

UCLA

UCLA Electronic Theses and Dissertations

Title

Structural Studies of Amyloid Fibrils and Oligomers

Permalink

<https://escholarship.org/uc/item/41c9x695>

Author

Boyer, David Robert

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Structural Studies of Amyloid Fibrils and Oligomers

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Biochemistry, Molecular and Structural Biology

by

David Robert Boyer

2021

© Copyright by
David Robert Boyer
2021

ABSTRACT OF THE DISSERTATION

Structural Studies of Amyloid Fibrils and Oligomers

by

David Robert Boyer

Doctor of Philosophy in Biochemistry, Molecular and Structural Biology

University of California, Los Angeles

Professor David S. Eisenberg, Chair

Abstract: Amyloid proteins play a critical role in both health and disease. Their unique fibrillar structure – termed the cross- β fold – is adopted by proteins ranging from the melanin-storing pre-melanosomal protein to the familiar amyloid- β plaques associated with Alzheimer’s Disease. Amyloid proteins also form smaller, non-fibrillar oligomers that are implicated in the pathology of numerous amyloid diseases. Despite years of effort to visualize the atomic structures of full-length amyloid proteins, their structures have evaded traditional structural biology techniques such as X-ray crystallography. It is only recently through the development of new techniques such as solid-state biomolecular NMR and cryo-electron microscopy with direct electron detectors that we have been able to visualize the atomic structures of amyloid fibrils. Using these techniques we can answer outstanding questions in amyloid structural biology such as, what are the stabilizing interactions in amyloid fibrils? What are the effect of hereditary mutations found in amyloid protein sequences on their fibril structures? Why do amyloid proteins only grow indefinitely in their length but not their width? What are the non-fibrillar, oligomeric structures of amyloid proteins? In this dissertation, I apply X-ray crystallography, cryo-electron microscopy, and computational structural analysis to answer these questions. In Chapter 2, I

propose a model for previously unseen long-range interactions between the far N-terminus of the protein tau with the fibril core in Alzheimer's Disease Paired Helical Filaments. In Chapters 3 and 4, I determine the structures of the amyloid fibrils of α -synuclein containing Parkinson's Disease hereditary mutation H50Q and Lewy Body Dementia hereditary mutation E46K, respectively, to investigate the role of hereditary disease mutations in modulating α -synuclein fibril structure. In Chapter 5, I analyze all the known amyloid fibril structures determined by cryo-electron microscopy to date in order to show how the helical properties and unique fold of amyloid fibrils place an upper limit on the width of the fibrils, which blocks them from growing in directions other than the fibril axis. In Chapter 6, I study the pre-fibrillar, oligomeric particles of amyloid- β S26s – a modified form of the amyloid- β peptide found in the amyloid plaques of Alzheimer's Disease – in order to understand what types of structures other than the cross- β fold could exist for non-fibrillar assemblies of amyloid proteins.

The dissertation of David Robert Boyer is approved.

Jose A. Rodriguez

Hong Zhou

Todd O. Yeates

David S. Eisenberg, Committee Chair

University of California, Los Angeles

2021

This dissertation is dedicated first and foremost to my parents: Mary Dallman Boyer, M.D. and Robert James Boyer, M.D. Not only are they responsible for giving me life, but for helping me to keep it when I needed help most. My father's constant search for and grasp of a vast wealth of knowledge and my mother's unrelenting desire to help others through medical practice as well as everyday kindnesses has hopefully been passed to me without too much dilution, allowing me to succeed as a scientist as well as to just be a good person. I also dedicate this dissertation to the swarm of extended family members arising from my maternal grandmother's two brothers and her own seven children and 16 grandchildren who have all left an indelible mark on me. I could not ask for a more supportive network of family. I also dedicate this dissertation to my partner and best friend, Natasha Mynhier. Her positivity, creativity, and support enriches my life by orders of magnitude and promises a lifetime full of excitement and happiness. Lastly, I dedicate this dissertation to all the scientists who push and pull against the boundaries of human knowledge, bringing new discoveries to life and enabling mankind's collective evolution.

TABLE OF CONTENTS

I.	Chapter 1: Introduction	1
	a. Overview	
	b. Methods to study the atomic structures of amyloid proteins	
	c. Important questions in amyloid structural biology and scientific contribution of dissertation research	
	d. Contributions to dissertation research	
	e. References	
II.	Chapter 2: A structure-based model for the electrostatic interaction of the N-terminus of protein tau with the fibril core of Alzheimer's Disease filaments	11
	a. Abstract	
	b. Introduction	
	c. Results	
	d. Discussion	
	e. Methods	
	f. Acknowledgments	
	g. References	
III.	Chapter 3: Structures of fibrils formed by α -synuclein hereditary disease mutant H50Q reveal new polymorphs	31
	a. Abstract	
	b. Introduction	
	c. Results	

- i. Cryo-EM structure and architecture of H50Q α -syn fibrils
 - ii. Differences between protofilaments A and B
 - iii. Cavities in α -syn fibril structures
 - iv. Wide fibril protofilament interface
 - v. Energetic and biochemical analysis of H50Q fibrils
 - d. Discussion
 - e. References
 - f. Extended Data
 - i. Methods
 - ii. References
 - iii. Acknowledgements
 - iv. Author contributions
 - v. Competing Interests
- IV. Chapter 4: The α -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure 55
- a. Abstract
 - b. Introduction
 - c. Results
 - i. Cryo-EM structure determination and Architecture of E46K α -Syn Fibrils
 - ii. Comparison of E46K and Wild-Type α -Syn Protofilament Folds
 - iii. Electrostatic Zippers Constitute the E46K Protofilament Interface
 - iv. E36K Fibrils Are More Pathogenic than Wild-Type
 - d. Discussion

- i. Structural Differences Help Explain Enhanced Pathogenicity of E46K α -Syn Fibrils
 - ii. Amyloid Polymorphism
 - iii. Kinetic Factors That Influence Amyloid Structure
 - iv. Growth Conditions That Influence Amyloid Structure
 - v. The Folding Landscape of Amyloid Proteins
 - e. Conclusion
 - f. Materials and Methods
 - g. Acknowledgments
 - h. References
 - i. Supporting Information
- V. Chapter 5: Why amyloid fibrils have a limited width 74
 - a. Abstract
 - b. Introduction
 - c. Results
 - d. Discussion
 - e. Author Contributions
 - f. Acknowledgements
 - g. Materials and Methods
 - h. References
 - i. Supplementary Information
 - j. Supplementary References

VI.	Chapter 6: Towards cryo-EM structure determination of amyloid- β S26s oligomers	128
	a. Abstract	
	b. Introduction	
	c. Results	
	d. Discussion	
	e. Materials and Methods	
	f. Author Contributions	
	g. References	
	h. Supplementary Information	

LIST OF FIGURES AND TABLES

I.	Chapter 2	
	a. Figure 1: Crystal structure of tau N-terminal segment $_{5RQEFEV_{10}}$	21
	b. Figure 2: Crystal structure of $_{5RQEFEV_{10}}$ reveals a wet and a dry interface	22
	c. Figure 3: Speculative model for $_{5RQEFEV_{10}}$ interaction with Alzheimer's Disease paired helical and straight filaments fibril cores	23
	d. Table 1: Data collection and refinement statistics	25
II.	Chapter 3	
	a. Figure 1: Comparison of wild-type and H50Q polymorphs	32
	b. Figure 2: Cryo-EM structures of H50Q polymorphs	33
	c. Table 1: Cryo-EM data collection, refinement and validation statistics	34
	d. Figure 3: Comparison of protofilaments A and B	35

e.	Figure 4: Solvation energy maps and biochemical characterization of H50Q and wild-type α -synuclein	36
f.	Table 2: Comparative solvation energy calculations	37
g.	Extended Data Fig. 1: Fourier Shell Analysis	42
h.	Extended Data Fig. 2: Cryo-EM images and processing	43
i.	Extended Data Fig. 3: Speculative Atomic Models for Islands 1 and 2	45
j.	Extended Data Fig. 4: Alternate conformations of K58 and T59 and potential solvent molecules in the α -syn β -arch cavity	47
k.	Extended Data Fig. 5: PreNAC homozipper Island 1 model and additional solvation energy maps	48
l.	Extended Data Fig. 6: Comparison of α -syn protofilament interfaces	49
m.	Extended Data Fig. 7: H50Q disrupts the wild-type rod polymorph preNAC protofilament interface	50
n.	Extended Data Fig. 8: H50Q fibrils disrupt PC12 cell membranes more than WT fibrils	52
o.	Extended Data Fig. 9: Structural alignment of different wild-type and mutant α -syn polymorphs	53
p.	Extended Data Fig. 10: Schematic illustrating possible secondary nucleation of Protofilament B by Narrow Fibrils	54
III.	Chapter 4	
a.	Figure 1: The E46K mutation leads to a repacked protofilament fold	56
b.	Figure 2: Overview of E46K cryo-EM structure	57
c.	Table 1: Cryo-EM data collection, refinement, and validation statistics	58

d.	Figure 3: Electrostatic residues and interactions of wild-type and E46K fibrils	59
e.	Table 2: Comparative solvation energies	60
f.	Figure 4: Comparison of protofilament interfaces of E46K and wild-type fibrils	61
g.	Figure 5: Biochemical analysis of E46K vs. wild-type fibrils	62
h.	Figure 6: Proposed misfolding landscape of α -syn rod and compact polymorphs	62
i.	Supplementary Figure 1: Schematic representation and free energy of stabilization maps for the wild-type and E46K polymorphs	66
j.	Supplementary Figure 2: Cryo-EM data collection and processing	68
k.	Supplementary Figure 3: Atomic model and density map highlight resolution of reconstruction and hydrogen bonding networks	69
l.	Supplementary Figure 4: Pairwise interaction analysis of wild-type and E46K protofilament folds	70
m.	Supplementary Figure 5: Cross-seeding of wild-type α -syn by E46K fibrils and SDS fibril stability assay	72
IV.	Chapter 5	
a.	Figure 1: Comparison between A-beta 20-34 isoasp23 crystal structure and A-beta 1-40 fibril structure	77
b.	Figure 2: The relationship between subunits in a helix can be described by an azimuthal twist around the helical axis and a translational rise along the helical axis	79

c. Figure 3: Simulated $i, i+1$ distances versus radial distance from the helical axis for different twist values	80
d. Figure 4: A theoretical amyloid fibril built by adding strands of peptide SVQIVY laterally outward from the helical fibril following the crystallographic symmetry of the SVIQVY peptide and then helicizing the model with a 1° /subunit twist and 4.8 Å rise	82
e. Figure 5: Plot of pitch versus radius for 57 cryo-EM structures demonstrates the general rule that fibrils with a larger radius have a larger pitch	84
f. Figure 6: Plots of $i, i+1$ distance for symmetrically related atoms (left), beta-sheet backbone hydrogen bond lengths (middle), and off-axis tilts (right) for four of the highest resolution cryo-EM fibril structures	87
g. Figure 7: Examples of fibril structures that do not display the expected trends in beta-sheet hydrogen bond length and off-axis tilt	89
h. Figure 8: Real space refinement of PDB 6lni with hydrogen bond length and angle restraints restores predicted trends	91
i. Supplementary Table 1: Q-score analysis of all 55 cryo-EM structures	107
j. Supplementary Figure 1: Plot of reported resolution from publication versus expected resolution from Q-score analysis (Q-score = $-0.178 * (\text{Expected Resolution}) + 1.119$)	108
k. Supplementary Figure 2: Plots of fibril structures showing $i, i+1$ distance between symmetry related atoms, backbone hydrogen bond distances and off-axis tilts between $i, i+1$ beta-strands	120
l. Supplementary Table 2: Information for all structures used in analysis	121

V.	Chapter 6	
a.	Figure 1: Initial Abeta S26s oligomer growth and cryo-EM data collection and processing	133
b.	Figure 2: Initial characterization of antibody binding and attempt at data collection of Abeta S26s oligomer-Fab 3D6 particles	136
c.	Figure 3: Seeded oligomer growth and improved Fab 3D6 labeling of oligomers	141
d.	Figure 4: Cryo-EM grid preparation and data processing of Abeta S26s oligomer-Fab 3D6 complex	143
e.	Supplementary Figure 1: Comparison of oligomer pulldown experiments	155
f.	Supplementary Figure 2: Cryo-EM images of A-beta S26s fibrils after extended incubation shows highly polymorphic fibrils	157

ACKNOWLEDGEMENTS

I acknowledge and thank the UCLA/NIH Cellular and Molecular Biology Training Grant and the National Science Foundation Graduate Research Fellowship for providing support during my graduate studies. Chapter 2 is a version of the bioRxiv pre-print: David R. Boyer and David S. Eisenberg. “A structure-based model for the electrostatic interaction of the N-terminus of protein tau with the fibril core of Alzheimer’s Disease filaments.” *bioRxiv* 484279; doi: <https://doi.org/10.1101/484279>. Chapter 3 is a reprint of the published article: Boyer DR, Li B, Sun C, Fan W, Sawaya MR, Jiang L, Eisenberg DS. “Structures of fibrils formed by α -synuclein hereditary disease mutant H50Q reveal new polymorphs.” *Nat. Struct. Mol. Biol.* 2019 Nov;26(11):1044-1052. doi: 10.1038/s41594-019-0322-y. Chapter 4 is a reprint of the published article: Boyer DR, Li B, Sun C, Fan W, Zhou K, Hughes MP, Sawaya MR, Jiang L, Eisenberg DS. “The α -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure.” *Proc. Natl. Acad. Sci. U S A.* 2020 Feb 18;117(7):3592-3602. doi: 10.1073/pnas.1917914117. All co-author contributions for Chapters 2-6 are mentioned in Chapter 1 as well as in the text of Chapters 2-6.

VITA

Education

University of Michigan
B.S. Biochemistry *with High Honors*

Sept. 2010 – Dec. 2014

Publications

1. Cao, Q. **Boyer, D.R.** Sawaya, M.R., Abskharon, R., Saelices, L., Lu, J., Lin, H., Kandeel, F., Eisenberg, D.S. “Cryo-EM structures of hIAPP fibrils seeded by patient-extracted fibrils reveal new polymorphs and conserved fibril cores.” *Submitted*.
2. Tayeb-Fligelman, E., Cheng, X., Tai, C., Bowler, J.T., Griner, S., Sawaya, M.R. Seidler, P.M. Jiang, Y.X., Lu, J., Rosenberg, G.M., Salwinski, L., Abskharon, R. Zee, C., Hou, K., Li, Y., **Boyer, D.R.**, Murray, K.A. Falcon, G., Anderson, D.H., Cascio, D., Saelices, L., Damoiseaux, R., Guo, F., Eisenberg, D.S. “Inhibition of amyloid formation of the Nucleoprotein of SARS-CoV-2.” *bioRxiv*. <https://doi.org/10.1101/2021.03.05.434000>.
3. Shipp, C. Kelly, R.H., Dahl, P.J., Yi, S.M., Vu, D. **Boyer, D.R.**, Glynn, C. Sawaya, M.R., Eisenberg, D.S. Batista, V.S. Malvankar, M.S. “Intrinsic electronic conductivity of individual atomically resolved amyloid crystals reveals micron-long hole hopping via tyrosines.” *PNAS*. Jan 2021, 118 (2).
4. Lu, J., Cao, Q., Hughes, M. P., Sawaya, M. R., **Boyer, D. R.**, Cascio, D. & Eisenberg, D. S. The cryoEM structure of the low-complexity domain of hnRNPA2 and its conversion to pathogenic amyloid. *Nature Communications* (2020) <https://doi.org/10.1038/s41467-020-17905-y>.
5. Seidler, P. M., **Boyer, D. R.**, Sawaya, M. R., Ge, P., Shin, W. S., DeTure, M. A., Dickson, D. W., Jiang, L. & Eisenberg, D. S. CryoEM reveals how the small molecule EGCG binds to Alzheimer’s brain-derived tau fibrils and initiates fibril disaggregation. *bioRxiv* 2020.05.29.124537 (2020) doi:[10.1101/2020.05.29.124537](https://doi.org/10.1101/2020.05.29.124537).
6. Cao, C., **Boyer, D. R.**, Sawaya, M. R., Ge, P., Eisenberg, D. S. Cryo-EM structure and inhibitor design of human IAPP (amylin) fibrils. *Nature Structural and Molecular Biology* (2020) <https://doi.org/10.1038/s41594-020-0435-3>.
7. **Boyer, D. R.***, Li, B.*, Sun, C., Fan, W., Zhou, K., Hughes, M. P., Sawaya, M. R., Jiang, L. # & Eisenberg, D. S. # The α -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. *PNAS* (2020) doi:10.1073/pnas.1917914117.
8. **Boyer, D. R.***, Li, B.*, Sun, C., Fan, W., Sawaya, M. R., Jiang, L.# & Eisenberg, D. S.# Structures of fibrils formed by α -synuclein hereditary disease mutant H50Q reveal new polymorphs. *Nature Structural and Molecular Biology*. 26, 1044–1052 (2019) doi: <https://doi.org/10.1038/s41594-019-0322-y>.
9. Seidler, P.M., **Boyer, D.R.**, Murray, K.A., Yang, T.P., Bentzel, M., Sawaya, M.R., Rosenberg, G., Cascio, D., Williams, C.K., Newell, K.L., Ghetti, B., DeTure, M.A., Dickson, D.W., Vinters, H.V., and Eisenberg, D.S. Structure-based inhibitors halt prion-like seeding by Alzheimer’s disease- and tauopathy-derived brain tissue samples. *Journal of Biological Chemistry* (2019) doi: 10.1074/jbc.RA119.009688.
10. Wormack RA, **Boyer DR**, Zee T, Richards L, Sawaya MR, Cascio D, Gonen T, Eisenberg DS, Clarke SG. Structure of amyloid- β (20-34) with Alzheimer’s-associated isomerization at Asp23 reveals a distinct protofibril interface. *Nature Communications* (2019).
11. Cao Q*, **Boyer DR***, Ge P, Sawaya, MR, Eisenberg, DS. Cryo-EM structures of Four Polymorphic TDP-43 Amyloid Cores. *Nature Structural and Molecular Biology* (2019).

12. **Boyer DR** and Eisenberg DS. A structure-based model for the electrostatic interaction of the N-terminus of protein tau with the fibril core of Alzheimer's disease filaments. *BioRxiv* (2018) <https://www.biorxiv.org/content/10.1101/484279v1>
13. Gallagher-Jones M, Ophus C, Bustillo KC, **Boyer DR**, Panova O, Glynn C, Zee T, Ciston J, Mancina KC, Minor AM, Rodriguez JA. Nanoscale mosaicity revealed in peptide microcrystals by scanning electron nanodiffraction. *Communications Biology* 2, (2019).
14. Li B*, Ge P*, Murray KA*, Sheth P, Zhang M, Nair G, Sawaya MR, Shin WS, **Boyer DR**, Ye S, Eisenberg DS#, Zhou ZH#, Jiang L#. Cryo-EM of full-length α -synuclein reveals fibril polymorphs with a common structural kernel. *Nature Communications* 9, (2018).
15. Hughes MP, Sawaya MR, **Boyer DR**, Goldschmidt L, Rodriguez JA, Cascio D, Chong L, Gonen T, Eisenberg DS. Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks. *Science* 359, 698–701 (2018).
16. Seidler, P.M., **Boyer DR**, Rodriguez, J.A., Sawaya, M.R., Cascio, D., Murray, K.A., Gonen, T., Eisenberg, D.S. Structure-based inhibitors of tau aggregation. *Nature Chemistry* 10, 170–176 (2018).
17. Gallagher-Jones J*, Glynn C*, **Boyer DR**, Martynowycz MW, Hernandez E, Miao J, Zee CT, Novikova IV, Goldschmidt L, McFarlane HT, Helguera GF, Evans JE, Sawaya MR, Cascio D, Eisenberg DS, Gonen T, Rodriguez JA. Sub-ångström cryo-EM structure of a prion protofibril reveals a polar clasp. *Nature Structural & Molecular Biology* 25, 131–134 (2018).
18. Hanczyc P, Mikhailovsky A, **Boyer DR**, Sawaya MR, Heeger A, Eisenberg D. Ultrafast Time-Resolved Studies on Fluorescein for Recognition Strands Architecture in Amyloid Fibrils. *Journal of Physical Chemistry B* 122, 8–18 (2018).
19. Guenther EL, Ge P, Trinh H, Sawaya MR, Cascio, **Boyer DR**, Gonen T, Zhou ZH, Eisenberg DS. Atomic-level evidence for packing and positional amyloid polymorphism by segment from TDP-43 RRM2. *Nature Structural & Molecular Biology* 25, 311–319 (2018).
20. Guenther EL*, Cao Q*, Trinh H, Lu J, Sawaya MR, Cascio D, **Boyer DR**, Rodriguez, Hughes MP, Eisenberg DS. Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. *Nature Structural & Molecular Biology* 1 (2018). doi:10.1038/s41594-018-0064-2
21. Lin F, Ferguson KL, **Boyer DR**, Lin XN, Marsh EN. Isofunctional enzymes PAD1 and UbiX catalyze formation of a novel cofactor required by ferulic acid decarboxylase and 4-hydroxy-3-polypropenylbenzoic acid decarboxylase. *ACS Chem Biol*. 2015 Apr 17;10(4):1137-44. PubMed PMID: [25647642](https://pubmed.ncbi.nlm.nih.gov/25647642/).

*Denotes equal contribution. #Co-corresponding author.

Patents

1. Eisenberg, D.S., Seidler, P.M., Griner, S., Murray, K.A., **Boyer D.R.** Provisional Patent Application. Peptide-based Inhibitors that Block Aggregation, Seeding and Cross-seeding of Amyloid B and Tau. *United States US 62/848,738*. 2019 May 05.
2. Seidler, P.M., **Boyer, D.R.**, and Eisenberg, D.S. Patent. Structure-based peptide inhibitors that target the Tau VQIINK fibrillization segment. *United States US 10,934,322*. Mar. 2, 2021
3. Minty JJ, Singer ME, Lin XN., **Boyer, D.R.** Patent Application. Compositions and Methods for Microbial Co-culture. *United States US 16/304,336*. 2017 May 26.
4. Nelson MC, Minty JJ, Kistler S, Gao Y, Wang HY, Lin XN, **Boyer DR.** Patent Application. Compositions and Methods for Generation of Biofuels. *United States US 14/931,107*. 2015 November 05.

CHAPTER 1: INTRODUCTION

A. Overview

Amyloid proteins occupy a unique niche in the world of protein structure and function¹. Usually, protein sequences fold into three-dimensional shapes that can be as drastically different from one another as the tasks that the proteins perform; on the other hand, amyloid proteins fold using a modular architecture called the cross- β fold that many protein sequences can adopt. The cross- β fold consists of two-dimensional layers of protein chains that stack upon one another indefinitely to form an extended, unbranched fibril – the amyloid fibril. The protein chains within each layer invariably have largely β -strand secondary structures where the backbone carbonyl oxygen and nitrogen atoms of the constituent amino acids point up or down the fibril axis. When the layers stack upon each other in the fibril, they then form extended, indefinite β -sheets. Furthermore, in each layer, the numerous β -strands within the protein chains interact tightly via their side chains in the direction orthogonal to the fibril axis. The cross- β fold architecture lends remarkable stability to amyloid fibrils often making them immune to the usual denaturants used to unfold other protein structures. This stability is likely the reason why amyloid fibrils are often found to be associated with and perhaps causative of many related diseases – termed amyloid diseases – including Alzheimer’s Disease, Parkinson’s Disease, and Type II Diabetes.

B. Methods to study the atomic structures of amyloid proteins

Although the cross- β fold of amyloid fibrils and their presence in numerous diseases has been known for more than 100 years, it is not until recently that we have been able to study their atomic structures at the same level of detail as other non-amyloid proteins. This is due to the main method for determining the atomic structures of proteins since the 1950’s: X-ray

crystallography. As its name implies, this technique relies on the formation of crystals, which are composed of ordered assemblies of many identical copies of a molecule. However, since the structure of the molecules in amyloid fibrils are defined by their interaction with other molecules in the amyloid fibril (in other words, a protein chain with a specific structure in the amyloid fibril would not hold its shape if it were taken out of the fibril), amyloid fibrils themselves would have to be crystallized in order to utilize X-ray crystallography to determine the structure of the protein chains in the fibril. This presents serious obstacles to crystallization as amyloid fibrils have flexible mechanical properties and can form non-identical structures from the same protein sequence (polymorphs) in the same fibril preparation. In addition to these intuitive reasons why amyloid fibrils oppose crystallization, a more technical reason is that even if an amyloid fibril had crystalline-like rigidity and no polymorphism, their extremely small helical twist leads to large translational repeat distances, which, in a crystal, would lead to an extremely large unit cell dimension. This fact, although not impossible to overcome, would create non-trivial problems in diffraction data collection and processing such as those encountered when studying large viruses².

To overcome the limits of traditional X-ray crystallography to study the atomic structures of amyloid fibrils, several methods have been devised. For years, the most prolific of these strategies is to use short segments of amyloid proteins that, when incubated at high concentrations, form crystals instead of fibrils³. These crystals range in size from less than a micron to centimeters – with electron diffraction being required for the smaller crystals, while traditional or microfocal X-ray sources can be used for the larger crystals. Crystal structures determined by X-ray or electron diffraction of short segments (ranging from 4-15 amino acids) have amyloid fibril-like qualities, including the formation of the cross- β structure where β -

strands stack upon each other to form indefinitely long β -sheets and the β -strands mate with the side chains of adjacent β -strands through a variety of so-called “steric zipper” arrangements. This allows a local view of segments of the amyloid protein that are likely forming important stabilizing interactions in the fibrils formed by full-length proteins. A key difference between these amyloid-like crystals and amyloid fibrils is the lack of a helical twist that allows the translational repeat distance along the fibril or crystal axis to be the distance between neighboring stacked β -strands – usually 4.8 Å – instead of the distance it takes for the helix to complete a full rotation ($\sim 400 - 2000$ Å). Despite the numerous insights that have been derived from crystallography of amyloid protein segments, to date, no full length amyloid protein has been crystallized in either an oligomeric or fibril form.

Solid-state nuclear magnetic resonance (ssNMR) offers another route to the atomic structures of amyloid fibrils⁴. Instead of relying on the crystallization of short segments of amyloid proteins, full-length amyloid fibrils can be studied directly. This is accomplished by forming amyloid fibrils of isotopically labeled proteins and measuring the interactions between the neighboring nuclei of the atoms of the protein composing the fibril. The measurements can be converted to distance restraints that can help to derive plausible atomic models for amyloid fibrils. Importantly, ssNMR was used to derive the first atomic models of full-length amyloid fibrils, including amyloid- β ⁵ and α -synuclein⁶. ssNMR can also be used to detect structural differences between different fibrils formed from the same protein sequence, which enables rapid fingerprinting of fibril structures without structure determination. This was applied to help reveal differences in amyloid- β fibril structures seeded by extracted fibrils from different Alzheimer’s patients – an important discovery that could lead to personalized treatments for Alzheimer’s Disease⁷. Despite its ground-breaking application to reveal structures of amyloid fibrils for the

first time, ssNMR has several limitations for the study of amyloid fibrils; namely, the presence of multiple polymorphs within the same sample would lead to the possible averaging of measurements arising from the same nuclei in different polymorphs, and the amyloid fibrils need to be isotopically labeled to allow measurements for structure determination. The former limitation could be potentially overcome with extensive sample optimization to ensure a single polymorph is present, although this could lead to other, biologically relevant polymorphs being excluded from analysis. The latter limitation presents technical challenges including differential isotopic labeling of segments of the protein chain to allow more accurate measurements, but also fundamental challenges as it disallows the direct visualization of amyloid fibrils directly extracted from patients – an important caveat given that amyloid fibrils formed in vivo have been shown to differ from those formed in vitro^{8,9}.

Cryo-electron microscopy (cryo-EM) has long been used to study biological molecules by taking images of many individual particles and then averaging their images into a three-dimensional picture of the molecule. However, it is not until recently that advances in direct electron detectors have made it feasible to routinely obtain three-dimensional pictures of biological molecules with high enough resolution to determine their atomic structures¹⁰. Currently, state-of-the-art cryo-EM allows the study of flexible and polymorphic amyloid fibrils due to the high volume of images that can be collected on a given sample and the sophisticated classification algorithms that can group similar particles together¹¹. This allows scientists to separate the images of different polymorphs that exist in the same fibril preparation and determine their atomic structures in parallel. Cryo-EM also has the added benefit that it can be readily applied to study amyloid fibrils extracted from patients, if they are extracted in sufficient

quantity¹². In other words, cryo-EM is forgiving of the properties of amyloid fibrils that make X-ray crystallography and ssNMR intractable.

C. Important questions in amyloid structural biology and scientific contribution of dissertation research

The three techniques mentioned above – X-ray crystallography, ssNMR, and cryo-EM – have enabled incredible insights into the amyloid fibril state of proteins from dozens of high-quality atomic structures. Important highlights include learning (i) the interactions that stabilize amyloid fibrils, (ii) the association of different amyloid fibril polymorphs with different amyloid diseases, (iii) the effect of fibril growth conditions on fibril polymorphism, (iv) the structural differences between pathogenic, irreversible fibrils and functional, reversible fibrils, (v) the effect of disease-associated hereditary mutations and post-translational modifications on fibril structure, and (vi) the differences between in vitro and in vivo generated fibrils.

In this dissertation, I attempt to both add understanding in some of these areas (Chapters 2-4), as well as address some unresolved questions. In Chapter 2, based on an X-ray crystal structure I determined of a six residue, N-terminal segment of the protein tau, I propose a model that helps to identify unresolved interactions in the cryo-EM structures of the tau amyloid fibrils extracted from Alzheimer's Disease patients. In Chapters 3 and 4, I demonstrate how the fibrils of α -synuclein bearing the disease-associated hereditary mutations H50Q (Parkinson's Disease) or E46K (Lewy Body Dementia) drastically differ from wild-type fibrils. This pair of chapters sheds light on how hereditary mutations may be promoting disease pathogenesis by creating different, more pathogenic fibrils.

Chapter 5 proposes an answer to the question, why do amyloid fibrils have a limited width? I hypothesize that the helical nature of the amyloid fibril, combined with its cross- β folding pattern, places a limit on the stability of the β -strands located farthest away from the helical axis. I survey all currently known amyloid fibril structures determined by cryo-EM to test this hypothesis. By examining the interactions in the fibril as a function of distance from the helical axis, I uncover a clear trend: the edges of the fibrils have the weakest interactions.

Chapter 6 addresses one of the most outstanding and technically challenging questions in amyloid structural biology: what are the atomic structures of amyloid oligomers? Although the cross- β architecture is ubiquitous for amyloid fibrils, amyloid proteins also form smaller assemblies (~5 to ~100 protein chains), termed oligomers, whose structural architecture is unknown. Nearly all amyloid fibrils are shown to form oligomers either on the pathway from unfolded monomer to amyloid fibril, or so-called “off-pathway” oligomers that do not convert into fibrils, but either are stable as oligomers, or disassemble into monomers. These smaller amyloid entities are thought to potentially be the most toxic amyloid species in the amyloid diseases, adding to the motivation to resolve their structures. I make progress towards determining the oligomeric structure of a modified form of amyloid- β by using an antibody to act as a fiducial marker on the small oligomeric particles to aid cryo-EM image alignment and structure determination.

D. Contributions to Dissertation Research

The research contained in this dissertation is the result of collaborative effort between numerous scientists of different labs and institutions. For the help that is freely given and for the opportunity to participate and help others answer their own research questions, I am eternally

grateful. There is almost no quicker way, in my experience, to learn something than from the generous guidance of someone farther along in a subject than oneself. Just as there is no better way to retain and enrich one's own knowledge, than to share it freely and frequently with others. I list in the following paragraphs my own specific contributions, as well as the contributions of others to the Chapters contained in this dissertation.

In Chapter 2, I identified peptide sRQEFEV_{10} from the ZipperDB database as a potential steric zipper in the tau protein whose crystal structure could be obtained to design new inhibitors of tau aggregation. I crystallized and determined the X-ray crystal structure of sRQEFEV_{10} . Although structure-based inhibitors were never developed against the peptide, I did notice it was predicted to possibly occupy unmodeled density in the Alzheimer's Disease tau amyloid fibrils determined by Fitzpatrick, et al¹². Therefore, I modeled the sRQEFEV_{10} structure into the unresolved density of the Alzheimer's disease paired helical and straight filaments to propose an integrative model using both X-ray and cryo-EM structures that explains how sRQEFEV_{10} could interact with the fibril core. David Eisenberg supervised the project and helped write the paper. My ability to determine the crystal structure of sRQEFEV_{10} was only possible through training on earlier failed and (some!) successful peptide X-ray crystallography projects. This training was imparted to me mainly through Duilio Cascio and Michael Sawaya, our resident masters of X-ray crystallography, during the preparation for and trips to the Argonne National Lab Advanced Photon Source synchrotron.

In Chapters 3 and 4, I worked with a team of scientists from the laboratories of Lin Jiang and David Eisenberg (co-supervisors of the projects). My role in the overall project of determining the structures of α -synuclein fibrils bearing hereditary disease-associated mutations H50Q and E46K was the screening of fibrils for cryo-EM data collection, high resolution Krios

data collection of the fibrils, data processing and structure determination of the two H50Q polymorphs and the one E46K polymorph, structural analysis and figure-making in both Chapters 3 and 4, as well as organizing and writing the chapters. Binsen Li and Chuanqi Sun purified the α -synuclein monomer and performed the fibril growth assays. Binsen Li also performed biochemical characterization of the fibrils including seeding and toxicity assays. Michael Sawaya and David Eisenberg contributed the software for calculating atomic solvation energies. Kang Zhou and Peng Ge aided in collecting high resolution Krios data.

In Chapter 5, I conceived of the idea why amyloid fibrils may have a limited width, namely that the helical arrangement of the fibril will make the distances between β -strands in the β -sheets running up and down the length of the fibril increase proportionally to their distance from the helical axis. This increase in distances will weaken the stabilizing interactions holding together the outside of the fibril. I led the analysis and figure-making of the project, as well as wrote the chapter. Michael Sawaya and Nikos Mynhier contributed software to calculate the hydrogen bond distances between β -strands, as well as calculate the various other parameters shown in Chapter 5 that examine interactions within the fibril as a function of distance from the helical axis.

In Chapter 6, I performed oligomer growth and characterization experiments as well as prepared all samples for EM and performed EM (both negative stain and cryo) data collection and processing. David Eisenberg supervised the project. Luis Alejandro Foley from the lab of Jevgenij Raskatov at UC Santa Cruz synthesized and purified the A- β S26s peptide used in our experiments. For the complexing of oligomers with antibody fragments, Romany Abshkaron was instrumental in identifying, purchasing/purifying, and characterizing the binding of oligomers by antibody fragments. Specifically, Romany identified Fab 3D6 as a potential binder to amyloid- β

S26s oligomers and performed experiments to validate its binding. He also performed all immunoprecipitation experiments that led to EM images of oligomers bound by Fabs. Sean Jiang performed oligomer preparation and purification experiments to identify ideal growth conditions and purification methods for the oligomers. Kang Zhou and Peng Ge assisted in automated negative stain- and cryo-EM data collection. Carter Lantz performed mass spectrometry experiments.

References

1. Eisenberg, D. & Jucker, M. The Amyloid State of Proteins in Human Diseases. *Cell* **148**, 1188–1203 (2012).
2. McPherson, A. & Larson, S. B. A guide to the crystallographic analysis of icosahedral viruses. *Crystallography Reviews* **21**, 3–56 (2015).
3. Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. Ø., Riek, C., Grothe, R. & Eisenberg, D. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
4. Tycko, R. Solid-state NMR studies of amyloid fibril structure. *Annu Rev Phys Chem* **62**, 279–299 (2011).
5. Paravastu, A. K., Leapman, R. D., Yau, W.-M. & Tycko, R. Molecular structural basis for polymorphism in Alzheimer's β -amyloid fibrils. *PNAS* **105**, 18349–18354 (2008).
6. Tuttle, M. D., Comellas, G., Nieuwkoop, A. J., Covell, D. J., Berthold, D. A., Kloepper, K. D., Courtney, J. M., Kim, J. K., Barclay, A. M., Kendall, A., Wan, W., Stubbs, G., Schwieters, C. D., Lee, V. M. Y., George, J. M. & Rienstra, C. M. Solid-state NMR structure of a pathogenic fibril of full-length human α -synuclein. *Nature Structural & Molecular Biology* **23**, 409–415 (2016).

7. Qiang, W., Yau, W.-M., Lu, J.-X., Collinge, J. & Tycko, R. Structural variation in amyloid- β fibrils from Alzheimer's disease clinical subtypes. *Nature* **541**, 217–221 (2017).
8. Zhang, W., Falcon, B., Murzin, A. G., Fan, J., Crowther, R. A., Goedert, M. & Scheres, S. H. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife* **8**, (2019).
9. Lövestam, S., Schweighauser, M., Matsubara, T., Murayama, S., Tomita, T., Ando, T., Hasegawa, K., Yoshida, M., Tarutani, A., Hasegawa, M., Goedert, M. & Scheres, S. H. W. Seeded assembly in vitro does not replicate the structures of α -synuclein filaments from multiple system atrophy. *FEBS Open Bio* **11**, 999–1013 (2021).
10. Kühlbrandt, W. The Resolution Revolution. *Science* **343**, 1443–1444 (2014).
11. Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat Struct Mol Biol* **26**, 619–627 (2019).
12. Fitzpatrick, A. W. P., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., Crowther, R. A., Ghetti, B., Goedert, M. & Scheres, S. H. W. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185–190 (2017).

A structure-based model for the electrostatic interaction of the N-terminus of protein tau with the fibril core of Alzheimer's Disease filaments

David R. Boyer¹ and David S. Eisenberg¹

¹ Department of Chemistry and Biochemistry and Biological Chemistry, UCLA-DOE Institute, and Howard Hughes Medical Institute, UCLA, Los Angeles CA 90095-1570

*Correspondence to: David S. Eisenberg, email: david@mbi.ucla.edu

Abstract:

Although portions of tau protein within the microtubule binding region have been shown to form the ordered core of tau filaments, the structural details of how other regions of tau participate in filament formation are so far unknown. In an attempt to understand how the N-terminus of tau may interact with fibril core, we crystallized and determined the structure of the N-terminal segment ${}^5\text{RQEFEV}_{10}$ of tau. Several lines of evidence have shown the importance of this segment for fibril formation. The crystal structure reveals an out-of-register Class 5 steric zipper with a wet and a dry interface. To examine the possible interaction of ${}^5\text{RQEFEV}_{10}$ with the tau fibril core, we modeled the binding of the wet interface of the ${}^5\text{RQEFEV}_{10}$ structure with the ${}^{313}\text{VDLSKVTSKC}_{322}$ region of the Alzheimer's Disease tau filament structures. This model is consistent with, and helps to explain previous findings on the possible interaction of these two segments, distant in sequence. In addition, we discuss the possible conservation of this interaction across multiple polymorphs of tau.

Introduction:

The aggregation of tau into amyloid fibrils is associated with some 25 neurological diseases, collectively termed tauopathies. Although scientists have for decades associated fibrous tau aggregates with disease for decades, the molecular events driving aggregation of tau into amyloid fibrils remain unknown. It is generally thought that tau remains in three pools in the cell: attached to microtubules to promote their stability(1,2), bound to molecular chaperones to protect nucleating sequences of tau from enabling aggregation(3), or in a fibrous state where each fiber contains hundreds to many thousands of tau molecules(4–6). Under what conditions the fibril state begins to dominate is unclear.

Previous studies have shown that soluble, monomeric tau largely lacks a defined 3-dimensional shape(7); however, other studies posit that tau adopts a “paper clip” conformation in solution(8) or a seed-competent conformation where amyloid nucleating sequences are exposed and able to seed fibril formation(9). In addition, the binding of different tau constructs to microtubules has been visualized by cryo-EM(2). Despite these findings, information on the structure of soluble, monomeric form of tau is limited due to its largely disordered nature; therefore, most structural studies have focused on the aggregated state of tau(10–14). Our laboratory first focused on the segments of tau shown to be essential for *in vitro* aggregation, the primary nucleating sequences VQIINK and VQIVYK, located at the beginning of tau microtubule binding repeats 2 and 3, respectively(15). The crystal structures of these segments revealed classical “steric zipper” structural features(10,11). Mutations to these segments inhibit full-length tau aggregation, and we have shown that inhibitors designed to “cap” the crystal structures

of VQIINK and VQIVYK segments also inhibit full-length tau aggregation, further demonstrating the importance of these segments(11,16).

Recently, cryo-EM studies of extracted tau filaments from Alzheimer's Disease and Pick's Disease patients have revealed several tau fibril polymorphs in near-atomic detail(12–14). In all of these structures, residues 306-378 spanning the length of Repeats 3 and 4 plus an additional six residues to the C-terminus of Repeat 4, are ordered in the fibril core, and in Pick's Disease, Repeat 1 residues 254-274 are also ordered(14). Although these landmark discoveries help illuminate the fold adopted by the microtubule binding region of tau, it is still unknown to what degree other parts of tau participate in the aggregation process.

In the AD fibril structures, there is additional density consistently seen near residues K317 and K321 that may indicate another region of tau is interacting with the fibril core(12,13). Fitzpatrick *et al.* hypothesize that this extra density belongs to the residues ${}_{7}\text{EFE}_9$, an N-terminal sequence of tau that is part of the Alz50/MC-1 antibody binding epitope(12,17). To better understand the potential interaction of the N-terminus and the AD fibril core, we sought to determine the structure of this N-terminal segment.

Results:

We first searched for segments containing ${}_{7}\text{EFE}_9$ that are likely to crystallize. Although no segment containing ${}_{7}\text{EFE}_9$ scored well on the structure-based ZipperDB server(18), the ability to form fibrils from segment ${}_{5}\text{RQEFEV}_{10}$ was previously predicted by a sequence-based method and demonstrated biochemically(19). Therefore, we

crystallized and determined the structure of the hexameric segment $_5\text{RQEFEV}_{10}$ (Figure 1 A-C).

The crystal structure of $_5\text{RQEFEV}_{10}$ revealed a Class 5 homozipper where beta-strands assemble in antiparallel sheets and these sheets mate together in distinct face-to-face and back-to-back interfaces. Notably, the sheets are out-of-register and are related to each other by a 2_1 “fibril axis” (20) (Figure 2 A, B). This combination of symmetry elements produces an $\sim 80^\circ$ crossing angle between strands of one sheet and its mated sheet (Figure 2 A)(21). The alternating sequence of charged and hydrophobic/uncharged residues leads to wet and dry interfaces in the crystal structure.

The wet interface features electrostatic interactions among polar, charged residues and water molecules. In particular, glutamates form an extensive hydrogen bond network with water molecules and arginines originating from the same sheet and from the opposing sheet (Figure 2 B). The dry interface features hydrophobic packing of phenylalanine, glutamine, and valine leading to the exclusion of water (Figure 2 A, B). Also, glutamine side chains clasp each other through a pair of hydrogen bonds, further stabilizing connections between neighboring strands in a sheet (Figure 2 A). This interaction is similar to the polar clasp described by Gallagher-Jones, *et al.*, with the distinction that glutamines in that study originated within the same strand(22). Similar to that polar clasp, neighboring aromatic residues restrict the glutamines to a conformation in which they bond to each other within a hydrophobic pocket (Figure 2 A). As stated by Gallagher-Jones, *et al.* the shielding of glutamines by neighboring aromatic residues may be essential for the formation of this polar clasp.

The crystal structure of ${}^5\text{RQEFEV}_{10}$ can account for the low resolution density found in the cryo-EM reconstructions of Alzheimer's Disease (AD) tau filaments near residues K317 and K321, much as suggested by Fitzpatrick, *et al* (12). The positioning of ${}^5\text{RQEFEV}_{10}$ near these residues in the tau filament conformation is supported by the binding of the MC-1 and Alz50 antibodies to a discontinuous epitope consisting of both ${}^7\text{EFE}_9$ and ${}^{313}\text{VDLSKVTSKC}_{322}$ (17).

In order to examine the potential interaction of the N-terminal ${}^7\text{EFE}_9$ segment with the AD fibril core, we first computationally docked the ${}^6\text{QEFEV}_{10}$ segment seen in the crystal structure into the low-resolution density shown to be adjacent to residues K317 and K321 in the AD Paired Helical Filament (PHF) (Figure 3 A-B)(12). In this model, the wet interface glutamates found in the crystal structure form electrostatic interactions with the exposed lysines in the PHF fibril, while the dry interface faces away from the PHF surface (Figure 3 A, B). Notably, we omitted Arg5 in this model due to steric clashes with Leu315 on the PHF. We speculate that Arg5 would have to adopt a different conformation in the fibril structure than in the crystal structure in order to maintain the interaction of Glu7 and Glu9 with Lys317 and Lys321.

To examine further the relevance of the ${}^7\text{EFE}_9$ and ${}^{313}\text{VDLSKVTSKC}_{322}$ interaction in tau fibrils, we searched the literature for other evidence that implicates the N-terminus of tau in fibril formation. Poorkaj, P. *et al*. described a missense mutation found in a Progressive Supranuclear Palsy (PSP) patient that changes R5 to a leucine(23). In addition, it has been shown that deletion of residues 2-18 produces less aggregated tau than the wild-type sequence whereas the inclusion of the R5L mutation increases the amount of aggregated tau in the presence of arachidonic acid(24). This is

consistent with our model of τ 7EFE₉ binding to τ 313VDLSKVTSKC₃₂₂ in the AD PHF; in that the deletion of residues 2-18 would abrogate the interaction of τ 7EFE₉ with τ 313VDLSKVTSKC₃₂₂. In addition, in our model the R5L mutation would result in a more stable interaction with Leu315 as discussed below.

To analyze if the R5L mutation might affect the binding of τ 5RQEFEV₁₀ to the τ 313VDLSKVTSKC₃₂₂ region in the AD filaments, we modeled the putative interaction of the sequence τ 5LQEFEV₁₀ with the AD PHF. To accomplish this, we mutated the R5 that was omitted in the wild-type model due a potential steric clash with L315 on the PHF, to a rotamer of leucine that would maximize its buried surface area and shape complementarity to L315 on the PHF (Figure 3 B). The model demonstrates that the mutation R5L would result in a more favorable interaction with the PHF than the native sequence, providing an explanation for R5L's ability to increase tau aggregation.

Our attempts to dock the τ 5RQEFEV₁₀ crystal structure into the τ 313VDLSKVTSKC₃₂₂ region on the cryoEM structure of the straight filaments (SFs) were hindered due to the tight packing protofilaments that occurs in this region. By truncating the residues present in the crystal structure to only τ 7EFE₉ it is possible to place these residues within hydrogen bonding distance of K317 on one protofilament and K321 on the other protofilament. This results in a binding site comprised of residues from two different tau monomers, as opposed to a binding site comprised of only one monomer as in the PHF (Figure 3D). However, this two-tau monomer model of τ 7EFE₉ bound to the SF would result in steric clashes if any other residues were added to the τ 7EFE₉ sequence (Figure 3D), particularly with L315, making it harder to assess whether there is enough space in the SF inter-protofilament interface for the N-terminal τ 7EFE₉

sequence. Likewise, it was difficult to examine the effect of the R5L mutation on this interaction due to the resulting steric clashes.

Discussion:

The initial proposal that ${}_{7}\text{EFE}_9$ interacts with ${}_{313}\text{VDLSKVTSKC}_{322}$ came from biochemical studies in which Jicha, *et al.* confirmed that two antibodies, MC-1 and Alz50, most likely bind a single epitope of tau comprised of discontinuous segments ${}_{7}\text{EFE}_9$ and ${}_{313}\text{VDLSKVTSKC}_{322}$. The idea of a single epitope comprising these two distal sequences was supported by antibody binding assays using a series of tau constructs containing truncations or mutations in these regions(19). Tau constructs missing either ${}_{7}\text{EFE}_9$ or ${}_{313}\text{VDLSKVTSKC}_{322}$ did not exhibit antibody binding, demonstrating that both sequences need to be present for antibody reactivity. In addition, a series of mutations to the ${}_{7}\text{EFE}_9$ segment (Glu7,9 -> Ala7,9; Phe8 -> Ser8) abrogated antibody binding to tau. Importantly, Jicha, *et al* showed that tau constructs missing ${}_{7}\text{EFE}_9$ or ${}_{313}\text{VDLSKVTSKC}_{322}$ could not be mixed in solution to recover the MC-1/Alz50 epitope, indicating that this epitope is formed intramolecularly.

In an attempt to examine which sequences might interact with the primary nucleating sequences of tau ${}_{275}\text{VQIINK}_{280}$ and ${}_{306}\text{VQIVYK}_{311}$, Moore *et al.* tested the ability of different tau sequences to accelerate and increase ${}_{275}\text{VQIINK}_{280}$ and ${}_{306}\text{VQIVYK}_{311}$ aggregation(19). Through these experiments, the authors predicted the heterozipper interaction formed between ${}_{306}\text{VQIVYK}_{311}$ and ${}_{375}\text{KLTFR}_{379}$. This predicted interaction was later confirmed by the AD tau filament structure(12). In addition, Moore, *et al* showed that ${}_{5}\text{RQEFEV}_{10}$ can form fibrils *in vitro*(19), although it did not affect the

aggregation of either ²⁷⁵VQIINK₂₈₀ or ³⁰⁶VQIVYK₃₁₁. This supports the idea that ⁵RQEFEV₁₀ instead interacts with ³¹³VDLSKVTSKC₃₂₂ in a different region of the fibril core. Further experiments similar to those performed by Moore, *et al*, including aggregation kinetics and circular dichroism of the individual peptides and a mixture of both peptides, could help strengthen evidence for the interaction of ⁵RQEFEV₁₀ and ³¹³VDLSKVTSKC₃₂₂ in the fibril state.

The results obtained by Jicha, *et al* and Moore, *et al* are consistent with the model proposed here where ⁵RQEFEV₁₀ occupies the un-modeled density that flanks residues K317 and K321 in the Fitzpatrick, *et al*. PHF cryo-em reconstruction(12). In particular, the abrogation of antibody binding by Glu7,9 -> Ala7,9 mutations performed by Jicha, *et al* can be explained by the disruption of the charge-charge interaction of glutamate and lysine residues in the proposed model (Figure 3 B)(17). The loss of this interaction would most likely greatly reduce the affinity of ⁷EFE₉ for the ³¹³VDLSKVTSKC₃₂₂ segment, leading to the loss of the MC-1 and Alz50 epitope. The loss of antibody binding from the Phe8 -> Ser8 can be explained in the proposed model given that Phe8 is facing away from the fibril, allowing it to remain exposed for antibody binding. Therefore, mutation of Phe8 may not prevent the far N-terminal segment from binding to the exposed lysines on the fibril core, but may still eliminate antibody reactivity. This suggests that the ⁷EFE₉ segment needs to be not only in a stacked conformation bound to K317 and K321 on the fibril core, but also needs F8 to be facing away from the fibril core and presented for antibody binding. A loss of either of these conditions would result in a loss of MC-1 reactivity.

The model of τ EF_{E9} interacting with K317 and K321 in the SF (Figure 3 C, D) suggests that either the τ EF_{E9} sequence binds in a different manner to the ₃₁₃VDLSKVTSKC₃₂₂ region on the SF or that the un-modeled density present in the Fitzpatrick, *et al* SF reconstruction does not result from the binding of the τ EF_{E9} motif, but perhaps some other anion. Tau AD filament structures from 3 additional cases seem to recapitulate the extra density seen at the SF inter-protofilament interface(13). This indicates that this density may be a common feature of the SF fold and necessary to interact with the four lysines resulting from K317 and K321 of each protofilament coming together at the SF inter-protofilament interface.

It is worth noting that the cryo-EM structures of AD tau fibrils display parallel, in-register beta-strands, whereas the RQEF_{E9}V crystal structure forms out-of-register, antiparallel beta-sheets. Because residues N-terminal to Val306 are not resolved in the cryo-EM structure, we cannot determine whether ₅RQEF_{E9}V₁₀ stacks into parallel or antiparallel sheets in the fibril. Our model used two strands of ₅RQEF_{E9}V₁₀ stacked in an anti-parallel beta-sheet as seen in the crystal structure. Although different from the crystal structure, parallel, in-register beta-sheets of ₅RQEF_{E9}V₁₀ would still form a wet and dry interface due to the alternating sequence of hydrophilic, charged residues and uncharged, mostly hydrophobic residues. Therefore, a parallel, in-register conformation of ₅RQEF_{E9}V₁₀ would still allow Glu7 and Glu9 to form electrostatic interactions with Lys317 and Lys321 in a manner similar to the model proposed in Figure 3 A-B.

Recently, a new polymorph of tau from the brain of a Pick's Disease case has been visualized by cryo-EM(14). This structure adopts a drastically different fold from the AD filaments; however, the Pick's Disease filaments are still MC-1 reactive,

indicating the preservation of the ${}_{7}\text{EFE}_9$ and ${}_{313}\text{VDLSKVTSKC}_{322}$ epitope(14). In this structure, K317 and K321 are exposed to the solvent in a beta-sheet conformation, which would allow the N-terminal ${}_{7}\text{EFE}_9$ segment to bind K317 and K321 through electrostatic interactions between the glutamates and lysines similar to the AD PHF model (Figure 3 A, B). This electrostatic interaction would preserve the MC-1 epitope and provide an explanation for why MC-1 recognizes both tau fibril polymorphs.

In addition, the potential strengthening of the N-terminal interaction with the fibril core through the R5L mutation and its discovery in a PSP patient, suggests that this interaction may also occur in the PSP tau fibril. Although there is evidence that so-called 4R tauopathies, where the dominant species found in aggregated tau are the 4R isoforms, PSP and Corticobasal Degeneration (CBD) form different tau polymorphs, their structures have not yet been determined(25). However, as long as the ${}_{313}\text{VDLSKVTSKC}_{322}$ region adopts a beta-sheet like fold, and K317 and K321 remain solvent-exposed, the long-range charge-charge interaction with ${}_{7}\text{EFE}_9$ could be preserved. In short, there may be a common interaction among the disparate folds of tau polymorphs.

In the past, our lab has developed inhibitors of tau aggregation by structure-based drug design(11,16). This requires detailed structural knowledge of a site of the tau protein in the aggregated state obtained by X-ray crystallography or MicroED. These inhibitors target segments of the tau protein in the microtubule binding region that is thought to participate in the fibril core of all tau filaments. However, given the structural evidence thus far that the microtubule binding region can adopt different folds in different diseases, it is likely that a spectrum of inhibitors will be necessary to most

effectively block aggregation or spreading of specific tau polymorphs. Immuno-labeling with MC-1 seems to indicate that the N-terminal interaction with the fibril core modeled here is preserved in both AD and Pick's Disease tau filaments. Therefore, an inhibitor targeted towards this interaction may be general to all tau filaments, providing another target for treating tauopathies.

Figures:

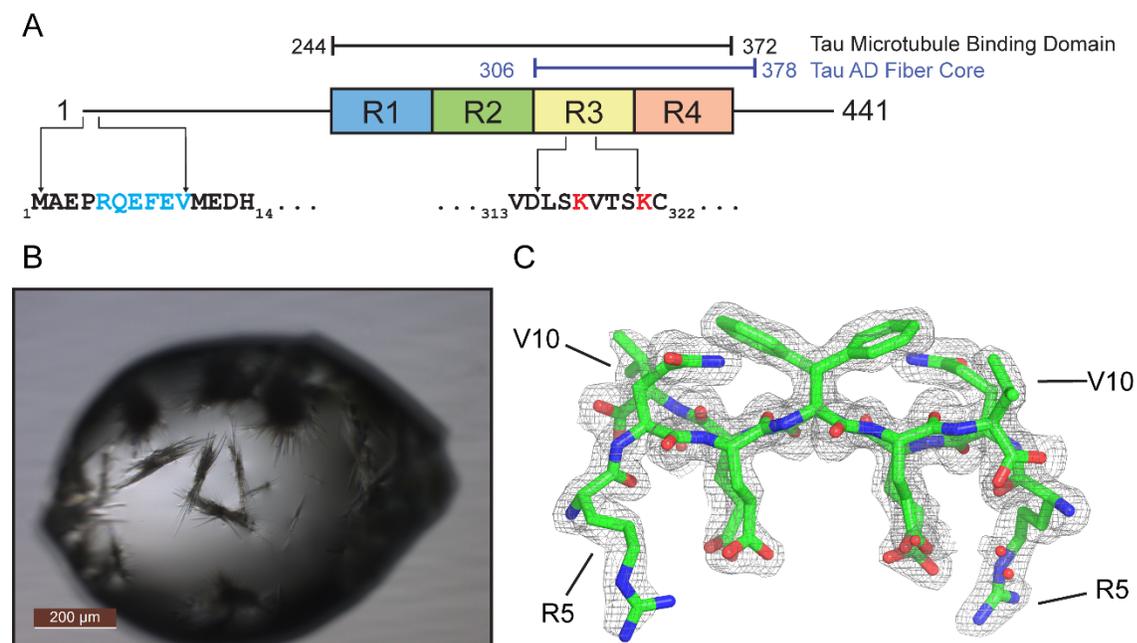


Figure 1: Crystal structure of tau N-terminal segment ${}^5\text{RQEFEV}_{10}$

A) Schematic of tau primary structure. B) Crystals of ${}^5\text{RQEFEV}_{10}$ grown using the hanging drop method. C) Atomic model and electron density of ${}^5\text{RQEFEV}_{10}$ demonstrating the quality of fit. The view is down the fibril axis, showing two anti-parallel strands.

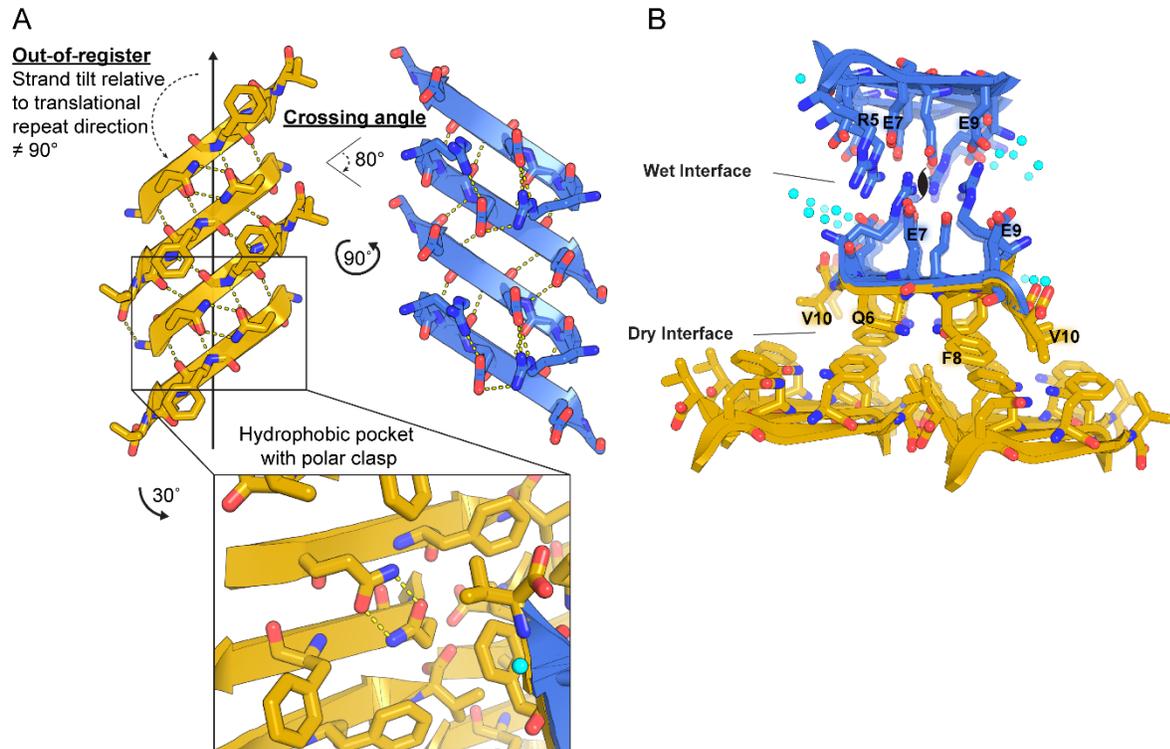


Figure 2: Crystal structure of $_5$ RQEFEV $_{10}$ reveals a wet and a dry interface

A) $_5$ RQEFEV $_{10}$ forms amyloid-like out-of-register protofilaments with wet and dry interfaces. Inset shows formation of a polar clasp with neighboring glutamines in the hydrophobic pocket of the dry interface. B) View down the fibril axis of $_5$ RQEFEV $_{10}$ highlighting the interactions between residues within the wet and dry interfaces. Water molecules are shown by aqua spheres.

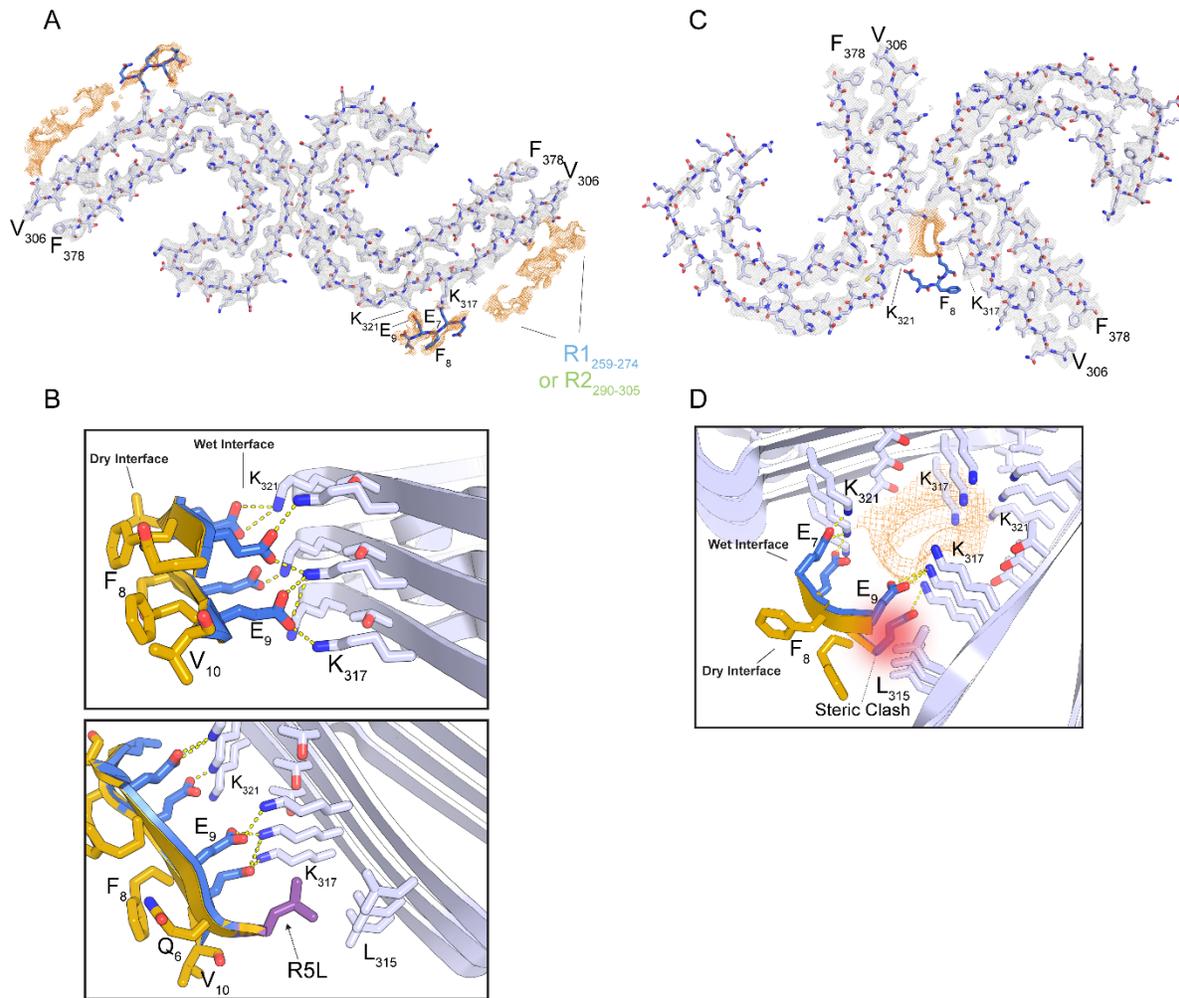


Figure 3: Speculative model for ${}^5\text{RQEFEV}_{10}$ interaction with Alzheimer's Disease paired helical and straight filaments fibril cores

A) Atomic model of Alzheimer's disease paired helical filaments (PHF) (5o3l.pdb) shown with electron density of modeled (grey) and un-modeled (orange) regions(12).

${}^5\text{RQEFEV}_{10}$ is docked into un-modeled density flanking the solvent-exposed K317 and K321 residues of the PHF. B) Detail (top) highlighting the interaction of the glutamates in the wet interface with K317 and K321 of the PHF. Detail (bottom) demonstrating the possible interaction of the R5L mutation with L315 of the PHF. C) Overview of potential

interaction of γ EFE₉ with straight filaments (SF) (5o3t.pdb) at the inter-protofilament interface. D) Detail of the potential hydrogen bonding of wet interface glutamates with K317 and K321 and potential steric clash with L315 of the SF.

Table 1. Data collection and refinement statistics.

	RQEFEV (PDB ID: 6N4P)
Wavelength	0.9792
Resolution range	16.08 - 1.851 (1.918 - 1.851)
Space group	P2 ₁
Unit cell	16.59 11.45 25.42 90 104.236 90
Total reflections	2416 (226)
Unique reflections	842 (80)
Multiplicity	2.9 (2.8)
Completeness (%)	97.2 (96.4)
Mean I/sigma(I)	4.2 (1.6)
Wilson B-factor	14.8
R-merge	0.16 (0.60)
R-meas	0.20 (0.72)
R-pim	0.11 (0.40)
CC1/2	0.97 (0.83)
CC*	0.99 (0.95)
Reflections used in refinement	838 (80)
Reflections used for R-free	85 (8)
R-work	0.19 (0.30)
R-free	0.27 (0.47)
CC(work)	0.96 (0.81)
CC(free)	0.91 (0.83)
Number of non-hydrogen atoms	118
macromolecules	114
solvent	4
Protein residues	12
RMS(bonds)	0.013
RMS(angles)	1.48
Ramachandran favored (%)	100.00
Ramachandran allowed (%)	0.00
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	4.55
Average B-factor	21.9
macromolecules	21.5
solvent	32.5

Statistics for the highest-resolution shell are shown in parentheses

Methods:

Crystallization and Data Collection: Synthetic peptide RQEFEV was ordered from GenScript. RQEFEV was crystallized using the hanging drop method with a 2:1 mixture

of 60 mg/mL RQEFEV and 0.2 M Ammonium Citrate Dibasic, 30% PEG 3350.

Diffraction data was collected at APS Beamline 24-ID-E using an Eiger detector.

Data Processing and Structure Determination: Diffraction data were indexed and integrated using XDS and scaled using XSCALE(26). Molecular replacement was performed using Phaser and an idealized beta-strand as a molecular replacement probe(27). Model-building and manual real-space refinement was performed in COOT(28). Automated reciprocal-space and real-space refinement was performed using Refmac and Phenix(29,30).

Modeling: Modeling was performed in COOT using the RQEFEV crystal structure and the cryo-em structures for the AD PHF (5o3l.pdb) and SF (5o3t.pdb) downloaded from the PDB. Cryo-em maps for the PHF (EMD-3741) and SF (EMD-3743) were also used for modeling and generating figures. All figures were made in Pymol (Schrodinger).

Acknowledgments:

We thank Michael Sawaya for discussion and Michael Collazo at UCLA-DOE Macromolecular Crystallization Core Technology Center for crystallization support. We thank NIH AG 054022. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which is funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). The Eiger 16M detector on 24-ID-E beam line is funded by a NIHORIP HEI grant (S10OD021527). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the

DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. DRB is funded by the National Science Foundation Graduate Research Fellowship Program.

References:

1. Kadavath H, Hofele RV, Biernat J, Kumar S, Tepper K, Urlaub H, et al. Tau stabilizes microtubules by binding at the interface between tubulin heterodimers. *Proc Natl Acad Sci U S A*. 2015 Jun 16;112(24):7501–6.
2. Kellogg EH, Hejab NMA, Poepsel S, Downing KH, DiMaio F, Nogales E. Near-atomic model of microtubule-tau interactions. *Science*. 2018 May 10;eaat1780.
3. Miyata Y, Koren J, Kiray J, Dickey CA, Gestwicki JE. Molecular chaperones and regulation of tau quality control: strategies for drug discovery in tauopathies. *Future Med Chem*. 2011 Sep;3(12):1523–37.
4. Goedert M, Wischik CM, Crowther RA, Walker JE, Klug A. Cloning and sequencing of the cDNA encoding a core protein of the paired helical filament of Alzheimer disease: identification as the microtubule-associated protein tau. *Proc Natl Acad Sci*. 1988 Jun 1;85(11):4051–5.
5. Wischik CM, Novak M, Thøgersen HC, Edwards PC, Runswick MJ, Jakes R, et al. Isolation of a fragment of tau derived from the core of the paired helical filament of Alzheimer disease. *Proc Natl Acad Sci U S A*. 1988 Jun;85(12):4506–10.
6. Wischik CM, Novak M, Edwards PC, Klug A, Tichelaar W, Crowther RA. Structural characterization of the core of the paired helical filament of Alzheimer disease. *Proc Natl Acad Sci U S A*. 1988 Jul;85(13):4884–8.

7. Narayanan RL, Dürr UHN, Bibow S, Biernat J, Mandelkow E, Zweckstetter M. Automatic Assignment of the Intrinsically Disordered Protein Tau with 441-Residues. *J Am Chem Soc.* 2010 Sep 1;132(34):11906–7.
8. Jeganathan S, von Bergen M, Brütlich H, Steinhoff H-J, Mandelkow E. Global Hairpin Folding of Tau in Solution. *Biochemistry.* 2006 Feb 1;45(7):2283–93.
9. Mirbaha H, Chen D, Morazova OA, Ruff KM, Sharma AM, Liu X, et al. Inert and seed-competent tau monomers suggest structural origins of aggregation. Akhmanova A, editor. *eLife.* 2018 Jul 10;7:e36584.
10. Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, et al. Atomic structures of amyloid cross- β spines reveal varied steric zippers. *Nature.* 2007 May;447(7143):453–7.
11. Seidler PM, Boyer DR, Rodriguez JA, Sawaya MR, Cascio D, Murray K, et al. Structure-based inhibitors of tau aggregation. *Nat Chem.* 2018 Feb;10(2):170–6.
12. Fitzpatrick AWP, Falcon B, He S, Murzin AG, Murshudov G, Garringer HJ, et al. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature.* 2017 Jul 13;547(7662):185–90.
13. Falcon B, Zhang W, Schweighauser M, Murzin AG, Vidal R, Garringer HJ, et al. Tau filaments from multiple cases of sporadic and inherited Alzheimer's disease adopt a common fold. *Acta Neuropathol (Berl)* [Internet]. 2018 Oct 1 [cited 2018 Oct 25]; Available from: <https://doi.org/10.1007/s00401-018-1914-z>
14. Falcon B, Zhang W, Murzin AG, Murshudov G, Garringer HJ, Vidal R, et al. Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature.* 2018 Sep;561(7721):137.

15. Bergen M von, Friedhoff P, Biernat J, Heberle J, Mandelkow E-M, Mandelkow E. Assembly of τ protein into Alzheimer paired helical filaments depends on a local sequence motif (306VQIVYK311) forming β structure. *Proc Natl Acad Sci*. 2000 May 9;97(10):5129–34.
16. Sievers SA, Karanicolas J, Chang HW, Zhao A, Jiang L, Zirafi O, et al. Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature*. 2011 Jun 15;475(7354):96–100.
17. Jicha GA, Bowser R, Kazam IG, Davies P. Alz-50 and MC-1, a new monoclonal antibody raised to paired helical filaments, recognize conformational epitopes on recombinant tau. *J Neurosci Res*. 1997 Apr 15;48(2):128–32.
18. Thompson MJ, Sievers SA, Karanicolas J, Ivanova MI, Baker D, Eisenberg D. The 3D profile method for identifying fibril-forming segments of proteins. *Proc Natl Acad Sci*. 2006 Mar 14;103(11):4074–8.
19. Moore CL, Huang MH, Robbenolt SA, Voss KR, Combs B, Gamblin TC, et al. Secondary nucleating sequences affect kinetics and thermodynamics of tau aggregation. *Biochemistry*. 2011 Dec 20;50(50):10876–86.
20. Eisenberg DS, Sawaya MR. Structural Studies of Amyloid Proteins at the Molecular Level. *Annu Rev Biochem*. 2017 20;86:69–95.
21. Liu C, Zhao M, Jiang L, Cheng P-N, Park J, Sawaya MR, et al. Out-of-register β -sheets suggest a pathway to toxic amyloid aggregates. *Proc Natl Acad Sci*. 2012 Dec 18;109(51):20913–8.

22. Poorkaj P, Muma NA, Zhukareva V, Cochran EJ, Shannon KM, Hurtig H, et al. An R5L τ mutation in a subject with a progressive supranuclear palsy phenotype. *Ann Neurol*. 2002 Oct 1;52(4):511–6.
23. Gamblin TC, Berry RW, Binder LI. Tau polymerization: role of the amino terminus. *Biochemistry*. 2003 Feb 25;42(7):2252–7.
24. Gibbons GS, Banks RA, Kim B, Changolkar L, Riddle DM, Leight SN, et al. Detection of Alzheimer Disease (AD)-Specific Tau Pathology in AD and NonAD Tauopathies by Immunohistochemistry With Novel Conformation-Selective Tau Antibodies. *J Neuropathol Exp Neurol*. 2018 Mar 1;77(3):216–28.
25. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr*. 2010 Feb 1;66(Pt 2):125–32.
26. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007 Aug 1;40(Pt 4):658–74.
27. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010 Apr;66(Pt 4):486–501.
28. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*. 1997 May 1;53(Pt 3):240–55.
29. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*. 2012 Apr 1;68(4):352–67.

Structures of fibrils formed by α -synuclein hereditary disease mutant H50Q reveal new polymorphs

David R. Boyer^{1,3}, Binsen Li^{2,3}, Chuanqi Sun², Weijia Fan², Michael R. Sawaya¹, Lin Jiang^{2*} and David S. Eisenberg^{1*}

Deposits of amyloid fibrils of α -synuclein are the histological hallmarks of Parkinson's disease, dementia with Lewy bodies and multiple system atrophy, with hereditary mutations in α -synuclein linked to the first two of these conditions. Seeing the changes to the structures of amyloid fibrils bearing these mutations may help to understand these diseases. To this end, we determined the cryo-EM structures of α -synuclein fibrils containing the H50Q hereditary mutation. We find that the H50Q mutation results in two previously unobserved polymorphs of α -synuclein: narrow and wide fibrils, formed from either one or two protofilaments, respectively. These structures recapitulate conserved features of the wild-type fold but reveal new structural elements, including a previously unobserved hydrogen-bond network and surprising new protofilament arrangements. The structures of the H50Q polymorphs help to rationalize the faster aggregation kinetics, higher seeding capacity in biosensor cells and greater cytotoxicity that we observe for H50Q compared to wild-type α -synuclein.

Several lines of evidence suggest that aggregation of α -synuclein (α -syn) into amyloid fibrils underlies the group of diseases termed synucleinopathies—Parkinson's disease (PD), Lewy body dementia and multiple systems atrophy. (1) α -Syn fibrils are found in the hallmark lesions of PD and Lewy body dementia (Lewy bodies), as well as in the hallmark glial and neuronal lesions in multiple systems atrophy^{1,2}. (2) Hereditary mutations in α -syn have been linked to PD and Lewy body dementia³. (3) Dominantly inherited duplications and triplications of the chromosomal region that contains wild-type SNCA—the gene that encodes α -syn—are sufficient to cause PD (refs. 4–6). (4) Recombinantly assembled α -syn fibrils show cross- β structure and their injection into the brains of wild-type mice induced PD-like Lewy body and Lewy neurite formation, as well as cell-to-cell spreading and motor deficits reminiscent of PD^{7,8}.

Advances in solid-state NMR and cryo-EM have greatly increased our knowledge of the structure of full-length amyloid proteins, allowing us to examine interactions beyond the local views provided by crystallographic methods^{9–13}. Therefore, we previously used cryo-EM to determine the structures of wild-type full-length α -syn fibrils (Fig. 1a,b, left). These structures reveal two distinct polymorphs—termed the rod and twister¹⁴. Both fibrils are wound from two identical protofilaments related by an approximate 2, fibril axis. The protofilaments that form the rod and twister polymorphs are distinct: the rod protofilaments contain ordered residues 38–97 whereas the twister protofilaments contain ordered residues 43–83. Both polymorphs share a similar structurally conserved β -arch formed by residues 50–77. However, the protofilament interfaces between the two polymorphs differ: in the rod polymorph, residues 50–57 from the preNAC region form the interface of the two protofilaments, whereas, in the twister polymorph, residues 66–78 from the NACore form the interface.

Hereditary mutations offer understanding of the link between protein structure and disease. H50Q is one such mutation that was discovered independently in two individuals with PD, with one patient having a known familial history of parkinsonism and dementia^{15,16}. The H50Q mutation enhances α -syn aggregation in vitro by reducing the solubility of the monomer, decreasing the lag time of fibril formation and increasing the amount of fibrils formed^{17,18}. Additionally, H50Q has been shown to be secreted at higher levels from SH-SY5Y cells and to be more cytotoxic to primary hippocampal neurons than wild-type α -syn¹⁹. Taken together these data suggest that patients harboring the H50Q mutation may develop fibrils more easily and that these fibrils may have different underlying structures than wild-type fibrils.

The structure of the wild-type rod polymorph suggests that the H50Q mutation may alter key contacts at the protofilament interface¹⁴. In this structure, two pairs of H50–E57 residues interact on opposing protofilaments, stabilizing the protofilament interface through charge–charge interactions. The mutation to the uncharged, polar glutamine may therefore disrupt the rod polymorph protofilament interface, leading to different polymorphs of α -syn, potentially explaining the different observed properties of H50Q versus wild-type α -syn^{17,18}. To examine the exact effect of the H50Q mutation on the structure of α -syn fibrils, we sought to determine the atomic structures of H50Q α -syn fibrils using cryo-EM and to compare aggregation kinetics, stability, seeding capacity and cytotoxicity to wild-type α -syn.

Results

Cryo-EM structure and architecture of H50Q α -syn fibrils. To pursue cryo-EM structure determination, we expressed and purified recombinant full-length α -syn containing the H50Q hereditary mutation and subsequently grew fibrils under the same conditions

¹Department of Chemistry and Biochemistry and Biological Chemistry, UCLA-DOE Institute, Molecular Biology Institute and Howard Hughes Medical Institute, UCLA, Los Angeles, CA, USA. ²Department of Neurology, David Geffen School of Medicine, Molecular Biology Institute, UCLA, Los Angeles, CA, USA. ³These authors contributed equally: David R. Boyer, Binsen Li. *e-mail: jianglin@ucla.edu; david@mbi.ucla.edu

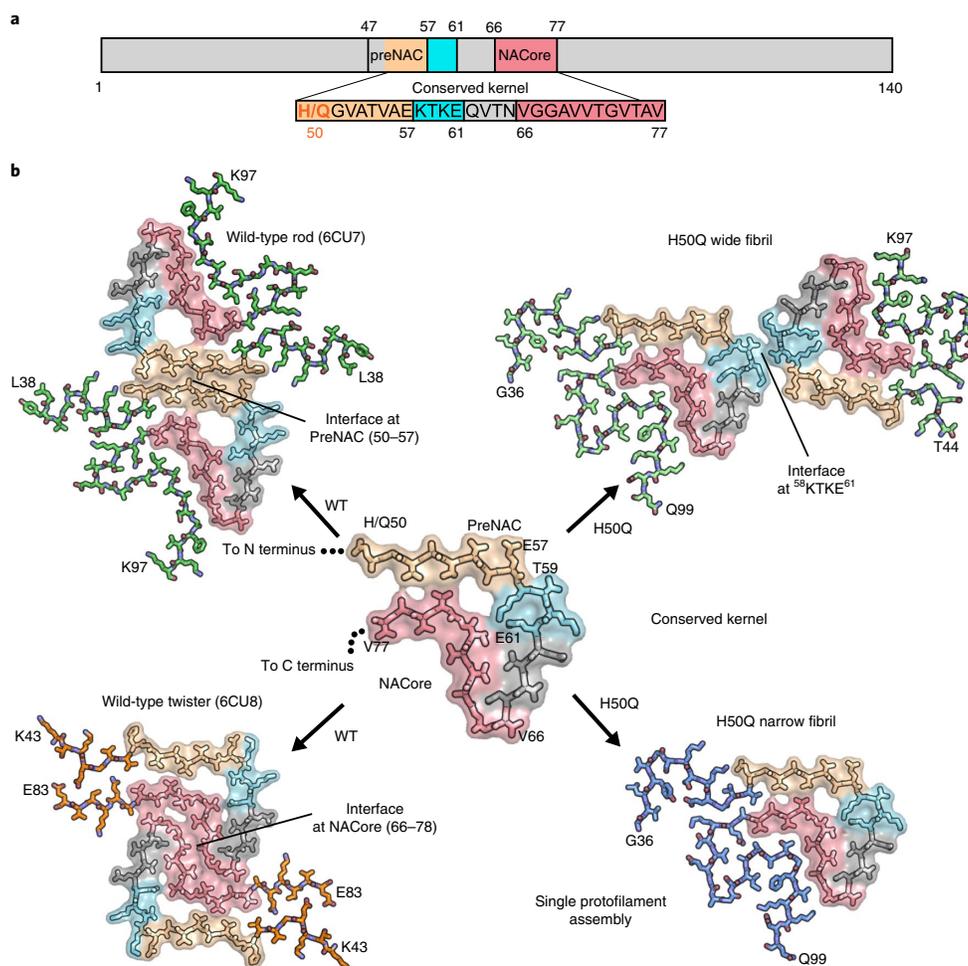


Fig. 1 | Comparison of wild-type and H50Q polymorphs. **a**, Primary structure schematic highlighting residues of the conserved kernel (50–77) that form protofilament interfaces in α -syn polymorphs. **b**, A conserved kernel is used to assemble wild-type (left) and H50Q (right) α -syn polymorphs in at least four different ways.

as those used for cryo-EM studies of wild-type α -syn (see Methods). After obtaining optimal cryo grid conditions and subsequent high-resolution cryo-EM data collection and processing, we determined the near-atomic structures of two polymorphs, which we term narrow and wide fibrils, to resolutions of 3.3 and 3.6 Å, respectively (Figs. 1b and 2a–d, Extended Data Figs. 1a,b and 2a and Table 1).

Both the narrow and wide fibrils have pitches of ~ 900 Å, as measured from the cross-over distances observed in electron micrographs as well as 2D classifications of box sizes containing helical segments that encompass an entire helical pitch (Extended Data Fig. 2b,c; see Methods). Narrow fibrils have a width that varies from 55 to 66 Å; wide fibrils have a width that varies from 67 to 116 Å (Fig. 2b and Extended Data Fig. 1a). The narrow fibril is wound from a single protofilament, which we designate protofilament A, while the wide fibril is wound from two slightly different protofilaments, both protofilament A and a second protofilament, which we term protofilament B (Fig. 2c,d). Both narrow and wide fibril reconstructions show identical densities flanking protofilament A, which we term islands 1 and 2 (Supplementary Note 1 and Extended Data Fig. 3). Narrow fibrils are roughly five times more abundant than

wide fibrils. Protofilament A contains ordered residues from G36 to Q99, while protofilament B contains ordered residues T44 to K97 (Fig. 2a,c,d).

The protein chains in the ordered cores of both protofilaments fold essentially within a two-dimensional layer (Fig. 2b–d), with stretches of straight β -strand regions interrupted by sharp turns (Fig. 3a,b). The fibrils are formed by the chains within the two-dimensional layers stacking on one another along the fibril axis every 4.8 Å, forming β -sheets that extend for hundreds of nanometers (Fig. 2b).

Differences between protofilaments A and B. To define the differences in protofilaments A and B observed in the two polymorphs, we aligned protofilament A with protofilament B (Fig. 3b). The alignment reveals that residues 47–97 adopt nearly identical conformations in both protofilaments A and B; however, protofilament A has an ordered β -arch formed by residues 36–46, whereas protofilament B becomes disordered after T44 (Fig. 3b,c). The β -arch in protofilament A features an extensive hydrogen-bond network among Y39, T44 and E46 that is not observed in any α -syn structures

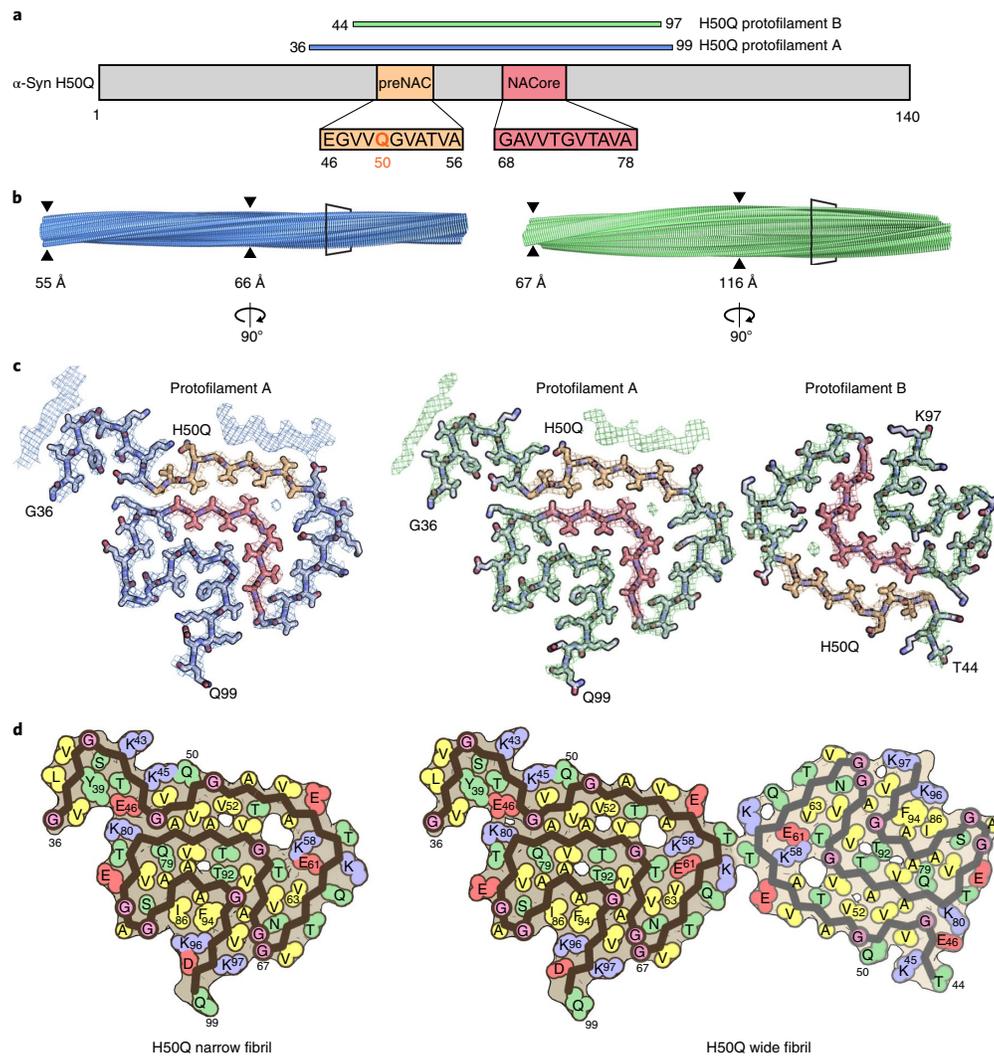


Fig. 2 | Cryo-EM structures of H50Q polymorphs. **a**, Schematic of α -syn primary structure demonstrating location of the ordered core of H50Q protofilaments A and B, preNAC and NACore and H50Q hereditary mutation. **b**, View perpendicular to the fibril axes of narrow and wide fibril cryo-EM reconstructions with the minimum and maximum widths of the fibrils labeled. **c**, View parallel to the fibril axes of sections of narrow and wide fibrils revealing one layer of each fibril. Narrow fibrils are composed of one protofilament, designated protofilament A. Wide fibrils are composed of two protofilaments: protofilament A and a less well-ordered chain designated protofilament B. Protofilament A is nearly identical in both fibril species, while protofilament B differs from protofilament A and is only found in the wide fibril species. **d**, Schematic representation of fibril structures with amino acid side chains colored as follows: hydrophobic (yellow), negatively charged (red), positively charged (blue), polar, uncharged (green) and glycine (pink).

determined so far (Fig. 3b,c). In addition, we note that in protofilament A, Q50 is hydrogen bonded to K45, whereas in protofilament B, K45 and Q50 are not hydrogen bonded and appear to be solvent facing (Fig. 3b,c). This alternative arrangement of K45 and Q50 may explain why protofilament A forms an ordered β -arch while protofilament B does not. This suggests that the conformation of K45 and Q50 can act as a switch whereby, as they become hydrogen bonded, the amino (N)-terminal residues 36–44 become ordered, forming the Y39–T44–E46 hydrogen bond triad. We also note that protofilament B maintains the E46–K80 hydrogen bond observed in wild-type structures, while participation of E46 in a hydrogen bond with T44 in protofilament A differs from that in

wild-type structures. Instead, K80 now hydrogen bonds with T81 in protofilament A (Fig. 3b,c). The differences observed here highlight the impact that hydrogen-bonding arrangements can have on fibril structure and explain the atomic basis for the asymmetry of the two protofilaments in the wide fibril polymorph. Indeed, previous studies on wild-type α -syn fibril structures have revealed fibrils composed of asymmetric protofilaments whose asymmetry could be explained by the types of alternative hydrogen-bond patterns that we observe here²⁰.

Cavities in α -syn fibril structures. We note that all α -syn fibril structures determined so far display a similar cavity at the center

Table 1 | Cryo-EM data collection, refinement and validation statistics

	H50Q narrow fibril (EMD-20328, PDB 6PEO)	H50Q wide fibril (EMD-20331, PDB 6PES)
Data collection and processing		
Magnification	130,000	130,000
Voltage (kV)	300	300
Camera	K2 Summit (Quantum LS)	K2 Summit (Quantum LS)
Frame exposure time (s)	0.2	0.2
Movie frames (no.)	30	30
Electron exposure (e ⁻ /Å ²)	36	36
Box size (pixel)	288	224
Inter-box distance (Å)	28.8	22.4
Micrographs collected (no.)	3,577	3,577
Segments extracted (no.)	1,183,284	137,395
Segments after Class2D (no.)	510,477	NA ^a
Segments after Class3D (no.)	30,133	28,016
Map resolution (Å)	3.3	3.6
FSC threshold	0.143	0.143
Refinement		
Initial model used	De novo	De novo
Model resolution (Å)	3.2	3.5
FSC threshold	0.5	0.5
Model resolution range (Å)	200–3.3	200–3.6
Map sharpening B factor (Å ²)	–148	–207
Model composition		
Nonhydrogen atoms	2,205	4,040
Protein residues	320	590
Ligands	NA	NA
B factors (Å ²)		
Protein	NA	NA
Ligand	NA	NA
R.m.s. deviations		
Bond lengths (Å)	0.003	0.009
Bond angles (°)	0.5	0.7
Validation		
MolProbity score	2.09	2.18
Clashscore	10.62	12.28
Poor rotamers (%)	0	0
Ramachandran plot		
Favored (%)	90.3	87.7
Allowed (%)	9.7	12.3
Disallowed (%)	0	0

^aNumber of segments after Class2D not available (NA) for the wide fibril as only 3D classification was performed.

of the β -arch, surrounded by residues T54, A56, K58, G73 and V74 (Figs. 1b and 2c and Extended Data Fig. 4a,b). However, in contrast to our previously published wild-type structures, K58 now faces inward towards the cavity, instead of outward towards the solvent, while T59 now flips away from the cavity (Extended

Data Fig. 4a). This inversion is possible because lysine and threonine have both hydrophobic and hydrophilic character, allowing them to be favorably positioned either facing the solvent or the cavity. Indeed, energetic calculations demonstrate that K58 and T59 can have a positive stabilization energy in both cases (Extended Data Fig. 5b,c).

Presumably the β -arch cavity is filled with disordered solvent that is not defined by cryo-EM or previously used solid-state NMR averaging methods. However, we note here that in our narrow and wide fibril polymorphs we visualize additional density in the cavity that is not accounted for by protein side chains (Fig. 2c and Extended Data Fig. 4b). We wonder whether this density arises from noise, or a back-projection artifact, but it is observed in the same location independently in three protofilaments (protofilament A from the narrow fibril and protofilaments A and B from the wide fibril) leading us to believe that it may come from a solvent molecule (Extended Data Fig. 4b). Since our fibril growth conditions contain only water and tetrabutylphosphonium bromide, whose long aliphatic groups make it too large to fit into the tight cavity, we examined if the density could come from a water molecule and if any of the surrounding residues could serve as hydrogen-bond partners. We observe that the γ -hydroxyl of T54, the ϵ -amino of K58 and the carbonyl oxygen of G73 are the only potential hydrogen-bonding partners for the putative water molecule (Extended Data Fig. 4b). In addition, there are several methyl groups from V74 and A56 that are proximal to the putative solvent molecule (Extended Data Fig. 4b). The distance between the density and potential hydrogen partners ranges from 4.2 to 5.4 Å, which is longer than usual for hydrogen bonds (Extended Data Fig. 4b). However, given that there are three hydrogen-bond partners, perhaps the density visualized in the reconstruction is an average of positions occupied by water molecules. Nonetheless, given the resolution of our maps we cannot unambiguously identify the molecule(s) occupying the density in the center of the cavity.

Wide fibril protofilament interface. The wide fibril contains a novel protofilament interface formed by residues ⁵⁸KTKKE⁶¹ not previously seen in wild-type α -syn structures¹⁴. Residues ⁵⁸KTKKE⁶¹ are located at a sharp turn in both protofilaments A and B, and, consequently, the interface between protofilaments is remarkably small (Extended Data Fig. 6a). Consistent with the minimal size of the interface, the shape complementarity of 0.57 and buried surface area of 50 Å² of the wide fibril interface are low compared to the more extensive preNAC and NACore interfaces seen in our previous wild-type structures (Extended Data Fig. 6b)¹⁴. Indeed, the C γ atom of T59 from each protofilament is the only atom to interact across the protofilament interface (Extended Data Fig. 6b), making this interface the smallest fibril protofilament interface observed to date.

We wondered why the H50Q polymorphs do not utilize the extensive preNAC or NACore protofilament interfaces found in the wild-type rod and twister polymorphs, respectively. Our structures reveal that, unlike the wild-type twister polymorph, the NACore is buried within the fibril core and is inaccessible as a protofilament interface (Fig. 2c). However, the preNAC region is more accessible as a potential protofilament interface in both of the H50Q polymorphs (Fig. 2c). To examine why the preNAC region does not form a protofilament interface, we first compared the preNAC protofilament interface in the wild-type rod polymorph with the same region in the H50Q narrow fibril (Extended Data Fig. 7a,b). We noticed that in the wild-type rod polymorph H50 on one protofilament interacts with E57 on the opposite protofilament, possibly through a charge–charge interaction (Extended Data Fig. 7b). The mutation to a polar, uncharged glutamine leads to a loss of interaction with E57 and instead Q50 forms an intramolecular hydrogen bond with K45 in H50Q protofilament A, thereby producing a single protofilament

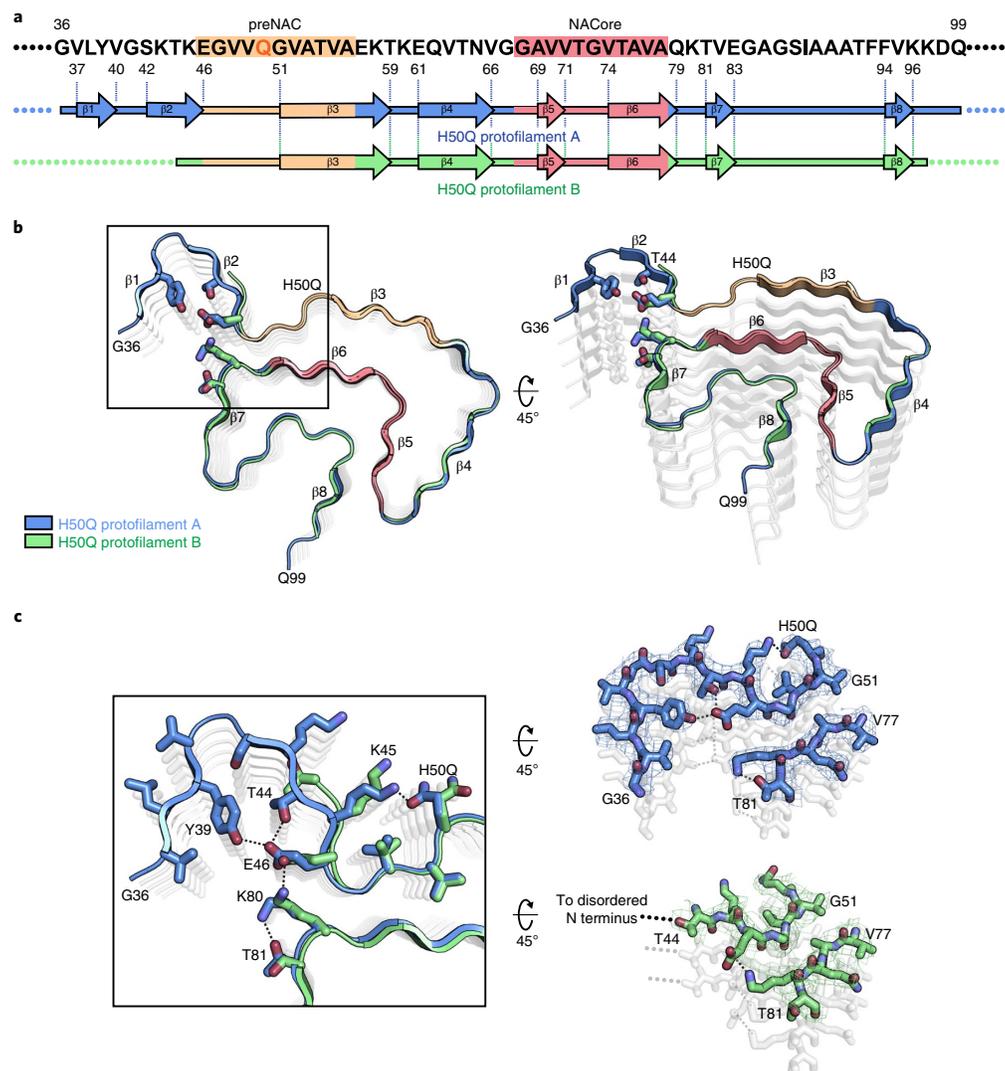


Fig. 3 | Comparison of protofilaments A and B. **a**, Schematic of primary and secondary structure of the fibril core of H50Q protofilaments A and B. Arrows indicate regions of protofilaments A and B that adopt β -strand conformations. **b**, Structural alignment of H50Q protofilaments A and B. **c**, Alignment of protofilaments A and B reveals that protofilament A has a β -arch formed by residues 36–46, whereas protofilament B is disordered before residue 44. Additionally, in protofilament A, H50Q hydrogen bonds with K45, while in protofilament B, H50Q and K45 adopt different conformations and do not hydrogen bond. In protofilament A, E46 participates in the Y39–T44–E46 hydrogen-bond triad and K80 hydrogen bonds with T81, while in protofilament B E46 hydrogen bonds with K80 as in the wild-type rod polymorphs.

assembly (Extended Data Fig. 7a). Further, speculative models where additional protofilaments are docked at the preNAC region in either protofilament A or B show severe steric clashes occurring, further supporting the idea that the H50Q mutation disallows assembly of protofilaments at the preNAC (Supplementary Note 2 and Extended Data Fig. 7c,d).

Energetic and biochemical analysis of H50Q fibrils. We next wondered if the structural differences that we observe in the H50Q versus wild-type fibrils affect their stabilities. To examine this, we first calculated modified atomic solvation energies for the H50Q mutant structures and wild-type α -syn structures (see Supplementary Note 3

for differences in stabilization energies for some residues in protofilaments A and B in the wide fibril). We find that both wild-type and H50Q fibrils are stabilized by energies comparable to a selection of known irreversible fibrils, including tau-paired helical filament structures from Alzheimer's disease²¹, serum amyloid A fibrils from systemic amyloidosis²² and TDP-43 SegA-sym fibrils²³, and significantly larger than the stabilization energies of the reversible FUS fibrils²⁴ (Fig. 4a and Table 2). To confirm this, we performed stability assays to measure the resistance of fibrils to heat and SDS, and we found that both wild type and H50Q show a similar resistance to denaturation, consistent with our energetic calculations (Fig. 4c). This is also consistent with the idea that both wild-type and H50Q

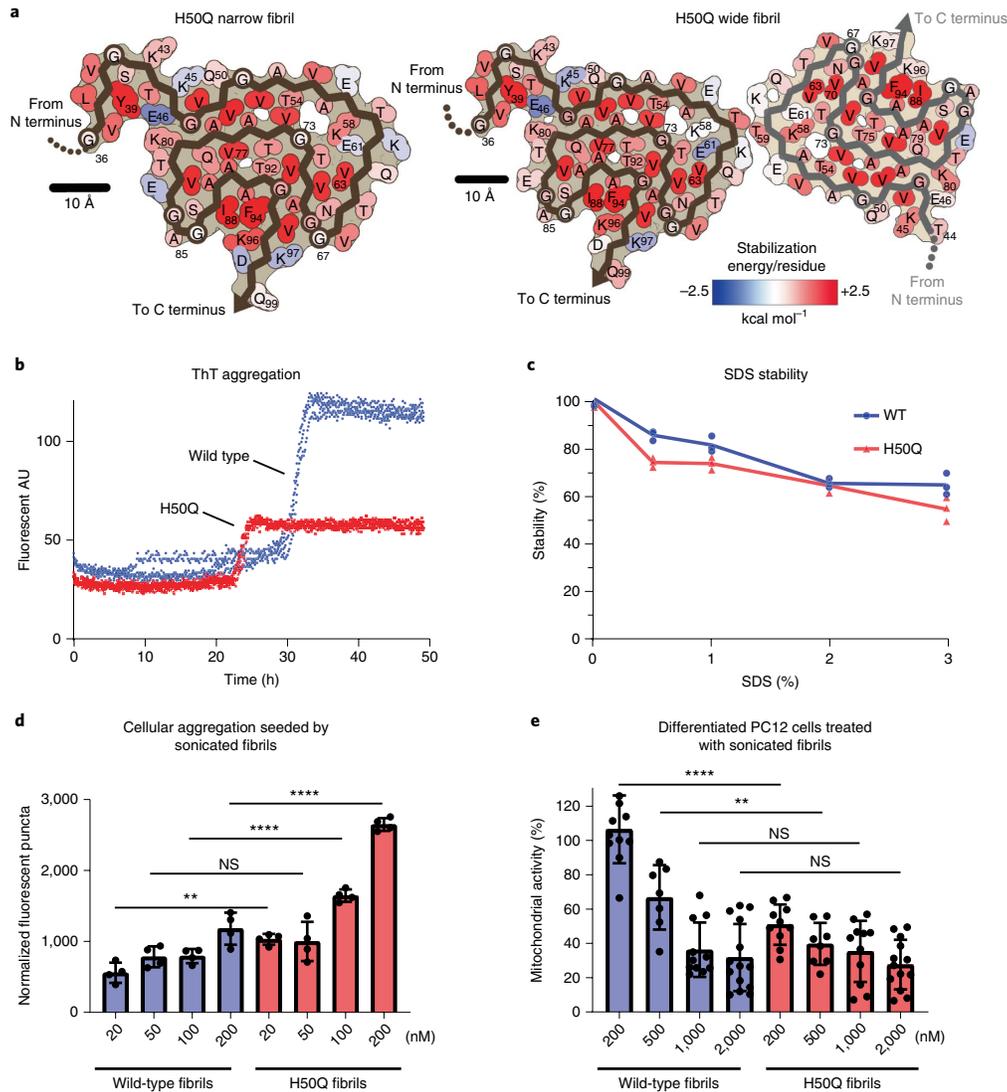


Fig. 4 | Solvation energy maps and biochemical characterization of H50Q and wild-type α -syn. **a**, Solvation energy maps of H50Q narrow and wide fibrils. The stabilizing residues are red; the destabilizing residues are blue. **b**, ThT assay measuring kinetics of H50Q and wild-type α -syn aggregation. Data from three independent replicates for each sample are shown for each time point. Wild-type aggregation plateaus at 30 h whereas H50Q aggregates faster and plateaus at 23 h. **c**, Stability assay of H50Q and wild-type α -syn. Fibrils were brought to the indicated concentration of SDS and heated at 70 °C for 15 min before ThT signal was measured. Both H50Q and wild-type fibrils are irreversible. **d**, Cell seeding assay of H50Q and wild-type α -syn performed in HEK293T α -syn-A53T-YFP biosensor cells. Sonicated fibrils were transfected into biosensor cells using Lipofectamine. After 48 h, the number of fluorescent puncta indicating aggregated endogenous α -syn-A53T-YFP was counted (see Methods). H50Q fibrils have a higher seeding capacity than wild-type fibrils. Error bars represent standard deviation of four independent measurements. **e**, MTT toxicity assay of H50Q and wild-type α -syn performed in differentiated PC12 cells. H50Q and wild-type α -syn fibrils were introduced to culture medium and after incubation, cell mitochondrial activity was measured via MTT assay. H50Q requires less fibrils than wild type to significantly disrupt mitochondrial activity. Error bars represent standard deviation of a minimum of 7 and a maximum of 14 independent measurements. **** $P \leq 0.0001$. ** $P \leq 0.01$. NS, not significant $P > 0.05$. P values were calculated using an unpaired, two-tailed t -test with a 95% CI.

α -syn are associated with pathogenic, irreversible fibrils formed in the synucleinopathies.

We next characterized the kinetics of H50Q and wild-type fibril growth and found that H50Q aggregation has a shorter lag phase and therefore forms fibrils more rapidly than wild-type fibrils, consistent

with other studies (Fig. 4c)^{18,19,25}. We note that although H50Q fibrils have a lower max thioflavin T (ThT) signal, this is likely explained by different fibril polymorphs having differential ThT binding and not as being due to overall less fibril formation. We also find that H50Q fibrils have higher seeding capacity in HEK293T α -syn-A53T-YFP

Table 2 | Comparative solvation energy calculations

Fibril structure	Atomic solvation standard free energy of stabilization (kcal mol ⁻¹)		
	Per layer	Per chain	Per residue
H50Q narrow fibril	58	58 (pf A)	0.90 (pf A)
H50Q wide fibril	114	58 (pf A)	0.90 (pf A)
		54 (pf B)	1.00 (pf B)
WT rod cryo-EM (PDB 6CU7)	112	56	0.93
WT twister cryo-EM (PDB 6CU8)	66	33	0.80
WT solid-state NMR (PDB 2N0A)	58	58	0.91
WT rod (PDB 6H6B)	101	51	0.86
WT rod (PDB 6A6B)	122	61	0.96
Tau PHF (PDB 530L)	128	64	0.88
Tau pick's (PDB 6GX5)	98	98	1.00
TDP-43 SegA-sym (PDB 6N37)	73	37	1.00
Serum amyloid A (PDB 6MST)	116	58	1.07
FUS (PDB 5W3N)	41	41	0.66

pf, protofilament; WT, wild type.

biosensor cells²⁶ and significantly higher cytotoxicity to differentiated neuron-like rat pheochromocytoma (PC12), as measured by a reduction of mitochondrial activity and cell membrane integrity (Fig. 4d,e and Extended Data Fig. 8). These results are similar to other studies demonstrating the enhanced pathogenic properties of H50Q versus wild-type α -syn^{17–19,25}, and overall our findings support the idea that the differences in the structures of the H50Q fibrils compared to wild-type fibrils result in higher pathogenicity.

Discussion

The structures of α -syn fibrils determined here demonstrate that the H50Q mutation results in two new polymorphs, which we term narrow and wide fibrils. Structural alignments of wild-type and H50Q polymorphs reveal that residues 50–77 in all structures adopt a largely similar β -arch-like fold, which we previously termed the conserved kernel (Extended Data Fig. 9)¹⁴. However, different sequence segments of the conserved kernel assemble protofilaments into distinct fibril polymorphs (Fig. 1a,b). The wild-type rod polymorph utilizes residues from the preNAC region^{14,27,28}, the wild-type twister polymorph utilizes residues from the NACore¹⁴ and the H50Q wide fibril utilizes residues ⁵⁸KTKE⁶¹ to assemble protofilaments, while the H50Q narrow fibril forms a single protofilament structure (Fig. 1a,b). Therefore, the structures determined to date of wild-type and mutant α -syn fibrils demonstrate that the conserved kernel formed by residues 50–77 acts a modular building block to assemble protofilaments into distinct polymorphs in at least four different ways.

In all seven α -syn fibril structures, the conserved kernel features a cavity, possibly a solvent channel surrounded by T54, A56, K58, G73 and V74 (Figs. 1b and 2c and Extended Data Fig. 4b)^{14,27–29}. However, the H50Q polymorphs determined here are, so far, the only α -syn fibril structures to resolve density in this cavity. A hydrophilic cavity is also observed in the human serum amyloid A cryo-EM structure and two structures of immunoglobulin light chain fibrils^{22,30,31}. Therefore, full-length amyloid fibrils share the property of crystal structures of amyloid segments where the majority of the protein packs in a manner excluding water, while some water molecules can be seen hydrogen bonding with the backbone or polar side chains³². Recent structures of tau fibrils extracted from the brains of patients with chronic traumatic encephalopathy suggest that a

hydrophobic molecule may occupy a hydrophobic cavity present in the fibril core³³. This is also similar to crystal structures of amyloid segments where molecules such as polyethylene glycol can occupy cavities in the crystal packing³⁴.

The structures of H50Q polymorphs help to explain the differences in aggregation kinetics, seeding capacity and cytotoxicity between H50Q and wild-type α -syn that we and others have observed (Fig. 4b,d,e)^{18,19,25}. First, the observation that the H50Q mutation results in a large proportion of narrow fibrils formed from a single protofilament may help explain the faster aggregation kinetics of the mutant fibrils. Given that multiple molecules must come together to nucleate amyloid fibril growth, fibrils composed of a single protofilament may therefore have a shorter lag time than fibrils composed of two protofilaments, as half as many molecules are required to form the nucleus. Thus, the H50Q mutation may lower the barrier to nucleation and more readily lead to fibril formation. Previous studies on wild-type α -syn have demonstrated that minor species of single protofilament fibrils exist in polymorphic preparations, suggesting that wild-type α -syn can also form single protofilaments²⁰. However, here we observe that single protofilament structures dominate, suggesting that the H50Q mutation may tip the balance to favor single protofilament fibrils and shorter lag times.

Second, the higher seeding capacity of H50Q versus wild-type fibrils in our biosensor cell assays may be explained by the presence of a secondary nucleation mechanism. We speculate that the wide fibril species represents a step in a secondary nucleation pathway whereby residues ⁵⁸KTKE⁶¹ of protofilament A serve as a surface to catalyze the formation of protofilament B (Extended Data Fig. 10)³⁵. Several observations support this idea. (1) Protofilament B is never observed alone, and is only observed with protofilament A in the wide fibril, whereas protofilament A is observed alone in the narrow fibril. (2) The structure of protofilament A in both the narrow and wide fibril is nearly identical (r.m.s. deviation = 0.26 Å) and the helical twist of both the narrow and wide fibril is nearly identical (cross-over distance ~900 Å), suggesting that protofilament A acts as an unperturbed scaffold for protofilament B to grow off its side. (3) The wide fibril protofilament interface is exceedingly small, perhaps making it a labile interface where protofilament B can nucleate and elongate but eventually fall off, forming individual single protofilament fibrils. Given that we do not observe protofilament B alone, we speculate that over time protofilament B may convert into protofilament A, perhaps initiated by the switching of H50Q into a conformation that allows hydrogen bonding with K45 and subsequent formation of the Y39–T44–E46 hydrogen-bond triad. We speculate that the conversion of protofilament B to protofilament A may occur immediately before disassembling from the wide fibril to form individual narrow fibrils or after disassembly from the wide fibril (Extended Data Fig. 10).

We note that other mechanisms may lead to the differential seeding capacities that we observe; for instance, there may be a different seed fibril length distribution due to differing sensitivities to sonication between H50Q and wild-type fibrils. In addition, the monomeric α -syn-A53T-YFP may be easier or harder to seed by H50Q or wild-type fibrils due to species barrier effects. Although for the latter point, we note that the wild-type fibrils have only one residue different from the α -syn-A53T-YFP construct, whereas the H50Q fibrils have two differing residues, making the increased seeding potency of H50Q fibrils more noteworthy despite the greater difference in sequence.

Third, the ultrastructural arrangement of protofilaments, the ordered β -arch in protofilament A featuring a unique Y39–T44–E46 hydrogen-bond triad and the presence of islands 1 and 2 represent major structural differences in the H50Q fibrils compared to previously determined wild-type α -syn structures. These structural differences create new ordered surfaces on the fibrils that may enable

their heightened cytotoxicity compared to wild type. Previously, it has been shown that polyQ inclusions can sequester essential cellular proteins and that resupply of sequestered proteins ameliorated toxicity and reduced inclusion size, presumably by coating the fibrils and rendering them inert³⁶. This highlights the importance of surfaces of fibrils in mediating cytotoxicity and suggests that different surface properties of amyloid fibrils may explain differential cytotoxicities. Indeed, others have shown that two polymorphs of wild-type α -syn could be homogeneously prepared and that these polymorphs had different cytotoxicities to SH-SY5Y cells, suggesting that differences in structure, including exposed fibril surfaces, result in different cytotoxicity³⁷. Therefore, we propose that the large structural differences observed in our H50Q fibrils compared to wild type could mediate the higher cytotoxicity of H50Q versus wild-type fibrils.

Hereditary mutations in α -syn are largely clustered in the pre-NAC region (residues 47–56) away from the NACore region (residues 68–78). The E46K hereditary mutation is predicted to disrupt a key salt bridge that forms between E46 and K80, potentially disturbing the fibril core. Consistent with this, NMR studies have shown large chemical shifts for residues in the fibril core of E46K fibrils³⁸. Interestingly, in H50Q protofilament A, E46 participates in a hydrogen-bond network with T44 and Y39; therefore, mutation to lysine may disrupt this network making the formation of the N-terminal 36–46 β -arch mutually exclusive with the E46K hereditary mutation. Further, the observation that the E46–K80 hydrogen bond is not maintained in protofilament A—unlike the previous wild-type rod polymorphs and protofilament B—demonstrates that this interaction is not necessary to maintain the overall wild-type fold, and that, in the case of the E46K hereditary mutation, it is the change to lysine and unfavorable juxtaposition of the positively charged E46K and K80 that explains the rearrangement of the fibril core as indicated by NMR³⁸. For mutations A30P and A53T, NMR studies of fibrils show small perturbations in chemical shifts and secondary structures at sites proximal to the mutation, suggesting that the overall fold of the fibril is largely unchanged^{38,39}. Here, the H50Q mutation seems to lie somewhere in the middle of the A30P, A53T and E46K mutations; in that, H50Q enforces a new conformation of the N terminus of the fibril core, disrupts previously observed protofilament interfaces and creates a new protofilament interface, while maintaining a conserved β -arch fold in the fibril core. Further work is needed to determine the exact structural effects of other α -synuclein hereditary mutations.

Overall, our results demonstrate that the H50Q hereditary mutation leads to new fibril polymorphs that have more rapid fibril-forming kinetics, higher seeding capacity and higher cytotoxicity. These findings provide a starting point for understanding the structural basis of mutation-enhanced pathogenesis in the synucleinopathies.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41594-019-0322-y>.

Received: 23 June 2019; Accepted: 30 September 2019;

Published online: 06 November 2019

References

- Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M. & Goedert, M. α -Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl Acad. Sci. USA* **95**, 6469–6473 (1998).

- Grazia Spillantini, M. et al. Filamentous α -synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies. *Neurosci. Lett.* **251**, 205–208 (1998).
- Spillantini, M. G. & Goedert, M. Neurodegeneration and the ordered assembly of α -synuclein. *Cell Tissue Res.* **373**, 137–148 (2018).
- Singleton, A. B. et al. α -Synuclein locus triplication causes parkinson's disease. *Science* **302**, 841–841 (2003).
- Chartier-Harlin, M.-C. et al. α -Synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* **364**, 1167–1169 (2004).
- Ibáñez, P. et al. Causal relation between α -synuclein gene duplication and familial Parkinson's disease. *Lancet* **364**, 1169–1171 (2004).
- Serpell, L. C., Berriman, J., Jakes, R., Goedert, M. & Crowther, R. A. Fiber diffraction of synthetic α -synuclein filaments shows amyloid-like cross- β conformation. *Proc. Natl Acad. Sci. USA* **97**, 4897–4902 (2000).
- Luk, K. C. et al. Pathological α -synuclein transmission initiates Parkinson-like neurodegeneration in nontransgenic mice. *Science* **338**, 949–953 (2012).
- Tycko, R. Solid-state NMR studies of amyloid fibril structure. *Annu. Rev. Phys. Chem.* **62**, 279–299 (2011).
- Bai, X., McMullan, G. & Scheres, S. H. W. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* **40**, 49–57 (2015).
- Sawaya, M. R. et al. Atomic structures of amyloid cross- β spines reveal varied steric zippers. *Nature* **447**, 453–457 (2007).
- Rodríguez, J. A. et al. Structure of the toxic core of α -synuclein from invisible crystals. *Nature* **525**, 486–490 (2015).
- Colvin, M. T. et al. Atomic resolution structure of monomeric $A\beta_{42}$ amyloid fibrils. *J. Am. Chem. Soc.* **138**, 9663–9674 (2016).
- Li, B. et al. Cryo-EM of full-length α -synuclein reveals fibril polymorphs with a common structural kernel. *Nat. Commun.* **9**, 3609 (2018).
- Proukakis, C. et al. A novel α -synuclein missense mutation in Parkinson disease. *Neurology* **80**, 1062–1064 (2013).
- Appel-Cresswell, S. et al. Alpha-synuclein p.H50Q, a novel pathogenic mutation for Parkinson's disease. *Mov. Disord.* **28**, 811–813 (2013).
- Rutherford, N. J., Moore, B. D., Golde, T. E. & Giasson, B. I. Divergent effects of the H50Q and G51D SNCA mutations on the aggregation of α -synuclein. *J. Neurochem.* **131**, 859–867 (2014).
- Porcari, R. et al. The H50Q mutation induces a 10-fold decrease in the solubility of α -synuclein. *J. Biol. Chem.* **290**, 2395–2404 (2015).
- Khalaf, O. et al. The H50Q mutation enhances α -synuclein aggregation, secretion, and toxicity. *J. Biol. Chem.* **289**, 21856–21876 (2014).
- Dearborn, A. D. et al. α -Synuclein amyloid fibrils with two entwined, asymmetrically associated protofibrils. *J. Biol. Chem.* **291**, 2310–2318 (2016).
- Fitzpatrick, A. W. P. et al. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185–190 (2017).
- Liberta, F. et al. Cryo-EM fibril structures from systemic AA amyloidosis reveal the species complementarity of pathological amyloids. *Nat. Commun.* **10**, 1104 (2019).
- Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat. Struct. Mol. Biol.* **26**, 619–627 (2019).
- Murray, D. T. et al. Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. *Cell* **171**, 615–627.e16 (2017).
- Ghosh, D. et al. The Parkinson's disease-associated H50Q mutation accelerates α -synuclein aggregation in vitro. *Biochemistry* **52**, 6925–6927 (2013).
- Prusiner, S. B. et al. Evidence for α -synuclein prions causing multiple system atrophy in humans with parkinsonism. *Proc. Natl Acad. Sci. USA* **112**, E5308–E5317 (2015).
- Li, Y. et al. Amyloid fibril structure of α -synuclein determined by cryo-electron microscopy. *Cell Res.* **28**, 897–903 (2018).
- Guerrero-Ferreira, R. et al. Cryo-EM structure of α -synuclein fibrils. *eLife* **7**, e36402 (2018).
- Tuttle, M. D. et al. Solid-state NMR structure of a pathogenic fibril of full-length human α -synuclein. *Nat. Struct. Mol. Biol.* **23**, 409–415 (2016).
- 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).
- Swuec, P. et al. Cryo-EM structure of cardiac amyloid fibrils from an immunoglobulin light chain AL amyloidosis patient. *Nat. Commun.* **10**, 1269 (2019).
- Nelson, R. et al. Structure of the cross- β spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
- Falcon, B. et al. Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019).
- Liu, C. et al. Out-of-register β -sheets suggest a pathway to toxic amyloid aggregates. *Proc. Natl Acad. Sci. USA* **109**, 20913–20918 (2012).
- Cohen, S. I. A. et al. Proliferation of amyloid- β 42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl Acad. Sci. USA* **110**, 9758–9763 (2013).

36. Hosp, F. et al. Spatiotemporal proteomic profiling of Huntington's disease inclusions reveals widespread loss of protein function. *Cell Rep.* **21**, 2291–2303 (2017).
37. Bousset, L. et al. Structural and functional characterization of two alpha-synuclein strains. *Nat. Commun.* **4**, 2575 (2013).
38. Lemkau, L. R. et al. Site-specific perturbations of alpha-synuclein fibril structure by the Parkinson's disease associated mutations A53T and E46K. *PLoS ONE* **8**, e49750 (2013).
39. Lemkau, L. R. et al. Mutant protein A30P α -synuclein adopts wild-type fibril structure, despite slower fibrillation kinetics. *J. Biol. Chem.* **287**, 11526–11532 (2012).

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Protein purification. Full-length α -syn wild-type and H50Q mutant proteins were expressed and purified according to a published protocol¹⁴. The bacterial induction started at an optical density (OD_{600}) of ~ 0.6 with 1 mM IPTG for 6 h at 30°C. The collected bacteria were lysed with a probe sonicator for 10 min in an iced water bath. After centrifugation, the soluble fraction was heated in boiling water for 10 min and then titrated with HCl to pH 4.5 to remove the unwanted precipitants. After adjusting to neutral pH, the protein was dialyzed overnight against Q Column loading buffer (20 mM Tris-HCl pH 8.0). The next day, the protein was loaded onto a HiPrep Q 16/10 column and eluted using elution buffer (20 mM Tris-HCl, 1 M NaCl, pH 8.0). The eluent was concentrated using Amicon Ultra-15 centrifugal filters (Millipore Sigma) to ~ 5 ml. The concentrated sample was further purified with size-exclusion chromatography through a HiPrep Sephacryl S-75 HR column in 20 mM Tris, pH 8.0. The purified protein was dialyzed against water, concentrated to 3 mg ml⁻¹ and stored at 4°C. The concentration of the protein was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, catalog no. 23225).

Fibril preparation and optimization. Both wild-type and H50Q fibrils were grown under the same condition: 300 μ M purified monomers, 15 mM tetrabutylphosphonium bromide, shaking at 37°C for 2 weeks.

Negative stain transmission electron microscopy (TEM). The fibril sample (3 μ l) was spotted onto a freshly glow-discharged carbon-coated electron microscopy grid. After 1 min, 6 μ l uranyl acetate (2% in aqueous solution) was applied to the grid for 1 min. The excess stain was removed by a filter paper. Another 6 μ l uranyl acetate was applied to the grid and immediately removed. The samples were imaged using an FEI T20 electron microscope.

ThT binding assay. Purified α -syn monomers (50 μ M) were adequately mixed with 20 μ M ThT and added into a 96-well-plate. Samples were incubated at 37°C for 2 d with 600 r.p.m. double orbital shaking. The ThT signal was monitored using the FLUOstar Omega Microplate Reader (BMG Labtech) at an excitation wavelength of 440 nm and an emission wavelength of 490 nm.

SDS stability. SDS was diluted in water to make SDS solutions at 2.5, 5, 10 and 15%. Fibrils at the end of the ThT assay were treated with SDS solution at 5:1 volume ratio to obtain SDS concentrations of 0.5, 1, 2 and 3%. Each solution was transferred to three microcentrifuge tubes and heated at 70°C for 15 min. After treatment, the ThT signal was obtained. The 0% SDS solution without heating was treated with an equal amount of water and used for normalization.

Cell lines. HEK293T biosensor cells expressing α -syn-A53T-YFP were a generous gift from the laboratory of M. Diamond. PC12 cells originate from the American Type Culture Collection (ATCC CRL-1721). Cells were not authenticated nor tested for mycoplasma infection in our hands.

Fibril seeding aggregation in cells. We performed the biosensor cell seeding assay based on a previously published protocol²⁶. Briefly, the assay works as follows: exogenous, unlabeled fibrils are transfected into HEK293T cells expressing α -syn-A53T-YFP. Seeded aggregation of endogenously expressed α -syn-A53T-YFP is monitored by formation of fluorescent puncta. The puncta represent a condensation of α -syn-A53T-YFP as a result of seeding by exogenous H50Q or wild-type fibrils.

HEK293T cells expressing full-length α -syn containing the hereditary A53T mutation were grown in DMEM (4 mM L-glutamine and 25 mM D-glucose) supplemented with 10% FBS, 1% penicillin/streptomycin. Trypsin-treated HEK293T cells were collected, seeded on flat 96-well plates at a concentration of 4×10^4 cells per well in 200 μ l culture medium per well and incubated in 5% CO₂ at 37°C for 18 h.

α -syn fibrils were prepared by diluting with Opti-MEM (Life Technologies) and sonicating in a water bath sonicator for 10 min. The fibril samples were then mixed with Lipofectamine 2000 (Thermo Fisher Scientific) and incubated for 15 min and then added to the cells. The actual volume of Lipofectamine 2000 was calculated based on the dose of 1 μ l per well. After 48 h of transfection, the cells were imaged with a Celigo Imaging Cytometer (Nexcelom Bioscience). Fluorescent images were processed in ImageJ to count the number of seeded cells. A buffer-treated control was used for normalization.

MTT mitochondrial activity assay. The addition of sonicated fibrils to nerve growth factor-differentiated PC12 cells is a well-established assay to measure cytotoxicity of amyloid fibrils^{40,41}. Use of this neuron-like cell line allows us to obtain a biologically relevant assay for cytotoxicity. Our MTT mitochondrial activity assay followed our previously published protocol¹⁴, which was adapted from the Provost and Wallert laboratories⁴². Thiazolyl blue tetrazolium bromide for the MTT cell toxicity assay was purchased from Millipore Sigma (catalog no. M2128-1G). PC12 cells were plated in 96-well plates with nerve growth factor. The cells were incubated for 2 d in an incubator with 5% CO₂ at 37°C. The cells were treated with different concentrations of α -syn fibrils (200 nM, 500 nM, 1,000 nM,

2,000 nM), which were sonicated in a water bath sonicator for 10 min before being added to the cells. After 18 h incubation, MTT was added to every well and the plate was returned to the incubator for 3.5 h. Absorbance was measured at 570 nm to determine the MTT signal and at 630 nm to determine background. The data were normalized to those from cells treated with 1% SDS to obtain a value of 0%, and to those from cells treated with PBS to obtain a value of 100%.

Lactate dehydrogenase assay. The lactate dehydrogenase viability assay was performed using the CytoTox-ONE Homogeneous Membrane Integrity Assay (Promega, catalog no. G7891). PC12 cells were cultured and differentiated with the same protocol as described in the MTT assay. Different concentrations of α -syn fibrils (200 nM, 500 nM, 1,000 nM, 2,000 nM) were added to the cells for 18 h of incubation with 5% CO₂ at 37°C. The assay was carried out in a 96-well plate and the fluorescence readings were taken in the FLUOstar Omega Microplate Reader (excitation 560 nm, emission 590 nm; BMG Labtech). PBS- and 0.2% Triton-X100-treated cells were used as negative and positive controls, respectively, for normalization.

Cryo-EM data collection and processing. Fibril solution (2 μ l) was applied to a baked and glow-discharged Quantifoil 1.2/1.3 electron microscope grid and plunge-frozen into liquid ethane using a Vitrobot Mark IV (FEI). Data were collected on a Titan Krios (FEI) microscope equipped with a Gatan Quantum LS/K2 Summit direct electron detection camera (operated with 300 kV acceleration voltage and slit width of 20 eV). Counting mode movies were collected on a Gatan K2 Summit direct electron detector with a nominal physical pixel size of 1.07 Å per pixel with a dose per frame 1.2 e⁻/Å². A total of 30 frames with a frame rate of 5 Hz were taken for each movie, resulting in a final dose 36 e⁻/Å² per image. Automated data collection was driven by the Leginon automation software package⁴³.

Micrographs containing crystalline ice were used to estimate the anisotropic magnification distortion using `mag_distortion_estimate`⁴⁴. CTF estimation was performed using CTFFIND 4.1.8 on movie stacks with a grouping of three frames and correction for anisotropic magnification distortion⁴⁵. Unblur⁴⁶ was used to correct beam-induced motion with dose weighting and anisotropic magnification correction, resulting in a physical pixel size of 1.065 Å per pixel.

All particle picking was performed manually using EMAN2 e2helixboxer.py⁴⁷. We manually picked two groups of particles for further data processing: the first group was composed of all fibrils and the second group was composed of wide fibrils. For the first group, particles were extracted in RELION using the 90% overlap scheme into 1,024- and 288-pixel boxes. Classification, helical reconstruction and three-dimensional (3D) refinement were used in RELION as described²³. For the first group of all particles, we isolated narrow fibrils during two-dimensional (2D) classification and subsequently processed them as a separate data set. 2D classifications of narrow fibril 1,024-pixel boxes were used to estimate helical parameters. We performed 3D classification with the estimated helical parameters for narrow fibrils and an elongated Gaussian blob as an initial model to generate starting reconstructions. We ran additional 3D classifications using the preliminary reconstructions from the previous step to select for particles contributing to homogenous classes (stable helicity and separation of β -strands in the x - y plane). Typically, we performed Class3D jobs with $K=3$ and manual control of the `tau_fudge` factor and `healpix` to reach a resolution of ~ 5 -6 Å to select for particles that contributed to the highest resolution class for each structure. We employed Refine3D on a final subset of narrow fibril particles with 288-pixel box size to obtain the final reconstruction. We performed the map-map Fourier shell correlation (FSC) with a generous, soft-edged solvent mask and high-resolution noise substitution in RELION PostProcess, resulting in a resolution estimate of 3.3 Å.

We extracted particles from the wide fibril dataset using 1,024- and 686-pixel boxes. 2D classifications of 1,024- and 686-pixel boxes were used to estimate helical parameters. 2D classifications of 686-pixel boxes were used to further isolate only wide fibril segments, since there were still some other fibril species that were included due to the fact that we could not separate all fibril species perfectly during manual picking. Once a homogenous set of wide fibrils was obtained during 2D classification of 686-pixel boxes, we performed a 3D reconstruction using an elongated Gaussian blob as an initial model. The asymmetry present in the 686 box, 2D class averages of the wide fibril (Extended Data Fig. 2d) prompted us to use a helical rise of 4.8 Å and C1 symmetry, due to the fact that if a twofold symmetry were present in the fibrils, 2D class averages would display a mirror symmetry across the fibril axis. After an initial 2D model was generated for the wide fibril, we re-extracted all tubes corresponding to those particles included in the final subset of wide fibril 686-pixel boxes with a box size of 224 pixels. All 224-pixel boxes were subjected to multiple rounds of 3D classification using the initial 686-pixel box wide fibril reconstruction as a reference. We refined the final subset of particles using Refine3D to a resolution of 3.6 Å. We performed resolution estimation as described above for the narrow fibrils.

Atomic model building. We sharpened both the narrow and wide fibril reconstructions using `phenix.auto_sharpen`⁴⁸ at the resolution cut-off indicated by the map-map FSC and subsequently built atomic models into the refined maps with Coot⁴⁹. We built the model for the narrow fibril de novo using previous

structures of wild-type α -synuclein fibrils as guides. To build the wide fibril model, we made a copy of one chain of the narrow fibril structure and rigid-body fit it into the second protofilament density observed in the wide fibril reconstruction. This resulted in the wide fibril being composed of one protofilament nearly identical to the narrow fibril, and one protofilament with less ordered N and carboxy (C) termini, therefore resulting in an asymmetric double protofilament structure.

For both narrow and wide fibrils, we generated a five-layer model to maintain local contacts between chains in the fibril during structure refinement. We performed automated structure refinement for both narrow and wide fibrils using phenix.real_space_refine⁵⁰. We employed hydrogen bond distance and angle restraints for backbone atoms participating in β -sheets and side chain hydrogen bonds during automated refinements. We performed comprehensive structure validation of all our final models in Phenix.

Although we did not include coordinates in our final models for additional residues that could occupy islands 1 and 2 neighboring protofilament A, because we could not be certain which residues occupy those densities, we built several speculative models (Extended Data Fig. 4). For island 2, we assumed that there was a short disordered linker between residue 36 of protofilament A and island 2, resulting in the residues occupying island 2 forming a tight interface with ³⁶GVLVYG⁴¹ of the fibril core. We noticed that the sequence ³²KTKE³⁵ immediately precedes the last ordered residue of protofilament A, G36. The sequence ³²KTKE³⁵ often forms the bends that connect straight β -strands in the ordered fibril core, so we assumed that this sequence would be a good candidate to form the tight bend that connects G36 to island 2. Therefore, we modeled in residues ²⁶VAEAAAG³¹ into island 2. These residues satisfied the requirements of having short, hydrophobic side chains forming the tight interface with residues V38 and I40 from the fibril core.

For island 1, we assumed that the densities come either from the N terminus of protofilament A or from the C terminus of protofilament B. For the former case, we assume that there is a minimum of approximately eight residues that are disordered between the end of island 2 and the beginning of island 1 (see Extended Data Fig. 4). This is because there is 27 Å between the end of island 2 (V26) and the beginning of island 1, and we assume a minimum of ~3.3 Å per residue. Therefore, we threaded eight residues at a time from the region ¹MDVFMKGLSKAKEGVVAAA¹⁹ onto a β -strand backbone placed in the island 1 density to identify candidate octamers. While most sequences could not plausibly occupy island 1, due to steric clashes with the preNAC region of the fibril core, or due to glycine residues occupying positions where there were obvious side chain densities, several candidate sequences were identified (Extended Data Fig. 3).

We followed a similar protocol to identify possible sequences from the C terminus. In this case, we assume that these residues could come from either a largely disordered protofilament B molecule in the narrow fibril or an ordered protofilament B molecule as seen in the wide fibril (see Extended Data Fig. 3). Here, we assume there is a minimum of ~15 residues from the last ordered residue of the C terminus of protofilament B—since we assume a minimum of ~3.3 Å per residue and there is ~45 Å between K97 of protofilament B and the beginning of island 1 (see Extended Data Fig. 3). Therefore, we threaded all possible octamers from the region ¹¹²ILEDMPVDPDNEAYEMPSEEGYQDYEP¹⁴⁰ onto a β -strand backbone occupying island 1 to identify candidate sequences following the same criterion as above.

We created the speculative model of an H50Q double protofilament containing a homomeric preNAC interface by aligning a single chain from the H50Q narrow protofilament with the helical axis of the wild-type ‘rod’ structure (6CU7) and applying a pseudo-2(1) helical symmetry to generate a symmetrically related second chain.

Energetic calculation. The stabilization energy is an adaptation of the solvation free energy described previously⁵¹, in which the energy is calculated as the sum of products of the area buried of each atom and its corresponding atomic solvation parameter (ASP). ASPs were taken from our previous work⁵¹. Area buried is calculated as the difference in solvent accessible surface area (SASA) of the reference state (that is, the unfolded state) and the SASA of the folded state. The SASA of residue i of the unfolded state was approximated as the SASA of residue i in the folded structure after removal of all other atoms except the main chain atoms of residues $i-1$ and $i+1$. The SASA of the folded state was measured for each atom in the context of all amyloid fibril atoms. Fibril coordinates were extended by symmetry by three to five chains on either side of the reported molecule, to ensure the energetic calculations were representative of the majority of molecules in a fibril, rather than a fibril end. To account for energetic stabilization of main chain hydrogen bonds, the ASP for backbone N/O elements was reassigned from -9 to 0 if they participated in a hydrogen bond. Similarly, if an asparagine or glutamine side chain participated in a polar ladder (two hydrogen bonds per amide), and was shielded from solvent (SASAfolded < 5 Å²), the ASPs of the side chain N and O elements were reassigned from -9 to 0. Last, the ASP of ionizable atoms (in, for example, Asp, Glu, Lys, His, Arg, N-terminal amine, or C-terminal carboxylate) were assigned the charged value (-37/-38), unless the atoms participated in a buried ion pair, defined as a pair of complementary ionizable atoms within 4.2 Å distance of each other, each with SASAfolded < 40 Å². In that case, the ASP of the ion pair was reassigned to -9. In the energy diagrams, a single color is assigned to each residue, rather than each atom. The color corresponds to the sum of solvation

free energy values of each of the atoms in the residue. The energy reported for FUS in Table 2 is the average over 20 NMR models. The standard deviation is 1.8 kcal mol⁻¹.

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

Data availability

All structural data have been deposited into the Protein Database (PDB) and the Electron Microscopy Data Bank (EMDB) with the following accession codes: H50Q narrow fibril (PDB 6PEO, EMD-20328) and H50Q wide fibril (PDB 6PES, EMD-20331). All other data are available from the authors upon reasonable request.

References

- Wakabayashi, M. & Matsuzaki, K. Formation of amyloids by A β -(1–42) on NGF-differentiated PC12 cells: roles of gangliosides and cholesterol. *J. Mol. Biol.* **371**, 924–933 (2007).
- Ono, K., Condron, M. M. & Teplow, D. B. Structure–neurotoxicity relationships of amyloid β -protein oligomers. *Proc. Natl Acad. Sci. USA* **106**, 14745–14750 (2009).
- van Meerloo, J., Kaspers, G. J. L. & Cloos, J. in *Cancer Cell Culture: Methods and Protocols* (ed. Cree, I. A.) 237–245 (Humana Press, 2011).
- Suloway, C. et al. Automated molecular microscopy: the new legion system. *J. Struct. Biol.* **151**, 41–60 (2005).
- Grant, T. & Grigorieff, N. Automatic estimation and correction of anisotropic magnification distortion in electron microscopes. *J. Struct. Biol.* **192**, 204–208 (2015).
- Rohou, A. & Grigorieff, N. CTFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
- Grant, T. & Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* **4**, e06980 (2015).
- Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
- Terwilliger, T. C., Sobolev, O. V., Afonine, P. V. & Adams, P. D. Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr. D* **74**, 545–559 (2018).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
- Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).
- Eisenberg, D. S., Wesson, M. & Yamashita, M. Interpretation of protein folding and binding with atomic solvation parameters. *Chem. Scr.* **29A**, 217–222 (1989).

Acknowledgements

We thank H. Zhou for use of Electron Imaging Center for Nanomachines (EICN) resources and P. Ge for assistance in cryo-EM data collection. We acknowledge the use of instruments at the EICN supported by NIH (grant nos. 1S10RR23057 and 1S10OD018111), NSF (grant no. DBI-1338135) and CNSI at UCLA. The authors acknowledge grant nos. NIH AG 060149, NIH AG 054022, NIH AG061847 and DOE DE-FC02-02ER63421 for support. D.R.B. was supported by the National Science Foundation Graduate Research Fellowship Program.

Author contributions

D.R.B. and B.L. designed experiments and performed data analysis. B.L. and C.S. expressed and purified the α -syn protein. B.L. grew fibrils of α -syn and performed biochemical experiments. D.R.B. and B.L. prepared cryo-EM samples and performed cryo-EM data collection. B.L. and W.F. selected filaments from cryo-EM images. D.R.B. performed cryo-EM data processing and built the atomic models. M.R.S. wrote the software for and D.R.B. carried out solvation energy calculations. All authors analyzed the results and D.R.B. wrote the manuscript with input from all authors. L.J. and D.S.E. supervised and guided the project.

Competing interests

D.S.E. is an advisor and equity shareholder in ADRx, Inc.

Additional information

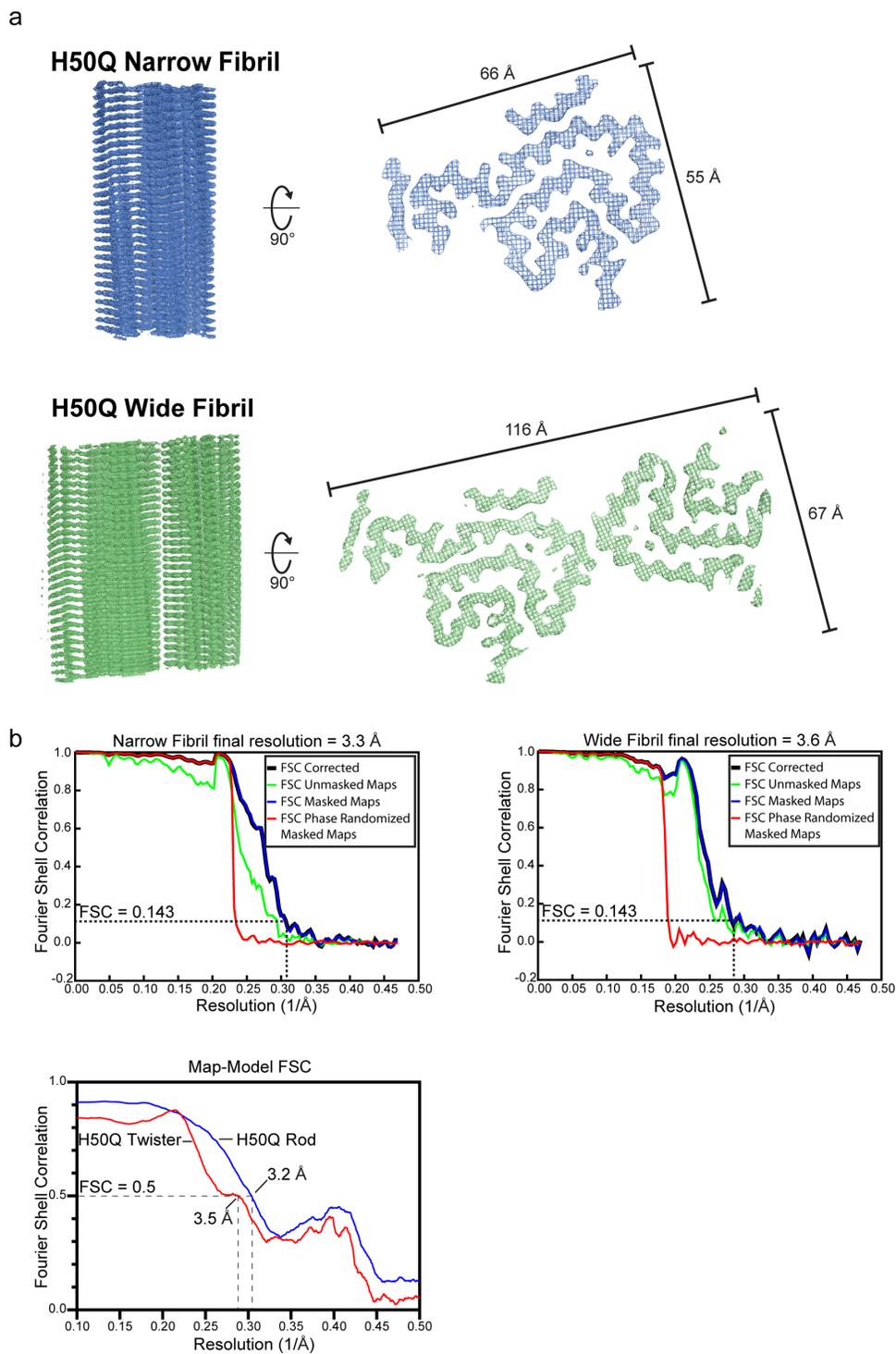
Extended data is available for this paper at <https://doi.org/10.1038/s41594-019-0322-y>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41594-019-0322-y>.

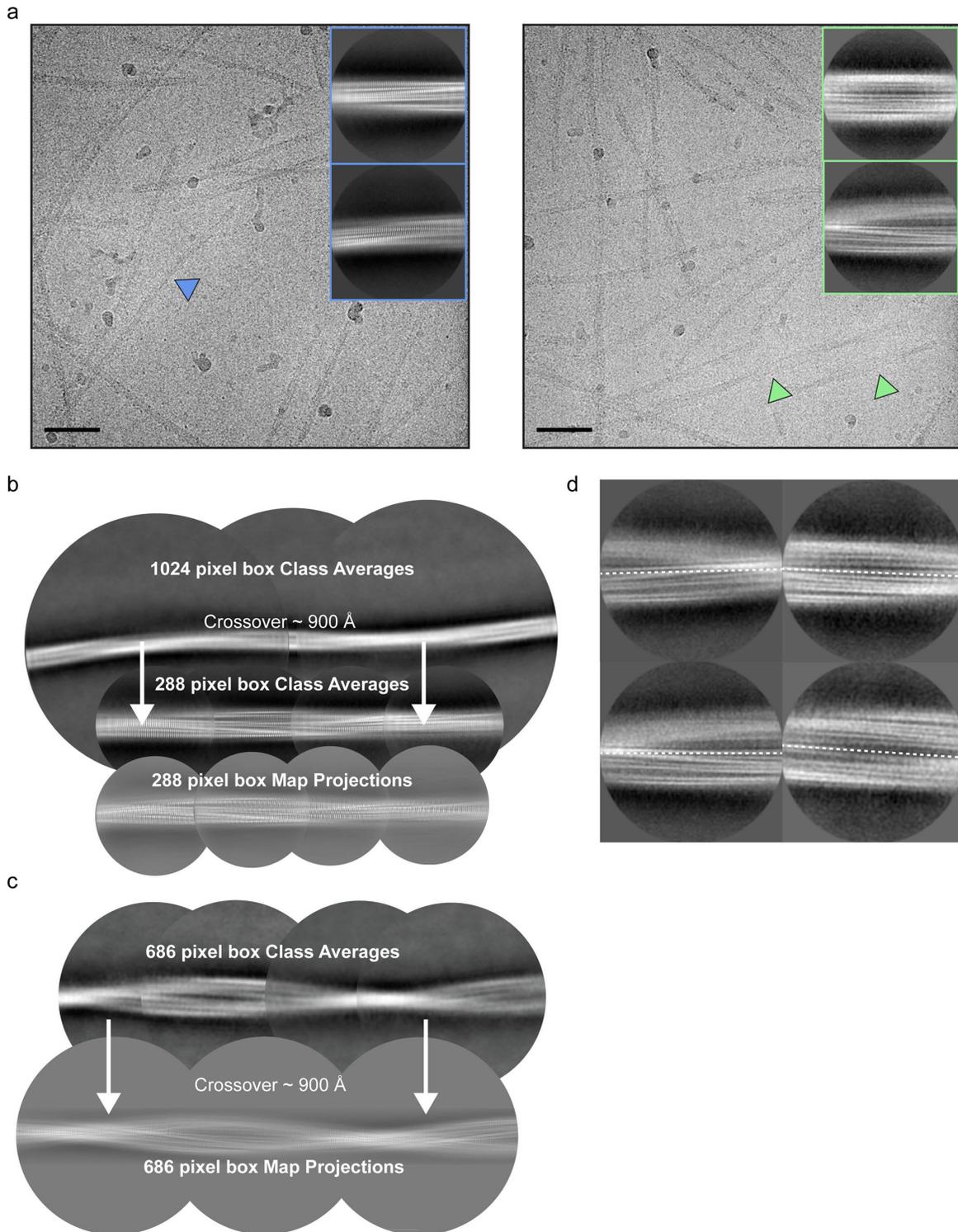
Correspondence and requests for materials should be addressed to L.J. or D.S.E.

Peer review information Inés Chen was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

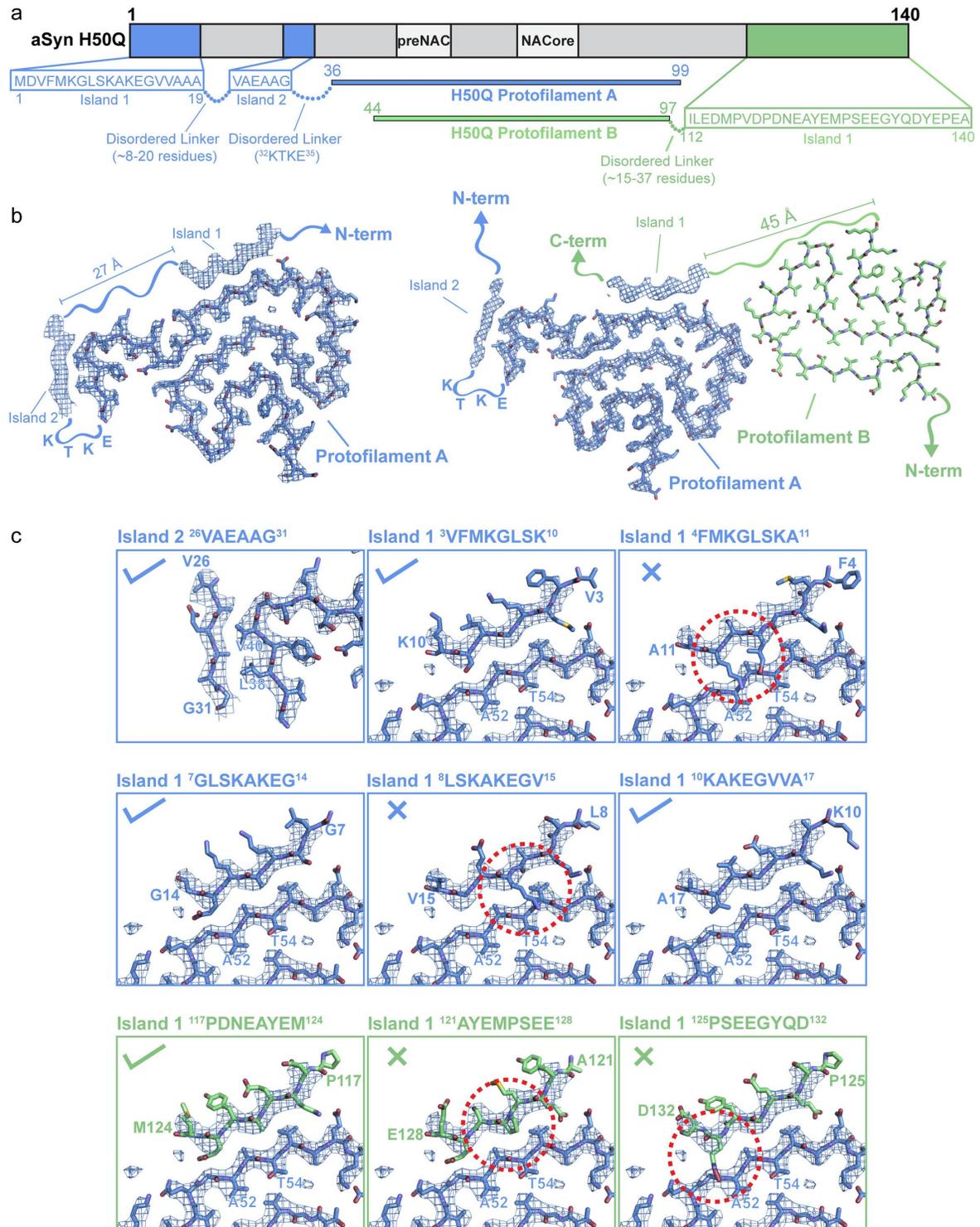


Extended Data Fig. 1 | Fourier Shell Analysis. a) Helical reconstructions of Narrow and Wide Fibrils with minimum and maximum widths labeled. b) Gold-standard half map FSC curves for Narrow (top, left) and Wide (top, right) Fibrils. Map-model FSC curve for Narrow and Wide Fibrils (bottom).



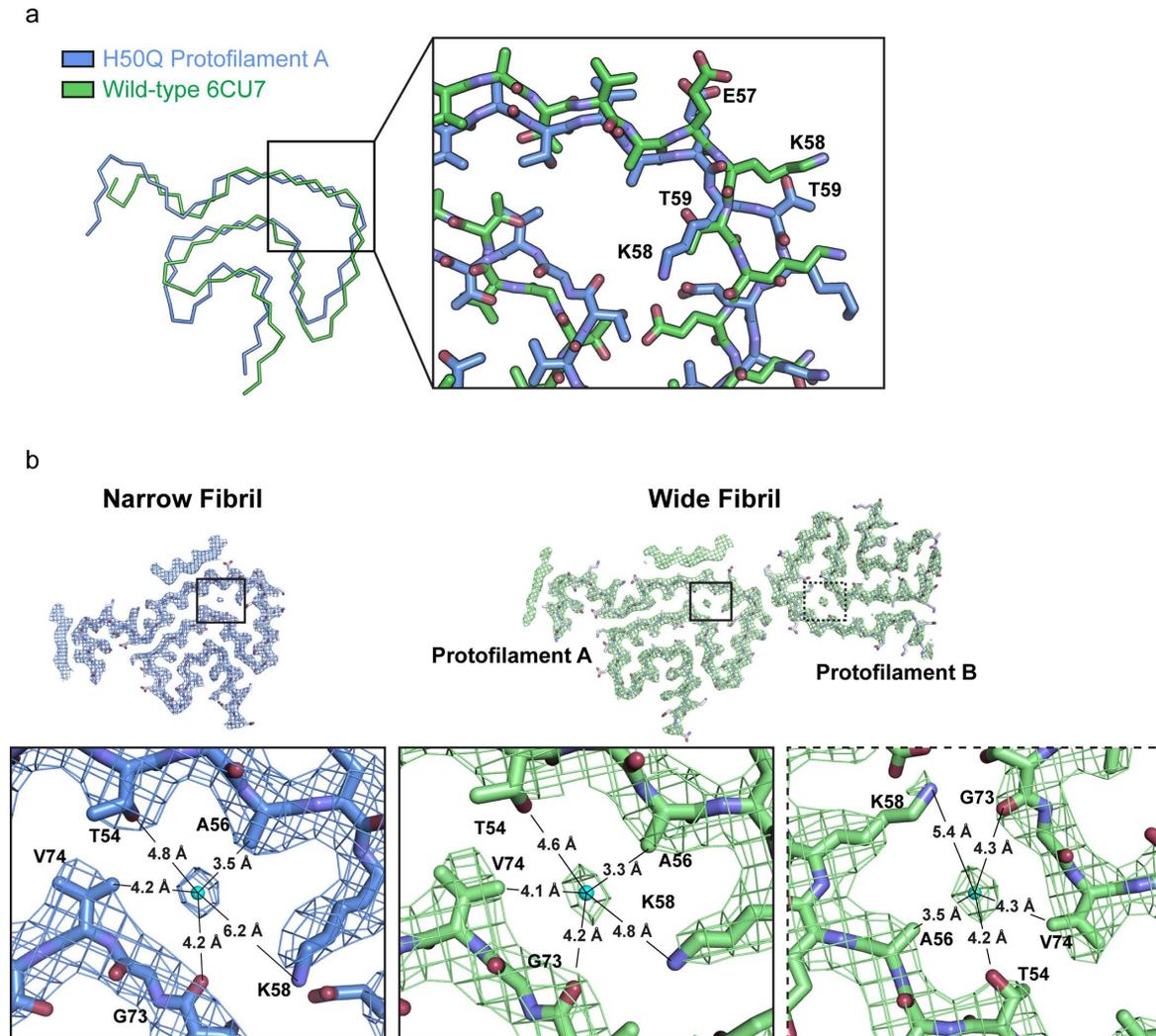
Extended Data Fig. 2 | see figure caption on next page.

Extended Data Fig. 2 | Cryo-EM images and processing. a) Cryo-EM micrographs and 2D class averages of Narrow (left) and Wide (right) Fibrils. Scale bar = 50 nm. b) 1024 and 288 pixel box size class averages of the Narrow Fibril used to determine crossover distance. 288 pixel box map projections match 2D class averages. c) 686 pixel box size class averages used to determine crossover distance. 686 pixel box map projections match 2D class averages. d) Wide Fibril class averages with a 320 pixel box demonstrate a lack of two-fold symmetry across the fibril axis.

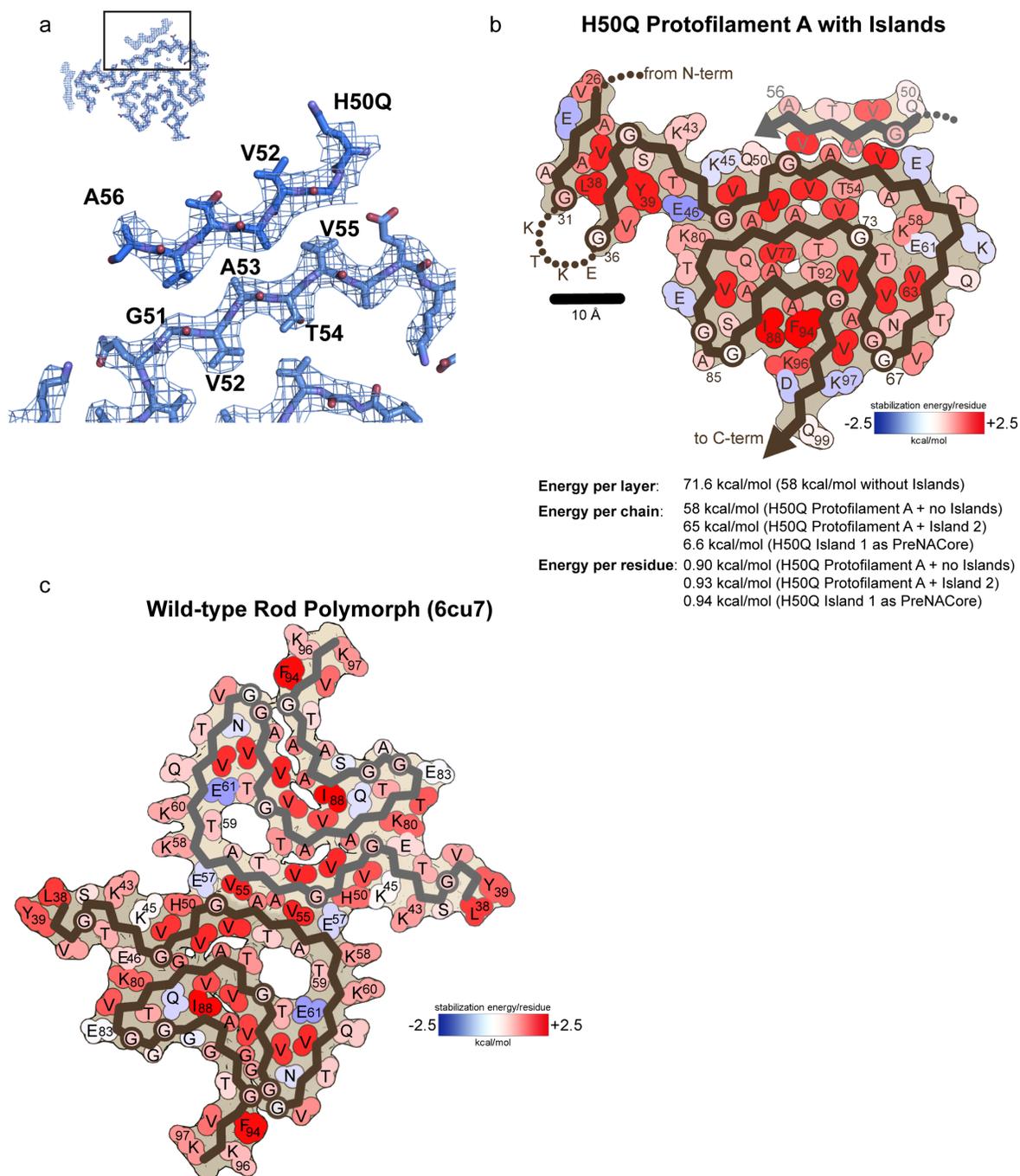


Extended Data Fig. 3 | see figure caption on next page.

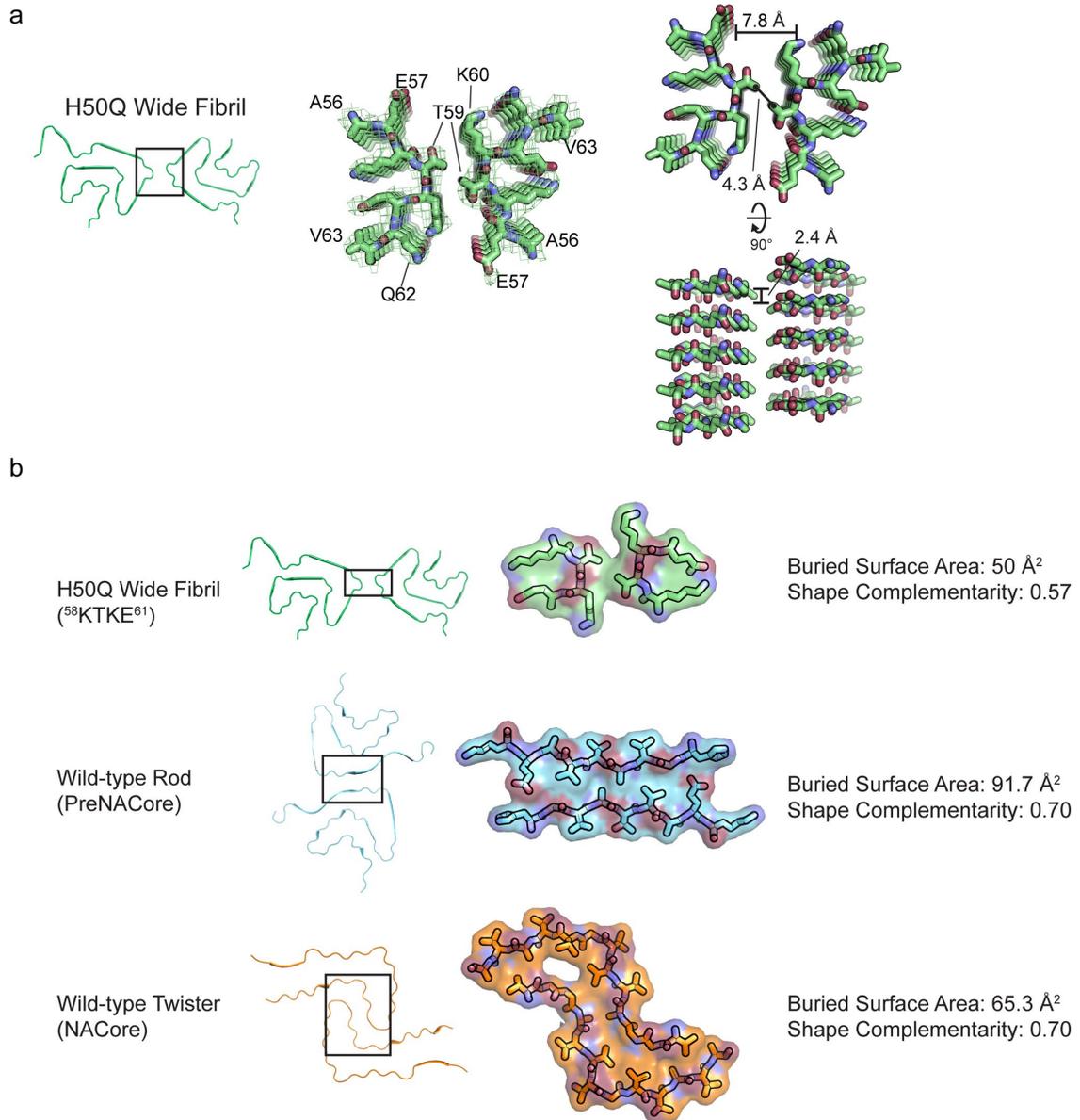
Extended Data Fig. 3 | Speculative Atomic Models for Islands 1 and 2. a) Schematic illustrating possible sequences occupying Islands 1 and 2. 8mers from residues 1-19 and 112-140 were considered as possibilities to occupy Island 1. Island 2 is considered to consist of residues 26-31 followed by a disordered linker formed by residues ³²KTKE³⁵. b) Illustration of possible regions from either Protofilament A (left) or Protofilament B (right) that could occupy Island 1. Note that residues from the N-terminus of Protofilament A could account for Island 1 in both the Narrow and Wide Fibril; however, only the Narrow Fibril model is shown here. Island 2 is thought to be formed by the N-terminus of Protofilament A in both Narrow and Wide Fibrils. c) Speculative models for Islands 1 and 2. Check marks indicate plausible models while X's indicate implausible models. Island 1 models are from either the N-terminus of Protofilament A (blue panels) or the C-terminus of Protofilament B (green panels). Examples of sequences that were found to not be allowed to occupy Island 1 are shown with red dashed circles highlighting steric clashes or β -strand breaking proline residues.



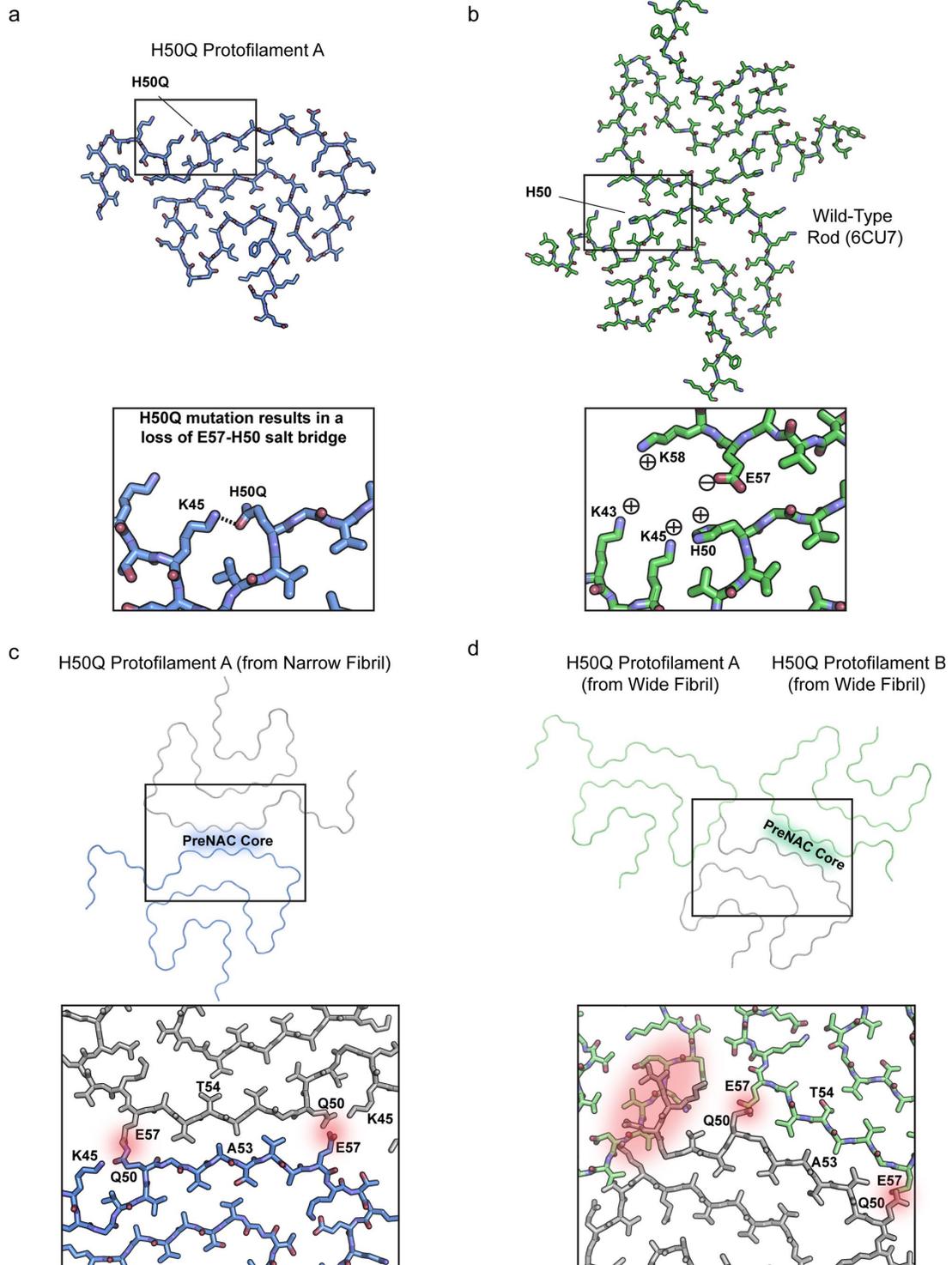
Extended Data Fig. 4 | Alternate conformations of K58 and T59 and potential solvent molecules in the α -syn β -arch cavity. a) Wild-type and H50Q fibrils display alternate conformations of K58 and T59. We note that in order for the Wide Fibril to form, T59 needs to be facing away from the fibril core. Therefore the formation of the Wide Fibril is mutually exclusive with our wild-type rod polymorph. b) Environmental distances of putative water molecule for Protofilament A and B in Wide Fibril and b) Protofilament A in Narrow Fibril.



Extended Data Fig. 5 | PreNAC homozipper Island 1 model and additional solvation energy maps. a) Speculative model of preNAC residues ⁵⁰QGVATVA⁵⁶ occupying Island 1 in Protofilament A. b) Atomic solvation map and energetic calculations for Protofilament A with Island 1 as ⁵⁰QGVATVA⁵⁶ and Island 2 as ²⁶VAEAAAG³¹. c) Atomic energy solvation map for Wild-type rod polymorph (6cu7). Notice that K58 and T59 can have favorable stabilization energies whether they are facing the solvent or facing the cavity in the β -arch.

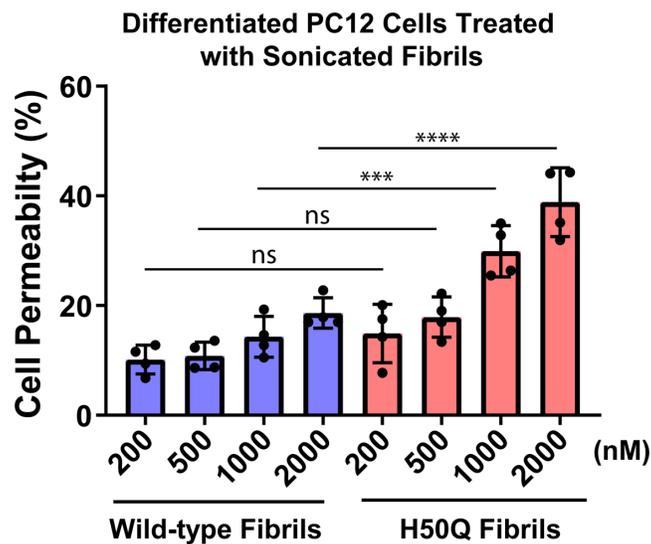


Extended Data Fig. 6 | Comparison of α -syn protofilament interfaces. a) Wide Fibril overview (left), ⁵⁶AEKTKEQV⁶³ homointerface with Wide Fibril electron density (middle), ⁵⁶AEKTKEQV⁶³ homointerface showing a 2.4 Å rise between mated strands from Protofilament A and Protofilament B and a distance of 7.8 Å between mated sheets of Protofilament A and B (right). b) Van der Waal's surface, buried surface area, and shape complementarity of ⁵⁸KTKE⁶¹ homointerface, preNAC interface, and NACore interface.

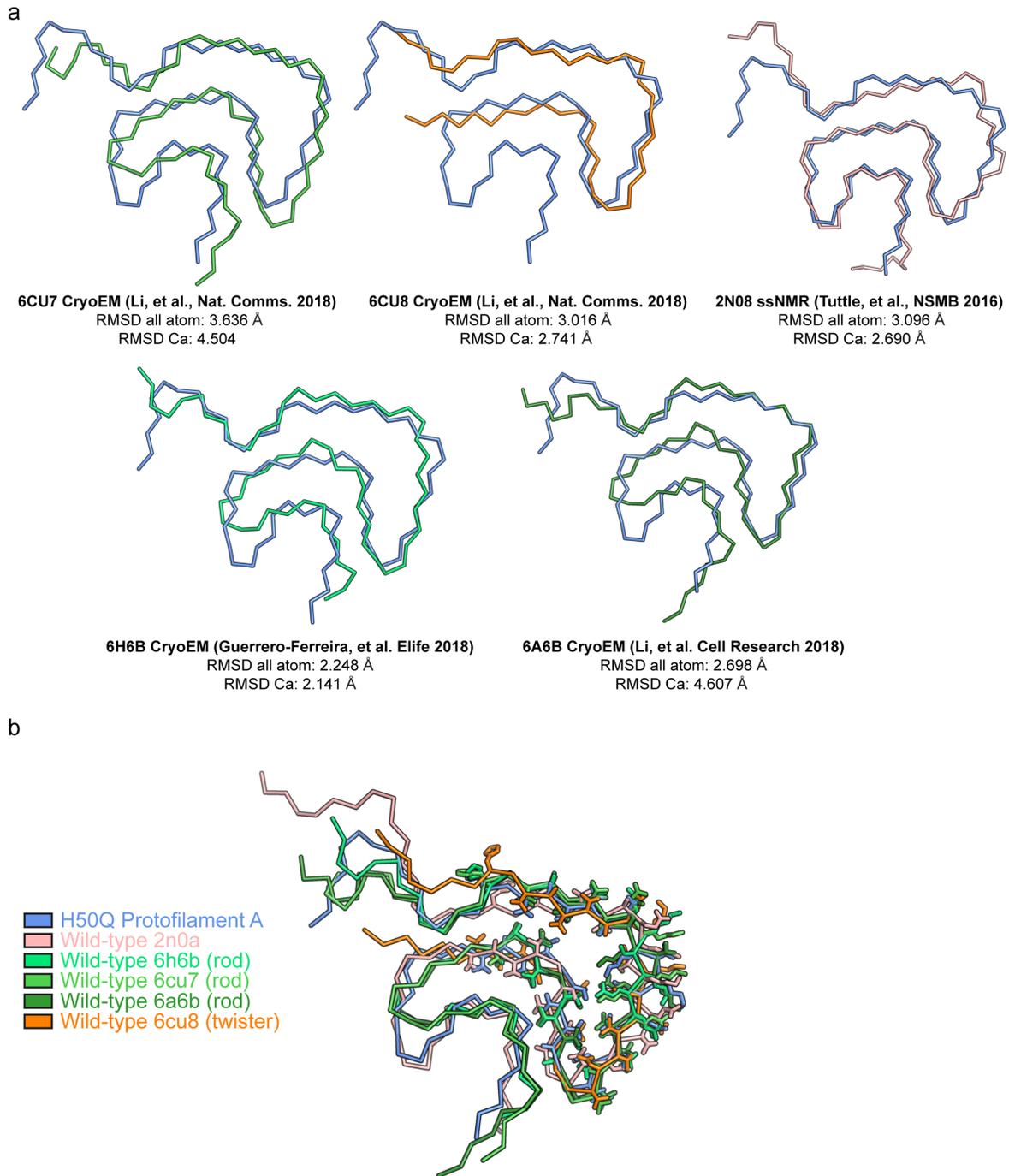


Extended Data Fig. 7 | see figure caption on next page.

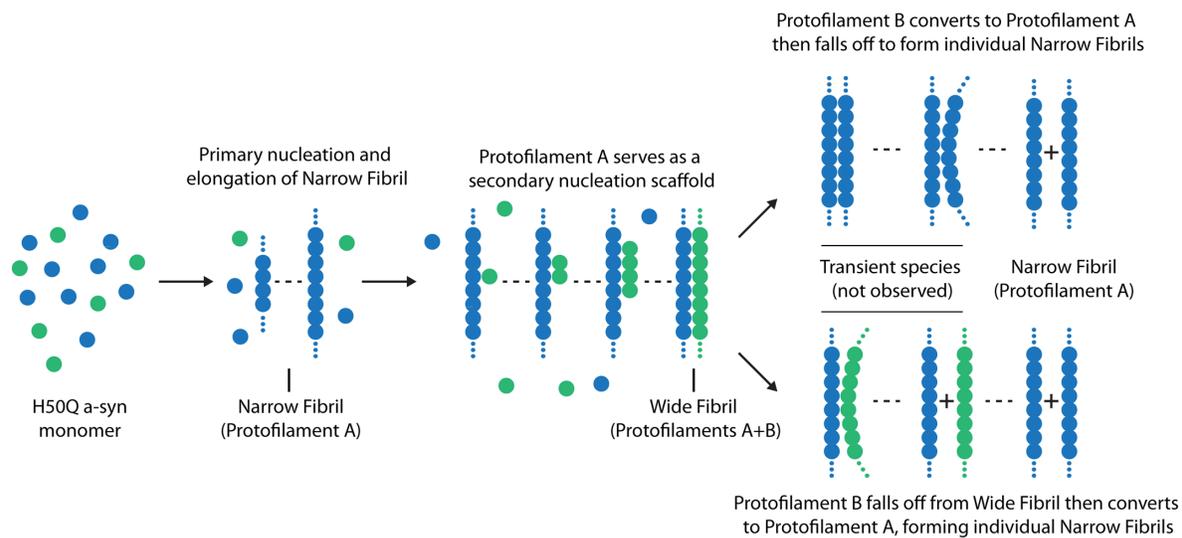
Extended Data Fig. 7 | H50Q disrupts the wild-type rod polymorph preNAC protofilament interface. a) Conformation of H50Q Protofilament A K45 and H50Q. b) Interaction of K45-H50-E57 in the wild-type rod polymorph protofilament interface. c) Hypothetical H50Q double protofilament using the preNAC of Protofilament A as a steric zipper interface. Notice that the H50Q mutation disfavors the protofilament interface due to steric clashes with E57. d) Hypothetical H50Q protofilament interface using preNAC of Protofilament B. Notice the steric clashes between H50Q and E57 at the hypothetical protofilament interface as well as clashes of other parts of the protofilament with Protofilament A.



Extended Data Fig. 8 | H50Q fibrils disrupt PC12 cell membranes more than WT fibrils. Differentiated PC12 cells were treated with sonicated WT and H50Q fibrils and cell permeability was measured via LDH activity in the media (see Methods). H50Q leads to significantly higher cell permeabilization at 1000 and 2000 nM than WT a-syn. Error bars represent standard deviation of four independent experiments. **** = p -value ≤ 0.0001 . *** = p -value ≤ 0.001 . ns = p -value > 0.05 . P -values were calculated using an unpaired, two-tailed t -test with a 95% CI.



Extended Data Fig. 9 | Structural alignment of different wild-type and mutant α -syn polymorphs. a) Structural alignment of H50Q Protofilament A with all wild-type structures determined thus far. b) Structural alignment of residues 50-57 in wild-type and mutant α -syn polymorphs reveals the kernel region is largely conserved while tail regions, especially the N-terminus, adopt variable conformations.



Extended Data Fig. 10 | Schematic illustrating possible secondary nucleation of Protofilament B by Narrow Fibrils.



The α -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure

David R. Boyer^{a,b,c,d,e,1} , Binsun Li^{d,f,1}, Chuanqi Sun^{d,f}, Weijia Fan^{d,f}, Kang Zhou^g, Michael P. Hughes^{a,b,c,d,e,2}, Michael R. Sawaya^{a,b,c,d,e}, Lin Jiang^{d,f,3}, and David S. Eisenberg^{a,b,c,d,e,3}

^aDepartment of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095; ^bDepartment of Biological Chemistry, University of California, Los Angeles, CA 90095; ^cDepartment of Energy Institute, University of California, Los Angeles, CA 90095; ^dMolecular Biology Institute, University of California, Los Angeles, CA 90095; ^eHoward Hughes Medical Institute, University of California, Los Angeles, CA 90095; ^fDepartment of Neurology, David Geffen School of Medicine, University of California, Los Angeles, CA 90095; and ^gCalifornia NanoSystems Institute, University of California, Los Angeles, CA 90095

Edited by William F. DeGrado, University of California, San Francisco, CA, and approved January 6, 2020 (received for review October 12, 2019)

Aggregation of α -synuclein is a defining molecular feature of Parkinson's disease, Lewy body dementia, and multiple systems atrophy. Hereditary mutations in α -synuclein are linked to both Parkinson's disease and Lewy body dementia; in particular, patients bearing the E46K disease mutation manifest a clinical picture of parkinsonism and Lewy body dementia, and E46K creates more pathogenic fibrils in vitro. Understanding the effect of these hereditary mutations on α -synuclein fibril structure is fundamental to α -synuclein biology. We therefore determined the cryo-electron microscopy (cryo-EM) structure of α -synuclein fibrils containing the hereditary E46K mutation. The 2.5-Å structure reveals a symmetric double protofilament in which the molecules adopt a vastly rearranged, lower energy fold compared to wild-type fibrils. We propose that the E46K misfolding pathway avoids electrostatic repulsion between K46 and K80, a residue pair which form the E46-K80 salt bridge in the wild-type fibril structure. We hypothesize that, under our conditions, the wild-type fold does not reach this deeper energy well of the E46K fold because the E46-K80 salt bridge diverts α -synuclein into a kinetic trap—a shallower, more accessible energy minimum. The E46K mutation apparently unlocks a more stable and pathogenic fibril structure.

α -synuclein | Parkinson's disease | Lewy body dementia | cryo-EM | hereditary mutations

The group of diseases termed the synucleinopathies—Parkinson's disease (PD), Lewy body dementia (LBD), and multiple systems atrophy (MSA)—are thought to be caused by the aggregation of α -synuclein (α -syn) into amyloid fibrils. The causal relationship between the formation of amyloid fibrils of α -syn and the synucleinopathies is supported by several observations. Aggregated α -syn is a major component of Lewy bodies, the hallmark lesion in PD and LBD, and the hallmark lesions of MSA (1, 2). Hereditary mutations in α -syn are linked to familial forms of PD and LBD (3). Overexpression of wild-type α -syn via dominantly inherited duplications and triplications of the gene that encodes α -syn, SNCA, are sufficient to cause PD (4–6). Further, the injection of fibrils of α -syn into the brains of mice induced PD-like pathology including Lewy body and Lewy neurite formation, cell-to-cell spreading of Lewy body pathology, and motor deficits similar to PD (7). Although it is never fully possible to establish causation, these combined observations suggest the case is strong for the linkage of aggregated α -syn to the synucleinopathies.

To gain a molecular level understanding of amyloid fibrils of α -syn, we previously applied cryo-electron microscopy (cryo-EM) to determine the near-atomic structures of fibrils of recombinantly assembled α -syn (9). We observed two distinct structures—termed the rod and the twister—that share a similar structural kernel formed by residues 50 to 77 but differ in their protofilament interfaces and flanking regions of the fibril core (Fig. 1B and *SI Appendix, Fig. S1 A and B*; also see Fig. 4B and C). The rod and twister structures display amyloid polymorphism—the formation of distinct fibril conformations by the same protein sequence.

Amyloid polymorphism has been observed in numerous other amyloid-forming proteins, including the tau protein and amyloid- β (7, 10–12). For tau protein, the different polymorphs observed correspond to distinct diseases, namely Alzheimer's, Pick's disease, and chronic traumatic encephalopathy (7, 10, 11). Although the structures of α -syn fibrils derived from human disease brain tissue have not yet been determined, previous research has shown that fibrils derived from PD have different biochemical properties than fibrils from MSA, including differential seeding activity and cell-type origin and infectivity, that may result from the formation of distinct polymorphs in the two diseases (13). Therefore, atomic structures of fibril polymorphs are key to both basic understanding of amyloid protein structure and the development of disease-specific therapeutics.

Although hereditary mutations offer a crucial link between α -syn and disease, it is unknown how they might exert their effects. One hypothesis is that hereditary mutations may encode new structures of α -syn with enhanced pathogenicity. We note that new structures formed as the result of hereditary mutations are not true polymorphs—different structures adopted by the same protein

Significance

Parkinson's is the second most prevalent neurodegenerative condition, leading to movement disorders, and dementia in some cases. Because of the strong association of this condition with amyloid aggregates of the protein α -synuclein, structural understanding of these amyloid aggregates may be the path to eventual therapies. Our study of the structure of a variant α -synuclein inherited in families afflicted with a clinical picture of parkinsonism and Lewy body dementia supplements recent structures of the wild-type structure and shows how a single residue change can result in a greatly changed structure that may underlie the inherited form of the disease.

Author contributions: D.R.B., B.L., L.J., and D.S.E. designed research; D.R.B., B.L., W.F., and K.Z. performed research; D.R.B., B.L., C.S., M.P.H., and M.R.S. contributed new reagents/analytic tools; D.R.B., B.L., L.J., and D.S.E. analyzed data; and D.R.B., B.L., L.J., and D.S.E. wrote the paper.

Competing interest statement: D.S.E. is an advisor and equity shareholder in ADRx, Inc.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: The cryo-EM map and model have been deposited in the Protein Data Bank, <https://www.rcsb.org/> (PDB ID code 6UFR) and the Electron Microscopy Data Bank, <https://www.ebi.ac.uk/pdbe/emdb> (accession no. EMD-20759).

¹D.R.B. and B.L. contributed equally to this work.

²Present address: Department of Cell & Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN 38105.

³To whom correspondence may be addressed. Email: jianglin@ucla.edu or david@mbi.ucla.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1917914117/-DCSupplemental>.

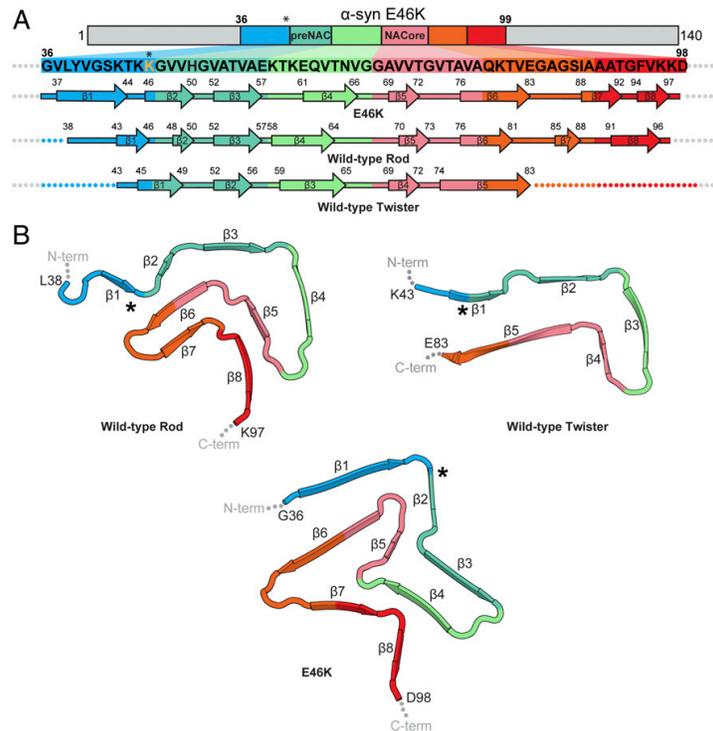


Fig. 1. The E46K mutation leads to a repacked protofilament fold. (A) Primary and secondary structure of α -syn wild-type and E46K fibrils. PreNAC comprises residues 47–56; NACore comprises residues 68–78 (8). Arrows indicate regions adopting β -strand conformation while other residues form loops and turns. The E46K mutation lies N-terminal to the preNAC region. (B) Protofilament folds of the previously determined wild-type folds (rod and twister) and E46K protofilament fold determined here (9). Asterisks indicate the location of residue 46.

sequence. Instead, we label them as “quasi-polymorphs”—different structures adopted by a protein pair with almost identical sequences.

Several hereditary mutations in α -syn may result in the formation of quasi-polymorphs. Mutations A30P, E46K, H50Q, G51D, A53E, and A53T have all been discovered to result in autosomal dominant synucleinopathies (3). Of these, E46K seems to be the only hereditary mutation that manifests in a clinical picture closer to LDB whereas other mutations are found in the context of PD, suggesting that E46K may have a unique effect on the structure of amyloid fibrils of α -syn (14). Indeed, solid state NMR studies of α -syn E46K show large chemical shift differences relative to wild-type fibrils, suggesting large scale rearrangements in the fibril structure as a result of the E46K mutation (15).

Consistent with the evidence that E46K alters α -syn fibril structure and disease manifestation, E46K has been shown to increase the pathogenicity of α -syn fibrils compared to wild-type. In vitro studies have shown that E46K results in an increase in α -syn’s phospholipid binding ability and an enhancement of fibril formation (16). In addition, E46K promotes higher levels of aggregation in cultured cells relative to wild-type, A53T, and A30P α -syn (17). Further, others have found that α -syn bearing E46K is more toxic to rat primary neurons compared to wild-type, A30P, and A53T (18).

Three previous structures of full-length wild-type non-acetylated as well as three structures of acetylated full-length and C-terminally truncated α -syn reveal that E46 participates in a conserved salt bridge with K80 (*SI Appendix, Fig. S1A*) (9, 19–22).

The E46K mutation must eliminate this salt bridge due to electrostatic repulsion that disfavors the proximity of K46 to K80, potentially leading to a different fibril structure (quasi-polymorph). A structural difference such as this may help explain the altered biochemical properties and solid state NMR (ssNMR) chemical shifts of E46K α -syn fibrils. Therefore, we sought to determine the structure of α -syn fibrils containing the E46K hereditary mutation. We find that the E46K mutation produces a homogenous sample composed of a single species whose structure differs radically from structures determined thus far. Consistent with prior studies demonstrating increased pathogenicity of E46K α -syn, we also find that E46K fibrils are more powerful seeds in α -syn HEK293T biosensor cells and more strongly impair mitochondrial activity in PC12 cells. Combining structural, energetic, and biochemical analysis, we attempt to understand the structure–function relationship of α -syn fibrils.

Results

Cryo-EM Structure Determination and Architecture of E46K α -Syn Fibrils

We purified α -syn bearing the E46K hereditary mutation and generated amyloid fibrils using the same growth conditions as in our previous studies (9). We then optimized cryo-EM grids of E46K fibrils and imaged them at 165,000 \times magnification on an energy-filtered Titan Krios equipped with a K2 Summit direct electron detector operating in superresolution mode. We also took advantage of image shift induced-beam tilt correction in SerialEM to maintain coma-free alignment during data collection (23). Using helical refinement procedures in Relion 3.0, we obtained a 2.5-Å resolution reconstruction of

E46K α -syn fibrils (Fig. 2 A–E, *SI Appendix*, Figs. S2 A–C and S3, and Table 1).

The high resolution allowed us to confidently assign side chain rotamer and carbonyl positions, confirming the left handedness and the internal C_2 symmetry of the helical reconstruction, and to build an unambiguous atomic model de novo (Fig. 2 B–E, *SI Appendix*, Fig. S3, and *Materials and Methods*). We were also able to visualize water molecules hydrogen bonded to polar and charged side chains such as threonine, serine, and lysine, as well as to backbone carbonyl oxygen atoms (Fig. 2 D and E and *SI Appendix*, Fig. S3). The atomic model reveals that the E46K fibrils are wound from two protofilaments formed by ordered residues Gly36 to Asp98 that come together at a “wet” two-fold symmetric interface formed by residues K45-E57 (Figs. 2 B and C and 3A).

Comparison of E46K and Wild-Type α -Syn Protofilament Folds.

Overall, the E46K mutation leads to a different protofilament fold than the previously observed wild-type rod and twister structures, resulting in a quasi-polymorph of α -syn. Fig. 1 A and B compares the secondary and tertiary structure of the E46K fibril with the wild-type rod and twister. Both the rod and twister form a similar structural kernel comprised of residues 50 to 77, in which β -strands β_3 , β_4 , and β_5 in the rod and β_2 , β_3 , and β_4 in the twister form a bent β -arch (Fig. 2). The twister has fewer ordered residues at the N and C termini of its fibril core compared to the rod (Fig. 1 A and B and *SI Appendix*, Fig. S1). The C-terminal residues in the rod structure form a Greek key-like fold comprising β_5 , β_6 , β_7 , and β_8 . However, in the E46K structure, the kernel and Greek key are not maintained, and a different packing arrangement is formed. This is likely due to the E46K mutation disrupting the wild-type E46-K80 salt bridge, allowing a rearrangement of the

backbone (Fig. 2B). Interestingly, K80 is now buried in the fibril core although we visualize an ordered molecule binding to the primary amine of the K80 side chain, suggesting that an extensive hydrogen bond network may help to counteract the apparent unfavorable placement of this polar, charged residue within the fibril core. Without the constraint of the E46-K80 salt bridge, a different set of residues form sheet-sheet interfaces within the protofilament (Fig. 2). Now tightly mated heterozippers are formed by β -strands β_1 and β_6 , β_3 and β_4 , and a roughly triangular-shaped bent β -arch fold is formed by β_4 , β_5 , β_6 , and β_7 (Fig. 1).

SI Appendix, Fig. S4 A–C helps to visualize the changes in interacting residues among the structures by displaying all pairwise interactions found within the protofilaments of wild-type and E46K quasi-polymorphs. This analysis of pairwise interactions not only codifies the differences in sets of interacting residues between the wild-type and E46K structures but also clearly reveals that the E46K structure has more interacting residues than its wild-type counterparts. Correspondingly, each chain within the E46K protofilament has a greater buried surface area than wild-type structures (7,944 \AA^2 for E46K; 7,605 \AA^2 for wild-type rod; 5,082 \AA^2 for wild-type twister). In line with this observation, energetic analysis indicates that the E46K structure has a lower standard free energy (greater stabilization) than wild-type structures (*SI Appendix*, Fig. S1 A–C and Table 2). To verify the energy estimate, we performed a sodium dodecyl sulfate (SDS) denaturation assay where both wild-type and E46K fibrils were incubated with various concentrations of SDS at 37 °C followed by thioflavin-T (ThT) fluorescence measurements. Our results demonstrate that E46K fibrils are more resistant to chemical denaturation than wild-type fibrils, consistent with our energetic calculation suggesting that E46K fibrils are more stable than wild-type fibrils (*SI Appendix*, Fig. S5B).

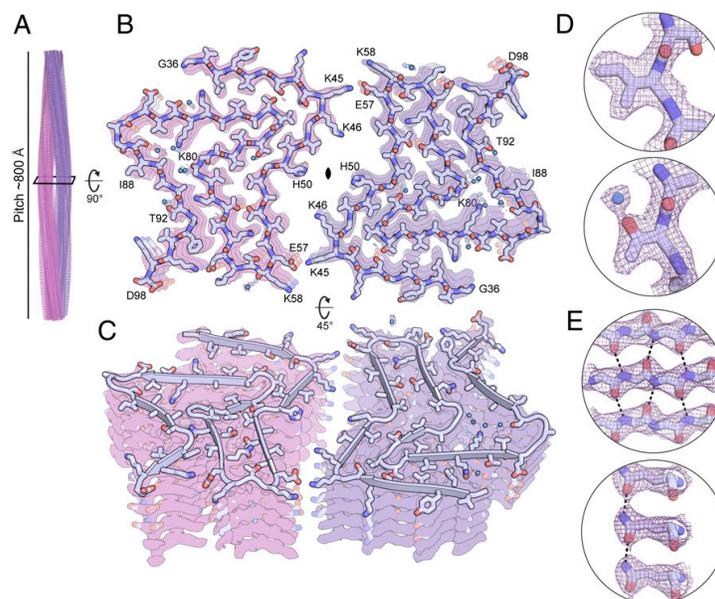


Fig. 2. Overview of E46K cryo-EM structure. (A) E46K fibril side view demonstrating a pitch of ~ 800 Å. (B) Cross-sectional view of one layer of the E46K fibril. The fibril is wound from two identical protofilaments related by a two-fold rotational symmetry axis. The E46K protofilament contains ordered residues 36 to 98. (C) Tilted view of E46K fibril cross-section demonstrating the stacking of identical layers. (D and E) Representative side chain and backbone densities highlight resolution of reconstruction and good map-to-model agreement. (D) Coulombic potential map of Ile88 and Thr92. An ordered water molecule is bound to the oxygen of the threonine side chain. (E) Coulombic potential map of representative backbone (Gln79-Val82) and side chain amide hydrogen bonding (Gln62).

Table 1. Cryo-EM data collection, refinement, and validation statistics

Name	E46K α -syn
PDB ID	6UFR
EMDB ID	EMD-20759
Data collection	
Magnification	$\times 165,000$
Defocus range, μm	0.75–3.37
Voltage, kV	300
Camera	K2 summit (quantum LS)
Frame exposure time, s	0.2
No. of movie frames	30
Total electron dose, $e^{-}/\text{\AA}^2$	36
Pixel size, \AA	0.838
Reconstruction	
Box size, pixels	400
Interbox distance, \AA	33.5
No. of micrographs collected	3,078
No. of segments extracted	210,593
No. of segments after Class2D	N/A
No. of segments after Class3D	114,260
Resolution, \AA	2.5
Map sharpening B-factor, \AA^2	–140
Helical rise, \AA	4.85
Helical twist, $^{\circ}$	178.92
Point group	C_2
Atomic model	
No. of nonhydrogen atoms	4330
No. of protein residues	630
rmsd bonds, \AA	0.012
rmsd angles, $^{\circ}$	0.958
Molprobrity clashscore, all atoms	4.36
Molprobrity score	1.64
Poor rotamers, %	0
Ramachandran outliers, %	0
Ramachandran allowed, %	6.6
Ramachandran favored, %	93.4
$C\beta$ deviations $> 0.25 \text{\AA}$, %	0
Bad bonds, %	0
Bad angles, %	0

Information for number of segments after Class2D is not available as only Class3D was performed for 400 pixel box segments (*Materials and Methods*)

A key difference between E46K and wild-type structures is their pattern of electrostatic interactions. The E46K fibril core has a greater number of charged pairs (interacting pairs of glutamate and lysine), creating a more balanced set of electrostatic interactions despite having a higher net charge than the wild-type fibrils (Fig. 3 A–C). This results from the E46K structure containing four electrostatic triads per layer featuring residues K45-E57-K58 and K60-E61-K96 while the wild-type rod and twister structures both contain only two electrostatic zippers featuring residues E46-K80 and E57-K58, respectively (Fig. 3 A–C). Therefore, in line with the observation above that the E46K structure has overall more interresidue contacts, a higher buried surface area, and a lower free energy than wild-type, it also has a richer set of electrostatic interactions.

Interestingly, although the wild-type and E46K structures differ, many residues adopt similar secondary structures. For instance, β -strand interrupting loops and turns are formed by KTKK pseudorepeats (residues 43 to 46, 58 to 61, and 80 to 83) and glycines (residues 51, 67 to 68, 73, 84, and 93) while other residues form β -strands (Fig. 1 A and B). Indeed, if one compares the structures in Fig. 1B, especially the wild-type rod and E46K, loops and turns between β -strands often lead to a similar

change in chain direction. For example, turns between $\beta 3$ and $\beta 4$ (right turn), $\beta 4$ and $\beta 5$ (right turn), $\beta 5$ and $\beta 6$ (left turn), $\beta 6$ and $\beta 7$ (left turn), and $\beta 7$ and $\beta 8$ (right turn) change the chain in generally similar manners. However, there are differences in the extent and radii of the turns that generate the structural diversity seen in the structures. For instance, the turn between $\beta 3$ and $\beta 4$ leads to an $\sim 90^{\circ}$ turn in the rod and an $\sim 180^{\circ}$ turn in the E46K structure. Also, the turn between $\beta 4$ and $\beta 5$ leads to an $\sim 180^{\circ}$ turn in the rod and an $\sim 100^{\circ}$ turn in the E46K structure. These observations highlight the critical importance of turn and loop regions in generating amyloid polymorphism.

Electrostatic Zippers Constitute the E46K Protofilament Interface.

Despite having a tighter protofilament fold that buries more surface area, the interface between E46K protofilaments contains fewer contacts than those in the wild-type rod and twister structures (Fig. 4 A–C). Instead of the two protofilaments meeting at a classical steric zipper interface in which beta-sheets from each protofilament tightly mate with interdigitating side chains excluding water, the E46K structure forms a largely solvent-filled interface spanning residues 45 to 57 (buried surface area = 47.3\AA^2). This is in contrast to the dry, steric zipper-like interfaces formed by the preNAC (residues 47–56, buried surface area = 91.7\AA^2) and NAC (residues 68–78, buried surface area = 65.3\AA^2) residues in the wild-type rod and twister, respectively. Although most residues are too far apart to interact in the E46K interface, two electrostatic zippers form on either side of the interface (Fig. 4A). Electrostatic zippers have previously been observed to enable counterion-induced DNA condensation whereby anionic phosphates and cations such as Mn^{2+} and spermidines alternate along the length of the DNA–DNA interface forming a “zipper” that “fastens” the molecules together (24). In the E46K protofilament interface, electrostatic zippers consist of carboxylate anions of E57 interleaving with the K45 side chain aminium cations, fastening the two protofilaments together (Fig. 4A). This interaction is not only repeated twice per protofilament interface due to the two-fold helical symmetry but extends for thousands of layers along the fibril axis. We note that the staggered nature of the electrostatic zipper leaves unpaired charges on both the top and bottom of the fibril, potentially attracting additional monomers to add to the fibril through long-range electrostatic interactions. We also observe that no single protofilaments were detected during class averaging, indicating that, despite its relative lack of interactions across the protofilament, the E46K interface is apparently strong enough to consistently bind together the two protofilaments of the fibril.

E46K Fibrils Are More Pathogenic than Wild-Type. We wondered if the differences in structure between the E46K and wild-type fibrils resulted in differences in biochemical properties. Therefore, we first examined the ability of sonicated E46K and wild-type fibrils to seed endogenously expressed α -syn-A53T-YFP in HEK293T biosensor cells (25). If seeding occurs, normally diffuse α -syn-A53T-YFP will aggregate into discrete puncta, which can be counted and used as a robust measure of seeding. Our results indicate that, at most concentrations tested, E46K fibrils are significantly more powerful seeds than wild-type fibrils (Fig. 5A). Whereas the biosensor cell assay supports that E46K fibrils are better seeds than wild-type fibrils, the biosensor cells express α -syn-A53T-YFP, and not wild-type α -syn. In addition, differences in lipofectamine uptake may influence which fibril species is able to induce more seeding in the biosensor cell assay. Therefore, in order to directly examine whether E46K fibrils can seed wild-type protein more strongly than wild-type fibrils, we next tested the seeding ability of E46K fibrils in an *in vitro* assay (*SI Appendix, Fig. S5A*). Similar to the biosensor cell assay, E46K fibrils were more efficient than wild-type fibrils in seeding growth of wild-type α -syn (*SI Appendix, Fig. S5A*). Further, the different

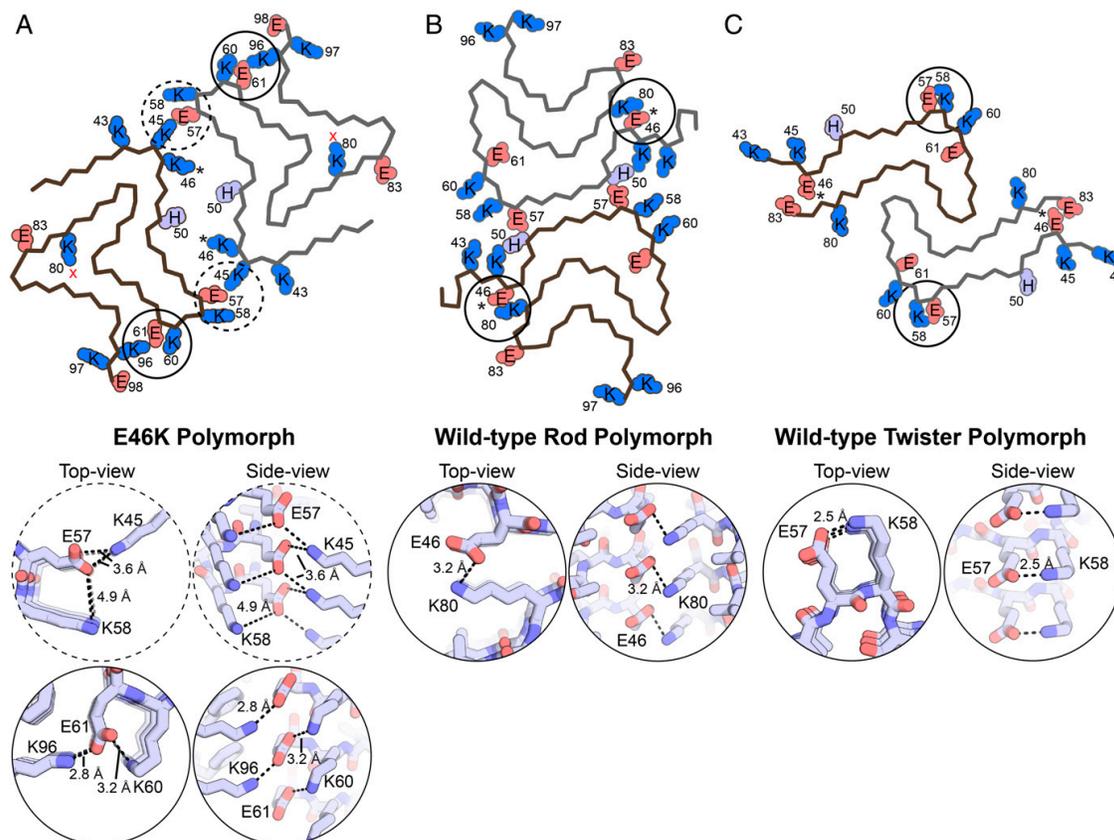


Fig. 3. Electrostatic residues and interactions of wild-type and E46K fibrils. (A) Overview of E46K quasi-polymorph fold with charged and ionizable residues shown (Top). Top and side view of the K45-E57-K58 and K60-E61-K96 electrostatic triads (Bottom). The E46K fibril core has a net charge of +11 per layer (2×8 lysine per chain + $2 \times 1/2$ histidine per chain - 2×3 glutamate per chain) and eight charged pairs per layer ($2 \times$ K45-E57 + $2 \times$ E57-E58 + $2 \times$ K60-E61 + $2 \times$ E61-K96). (B) Overview of the wild-type rod polymorph fold with charged and ionizable residues shown (Top). Top and side view of the E46-K80 salt bridge (Bottom). The wild-type rod fibril core has a net charge of +7 per layer (2×7 lysine per chain + $2 \times 1/2$ histidine per chain - 2×4 glutamate per chain) and two charge pairs per layer ($2 \times$ E46-K80 per chain). (C) Overview of wild-type twister polymorph fold with charged and ionizable residues shown (Top). Top and side view of the E57-K80 salt bridge (Bottom). The wild-type twister fibril core has a net charge of +3 per layer (2×5 lysine per chain + $2 \times 1/2$ histidine per chain - 2×4 glutamate per chain) and two charge pairs per layer ($2 \times$ E57-K58 per chain). Asterisks denote the location of residue 46. A red x denotes the location of the ordered water molecules hydrogen bonding to K80 in the E46K structure.

ThT binding ability of wild-type fibrils either grown without seeds or with wild-type seeds compared to wild-type fibrils seeded by E46K fibrils suggests that seeding by E46K fibrils can induce a different wild-type fibril structure, possibly one similar to the E46K fibril structure (*SI Appendix, Fig. S5A*). This is similar to what is seen in Watson and Lee where both non-acetylated and acetylated α -syn form fibrils to the same extent, yet acetylated α -syn has a substantially lower ThT signal (26). We also observe that unseeded E46K has a longer lag phase than unseeded wild-type α -syn; we speculate why E46K may form fibrils more slowly than wild-type in *Discussion*.

We next compared the ability of E46K and wild-type fibrils to impair mitochondrial activity of differentiated neuron-like rat pheochromocytoma (PC12) cells as a proxy for comparing cytotoxicities (27, 28). We observe significantly impaired mitochondrial activity by lower concentrations of E46K fibrils than of wild-type fibrils (Fig. 5B). Consistent with previous results showing increased aggregation in SH-SY5Y cultured cells, and higher toxicity to rat primary neurons, our results indicate that

the structure formed by E46K fibrils is more pathogenic than wild-type (16–18).

Discussion

Structural Differences Help Explain Enhanced Pathogenicity of E46K α -Syn Fibrils. The structural differences between E46K and wild-type fibrils help rationalize the differences in biochemical properties we observe. First, we have shown an increased seeding efficiency of E46K fibrils compared to wild-type in HEK293T biosensor cells and in vitro (Fig. 5A and *SI Appendix, Fig. S5A*). This difference may arise from the stronger electrostatic templating mechanism of E46K fibrils than wild-type fibrils: There are four electrostatic triads per layer in the E46K fibril structure compared to two E46-K80 salt bridges in the wild-type rod structure and two E57-K80 salt bridges in the wild-type twister structure, and each one of them forms a staggered electrostatic zipper, with overhanging, unsatisfied charges (Fig. 3A and C). Since all of these electrostatic residues are present in the α -syn-A53T-YFP construct in the biosensor cells and in wild-type α -syn

Table 2. Comparative solvation energies

Fibril structure	Atomic solvation standard free energy of stabilization, kcal/mol		
	Per layer	Per chain	Per residue
E46K	-73.8	-36.9	-0.59
WT cryo-EM rod (6CU7)	-59.6	-29.3	-0.49
WT cryo-EM twister (6CU8)	-25.0	-12.5	-0.30
Tau PHF (530L)	-62.8	-31.4	-0.43
Tau pick's (6GX5)	-44.8	-44.8	-0.48
Tau CTE type I (6NWP)	-67.4	-33.7	-0.48
TDP-43 SegA-sym (6N37)	-34.2	-17.1	-0.47
Serum amyloid A (6MST)	-68.8	-34.4	-0.65
FUS (5W3N)	-12.2	-12.2	-0.20

used in the *in vitro* assay, it is plausible that the electrostatic triads in the E46K fibril guide α -syn monomers to the ends of E46K fibrils through long-range electrostatic attraction. Second, the E46K fibril has a different set of ordered surfaces than wild-type fibrils, which may lead to a different set of interacting partners in the cellular milieu. It has previously been shown that aggregates of polyQ can siphon essential proteins into amyloid inclusions and that overexpression of these essential proteins can help alleviate toxicity and reduce aggregate size, presumably by rendering fibrils inert by coating their surface (29). In this way, the different ordered surfaces of E46K fibrils may interact more strongly than wild-type fibrils with certain essential proteins in the cell (for instance, those involved in mitochondrial homeostasis), and this may help to explain the greater reduction in mitochondrial activity we observe (Fig. 5B).

Amyloid Polymorphism. The differences in biological activity associated with structural differences between mutant and wild-type fibrils highlight the relevance of amyloid polymorphism to disease. Hereditary mutations may represent one important influence over the formation of different amyloid polymorphs. Here, we learn that the E46K hereditary mutation leads to a distinct α -syn quasi-polymorph by facilitating a large rearrangement in fibril structure, including a repacked protofilament fold and a protofilament interface. Furthermore, additional recent cryo-EM studies have captured yet another quasi-polymorph of E46K α -synuclein fibrils, which is different still from our own E46K fibril structure (30). The polymorphism displayed by E46K is reminiscent of modal polymorphism, first introduced by Caspar and Cohen (31), whereby identical units have different dispositions in different assemblies (in other words, two fibril structures of the same protein adopt different structures). Although wild-type and E46K fibrils are not strictly modal polymorphs because of the difference in sequence at one amino acid position and hence we refer to them as quasi-polymorphs, it has recently been shown that, under different buffer conditions from our own, the wild-type sequence can indeed form a structure—termed polymorph 2a—similar to the E46K structure determined here (32). This is consistent with the solvent-facing orientation of side chain 46 in the E46K structure, thereby compatible with either the negatively charged wild-type glutamate or the positively charged mutant lysine.

The vast difference in structure between the wild-type and E46K fibrils due to a change at a single amino acid position, or a change in fibril growth buffer conditions (32), highlights the large degree of sensitivity of α -syn fibril assembly to certain interactions. This vulnerability to large-scale rearrangements due to small changes in protein sequence or fibril growth environment reveals that the fibril misfolding landscape for α -syn is flat with many local minima (33). Here, a single amino acid change, the E46K mutation, shifts the fibril misfolding pathway of α -syn to a

different minimum in the folding landscape. Energetic analysis reveals that the E46K quasi-polymorph is significantly more stable than the wild-type folds (-0.59 kcal/mol/residue for E46K vs. -0.49 kcal/mol/residue for wild-type rod and -0.30 kcal/mol/residue for wild-type twister) (Table 1). This calculation, along with the fact that the wild-type sequence has been shown to be able to adopt the E46K structure (32), prompts the question why the wild-type sequence does not always form the more stable E46K structure?

Kinetic Factors That Influence Amyloid Structure. The predominance of one amyloid polymorph over another depends on stochastic nucleation events and kinetically driven growth processes, permitting less stable polymorphs in a sample to dominate if they form and replicate quicker than more stable polymorphs (34). We hypothesize that the formation of the E46-K80 salt bridge observed in all structures of the wild-type rod polymorph determined thus far is an early event in the fibril formation pathway, occurring in the transition state between prefibrillar and fibrillary structures, that lowers the energy barrier to forming the rod structure (9, 19, 20). In other words, the early formation of the E46-K80 salt bridge may divert α -syn into a kinetic trap. By constraining the residues connecting Glu46 and Lys80 to adopt a specific conformation that allows Glu46 and Lys80 to maintain their proximity, the E46-K80 salt bridge shifts the misfolding pathway away from the E46K structure and toward the rod structure, thus enforcing the formation of the less stable structure. This salt bridge could be one contributing factor to the observation of the similar kernels formed by residues 50 to 77 observed in wild-type rod structures (9, 19, 20). The mutation of residue 46 to lysine via the E46K hereditary mutation eliminates the potential to form the E46-K80 salt bridge. This raises the energy barrier to forming the rod and diverts α -syn to a different misfolding pathway, resulting in the more energetically stable E46K structure. These ideas are summarized in Fig. 6, which imagines the misfolding landscape of the α -syn rod and E46K quasi-polymorphs (we label the structure on the right the “compact” polymorph in order to avoid confusion as both E46K and other sequences could adopt this fold). This folding landscape is also consistent with our results that E46K α -syn aggregates more slowly than wild-type (*SI Appendix, Fig. S5A*) given that the transition state for the wild-type sequence to form the rod structure is predicted to be lower than the transition state for the E46K sequence to form the compact polymorph (Fig. 6).

To date, no other studies of wild-type or mutant α -syn have revealed the twister polymorph we determined in our initial cryo-EM study (9). It is especially surprising that E46K α -syn does not form the twister structure given that it was predicted that E46K would result in a favorable interaction with E83 in the twister conformation (9). We speculate that—given the role of kinetics in selecting amyloid polymorphs—the twister polymorph may be the result of a stochastic nucleation event leading to a rare fibril polymorph. Therefore, the twister structure may not be easily reproducible when mutations are added or buffer conditions (influence of fibril growth conditions discussed below) are changed, thereby providing an explanation for why we and others have not recapitulated the twister structure in more recent studies.

Growth Conditions That Influence Amyloid Structure. That the wild-type sequence can also adopt the E46K fold adds complexity to the argument that the formation of the E46-K80 salt bridge is an early event in fibril formation that determines the resulting fibril structure (32). Clearly, there are other factors at play. One such consideration is the role of the fibril growth environment. Guerrero-Ferreira et al. note that the juxtaposition of multiple positively charged lysines at positions 43, 45, and 58 seen in the wild-type rod structures may be allowed only in the presence of poly-anionic counter charges, such as phosphate (32). They

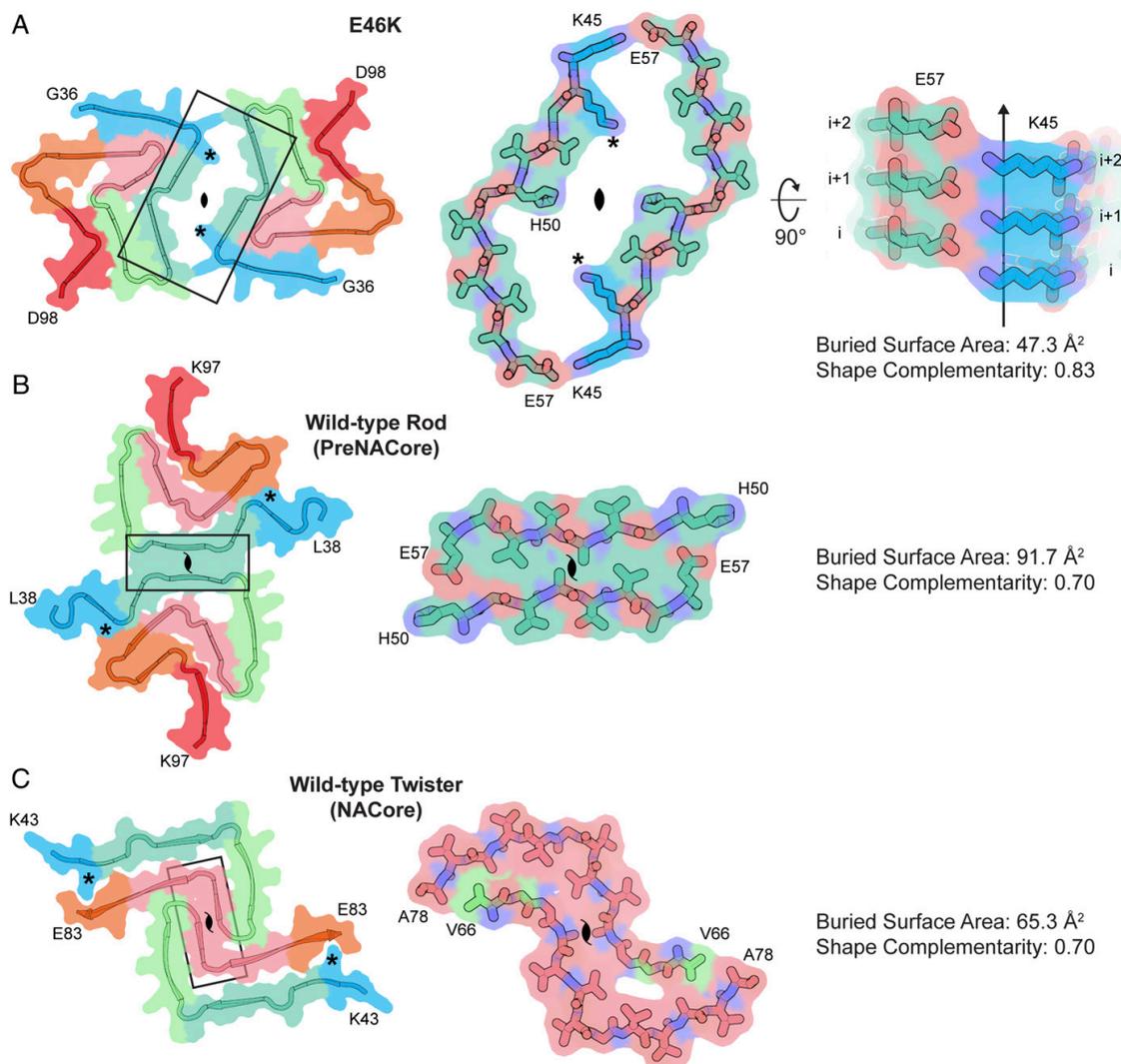


Fig. 4. Comparison of protofilament interfaces of E46K and wild-type fibrils. (A) The E46K quasi-polymorph forms a two-fold symmetric protofilament interface comprised of residues Lys45-Glu57. The interface is largely solvent-filled and is held together by electrostatic interaction of Lys45 and Glu57 at both ends of the interface. The Lys45-Glu57 salt bridge is repeated along the length of the fibril, creating an electrostatic zipper. (B) The wild-type rod polymorph forms a pseudo-2, symmetric dry steric zipper interface comprised of residues His50-E57 from the preNAC region (residues 47–56). (C) The wild-type twister polymorph forms a pseudo-2, symmetric dry steric zipper interface comprised of residues Val66-Ala78 from the NACore (residues 68–78). In all panels, the single asterisk indicates the location of residue 46.

observe that the removal of phosphate from the buffer conditions disallows this juxtaposition of lysine residues, facilitating the formation of the E46K-like polymorph they term 2a. This finding emphasizes again the sensitivity of α -syn fibril structure to perturbations in sequence or growth environment and that both of these factors play a role in producing varied folds. Further, our cross-seeding experiments demonstrating that E46K fibrils can seed wild-type monomer and the fact that the resulting fibrils have different ThT binding ability compared to unseeded wild-type—indicative of a different underlying structure—suggests that the E46-K80 salt bridge can also be disrupted by templated aggregation with seeds of a different structure (*SI*

Appendix, Fig. S5A). Together, these findings indicate that the force exerted by the E46-K80 salt bridge is relatively small and can be overcome by different buffer conditions or templated aggregation. However, these small forces have a large influence in selecting the polymorph because they are exerted early in the aggregation pathway.

Interestingly, Guerrero-Ferreira et al. also reconstruct a low-resolution density map of E46K α -syn grown in phosphate buffer that resembles the 2a polymorph and our E46K structure determined here (32). This is identical to the buffer they used previously to grow fibrils of wild-type, C-terminally truncated α -syn that produced a rod polymorph similar to the one we determined

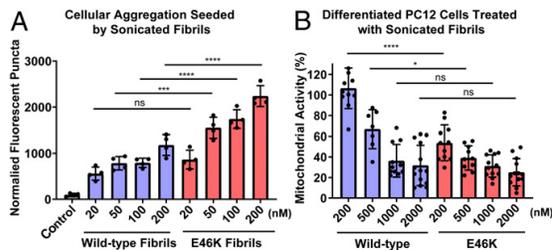


Fig. 5. Biochemical analysis of E46K vs. wild-type fibrils. (A) Fibrils of E46K and wild-type α -syn were sonicated and transfected into HEK293T α -syn-A53T-YFP biosensor cells, and aggregation is measured by the normally soluble and diffuse intracellular α -syn-A53T-YFP forming discrete puncta, which are then quantified through fluorescent image analysis (25). At all concentrations but 20 nM, E46K fibrils are significantly more powerful seeds than wild-type fibrils. Error bars represent the SD of four technical replicates. (B) In order to assay the toxicity of E46K and wild-type fibrils, we treated PC12 cells with sonicated fibrils and measured mitochondrial activity with an MTT assay (27, 28). E46K fibrils require a lower concentration to significantly impair mitochondrial activity compared to wild-type fibrils. Error bars represent the SD of 7 to 14 technical replicates. **** P value ≤ 0.0001 . *** P value ≤ 0.001 . * P value ≤ 0.05 . ns, P value > 0.05 . P values were calculated using an unpaired, two-tailed t test with a 95% CI.

(9, 19). This is similar to the case we describe here in which we grow E46K fibrils using identical buffer conditions as those used to grow fibrils of the wild-type rod polymorph, and the E46K mutation results in a different fold. Taking these results together, this implies that, under identical buffer conditions, E46K acts as a switch to shift the fibril folding pathway, likely through the disruption of the E46-

K80 kinetic trap discussed above, thereby unlocking a more stable polymorph.

Initial studies examining the effect of the E46K mutation on recombinantly assembled α -syn fibrils further emphasize the importance of fibril growth conditions on selecting polymorphs. Whereas we find that E46K α -syn forms fibrils more slowly than wild-type α -syn (*SI Appendix, Fig. S5A*), Choi et al. found that the E46K mutation accelerates α -syn fibril formation (16). However, the fibrils formed in their study appear to have a different morphology than our E46K fibrils (16). Specifically, the E46K fibrils prepared under their conditions have a faster twist and greater variation in width along the fibril, indicative of a different underlying molecular structure. This distinction is most likely due to the differences in buffer conditions between our two preparations: Choi et al. (16) grew their fibrils in Mops buffer whereas we grow our fibrils in water and tetrabutylphosphonium bromide, a fibrillation agent we identified in our initial screening to identify wild-type fibrils suitable for cryo-EM structure determination (9). The observed differences in fibril morphology and aggregation kinetics again highlight the sensitivity of α -syn fibrillation to growth conditions and hints at the ability of E46K α -syn to form still further polymorphs not determined in this work.

The Folding Landscape of Amyloid Proteins. As mentioned above, amyloid polymorphism is likely due to a flat protein folding landscape in which many local minima exist. This landscape contrasts with that of most proteins whose sequences have evolved to encode one structure with the lowest free energy in a funnel-shaped protein folding landscape (33, 35). The prodigious polymorphism observed not only in α -syn structures determined in our work but in other amyloid structures hints that evolution

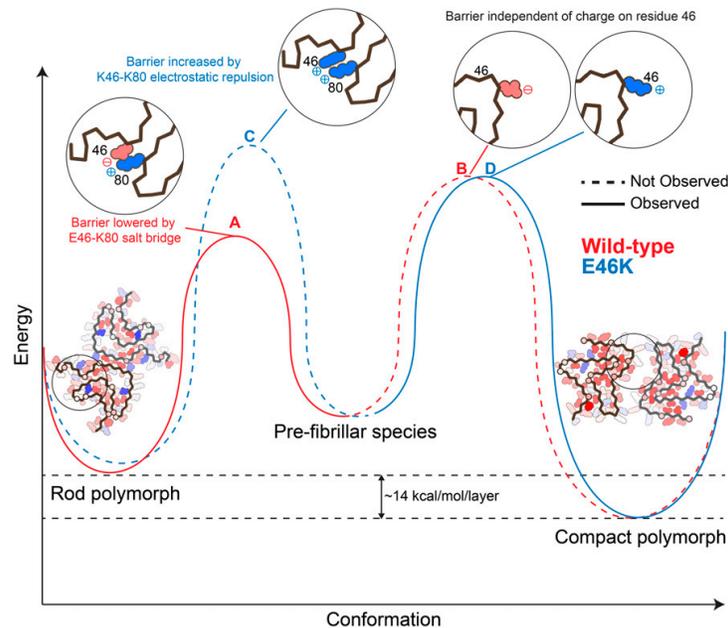


Fig. 6. Proposed misfolding landscape of α -syn rod and compact polymorphs. The E46-K80 salt bridge of the wild-type sequence (A) lowers the transition-state barrier to the rod polymorph. Thus the salt bridge kinetically favors formation of the rod polymorph, rather than the hypothetical, compact wild-type structure (B). That is, the compact polymorph we observe here is not observed for the wild-type sequence because the E46-K80 salt bridge acts as a kinetic trap, diverting the sequence to fold into the rod structure. The E46K mutation thereby serves to raise the transition state energy of forming the rod polymorph via electrostatic repulsion between K46-K80 (C), thus facilitating the formation of the compact polymorph (D).

has not played a role in specifying pathogenic amyloid structures. Therefore, many factors, such as concentration and intrachain interactions, may guide the sequence to forming a nonevolved structure. The E46-K80 salt bridge may represent one of these intrachain interactions that diverts wild-type α -syn into a local energy minimum (i.e., a kinetic trap), and the E46K hereditary mutation unlocks this constraint, allowing a more stable structure to form.

The release of local constraints leading to lower free energy structures could be a mechanism by which other hereditary mutations operate as well. This hypothesis is in line with our previous studies showing that pathogenic mutations and post-translational modifications can lead to more stable amyloid assemblies (36–38). That work was done on shorter peptide segments, however, and may not have captured all of the interactions present in a full-length amyloid fibril. Therefore, future work is needed to compare structures of full-length wild-type and hereditary mutant or posttranslationally modified amyloid fibrils to identify the possible mechanisms by which hereditary mutations and posttranslational modifications (PTMs) can alter amyloid protein pathogenicity. We have previously shown that the H50Q hereditary mutation alters α -syn's protofilament assembly, resulting in more pathogenic fibrils—a first step in understanding the role of hereditary mutations in fibril structure and activity (39). Recently, a pair of studies demonstrated that N-terminally acetylated α -syn has different ThT binding ability and different seeding properties compared to nonacetylated α -syn although the structures of nonacetylated and acetylated α -syn fibril cores remain largely similar (21, 26). Further, ssNMR studies of α -syn hereditary mutants A30P, which lies outside the fibril core of α -syn structures determined to date, and A53T, which lies at the protofilament interface of wild-type rod polymorphs, suggest only local perturbations compared to the wild-type fibril structure (15, 40). Also, hereditary mutations in transthyretin have been discovered that serve to destabilize transthyretin's native fold, thereby promoting fibril formation (41). Together, these data indicate that other modes of mutationally or PTM-encoded pathogenicity may exist, including effects localized to regions outside the fibril core or effects on the monomeric protein structure. Indeed, PTMs such as phosphorylation, oxidative stress, and truncation, have varying effects on α -syn aggregation and toxicity (42).

Conclusion

In summary, we have determined a 2.5-Å resolution reconstruction of recombinantly assembled E46K α -syn fibrils that provides the atomic structure of this hereditary mutation initially discovered in a family with a clinical diagnosis of parkinsonism and Lewy body dementia (14). The fibril structure of E46K α -syn greatly differs from, and has a lower free energy than, wild-type structures, and we attempt to use the structure to rationalize its higher seeding capacity and mitochondrial impairment compared to wild-type. We posit that, due to fibril α -syn's unevolved nature, the E46-K80 salt bridge in wild-type fibrils represents a local constraint that prevents the formation of the lower free energy fibril fold; E46K alleviates this constraint, allowing refolding to a more stable structure. The release of local constraints to allow repacking into more stable, pathogenic fibrils may be a mechanism by which other hereditary mutations operate in α -syn and other amyloid proteins.

Materials and Methods

Protein Purification. Full-length α -syn wild-type and E46K mutant proteins were expressed and purified according to a published protocol (9). Transformed bacteria were induced at an OD₆₀₀ of ~0.6 with 1 mM IPTG for 6 h at 30 °C. The bacteria were then lysed with a probe sonicator for 10 min in an iced water bath. After centrifugation, the soluble fraction was heated in boiling water for 10 min and then titrated with HCl to pH 4.5 to remove the

pellet. After adjusting to neutral pH, the protein was dialyzed overnight against Q Column loading buffer (20 mM Tris-HCl, pH 8.0). On the next day, the protein was loaded onto a HiPrep Q 16/10 column and eluted using elution buffer (20 mM Tris-HCl, 1 M NaCl, pH 8.0). The eluent was concentrated using Amicon Ultra-15 centrifugal filters (10 NMWL; Millipore Sigma) to ~5 mL. The concentrated sample was further purified with size-exclusion chromatography through a HiPrep Sephacryl S-75 HR column in 20 mM Tris, pH 8.0. The purified protein was dialyzed against water, concentrated to 3 mg/mL, and stored at 4 °C. The concentration of the protein was determined using the Pierce BCA Protein Assay Kit (cat. No. 23225; Thermo Fisher Scientific).

Fibril Preparation and Optimization. Both wild-type and E46K fibrils were grown under the same condition: 300 μ M purified monomers, 15 mM tetrabutylphosphonium bromide, shaking at 37 °C for 2 wk.

Fibril Seeding Aggregation in Cells. We performed the biosensor cell seeding assay based on a previously published protocol (25). Briefly, the assay works as follows: Exogenous, unlabeled fibrils are transfected into HEK293T cells expressing α -syn-A53T-YFP. Seeded aggregation of endogenously expressed α -syn-A53T-YFP is monitored by formation of fluorescent puncta. The puncta represent intracellular aggregation of α -syn-A53T-YFP as a result of seeding by exogenous E46K or wild-type fibrils.

Human embryonic kidney FRET Biosensor HEK293T cells expressing full-length α -syn containing the hereditary A53T mutation were grown in DMEM (4 mM L-glutamine and 25 mM D-glucose) supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin. Trypsin-treated HEK293T cells were harvested, seeded on flat 96-well plates at a concentration of 4×10^4 cells per well in 200 μ L of culture medium per well and incubated in 5% CO₂ at 37 °C for 18 h.

α -syn fibrils were prepared by diluting with Opti-MEM (Life Technologies, Carlsbad, CA) and sonicating in a water bath sonicator for 10 min. Fibril concentration was determined as monomer-equivalent concentration. The fibril samples were then mixed with Lipofectamine 2000 (Thermo Fisher Scientific) and incubated for 15 min and then added to the cells. The actual volume of Lipofectamine 2000 was calculated based on the dose of 1 μ L per well. After 48 h of transfection, the cells were trypsinized, transferred to a 96-well round-bottom plate, and resuspended in 200 μ L of chilled flow cytometry buffer (HBSS, 1% FBS, and 1 mM EDTA) containing 2% paraformaldehyde. The plate was sealed with Parafilm and stored at 4 °C for imaging. Fluorescent images were processed in ImageJ to count number of seeded cells.

Mitochondrial Activity Assay. The addition of sonicated fibrils to nerve growth factor-differentiated PC12 cells is a well-established assay to measure cytotoxicity of amyloid fibrils (9, 27, 28, 39). Use of this neuron-like cell line allows us to obtain a biologically relevant assay for cytotoxicity. For our MTT (3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide) mitochondrial activity assay, the protocol was adapted from the Provost and Wallert laboratories and was performed in an identical manner to our previous work (39, 43). Thiazolyl blue tetrazolium bromide for the MTT cell toxicity assay was purchased from Millipore Sigma (M2128-1G; Burlington, MA). PC12 cells were plated in 96-well plates at 10,000 cells per well in DMEM (Dulbecco's modification of Eagle's medium; 5% fetal bovine serum, 5% heat-inactivated horse serum, 1% penicillin/streptomycin, and 150 ng/mL nerve growth factor type-h factor 2.5S [Thermo Fisher Scientific]). The cells were incubated for 2 d in an incubator with 5% CO₂ at 37 °C. The cells were treated with different concentrations of monomer-equivalent α -syn fibrils (200 nM, 500 nM, 1,000 nM, and 2,000 nM). After 18 h of incubation, 20 μ L of 5 mg/mL MTT was added to every well, and the plate was returned to the incubator for 3.5 h. With the presence of MTT, the experiment was conducted in a laminar flow hood with the lights off, and the plate was wrapped in aluminum foil. The media was then removed with an aspirator, and the remaining formazan crystals in each well were dissolved with 100 μ L of 100% DMSO. Absorbance was measured at 570 nm to determine the MTT signal and at 630 nm to determine background. The data were normalized to those from cells treated with 1% SDS to obtain a value of 0%, and to those from cells treated with PBS to obtain a value of 100%.

In Vitro Aggregation and Seeding Assay. α -syn wild-type or E46K monomers (100 μ M) were mixed with 60 μ M thioflavin-T (ThT) and transferred into a 96-well plate. The signal was monitored using the FLUOstar Omega Microplate Reader (37 °C with 600 rpm double orbital shaking, ex. 440 nm, em. 490 nm; BMG Labtech). For the seeding groups, preformed wild-type or E46K fibrils

(10 μ M) after 10 min of water-bath sonication were added to the wild-type α -syn monomers immediately before beginning the aggregation assay.

SDS Stability Assay. The wild-type and E46K aggregated α -syn samples at the end of the ThT assay were treated with addition of 10% SDS to reach SDS final concentration of 0.5%. The ThT signal was measured after 5 min of incubation at 37 °C with 600 rpm double orbital shaking. The addition of SDS and ThT measurement was repeated with increments of 0.5% to a final SDS concentration of 3.5%. The initial ThT signals at 0% SDS were used for normalization.

Cryo-EM Data Collection and Processing. Two microliters of fibril solution was applied to a baked and glow-discharged Quantifoil 1.2/1.3 electron microscope grid and plunge-frozen into liquid ethane using a Vitrobot Mark IV (FEI). Data were collected on a Titan Krios (FEI) microscope equipped with a Gatan Quantum LS/K2 Summit direct electron detection camera (operated with 300-kV acceleration voltage and slit width of 20 eV). Counting mode movies were collected on a Gatan K2 Summit direct electron detector with a nominal physical pixel size of 0.843 Å per pixel with a dose per frame 1.2 e-/Å². A total of 30 frames with a frame rate of 5 Hz were taken for each movie, resulting in a final dose of 36 e-/Å² per image. Automated data collection was driven by the SerialEM automation software package, with image shift induced-beam tilt correction (23).

Micrographs containing crystalline ice were used to estimate the anisotropic magnification distortion using `mag_distortion_estimate` (44). CTF estimation was performed using `CTFFIND 4.1.8` on movie stacks with a grouping of three frames and correction for anisotropic magnification distortion (45). `Unblur` (46) was used to correct beam-induced motion with dose weighting and anisotropic magnification correction, resulting in a physical pixel size of 0.838 Å/pixel.

All particle picking was performed manually using `EMAN2 e2helixboxer.py` (47). All fibril particles were first extracted using 1,024-pixel box sizes and a 10% interbox distance and then subjected to two-dimensional (2D) class averaging. 2D class averages reveal that E46K α -syn forms fibrils of a single morphology with a pitch of \sim 800 Å (*SI Appendix, Fig. S2B*). We next extracted all fibrils with a 686-pixel box size and 10% interbox distance and again performed 2D class averaging. 2D class averaging of 686 pixel boxes resulted in clear separation of beta-strands along the length of the fibril. 2D class averages and their corresponding simulated diffraction patterns together indicate that the helical rise is \sim 4.8 Å with a C_n rotational symmetry due to the presence of a meridional reflection (*SI Appendix, Fig. S2B*). Due to the two-fold mirror symmetry present in the 2D class averages, we reasoned that the fibril had a C₂ rotational symmetry. Using a calculated helical twist of 178.92° (given pitch of 800 Å and rise of 4.8 Å) and C₂ rotational symmetry, we carried out three-dimensional (3D) class averaging with a single class and a featureless cylinder created by `reliion_helical_toolbox` as an initial

model (48). The featureless cylinder was refined to a reasonable model where separation of beta-sheets in the x-y plane of the fibril could be visualized. This model was then used to separate good and bad particles with a 3D class averaging job with three classes. Fibrils contributing to the best class in the previous 3D classification were reextracted with a 400-pixel box size and 10% interbox distance and phase flipped for subsequent classification and high resolution refinement. An additional 3D class averaging job was performed, and a final subset of 114,260 helical segments were selected for gold-standard auto-refinement in RELION. Refinement yielded a final 2.5-Å reconstruction (*SI Appendix, Fig. S2C*). We sharpened the map using `phenix.auto_sharpen` with a sharpening factor of -140 Å² and a resolution cutoff of 2.5 Å (49).

Atomic Model Building. We used `phenix.map_to_model` with an input sequence corresponding to E46K α -syn to build an initial model (49). `Phenix.map_to_model` correctly built a segment of the N terminus of the fibril core, and we manually built the rest of the structure into the density map in COOT (50). We generated a five-layer model to maintain local contacts between chains in the fibril during structure refinement. We performed automated structure refinement using `phenix.real_space_refine` (49). We employed hydrogen bond distance and angle restraints for backbone atoms participating in β -sheets and side chain hydrogen bonds during automated refinements. We performed comprehensive structure validation of all our final models in Phenix.

Energetic Calculation. The standard free energy of stabilization of a given amyloid chain is computed as the difference in atomic solvation energy of a pseudoextended, solvated chain and the folded chain in the center of five layers of the known structure of a protofilament (51). The atomic solvation parameters are those of Eisenberg et al. (52), with additional terms to describe the entropy change of side chains on folding, as calculated by Koehl and Delarue (53), scaled by the percentage of side chain surface area buried.

Data Availability Statement. All structural data have been deposited into the Worldwide Protein Data Bank (wwPDB) and Electron Microscopy Data Bank (EMDB) with the following accession codes: PDB 6UFR and EMD EMD-20759.

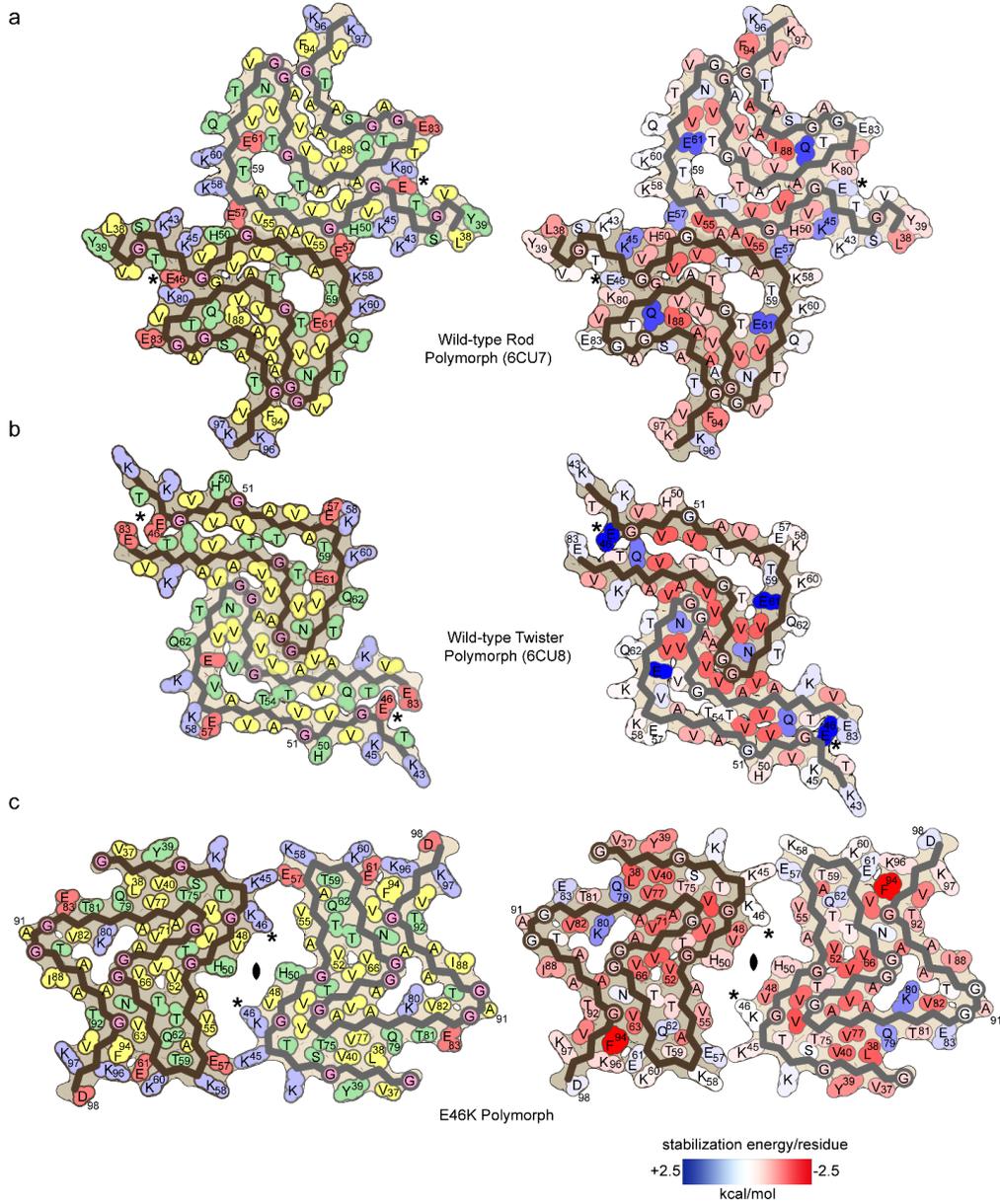
ACKNOWLEDGMENTS. We thank H. Zhou for use of the Electron Imaging Center for Nanomachines (EICN) resources and P. Ge for assistance in cryo-EM data collection. We acknowledge the use of instruments at the EICN supported by NIH Grants 1510RR23057 and 1510OD018111, NSF Grant DBI-1338135, and the California NanoSystems Institute at the University of California, Los Angeles. The authors acknowledge NIH Grant AG 060149, NIH Grant AG 054022, NIH Grant AG061847, and Department of Energy Grant DE-FC02-02ER63421 for support. D.R.B. was supported by the National Science Foundation Graduate Research Fellowship Program.

1. M. G. Spillantini, R. A. Crowther, R. Jakes, M. Hasegawa, M. Goedert, alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with lewy bodies. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6469–6473 (1998).
2. M. G. Spillantini et al., Filamentous α -synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies. *Neurosci. Lett.* **251**, 205–208 (1998).
3. M. G. Spillantini, M. Goedert, Neurodegeneration and the ordered assembly of α -synuclein. *Cell Tissue Res.* **373**, 137–148 (2018).
4. A. B. Singleton et al., α -Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841 (2003).
5. M.-C. Chartier-Harlin et al., Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* **364**, 1167–1169 (2004).
6. P. Ibáñez et al., Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. *Lancet* **364**, 1169–1171 (2004).
7. A. W. P. Fitzpatrick et al., Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185–190 (2017).
8. J. A. Rodriguez et al., Structure of the toxic core of α -synuclein from invisible crystals. *Nature* **525**, 486–490 (2015).
9. B. Li et al., Cryo-EM of full-length α -synuclein reveals fibril polymorphs with a common structural kernel. *Nat. Commun.* **9**, 3609 (2018).
10. B. Falcon et al., Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **561**, 137–140 (2018).
11. B. Falcon et al., Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019).
12. R. Tycko, Molecular structure of aggregated Amyloid- β : Insights from solid-state nuclear magnetic resonance. *Cold Spring Harb. Perspect. Med.* **6**, a024083 (2016).
13. C. Peng et al., Cellular milieu imparts distinct pathological α -synuclein strains in α -synucleinopathies. *Nature* **557**, 558–563 (2018).
14. J. J. Zarranz et al., The new mutation, E46K, of α -synuclein causes Parkinson and Lewy body dementia. *Ann. Neurol.* **55**, 164–173 (2004).
15. L. R. Lemkau et al., Site-specific perturbations of alpha-synuclein fibril structure by the Parkinson's disease associated mutations A53T and E46K. *PLoS One* **8**, e49750 (2013).
16. W. Choi et al., Mutation E46K increases phospholipid binding and assembly into filaments of human α -synuclein. *FEBS Lett.* **576**, 363–368 (2004).
17. N. Pandey, R. E. Schmidt, J. E. Galvin, The alpha-synuclein mutation E46K promotes aggregation in cultured cells. *Exp. Neurol.* **197**, 515–520 (2006).
18. I. Íñigo-Marco et al., E46K α -synuclein pathological mutation causes cell-autonomous toxicity without altering protein turnover or aggregation. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8274–E8283 (2017).
19. R. Guerrero-Ferreira et al., Cryo-EM structure of alpha-synuclein fibrils. *eLife* **7**, e36402 (2018).
20. Y. Li et al., Amyloid fibril structure of α -synuclein determined by cryo-electron microscopy. *Cell Res.* **28**, 897–903 (2018).
21. X. Ni, R. P. McGlinchey, J. Jiang, J. C. Lee, Structural insights into α -synuclein fibril polymorphism: Effects of Parkinson's disease-related C-terminal truncations. *J. Mol. Biol.* **431**, 3913–3919 (2019).
22. M. D. Tuttle et al., Solid-state NMR structure of a pathogenic fibril of full-length human α -synuclein. *Nat. Struct. Mol. Biol.* **23**, 409–415 (2016).
23. D. N. Mastrorade, E. M. Serial, A program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position. *Microsc. Microanal.* **9**, 1182–1183 (2003).
24. A. A. Kornyshev, S. Leikin, Electrostatic zipper motif for DNA aggregation. *Phys. Rev. Lett.* **82**, 4138–4141 (1999).
25. S. B. Prusiner et al., Evidence for α -synuclein prions causing multiple system atrophy in humans with parkinsonism. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5308–E5317 (2015).
26. M. D. Watson, J. C. Lee, N-terminal acetylation affects α -synuclein fibril polymorphism. *Biochemistry* **58**, 3630–3633 (2019).
27. M. Wakabayashi, K. Matsuzaki, Formation of amyloids by beta-(1-42) on NGF-differentiated PC12 cells: Roles of gangliosides and cholesterol. *J. Mol. Biol.* **371**, 924–933 (2007).
28. K. Ono, M. M. Condron, D. B. Teplow, Structure-neurotoxicity relationships of amyloid β -protein oligomers. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 14745–14750 (2009).

29. F. Hosp *et al.*, Spatiotemporal proteomic profiling of Huntington's disease inclusions reveals widespread loss of protein function. *Cell Rep.* **21**, 2291–2303 (2017).
30. K. Zhao *et al.*, Parkinson's disease associated mutation E46K of α -synuclein triggers the formation of a novel fibril structure. *bioRxiv* 870758, 10.1101/870758 (2019).
31. D. L. Caspar, C. Cohen, "Polymorphism of tropomyosin and a view of protein function" in *Nobel Symposium 11. Symmetry and Function of Biological Systems at the Macromolecular Level*, A. Engstrom, B. Strandberg, Eds. (Almqvist and Wiksell; John Wiley and Sons, Inc., 1969), pp. 393–414.
32. R. Guerrero-Ferreira *et al.*, Two new polymorphic structures of human full-length alpha-synuclein fibrils solved by cryo-electron microscopy. *eLife* **8**, e48907 (2019).
33. J. N. Onuchic, P. G. Wolynes, Z. Luthey-Schulten, N. D. Socci, Toward an outline of the topography of a realistic protein-folding funnel. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3626–3630 (1995).
34. R. Riek, D. S. Eisenberg, The activities of amyloids from a structural perspective. *Nature* **539**, 227–235 (2016).
35. C. B. Anfinsen, Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
36. E. L. Guenther *et al.*, Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. *Nat. Struct. Mol. Biol.* **25**, 463–471 (2018).
37. R. A. Warmack *et al.*, Structure of amyloid- β (20-34) with Alzheimer's-associated isomerization at Asp23 reveals a distinct protofilament interface. *Nat. Commun.* **10**, 3357 (2019).
38. Q. Cao, D. R. Boyer, M. R. Sawaya, P. Ge, D. S. Eisenberg, Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat. Struct. Mol. Biol.* **26**, 619–627 (2019).
39. D. R. Boyer *et al.*, Structures of fibrils formed by α -synuclein hereditary disease mutant H50Q reveal new polymorphs. *Nat. Struct. Mol. Biol.* **26**, 1044–1052 (2019).
40. L. R. Lemkau *et al.*, Mutant protein A30P α -synuclein adopts wild-type fibril structure, despite slower fibrillation kinetics. *J. Biol. Chem.* **287**, 11526–11532 (2012).
41. A. R. Hurshman Babbes, E. T. Powers, J. W. Kelly, Quantification of the thermodynamically linked quaternary and tertiary structural stabilities of transthyretin and its disease-associated variants: The relationship between stability and amyloidosis. *Biochemistry* **47**, 6969–6984 (2008).
42. H. A. Lashuel, C. R. Overk, A. Oueslati, E. Masliah, The many faces of α -synuclein: From structure and toxicity to therapeutic target. *Nat. Rev. Neurosci.* **14**, 38–48 (2013).
43. J. van Meerloo, G. J. L. Kaspers, J. Cloos, "Cell sensitivity assays: The MTT assay" in *Cancer Cell Culture: Methods and Protocols*, I. A. Cree, Ed. (Humana Press, 2011), pp. 237–245.
44. T. Grant, N. Grigorieff, Automatic estimation and correction of anisotropic magnification distortion in electron microscopes. *J. Struct. Biol.* **192**, 204–208 (2015).
45. A. Rohou, N. Grigorieff, CTFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
46. T. Grant, N. Grigorieff, Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* **4**, e06980 (2015).
47. G. Tang *et al.*, EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
48. S. He, S. H. W. Scheres, Helical reconstruction in RELION. *J. Struct. Biol.* **198**, 163–176 (2017).
49. P. V. Afonine *et al.*, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
50. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
51. D. Eisenberg, A. D. McLachlan, Solvation energy in protein folding and binding. *Nature* **319**, 199–203 (1986).
52. D. S. Eisenberg, M. Wesson, M. Yamashita, Interpretation of protein folding and binding with atomic solvation parameters. *Chem. Scr.* **29A**, 217–222 (1989).
53. P. Koehl, M. Delarue, Application of a self-consistent mean field theory to predict protein side-chains conformation and estimate their conformational entropy. *J. Mol. Biol.* **239**, 249–275 (1994).

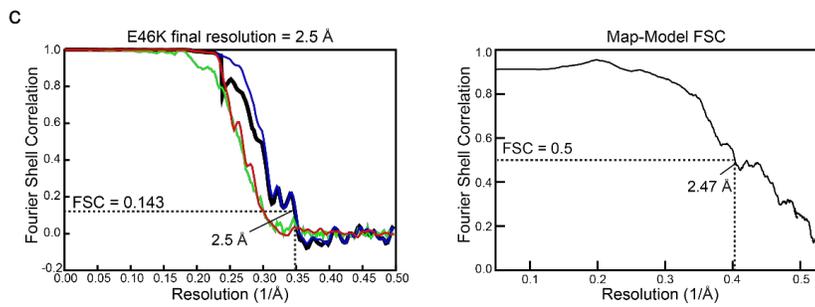
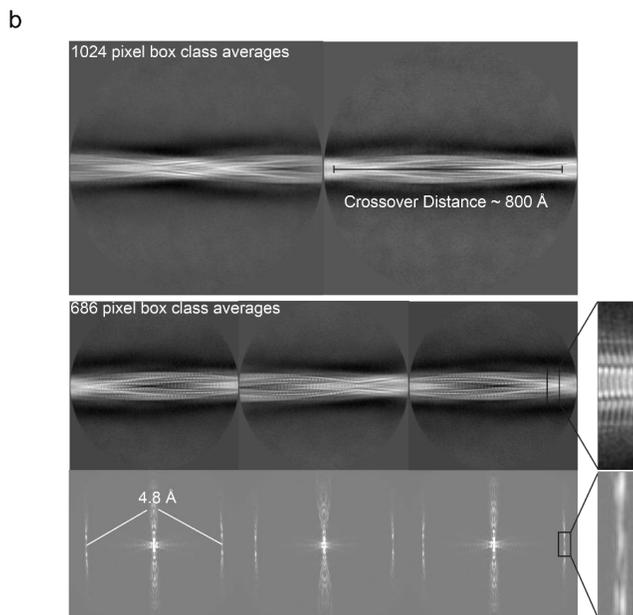
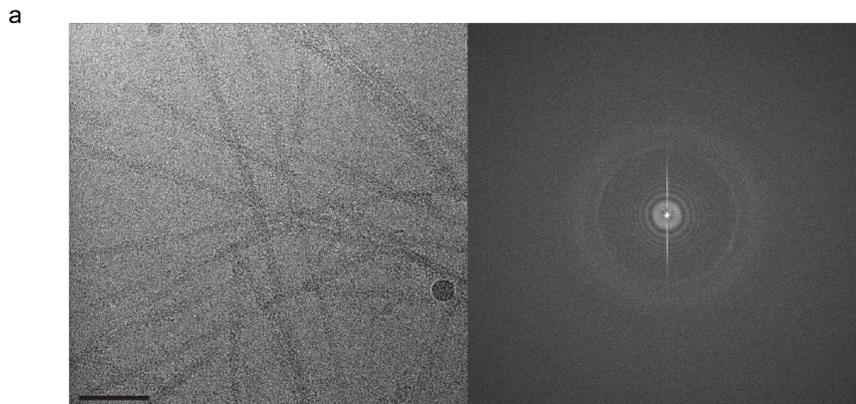
Supporting Information

Supplementary Figures:



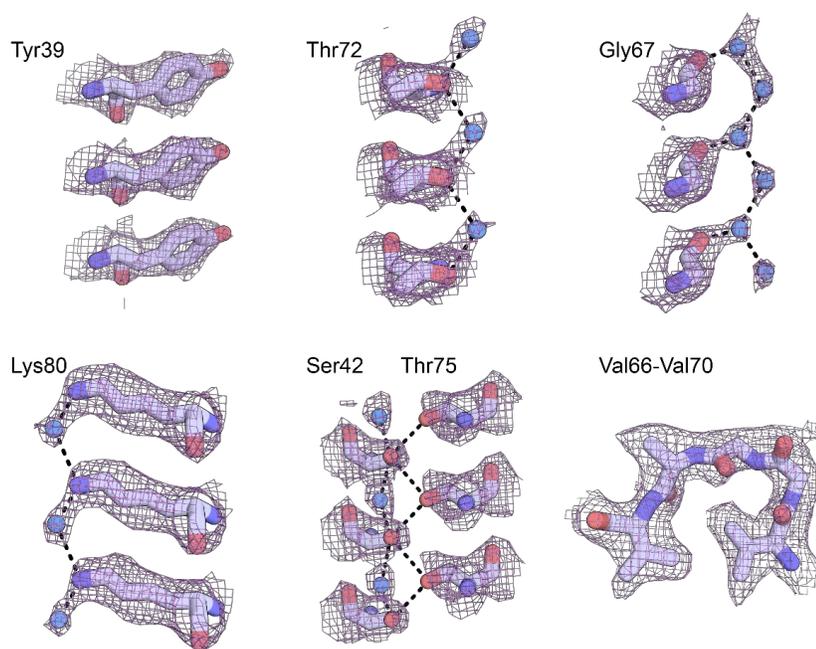
Supplementary Figure 1 Schematic representation and free energy of stabilization maps for the wild-type and E46K polymorphs

a-c) (left) Schematic representation of fibril structures with amino acid side chains colored as follows: hydrophobic (yellow), negatively charged (red), positively charged (blue), polar, uncharged (green), and glycine (pink). (right) Solvation energy maps of fibril structures. The stabilizing residues are red; the de-stabilizing residues are blue.



Supplementary Figure 2 Cryo-EM Data Collection and Processing

a) Representative cryo-EM image and power spectrum. Scale bar 50 nm. b) 1024 pixel box class averages reveal a helical pitch of ~ 800 Å. 686 pixel box class averages and corresponding simulated diffraction patterns reveal a rise of ~ 4.8 Å and a C_2 rotational symmetry. c) Gold-standard half-map-half-map FSC and map-to-model FSC.

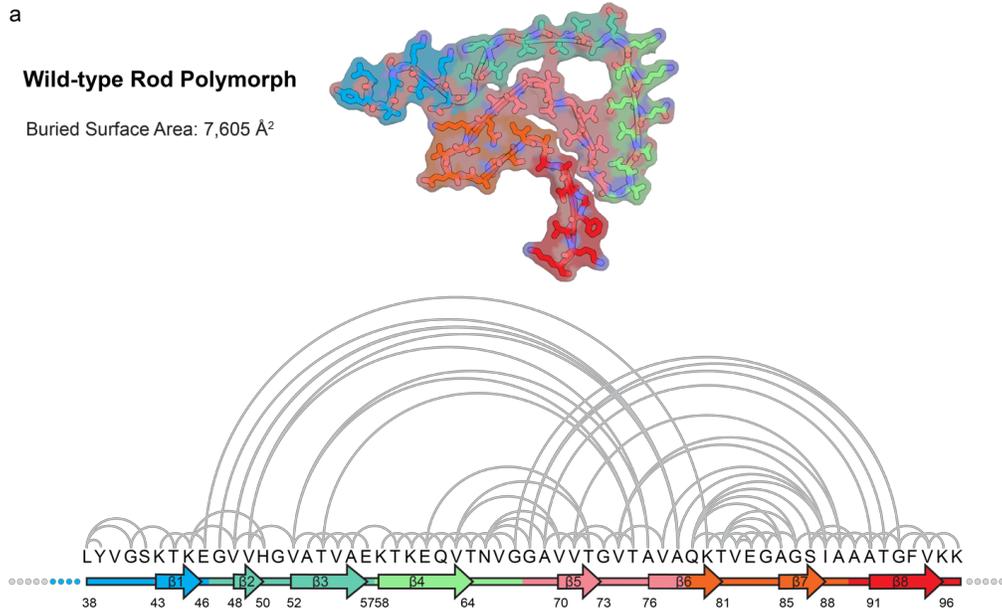


Supplementary Figure 3 Atomic model and density map highlight resolution of reconstruction and hydrogen bonding networks.

a

Wild-type Rod Polymorph

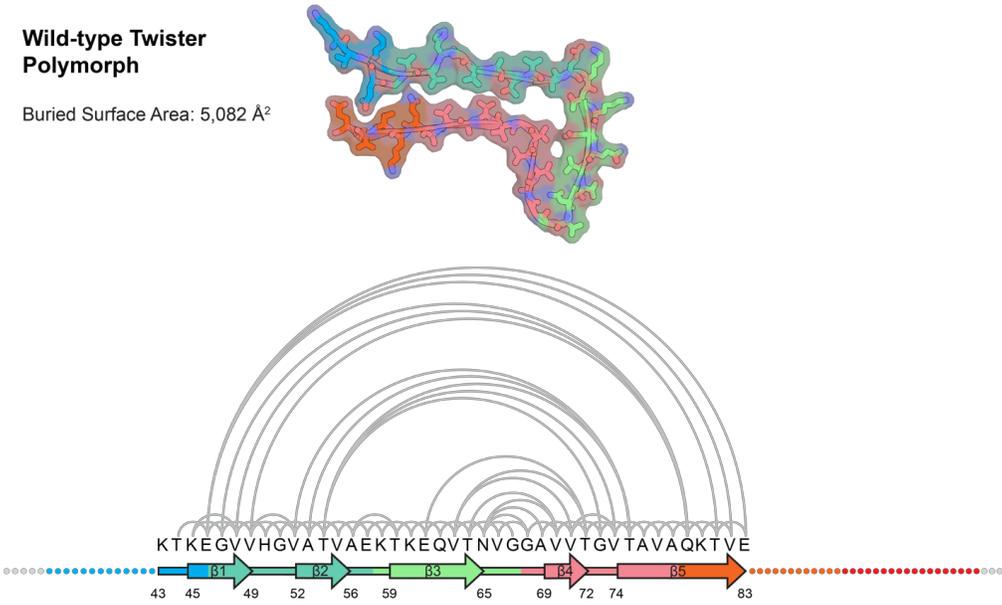
Buried Surface Area: 7,605 Å²

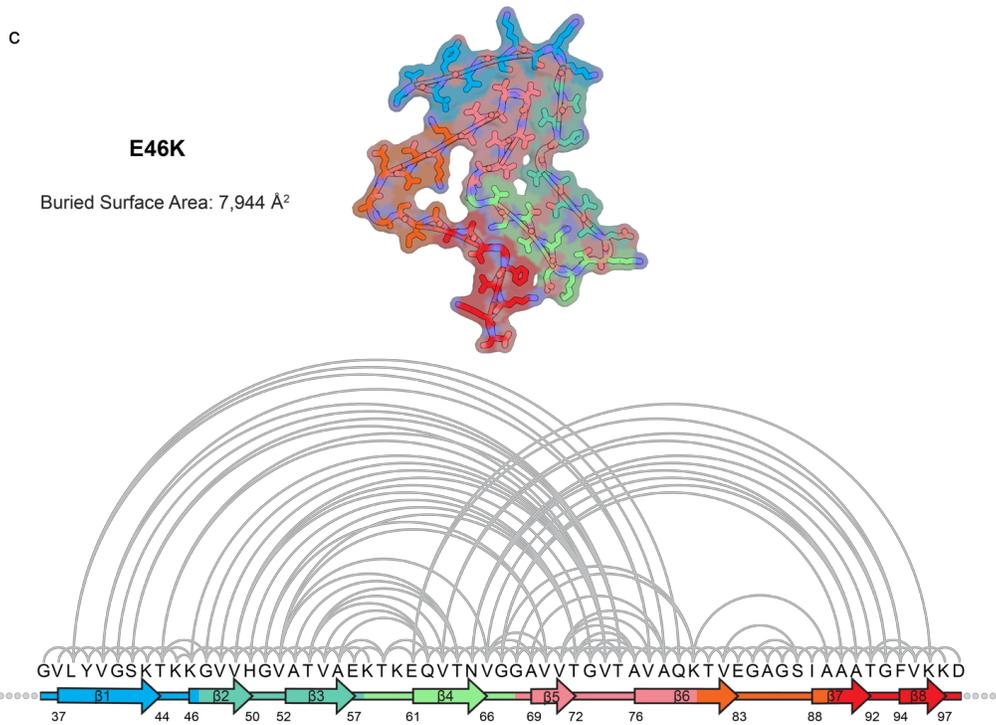


b

Wild-type Twister Polymorph

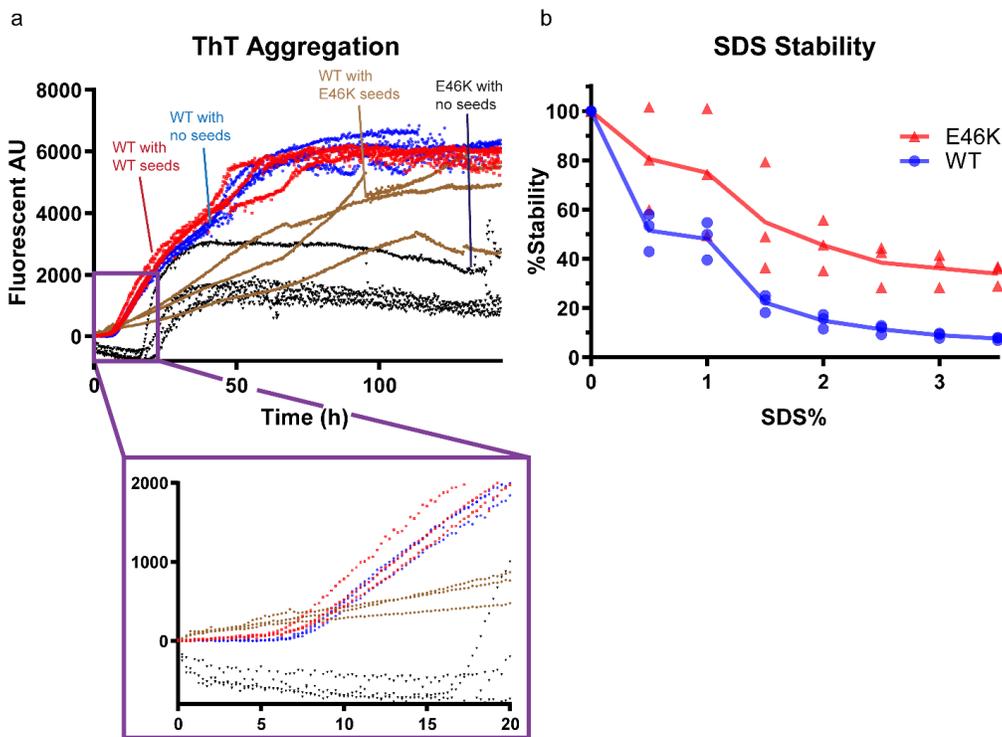
Buried Surface Area: 5,082 Å²





Supplementary Figure 4 Pairwise interaction analysis of wild-type and E46K protofilament folds

a-c) Protofilament fold and calculated buried surface area (top) and fibril core primary and secondary structure with pairs of interacting residues connected by half-circles (bottom) for a) wild-type rod, b) wild-type twister, and c) E46K polymorphs. E46K polymorph has a different set of interacting residues highlighting the difference in protofilament fold compared to wild-type. The E46K polymorph also has a larger number of interactions, which is reflected in its higher buried surface area.



Supplementary Figure 5 Cross-seeding of wild-type α -syn by E46K fibrils and SDS fibril stability assay

a) Seeding of wild-type α -syn by wild-type fibrils results in a modestly reduced lag-time while seeding by E46K fibrils eliminates the lag phase. In addition, both unseeded and self-seeded wild-type α -syn have similar ThT binding ability shown by their similar ThT aggregation curves; on the other hand, wild-type α -syn seeded by E46K fibrils has a different ThT fluorescence intensity, indicating a different underlying structure. Unseeded wild-type α -syn has a shorter lag phase and higher max ThT signal than E46K α -syn. Breaks in the ThT curves originate from the microplate-reader being interrupted and re-started to allow other experiments to be performed in separate wells in the same microplate. Due to normalizing the initial reading to zero, the E46K

lag phase dips below zero fluorescence AU. The reason for this slight dip after the initial ThT reading is unknown; however, the classical nucleation-elongation sigmoidal growth curve is still demonstrated for E46K α -syn. b) E46K and wild-type fibrils were heated to 37 °C and incubated with varying concentrations of SDS. E46K fibrils are more resistant to SDS than wild-type fibrils. Individual triplicate measurements are shown and the plotted line represents the average of the triplicates.

Title: Why amyloid fibrils have a limited width

David R. Boyer, Nikos A. Mynhier, Michael R. Sawaya

Abstract

Although amyloid fibrils can grow indefinitely along their ladder-like fibril axis by adding additional rungs of folded monomers to the top and bottom of the fibril, they do not grow laterally beyond ~10-20 nm. This prevents amyloid fibrils from growing into two-dimensional or three-dimensional arrays. However, given their ability to form cross- β structures where a primary tertiary structure element is the lateral association of β -sheets, the reasons for the limited width of amyloid fibrils are not immediately apparent. We hypothesize that the helical symmetry of amyloid fibrils limits their width because the separation of symmetrically related subunits in the helix becomes greater as a function of radial distance from the helical axis. Therefore, the backbone hydrogen bonds, as well as side chain interactions, that stabilize symmetrically related layers of the fibril become weaker towards the edge of the fibril, ultimately becoming too weak to remain ordered. To test our hypothesis, we examined 56 available cryo-EM amyloid fibril structures for trends in interstrand distance and β -sheet hydrogen bonding as a function of radial distance from the helical axis. We find that all fibril structures display an increase in interstrand distance as a function of radius and that most fibril structures have a discernible increase in β -sheet hydrogen bond distances as a function of radius. In addition, we identify a high resolution cryo-EM structure that does not follow our predicted hydrogen bonding trends and perform real space refinement with hydrogen bond distance and angle restraints to restore predicted hydrogen

bond trends. This highlights the potential to use our analysis to ensure realistic hydrogen bonding in amyloid fibrils when atomic resolution cryo-EM maps are not available.

Introduction

A wide variety of protein sequences can form amyloid fibrils, ranging from proteins that form amyloid fibrils during biological processes to proteins that only form fibrils under non-biological conditions in the test tube^{1,2}. However, even though the sequences and exact structures of individual amyloid fibrils can be quite different, all amyloid fibrils share a common blueprint³. To a first approximation, amyloid fibrils have a basic building block composed of one protein sequence that adopts a largely two-dimensional, serpentine-like fold. This fold has two open surfaces – the top and bottom of the two-dimensional layer – that interact repeatedly with other identical two-dimensional layers to form a helical lattice. The helix can extend indefinitely along the helical axis by adding further self-interacting protein chains to the “open” top and bottom layers of the fibril. This is opposed to most protein structures, which form a “closed” globular structure where other copies of the same protein do not add on to each other indefinitely, but exist as discrete, stable entities.

The architectural features of amyloid fibrils described above, as well as further details about the fibril structure, are encapsulated in the term cross- β fold. The cross- β fold describes not only the interactions that stabilize the subunits in the helix (quaternary structure), but also the interactions between different sections of the same protein chain (tertiary structure) and the conformation of contiguous sections of the protein chain (secondary structure). In the cross- β secondary structure, the amino acids in amyloid fibrils largely adopt β -strand conformations where the backbone carbonyl oxygens and amide hydrogens point alternately up and down the

fibril axis. In the tertiary structure, β -strands are punctuated by turns that allow the sequence to fold back on itself and form mated β -strands whose side chains interact in the direction orthogonal to the fibril axis. Finally, the cross- β quaternary structure is formed by individual two-dimensional protein chains stacking upon each other along the fibril axis, held together by the hydrogen bonding of the main chain carbonyl oxygens and amide hydrogens, into parallel, in-register β -sheets.

Given that lateral interactions between mated β -sheets in the amyloid fibril are so prominent in all amyloid fibrils, it is a mystery why amyloid fibrils do not grow wider than a certain amount. Indeed, in crystal structures of amyloid peptides, we often observe that the cross- β structure can extend many thousands of copies in the plane orthogonal to the fibril axis⁴. This is achieved through repeated lateral mating of β -sheets via their side chains in much the same way the β -sheets grow indefinitely along the fibril axis through repeated mainchain hydrogen bonding in both crystals and fibrils. Figure 1 illustrates the differences between crystals and fibrils by comparing the crystal structure of the longest amyloid peptide crystallized to date – A β 20-34 isoasp²³⁵ – with an amyloid fibril structure of a full-length amyloid protein – A- β 1-40⁶. In the crystal lattice, the peptides extend through side chain interactions in the entire plane orthogonal to the fibril axis, while in the fibril structure, the sides of the fibril structure are open for additional lateral interactions, yet no additional protein adds to the fibril laterally. In both cases, additional layers add along the fibril axis (coming in and out of the page) indefinitely through mainchain hydrogen bonding of β -sheets.

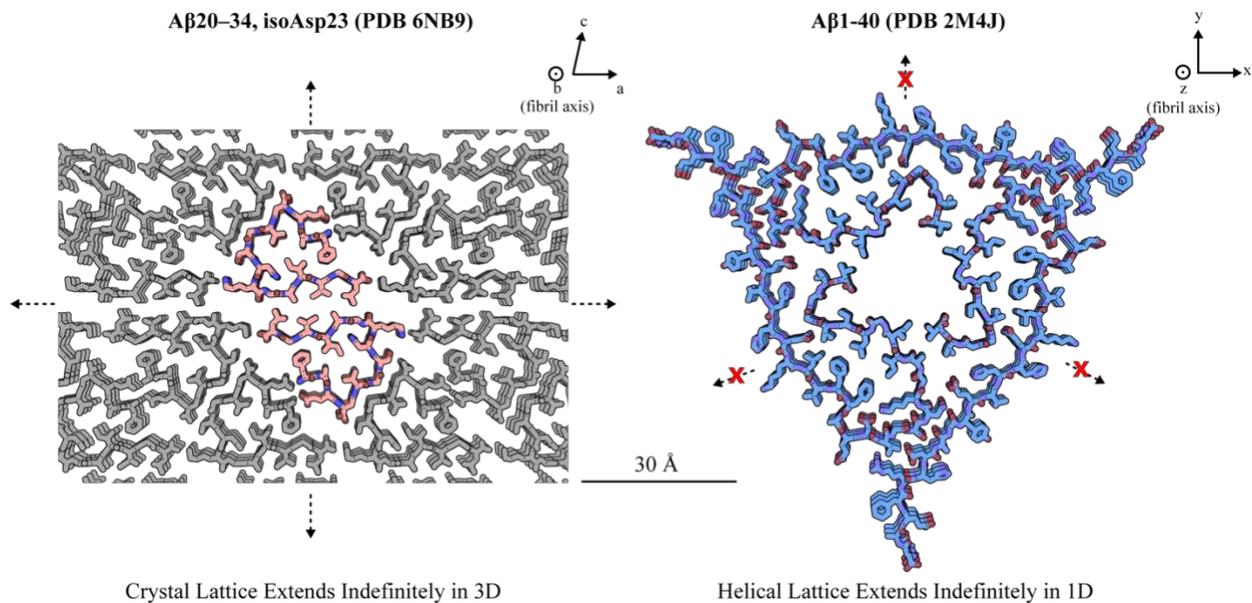


Figure 1. Comparison between A-β 20-34 isoasp23 crystal structure and A-β 1-40 fibril structure. The crystal lattice repeats indefinitely in all dimensions due to lateral side chain-side chain interactions as well as backbone hydrogen bonding along the fibril axis. The fibril does not grow laterally through repeated mating of additional protofilaments, and instead, extends indefinitely along the fibril axis through backbone hydrogen bonding, similar to the crystal lattice.

There are two possible mechanisms for amyloid fibrils to grow wider: i) additional protofilaments can be added to exposed surfaces of the fibril, ii) additional residues of the protein chain add to the cross-β fold. The former case is akin to that of a crystal lattice – addition of protein chains orthogonal to the fibril axis, while the latter case is only applicable to fibrils since in amyloid fibrils formed from full-length protein there are generally only a certain subset of amino acids in the protein chain that participate in forming the cross-β fold of the fibrils – the “fibril core”. Residues not in the fibril core form the “fuzzy coat” – a disordered tangle of amino acids coating the length of the fibril. The fact that additional protofilaments do not add

indefinitely to the fibril and that many residues are found in the fuzzy coat add to the mystery of why amyloid fibrils do not grow wider.

A key feature of amyloid fibrils is their helicity, which arises from the twisting of the β -strands in each protein chain. Twisted β -strands generate twisted β -sheets, due to the fact that the β -strands in a β -sheet have a repeated, asymmetric interface⁷. In general, a right-handed β -strand twist is thought to be more energetically stable for L-amino acids, and this gives rise to a left-handed β -sheet twist⁸. However, there are also examples of left-handed β -strands and corresponding right-handed β -sheets in nature⁹. Figure 2 shows a prototypical left-handed double helix where the relationship between symmetrically related subunits is given by two parameters: the helical *twist*, in $^\circ$ /subunit, which describes the incremental rotation of each subunit around the helical axis and the helical *rise*, in \AA /subunit, which describes the translation of the subunits along the helical axis. In cylindrical coordinates, the relationship between identical objects in a helix is given by:

$$\rho(r, \phi, z) = \rho(r, \phi + h * \textit{twist}, z + h * \textit{rise}) \textit{ for any integer value } h \quad (1)$$

Where r is the radial distance from the helical axis, ϕ is the angular coordinate, and z is the height. And the distance between two objects in cylindrical coordinates is:

$$l = \sqrt{r_1^2 + r_2^2 - 2r_1r_2(\cos(\phi_2 - \phi_1) + (z_2 - z_1)^2)} \quad (2)$$

Which reduces to Eq. 3 when calculating the distance between subunits i and $i+1$ (two consecutive layers in a helix, Figure XXX):

$$i, i + 1 \textit{ distance} = \sqrt{2 * \textit{radius}^2 (1 - \cos(\textit{twist})) + (\textit{rise})^2} \quad (3)$$

Eq. 3 shows that the distance between the $i, i+1$ subunits is proportional to their radial distance from the helical axis. In addition, Eq. 3 demonstrates that the $i, i+1$ distance will be dependent on the twist and rise parameters of the helix. In the case of amyloid fibrils, the rise will always be

~4.8 Å per layer; however, the twist can vary for each fibril. Fig. 3 shows the relationship between the $i, i+1$ distance and radius for different twist values. The $i, i+1$ distance increases as a function of radius according to Eq. 3, with the $i, i+1$ distance increasing more quickly for higher twist values.

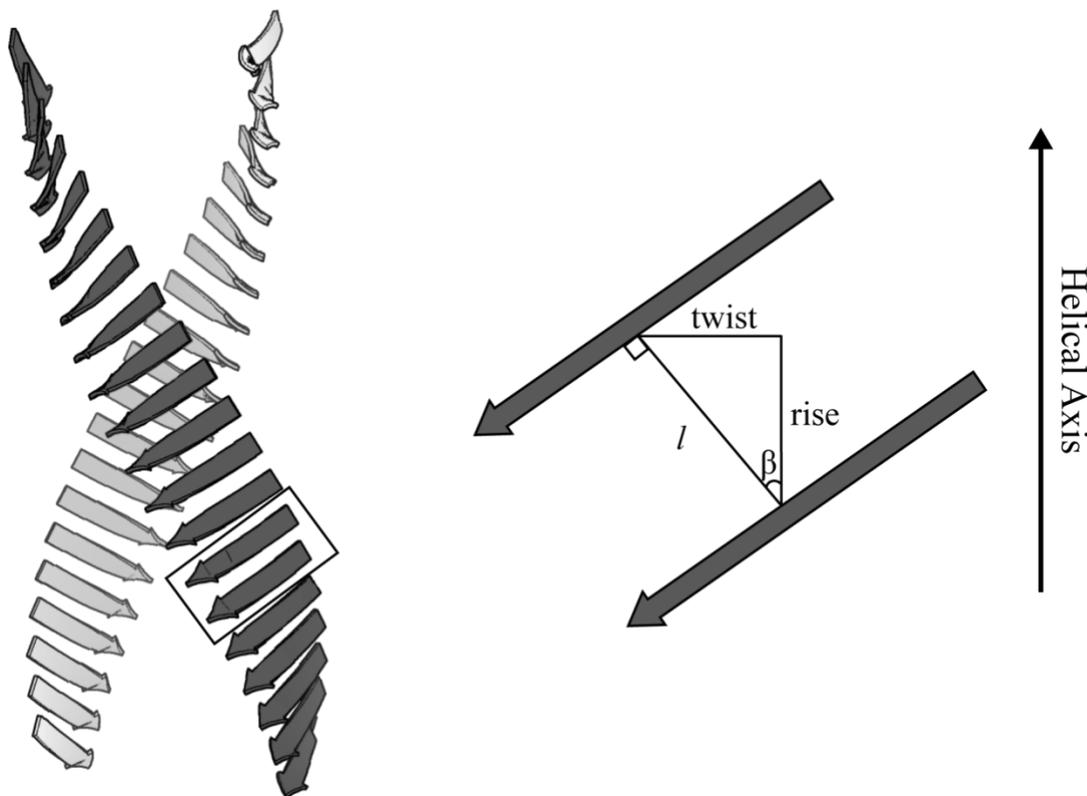


Figure 2. The relationship between subunits in a helix can be described by an azimuthal twist around the helical axis and a translational rise along the helical axis. The displacement between adjacent subunits is l , which can be calculated with Eq. 3. In order to maintain hydrogen bond geometry between adjacent beta-strands in an amyloid fibril, the strands will tilt according to Eq. 4.

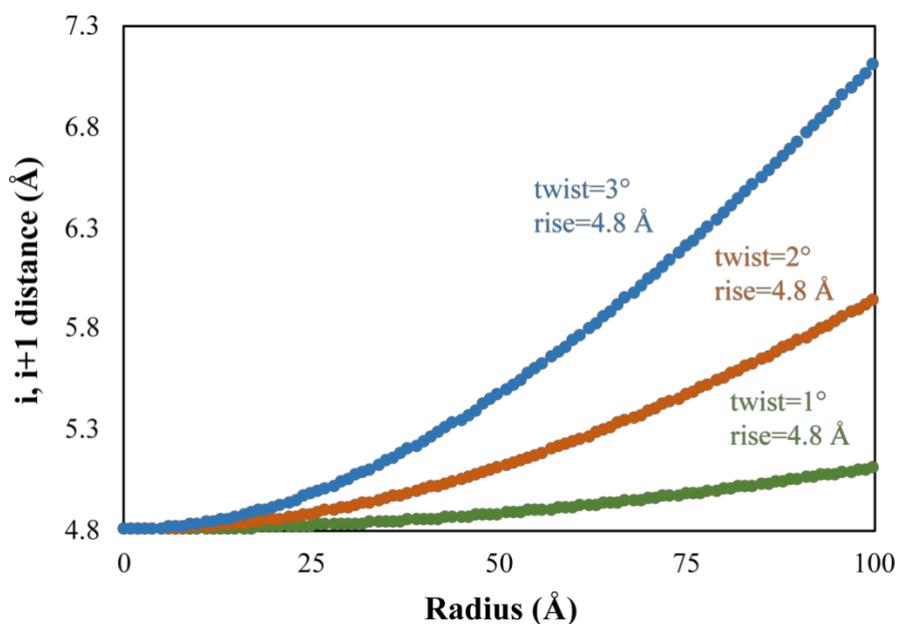


Figure 3. Simulated $i, i+1$ distances versus radial distance from the helical axis for different twist values. For helices with a larger twist, the distance between adjacent subunits increases more quickly as a function of radial distance from the helical axis.

We hypothesize the helical symmetry of amyloid fibrils places a restraint on the maximum width of the fibril. As the distance between the $i, i+1$ β -strands towards the edge of the fibril increases, they will not be close enough for β -sheet hydrogen bonding and will not be able to form part of the ordered core. Furthermore, as Eq. 3 and Figure 3 show, fibrils with a larger twist will have a smaller maximum width since the $i, i+1$ distance will increase more rapidly. Conversely, fibrils with a smaller twist can grow to a larger maximum width.

Another consequence of the helical symmetry of amyloid fibrils is that as the β -strands get farther from the helical axis, they will have to tilt about the axis orthogonal to the fibril axis in order to maintain the inter-strand hydrogen bond geometry. In Figure 2, this is the β angle

between the fibril axis and the line connecting identical parts of the subunits in the helix. As the radial distance from the helical axis grows, and the distance l between $i, i+1$ subunits grows, so will β according to Eq. 4:

$$\beta = \cos^{-1} \frac{\text{rise}}{l} \quad (4)$$

In order to illustrate how the helicity of amyloid fibrils affects the β -sheets composing the fibril as the distance from the helical axis increases, we took the crystal structure of tau peptide SVQIVY and made a theoretical helix (Figure 4) with a 1° /subunit twist (a twist typical of amyloid fibrils). To construct the helix, we used the crystallographic symmetry found in the SVQIVY crystal structure¹⁰ to add consecutive peptides laterally outward from the helical axis to a maximum radial distance of 150 \AA , giving the fibril a diameter of 250 \AA . Then, we applied helical symmetry to one layer of the fibril to generate the full helix. We made two copies of the helix: one where the strands in the starting layer were left un-tilted and one where the strands were tilted according to Eq. 4.

Figure 4 shows that as β -sheets get farther from the helical axis, the distance between the i and $i+1$ atom increases according to Eq. 3. Also, it is clear that without the tilting of the strands counterclockwise around the axis orthogonal to the fibril axis, the hydrogen bond geometry would not be maintained (Figure 4c). The top graph in Figure 4d shows the measured $i, i+1$ distances for all atoms as a function of radial distance from the helical axis while the middle and bottom graphs show the measured hydrogen bond distances and β -strand tilts as a function of radial distance from the helical axis, respectively. As expected, the hydrogen bond distances and tilts for the $i, i+1$ strands increase with increasing interstrand distance according to the rules of helical symmetry.

The theoretical helix of Fig. 4 is only a thought experiment. In the real life version of this helix with a twist of 1° /subunit, the strands would have likely stopped adding laterally to the helix at around $\sim 70 \text{ \AA}$ once the $i, i+1$ distance became greater than $\sim 4.9 \text{ \AA}$ (roughly $\sim 60 \text{ \AA}$ radial distance at 1° /subunit twist), the backbone hydrogen bonding in the β -sheets became greater than $\sim 2\text{-}2.2 \text{ \AA}$, and the strand tilt became greater than 15° . Therefore, at least in our thought experiment, the helicity of amyloid fibrils limits their ability to grow wider than a certain amount.

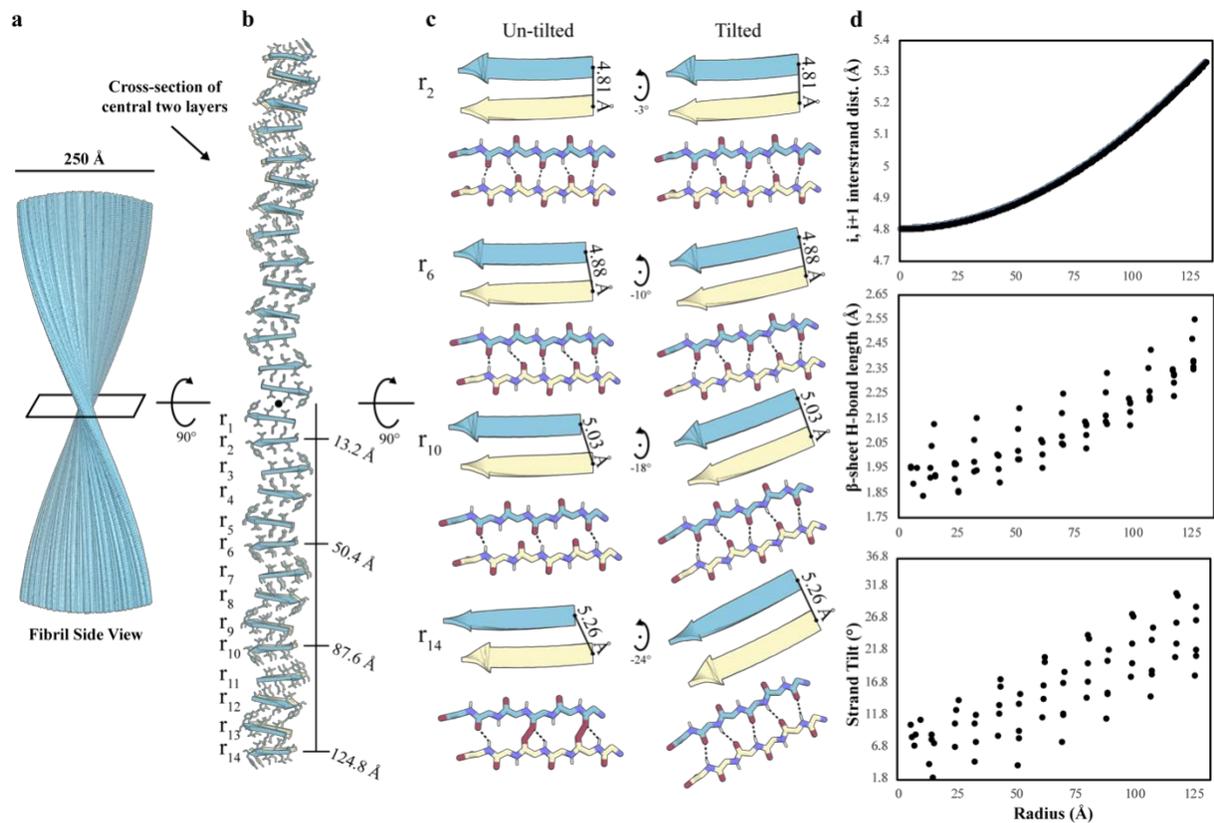


Figure 4 A theoretical amyloid fibril built by adding strands of peptide SVQIVY laterally outward from the helical fibril following the crystallographic symmetry of the SVQIVY peptide and then helicizing the model with a 1° /subunit twist and 4.8 \AA rise. a) Side view of the SVQIVY helix. b) Central cross-section of the helix highlighting the radii of different protofilaments of the fibril. c) Side-view of selected protofilaments at increasing radius from the

helical axis – r_2 (radius = 13.2 Å), r_6 (radius = 50.4 Å), r_{10} (radius = 87.6 Å), and r_{14} (radius = 124.8 Å) – demonstrates that the distance between strands grows larger as the radius increases and that beta-strands need to tilt to maintain hydrogen bond geometry as the radius increases. d) Quantification of symmetry related $i, i+1$ atoms, β -sheet hydrogen bond lengths and off-axis tilts as a function of radius from helical axis.

Results

Comparing Twist versus Fibril Radius

In order to examine whether the trends we predict are observed in real-life amyloid fibril structures, we conducted an analysis of 57 cryo-EM amyloid fibril structures determined to date (see Supplementary Table 2, Pick's Disease Wide Fibril used in this analysis, see Methods). We selected cryo-EM structures as opposed to those determined by solid state NMR since the helical twist and rise are known accurately in cryo-EM helical reconstructions. For the 55 structures, we first tested to see if there was any correlation between fibril width and fibril pitch (pitch is the translational distance along helical axis to complete a full 360° rotation; smaller twist values have longer pitch) since according to Eq. 3 and Figure 3, if fibrils have a large twist, the distance between symmetrically related $i, i+1$ atoms will become greater more quickly as the radial distance increases compared to fibrils with a smaller twist.

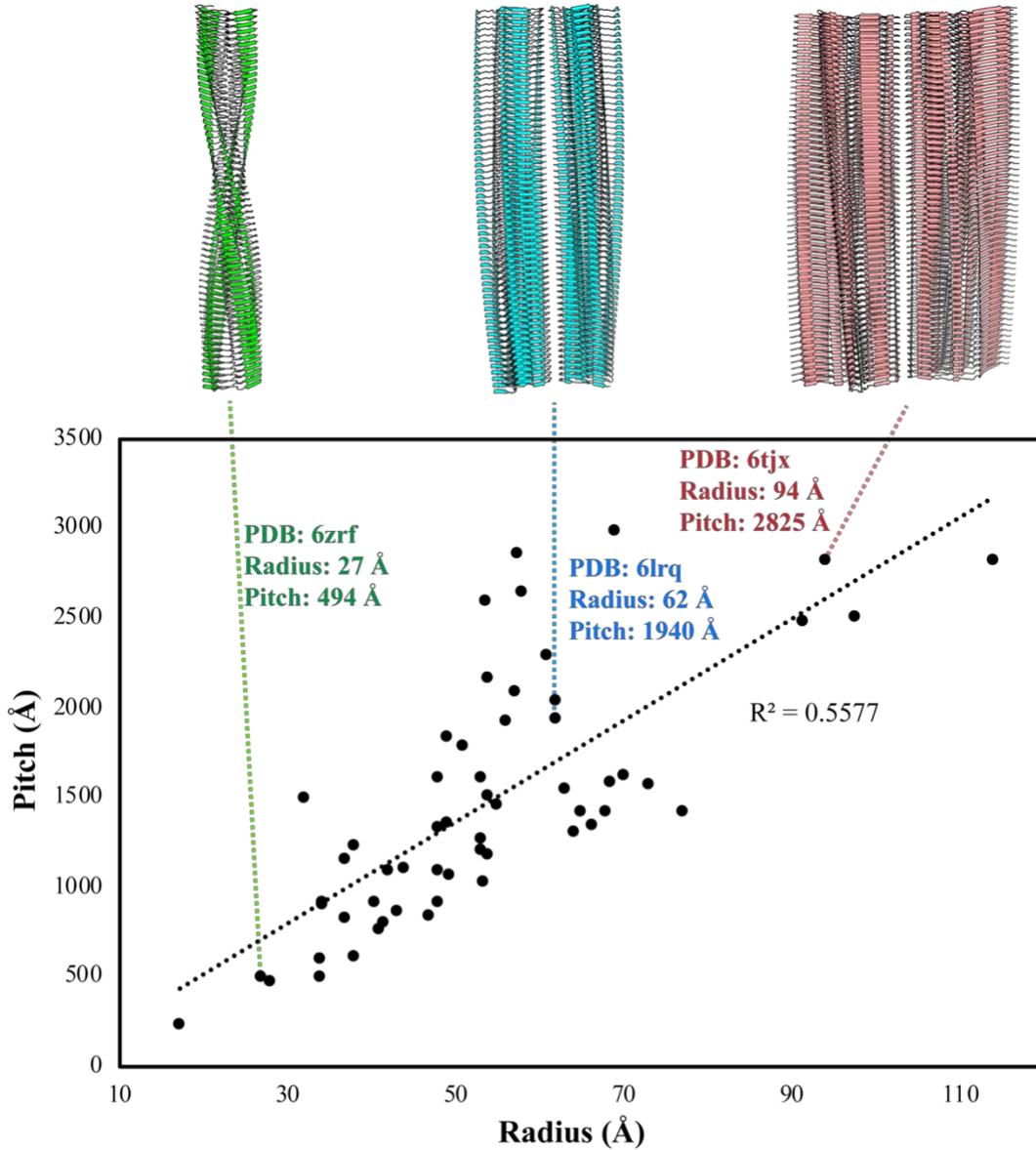


Figure 5 Plot of pitch versus radius for 57 cryo-EM structures demonstrates the general rule that fibrils with a larger radius have a larger pitch.

Figure 5 demonstrates that, in general, the wider a fibril is, the smaller its twist (and hence larger its pitch) must be. It also demonstrates that there is not a perfect inverse correlation to this finding; in that, there is no lower limit on the twist. Fibrils of any width may have a small

twist as judged by the fibrils that deviate from the trendline in the upper half of the plot. This finding supports our prediction that wider fibrils must have smaller twist values. In addition, this result is consistent with the correlation found by Wu, et al. between the number of residues in the fibril core (a proxy for fibril width) versus fibril pitch when they correctly identified that the very small core of RIPK3 fibrils in their study allowed for a large helical twist¹¹.

Computational analysis of all cryo-EM fibril structures

We next examined 55 fibril structures with atomic coordinates deposited in the PDB to test whether our predictions for the $i, i+1$ distance vs. radius, backbone hydrogen bonding length vs. radius, and backbone hydrogen bonding tilt vs. radius trends follow the trends we observed in our theoretical SVQIVY helix (Figure 4). Due to the fact that the hydrogen bond measurements we are making are very sensitive to the atomic coordinates of the fibril structures and that cryo-EM maps of amyloid fibrils have yet to reach unambiguous, atomic-resolution detail, we wanted a metric for how well the density supports each atomic coordinate. This would allow us to consider measurements from the different fibrils in light of the accuracy of the map and model. We therefore calculated and ranked all structures by their Q-score measurement in order to determine the most well resolved fibrils (Supplementary Table 1, Supplementary Figure 1) to aid our computational analysis¹².

Figure 6 shows example plots for four of the structures with Q-scores corresponding to a resolution of the model and map that is better than 2.5 Å (see Supplementary Table 1). For each structure analyzed in Figure 6 (6ufr¹³, 6xyq¹⁴, 6xyo¹⁴, and 6nwp¹⁵), the left most graph shows the predicted and measured distance between symmetrically related atoms in adjacent layers in the fibril versus the distance from the helical axis for that atom pair. As Eq. 3 predicts, this distance

increases at the wider parts of the fibril. The middle plot for each structure measures the backbone hydrogen bond lengths in β -sheets in adjacent layers in the fibril as a function of radius from the helical axis, while the right-most graph measures the tilt off the fibril axis of the same hydrogen bonds. Due to the increasing distance between symmetrically related atoms as a function of radius (left-most graph), we expect both the hydrogen bond length and tilt to increase in a manner similar to our theoretical SVQIVY helix (Figure 4). We observe that the structures in Figure 6 do indeed show modest increases in hydrogen bond length (middle plots) and off-axis tilt (right plots) as the distance from the helical axis increases. These trends support our predictions that hydrogen bonding between adjacent layers will become weaker at the wider parts of the fibril, potentially placing a limit on the maximum width a fibril can grow.

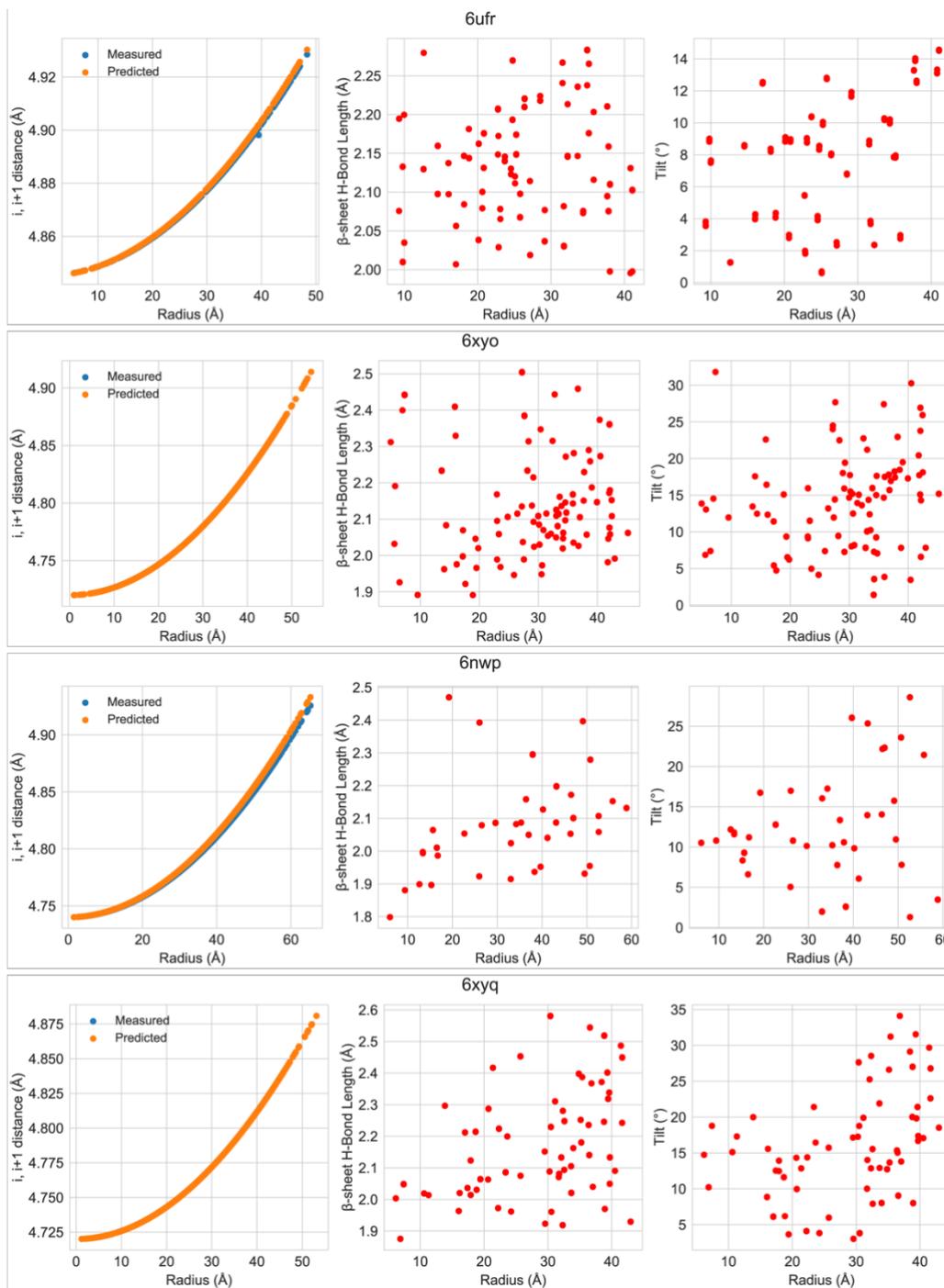


Figure 6 Plots of $i, i+1$ distance for symmetrically related atoms (left), β -sheet backbone hydrogen bond distances (middle) and off-axis tilts (right) for four of the highest resolution cryo-EM fibril structures.

In our survey of all 55 fibril structures included in our analysis, although all structures display the expected increase in distance between symmetry related atoms as the distance from the helical axis increases, there are some that show no discernible trend in the hydrogen bond length and off-axis tilt as a function of radial distance from the helical axis (Supplementary Figure 2). Figure 7 shows plots of example fibril structures with no apparent trend in the hydrogen bond length and off-axis tilt as a function of radius. Three of these structures (6zrq¹⁶, 6shs¹⁷, 6cu8¹⁸) have a lower predicted map/model resolution (3.4 Å, 4.0 Å, 3.5 Å, respectively) and a relatively small fibril core (34 Å, 43 Å, 48 Å, respectively). These characteristics potentially contribute to the fact they do not show obvious trends in backbone hydrogen bond length and tilt as a function of radius because the lower resolution of the map/model potentially limits the accuracy of the atomic coordinates while the smaller fibril core limits the number of backbone hydrogen bonds that can be used in the analysis.

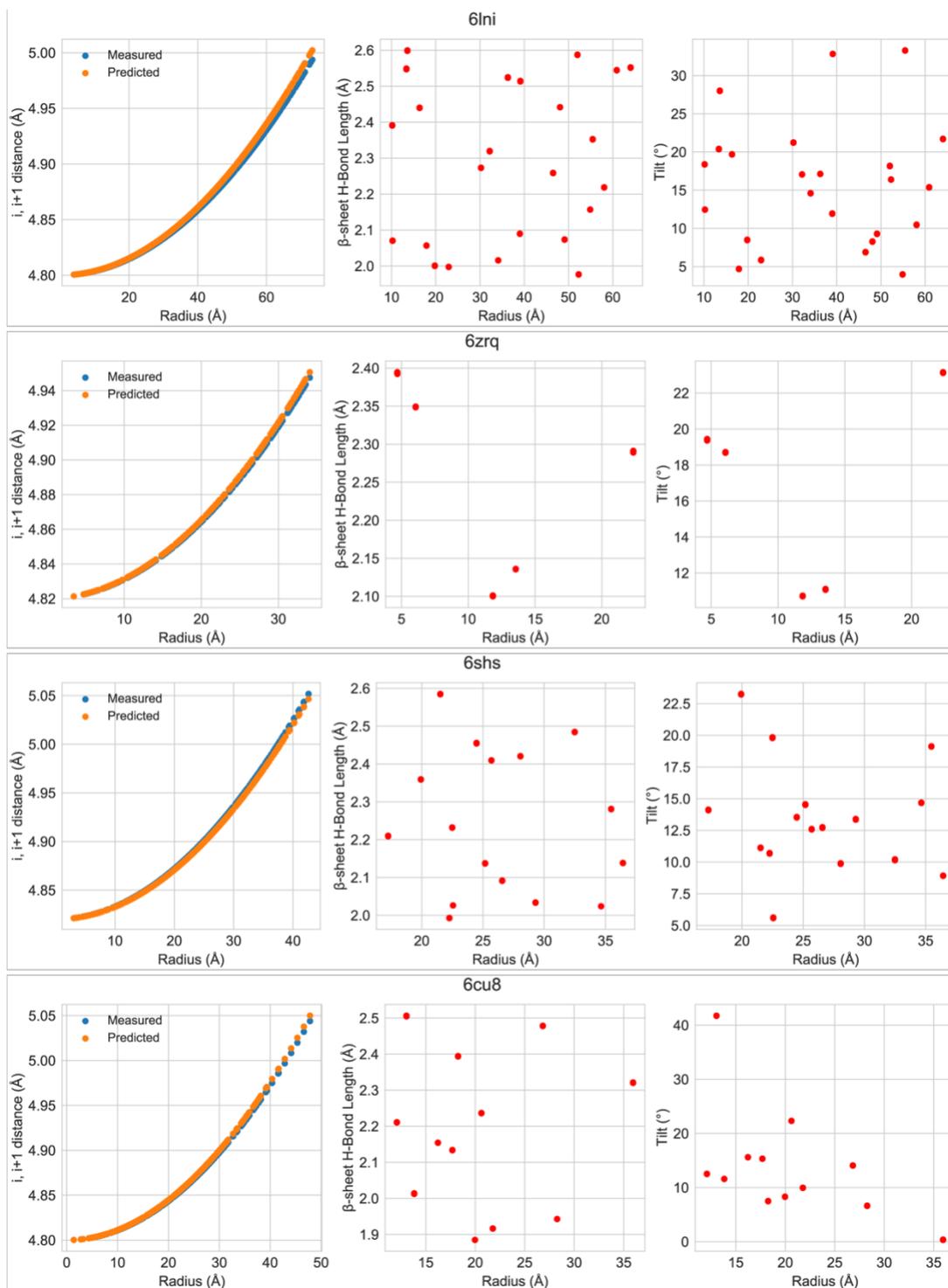


Figure 7 Examples of fibril structures that do not display the expected trends in β -sheet hydrogen bond length and off-axis tilt.

Restoring Hydrogen Bond Length and Tilt Trends of a High-Resolution Fibril Structure

Unlike the previous examples with low resolution and small fibril cores, PDB 6lni¹⁹ has a good Q-score predicted resolution (2.3 Å) and a relatively large fibril core (73 Å radius). Upon examination of the structure, it is clear that ideal restraints for hydrogen bond lengths and angles were not employed during structure refinement (Figure 8). This highlights the fact that even the higher resolution amyloid fibril cryo-EM maps determined to date do not have sufficient resolution to accurately model all non-covalent interactions and emphasizes the need to use geometric and chemical restraints for hydrogen bonding.

In order to test whether refinement of 6lni with hydrogen bond length and angle restraints led to the predicted increase in β -sheet hydrogen bond lengths and off-axis tilts as a function of radius, we added hydrogen bond length and angle restraints to the deposited PDB structure and refined the structure with restraints against the deposited map using phenix.real_space_refine²⁰. Figure 8 shows representative backbone β -sheet hydrogen bonding before and after refinement for two β -sheet regions, one close to the helical axis and one at the edge of the fibril. The refined structure shows backbone β -sheet hydrogen bonding lengths and angles closer to idealized parallel, in-register β sheets (Figure 8 b-d). We then re-analyzed the backbone β -sheet hydrogen bonding lengths and off-axis tilts as a function of radius and saw that our predicted patterns were more apparent (Figure 8 e) compared to the analysis of the structure from the PDB (Figure 7). This highlights the importance of using hydrogen bonding restraints during structural refinement of near-atomic models to model non-covalent interactions as accurately as possible. In addition, the analysis of the refined structure supports our prediction that backbone hydrogen bond lengths and off-axis tilts will increase proportional to their distance from the helical axis of the fibril. Furthermore, it is likely that the combination of limited cryo-EM map resolution and lack of

hydrogen bond restraints during refinement contributes to why some of the 55 structures in our analysis do not follow our predictions.

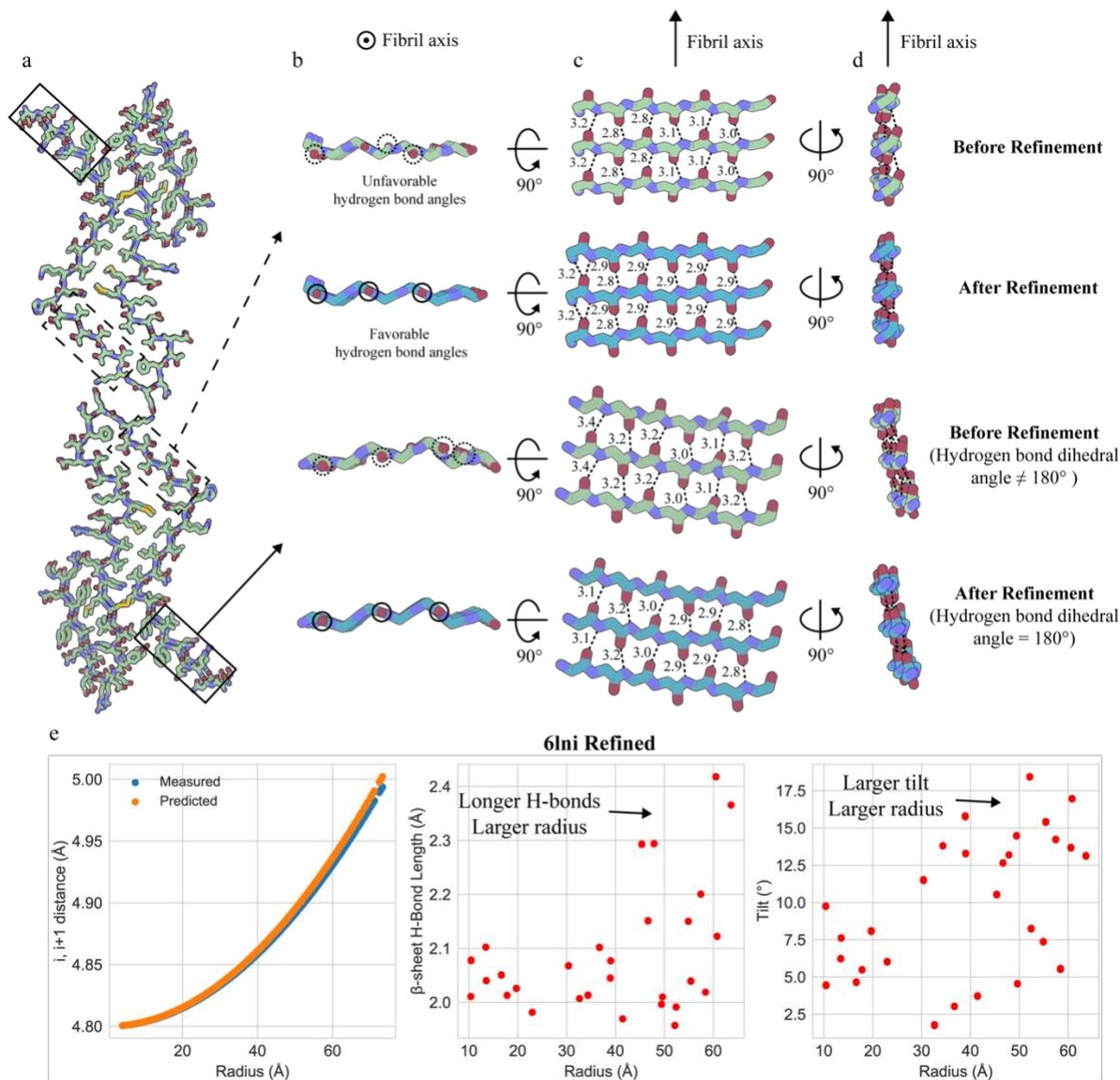


Figure 8 Real space refinement of PDB 6lni with hydrogen bond length and angle restraints restores predicted trends. a) Top-view of fibril 6lni formed by full-length human prion protein. b) Top-view of boxed sections showing backbone atoms before and after refinement with restraints. c) Side-view of boxed sections showing hydrogen bonding before and after refinement. d) Alternative side view of boxed sections highlighting both the correction of hydrogen bond

dihedral angles and the tilting of strands that are farther from the helical axis. e) Computational analysis of interstrand distance and β -sheet hydrogen bond length and tilt as a function of radius after refinement with hydrogen bond restraints.

Discussion

Through a combination of theory and structural analysis, we have shown that the distances between adjacent layers in amyloid fibrils increase as the fibril grows wider, and that for fibrils with a faster twist, the distances increase more quickly. This causes the β -sheet backbone hydrogen bonds that stabilize the fibril to grow longer, and hence weaker, towards the edges of the fibril, potentially causing fibrils to have a maximum width for a given fibril twist value. Using our analysis of hydrogen bond lengths and off-axis tilts, we identified a fibril structure with a high-resolution cryo-EM map that does not follow our predicted trends and showed that refinement with hydrogen bond length and angle restraints restored the expected hydrogen bonding patterns. This example highlights the potential utility of examining hydrogen bond lengths and off-axis tilts relative to radial distance for quality control during atomic model building into near-atomic cryo-EM maps.

Our analysis of 55 cryo-EM fibril structures indicates that the increase in hydrogen bond lengths and tilts is not as pronounced or noise-free as our theoretical SVQIVY helix (compare Figure 4d to Figure 6 and 7, Supplementary Figure 4). There are several possible reasons for this. First, the SVQIVY helix was constructed to have extra β -sheets added laterally to highlight the effects of helical symmetry on interstrand distances and hydrogen bonding. Therefore, the measurements that show the greatest effect of helical symmetry (when atoms are very far from the helical axis) will not exist in a realistic fibril since the hydrogen bonding between adjacent β -

strands will be too weak (interstrand distances $> \sim 5.0 \text{ \AA}$, Figure 4 c-d). Instead, only those within a range depending on the twist of the fibril will be observed in real life fibrils, leading to the more modest trends we observe in our analysis of cryo-EM fibril structures. Second, the SVQIVY helix was based on a high-resolution microED crystal structure where the position of the atoms involved in hydrogen bonds are known to a high accuracy. This is unlike cryo-EM fibril structures that have lower resolution, making the positions of atoms approximate and necessitating great care in modeling hydrogen bonds. Third, in crystal structures of amyloid-like peptides, almost all residues adopt a strict β -strand conformation due to the tight packing of the peptides in the crystal lattice as well as the non-twisting nature of the peptides. The lower resolution of the fibril structures, as well as the greater diversity of residue conformations in a full-length fibril with twisted β -strands make our backbone hydrogen bond measurements noisier in comparison to a theoretical helix based on a high resolution crystal structure.

In addition to the strain the helical twist imposes on the perimeter of fibrils, there may be other factors that limit the width of amyloid fibrils. One such factor is that not all amino acids are equally amyloidogenic. In other words, some sequences tend to not to be compatible with the cross- β fold²¹. For instance, proline residues disallow the typical Ramachandran angles of β -strand secondary structure, and charged residues may experience electrostatic repulsion in the cross- β fold due to the 4.8 \AA stacking of identical amino acids in parallel, in-register β -sheets along the fibril axis.

A good example for this situation is the protein tau, which forms amyloid fibrils in Alzheimer's and other diseases^{15,22-26}. The longest isoform of tau is 441 residues, with a large N-terminal sequence that is highly positively charged and proline rich, four C-terminal microtubule binding pseudo-repeat domains containing known amyloidogenic sequences, and an additional

~40 residue far C-terminal domain. In the fibril structures of tau, 60-90 residues of the repeat domains form the fibril core, while the remaining 300+ residues from the N- and far C-terminus are disordered in the fuzzy coat. This is consistent with the fact that the N-terminus is proline-rich and highly charged, but interestingly, the amyloid-segment predicting software ZipperDB²¹ indicates there are amyloidogenic segments in the ~40 residues of the far C-terminus. However, in all the tau fibril structures determined thus far, these additional amyloidogenic sequences do not contribute to the fibril core.

We demonstrated that there is a correlation between fibril twist and width; namely, that fibrils with a large twist must have smaller cores, while those with a smaller twist can have larger cores (Figure 5). However, it is not clear whether it is the twist that informs the size of the core, or if the size of the core determines the helical twist. Chou and Scheraga used computational chemistry techniques to show that β -strands favor a right-handed twist (more so than left-handed or flat) due to the steric interactions of $i, i+2$ side chains in L-amino acids; however, the twist of the β -strands is reduced in β -sheets – compared to isolated β -strands – due to interchain interactions (e.g., backbone hydrogen bonding, side chain-side chain interactions) “flattening” the twist of the constituent strands⁸. The large number of β -sheet residues in an amyloid fibril core therefore leads to many interchain interactions (e.g., backbone hydrogen bonding) that serve to flatten the twist of the constituent β -strands leading to extremely small twists ($\sim 0.5^\circ - 3.5^\circ$, Supplementary Table 2) compared to the large twists ($>10^\circ$) seen in β -sheets in globular proteins. Therefore, we hypothesize that it is the amount of residues that participate in forming the fibril core that determines the twist of the amyloid fibril. This leads to smaller cores having larger twists and larger cores having smaller twists. Then for a given fibril core, the helicity of the fibril will make the β -sheets at the edges of the fibril less stable, leading to either i) less residues from

the constituent protein chains being able to be added or ii) the inability of additional protofilaments to be added laterally to the fibril.

For scenario (i) above, one could imagine a theoretical protein sequence of 1000 residues where every residue could adopt the cross- β fold. Given the conclusion of Chou and Scheraga, if all 1000 residues of our imaginary protein formed part of the fibril core, the twist of the fibril could approach zero since the amount of interchain interactions favoring flattened β -strands would outweigh the $i, i+2$ side chain steric interactions that favor a twisted β -strand. Further study would be needed to see if this were indeed possible, or if there will exist a strain in the polypeptide chain that will cause it to twist, leading to a limited width for the fibril.

A potentially illuminating example for the above situation is the fused in sarcoma (FUS) protein. FUS is a 526 residue protein with a 214 residue N-terminal low complexity domain (LCD). Recent studies have shown that, when produced separately, both the N- (2-108) and C-terminal (111-214) halves of the LCD can form fibrils, while the full LCD forms fibrils with only the N-terminal core^{27,28}. This raises the question why a larger core, containing both the N- and C-terminal LCD, does not form. The authors in Lee, et al. correctly point out that if a putative fibril formed having both the N- and C-terminal LCD fibril cores present, and the fibril followed the same helical twist as the C-terminal-only fibril, the interstrand distances in the N-terminal core (the periphery of the fibril) would become too great²⁸. However, it is possible that in a fibril where both sequences contribute to the core, the helical twist may become smaller due to the additional interstrand interactions flattening the strands, thereby permitting a wider fibril. Or a fibril may form with a new fibril core structure where both N- and C-terminal LCD sequences interact and the twist similarly becomes smaller since there are more interchain interactions flattening the fibril. The fact that neither of these scenarios occur, at least under

tested experimental conditions, may be evidence that fibrils cannot grow indefinitely wide since there may be some amount of twist necessary to prevent straining the polypeptide chains, hence limiting the width of the fibril. This may be similar to tau where even in the largest diameter tau fibril²⁶ (PDB 6tjx; diameter 388 Å), with a twist of -0.61° , containing all four microtubule repeat domains (residues 274-380) there exist residues (381-441) just beyond the C-terminal end of the core that are predicted to be amyloidogenic by ZipperDB that are nonetheless not included in the fibril core.

Numerous studies have identified apparently non-twisting amyloid fibrils. For example, Schweighauser, et al. and Li, et al, identified non-twisting fibrils of alpha-synuclein from Parkinson's Disease and from recombinantly assembled wild-type fibrils, respectively^{14,18}. Due to their lack of helicity, their structures are unable to be determined by current cryo-EM helical reconstruction methods (although ssNMR would be ideal for structure determination since it does not rely on the helicity of fibrils). At present, it is unknown what factors lead these fibrils to not twist. Liberta, et al. demonstrated right- and left-handed serum amyloid A protein amyloid fibrils from human and mouse, respectively²⁹. Their analysis of Ramachandran angles showed that right-handed fibrils have a slight majority of residues in left-handed beta-strand conformation and vice versa for the left-handed fibrils. Therefore, one possible explanation for non-twisting fibrils is that there is an equal ratio of right- and left-handed beta-strands in the protein chains, although this has not yet been proven experimentally.

Our theory predicts that the helicity of amyloid fibrils constrains their width; however, this leads to the question: why do the observed non-twisting fibrils not associate laterally to create ordered two- or three dimensional arrays? One possible explanation for this is that the external surfaces of fibrils are not amenable to cross-beta association. This could be due to the

fact that charged residues tend to prefer to interact with water and remain solvent exposed rather than be buried inside the fibril core, as in globular proteins. However, there are numerous cases, including recombinantly assembled wild-type³⁰, E46K¹³, and Tyr39p³¹ alpha-synuclein, and full-length human prion protein¹⁹ fibrils where protofilaments assemble strictly through electrostatic interactions. This suggests that fibrils can indeed extend their width through lateral association of protofilaments mediated by electrostatic interactions, and that in these cases, the helical twist may impose the limit on width and not the surface characteristics of the fibril. Given these examples, it is unclear why non-twisting fibrils have not yet been observed to associate laterally. Perhaps due to their rarity, a fibril surface composition for non-twisting fibrils has not yet been sampled that may allow more extensive lateral growth. Further work determining the structures of non-twisting fibrils may illuminate the specific reasons these fibrils do not grow wider. In conclusion, our analysis suggests that fibril width is controlled by a multitude of factors, including helicity, the amyloidogenicity of the protein sequence, and the outward facing surfaces of the fibril.

Understanding the rules that determine the width of fibrils may enable the design of fibrils of arbitrary width. Here, we have shown that the twist of the fibril plays an important role in controlling the fibril width; thus, methods to control the twist may present a viable path to controlling fibril width. For instance, the use of alternating L- and D-amino acids may lead to fibril backbones that do not twist, therefore forming non-twisting fibrils that could have a larger width. It has already been shown that fibrils have comparable mechanical properties to that of steel and silk^{32,33}. The addition of more amino acids to a given fibril core would add to the network of interactions that stabilize fibrils and could potentially increase the mechanical stability of amyloid fibrils. Furthermore, fine-tuning the width of a fibril may affect its

assembly/disassembly kinetics which may be useful for drug delivery methods that rely on controlled release of monomers from fibrils³⁴. At present, it is not possible to rationally encode a fibril structure from a designed amino acid sequence, although meta-analyses such as the one we conducted here may hold clues to learning the design principles of amyloid fibrils, possibly enabling structure prediction and/or structure design as has been accomplished with globular proteins.

Author Contributions

D.R.B conceived of and oversaw the project. D.R.B., N.A.M., and M.R.S wrote the software for computational analysis. D.R.B. wrote the paper with input from all authors.

Acknowledgements

D.R.B was supported by the National Science Foundation Graduate Research Fellowship Program and the UCLA Graduate Division Dissertation Year Fellowship.

Material and Methods

Preparation of pdb and map files

All pdb and map files were downloaded from the Protein Data Bank (See Supplementary Table 2 for full list). In order to allow for easy computation, we curated the files as follows. We moved all fibrils so that their helical axis was coincident with the $z=0$ axis. We also removed all alternate rotamers. For many fibrils, we observed that the pdb file did not follow the exact symmetry of the map file. This is likely due to the fact that refinement programs such as phenix.real_space_refine, although they do use non-crystallographic symmetry, do not impose helical symmetry on the refined pdb. Therefore, we used the helical twist and rise values from the refined map for each structure, as listed from the corresponding publication and helicized the model before our analysis using pdbset.

For the analysis of the pitch and radius of the Pick's Disease Wide Fibril used in Figure 5, we downloaded the Wide Fibril map and Narrow Fibril model from the PDB and built a hypothetical Wide Fibril model as the authors propose in Falcon, et al. by rigid-body fitting two

Narrow Fibrils into the Wide Fibril map²³. We measured the maximum radius of the theoretical Wide Fibril and calculated its pitch using the Wide Fibril helical parameters from Falcon, et al.

Q-score analysis

We measured the Q-score for all maps and models after moving the fibril axis to the origin and helicizing all models according to their published helical parameters. We converted Q-scores into expected resolution by the equation $Q\text{-score} = -0.178 * (\text{Expected Resolution}) + 1.119$ from ref.¹² (Supplementary Table 1). We plotted the reported resolution from the publication versus the expected resolution from the Q-score in Supplementary Figure 1. This indicated that most structures have a better expected resolution than reported in the publication, possibly suggesting the gold-standard FSC slightly underestimates the resolution for amyloid fibril reconstructions or that the Q-score analysis needs to be re-calibrated for amyloid fibril structures compared to globular proteins.

Computational analysis

We wrote a FORTRAN program that measures the maximum distance from the helical axis for every atom in a fibril in order to create Figure 5. Pitch values for Figure 5 were calculated from the twist and rise values in the publications. In order to calculate the $i, i+1$ distance between symmetrically related atoms, we wrote a Python program that identified $i, i+1$ chains (symmetrically related chains immediately above or below one another in the helix) and calculated the distance between identical atoms in the two chains, as well as the distance from the helical axis for those atoms. In order to calculate the backbone-backbone hydrogen bonding distance, we added hydrogens to all fibril structures using phenix.reduce²⁰. We then wrote a Python 3 program to identify all backbone-backbone hydrogen bonds between amide hydrogens and carbonyl hydrogens with distance greater than 1.6 Å and less than 2.6 Å. We also measured

the off-axis tilt for each hydrogen bond by measuring the magnitude of the carbonyl C-O bond vector projected onto the X-Y plane. The arccos of this value divided by the magnitude of the C-O bond vector in all three dimensions, gives the complementary angle to the bond tilt relative to the fibril axis. We further calculated the Ramachandran angles of all residues in the fibrils so that we could explicitly examine β -sheet hydrogen bonding – where both hydrogen bond donor and acceptor are in β -strand conformations.

Code availability

All code used for calculating fibril radii, interstrand distances, hydrogen bond distances and tilts, centering cryo-EM maps and models, generating hydrogen bond distance and angle restraints for real space refinement, and helicizing models will be made available on a GitHub repository upon publication of the peer-reviewed paper associated with this dissertation chapter.

References

1. Eisenberg, D. & Jucker, M. The Amyloid State of Proteins in Human Diseases. *Cell* **148**, 1188–1203 (2012).
2. Otzen, D. & Riek, R. Functional Amyloids. *Cold Spring Harb Perspect Biol* **11**, a033860 (2019).
3. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B. & Blake, C. C. F. Common core structure of amyloid fibrils by synchrotron X-ray diffraction¹¹ Edited by F. E. Cohen. *Journal of Molecular Biology* **273**, 729–739 (1997).
4. Eisenberg, D. S. & Sawaya, M. R. Structural Studies of Amyloid Proteins at the Molecular Level. *Annual Review of Biochemistry* **86**, 69–95 (2017).

5. Warmack, R. A., Boyer, D. R., Zee, C.-T., Richards, L. S., Sawaya, M. R., Cascio, D., Gonen, T., Eisenberg, D. S. & Clarke, S. G. Structure of amyloid- β (20-34) with Alzheimer's-associated isomerization at Asp23 reveals a distinct protofilament interface. *Nat Commun* **10**, 3357 (2019).
6. Lu, J.-X., Qiang, W., Yau, W.-M., Schwieters, C. D., Meredith, S. C. & Tycko, R. Molecular Structure of β -Amyloid Fibrils in Alzheimer's Disease Brain Tissue. *Cell* **154**, 1257–1268 (2013).
7. Chothia, C. Conformation of twisted β -pleated sheets in proteins. *Journal of Molecular Biology* **75**, 295–302 (1973).
8. Chou, K.-C. & SCHERAGA, H. A. Origin of the right-handed twist of α 3-sheets of poly(LVal) chains. *Proc. Natl. Acad. Sci. USA* **5** (1982).
9. Cowen, L., Bradley, P., Menke, M., King, J. & Berger, B. Predicting the Beta-Helix Fold from Protein Sequence Data. *Journal of Computational Biology* **9**, 261–276 (2002).
10. Seidler, P. M., Boyer, D. R., Murray, K. A., Yang, T. P., Bentzel, M., Sawaya, M. R., Rosenberg, G., Cascio, D., Williams, C. K., Newell, K. L., Ghetti, B., DeTure, M. A., Dickson, D. W., Vinters, H. V. & Eisenberg, D. S. Structure-based inhibitors halt prion-like seeding by Alzheimer's disease-and tauopathy-derived brain tissue samples. *J Biol Chem* **294**, 16451–16464 (2019).
11. Wu, X., Ma, Y., Zhao, K., Zhang, J., Sun, Y., Li, Y., Dong, X., Hu, H., Liu, J., Wang, J., Zhang, X., Li, B., Wang, H., Li, D., Sun, B., Lu, J. & Liu, C. The structure of a minimum amyloid fibril core formed by necroptosis-mediating RHIM of human RIPK3. *PNAS* **118**, (2021).

12. Pintilie, G., Zhang, K., Su, Z., Li, S., Schmid, M. F. & Chiu, W. Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat Methods* **17**, 328–334 (2020).
13. Boyer, D. R., Li, B., Sun, C., Fan, W., Zhou, K., Hughes, M. P., Sawaya, M. R., Jiang, L. & Eisenberg, D. S. The α -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. *Proc Natl Acad Sci U S A* **117**, 3592–3602 (2020).
14. Schweighauser, M., Shi, Y., Tarutani, A., Kametani, F., Murzin, A. G., Ghetti, B., Matsubara, T., Tomita, T., Ando, T., Hasegawa, K., Murayama, S., Yoshida, M., Hasegawa, M., Scheres, S. H. W. & Goedert, M. Structures of α -synuclein filaments from multiple system atrophy. *Nature* **585**, 464–469 (2020).
15. Falcon, B., Zivanov, J., Zhang, W., Murzin, A. G., Garringer, H. J., Vidal, R., Crowther, R. A., Newell, K. L., Ghetti, B., Goedert, M. & Scheres, S. H. W. Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019).
16. Gallardo, R., Iadanza, M. G., Xu, Y., Heath, G. R., Foster, R., Radford, S. E. & Ranson, N. A. Fibril structures of diabetes-related amylin variants reveal a basis for surface-templated assembly. *Nat Struct Mol Biol* **27**, 1048–1056 (2020).
17. Kollmer, M., Close, W., Funk, L., Rasmussen, J., Bsoul, A., Schierhorn, A., Schmidt, M., Sigurdson, C. J., Jucker, M. & Fändrich, M. Cryo-EM structure and polymorphism of A β amyloid fibrils purified from Alzheimer's brain tissue. *Nat Commun* **10**, 4760 (2019).
18. Li, B., Ge, P., Murray, K. A., Sheth, P., Zhang, M., Nair, G., Sawaya, M. R., Shin, W. S., Boyer, D. R., Ye, S., Eisenberg, D. S., Zhou, Z. H. & Jiang, L. Cryo-EM of full-length α -synuclein reveals fibril polymorphs with a common structural kernel. *Nat Commun* **9**, (2018).

19. Wang, L.-Q., Zhao, K., Yuan, H.-Y., Wang, Q., Guan, Z., Tao, J., Li, X.-N., Sun, Y., Yi, C.-W., Chen, J., Li, D., Zhang, D., Yin, P., Liu, C. & Liang, Y. Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. *Nat Struct Mol Biol* **27**, 598–602 (2020).
20. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861–877 (2019).
21. Goldschmidt, L., Teng, P. K., Riek, R. & Eisenberg, D. Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci U S A* **107**, 3487–3492 (2010).
22. Fitzpatrick, A. W. P., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., Crowther, R. A., Ghetti, B., Goedert, M. & Scheres, S. H. W. Cryo-EM structures of tau filaments from Alzheimer’s disease. *Nature* **547**, 185–190 (2017).
23. Falcon, B., Zhang, W., Murzin, A. G., Murshudov, G., Garringer, H. J., Vidal, R., Crowther, R. A., Ghetti, B., Scheres, S. H. W. & Goedert, M. Structures of filaments from Pick’s disease reveal a novel tau protein fold. *Nature* **561**, 137 (2018).
24. Arakhamia, T., Lee, C. E., Carlomagno, Y., Duong, D. M., Kundinger, S. R., Wang, K., Williams, D., DeTure, M., Dickson, D. W., Cook, C. N., Seyfried, N. T., Petrucelli, L. & Fitzpatrick, A. W. P. Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. *Cell* **180**, 633–644.e12 (2020).

25. Zhang, W., Falcon, B., Murzin, A. G., Fan, J., Crowther, R. A., Goedert, M. & Scheres, S. H. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife* **8**, (2019).
26. Zhang, W., Tarutani, A., Newell, K. L., Murzin, A. G., Matsubara, T., Falcon, B., Vidal, R., Garringer, H. J., Shi, Y., Ikeuchi, T., Murayama, S., Ghetti, B., Hasegawa, M., Goedert, M. & Scheres, S. H. W. Novel tau filament fold in corticobasal degeneration. *Nature* **580**, 283–287 (2020).
27. Murray, D. T., Kato, M., Lin, Y., Thurber, K. R., Hung, I., McKnight, S. L. & Tycko, R. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* **171**, 615-627.e16 (2017).
28. Lee, M., Ghosh, U., Thurber, K. R., Kato, M. & Tycko, R. Molecular structure and interactions within amyloid-like fibrils formed by a low-complexity protein sequence from FUS. *Nat Commun* **11**, 5735 (2020).
29. Liberta, F., Loerch, S., Rennegarbe, M., Schierhorn, A., Westermark, P., Westermark, G. T., Hazenberg, B. P. C., Grigorieff, N., Fändrich, M. & Schmidt, M. Cryo-EM fibril structures from systemic AA amyloidosis reveal the species complementarity of pathological amyloids. *Nature Communications* **10**, 1104 (2019).
30. Guerrero-Ferreira, R., Taylor, N. M., Arteni, A.-A., Kumari, P., Mona, D., Ringler, P., Britschgi, M., Lauer, M. E., Makky, A., Verasdonck, J., Riek, R., Melki, R., Meier, B. H., Böckmann, A., Bousset, L. & Stahlberg, H. Two new polymorphic structures of human full-length alpha-synuclein fibrils solved by cryo-electron microscopy. *eLife* **8**, e48907 (2019).
31. Zhao, K., Lim, Y.-J., Liu, Z., Long, H., Sun, Y., Hu, J.-J., Zhao, C., Tao, Y., Zhang, X., Li, D., Li, Y.-M. & Liu, C. Parkinson's disease-related phosphorylation at Tyr39 rearranges α -

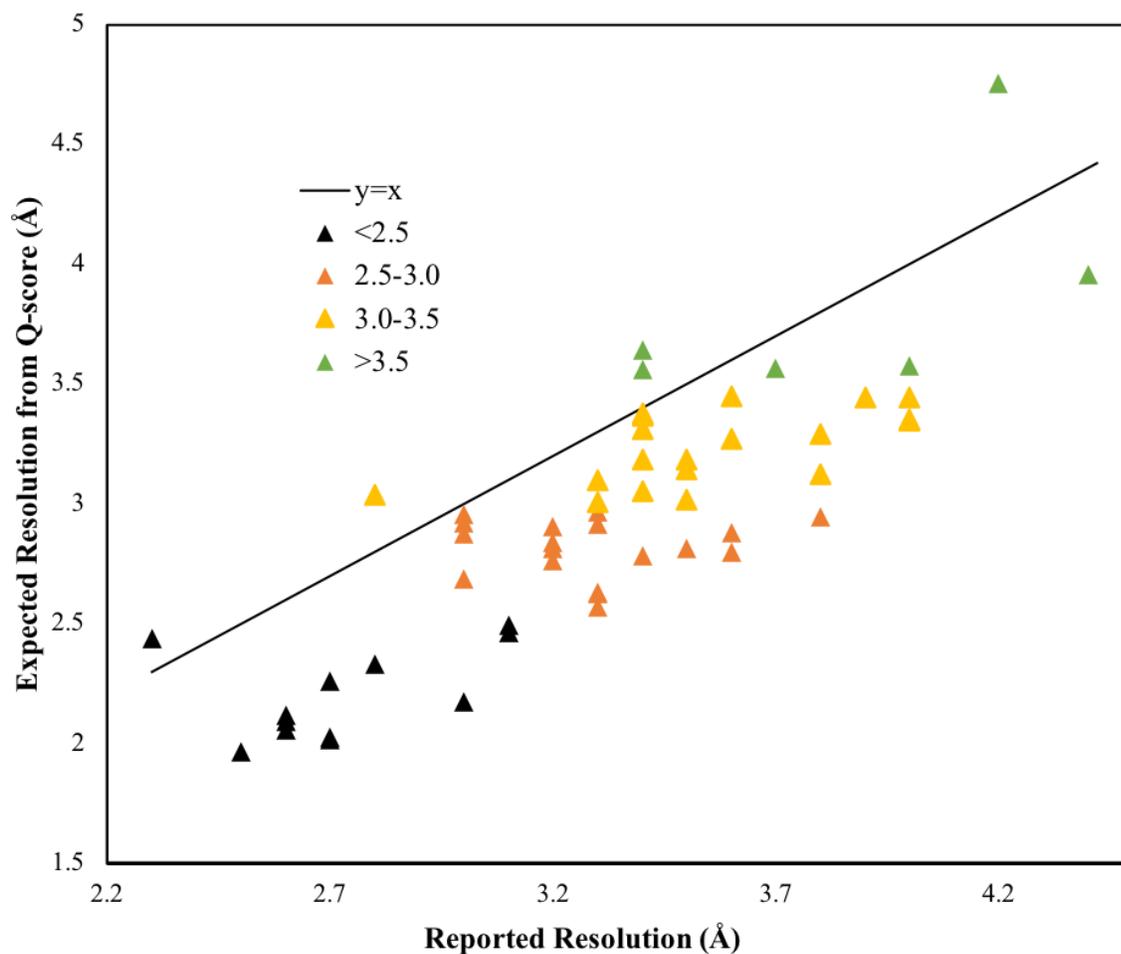
synuclein amyloid fibril structure revealed by cryo-EM. *Proc Natl Acad Sci U S A* **117**, 20305–20315 (2020).

32. Smith, J. F., Knowles, T. P. J., Dobson, C. M., Macphee, C. E. & Welland, M. E. Characterization of the nanoscale properties of individual amyloid fibrils. *Proc Natl Acad Sci U S A* **103**, 15806–15811 (2006).
33. Knowles, T. P., Fitzpatrick, A. W., Meehan, S., Mott, H. R., Vendruscolo, M., Dobson, C. M. & Welland, M. E. Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* **318**, 1900–1903 (2007).
34. Maji, S. K., Schubert, D., Rivier, C., Lee, S., Rivier, J. E. & Riek, R. Amyloid as a depot for the formulation of long-acting drugs. *PLoS Biol* **6**, e17 (2008).

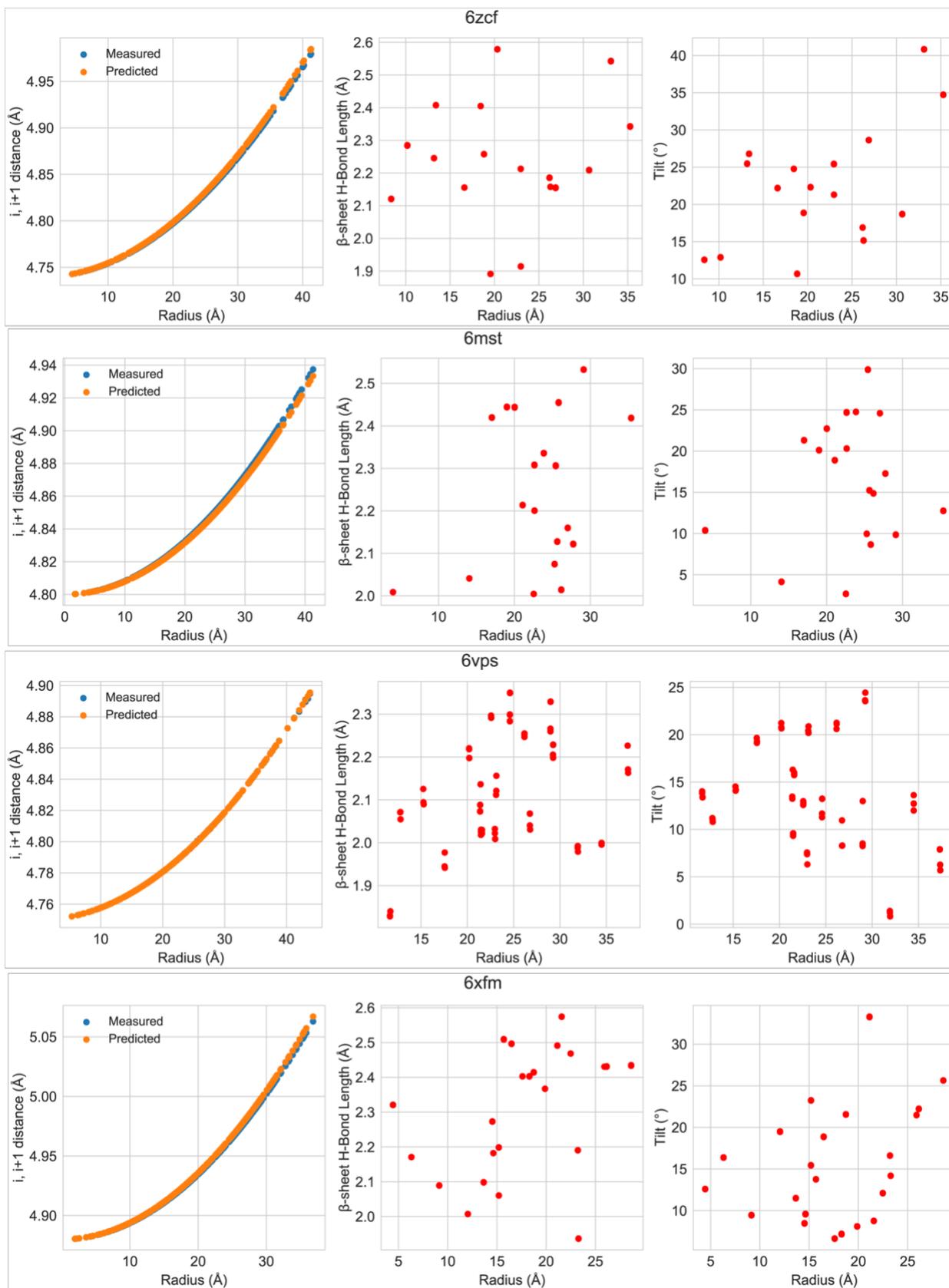
Supplementary Information

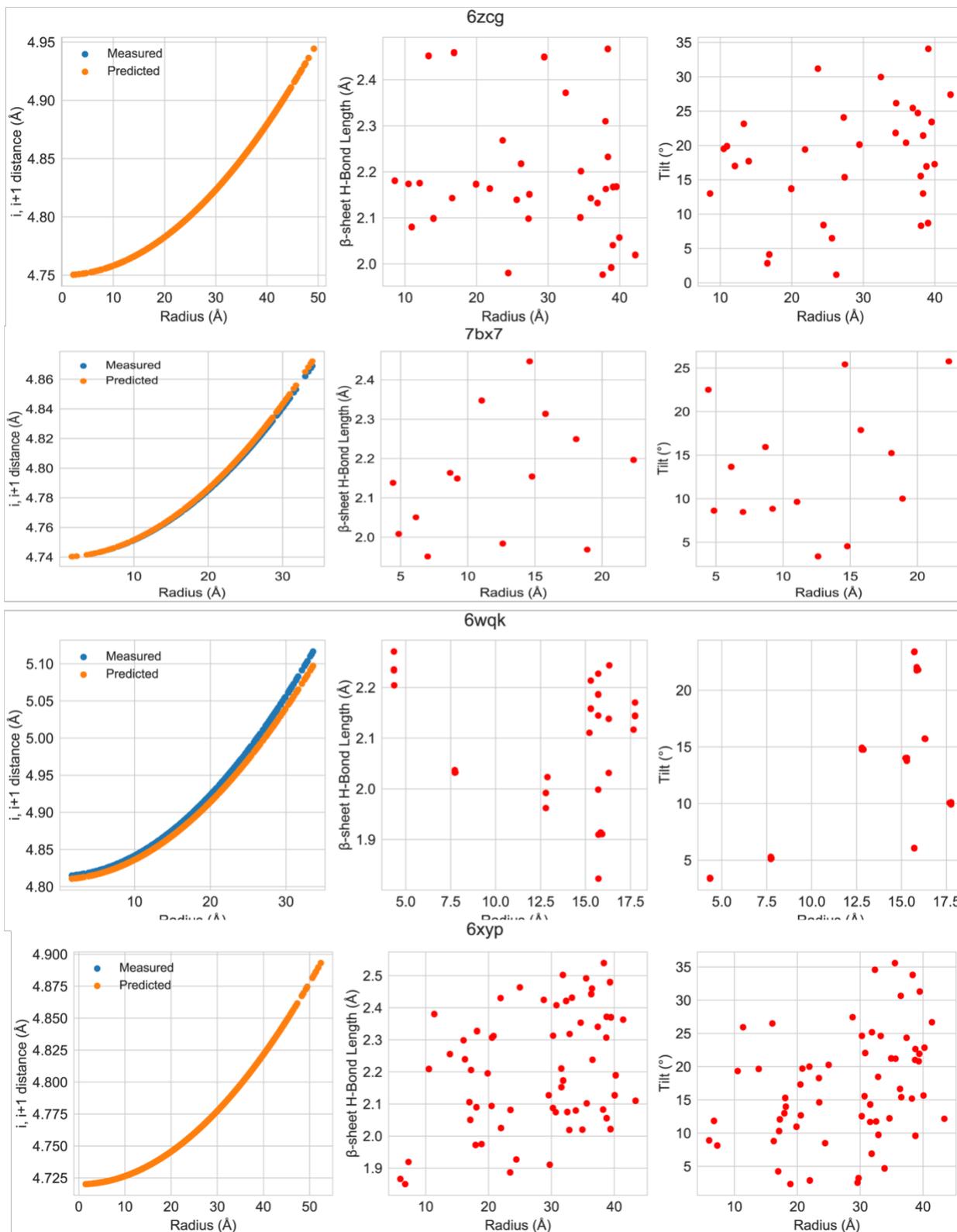
PDB	Reported Resolution (Å)	Q_score	Predicted Resolution (Å)	Pred. Res. Bin (Å)
6ufr	2.5	0.769	1.966292135	Res.<2.5
6zcf	2.7	0.76	2.016853933	
6mst	2.7	0.758	2.028089888	
6vps	2.6	0.753	2.056179775	
6xyo	2.6	0.747	2.08988764	
6xfm	2.6	0.742	2.117977528	
6zcg	3	0.732	2.174157303	
6lni	2.7	0.717	2.258426966	
7bx7	2.8	0.704	2.331460674	
6nwp	2.3	0.685	2.438202247	
6wqk	3.1	0.681	2.460674157	
6xyq	3.1	0.675	2.494382022	
6xyp	3.3	0.662	2.56741573	2.5<Res.<3.0
6n3c	3.3	0.652	2.623595506	
6peo	3.3	0.651	2.629213483	
6dso	3	0.641	2.685393258	
6gx5	3.2	0.627	2.764044944	
6llu	3.4	0.624	2.780898876	
6pes	3.6	0.621	2.797752809	
6uur	3.5	0.618	2.814606742	
6tjo	3.2	0.618	2.814606742	
6llt	3.2	0.614	2.837078652	
6ssx	3	0.607	2.876404494	
6zrf	3.6	0.606	2.882022472	
7kwz	3.2	0.602	2.904494382	
6qjm	3.3	0.6	2.915730337	
6tjx	3	0.599	2.921348315	
5w7v	3.8	0.595	2.943820225	
6sdz	3	0.593	2.95505618	
6qjh	3.3	0.591	2.966292135	
6n3a	3.3	0.584	3.005617978	3.0<Res.<3.5
6zch	3.5	0.582	3.016853933	
6w0o	2.8	0.578	3.039325843	
6l4s	3.4	0.576	3.050561798	
6ic3	3.3	0.568	3.095505618	
6n37	3.8	0.563	3.123595506	
6qjp	3.5	0.559	3.146067416	
6sst	3.4	0.552	3.185393258	
6lrq	3.5	0.552	3.185393258	
6cu7	3.6	0.537	3.269662921	
6n3b	3.8	0.533	3.292134831	
6r4r	3.4	0.53	3.308988764	
6hud	4	0.523	3.348314607	
6zrq	4	0.522	3.353932584	
6vw2	3.4	0.52	3.365168539	
5o3l	3.4	0.518	3.376404494	
6gk3	4	0.506	3.443820225	
6zrr	3.9	0.506	3.443820225	
6cu8	3.6	0.505	3.449438202	
6nwq	3.4	0.485	3.561797753	Res.>3.5
6qjq	3.7	0.484	3.56741573	
5oqv	4	0.483	3.573033708	
5o3t	3.4	0.471	3.640449438	
6shs	4.4	0.415	3.95505618	
6yla	4.2	0.273	4.752808989	

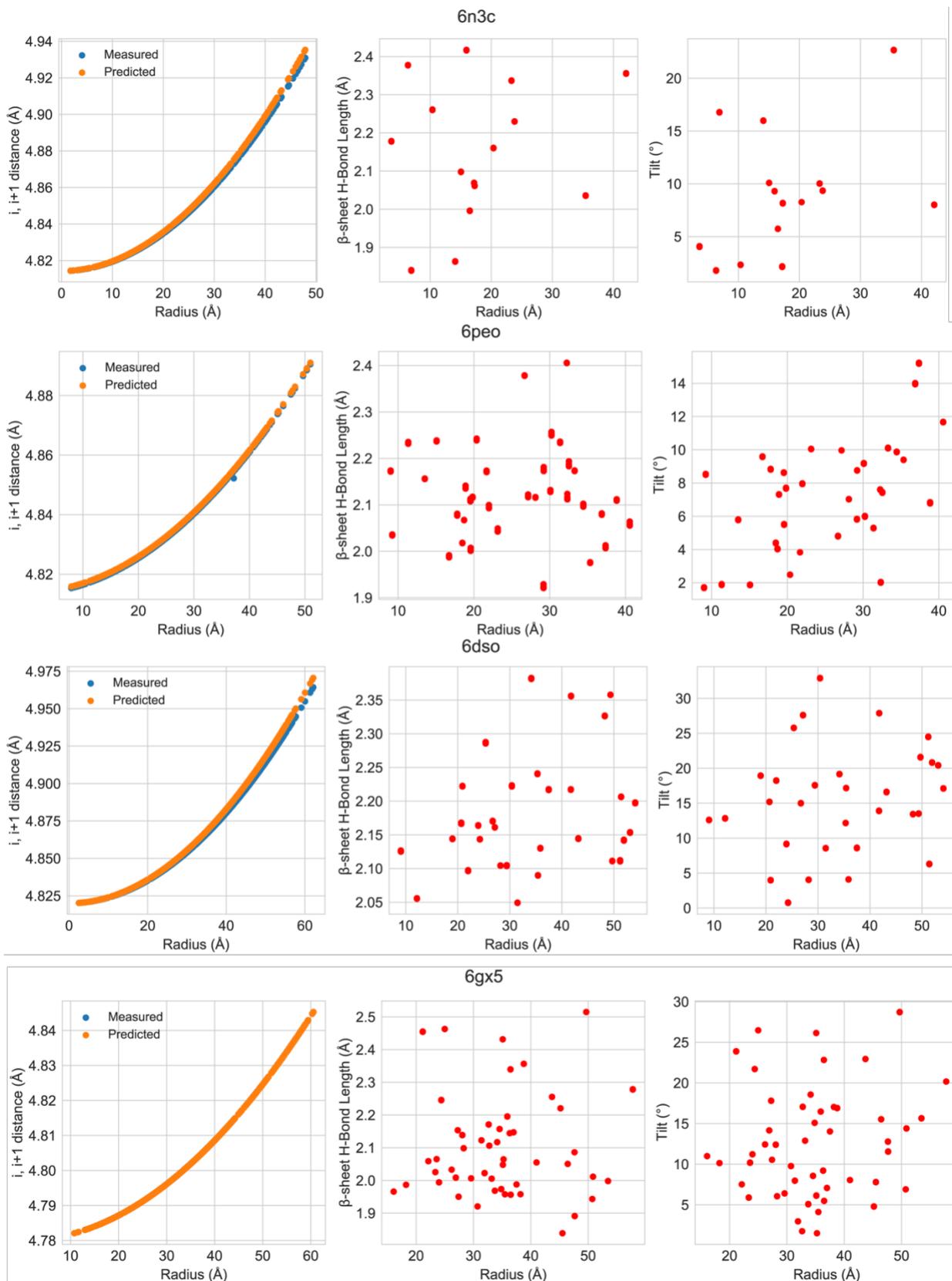
Supplementary Table 1 Q-score analysis of all 55 cryo-EM structures

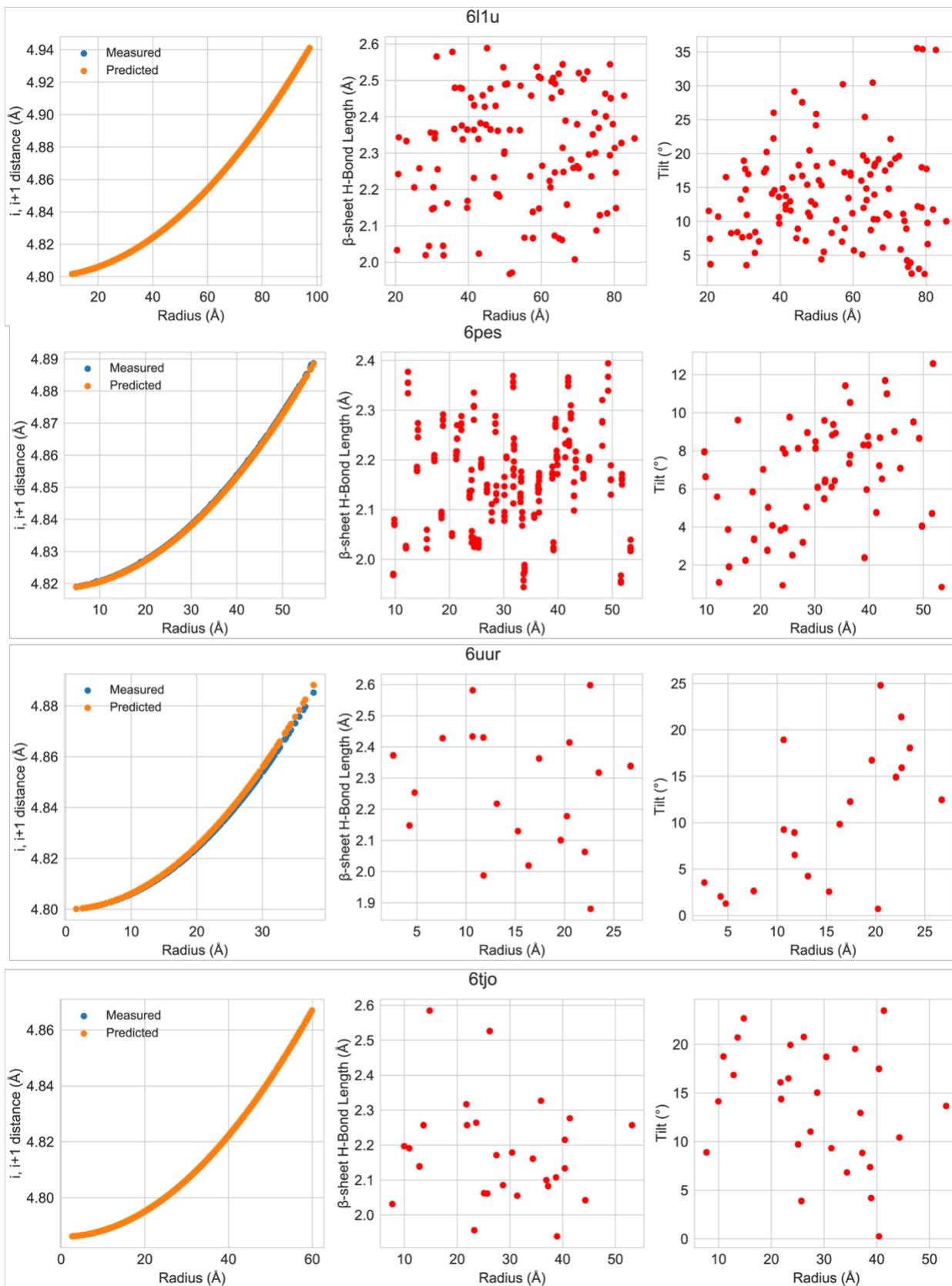


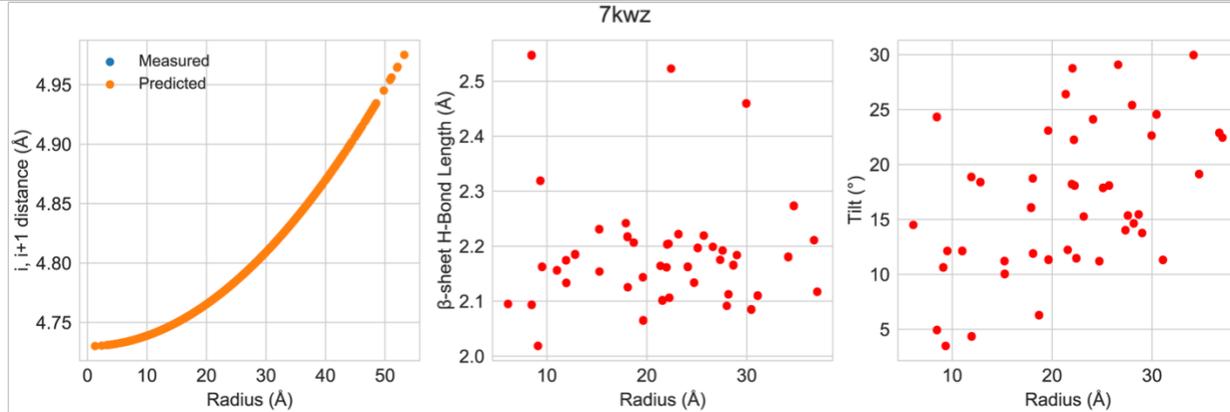
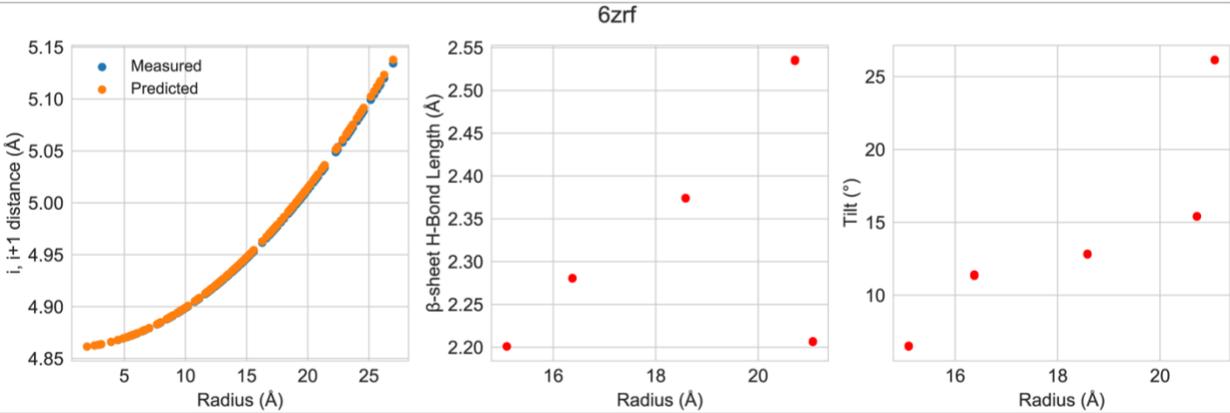
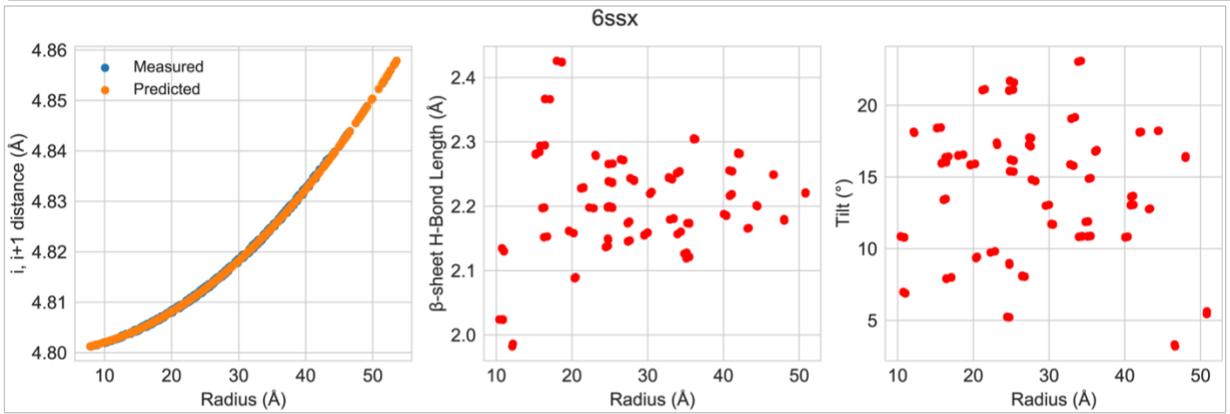
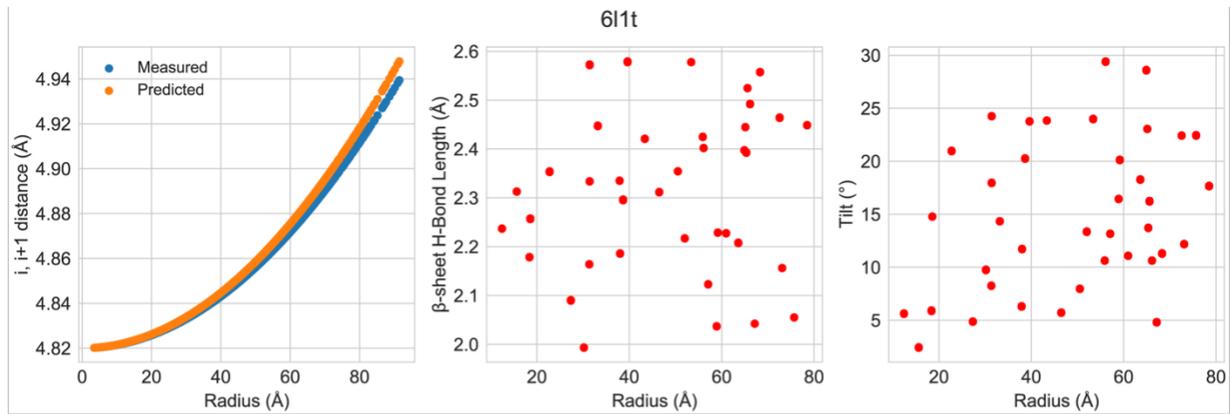
Supplementary Figure 1. Plot of reported resolution from publication versus expected resolution from Q-score analysis ($Q\text{-score} = -0.178 * (\text{Expected Resolution}) + 1.119$).

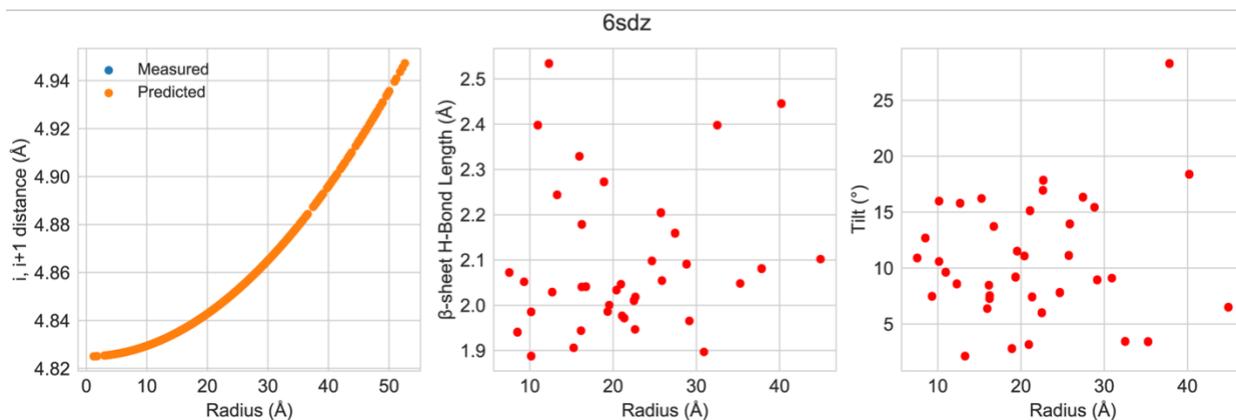
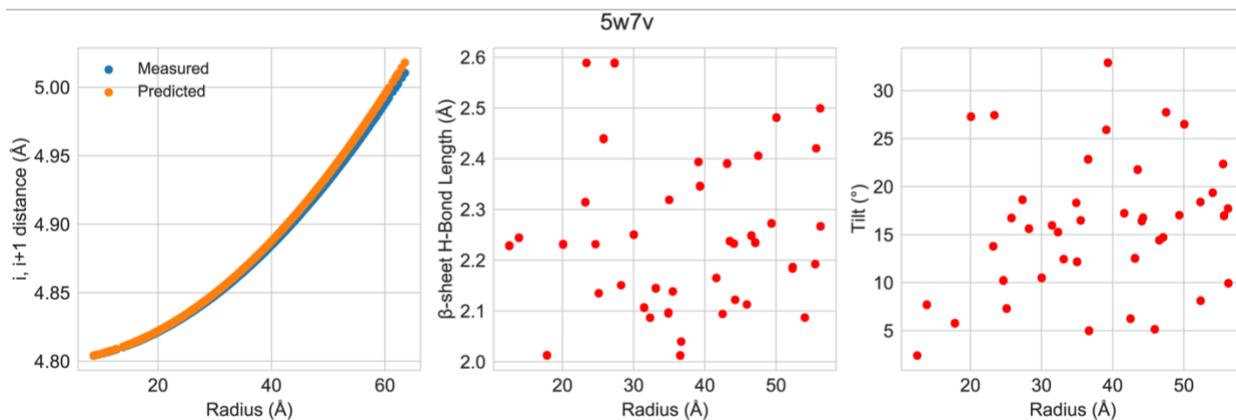
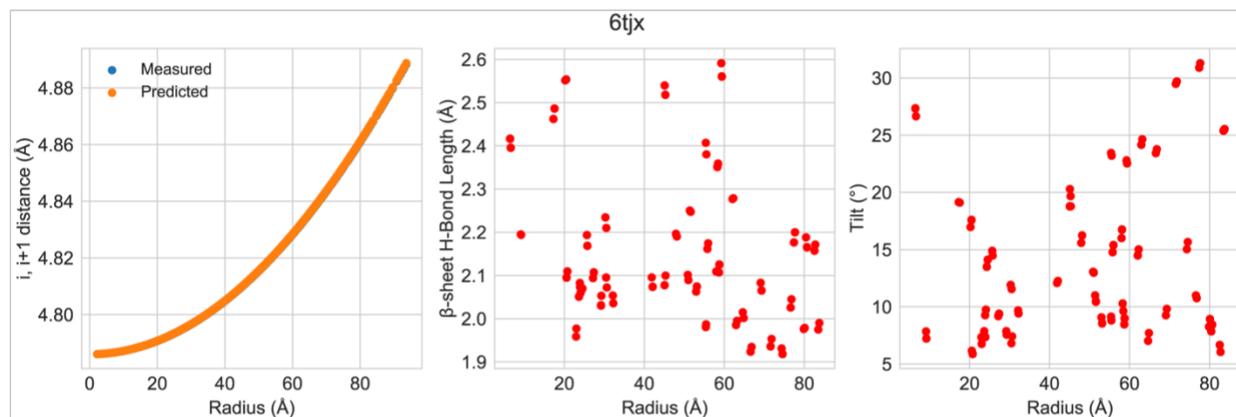
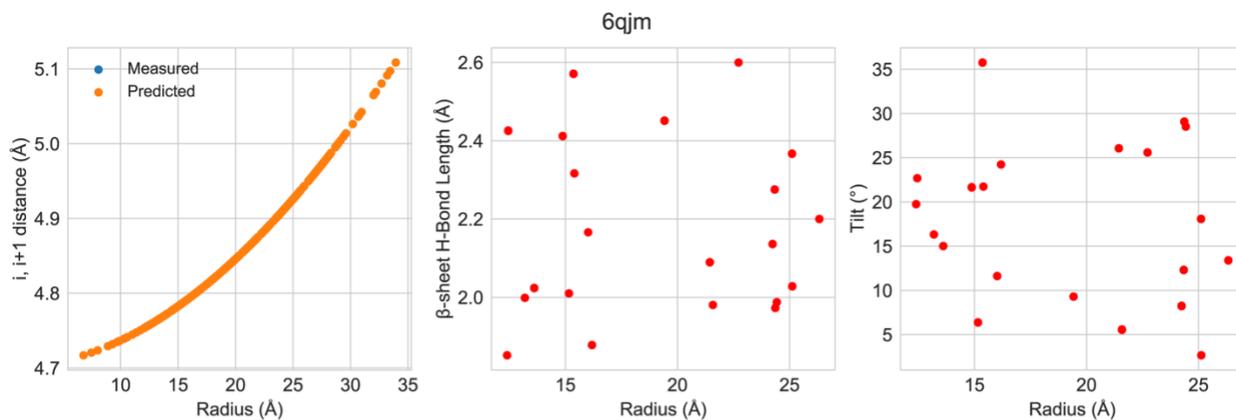


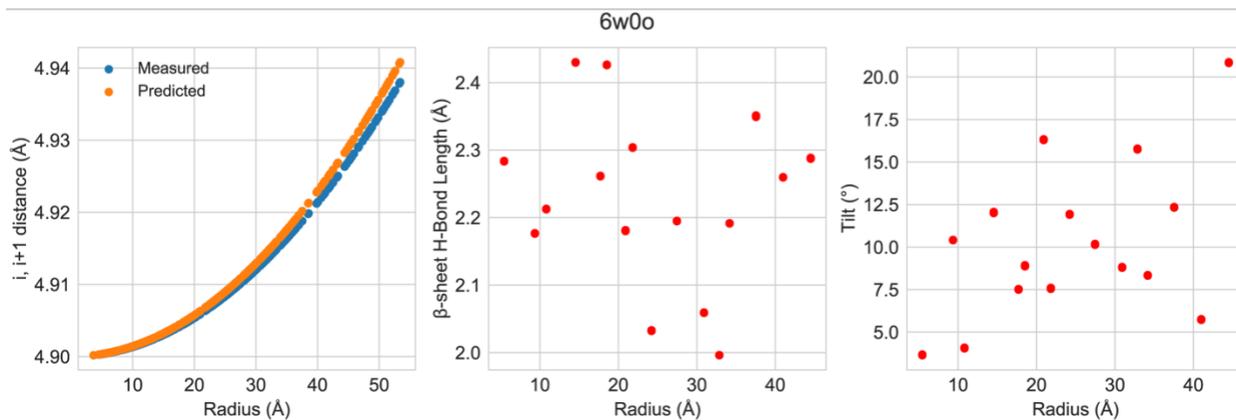
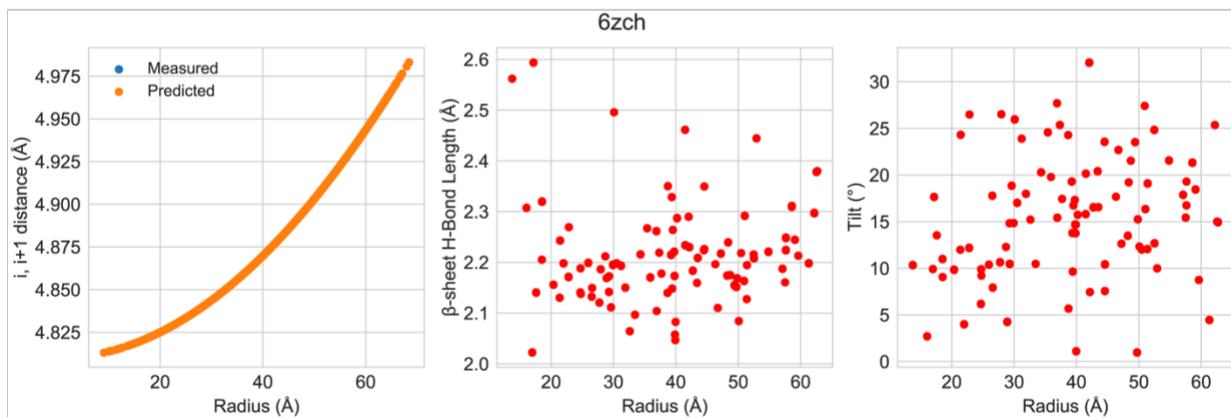
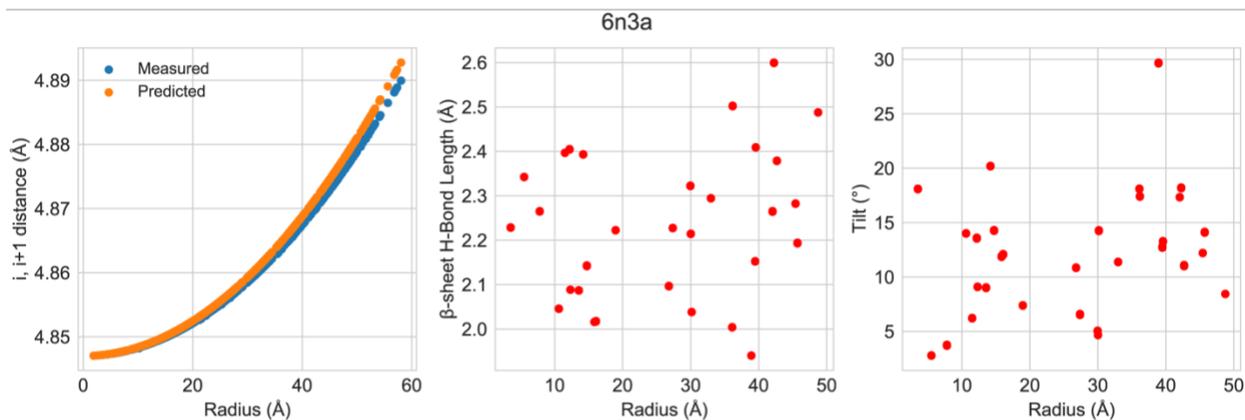
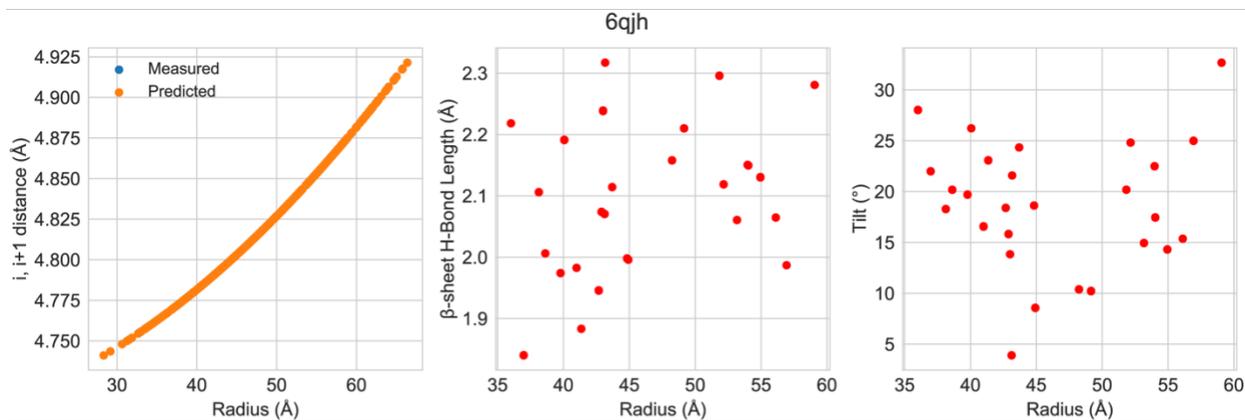


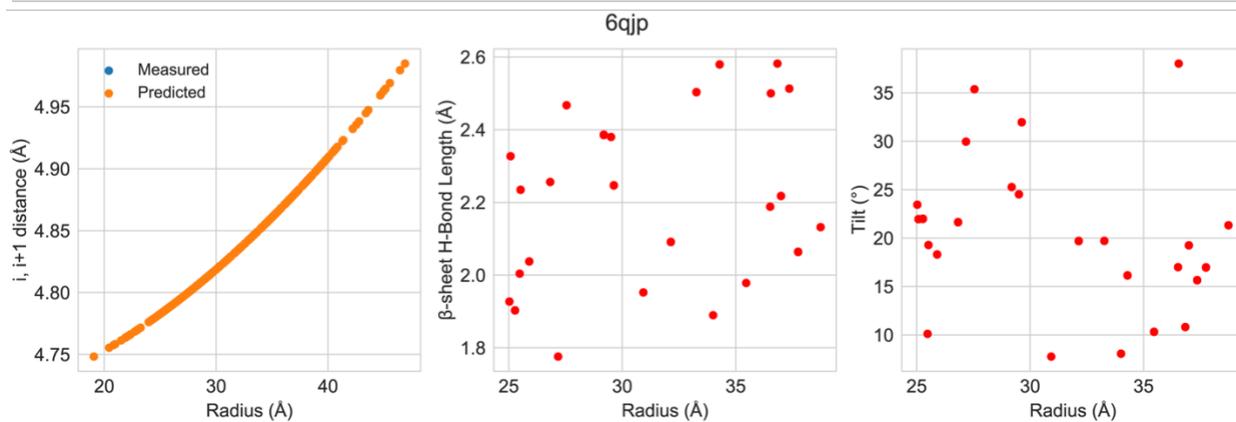
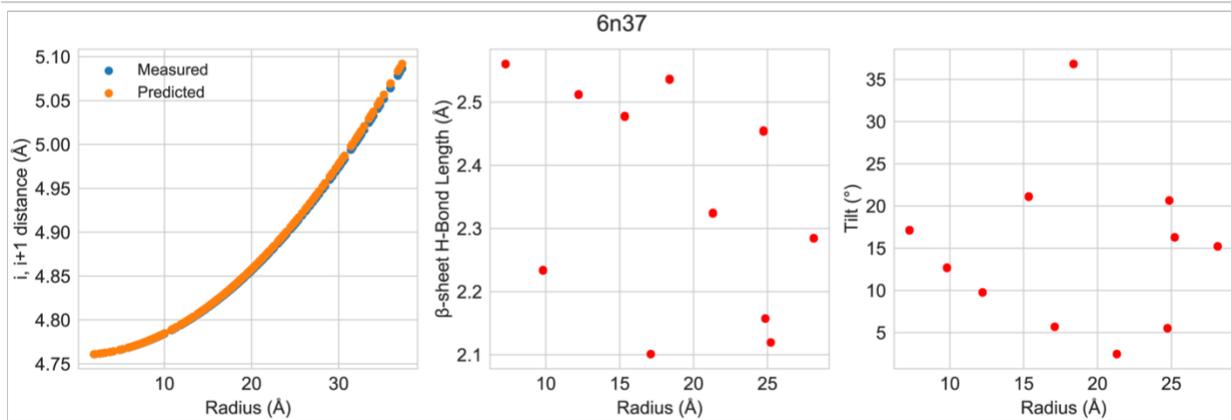
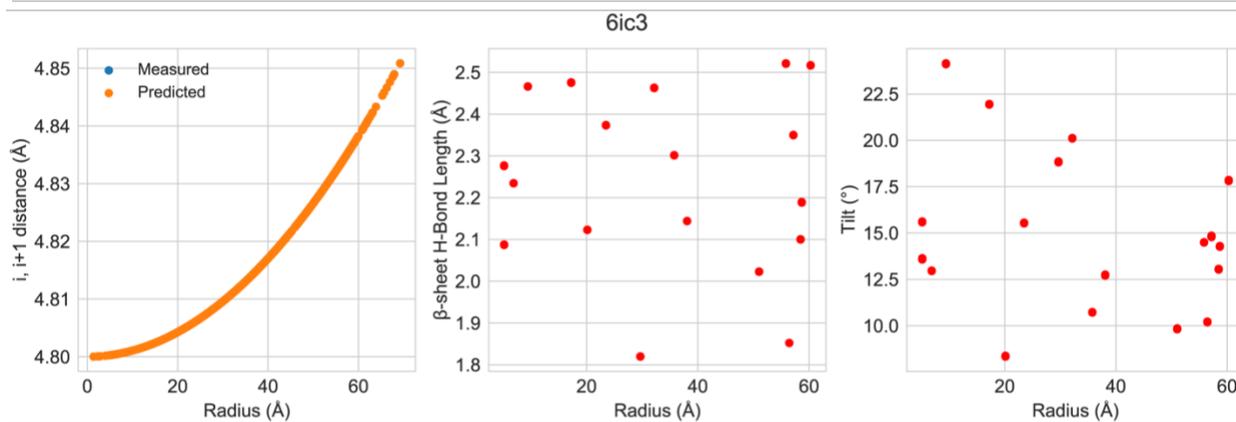
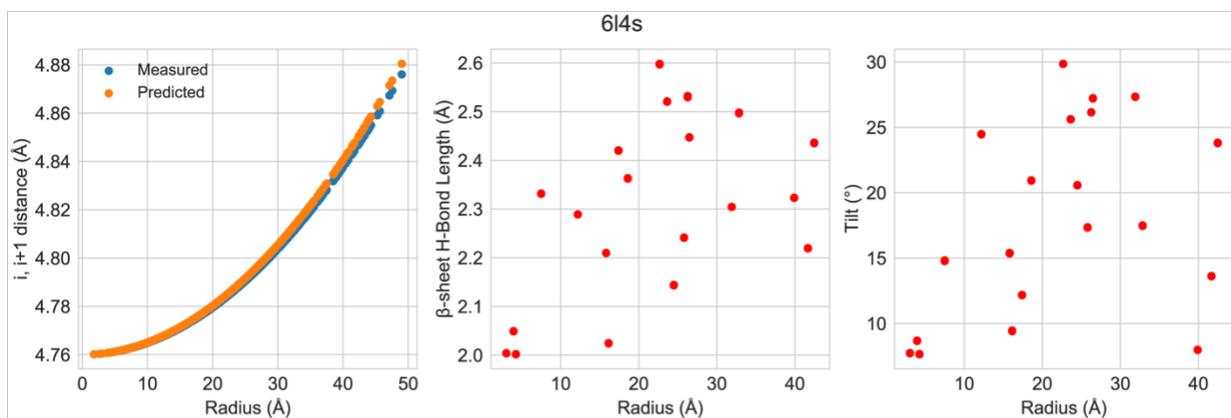


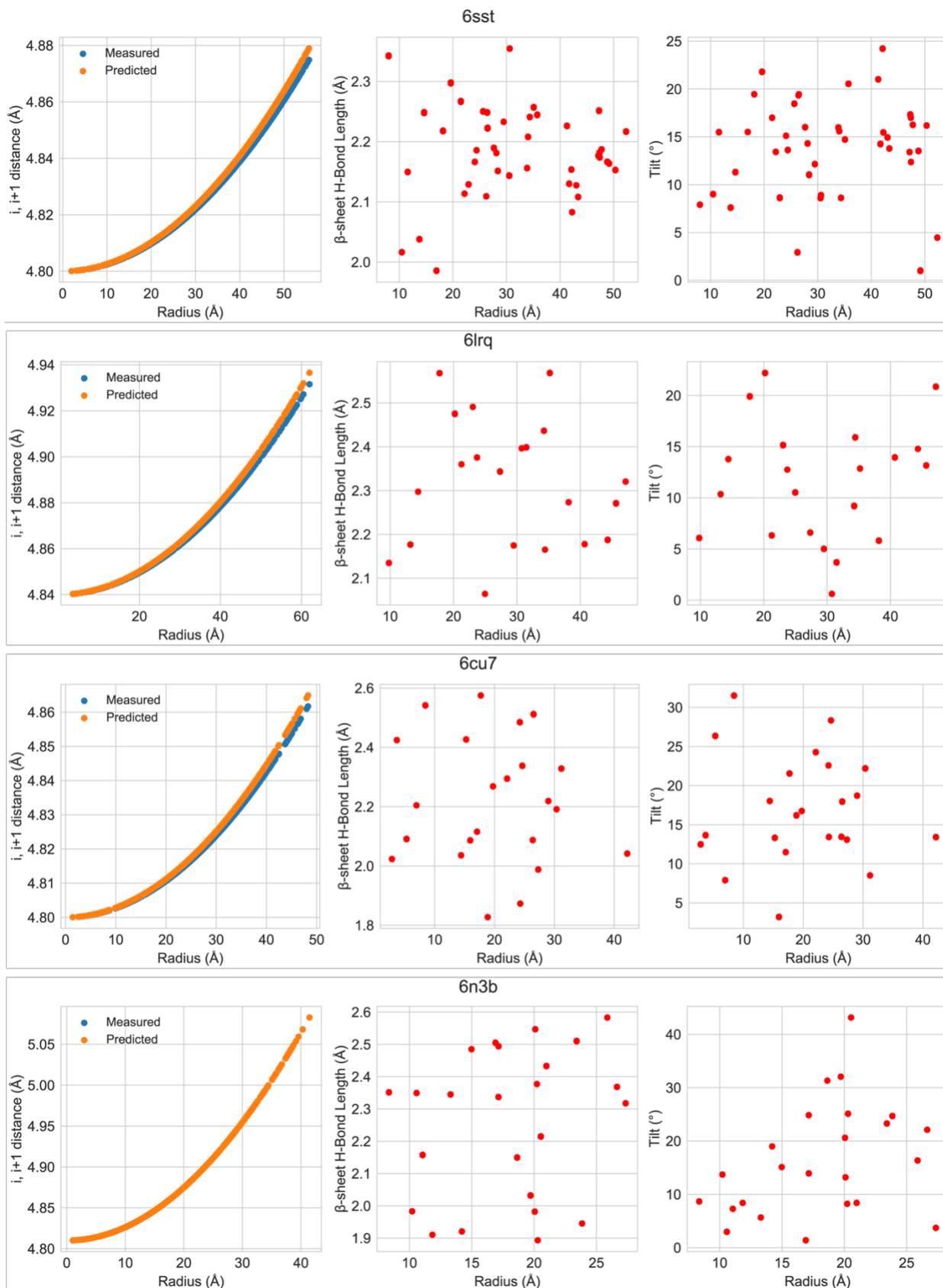


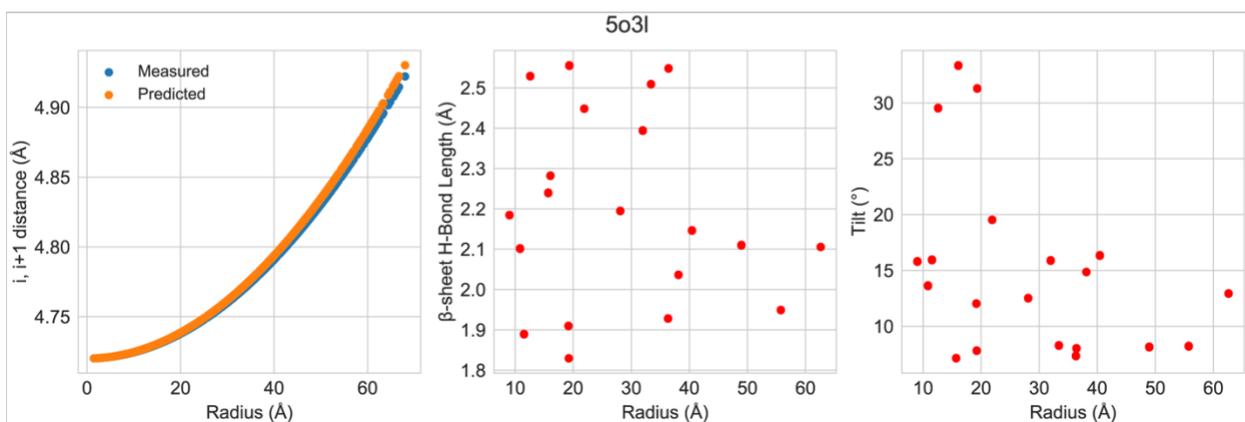
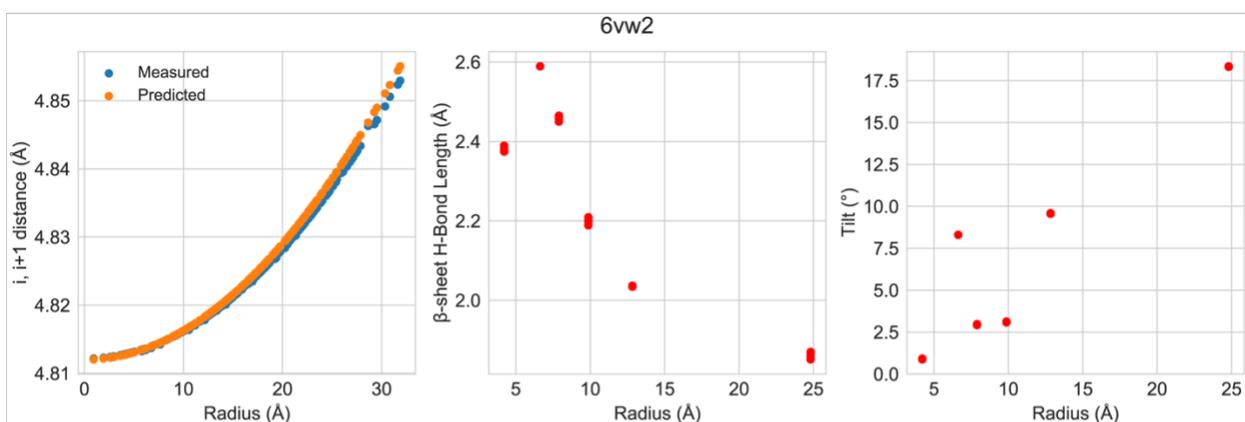
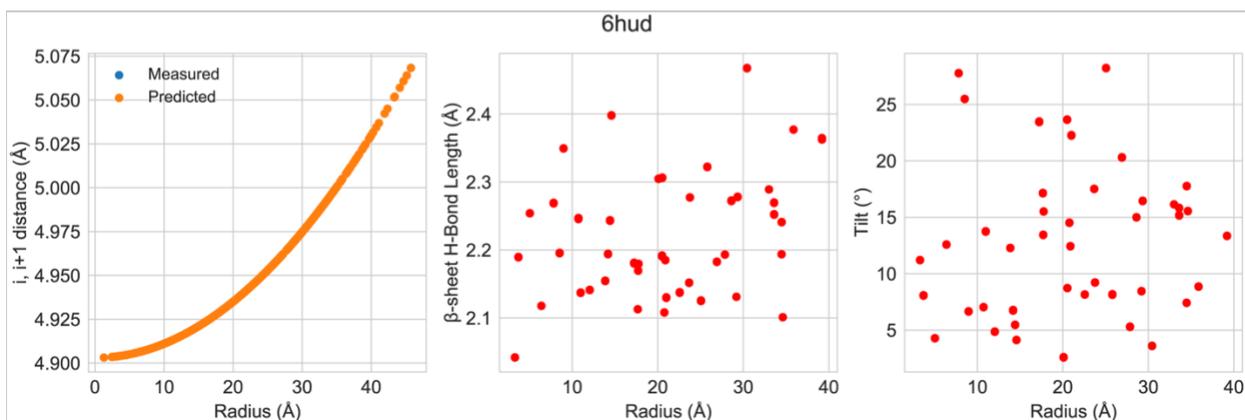
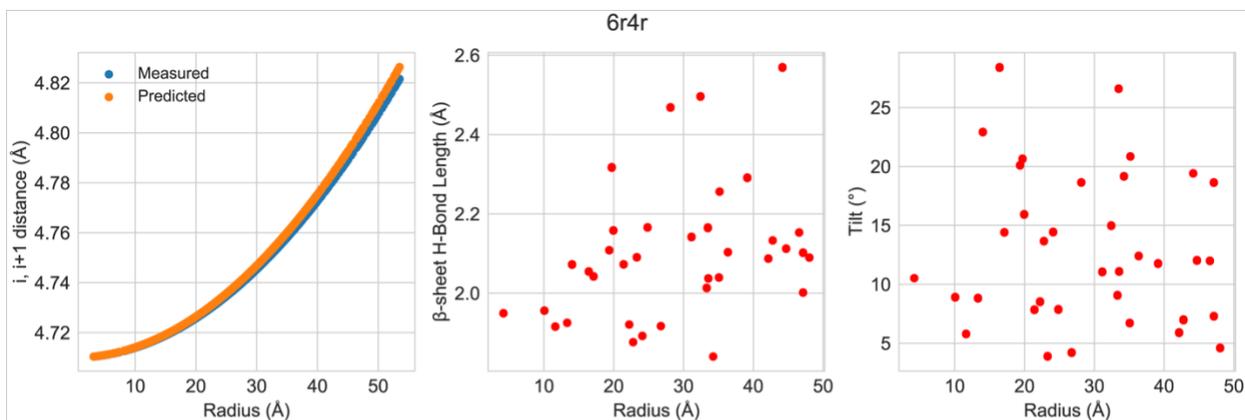


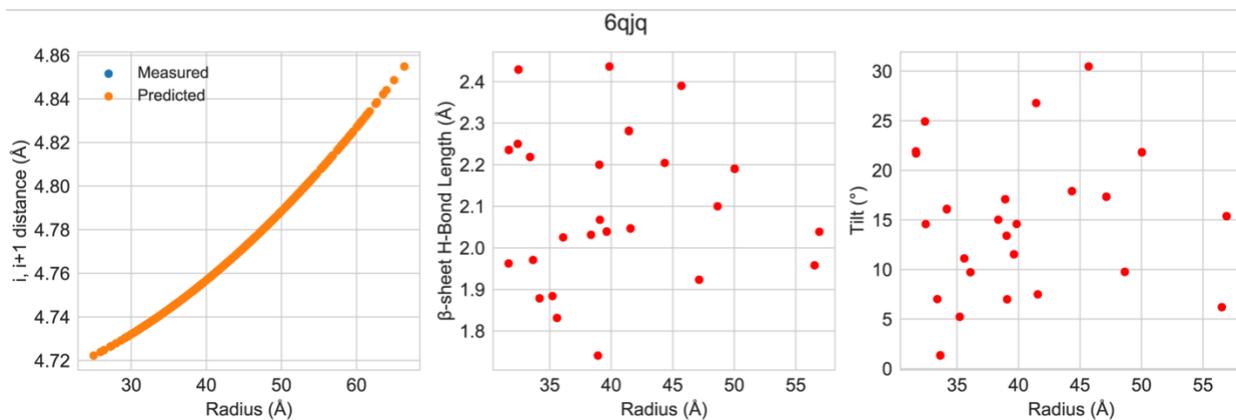
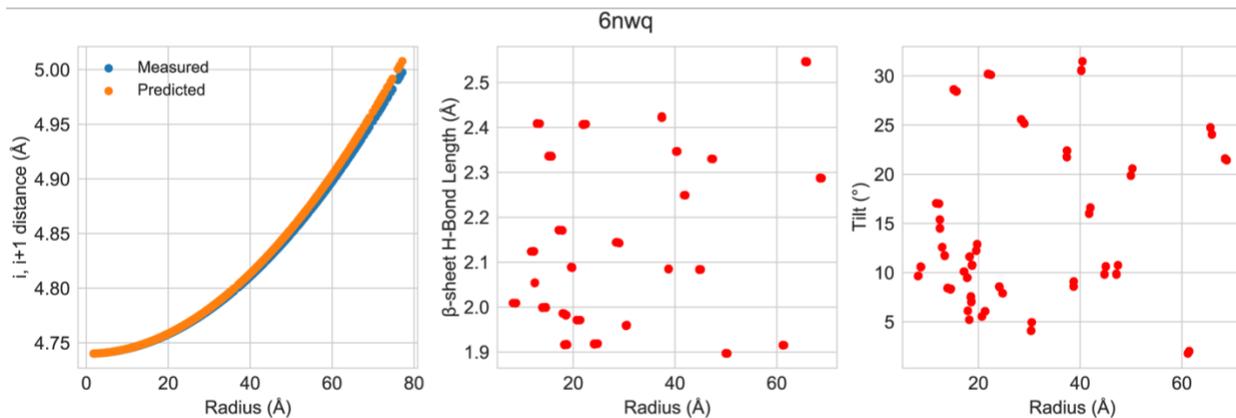
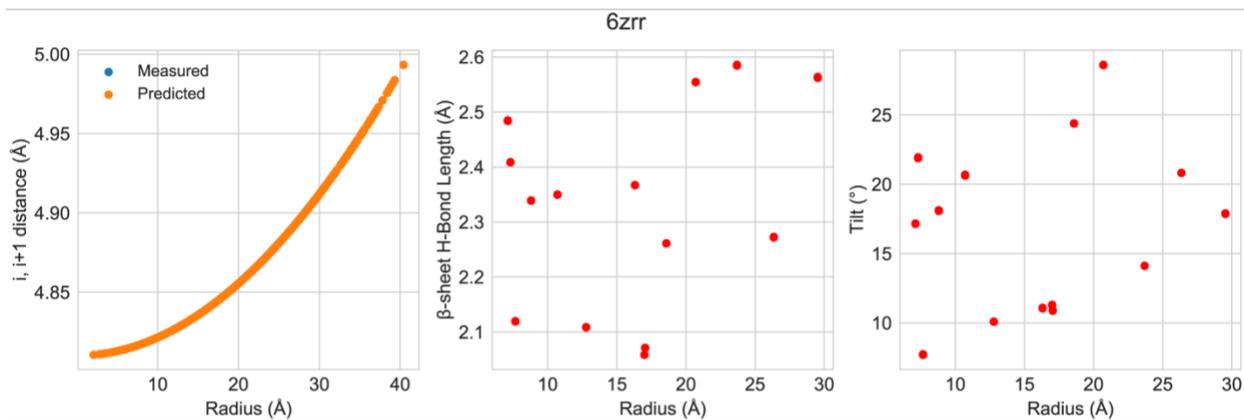
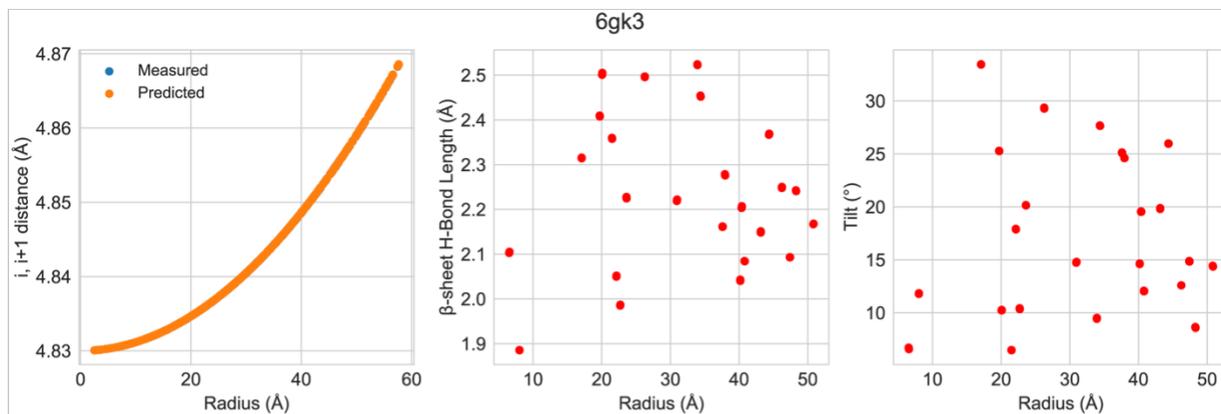


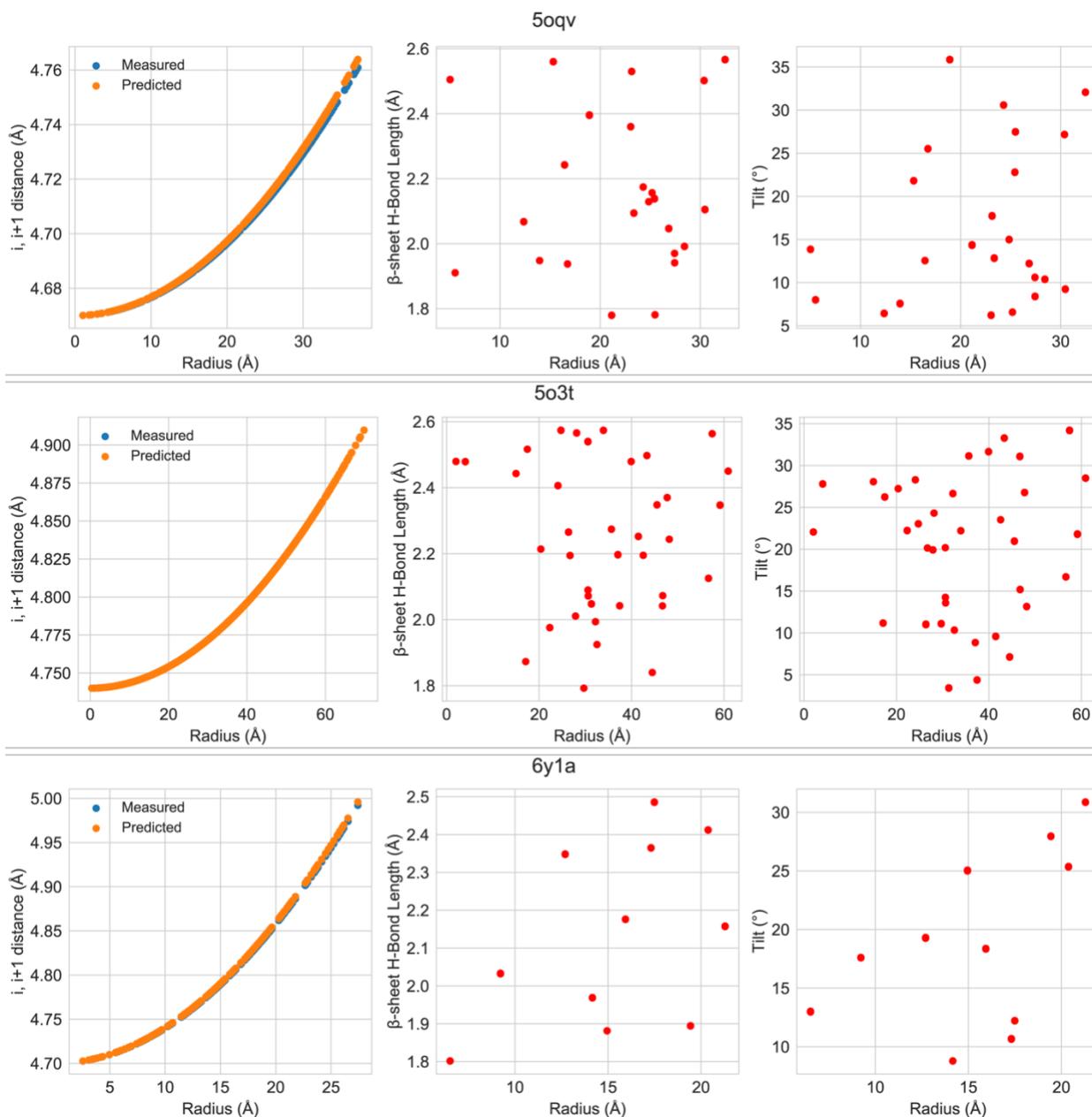












Supplementary Figure 2. Plots of fibril structures showing $i, i+1$ distance between symmetry related atoms, backbone hydrogen bond distances and off-axis tilts between $i, i+1$ β -strands.

PDB	# protofibrils	Citation	Ref.	Resolution	Symmetric	Point Group	Max Radius (Å)	Crossover (Å)	Pitch (Å)	Twist L-start (°)	Twist n-start (°)	Rise Å/subunit	Rise n-start (Å/subunit)
6w6o	2	https://dx.doi.org/10.1016/j.jmb.2019.08.011	1	2.8	Yes	C1	53.48	579.7	2594.1	179.66	0.68	2.45	4.9
A-beta 1-40 brain seeded eryoEM													
5oqv	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	2	4.0	Yes	C1	37	579.7	1159.4	179.275	1.45	2.35	4.67
A-beta 1-42													
6shs	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	3	4.4	Yes	C1	43	431.6	863.3	181.005	2.01	2.41	4.82
A-beta 1-40													
6euf	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	4	3.6	Yes	C1	49	919.1	1838.3	179.53	0.94	2.4	4.8
A-syn WT Rod													
6euh	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	4	3.6	Yes	C1	49	459.6	919.1	179.06	1.88	2.4	4.8
A-syn WT Twister													
6s8t	2	https://eifsc.de/doi/10.1002/eim.201900001	5	3.4	Yes	C1	56	960.0	1920.0	179.55	0.9	2.4	4.8
A-syn WT 2b													
6gk3	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	6	4.0	Yes	C2	58	1429.9	2859.9	-0.608	0.608	4.83	4.83
beta-2-microglobulin													
6yla	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	7	4.2	Yes	C1	28	239.0	478.0	-178.23	3.54	2.35	4.7
Islet Amyloid Polypeptide (IAPP)													
6vww	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	8	3.4	Yes	C1	32	746.7	1493.4	178.42	1.16	2.406	4.812
Islet Amyloid Polypeptide (IAPP)													
6zrf	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	9	3.6	Yes	C1	27	494.2	178.23	178.23	3.54	2.43	4.86
Islet Amyloid Polypeptide (IAPP) S20G 2PF													
6zrq	2	https://dx.doi.org/10.1016/j.jmb.2019.08.011	9	3.4	Yes	C1	34	456.6	913.3	179.05	1.9	2.41	4.82
Islet Amyloid Polypeptide (IAPP) S20G 2PF													
6vps	3	https://dx.doi.org/10.1016/j.jmb.2019.08.011	10	2.6	Yes	C3	44	367.7	1103.2	-1.55	1.55	4.75	4.75
Ohp2													
6d8o	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	11	3.0	Yes	C1	63	774.6	1549.3	179.44	1.12	2.41	4.82
Serum Amyloid A - mouse													
6mst	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	11	2.7	Yes	C1	42	546.8	1093.7	180.79	1.58	2.4	4.8
Serum Amyloid A - human													
6zcf	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	12	2.7	Yes	C1	41	398.7	797.4	178.93	1.14	2.37	4.74
Serum Amyloid A - mouse - in vitro													
6zcg	4	https://www.ncbi.nlm.nih.gov/pubmed/3113660	12	3.0	Yes	C2	49	534.4	1068.8	-1.6	1.6	4.75	4.75
Serum Amyloid A - mouse - in vitro													
6zcg	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	12	3.0	Yes	C2	49	534.4	1068.8	-1.6	1.6	4.75	4.75
Serum Amyloid A - mouse - in vitro													
6g4f	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	13	3.4	Yes	C1	54	751.6	1503.2	179.436	1.128	2.355	4.71
SH3 Domain of P18-kinase													
5o3l	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	14	3.4	Yes	C1	68	708.0	1416.0	179.4	1.2	2.36	4.72
tau AD PFF													
6nwp	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	15	2.3	Yes	C1	65	711.0	1422.0	179.4	1.2	2.37	4.74
tau CTE Type 1													
6nwx	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	15	3.4	Yes	C1	77	711.0	1422.0	179.4	1.2	2.37	4.74
tau CTE Type 2													
6nxx	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	16	3.0	Yes	C2	94	1412.3	2824.5	-0.61	0.61	4.786	4.786
Tau CBD Type 2													
5w7v	3	https://www.ncbi.nlm.nih.gov/pubmed/3113660	17	3.8	Yes	C1	64	436.4	1309.1	-120.44	1.32	1.6	4.8
TDP-43 (247-257)													
6m37	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	18	3.8	Yes	C1	38	308.2	616.4	178.61	2.78	2.38	4.76
TDP-43 (311-360) SegA-syn													
6m3a	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	18	3.3	Yes	C2	58	1323.9	2647.9	179.341006	0.658994	4.847049	4.847049
TDP-43 (311-360) SegA-slow													
6m3b	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	18	3.3	Yes	C2	58	666.6	1333.2	179.35	1.3	2.40717	4.81434
TDP-43 A315E (286-331) SegB													
N/A													
6uur	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	19	N/A	Yes	C2	114	1410.0	2820.0	-0.6	0.6	4.7	4.7
tau PID WPF													
6uur	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	19	3.5	Yes	C1	38	617.1	1234.3	179.3	1.4	2.4	4.8
hPFP 94178													
6uur	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	20	2.5	Yes	C2	48	806.4	1612.9	178.918583	1.081417	4.845015	4.845015
A-syn E46K													
6d4s	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	21	3.4	Yes	C1	49	680.0	1360.0	-179.37	1.26	2.38	4.76
A-syn ac-E46K													
6lq	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	22	3.5	Yes	C1	62	970.2	1940.3	179.551	0.898	2.42	4.84
A-syn A33T													
6l1l	2	https://doi.org/10.1016/j.jmb.2019.08.011	24	3.2	Yes	C1	91.46	1239.4	2478.9	179.65	0.7	2.41	4.82
A-syn Y39-phosph dimer													
6m1	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	25	2.7	Yes	C1	73	785.5	1570.9	179.45	1.1	2.4	4.8
hPFP 22-231													
6m1	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	26	2.8	Yes	C1	34.05	449.1	898.1	179.05	1.9	2.37	4.74
hPFP 22-231													
6m1	2	https://dx.doi.org/10.1016/j.jmb.2019.08.011	26	2.8	Yes	C1	36.88	414.3	828.7	178.94	2.12	2.44	4.88
hRNPA1													
6xhm	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	27	2.6	Yes	C1	51	891.5	1783.0	-0.972	0.972	4.814	4.814
Fus LC Domain													
6pco	1	https://www.ncbi.nlm.nih.gov/pubmed/3113660	28	3.3	No	C1	57	1042.6	2085.2	-0.831886	0.831886	4.81839	4.81839
A-syn H50Q Narrow													
6pes	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	28	3.6	No	C1	57	1042.6	2085.2	-0.831886	0.831886	4.81839	4.81839
A-syn H50Q Wide													
6hud	1	https://www.ncbi.nlm.nih.gov/pubmed/3113660	29	4.0	No	C1	48	548.8	1097.7	-1.608	1.608	4.903	4.903
light chain deltae, antibody													
6ic3	1	https://www.ncbi.nlm.nih.gov/pubmed/3113660	30	3.3	No	C1	69	1489.7	2979.3	0.58	0.58	4.8	4.8
Light Chain delta1, antibody													
5o3t	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	31	3.4	No	C1	70	812.6	1625.1	-1.05	1.05	4.74	4.74
tau AD SF													
6x5	1	https://www.ncbi.nlm.nih.gov/pubmed/3113660	31	3.2	No	C1	61	1147.2	2294.4	-0.75	0.75	4.78	4.78
tau PID NPF													
6lfo	1	https://dx.doi.org/10.1016/j.jmb.2019.08.011	16	3.2	No	C1	62	1019.5	2039.0	-0.845	0.845	4.786	4.786
Tau CBD Type 1													
6qjh	1	https://www.ncbi.nlm.nih.gov/pubmed/3113660	31	3.3	No	C1	66	671.4	1342.9	-1.26	1.26	4.7	4.7
Tau 4R, Snake													
6qjm	1	https://www.ncbi.nlm.nih.gov/pubmed/3113660	31	3.3	No	C1	34	250.3	500.6	-3.38	3.38	4.7	4.7
Tau 4R, Twister													
6qjp	1	https://www.ncbi.nlm.nih.gov/pubmed/3113660	31	3.3	No	C1	47	416.7	833.5	-2.03	2.03	4.7	4.7
Tau 4R, Jagged													
6qjq	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	31	3.7	No	C1	53	805.7	1611.4	-1.05	1.05	4.7	4.7
Tau 3R													
6n3b	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	31	3.8	No	C1	41	381.4	762.8	-2.27	2.27	4.81	4.81
TDP-43 (311-360) Seg A-syn													
6sdz	1	https://dx.doi.org/10.1016/j.jmb.2019.08.011	32	3.0	No	C1	55	729.8	1459.7	-1.19	1.19	4.825	4.825
Transferrin V30M													
6xyo	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	33	2.6	No	C1	54	590.0	1180.0	-1.44	1.44	4.72	4.72
A-syn MSA type I													
6xyq	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	33	3.3	No	C1	53	602.6	1205.1	-1.41	1.41	4.72	4.72
A-syn MSA type II(1)													
6xyq	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	34	3.1	No	C1	34	300.6	601.3	-2.88	2.88	4.81	4.81
A-syn MSA type II(2)													
6wqk	2	https://www.ncbi.nlm.nih.gov/pubmed/3113660	34	3.1	No	C1	34	300.6	601.3	-2.88	2.88	4.81	4.81

Supplementary References

1. Ghosh, U., Thurber, K. R., Yau, W.-M. & Tycko, R. Molecular structure of a prevalent amyloid- β fibril polymorph from Alzheimer's disease brain tissue. *Proc Natl Acad Sci U S A* **118**, (2021).
2. Gremer, L., Schölzel, D., Schenk, C., Reinartz, E., Labahn, J., Ravelli, R. B. G., Tusche, M., Lopez-Iglesias, C., Hoyer, W., Heise, H., Willbold, D. & Schröder, G. F. Fibril structure of amyloid- β (1-42) by cryo-electron microscopy. *Science* **358**, 116–119 (2017).
3. Kollmer, M., Close, W., Funk, L., Rasmussen, J., Bsoul, A., Schierhorn, A., Schmidt, M., Sigurdson, C. J., Jucker, M. & Fändrich, M. Cryo-EM structure and polymorphism of A β amyloid fibrils purified from Alzheimer's brain tissue. *Nat Commun* **10**, 4760 (2019).
4. Li, B., Ge, P., Murray, K. A., Sheth, P., Zhang, M., Nair, G., Sawaya, M. R., Shin, W. S., Boyer, D. R., Ye, S., Eisenberg, D. S., Zhou, Z. H. & Jiang, L. Cryo-EM of full-length α -synuclein reveals fibril polymorphs with a common structural kernel. *Nat Commun* **9**, (2018).
5. Guerrero-Ferreira, R., Taylor, N. M., Arteni, A.-A., Kumari, P., Mona, D., Ringler, P., Britschgi, M., Lauer, M. E., Makky, A., Verasdonck, J., Riek, R., Melki, R., Meier, B. H., Böckmann, A., Bousset, L. & Stahlberg, H. Two new polymorphic structures of human full-length alpha-synuclein fibrils solved by cryo-electron microscopy. *eLife* **8**, e48907 (2019).
6. Iadanza, M. G., Silvers, R., Boardman, J., Smith, H. I., Karamanos, T. K., Debelouchina, G. T., Su, Y., Griffin, R. G., Ranson, N. A. & Radford, S. E. The structure of a β 2 - microglobulin fibril suggests a molecular basis for its amyloid polymorphism. *Nature Communications* **9**, 4517 (2018).

7. Röder, C., Kupreichyk, T., Gremer, L., Schäfer, L. U., Pothula, K. R., Ravelli, R. B. G., Willbold, D., Hoyer, W. & Schröder, G. F. Cryo-EM structure of islet amyloid polypeptide fibrils reveals similarities with amyloid- β fibrils. *Nat Struct Mol Biol* **27**, 660–667 (2020).
8. Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structure and inhibitor design of human IAPP (amylin) fibrils. *Nat Struct Mol Biol* **27**, 653–659 (2020).
9. Gallardo, R., Iadanza, M. G., Xu, Y., Heath, G. R., Foster, R., Radford, S. E. & Ranson, N. A. Fibril structures of diabetes-related amylin variants reveal a basis for surface-templated assembly. *Nat Struct Mol Biol* **27**, 1048–1056 (2020).
10. Hervas, R., Rau, M. J., Park, Y., Zhang, W., Murzin, A. G., Fitzpatrick, J. A. J., Scheres, S. H. W. & Si, K. Cryo-EM structure of a neuronal functional amyloid implicated in memory persistence in *Drosophila*. *Science* **367**, 1230–1234 (2020).
11. Liberta, F., Loerch, S., Rennegarbe, M., Schierhorn, A., Westermark, P., Westermark, G. T., Hazenberg, B. P. C., Grigorieff, N., Fändrich, M. & Schmidt, M. Cryo-EM fibril structures from systemic AA amyloidosis reveal the species complementarity of pathological amyloids. *Nature Communications* **10**, 1104 (2019).
12. Bansal, A., Schmidt, M., Rennegarbe, M., Haupt, C., Liberta, F., Stecher, S., Puscalau-Girtu, I., Biedermann, A. & Fändrich, M. AA amyloid fibrils from diseased tissue are structurally different from in vitro formed SAA fibrils. *Nat Commun* **12**, 1013 (2021).
13. Röder, C., Vettore, N., Mangels, L. N., Gremer, L., Ravelli, R. B. G., Willbold, D., Hoyer, W., Buell, A. K. & Schröder, G. F. Atomic structure of PI3-kinase SH3 amyloid fibrils by cryo-electron microscopy. *Nat Commun* **10**, 3754 (2019).

14. Fitzpatrick, A. W. P., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., Crowther, R. A., Ghetti, B., Goedert, M. & Scheres, S. H. W. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185–190 (2017).
15. Falcon, B., Zivanov, J., Zhang, W., Murzin, A. G., Garringer, H. J., Vidal, R., Crowther, R. A., Newell, K. L., Ghetti, B., Goedert, M. & Scheres, S. H. W. Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019).
16. Zhang, W., Tarutani, A., Newell, K. L., Murzin, A. G., Matsubara, T., Falcon, B., Vidal, R., Garringer, H. J., Shi, Y., Ikeuchi, T., Murayama, S., Ghetti, B., Hasegawa, M., Goedert, M. & Scheres, S. H. W. Novel tau filament fold in corticobasal degeneration. *Nature* **580**, 283–287 (2020).
17. Guenther, E. L., Ge, P., Trinh, H., Sawaya, M. R., Cascio, D., Boyer, D. R., Gonen, T., Zhou, Z. H. & Eisenberg, D. S. Atomic-level evidence for packing and positional amyloid polymorphism by segment from TDP-43 RRM2. *Nature Structural & Molecular Biology* **25**, 311–319 (2018).
18. Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat Struct Mol Biol* **26**, 619–627 (2019).
19. Falcon, B., Zhang, W., Murzin, A. G., Murshudov, G., Garringer, H. J., Vidal, R., Crowther, R. A., Ghetti, B., Scheres, S. H. W. & Goedert, M. Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **561**, 137 (2018).
20. Glynn, C., Sawaya, M. R., Ge, P., Gallagher-Jones, M., Short, C. W., Bowman, R., Apostol, M., Zhou, Z. H., Eisenberg, D. S. & Rodriguez, J. A. Cryo-EM structure of a human prion fibril with a hydrophobic, protease-resistant core. *Nat Struct Mol Biol* **27**, 417–423 (2020).

21. Boyer, D. R., Li, B., Sun, C., Fan, W., Zhou, K., Hughes, M. P., Sawaya, M. R., Jiang, L. & Eisenberg, D. S. The α -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. *PNAS* (2020) doi:10.1073/pnas.1917914117.
22. Zhao, K., Li, Y., Liu, Z., Long, H., Zhao, C., Luo, F., Sun, Y., Tao, Y., Su, X.-D., Li, D., Li, X. & Liu, C. Parkinson's disease associated mutation E46K of α -synuclein triggers the formation of a distinct fibril structure. *Nat Commun* **11**, 2643 (2020).
23. Sun, Y., Hou, S., Zhao, K., Long, H., Liu, Z., Gao, J., Zhang, Y., Su, X.-D., Li, D. & Liu, C. Cryo-EM structure of full-length α -synuclein amyloid fibril with Parkinson's disease familial A53T mutation. *Cell Res* **30**, 360–362 (2020).
24. Zhao, K., Lim, Y.-J., Liu, Z., Long, H., Sun, Y., Hu, J.-J., Zhao, C., Tao, Y., Zhang, X., Li, D., Li, Y.-M. & Liu, C. Parkinson's disease-related phosphorylation at Tyr39 rearranges α -synuclein amyloid fibril structure revealed by cryo-EM. *Proc Natl Acad Sci U S A* **117**, 20305–20315 (2020).
25. Wang, L.-Q., Zhao, K., Yuan, H.-Y., Wang, Q., Guan, Z., Tao, J., Li, X.-N., Sun, Y., Yi, C.-W., Chen, J., Li, D., Zhang, D., Yin, P., Liu, C. & Liang, Y. Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. *Nat Struct Mol Biol* **27**, 598–602 (2020).
26. Sun, Y., Zhao, K., Xia, W., Feng, G., Gu, J., Ma, Y., Gui, X., Zhang, X., Fang, Y., Sun, B., Wang, R., Liu, C. & Li, D. The nuclear localization sequence mediates hnRNPA1 amyloid fibril formation revealed by cryoEM structure. *Nat Commun* **11**, 6349 (2020).
27. Lee, M., Ghosh, U., Thurber, K. R., Kato, M. & Tycko, R. Molecular structure and interactions within amyloid-like fibrils formed by a low-complexity protein sequence from FUS. *Nature Communications* **11**, 5735 (2020).

28. Boyer, D. R., Li, B., Sun, C., Fan, W., Sawaya, M. R., Jiang, L. & Eisenberg, D. S. Cryo-EM structures of alpha-synuclein fibrils with the H50Q hereditary mutation reveal new polymorphs. *bioRxiv* 738450 (2019) doi:10.1101/738450.
29. Swuec, P., Lavatelli, F., Tasaki, M., Papissoni, C., Rognoni, P., Maritan, M., Brambilla, F., Milani, P., Mauri, P., Camilloni, C., Palladini, G., Merlini, G., Ricagno, S. & Bolognesi, M. Cryo-EM structure of cardiac amyloid fibrils from an immunoglobulin light chain AL amyloidosis patient. *Nature Communications* **10**, 1269 (2019).
30. Radamaker, L., Lin, Y.-H., Annamalai, K., Huhn, S., Hegenbart, U., Schönland, S. O., Fritz, G., Schmidt, M. & Fändrich, M. Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis. *Nature Communications* **10**, 1103 (2019).
31. Zhang, W., Falcon, B., Murzin, A. G., Fan, J., Crowther, R. A., Goedert, M. & Scheres, S. H. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife* **8**, (2019).
32. Schmidt, M., Wiese, S., Adak, V., Engler, J., Agarwal, S., Fritz, G., Westermark, P., Zacharias, M. & Fändrich, M. Cryo-EM structure of a transthyretin-derived amyloid fibril from a patient with hereditary ATTR amyloidosis. *Nat Commun* **10**, 5008 (2019).
33. Schweighauser, M., Shi, Y., Tarutani, A., Kametani, F., Murzin, A. G., Ghetti, B., Matsubara, T., Tomita, T., Ando, T., Hasegawa, K., Murayama, S., Yoshida, M., Hasegawa, M., Scheres, S. H. W. & Goedert, M. Structures of α -synuclein filaments from multiple system atrophy. *Nature* **585**, 464–469 (2020).
34. Lu, J., Cao, Q., Hughes, M. P., Sawaya, M. R., Boyer, D. R., Cascio, D. & Eisenberg, D. S. CryoEM structure of the low-complexity domain of hnRNPA2 and its conversion to pathogenic amyloid. *Nature Communications* **11**, 4090 (2020).

35. Li, Q., Babinchak, W. M. & Surewicz, W. K. Cryo-EM structure of amyloid fibrils formed by the entire low complexity domain of TDP-43. *Nat Commun* **12**, 1620 (2021).
36. Wu, X., Ma, Y., Zhao, K., Zhang, J., Sun, Y., Li, Y., Dong, X., Hu, H., Liu, J., Wang, J., Zhang, X., Li, B., Wang, H., Li, D., Sun, B., Lu, J. & Liu, C. The structure of a minimum amyloid fibril core formed by necroptosis-mediating RHIM of human RIPK3. *PNAS* **118**, (2021).

Title: Towards cryo-EM structure determination of Amyloid- β S26s oligomers

David R. Boyer, Romany Abshkaron, Sean Jiang, Peng Ge, Kang Zhou, Carter Lantz, Luis Alejandro Foley, Jevgenij Raskatov, David S. Eisenberg

Abstract: Amyloid oligomers are generally thought to be transient assemblies of ~10-100 copies of an amyloid protein that either precede the formation of or form in parallel to amyloid fibrils. They have been implicated as the potential etiological agents for numerous amyloid diseases, including Alzheimer's. Nonetheless, despite decades of work, and the recent revolution in the visualization of the atomic structures of amyloid fibrils, there has been little progress in the structure determination of amyloid oligomers. Here, we make progress towards the cryo-EM structure determination of a variant of amyloid- β found in Alzheimer's amyloid plaques – A β S26s. A β S26s has a D-amino acid at Ser26 and forms stable oligomers in vitro. We find that a Fab that binds to the N-terminus of the A β peptide can be complexed with A β S26s oligomers and can act as a fiducial marker to aid image alignments during processing of electron microscopy images. In this chapter, we summarize current efforts towards determining the cryo-EM structure of the A β S26s oligomer-Fab 3D6 complex, as well as discuss parallel strategies to determine both A β S26s and other amyloid protein oligomer structures.

Introduction

Amyloid fibril formation is usually thought of as a two-step nucleation and elongation process¹. The nucleation step involves individual monomers of an amyloid protein colliding to form a small assembly, or nucleus, which then serves as a starting point for further monomer addition in the elongation step. The nuclei that precede fibril formation are also termed

oligomers. In addition to the two-step nucleation and elongation model, studies have posited additional mechanisms for the generation of oligomers, including the nucleation of oligomers along the sides of fibrils – termed secondary nucleation – as well as the release of oligomers from the tips of fibrils². Adding to this picture is the hypothesized existence of both on- and off-pathway oligomers³. The former being the nuclei that additional monomers extend from to form fibrils and the latter being oligomers that do not go on to form fibrils, but instead either persist as stable structures or disassociate back to monomer. Taken together, the amyloid aggregation process is thought to be a highly dynamic, where every step in the process is connected.

Visualization of the atomic structures of the different species involved in amyloid fibril formation is the most direct method of better understanding this process. For instance, determining the structures of oligomers and fibrils over time would help to resolve not only the structure of amyloid oligomers, but the possible structural differences between on-pathway and off-pathway oligomers. If an oligomeric structure resembled that of a small fibril, it would presumably be an on-pathway oligomer, whereas oligomeric structures incompatible with the fibril fold would presumably be off-pathway oligomers. Furthermore, studying the relative populations of different species over time would be enlightening: if oligomer structures resembling fibrils disappeared at a similar rate as fibril structures appeared, it is more conclusive these oligomers are on-pathway. Whereas if an oligomer structure that does not resemble a fibril persists throughout the aggregation process, it is likely this is a stable, off-pathway oligomer. Similarly, if an oligomer structure does not resemble a fibril and only appears during a certain period in the aggregation process, it is likely to be an unstable, off-pathway oligomer that forms and then disassociates back to monomer. Furthermore, late in the aggregation process, after the formation of amyloid fibrils, if we observe an increase in the population of oligomers, we can

answer questions relating to secondary nucleation: does secondary nucleation generate on-pathway or off-pathway oligomers? Are these structures the same or different to the oligomeric structures found before or at the beginning stages of fibril formation?

Despite the incredible potential that atomic structure determination of the different species formed during amyloid aggregation process would have, it remains a large technical challenge. To date, all methods for determining the atomic structures of biomolecules rely on averaging measurements from many identical molecules. For X-ray and electron crystallography, the diffraction of crystals composed of a large number of identical molecules is measured and used to calculate atomic structures. For nuclear magnetic resonance (NMR), the chemical shift of the individual nuclei of many copies of a biomolecule is measured to calculate the structure and dynamics of a protein. And in cryo-electron microscopy, many individual 2D or 3D images of biomolecules are averaged to calculate atomic structures. Since many different species exist simultaneously in the amyloid aggregation process, it is very difficult to purify a homogenous species for crystallization or NMR measurements. However, for cryo-EM since individual images of biomolecules are recorded before averaging, classification methods can be used to sort the particle images from a single sample into unique groups, provided the images of different species are sufficiently distinct⁴. This feature has made cryo-EM an ideal method to study heterogenous samples in recent years.

We decided to apply cryo-EM to study the atomic structures of oligomers of A β S26s – a modified version of the amyloid- β peptide found in Alzheimer’s disease plaques⁵. This variant of A β contains a D-amino acid at residue Ser26 and has been shown in vitro to have slower fibril growth and longer-lived oligomers compared to wild-type A β , which rapidly converts to fibrils. We hypothesized that these attributes make A β S26s a good starting point for cryo-EM structure

determination attempts since it facilitates trapping of different species along the aggregation pathway, especially those species that occur before extensive fibril formation. Here, we describe our attempts at structure determination of A β S26s oligomers, including the challenges these oligomers present to cryo-EM structure determination and the use of an antibody fragment to attempt to overcome these challenges.

Results

We first tested various growth conditions to obtain abundant, disperse, and homogeneously sized A β S26s oligomers by testing exhaustive combinations of: (i) the use of shaking or quiescent incubation, (ii) different incubation buffers, and (iii) a range of A β S26s starting concentrations (see Methods). We identified that 100 μ M A β S26s incubated quiescently at 37° C in 20 mM phosphate buffer pH 7.4 produced the best appearing oligomers by negative stain EM. We chose quiescent over shaking conditions since this allows slower fibril growth, allowing us to more easily capture pre-fibrillar oligomers for EM grid preparation. We used ThT fluorescence and negative stain EM to monitor oligomer and fibril growth over the course of several days, identifying that oligomers appeared stable over the course of 4-5 days, while fibrils start to appear variably at 1-3 days between different experiments (Figure 1 a-b). After many days or weeks of incubation, both fibrils and oligomers appear stable; however, oligomers become less spherical and more protofibrils appear. In addition, we observe that over time oligomers and protofibrils start to coat fibrils (Figure 1 b, 167 hrs.).

In order to determine the atomic structures of A β S26s oligomers, we made cryo-EM grids of oligomers after incubation for 20 hrs. and 44 hrs. Under our growth conditions, both of these time points have minimal fibrils and protofibrils present, with the majority of particles

being spherical oligomers. Furthermore, we selected two time points with the goal of determining if there are different oligomeric structures over the course of the incubation period. We took cryo-EM data sets on a Titan Krios (see Methods) for the two time points. A representative micrograph from this dataset shows discrete oligomers in a thin layer of vitreous ice, and the corresponding amplitude spectrum of this micrograph shows a ring at $1/4.8 \text{ \AA}$ indicating an isotropic arrangement of oligomers with β -sheet structure (Figure 1c). Independent 2D classifications of each dataset revealed the overall shapes of oligomers, but little internal structure (Figure 1d). In addition, the overall shapes of oligomers appeared largely similar at 20 and 44 hrs., suggesting that oligomers are stable over this time period. Ab initio reconstruction and subsequent 3D classification in both Cryosparc⁶ and RELION⁷ did not result in maps displaying secondary or tertiary structure.

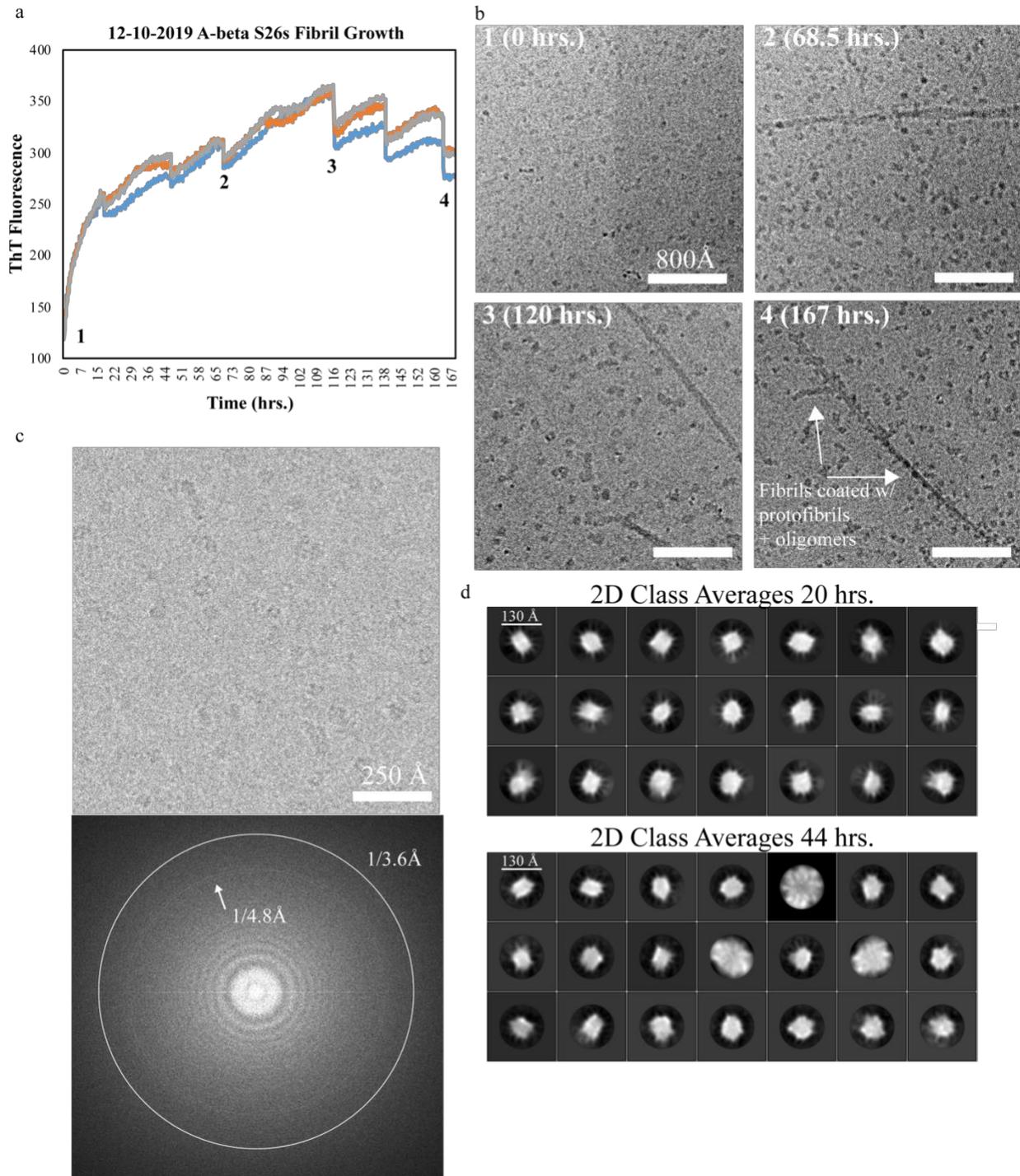


Figure 1 Initial A β S26s oligomer growth and cryo-EM data collection and processing. a) Fibril and oligomer formation assay monitored using ThT fluorescence. Samples were taken at various time points (dips in fluorescent curve) and examined by EM. b) Cryo-EM images of oligomers

and fibrils at various time points. Oligomers are stable over the course of days while fibrils start to appear between 1-3 days. After 7 days, oligomers and protofibrils visibly coat fibrils. c) Krios electron micrograph of oligomers in a thin layer of vitreous ice and amplitude spectrum of micrograph. Amplitude spectrum displays signal at $1/4.8 \text{ \AA}$ suggesting high β -sheet content of oligomers. d) Representative 2D classes of oligomers at 20 hrs. and 44 hrs. incubation show a lack of internal features in the oligomers, as well as no discernible difference in oligomer shape between the two time points (the large oligomers present at 44 hrs. are also present in 20 hrs. 2D classes, but not shown here).

We hypothesize that structure determination was hindered both by the heterogeneity of oligomeric species as well as their lack of medium resolution features (tertiary and secondary features such as protein domains, α -helices). Since cryo-EM structure determination relies on iterative orientation refinement starting at low resolution (aligning particle shapes) and progresses towards medium (tertiary structure) and high resolution (secondary and primary structure), the uniform shape of the largely spherical oligomers, as well as their high β -sheet content – which only becomes apparent at resolutions better than 4.8 \AA , as opposed to 9 \AA for α -helices – makes image alignment difficult. We therefore sought to add an “artificial” protein domain to the oligomer particles by complexing them with a Fab. We hypothesize that the Fab can act as a fiducial marker during image alignment in much the same way Fabs have aided image alignment and structure determination of membrane proteins⁸.

Mass spectrometry fragmentation experiments reveal that the N-terminus of A β S26s is most susceptible to forming short peptide fragments, while the C-terminal part of the peptide is more protected from fragmentation (Figure 3a). This indicates that the N-terminus is likely more

solvent-exposed, which is consistent with both A β fibril structures⁹ as well as ZipperDB¹⁰ predictions that indicate the amyloidogenic sequences are largely clustered in the C-terminal half of the peptide. We therefore selected Fab 3D6, which binds to the first five residues of the A β sequence, to complex with A β S26s oligomers¹¹. The crystal structure of 3D6 complexed with wild-type A β 1-40 displays strong electron density for A β residues 1-DAEFRH-6 between the heavy- and light-chain CDRs. This suggests high affinity between 3D6 and its epitope, making it a good candidate to assist in cryo-EM structure determination. Dot blot experiments reveal that Fab 3D6 binds to our A β oligomers in a concentration dependent manner (Figure 3b). We therefore mixed 3D6 with A β oligomers in a sub-stoichiometric ratio (~1:10 Fab:A β S26s monomer equivalent) assuming that the A β S26s monomer to oligomer conversion rate was relatively low and that a single Fab may bind a given oligomer that has many copies of the A β S26s peptide (Figure 3c). Cryo-EM images show Fabs decorating fibrils of A β S26s (Figure 3c), and therefore, assuming there was Fab decoration of oligomers as well, we collected a preliminary cryo-EM data set of the A β S26s-Fab mixture.

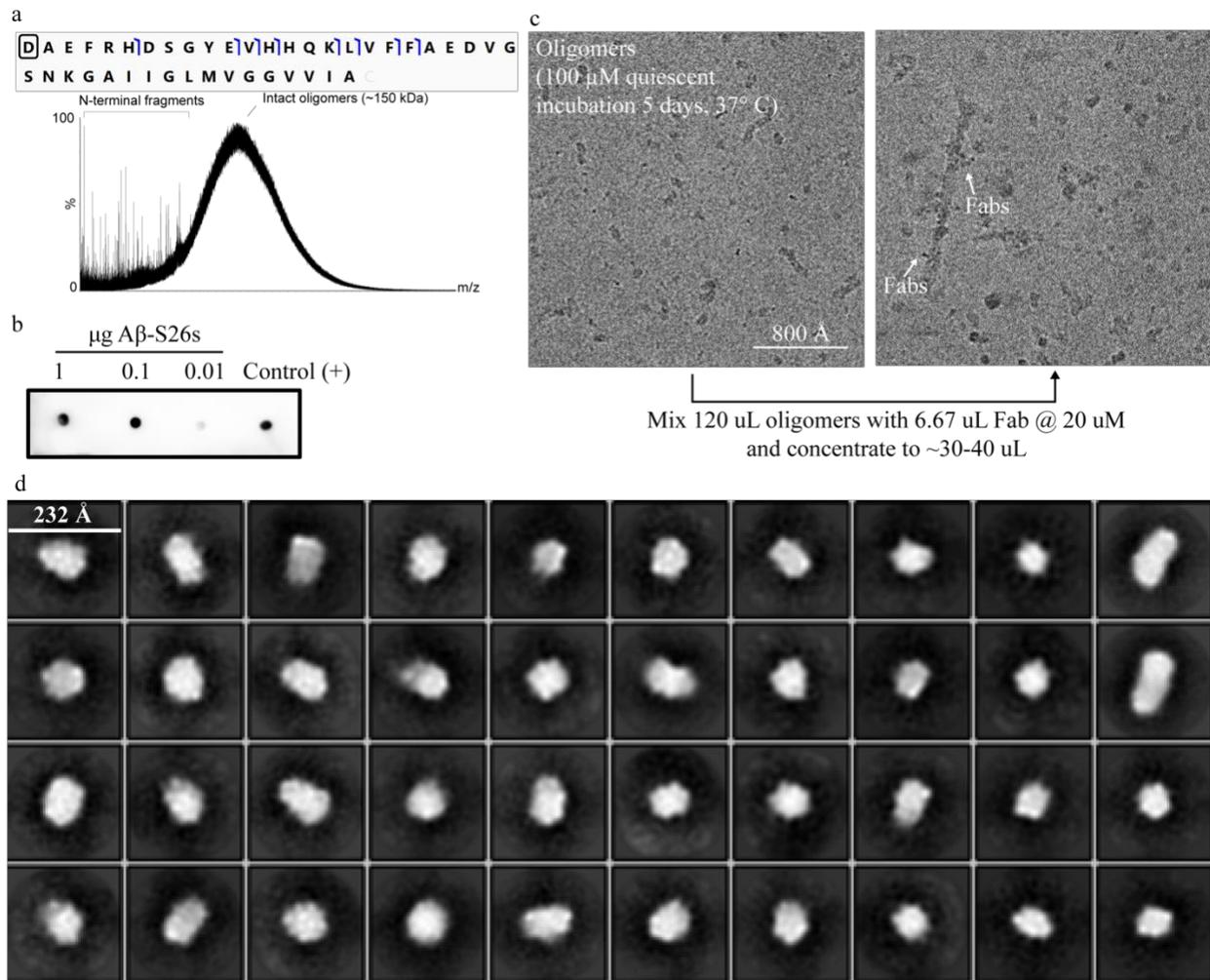


Figure 2 Initial characterization of antibody binding and attempt at data collection of A β S26s oligomer-Fab 3D6 particles. a) Mass spectrometry fragmentation experiment identifies that intact oligomers have a Gaussian distribution of m/z values, indicating an average MW of 150 kDa. In addition, fragmentation mapping indicates N-terminal portions of the peptide are most susceptible to fragmentation, suggesting they are the most solvent-exposed part of the peptide. b) Dot blot assay demonstrating that Fab 3D6 recognizes A β S26s oligomers. c) Cryo-EM images of oligomers before and after mixing with Fab 3D6. After mixing, Fab 3D6 can be seen decorating a fibril. d) 2D classification of oligomers reveals that there is an absence of Fab binding.

Surprisingly, cryo-EM data processing revealed that no Fab was bound to oligomers, and oligomer 2D class averages looked largely the same as our initial cryo-EM experiments of oligomers without Fab (Figure 3d). Ab initio reconstruction and 3D classification efforts did not result in viable maps for this oligomer preparation, similar to our initial cryo-EM experiment. Therefore, we sought to improve our A β S26s-3D6 complex sample preparation. We first attempted to produce oligomers in greater abundance and with greater apparent homogeneity than our initial condition. Our most successful attempt (see Methods) came by seeding A β S26s monomer with 5% (v/v) sonicated A β S26s fibrils, which, to our surprise, produced many spherical oligomers stable over many days and of more homogenous size compared to non-seeded preparations (Figure 3 a-b). We term this preparation “G2” and used it going forward in our attempts to produce better A β S26s oligomer-3D6 complexes.

We investigated several methods to improve our production and purification of the oligomer-Fab complex including size exclusion and ion exchange chromatography (see Methods); however, the best method we found was to perform a pulldown using Ni-conjugated magnetic beads. In this method, we incubated 3D6 bound to Ni-conjugated magnetic beads with A β S26 oligomers at various molar ratios, washed the sample with buffer to remove unbound Fab and oligomers, and eluted the oligomer-Fab complex from the beads with 500 mM imidazole buffer. To remove imidazole and excess Fab in the elution, we used Amicon 100 kDa spin filters to buffer-exchange the sample to 20 mM Tris-HCl pH 7.4, 150 mM NaCl (a “good” buffer for cryo-EM).

Supplementary Figure 1 highlights two different pulldown experiments; first using our original, un-seeded oligomer growth condition with less Fab (1:15 Fab:A β S26 monomer equivalent), and second using our G2 seeded condition with a larger amount of 3D6 (1:2.5

Fab:A β S26s monomer equivalent). In the first condition, more protofibrils and less isolated oligomers can be seen compared to the G2 seeded condition. In both conditions, Fab can be seen binding to individual oligomers after the pulldown. This results in one or two Fabs binding to the oligomers in the first condition where less Fab was used, and as many as six to eight Fabs binding to oligomers in the second condition when more Fab was used. We decided to proceed with the latter condition for cryo-EM since we reasoned that there would be less of a chance of Fabs partially occupying different epitope sites on the oligomer, potentially confounding image alignment.

We collected an initial negative stain data set of the A β S26s-3D6 particles to examine potential heterogeneity and also test whether the bound Fabs could be used to aid image alignment during data processing. 2D classification revealed oligomer classes with between two to eight bound Fabs, and ab initio reconstruction and 3D classification revealed 4 principle classes of particles, each with Fabs in different positions, suggesting point group symmetries of C3, C6, or D6 (Figure 3 b-c). In all classes, Fab 3D6 binds an inner spherical oligomer density around its equator. The inner spherical density in all classes is about 70 Å in diameter, suggesting the oligomer MW is ~200 kDa and is composed of ~50 A β S26s peptides in each oligomer (assuming 1.3 g/cm³ protein density and 4.5 kDa A β S26s MW), which is consistent with our mass spectrometry data demonstrating a Gaussian distribution of oligomer sizes centered around 150 kDa (Figure 3a). In addition, the 70 Å diameter for the central oligomer density includes the negative stain, making the true diameter for the oligomer likely less than 70 Å and hence even more in line with our mass spec data.

A potential model for the 3D6-bound A β S26s oligomers is presented in Figure 3d. Here, we assume the oligomer has six-fold rotational symmetry (C6), given the arrangement we see in

our 3D reconstructions. We also assume the peptides in the oligomer form a 15 layer stack of in-register, parallel β -sheets with a spacing of 4.8 Å and height of 70 Å. This architecture, similar to that of a fibril, would place the N-termini of the peptide facing laterally outward from the oligomer thereby satisfying the observation that antibody binding mainly occurs along the sides of the oligomer. Although this model is based on limited experimental constraints, it nevertheless helps us to envision the data processing strategies that could be employed to resolve the oligomer-Fab structure. For instance, the arrangement of Fabs can help us to determine which point group symmetries the complex may have and guide our testing of different symmetries during 3D classification. In a six-fold rotational arrangement like that in Figure 3d, employing C6 symmetry would help reduce angular searches to a range of 60°. However, as can be seen from the negative stain 3D reconstructions, particularly the bottom left reconstruction in Figure 3c, the Fabs can bind at different positions along the height of the oligomer. The bottom left reconstruction also demonstrates that two Fabs can stack upon one another along the height of the oligomer.

The closely packed N-termini protruding laterally outward in our oligomer model creates more potential Fab binding sites than can be satisfied given the size of the Fab. Therefore, likely up to two Fabs could bind along the side of the oligomer, or one Fab could bind at different heights. This may confound image averaging methods and lead to blurred reconstructions along the height of the oligomer-Fab complex. To overcome this, after obtaining an initial reconstruction of the complex, we could mask out the Fabs and perform local angular searches or skip angular searches while performing an offset search to try to iteratively align the inner oligomeric density.

Additional arrangements of the oligomer-Fab complex that may present challenges to structure determination are illustrated in Figure 3e. The left most arrangement is the easiest to tackle given that the outer symmetry of the Fabs (C3) is a subset of the inner oligomer symmetry (C6). If successful refinement is achieved using the apparent C3 symmetry of the Fabs, the internal symmetry of the oligomer will become apparent and the Fabs can be masked out or signal subtracted, allowing for refinement of the oligomer with C6 symmetry (this applies for different symmetry subsets as well, e.g., C2/C4/C6, C3/C6/C9). The middle left arrangement highlights the potential case of a symmetry mismatch between the apparent outer symmetry of the bound Fabs and the inner symmetry of the oligomer (C5/C6). In this case, we can proceed as in Liu, et al. where the structure of a C5/C12 portal vertex complex from herpes simplex virus 1 was determined using sub-particle extraction and symmetry expansion techniques¹². The right two arrangements highlight the potential problem of partial occupancies of the Fabs bound to the oligomers. This can be overcome either by initial refinement using C1 until the internal symmetry of the oligomer becomes apparent and then subsequent masking or signal subtraction of the Fabs for refinement with the internal symmetry of the oligomer. Alternatively, different symmetries could be tested during initial refinement, such as the highest apparent symmetry of the Fab molecules (e.g., if neighboring Fabs are 60° apart, C6 could be used) which would lead to lower resolution Fab densities due to partial occupancy at each binding site. If the apparent symmetry of the Fabs matches that of the internal oligomer symmetry, the oligomer density should become clear; if the apparent symmetry of the Fabs does not match the internal symmetry of the oligomer, sub-particle extraction and symmetry expansion techniques could be used as in the C5/C6 symmetry mismatch case.

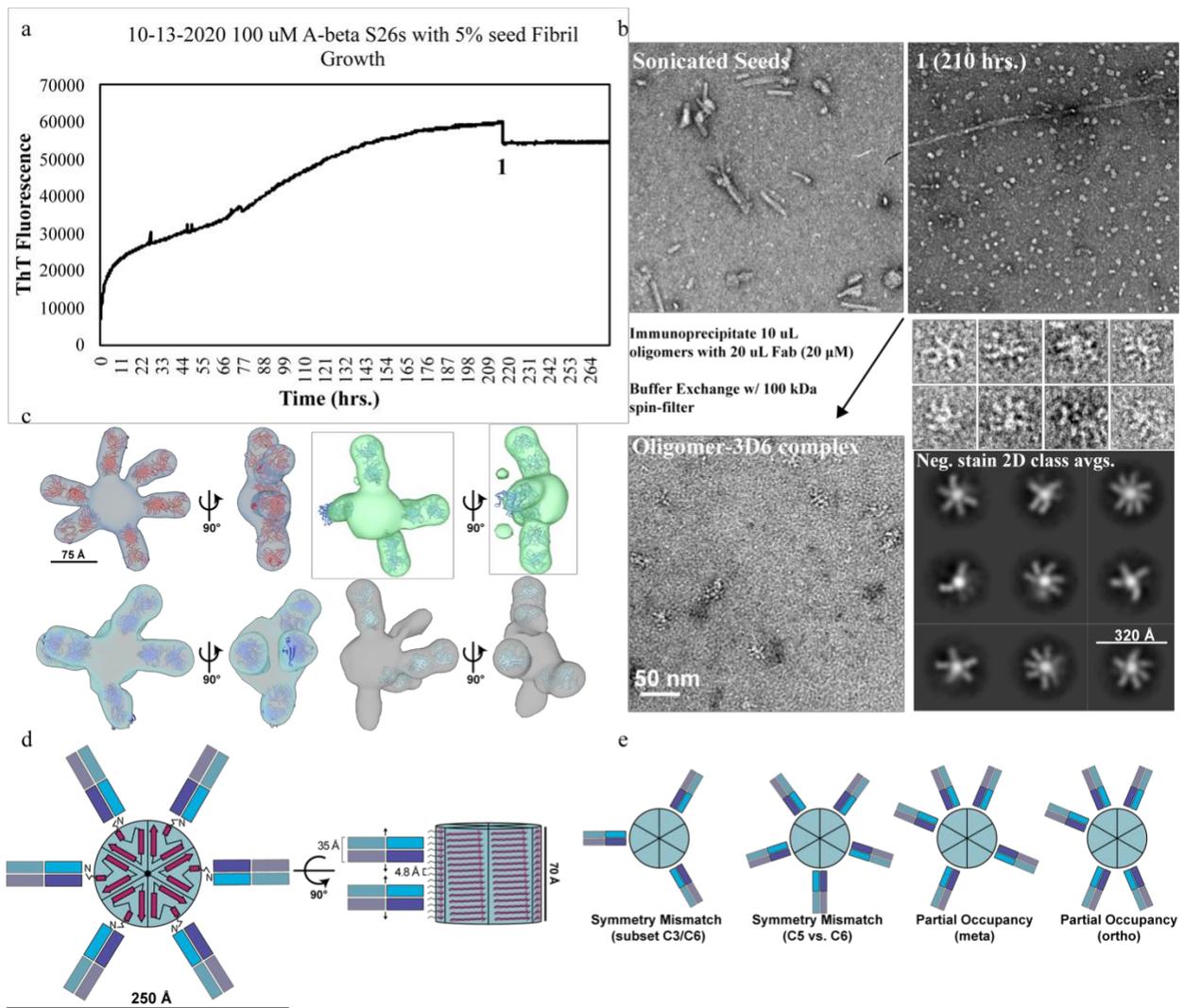


Figure 3 Seeded oligomer growth and improved Fab 3D6 labeling of oligomers. a) ThT assay monitoring growth of seeded A β S26s oligomers. ThT curve indicates an initial exponential growth phase, followed by an extended elongation phase. Samples were removed at 210 hrs. for analysis by EM. b) 5% v/v of sonicated seeds (top left) were used to seed oligomer and fibril growth. After 210 hrs. of quiescent incubation at 37 $^{\circ}$ C, abundant and well-dispersed oligomers can be seen (top right). Oligomers at this time point were immunoprecipitated with a large excess of Fab, buffer exchanged, and visualized by EM (bottom left). Negative stain A β S26s oligomer-Fab 3D6 particles and 2D class averages (bottom right) show oligomers decorated by between

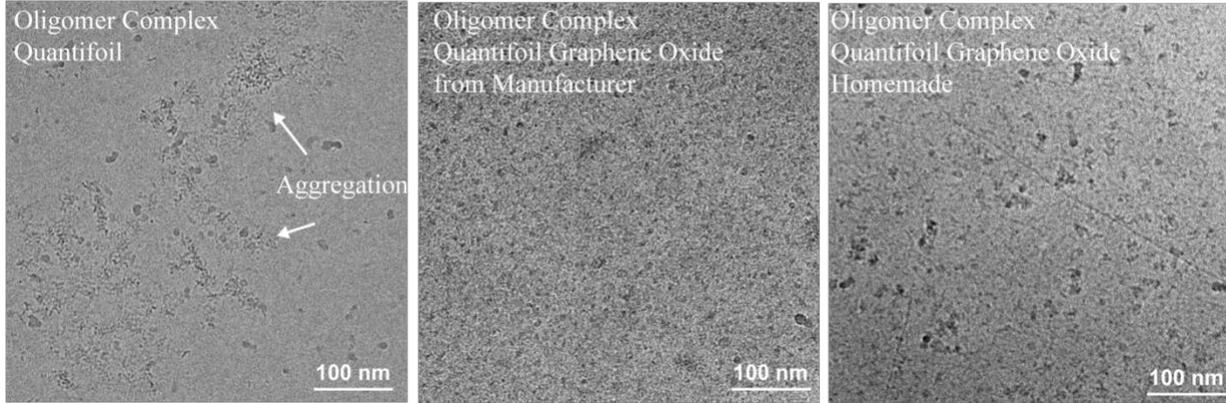
~2-8 Fab molecules. d) Hypothetical architecture of the oligomer-Fab complex showing six-fold rotational symmetry. The side view of the model highlights how steric constraints impose potential ambiguity of the Fab position along the height of the oligomer. e) Hypothetical arrangements of the oligomer-Fab complex that could present challenges during data processing.

Given the above possible arrangements of the oligomer-Fab complex, and the potential data processing strategies that could be used to determine the oligomer structure, we proceeded with preparing cryo-EM grids of the G2 seeded oligomers complexed with Fab 3D6. We discovered that the Fab changed the properties of the particles such that normal Quantifoil or UltrAuFoil grids with a variety of glow discharge and blotting conditions always gave rise to aggregated particles (Figure 4 a). We therefore tried both Quantifoil grids with single layer graphene oxide purchased from a manufacturer, as well as preparing single layer graphene oxide Quantifoil grids ourselves immediately before use (see Methods). We learned that preparing fresh graphene oxide coated Quantifoil grids led to the best hole coverage (highest percentage of holes with usable graphene) as well as generally led to better particle distribution and concentration than the graphene oxide grids purchased from the manufacturer.

We therefore prepared and collected preliminary cryo-EM data for oligomer-Fab complex particles to ensure that the complex does not disassociate due to forces involved during the blotting procedure. Our preliminary cryo-EM data set revealed that the oligomer-Fab complex withstood the shear forces present during grid blotting, so we proceeded to collect a high resolution data set on the CNSI Titan Krios. 2D class averaging from our Krios dataset reveals oligomer-Fab particles that appear similar to our negative stain class averages: a central

oligomer density decorated with ~6 Fabs. Ab initio reconstruction and 3D classification efforts are underway to determine the structure of the oligomer-Fab complex.

a



b

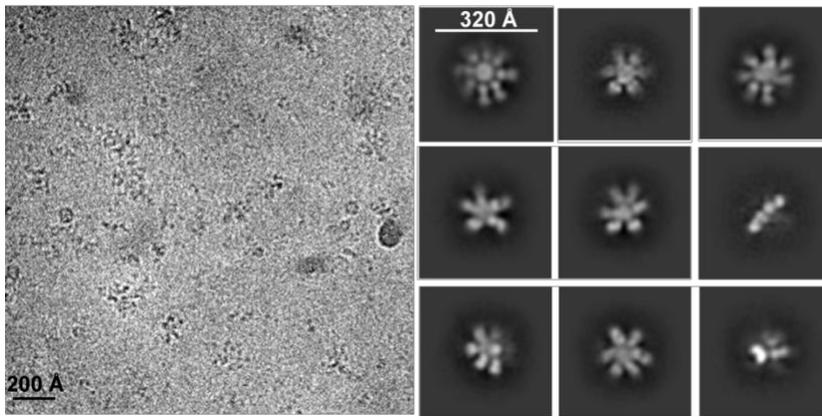


Figure 4 Cryo-EM grid preparation and data processing of A β S26s oligomer-Fab 3D6 complex.

a) Oligomer complex using standard Quantifoil (left) or gold grids (not shown) leads to aggregation. On Quantifoil GO R1.2/1.3 grids ordered from EMS, oligomers are not aggregated, but are crowded and difficult to identify (middle). Homemade graphene oxide Quantifoil grids (see Methods) lead to better particle distribution and ice thickness (right). b) Krios image (left) and initial 2D classification (right) of oligomer-Fab particles.

Discussion

We have demonstrated that although A β S26 oligomers have features making them amenable to study by cryo-EM – able to be produced in enough quantity, stable over a multi-day time scale, exist as discrete particles, do not aggregate during cryo-EM grid preparation – they nonetheless exhibit features prohibitive to structure determination – heterogeneity and lack of medium resolution sub-structure to facilitate image alignment. We therefore used Fab 3D6, which recognizes a linear epitope at the N-terminus of the A β peptide, to act as a fiducial marker to aid image alignment. We showed that during negative stain and cryo-EM image processing, the attached Fabs can aid in image alignment and classification, and may present a way forward to revealing the atomic structure of A β S26s oligomers. Efforts are currently underway to process high resolution images of oligomers decorated with Fab 3D6 in order to determine the atomic structure of the complex.

One possible drawback of our current strategy is the non-redundancy of the 3D6 epitope. Due to the fact that our mass spec and antibody labeling experiments show that the N-terminus of the A β S26s peptide is solvent-exposed, and that oligomers are composed of ~50 peptides, Fab 3D6 could be binding the oligomers in many possible unique permutations. Combining this with the fact that there may be many different oligomer species, each with a different structure, this creates a large combinatorial problem for cryo-EM image alignment and averaging. Although we discussed strategies for overcoming partial or non-identical labeling of oligomers by Fab 3D6, it is worth outlining other potential solutions to oligomer structure determination methods.

One potential solution is to use an antibody that binds to A β S26s oligomers in a unique manner. This could be achieved by employing an antibody that binds to the C-terminal region of A β S26s that is largely buried in the oligomer core. This would reduce the number of exposed

epitopes, possibly resulting in a few unique antibody-accessible binding sites making image alignment simpler. For instance, in the case of α -synuclein fibrils, there exist antibodies that only bind to one tip of the fibril structure. This is likely because it is only at one end of the fibril where the antibody epitope is available for binding, whereas in the rest of the fibril, it is covered by other layers. If the oligomer similarly had a single binding site for an antibody due to the fact that most antibody epitopes were buried in the oligomer core, this would greatly facilitate image alignment.

In order to attempt a different antibody labeling strategy, we obtained nanobodies Nb3 and Nb9 that are shown to bind with strong affinity to residues 17-28 of the wild-type A β peptide¹³. We also designed a bivalent version of Nb3, containing two Nb3 sequences connected by a linker. Our preliminary experiments show that all three nanobody constructs bind to A β S26s oligomers, indicating that the 17-28 epitope is still available for binding, despite a D-amino acid at Ser26. These three nanobodies will serve as backups in case structure determination with Fab 3D6 is unsuccessful. One potential drawback of using nanobodies is their smaller size (Nb3 and Nb9 are 15 kDa, bivalent Nb3 is 30 kDa), which potential limits visualization in cryo-EM images. However, their lack of a constant domain makes them less flexible than Fabs, which may aid image alignment. We will test the ability of Nb3, Nb9, and bivalent Nb3 to help image alignment of A β S26s oligomers in a similar manner to our oligomer-3D6 complex: we will first screen the oligomer-nanobody complexes by negative stain EM, and if the nanobodies can be visualized on the oligomer, we will attempt cryo-EM structure determination.

Another strategy that could overcome the intractability of oligomers for cryo-EM single particle analysis is cryo-electron tomography¹⁴. This method relies on obtaining partially complete 3D volumes (tomograms) of unique objects by collecting images of the sample tilted

+/- $\sim 60^\circ$. Although this leads to a reconstruction of the objects in the tomogram, it comes at a cost of lower resolution ($\sim 2-4$ nm) because the electron dose must be kept low enough to prevent radiation damage over the course of the tomogram. Since we are interested in the atomic arrangement of the A β S26s peptides in the oligomers, oligomer reconstructions from tomography would likely not reveal much more than the shape of our oligomers without revealing their internal structure. However, it may be possible to use sub-tomogram averaging to align and classify oligomer volumes found within the tomograms to achieve higher resolution. This may at the very least inform us of the symmetry of the oligomer, which could then be used in either our single particle analysis or subtomogram averaging to further improve resolution. Recent advances including EMClarity¹⁵ and M¹⁶, have pushed the boundary of subtomogram averaging to near atomic resolution for samples such as the ribosome and the immature HIV-1 Gag particle. It may therefore be possible to apply state-of-the-art cryo-ET methods to determine structures of A β S26s oligomers if single particle analysis fails.

Given that our goal is to determine the atomic structures of amyloid oligomers to reveal for the first time their atomic structure, and to see if there is a universal architecture to oligomer structures similar to the cross- β fold in fibrils, we are largely agnostic to which amyloid protein we study. Therefore, in addition to new antibody labeling approaches and cryo-ET of A β S26s oligomers, it may be expedient to examine additional amyloid protein oligomers. We initially experimented with A β S26s due to the fact that it forms long-lived oligomers, which are helpful for structural studies; however, we have shown that A β S26s has a large amount of heterogeneity that makes cryo-EM structure determination difficult. Indeed, cryo-EM images of A β S26s fibrils after an extended 30 day incubation show at least three fibril polymorphs (Supplementary Figure 2). Given that oligomers are precursors to fibrils, this suggests that A β S26s oligomers may have

at least three fibril polymorphs and perhaps more since not all oligomers may convert to fibrils, or certain oligomeric species may form that then disassociate. Unlike fibrils where multiple polymorphs can be either identified by eye during manual particle picking, or separated by 2D classification with large box sizes, oligomers of A β S26s are not easily identifiable by eye or by 2D classification. A more viable starting point for determining the structure of recombinantly assembled oligomers may be to start with a protein that forms monomorphic fibrils, which may potentially suggest more monomorphic oligomers. For instance, α -synuclein bearing the hereditary mutation E46K forms fibrils of a single type and therefore oligomers of this protein may be more likely to be homogenous¹⁷.

In conclusion, we have made efforts toward determining the structures of A β S26s oligomers and have shown that while the heterogeneity and lack of clear internal features of oligomers make image alignment and classification difficult, antibody labeling methods may provide a solution to this problem. We have outlined a clear strategy for producing and screening oligomer-antibody complexes and have applied it to a A β S26s oligomer-Fab 3D6 complex, producing reconstructions of 3D6 labeled oligomers and getting as far as high resolution cryo-EM data collection and preliminary data processing of A β S26s oligomer-3D6 particles. We have also discussed both potential pitfalls of and potential improvements to our method. Alternative strategies such as cryo-ET and a different amyloid protein starting point may also be worthwhile endeavors to reveal amyloid protein oligomers. Future work using a combination of these methods will hopefully lead to routine structure determination of amyloid oligomers, as has been accomplished for amyloid fibrils, furthering our understanding of amyloid protein folding and amyloid disease.

Materials and Methods

Peptide Reconstitution

Peptides were synthesized as in Foley, et al. and stored as a lyophilized powder⁵. In order to create smaller aliquots of peptide, we reconstituted lyophilized powder in 100 μ L 20 mM NH_4OH and re-lyophilized small aliquots of dissolved peptide for storage at -20°C and future use. For reconstitution of peptide immediately prior to use in oligomer growth experiments, we followed the procedure in Foley, et al.⁵ Briefly, peptide was dissolved in chilled 20 mM NaOH and kept on ice. Amount of NaOH added was calculated so that after dilution into oligomer growth buffer, final % v/v of NaOH was less than 3% in order to minimize potential peptide hydrolysis. Solubilized peptide solution was sonicated in a water bath sonicator for 30 seconds and its concentration was measured with a NanoDrop using absorbance at 280 nm and molar extinction coefficient at 280 nm of $1490\text{ M}^{-1}\text{cm}^{-1}$.

Oligomer Growth

We screened various oligomer growth conditions using a combination of different buffers, starting peptide concentration, and incubation conditions. Buffers tested included: PBS, MOPS, 20 mM phosphate buffer pH 7.4, and 20 mM Tris-HCl 150 mM NaCl. For 20 mM phosphate buffer pH 7.4, we also screened 0-200 mM NaCl. We screened starting peptide concentration ranging from 5-100 μ M, and incubation conditions were either quiescent at 37°C or using double orbital shaking (400 rpm) in an Omega microplate reader (BMG Labtech). All growth conditions were monitored using ThT fluorescence (20 μ M ThT) and 440/485 nm excitation/emission. For both non-seeded and seeded oligomer growth, the condition that produced the most oligomers with good dispersity and homogenous size was 100 μ M $\text{A}\beta$ S26s, 20 mM phosphate buffer pH 7.4, 20 μ M ThT, with quiescent incubation at 37°C .

For the “G2” seeded oligomer growth condition, we generated fibrils by incubating 50 μ M A β S26s in 20 mM phosphate buffer pH 7.4, 20 μ M ThT using 400 rpm double orbital shaking for 3 days. The resulting fibrils were sonicated to generate short, fragmented fibrils. The sonicated fibrils were used at 5% v/v in our standard oligomer growth condition (100 μ M A β S26s, 20 mM phosphate buffer pH 7.4, 20 μ M ThT, with quiescent incubation at 37° C).

Fab 3D6

Fab antibody was synthesized and obtained from BioIntron Biological Inc., China. Fab antibody was expressed and purified from CHO-K1 cell line (ECACC, Cat no.: 85051005).

Heavy chain:

EVKLVESGGGLVKPGASLKLSCAASGFTFSNYGMSWVRQNSDKRLEWVASIRSGGGRT
YYSDNVKGRFTISRENAKNTLYLQMSSLKSEDALYYCVRYDHYSGSSDYWGQGTTVT
VSSAKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAV
LQSDLYTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPRDCG

Light chain:

YVVMQTPLTLSVTIGQPASISCKSSQSLLDSDGKTYLNWLLQRPQGSPKRLIYLVSKLD
SGVPDRFTGSGSGTDFTLKISRIEAEDLGLYYCWQGTHFPRTFGGGTKLEIKRADAAPT
SIFPPSSEQLTSGGASVVCFLNFPKDIKWKIDGSRQNGVLNSWTDQDSKDYSTYS
MSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC

Immunoprecipitation of oligomers

We coupled Fab 3D6 to Ni-conjugated magnetic beads in 700 μ l PBST (PBS-tween) buffer for two hours at RT. We washed the beads twice with PBST buffer to remove excess of Fab. We then added oligomer solution to Fab-bound beads. We incubated the bead-Fab-oligomer solution for two hours at room temperature with mild shaking, washing the solution twice with 200 μ l

buffer before eluting using 50 mM tris, pH 7.5, 150 mM NaCl and 500 mM imidazole buffer.

We tested various molar ratios of Fab and oligomers by adjusting the volume of Fab or oligomers used in the immunoprecipitation. We buffer exchanged the elution into 20 mM Tris-HCl pH 7.4, 150 mM NaCl using an Amicon 100 kDa spin filter before preparing grids for electron microscopy.

Immunoprecipitation of oligomers

We mixed Fab 3D6 to Ni-conjugated magnetic beads in XXX buffer and then added oligomer solution to Fab-bound beads. We incubated the bead-Fab-oligomer solution for four hours at room temperature with mild shaking, washing the solution twice with XXX buffer before eluting using XXX buffer. We tested various molar ratios of Fab and oligomers by adjusting the volume of Fab or oligomers used in the immunoprecipitation. We buffer exchanged the elution into 20 mM Tris-HCl pH 7.4, 150 mM NaCl using an Amicon 100 kDa spin filter before preparing grids for electron microscopy.

Negative stain electron microscopy

2 μ L of sample was added to glow-discharged Formvar Carbon Cu grids (EMS) and incubated for 1 min. 90% of sample was blotted off using filter paper and 4 μ L of 2% uranyl acetate was added for 2 min. 90% of uranyl acetate was removed through blotting and a final 4 μ L was added and removed immediately (wash). Negative stain grids were imaged in either a FEI T12, T20, or TF20 microscope.

Cryo-electron microscopy

For our first two cryo-EM samples (Figures 1 and 2), we used UltraAuFoil R2/2 or Quantifoil R1.2/1.3 grids, respectively. Samples were aliquoted to glow-discharged grids, blotted in a Vitrobot Mark IV (FEI), and plunge-frozen in liquid ethane. For the A β S26s oligomer-Fab 3D6

complex, we used homemade graphene oxide over Quantifoil R1.2/1.3 grids prepared as follows: (i) we diluted 4 mg/mL graphene oxide (Sigma) to 0.2 mg/mL using ddH₂O and sonicated the diluted solution for 30s in a water bath; (ii) we centrifuged the sonicated solution at 300 rcf for 30s to remove large aggregates; (iii) we then diluted the sample further to ~0.03 mg/mL and added 1 μ L of the final solution to glow discharged grids until the solution evaporated; (iv) after drying, we examined the grids in the electron microscope for graphene oxide hole coverage and the presence of large aggregates, and adjusted the final dilution step (step iii), diluting more if there were too many multi-layer graphene oxide flakes over the holes or diluting less if there was too little single-layer graphene oxide coverage of holes. We collected cryo-EM images using a Titan Krios (FEI) equipped with a Gatan K3 Bioquantum direct electron detector/energy filter system and using an accelerating voltage of 300 kV and slit width of 20 eV. Automated data collection was performed using SerialEM or Legikon. Pre-processing for all data-sets was performed using a combination of WARP¹⁶, CTFFIND4¹⁸, and RELION⁷ for CTF estimation and motion correction. We used a combination of WARP, crYOLO¹⁹, and MicrographCleaner²⁰ to pick particles and Cryosparc⁶ and RELION for 2D classification, ab initio reconstruction, and 3D classification.

Author Contributions

D.R.B., R.A., and S.J. performed oligomer growth and characterization experiments. R.A. obtained antibodies and performed pulldown experiments. D.R.B, S.J., and R.A. performed negative stain EM. D.R.B. collected and processed negative stain EM dataset. K.Z. assisted in automated data collection of oligomers without antibodies. D.R.B. prepared cryo-EM grids and performed cryo-EM data collection and processing. P.G. assisted in automated data collection for

negative stain and cryo-EM of A-beta S26s oligomers-Fab 3D6 particles. L.A.F. synthesized and purified A-beta S26s peptide. C.L. performed mass spectrometry experiments. J.R. and D.S.E. oversaw the project.

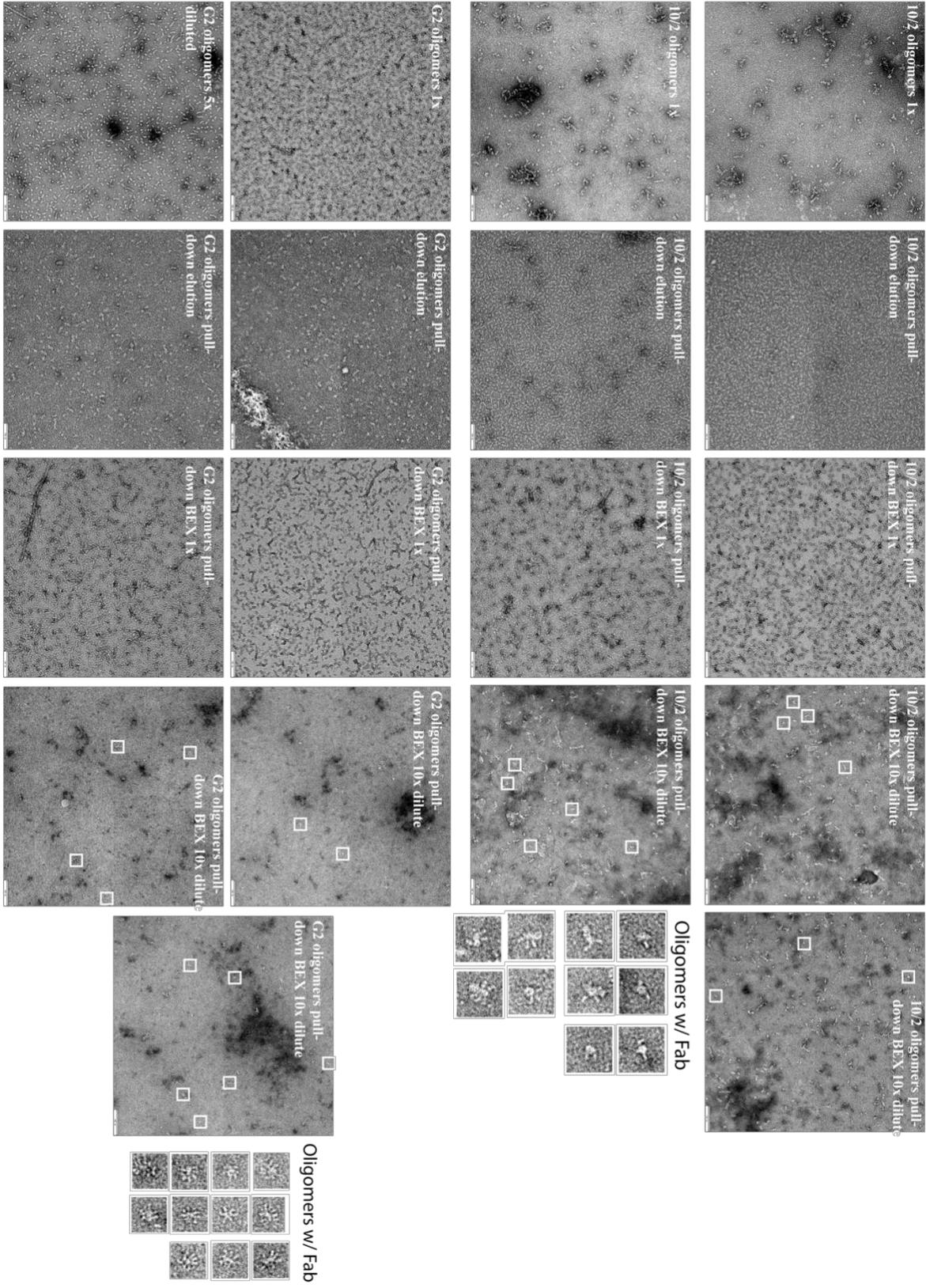
References

1. Auer, S., Ricchiuto, P. & Kashchiev, D. Two-Step Nucleation of Amyloid Fibrils: Omnipresent or Not? *Journal of Molecular Biology* **422**, 723–730 (2012).
2. Knowles, T. P. J., Waudby, C. A., Devlin, G. L., Cohen, S. I. A., Aguzzi, A., Vendruscolo, M., Terentjev, E. M., Welland, M. E. & Dobson, C. M. An Analytical Solution to the Kinetics of Breakable Filament Assembly. *Science* **326**, 1533–1537 (2009).
3. Michaels, T. C. T., Šarić, A., Curk, S., Bernfur, K., Arosio, P., Meisl, G., Dear, A. J., Cohen, S. I. A., Dobson, C. M., Vendruscolo, M., Linse, S. & Knowles, T. P. J. Dynamics of oligomer populations formed during the aggregation of Alzheimer's A β 42 peptide. *Nature Chemistry* **12**, 445–451 (2020).
4. Scheres, S. H. W. Chapter Six - Processing of Structurally Heterogeneous Cryo-EM Data in RELION. in *Methods in Enzymology* (ed. Crowther, R. A.) vol. 579 125–157 (Academic Press, 2016).
5. Foley, A. R., Finn, T. S., Kung, T., Hatami, A., Lee, H.-W., Jia, M., Rolandi, M. & Raskatov, J. A. Trapping and Characterization of Nontoxic A β 42 Aggregation Intermediates. *ACS Chem. Neurosci.* (2019) doi:10.1021/acchemneuro.9b00340.
6. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290–296 (2017).

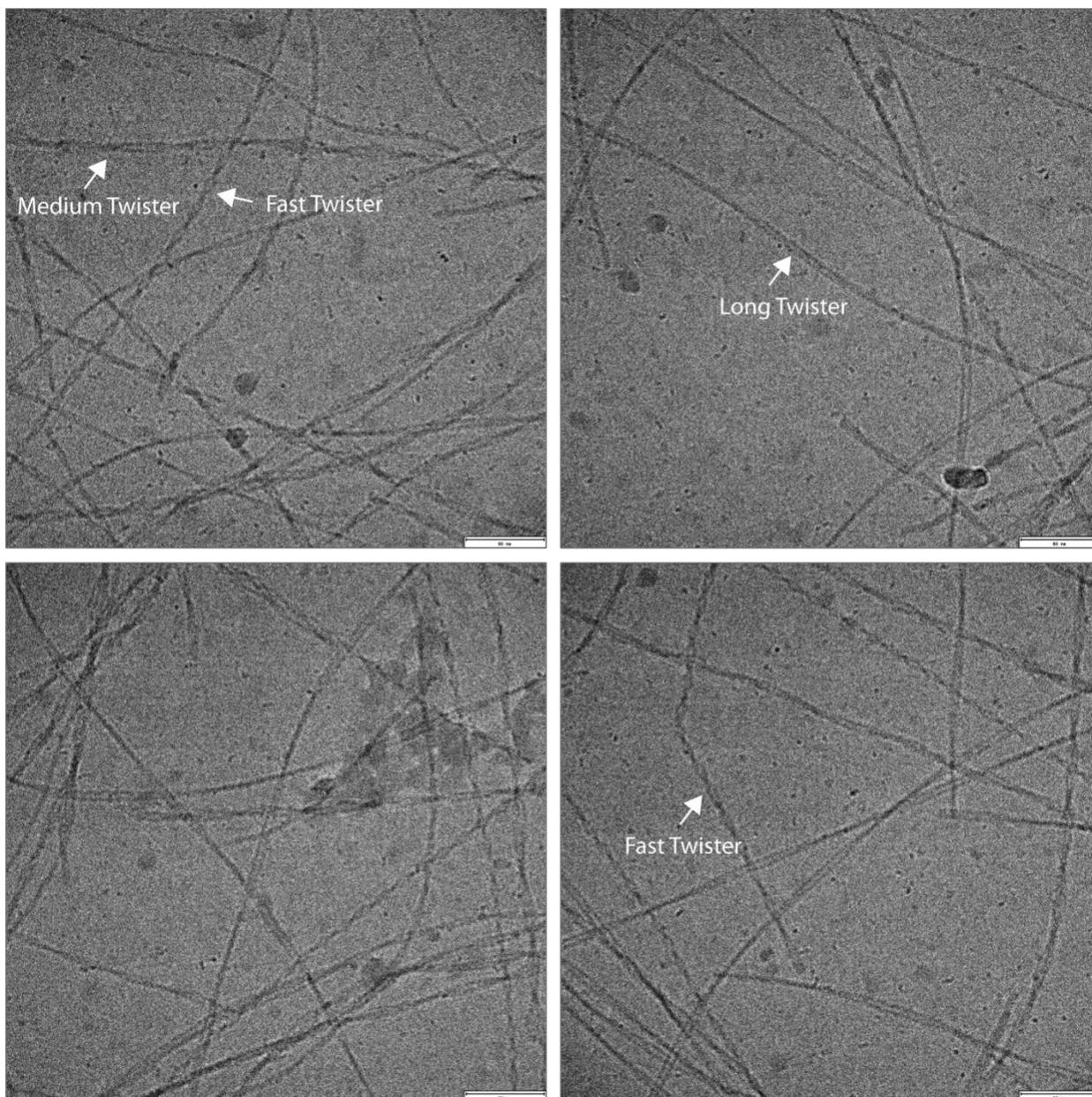
7. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519–530 (2012).
8. Mukherjee, S., Erramilli, S. K., Ammirati, M., Alvarez, F. J. D., Fennell, K. F., Purdy, M. D., Skrobek, B. M., Radziwon, K., Coukos, J., Kang, Y., Dutka, P., Gao, X., Qiu, X., Yeager, M., Eric Xu, H., Han, S. & Kossiakoff, A. A. Synthetic antibodies against BRIL as universal fiducial marks for single-particle cryoEM structure determination of membrane proteins. *Nature Communications* **11**, 1598 (2020).
9. Ghosh, U., Thurber, K. R., Yau, W.-M. & Tycko, R. Molecular structure of a prevalent amyloid- β fibril polymorph from Alzheimer's disease brain tissue. *Proc Natl Acad Sci U S A* **118**, (2021).
10. Goldschmidt, L., Teng, P. K., Riek, R. & Eisenberg, D. Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci U S A* **107**, 3487–3492 (2010).
11. Feinberg, H., Saldanha, J. W., Diep, L., Goel, A., Widom, A., Veldman, G. M., Weis, W. I., Schenk, D. & Basu, G. S. Crystal structure reveals conservation of amyloid- β conformation recognized by 3D6 following humanization to bapineuzumab. *Alzheimers Res Ther* **6**, 31 (2014).
12. Liu, Y.-T., Jih, J., Dai, X., Bi, G.-Q. & Zhou, Z. H. Cryo-EM structures of herpes simplex virus type 1 portal vertex and packaged genome. *Nature* **570**, 257–261 (2019).
13. Paraschiv, G., Vincke, C., Czaplewska, P., Manea, M., Muyldermans, S. & Przybylski, M. Epitope structure and binding affinity of single chain llama anti- β -amyloid antibodies revealed by proteolytic excision affinity-mass spectrometry. *Journal of Molecular Recognition* **26**, 1–9 (2013).

14. Turk, M. & Baumeister, W. The promise and the challenges of cryo-electron tomography. *FEBS Lett* **594**, 3243–3261 (2020).
15. Himes, B. A. & Zhang, P. emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. *Nat Methods* **15**, 955–961 (2018).
16. Tegunov, D., Xue, L., Dienemann, C., Cramer, P. & Mahamid, J. Multi-particle cryo-EM refinement with M visualizes ribosome-antibiotic complex at 3.5 Å in cells. *Nat Methods* **18**, 186–193 (2021).
17. Boyer, D. R., Li, B., Sun, C., Fan, W., Zhou, K., Hughes, M. P., Sawaya, M. R., Jiang, L. & Eisenberg, D. S. The α -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. *Proc Natl Acad Sci U S A* **117**, 3592–3602 (2020).
18. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
19. Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., Quentin, D., Roderer, D., Tacke, S., Siebolds, B., Schubert, E., Shaikh, T. R., Lill, P., Gatsogiannis, C. & Raunser, S. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications Biology* **2**, 218 (2019).
20. Sanchez-Garcia, R., Segura, J., Maluenda, D., Sorzano, C. O. S. & Carazo, J. M. MicrographCleaner: A python package for cryo-EM micrograph cleaning using deep learning. *Journal of Structural Biology* **210**, 107498 (2020).

Supplementary Information



Supplementary Figure 1 Comparison of oligomer pulldown experiments. The 10/02 oligomers were grown without seeding and the pulldown used less Fab (1:15 Fab:A-beta S26s monomer equivalent). Growing without seeding often leads to more protofibrils and more clumping of oligomers, while using less Fabs leads to incomplete labeling of oligomers. The G2 seeded condition grows more homogenous and discrete oligomers with less clumping, and the use of more Fab for the pulldown (1:2.5 Fab:A-beta S26s monomer equivalent) leads to more complete labeling of oligomers.



Supplementary Figure 2 Cryo-EM images of A-beta S26s fibrils after extended incubation shows highly polymorphic fibrils.