Supplementary Information for

# The Influence of Pathogenic Mutations in $\alpha$ -Synuclein on Biophysical and Structural Characteristics of Amyloid Fibrils

Francesco Simone Ruggeri<sup>1\*</sup>, Patrick Flagmeier<sup>1\*</sup>, Janet R. Kumita<sup>1</sup>, Georg Meisl,<sup>1</sup> Dimitri Y. Chirgadze<sup>2</sup>, Marie N. Bongiovanni<sup>1</sup>, Tuomas P.J. Knowles<sup>1,3#</sup>, Christopher M. Dobson<sup>1</sup>

<sup>1</sup> Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, United Kingdom

<sup>2</sup> Department of Biochemistry, University of Cambridge, Old Addenbrooke's Site, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom

<sup>4</sup> Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, United Kingdom

\*These authors contributed equally

Corresponding Authors: Tuomas P.J. Knowles, Christopher M. Dobson Email: <u>tpjk2@cam.ac.uk</u>

## This PDF file includes:

Supplementary Methods Figures S1 to S8 Tables S1 to S4 SI References

## **Supplementary Methods**

#### α-synuclein protein expression and purification

Recombinant  $\alpha$ -synuclein (WT protein and the mutational variants) was expressed in the Escherichia coli BL21 Gold (DE3) strain (Stratagene) and purified as described previously.<sup>1-3</sup> In short, the purification involved sonicating the cells, boiling the cell debris, and carrying out ion exchange chromatography. This step was followed by size exclusion chromatography in phosphate buffer (20 mM, pH 6.5). The protein concentration was determined from the absorbance at 275 nm using an extinction coefficient of 5600 M<sup>-1</sup>cm<sup>-1</sup>. Protein samples were flash frozen in liquid N<sub>2</sub> and stored at -80°C until use.

#### α-synuclein amyloid-like seed fibril preparation

Seed fibrils were prepared as previously described.<sup>3, 4</sup> Briefly, solutions of the monomeric protein were concentrated to 600 to 800  $\mu$ M and incubated, stirring with a Teflon bar at 37 °C for 48 to 72 h. The solutions of formed amyloid-like fibrils were diluted to a concentration of 200  $\mu$ M, aliquoted, flash frozen in liquid N<sub>2</sub> and stored at -80°C.

#### Kinetic measurement of amyloid fibril formation

Monomeric  $\alpha$ -synuclein (WT protein or its mutational variant) was incubated at a concentration of 100  $\mu$ M in the presence of sonicated fibrils of the respective variant at a concentration of 5  $\mu$ M and 50  $\mu$ M ThT at 37°C under quiescent conditions. The change in the ThT fluorescence signal was monitored using a Polarstar Omega fluorescence plate reader (BMG Labtech, Aylesbury, UK) in bottom reading mode. Corning 96 well plates with half-area (3881, polystyrene, black with clear bottom) non-binding surfaces sealed with metal sealing tape were used for each experiment.

## $\alpha$ -synuclein amyloid-like fibril preparation for the structural characterization

Monomeric  $\alpha$ -synuclein (WT protein or its mutational variant) was incubated at a concentration of 100  $\mu$ M in the presence of sonicated fibrils of the respective variant at a concentration of 5  $\mu$ M at 37°C under quiescent conditions. The kinetics of aggregation using Thioflavin T (ThT) fluorescence. Corning 96 well plates with half-area (3881, polystyrene, black with clear bottom) non-binding surfaces sealed with metal sealing tape were used for each preparation of amyloid-like fibrils. For the structural analysis of the resulting fibrils, however, we incubated the proteins in the absence of ThT to exclude any effects it could have on the structure of the final products.

## X-Ray diffraction pattern recording

The X-ray diffraction patterns were recorded at the Crystallographic X-ray Facility, Department of Biochemistry, University of Cambridge. Borosilicate glass capillaries were shortened to approximately 2 cm and sealed with liquid wax (T ~ 100°C). Two sealed capillaries were positioned in a petri dish in such a way that the ends with the wax faced each other, separated by a small gap (1 to 5 mm) that allows 5 to 10  $\mu$ L drops to be added that stay attached to both capillaries. 5 to 10  $\mu$ L of protein solutions of amyloid-like fibrils of the different variants were applied and incubated at room temperature for the

solution to evaporate. This procedure was repeated until a fibre attached to the end(s) was visible. The formed stalks of amyloid-like fibres of the different variants of α-synuclein were mounted on a Microstar microfocus rotating anode X-ray generator: X8 Proteum (Bruker AXS, Coventry, UK) equipped with a Platinum 135 CCD detector. X-ray diffraction patterns were collected and the data were analysed using Proteum 2 software (Bruker AXS, Coventry, UK).

# Transmission electron microscopy (TEM) imaging

Samples for TEM imaging of the fibrillar species of the WT protein and the mutational variants of αsynuclein (A30P, E46K, H50Q, G51D, A53T) were prepared on 400-mesh, 3-mm copper grid carbon support film (EM resolutions Ltd., Essex, UK) and stained with 2% uranyl acetate (wt/vol). The samples were imaged on a Tecnai G2 200 kV transmission electron microscope (FEI, part of Thermo Fisher Scientific, USA) at the Cambridge Advanced Imaging Centre, University of Cambridge. Images were analyzed using the SIS Megaview II Image Capture system (Olympus Soft Imaging Solutions GmbH, Germany).

# Bulk spectroscopic measurements

*Fluorescence measurements*. The fluorescence measurements were carried out on a Varian (Palo Alto, CA, USA) Model Cary Eclipse spectrofluorimeter equipped with a temperature-controlled cell holder. For our experiments we used a 2 mm x 10 mm path length cuvette.

8-anilo-1-naphtalene-sulfonic acid (ANS) binding was monitored by exciting the sample at 350 nm and recording the emission spectra from 400 to 650 nm. In this case, 5  $\mu$ M of each protein sample was incubated with 250  $\mu$ M ANS in sodium phosphate buffer (20mM, pH 6.5) for 30 min before recording the spectra.

Nile red binding was monitored by exciting the sample at a wavelength of 532 nm and the emission spectrum was recorded from 550 to 750 nm with a slide width of 5 nm.

*Far-UV Circular Dichroism (CD).* Far-UV CD spectra of the WT protein and the mutational variants of  $\alpha$ -synuclein (A30P, E46K, H50Q, G51D, A53T) in their monomeric and fibrillar state were acquired in phosphate buffer (20 mM, pH 6.5) at 30 °C. The protein solutions were diluted to a final concentration of 10  $\mu$ M and spectra were acquired using a 1 mm path length cuvette and a J-810 Jasco spectropolarimeter (Tokyo, Japan), equipped with a thermostat cell holder.<sup>5</sup>

# Fourier-Transform infrared spectroscopy (FTIR)

Attenuated total reflection infrared spectroscopy (ATR-FTIR) was performed using a Bruker Vertex 70 spectrometer equipped with a diamond ATR element. Spectra were acquired with a resolution of 4 cm<sup>-1</sup> and they were processed by means of Origin Pro software. They were first averaged (3 spectra with 128 co-averages) and then the second derivative was calculated applying a Savitzky-Golay filter (2<sup>nd</sup> order, 12 points).

#### Atomic force microscopy (AFM)

Atomic Force Microscopy was performed on positively functionalized mica substrates. We cleaved the mica surface and we incubated it for 1 minute with 10  $\mu$ l of 0.5% (v/v) 3-aminopropyl-triethoxysilane (APTES, from SIGMA) in Milli-Q water. Then, the substrate was rinsed three times with 1 ml of Milli-Q water and dried by gentle stream of nitrogen gas. Finally, for each sample, an aliquot of 10  $\mu$ l of the solution was deposited on the positively functionalized surface. The droplet was incubated for 10 minutes, then rinsed by 1 ml of Milli-Q water and dried by the gentle stream of nitrogen gas. The preparation was carried out at room temperature. AFM maps were realized by means of a NX10 (Park Systems) operating in tapping mode and equipped with a silicon tip (PPP-NCHR, 5 Nm<sup>-1</sup>) with a nominal radius <10 nm. Images flattening was performed by SPIP (Image metrology) software.

#### Cross-sectional packing of the fibrillar aggregates

The projection of the 3-D morphological space on the 2-D plane of *height vs. min-max height difference* shows fundamental morphological differences between the variants (**Fig. Sl6**). The larger variation of min-max height, reaching approximately 50% of maximal height, profile suggests that some of the E46K, H50Q and G51D variants can form fibrils in either a helical or a twisted ribbon arrangement. Whereas the WT, A30P and A53T variants form fibrils where the min-max height profile accounts for approximately 20% of the maximal height, thus suggesting a helical ribbon arrangement with more compact cross-sectional packing. This difference suggests that the fibrils formed by the H50Q and G51D variants possess a different geometrical packing of the protofilament units inside the fibril core when compared to those formed by the WT, A30P and A53T variants.

#### Cluster analysis of AFM morphology data

In Fig. 3 and Fig. SI5, we show the statistical analysis of the height, pitch and difference between maximum and minimum of the height for the amyloid-like fibrils formed from a-synuclein and its diseaseassociated mutational variants. Each single point in the graphs of height vs. pitch and height vs. minmax height represent a single fibril. It is clear that for several of the protein variants the fibrillar aggregates were occupying significantly different regions in these morphology spaces, which indicates the existence of different polymorphs. To demonstrate the existence of the polymorphs and differentiate robustly to which polymorph each fibril belonged, cluster analysis was performed using a rigorous unsupervised machine learning k-means algorithm from python's sklearn package.<sup>6</sup> The algorithm assigns each point (fibril) to one of the clusters (polymorph). To avoid biases in assigning each point to a cluster, the data were normalised along each coordinate, with the minimum value observed for that coordinate set to 0 and the maximum value set to 1. In the graphs of height vs. pitch and height vs. minmax height, this approach enabled us to distinguish the polymorphs formed within each individual mutant by assigning each point in Fig. 3 and Fig. SI5 to a cluster. We find that WT, E46K, H50Q and A53T form 2 periodic fibrillar polymorphs, G51D only forms a single polymorph, while A30P showed a heterogeneous polymorph. Then, we applied the cluster analysis in the 3-D space of height, pitch, minmax height to study the inter-variants subgrouping (Fig. SI5-6 and Fig. 4). The clustering analysis

reveals that WT (95%), A53T (90%) and A30P (80%) belong to a first cluster, while the other variants G51D (100%), E46K (100%) and H50Q (65%) belong to a separate second cluster; where in parenthesis we indicate the percentage of points belonging to the assigned cluster.

## **Supplementary Figures**



Figure S1: Biophysical characterization of amyloid-like fibrils formed from  $\alpha$ -synuclein and its mutational variants (A30P. E46K. H50Q. G51D and A53T). (Top) The amino acid sequence of the 140 residue protein with the residues substituted as a result of mutations in familial cases of Parkinson's disease highlighted (orange: A30, green: E46, blue: H50, red: G51 and purple A53). The structure of  $\alpha$ -synuclein can be divided into three regions: the amphipathic N-terminal domain (1-60), the non-amyloid B-component domain (NAC, 61-94) and the acidic C-terminal domain (95-140). The seven imperfect repeat sequences in α-synuclein are highlighted (light blue bars) and the regions involved in interactions with membrane surfaces with the anchor region (blue, 6-25), the sensor region that modulated the strength of the interaction (grey, 26-97) and the C-terminal fragment that is not bound (green, 98-140). For the figure we follow the colour code used in the paper (black: WT, orange: A30P, green: E46K, blue: H50Q, red: G51D, purple: A53T). (a) Change in ThT fluorescence intensity when monomeric α-synuclein at a concentration of 100 µM was incubated in the presence of 5 µM pre-formed fibrils of the respective variant and 50 µM ThT in phosphate buffer (20 mM, pH 6.5) under quiescent conditions at 37 °C. (b) ANS and (c) Nile Red binding of amyloid-like fibrils formed from α-synuclein and its mutational variants (grey: dye in solution). (d) Far-UV circular dichroism (CD) spectra of WT α-synuclein and its mutational variants. CD spectra of the WT protein and its mutational variants were recorded at a protein concentration of 10 µM in its monomeric (light colour) or fibrillar state (dark colour).



Figure S2: FTIR spectra of amyloid-like fibrils formed from  $\alpha$ -synuclein WT and its mutational variants (A30P, E46K, H50Q, G51D, and A53T). Top row: Recorded spectra of amyloid-like fibrils of  $\alpha$ -synuclein and its disease-associated mutational variants. Middle row: Second derivative of the recorded FTIR spectra. Bottom row: Zoom on the second derivative in the wavelength range of the Amide I band.



Figure S3: The relative secondary structure content and TEM imaging of amyloid-like fibrils formed from  $\alpha$ -synuclein and its mutational variants (A30P, E46K, H50Q, G51D, and A53T). (a) Baselined and normalised IR spectra in the Amide band I region (solid line); and difference spectra obtained by subtracting the spectrum of the mutation from the WT one. The normalised area of the difference spectrum for each mutation is the *structural difference* parameter used in *Fig. 1g* of the main text. (b) Second derivative of the Amide band I was calculated and analysed to evaluate the secondary structure of each variant. (c) Ratio of difference secondary structure contributions were calculated and normalized based on the data in Fig. 1h considering the relative structure content of anti- and parallel  $\beta$ -sheet compared to  $\alpha$ -helix and  $\beta$ -turn. (d) TEM images of the fibrils. (e-g) TEM images of fibrils formed by the E46K variant to demonstrate the heterogeneity of the fibrils (e) rod-like, (f) tightly twisted rope-like and (g) twisted rope-like fibrils. In the case of the G51D variant, a small number of rod-like fibrils are visible in the TEM images. All scale bars represent 100 nm.



Figure S4: Statistical analysis of rod-like polymorph of  $\alpha$ -synuclein variants. The WT (top left), A53T (bottom) and A30P (top-right) mutation showed two rod-like variants (p<0.001).



Figure S5: Morphological features of amyloid-like fibrils formed from  $\alpha$ -synuclein and its disease-associated mutational variants and intra variant clustering. Analysis of the (a) morphological properties of the periodic fibrils and the (b) Fourier Transform of their pitch. (c) Analysis of fibrillar polymorphism. Each point in the graph represent a fibril. For each mutant we analysed the 3-D morphology of at least n > 45 fibrils. Clustering analysis determined that WT, E46K, H50Q and A53T form 2 clusters. Each point was assigned to a cluster by the algorithm and the dashed circle are used as guidelines to visualise the points assigned to each cluster (the points belonging to different cluster also have different colour intensity).



Figure S6: Inter-variant Clustering Analysis on the 3-D Morphological features of amyloid-like fibrils formed from  $\alpha$ -synuclein and its disease-associated mutational variants. We represent each single fibril as a point in the 3-D space of morphology. The unsupervised machine learning *K*-means clustering distinguishes between two major clusters of mutants (indicated by spheres and cubes): 1) WT, A53T, A30P; 2) G51D, E46K and H50Q.



Figure S7: Average height of the fibril polymorphs versus min-max height.



**Figure S8. (a) Three-dimensional diagram of the morphological features of amyloid fibrils formed under identical conditions**. The fibrils formed by the WT protein and the A30P and A53T variants share morphological space with regards to their height, whereas the fibrils formed by the H50Q, E46K and the G51D variants occupy different regions of morphological space. The circle are used to indicate the mutations belonging to each cluster, which is defined by an unsupervised machine learning K-means clustering (**b**, **c**) The relative rates of the lipid induced aggregation and secondary nucleation (from reference<sup>12</sup>). The WT (black), A30P (orange) and A53T (purple) variants share similar fibril families, and secondary nucleation and lipid-induced aggregation were observed. In the case of the G51D variant (red) a single dominant fibril family was found and neither secondary nucleation nor lipid-induced aggregation was observed. The H50Q variant was found to form fibril families of a smaller diameter. Secondary nucleation was not observed at a detectable rate and the lipid-induced aggregation was found to be slow and to form fibrils that differed from those of the WT protein. Secondary nucleation was observed as for the WT, A30P and A53T variants but the lipid-induced aggregation was similar to that of H50Q. (**b**) Relative rate of secondary nucleation plotted *versus* the pitch of the fibrils. (**c**) Plot of the relative lipid-induced aggregation *versus* the pitch of the fibrils.

Table S1: Deconvolution of the CD spectra of the amyloid-like fibrils formed from WT  $\alpha$ -synuclein and its mutational variants. The online platform DichroWeb<sup>5, 7</sup> was used for the deconvolution of the spectra into the secondary structure content with the K2d algorithm<sup>8</sup> and the CDSSTR algorithm<sup>9</sup>.

	K2d			CDSSTR					
	α	β	r	H1	H2	S1	S2	Т	U
WT	0.05	0.47	0.48	0.01	0.05	0.21	0.12	0.27	0.32
A30P	0.05	0.47	0.48	0.00	0.05	0.27	0.14	0.23	0.30
E46K	0.05	0.48	0.48	0.00	0.06	0.26	0.14	0.22	0.30
H50Q	0.05	0.47	0.48	0.00	0.05	0.26	0.13	0.21	0.34
G51D	0.09	0.43	0.48	-	-	-	-	-	-
A53T	0.05	0.47	0.48	0.04	0.07	0.25	0.15	0.21	0.27

Table S2: The relative secondary structure content of amyloid-like fibrils formed from  $\alpha$ synuclein and its mutational variants (A30P, E46K, H50Q, G51D and A53T). FTIR spectra were recorded for the formed amyloid-like fibrils, the Amide I band was analysed by computing its second derivative and the peaks of the Amide I band were integrated. The values shown here were rounded which is why the sum is not always 1.0. The error in the determination of the relative secondary structure content is 5%. The data is plotted in Figure 2f.

	Parallel β-sheet	α-helix	β-turn	Antiparallel β-sheet
WT	0.50	0.11	0.24	0.14
A30P	0.51	0.09	0.25	0.15
E46K	0.53	0.07	0.24	0.16
H50Q	0.50	0.10	0.21	0.19
G51D	0.47	0.16	0.21	0.15
A53T	0.51	0.10	0.24	0.15

Table S3: Ratios of the relative secondary structure content of amyloid-like fibrils formed from  $\alpha$ -synuclein and its mutational variants (A30P, E46K, H50Q, G51D and A53T). The ratios were calculated based on the relative structure content in Supplementary Table 2.

	Parallel β-sheet Antiparallel β-sheet		Parallel and antiparallel β-sheet			
Ratio	a-helix	a-helix	α-helix	β-turn	$\alpha$ -helix and $\beta$ -turn	
WT	4.53 (±0.23)	1.30 (±0.06)	5.83 (±0.30)	2.67 (±0.13)	1.83 (±0.09)	
A30P	5.48 (±0.27)	1.66 (±0.08)	1.66 (±0.08)	2.70 (±0.13)	1.96 (±0.10)	
E46K	7.33 (±0.37)	2.17 (±0-11)	2.17 (±0.11)	2.81 (±0.14)	2.17 (±0.11)	
H50Q	5.17 (±0.26)	1.96 (±0.10)	1.96 (±0.10)	3.24 (±0.16)	2.23 (±0.11)	
G51D	2.91 (±0.15)	0.94 (±0.05)	0.94 (±0.05)	2.96 (±0.15)	1.67 (±0.08)	
A53T	5.20 (±0.26)	1.57 (±0.08)	1.57 (±0.08)	2.83 (±0.14)	1.99 (±0.10)	

Table 4: Morphological features of amyloid-like fibrils of WT  $\alpha$ -synuclein and its diseaseassociated mutational variants. Aliquots were taken from aggregation reactions and deposited on functionalized mica. AFM images were acquired and statistically analysed. Both periodic and nonperiodic fibrils were found. The non-periodic fibrils are highlighted with light-blue.

	Average Height (nm)	Pitch (nm)	Maximum – Minimum (nm)
	7.9 (±0.2)	109 (±18)	2.1 (±0.2)
WT	10.4 (±0.2)	115 (±17)	3.6 (±0.4)
	5.8 (±0.3)		
	9.4 (±0.6)		
	7.8 (±0.2)	98 (±17)	1.9 (±0.2)
A30P	6.5 (±0.3)		
	8.1 (±0.3)		
	6.5 (±0.2)	46 (±15)	2.6 (±0.2)
E46K	7.7 (±0.2)	69 (±15)	1.7 (±0.2)
	7.0 (±0.3)		
	5.4 (±0.2)	59 (±15)	1.5 (±0.3)
H50Q	7.1 (±0.3)	89 (±16)	2.8 (±0.3)
	5.5 (±0.2)		
G51D	6.3 (±0.2)	61 (±15.0)	2.6 (±0.2)
	7.9 (±0.2)	95 (±16)	1.9 (±0.3)
	9.8 (±0.3)	101 (±16)	2.7 (±0.4)
A53T	6.5 (±0.3)		
	8.6 (±0.4)		

# SUPPLEMENTARY REFERENCES

- 1. Hoyer, W.; Antony, T.; Cherny, D.; Heim, G.; Jovin, T. M.; Subramaniam, V., Dependence of α-Synuclein Aggregate Morphology on Solution Conditions. *J Mol Biol* **2002**, *322*, 383-393.
- 2. Galvagnion, C.; Brown, J. W.; Ouberai, M. M.; Flagmeier, P.; Vendruscolo, M.; Buell, A. K.; Sparr, E.; Dobson, C. M., Chemical Properties of Lipids Strongly Affect the Kinetics of the Membrane-Induced Aggregation of α-Synuclein. *Proc Natl Acad Sci U S A* **2016**, *113*, 7065-7070.
- 3. Flagmeier, P.; Meisl, G.; Vendruscolo, M.; Knowles, T. P.; Dobson, C. M.; Buell, A. K.; Galvagnion, C., Mutations Associated with Familial Parkinson's Disease Alter the Initiation and Amplification Steps of α-Synuclein Aggregation. *Proc Natl Acad Sci U S A* **2016**, *113*, 10328-10333.
- Brown, J. W.; Buell, A. K.; Michaels, T. C.; Meisl, G.; Carozza, J.; Flagmeier, P.; Vendruscolo, M.; Knowles, T. P.; Dobson, C. M.; Galvagnion, C., β-Synuclein Suppresses both the Initiation and Amplification Steps of α-Synuclein Aggregation *via* Competitive Binding to Surfaces. *Sci Rep* **2016**, *6*, 36010, doi: 10.1038/srep36010.
- 5. Whitmore, L.; Wallace, B. A., DICHROWEB, an Online Server for Protein Secondary Structure Analyses from Circular Dichroism Spectroscopic Data. *Nucleic Acids Res* **2004**, *32*, W668-673.
- 6. Rousseeuw, P. J., Silhouettes A Graphical Aid to the Interpretation and Validation of Cluster-Analysis. *J Comput Appl Math* **1987**, *20*, 53-65.
- 7. Whitmore, L.; Wallace, B. A., Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers* **2008**, *89*, 392-400.
- Andrade, M. A.; Chacon, P.; Merelo, J. J.; Moran, F., Evaluation of Secondary Structure of Proteins from UV Circular Dichroism Spectra Using an Unsupervised Learning Neural Network. *Protein Eng* 1993, *6*, 383-390.
- 9. Johnson, W. C., Analyzing Protein Circular Dichroism Spectra for Accurate Secondary Structures. *Proteins* **1999**, *35*, 307-312.