


# Cryo-EM structure of a lysozyme-derived amyloid fibril from hereditary amyloidosis

Received: 20 November 2023

Accepted: 1 November 2024

Published online: 07 November 2024

 Check for updates

Sara Karimi-Farsijani<sup>1</sup>✉, Kartikay Sharma<sup>1</sup>, Marijana Ugrina<sup>2</sup>, Lukas Kuhn<sup>1</sup>, Peter Benedikt Pfeiffer<sup>1</sup>, Christian Haupt<sup>1</sup>, Sebastian Wiese<sup>3</sup>, Ute Hegenbart<sup>4</sup>, Stefan O. Schönland<sup>4</sup>, Nadine Schwierz<sup>2</sup>, Matthias Schmidt<sup>1</sup> & Marcus Fändrich<sup>1</sup>

Systemic ALys amyloidosis is a debilitating protein misfolding disease that arises from the formation of amyloid fibrils from C-type lysozyme. We here present a 2.8 Å cryo-electron microscopy structure of an amyloid fibril, which was isolated from the abdominal fat tissue of a patient who expressed the D87G variant of human lysozyme. We find that the fibril possesses a stable core that is formed by all 130 residues of the fibril precursor protein. There are four disulfide bonds in each fibril protein that connect the same residues as in the globularly folded protein. As the conformation of lysozyme in the fibril is otherwise fundamentally different from native lysozyme, our data provide a structural rationale for the need of protein unfolding in the development of systemic ALys amyloidosis.

Systemic ALys amyloidosis is a disease that is based on the misfolding of human C-type lysozyme<sup>1</sup>. The disease affects individuals of different age groups, and patients show a wide range of symptoms from weight loss and gastro-intestinal bleeding, renal dysfunction, and sicca syndrome<sup>2</sup>. It is a hereditary disease that arises because of mutational changes in the lysozyme gene that encodes for lysozyme<sup>3,4</sup>. The disease is relatively rare even within the group of systemic amyloidosis. A study from 2020 reports less than 1 % of the patients with non-neuronal amyloidosis to be affected by systemic ALys amyloidosis<sup>5</sup>. The disease is monogenetic and depends invariably on mutational changes in lysozyme. Different lysozyme mutations have been described to underlie the disease in different patients, including the single-site mutational changes Y54N<sup>6</sup>, I56T<sup>3</sup>, F57I<sup>7</sup>, W64R<sup>8</sup>, D67G<sup>9</sup>, D67H<sup>3</sup>, L84S<sup>10</sup>, D87G<sup>2</sup> and the double mutations T70N/W112R<sup>11</sup> and F3L/T70N<sup>12</sup>. Hence, the occurrence of any of these pathogenic mutations is small in the human population.

Lysozyme is an important component of the innate immune system that helps to neutralize pathogenic microbes. The protein is one of the best studied proteins in biochemistry and a textbook example of enzymatic function<sup>13</sup>. Natively folded lysozyme possesses a compact, globular structure and has been the archetype example of a protein

from the  $\alpha + \beta$  class of protein folds. The protein contains two domains, an  $\alpha$ -domain with mainly  $\alpha$ -helical conformation and a  $\beta$ -domain containing mainly  $\beta$ -sheet structure<sup>13</sup>. Its native folding and misfolding characteristics have been studied extensively by several laboratories including the laboratory of the late Chris Dobson<sup>14</sup>.

Biophysical analysis of the effect of disease-associated mutations revealed a mechanism, in which the native structure of lysozyme becomes destabilized by the mutations so that the mutant protein is prone to unfold into highly aggregation competent protein states<sup>15,16</sup>. While a native state-destabilizing activity of hereditary mutations has been well-documented in ALys and other forms of systemic amyloidosis<sup>17,18</sup>, it is not known whether the mutations may contribute through further mechanisms to pathogenesis. For example, it may be possible that the mutations could modify the misfolding pathway and promote the formation of a specific fibril conformation. In absence of coordinate files showing the three-dimensional (3D) structures of the patient-derived ALys fibrils, it is difficult to further examine this possibility.

Here, we present the cryo-electron microscopy (cryo-EM) structure of an ALys amyloid fibril that was isolated from the amyloidotic tissue of a patient with systemic ALys amyloidosis. The patient suffered

<sup>1</sup>Institute of Protein Biochemistry, Ulm University, Ulm, Germany. <sup>2</sup>Institute of Physics, University of Augsburg, Augsburg, Germany. <sup>3</sup>Core Unit Mass Spectrometry and Proteomics, Medical Faculty, Ulm University, Ulm, Germany. <sup>4</sup>Medical Department V, Amyloidosis Center, Heidelberg University Hospital, Heidelberg, Germany. ✉ e-mail: [sara.karimi-farsijani@uni-ulm.de](mailto:sara.karimi-farsijani@uni-ulm.de)

from hereditary amyloidosis and expressed the LysD87G<sup>2</sup>. Analysis of the isolated fibrils with cryo-EM showed that they belong to a single fibril morphology and contain a single stack of fibril proteins. All residues of lysozyme contribute to formation of the stable fibril structure, which lack conformationally disordered fibril protein segments. The conformation of the fibril protein differs fundamentally from natively folded lysozyme, demonstrating the need of protein unfolding for the formation of the pathogenic fibril conformation.

## Results

### Isolation and characterization of LysD87G amyloid fibrils

ALys amyloid fibrils were isolated from the abdominal fat tissue of a previously described patient with systemic ALys amyloidosis arising from the D87G variant of human C-type lysozyme<sup>2</sup>. The patient was diagnosed at the age of 33 years and showed amyloid deposits in kidneys and abdominal fat tissue. Using a previously described isolation method<sup>19</sup>, we isolated ALys fibrils from fat tissue to facilitate biophysical analysis. The fibrils were relatively pure, as indicated by a single protein band on a denaturing protein electrophoresis gel (Supplementary Fig. 1). The protein migrated at ~14 kDa, similar to previously described ALys amyloid fibril proteins<sup>2</sup>. Also, mass spectrometry shows that the fibrils consist mainly of full-length D87G lysozyme (Fig. Si 2). Transmission electron microscopy (TEM) showed well-resolved cross-overs in the fibril structure (Supplementary Fig. 3a). Scanning electron microscopy (SEM) and platinum side-shadowing additionally demonstrated the left-hand twist of the fibrils (Supplementary Fig. 3b).

### Reconstruction of the cryo-EM structure of the fibril

The fibrils were cryo-frozen and imaged with cryo-EM at 300 kV (Fig. 1a). There was an essentially monomorphic distribution of fibrils with no discernible polymorphism. The fibril width was determined at  $106 \pm 9 \text{ \AA}$ , while the fibril crossovers occurred at a regular distance of  $584 \pm 58 \text{ \AA}$  ( $n = 30$ ; Supplementary Fig. 4). The 2D (Supplementary Fig. 5a) and 3D classes (Supplementary Fig. 5b) also showed no particle heterogeneity. The three-dimensional (3D) map was reconstructed at a final resolution of  $2.8 \text{ \AA}$ , based on the 0.143 Fourier-shell correlation (FSC) criterion (Supplementary Fig. 6b). The atomic model, which we derived from the 3D map (Fig. 1a), depicts a polar fibril with C1 helical symmetry. The fibril consists of a single stack of fibril proteins; this is, the fibril contains a single protofilament (Fig. 1b). All 130 amino acid residues of native lysozyme could be fitted to the 3D map and formed

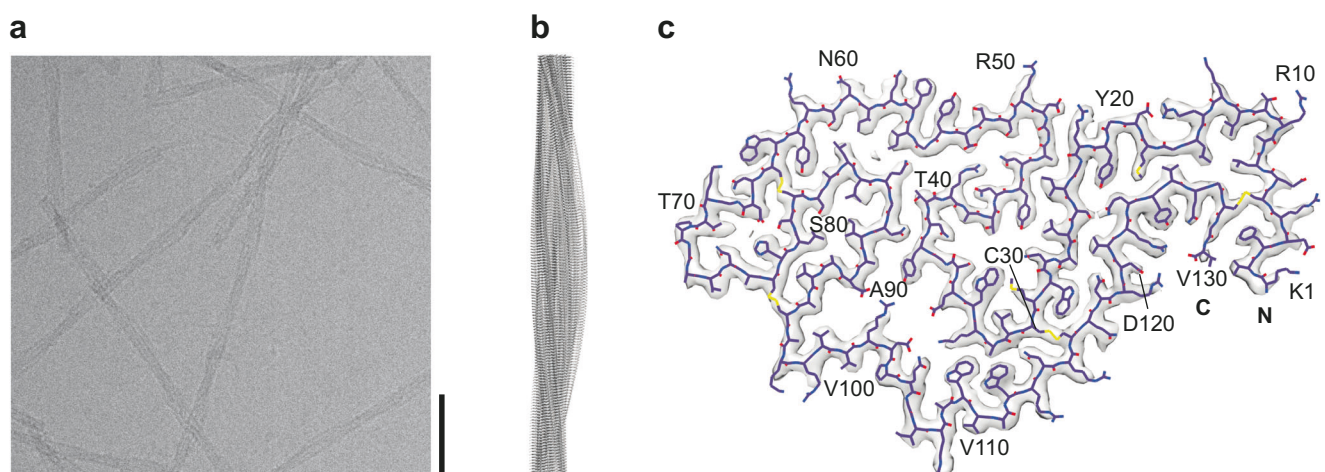
the fibril core (Fig. 1c). The absence of conformationally disordered segments in this fibril contrasts to the vast majority of fibril structures from systemic amyloidosis, which typically contain conformationally disordered segments, irrespective of whether the fibrils were purified from human or animal tissue<sup>20–29</sup>.

### Fibril protein fold and enclosed structural cavities

The fibril protein shows an all- $\beta$  fold that lacks elements of  $\alpha$ -helical conformation (Fig. 2a,b). This conformation differs starkly from the structure of natively folded lysozyme, which contains elements of both  $\alpha$ -helical and  $\beta$ -sheet structure (Fig. 2a,c). The fibril encompasses fifteen cross  $\beta$ -sheets ( $\beta 1$ – $\beta 15$ ) with strands varying from two to nine amino acid residues (Fig. 2a, b). The relative orientation of adjacent  $\beta$ -strands is uniformly parallel in the direction of the fibril main axis (Fig. 2b). There are three internal cavities that are termed here A, B and C (Supplementary Fig. 7a). Cavity A is surrounded by hydrophobic chemical groups and harbors a small density feature that could not be assigned to lysozyme (Supplementary Fig. 7a). This density feature is continuous in the direction of the fibril z-axis (Supplementary Fig. 7b) and could represent a non-polar molecular inclusion, such as a lipid, that runs across multiple layers of fibril proteins. Similar types of inclusions have been reported for several other ex vivo amyloid fibril structures derived for example from immunoglobulin light chain<sup>21</sup> or tau protein<sup>30</sup>. Cavities B and C are lined with polar amino acid residues and probably water-filled (Supplementary Fig. 7a). A computational evaluation of the fibril structure with the Amber function in Chimera<sup>31</sup> confirms this view and shows that the cavities are large enough to accommodate water molecules (Supplementary Fig. 8).

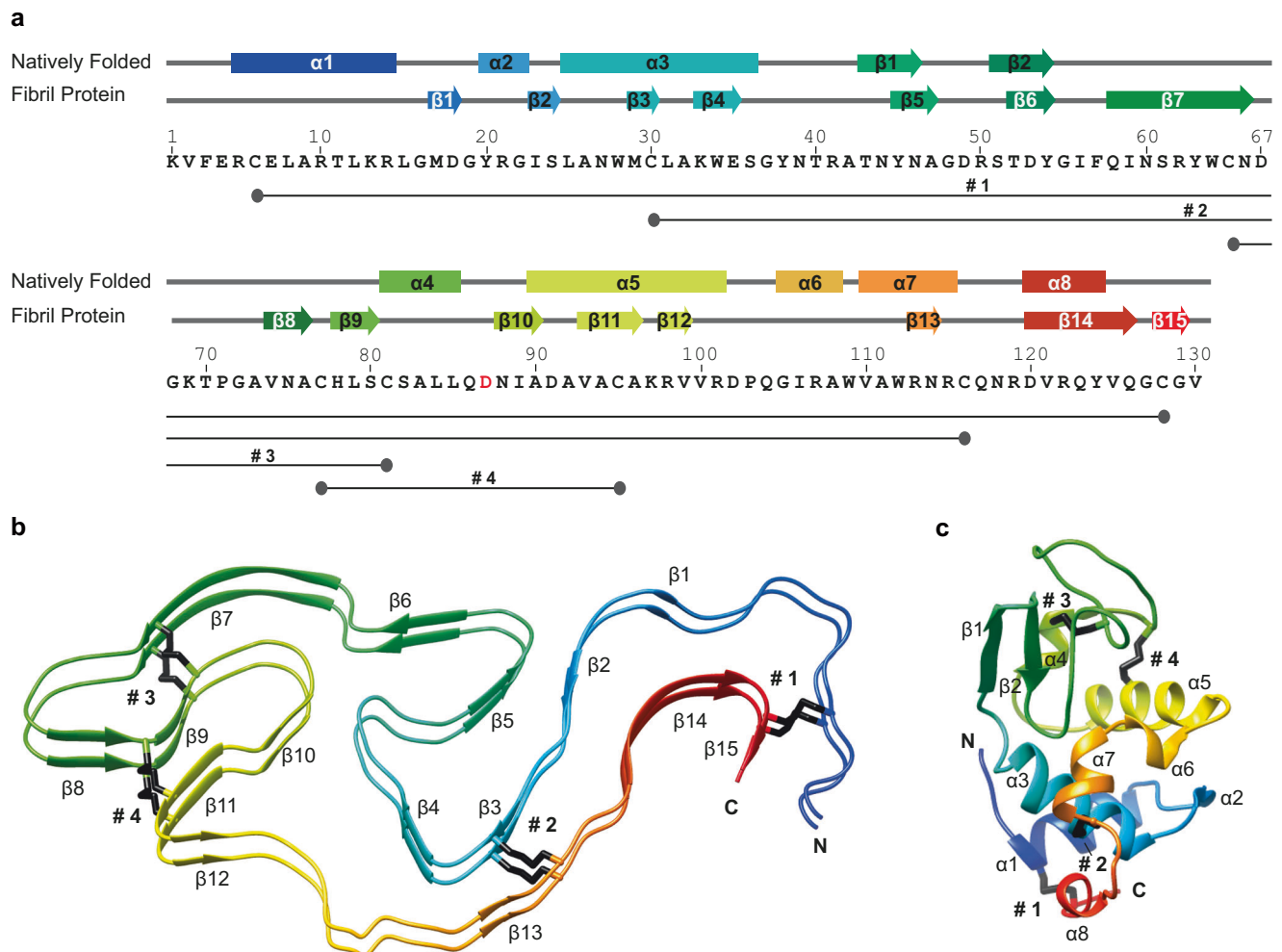
### Molecular interactions stabilizing the fibril structure

The fibril protein fold is stabilized by four intramolecular disulfide bonds that occur at residues Cys6-Cys128, Cys30-Cys116, Cys65-Cys81 and Cys77-Cys95 (Fig. 2b). These disulfides correspond to the four disulfide bonds in natively folded lysozyme (Fig. 2c). The disulfides impose severe, covalent constraints on the structure that help to prevent the unfolding of the fibril. There are a range of non-covalent interactions, such as buried clusters of hydrophobic residues - one such cluster is formed by residues Leu8, Ala9, Leu12, Leu15 and Val125 (Supplementary Fig. 9) - or buried pairs of residues with opposing charge, e.g. Asp67 and His78 (Supplementary Fig. 10a). Some non-covalent interactions run across different molecular layers of fibril proteins, such as the interaction between Arg41 from layer  $i$  and Asp53



**Fig. 1 | Cryo-EM structure of an ex vivo ALys amyloid fibril. a** Representative cryo-EM image of a total of 2013 micrographs of the isolated fibrils ( $n = 1$ , biological sample). Scale bar: 50 nm. **b** Side-view of the reconstructed 3D map. **c** Cross-section

of the 3D map (gray) overlaid with the atomic model (purple). The protein N and C-terminal ends as well as every tenth amino acid are labeled in the figure.



**Fig. 2 | Secondary structure of lysozyme fibrils and the natively folded protein.**

**a** Sequence of human LysD87G. The mutant residue is marked in red. Above the sequence is a schematic representation of the secondary structural elements of the fibril protein and natively folded lysozyme. Arrows:  $\beta$ -sheets, boxes:  $\alpha$ -helices. The four disulfide bonds (#1 to #4) are drawn as bars below the sequence. **b** Ribbon

diagram of two molecular layers of the fibril. **c** Ribbon diagram of native human WT lysozyme (protein data bank: 1JSF<sup>62</sup>), scaled equally to the fibril protein. N- and C-terminal ends of the protein, secondary structural elements and disulfide bonds are labeled.

from layer  $i - 1$  (Supplementary Fig. 10b). Cross-layer interactions occur because the fibril protein extends by 19 Å in the direction of the fibril z-axis, as measured as the distance between the highest and lowest C $\alpha$  atom along the fibril z-axis (Supplementary Fig. 11a). Evaluation of the intermolecular interfaces of the fibril with the program PDBePISA<sup>32</sup> shows that each layer is in contact with nine other layers or fibril protein molecules (Supplementary Fig. 11b). The associated non-covalent interactions across the molecular layers of the fibril help to sterically interlock its architecture.

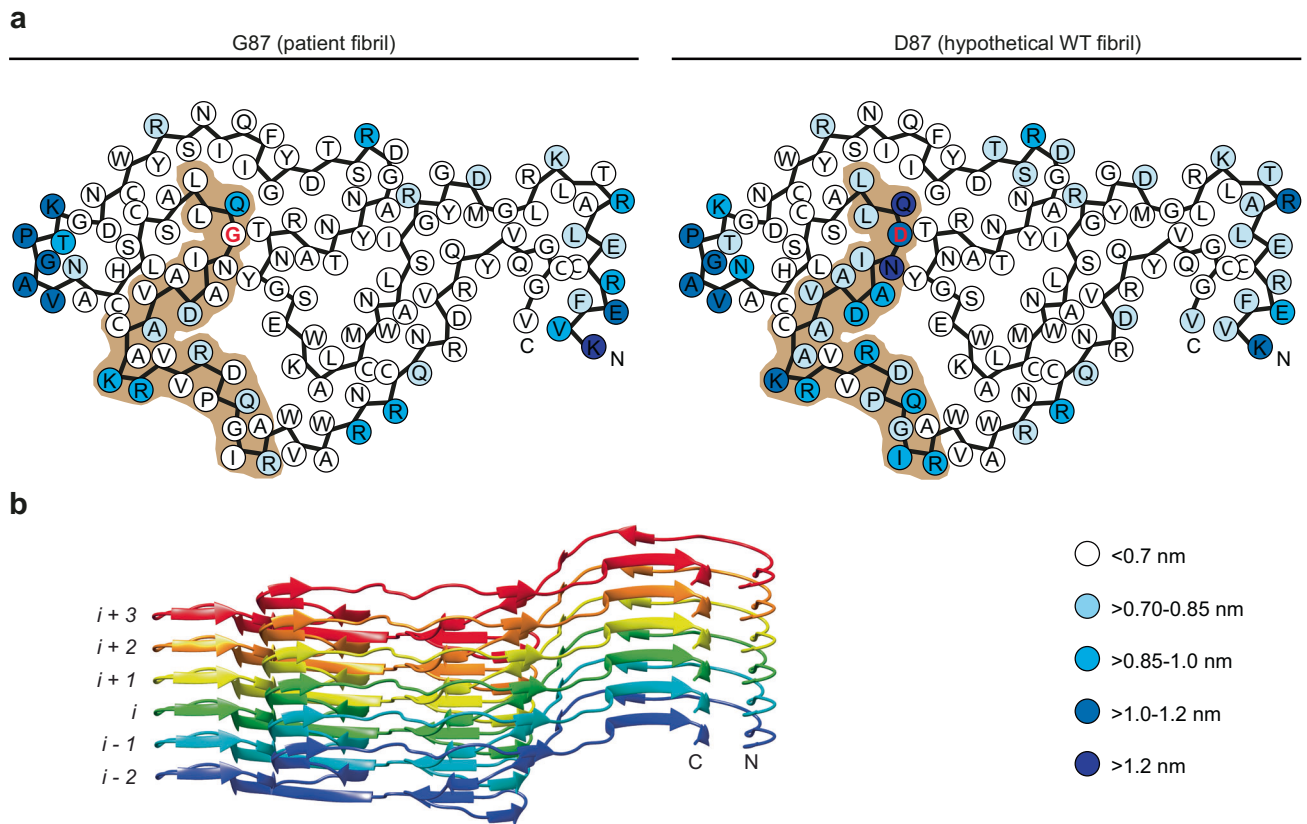
#### Location of aggregation-prone segments in the fibril structure

Analysis of the fibril protein sequence with computational methods shows that the most aggregation-prone segments do not always match the segments that form the fibril cross- $\beta$  structure (Supplementary Fig. 12a). This observation indicates that the intrinsic aggregation propensity of the protein sequence does not suffice to determine the observed fibril structure, although it may help to kinetically promote fibril formation or to form a certain fibril nucleus. The most aggregation-prone segments occur in the vicinity of the disulfide bonds Cys30-Cys116, Cys65-Cys81 and Cys77-Cys95. This association suggests that the three disulfide bonds occur at a site that may be important to trigger the aggregation process; and interestingly, many

ALys amyloidosis-associated mutations are also located within these segments (Supplementary Fig. 12b).

#### Importance of the D87G mutation for the fibril

A previous analysis of the effect of the D87G mutation on the native protein conformation could not detect any strong destabilizing effect of this mutation on the globular lysozyme fold<sup>2</sup>. We thus analyzed the possible effect of this mutation on the fibril structure. The mutant residue (position 87) occurs within a tightly packed, hydrophobic  $\beta$ -arc of the fibril that is formed by residues Cys77-Cys95 (Supplementary Fig. 9). Residue 87 is a small glycine in the patient but a bulky and charged aspartate residue in wildtype (WT) lysozyme, suggesting that the WT protein is not compatible with the observed fibril structure. To test this possibility, we analyzed the experimentally determined fibril structure and a computationally generated homology model, which contains an aspartate residue at position 87 as in WT lysozyme, with different methods. Using programs such as Foldx<sup>33</sup>, Amyloid Illustrator<sup>34</sup> and PDBePISA<sup>32</sup>, there were some differences in the computed energy values of the patient fibril and the hypothetical WT counterpart (Fig. Si 13). However, the different programs varied considerably concerning the absolute height of the calculated stabilities, the stability contributions included in each program and the relative



**Fig. 3 | The WT D87 residue leads to a local loosening of the fibril structure.** **a** Schematic representation of the fibril protein with each residue being color-coded by the maximum distance between layers  $i + 2$  and  $i + 3$  adopted in the course of the MD simulation (blue scale and indicated in the panel). Brown: main region (Leu84 to Arg107) affected by the mutation. The figure was constructed based on

the average values plotted in Fig. Si 13. Left: patient protein fibril with a glycine at position 87; right: hypothetical WT fibril structure, containing a D87 residue.

**b** Ribbon diagram of a six-layer stack of the patient fibril indicating the definition of layers  $i - 2$  to  $i + 3$ .

stability order of the two fibril forms (Supplementary Fig. 13). Therefore, we were not able to discern any consistent trend whether or how the WT D87 residue influences the fibril structure.

Molecular dynamics (MD) simulations carried out with an explicit solvent model showed only small differences between the WT and the patient fibril. The most profound difference was seen for the molecules at one of the terminal ends of the fibril and in a segment of residues in the vicinity of the mutant position. This difference could be revealed by analyzing the distances between corresponding residues in successive layers of the fibril adopted in the course of the MD simulation. In a sequence segment extending from residue Leu84 to Arg107 of the WT fibril, we find many residues to show a distance of more than 0.7 nm to the corresponding residues in the adjacent fibril layer (Fig. 3). That is, the WT fibril shows a local loosening of the structure, which is not seen in the patient fibril, that contains a G87 residue, and in the remaining layers of the WT fibril (Supplementary Fig. 14).

#### Location of disease-associated mutations in the fibril structure

Analysis of the various known ALys amyloidosis-associated mutations does not indicate any obvious preference of the mutations for a specific secondary structural element of the native protein or fibril structure (Supplementary Fig. 15a–c). Several mutations seem compatible with the presently described fibril. Examples are the three mutations Y54N, F57I and W64R, which occur at solvent-exposed positions of the fibril (Supplementary Fig. 15c), or the relatively conservative mutations F3L and T70N that affect buried positions of the fibril. Other mutations seem more difficult to reconcile with our fibril. Examples are the mutations I56T and L84S that occur at buried sites of the fibril where they surround the hydrophobic cavity A. Both

mutations replace a hydrophobic surface residue with a polar one (Supplementary Fig. 15c). The remaining three mutations (D67G, D67H and W112R) are buried and alter the chemical properties of the residue quite substantially. Plotted onto our fibril structure, they disrupt a buried ion pair or insert an uncompensated charge into a hydrophobic pocket. All of these mutations occur in the context of lysozyme molecules containing a D87 residue, which could be at least locally destabilizing to the fibril structure, as indicated above (Fig. 3, Supplementary Fig. 13,14).

#### Discussion

Lysozyme is a vital component of the human innate immune system, although it may also form the basis of a debilitating protein misfolding disease. Using cryo-EM, we here obtained the structure of a fibril, that underlies the development of amyloidosis in one such patient. The fibril sample contained one dominant ALys fibril morphology that consisted of a single stack of fibril proteins. It was purified from abdominal fat tissue, which is an important amyloid deposition site within the human body that is even used for diagnostic purposes<sup>5</sup>. Previous studies established that the fibril morphology is consistent in fat tissue and other deposition sites within the human body<sup>35</sup>. Indeed, there is evidence from different amyloid diseases and from animals and humans that the fibril structure is consistent at different deposition sites of the same patient or animal<sup>29,35</sup>, all suggesting that the observed fibril structure is representative for the fibrils in this patient.

So far, it is not clear whether the observed fibril structure occurs also in ALys patients, who are associated by other lysozyme mutations. High-resolution structures of such fibrils are missing, and the available data provide a mixed picture. The absence of WT protein in the fibril

(Supplementary Fig. 2) suggests that lysozyme variants containing a D87 residue are not compatible with the fibril structure. Yet, there was no strong evidence from our theoretic examinations that a buried aspartate at position 87 cannot be tolerated by our fibril structure (Fig. 3, Supplementary Fig. 13, 14). Furthermore, several previous fibril structures show that charged residues that are not part of ion pairs can exist in amyloid fibril structures, such as the protein data bank (PDB) entries 8CG3<sup>36</sup>, 7SAQ<sup>37</sup> and 6CU8<sup>38</sup>.

Analysis of the nature of the lysozyme mutations in other ALys patients suggests that some are compatible with our fibril structure, while others may be unfavorable (Fig. Si 15). Some patients were reported to have fibril proteins that consist of full-length lysozyme<sup>3</sup>, similar to the D87G lysozyme fibrils in our study. As the length of the fibril protein and its fragmentation is indicative of the fibril morphology<sup>26,39</sup>, it is possible that the respective patients are affected by our fibril morphology as well. Still, other fibril structures than the ones reported here may be involved in systemic ALys amyloidosis, as other ALys patients have fibrils with fragmented lysozyme<sup>11</sup>, suggesting that their fibrils are differently structured from ours.

The strong predominance of a single fibril morphology implies that this structure is particularly stable (biologically and/or thermodynamically) compared with other conformational states that may be formed from lysozyme. Indeed, several previous studies have suggested that pathogenic amyloid fibril structures are relatively stable in terms of their thermodynamic stability and/or resistance to proteolytic degradation<sup>40,41</sup>; and it is because of their biological stability that only specific amyloid fibril structures are able to accumulate and to cause damage inside the body. This selection of specific fibril morphologies because of their high proteolytic stability in vivo was termed 'proteolytic selection mechanism'<sup>22</sup>.

The stability of the ALys fibril arises, on the one hand, from a number of non-covalent interactions that occur either within or across the molecular layers of the fibril (Fig. Si 10). On the other hand, there are four covalent cross-links provided by the protein's disulfide bonds that significantly constrain the fibril structure (Fig. 2b). These four disulfides connect the same eight residues that are also cross-linked in the natively folded structure of lysozyme (Fig. 2c), indicating that misfolding starts from a native protein conformation, which then misfolds into a specific, and pathogenic amyloid fibril state. The stark differences between the native protein and the fibril protein conformation demonstrate the need of unfolding of the native conformation for amyloid fibril formation. Our structural data thus explain the previous observation that the unfolding of lysozyme, in particular by mutational changes, is crucial to the development of systemic ALys amyloidosis<sup>15,16</sup>.

The model for the transition of native lysozyme into a pathogenic amyloid fibril structure that emerges from our data emphasizes the need of unfolding for fibril formation. The predominance of a single fibril morphology in our samples further indicates the involvement of one or several mechanisms that prevent the formation or accumulation of other fibril morphologies inside the body. That is, it is either unfavorable for D87G lysozyme to form other fibril structures or it forms alternative fibril structures that are not stable in a biological environment, preventing them from being accumulated and becoming pathogenic. These conclusions are consistent with observations that amyloidosis is associated with fibril morphologies<sup>41</sup> of high proteolytic stability and the idea of pathogenic amyloid fibrils being proteolytically selected<sup>22,41</sup>. The great importance of the relative stabilities of the involved protein states further supports the idea that stabilizing the native state could represent a promising target to block the first step in the pathway and to ultimately combat the disease in humans. Indeed, this strategy has been employed with great success for the treatment of patients with systemic ATTR amyloidosis<sup>42</sup>, a disease caused by the misfolding of transthyretin. Transthyretin is lipid-transporter protein of the serum that naturally binds the thyroid hormone thyroxine; and

the thyroxine binding site was used to design a complementary ligand that binds to the native protein and prevents its unfolding and further misfolding into amyloid fibrils. Since lysozyme also possesses a ligand binding cavity, that naturally binds bacterial peptidoglycans, it may be possible that this site may also be used to develop complementary ligands that stabilize the native protein fold and ultimately help to alleviate the problems associated with lysozyme misfolding inside the human body.

## Methods

### Ethical statement

The patient's written informed consent was obtained before the biopsy was taken. The collection and use of methods were approved by the ethics committees of the University of Heidelberg (123/2006) and the University of Ulm (203/18).

### Patient description

The fibrils used in this study were extracted from the subcutaneous abdominal fat tissue of a female patient in her 50s suffering from systemic D87G ALys amyloidosis, confirmed by kidney biopsy. The abdominal fat tissue was removed from the patient with a 16-gauge needle and stored at  $-80^{\circ}\text{C}$  until fibril purification. The patient presented with a dominant renal involvement, but also with gastrointestinal symptoms and polyneuropathy<sup>2</sup>. The D87G mutation was also detected in the child of the patient, who is also suffering from amyloidosis. Later in life, the patient reached end-stage renal failure and was listed for a kidney transplantation.

### Fibril extraction from abdominal fat tissue

ALys amyloid fibrils were extracted from  $\sim 125$  mg amyloid-laden abdominal fat tissue, following a previously described method with some modifications<sup>19</sup>. In brief, the tissue was placed in 0.25 mL Tris-(hydroxymethyl)-aminomethane (Tris) calcium-buffer (TCB), pH 8.0, which consisted of 20 mM Tris, 138 mM NaCl, 2 mM  $\text{CaCl}_2$  and 0.1% (w/v)  $\text{NaN}_3$ . The sample was vortexed and spun down for 5 min at  $3100 \times g$  and  $4^{\circ}\text{C}$ . The washing step was repeated eight more times and the supernatants from each step were retained. We then dissolved one tablet of ethylenediamine tetra-acetic acid (EDTA)-free protease inhibitor (Roche) in 7 mL TCB and mixed 0.25 mL of this solution with 1.25 mg *Clostridium histolyticum* collagenase (Sigma-Aldrich) to obtain a final concentration of 5 mg/mL. The pellet from the last washing step was resuspended in 0.25 mL of the freshly prepared collagenase solution in TCB and incubated overnight at  $37^{\circ}\text{C}$ . In this incubation step the sample was constantly agitated at 700 rpm in an IKA MTS 2/4 digital table shaker. The sample was spun down for 30 min at  $3100 \times g$  and the top fat layer was removed. The pellet was washed three times in 125  $\mu\text{L}$  Tris-EDTA-buffer (TEB, 20 mM Tris, 140 mM NaCl, 10 mM EDTA, 0.1 % (w/v)  $\text{NaN}_3$ , pH 8). The pellet obtained after the last TEB washing step was resuspended in 125  $\mu\text{L}$  ice-cold water and centrifuged. The fibril-containing supernatant was retained and the water extraction was repeated eight more times. The supernatants from all water extraction steps were analyzed by gel electrophoresis or electron microscopy for the presence of fibrils. The fractions containing most fibrils were retained and used for further analysis.

### Denaturing protein gel electrophoresis

A 6  $\mu\text{L}$  aliquot of the respective protein sample was mixed with 1.2  $\mu\text{L}$  of  $10 \times$  NuPAGE reducing agent (Thermo Fisher Scientific), 3  $\mu\text{L}$  of  $4 \times$  NuPAGE LDS sample buffer (Thermo Fisher Scientific) and 1.8  $\mu\text{L}$  water. The mixture was heated at  $95^{\circ}\text{C}$  for 10 min and loaded onto a 4–12% NuPAGE Bis-Tris gel (Thermo Fisher Scientific), which was run in NuPAGE MES LDS running buffer (Thermo Fisher Scientific) as recommended by the vendor. BlueEasy prestained protein ladder (Genetics) was used as a size marker. The gel was stained in a solution containing 30 % (v/v) ethanol, 10 % (v/v) acetic acid and 0.25 % (w/v)

Coomassie brilliant blue and destained with a solution containing 20 % (v/v) ethanol and 10 % (v/v) acetic acid.

### Mass analysis by mass spectrometry

A 100  $\mu$ L aliquot of the fibril sample was lyophilized and dissolved in 4  $\mu$ L 6 M guanidinium hydrochloride, 50 mM Tris, pH 8.0. After incubation at 4 °C over night, 11  $\mu$ L of 0.1 % (v/v) trifluoroacetic acid (TFA) was added to the sample and analyzed by a U3000 RSLCnano high performance liquid chromatography (HPLC) system (Thermo Fisher Scientific). The sample was injected onto an Acclaim PepMap 100 pre-column (0.3 mm $\times$ 5 mm) (Thermo Fisher Scientific) equilibrated with 0.1 % (v/v) TFA. In order to remove contaminants, the sample was washed with 0.1% (v/v) TFA for 5 minutes at a flow rate of 30  $\mu$ L/min. Subsequently, the pre-column was switched in line with an Acclaim PepMap RSLC nanoViper column (0.075  $\times$  500 mm) (Thermo Fisher Scientific). Proteins were eluted by applying a linear gradient from 5% to 40% HPLC solvent B (0.1% (v/v) formic acid, 86% (v/v) acetonitrile) in HPLC solvent A (0.1% (v/v) formic acid) over 30 min at a flow rate of 250 nL/min.

The HPLC was online coupled to an Orbitrap Elite mass spectrometer using the nano-Spray interface (both Thermo Scientific). For ionization, distal coated Silica-Tips (Bruker Daltonics) and the following parameter were used: spray voltage: 1.5 kV, capillary temperature: 250 °C; S-Lens radio frequency level: 68.9%. The mass spectra were recorded in the range from m/z 400 to m/z 1700 at a resolution of 30,000. The mass spectrometer was externally calibrated. The recorded mass spectra were summed over the chromatographic peak and deconvoluted with the FreeStyle 1.8 (ThermoScientific).

### Negative staining TEM

To prepare TEM negative samples, 3.5  $\mu$ L of the sample was applied to a previously glow-discharged 200-mesh copper grid coated with formvar and carbon (Electron Microscopy Sciences). After incubating the sample at room temperature for 1 min, the excess solvent was carefully removed from the edge by filter paper. The copper grid was stained three times with a 2% (w/v) uranyl acetate and then washed three times with Millipore water. Afterwards, it was allowed to dry thoroughly and examined with a JEM-1400 transmission electron microscope (JEOL) at 120 kV.

### Platinum side shadowing

The handedness of the fibrils was determined by platinum side shadowing. A formvar and carbon-coated 200 mesh copper grid (Plano) was glow discharged for 30 s at 40 mA using a PELCO easiGlow glow-discharge cleaning system (Ted Pella). 1.5  $\mu$ L fibril solution were applied on the grid and incubated for 30 min until it was dried up. Platinum particles were evaporated at an angle of 30° onto the grid to form a 1-nm-thick layer by use of a Balzers BAF 300 (Firma) coating device. Grids were imaged using a Hitachi S-5200 scanning electron microscope (Hitachi) provided with an acceleration voltage of 30 kV and operated with the secondary electron detector.

### Cryo-EM sample preparation and data collection

C-flat 1.2/1.3 400-mesh holey carbon-coated grids (Science Services) were glow-discharged at 40 mA for 30 s using a PELCO easiGlow glow-discharge cleaning system (Ted Pella). The fibril solution (i.e. supernatant fractions from the water extraction step) was diluted 1:3 with pure water and 3.5  $\mu$ L of it was placed on the grid. The grid was plunge-frozen in liquid ethane after 8 s with a Leica G 2 (Thermo Fisher Scientific) plunging device using an incubation time of 30 s, >95 % humidity and back-side blotting with filter paper (Whatman). The grid was examined in at 200 kV in a JEM 2100 F (JEOL) transmission electron microscope to optimize the sample quality. The data set for cryo-EM reconstruction was recorded at 300 kV with a Titan Krios (Thermo

Fisher Scientific) that was equipped with K2-Summit detector (Gatan). See Supplementary Table 1 for the data acquisition parameters.

### Reconstruction of the 3D map

The raw data given in tiff were converted into mrcs by IMOD<sup>43</sup>. The mrcs files were then imported into the software Relion 3.1.1<sup>44</sup>, which was used for helical reconstruction. Motion correction was performed using MotionCor2<sup>45</sup>. For refining and correction of the contrast transfer function CTFFIND-4.1<sup>46</sup> was used. The particles were manually picked and extracted with a box-size of 300 Å and an interbox-distance of 33.5 Å, resulting into a total number of 120,356 particles. Afterwards a reference-free 2D classification was performed using 100 classes and a regularization parameter T=2. No classes were dismissed and all particles were used for 3D classification. 3D classification consisted of one round using a cylinder as the initial model and the following other parameters: one class, T = 4, and a low-pass filter of 60 Å. The resulting 3D class was used as a reference model for the next rounds of classification, using 6 classes and T = 4. All classes seemed to belong to a common fibril morphology, and all contained particles were combined for further processing. The T value for following 3D classification with one class was increased to 20 and the result was used as improved reference model for the subsequent 3D auto-refinement and post-processing. This reconstruction was further improved using Bayesian polishing to achieve the final resolution of 2.8 Å, based on the value of the FSC curve for two independently refined half-maps at 0.143. An estimated map-sharpening B-factor factor of -75.9 Å<sup>2</sup> was applied by Relion post-processing. The final map had a twist of -1.4° and a helical rise of 4.8 Å. The reconstruction parameters are listed in Supplementary Table 1.

### Model building

The reconstructed 3D map was used for de-novo model building using Coot v0.9.8<sup>47</sup> software. In a first step, a poly-L-alanine chain was generated and threaded into one molecular layer of the fibril. The poly-L-alanine sequence was replaced with the lysozyme sequence of the patient protein. This model was validated in Phenix version v1.20.1<sup>48</sup> and further refined manually in Coot v0.9.8<sup>47</sup> with the aim to reduce the Ramachandran outliers, geometry outliers, rotamer outliers, and the clash score. After refining one molecular layer of the fibril, a fibril stack of nine layers was generated and refined. Multiplication of the layers was done using the pdsymm tool from Situs<sup>49</sup>. The modeling parameters are listed in Supplementary Table 2.

### Aggregation score

Aggregation scores were calculated using TANGO<sup>50</sup>, WALTZ<sup>51</sup>, FoldAmyloid<sup>52</sup>, Aggrescan<sup>53</sup>, and PASTA2.0<sup>54</sup>. These programs reveal aggregation-prone segments. For this purpose, a value is calculated for each amino acid residue. The following settings were chosen for each program: WALTZ: threshold was set to best overall performance at pH of 7 and values above 75% are considered as hits; TANGO: Beta sheet aggregation values above 0.1 were considered as hits; FoldAmyloid: the scale was set to triple hybrid and amino acid residues having values above 0.062 for five consecutive residues are considered as hits and PASTA 2.0: region detection was set on 90% specificity and values below -2.8 PASTA Energy Units are considered as hits. An aggregation value of 0 means that none of the programs predict a high aggregation value for a particular amino acid residue, and conversely, an aggregation value of 5 means that all five programs predict a high aggregation value for that residue. The values from 0 to 5 were each coded with a color (Fig. S12), and the residues in the model were colored according to the value achieved by these 5 programs.

### Fibril stability calculations

The stabilities of the patient fibril with a G87 residue and the hypothetical WT fibril containing a D87 side chain, were calculated with

different computational tools. The Gibbs free energy of one molecular layer of the structures was calculated with the Amyloid Illustrator server<sup>34</sup>. Three-layer stacks of both fibrils were uploaded to the server and the free energy of one layer was obtained using the standard settings. Next, the energy difference between the unfolded and the folded state of the fibril proteins was calculated with FoldX, using the “stability of object” feature<sup>33</sup>. The obtained energy values for one molecular layer of the fibrils were inverted to reflect the energy gain upon the formation of the folded state. Finally, dissociation energies between two molecular layers of the fibrils along the z-axis were calculated with the program PDBEPIA<sup>32</sup>. The obtained values for the energy needed to dissociate two adjacent layers were inverted to obtain the energy gain upon the association of two layers.

### MD simulations

All-atom MD simulations were performed of the experimentally resolved lysozyme fibril structure and a homology model of this fibril, in which the sequence was reverted back to WT. To obtain the homology model, Gly87 was replaced by Asp using the PDB Manipulator tool of the CHARMM-GUI web-interface<sup>55</sup>. The algorithm replaced G87 with D87 such that the backbone dihedrals remained unchanged and the NH<sub>2</sub> atoms, which are identical in both residues, were placed at the same positions. The fibrils consisted of 6 protein layers, which formed a stable cross- $\beta$  structure as in previous work<sup>56</sup>. The experimental structure and the homology model were placed in cubic simulation boxes with edge lengths of 146.5 Å and simulated at a temperature of 300 K. The boxes were filled with water and 0.15 M NaCl, leading to a total system size of about 412,950 atoms per box. The force field parameters for the proteins were taken from the Amber99sb-star-ildn force field<sup>57</sup> and the TIP4P-Ew<sup>58</sup> model was used for water. For NaCl, we used the Mamatkulov-Schwierz force field parameters<sup>59</sup>.

The MD simulations were performed using the Gromacs simulation package, version 2020.6<sup>60</sup>. To prepare the starting structures, we used the steepest descent algorithm for energy minimization. This method iteratively adjusts atomic positions to lower the potential energy of the system, relaxing the structure to a local minimum. Convergence criteria were set to forces below 1000 kJ/mol/nm and convergence was achieved within 50,000 steps. Subsequently, we performed simulations at 300 K with a constant number of particles N, pressure P and temperature T. Periodic boundary conditions were applied, and the particle-mesh Ewald method was used for the periodic treatment of Coulombic interactions. Bonds to hydrogen atoms were constrained using the LINCS algorithm and a 2 fs time step was used. The systems were equilibrated for 1 ns in the NVT ensemble and then in the NPT ensemble using the Berendsen thermostat and barostat. For the production run, we performed 200 ns long simulations employing the velocity-rescaling thermostat with a stochastic term and a time constant  $\tau_T = 0.1$  ps and isotropic Parrinello–Rahman pressure coupling with a time constant of  $\tau_p = 5$  ps. To obtain sufficient statistics, each production run was repeated six times with the starting velocities drawn randomly from a Maxwell-Boltzmann velocity distribution. In total 2.4  $\mu$ s of simulation time was used.

To quantify the dissociation of terminal chains, we calculated the maximum distance between two adjacent layers at the fibril tips over the course of the 200 ns simulations. The calculations were performed as follows: First, time dependent distances between equivalent amino acids in the two neighboring chains were calculated. Subsequently, maximum distances and standard error of the mean were calculated from the time dependent data. The first 50 ns of each trajectory were skipped in the calculations.

### Statistics

Whenever seen as appropriate the mean value and corresponding standard deviation was reported in this paper.

### Imaging and Morphology analysis

The images of the 3D map and protein model were prepared with the software UCSF Chimera v1.16<sup>31</sup>. Measurements of fibril width and cross-over distance was done with Fiji<sup>61</sup>.

### Reporting summary

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

### Data availability

The reconstructed 3D map was deposited in the Electron Microscopy Data Bank with the accession code [EMD-18883](https://www.rcsb.org/). The coordinates files of the corresponding atomic model were deposited in the Protein Data Bank (<https://www.rcsb.org/>) under the accession code [8R4A](https://www.rcsb.org/). The following published PDB structure was used in the paper: PDB [IJSF](https://www.rcsb.org/). The cryo-EM raw data of the lysozyme derived fibrils were deposited on Electron Microscopy Public Image Archive (<https://www.ebi.ac.uk/empiar/>) with the accession code [EMPIAR-I1785](https://www.ebi.ac.uk/empiar/). Mass spectrometry data was deposited at MassIVE [<https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=bc5cb0cfe8344f1494820fe9dccc289d>], a member of the ProteomeXchange consortium, under the accession code MSV000096142. Source data are provided with this paper.

### References

1. Dumoulin, M., Kumita, J. R. & Dobson, C. M. Normal and aberrant biological self-assembly: insights from studies of human lysozyme and its amyloidogenic variants. *Acc. Chem. Res.* **39**, 603–610 (2006).
2. Anker, S. et al. Lysozyme amyloidosis—a report on a large German cohort and the characterisation of a novel amyloidogenic lysozyme gene variant. *Amyloid* **29**, 245–254 (2022).
3. Pepys, M. B. et al. Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature* **362**, 553–557 (1993).
4. Buxbaum, J. N. et al. Amyloid nomenclature 2022: update, novel proteins, and recommendations by the International Society of Amyloidosis (ISA) nomenclature Committee. *Amyloid* **29**, 213–219 (2022).
5. Dasari, S. et al. Amyloid typing by mass spectrometry in clinical practice: a comprehensive review of 16,175 samples. *Mayo Clin. Proc.* **95**, 1852–1864 (2020).
6. Girnius, S. et al. A new lysozyme tyr54asn mutation causing amyloidosis in a family of Swedish ancestry with gastrointestinal symptoms. *Amyloid Int. J. Exp. Clin. Investig. J. Int. Soc. Amyloidosis* **19**, 182–185 (2012).
7. Yazaki, M., Farrell, S. A. & Benson, M. D. A novel lysozyme mutation Phe57Ile associated with hereditary renal amyloidosis. *Kidney Int* **63**, 1652–1657 (2003).
8. Valleix, S. et al. Hereditary renal amyloidosis caused by a new variant lysozyme W64R in a French family. *Kidney Int* **61**, 907–912 (2002).
9. Wooliver, et al. Familial amyloidosis associated with a novel mutation (D68G) in the lysozyme gene. *XIth Int Symp Amyloidosis*. p. 208–210 (2008).
10. Nasr, S. H. et al. Hereditary lysozyme amyloidosis variant p.Leu102Ser associates with unique phenotype. *J. Am. Soc. Nephrol. JASN* **28**, 431–438 (2017).
11. Röcken, C. et al. ALys amyloidosis caused by compound heterozygosity in Exon 2 (Thr70Asn) and Exon 4 (Trp112Arg) of the lysozyme gene. *Hum. Mutat.* **27**, 119–120 (2006).
12. Suvorina, M. Y. et al. First report of lysozyme amyloidosis with p.F21L/T88N amino acid substitutions in a Russian family. *Int. J. Mol. Sci.* **24**, 14453 (2023).
13. Blake, C. C. F. et al. Structure of hen egg-white lysozyme: a three-dimensional fourier synthesis at 2 Å resolution. *Nature* **206**, 757–761 (1965).

14. Matagne, A. & Dobson, C. M. The folding process of hen lysozyme: a perspective from the 'new view'. *Cell. Mol. Life Sci. CMLS* **54**, 363–371 (1998).
15. Booth, D. R. et al. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **385**, 787–793 (1997).
16. Canet, D. et al. Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme. *Nat. Struct. Biol.* **9**, 308–315 (2002).
17. Hammarström, P., Jiang, X., Hurshman, A. R., Powers, E. T. & Kelly, J. W. Sequence-dependent denaturation energetics: a major determinant in amyloid disease diversity. *Proc. Natl Acad. Sci.* **99**, 16427–16432 (2002).
18. Blancas-Mejía, L. M. & Ramirez-Alvarado, M. Systemic amyloidoses. *Annu. Rev. Biochem.* **82**, 745–774 (2013).
19. Baur, J. et al. Identification of AL proteins from 10  $\lambda$ -AL amyloidosis patients by mass spectrometry extracted from abdominal fat and heart tissue. *Amyloid* **30**, 27–37 (2023).
20. Liberta, F. et al. Cryo-EM fibril structures from systemic AA amyloidosis reveal the species complementarity of pathological amyloids. *Nat. Commun.* **10**, 1104 (2019).
21. Radamaker, L. et al. Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis. *Nat. Commun.* **10**, 1103 (2019).
22. Bansal, A. et al. AA amyloid fibrils from diseased tissue are structurally different from in vitro formed SAA fibrils. *Nat. Commun.* **12**, 1013 (2021).
23. Iakovleva, I. et al. Structural basis for transthyretin amyloid formation in vitreous body of the eye. *Nat. Commun.* **12**, 7141 (2021).
24. Radamaker, L. et al. Cryo-EM reveals structural breaks in a patient-derived amyloid fibril from systemic AL amyloidosis. *Nat. Commun.* **12**, 875 (2021).
25. Radamaker, L. et al. Role of mutations and post-translational modifications in systemic AL amyloidosis studied by cryo-EM. *Nat. Commun.* **12**, 6434 (2021).
26. Banerjee, S. et al. Amyloid fibril structure from the vascular variant of systemic AA amyloidosis. *Nat. Commun.* **13**, 7261 (2022).
27. Schulte, T. et al. Cryo-EM structure of ex vivo fibrils associated with extreme AA amyloidosis prevalence in a cat shelter. *Nat. Commun.* **13**, 7041 (2022).
28. Steinebrei, M. et al. Cryo-EM structure of an ATTRwt amyloid fibril from systemic non-hereditary transthyretin amyloidosis. *Nat. Commun.* **13**, 6398 (2022).
29. Puri, S. et al. The Cryo-EM STRUCTURE of renal amyloid fibril suggests structurally homogeneous multiorgan aggregation in AL amyloidosis. *J. Mol. Biol.* **435**, 168215 (2023).
30. Falcon, B. et al. Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019).
31. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
32. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
33. Vanhee, P. et al. BriX: a database of protein building blocks for structural analysis, modeling and design. *Nucleic Acids Res* **39**, D435–D442 (2011).
34. Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R. & Eisenberg, D. S. The expanding amyloid family: structure, stability, function, and pathogenesis. *Cell* **184**, 4857–4873 (2021).
35. Annamalai, K. et al. Common fibril structures imply systemically conserved protein misfolding pathways in vivo. *Angew. Chem.* **129**, 7618–7622 (2017).
36. Arseni, D. et al. TDP-43 forms amyloid filaments with a distinct fold in type A FTLD-TDP. *Nature* **620**, 898–903 (2023).
37. Jiang, Y. X. et al. Amyloid fibrils in FTLD-TDP are composed of TMEM106B and not TDP-43. *Nature* **605**, 304–309 (2022).
38. Li, B. et al. Cryo-EM of full-length  $\alpha$ -synuclein reveals fibril polymorphs with a common structural kernel. *Nat. Commun.* **9**, 3609 (2018).
39. Ihse, E. et al. Amyloid fibril composition is related to the phenotype of hereditary transthyretin V30M amyloidosis. *J. Pathol.* **216**, 253–261 (2008).
40. Raimondi, S. et al. Comparative study of the stabilities of synthetic in vitro and natural ex vivo transthyretin amyloid fibrils. *J. Biol. Chem.* **295**, 11379–11387 (2020).
41. Schönfelder, J. et al. Protease resistance of ex vivo amyloid fibrils implies the proteolytic selection of disease-associated fibril morphologies. *Amyloid* **28**, 243–251 (2021).
42. Bulawa, C. E. et al. Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *Proc. Natl Acad. Sci.* **109**, 9629–9634 (2012).
43. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
44. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
45. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
46. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
47. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
48. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877 (2019).
49. Wriggers, W. Conventions and workflows for using Situs. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 344–351 (2012).
50. Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. & Serano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* **22**, 1302–1306 (2004).
51. Maurer-Stroh, S. et al. Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat. Methods* **7**, 237–242 (2010).
52. Garbuzynskiy, S. O., Lobanov, M. Y. & Galzitskaya, O. V. FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. *Bioinforma. Oxf. Engl.* **26**, 326–332 (2010).
53. Conchillo-Solé, O. et al. AGGRESCAN: a server for the prediction and evaluation of 'hot spots' of aggregation in polypeptides. *BMC Bioinforma.* **8**, 65 (2007).
54. Walsh, I., Seno, F., Tosatto, S. C. E. & Trovato, A. PASTA 2.0: an improved server for protein aggregation prediction. *Nucleic Acids Res* **42**, W301–W307 (2014).
55. Park, S.-J., Kern, N., Brown, T., Lee, J. & Im, W. CHARMM-GUI PDB manipulator: various PDB structural modifications for biomolecular modeling and simulation. *J. Mol. Biol.* **435**, 167995 (2023).
56. Schwierz, N., Frost, C. V., Geissler, P. L. & Zacharias, M. Dynamics of seeded A $\beta$ 40-Fibril growth from atomistic molecular dynamics simulations: kinetic trapping and reduced water mobility in the locking step. *J. Am. Chem. Soc.* **138**, 527–539 (2016).
57. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
58. Horn, H. W. et al. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **120**, 9665–9678 (2004).



59. Mamatkulov, S. & Schwierz, N. Force fields for monovalent and divalent metal cations in TIP3P water based on thermodynamic and kinetic properties. *J. Chem. Phys.* **148**, 074504 (2018).
60. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
61. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
62. Harata, K., Abe, Y. & Muraki, M. Full-matrix least-squares refinement of lysozymes and analysis of anisotropic thermal motion. *Proteins* **30**, 232–243 (1998).

## Acknowledgements

The authors thank Paul Walther (Ulm University) for technical support. Funding was partially provided by the Deutsche Forschungsgemeinschaft (FA 456/28 to M.F.). The collection of the cryo-EM datasets was funded by iNEXT Discovery (18586; Horizon 2020, European Union) and was done in the European Molecular Biology Laboratory in Heidelberg (Germany).

## Author contributions

S.K., K.S., L.K., M.U., N.S., P.B.P., C.H., S.W. and M.S. carried out research. S.K., K.S., M.U., N.S., M.S., C.H., S.W. and M.F. analyzed data. U.H., and S.O.S. contributed materials. M.F. designed research. S.K., M.S. and M.F. wrote the paper with contributions from all other authors.

## 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/s41467-024-54091-7>.

**Correspondence** and requests for materials should be addressed to Sara Karimi-Farsijani.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024