1

Structural identification of individual helical amyloid filaments by integration of cryo-electron microscopy-derived maps in comparative morphometric atomic force microscopy image analysis

Liisa Lutter [1], Youssra K. Al-Hilaly [2,3], Christopher J. Serpell [4], Mick F. Tuite [1], Claude M. Wischik [5,6], Louise C. Serpell [2], Wei-Feng Xue [1]*

[1] School of Biosciences, Division of Natural Sciences, University of Kent, CT2 7NJ, Canterbury, UK; [2] Sussex Neuroscience, School of Life Sciences, University of Sussex, BN1 9QG, Falmer, Brighton, UK; [3] Chemistry Department, College of Science, Mustansiriyah University, Baghdad, Iraq; [4] School of Physical Sciences, Division of Natural Sciences, University of Kent, CT2 7NH, Canterbury, UK; [5] Institute of Medical Sciences, University of Aberdeen, UK; [6] TauRx Therapeutics Ltd., Aberdeen, UK

* Correspondence to: W.F.Xue@kent.ac.uk

Key words: atomic force microscopy / amyloid / polymorphism / structural biology / cryo-electron microscopy / image analysis / nano-morphometry / tau / dGAE

23

ABSTRACT

25 The presence of amyloid fibrils is a hallmark of more than 50 human disorders, including

26 neurodegenerative diseases and systemic amyloidoses. A key unresolved challenge in

27 understanding the involvement of amyloid in disease is to explain the relationship between

28 individual structural polymorphs of amyloid fibrils, in potentially mixed populations, and the

29 specific pathologies with which they are associated. Although cryo-electron microscopy (cryo-

30 EM) and solid-state nuclear magnetic resonance (ssNMR) spectroscopy methods have been

31 successfully employed in recent years to determine the structures of amyloid fibrils with high

32 resolution detail, they rely on ensemble averaging of fibril structures in the entire sample or

33 significant subpopulations. Here, we report a method for structural identification of individual

34 fibril structures imaged by atomic force microscopy (AFM) by integration of high-resolution

35 maps of amyloid fibrils determined by cryo-EM in comparative AFM image analysis. This

36 approach was demonstrated using the hitherto structurally unresolved amyloid fibrils formed

37 *in vitro* from a fragment of tau (297-391), termed 'dGAE'. Our approach established

38 unequivocally that dGAE amyloid fibrils bear no structural relationship to heparin-induced tau

39 fibrils formed *in vitro*. Furthermore, our comparative analysis resulted in the prediction that

40 dGAE fibrils are closely related structurally to the paired helical filaments (PHFs) isolated from

41 Alzheimer's disease (AD) brain tissue characterised by cryo-EM. These results show the utility

42 of individual particle structural analysis using AFM, provide a workflow of how cryo-EM data

43 can be incorporated into AFM image analysis and facilitate an integrated structural analysis of

44 amyloid polymorphism.

45

46

INTRODUCTION

Amyloid is a conformational state of polypeptides, which may be accessed by various natively folded or intrinsically disordered proteins (Iadanza et al., 2018). Conversion of protein and peptides into the amyloid state results in an aggregation-prone, β-sheet-rich fold (Eisenberg and Sawaya, 2017), followed by self-assembly into twisted fibrils via a nucleation-dependent mechanism, with numerous intermediate, self-propagating and coexisting species (Chuang et al., 2018). These assemblies, in some cases, are required for the normal physiological functioning of an organism, and in others lead to the formation of cytotoxic species that are a hallmark of neurodegenerative diseases and systemic build-up of misfolded proteins (Dobson, 2017). The connection between pathology and amyloid structures remains of particular interest. All amyloid fibrils share a cross-β core architecture comprised of extended stacks of β-strands that run perpendicular to the fibril axis, but which, importantly, can differ in cross-sectional folds and lateral organisation of protofilaments (Lutter et al., 2021). This diversity has been elucidated recently by the development of high-resolution structural analysis of fibrillar amyloid by cryo-electron microscopy (cryo-EM) and solid-state nuclear magnetic resonance (ssNMR) (Lutter et al., 2021; Willbold et al., 2021), as well as X-ray and electron diffraction of small amyloid peptide crystals (Sawaya et al., 2007, 2016). Proteins with an identical sequence are able to form different core or filament arrangements (Gremer et al., 2017; Wälti et al., 2016). Differences have also been demonstrated for *ex vivo* extracted and *in vitro* assembled fibrils (Bansal et al., 2021), with further differences arising from varying *in vitro* sample preparation conditions and techniques (Colvin et al., 2016; Schmidt et al., 2015). In particular, recent advances in cryo-EM have led to the elucidation of an array of structures of amyloid filaments formed by the tau protein (Shi et al., 2021). Tau is well known for its involvement in Alzheimer's disease, but is also found in chronic traumatic encephalopathy,

72   Pick's disease and numerous other neurodegenerative diseases (Spillantini and Goedert, 2013).

73   Cryo-EM studies have revealed that structures of tau fibrils may be disease-specific, although

74   different diseases can also share the same tau folds (Hallinan et al., 2021; Shi et al., 2021).

75   Importantly, cryo-EM-based structural models of heparin-induced tau filaments revealed

76   significant differences between fibrils extracted from patient tissues and those generated

77   synthetically *in vitro* (Arakhamia et al., 2020; Falcon et al., 2018b; Zhang et al., 2019)

78

79   Although cryo-EM and ssNMR methods can provide detailed structural information of the

80   amyloid fibril core, both require measurements of samples with homogenous populations, or

81   significantly populated subpopulations. ssNMR requires a ^{15}N- and ^{13}C-labeled homogenous

82   sample, often achieved through rounds of consecutive seeding of fibrils with fresh monomer

83   (Paravastu et al., 2009). However, seeding *ex vivo* extracted fibrils with recombinant monomer

84   *in vitro* may not result in structures identical to those of the original *ex vivo* material (Koloteva-

85   Levine et al., 2021; Lövestam et al., 2021). Non-seeded *in vitro* self-assembled fibril samples

86   can also yield well-resolved spectra if the labelled starting material is mostly homogenous

87   (Colvin et al., 2016), although *ex vivo* fibril structures cannot be studied without seeding with

88   labelled monomer to produce fibrils amenable for ssNMR analysis. Cryo-EM can achieve high

89   resolution structural models for significantly populated polymorphs of a heterogenous sample

90   by classification of sub-populations. However, reconstruction requires averaging of tens to

91   even hundreds of thousands of identical fibril segments. Therefore, it may not be possible to

92   assess the polymorphic extent of the sample or to reconstruct polymorphs which are present in

93   low numbers, but which may still be biologically relevant. Additionally, the analysis of

94   intrafibrillar variation of structure and morphology of fibrils is limited due to the low signal-

95   to-noise ratio from individual cryo-EM micrographs.

96

97   Atomic force microscopy (AFM) has recently emerged as a method capable of addressing

98   amyloid structural polymorphism on an individual fibril level (e.g. Adamcik and Mezzenga,

99   2018; Aubrey et al., 2020). AFM is a technique in which the sample surface is probed with a

100   sharp tip attached to a cantilever, creating 3-dimensional topographs at nanometre image

101   resolution with a high signal-to-noise ratio. It enables the surfaces of individual amyloid fibrils

102   to be reconstructed and the polymorphism landscape of the sample to be analysed on a single-

103   molecule basis (Aubrey et al., 2020; Lutter et al., 2020). Therefore, AFM is highly

104   complementary to cryo-EM and ssNMR, which rely on particle or ensemble averaging to

105   produce reconstructions of fibril core structures. The application of AFM to structural biology

106   has been hampered by the imaging artefact that stems from the physical finite size of the AFM

107   probe, resulting in dilation and distortion of surface features. However, we have recently shown

108   that a computational deconvolution approach is able to correct for this artefact, and the

109   recovered surface sampling can result in an increased local image resolution. Furthermore,

110   samples exhibiting helical symmetry, including amyloid fibrils, can be reconstructed as 3D

111   models, facilitating structural analysis of the fibril surface envelopes (Lutter et al., 2020). AFM

112   image deconvolution and 3D reconstruction of helical filaments can then be employed to carry

113   out a quantitative mapping of amyloid fibril polymorphism at an individual fibril level (Aubrey

114   et al., 2020). This deconvolution and reconstruction approach opens up the possibility of

115   integrating high signal-to-noise individual filament 3D envelope structures from AFM data

116   with high resolution ensemble-averaged core structural information available from cryo-EM

117   and ssNMR methodologies. To date, AFM imaging of amyloid fibrils has been carried out

118   alongside cryo-EM and ssNMR studies to determine fibril handedness or visualise the sample

119   supra-structural morphology (Ghosh et al., 2021; Li et al., 2018; Röder et al., 2019). However,

120   a fully integrated approach that links AFM data with cryo-EM derived structural maps has the

121   potential to enhance the structural analysis of amyloid fibril polymorphism as AFM provides

122   structural data on an individual fibril level, which is highly complementary to the averaged

123   structural models from cryo-EM and ssNMR data. Nevertheless, there is currently no unifying

124   workflow which would enable comparison of 3D-information from AFM surface envelope

125   reconstructions with the structural maps provided by cryo-EM and ssNMR.

126

127   Here, we report a new approach to quantitative structural comparison of individual amyloid

128   fibrils imaged by AFM with ensemble-averaged structural data derived from cryo-EM. This

129   approach integrates AFM 3D envelope models of individual fibrils with cryo-EM maps and

130   has been applied to the structural characterisation and identification of individual filaments

131   formed by tau (297-391) known as 'dGAE' (Al-Hilaly et al., 2020), which has not yet been

132   structurally resolved to high resolution by cryo-EM or by ssNMR. dGAE tau self-assembles

133   spontaneously *in vitro* without the presence of anionic cofactors, such as heparin, to form

134   filaments that share the height and cross-over characteristics observed in paired helical

135   filaments (PHFs) isolated from Alzheimer's disease tissue (Al-Hilaly et al., 2020). We

136   compared the 3D surface envelope models of synthetic, *in vitro* assembled dGAE tau filaments

137   constructed from previously published AFM image data with existing cryo-EM density maps

138   of both *ex vivo* extracted tau fibrils and those formed *in vitro* by the addition of heparin, found

139   in the Electron Microscopy Data Bank (EMDB). This analysis showed that the surface of

140   dGAE amyloid fibrils closely resembles that of the cryo-EM density maps arising from some

141   of the tauopathy tissue-extracted tau filaments, notably those of PHFs isolated from *ex vivo*

142   Alzheimer's disease tissue. Thus, our cryo-EM data-integrated AFM image analysis ruled out

143   structural similarity of dGAE amyloid fibrils assembled *in vitro* to *in vitro* formed heparin-

144   induced tau filaments, and has provided evidence that identified dGAE fibrils are structurally

145   closely related to PHFs isolated from AD brain tissue. Individual fibril level structures in mixed

146   polymorphic amyloid populations are likely to underpin properties such as overall

147   fragmentation propensity, assembly kinetics and cytotoxic potential. Thus, by developing an

148   integrative AFM image analysis, we reveal the potential for this form of AFM analysis to be

149   used for identification of individual fibril structures in heterogeneous and polymorphic amyloid

150   populations.

151

152

153   **RESULTS**

154

155   **Three-dimensional reconstruction using AFM image data enabled structural analysis of**

156   **individual dGAE tau amyloid filaments**

157

158   AFM imaging of *in vitro* assembled dGAE filaments (Al-Hilaly et al., 2020) revealed a mixture

159   of long individual fibrils and laterally associated, twisted bundles (**Fig 1a**). All fibrils and

160   bundles showed a left-handed twist, with variation in the cross-over distance (*cod*, defined

161   according to Supplementary **Fig S1a**). dGAE fibrils demonstrated a high propensity to bundle

162   (**Fig 1a-b**). This may be indicative of surface properties of the dGAE protofilaments or fibrils,

163   which favour the formation of large, twisted bundles of various widths. To analyse the structure

164   of individual dGAE fibrils, 3D envelopes of individually selected single, non-bundled dGAE

165   fibrils can be reconstructed from the AFM topographs using previously developed methods

166   (Lutter et al., 2020). Here, to demonstrate this approach, a 3D surface envelope model of a

167   single selected well-separated canonical dGAE fibril was reconstructed (**Fig 1c-d**). The central

168   height profile of this dGAE filament showed peaks with the average height of 8.9 nm and

169   grooves with the average height of 5.9 nm, with cross-over distances ranging from 43.1-58.6

170   nm, with an average of 47.7 nm (**Fig 1d**). These morphometric parameters are in line with the

171   ranges expected of a typical non-bundled dGAE fibril observed by AFM (Al-Hilaly et al.,

172 2020). During the 3D reconstruction workflow (Lutter et al., 2020), the fibril was determined

173 to have 2-fold helical symmetry, suggesting that it is composed of 2-protofilaments arranged

174 symmetrically about the central helical axis. The reconstructed 3D envelope model also

175 enabled the evaluation of the cross-sectional area and shape, perpendicular to the fibril axis

176 (defined in Supplementary **Fig S1a**), which was subsequently used in comparison with cryo-

177 EM density maps of a range of different tau fibril structures.

178

179 **Simulation of topographic images from volumetric cryo-EM maps allowed integration of**

180 **cryo-EM data in comparative AFM image analysis**

181

182 AFM imaging is able to generate image data in the form of 3D topographic raster scans, which

183 are typically displayed as coloured 2D images where the pixel value represents surface height.

184 As the images have a high signal-to-noise ratio, topographs of individual filaments can be

185 traced and extracted from the surface scan data (**Fig 1**). In order to compare the 3D structural

186 information from cryo-EM and AFM, simulated topographic AFM height images were

187 generated from cryo-EM density maps (examples in **Fig 2**). This enabled the comparative

188 analysis of the dGAE fibril imaged by AFM (**Fig 1c-d**) and cryo-EM data of patient derived

189 tau fibrils and *in vitro* self-assembled heparin-induced tau fibrils.

190

191 Volumetric cryo-EM maps of tau amyloid fibrils were first downloaded from the Electron

192 Microscopy Data Bank, followed by extraction of the map iso-surface 3D coordinates at the

193 iso-values specified by the authors of each cryo-EM entry. Where necessary, the individual

194 maps were then denoised and the maps of the short fibril segments were extended by finding

195 the helical axis of the map, followed by rotation and translation of multiple copies of the maps

196 about their helical axes, using author reported twist and rise values, to a total map length

8

197   equivalent to that of the AFM imaged dGAE fibril. Although helical twist and rise values are

198   provided by the authors in experimental metadata, the position and angle of the fibril map

199   screw-axis is unknown for all cryo-EM maps. Therefore, an iterative approach was used to find

200   the rotation and translation parameters that provide the best alignment of the cryo-EM fibril

201   maps to their respective helical axis positions (**Fig 2a, c**). In some cases, for example heparin-

202   induced recombinant tau 'snake' fibrils, the resulting screw-axis position lies off the fibril

203   centre-of-mass, resulting in a hollow centre (**Fig 2c**), in line with previous analyses (Zhang et

204   al., 2019). In order to simulate topographic AFM images of the extended iso-surfaces, a model

205   of the AFM cantilever probe tip, which interacts with the filament surface, is then needed. The

206   tip angle and geometry were estimated using the values provided by the manufacturer of the

207   AFM probe used experimentally. The tip radius used in the simulations was matched to that

208   used to image the dGAE fibril, which was estimated using AFM image data from the fibril (**Fig

209   1c-d**) according to a previous published algorithm (Lutter et al., 2020). A simulated AFM tip

210   constructed from these geometry, angle, and radius parameters was then used to find the tip-

211   sample contact points with the extended fibril iso-surface, to account for, and to reproduce, the

212   tip-sample convolution effect in the simulated images. The sampling rate was also set to be

213   equivalent to that of the experimental AFM data (**Fig 2b, d**). As both the tip radius and sampled

214   pixel density of the simulated images match that of the individual fibril from the AFM image

215   data, it was possible to make a direct comparison between the two.

216

217   **Simulated topographic AFM images from cryo-EM maps of helical fibrils enabled**

218   **validation of the AFM based 3D reconstruction methodology**

219

220   Prior to a full comparative analysis of the AFM derived 3D model of the dGAE fibril with

221   simulated AFM images from published cryo-EM structural data of tau fibrils, the simulated

9

222   AFM images were first used to validate the AFM-based 3D reconstruction methodology used

223   to construct the dGAE fibril model (**Fig 1d**). In order to evaluate the 3D reconstruction method

224   that we developed to reconstruct fibril 3D envelopes from AFM data and to determine their

225   cross-sectional shapes and helical symmetries, reconstructions from simulated images of fibrils

226   with known 3D structures were carried out, using cryo-EM maps of selected tau fibril structures

227   (Supplementary **Fig S2** and **S3**). In this validation process, AFM topographs were simulated

228   as described above, using the same pixel size and AFM tip radius as seen in the experimental

229   AFM image of the dGAE tau fibril. Reconstruction of 3D envelopes was then carried out from

230   these simulated images in the same way as from experimental AFM images, with the aim of

231   comparing the final 3D reconstructions to the known fibril structures from which the input

232   images were simulated. This allowed the validation of the reconstruction method regarding the

233   accuracy of the cross-sectional shape, symmetry assignment, and positioning of the screw-axis,

234   which the comparative image and morphometric analyses of dGAE tau fibrils requires. During

235   the reconstruction workflow (Lutter et al., 2020), fibril symmetry is initially unknown so

236   envelopes of fibrils with various symmetries, as well as an asymmetrically cross-sectioned

237   fibril, were reconstructed from the simulated image data. The correct symmetry can then be

238   identified by the angle of the twisting pattern on the fibril surface by comparison of topographs

239   re-simulated from various symmetry models with the input topograph of the experimental or

240   simulated image, in direct, as well as in Fourier space (Supplementary **Fig S2f** and **S3f**).

241

242   The cryo-EM density iso-surfaces for PHF (EMD-0259) and heparin-jagged (EMD-4565)

243   fibrils were used to validate the 3D reconstruction method by comparison of the final 3D model

244   with the known input structures. PHF and heparin-jagged maps were chosen to evaluate the

245   reconstruction method, as the structures have C2 and C1 rotational symmetry, respectively,

246   allowing the reconstruction of different symmetry assignments to be assessed. In addition,

10

247 heparin-jagged has a helical axis that lies off the centre-of-mass of the core density. Evaluation

248 of the approach demonstrated that the 3D reconstruction method can accurately and

249 unambiguously identify both the C2 and C1 helical symmetries from fibril topographs

250 (Supplementary **Fig S1f** and **S2f**). Furthermore, for PHF, the shape of the fibril cross-section

251 closely matched that of the AFM tip-accessible cross-section from the original cryo-EM

252 density map. The root-mean-square deviation (RMSD) between the tip-accessible PHF density

253 map cross-section, which was obtained by tracing a circle with the radius equivalent to that of

254 the AFM tip radius around the iso-surface cross-section to account for the topographic

255 resolution imposed by the finite size of the AFM probe (Supplementary **Fig S1d**), and 3D

256 reconstructed envelope (Supplementary **Fig S2e**) was 0.08 nm. For heparin-jagged fibril, with

257 an off-centre screw-axis, cross-sectional RMSD was 0.23 nm (Supplementary **Fig S3e**),

258 although the helical axis was correctly positioned to be lying off the fibril centre-of-mass (cross

259 in Supplementary **Fig S3e**). These results validated the AFM based 3D reconstruction and

260 simulation approach, although the RMSD values showed potential for improvement for the

261 reconstruction workflow for asymmetric fibrils with off centre helical axes. More importantly,

262 these results demonstrated that the approach correctly identified the symmetry and screw-axis

263 position using known structural maps as input. This gave confidence that the 2-fold symmetry

264 identified for the dGAE fibril from the experimental image data is accurate, indicating that it

265 is composed of 2-protofilaments arranged symmetrically about a central helical axis, like the

266 PHF from AD brain. Furthermore, as the RMSD between the known input and the 3D

267 reconstruction was sub-Ångström for the 2-fold symmetric PHF example, the shape of dGAE

268 fibril is also not likely to be significantly misrepresented by the 3D reconstruction process from

269 the AFM data, although limitations can still potentially arise from the sample structure's local

270 mechanical stability and imaging conditions.

271

Comparison of cryo-EM derived topographs with dGAE tau AFM data ruled out structural similarity with heparin-induced tau fibrils

Simulated topographic images from helical axis-aligned and lengthened cryo-EM density maps of *ex vivo* tissue-extracted and *in vitro* assembled heparin-induced tau filaments were subsequently prepared for comparison with the topographic AFM image of the canonical dGAE tau filament (**Fig 3**). This was achieved by carrying out topographic image simulations for all unique tau fibril cryo-EM maps available in the EMDB up to August 2021, which included structural maps of fibrils purified from tissues of patients with corticobasal degeneration (CBD), chronic traumatic encelopathy (CTE), Pick's disease (PiD), Alzheimer's disease (AD), as well as four fibril polymorphs formed *in vitro* by addition of heparin to recombinant tau monomers. In addition to the selected *ex vivo* fibril data, entries in the EMDB of near-identical structures to the selected PHFs and straight filaments (SFs) were also determined for filament structures from various other neurodegenerative diseases, including prion protein cerebral amyloid angiopathy (PrP-CAA) and primary age-related tauopathy (PART) (Hallinan et al., 2021; Shi et al., 2021). Where several near-identical structures have been reported for the same disease, the one with the highest map resolution was used. The helical twist of amyloid fibrils within a population of fibrils is variable even when the core structures are identical (Zhang et al., 2020), but the helical twist parameters presented in the population averaged cryo-EM maps represent only one of a continuum of possible values within a range, while AFM images report the full structural variation within an individual filament. To account for this difference in the information content, the periodicities of the simulated fibril images were optimised to maximise the normalised cross-correlation of the simulated and experimental fibril images by allowing the long fibril axis of the simulated image to contract or dilate within a range (see Methods). This allowed for slight adjustments of the

12

297   simulated fibril periodicity, which aids the assessment of image similarity between

298   experimental AFM data and simulated images derived from cryo-EM maps.

299

300   The digitally straightened dGAE fibril (**Fig 1c-d**) is shown in **Fig 3** together with all of the

301   simulated images of tau fibrils from published cryo-EM maps. All simulated tau fibrils are left-

302   handed, as is the dGAE filament. As with any topographic AFM height image, the intensity of

303   the colour also represents the height value in the simulated topographs. The comparison of

304   fibril heights (colour intensity in **Fig 3)** immediately highlights the differences between the

305   dGAE fibril and CBD wide filaments, as well as wide Pick filaments (WPF), both consisting

306   of two intertwined protofilaments with larger cross-sectional areas compared to the dGAE

307   fibril. However, CBD narrow filaments, narrow Pick filaments (NPF), and straight filaments

308   (SF) from AD, which appear to be similar to dGAE in height, exhibit a difference in the

309   repeating pattern compared to dGAE, due to their asymmetric cross-sections. These types of

310   asymmetric cross-sections lead to a complex pattern of peaks and grooves along the length of

311   the fibril, compared to the single visible peak and groove of the dGAE with 2-fold screw-

312   symmetry, which in comparison is more similar to PHFs from AD. A similar pattern, although

313   with a repeating unit consisting of a major and a minor peak, resulting from tip-sample contact

314   with the rhombic cross-sections, are also seen on CTE type I and CTE type II filaments. Finally,

315   the morphologies of all four heparin-induced tau fibrils differ from those of the other simulated

316   images, and from the dGAE fibril image, due to their off-centre helical axis and asymmetric

317   cross-sections. Although the dGAE fibril shown here matches the heparin-twister fibril well in

318   height, the displayed repeating pattern differs substantially from the experimental dGAE fibril

319   image due to differences in the cross-sectional shape and position of the screw-axis.

320

321 To quantify the similarity between the experimental dGAE fibril image and the simulated

322 topographical AFM images derived from cryo-EM maps, a correlation distance based score

323 (d_{img}, see Methods) was calculated between the dGAE fibril image and each of the simulated

324 images (Supplementary **Table S1**). Because the pixel sampling rate in the simulations was set

325 to be identical to that of the experimental AFM image, each pixel value (height value) in the

326 data image can be compared to the pixel value in a simulated image at the same coordinate

327 location. The correlation distance score, d_{img}, is then one minus the correlation coefficient

328 between the data pixel values and the simulated pixel values, with smaller distance scores

329 signifying more similar images. The d_{img} score was calculated using the pixel value at the

330 coordinates in the experimental AFM data image where the height values are above the

331 estimated height of the fibril axis to minimise complex contributions from tip-sample

332 convolution. The resulting d_{img} scores (Supplementary **Table S1**) confirmed that the

333 experimental dGAE fibril image showed closest similarity to the simulated AFM images

334 derived from cryo-EM data of patient-tissue purified *ex vivo* PHF and CTE type II filaments.

335

336 **Quantitative and comparative morphometric analysis of dGAE tau amyloid fibrils with**

337 **cryo-EM derived structural maps confirms their structural similarity to patient-derived**

338 **tau structures**

339

340 To investigate further whether dGAE tau amyloid fibrils are structurally similar to the tau

341 amyloid purified from patient tissues, a quantitative and comparative morphometric analysis

342 was performed next. First, the rotational cross-section of the dGAE tau fibril from the 3D model

343 constructed using AFM data (**Fig 1c-d**), and those of the cryo-EM tau fibril density maps, were

344 quantified in terms of their area and shape differences (cross-sectional area, *csa*, and cross-

345 sectional difference, *csd*, respectively, defined according to Supplementary **Fig S1c-d)**. For

346   cryo-EM density maps, the rotational cross-sections were calculated by de-rotating the cross-

347   sections along their helical axes, followed by tracing of the boundaries of the resulting 2D-

348   maps with a circle of a radius equivalent to that of the AFM tip of the experimental dGAE fibril

349   image. This procedure made the cross-sections from cryo-EM and AFM data comparable by

350   accounting for the finite size of the tip during AFM imaging and thus estimating 'tip-

351   accessible' cross-sections from the cryo-EM maps. A visual comparison of the dGAE fibril

352   cross-section with CTE type II, PHF, and heparin-twister filament cross-sections (**Fig 4**)

353   showed that although CTE II and PHF fibrils both display two-fold helical symmetry and show

354   similarity in their simulated topographs to the dGAE fibril image data (**Fig 3**), the dimensions

355   of the dGAE fibril model cross-section are smaller. However, although the area of the heparin-

356   twister fibril cross-section better matched that of the dGAE fibril, it has an asymmetric cross-

357   sectional shape with the helical axis positioned off the centre-of-mass, like other heparin-

358   induced fibrils, which are dissimilar to the dGAE fibril with two-fold helical symmetry.

359   Therefore, the differences in the helical symmetry and average cross-over distance were also

360   taken into account in the comparative analysis, in addition to the cross-sectional difference and

361   area parameters. The differences for all of these morphometric parameters between the dGAE

362   tau fibril and patient tissue-extracted tau filament structures, or *in vitro* heparin-induced

363   filaments, as determined by cryo-EM, were compiled and listed in Supplementary **Table S1**.

364   These morphometric distance-based similarity scores, together with the image similarity score,

365   d_{img}, allowed the calculation of a combined score of similarity, d_{Σ} (see Methods), between the

366   dGAE fibril imaged by AFM and each of the unique previously determined cryo-EM tau fibril

367   structural maps. This combined score enabled ranking of overall structural similarity

368   (Supplementary **Table S1**), showing that the canonical dGAE fibril (**Fig 1c-d**) exhibits the

369   highest overall structural similarity to PHF from Alzheimer's disease, followed closely by CTE

370 type II filaments, while unequivocally ruling out any close structural similarity to *in vitro*

371 heparin-induced tau fibrils, as well as other disease-associated tau fibrils (**Fig 5**).

372

373

374 **DISCUSSION**

375

376 Structural analysis of amyloid fibrils at the level of individual filaments is key to the

377 understanding of the structural polymorphism displayed by amyloid assemblies, which in turn

378 may relate to the varying physiological effects of amyloid *in vivo*. AFM imaging is

379 complementary to high-resolution, ensemble-averaging techniques like cryo-EM, due to its

380 capability for the structural analyses of single, individual molecules, e.g. membrane proteins

381 (Heath et al., 2021), DNA (Pyne et al., 2021) and amyloid filaments (Aubrey et al., 2020), as

382 well as their chemical signatures, response to force, interactions and dynamics (e.g. Maity and

383 Lyubchenko, 2019; Ruggeri et al., 2015; Watanabe-Nakayama et al., 2016) Here using dGAE

384 tau amyloid fibrils, we show that the 3D information from AFM and cryo-EM can be integrated

385 with each other by simulation of topographic images from cryo-EM density maps or by 3D

386 reconstruction of fibril surface envelopes from AFM topographic data. As exemplified by the

387 analysis of the dGAE tau amyloid structure, these workflows enable direct comparison between

388 experimental AFM images and the increasing numbers of cryo-EM data deposited in the

389 EMDB. We demonstrate this approach by comparing high-resolution structural maps of

390 various tau amyloid fibril cores determined by cryo-EM with a single AFM topographic scan

391 of a sample of dGAE tau fibrils formed *in vitro* without the addition of anionic cofactors. In

392 this case, the quantitative comparison of simulated topographs derived from the cryo-EM

393 density maps with an experimental AFM fibril image of a canonical dGAE tau fibril shows that

394 the structure of dGAE fibrils differs unambiguously from the structures of *in vitro* assembled

395 heparin-induced recombinant tau fibrils. The analysis also predicts a close structural match

396 between dGAE tau fibrils and both PHFs and CTE type II fibrils extracted from patient brain

397 tissues.

398

399 The biological implications of our findings are significant because it is necessary to establish

400 conditions of tau self-assembly *in vitro* that could result in the formation of disease-relevant

401 fibril structures, in order to enable studies of these structures along a disease-relevant

402 aggregation pathway, as well as to facilitate accessibility to physiologically relevant fibril

403 structures for the development of aggregation modulators. Like dGAE fibrils, PHFs and CTE

404 II fibrils also have 2-fold helical symmetry and have a similar cross-sectional shape. The cross-

405 section of dGAE fibrils as estimated by the AFM derived 3D model is, however, smaller than

406 those derived from CTE II fibrils and PHFs (**Fig 4**). This may be due to the physical interactions

407 of the AFM probe tip with the fibril surface despite the low force applied, the deposition of the

408 fibrils for AFM scanning in air or underlying molecular differences. However, the similarity

409 between dGAE fibrils and PHFs as well as CTE type II filaments is also supported by sequence

410 comparison. The dGAE tau sequence contains residues that correspond to the repeat regions 3

411 (R3) and 4 (R4) of the 4-repeat tau isoform, as well as 9 residues from the C-terminus of the

412 repeat region 2 (R2), and thus it is expected the morphology of the fibrils would more closely

413 resemble that of the fibrils which also contain R3 and R4 residues in their cores but few or

414 none R1 or R2 residues. Such fibrils include PHFs and SFs AD filaments, as well as the two

415 CTE filaments (Falcon et al., 2019, 2018b). In contrast, filaments from CBD, PiD, and heparin-

416 induced recombinant tau also contain a larger numbers of residues from the R2 or R1 regions

417 within their core structures (Falcon et al., 2018a; Zhang et al., 2020, 2019).

418

419   Although the sampling pixel density and tip-sample convolution effect are controlled for and

420   matched to experimental values during the simulation of height images using cryo-EM maps,

421   other factors arising from experimental imaging conditions such as the sample deposition

422   process or residual monomeric or small aggregated species sticking onto the imaged fibril or

423   the probe tip will affect the quality of experimental AFM data. Further improvements to the

424   resolution and imaging parameters of the experimental dataset, either during data acquisition

425   or by computational post-acquisition algorithms, could improve the accuracy of the

426   quantitative comparison and the ranking of the similarity between experimental and simulated

427   fibril images. The method presented here can also be extended to include ssNMR derived

428   structural models by using atomic coordinates deposited to the Protein Data Bank instead of

429   cryo-EM density maps for comparison. Thus, the approach presented here demonstrates the

430   vast potential of integrated structural analysis that utilises the complementary 3D information

431   acquired by different physical principles.

432

433

434   **METHODS**

435

436   **AFM image data**

437

438   The AFM image data analysed here was collected on a Bruker MultiMode 8 scanning probe

439   microscope with a Nanoscope III controller, and is part of a previously published data set (Al-

440   Hilaly et al., 2020). The images were acquired using the ScanAsyst in Air mode using Bruker

441   Scan Asyst Air probes with nominal spring constant of 0.04 N/m. The AFM image shown in

442   **Fig 1** represents topographical height scans 3 μm × 3 μm in size at a resolution of 2048 × 2048

443   pixels. The height image was flattened using Bruker Nanoscope Analysis software to remove

444   tilt and bow. All further image analysis was carried out in Matlab.

445

446   **AFM image analysis**

447

448   Fibrils were traced (Xue, 2014; Xue et al., 2009) and digitally straightened (Egelman, 1986)

449   using an in-house application. The radius of the AFM tip during scanning was estimated by a

450   least-squares fit of tip-sample contact-point model on the mean cross-section of a straightened

451   fibril across its length, with tip radius as fitted parameter (Lutter et al., 2020). A tip model was

452   then constructed from the tip radius, and the tip angle and geometry provided by the

453   manufacturer. The same tip model was later used for simulating AFM topographs from cryo-

454   EM data. The average fibril cross-over distance was found as the highest frequency peak from

455   the Fourier transform of the fibril centre line height profile. The 3D reconstruction of the traced

456   and straightened dGAE fibril was carried out as previously described (Lutter et al., 2020).

457

458   **Cryo-EM volumetric data pre-processing**

459

460   Cryo-EM density maps were downloaded from the Electron Microscopy Data Bank (EMDB).

461   The voxel-based volumetric Coulomb density map data were imported into Matlab where all

462   subsequent processing and analysis was carried out. The helical rise, twist and handedness

463   values were obtained from the information provided in the EMDB metadata of each entry. The

464   iso-surface of a density map was found by applying the isovalue provided by the authors of

465   each of the EMDB data entries. The iso-surface vertices were rotated and translated, where

466   necessary, so that the long fibril axis runs along the z-axis. Where necessary, the density maps

467   were denoised by finding the nearest neighbours for all iso-surface vertices and subsequently

19

468 applying a threshold value to remove vertices with a significantly larger mean distance from

469 neighbours. The number of nearest neighbours and threshold were changed according to the

470 noise level of each map by manual inspection. All maps were then further denoised by finding

471 the alpha-shape of the iso-surface using an alpha radius of twice the voxel size of the EM

472 density map, while suppressing holes and retaining only the vertices that belong to the largest

473 single region of the isosurface.

474

475 **Identifying and aligning the fibril screw-axis**

476

477 The position of the fibril screw-axis was estimated by iterative optimisation of the rotation and

478 translation parameters while maximising the overlap between a voxelised isosurface and the

479 same isosurface rotated and translated about the Z-axis by applying the helical rise and twist

480 values from the EMDB entry of each density map. Depending on the map features, rotation

481 and translation parameters of the map isosurface coordinates relative to the screw-axis were

482 optimised jointly or either rotation or translation was initially optimised separately for an initial

483 estimate, followed by joint optimisation of all parameters.

484

485 **Simulation of AFM topographs from cryo-EM structural data**

486

487 The axis-aligned iso-surface coordinates were extended to double the length of an experimental

488 AFM image fibril to allow for periodic alignment. This was done by rotating and translating

489 the iso-surface vertices along the long fibril axis (the z-axis in axis-aligned iso-surface) using

490 the helical twist and rise values from the EMDB metadata, followed by a 90° rotation so that

491 the filament axis aligned with the y-axis. The tip model with the radius specific to the

492 experimentally imaged fibril was then used to find the contact points between the tip model

493 and the fibril surface on an XY-grid constructed with the pixel size equivalent to that of the

494 experimental AFM image. The periodicity pattern of the simulated fibril images was

495 subsequently optimised to take into account possible structural variations in the twist of

496 filaments by finding a possible best match to that of the experimental AFM fibril image within

497 a plausible bounded range. This is achieved by an iterative optimisation of a parameter which

498 dilates or contracts the simulated fibril image along the length of the fibril within a bounded

499 interval of 0.5-2 times the nominal pixel size and maximises the cross-correlation between the

500 transformed simulated image and the straightened experimental fibril image.

501

502 **Calculation of cross-sectional area and shape differences**

503

504 Filament cross-sections perpendicular to the fibril axis were found for both the AFM 3D

505 envelope model of the dGAE tau amyloid fibril and tau amyloid fibril cryo-EM density maps

506 by untwisting the coordinates to the same 2D plane along the length of the fibril, using the twist

507 and rise values provided in the EMDB metadata for cryo-EM models, or the analysis of

508 periodicity for AFM imaged fibrils. For EM density maps, the tip accessible cross-section was

509 estimated as the alpha shape of the cross-section, which was then found with the alpha radius

510 equivalent to the tip radius of the fibril on the experimental AFM image. The alpha shape was

511 then rotated by 1° intervals for a total of 360° and for each interval, RMSD to the cross-section

512 of the experimental AFM fibril envelope was calculated. The rotation angle that resulted in the

513 lowest RMSD indicated the best rotational alignment of the two cross-sections. The cross-

514 sectional area or difference area were then calculated as defined in **Fig S1 c-d**.

515

516 **Similarity scores and comparative structural ranking**

517

21

518   A score to quantify the similarity between the digitally straightened experimental AFM image

519   of the dGAE fibril and simulated AFM images using cryo-EM maps of different polymorph of

520   tau filaments found in the EMDB was defined as the correlation distance, d_{img}, between the

521   pixel height values, which is one minus the correlation coefficient, r, between the data pixel

522   values and the simulated pixel values (**Eq. 1**)

523    $$d_{img} = 1 - r$$ Eq. 1

524   The d_{img} score was calculated using only the pixel value at the coordinates in the experimental

525   AFM data image where the height values are above the estimated height of the fibril axis to

526   minimise complex contributions from tip-sample convolution. Scores to quantify the

527   differences in the morphometric parameters (defined according to Supplementary **Fig S1**)

528   helical symmetry, sym, twist handedness, hnd, average cross-over distance, cod, cross-

529   sectional area, csa and cross-sectional difference area, csd, between each of the cryo-EM

530   derived tau filament maps and the experimentally AFM-imaged dGAE fibril are defined as

531   their individual Euclidian distances d (Aubrey et al., 2020), normalised to their observed

532   maximum values, or the theoretical maximum value of 2 for hnd.

533    $$d_{par} = \frac{|par_{EM\ map} - par_{dGAE}|}{|par_{EM\ map} - par_{dGAE}|_{max}}$$

534   Eq. 2

535   In **Eq. 2**, 'par' represents the parameters sym, hnd, cod, csa or csd. Defining of all the similarity

536   scores using standardised distance measures means that lower numerical values indicate higher

537   degree of similarity. The overall score of similarity, d_{Σ}, was then defined as the sum of the

538   individual image and morphometric similarity scores (**Eq. 3**).

539    $$d_{\Sigma} = d_{img} + d_{sym} + d_{hnd} + d_{cod} + d_{csa} + d_{csd}$$ Eq. 3

540

541

ACKNOWLEDGMENTS

REFERENCES

Adamcik, J., Mezzenga, R., 2018. Amyloid polymorphism in the protein folding and aggregation energy landscape. Angew. Chem. Int. Ed. 57, 8370–8382. https://doi.org/10.1002/anie.201713416

Al-Hilaly, Y.K., Foster, B.E., Biasetti, L., Lutter, L., Pollack, S.J., Rickard, J.E., Storey, J.M.D., Harrington, C.R., Xue, W.-F., Wischik, C.M., Serpell, L.C., 2020. Tau (297-391) forms filaments that structurally mimic the core of paired helical filaments in Alzheimer's disease brain. FEBS Lett. 594, 944–950. https://doi.org/10.1002/1873-3468.13675

Arakhamia, T., Lee, C.E., Carlomagno, Y., Duong, D.M., Kundinger, S.R., Wang, K., Williams, D., DeTure, M., Dickson, D.W., Cook, C.N., Seyfried, N.T., Petrucelli, L., Fitzpatrick, A.W.P., 2020. Posttranslational modifications mediate the structural

566   diversity of tauopathy strains. Cell 180, 633-644.e12.

567   https://doi.org/10.1016/j.cell.2020.01.027

568 Aubrey, L.D., Blakeman, B.J.F., Lutter, L., Serpell, C.J., Tuite, M.F., Serpell, L.C., Xue, W.-

569   F., 2020. Quantification of amyloid fibril polymorphism by nano-morphometry

570   reveals the individuality of filament assembly. Commun. Chem. 3, 1–10.

571   https://doi.org/10.1038/s42004-020-00372-3

572 Bansal, A., Schmidt, M., Rennegarbe, M., Haupt, C., Liberta, F., Stecher, S., Puscalau-Girtu,

573   I., Biedermann, A., Fändrich, M., 2021. AA amyloid fibrils from diseased tissue are

574   structurally different from *in vitro* formed SAA fibrils. Nat. Commun. 12, 1013.

575   https://doi.org/10.1038/s41467-021-21129-z

576 Chuang, E., Hori, A.M., Hesketh, C.D., Shorter, J., 2018. Amyloid assembly and

577   disassembly. J. Cell Sci. 131. https://doi.org/10.1242/jcs.189928

578 Colvin, M.T., Silvers, R., Ni, Q.Z., Can, T.V., Sergeyev, I., Rosay, M., Donovan, K.J.,

579   Michael, B., Wall, J., Linse, S., Griffin, R.G., 2016. Atomic resolution structure of

580   monomorphic Aβ_{42} amyloid fibrils. J. Am. Chem. Soc. 138, 9663–9674.

581   https://doi.org/10.1021/jacs.6b05129

582 Dobson, C.M., 2017. The amyloid phenomenon and its links with human disease. Cold

583   Spring Harb. Perspect. Biol. 9. https://doi.org/10.1101/cshperspect.a023648

584 Egelman, E.H., 1986. An algorithm for straightening images of curved filamentous

585   structures. Ultramicroscopy 19, 367–373. https://doi.org/10.1016/0304-

586   3991(86)90096-3

587 Eisenberg, D.S., Sawaya, M.R., 2017. Structural studies of amyloid proteins at the molecular

588   level. Annu. Rev. Biochem. 86, 69–95. https://doi.org/10.1146/annurev-biochem-

589   061516-045104

590   Falcon, B., Zhang, W., Murzin, A.G., Murshudov, G., Garringer, H.J., Vidal, R., Crowther,

591   R.A., Ghetti, B., Scheres, S.H.W., Goedert, M., 2018a. Structures of filaments from

592   Pick's disease reveal a novel tau protein fold. Nature 1.

593   https://doi.org/10.1038/s41586-018-0454-y

594   Falcon, B., Zhang, W., Schweighauser, M., Murzin, A.G., Vidal, R., Garringer, H.J., Ghetti,

595   B., Scheres, S.H.W., Goedert, M., 2018b. Tau filaments from multiple cases of

596   sporadic and inherited Alzheimer's disease adopt a common fold. Acta Neuropathol.

597   (Berl.) 136, 699–708. https://doi.org/10.1007/s00401-018-1914-z

598   Falcon, B., Zivanov, J., Zhang, W., Murzin, A.G., Garringer, H.J., Vidal, R., Crowther, R.A.,

599   Newell, K.L., Ghetti, B., Goedert, M., Scheres, S.H.W., 2019. Novel tau filament fold

600   in chronic traumatic encephalopathy encloses hydrophobic molecules. Nature 568,

601   420–423. https://doi.org/10.1038/s41586-019-1026-5

602   Ghosh, U., Thurber, K.R., Yau, W.-M., Tycko, R., 2021. Molecular structure of a prevalent

603   amyloid-β fibril polymorph from Alzheimer's disease brain tissue. Proc. Natl. Acad.

604   Sci. 118, e2023089118. https://doi.org/10.1073/pnas.2023089118

605   Gremer, L., Schölzel, D., Schenk, C., Reinartz, E., Labahn, J., Ravelli, R.B.G., Tusche, M.,

606   Lopez-Iglesias, C., Hoyer, W., Heise, H., Willbold, D., Schröder, G.F., 2017. Fibril

607   structure of amyloid-β(1–42) by cryo–electron microscopy 5.

608   Hallinan, G.I., Hoq, M.R., Ghosh, M., Vago, F.S., Fernandez, A., Garringer, H.J., Vidal, R.,

609   Jiang, W., Ghetti, B., 2021. Structure of tau filaments in prion protein amyloidoses.

610   Acta Neuropathol. (Berl.). https://doi.org/10.1007/s00401-021-02336-w

611   Heath, G.R., Kots, E., Robertson, J.L., Lansky, S., Khelashvili, G., Weinstein, H., Scheuring,

612   S., 2021. Localization atomic force microscopy. Nature 594, 385–390.

613   https://doi.org/10.1038/s41586-021-03551-x

614   Iadanza, M.G., Jackson, M.P., Hewitt, E.W., Ranson, N.A., Radford, S.E., 2018. A new era

615   for understanding amyloid structures and disease. Nat. Rev. Mol. Cell Biol. 19, 755–

616   773. https://doi.org/10.1038/s41580-018-0060-8

617   Koloteva-Levine, N., Aubrey, L.D., Marchante, R., Purton, T.J., Hiscock, J.R., Tuite, M.F.,

618   Xue, W.-F., 2021. Amyloid particles facilitate surface-catalyzed cross-seeding by

619   acting as promiscuous nanoparticles. Proc. Natl. Acad. Sci. 118.

620   https://doi.org/10.1073/pnas.2104148118

621   Li, Y., Zhao, C., Luo, F., Liu, Z., Gui, X., Luo, Z., Zhang, X., Li, D., Liu, C., Li, X., 2018.

622   Amyloid fibril structure of α-synuclein determined by cryo-electron microscopy. Cell

623   Res. 1. https://doi.org/10.1038/s41422-018-0075-x

624   Lövestam, S., Schweighauser, M., Matsubara, T., Murayama, S., Tomita, T., Ando, T.,

625   Hasegawa, K., Yoshida, M., Tarutani, A., Hasegawa, M., Goedert, M., Scheres,

626   S.H.W., 2021. Seeded assembly *in vitro* does not replicate the structures of α-

627   synuclein filaments from multiple system atrophy. FEBS Open Bio 11, 999–1013.

628   https://doi.org/10.1002/2211-5463.13110

629   Lutter, L., Aubrey, L.D., Xue, W.-F., 2021. On the structural diversity and individuality of

630   polymorphic amyloid protein assemblies. J. Mol. Biol. 167124.

631   https://doi.org/10.1016/j.jmb.2021.167124

632   Lutter, L., Serpell, C.J., Tuite, M.F., Serpell, L.C., Xue, W.-F., 2020. Three-dimensional

633   reconstruction of individual helical nano-filament structures from atomic force

634   microscopy topographs. Biomol. Concepts 11, 102–115. https://doi.org/10.1515/bmc-

635   2020-0009

636   Maity, S., Lyubchenko, Y.L., 2019. Force clamp approach for fharacterization of nano-

637   assembly in amyloid beta 42 dimer. Nanoscale 11, 12259–12265.

638   https://doi.org/10.1039/c9nr01670h

639　Paravastu, A.K., Qahwash, I., Leapman, R.D., Meredith, S.C., Tycko, R., 2009. Seeded

640　　　growth of β-amyloid fibrils from Alzheimer's brain-derived fibrils produces a distinct

641　　　fibril structure. Proc. Natl. Acad. Sci. 106, 7443–7448.

642　　　https://doi.org/10.1073/pnas.0812033106

643　Pyne, A.L.B., Noy, A., Main, K.H.S., Velasco-Berrelleza, V., Piperakis, M.M., Mitchenall,

644　　　L.A., Cugliandolo, F.M., Beton, J.G., Stevenson, C.E.M., Hoogenboom, B.W., Bates,

645　　　A.D., Maxwell, A., Harris, S.A., 2021. Base-pair resolution analysis of the effect of

646　　　supercoiling on DNA flexibility and major groove recognition by triplex-forming

647　　　oligonucleotides. Nat. Commun. 12, 1053. https://doi.org/10.1038/s41467-021-

648　　　21243-y

649　Röder, C., Vettore, N., Mangels, L.N., Gremer, L., Ravelli, R.B.G., Willbold, D., Hoyer, W.,

650　　　Buell, A.K., Schröder, G.F., 2019. Atomic structure of PI3-kinase SH3 amyloid fibrils

651　　　by cryo-electron microscopy. Nat. Commun. 10, 3754.

652　　　https://doi.org/10.1038/s41467-019-11320-8

653　Ruggeri, F.S., Longo, G., Faggiano, S., Lipiec, E., Pastore, A., Dietler, G., 2015. Infrared

654　　　nanospectroscopy characterization of oligomeric and fibrillar aggregates during

655　　　amyloid formation. Nat. Commun. 6, 7831. https://doi.org/10.1038/ncomms8831

656　Sawaya, M.R., Rodriguez, J., Cascio, D., Collazo, M.J., Shi, D., Reyes, F.E., Hattne, J.,

657　　　Gonen, T., Eisenberg, D.S., 2016. Ab initio structure determination from prion

658　　　nanocrystals at atomic resolution by MicroED. Proc. Natl. Acad. Sci. 113, 11232–

659　　　11236.

660　Sawaya, M.R., Sambashivan, S., Nelson, R., Ivanova, M.I., Sievers, S.A., Apostol, M.I.,

661　　　Thompson, M.J., Balbirnie, M., Wiltzius, J.J.W., McFarlane, H.T., Madsen, A.Ø.,

662　　　Riekel, C., Eisenberg, D., 2007. Atomic structures of amyloid cross-β spines reveal

663　　　varied steric zippers. Nature 447, 453–457. https://doi.org/10.1038/nature05695

664   Schmidt, M., Rohou, A., Lasker, K., Yadav, J.K., Schiene-Fischer, C., Fändrich, M.,

665   Grigorieff, N., 2015. Peptide dimer structure in an Aβ(1–42) fibril visualized with

666   cryo-EM. Proc. Natl. Acad. Sci. 112, 11858–11863.

667   https://doi.org/10.1073/pnas.1503455112

668   Shi, Y., Zhang, W., Yang, Y., Murzin, A., Falcon, B., Kotecha, A., Beers, M. van, Tarutani,

669   A., Kametani, F., Garringer, H.J., Vidal, R., Hallinan, G.I., Lashley, T., Saito, Y.,

670   Murayama, S., Yoshida, M., Tanaka, H., Kakita, A., Ikeuchi, T., Robinson, A.C.,

671   Mann, D.M.A., Kovacs, G.G., Revesz, T., Ghetti, B., Hasegawa, M., Goedert, M.,

672   Scheres, S.H.W., 2021. Structure-based classification of tauopathies. bioRxiv

673   2021.05.28.446130. https://doi.org/10.1101/2021.05.28.446130

674   Spillantini, M.G., Goedert, M., 2013. Tau pathology and neurodegeneration. Lancet Neurol.

675   12, 609–622. https://doi.org/10.1016/S1474-4422(13)70090-5

676   Wälti, M.A., Ravotti, F., Arai, H., Glabe, C.G., Wall, J.S., Böckmann, A., Güntert, P., Meier,

677   B.H., Riek, R., 2016. Atomic-resolution structure of a disease-relevant Aβ(1–42)

678   amyloid fibril. Proc. Natl. Acad. Sci. 113, E4976–E4984.

679   https://doi.org/10.1073/pnas.1600749113

680   Watanabe-Nakayama, T., Ono, K., Itami, M., Takahashi, R., Teplow, D.B., Yamada, M.,

681   2016. High-speed atomic force microscopy reveals structural dynamics of amyloid β_{1-}

682        $_{42}$ aggregates. Proc. Natl. Acad. Sci. 113, 5835–5840.

683   https://doi.org/10.1073/pnas.1524807113

684   Willbold, D., Strodel, B., Schröder, G.F., Hoyer, W., Heise, H., 2021. Amyloid-type protein

685   aggregation and prion-like properties of amyloids. Chem. Rev.

686   https://doi.org/10.1021/acs.chemrev.1c00196

687   Xue, W.-F., 2014. Chapter 2 - Amyloid fibril length quantification by atomic force

688   microscopy, in: Uversky, V.N., Lyubchenko, Y.L. (Eds.), Bio-Nanoimaging.

689   Academic Press, Boston, pp. 17–25. https://doi.org/10.1016/B978-0-12-394431-

690   3.00002-X

691   Xue, W.-F., Homans, S.W., Radford, S.E., 2009. Amyloid fibril length distribution quantified

692   by atomic force microscopy single-particle image analysis. Protein Eng. Des. Sel. 22,

693   489–496. https://doi.org/10.1093/protein/gzp026

694   Zhang, W., Falcon, B., Murzin, A.G., Fan, J., Crowther, R.A., Goedert, M., Scheres, S.H.,

695   2019. Heparin-induced tau filaments are polymorphic and differ from those in

696   Alzheimer's and Pick's diseases. eLife 8, e43584. https://doi.org/10.7554/eLife.43584

697   Zhang, W., Tarutani, A., Newell, K.L., Murzin, A.G., Matsubara, T., Falcon, B., Vidal, R.,

698   Garringer, H.J., Shi, Y., Ikeuchi, T., Murayama, S., Ghetti, B., Hasegawa, M.,

699   Goedert, M., Scheres, S.H.W., 2020. Novel tau filament fold in corticobasal

700   degeneration. Nature 580, 283–287. https://doi.org/10.1038/s41586-020-2043-0

701

702 **FIGURES**

703

704

705
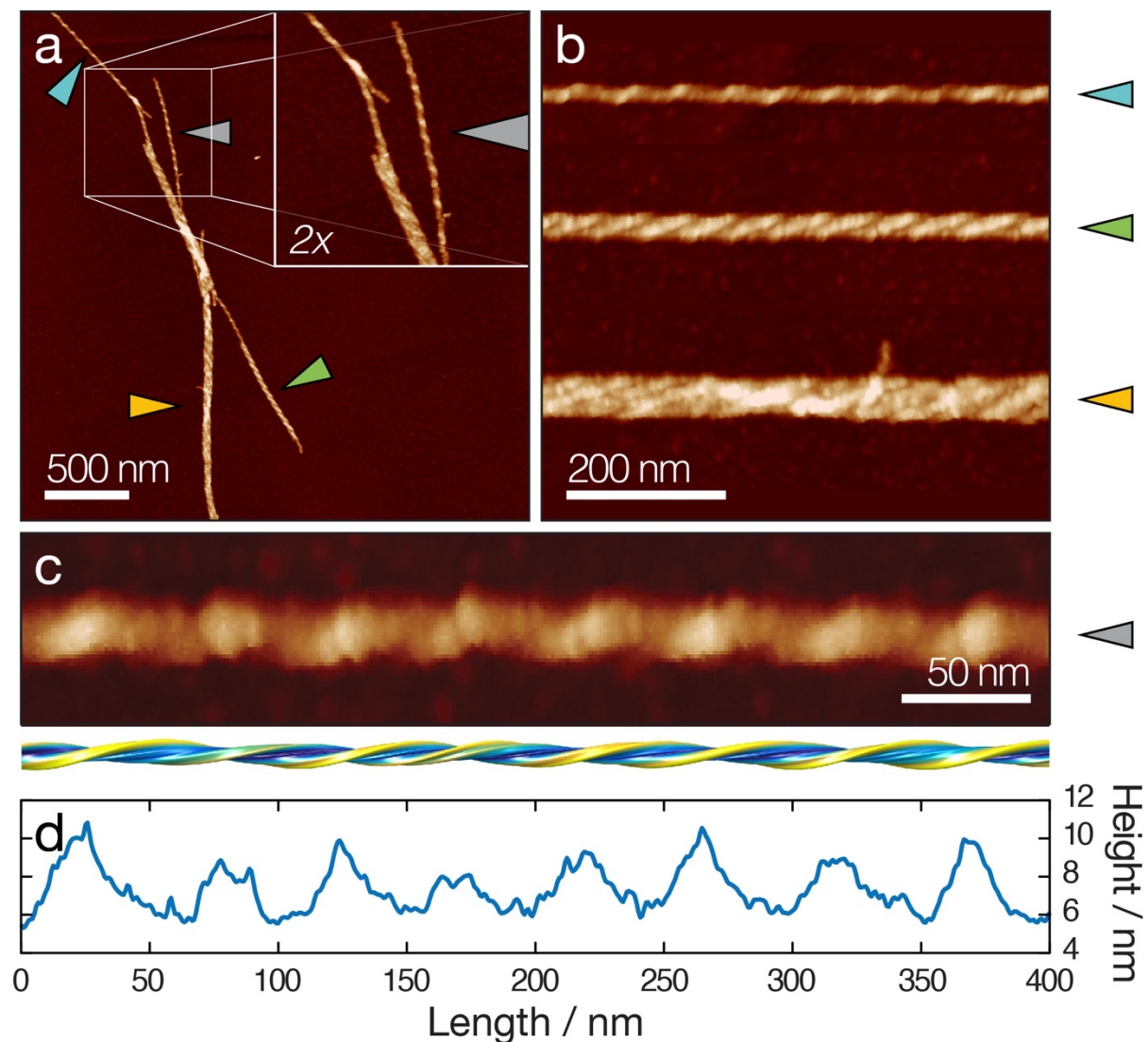
706 **Figure 1.** Analysis of individual dGAE filaments from a topographical AFM height image. (a)

707 A typical AFM image of dGAE fibrils (Al-Hilaly et al., 2020) with an inset showing a 2x

708 magnified view. Arrows point to the fibrils shown in (b) and (c). Scale bar represents 500 nm.

709 (b) Digitally straightened images showing a typical dGAE fibril and typical fibril bundles.

710 Scale bar represents 200 nm. (c) Digitally straightened AFM image of a canonical dGAE tau

711 fibril together with a 3D reconstructed surface envelope model. Scale bar represents 50 nm. (d)

712 The central line height profile of the fibril image shown in (c).

713

714

715
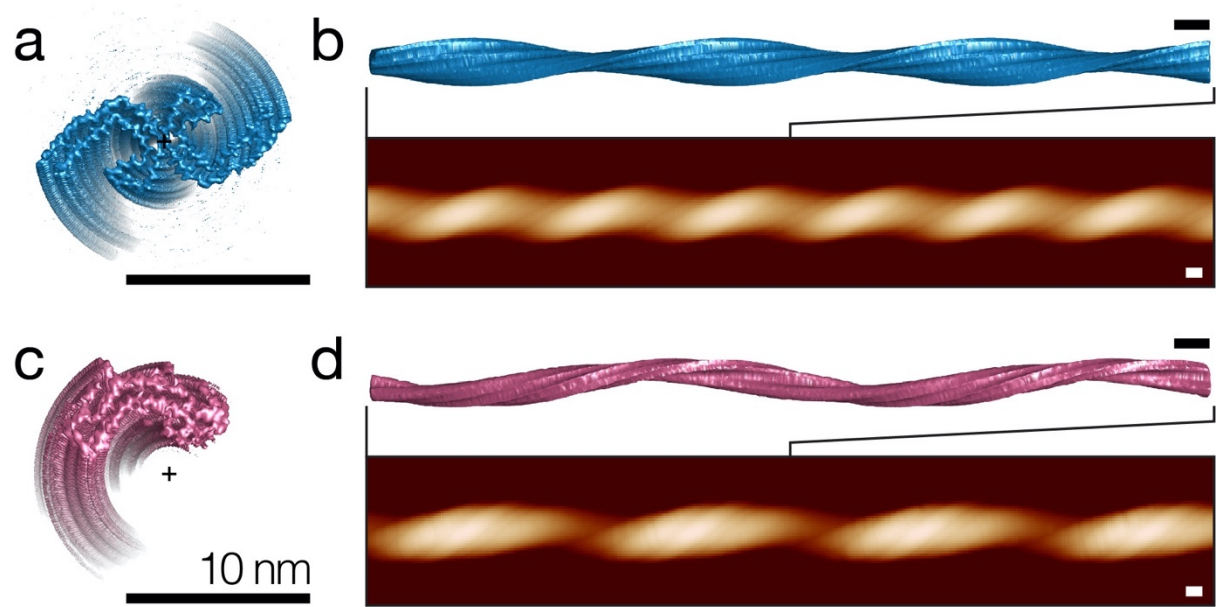
716 **Figure 2.** Simulation of topographic AFM height images from cryo-EM density maps. Cross-

717 sectional views of axis-aligned cryo-EM density map iso-surfaces for (a) paired helical

718 filaments (EMD-0259) and (c) 'heparin-snake' (EMD-4563) tau fibrils are shown with the

719 position of the helical axis (+). Simulated topographic images of 500 nm fibril segments and

720 the top view of extended 250 nm segments of iso-surfaces aligned to the simulated images are

721 shown for PHF in (b) and 'heparin-snake' in (d). All scale bars represent 10 nm.

722

723

724
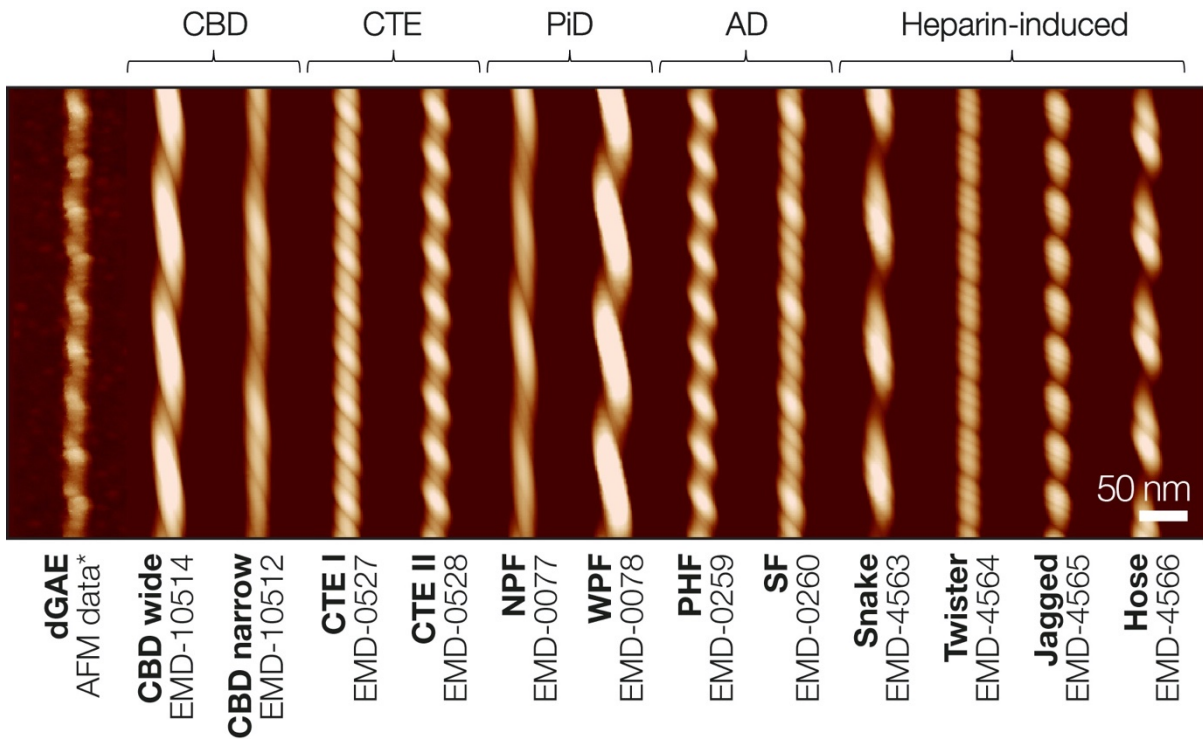
Figure 3. Comparison of the typical experimental AFM height image of dGAE tau fibril with topographical images of tau amyloid fibril polymorphs. The tau fibril polymorphs were simulated with identical tip (here estimated to be 11.7 nm) and sampling parameters using extended and axis-aligned cryo-EM map iso-surfaces. The '*' indicate image data shown in **Fig 1c**. Length of the scale bar is 50 nm. The simulated images have been aligned to the same estimated helical axis height position as the experimental data image and, therefore, have comparable height colour coding. The EMDB accession codes of the cryo-EM density maps from which the topographs were simulated from are listed with the abbreviated names of the individual filaments. The sample origin is noted above the filaments' images. The following abbreviations were used for disease diagnoses and filament labels: CBD – corticobasal degeneration, CTE – chronic traumatic encephalopathy, PiD – Pick's disease, AD – Alzheimer's disease, NPF – narrow Pick filament, WPF – wide Pick filament, PHF – paired helical filament, SF – straight filament.

738
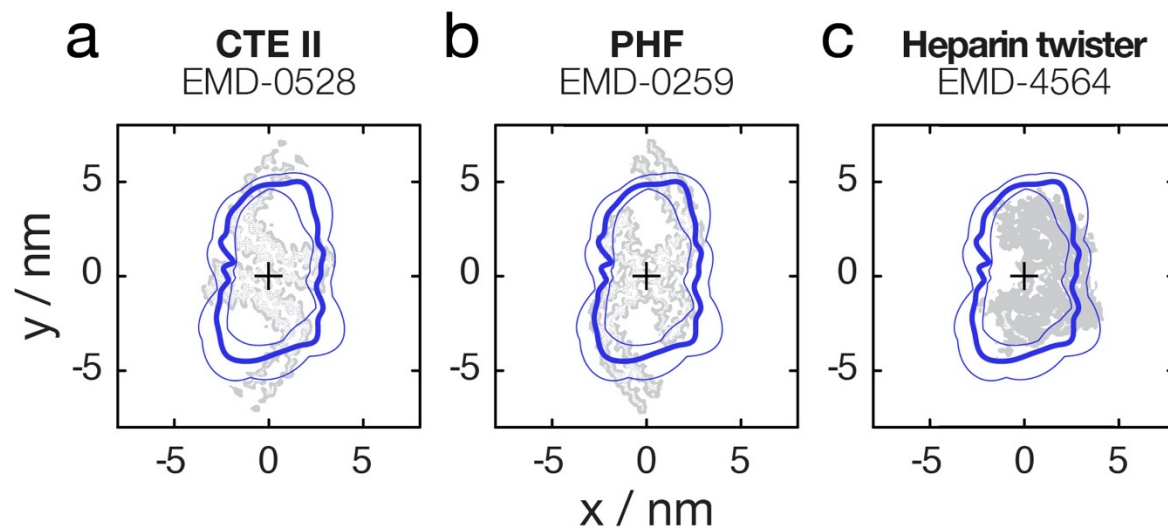
Figure 4. Comparison of the cross-section of the surface envelope model derived from the experimental AFM dGAE filament image and cryo-EM density maps cross-sections. The cross-sections for the CTE type II filaments (a) and PHF from AD patient tissue (b), as well as *in vitro* assembled heparin-induced 'twister' filaments (c) are shown. The average cross-section of the AFM image derived 3D model is shown as thick blue lines, and the structural variation observed along the fibril in the AFM image is indicated by the thin blue line denoting the 2.5 and 97.5 percentile bounds. The cryo-EM density cross-sections are shown in grey.
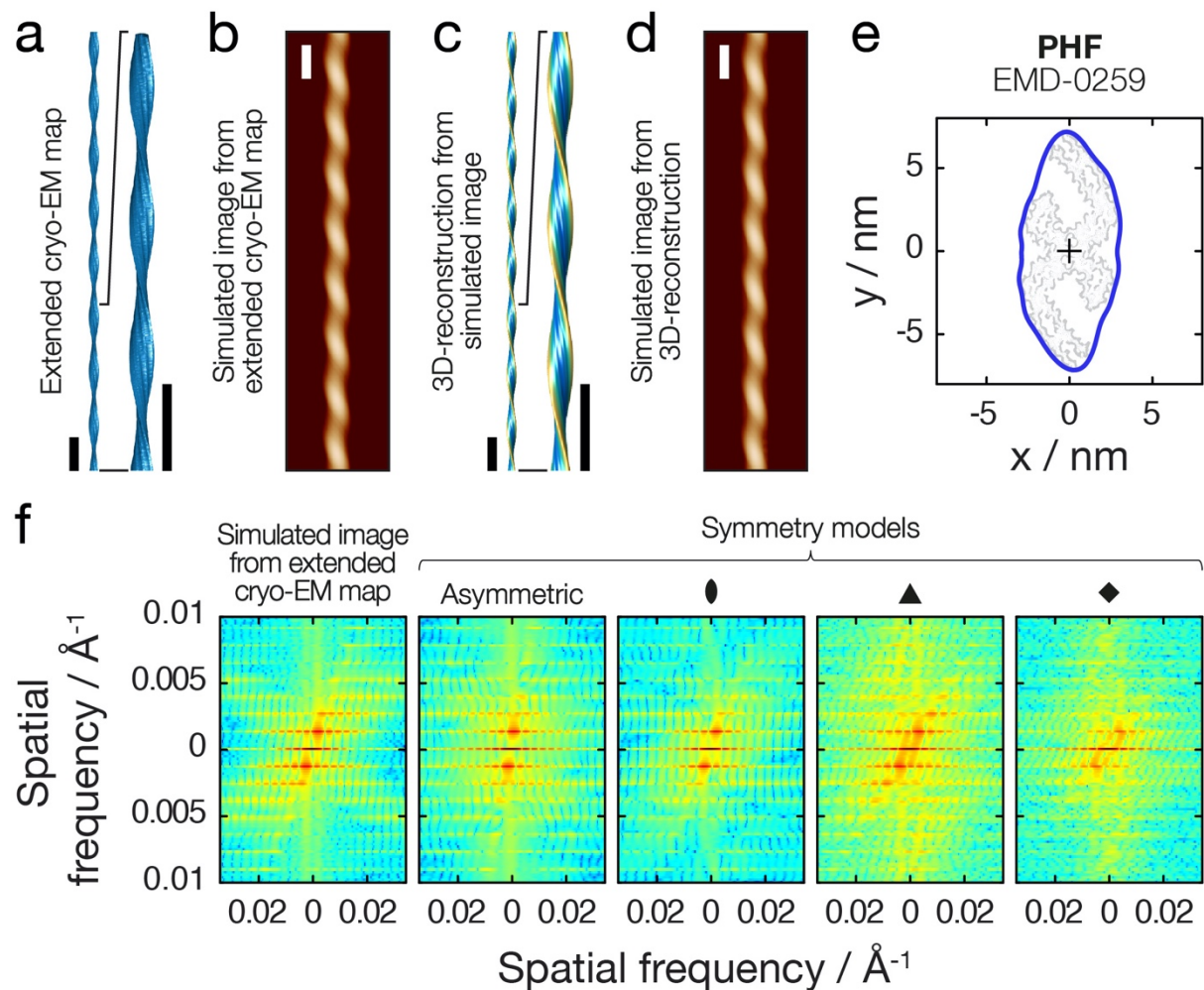
749

750
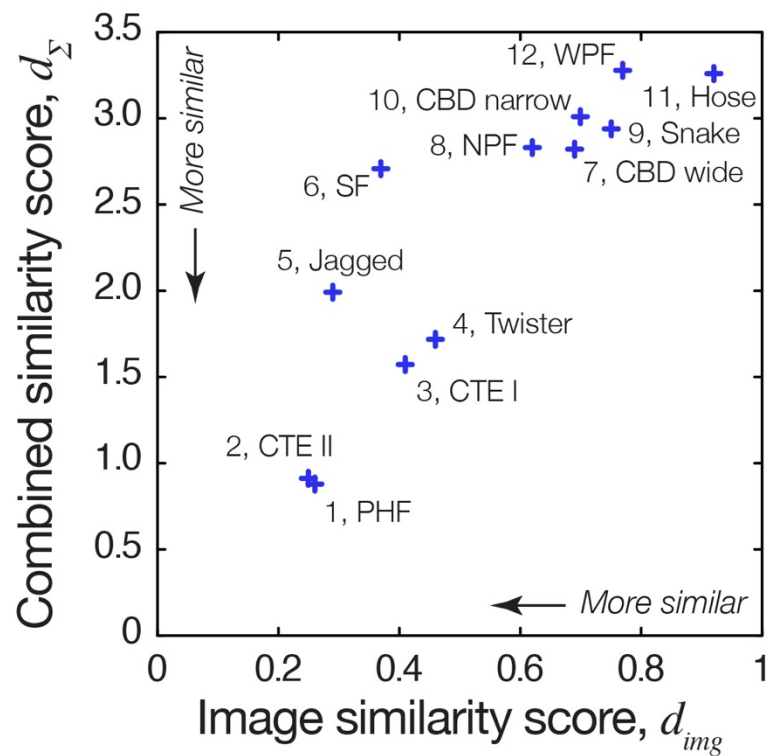
751 **Figure 5.** The image similarity score and the combined score of similarity between the dGAE

752 fibril AFM image data and cryo-EM derived density maps of tau fibril polymorphs. Lower

753 score denotes higher degree of similarity. The overall comparative rankings (from highest to

754 lowest) based on the combined score of similarity (Supplementary **Table S1**) are labelled for

755 each entry.

756

757

758 **SUPPLEMENTARY FIGURES**
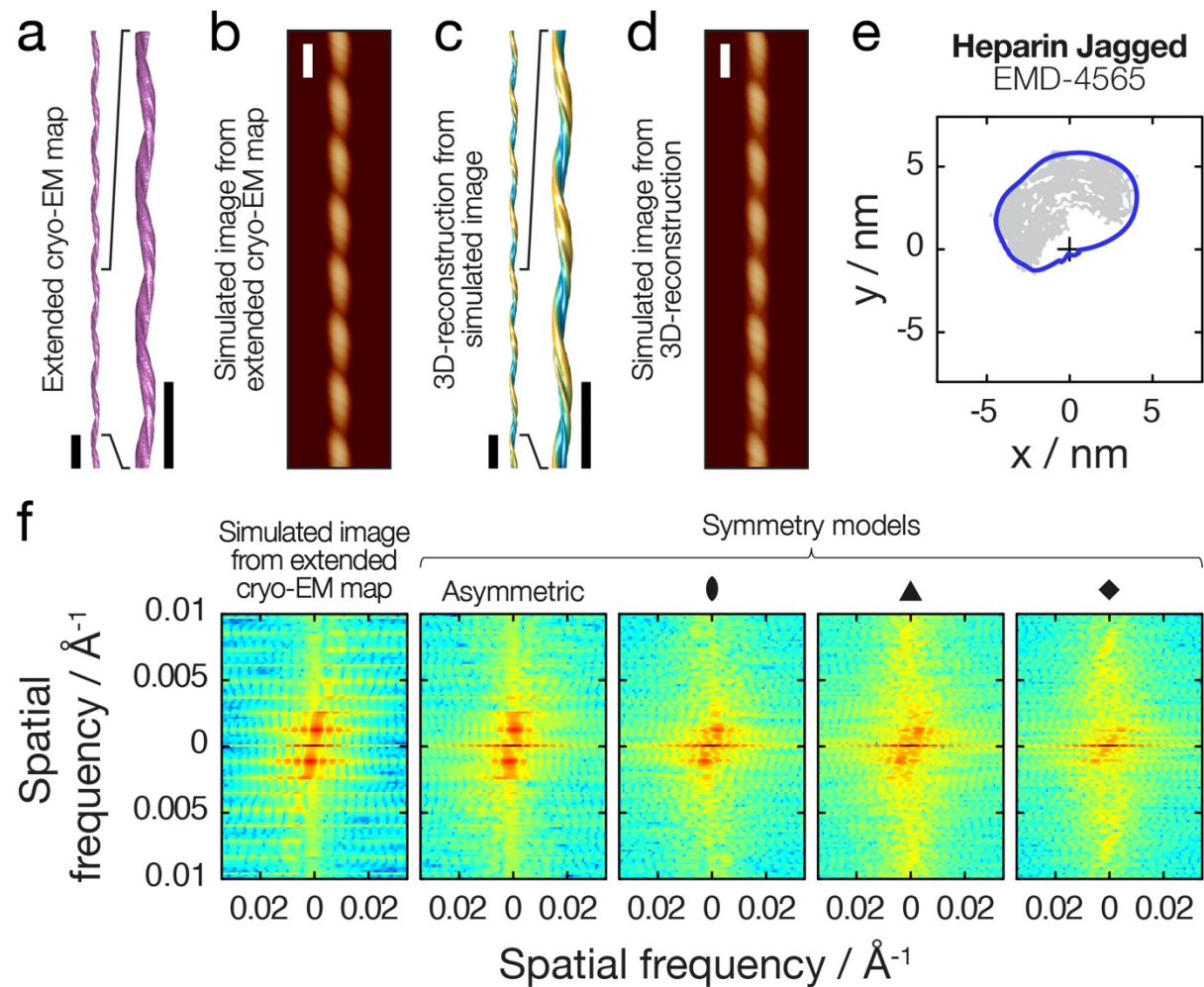
759

760

761
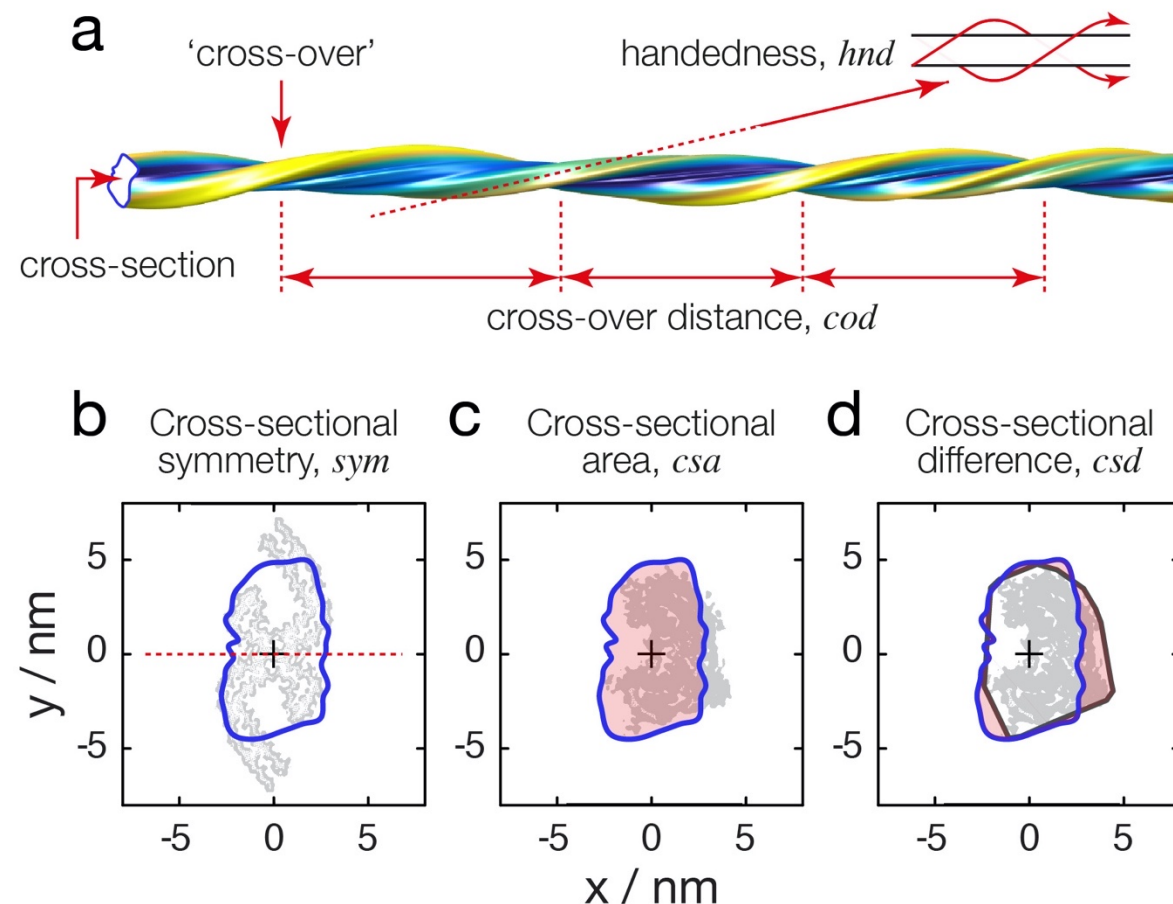
762 **Figure S1**. Schematic illustrations of the definitions of the morphometric parameters used for

763 quantitative comparative structural analysis. (a) Schematic diagram illustrating the structural

764 properties analysed, exemplified using a segment of the AFM derived dGAE fibril model

765 shown in **Fig 1c**. The comparative morphometric parameters calculated by analysis of the

766 cross-section are illustrated by the red line or areas in (b) for symmetry, (c) for cross-sectional

767 area and (d) for cross-sectional difference area. Grey shapes show the cross-sections of cryo-

768 EM density maps and the black line in (d) shows the tip-accessible cross-section, found by

769 tracing a circle with the radius equivalent to that of the AFM tip radius around the cryo-EM

770 density cross-section.

771

772

773

Figure S2. Validation of the 3D reconstruction and topograph simulation algorithms and workflow using a cryo-EM density map of PHF (EMD-0259). (a) Filament axis-aligned and lengthened cryo-EM density map of PHF. (b) Topographic AFM height image simulated from the density map in (a). (c) 3D surface envelope model reconstructed using the simulated image in (b). (d) Topographic AFM height image simulated from the 3D reconstructed model in (c). All scale bars represent 50 nm. (e) Comparison of the cross-sections of the EM density map (grey) and the AFM derived 3D model (blue). (f) Symmetry estimation during the 3D reconstruction process. As the symmetry is unknown, 3D reconstruction is performed with various symmetry estimates on the same AFM image data (here showing asymmetric, 2, 3, and

783   4-fold helical symmetries) and the 2D power spectra of the simulated topograph is compared

784   to that of the original image (i.e. image shown in b). Here 2-fold symmetry gave the best match

785   and was used for the final reconstruction.

786

787

788

789

790   **Figure S3**. Validation of the 3D reconstruction and topograph simulation algorithms and

791   workflow using a cryo-EM density map of a 'heparin-jagged' filament (EMD-4565). (a)

792   Filament axis-aligned and lengthened cryo-EM density map of 'heparin-jagged'. (b)

793   Topographic AFM height image simulated from the density map in (a). (c) 3D surface envelope

794   model reconstructed using the simulated image in (b). (d) Topographic AFM height image

795   simulated from the 3D reconstructed model in (c). All scale bars represent 50 nm. (e)

796   Comparison of the cross-sections of the EM density map (grey) and the AFM derived 3D model

797   (blue). (f) Symmetry estimation during the 3D reconstruction process. As the symmetry is

798   unknown, 3D reconstruction is performed with various symmetry estimates on the same AFM

799   image data (here showing asymmetric, 2, 3, and 4-fold helical symmetries) and the 2D power

800   spectra of the simulated topograph is compared to that of the original image (i.e. image shown

801   in b). Here the asymmetric model gave the best match and was used for the final reconstruction

802

803

804   **SUPPLEMENTARY TABLE**

805

806   **Table S1.** Quantitative comparison of image similarity and similarity in morphometric parameters between AFM image data of the dGAE fibril

807   shown in **Fig 1c** and cryo-EM derived tau fibril maps using the distance measure-based scoring system described in the Methods section. The

808   similarity rank is based on the combined score d_Σ, from most similar with lowest d_Σ value to least similar with highest d_Σ value.

809

Filament type	sym	hnd§	cod / nm	csa / nm²	csd / nm²	d_{img}	d_{sym}	d_{hnd}	d_{cod}	d_{csa}	d_{csd}	d_Σ	Rank
dGAE, AFM data	2	-1	49.1†	44.1	0.0								
CBD wide, EMD-10514	2	-1	141.2	106.1	62.8	0.69	0	0	0.51	0.82	0.80	2.82	7
CBD narrow, EMD-10512	1	-1	203.9	54.8	24.3	0.70	1	0	0.86	0.14	0.31	3.01	10
CTE type I, EMD-0527	2	-1	71.1	83.5	40.2	0.41	0	0	0.12	0.52	0.51	1.57	3
CTE type II, EMD-0528	2	-1	71.1	64.4	21.3	0.25	0	0	0.12	0.27	0.27	0.91	2
Narrow Pick's filament (NPF) , EMD-0077	1	-1	229.4	39.6	11.5	0.62	1	0	1.00	0.06	0.15	2.83	8
Wide Pick's filament (WPF) , EMD-0078	2	-1	141.0	119.9	78.2	0.77	0	0	0.51	1.00	1.00	3.28	12
Paired helical filament (PHF) , EMD-0259	2	-1	77.6	61.3	18.4	0.26	0	0	0.16	0.23	0.23	0.88	1
Straight filament (SF) , EMD-0260	1	-1	164.8	70.5	27.9	0.37	1	0	0.64	0.35	0.36	2.71	6
Heparin snake, EMD-4563	1	-1	134.3	48.0	52.2	0.75	1	0	0.47	0.05	0.67	2.94	9
Heparin twister, EMD-4564	1	-1	50.1	37.3	12.9	0.46	1	0	0.01	0.09	0.16	1.72	4
Heparin jagged, EMD-4565	1	-1	83.3	45.2	38.8	0.29	1	0	0.19	0.01	0.50	1.99	5
Heparin hose, EMD-4566	1	-1	161.1	30.6	42.6	0.92	1	0	0.62	0.18	0.54	3.26	11

810   §. Left-hand twisted filaments are assigned a *hnd* value of -1 and right-hand twisted filaments are assigned a *hnd* value of 1

811   †. The mean cross-over distance of the selected dGAE fibril shown in **Fig 1c-d**. The overall comparative similarity rank is unchanged if the

812   population mean *cod* value of dGAE fibrils of 72.7 nm (Al-Hilaly et al., 2020) is used instead.

813

814