Showcasing research from Dr Sinnige's laboratory, Department of Chemistry, Utrecht University, The Netherlands.

Molecular mechanisms of amyloid formation in living systems

The molecular mechanisms of amyloid formation have been studied extensively in test tube reactions, but it remains unclear how these translate to the complex environment of living cells and organisms. Model organisms such as C. elegans can be used to bridge the gap between in vitro studies and human disease. The image shows C. elegans expressing an aggregation-prone polyglutamine protein in the muscle cells, and a structure of amyloid-beta fibrils derived from Alzheimer’s patient material (pdb 7q4m). Harold van de Kamp is acknowledged for the graphic design.
Molecular mechanisms of amyloid formation in living systems

Tessa Sinnige

Fibrillar protein aggregation is a hallmark of a variety of human diseases. Examples include the deposition of amyloid-β and tau in Alzheimer’s disease, and that of α-synuclein in Parkinson’s disease. The molecular mechanisms by which soluble proteins form amyloid fibrils have been extensively studied in the test tube. These investigations have revealed the microscopic steps underlying amyloid formation, and the role of factors such as chaperones that modulate these processes. This perspective explores the question to what extent the mechanisms of amyloid formation elucidated in vitro apply to human disease. The answer is not yet clear, and may differ depending on the protein and the associated disease. Nevertheless, there are striking qualitative similarities between the aggregation behaviour of proteins in vitro and the development of the related diseases. Limited quantitative data obtained in model organisms such as Caenorhabditis elegans support the notion that aggregation mechanisms in vivo can be interpreted using the same biophysical principles established in vitro. These results may however be biased by the high overexpression levels typically used in animal models of protein aggregation diseases. Molecular chaperones have been found to suppress protein aggregation in animal models, but their mechanisms of action have not yet been quantitatively analysed. Several mechanisms are proposed by which the decline of protein quality control with organismal age, but also the intrinsic nature of the aggregation process may contribute to the kinetics of protein aggregation observed in human disease.

1. Introduction

More than fifty human diseases are characterised by the deposition of fibrillar protein aggregates.1–3 These include neurodegenerative disorders such as Alzheimer’s, Parkinson’s, and Huntington’s diseases. Proteins also aggregate in organs other than the brain, for example in systemic amyloidosis and type II diabetes. Early investigations into the nature of the deposits found in disease tissue showed that they could be stained with iodine, leading to the name ‘amyloid’ derived from the Greek word for starch.4 The different proteins that aggregate in human diseases are generally unrelated in terms of sequence or native structure (see Table 1 for the main proteins discussed in this article). However, they have in common that in disease they fold into a fibrillar structure with a cross-β architecture, referred to as ‘amyloid fibril’.

The relationship between amyloid formation and toxicity is still unclear, leading some to question whether protein aggregation plays a causal role in human disease.4 Amyloids can also have functional roles, for example in bacterial biofilm formation5 and the storage of peptide hormones.6 However, a large body of evidence suggests that the process of pathogenic fibril formation contributes to cellular dysfunction and death.

Toxicity associated with amyloid formation has been attributed to intermediate or off-pathway oligomeric species, which may engage in a variety of aberrant interactions and thereby disrupt cellular pathways.7–9 The molecular features of oligomers remain largely elusive, given their heterogeneous and – in the case of intermediate species – transient nature. The process of fibril growth may in addition cause significant perturbation to cellular structures such as membranes.10,11 Mature fibrils are considered to be relatively inert, but they have also been shown to engage in non-physiological interactions and to sequester essential cellular components.12,13

The precise relationship between amyloid formation and toxicity is difficult to establish without knowledge of the molecular mechanisms of protein aggregation in a cellular context. Great progress has been made in elucidating the molecular mechanisms of fibril formation using test tube experiments with purified proteins and peptides, where detailed kinetic analysis has quantitatively revealed the aggregation processes of e.g. amyloid-β14,15 and α-synuclein.16,17 However, it is not clear to what extent these mechanisms are relevant for the situation in vivo. Even though the biophysical properties of a protein remain fundamentally the same, its aggregation behaviour may be modulated in living cells and organisms, for example by the presence of molecular chaperones. It is not clear whether these factors merely change the rates of the different aggregation steps, or whether the aggregation mechanisms are
more extensively remodelled. This perspective explores the question to what extent the mechanisms elucidated in vitro apply to human disease, by combining our knowledge on the biophysics of amyloid formation with the genetics of human disease and information from model organisms.

2. Lessons learnt from in vitro studies

In vitro studies of amyloid formation using isolated peptides and proteins have the advantage that the intrinsic aggregation behaviour is observed, and that the conditions can be carefully controlled. Even so, the process of amyloid formation in itself is rather complex, comprising several microscopic steps (Fig. 1A) that are not immediately evident from the macroscopic aggregation data. The nature and contributions of the microscopic steps can be elucidated by means of kinetic analysis, based on global fitting of experimental data to mathematical models describing the aggregation processes. The experimental data are typically based on Thioflavin T fluorescence or other measurements that provide a read-out for the fibril mass, monitored for a range of protein concentrations over time.

### Table 1 Characteristics of the main disease-associated proteins discussed in this perspective

<table>
<thead>
<tr>
<th>Protein</th>
<th>Disease</th>
<th>Isoforms and post-translational modifications</th>
<th>Aggregation behaviour in vitro</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amyloid-β</td>
<td>Alzheimer’s disease</td>
<td>Cleavage product of APP, different lengths including 1–38, 1–40, 1–42, 1–43</td>
<td>Spontaneous under physiological conditions; dominated by secondary nucleation</td>
<td>15 and 16</td>
</tr>
<tr>
<td>Tau</td>
<td>Alzheimer’s disease and other tauopathies</td>
<td>6 isoforms, heavily phosphorylated in disease</td>
<td>Promoted by the addition of polyanions; dominated by secondary nucleation and/or fragmentation</td>
<td>19 and 38</td>
</tr>
<tr>
<td>α-Synuclein</td>
<td>Parkinson’s disease and other synucleinopathies</td>
<td>C-terminal cleavage and phosphorylation of S129 and other sites in disease</td>
<td>Slow intrinsic aggregation; promoted by the addition of negatively charged lipid bilayers (primary nucleation), low pH (secondary nucleation) or shaking (fragmentation)</td>
<td>17, 18 and 20–22</td>
</tr>
<tr>
<td>Huntingtin ( htt) fragments with expanded polyQ</td>
<td>Huntington’s disease&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Aberrant splicing and/or cleavage give rise to N-terminal fragments including exon 1</td>
<td>Simple polyQ stretches and htt exon 1 undergo spontaneous aggregation under physiological conditions; dominated by secondary nucleation; strong dependence on polyQ length</td>
<td>23–25</td>
</tr>
<tr>
<td>Islet amyloid polypeptide (IAPP)</td>
<td>Type II diabetes</td>
<td>Mature IAPP is produced from a precursor protein and secreted</td>
<td>Spontaneous aggregation; accelerated by membranes containing anionic lipids</td>
<td>26–30</td>
</tr>
<tr>
<td>Superoxide dismutase 1 (SOD1)</td>
<td>Amyotrophic lateral sclerosis (ALS)</td>
<td>Mature protein is a disulfide-crosslinked dimer with copper and zinc ions bound</td>
<td>The apoSOD1 monomer forms fibrils via an unfolded intermediate; fragmentation-driven mechanism under agitation</td>
<td>31–33</td>
</tr>
<tr>
<td>Prion protein (PrP)</td>
<td>Creutzfeldt-Jakob disease</td>
<td>Glycosylated and lipid-anchored</td>
<td>Natively folded PrP&lt;sup&gt;b&lt;/sup&gt; is converted into PrP&lt;sup&gt;c&lt;/sup&gt; through a fragmentation-based mechanism</td>
<td>22</td>
</tr>
</tbody>
</table>

---

<sup>a</sup> 9 polyQ expansion diseases are known, which are each associated with polyQ stretches in a different protein.

### 2.1 Mechanisms of amyloid formation established in vitro

The formation of fibrils from purely monomeric protein depends on an initial conformational change, referred to as primary nucleation. The critical nucleus can be seen as a transition state onto which the addition of monomers is energetically favourable. Nucleation combined with elongation represents the simplest mechanism of amyloid formation starting from monomeric protein, and is characterised by a sigmoidal aggregation curve (Fig. 1B). A more strongly sigmoidal curve with a long lag phase and a rapid increase towards the plateau is indicative of secondary processes, which accelerate aggregation as a function of the concentration of the fibril mass already present (Fig. 1C). Secondary processes include fibril fragmentation, leading to the generation of new elongation-competent fibril ends, and secondary nucleation, the catalysis of nucleation on the surface of existing fibrils (Fig. 1A). The lag phase can be shortened by adding pre-formed fibril seeds, bypassing the need for primary nucleation (Fig. 1C).

Secondary nucleation has been found to dominate the overall process of fibril formation for many disease-associated...
peptides and proteins \textit{in vitro}, including amyloid-\(\beta\),\(^{15}\) a fragment of tau,\(^{27}\) huntingtin exon 1 with expanded polyglutamine,\(^{25}\) and islet amyloid polypeptide (IAPP)\(^{30,38}\) (Table 1). Single-molecule studies on a full length tau isoform, \(\alpha\)-synuclein and mouse prion protein highlighted fragmentation as the dominant mechanism producing new fibrils (Table 1).\(^{19,22}\) Given that fragmentation is independent of the monomer concentration, it may play a more significant role relative to secondary nucleation at the low protein concentrations expected \textit{in vivo}.\(^{20}\) Secondary processes can explain aggregate spreading across tissues by the transfer of fibril seeds, and may contribute to the long lag phase of aggregation observed in human disease, as further discussed in Section 6.

2.2 Experimental conditions affect aggregation behaviour

It is likely that the biological environment changes the rates of the different aggregation steps to varying extents, thereby shifting the mechanism to a different dominant process. Even \textit{in vitro}, experimental conditions greatly affect the aggregation mechanism. A good example is presented by \(\alpha\)-synuclein (Table 1), which accumulates in Lewy Bodies in Parkinson’s disease.\(^{39}\) Under quiescent conditions and physiological pH, \(\alpha\)-synuclein fibrillization occurs very slowly, taking weeks until completion at 37 °C.\(^{20}\) The aggregation process can be accelerated by vigorous shaking,\(^{21}\) leading to fragmentation being the main driver of fibril formation. The addition of purely anionic lipid vesicles strongly promotes aggregation by accelerating primary nucleation.\(^{18}\) Other lipid compositions and higher lipid ratios were found to have the opposite effect, however.\(^{40}\) At pH values below 6.0, \(\alpha\)-synuclein aggregation occurs readily due to a major increase in the rate of secondary nucleation.\(^{17}\) \(\alpha\)-Synuclein engages in functional interactions with synaptic vesicles \textit{in vivo},\(^{41,42}\) and is thought to traffic via acidic cellular organelles such as endosomes.\(^{43-45}\) Whether aggregation is promoted and which of the mechanisms is dominant may thus strongly depend on the intracellular localisation of the protein.

The confinement of proteins inside the cell by itself has major implications for the aggregation mechanisms, and these small volume effects have also been explored \textit{in vitro}. Nucleation is a stochastic process, yet in bulk experiments comprising many molecules this behaviour is not observable. Encapsulation of insulin into nanodroplets, however, has revealed highly stochastic behaviour, whereby aggregation is initiated in different droplets at different points in time.\(^{46}\) Once aggregation has started, it spreads rapidly across the droplet, suggesting that stochastic nucleation events are followed by secondary processes that lead to the propagation of fibril formation within the droplet. The volume range in which
stochastic nucleation is noticeable (fl-PL) encompasses that of most living cells, and this behaviour has indeed been observed in cultured cells as well as a multicellular animal model as further discussed in Section 4.

Liquid–liquid phase separation has recently emerged as a way to further control protein localisation inside cells. Given that the local protein concentration within the phase-separated compartment is orders of magnitude higher than in the dilute surroundings, this phenomenon can greatly increase aggregation rates. Liquid–liquid phase separation has functional roles, but has also been shown to lead to the aggregation of a range of disease-associated proteins in vitro, including FUS, α-synuclein, and tau. In the case of α-synuclein, the aggregation of phase-separated protein was found to be independent of the bulk monomer concentration, showing that nucleation is initiated within the droplets in which the concentration remains constant. Translating this behaviour to the situation in vivo, phase-separation may thus induce aggregation even when the overall concentration of a particular protein would be too low to lead to significant aggregation in a dispersed state.

3. Biophysical principles of aggregation play a role in human disease

Despite the multitude of factors that may affect aggregation pathways and rates in a biological environment, key aspects of the behaviour of aggregation-prone proteins as established in vitro appear to be preserved in vivo. Genetic forms of protein aggregation diseases are often associated with increased protein levels or the production of more aggregation-prone variants, as detailed below. When comparing different diseases, the trend emerges that proteins that are more intrinsically aggregation-prone in vitro (Table 1) cause a more robust pathology in vivo. Altogether, aggregation propensity appears to be an important factor for the onset and severity of human disease, even though the precise relationship is not yet clear and there are notable exceptions.

3.1 Concentration dependence of amyloid formation in human disease

The aggregation kinetics of amyloid formation in vitro are dependent on the starting concentration of monomeric protein. The identification of gene multiplications in familial forms of Alzheimer’s and Parkinson’s diseases suggests that this may also be the case in humans. An early-onset familial form of Parkinson’s disease with dementia was found to arise from a triplication of the SNCA gene encoding α-synuclein, resulting in an approximate doubling of mRNA and soluble protein levels. Gene duplication of SNCA is associated with a less aggressive, late-onset form of the disease, suggesting a direct link between α-synuclein levels and disease progression.

The APP gene encoding the amyloid precursor protein is located on chromosome 21, and people with Down syndrome who carry an additional copy of this chromosome are at high risk to develop early-onset Alzheimer’s disease. Duplication of the APP gene itself has also been associated with a rare early-onset form of the disease. Other familial forms of Alzheimer’s disease are linked to the cleavage of APP by the γ-secretase complex, releasing amyloid-β. Typically, these mutations result in higher levels of the total pool of amyloid-β peptides, or specifically of the more aggregation-prone 1–42 isofrom compared to 1–40. Sporadic Alzheimer’s disease may in fact also be linked to increased amyloid-β levels, given that the main genetic risk factor, APOE, is involved in its clearance.

3.2 Disease mutations affect aggregation propensity

Disease mutations in the part of APP that encodes amyloid-β generally enhance the aggregation kinetics of the peptide in vitro, in particular due to an increased contribution of secondary nucleation. Mutations in tau are typically not found in Alzheimer’s disease, but are associated with the development of frontotemporal dementia and other tauopathies. Certain variants such as the P301L mutation have been shown to aggregate faster than wild-type under a range of experimental conditions and to induce a strong phenotype in animal models. Other mutations seem to have a stronger effect on microtubule binding, which is the physiological role of tau. In a cellular context, this may lead to an increased concentration of free tau, which also has the effect of promoting protein aggregation. Yet other mutations have been found to affect the relative abundance of the different tau isoforms, but their respective aggregation propensities remain to be determined.

When looking at amino acid substitutions in α-synuclein that underly familial forms of Parkinson’s disease, the picture is rather diverse. Kinetic analysis in vitro revealed that the A53T mutant protein has increased rates of both anionic lipid-catalysed aggregation and secondary processes occurring at mildly acidic pH. For A30P only the latter were accelerated, whereas the E46K, G51D and H50Q mutations had either unaffected or lower aggregation rates compared to wild-type α-synuclein. These mutants might display different aggregation behaviour in the human brain compared to the test tube. For example, A30P has a reduced affinity for neutral membranes, which may lead to an increase in free protein thereby promoting aggregation, similarly as described above for tau. It is also conceivable that α-synuclein fibril formation is not the factor driving these familial disease forms. In support of this notion, certain genetic forms of Parkinson’s disease with mutations in parkin or LRKK2 present cases without Lewy Body pathology.

Perhaps the most direct evidence for a correlation between protein aggregation and disease is provided by Huntington’s and other polyglutamine (polyQ) expansion diseases, which all have a genetic cause. The expansion of CAG repeats leads to mutant proteins with an extended polyQ tract, causing intracellular protein inclusions and disease symptoms when the pathogenic threshold is crossed. The longer the polyQ stretch, the faster the aggregation kinetics in vitro, and the earlier disease symptoms occur. In the case of Huntington’s...
disease, the typical polyQ threshold of ~40 residues causes
disease in mid-life, whereas expansions of over 60 residues are
associated with juvenile Huntington’s disease affecting patients
younger than 21 years old.92,93

The peptides and proteins considered above are largely
intrinsically disordered, although structural segments may
be formed upon physiological interactions, e.g. in the case of tau
binding to microtubules. Proteins with a globular fold are also
capable of forming amyloid fibrils, which occurs via partial
unfolding. In these cases, destabilisation of the folded state by
mutations typically promotes amyloid formation and disease.
This has been shown for SOD1, which aggregates in familial
forms of ALS (Table 1),95 but also for proteins associated with
systemic amyloidosis including lysozyme,96 β-2 microglobulin,97
and transthyretin.98 Drugs that stabilise the native transthyretin
tetramer slow the progression of disease forms associated with
mutant as well as wild-type proteins, providing strong evidence
for a causal link between protein aggregation and the disease.99

3.3 Correlation between aggregation propensities in vitro
and in disease

It is difficult to compare the aggregation kinetics of different
proteins in vitro, for which the experiments have been carried
out at varying conditions, and even more so in vivo, where
limited quantitative information is available. Still, it is interesting
to qualitatively consider how the aggregation propensities
in vitro correlate with the onset and extent of aggregation in
different diseases (Table 1). For example, amyloid-β 1–42 is so
aggregation-prone that it is difficult to handle in test tube
experiments. Amyloid plaques appear decades before the onset
of symptoms in Alzheimer’s disease, and are even found in
cognitively normal aged individuals.90,91 Thus, it is tempting to
speculate that amyloid-β aggregation in vivo happens sponta-
aneously once sufficiently high concentrations are reached (see
Sections 3.1 and 6.2).

A similar argument can be made for polyQ diseases. Simple
peptides comprising as few as 15 glutamine residues spontane-
ously aggregate into fibrils in the test tube.92 A physical
explanation for the higher threshold for polyQ aggregation
observed in human disease is the transition from a tetrameric
or dimeric nucleus for shorter polyQ lengths, towards a mono-
ermic nucleus at longer polyQ lengths.93 The longer polyQ
peptides are able to form a β-hairpin structure within a single
chain, accelerating fibril formation by orders of magnitude. The
aggregation kinetics of polyQ proteins are modulated by the
flanking regions,94,95 and consistently the threshold is some-
what different between the different polyQ diseases.96 The
diseases are not fully penetrant at intermediate polyQ lengths
(e.g. 35–39 for Huntington’s disease), suggesting that in this
regime the cellular environment has a major influence on the
aggregation process. At polyQ lengths above the pathogenic
threshold (40 for Huntington’s disease), the disease will
however invariably manifest, accompanied by the characteristic
protein inclusions.

Contrastingly, wild-type tau is much less prone to sponta-
neous aggregation in vitro, and most assays include negatively
charged molecules such as heparin to drive tau self-association.
In humans, the protein occurs in six alternatively spliced iso-
forms, and in Alzheimer’s brain it is heavily phosphorylated.96,97
The effects of the amino acid sequence and posttranslational
modifications on the aggregation propensity of tau are not yet
clear, given that most in vitro studies have been performed using
short and unmodified tau fragments. However, it appears that
the aggregation of wild-type tau in Alzheimer’s disease requires
certain biological triggers, as evidenced by the fact that it occurs
downstream of amyloid-β deposition.98 Most of the mouse
models for Alzheimer’s disease based on APP and γ-secretase
mutations fail to develop tau tangles in spite of extensive amyloid
plaque formation, and the relationship between the aggregation
of the two proteins remains to be determined.99

Wild-type α-synuclein is also less intrinsically aggregation-
prone than for example amyloid-β, and in vitro aggregation
assays typically rely on the addition of seeds, anionic lipid vesi-
cles or a shift to acidic pH as discussed in Section 2.2. In cultured
cells, α-synuclein is highly soluble and found as a disordered
monomer99 or a β-helical tetramer.100 Intriguingly, Lewy Bodies
in human brain tissue have been suggested to be a medley of
cellular organelles and membranes, in the apparent absence of
α-synuclein fibrils.101 However, α-synuclein aggregation and
neuronal dysfunction can be robustly induced by introducing
pre-formed fibrils in cellular and mouse models,102–107 leading to
the hypothesis that α-synuclein pathology in the human brain
develops through the spreading of fibril seeds. Seeded aggrega-
tion is highly efficient in vitro by bypassing the need for primary
nucleation (Fig. 1C), and in this way aggregation could proceed in vivo
even for proteins with very low intrinsic aggregation
propensity. The hypothesis of aggregation via the cell-to-cell
transfer of seeds in combination with secondary processes will
be further discussed in Section 6.3.

4. In vivo studies on the kinetics of
amyloid formation

A more quantitative and mechanistic comparison of protein
aggregation in vitro and in vivo is more challenging to achieve,
although some progress has been made in analysing protein
aggregation in humans and animal models. The development
of PET tracers has made it possible to investigate protein
aggregation in the human brain in a longitudinal fashion,108,109
although at limited resolution. High-resolution cryoEM struc-
tures of fibrils extracted from the human brain,110–112 on the
other hand, are based on the post-mortem state and lack kinetic
information. As such, mechanistic studies largely depend on
the use of animal models, and even simple models can be
appropriate to address the fundamental aspects of protein
aggregation in vivo. A limited number of studies on the kinetics
of amyloid formation in living organisms are available, as dis-
cussed in the following.

4.1 Aggregation kinetics in the nematode C. elegans

The nematode Caenorhabditis elegans is a relatively simple
multicellular animal model, composed of 959 somatic cells for
the hermaphrodite adult. These cells comprise different tissues including the muscles, the intestine, and a basic nervous system. The main advantages of using *C. elegans* to study protein aggregation mechanisms are its fast reproductive cycle of ca. 3 days and short lifespan of 2–3 weeks, and its optical transparency, which enables the visualisation of protein aggregation in live animals.

In *C. elegans* models that were developed to identify genetic modifiers of aggregation and toxicity, the disease-associated proteins are typically fluorescently labelled and overexpressed in the large muscle cells for clear visualisation and correlation with a motility phenotype.\(^\text{116,117}\) As such, these models are relatively artificial, yet they provide the opportunity to characterise the behaviour of disease-related proteins in a living and ageing animal. For example, using fluorescence lifetime imaging (FLIM), the aggregation kinetics and the nature of the inclusions formed by polyQ and \(\alpha\)-synuclein were shown to be clearly distinct (Fig. 2). Whereas polyQ aggregates rapidly into amyloid-like inclusions, \(\alpha\)-synuclein forms a long-lived intermediate state that appears liquid-like.\(^\text{55,118}\) Given that the proteins were studied in the same biological environment, these differences can be attributed to their intrinsic biophysical properties. The ability of \(\alpha\)-synuclein to undergo liquid–liquid phase separation, after which the droplets mature into amyloid-like aggregates, has indeed also been shown *in vitro*.\(^\text{55,57}\)

In the case of polyQ aggregation, the experimental tractability of the *C. elegans* system allowed for a quantitative analysis similar to that performed *in vitro* (Fig. 3).\(^\text{49}\) Fitting the kinetics for multiple strains expressing the protein at different levels in the muscle cells revealed that aggregation is governed by independent nucleation events in individual cells (Fig. 3B). Cell-to-cell spreading or other non-cell-autonomous effects do not noticeably contribute to the aggregation kinetics in this *C. elegans* model (Fig. 3C). Instead, the muscle cells essentially operate as ‘individual test tubes’, similarly as found in a kinetic study using cultured neuronal cells.\(^\text{48}\) Moreover, the nucleation rate is effectively constant across the investigated timespan, and contributions arising from ageing processes are not detectable. Future studies are required to confirm whether this spontaneous aggregation behaviour holds true for more disease-relevant protein constructs, cell types and expression levels.

### 4.2 Aggregation kinetics in flies, mice, and the human brain

It is more challenging to obtain high-quality experimental data in more complex organisms such as fruit flies and mice, yet a number of noteworthy studies have provided mechanistic insights into aggregation pathways and rates in these systems. In *Drosophila*, a striking correlation was found between the aggregation propensity of amyloid-\(\beta\) mutants and the neurotoxicity assessed by longevity and motor function.\(^\text{119}\) A study on mice expressing SOD1 mutants found in ALS reported that survival times were directly correlated with the levels of unfolded protein.\(^\text{120}\) In the case of this globular protein, unfolding is required for amyloid formation to occur, and decreased protein stability is expected to lead to faster aggregation. The kinetic profiles obtained in the mice were similar to those *in vitro* and consistent with a fragmentation-based mechanism.

Also in the case of a mouse model of prion disease, the scaling of the aggregation kinetics with the protein levels in different strains pointed towards a fragmentation-dominated mechanism.\(^\text{131}\) The aggregation rates were found to be orders of magnitude lower than determined for the prion protein (PrP) *in vitro*. It is perhaps expected that aggregation rates are lower *in vivo* given the presence of molecular chaperones, as discussed in Section 5. However, analysis of tau P301S aggregation in mice yielded very similar rates as *in vitro*.\(^\text{122}\) The overexpression as well as the highly aggregation-prone mutation may favour the spontaneous aggregation of this protein in mice, although chaperones should remain effective against protein aggregation under these circumstances, as detailed in the next section.

The experimental possibilities to elucidate protein aggregation mechanisms in the human brain are evidently more limited compared to model organisms. Nevertheless, the progression of tau aggregation in Alzheimer’s brain was recently quantitatively analysed based on different datasets, including antibody staining, tau seeding assays and longitudinal PET data.\(^\text{123}\) The determined rates for tau aggregation in the human brain were orders of magnitude lower than those *in vitro*, which could again be attributed to the presence of molecular chaperones.

Notably, the recent cryo-EM structures of tau, \(\alpha\)-synuclein and amyloid-\(\beta\) fibrils extracted from patient brain material do not typically match those formed by the recombinant proteins *in vitro*.\(^\text{114,115,121}\) Although these structures represent kinetic end points, it can be inferred that the aggregation processes must differ. The structure of amyloid-\(\beta\) fibrils purified from a mouse model does correspond to that of the main fibril type identified in familial Alzheimer’s disease.\(^\text{115}\) Thus, further studies into the aggregation mechanisms in animal models have the potential to bridge the gap with *in vitro* experiments and provide insights relevant to amyloid formation in human disease.

---

**Fig. 2** FLIM shows the difference in kinetics and the nature of the inclusions formed by polyQ versus \(\alpha\)-synuclein. Shown are the head regions of *C. elegans* with expression in the muscle cells of YFP (control), 40 glutamines tagged with YFP (Q40) and \(\alpha\)-synuclein tagged with YFP (\(\alpha\)-syn). Amyloid formation is accompanied by a reduction in the fluorescence lifetime (blue–green colours). Q40 forms amyloid-like inclusions directly, whereas \(\alpha\)-synuclein accumulates in inclusions that are not amyloid-like until the worms have reached old age. Based on ref. 118.
5. Role of molecular chaperones in controlling amyloid formation

Although many proteins spontaneously fold into their native structure when diluted into buffer from a denatured state, efficient protein folding inside the cell is dependent on the presence of molecular chaperones. Various classes of chaperones and co-chaperones exist (see e.g. ref. 124–128 for reviews) and many of these have been shown to inhibit amyloid formation in test tube reactions (Fig. 4). Different microscopic steps of the aggregation process may be affected, depending on the species the molecular chaperone interacts with. Affinity for the monomeric protein reduces several aggregation steps including primary nucleation, whereas binding to the fibril surface is a potent way to block secondary nucleation, and binding to fibril ends specifically interferes with elongation.129

5.1 Molecular chaperones inhibit amyloid formation in model organisms

Many studies have shown that chaperones also inhibit protein aggregation in cultured cells and animal models (see ref. 130 for a more extensive review), although mechanistic investigations in vivo are limited. A notable exception is presented by a study on the J-protein DNAJB6, which inhibits polyQ aggregation in vitro, in cells and in mouse models.25 Kinetic analysis of the in vitro data revealed that DNAJB6 predominately inhibits primary nucleation. Consistently with this mechanism, the chaperone prevented polyQ aggregation in cultured cells, but failed to do so in the presence of added fibril seeds. Aggregation kinetics are difficult to monitor in mouse models (see Section 4), but at the examined time points the numbers of aggregates in mice overexpressing DNAJB6 were lowered, and the onset of symptoms was delayed. These data suggest that indeed the early aggregation events are inhibited by DNAJB6 in vivo.

The importance of the chaperone network in preventing protein aggregation and toxicity is evident from the results of genetic screens in C. elegans models, which have implicated major chaperone classes such as Hsp70, Hsp90, TRiC (the eukaryotic chaperonin), and several co-chaperones (Fig. 4).131,132 It is striking that these screens are based on strains overexpressing aggregation-prone proteins such as polyQ, which might be expected to overwhelm the available chaperone capacity. However, the knock-down of the identified chaperones exacerbates protein aggregation compared to control conditions, meaning that these chaperones are effective even at low stoichiometric ratios. The recently developed kinetic analysis of protein aggregation in C. elegans discussed above may prove useful to dissect and quantify the contributions of individual chaperones in this model system.133

Chaperones typically act to prevent protein aggregation and not to promote it. The knock-down of a number of co-chaperones was found, perhaps surprisingly, to decrease polyQ aggregation in a genetic screen in C. elegans.133 This result was suggested to arise from negative regulation of chaperone function by these co-chaperones, so that their knock-down actually increases protein folding capacity. The Hsp90 co-chaperone Aha1 was shown to promote tau aggregation when the two components were added to in vitro assays, and the accumulation of insoluble tau was found to be increased in mice overexpressing Aha1.134 These data indeed provide evidence that co-chaperones may modulate the interaction with chaperones in a manner that reduces their anti-aggregation effects both in vitro and in vivo.

5.2 Disaggregation, remodelling and degradation

Another intriguing way in which chaperones may effectively promote aggregation, is by disassembling existing fibrils into seeding-competent species (Fig. 4B). A metazoan disaggregation machinery consisting of a Hsp70, a Hsp110 nucleotide exchange factor and a specific J-protein co-chaperone was shown to break down preformed α-synuclein fibrils in vitro via a combination of fragmentation and depolymerisation, finally releasing non-toxic monomers.135,136 Other studies came to slightly different conclusions, suggesting an “all-or-none” mechanism of protofilament unzipping followed by depolymerisation,137 or even simpler the direct release of monomers...
Molecular chaperones preventing amyloid formation. (A) Representative structures of different classes of chaperones: Hsp70 (Escherichia coli DnaK, PDB 2KHO), Hsp90 (human Hsp90α, ref. 153), J-protein (Thermus Thermophilus DnaJ dimer, PDB 4J80), chaperonin (human TRiC/CCT complex, PDB 6NRA), small heat shock protein (sHsp; human Hsp27 oligomer, PDB 6DV5). (B) Schematic of the aggregation pathway with the action of molecular chaperones, derived from a combination of in vitro and in vivo data as described in the main text.

Fig. 4

from fibril ends.138 However, knocking down the disaggregation component Hsp110 in C. elegans reduced foci formation of both a-synuclein and polyQ, suggesting that in vivo the disaggregation machinery promotes aggregation by creating seeding-competent species.139 The spreading of a-synuclein to neighbouring tissues was also reduced by the knock-down of Hsp110 in C. elegans,139 implying that the disaggregation machinery may drive an aggregation mechanism of fragmentation coupled to the cell-to-cell transmission of seeds. It will be interesting to also analyse the effects of the disaggregation machinery in a quantitative manner in vivo to establish its contribution to the overall aggregation process.

Altogether, chaperones are likely to affect the rates of individual aggregation processes in vivo and may thus shift the dominant mechanism by which fibrils arise, compared to the situation in vitro. Another way of remodelling the aggregation pathway by changing the morphology of the final aggregate species (Fig. 4B). This possibility has not yet been widely investigated, but studies in vitro have shown that a small heat shock protein reroutes the aggregation pathway of amyloid-β 1–40 towards amorphous, rather than fibrillar, aggregates.140 Along the same lines, overexpression of the co-chaperone Sis1 in yeast induces the cloud-like condensation of polyQ, rather than the typical compact fibrillar inclusions, also suggesting a change in aggregate morphology.141 Studies on the remodelling of aggregate structures in more complex organisms are lacking thus far, although small heat shock proteins have been suggested to promote aggregation into less toxic species in ageing C. elegans.142 Whether chaperones contribute to the diversity of fibril morphologies as observed e.g. across different tauopathies110–115 remains to be determined. Recent advances in cryo-electron tomography will be particularly helpful to reveal the molecular nature of aggregates under the influence of different chaperone levels in single cells12,107,143–144 or potentially in small model organisms such as the nematode.145,146

Perhaps the most effective way for chaperones to deal with aggregation-prone proteins is to target them for degradation (Fig. 4B). Hsp70 and Hsp90 cooperate with the co-chaperone CHIP, which is an E3 ubiquitin ligase that mediates proteasomal degradation.147 Increasing the levels of Hsp70 in a tauopathy mouse model decreases tau aggregation by reducing its levels, which was shown to be accomplished via CHIP.148 Blocking the folding capacity of Hsp90 similarly reroutes phosphorylated tau for degradation in cultured cells and in mice.149 The C. elegans homolog of CHIP was not identified in the genetic screens mentioned above, which were carried out in polyQ and amyloid-β expressing animals, and it will be interesting to shed more light on the substrate specificity of this system. Chaperones are also connected with the other main degradation pathway next to the ubiquitin-proteasome system, which is autophagy. The Hsp70 co-chaperone BAG3 promotes the autophagy of ubiquitinated substrates when proteasomal degradation is impaired, and together with the small heat shock protein HSPB8 reduces the levels of expanded huntingtin in cultured cells.150 Promoting degradation may in general be a good strategy to prevent protein aggregation in disease by...
lowering the levels of the soluble protein, as well as clearing aggregates (see e.g. ref. 151 and 152 for further reviews on this topic).

6. What determines the kinetics of aggregation in human disease?

The late onset of sporadic forms of protein aggregation diseases such as Alzheimer's and Parkinson's is often attributed to the decline in protein homeostasis (proteostasis) with age. However, the accumulation of fibril mass also intrinsically takes time to develop, as obvious from in vitro experiments (Fig. 1B and C). The question thus emerges to what extent the time course of amyloid-like protein aggregation in vivo is determined by the decline in proteostasis during ageing, versus the kinetics of spontaneous amyloid formation as observed in vitro. In other words, is protein aggregation a function of biological ageing, or of age as simply “the passing of time”? The typical course of human protein aggregation diseases with a long lag phase and an onset of aggregation and symptoms in mid-to-late life can be explained by several reasons, both from the point of a decline in proteostasis as well as from the intrinsic kinetics of amyloid formation. Some possible scenarios will be discussed below. These are by no means exhaustive, and different combinations may be at play depending on the protein and the associated disease.

6.1 Proteostasis decline during ageing

Proteostasis encompasses the pathways that collectively ensure that proteins are functional, including biogenesis, folding, translocation and degradation. The use of C. elegans has been instrumental in demonstrating the fact that proteostasis capacity is limited. In a hallmark study, it was shown that the expression of metastable proteins alongside polyQ results in premature aggregation and loss of function. Metastable proteins furthermore start to aggregate and lose their function over the course of adulthood of the animal, suggesting that proteostasis capacity declines with age. Similar discoveries were recently made in the mouse brain. Widespread changes in the proteome have also been found to take place during the reproductive phase of C. elegans, followed by the aggregation of many endogenous proteins at a later stage. The ability to mount robust stress responses such as the heat shock response deteriorates even earlier in the life cycle of the nematode, at the onset of the reproductive phase. Although the decline of proteostasis with age encompasses changes in a variety of pathways, the focus here is on chaperone function and degradation, since these most notably modulate the propensity for amyloid formation.

It is important to note that amyloid formation and amorphous aggregation are two distinct processes. The metastable proteins that have been used as proteostasis sensors in animal models are natively globular proteins that precipitate when they fail to reach their native state, due to the exposure of hydrophobic patches and loss of solubility. Many amyloid-forming proteins such as tau and α-synuclein are however intrinsically disordered, and relatively soluble in their unfolded native states. In contrast to amorphous aggregation, amyloid formation is a highly ordered process, which is limited by an energetically unfavourable nucleation event. As for every phase transition, amyloid-like aggregation only occurs at levels above the protein’s solubility, here defined as the critical protein concentration. Beyond this threshold, fibril elongation by monomer addition takes place spontaneously, so that aggregation will invariably proceed once nucleation has occurred. It should be noted that once the critical concentration has been surpassed, it still takes time for the fibril mass to build up via the mechanisms described in Section 2 (Fig. 1), in contrast to amorphous precipitation which takes place almost immediately.

6.2 Aggregation mechanisms driven by a decline in proteostasis

Considering this framework, a first possibility explaining late-onset protein aggregation is that a decline in chaperone capacity lowers the critical concentration for amyloid formation, causing the disease-associated protein to cross this barrier (Fig. 5A). A reduction of the levels of (active) chaperones that transiently interact with amyloid-prone proteins would increase the levels of protein free to undergo amyloid aggregation, effectively lowering its critical concentration. Changes in chaperone levels during ageing in C. elegans are relatively modest, but some J-proteins are notably reduced. In brain tissue of Alzheimer’s patients, most chaperones also remain unchanged, but Hsp90 paralogs are selectively down-regulated. For specific client proteins, such disturbances may be sufficient to exceed their solubility limits. Computational predictions have revealed that many disease-associated proteins are expressed at levels close to their solubility, meaning that subtle perturbations may be sufficient to cross this threshold.

Another way to cross the critical concentration is an increase in the level of the amyloid-prone protein (Fig. 5B). This may be either due to decreased clearance or increased production, the latter also including the generation of cleavage products or aggregation-prone post-translationally modified species. The genetics behind Alzheimer’s disease support several of these points, given that familial mutants lead to the generation of more and/or longer amyloid-β peptides, and the main genetic risk factor for sporadic disease is associated with impaired clearance. Proteins may also accumulate in a particular subcellular compartment, as exemplified by FUS and TDP-43 for which the mislocalisation from the nucleus to the cytoplasm is thought to be one of the triggers for aggregation and disease. In the cytoplasm, these proteins undergo liquid–liquid phase separation and subsequently aggregate, whereas the high RNA concentrations in the nucleus are thought to prevent their phase separation and aggregation.

A decline in proteostasis may also contribute to the kinetic profile observed in human disease in a more subtle manner. Perhaps, certain proteins are already present above their critical concentrations, but the aggregation rates are too low to lead to significant problems and/or can be dealt with by the degradation machinery. However, changes in the chaperone network may
lead to increased aggregation propensities, for example by promoting nucleation or by enhancing fibril fragmentation (Fig. 5C). Such effects may in particular lead to strongly sigmoidal curves in the case of a positive feedback loop, whereby protein aggregation induces further disruption of proteostasis. The critical concentration hypothesis is somewhat more appealing because it precludes any incidence of aggregation prior to crossing the barrier (Fig. 5A and B). Given that amyloid formation can proceed quickly once nucleation has occurred, especially when secondary processes are at play, the possibility of the cell to degrade these aggregates may be limited.

As a variation to the mechanism described above, small molecules such as metabolites or lipids that can act as co-factors for nucleation may also accumulate during ageing and hence effectively increase the nucleation rate constant (Fig. 5C). Unfortunately, not much is known about factors that directly promote aggregation in vivo, although it is likely that nucleation is heterogeneous in the complex biological environment and involves interplay with other molecules and cellular structures. Interestingly, a cryoEM structure of tau fibrils extracted from brain tissue of chronic traumatic encephalopathy (CTE) patients contains additional density in a hydrophobic cavity, which may correspond to a sterol or a fatty acid. The build-up of a co-factor or a catalyst may be a particularly important trigger for the aggregation of proteins that do not have a high intrinsic aggregation propensity, such as tau, which also requires additives to undergo aggregation in vitro. Whereas CTE has a clear external cause, namely repetitive head impact, other neurodegenerative diseases are also associated with lifestyle factors and changes in metabolism, and it will be interesting to further explore this possibility.

6.3 Aggregation mechanisms driven by secondary processes

Alternatively, the long lag phase observed in human disease can also be attributed solely to an aggregation mechanism that is dominated by secondary processes, which by itself leads to a strongly sigmoidal aggregation curve (Fig. 5D, see Fig. 1C and 3C). Both the rates of fragmentation and secondary nucleation increase with increasing fibril mass, and in particular when the monomer concentrations are low and/or primary nucleation is very slow compared to the secondary process, these mechanisms induce a long lag phase followed by a rapid increase in fibril formation. In the case of intracellular aggregation as observed for example in the case of tau, α-synuclein and polyQ proteins, these mechanisms would have to involve cell-to-cell spreading of seeds for aggregation to propagate across the tissue.

The spreading of protein aggregation following neuronal connectivity was already suggested from the Braak stages of pathology in Alzheimer’s and Parkinson’s diseases. Over the past years, a wide range of evidence for aggregate spreading of tau, α-synuclein and other proteins has been gathered in a range of experimental systems (see ref. 169–173 for reviews). A mechanism dominated by secondary processes combined with spreading is an attractive explanation for the kinetics observed in human disease, although much remains to be uncovered. In particular, it is not clear to what extent fragmentation and secondary nucleation occur in vivo, with respect to in vitro measurements. The fact that disease brains typically contain only one or two structural polymorphs suggests that the fibril structure is preserved throughout the process of spreading and replication, yet this can be the consequence of either one of the secondary processes. The rate of fragmentation is fundamentally independent of the concentration of monomeric protein, perhaps rendering this process somewhat more likely as the main mechanism to generate seeds compared to secondary nucleation at low physiological protein concentrations. The disaggregate machinery can further promote the breakage of fibrils (see Section 5.2), although the contribution of this activity to the overall aggregation process is not yet clear.
In the case of PrP as well as SOD1, kinetic analysis in mice (Section 4.2) pointed towards fragmentation as the dominant mechanism. Direct evidence for the occurrence of secondary nucleation in vivo is not yet available. The fibril structures of e.g. tau from patient brains have revealed a fuzzy coat formed by disordered regions of the protein surrounding the amyloid core, but it is not clear whether this would interfere with or enhance the absorption of monomers followed by their nucleation. Other interactors, including molecular chaperones, may also decorate the surface of fibrils making them unavailable for monomer interactions. Fibrils are moreover typically found in dense clusters in vivo. This may limit secondary nucleation on the one hand, but on the other hand may also be a consequence of exactly this process. Further quantitative analysis of (seeded) aggregation in model organisms will be critical to resolve these outstanding questions.

7. Conclusions and outlook

The biophysical principles driving amyloid formation in vitro fundamentally hold true in vivo, despite the complexity of the cellular and organismal environment. However, each of the microscopic processes, for example primary nucleation or fibril elongation, is modulated by the surroundings. Chaperones have been shown to affect the rates of several microscopic aggregation steps in vitro, but it is not yet clear to what extent these mechanisms remodel the aggregation pathways in vivo. Physical constraints such as cellular confinement and phase separation furthermore lead to different processes being dominant, compared to bulk in vitro studies. Intracellular protein aggregation is strongly limited by primary nucleation, which needs to occur in every individual cell in the absence of aggregate spreading. The spreading of seeds combined with secondary nucleation or fragmentation is likely to dominate the overall aggregation process for many cells in a tissue to be affected when primary nucleation rates are low. Phase separation, on the other hand, efficiently promotes primary nucleation due to the high local protein concentration within the phase-separated state.

The answer to what extent the in vitro mechanisms are relevant to human disorders depends on the specifics of the protein and the associated disease. For highly aggregation-prone protein variants, such as long polyQ proteins or aggregative point mutants, the aggregation process may be more ‘spontaneous’ and resemble the in vitro mechanism. The case of juvenile Huntington’s diseases exemplifies the fact that protein aggregation diseases are not necessarily age-related, and can manifest at an age where protein quality control is expected to be intact. In these cases, the lag phase observed in disease could be due to the fact that secondary processes dominate the aggregation mechanism, but an increase in protein levels over time may also contribute. In contrast, for proteins that are less intrinsically aggregation-prone, such as wild-type α-synuclein or tau, cellular handling including localisation and post-translational modifications may trigger the onset of aggregation.

A problem with elucidating the contributions of spontaneous aggregation versus changes in the biological environment is that most animal models used to study protein aggregation diseases, including the C. elegans models discussed in this perspective as well as most mouse models, are based on high levels of overexpression of the disease-associated protein. These models typically display aggregation starting from early adulthood or even during development (reviewed in ref. 174). It will be important to investigate aggregation mechanisms in animal models expressing more modest protein levels. In such models, it will be possible to quantify the critical concentration for aggregation and to elucidate which aspects of ageing may trigger amyloid formation. In addition, obtaining more quantitative data about the concentrations of disease-associated proteins in humans at different time points would be extremely helpful. Altogether, bridging the gap between biophysics and the biology of ageing and disease is an exciting avenue that will hopefully contribute to a molecular understanding of the mechanisms underlying protein aggregation diseases, and the development of novel therapeutics.

Author contributions

T. S. wrote the article.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

I would like to thank Stefan Rüdiger and the members of my lab for critical reading of the manuscript, and Utrecht University for a start-up grant.

References


59 M. Farrer, J. Kachergus, L. Forno, S. Lincoln, D. S. Wang, M. Hulihan, D. Maraganeor, K. Gwinn-Hardy, Z. Wszolek,


84 B. P. C. Van De Warrenburg, H. Hendriks, A. Dürr, M. C. A. Van Zuijlen, G. Stavenin, A. Camuzat, R. J. Sinke, A. Brice and B. P. H. Kremer, Age at onset variance...


