Title
Designed peptides that assemble into cross-α amyloid-like structures

Authors
Zhang, Shaoqing
Huang, Hai
Yang, Junjiao
et al.

Publication Date
2018-07-01

Peer reviewed

Designed peptides that assemble into cross-α amyloid-like structures

Shao-Qing Zhang[1,2,3], Hai Huang[3,7], Junjiao Yang[2,3,7], Huong T. Kratochvil[2,3], Marco Lolicato[3], Yanxin Liu[4], Xiaokun Shu[2,3,*], Lijun Liu[3,5,*], William F. DeGrado[2,3,6,*]

[1]Department of Chemistry, University of Pennsylvania, Philadelphia, PA, 19014, United States
[2]Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA 94158, United States
[3]Cardiovascular Research Institute, University of California at San Francisco, San Francisco, CA 94158, United States
[4]Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, CA 94158, United States
[5]DLX Scientific, Lawrence, KS, 66049, United States
[6]Institute for Neurodegenerative Diseases, University of California at San Francisco, San Francisco, CA 94158, United States
[7]These authors contributed equally to this work.
[*]Email: xiaokun.shu@ucsf.edu, lijunliuks@gmail.com or william.degrado@ucsf.edu.

Abstract

Amyloids adopt "cross-β" structures composed of long twisted fibrils with β-strands running perpendicular to the fibril axis. Recently, a toxic peptide was proposed to form amyloid-like cross-α structures in solution, with a planar bilayer-like assembly observed in the crystal structure. Here we crystallographically characterize designed peptides that assemble into spiraling cross-α amyloid-like structures, which resemble twisted β-amyloid fibrils. The peptides form helical dimers, stabilized by packing of small and apolar residues, and the dimers further assemble into cross-α amyloid-like fibrils with superhelical pitches ranging from 170 Å to 200 Å. Converting a small residue that appeared critical for packing to Leu, resulted in structural rearrangement to a helical polymer. Fluorescently tagged versions of the designed peptides form puncta in mammalian cells, which recover from photobleaching with markedly different kinetics – potentially useful for directing *in vivo* protein assemblies with predetermined spacings and stabilities.

Introduction

Understanding the principles by which peptides organize into higher order assemblies is a topic of considerable interest in the soft matter, biological and chemical communities[1]. The emerging principles inform our knowledge of normal and pathological processes in biology, and this knowledge is translating to the construction of soft materials with diverse functions, including stimulus-responsive hydrogel sensors[2] and vehicles for drug delivery[3]. In particular, the study of β-amyloids[4] represents a large area of science with implications for understanding neurodegeneration[5], amyloid diseases[6] and epigenetic phenomena[7]. β-Amyloids have also been widely used for designing nanomaterials[8] and catalysts[9, 10, 11]; they also are proposed to represent key steps in the molecular evolution of proteins[12, 13]. Classically, amyloids have "cross-β" structures, in which the β-strands align perpendicular to the long axis of an infinite fibril. By contrast, in self-assembling elongated helical peptides such as coiled coils the helices generally align nearly parallel to the fiber axis. Therefore, it was surprising to discover that a toxic peptide PSMα3 from *Staphylococcal aureus* formed cross-α amyloid-like structures in which the helices were proposed to align perpendicular rather than parallel to the fibril axis[14]. The assembly was demonstrated by negative stain electron microscopy (EM) and the ability to bind an amyloid-staining dye. Furthermore, the peptide crystallized in a bilayer-like arrangement with the helices interacting laterally. Prior to this work, designed peptides had been found to crystallize as bilayers[15, 16], but the same peptides did not appear to form fibrils in solution.

Given that the cross-α structure has only recently been proposed as an important organizing principle in Nature, very little is known about the general physical principles by which cross-α amyloid-like form, or the extent to which the morphology of the cross-α amyloid-like structures can be manipulated by design. Peptide nanotubes have been designed based on a coiled coil repeat, which aligned either perpendicular or diagonal to the long axis of the tube as seen by EM at near-atomic resolution[17]. Also, a number of repeat proteins, which incorporate loops between helices, have been prepared[18-20]. Nevertheless, structures of self-assembling spiraling cross-α amyloid-like structures have yet to be observed. Here, we describe the structure of a membrane-interactive peptide that surprisingly forms a long, twisted cross-α spiral. We then analyze the structural features stabilizing the spiral to guide the design of water-soluble cross-α amyloid-like structures that assemble *in vitro* and *in vivo*. Finally, we demonstrate how small sequence changes can translate into large changes in supramolecular structure, and how these structural changes influence their ability to assemble and remodel within living cells.

Results

X-ray structure of a cross-α spiral hydrophobic peptide

Our discovery of a structurally well-defined cross-α structures arose from a serendipitous discovery of the packing in the crystal structure of αAm_{mem}, an analogue of the membrane-spanning Zn^{2+}-transporting peptide Rocker[21] based on the backbone of a mononuclear Zn^{2+}-binding four-helix bundle[22] (Supplementary Fig. 1a). αAm_{mem} adopted an antiparallel dimer of straight α-helices, which further assembled into long counter-clockwise twisted fibrils in the crystal lattice (Supplementary Fig. 1b). Similar to cross-β structures, the axes of the α-helices lie perpendicular to the main superhelical axis of this structure (Fig. 1a). The superhelix is formed from secondary structural units that interact across the fibril axis

to create a two-layered structure, shown in Fig. 1b. In αAm_{mem} the helices form tight parallel interactions across the superhelical axis (Fig. 1c), creating a series of parallel dimeric helical pairs. Progressing along the superhelical axis, each dimer is rotated by -160° creating a spiraling set of left-handed antiparallel helical pairs (Fig. 1a). The resulting up-up-down-down topology corresponds to class 5 packing in the amyloid classification scheme[23] (Supplementary Fig. 2). Both the parallel and antiparallel helical pairs have approximate two-fold symmetry (Fig. 1c,d). Unlike in coiled coils, the helices are straight rather than curving around neighboring α-helices.

Design of water-soluble cross-α spiral peptides

There is considerable geometric complementarity in the sidechain packing along the entire superhelical assembly of αAm_{mem}. Each α-helix in the superhelix interacts in three geometrically distinct manners: namely antiparallel interactions with helices within a sheet above and below the reference helix, as well as the antiparallel interaction across the fibril axis (Fig. 2a). Each of the three geometrically distinct helical pairs has left-handed crossing angles, which propagate to create the spiraling superhelix. Small Ala or Ser residues (sticks in Fig. 2b) positioned on three faces of each α-helix appear critical for achieving the tight packing of the structure. They pack near the point of closest approach of each of the three helix-helix packing interfaces where they facilitate close inter-helical contacts; larger residues (shown as sticks in Fig. 2b) line the interfaces as the helices diverge from a point of closest approach.

Although the dimers feature straight helices (rather than curving around one another), the packing shares similarities to knobs-into-holes packing of left-handed coiled coils near the point of closest approach. We therefore use the familiar heptad repeat nomenclature (Fig. 2c) to discuss packing of the helices. Ala17 facilitates a close contact between the parallel pairs of helices across the fibril axis in the *d-d'* interface (defined in Fig. 2b), and Ser11 and Ala13 mediate close packing between the two geometrically distinct antiparallel helical pairings (Fig. 2b). Thus, the sequence of αAm_{mem} satisfies the requirements for mutual stabilization of three distinct packing arrangements of a single α-helix. Thus, the satisfaction of these multiple packing requirements underlies the tendency of the helices to remain straight rather than coiling about one another (which would optimize packing about only one or two helix-helix interfaces in this scenario in which all helices have a left-handed crossing angle).

We used the above-mentioned features to engineer the sequences of water-soluble peptides (Fig. 2d) capable of assembling into spiraling cross-α structures, as described in Online Methods. Briefly, the surface residues (Fig. 2e) in the αAm_{mem} structure were replaced with water-solubilizing Glu, Lys and Arg residues, capable of forming stabilizing electrostatic and hydrogen-bonded interactions with the charged sidechains (Fig. 2f). The remaining interior positions were retained as in αAm_{mem}. To test the importance of the small Ser11 sidechain in mediating interhelical interactions (Fig. 2d), we synthesized a set of peptides in which this residue was varied to other small residues Gly, Ala and Ser, as well as a larger hydrophobic Leu sidechain (Fig. 2d) in peptides designated αAm_G, αAm_A, αAm_S and αAm_L.

Assembly and crystal structures of αAm peptides

The αAm peptides showed good solubility in deionized water, but were found to assemble in a time-dependent manner when incubated in buffer at a broad range of pH from 3.5-9.5. The formation of amyloid is monitored by measuring the fluorescence enhancement of the fluorogenic dye thioflavin T (ThT), which is known to stain cross-α amyloid-like structures[14] and β-amyloid assemblies. Following dissolution in aqueous buffers, all of the αAm peptides showed time-dependent increases in the fluorescence intensity of ThT, with no apparent lag time in the kinetic traces (Fig. 3a). The toxic fibril-forming peptide PSMα3 has a similar aggregation behavior[14]. The half-time for assembly varied from approximately 7 min for the most hydrophobic peptide, αAm_L, to 24 hr for αAm_A (Fig. 3a). αAm_G showed relatively rapid kinetics ($t_{1/2}$ = 30 min) while αAm_S was slow ($t_{1/2}$ = 8 hr), indicating that hydrophobicity alone did not explain the variation in the assembly kinetics. αAm_G, αAm_A and αAm_S form thin fibrils as shown by negative-stain EM, while αAm_L form much wider and longer fibrils (Fig. 3b). Infrared (IR) and circular dichroism (CD) spectroscopies was used to determine the secondary structure of the peptides. Each peptide showed a well-resolved spectrum, characteristic of the α-helix (Supplementary Fig. 3).

The peptides in Fig. 3b were crystallized, and high-resolution structures (Supplementary Table 1) were determined for all except αAm_A, whose crystals diffracted to only 4.0 Å. The geometric parameters of the amyloid-like assemblies are listed in Supplementary Table 2. The superhelical structure of αAm_G was solved by molecular replacement in two space groups ($P4_322$, 2.49 Å resolution and $P6_122$, 3.30 Å resolution), and the structures were found to be nearly identical (0.4 Å Cα RMSD computed over the 450 residues in the asymmetric unit (ASU)). There are nine parallel dimers in the asymmetric unit, which form a half turn of a superhelix (i.e., 18 parallel dimers/turn; pitch = 172.8 Å) (Fig. 3c). The local packing interactions in the αAm_G structures are nearly identical to those in αAm_{mem}, including the inclusion of small residues at positions where the helices approach most closely (Supplementary Fig. 4).

For αAm_S, two superhelical arrangements were observed running in different directions through the crystal lattice: one had 20 parallel dimers/turn while the other had 19.5 dimers/turn. αAm_S formed crystals in space group $P2_1$ (diffraction limit, 2.5 Å) with a large unit cell (a = 161.166 Å, b = 160.159 Å, c = 198.502 Å), which presented a significant challenge for structure determination. Ultimately the structure was solved by molecular replacement as described in the supplement. The asymmetric unit contains 236 crystallographically non-equivalent straight α-helices. Together, they form three crystallographically distinct but structurally related superhelices. The first two superhelices (designated Superhelices "1" and "2") have 20 dimers per turn (pitch = 198.5 Å) (Fig. 3d). The remaining 78 parallel dimers form four complete turns of "Superhelix 3" which repeats over a length of 775.6 Å (19.5 parallel dimers/turn, pitch = 193.9 Å) (Fig. 3e). Superhelix 3 can be conceptually subdivided into four subgroups, designated 3A-3D, each representing approximately one superhelical turn similar to groups 1 and 2. A unit cell contains multiple copies of the subgroups, and subgroups 3A-3D stack between adjacent unit cells to generate the four turns in superhelix 3, which repeats through the crystal lattice (Supplementary Fig. 5). The small-residue packing patterns of αAm_S resemble are very

similar to those seen in αAm_{mem} (Supplementary Fig. 4), although the limited resolution of αAm_{mem} (3.5 Å) precluded a detailed comparison.

In summary, αAm_S and αAm_G have a limited degree of pitch diversity, ranging from 18 parallel dimers/turn seen in two independent crystal structures for αAm_G to 19.5 and 20 dimers/turn observed within a single crystal lattice for αAm_S. Structurally, these changes represent only small differences in the crossing angle of the antiparallel helical pairs that define the superhelical pitch, ranging only slightly from 20° for αAm_G to 18°, 18° and 18.5° for the three superhelices observed for αAm_S (angles measured projected onto a plane normal to the superhelical axis). The minor increase in packing angle for αAm_G is likely related to differences in packing of the small Gly and Ser residues. Also, electron density from one or more solvent molecules was observed between abutting Gly residues on adjacent helices in αAm_G, although it could not be assigned with confidence at a diffraction limit of 2.5 Å.

More deep-seated structural differences were seen in two variants, which we synthesized to evaluate the effect of small-to-large variations at the packing interface; the variant αAm_L and αAm_F have a large Leu and an even larger Phe sidechain at position 11, respectively, while all three critical small positions at 11, 13 and 17 were simultaneously changed to Leu in αAm_{3L} (Fig. 2d). αAm_F and αAm_{3L} were helical in solution (Supplementary Fig. 3), but had limited solubility in buffers and failed crystallization attempts, and were as observed to form amorphous aggregates by negative-stain EM (Supplementary Fig. 6); they therefore were not studied further. By contrast, αAm_L, in which only one small residue at position 11 is mutated to intermediate-sized Leu, rapidly assembled into a structure that bound ThT (Fig. 3a), and formed crystals that diffract to 2.0 Å (Supplementary Table 1). Interestingly, the peptide adopts an entirely different structure from the other peptides with small residues at this position (Fig. 4a). The large Leu sidechain apparently disrupts the cross-α packing, and it instead defaults to form a canonical antiparallel four-helix coiled coil[24] (Fig. 4b). The structure of αAm_L can be related to that of αAm_S by a 48° rotation of the α-helices about their α-helical axes, and an approximately 5 Å net translation of the parallel helices, such that they come to occupy more distant diagonal positions in the four-helix bundle (Fig. 4b). These geometric changes place the small Ala residues at positions 13 and 17 into "*d*" and "*a*" positions of a heptad repeat, respectively (Supplementary Fig. 7a,b). In the tetramer, these small residues pack together with larger residues in a geometrically complementary jigsaw puzzle-like manner similar to that first seen in the protein ROP[25].

The αAm_L tetramers were found to form a much open and wide helical polymer rather than a cross-α amyloid-like structure. The large Leu11 sidechains project from the surface of the bundle, where they mediate hydrophobic lateral contacts between individual tetramers (Fig. 4a). To probe the role of Leu11 and other hydrophobic residues at the tetramer-tetramer interface in mediating its higher order assembly, we converted the apolar residues that mediate the assembly to polar sidechains, while maintaining the identities of the remaining residues (Fig. 2d and Supplementary Fig. 7c,d). The crystal structure of this peptide, designated αTet, was nearly identical to the tetrameric unit of αAm_L (Cα RMSD 0.7 Å over the full tetramers). However, αTet did not show a high order assembly in

solution as assessed by the ThT assay (Fig. 3a). Also, there were few inter-tetramer contacts in the lattice, and those that were formed were primarily solvent-mediated (Supplementary Fig. 8).

Association of αAm peptides in mammalian cells

The aggregation of amyloid-forming proteins such as tau and α-synuclein in mammalian cells is often assessed by fluorescence microscopy, using genetically encoded fusions of the protein of interest and a fluorescent protein[26, 27]. The formed protein inclusions appear as bright intracellular puncta accompanied by the loss of the more diffuse staining from soluble proteins (Supplementary Fig. 9). To determine whether the αAm series of peptides similarly formed intracellular inclusions, four peptides (αAm_G, αAm_A, αAm_S and αAm_L) were tagged with the enhanced GFP (EGFP) and expressed in HEK 293T cells. All of these αAm peptides aggregated in the cytosol as assessed from the presence of bright puncta (diameter < 5 μm) (Fig. 5a). To pinpoint the physical properties of the inclusions, we performed fluorescence recovery after photo-bleaching (FRAP) experiments, which examines protein motility in the inclusions (Fig. 5b). Interestingly, both αAm_G and αAm_S, which had been shown by crystallography to form cross-α structures, showed irreversible photo-bleaching over a period of 10 min (during which time the fluorescent intensity of the bleached area stayed below 50% of pre-bleach level). The results suggest that the fusion proteins in the inclusions are immobile within this time range (Fig. 5b). The αAm_A fusion, which presumably also forms a cross-α spiral (given the similarity of its sequence to αAm_G and αAm_S), behaved similarly. In contrast, αAm_L-expressing cells showed fluorescence recovery after photo-bleaching, indicating that this protein is mobile in the puncta structure (Fig. 5c). The result suggests that this peptide forms liquid droplet-like phases in the cells, which is consistent with the more open and less tightly packed helical polymer formed by this peptide. As a negative control, there is no puncta in αTet-expressing cells (Supplementary Fig. 10).

Discussion

Molecular assembly of peptides and proteins is used throughout biology for compartmentalization and display purposes. Protein design provides an approach to test and extend our understanding of assembly, and to engineer artificial molecules that can direct assembly with precisely defined stoichiometries and spacings. For example, Keating, Woolfson and others have developed principles for design of self-assembly of coiled coils, which can be used to test the role of dimerization or oligomerization in diverse cellular processes[28, 29]. Much larger polymeric assemblies are also ubiquitous throughout nature, and they range from the very rigid, precisely ordered and closely spaced spiral arrays formed by amyloid-forming sequences to highly mobile liquid droplet-like phases[30]. Here, we describe the design of two types of self-assembling systems: densely packed cross-α amyloid-like materials as well as a less densely packed reversibly assembling helical polymer. Both might prove useful for diverse applications from nanotechnology to cell biology, where they could provide useful modules for inducing assembly of protein domains into well-defined and predetermined arrays.

The spiraling cross-α amyloid-like conformation is particularly interesting, as to the best of our knowledge it had not been crystallographically characterized in natural or synthetic

systems. Its conformation contrasts with the more planar bilayer-like packing arrangement seen in designed helical peptides[15, 16] and a natural toxic peptide[14] (Supplementary Fig. 11a). The features required for assembly into this cross-α amyloid-like spiral structures are relatively simple: small residues positioned on three faces of an α-helix mediate close contacts with neighboring helices (Fig. 2 and Supplementary Fig. 4). Together with larger apolar residues aligned along the three packing interfaces, the small residues mediate packing between straight helices with a small (15° to 20°) left-handed crossing angle. The uniform left-handed crossings give rise to a progressive left-handed screw that generates the spiraling amyloid-like structures (Supplementary Fig. 11b). Menwhile, it is interesting to compare the helical polymer structure of αAm_L to that of the SAM oligomerization domain[31] (Supplementary Fig. 11c), which is widely utilized oligomerization motif used to assemble a variety of domains for diverse signaling functions[32, 33]. Disruption of this packing motif by increasing the bulk of even one of the small residues in αAm_L resulted in the formation of classical antiparallel four-helix bundles that assembled in a less regular and dense manner through association of the faces of tetrameric units (Supplementary Fig. 11d). Thus, the novel folds described in this work provide a range of assemblies that have not yet been discovered in Nature. Furthermore, these folds potentially provide both packing densities and the ability to direct patterned linear arrays of fused domains that aptly reflect phenomena found in Nature.

Acknowledgments
We thank Drs. David Bulkley, Peng Jin, Sam Li, Xi Liu, Nicholas Polizzi, Nathan Schmidt and Haifan Wu for technical help. This work was primarily supported by NIH grant R35GM122603 to W.F.D., with additional support from the NSF (CHE1413295) for the MRSEC program to the LRSM at the University of Pennsylvania. Y.L was supported by a Howard Hughes Medical Institute-Helen Hay Whitney Foundation Postdoctoral Fellowship.

Author contributions
S.-Q.Z. and W.F.D. conceived the project. S.-Q.Z. designed all the peptide sequences and performed *in-vitro* experiments with H.T.K., M.L. and Y.L. H.H. and J.Y. conducted cellular *in-vivo* experiments. L.L. solved and refined all the crystal structures. S.-Q.Z., H.H, J.Y., H.T.K., Y.L., X.S., L.L. and W.F.D. analyzed the data. S.-Q.Z. and W.F.D. prepared the manuscript with contributions from all the authors.

Competing financial interests
The authors declare no competing financial interests.

Accessions
The crystal structures of αAm_{mem}, αAm_G, αAm_G (low resolution), αAm_S, αAm_L and αTet have been deposited in the RCSB Protein Database Bank under the codes 6C4X, 6C4Y, 6C4Z, 6C50, 6C51 and 6C52, respectively.

References

1. Gazit, E. Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chemical Society reviews* **36**, 1263-1269 (2007).
2. Shigemitsu, H. & Hamachi, I. Design Strategies of Stimuli-Responsive Supramolecular Hydrogels Relying on Structural Analyses and Cell-Mimicking Approaches. *Acc Chem Res* **50**, 740-750 (2017).
3. Shu, J.Y., Panganiban, B. & Xu, T. Peptide-polymer conjugates: from fundamental science to application. *Annu Rev Phys Chem* **64**, 631-657 (2013).
4. Riek, R. & Eisenberg, D.S. The activities of amyloids from a structural perspective. *Nature* **539**, 227-235 (2016).
5. Prusiner, S.B. Cell biology. A unifying role for prions in neurodegenerative diseases. *Science* **336**, 1511-1513 (2012).
6. Sacchettini, J.C. & Kelly, J.W. Therapeutic strategies for human amyloid diseases. *Nature reviews. Drug discovery* **1**, 267-275 (2002).
7. Toyama, B.H. & Weissman, J.S. Amyloid structure: conformational diversity and consequences. *Annual review of biochemistry* **80**, 557-585 (2011).
8. Knowles, T.P. & Buehler, M.J. Nanomechanics of functional and pathological amyloid materials. *Nat Nanotechnol* **6**, 469-479 (2011).
9. Rufo, C.M. *et al.* Short peptides self-assemble to produce catalytic amyloids. *Nat Chem* **6**, 303-309 (2014).
10. Makhlynets, O.V., Gosavi, P.M. & Korendovych, I.V. Short Self-Assembling Peptides Are Able to Bind to Copper and Activate Oxygen. *Angew Chem Int Ed Engl* **55**, 9017-9020 (2016).
11. Tena-Solsona, M. *et al.* Emergent Catalytic Behavior of Self-Assembled Low Molecular Weight Peptide-Based Aggregates and Hydrogels. *Chemistry* **22**, 6687-6694 (2016).
12. Friedmann, M.P. *et al.* Towards Prebiotic Catalytic Amyloids Using High Throughput Screening. *PLoS One* **10**, e0143948 (2015).
13. Childers, W.S., Ni, R., Mehta, A.K. & Lynn, D.G. Peptide membranes in chemical evolution. *Curr Opin Chem Biol* **13**, 652-659 (2009).
14. Tayeb-Fligelman, E. *et al.* The cytotoxic Staphylococcus aureus PSMalpha3 reveals a cross-alpha amyloid-like fibril. *Science* **355**, 831-833 (2017).
15. Prive, G.G., Anderson, D.H., Wesson, L., Cascio, D. & Eisenberg, D. Packed protein bilayers in the 0.90 A resolution structure of a designed alpha helical bundle. *Protein Sci* **8**, 1400-1409 (1999).
16. Patterson, W.R., Anderson, D.H., DeGrado, W.F., Cascio, D. & Eisenberg, D. Centrosymmetric bilayers in the 0.75 A resolution structure of a designed alpha-helical peptide, D,L-Alpha-1. *Protein Sci* **8**, 1410-1422 (1999).
17. Egelman, E.H. *et al.* Structural plasticity of helical nanotubes based on coiled-coil assemblies. *Structure* **23**, 280-289 (2015).
18. Brunette, T.J. *et al.* Exploring the repeat protein universe through computational protein design. *Nature* **528**, 580-584 (2015).
19. Main, E.R., Jackson, S.E. & Regan, L. The folding and design of repeat proteins: reaching a consensus. *Curr Opin Struct Biol* **13**, 482-489 (2003).
20. Pluckthun, A. Designed ankyrin repeat proteins (DARPins): binding proteins for research, diagnostics, and therapy. *Annu Rev Pharmacol Toxicol* **55**, 489-511 (2015).
21. Joh, N.H. *et al.* De novo design of a transmembrane Zn(2)(+)-transporting four-helix bundle. *Science* **346**, 1520-1524 (2014).

22. Zhang, S.-Q. *et al.* De Novo Design of Tetranuclear Transition Metal Clusters Stabilized by Hydrogen-Bonded Networks in Helical Bundles. *Journal of the American Chemical Society* **140**, 1294-1304 (2018).
23. Eisenberg, D.S. & Sawaya, M.R. Structural Studies of Amyloid Proteins at the Molecular Level. *Annual Review of Biochemistry* **86**, 69-95 (2017).
24. Szczepaniak, K., Lach, G., Bujnicki, J.M. & Dunin-Horkawicz, S. Designability landscape reveals sequence features that define axial helix rotation in four-helical homo-oligomeric antiparallel coiled-coil structures. *J Struct Biol* **188**, 123-133 (2014).
25. Banner, D.W., Kokkinidis, M. & Tsernoglou, D. Structure of the ColE1 rop protein at 1.7 A resolution. *J Mol Biol* **196**, 657-675 (1987).
26. Sanders, D.W. *et al.* Distinct tau prion strains propagate in cells and mice and define different tauopathies. *Neuron* **82**, 1271-1288 (2014).
27. Prusiner, S.B. *et al.* Evidence for alpha-synuclein prions causing multiple system atrophy in humans with parkinsonism. *Proc Natl Acad Sci U S A* **112**, E5308-5317 (2015).
28. Thompson, K.E., Bashor, C.J., Lim, W.A. & Keating, A.E. SYNZIP protein interaction toolbox: in vitro and in vivo specifications of heterospecific coiled-coil interaction domains. *ACS Synth Biol* **1**, 118-129 (2012).
29. Fletcher, J.M. *et al.* A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology. *ACS Synth Biol* **1**, 240-250 (2012).
30. Banani, S.F., Lee, H.O., Hyman, A.A. & Rosen, M.K. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285-298 (2017).
31. Kim, C.A., Sawaya, M.R., Cascio, D., Kim, W. & Bowie, J.U. Structural organization of a Sex-comb-on-midleg/polyhomeotic copolymer. *J Biol Chem* **280**, 27769-27775 (2005).
32. Kim, C.A. & Bowie, J.U. SAM domains: uniform structure, diversity of function. *Trends Biochem Sci* **28**, 625-628 (2003).
33. Wu, H. & Fuxreiter, M. The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell* **165**, 1055-1066 (2016).

Online Methods

Protein Design

The sequences of water-soluble αAm peptides were designed based on the crystal structure of αAm_{mem}. The residues on the positions *b*, *c* and *f* in the heptad repeats as shown in **Fig. 2c** were designed with the models generated by the Rosetta fixbb module[34], and residues on other positions are fixed. The residues on these three positions allowed only charged residues Arg, Asp, Lys or Glu. Rosetta generated models were used to discover combinations that could form interchain hydrogen bonds, and we then manually selected pairs for the final sequence of αAm_S. The sequences of αAm_G, αAm_A and αAm_L were obtained by changing the residues on position *e* to modulate the interface size. The sequence of αAm_{3L} is based on that of the water-soluble αAm peptides for a systematic change of all three small residues to test their importance. The sequence of αTet is obtained by manually changing the residues at position c and e in the heptad designation for the antiparallel tetramer subunit of αAm_L for intrachain hydrogen bonds between the residues i→i+2 and i→i+3.

Peptide Synthesis and Purification

The peptides were synthesized and purified according to the procedures previously described[22].

Thioflavin-T Kinetics Assay

The peptides were prepared to a final volume of 100 µL at a final concentration of 200 µM peptides and 200 µM ThT in 1X PBS in the 96-well bottom-clear non-binding plates (Greiner). Each peptide was tested with four replicates. The plate was sealed with a clear film (Nunc), placed in a Spectramax M5 plate reader (Molecular Devices) set at 37 ºC, and subjected to repeated rounds of 1-min rest and 4-min shaking. The reading of the ThT fluorescence was top-read at λ_{ex} = 444 nm and λ_{em} = 485 nm, and recorded in an interval of 5 min for 96 hours.

Formation of fibrils

The peptides were prepared at 100 µM in 1X PBS in 1.5-mL Eppendorf tubes. Samples were incubated at 37 °C for 96 hr under constant agitation at 900 r.p.m.

Negative-stain Electron Microscopy

A suspension of 5 µl samples prepared as described above was briefly vortexed and then applied to 300 mesh Cu grids coated with thin carbon, and incubated for ~20 min. Following sample incubation, the grids were stained twice with uranyl formate (for αAm_G) or uranyl acetate (for other peptides). The excess stain was removed by blotting from the side and vacuum dried. Prepared grids were imaged with a FEI TECNAI 20 operated at 200 kV. Images were recorded using an 8k × 8k TemCam-F816 CMOS camera from TVIPS at a magnification of 11000x (αAm_L), 62000x (αAm_G, αAm_A, and αAm_S) and 29000x (αAm_F and αAm_{3L}) with -1.5 µm defocus.

Infrared Spectroscopy

5 µL of the fibril solution was slowly dried onto a ZnSe single reflection ATR crystal plate for the Smart iTX optical module of the Thermo Fisher Nicolet iS10 FTIR spectrometer.

For each spectrum, 500 scans were taken at room temperature with a nominal spectral resolution of 4 cm^{-1}. The ATR crystal was previously cleaned with water and isopropanol.

Circular Dichroism Spectroscopy

CD spectra of at 40 μM peptides dissolved in PBS buffer were collected at room temperature using a Jasco J-810 CD and 1 mm path length cuvettes. The spectra were collected after dilution from deionized water into buffer.

Crystallography

The peptide αAm_{mem} was dissolved at 5mg/ml in 50 mM octyl-beta-glucoside in water, and other water-soluble peptides were dissolved at 10-15 mg/ml in water. The hanging-drop vapor-diffusion method at room temperature was used for crystallization. The crystallization conditions for the different peptides are as follows: (1) αAm_{mem}: 35% MPD, 0.2 M $MgCl_2$ and imidazole 0.1 M pH 8; (2) αAm_G (#1) 0.2 M ammonium acetate, 0.1 M tri-sodium citrate pH 5.6, 30% (w/v) MPD; (3) αAm_G (#2): 40% MDP, 0.2 M K/Na tartrate; (4) αAm_A: 45% MPD, 0.25 M NaH_2PO_4; (5) αAm_S: 45% MPD, 0.2 M Na formate; (6) αAm_L: 45% MPD, 0.6 M NaH_2PO_4; (7) αTet: 2.0 M Na formate, 0.1 M Na acetate. Crystals were flash frozen with liquid N_2, and data collection temperature was 100 K. No extra/additional cryoprotectant was required for the peptides except αTet, for which 30% glycerol was used. The data of αAm_{mem}, αAm_G and αTet were collected at the Advanced Light Source, Lawrence Berkeley National Laboratory at the Beamline 8.3.1 on a Pilatus3 6M detector with X-ray wavelength of 1.11584 Å. Those for αAm_A, αAm_S, αAm_L and αAm_L-2 were recorded on a Pilatus3 6M detector with wavelength of 1.03320 Å at the Beamline 23ID-D of the Argonne National Laboratory. Data were processed with HKL2000[35] and/or XDS[36] packages. Statistics for data processing and structural refinement were shown in Supplementary Table 1.

Molecular replacement method was used to solve the structures. COOT[37] was used for modeling and rebuilding. For αAm_{mem}, a single helix of the Zn^{2+}-binding helical bundle 4EH1[22] (PDB code: 5WLJ) was truncated to a poly-alanine search model; the searching was done with Phaser[38]. In total 8 copies of helices were located in the asymmetric unit. Different from the antiparallel bundles of 4EH1, however, the peptides were stacked by alternately reversed parallel helical pairs. Due to low resolution, the orientations of helices were determined by refinement with different combinations of helical directions; difference Fourier analyses and identifying interaction between polar residues from neighboring chains played key roles in removal ambiguity. For αAm_L, similarly, the same single helix was used as search model, and all 4 helices were located with Phaser[38]; the αAm_L subunit appeared to be antiparallel bundles, similar to that in 4EH1. Similar strategy was applied in solving and refining the structure of αTet. These structures were refined with REFMAC[39] in the CCP4 packages[40] or Phenix[41].

For αAm_S, the molecular replacement was challenging since there calculated to be ~332 copies of the αAm_S peptides in the asymmetric unit, assuming 50% solvent content in the unit cell. Helical dimers from 4EH1 and the 4-helix bundle from 4EH1 were used to discover potential orientations within the large unit. During the molecular replacement test, among the random outputs, very occasionally solutions were observed in which two

bundles were docked with a relative rotation of ~30-40 degrees along the primary non-crystallographic symmetry (NCS) axis (parallel to cell edge ***c***, calculated by self-rotation function with MOLREP in CCP4[40]). However, the NCS-fold was hard to assign and the distance between these two bundles varied from 18 to 25 Å. The NCS-fold was shown to be compatible with possible numbers from 2 to 12 or even higher, owing to too many molecules in the asymmetric unit. However, the strong NCS signal along ***c*** enabled us a rational scanning on potential molecular packing in this direction. To make facilitate the calculation, the monoclinic unit cell was first re-indexed to switch the ***a*** and ***c*** axes. Molecular replacement was then re-done to acquire the same aforementioned "dimer" (dimer of four-helix bundles) with the rotational NCS along the new "***a***" direction. Thereafter, a single 4-helical bundle of the "dimer" was translated and fixed at the unit-cell origin. A more sophisticated dimeric model was then generated through expanding the model from the first bundle by rotating a 360/n degree (***Chi*** angle in Polar angle convention) plus a translation of ± a/(n*k) [a is the cell constant of new "***a***", n the postulated NCS-fold, k an integer that keeps D in the reasonable interaction distance of two 4-helical bundles and restricts D in the range of 15-25 Å]. When one direction is fixed as the primary NCS direction for expanding model along the new "***a***" axis, two other directions (denoted by two other polar angles, ***Phi*** and ***Psi***) of the first placed bundle at origin were scanned in ± 5 degrees with 1-degree intervals, and this scanning was incorporated into the model expansion and molecular replacement test. The best dimeric model was found with n=10, i.e., the two bundles was related by 36 degrees along the new "***a***" and distance of 19.85 Å. After this, further elongating the model to a string of 5 bundles (still antiparallel αAm_L) resulted in the discovery of 20 bundles with MOLREP; each 10 bundles in a string was exactly fit into a unit distance of the new "***a***". The 10-bundle was then used as model and 4 extra 10-bundles were found (with one bundle overlapped in total). Overall, in total 59 4-helical bundles, or 236 helices, were replaced in the asymmetric unit. After all peptides were placed in ASU, the cells was re-indexed back to the original unit cell convention by switch ***"a"***/***"c"*** again, and the coordinates were transformed accordingly. During refinement, a half number of the helices were revealed to be reversed and were corrected. During the refinement thereafter, the NCS was turned off, as a 236-fold of NCS simply slowed down the refinement for more than 10 folds. Rigid body and TLS domains were reduced down to single helices. The refinement was performed with Phenix[41] for its great advantage in handling 236 chains.

.

For the αAm_G ($P4_322$ space group), the molecular replacement was started with the antiparallel αAm_L bundle as model too (in parallel to the αAm_S project). In total 4.5 bundles (18 helices) were found. Similarly, half number of αAm_G were revealed to be reversed. The refinement was done with REFMAC[39]. For the αAm_G in another space group ($P6_122$), the previous solution (18 peptides) was used as model, and the exactly same peptide content was located in the asymmetric unit. The refinement strategy was similarly used as mentioned above.

For αAm_A, the data were constantly restricted to ~4.0 Å. Difficulty persists when removing the huge bias and ambiguity of MR solutions, and we are struggling for data of better quality.

Cellular Assay

Each αAm peptide was cloned into a pcDNA3-EGFP vector which contains EGFP. The sequence of each αAm peptide was inserted at the 3' end of EGFP gene with a long flexible linker containing Gly, Ala and Ser repeats (GSGSAGGSAGGSAGGSAGGSAGGSAGGSAGGS). The fusion proteins (EGFP-linker-αAm) were expressed in HEK 293T cells by transient transfection. HEK 293T cells were transfected 48 hr before imaging and photobleaching. UV light (405 nm) was applied at 80% power to each sample at 20X. Samples were illuminated for 1300 milli-second to bleach GFP fluorescence in the regions of interest (ROIs). All samples were imaged for two frames before photobleaching, and were monitored afterwards for 10 min with a 5-sec interval. Images were captured by an Olympus FV3000 inverted confocal laser scanning microscope. The intensity of 488 nm fluorescence in ROIs (with background fluorescence subtracted) was measured using NIH imageJ. Values in **Fig. 5** were plotted as percentage of the first frame before photobleaching.

Data availability

The datasets in the current study are available from the corresponding authors upon request.

References

34. Kuhlman, B. & Baker, D. Native protein sequences are close to optimal for their structures. *Proc Natl Acad Sci U S A* **97**, 10383-10388 (2000).
35. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* **276**, 307-326 (1997).
36. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* **66**, 133-144 (2010).
37. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
38. McCoy, A.J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).
39. Murshudov, G.N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367 (2011).
40. Winn, M.D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242 (2011).
41. Adams, P.D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).

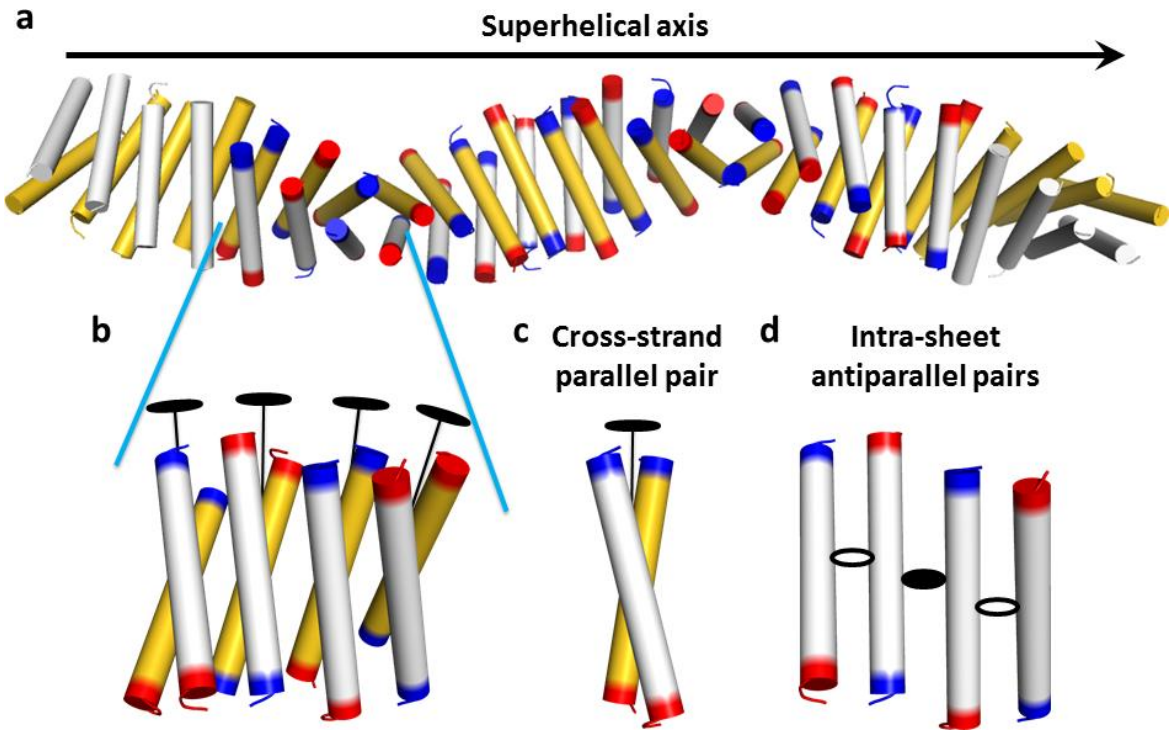

Fig. 1. The amyloid-like structure of αAm$_{mem}$. In (**a**) the superhelical repeat consists of 16 dimers, which are colored blue and red at the N- and C-termini, respectively. The helix dimer subunits progress perpendicular to the superhelical axis. The crystallographic asymmetric unit (**b**) consists of two twisted sheets of helical dimers, which are colored white and golden. The basic unit of the whole amyloid-like assembly is a pseudo-2-fold symmetric cross-strand parallel helix dimer, with 2-fold rotational axis running along the long axis of the dimer and indicated by a black oval and line (**c**). In each sheet, there are two different types of 2-fold symmetric antiparallel helix pairs (shown in more detail in **Fig. 2**). These pairs have approximate two-fold symmetry, but this time with the two-fold symmetry axis directed between the helices as indicated by open and solid ovals in (**d**). Note that the non-crystallographic pseudo-symmetry axes shown in **b-d** are all directed orthogonal to the main superhelical axis of the overall structure shown in (**a**).

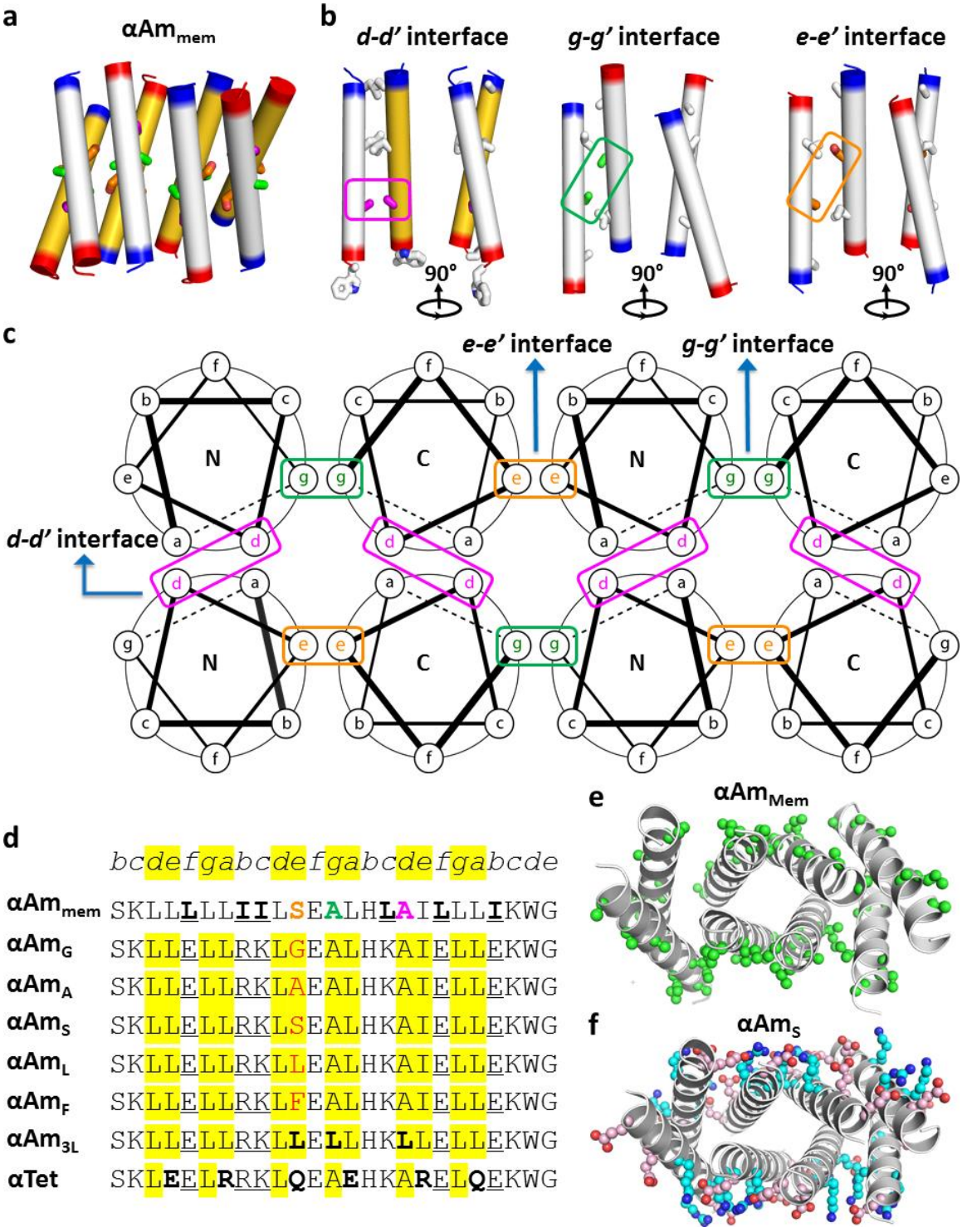

Fig. 2. Design of cross-α amyloid-like assembly. Three types of helix-helix interfaces with small-residue (as sticks) packing exist in the αAm$_{mem}$ amyloid-like structure shown in (**a**). Ala17 (magenta), Ala13 (green) and Ser11 (carbon colored orange) as small residues are involved in the interhelical packing with larger hydrophobic residues (carbon as white

sticks) occurring positions filling the space as the helices diverge from the point of closest approach near the small residues. Their interfaces are designated as *d-d'*, *g-g'* and *e-e'* interfaces, according to the helix wheels in (**c**) for the amyloid-like structure with a parallel dimer as the subunit (the orientation of the N- versus C-terminus towards the viewer is denoted as N and C, respectively). The small residues and the corresponding interfaces in the helix wheels are boxed in (**b**) and (**c**), respectively. The designed sequences intended to form water-soluble amyloid-like structures is compared to αAm_{mem} in (**d**). Panel **e** and **f** show the sequence changes between the crystal structures of αAm_{mem} vs. $\alpha Am_{S,}$ respectively in ball-in-sticks. Their sequences are designed by keeping the residues at the hydrophobic core intact, but modifying the residues (underlined positions) facing the solvent with charged residues for enhanced electrostatic and hydrogen-bonded interactions. The hydrophobic residues on the surface of αAm_{mem} are colored green, while the designed charged residues at the same locations of αAm_S are colored cyan and pink for positively and negatively charged, respectively. As shown in (**d**), the residue on position 11 in the *e-e'* interface is varied to examine its size effect, as shown in red in $\alpha Am_{G,}$ $\alpha Am_{A,}$ αAm_S, αAm_L and $\alpha Am_{F.}$ The effects of varying three small residues to Leu is tested by αAm_{3L}. A non-aggregating water-soluble αTet is also designed.

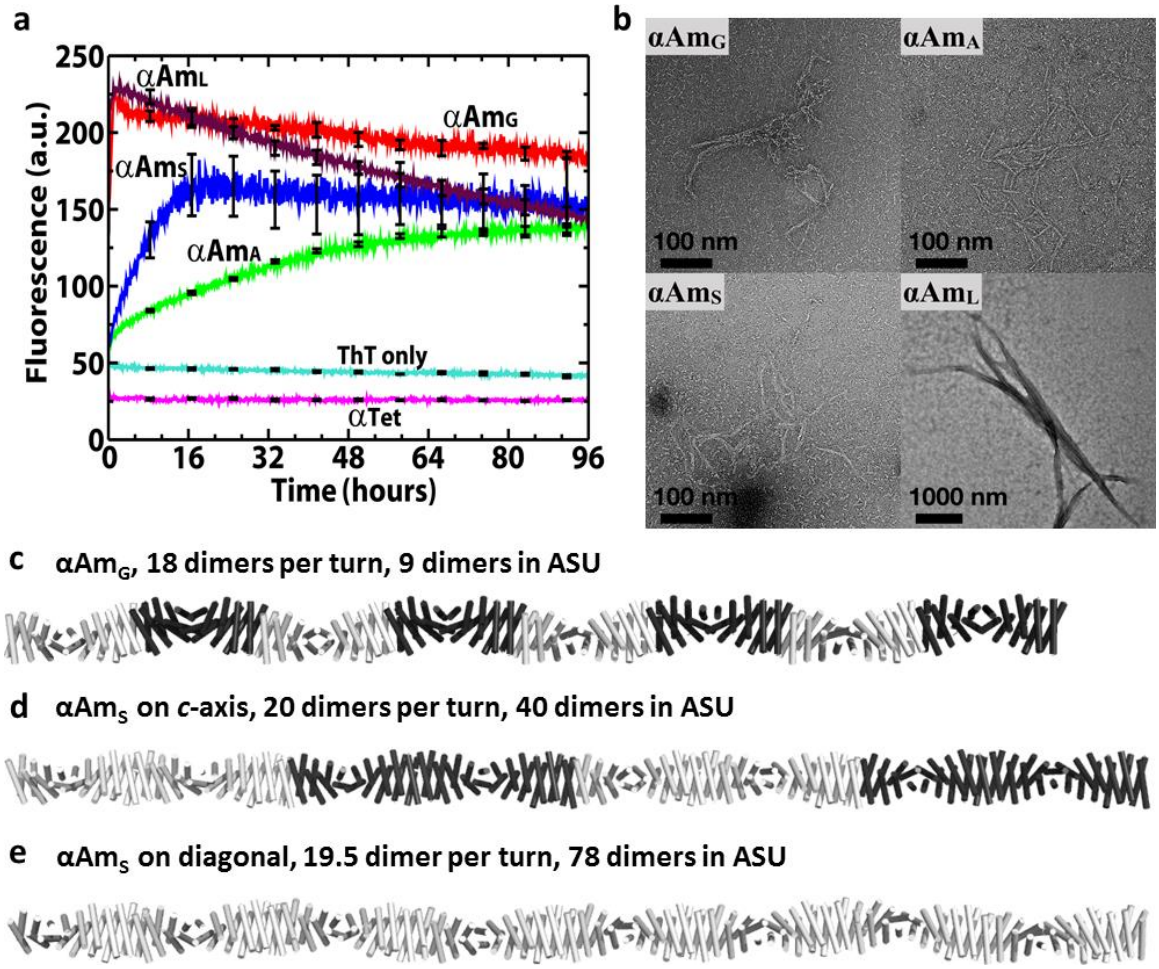

Fig. 3. **The aggregation behavior and fibril formation of the water-soluble peptides and the crystal structures of the cross-α amyloid-like fibrils.** The aggregation kinetics of the peptides monitored by ThT fluorescence is shown in (**a**). The standard errors of 3 replicates are shown by error bars. Negative-stain EM images of the fibrils formed by the designed peptides αAm_G, αAm_A, αAm_S, and αAm_L are displayed in (**b**). Four turns of the αAm_G, αAm_S on the crystallographic *c*-axis and αAm_S on the diagonal in the unit cell are shown in (**c**), (**d**) and (**e**), respectively. In (**c**) eight repeats of the peptide assembly of αAm_G in the asymmetric unit are shown; in (**d**) four repeats of 20 dimers of αAm_S on the *c*-axis are shown; in (**e**) a single repeat of 78 dimers on the diagonal in the asymmetric unit is shown.

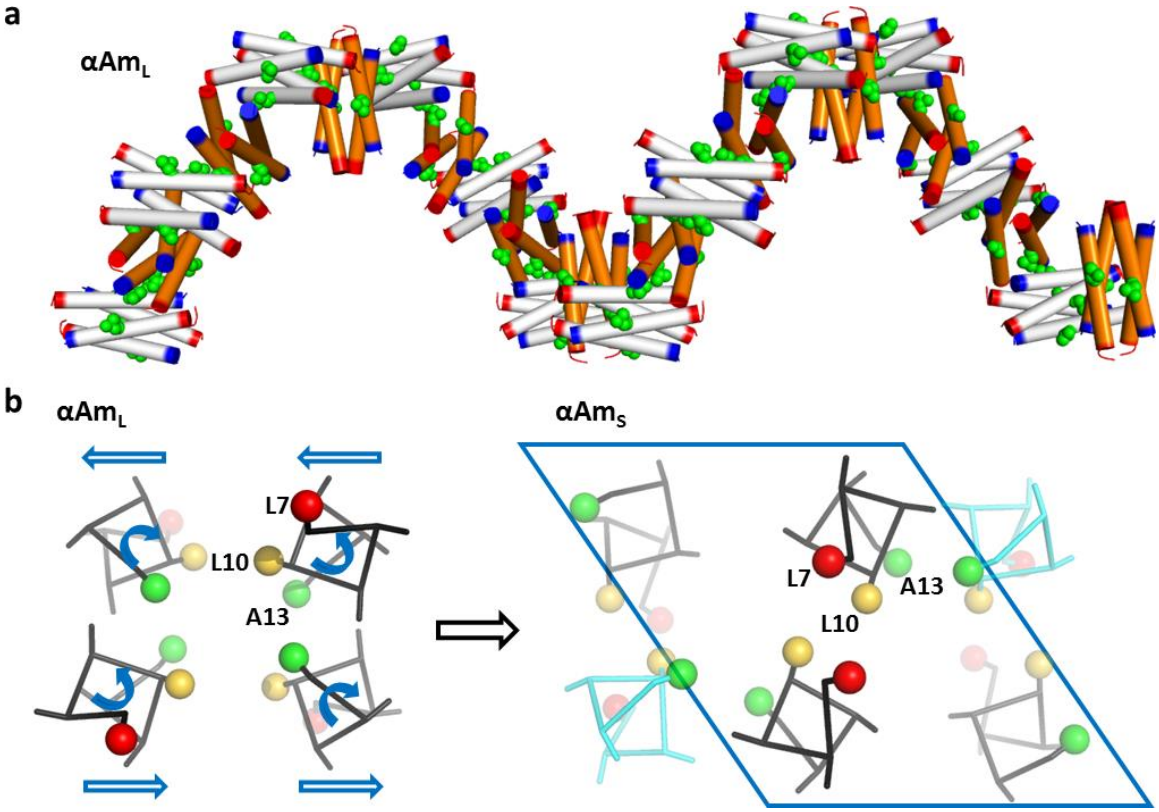

Fig. 4. Packing of αAm$_L$ as a helical polymer composed of helix tetramers and its configurational relationship with the cross-α amyloid-like assembly of αAm$_G$ and αAm$_S$. (**a**) Two turns of αAm$_L$ superhelices; the N- and C-termini of the peptides are colored blue and red, respectively, and the sidechain of Leu11 is colored green. The helical tetramer subunits are colored white and orange alternately. (**b**) The geometric relationship between the structure of an αAm$_L$ tetramer, and the cross-α spiral of αAm$_S$. The left panel shows a slice of the crystal structure of αAm$_L$, which forms a canonical four-helix bundle**,** which assemble into the helical polymer shown in (**a**). Thus, the repeating unit αAm$_L$ is a tetramer, instead of the parallel dimer seen in αAm$_S$. The tetramer of αAm$_L$ can be converted to the repeating structure of αAm$_S$ by the indicated rotations of the helices in αAm$_L$ by 48° and the translation of the helices by approximately 5 Å as indicated. In both αAm$_L$ and αAm$_S$, the balls in red, yellow and green are residues Leu7, Leu10 and Ala13, respectively. The two helices uninvolved in configurational transformation in αAm$_S$ are colored cyan.

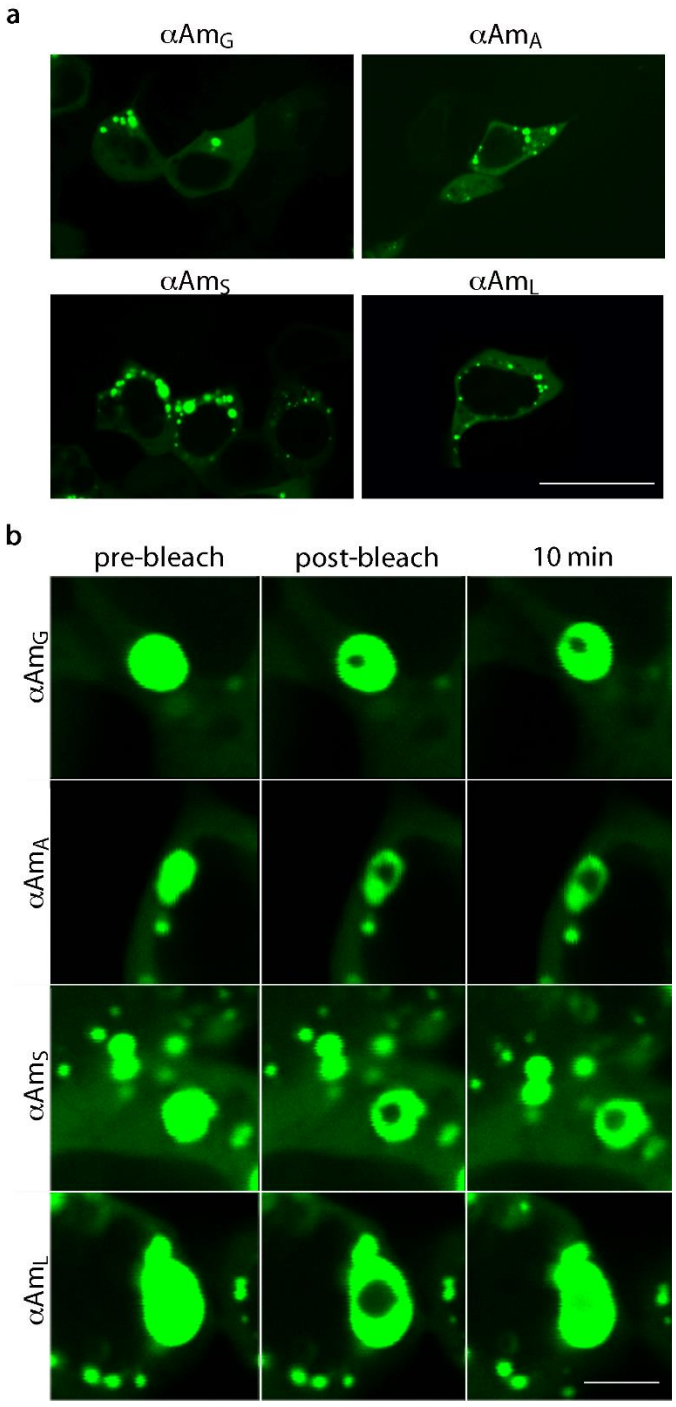

Fig. 5. EGFP tagged αAm peptides form inclusions in the cytosol of mammalian cells; αAm_G, αAm_A and αAm_S form a solid-like phase, while αAm_L is more mobile. (**a**) Fluorescent images of HEK 293T cells expressing EGFP-fused αAm peptides. Scale bar is 50 μm. (**b**) Confocal images show droplets of αAm_G, αAm_A, αAm_S and αAm_L in HEK 293T cells before and after photobleach. Scale bar is 5 μm.