Neurodegenerative and other protein misfolding diseases are associated with the aggregation of a protein, which may be mutated in genetic forms of disease, or the wild type form in late onset sporadic disease. A wide variety of proteins and peptides can be involved, with aggregation originating from a natively folded or a natively unstructured species. Large deposits of amyloid fibrils are typically associated with cell death in late stage pathology. In this review, we illustrate the contributions of cryo-EM and related methods to the structure determination of amyloid fibrils extracted post mortem from patient brains or formed in vitro. We also discuss cell models of protein aggregation and the contributions of electron tomography to understanding the cellular context of aggregation.

Addresses
1 Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, 3227 Broadway, Quad 4C, New York, NY 10027, USA
2 Institute of Structural and Molecular Biology, Birkbeck College London, Malet St, London WC1E 7HX, UK

Corresponding authors: Fitzpatrick, Anthony WP (awpfitzpatrick@gmail.com), Saibil, Helen R (h.saibil@mail.cryst.bbk.ac.uk)

Introduction: amyloid diseases
An alternative fold of polypeptides that is not strongly coupled to amino acid sequence violates the general notion that a unique, stable three-dimensional (3D) structure is associated with each gene sequence encoding a protein. Moreover, this alternate fold, the cross-β structure, is strongly associated with invariably fatal degenerative diseases, for which there are few diagnostic tools and hardly any treatments [1]. A wide variety of different peptides and proteins can be converted into fibrils containing long ribbons of cross-β structure, generally known as amyloid fibrils.

Alzheimer’s disease (AD) is the commonest form of dementia, the progressive decline of cognitive processes that affects about 20% of people over 75. Extracellular plaques of the Amyloid-β (Aβ) peptide and intraneuronal tangles of hyperphosphorylated tau protein (a microtubule associated protein) in the brains of AD patients are associated with neuronal death and loss of cholinergic pathways [2]. AD is not usually inherited, but some genetic factors predispose individuals to the disease and lower the age of onset. Most of the genes known to be involved play a role in the metabolism of APP (β-amyloid precursor protein), a transmembrane protein which is cleaved by proteases termed secretases to release the Aβ peptide. Aβ is normally 40 residues long, but abnormal cleavage produces Aβ 1–42, which has a strong tendency to form amyloid.

The next most common neurodegenerative disease is Parkinson’s (PD). PD and the related Lewy body disease, which accounts for about 20% of dementia cases, cause movement abnormalities and dementia. Lewy bodies are intracellular deposits of the protein α-synuclein, a highly conserved synaptic protein (named for both synaptic and nuclear localization) that plays a role in synaptic vesicle recycling [3]. The most conserved region of the protein is predicted to form an amphipathic helix, consistent with a role in binding lipids or vesicles [4].

A group of severe neurodegenerative diseases with a clear genetic cause, including Huntington’s disease (HD), arise from expansion of glutamine repeats (CAG codon), and are known as CAG repeat diseases. The protein huntingtin (>3000 residues, with long stretches of glutamine and also proline repeats near the N terminus), essential for embryonic neural development, is found in intranuclear and cytoplasmic deposits of degenerating neurons [5]. The repeat region is in exon 1 of the protein, and the aggregates contain significant amounts of isolated exon 1. Individuals with less than 37–42 repeated glutamines do not develop the disease, but those with over 42 glutamines do, in correlation with its in vitro aggregation into amyloid-like fibrils and ribbons. The longer the repeat, the earlier the disease onset, and the more aggregation prone the protein. A single particle cryo-EM structure revealed the fold of soluble, full length huntingtin, which consists largely of HEAT repeats, but does not contain density for exon1, owing to disorder [6].

The prion diseases (transmissible spongiform encephalopathies) are the only amyloid diseases that have been clearly demonstrated to be transmissible. Ritual cannibalism formerly practiced by the Fore people of New Guinea spread the fatal disease kuru through the population. Creutzfeld Jacob disease (CJD) has been transmitted to patients receiving human growth hormone, and about 200 people died of variant CJD, derived from mad cow disease, itself propagated in cows via contaminated feed in the 1980s. The prion protein PrP is a brain
glycoprotein on the cell surface (GPI-anchored) that may play a role (so far unknown) in synaptic function. In its normal (cellular) form, PrP$$^{\text{C}}$$, it has a mainly α-helical structure. The β-sheet containing PrP$$^{\text{SC}}$$ (the scrapie form, named after the original ovine disease) from diseased brain can induce PrP$$^{\text{C}}$$ in previously healthy individuals to convert to the scrapie form through an unknown mechanism. Consistent with the clear link between the amyloid forming protein and disease, PrP knockout mice are not susceptible to infection. Different strain types can be detected by different glycosylation patterns, and these are propagated in experimental animals [7].

There are many major, unanswered questions about amyloid disease. Although pathology always involves cell death, related to aggregation of the protein(s) involved, the toxic species, whether fibril fragments, assembly intermediates or products, remain unknown, as do their precise mechanisms of inducing toxicity. In prion disease, the infectious species responsible for transmission of the prion conformation is distinct from the species causing toxicity, which begins to develop only after the infectivity has reached its maximum level [7]. In general, fibril structures extracted from patient brain differ from those grown in vitro from the same precursor protein, but the structural differences have not so far provided an explanation for the pathology. In addition, the tissue specificity and delayed onset characteristic of each disease variant are not understood. In the case of intrinsically disordered proteins, hints of physiological function are emerging, through reversible liquid–liquid phase separation providing a protective environment to shield the less stable proteins from temporary stress. But at least in some cases, the liquid phase can progress to irreversible fibrillar aggregates.

**Amyloid fibril structure and assembly**

Aggregation often occurs after exposure of the monomer to denaturing conditions and low pH, and requires partial or total unfolding of the native structure to allow the formation of intermolecular β-sheets. Mutations that destabilise the native fold can favour the conformational conversion. Despite the great variety in length, composition and sequence of proteins and peptides capable of forming amyloid aggregates, the resulting fibrils are characterized by their archetypal all-β tertiary structure that is a common feature of protein molecules in the amyloid state [8–10]. Their common core structure is the cross-β fold, shown in the first panel of Figure 1, in which β-strands form long ribbon-like sheets that pack against each other to form long protofilaments that assemble together into fibrils. They share a cross-β X-ray diffraction pattern arising from the 4.8 Å β-strand spacing and the ~10 Å inter-sheet spacing. They have a filamentous morphology, and bind thioflavin-T (ThT), Congo red and other amyloidophilic compounds [11,12]. Given that the accumulation of protein aggregates is a common characteristic of many clinically, genetically and pathologically distinct neurodegenerative diseases, determining the atomic structures of these fibrous aggregates, which are strongly associated with selective neurodegeneration, is likely to be important in understanding their formation, spreading and clearance in the human body.

Until recently, few amyloid structures were determined at high resolution. Recent advances in solid state NMR [13,14] and particularly in cryo-EM [15**] have generated substantial advances in structural information. A key element has been the dramatic improvement in resolution achievable by single particle cryo-EM, because of its power in dealing with heterogeneous assemblies. Amyloid fibrils are typically variable in protofilament structure, packing and in long range helical twist. Such variations limit the resolution achievable by solution methods that average the fibril structures present in the sample. However, single particle cryo-EM can be used to computationally extract short segments of the fibrils and to classify them into structurally homogeneous subsets (e.g. Ref. [16]).

With current electron detectors and software, this approach has yielded a series of near atomic structures revealing the subunit fold in fibrils of physiologically relevant amyloid proteins including yeast prions [17], α-synuclein [18*], Aβ42 [19*] and tau, the latter extracted from post mortem human brain (Figure 1). These structures have made clear that specific conformers of a protein molecule manifest in diseased tissue, as the folds of the tau filament core for paired helical and straight filaments in AD [15**] (Figure 1) are significantly different from its fold in Pick’s disease [20**]. As mentioned, structures derived from fibrils produced in vitro often differ those that form in the brain; for example, infectious mammalian PrP fibrils extracted from mice have been characterized using negative stain EM and found to adopt a short, paired double helical structure, distinct from in vitro formed fibrils [21].

Collectively, these reports reveal a great variety of cross-β molecular folds [22], as well as polymorphism in subunit structure and packing interfaces [15**,20**,23*]. Specific folds are likely to be influenced by the environment in which the aggregates are formed. Indeed, the in vitro aggregation kinetics depend strongly on environmental conditions, such as the presence of the corresponding seed fibrils [24,25] or lipid vesicles [26]. Moreover, the interface between protofilaments has been proven variable for disease-specific conformers, which may provide a molecular basis for the emergence of fibril strains (Figure 1). This is particularly evident for paired versus straight helical filaments of tau in AD [15**] (Figure 1).

**β-strand stacking and sequence specificity**

In general, fibrils tend to adopt a parallel, in-register arrangement of their constituent β-strands [27]. This
Cross-β structure and tau deposits in Alzheimer’s disease brain. From left to right, the upper panels show a schematic of the β-strand arrangement in the cross-β core of an amyloid fibril; the AD brain used for cryo-EM; and a Thioflavin-S stained light microscopy image showing abundant neurofibrillary tangles in temporal cortex. Lower panels, an electron micrograph of negatively stained filaments with a blue arrow indicating a paired helical filament (PHF) and a green arrow indicating a straight filament (SF); cryo-EM reconstructions of PHFs (blue) and SFs (green) with detailed cross sections; and de novo atomic models of filaments showing C-shaped subunits stacked to form each protofilament, with protofilaments paired into twisted polymorphic fibrils (Tau panels adapted from Ref. [15†]).

Current Opinion in Structural Biology

may be because antiparallel β-strands have an intrinsic limit on β-strand length ([28] and that parallel β-sheets tend to contain more β-strands than antiparallel β-sheets [29], making in-register parallel β-strand stacking of identical residues more suitable for templated amyloid fibril growth [30]. However, little consensus is evident among the atomic models of the few filament cores that have been solved to date, raising the question as to how and why a protein adopts a specific fold within the amyloid state in different pathologies.

The two all-β structural motifs most closely associated with amyloid and prion structure are cross-β packing [31] and β-helices [17] respectively (Figure 2). The generic cross-β structure, widely accessible to peptides and proteins of varying lengths with unrelated amino acid sequences, is a polypeptide scaffold characterized by arrays of continuous β-sheets, separated by an inter-sheet distance of 8–12 Å, running parallel to the long axis of the fibrils (Figure 2a). The importance of sequence-independent hydrogen bonding in defining the cross-β fold is underscored by the observation that polar, non-polar and even homopolymeric sequences of amino acids can form cross-β amyloid fibrils [32].

By contrast, β-helices, and in particular β-solenoids, are intricate structures composed of three β-sheets arranged in a triangular fashion, mediated by specific interactions not accessible to most amino acid sequences [33] (Figure 2b). A 21 amino acid sequence capable of forming a four amino acid β-strand, a two-residue β-arch, a five amino acid β-strand, a two-residue or three-residue β-arc,
and a six amino acid β-strand all closed by a ~70° glycine conformation is likely to have the following polar (P) / nonpolar (H) periodicity: HPH/PPHPPHPHPHHPH/PHGPH/P. Such hydrophobic patterning, with the occurrence of a glycine residue at position 19, is non-universal, requiring specific packing of side-chains (Figure 2b).

The tight regulation of β-helices by geometric and amino acid sequence restraints that do not constrain generic cross-β arrangements implies that these two all-β structural motifs display different behaviours. While cross-β packing (Figure 2a) is largely unaffected by point mutation, which typically only alters the kinetics of cross-β fibril formation [34], the mutation of the universally conserved β-arch-forming glycine (Figure 2b) to alanine is sufficient to disrupt the formation of the β-helix, preventing seeding and infectivity of the [Het-s] prion of the filamentous fungus Podospora anserina [35]. The sequence-independent nature of cross-β filament formation [32] makes it more widely accessible to many peptides and proteins than the sidechain specificity of β-helical assembly. The promiscuity of the cross-β fold allows this structural scaffold to template dissimilar protein molecules and cross-seed, making amino acid sequences that are more than 70% identical highly prone to coaggregation [36].

Both sidechain-specific β-helical and generic cross-β structural elements are present in the tau protofilament formed in human AD brains (Figure 2c). The core of the paired helical filaments and straight filaments (Figure 1) are composed of eight β-sheets that run along the length of the protofilament, adopting a C-shaped architecture (Figure 2c). Each C consists of a β-helix region with three β-sheets are arranged in a triangular fashion, and two regions with a cross-β architecture, where pairs of β-sheets pack anti-parallel to each other [15**].

Importantly, the tau protofilament in human Pick’s disease brains adopts a different fold, despite having many β-structure residues in common with the AD tau protofilament, only containing generic cross-β structural elements [20**]. The discovery that tau residues 337–357 can adopt both a canonical β-helix structure in an Alzheimer’s disease brain and cross-β packing in Pick’s disease...
highlights the molecular polymorphism [38] in generating disease-specific folds in the ordered cores of tau filaments. It is clear that a combination of molecular and assembly polymorphism is responsible for the conformational diversity of fibril polymorphs across all proteinopathies. Different amyloid fibril polymorphs have distinct physical properties, such as surface hydrophobicity [13], kinetics of self-assembly [39], reversibility [40], and thermodynamic stability [41]. Thus, the sequential, structural and physical determinants of fibril polymorphs lead to a variety of behaviours in vivo, possibly resulting in a broad spectrum of disease phenotypes [42].

**Amyloid aggregation in cells**

Protein aggregation into amyloid fibrils has been extensively studied by biophysical experiments in vitro. Fibril formation begins with a slow nucleation phase, followed by fibril elongation, and is greatly accelerated by seeding with preformed fragments. In the physiological setting, however, many factors modify the course and outcome of aggregation, notably the protein quality control machinery for fibril fragmentation and disaggregation [43]. Deposition of aggregates in cells can take many forms, and is typically characterised by the observation of fluorescent foci. EM of these structures reveals a wide range of assemblies, interacting in different ways with other cellular components.

**An emerging physiological pathway for protein segregation**

In budding yeast, the [PSY+] prion forming protein, sup35, consists of a folded, translation termination factor preceded by a low complexity N-terminal, prion domain and a charged, unstructured middle domain. A recent study shows that stress conditions such as low pH trigger liquid phase separation driven by the prion domain, which reversibly recruits the molecule into liquid droplets. After the stress is relieved, it is released into its viable form [44**]. Deletion of the prion domain results in irreversible aggregation of the folded domain upon stress. Another recent study, in a mammalian system, shows that reversible aggregation of the RNA-binding protein TDP-43 functions in muscle regeneration [45**]. With the findings of physiological roles for liquid phase separation, the first hints are beginning to emerge about

---

**Figure 3**

Aggregation pathway of Huntingtin exon 1 in cells and in vitro. (a) Huntingtin exon 1 aggregates (43Q) in the course of conversion from reversible, liquid phase droplets to irreversible fibrous aggregates in a HEK293 cell (reproduced from Ref. [48]). (b,c) The in vitro counterpart of this conversion, showing the start and end points of conversion from liquid droplets to fibrous outgrowth, which consumes all the fluorescence from the original droplets [48]. The in vitro experiment was done with 25Q, with its phase separation induced by addition of dextran, because aggregation occurs too rapidly to follow with pathogenic glutamine repeat lengths.
why such potentially damaging domains have been retained in evolution. However, adaptive phase separation may come at a cost, as liquid-like compartments may have the capacity to convert to irreversible solids or promote amyloid aggregation in disease. Several proteins are known to phase separate into liquid-like assemblies that mature into solid aggregates \textit{in vitro}, particularly when the proteins carry disease associated mutations [46,47]. In the case of Huntingtin exon1 \textit{in vitro} and as expressed in mammalian and yeast models, the protein can separate into a reversible liquid phase, which promotes its conversion into irreversible amyloid aggregates (Figure 3) [48*].

**Interactions of aggregating species with cellular components**

A model of [PSI+] prion formation and transmission, in which the translation terminal factor domain is replaced by YFP and the fluorescent construct overexpressed, shows a regimented arrangement of short bundles of fibrils [49]. Modulation of the Hsp70 chaperone network revealed that the disaggregation systems of Hsp104, a AAA+ protein involved in extraction of proteins from large aggregates, and Hsp110, a nucleotide exchange factor for Hsp70 that confers disaggregation activity on the Hsp70 system, are both involved in determining the structure of the prion deposits. Hsp104 is required for [PSI+] propagation, but its overexpression converts the fibrils into amorphous, but still irreversible, aggregates. The level of Hsp110 appears to set the fibril length. \textit{In vivo}, amyloid fibrils tend to be much shorter than the ones grown \textit{in vitro}, likely due to the actions of the chaperone systems.

Cryo electron tomography (ET) explores the nanostructures of fibrillar deposits and their interactions with other components in cells. The difficulty is to gain access to the appropriate cellular regions—the aggregates tend to be in the thinner parts of cells, and some form of sectioning is needed to produce layers of vitrified cell structure containing the aggregates but thin enough for transmission EM, that is, well under a micron in thickness. Cryo-sectioning provides a view of the structures in very thin sections (e.g. Ref. [49]) but it is mechanically damaging and restricts the section thickness to little more than 50 nm, giving a very limited window into the cell. Although it is useful for correlative fluorescence and EM, and for verifying features seen by the easier method of freeze substitution (see Stahlberg chapter), it does not provide an undisturbed, high resolution view of cell structure. Focussed ion beam (FIB) milling offers a way to produce relatively undisturbed layers of suitable thickness for electron tomography, and two recent studies show Huntingtin exon 1 and poly GA aggregates in mammalian cell models by cryo-ET of FIB milled cells [50*,51*]. The huntingtin inclusions consist of jumbled, short fibrils in either nuclear or cytoplasmic aggregates. In the cytoplasm, there is ER recruited around the edges of the inclusions, interacting with fibril ends and distorted in a way similar to that previously reported with vesicles and short amyloid fibrils \textit{in vitro} (Figure 4) [52]. In the case of poly GA, a Gly-Ala repeat found in an amyloidogenic protein in ALS and frontotemporal dementia, irregular twisted fibrils interact with and appear to stall proteasomes [51*]. These studies suggest mechanisms by which the aggregates might cause damage to cells, but the initial causes of pathology and damage remain obscure.

---

**Figure 4**

Huntingtin aggregates in a cell model. (a) Tomogram section and (b) corresponding rendered view of a cryo-tomogram of a Huntingtin exon 1 aggregate from a FIB-milled lamella of a Hela cell (reproduced from Ref. [50*]). Fibrils, blue; ribosomes, green, ER, red; vesicles, white. Scale bar, 400 nm.
Disaggregation of amyloid fibrils by the Hsp70 system

In bacteria, yeast and plants, the AAA+ Hsp100 chaperones cooperate with the Hsp70 system to extract protein from aggregates or from prion fibrils, and return them to the unfolded state. This activity involves threading of the protein chain through the unfoldase/disaggregase Hsp104 or ClpB in yeast or bacteria respectively. However, the Hsp100 family is not found in the cytosol of metazoa. Disaggregation activity is provided by the Hsp70 system itself, with specific forms of the Hsp40 and Hsp110 partner proteins. It has been demonstrated for two in vitro amyloid systems, α-synuclein and Huntingtin exon 1, that a specific combination of Hsp70-40-110 components can depolymerise the fibrils and return the subunits into solution as monomers [53,54]. In vivo studies have demonstrated that fibril deposition is inversely related to the cellular levels of these chaperones [43].

Concluding remarks

In summary, recent advances in single-particle cryo-EM have enabled the structure determination of amyloid fibrils from post-mortem human brain tissue. These findings demonstrate that cryo-EM allows near-atomic characterization of amyloid filaments from patient-derived material, paving the way for the investigation of a range of human pathologies. Current evidence suggests that each neurodegenerative disease has its own unique amyloid fibril structural signature that can be recognized by cryo-EM. As more clinically derived structures emerge, we should begin to understand the hypothesized link between protein aggregates and clinical phenotype, enabling a targeted approach to structure-based drug design and the development of biomarkers for early detection of diseases. In the physiological context, electron tomography can provide high resolution structural information within intact regions of cells to investigate the interplay between amyloid assemblies and other cellular components.

Conflict of interest statement

Nothing declared.

Acknowledgements

We thank Tom Peskett for helpful comments on the manuscript. This work was supported by Wellcome Trust grant 106252/Z/14/Z to H.R.S and N.I.H. grant U01NS110438 to A.W.P.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest


The first near-atomic structure of amyloid fibrils extracted from post-mortem brains of Alzheimer’s disease patients. Two filament structures are found, differing in twist and in inter-protofibril contacts.


The atomic structure of an alpha-synuclein fibril shows an extended interface with hydrophobic contacts between the two protofibrils in a polar fibril structure. Many of the mutations in familial Parkinson’s disease are involved in stabilising the fibril structure.


The atomic structure of an Aβ fibril by cryo EM and solid-state NMR reveals a compact pair of protofibrils with an intertwined interface burying both termini in a polar fibril.


The cryo-EM structure of amyloid filaments of tau from a patient with Pick’s disease shows, surprisingly, that they have a different protofila-
mament fold from the ones found in Alzheimer’s disease. The differences can be related to a different number of microtubule binding repeats and different phosphorylation between the two forms.


Fibril structures of β2-microglobulin reveal significant coincidence of β-strands between the native immunoglobulin fold and the amyloid state. The interprotofilament interface is formed by α-stacking, a feature not previously reported.


Stress-induced segregation of the yeast prion sup35 is driven by its prion domain, which sequesters the protein in a reversible, phase separated state. If the prion domain is deleted, the catalytic domain aggregates irresponsibly in response to stress, suggesting a protective, physiological role for the prion domain.


A physiological role for phase separation is demonstrated by the transient formation of granules of the RNA-binding protein in regenerating muscle fibres. These myo-granules can seed amyloid fibril formation by TDP-43, whose aggregation is associated with amyotrophic lateral sclerosis.


Reversible, liquid phase aggregates of Huntingtin exon 1 is observed in the process of conversion to irreversible fibrillar aggregates, in cell models as well in vitro. The cellular observations were done by correlative fluorescence and electron tomography, combined with video microscopy.


Cryo electron tomography of FIB-milled cells expressing pathogenic Huntington exon 1 reveals fibrillar, cytoplasmic aggregates appearing to recruit, and sometimes distort, vesicles of the ER.
Cryo electron tomography of cellular aggregates of fibrils formed by a poly GA repeat involved in ALS reveal recruitment and stalling of proteasomes, suggesting a potential pathway of toxicity in this cell model.

