On the pH-dependence of α-synuclein amyloid polymorphism and the role of secondary nucleation in seed-based amyloid propagation

Lukas Frey[1†], Dhiman Ghosh[1†], Bilal M. Qureshi[2†], David Rhyner[1], Ricardo Guerrero-Ferreira[3] Aditya Pokharna[1], Witek Kwiatkowski[1], Tetiana Serdiuk[4], Paola Picotti[4], Roland Riek[1,*], Jason Greenwald[1,*]

[1]Institute of Molecular Physical Science, Vladimir-Prelog-Weg 2, ETH Zürich, Zürich, Switzerland
[2]Scientific Center for Optical and Electron Microscopy, Otto-Stern-Weg 3, ETH Zürich, Zürich, Switzerland
[3]Robert P. Apkarian Integrated Electron Microscopy Core, Emory University, Atlanta, Georgia, USA
[4]Institute of Molecular Systems Biology, ETH Zürich, Zürich, Switzerland

[†]These authors contributed equally to this work

Keywords: amyloids, Parkinson's disease, α-synuclein, polymorphism

*Correspondence to: roland.riek@phys.chem.ethz.ch; jason.greenwald@phys.chem.ethz.ch

Abstract
The aggregation of the protein α-synuclein is closely associated with several neurodegenerative disorders and as such the structures of the amyloid fibril aggregates have high scientific and medical significance. However, there are dozens of unique atomic-resolution structures of these aggregates, and such a highly polymorphic nature of the α-synuclein fibrils hampers efforts in disease-relevant *in vitro* studies on α-synuclein amyloid aggregation. In order to better understand the factors that affect polymorph selection, we studied the structures of α-synuclein fibrils *in vitro* as a function of pH and buffer using cryo-EM helical reconstruction. We find that in the physiological range of pH 5.8-7.4 a pH-dependent selection between Types 1, 2 and 3 polymorphs occurs. Our results indicate that even in the presence of seeds, the polymorph selection during aggregation is highly dependent on the buffer conditions, attributed to the non-polymorph-specific nature of secondary nucleation. We also uncovered two new polymorphs that occur at pH 7.0 in phosphate-buffered saline. The first is a monofilament Type 1 fibril that highly resembles the structure of the juvenile-onset synucleinopathy polymorph found in patient-derived material. The second is a new Type 5 polymorph that resembles a polymorph that has been recently reported in a study that used diseased tissues to seed aggregation. Taken together, our results highlight the shallow amyloid energy hypersurface that can be altered by subtle changes in the environment, including the pH which is shown to play a major role in polymorph selection and in many cases appears to be the determining factor in seeded aggregation. The results also suggest the possibility of producing disease-relevant structure *in vitro*.

Introduction

Amyloid fibrils are a unique class of self-assembled filamentous protein/peptide aggregates that have been associated with dozens of human diseases including Alzheimer's, Parkinson's, type II diabetes, and prion diseases[1]. It has long been established that all amyloids contain a common structural scaffold known as the cross β-sheet motif, yet only more recently, with the advent of solid-state nuclear magnetic resonance spectroscopy (NMR)[2] and cryo-electron microscopy (cryo-EM) helical reconstruction[3], has it become clear that this shared repetitive motif can generate a wide range of structural folds[4,5]. Yet unlike globular proteins, whose amino acid sequence typically guides the protein folding process into a unique 3-dimensional structure, the amyloid structures derived from a single peptide sequence are often polymorphic: they can exist in various stable three-dimensional architectures. Another unique property of amyloids is their tendency to induce their own formation in a self-replicative manner. Amyloid formation is a nucleated growth process analogous to the growth of a 1D crystal[6] and can therefore be kinetically enhanced through templating on preformed fibrillar structures, termed seeds. Thus, in seeding, the amyloid fold itself acts as a structural template for nucleation and growth, whereby protein monomers are induced into the amyloid conformation and contribute to fibril elongation. Additionally, as revealed by kinetic experiments, there are secondary nucleation processes, which include *de novo* nuclei formation on the lateral surface of the amyloid fibrils[7]. Structural polymorphism and seeding-enhanced growth are two hallmarks of amyloid structural biology. However, despite their likely importance in the etiology of many diseases, the factors that control polymorphs and the molecular mechanisms of seeding are poorly understood.

Amyloid polymorphs whose differences lie in both their tertiary structure (the arrangement of the β-strands) and the quaternary structure (protofilament-protofilament assembly) have been found to display distinct biological activities[8–12]. In the case of the protein tau, cryo-EM has been used to correlate 14 different tauopathies with distinct combinations of 8 different amyloid polymorphs[13]. In the case of α-synuclein (α-Syn), the structures of many *in vitro* polymorphs have been determined[14–37] and yet, they are all distinct from the polymorphs that have been sarkosyl-extracted from the brains of patients of both multiple systems atrophy (MSA)[24] and Parkinson's disease (PD)[34]. Understanding how these disease-relevant polymorphs form and reproducing them in vitro will be critical in the study of these synucleinopathies. Some clues

are present in the sarkosyl-insoluble brain-derived structures: two MSA polymorphs contain an unidentified molecule at the interface of the protofilaments and in the PD polymorph an unidentified molecule appears to be stabilizing the structure of the monofilament fibril. These findings support the idea that the polymorph type is in some way associated with the disease, and that the environment plays a critical role in amyloid formation, essentially dictating which polymorphs are accessible.

Concerning the seeding of polymorphs, it was demonstrated that prion strains with their specific pathologies can be serially transmitted from animal to animal[38,39] and that strains can be amplified *in vitro* by protein misfolding cyclic amplification (PMCA)[40,41] or similar approaches[42]. However, *in vitro,* amyloid seeding is not always faithful in reproducing the seed polymorph. In the case of α-Syn, PMCA-based seeding has been proposed as a diagnostic tool that could also differentiate between MSA and PD[41,43]. However, attempts to amplify the MSA and the PD patient-derived polymorphs *in vitro* with recombinant α-Syn have so far resulted in structures distinct from those that were extracted with sarkosyl from patient brains[29,44,45]. These results could be due to the strong role of secondary nucleation in the *in vitro* seeding processes, which in the case of α-Syn have been shown to involve flexible interactions that might not be able to transmit polymorph-specific structural information[43]. However, it could also be that multiple polymorphs exist in diseased tissues and that the sarkosyl-based procedure itself selects different polymorphs compared to the PMCA-based procedures. Nonetheless, seeding with α-Syn disease-derived amyloids has, at least for cerebrospinal fluid (CSF)-derived seeds, some diagnostic utility as it been shown that under conditions of low α-Syn monomer, patient-derived CSF can seed several polymorphs while healthy CSF does not seed amyloid formation at all[37]. On the other hand, seeding experiments with *in vitro*-derived polymorphs have yielded mixed results in that some polymorphs are seemingly easier to reproduce via seeding. For example, the Type 4 polymorph from E46K[25] or G51D[26] mutants can be transmitted to wild-type α-Syn, and a polymorph that forms normally only at low salt can be produced at high salt and vice versa by using very high seed concentrations (10%) and by limiting primary and secondary nucleation processes[12,46].

Having observed the plethora of α-Syn fibril polymorphs published in just the last few years, we set out to investigate the relationship between α-Syn polymorphs and their aggregation conditions. The speed and abundance of polymorphs being discovered have led to inconsistent naming conventions, and in this work, we expand upon previous nomenclature[16,22] which is illustrated in Figure 1, whereby polymorphs with distinct folds are assigned different Arabic numerals and sub-polymorphs formed through different proto-filament interfaces given different letters (with the exception of polymorph 1B whose name is kept for consistency with previous publications despite itself being a distinct fold, see Figure 2). We find that the pH of the aggregation condition is a major but not the only factor in determining which polymorph is formed *in vitro*. While attempting to reproduce the published conditions for the Type 2 polymorph, we found a new polymorph (here termed Type 5) and a monofilament Type 1 polymorph, providing further evidence that the energetic landscape of α-Syn aggregation is rough with many interconnected local minima. In addition, our findings with cross-pH seeding experiments support the notion that seeding with α-Syn is non-polymorph-specific under conditions of secondary nucleation-dominated aggregation[46], consistent with difficulties in recreating the *in-vivo* observed fibrils through seeding[29].

Results and Discussion

The range of accessible polymorphs is controlled by pH

There are more than 60 *in vitro* amyloid structures of α-Syn deposited in the PDB, representing over a dozen different structural polymorphs. The authors of these structures often try to frame newly observed polymorph(s) in the context of the diseases associated with α-Syn aggregation. However, given so many published structures, an examination of the conditions under which the polymorphs were formed also offers hints into the environmental influences

of polymorph selection. In particular, we noted that Type 1 structures were observed almost exclusively above pH 7.0, while Type 2 and 3 structures were observed below pH 7.4 and down to pH 6.5. To systematically analyze the effect of buffer and pH we studied the structures of α-Syn amyloids of the full-length, N-terminally acetylated protein formed under a variety of conditions. In a first series, we produced α-Syn amyloids in PBS (10 mM PO_4, 137 mM NaCl, 2.7 mM KCl) at the physiological pH values of 7.4, 7.0, 6.5, and 5.8.

At pH 5.8 we observed exclusively the Type 3 structures 3B and 3C (Figures 3 and 4). Considering the similarities between Type 2 and 3 structures (see below), it is noteworthy that an A-interface polymorph was not observed for Type 3. In its place, we find the 3C polymorph for which the interface contacts are shifted by one residue with respect to 2A and are now defined by a salt bridge between E46 and K58 instead of K45 and E57 (Figure 5B). In the previously reported Type 2A and 2B structures solved from a single sample aggregated in PBS at pH 7.4, the major polymorph (83%) was 2A. In our first sample prepared at pH 5.8, we observed 3B as a major polymorph (> 90%) with 3C as a minor polymorph. We hypothesized that the absence of a "3A" polymorph and the presence of the 3B as the major polymorph could be related to the ionic strength of the buffer used in aggregation. Therefore, we also pursued cryo-EM structures under similar conditions but with an additional 50 mM or 100 mM NaCl. While we did see a difference in the relative abundance of the 3B and 3C polymorphs in these samples (Table 1), "3A" was never detected nor was there an obvious correlation between the NaCl concentration and the 3B:3C polymorph population. In fact, in a second sample prepared under the same original PBS conditions which had yielded an approximately 93%:7% 3B:3C ratio, we encountered yet another ratio of 54%:46%, leading us to the conclusion that stochastic or not well-controlled conditions affect the selection between the 'B' and 'C' interface variants of Type 3 fibrils.

Previous reports of non-acetylated α-Syn fibrils seeded from MSA patient material and grown at pH 6.5[29] contained a mixture of 2A, 2B, 3B, and 3C polymorphs as well as hybrid 2/3B fibrils. It is therefore noteworthy that in the five independent amyloid samples that we prepared at pH 5.8, none of them contained a detectable level of Type 2 fibrils. However, in the sample prepared at pH 6.5, we did indeed see a mixture of Type 2 and 3 (2A:2B:3B:3C at 25%:18%:31%:26%) suggesting that Type 2 is not accessible below pH 6 (Figure 3). The structural basis for the Type 3 preference at lower pH may lie in the protonation state of glutamic acid residues in the vicinity of the C-terminal β-strand (i.e. E83) or of H50, two regions where polymorphs 2 and 3 are most distinct from each other (discussed below). The fact that we observe the same set of polymorphs in the absence of MSA seed material suggests that not only was the seeding described above non-polymorph specific but also that pH was the main polymorph determinant. Seeding experiments have also been performed with cerebrospinal fluid (CSF) of Parkinson's disease patients at pH 6.5, this time with N-terminally acetylated α-Syn, yielding once again the 3B and 3C polymorphs[37].

In both the MSA and Parkinson's-CSF seeding experiments, the authors also reported on some samples that produced Type 1-like polymorphs which, unlike most *in vitro* polymorphs, existed primarily as single monofilament fibrils. These are noteworthy because the large majority of *in vitro* and *in vivo* α-Syn polymorphs are composed of more than one protofibril, most often with a C2 or pseudo 2_1 helical symmetry. Therefore, we were surprised to find that in several pH 7.0 aggregation samples (Table 1, Figure 4), we also observed a Type 1 monofilament fibril polymorph (termed 1_M), similar to those from the MSA and Parkinsons-CSF seeding experiments and very similar to the Juvenile-onset synucleinopathy (JOS) polymorph[29,44,45] (Figure 6, discussed below). This again indicates that seeding experiments may be influenced by environment-dictated polymorph formation. Also puzzling to us is that we could not produce Type 2 polymorphs at pH 7.0 or 7.4 despite previous reports of 2A and 2B being formed under these conditions[22,30]. In yet another of the pH 7.0 aggregation samples we found a completely new polymorph which we have termed Type 5 (Figure 7). Apart from its novel fold, it is unusual in the number of residues that are visible/ordered in the structure, starting from the acetylated N-terminal methionine (buried at the dimer interface) up until

residue 96. Interestingly in a recent BioRxiv manuscript[45] (for which the coordinates are currently on hold), a similar polymorph is found upon seeding with MSA and PD material, again suggesting that the environment could play a role in the selection of polymorphs.[14]

In the original reports of Type 2 structures (2A and 2B), fibrils were formed by non-acetylated α-Syn at pH 7.5 in Tris buffer as well by three other constructs at pH 7.0-7.3 in DPBS: (i) the E46K mutant, (ii) phosphorylation at S129 and (iii) an N-terminally acetylated α-Syn. Therefore, we were somewhat surprised that after measuring cryo-EM data for six independent fibrilization samples at pH 7.0 and 7.4 in PBS or Tris (Table 1), we could not detect any Type 2 polymorphs. Instead, we found that at pH 7.0 in PBS, the samples exhibited a multitude of polymorphs including one having mostly flat (not twisted) fibrils of variable width and the other two each yielding a new polymorph (5A and 1_M). However, in Tris buffer at pH 7.0 we did produce polymorph 1A as the major species. Considering the strong temperature dependence of the Tris pKa, it is worth noting that our Tris buffer was prepared for pH 7.0 at the fibrillization temperature of 37 °C. Returning to PBS but at pH 7.4, polymorph 1A was the dominant species in both samples that we prepared, even though one was seeded by 5% Type 3 fibrils that had grown at pH 5.8 (the latter discussed in more detail below). While in our hands 1A is the preferred polymorph above pH 7.0, regardless of the buffer used, we also observed a higher degree of sample variability in the samples grown at pH 7.0 and 7.4 (Table 1, Figure 4). In order to further explore the variability/reproducibility of the pH 7.0 polymorphs we measured cryo-EM data for 6 more independent fibril samples for a total of 10 samples, all aggregated independently at pH 7.0 and originating from 4 different batches of purified α-Syn (datasets 7-16, Table 1). In three of these pH 7.0 samples we did finally reproduce the type 2A/B polymorphs (datasets 11-13), two of which had also a significant amount of type 1_M fibrils. In yet two others datasets (14-15) we observed only type 1A. Several samples contained predominantly non-twisted fibrils and could therefore not be used for helical reconstruction. In most samples, we also observed fibrils that were thinner or wider than the known polymorphs, but these were not pursued due to a poor homogeneity between fibrils, lack of twist, or an insufficient number of specimens and so were excluded either during manual picking or later in 2D and 3D classification.

On the structural heterogeneity of Type 1 polymorphs

27 structures of Type 1 polymorphs have been reported, mostly in the pH range of 7.0-7.4 in distinct buffers including Tris or phosphate and comprised of full-length α-Syn, N-acetylated α-Syn, shorter constructs (i.e. 1-103, 1-122), familial mutants (i.e. A53T, A53E, H50Q, and a seven residue insertion associated with JOS), patient-derived material (i.e. MSA and JOS), and fibrils that were seeded by MSA, PD, and dementia with Lewy bodies material (including brain or cerebrospinal fluid). The Type 1 polymorphs display a large structural heterogeneity both at the tertiary structure level (Figure 8) as well as the quaternary level, having distinct protofilament interfaces or even existing as a monofilament fibril[14–16,19,20,24,29,31–33,35–37]. This can be rationalized in part, for example with the familial mutants A53T, A53E, and H50Q which all have been found in polymorph 1C with its smaller protofilament interface (residues 59-60) compared to 1A (residues 50-57). The mutants lie at the 1A interface, destabilizing it and favoring the smaller interface of 1C or even the monofilament 1_M (illustrated in Figure 2). In the context of how buffer conditions influence the polymorph structure, it is interesting to mention that in the presence of phosphate, the K58 side chain is often facing out towards the 1A protofilament interface, thereby forming a triad with K43 and K45 of the other protofilament to surround some extra density that is interpreted to be phosphate (Figure 8). This constellation of lysine residues is found in about 2/3 of the dozen structures of amyloids grown in the presence of phosphate. In contrast, in phosphate-free conditions (such as Tris, PIPES, or unbuffered) K58 is facing into the core in about 2/3 of a dozen structures, forming a salt bridge/hydrogen bond with E61. Since in the presence and absence of phosphate both K58 orientations have been documented, other buffer details such as pH may also be relevant for this and other variations. Furthermore, it points to a rather shallow amyloid energy hypersurface that can be altered by subtle changes in the environment.

In the context of polymorph 1 and the interfacial lysine triad discussed above, it is worth mentioning the structure of the disease-relevant JOS brain-derived polymorph. In this polymorph, the 1A protofilament interface is disrupted by an additional β-strand comprising the N-terminal residues E12-E25, yielding a monofilament structure that we term Type 1_M as well as a double protofilament structure with a unique protofilament interface. Consequently, the lysine triad of 1A is lost (with K58 still facing outward), but now K43 and K45 together with K21 and K23 of the N-terminal additional β-strand embrace a small molecule that is likely negatively charged, yielding overall a similar local structural arrangement around K43 and K45 (and H50). An attempt to produce the JOS polymorphs *in vitro* using a mixture of wild-type α-Syn and variant containing the JOS-associated seven-residue insertion MAAAEKT after residue 22 yielded instead polymorph 1A structures with the typical protofilament interface[36]. However, in our structure-buffer relationship screening, a JOS-like 1_M polymorph containing the additional N-terminal strand of the JOS polymorph was obtained with wild-type N-terminally acetylated α-Syn. There is only a small difference in the conditions, pH 7.4 versus pH 7.0 in PBS, that lead to the 1A versus the JOS-like 1_M polymorphs, again illustrating the shallow energy surface between the polymorphs and the critical role of the buffer conditions in polymorph selection. The *in vitro* prepared JOS-like 1_M polymorph superimposes with the Juvenile-onset synucleinopathy purified material almost entirely except for the aforementioned K58 side chain facing inward in the former and outward in the latter, and the unassigned peptide density in the latter (Figure 6B). The cryo-EM map for the JOS-like 1_M fibrils also has an additional sausage-like density, likely for a beta-strand, but it was not used for model building because it lacked enough features (Figure 6A).

On the structural heterogeneity of Type 2 and 3 structures

The structures of Type 2 and 3 polymorphs share some similarities, however, there are significant differences that justify their assignment to unique polymorph classes (Figure 5AB). First, there is a displacement of the structural fragment comprising residues V52-T72 which appears to move on the G51 and G73 hinges. The turn with residues G73-A78 is again similar between the two structures but beyond this, there is an “inside-out” rearrangement of the C-terminal β-strand (A85-G86-S87-I88-A89-A90-A91-T92) such that only in the Type 2 polymorph can the N-terminal fragment (V15-K21) form an additional β-strand packed against the β-strand A85-A91[22]. However, also within Type 2 and 3 polymorphs exists a large structural heterogeneity: some polymorph 3 structures have an “inside-out” C-terminal β-strand but still without the N-terminal fragment present. In terms of the interface variants, there is a difference between our Type 3B structure and that of from the MSA-seeded Type 3B[29]. Our data shows the characteristics of a pseudo-2_1 helical symmetry (ca. 179.5º twist and 2.35 Å rise) when refined under C1 symmetry with a ca. -1º twist and 4.7 Å rise (Figure 9) while the MSA-seeded Type 3B structure (PDB entry 7NCG) has a C2 symmetry (-0.95º twist 4.75 Å rise).

The interfaces present in the 2A and 3C polymorphs also appear to be similar at first glance, consisting of a pair of symmetric salt bridges arranged to create a large solvent-filled cavity at the interface. However, a close inspection reveals that they are different, comprising neighboring sets of E/K residues: K45/E57 in 2A and E46/K58 in 3C. This is remarkable because there is no obvious reason that the “C” interface is preferred over the “A” interface. An example of a 3A polymorph has been reported but it is in the context of the E46K mutant[23] and formed in an unbuffered solution. The familial Parkinson’s disease mutant α-Syn (E46K)[47] has been extensively studied and the structures of fibrils of this mutant have been reported at pH 7.5 in Tris[18], at pH 7.4 in PBS[22] and as mentioned above, unbuffered in 15 mM tetrabutylphosphonium bromide[23]. The Tris condition produced a Type 4 polymorph, the PBS a Type 2A, and the unbuffered condition a Type 3A. Interestingly, since the 2A and 3A polymorphs share the same dimer interface with symmetric K45/E57 salt bridges it is clear that the “A” interface is not mutually exclusive with the Type 3 polymorph. However, it is still not clear as to why Type 3B and 3C polymorphs, both with E46 in their salt-bridged dimer interfaces, are preferred for the wild-type protein but we suspected that it was only through

the elimination of E46 that the E46K mutant could access the 3A polymorph. To test this further, we produced E46K fibrils at pH 5.8, the condition where until now only Type 3 has been observed and uncovered yet another Type 3 polymorph with a new interface termed 3D formed by an E57/K58 salt bridge (Figure 5B). This was surprising considering the previously mentioned Type 3A structure of this mutant and suggests that the Type 3 fibrils formed at pH 5.8 are somehow distinct from those formed, until now only by mutants, above pH 7. While researching deposited α-Syn structures we found another Type 3D polymorph (PDB entry 7WO0, not yet published) that is formed by the A53T mutant in the presence of 20 mM $CaCl_2$ at pH 7.6. Finally, the coexistence of polymorphs 2A/2B and 3B/3C at pH 6.5 indicates either that they all have a similar thermodynamic stability or that some of them are kinetically trapped, the latter possibility having been observed in one case of two distinct α-Syn polymorphs grown in a single condition (10 mM Tris, 0.02% NaN_3, pH 7.0 at 37 °C)[48].

The Type 5 structure

Apart from its novel fold, polymorph 5A is unusual in the number of residues that are visible/ordered in the structure, starting from the acetylated N-terminal methionine (buried at the dimer interface) up until residue E96. As is often the case, there are shared structural motifs between Type 5 and other α-Syn polymorphs. In particular, residues 47-72, comprising a two-beta strand motif with a turn in between, overlay very well with the polymorph formed by a Y39 phosphorylated α-Syn[17] and to a lesser extent in Types 2, 3, and 4 as well. However, the first 35 residues are distinct in that they are usually disordered (Types 1 and 3), or only a segment thereof binds to the core in Type 2 or polymorph 1_M. Within this segment, there are seven lysine and five glutamic/aspartic acid residues and five of the lysine residues are buried inside the structure and form salt bridges with buried glutamic/aspartic acid residues. This yields two large cavities separated only by the very polar segment G31-K32-T33-K34-E35-G36, and it is likely that these two cavities are filled with water. In comparison to the other known polymorph structures the size of the cavities and the presence of five internal salt bridges are unique. Only polymorph 1 shows a small significant cavity which as discussed above usually appears in the presence of phosphate buffer. However, even the more compact Type 3 structures which do not have any obvious cavities contain networked channels of waters that are shared between the wild-type Type 3B and E46K Type 3D structures (Figure 5C,D). The occurrence of the Type 5A polymorph thus reflects again on the promiscuity of α-Syn polymorph formation, the influence of slight changes in buffer conditions, and more specifically the potential role of lysine residues in polymorph selection. It is also worth mentioning that in the 10 independent aggregations that we analyzed at pH 7.0, the Type 5 structure was found in only one sample, and then, in only 10-20% of the observed fibrils in the sample. The remainder of the fibrils in the sample were non-twisted and therefore could not be analyzed by 3D reconstruction.

The buffer environment can dictate polymorph during seeded nucleation

While for many amyloids, seeded growth results in replication of the structural polymorph of the seeds, in at least one instance, the seeding with MSA brain-derived fibrils failed to reproduce the seed polymorphs[29]. One explanation could be the absence of the as yet unknown cofactor that appears to be bound in MSA polymorphs. However, as we have observed, the buffer environment/condition may play such an important role in polymorph selection that we wondered whether, for pure *in vitro* seeding experiments with *in vitro*-derived seeds, it would be the environment or the seed which dictates the resulting polymorph. To address this, we used fibrils of polymorph 3B/3C produced at pH 5.8 as seeds (5%) in seeding experiments at pH 7.4 and as a control at pH 5.8. The inverse experiment was also carried out with fibrils of the Type 1 polymorph produced at pH 7.4 used as seeds (5%) in seeding experiments at pH 5.8. In both cases, control seeding experiments were performed at the pH of the original seed fibrils. For these analyses, we monitored the effect of seeding on the kinetics of fibril formation using thioflavin T (ThT) fluorescence[49,50] and probed the structures of the resulting seeded fibrils by cryo-EM.

In terms of kinetics, the presence of seed fibrils consistently accelerated the aggregation of α-Syn (Figure 10A). However, in the Cryo-EM analyses of the fibrils produced at pH 5.8 with pH 7.4 seeds, only the polymorphs 3B and 3C are detected (Figure 10B) and in the inverse case, with fibrils produced at pH 7.4 with pH 5.8 seeds only polymorph 1A is detected (Figure 10B). We were unable to detect any traces of the seed polymorphs in either of the cross-seeded aggregations despite the fact that 5% seeds were used. Limited proteolysis-coupled mass spectrometry (LiP-MS)[51,52], a quantitative technique that enables the detection of nearly single residue resolution structural differences/alterations in proteins in bulk solution, confirmed these findings (Figure 10C). In our LiP-MS analysis, we digested the fibril samples with the sequence-unspecific protease proteinase K for a limited amount of time to generate structure-specific proteolytic fingerprints and then identified and differentially quantified peptides between the different fibril states by MS analysis. To obtain single amino acid resolution in our comparison of the structural features of different types of fibrils, we first assigned a score accounting for the extent of change between samples and a measure of statistical significance to each detected peptide and then aligned the peptides with overlapping regions to calculate the mean score value for each amino acid (see methods for details). Thus, larger differences between two structures yield higher per-residue scores. LiP-MS analysis revealed that fibrils generated at pH 7.4 (F3) and pH 5.8 (F4), in the presence of unfragmented fibrillar seeds generated at pH 5.8 and pH 7.4 respectively, are on average more similar in their structure to the fibrils generated without seeds at pH 7.4 (F1) and pH 5.8 (F2), respectively (Figure 10C). These data suggest that fibrillar conformations resulting from seeding monomeric alpha-synuclein with unfragmented fibrillar seeds are driven by pH and not by the structure of the seed.

Amyloid aggregation can be accelerated by two seed-based processes: (i) fibril elongation and (ii) secondary nucleation. In fibril elongation, the incoming protein adopts the conformation of the seed by attaching itself to the end of the fibril. In this process of growth acceleration, which is proportional to the number of fibril ends, the polymorph structure is expected to be preserved[53,54]. In the secondary nucleation process, *de novo* aggregation occurs on the sides of the fibril and therefore the kinetics of secondary nucleation is proportional to the mass of the fibril[55]. In the case of α-Syn, the secondary nucleation mechanism has been proposed to be based on the interaction of the positively charged N-terminal region of monomeric α-Syn and the disordered, negatively charged C-terminal region of the α-Syn amyloid fibrils[56]. The interaction is transient but causes a further unfolding of the monomeric α-Syn while increasing its local concentration and possibly aligning it in a way that promotes aggregation. Since all known amyloid polymorphs of α-Syn are disordered in their C-terminal regions, we propose that secondary nucleation in α-Syn aggregation is generally polymorph unspecific. This conclusion is consistent with those of Peduzzo et al.[46] who showed that at lower seed concentrations under which secondary nucleation dominates over fibril elongation, the environment dictates the resulting polymorph. It is worth noting that unfragmented fibrils will likely favor secondary nucleation, while with shorter fibrils and more available ends per unit mass, elongation could dominate.

In addition to the α-Syn seeding experiments discussed until now, Strohäcker et al.[44] used NMR to analyze the results of PMCA using both PD and MSA seed material in PBS buffer at pH 7.4 in the presence of 1% triton X-100. Both samples yielded indistinguishable polymorphs. Hence, it appears that their protocol amplified the seeds via secondary nucleation for which a single polymorph can be expected and whose identity is dependent on the environment; here that includes 1% triton X-100 and 5% w/v brain homogenate. In another PMCA study that is published currently only on the BioRxiv server, Burger et al.[45], use both MSA and PD patient-derived material to produce a new polymorph which they name Type 6. A similar Type 6 fold was obtained irrespective of whether the starting material was from an MSA or PD patient brain, however protofilaments were arranged in a disease-specific manner. Interestingly, the Type 6 fold resembles the Type 5 that we found at pH 7.0, deviating however in its three different protofilament interfaces yielding polymorphs 6A, 6B, and 6C as well as extra density that the authors have assigned to the C-terminal residues. Without the deposited coordinates

for the Type 6 structures it is difficult to ascertain how similar these polymorphs are to Type 5, however, their obvious similarity suggests that the environment of the PMCA seeding could have played a role in determining the amplified polymorph. Here the environment includes the brain homogenate at 2% w/v and the PMCA buffer: 150 mM KCl, 50 mM Tris, pH 7.5[45] at 37 °C, which considering the temperature dependence of the Tris pK_a could actually be close to pH 7 and thus similar to the buffer conditions that produced Type 5. However, it must also be stated that in addition to the cryo-EM structures[45], previous cell biological and biochemical support for polymorph-specific seeding of the MSA and PD patient-derived material under the same buffer conditions has been reported[11].

The general impact of the environment for seeding assays was recently investigated in a RT-QuIC assay for which 168 different conditions were tested to find those optimal for α-Syn aggregate detection[57]. The results indicate that seeding experiments with α-Syn can be very sensitive to the environment and that it is a multifactorial problem where there is not a single optimal pH or optimal salt condition. Another recent report compared the biological activity of (non-seeded) *in vitro* fibrils versus those amplified from Lewy body diseased brains upon injection into mice brains, with the finding that only the latter faithfully models key aspects of Lewy body disease[8]. The fact that a pH of 7.4 was used during the seeding process further suggests that there are other environmental factors at play in polymorphs selection. Considering the importance of reproducing the disease-relevant polymorphs and the often irreproducibility of amplification assays it is apparent that approaches need to be found that can establish robust methods for α-Syn polymorph-specific amplification. In addition to finding appropriate buffer conditions, we propose that any approach that reduces or eliminates secondary nucleation is worth pursuing. This could be potentially achieved by using shorter, fragmented fibrils as seeds thereby increasing the fibril end-to-side ratio[11] or alkaline fibrillation conditions and higher salt concentration[56] where secondary nucleation appears to be less prominent[58]. The importance of salt for seeded polymorph replication has also been reported for the polymorph-specific seed replication between the non-ThT-sensitive ribbon polymorph and a fibril polymorph that succeeds in 150 mM KCl (with 50 mM Tris, pH 7.5, 37 °C) but not in the absence of KCl[12]. Another approach would be to use an N-terminal-truncated version of α-Syn or to proteolytically remove the disordered C-terminus from the seeds. While it has been shown that with a larger proportion of seeds, it is possible to drive the growth by elongation instead of secondary nucleation[12,46], this may be difficult to achieve by starting from *in vivo* material due to limited sample amounts. Finally, the use of molecules that inhibit secondary nucleation by binding to the C-terminus of α-Syn such as the protein Brichos could be explored.

Conclusion

The presented study, which focuses on the pH-dependence of fibril polymorph selection in both *de novo* as well as seeded fibril growth, shows the complexity of the fibril formation process, yielding many distinct polymorphs whose diversity is attributed to the absence of an evolutionary selection towards aggregation. On the other hand, it also highlights a rationale for the influence of the environment on this chameleon-like property of alpha-synuclein. With respect to disease, the *in vitro* generation of a JOS-like polymorph could be obtained, and the problem of polymorph-unspecific seeding could in many experimental settings be attributed to secondary nucleation, opening an avenue to resolve this open problem in cell and animal disease studies and diagnosis.

Methods

Recombinant protein expression and purification of α-Syn

Recombinant N-terminally acetylated, human WT α-Syn was produced by co-expressing the pRK172 plasmid with the yeast N-acetyltransferase complex B (NatB)[59] in *E. coli* BL21(DE3*). α-Syn was expressed and purified according to the previously published protocol[60] with slight modification. In brief, a single colony from the overnight transformed plate was picked up. It was grown at 37 °C in 100 ml LB media containing 100 µg/ml Ampicillin and

50 µg/ml chloramphenicol. Overnight primary growth culture was added to freshly prepared 1 lit LB media at a final concentration of 2.5% (v/v) containing 100 µg/ml Ampicillin and 50 µg/ml chloramphenicol. It was further grown till O.D reached 0.8. Protein overexpression was initiated by adding IPTG at a final concentration of 1 mM. Thereafter, cells were grown at 37 °C for more than 4 hours and harvested. α-Syn was obtained from a periplasmic extract using a non-denaturing protocol. Finally, pure α-Syn was obtained by ion exchange chromatography followed by purification on a hydrophobic HiPrep phenyl FF 16/10 column (Cytiva). The pure protein was dialyzed extensively against water, lyophilized, and stored at -20 °C until further use.

Amyloid aggregation

Lyophilized α-Syn was dissolved in PBS buffer prepared from commercially available tablets (Sigma-Aldrich). The pH of the dissolved α-Syn was adjusted to pH 7.4 and then separated from aggregates on a Superdex 75 10/60 column (Cytiva) pre-equilibrated with the same buffer and pH. Monomeric α-Syn obtained from SEC was free from any preformed aggregates and its concentration was adjusted to 300 µM. 500 µl of the protein solution was incubated in a 1.5 ml Eppendorf Protein LoBind tube at gentle agitation via tumbling (~50 rpm) at 37 °C for 5 days. The fibrils formed in different pH/buffer conditions were prepared by buffer exchanging the monomeric α-Syn (the elution from the Superdex 75 column) over a PD-10 desalting column (Cytiva) pre-equilibrated in the appropriate buffer and/or pH before incubation at 37 °C with agitation. During every fibrillization procedure, protein concentration and incubated volume were kept constant at 300 µM and 500 µl, respectively, to ensure that different starting conditions do not affect the aggregation process. Similarly, the seeded and cross-seeded fibrils were produced by incubating 300 µM, 500 µl monomeric α-Syn in the presence of 5% (v/v) preformed fibrils (from a previous aggregation of 300 µM α-Syn).

α-Syn aggregation kinetics monitored by ThT fluorescence

α-Syn aggregation kinetics in the presence and absence of preformed seeds was monitored by ThT fluorescence. Monomeric α-Syn prepared and incubated as above and at regular time intervals, 5 µl of the sample was taken and diluted to 200 µl by in the same aggregation buffer. ThT was added at a final concentration of 10 µM and its fluorescence was measured by excitation at 450 nm, and emission at 480 nm with a 5 nm bandpass for both excitation and emission on a Fluoromax-4 (Jobin Yvon) spectrofluorometer. ThT fluorescence intensity at 480 nm was plotted against different time points and three independent measurements were performed for each sample. Here, only 1% of preformed aggregates were used as seeds for ThT measurement in contrast to 5% used for producing fibrils for structural studies.

Electron microscopy grid preparation and data collection

Cu R2/2 300 mesh or Au R2/2 300 mesh grids (Quantifoils) were glow discharged at 25 mA for 30 sec and 120 sec, respectively. Freshly glow-discharged grids were used in a Vitrobot Mark IV (Thermo Fisher Scientific) with its chamber set at 100% humidity and at a temperature of 22 °C. Fibrils (4 uL aliquots) were applied to the grid and blotted for 8 to 15 seconds after a 30-60 second wait time, and subsequently plunge-frozen into a liquid Ethane/Propane mix. The grids were clipped and immediately used or stored in liquid nitrogen dewar.

Data acquisition was performed on a Titan Krios (Thermo Fisher Scientific) operating at 300 kV equipped with a Gatan Imaging Filter (GIF) with a 20 eV energy slit using Gatan's K3 direct electron detector in counting mode. Movies were collected using EPU software (Thermo Fisher Scientific) at a magnification of 130kx (see Table 1 for physical pixel size) and a dose rate of approximately 8 e/pix/sec and total dose ca. 62-74 e/Å^2 (see Table 2 for details).

Image processing

Image processing and helical reconstruction were carried out with RELION 4.0, following the procedure for amyloid structures described by Scheres[61]. Movies were gain-corrected and

RELION's own motion correction was used to correct for drift and dose-weighting and CTF estimation was done using Ctffind4.1[62]. Individual filaments were either manually selected or selected with the Topaz implementation in RELION. The segments were extracted in a 333 Å box (except for dataset 13, Table 1 for which a 260 Å box was used) with an inter-box distance of 33 Å and initially binned 4x to a pixel size of 2.6 Å for 2D classification. In the case of mixed polymorphs, larger box sizes were also used to help visually separate the classes and extract the crossover distances. For mixed polymorph datasets, the segments belonging to different classes were separated either by inspection of the 2D classes or by the new filament subset selection tool in RELION (version 5). For the new polymorphs (1_M and 5) we then used *relion_helix_inimodel2d* to generate an initial model that could be used for the subsequent 3D refinement steps. In some datasets, the polymorphs were separated purely by 3D classification and for others purely by filament subset selection (indicated in Table 1). As the refinement progressed, the refined particles were re-extracted at either 1.3 Å/pixel (2x binning) or 0.65 Å/pixel (no binning) depending on the current resolution. For the deposited datasets, Bayesian polishing[63] and CTF refinement[64] were performed to improve the resolution of the reconstructions. The deposited reconstructions were sharpened using standard postprocessing procedures in RELION. For the data for which maps have been deposited in the EMDB (Table 2), representative 2D classes and the postprocess-generated FSC curves are shown in Figures 11-17.

Model building and refinement

Models were built into the RELION post-processed maps with ModelAngelo[65] the output of which was manually adjusted in COOT[66] and further adjusted by real-space refinement as a 9-layer fibril in ISOLDE[67] with symmetry and secondary structure restraints. The outer four layers, which often diverge slightly in structure due to their placement at the edges of the model were removed and the central 5 layers further refined in PHENIX[68] to obtain reasonable b-factors before deposition in the PDB. Figures were prepared with CCP4MG[69] and UCSF Chimera[70]. For the data for which coordinates have been deposited in the PDB (Table 2), model-map FSC curves generated with the validation tools in PHENIX[71] are shown in Figures 11-15. For the model-map analyses, the maps were first masked to contain just the 5-layer PDB file.

Limited proteolysis mass spectrometry (LiP-MS).

LiP-MS experiments were conducted as previously described[72,73]. Briefly, Proteinase K (PK) was added at a 1:100 enzyme-to-substrate ratio to 2 μg of alpha-synuclein in 100 μl of LiP-buffer. Four independent PK digestions (technical replicates) were performed for each condition at 37°C for 5 min. Subsequently, PK activity was halted by incubating the samples at 99°C for 5 min. After the PK heat-inactivation step, all the samples were incubated for 5 minutes on ice. Sodium deoxycholate (DOC, 10%) was added to each sample to achieve a final DOC concentration of 5%. Next, the samples were reduced with tris(2-carboxyethyl)phosphine hydrochloride (TCEP) at a final concentration of 5 mM at 37 °C for 40 minutes with shaking at 800 rpm in a thermomixer (Eppendorf). Alkylation was performed with iodoacetamide (IAA) at a final concentration of 40 mM in the dark for 30 min at room temperature. To decrease the DOC concentration to 1%, the samples were diluted with 10 mM ammonium bicarbonate. Next, sequencing-grade porcine trypsin (Promega) and Lysyl endopeptidase LysC (Wako Chemicals) were added to the samples at an enzyme-to-substrate ratio of 1:100. The samples were digested overnight in a thermomixer at 37 °C with continuous agitation at 800 rpm. After overnight incubation, the digestion was stopped by adding 100% formic acid (FA) (Carl Roth GmbH) to achieve a final concentration of 2%. Finally, precipitated

DOC was removed by three cycles of 20 min centrifugation at 21,000 rcf. The samples were loaded onto Sep-Pak C18 cartridges (Waters), desalted, and subsequently eluted with 80% acetonitrile, 0.1% formic acid. The samples were analyzed by Orbitrap Eclipse Tribrid Mass Spectrometer (ThermoFisher Scientific) in DIA mode[74,75]. The raw MS data were searched with Spectronaut (Biognosys AG, version 15) and statistical data analysis was performed in R version 4.3.1. A two-tailed t-test was used to calculate p-values. Scores were designated to each detected peptide, calculated as a product of the $-\log_{10}$(p-value) and $|\log_2$(fold change)$|$. Subsequently, we calculated the mean value of this score for each amino acid based on overlapping peptides.

Acknowledgements
We would like to thank the staff at the ScopeM ETHZ for assistance with the cryo-EM data-collection analysis. This research was supported by the Swiss National Science Foundation (grant number 205320_182800).

Table 1: Cryo-EM samples analyzed for this manuscript

Dataset	Construct	Batch	Aggregation Condition	pH[3]	Polymorph(s)	Relative abundance[4]	Refinement stage
1	Ac-1-140[1]	2	PBS[2]	5.8	3B[5]:3C	93%:7%	3B, see Table 2 : 3C, 3D-Classification
2	Ac-1-140	6	PBS	5.8	3B:3C	54%:46%	3D-Classification
3	Ac-1-140	3	PBS + 50 mM NaCl	5.8	3B:3C	47%:53%	3C, see Table 2 : 3B, 3D-Classification
4	Ac-1-140	4	PBS + 100 mM NaCl	5.8	3B:3C	52%:48%	3D-Classification
5	Ac-1-140 (E46K)		PBS + 50 mM NaCl	5.8	3D		See Table 2
6	Ac-1-140	7	PBS	6.5	2A:2B:3B:3C	25%:18%:31%:26%	[7]Filament subset selection followed by 3D-Refinement to 3.51Å, 4.73Å, 4.56Å, 4.06Å respectively
7	Ac-1-140	4	20 mM Tris, 140 mM NaCl	7.0	1A		3D-Refinement 3.65 Å
8	Ac-1-140	4	PBS	7.0	5A		See Table 2
9	Ac-1-140	5	PBS	7.0	1_M		See Table 2
10	Ac-1-140	5	PBS	7.0	[6]non-twisted		
11	Ac-1-140	7	PBS	7.0	2A:2B:1_M	[8]30%:20%:50%	Filament subset selection followed by 3D-Refinement to 4.2 Å, 4.9 Å and 4.8 Å respectively
12	Ac-1-140	7	PBS	7.0	2A:2B:1_M	[8]16%:9%:75%	[7]Filament subset selection followed by 3D-Refinement to 5.4 Å, 6.4 Å and 5.8 Å respectively
13	Ac-1-140	8	PBS	7.0	2A:2B	78%:22%	[7]Filament subset selection - See Table 2
14	Ac-1-140	8	PBS	7.0	1A		3D-Refinement to 2.95 Å
15	Ac-1-140	8	PBS	7.0	1A		3D-Refinement to 3.94 Å
16	Ac-1-140	8	PBS	7.0	[6]non-twisted and clumped		
17	Ac-1-140	4	PBS	7.4	1A		3D-Refinement to 3.73 Å
18 (F3)	Ac-1-140	5	PBS + 5% pH 5.8 seeds	7.4	1A		3D-Refinement to 3.50 Å
19 (F4)	Ac-1-140	7	PBS + 5% pH 7.4 seeds	5.8	3B:3C	52%:48%	3D-Classification

[1] All constructs are full-length, N-terminally acetylated (Ac)

[2] PBS is a 10 mM phosphate buffer solution with 137 mM NaCl and 2.7 mM KCl
[3] The pH was adjusted after dissolving the PBS tablet (Sigma-Aldrich) by the addition of HCl
[4] In samples for which more than one polymorph could be identified by 3D-classification or Filament subset selection in RELION
[5] Underlining indicates data that was used for 3D refinement of the deposited maps and for building the deposited coordinates
[6] This data could not be analyzed by helical reconstruction as only non-twisted fibrils were present
[7] Filament subset selection was run after 2D Classification of the entire set of auto-picked particles. The identified filament classes were individually 2D classified and these classes used to create an initial model with relion_helix_inimodel2d which was used for automated 3D refinement. In the case of dataset 13 followed by CTF refinement and Bayesian polishing in RELION
[8] Due to the small size of these datasets, the relative abundances are not likely to be precise.

Table 2: Cryo-EM structure determination statistics

Polymorph	1_M	2A	2B	3B	3C	5A	3D (E46K)
Data Collection							
Pixel size [Å]	0.65	0.65	0.65	0.65	0.65	0.65	0.65
Defocus Range [µm]	-0.8 to -2.5	0.8 to -2.5	0.8 to - 2.5	-0.8 to - 2.5	-0.8 to - 2.5	-0.8 to - 2.5	-0.8 to - 2.5
Voltage [kV]	300	300	300	300	300	300	300
Number of frames	40	40	40	40	40	40	40
Total dose [e- /Å^2]	69	58.5	58.5	62	67	75	65
Reconstruction							
Reconstruction Box width [pixels]	256	200	200	512	256	256	512
Inter-box distance [Å]	33	33	33	33	33	33	33
Reconstruction Pixel size [Å]	1.3	1.3	1.3	0.65	1.3	1.3	0.65
Micrographs	1,624	1339	1339	7,729	3,127	1,850	4,666
Initially extracted segments	84,666	238,570	238,570	279,929	159,308	109,817	287,018
Segments after 2D and 3D classification	19,800	91,238	26,426	178,710	28,022	86,219	40,181
3D refinement Resolution [Å] (FSC > 0.143)	3.58	2.99	3.13	2.64	3.41	3.40	2.40
Final resolution [Å] (FSC > 0.143)	3.26	2.86	2.95	2.23	3.41	3.30	2.31
Estimated map sharpening B-factor [Å^2]	-43.8	-73.0	-74.9	-51.7	-101.7	-87.4	-32.8
Axial symmetry	C1	C2	C1	C1	C2	C1	C1
Helical rise [Å]	4.79	4.82	2.39	2.37	4.77	2.42	2.41
Helical twist [°]	-0.95	-0.80	179.6	179.5	-0.995	179.6	179.5
Model composition and validation							
Non-hydrogen atoms (5 layers)	2515	----	----	4550	4500	6690	4450
Protein residues (5 layers)	365	----	----	640	650	960	630
R.m.s. devations bond length [Å]	0.008	----	----	0.005	0.007	0.006	0.007
R.m.s. devations bond angles [°]	1.166	----	----	0.637	0.976	0.863	1.093
MolProbity score	1.66	----	----	2.01	1.33	1.64	1.60
Clashscore	3.71	----	----	7.42	1.62	4.72	3.57
Rotamer outliers [%]	0	----	----	2.22	1.3	0	0.68
Ramachandran plot favored [%]	91.6	----	----	95.2	95.2	94.2	92.8
Ramachandran plot allowed [%]	81.4	----	----	4.8	4.8	5.8	7.2
Ramachandran plot disallowed [%]	0	----	----	0	0	0	0
Model Resolution [Å] (FSC > 0.143 / 0.5)	3.1 / 3.3	----	----	2.0 / 2.2	3.1 / 3.3	3.1 / 3.4	2.0 / 2.4
PDB code	8PK2	----	----	9FYP	8PIX	8PK4	8PJO
EMDB-ID	17723	50860	50077	50888	17693	17726	17714

Bibliography

1. Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* **75**, (2006).
2. Meier, B. H., Riek, R. & Böckmann, A. Emerging Structural Understanding of Amyloid Fibrils by Solid-State NMR. *Trends Biochem Sci* **42**, (2017).
3. He, S. & Scheres, S. H. W. Helical reconstruction in RELION. *J Struct Biol* **198**, (2017).
4. Iadanza, M. G., Jackson, M. P., Hewitt, E. W., Ranson, N. A. & Radford, S. E. A new era for understanding amyloid structures and disease. *Nat Rev Mol Cell Biol* **19**, (2018).
5. Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R. & Eisenberg, D. S. The expanding amyloid family: Structure, stability, function, and pathogenesis. *Cell* **184**, (2021).
6. Riek, R. & Eisenberg, D. S. The activities of amyloids from a structural perspective. *Nature* **539**, (2016).
7. Törnquist, M. *et al.* Secondary nucleation in amyloid formation. *Chemical Communications* **54**, (2018).
8. Uemura, N. *et al.* α-Synuclein aggregates amplified from patient-derived Lewy bodies recapitulate Lewy body diseases in mice. *Nat Commun* **14**, (2023).
9. Peelaerts, W. *et al.* α-Synuclein strains cause distinct synucleinopathies after local and systemic administration. *Nature* **522**, (2015).
10. Guo, J. L. *et al.* Distinct α-Synuclein Strains Differentially Promote Tau Inclusions in Neurons. *Cell* **154**, (2013).
11. Van der Perren, A. *et al.* The structural differences between patient-derived α-synuclein strains dictate characteristics of Parkinson's disease, multiple system atrophy and dementia with Lewy bodies. *Acta Neuropathol* **139**, (2020).
12. Bousset, L. *et al.* Structural and functional characterization of two alpha-synuclein strains. *Nat Commun* **4**, (2013).
13. Shi, Y. *et al.* Structure-based classification of tauopathies. *Nature* **598**, (2021).
14. Li, Y. *et al.* Amyloid fibril structure of α-synuclein determined by cryo-electron microscopy. *Cell Res* **28**, (2018).
15. Li, B. *et al.* Cryo-EM of full-length α-synuclein reveals fibril polymorphs with a common structural kernel. *Nat Commun* **9**, (2018).
16. Guerrero-Ferreira, R. *et al.* Cryo-EM structure of alpha-synuclein fibrils. *Elife* **7**, (2018).
17. Zhao, K. *et al.* 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**, (2020).
18. Zhao, K. *et al.* Parkinson's disease associated mutation E46K of α-synuclein triggers the formation of a distinct fibril structure. *Nat Commun* **11**, (2020).
19. Sun, Y. *et al.* Cryo-EM structure of full-length α-synuclein amyloid fibril with Parkinson's disease familial A53T mutation. *Cell Res* **30**, (2020).
20. Ni, X., McGlinchey, R. P., Jiang, J. & Lee, J. C. Structural Insights into α-Synuclein Fibril Polymorphism: Effects of Parkinson's Disease-Related C-Terminal Truncations. *J Mol Biol* **431**, (2019).
21. Boyer, D. R. *et al.* Structures of fibrils formed by α-synuclein hereditary disease mutant H50Q reveal new polymorphs. *Nat Struct Mol Biol* **26**, (2019).
22. Guerrero-Ferreira, R. *et al.* Two new polymorphic structures of human full-length alpha-synuclein fibrils solved by cryo-electron microscopy. *Elife* **8**, (2019).
23. Boyer, D. R. *et al.* The α-synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. *Proc Natl Acad Sci U S A* **117**, (2020).
24. Schweighauser, M. *et al.* Structures of α-synuclein filaments from multiple system atrophy. *Nature* **585**, (2020).

25. Long, H. *et al.* Wild-type α-synuclein inherits the structure and exacerbated neuropathology of E46K mutant fibril strain by cross-seeding. *Proceedings of the National Academy of Sciences* **118**, (2021).
26. Sun, Y. *et al.* The hereditary mutation G51D unlocks a distinct fibril strain transmissible to wild-type α-synuclein. *Nat Commun* **12**, (2021).
27. Hojjatian, A. *et al.* Distinct cryo-EM Structure of α-synuclein Filaments derived by Tau. *bioRxiv* (2021).
28. McGlinchey, R. P., Ni, X., Shadish, J. A., Jiang, J. & Lee, J. C. The N terminus of α-synuclein dictates fibril formation. *Proceedings of the National Academy of Sciences* **118**, (2021).
29. Lövestam, S. *et al.* Seeded assembly in vitro does not replicate the structures of α-synuclein filaments from multiple system atrophy. *FEBS Open Bio* **11**, (2021).
30. Frieg, B. *et al.* α-Synuclein polymorphism determines oligodendroglial dysfunction. *bioRxiv* (2021).
31. Sun, C. *et al.* Cryo-EM structure of amyloid fibril formed by α-synuclein hereditary A53E mutation reveals a distinct protofilament interface. *Journal of Biological Chemistry* **299**, (2023).
32. Zhao, Q. *et al.* Structural Insights of Fe3+ Induced α-synuclein Fibrillation in Parkinson's Disease. *J Mol Biol* **435**, (2023).
33. Zhang, S. *et al.* Conformational Dynamics of an α-Synuclein Fibril upon Receptor Binding Revealed by Insensitive Nuclei Enhanced by Polarization Transfer-Based Solid-State Nuclear Magnetic Resonance and Cryo-Electron Microscopy. *J Am Chem Soc* **145**, (2023).
34. Yang, Y. *et al.* Structures of α-synuclein filaments from human brains with Lewy pathology. *Nature* **610**, (2022).
35. Frieg, B. *et al.* The 3D structure of lipidic fibrils of α-synuclein. *Nat Commun* **13**, (2022).
36. Yang, Y. *et al.* New SNCA mutation and structures of α-synuclein filaments from juvenile-onset synucleinopathy. *Acta Neuropathol* **145**, (2023).
37. Fan, Y. *et al.* Conformational change of α-synuclein fibrils in cerebrospinal fluid from different clinical phases of Parkinson's disease. *Structure* **31**, (2023).
38. Prusiner, S. B. Prions. *Proc Natl Acad Sci U S A* **95**, (1998).
39. Aguzzi, A. & Rajendran, L. The Transcellular Spread of Cytosolic Amyloids, Prions, and Prionoids. *Neuron* **64**, (2009).
40. Saborio, G. P., Permanne, B. & Soto, C. Sensitive detection of pathological prion protein by cyclic amplification of protein misfolding. *Nature* **411**, (2001).
41. Soto, C., Saborio, G. P. & Anderes, L. Cyclic amplification of protein misfolding: Application to prion-related disorders and beyond. *Trends Neurosci* **25**, (2002).
42. Srivastava, A., Alam, P. & Caughey, B. RT-QuIC and Related Assays for Detecting and Quantifying Prion-like Pathological Seeds of α-Synuclein. *Biomolecules* **12**, (2022).
43. Gerez, J. A. & Riek, R. Neurodegenerative diseases distinguished through protein-structure analysis. *Nature* **578**, (2020).
44. Strohäker, T. *et al.* Structural heterogeneity of α-synuclein fibrils amplified from patient brain extracts. *Nat Commun* **10**, (2019).
45. Burger, D., Fenyi, A., Bousset, L., Stahlberg, H. & Melki, R. Cryo-EM structure of alpha-synuclein fibrils amplified by PMCA from PD and MSA patient brains. *bioRxiv* (2021).
46. Peduzzo, A., Linse, S. & Buell, A. K. The Properties of α-Synuclein Secondary Nuclei Are Dominated by the Solution Conditions Rather than the Seed Fibril Strain. *ACS Chem Neurosci* **11**, (2020).
47. Zarranz, J. J. *et al.* The New Mutation, E46K, of α-Synuclein Causes Parkinson and Lewy Body Dementia. *Ann Neurol* **55**, (2004).
48. Pálmadóttir, T. *et al.* Morphology-Dependent Interactions between α-Synuclein Monomers and Fibrils. *Int J Mol Sci* **24**, (2023).

49. Rogers, D. R. Screening for amyloid with the thioflavin-T fluorescent method. *Am J Clin Pathol* **44**, (1965).
50. LeVine, H. Quantification of β-sheet amyloid fibril structures with thioflavin T. in *Amyloid, Prions, and Other Protein Aggregates* vol. 309 (Academic Press, 1999).
51. Schopper, S. *et al.* Measuring protein structural changes on a proteome-wide scale using limited proteolysis-coupled mass spectrometry. *Nat Protoc* **12**, (2017).
52. de Souza, N. & Picotti, P. Mass spectrometry analysis of the structural proteome. *Curr Opin Struct Biol* **60**, (2020).
53. Jarrett, J. T. & Lansbury, P. T. Seeding 'one-dimensional crystallization' of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **73**, (1993).
54. Jarrett, J. T. & Lansbury, P. T. Amyloid Fibril Formation Requires a Chemically Discriminating Nucleation Event: Studies of an Amyloidogenic Sequence from the Bacterial Protein OsmB. *Biochemistry* **31**, (1992).
55. Knowles, T. P. J. *et al.* Kinetics and thermodynamics of amyloid formation from direct measurements of fluctuations in fibril mass. *Proc Natl Acad Sci U S A* **104**, (2007).
56. Kumari, P. *et al.* Structural insights into α-synuclein monomer–fibril interactions. *Proc Natl Acad Sci U S A* **118**, (2021).
57. Martinez-Valbuena, I. *et al.* Alpha-synuclein seeding shows a wide heterogeneity in multiple system atrophy. *Transl Neurodegener* **11**, (2022).
58. Buell, A. K. *et al.* Solution conditions determine the relative importance of nucleation and growth processes in α-synuclein aggregation. *Proceedings of the National Academy of Sciences* **111**, (2014).
59. Johnson, M., Coulton, A. T., Geeves, M. A. & Mulvihill, D. P. Targeted amino-terminal acetylation of recombinant proteins in E. coli. *PLoS One* **5**, (2010).
60. Campioni, S. *et al.* The presence of an air-water interface affects formation and elongation of α-synuclein fibrils. *J Am Chem Soc* **136**, (2014).
61. Scheres, S. H. W. Amyloid structure determination in RELION-3.1. *Acta Crystallogr D Struct Biol* **76**, (2020).
62. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, (2015).
63. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, (2019).
64. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, (2020).
65. Jamali, K. *et al.* Automated model building and protein identification in cryo-EM maps. *bioRxiv* (2023).
66. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, (2004).
67. Croll, T. I. ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, (2018).
68. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, (2019).
69. McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallogr D Biol Crystallogr* **67**, (2011).
70. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* **25**, (2004).
71. Afonine, P. V. *et al.* New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D Struct Biol* **74**, (2018).
72. Schopper, S. *et al.* Measuring protein structural changes on a proteome-wide scale using limited proteolysis-coupled mass spectrometry. *Nat Protoc* **12**, (2017).
73. Malinovska, L. *et al.* Proteome-wide structural changes measured with limited proteolysis-mass spectrometry: an advanced protocol for high-throughput applications. *Nat Protoc* **18**, (2023).

74. Gillet, L. C. *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. *Molecular and Cellular Proteomics* **11**, (2012).
75. Cappelletti, V. *et al.* Dynamic 3D proteomes reveal protein functional alterations at high resolution in situ. *Cell* **184**, (2021).

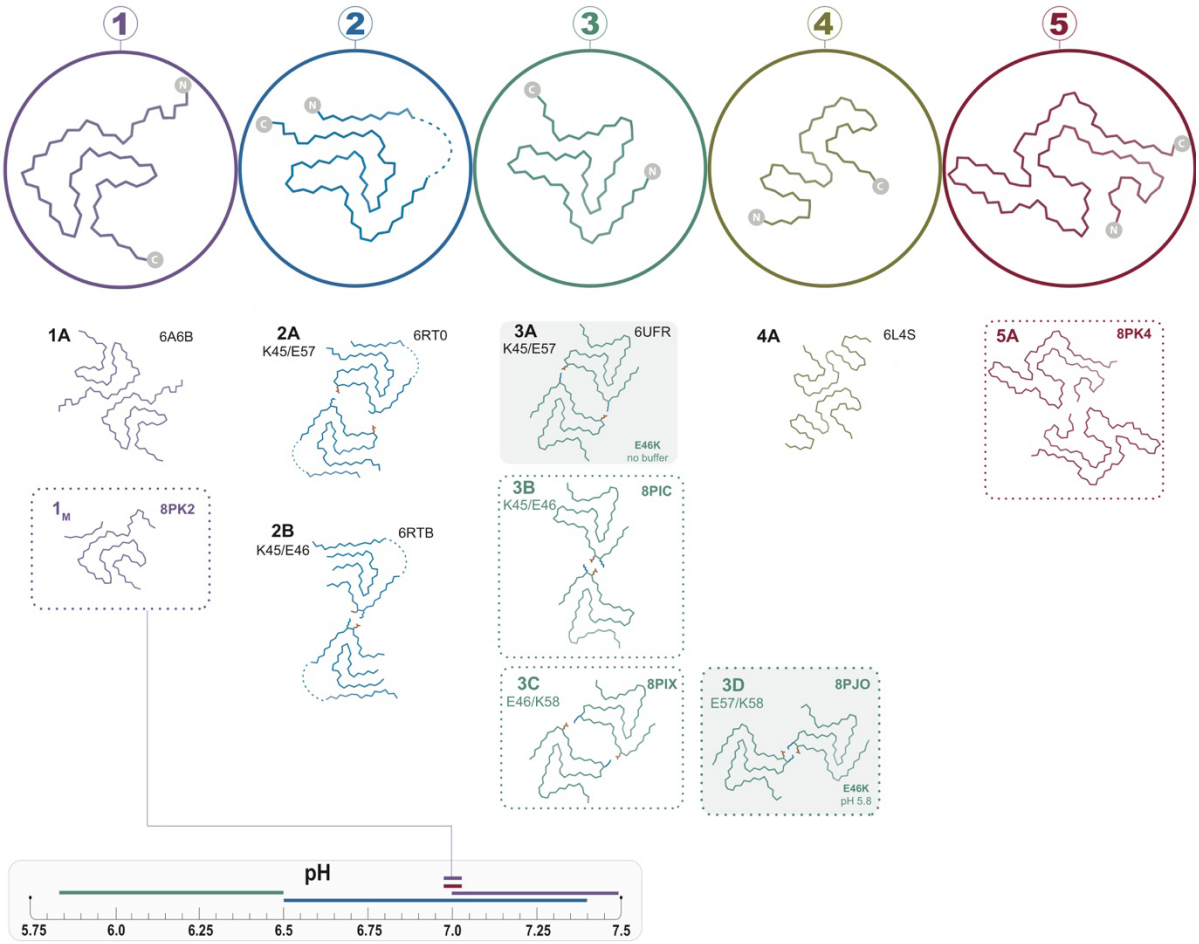

Figure 1. The polymorphs formed *in vitro* by wild-type α-Syn. The numbered backbone Cα traces depicting a single chain from each of the five Types of polymorphs that have been reported to form *in vitro* with wild-type α-Syn but without co-factors. The interface variants of each type are also shown by a representative structure from the PDB or from one of the structures reported in this manuscript (boxed in dotted lines). Two mutant structures for the Type 3 fold are included in order to show the special 3A and 3D interfaces that form for these mutants (shaded background). The observed pH-dependence of each of the polymorphs described in this manuscript is indicated by the colored line above the pH scale (or, in the case of the 1_M polymorph reported here, by a line connecting it to the scale).

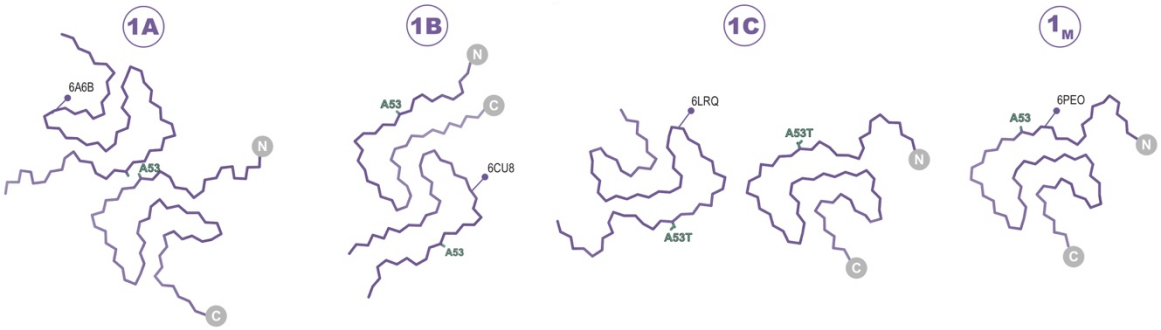

Figure 2. The Type 1 polymorphs. Backbone Cα traces showing the 4 different interface polymorphs of Type 1, including 1B whose fold is not classically Type 1 but whose name is kept for consistency with previous publications. The 1C polymorph is only formed by mutants at the 1A interface (here A53T) and so residue 53 is highlighted in each structure.

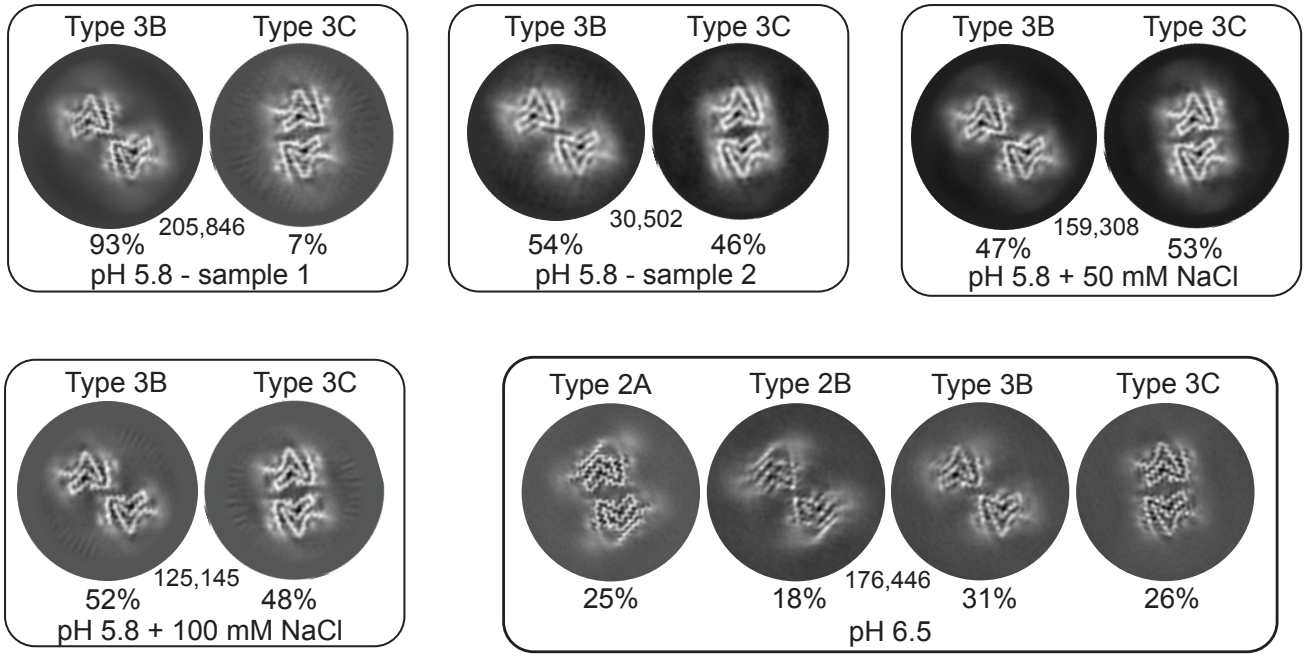

Figure 3. 3D classification of particles in fibril samples from various conditions. The 3D-classes obtained in RELION for five different fibril samples grown at pH 5.8 and 6.5 in PBS with or without additional NaCl are shown. Each box presents the output classes from one sample with the ratio of the polymorphs determined by the number of particles assigned to each class and the total number of particles used indicated in the center of the box. Classifications were performed with input models for each of the output classes plus one additional cylindrical model (not shown) to provide a "junk" class for particles that do not fit any of the main classes. The particles that went into the "junk" class were not included in determining that ratios of particles in the main classes. In the case of the pH 6.5 data, the fibril classes were better separated by filament subset selection in RELION (see methods and Table 1) and the 4 classes shown represent the output from individual 3D refinements.

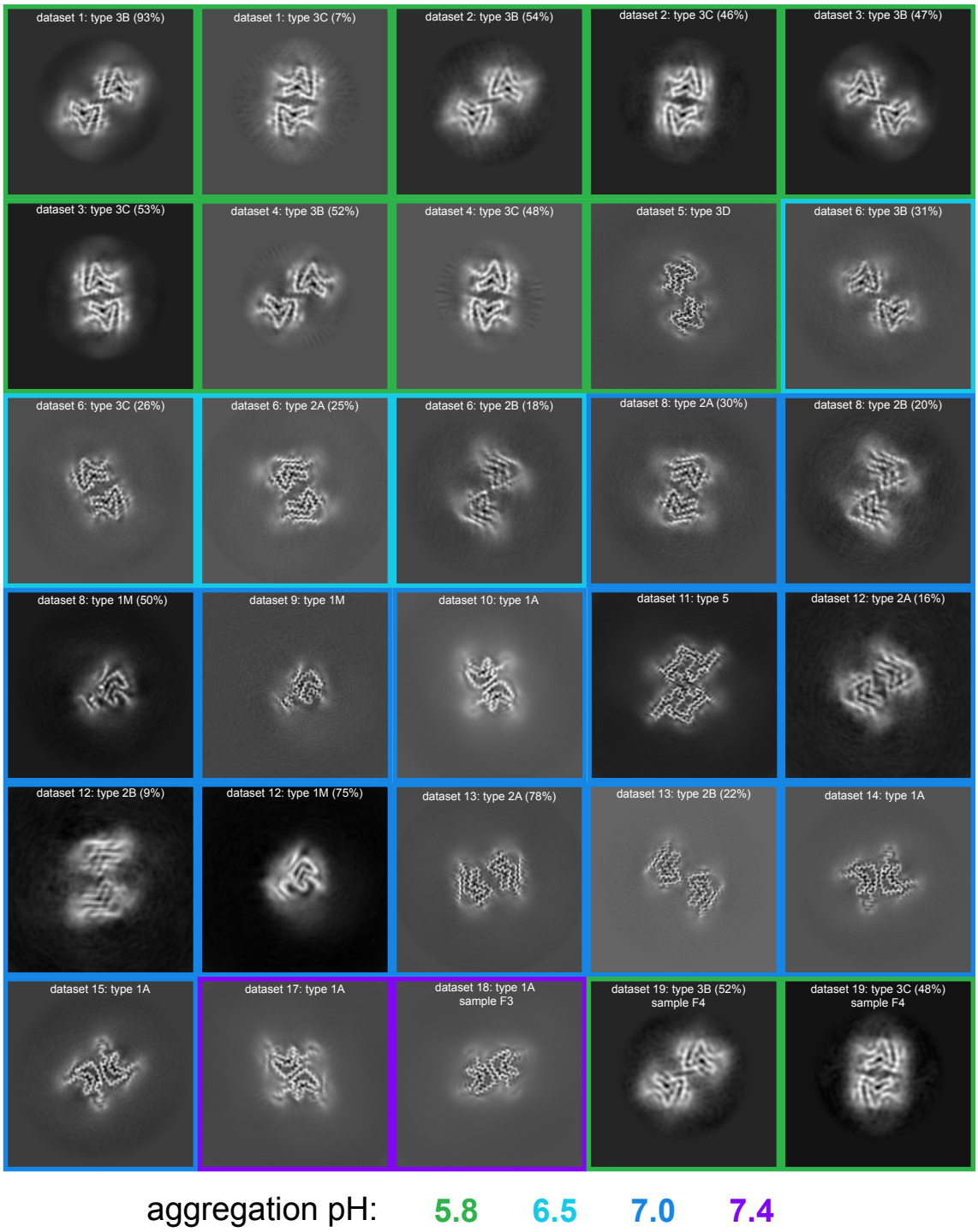

Figure 4. pH dependence of observed α-Syn polymorphs. An overview off all of the Cryo-EM-resolved polymorphs in this study. The images represent a 9.5 Å slice (about 2 layers of the amyloid) of the 3D map projected in the Z direction. The 3D maps were obtained either from 3D classification or 3D refinement in RELION. The datasets are ordered and numbered as in Table 1. The colored borders indicate the pH at which the fibrils were grown. For datasets from which multiple polymorphs were resolved, the relative percentage of particles corresponding to each polymorph is indicated in parentheses.

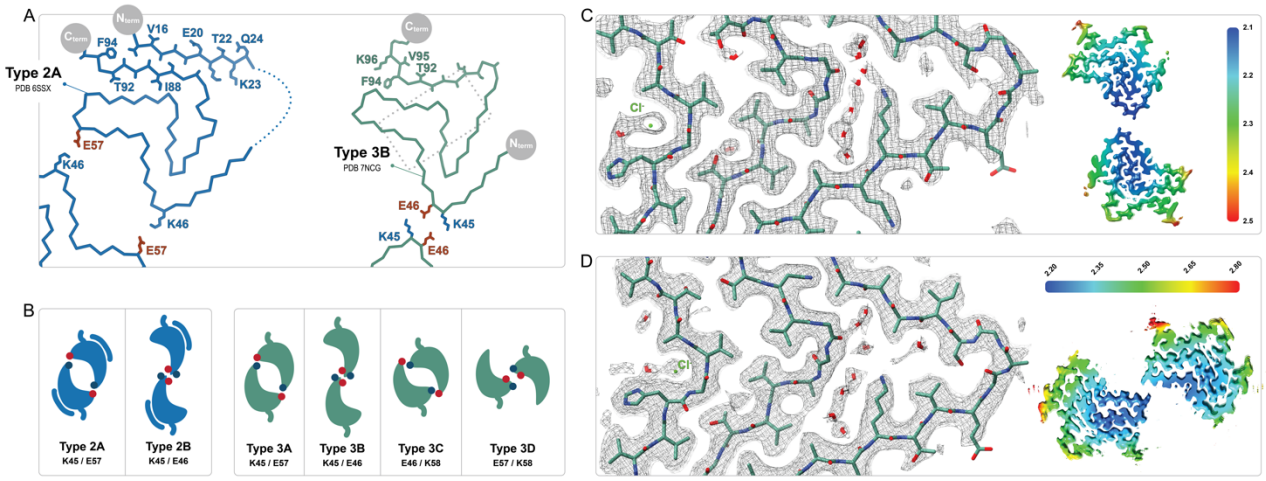

Figure 5. Comparison of the Types 2 and 3 polymorphs. A) Protofilaments of Polymorphs 2A (blue) and 3B (green) are depicted as Cα traces. The sidechains are included on the C-terminal β-strand and the N-terminal segment in the 2A structure in order to highlight the in-side-out flipped orientations of their C-terminal β-strands. The dashed box indicates the region highlighted in panels **C** and **D. B**) The 2 interfaces of Type 2 and the 4 interfaces of Type 3 polymorphs are shown as a cartoon schematic. The charged interfacial residues are indicated by blue/red dots for K/E and listed below each interface. The C-term is indicated as a short tail and Type 2 has the extra N-terminal segment. **C**) A close-up view of the Type 3B structure (PDB:8PIC) and **D**) Type 3D (E46K) structure (PDB:8PJO) and their EM density, showing the shared set of immobilized water molecules and strong density that has been modeled as a chloride ion. The local resolution maps for each structure are shown to the right with the color scale indicating the resolution range in Å.

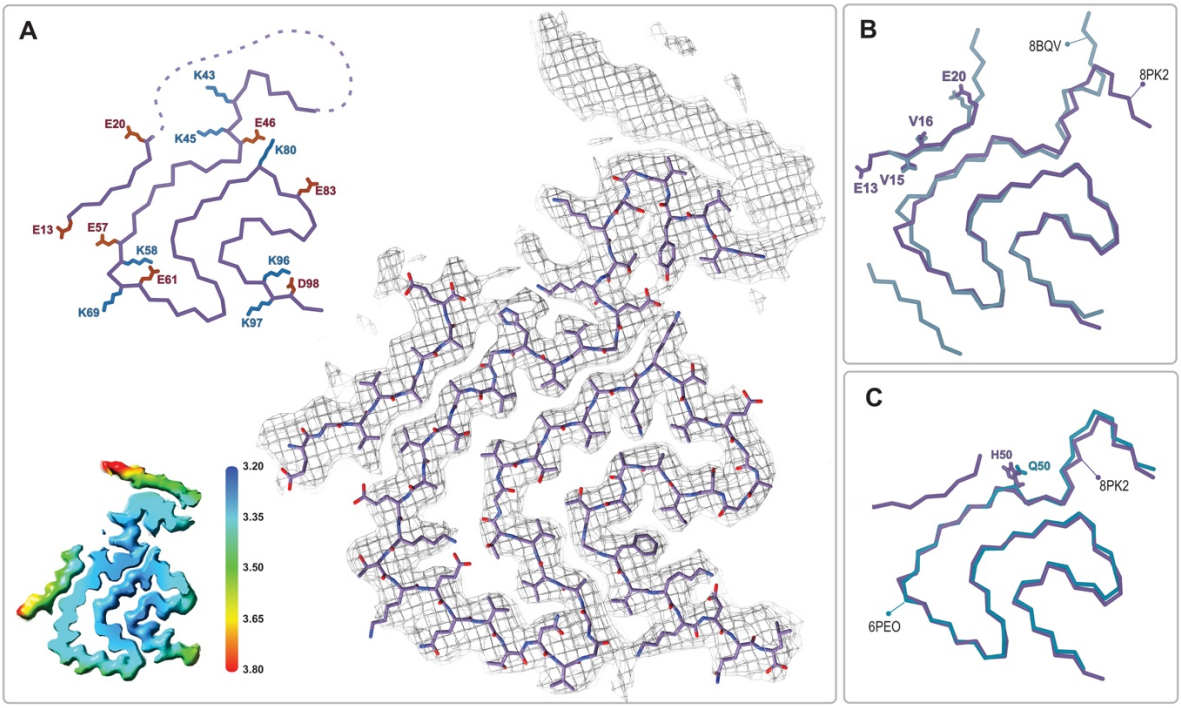

Figure 6. The JOS-like Type 1_M polymorph. A) The Type 1_M monofilament structure (PDB:8PK2) overlaid with its EM density (including the unmodeled density). The Cα trace is also shown on the left with all K and E side chains indicated as well as the local resolution map for the density with color scale showing the resolution range in Å. **B**) An overlay of the Cα trace of the Type 1_M and the JOS polymorph (PDB:8BQV) from which the identity of residues 13-20 in the N-terminal strand were assigned. **C**) An overlay of the Cα trace of the JOS-like Type 1_M to the 1_M structure of the H50Q mutant (PDB:6PEO). The structural alignments were done by an LSQ-superposition of residues 51-66.

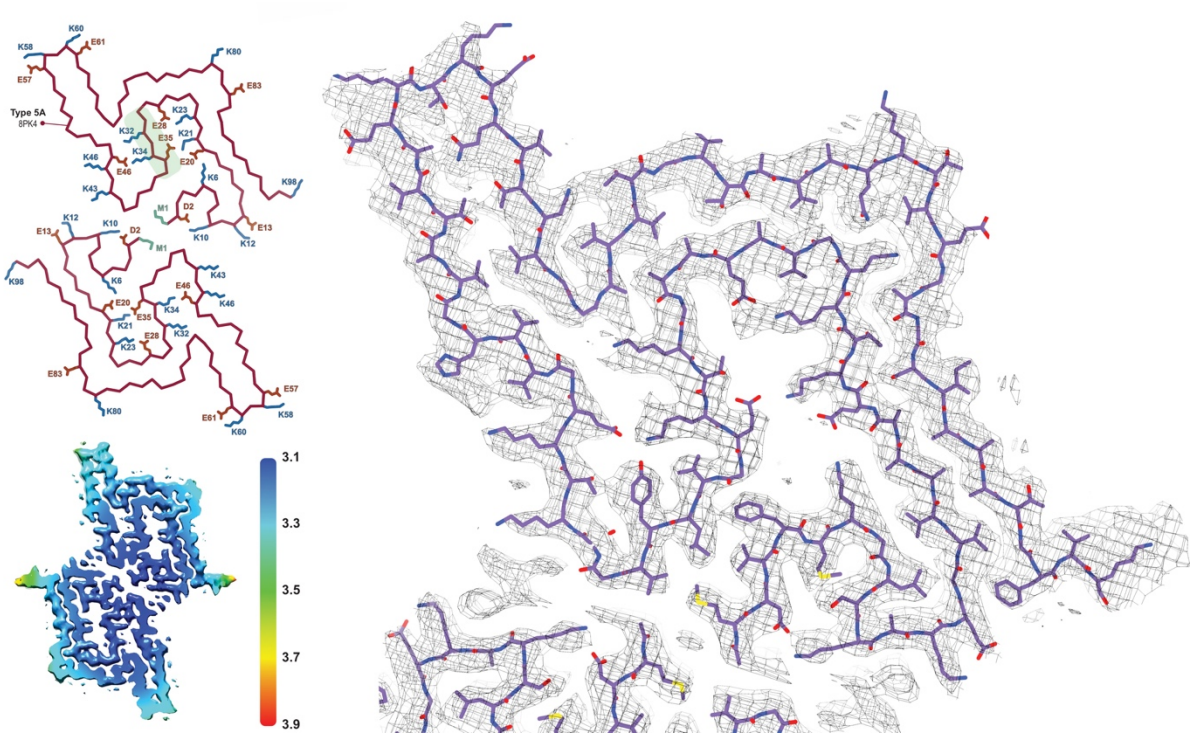

Figure 7. The Type 5 polymorph. The structure of a Type 5 protofilament (PDB:8PK4) overlaid with its EM density. The Cα trace for the two filaments of the Type 5A fibril is also shown to the left. All of the charged (E/D/K) residues and the acetylated N-terminal Met are labeled and the polar segment that dissects the two cavities is shaded green in the upper chain. The local resolution map is also shown with the color scale indicating the resolution in Å.

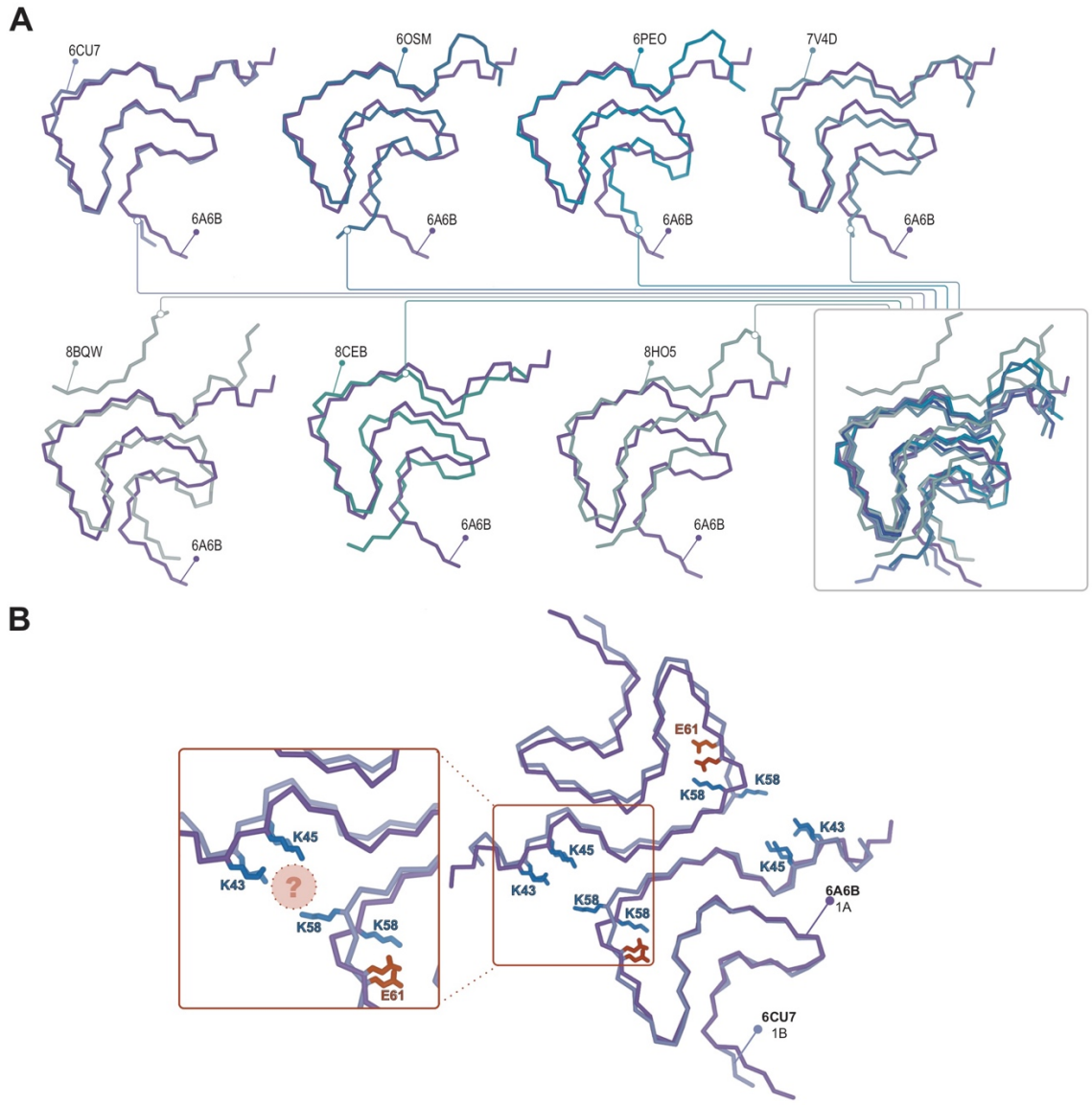

Figure 8. Structural variability within the Type 1 polymorphs. A) Seven pairwise overlays of the wild-type 1A structure (PDB:6A6B) with other wild-type and mutant structures depicting the range of structural variability that is found in the Type 1 fold. The structural alignments were done by an LSQ-superposition of the Cα atoms of residues 51-66. The other Type 1 structures from left to right, top to bottom are wild-type 1A without buffer (PDB:6CU7), N-terminally acetylated wild-type 1A residues 1-103 (PDB:6OSM), the H50Q mutant 1_M without buffer (PDB:6PEO), wild-type 1A in Tris buffered saline (PDB:7V4D), the patient-derived JOS 1_M polymorph (PDB:8BQV), mix of wild-type and seven residue JOS-associated insertion mutant 1A in PBS (PDB:8CEB) and N-terminally acetylated wild-type 1_M seeded from PD patient CSF in PIPES with 500 mM NaCl. **B**) Same overlay shown at the top right in **A**, showing the triad of lysine residues often formed at the interface in the presence of phosphate (PDB:6CU7) compared the inward facing orientation of K58 in the absence of phosphate (PDB:6A6B). The location of often observed density, thought to be PO_4^{2-} is indicated with the pink sphere.

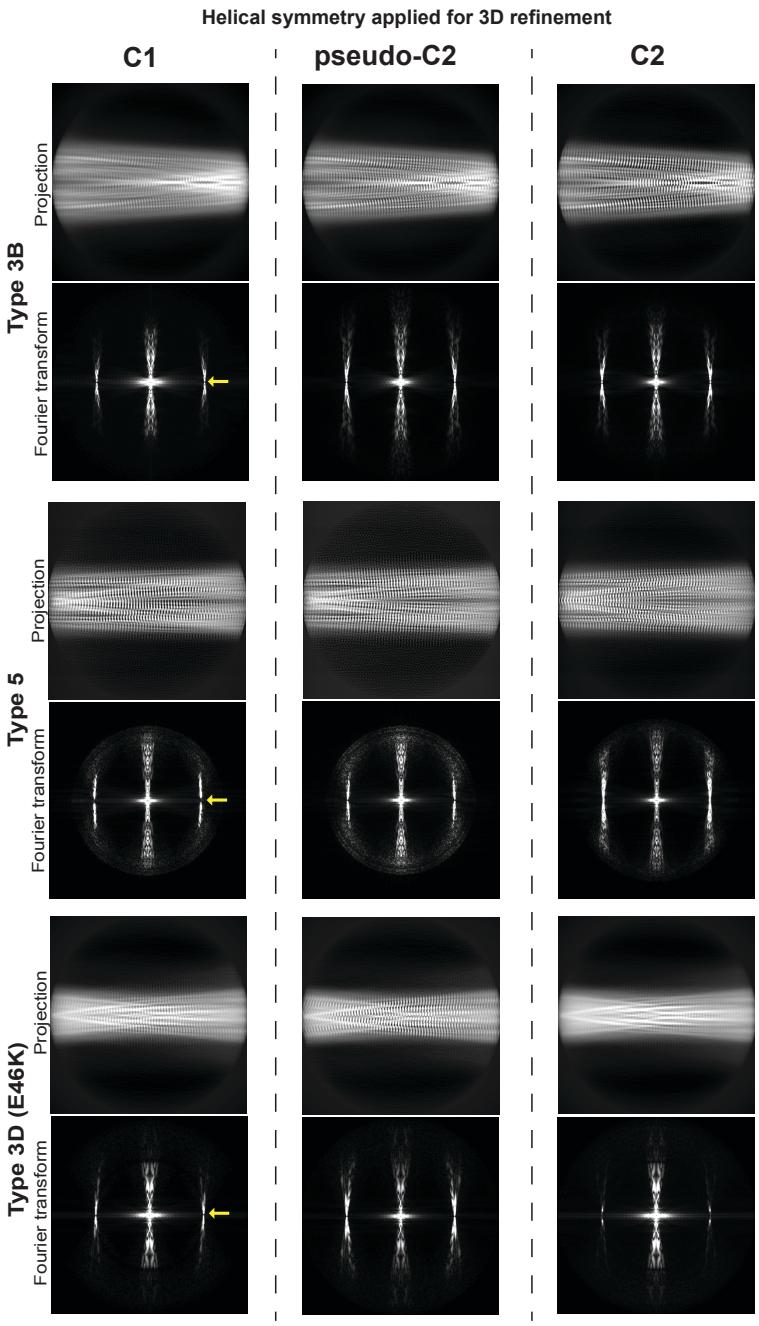

Figure 9. Determination of helical symmetry in Type 3B, 3D and 5A polymorphs. 2D projections and their Fourier transforms for three different structures (rows) refined with three different types of symmetry (columns). Each of the structures studied was refined as far as possible in a C1 symmetry with a ca. -1° twist and 4.7 Å rise in order to examine the higher order symmetry expected to be present: either C2 with a 4.7 Å rise or a pseudo-2_1 helical symmetry with ca. -179.5 twist and 2.35 Å rise. A comparison of the projections of the refined volumes indicated that all three of these structures have a pseudo-2_1 helical symmetry. This can be seen in the C1 projections which lack a mirror symmetry down their middle and in their Fourier transforms which lack an n=0 Bessel function meridional peak in the layer line at 1/4.7 Å (location marked with a yellow arrow).

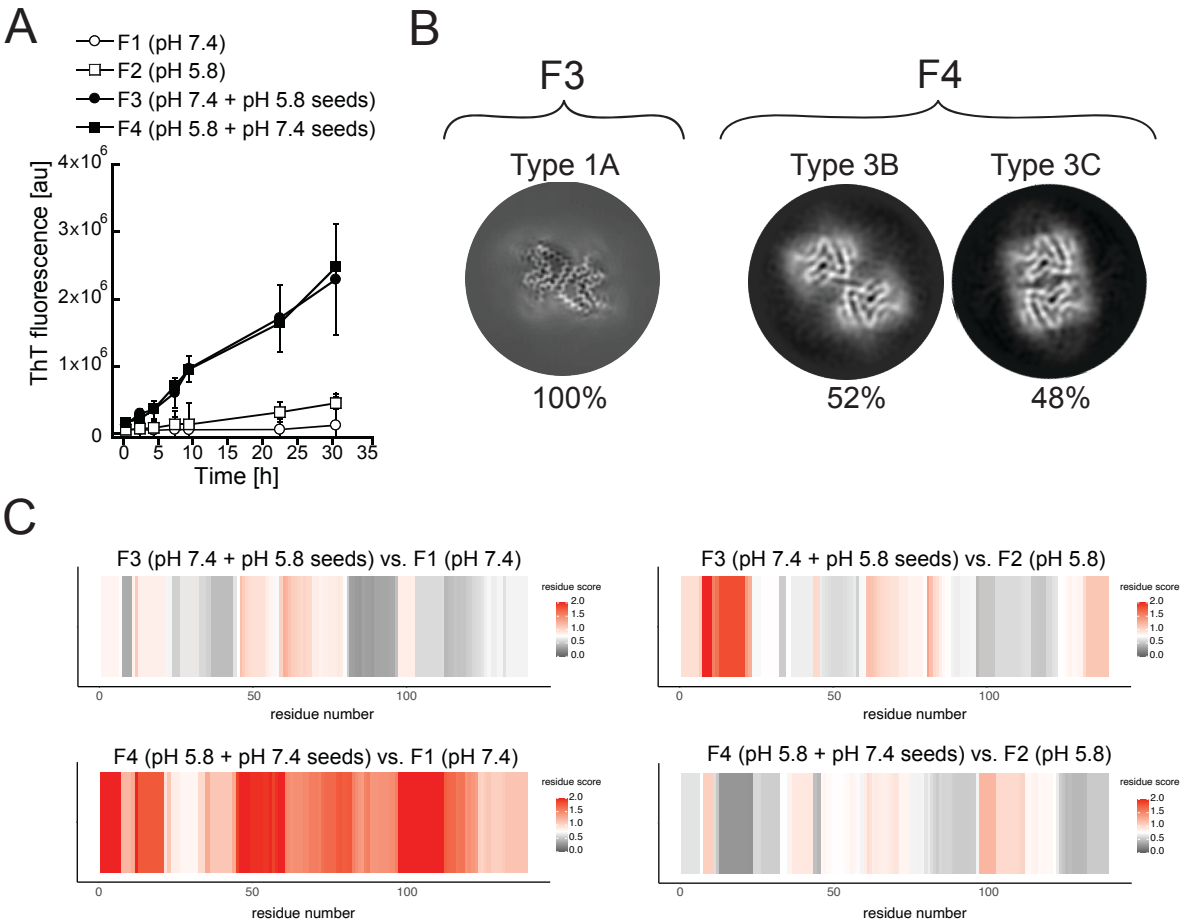

Figure 10. Cross-seeding does not preserve the seed polymorph. Fibrils produced at pH 7.4 (F1) and pH 5.8 (F2) were used as seeds to generate new fibril samples with fresh alpha-synuclein monomer: at pH 7.4 with 5% unfragmented pH 5.8 seeds (F3) or at pH 5.8 with pH 7.4 seeds (F4). **A**) The effect of cross-seeding with fibrils produced in a different pH from that of the aggregation condition is shown in the aggregations kinetics, as monitored by ThT fluorescence. **B**) Cryo-EM analyses of the seeded samples yielded a single 3D class representing polymorph 1A. (here depicted as a Z-section) for the F3 fibrils and two classes representing Types 3B and 3C for the F4 fibrils. **C**) LiP-MS analysis showing the differences in fibrillar structures in bulk solution. The comparison of seeded fibrils (F3 and F4) with F1 is presented on the left, while the comparison of seeded fibrils with F2 is presented on the right. Differences in structures per residue are plotted along the sequence of alpha-synuclein in a form of scores ($-\log_{10}$(p-value) $\times$ $|\log_2$(fold change)$|$). The more intense the red color, the greater the difference between the two structures in each region. Grey indicates no significant differences, while white indicates the significance threshold corresponding to ($-\log_{10}(0.05) \times |\log_2(1.5)|$).

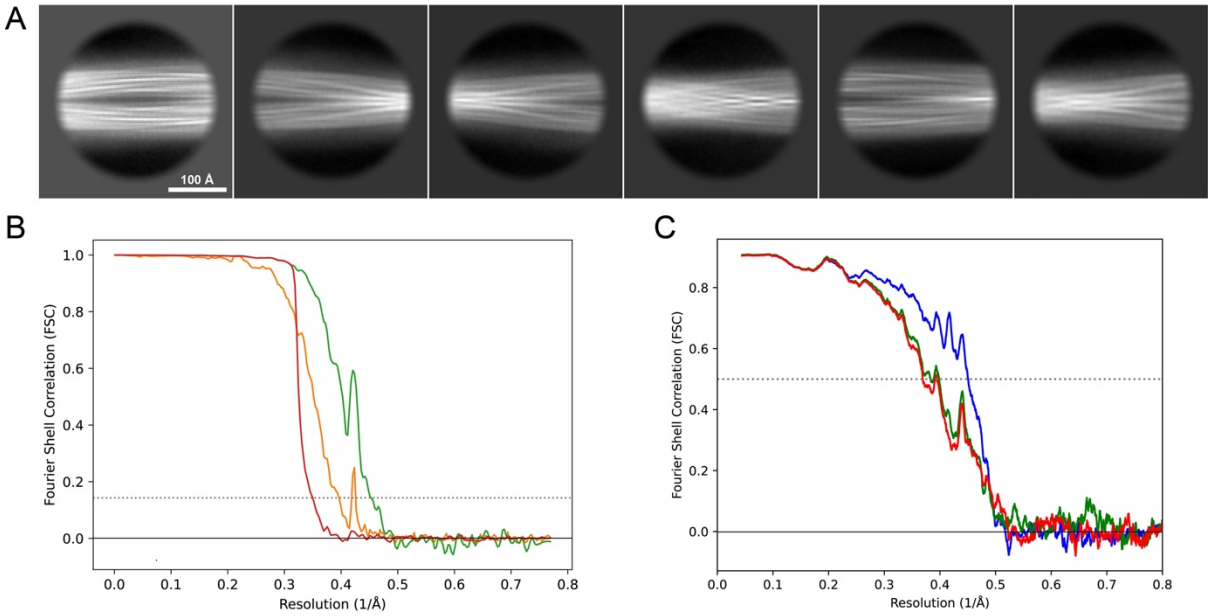

Figure 11. 2D classes and half-map and model-map FSC curves for dataset 1, Type 3B. A) Representative 2D classes of the segments that were used for the 3D reconstruction. **B**) The FSC curves produced during postprocessing in RELION with red showing the plot for the phase randomized, orange the unmasked and green the masked maps. **C**) The model-map FSC curves produced in PHENIX. The blue curve is for the deposited coordinates and full postprocessed map against which it had been fit by real space refinement in PHENIX. Coordinates were similarly generated by refining against the first half map and then compared to the same half-map (green) or the second half-map (red)

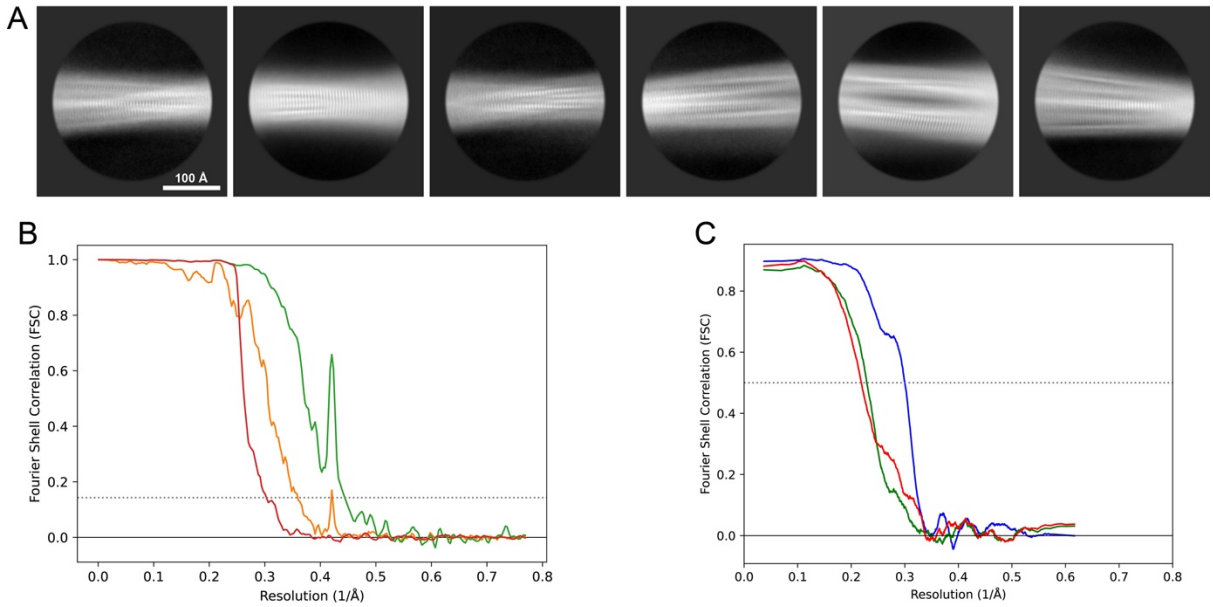

Figure 12. 2D classes and half-map and model-map FSC curves for dataset 3, Type 3C. A) Representative 2D classes of the segments that were used for the 3D reconstruction. **B**) The FSC curves produced during postprocessing in RELION with red showing the plot for the phase randomized, orange the unmasked and green the masked maps. **C**) The model-map FSC curves produced in PHENIX. The blue curve is for the deposited coordinates and full postprocessed map against which it had been fit by real space refinement in PHENIX. Coordinates were similarly generated by refining against the first half map and then compared to the same half-map (green) or the second half-map (red)

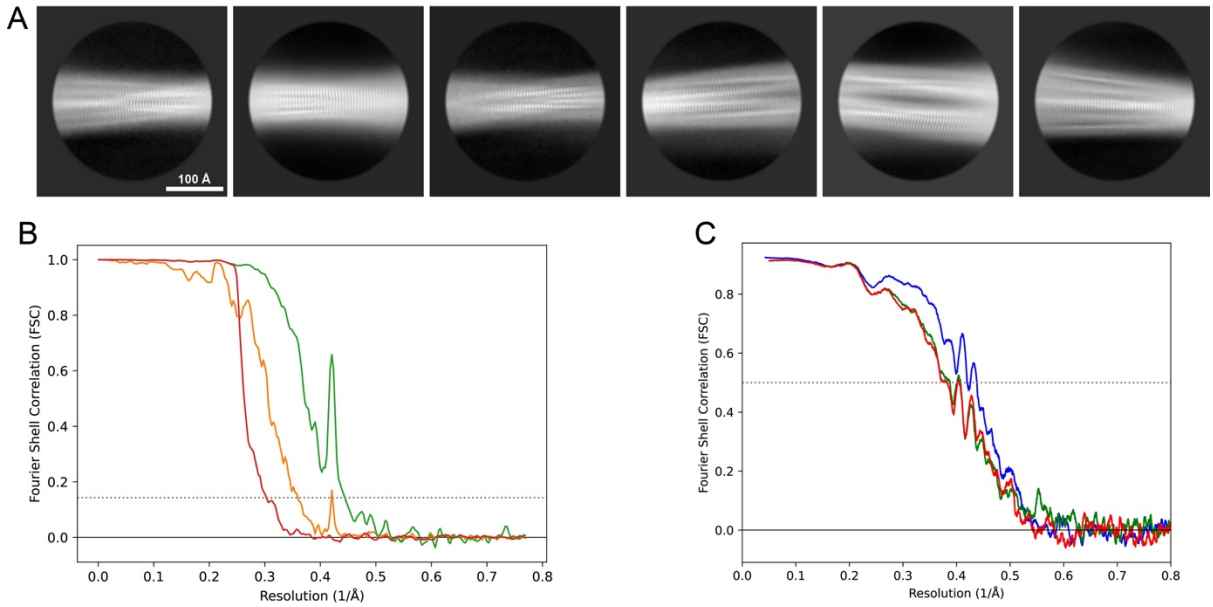

Figure 13. 2D classes and half-map and model-map FSC curves for dataset 5, Type 3D. A) Representative 2D classes of the segments that were used for the 3D reconstruction. **B**) The FSC curves produced during postprocessing in RELION with red showing the plot for the phase randomized, orange the unmasked and green the masked maps. **C**) The model-map FSC curves produced in PHENIX. The blue curve is for the deposited coordinates and full postprocessed map against which it had been fit by real space refinement in PHENIX. Coordinates were similarly generated by refining against the first half map and then compared to the same half-map (green) or the second half-map (red)

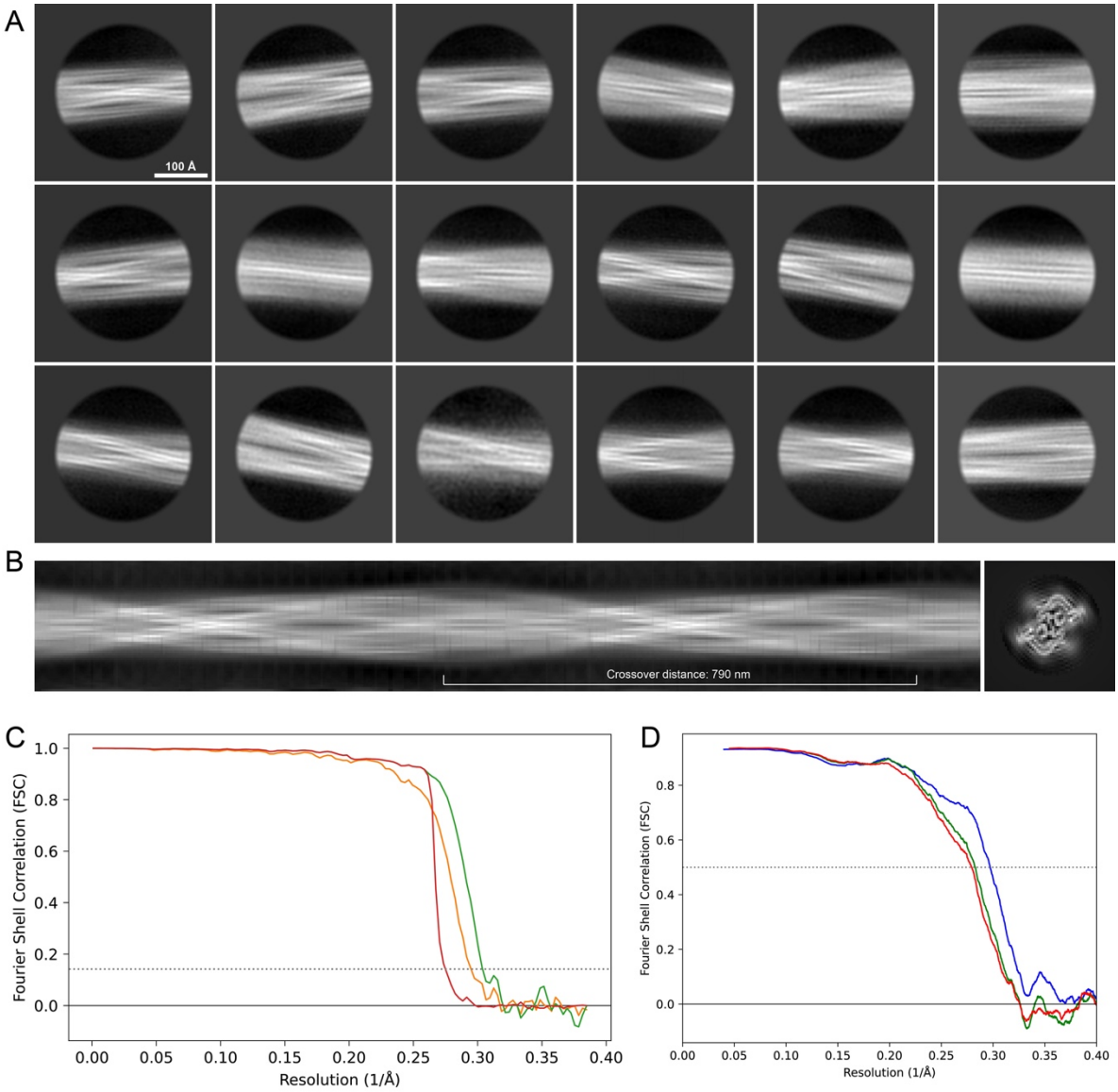

Figure 14. 2D classes and half-map and model-map FSC curves for dataset 8, Type 5. A) The 2D classes of the segments that were used to produce the initial model in relion_helix_inimodel2d. **B**) The output of relion_helix_inimodel2d shown as the summed 2D classes and a Z-projection of the reconstructed 3D model used as input for refinement in the 3D refinement. **C**) The FSC curves produced during postprocessing in RELION with red showing the plot for the phase randomized, orange the unmasked and green the masked maps. **D**) The model-map FSC curves produced in PHENIX. The blue curve is for the deposited coordinates and full postprocessed map against which it had been fit by real space refinement in PHENIX. Coordinates were similarly generated by refining against the first half map and then compared to the same half-map (green) or the second half-map (red)

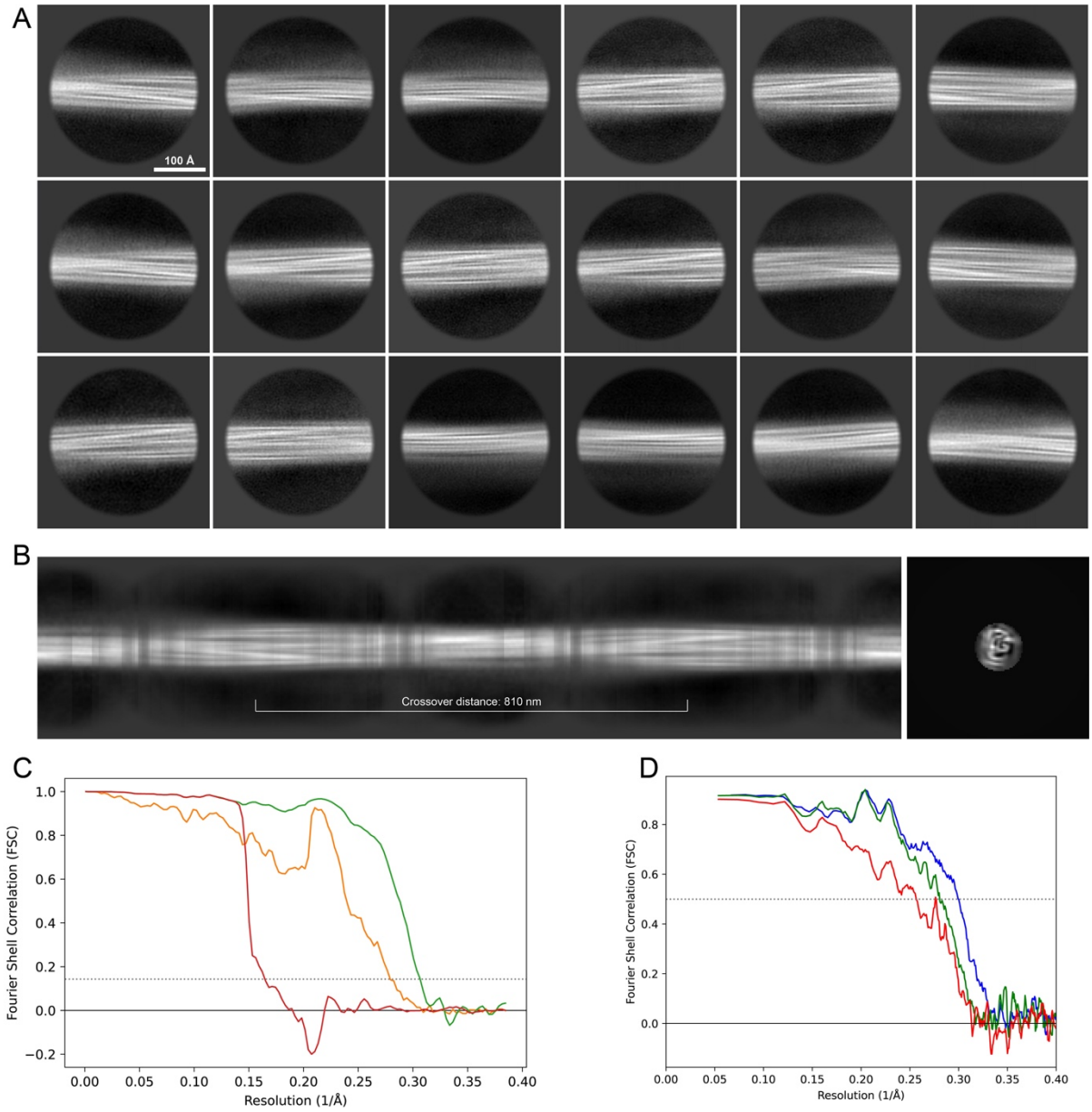

Figure 15. 2D classes and half-map and model-map FSC curves for dataset 9, Type 1_M. A) The 2D classes of the segments that were used to produce the initial model in relion_helix_inimodel2d. **B**) The output of relion_helix_inimodel2d shown as the summed 2D classes and a Z-projection of the reconstructed 3D model used as input for refinement in the 3D refinement. **C**) The FSC curves produced during postprocessing in RELION with red showing the plot for the phase randomized, orange the unmasked and green the masked maps. **D**) The model-map FSC curves produced in PHENIX. The blue curve is for the deposited coordinates and full postprocessed map against which it had been fit by real space refinement in PHENIX. Coordinates were similarly generated by refining against the first half map and then compared to the same half-map (green) or the second half-map (red)

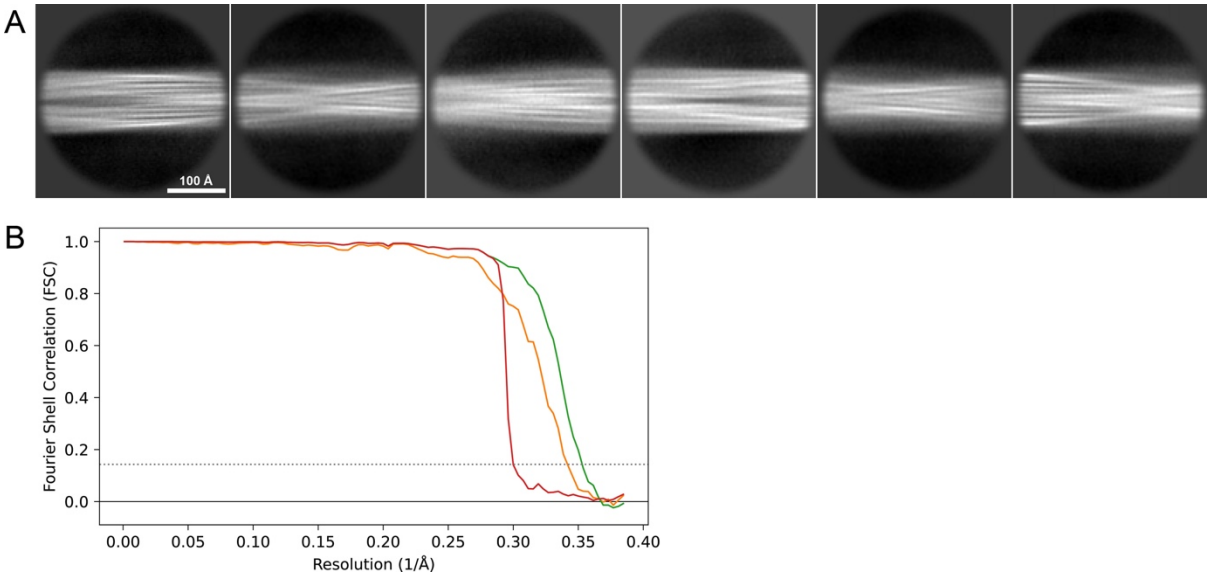

Figure 16. 2D classes and half-map FSC curves for dataset 13, Type 2A. A) Representative 2D classes of the segments that were used for the 3D reconstruction. **B**) The FSC curves produced during postprocessing in RELION with red showing the plot for the phase randomized, orange the unmasked and green the masked maps.

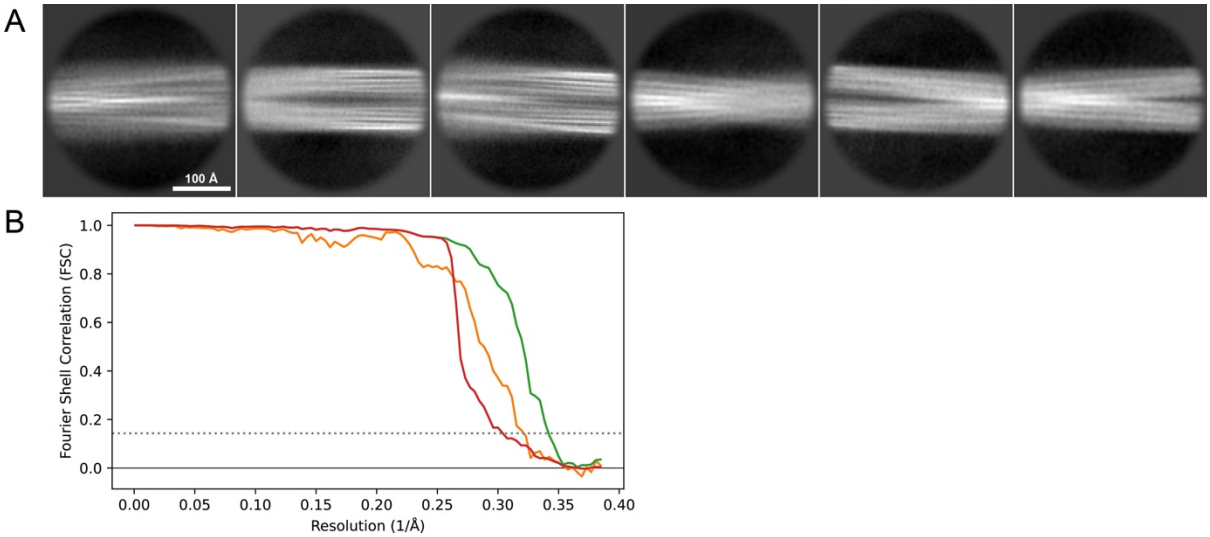

Figure 17. 2D classes and half-map FSC curves for dataset 13, Type 2B. A) Representative 2D classes of the segments that were used for the 3D reconstruction. **B**) The FSC curves produced during postprocessing in RELION with red showing the plot for the phase randomized, orange the unmasked and green the masked maps.