The University Of Sheffield.

Coarse grained modelling of amyloid fibril formation, inhibition and disruption using Monte Carlo simulations

Lianne Dawn Gahan

A thesis presented for the degree of

Doctor of Philosophy

School of Biosciences

Department of Physics & Astronomy

The University of Sheffield

September 2022

Amyloid Proteins and their hallmark fibrous structures are central to many disease pathologies such as Alzheimers[1], Type 2 Diabetes and Parkinsons[2]. Many proteins are capable of forming amyloid fibres, and above a critical concentration, the fibrillar structure is the most energetically stable structure that amyloid proteins can form. We use an in-house dynamic Monte Carlo simulation scheme[3, 4] in the NVT ensemble to model amyloid fibre assembly and its inhibition. In the context of this work, we represent amyloid protein monomers with a single rod-shaped particle. These protein monomers are modelled as spherocylinders interacting via attractive patches and a Van der Waals-like potential. We model inhibitors as isotropically interacting spheres that are not self-interacting but interact with the spherocylinders via a similar potential, requiring the attractive patch to point towards the sphere.

We have identified several regions of parameter space where significant fibril inhibition occurs. By varying properties such as inhibitor radius, inhibitor-protein interaction strength, protein aspect ratio and relative protein population, we observe mechanisms such as fibril capping, monomer and small cluster binding and surface coverage of clusters. We have observed key relationships between rod aspect ratio and the level of structural order in large clusters. Large inhibitory molecules can disrupt this order in almost all conditions considered. We have found that inhibitors can both aid and prevent fibrillisation in different regions of parameter space by modifying the strength of the interaction between the protein monomer and inhibitory particles. These mechanisms have implications for drug design to prevent further fibril growth. In preventing fibril growth, we could prevent the production of neurotoxic complexes in the following cases: (a) in an excess of free monomers by disrupting fibril assembly as it occurs, potentially preventing further brain degradation caused by downstream effects of additional fibrillisation. And (b) by disrupting and disaggregating existing fibril populations in a diseased brain, providing an opportunity for clearance of smaller, less ordered structures.

Acknowledgements

Let's start at the beginning: Caitlin and Lanz who provided that key emotional support during my master's year. They witnessed first hand a poor physicist trying to find her way through biology for the first time in 7 years. To Pete, Callum, Alex and Xander who all became part of Rosie's wee academic family with me. Alex without whom I'm sure I would have finished this PhD three years ago because I wouldn't be forever distracted with who's triangle it is, or what I'm taking to the moon. To Rosie who always manages to say the right thing to make it all seem fine. You are always so supportive, I can't imagine having gone through this process with anyone else a supervisor. Thank you all in the NMR group for all the lovely Christmas lunches, walks in the peaks and delicious cake mornings on a Friday. It's such a shame COVID put a bit of a stop to that, I hope it's a tradition you pick back up again soon. Aaron and Henry for the many drinks shared on a Friday evening, who I'm sure will both go on to become fabulous scientists in their own right. I have to make a special mention to Jessie for just being one of those people who's amazing at everything. I'm so excited to see where your career goes!

Rhoda I cannot thank you enough for steering this ship in the later years of my PhD. You have been the physicist voice of reason just when I needed it most, and I often wonder about where the project would be now (or where it wouldn't!) without you. Thank you so much for taking me on and for your patience. Thank you also to Tash, Helen and Medhi who always made me feel part of the group despite living 300 miles away.

To Nick, Lewis and Adam who I can always rely on for a good computational chat, or a shoulder to cry on when my code hasn't worked (again). Nick especially has been such a lynch pin in my PhD

and has taught me what I need to do to succeed, and gave me the tools to learn what I needed to. He has always taken time out of his schedule to meet with me to discuss how I can fix something, or just to lament about not having fixed it. I owe a great deal to Nick, and I hope to have made a friend in him over the years.

My partner, Alex, who has been completely unwavering in his support of me over the years and has always believed I could do it. Finally of course, my family who think I'm going to be at school for the rest of my life, and are very excited for me to get a "real job". My lovely parents and grandparents who have had my back since day 1, I can't express enough my love and gratitude for you, and the support you have given me to advance so far in my studies.

Contents

Chapter 1

Neurodegenerative Disease and Relevant Biological Context

1.1 Motivation

Neurodegenerative diseases are one of the major societal problems in an ageing population. Among these, Alzheimer's disease[1], Parkinson's[5] and other dementias are all associated with protein misfolding. There are 50 diseases associated with amyloidoses, such as type 2 diabetes and Sickle Cell Anemia[6]. Amyloid fibril formation is a significant hallmark of such disease pathologies. Amyloid-β is a small amyloidogenic peptide known to be involved in the pathology of Alzheimer's disease[1, 7]. Aβ is cleaved by β- and γ-secretase from amyloid precursor protein (APP) - an integral membrane protein expressed in many tissues and concentrated in the synapses of neurons. Monomers of Aβ peptide aggregate into large fibrillar structures (long one-dimensional biopolymers formed via extensive hydrogen bonding), which are found in the brains of Alzheimer's patients and ageing patients generally[8]. The corresponding length scales for these different structures are highlighted in figure 1.1. Current research determines that oligomeric intermediate species are the candidate for neurotoxicity[9–11] and that the presence of Aβ fibrils catalyses the production of such toxic species.

Alzheimer's disease (AD) is a neurodegenerative disease which causes 60-70% of cases of dementia[12, 13]. Though the precise cause of the disease is unknown, the β-amyloid peptide (Aβ) has

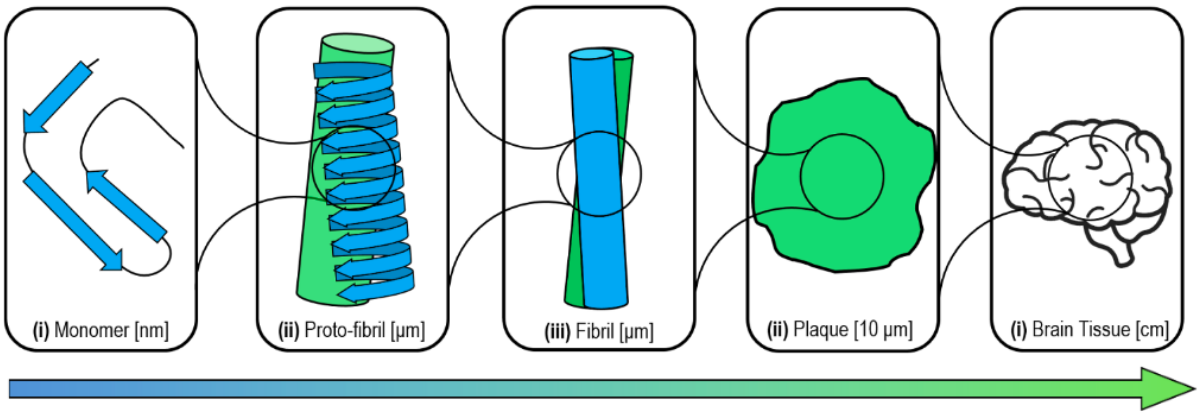

Figure 1.1: Corresponding length scales for comparison between structures. Fibrils form insoluble plaques (iv) on the brain (v). Fibrils (iii) are comprised of 2 or more proto-fibrils (ii) which are each in turn made up of thousands of monomers (i). While plaques are observed in the brains of AD patients, they do not necessarily imply a diseased state.

long been associated with disease progression due to high volumes of ordered fibril structures and plaques, found in the brains of AD patients[14]. However, the presence and formation of plaques do not necessarily imply a diseased state. Thus, the correlation between fibrillar aggregates that lead to plaques and AD is currently debated[15]. Plaques are present in older brains, but larger quantities of these depositions are observed in diseased brains[16, 17]. Rodrigue suggests that the poor link between amyloid deposition and cognitive function is because the amyloid deposition is a very early event in a series of downstream effects which will later result in detriment to cognitive function.[17].

Protein misfolding is believed to cause many neurodegenerative diseases. In amyloid formation, when one protein misfolds into a non-standard conformation, it is energetically favourable (in this case) for other proteins to fold into the same conformation, which leads to the build-up of these into sizeable insoluble aggregate beta sheets. The fact that the beta-sheet has no end point where fibrillisation can no longer take place makes the build-up so dangerous - the misfolding is self-perpetuating. The build-up of this amyloid plaque on the brain is indicative of an Alzheimer's disease brain. We can gain insight into this process by modelling mathematically and computationally. Determining a mechanistic understanding of the fibrillisation process allows us to understand better *in-vitro* experimental results and eventual prevention of the disease and other diseases like it.[18–20]

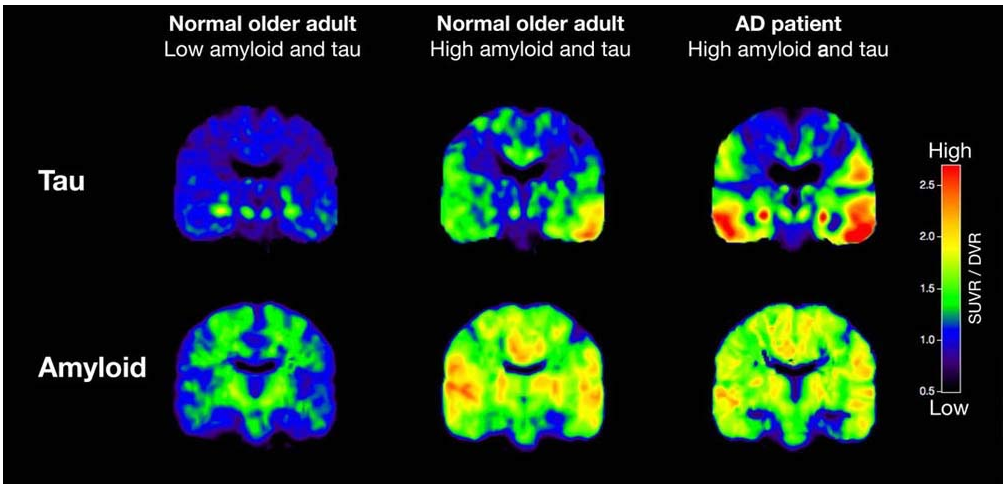

Figure 1.2: PET scans of a brain at varying levels of degradation due to Alzheimer's. Each row shows the effects of Tau and amyloid-β content in a normal adult with low content, a normal adult with high content and an Alzheimer's patient with high levels of amyloid content. Reproduced from the NIH[21].

In advanced Alzheimer's disease, it is easy to determine post-mortem if a brain is diseased due to brain shrinkage and atrophy. In more recent history, PET scans can highlight amyloid plaque build-up inside the brain as shown in figure 1.2 and also decreased areas of brain activity via glucose levels.

At present, there is only one drug which directly effects the disease pathology directly, lecanemab. Lecanemab prevents amyloid plaques in the brain from clumping. This study was the largest so far to look at whether clearing clumps of amyloid plaques from the brain can slow the disease[22]. However, there are medications available to alleviate symptoms. For example, certain drugs enhance neuronal activity by increasing the functionality of neurotransmitters[23, 24] thereby retarding disease progression, allowing the patient to maintain their independence for an additional $\approx$ 6-12 months. After this period, however, the patient deteriorates quickly as the brain degeneration is only masked rather than prevented by medication. Drugs that directly modify AD pathology have reached clinical trials, but no drug has been proven to significantly prevent or stop disease progression[11, 25, 26]. Some Aβ targeting drugs have hit the headlines in recent years, such as treatments which trigger immune responses that clear Aβ from cells or degrade the monomeric form of the protein. Some techniques lower Aβ quantities outside the brain and in the brain - suggesting that this process affects the blood-brain barrier. (A system of tight junctions that inhibit transport

across the vascular and central nervous systems) Alternatively, species which move Aβ around the body (carriers). Despite these studies, the precise mechanisms of the disease pathology of brain degeneration are widely unknown.

1.2 Amyloid Cascade Hypothesis

The pathology of Alzheimer's disease is governed primarily by the β-amyloid(Aβ) and tau proteins[12–14, 27, 28]. Monomers of Aβ protein aggregate into large fibrillar species, which have been found in the brains of Alzheimer's patients and older brains generally. However, there is a conformational difference between inert fibrils and those which cause the formation of insoluble plaques in AD patients.[8, 15, 29–31] It is thought that oligomeric intermediate species are neurotoxic[9–11, 32] and that the presence of Aβ fibrils may catalyse the production of such toxic species. This connects with literature from Rodrigue[17], which highlights the presence of high volumes of amyloid deposition in AD, but a poor link with cognitive function. A second protein, Tau, forms fibrils which tangle. The formation of tau fibrils is driven by hyperphosphorylation, a process caused by producing toxic Aβ structures, causing neurone cell death[33–36]. These tau aggregates are linked to cell death.

The amyloid cascade hypothesis[1] can be summarised to the idea that mutations in genes which produce amyloid precursor protein (APP), PSEN1 or PSEN2 can lead to the overproduction of amyloid-β protein. As shown in figure 1.3, α-secretase cleaves APP into two smaller proteins. An alternative process can also occur wherein β- and γ- secretase enzymes cleave APP in two places, forming three proteins, one of which is the β-amyloid peptide. PSEN1 and PSEN2 are the genes which govern the production of these secretases. Aβ cleavage can arise at different points in APP, which causes the formation of, amongst others, the 40 and 42 residue species of Aβ. These processes occur regularly in the body, only producing small quantities of β-amyloid peptide. Evidence suggests that unfolded Aβ may be useful for long-term memory[38, 39], or anti-inflammatory pathways in the brain. The anti-inflammatory pathways could then lead to a chain reaction of excess amyloid β protein accumulation, with a higher probability for misfolding events, which kickstarts the amyloid formation process and further disrupts brain function through downstream issues.[40,

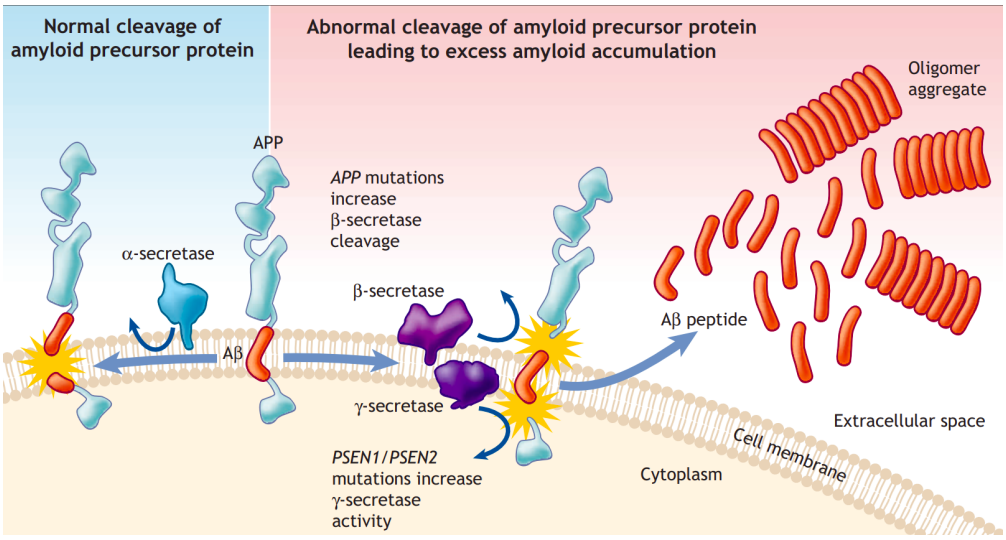

Figure 1.3: The amyloid cascade hypothesis: Formation pathways from APP into non-amyloid products via cleaving by α-secretase enzyme, and amyloidogenic products (i.e. Aβ peptide) via cleaving by β and γ-secretase enzymes leading to the formation of aggregates of varying size in extracellular space. Mutations in APP and PSEN1/PSEN2 (genes which encode cleaving enzymes) lead to the over-production of Aβ. Reproduced from Patterson et al.[37]

41] This process is highlighted in figure 1.3.

Mutations in the secretases can also affect the proportion of each alloform produced[42], which can affect disease pathology. The 42 residue alloform has quicker half times of fibrillisation than the 40 residue counterpart at the same concentration. Amyloid-β(1-42) is the main component of aggregates found in the brain, and there is clear evidence that high ratios of Aβ(1-42) to Aβ(1-40) make it more likely that someone has AD[43]. The 11-42 fragment is also of interest. There are also a range of other modifications at the N-terminal (acetylation, causing truncations of 11-40 and 11-42) and different lengths at the C-terminal (1-39, 1-40, 1-42 and 1-43), which all have different aggregation behaviours. Aggressive aggregation of specific mutants of Aβ(1-42) is associated with familial Alzheimer's. Mutations in these specific genes directly link to early onset/familial Alzheimer's and specific region variants of the disease. The gene for APP is found in chromosome 21, an additional copy of which is present in the genetic code of affected individuals[44]. Mutations in genes which encode the structure of APP, and the enzymes which cleave it, are considered risk factors in AD[42]. The removal of APP in transgenic mice[9, 45] has prevented AD pathology from occurring, and the addition of mutations leads to an AD-like pathology. The genetic argument puts

amyloidosis at the centre of many diseases, particularly AD. However, it is not possible to do this in humans. The removal of APP in transgenic mice can, however, result in different amyloidoses. If the fibrillisation and production of plaques of Aβ did not occur as a direct link of previously mentioned mutations, then there would be far more ways of getting AD and similar diseases.

Perhaps the strongest argument linking Abeta to disease causation is not just that altering the level or aggregation propensity of the Aβ peptide leads to disease, but also that there are no other ways of doing this. The latest, and most extensive, genetic population studies (GWAS: genome wide association study) put amyloid-β squarely at the centre of disease pathology.[46] Given the complexity of cell death mechanisms in the cell, and therefore the number of candidate genes, this makes for a convincing argument.

The genetic evidence supports the amyloid cascade hypothesis that amyloid formation is causative of disease pathologies. AD has been found to occur more frequently in older generations and people with Down syndrome[47, 48].

Whilst the amyloid cascade hypothesis is likely true for familial or hereditary Alzheimer's disease, this cannot be true for "sporadic" or spontaneous cases of AD without broadening the theory. There must be a process that imbalances the quantities of Aβ or tau protein monomers in the brain such that the body can no longer process aggregates and the amyloid formation process dominates over any clearing mechanisms the body has. Some possible triggers for this process are as follows:

- Aβ is an antimicrobial agent:

  Age leads to changes in the bacterial and viral load of the brain. Infections lead to an inflammatory response which kickstarts the process of producing Aβ from APP. Aβ acts as an antimicrobial peptide and induces the break-up of bacteria; an excess of Aβ leads to aggregation. The infection leads to an inflammatory response which causes APP to be processed into amyloid-β. Evidence suggests that an imbalance in Aβ concentration leads to toxicity and an inflammatory response[49]. However, it is believed that a cyclical process where Aβ is used in the inflammation response could drive aggregation of Aβ. However, the aggregates can cause neuroinflammation, further driving the production of Aβ. This

discussion creates a chicken and the egg argument over which causes which. Regardless, the cyclical process leads to a breakdown in the blood-brain barrier and leaves the brain more able to be permeated with an infection. This breakdown furthers the cyclical issue where the inflammation response to said infection leads to further production of Aβ.

- Concentration of protein changing with environment:

  The most prominent risk factors for increasing the probability of fibril nucleation come from increased concentrations of amyloid proteins such as Aβ, as well as changes to the environment that these proteins exist in. Ageing, obesity and smoking are all known to cause fluctuations in proteostasis which will affect the conditions in the body[50]. Such changes may also affect the nucleation surfaces for Aβ. e.g. Changes in populations of other proteins may provide a better catalytic surface for the nucleation of amyloid fibrils - the process of a critical number of amyloidogenic proteins coming together to form a new fibril (nucleation) on an attractive surface (heterogeneous).

- Cholesterol:

  The toxicity of Aβ is enhanced by a higher ratio of cholesterol to phospholipid levels. Cholesterol is an essential component of cell membranes and has important function in the the brain. It regulates fluidity, affecting key processes upstream and downstream of Aβ production. The processing of APP, which occurs at the membrane, is enhanced by high cholesterol:phospholipid levels. Once it aggregates, Aβ causes membrane damage directly, but also indirectly via binding to a number of membrane protein receptors which leads to neuronal dysfunction. Another link is via Apolipoprotein E (APOE) is a protein involved in the metabolism of fats in the body of mammals. APOE is the leading genetic risk factor for late-onset Alzheimer's and is a major cholesterol carrier in the brain because cholesterol may compete with Aβ peptide for its carrier[51, 52]. Carriers include APOE, serum albumin (HSA) in the cerebral spinal fluid and blood and membrane transporters to remove Aβ from the blood. If the body has higher levels of cholesterol, then it results in higher levels of Aβ not being processed, potentially leading to fibrillisation.

1.3 Protein Folding

Protein folding is the process of a chain of amino acids folding into a 3D structure. The physics of the protein folding problem has been well-studied[53]. The native state of a protein is typically a folded configuration. The folded state is the result of a delicate balance of opposing forces. The unfolded state allows hydrogen bonds to form between the solvent and the protein backbone or with side chains that may otherwise be buried in the folded state. The folded protein must form intramolecular interactions to stabilise the configuration to account for the loss of energetically favourable hydrogen bonds with the aqueous media. Regions of the protein backbone form hydrogen bonds with one another other, producing secondary structures (α-helices and β-sheets shown in figure 1.4), with combinations of these resulting in tertiary structures. Electrostatic interactions will form stronger ionic bonds when there is less or no screening of charges by the solvent. There is a contribution to folding and stability arising from van der Waals interactions, though they are very weak. In the near-crystalline environment of the folded protein interior, all atoms contribute to this distance-dependent force such that the effect is cumulative and, therefore, still has a significant impact. The folding process is therefore favoured due to a combination of these effects. A fundamental question in protein folding is as follows: How do proteins repeatedly fold to a specific conformation in such short timescales - despite the infinitely many conformations any protein could take? This has been explored by many authors[53–55] and is referred to as Levinthal's paradox. In reality, the free energy landscapes of proteins are often extremely complex. Proteins can find specific conformational structures due to funnel-like channels in the landscape, which guide proteins to specific structures[53, 54]. Again, the details of these mechanisms are highly complex and are active questions in the research community.

The folding process is encouraged when the free energy change is negative. Free energy is made up of enthalpy and entropy. The entropy of the protein typically decreases when folding occurs, so enthalpy (H-bonds) or entropy elsewhere must counter that so it becomes favourable for folding to occur. The following factors must be considered: hydrophobic effect, hydrogen bonds, and chaperones (some proteins need other proteins to fold). Entropy is defined by the number of different conformations a molecule can have based on its degrees of freedom. Chain entropy is limited

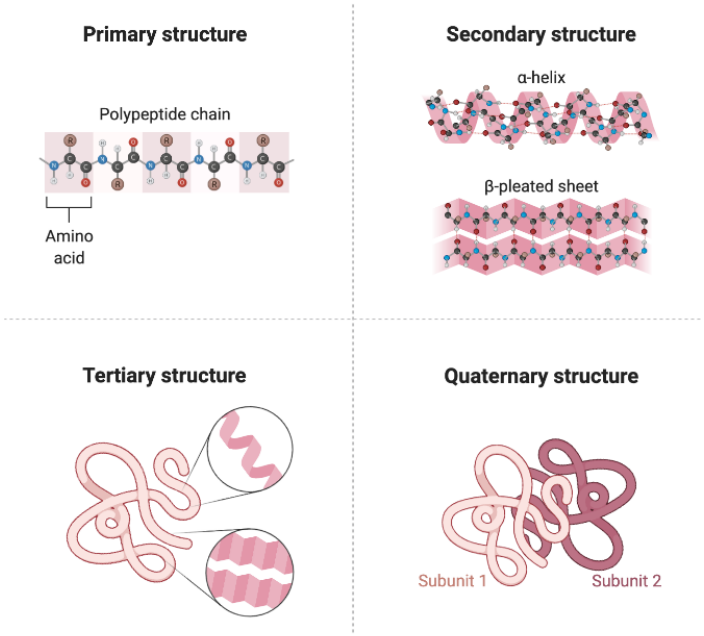

Figure 1.4: Schematic of the structure heirachy of proteins. This includes primary structure (top left), secondary structures of α-helices and β-sheets (top right), tertiary structure (bottom left) and quaternary structure (bottom right). Reproduced from Biorender.[56]

by backbone angle preference and steric exclusion. The entropy of a disordered unfolded random coil is much higher than the folded conformation of a protein and is seen as a barrier opposing folding[57, 58]. The collapse of the protein chain and the loss of interactions with the solvent are enabled because water-water interactions are much more favourable than those between water and a hydrocarbon. The minimisation occurs because water molecules can create H-bonds with other water molecules but not with hydrophobic groups in protein. i.e. Non-polar groups (e.g. specific side chains) prefer not to be in contact with polar environments (e.g. Water). Water also gains a massive amount of entropy throughout this process. Hydrophobic side-chains group together to minimise their contact with the solution. In many proteins, hydrophilic side chains can be found on the surface of a protein in contact with water whilst encasing hydrophobic groups. This type of behaviour is characteristic of quaternary protein structures including fibrils. The burial of hydrophobic side chains invariably leads to the burial of part of the backbone, which can no longer form H-bonds with water. Hydrophobic collapse thus promotes secondary structure formation to maintain hydrogen bonding potential during folding in the form of main chain hydrogen bonding. Therefore, folding can result in a reduced contribution of hydrogen bonds overall, decreasing system

enthalpy and reducing protein entropy which is approximately balanced by the entropy of the water. The hydrogen bonds forming between sections of the protein backbone will be shorter-ranged and, therefore, stronger than those formed between the protein and the solute. There is a delicate balance between entropy, enthalpy and free energy, and a slight change in conditions can result in different forces stabilising the protein structure.

In the case of many functional proteins, folding is supported by chaperone proteins. Chaperone proteins assist or catalyse the folding of other proteins. From a free energy perspective, they guide proteins through their free energy landscape to specific arrangements. Chaperones also assist a protein transition between several states required for their function. As mentioned previously, protein folding is a highly complex process. Chaperones make the process of correct folding more efficient.

1.4 Properties of Amyloid Proteins

An amyloid fibril is a long one-dimensional biopolymer composed of repeating protein monomers bound together by hydrogen bonds and other non-covalent interactions. In the case of fibril formation through amyloid protein aggregation, the process is often referred to as misfolding. Misfolding is an event where a protein (or proteins) finds a new conformation erroneously - which in neurodegenerative disease results in a high probability of aggregation[59]. Misfolding may happen when folding occurs incorrectly without the presence of a chaperone. The theory surrounding the kinetics of the protein states that there is a critical concentration above which aggregation into fibrils will occur[60]. An example of a cryo-EM fibril structure of Aβ is shown in figure 1.5. There are two protofibrils shown in figure 1.5(C-D). There are several distinguishing characteristics of amyloid fibrils listed below:

- Cross-β structure:

  A motif shared by all amyloid proteins involves the lamination of many beta-sheet layers. This structure is a result of all the other structural properties of fibrils[61]. The unifying features of amyloid proteins is their cross beta structure, an X-Ray fibre diffraction pattern ring at $d = 4.7\mathring{A}$,[62] and their visibility with congo red stain under polarised light. Observing

these different features is key in "diagnosing" if fibrils are amyloid fibrils.

- Hydrogen bonded beta strands in parallel stacks:

  Hydrogen bonding is observed between backbone amides and carbonyl groups of adjacent monomers along the fibril axis. The bond is non-rotational and planar, with the donor pointing along the acceptor's lone pair orbital and is the strongest form of non-covalent bond. Hydrogen bonding results in secondary, tertiary and quaternary protein structures where the hydrogen bonding occurs between the (peptide bond) carbonyl (oxygen) and the amide groups in backbone amino acids. Secondary structures include α-helices and β-sheets, which are then seen in higher order protein structures shown in figure 1.4. Hydrogen bonding is present along the protein backbone in fibrillar Aβ between parallel monomers forming a ladder-like pattern of bonds where amide and acid groups "zip" along the protein backbones. Hydrogen bonds can also contribute to salt bridges[63] between side chains, these are bonds made up of a hydrogen bond and an electrostatic interaction if the distance between two atoms is less than $4\mathring{A}$.

- In register side chains and aromatic groups:

  π stacks are a subcategory of electrostatically interacting structures that occur between aromatic groups from amino acids in different (adjacent) monomers within the amyloid fibril structure. There are several ways that this can occur. However, the most common is where one ring will be stacked but slightly offset from the one below it so that the aromatic ring's positive top or bottom region can bond ionically with the middle electron-dense ring. A hydrogen atom from a side-facing aromatic ring bonded to the face of another aromatic ring can be treated as a hydrogen bond[64]. While π-stacking is an effect broadly observed in many aspects of biochemistry, the extent of the stacking is unique in amyloids due to the length of the repeated structure.

- Burial of hydrophobic residues/hydrophilic residues depending upon the surrounding media

- The secondary structure is typically a feature of intermolecular interactions in a fibril:

  One of the significant structural differences between a standard folded protein and an amyloid fibril is that the secondary structure is predominantly based on the intermolecular interactions

in a fibril (figure 1.4). There may be interactions between different proteins, but primarily these are tertiary and quaternary properties of folded proteins rather than secondary. In fibrils, layers of β strands in amyloid fibrils result from hydrogen bonds between adjacent monomers rather than intramolecular interactions within a given monomer. It boils down primarily to how strands stack up by forming intramolecular H-bonds and other interactions such as π stacking. General folding rules apply to amyloids the same as to other protein structures. Therefore, burial of residues occurs here as well as other protein folding principles discussed in section 1.3.

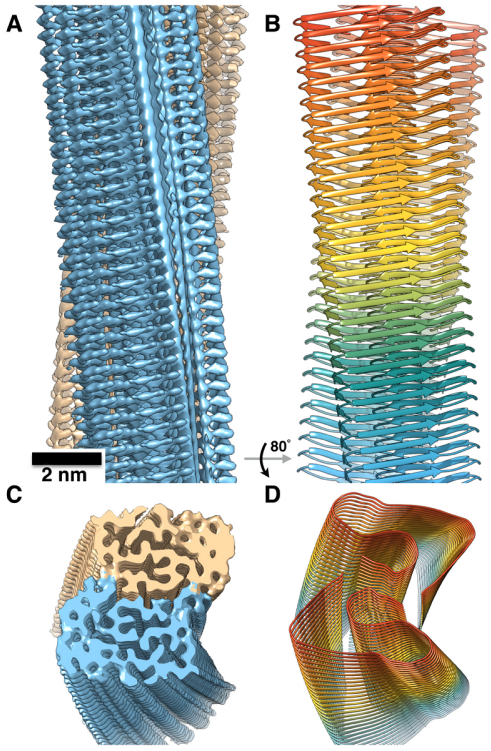

Figure 1.5: Aβ(1-42) fibril structure reproduced from Gremer et al.[65]. (A) 3D reconstruction from cryo-EM images showing density of two protofilaments (brown and blue) and the clear separation of the β-strands. (B) Atomic model of the fibril with parallel cross-β structure. (C) and (D) Tilted views of the cross-section of the EM density and the backbone model.

It is worth noting that amyloid-forming proteins can vary highly in length. Often, particularly in functional amyloids[66], only a small section of the protein chain contributes to the amyloid structure. Specific motifs are known for forming beta strands within fibrils such as KLVFFA and GNNQQNY.[67–69]

1.4.1 Fibril Stability

Alzheimer's disease and other protein misfolding diseases have pathologies centred about the amyloid state of proteins[13, 70–72]. (Figure 1.6) The amyloid state is typically the global energetic minimum state a protein can exist in. However, a significant barrier prevents most proteins from forming amyloid without a sizeable energetic cost. The native state of a protein is typically a local minimum on its free energy landscape. Free energy landscapes map a protein's conformational space in intermolecular and intramolecular contacts. This type of graph is generated via free energy calculations for a variety of different conformations and macromolecular structures - often, the space between these points is interpolated. A cartoon free energy landscape for an amyloidogenic protein is shown in figure 1.6.

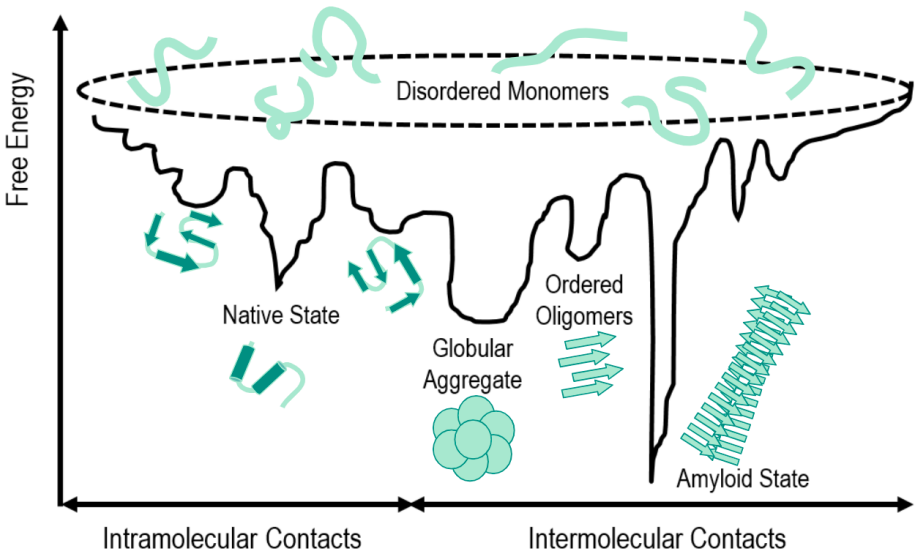

Figure 1.6: A generalised free energy landscape of an amyloidogenic protein. The inter- and intramolecular contacts formed due to conformational change within the monomer are included. Varying intramolecular contacts result in different conformations within a protein monomer, such as the native state of a given protein or a variety of folded states. Intermolecular contacts result in the formation of aggregates of different order parameters. These show the energy wells associated with structures such as amyloid fibrils, ordered oligomers and globular aggregates. The amyloid state is highly stable due to the huge free energy barrier, and once formed, it is improbable that the amyloid will entirely dissociate into its constituent monomers. It is the global minimum of free energy landscapes of many proteins. A protein can transition between these states with the assistance of chaperones which act as catalysts for such reactions. Some examples of disordered monomers are shown at the smallest free energies to highlight how these can collapse into local free energy minima. Reproduced from Muntau et al.[73]

Although most proteins can form amyloid structures, proteins are typically observed in a local minimum in their native conformation. Proteins such as Aβ can misfold into the amyloid state, forming large fibrillar structures comprised of one or more protofilaments. Protofilaments are long intermolecular *beta* sheets of Aβ monomers as highlighted in the structure hierachy diagram (Figure 1.1). There is an abundance of favourable hydrogen bonding and burial of hydrophobic residues in the core of fibrils.

Proteins in solution are subject to thermal fluctuations of the environment, implying a conformation is dynamic. The energy scale associated with such fluctuations is the thermal energy, $\approx k_B T$. A protein acquires its native state via sequential steps to minimise the system free energy. This native conformation, however, is not necessarily the energetic global minimum - in nature, there is typically a steep energy barrier surrounding the native state of any protein. The global minimum for many proteins at physiological concentrations is the amyloid state[72] as shown in figure 1.6. However, in vivo recycling of proteins results in insufficient lifetimes to find the amyloid state. Due to the concentration dependence of amyloid structures, diluting a solution containing fibrillar structures can cause the amyloid to disassemble into monomers. Figure 1.6 shows the energy of a protein as a function of its conformation. As can be seen, there are many energetic wells where different structures of Aβ can become kinetically trapped. These include intra- and intermolecular structures - some of which can be reversed by the presence of chaperone molecules that act as a catalyst to promote transitions between conformations over energetic barriers separating different conformations. Many conformations are available to any individual protein, each with an associated energetic state[55]. As such, to move over an energy barrier is to change conformation or yield aggregation/dissociation.

1.4.2 Oligomers

Oligomer is a generic term used in the amyloid field which encompasses a wide variety of structures even in the context of Aβ and is often the name given to polymers made up of small assemblies or clusters of monomers. There are a variety of different stable oligomeric structures of Aβ. Oligomers of Aβ take a variety of forms structurally. Such examples observed in simulations are discussed in section 1.8.2.2. These can include β-sheet rich structures (typical of amyloid structures[8]) or amor-

phous aggregates[73] made up of randomly orientated monomers. Not all oligomeric species lead to the formation of mature fibrils. These structures can be simply small multimers, up to a protofibril structure, or even completely different such as several monomers forming a *beta*-barrel[74] structure, which is hypothesised to pierce cell membranes and act as a channel[10]. Oligomeric intermediates observed in reaction pathways to fibril structures are hypothesised as the primary candidate for toxicity in Alzheimer's disease[9–11, 75–77].

Oligomer formation, polymorphism and conversion are largely not accounted for in current analytical models of amyloid formation. Similarly, it is difficult to experimentally measure or observe oligomeric structures or quantities. Producing a reliable analytical mathematical model which can be used against measurable quantities such as toxicity in cell models is critical in this field. Structural models and prediction of assembly mechanisms by including Aβ specific properties can inform said, mathematical models[78]. The most effective way of doing this is to compare existing structural models[65, 79–83] for Aβ and to extract common structural elements and conformational constraints.

Oligomers of amyloid-β have been found in the brains of Alzheimer's patients and appear more strongly correlated with the severity of AD symptoms[25, 26]. For example, soluble oligomers isolated from patients' brains were found to inhibit synapse responses[11]. Additionally, targeting Aβ oligomers in rodents has led to a halt in disease progression[9].

1.5 Studying Amyloid Formation

Thioflavin T (ThT) assays are the backbone of biochemistry work on the kinetics of amyloid-β[84, 85]. ThT does not bind to the Aβ monomer. However, it can bind to the fibril, allowing the observation of the mass concentration of fibrils via enhanced fluorescence from bound ThT at 480nm. The hallmark of kinetics in amyloid proteins is the sigmoidal curves produced from such assays, which can be fitted by the Oosawa nucleated polymerisation model[86], or more suitably complex models. The primary outcome of most papers - in theory, simulation and experiment - for amyloid peptide models is the characteristic sigmoidal growth curves[87–89].

There are gaps in our knowledge of the pathology of such diseases. Not only this, but working experimentally with amyloid proteins often comes with great difficulty due to - among other issues - primarily their tendency to aggregate. Modelling protein aggregation using theory and simulation allows the scientific community to make predictions and gain insight into the behaviour of said proteins on various length scales.

1.5.1 Kinetic Modelling

Using nucleation theory and mass action equations, we can use mathematics to model kinetics. More sophisticated predictions of system behaviour in different conditions come with increasingly complex models e.g. if the system is seeded or if specific processes are included. Any multi-step reaction where the end product is the only observable will retrieve sigmoidal kinetics[90]. These reactions can be characterised by a lag phase, a growth phase and a plateau. The lag phase present in amyloid aggregation is characterised to be larger for reactions with more intermediate steps due to the time taken for intermediate reactions to occur before the observation of the end product[91]. Following this is a growth phase wherein there is a marked increase in production of the end product - the fibril, in this case. This growth phase then exhibits a plateau as the amount of monomer available for polymerisation or nucleation events into new fibrils is depleted. Lansbury et al.[92–94] first hypothesised that amyloid formation is a nucleated reaction due to its dependence on concentration, seeding and the presence of a surface.

Mathematically, we can treat homogeneous and heterogeneous nucleation identically in the limit of weak binding to a surface for the heterogeneous case. In the strong binding limit, the high local concentration leads to over-saturation, leading to a plateau in nucleation rate wherein monomers run out of the said surface to bind to.

Aβ nucleation is progressive, this is where nucleation events occur continuously throughout crystallisation or polymerisation rather than instantaneously - both of which are shown in figures 1.7**a** and 1.7**b**. i.e. A singular event in crystallisation where all nuclei are formed at $t = 0$[95]. After

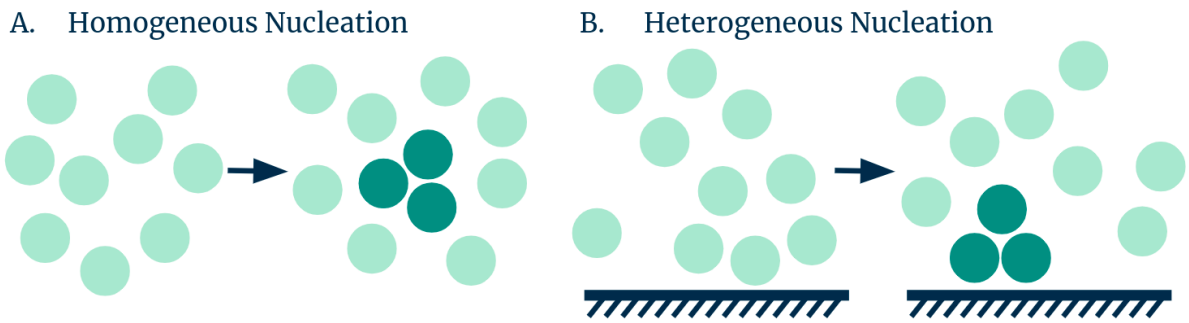

Figure 1.7: Diagram showing (a) Homogeneous nucleation of a new phase with nucleus size $n = 3$ which is in contact with the old phase only and (b) Heterogeneous nucleation of a new phase with nucleus size $n = 3$ in contact with both a surface and the old phase.

this point, there are no nucleation events.

Analysis of kinetics in monomer dependent secondary nucleation was pioneered by Eaton and Ferrone[96–98] for sickle cell haemoglobin. A single amino acid change in haemoglobin causes it to aggregate into large rods, preventing red blood cells from passing through capillaries, thereby preventing oxygen from getting where it needs to in the human body. The Knowles group have since contributed a large amount of work to nucleated polymerisation mass action models in their field[60, 99–109]. They have carried out an extended analysis of existing models such as Oosawa's[104] and have produced numerous models, which include polymerisation, depolymerisation, fragmentation, primary nucleation and secondary nucleation processes. While these models have analytical solutions found via fixed point analysis and moment closure, they do not account properly for the presence of oligomers or polymorphism in fibrillar species. Other groups[110–112] have also produced models of this kind and solved using moment closure wherein results are directly compared with similar Knowles models. Hall et al.[112] present a similar mass action model without secondary nucleation processes, which encompass a degree of polymorphism via the formation of both fibrils and amorphous aggregates. The general form of the master equation to model the time dependent fibril concentration $f(t, j)$ of length j, produced by Cohen et al. [60, 99, 100] is defined by equation 1.5.1.

$$\frac{\partial f(t,j)}{\partial t} = 2m(t)k_+ f(t,j-1) - 2m(t)k_+ f(t,j) + 2k_{off} f(t,j+1) - 2k_{off} f(t,j) \\ - k_-(j-1)f(t,j) + 2k_- \sum_{i=j+1} f(t,i) + k_2 m(t)^{n_2} \sum_{i=n_c} f(t,i)\delta_{j,n_2} + k_n m(t)^{n_c} \delta_{j,n_c} \tag{1.1}$$

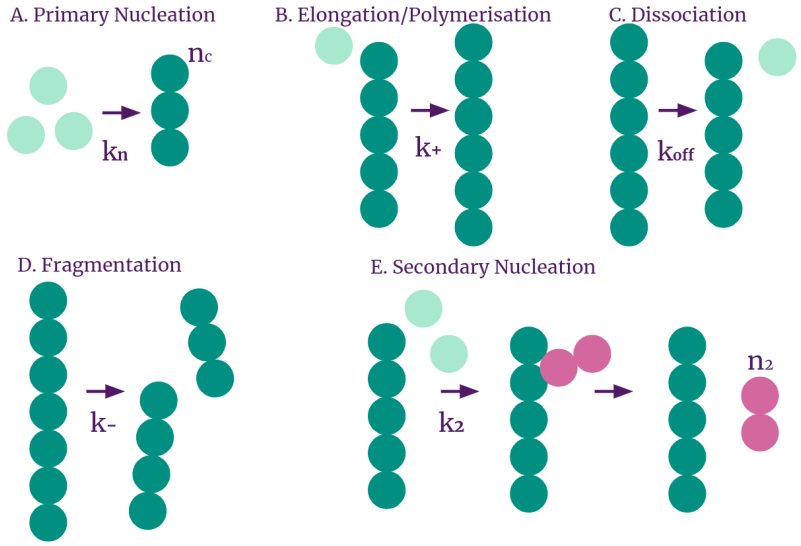

Figure 1.8: Examples of each type of process used in kinetic modelling of amyloids including (a) primary nucleation of a fibril size n_c with rate constant k_n, (b) elongation or polymerisation k_+, (c) dissociation or depolymerisation k_{off}, (d) fragmentation k_-, and (e) monomer dependent secondary nucleation of a fibril size n_2 on the surface of an existing fibril k_2. Reproduced from Cohen et al.[100]

Where $m(t)$ is the free monomer concentration, k_+ is the growth or polymersiation rate, k_- is the depolymerisation rate, k_{off} is the fragmentation rate, k_n is the primary nucleation rate and k_2 is the secondary nucleation rate. n_c and n_2 are the minimal nuclei sizes for primary and secondary nucleation respectively. This model builds upon the Oosawa nucleated polymerisation model by including several different processes. These are the dissociation of monomer from a fibril with rate constant k_{off}, the fragmentation of fibrils into two smaller fragments with rate constant k_- and monomer-dependent secondary nucleation, wherein a new fibril will nucleate and detach from the surface of an existing fibril with rate constant k_2. Examples of these processes are shown in figure 1.8(a-e). Dear et al.[109] follow this work and applies the Oosawa[86] model to a heterogeneous fibril assembly. They identify a phase diagram of the resulting possible structures from a system containing mixtures of two different species. Whilst this is not directly applicable to β-amyloid

fibrillisation, this is common in other areas of protein aggregation and polymer formation.

There are a great number of models that encompass these key steps in amyloid formation[113, 114]. Some have a focus on nucleation behaviours[115–121] which is a broad area of research that is important to many biopolymer systems. Others focus on conformational conversion which is a process which occurs when there are conformational changes within monomeric units in a fibril, which propogate along the length of the fibril[122]. These models are important as they provide crucial information regarding the rate constants of the individual processes. These kinetic models are often highly non-linear and require complex mathematics to solve - in the few special cases where they can be solved analytically[60, 99–101, 110, 123]. For this reason, many authors produce these complex master equations to describe the aggregation behaviour and solve them numerically for specific cases[60, 113, 124, 125].

An important output of kinetic models in addition to their ability to predict rate constants for the micro-processes within biopolymer growth, are length distributions. Whilst it is simpler to imagine the nucleation and growth of an individual fibril, in reality there will be a population of fibrils in any system at different stages of their development. Length distributions provide insight into a population of structures at a given point in time. Typically, analytical models can only provide equilibrium or steady state distributions due to the many assumptions required to solve these highly non-linear systems. However, when the model is solved numerically, it may be possible to determine the dynamics of how a population changes with time. An important downside to consider with the majority of kinetic models is that they do not provide structural information. Whilst we can gain key insight with a size distribution, few models differentiate between multiple states or conformers[122, 126–128]. Even in models where there is differentiation between species or conformers of the monomer, there is no differentiation between the different aggregates that even a single conformer could produce. The most simplified consideration here is that a number of identical monomers could form a spherical, globular aggregate or a fibril structure - which a size distribution would not be able to discern between.

In summary, kinetic modelling is vital to understanding amyloid formation. It is an important tool

we can utilise in fitting experimental data, and determining rate constants and complex outputs such as size distributions. They provide information that is either extremely challenging, or not possible to obtain through qualitative experimental methods including information about oligomeric species. However, they do not contain crucial structural information regarding the aggregate morphology. To do this, different computational approaches must be used. These will be discussed in detail in section 1.8.

1.6 The roles of Inhibitors

Inhibitors encompass a wide variety of structures with different sizes, shapes, binding patterns and functions. In AD, there is still plenty of discussion about the neurotoxic species. Fibrils appear to be the natural target. As fibrils mature, they tangle with other fibrils and develop a "halo" of oligomeric structures around them. Fibrils are believed to undergo secondary nucleation processes that produce neurotoxic oligomers and further fibrils[129], there are a huge number of downstream effects of fibril formation in the disease context. Therefore, it is helpful to consider drug or natural inhibitor intervention at multiple points along the biopolymer formation pathway. Examples of such interruption points are highlighted in figure 1.9.

Drug companies often screen huge libraries of compounds for disease treatment without mechanistic knowledge of their function. Therefore, there is little insight into the target species for some anti-AD drugs and how the anti-AD drugs interact with the target species. In the following sections, the current leading anti-amyloid drugs are explored as well ass natural amyloid inhibitors found in the body, such as protein chaperones used to clear amyloid proteins. Creating drugs that interact specifically without interacting with other brain functions is essential.

One of the first proposed treatments for AD was to create a vaccine against Aβ, an active immunisation process where the body creates antibodies against Aβ. Essentially this stimulates the immune system to attack amyloid-beta. If an AD brain is full of Aβ, this can lead to lethal encephalitis. Whilst this is probably the most efficient way of clearing amyloid, the consequences and side effects were quite drastic. Most research since has been under the passive immunisation

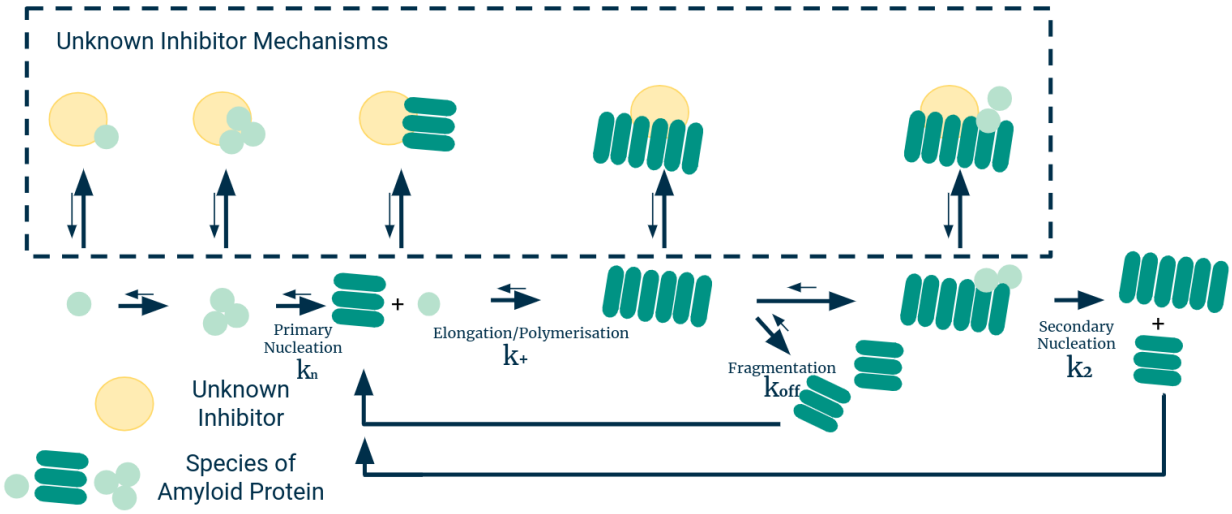

Figure 1.9: Examples of processes which occur on the pathway to fibril assembly in amyloid including (left to right) primary nucleation of a small oligomer into a fibrillar conformation, polymerisation growth of further monomers to existing fibrils, fragmentation events which split existing fibrils into multiple pieces and secondary nucleation events where new fibrils are nucleated on the surface of an existing fibril. Fragmentation and secondary nucleation events lead to positive feedback loops of further biopolymer assembly. Example inhibitors are included as byproducts with different amyloid protein structures at each stage of assembly.

umbrella, where the antibodies are made and administered. However, the body cannot replicate these nor do the antibodies stimulate the body to produce more antibodies. However, they can stimulate the immune system to attack a certain target. The effectiveness issues are that passive immunisation is gentler; it is also administered outside the brain and into the bloodstream. Despite this, several antibodies are quite effective at clearing amyloid. (section 1.7.1).

Classifying antibodies and other inhibitor behaviours can be complex. Little is known about their proper function and which part of the fibrillisation process they interfere with because it is often difficult to characterise this experimentally. Instead, inhibitors can be characterised by their size, binding behaviour, the number of binding sites and their binding affinity, and the stochiometric ratios required to function effectively.

1.7 Fibril and Oligomer Inhibitors

1.7.1 Currently Available Drugs and Fibril Inhibitors

According to Alzforum, there are currently only seven therapeutics for Alzheimer's Disease which are FDA-approved and have not been discontinued or are considered inactive. These are Lecanemab[22] Aducanumab[130], Donepezil[131], Galantamine, Memantine, Rivastigmine, Suvorexant and Tacrine.

Donepezil is one of the world's most currently used therapeutics and is used to combat mild to moderate AD, and has been improved in several countries for dementia with Lewy bodies and Parkinson's. This drug is not known to affect the disease pathology but does delay the worsening of cognitive symptoms in AD[133, 134]. Similarly Galantamine[135] and Rivastigmine[136] are also used to target the cholinergic system. Galantamine and Rivastigmine are acetylcholinesterase inhibitors - an enzyme found at neuromuscular junctions. Neuronal loss during AD progression leads to lower levels of acetylcholine (ACh) which then goes to weaken synaptic function and exacerbates AD symptoms[137]. Like Donepezil, there is no evidence to suggest they directly affect the disease pathology. However, there is evidence to suggest they improve cognitive symptoms and prevent the worsening of such symptoms. They do, however, have an extensive side effect list including nausea, vomiting and diarrhoea - although Rivastigmine appears to be less severe and is also available for Parkinson's disease treatment and mild dementias, as shown in figure 1.10. Tacrine is also on this list, which has largely been discontinued in favour of other more effective acetylcholinesterase inhibitors which have since been produced. This brings the number of FDA-approved drugs to just 6.

Memantine is a small molecule drug designed to tackle symptoms of moderate to severe AD in many countries. It is also known as Akatinol and is used to treat dementia in Germany by alleviating symptoms and preventing behavioural changes. Memantine's principal mechanism of action is believed to prevent the current flow through channels of NMDA receptor-operated ion channels (the NMDA receptor is a glutamate receptor with an ion channel found in neurons). It does this by reducing the effects of excitotoxic glutamate release. Memantine has higher binding affinity than Mg^{2+} ions at the NMDA receptor, therefore the Ca^{2+} influx is blocked while preserving

Name	Synonyms	FDA Status	Company	Target Type	Therapy Type	Approved For
Aduhelm	Aducanumab, BIIB037	Alzheimer's Disease (Approved)	Biogen, Neurimmune	Amyloid-Related	Immunotherapy (passive)	
Donepezil	Aricept™, Donepezil hydrochloride, Eranz®, E 2020	Alzheimer's Disease (Approved), Dementia with Lewy Bodies (Inactive), Down's Syndrome (Inactive), Parkinson's Disease Dementia (Inactive)	Corium, Inc., Eisai Co., Ltd., Pfizer	Cholinergic System	Small Molecule	Alzheimer's Disease, Dementia with Lewy Bodies (Japan)
Galantamine	Razadyne™, Reminyl™, Nivalin®	Alzheimer's Disease (Approved)	Janssen, Ortho-McNeil Pharmaceutical, Sanochemia Pharmazeutika, Shire, Takeda Pharmaceutical Company	Cholinergic System	Small Molecule	Mild to Moderate Alzheimer's disease
Leqembi	Lecanemab-irmb, BAN2401, mAb158	Alzheimer's Disease (Approved)	BioArctic AB, Biogen, Eisai Co., Ltd.	Amyloid-Related	Immunotherapy (passive)	
Memantine	Ebixa™, Namenda™ , Axura®, Akatinol®, Memary®	Alzheimer's Disease (Approved)	Forest Laboratories, Inc., Lundbeck, Merz Pharma	Other Neurotransmitters	Small Molecule	Alzheimer's Disease
Rivastigmine	Exelon™, Rivastigmine tartrate , Rivastach® Patch, Prometax®, SDZ ENA 713	Alzheimer's Disease (Approved), Parkinson's Disease Dementia (Approved)	Novartis Pharmaceuticals Corporation	Cholinergic System	Small Molecule	Mild to moderate Alzheimer's disease and mild to moderate dementia related to Parkinson's disease
Suvorexant	Belsomra, MK-4305	Alzheimer's Disease (Approved)	Merck	Other	Small Molecule	Insomnia, Insomnia in mild to moderate Alzheimer's disease
Tacrine	Cognex™	Alzheimer's Disease (Approved)	Pfizer, Shionogi Pharma	Cholinergic System	Small Molecule	

Figure 1.10: Table of available Alzheimer's Disease therapeutics showing their name, synonyms, FDA status, Company of Origin, Chemical Target, Type of therapy and for what diseases the drug is approved. There are only eight drugs which are currently FDA-approved and in use for AD treatment.[132]

transient physiological activation of the ion channels by activity-dependent, synaptically released glutamate[138]. There is debate surrounding the drugs' effectiveness for mild AD[139, 140] even though it is often prescribed at this stage of disease progression. Studies suggest that using Memantine and a cholinesterase inhibitor is more cost-effective than no treatment as it reduces the number of caregiver hours required by slowing cognitive decline in AD and similar conditions[141].

There is also a drug, Suvorexant, which was initially designed to tackle the sleep disorder insomnia. In recent years the drug has been prescribed to alleviate disruptions to the circadian rhythm in AD patients to help them sleep. Suvorexant has also been shown to improve cognitive function, restore hippocampal synaptic plasticity, and reduce Aβ plaque deposition in the hippocampus and cortex[142].

All of these FDA-approved therapeutics are small molecule targets except for one: Aduhelm. In May 2022, the drug moved into its 4th phase trial, with results expected in 2026. Aduhelm - more commonly known as Aducanumab, is an antibody inhibitor which binds to aggregated forms of Aβ rather than the monomer. It has been found to reduce the size of plaques in the brain (parenchymal amyloid) but with no change in size to vascular amyloid aggregates. Cerebral amyloid angiopathy (CAA) or vascular dementia is often associated with Alzheimer's. It generally results in plaques rich in Aβ(1-40) in the blood vessels, leading to symptoms of multiple ministrokes, increasing in severity with disease progression. It is surprising to observe a drug with such success despite not targeting the causative aggregates in blood vessels. However, clinical trials evidenced a slowed rate of decrease in Clinical Dementia Rating and slowed the rate of decrease in cognitive function[130]. A very recent breakthrough has also resulted in the FDA approval of Lecanemab, which has a high affinity to binding with Aβ(42) protofibrils, which are a form of oligomer that are known to have a high neurotoxicity rating[143].

1.7.2 Antibody Inhibitors in Clinical Trials

Gantenerumab[144] and solanezumab[145] are antibody drugs in their 3rd phase of medical trials. Gantenerumab is proposed to disassemble and degrade the brain's parenchymal and vascular Aβ plaques. Roche is currently developing new drug delivery systems called "brain shuttle" technology to increase the ability of Gantenerumab to enter the brain as shown with another drug, the protease neprilysin[146]. Gantenerumab also neutralises the inhibitory effects of oligomeric Aβ(42) on long-term potentiation in rat brains. i.e. Signalling between neurons is improved.

Solanezumab is another antibody drug that sequesters Aβ by removing small soluble species believed to be toxic to synapse function.[145] It is a monomer-selective antibody that interacts with an epitope of Aβ(residues 16-26). (Figure 1.11) Despite the remaining amyloid plaques, the drug has been shown to reverse memory deficits in APP transgenic mouse models. This drug will not be considered for mild AD but is still under consideration for prodromal AD. i.e. The period between the appearance of initial symptoms and development to a more severe case. A trial is being carried out on people without memory loss who have elevated brain amyloid levels.[147, 148]

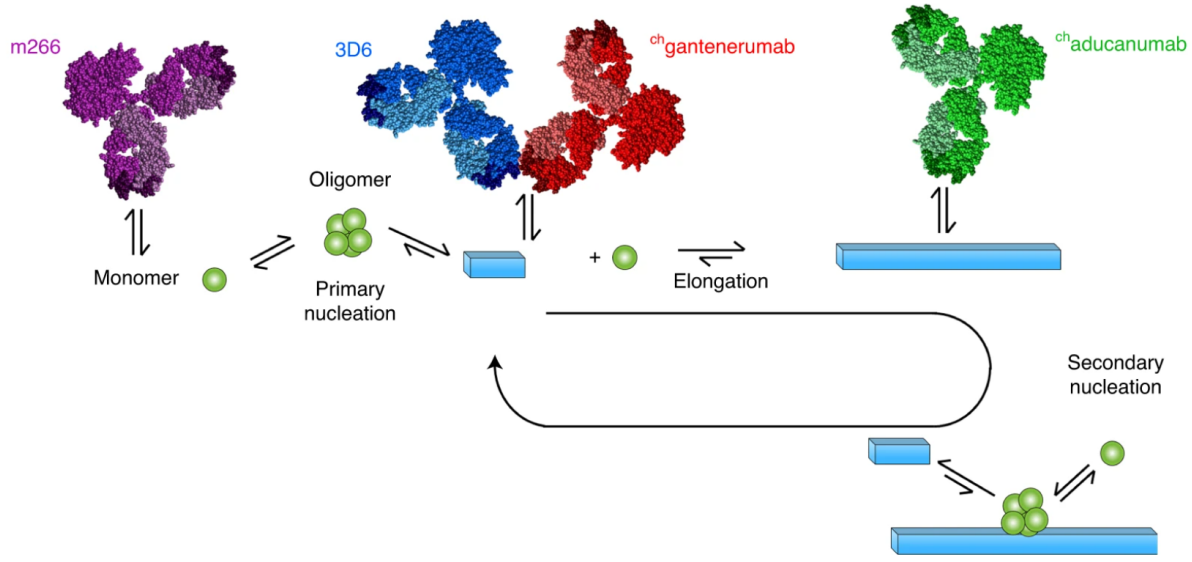

Figure 1.11: Schematic illustration of the different microscopic steps in the aggregation primarily affected by the four antibodies. Structural models were prepared using MOLMOL[149]. Reproduced from Cohen et al. [150]

.

There is also the antibody inhibitor Bapineuzumab which has since been discontinued from the use by the FDA due to the lack of evidence of treatment effect on cognitive or functional outcomes in patients. Evidence suggests that Bapineuzumab interacted with its target but did not clinically benefit the patients. The trials were carried out at a low dosage due to the extreme side effects and given to patients with late-stage AD, so this might be a matter of "too little too late".

Many of these antibodies do not appear to significantly change the amount of soluble oligomer despite clearly having some impact on symptoms in clinical trials. Kinetic studies suggest that

Aducanumab is best able to tackle secondary nucleation behaviour[150], which is believed to drive the production of neurotoxic oligomers. (Figure 1.11)

1.7.3 Chaperone Proteins

BRICHOS is a protein domain of about 100 amino acids originally found in BRI - a precursor protein to familial dementia[151] which has since been found in several protein families associated with diseases such as respiratory distress syndrome and cancer[152, 153]. There are approximately 50 diseases known to be associated amyloid aggregation processes. Almost all proteins have regions capable of forming amyloid, but the majority do not - suggesting that there are protective mechanisms to prevent amyloid formation *in-vivo*[154]. The transmembrane region of lung surfactant protein C precursor (proSP-C) forms amyloid-like fibrils in vitro. ProSP-C contains a BRICHOS domain, in which many interstitial lung disease (ILD)-associated mutations are localised, and the BRICHOS domain can prevent SP-C from forming amyloid-like fibrils[153].

There is evidence that shows Bri2 (a specific domain in the BRICHOS protein) affects the processing of amyloid precursor protein (APP) and that Bri2 reduces Aβ aggregation[155]. However, the mechanism of this is not known. Cohen[156, 157] proposes that BRICHOS inhibits the production of neurotoxic oligomers associated with Aβ(42) aggregates as shown in figure 1.12. This figure shows a disruption to the positive feedback mechanism of secondary nucleation events, which produce neurotoxic oligomers. Adding BRICHOS disrupts this feedback loop, leaving brichos bound to surfaces of existing fibrils. Cohen et al. propose that the inhibition behaviour of the BRICHOS domain is highly selective to inhibiting the secondary nucleation mechanisms in fibril aggregation - it does not interact with monomeric amyloid-β, only fibrils.

The Knowles group have worked extensively on fibril formation kinetics and the rate constants that underpin such processes[60, 70, 99, 100, 102]. Their work in recent years has moved towards inhibition of such processes and using their analytical models to determine the effects of drugs on the aggregation behaviours of amyloid proteins such as Aβ[150]. Brichos shows a “kinetic fingerprint” similar to the antibody inhibitor aducanumab. The critical difference is that while aducanumab decreases the secondary nucleation rate, BRICHOS can completely suppress 2° nucleation events

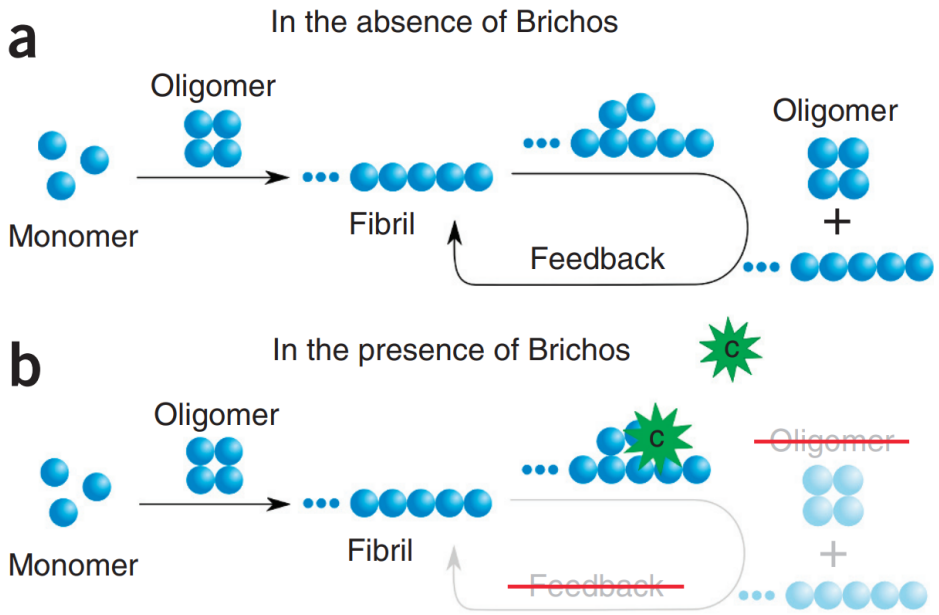

Figure 1.12: Brichos inhibits the catalytic cycle that generates toxic Aβ(42) oligomers. Schematic diagram showing the molecular pathways primary and secondary nucleationinvolved in oligomer formation in Aβ(42) aggregation (a) and the mechanism by which Brichos suppresses the formation of toxic oligomers, in which the secondary nucleation pathway is specifically inhibited to remove the dominant source of oligomers (b). Reproduced from Cohen et al.[157]

under high concentrations[156, 157]. There does not appear to be a benefit to adding Brichos and aducanumab together regarding how they affect amyloid kinetics. However, there is scope for combining therapies of Brichos and another antibody which inhibits a different kinetic pathway in amyloid aggregation[150].

Linse et al.[150] suggest that inhibition mechanisms that target the monomer would be challenging to achieve, as it would require stochiometric levels of the inhibitor to be present in the brain. Inhibitors of this type would reduce primary nucleation events and decrease the rate constant k_n. Similarly, by disrupting the elongation of fibrils by preventing further monomer addition e.g. Fibril capping. This would reduce the elongation rate k_+ and, in turn, potentially reduce the propensity for secondary nucleation events if the elongation processes are stopped early enough in the aggregation process. Finally, it is also essential to try and reduce the rate of secondary nucleation events k_2. Secondary nucleation is an autocatalytic feedback mechanism that accelerates further assembly. It occurs by catalysing nucleation on the surface of existing fibres (heterogeneous nucle-

ation). Secondary nucleation most significantly contributes to small (presumably toxic) oligomer formation. By covering the surface of fibrils, we might reduce the secondary nucleation rate k_2.

Most chaperones are intracellular because they are ATP-ases. There is no ATP outside the cell i.e. None of these standard systems works outside the cell. BRICHOS is an extracellular chaperone, and another example is clusterin. The original idea of looking at cystatins and transthyretin and how they inhibit Aβ was another chaperone-based idea. They also make amyloids, which have similar features and bind[158, 159]. Other chaperone proteins such as HSP70 have been found to disaggregate fibrous amyloid[344].

1.7.4 Small Molecule Inhibitors

Small molecule encompasses a variety of functional molecules which are on the order of size of amyloid protein monomers. This definition is broad due to the variance in size of proteins with a high propensity to form amyloid structures. However many of these compounds are smaller than a monomer of a small amyloidogenic peptide such as Aβ. (Figure 1.1) There are a number of natural compounds which are known to interact with amyloid proteins[160].

Small molecules tend to be quite generic rather than being appropriate to target specific amyloid. There is a delicate balance to be found here, where if a molecule is too generically binding, it could potentially bind to many different proteins in the body. For example, Epigallocatechin gallate (EGCG), a small molecule natural product in green tea that interacts with amyloid fibres, is not harmful to the body. However, it can bind to many different proteins, making it difficult for any part of the body to receive an effective dosage. In addition to this, the body can break down EGCG upon ingestion. This concept of the amount of a drug that can reach its target site is called bioavailability. The opposite effect would be to use something that is tightly binding such antibodies. However, these are typically highly specific to proteins or specific structures of a given protein. Usually, with drug design, researchers aim for them to have the highest binding capacity, which usually implies specificity, but this might not always be the best approach.

Polyphenols are a group of compounds (more commonly known as antioxidants) found in fruit

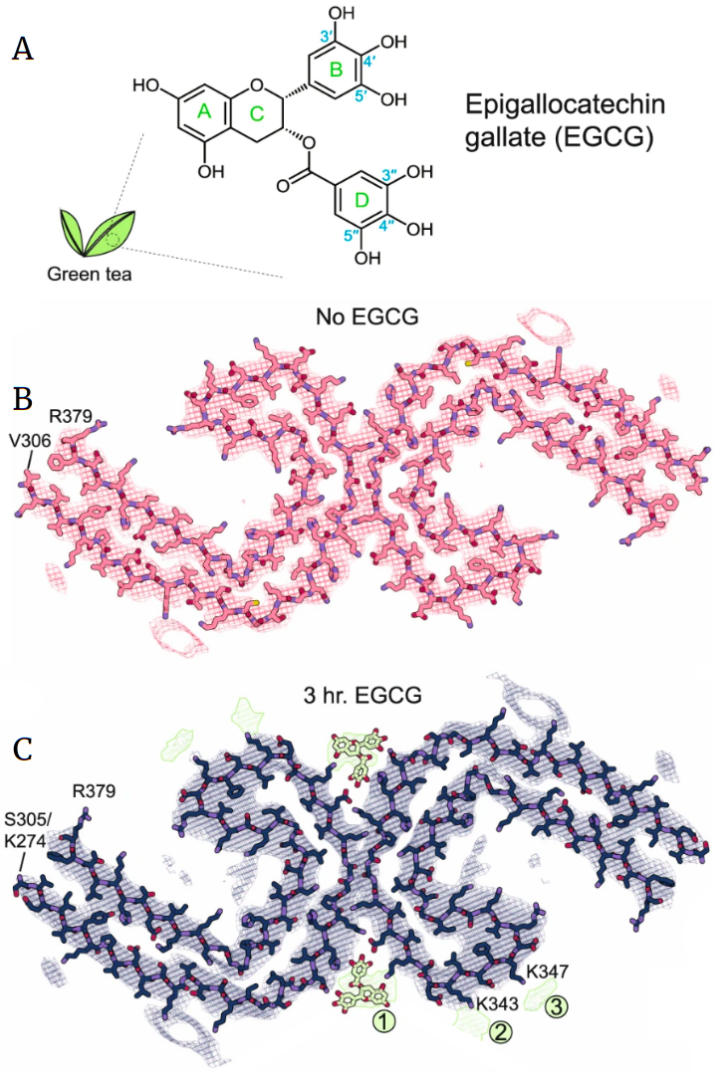

Figure 1.13: Schematic of EGCG binding to AD-tau fibrils with EM data. Reproduced from Seidler et al.[161]. **A** Epigallocatechin gallate (EGCG) is the most abundant polyphenol in green tea. It includes a benzenediol ring (A) adjoined to a tetrahydropyran moiety (C), which are connected to a galloyl ring (D) and pyrogallol ring (B). The 8 hydroxyl groups allow EGCG to engage in hydrogen bonding and other polar interactions with numerous biomolecules. **B** Cross-sectional view of the AD patient brain-derived tau PHF cryoEM structure before the addition of EGCG. **C** Tau PHF structure following 3-h incubation with EGCG. Three new regions of density become apparent with the addition of EGCG (Sites 13). Site 1 is located in the polar cleft at the intersection of the two protofilaments composing the PHF. Sites 2 and 3 of new density are observed adjacent to K343 and K347 near the -helix of the fibril. Both Sites 2 and 3 display weaker density than Site 1.

and vegetables. EGCG has been found to inhibit many types of amyloid fibre formation, and disaggregates Aβ as discussed above, α-SN and tau fibrils[161–164]. There is usually very little cryo-EM information because the drug needs to be bound uniformly down the fibril to build enough

electron density to show up. One can only image these types of binding events when the binding is uniform and tightly bound. EGCG is one such compound wherein, cooperatively, thousands of EGCG molecules bind in repeating units of a cleft in the amyloid structure. An example of one such process is highlighted in figure 1.13.[161]. Adjacent EGCG molecules form stabilising π stacking interactions with one another. This repeating unit provides enough electron density to be imaged with the same methods as the amyloid protein. Anti-amyloids destabilise the amyloid. Therefore once the EGCG has bound to the fibril, for example, the fibril itself begins to be disrupted, which will change the electron density. EGCG's mechanism of disaggregation[161] is proposed to drive a wedge between two protofilaments. The EGCG molecules then decrease the contact between adjacent protofibrils or even break some bonds, causing the fibril to disaggregate. Other studies highlight that EGCG is able to interact non-specifically with Aβ monomers[165], and remodel protofibrils into spherical oligomers[163]. Evidence also suggests EGCG also interferes with secondary nucleation processes which can lead to production of neurotoxic olgiomers[165].

Other polyphenols, specifically OLEA and quercetin have been found to interact with Aβ, Tau and α-syn[166]. They are neuroprotective and they able to modulate proteostasis machinery. Quercetin is found to counteract oxidative stress which can lead to neuronal cell damage in animal models. Quercetin is found in this study to prevent fibrillisation of several amyloid proteins. Another compound, OLEA, which is found in olive plants, is found to disrupt Aβ aggregation[167]. Studies show that OLEA is able to disrupt Aβ aggregation through disrupting π stacking[168] and is able to prevent the formation of toxic oligomers[166]. Again, their limited bioavailability due to the body's ability to metabolise these compounds is a limiting factor in their viability as therapeutics.

Many low molecular weight compounds ($<$ 500 Daltons) found to inhibit Aβ are able to cross the blood-brain-barrier (BBB)[169]. However, a great number of these are not able to reach the brain in significant enough concentrations to have a significant therapeutic effect. Compounds such as EGCG have great success in remodelling fibril and oligomeric structures *in-vitro*[161, 162, 170]. However, their low bioavailability in current approaches hampers any inhibition effect they incur *in-vivo*. Using natural compounds as a basis for anti-amyloid drugs is a very sensible choice, however the drug delivery mechanism may need adjusting so that these compounds reach the target

species with the necessary concentrations.

1.7.5 G3P

A general amyloid interaction motif (GAIM) is a protein domain that has been identified on the surface of a bacteriophage particle (M13) which can remodel amyloid structures[171]. GAIM constructs have shown promise in animal models and is currently in clinical trials[172].

M13 phage can remodel amyloid structures of various proteins, including Aβ, tau, α-syn and NM yeast prion domain. The M13 phage is a tiny molecule of single-stranded DNA wrapped in a protein coating. One of these proteins is G3P, which sits on the tip of the coat. It has two N-terminal domains N1 and N2, and there are usually $\approx$3 copies of G3P within a given M13. M13, or the G3P (Gene 3 protein) domain within M13, is key to amyloid binding and remodelling. The G3P protein is not very efficient on its own, but multivalent constrcuts rescue the activity displayed by the phage. The physiological role of these N1N2 domains within G3P may be to provide a surface to attach themselves to the curli fibrils, which form part of the biofilm around E-coli bacteria. Whilst a phage is a virus; it is not a virus that will attack the human body. In experiments where the M13 has been heated above a specific temperature, the N1N2 domains uncouple and come apart more quickly, which aids the anti-amyloid activity. M13 is a useful molecular biology tool that has been exceptionally well-characterised and has only recently been discovered to have anti-amyloid effects.

The G3P domain of M13 phage has been proposed as a potential drug treatment for several neurodegenerative diseases due to its ability to clear existing amyloid structures. G3P was proposed as an alternative to using the complete M13 phage as delivery of the phage and dosage control would be difficult.[171, 173]. Coupling to antibody constructs generates polyvalency and allows the manufacturing pipeline to be standardised. The M13/G3P protein indicates a high specificity for fibrillar conformations, and results indicated that interactions between the phage and amyloid-β aggregates did not result in free low molecular weight oligomers or monomers on the experimental timescale examined. This is because single monovalent G3P will be much slower, and is useful to investigate intermediates. There are, however, multivalent G3P constructs which are much more effective than G3P alone, and producing such constructs is much more efficient and commercially

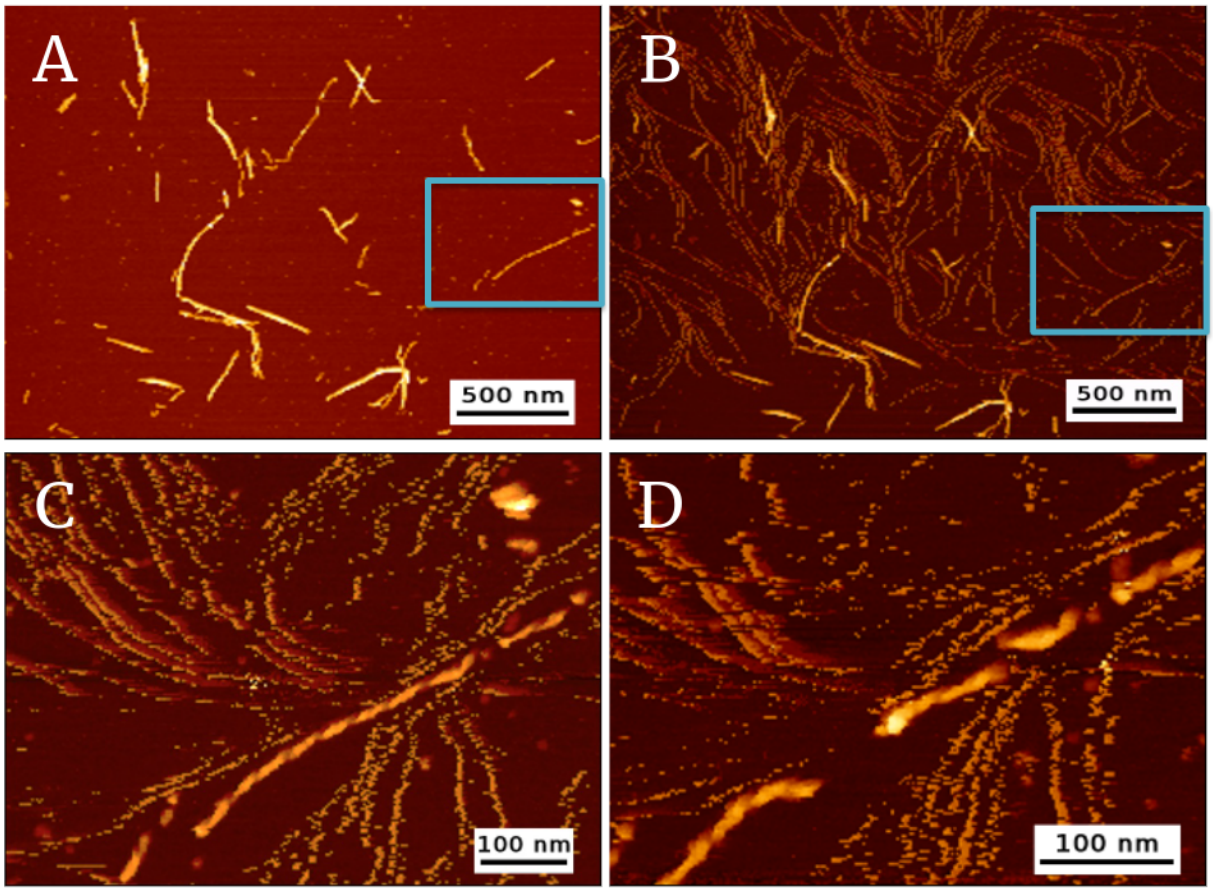

Figure 1.14: AFM images from the Staniforth lab of (A) Stable amyloid-β fibres imaged for 3 hours in pH 6.5 buffer in QI mode and remained stable with length scale included for 500nm. Specific fibre highlighted with a box. Added M13 phage to the Aβ for (B-D). (B) A 9-minute scan highlights low phage concentration around the selected fibre. (C) Subsequent 10-minute scan to (B) of the highlighted region only with increased phage activity in the region and visible twists on the amyloid fibre. (D) Further 10 minutes from (C) highlighting the broken fibre along its length.

viable than producing the entire phage[171].

1.7.6 Is amyloid the only valid drug target?

Anti-amyloid drugs form more than 20% of drugs in current clinical trials aimed at tackling AD, and similar statistics exist for Parkinson's disease. Simply using anti-amyloid drugs might be too specific, and similarly, using a drug which only targets one kind of amyloid might also be too specific. Alzheimer's disease is known to cause aggregation behaviour in tau proteins and even α-synuclein in some cases, which is typically associated with Parkinson's disease. We cannot simply treat AD by only tackling the initial aggregates formed. This approach might work as a preven-

tative measure to stop further amyloid aggregates from forming. However, this will not work for someone who already has mid-stage Alzheimer's disease and has multiple amyloidoses of Aβ, tau and α-syn fibrils. Aβ oligomers appear to affect a huge number of pathways, including tau aggregation into neurofibrillary tangles. There are too many "side effects" or downstream effects of amyloid aggregation behaviour that are issues that also need treating in their own right.

A possible treatment option is to target the unfolded protein response (UPR). An excess of misfolded protein (i.e. Aβ or Tau can activate the UPR.[174, 175] Inflammatory response all targets UPR. Excessive UPR signalling eventually leads to apoptosis, a long-term effect of UPR activation. The UPR cannot be targeted early in development because the UPR is needed for neurogenesis. Therefore, if we inhibit the UPR, we prevent regular gene expression and proper development. Drug screens from $\approx$ 1000 drugs to look for drugs that inhibit UPR, resulting in roughly 20 suitable compounds. Cell culture assays[176] looked at the impact of their compounds, which resulted in two drugs that inhibit the UPR currently in clinical trials for AD treatment. It is easier to screen drugs that already exist and are approved for treating other diseases or conditions rather than moving through the whole clinical trial phases from scratch because it is a slow and laborious process.

Frontotemporal dementia (FTD) is caused by tau deposits without the presence of Aβ. Typically occurs earlier than AD and is easier to see in mouse models. Mice can recover from prion and FTD despite having high volumes of amyloid deposits in the brain. In prion disease, there is no clearance of amyloid, whereas, in FTD, there is evidence of a reduction in phosphorylation of Tau, which leads to amyloid deposition. Tau might be downstream of UPR. i.e. We can prevent Tau aggregates by targeting the UPR[177].

Therefore, we either need to tackle AD with a multi-pronged approach to treatment or find an inhibitor that will tackle multiple fibrillar species, or simply get better at targeting the neurotoxic elements of the aggregation process. i.e. The neurotoxic oligomers, which are doing damage inside the brain. Alternatively, one better: if we can prevent amyloid formation from ever occurring, we can prevent the disease. However, this is unlikely to be feasible with current knowledge of the dis-

ease. While prevention is an excellent idea in theory, unless it is possible to identify who is at risk of developing AD or other neurodegenerative conditions, it would mean that the entire population would have to take anti-AD drugs from their 30s-40s onwards. Therefore, a combinatoric treatment approach for those who already have the disease is likely to be the most successful with current knowledge and technologies.

1.7.7 Summary

In summary, we have identified key areas of research in amyloid inhibition with a key focus on Aβ and the Alzheimer's pathology. A combinatoric therapeutic approach needs to be used to tackle downstream issues in AD patients as well as the root cause. Getting better generic anti-amyloid drugs would be incredibly useful, as well as a more fundamental understanding of how to impact many different areas of the assembly kinetics. Whilst targetting secondary nucleation is a credible option due to the toxicity arguments, it is not the be-all and end-all of inhibiting amyloid assembly.

There are plenty of promising options in amyloid inhibition. There are several viable naturally occurring small molecules such as EGCG[161, 165] and OLEA[160, 166] which are found to have extensive anti-amyloid effects by remodelling fibrillar species and inhibiting secondary nucleation events. However, due to their limited bioavailability it is not currently possible to deliver doses at high enough concentrations to enact their anti-amyloid effects[169].

There are also several antibody inhibitors which inhibit secondary nucleation behaviours in a similar fashion to the chaperone protein BRICHOS which is neuroprotective. Aducanumab[130] is the most promising of these inhibitors which binds to fibrillar Aβ and prevents secondary nucleation. The issue with these antibody drug candidates is that they appear to prevent further aggregation specific to Aβ, but do not tackle any of the downstream effects of Aβ fibrillisation. i.e. Further amyloidoses of tau or α-syn, and issues with signalling pathways in the brain.

G3P is the most generic protein based anti-amyloid which has been found to interact with many amyloid proteins[171]. It highly specific to fibrillar conformations, and therefore will not dis-

rupt functional Aβ[171, 173] or other monomeric amyloid peptides. However, comparably to the molecules such as EGCG and polyphenols, more development is required to ensure there is an effective and efficent dosage.

It is clear that there is limited mechanistic information surrounding the functionality of the majority of inhibitors currently available. By creating a simulation model capable of producing hypothesised inhibition mechanisms, we could gain insight into the mechanisms of inhibition of available inhibitors, and potentially inform future drug design of anti amyloids. Such a model would provide structural information that cannot be determined through current experimental methodologies.

1.8 Molecular Simulation Techniques

In recent years, with exponentially increasing computational power, more researchers are utilising computational approaches to explore their hypotheses. Computational or *in-silico* methods allow us to make observations that are simply not possible through experimental methods. For example, in the case of amyloid proteins, small aggregates are difficult to observe in terms of their structure and dynamics[18, 178, 179]. Therefore, many have taken to using computational models to observe such species. Similarly, whilst we can observe fibrillar structures using cryo-EM and NMR, we cannot observe their formation structurally. Aggregation is also difficult to observe in computation due to the extremely high numbers of particles required for simulation. In such cases, we reduce the degrees of freedom by using a more simplified model. In this section, we will provide a non-exhaustive overview of the available simulation techniques and their applications to amyloid formation and inhibition by providing examples.

1.8.1 Length and Time scales

The choice of modelling technique used is solely dependent on the scientific problem at hand. Different modelling techniques naturally have their pros and cons, but the length scales of the objects and the timescales of the effects in question inform the technique used.

Almost all molecular simulations can be described by the categories highlighted in figure 1.15:

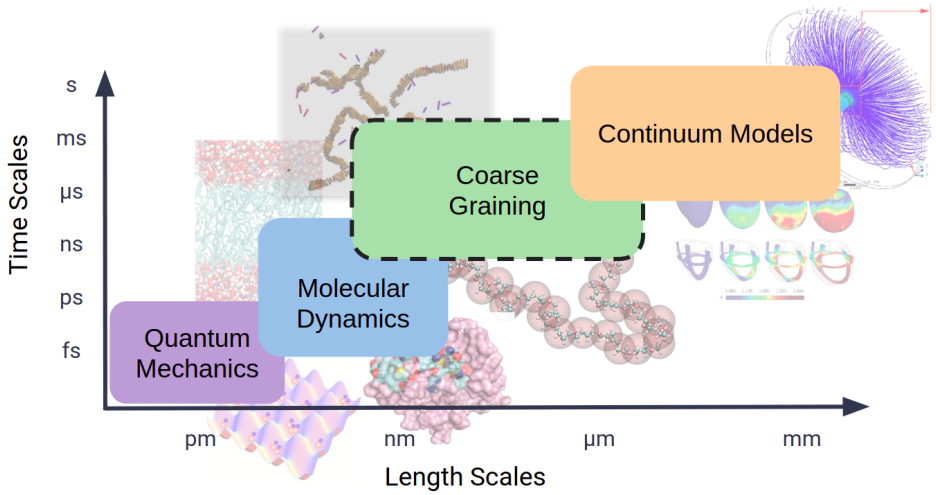

Figure 1.15: A representation of approximate time scales alternative simulation techniques cover and their corresponding length scales.

quantum mechanical calculations, molecular dynamics, coarse-graining and continuum models. All of these techniques have their benefits when simulating specific systems. With current technologies and resources, it is simply not possible to use high-level quantum calculations on anything larger than small molecules or regions of larger molecules[180]. As explained by the title, when using quantum techniques, we are primarily interested in processes on the quantum scale, which are often extremely short-lived on the picosecond ps scale or even the femtosecond fs scale. Therefore if we are interested in fibrillisation processes where the monomer is on the nanometer nm scale and the aggregation process occurs at the very least on the nanosecond μs scale, such as fluctuations in beta-structures, using quantum mechanics is not feasible or realistic here. At the other extreme, continuum models are often used in materials science, as well as to simulate processes on the biological cell scale or larger (a minimum of micrometres μm). Using these techniques, we lose too much of the fine-grained molecular detail of the amyloid protein monomers. A continuum model may be appropriate for simulations of mature fibrils but not for modelling the aggregation process.

Boundary conditions in simulations are also necessary. The most commonly utilised are periodic. Periodic boundaries negate most boundary issues as biopolymers would not feel any boundary effect as it acts as if the simulation is a volume element from bulk media. It would simply pass through a surface and appear on the opposite side. A limiting factor here is that the box needs to be sufficiently large that opposing ends of a large molecule are not interacting with one other.

Similarly, if there is some boundary condition, the box also needs to be large enough that said condition does not significantly alter the dynamics of the simulation.

Molecular Dynamics[181] is a broad term used to describe simulations of particles over some time period by calculating trajectories with Newton's equations of motion. There are multiple subcategories under the umbrella of molecular dynamics, which can affect the algorithms used to calculate particle trajectories, what is included in the simulation and what is simulated. It is also possible to drive simulations in a specific way. i.e. The user could fix the positions of particles or apply some additional external force to specific particles in a simulation to observe behaviour in specific scenarios. In particular, NAMD[182] and GROMACS[183–188] are known for their Steered Molecular Dynamics (SMD) features. Molecular dynamics is typically a descriptor for all-atom or atomistic simulations, but it is also possible to carry out grained simulations[180, 189] - each of these are discussed in further detail here.

1.8.2 Molecular Dynamics

Recent advances in computing techniques, such as the implementation of simulation software with GPUs rather than CPUs, have made it possible to use increasingly large particle numbers across increasingly large simulation timescales. Many of these techniques are only available to a very restricted number of research groups with considerable computing resources, often with complex algorithms that implement simulations with high levels of parallelism. Groups[190] have begun simulating entire large complexes or complete chemical systems rather than small sections of proteins, which has been a common focus of molecular dynamics simulations in recent history[191].

For the case of amyloid proteins, these computational advances are not significant enough to justify the simulation runtime for the number of particles available to simulate hundreds or thousands of monomers in all-atom simulations with solvent included. Therefore, many groups choose to simulate essential fragments of a given protein to understand the aggregation process better or instead reduce the number of degrees of freedom a particular monomer has via a technique called coarse-graining.

1.8.2.1 Force Fields

In molecular dynamics simulations, force fields model interactions between particles or atoms. These can vary from modelling specific interactions, such as harmonic covalent bonds, to continuum modelling for solvents. Many molecular dynamics packages have force field packages associated with them. Some commonly used examples of such software are CHARMM[192], Amber[193], LAMMPS[194] and GROMACS[183–188, 195, 196], which all have their own force fields with specific parameter sets dependent upon the type of molecules which are simulated.

A huge amount of scientific work is placed in parameterising all atom forcefields in molecule dynamics. An alternative technique called "coarse-graining" simulations, proposes to use custom potentials depending on the level of coarse-graining. Some groups encompass multiple potentials into one potential depending on the model's assumptions and simplifications. All interactions can be characterised into bonded or non-bonded interactions, as shown in equation array (1.2) below.

$$\begin{aligned} V &= V_{BONDED} + V_{NON-BONDED} \\ V_{BONDED} &= V_{STRETCH} + V_{BEND} + V_{DIHEDRAL} \\ V_{NON-BONDED} &= V_{COULOMB} + V_{VDW} \end{aligned} \tag{1.2}$$

Coarse graining reduces the complexity of the atoms and/or the interactions between them. Using a simplified approach trades complexity for computational expense. Intramolecular contacts can be modelled by potentials for bond stretching, bending and dihedral angles. Bond stretching is represented by a 1D Hookian spring potential along the axis of adjacent particles or a finite extensible non-linear elastic potential (FENE bond)[197]. Bond bending is typically a potential with angular dependence, which contributes toward rigidity, steric exclusion and reduction in conformational space. In protein chemistry, dihedral angles along the protein backbone play a huge role in conformational protein structure and steric exclusion. These two features limit the conformational phase space available to a protein. Dihedral angles are often separated into dependence upon torsion and improper dihedrals in simulation software[182–188, 192–195].

Long-ranged interactions can be difficult and computationally expensive. Multiple algorithms can be used in MD software to calculate electrostatics and Van der Waals forces in the least computationally expensive manner. Examples include the Particle Mesh Ewald (PME) algorithm[198] and Verlet integration[199]. PME works by separating one slowly converging sum into two rapidly converging sums. Van der Waals forces can be approximated using various methods but are typically modelled with either the standard Lennard Jones[200] potential or the Buckingham potential[201].

Hydrogen bonding is a notably tricky phenomenon to simulate due to its highly directional nature and is primarily electrostatic [53]. Hydrogen bonding has the largest cumulative effect in solvent interactions and plays a significant role in aggregation kinetics[202] due to cooperativity in hydrophobic burial[203, 204].

1.8.2.2 Atomistic Simulations

In amyloid research, many choose to only simulate small regions of the protein as the monomer rather than the entire sequence to reach longer time scales whilst compensating for full molecular detail. By focusing on key regions in a given protein sequence, we can negate the computational expense of the rest of the protein. Only using a section of a protein or peptide naturally has the downside that the simulation may culminate in behaviour not typically observed experimentally without the presence of residues that would usually drive a protein to find other conformations.

Schor et al.[205] simulate amyloid fibril formation at different pH values using atomistic structures of TTR(105-115), finding multiple fibril formation pathways which vary in weighting depending on pH. Baftizadeh et al[206]. simulate the monomers of the C terminal region in Aβ and the formation of fibrils, highlighting rate-limiting steps in nucleation and slow timescales in the kinetics. Nguyen and Derreumaux[207] utilise REMD of the 6 C terminal residues in Aβ, highlighting a variety of oligomeric structures. The evaluation of β sheet content of such structures shows that there is packing polymorphism in strands rather than just the in-register parallel strands observed experimentally. Baumketner and Shea[208] also find protofilament polymorphism through simulations

of a central region, Aβ(12-28), in explicit solvent. Most of the complete protein is missing from these simulations, which would likely further stabilise parallel strands. Simona et al. [209] evaluate secondary structure content in the same region in all-atom MD simulations, observing that hydrophobic contacts are important in the aggregation of β hairpin structured monomers. However, hydrogen bonding and electrostatics provide stabilisation in said structures[210]. In tetrameric form, simulations of varying segments of Aβ peptide[211] show that non-fibrillar oligomers are more stable in longer strands of or complete sequence Aβ. The stability of these structures is strongly correlated with the length of the C terminal section. Secondary structural characteristics depend heavily on the primary sequence of residues[212, 213]. Simulations of the (16-22) region of Aβ[214] show that monomers tend to a helical state, but multimers prefer β-sheet conformations. There is little experimental evidence to support these claims. Therefore the helical content may be the result of a poorly optimised forcefield. Simulation forcefields commonly overpredict secondary structures. Overproduction of helices may result from overweighted electrostatic interactions. Meinke et al. also show that these regions prefer to form antiparallel beta sheets, which is not typically observed experimentally for the unmodified peptide but is seen for the same Aβ fragments, suggesting that this is the result of only using a small segment of the complete sequence. There are experimental data of the same Aβ fragments, which also evidence a tendency to produce antiparallel sheets. Simulating or carrying out experiments on short fragments allows us to understand critical regions of proteins better. However, if they do not replicate the same behaviours as full amyloid sequences, perhaps a different approach to simulation is necessary.

REMD of low molecular weight oligomers find β barrel structures[215]. The results suggest that β-barrels are a common theme in aggregation pathways of amyloidogenic proteins. β-hairpin structures in small proteins are found to straighten out and form cooperative, interchain hydrogen bonds during aggregation[216]. Whilst we can generalise the properties of amyloid formation to some extent, the specific dynamics of amyloid peptide oligomerisation are sequence dependent[217] and largely dependent on the level of hydrophobicity in a specific peptide.

Xu et al.[218] simulate Aβ(9-40), which forms oligomer structures that become more fibrillar with additional monomers. The monomer favours some random coil conformation. Xu also simulates

Aβ(40) with Zn2+ in different conformations, finding that Aβ oligomers in β-hairpin conformations stabilised with Zinc could be important in the fibrillisation of Aβ. Yu and Zheng[219] assemble a variety of oligomeric structures (globulomers) from monomers and dimers of Aβ and observe their structural stabilities to determine which could be off-aggregation-pathway conformations in the production of amyloid fibrils. Such structures are compact with curved surfaces and minimal β sheet content and have been observed in poor solvents[220]. Conversely, dimers are found to have higher free energies than their monomeric counterparts[75] and are not considered to be included in the pathogenesis of oligomers. Gurry and Stultz[221] show that Aβ(40) and Aβ(42) both share β-hairpin intermediates which bind to fibrils and hypothesise that fibrils could serve as secondary nucleation sites which catalyse the formation of soluble oligomers[221]. Barz et al.[222] also compares Aβ(40) and Aβ(42) by producing free energy landscapes for both alloforms. Barz finds that the 42 residue alloform occupies a broader variety of conformations, but both proteins produce very similar energetic landscapes. The reduction in the number of residues leads to a shorter C terminal region and reduces the β-hairpin formation[223] and, therefore, the stability of oligomeric structures[224–226]. Zheng et al.[224] show that solubility in Aβ(42) is 10x smaller than in the 40 residue alloform, with a steeper gradient in the free energy profile. Solubility depends on the thickness of the fibril and other factors such as temperature and backbone hydrogen bonding[227, 228]. Auer concludes that the concentration dependence of the nucleation rate of fibrils is linked with the solubility of a protein[121]. Lin et al. [229] also include monomers of the Italian mutant Aβ(42)-E22K for comparison and find that regions have a high α-helix propensity which results in changes to the kinetics.

Stability of protofibrils in Aβ(42) and other superstructures is found to depend on salt bridges in the core of the structure[230, 231]. The level of packing in the core is also used to modulate the level of solvent in the structure's core, allowing for less electrostatic screening and strengthening the Asp23-Lys28 salt bridge. Hydrophobic interactions are essential in the stability of proteins in the burial of hydrophobic residues and the accessibility of hydrophilic side chains to solvent. Electrostatics can vary in importance depending on the SASA surrounding charged residues[232]. Strength and stability with ionic strength can be observed by increasing the salt concentration in the solvent surrounding Aβ. Zidar and Merzel[233] find that electrostatic interactions result in

polymerisation occurring at the fibril ends only with low salt concentration. At high salt concentrations, hydrophobic effects dominate, allowing for polymerisation on all sides of the fibril. Soluble proteins which have unshielded electrostatic interactions are prone to form aggregated species[234]

In silico experiments of oligomers comprised of fragments of Aβ(17-42)[235] show that oligomers in conformations similar to fibrils can act as nucleation sites as only minimal energetic expense is required to find fibrillar conformations. Horn and Sticht[235] suggest that Aβ dimer toxicity could be explained by its highly stable structure, which is difficult to incorporate into fibrils. Han and Schulten identify the strand-loop-strand conformation of Aβ monomers to be important conformational intermediates[236]. One fibril end has exposed L17-A21 residues, hypothesised to explain the experimental preferential growth end. Rojas et al.[237] find that aggregation cannot occur unless an α-β transition occurs across the whole chain. In their simulation, residues 17-21 are forced into an α helix conformation, which prevents the aggregation process but allows for forming Aβ dimers. Analysis of β-sheet content through MD of Aβ oligomers[238] suggests that the trimer could be the minimum size super nucleus for the continued growth of filaments.

Zheng et al.[77] simulate atomistic Aβ(40) monomers to octamers with a coarse-grained forcefield [AWSEM]. A free energy landscape is constructed via umbrella sampling to identify the conversion pathway between pre-fibrillar species and mature fibrillar structures. Perturbation theory allows insight into the changes to aggregation pathways made by small mutations in Aβ, i.e. In the Dutch and Arctic mutations. Lu et al.[239] observe changes in stability and the aggregation kinetics in mutants of Aβ(29-42), benchmarked against the wild-type.

In the case of modelling ion channels, it is sensible to use all-atom molecular dynamics as the scientific questions here require fine-grained structural characteristics. Often in cases such as this, the driving behaviours lie in crucial structural changes that need to be observed on the all-atom scale. Jang et al.[240, 241] model Aβ ion channels in lipid bilayers using NMR-determined structures for the Aβ monomer. Ion channels were found to present the highest toxic ion flux when made up of 16-24 monomers. Jang[242] also determines there are multiple aggregation pathways by simulating Aβ oligomers and observing the variety of different structures. Similarly, Strodel et al.[243] deter-

mine stable oligomeric structures of Aβ(42) in the lipid bilayer, suggesting that membrane pores may be built up from tetrameric and hexameric subunits with β-sheet structures.
Whilst atomistic simulation techniques are an incredibly useful resource to groups investigating biomolecular systems, in the case of the amyloid field, it is incredibly computationally expense to simulate the necessary number of atoms to observe the aggregation process in this way. The findings of the research discussed in this section show that molecular dynamics still has many uses within the amyloid field. However, these use cases are far more specific to looking at specific aspects of aggregation, cytotoxicity and oligomer formation with relatively low particle numbers. An alternative technique must be used to simulate aggregation on a larger length scale.

1.8.3 Monte Carlo Techniques

The foundation of most Monte Carlo simulations is the Metropolis-Hastings algorithm wherein a potential new state is generated, and using some ratio of the probabilities of finding a system in the current and new states, we can accept or reject this new state based upon an acceptance ratio[244]. This algorithm is most commonly used in Monte Carlo simulations for lattice-based models. A similar algorithm is the Gillepsie Algorithm[245, 246] which is used on inherently stochastic systems where it generates variable time steps and changes to a system. This algorithm is also known as Kinetic Monte Carlo, a subcategory of Monte Carlo simulation used primarily for natural processes with known transition rates between states.

Monte Carlo encompasses a variety of random sampling techniques[244, 247] used to minimise or maximise a given property in a system. Frequently, they are used to sample a system's parameter space using rejection hypotheses dependent upon energy calculations with Boltzmann probabilities. Put more simply, Monte Carlo simulations are used in complex systems with many degrees of freedom to sample the available parameter space. Simulations of this type are used widely for biopolymers to sample many states. However, it is limited by cumulative integration errors in simulations with large numbers of time steps.

Many groups utilise Monte Carlo type simulations in their research[87, 88, 113, 118, 204, 210, 213, 214, 227, 248–251]. Cabriolu et al.[118], for example, use a lattice model with nearest neighbour

interactions to model nucleation in 2D crystals. Zhang & Muthukumar[248] originally applied the same model to nucleated polymerisation in amyloids. Their simulations showed the typical lag and growth phases observed experimentally in fibrillisation, including the phenomenon of Ostwald Ripening after the growth phase[95]. They[248] also showed that the lag time could arise from two different mechanisms.

Another simulation style, Replica Exchange Molecular Dynamics (REMD), utilises Monte Carlo style sampling among multiple very slightly different simulations via the Metropolis algorithm wherein the acceptance ratio is the likelihood of two systems exchanging coordinates. An example of the use of REMD is to make configurations in a system at a high temperature available to the same system at a lower temperature. Jang and Shin[242] use REMD in an atomistic study of Aβ(10-35) i-mers to illustrate the variety of oligomeric structures, showing that these structures are likely to be intermediates in specific aggregation pathways for Aβ. Rojas et al.[252] observe a two-step dock-lock mechanism throughout a variety of assembly pathways in coarse-grained Aβ. A dock and lock mechanism is where some conformational change is required before a monomer binds fully or "locks" into a structure. REMD uses rare event sampling - a technique used to sample chosen regions of a parameter space which would otherwise rarely occur. Stochastic processes often benefit from this type of sampling as it negates unnecessary computational expense. Nucleation is another example of a rare event relative to other time and length scales in the motion of Aβ monomers[253] where REMD has been utilised.

1.8.4 Coarse Graining

Whilst atomistic simulations are beneficial, they are incredibly computationally expensive. Atomistic data is often used to inform force field models for coarse-graining[254, 255]. Simplification can benefit simulations; by trading computational expense for model detail, we can observe a system's key features. Coarse graining (CG) reduces the number of degrees of freedom a system has by representing molecules in a simplified way. Figure 1.16 highlights different degrees of coarse-graining of a simple protein structure. The CG representation can be executed on a variety of length scales, from representing small functional groups to entire proteins with singular particles. Figure 1.16(A) shows a bead representation where singular particles represent sections of the protein backbone

and functional groups. Figure 1.16(B) highlights a model with more degrees of freedom where the backbone uses separate pseudo-atoms for the functional groups/side chains. Finally, figure 1.16(C) highlights a model design where a singular pseudo-atom represents the entire peptide. The most suitable degree of coarse-graining depends on the system in question. Sometimes the most simplified models can still explore a significant portion of the parameter space for a given system by negating intramolecular detail, such as the one we present later in this thesis. Similarly, we can also coarse-grain the force field to a more simplified version of what is used for an atomistic simulation to account for the simplified nature of the system. This scaling and removal of certain interactions can drastically reduce computational time for running simulations.

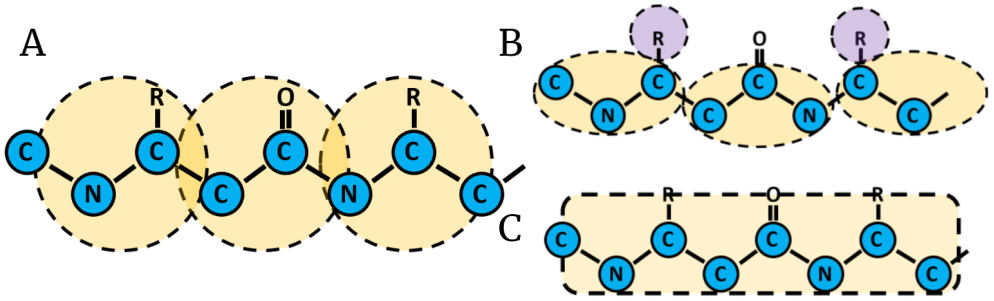

Figure 1.16: A bottom-up coarse-grained representation of a peptide where pseudo-atoms represent groups of atoms. Hydrogen atoms are omitted from this figure for simplicity. (A) Representation where a single pseudo-atom encompasses sections of the protein backbone and side chains. (B) A representation where the backbone is represented by pseudo-atoms and the side chains are also represented by an additional species. (C) A representation where a single pseudo-atom or particle represents larger regions of a protein.

Coarse-graining has many benefits but also some significant faults. By simplifying a model and - by definition, reducing the number of degrees of freedom, we cannot expect to observe fine-grained details of proteins such as in ion channels[240, 241]. Interaction site per residue models are popular - notably in large biomolecules. Another standard method is to use multiple beads per residue to model the protein backbone effectively whilst accounting for sidechain mobility[256–258].

An example of a commonly used coarse-graining regime is MARTINI[256, 257] - a package used to simplify protein structures (.pdb files) into coarse-grained models - the outputs are coordinate files to be used in molecular dynamics packages. It roughly uses four-to-one mapping wherein four

large atoms are modelled as one site[258]. Coarse-grained representations of the amino acids are shown in figure 1.17. This package is used to simplify atomistic structures in conjunction with its forcefield. Extensions of this model exist, notably one suited to proteins with amyloid-like or elastic-like properties[259].

Another method of reducing the degrees of freedom in a simulation is to keep the simulation fully atomistic but coarse-grain the interactions between particles/atoms. This is done to remove unnecessary computational expense without reducing the conformational phase space of the protein.

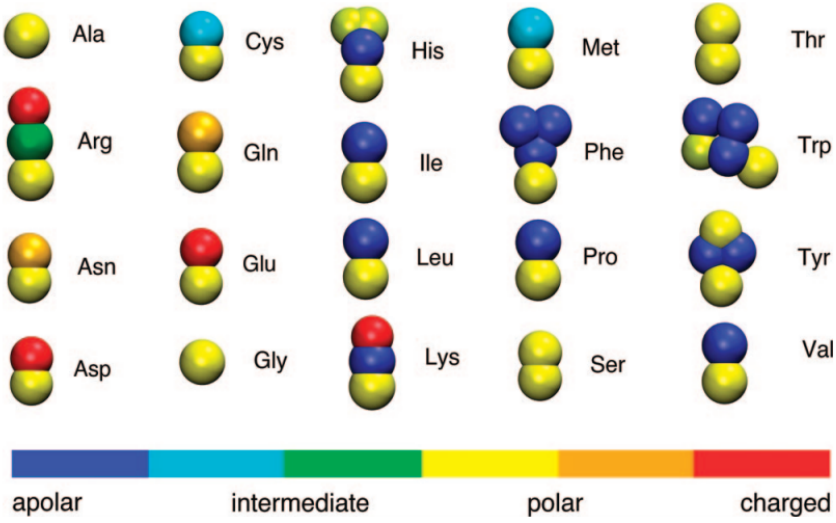

Figure 1.17: Coarse-grained representations of common amino acids using an applicable protein extension of the MARTINI package wherein each type of particle is represented by a different colour. Reproduced from Monticelli et al.[258]

Different approaches to coarse-graining can be taken depending on the aims of the simulation and the biological questions one might be looking to answer. Figure 1.18 provides an overview of the various top down or "minimal" coarse-grained models used to tackle different subsets of biological problems. Taking a known structure from an atomistic level to some more simplified level, building a generalised model to represent specific physical characteristics without taking into account chemical features of any specific protein, and a third knowledge-based approach whereby a model is built based on the structure and native conformation. It is often favourable to keep a model as general as possible when working on whole types of protein - as a result of this, there is a variety

of literature focused on the properties of amyloidogenic proteins[3, 113, 250, 260–274]. This allows us to address interesting questions regarding the general determinants of amyloid formation. Conversely, many groups are working on Aβ to outline the chemical pathology of Alzheimer's disease, therefore, taking the alternative knowledge & structure-based approach with physical properties specific to Aβ.[75, 77, 225, 239, 252, 275–280]

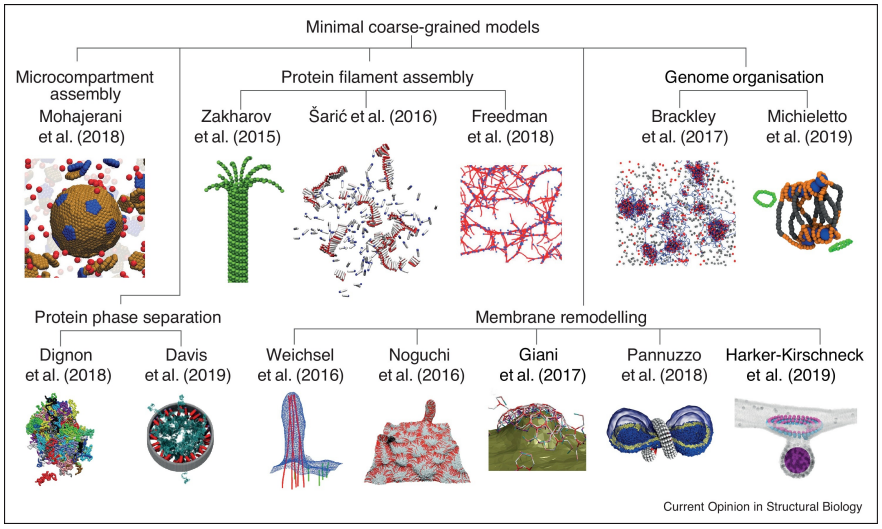

Figure 1.18: Top-down coarse-grained models taken from a review from Hafner et al.[281]. The models are separated into microcompartment assembly, protein phase separation, protein filament assembly, membrane remodelling and genome organisation.[113, 282–287]

Hu et al.[276] present a coarse-grained model for Aβ(42) made up of a protein backbone and beads representing side chains. This model produces amorphous oligomers and later forms large amorphous aggregates. Theory and experiment predict conformational conversion or reorganisation into a fibrillar structure at a later time, but this is outside the scope of the simulation in this case[276].

Bellesia and Shea[267] highlight this reorganisation method[276] as one of three main aggregation pathways using a coarse-grained model of Aβ. A second pathway involves directly aggregating into a fibril from a monomer. The third includes some oligomeric intermediate - both of these are observed in a series of other *in silico* CG experiments wherein β-sheet forming propensity is altered (represented by small changes in free energy, which is physically equivalent to changing a protein sequence slightly[267]). By including a fibrillar seed and varying the β-sheet propensity of

the free monomer, Bellesia et al.[272] show that a fibril can polymerise through two mechanisms. For low β-propensity, a dock and lock mechanism exists where a conformational change is required before a monomer is fully bound to a fibril. In contrast, for high β propensity, the growth occurred directly without conformational rearrangement. The dock and lock mechanism is found to be the dominant assembly process[288]. This work is extended to aggregation on lipid bilayers wherein a fibril structure will cause local structural change in the surface[273]. The conformation of the fibril will also be different to if it was produced on a solid surface.

Bellesia et al. have also produced a coarse-grained model[267, 268, 289] which is a generic representation of a short peptide with two interaction sites per residue and where residues form an alternating pattern of hydrophobic and hydrophilic bonding. Each end is also charged to represent the termini. Bellesia et al. perform a series of simulations and analyses to characterise the model and to establish insight into the properties of β sheet fibrils[267, 268, 289]. They observe and characterise a variety of different morphologies which are able to produce oligomeric structures discussed in the literature. They also discuss similarities to phase changes in liquid crystal systems.

As mentioned previously, a common simulation technique in CG modelling used by many authors is the varying of β-sheet forming propensities in their respective models. This allows them to show different reaction pathways to fibrillar species. Pellarin, Caflisch & Bieler [249, 251, 265, 267, 272] show that by varying beta-sheet propensity, there are two distinct reaction pathways to fibrillar species. High propensity yields a direct one-step pathway from monomer to fibril, whereas lower β-sheet propensities lead to observations of oligomeric intermediates before converting into fibrils. Free energy profiles of aggregation of amyloids using a 2-state model suggest that the size of the nucleation barrier determines the size of the populations of different fibril polymorphs[249]. Mutations that would reduce β propensity would increase the frequency of oligomer formation, which would match data for mutations such as the Arctic mutant in Aβ, for example. Similar results are found in papers proposing a two-state model[3, 113]. Saric et al.[113] use a patchy spherocylinder model (discussed in more detail in the following paragraphs) to determine that the rate-limiting step in aggregation takes place on the fibril surface, to show that primary and secondary nucleation are, in fact, separate processes. Ilie et al.[290] use a coarse-grained 2-state patchy particle to model

aggregation of α-synuclein in Parkinson's disease - we can draw parallels between the two due to the different nucleation pathways observed in such simulations, i.e. one and two-step nucleation.

Xu et al[279]. use a per residue CG model of Aβ(17-42) to establish oligomeric species and insight into secondary nucleation processes. They take a "one pot" approach from various starting points. i.e. no seed, seeded, monomers with a β-sheet secondary structure motif and monomers in a random coil. Each monomer within a given fibril forms a β-hairpin motif. Urbanc et al.[278] present a four-bead per amino acid protein model. They compare Aβ(40) and Aβ(42) to observe structural differences in oligomers, observing a unimodal size distribution for Aβ(40) and a trimodal distribution for oligomers of the 42 residue alloform. The extended C terminal region in Aβ(42) plays a role in the oligomer structures, whereas Aβ(40) structures appear to be driven primarily by hydrophobics[275].

Nguyen and Hall[254] use a Protein Intermediate-Resolution Model (PRIME)[291] to coarse-grain polyalanine chains and observe fibril formation via an oligomeric intermediate. Further work includes force field parameterisation, which can be used for on-lattice protein models[280]. Here, they present a lattice-based model of Aβ(16-22), which explicitly incorporates hydrogen bonds and directions of side chains. Nguyen parameterises the force by fitting it against existing atomistic data. This process is useful because it extends the timescales a simulation can undergo by approximating behaviour to existing simulation data, so a more considerable amount of given parameter space can be explored with the same level of computational expense. They characterise the energy landscapes for aggregates, which agree well with crystal structures.

Bereau and Deserno[269] use a more coarse-grained approach, presenting a proof of concept coarse-grained model with no bias to either secondary structure. The model coarse-grains to four beads per amino acid, allowing for many local conformations. An effective nearest neighbour dipole interaction is included. This system is tested by simulating a three-helix bundle with an implicit solvent. An alternative rod model for peptide self-assembly is presented whereby each molecule is modelled as a rigid rod which is decorated with spherical bead binding sites along its length.[271]. Mondal et al. conclude that the strength of short-ranged interactions can alter the self-assembly process

drastically, and the size of the binding site beads can affect the steric confinements of a system. Whilst these models are not applied to neurodegenerative disease directly, the model results in the self-assembly of fibrous structures. It is also important to consider the key structural factors that drive self-assembly behaviours.

Using a self-avoiding tube model - a more coarse-grained approach, again - Auer et al.[263] show that oligomeric intermediates form in simulations with more hydrophobic monomers and higher monomer concentrations. In contrast, monomers aggregate directly into fibrils at lower concentrations. Other works[260–262] suggest that oligomeric intermediate formation is not mandatory, but some oligomer structures can serve as nuclei for fibril formation for more hydrophobic peptides. This group characterises β-sheets and therefore amyloid structures as the most energetically stable[261].

The "patchy spherocylinder" (PSC) model, where cylinders with hemispherical caps are used with attractive patches, is commonly used as it is effective yet simple. Each monomer of amyloid protein (or peptide) is represented by a spherocylinder, with varying states dependent upon the publication in question. Publications from Daan Frenkel and Robert Vacha[4] introduce the spherocylinder model with two different patch morphologies as highlighted in figure 1.19. Each of these patch morphologies produced distinctly different aggregate morphologies after sampling parameter space of the patch width (proportion of the surface area covered by the attractive patch region) and aspect ratio (ratio of the length and diameter of a spherocylinder). The model on the right-hand side of figure 1.19 goes on to form the basis of future models[3, 251, 270, 292]. Robert Vacha[270] presents a similar, further developed model, which includes a coiled state and a fibril-competent state to observe the effect of surfaces on aggregation kinetics. The effect a given surface has depends upon the association propensity of a given peptide's surface relative to the bulk propensity and the concentration of protein available. i.e. an attractive surface accelerates fibril growth for low fibril forming propensity but retards growth for high propensity[270]. Bieler achieves the characteristic sigmoidal growth curves for mass concentration via Dynamic Monte Carlo simulations of a two-state spherocylinder model[251].

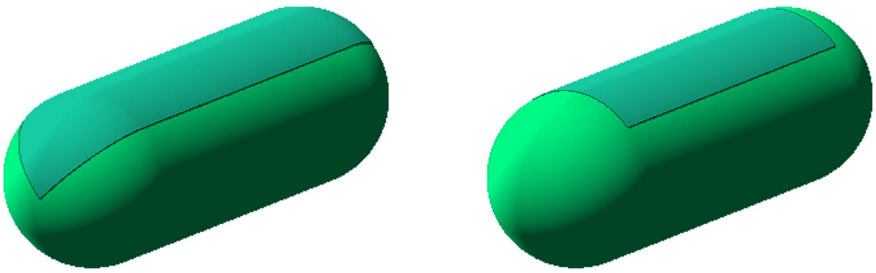

Figure 1.19: Spherocylinder models reproduced from Vacha et al. [4] Graphical representation of the two patchy spherocylinder models. (Green) The part of the surface that interacts as a hard spherocylinder. (Blue) Attractive patches. (Left) PSC-AE (attractive endcaps). (Right) PSC-NE (nonattractive endcaps).

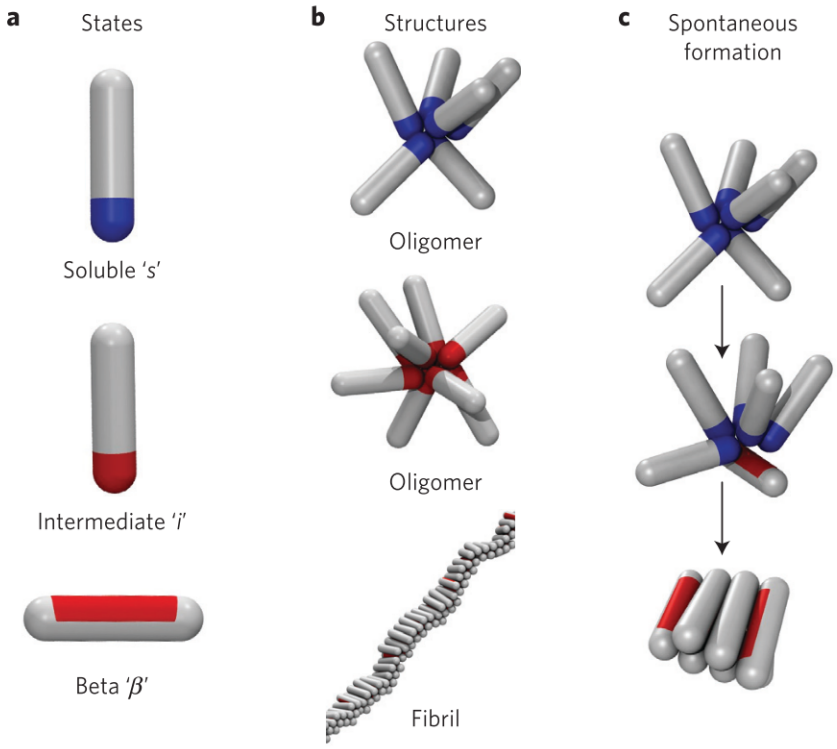

Figure 1.20: Spherocylinder model reproduced from Saric et al.[292] (a) A protein is allowed to exist in three conformations. From top to bottom: soluble state ('s'), intermediate conformation on the fibril surface ('i'), and the β-sheet-prone state ('β'). (b) Aggregated proteins. From top to bottom: oligomer made of soluble proteins, oligomer made of proteins in the intermediate state, and the fibril made of proteins in the β-sheet-prone state. (c) Primary nucleation takes place in two steps. Soluble proteins form finite oligomers (top), which can convert into a nucleus rich in β-sheet (bottom) that continues growing.

Saric et al.[3, 292, 293] show that fibril formation is a two-step process which occurs via the formation of oligomeric intermediates. Their model uses attractive patches on spherocylinders of two states with a fixed aspect ratio $AR = 4$. Figure 1.20 depicts the three different states utilised

in their model. The first is a soluble monomer state with an attractive patch on one cap of the spherocylinder. The second is a separate β-sheet-like state where there is a 180° wide patch along the cylindrical portion of the spherocylinder. This patch morphology is an extension of the right-hand side model presented by Vacha and Frenkel[4] in figure 1.19. The formation of oligomers and fibrils is controlled by the strength of said non-specific interactions in the model and the transitions between them. Saric also suggests that globular oligomers facilitate the conformational change in monomers to a β sheet form suitable for fibril growth[3]. The model is then developed further[292], which introduces an intermediate state which facilitates secondary nucleation events, discusses one- and two-step fibril nucleation mechanisms and oligomer production[294, 295].

At the most significant coarse-graining, we use continuum models. Extracting material properties of fibrils and other protein structures can be complex experimentally. Approximating fibrils as a continuum model leads to insight into material properties[296]. Similarly, CG Normal mode analysis can be used to measure structural properties[297]. Rate constants and structural information can inform mathematical modelling and experiments by simulating the filament depolymerisation and fragmentation[274].

We conclude that modelling the amyloid monomer with a single particle such as a rod has a unique ability to simulate the key elements of protein polymerisation processes such as amyloid formation. The aspherical nature of the particle provides additional model detail whilst retaining the key elements of the interactions required for polymerisation to occur. In this thesis, we plan to examine how other molecules can modulate the aggregation process and thus continue our literature search to examine models with more than one particle morphology.

1.8.5 Mixed Particle Simulations

Here, we examine the existing literature where alternative spherical particles have been added to simulations of rods. While no existing literature specifically aims to mimic protein polymerisation reactions, it is informative to explore related systems. Simulations on hard rods (strictly no overlaps between particles) and spheres have predominantly been carried out at high densities in a more soft matter-like context. At higher densities/concentrations, the observed phases will be quite different

to those observed at concentrations closer to physiological conditions. Early simulation research by Daan Frenkel, Peter Bolhuis and Henk Lekkerkerker[298–305] establish key transitions in colloidal simulations of rods and mixtures of rods and spheres - particularly at higher densities. Several research groups observe these high-density phases[306–308] such as lamellar phases, nematic and isotropic phases, as shown in figure 1.21. By including spheres, Urakami et al. [308] describe their aggregates with combinations of miscible, immiscible, nematic and isotropic. Miscible is used to describe a mixed rod sphere phase, and immiscible is used to describe phase-separated rods and spheres. (Shown in figure 1.22.) This language is commonly used in describing aggregate morphologies of rods and spheres.

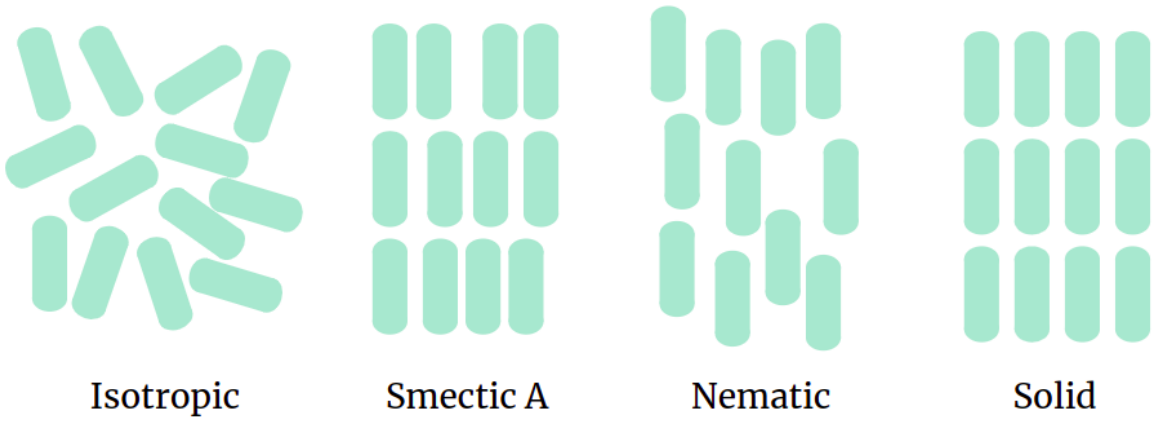

Figure 1.21: Schematic of the isotropic, smectic-A, nematic and solid phase behaviours of rods.

Dogic et al.[298] highlight that for hardcore rods, increasing the rod length increases the stability of layered phases, which form above a critical volume fraction. Similarly, for mixtures of hard rods and spheres, adding spheres smaller than the rod length decreases the volume fraction of rods required to form these layered phases. Wu[309] et al. study phases of high concentrations of hard-core repulsed spherocylinders and spheres in a Monte Carlo simulation system, finding isotropic and nematic phases and estimating the location in parameter space for the isotropic-nematic phase transition. For a system of pure spherocylinders of aspect ratio (length divided by width) $AR = 5$ with two rigid walls, the system produces pre-established isotropic, nematic, smectic-A, and solid phases as the density of the system is increased[310, 311]. Their work continues[312] to find that the system is fully isotropic at low-pressure (low-density) states. Only spheres of equal diameter

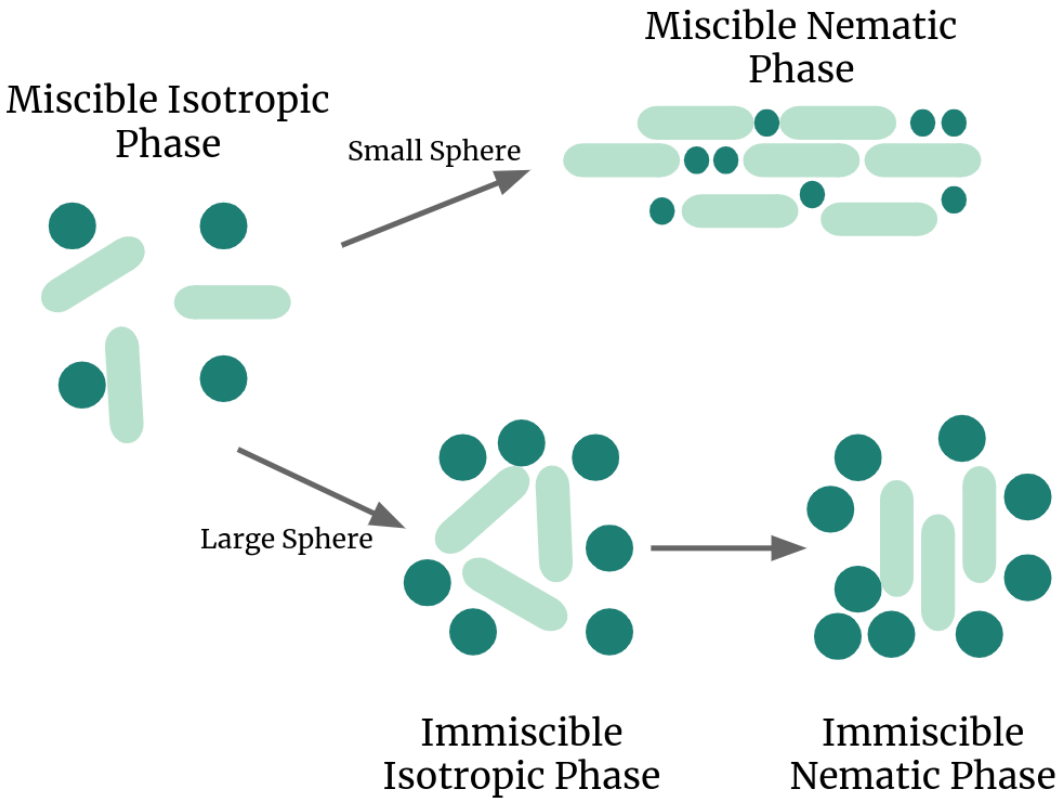

Figure 1.22: Schematic of the phase behaviours obtained in simulations in mixtures of rods and spheres. Reproduced from Urakami et al.[308]

to the spherocylinders are considered in this study[309], and they find that their simulation results agree well with the theory surrounding the mixing and de-mixing of hard rod-sphere mixtures[313]. Further to this, Wu et al.[312] find that hard spheres destabilise the surface nematisation generally, and increasing the concentration of the hard sphere enhances de-mixing effects.

1.9 Determining a Research Question

Throughout this review, we have discussed the impact of neurodegenerative disease with a specific focus on Alzheimer's disease. We have discussed aspects of protein folding and highlighted the key properties of amyloid fibrils and the kinetic processes associated with their assembly. Following a discussion of the currently available therapeutics for the treatment of AD, we have highlighted a necessity for mechanistic information for inhibitor binding to structures of proteins associated

with neurodegenerative disease. Not only this but there is also a requirement for the identification of toxic species and the ability to treat disease with these toxic species as a drug target. It is expensive and time-consuming to routinely screen drug compounds for their efficacy in treating neurodegeneration. Therefore, if we can build a model of amyloid fibre assembly and inhibitors of a variety of different properties, then we can inform the drug design process.

We have summarised several simulation techniques used to describe the properties of fibrils and to model fibril assembly. However, we have identified that a more coarse-grained approach is required to simulate a suitable number of particles where fibril formation is observable to reach the appropriate time scales on which fibril formation occurs. Fibrous structures are highly energetically stable and represent a global minimum in the free energy landscape of protein folding. Therefore, we can use an energy minimisation simulation technique - Monte Carlo - to simulate and disrupt the aggregation process.

By producing a model capable of producing fibril-like aggregates, we can gain insight into the aggregation process and identify the minimal conditions required for aggregation to occur within the scope of our model. The model must be simple enough that it is not computationally demanding or expensive to simulate but complex enough to retain enough detail to represent the amyloid system accurately. Therefore we choose to use a coarse-grained representation where the specific molecular properties of the monomer are lost. However, we can minimally represent the building blocks in fibril formation with non-specific interactions and simple shape parameters. We see that authors such as Saric[3] and Frenkel[4] demonstrate the need for aspherical particles with attractive patches as a minimal model for fibril formation. We utilise this same approach of using a spherocylinder.

We have also identified a need for an improved mechanistic understanding of aggregation prevention mechanisms and inhibitors-fibril binding mechanisms. Therefore we will also include a minimal representation of an inhibitor. The inhibitor will have simple shape and interaction constants, providing a multidimensional parameter space to explore. We hypothesise that by varying these parameters, we can identify critical regions of the phase space, which result in different inhibition mechanisms. Insight into the properties of inhibitors which result in specific inhibition mechanisms

can be compared with the properties of pre-existing drugs and known fibril inhibitors.

Chapter 2

Developing a Model System

In this section, we will describe the development of a model system for amyloid fibre formation and the inhibition of that process. We will discuss modelling strategies at different levels of theory before describing the proposed model used to determine the results presented in this thesis.

Biological systems are complex with many degrees of freedom. Amyloid protein aggregation is the process of thousands of monomers aggregating into huge structures wherein the monomers can have greatly varying degrees of freedom depending on the protein. See chapter 1 for more detail. Amyloid-β in Alzheimer's disease - the specific focus of this project - is tiny compared to other amyloid-forming proteins involved in AD pathology. One example is tau - which has 352-441 amino acids dependent upon the isoform. Aβ is only a 39-45 residue protein, yet it still can find many different conformations given that it is an Intrinsically Disordered Protein (IDP). Preferred backbone angle orientations largely determine the spatial arrangement of a protein with respect to neighbouring backbone molecules. Immediately a significant proportion of arrangements are excluded to steric exclusion (volume of a given amino acid reducing the available space its neighbour could move into), but this still leaves a region of interest where the bonds with its neighbours are strongest. A useful way to illustrate this is with Ramachandran plots[314], which show all of the energetically viable regions in the phase space of the two angles, ϕ and Ψ between atoms when the peptide bond ω is planar and therefore fixed at 180°.

Many continuum models remove the geometric information of the individual particles, which we

believe is key to simulating the aggregation process and would therefore remove the ability to obtain particle-specific information within a more extensive cluster or fibril. Conversely, using quantum calculations even for small molecules is incredibly complex, computationally demanding and time-consuming and therefore is inappropriate for this biological system. This leaves us with atomistic simulations or coarse-grained simulations. Simulating systems which are this large atomistically with current technologies would be incredibly complex and computationally demanding due to the high computational expense of simulating even hundreds of monomers of protein for the time scales we are interested in. State-of-the-art simulations have been carried out with up to 1×10^9 atoms. However, these are multiscale simulations carried out on large numbers of GPU resources. The number of particles required to simulate enough monomers of even small peptides such as amyloid-beta and the surrounding solution is on the order of millions. By using a coarse-grained model, we reduce the number of particles and, therefore, the degrees of freedom within the simulation in favour of reduced computational expense for a single timestep, thereby increasing the timescales we can access with simulations.

In recent years, the importance of in-silico and mathematical modelling in developing a mechanistic understanding of biological processes underlying diseases has been highlighted. By producing a coarse-grained model[3, 315, 316] of amyloid-forming proteins with the inclusion of interaction inhibiting proteins, we can observe a variety of macroscopic behaviours which affect fibril formation and make connections to experimental kinetics. Whilst our primary interest is the context of Alzheimer's disease; the model will apply to a wide variety of amyloid-forming proteins due to the simplicity of the model.

Coarse graining is a technique used to minimise computational expense by reducing a system down to its minimally constituent parts or by representing proteins/sections of protein as a particle. The length scale of the coarse-graining is essential as it allows us to choose the minimum scale our fundamental particles exist on. Our simulation method is primarily Dynamic Monte Carlo simulations in the canonical ensemble. This technique is founded upon the standard MC metropolis algorithm[244, 317], but translational and rotational moves are calculated based upon existing diffusion data for a specific protein. Monte Carlo is a random sampling method with applications

in all major scientific disciplines. It uses random numbers to generate small perturbations within a system to reduce the total system energy as far as possible. Our Monte Carlo scheme is an off-lattice Metropolis Monte Carlo method in the NVT ensemble. i.e. a constant number of particles, constant volume and constant temperature. Many simplified systems are carried out on a lattice. However, in application to biological systems, many codes are written to observe the dynamics of a system with accurate diffusion information. These simulations are carried out to use information from experiments and atomistic simulations of Aβ, or more generally, of amyloidogenic proteins for comparison against the simulation data. The model detail is shown in figure 2.1, where each monomer of amyloid-forming protein is modelled as a spherocylinder with an attractive patch along the cylindrical portion. It interacts via two different potentials, discussed in further detail in section 2.1, one mimics hydrogen bonding, and one represents Van der Waals interactions. Monte Carlo simulations are generally carried out with typically $10^6 - 10^8$ MC steps[113]. Each step does not necessarily have a specific timescale, as is the case for most simulations of this nature, as the primary aim is to minimise the system's free energy[316]. However, it may be possible to estimate it upon completion using an effective reduced units scheme. Our Monte Carlo simulations are typically run for $\approx 10^6 - 10^7$ MC steps, which we find is a sufficient number of sweeps for our simulation systems to reach a steady state.

We are looking to model the effects of inhibitors on the growth of fibrils and oligomeric species in the pathologies of neurodegenerative diseases. The Staniforth lab has a wealth of experimental data, including assays of Aβ fibril formations in the presence of inhibiting molecules. There is also a wealth of existing literature on amyloid protein inhibitors[156, 165, 171, 318–321]. However, there is limited knowledge on how inhibiting molecules work, so creating approximate mechanisms for how and where such inhibitors bind is valuable information. By estimating the properties of possible inhibiting targets even from such a coarse-grained model, we are offering useful information for researchers closer to clinical application.

The phase space on offer due to the huge number of different variables, the phase space on offer is vast, so the most crucial step in this process is deciding which variables will have key effects on what we will observe. Frenkel[322] has shown significant changes in aggregate morphology across

two spherocylinder properties: aspect ratio and design of the attractive patch. There are several studies, including rods and spheres[298, 301, 305, 308, 312, 323–327] which have confidently outlined the phase behaviours across many properties of rod sphere mixtures. Many of these, however, are carried out with no interactions or at a high density. Higher density simulations better describe soft matter systems, which are also of peripheral interest to us in this project.

By representing each amyloid protein monomer as one particle, we remove the internal degrees of freedom of the protein in favour of observing intermolecular interactions on a longer time scale. Using this representation immediately removes time for monomers to fluctuate and fold into a fibril nucleating or elongating conformation and instead assumes that the particle is already in such a conformation where it can form fibrils as well as time taken to develop a force field to drive these processes. A similar approach is taken with the inhibitory molecules to keep the model as generalised as possible so that it can be applied to a variety of systems where amyloid fibrils form.

2.1 Model Description

We have chosen to model the amyloid protein monomer as a spherocylinder. The spherocylinder is defined as a cylinder with hemispherical caps, as shown in figure 2.1. Aspherical particles such as these have been well characterised in soft matter systems as a method of simulating nematic and smectic crystals[313, 328–332]. The spherocylinders in our system also have an attractive 180° wide patch which is $L_p = 0.6L$ long, where L is the total length of the spherocylinder. From this point onward, we will refer to the spherocylinders interchangeably as rods, spherocylinders and PSCs (Patchy Spherocylinders). We have chosen a spherocylinder as the minimal building block to remove the simulation time for finding specific intramolecular conformations of any given monomer and provided a single conformation for these monomers such that any spherocylinder can interact with any other spherocylinder. Previous simulations using a similar model of spherocylinders[4] have highlighted the ability of particles to form fibril-like stacks with two protofibrils. i.e. Parallel stacks of pairs of monomers. The aspherical nature of the particle provides multiple possible shape parameters to change the characteristics of a given monomer, as well as interaction strengths and other such properties which change the properties of the simulation, such as volume or temperature.

Some of these equate to equivalent real properties in *in-vitro* systems, e.g. temperature. Other changes such as how one would change properties such as pH *in-vitro*, are incorporated into the small numbers of interaction strength parameters or similar.

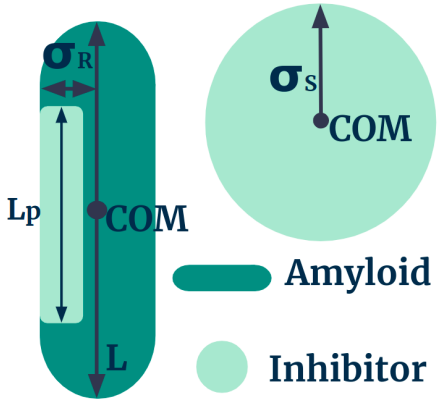

Figure 2.1: A 2D schematic of the particles included in the model. The patchy spherocylinder depicts the position of patch location (light green) and patch length L_p relative to the body axis. The body axis spherocylinder has properties including the radius σ_R, length L, patch length L_p Centre of Mass, **COM** each centred around the origin. Each sphere also has a radius σ_S and centre of mass **COM**, which is centred about the origin in the body axis.

The original implementation of the model considers the spherocylinder as a whole monomer of amyloid-forming protein. However, in the results of the simulations, we discuss a second possible implementation of the model where we describe the monomer as a beta strand within a given amyloid protein monomer. Rescaling of the radius of the spherocylinder - which is used as the minimal length unit in the simulation scheme - will be fixed according to the average hydrogen bond length between adjacent strands in a fibril. In addition to this, the lengths of these segments can be discretised to the numbers of amino acids in a given beta-strand, which would equate to the length or aspect ratio of the individual spherocylinders. This second implementation considers a single beta strand per monomer only, which requires us to negate the remainder of the sequence of the monomer or to consider short-chain amyloid proteins in the range of 3-25 amino acids. The two proposed model suggestions can be found in figure 2.2 (B-C).

For simplicity in the model, we have modelled inhibitory molecules as spheres which are isotropically interacting with one another and are attracted to the patchy region of the spherocylinder. We expect

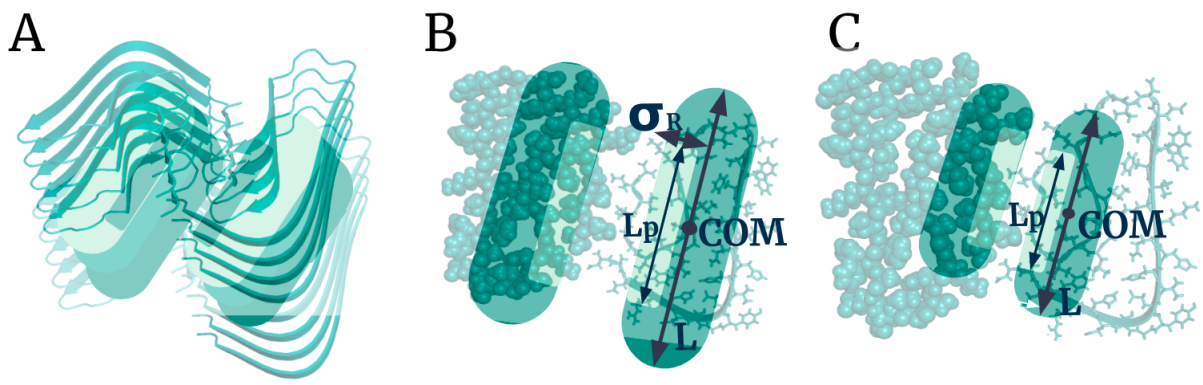

Figure 2.2: A schematic of the proposed model of the spherocylinder as an entire monomer of amyloid protein. (A) A fibril conformation of Aβ reproduced from Gremer et al. [65] with inserted spherocylinders. (B) A single layer of Aβ fibril conformers from Gremer et al. [65] with overlaid spherocylinders used in our model. (C) A single layer of Aβ fibril conformers from Gremer et al. [65] with an alternative model implementation of one spherocylinder equating to a single beta strand within a monomer.

to observe a wide variety of inhibition mechanisms across all areas of the parameter space available. Current research foci in anti-amyloid drug design are in small molecules which are of comparable size to amyloid monomers, larger protein-like inhibitors such as chaperone proteins found in the body which are known to interact with amyloid proteins such as Aβ[153], and also antibodies which are much larger than amyloid peptides that target a variety of structures including the amyloid monomer. These known examples of inhibitors are discussed in section 1.7. Many pharmaceutical companies in the field of drug design screen many compounds without fundamental mechanistic knowledge of how they function, what molecules they might target and what complexes they form with said drug targets. Whilst these screening processes do have successes; if we can inform the drug design process with mechanistic information/properties of the geometries, valence and size of complexes at different regions of many-dimensional parameter space, then we can further the understanding of how these drugs interact with proteins in the body in different disease pathologies. The model is highly coarse-grained, which means that while we lose the specificity of properties such as charged groups, aromatic side chains and functional sequences of specific amino acids, our model can gain insight into a huge variety of amyloid and other biopolymer aggregation processes and how even simple shape parameters can influence the mechanisms of different inhibition behaviours.

2.1.1 Potentials and Attractive patches

We can use mathematical potentials to describe interactions between particles. We would use such potentials in traditional atomistic molecular dynamics simulations to describe specific effects. For example, a Lennard Jones potential could describe Van der Waals interactions or Coulomb's law to describe interactions between charged particles. The interaction constants for these potentials are then parameterised to create something we call a forcefield which may be specific to describing interactions for a specific group of proteins or even as specific as parameterising for a singular protein. These interactions are often parameterised by calculating binding affinities from experimental data or by using machine learning potentials on vast quantities of experimental and simulation data in more recent scientific research. In the case of a coarse-grained model, we use non-specific interactions to capture the character of the overall behaviour or to model only the key interactions in a system. Hence, the parameterisation process is equally less specific. Our model uses two potentials. One to model volume exclusion, and the other to model the attractive interactions between monomers of amyloid protein with one all-encompassing potential. Both are described below in the coming sections, but an overview is given in figure 2.3.

2.1.1.1 Volume Exclusion Interactions

We use a Weeks Chandler Andersen potential to model the volume-exclusion interactions between any pair of particles. The potential takes the form shown in equation 2.1

$$V_{WCA} = \begin{cases} 4E_{WCA}\left[\left(\frac{\sigma_i+\sigma_j}{d}\right)^{12} - \left(\frac{\sigma_i+\sigma_j}{d}\right)^{6} + \frac{1}{4}\right], & \text{if } d <= \sqrt[6]{2}(\sigma_i+\sigma_j) \\ 0, & \text{otherwise} \end{cases} \tag{2.1}$$

Equation 2.1 is also highlighted in figure 2.3 where σ_i and σ_j are the radii of a given pair of particles i and j. It is a Lennard Jones potential that has been shifted. This potential is used in many other coarse-grained simulation schemes[3] at a short range to effectively simulate hard core repulsion with a small amount of flexibility at the particle boundaries. The potential has the cut-off distance $d <= \sqrt[6]{2}(\sigma_i+\sigma_j)$, such that the potential is always repulsive within the cut-off distance.

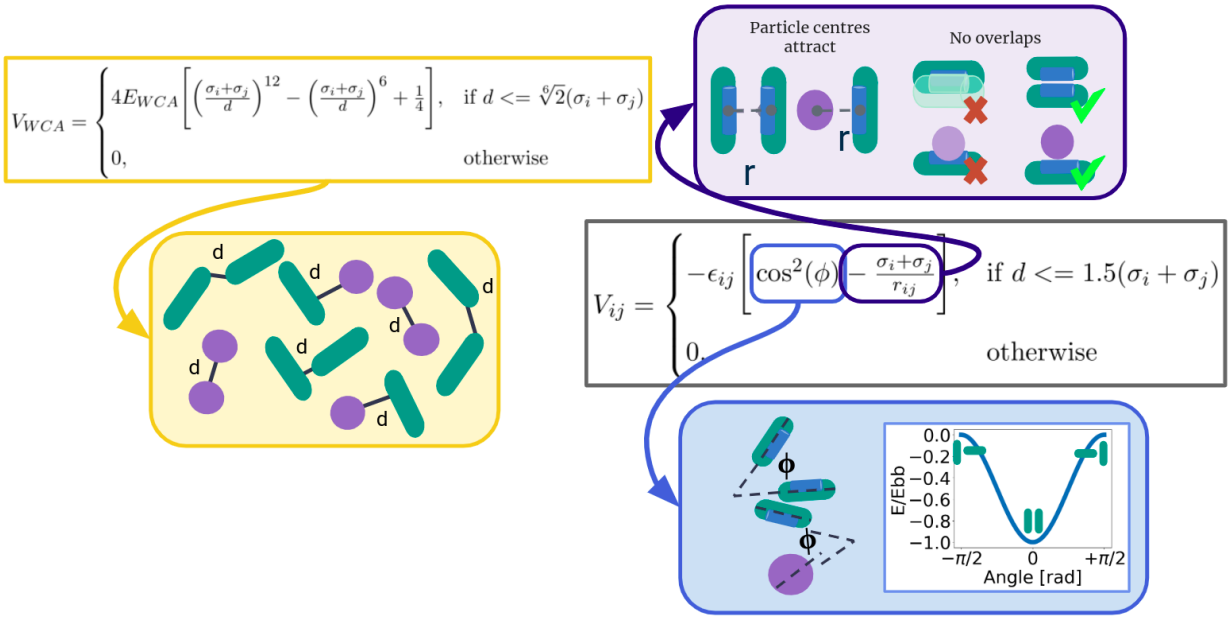

Figure 2.3: Complete depiction of all rod-rod interactions present in the system. The left-hand side [yellow] depicts the Week's Chandler Andersen potential evaluated at the shortest distance d between a pair of rods. This is non-specific to the patch. The right-hand side depicts the beta-sheet-like interaction which is specific to the patch. The first term (blue) depends on the angle ϕ between the axes of a given pair of particles, and the second (purple) favours close packing of rods where σ is the radius of a rod, and r is the distance between the centres of the rods.

The cutoff distance for this interaction is $r_c = \sqrt[6]{2}d$ where d is the shortest distance between a given pair of particles e.g. This is simply the centre of mass distance for two inhibitors but is more complicated for pairs of particles involving a rod. There are separate algorithms used to establish the shortest distance between a pair of rods or a rod and a sphere discussed in section 2.3.2. The repulsive interaction shows that the particles are not perfectly hard-core repulsed. However, the potential energy grows extremely fast. Therefore only significant overlaps would occur if the attractive interactions were many orders of magnitude larger.

2.1.1.2 Attractive Interactions

We use an equation from Saric et al.[292] also highlighted in figure 2.3 defined as:

$$V_{ij} = \begin{cases} -\epsilon_{ij}\left[\cos^2(\phi) - \frac{\sigma_i+\sigma_j}{r_{ij}}\right], & \text{if } d <= 1.5(\sigma_i+\sigma_j) \\ 0, & \text{otherwise} \end{cases} \tag{2.2}$$

In order to model the attractive interactions between amyloid protein monomers, the interactions in amyloid fibrils are predominantly driven by hydrogen bonding between adjacent beta strands. This feature reflected in the design of the potential, which takes the form shown in equation 2.2. ϵ_{ij} is the interaction strength between any pair of particles, ϕ the angle between the two unit vectors for each individual spherocylinder. For the case where the selected particle is a sphere, the angle is fixed such that as long as a spherocylinder patch is facing the sphere, the $\cos^2(\phi)$ term is maximised. σ_i is the radius of particle i, σ_j is the radius of particle j, and r_{ij} (equation 2.3) is the centre of mass separation between the two particles with type i and j i.e. V_{RR} is the notation for an interaction for two rods, and V_{RS} is the interaction for a given rod and sphere. We use the general term V_{ij} as the potential in equation 2.1.1.2 encompasses the behaviour of rod-rod interactions and rod-sphere interactions.

$$\vec{r_{ij}} = C\vec{O}M_i - C\vec{O}M_j \tag{2.3}$$

This interaction only occurs when the attractive patches on each spherocylinder face each other and fall within the cut-off distance, which mimics the directional dependence of hydrogen bonding. There is also a directional term in equation 2.2, which is most energetically favourable when a pair of rods are parallel to one another with their respective patches facing the other. Hydrogen bonds are also very short-ranged interactions, which is reflected in the cut-off radius for the interaction where the shortest distance d between a pair of particles must be less than or equal to $r_c <= 1.5d$. The interaction is not explicitly designed to model hydrogen bonds but is a non-specific potential to encompass the general binding behaviour of amyloid protein monomers in β sheet conformations. However, it is useful to consider the properties of hydrogen bonds in beta sheets and beta strands in this process due to the directionality of the interaction. There is a second term in the equation which has a r^{-1} dependence where r_{ij} is the centre-centre distance between two patchy spherocylinders (PSCs) or a PSC and a sphere. This term is designed to favour close packing of particles. However, the core-repulsion potential will tend to infinity at distances close to 0. This distance is intended as the distance between adjacent beta strands in a given amyloid fibril which

is determined by the average hydrogen bond length between these strands.[333]

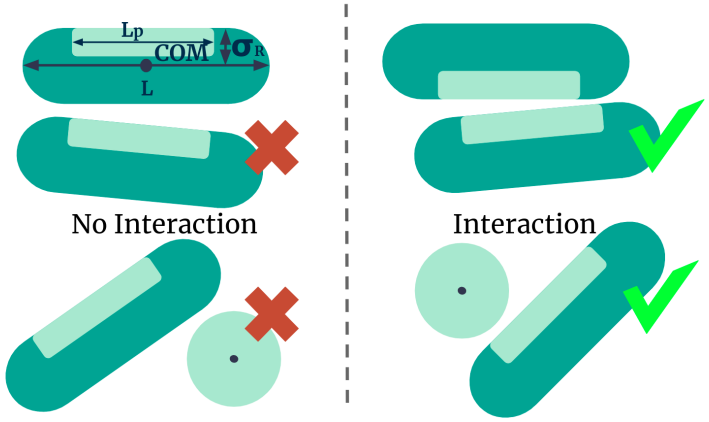

Figure 2.4: A 2D schematic of the particles included in the model. The patchy spherocylinder is labelled with its length L, the centre of mass COM and patch length L_p. The left-hand side depicts example cases where the patch is not facing the other particle and, therefore, would not interact. Conversely, on the right-hand side, there are example orientations where the rod is able to interact with another rod or sphere.

As shown in figure 2.4, in the spherocylinder-sphere case, the attractive patch on the spherocylinder must be facing the sphere and fall within the interaction cut-off distance in order for the interaction to occur. The sphere interacts isotropically and therefore is not direction dependent. The actual calculation of this function in both cases is less trivial than it first appears. In order to check if the attractive patch on the spherocylinder is facing the other particle - one of the criteria for the interaction to take place - is computationally demanding due to the millions of times it will be called throughout the simulation. A particularly demanding aspect of this calculation is the shortest distance calculations required to establish the shortest distance between a pair of rods and a rod-sphere pair. We discuss this calculation in more detail under section 2.3.2. The mathematical routines for the attractive potential to establish if two particles' patches are facing one another is shown below.

2.1.1.3 Determining Mathematically if Two Particles or Patches are Facing Another

A given particle will have a unit vector $\vec{u}_i$ which acts along the length of the rod and a patch normal vector, $\vec{p}_i$, which points outward from the centre of the patch. By definition $\vec{p}_i$ and $\vec{u}_i$ are

perpendicular. For this calculation, the centre of mass distance between particle i and particle j is defined as $\vec{r}_{ij}$ (shown in equation 2.3).

Firstly there is the calculation of x_1 defined by equation 2.4, x_1 is a dot product between the unit vectors of particles i and j to determine their alignment. Secondly, there is x_2 and x_3 shown in equations 2.5 and 2.6. Each of these is a dot product between the distance vector r_{ij} and the unit vector of particles i and j separately normalised with r_{ij} as shown in equations 2.5 and 2.6.

$$x_1 = \vec{u_i} \cdot \vec{u_j} \tag{2.4}$$

$$x_2 = (\vec{r_{ij}} \cdot \vec{u_i})/|\vec{r_{ij}}| \tag{2.5}$$

$$x_3 = (\vec{r_{ij}} \cdot \vec{u_j})/|\vec{r_{ij}}| \tag{2.6}$$

We can consider $\vec{\omega}_i$ and $\vec{\omega}_j$ by definition to be perpendicular to the patch unit vector. They are the ideal directions for the patch of one particle to point towards to interact with the other spherocylinder.

$$\vec{\omega}_i = \vec{r_{ij}} - |\vec{r_{ij}}| x_2 \cdot \vec{u_i} \tag{2.7}$$

$$\vec{\omega}_j = \vec{r_{ij}} - |\vec{r_{ij}}| x_3 \cdot \vec{u_j} \tag{2.8}$$

The patch vectors, $\vec{p}_i$ and $\vec{p}_j$ live in the same plane as $\vec{\omega}_i$ and $\vec{\omega}_j$, respectively. So now we must check if the patches point enough in the right direction of the other spherocylinder. For this we compute dot products of $\vec{\omega}_i$ and $\vec{\omega}_j$ with the patch direction vectors p_i and p_j, which results in ξ_i and ξ_j:

$$\xi_i = \vec{\omega}_i \cdot \vec{p_i} \tag{2.9}$$

$$\xi_j = \vec{\omega}_j \cdot \vec{p_j} \tag{2.10}$$

These calculations of ξ_i and ξ_j check how well the true orientation of the patch p_i matches the supposed "perfect" orientation, ω_i. It is possible to control the patch opening angle (the width of the patch in degrees). However, for the purposes of our project, the patch angle is $\theta = 180° = \pi$. Only

under that specific patch morphology do we observe the specific parallel stacked fibril morphology described by Vacha et al[4]. Then for interactions to be possible, we need that $\xi_i > 0$ and $\xi_j > 0$, such that it points in the direction of the other particle. For the interaction to occur, the patch has to point at least partly in the ideal direction toward the other spherocylinder. If the ideal orientation $\vec{\omega}_i$ and patch vector are counter-aligned, we obtain $\xi_i = -1$, if they are perfectly aligned $\xi_i = 1$, therefore any value above 0 shows that the particle is at least partially facing the direction of the other particle. In the case that the patches are facing one another, we can make the following relation:

$$\begin{aligned} x_1^2 &= \cos(\phi)^2 \\ &= (\vec{u_i} \cdot \vec{u_j})^2 \end{aligned} \tag{2.11}$$

This is then substituted into the original equation 2.2 as the angular contribution to the potential.

2.2 Differences from other models

The primary difference between Andela Saric's model and the model created in our model is that the simulations here are carried out in the NVT ensemble. Papers from the Saric lab also use the canonical ensemble[3], but move on to use the grand canonical[292]. Our model only has one state, whereas Saric's model has three separate states to describe the soluble state, a beta-sheet state and an intermediate state. We choose to use only one state to not explicitly force any behaviour and keep the model as generic as possible to apply to many amyloid proteins. The shortfall of this is that the nucleation behaviour is not described with as much accuracy as other models can capture. There is no barrier to nucleation provided that the attractive patches face each other and the particles are within the interaction cut-off distance. However, there are advantages to using a single-state model. The model becomes even more generalised, and any unique behaviours that do occur can be considered a by-product of the delicate balance of potentials which reflects the biology. There is a high degree of specificity in the narrow range of parameters where certain behaviours are observed with the model - which is not a result of explicit nucleation mechanisms.

In other models, multiple states are designed to simulate nucleation processes more accurately. Therefore, our model falls short of capturing detail correctly in the early time. However, the steady-state results if our model can be considered to be representative of fibril formation in *in-vitro* systems.

Daan Frenkel's model uses only one state, the 2011 paper[4] describes the available phase space for two different types of attractive patches. (Section 1.8.4, figure 1.19) One of which is identical to that used in our model where the patch extends across the cylindrical portion of the spherocylinder only. Their work describes a variety of patch sizes based on an opening angle, however, their work is not amyloid protein specific. Our model focuses on the case where the patch extends across 180° of the cylindrical portion of the spherocylinder. This parameter also determines the morphology of the aggregate. The other patch design extends the patch over the hemispherical portions of the spherocylinder, which arises in a much greater variety of different structures that are not biologically relevant for this context. The patch morphology chosen better represents *in-vitro* and *in-vivo* fibril structures. Our model is capable of forming such structures described as parallel stacks in Vacha et al. [4]. Andela Saric's[292] and Noah Bieler's[251] models also use a simplified patch description of this patch geometry which retrieves the traditional fibril morphology as the focus of their work is also amyloid fibril formation. This simplified interaction is also implemented in our model in favour of the more complex mathematical description given in Vacha et al.[4].

The most obvious difference here is that our model also encapsulates inhibitor behaviours whereas the others do not. Wu et al.[309, 312] do however simulate systems with spherocylinders and spheres and establish the phase behaviour of the system.

2.3 Detailed outline of code structure

Monte Carlo (MC) simulations are a process of random changes to a system which are accepted or rejected based on a set of criteria. The simulations carried out, in this case, are Dynamic Monte Carlo - an algorithm grounded in randomised behaviours in the same way as Metropolis Monte

Carlo, which includes movements and rotations that map to those observed in equivalent physical systems.

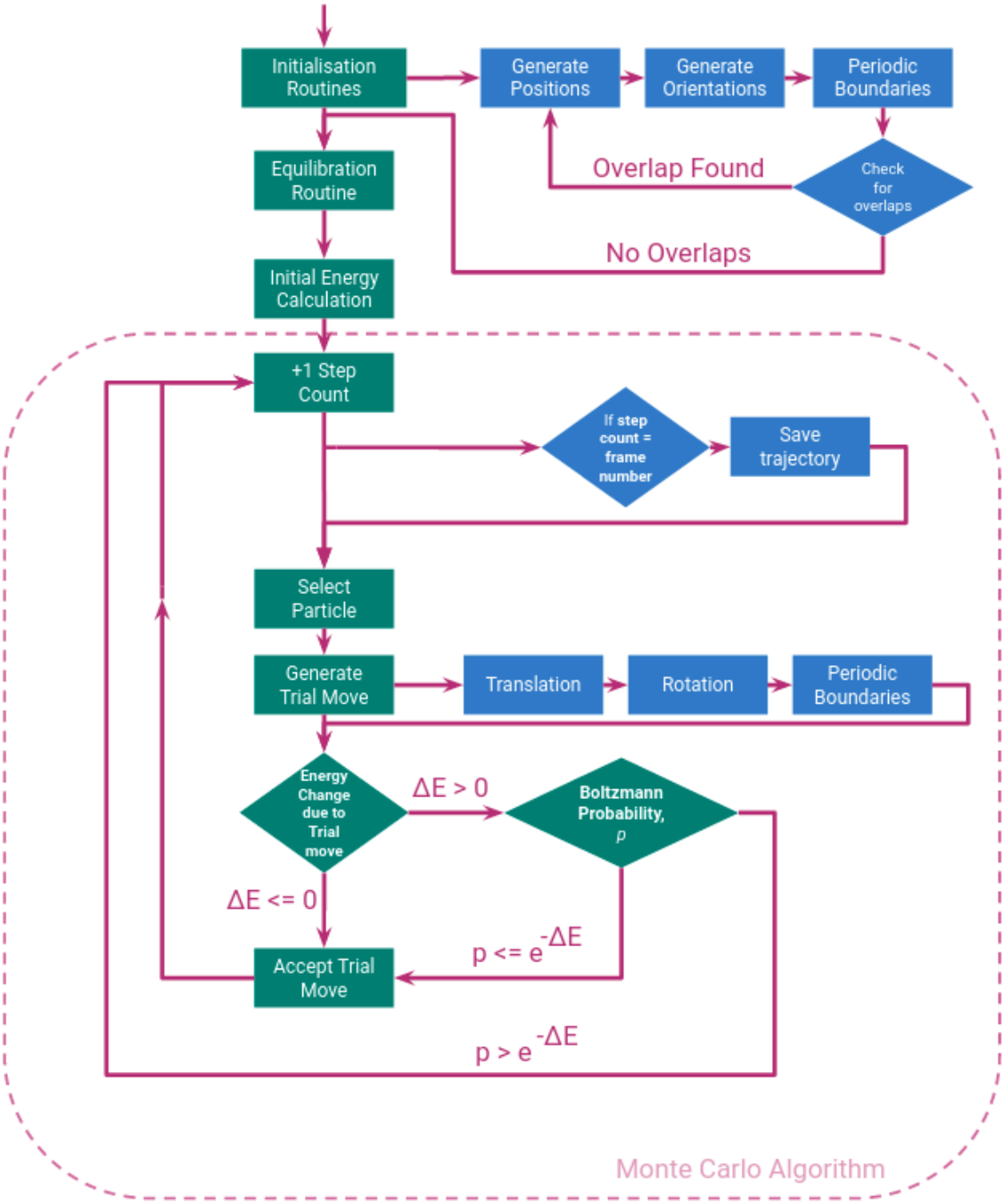

Figure 2.5: A flow diagram for the written code produced for this thesis. Rectangular boxes represent functions or groups of functions in different areas of the simulation. These are joined by arrows indicating the order of calculation in the code and indicate looping regions in some cases.

The main algorithm works such that a particle is selected at random to undergo a translation and rotation as a trial move. This trial move can result in a change in energy dependent upon what that particle is interacting with, if the overall energy decreases, then the move is immediately accepted. In the case of an increase in energy, a Boltzmann probability is calculated of the form $e^{-\triangle E}$ where $\triangle E$ is the change in energy - in units of k_BT before and after the trial move. A random number p is generated in the range $[0, 1)$, if $p <= e^{-\triangle E}$, the move is accepted and otherwise rejected. This is repeated for N times for the number of particles which equates to 1 MC move. All simulations run for 1-10 million MC moves, depending on how long they take to reach a steady-state solution.

2.3.1 Generating Particles and Initialisation Routines

Each object is generated based on the necessary numbers of each particle type and state. Each particle has its own id `ID`, particle type `PARTICLE TYPE`, centre of mass `COM`, unit vector `UNIT`, quaternion `Q`, patch location `PATCH`, and state `STATE`. Generating patch locations is done in two steps. The first passes the state through a function `PATCH body initialisation` which assigns the patch location by rotating the body axis patch location using the **Q Rotation Matrix** function and adding the centre of mass to each element. Further descriptions of rotation mechanics are described in section 2.3.1.5

This generation function takes the newly created object `particle` and generates all necessary information concerning initial conditions given to the particle. i.e. state and particle type. The majority of the code uses an object-oriented scheme where each rod or sphere is represented by an object as part of a class. The `amyloid` class has objects which each have the following properties:

- **Center of Mass** `COM`: defined as the radial centre of the cylindrical portion of a given spherocylinder positioned halfway along the length of the particle. In the body axis (particle frame of reference) $COM = [0, 0, 0]$
- **Unit vector** `UNIT`: defined as a vector along the length of the particle. In the body axis, this is pointing vertically in the z axis. $UNIT = [0, 0, 1]$
- **Quaternion** `Q`: defined in accordance with Ovito and LAMMPS, where the quaternion is a

representation of the rotation of a spherocylinder from the z axis. The default quaternion, $Q = [1, 0, 0, 0]$ which results in no rotation.

- **Patch Unit vector**, `PATCH UNIT`: defined as perpendicular to the spherocylinders' unit vector, pointing radially from the centre of mass to signify the centre of the attractive beta-sheet patch. In the body axis this is a unit vector in the x direction, $PATCHUNIT = [1, 0, 0]$.

- **Patch Location** `PATCH` The patch location is generated by rotating the body axis patch location - which varies depending upon what state the particle is in - using the **Rotation Q** function and the quaternion generated in the previous step. The centre of mass location is then added to the patch location using vector addition to align the patch to the particle it is a part of.

There are three different ways of generating particles and their respective properties in the simulation scheme.

- **Generalised**

  This will generate an ID, particle type, the state that the particle is in, a quaternion, a unit vector, a centre of mass, a coordinate for the patch, and a normalised patch unit vector which by definition is perpendicular to the spherocylinder unit vector, and finally two coordinates for corresponding to the end of each rod.

- **Fibril Conformation**

  Generates particles in a fibril conformation and will supply - with a given state a given particle type, seed ID and particle ID - the centre of mass, quaternion on the unit vector along with the patch location and the normalised patch unit vector.

- **Feed in Old Simulation Data**

  This is used to restart simulations from their previous endpoints. This function reads in trajectory information from a previous simulation and recreates those objects.

There is also a function that inserts spheres into the simulation. Much of this information is the same for properties of the spherocylinder because we need the directional and orientational properties to establish a directional dependence on the interaction for the sphere where the interaction

only takes place if the attractive patch of the spherocylinder is facing the sphere. This must occur so that the attractive patch and the sphere do not interact through the reverse side of the spherocylinder. We can also feed in sphere simulation data or add more particles to a preexisting simulation trajectory. In addition to this for potential further extensions of this code, the choice was made to keep directional information for spheres as well as spherocylinders in order to make further code developments more accessible.

2.3.1.1 Box Size

The size of the box is determined by a volume fraction chosen by the user. This volume fraction is the total particle volume divided by the box volume. Therefore the box size can be calculated from a known volume fraction. The rod volume is given by equation 2.12 and sphere volume is given by equation 2.14.

$$V_{ROD} = 2 * V_{HEMISPHERE} + V_{CYLINDER} \tag{2.12}$$

$$= \frac{4}{3}\pi\sigma_{ROD}^3 + \pi L\sigma_{ROD}^2 \tag{2.13}$$

$$V_{SPHERE} = \frac{4}{3}\pi\sigma_{SPHERE}^3 \tag{2.14}$$

The box volume is then calculated with the following relationship:

$$V_{BOX} = L_{BOX}^3 = \frac{V_{ROD} + V_{SPHERE}}{VF} \tag{2.15}$$

2.3.1.2 Check for Overlaps and Boundary Conditions

The whole system then undergoes an overlaps check, which is an expensive but necessary process. This checks if any particle is overlapping with another, and exits if an overlap is found. In this case, the position, patch location and orientation are recalculated, but the state and particle type is retained. Once this has taken place, the simulation can start. This process contains a `while` loop, which will continue to run until it finds an orientation which returns no overlaps or until the process has been repeated several thousand times - at which point it is improbable that a suitable arrangement of particles exists in this method. Other simulation systems may opt for a lattice-like arrangement of particles which requires a long equilibrating phase to find a truly random arrangement. By using a random orientation and position placement of particles, we can run a shorter equilibrating phase than other particle placement methods. This process is an $O(N^2)$ process, which is not computationally efficient. This means that by doubling the number of particles, the process would take four times as long. Using the random placement method, however, is faster than running as many Monte Carlo time steps as required to fully equilibrate the system for the number of particles we typically use. (≈ 1000)

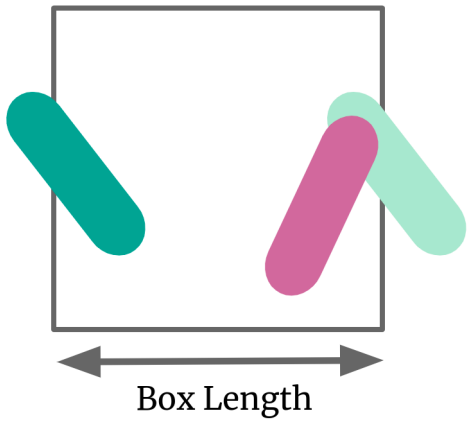

Figure 2.6: Example of a spherocylinder passing across a periodic boundary (green) overlapping with another spherocylinder (purple).

A particle may have its centre of mass near a given boundary. However, the remainder of the particle may extend over that boundary and overlap with a particle (highlighted in figure 2.6.) The boundaries in the simulation are periodic in all directions, so we must consider this in the process of checking for overlaps.

2.3.1.3 Equibrilation Phase

The equilibration phase ensures a completely random dispersion of particles regarding their position and orientation. It is usually more important to include this for non-rigid particles with more internal degrees of freedom, but it is an important step here, too. The process is carried out for a fixed number of steps proportionate to the number of particles in the system. Translations and rotations are generated, but there are no interactions between any particles other than to ensure that they are hard core repulsed, i.e. That they don't interact.

2.3.1.4 Translation and Rotation Function

The translation function takes in the information for a given particle and passes the object through several other functions. Firstly a new centre of mass is generated by generating a translation in x, y and z coordinates. A random number is generated from a uniform distribution in the range $-D^T_{MAX} <= x <= D^T_{MAX}$, which is set according to the literature for amyloid β specifically. D^T_{MAX} is the maxmimum step size for translation of a particle. The newly calculated centre of mass is then passed through a periodic boundary function to check that the new centre of mass is not outside the simulation box.

A new orientation p is then generated using the `Generate Q` function, which is then multiplied with the existing quaternion for the particle with the method discussed below in section 2.3.1.5. The new unit vector for the particle is then calculated by rotating the original unit vector by the new quaternion with the rotation matrix $\mathbf{R}$. The same method is used for the patch unit vector, the patch centre of mass and the endpoints of the spherocylinder because all of these objects need to be rotated from the body axis to the simulation box axis or the current axis of the particle. A similar calculation takes place for sphere objects without properties such as the endpoints of the particle.

The Body Axis The body axis is used primarily in dealing with quaternion rotations and Ovito. Ovito reads quaternions as rotations from a body axis. The body axis for a rod points directly in the z direction with no contribution from x or y as shown in figure 2.7. Therefore each time a particle is rotated, its patch location must be recalculated from its original position relative to

the body axis. Similarly, this must be done with unit vectors to ensure orientations are true to position in comparison to patch location. i.e. The unit vector $\vec{u}$ always points along the z axis in the reference frame (referred to as the body axis). Similarly, in the body axis, the patch unit vector is defined such that it is always perpendicular to $\vec{u}$. The unit vector changes when the rod is rotated in the simulation axis/lab frame. Consequently, the patch unit vector $\vec{p}$ must be rotated according to the relationship to $\vec{u}$.

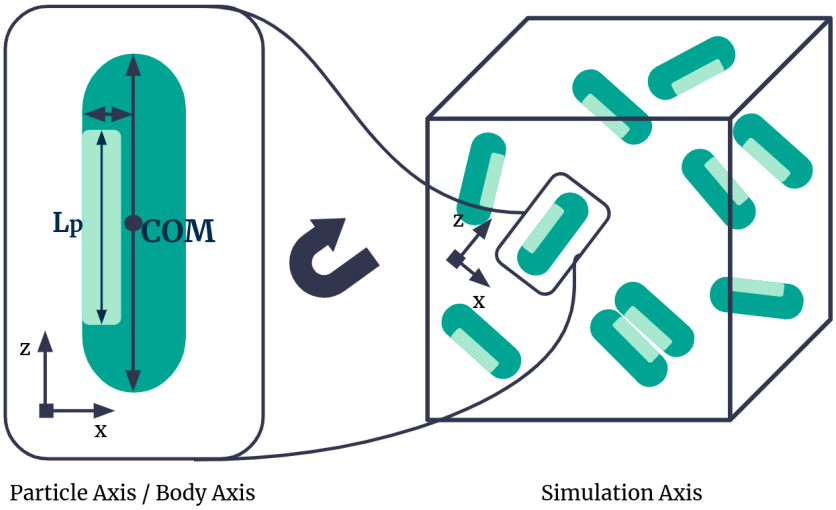

Figure 2.7: A 2D schematic of an example simulation full of spherocylinders labelled "Simulation Axis" with a highlighted particle showing its rotation from the particle axis to the simulation axis.

We use the body axis for all translations and rotations in the system to calculate new orientations based on their rotation from the reference/body axis. This does not significantly increase the computational expense over simply rotating from the previous orientation and decreases the effect of numerical rounding errors from carrying out millions of rotations on the same unit vector. If we recalculate from the body axis each time, the quaternion rotations and unit vectors will be consistent with one another. This also prevents rounding errors in calculation. This effect is further reduced by using a reduced units scheme as described in section 2.4.

2.3.1.5 Rotations with Quaternions

The reader might expect unit vectors and Euler angles to be used to calculate rotations of particles in 3D space. Under this scheme, a unit vector is rotated via a rotation matrix that contains information about the axis of rotation and degree of rotation. Euler angles - a series of 3 angles determining how much a given vector will rotate in each degree of freedom. Quaternions are a more sensible choice of shorthand for rotation and orientation as they avoid the "Gimbal Lock" issues which arise with using Euler angles. This is where there is a reduction of a degree of freedom when axes of rotation become aligned and therefore indistinguishable from one another.

Quaternion notation is useful as the algebra associated with it is well established and a robust method. Quaternions take the form $\mathbf{Q} = (q_w, q_i, q_j, q_k)$. The components do not directly translate to a physical quantity. However, they form a useful rotation matrix $\mathbf{R}$ as defined.

$$\mathbf{R} = \begin{bmatrix} 1 - q_2^2 - q_3^2 & 2(q_1 q_2 + q_0 q_3) & 2(q_1 q_3 - q_0 q_2) \\ 2(q_1 q_2 - q_0 q_3) & 1 - q_1^2 - q_3^2 & 2(q_2 q_3 + q_0 q_1) \\ 2(q_1 q_3 + q_0 q_2) & 2(q_2 q_3 - q_0 q_1) & 1 - q_1^2 - q_2^2 \end{bmatrix}$$

Quaternions can be used to represent an orientation or a rotation. In this case, they are used as a rotation from the z-axis and form a rotation matrix used to rotate the locations of patches. Rotations are generated by choosing an angle θ to rotate by and a unit vector $\mathbf{u}$ to rotate about. The quaternion $\mathbf{q}$ is then calculated with the following relationship:

$$\mathbf{q} = \begin{bmatrix} \cos(\theta/2) \\ u_0 \sin(\theta/2) \\ u_1 \sin(\theta/2) \\ u_2 \sin(\theta/2) \end{bmatrix}$$

In order to match the input for the visualisation software used for visualising aspherical particles, Ovito[334], the quaternions are mapped as a rotation from the z axis - the default orientation for a spherocylinder in our model. As discussed in the previous section, this is also useful in reducing numerical rounding errors. Treating quaternions as rotations can be confusing as quaternions can

be used to represent an orientation **or** a rotation from a specific axis. Therefore, each generated quaternion must be combined with the existing quaternion rotation for a given particle to produce the net rotation of the particle. For two rotation quaternions q and p made up of a scalar component and a vector component:

$$q = (s, \vec{v}) \tag{2.16}$$

$$p = (t, \vec{w}) \tag{2.17}$$

It is possible to multiply two quaternions as follows:

$$qp = (s, \vec{v})(t, \vec{w}) = (st - \vec{v} \cdot \vec{w}) + (s\vec{w} + t\vec{v} + \vec{v} \times \vec{w}) \tag{2.18}$$

The resulting quaternion qp shows the combined rotation from the z axis. By substituting this quaternion and the unit vector into the quaternion rotation matrix, we obtain the final unit vector orientation of a particle as shown in equation 2.19:

$$\vec{u'} = \mathbf{R}\vec{u} \tag{2.19}$$

It is also necessary to do this for coordinates of the rod endpoints which are normalised in the body axis and then multiplied by their length following transformation. This extra step is unnecessary for a unit vector as it is already normalised.

$$\vec{p'} = |p|\mathbf{R}\hat{p} \tag{2.20}$$

2.3.2 Energy Calculation

Once a particle is selected, and the translational and rotational moves take place, we evaluate the interactions between this selected particle and all others in the system. Two algorithms are imported from the C++ library GTE Mathematics: one which evaluates the shortest distance between two line segments that is used to calculate the shortest distance between a pair of rods, and a second which calculates the shortest distance between a line segment and a point which is used to evaluate the distance between a rod and a sphere as shown in figure 2.8. These algorithms are

fast and do not have issues with calculating distances between nearly parallel rods like many similar algorithms. The distance calculation algorithms are one of the most computationally demanding aspects of the code, not because it is a large function but because the algorithm is called $2 \times N_{PARTICLES}^2 \times N_{STEPS}$ times per simulation, and they require unique vector-like data types as input which become expensive to generate.

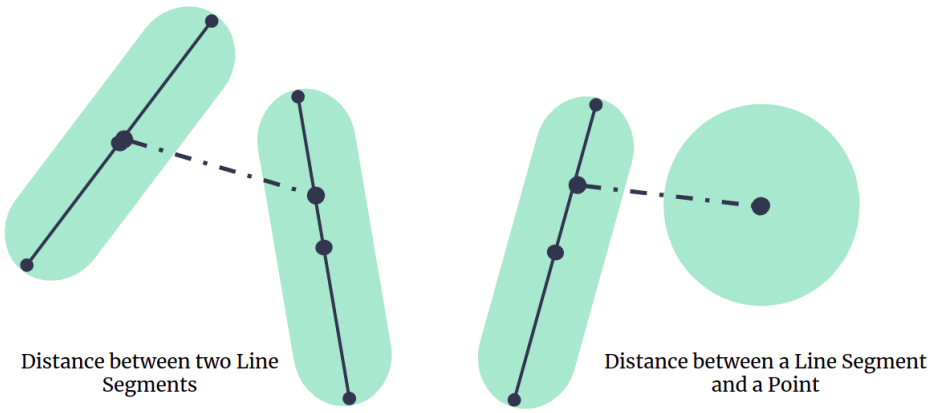

Figure 2.8: A 2D schematic of example shortest distances d between two line segments inside two rods and a line segment and a point inside a rod and a sphere.

In order to save the computational expense of generating different data types for these algorithms, there are simple distance checks between the ends of spherocylinders (or the ends of a spherocylinder and a sphere) to see if the two particles are close enough to likely meet other interaction criteria. e.g. other smaller cut-off distances and directional dependencies. To do this, we treat the rods as spheres with radius L where L is the length of the rod and check if the centre of mass distance between neighbouring particles is less than $2L$. If this is the case, then we use the more fine-grained distance calculation algorithms described above.

A total pairwise energy calculation is completed in the first iteration of the simulation outside of the equilibration phase. Following this, all subsequent energy calculations are completed by evaluating the change in energy due to moving a particle instead of fully recalculating the energies. This requires one loop through the total number of particles. These energies are calculated pairwise where the specific interactions are listed in the previous section 2.1. The WCA potential is evaluated between all pairs of particles, and the attractive interactions only occur between pairs of

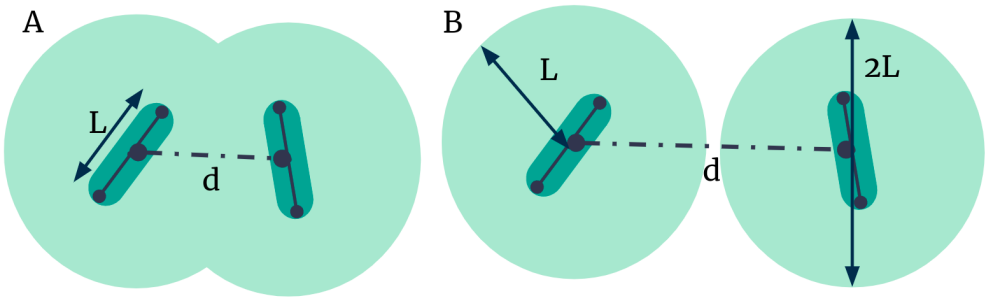

Figure 2.9: A 2D schematic of example distance calculations between pairs of rods. We assume each spherocylinder is a sphere of radius $\sigma = L$ and check that centre to centre distance is (A) less than $2L$ and (B) greater than $2L$.

spherocylinders or a spherocylinder-sphere pair.

2.4 Reduced Temperature Scheme

We must be able to control features such as temperature, T in the canonical (NVT) ensemble. However, to improve the numerical accuracy in the simulations, it is favourable to utilise a reduced unit scheme. Such a scheme essentially reduces (or increases) mass, m, energy ε and distance σ to the minimal building blocks all other quantities are calculated from. This is used so that within the simulation, all calculations are completed using numbers of close orders of magnitude - information can be lost in extensive repeated calculations of extremely small or extremely large numbers. If said minimal quantities are set as unity, then it is not even necessary to include them in the calculations in some cases. For the understanding of the user, however, it is important that they are able to switch between the reduced units scheme and SI units.

The following equations dictate the conversion between energy, E, mass M, distance or length quantity D and temperature T and their reduced quantities E^*, M^*, D^* and T^*:

$$\begin{aligned} E^* &= \frac{E}{\varepsilon} \\ T^* &= \frac{k_B T}{\varepsilon} \end{aligned} \tag{2.21}$$

Where ε is the minimal unit used to convert between the real and reduced quantities which will be used in simulations.

2.4.0.1 Setting Reducing Constants

Distance σ will be used thoroughly in the simulations because a new movement must be generated on every MC move. The most sensible unit here is to reduce the radius of the spherocylinder σ from 1×10^{-9} to 1. The length of the rods, L can vary between $L = 2\sigma$ and $L = 32\sigma$ depending on the aspect ratio of the particle.

$$\begin{aligned} \sigma_{real} &= 1nm \\ \sigma &= 1 \end{aligned} \tag{2.22}$$

Energy and temperature are the most complex issues here. In order to vary the temperature in the system and to observe any effects, we must evaluate the interaction strengths according to some reference temperature T_{REF}. Effective energy can be calculated using $E = k_B T_{REF}$. An example is shown below:

$$V_{AB} = Ak_B T_{REF} \tag{2.23}$$

Where V_{AB} is some example potential that occurs with interaction constant A. We must then evaluate the energies using the reduced temperature, T. The reduced temperature is a dimensionless ratio of the reference temperature, T_{REF} and the absolute (or real) temperature of the simulation, T_{ABS}.

$$T = \frac{T_{ABS}}{T_{REF}}$$

Similarly, the reduced energy ε, which is used to reduce energies, is given by:

$$\varepsilon = k_B T_{ABS}$$

We can then convert the interaction strengths using the following relation:

$$\begin{aligned} E_{AB} &= A\frac{E_{REF}}{\varepsilon} \\ &= A\frac{k_B T_{REF}}{k_B T_{ABS}} \\ &= A\frac{T_{REF}}{T_{ABS}} \end{aligned}$$

To eliminate the need and computational expense of repeatedly dividing by $1/k_B T_{ABS}$ or β in evaluating Boltzmann energies in the MC algorithm, we can do this before the start of the simulation.

$$\begin{aligned} E_{BOLTZMANN} &= \frac{k_B T_{ABS}}{\varepsilon} \\ &= \frac{k_B T_{ABS}}{k_B T_{ABS}} \\ &= 1 \end{aligned} \tag{2.24}$$

For each quantity, we rearrange our equations. This is simple for mass and distance as we have reduced each to 1. The most important part of this scheme is that we can transform back to real units from the reduced units at the end of the simulation or whenever we need to perform calculations related to energy to file.

$$\sigma_{ABS} = \sigma\sigma_X$$

$$E_{ABS} = \varepsilon E$$

The step sizes for translation and rotation used are scaled according to the minimal unit and will therefore scale linearly with the size of the minimal unit.

2.5 Constants and Variable Definitions

- **Box dim min** The smallest coordinate of the simulation box
- **Box dim max** The largest coordinate of the simulation box
- σ_R The radius of rod-like or spherocylinder particles such as amyloid protein monomer

- σ_S The radius of spherical particles such as inhibitor protein monomer
- **L** The length of rod-like particles expressed in multiples of the radius of the rods
- **D** The diameter of rod-like particles
- **Lp** The length of the attractive patch
- **Writing Frequency** Every set number of steps, the script will dump a set of files marking trajectories of each type of particle and trajectories of patches associated
- **Write Count** A count to compare against the **writing frequency** to ensure file writing takes place
- **N steps** The total number of MC sweeps in the simulation
- **N rods** The total number of rod-like particles in the simulation
- **N spheres** The total number of spherical particles in the simulation
- **N particles** The total number of particles available in a simulation (= **N rods** + **N spheres**)
- **Max displacement** The maximum translational displacement which can take place in any direction per step
- **Max rotation** The maximum rotational displacement which can take place per quaternion element per step
- **min dist rr2** = (2 sigma rod)^2 The minimum squared distance that two rods can be apart without overlapping
- **min dist rs2** = (sigma rod + sigma sphere)^2 The minimum squared distance that a rod and a sphere can be apart without overlapping
- **min dist ss2** = (2 sigma sphere)^2 The minimum squared distance that two spheres can be apart without overlapping
- **N beta** The number of beta state particles in a simulation
- E_{RR} The interaction strength between two-rod patches

- E_{RS} The interaction strength between the attractive rod patch and a sphere

Chapter 3

Benchmarking Fibril Growth and Exploration of Parameter Space

This chapter is about exploring the available phase space of the spherocylinders. By looking at the extreme regions of the parameter space, we can establish that the system is physical and provides a suitable benchmark of the system before including inhibitory molecules. We must first parameterise the system and gain insight into the parameters which produce fibril assembly. Several models of patchy spherocylinder systems are already in existence[4, 113], so this chapter highlights the similarities and differences between this model and others as discussed previously in section 2.2.

There are key parameters used to benchmark and understand how the patchy spherocylinder system functions in terms of its physical limits, the aggregation behaviour(s) and scaling behaviours. We firstly address constants set such as step size, relative temperature and other reduced unit quantities. This is followed by several 2D parameter sweeps, including temperature vs interaction strength to identify the region where fibril assembly occurs, temperature vs rod aspect ratio to determine the effect of shape parameters on aggregation, and temperature vs volume fraction to determine how high concentrations can reach to improve simulation time without significant overcrowding events and finally a sweep of rotational and translational diffusion parameters. We follow this process to determine how rods with particular characteristics behave and to draw implications for fibril assembly.

Each of these sweeps has important modelling implications. By changing the shape parameters or interaction strength parameters, we can tune the degree of aggregation behaviour and the morphology of the aggregates. In doing so, we can connect these aggregates to different stages of the amyloid aggregation process and observe how one might move between these different states via conformational changes or growth processes. Equally, by adjusting the translational and rotational stepsizes, we can use the model for different amyloid proteins and other fibrous non-biological applications.

3.1 Analysis Methods

In key tools employed to analyse simulation trajectories are grids of time-dependent data, including the mean cluster size, largest cluster size, populations of specific species such as the monomer count, percentage fibril content and order parameters. The evaluation of which particles are in a given cluster is calculated by the python package for the visualisation software Ovito. This data is then handled by an in-house python code I developed for the analysis of these simulations, which calculates the following for a given time point:

(i) Mean Rod cluster size:

$$\bar{X} = \frac{1}{N}\sum_{i=0}^{N} X_i \tag{3.1}$$

Where X_i is the size of a given rod cluster in terms of the number of monomers and N is the number of rod clusters.

(ii) Largest Rod cluster size

Calculated using the clustering algorithm and observing the cluster with the largest number of rods.

(iii) Median Rod Cluster Size:

$$\tilde{X} = \begin{cases} X[N/2], & \text{if N is even} \\ \frac{1}{2}(X[\frac{N-1}{2}] + X[\frac{N+1}{2}]), & \text{otherwise} \end{cases}$$

Calculated using the clustering algorithm.

(iv) Number of Free Monomers:

Calculated using the clustering algorithm and observed the number of clusters with a cluster size $X = 1$.

(v) Proportion of Monomers in Clusters:

Calculated using the clustering algorithm and observed using the number of clusters with a cluster size $X > 1$.

(vi) Size distribution of Rod Clusters:

Calculated using the clustering algorithm with an input of rod trajectories, then produced a histogram of the size distribution weighted to the same distribution. This is done in order to observe larger clusters better, as the data clearly show where the majority of rods in a given system are. Discussed in more depth in section 3.1.2

(vii) Orientational Order Parameter:

$$S = \frac{1}{2}\Big\langle 3\cos^2(\phi) - 1 \Big\rangle \tag{3.2}$$

Where ϕ is the angle between a given rod and its neighbour. S is an averaged quantity over all clusters and is weighted according to size i.e. a larger cluster will have a larger contribution to the total S. Discussed in more depth in section 3.1.1

3.1.1 Liquid Crystal Parameters

The nematic order parameter is a measure of how ordered a group of symmetric aspherical particles are. Typically these liquid crystal parameters are used in soft matter systems. We apply this measure differently by measuring the level of order within single clusters. Nematic order is an averaged quantity given by the expression in equation 3.2.

The angle ϕ is the angle between a given particle and the director for the cluster. As shown in figure 3.1, the order parameter $S = 0$ when there is a complete lack of order and the mix of directions of rods is isotropic and conversely, $S = 1$ when all rods point in the same direction. For the case of

these simulations, the director is non-trivial to calculate, given that all particles are symmetric but contain a unit vector that must point along the rod, making them mathematically asymmetric.

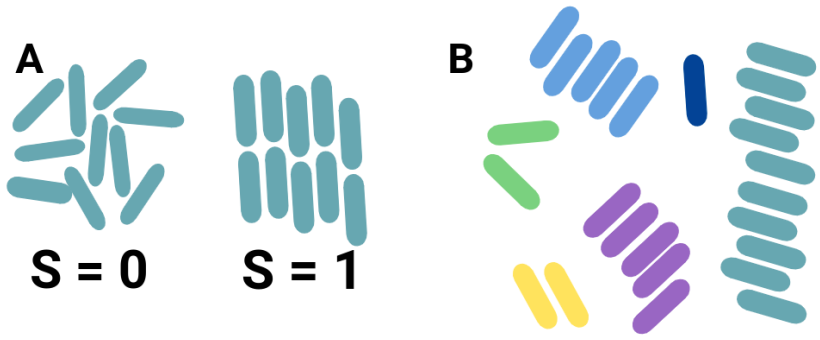

Figure 3.1: (A)The order parameter, S, is 0 when the rods have randomised orientation, and there is no order. There is total order, and all particles are perfectly aligned when $S = 1$. This is orientational order only and not spatial. (B) A graphic showing different clusters in different colours to highlight that order parameter calculation for each cluster is independent of other clusters.

For this reason, here I use an alternative calculation for an orientational order parameter which takes the same mathematical form of the nematic order parameter, but where instead the angle ϕ is the angle between neighbouring spherocylinders i and j within a cluster. The order parameter S is an averaged quantity which varies between 0 - isotropic, and 1 - completely ordered as shown in figure 3.1. For the case of these simulations, the order parameter is calculated on a per cluster basis and is incorporated into a weighted average according to cluster size. As a result of this, the orientational order parameter, S, is a measure of the order of rods relative to their neighbours, irrespective of the cluster size. The fact that the filament-like clusters can bend means that the rod angle is relative to a curve rather than a straight line. In this case, the director changes direction along the bent filament, so we can use this alternative order parameter to better estimate the level of local order within a filament.

3.1.2 Length Distributions throughout simulation

Many of the analyses for this project use length distributions. We are interested in length distributions in order to properly characterise all of the different structures produced in the simulations. The populations of small clusters or polymers are as important as those of extended fibril structures.

The amount of protein that is in a fibrillar state can be observed experimentally via Thioflavin T assays. Thio-T binds to highly ordered amyloid-β structures, so it measures the amount of ordered aggregate present in a sample. These may not just be mature fibrils but can also be protofibrils and fibril-like oligomeric structures.[84, 85]

It is impossible to observe the full-length distribution of a sample experimentally. We can only observe areas of it. For example, ongoing work within the Staniforth lab uses asymmetric flow field flow fractionation techniques (AF4) combined with a multi-angle light scattering (MALS) detector to observe smaller structures such as dimers, other low multimers and some oligomeric structures. It is possible to predict the shape of such populations analytically under specific conditions. Due to the non-linearity of mathematical descriptions of biopolymer formation, analytical solutions are extremely difficult to solve and are often limited to specific sets of conditions and involve many assumptions and simplifications. These analytical solutions are incredibly useful and allow us to predict relative populations of different-sized aggregates such as oligomers. However, there is no knowledge of the shape and arrangement of monomers within said oligomers from such distributions. The simulations carried out in this thesis allow us to obtain a full picture of the size, shape and arrangement of monomers across the entire size distribution throughout the entire time course of the simulation. These simulations, therefore, provide geometric information that we cannot obtain experimentally or through kinetic models.

The length distributions we have used are weighted to the numbers of monomers within them to highlight clearly how much monomer is in a particular structure of size N. A regular histogram would only add a value of 1 to the frequency occurrence for the presence of a cluster of a certain size. In contrast, here, it is clearer to see a peak of size 100 where there is a cluster of size $N = 100$, or similarly a peak of size 50 where there are 5 clusters of size $N = 10$ present. Using the fraction of monomers in clusters rather than the number of clusters is a better measure of how aggregated the system is.

3.2 Choices of constants

3.2.1 Rod Radius

For most parameter sweeps, the radius of the spherocylinder is $\sigma = 1$nm. This value is taken from the literature of a similar spherocylinder model applied to Aβ aggregation[3]. However, this is used as the minimal length unit of the simulation, where all other lengths are calculated as multiples of this minimal length unit. Therefore we can set this radius as any number, and the results of the simulations will not be affected. For further discussion of the impact on the step sizes see section 3.2.3.

3.2.2 Weeks-Chandler-Andersen Potential

The interaction constant used for the WCA potential is set to $1k_BT_0$ so that it acts as the minimal unit for the simulation system. All other interaction strengths are calculated relative to this minimal unit.

3.2.3 Step sizes

We have taken the value for translational diffusion from the literature[3]. At present, diffusion does not depend on direction i.e. The constant is the same in x, y and z. The rotational step size is also taken from the literature [3]. These values are based on atomistic simulation data of a fibrillar conformation of Aβ protein. The translational and rotational step sizes are shown below:

$$\begin{aligned} D^T_{MAX} &= 0.212\sigma_{ROD} \\ D^R_{MAX} &= 0.122\sigma_{ROD} \end{aligned}$$

In this model, we do not have a different translational step size for motion that is parallel and perpendicular to the rod. Instead, we take the singular value from the literature for amyloid beta specifically. We estimate that since our monomers are relatively short, their aspect ratio is small enough that the asymmetry of diffusion is not a dominant factor; therefore, we neglect it for simplicity. Given that the step sizes are set to parameterise a specific protein, this will impact

further results if the radius of the minimal unit (spherocylinder) is changed. However, the step sizes scale linearly with the spherocylinder radius (assuming uniform density across different proteins, the spherocylinder mass). Therefore given the highly coarse-grained nature of the system, these diffusion parameters should fall within the region of error if changed - provided that the new radius is not changed over orders of magnitude.

3.2.4 Temperature Reduced Unit

The chosen reference temperature is $T_0 = 310K$. This has been chosen as this is representative of physiological conditions in the human body, and therefore all temperatures are calculated as a multiple of this number. The Boltzmann constant k_B is set to 1 for these simulations to make the interaction strength and temperature directly comparable. Reduced temperatures and other reduced units are discussed in further detail in section 2.4

3.3 Examining the Parameter Space

3.3.1 Interaction Strength and Temperature

The interaction strength and temperature units have been reduced such that both parameters are in terms of the reference temperature T_0. Therefore it is easiest to consider the two parameters as a balance of interaction energy and thermal energy or thermal noise. The temperature in Monte Carlo simulations is essentially boiled down to the likelihood that an unfavourable move will be accepted. A high-temperature results in a higher probability that energetically unfavourable moves will be accepted, and a lower temperature decreases that probability. Therefore when we introduce attractive interactions into the system, we can treat the interaction strength as a balance with temperature i.e. In the case where the temperature is higher than the interaction strength, the thermal energy is higher than the interaction strength. Therefore the attractive interactions are cancelled out by the thermal energy.

By sampling the interaction energy space and holding all other parameters constant, aggregation behaviours are recorded. Varying the interaction energy can represent the sampling of different conformations of amyloid monomers able to undergo structural changes. This is because confor-

mation affects the hydrogen bonding potential of one monomer and, therefore, its interactions with another monomer in that conformation. We also vary the interaction strength with temperature, so it is possible to determine further the most suitable conditions for the fibrillisation of rods. We first run a wide parameter sweep before focusing on the region of biological interest to get an overall picture of the phase space behaviour.

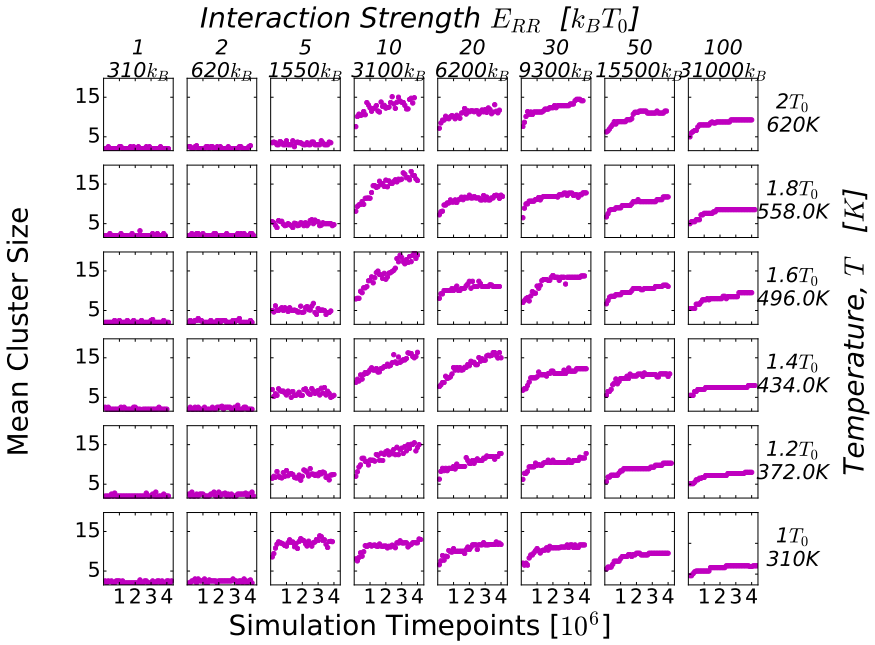

Figure 3.2: Grid of mean cluster size against time for temperatures in the range $T_0 \leq T \leq 2T_0$ ($310K < T < 620K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $1 < E_{RR} < 100$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 4 million Monte Carlo time steps.

Time Dependence of Cluster Averages In all simulations, the mean cluster size is used as an approximate indicator of whether a simulation has reached a steady state or dynamic equilibrium. This measure is a good indicator of, on average, how much aggregate content is present at a given time in a simulation. This may not always be fibril content, but in certain regions of the parameter space, there is only one expected aggregation morphology. If the mean cluster size has plateaued, then this is one indicator to suggest that simulation has reached its steady state as it shows there is no further growth in aggregate size except for rare events such as large cluster merging. However, it is often the case that there are few extremely large structures which skew a mean value. This is

why we also observe the length distribution for simulation data.

The first set of simulation results represents the time dependence of mean cluster size as the interaction strength and temperature is varied, as shown in figure 3.2. The parameter space can be divided into regions of different interaction strengths E_{RR} in the range $1k_BT_0 \leq E_{RR} \leq 100k_BT_0$. We also explore temperature incrementally in the range $1T_0 \leq T \leq 2T_0$ ($310K \leq T \leq 620K$). Each box in a grid represents the time-dependent data of a single simulation i.e. One can pick a pair of parameters (an interaction strength and a temperature) and observe how the mean cluster size changes over time.

For low rod-rod interaction strengths, where $E_{RR} = 1kT_0, 2kT_0$, the mean cluster size is consistently less than 5, meaning that there are very low levels of cluster formation under these conditions. In this region of phase space, the interaction strength is of the order of the thermal energy for all temperatures. Therefore each of these systems will be largely unable to form stable clusters and will only form short-lived clusters which rapidly dissociate.

For high interaction strengths ($> 20k_BT_0$), the mean cluster size reaches a plateau at a given value. This value increases slightly with temperature but not significantly. However, we observe that the plateau mean cluster size value decreases with increasing interaction strength. The interaction energy for larger interaction strengths will be significantly larger than the thermal energy, which decreases the chance that a fluctuation away from a "perfect" parallel binding move will be accepted.

The largest mean cluster size is found in the intermediate region where $5k_BT_0 < E_{RR} < 20k_BT_0$ for all temperatures, and the overall behaviour changes rapidly, therefore it is of interest to explore this region further. For constant temperatures, the behaviour is comparable to sweeps in section 3.3.2 where the mean cluster size plateau peaks and then decreases for increasing interaction strengths. Conversely, for constant interaction strengths and varying temperatures, the mean cluster size appears to be roughly the same except for $E_{RR} = 5k_BT_0$, where there is a distinct drop in mean cluster size with increasing temperatures.

Structurally, we only observe one major morphology for large aggregates of $AR = 4$ rods, which we have used in this parameter sweep. However, this will be discussed in further detail in a later section 3.3.2. This expected morphology is called a parallel stack in line with Vacha et al.[4], where there are many layers of pairs of monomers bound together to form a chain. The lengths of these aggregates can vary greatly, and each simulation will end with a length distribution of differently sized aggregates primarily in this conformation. In almost all the simulations where we observe aggregation, there is no typical lag phase that we would expect in experimental data collection on fibril aggregation due to the lack of an explicit nucleation mechanism. The growth period is distinct with a steep gradient consisting of monomer and dimer binding to existing clusters. This then leads to a plateau once the monomer in the system has depleted. This happens on quite short timescales (within the first 1-2M MC steps) due to the number of different clusters. In a model with an explicit nucleation mechanism, there would be a distinct lag period wherein we are waiting for the nucleation events to occur, and there would be fewer fibrils due to this. This model has no energy barrier to nucleation, meaning we move straight to the growth phase. As well as saving computational expense in reaching the steady state, an advantage of this model is that it does not make assumptions concerning the nucleation mechanism. On the other hand, it results in a less biologically accurate model in the short timescales of the system. Therefore we will discuss the dynamics of different multimer formations while keeping in mind the model's limitations that the most biologically accurate aspect is the steady state.

Time Dependence of Monomer Content In this section we highlight the time dependence of monomer content and overall fibril content across figures 3.3 and 3.4. This data is from the same set of simulations as figure 3.2, where we explore the interaction strength and temperature phase space incrementally. It is useful to observe these population sizes as an indicator of when each simulation system has reached a steady state - in combination with the mean cluster size. The monomer population also measures how stable assemblies are under different conditions. If there is a consistently significant monomer population, this may be indicative that there is a dynamic equilibrium of association and dissociation events.

There are significant monomer populations for all temperatures at interaction strengths $E_{RR} =$

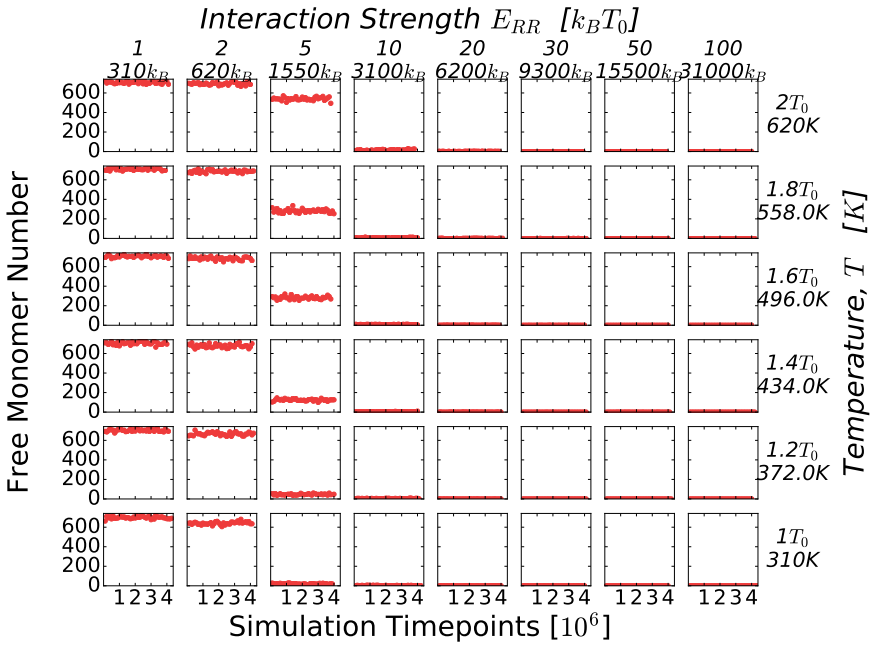

Figure 3.3: Population of monomers against time for temperatures in the range $T_0 <= T <= 2T_0$ ($310K < T < 620K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $1 < E_{RR} < 100$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 4 million Monte Carlo time steps.

$1kT_0$ and $E_{RR} = 2kT_0$ as shown in figure 3.3. The transition from monomer dominant to cluster dominant occurs at $E_{RR} = 5kT_0$. The monomer population increases with increasing temperatures because the thermal energy becomes comparable to the interaction strength as temperature increases, making clusters less stable. All interaction strengths above this are high enough compared to the thermal energy that clusters will form and will be more likely to be stable upon formation.

The converse is true for the population of clusters or fibrils (Figure 3.4) wherein there is almost entirely cluster or fibril content for interaction strengths of $E_{RR} > 5kT_0$ for all temperatures, a transition with mixed clusters and monomer populations at $E_{RR} = 5kT_0$, and no significant fibril content for $E_{RR} < 5kT_0$.

Time Dependence of Small "Oligomers" For low interaction strengths, the dimer population dominates with a high variance in population size throughout the simulation. This is due to the interaction strength being high enough to incur binding events but not enough that they are stable

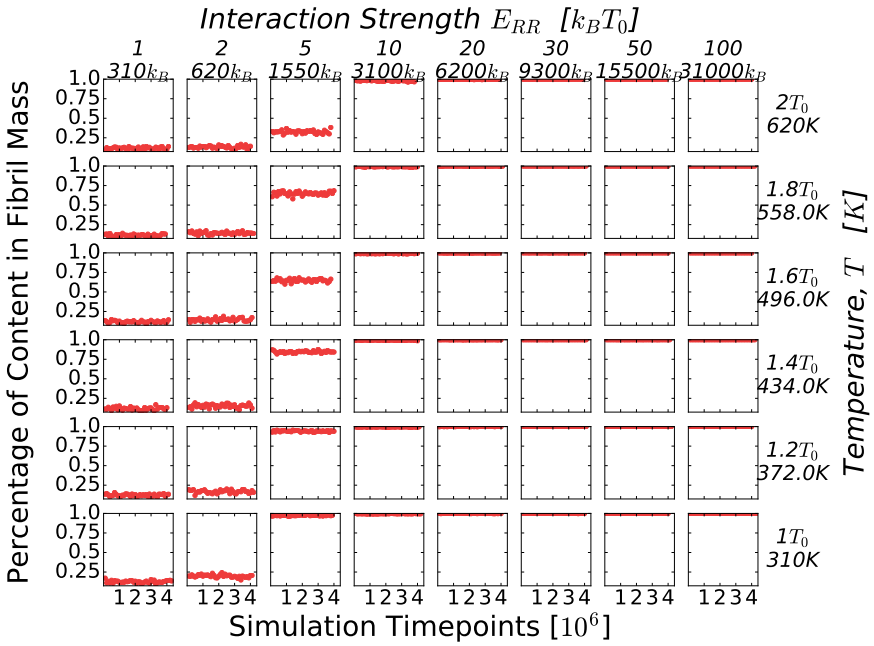

Figure 3.4: Population of fibrils as a fraction of total particle number against time for temperatures in the range $T_0 \leq T \leq 2T_0$ ($310K < T < 620K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $1 < E_{RR} < 100$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 4 million Monte Carlo time steps.

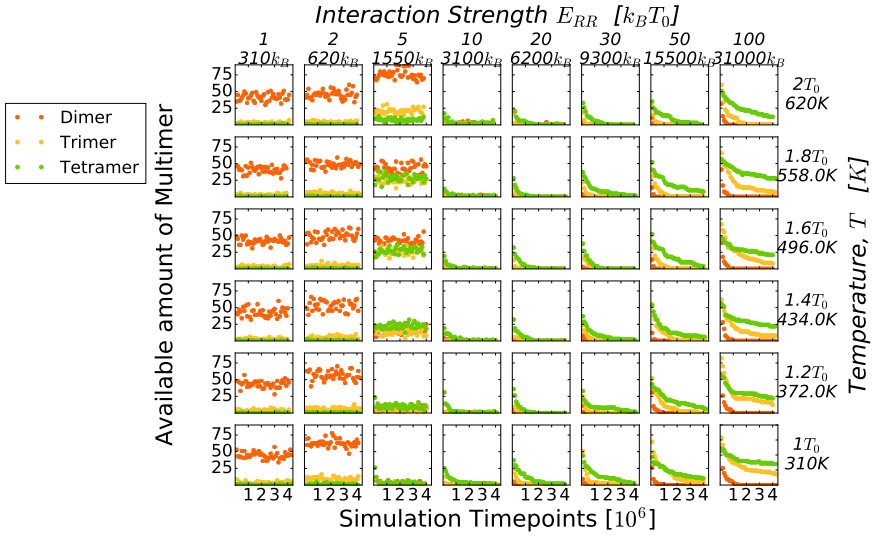

Figure 3.5: Populations of dimers, trimers and tetramers against time for temperatures in the range $T_0 \leq T \leq 2T_0$ ($310K < T < 620K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $1 < E_{RR} < 100$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 4 million Monte Carlo time steps.

structures that grow into extended fibrillar structures. When interaction strength $E_{RR} = 5kT_0$ there is a significant tetramer population for higher temperatures.

At the highest interaction strengths, and for the majority of the central region of this parameter space, there is an initial value of all populations which decreases throughout the simulation. The dimer population is consistently the first to disappear, followed by the trimer and then tetramer populations. This occurs because, in most simulations, we observe larger clusters and fibrils that will transiently pass these smaller multimer populations on the pathway to forming larger structures.

The dimer population at $E_{RR} = 5kT_0$ increases with temperature. This occurs because larger structures become less stable with increasing temperature due to the relative balance between interaction energy and thermal energy, resulting in more dimers which cannot grow into larger structures for high temperatures. For lower temperatures, there is a crossover in relative multimer populations where the tetramer population becomes the most stable of the three at $T = 1.4T_0$. There is an interesting set of behaviours for interaction strength $E_{RR} = 5k_BT_0$. The pentamer is the dominating species for the highest temperatures in this phase space ($T >= 1.6T_0K$). With decreasing temperatures, the larger species become more stable; therefore, their populations dominate e.g. Hexamers for $T = 1.2T_0K$ and heptamers for $T = T_0K$.

In the central region of this parameter space, where the interaction strength varies from $5kT_0 < E_{RR} < 30kT_0$, there are very low numbers of all multimers throughout the whole simulation. In line with later data, this region is where most elongated fibril formation occurs. Further studies later in this section narrow the search for the most suitable interaction strength for fibril formation.

Consistent with other data for this set of simulations, there are no stable populations of pentamers, hexamers or heptamers for low interaction strengths. The noisy data show this around the x axis, where the population size is 0. (Figure 3.6.)
For all interaction strengths $E_{RR} >= 10k_BT_0$, all multimer populations decrease over time with an initial value. These populations are likely intermediate species indicated by the mean cluster size plateau values for each simulation, which are all larger than 7 (heptamer). This indicates

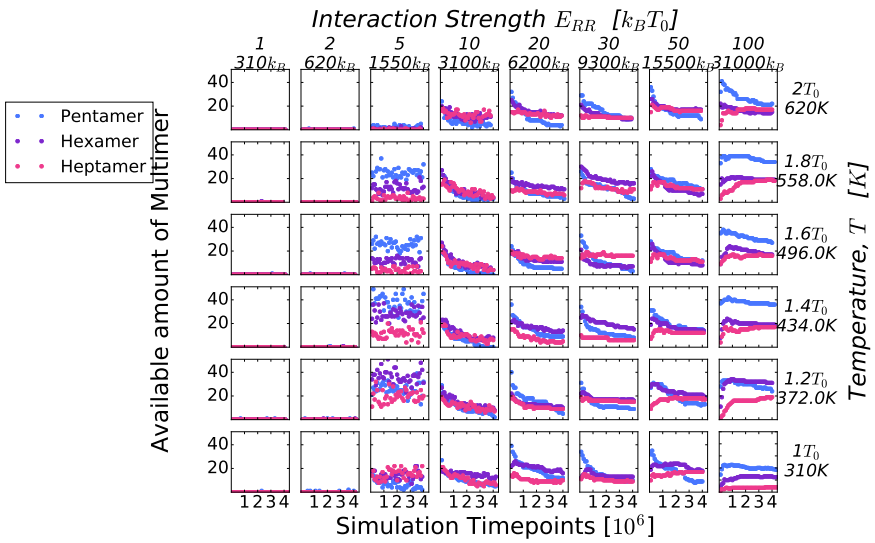

Figure 3.6: Populations of pentamers, hexamers and heptamers against time for temperatures in the range $T_0 \leq T \leq 2T_0$ ($310K < T < 620K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $1 < E_{RR} < 100$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 4 million Monte Carlo time steps.

that the majority of these simulations have steady states containing stable fibril or fibril-like species.

This observation starts to break down for the highest interaction strength $E_{RR} = 100k_BT_0$ where the mean cluster size is small for all temperatures. This is due to low mobility as a result of extremely high interaction strength. In cases such as this, where the mobility of clusters is low, it decreases the likelihood that small clusters will bind together and grow into larger clusters/fibrils. This may also be an artefact of the simulation scheme - the ensemble includes a constant particle number. Therefore the number of monomers available at the start of the simulation is finite, which would not be the case *in-vivo*. This means that after an initial series of nucleation events, there is frequently a very low amount of monomer available - this might lead to non-physical simulation results. This is not a large issue for this specific case as the interaction energies are very high compared to the thermal energy, which is not of direct biological interest to this work.

Cluster Averages and Summary of Data In this subsection, we consider two final figures. Figure 3.7, shows a length distribution for the steady state of each simulation in the previous sweep

of interaction strength and temperature parameter space. These distributions are one of the most useful tools in our analysis toolkit, as it allows us to fully observe the distribution of aggregates at equilibrium. The other key analysis tool is shown in figure 3.8, which summarises steady state behaviours across several subfigures. These subfigures are heatmaps of aspects of a given size distribution shown in figure 3.7, including mean cluster size, largest cluster size, median cluster size and number of clusters. In addition to this, heat maps of the concentrations of spherocylinders are also shown, as well as a heatmap of the average orientational order parameter, which is weighted according to aggregate size. All of these observable parameters provide insight into the overall trends of a given phase space, as well as knowledge about the level of order in a given system. A summary of all cluster properties over the range of E_{RR} and temperature is shown using heatmaps (Figure 3.8 and 2D curves (Figure 3.9).

The length distributions in figure 3.7 show two transitional regions that depend primarily on the interaction strength. Temperature changes the distributions' shapes slightly, but generally, the driving parameter is interaction strength. These two transition regions occur at $E_{RR} = 5kT_0$ and $E_{RR} = 50kT_0$. These two points are the boundaries for the parameter space's central region, resulting in the formation of larger structures of varying size and shape. Below $E_{RR} = 5kT_0$, we find that the system is dominantly monomeric with low numbers of dimers that are unlikely to be stable. This can be explained by the interaction strength being comparable to or lower than the thermal energy. The second transition in behaviour occurs for extremely large interaction strengths where the interaction energy is so strong in comparison to the thermal energy that thermal fluctuations are much less likely to occur. Therefore it becomes less probable that small clusters will be able to bind together to form more fibril like clusters. This region's orientational order parameter, S, is very close to 1. However, the mean cluster size is small, resulting in highly ordered but very small clusters.

Figure 3.9 highlights the importance of $E_{RR} = 5k_BT_0$ as a key parameter highlighting the transition between broad aggregation behaviour for lower and higher interaction strengths. Therefore, we must explore the region of parameter space around $E_{RR} = 5k_BT_0$ further. We have identified many regions are capable of forming fibril-style aggregates, but the region of parameter space where

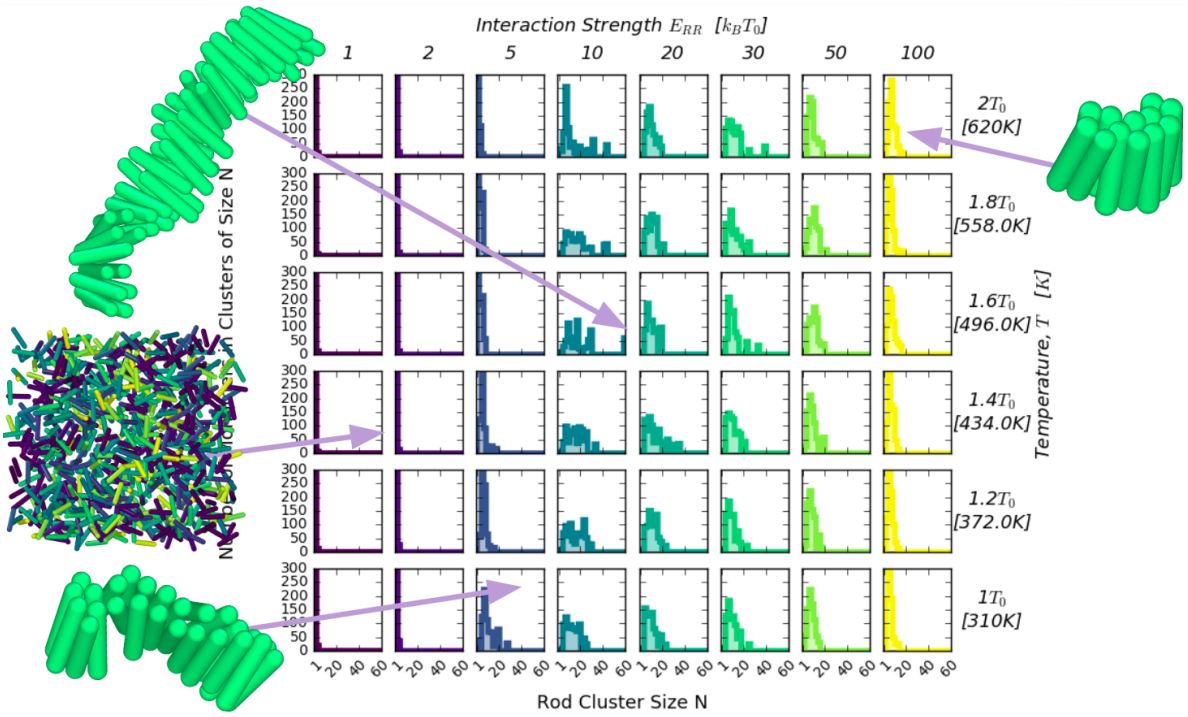

Figure 3.7: Cluster size distributions for the steady state assemblies as a function of interaction strength and temperature. For temperatures in the range $T_0 \leq T \leq 2T_0$ ($310K < T < 620K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $1 < E_{RR} < 100$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 4 million Monte Carlo time steps. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{rod}$. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$. Each bin equates to a discrete cluster size, N, which is filled with the total number of monomers that are found in a cluster of size N. Arrows point to example structures found in specific areas of the parameter space.

extended fibril formation occurs is narrow.

3.3.1.1 Magnified Interaction Strength Sweep Close to Physioligcal Conditions

For this magnified search of the parameter space in the region of interest determined in the previous section (i.e. Around $E_{RR} = 5k_BT_0$ and $T = 310K$), we can establish that the simulations have reached a steady state by using the time dependent mean cluster size (Figure 3.10). In addition to this, we incorporate the summary tools employed in the previous section. A summary of all cluster properties over the range of E_{RR} and temperature is shown using heatmaps (Figure 3.11 and 2D curves (Figure 3.12).

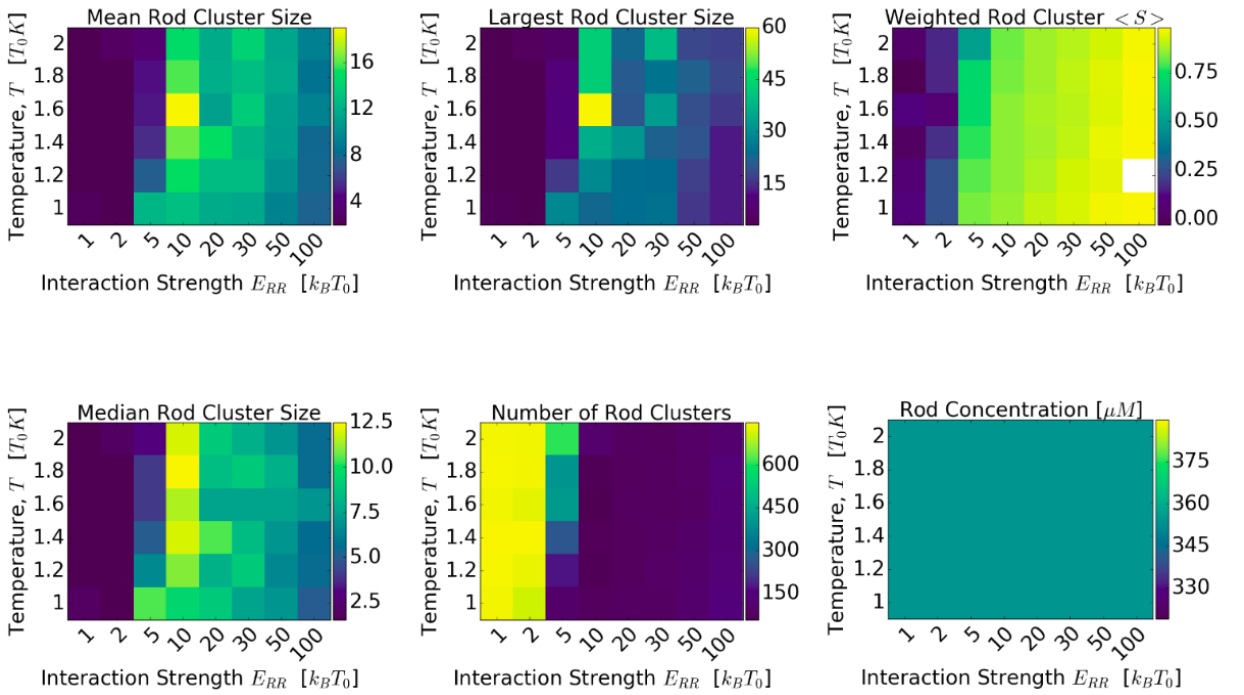

Figure 3.8: A series of heatmaps depicting the steady state results of 2D parameter sweeps for temperatures in the range $T_0 \leq T \leq 2T_0$ ($310K < T < 620K$) along the y axis (right-hand side) and rod-rod interaction strengths in the range $1 < E_{RR} < 100$ along the x-axis (top). Heatmaps from left top right. Top row: Mean rod cluster size, largest rod cluster size, weighted orientational order parameter. Bottom row: Median rod cluster size, number of rod clusters, rod concentration. Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 4 million Monte Carlo time steps.

There is a significant region of interest in figure 3.10 where the interaction strength is $E_{RR} = 6kT_0$ which results in the largest mean cluster size. Along the main diagonal of the heatmaps, we observe an apparent "sweet spot" in balancing interaction energy and thermal fluctuations to create the largest fibril structures. This diagonal equals approximately $E_{RR} = 6kT$ (wherever the interaction strength $E_{RR} = 6 \times T$, observe that E_{RR} are in units of $k_B T_0$.)

The endpoint mean cluster size data in figure 3.11 show a clear peak in cluster size in the central regions where temperatures $0.95T_0 <= T <= 1.0T_0$ for interaction strengths $E_{RR} = 6kT_0, 7kT_0$. This is also highlighted in the largest cluster size heatmap - the data is noisier here. However, it is clear that this region of parameter space produces the largest fibrillar structures. The column region where $E_{RR} = 4kT_0$ shows a significant decrease in all cluster size averages and an uptick in

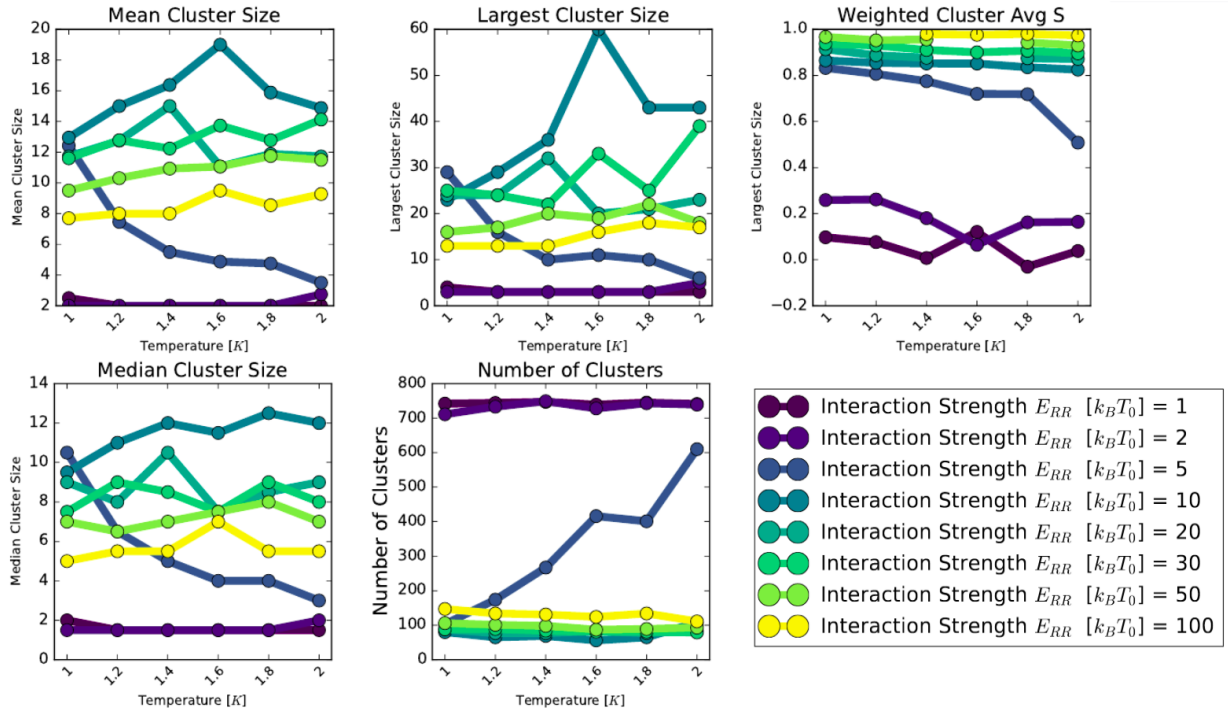

Figure 3.9: A series of data for simulations with temperatures in the range $T_0 \leq T \leq 2T_0$ ($310K < T < 620K$) along the x axis and rod-rod interaction strengths in the range $1 < E_{RR} < 100$ in a series of colours (purple - green - yellow). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 4 million Monte Carlo time steps. From left to right, top row: Mean rod cluster size, largest rod cluster size, weighted average orientational rod order parameter. Bottom row: Median rod cluster size, number of rod clusters.

the total number of clusters in the system, which is a clear indicator that this interaction strength is not large enough in comparison to the thermal energies in each simulation to produce larger stable clusters. This is confirmed by the specific case where $T = 1.2T_0$ - the largest thermal energy in this parameter sweep - results in the largest number of clusters, the smallest mean cluster sizes, and the least ordered structures.

Perhaps the most interesting result here is that the largest structures occur close to physiological conditions. In addition to this, the interaction strength E_{RR} for these structures is also close to the interaction energy for small protein folding.[335]. A second result is that the observed large structures are not necessarily the most ordered. The orientational order parameter, S, in figures 3.11 and 3.12 shows a clear trend that the higher the interaction strength, the more ordered the average cluster is. This makes physical sense that a higher interaction strength would result

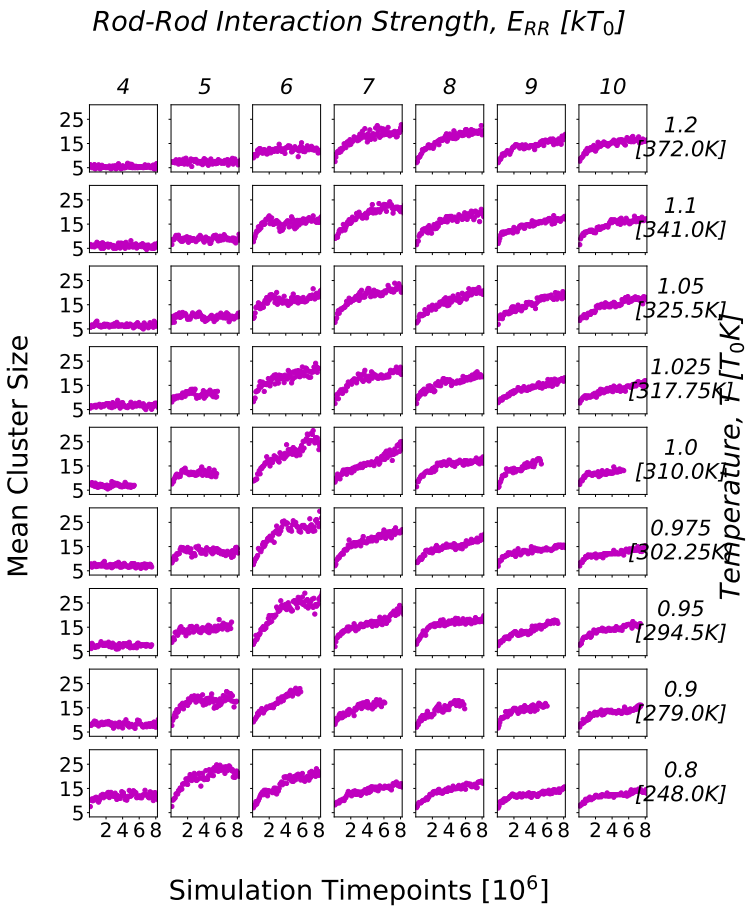

Figure 3.10: Grid of mean cluster size against time for temperatures in the range $T_0 \leq T \leq 2T_0$ ($310K < T < 620K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $4kT_0 < E_{RR} < 10kT_0$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 8 million Monte Carlo time steps.

in an interaction energy that dominates over the thermal fluctuations of a rod in a cluster and would keep rods more aligned. These higher interaction strengths, however, prevent the necessary thermal fluctuations for rods to diffuse enough to continue to bind to each other and grow. There is a question here as to whether we are observing a physical effect or whether this is a byproduct of the simulation set-up.

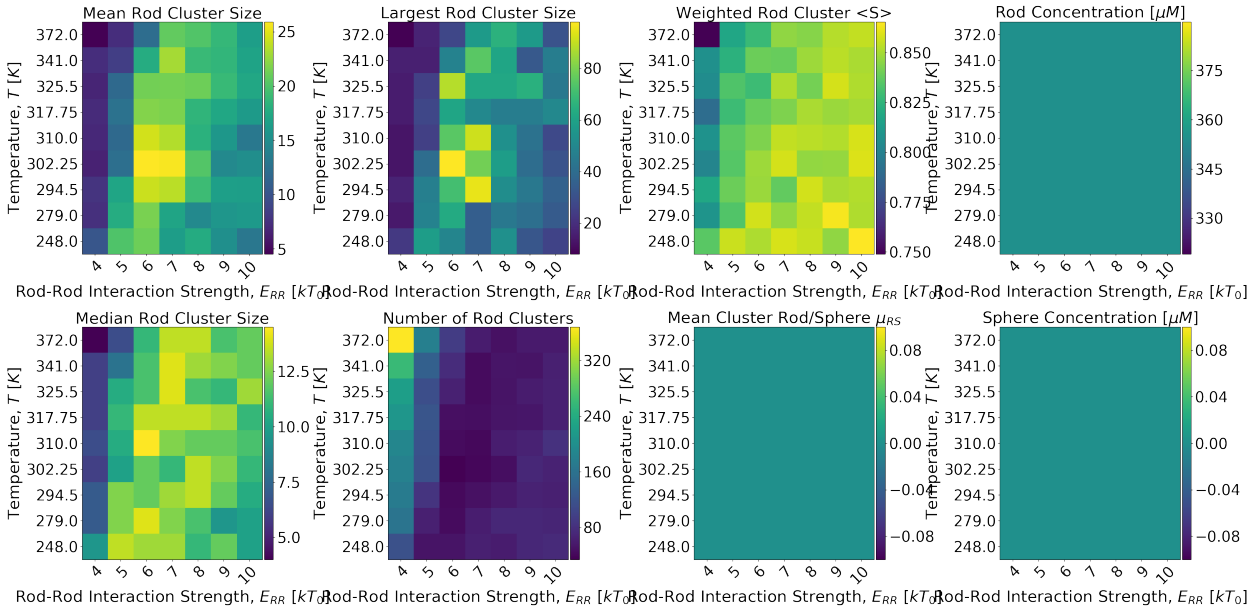

Figure 3.11: A series of heatmaps depicting the steady state results of 2D parameter sweeps Heatmaps from left top right. Top row: Mean cluster size, largest cluster size, weighted orientational order parameter, rod concentration. Bottom row: Median cluster size, number of clusters, mean cluster rod/sphere and sphere concentration. For temperatures in the range $0.8T_0 \leq T \leq 1.2T_0$ ($248K < T < 372K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $4kT_0 < E_{RR} < 10kT_0$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 8 million Monte Carlo time steps.

3.3.2 Shape Parameters and Temperature

Biological processes only function in the expected fashion for an extremely narrow set of conditions or parameters. For example, all proteins are functional at 37°C in the body, but begin to denature at temperatures as low as 41°C, merely 4°C higher. The melting points will differ for all proteins, but hydrogen bonds begin to break down within many proteins at around 41°C. This demonstrates the severe effects of even slight temperature increases in the body due to a fever could be[336]. Comparably, many physics research areas, such as astrophysics, explore a considerable variance of temperatures. It is easy to consider that many solar bodies have different temperatures that vary by thousands of degrees rather than by tens. However, these temperature variances are much more significant because astrophysicists study huge astronomical bodies and not microscopic processes within the human body. In simulation systems, it is essential to look at extreme regions of the parameter space to thoroughly benchmark the system and, in turn, to ensure the physics of the system behaves as expected.

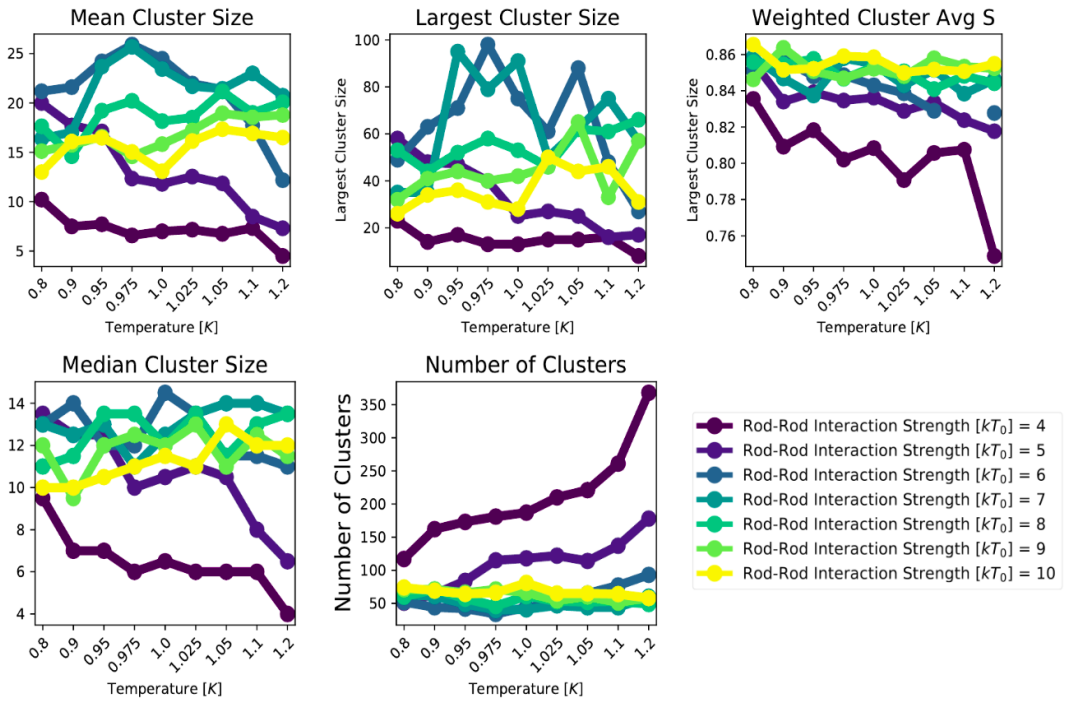

Figure 3.12: Summary of steady state cluster properties over a focused range of E_{RR} and T. For temperatures in the range $0.8T_0 \leq T \leq 1.2T_0$ ($248K < T < 372K$) along the y axis (right hand side) and rod-rod interaction strengths in the range $4kT_0 < E_{RR} < 10kT_0$ along the x-axis (top). Each simulation is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ and aspect ratio $AR = 4$ at volume fraction $VF = 0.05$ and runs for approximately 8 million Monte Carlo time steps.

We incorporate a reduced units scheme centred about $T_0 = 310K$. All interaction strengths are multiples of k_BT_0 to be consistent in comparing simulation parameters and for ease of understanding. We found in section 3.3.1 that the rod-rod interaction strength which produces the most numerous and largest fibrils is $E_{RR} = 6k_BT_0$. Therefore this is the interaction strength which will be utilised for all simulations following.

Aspect ratio is one of, if not the most important variable(s) in this thesis. Phase behaviour can change drastically by changing the aspect ratio[4]. Defined as $AR = L/D$, the aspect ratio is simply a ratio of the length of a particle, L divided by its diameter D. The aspect ratio parameter is a shape parameter that would not be observed or used within a kinetic model. The geometric parameters available are the critical difference between kinetic or numerical modelling and simulations of

physical objects. Numerical modelling using kinetic models simply cannot capture steric effects and the variation in orientation the way that DMC simulations can. For the case of spherocylinders, if the aspect ratio is 1, then the particle becomes a sphere. Larger aspect ratios can be needle-like and, in extreme examples, are essentially lines. We set the aspect ratio in all simulations as part of the initialisation process. We are interested primarily in lower order aspect ratios in the range $2 < AR < 16$ to stay in the biologically relevant regime. Simulations so far have used a geometry where $AR = 4$; this is more specific than one might realise. Beta strands in many amyloid-forming proteins can be approximated to aspect ratio 4. It is important to look beyond this specific parameter space to understand what is important about this key area of the parameter space and what makes this region able to form extended fibril structures.

In the following phase diagrams, this chapter shows that the most favourable fibril forming conditions exist where the aspect ratio (particle length/diameter) $AR = 4$ and the temperature is $T = 310k$. Therefore the majority of simulation data presented in this thesis uses these conditions. It is important from a physics perspective to simulate a much wider variance of temperatures than is important to the biology specifically to benchmark the system's physics appropriately. It is easier to highlight issues during testing by first screening extremes. Therefore simulations in this section and in section 3.3.1 are carried out at temperatures of up to thousands of kelvin. The level of coarse-graining in this system is such that the aspect ratio of the amyloid monomers is one of the few quantities which can be tuned to approximate specific amyloidogenic protein monomers. The size of amyloid-forming monomers can vary greatly. However, the regions of beta-sheet content, which represent the structural units within sometimes much larger, flexible precursors, tend to be of comparable size. Therefore the aspect ratio of many amyloidogenic "structural units" within their respective fibrillar conformations is similar and interestingly close to 4, where these data sets peak.

3.3.2.1 Exploration of Aspect Ratio - Temperature Parameter Space

Figure 3.13 broadly shows the overall steady-state behaviours of the parameter sweep between the shape parameter aspect ratio and temperature. In order to keep the volume fraction constant,

the rod concentration is highest for aspect ratio $AR = 2$ and lowest for $AR = 16$. However, the concentrations for all of these simulations are artificially higher than *in-vitro* experiments to reduce computational expense by reaching a steady state sooner.

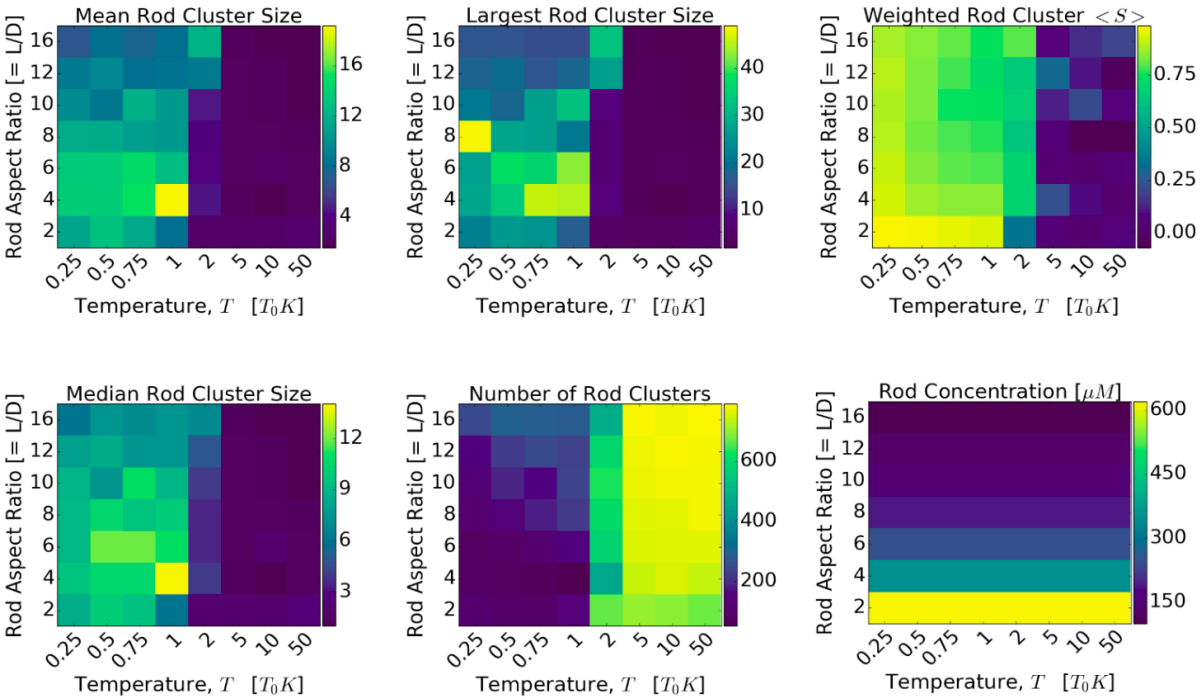

Figure 3.13: Heatmaps for the properties of clusters obtained from simulations where AR increased from 2 to 16 along the y axis and increasing temperatures from $77.5K$ to $15500K$ ($0.25T_0 \leq T \leq 50T_0$) along the x-axis. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{ROD}$. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$. All data taken from endpoints of simulations ≈ 8 million Monte Carlo time steps long. Top (left to right): Heatmaps of mean rod cluster size, largest rod cluster size and average orientational rod order parameter weighted to the cluster size S. Bottom (left to right): Heatmaps of median rod cluster size, number of rod clusters and rod concentration.

The mean cluster size for temperatures lower than $T = 2T_0$ has a large variance. There is a peak in all cluster size estimates around $AR = 4 - 6$ with a distinct outlier at $AR = 4$, $T = T_0$ showing a mean cluster size of 19. There is a marked change in behaviour for temperatures greater than or equal to $2T_0$ (Figure 3.13.) Almost all the data shows systems comprised of size five monomers or lower clusters. For the highest temperature, this small cluster size peak narrows and is almost exclusively monomeric or dimeric. There is one specific exception to this at $AR = 16$, where the

distribution skews into a small quantity of larger clusters containing up to 25 monomers. There is a small skew for all aspect ratios at this temperature which becomes more pronounced with increasing aspect ratio.

The order parameter S also reflects a distinct change in behaviour between $T = 2T_0$ and $T = 5T_0$. For lower temperatures, the order parameter is $\approx 0.6 - 0.8$, and in higher temperatures, there is a significant drop to near 0 in some specific simulations. Overall the order parameter is less than half the values observed for lower temperatures.

Overall, this means that in the low-temperature regime, we observe highly ordered aggregates of sizes that vary greatly with the aspect ratio of the monomer. We observe little to no aggregation in the high-temperature regime, where any aggregates formed are highly disordered. The mean cluster size in this region is consistently between 1 and 5, so it is likely that a small number of small multimers are forming, which are disordered due to thermal fluctuations. The most ordered aggregates are found when $AR = 2$, meaning that the most ordered aggregates are not necessarily the largest. The fibrillar structures observed in these simulations often bend and twist in sections the larger they get, so it might be the case that the $AR = 2$ aggregates are in a smaller, more rod-like domain. In contrast, the $AR = 4$ aggregates are in the longer, semiflexible regime.

Figure 3.14 shows a length distribution at the endpoint of each simulation for parameter sweeps of aspect ratio versus temperature. As illustrated by the previous figure 3.13, there are distinct regions of cluster formation ($T < 2T_0$) and no cluster formation $T > 2T_0$. Similarly to other analyses for this dataset, there is a transition at $T = 2T_0$ where there is some cluster formation. However, producing fibril-like structures is impossible as the thermal energy is too close to the interaction strength for small clusters to stay bound for long enough to attract additional monomer. The exclusions to the rule are for high aspect ratios $AR = 12$ and $AR = 16$, which achieve the formation of unusual clusters made up of cross-hatched ordered layers of monomers. It is possible to create the fibril and fibril-like structures in all of the lower range temperatures; however, the most uniform-like distribution of different cluster sizes is found at $AR = 4$, $T = 1T_0$.

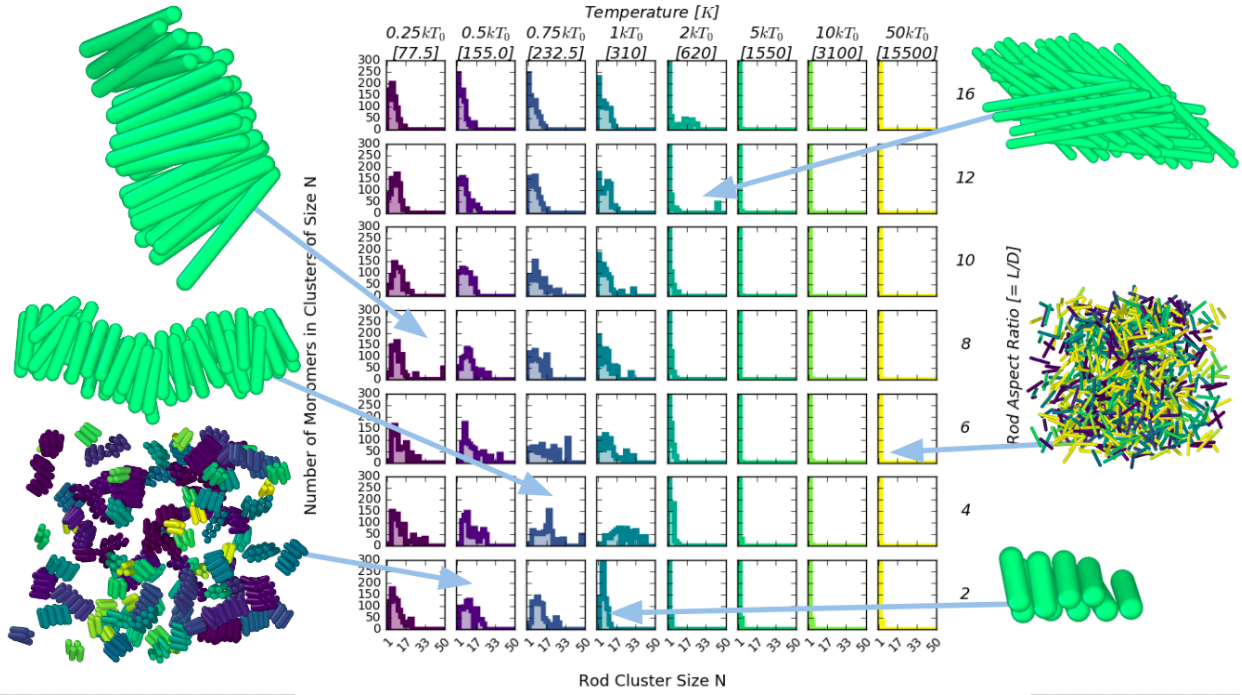

Figure 3.14: Histograms for the size distribution of clusters at the end point of each simulation. For increasing aspect ratios from 2 to 16 along the y axis (right hand side) and increasing temperatures from $77.5K$ to $15500K$ ($0.25T_0 \leq T \leq 50T_0$) along the x-axis (top). Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{rod}$. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$. Each bin equates to a discrete cluster size, N, which is filled with the total number of monomers found in a cluster of size N. Arrows point to example structures found in specific areas of the parameter space.

The resulting distributions at $0.25T_0$ and $0.5T_0$ are surprisingly similar to their higher temperature counterparts, with a strong peak for small cluster sizes and an extended tail into larger clusters. Some of these have distinct peaks away from the y-axis, showing a complete lack of monomer content in their respective systems. (Notably, $AR = 4$ and $AR = 6$ where the peak has moved). The most interesting results are for $T = 310k$ and $AR = 4$ again, simply because the distribution could be approximated as uniform, showing large populations in all cluster sizes with no monomer. It is also interesting how rapidly the behaviour changes by simply doubling the length of the particle. The difference in results at $T = 310k$ between $AR = 2, 4$ is staggering: a complete change in populations and the entire shape of the distribution, which for increasing aspect ratios appears to converge back to a distribution in a comparable shape as the $AR = 2$ system. There is a clear "goldilocks zone" in terms of the region of interest, which aligns well with the biological context of the system.

The trend with temperature for constant aspect ratio is also of interest. In each aspect ratio case, for increasing temperatures, there is a shift in the peak of the distribution until temperatures are reached where clusters can no longer form, i.e. at $T = 2T_0$. This trend is evident for mid-range aspect ratios from $AR = 4$ to $AR = 10$. The range of conditions that form long extended fibrillar structures appears to be surprisingly narrow given the model's simplicity. However, this is broadly comparable to many biological systems where the expected function results from a delicate balance of the biological processes and chemical interactions at play.

There is an apparent drop-off in order parameter S between temperatures $2kT_0 - 5kT_0$. This correlates well with the mean cluster size and also inversely with the number of clusters. The most ordered structures are found for the lowest temperatures and the lowest aspect ratios. However, the largest fibrils are found for aspect ratio $AR = 4$ spherocylinders.

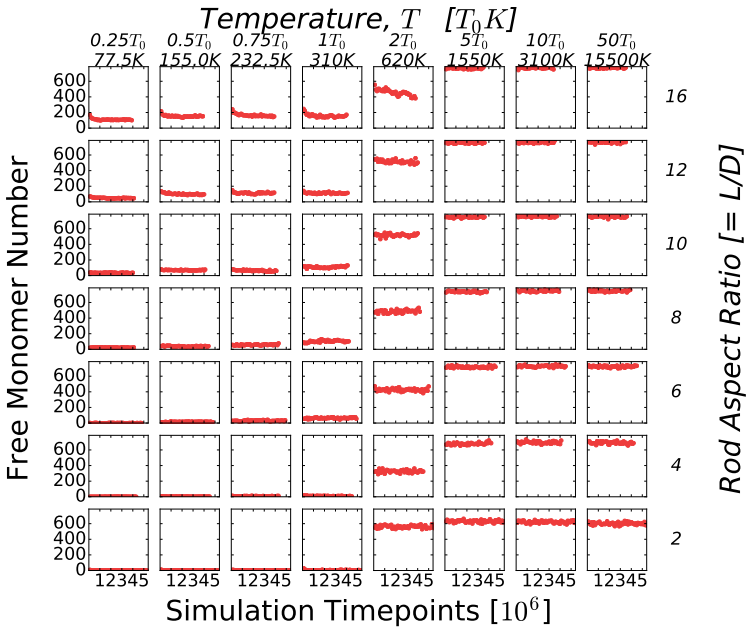

Figure 3.15: For increasing aspect ratios from 2 to 16 along the y axis (right-hand side) and increasing temperatures from $77.5K$ to $15500K$ ($0.25T_0 \leq T \leq 50T_0$) along the x-axis (top), each subplot shows the number of free monomers available in a system and how that varies with time. Time units are relative to Monte Carlo time steps wherein one sweep of N moves, where N is the total number of particles in the system, is considered one MC time step. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$ and interaction strength $E_{RR} = 6k_BT_0$.

Time Dependence of Free Monomer Content Figure 3.15 shows the time dependence of the free monomer content for each simulation in the 2D parameter sweep of aspect ratio and temperature. The bottom left of this phase space where aspect ratio and temperature are lowest shows that throughout the time course, there is little to no monomer available at any point in the simulation but the very beginning, meaning almost all protein content is made up into clusters for the whole simulation. This is the opposite case for all simulations at temperatures greater than $310K$. In these cases, a significant amount of free monomer is always available. This is simply because for higher temperatures, spherocylinders are more likely to make energetically unfavourable moves translationally or rotationally and are, therefore, far less likely to stay bound to other spherocylinders. For shorter rods, the temperature change is more drastic than for longer rods. However, this is an artificial result. The method of determining clusters is calculated based on the interaction length of the V_{RR} potential, which is calculated as a multiple of the diameter of the rod. The longer rods are far more likely to undergo steric clashes with surrounding particles for the same volume fraction as their length will be closer to the length of half the box size than a lower volume fraction. This leads to a slower aggregation process when compared with lower volume fractions.

For cases where the temperature is constant and the aspect ratio is varied (Figure 3.15), the monomer content increases with the aspect ratio. The most exciting result is that the monomer population is consistently at the same level throughout each simulation. In low-temperature cases, the monomer population is consistently deficient for all aspect ratios, implying that almost instantaneously, all monomer is sequestered into clusters of some form, whether this is dimeric or larger, as shown in other later figures. There is notable behaviour at $2T_0$ where around 50% of the monomers are unbound throughout the time courses. A consistent population of available monomers shows an ability for these systems to undergo a dynamic equilibrium of binding and unbinding of monomers to existing clusters and the formation and dissociation of small clusters. As this is simply an initial sweep of the system's phase space, it is clear that more exploration is needed between $1T$ and $5T$, where the monomer populations drastically change with temperature for all aspect ratios.

Time Dependence of Cluster Averages Simulations are often considered complete or to have reached a steady state when the mean cluster size for that simulation plateaus. High-temperature

simulations (Figure 3.16, right-hand side) consistently plateau early or instantaneously due to the inability of rods to form large clusters or fibrils. Larger aspect ratios more favourably form fibrils at higher temperatures than shorter rods. The mean cluster size does not appear very large in any region of this diagram. However, this is often because there are a small number of extremely large fibrils and a much larger number of small clusters. There is a clear peak in the mean at $1T$ and aspect ratio $AR = 4$, meaning that these conditions are the most favourable to forming fibrils. $AR = 4$ simulations consistently have the highest mean plateau for all temperatures. Higher aspect ratio simulations tend to take longer to plateau; this might be due to the artificial steric effects mentioned in the previous section. In order to obtain a more detailed description of the population sizes at the end points of the simulation, length distributions should be evaluated at the end of each simulation.

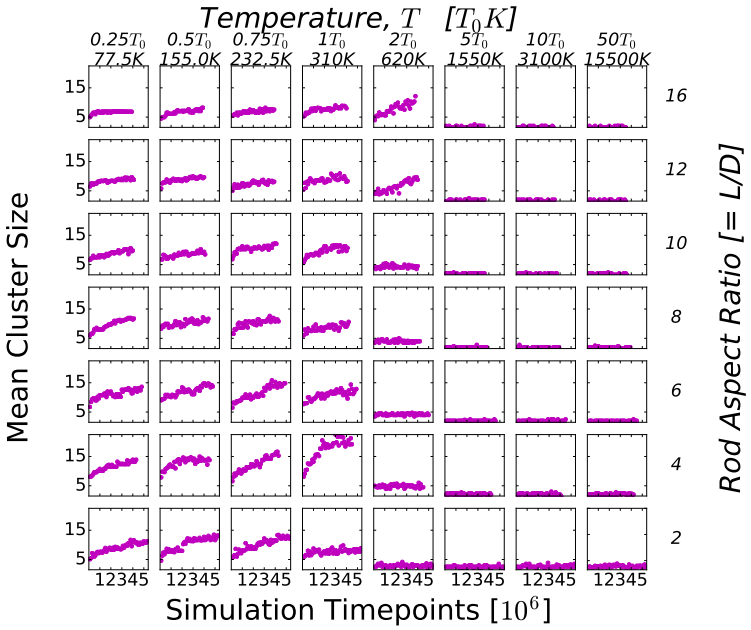

Figure 3.16: Time dependence of the mean cluster size as a function of aspect ratio and temperature. For increasing aspect ratios from 2 to 16 along the y axis (right-hand side) and increasing temperatures from $77.5K$ to $15500K$ ($0.25T_0 \leq T \leq 50T_0$) along the x-axis (top), each subplot shows the mean cluster size and how that varies with time. Time units are relative to Monte Carlo time steps wherein one sweep of N moves, where N is the total number of particles in the system, is considered to be one MC time step. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{ROD}$. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$ and interaction strength $E_{RR} = 6k_BT_0$.

Time Dependence of Largest Cluster Size Figure 3.17 presents the time dependence of the largest cluster size across a range of aspect ratios and temperatures. The observed trend is very similar but exaggerated concerning the mean cluster size in figure 3.16. The data shown here are noisier than in figure 3.16. The largest cluster size is not an averaged quantity and therefore fluctuates far more than the mean size. However, this is a useful tool to analyse which conditions have the highest propensity to form fibrils. These data often have secondary or tertiary plateaus due to late-time cluster-cluster binding, which is considered a relatively rare event.

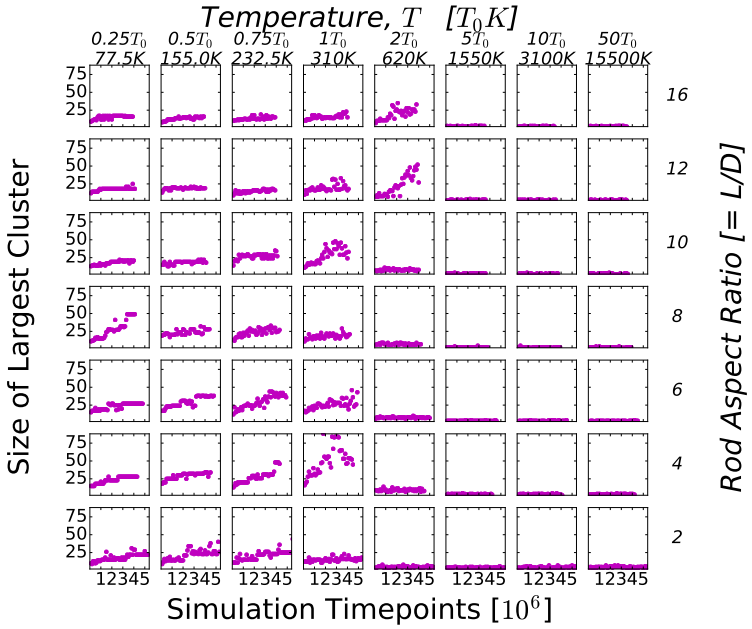

Figure 3.17: Time dependence of the largest cluster size as a function of aspect ratio and temperature. For increasing aspect ratios from 2 to 16 along the y axis (right-hand side) and increasing temperatures from $77.5K$ to $15500K$ ($0.25T_0 \leq T \leq 50T_0$) along the x-axis (top), each subplot shows the size of the largest cluster observed in a simulation and how that varies with time. Time units are relative to Monte Carlo time steps wherein one sweep of N moves, where N is the total number of particles in the system, is considered to be one MC time step. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{ROD}$. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$ and interaction strength $E_{RR} = 6k_BT_0$.

Overall, therefore, the longest fibrils are observed at low temperatures and aspect ratios lower or equal to 10, with a sweet spot at aspect ratio 4 and temperature $T = T_0$.

There is a significant uptick in cluster sizes for $AR = 16$, which falls out of trend with the other

data. At this aspect ratio, and for $AR = 12$, a different kind of cluster exists than the parallel stacked fibrils we will attempt to disrupt the aggregation process of in later chapters. These are cross-hatched stacks of rods that maximise the rods' apparent valency. Rather than simply stacking in pairs of parallel-oriented rods, each layer is perpendicular to the layers above and below it and the size of each layer is determined by the number of rods that can lie perpendicular to the rod(s) below it and still be in contact with the attractive patch.

Time Dependent Content of Small Oligomers As for interaction strengths, it is interesting to explore regions of aspect ratio versus temperature space where small oligomers are stable. Dimers, trimers and tetramers are presented in figure 3.18, and pentamers, hexamers and heptamers are shown in figure 3.19. There are two distinct behaviour regions comparable to other analyses for this parameter sweep. For temperatures less than or equal to $1T_0$ or $310K$, there is consistently a significant population of tetramers across all aspect ratios. There is a small exception for $AR = 16$ where the dimer population is slightly larger, but there is still a significant trimer and tetramer population.

Populations of pentamers, hexamers and heptamers decrease with time. This is expected in this region. All systems are producing significant and large fibril populations. Therefore, the populations of multimers decrease as the clusters continue to grow. There is a single exception at $AR = 2$, $T = T_0$, where the pentamer and hexamer populations increase with time. As discussed previously, the most favourable fibril forming conditions are found when $T = T_0K$ for aspect ratios $AR = 4$ and $AR = 6$. The populations of these multimers decrease rapidly as these multimers are simply intermediates which grow rapidly in these systems. The $AR = 4$ results show almost no population in any of these small clusters. This agrees with other analyses of these simulations, as other results show that this aspect ratio holds the largest and the largest mean cluster size. Therefore more of the particles exist in much larger clusters/fibrils. The $AR = 2$ simulations show the only significant population changes throughout a simulation - an overall decrease in all three populations throughout the time course for this region of temperatures. Almost all other simulations undergo a change very early and maintain a near-constant level for the remainder of the time steps. Whilst it is interesting to discuss the changes in population sizes throughout a time course, this model is

most reliable at the steady state due to the lack of a significant nucleation barrier. It would be of significant interest to pursue a model with a different nucleation mechanism in further work.

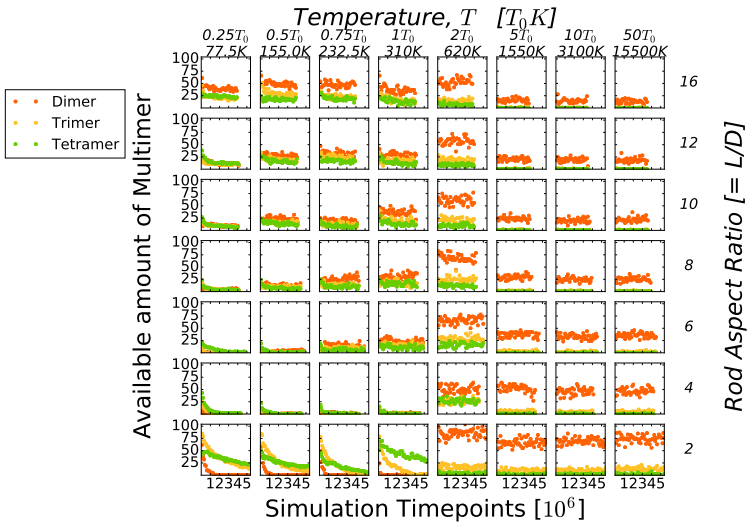

Figure 3.18: Dimer, trimer and tetramer population time dependence for increasing aspect ratios from 2 to 16 along the y axis (right-hand side) and increasing temperatures from $77.5K$ to $15500K$ ($0.25T_0 \leq T \leq 50T_0$) along the x-axis (top), each subplot shows the population of each multimer, how that varies with time. Time units are relative to Monte Carlo time steps wherein one sweep of N moves, where N is the total number of particles in the system, is considered to be one MC time step. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{ROD}$. A number of clusters of size 2 (orange), 3 (yellow) and 4(green) are shown on shared axes for each simulation of ≈ 5 million Monte Carlo time steps. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$ and interaction strength $E_{RR} = 6k_BT_0$.

For temperatures $2T_0$ and higher, the dominating population is dimeric across all aspect ratios. As discussed in other results, this is due to the decreased ability or likelihood of PSCs forming stable clusters. Therefore these clusters will be rapidly binding, unbinding and binding with other monomers throughout the timecourse. This is consistent with temperature, wherein the rightmost columns show a decreased ability to form trimers and tetramers when compared with lower temperatures. The dimers formed are transient in this regime and are a result of particles passing close to each other. At $T = 2T_0$, the temperature is high enough to prevent large fibril structures from forming but low enough to allow for some stable binding of monomers. The results of simulations at this temperature are the most significant as it allows the existence of the largest variance in species, each in significant populations to exist at one time i.e. There are some examples of short fibrils,

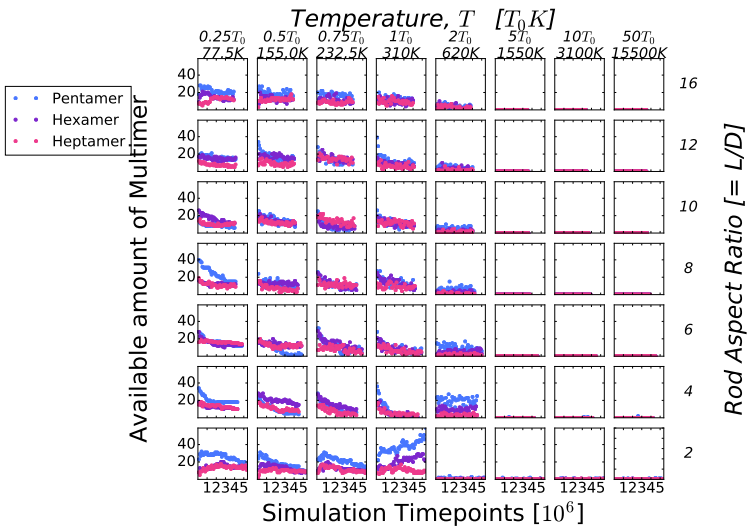

Figure 3.19: Pentamer, hexamer and heptamer population time dependence for increasing aspect ratios from 2 to 16 along the y axis (right-hand side) and increasing temperatures from $77.5K$ to $15500K$ ($0.25T_0 \leq T \leq 50T_0$) along the x-axis (top), each subplot shows the population of each multimer, how that varies with time. Time units are relative to Monte Carlo time steps wherein one sweep of N moves, where N is the total number of particles in the system, is considered to be one MC time step. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{rod}$. A number of clusters of size 5 (blue), 6 (purple) and 7(pink) are shown on shared axes for each simulation of ≈ 5 million Monte Carlo time steps. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$ and interaction strength $E_{RR} = 6k_BT_0$.

clusters of small multimers and a significant monomer population throughout the time course.

For increasing aspect ratio and constant temperature for all higher temperatures (RHS), it is more difficult to form trimers and tetramers. The $AR = 2$ simulations show some consistent populations, whereas, at the other extreme $AR = 16$, there is almost none of these species throughout the timecourses. The $AR = 2$ results more closely replicate the results for all aspect ratios at $T = 2T_0$. For temperatures larger than $T = 2T_0$, there are no clusters of this size except for a small population of pentamers for all T for smaller aspect ratios. This is expected due to the extremely low probability of larger multimers forming. This data shows that systems under these conditions favour the formation of small multimers and cannot support the formation of any structure larger than a pentamer. At $T = 2T_0$, the pentamer population is the largest for all aspect ratios, but all aspect ratios are also capable of producing hexamers and heptamers at this temperature. The size of each of these populations decreases with increasing aspect ratio, with the exception of $AR = 2$.

Figure 3.19 shows an opposing case of the previous figure showing dimer, trimer and tetramer populations. The left-hand side shows consistently large populations of pentamers, hexamers and heptamers, which in this case, do change with time.

3.3.2.2 Magnified Temperature Sweep Close to Physical Conditions

Fibrils and smaller oligomeric assemblies of interest are most highly represented at biologically relevant temperatures. A more focused sweep allows a closer look at this area of parameter space, where temperature is varied from $248K$ ($0.8T_0K$) to $620K$ ($2T_0K$) (figures 3.21 and 3.20.) For this set of parameters, the temperature changes are much smaller between each system, and the overall behaviour is similar in most cases. We chose these parameters to properly understand the relationship between aggregation behaviour and temperature much closer to physiological temperatures. The previous parameter sweep was used to benchmark the system and ensure the physics of the system was producing the expected results in accordance with the interaction strengths chosen.

A summary of all cluster properties over the range of E_{RR} and temperature is shown using heatmaps (Figure 3.21 and 2D curves (Figure 3.20).

Generally, with increasing temperature for aspect ratios $AR = 10$ or lower, the mean cluster size plateau value increases, peaks at some temperature and then decreases. For example, at $AR = 4$, there is a peak in mean cluster size at $T = T_0K$ and decreases for all increasing temperatures following that. Similarly, at $AR = 2$, the peak value of mean cluster size is at $T = 0.8T_0K$. Due to the similarity of this data, it is easier to observe the plateau values by visualising them differently. Amyloid-β oligomers, which may not have as significant regions of beta-strands within their constituent monomers when compared with fibrillar species, are prepared in the lab in a refrigerated environment. $248K$ equates to $-25°C$, representing such an environment. Therefore, this peak for $AR = 2$ is indicative of oligomeric structures and their preference to aggregate under cooler conditions than longer beta strands.

There is a distinct peak at $AR = 4$ in several of the plots shown. More specifically mean cluster

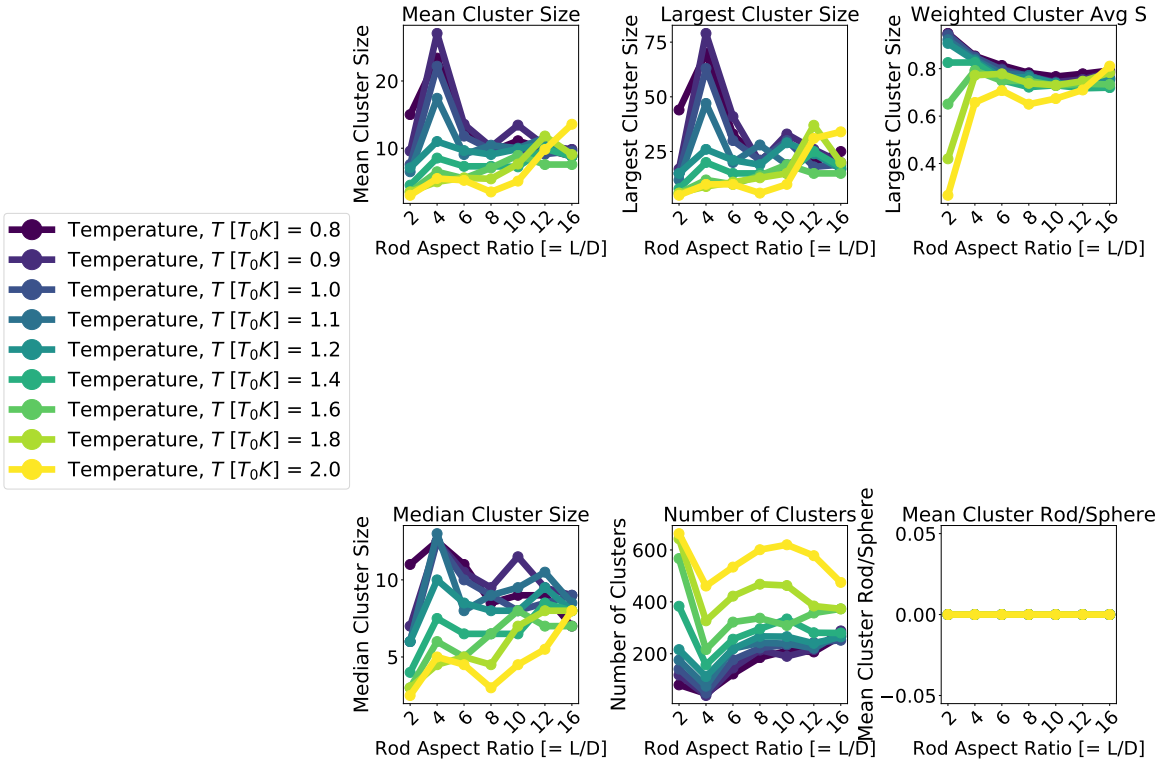

Figure 3.20: Summary of steady state cluster properties over a focused range of AR and T. For increasing aspect ratios from 2 to 16 along the y axis (right-hand side) and increasing temperatures from $248K$ to $620K$ ($0.8T_0 \leq T \leq 2T_0$) along the x-axis (top), each subplot shows a histogram for the size distribution of clusters at the end point of each simulation. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{ROD}$. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$ and interaction strength $E_{RR} = 6k_BT_0$.

size, largest cluster size and median cluster sizes all follow the same trends. The median cluster size has a smaller variance than the other average quantities, showing that the majority of conditions in this phase space are capable of forming clusters. This heatmap follows the same trend as previously discussed: the higher the temperature, the smaller the average cluster sizes for a given aspect ratio AR. The mean and largest cluster sizes show an almost identical data set, with the exception that the scale of the largest cluster size heatmap is different. Unsurprisingly, the region of interest producing the largest clusters also yields the largest mean cluster size. The number of clusters follows the same trend inversely. The lowest cluster numbers are found when $AR = 4$ irrespective of temperature, but the data does follow the same temperature trend in that there are

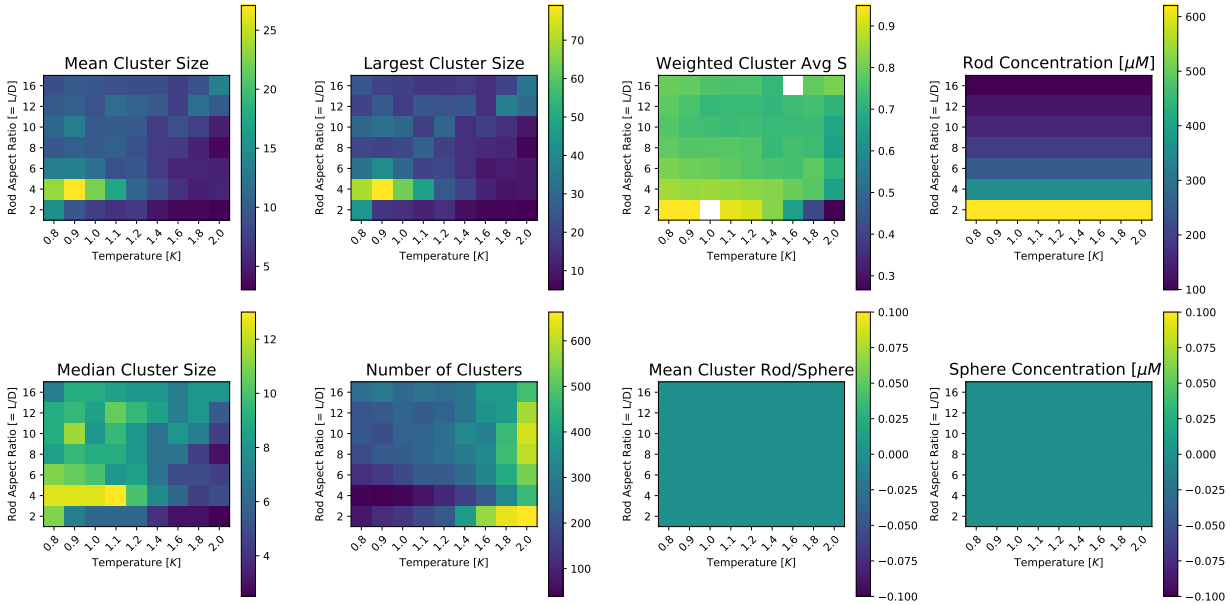

Figure 3.21: Heatmaps are showing steady-state properties of clusters for the properties of clusters obtained from simulations where aspect ratios vary from 2 to 16 along the y axis and increasing temperatures from $248K$ to $620K$ ($0.8T_0 \leq T \leq 2T_0$) along the x-axis. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{ROD}$. The total number of particles is $N = 800$ for a volume fraction $VF = 0.05$ and interaction strength $E_{RR} = 6k_BT_0$. All data taken from endpoints of simulations ≈ 5 million Monte Carlo time steps long. Heatmap measurements from left to right. Top row: Mean cluster size, largest cluster size, Orientational order parameter as a weighted average by cluster size, concentration of rods. Bottom Row: median cluster size, number of clusters at that timepoint, Ratio of mean rod cluster size and mean sphere cluster size, sphere concentration.

lower cluster numbers for lower temperatures. Likely because the clusters in this region are bigger, and therefore there are fewer clusters overall.

There are examples in the high aspect ratio and high temperature (top left figure 3.20) region of the phase space where different types of cluster morphology are observed. These clusters are cross-stacked so that each layer is perpendicular to the one before it. Structures such as this are observed in plant cells, but the aspect ratio would have to be much larger to accurately simulate a system such as this. Therefore this will not be the focus and is simply an aside. $AR = 16$ simulations will be omitted from future sweeps.

3.3.2.3 Order in Clusters

In figure 3.20 the orientational order parameter S shows two distinct peaks at $AR = 4$ and $AR = 16$ for all temperatures. There is a significant change in behaviour across all temperatures for aspect ratio $AR = 2$, which can be attributed to the equally drastic change in the range of cluster sizes for this aspect ratio. This data is also shown in the heat map in figure 3.21 where there is a clear variance in S across the line $AR = 2$. This variance in $AR = 2$ may be indicative of oligomeric properties. Oligomeric structures encompass a wide variety of structures that have unique order parameters and sizes. We can consider increased temperatures to decrease the relative interaction energy between pairs of rods. Due to the variety of potential oligomeric structures, it is not unreasonable to expect that the aggregation behaviours and observations will have the same variance in the regions of parameter space where these structures are expected to form i.e. $AR = 2$.

The orientational order parameter does not vary significantly in the remaining regions of this parameter space. There is a general decrease with increasing aspect ratio except for $AR = 16$, showing that the length of the rod does not necessarily improve the level of order within an aggregate. Increasing temperature does decrease the order parameter, but not significantly - even for the highest temperatures in this sweep, the order parameter only significantly drops for the $AR = 2$ case. The increase at the highest aspect ratio $AR = 16$ can be attributed to the alternative structures formed under these conditions, which are more populated at $T = 2T_0$. These structures would not produce a large order parameter as many particle neighbours are perpendicular to the rod. However, rods still lie parallel to each other within these aggregates, which would contribute mainly to the order parameter. This leads to a high order parameter but not as high as the parallel stacked aggregates.

It is notable that the largest fibrils are not necessarily the most ordered. This is a common occurrence across several data sets. This could be attributed to larger structures having more bends and twists than the shorter clusters because they are long enough to be considered semiflexible or even flexible. In contrast, the shorter clusters may be rod-like or close to the semiflexible regime with only small deviations away from the normal line of that cluster. This idea of flexibility moves away

from mature fibrils' physical behaviour *in-vitro*. EM and AFM data[337] within the Staniforth lab suggest that fibrils have an extremely high persistence length, so the simulation results may more accurately represent a proto-fibril structure or a fibril that has not reached full maturity. *In vitro* a fibril may undergo more long-range interactions that are considered in the scope of this project, so a further extension to the work might be to extend the code to model long-range interactions along the fibril for additional stability such as electrostatic interactions or stabilising interactions along aromatic groups. Alignment in structures such as the cryo-EM Aβ(42) in Gremer et al.[65] shows that a given monomer inside a fibril might interact with several monomers above and below it within the fibril.

3.3.3 Concentration Dependence

Typical concentrations for fibrillisation experiments with Aβ range from $1\mu M$ to $100\mu M$[8, 19, 70]. It is often more computationally efficient to simulate at higher concentrations than one would use in the lab to force relevant events to occur on shorter timescales to reduce the number of simulation hours required.

We convert between volume fraction VF and concentration c with the following relationships:

$$V_{ROD} = \frac{4}{3}\pi\sigma_R^3 + \pi\sigma_R L \tag{3.3}$$

$$VF = N_{ROD}V_{ROD}/L_{BOX}^3 \tag{3.4}$$

$$c = \frac{VF \times N_{ROD}}{L_{BOX}^3 \times N_A} \tag{3.5}$$

Where L_{BOX} is the length of one side of the cubic simulation box volume, initially calculated at the start of the simulation relative to the volume fraction of the particles, and N_A is Avogadro's constant. The concentration is now in units of mol/m^3; we can convert to micromolar simply by multiplying by 1000, where $1mol/m^3 = 1000uM$. We determine the box size of a given simulation from the desired volume fraction. We choose to keep the volume fraction the same at the expense of a changing concentration to prevent significant crowding events.

In this parameter sweep we vary the volume fraction of rods in the range $0.0001(0.01\%) \leq VF \leq 0.1(10\%)$ which corresponds to concentrations in range $50\mu M \leq c \leq 700\mu M$. Therefore we have artificially increased the concentrations compared with those used experimentally. Although experimental systems are relatively low in concentration, there is extensive evidence for fibrillisation requiring catalysis at surfaces. A prime function of such attractive surfaces is to increase the local concentration of particles[338]. Therefore, we use artificially high concentrations in these simulations to negate the computational expense of waiting for the local concentrations to increase and to increase the probability of fibril nucleation events. We revert back to a wide parameter sweep of temperatures in the range $77.5K \leq T \leq 15500K$ $(0.25T_0K \leq T \leq 50T_0K)$, and set the

interaction strength to $E_{RR} = 6k_BT_0$ based on the simulations in the previous section. (Section 3.3.1.) We fix the rod aspect ratio to $AR = 4$ based on the parameterisation carried out in section 3.3.2. We present several figures for our discussion: figure 3.22 represents the time dependence of the mean cluster size across the 2D parameter space of volume fraction and temperature, figure 3.23 and 3.26 indicate the time dependence of monomer and cluster content respectively. We also present figure 3.24 and figure 3.25 to highlight the time dependence of oligomer populations.

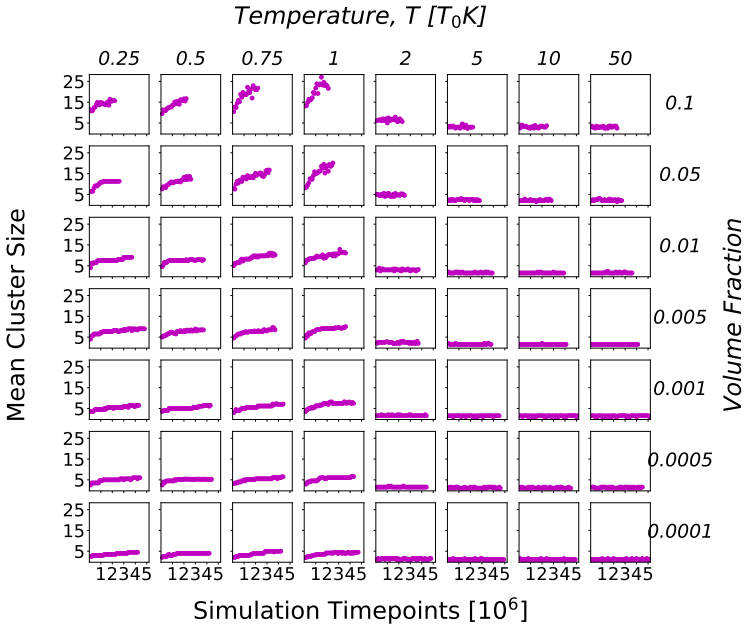

Figure 3.22: Grid of mean cluster size against time for temperatures in the range $0.25T_0 \leq T \leq 50T_0$ ($77.5K < T < 15500K$) along the x-axis (top) and fraction of volume that rods occupy in the range $0.01\% < VF < 10\%$ along the y axis (right). Each simulation is made up of $N = 1000$ rods with a patch length $L_p = 0.6L$ at aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$ and runs for approximately $4 - 5$ million Monte Carlo time steps.

All of these figures both show four distinct regions of behaviour in each of the four corners of this parameter space. Each area of parameter space will be discussed in detail in the following sections. (A. Low temperature, low volume fraction B. Low temperature, high volume fraction C. High temperature, low volume fraction D. High temperature, high volume fraction) For all regions of the parameter space, except for the low volume fraction and high-temperature region, we observe populations of multimers. For all low-temperature cases, we see varying populations of multimers that change with time, indicating growth throughout the simulations. This is particularly clear for low volume fraction cases where fibril growth is happening slowly, so we can clearly see the

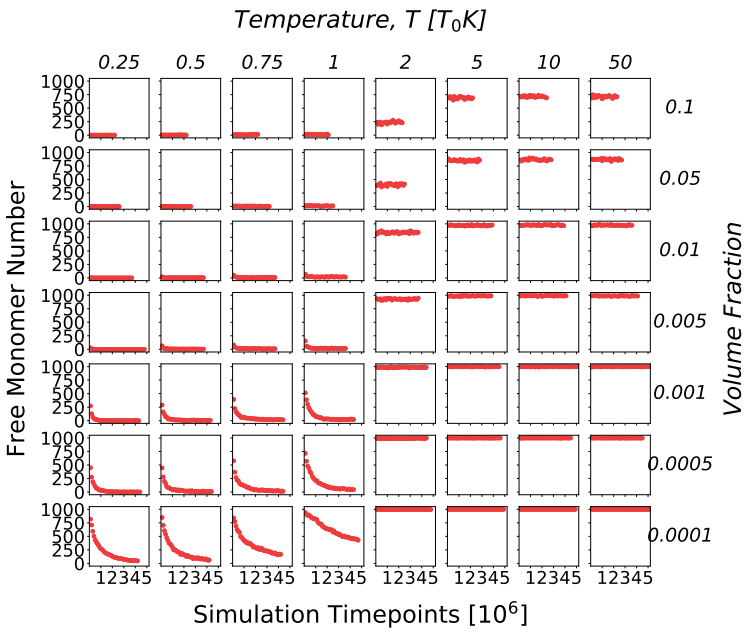

Figure 3.23: Population of monomers against time for temperatures in the range $0.25T_0 \leq T \leq 50T_0$ ($77.5K < T < 15500K$) along the x-axis (top) and fraction of volume that rods occupy in the range $0.01\% < VF < 10\%$ along the y axis (right). Each simulation is made up of $N = 1000$ rods with a patch length $L_p = 0.6L$ at aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$ and runs for approximately $4 - 5$ million Monte Carlo time steps.

transitions of initially high dimer populations progressing into other larger multimers.

3.3.3.1 Low temperature, low volume fraction

The monomer population 3.23 shows a sharp decrease during the growth phase and, comparably, the fibril population shows a converse sharp increase. Neither of these has plateaued, which matches other data such as the mean cluster size (figure 3.22), which shows slow increases with no sign of plateau. These simulations will take the longest to finish as the concentration of rods is lowest in this region, meaning the chance of interaction is much lower. The low values of mean cluster size correlate well with high levels of primary nucleation events but very little elongation of fibrils. This is the result of many small clusters nucleating. However, they have not grown significantly throughout the simulations because the concentrations are low, meaning the chance for any small clusters to bind is decreased compared to a high concentration system. In the region with low temperature and low volume fraction, there is an initial dimer peak with transitions into trimers and tetramers, which mostly plateau. In figure 3.25 the multimers in this region are steadily

increasing, showing that this is the region where there is slow but continuing cluster growth as reflected in the mean cluster size in figure 3.22.

3.3.3.2 Low temperature, high volume fraction

Increasing the volume fraction greatly increases the likelihood that a rod will interact with other rods. There is very little or no monomer available from very early time points in the simulation because the simulation box is more crowded, so clusters form almost immediately. The large mean cluster sizes also indicate this that lots of fibril growth occurs after initial nucleation events. The mean cluster size (Figure 3.23) at the endpoints of simulations are high, meaning there are lots of large clusters. In this region, much larger structures will form in crowded environments, so multimers are simply intermediate structures. There is varying behaviour of the larger multimers (Figure 3.25) dependent upon volume fraction. The highest volume fractions see similar transitions into likely larger and larger multimers or fibril species (Figure 3.26), whereas at lower volume fractions, populations continue to grow and then plateau. The higher volume fractions increase the probability that multiple small clusters (Figure 3.24) or monomers (Figure 3.23) will bind and cause fibril growth.

3.3.3.3 High temperature, low volume fraction

In this region of the parameter space, there are few or no clusters (Figure 3.26) forming due to thermal energy being close to or higher than the interaction energy. There may be some small clusters (Figure 3.24) forming most likely due to infrequent binding events, but these are very unlikely to be stable over long time scales. As shown in figure 3.23, the system is almost entirely monomeric because the thermal energy is too high for stable clusters to form. It is therefore unsurprising that there are no significant multimer populations (Figures 3.24 and 3.25) observed in simulations in this region.

3.3.3.4 High temperature, high volume fraction

Similarly to the previous case, few or no clusters are forming due to thermal energy being of the order or larger than the interaction energy. Low levels of small clusters forming, indicated by the mean cluster size in figure 3.22, most likely due to crowding events in the highest concentrations

keeping small clusters bound for longer than in the lower density case. Some of these structures may be stable, but it is unlikely that there will be a stable fibril-like conformation under these conditions if the simulation is left to run indefinitely. The populations of multimers (figures 3.24 and 3.25) are decreasing with time showing a faster growth into fibril-like structures. The simulations do not start with large values. However, most of this small multimer growth will occur before the first data collection time-point, which gives the appearance of initially large values. For the high temperature, high volume fraction case, the populations of multimers (figures 3.24 and 3.25) are likely forming due to crowding events that keep small clusters bound for long periods, or frequent momentary binding events of small clusters that immediately disassemble due to the high thermal energies. This is particularly the case for small multimers (Figure 3.24), shown by the high dimer populations at the largest volume fractions.

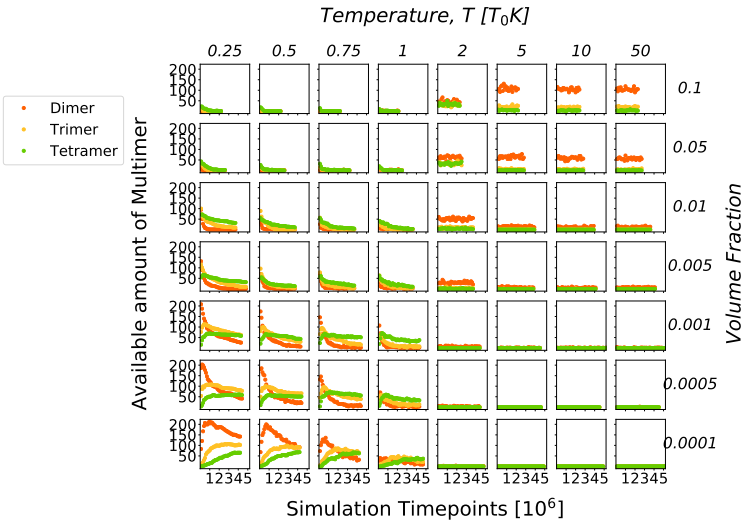

Figure 3.24: Populations of dimers, trimers and tetramers against time for temperatures in the range $0.25T_0 \leq T \leq 50T_0$ ($77.5K < T < 15500K$) along the x-axis (top) and fraction of volume that rods occupy in the range $0.01\% < VF < 10\%$ along the y axis (right). Each simulation is made up of $N = 1000$ rods with a patch length $L_p = 0.6L$ at aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$ and runs for approximately $4 - 5$ million Monte Carlo time steps.

For the case of the volume fraction and temperature sweep for rods only, the lowest density or volume fraction of 0.0001 equates to the upper limit of what is typically carried out *in-vitro* experimentally. In our system and many other simulation schemes, it is typical to use artificially high concentrations or volume fractions to speed up the simulation process. This is shown here in the

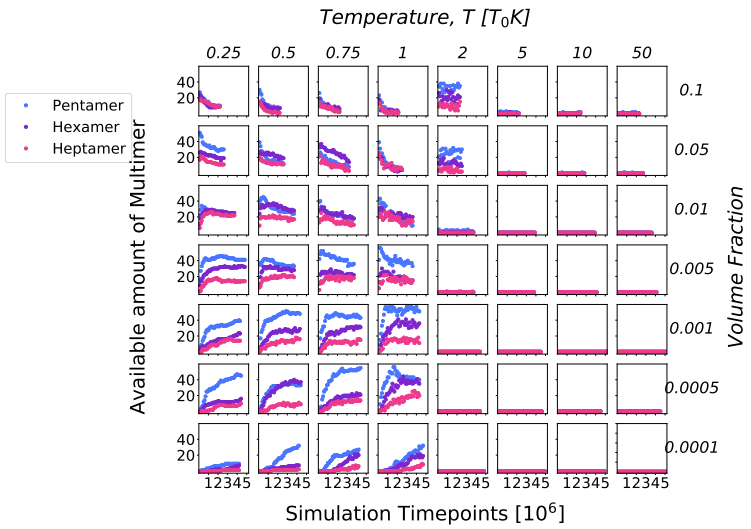

Figure 3.25: Populations of pentamers, hexamers and heptamers against time for temperatures in the range $0.25T_0 \leq T \leq 50T_0$ ($77.5K < T < 15500K$) along the x-axis (top) and fraction of volume that rods occupy in the range $0.01\% < VF < 10\%$ along the y axis (right). Each simulation is made up of $N = 1000$ rods with a patch length $L_p = 0.6L$at aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$ and runs for approximately $4 - 5$ million Monte Carlo time steps.

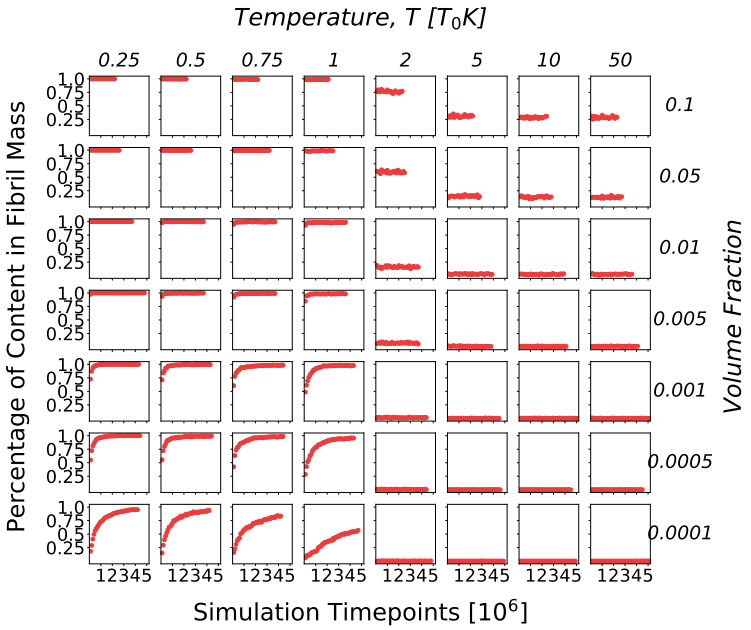

Figure 3.26: Population of fibril content as a fraction of total particle number against time for temperatures in the range $0.25T_0 \leq T \leq 50T_0$ ($77.5K < T < 15500K$) along the x-axis (top) and the fraction of volume that rods occupy in the range $0.01\% < VF < 10\%$ along the y axis (right). Each simulation is made up of $N = 1000$ rods with a patch length $L_p = 0.6L$ at aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$ and runs for approximately $4 - 5$ million Monte Carlo time steps.

sweep by the fact that the same trend is occurring for high volume fractions, albeit much faster than the case for the lower volume fractions, which are close to what is happening *in-vitro*. However, it would take too long to simulate a sensible number of particles at this kind of concentration until it reaches a steady state. Therefore, we artificially increased the concentration to observe the same behaviour but sped up.

The highest rod concentration or volume fraction is 10%. We can see that crowding effects are starting to take place as the data is noisier. However, at volume fractions of around 1%, it is clear that the trends are identical in that we still observe a plateau in the mean cluster size and that if the simulations were run for long enough, eventually, the lowest volume fractions would plateau in a similar way. The mean cluster size (Figure 3.22) plateau value increases with volume fraction until it reaches 10% where the mean cluster size is determined by crowding effects and the increased chance for large clusters to bind together, which experimentally would be considered a rare event.

We conclude that the aggregation behaviour remains constant over a broad range of volume fractions, which validates the artificially high volume fraction of 5% chosen for future parameter sweeps in order to save computational expense and computing time. Using a broad temperature range allows us to ensure that the system operates physically for a broad range of parameters. However, now we can observe the aggregation behaviours in a physiologically appropriate range.

3.3.3.5 Temperatures closer to Physical Parameters

Now that we have fully explored the physics of the system, we repeat the parameter sweep for temperatures closer to physical temperatures. We choose a range of temperatures $0.8T_0K \leq T \leq 2T_0K$. The variance of behaviours is reduced significantly compared with the previous sweep of temperatures and volume fractions. We present our data in two figures 3.27 and 3.28. Figure 3.27 shows a series of heatmaps which describe the steady state behaviours of clusters, and figure 3.28 shows a series of length distributions at sampled coordinates in the parameter space which indicate the number of monomers in a cluster of size N. This figure also highlights typical structures produced under different conditions.

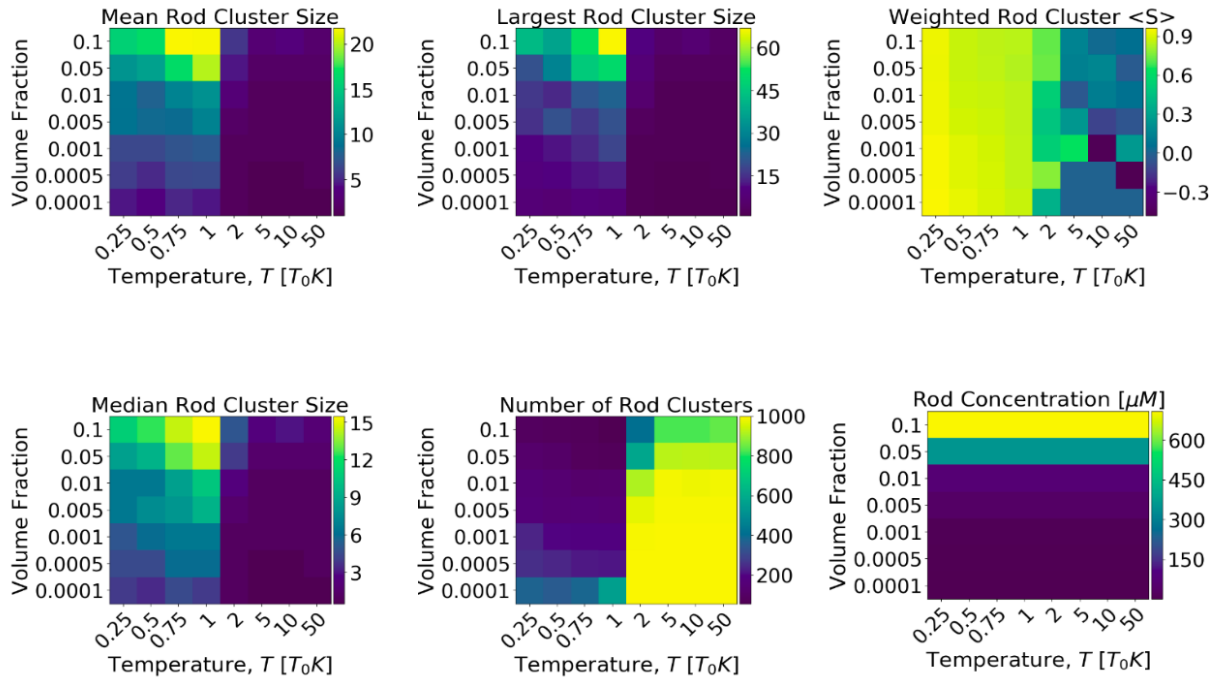

Figure 3.27: Examination of the impact of volume fraction on the steady-state properties of rod clusters over a focussed temperature range. Heatmaps for (top row, left to right) mean rod cluster size, largest rod cluster size and orientational rod order parameter weighted by cluster size, (bottom row, left to right) median rod cluster size, and the number of rod clusters. For temperatures in the range $0.8T_0 < T < 2T_0$ along the x-axis (right hand side) where $T_0 = 310K$ and spherocylinder particle volume fraction in the range $0.0001 < VF < 0.1$ along the y axis (top). Each simulation is made up of $N = 1000$ rods with a patch length $L_p = 0.6L$ at aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$ and runs for approximately 12 million Monte Carlo time steps. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{rod}$.

In comparison to the previous sweep, there are only two extremes with these data rather than four distinct regions of behaviour. Little to no cluster formation when temperature is high ($T \geq 1.4T_0K$) and the volume fraction is very low at $VF = 0.0001 = 0.01\%$ as indicated by figure 3.27. The number of clusters is almost the same as the total number of particles (1000), and the mean cluster size is 1, meaning the steady states of the simulations at and near these conditions are entirely monomeric. This is also indicated in figure 3.28, which shows the steady state size distributions for the same parameter sweep. Most simulations have a very narrow spread of population sizes, with only a very small skew to the large cluster sizes. Almost all of the simulations in this region fail to reach a cluster size larger than 30, which is more clearly indicated by the largest cluster size heat map in figure 3.27. The level of skew and, therefore, the largest cluster size in each distribution increases

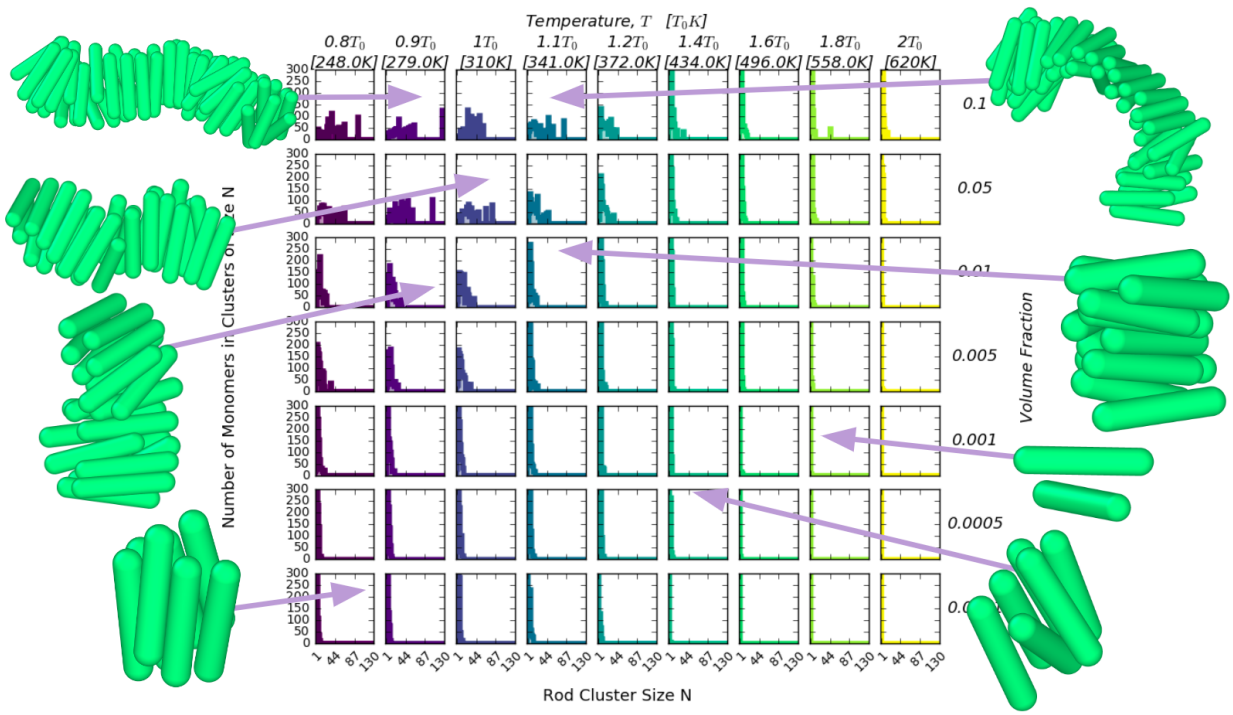

Figure 3.28: Series of length distributions where each bin equates to a discrete cluster size, N, which is filled with the total number of monomers found in a cluster of size N. The colours of each distribution are according to the temperature axis. For temperatures in the range $0.8T_0 < T < 2T_0$ along the x-axis (right hand side) where $T_0 = 310K$ and spherocylinder particle volume fraction in the range $0.0001 < VF < 0.1$ along the y axis (top). Each simulation is made up of $N = 1000$ rods with a patch length $L_p = 0.6L$ at aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$ and runs for approximately 12 million Monte Carlo time steps. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{rod}$.

slowly with volume fraction until we reach $VF = 0.05 = 5\%$, where the distribution changes. This other extreme behaviour is where the volume fraction is high, and the temperature is lower in the range $0.8T_0 <= T <= T_0$. In this region, significant cluster formation is shown by the large mean cluster sizes and the low number of clusters. We find that the aggregation behaviour scales well with volume fraction. We observe larger aggregates on average for the larger mean cluster size, which scales linearly with volume fraction. The median cluster size is highest for $T = T_0$ when the volume fraction is 5%. (Figure 3.27) Under these conditions, we observe many large clusters wherein the time dependence of the simulations presented here, of the mean cluster size changes with time, which occur without a significant level of noise in the data (see figure 3.22 and compare $VF = 5\%$ and $VF = 10\%$). Figure 3.28 in this region shows a more uniform distribution of the monomer into clusters of many different sizes.

These trends are also reflected in the order parameter S. In regions with little to no cluster formation, the order parameter is close to 0, and the data has a high variance due to thermal noise. The number of clusters in this region is close to the total number of particles, and therefore, there will be very few clusters which will contribute to the order parameter, which in itself will create noisy data. The order parameter S is above 0.7 for almost all regions with cluster formation, which means that those clusters are highly ordered even in regions where the mean cluster size is smaller. The order parameter has the lowest variance in the extreme region with low to mid T and high volume fraction, meaning that this region of parameter space consistently produces highly ordered fibrillar aggregates. This is also reflected in figure 3.28 for these conditions where we observe many large aggregates.

The concentrations translate to the range $70\mu M$ to $700\mu M$. This range is extreme compared to the range of concentrations used *in-vitro*, and only the lowest two volume fractions correspond to a concentration that might be used experimentally. The simplest justification for this has already been mentioned, where we use artificially high concentrations of amyloid protein to reduce the number of simulation hours to achieve the same steady-state solution. Below a critical concentration for a given amyloid protein, there is little to no aggregation. It is believed that $in-vivo$ the overall concentrations are extremely low and are on the pM scale[339], which would result in an extremely low nucleation rate in the absence of a catalyst. However, the local concentrations of amyloid beta protein in a diseased brain are high enough to aggregate.[340, 341] We propose that this model encompasses these higher local concentrations or "catalysed" systems because of the high concentrations used in these simulations and the number of particles used.

3.3.4 Translational and Rotational Stepsize

Changing the step sizes will change how far a particle can move in a single Monte Carlo move. In this simulation scheme, these constants act as maxima and minima for a uniform distribution centred around 0. This allows particles to translate backwards and forwards along x, y and z and rotate clockwise and anticlockwise. The values for the step sizes in this work are taken from other literature[292] from atomistic simulations of a fibrillar conformation of Aβ. However, it is still a useful benchmarking exercise to observe fibril growth for a variety of different step sizes (translational vs rotational in a 2D parameter space), even if this parameter sweep is to confirm existing values rather than to determine parameters for future simulations.

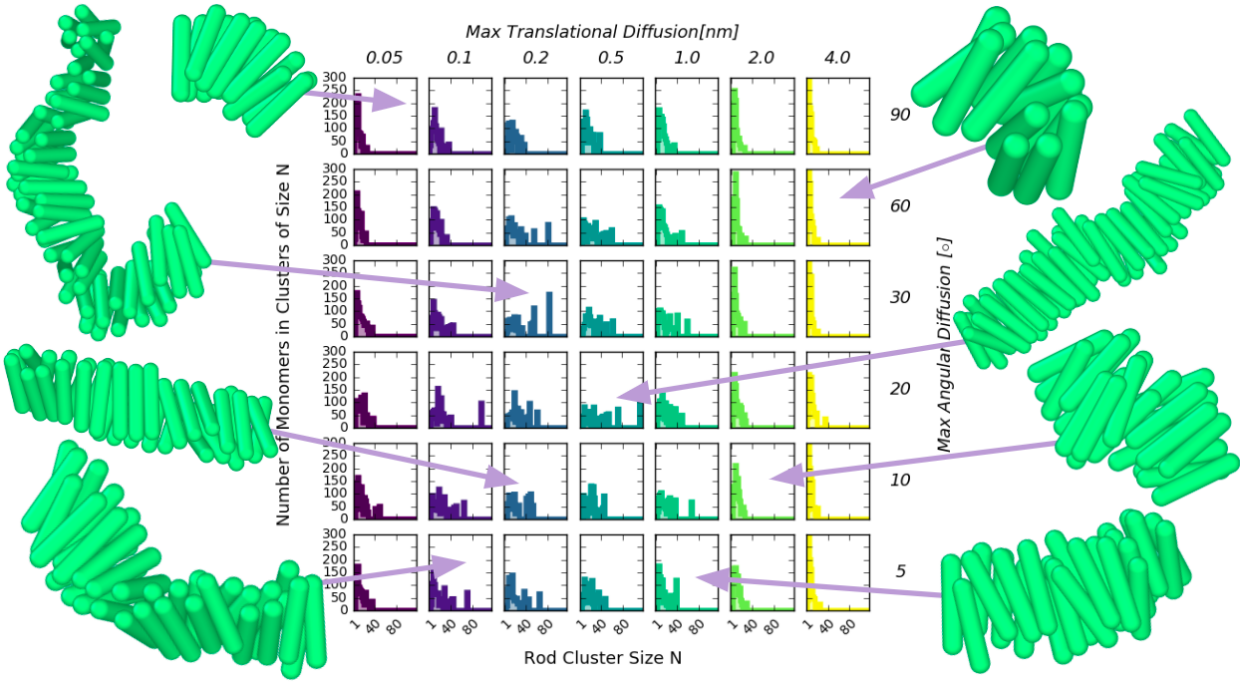

Figure 3.29: Series of steady state length distributions for assemblies populated when varying the maximise translational and rotational diffusion values. Each bin equates to a discrete cluster size, N, which is filled with the total number of monomers found in a cluster of size N. The colours of each distribution change according to the temperature axis. For increasing maximum translational diffusion values from 0.05nm to 4.0nm along the x axis (top) and increasing rotational diffusion maxima from 5 degrees to 90 i.e. on the y axis (right). Simulations are carried out at volume fraction $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1T_0 = 310K$ and the interaction strength $E_{RR} = 6k_BT_0$ where $T_0 = 310K$. Each simulation comprises $N = 1000$ rods with a patch length $L_p = 0.6L$ at aspect ratio $AR = 4$ and runs for approximately 12 million Monte Carlo time steps. Clusters are evaluated using the cluster analysis function from the Ovito python package [334] where the cutoff radius $r_c = 1.5D_{rod}$.

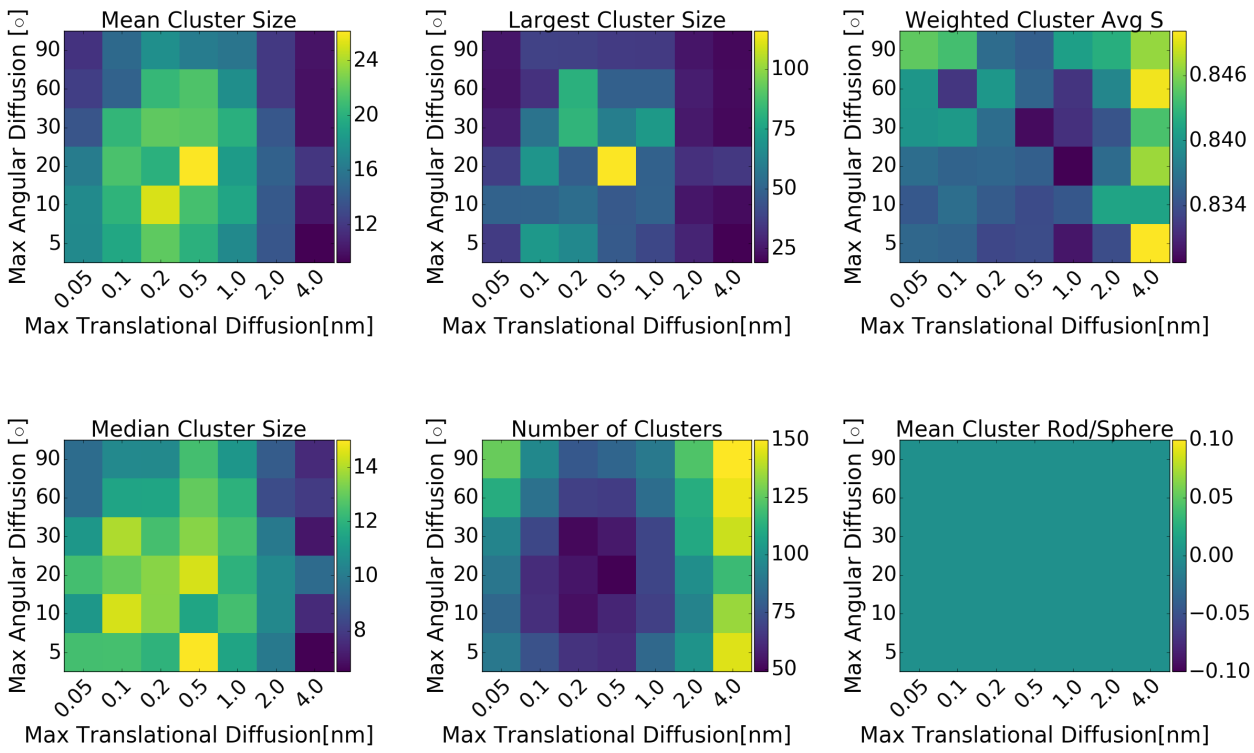

Figure 3.30: Steady state properties of clusters for a 2D parameter sweep of translational and rotational diffusion parameters. Heatmaps for (top row, left to right) mean rod cluster size, largest rod cluster size and orientational rod order parameter weighted by cluster size, (bottom row, left to right) median rod cluster size, and the number of rod clusters. For increasing maximum translational diffusion values from 0.05nm to 4.0nm along the x axes and increasing rotational diffusion maxima from 5 i.e. to 90 i.e. along the y axes. The rod aspect ratio AR is fixed at $AR = 4$ where the rod radius is held constant at $\sigma_{ROD} = 1$nm Simulations are carried out at volume fraction $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1k_BT = 310K$ and the interaction strength $E_{RR} = 6k_BT_0$ where $T_0 = 310K$.

We observe a "halo" like pattern in the phase space for average cluster sizes in figure 3.30 with a large range of different cluster sizes, which suggests a successful coarse sampling of the parameter space which encompasses different aggregation behaviours. There are regions where we observe smaller aggregates; these include all extremes of this parameter sweep with an apparent sweet spot in the central region which is also reflected in figure 3.29 by the presence of a wider variety of cluster sizes in the central region of the phase space. The decreased cluster size regions result from different extremes of the same behaviour. For low diffusion maxima, the possible range of generated movements is much smaller; therefore, the resulting energy change from a very small translation or rotation is smaller and more likely to be accepted by the Monte Carlo algorithm. This means that dissociation events are more likely, which will keep the mean cluster size smaller. Conversely, the

range of possible translation and rotation values is much larger in the large diffusion maxima regions. This results in the more frequent generation of moves that create large unfavourable changes in energy that, in turn, are less likely to be accepted. Aggregates will then move less despite the monomer translational step size being higher. The effects of changing the rotational step size do not appear as prominent as changing the translational step size. This is likely because rotating a rod in contact with another rod results in a less prominent energy change than physically moving a rod away from another rod. Another reason might be that rotating a rod might be more likely to result in a steric clash, and the moves are therefore not accepted.

Figures 3.31 and 3.30 summarise the key properties of clusters observed at the endpoint of the diffusion parameters sweep. The mean cluster size shows a peak when the translational diffusion maximum is $D^T_{MAX} = 0.2 - 0.5$nm for angular diffusion maxima $D^R_{MAX} = 10 - 20°$. The number of rod clusters troughs in a similar region of the parameter space, which implies that these simulations are producing lots of large clusters under these conditions. This is also reflected in figure 3.29 by the wide distributions with many larger clusters present. These values are close but not exact to the conditions we have taken from the literature[292]. This is not unexpected because the step sizes for amyloid-beta will not be the same as for all amyloid proteins due to different properties such as mass and hydrodynamic radii. The "ideal case" values here would work towards designing the perfect amyloid fibril forming protein, whereas, in reality, the human body does not want to form amyloids. There are functional amyloid proteins, but we focus on disease-causing amyloids, which involve the production of amyloid fibrils on the pathway to neural degeneration. The translational diffusion value $D^T_{MAX} = 0.212$nm falls within the range of successful fibril formation, but the angular diffusion value $D^R_{MAX} = 7.5°$ falls just outside of this region of parameter space. However, the parameter selection for this sweep is coarse, so the literature value may fall within this range - a more fine-grained parameter sweep would need to be completed to confirm this.

There is a clear peak in the largest cluster size in figure 3.30 where $D^T_{MAX} = 0.5nm$ and $D^R_{MAX} = 20°$. This agrees with the previous statements about a preferred region of phase space for fibril formation. However, an outlier of this size is likely due to late time large cluster binding in addition to being in the region of interest because the cluster size is close to double the largest cluster in

simulations with neighbouring parameters. Therefore, in this case, the mean is a more accurate measure of the aggregates in a given simulation.

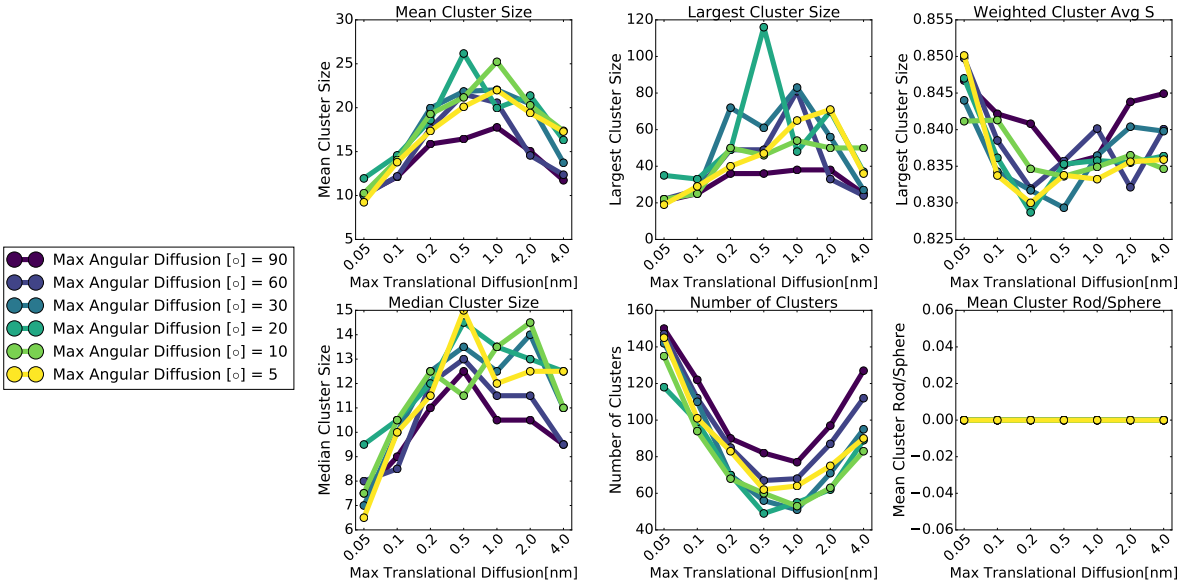

Figure 3.31: Steady state properties of rod clusters over a range of translational and rotational step sizes. Graphs for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size, (bottom row, left to right) median cluster size, number of clusters and ratio of mean rod cluster size and mean sphere cluster size. For increasing maximum translational diffusion values from 0.05nm to 4.0nm along the x axis and lines of increasing rotational diffusion maxima from 5 i.e. (yellow) to 90 i.e. (purple). The rod aspect ratio AR is fixed at $AR = 4$ where the rod radius is held constant at $\sigma_{ROD} = 1$nm Simulations are carried out at volume fraction $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1k_BT = 310K$ and the interaction strength $E_{RR} = 6k_BT_0$ where $T_0 = 310K$.

The orientational order parameter S peaks for the highest translational diffusion maximum for all possible rotational diffusion maxima. Rods with a higher maximum diffusion have a higher range of movements that have a higher probability of making rods more aligned. Similarly, smaller unfavourable movements away from the ideal conformation are more likely to be accepted than a large move, as discussed earlier in this section, so that we will see more local fluctuations for lower translational diffusion maxima. However, the total variance in the order parameter is very small (≈ 0.015) for this parameter sweep, so all relative changes in order are very small and only have minor effects. Interestingly, the number of clusters is significantly higher in the region where S is highest; therefore, despite the clusters being much smaller on average, they are still highly ordered

clusters in this region. Because we are sampling step sizes from a distribution, a change in step size should not affect the mean cluster size. Therefore our observations can be attributed to a delay in reaching steady state for very small or very large step sizes in translation and rotation.

3.4 Discussion and Summary

Following several parameter sweeps, we have found that fibril formation occurs in a narrow range of conditions. We achieve the largest and most numerous fibrils where the rod-rod interaction strength $E_{RR} = 6k_BT_0$, the temperature is $T = 310K$ and the volume fraction is $VF = 5\%$.

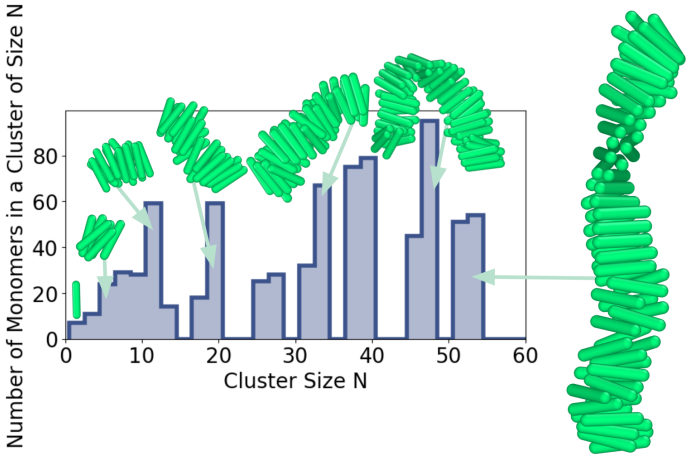

Figure 3.32: A size distribution of clusters at the end point of the simulation. Clusters are evaluated using the cluster analysis function from the Ovito python package[334] with a cutoff radius $r_c = 1.5D_{ROD}$. Example structures from the distribution are shown according to their cluster size. The distribution is weighted against the same distribution to highlight the number of monomers in a cluster of size N. The simulation is carried out at $T = 310K$ and $E_{RR} = 6kT_0$ and is made up of $N = 800$ rods with a patch length $L_p = 0.6L$ at volume fraction $VF = 0.05$ and runs for approximately 7 million Monte Carlo time steps.

Each of these conditions is important in its own right. An aspect ratio of 4 is reminiscent of short amyloid-aggregating polypeptides such as KLVFFA, a known beta-sheet-rich region of amyloid β protein. These simulations encompass beta-strand regions of proteins with lengths between 3 and 25 residues. The latter value, 25, is too long and would not be found to produce amyloid fibres and would correspond to aspect ratio $AR = 16$, which we have found from our simulations produces an

aggregate morphology inconsistent with amyloid formation. Similarly, the region $AR = 2$, which would equate to a length of three residues, is also too short to act as beta strands in fibrils. However, the idea that it may represent the emerging structure in otherwise disordered proteins is intriguing. The range of expected aspect ratios which would equate to real, physical beta strands in fibrils, are found in the range $4 <= AR <= 10$, which agrees well with the simulated aggregation behaviours of aspect ratios in this range. We have found that $AR = 4$ is the most prominent region for fibril formation, which is consistent with a beta-strand length of approximately 6-7 residues. This is a significant result as many beta strands in fibrils are around this size e.g. KLVFFA and GNNQQNM.[67, 68]

We have found in our simulations that the most numerous and largest fibrils are found in the region $0.95T_0 <= T <= T_0$ where $T_0 = 310K$ i.e. $21°C <= T <= 37°C$. This is a very exciting result for such a coarse-grained model as naturally, from a biological perspective, these proteins will be found in a body temperature environment of $37°C$. The largest fibrils were found at $T = 20-29°C$, which is a common range of temperatures for *in-vitro* experiments with amyloid proteins. This was found during an exploration of the rod-rod interaction strength parameter space. When $E_{RR} = 6k_BT_0$, fibril growth is significant near T_0. This does mean that aggregation behaviour is found near T_0 because we have engineered it to do so. However, we believe the result is still exciting because the range of temperatures where aggregation behaviour dominates is specific to experimental procedures and body temperature. The interaction strength is found to be of comparable strength to other similar simulation schemes[3] and is representative of the weak intermolecular forces stabilising protein structures[335].

We have chosen to fix the volume fraction as opposed to the concentration of our simulations. As a result, the equivalent concentrations are very high when compared with *in-vitro* conditions. For example, the conditions of the simulation from figure 3.32 are on the order of $300\mu M$, which would be extremely high for an experiment. In simulation work, concentrations are expected to be higher than in equivalent *in-vitro* experiments to decrease simulation timescales artificially. This artificial increase can have side effects due to crowding events such as large cluster merging, which would typically be considered an extremely rare event. The volume fraction for the simulation in figure

3.32 is fixed at 5%. This is high, and the data is noisier for choosing a high volume fraction such as this. However, we observe rapid cluster formation and reach a steady state sooner due to this. In our model, whilst we are interested in the dynamics and aggregation behaviours on the pathway to the steady state, the steady state is the most biologically accurate aspect of the simulations.

It is possible to consider the model not as a variety of different amyloidogenic monomers with different properties but as a variety of states that a given protein might occupy. For example, $AR = 2$ simulations can be considered a model of a protein with a short β strand region responsible for early binding behaviour within a fibril. This protein could then undergo a conformational change, causing the strands to lengthen and the structure to stabilise, therefore reaching a state comparable to $AR = 4$. In this line of thought, we would only consider the beta-strand region as the spherocylinder, and there would be large worm-like or partially folded regions of a given protein not considered by the model. Similarly, changes in the interaction strength could be considered the result of a conformational change from a weakly binding regime to a more tightly bound conformation with a high interaction strength. Further investigation into behaviour at $AR = 2$ may highlight further regions of interest concerning the behaviour of oligomeric intermediates. Furthermore, it would be interesting to consider an extension to the spherocylinder model where one represents the protein monomer as a pair of spherocylinders bound by a linker, where we could explore simple conformational changes within a given monomer, as well as fibril growth.

Chapter 4

Inclusion of Inhibitory Molecules

Now that we have identified areas of parameter space where our rods assemble into long, fibrillar clusters, we wanted to find the region of space which maximises their disruption by including spherical particles to mimic biological inhibitors. We wanted to find the region(s) of parameter space which minimise the formation of large clusters of rods. We add spheres to our rod simulations to mimic inhibitory molecules. We will now search for the properties of these spheres which minimise rod cluster formation, which will give us insight into the properties required for the design of potential inhibitory molecules. The aim of this work is to guide drug design to minimise the formation of amyloid fibrils in disease. We will explore a multidimensional parameter space to gain insight into the wide variety of aggregate morphologies by varying key shape parameters and interaction strengths. We first discuss the additional analysis methods used to encompass the behaviour of spheres as well as rods. This includes extending size distributions to include multiple cut-off distances, as well as including aggregate averages for mixed assemblies containing both rods and spheres, as well as rod cluster sizes within these larger, mixed aggregates. We then explore the results of several pairs of parameters through a series of 2D parameter sweeps intending to explore as much of the "useful" parameter space as possible. We fix all simulation temperatures at $T = 310K$ and hold the total volume fraction of all particles at 5% to prevent crowding issues.

The most biologically relevant applications can be found at the extremes of this system. The main foci of work on anti-amyloid drug design are in small molecule work[**Siedler2022**, 163, 165]. EGCG, a compound found in tea, redirects amyloid-forming proteins into non-toxic oligomers.

There are many subcategories of such small molecule drugs, often polyphenolic compounds such as curcumin[342] and resveratol.[343] These tend to bind very tightly and will therefore be modelled with high interaction strengths and with sphere sizes equal to or smaller than the radii of the spherocylinders.

The other extreme is the case of chaperones[153, 344], and biological macromolecules (proteins including lipoproteins and glycoproteins), which naturally crowd the environment *in-vivo*, which are modelled with large spheres. These, in reality, will bind more specifically than an isotropic interaction. However, it is still useful to observe such binding mechanisms *in-silico*. Which, for example, cover the surfaces of fibrils to prevent secondary nucleation behaviours. These will also have a range of interaction strengths, but due to the large surface area of these spheres, it will be possible for multiple rods to bind cooperatively to these surfaces. In some examples of simulations in this thesis, the spheres appear to act as a heterogeneous nucleation surface for the spherocylinders.

The primary parameters we choose to sample in this system are sphere radius to rod radius ratio, rod-to-sphere particle number ratios, rod aspect ratio and rod-sphere interaction strength as a multiple of the rod-rod interaction strength. (Figure 4.1) The volume fraction remains the same, but the concentration is subject to change implicitly by explicitly changing the size of each particle - for example, the box size changes by increasing the sphere-to-rod radius ratio or the rod aspect ratio. Depending on the ratio of particle numbers, this will either increase or decrease the relative concentrations of each particle. For this reason, the range of concentrations examined here is somewhat higher than for our rod-only work. Many other parameters are available to change, but we believe these four, and implicitly the fifth, provide the greatest variance in phase behaviour. There are also a huge variety of ways to extend this model further. However, it is wise first to characterise the most simplified version of this system to understand more complex iterations later.

We begin with variations of sphere radius and rod-sphere interaction strength. By establishing relationships between these two parameters in the presence of fixed aspect ratio ($AR = 4$) rods, we can use the results to inform choices for the subsequent parameter sweeps. These simulations

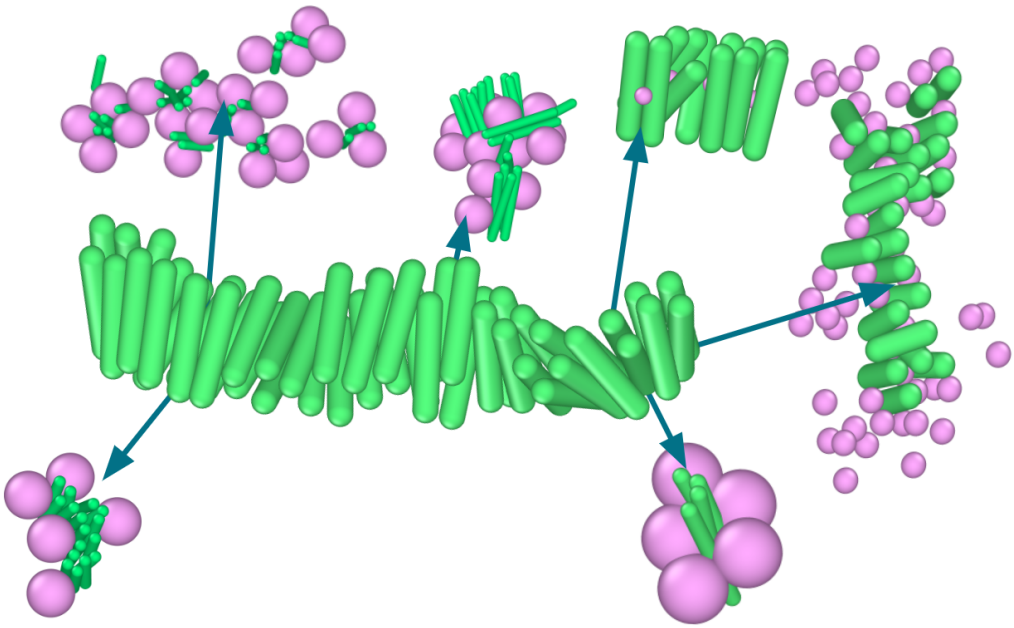

Figure 4.1: A series of example inhibition mechanisms based on the variation of different parameters: sphere radius, rod-sphere interaction strength, relative population ratio and rod aspect ratio.

are carried out at three different population ratios of rods:spheres. These are 1:1, 1:20 and 20:1. They have been selected to sample the extremes of this particular dimension at 5%, 50% and 95% sphere content. We expect different behaviours at different stoichiometric ratios due to crowding events, which occur at rod concentrations which exceed mM, as well as considering that potential amyloid inhibitors will require different dosing to inhibit the fibril formation process effectively.

We then move on to variations of rod aspect ratio and rod-sphere interaction strength. Given that we have already thoroughly explored the effect of aspect ratio for the rod-only system, we can use this as a starting point to understand the impact of including spheres in the system and the impact of changing the binding affinity/interaction strength between the sphere and the attractive patches on the rods. For this, we use two distinct sphere radii which exhibit very different phase behaviour, to explore two "slices" (specific 2D planes in multidimensional space) of the parameter space.

Finally, we will look at the effect of changing sphere radius and the rod aspect ratio. The rod-sphere interaction strength will be fixed at $E_{RS} = 10E_{RR}$ within this section. In a similar fashion to previous sections, we observe three slices of the population ratio parameter space at 5%, 50%

and 95% sphere content. This final sweep is a direct comparison of the shape parameters only so that we can gain insight into the effects of steric clashes, competitive binding and changes in the size of the binding site.

4.1 Analysis Techniques

Whilst the majority of our analysis techniques remain the same as in chapter 3, in order to observe the relative changes before and after the addition of spheres, we must also broaden our analysis to encompass the aggregation behaviour of the inhibitory particles. In addition to the analysis techniques used in the previous chapter (section 3.1), we also utilise the following:

- Size distribution Clusters

  Size distributions are calculated for any particle type, including clusters that are mixtures of rods and spheres. This is calculated at a steady state. Input data is made up of trajectory files for rods as well as spheres. Only one cut-off distance can be used, so it is the larger of either $1.5D_{sphere}$ or $1.5D_{rod}$. Distributions are then placed into histograms and weighted against themselves for the total graph area to be equal to the number of particles ($N = N_{rod} + Nsphere$), i.e. each peak at point X along the x axis equates to a cluster of size X, and the size of the peak Y will equate to the number of particles (rods or spheres) in a cluster of that size. For example, a system with 5 clusters, each containing six particles, will result in the peak $(6, 30)$. We can also further separate the number of rods and spheres in a given cluster.

- Mean Rod Cluster and Sphere Cluster Ratio

$$\bar{\mu}_{RS} = \frac{\bar{X}}{\bar{Y}} = \frac{\frac{1}{N}\sum_{i=0}^{N} L_i}{\frac{1}{M}\sum_{i=0}^{M} L_i} \tag{4.1}$$

  The mean μ_{RS} acts as a ratio of the mean number of rods $\bar{X}$ in a cluster across the size distribution, and the mean number of spheres $\bar{Y}$ in a cluster across the size distribution. These quantities are calculated from the size distribution described above.

4.2 Limitations of the Model

It is impossible to simulate spheres with radii $\sigma_S \leq \sigma_R/2$ in this simulation system. The rod-rod interaction cutoff distance is dependent on the diameter of the rod, D_{rod} where the cutoff distance is $r_c = 1.5D_{rod}$. A rod could interact with the sphere and a second particle "through" the sphere for spheres in the given range. This is problematic as the interaction is not screened through the sphere in the simulation and will be evaluated as if it is not present. Therefore this leads to unphysical behaviour that would not be observed in the biological systems of interest. A second reason not to pursue this region of the overall phase space is that the radii of spheres in this range will potentially be on the order of $10^{-10}m$. Small molecules that can be approximated as spherical do not exist on this length scale. Even the tiniest molecules would not be approximated in this way and would require an atomistic or quantum approach, which is not within this project's scope. A diagram highlighting the issue of particles interacting with one another is shown in figure 4.2. Bii shows the correct series of interactions between particles, i.e. an interaction between the sphere and each rod, but no interaction between the nearby rods through the sphere. (Figure 4.2Bi)

The simulation set-up for rods and spheres is such that the particles are all inserted randomly. However, the spheres are inserted first if their radius is larger than the rod radius, or vice versa. This is completed in such a way as to maximise the chance of finding orientations for all particles with no overlaps, i.e. it is easier for smaller particles to find free space in the gaps between larger particles than it is for larger particles to find space amongst randomly distributed small particles. However, some issues with inserting large spheres into the simulation scheme exist. For spheres greater than or equal to $\sigma_s => 16\sigma_R$, the simulation set-up algorithms struggle to find a suitable configuration with no overlaps. For these reasons we confine our variation of the sphere radius to the region $\sigma_R/2 \leq \sigma_s \leq 16\sigma_R$.

4.3 Interaction Strength vs Sphere Radius

In this section, we observe several 2D parameter sweeps of rod-sphere interaction strength E_{RS} and sphere radius σ_S. These parameter sweeps will be undertaken in 3 special cases of rod:sphere populations to vary the relative populations of protein and inhibitor. We choose to keep the aspect

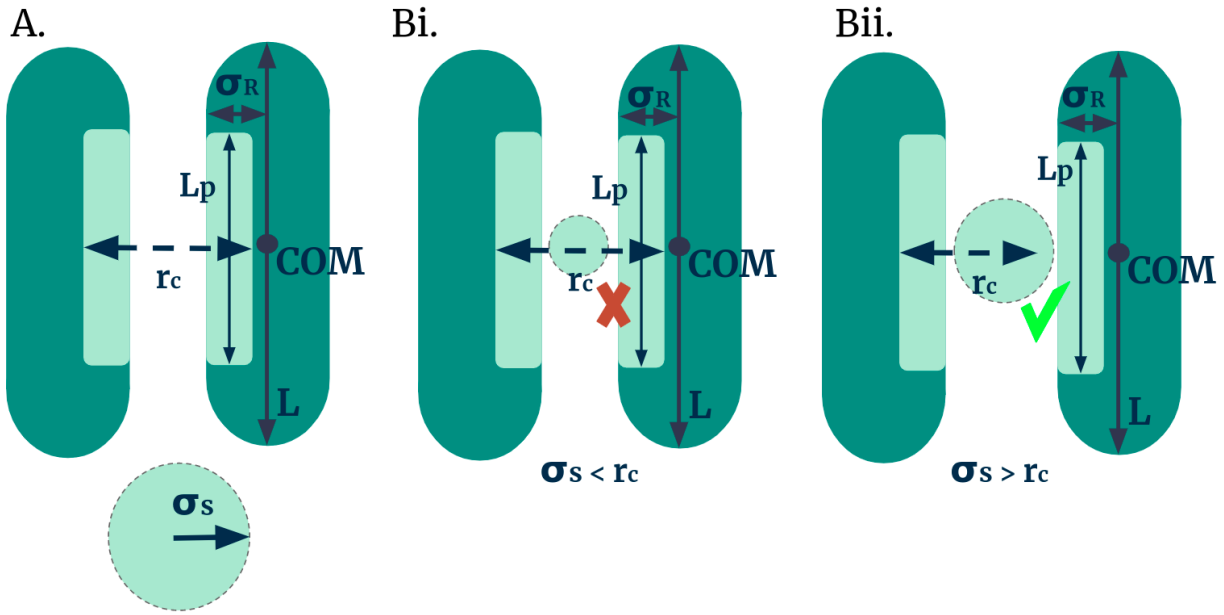

Figure 4.2: Limitations of rod-sphere particle interactions. A. An example pair of spherocylinders interacting with cut off distance $r_c = 1.5D_R = 3\sigma_R$ where σ_R is the spherocylinder radius, and a sphere of radius σ_S. B. Examples of rod-rod interactions (i) incorrectly occurring *through* a sphere with radius $\sigma_S < r_c$ and (ii) correctly not occurring through a sphere of radius $\sigma_S > r_c$.

ratio $AR = 4$ constant in this first set of parameter sweeps to better understand the effects of changing properties of the spheres. $AR = 4$ was shown to stimulate the formation of amyloid-like fibrillar clusters in chapter 3. (Section 3.3.2) Changing the sphere radius σ_S gives rise to some of the most extreme changes in phase behaviour throughout the observed regions of the parameter space. We fix the rod radius σ_R at $\sigma_R = 1$nm, and all sphere radii are multiples of this rod radius. We also fix the rod-rod interaction strength E_{RR} at $E_{RR} = 6k_BT$. We found in section 3.3.1 that $E_{RR} = 6k_BT$ results in the formation of fibril clusters at physiological temperatures.

20:1 Rod:Sphere Population Ratio

Here, we sample incrementally across a 2D parameter space of rod-sphere interaction strength and sphere radius. We hold the total volume fraction of particles at 5% and set the population ratio of rods:spheres at 20:1. This means that there are 950 rods and just 50 spheres. Figure 4.3 is a 2D series of size distributions across incrementally sampled rod-sphere interaction strengths in the range $0.1E_{RR} \leq E_{RS} \leq 128E_{RR}$. The distributions contain aggregates of all types, mixed or single

species, as defined in section 4.1. We also sample incrementally across sphere radii in the range $0.75\sigma_R \leq \sigma_S \leq 16\sigma_R$. Each of these is a multiple of the equivalent rod-only property, i.e. rod-rod interaction strength E_{RR} and rod radius σ_R. We set up parameter sweeps in this way for ease of comparison to the rod-only properties and because it is easier to conclude with experimental work by observing the relative changes in each system rather than using the absolute values. We also contribute to the summary data in figure 4.4 which highlights a series of steady-state observations of cluster averages such as rod cluster size averages - which includes the size of rod clusters within larger, mixed clusters - concentrations of rod and sphere particles, and order in rod clusters S.

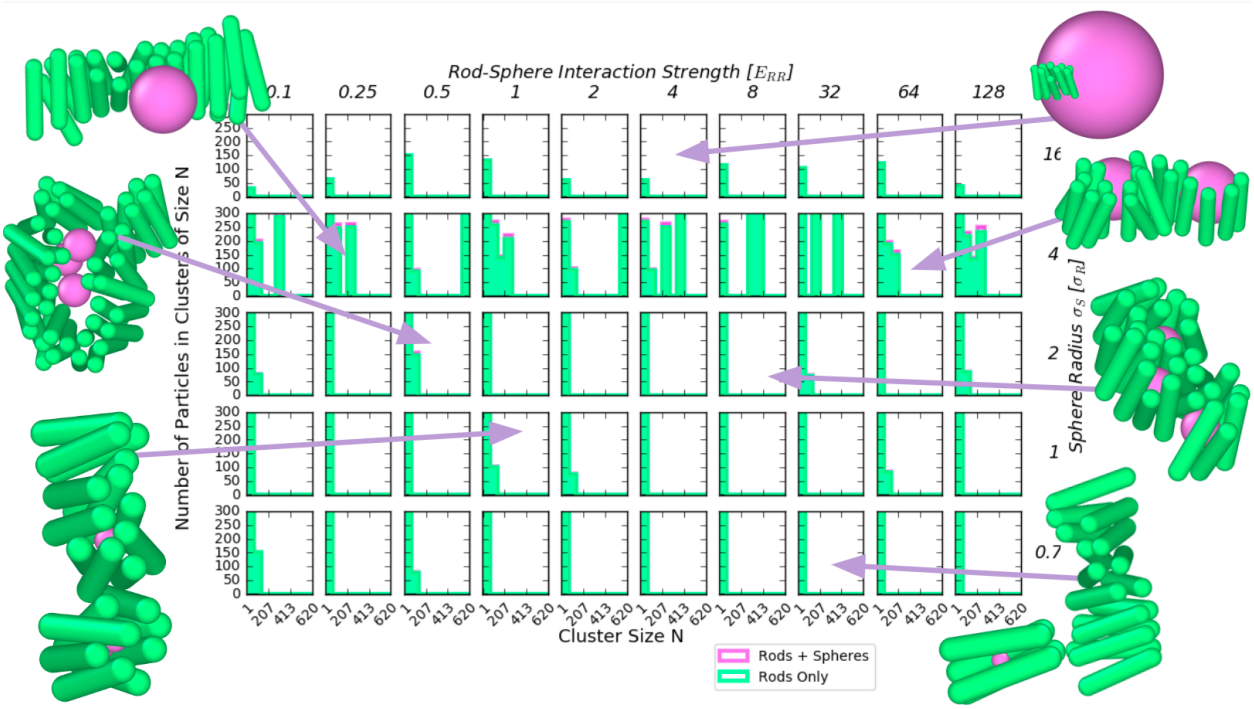

Figure 4.3: A series of steady-state size distributions across a 2D parameter sweep of rod-sphere interaction strength E_{RS} and sphere radius σ_S. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. Rod sphere interaction strength varies along the x axes between $E_{RS} = 0.1E_{RR}$ and $E_{RS} = 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. The ratio of rods to spheres is $20:1$ in a simulation box of 1000 particles. Rod aspect ratio is fixed at $AR = 4$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1k_BT = 310K$

In the case where the rods or sphere ratio is 20 to 1, in that there are 950 rods to 50 spheres, the concentration of rods is significantly higher than the concentration of spheres. This is reflected on

the right hand side, figure 4.4, where the concentration reaches $2800\mu M$. Due to this, there can be a crowding effect from the rods - in section 3.3.3 the rod concentrations for the currently used volume fraction (5%) are much lower than those observed across this parameter space for sphere radii $\sigma_S \leq 4\sigma_R$. Figure 4.4 highlights steady state observations on cluster averages for rods and spheres, including the rods and spheres' concentrations which remain constant throughout the simulations. These higher concentrations observed due to the increased rod quantity and low sphere radii result in crowding effects throughout the simulation.

Inspection of figure 4.3 suggests immediately that a ratio of sphere to rod radius size of 4 ($\sigma_S/\sigma_R = 4$) leads to the largest range of cluster sizes. This equates to a sphere size analogous to the rod length. All simulations reveal large changes in the morphology of the assemblies, as shown by the example images. Regular fibrillar assemblies are disrupted, and spheres intercalate between the rods and generate more amorphous structures. Smaller spheres can sandwich themselves between sheets of rods, whereas larger spheres break these up. Figure 4.4 explores the impact of the spheres on the rod clusters observed within the now heterogeneous assemblies. There is a distinct change in behaviour when the sphere radius is 16 times larger than the rod radius. This behaviour change is observed across several data sets in figure 4.4, including the median rod cluster size, the mean rod cluster size and the largest rod cluster size. The values for $\sigma_S = 16\sigma_R$ are much lower than at any other point in this parameter space. Conversely, the number of clusters in this region is also significantly higher than in other regions of the same phase space. There is a slight uptick in the weighted cluster average of the orientational order parameter S. However, this is not significantly larger than in other regions of the same parameter space.

The ratio of mean cluster size for rods and spheres μ_{RS} shows two distinct regions of behaviour as we increase the rod sphere interaction strength. (Figure 4.4) When the rod sphere interaction strength is low, we observe a higher mean cluster rod sphere ratio μ_{RS}. To the right, when the rod sphere interaction strength is four times larger or higher than the rod-rod interaction strength ($E_{RS} >= 4E_{RR}$), we observe a lower μ_{RS}. The mean cluster rod sphere ratio μ_{RS} (Figure 4.4) will be higher when the interaction strength is two or lower ($E_{RS} \leq 2E_{RR}$). In this region of the parameter space, figure 4.3highlights example structures with fibrillar structures and three differ-

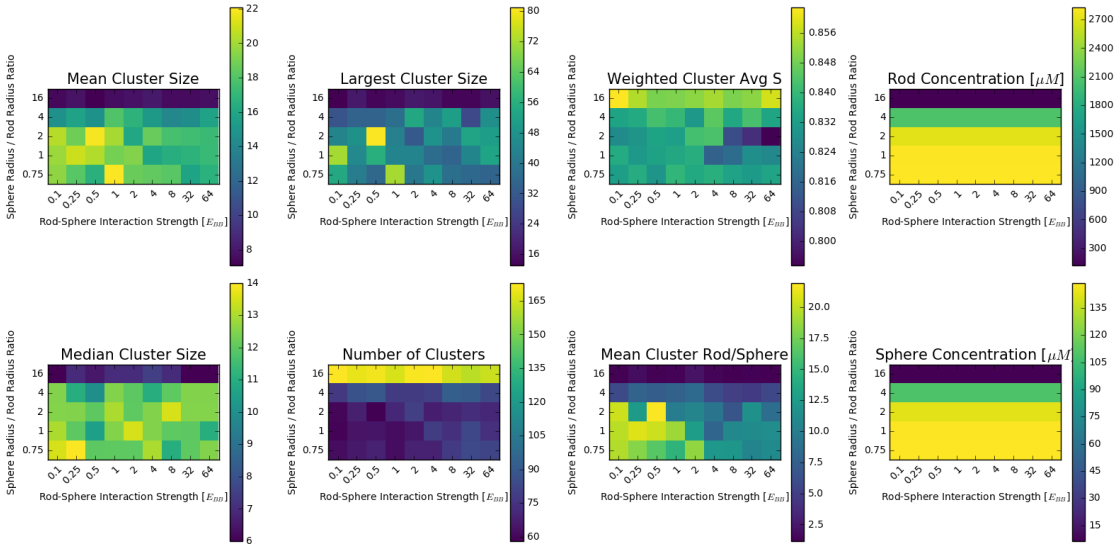

Figure 4.4: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod-sphere interaction strength E_{RS} and sphere radius σ_S, where the ratio of the rod to sphere population is fixed at 20:1. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $0.1E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. The ratio of rods to spheres is $20:1$ in a simulation box of 1000 particles. Rod aspect ratio is fixed at $AR = 4$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1k_BT = 310K$

ent inhibitory events dependent upon sphere radius. The inhibitors can create multiple rod-sphere interactions for small sphere radii by filling the space inside fibrils. As these spheres increase in radius, we find fibrillar structures binding or even wrapping around inhibitors. These behaviours occur because the interactions between rods and spheres are not strong enough that they will cause pairs of rods to dissociate from one another. Spheres cannot form clusters without rods because there is no attractive interaction between spheres. The only way that spheres will form clusters is if they are within the interaction length of other spheres, and this will only be the case when there are mixed aggregates of rods and spheres. The structures observed for large spheres and low interaction strengths have high order parameters comparable to those observed for the rod-only systems (See section 3.3.2, figure 3.16) and appear to increase the ability of a fibril to elongate by

providing an attractive surface for monomers to bind to, and thereby elongating the existing fibril.

This may be because the sphere radius is significantly larger than the rod radius. Therefore, the calculations of sphere clusters will result in low cluster numbers and sizes, whereas the rod clusters may remain the same. The rod mean cluster size, the largest cluster size and the median cluster size for rods show there are low numbers of large rod clusters and high numbers of clusters, and therefore, with the presence of very few or almost no sphere clusters, then this mean cluster size ratio is also going to be very low.

Therefore, conversely, when the interaction strength between rods and spheres is higher (Figure 4.4) - such that the interaction strength is significantly large enough that it is forming mixed aggregates of rods and spheres - the mean cluster rod sphere ratio μ_{RS} decreases proportionately to the size of the interaction strength. Higher interaction strengths result in more fibril inhibition effects (As shown in example structures in figure 4.3) where once again we find spheres filling the central regions of aggregates. When interaction strengths are increased, larger sphere radii can no longer fill these aggregates without breaking the fibrils apart into smaller pieces (An example of this breaking apart behaviour is highlighted in figure 4.3 where $E_{RS} = 64E_{RR}$ and $\sigma_S = 4\sigma_R$).

To conclude, we have found that the mechanism of fibril inhibition is dependent upon the relative size and interaction strength of the inhibitor relative to the properties of the amyloid protein. Weakly interacting inhibitors appear to act as heterogeneous nucleation surfaces which drive fibrillisation rather than inhibit it. Similarly, spheres with an increased interaction strength can cause significant fibril defects (As shown in example structures in figure 4.3). The degree of the defect is dependent upon the size of the sphere relative to the rod. However, this relationship between shapes is to be explored further in section 4.5.

1:1 Rod:Sphere Population Ratio

In this section, we observe a 2D parameter sweep of the interaction strength between rods and spheres as a multiple of the rod-rod interaction strength and sphere radius, where each radius is a multiple of the rod radius. For our analysis, we use two figures. Figure 4.5 is a 2D se-

ries of length distributions across incrementally sampled rod-sphere interaction strengths in the range $0.1E_{RR} \leq E_{RS} \leq 128E_{RR}$, we also sample incrementally across sphere radii in the range $0.75\sigma_R \leq \sigma_S \leq 16\sigma_R$. These length distributions provide an overall picture of the relative aggregate populations across different regions of the parameter space at a steady state. We also use a series of heatmaps (Figure 4.6) to observe rod cluster averages, often found within the larger mixed aggregates, degree of rod cluster order and concentrations of rods and spheres, respectively. These heatmaps are determined at the steady state of each simulation. For this case, where the number of rods is equal to the number of spheres - the rod concentration and sphere concentration at a given point in the parameter space are always of comparable order. (Figure 4.6) It is lowest when the sphere radius is at its largest at 16 times the rod radius ($\sigma_S = 16\sigma_R$) and is at its highest when the rod radius is at its smallest, which is 0.75 times the rod radius ($\sigma_S = 0.75\sigma_R$).

The weighted cluster average orientational order parameter S in figure 4.6 is lower but remains consistently high across most of this phase space. There is a slight decrease where the sphere radius is equal to 16 times the rod radius ($\sigma_S = 16\sigma_R$) and also for the most extreme interaction strengths. These very extreme interaction strengths are where we observe significant inhibition behaviours, which vary significantly depending on the sphere radius. (See right hand side example structures in figure 4.5)

The mean cluster size also has a distinct peak for the lowest interaction strengths and the lowest sphere radii. This is most likely because there are little to no inhibition effects occurring in this region, and therefore, the rods can be considered to be under normal fibril-forming conditions. i.e. $AR = 4$. Given that the sphere radius is so small, it is unlikely that there will be any steric effects on the fibril formation process, and therefore, we would expect to see the expected distribution of rods if there were no inhibitors present in that there will be numerous and elongated fibrils. Differences are also observed with respect to the ratio of mean rod to sphere cluster size μ_{RS}. There is a clear peak in the mean cluster ratio of rods to spheres, μ_{RS} (Figure 4.6) where the rod sphere interaction strength is $E_{RS} = 0.25$ and where the sphere radius is $\sigma_S = 0.75$.

The data presented here are different to the previous case where the stoichiometric ratio of rods to

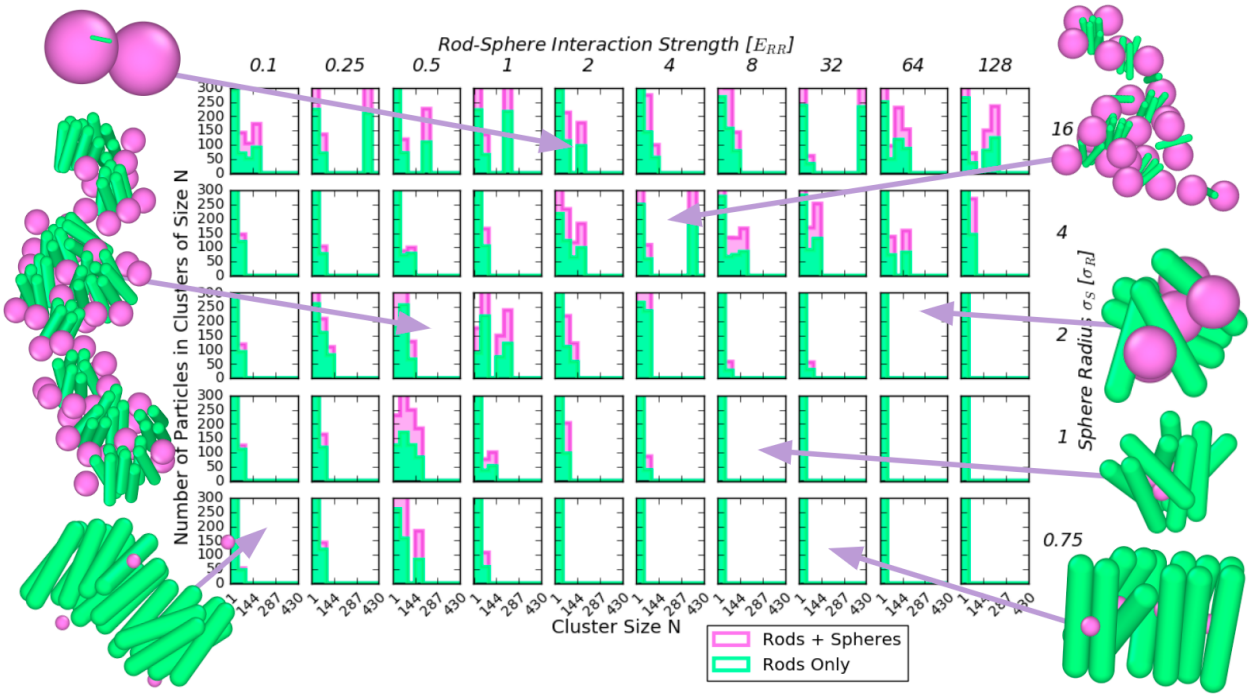

Figure 4.5: A series of steady-state size distributions across a 2D parameter sweep of rod-sphere interaction strength E_{RS} and sphere radius σ_S, when the ratio of the rod to sphere populations 1:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. Rod sphere interaction strength varies along the x axes in the range $0.1E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. The ratio of rods to spheres is $1:1$ in a simulation box of 1000 particles. Rod aspect ratio is fixed at $AR = 4$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1k_BT = 310K$

spheres was 20:1. There is a large central region of this parameter space which results in globular aggregates with varying order parameters. Particularly for smaller sphere radii, the low rod-sphere interaction strengths result in highly ordered local regions of ordered rods amongst significant surface coverage of said aggregates. The converse case results in highly disordered aggregates where the rod-sphere interaction energies dominate over the rod-rod interaction energies. The smallest radii are capable of filling the gaps within fibrillar aggregates, disrupting the order of clusters by pushing rods apart. This data is expected to be noisier than the previous set of data because there are fewer rods. However, there are also more spheres, so any data which includes the calculation of spheres will be less noisy than the previous dataset.

The median cluster size (Bottom left figure 4.6) does not correlate to the same strong peak as the mean rod cluster size (Top left figure 4.6). The overall trend is the same, but there is no distinct

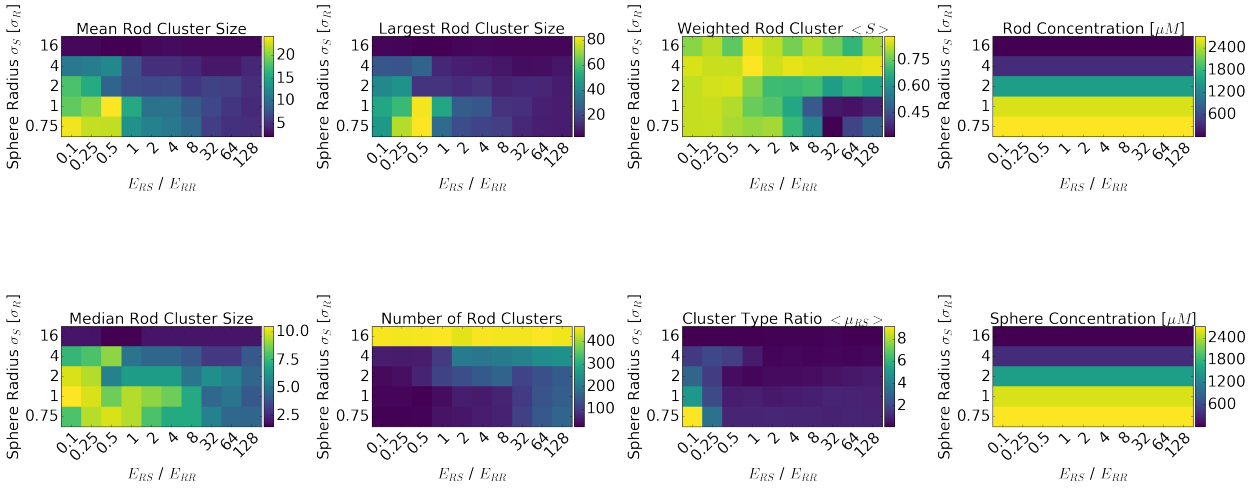

Figure 4.6: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod-sphere interaction strength E_{RS} and sphere radius σ_S. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. Rod sphere interaction strength varies along the x axes between $0.1E_{RR} \leq E_{RS} \leq 64E_{RR}$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. The ratio of rods to spheres is $1:1$ in a simulation box of 1000 particles. Rod aspect ratio is fixed at $AR = 4$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1k_BT = 310K$

peak for the lowest sphere radius and lowest rod sphere interaction strength. However, there is a general trend that with increasing sphere rod radius, there is a distinct decrease in the average cluster size of rods with a sharp drop-off for the sphere radius $\sigma_S = 16\sigma_R$, which is consistent with the previous data set in section 4.3. These decreasing order and cluster size trends with increasing interaction strength indicate inhibitory behaviours. Figure 4.5 highlights example structures on the right hand side (high interaction strengths) which are much smaller than equivalent sample structures for low interaction strengths. Similar to the results for stoichiometric ratio 20:1 rods:spheres, the sphere radius is the parameter which determines the inhibition behaviour, whereas the increasing interaction strength appears to drive that behaviour.

To conclude, the results of this parameter sweep again suggest there is a relationship between sphere size and mechanism. i.e. Can larger spheres disrupt the aggregates from the inside? However, further exploration of the shape parameter space is required. This will be discussed in further

detail in section 4.5.

1:20 Rod:Sphere Population Ratio

In this section we repeat our exploration of the a 2D parameter space of interaction strength in the range $0.1E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength, and sphere radii in the range $0.75 \leq \sigma_S/\sigma_R \leq 16$. This time, we fix the rod-to-sphere ratio populations at 1:20. The volume fraction of all particles combined is fixed at 5%, resulting in different rods and spheres concentrations throughout the parameter space. These concentrations can be observed, along with steady-state observations on cluster averages and rod cluster order in figure 4.8. This is used to provide an overview of steady-state behaviours. Figure 4.7 highlights steady-state size distributions of mixed aggregates of amyloids and inhibitors together with example images of endpoint aggregates. The population ratio in this region is 1:20 rods:spheres, resulting in 50 rods and 950 spheres. For this reason, rod-only cluster observations are likely to be noisier. Sample structures may provide more accurate depictions of the aggregate morphologies observed.

The mean cluster size (Top left, figure 4.8) is consistently very low with one exception where the sphere radius $\sigma_s = 0.75\sigma_R$ and rod sphere interaction strength $E_{RS} = 0.25E_{RR}$. We consider this point an outlier where the mean rod cluster size is the same as the total number of rods in the system. This is an artefact of the cut-off radii used to determine cluster sizes and the high concentration of spheres present in the system, which can result in extremely large clusters of particles.

Generally, the mean rod cluster size in the top left in figure 4.8 is decreasing for increasing interaction strength and sphere radius. i.e. The smallest rod cluster sizes are found at the highest interaction strength and largest sphere radius. This trend is more clearly reflected in the median rod cluster size in the bottom left of figure 4.8 where we observe median rod cluster sizes in the range $0.8 \leq x \leq 3.2$. This range is very small compared with the total number of rods in each simulation (50), but given the high population ratio of rods to spheres (1:20), the fact that most rods are not forming large clusters with one another can be explained because it is much more likely that they will bind with spheres. This is also reflected in figure 4.7 where we see that in the majority of cases, only rods are found in the smallest category of cluster sizes. The number of

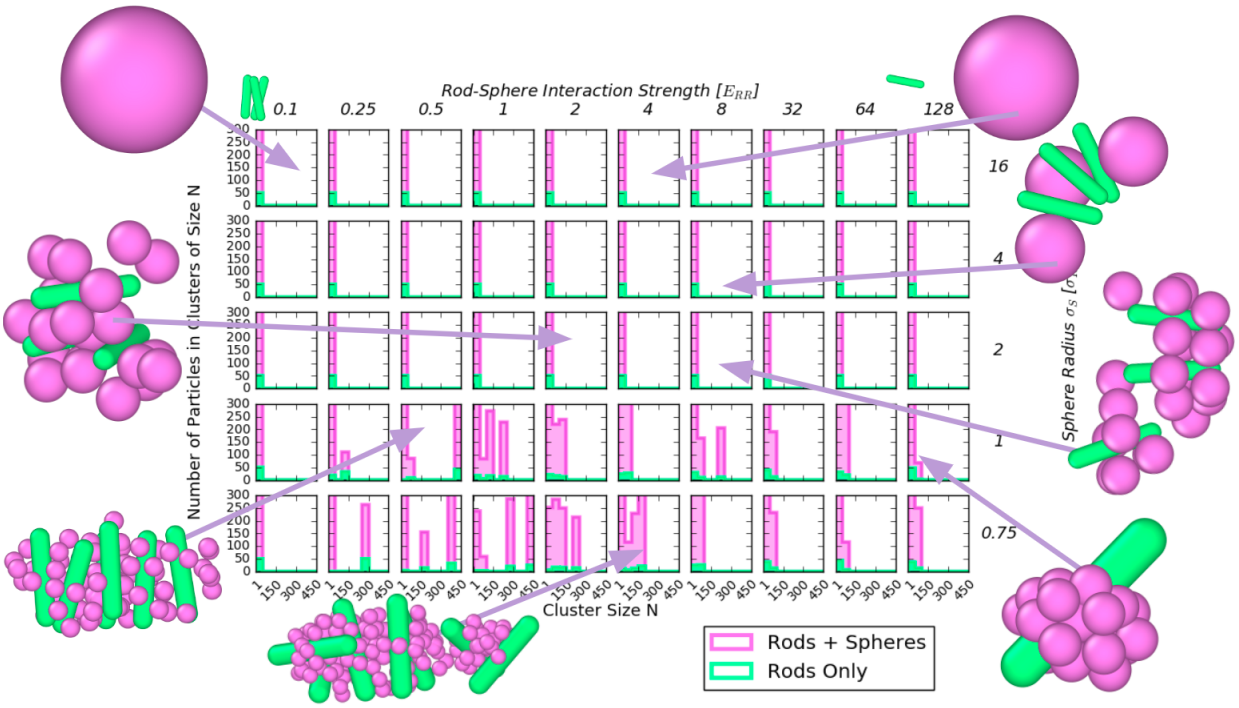

Figure 4.7: A series of steady-state size distributions across a 2D parameter sweep of rod-sphere interaction strength E_{RS} and sphere radius σ_S where the ratio of the rod to sphere populations is 1:20. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. Rod sphere interaction strength varies along the x axes in the range $0.1E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. The ratio of rods to spheres is $1:20$ in a simulation box of 1000 particles. Rod aspect ratio is fixed at $AR = 4$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1k_BT = 310K$

clusters in figure 4.8 is close to the total number of rods in the system for these same conditions (high sphere radius and interaction strength), which shows that the majority of rods are not binding with other rods, and the inhibition behaviour is strong. Most clusters tend to be dimers, trimers or monomers, suggesting that the growth rate has been decreased and likely the primary nucleation rate. Nucleation is discussed in section 1.5.1. In the case of these simulations, it is the formation of a new rod cluster.

The data for the orientational order parameter S (Figure 4.8) are quite noisy. However, the data still shows an increased order parameter for lower interaction strengths, providing evidence to support that increasing the interaction strength is disrupting the level of order in a given cluster and the size. The clusters in the high interaction strength region of the parameter space are much more

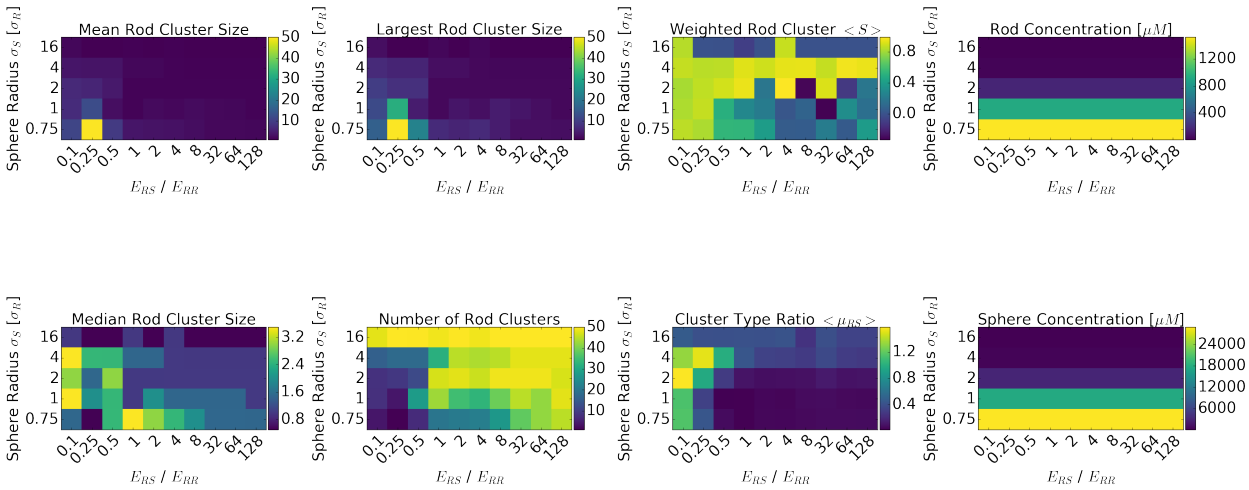

Figure 4.8: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod-sphere interaction strength E_{RS} and sphere radius σ_S where the ratio of the rod to sphere populations is fixed at 1:20. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $0.1E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. The ratio of rods to spheres is $1:20$ in a simulation box of 1000 particles. Rod aspect ratio is fixed at $AR = 4$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. Temperature is fixed at $T = 1k_BT = 310K$

likely to be globular mixed aggregates due to the decreased order parameter. This is reflected in the example structures highlighted in figure 4.7 The sphere radius also impinges on this effect. A smaller sphere which is closer to the rod radius is more likely to affect the cluster order as it will be able to push rods apart whilst still allowing the rod patches to remain close enough to interact with one another.

The morphologies of different aggregates across the phase space are highlighted in figure 4.7. The general trend is that the larger the inhibitor radius, the lower the number of inhibitors that can bind to a given rod. The largest sphere radius $\sigma_S = 16\sigma_R$ simulations result in very low numbers binding events indicated by the consistently narrow distributions at low cluster sizes in figure 4.7. Due to the high stoichiometric ratio of spheres:rods and the large radii of spheres, the spheres appear to inhibit fibril production simply by steric effects. Whilst this result is important, it is not representative of a true treatment option for amyloidosis because it would not be possible to

deliver a drug of this size with this stoichiometry. Even individual amyloid fibrils can contain up to 10s of thousands of amyloid protein monomers. Therefore, treating amyloidoses drugs at this stoichiometric ratio would require 20 times more drug molecules than amyloid monomers, which are likely to be highly disruptive to brain chemistry. To parameterise a potential drug with these properties, an exploration of more stoichiometries close to those currently used in anti-amyloid drugs must be carried out. On the other hand, under non-pathological conditions, the number of endogenous chaperone molecules may greatly exceed the number of amyloid precursors, thus illustrating a potential mechanism of action for chaperone molecules such as BRICHOS.

As we decrease the sphere radius, we see mixed aggregates with much larger numbers of spheres within them. (Figure 4.7) The larger clusters also indicate this in the size distributions for lower sphere radii. Steric effects can explain this; one large sphere is much more able to bind many rods rather than many spheres binding to a singular rod. The reverse is the case for the smallest spheres, where many spheres can bind to a single rod. The sphere can bind isotropically, but the patch size is finite, so only so many spheres can be within the interaction distance at once, dependent upon the sphere radius. For the smallest sphere radius, the aggregates are predominantly spheres which are unsurprising given the stochiometric ratio of 95% spheres. The smallest sphere radius is less than the rod radius, meaning that several spheres can bind to a given rod's patch. For the lower rod-sphere interaction strengths, this leads to some interesting mixed aggregates. Even though it is more favourable for a pair of rods to bind than a rod and a sphere, the spheres bind cooperatively. Indicated by the presence of larger cluster sizes in interaction strengths larger than $E_{RS} \geq 0.25E_{RR}$ for $\sigma_S = 0.75\sigma_R$, the sum of X rod-sphere interaction energies for example, is greater than a single rod-rod interaction. X would need to be ≈ 10 for $E_{RS} = 0.1E_{RR}$, which is not possible for the sphere radii used in these simulations.

Particularly for sphere radii $\sigma_S \leq 2\sigma_R$, we observe a global change in aggregate morphologies as we increase the interaction strength, as illustrated by figures 4.7 and 4.8. As rod-sphere interaction strengths become orders of magnitude larger than the rod-rod interaction strength, we see more tightly bound spheres with smaller numbers of rods. Spheres become tightly bound, which covers all available interaction areas of the attractive patches and leaves no space for other rods, which by

comparison bind very weakly. For the lower interaction strengths, we observe more miscible phases of rods and spheres, which contain multiple rods, still relatively ordered to one another, with the space between them filled with spheres. While the rods are ordered relative to one another, they are often not bound to one another directly and will not always contribute to an order parameter calculation. The reason they are ordered is that this means the many rod-sphere energies will be maximised by spheres binding to 2 rods at once along both of their patches. With multiple spheres doing this, it forces the rods into the same orientations. For the lowest rod-sphere interaction strengths, there is also a high order parameter indicated by figure 4.8. This is less the case as we increase interaction strength, as the spheres bind more tightly, forcing the spherocylinders into seemingly more random orientations to maximise their interaction energies. The disordered structures are most apparent when the sphere radius is close to the rod radius. The spheres fill the space between adjacent rods and also bind very tightly when the interaction strength is higher, creating amorphous globular aggregates.

Whilst a strongly binding inhibitor decreases the rod cluster size the most; it does not necessarily mean it is the most suitable option for all inhibitors from a drug design perspective. In a physiological system, there is a very dense environment of different species, so a quite specific binding is often needed for cell machinery to work to prevent it from binding to the wrong species, for example. Conversely, small molecule drugs tend to be more non-specific and can bind to a variety of species. A simplified version of such small molecule binding events is evidenced here, where many small molecules act cooperatively to have an inhibitory effect even at relatively low interaction energies. Perhaps this indicates that small molecule drugs could bind to many known motifs that are known to form amyloid fibres readily[67, 68], and this would be enough to prevent full fibril formation. This, of course, would need to be tested by including the heterogeneity of spherocylinders. There is a delicate balance at play here, however, because beta-sheet-rich regions of proteins are very common, so any therapeutic would need to be able to bind to the amyloid-forming proteins, but not other functional proteins in the brain. The larger inhibitors aim to model antibody-like inhibitors, which tend to bind more specifically. This parameter sweep is indicative that with high stoichiometric ratios, simply increasing steric clashes appears to have an inhibitory effect. Under this stochiometry, we are not observing a physical system for antibody drug candidates, and so the

more biologically relevant results will be found at a lower ratio.

4.4 Aspect Ratio vs Interaction Strength

Two of the most important quantities in this system are rod aspect ratio and the relative interaction strengths of differently shaped particles. From our results on the rod-only system (Section 3.3.1), we established that the most favourable conditions for filament formation were found when $E_{RR} = 6k_BT$. We maintain this interaction strength for the simulations in this section as constant and vary only the rod sphere interaction strength E_{RS}. By varying the rod-sphere interaction strength relative to a given rod-rod interaction strength, we hope to draw more generalist conclusions regarding the properties of inhibitory molecules rather than specific free energies of structures produced. In doing this, the model is more applicable to a variety of systems, and we are better able to inform experimental and clinical scientists about the relative properties of inhibitors to their target species.

The other important varied quantity, in this case, is the rod aspect ratio AR. In the rod-only system (section 3.3.2), we found that rods with $AR = 4$ produced the most numerous and elongated fibrils in the region of interest within the parameter space. We also hinted that $AR = 2$ was a condition where small oligomers would form. The shorter length of these rods may reflect the less folded nature of oligomeric species. This quantity, aspect ratio will be varied in this section. In other words, we can consider the aspect ratio of a rod to be a function of the number of residues in a beta-strand for a given monomer. Therefore it is still of interest to consider how strands of different lengths (or different numbers of residues) can be prevented from aggregating into fibrillar species.

The number of simulations required to sample this parameter space snowballs. Therefore we have chosen to sample with large steps between adjacent samples of the given parameter space. We have chosen 6 aspect ratios in the range $2 \leq AR \leq 10$, and 6 interaction strengths in the range $1k_BT \leq E_{RS} \leq 128k_BT$. Given that we are also interested in the relative quantities of rods and spheres, we have chosen three regions of interest in this parameter space in the following ratios of

rods:spheres - 1 : 20, 1 : 1 and 20 : 1. These particular regions are extreme relative to each other and should therefore retrieve a large variance in properties of structures formed.

4.4.1 Equal Size Rods and Spheres

In this first section, the rod radius is equal to the sphere radius $\sigma_R = \sigma_S$. It mimics small molecule inhibitors with a lower probability of steric effects due to their relatively small diameters.

1:20 Rod:Sphere Population Ratio

In this 2D parameter sweep of rod-sphere interaction strengths in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$, and rod aspect ratios which are sampled in the range $2 \leq AR \leq 16$, we observe the inhibitory effects of spheres with radius equal to the rod radius ($\sigma_S = \sigma_R$) at a population ratio of 1 rod to every 20 spheres. To do this, we utilise two analysis tools, one (Figure 4.10) which displays heatmaps of steady state quantities, providing a complete characterisation of average fibrillisation and inhibition behaviours. The second (Figure 4.9) is a series of size distributions for each sampled point in the parameter space which highlights the rod and sphere contribution to each cluster. This figure also provides a series of sample structures observed at different points in the parameter space, highlighted by arrows.

For a system of 1 rod for every 20 spheres when there are 1000 particles total, there are only 50 rods. This makes observations on average values more difficult than for other cases because data will be much noisier due to the smaller sample size. Regardless of this, there are still observations to be made. In this case, it becomes more reliable to look at the trajectory data directly to observe mechanisms rather than to infer it from the observables e.g.Mean rod cluster size and order parameters. In figure 4.10, there are two distinct regions of behaviour in the mean rod cluster-sphere cluster ratio μ_{RS}. Although overall, the available range of these values is very small, there is a divide in behaviours for interaction strengths $E_{RS} \leq 4E_{RR}$ and $E_{RS} > 4E_{RR}$. This behaviour change is most distinct for high aspect ratios, which will become a clear feature of the parameter space even as we change the sphere radius and relative population of rods and spheres.

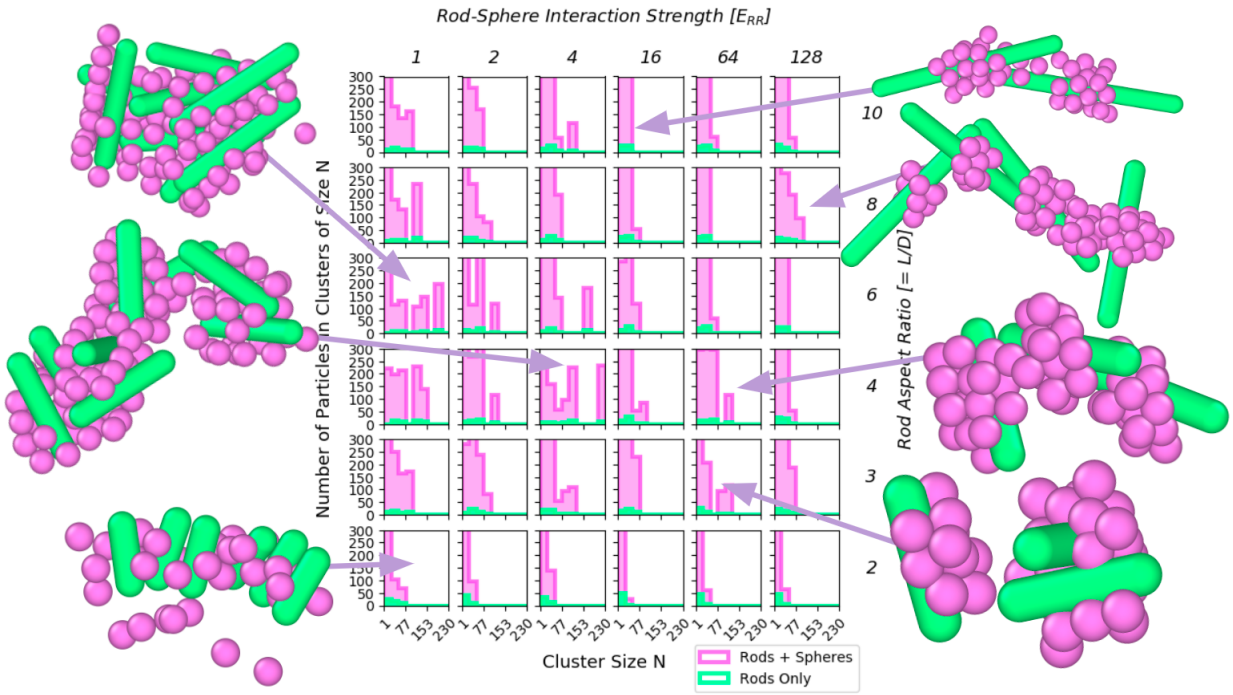

Figure 4.9: A series of steady-state size distributions across a 2D parameter sweep of rod-sphere interaction strength E_{RS} and aspect ratio AR where rod to sphere population ratios are fixed at 1:20. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. Rod sphere interaction strength varies along the x axes in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 10$ where the rod radius is held constant at $\sigma_{ROD} = 1$nm. Rod radius equals sphere radius with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1:20$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT = 310K$

The number of clusters increases with increasing rod-sphere interaction strengths, which highlights the inhibition effect on the growth rate of the fibrils. Rather than many small cluster binding events, the monomers or small multimers of rods are sequestered by inhibitory molecules. Under these parameters, we can see that various mixed cluster morphologies are observed. These clusters predominantly contain inhibitors, largely due to the huge excess of inhibitors in this sweep. Therefore we observe very globular clusters with lower order parameters as the inhibitor binding strength is increased. This is evidenced in figure 4.9 in example structures found for large rod-sphere interaction strengths. The structures contain small numbers of rods with many tightly bound spheres, which disrupt the level of order in these structures. Conversely, the structures observed at lower interaction strengths still contain many spheres crowding the interaction areas of the spherocylinders. However, the clusters are better able to retain their order, as in this region, it

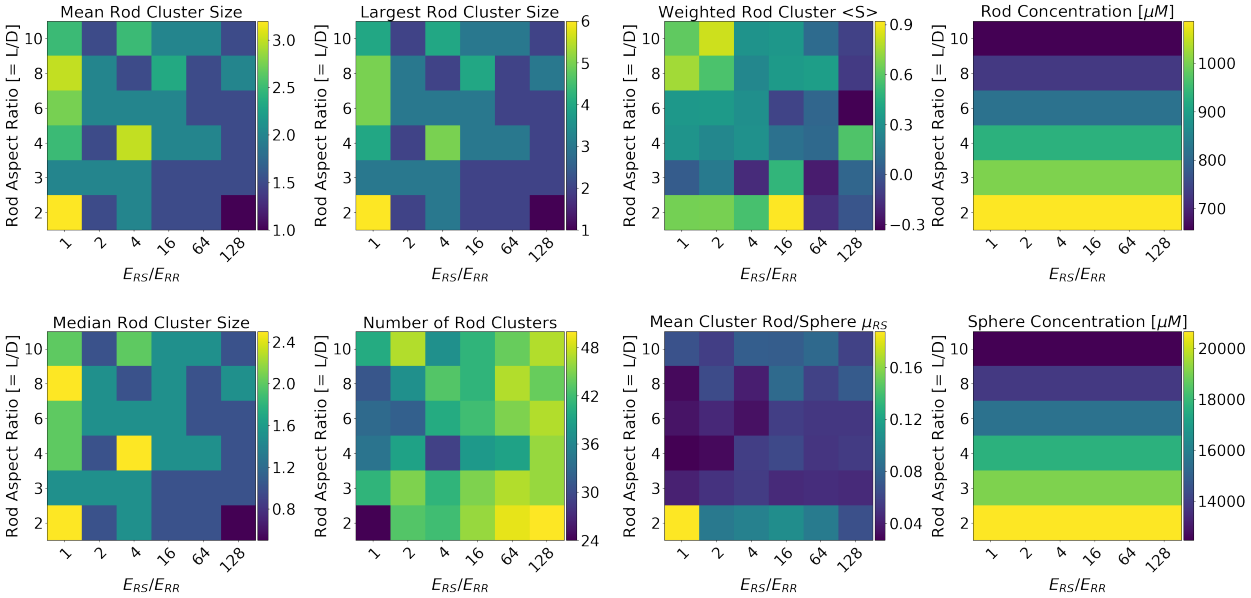

Figure 4.10: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod-sphere interaction strength E_{RS} and aspect ratio AR where rod to sphere population ratios are fixed at 1:20. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 10$ where the rod radius is held constant at $\sigma_{ROD} = 1$nm. Rod radius is equal to sphere radius with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1 : 20$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT = 310K$

is more energetically favourable for spherocylinders to form multiple interactions with one another, as well as many weaker interactions between rods and spheres on the surface of the rod aggregates.

The mean rod-sphere cluster ratio, in fact, increases when the inhibitor binding strength is higher, which might even suggest that the inhibitors are binding competitively with one another i.e.Fewer spheres are binding but much closer to maximising their interactions with the rods, which produce larger free energies than for many weakly binding spheres. Interestingly, this behaviour becomes more pronounced with larger interaction strengths. The mean cluster size, largest cluster size, and median cluster sizes do not change significantly with parameters in this phase space except for the weakest rod-sphere interaction strength $E_{RS} = E_{RR}$. In this region, it is as likely that a rod will bind with a sphere as it is that a rod will bind with another rod. Therefore the rods may bind with

each other as if it were a rod-only system, but with additional steric effects due to the high volume of inhibitors and lower rod particle numbers. We observe trends similar to the rod-only case but with a decreased variance of values due to the low particle numbers.

To conclude, the data, as well as by observation of the trajectories of these simulations (example structures shown in figure 4.9), we observe two distinct types of clusters. The morphology of the aggregates appears to be driven solely by the interaction strength. This parameter drives the tightness of binding and appears to disrupt the size as well as the shape of the aggregates - increasing aspect ratio results in more visually disordered clusters, similar to the rod-only case. However, the order parameter is calculated by calculation of rod clusters only, so bundles of rods separated by spheres in the same rod-sphere cluster will be considered unique clusters. Therefore, these changes are not necessarily reflected in the numeric value of the order parameter in figure 4.10. However, we have observed two regions of mean rod cluster-sphere cluster ratio, which is reflected in the two morphologies. One, where the spheres bind more weakly, but in higher numbers, and in another, the ratio is higher due to fewer rods per cluster and more tightly binding spheres.

1:1 Rod:Sphere Population Ratio

Here, again we explore a 2D parameter sweep of rod-sphere interaction strength E_{RS} and rod aspect ratio AR. In this section, the stochiometric ratio of rods:spheres is increased such that for every 1 rod, there is one sphere. Again, the sphere radius is equal to the rod radius, and the total volume fraction of all particles is fixed at 5%. Due to the changing aspect ratio and the changing volume of rods for increasing aspect ratios, the relative concentrations of rods and spheres will change. These are reflected on the right-hand side in figure 4.12 which contains a series of steady-state observations of cluster averages. We also utilise size distributions (Figure 4.11) which highlight the individual rod and sphere contributions to each cluster bin. This figure also provides examples of typical aggregates formed across the parameter space.

Overall, the median cluster size in figure 4.12 is highest for lower aspect ratios and lower interaction strengths, which corresponds well with rod-only data, suggesting that the inhibition behaviours in this region of the parameter space are not as significant as for increased interaction strengths. In-

creasing rod-sphere interaction strength decreases the mean cluster size, median cluster size, largest cluster size and the orientational order parameter but increases the number of clusters and the mean rod cluster-sphere cluster ratio μ_{RS}. These trends indicate increased inhibition behaviour and the formation of small disordered rod-sphere aggregates.

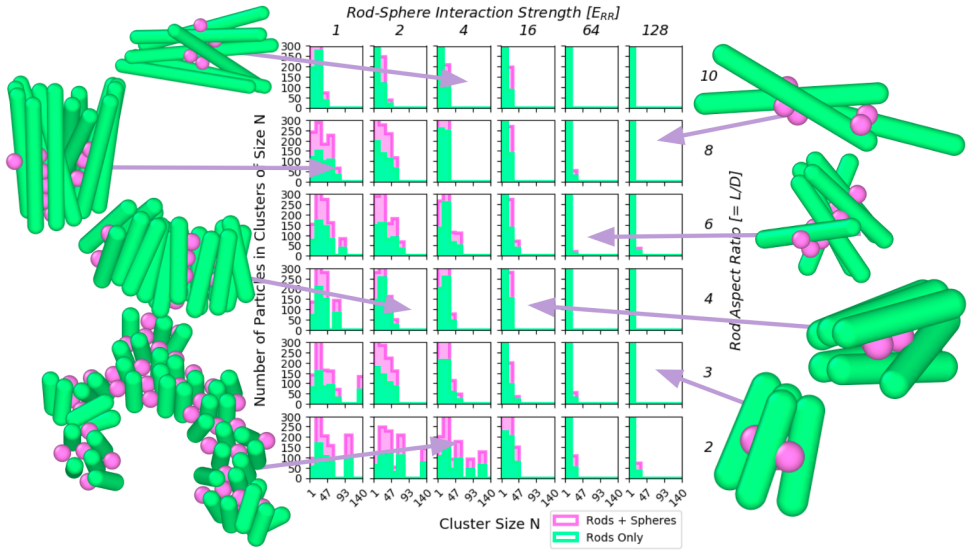

Figure 4.11: A series of steady-state size distributions across a 2D parameter sweep of rod-sphere interaction strength E_{RS} and aspect ratio AR where the ratio of the rod to sphere populations is fixed at 1:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. Rod sphere interaction strength varies along the x axes in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 10$ where the rod radius is held constant at $\sigma_{ROD} = 1$nm. Rod radius equals sphere radius with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT = 310K$

The peak in median cluster sizes (Figure 4.12) best matches the rod-only data when $E_{RS} = E_{RR}$ which agrees well with the equivalent rod-only system as, under these conditions, the chance that a rod will bind with another rod is equally as probable as binding to a sphere. Under all other conditions in this parameter sweep, it is more likely that we will observe rod-sphere binding events - however that is only true if we only consider pairwise binding events and not combined effects from cluster formation. Therefore, for this sweep, with these available conditions, when $E_{RS} = E_{RR}$ we are as close to rod-only conditions as possible and the aggregation behaviours are most likely to map to the rod-only case for $T = 310K$ in section 3.3.2. The data are similar, with a peak for a

mid-sized aspect ratio $AR = 3 - 4$ and a high order parameter across all aspect ratios. However, there are still some binding events indicated by the mean rod cluster size and the mean rod cluster-sphere cluster ratio μ, which is larger than 1 for $AR = 2$ and $AR = 4$. In addition to this, all other aspect ratios have a μ_{RS} value of less than 1, which means there is still some binding to spheres, but it is not the dominating particle type in a given cluster.

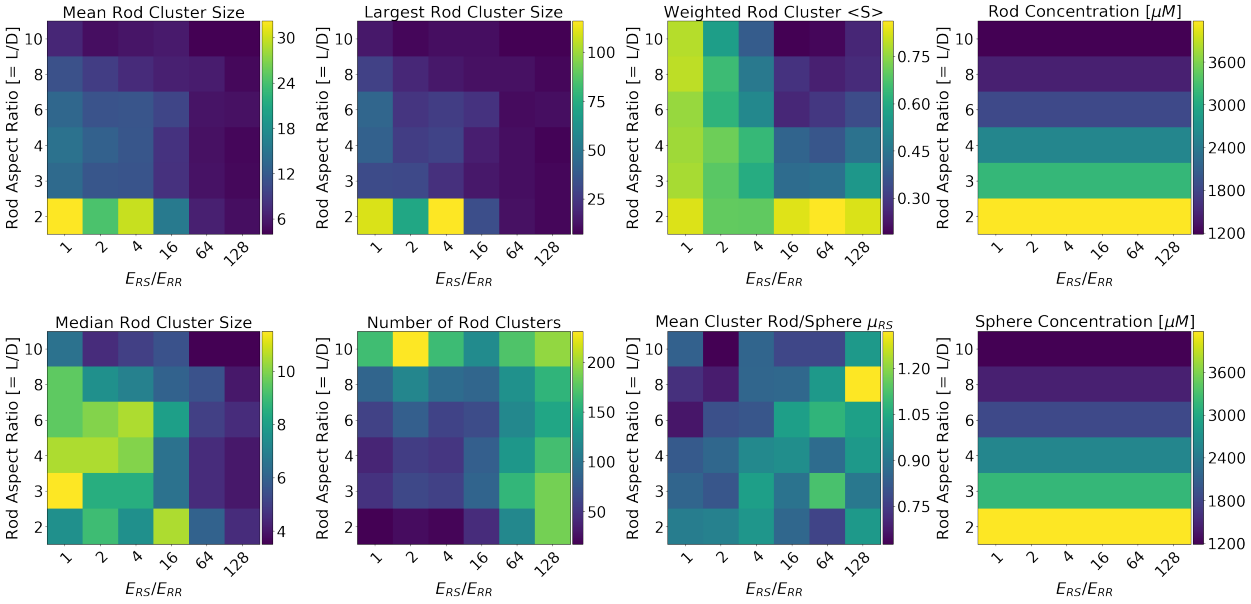

Figure 4.12: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod-sphere interaction strength E_{RS} and aspect ratio AR. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 10$ where the rod radius is held constant at $\sigma_{ROD} = 1\text{nm}$. Rod radius is equal to sphere radius with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT = 310K$

The mean rod cluster size is lowest for the highest rod-sphere interaction strengths as shown by figure 4.12. In this region of parameter space, the binding events are extremely tight and significantly outweigh rod-rod binding events in magnitude. This small mean and largest cluster size indicate that significant inhibition behaviours are occurring in this region. This is further evidenced by a mean rod cluster-sphere cluster ratio which is close to 1 for the majority of high interaction strength cases. A value for $\mu_{RS} > 1$ means that the average cluster contains more rods than spheres. There-

fore in this region, it is likely that there are spheres bound to small clusters of rods. However, the order parameter S undergoes a transition with an increasing aspect ratio in this region. Mean rod cluster size μ_R is also low for the largest aspect ratios $AR = 10$. (Figure 4.12) Rods with these dimensions already had a decreased likelihood of forming multiple large fibrils without the presence of inhibitory molecules. However, the cluster sizes here are much smaller than those formed in amyloid-only simulations, so there is a confirmed inhibitory effect indicated by the change in order parameter for aspect ratio $AR = 10$ simulations as we increase the rod sphere interaction strength. The higher the interaction strength, the lower the order parameter, showing that the inhibitors are disrupting cluster order in this case rather than size. However, the presence of inhibitors in this ratio may be providing steric effects, decreasing the aggregation propensity of these longer rods in the given timescale - the lower the aspect ratio, the larger the order parameter. There are highly ordered structures for $AR = 2$ in particular, which for $E_{RS} = 128E_{RR}$ is paired with a minimal mean cluster size. Therefore we can determine that the structures formed are highly ordered and oligomeric under these parameters. These structures are also likely to have a close to 1:1 rod:sphere binding ratio. Therefore the surface of these aggregates is likely to bind multiple spheres.

$AR = 2$ does not follow the same trends as the other aspect ratios in this system. The most likely reason for this is that rods of this aspect ratio are closest to spherical, and other effects are at play. (what they are, we need to look at specific trajectories to find out). The mean cluster size is largest when rod-sphere interaction strength is lowest and steadily decreases when interaction strength is increased, similarly to the other aspect ratios. However, the drop-off in cluster size is much steeper, and the order parameter of the aggregates remains high, which suggests the spheres/inhibitors are decreasing cluster size but not the order of the monomers in a given cluster. Given that these rods are very short, a sphere of radius equal to the rod radius might not be able to bind inside a fibril structure but might instead break apart structures or cover the surface of the aggregates instead, preventing further growth.

Given that the order parameter broadly increases for decreasing aspect ratios - an effect which becomes more pronounced the more that the interaction strength increases - there appears to be an ability for smaller aspect ratios to retain their order under inhibition. This might be because the

longer rods are better able to still bind to other rods, whilst having inhibitors find and exacerbate gaps and defects within rod clusters. The idea that clusters of shorter rods can retain their order is likely more the case that spheres cannot as easily find defects within said clusters and instead bind to the surface of those aggregates. Therefore we have established that there are several possible inhibition mechanisms even in this one parameter sweep, and we can use size distributions and trajectory information to further our understanding.

4.4.2 Larger Spheres

In this section the sphere radius is 4 times larger than the rod radius $\sigma_S = 4\sigma_R$. Similarly to the cases where radii are equal (section 4.4.1), the results from section 4.3 inform the choice for sphere radius. At a radius $\sigma_S = 4\sigma_R$, the inhibitor properties changed markedly as a single sphere is able to cover the rod binding patch, and inhibition "power" increases, as measured by the impact on the order parameter S. This section explores what happens when AR is varied, so the relative patch size changes.

1:1 Rod:Sphere Population Ratio

In this section we complete a 2D parameter sweep of rod aspect ratio AR and rod-sphere interaction strength E_{RS} for a sphere radius $\sigma_S = 4\sigma_R$. Figure 4.13 is a 2D series of length distributions across incrementally sampled rod-sphere interaction strengths in the range $0.1E_{RR} \leq E_{RS} \leq 128E_{RR}$, we also sample incrementally across rod aspect ratios in the range $2 \leq AR \leq 16$. Rod-sphere interaction strength E_{RS} is a multiple of the equivalent rod-only property i.e.Rod-rod interaction strength E_{RR}. Figure 4.14 highlights a series of steady-state observations of cluster averages such as rod cluster size averages, concentrations and order in rod clusters S.

The orientational order parameter S is highest for the lowest aspect ratios as shown in figure 4.14. However, the mean cluster size is ≈ 5 for $AR = 2$. (Figure 4.14) Therefore, in this parameter space region, we observe highly ordered but small clusters. The inhibitor size in this sweep is larger than the rod size, therefore, the inhibition effect is such that individual spheres sequester small clusters. This is reflected by the peak in low cluster sizes in figure 4.13 for the same conditions. This is the case for all interaction strengths between rods and spheres. The order parameter S decreases

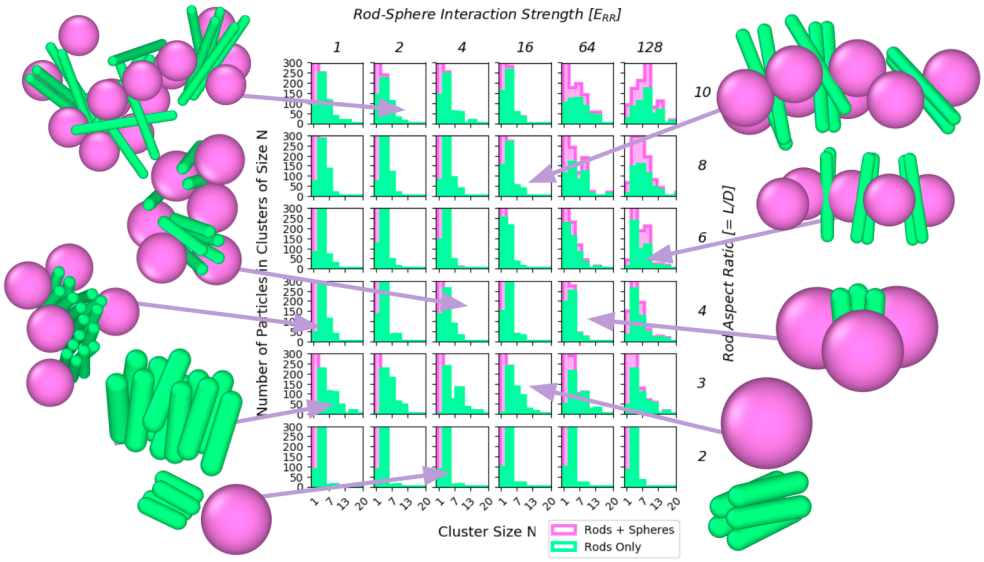

Figure 4.13: A series of steady-state size distributions across a 2D parameter sweep of rod-sphere interaction strength E_{RS} and aspect ratio AR where the ratio of the rod to sphere populations is fixed at 1:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. Rod sphere interaction strength varies along the x axes in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 10$ where the rod radius is held constant at $\sigma_{ROD} = 1$nm. Rod radius, $\sigma_{SPHERE} = 4\sigma_{ROD}$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT = 310K$.

as the aspect ratio increases. This trend also exists for the rod-only data in section 3.3.2, but is exacerbated in the presence of inhibitors.

The mean cluster size is small compared to rod-only data across the whole parameter sweep, which implies there is a global inhibition effect regardless of specific parameters. Therefore, we can determine that the presence of large spheres has an inhibitory effect on the aggregation process due to steric hindrance. Local crowding events decrease the probability that two rods will bind together and increase the probability that a given rod will instead interact with a sphere. However, this is not the only inhibitory effect. For increasing interaction strengths, there is an additional behaviour which is most pronounced for higher aspect ratios as evidenced by the decreasing order parameter and increased number of clusters in figure 4.14. In addition, the distributions in the high aspect ratio, high interaction strength region in figure 4.13 contain large mixed aggregates, suggesting that rod clusters are binding or forming around inhibitors. There is a peak in mean cluster size for

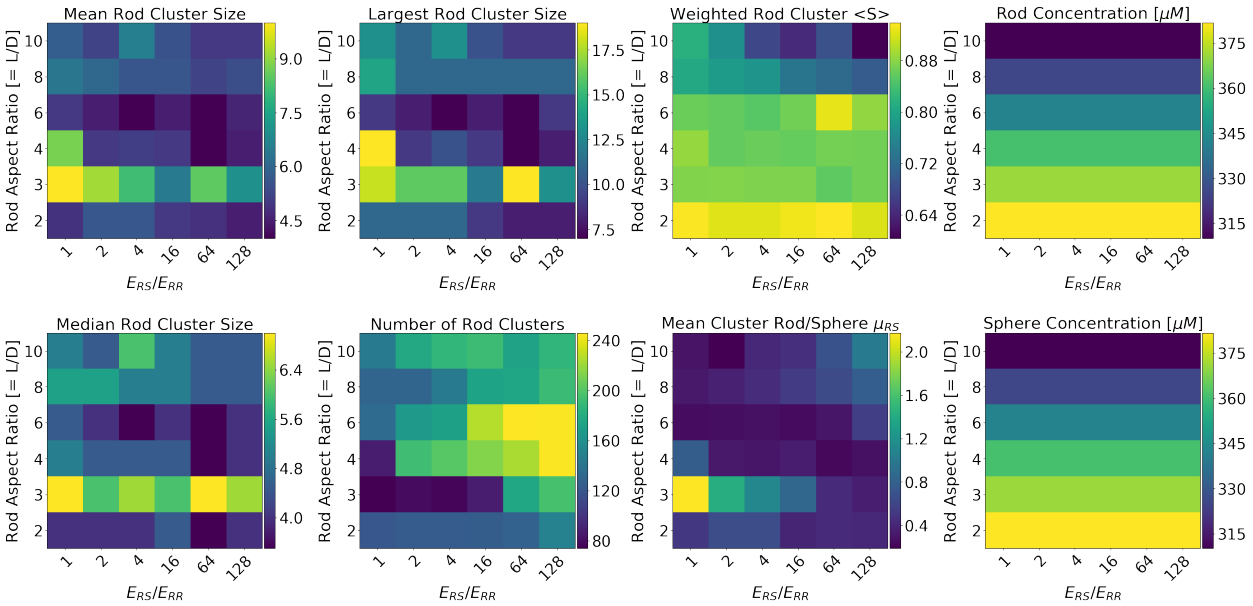

Figure 4.14: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod-sphere interaction strength E_{RS} and aspect ratio AR where the ratio of the rod to sphere populations is fixed at 1:1. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 10$ where the rod radius is held constant at $\sigma_{ROD} = 1\text{nm}$. Rod radius, $\sigma_{SPHERE} = 4\sigma_{ROD}$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT = 310K$.

$AR = 3$. However, this is still lower than the equivalent rod-only case. According to figure 4.13 there are many rod only clusters for $AR = 3$ with no spheres within those clusters. The sphere content is only significant in the smallest cluster sizes, which means that the spheres are either unbound or are forming small clusters with rods. There are still larger clusters of rods forming in this region which are fibril-like. However, the inhibitor content in larger clusters increases with increasing interaction strength which causes a decrease in the mean rod cluster size and a slight decrease in the order parameter. The strongest interaction strengths lead to more globular aggregates due to the binding strength being much higher than for a rod-rod interaction, but specifically damps the elongation rate of biopolymer growth k^+. This is also reflected in the increased number of rod clusters for $AR = 3$ as we increase interaction strength, which suggests that while there is a decrease in the elongation rate, there may be an increase in the nucleation rate. Similarly, this

might be an artefact of the decreased elongation rate, which results in less small rod cluster binding during the inhibition process and leaves many small clusters sequestered by inhibitors rather than available to form fibrils.

The inhibitors appear to be exclusively in small clusters for all regions except for high interaction strength and high aspect ratio. This does not necessarily imply that there is no inhibitory effect. Given that the sphere radius is much larger than the rod radius, it is simply the case that small clusters of rods or monomeric rods are binding to spheres rather than fibril-like clusters. This is significant when we also consider the shape of rod contribution to given cluster sizes in figure 4.13. Again, for most of the parameter space, there are only small clusters with a relatively small number of larger, more fibril-like structures. Therefore we can conclude that there is a significant inhibition effect on fibril formation. Primarily the elongation rate of fibril formation k^+, however it is also possible the nucleation rate k^n will also be decreased under these conditions.

20:1 Rod:Sphere Population Ratio

In this section, we complete a similar parameter sweep of rod-sphere interaction strength E_{RS} and the aspect ratio AR. This time, while the sphere radius remains fixed at $\sigma_S = 4\sigma_R$, the ratio of rod to sphere populations is increased further to 20:1. The two analysis tools we utilise in this section are figures 4.16 and 4.15. Figure 4.16 provides general insight into the steady state properties of each system in the parameter space via heatmaps of average cluster properties, and figure 4.15 provides full-size distributions for each simulation steady state where the individual rod and sphere contributions are highlighted. Between these two forms of analysis, we can gain insight into the different kinds of aggregates formed and their relative population sizes.

The data for a simulation with a low number of spheres will be similar to data without inhibitors unless the interaction strength is high enough to disrupt the fibrillisation process. Therefore we expect to see more significant inhibition for more considerable interaction strengths. This is true for interaction strengths of $E_{RS} => 16E_{RR}$ where there is a decrease in the orientational order parameter S, and an increase in the mean ratio of rod cluster sphere cluster size μ_{RS}.

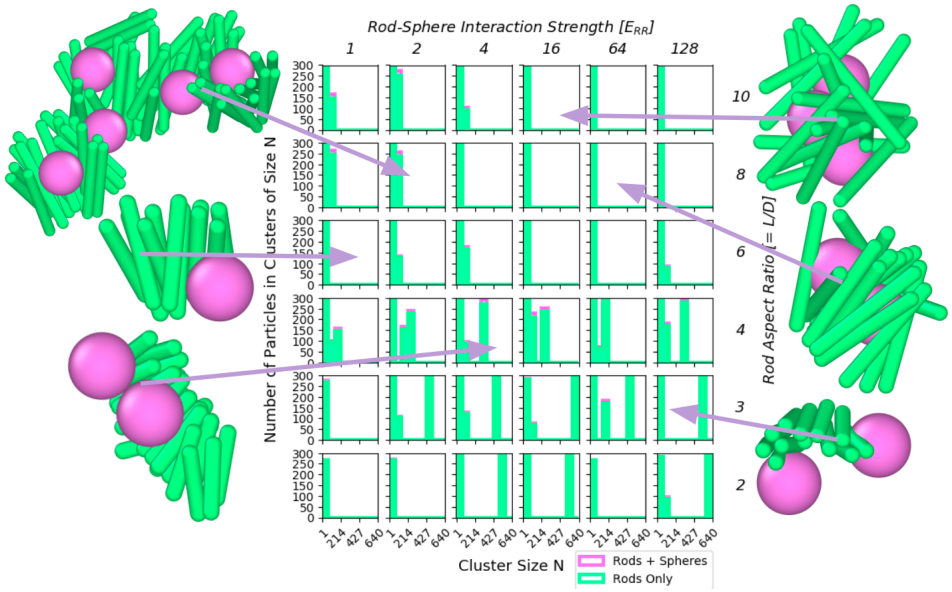

Figure 4.15: A series of steady-state size distributions across a 2D parameter sweep of rod-sphere interaction strength E_{RS} and aspect ratio AR where the ratio of the rod to sphere populations is fixed at 20:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. Rod sphere interaction strength varies along the x axes in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 10$ where the rod radius is held constant at $\sigma_{ROD} = 1$nm. Rod radius, $\sigma_{SPHERE} = 4\sigma_{ROD}$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $20:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT = 310K$

The mean rod cluster size is largest for mid-range, biologically relevant aspect ratios $AR = 3-4$. This is where the clusters appear to be ordered, and there is a low number of clusters which indicates that fibril formation is not inhibited in this central region of the parameter space. However, there is a step change between $AR = 4$ and $AR = 6$. This is indicated in figure 4.16 by a decreased order parameter S and an increased μ_{RS} value. μ_{RS} is not crossing the threshold between rod dominating ($\mu_{RS} > 1$) and sphere dominating ($\mu_{RS} < 1$) here but is distinctly doubling the number of spheres that are bound to a given rod cluster with a single increase in aspect ratio. This may be due to the increased patch surface area provided by a longer spherocylinder, which means that there is room for more than one large sphere to bind. The radius of the spheres in this case is $\sigma_S = 4\sigma_R = 8$nm, whereas the length of the patch is $L_p = 0.6L = AR \times 1.2\sigma_R$. For $AR = 4$, $L_p = 4.8$nm and for $AR = 6$, $L_p = 7.2$nm. The $AR = 6$ patch length is much closer to the radius of a given inhibitor and could feasibly bind more spheres given an interaction radius of $r_c = 1.5(\sigma_R + \sigma_S) = 7.5$nm. A similar step change occurs between $AR = 2$ and $AR = 3$, which is reflected in the decreased order

parameter and increased μ_{RS} for the larger aspect ratio. However, there is also a distinct change in the number of rod clusters for $AR = 2$ which is significantly higher than for $AR \geq 3$. In this case, we have a low mean, median and largest cluster size, a huge increase in the number of clusters and a very low μ_{RS}, which is close to 1. Therefore, the aggregation behaviour of rods with these shape parameters seems to be extremely well inhibited by spheres in this parameter space irrespective of rod sphere interaction strength E_{RS}. The size of the aggregates is significantly affected, however, the $AR = 2$ aggregates remain the most ordered in this given parameter space.

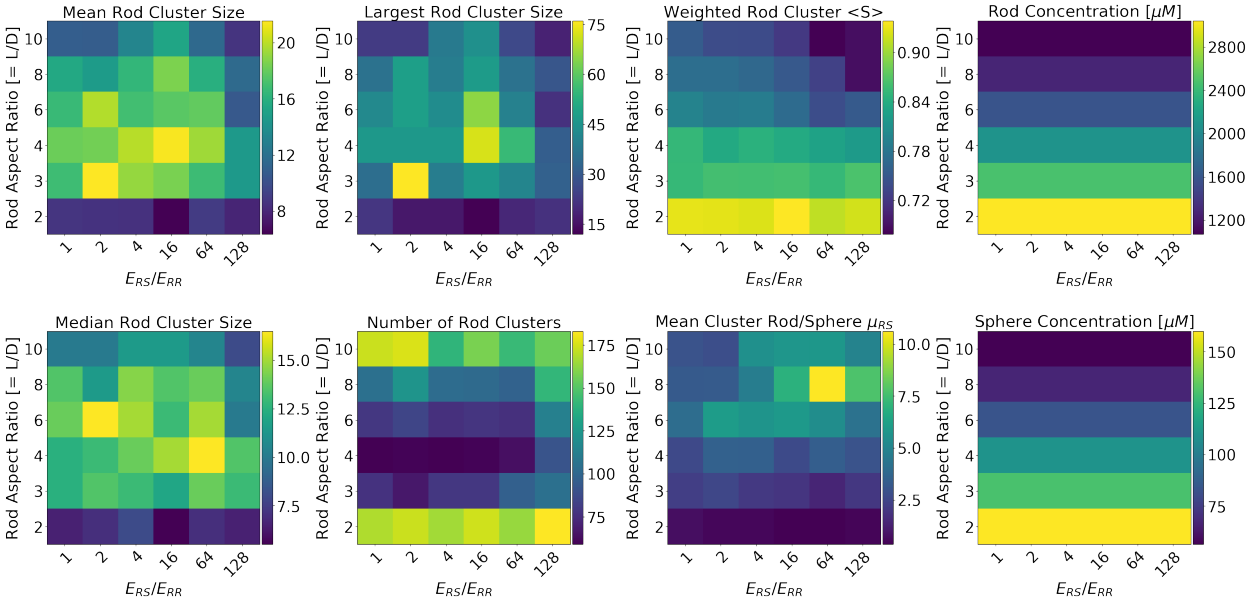

Figure 4.16: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod-sphere interaction strength E_{RS} and aspect ratio AR where the ratio of the rod to sphere populations is fixed at 20:1. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $E_{RR} \leq E_{RS} \leq 128E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 10$ where the rod radius is held constant at $\sigma_{ROD} = 1$nm. Rod radius, $\sigma_{SPHERE} = 4\sigma_{ROD}$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $20:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT = 310K$

It is likely that for the larger sphere sizes used in this parameter space and for weakly interacting rods and spheres, we are observing heterogeneous nucleation events where the surface of the sphere acts as an attractive surface for fibrils to grow from. Despite the rod-sphere interaction strength

being larger than the rod-rod interaction strength, in this weakly interacting regime, one rod-sphere interaction can be counteracted with 2-4 rod-rod interactions. Therefore, the sphere can still bind to a small number of rods, but it is more favourable for nearby rods to form fibrils by bonding with several rods near the sphere rather than binding to the sphere itself. To conclude, there is a narrow range of parameter space separating a diversity of behaviours from inhibition and disruption all the way to fibrillisation catalysts. Simple properties such as dosage and environmental conditions which affect fold (AR and interaction energy E_{RS}) can cause unwarranted changes in biological behaviours.

4.5 Sphere Radius vs. Aspect Ratio

The final parameter sweep in this section is sphere radius and rod aspect ratio. Given that we have already discussed the effect of interaction strengths on these parameters, for completion, we will discuss in this section the effect on sphere radius when we change the aspect ratio of rods. The total volume fraction of particles is fixed at 5%, but we will explore this set of parameters at three points in the rod:sphere population ratio dimension - 20:1, 1:1 and 1:20 rods:spheres. We choose these specific parameters to examine the effects of these extreme stoichiometries which are far from one another yet within the realms of biologically relevant treatments.

In these parameter sweeps, we fix the rod-sphere interaction strength $E_{RS} = 10E_{RR}$ in order to explore the effects of changing shape parameters on inhibition behaviour at a significantly high enough rod-sphere interaction strength such that there will be observable inhibitory effects. In future work, it would also be interesting to test whether the effects observed in this section change for higher rod-sphere interaction strengths or whether the same effects occur on shorter time scales.

20:1 Rod:Sphere Population Ratio

In this section we will explore the parameter space by incrementally changing the sphere radii in the range $0.75\sigma_R \leq \sigma_S \leq 16\sigma_R$, and the rod aspect ratio ($= L/D$) in the range $2 \leq AR \leq 16$. We carry out this parameter sweep with fixed rod-sphere interaction strength $E_{RS} = 10E_{RR}$, and the population ratio of rods to spheres is set to $20 : 1$. This population ratio equates to 950

rods and only 50 spheres, meaning that sphere observations are susceptible to noise. We aim to minimise this noise by using averaged quantities across all clusters. As in our previous sweeps, we use two key analysis tools. Figure 4.17 shows a series of size distributions of mixed clusters across the 2D parameter space of sphere radius and rod aspect ratio. The individual rod and sphere contribution to cluster size is highlighted to gain insight into the variation of particle populations amongst different-sized clusters. Representative structures are also illustrated. Figure 4.18 shows a series of heatmaps depicting steady-state averaged properties of rod clusters and sphere clusters, often found within larger mixed aggregates, as well as input concentrations of rods and spheres separately.

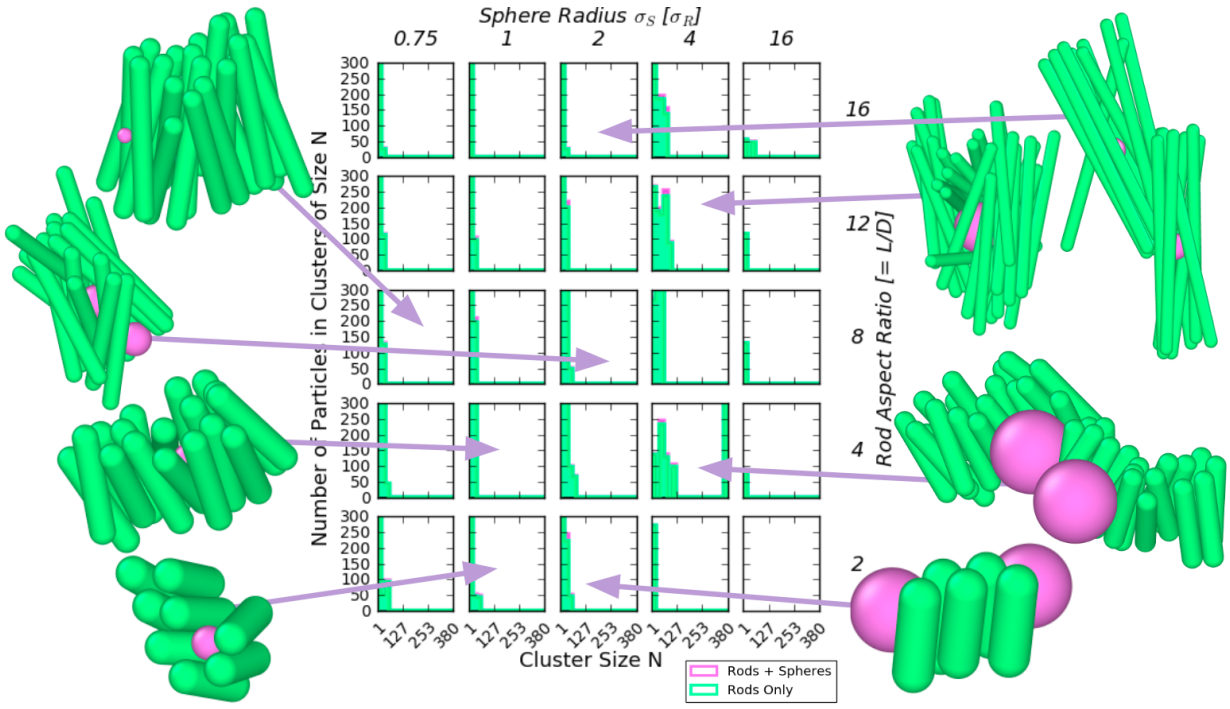

Figure 4.17: A series of steady-state size distributions across a 2D parameter sweep of aspect ratio AR and sphere radius σ_S where the ratio of the rod to sphere populations is fixed at 20:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. $E_{RS} = 10E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $20:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT_0 = 310K$

According to figure 4.18, the largest mean cluster sizes are found where $AR = 4$. Given the high sto-

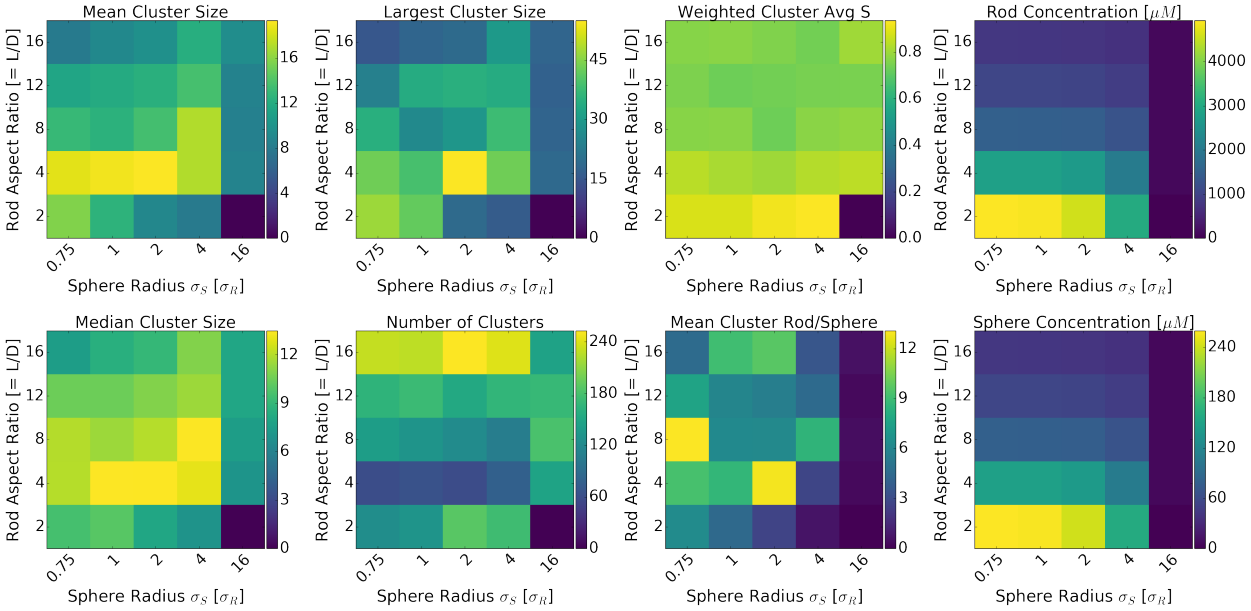

Figure 4.18: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod aspect ratio AR and sphere radius σ_S where the ratio of the rod to sphere populations is fixed at 20:1. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1\text{nm}$. $E_{RS} = 10E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 16$ where the rod radius is held constant at $\sigma_R = 1\text{nm}$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $20:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT_0 = 310K$

ichiometric ratio of rods:spheres, where particles are predominantly amyloid monomers, it is likely that only the largest inhibitory molecules will have a significant inhibition effect at this population ratio. This inhibition effect is reflected in the mean cluster size decreasing on average with increasing sphere radius. We also observe this behaviour in figure 4.17 along the axis of $AR = 4$, although there is some discrepancy between the rod-only mean cluster size in the top left of figure 4.18 and the shapes of the distributions in figure 4.17. For example, the distribution for $AR = 4$, $\sigma_S = 4\sigma_R$ shows an extended distribution with extremely large cluster sizes not reflected in the equivalent mean rod cluster size heatmap. For larger sphere radii, particularly where the rod concentrations are on the order of $10^3\mu M$, (Figure 4.18 top right), spheres tend to conglomerate rod clusters into larger aggregates. It is also possible that the clustering algorithm tends to over-calculate clusters when used on rods and spheres simultaneously. Regardless, it is clear from the example structures

in figure 4.17 that lower sphere radii are better able to fill the space inside fibril-like aggregates and that the large spheres for $AR = 4$ tend to cause fibrillar defects. The example shown highlights a fibril bending around inhibitors. Such effects can occur when the rod-sphere interaction strength is stronger than the thermal energy but not strong enough to overcome multiple rod-rod interaction energies, collectively leading to fibril aggregation behaviours. Moderate cluster sizes can indicate such aggregates with low levels of inhibitor contributions shown in the size distribution.

Figure 4.17 also highlights example structures for long aspect ratios, showing that even for sphere sizes larger than the rod radius, large inhibitors can fill the space in the central region of the aggregate. This does not appear to be possible for inhibitors of radius $\sigma_S = 16\sigma_R$ in the simulations we have carried out. However, we hypothesise that this could be possible if the aspect ratio were increased enough. This work thus provides a rationale for inhibitors' observed stabilisation of "off-pathway" oligomeric structures.

Figure 4.18 shows there is little effect on order in the parameter space. There is an outlier at $AR = 2$, $\sigma_S = 16\sigma_R$ as there is no data for this simulation because the simulation setup could not find an arrangement of particles which do not overlap. This is the case for all parameter sweeps with these shape parameters. The order parameter shown in figure 4.18 is not changing significantly because of the high population ratio. The largest number of clusters is 240, but the total number of spheres is 50. It is likely that each cluster is only bound to one sphere or that multiple clusters are bound to a large sphere. For an intermediate rod-sphere interaction strength $E_{RS} = 10kT_0$, a sphere-rod cluster binding event is not more energetically favourable than multiple rods bound together. Therefore, the inhibitors in this region of the parameter space might disrupt further aggregation, but they do not appear to disrupt the order of existing aggregates at low aspect ratios. This provides further evidence to suggest that the size of the spheres able to fill the space in fibrillar aggregates is proportional to the aspect ratio of the rods in those aggregates. The order parameter in figure 4.18 decreases at higher aspect ratios because spheres can better fill the gaps in the aggregate whilst maintaining rod-rod bonding.

1:1 Rod:Sphere Population Ratio

In this section we explore a 2D sweep of parameters where we vary the rod aspect ratio in the range $2 \leq AR \leq 16$, and the sphere radius in the range $0.75\sigma_R \leq \sigma_S \leq 16\sigma_R$. We hold the total volume fraction of rods and spheres at 5% but set the population ratio at 1:1. This means that each simulation contains 500 rods and 500 spheres. We expect more cooperative inhibition behaviour for this population ratio than for the previous section, which used a ratio of 20:1. Figure 4.19 shows a series of size distributions of clusters across the 2D parameter space of sphere radius and rod aspect ratio. The individual rod and sphere contribution to cluster size is highlighted to gain insight into the variation of particle populations amongst different-sized clusters. Figure 4.20 shows a series of heatmaps depicting steady-state averaged properties of clusters and input concentrations of rods and spheres separately.

In figure 4.20 the mean cluster size and largest cluster size are largest for lower aspect ratios and for smaller sphere radii. There are two effects at play here:

- Aspect ratio changes aggregation behaviour:

  As discussed in section 3.3.2, the largest aggregates are found for lower aspect ratios, and different aggregate morphologies occur for the largest aspect ratios used here. This accounts for an increased mean cluster size for the smaller aspect ratios.

- Smaller Inhibitors must act cooperatively to cause inhibition effectively:

  The mean cluster size is largest for the smallest inhibitor radii, and the number of clusters is the lowest. However, the order parameter S is also decreased in this region, meaning there are large but disordered clusters. Small inhibitory molecules will bind to existing clusters. However, it is still possible for amyloid monomers to continue interacting with these clusters because the inhibitors are not large enough to provide a significant steric effect. The result of this is mixed aggregates in fibrillar conformations where small inhibitors decorate the cluster and prevent "perfect" rod-rod binding, disrupting and therefore decreasing the cluster order parameter.

There is a significant peak in mean rod cluster size and largest rod cluster size in figure 4.20 when $AR = 2$, $\sigma_S = \sigma_R$. However, the order parameter S for these shape parameters is lower than we

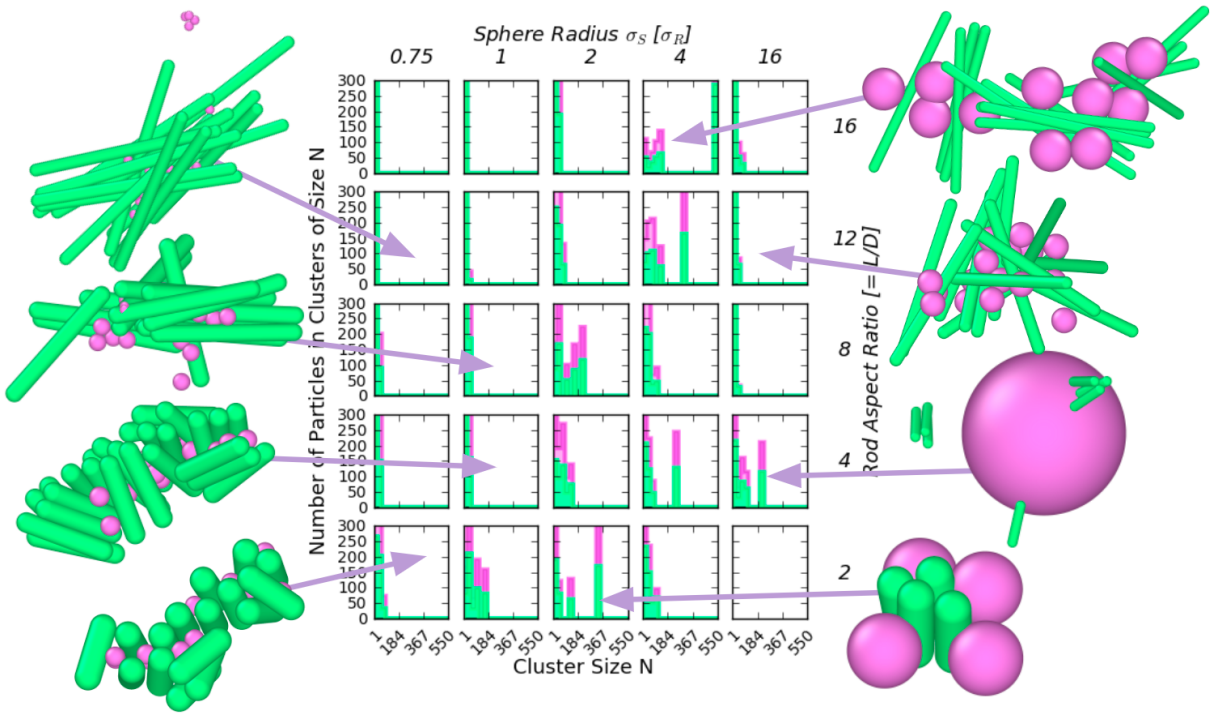

Figure 4.19: A series of steady-state size distributions across a 2D parameter sweep of aspect ratio AR and sphere radius σ_S where the ratio of the rod to sphere populations is fixed at 1:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1\text{nm}$. $E_{RS} = 10E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 16$ where the rod radius is held constant at $\sigma_R = 1\text{nm}$ with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT_0 = 310K$

expect for a typical fibrillar cluster. Rod simulations in section 3.3.2 typically showed an order parameter of $0.8 \leq S \leq 0.9$.

The mean rod cluster and mean sphere data in figure 4.20 is interesting when the cluster size ratio μ_{RS} is 1. When the ratio is 1, ($\mu_{RS} = 1$) the mean number of rod cluster size and mean sphere cluster size are equal. We can assume that this means that, on average, the number of rods and spheres interacting with each other or the number of each in a given mixed aggregate is equal. We can assume this is because the spheres do not interact with one another but do interact with rods. Therefore, sphere clusters can only exist when they are bound closely to rods. Similarly, when the ratio is less than 1, this means that, on average more spheres are in a given cluster than rods,

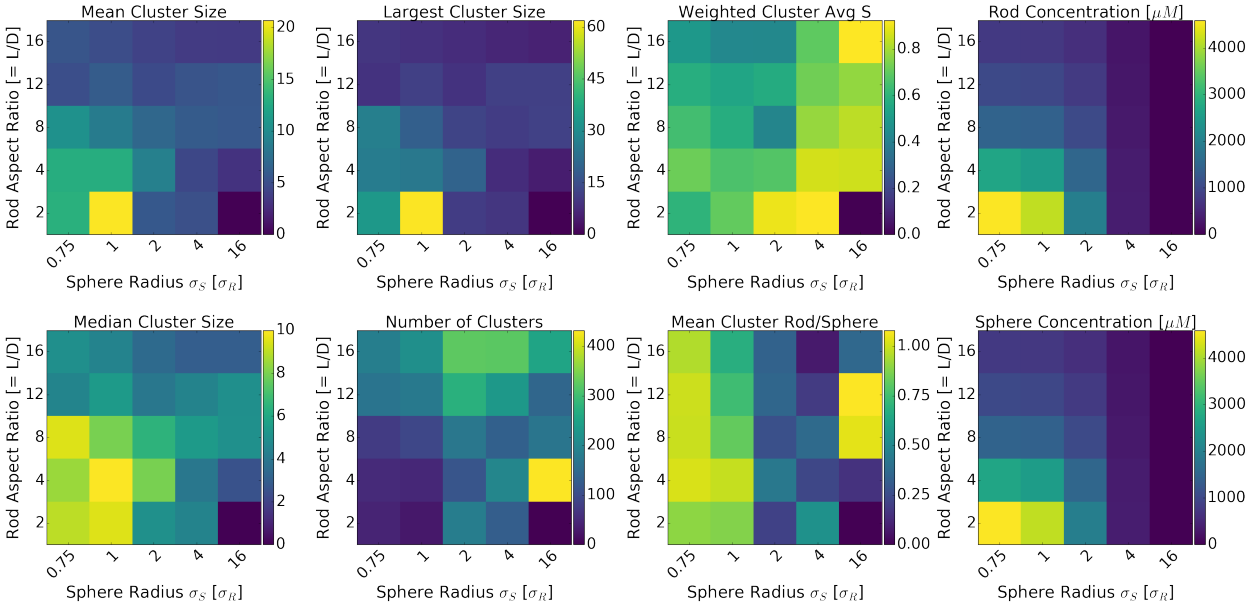

Figure 4.20: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod aspect ratio AR and sphere radius σ_S where the ratio of the rod to sphere populations is fixed at 1. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. $E_{RS} = 10E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1:1$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT_0 = 310K$

implying a high valence of spheres to rods. Finally, when the ratio is larger than 1, on average, each cluster will contain more rods than spheres. This last case is relevant to the small sphere size parameter space, as it shows that the mean cluster ratio in this region is close to or greater than 1. In this region, particularly for lower aspect ratios where the cluster size is large, there are large but disordered clusters due to cooperative small inhibitor binding.

Conversely for sphere sizes $\sigma_S >= 2$, the mean cluster ratio in figure 4.20 is much lower and remains firmly less than 0.5. The order parameter S in this region is also higher, but there is a small decrease in comparison to a rods-only simulation. The rod mean cluster size is also smaller with increasing sphere radius. Therefore we can conclude that in this region, spheres are binding to rods in 2:1 to 4:1 ratios, which are large enough to prevent fibril growth but only slightly decrease the

level of order. The inhibitors with these shape parameters cover the surface of aggregates with a limit of binding numbers related to their volume, whereas the smaller inhibitors are more able to disrupt fibril order by filling and creating gaps in aggregates. It is interesting to speculate that the inherent complexity and flexibility of biological macromolecules may allow a protein, for example, to adopt properties of both small and large molecules. After initial binding as a large sphere, the conformational change would permit local unravelling and insertion of smaller, flexible parts.

The order parameters in figure 4.20 observed when sphere radius $\sigma_S \geq 2\sigma_R$ are much higher than expected when compared to the equivalent example structures in figure 4.19. As discussed in section 4.5, there may be local pockets of order with adjacent rods that register high order parameters - but inhibitory molecules may separate these pockets. This is a discrepancy in the analysis because the order parameter is calculated based on clusters of rods only, so it will register several rod clusters within a mixed aggregate separately, resulting in an artificially high order parameter. By observation, aggregates containing large spheres are, therefore, highly disordered.

To conclude, the aggregation behaviour is changed primarily by the sphere radius in this region of parameter space. The changes due to aspect ratio mimic that of the equivalent rod-only system discussed in section 3.3.2, where $AR = 4$ has the highest propensity to form aggregates, but other aspect ratios lower than 16 are also capable of forming fibril content. The key parameter here is sphere radius, which determines whether the spheres will fill (smaller spheres) or cover (larger spheres) aggregates. (See figure 4.19) The level of order is also highly dependent on this parameter but is not always readily observed in the data (Figure 4.20).

1:20 Rod:Sphere Population Ratio

This population ratio is the most difficult to observe trends with as there are only 50 rods in a given system, meaning it becomes less reliable to observe averages. It is expected that, at this population ratio, there will be significant crowding effects and extremely high numbers of spheres bound to any rod or rod cluster. Figure 4.21 shows a series of size distributions of clusters across the 2D parameter space of sphere radius and rod aspect ratio. The individual rod and sphere contributions to cluster size are highlighted to gain insight into the variation of particle populations amongst

different-sized clusters. Figure 4.22 shows a series of heatmaps depicting steady-state averaged properties of clusters and input concentrations of rods and spheres separately.

In figure 4.22, the number of rod clusters varies between 15 and 50. 50 is the total number of rods in these simulations, and therefore in the cases where the number of clusters is high, we can conclude that the majority of rods are only bound to spheres and not other rods. This is reflected in the mean cluster ratio. For example, the number of clusters is largest when the sphere radius is largest. This is supported by figure 4.21 where in this region of the parameter space, we observe predominantly sphere-based aggregates which surround monomers or small multimers.

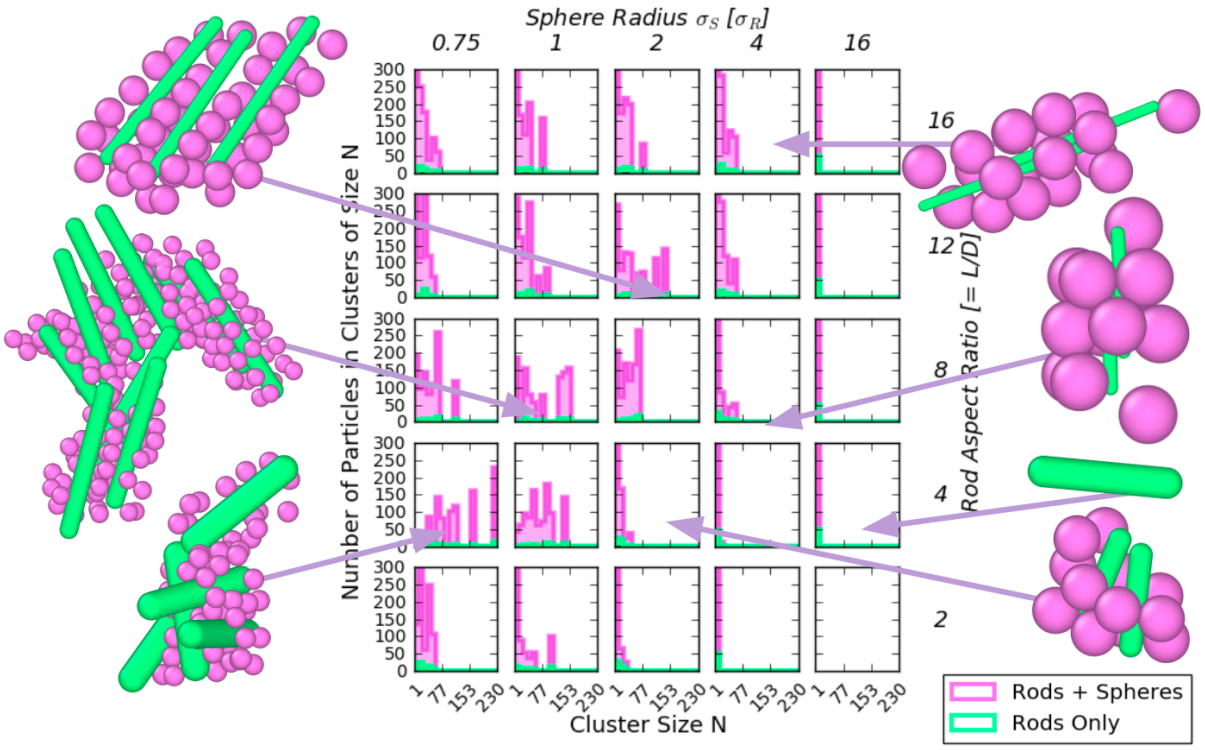

Figure 4.21: A series of steady-state size distributions across a 2D parameter sweep of aspect ratio AR and sphere radius σ_S where the ratio of the rod to sphere populations is fixed at 1:20. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. $E_{RS} = 10E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1 : 20$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT_0 = 310K$

Figure 4.21 encompasses a wide variety of aggregation behaviours. We observe many different cluster sizes and morphologies despite the extreme stochiometric ratio. There are two global trends in behaviour. For larger aspect ratios, the patch length is longer; therefore, more spheres of the same size can bind to the rod simultaneously. Spheres can bind to the full length of the patch. However, it is most energetically favourable for them to bind to the centre. Therefore as we increase the aspect ratio of the rod, we increase the potential valence of the rod as a function of the sphere size. We can see, for example, when $AR = 8$ and $\sigma_S = \sigma_R$ that a larger number of spheres can bind along a single rod than when $AR = 12$ and $\sigma_S = 2\sigma_R$. However, the morphology of these two aggregate examples is very similar. These aggregates are similar to those found in section 4.3, where there is relative order between adjacent rods, not because they are necessarily bound to one another, but because this is a stable energetic minimum for multiple spheres to bind to more than one rod at once. For smaller sphere radii which are less than or equal to the rod radius, there may be cases where the shortest distance between a pair of rods falls within the cutoff distance despite being separated by spheres, which allows the rods to form a bond with one another in addition to the high number of rod-sphere interactions. The level of order in the clusters begins to break down for lower aspect ratios - likely because the patch length is smaller, and therefore the available binding surface for spheres is lower. As well as this, the projection of a short rod from its centre of mass is smaller than that of a large rod, so there is a decreased likelihood that two rods can be separated by a sphere whilst their rod/patch ends are close enough to also interact. This will also contribute to a decreased order parameter in lower aspect ratios, reflected in figure 4.22, despite the noisy data.

We observe a change with increasing sphere radius that results in a change from aggregates with multiple rods to aggregates with only a small number of rods (Figure 4.21). This indicates that there are some stabilising interactions between adjacent rods in clusters containing smaller spheres. Larger sphere radii take up a larger fraction of the interaction volume for a given spherocylinder than a small sphere will. When spheres are large enough in comparison to the rod radius (this appears to be $\sigma_S > 2\sigma_R$), their diameter becomes large enough that adjacent rods separated by a sphere cannot interact, even at random over thermal fluctuations. The spheres coat the interaction area of the patch less efficiently than the small spheres due to their increased volume, resulting

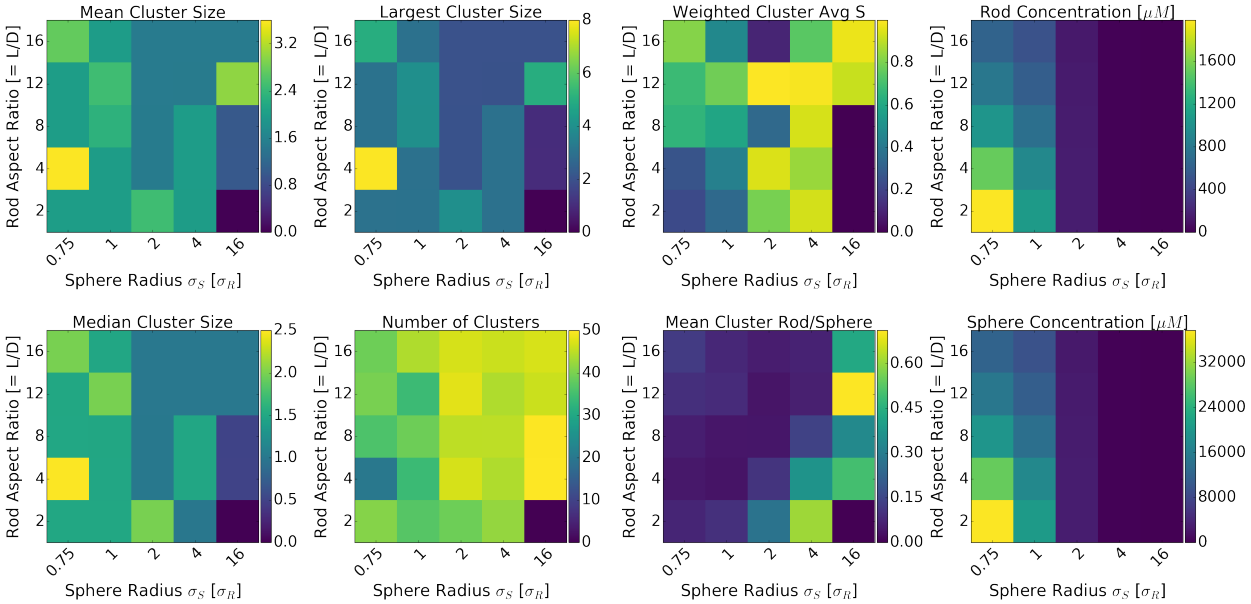

Figure 4.22: A series of heatmaps depicting the steady state results of a 2D parameter sweep of rod aspect ratio AR and sphere radius σ_S where the ratio of the rod to sphere populations is fixed at 1:20. Heatmaps for (top row, left to right) mean rod cluster size, largest cluster size, orientational order parameter weighted by cluster size and rod concentration, (bottom row, left to right) median cluster size, number of clusters, the ratio of mean rod cluster size and mean sphere cluster size and sphere concentration. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_R/\sigma_S \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm. $E_{RS} = 10E_{RR}$ where E_{RR} is the rod-rod interaction strength. The rod aspect ratio AR changes along the y axes from $2 \leq AR \leq 16$ where the rod radius is held constant at $\sigma_R = 1$nm with a combined volume fraction for rods and spheres of $VF = 0.05 = 5\%$. The ratio of rods to spheres is $1:20$ in a simulation box of 1000 particles. Temperature is fixed at $T = 1k_BT_0 = 310K$

in a larger free volume which additional spheres of the same size cannot take up due to packing restraints. These spheres coat the surface of the rod or a small multimer of rods, and due to the high stoichiometric ratio of spheres to rods, it is likely the case that spheres coat these small bundles and decrease the probability that any two-rod clusters will meet. The rod and sphere concentrations shown in figure 4.22 are lowest for these large sphere sizes. This is an implicit fifth dimension of the parameter space, which changes as we make changes to other parameters, given the choice to hold the volume fraction of particles constant at 5%. These changes in morphology can at least partially be attributed to decreases in concentration and, therefore, a decreased probability of rod cluster-cluster contacts as a result of increasing the sphere radii rather than as a direct result of the increased radii themselves. In other sections where the sphere radius is larger than the rod radius, we observe clusters with the morphology shown (in the high aspect ratio, small sphere size

region of this parameter space shown in figure 4.21). Therefore, we can conclude that these effects are most likely due to changes in concentration.

4.6 Discussion and Conclusion

4.6.1 Fibril Disruption

The primary aim of inhibitors in this system is to disrupt fibril formation. In our simulations, this has been achieved in many different ways, including by forming defects within a fibril structure and disrupting the highly ordered structure, preventing further fibril growth or fibril formation from occurring in the first place by sequestering monomer or forming a variety of mixed aggregates.

4.6.1.1 Fibril Capping

Fibril capping is one of the more specific examples of fibril disruption. It is a special case where the fibril ends are bound by the inhibitor or inhibitors, which prevents further fibril growth. The prevention of fibrils from growing into extended structures in the brain could prevent the fibril from maintaining its highly ordered structure and would therefore aid the body in the breakdown of such structures. The larger the fibril, the higher the level of cooperativity of monomers within that structure, and it will become more difficult to disassemble with increasing length.

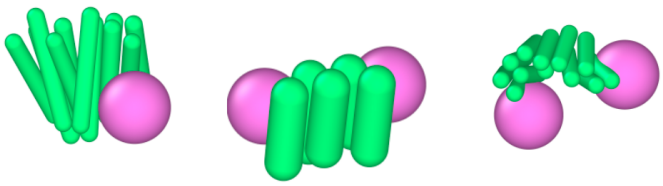

Figure 4.23: Example structures from earlier in the chapter (Figures 4.17 and 4.15) found by changing the available parameters which fall under small cluster formation, inhibiting primary nucleation and elongation rates in an equivalent *in-vitro* system.

We have observed several aggregate morphologies that impact the elongation rate of aggregates using fibril capping - which in this case is a specific case of surface coverage which only considers the ends of the fibrils. By impacting the elongation rate of the aggregates, we can prevent the existing aggregates from growing. For the case of shorter aggregates which have been capped or

prevented from elongating otherwise, this also impacts the secondary nucleation rate. Suppose we consider that a fibril of a given length has a given surface area. By preventing the fibril or aggregate from growing further, we remove the potential for the surface area of the aggregate to increase, thereby capping the secondary nucleation rate at its current rate.

Fibril capping events are rare in our simulation scheme due to the high concentration of rods and spheres. Crowding effects may prevent cases where fibrils are capped and could instead drive the production of more amorphous, globular aggregates. We expect such structures to be more likely observed in systems with low sphere concentration. These structures are also more commonly observed where the sphere radius is larger than the rod radius by a small multiple. e.g. 2-4. This sounds specific, but other behaviours are likely to take precedence when the sphere radius is much greater than the rod radius. When the sphere radius is $\sigma_S \approx 2\sigma_R$, the sphere diameter is close to the diameter of the fibril and, therefore, will effectively cover the growing ends of a given fibril. This behaviour is more commonly observed for high ratios of rods:spheres. As this is a specific subset of surface coverage behaviour, there needs to be a limited number of spheres relative to the number of rods.

Another potential region of parameter space where this behaviour takes place is when the sphere radii are less than the rod radius. This is likely to take place at around stoichiometric ratios of rods and spheres as similar to the latter case, other more probable behaviours will dominate in regions of parameter space where the number of spheres is much greater than the number of rods - such as full aggregate surface coverage (Section 4.6.2). More than one sphere is often required to cap a given fibril if $\sigma_S < \sigma_R$; therefore, this behaviour may occur at higher stoichiometries. i.e. 20:1 rods:spheres up to 1:1 rods:spheres.

4.6.1.2 Fibril Defects

The term "fibril defects" encompasses a huge variety of different behaviours in this system. Many of these proposed mechanisms result in a decreased order parameter compared to a rod-only system. Disrupting the level of order in fibrils is an important step in reducing the binding energy between pairs of rods within fibril structures and their potential disassembly. Once a fibril has formed, it is

almost impossible to disassemble without external stimuli. By disrupting the order level within a given fibril, we are also increasing the entropy for that fibril.

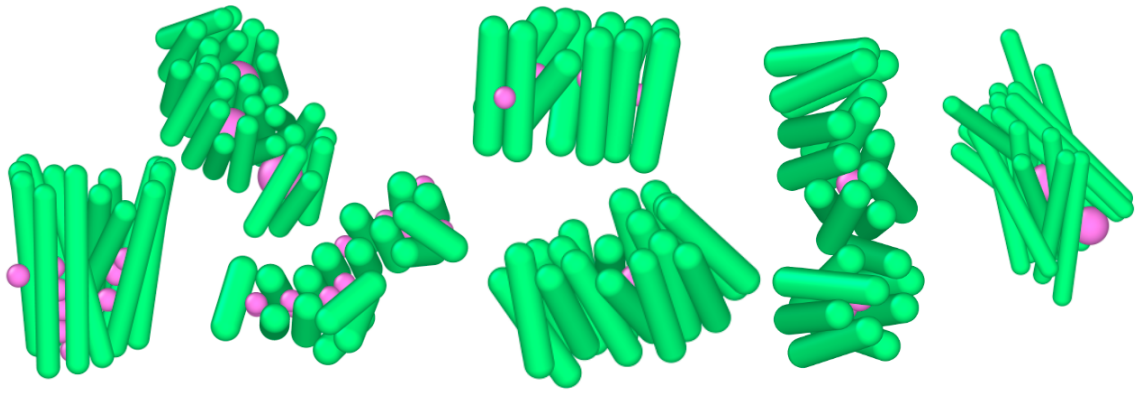

Figure 4.24: Example structures from earlier in the chapter (Figures 4.3, 4.5, 4.11 and 4.17) found by changing the available parameters which fall under small cluster formation, inhibiting primary nucleation and elongation rates in an equivalent *in-vitro* system by creating disordered aggregates better able to be removed by the body in an equivalent *in-vivo* system.

One of the most commonly observed fibril defects in our simulation scheme is small spheres filling the central space in the centre of the fibril along its full length. Figure 4.24 shows some examples of such structures. This can only occur when the rod-to-sphere radius ratio is less than or equal to 1. Naturally, it is easier for smaller spheres to create this type of space inside the fibril without breaking it apart. Therefore these aggregates occur more favourably when the radius is less than the rod radius. *Ex-vivo* cryoEM shows that this occurs. However, the observed structures are not regular enough to identify the small molecules. These structures are still typically quite ordered.

With increasing sphere size, the order in the structure decreases. The associated decrease in order disrupts the fibril surface. Therefore in an equivalent *in-vitro* or *in-vivo* system, the available surface for elongation or secondary nucleation behaviours is disrupted. These structures tend to be shorter than fibrils without inhibitors, suggesting that the presence of the inhibitors affects the available attractive patch interaction volume at the ends of fibrils for further elongation. i.e. This interaction volume might be decreased due to the presence of inhibitors inside the fibrils. The attractive regions may be drawn to the inhibitor rather than a monomer available for addition. Aggregates such as these are typically found when the rod-sphere interaction strength is larger

than the rod-rod interaction strength ($E_{RS} > E_{RR}$) but in approximately the same order of magnitude. For rod-sphere interaction strengths much greater than the rod-rod interaction strength ($E_{RS} >> E_{RR}$), the high rod-sphere interaction strength tends to drive the size of the aggregates down further. Furthermore, as the structure loses stability, it is expected that these aggregates will become susceptible to breakdown by the cellular machinery, i.e. proteases and chaperones.

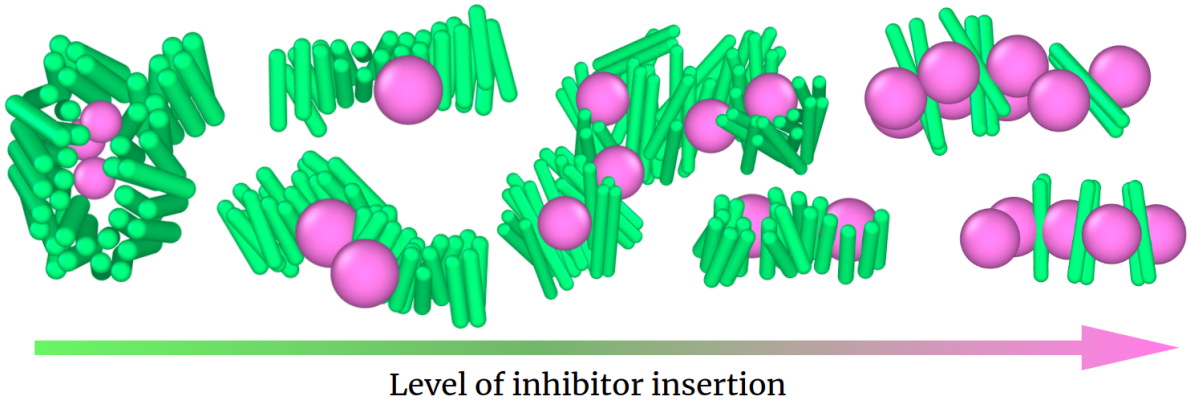

Figure 4.25: Example structures from earlier in the chapter (Figures 4.3, 4.13, 4.15, and 4.17) found by changing the available parameters which fall under small cluster formation, inhibiting primary nucleation and elongation rates in an equivalent *in-vitro* system.

Further fibril defects are shown in figure 4.25. These aggregates encompass degrees of severity of insertion of inhibitors into fibril structures. We consider "severity of insertion" to be a qualitative measure of how much inhibitors have disrupted the order of a fibrillar aggregate by pushing adjacent "layers" of rods apart along the fibril axis. On the left-hand side of figure 4.25, inhibitors much larger than the amyloid monomer bind to sections of a longer fibril-like structure and appear to cause the fibril to bend and twist around the inhibitor (left-most image in figure 4.25). In the other extreme, shown on the right in figure 4.25, the inhibitors have been inserted into the fibril structure, creating large gaps between adjacent pairs of fibrils. There is a glut of examples of the mid-range behaviour found in the central region of the figure. In this region, the order level is disrupted due to short regions of the fibril wrapping around a given inhibitor. The rod-sphere interaction strength determines the severity of this type of inhibition effect and how strong it is compared to the rod-rod interaction strength. In summary, a weak binding between rods and spheres results in fibril structures wrapping around and, in some cases enclosing a group of inhibitors. Increasing this

interaction strength results in less ordered structures and eventually breakages in fibril lengths.

4.6.2 Aggregate Surface Coverage and Secondary Nucleation

An important process to limit amyloid formation is the inhibition of secondary nucleation. As discussed in chapter 1, secondary nucleation is the process of further nucleation events, specifically on the surface of a pre-existing fibril. The current consensus in the literature is that these secondary nucleation events drive the formation of neurotoxic oligomeric structures[60, 105, 345]. Whilst this is not the sole focus for hypotheses surrounding neurotoxic oligomer formation - it is a major one. Regardless of whether the oligomers are toxic, it is in our interest to prevent further nucleation of new fibrils. If, in examples such as in figure 4.26, we observe surface coverage of fibril species. Whilst we do not have an explicit secondary nucleation mechanism, we can see that we have prevented further rod binding to the surface of rods already in a fibril-like conformation. In reducing the available surface for further aggregation, we can reduce the secondary nucleation rate in an equivalent *in-vitro* or *in-vivo* system. Some examples of such structures are found below in figure 4.26.

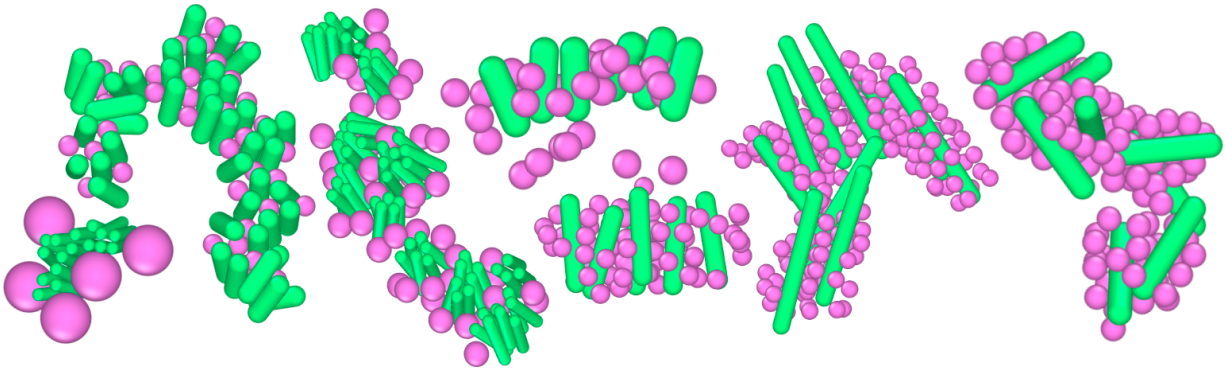

Figure 4.26: Example structures from earlier in the chapter (Figures 4.5, 4.7, 4.9, 4.13 and 4.21) found by changing the available parameters which fall under surface coverage of fibrils or inhibition of secondary nucleation behaviours in an equivalent *in-vitro* system.

4.6.3 Monomer Sequestration

Current theories on the concentration dependence of Aβ aggregation suggest that the aggregation process starts rapidly once the local concentration of amyloid protein monomers reaches a critical

concentration. It is believed that the body stops being able to clear the Aβ monomers quickly enough, and this is the beginning of fibril formation in the brain. A potential pre-AD solution to this would be to use a drug that sequesters individual monomers to prevent extended fibril formation. This could happen in many regions of the available parameter space, however, it is considered more likely at lower concentrations of rods and spheres than in the simulations we have presented in this chapter. The sphere size would need to of comparable order of magnitude in size to the amyloid monomer.

Monomer sequestration can occur in all sphere-to-rod radii ratios where the stochiometric ratio is inversely relatable to the radii ratio. e.g. Regions where spheres are large but low in abundance, where many rods would bind to the surface of the sphere. This is an interesting, specific case as the rod-sphere interaction strength would have to be high enough that it is not more favourable for a fibril to form on the surface of the sphere. Another case to consider would be for small spheres, which are highly abundant and could lead to patch coverage of rods. Again, the rod-sphere interaction strength must be high enough such that it is not more favourable for rods to bind with each other. However, these conditions are very similar to those where we observe small cluster formation. Therefore within the scope of our model, where the binding to a monomeric rod is identical to the binding to a monomer within a cluster, we can treat the monomer-inhibitor binding event as a special case of the small cluster-inhibitor binding event. Figure 4.28 shows an example on the left.

4.6.4 Small Cluster Formation and Amorphous Aggregates

Due to the ambiguity of the term "amorphous aggregate," we can conclude that this would account for a large central region of the given parameter space. Amorphous aggregates in our system are defined as mixed, miscible aggregates of rods and spheres with no defined shape. e.g. A fibril has a defined shape and high-order parameter. Many such mixed aggregates form where the rod-to-sphere radius ratio is approximately one due to the high probability of defects forming which can not be recovered. i.e. Many of these structures may result from the dissociation of large fibril-like structures, which have been disrupted with defects and broken apart into fragments as a result.

Many mixed aggregates also form in regions where the rod aspect ratio is high due to the increased interaction area per particle available. An example is shown in figure 4.27. A rod would be able to bind to several other rods as well as spheres if the sphere radius is not significantly larger than the rod radius. However, this tends to result in a different kind of aggregate.

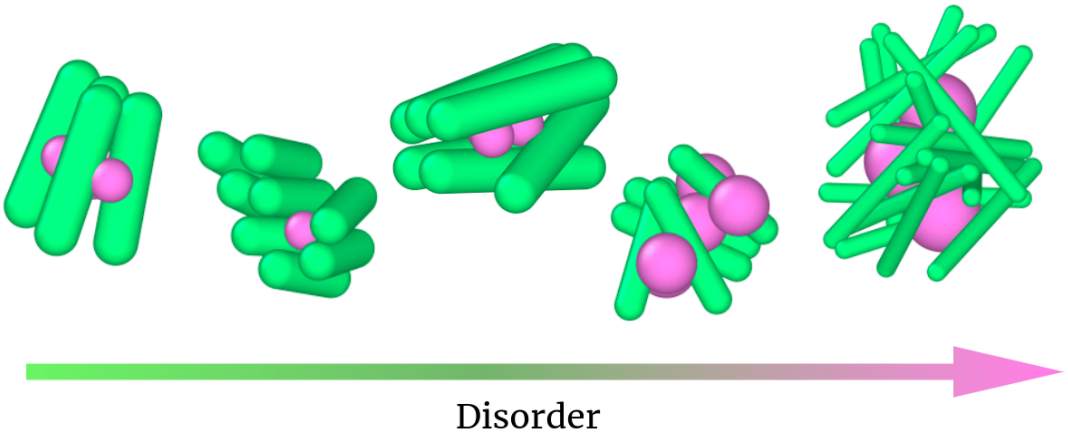

Figure 4.27: Example structures from earlier in the chapter (Figures 4.5, 4.11, 4.15 and 4.17)found by changing the available parameters which fall under small cluster formation and amorphous aggregates, inhibiting primary nucleation and elongation rates, and by extension, dampening other kinetic processes in an equivalent *in-vitro* system.

The common theme between all of these aggregates is that the rod-sphere interaction strength is large in comparison to the rod-rod interaction strength. In addition to this, these much smaller aggregates tend to occur when the stochiometric ratio of rods:spheres is at least 1:1, as in there must be at least as many spheres as there are rods. We observe in figure 4.27 that rods with different aspect ratios are capable of covering spheres of different sizes. The larger the sphere, the less ordered the overall structure. As discussed, the smaller inhibitors can better sit in gaps within a rod structure without great disruption to the degree of parallelism between adjacent rods. These structures tend to be observed at equal stoichiometries. We still observe mixed aggregates for low numbers of spheres, however they tend to be more fibrillar in morphology, unless the interaction strength between the rod and sphere is sufficiently large.

Comparatively, we observe small mixed aggregates where rods are found at the centre, at stochiometries where inhibitors are the dominating species - examples of which are shown in figure 4.28. The

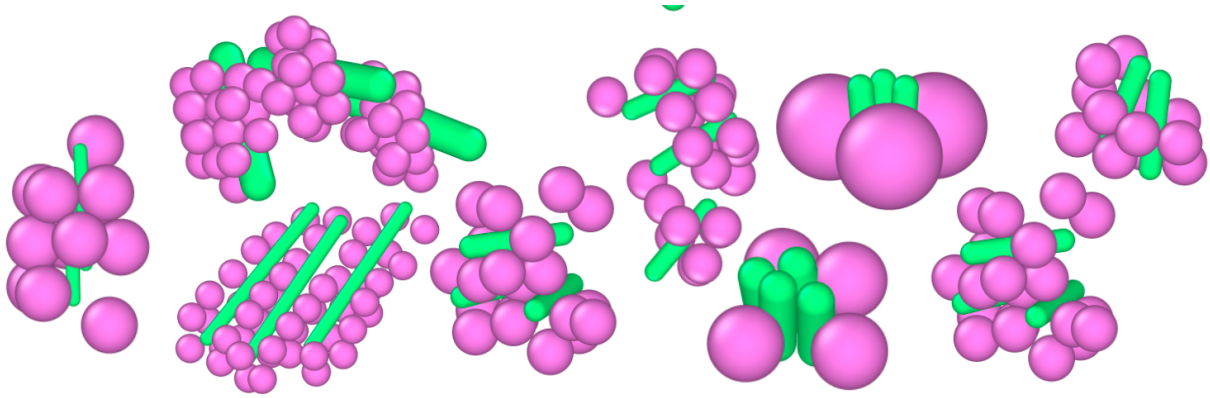

Figure 4.28: Example structures from earlier in the chapter (Figures 4.7, 4.9, 4.13, 4.19 and 4.21) found by changing the available parameters which fall under small cluster formation, inhibiting primary nucleation, elongation and secondary nucleation rates in an equivalent *in-vitro* system.

spheres most effectively cover larger aspect ratio rods due to the increased interaction volume per rod. Dependent upon the rod-sphere interaction strength, we observe different morphologies of these small aggregates. One of which is separated rods which are separated by spheres. This is often the case for longer rods due to the favourability of a rod to either form multiple interactions with spheres or just one or two interactions with other rods. These aggregates are commonly observed when the sphere radii are close to the rod radii. Comparatively, when the sphere radius is larger, it is more favourable for a small group of rods to bind together with strong alignment in favour of a small number of rod-sphere bonds. This region of the parameter space results in small groups of rods surrounded by inhibiting particles which prevent further growth.

There is scope for an argument that by creating small aggregates with mixed surfaces, that these could still go on to produce further fibrillar aggregates in the presence of more amyloid monomers. This would need to be tested further, and some preliminary work is done to this effect in the following chapter. However, these smaller aggregates, in an *in-vivo* environment would have a much higher chance of being processed into waste and/or removed by the body than their extended fibrillar counterpart.

4.6.5 Surface Nucleation

For some cases where the rod-sphere interaction strength is higher than the rod-rod interaction but still weakly interacting, if the sphere radius is large enough, it can act as a surface for fibril

nucleation and growth. This is the opposite behaviour of what this model is intended for. However, it is an extremely common biological process that takes advantage of a high local concentration of protein, highlighting the scope of even a generalised model such as this. Some examples of this are shown above in figure 4.25, where the fibrils are weakly attracted to the inhibitors. The rods are initially attracted to the spheres due to the longer cut-off distance but find it is more favourable to form multiple interactions with other rods, resulting in fibril growth. In this case, the inhibitors do not inhibit and result in long fibrillar aggregates. This is argued as a major mechanism of primary nucleation[122].

Chapter 5

Preliminary Screening of Inhibition behaviour in Seeded Fibril Conformations

In the previous chapter, we explored the parameter space of interactions and shape properties of spherocylinders (rod-shaped amyloid monomers), which form fibrillar structures, and isotropically interacting spheres (fibril inhibitors) in order to find regions of the parameter space which inhibit fibril formation. We sampled rod aspect ratio AR, which we parameterised in section 3.4, section 3.3.2, where we determined that $AR = 4$ is capable of producing the largest fibril populations at physiological temperatures. In addition to this, in a rescaling of parameters where each spherocylinder represents a beta-strand region of an amyloid protein rather than the entire peptide or protein monomer, we determined that different aspect ratios can represent lengths of beta-strand regions and, in turn, numbers of amino acids in a given beta strand motif. The results show that specific lengths of beta-strand motif have a higher propensity to aggregate. However, we are still interested in several aspect ratios and their ability or inability to aggregate in the presence of inhibitors. Therefore we included the aspect ratio in our multidimensional parameter space in chapter 4.2. The properties we sampled in addition to aspect ratio were sphere radius σ_S, rod-sphere interaction strength E_{RS} and population ratio of rods to spheres. While we fixed the total volume fraction that the particles can occupy to minimise severe crowding effects and issues with the simulation

set-up, rod and sphere concentration changed implicitly whilst we explicitly varied the other four parameters. The results of chapter 4.2 highlight a variety of inhibition mechanisms including fibril capping (Figure 4.23), fibril defects (Figure 4.25), small cluster formation of monomer centric (Figure 4.27) and inhibitor centric clusters (Figure 4.28) respectively and small molecule insertion into fibril structures which were found to disrupt the level of order (figure 4.24).

Whilst gaining insight into these inhibition mechanisms and the properties of the inhibitors that drive them is of great importance to our project and has implications for principles of drug design when combating neurodegenerative disease, they are not representative of a diseased state. A diseased brain with Alzheimer's disease, for example, would already have significant fibril populations, insoluble plaques and neurotoxic assemblies, which would be causing a variety of downstream issues. This does not mean that our previous results are invalid. Early diagnosis of neurodegenerative disease would be the most effective time to treat such diseases, where there are minimal assemblies or downstream issues. Therefore the results of chapter 4.2 best characterise inhibitors/drugs which can target excess free monomer before fibril formation or disrupt ongoing fibril formation and assembly. Therefore, to more appropriately represent the disease state conditions in simulation, we seed a distribution of steady-state fibril structures and add inhibitors. If we can determine the properties of inhibitors which can disassemble or disrupt the pre-existing fibrils, then this could have implications for reversing the assembly process in the brain, which could prevent further damage to the brain and even reverse aspects of the degeneration.

By restarting simulations with fully formed fibrils and adding in inhibitory molecules, we can observe a variety of inhibitory effects. We sample the parameter space of rod-sphere interaction strength E_{RS} and sphere radius σ_S in the following parameter sweeps. As in the previous chapter, changing these properties encompasses a wide variety of inhibition behaviours (section 4.3). Therefore we use a similar parameter sweep for these simulations. We established in section 3.3.2 that aspect ratio 4 corresponds to the length of common beta strand motifs in amyloid proteins and produces fibril assemblies the most readily at physiological temperatures. To reduce the dimensionality of the parameter space, we choose a fibril distribution to seed from, i.e. a population of fibrillar assemblies from a simulation in chapter 3, section 3.3.2 with these properties (an aspect

ratio of $AR = 4$). The data seeded into the following simulations were initially carried out at volume fraction $VF = 0.05 = 5\%$. We established in section 3.3.3 that this volume fraction corresponds to concentrations that are artificially high but can be justified to increase the rate of fibril assembly in simulation to reduce computational expense. In this chapter, we do not need to reduce computational expense to form fibrils faster, as there is already a distribution of fibrils. Therefore we decrease the volume fraction of rods from 5% to 0.5%. This volume fraction is equivalent to $\approx 35\mu M$ and better replicates the concentrations of amyloid protein *in-vitro*.[8, 19, 70] However, this means we place the seeded distribution into a larger box than we initially simulated. This results in a non-steady state starting point.

In this chapter, we simulate several parameter sweeps with a seeded distribution of fibrils and incrementally increase the population of inhibitors until we reach a 1:1 stoichiometric ratio of rods to spheres. The stoichiometries we will be exploring are 8:1, 2:1 and 1:1 rods:spheres. We discussed in the previous chapter that stoichiometries, where there is more than one inhibitor per monomer of amyloid protein, were unlikely to be utilised from a drug design perspective due to the danger to the brain chemistry of including such large quantities of any molecule. Therefore, we have selected several stoichiometric ratios where the numbers of inhibitors are less than or equal to the number of amyloid monomers. In this chapter, we utilise the same analysis tools described at the start of chapter 4.2.

For completeness, we start with the seeded distributions of fibrils into a box with no inhibitors present. The structures are seeded into one section of the box and are left to diffuse. The inhibitors will be distributed randomly about the simulation box and, upon insertion, may displace individual rods, resulting in the redistribution and random placement of said rods. Therefore, the results presented in this chapter are preliminary. A more rigorous set-up procedure will be used for future iterations, resulting in the random distribution of individual fibrils throughout the new, larger simulation box. An example of this is shown in the control case section below in figure 5.1.

5.1 Control Case: No Inhibitors

Here, we have seeded a distribution of fibril structures into a large box of volume fraction $VF = 0.005 = 0.5\%$. In this section, we have not included inhibitors, as it is important to determine first that the distribution of structures is stable over the expected simulation time period ($\approx 8 \times 10^6$ MC Steps). The seeded structures fill one corner of the simulation box due to the reduced rod volume fraction and increased box size. However, with periodic boundaries, they will diffuse randomly over time through the box or over the boundary, with no directional bias. Most fibrils in these simulations contain tens of monomers, so they are expected to diffuse slowly.

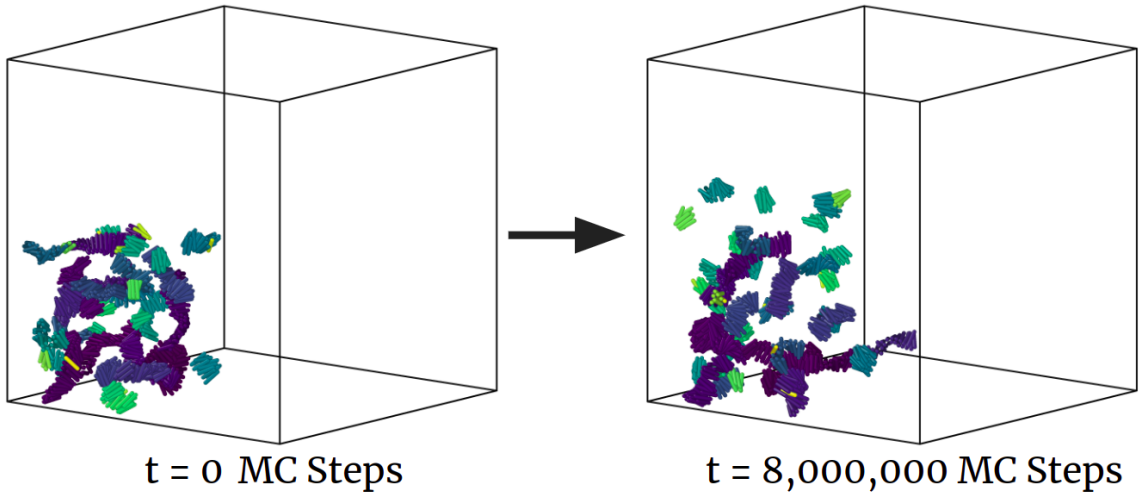

Figure 5.1: Example simulation boxes at $t = 0$ MC Steps and $t = 8 \times 10^6$ MC Steps. Seeded rods from a previous simulation where rods have an aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$. Temperature is fixed at $T = 1k_BT_0 = 310K$.

Figure 5.1 shows an example of the initially seeded fibril distribution at $t = 0$ MC Steps and the final distribution after the simulation time period. We find that in previous cases (for example, section 3.3.2) that 8×10^6 MC Steps is a sufficient period for fibril assembly. Therefore, whilst the distribution has not diffused fully throughout the simulation box, we expect that if disassembly were to occur, it would have happened within this period.

Figure 5.2 shows the size distributions (top) and simulation snapshots (bottom) for two repeats

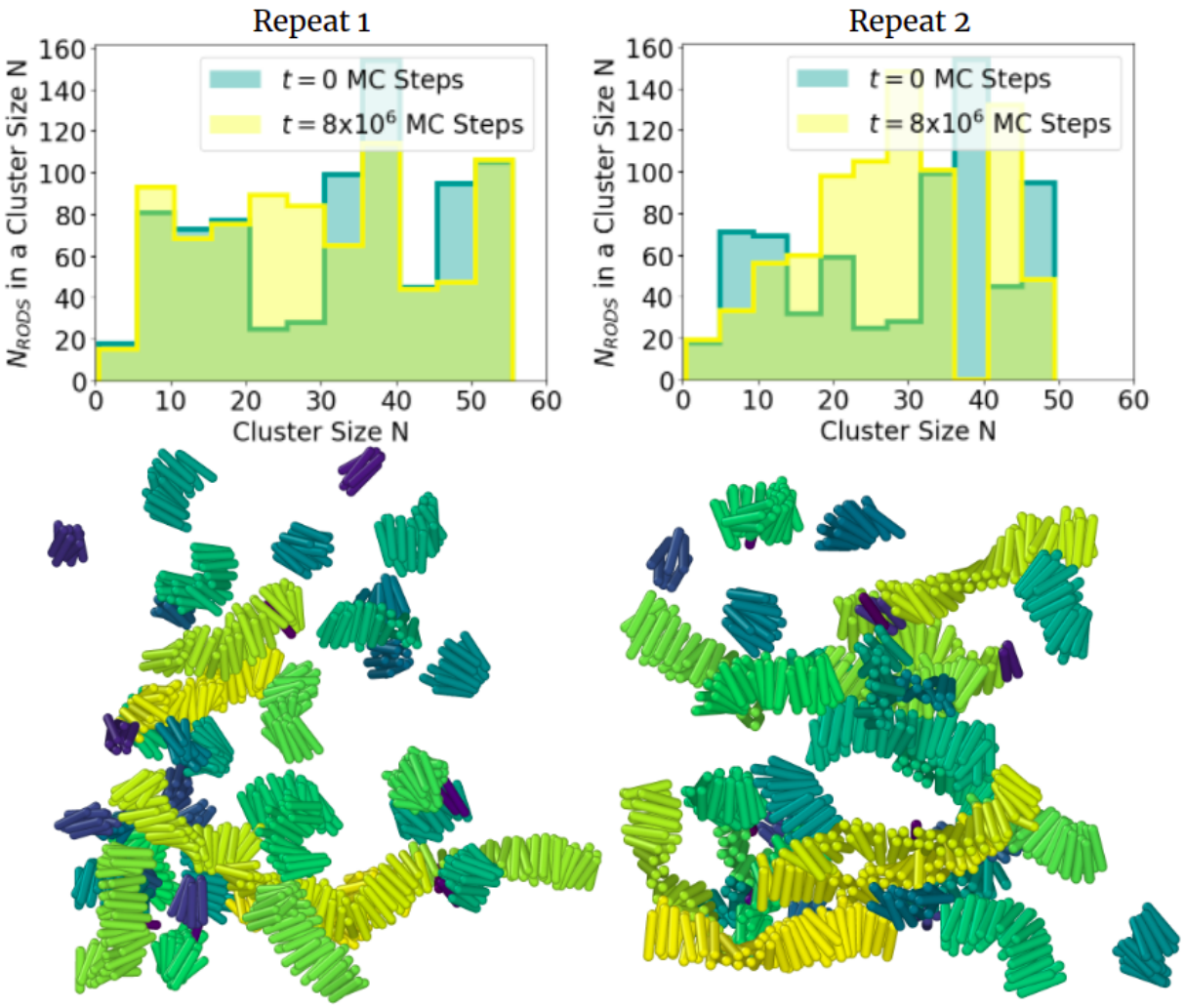

Figure 5.2: Example simulation boxes at $t = 0$ MC Steps and $t = 8 \times 10^6$ MC Steps. Seeded rods from a previous simulation where rods have an aspect ratio $AR = 4$ and interaction strength $E_{RR} = 6k_BT_0$. Temperature is fixed at $T = 1k_BT_0 = 310K$.

respectively. The size distributions undergo small changes in distribution shape throughout the time course. The mean cluster sizes are $\bar{x} = 23, 24$ respectively for the two repeats at the endpoints of each simulation. Therefore, despite some movement in each distribution in figure 5.2 throughout the simulations, it is clear that the mean cluster size is not moving away significantly from the value of $\bar{x} = 24$ at $t = 0$ MC Steps as the clusters diffuse into the available space. The bottom of figure 5.2 shows the simulation endpoints for each of the repeats. Each contains a distribution of fibrillar aggregates which twist and bend, as discussed in chapter 3. We conclude that the differences between the originally seeded distribution before and after the simulation are insignificant.

Therefore, we can proceed with the inclusion of inhibitors to measure the level of disruption away from this steady state.

5.2 Adding 100 Inhibitors

In this section we simulate a 2D parameter sweep varying the rod-sphere interaction strength E_{RS} in the range $0.5 \leq E_{RS}/E_{RR} \leq 128$ where E_{RR} is the rod-rod interaction strength, and sphere radius σ_S in the range $0.75 \leq \sigma_S/\sigma_R \leq 4$ where σ_R is the rod radius. We seed a distribution of fibrils with aspect ratio $AR = 4$ and add inhibitors with properties sampled in the aforementioned 2D parameter space. We use two analysis tools to discuss the results of the parameter sweep, displayed in figures 5.4 and 5.3. Figure 5.3 presents a series of size distributions at different points in the parameter space, highlighting the contributions of each cluster size from rods and spheres, respectively. However, at this stoichiometric ratio (8:1 rods:spheres), there are significantly fewer spheres than rods; therefore, sphere contributions to clusters are potentially difficult to observe. Figure 5.4 shows a series of steady-state properties of the clusters. Many of these are specific observations of rod clusters, but we also include the ratio of the mean rod cluster size and mean sphere cluster size μ_{RS}, which is calculated using the average contributions to each mixed cluster size in the distributions highlighted in figure 5.3.

The broad trends in this sweep are highlighted in the mean rod cluster size (a more apparent inverse trend highlighted by the number of rod clusters heatmap) and the rod order parameter (Figure 5.4). Despite the noise in the data, the heatmaps highlight a trending decrease in mean fibril size and fibril order as we increase the rod-sphere interaction strength E_{RS}. This suggests an increased inhibition effect as we increase the inhibitor binding strength, which is a repeated result from section 4.3 in the previous chapter. This time, we are not looking at the inhibition of fibril growth; we are looking for disassembly behaviours or changes in order parameter. In this region, an increased number of rod clusters (Figure 5.4) suggests fibril breakage/fragmentation or dissociation events are more probable when the inhibitor binding affinity increases. This trend is exaggerated as the inhibitor radius increases, with the largest number of rod clusters when the inhibitor radius is at its largest in this sweep ($\sigma_S = 4\sigma_R$). The top right of figure 5.3 highlights

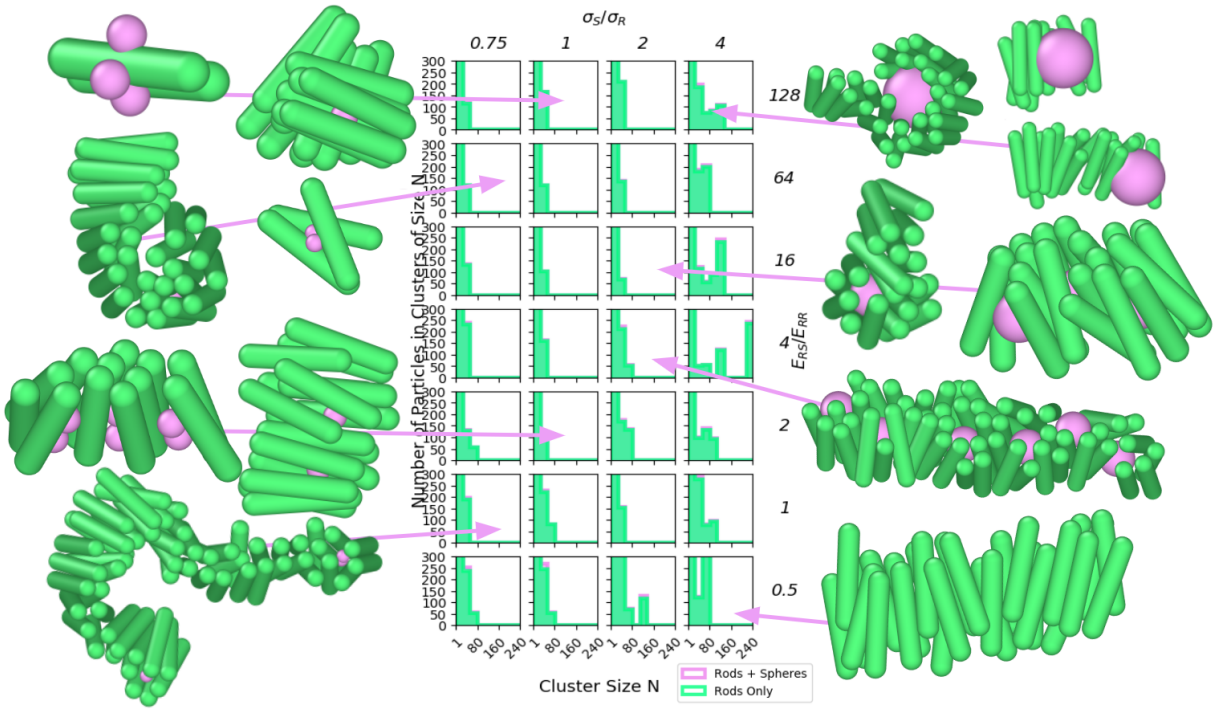

Figure 5.3: A series of steady-state size distributions across a 2D parameter sweep of aspect ratio AR and sphere radius σ_S where the population ratio of rods to spheres is 8:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_S/\sigma_R \leq 4$ where the rod radius is held constant at $\sigma_R = 1$nm. Rod sphere interaction strength E_{RS} varies along the y axis in the range $0.5 \leq E_{RS}/E_{RR} \leq 128$ where E_{RR} is the rod-rod interaction strength. Temperature is fixed at $T = 1k_BT = 310K$.

example structures produced in this region of the parameter space, indicating fibril capping mechanisms and sequestering of fibril structures by the fibril wrapping around a large sphere. This latter example suggests that inhibitors with these properties might not be appropriate as inhibitors, as the structure still has growing ends available for monomer binding. The sphere may have only provided a highly attractive surface for further fibrillisation. However, the mean rod cluster size is significantly decreased for $E_{RS} = 128E_{RR}$ and $\sigma_S = 4\sigma_R$ (Figure 5.4). The other given example structures highlight fibril capping, preventing further elongation of the fibrils on one end due to the inhibitor covering the attractive surface on the individual monomers. Therefore the fibril wrapping around the inhibitor is a low probability event.

Throughout the parameter space, the median cluster size (Bottom left figure 5.4) by inspection is noisy. However, the range of these values is ≈ 3; therefore, we can determine that there are

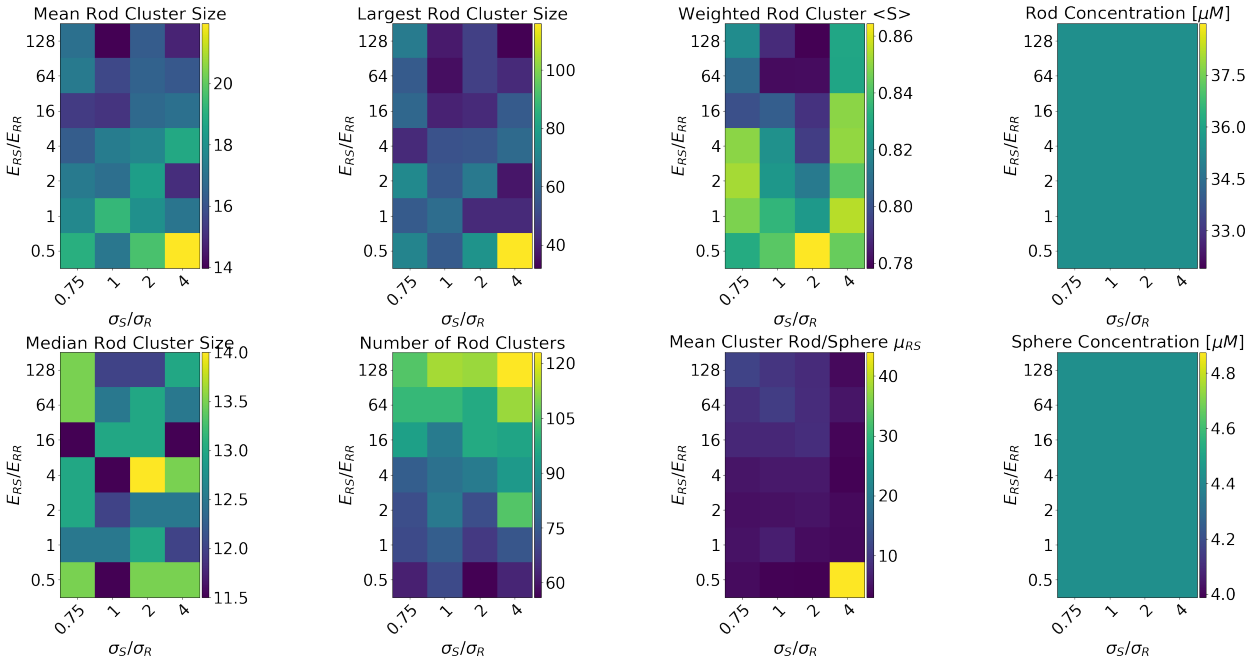

Figure 5.4: A series of steady-state observations of clusters in the form of heatmaps for a 2D parameter sweep of rod-sphere interaction strength and sphere radius where the population ratio of rods to spheres is 8:1. Heatmaps for (top row, left to right) mean rod cluster size, largest rod cluster size, orientational rod order parameter weighted by rod cluster size and rod concentration, (bottom row, left to right) median rod cluster size, number of rod clusters, the ratio of mean rod cluster size and mean sphere cluster size μ_{RS} and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $0.5 \leq E_{RS}/E_{RR} \leq 128$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 <= \sigma_S/\sigma_R <= 4$ where the rod radius is held constant at $\sigma_R = 1$nm. Temperature is fixed at $T = 1k_BT = 310K$.

inhibition effects for small and large aggregates across the whole parameter space, given that other observable quantities are changing as the properties of the inhibitors change. The redistribution of monomers can explain the changes in the distributions' small cluster/small multimer range at the start of each of the simulations, which result in the formation of small numbers of rods bound to spheres across the majority of the parameter space. (See examples in figure 5.3.)

There appears to be an anomaly in the value of the ratio of mean rod and sphere cluster sizes, μ_{RS} (Middle bottom right figure 5.4) when $E_{RS} = 0.5E_{RR}$ and $\sigma_S = 4\sigma_R$. This anomaly occurs in a region of parameter space where there is no apparent inhibition effect, as highlighted by the bottom right of figure 5.3, where the sample structure is purely fibrillar and represents all structures in that given simulation. Inhibitors with radii less than or equal to the monomer radius are able to enter

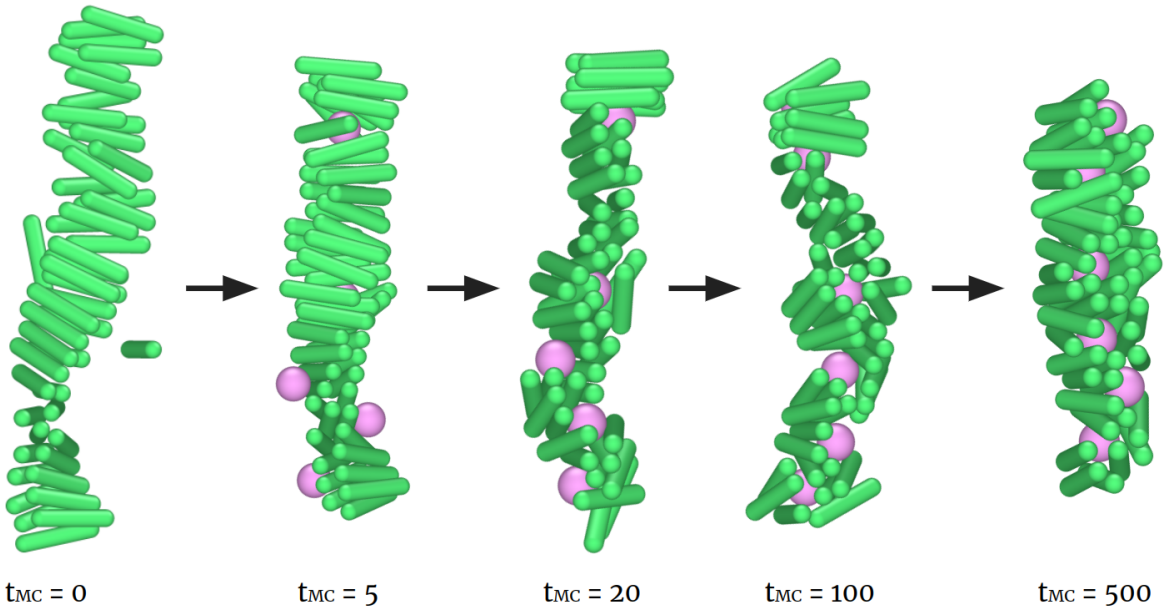

Figure 5.5: Snapshots of inhibitor insertion into a fibril structure over a time course of 5×10^6 MC sweeps. Timepoints in the diagram are multiples of 10^4. The rod-sphere interaction strength is $E_{RS} = 4E_{RR}$ and the inhibitor radius is $\sigma_S = 2\sigma_R$ where the rod radius is held constant at $\sigma_R = 1$nm. Temperature is fixed at $T = 1k_BT = 310K$.

fibril structures and interact with multiple monomers (see bottom left figure 5.4) even when the interaction strength is low. Larger spheres cannot undergo this behaviour as they cannot overcome the energetic cost of disrupting the rod-rod interactions and, therefore, the order. Therefore, when $\sigma_S = 4\sigma_R$ and $E_{RS} = 0.5E_{RR}$, there are no inhibitor-amyloid clusters resulting in a division by 0 for the calculation of the erroneously high μ_{RS}. This anomaly in μ_{RS} is, in fact, indicative of a lack of inhibition behaviour.

The remaining values of μ_{RS} (see middle bottom right figure 5.4) are lowest for low rod-sphere interaction strengths E_{RS} and large sphere radii σ_S. The largest value is found for high interaction strengths and small sphere radii. i.e. When $E_{RS} = 128E_{RR}$ and $\sigma_S = 0.75\sigma_R$. This indicates that there are more spheres per cluster than in any other region of the parameter space, which is reflected in the sample structures in the top left of figure 5.3. Similar events occur for all small sphere radii indicated by the low μ_{RS} and sample structures. These small inhibitors can insert themselves into fibril structures, causing structure remodelling and disruption to the level of order, indicated by the order parameter S. (Figure 5.4.) Although this is not a very large decrease in the

total average value of S ($S \approx 0.78$ compared with $S \approx 0.85$ in rod-only cases), the visual change in the sample aggregates (figure 5.3) indicate a remodelling behaviour.

Finally, we explore intermediate sphere sizes in the central region of the sampled parameter space. In figure 5.5 we observe several snapshots over the course of the simulation where $\sigma_S = 2\sigma_R$ and $E_{RS} = 4E_{RR}$. The equivalent steady state structure is shown in figure 5.3. These intermediate parameters show that larger sphere radii are capable of inserting themselves into fibrils and disrupting their order, however, the energetic barrier necessary for such large spheres to insert themselves is much larger than for smaller spheres, therefore requiring larger interaction strength constants. This is why we do not observe this behaviour at lower interaction strengths. According to figure 5.5, the spheres insert themselves into an elongated fibril structure which causes significant disruption to the order. The fibril structure then undergoes significant reordering in the central stages, which results in a shorter, thicker cluster. The cluster is still in a fibril-like conformation but, by inspection, is far less flexible. There is less twist between adjacent rod layers in the fibril and a more rod-like regime of persistence length. This behaviour is disruptive to fibril order, but it is questionable whether this behaviour will prevent further elongation of the structure.

To conclude, in this stoichiometric ratio, increasing the interaction strength disrupts the fibril order irrespective of sphere radius. However, the sphere radius determines the mechanism similarly to section 4.3. Large sphere radii typically result in fibril capping or wrapping of fibrils around the inhibitors. Smaller radii are much more likely to insert themselves into fibrils and disrupt them by pushing adjacent rods apart. Larger inhibitor sizes are also capable of this behaviour (Figure 5.5) but require an increased rod-sphere interaction strength when compared with smaller inhibitors. Therefore, there is a range of appropriate properties for fibril inhibition in the extremes of this parameter space, dependent upon the desired mechanism of inhibition. However, whilst the central region of the parameter space undergoes aggregate transformation and interference by inhibitors, it does not always indicate that these are inhibition mechanisms. Inhibitors with these properties can remodel fibril structures into alternative assembly morphologies.

5.3 Adding 400 Inhibitors

In this section, we repeat the parameter sweep described in the previous section. We complete a 2D parameter sweep of rod-sphere interaction strength E_{RS} in the range $0.5 \leq E_{RS}/E_{RR} \leq 128$ and sphere radius in the range $0.75 <= \sigma_S/\sigma_R <= 4$. In this case, we add 400 inhibitors to a steady-state system of amyloid fibres containing 800 monomers. This results in a stoichiometric ratio of 2:1 rods:spheres. We have chosen this intermediate stoichiometric ratio to explore the available parameter space further. In chapter 4, we used three stoichiometries, 1:20, 20:1 and 1:1. We found that simulations at stoichiometries of 1:20 rods:spheres were unlikely to inform drug design because the quantities required to reproduce the inhibition effect would be unrealistic and are highly likely to cause further issues within the brain. Therefore, we have chosen to explore regions of the parameter space which are much more likely to have implications for clinical use. We use two tools to describe the data. Firstly, we present a series of size distributions which indicate the degree of monomer and inhibitor content, respectively. (Figure 5.6.) These size distributions are also decorated with sample structures from the steady states of simulations with different parameters to provide insight into structural morphology and size. Figure 5.7 contains steady-state observations of clusters. Many of these observations are carried out with a cluster analysis which only includes rods. We do the calculations by this method to better compare with experiments and establish what is happening to the amyloid *in-vivo*. It is likely that disrupted assemblies such as these would become susceptible to the cell's degradation machinery. Often we can only observe the change in the fibril content rather than observe inhibitors directly. Each of the observations in figure 5.7 are displayed as heatmaps to highlight the change of behaviour as the sphere radius and rod-sphere interaction strength are varied.

The overall trends in rod cluster calculations (figure 5.7) are similar to those discussed in section 5.2. Again, we observe a decrease in mean rod cluster size and largest rod cluster size as we increase the rod-sphere interaction strength E_{RS}. In this sweep, the median rod cluster size (bottom left figure 5.7) follows the same trend as the mean rod cluster size. This differs from section 5.2, where there was no clear trend. In this case, the population ratio of rods:spheres is decreased from 1:8 to 1:2 due to an increase in inhibitor particles. The increase in inhibitors results in better-defined

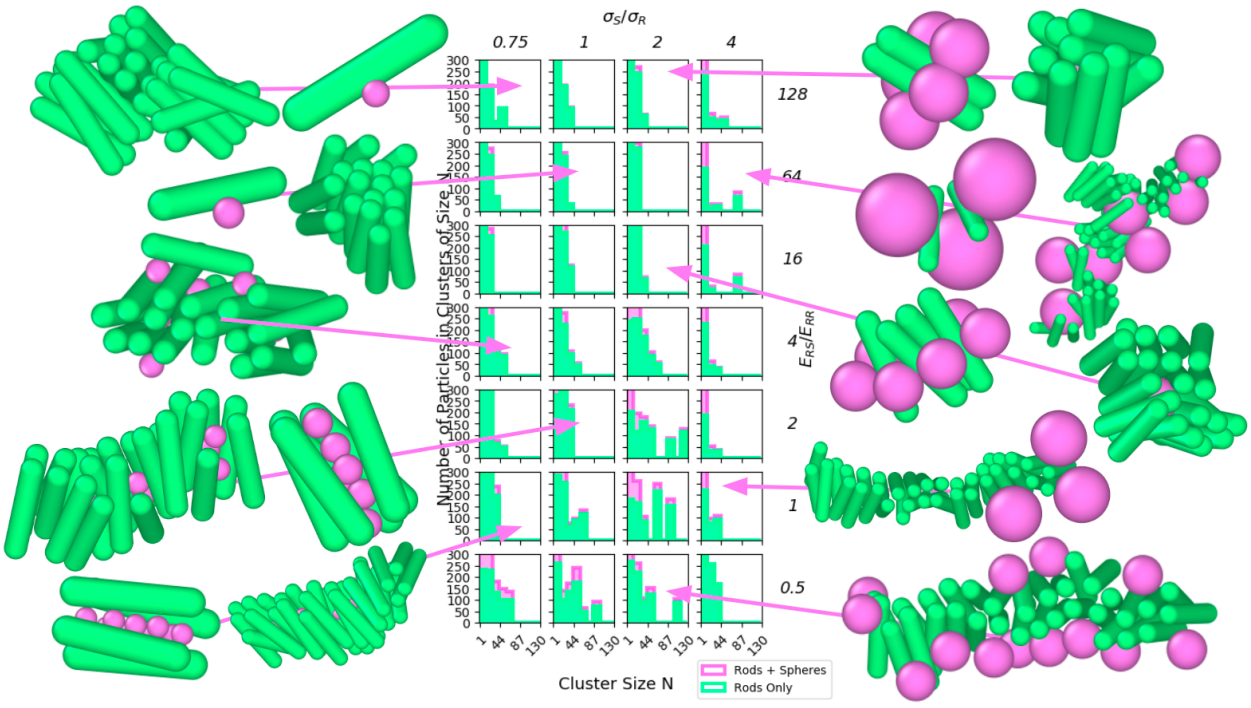

Figure 5.6: A series of steady-state size distributions across a 2D parameter sweep of aspect ratio AR and sphere radius σ_S where the population ratio of rods to spheres is 2:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_S/\sigma_R \leq 4$ where the rod radius is held constant at $\sigma_R = 1$nm. Rod sphere interaction strength E_{RS} varies along the y axis in the range $0.5 \leq E_{RS}/E_{RR} \leq 128$ where E_{RR} is the rod-rod interaction strength. Temperature is fixed at $T = 1k_BT = 310K$.

behaviours than in the previous section. The number of rod clusters has a larger peak when compared with the 8:1 case (section 5.2), suggesting the increased number of spheres here induces a stronger inhibition effect.

After we seed the fibrils, we add 400 inhibitors. We distribute these inhibitors randomly. During initialisation processes for the simulation to check for overlapping particles, this can result in the redistribution of some monomers amongst the volume of inhibitors. Therefore, a small amount of free monomer is available at the start of the simulation, resulting in the formation of small clusters similar to those found in chapter 4. These small clusters are shown in figure 5.6, which show a small number of rods bound together with the central region filled with spheres to produce a rod-like shape. These aggregates are present across all regions of the parameter space, indicated by the peak in the lower end of all distributions.

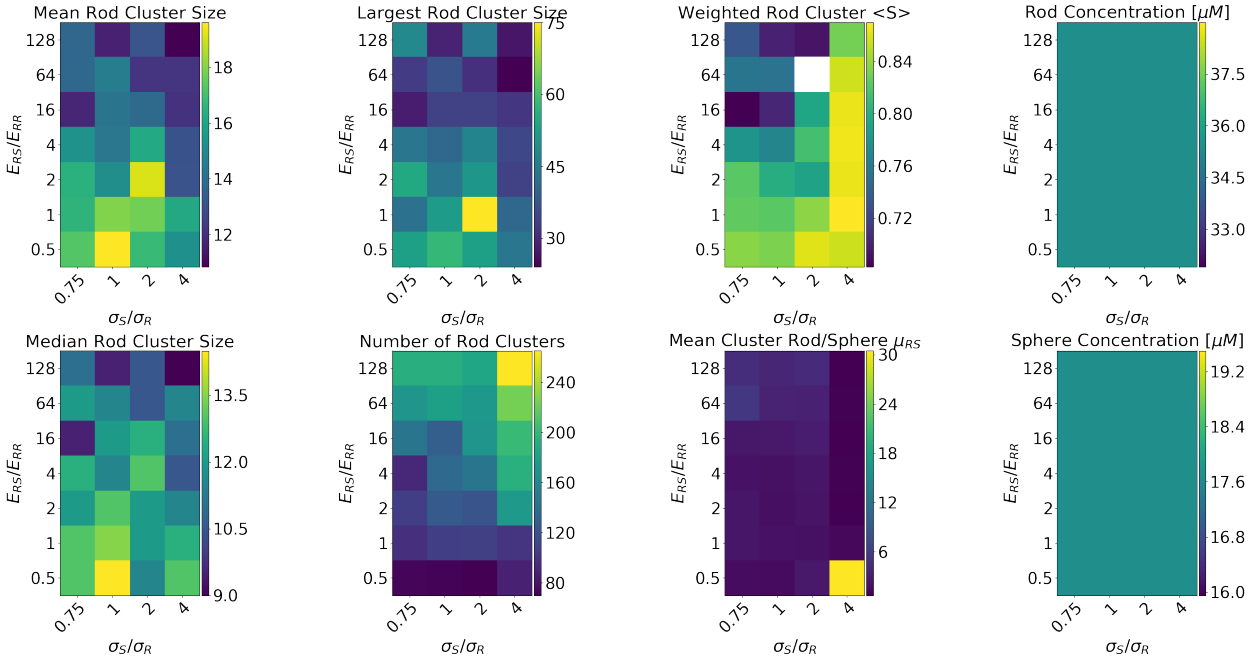

Figure 5.7: A series of steady-state observations of clusters in the form of heatmaps for a 2D parameter sweep of rod-sphere interaction strength and sphere radius, where the population ratio of rods to spheres is 2:1.Heatmaps for (top row, left to right) mean rod cluster size, largest rod cluster size, orientational rod order parameter weighted by rod cluster size and rod concentration, (bottom row, left to right) median rod cluster size, number of rod clusters, the ratio of mean rod cluster size and mean sphere cluster size μ_{RS} and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $0.5 \leq E_{RS}/E_{RR} \leq 128$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 <= \sigma_S/\sigma_R <= 4$ where the rod radius is held constant at $\sigma_R = 1$nm. Temperature is fixed at $T = 1k_BT = 310K$.

As discussed in section 5.2 and the previous chapter in section 4.3, the inhibition effect becomes more prominent with increased interaction strength E_{RS}. However, the mechanism is determined by the sphere radius. The behaviour here can be described by four regions in the following subsections.

Low Rod-Sphere Interaction Strength E_{RS}, Small Sphere radius σ_S Information about aggregates in this region are found in the bottom left of figure 5.6 and bottom left of individual heatmaps in figure 5.7. Here, small spheres fill up the space inside fibrillar aggregates. The mean rod cluster size and order parameter (figure 5.7) are decreased compared with the rod-only case. (Section 3.3.2) This indicates that the spheres disrupt the order of fibrillar clusters and cause monomers to dissociate or fibrils to break apart into smaller fibril-like aggregates. We conclude

that in this region of parameter space, the presence of inhibitors only serves to disrupt order. This suggests a low level of remodelling ability.

High Rod-Sphere Interaction Strength E_{RS}, Small Sphere radius σ_S Aggregates in this region undergo significant remodelling. The average rod order parameter S is significantly decreased (Top right figure 5.7) compared with the rod-only case, creating amorphous aggregates of rods surrounding small numbers of spheres. There is very little fibril content here, shown by the distributions in figure 5.6, indicating that the inhibitors have a significant effect on the aggregates when added to a pre-existing distribution. The free monomers bind at a 1:1 ratio to inhibitors in this region, resulting in many rod clusters. (Figure 5.7).

Low Rod-Sphere Interaction Strength E_{RS}, Large Sphere radius σ_S These larger spheres cannot fit inside fibrillar aggregates but instead, cover their surfaces at low interaction strengths E_{RS}. This surface coverage of fibrils would prevent secondary nucleation behaviours in *in-vivo* or *in-vitro* and, in turn, the prevention of toxic assemblies. We also observe fibril capping mechanisms. However, these are discussed in section 4.6.2 to be a specific case of the surface coverage mechanism within the scope of this model. However, it is still of interest because the assemblies remain very similar to the rod-only case here. The mean rod cluster size and the rod order parameter S are almost identical to the rod-only case. Therefore, only the elongation rate is affected, preventing further growth of fibrillar aggregates. These fibril capping aggregates are found in the 8:1 rod:sphere sweep in the previous section, but for higher interaction strengths. Interestingly, this behaviour is driven to occur at lower interaction strengths when there is a more minor difference in stoichiometries of rods and spheres.

High Rod-Sphere Interaction Strength E_{RS}, Large Sphere radius σ_S Similarly to the high interaction strength and small sphere radius (section 5.3), there are several examples of complete fibril remodelling in this region of the parameter space. (Top right of figure 5.6) The rod cluster calculations (top right of each heatmap, figure 5.7) indicate large numbers of rod clusters, a low mean rod cluster size and a high variation in order parameter S. Many sample aggregates contain local regions of ordered rods within larger, mixed aggregates of rods and spheres. The example

structures evidence this for high rod-sphere interaction strengths (top of figure 5.6), particularly for the largest sphere sizes $\sigma_S = 4\sigma_R$ where there is a decreased level of fibrillar aggregates. By inspection, the increased number of inhibitors has increased the ability of said inhibitors to remodel fibrillar aggregates into globular oligomeric structures. This is also the hypothesis of how G3P can remodel fibril structures[171]. It breaks fibrils apart and remodels each section into amorphous oligomers. In the case of this latter protein, the resulting species seem to lose toxicity, but this is not always the case.

Summary

The increased number of inhibitors increases the amount of inhibition behaviours observed. Some of the mechanisms we observed at high interaction strengths in the 8:1 case (For example, fibril capping) are now found at lower interaction strengths in this case. There is significantly more remodelling of fibril structures at this stoichiometry, which increases probability as the rod-sphere interaction strength E_{RS} increases. As mentioned previously, the sphere radius is the parameter which determines the inhibition mechanism. Small sphere radii ($\sigma_S = 0.75\sigma_R, \sigma_R$) can fit inside fibrillar aggregates, where the interaction strength determines the degree of remodelling. Conversely, larger sphere radii ($\sigma_S = 2\sigma_R, 4\sigma_R$) predominantly prevent further nucleation and elongation for lower interaction strengths. Comparatively to the lower radii, increasing the interaction strength increases the inhibitors' ability to remodel fibril structures. Therefore we conclude that at the highest interaction strengths, the largest radii are best able to completely change the aggregate morphology through various effects, which shows equivalence to experimental data of inhibitor candidates. A drug candidate which starts with these properties could be capable of complete fibril disassembly into amorphous oligomeric structures, which cell machinery may be better able to process and remove.

5.4 Adding 800 Inhibitors

Here, we repeat the parameter sweep carried out in the previous sections of sphere radius σ_S and rod-sphere inhibitor strength E_{RS}. In this case, we include 800 inhibitors modelled as isotropically interacting spheres, which results in a population ratio of rods:spheres of 1:1. We use the same

analysis techniques as in previous sections. (Figures 5.8 and 5.9). Figure 5.8 shows a series of steady-state size distributions of clusters that individually depict the rod and sphere content. The distributions are decorated with typical aggregates across all regions of the parameter space. Figure 5.9 shows a series of steady-state analyses of rod cluster averages which provide an overview of the changes in aggregate morphologies observed in different regions of the parameter space.

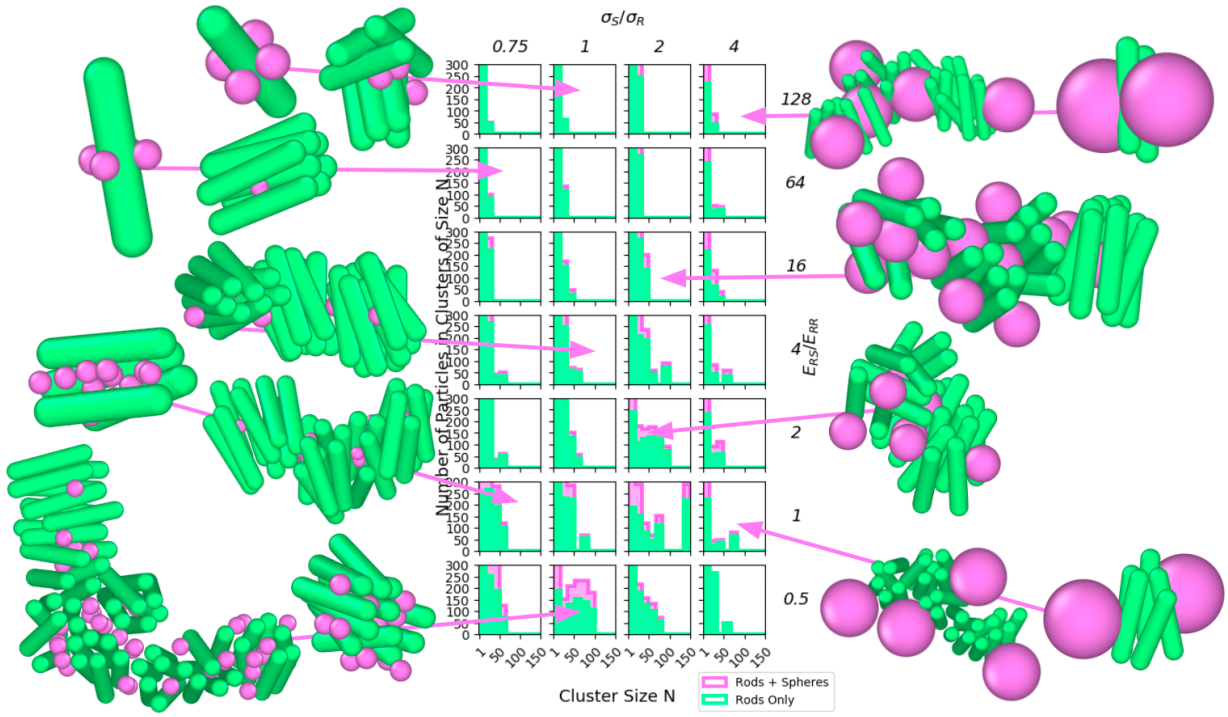

Figure 5.8: A series of steady-state size distributions across a 2D parameter sweep of aspect ratio AR and sphere radius σ_S, where the population ratio of rods to spheres is 1:1. Rod contributions to specific cluster sizes and sphere contributions are highlighted according to colour. The ratio of sphere radius to rod radius changes along the x axis from $0.75 \leq \sigma_S/\sigma_R \leq 4$ where the rod radius is held constant at $\sigma_R = 1$nm. Rod sphere interaction strength E_{RS} varies along the y axis in the range $0.5 \leq E_{RS}/E_{RR} \leq 128$ where E_{RR} is the rod-rod interaction strength. Temperature is fixed at $T = 1k_BT = 310K$.

In this parameter sweep, the inhibitory effects are more significant again when compared with previous sweeps. We observe the same broad trends in the heatmaps (Figure 5.9) where rod cluster averages decrease for increasing rod-sphere interaction strength E_{RS}. We also observe that the rod order parameter S is lower for the smallest sphere radii and highest interaction strengths. In this parameter space, we observe similar aggregate morphologies as for the previous sweeps (Figure 5.8), where again, the effects are more drastic due to the change in stoichiometric ratio to 1:1 rods:spheres.

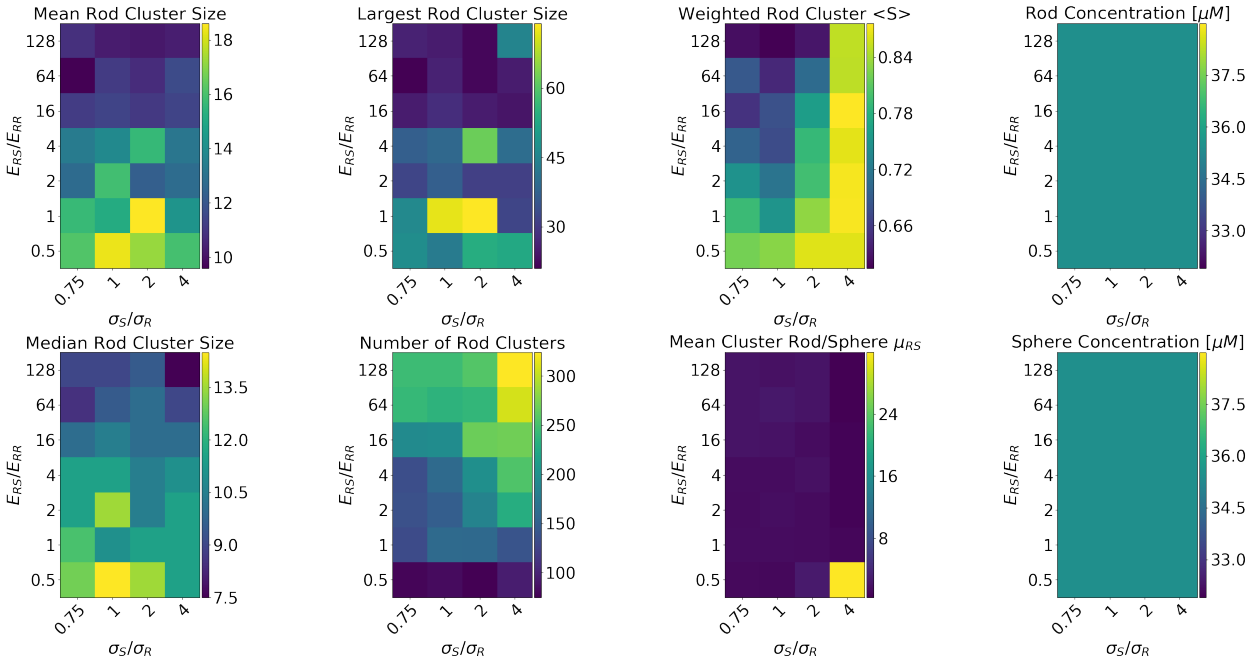

Figure 5.9: A series of steady-state observations of clusters in the form of heatmaps for a 2D parameter sweep of rod-sphere interaction strength and sphere radius, where the population ratio of rods to spheres is 1:1. Heatmaps for (top row, left to right) mean rod cluster size, largest rod cluster size, orientational rod order parameter weighted by rod cluster size and rod concentration, (bottom row, left to right) median rod cluster size, number of rod clusters, the ratio of mean rod cluster size and mean sphere cluster size μ_{RS} and sphere concentration. Rod sphere interaction strength varies along the x axes in the range $0.5 \leq E_{RS}/E_{RR} \leq 128$ where E_{RR} is the rod-rod interaction strength. The ratio of sphere radius to rod radius changes along the y axes from $0.75 <= \sigma_S/\sigma_R <= 4$ where the rod radius is held constant at $\sigma_R = 1$nm. Temperature is fixed at $T = 1k_BT = 310K$.

For small sphere radii ($\sigma_S = 0.75\sigma_R$ and $\sigma_S = \sigma_R$), we observe a variety of aggregate morphologies. The key observation for all of these is that the spheres fill the centres of the aggregates. At this stoichiometry, even the lowest interaction strengths result in extreme disruptions to fibril order, showing that even weakly binding inhibitors can have significant inhibition effects. For example, the specific case where $E_{RS} = 0.5E_{RR}$ and $\sigma_S = \sigma_R$ shows a distribution where there is a significant inhibitor content in large aggregate sizes. (Figure 5.8) The typical structures shown for this case indicate a high level of disruption to order visually, as well as a slight decrease in the calculated average order parameter S. (Figure 5.9) As the interaction strength increases, the average rod cluster size decreases, indicating a disruption to fibril content. This is further evidenced by the sample structures in figure 5.8 on the top left, as well as very narrow distributions. These disruption events

are common in our parameter space, and we hypothesise they lead to fibril breakage. However, we only observe full remodelling into amorphous structures at the highest interaction strengths, where fibrils are broken apart into much smaller structures which contain inhibitors at their centres. Behaviours along the interaction strength axis reflect the proposed behaviour of EGCG.[161–164]

In the larger sphere radii, the crowding effects will increase at this stoichiometry (1:1) when compared with the previous sections. The most significant inhibition behaviours occur in the top right of the parameter space where interaction strength is highest (Figure 5.8). In this region, we observe many mixed miscible aggregates. However, similarly to previous sections (5.2 and 5.3), these aggregates have local regions of order within them, resulting in high order parameters S (figure 5.9) which does not reflect the appearance of the aggregates. In the low interaction strength region, we observe surface binding of inhibitors to fibrillar clusters. As we increase the rod-sphere interaction strength, we see increasingly significant mixing until we observe alternating rod clusters and spheres within larger aggregates for $E_{RS} = 128E_{RR}$. The disruption and formation of mixed, amorphous structures are comparable to the proposed mechanism of G3P remodelling fibril structures.[171]

5.5 Discussion and Conclusion

Following the seeding of a distribution of fibrils, we found that in the case of no inhibitors, there were no significant changes to the distribution or the rod cluster averages typically used to measure aggregation behaviour in our simulations. We determined that we could proceed with inserting inhibitors in 2D parameter sweeps of rod-sphere interaction strengths E_{RS} and sphere radii σ_S at three stoichiometries relevant to drug treatments.

We have identified several mechanisms of inhibition which have been observed in other regions of the parameter space discussed in chapter 4. These include surface coverage, small molecule insertion and monomer sequestration. These have been discussed in detail in the summary of the previous chapter. (Section 4.6) However, in this case, we find that the inhibitors that drive these mechanisms can also disrupt pre-formed fibrils similarly. It may, therefore, also be the case that

fibrils form before undergoing transformation by inhibitors in chapter 4. Further analysis would need to be carried out considering the dynamics of the systems as well as the steady state aggregates.

We also observe remodelling of fibrils which has only been observed within this chapter, comparable to the proposed behaviour of the anti-amyloid protein G3P[171]. With this in mind, we can use the inhibitor parameters required to induce specific mechanisms to inform the drug design process.

Chapter 6

Discussion and Conclusion

In section 2.5, we presented a model of amyloid proteins where each monomer was to be represented by a spherocylinder with an attractive patch which spans 180° of the cylindrical portion of the particle. The spherocylinders interact via two potentials, one which prevents significant overlaps of particles and another which mimics the hydrogen bonding which occurs between beta strands in adjacent layers of fibrillar structures. The details of the model can be found in figure 2.3. The model aimed to model fibril formation and aggregate assembly for amyloid proteins and then disrupt these processes with inhibitory molecules. Our model describes an inhibitor as an isotropically interacting sphere interacting with the attractive patch on the spherocylinder.

In chapter 3, we search the parameter space of the spherocylinder model to find the most appropriate conditions and suitable parameter values for fibril assembly. We perform a series of 2D parameter sweeps of aspect ratio $AR = L/D$ and temperature T, rod-rod interaction strength E_{RR} and temperature, volume fraction VF and temperature, and finally, a sweep of the translational and rotational diffusion constants. We determined a proportional relationship between interaction strength and temperature in the 2D parameter sweep of rod-rod interaction strength and temperature described in section 3.3.1. The key region for large and numerous assemblies occurs when the interaction strength is approximately six times the thermal energy. i.e. $E_{RR} = 6k_BT$. We find that at physiological temperature $T_0 = 310K$, the interaction constant which produces a broad length distribution of fibrillar and fibril-like aggregates is $E_{RR} = 6k_BT_0$. This interaction strength corresponds to interaction energies of the order of the energy required for a small protein to fold.

We use this chosen parameter for future parameter sweeps.

In the aspect ratio and temperature sweep described in section 3.3.2, we found that aspect ratios up to $AR \approx 10$ can produce fibril-like species at physiological temperatures. Significant fibril assembly in our model has only been found to occur in a small range of physiological temperatures. In rescaling parameters where each spherocylinder represents a beta-strand region of an amyloid protein rather than a total monomer, we determine that different aspect ratios can represent lengths of beta-strand regions and, in turn, numbers of amino acids in a given beta-strand motif. Aspect ratio $AR = 4$ has the highest propensity to produce fibril assemblies, which corresponds to a number of common beta strand motifs in amyloid proteins, suggesting that even simple parameters such as shape can significantly impact the type of aggregation behaviour which occurs.

The concentration and temperature sweep in section 3.3.3 samples a series of volume fractions which the rods will occupy in a given simulation box. Changing this fraction, in turn, changes the concentration of rods or protein monomers. Amyloid-β experiments typically use concentrations in the range $1 - 100\mu M$, and our simulations were carried out in the range $50\mu M \leq c \leq 700\mu M$, stretching far beyond what is typically used experimentally. We use artificially high concentrations to reduce the computational resource and compute the waiting time for nucleation events to occur and for stable populations to emerge. We determined that 5% was an appropriate volume fraction for future parameterisation as this satisfies the trade-off between compute time and crowding events. The mean cluster size fluctuates more at higher concentrations than at lower concentrations. However, we still observe a plateau in mean cluster size, indicating that simulations reach a steady state in a physiological temperature range. Albeit much sooner than for lower concentrations.

Finally, the translational and rotational diffusion parameters are sampled around the values for amyloid-β supplied from the literature for similar models[3, 292], which are used for all further simulations. We find that the parameters from the literature are close to the region which produces the most significant fibril distribution. However, this does not imply the model is incorrect, simply that amyloid-β may not be the "ideal" amyloidogenic protein. Similarly, we could consider a different application to the model where each spherocylinder only represents a beta strand region of

each amyloid protein monomer. The diffusion parameters are currently representative of a complete amyloid monomer and would need to be adjusted as such.

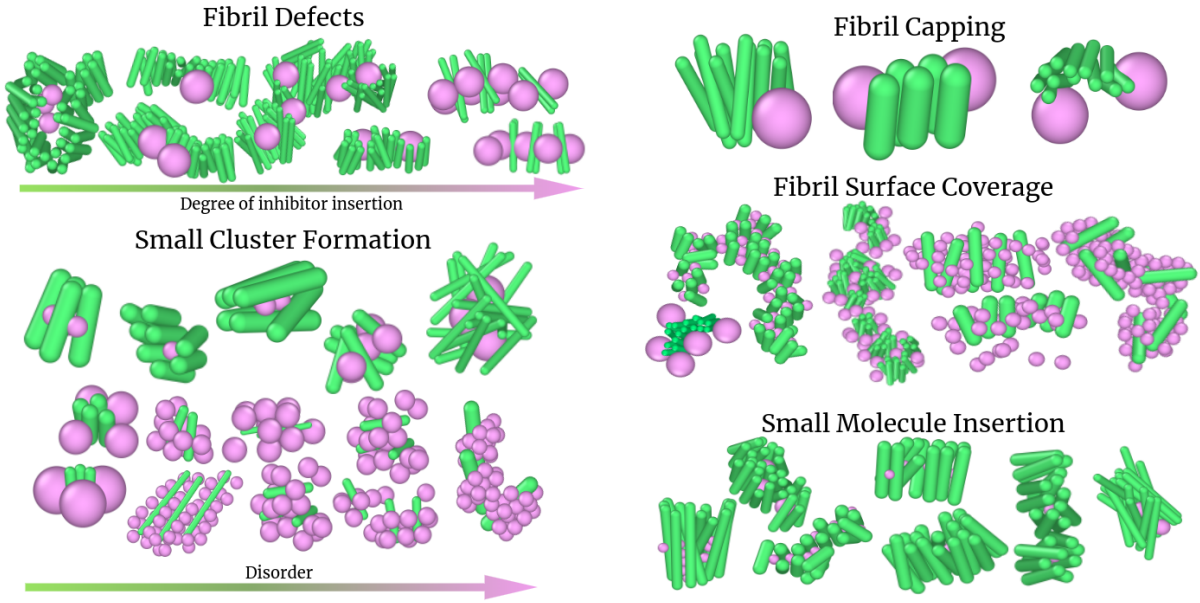

Figure 6.1: Categorisation of the different mixed aggregate types produced by exploring the rod-sphere parameter space.

Following the parameter sweeps in chapter 3, where we identified key regions of multidimensional parameter space where fibril assembly occurs, we inserted particles to disrupt and inhibit such processes. These inhibitors open the doors to more parameters, widening the phase space further. The multidimensional parameter space for inhibition has been explored based on parameters related to the properties of the amyloid monomer (spherocylinder or rod). We include the aspect ratio of the rod as one of our parameters as we are interested in disrupting assemblies of many different amyloid proteins with different shape morphologies. The other parameters of interest are rod-sphere interaction strength E_{RS}, sphere radius σ_S and the population ratio of rods to spheres. We fix the volume fraction of particles to $VF = 5\%$ to balance crowding events with reducing the computational expense of reaching a steady state. Therefore, the concentrations of rods and spheres vary implicitly as we change the shape parameters of the particles in a given simulation.

In chapter 4, we completed 2D pairwise parameter sweeps of the phase space (rod-sphere interaction strength and sphere radius in section 4.3, rod aspect ratio and rod-sphere interaction strength in

section 4.4 and finally, sphere radius and aspect ratio in section 4.5) were used to establish broad trends. Then aggregates were categorised into different morphologies. (Figure 6.1.) These different cluster morphologies include:

- Fibril defects: inhibitors with large radii (compared with monomer radius) cause fibrils to wrap around and, in specific cases, even begin to break apart in favour of forming additional bonds with the inhibitors rather than with other monomers. The top left of figure 6.1 highlights an increase in the severity of inhibitor insertion until we observe a miscible aggregate of rods and spheres. This increase in sphere insertion correlates with an increase in rod-sphere interaction strength E_{RS} and a relative increase in the population of inhibitors.

- Fibril Capping: inhibitors cover a fibril's end(s) by binding to the attractive patches on monomers at either end of a given fibril, preventing further elongation of fibril structures. For the case of our model, this is likely to be a specific case of the fibril defect mechanism where the spheres only bind to the end(s) of a fibril, which is potentially why it is not a commonly observed mechanism and often occurs in the same regions of parameter space where we observe fibril defects. Larger spheres are more likely to cause fibril capping where the sphere diameter is close to the length of the attractive patch on the monomer.

- Fibril surface coverage: where aggregate surfaces are covered in inhibitors, which we hypothesise would prevent secondary nucleation behaviours. Chaperone protein BRICHOS[150] may prevent secondary nucleation events in this fashion. In our parameter space, this behaviour is typically observed for spheres with radii close to the rod radii at stoichiometries of 1:1 or higher, such that there needs to be a greater than or an equal number of spheres than rods. These structures also often produce decreased order parameters. However, inhibitors like Brichos[150, 156] may have an affinity for the surface of fibrils that is not β-strand rich.

- Small Cluster Formation: These clusters are found in many regions of the parameter space. There are two categories of these assemblies: monomer-centric and inhibitor centric. The inhibitor centric tends to be comprised of small multimers of amyloid protein with a small number of inhibitors creating some disruption. Therefore, the order parameter is dependent on the sphere radius. Increasing the sphere radius increases the likelihood that rods will

reorder around the inhibitor(s). Conversely, the monomer-centric assemblies (bottom left, figure 6.1) are small numbers of protein monomers surrounded or coated in inhibitors. These aggregates are found in the same region of parameter space as the fibril surface coverage and, therefore, likely the same mechanism occurring on smaller multimer assemblies of amyloid protein.

- Small Molecule Insertion: small inhibitors typically of radius $\sigma_S \leq \sigma_R$ fill the centres of fibrils, disrupting order parameters and possibly causing fibril disassembly. This mechanism can occur at all stoichiometries. However, it is more likely to disrupt order with higher inhibitor quantities. There is experimental evidence[161] that small molecules can insert themselves cooperatively into fibrils, breaking H-bonds and causing fibril disassembly.

The different morphologies are the result of an extensive sampling of different regions of the multidimensional parameter space. These aggregates and mechanisms of fibril inhibition represent a system with an excess of amyloid protein monomers before the production of the distribution of fibrils. Therefore, the drug designing implications of chapter 4 would need to be applied to prevent the formation of further aggregates in an AD brain or as a preventative treatment for those likely to develop Alzheimer's disease and other neurodegenerative diseases. For example, patients with hereditary genetic mutations are known to increase the probability of protein misfolding and aggregation. In summary, the findings of chapter 4 are geared toward targeting protein monomers or interrupting the aggregation process whilst it is happening.

In chapter 5, we explore several preliminary parameter sweeps where we seed a fibril distribution and insert inhibitors to disrupt existing assemblies. This is a more appropriate representation of assemblies in a diseased brain. Disassembling or disrupting aggregates could potentially treat the disease and reverse aspects of the damage caused by the presence of aggregates rather than preventing a diseased brain from worsening (Chapter 4). We seed fibril distributions at concentrations used *in-vitro* ($\approx 35\mu M$) and add inhibitors at several stoichiometries and vary their radii σ_S and monomer-inhibitor interaction strengths E_{RS}. We find that various mechanisms occur, comparable to the results of chapter 4. However, not all of them are likely to prevent further aggregation behaviours.

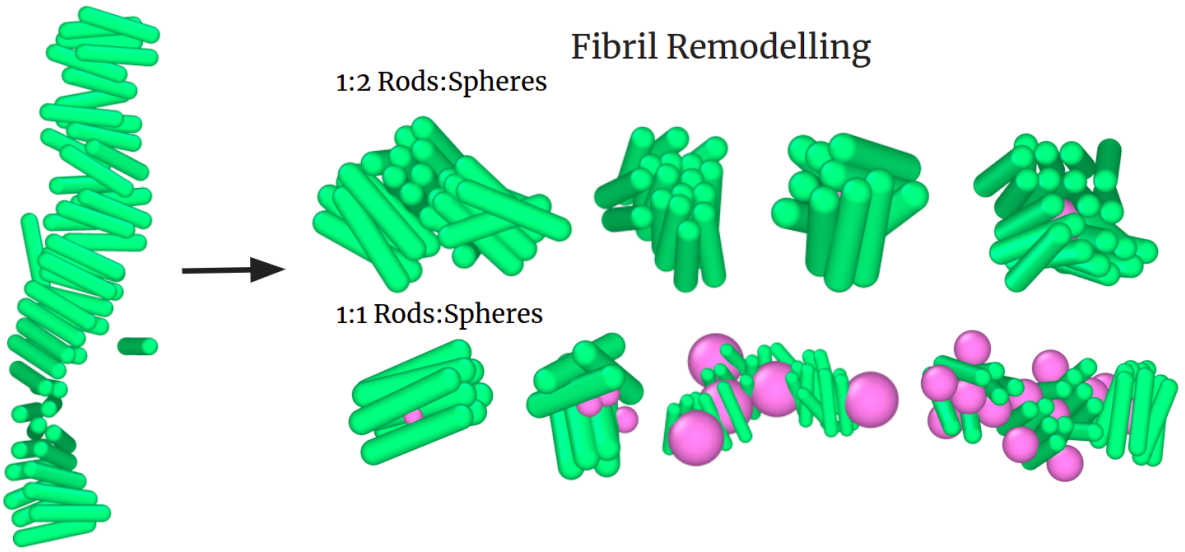

Figure 6.2: Sample aggregates formed via fibril remodelling from sections 5.3 and 5.4.

We also find regions of parameter space for stoichiometries of 1:2 rods:spheres or higher where the inhibitors can completely remodel fibril structures into globular mixed aggregates of monomers and inhibitors. Examples of such aggregates are shown in figure 6.2 and agree well with the proposed mechanisms of anti-amyloid protein G3P, which can remodel fibrils into amorphous clusters containing monomers of amyloid protein and inhibitors.

In summary, after extensive exploration of the available parameter space, we have found several regions of parameter space which are found to (a) inhibit the production of fibrillar aggregates. (b) Prevent further growth of existing aggregates and, in some cases, disaggregate amyloid assemblies. Inhibitor candidates validate the properties of these inhibitors in the literature, which are naturally occuring[160, 162, 163], chaperone proteins that assist with amyloid clearance[150], and proposed antibody drugs[22, 130, 139, 144]. We hope that the mechanisms in our simulations induced by varying shape parameters, interaction strength, and stoichiometry will inform future drug design to prevent specific kinetic processes in the amyloid formation processes. The information gathered here on steady states is extensive. However, there is still much to obtain from this inhibition model regarding the dynamics of aggregate formation.

6.1 Future Work

6.1.1 Changes to the Monomer Model

There are several different avenues to explore in the future work of this project. We can extend the original single patchy spherocylinder amyloid model to include more flexibility by modelling the amyloid protein monomer as two linked patchy spherocylinders. (Figure 6.3) We will gain further insight into the role of conformation in amyloid formation. It is already well-established that several conformations of a given amyloid protein can form fibrils. By extending the number of degrees of freedom a given amyloidogenic monomer has, we can observe a greater number of conformational states whilst retaining the system's simplistic and highly coarse-grained nature. This simulation model can also be validated against analytical work on conformational change in biopolymer formation[122].

Experimental and analytical data have highlighted the importance of secondary nucleation behaviours[292]. Therefore, a potential further development to the model could also encompass such behaviours and an explicit primary nucleation mechanism as have other simulation models.[4]. However, this model (Figure 6.3) will have more conformational degrees of freedom.

Further extensions of this work would include multiple fibrous species that are known to interact, create simulations that better model the *in-vivo* environment, and gain insight into the cooperative behaviours between such proteins and their effect on the amyloid aggregation process. Given that the model is highly coarse-grained, there is also scope to adapt the length scales of the model to describe other systems. One example is microtubule interactions with tau fibrils. Tau aggregates are seen to stabilise microtubules - exploring how this occurs using this model will lead to novel insights into the binding behaviour of these structures.

6.1.2 Additional Inhibitor Models

The current model we have proposed for the inhibitor is the most simplistic: an isotropically attractive interacting sphere. There are several proposed models that we have considered for future work shown in figure 6.4. These models propose attractive patches to better model the inhibitor

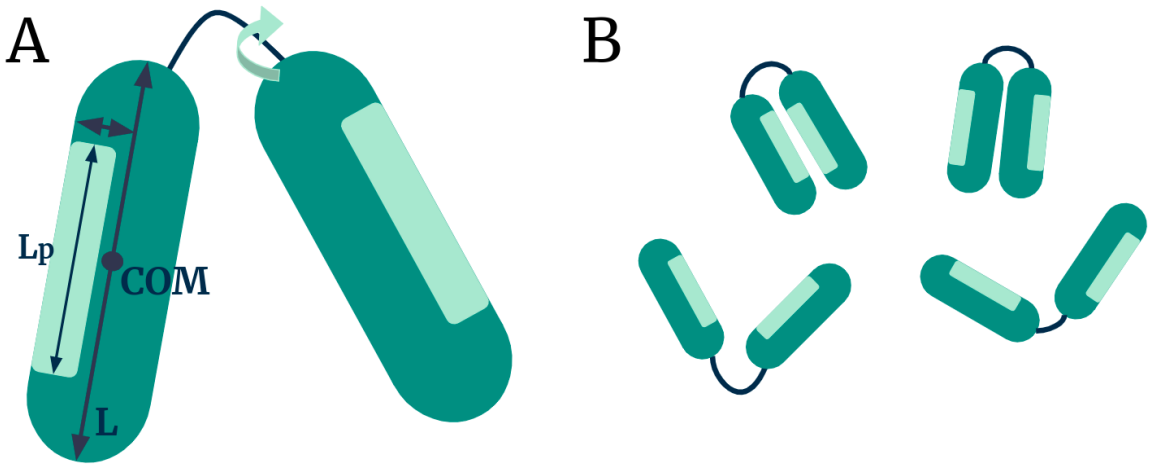

Figure 6.3: A potential new model design for the monomer of the amyloid-forming protein. (A) A schematic of two spherocylinders with a spring-like linker can rotate and change length. (B) Several example conformations of the proposed monomer

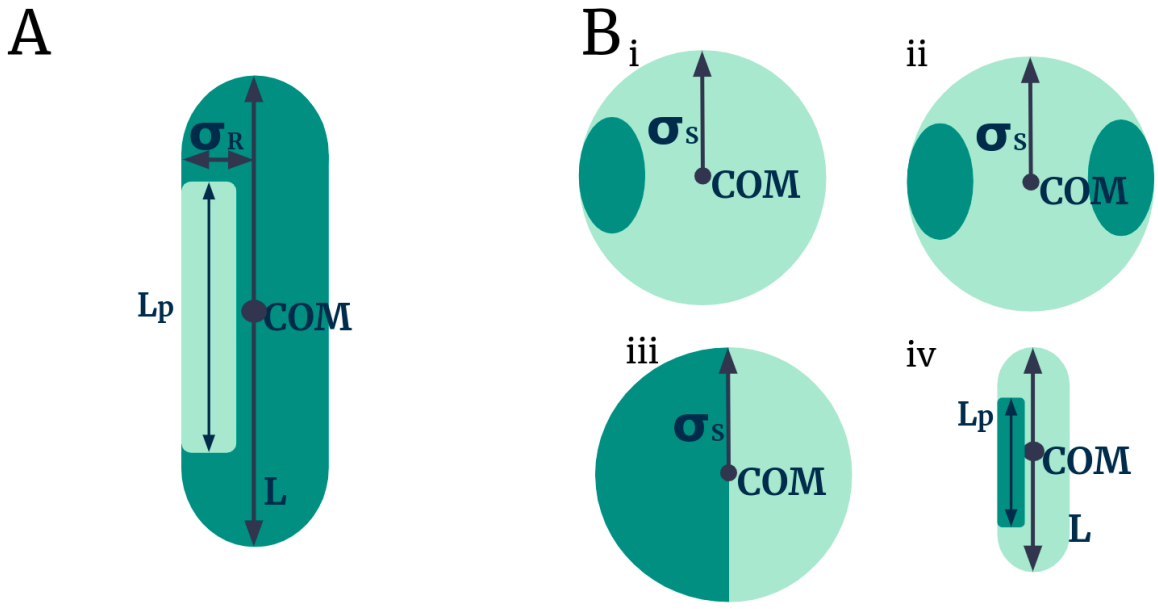

Figure 6.4: An potential new model design for the monomer of the amyloid-forming protein. (A) A schematic of the amyloid-forming monomer. (B) Several example conformations of the proposed inhibitor. (i) A sphere with a singular attractive binding site. (ii) A sphere with N attractive binding sites. (iii) A Janus particle inhibitor where one-half of the sphere interacts and one-half is either non-interacting or repulsive. (iv) A spherocylinder-shaped inhibitor.

with specific binding sites. Within these proposed models, we unlock several new tunable parameters, such as patch shape and size parameters, as well as the potential for creating a repulsive potential when rods are within the interaction distance of the non-patch area on the sphere. A final proposed model would be to entirely change the inhibitor's morphology to a second patchy spherocylinder with unique shape parameters to that of the amyloid protein monomer. Adding this functionality to the code would also allow for heterogeneity in the amyloid monomers, which has further applications for studying the amyloid formation process in more detail. With a model with such capability, further insight could be gained into existing kinetic studies which describe conformational changes along fibrils.

There is a wealth of literature on the aggregation behaviour of patchy colloids, but this becomes more limited under physiological (or close to) concentrations when mixed with patchy spherocylinders and other patchy colloids. Antibody-like inhibitors tend to bind tightly with a specific binding site(s) on a large body; an improvement to the current model would include the proposed patches to mimic specific binding sites. In the case of multiple binding patches, tuning the interaction strengths or potentials at each patch would also be possible. This massively increases the number of possible combinations of simulation conditions, increasing an already 5D parameter space into one which is N-dimensional.

Bibliography

[1]J. A. Hardy and G. A. Higgins, *Alzheimer's disease: The amyloid cascade hypothesis*, 1992.

[2]M. G. Spillantini, A. R. Crowther, R Jakes, N. J. Cairns, P. L. Lantos, and M Goedert, "Filamentous α-synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies", Neurosci. Lett. **251**, 205–208 (1998).

[3]A. Šarić, Y. C. Chebaro, T. P. J. Knowles, and D. Frenkel, "Crucial role of nonspecific interactions in amyloid nucleation", Proceedings of the National Academy of Sciences **111**, 17869–17874 (2014).

[4]R. Vácha and D. Frenkel, "Relation between Molecular Shape and the Morphology of Self-Assembling Aggregates: A Simulation Study", Biophysical Journal **101**, 1432–1439 (2011).

[5]M. Grazia Spillantini, R. Anthony Crowther, R. Jakes, N. J. Cairns, P. L. Lantos, and M. Goedert, "Filamentous α-synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies", Neuroscience Letters **251**, 205–208 (1998).

[6]L Pauling, H. A. Itano, S. J. Singer, and I. C. Wells, "Sickle cell anemia, a molecular disease", Science **110**, 543–548 (1949).

[7]D. J. Selkoe, "Alzheimer's Disease: Genes, Proteins, and Therapy", Physiological Reviews **81**, 741–766 (2001).

[8]M. Fändrich, "On the structural definition of amyloid fibrils and other polypeptide aggregates", Cellular and Molecular Life Sciences **64**, 2066–2078 (2007).

[9]U. Sengupta, A. N. Nilson, and R. Kayed, "The Role of Amyloid-β Oligomers in Toxicity, Propagation, and Immunotherapy.", EBioMedicine **6**, 42–49 (2016).

[10]M. Serra-Batiste, M. Ninot-Pedrosa, M. Bayoumi, M. Gairí, G. Maglia, and N. Carulla, "Aβ42 assembles into specific β-barrel pore-forming oligomers in membrane-mimicking environments.", Proceedings of the National Academy of Sciences of the United States of America **113**, 10866–71 (2016).

[11]B. Mroczko, M. Groblewska, A. Litman-Zawadzka, J. Kornhuber, and P. Lewczuk, "Amyloid β oligomers (AβOs) in Alzheimer's disease", Journal of Neural Transmission **125**, 177–191 (2018).

[12]A. Burns and S. Iliffe, "Dementia", BMJ: British Medical Journal **338**, 405–409 (2009).

[13]H. Hippius and G. Neundörfer, "The discovery of Alzheimer's disease.", Dialogues in clinical neuroscience **5**, 101–8 (2003).

[14]I. W. Hamley, "The Amyloid Beta Peptide: A Chemist's Perspective. Role in Alzheimer's and Fibrillization", Chemical Reviews **112**, 5147–5192 (2012).

[15]M. R. Brown, S. E. Radford, and E. W. Hewitt, "Modulation of β-Amyloid Fibril Formation in Alzheimer's Disease by Microglia and Infection", Frontiers in Molecular Neuroscience **13**, 228 (2020).

[16]D. W. Dickson, H. A. Crystal, L. A. Mattiace, D. M. Masur, A. D. Blau, P. Davies, S. H. Yen, and M. K. Aronson, "Identification of normal and pathological aging in prospectively studied nondemented elderly humans", Neurobiology of aging **13**, 179–189 (1992).

[17]K. M. Rodrigue, K. M. Kennedy, and D. C. Park, "Beta-Amyloid Deposition and the Aging Brain", Neuropsychology review **19**, 436 (2009).

[18]J. D. Sipe and A. S. Cohen, *Review: History of the amyloid fibril*, 2000.

[19]A. M. Morris, M. A. Watzky, and R. G. Finke, "Protein aggregation kinetics, mechanism, and curve-fitting: A review of the literature", Biochimica et Biophysica Acta - Proteins and Proteomics **1794**, 375–397 (2009).

[20]J Straub and D Thirumalai, "Toward a Molecular Theory of Early and Late Events in Monomer to Amyloid Fibril Formation", Annu. Rev. Phys. Chem. **67**, 437–463 (2011).

[21]P. N. Institute, *Brain Imaging for Alzheimer's & Dementia — Pacific Brain Health Center.*

[22]C. H. van Dyck, C. J. Swanson, P. Aisen, R. J. Bateman, C. Chen, M. Gee, M. Kanekiyo, D. Li, L. Reyderman, S. Cohen, L. Froelich, S. Katayama, M. Sabbagh, B. Vellas, D. Watson, S. Dhadda, M. Irizarry, L. D. Kramer, and T. Iwatsubo, "Lecanemab in Early Alzheimer's Disease.", The New England journal of medicine (2022) 10.1056/NEJMOA2212948/SUPPL_FILE/NEJMOA2212948_DATA-SHARING.PDF.

[23]J. S. Birks and J. Grimley Evans, "Rivastigmine for Alzheimer's disease", in *Cochrane database of systematic reviews*, edited by J. S. Birks (John Wiley & Sons, Ltd, Chichester, UK, 2015).

[24]J. S. Schneider, E. Y. Pioli, Y. Jianzhong, Q. Li, and E. Bezard, "Effects of memantine and galantamine on cognitive performance in aged rhesus macaques", Neurobiology of Aging **34**, 1126–1132 (2013).

[25]J Ojaimi, C. L. Masters, C McLean, K Opeskin, P McKelvie, and E Byrne, "Irregular distribution of cytochrome c oxidase protein subunits in aging and Alzheimer's disease.", Annals of neurology **46**, 656–60 (1999).

[26]C. Haass and D. J. Selkoe, "Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide", Nature Reviews Molecular Cell Biology **8**, 101–112 (2007).

[27]M. Goedert, M. G. Spillantini, and R. A. Crowther, "Tau Proteins and Neurofibrillary Degeneration", Brain Pathology **1**, 279–286 (1991).

[28]D. W. Cleveland, S.-Y. Hwo, and M. W. Kirschner, "Physical and chemical properties of purified tau factor and the role of tau in microtubule assembly", Journal of Molecular Biology **116**, 227–247 (1977).

[29]M. B. Pepys, "Amyloidosis", Annual review of medicine **57**, 223–241 (2006).

[30]M. Meyer-Luehmann, J. Coomaraswamy, T. Bolmont, S. Kaeser, C. Schaefer, E. Kilger, A. Neuenschwander, D. Abramowski, P. Frey, A. L. Jaton, J. M. Vigouret, P. Paganetti, D. M. Walsh, P. M. Mathews, J. Ghiso, M. Staufenbiel, L. C. Walker, and M. Jucker, "Exogenous induction of cerebral beta-amyloidogenesis is governed by agent and host", Science (New York, N.Y.) **313**, 1781–1784 (2006).

[31]C. Goldsbury, P. Frey, V. Olivieri, U. Aebi, and S. A. Müller, "Multiple assembly pathways underlie amyloid-beta fibril polymorphisms", Journal of molecular biology **352**, 282–298 (2005).

[32]A. L. Gharibyan, V. Zamotin, K. Yanamandra, O. S. Moskaleva, B. A. Margulis, I. A. Kostanyan, and L. A. Morozova-Roche, "Lysozyme amyloid oligomers and fibrils induce cellular death via different apoptotic/necrotic pathways", Journal of molecular biology **365**, 1337–1349 (2007).

[33]F. Kametani and M. Hasegawa, "Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease.", Frontiers in neuroscience **12**, 25 (2018).

[34]A. Takashima, K. Noguchi, K. Sato, T. Hoshino, and K. Imahori, "Tau protein kinase I is essential for amyloid beta-protein-induced neurotoxicity", Proceedings of the National Academy of Sciences of the United States of America **90**, 7789–7793 (1993).

[35]M. Hoshi, M. Sato, S. Matsumoto, A. Noguchi, K. Yasutake, N. Yoshida, and K. Sato, "Spherical aggregates of -amyloid (amylospheroid) show high neurotoxicity and activate tau protein kinase I/glycogen syn-

thase kinase-3", Proceedings of the National Academy of Sciences **100**, 6370–6375 (2003).

[36]J. Busciglio, A. Lorenzo, J. Yeh, and B. A. Yankner, "Beta-amyloid fibrils induce tau phosphorylation and loss of microtubule binding", Neuron **14**, 879–888 (1995).

[37]C. Patterson, J. W. Feightner, A. Garcia, G.-Y. R. Hsiung, C. MacKnight, and A. D. Sadovnick, "Diagnosis and treatment of dementia: 1. Risk assessment and primary prevention of Alzheimer disease", Canadian Medical Association Journal **178**, 548–556 (2008).

[38]A. Palmeri, R. Ricciarelli, W. Gulisano, D. Rivera, C. Rebosio, E. Calcagno, M. R. Tropea, S. Conti, U. Das, S. Roy, M. A. Pronzato, O. Arancio, E. Fedele, and D. Puzzo, "Amyloid-β Peptide Is Needed for cGMP-Induced Long-Term Potentiation and Memory", The Journal of neuroscience : the official journal of the Society for Neuroscience **37**, 6926–6937 (2017).

[39]A. Garcia-Osta and C. M. Alberini, "Amyloid beta mediates memory formation", Learning & Memory **16**, 267 (2009).

[40]C. Priller, T. Bauer, G. Mitteregger, B. Krebs, H. A. Kretzschmar, and J. Herms, "Neurobiology of Disease Synapse Formation and Function Is Modulated by the Amyloid Precursor Protein", 10.1523/JNEUROSCI.1450-06.2006.

[41]M. Citron, T. Oltersdorf, C. Haass, L. McConlogue, A. Y. Hung, P. Seubert, C. Vigo-Pelfrey, I. Lieberburg, and D. J. Selkoe, "Mutation of the β-amyloid precursor protein in familial Alzheimer's disease increases β-protein production", Nature **360**, 672–674 (1992).

[42]R. Delabio, L. Rasmussen, I. Mizumoto, G.-A. Viani, E. Chen, J. Villares, I.-B. Costa, G. Turecki, S. A. Linde, M. C. Smith, and S.-L. Payão, "PSEN1 and PSEN2 gene expression in Alzheimer's disease brain: a new approach.", Journal of Alzheimer's disease : JAD **42**, 757–60 (2014).

[43]J. D. Doecke, V. Pérez-Grijalba, N. Fandos, C. Fowler, V. L. Villemagne, C. L. Masters, P. Pesini, and M. Sarasa, "Total Aβ42/Aβ40 ratio in plasma predicts amyloid-PET status, independent of clinical AD diagnosis", Neurology **94**, e1580–e1591 (2020).

[44]G. G. Glenner and C. W. Wong, "Alzheimer's disease and Down's syndrome: Sharing of a unique cerebrovascular amyloid fibril protein", Biochemical and Biophysical Research Communications **122**, 1131–1135 (1984).

[45]S Heber, J Herms, V Gajic, J Hainfellner, A Aguzzi, T Rülicke, H von Kretzschmar, C von

Koch, S Sisodia, P Tremml, H. P. Lipp, D. P. Wolfer, and U Müller, "Mice with combined gene knock-outs reveal essential and partially redundant functions of amyloid precursor protein family members.", The Journal of neuroscience : the official journal of the Society for Neuroscience **20**, 7951–63 (2000).

[46]A. Buniello, J. A. Macarthur, M. Cerezo, L. W. Harris, J. Hayhurst, C. Malangone, A. McMahon, J. Morales, E. Mountjoy, E. Sollis, D. Suveges, O. Vrousgou, P. L. Whetzel, R. Amode, J. A. Guillen, H. S. Riat, S. J. Trevanion, P. Hall, H. Junkins, P. Flicek, T. Burdett, L. A. Hindorff, F. Cunningham, and H. Parkinson, "The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019", Nucleic acids research **47**, D1005–D1012 (2019).

[47]M. Nistor, M. Don, M. Parekh, F. Sarsoza, M. Goodus, G. Lopez, C. Kawas, J. Leverenz, E. Doran, I. Lott, M. Hill, and E. Head, "Alpha- and beta-secretase activity as a function of age and beta-amyloid in Down syndrome and normal brain", Neurobiology of Aging **28**, 1493–1506 (2007).

[48]I. T. Lott and E. Head, "Alzheimer disease and Down syndrome: factors in pathogenesis", Neurobiology of Aging **26**, 383–389 (2005).

[49]F. Su, F. Bai, H. Zhou, and Z. Zhang, "Microglial toll-like receptors and Alzheimer's disease", Brain, behavior, and immunity **52**, 187–198 (2016).

[50]J. Labbadia and R. I. Morimoto, "The biology of proteostasis in aging and disease", Annual review of biochemistry **84**, 435–464 (2015).

[51]D. C. Bode, H. F. Stanyon, T. Hirani, M. D. Baker, J. Nield, and J. H. Viles, "Serum Albumin's Protective Inhibition of Amyloid-β Fiber Formation Is Suppressed by Cholesterol, Fatty Acids and Warfarin", Journal of Molecular Biology **430**, 919–934 (2018).

[52]Z. X. Yao and V. Papadopoulos, "Function of beta-amyloid in cholesterol transport: a lead to neurotoxicity", FASEB journal : official publication of the Federation of American Societies for Experimental Biology **16**, 1677–1679 (2002).

[53]K. A. Dill, S. B. Ozkan, M. S. Shell, and T. R. Weikl, "The protein folding problem.", Annual review of biophysics **37**, 289–316 (2008).

[54]P. G. Wolynes, "Water Mediated Interactions in Protein Folding and Structure", Biophysical Journal **100**, 5a (2011).

[55]M. Karplus, "The Levinthal paradox: yesterday and today", Folding and Design **2**, S69–S75 (1997).

[56]Biorender, *Reprinted from Protein Structure, by BioRender.com*, 2022.

[57]S. D. Pickett and M. J. Sternberg, "Empirical scale of side-chain conformational entropy in protein folding", Journal of molecular biology **231**, 825–839 (1993).

[58]A. J. Doig' And, M. J. E. Sternberg', and A. J. Doig, "Side-chain conformational entropy in protein folding", Protein Science **4**, 2247–2251 (1995).

[59]C. M. Dobson, "Protein folding and misfolding", Nature 2003 426:6968 **426**, 884–890 (2003).

[60]S. I. A. Cohen, M. Vendruscolo, M. E. Welland, C. M. Dobson, E. M. Terentjev, and T. P. J. Knowles, "Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments", The Journal of Chemical Physics **135**, 065105 (2011).

[61]L. C. Serpell and J. M. Smith, "Direct visualisation of the β-sheet structure of synthetic Alzheimer's amyloid", Journal of Molecular Biology **299**, 225–231 (2000).

[62]M. Sunde, L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys, and C. C. Blake, "Common core structure of amyloid fibrils by synchrotron X-ray diffraction", Journal of molecular biology **273**, 729–739 (1997).

[63]S. Kumar and J. B. Udgaonkar, "Conformational Conversion May Precede or Follow Aggregate Elongation on Alternative Pathways of Amyloid Protofibril Formation", J. Mol. Biol. **385**, 1266–1276 (2009).

[64]C. A. Hunter and J. K. Sanders, "The Nature of π-π Interactions", Journal of the American Chemical Society **112**, 5525–5534 (1990).

[65]L. Gremer, D. Schölzel, C. Schenk, E. Reinartz, J. Labahn, R. B. G. Ravelli, M. Tusche, C. Lopez-Iglesias, W. Hoyer, H. Heise, D. Willbold, and G. F. Schröder, "Fibril structure of amyloid-β(1-42) by cryo-electron microscopy.", Science (New York, N.Y.) **358**, 116–119 (2017).

[66]M. S. Rubel, S. A. Fedotov, A. V. Grizel, J. V. Sopova, O. A. Malikova, Y. O. Chernoff, and A. A. Rubel, "Functional Mammalian Amyloids and Amyloid-Like Proteins", Life 2020, Vol. 10, Page 156 **10**, 156 (2020).

[67]K. L. Osborne, M. Bachmann, and B. Strodel, "Thermodynamic analysis of structural transitions during GNNQQNY aggregation", Proteins: Structure, Function and Bioinformatics **81**, 1141–1155 (2013).

[68]A. S. Reddy, M. Chopra, and J. J. De Pablo, "GNNQQNYInvestigation of Early Steps during Amyloid Formation", Biophysical Journal **98**, 1038 (2010).

[69]J. Fiori, M. Naldib, and V. Andrisano, "Mass spectrometry as an efficient tool for the characterization of amyloid β peptide 25-35 self-assembly species in aggregation and inhibition studies", European journal of mass spectrometry (Chichester, England) **19**, 483–490 (2013).

[70]T. P. J. Knowles, M. Vendruscolo, and C. M. Dobson, "The amyloid state and its association with protein misfolding diseases.", Nature reviews. Molecular cell biology **15**, 384–96 (2014).

[71]J. Hardy and D. J. Selkoe, "The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.", Science (New York, N.Y.) **297**, 353–6 (2002).

[72]D. Eisenberg and M. Jucker, "The Amyloid State of Proteins in Human Diseases", Cell **148**, 1188–1203 (2012).

[73]A. C. Muntau, J. Leandro, M. Staudigl, F. Mayer, and S. W. Gersting, "Innovative strategies to treat protein misfolding in inborn errors of metabolism: pharmacological chaperones and proteostasis regulators", Journal of Inherited Metabolic Disease **37**, 505–523 (2014).

[74]C. Lendel, M. Bjerring, A. Dubnovitsky, R. T. Kelly, A. Filippov, O. N. Antzutkin, N. C. Nielsen, and T. Härd, "A hexameric peptide barrel as building block of amyloid-β protofibrils", Angewandte Chemie - International Edition **53**, 12756–12760 (2014).

[75]B. Urbanc, L. Cruz, F. Ding, D. Sammond, S. Khare, S. V. Buldyrev, H. E. Stanley, and N. V. Dokholyan, "Molecular dynamics simulation of amyloid β dimer formation", Biophysical Journal **87**, 2310–2321 (2004).

[76]G. Bellesia and J. E. Shea, "Diversity of kinetic pathways in amyloid fibril formation", Journal of Chemical Physics (2009) 10.1063/1.3216103.

[77]W. Zheng, M.-Y. Tsai, M. Chen, and P. G. Wolynes, "Exploring the aggregation free energy landscape of the amyloid-β protein (140)", Proceedings of the National Academy of Sciences (2016) 10.1073/pnas.1612362113.

[78]S. I. A. Cohen, R. Cukalevski, T. C. T. Michaels, A Šarić, M. Törnquist, M. Vendruscolo, C. M. Dobson, A. K. Buell, T. P. J. Knowles, and S. Linse, "Distinct thermodynamic signatures of oligomer generation in the aggregation of the amyloid-β peptide", Nat. Chem. **10**, 523–531 (2018).

[79]R Tycko, "Solid state NMR studies of amyloid fibril structure", Annu. Rev. Phys. Chem. **62**, 279–299 (2011).

[80]W. Qiang, W.-M. Yau, Y. Luo, M. P. Mattson, and R. Tycko, "Antiparallel β-sheet architecture in Iowa-mutant β-amyloid fibrils.", Proceedings of the National Academy of Sciences of the United States of America **109**, 4443–8 (2012).

[81]Y. Xiao, B. Ma, D. McElheny, S. Parthasarathy, F. Long, M. Hoshi, R. Nussinov, and Y. Ishii, "Aβ(142) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease", Nature Structural & Molecular Biology **22**, 499–505 (2015).

[82]M. A. Wälti, F. Ravotti, H. Arai, C. G. Glabe, J. S. Wall, A. Böckmann, P. Güntert, B. H. Meier, and R. Riek, "Atomic-resolution structure of a disease-relevant Aβ(1-42) amyloid fibril.", Proceedings of the National Academy of Sciences of the United States of America **113**, E4976–84 (2016).

[83]M. T. Colvin, R. Silvers, Q. Z. Ni, T. V. Can, I. Sergeyev, M. Rosay, K. J. Donovan, B. Michael, J. Wall, S. Linse, and R. G. Griffin, "Atomic Resolution Structure of Monomorphic Aβ(42) Amyloid Fibrils", Journal of the American Chemical Society **138**, 9663–9674 (2016).

[84]M. Biancalana and S. Koide, "Molecular mechanism of Thioflavin-T binding to amyloid fibrils.", Biochimica et biophysica acta **1804**, 1405–12 (2010).

[85]C. Xue, T. Y. Lin, D. Chang, and Z. Guo, "Thioflavin T as an amyloid dye: fibril quantification, optimal concentration and effect on aggregation.", Royal Society open science **4**, 160696 (2017).

[86]F Oosawa and S Asakura, "Thermodynamics of the Polymerization of Protein", (1975).

[87]B. Linse and S. Linse, "Monte Carlo simulations of protein amyloid formation reveal origin of sigmoidal aggregation kinetics", Molecular BioSystems **7**, 2296–2303 (2011).

[88]A. Irbäck, S. Jónsson, N. Linnemann, B. Linse, and S. Wallin, "Aggregate geometry in amyloid fibril nucleation", Physical Review Letters **110**, 1–4 (2013).

[89]S. Eugïne, W. F. Xue, P. Robert, and M. Doumic, "Insights into the variability of nucleated amyloid polymerization by a minimalistic model of stochastic protein assembly", Journal of Chemical Physics **144** (2016) 10.1063/1.4947472.

[90]H. Gutfreund, *Kinetics for the life sciences : receptors, transmitters, and catalysts* (Cambridge University Press, 1995), p. 346.

[91]M. Wolff, D. Unuchek, B. Zhang, V. Gordeliy, D. Willbold, and L. Nagel-Steger, "Amyloid

β oligomeric species present in the lag phase of amyloid formation", PLoS ONE **10**, 1–14 (2015).

[92]E. H. Koo, P. T. Lansbury, J. W. Kelly, and J. W. Kelly, "Amyloid diseases: abnormal protein aggregation in neurodegeneration.", Proceedings of the National Academy of Sciences of the United States of America **96**, 9989–90 (1999).

[93]J. H. Come, P. E. Fraser, and P. T. Lansbury, "A kinetic model for amyloid formation in the prion diseases: importance of seeding.", Proceedings of the National Academy of Sciences of the United States of America **90**, 5959–63 (1993).

[94]P. T. Lansbury, "Evolution of amyloid: what normal protein folding may tell us about fibrillogenesis and disease.", Proceedings of the National Academy of Sciences of the United States of America **96**, 3342–4 (1999).

[95]D. Kashchiev, *Nucleation : basic theory with applications* (Butterworth Heinemann, 2000), p. 529.

[96]F. A. Ferrone, J Hofrichter, H. R. Sunshine, and W. A. Eaton, "Kinetics studies on photolysis-induced gelation of sickle-cell hemoglobin suggest a new mechanism.", Biophys. J. **32**, 361–380 (1980).

[97]F. A. Ferrone, J. Hofrichter, and W. A. Eaton, "Kinetics of sickle hemoglobin polymerization: II. A double nucleation mechanism", Journal of Molecular Biology **183**, 611–631 (1985).

[98]M. Bishop and F. Ferrone, "Kinetics of nucleation-controlled polymerization. A perturbation treatment for use with a secondary pathway", Biophysical Journal **46**, 631–644 (1984).

[99]S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, "Nucleated polymerization with secondary pathways. III. Equilibrium behavior and oligomer populations", The Journal of Chemical Physics **135**, 065107 (2011).

[100]S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, "Nucleated polymerization with secondary pathways. II. Determination of self-consistent solutions to growth processes described by non-linear master equations", The Journal of Chemical Physics **135**, 065106 (2011).

[101]G. A. Garcia, S. I. A. Cohen, C. M. Dobson, and T. P. J. Knowles, "Nucleation-conversion-polymerization reactions of biological macromolecules with prenucleation clusters", Phys. Rev. E **89**, 32712 (2014).

[102]T. P. J. Knowles, C. A. Waudby, G. L. Devlin, S. I. A. Cohen, A. Aguzzi, M. Vendruscolo, E. M. Terentjev, M. E. Welland, and C. M. Dobson, "An Analytical Solution to the Kinetics of Breakable Filament Assembly", Science **326**, 1533–1537 (2009).

[103]T. C. T. Michaels, G. A. Garcia, and T. P. J. Knowles, "Asymptotic solutions of the Oosawa model for the length distribution of biofilaments", Proc. Natl. Acad. Sci. U. S. A. **140**, 194906 (2014).

[104]T. C. Michaels, G. A. Garcia, and T. P. Knowles, "Asymptotic solutions of the Oosawa model for the length distribution of biofilaments", Journal of Chemical Physics **140** (2014) 10.1063/1.4875897.

[105]T. C. Michaels, H. W. Lazell, P. Arosio, and T. P. Knowles, "Dynamics of protein aggregation and oligomer formation governed by secondary nucleation", Journal of Chemical Physics **143** (2015) 10.1063/1.4927655.

[106]T. C. Michaels, P. Yde, J. C. Willis, M. H. Jensen, D. Otzen, C. M. Dobson, A. K. Buell, and T. P. Knowles, "The length distribution of frangible biofilaments", Journal of Chemical Physics **143** (2015) 10.1063/1.4933230.

[107]T. C. T. Michaels, A. J. Dear, and T. P. J. Knowles, "Scaling and dimensionality in the chemical kinetics of protein filament formation", Int. Rev. Phys. Chem. **35**, 679–703 (2016).

[108]T. C. Michaels, S. I. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. Knowles, "Hamiltonian Dynamics of Protein Filament Formation", Physical Review Letters **116**, 038101 (2016).

[109]A. J. Dear, T. C. Michaels, and T. P. Knowles, "Dynamics of heteromolecular filament formation", Journal of Chemical Physics **145** (2016) 10.1063/1.4966571.

[110]L. Hong and W. A. Yong, "Simple moment-closure model for the self-assembly of breakable amyloid filaments", Biophysical Journal **104**, 533–540 (2013).

[111]A. Zaccone, J. J. Crassous, and M. Ballauff, "Colloidal gelation with variable attraction energy", The Journal of Chemical Physics **138**, 104908 (2013).

[112]D. Hall, J. Kardos, H. Edskes, J. A. Carver, and Y. Goto, "A multi-pathway perspective on protein aggregation: Implications for control of the rate and extent of amyloid formation", FEBS Letters **589**, 672–679 (2015).

[113]A. Šarić, T. C. T. Michaels, A. Zaccone, T. P. J. Knowles, and D. Frenkel, "Kinetics of spontaneous filament nucleation via

oligomers: Insights from theory and simulation", J. Chem. Phys. **145**, 211926 (2016).

[114]T. Steckmann, Z. Awan, B. S. Gerstman, and P. P. Chapagain, "Kinetics of peptide secondary structure conversion during amyloid -protein fibrillogenesis", Journal of Theoretical Biology **301**, 95–102 (2012).

[115]R. Cabriolu, D. Kashchiev, and S. Auer, "Atomistic theory of amyloid fibril nucleation", Journal of Chemical Physics **133** (2010) 10.1063/1.3512642.

[116]D. Kashchiev, R. Cabriolu, and S. Auer, "Confounding the paradigm: Peculiarities of amyloid fibril nucleation", Journal of the American Chemical Society **135**, 1531–1539 (2013).

[117]S. Auer, P. Ricchiuto, and D. Kashchiev, "Two-step nucleation of amyloid fibrils: Omnipresent or not?", J. Mol. Biol. **422**, 723–730 (2012).

[118]R. Cabriolu, D. Kashchiev, and S. Auer, "Breakdown of nucleation theory for crystals with strongly anisotropic interactions between molecules", Journal of Chemical Physics **137** (2012) 10.1063/1.4767531.

[119]S. Auer, "Amyloid fibril nucleation: Effect of amino acid hydrophobicity", Journal of Physical Chemistry B **118**, 5289–5299 (2014).

[120]R. Cabriolu and S. Auer, "Amyloid fibrillation kinetics: Insight from atomistic nucleation theory", Journal of Molecular Biology **411**, 275–285 (2011).

[121]S. Auer, "Nucleation of polymorphic amyloid fibrils", Biophysical Journal **108**, 1176–1186 (2015).

[122]A. I. Taylor, L. D. Gahan, B. Chakrabarti, and R. A. Staniforth, "A two-step biopolymer nucleation model shows a nonequilibrium critical point", The Journal of Chemical Physics **153**, 025102 (2020).

[123]M. M. Pallitto and R. M. Murphy, "A mathematical model of the kinetics of b-amyloid fibril growth from the denatured state", Biophysical Journal **81**, 1805–1822 (2001).

[124]J. D. Schmit, "Kinetic theory of amyloid fibril templating", Journal of Chemical Physics **138** (2013) 10.1063/1.4803658.

[125]J. P. Bernacki and R. M. Murphy, "Model Discrimination and Mechanistic interpretation of Kinetic Data in Protein Aggregation Studies", Biophy. J. **96**, 2871–2887 (2009).

[126]F. Y. Wu, "The Potts model", Reviews of Modern Physics **54**, 235–268 (1982).

[127]J. S. Schreck and J. M. Yuan, *Statistical mechanical treatments of protein amyloid formation*, Vol. 14, 9 (2013), pp. 17420–17452.

[128]A. J. Dear,. Anela, S.. Aric Aric, T. C. T. Michaels, C. M. Dobson, and T. P. J. Knowles, "Statistical Mechanics of Globular Oligomer Formation by Protein Molecules", (2018) 10.1021/acs.jpcb.8b07805.

[129]K. W. Tipping, P. van Oosten-Hawle, E. W. Hewitt, and S. E. Radford, "Amyloid Fibres: Inert End-Stage Aggregates or Key Players in Disease?", Trends in biochemical sciences **40**, 719–727 (2015).

[130]J. Sevigny, P. Chiao, T. Bussière, P. H. Weinreb, L. Williams, M. Maier, R. Dunstan, S. Salloway, T. Chen, Y. Ling, J. O'Gorman, F. Qian, M. Arastu, M. Li, S. Chollate, M. S. Brennan, O. Quintero-Monzon, R. H. Scannevin, H. M. Arnold, T. Engber, K. Rhodes, J. Ferrero, Y. Hang, A. Mikulskis, J. Grimm, C. Hock, R. M. Nitsch, and A. Sandrock, "The antibody aducanumab reduces Aβ plaques in Alzheimer's disease", Nature **537**, 50–56 (2016).

[131]S. Z. Levine, Y. Goldberg, K. Yoshida, M. Samara, A. Cipriani, T. Iwatsubo, S. Leucht, and T. A. Furukawa, "Early- and subsequent- response of cognitive functioning in Alzheimer's disease: Individual-participant data from five pivotal randomized clinical trials of donepezil", Journal of psychiatric research **148**, 159–164 (2022).

[132]Alzforum, *ALZFORUM — NETWORKING FOR A CURE*.

[133]S. L. Rogers, M. R. Farlow, R. S. Doody, R. Mohs, and L. T. Friedhoff, "A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group", Neurology **50**, 136–145 (1998).

[134]A. Burns, M. Rossor, J. Hecker, S. Gauthier, H. Petit, H. J. Möller, S. L. Rogers, and L. T. Friedhoff, "The effects of donepezil in Alzheimer's disease - results from a multinational trial", Dementia and geriatric cognitive disorders **10**, 237–244 (1999).

[135]G. Wilcock, I. Howe, H. Coles, S. Lilienfeld, L. Truyen, Y. Zhu, R. Bullock, and P. Kershaw, "A long-term comparison of galantamine and donepezil in the treatment of Alzheimer's disease", Drugs & aging **20**, 777–789 (2003).

[136]J. Cummings, G. Lefèvre, G. Small, and S. Appel-Dingemanse, "Pharmacokinetic rationale for the rivastigmine patch", Neurology **69** (2007) 10.1212/01.WNL.0000281846.40390.50.

[137]N. Balázs, D. Bereczki, and T. Kovács, "Cholinesterase inhibitors and memantine for the treatment of Alzheimer and non-Alzheimer dementias", Ideggyogyaszati szemle **74**, 379–387 (2021).

[138]B. Reisberg, R. Doody, A. Stöffler, F. Schmitt, S. Ferris, and H. J. Möbius, "A 24-week open-label extension study of memantine in moderate to severe Alzheimer disease", Archives of neurology **63**, 49–54 (2006).

[139]R. S. Doody, P. N. Tariot, E. Pfeiffer, J. T. Olin, and S. M. Graham, "Meta-analysis of six-month memantine trials in Alzheimer's disease", Alzheimer's & dementia : the journal of the Alzheimer's Association **3**, 7–17 (2007).

[140]L. S. Schneider, K. S. Dagerman, J. P. Higgins, and R. McShane, "Lack of evidence for the efficacy of memantine in mild Alzheimer disease", Archives of neurology **68**, 991–998 (2011).

[141]A. M. Pfeil, R. W. Kressig, and T. D. Szucs, "Alzheimer's dementia: budget impact and cost-utility analysis of a combination treatment of a cholinesterase inhibitor and memantine in Switzerland", Swiss medical weekly **142** (2012) 10.4414/SMW.2012.13676.

[142]F. Zhou, X. D. Yan, C. Wang, Y. X. He, Y. Y. Li, J. Zhang, Z. J. Wang, H. Y. Cai, J. S. Qi, and M. N. Wu, "Suvorexant ameliorates cognitive impairments and pathology in APP/PS1 transgenic mice", Neurobiology of aging **91**, 66–75 (2020).

[143]B. O'Nuallain, D. B. Freir, A. J. Nicoll, E. Risse, N. Ferguson, C. E. Herron, J. Collinge, and D. M. Walsh, "Amyloid β-Protein Dimers Rapidly Form Stable Synaptotoxic Protofibrils", Journal of Neuroscience **30**, 14411–14419 (2010).

[144]B. Bohrmann, K. Baumann, J. Benz, F. Gerber, W. Huber, F. Knoflach, J. Messer, K. Oroszlan, R. Rauchenberger, W. F. Richter, C. Rothe, M. Urban, M. Bardroff, M. Winter, C. Nordstedt, and H. Loetscher, "Gantenerumab: a novel human anti-Aβ antibody demonstrates sustained cerebral amyloid-β binding and elicits cell-mediated removal of human amyloid-β", Journal of Alzheimer's disease : JAD **28**, 49–69 (2012).

[145]R. S. Doody, R. G. Thomas, M. Farlow, T. Iwatsubo, B. Vellas, S. Joffe, K. Kieburtz, R. Raman, X. Sun, P. S. Aisen, E. Siemers, H. Liu-Seifert, and R. Mohs, "Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease", The New England journal of medicine **370**, 311–321 (2014).

[146]C. R. Campos, A. M. Kemble, J. Niewoehner, P. O. Freskgård, and E. Urich, "Brain Shuttle Neprilysin reduces central Amyloid-β levels", PloS one **15** (2020) 10.1371/JOURNAL.PONE.0229850.

[147]R. J. Bateman, T. L. Benzinger, S. Berry, D. B. Clifford, C. Duggan, A. M. Fagan, K. Fanning, M. R. Farlow, J. Hassenstab, E. M.

McDade, S. Mills, K. Paumier, M. Quintana, S. P. Salloway, A. Santacruz, L. S. Schneider, G. Wang, and C. Xiong, "The DIAN-TU Next Generation Alzheimer's prevention trial: Adaptive design and disease progression model", Alzheimer's & Dementia **13**, 8–19 (2017).

[148]L. S. Honig, B. Vellas, M. Woodward, M. Boada, R. Bullock, M. Borrie, K. Hager, N. Andreasen, E. Scarpini, H. Liu-Seifert, M. Case, R. A. Dean, A. Hake, K. Sundell, V. Poole Hoffmann, C. Carlson, R. Khanna, M. Mintun, R. DeMattos, K. J. Selzler, and E. Siemers, "Trial of Solanezumab for Mild Dementia Due to Alzheimer's Disease", New England Journal of Medicine **378**, 321–330 (2018).

[149]R. Koradi, M. Billeter, and K. Wüthrich, "MOLMOL: A program for display and analysis of macromolecular structures", Journal of Molecular Graphics **14**, 51–55 (1996).

[150]S. Linse, T. Scheidt, K. Bernfur, M. Vendruscolo, C. M. Dobson, S. I. Cohen, E. Sileikis, M. Lundqvist, F. Qian, T. O'Malley, T. Bussiere, P. H. Weinreb, C. K. Xu, G. Meisl, S. R. Devenish, T. P. Knowles, and O. Hansson, "Kinetic fingerprints differentiate the mechanisms of action of anti-Aβ antibodies", Nature structural & molecular biology **27**, 1125–1133 (2020).

[151]H. Akiyama, H. Kondo, T. Arai, K. Ikeda, M. Kato, E. Iseki, C. Schwab, and P. L. McGeer, "Expression of BRI, the normal precursor of the amyloid protein of familial British dementia, in human brain", Acta Neuropathologica 2003 107:1 **107**, 53–58 (2003).

[152]J. Hedlund, J. Johansson, and B. Persson, "BRICHOS - A superfamily of multidomain proteins with diverse functions", BMC Research Notes **2**, 1–10 (2009).

[153]S. D. Knight, J. Presto, S. Linse, and J. Johansson, "The BRICHOS domain, amyloid fibril formation, and their relationship", Biochemistry **52**, 7523–7531 (2013).

[154]J. D. Sipe, M. D. Benson, J. N. Buxbaum, S. I. Ikeda, G. Merlini, M. J. Saraiva, and P. Westermark, "Amyloid fibril protein nomenclature: 2012 recommendations from the Nomenclature Committee of the International Society of Amyloidosis", Amyloid : the international journal of experimental and clinical investigation : the official journal of the International Society of Amyloidosis **19**, 167–170 (2012).

[155]L. U. Peng, L. I. Shuixiang, J. Zhao, and M. Lingyi, "SCIENCE CHINA Physics, Mechanics & Astronomy A computational in-

vestigation on random packings of sphere-spherocylinder mixtures", **53**, 2284–2292 (2010).

[156]S. I. Cohen, P. Arosio, J. Presto, F. R. Kurudenkandy, H. Biverstål, L. Dolfe, C. Dunning, X. Yang, B. Frohm, M. Vendruscolo, J. Johansson, C. M. Dobson, A. Fisahn, T. P. Knowles, and S. Linse, "A molecular chaperone breaks the catalytic cycle that generates toxic Aβ oligomers", Nature Structural and Molecular Biology **22**, 207–213 (2015).

[157]S. I. A. Cohen, P. Arosio, J. Presto, F. Roshan Kurudenkandy, H. Biverstål, L. Dolfe, C. Dunning, X. Yang, B. Frohm, M. Vendruscolo, J. Johansson, C. M. Dobson, A. Fisahn, T. P. J. Knowles, and S. Linse, "A molecular chaperone breaks the catalytic cycle that generates toxic Aβ oligomers", (2015) 10.1038/nsmb.2971.

[158]J. N. Buxbaum, Z. Ye, N. Reixach, L. Friske, C. Levy, P. Das, T. Golde, E. Masliah, A. R. Roberts, and T. Bartfai, "Transthyretin protects Alzheimer's mice from the behavioral and biochemical effects of Abeta toxicity", Proceedings of the National Academy of Sciences of the United States of America **105**, 2681–2686 (2008).

[159]X. Li and J. N. Buxbaum, "Transthyretin and the brain re-visited: is neuronal synthesis of transthyretin protective in Alzheimer's disease?", Molecular neurodegeneration **6** (2011) 10.1186/1750-1326-6-79.

[160]K. Pagano, S. Tomaselli, H. Molinari, and L. Ragona, "Natural Compounds as Inhibitors of Aβ Peptide Aggregation: Chemical Requirements and Molecular Mechanisms", Frontiers in Neuroscience **14**, 1384 (2020).

[161]P. M. Seidler, K. A. Murray, D. R. Boyer, P. Ge, M. R. Sawaya, C. J. Hu, X. Cheng, R. Abskharon, H. Pan, M. A. DeTure, C. K. Williams, D. W. Dickson, H. V. Vinters, and D. S. Eisenberg, "Structure-based discovery of small molecules that disaggregate Alzheimer's disease tissue derived tau fibrils in vitro", Nature Communications 2022 13:1 **13**, 1–12 (2022).

[162]K. Andrich and J. Bieschke, "The effect of (-)-epigallo-catechin-(3)-gallate on amyloidogenic proteins suggests a common mechanism", Advances in Experimental Medicine and Biology **863**, 139–161 (2015).

[163]J. Bieschke, J. Russ, R. P. Friedrich, D. E. Ehrnhoefer, H. Wobst, K. Neugebauer, and E. E. Wanker, "EGCG remodels mature α-synuclein and amyloid-β fibrils and reduces cellular toxicity", Proceedings of the National Academy of Sciences of the United States of America **107**, 7710–7715 (2010).

[164]D. Townsend, E. Hughes, G. Akien, K. L. Stewart, S. E. Radford, D. Rochester, and D. A. Middleton, "Epigallocatechin-3-gallate remodels apolipoprotein A-I amyloid fibrils into soluble oligomers in the presence of heparin", The Journal of biological chemistry **293**, 12877–12893 (2018).

[165]R. Ahmed, B. Vanschouwen, N. Jafari, X. Ni, J. Ortega, and G. Melacini, "Molecular mechanism for the (-)-Epigallocatechin gallate-induced toxic to nontoxic remodeling of Aβ oligomers", Journal of the American Chemical Society **139**, 13720–13734 (2017).

[166]G. Henríquez, A. Gomez, E. Guerrero, and M. Narayan, "Potential Role of Natural Polyphenols against Protein Aggregation Toxicity: In Vitro, In Vivo, and Clinical Studies", (2020) 10.1021/acschemneuro.0c00381.

[167]A. Daccache, C. Lion, N. Sibille, M. Gerard, C. Slomianny, G. Lippens, and P. Cotelle, "Oleuropein and derivatives from olives as Tau aggregation inhibitors", Neurochemistry International **58**, 700–707 (2011).

[168]B. Cheng, H. Gong, H. Xiao, R. B. Petersen, L. Zheng, and K. Huang, "Inhibiting toxic aggregation of amyloidogenic proteins: A therapeutic strategy for protein misfolding diseases", Biochimica et Biophysica Acta (BBA) - General Subjects **1830**, 4860–4871 (2013).

[169]M. J. Ramalho, S. Andrade, J. A. Loureiro, and M. do Carmo Pereira, "Nanotechnology to improve the Alzheimer's disease therapy with natural compounds", Drug Delivery and Translational Research 2019 10:2 **10**, 380–402 (2019).

[170]J. Rivera-Pérez, M. Martínez-Rosas, C. A. Conde-Castañón, J. D. Toscano-Garibay, N. J. Ruiz-Pérez, P. L. Flores, E. M. Jiménez, and J. Flores-Estrada, "Epigallocatechin 3-Gallate Has a Neuroprotective Effect in Retinas of Rabbits with Ischemia/Reperfusion through the Activation of Nrf2/HO-1", International journal of molecular sciences **21** (2020) 10.3390/IJMS21103716.

[171]R. Krishnan, H. Tsubery, M. Y. Proschitsky, E. Asp, M. Lulu, S. Gilead, M. Gartner, J. P. Waltho, P. J. Davis, A. M. Hounslow, D. A. Kirschner, H. Inouye, D. G. Myszka, J. Wright, B. Solomon, and R. A. Fisher, "A Bacteriophage Capsid Protein Provides a General Amyloid Interaction Motif (GAIM) That Binds and Remodels Misfolded Protein Assemblies", Journal of Molecular Biology **426**, 2500–2519 (2014).

[172]J. M. Levenson, S. Schroeter, J. C. Carroll, V. Cullen, E. Asp, M. Proschitsky, C. H. Chung, S. Gilead, M. Nadeem, H. B. Dodiya, S. Shoaga, E. J. Mufson, H. Tsubery, R. Kr-

ishnan, J. Wright, B. Solomon, R. Fisher, and K. S. Gannon, "NPT088 reduces both amyloid-β and tau pathologies in transgenic mice", Alzheimer's & dementia (New York, N. Y.) **2**, 141–155 (2016).

[173]J. Messing, "Phage M13 for the treatment of Alzheimer and Parkinson disease", Gene **583**, 85–89 (2016).

[174]Y. Poirier, A. Grimm, K. Schmitt, and A. Eckert, "Link between the unfolded protein response and dysregulation of mitochondrial bioenergetics in Alzheimer's disease", Cellular and molecular life sciences : CMLS **76**, 1419–1431 (2019).

[175]H. C. Murray, B. V. Dieriks, M. E. Swanson, P. V. Anekal, C. Turner, R. L. Faull, L. Belluscio, A. Koretsky, and M. A. Curtis, "The unfolded protein response is activated in the olfactory system in Alzheimer's disease", Acta Neuropathologica Communications **8**, 1–15 (2020).

[176]J. M. Grandjean and R. Luke Wiseman, "Small molecule strategies to harness the unfolded protein response: where do we go from here?", The Journal of Biological Chemistry **295**, 15692 (2020).

[177]J. M. Van Der Harg, A. Nölle, R. Zwart, A. S. Boerema, E. S. Van Haastert, A. M. Strijkstra, J. J. Hoozemans, and W. Scheper, "The unfolded protein response mediates reversible tau phosphorylation induced by metabolic stress", Cell Death & Disease **5**, e1393 (2014).

[178]G. P. Lotz and J. Legleiter, "The role of amyloidogenic protein oligomerization in neurodegenerative disease", Journal of Molecular Medicine **91**, 653–664 (2013).

[179]L. Breydo and V. N. Uversky, "Structural, morphological, and functional diversity of amyloid oligomers", FEBS Letters **589**, 2640–2648 (2015).

[180]S. Kmiecik, D. Gront, M. Kolinski, L. Wieteska, A. E. Dawid, and A. Kolinski, "Coarse-Grained Protein Models and Their Applications", Chemical Reviews **116**, 7898–936 (2016).

[181]M. Levitt and A. Warshel, "Computer simulation of protein folding", Nature **253**, 694–698 (1975).

[182]J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD", Journal of Computational Chemistry **26**, 1781–1802 (2005).

[183]H. Berendsen, D. van der Spoel, and R. van Drunen, "GROMACS: A message-passing parallel molecular dynamics implementa-

tion", Computer Physics Communications **91**, 43–56 (1995).

[184]E. Lindahl, B. Hess, and D. van der Spoel, "GROMACS 3.0: a package for molecular simulation and trajectory analysis", Journal of Molecular Modeling **7**, 306–317 (2001).

[185]B. Hess*, C. Kutzner, D. van der Spoel, and E. Lindahl, "GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation", Journal of Chemical Theory and Computation **4**, 435–447 (2008).

[186]S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, "GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit", Bioinformatics **29**, 845–854 (2013).

[187]S. Páll, M. J. Abraham, C. Kutzner, B. Hess, and E. Lindahl, "Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS", in *Lecture notes in computer science* (Springer, Cham, 2015), pp. 3–27.

[188]M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers", SoftwareX **1-2**, 19–25 (2015).

[189]H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. de Jong, S. M. Gopal, X. Periole, and S. J. Marrink, "The power of coarse graining in biomolecular simulations.", Wiley interdisciplinary reviews. Computational molecular science **4**, 225–248 (2014).

[190]M. I. Zimmerman, J. R. Porter, M. D. Ward, S. Singh, N. Vithani, A. Meller, U. L. Mallimadugula, C. E. Kuhn, J. H. Borowsky, R. P. Wiewiora, M. F. D. Hurley, A. M. Harbison, C. A. Fogarty, J. E. Coffland, E. Fadda, V. A. Voelz, J. D. Chodera, and G. R. Bowman, "SARS-CoV-2 Simulations Go Exascale to Capture Spike Opening and Reveal Cryptic Pockets Across the Proteome", bioRxiv (2020) 10.1101/2020.06.27.175430.

[191]A. Montagud, M. Ponce-de Leon, and A. Valencia, "Systems biology at the giga-scale: Large multiscale models of complex, heterogeneous multicellular systems", Current Opinion in Systems Biology **28**, 100385 (2021).

[192]B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov,

E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, "CHARMM: The biomolecular simulation program", Journal of Computational Chemistry **30**, 1545–1614 (2009).

[193]D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, "The Amber biomolecular simulation programs", Journal of Computational Chemistry **26**, 1668–1688 (2005).

[194]S. Plimpton, *Fast Parallel Algorithms for Short-Range Molecular Dynamics*, tech. rep. (1995), pp. 1–19.

[195]D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. Berendsen, "GROMACS: Fast, flexible, and free", Journal of Computational Chemistry **26**, 1701–1718 (2005).

[196]M. J. Abraham, D van der Spoel, E Lindahl, and B Hess, *GROMACS User Manual version 5.0.4*, 2014.

[197]N. W. Lima, L. I. Gutierres, R. I. Gonzalez, S. Müller, R. S. Thomaz, E. M. Bringa, and R. M. Papaléo, "Molecular dynamics simulation of polymerlike thin films irradiated by fast ions: A comparison between FENE and Lennard-Jones potentials", Physical Review B **94**, 195417 (2016).

[198]P. P. Ewald, "Die Berechnung optischer und elektrostatischer Gitterpotentiale", Ann. Phys. **64**, 253–287 (1921).

[199]L. Verlet, "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules", Physical Review **159**, 98–103 (1967).

[200]J. E. Jones, "On the Determination of Molecular Fields. II. From the Equation of State of a Gas", Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **106**, 463–477 (1924).

[201]R. A. Buckingham, "The classical equation of state of gaseous helium, neon and argon", Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **168**, 264–283 (1938).

[202]S. H. Chong and S. Ham, "Distinct role of hydration ater in protein misfolding and aggregation revealed by fluctuating thermodynamics analysis", Accounts of Chemical Research **48**, 956–965 (2015).

[203]R. D. Hills and C. L. Brooks, "Hydrophobic Cooperativity as a Mechanism for Amyloid Nucleation", Journal of Molecular Biology **368**, 894–901 (2007).

[204]A. W. Fitzpatrick, T. P. Knowles, C. A. Waudby, M. Vendruscolo, and C. M. Dobson, "Inversion of the balance between hydrophobic and hydrogen bonding interactions in protein folding and aggregation", PLoS Computational Biology **7** (2011) 10.1371/journal.pcbi.1002169.

[205]M. Schor, A. S.J. S. Mey, F. Noé, and C. E. MacPhee, "Shedding Light on the DockLock Mechanism in Amyloid Fibril Growth Using Markov State Models", The Journal of Physical Chemistry Letters **6**, 1076–1081 (2015).

[206]F. Baftizadeh, F. Pietrucci, X. Biarnés, and A. Laio, "Nucleation process of a fibril precursor in the C-terminal segment of amyloid-β", Physical Review Letters **110**, 1–5 (2013).

[207]P. H. Nguyen and P. Derreumaux, "Conformational ensemble and polymorphism of the all-atom alzheimer's Aβ37-42amyloid peptide oligomers", Journal of Physical Chemistry B **117**, 5831–5840 (2013).

[208]A. Baumketner and J. E. Shea, "Folding Landscapes of the Alzheimer Amyloid-β(12-28) Peptide", Journal of Molecular Biology **362**, 567–579 (2006).

[209]F. Simona, G. Tiana, R. A. Broglia, and G. Colombo, "Modeling the α-helix to β-hairpin transition mechanism and the formation of oligomeric aggregates of the fibrillogenic peptide Aβ(12-28): Insights from all-atom molecular dynamics simulations", Journal of Molecular Graphics and Modelling **23**, 263–273 (2004).

[210]A. Irbäck and J. Wessén, "Thermodynamics of amyloid formation and the role of intersheet interactions", Journal of Chemical Physics (2015) 10.1063/1.4930280.

[211]E. Socher, H. Sticht, and A. H. C. Horn, "The conformational stability of nonfibrillar amyloid-β peptide oligomers critically depends on the C-terminal peptide length", ACS Chemical Neuroscience **5**, 161–167 (2014).

[212]L. Liang, L. W. Wang, and J. W. Shen, "The self-assembly mechanism of tetra-peptides from the motif of β-amyloid peptides: A combined coarse-grained and all-atom molecular dynamics simulation", RSC Advances (2016) 10.1039/c6ra18204f.

[213]T. X. Hoang, L. Marsella, A. Trovato, F. Seno, J. R. Banavar, and A. Maritan, "Common attributes of native-state structures of proteins, disordered proteins, and amyloid.", Proceedings of the National Academy of Sciences of the United States of America **103**, 6883–8 (2006).

[214]J. H. Meinke and U. H. Hansmann, "Aggregation of -amyloid fragments", Journal of

Chemical Physics **126** (2007) 10.1063/1.2423013.

[215]A. De Simone and P. Derreumaux, "Low molecular weight oligomers of amyloid peptides display -barrel conformations: A replica exchange molecular dynamics study in explicit solvent", Journal of Chemical Physics **132** (2010) 10.1063/1.3385470.

[216]T. Steckmann, Y. R. Bhandari, P. P. Chapagain, and B. S. Gerstman, "Cooperative structural transitions in amyloid-like aggregation", Journal of Chemical Physics **146** (2017) 10.1063/1.4979516.

[217]Y. Lu, P. Derreumaux, Z. Guo, N. Mousseau, and G. Wei, "Thermodynamics and dynamics of amyloid peptide oligomerization are sequence dependent", Proteins: Structure, Function and Bioinformatics **75**, 954–963 (2009).

[218]J Xu, J. Z. Zhang, and Y Xiang, "Molecular dynamics simulation and computational two-dimensional infrared spectroscopic study of model amyloid beta-peptide oligomers", The Journal of Physical Chemistry A **117**, 6373–6379 (2013).

[219]X. Yu and J. Zheng, "Polymorphic structures of alzheimer's β-amyloid globulomers", PLoS ONE **6** (2011) 10.1371/journal.pone.0020575.

[220]S. Kim, T. Takeda, and D. K. Klimov, "Globular state in the oligomers formed by A peptides", Journal of Chemical Physics **132** (2010) 10.1063/1.3447894.

[221]T. Gurry and C. M. Stultz, "Mechanism of amyloid-β fibril elongation", Biochemistry **53**, 6981–6991 (2014).

[222]B. Barz and B. Urbanc, "Dimer formation enhances structural differences between amyloid β-protein (1-40) and (1-42): An explicit-solvent molecular dynamics study", PLoS ONE **7** (2012) 10.1371/journal.pone.0034345.

[223]Y. S. Lin, G. R. Bowman, K. A. Beauchamp, and V. S. Pande, "Investigating how peptide length and a pathogenic mutation modify the structural ensemble of amyloid beta monomer", Biophysical Journal **102**, 315–324 (2012).

[224]W. Zheng, N. P. Schafer, and P. G. Wolynes, "Free energy landscapes for initiation and branching of protein aggregation", Proceedings of the National Academy of Sciences **110**, 20515–20520 (2013).

[225]S. Côté, R. Laghaei, P. Derreumaux, and N. Mousseau, "Distinct dimerization for various alloforms of the amyloid-beta protein: Aβ1-40, Aβ1-42, and Aβ1-40(D23N)", Journal of

Physical Chemistry B (2012) 10 . 1021 / jp2126366.

[226]K. Kassler, A. H. C. Horn, and H. Sticht, "Effect of pathogenic mutations on the structure and dynamics of Alzheimer's A??42-amyloid oligomers", Journal of Molecular Modeling **16**, 1011–1020 (2010).

[227]L. G. Rizzi and S. Auer, "Amyloid Fibril Solubility", Journal of Physical Chemistry B **119**, 14631–14636 (2015).

[228]T. R. Rieger, R. I. Morimoto, and V. Hatzimanikatis, "Bistability explains threshold phenomena in protein aggregation both in vitro and in vivo", Biophysical Journal **90**, 886–895 (2006).

[229]E. I. Lin and M. S. Shell, "Can Peptide Folding Simulations Provide Predictive Information for Aggregation Propensity?", The Journal of Physical Chemistry B **114**, 11899–11908 (2010).

[230]J. a. Lemkul and D. R. Bevan, "Assessing the Stability of Alzheimer ' s Amyloid Protofibrils Using Molecular Dynamics (Supporting Information) (a) (b)", Journal of Physical Chemistry B **114**, 100 (2010).

[231]V. Foderà, A. Zaccone, M. Lattuada, and A. M. Donald, "Electrostatics controls the formation of amyloid superstructures in protein aggregation", Physical Review Letters **111**, 1–5 (2013).

[232]S. Rao Jampani, F. Mahmoudinobar, Z. Su, and C. L. Dias, "Thermodynamics of A??16-21 dissociation from a fibril: Enthalpy, entropy, and volumetric properties", Proteins: Structure, Function and Bioinformatics **83**, 1963–1972 (2015).

[233]J. Zidar and F. Merzel, "Probing amyloid-beta fibril stability by increasing ionic strengths", Journal of Physical Chemistry B **115**, 2075–2081 (2011).

[234]A. Fernández, "What factor drives the fibrillogenic association of β-sheets?", FEBS Letters **579**, 6635–6640 (2005).

[235]A. H. Horn and H. Sticht, "Amyloid-β42 oligomer structures from fibrils: A systematic molecular dynamics study", Journal of Physical Chemistry B **114**, 2219–2226 (2010).

[236]W. Han and K. Schulten, "Fibril Elongation by Aβ 1742 : Kinetic Network Analysis of Hybrid-Resolution Molecular Dynamics Simulations", Journal of the American Chemical Society **136**, 12450–12460 (2014).

[237]A. V. Rojas, A. Liwo, and H. A. Scheraga, "A study of the α-helical intermediate preceding the aggregation of the amino-terminal fragment of the β amyloid peptide (Aβ 1-28)",

Journal of Physical Chemistry B **115**, 12978–12983 (2011).

[238]J. Zhao, Q. Wang, G. Liang, and J. Zheng, "Molecular dynamics simulations of low-ordered Alzheimer β-amyloid oligomers from dimer to hexamer on self-assembled monolayers", Langmuir **27**, 14876–14887 (2011).

[239]Y. Lu, G. Wei, and P. Derreumaux, "Effects of G33A and G33I mutations on the structures of monomer and dimer of the amyloid-β fragment 29-42 by replica exchange molecular dynamics simulations", Journal of Physical Chemistry B **115**, 1282–1288 (2011).

[240]H. Jang, F. Teran Arce, S. Ramachandran, R. Capone, R. Lal, and R. Nussinov, "Structural convergence among diverse, toxic β-sheet ion channels", Journal of Physical Chemistry B **114**, 9445–9451 (2010).

[241]H. Jang, J. Zheng, and R. Nussinov, "Models of β-amyloid ion channels in the membrane suggest that channel formation in the bilayer is a dynamic process", Biophysical Journal **93**, 1938–1949 (2007).

[242]S. Jang and S. Shin, "Computational study on the structural diversity of amyloid Beta Peptide (abeta(10-35)) oligomers.", The journal of physical chemistry. B **112**, 3479–3484 (2008).

[243]B. Strodel, J. W. Lee, C. S. Whittleston, and D. J. Wales, "Transmembrane structures for Alzheimer's Aβ1-42oligomers", Journal of the American Chemical Society **132**, 13300–13312 (2010).

[244]N. Metropolis, "The beginning of the Monte Carlo method", Los Alamos Science, 125–130 (1987).

[245]D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", Journal of Computational Physics **22**, 403–434 (1976).

[246]D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions", The Journal of Physical Chemistry **81**, 2340–2361 (1977).

[247]K. A. Berman and J. L. Paul, *Algorithms : sequential, parallel, and distributed* (Thomson/-Course Technology, Boston, 2005), p. 962.

[248]J. Zhang and M. Muthukumar, "Simulations of nucleation and elongation of amyloid fibrils", Journal of Chemical Physics (2009) 10.1063/1.3050295.

[249]R. Pellarin, P. Schuetz, E. Guarnera, and A. Caflisch, "Amyloid fibril polymorphism is under kinetic control", Journal of the American Chemical Society **132**, 14960–14970 (2010).

[250]S. Auer and D. Kashchiev, "Phase diagram of α-helical and β-sheet forming peptides", Physical Review Letters **104**, 2–5 (2010).

[251]N. S. Bieler, T. P. J. Knowles, D Frenkel, and R Vácha, "Connecting Macroscopic Observables and Microscopic Assembly Events in Amyloid Formation Using Coarse Grained Simulations", PLoS Comput. Biol. **8**, e1002692 (2012).

[252]A. Rojas, A. Liwo, D. Browne, and H. A. Scheraga, "Mechanism of Fiber Assembly: Treatment of Aβ Peptide Aggregation with a Coarse-Grained United-Residue Force Field", Journal of Molecular Biology **404**, 537–552 (2010).

[253]P. Hortschansky, V. Schroeckh, T. Christopeit, G. Zandomeneghi, and M. Fändrich, "The aggregation kinetics of Alzheimer ' s b-amyloid peptide is controlled by stochastic nucleation", Protein Science **14**, 1753–1759 (2005).

[254]H. D. Nguyen and C. K. Hall, "Phase diagrams describing fibrillization by polyalanine peptides", Biophysical Journal **87**, 4122–4134 (2004).

[255]A. De Simone, C. Kitchen, A. H. Kwan, M. Sunde, C. M. Dobson, and D. Frenkel, "Intrinsic disorder modulates protein self-assembly and aggregation", Proceedings of the National Academy of Sciences **109**, 6951–6956 (2012).

[256]S. J. Marrink, A. H. de Vries, and A. E. Mark, "Coarse Grained Model for Semiquantitative Lipid Simulations", The Journal of Physical Chemistry B **108**, 750–760 (2004).

[257]S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, "The MARTINI force field: coarse grained model for biomolecular simulations.", The journal of physical chemistry. B **111**, 7812–24 (2007).

[258]L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink, "The MARTINI coarse-grained force field: Extension to proteins", Journal of Chemical Theory and Computation **4**, 819–834 (2008).

[259]M. Seo, S. Rauscher, R. Pomès, and D. P. Tieleman, "Improving Internal Peptide Dynamics in the Coarse-Grained MARTINI Model: Toward Large-Scale Simulations of Amyloid- and Elastin-like Peptides", Journal of Chemical Theory and Computation **8**, 1774–1785 (2012).

[260]S. Auer, C. M. Dobson, M. Vendruscolo, and A. Maritan, "Self-templated nucleation in peptide and protein aggregation", Phys. Rev. Lett. **101**, 258101 (2008).

[261]S. Auer, "Phase diagram of polypeptide chains.", The Journal of chemical physics **135**, 175103 (2011).

[262]S. Auer, F. Meersman, C. M. Dobson, and M. Vendruscolo, "A generic mechanism of emergence of amyloid protofilaments from disordered oligomeric aggregates", PLoS Computational Biology **4** (2008) 10.1371/journal.pcbi.1000222.

[263]S. Auer, C. M. Dobson, and M. Vendruscolo, "Characterization of the nucleation barriers for protein aggregation and amyloid formation", HFSP J. **1**, 137–146 (2007).

[264]P. Ricchiuto, A. V. Brukhno, and S. Auer, "Protein aggregation: Kinetics versus thermodynamics", Journal of Physical Chemistry B **116**, 5384–5390 (2012).

[265]R. Pellarin and A. Caflisch, "Interpreting the Aggregation Kinetics of Amyloid Peptides", Journal of Molecular Biology **360**, 882–892 (2006).

[266]R. Pellarin, E. Guarnera, and A. Caflisch, "Pathways and Intermediates of Amyloid Fibril Formation", J. Mol. Biol. **374**, 917–924 (2007).

[267]G. Bellesia and J. E. Shea, "Effect of -sheet propensity on peptide aggregation", Journal of Chemical Physics **130** (2009) 10.1063/1.3108461.

[268]G. Bellesia and J. E. Shea, "Self-assembly of -sheet forming peptides into chiral fibrillar aggregates", Journal of Chemical Physics **126** (2007) 10.1063/1.2739547.

[269]T. Bereau and M. Deserno, "No Title", **96**, 405a (2009).

[270]R. Vácha, S. Linse, and M. Lund, "Surface effects on aggregation kinetics of amyloidogenic peptides.", J Am Chem Soc **136**, 11776–11782 (2014).

[271]J. Mondal, B. J. Sung, and A. Yethiraj, "Sequence dependent self-assembly of -peptides: Insights from a coarse-grained model", Journal of Chemical Physics **132** (2010) 10.1063/1.3314309.

[272]A. Morriss-Andrews, G. Bellesia, and J. E. Shea, "β-Sheet propensity controls the kinetic pathways and morphologies of seeded peptide aggregation", Journal of Chemical Physics **137** (2012) 10.1063/1.4755748.

[273]A. Morriss-Andrews, F. L. H. Brown, and J. E. Shea, "A coarse-grained model for peptide aggregation on a membrane surface", Journal of Physical Chemistry B (2014) 10.1021/jp502871m.

[274]A. Zaccone, I. Terentjev, T. W. Herling, T. P. Knowles, A. Aleksandrova, and E. M. Terentjev, "Kinetics of fragmentation and dissociation of two-strand protein filaments: Coarse-

grained simulations and experiments", Journal of Chemical Physics **145** (2016) 10.1063/1.4962366.

[275]B Urbanc, M Betnel, L Cruz, G Bitan, and D. B. Teplow, "Elucidation of amyloid β-protein oligomerization mechanisms: Discrete molecular dynamics study", J. Am. Chem. Soc. **132**, 4266–4280 (2010).

[276]D. Hu, W. Zhao, Y. Zhu, H. Ai, and B. Kang, "Bead-Level Characterization of Early-Stage Amyloidβ ¡sub¿42¡/sub¿ Aggregates: Nuclei and Ionic Concentration Effects", Chemistry - A European Journal **23**, 16257–16273 (2017).

[277]D. C. Latshaw, M. Cheon, and C. K. Hall, "Effects of macromolecular crowding on amyloid beta (16-22) aggregation using coarse-grained simulations", Journal of Physical Chemistry B (2014) 10.1021/jp508970q.

[278]S. Yun, B. Urbanc, L. Cruz, G. Bitan, D. B. Teplow, and H. E. Stanley, "Role of electrostatic interactions in amyloid β-protein (Aβ) oligomer formation: A discrete molecular dynamics study", Biophysical Journal **92**, 4064–4077 (2007).

[279]L. Xu, Y. Chen, and X. Wang, "Assembly of amyloid β peptides in the presence of fibril seeds: One-pot coarse-grained molecular dynamics simulations", Journal of Physical Chemistry B **118**, 9238–9246 (2014).

[280]T. T. Tran, P. H. Nguyen, and P. Derreumaux, "Lattice model for amyloid peptides: OPEP force field parametrization and applications to the nucleus size of Alzheimer's peptides", The Journal of Chemical Physics **144**, 205103 (2016).

[281]A. E. Hafner, J. Krausser, and A. Šarić, "Minimal coarse-grained models for molecular self-organisation in biology", Current Opinion in Structural Biology **58**, 43–52 (2019).

[282]F. Mohajerani, M. F. Hagan, and M. A. Fisher, "The role of the encapsulated cargo in microcompartment assembly", (2018) 10.1371/journal.pcbi.1006351.

[283]S. L. Freedman, G. M. Hocky, S. Banerjee, and A. R. Dinner, "Nonequilibrium phase diagrams for actomyosin networks", Soft Matter **14**, 7740–7747 (2018).

[284]J. Weichsel and P. L. Geissler, "The More the Tubular: Dynamic Bundling of Actin Filaments for Membrane Tube Formation", PLOS Computational Biology — OPEN ACCESS Citation: Weichsel J, Geissler PL **12**, 1004982 (2016).

[285]M. Giani, W. K. Den Otter, and W. J. Briels, "Early stages of clathrin aggregation at a membrane in coarse-grained simulations", The Journal of Chemical Physics **146**, 155102 (2017).

[286]M. Pannuzzo, Z. A. McDargh, and M. Deserno, "The role of scaffold reshaping and disassembly in dynamin driven membrane fission", eLife **7** (2018) 10.7554/ELIFE.39441.

[287]L. Harker-Kirschneck, B. Baum, and A. ela Šarić, "Changes in ESCRT-III filament geometry drive membrane remodelling and fission in silico", BMC Biology **17**, 1–8 (2019).

[288]S. Sasmal, N. Schwierz, and T. Head-Gordon, "Mechanism of nucleation and growth of Aβ40 fibrils from all-atom and coarse-grained simulations", Journal of Physical Chemistry B **120**, 12088–12097 (2016).

[289]G. Bellesia, M. V. Fedorov, and E. G. Timoshenko, "Structural transitions in model -sheet tapes", Journal of Chemical Physics **128** (2008) 10.1063/1.2909550.

[290]I. M. Ilie, W. K. Den Otter, and W. J. Briels, "A coarse grained protein model with internal degrees of freedom. Application to α-synuclein aggregation", Journal of Chemical Physics (2016) 10.1063/1.4942115.

[291]A. V Smith and C. K. Hall, "1", Proteins Struct. Funct. Genet **44**, 344–360 (2001).

[292]A Šarić, A. K. Buell, G Meisl, T. C. T. Michaels, C. M. Dobson, S Linse, T. P. J. Knowles, and D Frenkel, "Physical determinants of the self-replication of protein fibrils", Nature Phys. **12**, 874–880 (2016).

[293]A. Šarić, T. C. Michaels, A. Zaccone, T. P. Knowles, and D. Frenkel, "Kinetics of spontaneous filament nucleation via oligomers: Insights from theory and simulation", Journal of Chemical Physics **145** (2016) 10.1063/1.4965040.

[294]T. C. Michaels, A. Šarić, G. Meisl, G. T. Heller, S. Curk, P. Arosio, S. Linse, C. M. Dobson, M. Vendruscolo, and T. P. Knowles, "Thermodynamic and kinetic design principles for amyloid-aggregation inhibitors", Proceedings of the National Academy of Sciences of the United States of America **117**, 24251–24257 (2020).

[295]T. C. Michaels, A. Šarić, S. Curk, K. Bernfur, P. Arosio, G. Meisl, A. J. Dear, S. I. Cohen, C. M. Dobson, M. Vendruscolo, S. Linse, and T. P. Knowles, "Dynamics of oligomer populations formed during the aggregation of Alzheimer's Aβ42 peptide", Nature Chemistry **12**, 445–451 (2020).

[296]J. Park, B. Kahng, R. D. Kamm, and W. Hwang, "Atomistic simulation approach to a continuum description of self-assembled β-sheet filaments", Biophysical Journal **90**, 2510–2524 (2006).

[297]G. Yoon, J. Kwak, J. I. Kim, S. Na, and K. Eom, "Mechanical characterization of amyloid fibrils using coarse-grained normal mode analysis", Advanced Functional Materials **21**, 3454–3463 (2011).

[298]Z. Dogic, D. Frenkel, and S. Fraden, "Enhanced stability of layered phases in parallel hard spherocylinders due to addition of hard spheres", Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics **62**, 3925–3933 (2000).

[299]A Stroobants, H. N. W. Lekkerkerker, and D Frenkel, *Evidence for one-, two-, and three-dimensional order in a system of hard parallel spherocylinders*, tech. rep. 6 (1987).

[300]A. Stroobants, "Liquid crystal phase transitions in bidisperse hard-rod systems", Journal of Physics: Condensed Matter **6** (1994) 10.1088/0953-8984/6/23A/045.

[301]P. Bolhuis and D. Frenkel, "Numerical study of the phase diagram of a mixture of spherical and rodlike colloids", The Journal of Chemical Physics **101**, 666 (1994).

[302]P. Bolhuis and D. Frenkel, "Tracing the phase boundaries of hard spherocylinders", The Journal of Chemical Physics **106**, 666–687 (1997).

[303]P. G. Bolhuis, A. Stroobants, D. Frenkel, and H. N. W. Lekkerkerker, "Cite as", The Journal of Chemical Physics **107**, 244909 (1997).

[304]P. Bolhuis and D. Frenkel, "Cite as", The Journal of Chemical Physics **106**, 1551 (1997).

[305]G. A. Vliegenthart and H. N. Lekkerkerker, "Phase behavior of colloidal rod-sphere mixtures", Journal of Chemical Physics **111**, 4153–4157 (1999).

[306]G. Cinacchi and L. De Gaetani, "Diffusion in the lamellar phase of a rod-sphere mixture", Journal of Chemical Physics **131** (2009) 10.1063/1.3207951.

[307]D. Antypov and D. J. Cleaver, "The role of attractive interactions in rod-sphere mixtures", Journal of Chemical Physics **120**, 10307–10316 (2004).

[308]N. Urakami and M. Imai, "Dependence on sphere size of the phase behavior of mixtures of rods and spheres", Journal of Chemical Physics **119**, 2463–2470 (2003).

[309]L. Wu, A. Malijevský, G. Jackson, E. A. Müller, and C. Avendaño, "Orientational ordering and phase behaviour of binary mixtures of hard spheres and hard spherocylinders", The Journal of Chemical Physics **143**, 044906 (2015).

[310]P. G. Bolhuis, A. Stroobants, D. Frenkel, and H. N. Lekkerkerker, "Numerical study of the phase behavior of rodlike colloids with attractive interactions", Journal of Chemical Physics **107**, 1551–1564 (1997).

[311]A. Stroobants, H. N. Lekkerkerker, and D. Frenkel, "Evidence for Smectic Order in a Fluid of Hard Parallel Spherocylinders", Physical Review Letters **57**, 1452 (1986).

[312]L. Wu, A. Malijevský, C. Avendaño, E. A. Müller, and G. Jackson, "Demixing, surface nematization, and competing adsorption in binary mixtures of hard rods and hard spheres under confinement", Journal of Chemical Physics **148** (2018) 10.1063/1.5020002.

[313]A. Cuetos, B. Martínez-Haya, S. Lago, and L. F. Rull, "Use of Parsons-Lee and Onsager theories to predict nematic and demixing behavior in binary mixtures of hard rods and hard spheres", Physical review. E, Statistical, nonlinear, and soft matter physics **75** (2007) 10.1103/PHYSREVE.75.061701.

[314]S. A. Hollingsworth and P. A. Karplus, "A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins", Biomolecular concepts **1**, 271 (2010).

[315]J. W. Wagner, J. F. Dama, A. E. Durumeric, and G. A. Voth, "On the representability problem and the physical meaning of coarse-grained models", Journal of Chemical Physics **145** (2016) 10.1063/1.4959168.

[316]D. Frenkel and B. Smit, *Understanding Molecular Simulation*, Vol. 11, 4 (AIP Publishing, 1997), p. 351.

[317]M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Clarendon Press, 1989), p. 385.

[318]H. F. Stanyon and J. H. Viles, "Human Serum Albumin Can Regulate Amyloid- Peptide Fiber Growth in the Brain Interstitium", (2012) 10.1074/jbc.C112.360800.

[319]S. M. Johnson, S. Connelly, C. Fearns, E. T. Powers, and J. W. Kelly, "The Transthyretin Amyloidoses : From Delineating the Molecular Mechanism of Aggregation Linked to Pathology to a Regulatory-Agency-Approved Drug", 185–203 (2012).

[320]J. A. Hebda, M. Magzoub, and A. D. Miranker, "Small molecule screening in context: Lipid-catalyzed amyloid formation", Protein Science **23**, 1341–1348 (2014).

[321]C. Galvagnion, A. K. Buell, G. Meisl, T. C. Michaels, M. Vendruscolo, T. P. Knowles, and C. M. Dobson, "Lipid vesicles trigger α-synuclein aggregation by stimulating primary nucleation", Nature Chemical Biology **11**, 229–234 (2015).

[322]R. Vá and D. Frenkel, "Relation between Molecular Shape and the Morphology of Self-Assembling Aggregates: A Simulation Study", 10.1016/j.bpj.2011.07.046.

[323]G. Cinacchi and L. De Gaetani, "Diffusion in the lamellar phase of a rod-sphere mixture", Journal of Chemical Physics **131** (2009) 10.1063/1.3207951.

[324]S. Lago, A. Cuetos, B. Martínez-Haya, and L. F. Rull, "Crowding effects in binary mixtures of rod-like and spherical particles", in Journal of molecular recognition, Vol. 17, 5 (2004), pp. 417–425.

[325]M. Schmidt and J. M. Brader, "Hard sphere fluids in random fiber networks", Journal of Chemical Physics **119**, 3495–3500 (2003).

[326]L. Wu, A. Malijevský, G. Jackson, E. A. Müller, and C. Avendaño, "Orientational ordering and phase behaviour of binary mixtures of hard spheres and hard spherocylinders", Journal of Chemical Physics **143** (2015) 10.1063/1.4923291.

[327]N. Yasarawan and J. S. Van Duijneveldt, "Arrested phase separation of colloidal rod-sphere mixtures", Soft Matter **6**, 353–362 (2010).

[328]P. G. Bolhuis, A. Stroobants, D. Frenkel, and H. N. W. Lekkerkerker, "Numerical study of the phase behavior of rodlike colloids with attractive interactions", The Journal of Chemical Physics **107**, 1551–1564 (1997).

[329]S. V. Savenko and M. Dijkstra, "Sedimentation and multiphase equilibria in suspensions of colloidal hard rods", Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics **70**, 7 (2004).

[330]R. P. Sear, "Nucleation: theory and applications to protein solutions and colloidal suspensions", Journal of Physics: Condensed Matter **19**, 033101 (2007).

[331]T. Strübing, A. Khosravanizadeh, A. Vilfan, E. Bodenschatz, R. Golestanian, and I. Guido, "Wrinkling Instability in 3D Active Nematics", Nano Letters **20**, 6281–6288 (2020).

[332]S. Lee, "A numerical investigation of nematic ordering based on a simple hardrod model", The Journal of Chemical Physics **87**, 4972–4974 (1987).

[333]A. Gorkovskiy, K. R. Thurber, R. Tycko, and R. B. Wickner, "Locating folds of the in-register parallel β-sheet of the Sup35p prion domain infectious amyloid", Proceedings of the National Academy of Sciences of the United States of America **111**, E4615–E4622 (2014).

[334]A. Stukowski, "Visualization and analysis of atomistic simulation data with OVITO-the Open Visualization Tool", MODELLING

AND SIMULATION IN MATERIALS SCIENCE AND ENGINEERING **18** (2010) 10.1088/0965-0393/18/1/015012.

[335]T. E. Creighton, "Proteins : structures and molecular properties", 507 (1993).

[336]C. Ó. Fágáin, "Understanding and increasing protein stability", Biochimica et biophysica acta **1252**, 1–14 (1995).

[337]C. C. Vandenakker, M. F. Engel, K. P. Velikov, M. Bonn, and G. H. Koenderink, "Morphology and persistence length of amyloid fibrils are correlated to peptide molecular structure", Journal of the American Chemical Society **133**, 18030–18033 (2011).

[338]J. Habchi, S. Chia, C. Galvagnion, T. C. T. Michaels, M. M. J. Bellaiche, F. S. Ruggeri, M. Sanguanini, I. Idini, J. R. Kumita, E. Sparr, S. Linse, C. M. Dobson, T. P. J. Knowles, and M. Vendruscolo, "Cholesterol catalyses Aβ42 aggregation through a heterogeneous nucleation pathway in the presence of lipid membranes", Nat. Chem. **10**, 673–683 (2018).

[339]V. Lazarevic, S. Fieńko, M. Andres-Alonso, D. Anni, D. Ivanova, C. Montenegro-Venegas, E. D. Gundelfinger, M. A. Cousin, and A. Fejtova, "Physiological Concentrations of Amyloid Beta Regulate Recycling of Synaptic Vesicles via Alpha7 Acetylcholine Receptor and CDK5/Calcineurin Signaling", Frontiers in Molecular Neuroscience **10** (2017) 10.3389/FNMOL.2017.00221.

[340]X. Hu, S. L. Crick, G. Bu, C. Frieden, R. V. Pappu, and J. M. Lee, "Amyloid seeds formed by cellular uptake, concentration, and aggregation of the amyloid-beta peptide", Proceedings of the National Academy of Sciences of the United States of America **106**, 20324–20329 (2009).

[341]C. M. Yuede, H. Lee, J. L. Restivo, T. A. Davis, J. C. Hettinger, C. E. Wallace, K. L. Young, M. R. Hayne, G. Bu, C. zhong Li, and J. R. Cirrito, "Rapid in vivo measurement of β-amyloid reveals biphasic clearance kinetics in an Alzheimer's mouse model", The Journal of Experimental Medicine **213**, 677 (2016).

[342]F. Yang, G. P. Lim, A. N. Begum, O. J. Ubeda, M. R. Simmons, S. S. Ambegaokar, P. Chen, R. Kayed, C. G. Glabe, S. A. Frautschy, and G. M. Cole, "Curcumin inhibits formation of amyloid beta oligomers and fibrils, binds plaques, and reduces amyloid in vivo", The Journal of biological chemistry **280**, 5892–5901 (2005).

[343]A. R. A. Ladiwala, J. C. Lin, S. S. Bale, A. M. Marcelino-Cruz, M. Bhattacharya, J. S. Dordick, and P. M. Tessier, "Resveratrol selectively remodels soluble oligomers and fib-

rils of amyloid Abeta into off-pathway conformers", The Journal of biological chemistry **285**, 24228–24237 (2010).

[344]A. S. Wentink, N. B. Nillegoda, J. Feufel, G. Ubartait, C. P. Schneider, P. De Los Rios, J. Hennig, A. Barducci, and B. Bukau, "Molecular dissection of amyloid disaggregation by human HSP70", Nature 2020 587:7834 **587**, 483–488 (2020).

[345]K Ono, M. M. Condron, and D. B. Teplow, "Structure-neurotoxicity relationships of amyloid β-protein oligomers", Proc. Natl. Acad. Sci. U. S. A. **106**, 14745–14750 (2009).

Appendix A

Appendices from Methods Section

A.1 Additional Programming Packages

A.1.1 Ovito

The main file writing function **File Writing** finds particles of each particle type, and in turn, state and puts copies of these objects into lists. These lists are then passed to other file writing functions, **soluble state filewriter, beta state filewriter, intermediate state filewriter, trajectory file writer and sphere file writer**. Some of these could be doubled up with a little work but for the time, each is a different function depending on whether the object should be put into ovito format for a rod or a sphere. i.e. The **trajectory file writer** and the **beta state file writer** both require spherocylinder or rod configurations which include information on orientation and the shape properties of the objects. All other trajectory saving is a more simple save with spherical properties. It is necessary to save data surrounding patches to troubleshoot, check and benchmark simulations.

There are few softwares capable of visualising aspherical particles such as spherocylinders. Ovito is designed specifically for LAMMPS outputs which supports simulations with aspherical particles. Formatting output files for Ovito is identical to the output of .dump files for LAMMPS. There are specific properties required for anisotropic particles such as spherocylinders.

The required quantities are as follows: a centre of mass, COM, a unit vector in quaternion format,

(x, y, z, w) and a vector denoting the size of the quaternion, $(\sigma, 0, L)$ where σ is the radius of the particle, the y coordinate is ignored and L is the length of the particle along the z axis. Ovito formatting is different to the output of LAMMPS, or similarly, the simulations carried out in this project - therefore a column mapping must take place to match each quantity correctly. Patch locations are denoted by cylinders or spheres depending on the state the spherocylinder is in.

A.1.2 GTE Mathematics

This is a C++ library with a variety of mathematical functions and algorithms. -Unsure of whether to write information about shortest distance algorithms here or in the other functions where these algorithms are called.

A.2 Random Number Generator

The random number generator used is the built in C++ generator from the "random" package. As shown in figure A.1, the distribution of random numbers over 1000 bins is uniform, with a variance of CALCULATE VARIANCE which is not statistically significant.

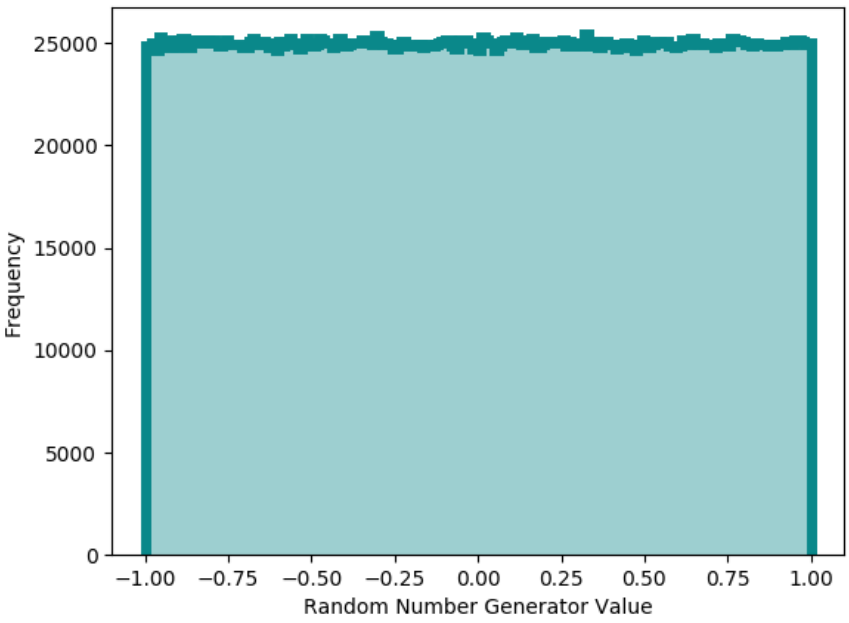

Figure A.1: Histogram of random number generator of a uniform distribution $[0, 1)$ with a fixed bin width of 0.01.

By executing 5000000 rotations on a single rod we can show that there is no bias in orientation. Each of the unit vector directions x, y, z finds orientations within the expected range $-1 < s < 1$.

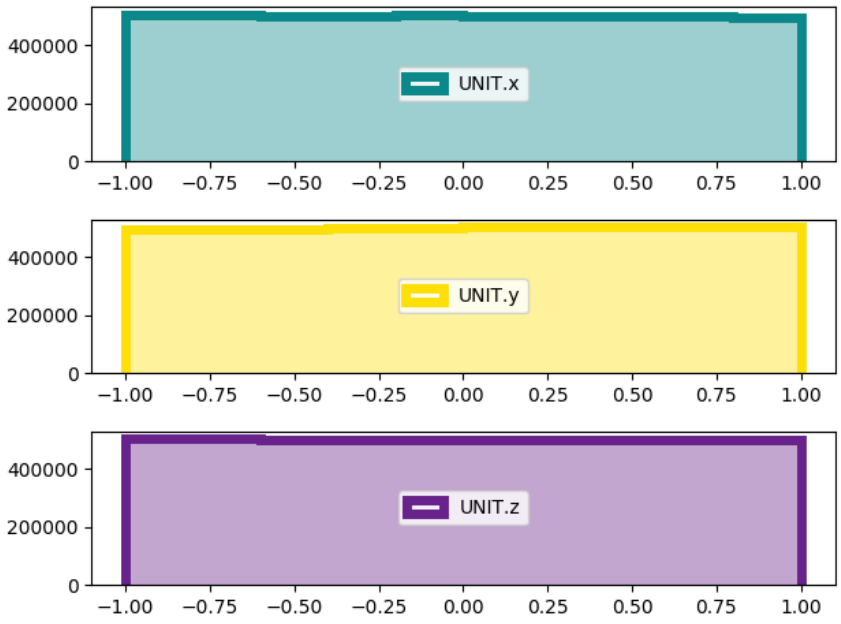

Figure A.2: Histograms of unit vectors for a single particle rotated 5 million times.

A.3 Code Structure

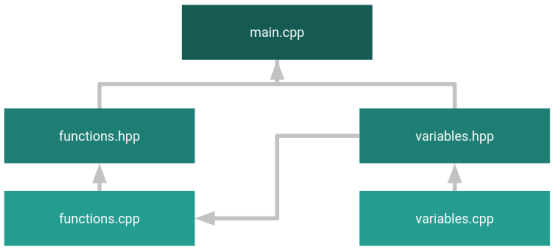

Figure A.3: The overall code structure and workflow for the DMC code we have written. Arrows denote an "include" in another file. e.g. *"variables.hpp"* is included in *"functions.cpp"*. All ".cpp" must be included in their respective ".hpp" files in order for the code to function.

The code structure is displayed in figure A.3. This structure is quite simple for the time being, however in future I would like to separate this further into files containing functions that specifically

manage translation and rotations, for file writing and for any pre processing.

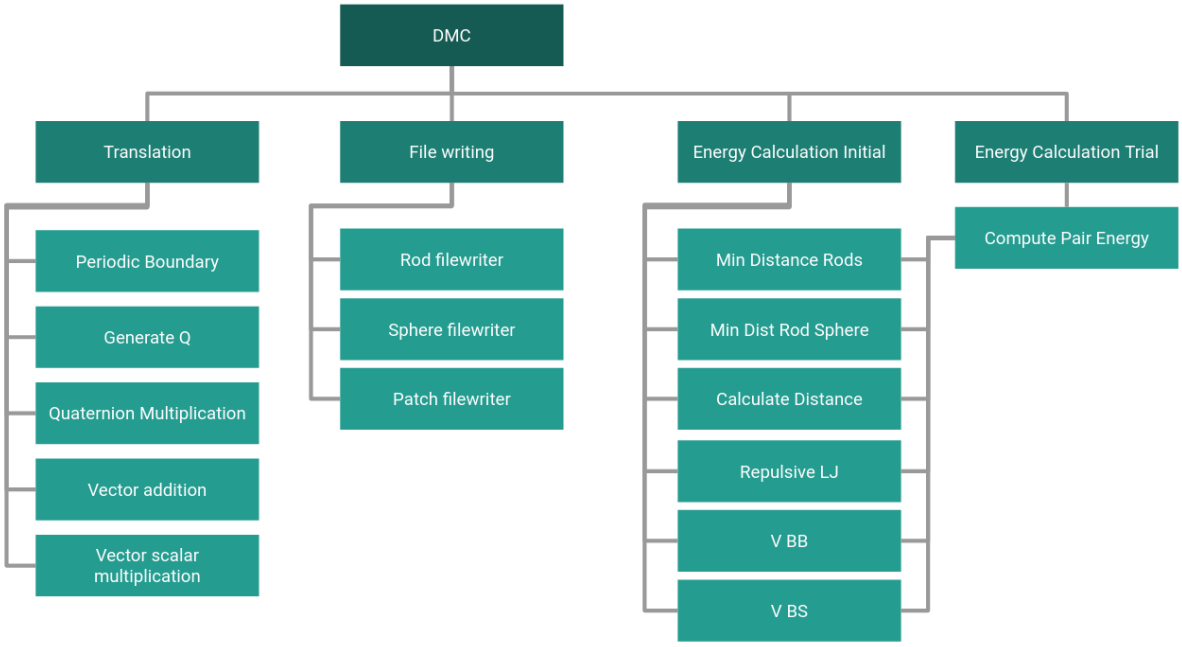

Figure A.4: A dependency tree for the functions called in the "DMC" function of the code which executes the main simulation. Several functions are called by more than on parent function and are highlighted by lines from the functions(s) to their parent function(s).