Non-monotonic fibril surface occlusion by GFP tags from coarse-grained molecular simulations

Julian C. Shillcock [a,b,*], Janna Hastings [b,c], Nathan Riguet [b], Hilal A. Lashuel [b,*]

[a] Blue Brain Project, Ecole polytechnique fédérale de Lausanne, CH-1202 Geneva, Switzerland
[b] Laboratory of Molecular and Chemical Biology of Neurodegeneration, Ecole polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland
[c] Bioinformatics Competence Center, Ecole polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland

A B S T R A C T

The pathological growth of amyloid fibrils in neurons underlies the progression of neurodegenerative diseases including Alzheimer's and Parkinson's disease. Fibrils form when soluble monomers oligomerise in the cytoplasm. Their subsequent growth occurs via nucleated polymerization mechanisms involving the free ends of the fibrils augmented by secondary nucleation of new oligomers at their surface. Amyloid fibrils possess a complex interactome with diffusing cytoplasmic proteins that regulates many aspects of their growth, seeding capacity, biochemical activity and transition to pathological inclusions in diseased brains. Changes to their surface are also expected to modify their interactome, pathogenicity and spreading in the brain. Many assays visualise fibril formation, growth and inclusion formation by decorating monomeric proteins with fluorescent tags such as GFP. Recent studies from our group suggest that tags with sizes comparable to the fibril radius may modify the fibril surface accessibility and thus their PTM pattern, interactome and ability to form inclusions. Using coarse-grained molecular simulations of a single alpha synuclein fibril tagged with GFP we find that thermal fluctuations of the tags create a non-monotonic, size-dependent sieve around the fibril that perturbs its interactome with diffusing species. Our results indicate that experiments using tagged and untagged monomers to study the growth and interactome of fibrils should be compared with caution, and the confounding effects of the tags are more complex than a reduction in surface accessibility. The prevalence of fluorescent tags in amyloid fibril growth experiments suggests this has implications beyond the specific alpha synuclein fibrils we model here.

© 2021 Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Amyloid fibrils formed of misfolded proteins underlie many neurodegenerative diseases [1]. They are a component of Lewy bodies present in Parkinson's disease [2–4], amyloid plaques and neurofibrillary tangles in Alzheimer's disease and cellular inclusions formed in Huntington's disease among others [5]. The composition of these inclusions is complex and involves not only fibrils, but also lipids, membranous organelles, and other proteins [4,6]. Fibrils may interfere with a cell's homeostasis by their presence as rigid bodies in the cell and as the result of their complex interactome with proteins [7], and cellular membranes and organelles [3] or their ability to activate cell death pathways and

biochemical reactions occurring at specific surface domains or interaction hubs. Within pathological inclusions, the fibrils are typically heavily modified by enzymes that access their surface or the flexible termini of monomeric subunits that project into the surroundings (e.g., N- and C-terminal domains of alpha synuclein). Post-translational modifications also influence fibril structure and surface properties [8–11] as well as their propensity to self-associate [12].

Fibrils are nucleated from soluble monomeric proteins via two distinct mechanisms: 1) In the primary mechanism, short oligomers form in bulk solution and elongate by addition of monomers at their ends; 2) Secondary nucleation occurs when oligomers form at the surface of existing fibrils from monomers that have diffused onto them (Fig. 1). These oligomers may then form branches from the existing fibril or detach and grow by the primary mechanism into new fibrils. Reaction rate models predict that secondary nucleation is a major contributor to fibril elongation [13,14]. The fibril
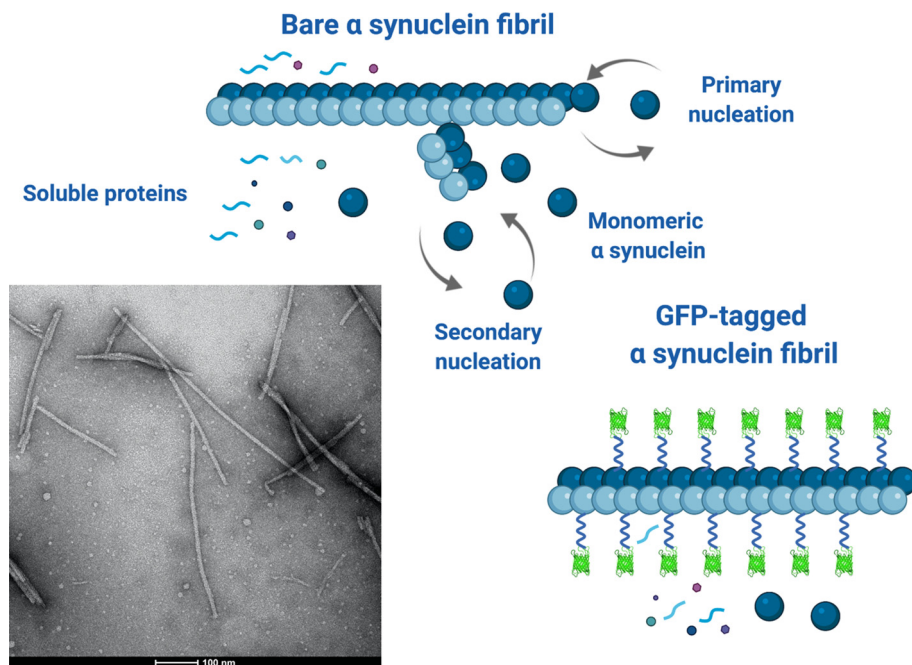

Fig. 1. Cartoon of a bare aSyn fibril illustrating the primary and secondary nucleation mechanisms by which it elongates and forms new fibrils at its surface respectively. Soluble proteins and aSyn monomers/oligomers can diffuse to, and interact with, the surface of the fibrils giving rise to a dynamic interactome. The GFP-decorated fibril is shown undergoing less surface-mediated interactions due the occluding effects of the tags that can modify the fibril's interactome. The EM image shows several bare aSyn fibrils that are straight, rigid rods on length scales below 100 nm, which is the range of the simulations. (Created with BioRender.com).

surface is also an important target for drugs to inhibit pathological amyloid fibrils as it contains multiple domains or binding sites for small molecules [15]. Access to the fibril surface is therefore a key determinant of pathological fibril growth, inclusion formation and potential therapeutic interventions. The ability of other proteins to modify or interact with the surface depends on the ease with which they can access it via diffusion. Understanding how this access is modified by experimental protocols or natural variations in fibril structure due to post translational modifications or unnatural modifications that are commonly introduced to facilitate the detection and monitoring of these proteins in cells and in vivo, is a prerequisite for a better understanding of fibril involvement in disease. Here we use molecular simulations to reveal how the accessibility of the surface of an amyloid fibril composed of the protein alpha synuclein is modified when the individual monomers are tagged by Green Fluorescent Protein (GFP).

Alpha synuclein (aSyn) is an intrinsically disordered protein (IDP) that is genetically and biochemically linked to Parkinson's disease (PD) pathology and pathogenesis [16,17]. aSyn fibrils are one of the main protein components of the pathological inclusions, called Lewy Bodies, found in the brain of PD and other neurodegenerative diseases that are collectively referred to as synucleinopathies, including Dementia with Lewy Bodies [2]. aSyn is a 140-residue sequence with disordered N and C terminal domains. Under pathological conditions it spontaneously assembles into rigid fibrils whose growth is enhanced by interactions of the monomeric protein with the growing ends or lateral surface of the fibrils [18]. Cryo-EM studies have shown that aSyn fibrils are composed of two protofilaments, and have diameters around 10 nm and lengths of 20–500 nm (see Fig. 1) [19]. The protofilaments contain a rigid core formed of the central part of the protein while the final $\sim$ 40 residues at the N and C termini extend in a disordered manner into the surrounding fluid. The C-terminal region is highly negatively charged, harbors most of the disease-associated post-translational modifications (PTM) of the protein and represents an exposed interactome hub. Increasing evidence suggests that

many of these PTMs occur post fibrillization and contribute to the packing of fibrils into dense aggregates within Lewy Bodies [4].

Experiments on fibril growth and interactions in cell models and *in vitro* frequently attach fluorescent labels to amyloid-forming proteins to make them visible in microscopy [20]. These labels include ligand-coated gold nanoparticles [21], with diameters in the range 2.5 – 4 nm, and GFP [22], with a linear dimension of 4 nm. GFP is typically covalently bound to fibril monomers by flexible peptide linkers containing 10–15 residues (equivalent to several nm [23]), usually to sequences that decorate the surfaces of the fibrils [24]. These sizes are comparable to the fibril diameters, which are in the range of 5–10 nm, allowing the morphology of the fibrils to be visualised. Fig. 1 shows a cartoon that illustrates our hypothesis that the linked GFP tags partially occlude the surface of the fibrils and may interfere with biochemical reactions at its surface, including secondary nucleation and interactions with soluble proteins or other molecular species. The cryo-EM image in Fig. 1 shows that the fibrils are typically rigid straight rods on length scales below $\sim$ 100 nm. We note here that the disordered termini of the aSyn protofilaments also protrude from the fibril surface and may also modify its accessibility, but this is not shown in Fig. 1 for clarity and we do not address their effects in this work.

The process by which a diffusing molecular species interacts at a fibril surface consists of two sequential steps: 1) the molecules approach the fibril by diffusion; 2) closely-apposed molecules undergo conformational fluctuations leading to a complex, time-dependent interaction with the surface. The presence of the GFP tag interferes with the first process because it hinders diffusion to the fibril surface. Whether it affects the second process is a complex function of the diffusing particle's size and conformational ensemble, and the GFP linker length. We do not attempt to address the full growth process here, but use coarse-grained molecular simulations to study the first stage in which untagged monomers diffuse to a fibril's surface. Computer simulations provide a powerful tool to explore the kinetics of interactions of diffusing particles, oligomers, and amyloid fibrils [25]. However, many soluble disor-

dered proteins have hydrodynamic radii of a few nanometers and require hundreds of nanoseconds to diffuse distances comparable to the diameter of aSyn fibrils [26]. Lipid vesicles and organelles are larger still, with diameters in the range of many 10 s of nanometers. This makes the use of Atomistic Molecular dynamics computationally prohibitive, although it has been used to explore conformational fluctuations of single disordered proteins [27]. Coarse-grained simulation techniques are able to follow the diffusion of particles around a fibril over the length and time scales required here as discussed in Section 4.1 of Materials and Methods. To retain near-molecular detail, we use the coarse-grained technique of dissipative particle dynamics [28–30]. A single aSyn fibril is represented in the simulations as a rigid cylinder to which GFP tags can be attached by flexible linkers (Fig. 1). The linkers represent peptide chains with sizes in the range 2 – 4 nm (∼7 – 15 residues), but we do not assume a specific sequence beyond its flexibility, which typically requires glycine, proline or serine residues [23]. The diffusing monomers are modelled as spheres and ellipsoids with a hydrodynamic radius comparable to those of monomeric aSyn [26]. Although these proteins have a large conformational ensemble in solution [31], their experimental diffusion is frequently modelled by spheres of an equivalent hydrodynamic radius [26,32]. We do not attempt to capture atomistic details of these particles, which allows them to represent any diffusing object of equivalent hydrodynamic radius. These could be aSyn monomers or other proteins known to interact with its surface, small lipid micelles or membranous clusters.

Our results show that fluctuating GFP tags significantly reduce the residence time of diffusing particles at the fibril surface, and the magnitude of the effect has a complex dependence on the particle size and linker length. A non-monotonic variation in the residence time is found for particles whose hydrodynamic radius is 1–2 nm as the GFP linker length increases. Although the GFP tags hinder the approach of small particles to the fibril surface, once they have reached it they stay longer in its vicinity because the tags retard their diffusion away. Finally, we caution that experiments investigating the growth, interactome, and toxicity of GFP-labelled amyloid fibrils should be interpreted with care to eliminate the artifacts arising from the complex occlusion of the fibril's surface by the tags.

2. Results

2.1. How do particle size and linker length influence fibril surface accessibility?

The physical dimensions of GFP tags are comparable to those of the hydrodynamic radius of monomeric intrinsically-disordered proteins such as aSyn. Our first aim is to explore how the presence of tags on a pre-formed fibril affects the ability of diffusing particles to access the fibril surface. We use the coarse-grained simulation technique of dissipative particle dynamics (Materials and Methods, Section 4.1) to follow the diffusion of small nanoparticles around the model aSyn fibril and measure the amount of time they spend near its surface. Cryo-EM data from Guerrero-Ferreira et al. [19] is used to set the diameter of the fibril to 10 nm and the in-register monolayers have thickness 0.5 nm representing paired protofilaments. The fibril is circular because the atomic structure of its surface is not observable at the resolution of the coarse-grained simulations (∼1 nm). A single fibril is preassembled in the centre of the $40 \times 40 \times 30$ nm^3 simulation box with its long axis oriented along the Z axis of the box (Fig. 2). The GFP tag (when present) is represented as a rigid cylinder of dimensions 4×3 nm attached by a flexible linker of length 2 or 4 nm to the C terminal of each monomer in the fibril [22]. The remainder of the box is filled

with solvent particles and a number of rigid diffusing nanoparticles that represent untagged, monomeric aSyn proteins, or other proteins/particles of comparable size, in dilute solution around the fibril. We use the results of Tomasso et al. to represent monomeric aSyn as spheres with a hydrodynamic radius of approximately 3 nm [26], and this is taken as the upper limit of the particles. The first two million time steps of each simulation are discarded to allow the system to reach equilibrium, and results sampled from the subsequent two million steps. Further details of the simulations are given in Section 4.2 of Materials and Methods.

Snapshots from simulations of 10 spherical particles with different radii diffusing around the fibril are shown in Fig. 2. The upper row shows the fibril decorated with GFP molecules via 2 nm flexible linkers. Panel 2a has particles of radius 1 nm, and panel 2b has particles of radius 2 nm (cp. Supplementary Movie SM1). Panel 2c (and Supplementary Movie SM2) shows particles of radius 3 nm. The wide variety of conformations accessible to GFP tags on the flexible linkers is clear, and thermal motion of the tags causes them to fluctuate over an area greater than their own dimensions depending on the linker length, an effect not apparent in the static snapshots in Fig. 2 but visible in the Supplementary Movie SM1. The bottom row of snapshots in Fig. 2 shows the same number of spherical particles of radius 1 nm (d), 2 nm (e), and 3 nm (f) diffusing around the bare fibril. Supplementary Figure S1 shows the equivalent snapshots for GFP on 4 nm linkers. Supplementary Figure S2 shows snapshots from simulations of ellipsoidal nanoparticles of similar dimensions.

Previous computational studies have shown that globular domains, including GFP, linked to model Amyloid Beta monomers by short peptides interfere with their ability to oligomerise by creating a steric hindrance between monomers [33]. The linker used in this study was a flexible Ser-Pro-Ser chain. A minimal length of 7 residues (∼2 nm) was predicted to allow fibril elongation while ensuring the GFPs did not sterically intersect. In that work, the GFP molecules bound via a linker to the fibril monomers were stationary in space. However, the flexible peptide linkers used in experimental assays allow the GFPs to fluctuate around the fibril core [23]. In the case of aSyn, previous experimental studies have shown that the fusion of GFP to the C-terminus of aSyn with linkers ranging from 6 to 13 amino acids did not interfere with the ability of the protein to form fibrils *in vitro* [34,35]. In this study, we estimate the consequences of thermal fluctuations on the occluding power of the tags as follows.

The ease with which the particles can approach the fibril surface by diffusion is intuitively expected to be harder when the GFP tags are present compared to the bare fibril. We quantify this intuition by calculating the probability for the particles to lie within successive circular shells of constant thickness (0.5 nm) as a function of their distance from the fibril with and without the decorating GFPs in independent simulations (see Figure S3 of the Supplementary Material for an illustration of the shells used in the histogram calculation). The probability is obtained from the amount of simulation time the particles spend in each shell normalised by dividing by the total simulation time, the number of particles, and the area of the shell. From the normalised probability distributions, we define a surface occlusion factor as the ratio of the amount of time the particles spend near the fibril core with the GFP tags present to the time spent near the bare fibril (Materials and Methods, Section 4.3). This quantifies the change in the equilibrium probability for the particles to be at the fibril surface.

The effects of the GFP tags on the probability of the diffusing particles approaching the fibril are quantified in Fig. 3. The top row shows results for spherical particles of radius 1 nm, and the bottom row for those of radius 2 nm. Results are not shown for 3 nm radius particles because at a concentration that generates results of reasonable accuracy, the crowding of the particles in
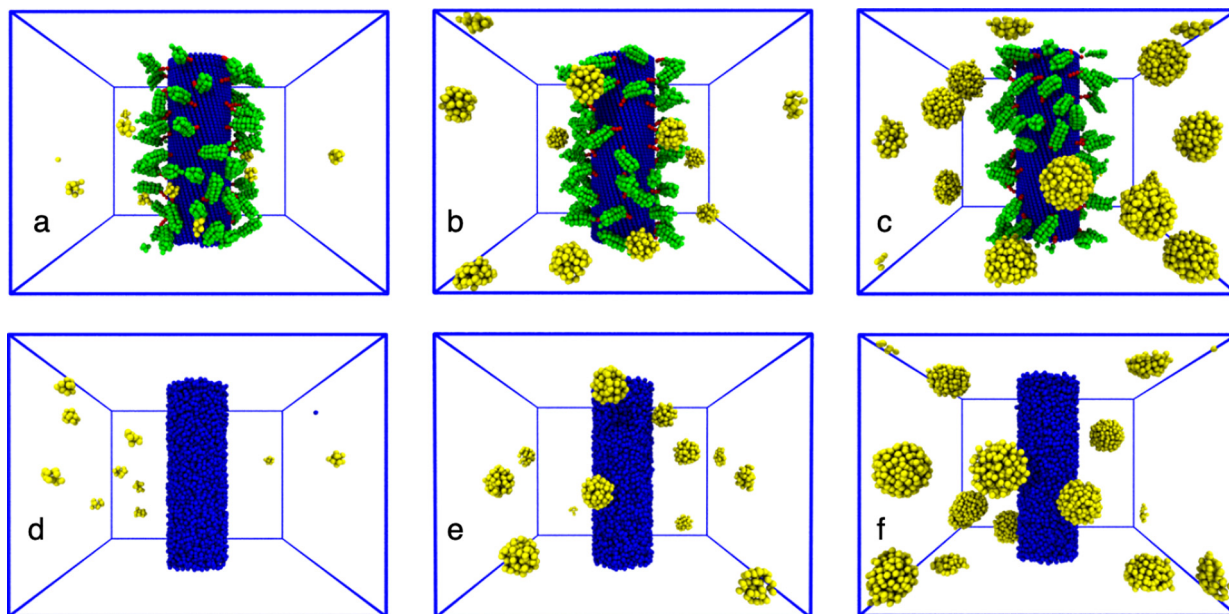
Fig. 2. Snapshots of 10 spherical nanoparticles diffusing around a 30 nm fibril of diameter 10 nm (solvent particles are invisible for clarity). Top row shows the tagged fibril with particles of radius: (**a**) 1 nm, (**b**) 2 nm, (**c**) 3 nm. The bottom row shows the bare fibril with the same number of nanoparticles used to define the baseline surface accessibility with particle radius: (**d**) 1 nm, (**e**) 2 nm, (**f**) 3 nm. Particles apparently cut by the simulation box boundary are connected via the periodic boundary conditions.
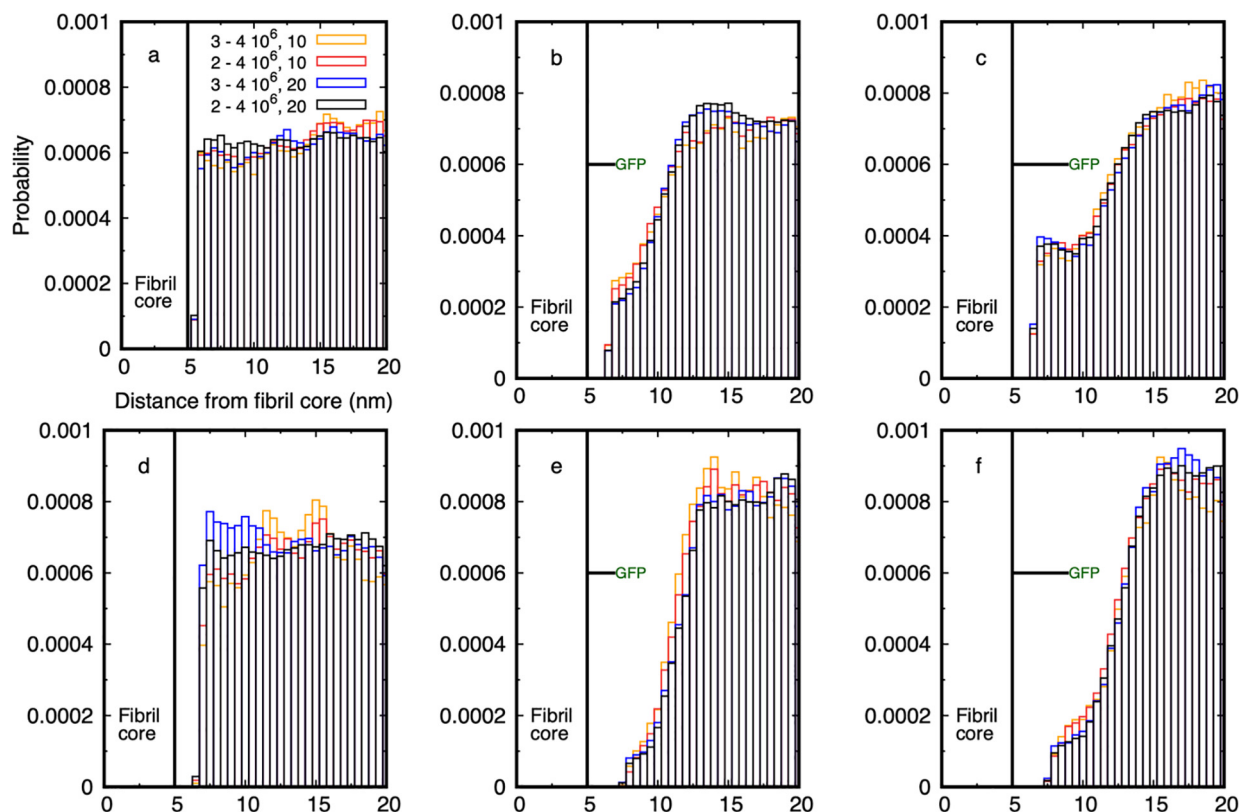

Fig. 3. Probability histograms for diffusing spheres to lie within cylindrical shells around the fibril. Four histograms are shown to illustrate the size of statistical errors. The legend and axis labels in panel a apply to all panels. The left column is for a bare fibril and the GFP linker length in the middle and right columns is indicated by the horizontal line. (**a**) The probability distribution for 1 nm radius spheres is flat around a bare fibril (cp. left column of Fig. 2). (**b**) The distribution for 1 nm spheres around a GFP-tagged fibril using a 2 nm linker shows a strong reduction in near-fibril probability. (**c**) The equivalent distribution for GFP tags on a 4 nm linker shows an enhanced probability at ~ 2 nm from the fibril surface. (**d**) The distribution for 2 nm radius spheres around a bare fibril is also flat but displaced outwards by the sphere radius (cp. middle column of Fig. 2). (**e**) The distribution for 2 nm radius spheres around a fibril decorated with GFP on a 2 nm linker shows a larger reduction in near-fibril probability compared to the 1 nm spheres in (b). (**f**) No enhancement in the near-fibril probability is seen for 2 nm radius spheres when the linker is 4 nm.

the simulation box modifies the probability distribution so that it cannot be compared with those of the smaller particles (see Supplementary Movie SM2). We expect that the displacement of the particles away from the fibril seen for 2 nm radius particles will be even stronger for larger particles. We plot the data out to 20 nm from the fibril centre, which is half the width of the simulation box (the full histograms are shown in Figure S3 of the Supplementary Material). Four histograms are shown in each panel taken from two independent simulations of 10 and 20 nanoparticles, and their variability shows the magnitude of the statistical errors of the sampling. Panel 3a shows that the probability of finding 1 nm particles around the fibril is uniform beyond the sum of the fibril and particle radii $R_{fibril} + R_{particle}$. Panel 3b shows that GFP tags attached to the fibril by 2 nm linkers strongly reduce the near-fibril probability because the particles find it harder to diffuse past the fluctuating tags to the fibril surface. However, when the linker length is increased to 4 nm (panel 3c), which is significantly larger than the particle size, there is a smaller reduction in probability out to 10 nm and a peak in the probability appears around 2 nm from the fibril surface. There is therefore a non-monotonic change of the probability for these particles to be near the fibril surface as the linker length is increased. The corresponding results for 2 nm particles are different. Panel 3d shows again a flat probability distribution for the larger particles when no GFP tags are present, and panel 3e shows an expected greater reduction in the probability compared to that in 3b because of the larger particle size and linker length of 2 nm. But although panel 3f shows a slight enhancement in the probability out to 10 nm, no peak is observed in the distribution as was observed in 3c for the smaller particles. The larger particles exhibit a monotonic decrease in their probability distribution near the fibril surface when the linker length is increased from 2 to 4 nm. Supplementary movie SM3 shows typical simulations of 20 spherical nanoparticles of radius 2 nm for the case of GFP attached by 2 nm and 4 nm linkers. The similarity of the histograms in Fig. 3 shows that the change in probability for different sized particles and linker lengths is independent of the number of particles.

2.2. Dependence of fibril surface accessibility on particle shape

The degree of occlusion of the fibril surface most likely depends on the shape as well as the size of the diffusing particles. Fig. 4 compares the probability histograms for 10 and 20 spheres and ellipsoids with similar dimensions (Supplementary Figure S4 shows the histograms for ellipsoidal particles and Supplementary Movies SM4 and SM5 show 10 spherical and ellipsoidal particles with sizes of 1 and 2 nm respectively diffusing around a fibril decorated with GFP on 2 nm linkers). Four histograms are shown for each shape taken from independent simulations of 10 and 20 particles. The distributions around the bare fibril are similar for both sizes and types of particle (panels a and d); and also for the smaller particles with both linker lengths (panels b and c). Note that the smallest ellipsoids have dimensions 1.5 × 1 nm, which is quite similar to spheres of radius 1 nm. Given the resolution of the coarse-grained particles in the simulations, we are unable to construct ellipsoids of this size with a more precise shape.

Comparing panels e and f shows that the larger ellipsoids, of dimensions 2 × 1 nm, show an enhanced probability (blue boxes are higher than red boxes) to be near the fibril for both linker lengths compared to spheres of radius 2 nm, the difference being larger for the longer linker. This indicates that the ellipsoids are able to diffuse more readily to the surface in the presence of the tags than spheres with equivalent dimensions. This result is expected geometrically as an ellipsoid whose semi-major axis is equal to the diameter of a sphere is smaller in the transverse dimensions, and can more easily diffuse between the tags.

We have shown so far that the presence of GFP tags attached to a fibril by flexible linkers of length 2 and 4 nm reduces the accessibility of its surface to diffusing particles in all cases studied compared to the bare fibril. The range of linker lengths and particle sizes examined is comparable to the hydrodynamic radius of monomeric aSyn and similar IDPs. We now quantify the surface occlusion as follows. The particles are sterically unable to penetrate the fibril and, far from the surface, are relatively unaffected by the GFP tags. For a given linker length and particle radius, there is a range over which the probability histogram is modified by the GFP tags. We integrate the probability over this range for the decorated fibril and the bare fibril and use their ratio as a measure of the occluding effect of the tags. This measure depends on both the linker length and particle radius, but is normalised to be independent of the number of particles and total simulation time (see Materials and Methods, Section 4.3).

Table 1 shows the baseline probability for particles to be within a fixed distance of the fibril surface (column 3), and the ratio of this probability integrated over the same range for both linker lengths is shown in columns 4 and 5. For spheres and ellipsoids of dimensions 1 nm, the short linker reduces the surface accessibility to 30–40% of its bare value, and the longer linker to 40–50%. This drops to 10% and 15% respectively for particles of dimension 2 nm. The values for larger spheres are too small to be significant as they are effectively completely excluded from the fibril's surface, but are shown for completeness. However, ellipsoids of semi-major axis 2 and 3 nm show a greater probability of being near the surface than the equivalent size spheres. The enhancement in the surface accessibility as a result of increasing the linker length is defined as the ratio of the two occlusion fractions, and is shown in the final column. It is clear that for spheres and ellipsoids, the enhancement can be large. The 1 nm spheres can spend more than 45% more time closer to the fibril surface when the linker is 4 nm compared to 2 nm, and ellipsoids of size 2 and 3 nm can spend up to 80% more time at the surface, albeit from a lower baseline. Note that although the final column of Table 1 shows that spheres of radius 2 nm have a greater enhancement at the surface for the longer linker than the 1 nm spheres, Fig. 3 and the third column of Table 1 show that they actually spend much less time there than the 1 nm spheres, and the apparent increase is from a lower baseline value.

It might be expected that the occluding effect of the tags would decrease with increasing linker length because of the greater free space around the fibril by which diffusing particles can approach its surface. We have tested this hypothesis by performing simulations in which the GFP tags are connected by linkers that are 10 nm long in a larger simulation box ($50 \times 50 \times 30$ nm^3). Because these simulations are computationally expensive, we present only the histograms of the probability distributions of particles around the fibril. Fig. 5 shows that the fibril surface is still significantly occluded for particles of radius 1, 2, and 4 nm even for a 10 nm linker. Although the 1 nm radius particles penetrate to the fibril surface, the probability of particles with a radius of 4 nm being closer than 8 nm of its surface is negligible. Supplementary Movie SM6 shows the particles diffusing around the tagged fibril. This result indicates that the molecular sieve still operates even when the linker length is several times larger than the GFP and particles.

3. Discussion

3.1. GFP creates a molecular-size dependent sieve around a tagged fibril

We have used coarse-grained simulations to explore how fluorescent protein tags bound to a model aSyn fibril by flexible linkers
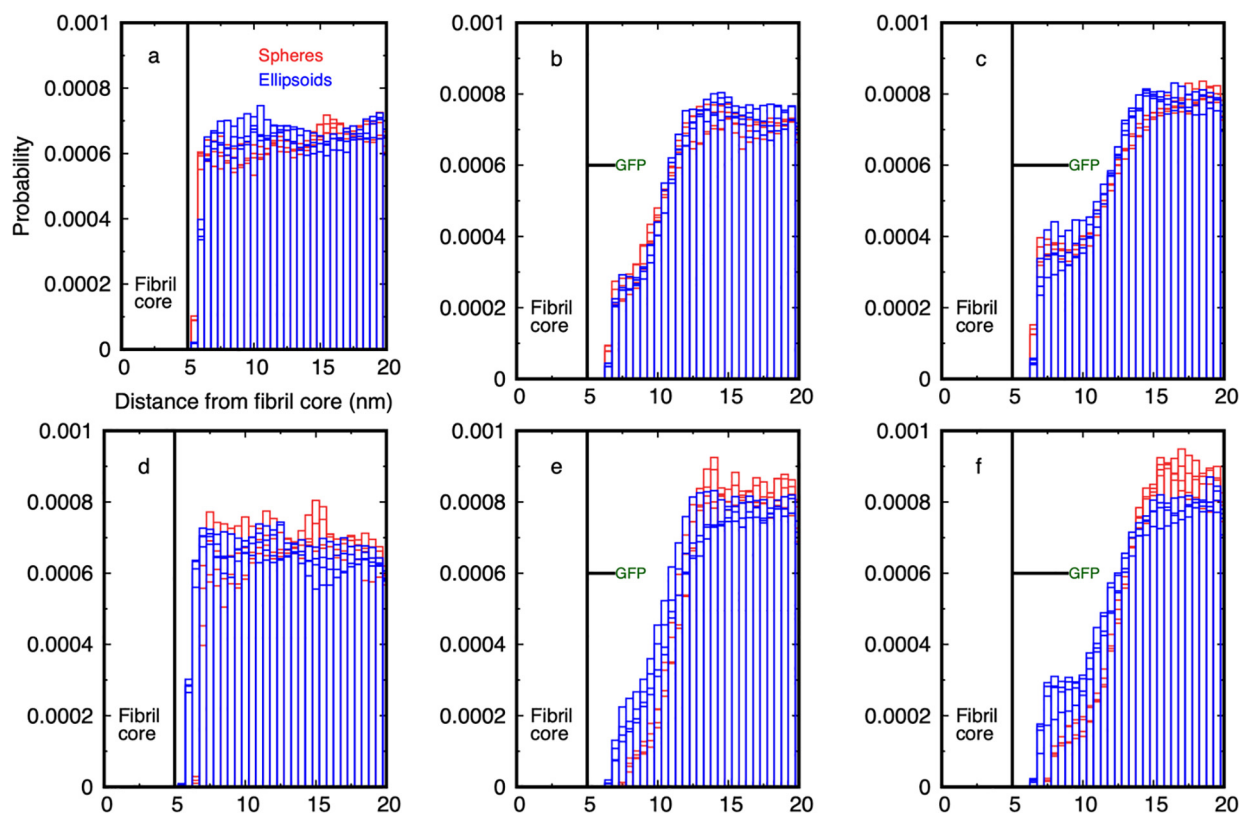
Fig. 4. Comparison of the probability for spheres (red boxes) and ellipsoids (blue boxes) to be within cylindrical shells around the fibril. The legend applies to all graphs. **(Top row)** Probability distribution for spheres of radius 1 nm and ellipsoids of size 1.5 × 1 nm around: (**a**) the bare fibril, (**b**) GFP-tagged fibril with 2 nm linker, and (**c**) GFP-tagged fibril with 4 nm linker. **(Bottom row)** Equivalent probability distributions for spheres of radius 2 nm and ellipsoids of dimension 2 × 1 nm around: (**d**) the bare fibril, (**e**) GFP on 2 nm linker, and (**f**) GFP on 4 nm linker. While the smaller ellipsoids and spheres have similar distributions, the larger ellipsoids have an enhanced probability to be near the fibril surface compared to spheres of comparable size when the GFP is attached by the 2 and 4 nm linker.

Table 1

Effect of the linker length on the fibril surface accessibility for spherical and ellipsoidal particles of different sizes. The first column gives the radius of the spherical particles and the semi-major axis of ellipsoidal particles. The fibril diameter is 10 nm in all cases. The bare fibril accessibility is the integral of the histogram for the particles to be within 4 nm of the fibril surface (9 nm from its axis, cf. Fig. 3). The occlusion ratio is the fraction of the bare fibril accessibility remaining when the linker/GFP combination is present: a value of 1 means there is no occlusion while a value of 0 means access to the surface is entirely blocked. These are values taken from independent simulations of each type. The final column shows the occlusion ratio for the 4 nm linker case divided by that for the 2 nm case, and quantifies the magnitude of the enhancement due to the larger space behind the GFP on the long linker. The occlusion ratios are calculated by integrating the radial distribution function from the fibril surface out to a given distance as described in the Supplementary Material. na = not applicable as the values are too small to be determined accurately.

Particle size/nm	Number/Type of particles (S = sphere, E = ellipsoid)	Bare fibril accessibility	Occlusion ratio (2 nm)	Occlusion ratio (4 nm)	Ratio of occlusion ratio for 4 nm / 2 nm linkers
1	10S	0.0042	0.375	0.456	1.22
	20S	0.0045	0.302	0.439	1.45
2	10S	0.0029	0.084	0.138	1.64
	20S	0.0032	0.077	0.11	1.43
3	10S	0.0023	0.0066	0.024	na
	20S	0.0022	0.0092	0.020	na
1.5	10E	0.0041	0.328	0.426	1.30
	20E	0.0046	0.281	0.297	1.06
2	10E	0.0043	0.181	0.326	1.80
	20E	0.0047	0.211	0.268	1.27
3	10E	0.0030	0.091	0.160	1.76
	20E	0.0031	0.106	0.184	1.74

occlude its surface to particles approaching by diffusion. Soluble proteins such as monomeric aSyn are treated as rigid spheres or ellipsoids that diffuse with the same hydrodynamic radius as that obtained experimentally. The accessibility of the fibril surface is always reduced by the presence of bound fluorescent tags compared to the bare fibril, as expected intuitively, and the effect is generally stronger as the particle size and linker length increase within the range of 2 – 4 nm. However, the surface occlusion is a non-monotonic function of the GFP linker length and particle size,

and the residence time of small particles at the surface is counter-intuitively *increased* by the presence of GFP tags when the particles are smaller than the linker length. The linker lengths used to attach fluorescent groups to monomers in amyloid studies are often 10 – 14 residues, which corresponds to 3–4 nm (taking the average length of a residue to be 0.3 nm) and this is the range we have studied here.

Surface occlusion occurs because thermal fluctuations of the GFP tags around the fibril surface create a steric barrier to the
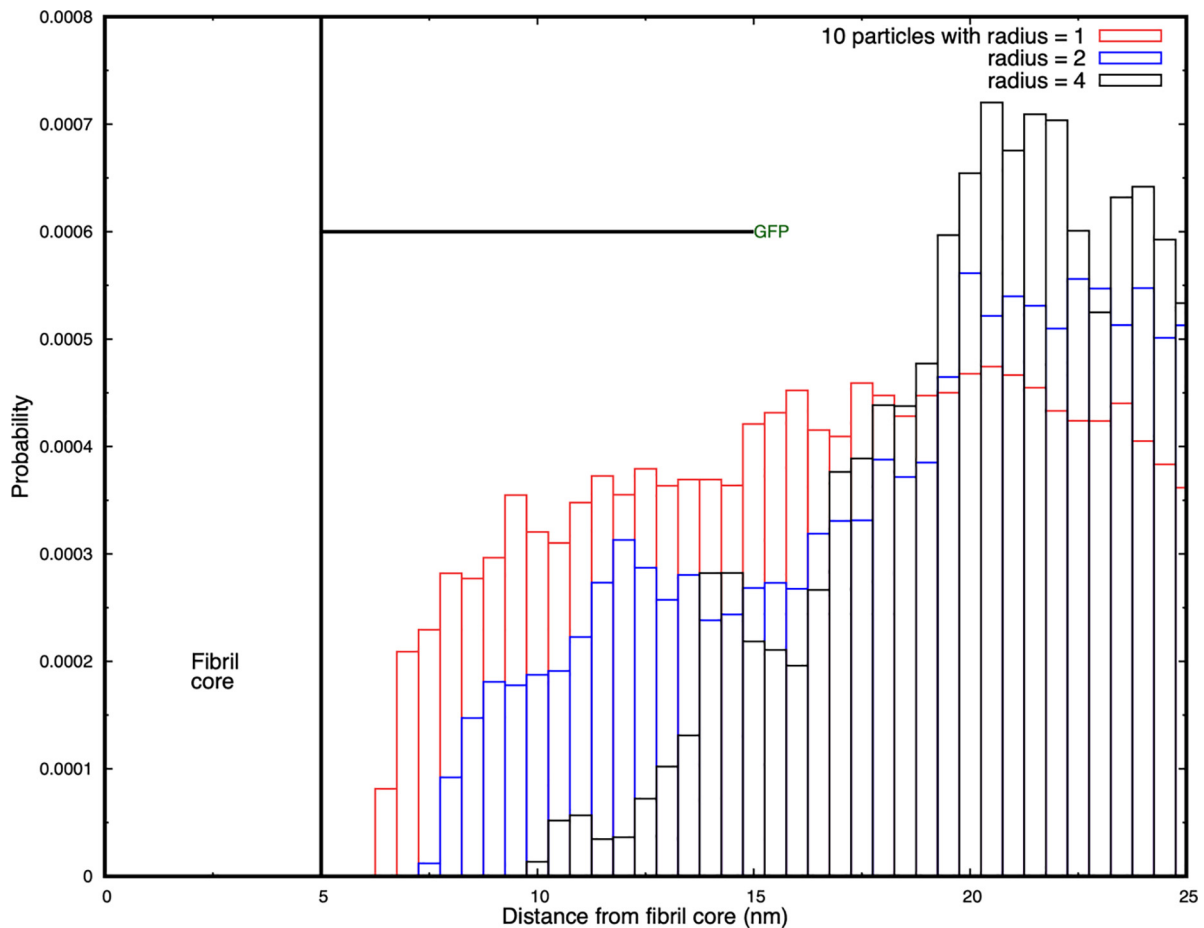
Fig. 5. Probability histograms for diffusing spheres to have their centre of mass in cylindrical shells around the fibril decorated with GFP tags on a 10 nm linker. Results are shown for 10 particles of radius 1, 2, and 4 nm, and the occlusion increases with increasing particle size.

approach of diffusing particles. We hypothesize that the increase in residence time for particles smaller than the linker length occurs because although the GFP tags interfere with the approach of the particles to the fibril, they also transiently hinder their diffusive escape. This results in a significant enhancement in their residence probability at the surface. We find that this is not an insignificant effect: spheres of radius 1 nm spend $\sim$ 50% more time at the surface when the GFP is attached by a 4 nm linker than for a 2 nm linker. Ellipsoidal particles show a similar pattern of enhanced residence at the surface. Compared to previous work in the literature in which the steric effects of static tags were explored [33], we find that thermal fluctuations of the tags gives rise to more complex occlusion of the fibril surface.

The question arises how relevant it is to approximate the steric interactions of a conformationally-fluctuating, soluble protein by those of a rigid sphere or ellipse? Molecular dynamics simulations predict that aSyn and other IDPs sample a wide range of conformations in solution [36], which suggests that it may be inaccurate to use their equivalent hydrodynamic radius to characterise their diffusion in the presence of GFP-decorated fibrils.

Although intrinsically-disordered proteins are not rigid, as are folded proteins, their hydrodynamic radius is often used to calculate their diffusion. Marsh and Forman-Kay[32] and Tomasso et al. [26] have tabulated how the hydrodynamic radius of an IDP scales with the number of residues. An aSyn monomer has 140 residues, and is predicted by Tomasso et al. to have a hydrodynamic radius of 2.7 nm. This may be compared to experimental values of 3.17 nm (pulse-field gradient NMR) [37] and 3.27 nm

(FCS) [38]. Experimental values depend on pH, the molecule becoming more compact at lower pH. The relation between the hydrodynamic diffusion of an IDP and its molecular size is therefore also of experimental interest. The steric interactions of a fluctuating polymer arise from direct contact of its monomers whose average spatial distribution is described by its radius of gyration. Dünweg et al. used Monte Carlo simulations to show that the radius of gyration of a self-avoiding polymer is 60% larger than its hydrodynamic radius [39]. This is in contrast to a uniform sphere, for which the radius of gyration and hydrodynamic radius are related by the familiar formula $R_g^2 = \frac{3}{5} R_h^2$. This implies that the steric interactions of fluctuating disordered proteins are *larger* than those of a uniform sphere of the same hydrodynamic radius. Steric interactions between soluble monomeric IDPs and fibrils in experiments are likely to be greater than those present in our simulations, which represents a lower limit to the occluding effect. Other typical fluorescent dye molecules such as rhodamines, oxazines, and fluorescein have dimensions around 0.7 – 1 nm, but can be larger when attached to proteins. These are smaller than GFP and are expected to produce a smaller occluding effect.

3.2. Implications for the mechanisms of pathological aggregate formation and toxicity

Our modeling and experimental observations suggest that the presence of GFP on the surface of the fibrils could significantly modify their surface properties, the potential for post-translational modifications, and interactions with other proteins

and organelles. Recent studies from our group and others have shown that post-aggregation PTMs on alpha synuclein fibrils play critical roles in regulating their packing, processing, and transition to Lewy body-like inclusions. Most pathology-associated PTMs (e.g., phosphorylation at S129, phosphorylation, and nitration at Y125, Y133, and Y136, and C-terminal truncations) occur on the C-terminal domain (residues 120–140) of the protein to which fluorescent proteins are usually fused. Therefore, we speculate that the presence of GFPs on the surfaces of fibrils and close to the PTM sites could interfere with the interactions between alpha-synuclein fibrils and the enzymes responsible for their modifications and other proteins that regulate their clearance [3,4,6,12,40].

We can derive several hypotheses from our results. First, secondary nucleation of aSyn filaments is predicted from reaction rate models to be a significant contributor to filament creation [13,14]. If aSyn monomers linger near the surface of GFP-tagged fibrils longer than untagged fibrils, this should increase the rate of secondary nucleation and so also the total fibril mass. This prediction could be tested by measuring the time-dependent fibril mass over time for tagged and untagged aSyn monomers.

Second, biochemical reactions and the propensity for post-translational modification at the fibril surface should be sensitive to the residence time of diffusing proteins/kinases/phosphatases. Reaction rates are necessary inputs to models of fibril nucleation and growth [13]. The rate constant of a unary reaction in which a protein interacts with the surface is predicted to drop to $\sim 45\%$ of its well-mixed value for an untagged fibril for a protein with a 1 nm hydrodynamic radius, and to 10% for 2 nm radius. A binary reaction requires two proteins to meet at the surface. But we have only the single-particle probability to be within a 2 nm radius of the fibril's surface. If we assume the probability distribution has translational symmetry along the fibril and circular symmetry around the fibril, and the diffusion of two particles to the surface are independent events, then the probability of two soluble proteins being within any small volume (of the size of the proteins) is the square of the single-particle probability scaled by the ratio of the small volume to the volume of the circular shell around the fibril. As the latter ratio is the same for the bare fibril as the tagged one, the reaction rate constant is just reduced by the square of the single-particle probability. Therefore, the rate constant will be reduced to $0.45^2 \sim 0.2$ of its untagged value for 1 nm radius proteins, and to $0.1^2 \sim 0.01$ for 2 nm radius proteins. These are large reductions even given the statistical errors in the simulations, and should be apparent in experiments.

Third, we have assumed here that every monomer in the fibril is tagged by a GFP moiety. We expect from Fig. 5 that even GFP on 10 nm linkers has a significant occluding effect. However, if there are regions of the fibril enriched in untagged monomers, this will increase the surface accessibility there and reduce the size-dependent sieve effect. This effect should vary along the fibril according to the local GFP tag density. PTM sites along the fibrils would then also be predicted to be differentially activated according to the local GFP tag density. We are not aware of any published data that measures the changes in surface mediated reactions, secondary nucleation, or PTM patterns along fibrils due to the uneven presence of GFP tags. We hope that our work here is a spur to experimentalists to measure the molecular sieve effect of GFP tags on surface reactions at fibrils.

3.3. The fusion of GFP influences the biophysical properties of amyloid fibrils

Recent studies from our group and others have shown that surface reactions play a central role in the biogenesis of pathological inclusion formation [3,4,6,12,40]. Next, we reflect on the implications of our findings on the mechanisms of amyloid formation

and toxicity based on the experimental data available in the literature today. We attempt to explain how GFP could influence the properties of amyloid aggregates and our ability to model critical processes linked to the formation and maturation of pathological inclusions associated with PD and other neurodegenerative diseases.

Fibril growth is a complex process, and GFP tags may interfere with the primary elongation mechanism or secondary nucleation at the surface, or both. It is known that fluorescent tags modify the size distribution of oligomers of the Alzheimer Aβ peptide [20], and may also modify its interactome by creating a steric hindrance to the approach of diffusing molecular species, and the lateral association between decorated fibrils, which is important for their pathological inclusion formation and maturation [41].

Several amyloid proteins have been expressed and purified as GFP fusion proteins, including amyloid-beta (Aβ, 4 kDa) [42], alpha synuclein (aSyn, 14 kDa) [34,35], Tau [43], Tau fragments and mutant Huntingtin fragments (Htt, exon1 11 KDa) [6,44,45]. Despite the fact that GFP is much bigger in size (27 kDa) compared to most of these proteins, it did not interfere with their ability to form fibrils in vitro [43], or in cells [4,6,45], except for Aβ, where the addition of GFP resulted in complete inhibition of Aβ fibril formation in vitro [42]. These observations have led to the use of GFP-fusion proteins to investigate many aspects of the kinetics and mechanisms of amyloid formation in cellular assays and animal models. The assumption in many of these studies is that GFP does not alter the biophysical properties of the final fibrils or their ability to evolve and mature to the final pathological inclusions found in the brain, amyloid plaque (Aβ), Lewy bodies (aSyn), or neurofibrillary tangles (Tau). However, recent biophysical studies showed that the fusion of GFP to full length Tau (Tau_FL-GFP) or the short repeat domain Tau containing a pro-aggregating mutation (Tau_RDΔK-GFP), separated with a 13–14 amino acids linker, significantly alters the β-strands packing within the fibrils [43]. In addition, atomic force microscopy (AFM) showed that the fibrils formed by Tau_FL-GFP were wider and characterized by the presence of an additional halo of height corresponding to the size of GFP (≈3 nm). Similar observations were made for mutant Httex1 fused to GFP (Httex1-GFP) [44]. The increase of fibril width by 3 nm is also consistent with the size of the GFP protein. Similarly, the generation of aSyn-GFP fibrils in vitro resulted in a significantly delayed aggregation kinetics and the formation of wider fibrils [46]. In the study by Afitska et al. [47], the authors even reported that the fusion of GFP completely inhibited the primary nucleation of aSyn. In all cases, the resulting fibrils do not share the morphological properties of the fibrils found in AD, PD or HD brains. These observations demonstrate that the GFP subunits decorate the surface of the fibrils, change the surface properties of the fibrils and limit access to their core structure, thus altering their interactome or ability to catalyse surface-mediated secondary nucleation events. We summarize in Fig. 6 the biophysical and cellular influence of GFP on amyloid fibrilization and inclusion formation in vitro and in cells.

Many proteins are known to interact with aSyn fibrils [15] and their ability to access the fibril surface could influence fibril growth, post-translational modification, morphology and interactome, all of which influence their toxicity and ability to transition or mature into the pathological aggregates found in PD brains. Several studies, including work from our group have shown that some of the toxicity associated with inclusion formation is mediated by the recruitment of functional proteins and organelles into the inclusions. Furthermore, our work on Huntingtin inclusions shows that mutant Httex1 aggregates fused to GFP exhibit different protein interactomes and toxic properties compared to the inclusions formed by the untagged mutant Httex1 protein [48].
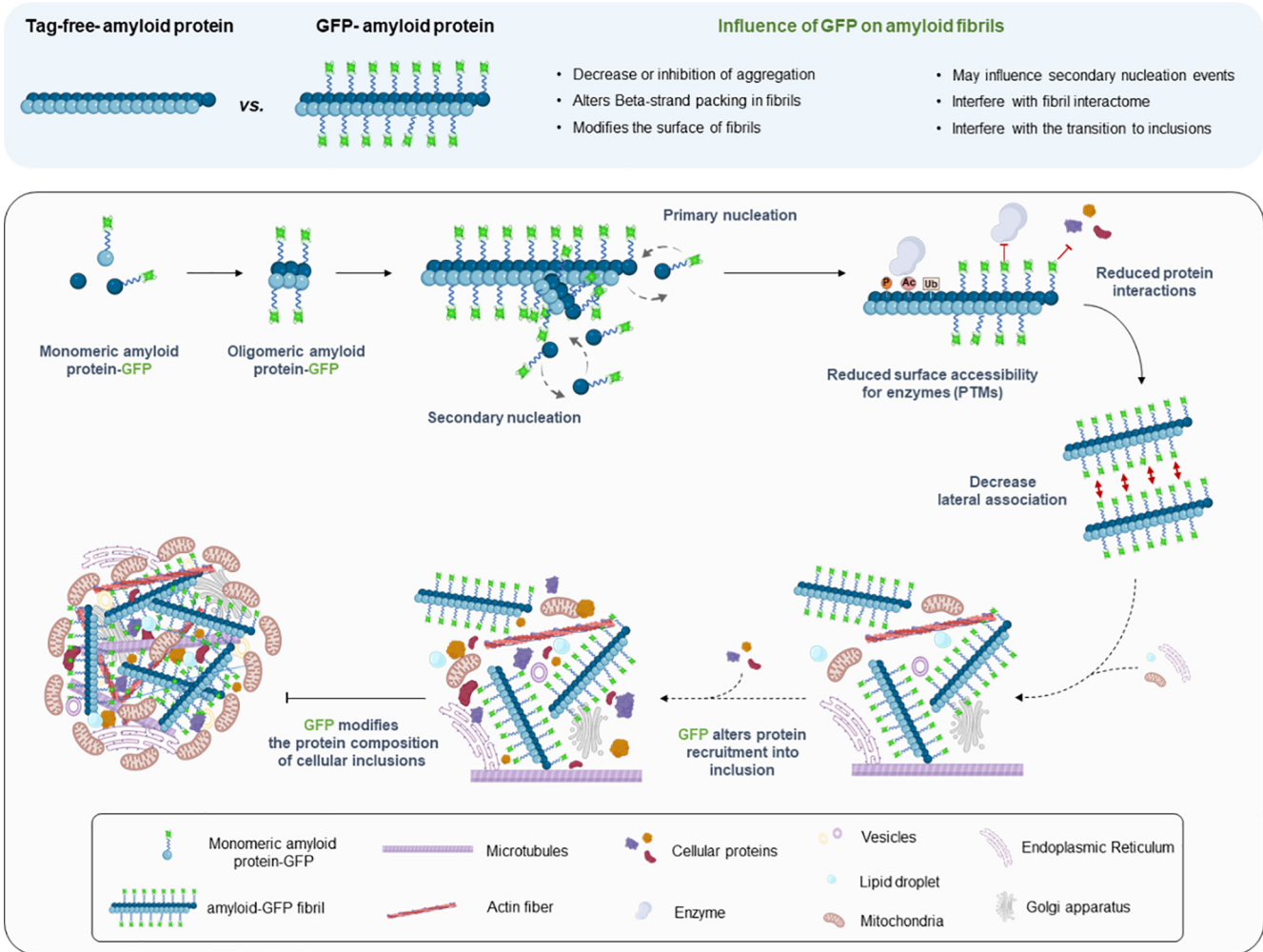
Fig. 6. Schematic illustration of the effects of GFP on the various stages of protein aggregation and inclusion formation based on published studies and predictions from our work. The depicted mechanisms illustrate the various stages associated with the mechanisms of aSyn oligomerization, fibrilization and LB formation [3,4]. (Created with BioRender.com).

Our results predict that the fluorescent tags act as a molecular "sieve" that differentially restricts access to the fibril surface for proteins of different sizes. We are not aware of any experimental data that compares the protein size-dependent interactome for aSyn fibrils with and without GFP proteins. However, we have recently determined the enrichment of soluble proteins in cytoplasmic inclusions in cells overexpressing mutant forms of exon 1 of the huntingtin (Htt) protein, with and without GFP fused to their C-terminal domain. Our correlative light electron microscopy studies confirmed that Htt fibrils are the primary component of these inclusions, confirming that the presence of GFP does not interfere with the ability of mutant Htt to form fibrils [6]. Despite this, a careful comparison of the composition of Htt inclusions and their associated toxicity revealed that the presence of GFP strongly influenced the ultrastructure, proteome and lipid composition of the inclusions and their toxicity.

Given the availability of the proteome data for the Httex1 72Q and Httex1 72Q-GFP cytoplasmic inclusions from this study (https://www.ebi.ac.uk/pride/archive/projects/PXD021742), we sought to determine if the presence of GFP influences the size distribution of the proteins that are recruited into Htt inclusions in cells. Fig. 7 shows the molecular weight distribution of proteins that are enriched in Htt inclusions in HEK cells. The right bar shows the composition in inclusions formed of bare Htt fibrils with a
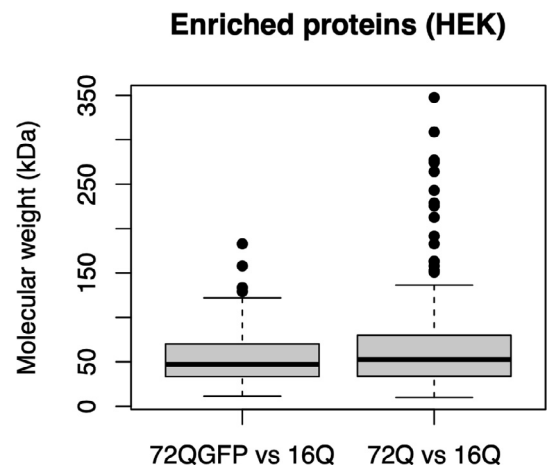

Fig. 7. Comparison of the protein composition of cytoplasmic Htt inclusions in an experimental model of Huntington's disease Htt aggregate inclusion formation from Ref. 6. HEK cells with inclusions formed after seeding with fibrils of Htt with a poly-Q repeat length of 72Q (as compared to a baseline of 16Q) decorated with GFP at the C terminus (left) and for unlabelled Htt (right). Inclusions containing label-free Htt fibrils are enriched in a wide range of proteins up to high molecular weights. GFP-tagged Htt fibrils have a similar distribution of low molecular weight proteins (∼50 kDa), but no proteins with molecular weights above 160 kDa are present.

polyQ length of 72, which is above the value at which pathological fibrils form. The left bar shows the composition when the Htt fibrils are decorated with GFP tags at their C-terminal end. It is clear that while low molecular weight proteins are present in both cases, no proteins with masses above ~ 150 kDa are found in the inclusions where GFP is present. The difference in the means between the two distributions is statistically significant (with a *t*-test for difference of means, p < 0.05) but what is most striking is the absence of proteins with high molecular weights in the GFP-tagged inclusions. Although the experimental situation is more complex than the simulations, it is consistent with the prediction of our simulations that GFP tags on flexible linkers preferentially exclude higher molecular weight proteins from the fibril surface.

Finally, the presence of GFP on the surfaces of fibrils seems to also influence their self-association and packing during the formation and maturation of inclusions. Bäuerlein *et al.* reported that the fusion of GFP to the C-terminal part of Httex1 resulted in a 50% reduction in fibril density inside cytoplasmic inclusion formed in primary neurons and a 25% increase in fibril stiffness due to the GFP decoration along fibrils [45]. A review of the aSyn seeding models revealed that seeding mediated aggregation in cells or transgenic mice overexpressing aSyn-GFP resulted in the formation of filamentous aggregates but not LB-like inclusions [34,35]. In Schaser *et al.* [49], A53T mutant aSyn-GFP mice were injected with PFFs which resulted in the formation of pS129, and ubiquitin positive filaments but did not form round LB-like inclusions. Similarly, Trinkaus and colleagues performed cryo-electron tomography (cryo-ET) from neurons expressing A53T aSyn-GFP and treated with recombinant or brain derived PFFs [50]. In both cases, the detected aSyn neuronal aggregates were predominantly composed of aSyn fibrils in the middle of cellular organelles and membranes, but they did not observe Lewy-body-like spherical inclusions similar to those formed by endogenous untagged aSyn. These observations suggest that the presence of GFP interferes with aSyn fibril lateral association and interactions with cytoplasmic proteins and organelles, processes that are tightly linked to the formation of LB-like inclusions. Although aSyn-GFP can form fibrils, whether these fibrils can seed aSyn aggregation as native (tag-free) aSyn PFFs has not been investigated [4,51]. Altogether, these observations and our data demonstrate that in addition to altering the kinetics of fibrilization of amyloid proteins and the biophysical properties of amyloid fibrils, the presence of large tags will also change the final structure and composition of protein aggregates and inclusions in cells.

3.4. Implications for drug discovery and identification of modifiers of amyloid formation and clearance

High throughput screening of small drug molecules has uncovered compounds that interfere with the initial formation of aSyn fibrils or secondary nucleation events at the surface of existing fibrils [52]. The molecule ZPD-2 inhibits the initial seeding of new filaments and was found to be most active when added early in the aggregation reaction [53]. SynuClean-D, by contrast, is a small molecule that inhibits aSyn fibril aggregation and acts to disaggregate mature fibrils in human cell and *C. elegans* model systems, but does not strongly interact with monomeric aSyn. It was predicted by computational analysis to bind to small cavities in the fibril surface supporting the importance of the surface [54].

Similarly, many of the therapeutic antibodies designed to target pathological aggregates or facilitate their clearance are designed to bind to sequences that decorate the fibril surfaces. Therefore, it is important that such therapeutic agents are validated in models expressing native protein sequences. The variation in the degree of occlusion with linker length and particle sizes suggests that experiments exploring the interactome of GFP tagged-amyloid fibrils with proteins or drugs, and the lateral association of multiple fibrils, should take account of the non-monotonic steric effects of tag and linker when comparing to label-free experiments. Additionally, we propose that *in vitro* experiments that measure reaction rates at fibril surfaces should be corrected for the effects of GFP tags when used to determine rate constants for theoretical modelling [13]. Accurately understanding fibril surface occlusion is therefore important for interpreting experimental data, building kinetic models of fibril elongation, and drug discovery studies. There is, unfortunately, very little experimental data on the identity and duration of proteins at fibril surfaces with which to compare our predictions. Fig. 7 shows recent results from our laboratory for Huntingtin fibrils, but permits only the size not the residence time of the proteins at the fibril surface to be extracted. We are planning to address this question experimentally in the future.

Collectively, our results indicate that experiments that use tagged and untagged monomers to study the growth and interactome of fibrils should be compared with caution, and the confounding effects of the tags are more complex than a simple reduction in surface accessibility. The prevalence of fluorescent tags in amyloid fibril growth experiments suggests that this has implications beyond the specific alpha synuclein fibrils we model here. Finally, given the increasing use of cellular assays and biosensors based on the expression of the amyloid protein to fluorescent proteins in drug discovery, it is essential first to determine which aspects of the pathological protein aggregation process of interest are recapitulated in these assays [55–61]. This requires detailed characterization of the aggregates and inclusions formed at the ultrastructural and biochemical levels. Furthermore, compounds and antibodies identified using these assays should always be validated in cellular and animal models expressing untagged native proteins.

4. Materials and methods

4.1. Dissipative particle dynamics simulation technique

We use the Dissipative Particle Dynamics technique (DPD) to study the diffusive approach of rigid nanoparticles to a stationary model aSyn fibril. The source code for the DPD simulations carried out in this work is available on GitHub: https://github.com/Osprey-DPD/osprey-dpd. DPD is a coarse-grained, explicit-solvent, molecular simulation technique designed to study the hydrodynamic behaviour of complex fluids [28–30], and soft materials [62–64]. Its advantage over both atomistic and coarse-grained molecular dynamics are its speed of execution and retention of the correct hydrodynamic behaviour of the solvent. The speedup is obtained by grouping several atoms or atomic groups into *beads* that interact via soft forces. This allows a larger integration step size to be used in the equations of motion. The large system size and long simulation times required to observe the diffusive approach of the particles to the fibril surface preclude the use of atomistic and coarse-grained Molecular dynamics (such as the Martini force field - http://www.cgmartini.nl/index.php/martini) because the simulations would require hundreds of days per run. This makes coarse-grained techniques, such as DPD, the only suitable tools. Although it means that sub-nanometer surface structure of the fibrils is not resolved in the simulations, we expect that the steric interference of the fluctuating GFP tags is accurately captured by the DPD technique because it does not depend on these surface details.

Atoms and molecular groups are represented in DPD by beads that interact via three non-bonded interactions that are soft, short-ranged (vanish beyond a fixed length-scale r_0) and pairwise

additive, conserving linear momentum. One force is conservative and gives each bead an identity such as hydrophilic or hydrophobic. Its magnitude is set by the parameter a_{ij}, which is the maximum force between beads of type i and j. The other two forces are a dissipative and a random force that together provide a thermostat that maintains a constant system temperature. The magnitude of the dissipative force is set by the parameter γ_{ij}. The masses of all beads are equal and set to unity. Molecules in DPD are constructed by connecting beads by Hookean springs defined by a spring constant k_2 and unstretched length l_0 that may depend on the bead types. A bending stiffness potential may be associated with adjacent bonds in a molecule that has the form $k_3(1 - cos(\varphi - \varphi_0))$ where k_3 is the bending constant (in units of k_BT) and Φ_0 is the preferred angle, which is zero if the bonds prefer to align parallel. Once all the forces have been defined, the simulation is evolved by integrating Newton's laws of motion for all the beads as described in the literature [28,62].

The aSyn fibril is composed of a central core made up of beads of type C that are constrained to be stationary; the linker is a short, linear chain of beads L, and the GFP moiety is a rigid cylindrical structure made of beads G. The solvent particles are represented by a single bead W that represents several water molecules. All the interaction parameters for the bead types and bonds are specified in Table 2.

4.2. Constructing the fibril and nanoparticles

The simulation length scale is set by the experimental value for the aSyn fibril diameter, which we take as 10 nm and the paired protofilament thickness 0.5 nm to correspond with cryo-EM experiments of aSyn fibrils [19]. The smallest nanoparticle has a 1 nm radius, and we set the range of the DPD non-bonded forces to this value, $r_0 = 1 nm$. A model aSyn fibril is preassembled in the simulation box from circular disks that represent the paired aSyn protofilaments or *monomers*. The monomers are bound together with strong Hookean springs to give the fibril a high rigidity. Although aSyn fibrils have distinct polymorphs, including twisted structures [65,66], these sub-nanometer details are not resolved in our coarse-grained simulations. We expect that the diffusive approach of a protein to the fibril surface is not greatly affected by atomic details of the surface until it approaches closer than one nanometer, which is below the accessible length scale here. We also ignore the disordered parts of the aSyn termini that protrude from the fibril's surface.

When simulating the decorated fibril, every monomer has a single GFP attached to it via a flexible linker that is bound to one point on the monomer. This corresponds to the typical experimental situation where all aSyn monomers are tagged. The GFP is represented as a rigid cylinder of diameter 3 nm and length 4 nm [22]. The point of attachment of the linker for successive GFPs is rotated by 51 degrees in a spiral around the core [67]. The tags are sterically excluded from the fibril and each other. The flexible linker allows them to fluctuate in position in response to thermal

noise subject to not intersecting other rigid objects in the simulation.

The simulation box is $40 \times 40 \times 30$ nm^3 and the fibril is 30 nm in length oriented along the Z axis. This is a compromise between a sufficiently long fibril to minimise the effects of the boundary conditions at its ends and a reasonable computational cost of the simulations. The lateral dimensions of the simulation box are four times the fibril diameter also to minimise the effects of the system size on the results.

Because we are interested in equilibrium properties, we do not attempt to fix the simulation time scale precisely. But an approximate value can be obtained as follows. Stokes law predicts that the diffusion constant for a rigid particle undergoing Brownian motion in a medium is $D = k_B T/6\pi\eta a$ where k_B is Boltzmann's constant, T is the temperature, η is the medium viscosity (0.001 Pa.sec for water), and a is the hydrodynamic radius of the particle. By comparing this with the dimensionless quantity ($D_{sim} = D\tau/r_0^2$) measured in the simulations, the DPD time-scale is τ = 0.3 ns. Each simulation is run for 4 10^6 steps with an integration step-size of 0.025 τ, which corresponds to 30 μsec, and requires 12 cpu-days on a single core of an AMD Ryzen Threadripper 3970X processor.

In order to allow the systems to reach equilibrium before we start measuring the probability distributions, we discard the first 2 million steps of the total simulation time of 4 million steps. The Supplementary Movies show that the particles diffuse freely throughout all regions of the simulation box in this time-frame, showing that they have reached equilibrium. Additionally, we illustrate the magnitude of the statistical errors in the distributions by taking samples from two time periods 2 – 4 10^6 and 3 – 4 10^6 time-steps for 10 and 20 particles. The similarity of the histograms in each panel of Figs. 3 and 4 show that the probabilities are independent of time and the number of particles (for 10 and 20 particles).

Finally, a number of nanoparticles are created in the solvent region of the box that represent diffusing particles such as aSyn monomers. They are geometric objects (spheres and ellipsoids) whose dimensions are in the range of typical intrinsically-disordered proteins [26]. The nanoparticles are constructed of a different bead type than the other structures in the simulation (GFP, linker, and fibril core) in order to allow their locations to be followed throughout a simulation. They are constructed as follows. A set of points on a three-dimensional rectangular lattice are defined in the simulation box distant from the fibril and GFP/linker according to the number of nanoparticles desired. The particles are then constructed at each of these points. Spherical nanoparticles are constructed by selecting all water beads within a specified radius around each point and tying them together with stiff Hookean springs to create a near-rigid body. Elliptic nanoparticles are created similarly by selecting all water beads within an ellipsoidal volume with given semi-major and semi-minor axes. Once they have been created, the nanoparticles are given a new bead type that distinguishes them from all other bead types in the simulation. The simulations place 10 or 20 nanoparticles at the points of the lattice within the solvent region of the box ensur-

Table 2

Bead-bead conservative force parameters a_{ij} (in units of k_BT/r_0) and dissipative force parameters γ_{ij} (in units of $\sqrt{m_0 k_B T/r_0^2}$) for all bead pairs, and Hookean bond potential parameters (in units of k_BT/r_0^2 and r_0 respectively). The water beads have the same conservative interaction with all bead types. NA = not applicable. The bending stiffness parameters for the linker (LLL) and GFP beads (GGG) are $k_3 = 20 \ k_BT$ and $\varphi_0 = 0$. Further details of the simulation parameters are given in the literature [62,64].

Bead Pairs	a_{ij}	γ_{ij}	k_2	l_0
WW, WC, WL, WG	25	4.5	NA	NA
CC	25	4.5	128	0.5
LC, LL, GC, GL, GG	50	4.5	128	0.5

ing they do not intersect with the fibril or GFP/linkers if present. The nanoparticles subsequently diffuse freely in the solvent as rigid bodies and are sterically unable to penetrate each other, the GFP tags or the fibril.

4.3. Quantifying the occlusion of the fibril surface by linked GFP tags

The probability that a particle will have its centre of mass in the circular shell from R to R + dR measured from the fibril axis is proportional to the amount of simulation time that it spends in this shell. We have chosen the box size to be sufficiently large that the diffusion of the particles is largely independent. The (unnormalized) histogram of this probability is obtained by summing the number of time-steps in which the particles are in each shell over the simulation time from $2 - 3 \ 10^6$, $3 - 4 \ 10^6$, and $2 - 4 \ 10^6$ time-steps, although we only show a subset of these results for clarity. We discard the first $2 \ 10^6$ steps to allow the system to equilibrate. The histogram is normalised by dividing it by the area of each shell (of constant thickness dR = 0.5 nm), the number of samples taken, and the number of particles (see Figure S3 for the limitation due to the simulation box size). It still depends on the length of the fibril and simulation box size. To remove this dependency, we calculate the histogram for the decorated fibril and the corresponding bare fibril (no GFP, no linker). We integrate each histogram over a user-defined region of space around the fibril. The ratio of this integral with the GFP tag to the corresponding integral for the bare fibril defines our measure of the occluding power of the tags. The selection of the precise region of integration is described next.

A dimensionless measure of the fibril surface accessibility is the ratio of the probability of a diffusing particle being within a certain distance of the fibril surface with the GFP tags present to the bare fibril value. This measure depends on the precise region over which the probability is integrated, and requires the lower and upper bounds to be chosen carefully. The lower limit of the integration range is the fibril radius, as the diffusing species are sterically excluded from penetrating the fibril. The upper limit of the integration range is set by the following condition. Far enough from the surface of the fibril, the probability of a particle lying within a given cylindrical shell around the fibril is unaffected by the presence of the GFP tags/linkers. The probability for particles to be in this region should not be included in the occlusion measure as it will overwhelm the signal from the (smaller) region where the GFP tags influence the particle's motion. The farthest distance at which the GFP tags can sterically interact with the diffusing particles depends on the fibril radius (R_{fibril}), linker length (L), and GFP length (4 nm from Yang et al.) [22], $R_{fibril} + L + R_{GFP}$. Thermal fluctuations will reduce this upper limit, so in practise we choose a smaller range by visual inspection of the histograms where the probability distribution is changing most rapidly. We have chosen the upper limit to be 4 nm from the fibril surface. This does not extend to the distance at which the bare fibril probability becomes flat because the signal we are attempting to measure would then be smothered by the probability unaffected by the GFP tag. We have explored the dependence of the occlusion factor when this distance is varied slightly, and our results are not significantly different.

The surface occlusion measure is defined as the ratio of the integrated probability of the nanoparticles' probability distribution over the predefined range with the GFP tags present to the value for the bare fibril. A value of unity indicates no occlusion, while a value of zero corresponds to the tags completely preventing access to the fibril surface.

CRediT authorship contribution statement

Julian C. Shillcock: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing –

review & editing, Visualization. **Janna Hastings:** Formal analysis, Data curation, Writing – review & editing, Visualization. **Nathan Riguet:** Data curation, Writing – review & editing, Visualization. **Hilal A. Lashuel:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hilal Lashuel has received funding from industry to support research on neurodegenerative diseases, including from Merck Serono, UCB, Idorsia and Abbvie. These companies had no specific role in the in the conceptualization and preparation of and decision to publish this work. H.A.L is also the co-founder and Chief Scientific Officer of ND BioSciences SA, a company that develops diagnostics and treatments for neurodegenerative diseases based on platforms that reproduce the complexity and diversity of proteins implicated in neurodegenerative diseases and their pathologies.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.12.017.

References

[1] Chiti F, Dobson CM. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. Annu Rev Biochem 2017;86:27–68.

[2] Goedert M, Spillantini MG, Del Tredici K, Braak H. 100 years of Lewy pathology. Nature Rev Neurol 2013;9:13–24.

[3] Shahmoradian SH, Lewis AJ, Genoud C, Hench J, Moors TE, Navarro PP, et al. Lewy pathology in Parkinson's disease consists of crowded organelles and lipid membranes. Nat Neurosci 2019;22:1099–109.

[4] Mahul-Mellier A-L, Burtscher J, Maharjan N, Weerens L, Croisier M, Kuttler F, et al. The process of Lewy body formation, rather than simply α-synuclein fibrillization, is one of the major drivers of neurodegeneration. Proc Natl Acad Sci 2020;117:4971.

[5] Ross CA, Aylward EH, Wild EJ, Langbehn DR, Long JD, Warner JH, et al. Huntington disease: natural history, biomarkers and prospects for therapeutics. Nature Rev Neurol 2014;10:204–16.

[6] Riguet N, Mahul-Mellier A-L, Maharjan N, Burtscher J, Croisier M, Knott G, et al. Nuclear and cytoplasmic huntingtin inclusions exhibit distinct biochemical composition, interactome and ultrastructural properties. bioRxiv. 2021:2020.07.29.226977.

[7] Leitão ADG, Rudolffi-Soto P, Chappard A, Bhumkar A, Lau D, Hunter DJB, et al. Selectivity of Lewy body protein interactions along the aggregation pathway of α-synuclein. Commun Biol 2021;4:1124.

[8] Burai R, Ait-Bouziad N, Chiki A, Lashuel HA. Elucidating the role of site-specific nitration of α-synuclein in the pathogenesis of Parkinson's disease via protein semisynthesis and mutagenesis. J Am Chem Soc 2015;137:5041–52.

[9] Zhao K, Lim Y-J, Liu Z, Long H, Sun Y, Hu J-J, et al. Parkinson's disease-related phosphorylation at Tyr39 rearranges α-synuclein amyloid fibril structure revealed by cryo-EM. Proc Natl Acad Sci 2020;117:20305.

[10] Arakhamia T, Lee CE, Carlomagno Y, Duong DM, Kundinger SR, Wang K, et al. Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. *Cell.* 2020;180:633-44.e12.

[11] Lashuel HA. Rethinking protein aggregation and drug discovery in neurodegenerative diseases: Why we need to embrace complexity? Curr Opin Chem Biol 2021;64:67–75.

[12] Mahul-Mellier A-L, Altay MF, Burtscher J, Maharjan N, Ait-Bouziad N, Chiki A, et al. The making of a Lewy body: the role of α-synuclein post-fibrillization modifications in regulating the formation and the maturation of pathological inclusions. bioRxiv. 2018:500058.

[13] Michaels TCT, Saric A, Curk S, Bernfur K, Arosio P, Meisl G, et al. Dynamics of oligomer populations formed during the aggregation of Alzheimer's Aβ42 peptide. Nat Chem 2020;12:445–51.

[14] Joseph J, Maji SK, Padinhateeri R. Computational model for studying breakage-dependent amyloid growth. ACS Chem Neurosci 2020;11:3615–22.

[15] Hsieh CJ, Ferrie JJ, Xu K, Lee I, Graham TJA, Tu Z, et al. Alpha synuclein fibrils contain multiple binding sites for small molecules. ACS Chem Neurosci 2018;9:2521–7.

[16] Spillantini MG, Schmidt ML, Lee V-M-Y, Trojanowski JQ, Jakes R, Goedert M. α-Synuclein in Lewy bodies. Nature 1997;388:839–40.

[17] Lautenschläger J, Kaminski CF, Kaminski Schierle GS. α-Synuclein – regulator of exocytosis, endocytosis, or both? Trends Cell Biol 2017;27:468–79.

[18] Yang X, Wang B, Hoop CL, Williams JK, Baum J. NMR unveils an N-terminal interaction interface on acetylated-α-synuclein monomers for recruitment to fibrils. Proc Natl Acad Sci 2021;118:e2017452118.

[19] Guerrero-Ferreira R, Taylor NM, Mona D, Ringler P, Lauer ME, Riek R, et al. Cryo-EM structure of alpha-synuclein fibrils. eLife 2018;7:1–18.

[20] Wagele J, De Sio S, Voigt B, Balbach J, Ott M. How fluorescent tags modify oligomer size distributions of the Alzheimer peptide. Biophys J 2019;116:227–38.

[21] Cendrowska U, Silva PJ, Ait-Bouziad N, Müller M, Guven ZP, Vieweg S, et al. Unraveling the complexity of amyloid polymorphism using gold nanoparticles and cryo-EM. Proc Natl Acad Sci U S A. 2020;117:6866–74.

[22] Yang F, Moss LG, Phillips GN. The molecular structure of green fluorescent protein. Nature 1996;14:1246–51.

[23] Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev 2013;65:1357–69.

[24] Mitrea DM, Chandra B, Ferrolino MC, Gibbs EB, Tolbert M, White MR, et al. Methods for physical characterization of phase-separated bodies and membrane-less organelles. J Mol Biol 2018;430:4773–805.

[25] Ilie IM, Caflisch A. Simulation studies of amyloidogenic polypeptides and their aggregates. Chem Rev 2019;119:6956–93.

[26] Tomasso ME, Tarver MJ, Devarajan D, Whitten ST. Hydrodynamic radii of intrinsically disordered proteins determined from experimental polyproline II propensities. PLoS Comput Biol 2015;12.

[27] Balupuri A, Choi K-E, Kang NS. Computational Insights into the Role of α-Strand/Sheet in Aggregation of α-Synuclein. *Scientific Reports*. 2019;9:59-1 - -13.

[28] Groot RD, Warren PB. Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulations. J Chem Phys 1997;107:4423–35.

[29] Espagnol P, Warren PB. Perspective: dissipative particle dynamics. J Chem Phys 2017;146:150901–17.

[30] Hoogerbrugge PJ, Koelman JMVA. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. Europhys Lett 1992;19:155–60.

[31] Hofmann H, Soranno A, Borgia A, Gast K, Nettels D, Schuler B. Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. Proc Natl Acad Sci 2012;109:16155.

[32] Marsh JA, Forman-Kay JD. Sequence determinants of compaction in intrinsically disordered proteins. Biophys J 2010;98:2383–90.

[33] Azizyan RA, Garro A, Radkova Z, Anikeenko A, Bakulina A, Duma C, et al. Establishment of constraints on amyloid formation imposed by steric exclusion of globular domains. J Mol Biol 2018;430:3835–46.

[34] Jarvela TS, Chaplot K, Lindberg I. A protease protection assay for the detection of internalized alpha-synuclein pre-formed fibrils. PLoS ONE 2021;16:e0241161.

[35] Karpowicz RJ, Haney CM, Mihaila TS, Sandler RM, Petersson EJ, Lee VMY. Selective imaging of internalized proteopathic α-synuclein seeds in primary neurons reveals mechanistic insight into transmission of synucleinopathies. J Biol Chem 2017;292:13482–97.

[36] Sanjeev A, Mattaparthi VSK. Computational investigation on the effects of H50Q and G51D mutations on the α-Synuclein aggregation propensity. J Biomol Struct Dyn 2018;36:2224–36.

[37] Wu K-P, Weinstock DS, Narayanan C, Levy RM, Baum J. Structural Reorganization of α-Synuclein at Low pH Observed by NMR and REMD Simulations. J Mol Biol 2009;391:784–96.

[38] Li X, Dong C, Hoffmann M, Garen CR, Cortez LM, Petersen NO, et al. Early stages of aggregation of engineered α-synuclein monomers and oligomers in solution. Sci Rep 2019;9:1734.

[39] Dünweg B, Reith D, Steinhauser M, Kremer K. Corrections to scaling in the hydrodynamic properties of dilute polymer solutions. J Chem Phys 2002;117:914–24.

[40] Nazarov S, Chiki A, Boudeffa D, Lashuel HA. The structural basis of huntingtin (Htt) fibril polymorphism, revealed by cryo-EM of exon 1 Htt fibrils. bioRxiv. 2021:2021.09.23.461534.

[41] Stephens AD, Zacharopoulou M, Moons R, Fusco G, Seetaloo N, Chiki A, et al. Extent of N-terminus exposure of monomeric alpha-synuclein determines its aggregation propensity. Nat Commun 2020;11:2820.

[42] Ochiishi T, Doi M, Yamasaki K, Hirose K, Kitamura A, Urabe T, et al. Development of new fusion proteins for visualizing amyloid-β oligomers in vivo. Sci Rep 2016;6:22712.

[43] Kaniyappan S, Tepper K, Biernat J, Chandupatla RR, Hübschmann S, Irsen S, et al. FRET-based Tau seeding assay does not represent prion-like templated assembly of Tau filaments. Mol Neurodegener 2020;15:39.

[44] Dahlgren PR, Karymov MA, Bankston J, Holden T, Thumfort P, Ingram VM, et al. Atomic force microscopy analysis of the Huntington protein nanofibril formation. Nanomedicine. 2005;1:52–7.

[45] Bäuerlein FJB, Saha I, Mishra A, Kalemanov M, Martínez-Sánchez A, Klein R, et al. In Situ Architecture and Cellular Interactions of PolyQ Inclusions. *Cell*. 2017;171:179-87.e10.

[46] Caputo A, Liang Y, Raabe TD, Lo A, Horvath M, Zhang B, et al. Snca-GFP Knock-In Mice Reflect Patterns of Endogenous Expression and Pathological Seeding. *eneuro*. 2020;7:ENEURO.0007-20.2020.

[47] Afitska K, Fucikova A, Shvadchak VV, Yushchenko DA. Modification of C terminus provides new insights into the mechanism of α-synuclein aggregation. Biophys J 2017;113:2182–91.

[48] Riguet N, Mahul-Mellier A-L, Maharjan N, Burtscher J, Croisier M, Knott G, et al. Nuclear and cytoplasmic huntingtin inclusions exhibit distinct biochemical composition, interactome and ultrastructural properties. Nat Commun 2021;12:6579.

[49] Schaser AJ, Stackhouse TL, Weston LJ, Kerstein PC, Osterberg VR, López CS, et al. Trans-synaptic and retrograde axonal spread of Lewy pathology following pre-formed fibril injection in an in vivo A53T alpha-synuclein mouse model of synucleinopathy. Acta Neuropathol Commun 2020;8:150.

[50] Trinkaus VA, Riera-Tur I, Martínez-Sánchez A, Bäuerlein FJB, Guo Q, Arzberger T, et al. In situ architecture of neuronal α-Synuclein inclusions. Nat Commun 2021;12:2110.

[51] Volpicelli-Daley LA, Luk KC, Lee VM. Addition of exogenous α-synuclein preformed fibrils to primary neuronal cultures to seed recruitment of endogenous α-synuclein to Lewy body and Lewy neurite-like aggregates. Nat Protoc 2014;9:2135–46.

[52] Pena-Diaz S, Pujols J, Ventura S. Small molecules to prevent the neurodegeneration caused by alpha-synuclein aggregation. Neural Regener Res 2020;15:2260–1.

[53] Peña-Díaz S, Pujols J, Conde-Giménez M, Čarija A, Dalfo E, García J, et al. ZPD-2, a Small Compound That Inhibits α-Synuclein Amyloid Aggregation and Its Seeded Polymerization. Front Mol Neurosci 2019;12.

[54] Pujols J, Peña-Díaz S, Lázaro DF, Peccati F, Pinheiro F, González D, et al. Small molecule inhibits α-synuclein aggregation, disrupts amyloid fibrils, and prevents degeneration of dopaminergic neurons. Proc Natl Acad Sci U S A. 2018;115:10481–6.

[55] Apostol BL, Kazantsev A, Raffioni S, Illes K, Pallos J, Bodai L, et al. A cell-based assay for aggregation inhibitors as therapeutics of polyglutamine-repeat disease and validation in Drosophila. Proc Natl Acad Sci 2003;100:5950.

[56] Nagai Y, Tucker T, Ren H, Kenan DJ, Henderson BS, Keene JD, et al. Inhibition of polyglutamine protein aggregation and cell death by novel peptides identified by phage display screening*. J Biol Chem 2000;275:10437–42.

[57] Zhang X, Smith DL, Meriin AB, Engemann S, Russel DE, Roark M, et al. A potent small molecule inhibits polyglutamine aggregation in Huntington's disease neurons and suppresses neurodegeneration in vivo. PNAS 2005;102:892.

[58] Zhou Q, Lehmer C, Michaelsen M, Mori K, Alterauge D, Baumjohann D, et al. Antibodies inhibit transmission and aggregation of C9orf72 poly-GA dipeptide repeat proteins. EMBO Mol Med 2017;9:687–702.

[59] Braun AR, Liao EE, Horvath M, Kalra P, Acosta K, Young MC, et al. Potent inhibitors of toxic alpha-synuclein identified via cellular time-resolved FRET biosensors. npj Parkinson's Dis 2021;7:52.

[60] Lo CH, Lim CK, Ding Z, Wickramasinghe SP, Braun AR, Ashe KH, et al. Targeting the ensemble of heterogeneous tau oligomers in cells: a novel small molecule screening platform for tauopathies. Alzheimers Dement. 2019;15:1489–502.

[61] Ahmadi S, Kerman K, Uppal M, Kraatz H-B. Protein-Based Biosensor for Screening of Tau Aggregation Inhibitors, a Pharmaceutical Application. ECS Meeting Abstracts 2020. MA2020-01:2474-.

[62] Shillcock J, Lipowsky R. Equilibrium structure and lateral stress distribution of amphiphilic bilayers from dissipative particle dynamics simulations. J Chem Phys 2002;117:5048–61.

[63] Lipowsky R, Brinkmann M, Dimova R, Haluska C, Kierfeld J, Shillcock J. Wetting, budding, and fusion—morphological transitions of soft surfaces. J Phys: Condens Matter 2005;17:S2885–902.

[64] Shillcock JC, Brochut M, Chénais E, Ipsen JH. Phase behaviour and structure of a model biomolecular condensate. Soft Matter 2020;16:6413–23.

[65] Fändrich M, Schmidt M, Grigorieff N. Recent progress in understanding Alzheimer's β-amyloid structures. Trends Biochem Sci 2011;36:338–45.

[66] Li B, Ge P, Murray KA, Sheth P, Zhang M, Nair G, et al. Cryo-EM of full-length α-synuclein reveals fibril polymorphs with a common structural kernel. Nat Commun 2018;9:3609.

[67] Azizyan RA, Garro A, Radkova Z, Anikeenko A, Bakulina A, Dumas C, et al. Establishment of constraints on amyloid formation imposed by steric exclusion of globular domains. J Mol Biol 2018;430:3835–46.

[68] Humphrey W, Dalke A, Schulten K. VMD - visual molecular dynamics. J Mol Graph 1996;14:33–8.