









ARTICLE

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

OPEN

# Atomic structure of PI3-kinase SH3 amyloid fibrils by cryo-electron microscopy

Christine Röder <sup>1,2</sup>, Nicola Vettore <sup>2,6</sup>, Lena N. Mangels<sup>2,6</sup>, Lothar Gremer <sup>1,2</sup>, Raimond B.G. Ravelli <sup>3</sup>, Dieter Willbold <sup>1,2</sup>, Wolfgang Hoyer <sup>1,2</sup>, Alexander K. Buell <sup>2,4</sup> & Gunnar F. Schröder <sup>1,5</sup>

High resolution structural information on amyloid fibrils is crucial for the understanding of their formation mechanisms and for the rational design of amyloid inhibitors in the context of protein misfolding diseases. The Src-homology 3 domain of phosphatidylinositol-3-kinase (PI3K-SH3) is a model amyloid system that plays a pivotal role in our basic understanding of protein misfolding and aggregation. Here, we present the atomic model of the PI3K-SH3 amyloid fibril with a resolution determined to 3.4 Å by cryo-electron microscopy (cryo-EM). The fibril is composed of two intertwined protofilaments that create an interface spanning 13 residues from each monomer. The model comprises residues 1–77 out of 86 amino acids in total, with the missing residues located in the highly flexible C-terminus. The fibril structure allows us to rationalise the effects of chemically conservative point mutations as well as of the previously reported sequence perturbations on PI3K-SH3 fibril formation and growth.

<sup>1</sup>Institute of Complex Systems, Structural Biochemistry (ICS-6) and JuStruct, Jülich Center for Structural Biology, Forschungszentrum Jülich, 52425 Jülich, Germany. <sup>2</sup>Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany. <sup>3</sup>The Maastricht Multimodal Molecular Imaging Institute, Maastricht University, Universiteitssingel 50, 6229 ER Maastricht, The Netherlands. <sup>4</sup>Department of Biotechnology and Biomedicine, Technical University of Denmark, Søtofts Plads, 2800 Kgs Lyngby, Denmark. <sup>5</sup>Physics Department, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany. <sup>6</sup>These authors contributed equally: Nicola Vettore, Lena N. Mangels. Correspondence and requests for materials should be addressed to A.K.B. (email: [alebu@dtu.dk](mailto:alebu@dtu.dk)) or to G.F.S. (email: [gu.schroeder@fz-juelich.de](mailto:gu.schroeder@fz-juelich.de))

Alzheimer's and Parkinson's Disease as well as spongiform encephalopathies are prominent examples of protein misfolding diseases<sup>1</sup>. These disorders are characterised by the presence of amyloid fibrils<sup>2</sup>. Amyloid fibrils are straight and unbranched thread-like homopolymeric protein assemblies, which are stabilised by backbone hydrogen bonding between individual peptide molecules. These interactions lead to a highly ordered, repetitive cross- $\beta$  architecture, in which the  $\beta$ -strands run perpendicularly to the fibril axis.

It has been shown that in the case of neurodegenerative protein misfolding diseases, the final amyloid fibrils are often not the most cytotoxic species, but that small, oligomeric precursors are more hydrophobic and more mobile and hence more prone to deleterious interactions with cellular components<sup>3</sup>. However, recent progress in the mechanistic understanding of amyloid fibril formation shows that the mature fibrils can be the main source of toxic oligomers, due to their role as catalytic sites in secondary nucleation processes<sup>4</sup>. Furthermore, in the case of systemic amyloidosis diseases, where amyloid fibrils form in organs other than the brain, the amyloid fibrils themselves are the deleterious species, as their presence in large quantities can disrupt organ functions<sup>5</sup>.

Until recently, structural information on amyloid fibrils could only be obtained from relatively low-resolution methods, such as X-ray fibre diffraction<sup>6</sup>, limited proteolysis<sup>7</sup> and H/D exchange<sup>8</sup>. High-resolution structural information on amyloid fibrils has only become available in recent years through progress in solid state NMR spectroscopy (ssNMR)<sup>9–11</sup> and cryo-electron microscopy (cryo-EM). In particular cryo-EM enables atomic resolution structures of amyloid fibrils to be determined, and this has indeed been achieved in a few cases so far<sup>12–23</sup>. Such detailed structural information is crucial for the understanding of amyloid formation mechanisms, as well as for the rational design of inhibitors of the individual mechanistic steps, such as fibril nucleation and growth<sup>24</sup>.

Here we present the high-resolution cryo-EM structure of amyloid fibrils of the Src-homology 3 domain of phosphatidylinositol-3-kinase (PI3K-SH3). SH3 domains are kinase subdomains of usually <100 amino acids length and have been found to be part of more than 350 proteins, ranging from kinases and GTPases to adaptor and structural proteins, within various organisms<sup>25</sup>. SH3 domains are known to play a significant role in several signalling pathways where they mediate protein-protein interactions by recognising PxxP sequence motifs<sup>26,27</sup>. The structure of natively folded PI3K-SH3, a domain consisting of 86 amino acids from bovine PI3K, has been well-characterised by X-ray crystallography and NMR spectroscopy<sup>26,28</sup>.

Used initially as a model system for protein folding studies<sup>29</sup>, PI3K-SH3 was among the first proteins discovered to form amyloid fibrils in the test tube, while not being associated to any known human disease<sup>30</sup>. Fibril formation was observed at acidic pH, where in contrast to the native fold at neutral pH<sup>27,28</sup>, monomeric PI3K-SH3 lacks a well-defined secondary structure<sup>30–32</sup>. Since this discovery, PI3K-SH3 has played a pivotal role in advancing our fundamental understanding of the relationships between protein folding, misfolding and aggregation. Indeed, the hypothesis of the amyloid fibril as the most generic 'fold'<sup>1</sup> that a polypeptide can adopt was significantly shaped by the finding that PI3K-SH3 forms amyloid fibrils. Many pioneering studies on the basic biochemical, structural and mechanistic features of amyloid fibrils have been performed with PI3K-SH3. Early cryo-EM measurements highlighted the need for conformational rearrangement of the sequence within the fibril<sup>33</sup>. The important role of the destabilisation of native secondary structure elements and the need for non-native contacts and extensive structural rearrangements during the formation of

fibrillar aggregates was also observed for a related SH3 domain<sup>34</sup>. Despite not being related to any human disease, PI3K-SH3 aggregates were shown to be cytotoxic, suggesting sequence-independent toxic properties of amyloid fibrils and their precursors<sup>35</sup>. PI3K-SH3 also provided insight into the kinetics of molecular recycling of the monomeric building blocks of the fibril<sup>36</sup>, as well as into the dynamics of the formation of oligomeric precursors of amyloid fibrils<sup>37</sup>.

The structure of PI3K-SH3 fibrils we present here is in agreement with previous ssNMR data<sup>38</sup>, but we find that its interface is orthogonal to that suggested previously based on a low resolution reconstruction<sup>33</sup>. Indeed, the inter-filament interface in PI3K-SH3 fibrils is large compared to that of other amyloid fibrils determined to-date, and is formed from residues distant in primary sequence. With the atomic model we can rationalise the effect of newly designed, as well as previously reported sequence variants of PI3K-SH3 on the kinetics of fibril growth. Our study therefore not only adds important insight into the structural variety of amyloid fibrils, but also demonstrates how such structures can be used to rationalise the dynamics of protein assembly processes.

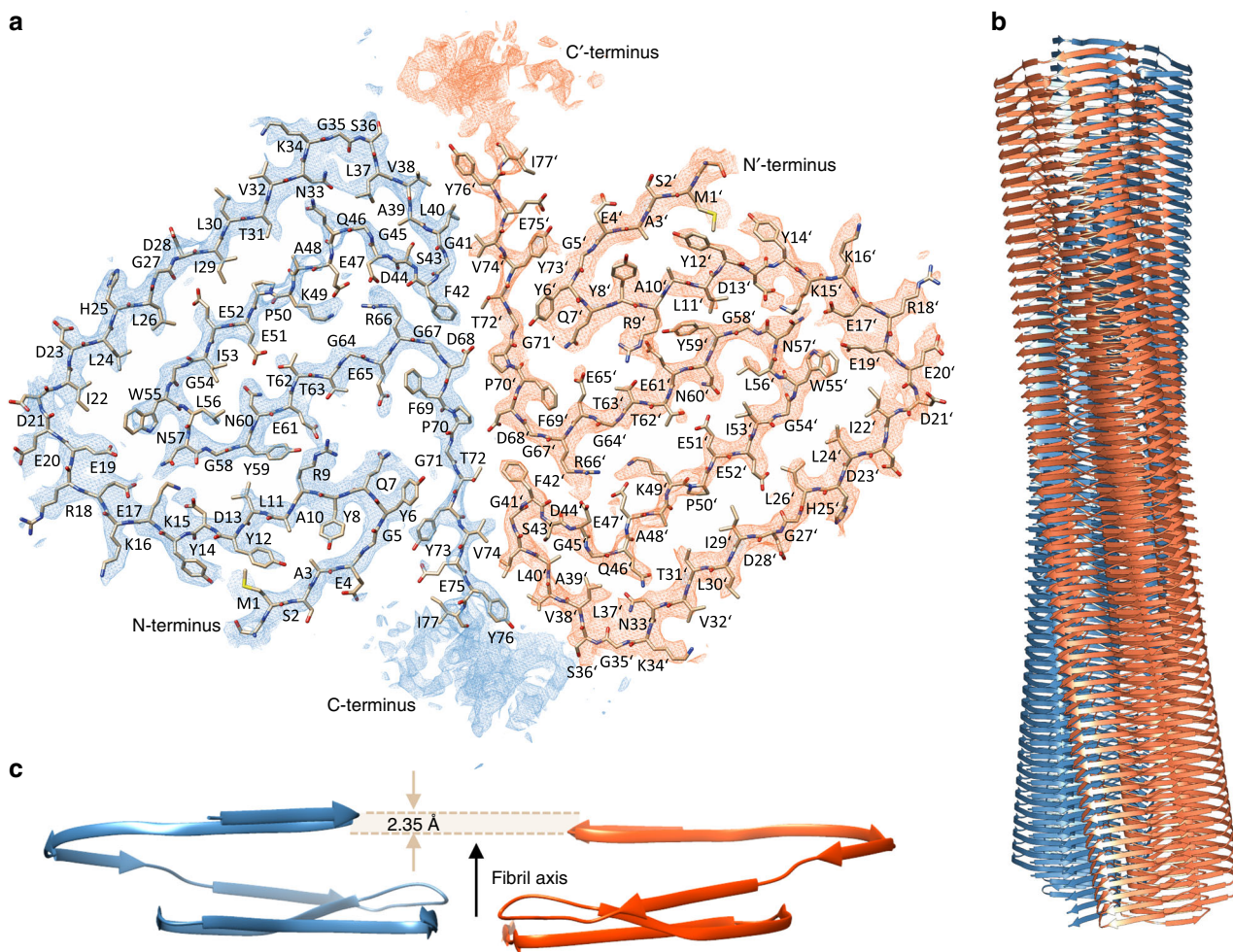
## Results

**Structure determination by cryo-EM.** Fibril formation by the full-length PI3K-SH3 domain under acidic solution conditions<sup>30</sup> led to long, straight fibrils of which the main population could be structurally determined by cryo-EM (Fig. 1). High overall homogeneity of the preparation has been shown by atomic force microscopy (AFM) and negative staining EM imaging (Fig. 2, Supplementary Figs. 1 and 2). Nevertheless, two different morphologies could be distinguished in both AFM and EM images (Fig. 2, Supplementary Fig. 1). The predominant, thick morphology has a diameter of 7–8 nm while the other, thinner morphology exhibits about half the diameter of the thick fibril (Supplementary Fig. 2).

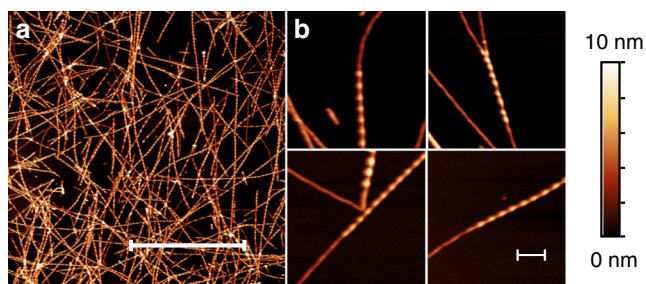
For cryo-EM imaging, samples of PI3K-SH3 amyloid fibrils were flash-frozen on Quantifoil cryo-EM grids and imaged with a Tecnai Arctica microscope (200 kV) equipped with a Falcon 3 direct electron detector (Supplementary Fig. 3). Image processing and helical reconstruction were performed with RELION-2<sup>39–41</sup>. A three-dimensional density map for the thick PI3K-SH3 fibril could be reconstructed to an overall resolution of 3.4 Å. The clear density of the fibril allowed us to build an atomic model for residues 1–77 out of a total of 86 amino acids (Fig. 1). The missing nine residues are located at the C-terminus, which shows blurred density likely due to substantial flexibility. Previously reported low-resolution cryo-EM<sup>33</sup> data are in good agreement with our structure (Supplementary Fig. 4).

**Architecture of the PI3K-SH3 amyloid fibril.** The thick PI3K-SH3 fibril is a left-handed helical structure consisting of two intertwined protofilaments, and is thus called DF (double filament) fibril. The handedness of the density reconstruction was determined by comparison with AFM images (see Methods and Supplementary Fig. 5). From an analysis of the fibril height profiles in AFM images, we determined the helical pitch to be  $170 \pm 10$  nm in reasonable agreement with a pitch of 140 nm obtained from the cryo-EM structure.

Protofilament subunits (PI3K-SH3 monomers) are stacked in a parallel, in-register cross- $\beta$  structure. The spacing between the layers of the cross- $\beta$  structure is around 4.7 Å and well visible in the density (Supplementary Fig. 6). The subunits in the two opposing protofilaments are not on the same z-position along the fibril axis but are arranged in a staggered fashion (Fig. 1c). The helical symmetry is therefore described by a twist of  $179.4^\circ$  and



**Fig. 1** Double Filament SH3 fibril structure. **a** Cross-section of the double filament (DF) PI3K-SH3 fibril (two monomers). The different protofilaments are colored blue and orange. Density maps are shown at a contour level of  $1.4 \sigma$ . The density map was sufficiently clear to model residues 1-77. **b** Side view of filaments twisting around each other displaying a total of 125 layers. **c** Relative arrangement of two adjacent PI3K-SH3 monomers within the fibril, showing a 2.35 Å shift between the protofilaments



**Fig. 2** AFM images of PI3K-SH3 amyloid fibrils. **a** AFM image of PI3K-SH3 fibrils (scale bar, 3 µm). **b** Four different close-up views of fibrils showing both the thick, double filament (DF) fibril and thin, single filament (SF) fibril morphology (scale bar, 300 nm)

rise of 2.35 Å, which corresponds to an approximate  $2_1$ -screw symmetry. The same staggering arrangement has been observed previously for other fibril structures, such as amyloid- $\beta$ (1-42)<sup>13</sup> and paired-helical filaments of tau<sup>23</sup>.

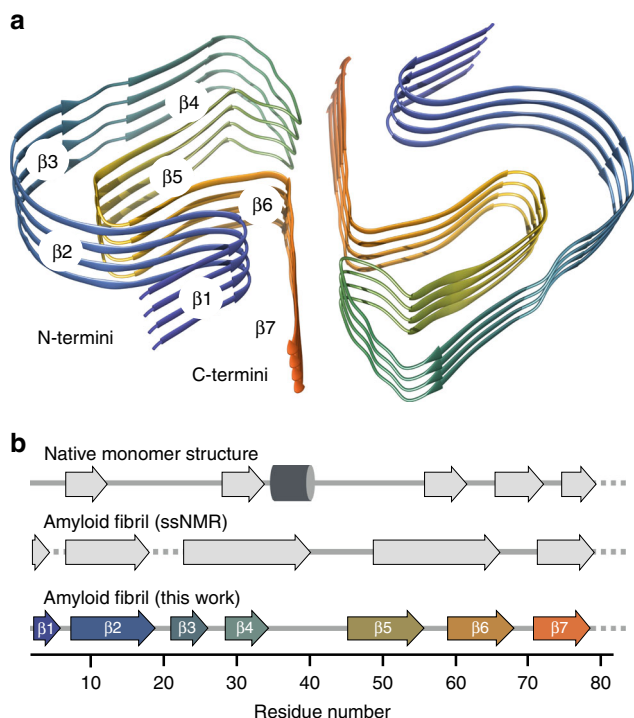
Each protofilament subunit, or PI3K-SH3 monomer, consists of seven parallel in-register  $\beta$ -strands ( $\beta$ 1, aa1-5;  $\beta$ 2, aa7-19;  $\beta$ 3, aa22-26;  $\beta$ 4, aa28-34;  $\beta$ 5, aa46-56;  $\beta$ 6, aa59-68;  $\beta$ 7, aa72-77) that

are interrupted by either sharp kinks, glycine or proline residues—or a combination of those (Figs. 1a and 3, Supplementary Fig. 6). In particular glycine residues at kink positions have also been observed in other amyloid structures<sup>17</sup>. The total of seven kinks and turns (Fig. 3a) results in an amyloid key topology<sup>20</sup>, which includes a structural motif similar to the bent  $\beta$ -arch described by Li et al.<sup>17</sup> for the  $\alpha$ -synuclein fibril.

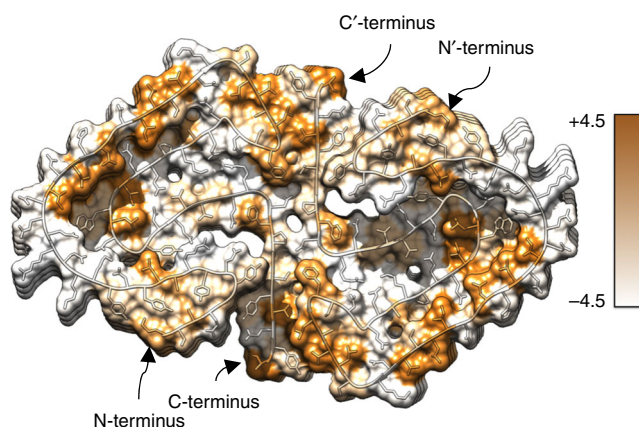
By encompassing 13 residues of each monomer, 26 in total (Supplementary Fig. 7), the inter-filament interface of the PI3K-SH3 DF fibril is very large. The protofilaments mainly interact through two identical symmetry-related hydrophobic patches (Fig. 4) at the periphery of their interface, between the bottom part of the bent  $\beta$ -arch motif and the respective C-terminal part of the opposing protofilament. The stability is provided by the hydrophobic clusters of Val<sup>38</sup>, Leu<sup>40</sup>, Gly<sup>41</sup>, Phe<sup>42</sup>, Val<sup>74</sup>, Tyr<sup>67</sup> and Val<sup>38</sup>, Leu<sup>40</sup>, Gly<sup>41</sup>, Phe<sup>42</sup>, Val<sup>74</sup>, Tyr<sup>67</sup> (Figs. 1a and 4).

The amyloid-characteristic cross- $\beta$  motif composed of parallel in-register  $\beta$ -sheets connects the different DF fibril layers and therefore contributes the largest share of intermolecular contacts (Fig. 5, Supplementary Fig. 6). This cross- $\beta$  stacking is complemented by multiple inter- and intramolecular contacts including side chain interactions in homosteric and heterosteric zippers<sup>42</sup>. A further noteworthy feature of the structure is the fact that the PI3K-SH3 subunits are not planar but extend along





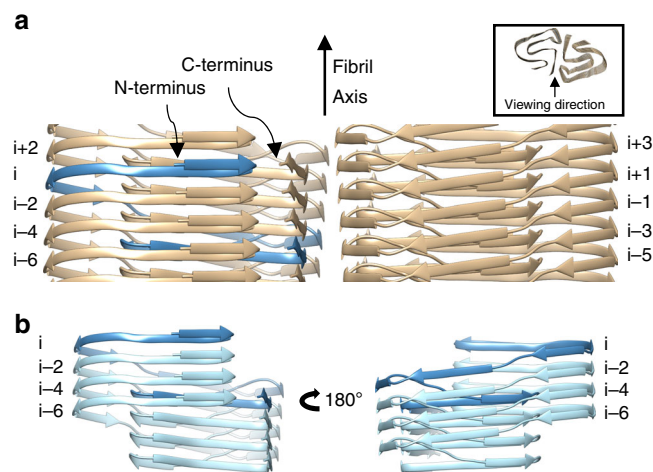
**Fig. 3** Secondary structure comparison of the PI3K-SH3 DF fibril. Secondary structure of the presented cryo-EM structure compared to data obtained previously by solid-state NMR (fibrils) or liquid-state NMR (monomeric, native state). **a** Tilted cross-section of four SH3 DF fibril layers. Secondary structure is predominantly formed by seven cross- $\beta$  sheets. **b** Secondary structure comparison to the native solution structure (PDB: 1PKS<sup>28</sup>) and ssNMR results (fibrils) modified according to Bayro et al.<sup>38</sup>. Flexible regions are shown as dashed lines,  $\beta$ -sheets as arrows, and the helix as a cylinder



**Fig. 4** Hydrophobicity of the fibril cross-section. Hydrophobicity levels of the SH3 DF fibril cross-section are coloured according to Kyte-Doolittle<sup>61</sup>. Hydrophobic residues are mainly packed within the fibril core, while hydrophilic residues point towards the solvent. Hydrophobic patches in both monomers are clearly visible and are spanning several  $\beta$ -sheets

the fibril axis (Fig. 5). The subunits within a protofilament therefore interact not only with the layer directly above ( $i + 2$ ) and below ( $i - 2$ ), but also with layers up to ( $i + 6$ ) and ( $i - 6$ ); the subunits are interlocked.

A single PI3K-SH3 monomer in the DF fibril exhibits an amyloid key topology (Fig. 1a), which is stabilised mainly by hydrophobic patches stretching between the strands  $\beta 3$ – $\beta 5$  (Ile<sup>22</sup>( $i$ ), Trp<sup>55</sup>( $i - 4$ ), Leu<sup>24</sup>( $i$ ), Leu<sup>26</sup>( $i$ )) (Fig. 4), and hydrogen bonds



**Fig. 5** Side view of the secondary structure of the atomic model. **a** Single subunit  $i$  highlighted in blue with adjacent subunits in beige, described by even numbers, while the subunits of the opposite protofilament are described by odd numbers. **b** View of the minimal fibril unit ( $i$ ) to ( $i - 6$ ) in one filament. The minimal fibril unit is displayed from two views by a turn of 180° highlighting the course of one monomer spanning several fibril layers

or salt bridges between strands  $\beta 5$  and  $\beta 6$  (Asp<sup>44</sup>( $i$ ), Arg<sup>66</sup>( $i - 2$ ), Glu<sup>47</sup>( $i$ )) and  $\beta 2$ – $\beta 6$  (Arg<sup>9</sup>( $i$ ), Glu<sup>61</sup>( $i - 6$ )) (Figs. 1a and 3a). In the turn between strands  $\beta 5$  and  $\beta 6$ , Asp<sup>13</sup>( $i + 4$ ) might bind to Lys<sup>15</sup>( $i + 4$ ) while Glu<sup>17</sup>( $i + 4$ ) might interact with Asn<sup>57</sup>( $i$ ) (Figs. 1a and 3a). The bent  $\beta$ -arch motif between strands  $\beta 4$  and  $\beta 5$  is potentially strengthened by a contact between Asn<sup>33</sup>( $i$ ) and Gln<sup>46</sup>( $i$ ) that would tie the motif together (Figs. 1a and 3a). Further possible electrostatic interactions can be observed between strands  $\beta 2$  and  $\beta 6$  where the amino-group of Arg<sup>9</sup>( $i + 4$ ) might exhibit a salt bridge to Glu<sup>61</sup>( $i - 2$ ) (Figs. 1a and 3a). In addition, aromatic side chains are located in close proximity to glycine residues Tyr<sup>6</sup>( $i$ )-Gly<sup>71</sup>( $i + 6$ ), Phe<sup>42</sup>( $i$ )-Gly<sup>67</sup>( $i + 2$ ), Tyr<sup>73</sup>( $i$ )-Gly<sup>5</sup>( $i - 6$ ), indicating a potential involvement of glycine-aromatic  $\text{Ca-H}\cdots\pi$ -interactions<sup>43,44</sup>. Further aliphatic-aliphatic and aromatic-aliphatic interactions comprise Ile<sup>29</sup>( $i$ )-Pro<sup>50</sup>( $i - 2$ ,  $i - 4$ )-Ala<sup>48</sup>( $i - 2$ ,  $i - 4$ ), Met<sup>1</sup>( $i$ )-Tyr<sup>12</sup>( $i$ ), and Leu<sup>11</sup>( $i$ )-Tyr<sup>59</sup>( $i - 4$ ). We could also observe possible interactions in-between monomer layers, so-called hydrogen bond ladders, e.g. with Gln<sup>7</sup>, Gln<sup>46</sup>, Asn<sup>57</sup> and Asn<sup>60</sup>. Additionally, these ladders could potentially be formed as well by side chains pointing outside of the fibril such as Glu<sup>4</sup>, Glu<sup>20</sup>, Asp<sup>21</sup>, Asp<sup>23</sup> and Asp<sup>28</sup>. Remarkably, nearly all side chains in the outer Ca-chain that are oriented towards the solvent are non-hydrophobic (Fig. 1a) (Ser<sup>0</sup>-Glu<sup>4</sup>, Tyr<sup>14</sup>-Ser<sup>36</sup>). Hence, the polar outside of the fibril shields the hydrophobic interface (Fig. 4). Only the hydrophobic residues Leu<sup>30</sup> and Val<sup>32</sup>, which are located on the fibril surface next to the bent  $\beta$ -arch motif, disrupt this pattern (Figs. 1a and 4).

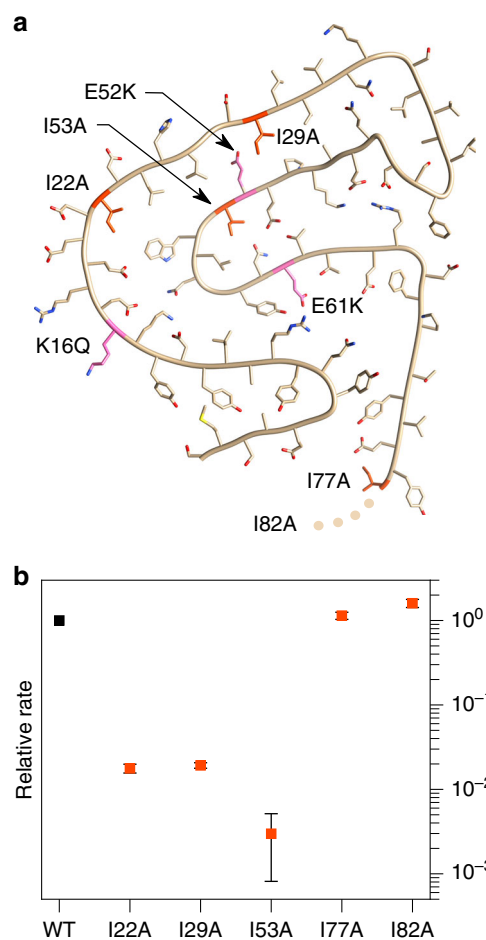
Further hints towards the role of electrostatic interactions in maintaining the structure, comes from a pH-shift experiment. By changing the pH from 2 to 7.4, we observed that fibrils depolymerise almost completely after 1 h (Supplementary Fig. 8), in agreement with the highly dynamic nature of PI3K-SH3 fibrils reported previously<sup>36</sup>. We propose three clusters that may be highly influenced by this pH shift: (a) the space between strands  $\beta 5$  and  $\beta 6$  (Asp<sup>44</sup>( $i$ ), Arg<sup>66</sup>( $i - 2$ ), Glu<sup>47</sup>( $i$ )); (b) the interactions between Asp<sup>13</sup>( $i$ ) and Lys<sup>15</sup>( $i$ ), and between Glu<sup>17</sup>( $i + 4$ ) and Asn<sup>57</sup>( $i$ ); (c) the solvent-exposed and protonatable patch involving Asp<sup>21</sup>( $i$ ), Asp<sup>23</sup>( $i$ ) and His<sup>25</sup>( $i$ ) (also highlighted as having a fundamental role in amyloid formation<sup>45</sup>).

The secondary structure of PI3K-SH3 fibrils and monomers has been analysed by comparing available solution NMR

(monomeric native fold, PDB: 1PKS)<sup>28</sup> and ssNMR data (amyloid fold)<sup>38</sup> with our model (Fig. 3b). The only feature that is shared by all models is the flexible C-terminus starting around residue 80. Analysis of protein contact maps via the Contact Map WebViewer<sup>46</sup> of the DF fibril compared to the monomeric native structure<sup>28</sup> showed no consistent residue contacts in both structures, illustrating the substantially different conformation that the monomer unit has to adopt in order to incorporate into the fibril. The secondary structure of the native and amyloid fold differs substantially apart from a  $\beta$ -sheet between residues 70–80. While the monomeric native structure is characterised by one helix between residues 34 and 39,  $\beta$ -sheets and multiple flexible loop-regions, the DF fibril consists of seven  $\beta$ -sheets exclusively, that are almost uninterrupted (Fig. 3b). The longest break in the  $\beta$ -sheet pattern of the fibril is the bent  $\beta$ -arch motif leading to a rigid loop between Gly<sup>35</sup> and Gly<sup>45</sup> (Figs. 1a and 3a). Our findings are consistent with former results by Bayro et al.<sup>38</sup> who proposed a PI3K-SH3 amyloid model based on solid-state NMR data. Both structures show  $\beta$ -sheets as the only secondary structure motif with most of the  $\beta$ -sheet regions corresponding (Fig. 3b). The main differences between the ssNMR and cryo-EM structures are found in the region between residues 25 and 60. Here, ssNMR data suggest two wide-spanning  $\beta$ -sheets while according to cryo-EM data this region consists of not two but four  $\beta$ -sheets, that are disrupted by a glycine residue, Gly<sup>27</sup> ( $\beta$ 3– $\beta$ 4) (Supplementary Fig. 6) and a sharp kink, Gly<sup>54</sup>-Tyr<sup>59</sup> ( $\beta$ 5– $\beta$ 6) (Fig. 3a).

**Impact of mutations on the SH3 aggregation.** We probed the sensitivity of the fibril growth kinetics towards chemically conservative single point mutations by substituting isoleucine residues for alanine across the protein sequence. We expressed and purified the sequence variants and experimentally quantified the rates at which five different variants (I22A, I29A, I53A, I77A, I82A) elongated wildtype (WT) fibrils by quartz crystal microbalance (QCM) measurements (Supplementary Fig. 9)<sup>47,48</sup>. This technology is ideally suited for such cross-seeding experiments. The fibril growth rates of WT and variant proteins can be directly compared, given that the same, constant ensemble of fibrils is monitored. An additional big advantage is that the use of WT seeds as templates ensures that the sequence variants adopt the same fibril structure as the WT, and therefore the change in fibril growth kinetics with respect to the wild type can be interpreted in terms of the perturbation induced by the sequence modification.

We measured the rates of WT fibril elongation by the different variants and expressed the rates relative to that of the elongation by WT monomer (Fig. 6, Supplementary Fig. 9). We found that the relative elongation rates differ by more than two orders of magnitude, with the mutations in the first third of the sequence, as well as around the middle of the sequence, displaying slow elongation rates (I22A =  $0.02 \pm 10.9\%$ , I29A =  $0.02 \pm 7.5\%$ , I53A =  $0.003 \pm 72.7\%$ ), whereas the mutations close to or within the disordered C-terminus display approximately the same rates as the wild type (I77A =  $1.13 \pm 9.3\%$ , I82A =  $1.59 \pm 10.3\%$ ). We also took AFM images of the different single point mutants at a concentration of  $100 \mu\text{M}$  that were all seeded with WT-derived fibrils and incubated for 2 days at room temperature under quiescent conditions (Supplementary Fig. 10). From these images, it can be seen that all fibrils at the end of the experiment have a very similar morphology and length, except for those in the sample with I53A, where shorter fibrils are observed. The fibril growth of all the mutants seems to have come to completion within this time scale leading to very similar fibril lengths. Only in the case of I53A, the fibril growth rate is so slow that the available monomer was only partly incorporated into the seed fibrils



**Fig. 6** Point mutations and their effect on fibril elongation rates. **a** Top view of a PI3K-SH3 monomer derived from DF fibrils. Highlighted residues: Isoleucine residues mutated to alanine (I22A, I29A, I53A, I77A, I82A), from this work (red) and charged residues mutated by Buell et al.<sup>47</sup> (K16Q, E61K, E52K, pink). Mutation I82A is missing in the model due to the flexible C-terminus. **b** Elongation rate of Ile-to-Ala mutants. The elongation rates are normalised to a WT rate of one. Source data are provided as a Source Data file

during the course of the experiment. However, the very high mass sensitivity of the QCM permits to resolve even the growth rate of this slow growing mutant, and hence we base our analysis of the relative growth rates exclusively on the QCM experiments (Supplementary Fig. 9).

## Discussion

Concerning fibril architecture and polymorphism, we observed that, in addition to the thick 7–8 nm SH3 DF fibril consisting of two filaments, also thin fibrils of about half the diameter are present in both AFM (Fig. 2, Supplementary Fig. 2) and negative stain EM images (Supplementary Fig. 1). We therefore hypothesise that the thin fibrils correspond to single filament fibrils (SF fibrils) of the same type that make up the DF fibrils. Our hypothesis is further supported by AFM images that suggest that long SF fibrils can contain stretches that appear to be identical to DF fibrils (Supplementary Fig. 2). It therefore seems likely that fragments of an additional protofilament can attach to or grow on a given SF fibril, to form DF fibrils. The incomplete cooperativity between the elongation of the individual filaments in a DF fibril suggests relatively weak interactions between the filaments. This, in turn, implies that one filament could well be stable without

contact to another filament. We conclude that, at least in the case of PI3K-SH3, inter-filament interface contacts are not necessary for fibril formation.

A similar observation has been made for  $\beta$ 2-microglobulin fibrils<sup>15</sup>, which can comprise a single protofilament as well as two or more filaments resulting in at least six different polymorphs without major differences in the filament structure. Accordingly, a mixture of all polymorphs yielded only a single set of NMR resonances<sup>15</sup>. The stability of individual PI3K-SH3 filaments is particularly noteworthy, inasmuch as the inter-filament interface in the DF fibrils is comparatively large (Supplementary Fig. 7). A large inter-filament interface does therefore not indicate stable inter-filament interactions. We note here also that the identical conformation of every subunit along the fibril axis is in line with fibril growth by PI3K-SH3 monomer addition<sup>49,50</sup>.

A single subunit in the PI3K-SH3 DF fibril is not planar but winds itself along the protofilament axis (Fig. 5), which leads to interlocking within a protofilament. In contrast, in a planar subunit, the inter-subunit interactions within one protofilament would consist exclusively of the cross- $\beta$  pattern, while all transverse interactions would be intramolecular. Here, however, one subunit within one protofilament is not only in contact with its direct neighbours (above and below) but with four other monomers in total, i.e. in addition to the longitudinal hydrogen bonds in the cross- $\beta$  structure, there are other transverse interactions between the subunits. This staggered arrangement leads to an interlocking mechanism connecting several monomers within a protofilament, which very likely further stabilises the structure (Fig. 5). Interestingly, this interlocking mechanism is commonly seen in other amyloid fibril structures determined by cryo-EM<sup>13,17,18,20,21</sup>. It should be noted that NMR data can only distinguish between intra- and intermolecular contacts but cannot directly reveal a potential staggering of subunits along the fibril axis.

Given that the interlocking is observed in other fibril structures as well, it might in general contribute to the formation of stable fibrils by optimisation of the side chain packing<sup>10</sup>. In addition, the staggered architecture might in part be responsible for the templating effect during fibril elongation, as it establishes a rugged binding interface that may guide the incoming monomer into the fibril conformation, engaging it in more intermolecular contacts than a flat interface could.

As a measure for the interlocking of subunits we have previously defined the concept of a minimal fibril unit<sup>13</sup>, which is the smallest fibril structure fragment in which the capping subunits at both ends would have established the same full contact interface with other constituting monomers as the capping subunits of an extended fibril. Since we hypothesise that one protofilament of the DF fibril could exist on its own and is identical to the SF fibril, here we describe the minimal fibril unit also for a single protofilament. For the PI3K-SH3 fibril, the minimal fibril unit has a size of four subunits when considering an individual protofilament, and a size of eight subunits in the case of the DF fibril (Fig. 5b).

In order to rationalise kinetic data of fibril formation and growth, we substituted five different isoleucine residues with alanines, probing how these chemically conservative single point mutations at different positions affect the addition of new monomers to the fibrillar structure formed by the wild type sequence. Three mutations (I22A, I29A, I53A) showed a strong decrease of the elongation rate of two to three orders of magnitude compared to the wild type sequence (Fig. 6). In these positions, the side chains of the three isoleucines are all pointing towards the fibrillar core (Fig. 6). While the chemical nature of the amino acid substitution we chose is conservative (aliphatic to aliphatic), the bulkiness of the side chain decreases. The ability of

these variants to elongate the WT structure, albeit significantly slower than the WT sequence itself, suggests that the formation of a cavity due to the reduced bulkiness is energetically tolerated. However, the free energy difference between the monomeric state and the transition state (structural ensemble) of the elongation reaction appears to be increased. Such an increase in energy difference could come either from a stabilisation of the monomeric state, or from a destabilisation of the transition state. The former seems less likely, due to the mostly disordered nature of PI3K-SH3 at pH 2, while the latter possibility could be caused by the reduction in hydrophobic contacts between the monomer and the fibril end. This conclusion, which assumes some degree of contact between the monomer and the fibril end in the transition state for fibril growth, is in excellent agreement with previous results that underline the importance of the sequence hydrophobicity for the magnitude of the elongation free energy barrier<sup>51</sup>. The remaining two mutations (I77A, I82A) (Fig. 6) are instead located close to or within the flexible C-terminus and show indeed a much weaker or no effect on the elongation rate. In both cases a chemical modification to alanine does not perturb any interactions crucial for the energetics of the transition state.

The availability of the high-resolution structure also allows us to rationalise the influence of previously reported single point mutations of PI3K-SH3 on the kinetics of fibril elongation. In a previous study<sup>47</sup>, the effect of changes in charge at three different positions (Fig. 6; K16Q, E52K, E61K; residues depicted in pink) led to very different effects on the elongation rate. While these mutations, similar to the ones we have designed and studied in the present work, could lead to different fibrillar structures if induced to form fibrils *de novo*, the use of WT seed fibrils in both studies allows us to discuss here the effect of these mutations in the light of the present fibril structure. This is because the well-known templating effect in amyloid fibril growth imprints the structure of the fibril template onto the monomeric protein that adds onto the fibril end. The mutation E52K probably leads to the creation of a positive charge inside the hydrophobic fibril core, a highly unstable arrangement. In the WT fibril, Glu<sup>61</sup> can form a salt bridge with Arg<sup>9</sup>, which leads to an additional driving force for the deprotonation of Glu<sup>61</sup>. If Glu<sup>61</sup> is substituted by a lysine (E61K), two positive charges come into close proximity, again leading to a highly unstable situation. On the other hand, the mutation K16Q does not lead to any major change in the kinetics of fibril elongation, which is most likely due to the side chain pointing towards the outside of the fibril<sup>52</sup>.

Our structural model can also help to understand the effects of further sets of previously investigated sequence changes<sup>45</sup>. With the aim of understanding which part of the sequence plays a major role in the amyloidogenicity of the protein, different portions of the sequence were replaced or mutated. In the vast majority of cases, the amyloidogenic behaviour was completely abolished, which we can now explain with our structural model: the introduction of bulky or charged residues facing the inner core of the structure destabilises the present amyloid fold, as evidenced by the mutants E17R/D23R and Q7E/R9K/E17R/D23R<sup>45</sup>. The only mutation that does not show a significant decrease in amyloidogenicity does not involve changing the charge of a buried residue: the mutant referred to as PI3-QMR (E17Q/D23M/H25R) modifies a single charge of outward pointing residue 25 (assuming Glu<sup>17</sup> and Asp<sup>23</sup> to be protonated at pH 2<sup>45</sup>).

It should also be noted that the circularisation of the PI3K-SH3 sequence through the use of disulphide bridges causes a decrease of the elongation rate but does not prevent the circularised mutant to acquire the amyloid conformation<sup>53</sup>. The close proximity of the N- and C-termini (Fig. 1a), in combination with the flexibility of the C-terminal nine residues, likely allows the



cyclised (disulfide bridge between Cys<sup>3</sup> and Cys<sup>82</sup>) sequence to form an amyloid structure very similar, if not identical to the one presented here.

A noteworthy example of substantial sequence modification is the grafting of the N-Src loop of SPC-SH3 onto PI3K-SH3, while simultaneously removing the stretch from residue 31 to 53 from PI3K-SH3<sup>54</sup>. This operation did not remove the amyloidogenic properties from the modified PI3K-SH3 domain. The removed sequence stretch is part of the bent  $\beta$ -arch motif in the amyloid conformation. The removal of this prominent motif could have been expected to lead to a more significant impairment of amyloid fibril formation. However, the altered sequence is likely to be able to respond to this strong perturbation by forming an alternative structure.

In summary, we have determined the structure of a PI3K-SH3 amyloid fibril. The PI3K-SH3 fibril has been extensively studied in the past and the effect of many mutations on the kinetics of amyloid formation has been described. The atomic structure of the PI3K-SH3 fibril enables us now to rationalise the effect of these mutations, which is the basis for understanding the sequence-dependence of amyloidogenicity and ultimately the determinants of amyloid formation in general.

## Methods

**Protein production.** WT and mutants of the bovine PI3K-SH3 domain were purified according to the protocol of Zurdo et al.<sup>31</sup>. All constructs contain a 6xHis-tag linked to the protein by a thrombin cleavage site. The sequence of the WT protein after cleavage is the following, with the peptide Gly-Ser remaining as overhang from the cleavage:

GSMSAEGYQYRALYDYKKEREEDIDLHLGDLITVNGKSL  
VALGFSDDGQEAKEPEIGWLNQYNETTGERGDFPGTYVEYIGRK  
KISF

The protein was expressed in a BL21 *E. coli* strain with TB medium for auto-induction containing 0.012% glucose and 0.048% lactose. The cells were grown for over 24 h and then harvested by centrifugation. After resuspension in sodium phosphate buffer (50 mM sodium phosphate pH 8, 5 mM imidazole and 100 mM NaCl), the cells were disrupted by sonication, in presence of protease inhibitors and DNase. The lysate was centrifuged, and the supernatant loaded on a Ni-NTA Superflow Cartridge (Qiagen, Venlo, Netherlands) equilibrated in 50 mM sodium phosphate pH 8, 5 mM imidazole and 100 mM NaCl. The protein was eluted with a linear gradient from 5 to 300 mM imidazole in 50 mM sodium phosphate pH 8, 100 mM NaCl in 25 ml elution volume. Fractions containing the protein were collected and cleaved overnight at 7°C with 1 unit of thrombin (from bovine plasma, Sigma-Aldrich Saint Louis, Missouri, USA) per 1 mg of protein. The cleaved solution was then concentrated and loaded on a SEC HiLoad 26/60 Superdex 75 column (GE Healthcare, Chicago, Illinois, USA) equilibrated with 5 mM ammonium acetate pH 7. Fractions containing the PI3K-SH3 domain were collected and lyophilised for further use.

**Fibril formation.** The lyophilised protein was resuspended in 10 mM glycine-hydrochloride pH 2.5 buffer at a final concentration of ca. 200  $\mu$ M. The solution was shaken in an Eppendorf tube at 1400 rpm at 42°C for 24 h to form seeds. These seeds were then sonicated in an Eppendorf tube in a volume of ca. 500  $\mu$ l for 15 s (1 s 'on', 2 s 'off', 10% amplitude) with a Bandelin Sonopuls using a M72 probe. To prepare the twisted fibrils, a new solution with ca. 100  $\mu$ M monomer in 10 mM glycine-hydrochloride pH 2.5 was then mixed with 5  $\mu$ M of equivalent seeds mass and incubated without stirring at 50°C overnight.

**AFM imaging.** The fibril samples were diluted in 10 mM glycine-hydrochloride, pH 2.5 to a concentration of 5  $\mu$ M and 10  $\mu$ l were pipetted on a mica substrate. After 10 min of incubation, the mica was washed extensively with milliQ water and dried under a nitrogen gas flush. The pictures were taken in tapping mode on a Bruker Multimode 8 (Billerica, Massachusetts, USA) using OMCL-AC160TS cantilevers (Shinjuku, Tokyo, Japan).

**Fibril elongation measurements with QCM.** The elongation rate of PI3K-SH3 fibrils was measured through immobilisation of fibrils on a QCM sensor and subsequent incubation with monomer solution<sup>50</sup>. To immobilise the fibrils on the sensor, chemical modification is necessary. To achieve that, the fibrils were mixed at a final concentration of 50  $\mu$ M in buffer (10 mM glycine-hydrochloride, pH 2) with EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride) (1 M) and cystamine hydrochloride (0.5 mg ml<sup>-1</sup>). After pelleting and washing the chemically modified fibrils, they were sonicated in an Eppendorf tube in a volume of ca. 500 microlitre with a Bandelin Sonopuls using a MS72 probe (10%

amplitude, 15 s, 1 s 'on', 2 s 'off'). The gold sensors (Biolin Scientific, Gothenburg, Sweden) were then incubated with the above-mentioned solution overnight in a 100% humidity environment. The measurements were performed with a QSense Pro (Biolin Scientific, Gothenburg, Sweden) by measuring the elongation rate as change in resonant frequency over time. With the temperature set at 25°C, the monomer solutions were injected for 30 s at a flow rate of 100  $\mu$ l per second and the measurement lasted until a stable slope was reached. To obtain the relative rates, the protein solutions were injected in different sensor chambers after a WT injection, the latter being used as normalization reference. Two different triplicate measurements were performed for I53A. Two different duplicate measurements were performed for all the other mutants. The rate was measured as slope of the 3rd overtone and averaged among the multiple injections. The data are presented as average values with error bars indicating the standard deviation.

**Fibril dissociation at pH 7.4.** Fibril dissociation at pH 7.4 was probed by measuring ThT and intrinsic fluorescence change over time in two series of triplicates. The ThT measurements were performed by mixing 40  $\mu$ l of 100  $\mu$ M PI3K-SH3 fibrils and 50  $\mu$ M ThT in 10 mM glycine-hydrochloride pH 2 with 60  $\mu$ l of 100 mM sodium phosphate pH 7.4. The mixing was carried out using the injection system of a CLARIOstar plate reader (BMG LABTECH, Ortenberg, Germany) and measuring immediately afterwards by exciting at 440 nm and recording the signal intensity at 480 nm. The intrinsic tryptophan fluorescence measurements were carried out by mixing the same two solutions (without ThT) by pipetting, followed by the measurement of fluorescence spectra every 15 s by exciting at 290 nm and recording between 300 and 380 nm in 2 nm intervals.

The analysis of soluble peptide by concentration determination at the end of the dissolution experiment was performed after one night of equilibration after the mixing of the two solutions mentioned above (without ThT). The samples were spun down for 30 min at 16,100 g. The protein concentration in the supernatant was determined by measuring the absorbance at 280 nm together with the extinction coefficient of PI3K-SH3 of  $\epsilon_{280} = 15,930 \text{ M}^{-1}\text{cm}^{-1}$  using a V650 UV-Vis spectrophotometer (Jasco, Easton, MD, USA).

**Negative stain and cryo-EM image acquisition.** Negatively stained fibrils were prepared on 400 mesh carbon-coated copper grids (S160-4, Plano GmbH, Germany), stained with 1% uranyl acetate, and imaged using a Libra120 electron microscope (Zeiss) operated at 120 kV. Cryo-preparation was performed on glow-discharged holey carbon films (Quantifoil R 1.2/1.3, 300 mesh). The sample containing 50  $\mu$ M PI3K-SH3 was 4 $\times$ 10 $\times$ 20 $\times$  diluted with 10 mM glycine-hydrochloride (pH 2) to a final concentration of 12.5, 5 or 2.5  $\mu$ M monomer equivalent. A total sample volume of 2.5  $\mu$ l was applied onto the carbon grid and blotted for 3.5 s before being cryo-plunged using a Vitrobot (FEI). With 110,000-fold nominal magnification 622 micrographs have been recorded on a Tecnai Arctic electron microscope operating at 200 kV with a field emission gun using a Falcon III (FEI) direct electron detector in electron counting mode directed by EPU data collection software. Each micrograph was composed of 60 fractions. Each fraction contained 42 frames, i.e. in total 2520 frames were recorded per micrograph. The samples were exposed for 65 s to an integrated flux of 0.4 e<sup>-</sup>/Å<sup>2</sup>/s. Applied underfocus values ranged between 1.5 and 2.25  $\mu$ m. The pixel size was 0.935 Å, as calibrated using gold diffraction rings within the powerspectra of a cross grating grid (EMS, Hatfield).

**Cryo-EM image processing and helical reconstruction.** MotionCor2<sup>55</sup> was used for movie correction. Fitting CTF parameters for all 622 micrographs was performed using CTFFIND4<sup>56</sup>. Further image processing and 3D reconstructions were done with RELION-2<sup>39,40</sup>. Selection of 256 micrographs was done with CCTFFIND by estimating the maximum resolution at which Thon rings could be detected to be better than 5 Å. From these micrographs, 4540 fibrils were manually picked. From these fibrils, 103,733 segments were extracted using an overlap of 90% between neighbouring segments. The size of the segment images is 220 pixels. For data set characterisation we performed 2D classification (Supplementary Fig. 11). As an initial model for the refinement we used a noise-filled cylinder.

After several rounds of 3D refinements with helical symmetry search, we found a problem with the tilt priors: the tilt angle distribution became bimodal with maxima at 85° and 95°. However, we would expect the tilt angles to show a unimodal distribution around 90°. The  $\sim$ 4.7 Å cross- $\beta$  pattern is a strong signal and substantially affects the alignment. If the helical rise parameter is slightly smaller than the correct value, the cross- $\beta$  pattern can still be aligned by changing the tilt angles to higher or lower values (which accordingly reduces the spacing of the cross- $\beta$  pattern). To overcome this problem, we fixed the tilt prior to 90° by usage of the RELION option *helical\_keep\_tilt\_prior\_fixed*, and then first optimized the helical parameters. In subsequent refinements, the helical symmetry parameters were fixed and the tilt angles (together with the other angles) were optimized.

Since the automated 3D refinement in RELION did not yield high-resolution reconstructions, we performed gold-standard refinements by splitting the data into an even and an odd set by selecting entire fibrils (not just segments, as they are overlapping). The FSC curve (Supplementary Fig. 12) was computed between the two half-maps and yields a resolution (with the 0.143 criterion) of 3.4 Å.

The handedness of the fibril structure was determined by comparing the reconstructed density with AFM images (Supplementary Fig. 5). For this comparison, the 3D density map of the fibril was converted to a height profile using Chimera<sup>57</sup> as follows: set surface color by height and set the color scale to gray. Then set camera projection mode to orthographic and save the image. The alignment of the AFM image with the calculated height profiles yields a cross-correlation coefficient of 0.943 for the left-handed and 0.914 for the right-handed helix.

**Model building and refinement.** A single chain atomic model of PI3K-SH3 was built with Coot<sup>58,59</sup>. Subsequently, seven copies of a single chain were placed into the EM density map. At residues Gln<sup>46</sup> and Glu<sup>47</sup>, between strands  $\beta$ 4 and  $\beta$ 5, the density map is slightly ambiguous and could possibly be in agreement with an alternative interpretation for the trace of the Ca-chain (Supplementary Fig. 13).

The final model containing seven helical symmetry-related chains was used for further real space refinement in PHENIX<sup>60</sup>. Refinement was carried out using a resolution cut-off of 3.4 Å and NCS restraints between all seven subunits. At later stages of the refinement, hydrogen-bond restraints were defined for the cross- $\beta$  sheets and Ramachandran restraints were used. The model-map FSC curve as obtained from *phenix.real\_space\_refine* is shown in Supplementary Fig. 12 (dashed line). The final statistics on the details of the refinement are shown in Supplementary Table 1. Molecular graphics and analyses were performed with Chimera<sup>57</sup>.

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

### Data availability

The structure of the PI3K-SH3 fibril has been deposited in the Protein Data Bank under accession code 6R4R [<https://doi.org/10.2210/pdb6R4R/pdb>]. The 3.4 Å cryo-EM density map has been deposited in the Electron Microscopy Data Bank under accession code EMD-4727. The source data underlying Fig. 6b and Supplementary Fig. 8 are provided as a Source Data file. Other data are available from the corresponding authors upon reasonable request.

Received: 26 March 2019 Accepted: 9 July 2019

Published online: 21 August 2019

### References

- Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890 (2003).
- Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 (2014).
- Kayed, R. et al. Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**, 486–489 (2003).
- Cohen, S. I. A. et al. Proliferation of amyloid- $\beta$ 42 aggregates occurs through a secondary nucleation mechanism. *PNAS* **110**, 9758–9763 (2013).
- Pepys, M. B. et al. Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature* **362**, 553 (1993).
- Sunde, M. et al. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273**, 729–739 (1997).
- Kheterpal, I., Williams, A., Murphy, C., Bledsoe, B. & Wetzel, R. Structural features of the Abeta amyloid fibril elucidated by limited proteolysis. *Biochemistry* **40**, 11757–11767 (2001).
- Hoshino, M. et al. Mapping the core of the  $\beta$ 2-microglobulin amyloid fibril by H/D exchange. *Nat. Struct. Mol. Biol.* **9**, 332–336 (2002).
- Fitzpatrick, A. W. P. et al. Atomic structure and hierarchical assembly of a cross- $\beta$  amyloid fibril. *PNAS* **110**, 5468–5473 (2013).
- Tycko, R. Molecular structure of aggregated amyloid- $\beta$ : insights from solid state nuclear magnetic resonance. *Q. Rev. Biophys.* **39**, 1–55 (2016).
- Wasmer, C. et al. Amyloid fibrils of the HET-s(218–289) prion form a  $\beta$  solenoid with a triangular hydrophobic core. *Science* **319**, 1523–1526 (2008).
- Radamaker, L. et al. Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis. *Nat. Commun.* **10**, 1103 (2019).
- Gremer, L. et al. Fibril structure of amyloid- $\beta$ (1–42) by cryo-electron microscopy. *Science* **358**, 116–119 (2017).
- Zhang, W. et al. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *eLife* **8**, e43584 (2019).
- Iadanza, M. G. et al. The structure of a  $\beta$ 2-microglobulin fibril suggests a molecular basis for its amyloid polymorphism. *Nat. Commun.* **9**, 4517 (2018).
- Falcon, B. et al. Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **561**, 137–146 (2018).
- Li, B. et al. Cryo-EM of full-length  $\alpha$ -synuclein reveals fibril polymorphs with a common structural kernel. *Nat. Commun.* **9**, 3609 (2018).
- Guerrero-Ferreira, R. et al. Cryo-EM structure of alpha-synuclein fibrils. *eLife* **7**, e360402 (2018).
- Li, Y. et al. Amyloid fibril structure of  $\alpha$ -synuclein determined by cryo-electron microscopy. *Cell Res.* **28**, 897–903 (2018).
- Liberta, F. et al. Cryo-EM fibril structures from systemic AA amyloidosis reveal the species complementarity of pathological amyloids. *Nat. Commun.* **10**, 1104 (2019).
- Swuec, P. et al. Cryo-EM structure of cardiac amyloid fibrils from an immunoglobulin light chain (AL) amyloidosis patient. *Nat. Commun.* **10**, 1269 (2019).
- Falcon, B. et al. Tau filaments from multiple cases of sporadic and inherited Alzheimer's disease adopt a common fold. *Acta Neuropathol.* **136**, 699–708 (2018).
- Fitzpatrick, A. W. P. et al. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185–190 (2017).
- Arosio, P. et al. Kinetic analysis reveals the diversity of microscopic mechanisms through which molecular chaperones suppress amyloid formation. *Nat. Commun.* **7**, 10948 (2016).
- Koch, C. A., Anderson, D., Moran, M. F., Ellis, C. & Pawson, T. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. *Science* **252**, 668–674 (1991).
- Yu, H. et al. Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933–945 (1994).
- Batra-Safferling, R., Granzin, J., Mödder, S., Hoffmann, S. & Willbold, D. Structural studies of the phosphatidylinositol 3-kinase (PI3K) SH3 domain in complex with a peptide ligand: Role of the anchor residue in ligand binding. *Biol. Chem.* **391**, 33–42 (2010).
- Koyama, S. et al. Structure of the PI3K SH3 domain and analysis of the SH3 family. *Cell* **72**, 945–952 (1993).
- Guijarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D. & Dobson, C. M. Folding kinetics of the SH3 domain of PI3 kinase by real-time NMR combined with optical spectroscopy. *J. Mol. Biol.* **276**, 657–667 (1998).
- Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D. & Dobson, C. M. Amyloid fibril formation by an SH3 domain. *PNAS* **95**, 4224–4228 (1998).
- Zurdo, J., Guijarro, J. I., Jiménez, J. L., Saibil, H. R. & Dobson, C. M. Dependence on solution conditions of aggregation and amyloid formation by an SH3 domain. *J. Mol. Biol.* **311**, 325–340 (2001).
- Polverino de Lauroto, P. et al. Protein aggregation and amyloid fibril formation by an SH3 domain probed by limited proteolysis. *J. Mol. Biol.* **334**, 129–141 (2003).
- Jiménez, J. L. et al. Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *EMBO J.* **18**, 815–821 (1999).
- Neudecker, P. et al. Structure of an intermediate state in protein folding and aggregation. *Science* **336**, 362–366 (2012).
- Bucciantini, M. et al. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507–511 (2002).
- Carulla, N. et al. Molecular recycling within amyloid fibrils. *Nature* **436**, 554–558 (2005).
- Orte, A. et al. Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *PNAS* **105**, 14424–14429 (2008).
- Bayro, M. J. et al. High-resolution MAS NMR analysis of PI3-SH3 amyloid fibrils: backbone conformation and implications for protofilament assembly and structure. *Biochemistry* **49**, 7474–7484 (2010).
- Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
- He, S. & Scheres, S. H. W. Helical reconstruction in RELION. *J. Struct. Biol.* **198**, 163–176 (2017).
- Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016).
- Riek, R. & Eisenberg, D. S. The activities of amyloids from a structural perspective. *Nature* **539**, 227–235 (2016).
- Zondlo, N. J. Aromatic-proline interactions: electronically tunable CH/ $\pi$  interactions. *Acc. Chem. Res.* **36**, 1039–1049 (2013).
- Unterreitmeier, S. et al. Phenylalanine promotes interaction of transmembrane domains via GxxxG motifs. *J. Mol. Biol.* **374**, 705–718 (2007).
- Ventura, S. et al. Short amino acid stretches can mediate amyloid formation in globular proteins: the Src homology 3 (SH3) case. *PNAS* **101**, 7258–7263 (2004).
- Kozma, D., Simon, I. & Tusnády, G. E. CMWeb: an interactive on-line tool for analysing residue-residue contacts and contact prediction methods. *Nucleic Acids Res.* **40**, 329–333 (2012).
- Buell, A. K. et al. Position-dependent electrostatic protection against protein aggregation. *ChemBioChem* **10**, 1309–1312 (2009).
- Knowles, T. P. J. et al. Kinetics and thermodynamics of amyloid formation from direct measurements of fluctuations in fibril mass. *PNAS* **104**, 10016–10021 (2007).
- Collins, S. R., Douglass, A., Vale, R. D. & Weissman, J. S. Mechanism of prion propagation: amyloid growth occurs by monomer addition. *PLoS Biol.* **2**, e321 (2004).



50. Buell, A. K. et al. Frequency factors in a landscape model of filamentous protein aggregation. *Phys. Rev. Lett.* **104**, 228101 (2010).
51. Buell, A. K. et al. Detailed analysis of the energy barriers for amyloid fibril growth. *Angew. Chem. Int. Ed.* **51**, 5247–5251 (2012).
52. Buell, A. K. et al. Electrostatic effects in filamentous protein aggregation. *Biophys. J.* **104**, 1116–1126 (2013).
53. Graña-Montes, R. et al. Contribution of disulfide bonds to stability, folding, and amyloid fibril formation: the PI3-SH3 domain case. *Antioxid. Redox Signal.* **16**, 1–15 (2011).
54. Ventura, S., Lacroix, E. & Serrano, L. Insights into the origin of the tendency of the PI3-SH3 domain to form amyloid fibrils. *J. Mol. Biol.* **322**, 1147–1158 (2002).
55. Zheng, S. Q., Palovcak, E., Armache, J.-P., Cheng, Y. & Agard, D. A. MotionCor2: anisotropic correction of beam-induced motion for improved single-particle electron cryo-microscopy. *Nat. Methods* **14**, 331–332 (2017).
56. Rohou, A. & Grigorieff, N. CTFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
57. Pettersen, E. F. et al. UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
58. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 486–501 (2010).
59. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
60. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 213–221 (2010).
61. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).

## Acknowledgements

We thank P.J. Peters and C. López-Iglesias for their advice and helpful discussions, H. Duimel for help with sample preparation, and the M4I Division of Nanoscopy of Maastricht University for microscope access and support. We thank P. Neudecker for helpful comments on the paper. We gratefully acknowledge the computing time granted by the Jülich Aachen Research Alliance High-Performance Computing (JARA-HPC) Vergabegremium and VSR commission on the supercomputer JURECA at Forschungszentrum Jülich. Computational support and infrastructure was provided by the Center for Information and Media Technology (ZIM) at the University of Düsseldorf (Germany). A.K.B. and N.V. thank the Deutsche Forschungsgemeinschaft (DFG) for funding. D.W. was supported by grants from the Portfolio Technology and Medicine, the Portfolio Drug Design, and the Helmholtz-Validierungsfonds of the Impuls und Vernetzung-Fonds der Helmholtzgemeinschaft. This study was funded in part by the DFG SFB 974 and SFB 1208 (to D.W.). Support from a European Research Council

(ERC) Consolidator Grant (grant agreement no. 726368) to W.H. is acknowledged. A.K.B. thanks the Novo Nordisk Foundation for support through a Novo Nordisk Foundation Professorship.

## Author contributions

A.K.B. and G.F.S. conceived the study. N.V. and L.N.M. performed the biochemical experiments. A.K.B., N.V., L.N.M. analysed the biochemical and kinetics data. R.G.B.R. performed cryo-EM experiments and initial data analysis. C.R. and G.F.S. performed image processing, reconstruction and model building. C.R., A.K.B., N.V., G.F.S., W.H. and L.G. wrote the paper. D.W. and all other authors discussed results and commented on the paper.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-11320-8>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Peer review information:** *Nature Communications* thanks Henning Stahlberg, Stefano Ricagno and other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019