was further characterized by NMR spectroscopy. ^{15}N-^{1}H HSQC spectra of the C_L^{wt} kinetic intermediate and C_L^{P35A} were almost completely superimposable (Fig. 3*A*), demonstrating equivalent secondary and tertiary structure in both species. Crucially, the trapping of the intermediate state allowed an NMR assignment to be carried out (Fig. 3*A*), which was not feasible for the transiently populated kinetic folding intermediate observed in the folding of C_L^{wt}. To further compare the native and the intermediate states, a set of NMR measurements was recorded for both proteins (C_L^{wt} and C_L^{P35A}). These included, besides

BIOPHYSICS

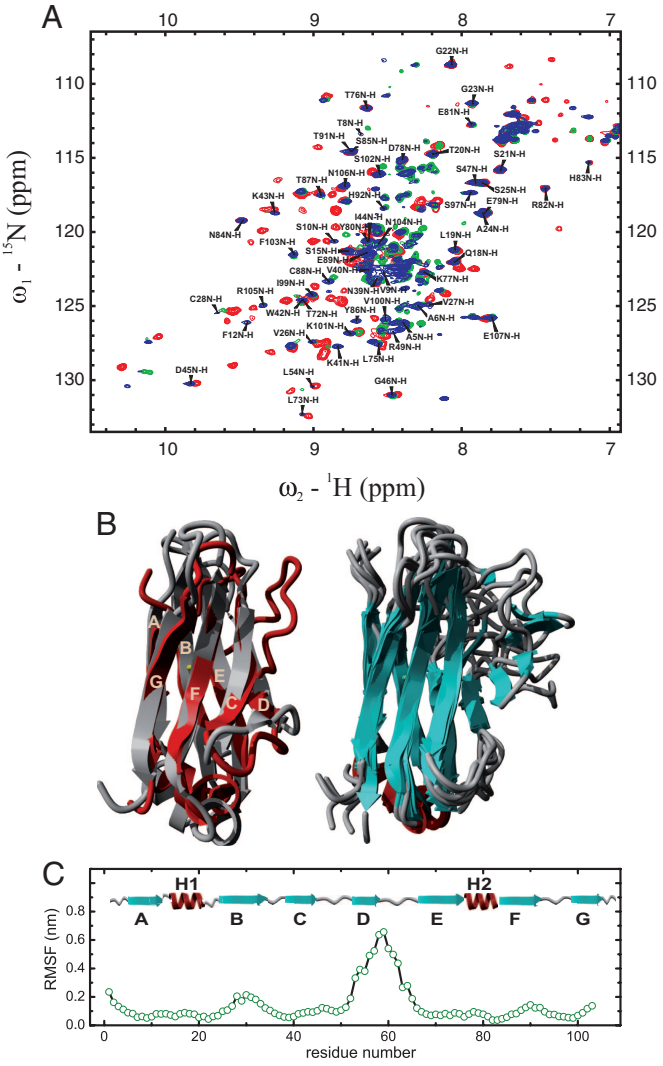

Fig. 3. NMR-spectroscopic characterization of the C_L^{P35A} mutant and structural comparison to the kinetic folding intermediate and C_L^{wt}. The ^{15}N-^{1}H HSQC spectrum of C_L^{P35A} including the backbone resonance assignment is shown in blue (*A*). For comparison, the HSQC spectrum of the kinetic folding intermediate is shown in green and the one of the wild type in red. All spectra were recorded at 2°C and protein concentrations of 0.5–1.0 mM in PBS buffer with a GdmCl concentration of 0.2 M. (*B Left*) Overlay of the energy-minimized native C_L structure (gray) with the average of 30 structures from the NMR-restrained MD simulations of C_L^{P35A} (red). (*B Right*) Eight structures from the restrained C_L^{P35A} MD simulations are shown to highlight flexible parts. (*C*) Root mean square fluctuations (RMSF) for the last 30 C_L^{P35A} structures derived from the restrained simulations.

the standard backbone experiments and triple resonance experiments for aliphatic side-chain assignment, 3D-NOESY spectra. These data revealed that large parts of the proteins possess almost identical carbon chemical shifts and, in addition, corresponding NOESY strips within the same range (Fig. S2 and data not shown), indicating that the intermediate already adopts a highly ordered structure. Notably, the two helices are fully formed, as judged from the NOESY pattern and stabilized by distinct helix capping motifs (34). Regions in C_L^{P35A} with significant differences in the carbon chemical shifts and the NOESY pattern from C_L^{wt} are all located around A35 and strands D and E. To gain further dynamic and structural information that cannot be deduced from the experiments alone, we performed molecular dynamics (MD) simulations with NMR-derived restraints (see *Materials and Methods* for details). Low-energy conformations in which restraints are collectively minimally violated can be sampled by this approach and provide an ensemble of structures that best represent the partially folded state. The simulations show that the overall topology of C_L^{P35A} is on average well retained (Fig. 3*B*). Additionally, both helices were fully structured in the simulations and showed only minor fluctuations (Fig. 3 *B* and *C*). Strikingly, the edge strands A and G were found to be native in all simulations (Fig. 3 *B* and *C*). However, we observe considerable heterogeneity at the ends of some strands in the ensemble of structures (Fig. 3*C*). This is especially pronounced around A35, arguing for a partial distortion of one edge of the protein by the *trans* state of the peptide bond preceding residue 35. The only strand which was highly distorted was strand D (Fig. 3 *B* and *C*). Only residues V53 and L54 of this strand were relatively native-like in the ensemble of structures. This is in good agreement with NOE signals observed exclusively for these residues within strand D (data not shown). The N terminus of strand E and strand G were found to be flexible—i.e., a mixture of structures with full or partial ordering of the respective strands was observed in the MD simulations (Fig. 3 *B* and *C*). Interestingly, the partial disordering of strand E and one of its ends provides a rationale for the missing NMR assignment of some residues from its flanking strand B, which itself is found to be highly structured in the simulations (Fig. 3 *B* and *C*). The overall solvent-accessible surface area of C_L^{P35A} was found to be on average 10% larger than for the wild type, which is in good agreement with the observed 20% decrease in the cooperativity parameter, m_{eq}, for unfolding. Taken together, the NMR experiments in combination with the simulations provide a detailed picture of the major C_L folding intermediate. The two small helices and their local environment are completely folded, and the intermediate exhibits a native-like core structure despite the presence of flexible regions that are able to adopt a variety of conformations.

Dissection of the Amyloidogenic Properties of the C_L Domain and β_2-Microglobulin. The constant domain of the antibody light chain has never been reported to be directly responsible for amyloidogenic processes even though it possesses the same topology as amyloidogenic variable antibody domains (V_L) or β_2-microglobulin (β_2m). In both cases, amyloid formation is assumed to proceed from a partially folded intermediate state (27, 35). Interestingly, neither amyloidogenic protein possesses the short strand-connecting helices that we identified as highly structured elements in the C_L folding intermediate. Accordingly, the sequence or structure of these helical elements might play a role in the inhibition of amyloid formation. To test this, we exchanged the unstructured loop regions connecting strands A and B as well as strands E and F in β_2m against the corresponding helical elements of the C_L domain (β_2m^{toCL}) (see Fig. S3). The β_2m^{toCL} exchange mutant folds to a well defined structure with similar far-UV CD-spectroscopic properties as wild-type β_2m (see Fig. S3 for details). The helical elements destabilize β_2m^{toCL} against thermal denaturation but have only a minor effect on its pH stability as compared to β_2m (see Fig. S3). According to the TANGO algorithm (36), the aggregation propensity of its primary sequence is left unaffected by the mutations (data not shown). To assess the amyloidogenicity of the different proteins, we used established reaction conditions for β_2m (37) and monitored fibril formation by atomic force microscopy (AFM). As expected, C_L^{wt} and the C_L^{P35A} mutant were not prone to fibril formation. In only one out of seven individual experiments were fibrils observed at pH 1.5, yet no fibrils were found at pH 3.0 or under physiological conditions, whether seeded or not (see Fig. S4). Importantly, a clear difference in amyloidogenicity is observed for wild-type β_2m and the β_2m^{toCL} mutant. Whereas β_2m readily formed fibrils under all conditions tested, β_2m^{toCL} only formed short fibrils at pH 1.5 (Fig. 4). At pH 3.0, fibrils were only detected in two out of seven individual experiments for β_2m^{toCL}, and no fibrils were detected under physiological conditions for this protein, even when β_2m fibrils were used for cross-seeding experiments (data not shown).

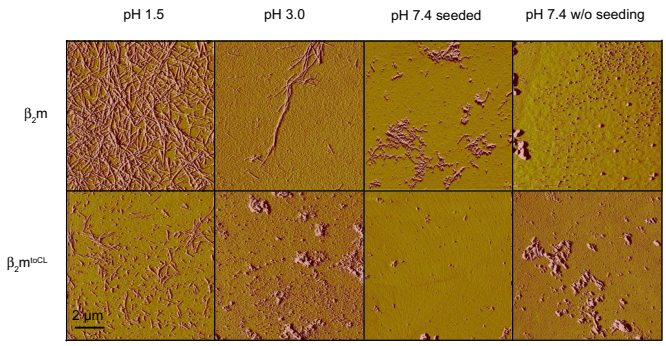

Fig. 4. Amyloidogenic properties of β_2m and the β_2m^{toCL} exchange mutant. Both proteins were incubated at pH 1.5 or 3.0 at 37°C for 7 days at a concentration of 50 μM. Additionally, the proteins were incubated under physiological conditions (PBS, 37°C) either seeded with β_2m or β_2m^{toCL} fibrils or not. Formation of amyloid fibrils was assessed by AFM measurements. Representative pictures of each sample are shown.

The data clearly show that transplanting the sequences corresponding to the C_L helices into the β_2m framework significantly reduces its amyloidogenicity.

Discussion

From a combination of NMR experiments and MD simulations we determined the ensemble of structures making up the major kinetic folding intermediate on the C_L folding pathway at atomic resolution. The structures provide hints on what sets an aggregation-prone folding intermediate apart from a structurally similar yet productive intermediate. On average, the overall β-sheet topology is well established for a major part of the protein, but most aromatic residues are still solvent-exposed or adjacent to dynamic structural elements. The only strand that seems to be highly disordered is strand D, despite native interactions of V53 and L54 with strand E. All other strands and in particular strands B, C, E, and F, which constitute the folding nucleus of Ig proteins (38, 39), exhibit some dynamics but are already well structured.

The most striking structural features of the intermediate are the two completely folded small helices. Although they are strongly conserved in constant antibody domains, their role in the folding process has not yet been recognized. These helices seem to fulfill a spacer and orienting function between strand pairs A–B and E–F and provide hydrogen bond donors and acceptors for adjacent strands and loops. In addition, the helices appear to position hydrophobic residues (e.g., Y80 in helix 2) so that they can participate in the formation of the hydrophobic core. Our data suggest that the two helices are able to fold efficiently and autonomously to their native structures in the context of the intermediate. Hence, these two helices can be regarded as a scaffold within the C_L intermediate favoring the formation of a native-like topology by correctly positioning important parts of the molecule.

An important protective role against amyloid formation has been attributed to the edge strands of β-sheet proteins (40, 41). Our finding that the edge strands A and G on one side of the C_L intermediate are highly structured provides one possible explanation for the marked difference in amyloidogenicity between C_L and β_2m. Grafting the sequences corresponding to the C_L helices onto the corresponding positions in β_2m significantly reduces the propensity of β_2m to form amyloids. This reduction is most pronounced at physiological pH, and the helices seem to be robust folding elements, which suggests a structural explanation for this effect but cannot completely rule out sequence effects of the transplanted elements themselves. Similar approaches to study the determinants of amyloid formation have been used previously but were targeted at increasing the amyloidogenicity of the "host" proteins rather than suppressing it as done here (42, 43). The edge strands A, D, and G have been reported to be disturbed in the amyloidogenic intermediate of β_2m (24–27, 44). This leaves edge strands on both sides of the protein molecule unprotected, making a linear arrangement of monomers into fibrils more likely than when just one side of the protein is partially distorted, as observed for C_L. For the amyloidogenic TTR, one complete sheet of the native β-sandwich structure is destabilized in the amyloidogenic precursor (45), and for V_L, the major folding intermediate is less structured than the C_L intermediate (38). In this context, we also note that β_2m forms fibrils more readily than C_L even though it is significantly more stable than C_L (24, 30). Accordingly, it is not the stability of the native state *per se* that is the most important factor determining the amyloid forming propensity of a protein, but rather the sequence of the protein (46) and structural characteristics of partially folded species that may be populated along the folding pathway (37).

In conclusion, our data show how a high degree of local structuring in a protein folding intermediate can significantly influence the folding landscape and favor robust folding over harmful misfolding. The different characteristics of C_L and β_2m can be understood in evolutionary terms. Selection of antibodies took place under harsh extracellular conditions with high concentrations of the multimeric protein present (47), whereas β_2m is found at much lower concentrations and usually associated with the MHCI complex (21). Thus, small differences, acquired over the course of evolution, between members of the same protein superfamily can lead to the avoidance of pathogenic misfolding reactions while preserving an identical protein topology.

Materials and Methods

Protein Production and Purification. Proteins were expressed and purified as described in *SI Methods*.

Optical Spectroscopy. CD kinetics and spectra were measured as described in *SI Methods*. Equilibrium unfolding transitions, stopped-flow, and interrupted refolding experiments were performed as described in ref. 30. For the interrupted refolding experiments, C_L^{P35A} was unfolded in 1.5 M GdmCl, refolded in 136 mM GdmCl for different times, and finally unfolded again in 1.5 M GdmCl.

NMR Spectroscopy. If not stated otherwise, all spectra were recorded at 25°C on Bruker DMX600, DMX750, and AVANCE900 spectrometers as described in *SI Methods*. For folding studies, ^{15}N-labeled unfolded C_L in PBS containing 2 M GdmCl was diluted 10-fold by adding ice-cold PBS without GdmCl. Real-time ^{15}N-^{1}H HSQC spectra were recorded at 2°C every 14 min by using selective proton flip-back pulses (48). Identical processing of all of the spectra was performed by using the program TOPSPIN 1.3 (Bruker BioSpin). To obtain the ^{15}N-^{1}H HSQC of the intermediate, 10% of the final spectrum was subtracted from the first recorded HSQC spectrum. Peak intensities were analyzed by using the program SPARKY (www.cgl.ucsf.edu/home/sparky). For kinetic studies, the intensities of every amino acid during the folding process were corrected for 10% native molecules and normalized to the corresponding intensity in the final spectrum after 7 h. Backbone resonance assignments were transferred from 25°C to 2°C recording a temperature series of spectra referenced to the internal standard TSP.

MD Simulations. The C_L^{P35A} mutant was created from the crystal structure of the C_L domain (PDB entry 1FH5). The bond preceding A35 was set to *trans* and the protein was energy-minimized with the steepest-descent and the adopted Newton–Rhapson methods (49). Distance and dihedral restraints derived from a comparison of the NMR chemical shifts for C_L^{wt} and the C_L^{P35A} mutant were used in conjunction with a simulated annealing (SA)-like protocol to derive a set of structures best representing the C_L^{P35A} mutant. One hundred cycles of SA were performed, and the final structures from the last 30 cycles were used for further analysis. Details about the restraints and the SA protocol are presented in *SI Methods*.

All simulations were performed with CHARMM (49), using the CHARMM19 force field, the EEF1 implicit solvent model (50), a Langevin dynamics scheme with a friction coefficient of 1 ps^{-1}, a time step of 2 fs, and holonomic constraints (SHAKE) on all bonds involving hydrogen atoms.

AFM Measurements. Fibrillization experiments and AFM measurements were performed as described in *SI Methods*.

ACKNOWLEDGMENTS. We thank Helmut Krause for performing the MS analysis and Emma R. Simpson for helpful comments on the manuscript. We are grateful for financial support of the Max-Buchner Forschungsstiftung, Deutsche Forschungsgemeinschaft Sonderforschungsbereich 749, and the Fonds der Chemischen Industrie. M.J.F. and M.M. acknowledge a Ph.D. scholarship from the Studienstiftung des deutschen Volkes. Z.T.Y. acknowledges a Ph.D. scholarship from the Wellcome Trust and the University of Leeds.

1. Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* 4:10–19.
2. Sanchez IE, Kiefhaber T (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J Mol Biol* 325:367–376.
3. Korzhnev DM, *et al.* (2004) Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* 430:586–590.
4. Neuweiler H, Doose S, Sauer M (2005) A microscopic view of miniprotein folding: Enhanced folding efficiency through formation of an intermediate. *Proc Natl Acad Sci USA* 102:16650–16655.
5. Bai Y (2006) Energy barriers, cooperativity, and hidden intermediates in the folding of small proteins. *Biochem Biophys Res Commun* 340:976–983.
6. Brockwell DJ, Radford SE (2007) Intermediates: Ubiquitous species on folding energy landscapes? *Curr Opin Struct Biol* 17:30–37.
7. Jahn TR, Radford SE (2008) Folding versus aggregation: Polypeptide conformations on competing pathways. *Arch Biochem Biophys* 469:100–117.
8. Balbach J, *et al.* (1995) Following protein-folding in real-time using NMR-spectroscopy. *Nat Struct Biol* 2:865–870.
9. Mayor U, *et al.* (2003) The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* 421:863–867.
10. Feng HQ, Zhou Z, Bai YW (2005) A protein folding pathway with multiple folding intermediates at atomic resolution. *Proc Natl Acad Sci USA* 102:5026–5031.
11. Religa TL, Markson JS, Mayor U, Freund SMV, Fersht AR (2005) Solution structure of a protein denatured state and folding intermediate. *Nature* 437:1053–1056.
12. Nishimura C, Dyson HJ, Wright PE (2006) Identification of native and non-native structure in kinetic folding intermediates of apomyoglobin. *J Mol Biol* 355:139–156.
13. Mizuguchi M, Kroon GJ, Wright PE, Dyson HJ (2003) Folding of a β-sheet protein monitored by real-time NMR spectroscopy. *J Mol Biol* 328:1161–1171.
14. Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75:333–366.
15. Hou X, Aguilar MI, Small DH (2007) Transthyretin and familial amyloidotic polyneuropathy—Recent progress in understanding the molecular mechanism of neurodegeneration. *FEBS J* 274:1637–1650.
16. Bellotti V, Mangione P, Merlini G (2000) Review: Immunoglobulin light chain amyloidosis—The archetype of structural and pathogenic variability. *J Struct Biol* 130:280–289.
17. Eakin CM, Miranker AD (2005) From chance to frequent encounters: Origins of β_2-microglobulin fibrillogenesis. *Biochim Biophys Acta Proteins Proteomics* 1753:92–99.
18. Kelly JW (1998) The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr Opin Struct Biol* 8:101–106.
19. Fandrich M (2007) On the structural definition of amyloid fibrils and other polypeptide aggregates. *Cell Mol Life Sci* 64:2066–2078.
20. Dobson CM (2003) Protein folding and misfolding. *Nature* 426:884–890.
21. Guo HC, *et al.* (1992) Different length peptides bind to Hla-Aw68 similarly at their ends but bulge out in the middle. *Nature* 360:364–366.
22. Bork P, Holm L, Sander C (1994) The immunoglobulin fold—Structural classification, sequence patterns and common core. *J Mol Biol* 242:309–320.
23. McParland VJ, *et al.* (2000) Partially unfolded states of β_2-microglobulin and amyloid formation in vitro. *Biochemistry* 39:8735–8746.
24. Jahn TR, Parker MJ, Homans SW, Radford SE (2006) Amyloid formation under physiological conditions proceeds via a native-like folding intermediate. *Nat Struct Mol Biol* 13:195–201.
25. Kameda A, *et al.* (2005) Nuclear magnetic resonance characterization of the refolding intermediate of β_2-microglobulin trapped by non-native prolyl peptide bond. *J Mol Biol* 348:383–397.
26. Eakin CM, Berman AJ, Miranker AD (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nat Struct Mol Biol* 13:202–208.
27. McParland VJ, Kalverda AP, Homans SW, Radford SE (2002) Structural properties of an amyloid precursor of β_2-microglobulin. *Nat Struct Biol* 9:326–331.
28. Fandrich M, Fletcher MA, Dobson CM (2001) Amyloid fibrils from muscle myoglobin—Even an ordinary globular protein can assume a rogue guise if conditions are right. *Nature* 410:165–166.
29. Augustine JG, de la Calle A, Knarr G, Buchner J, Frederick CA (2001) The crystal structure of the Fab fragment of the monoclonal antibody MAK33—Implications for folding and interaction with the chaperone BiP. *J Biol Chem* 276:3287–3294.
30. Feige MJ, Hagn F, Esser J, Kessler H, Buchner J (2007) Influence of the internal disulfide bridge on the folding pathway of the CL antibody domain. *J Mol Biol* 365:1232–1244.
31. van Rhee F, *et al.* (2007) High serum-free light chain levels and their rapid reduction in response to therapy define an aggressive multiple myeloma subtype with poor prognosis. *Blood* 110:827–832.
32. Goto Y, Hamaguchi K (1982) Unfolding and refolding of the constant fragment of the immunoglobulin light chain. *J Mol Biol* 156:891–910.
33. Pappenberger G, *et al.* (2001) Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics. *Nat Struct Biol* 8:452–458.
34. Aurora R, Rose GD (1998) Helix capping. *Protein Sci* 7:21–38.
35. Qin ZJ, Hu DM, 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.
36. Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* 22:1302–1306.
37. Smith DP, Jones S, Serpell LC, Sunde M, Radford SE (2003) A systematic investigation into the effect of protein destabilisation on β_2-microglobulin amyloid formation. *J Mol Biol* 330:943–954.
38. Freund C, Honegger A, Hunziker P, Holak TA, Pluckthun A (1996) Folding nuclei of the scFv fragment of an antibody. *Biochemistry* 35:8457–8464.
39. Hamill SJ, Steward A, Clarke J (2000) The folding of an immunoglobulin-like Greek key protein is defined by a common-core nucleus and regions constrained by topology. *J Mol Biol* 297:165–178.
40. Richardson JS, Richardson DC (2002) Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc Natl Acad Sci USA* 99:2754–2759.
41. Monsellier E, Chiti F (2007) Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep* 8:737–742.
42. Otzen DE, Kristensen O, Oliveberg M (2000) Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: A structural clue to amyloid assembly. *Proc Natl Acad Sci USA* 97:9907–9912.
43. Ventura S, *et al.* (2004) Short amino acid stretches can mediate amyloid formation in globular proteins: The Src homology 3 (SH3) case. *Proc Natl Acad Sci USA* 101:7258–7263.
44. Hoshino M, *et al.* (2002) Mapping the core of the β_2-microglobulin amyloid fibril by H/D exchange. *Nat Struct Biol* 9:332–336.
45. Liu K, Cho HS, Lashuel HA, Kelly JW, Wemmer DE (2000) A glimpse of a possible amyloidogenic intermediate of transthyretin. *Nat Struct Biol* 7:754–757.
46. Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 424:805–808.
47. Han JH, Batey S, Nickson AA, Teichmann SA, Clarke J (2007) The folding and evolution of multidomain proteins. *Nat Rev Mol Cell Biol* 8:319–330.
48. Diercks T, Daniels M, Kaptein R (2005) Extended flip-back schemes for sensitivity enhancement in multidimensional HSQC-type out-and-back experiments. *J Biomol NMR* 33:243–259.
49. Brooks BR, *et al.* (1983) CHARMM—A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217.
50. Lazaridis T, Karplus M (1999) Effective energy function for proteins in solution. *Proteins Struct Funct Genet* 35:133–152.