

**Title: The  $\alpha$ -synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure**

**Authors:** David R. Boyer<sup>1,†</sup>, Binsen Li<sup>2,†</sup>, Chuanqi Sun<sup>2</sup>, Weijia Fan<sup>2</sup>, Kang Zhou<sup>3</sup>, Michael P. Hughes<sup>1,4</sup>, Michael R. Sawaya<sup>1</sup>, Lin Jiang<sup>2,\*</sup>, David S. Eisenberg<sup>1,\*</sup>

**Affiliations:**

<sup>1</sup> Departments of Chemistry and Biochemistry and Biological Chemistry, UCLA-DOE Institute, Molecular Biology Institute, and Howard Hughes Medical Institute, UCLA, Los Angeles, CA 90095

<sup>2</sup> Department of Neurology, David Geffen School of Medicine, Molecular Biology Institute, UCLA, Los Angeles, CA 90095 USA

<sup>3</sup> California Nano Systems Institute, UCLA, Los Angeles, CA 90095 USA

<sup>4</sup> Current address: Department of Cell & Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN 38105 USA

†These authors contributed equally to this work.

\*Correspondence to: Lin Jiang, email: [jianglin@ucla.edu](mailto:jianglin@ucla.edu); David S. Eisenberg, email:

[david@mbi.ucla.edu](mailto:david@mbi.ucla.edu)

**Significance Statement:**

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  $\alpha$ -synuclein, structural understanding of these amyloid

aggregates may be the path to eventual therapies. Our study of the structure of a variant  $\alpha$ -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.

### **Abstract:**

Aggregation of  $\alpha$ -synuclein is a defining molecular feature of Parkinson's disease, Lewy Body Dementia, and Multiple Systems Atrophy. Hereditary mutations in  $\alpha$ -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  $\alpha$ -synuclein fibril structure is fundamental to  $\alpha$ -synuclein biology. We therefore determined the cryoEM structure of  $\alpha$ -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 forms 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  $\alpha$ -synuclein into a kinetic trap – a shallower, more accessible energy minimum. The E46K mutation apparently unlocks a more stable and pathogenic fibril structure.

### **Introduction:**

The group of diseases termed the synucleinopathies - Parkinson's Disease (PD), Lewy Body Dementia (LBD), and Multiple Systems Atrophy (MSA) – is thought to be caused by the aggregation of  $\alpha$ -synuclein ( $\alpha$ -syn) into amyloid fibrils. The causal relationship between the formation of amyloid fibrils of  $\alpha$ -syn and the synucleinopathies is supported by several observations. Aggregated  $\alpha$ -syn is a major component of Lewy Bodies, the hallmark lesion in PD and LBD, and the hallmark lesions of MSA<sup>1,2</sup>. Hereditary mutations in  $\alpha$ -syn are linked to familial forms of PD and LBD<sup>3</sup>. Overexpression of wild-type  $\alpha$ -syn via dominantly inherited duplications and triplications of the gene that encodes  $\alpha$ -syn, SNCA, are sufficient to cause PD<sup>4-6</sup>. Further, the injection of fibrils of  $\alpha$ -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<sup>8</sup>. Although it is never fully possible to establish causation, these combined observations suggest the case is strong for the linkage of aggregated  $\alpha$ -syn to the synucleinopathies.

To gain a molecular level understanding of amyloid fibrils of  $\alpha$ -syn, we previously applied cryo-EM to determine the near-atomic structures of fibrils of recombinantly assembled  $\alpha$ -syn<sup>7</sup>. We observed two distinct structures – termed the rod and the twister – that share a similar structural kernel formed by residues 50-77 but differ in their protofilament interfaces and flanking regions of the fibril core (Figure 1 b, Figure 4 b-c, Supplementary Figure 1 a-b). 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- $\beta$ <sup>8-11</sup>. For tau protein, the different polymorphs observed correspond to distinct diseases, namely Alzheimer's, Pick's disease, and chronic traumatic encephalopathy<sup>8-10</sup>. Although the structures of  $\alpha$ -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<sup>12</sup>. 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  $\alpha$ -syn and disease, it is unknown how they might exert their effects. One hypothesis is that hereditary mutations may encode new structures of  $\alpha$ -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 sequence. Instead, we label them as “quasi-polymorphs” – different structures adopted by a protein pair with almost identical sequences.

Several hereditary mutations in  $\alpha$ -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<sup>3</sup>. 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  $\alpha$ -syn<sup>13</sup>. Indeed, solid state NMR studies of  $\alpha$ -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<sup>14</sup>.

Consistent with the evidence that E46K alters  $\alpha$ -syn fibril structure and disease manifestation, E46K has been shown to increase the pathogenicity of  $\alpha$ -syn fibrils compared to wild-type. *In vitro* studies have shown that E46K results in an increase in  $\alpha$ -syn's phospholipid binding ability and an enhancement of fibril formation<sup>15</sup>. In addition, E46K promotes higher levels of aggregation in cultured cells relative to wild-type, A53T, and A30P  $\alpha$ -syn<sup>16</sup>. Further, others have found that  $\alpha$ -syn bearing E46K is more toxic to rat primary neurons compared to wild-type, A30P, and A53T<sup>17</sup>.

Three previous structures of full-length wild-type non-acetylated as well as three structures of acetylated full-length and C-terminally truncated  $\alpha$ -syn reveal that E46 participates in a conserved salt-bridge with K80 (Supplementary Figure 1 a)<sup>7,18-21</sup>. 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 ssNMR chemical shifts of E46K  $\alpha$ -syn fibrils. Therefore, we sought to determine the structure of  $\alpha$ -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  $\alpha$ -syn, we also find that E46K fibrils are more powerful seeds in  $\alpha$ -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  $\alpha$ -syn fibrils.

## **Results:**

### **Cryo-EM structure determination and architecture of E46K $\alpha$ -syn fibrils**

We purified  $\alpha$ -syn bearing the E46K hereditary mutation and generated amyloid fibrils using the same growth conditions as in our previous studies<sup>7</sup>. We then optimized cryo-EM grids of E46K fibrils and imaged them at 165,000x magnification on an energy-filtered Titan Krios equipped with a K2 Summit direct electron detector operating in super-resolution mode. We also took advantage of image shift induced-beam tilt correction in SerialEM during data collection to maintain coma-free alignment during data collection<sup>22</sup>. Using helical refinement procedures in Relion 3.0, we obtained a 2.5 Å resolution reconstruction of E46K  $\alpha$ -syn fibrils (Figure 2 a-e, Supplementary Figure 2 a-c, Supplementary Figure 3, Table 1).

The high resolution allowed us to confidently assign side chain rotamer and carbonyl positions, confirming the left handedness and the internal C<sub>2</sub> symmetry of the helical reconstruction, and to build an unambiguous atomic model *de novo* (Figure 2 b-e, Supplementary Figure 3, see *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 (Figure 2 d-e, Supplementary Figure 3). 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 (Figure 2 b-c, Figure 3 a).

## Comparison of E46K and wild-type $\alpha$ -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 new quasi-polymorph of  $\alpha$ -syn. Figure 1 a-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 50-77, in which  $\beta$ -strands  $\beta$ 3,  $\beta$ 4, and  $\beta$ 5 in the rod and  $\beta$ 2,  $\beta$ 3, and  $\beta$ 4 in the twister form a bent  $\beta$ -arch (Figure 2). The twister has fewer ordered residues at the N- and C-termini of its fibril core compared to the rod (Figure 1 a-b, Supplementary Figure 1). The C-terminal residues in the rod structure form a Greek-key like fold comprising  $\beta$ 5,  $\beta$ 6,  $\beta$ 7, and  $\beta$ 8. However, in the E46K structure, the kernel and Greek-key are not maintained and a new 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 (Figure 2 b). 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 new set of residues form sheet-sheet interfaces within the protofilament (Figure 2). Now tightly mated heterozippers are formed by  $\beta$ -strands  $\beta$ 1 and  $\beta$ 6,  $\beta$ 3 and  $\beta$ 4, and a roughly triangular-shaped bent  $\beta$ -arch fold is formed by  $\beta$ 4,  $\beta$ 5,  $\beta$ 6, and  $\beta$ 7 (Figure 1).

Supplementary Figure 4 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 Å<sup>2</sup> for E46K; 7,605 Å<sup>2</sup> for wild-type rod; 5,082 Å<sup>2</sup> 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 (Supplementary Figure 1 a-c, Table 2). To verify the energy estimate, we performed an SDS denaturation assay where both wild-type and E46K fibrils were incubated with various concentrations of SDS at 37 °C followed by 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 (Supplementary Figure 5 b).

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 (Figure 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 (Figure 3 a-c). Therefore, in line with the observation above that the E46K structure has overall more inter-residue 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,  $\beta$ -strand interrupting loops and turns are formed by KTKE pseudo-repeats (residues 43-46, 58-61, 80-83) and glycines (residues 51, 67-68, 73, 84, and 93) while other residues form  $\beta$ -strands (Figure 1 a-b). Indeed, if one compares the structures in Figure 1 b, especially the wild-type rod and E46K, loops and turns between  $\beta$ -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),  $\beta$ 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 a  $\sim 90^\circ$  turn in the rod and a  $\sim 180^\circ$  turn in the E46K structure. Also, the turn between  $\beta$ 4 and  $\beta$ 5 leads to a  $\sim 180^\circ$  turn in the rod and a  $\sim 100^\circ$  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 (Figure 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 45-57 (buried surface area =  $47.3 \text{ \AA}^2$ ). This is in contrast to the dry, steric zipper-like interfaces formed by the preNAC (buried surface area =  $91.7 \text{ \AA}^2$ ) and NAC (buried surface area =  $65.3 \text{ \AA}^2$ ) 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 (Figure 4 a). Electrostatic zippers have previously been observed to enable counterion-induced DNA condensation whereby anionic phosphates and cations such as  $\text{Mn}^{2+}$  and spermidines alternate along the length of the DNA-DNA interface forming a “zipper” that “fastens” the molecules together<sup>23</sup>. 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 (Figure 4 a). 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  $\alpha$ -syn-A53T-YFP in HEK293T biosensor cells<sup>24</sup>. If seeding occurs, normally diffuse  $\alpha$ -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 (Figure 5 a). Whereas the biosensor cell assay supports that E46K fibrils are better seeds than wild-type fibrils, the biosensor cells express  $\alpha$ -syn-A53T-YFP, and not wild-type  $\alpha$ -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 (Supplementary Figure 5 a). Similar to the biosensor cell assay, E46K fibrils were more efficient than wild-type fibrils in seeding growth of wild-type  $\alpha$ -syn (Supplementary Figure 5 a). Further, the different 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 (Supplementary Figure 5 a). This is similar to what is seen in Watson *et al.* where both non-acetylated and acetylated  $\alpha$ -syn form fibrils to the same extent, yet acetylated  $\alpha$ -syn has a substantially lower ThT signal<sup>25</sup>. We also observe that unseeded E46K has a longer lag phase than unseeded wild-type  $\alpha$ -syn; we speculate why E46K may form fibrils more slowly than wild-type in the 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<sup>26,27</sup>. We observe significantly impaired mitochondrial activity by lower concentrations of E46K fibrils than of wild-type fibrils (Figure 5 b). 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<sup>15-17</sup>.

## **Discussion:**

### **Structural differences help explain enhanced pathogenicity of E46K $\alpha$ -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* (Figure 5 a, Supplementary Figure 5 a). 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 (Figure 3 a, c). Since all of these electrostatic residues are present in the  $\alpha$ -syn-A53T-YFP construct in the biosensor cells and in wild-type  $\alpha$ -syn used in the *in vitro* assay, it is plausible that the electrostatic triads in the E46K fibril guide  $\alpha$ -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<sup>28</sup>. 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 (Figure 5 b).

### **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 new  $\alpha$ -syn quasi-polymorph by facilitating a large rearrangement in fibril structure including a re-packed protofilament fold and a new protofilament interface. The polymorphism displayed by E46K is reminiscent of modal polymorphism, first introduced by Caspar and Cohen<sup>29</sup>, 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<sup>30</sup>. This is consistent with the solvent-facing orientation of sidechain 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<sup>30</sup>, highlights the large degree of sensitivity of  $\alpha$ -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  $\alpha$ -syn is flat with many local minima<sup>31</sup>. Here, a single amino acid change, the E46K mutation, shifts the fibril misfolding pathway of  $\alpha$ -syn to a different minimum in the folding landscape – the first atomic evidence of a hereditary mutation in an amyloid protein to do so. 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, see Table 1). This calculation, along with the fact that the wild-type sequence has been shown to be able to adopt the E46K structure<sup>30</sup>, 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 is 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<sup>32</sup>. 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 pre-fibrillar and fibrillary structures, that lowers the energy barrier to forming the rod structure<sup>7,18,19</sup>. In other words, the early formation of the E46-K80 salt bridge may divert  $\alpha$ -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 50-77 observed in wild-type rod structures<sup>7,18,19</sup>. 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  $\alpha$ -syn to a different misfolding pathway, resulting in the more energetically stable E46K structure. These ideas are summarized in Figure 6, which imagines the misfolding landscape of the  $\alpha$ -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  $\alpha$ -syn aggregates more slowly than wild-type (Supplementary Figure 5 a) 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 (Figure 6).

To date, no other studies of wild-type or mutant  $\alpha$ -syn have revealed the twister polymorph we determined in our initial cryo-EM study<sup>7</sup>. It is especially surprising that E46K  $\alpha$ -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<sup>7</sup>. 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<sup>30</sup>. 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<sup>30</sup>. They 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  $\alpha$ -syn fibril structure to perturbations in sequence or growth environment and that both of these factors play a role in producing new 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 (Supplementary Figure 5 a). 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  $\alpha$ -syn grown in phosphate buffer that resembles the 2a polymorph and our E46K structure determined here<sup>30</sup>. This is identical to the buffer they used previously to grow fibrils of wild-type, C-terminally truncated  $\alpha$ -syn that produced a rod polymorph similar to the one we

determined<sup>7,18</sup>. 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 new 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 new, more stable polymorph.

Initial studies examining the effect of the E46K mutation on recombinantly assembled  $\alpha$ -syn fibrils further emphasize the importance of fibril growth conditions on selecting polymorphs. Whereas we find that E46K  $\alpha$ -syn forms fibrils more slowly than wild-type  $\alpha$ -syn (Supplementary Figure 5 a), Choi, *et al.* found that the E46K mutation accelerates  $\alpha$ -syn fibril formation<sup>15</sup>. However, the fibrils formed in their study appear to have a different morphology than our E46K fibrils<sup>15</sup>. 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.* 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<sup>7</sup>. The observed differences in fibril morphology and aggregation kinetics again highlight the sensitivity of  $\alpha$ -syn fibrillation to growth conditions and hints at the ability of E46K  $\alpha$ -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<sup>31,33</sup>. The prodigious polymorphism observed not only in  $\alpha$ -syn structures determined in our work but in other amyloid structures hints that evolution has not played a role in specifying pathogenic amyloid structures. Therefore, many factors, such as concentration and intra-chain interactions, may guide the sequence to forming a non-evolved structure. The E46-K80 salt bridge may represent one of these intra-chain interactions that diverts wild-type  $\alpha$ -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<sup>34-36</sup>. That work was done on shorter peptide segments however and may not have captured all 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 post-translationally modified amyloid fibrils to identify the possible mechanisms by which hereditary mutations and PTM's can alter amyloid protein pathogenicity. We have previously shown that the H50Q hereditary mutation alters  $\alpha$ -syn's protofilament assembly, resulting in more pathogenic fibrils – a first step in understanding the role of hereditary mutations in fibril structure and activity<sup>37</sup>. Recently, a pair of studies demonstrated that N-terminally acetylated  $\alpha$ -syn has different ThT binding ability and different seeding properties compared to non-

acetylated  $\alpha$ -syn although the structures of non-acetylated and acetylated  $\alpha$ -syn fibril cores remain largely similar<sup>20,25</sup>. Further, ssNMR studies of  $\alpha$ -syn hereditary mutants A30P, which lies outside the fibril core of  $\alpha$ -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<sup>14,38</sup>. Also, hereditary mutations in transthyretin have been discovered that serve to de-stabilize transthyretin's native fold thereby promoting fibril formation<sup>39</sup>. 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, PTM's such as phosphorylation, oxidative stress and truncation have varying effects on  $\alpha$ -syn aggregation and toxicity<sup>40</sup>.

## Conclusion

In summary, we have determined a 2.5 Å resolution reconstruction of recombinantly assembled E46K  $\alpha$ -syn fibrils that provides the first atomic structure of this hereditary mutation initially discovered in a family with a clinical diagnosis of parkinsonism and Lewy Body Dementia<sup>13</sup>. The fibril structure of E46K  $\alpha$ -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  $\alpha$ -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 re-folding to a more stable structure. The release of local constraints to allow re-packing into more stable, pathogenic fibrils may be a mechanism by which other hereditary mutations operate in  $\alpha$ -syn and other amyloid proteins.

## **Acknowledgments**

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 (1S10RR23057 and 1S10OD018111), NSF (DBI-1338135), and CNSI at UCLA. The authors acknowledge 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  $\alpha$ -syn protein. B.L. grew fibrils of  $\alpha$ -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. M.P.H. wrote the software for and D.R.B. carried out the pairwise interaction analysis. 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.

## **Data Availability Statement:**



All structural data have been deposited into the wwPDB and EMDB with the following accession codes: PDB 6UFR and EMDB EMD-20759. All other data are available from the authors upon reasonable request.

## References:

1. Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M. & Goedert, M. 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. Grazia Spillantini, M. *et al.* Filamentous  $\alpha$ -synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies. *Neuroscience Letters* **251**, 205–208 (1998).
3. Spillantini, M. G. & Goedert, M. Neurodegeneration and the ordered assembly of  $\alpha$ -synuclein. *Cell Tissue Res* **373**, 137–148 (2018).
4. Singleton, A. B. *et al.*  $\alpha$ -Synuclein Locus Triplication Causes Parkinson's Disease. *Science* **302**, 841–841 (2003).
5. Chartier-Harlin, M.-C. *et al.* Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* **364**, 1167–1169 (2004).
6. Ibáñez, P. *et al.* Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. *Lancet* **364**, 1169–1171 (2004).
7. Li, B. *et al.* Cryo-EM of full-length  $\alpha$ -synuclein reveals fibril polymorphs with a common structural kernel. *Nat Commun* **9**, (2018).
8. Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185–190 (2017).

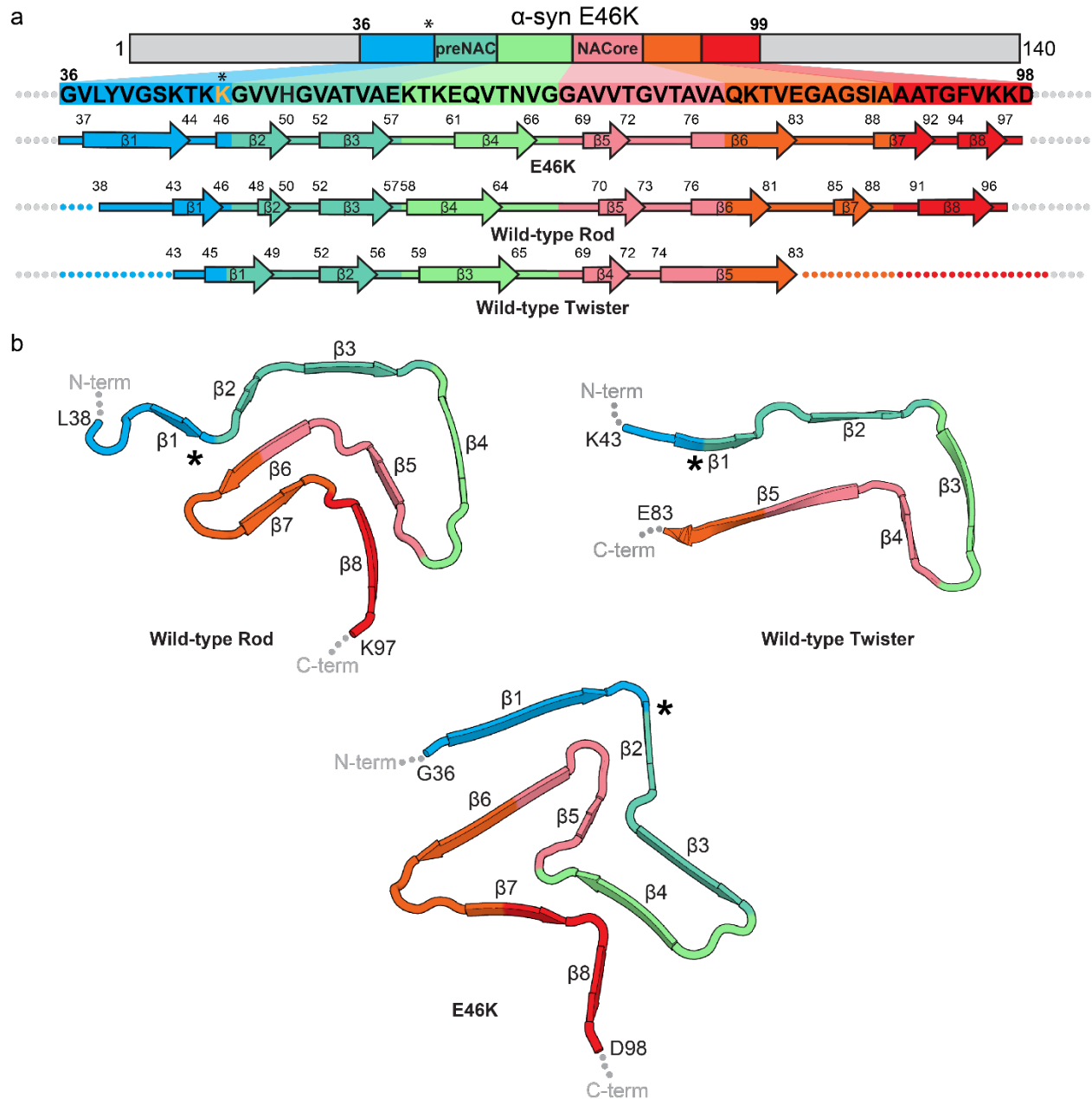
9. Falcon, B. *et al.* Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **561**, 137 (2018).
10. Falcon, B. *et al.* Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019).
11. Tycko, R. Molecular Structure of Aggregated Amyloid- $\beta$ : Insights from Solid-State Nuclear Magnetic Resonance. *Cold Spring Harb Perspect Med* **6**, a024083 (2016).
12. Peng, C. *et al.* Cellular milieu imparts distinct pathological  $\alpha$ -synuclein strains in  $\alpha$ -synucleinopathies. *Nature* **557**, 558–563 (2018).
13. Zarranz, J. J. *et al.* The new mutation, E46K, of  $\alpha$ -synuclein causes parkinson and Lewy body dementia. *Annals of Neurology* **55**, 164–173 (2004).
14. 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).
15. Choi, W. *et al.* Mutation E46K increases phospholipid binding and assembly into filaments of human  $\alpha$ -synuclein. *FEBS Letters* **576**, 363–368 (2004).
16. Pandey, N., Schmidt, R. E. & Galvin, J. E. The alpha-synuclein mutation E46K promotes aggregation in cultured cells. *Experimental Neurology* **197**, 515–520 (2006).
17. Íñigo-Marco, I. *et al.* E46K  $\alpha$ -synuclein pathological mutation causes cell-autonomous toxicity without altering protein turnover or aggregation. *PNAS* **114**, E8274–E8283 (2017).
18. Guerrero-Ferreira, R. *et al.* Cryo-EM structure of alpha-synuclein fibrils. *Elife* **7**, (2018).
19. Li, Y. *et al.* Amyloid fibril structure of  $\alpha$ -synuclein determined by cryo-electron microscopy. *Cell Res.* **28**, 897–903 (2018).

20. Ni, X., McGlinchey, R. P., Jiang, J. & Lee, J. C. Structural Insights into  $\alpha$ -Synuclein Fibril Polymorphism: Effects of Parkinson's Disease-Related C-Terminal Truncations. *J. Mol. Biol.* **431**, 3913–3919 (2019).
21. Tuttle, M. D. *et al.* Solid-state NMR structure of a pathogenic fibril of full-length human  $\alpha$ -synuclein. *Nature Structural & Molecular Biology* **23**, 409–415 (2016).
22. Mastronarde, D. N. SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. *Microscopy and Microanalysis* **9**, 1182–1183 (2003).
23. Kornyshev, A. A. & Leikin, S. Electrostatic Zipper Motif for DNA Aggregation. *Phys. Rev. Lett.* **82**, 4138–4141 (1999).
24. Prusiner, S. B. *et al.* Evidence for  $\alpha$ -synuclein prions causing multiple system atrophy in humans with parkinsonism. *PNAS* **112**, E5308–E5317 (2015).
25. Watson, M. D. & Lee, J. C. N-Terminal Acetylation Affects  $\alpha$ -Synuclein Fibril Polymorphism. *Biochemistry* **58**, 3630–3633 (2019).
26. Wakabayashi, M. & Matsuzaki, K. Formation of Amyloids by A $\beta$ -(1–42) on NGF-differentiated PC12 Cells: Roles of Gangliosides and Cholesterol. *Journal of Molecular Biology* **371**, 924–933 (2007).
27. Ono, K., Condron, M. M. & Teplov, D. B. Structure–neurotoxicity relationships of amyloid  $\beta$ -protein oligomers. *PNAS* **106**, 14745–14750 (2009).
28. Hosp, F. *et al.* Spatiotemporal Proteomic Profiling of Huntington's Disease Inclusions Reveals Widespread Loss of Protein Function. *Cell Reports* **21**, 2291–2303 (2017).
29. Caspar, D. L. & Cohen, C. Polymorphism of tropomyosin and a view of protein function; in *Nobel Symposium 11. Symmetry and Function of Biological Systems at the Macromolecular*

- Level* (eds. Engstrom, A. & Strandberg, B.) 393–414 (Almquist and Wiksell; John Wiley and Sons, Inc., 1969).
30. Guerrero-Ferreira, R. *et al.* Two new polymorphic structures of alpha-synuclein solved by cryo-electron microscopy. *bioRxiv* 654582 (2019) doi:10.1101/654582.
  31. Dill, K. A. & MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science* **338**, 1042–1046 (2012).
  32. Riek, R. & Eisenberg, D. S. The activities of amyloids from a structural perspective. *Nature* **539**, 227–235 (2016).
  33. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
  34. Guenther, E. L. *et al.* 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.
  35. Warmack, R. A. *et al.* Structure of amyloid- $\beta$  (20-34) with Alzheimer's-associated isomerization at Asp23 reveals a distinct protofilament interface. *Nat Commun* **10**, 1–12 (2019).
  36. 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).
  37. Boyer, D. R. *et al.* Structures of fibrils formed by  $\alpha$ -synuclein hereditary disease mutant H50Q reveal new polymorphs. *Nat Struct Mol Biol* 1–9 (2019) doi:10.1038/s41594-019-0322-y.
  38. Lemkau, L. R. *et al.* Mutant protein A30P  $\alpha$ -synuclein adopts wild-type fibril structure, despite slower fibrillation kinetics. *J. Biol. Chem.* **287**, 11526–11532 (2012).

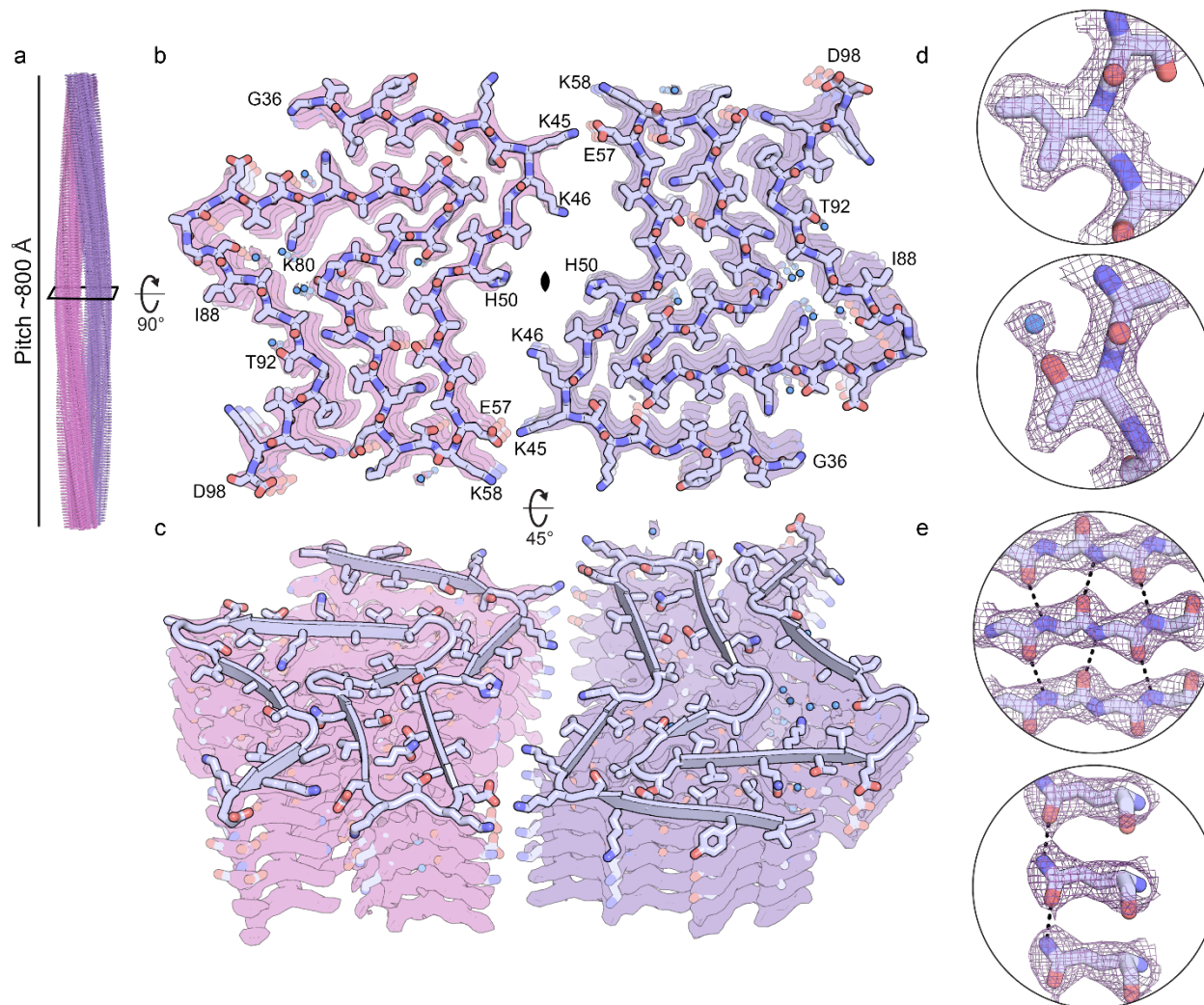
39. Hurshman Babbes, A. R., Powers, E. T. & Kelly, J. W. Quantification of the thermodynamically linked quaternary and tertiary structural stabilities of transthyretin and its disease-associated variants: the relationship between stability and amyloidosis. *Biochemistry* **47**, 6969–6984 (2008).
40. Lashuel, H. A., Overk, C. R., Oueslati, A. & Masliah, E. The many faces of  $\alpha$ -synuclein: from structure and toxicity to therapeutic target. *Nature Reviews Neuroscience* **14**, 38–48 (2013).

## Figures and Tables:



determined wild-type folds (rod and twister) and E46K protofilament fold determined here<sup>7</sup>.

Asterisks indicate the location of residue 46.



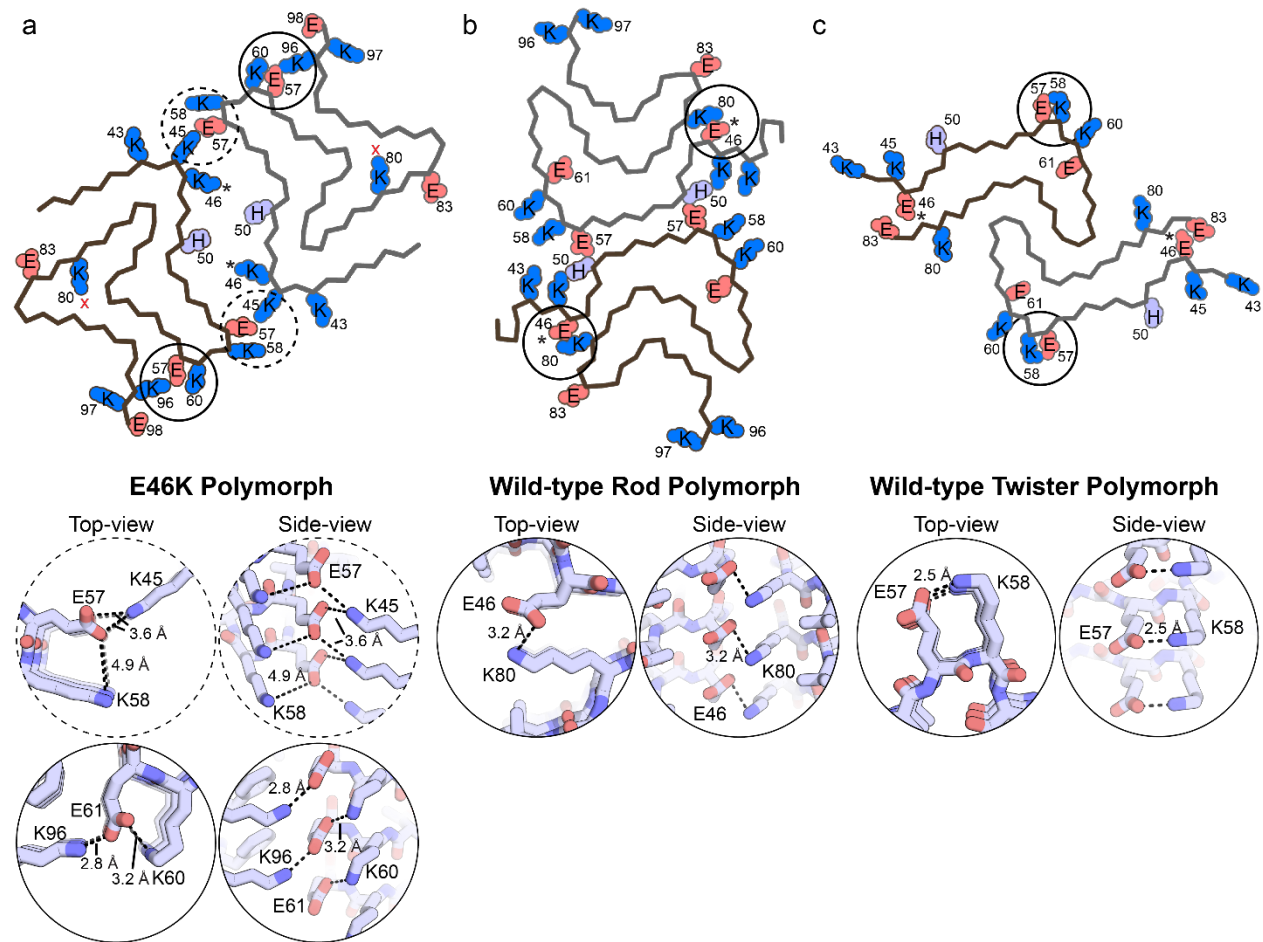
**Figure 2** Overview of E46K cryo-EM structure

a) E46K fibril side-view demonstrating a pitch of  $\sim 800$  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 36-98. c) Tilted view of E46K fibril cross-section demonstrating the stacking of identical layers. d-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).



**Figure 3** Electrostatic residues and interactions of wild-type and E46K fibrils

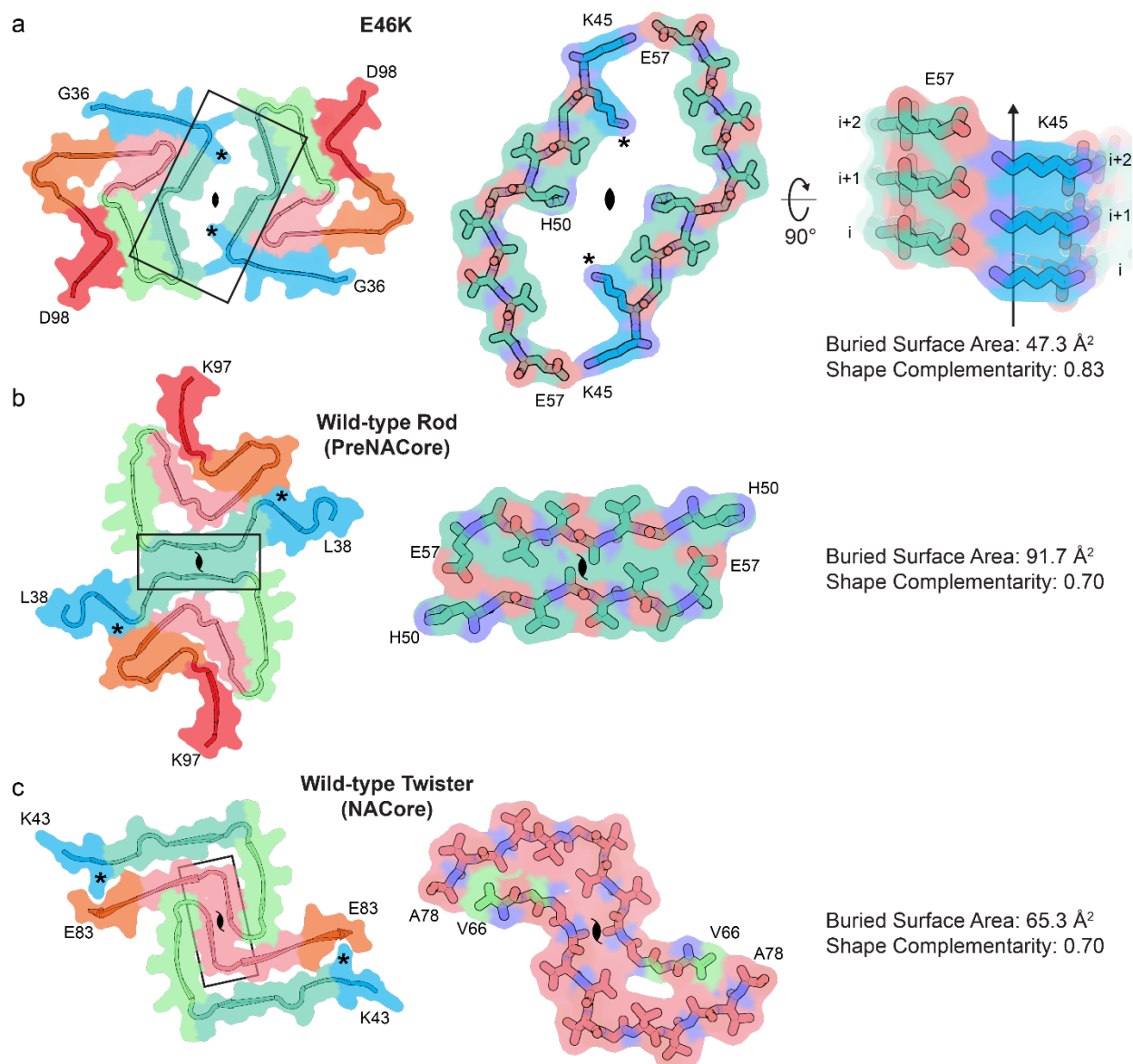
a) Overview of E46K quasi-polymorph fold with charged and ionizable residues shown (top). Top- and side-view of the K45-E57-K58 and K60-E61-K96 electrostatic triads (bottom). The E46K fibril core has a net charge of +11 per layer ( $2 \times 8$  lysine/per chain +  $2 \times 1/2$  histidine per chain –  $2 \times 3$  glutamate per chain) and 8 charged pairs per layer ( $2 \times \text{K45-E57} + 2 \times \text{E57-E58} + 2 \times \text{K60-E61} + 2 \times \text{E61-K96}$ ). b) Overview of the wild-type rod polymorph fold with charged and ionizable residues shown (top). Top- and side-view of E46-K80 salt bridge (bottom). The wild-



type rod fibril core has a net charge of +7 per layer ( $2 \times 7$  lysine/per chain +  $2 \times 1/2$  histidine per chain –  $2 \times 4$  glutamate per chain) and two charge pairs per layer ( $2 \times \text{E46-K80}$  per chain). c)

Overview of wild-type twister polymorph fold with charged and ionizable residues shown (top).

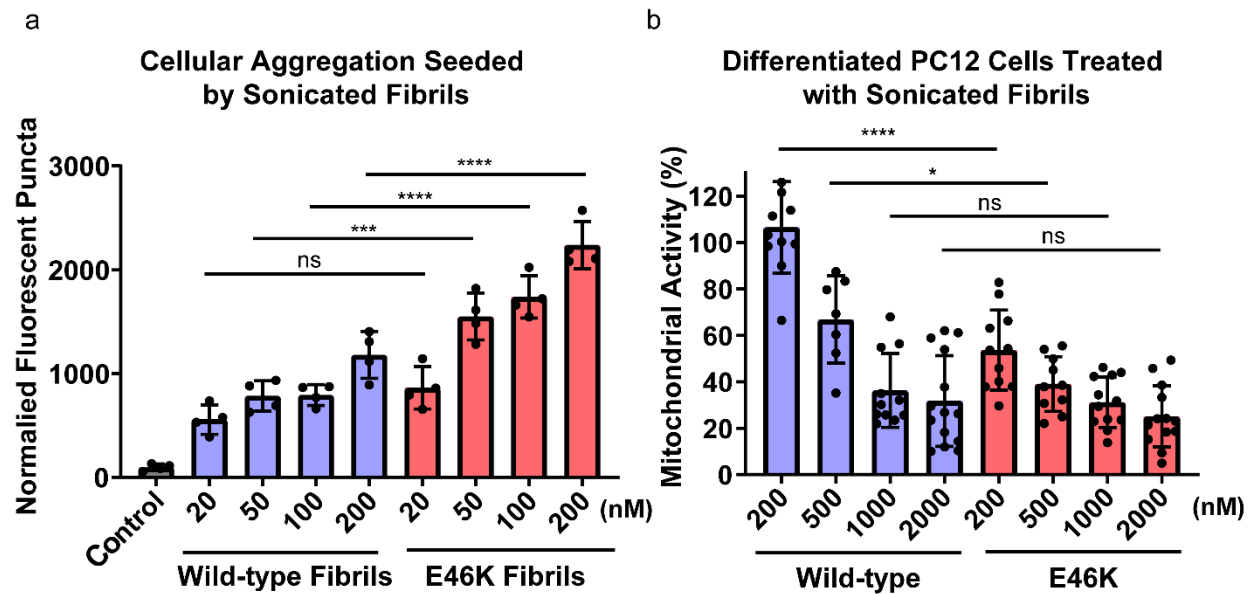
Top- and side-view of E57-K80 salt bridge (bottom). The wild-type twister fibril core has a net charge of +3 per layer ( $2 \times 5$  lysine/per chain +  $2 \times 1/2$  histidine per chain –  $2 \times 4$  glutamate per chain) and two charge pairs per layer ( $2 \times \text{E57-K58}$  per chain). Asterisks denote the location of residue 46. Red x's denote location of the ordered water molecules hydrogen bonding to K80 in the E46K structure.



**Figure 4** Comparison of protofilament interfaces of E46K and wild-type fibrils

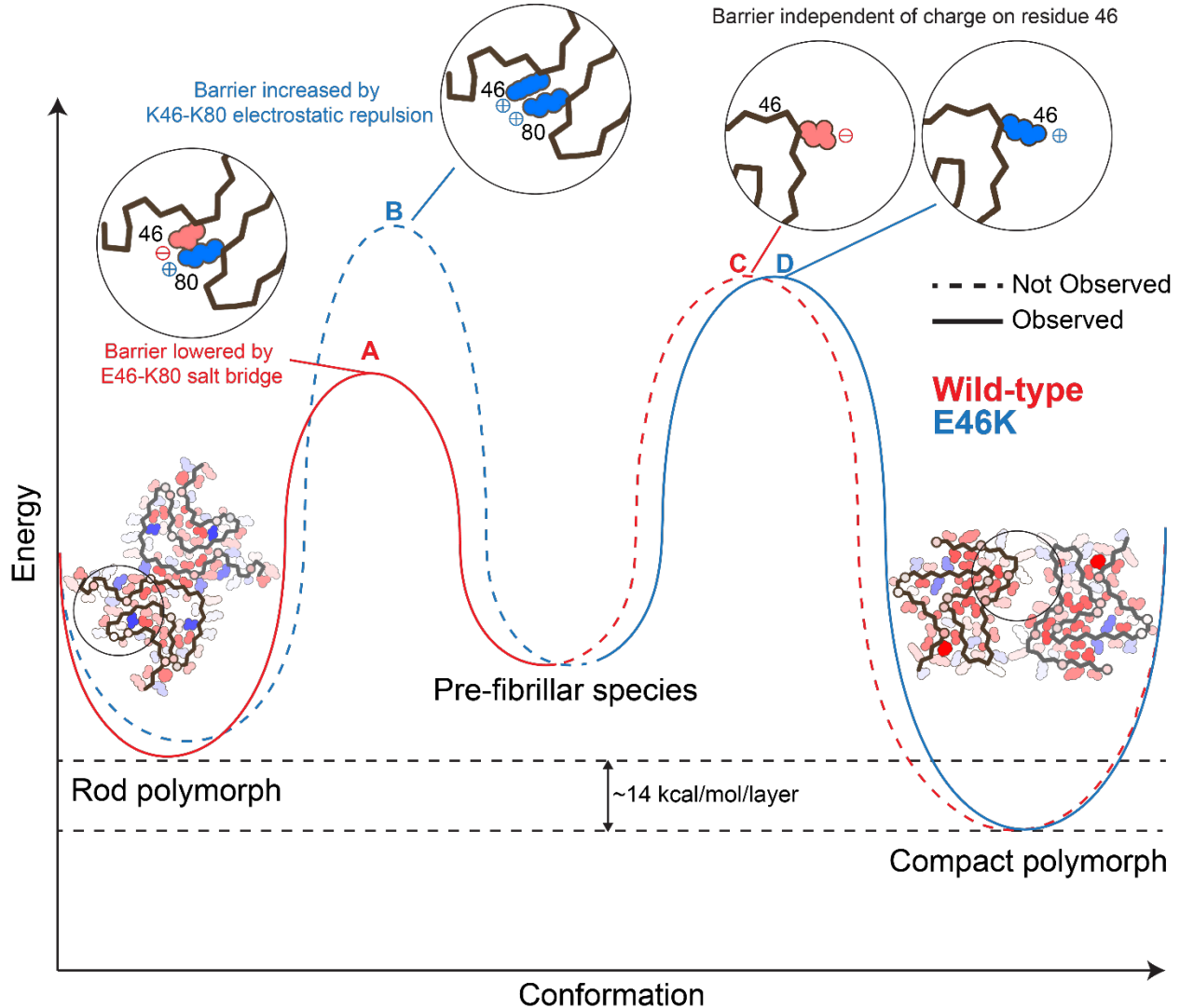
a) The E46K quasi-polymorph forms a two-fold symmetric protofilament interface comprised of residues Lys45-Glu57. The interface is largely solvent-filled and is held together by electrostatic interaction of Lys45 and Glu57 at both ends of the interface. The Lys45-Glu57 salt bridge is repeated along the length of the fibril creating an electrostatic zipper. b) The wild-type rod polymorph forms a pseudo-2<sub>1</sub> symmetric dry steric zipper interface comprised of residues His50-

E57 from the preNAC region. c) The wild-type twister polymorph forms a pseudo- $2_1$  symmetric dry steric zipper interface comprised of residues Val66-Ala78 from the NACore.



**Figure 5** Biochemical analysis of E46K vs. wild-type fibrils

a) Fibrils of E46K and wild-type  $\alpha$ -syn were sonicated and transfected into HEK293T  $\alpha$ -syn-A53T-YFP biosensor cells, and aggregation is measured by the normally soluble and diffuse intracellular  $\alpha$ -syn-A53T-YFP forming discrete puncta, which are then quantified through fluorescent image analysis<sup>24</sup>. At all concentrations but 20 nM, E46K fibrils are significantly more powerful seeds than wild-type fibrils. Error bars represent the standard deviation 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<sup>26,27</sup>. E46K fibrils require a lower concentration to significantly impair mitochondrial activity compared to wild-type fibrils. Error bars represent the standard deviation of 7-14 technical replicates. \*\*\*\* = p-value  $\leq 0.0001$ . \*\* = p-value  $\leq 0.01$ . ns = p-value  $> 0.05$ . P-values were calculated using an unpaired, two-tailed t-test with a 95% CI.



**Figure 6** Proposed misfolding landscape of  $\alpha$ -syn rod and compact polymorphs

The proposed misfolding landscape of  $\alpha$ -syn rod and compact (including the E46K quasi-polymorph) polymorphs demonstrates that for the wild-type sequence, it is kinetically more feasible to form the lower energy rod structure due to a transition state whose energy is lowered by the E46-K80 salt-bridge (A). The compact polymorph determined as part of this work is also accessible to the wild-type sequence (C), although it is not observed due to the E46-K80 salt bridge acting 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 (B), thus facilitating the formation of the compact polymorph (D).

Name	E46K
PDB ID	6UFR
EMDB ID	EMD-20759
<b>Data collection</b>	
Magnification	×165,000
Defocus range (um)	0.75-3.37
Voltage (kV)	300
Camera	K2 Summit (Quantum LS)
Frame exposure time (s)	0.2
# movie frames	30
Total electron dose (e <sup>-</sup> / Å <sup>2</sup> )	36
Pixel size (Å)	0.838
<b>Reconstruction</b>	
Box size (pixel)	400
Inter-box distance (Å)	33.5
# micrograph collected	3,078
# segments extracted	210,593
# segments after Class2D	N/A
# segments after Class3D	114,260
Resolution (Å)	2.5
Map sharpening B-factor (Å <sup>2</sup> )	-140
Helical rise (Å)	4.85
Helical twist (°)	178.92
Point Group	C <sub>2</sub>
<b>Atomic model</b>	
# non-hydrogen atoms	4330
# protein residues	630
R.m.s.d. bonds (Å)	0.012
R.m.s.d. angles (°)	0.958
Molprobity clashscore, all atoms	4.36
Molprobity score	1.64
Poor rotamers (%)	0
Ramachandran outliers (%)	0
Ramachandran allowed (%)	6.6
Ramachandran favoured (%)	93.4
Cβ deviations > 0.25 Å (%)	0
Bad bonds (%)	0
Bad angles (%)	0

**Table 1** Cryo-EM data collection, refinement, and validation statistics. Information for number of segments after Class2D is not available as only Class3D was performed for 400 pixel box segments (see *Methods*).

<b>Fibril Structure</b>	<b>Atomic solvation standard free energy of stabilization (kcal/mol)</b>		
	<b>Per Layer</b>	<b>Per Chain</b>	<b>Per Residue</b>
E46K	-73.8	-36.9	-0.59
WT CryoEM Rod (6CU7)	-59.6	-29.3	-0.49
WT CryoEM 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

**Table 2** Comparative solvation energies

## **Supporting Information**

### **Methods**

#### **Protein purification**

Full-length  $\alpha$ -syn wild-type and E46K mutant proteins were expressed and purified according to a published protocol<sup>1</sup>. 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 minutes in an iced water bath. After centrifugation, the soluble fraction was heated in boiling water for 10 minutes 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, 1M NaCl, pH 8.0). The eluent was concentrated using Amicon Ultra-15 centrifugal filters (Millipore Sigma, 10 NMWL) 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  $\mu$ M purified monomers, 15mM tetrabutylphosphonium bromide, shaking at 37°C for 2 weeks.

### Fibril seeding aggregation in cells

We performed the biosensor cell seeding assay based on a previously published protocol<sup>2</sup>.

Briefly, the assay works as follows: exogenous, un-labeled fibrils are transfected into HEK293T cells expressing  $\alpha$ -syn-A53T-YFP. Seeded aggregation of endogenously expressed  $\alpha$ -syn-A53T-YFP is monitored by formation of fluorescent puncta. The puncta represent intracellular aggregation of  $\alpha$ -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  $\alpha$ -syn containing the hereditary A53T mutation were grown in DMEM (4mM L-glutamine and 25mM D-glucose) supplemented with 10% FBS, 1% penicillin/streptomycin. Trypsin-treated HEK293T cells were harvested, seeded on flat 96-well plates at a concentration of  $4 \times 10^4$  cells/well in 200  $\mu$ L culture medium per well and incubated in 5% CO<sub>2</sub> at 37°C for 18 hours.

$\alpha$ -syn fibrils were prepared by diluting with Opti-MEM™ (Life Technologies; Carlsbad CA) and sonicating in a water bath sonicator for 10 minutes. Fibril concentration was determined as monomer-equivalent concentration. The fibril samples were then mixed with Lipofectamine™ 2000 (Thermo Fisher Scientific) and incubated for 15 minutes and then added to the cells. The actual volume of Lipofectamine™ 2000 was calculated based on the dose of 1  $\mu$ L per well. After 48 hours of transfection, the cells were trypsinized, transferred to a 96-well round-bottom plate and resuspended in 200  $\mu$ L 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.



### 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<sup>1,3-5</sup>. Use of this neuron-like cell line allows us to obtain a biologically relevant assay for cytotoxicity. For our MTT mitochondrial activity assay, the protocol was adapted from the Provost and Wallert laboratories and was performed in an identical manner to our previous work<sup>5,6</sup>. 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/well in DMEM (Dulbecco's modification of Eagle's medium; 5% fetal bovine serum [FBS], 5% heat-inactivated horse serum, 1% penicillin/streptomycin and 150 ng/mL nerve growth factor 2.5S (Thermo Fisher Scientific). The cells were incubated for 2 days in an incubator with 5% CO<sub>2</sub> at 37°C. The cells were treated with different concentrations of monomer-equivalent  $\alpha$ -syn fibrils (200 nM, 500 nM, 1000 nM, and 2000 nM). After 18 hours of incubation, 20  $\mu$ l of 5 mg/ml MTT was added to every well and the plate was returned to the incubator for 3.5 hours. 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  $\mu$ 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

$\alpha$ -syn wild-type or E46K monomers (100  $\mu$ M) were mixed with 60  $\mu$ M thioflavin-T (ThT) and transferred into a 96-well plate. The signal was monitored using the FLUOstar Omega

Microplate Reader (BMG Labtech, 37 °C with 600 r.p.m. double orbital shaking, ex. 440 nm, em. 490 nm). For the seeding groups, preformed wild-type or E46K fibrils (10 μM) after 10 minutes of water-bath sonication were added to the wild-type  $\alpha$ -syn monomers immediately before beginning the aggregation assay.

### SDS Stability Assay

The wild-type and E46K aggregated  $\alpha$ -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 is measured after 5 minutes of incubation at 37 °C with 600 r.p.m. double orbital shaking. The addition of SDS and ThT measurement is 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

2 μl 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/pixel with a dose per frame 1.2 e-/Å<sup>2</sup>. 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-/Å<sup>2</sup> per image. Automated data collection was driven by the SerialEM automation software package with image shift induced-beam tilt correction<sup>7</sup>.

Micrographs containing crystalline ice were used to estimate the anisotropic magnification distortion using `mag_distortion_estimate`<sup>8</sup>. CTF estimation was performed using CTFFIND 4.1.8 on movie stacks with a grouping of 3 frames and correction for anisotropic magnification distortion<sup>9</sup>. Unblur<sup>10</sup> 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 `e2heliboxer.py`<sup>11</sup>. All fibril particles were first extracted using 1024 pixel box sizes and a 10% inter-box distance then subjected to 2D class averaging. 2D class averages reveal that E46K  $\alpha$ -syn forms fibrils of a single morphology with a pitch of  $\sim 800$  Å (Supplementary Figure 2b). We next extracted all fibrils with a 686 pixel box size and 10% inter-box 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 (Supplementary Figure 2b). Due to the two-fold mirror symmetry present in the 2D class averages, we reasoned that the fibril had a  $C_2$  rotational symmetry. Using a calculated helical twist of  $178.92^\circ$  (given pitch of 800 Å and rise of 4.8 Å) and  $C_2$  rotational symmetry, we carried out 3D class averaging with a single class and a featureless cylinder created by `reliion_helical_toolbox` as an initial model<sup>12</sup>. 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 re-extracted with a 400 pixel box size, 10% inter-box 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 (Supplementary Figure 2c). We sharpened the map using phenix.auto\_sharpen with a sharpening factor of  $-140 \text{ \AA}^2$  and a resolution cutoff of  $2.5 \text{ \AA}$ <sup>13</sup>.

### Atomic model building

We used phenix.map\_to\_model with an input sequence corresponding to E46K  $\alpha$ -syn to build an initial model<sup>13</sup>. 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<sup>14</sup>. We generated a 5-layer model to maintain local contacts between chains in the fibril during structure refinement. We performed automated structure refinement using phenix.real\_space\_refine<sup>13</sup>. We employed hydrogen bond distance and angle restraints for backbone atoms participating in  $\beta$ -sheets and side chain hydrogen bonds during automated refinements. We performed comprehensive structure validation of all our final models in Phenix.

### Energetic calculation

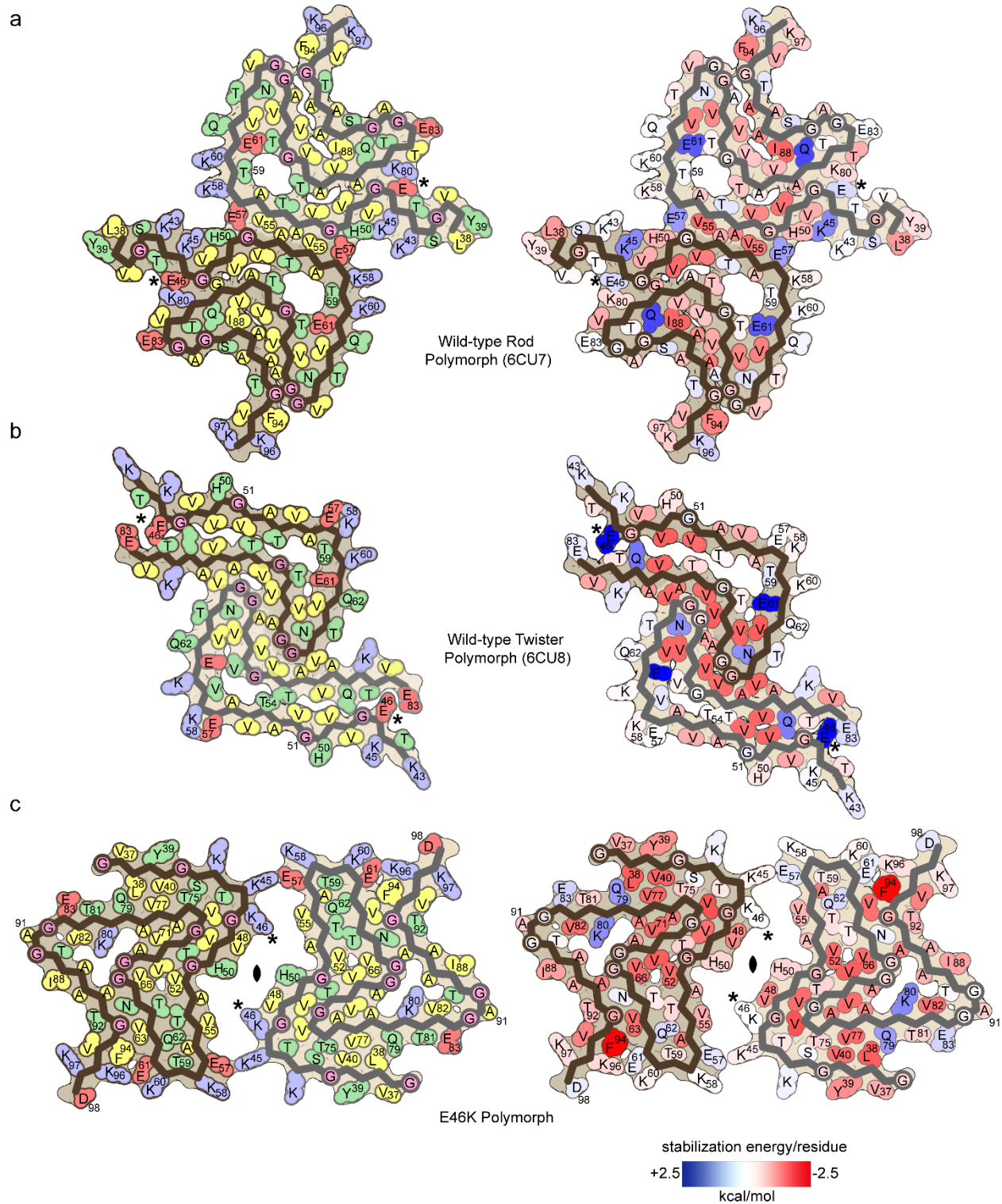
The standard free energy of stabilization of a given amyloid chain is computed as the difference in atomic solvation energy of a pseudo-extended, solvated chain and the folded chain in the center of five layers of the known structure of a protofilament<sup>15</sup>. The atomic solvation parameters are those of Eisenberg *et al.* (1989), with additional terms to describe the entropy change of sidechains on folding, as calculated by Koehl and Delarue (1994), scaled by the percentage of side chain surface area buried<sup>16,17</sup>.

## References:

1. Li, B. *et al.* Cryo-EM of full-length  $\alpha$ -synuclein reveals fibril polymorphs with a common structural kernel. *Nat Commun* **9**, (2018).
2. Prusiner, S. B. *et al.* Evidence for  $\alpha$ -synuclein prions causing multiple system atrophy in humans with parkinsonism. *PNAS* **112**, E5308–E5317 (2015).
3. Wakabayashi, M. & Matsuzaki, K. Formation of Amyloids by A $\beta$ -(1–42) on NGF-differentiated PC12 Cells: Roles of Gangliosides and Cholesterol. *Journal of Molecular Biology* **371**, 924–933 (2007).
4. Ono, K., Condron, M. M. & Teplow, D. B. Structure–neurotoxicity relationships of amyloid  $\beta$ -protein oligomers. *PNAS* **106**, 14745–14750 (2009).
5. Boyer, D. R. *et al.* Structures of fibrils formed by  $\alpha$ -synuclein hereditary disease mutant H50Q reveal new polymorphs. *Nat Struct Mol Biol* 1–9 (2019) doi:10.1038/s41594-019-0322-y.
6. van Meerloo, J., Kaspers, G. J. L. & Cloos, J. Cell Sensitivity Assays: The MTT Assay. in *Cancer Cell Culture: Methods and Protocols* (ed. Cree, I. A.) 237–245 (Humana Press, 2011). doi:10.1007/978-1-61779-080-5\_20.
7. Mastronarde, D. N. SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. *Microscopy and Microanalysis* **9**, 1182–1183 (2003).
8. Grant, T. & Grigorieff, N. Automatic estimation and correction of anisotropic magnification distortion in electron microscopes. *J. Struct. Biol.* **192**, 204–208 (2015).

9. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
10. 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).
11. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
12. He, S. & Scheres, S. H. W. Helical reconstruction in RELION. *Journal of Structural Biology* **198**, 163–176 (2017).
13. Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Cryst D* **74**, 531–544 (2018).
14. Emsley, P., Lohkamp, B., Scott, W. G. & Cowild-typean, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
15. Eisenberg, D. & McLachlan, A. D. Solvation energy in protein folding and binding. *Nature* **319**, 199–203 (1986).
16. Eisenberg, D. S., Wesson, M. & Yamashita, M. Interpretation of protein folding and binding with atomic solvation parameters. *Chemica Scripta* **29A**, 217–222.
17. Koehl, P. & Delarue, M. 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).

## Supplementary Figures and Tables:

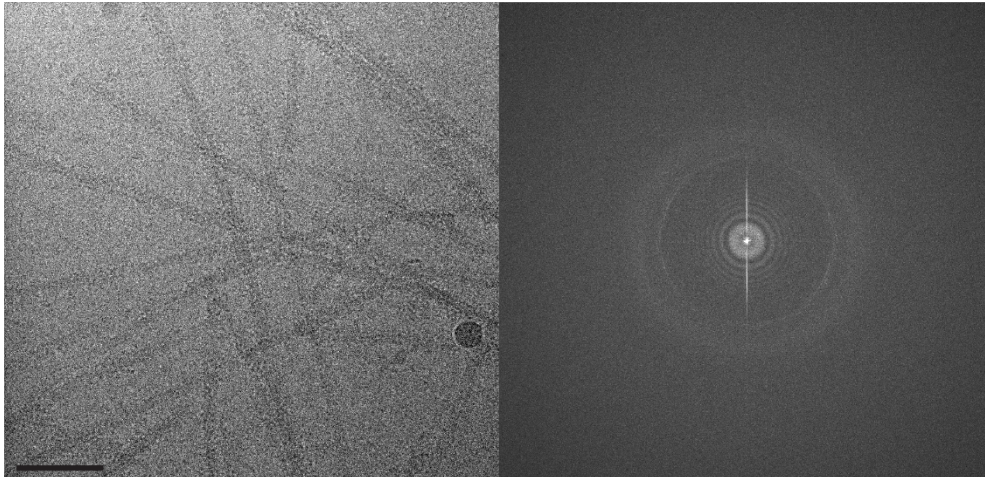


Supplementary Figure 1 Schematic representation and free energy of stabilization maps for the wild-type and E46K polymorphs

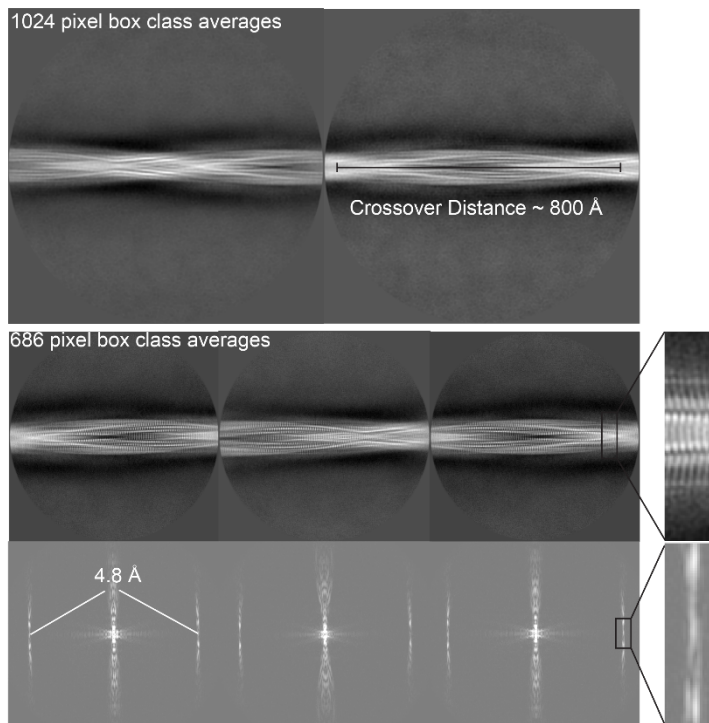
a-c) (left) Schematic representation of fibril structures with amino acid side chains colored as follows: hydrophobic (yellow), negatively charged (red), positively charged (blue), polar, uncharged (green), and glycine (pink). (right) Solvation energy maps of fibril structures. The stabilizing residues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



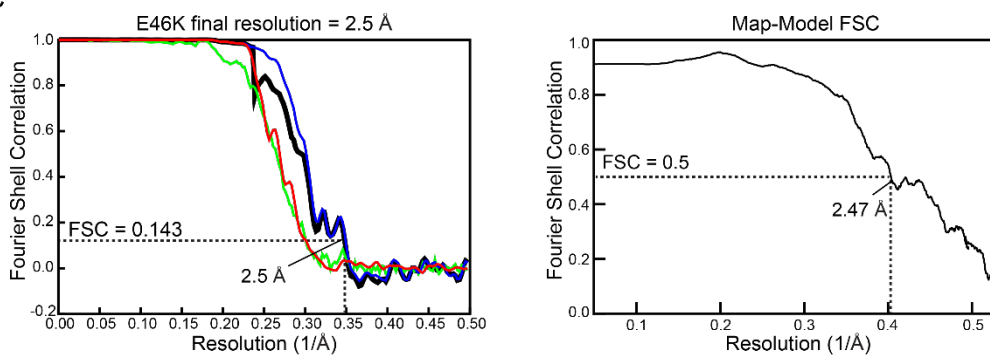
a



b

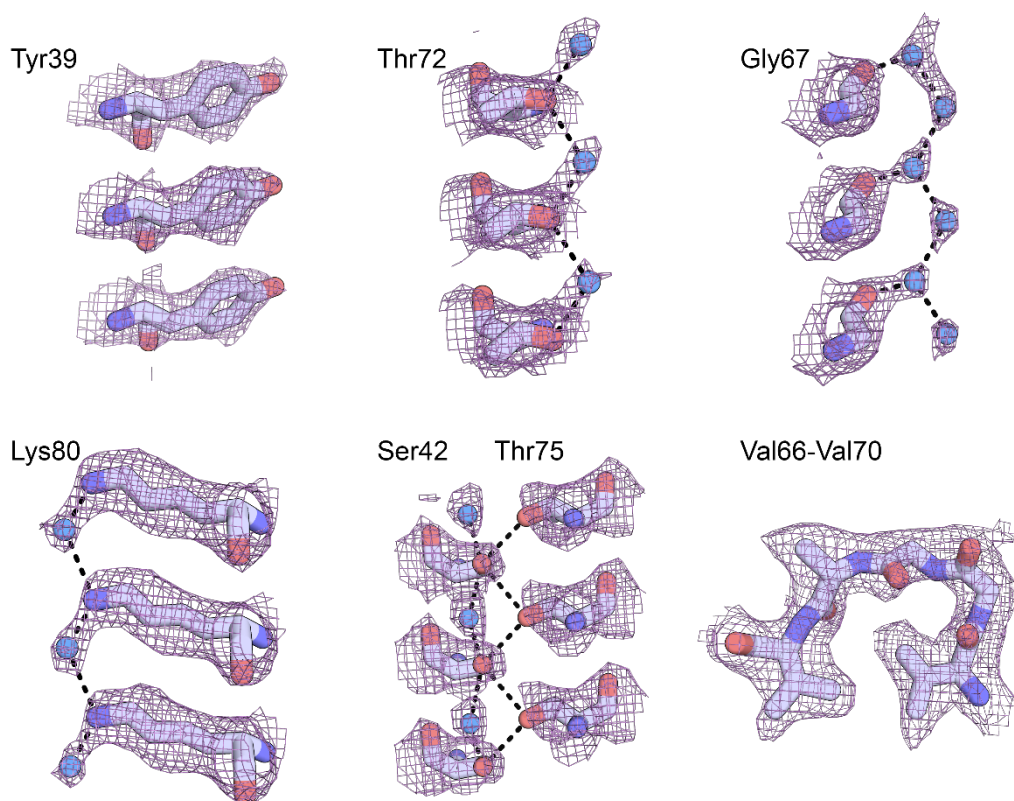


c



### 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  $\sim 800$  Å. 686 pixel box class averages and corresponding simulated diffraction patterns reveal a rise of  $\sim 4.8$  Å and a  $C_2$  rotational symmetry. c) Gold-standard half-map-half-map FSC and map-to-model FSC.

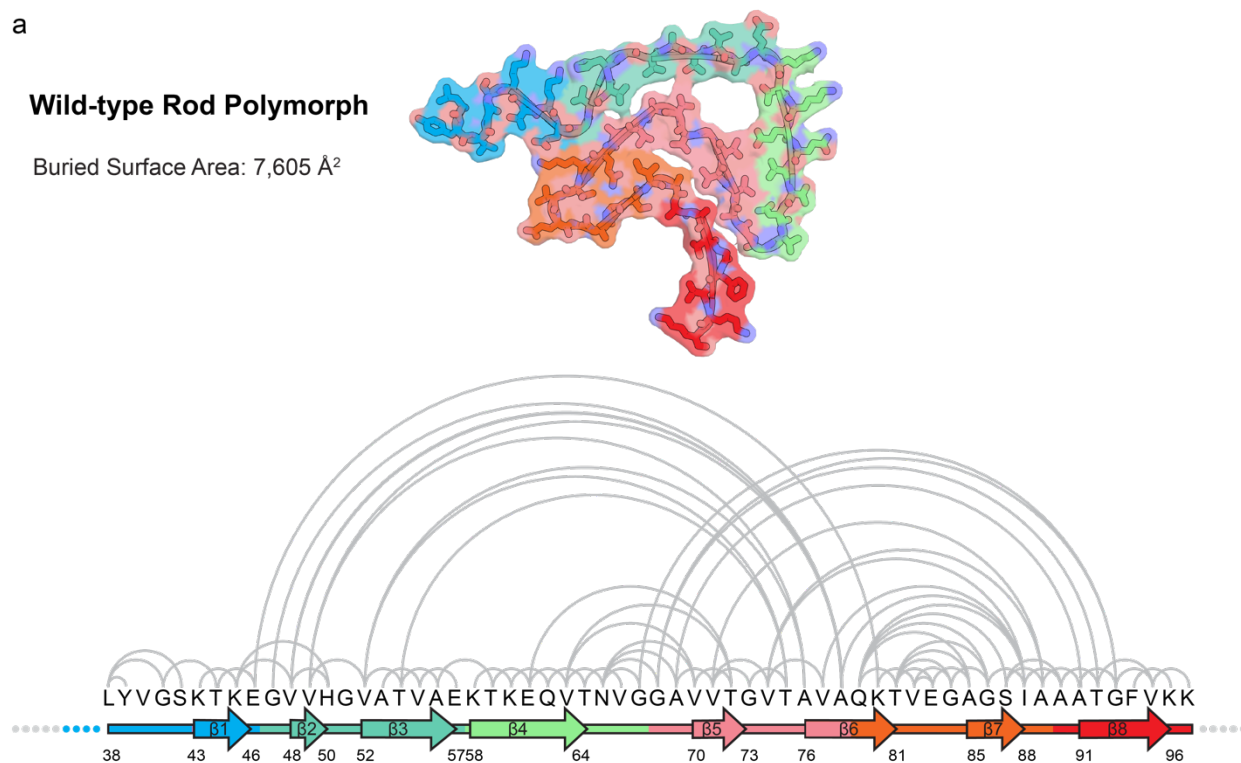


### Supplementary Figure 3 Atomic model and density map highlight resolution of reconstruction and hydrogen bonding networks.

a

### Wild-type Rod Polymorph

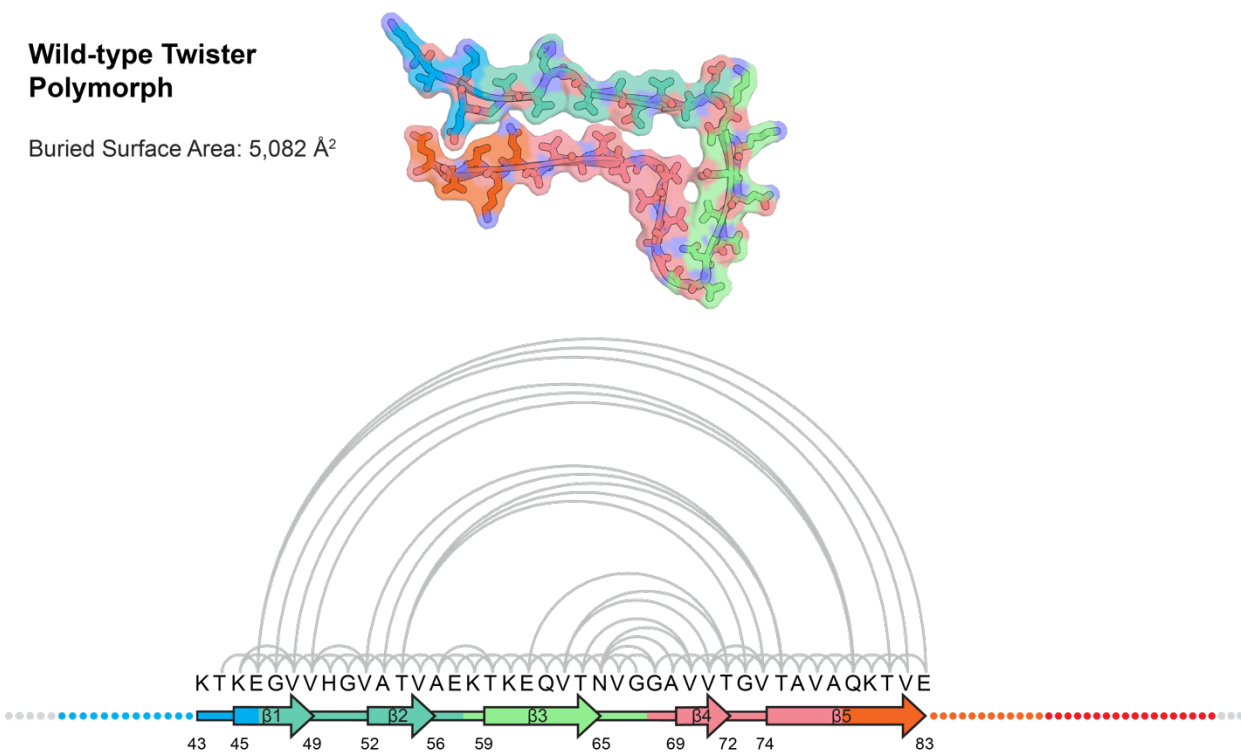
Buried Surface Area: 7,605 Å<sup>2</sup>



b

### Wild-type Twister Polymorph

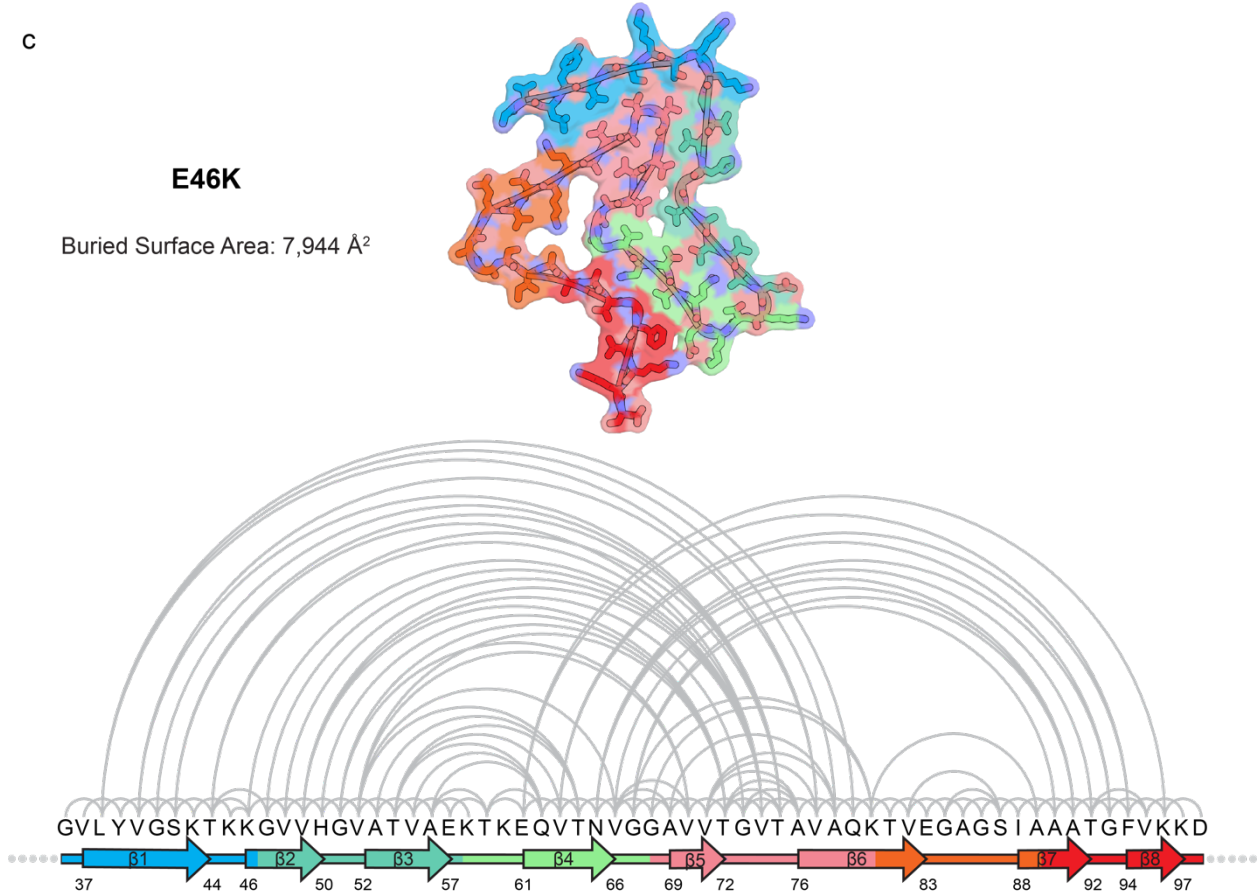
Buried Surface Area: 5,082 Å<sup>2</sup>



c

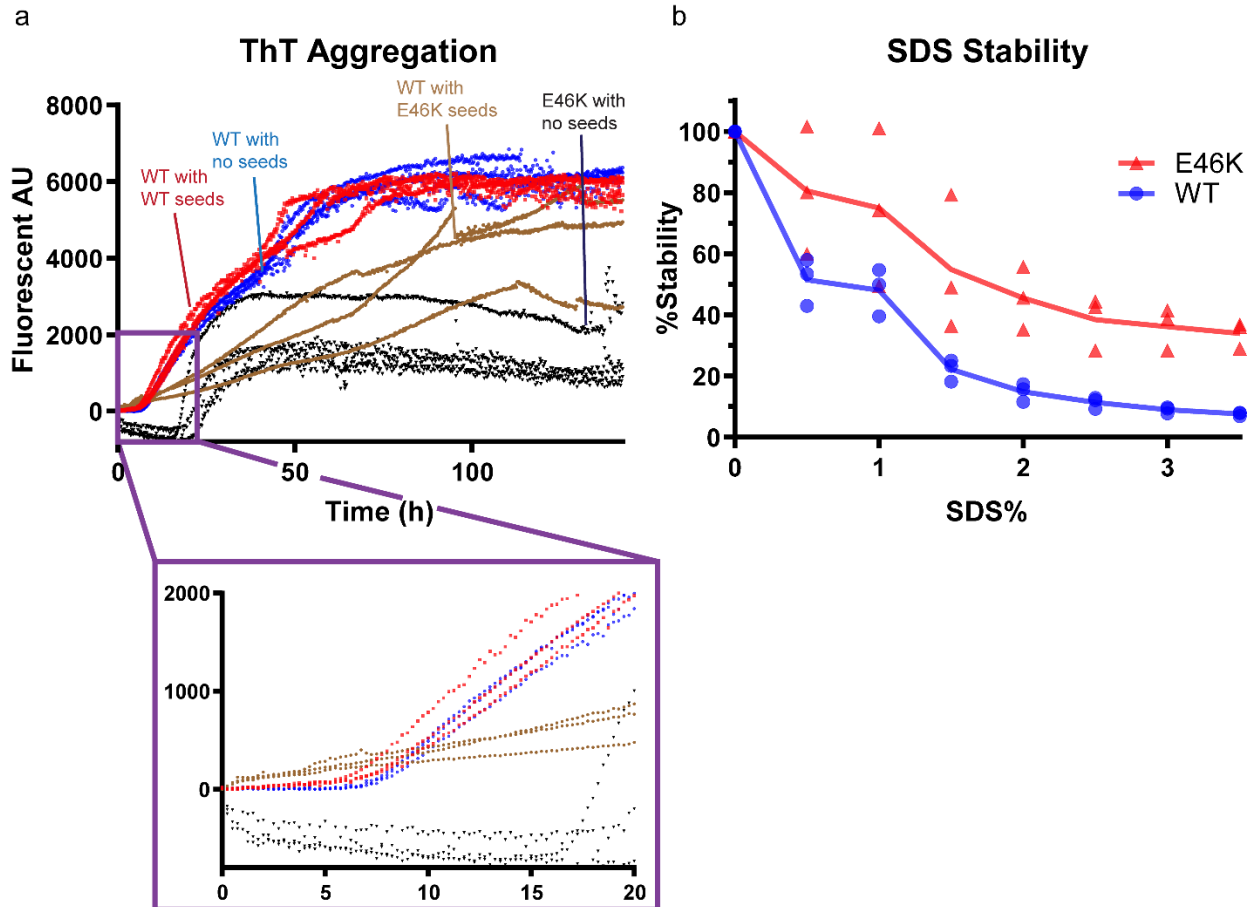
**E46K**

Buried Surface Area: 7,944 Å<sup>2</sup>



#### 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  $\alpha$ -syn by E46K fibrils and SDS fibril stability assay

a) Seeding of wild-type  $\alpha$ -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  $\alpha$ -syn have similar ThT binding ability shown by their similar ThT aggregation curves; on the other hand, wild-type  $\alpha$ -syn seeded by E46K fibrils has a different ThT fluorescence intensity, indicating a different underlying structure. Unseeded wild-type  $\alpha$ -syn has a shorter lag phase and higher max ThT signal than E46K  $\alpha$ -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  $\alpha$ -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.