






Cryo-EM reveals structural breaks in a patient-derived amyloid fibril from systemic AL amyloidosis

Lynn Radamaker ¹, Julian Baur¹, Stefanie Huhn², Christian Haupt¹, Ute Hegenbart ³, Stefan Schönland³, Akanksha Bansal ¹, Matthias Schmidt ¹ & Marcus Fändrich ¹✉

Systemic AL amyloidosis is a debilitating and potentially fatal disease that arises from the misfolding and fibrillation of immunoglobulin light chains (LCs). The disease is patient-specific with essentially each patient possessing a unique LC sequence. In this study, we present two *ex vivo* fibril structures of a λ 3 LC. The fibrils were extracted from the explanted heart of a patient (FOR005) and consist of 115-residue fibril proteins, mainly from the LC variable domain. The fibril structures imply that a 180° rotation around the disulfide bond and a major unfolding step are necessary for fibrils to form. The two fibril structures show highly similar fibril protein folds, differing in only a 12-residue segment. Remarkably, the two structures do not represent separate fibril morphologies, as they can co-exist at different z-axial positions within the same fibril. Our data imply the presence of structural breaks at the interface of the two structural forms.

¹Institute of Protein Biochemistry, Ulm University, Ulm, Germany. ²Medical Department V, Section of Multiple Myeloma, Heidelberg University Hospital, Heidelberg, Germany. ³Medical Department V, Amyloidosis Center, Heidelberg University Hospital, Heidelberg, Germany. ✉email: marcus.faendrich@uni-ulm.de

Systemic AL amyloidosis is a protein misfolding disease that is newly diagnosed in 4–15 persons per one million per year in the United States of America¹ and other parts of the world^{2,3}. The amyloid deposits underlying this disease frequently occur in heart and kidneys, with cardiac involvement being the most important prognostic factor for the patient survival⁴. The LC amino acid sequence is highly variable, as a consequence of the recombination of different variable (V), joining (J), and constant (C) germ line (GL) segments, as well as somatic hypermutation⁵. Out of the resulting LC variants, the subtypes $\lambda 1$, $\lambda 2$, $\lambda 3$, $\lambda 6$, and $\kappa 1$ are in particular associated with AL amyloidosis^{6,7}.

It is well established that amyloid fibrils and other LC aggregates play a defining role in the pathogenicity of this disease^{4,5}. However, except for the chemotherapeutic removal of the pathogenic plasma cell clone, no pharmacological treatment options exist which directly prevent fibril formation or reverse fibril-induced organ damage⁸. One reason for this paucity of treatment options is the lack of knowledge about the mechanism of LC misfolding and the structure of pathogenic amyloid fibrils in vivo. Another reason is the patient-specific nature of systemic AL amyloidosis, with each patient presenting an essentially unique LC precursor and fibril protein⁹.

To provide insight into the fibril structure and LC misfolding mechanism in vivo, we recently set up a research strategy in which AL amyloid fibrils were extracted from diseased tissue and subjected to biochemical analysis^{10,11}. The fibril proteins are mainly derived from the LC variable (V_L) domain of the fibril protein precursor^{10,11}, consistent with earlier observations¹². They contain the intramolecular disulfide bond that is also present within the natively folded V_L domain^{10,13}. The fibrils are polymorphic¹¹, but consistent amyloid fibril morphologies are found in different organs/deposition sites within the same patient¹⁰. Different AL patients present different fibril morphologies^{10,11}, suggesting that the variability of the LC sequence leads to different, or even patient-specific fibril structures.

To obtain insight into their molecular conformations, we and others recently started to employ cryo-electron microscopy (cryo-EM) combined with three-dimensional (3D) reconstruction. So far, two AL amyloid fibrils were analyzed with this combination of methods, one derived from a $\lambda 1$ (ref. ¹³), termed hereafter FOR006, and one from a $\lambda 6$ LC¹⁴. The two fibril proteins showed markedly different folds, and their conformations differed fundamentally from natively folded LCs. In the present study, we analyze the structure of fibrils that were purified from the heart muscle tissue of a patient (FOR005) with $\lambda 3$ LC-derived amyloid fibrils. Using cryo-EM, we obtained two different fibril structures, termed here A and B. The two structures coexist at different z-axial positions within the same fibril, which implies the presence of structural breaks in these patient-derived amyloid fibrils.

Results

Extraction of the fibril protein and sequence analysis. The analyzed AL amyloid fibrils were extracted from the explanted heart of a female patient (FOR005) with systemic AL amyloidosis. The patient suffered from severe cardiomyopathy and underwent heart transplantation at the age of 50 years. We previously obtained the amino acid sequence of the fibril protein by protein sequencing, and the nucleotide sequence of the precursor LC by cDNA sequencing¹⁰. The tissue-deposited fibril protein consists of residues Ser2–Ser116 of the $\lambda 3$ precursor LC, which correspond to the V_L domain and a few residues (Gly109–Ser116) of the LC constant (C_L) domain. Bioinformatic analysis of the LC cDNA sequence indicated that it originates from the GL segments

*IGLV3-19*01*, *IGLJ2*01*, and *IGLC2*, which agrees with previous data showing that the *IGLV3-19* GL segment is linked to heart involvement^{6,7}. The amino acid sequence of the fibril protein differs from the protein sequence of the translated GL segments in several positions, probably as a result of the B-cell clone-specific somatic hypermutation. The fibril protein sequence contains five mutations with respect to the GL protein sequence within the *IGLV3-19* segment (Tyr31Ser, Tyr48Phe, Gly49Arg, Asn51Ser, and Gly94Ala), one within *IGLJ2*01* (Val97Gln), and one within *IGLC2* (Val135Gly). C_L mutations are rarely reported for patients with AL amyloidosis, possibly because the cDNA-based gene sequencing of the fibril protein precursor is often confined to the V_L domain.

Observation of two fibril structures in the fibril extracts. The extracted fibrils were subjected to cryo-EM and imaged at 300 kV (Supplementary Fig. 1a). Visual inspection of the recorded images revealed, despite evidence for polymorphism¹⁰, one apparently dominant fibril morphology that corresponded to >95% of fibrils visible in the fibril extracts. Picking this fibril morphology for 3D reconstruction and performing two-dimensional (2D) and 3D classification resulted in two 3D classes showing two different fibril structures, termed here A and B (Fig. 1 and Supplementary Fig. 1b). The corresponding reconstructions were refined to spatial resolutions of 3.2 Å for fibril structure A and 3.4 Å for B (Supplementary Table 1), based on the 0.143 Fourier shell correlation (FSC) criterion (Supplementary Fig. 2a). Their local resolution varied in the fibril cross sections, with higher resolution occurring at the fibril center and lower resolution toward the edges (Supplementary Fig. 2b). Additional rounds of 3D classification did not further subdivide the data sets in a meaningful fashion.

After the initial 3D classification, the data set contained 64,652 segments classified as fibril structure A and 36,667 as fibril structure B. The final reconstructions contained 11,003 segments for fibril structure A and 12,122 for fibril structure B (Supplementary Table 1). We interpreted the two reconstructions with molecular models (Fig. 1c, d) and obtained model resolutions of 3.1 Å for reconstruction A and 3.2 Å for B (Supplementary Table 1 and Supplementary Fig. 2c). 2D projections of the models correspond well to the 2D class averages of the original segments (Supplementary Fig. 3). Both models depict polar fibrils with C1 helical symmetry (Supplementary Table 1), consisting of only one protofilament and a single stack of fibril proteins (Fig. 1b). All peptide bonds of the fibril proteins, including the two X-Pro bonds, are modeled as *trans* isomers.

Both 3D maps contain diffuse density decorating the ordered fibril core (Fig. 1c, d, blue star), reminiscent of the two previously reported cryo-EM structures of ex vivo AL amyloid fibrils^{13,14}. These diffuse density regions may represent disordered parts of the fibril protein or non-fibril components. In addition, there is a well-defined density feature (Fig. 1c, d, red star) that appears to stem from a peptide segment in β -sheet conformation, owing to the zigzag pattern and a 4.8 Å rise along the fibril axis (Supplementary Fig. 4a, b). Similar well-defined density islands were previously observed with in vitro formed fibril structures^{15,16}. One study suggested that the density islands were formed from a segment of the fibril protein that was protruding from the main fibril core¹⁵. In another study, the density island originated from a peripherally attached fibril protein that adopted a single, short cross- β -strand but that was otherwise conformationally disordered¹⁶. As all segments outside the FOR005 fibril core are too short to reach our density islands (Supplementary Fig. 4c), we conclude that non-covalently attached fibril proteins, or fibril protein fragments, are the most plausible explanation of the density islands in the FOR005 fibril structure.

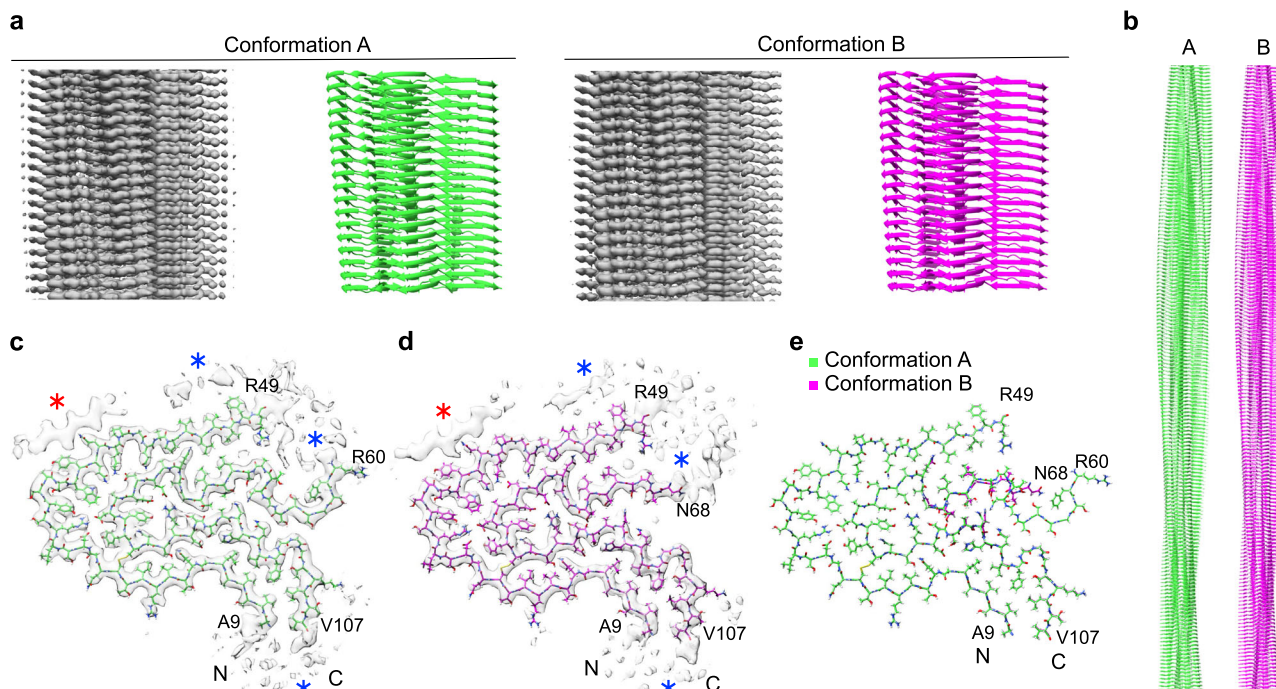


Fig. 1 Two different fibril protein conformations underlie the FOR005 amyloid fibrils. **a** Side views of the 3D maps of fibril structures A and B (left, gray), and corresponding molecular models (right, green/magenta). **b** Side view of longer segments of the two molecular models. **c, d** Cross-sectional views of the fibril protein conformations A (**c**) (EMD-11031) and B (**d**) (EMD-11030). Blue asterisk: region with blurry density surrounding the fibril core. Red asterisk: extra density decorating the fibril core, indicating an ordered peptide conformation. **e** Overlay of the molecular models of fibril structures A (PDB: 6Z10) and B (PDB: 6Z11). The N- and C-terminal residues of the model are highlighted.

Comparison of the fibril protein conformations A and B. The two fibril structures arise from similar but slightly different protein conformations. The fibril proteins are essentially indistinguishable at residues Ala9–Arg49 and Leu72–Val107. Residues Ser2–Pro8 and Leu108–Ser116 could not be assigned to any well-defined density in the 3D map, which implies that these segments are structurally heterogeneous or disordered. The main difference between the two structures lies in the segment Arg60–Ser71 (Fig. 1e). In fibril A, residues Arg60–Ser71 are in a stable conformation encompassing an arch, while residues Lys50–Asp59 are not well defined in the 3D map (Fig. 1c). In fibril B, residues Asn68–Ser71 are in a relatively extended conformation and the region of structural disorder occurs between residues Lys50–Gly67 (Fig. 1d). Importantly, mass spectrometry previously demonstrated the fibril protein to be continuous and to extend from Ser2 to Ser116 (ref. 10). Thus, the fibril core as seen in our 3D map is not made up of two fibril protein fragments, but instead it consists of two structurally ordered segments (Ala9–Arg49 and Arg60/Asn68–Val107) that are linked by a structurally heterogeneous region (Lys50–Asp59/Gly67).

The fibril protein shows β -strand conformation at residues Val10–Leu14, Thr17–Gln23, Asp25–Ser26, Arg28–Ser31, Trp34–Gln37, Pro43–Ile47, Leu72–Thr75, Ala79–Glu82, Tyr85–Tyr86, Asn88–Asp91, Asn95–Gln97, and Thr103–Thr106 in both fibrils (Fig. 2a). We refer to these segments as β 1 to β 12. Structure A contains two additional β -strands in a segment that is disordered in structure B (Arg60–Gly67). These strands are formed by residues Arg60–Ser62 and Ser64–Ser65 and are termed β 6' and β 6'' because they are in between the strands β 6 and β 7. All strands form cross- β -sheets with parallel, hydrogen bonded strand–strand interactions (Fig. 2b and Supplementary Fig. 5a). The protein fold is compact and devoid of large internal cavities.

Molecular interactions defining the fibril structure. The fibril proteins interact along the fibril axis through backbone hydrogen

bonds that extend between the strands of the cross- β -sheets. In addition, there are side chain–side chain interactions, such as polar ladders of asparagine or glutamine residues, or stacked hydrophobic or aromatic groups. These features are shown for residues Gln37 and Phe48 in the Supplementary Fig. 6a, b. The fibril backbones show axial height changes of 7.5 Å (fibril A) and 7.0 Å (fibril B), which lead to polar fibril topologies and sterically interlock the fibril layers. Each fibril protein layer interacts only with the layers above and below, e.g., Lys38 from layer i interacts with Asp81 from layer $i + 1$ (Supplementary Fig. 6c).

The surfaces of both fibrils are rich in charged and polar amino acids (Fig. 2c and Supplementary Fig. 5b). The fibril cores contain small hydrophobic patches, such as the one formed by residues Val10, Val12, Leu14, Val98, Phe99, Leu105, and Val107 (Fig. 2c and Supplementary Fig. 5b), as well as patches of buried polar residues. The structure buries a number of compensating charge–charge interactions, for example, at residues Asp25 and Arg28, Arg28 and Asp84, Lys38 and Asp81, Glu80 and Arg90 (Fig. 2c and Supplementary Fig. 5b), as well as an acidic moiety, which is not fully charge compensated. This moiety is formed by residues Glu80, Asp81, Glu82, and Asp84 (Fig. 2d), resembling the partially uncompensated acidic moiety in the previously described λ 1 fibril structure¹³. In contrast to the previous λ 1 fibril, however, there is no water-filled cavity around the acidic moiety in our fibril.

Location of aggregation-prone segments and mutations. The mutagenic changes of the amyloidogenic LCs compared with the GL sequences are widely believed to trigger amyloidosis in the respective patients^{9,17}. However, analysis of the mutated positions within our structure does not readily offer an explanation for their pathogenicity. Some mutations, such as Asn51Ser and Val135Gly, lie within a part of the precursor protein that is disordered or cleaved off in the fibril (Fig. 2a). In addition, none of

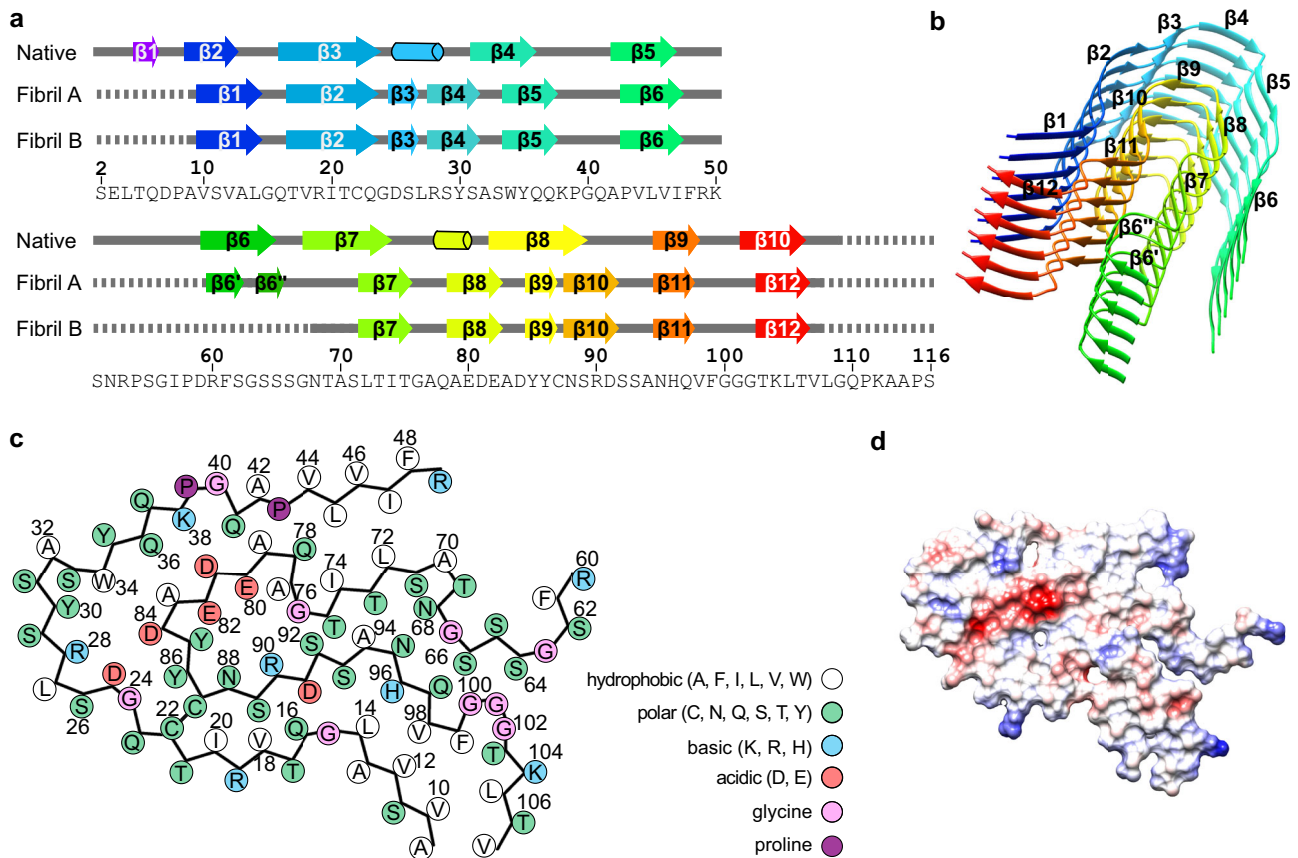


Fig. 2 Secondary structure and folding of the fibril proteins. a Schematic representation of the secondary structure of the fibril proteins A and B, and of a crystal structure of the refolded fibril protein (native, PDB: 5L6Q¹⁰). Arrows indicate β -strands and cylinders α -helical conformations. Continuous lines indicate ordered conformation, dotted lines indicate disordered segments. The definition of secondary structural elements follows the definition in the respective manuscripts. **b** Ribbon diagram of a stack of six fibril proteins (conformation A). β -strands have been colored in rainbow palette from the N- to the C-terminus. **c** Schematic representation of the amino acid positions in conformation A. **d** Electrostatic surface representation of the fibril protein conformation A. Red indicates negative charge, blue positive, and white neutral. Supplementary Figure 5 shows the corresponding images for fibril conformation B.

the mutations affecting the fibril core adds an obviously favorable interaction. The nonconservative Gly49Arg mutation even leads to a buried charge that is not compensated by a nearby opposite charge (Fig. 3a), suggesting that this mutation may even be unfavorable to the fibril structure. Moreover, analysis of the location of the mutations within the native LC does not readily provide evidence that they might be destabilizing to the native protein conformation (Supplementary Fig. 7a, b). The mutations do not remove an obviously stabilizing interaction and do not affect internal residues that might be considered to be crucial for protein stability. Instead, all mutations are located on the surface of the globularly folded LC (Supplementary Fig. 7a, b).

Computer-based predictions of the aggregation propensity of the FOR005 LC identified the highest aggregation score in the V_L domain at residues Val44–Arg49 (Fig. 3b and Supplementary Fig. 8). These residues form a hydrophobic patch on the fibril surface, which is decorated with the extra density region described above (Fig. 1c, d, blue star). The three disordered protein segments in the fibril protein (Ser2–Pro8, Lys50–Gly67, and Leu108–Ser116) correlate with regions having low aggregation scores (Supplementary Fig. 8). However, comparing the aggregation score of the FOR005 LC to that of the GL protein sequence (Supplementary Fig. 8) does not reveal any clear trend whether the FOR005 LC or its putative GL precursor is more aggregation prone (Supplementary Fig. 8). We conclude that the effect of mutation is subtle and not readily evident by the above

analysis. Support for this view comes from a recent study in which the rather counterintuitive observation was reported that a conservative leucine to valine mutation on the surface of a patient-derived V_L domain is strongly destabilizing to the native protein structure, and promotes the formation of amyloid fibrils *in vitro*¹⁸.

Conformations A and B coexist within the same fibrils. Finally, we sought to determine whether the two reconstructed 3D maps A and B represent two different fibril morphologies, or whether the two structures coexist within the same fibril particle. By visual inspection of the cryo-EM micrographs and measurement of global parameters, such as fibril width or crossover distance, we could not categorize the fibrils in our sample into separate structures A and B. Also, the reconstructed 3D maps have identical helical parameters, such as fibril symmetry, polarity, axial rise, and twist value, as well as a fibril pitch of 155 nm (Supplementary Table 1). The difference between the two structures could only be revealed when the fibril images were cropped into segments that were then aligned independently of their structural context during 3D classification.

Analyzing the origin of the fibril segments in the respective data sets producing reconstructions A and B, we would have expected, for separate morphologies, that each fibril contains segments belonging to only one of the two data sets A or B

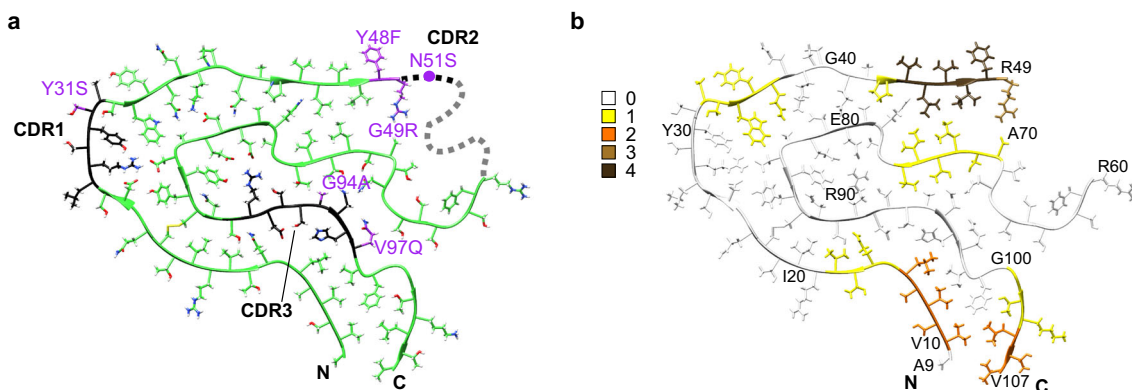


Fig. 3 Location of the mutational positions and aggregation-prone regions in the fibril structure. **a** Mutations with respect to the GL protein sequence (purple) and CDRs (black) marked in fibril structure A (green). The dotted line represents the intermedial disordered segment. **b** Molecular model of conformation A colored according to aggregation score.

(except for minor classification errors). Surprisingly, however, we found that most fibrils in our sample showed a mixture of A and B segments (Fig. 4a and Supplementary Fig. 9a). Fractions of type B segments per fibril varied almost continuously from 0 to 1 (Fig. 4b and Supplementary Fig. 9b). The assignment of segments to the fibril protein conformations A and B could not readily be correlated with certain positions on the fibril helix on the cryo-EM micrographs, for example, the crossover or the segment in between two crossovers, which might have suggested problems in their alignment. Moreover, the segments are not randomly distributed across the fibrils, but mostly separated into distinct regions along the fibril axis, in which all segments correspond to either structure A or B (Fig. 4a). These results were obtained consistently across different data sets, including the data set resulting from the initial 3D classification (Fig. 4a, b), as well as the data set from the final reconstructions (Supplementary Fig. 9a, b). In conclusion, the fibrils in our data set cannot simply be divided into two fibril morphologies A and B. Instead, the two fibril structures A and B occur simultaneously within a fibril protein stack. This observation indicates that there are structural breaks at the interface of fibril regions corresponding to structures A or B (Fig. 4c, d).

Discussion

We here present the cryo-EM structures of two amyloid fibrils (A and B) that were extracted from the explanted heart of a patient (FOR005) with systemic AL amyloidosis. The spatial resolutions are 3.2 Å for fibril structure A and 3.4 Å for fibril structure B (Supplementary Table 1). These resolutions are sufficient to establish the overall fibril topology and the fibril protein fold. However, uncertainty remains in the exact conformation of the backbone and side chains. This problem is further exacerbated by the known artifacts of cryo-EM structures, such as a loss of side chain density due to beam damage¹⁹, which could be relevant in our reconstructions, e.g., at residues Glu80 or Lys104 (Fig. 1c, d).

Systemic AL amyloidosis is a patient-specific disease⁹. Identification of common structural features in different patient-derived amyloid fibrils is potentially informative about common steps in the misfolding pathways across patient cases. Based on the available cryo-EM structures of ex vivo fibrils from systemic AL amyloidosis (Fig. 5 and Supplementary Fig. 10), the following commonalities can now be identified: the extracted fibril samples contain a dominant fibril morphology that consists of a single, polar protofilament. The fibril core is formed by the V_L domain of the precursor λ -LC in all cases. The C_L domain is structurally disordered and/or lost by proteolysis. The fibril proteins retain

the intramolecular disulfide bond of the native V_L domain, indicating that LC misfolding happens in an oxidative environment, such as the extracellular space or an endocytic compartment, that retains the disulfide bond of the native V_L domain. The fibril proteins show an antiparallel N-to-C orientation at the disulfide that is flipped by 180° relative to the native state. The fibril protein conformations differ substantially from the natively folded LC, demonstrating that a global structural rearrangement and/or unfolding reaction takes place during the conversion of the native LC, or of a LC fragment, into a fibril. The fibrils are decorated by blurry density of uncertain origin that may arise from fibril protein segments outside the ordered core, or cellular factors attached to the fibril surface.

Patient-specific features of the fibrils include the exact fold of the fibril protein (Fig. 5), and the location of the β -strands and disordered segments within the sequence (Supplementary Fig. 10). The current λ 3- and the previous λ 1-derived fibril proteins possess solvent-exposed and conformationally disordered N-termini, while the N-terminal segment of the λ 6-derived fibril is buried in the fibril core and part of a β -strand (Fig. 5 and Supplementary Fig. 10). The C-termini are disordered in each of these fibril proteins. Our current fibril structures and the previous λ 6 fibril structure each contain an internal, disordered segment interrupting the fibril protein fold. In contrast, the fold of the λ 1 fibril protein is continuous (Fig. 5 and Supplementary Fig. 10). The λ 1 fibril possesses three large channels, two of which are thought to be water-filled, while the third one contains an apolar molecular inclusion¹³. No such channels or inclusions were identified in the other fibril structures. A feature unique to the current λ 3 fibrils is a well-resolved density island attached to the fibril core (Supplementary Fig. 4a, b).

Our structures also differ from a number of studies which used nuclear magnetic resonance (NMR) spectroscopy to investigate the structure of LC-derived fibrils formed in vitro. These fibrils were formed from V_L domain constructs and include murine κ (ref. 20), human κ 1 (ref. 21), as well as human λ 3 (ref. 22) and λ 6 sequences²³. Particularly relevant in this case is the comparison of our structures to the NMR analysis of recombinant FOR005 V_L domain fibrils²². These fibrils were seeded in vitro with amyloid fibrils that were extracted from the heart of the patient (FOR005) with the aim to propagate the ex vivo fibril structure in the in vitro seeded fibrils²². Comparison of the in vitro seeded fibrils with our cryo-EM structures of patient fibrils revealed several differences. First, the ex vivo fibrils possess a stable β -strand at residues Thr103–Thr106 that are outside the ordered core of the in vitro seeded fibrils. Second, the in vitro seeded fibrils contain a salt bridge between residues Arg49 and

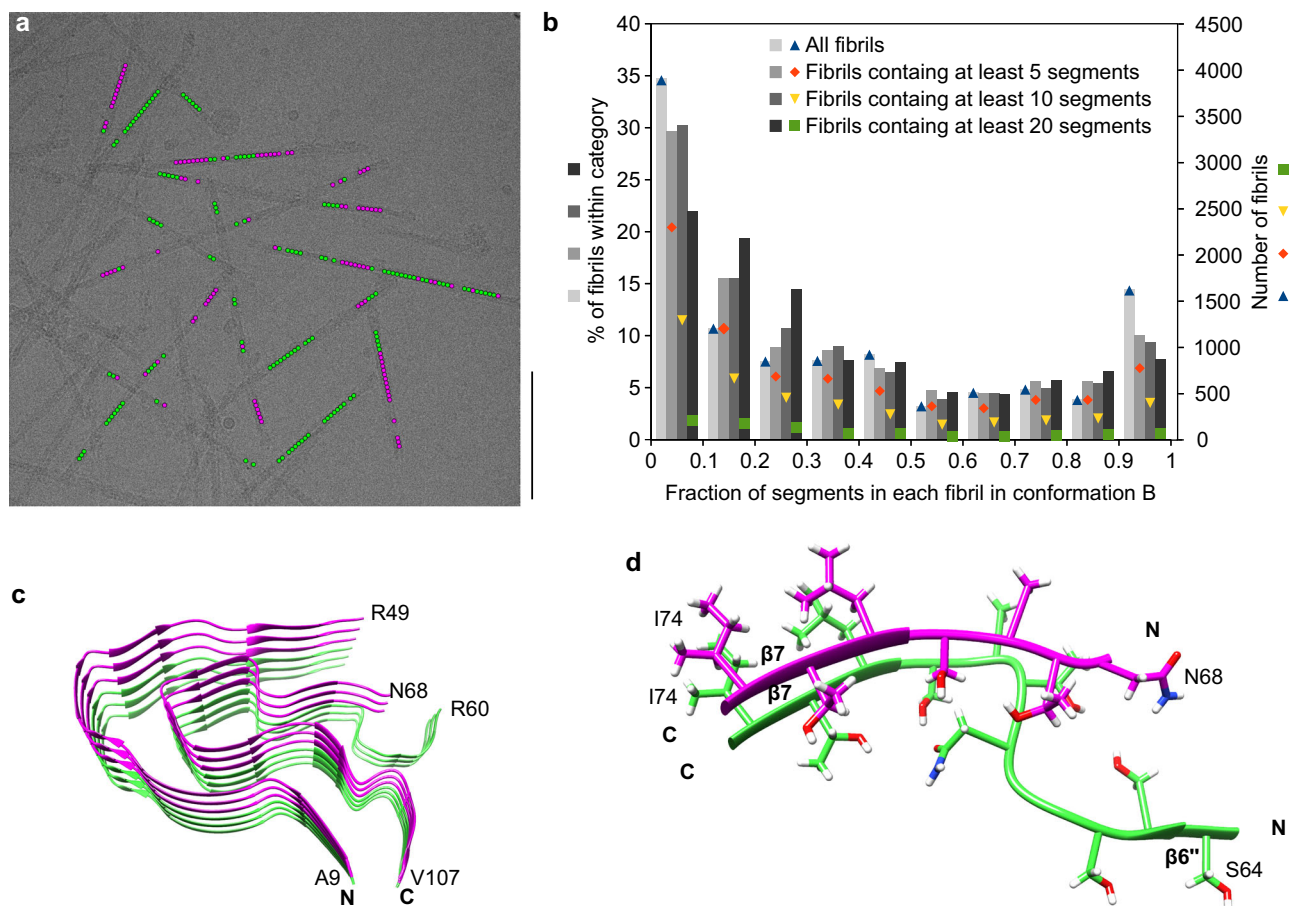


Fig. 4 Evidence for structural breaks in FOR005 fibrils. **a** Representative cryo-EM micrograph showing the location of segments classified after the first 3D classification as fibril structure A (green) and fibril structure B (magenta). Scale bar: 100 nm. Data were collected from 1,378 micrographs from one fibril sample. Supplementary Fig. 9a shows the same image, highlighting only the segments used in the final fibril reconstruction. **b** Histogram of the fraction of segments classified as fibril structure B, per fibril, after the first 3D classification. Different thresholds were chosen for the minimum number of segments per fibril, resulting in four categories: all fibrils (11,194 fibrils), fibrils containing a minimum of 5 segments (7,738 fibrils), fibrils containing a minimum of 10 segments (4,278 fibrils), and fibrils containing a minimum of 20 segments (951 fibrils). The percentages were normed using the total number of fibrils in each category. Colored points show the absolute number of fibrils in each category and group (fraction in conformation B). In total, the data set ($n = 101,319$) contained 64,652 segments classified as fibril structure A and 36,667 as fibril structure B. Supplementary Fig. 9b shows an analogous histogram, but including only the segments used for the final reconstructions. **c** Stack of three fibril proteins in conformation B (magenta) on top of three fibril proteins in conformation A (green), illustrating the presence of structural breaks within the patient amyloid fibrils. **d** Detailed view of a structural break, including side chains.

Asp25 (ref. ²²), which are far apart in the protein fold of the patient fibril (Fig. 2c). Third, the in vitro seeded fibrils show an electrostatic interaction between Lys50 and Asp81 (ref. ²²), whereas in the ex vivo fibrils, Lys50 is in the internal disordered region, and Asp81 is far away from this segment (Fig. 2c) and has an electrostatic interaction with Lys38 (Supplementary Fig. 6).

These data demonstrate that the in vitro fibrils are structurally different from the ex vivo fibrils analyzed here with cryo-EM. In vitro seeding with ex vivo FOR005 fibrils did not propagate, in this case, the seed structure to the daughter fibrils, although it modified the fibril structure compared with unseeded fibrils²². These observations imply that the seeding mechanism did not involve a replication of the seed structure. Early work with sickle cell hemoglobin identified two possible seeding mechanisms: homogeneous and heterogeneous nucleation. Homogeneous nucleation involves the attachment of the soluble fibril precursor proteins to the fibril tip, while heterogeneous nucleation involves the nucleation of new fibrils on the lateral cylindrical surface of an existing fibril²⁴. While attachment of the fibril precursor protein to the tip of an amyloid fibril would be expected to lead to a

replication of the fibril protein fold, this replication of the seed protein structure may not necessarily occur during heterogeneous seeding. Indeed, the observation of density islands on the outside of the FOR005 fibril core structure (Fig. 1c, d, red star and Supplementary Fig. 4a, b) suggests that fibril proteins may have attached on the fibril surface, but do not fully replicate the fold of the fibril protein. Hence, it is important to use patient-derived fibrils when investigating the structural basis of disease. A similar conclusion was obtained previously when FOR005 fibril protein was extracted from the patient's heart, denatured in guanidine, refolded, and fibrillated in vitro (without seeds). These in vitro fibrils also showed a different morphology than the fibrils that were purified from FOR005 patient tissue¹⁰, as judged by transmission electron microscopy.

A particularly interesting finding in the present study is the observation of structural breaks. So far, it has been part of our general understanding of amyloid fibril structures that these are conformationally uniform along the fibril axis. Occasionally, fibrils were reported that differed morphologically at its two ends^{25–28}. Some of these cases could be attributed to a fibril cross-

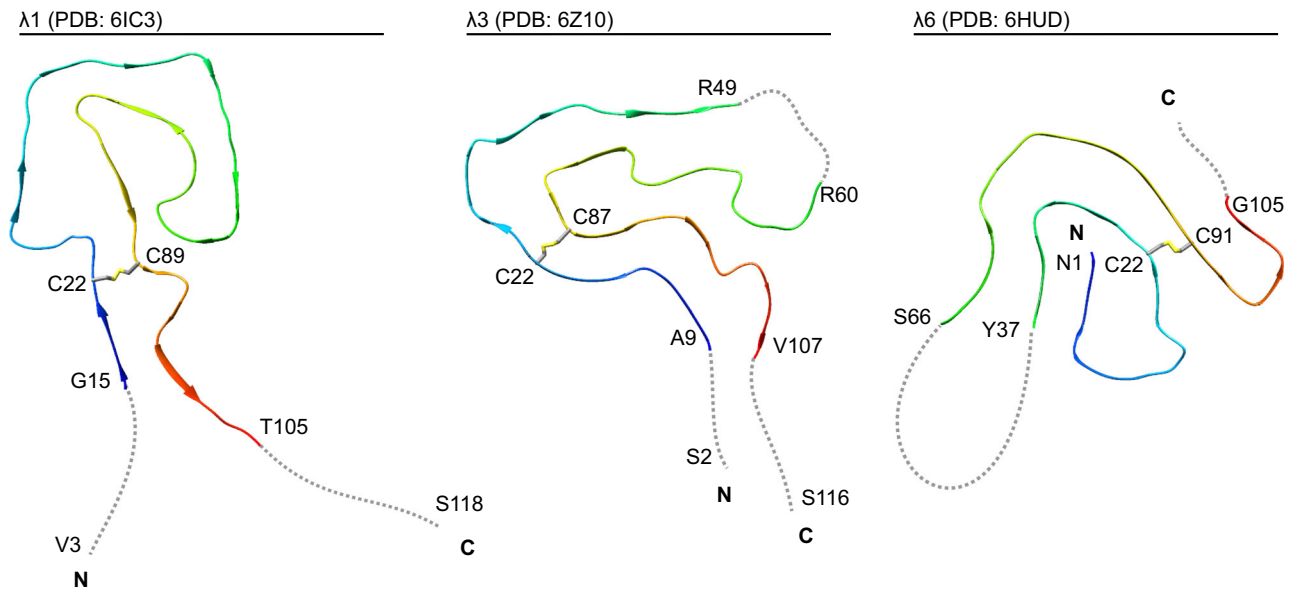


Fig. 5 Comparison of the available cryo-EM structures of ex vivo AL amyloid fibrils. Ribbon diagrams of a $\lambda 1$ fibril (PDB: 6IC3¹³), the current $\lambda 3$ fibril (conformation A, PDB: 6Z10), and a $\lambda 6$ fibril (PDB: 6HUD¹⁴). The fibrils are shown in a cross-sectional view. For all structures, the location of the disulfide bond forming cysteine residues is marked. Disordered segments are represented as dotted gray lines and depicted in an arbitrary conformation. The first and the last residue of the ordered segments, as well as the first and the last residue of the fibril protein are indicated, if known.

seeding, that is, the attachment of a different fibril precursor protein to a fibril tip, or to a splintering of a multi-protofilament fibril into fibril morphologies with a smaller number, or a different arrangement of protofilaments. In other cases, it was unclear whether two fibril morphologies may have annealed after their formation. In our samples, however, the majority of fibrils show a mixture of conformations, and show multiple, seemingly arbitrary switching between the conformations A and B in each fibril. Therefore, short segments, possibly down to a single protein layer, may be able to adopt a conformation different from that of the surrounding layers. The fibril breaks and the two fibril structures defining the breaks emerged at the 3D classification stage in our analysis and resolved a previously blurry density region into two distinct density paths (Supplementary Fig. 1b). Unresolved density regions resulting from one or more disordered segments of the protein chain are reported for the majority of cryo-EM structures of in vitro and ex vivo amyloid fibrils²⁹. Therefore, structural heterogeneity such as described here could be relevant to other fibril structures as well. Furthermore, it is possible that our fibrils contain fibril protein structures which we were unable to resolve so far. We originally extracted 194,502 fibril segments from the cryo-electron micrographs and used only 11,003 (A) and 12,122 (B) of these segments for the final 3D reconstructions (Supplementary Table 1).

Two possible scenarios can be envisioned to explain the mechanism of the formation of structural breaks. One scenario is that they appear during fibril assembly due to an imperfect replication of the seed structure, as a new molecule attaches to the fibril end (Fig. 6). Consistent with this idea, real-time microscopy studies explained the stop-and-go kinetics during fibril growth with irregularities in the addition of molecules to the tip of a growing fibril³⁰. The other scenario is that breaks emerge after fibril assembly, for example, because initially disordered segments adopt different stable conformations, which then proliferate along the fibril axis (Fig. 6). Support for the latter mechanism is provided by the fact that the type A and B fibril proteins are mostly identical, and that the differences are confined to a small segment that lies in the vicinity of an unstructured region.

While more work is necessary to discriminate between these two mechanisms, our observations lead to an important change in our understanding of the assembly of polypeptide chains into amyloid fibrils. They demonstrate that these linear aggregates are not as perfectly regular and uniform as has generally been assumed by most previous studies. Considering that the breaks were revealed in the FOR005 fibril samples only at an advanced stage of the analysis, we would predict that they will be observed more frequently in the future, as the methods of structural biology become more powerful and will be able to resolve such fine details more routinely. Structural breaks and other structural defects in cross- β -sheets could have significant ramifications for the biological properties of amyloid fibrils. Examples hereof include the fragility³¹ and the loss of torsional coherence of amyloid fibrils^{32,33}, the branching of amyloid fibrils during fibril outgrowth³⁴ and the ability of molecular chaperones to bind to, sever, and to break down amyloid aggregates³⁵.

Methods

Source of AL fibrils. Heart tissue was collected from a female patient (FOR005) at the age of 50, suffering from AL amyloidosis and consequent advanced heart failure. A monoclonal gammopathy was the underlying condition. The patient was treated within the heart transplant program of the University Hospital Heidelberg. The explanted heart tissue was stored at -80°C . The study was approved by the ethical committees of the University of Heidelberg (123/2006) and of Ulm University (203/18). Informed consent was obtained from the patient for the analysis of the amyloid deposits.

Fibril extraction from patient tissue. Applying a previously established protocol¹¹ for fibril extraction, 250 mg of patient heart tissue were diced finely and 0.5 mL of ice-cold Tris calcium buffer (20 mM Tris, 138 mM NaCl, 2 mM CaCl_2 , 0.1 (w/v) % NaN_3 , pH 8.0) added. The sample was homogenized using a Kontes Pellet Pestle, after which it was centrifuged for 5 min at $3100 \times g$ at 4°C . The washing step was repeated five times and each supernatant was stored for further analysis. Afterward, 1 mL of freshly prepared 5 mg mL^{-1} *Clostridium histolyticum* collagenase (Sigma) in Tris calcium buffer with ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor (Roche) were added and the pellet resuspended. Overnight incubation at 37°C was followed by a 30 min centrifuge cycle at $3100 \times g$. Ten further washing steps with 20 mM Tris, 140 mM NaCl, 10 mM EDTA, and 0.1 % (w/v) NaN_3 and ten subsequent steps with ice-cold water

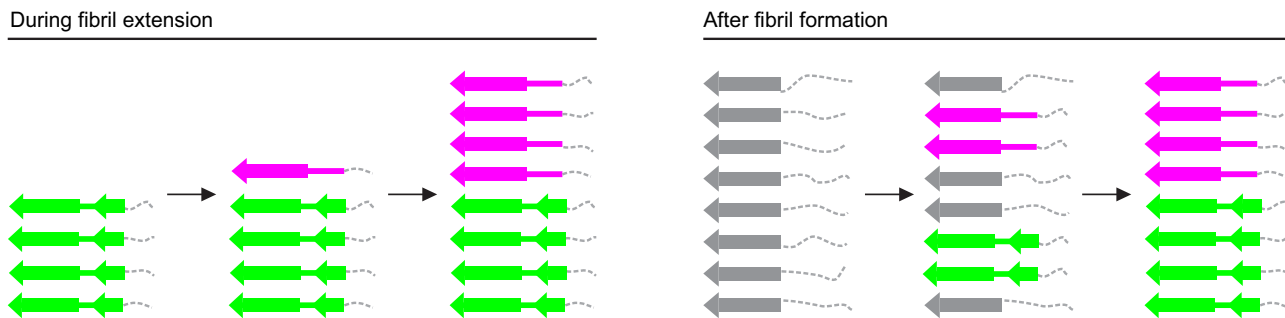


Fig. 6 The origin of structural breaks: two possible scenarios. Schematic representation of a stack of fibril proteins, illustrating two different hypotheses on how structural breaks form: during fibril extension (left) or after fibril formation (right). Conformation A is represented by two β -sheets (green). Conformation B is represented by one β -sheet (magenta). Disordered segments are represented by dotted gray lines. Gray arrows represent the immature fibril proteins, before the mature conformations A and B are adopted.

were then performed on the pellet, using a pipette for homogenization. One of the water supernatants was selected for cryo-EM.

Cryo-EM. Holey carbon-coated grids (C-flat 1.2/1.3 400 mesh) were glow-discharged using 40 mA for 40 s. Using a VitroBot (Thermo Fisher Scientific), 3.5 μ L of the extracted fibril sample were incubated on each grid for 30 s at a humidity of >95%, the excess fluid was blotted off and the grid plunged into liquid ethane, then transferred to a grid box. After plunging, one grid from each grid box (containing four grids) was checked using a 200 kV Jeol JEM 2100 F electron microscope (Ulm University). The remaining grids in the grid boxes were kept at liquid nitrogen temperature. Cryo-electron microscopic image acquisition of one selected grid was performed using a Titan Krios transmission electron microscope (Thermo Fisher Scientific) at 300 kV equipped with a K2-Summit detector (Gatan) in counting mode. A Gatan imaging filter with a 20 eV slit was applied. The software package SerialEM v3.7 was used for data collection. The data acquisition parameters can be found in Supplementary Table 1. Global parameters of the fibril morphologies were measured using Fiji³⁶. No clearly identifiable second morphology was found and the occurrence of the main morphology was estimated at over 95%.

Helical reconstruction. Helical reconstruction was performed using Relion 2.1 (ref. 37). The raw data were converted from TIFF to mrcs format using IMOD³⁸. Motion and gain corrections, as well as dose-weighting were performed using Motioncor2 (ref. 39). The non-dose-weighted micrographs were used to determine the defocus values with Gctf⁴⁰, whereas the dose-weighted micrographs were used for all subsequent steps. Fibrils were manually selected from 1,964 micrographs and segments were extracted with a box size of 300 pixels (312 Å) with an inter-box distance of 33.6 Å (~11%). Two rounds of reference-free 2D classification were performed with a regularization value of $T = 2$. Class averages which showed the helical repeat along the fibril axis were selected, whereas classes showing artifacts and noise were discarded, resulting in a selection of 101,319 particles. The selection was confirmed by manually arranging the class averages into a full 2D fibril side view. An initial model from a previous reconstruction was filtered to 60 Å resulting in a rod-like structure. This initial model was used as a reference to create a first 3D map using 3D classification with 553 particles picked from nine micrographs. The resulting map was used as a reference for 3D classification with six classes and $T = 3$ using the 101,319 particles obtained from the 2D classification selection. Three of the six classes showed a clear backbone and revealed the presence of two different conformations. These classes were selected (60,044 particles) and further rounds of 3D classification with four classes and step-wise increasing T -values from 3 to 20 were performed. Approximately 11,003 particles were selected for conformation A and 12,122 particles for conformation B. Final round of 3D classification with increasing T -values from 80 to 200, followed by 3D auto-refinement and post-processing with a soft-edged mask and an estimated sharpening B -factor of -33.1013 \AA^2 (conformation A) and -55.1958 \AA^2 (conformation B) for each of the two conformations led to the post-processed maps. Helical z -percentages used for 3D classification and 3D refinement varied between 0.1 and 0.3. The twist and rise were 1.11 and 4.79, respectively, for both conformations. These values agree with the measured crossover distances on the motion-corrected cryo-EM micrographs, and the corresponding power spectra. A left-handed twist was assumed. The resolution of each map was estimated from the value of the FSC curve for two independently refined half-maps at 0.143.

Model building and refinement. The model was manually built using the software Coot⁴¹. The process was initiated by tracing a poly-Ala chain along the 3D map. The alanine residues were mutated to the LC sequence as determined previously¹¹. The model obtained was subjected to manual, as well as automated refinement using phenix.real_space_refine^{42,43} tool, as implemented in phenix with non-

crystallographic symmetry and secondary structure restraints imposed. This process was repeated until a satisfactory map to model fit was obtained. This model was subsequently used to perform model based automated sharpening of the map (phenix.auto_sharpen^{42,43}). The sharpened map was used for improving the model further. The final refined model was evaluated for its quality using the MolProbity⁴⁴ validation report.

Sequence analysis. The amino acid sequence of the LC investigated in this study was taken from the gene bank entry KX290463 (ref. 10), which was obtained by cDNA sequencing of FOR005 patient material. The residue numbering throughout this article refers to the precursor LC sequence (GenBank ANN81988.1) from the patient, starting with the first residue (Ser1) after the signal sequence, which is cleaved off in the fibril protein. All mutations in this manuscript are represented in the direction GL to FOR005 fibril protein. The sequence elements were defined as follows. First, the patient cDNA was translated to the putative amino acid sequence of the fibril precursor protein. Then, the cDNA of the patient and the corresponding amino acid sequence were analyzed to determine the most probable GL segments, using the vbase2 (ref. 45) and BLAST/BLAT search tools⁴⁶ (<http://www.ensembl.org>). This analysis yielded several hits for possible GL segments. The cDNA sequences and corresponding amino acid sequences of these V, J, and C GL segments were retrieved from the vbase2 (ref. 45), ENSEMBL⁴⁷, and <http://www.imgt.org>⁴⁸ databases, and genetic distances to the patient sequence were calculated by maximum composite likelihood to confirm the most probable V, J, and C GL segments. Finally, the cDNA sequence of the patient and the corresponding amino acid sequence were aligned to these GL segments (fit > 80%), using the MEGA multiple sequencing alignment tool “MUSCLE”⁴⁹ and Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>)⁵⁰ to identify patient-specific mutations. The definitions of the CDRs were taken from a previous manuscript¹⁰.

LC aggregation propensities for every residue were calculated using the programs TANGO⁵¹, Foldamyloid⁵², Aggrescan⁵³, and PASTA 2.0 (ref. 54). For every prediction tool standard parameters were chosen. For Tango (Version 2.1), every amino acid which scores above 5% was counted as a hit. For Foldamyloid, the “triple hybrid” scale was used, meaning that five successive amino acids with a score above 21.4 were counted as hits. PASTA 2.0 calculated aggregation with 90% sensitivity and predicted hits if the energy cutoff fell below -2.8 PASTA Energy Units. For Aggrescan, all amino acids with values above -0.02 were counted as hot spots. In this study, an aggregation score of 0 means that none of these programs identified the corresponding residue as aggregation prone. An aggregation score of 4 means that all four programs identified the corresponding residue as prone to aggregation.

Protein structure representation. UCSF Chimera⁵⁵ was used for creating the images of the density maps and protein models. The structure of the refolded FOR005 fibril protein was obtained previously using protein X-ray crystallography and has the protein data bank (PDB) entry 5L6Q¹⁰. The native C₁ domain is the IGLC2 segment of PDB 4EOW⁵⁶, where the side chain of residue Val135 was not shown in order to represent the mutated residue Gly135 of the fibril protein. The following previously published coordinates were used in Fig. 5: 6IC3 and 6HUD.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The reconstructed cryo-EM maps were deposited in the Electron Microscopy Data Bank with accession codes EMD-11031 (structure A) and EMD-11030 (structure B). The coordinates of the fitted atomic model were deposited in the PDB under the accession codes 6Z10 (structure A) and 6Z11 (structure B). The cryo-EM data were deposited on EMPIAR with the accession code EMPIAR-10457. The data sets used during the current

study are available from the corresponding author on reasonable request. The unique materials used in the current study are available from the authors on reasonable request. Source data are provided with this paper.

Received: 5 August 2020; Accepted: 21 December 2020;

Published online: 08 February 2021

References

- Quock, T. P., Yan, T., Chang, E., Guthrie, S. & Broder, M. S. Epidemiology of AL amyloidosis: a real-world study using US claims data. *Blood Adv.* **2**, 1046–1053 (2018).
- Shimazaki, C. Current diagnosis and treatment of AL amyloidosis in Japan: a nationwide epidemiological survey. *Rinsho Ketsueki* **60.8**, 973–978 (2019).
- Wisniewski, B. et al. The epidemiology of amyloidosis in Queensland, Australia. *Br. J. Haematol.* **186.6**, 829–836 (2019).
- Merlini, G. et al. Systemic immunoglobulin light chain amyloidosis. *Nat. Rev. Dis. Prim.* **4**, 1–19 (2018).
- Blancas-Mejia, L. M. & Ramirez-Alvarado, M. Systemic amyloidoses. *Annu. Rev. Biochem.* **82**, 745–774 (2013).
- Perfetti, V. et al. The repertoire of lambda light chains causing predominant amyloid heart involvement and identification of a preferentially involved germline gene, IGLV1-44. *Blood* **119**, 144–150 (2012).
- Kourelis, T. V. et al. Clarifying immunoglobulin gene usage in systemic and localized immunoglobulin light-chain amyloidosis by mass spectrometry. *Blood* **129**, 299–306 (2017).
- Vaxman, I., Dispenzieri, A., Muchtar, E. & Gertz, M. New developments in diagnosis, risk assessment and management in systemic amyloidosis. *Blood Rev.* **40**, 100636 (2020).
- Bellotti, V., Mangione, P. & Merlini, G. Review: immunoglobulin light chain amyloidosis—the archetype of structural and pathogenic variability. *J. Struct. Biol.* **130**, 280–289 (2000).
- Annamalai, K. et al. Common fibril structures imply systemically conserved protein misfolding pathways in vivo. *Angew. Chem. Int. Ed.* **56.26**, 7510–7514 (2017).
- Annamalai, K. et al. Polymorphism of amyloid fibrils in vivo. *Angew. Chem. Int. Ed.* **55.15**, 4822–4825 (2016).
- Glenner, G. G., Harbaugh, J., Ohms, J. I., Harada, M. & Cuatrecasas, P. An amyloid protein: the amino-terminal variable fragment of an immunoglobulin light chain. *Biochem. Biophys. Res. Commun.* **41**, 1287–1289 (1970).
- Radamaker, L. et al. Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis. *Nat. Commun.* **10.1**, 1–8 (2019).
- Swuec, P. et al. Cryo-EM structure of cardiac amyloid fibrils from an immunoglobulin light chain AL amyloidosis patient. *Nat. Commun.* **10.1**, 1–9 (2019).
- Boyer, D. R. et al. Structures of fibrils formed by α -synuclein hereditary disease mutant H50Q reveal new polymorphs. *Nat. Struct. Mol. Biol.* **26**, 1044–1052 (2019).
- Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat. Struct. Mol. Biol.* **26**, 619–627 (2019).
- Blancas-Mejia, L. M. et al. Immunoglobulin light chain amyloid aggregation. *Chem. Commun.* **54.76**, 10664–10674 (2018).
- Kazman, P. et al. Fatal amyloid formation in a patient’s antibody light chain is caused by a single point mutation. *Elife* **9**, e52300 (2020).
- Wang, J. & Moore, P. B. On the interpretation of electron microscopic maps of biological macromolecules. *Protein Sci.* **26.1**, 122–129 (2017).
- Hora, M. et al. MAK33 antibody light chain amyloid fibrils are similar to oligomeric precursors. *PLoS ONE* **12.7**, e0181799 (2017).
- Piehl, D. W., Blancas-Mejia, L. M., Ramirez-Alvarado, M. & Rienstra, C. M. Solid-state NMR chemical shift assignments for AL-09 V L immunoglobulin light chain fibrils. *Biomol. NMR Assign.* **11**, 45–50 (2017).
- Pradhan, T. et al. Seeded fibrils of the germline variant of human λ -III immunoglobulin light chain FOR005 have a similar core as patient fibrils with reduced stability. *J. Biol. Chem.* **295**, 18474–18484 (2020).
- Lecoq, L. et al. A substantial structural conversion of the native monomer leads to in-register parallel amyloid fibril formation in light-chain amyloidosis. *Chembiochem* **20.8**, 1027–1031 (2019).
- Ferrone, F. A., Hofrichter, J. & Eaton, W. A. Kinetics of sickle hemoglobin polymerization: I. Studies using temperature-jump and laser photolysis techniques. *J. Mol. Biol.* **183**, 591–610 (1985).
- Crowther, R. A. Straight and paired helical filaments in Alzheimer disease have a common structural unit. *Proc. Natl Acad. Sci. USA* **88**, 2288–2292 (1991).
- Makarava, N., Ostapchenko, V. G., Savtchenko, R. & Baskakov, I. V. Conformational switching within individual amyloid fibrils. *J. Biol. Chem.* **284**, 14386–14395 (2009).
- Zhang, W. et al. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer’s and Pick’s diseases. *Elife* **8**, e43584 (2019).
- Kollmer, M. et al. Cryo-EM structure and polymorphism of A β amyloid fibrils purified from Alzheimer’s brain tissue. *Nat. Commun.* **10.1**, 1–8 (2019).
- Gallardo, R., Ranson, N. E. & Radford, S. E. Amyloid structures: much more than just a cross- β fold. *Curr. Opin. Struct. Biol.* **60**, 7–16 (2020).
- Ferkinghoff-Borg, J. et al. Stop-and-go kinetics in amyloid fibrillation. *Phys. Rev. E* **82**, 010901 (2010).
- Ndlovu, H., Ashcroft, A. E., Radford, S. E. & Harris, S. A. Molecular dynamics simulations of mechanical failure in polymorphic arrangements of amyloid fibrils containing structural defects. *Beilstein J. Nanotechnol.* **4**, 429–440 (2013).
- Knowles, T. P., Smith, J. F., Craig, A., Dobson, C. M. & Welland, M. E. Spatial persistence of angular correlations in amyloid fibrils. *Phys. Rev. Lett.* **96**, 238301 (2006).
- Knowles, T. P., Smith, J. F., Devlin, G. L., Dobson, C. M. & Welland, M. E. Analysis of structural order in amyloid fibrils. *Nanotechnology* **18**, 044031 (2006).
- Andersen, C. B. et al. Branching in amyloid fibril growth. *Biophys. J.* **96**, 1529–1536 (2009).
- Gao, X. et al. Human Hsp70 disaggregase reverses Parkinson’s-linked α -synuclein amyloid fibrils. *Mol. Cell* **59**, 781–793 (2015).
- Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
- He, S. & Scheres, S. H. W. Helical reconstruction in RELION. *J. Struct. Biol.* **198**, 163–176 (2017).
- Kremer, J. R., Mastrorade, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
- Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
- Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
- Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D Struct. Biol.* **74**, 814–840 (2018).
- Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
- Williams, C. J. et al. MolProbity: more and better reference data for improved all atom structure validation. *Protein Sci.* **27.1**, 293–315 (2018).
- Retter, I., Althaus, H. H., Munch, R. & Muller, W. VBASE2, an integrative V gene database. *Nucleic Acids Res.* **33**, D671–D674 (2005).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- Hunt, S. E. et al. Ensembl variation resources. *Database* **2018**, bay119 (2018).
- Ruiz, M. et al. IMGIT, the international ImMunoGeneTics database. *Nucleic Acids Res.* **28**, 219–221 (2000).
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A. & Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
- Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
- Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. & Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* **22**, 1302–1306 (2004).
- Garbuzynskiy, S. O., Lobanov, M. Y. & Galzitskaya, O. V. FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics* **26**, 326–332 (2010).
- Conchillo-Solé, O. et al. AGGRESCAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC Bioinformatics* **8**, 65 (2007).
- Walsh, I., Seno, F., Tosatto, S. C. & Trovato, A. PASTA 2.0: an improved server for protein aggregation prediction. *Nucleic Acids Res.* **42**, W301–W307 (2014).
- Petersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Blech, M. et al. Molecular structure of human GM-CSF in complex with a disease-associated anti-human GM-CSF autoantibody and its potential biological implications. *Biochem. J.* **447.2**, 205–215 (2012).

Acknowledgements

We would like to thank the Deutsche Forschungsgemeinschaft for funding of the Research Unit FOR 2969, projects FA 456/27, HA 7138/3, HE 8472/1-1, HU 2400/1-1, and SCHO 1364/2-1. We are grateful to Prof. Dr. Bernd Reif (Technical University Munich) for helpful discussions and to the students who participated in the practical course “Protein Biochemistry and Structural Biology” in the winter semester 2019, for assisting in the manual picking of the fibril segments. All cryo-EM data were collected at the European Molecular Biology Laboratory, Heidelberg (Germany), funded by iNEXT (Horizon 2020, European Union).

Author contributions

L.R. and J.B. carried out experiments. L.R., J.B., S.H., C.H., A.B., M.S., and M.F. analyzed data. U.H. and S.S. contributed tools and reagents. M.F. designed research. L.R. and M.F. wrote the paper.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-21126-2>.

Correspondence and requests for materials should be addressed to M.F.

Peer review information *Nature Communications* thanks Luis del Pozo-Yauner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021