TITLE

Authorlist

Hospital for Sick Children, …

Outline for results:

1) **irreversible surface binding**
2) **Dehydration (heatmaps of SC vs BB hydration)**
3) **Peptide extension (end-to-end distance plots) – ok**
4) **Hydrogen bonding (time series, replacement of bound waters, preference for longer d_ete, etc.)**
   → compare to trend without octane (aggregation at high concentration in water)
5) **Start with aggregates here (put graphs of ete vs. HB type)**
   (Aggregate size/shape/morphology/etc – basic graph with # of aggregated vs. # of unaggregated peptides / number of total aggregates per time) **
   Insert heatmaps here of the backbone hydration vs. inter HB
6) **Secondary structure (beta structure formation, structure of sheets)**
   → compare to trend without octane

Intro bits and pieces:

On rationale:

→ discuss amyloid formation at interfaces (including the recent paper by Lee et al)

Discussion:

Add Lee et al 2008

Discuss Knecht 2008 (hairpin folding @ air-water interface)

Abstract

Introduction

The β-sheet is one of the most common structural elements in folded proteins. Sheets are diverse, and able to form a variety of structural elements, such as β-barrels, greek key motifs, and extended sheets, so understanding the driving forces behind their formation is an important task. The determinants of β-sheet formation are made even more interesting because of the role that β-sheet structures formed by misfolded peptides play in amyloid diseases, including Alzheimer's disease (Aβ and Tau), Huntington's disease (poly-glutamine), and Parkinson's disease (α-synuclein)[1].

Many different patterns of amino acid polarity are observed within β-sheets, but the most common patterns typically have a periodicity of two. Of all of these patterns, the pattern with alternating polar and non-polar residues has been found to be particularly significant[2]. These types of patterns are generally found on β-sheets found at the interface of water and the protein's hydrophobic core. Because these regions are essentially interfaces between a polar phase (the water) and a comparatively non-polar phase (the hydrophobic core), these sorts of interfaces may play some role in helping these segments of proteins to attain the correct conformation. Another interesting aspect of alternating sequences is that they have also been shown to be capable of aggregating and forming amyloid-like fibrils [3]. Because of this, alternating polar-nonpolar sequences are an interesting test system, as they appear to be relevant both to β-sheets in correctly folded, non-pathological proteins, and to β-sheets as found in pathogenic aggregation with amyloid fibrils.

Amyloid fibrils are thin, elongated, insoluble structures often found in extracellular plaques around affected cells [1]. Structural studies have shown that mature amyloid fibrils consist of many copies of aligned short peptide fragments, usually between 6 and 12 residues in length,

forming the characteristic structure of the amyloid fibril, known as a cross-β sheet. These structures often involve many intertwined strands of cross-β sheets [4, 5, 6]. Fragments of many known amyloid disease peptides, including ones from Aβ and α-synuclein, have been shown to form this same structure, despite having vastly different amino acid sequences [7]. Furthermore, stretches of as few as six residues from amyloidogenic peptides have been found to induce apoptosis in cell culture [8]. In the case of calcitonin, a four residue peptide was found to have amyloid-forming activity[9]. Furthermore, recently proposed models for IAPP/amylin and α-synuclein both suggest that the amyloid forms of these peptides consist of short stretches of β-sheet connected by hairpin loops[10,11]. All of this evidence suggests that short sequences are able to form amyloid, and that short sequences provide good models of β-sheet assembly and amyloidogenesis.

Specific properties of the cross β-sheet, such as its diameter and the number of strands wound together into the mature fibril, are dependent on the peptide sequence, although the general structural motif remains unchanged. Structural studies of fibrillization of different short peptides featuring repeats of only one amino acid (such as polylysine) have found that all these were capable of aggregating and forming fibrils, and that the sequence only affected the packing and the register of the mature fibril, while retaining the characteristic cross-β structure[14]. Recent crystallization experiments with fragments of several amyloidogenic peptides showed all of them to have the same amyloid structure, except with sequence-specific differences in packing and register[7]. Because there is a vast body of evidence showing that amyloid formation is triggered by short sequences, the behaviour of these short sequences can be used as an appropriate model for understanding the overall processes and fundamental mechanisms governing amyloid formation and toxicity.

One factor that has been found to affect the rates at which amyloid fibrils form is the presence of membranes, which are essentially a more complex case of a water-nonpolar interface. For example, the presence of sub-micellar concentrations of SDS was found to promote conversion of PrP (Prion protein) into its infectious form and increase its aggregation [15]. Several studies have found that lipids are capable of driving the formation of amyloid fibrils. In particular, liposomes with a high content of negatively charged lipids have been found to promote amyloidogenesis in proteins as diverse as insulin, G3P dehydrogenase, myoglobin, and Aβ [12,16]. Acidic liposomes were also found to accelerate fibril formation in a fragment of the immunoglobin light chain, as well as in the Aβ peptide \cite{zhu2002sca} \cite{mclaurin1996} \cite{yip2002apa}.

Membranes are more than just scaffolds: studies show that the toxicity of several amyloid species, including α-synuclein, has been associated with binding and permeabilization of liposome vesicles[17]. Membrane binding was also found to increase the rates of α-synuclein aggregation as well as induce cell death ~\cite{lee2002mbalpha}\cite{jo2000alphasm}. Aβ has also been found to affect membrane properties by facilitating formation of domains, rather than through direct formation of channels[18]. Fibrillization of IAPP, a protein found in diabetes, has also been shown to involve uptake of lipids, leading to cytotoxicity \cite{sparr2004iap}. The E. coli protein HypF has also been found to interact with phospholipid bilayers and cause destabilization, especially in the presence of anionic lipids[19]. In nearly all of these cases, however, mature fibrils have been found to be inert. Membranes may not always catalyze amyloid formation, as there is also some evidence that lipid binding by peptides like α-synuclein can, in certain cases, stabilize the peptide and inhibit toxicity, suggesting that membranes do not always lead to unfolding [20]. This suggests that the toxicity of amyloid is intimately related to

membranes and to oligomers formed in the early stages of amyloidogenesis. Furthermore, it suggests that interactions with interfaces may also be related to the observed toxicity of some amyloid species.

Although the final structure of amyloid fibrils is known, the early steps of amyloid fibril formation remain unclear. These are of interest as studies have shown that the toxicity of oligomers, that is, the early intermediates of aggregation, is far greater than that of mature fibrils[21]. Oligomer toxicity has also been shown for oligomers of fragments of amyloidogenic peptides that are only six residues in length, suggesting that amyloid toxicity results from a shared structural property, rather than from specific properties of particular peptides [8].

Though much is known about the structure of mature amyloid fibrils, the mechanism by which they form, the factors that help in their formation, and the intermediates through which they must pass are still unclear. Models of amyloid formation are generally based on the principle of nucleation and propagation, and require the formation of a so-called "critical nucleus"[1]. These models often include the concept of seeded aggregation, that is, the requirement to form an aggregate of a particular size before substantial amounts of aggregation can occur[22]. In vitro, amyloid formation has been experimentally shown to be partly reversible, especially in its early stages, suggesting that aggregation is a reversible equilibrium reaction[23]. Also, recent work by Xue and coworkers suggests that rearrangement via fragmentation plays an important role in the kinetics of amyloid formation[24]. There is also experimental evidence that elongation may proceed by the linear addition of monomers to the end of the elongating fibril[25]. Although experimental data for aggregation and amyloid formation on lipid-water interfaces are limited, existing studies of this subject suggest that interfaces amplify kinetics, although they are limited in the amount of amyloid they can form because of density-dependence and limited surface

area[26]. Understanding the ways in which interfaces modulate the formation of β-sheet could help better understanding of the *in vivo* mechanisms by which amyloid is formed, as the presence of lipid membranes and crowding within the cells could potentially have significant repercussions on the mechanism by which amyloid forms.

Small peptide fragments appear capable of forming amyloid in a relatively sequence-independent manner, and there is evidence that membranes may affect the kinetics of this process. As amyloid may have a common mechanism of formation, elucidating what is happening at the atomic level in the early stages of fibril formation would improve our understanding of the structure of early aggregates, and potentially help us to glean insight into their mechanisms of toxicity. Since early aggregates are transient and very small in size, molecular dynamics is an ideal technique for studying this aspect of amyloid formation, as MD can give us insight into what is happening at the atomic level. Because of the level of detail attainable with MD, it may also help us to understand the interactions between peptides and membranes, and help clarify the role of hydrophobic interfaces in the process of amyloidogenesis.

Molecular dynamics simulations of amyloid have revealed much information about the nature of mature fibril structure and stability, and the formation of small oligomers in water. Furthermore, results from coarse-grained simulations have shown support for the nucleation-propagation mechanisms suggested by in vitro data. Many of these simulations have been performed not only on full-length amyloidogenic peptides such as Aβ(1–42), but also on various short amyloid-forming peptides. Although long-scale studies have been useful for evaluating the kinetic relevance of nucleation-propagation models[27] and improving our understanding of fibril structure, the molecular basis of interactions between interfaces and peptides is still not well

understood. Although recent simulations have investigated the interactions of amyloid-forming peptides with membranes or interfaces, they did not examine aggregation dynamics[28,29]. Our work is the first thorough molecular investigation into the role of interfaces in the formation of β-sheets.

Although there are many different short fragments that we can study, we decided to focus on the sequence GV4. This sequence has a very similar motif to the sequence $(GGVGV)_n$, a fragment of elastin that has been experimentally shown to form fibrils [30]. Moreover, GV4 is a similar structural motif to the sequence $(GA)_n$, a common component of many rigid spider silks that are also known to have an amyloid-like structure [31]. Furthermore, as it consists of an alternating polar-non-polar sequence, it represents a pattern commonly found in natively β-sheet structured proteins[2]. Previous simulations from our lab have predicted that an extended $(GV)_n$ repeat is capable of forming amyloid [32]. Herein we intend to investigate the binding to an octane slab and the aggregation of the GV4 peptide. Through all-atom molecular dynamics simulations, we hope to glean more insight into the role of interfaces in intermolecular β-sheet assembly, a process that is related not only to amyloid pathology, but which may also be relevant to proper folding of β-sheets in proteins.

Methods

Simulation parameters:

All simulations were performed using GROMACS 3.3.1[33,34]. Proteins and octane were modeled using the OPLS-AA forcefield[35] and the TIP3P[36] water model was used. The simulation was run using a leapfrog algorithm with a 2 fs integration timestep. Electrostatic interactions were calculated using PME, with a gridsize of 0.12 and updating of the neighbourlist every 10 steps [37]. Lennard-Jones interactions were treated using a twin-range cut-off with a short-range and long-range cut-offs of 0.9 and 1.4 nm respectively. Pressure coupling was semi-isotropic, using the Berendsen barostat and a coupling constant of 4 ps, with a fixed Z-axis to prevent contraction of octane [38]. Protein, octane and water atoms were placed in separate temperature coupling groups, and temperatures were coupled independently with a coupling constant of 0.1 ps. Constraints were treated with LINCS[39], and water was treated with the SETTLE algorithm [40]. Each simulation was position restrained and equilibrated for 400 ps at constant pressure before the production runs were begun. Simulations were performed as specified in *Table 1.*

Generation of the Octane Slab

The octane slab used in Simulations B through D was generated by tiling octane atoms to form a slab at the center of the box, and then hydrating the system. The resultant system was energy-minimized and simulated for 600 ps under isotropic pressure coupling. This was enough time to allow waters that had been placed within the slab to return to the aqueous phase and for the octane molecules to reach equilibrium.

System preparation:

All GV_4 peptides were capped with amide and acetyl groups to stabilize the peptides and reduce the repulsive effects of terminal charges on dimer formation. The initial conformations

for the peptides were taken from an ensemble generated at 296 K by simulated tempering within our laboratory. Peptides and octane were joined together and hydrated using either the GROMACS utilities (for most systems) or an in-house method (for Simulation A) to keep water out of the slab. The resultant systems were energy minimized and position restrained at NPT for 400 ps. After this, the production runs were started.

To allow peptides to accumulate on the surface while minimizing the effects of aggregation in water (e.g. in Simulation C), peptides were gradually titrated into the solution. Simulation C was generated by titrating 8, then 16, then 18 peptides onto the slab, allowing them to equilibrate until all peptides had adsorbed to the slab (which typically took about 10 ns). Measurements were started after 32 ns and 3 additions of peptide.

The systems in Simulation D were generated by removing the slab from the starting conformations of runs in Simulation B1 and translating water from the bottom edge of the box upward to fill the space created by the removal of the space. This way we created a system that had a smaller Z axis but was otherwise equivalent to the water phase of the slab systems of Simulation B1. Although the system was equivalent to the water phase of B1, its concentration was effectively double of that of the slab phase of B1, as there is only one effective surface to the two approximately equal-sized surfaces of the slab.

Analysis

i) Distance

Two different distance metrics were calculated for the peptides: the distance of the center of mass of the peptide from the slab on the Z axis (a measure of adsorption) and the end-to-end distance d_{ete}, which is the three-dimensional Euclidian distance between the first and last Cα atoms within the peptide, and provides a measure of extension. To determine the location of the

slab, the density of octane was measured in sections along the Z axis, normal to the slab interface, and the locations where the density crossed 650 kg/m^3 were used as markers for the starting and ending points of the slab.

ii) Axial distribution

In order to compensate for potential ruggedness on the octane surface, axial distributions were computed by computing the distance from the octane slab surface using a Voronoi tessellation-based method similar to that of Pandit et al[41]. Because octane is not ordered in the same way as a bilayer, as its atoms are uniformly distributed throughout the octane phase, the octane C atoms were put on a grid of a set size (in this case 1 nm x 1 nm). Then, the highest and lowest octane C on each point in the grid was chosen, and those atoms were used to construct the tessellation. This method provides an improved estimate of distances from the surface of the octane compared to methods that use the center of mass of the slab phase to calculate the axial distribution of groups to the slab surface.

iii) Hydrogen Bonding

Putative hydrogen bonds were identified as any contacts where the distance between putative donor and acceptor molecules was less than 0.35 nm and that between the hydrogen and the acceptor was less than 0.25 nm. Furthermore, the angle Donor-Hydrogen-Acceptor was calculated and only accepted if it was less than 60°. In addition to these geometrical restraints, the DSSP hydrogen bonding energy criterion was considered.[42] A bond matches this criterion if its bond energy is in the range of –0.5 to –9.9 kcal/mol. If a putative hydrogen bond matched all of the above criteria, it was accepted as a hydrogen bond.

iv) Secondary Structure

The levels and types of secondary structure in the samples were predicted using the DSSP algorithm[42], with adjustments to distance calculations in order to account for the presence of periodic boundary conditions.

v) Hydration Analysis

Backbone hydration:

The hydration of the peptide backbone was analysed by calculating the number of hydrogen bonds to water for each peptide backbone group (either CO or NH). The hydrogen bonds were calculated using the same criteria outlined for backbone peptide-peptide hydrogen bonds, and were subsequently binned by simulation timepoint, peptide, and backbone group. The data, segmented in this way, were used to compare backbone hydration with other peptide properties, such as d_{ete} and secondary structure.

Side chain hydration:

For GV4, hydration of the valine side chains was also measured, by measuring the hydration of the Cγ atoms (the two terminal atoms of the valine side chain). A cut-off of 0.43 nm was chosen because it corresponds to the first minimum in the radial distribution function of water oxygen to the Cγ atoms of valine (data not shown), and encompasses the first hydration shell around the methyl group. Because carbon atoms cannot form hydrogen bonds to water, the hydration analysis for the side chains consisted of measuring the number of water oxygen atoms within the cut-off for each Cγ.

vi) Other analyses

Tools from the GROMACS toolkit were used to perform all other analyses and distance calculations (except d_{ete} in Simulation A, which was calculated without PBC because of the small system size). Images were generated with UCSF Chimera[43].

Kinetics were calculated …

vii) Error Analysis

Standard error of histograms was calculated using block averaging[44], and the error bars represent the upper and lower bounds of the 95% confidence interval of the mean.

Results

General Results

All of the simulations we performed with GV4 in the presence of the octane slab showed an irreversible binding behaviour of GV4 to the slab. [INSERT Sim A results] Because GV4 is a disordered peptide, it was found to adopt a large number of different conformations when it bound to the slab. Some of those conformations are shown in *Figure 1.* These can be loosely separated into three categories based on the separation of the ends of the peptide, hereafter referred to as the end-to-end distance (d_{ete}). Conformations with a very narrow separation of the ends (left panel), exhibit some hairpin-like conformations, although they do not always have hydrogen bonds. Those with intermediate end-to-end separation show a diverse set of conformations, some of which show hydrogen bonding and resemble loose turns of various sorts, and others, which do not show hydrogen bonding at all. The last class of conformations (right panel) are extended, and do not have any hydrogen bonding.

After testing the binding of monomers to the slab, we decided to investigate the effects of placing several peptides on the same leaflet of the slab, and looked at aggregation over a longer timescale. Figure 14 shows some snapshots from one of the runs of Simulation B2. As the simulation progresses, hairpins and short β-structure segments consisting of either two extended peptides or an extended peptide and a hairpin become visible. Qualitatively, β-structure increases as the simulation progresses – this increase is most notable when comparing early snapshots in the simulation, such as (a) and (b) to later ones, including (c) and (d).

Figure 13 shows some of the extended beta sheets observed after 219 ns of simulation time. There are numerous extended β-sheets, forming higher-order structures including trimers and tetramers. The GV4 peptide is shown to form both parallel and antiparallel β-sheets. The

amounts of β-sheet shown in Figure 13 and Figure 14 are much higher than that observed in Figure 15, which is the result of one of the runs of Simulation D, after it was extended for another 160 ns. Despite the timescale being similar to that in Simulation C2, the amount of β-sheet that is visible in the cluster is less, suggesting that the efficiency of formation of β-sheets in water is not particularly high.

→ snapshots from A,B,C,D [move up from second part into this part] --- water and slabs

i) GV4 binds to a model octane slab

Simulations A and B1 revealed that the GV4 peptide has a characteristic binding behaviour to the octane slab. Binding to the slab is generally initiated by interactions between one of the valine residues and the slab (Figure 1a). Once the peptide is fully bound all four valine residues in the GV4 peptide interact with the slab (Figure 1b). Furthermore, continued interaction with the slab promotes formation of extended conformations by the peptide. Most of the peptide adsorption to the slab occurs fairly rapidly, with 99% of all chains bound by the end of Simulation B1 at 10 ns (Figure 2). Interestingly, the number of chains adsorbed to the slab follows standard first-order kinetics. Fitting the curve to a first-order kinetics equation suggests a value of K of between 120-200 (!!!), suggesting that the binding to the slab is a weakly favourable process with a ΔG of about 12 kJ/mol. This suggests that even though binding is weakly favourable, possible cooperativity between the valine residues makes the process effectively irreversible due to the difficulty of four valines dissociating from the slab within a short enough period for the peptide to detach.

Figure 3a shows a sample density profile, taken from Simulation C2. The peptide chains are all distributed very close to the octane, in the interface between the octane and water. Axial distributions of the valine Cγ atoms and the Cα atoms of the chains for Simulation C2 show that

the valines are distributed asymmetrically, and that their distribution is skewed toward the slab, with fewer than expected on the other side of the Cα atoms. In comparison to the Cγ, the Cα atoms are distributed in a fairly symmetric way, centered at 0.5 nm from the slab. Furthermore, the main chain atoms are skewed slightly away from the octane in comparison to the Cα atoms, likely due to the relatively polar character of the peptide backbone.

ii) Slab binding dehydrates the GV4 peptides

We decided to also investigate the effects of slab binding on hydration of the backbone and the valine sidechains. A cumulative comparison of the sidechain and backbone hydration is shown in Figure 4. Monomers on the slab (a) show complete dehydration of side chains, which suggests that dehydration of side chains is independent of aggregation, and may be a general characteristic of peptide-interface binding. In comparison, the aggregated peptides in water show some loss of backbone hydration, having similar levels of hydration to the monomers on the slab, but their side chains are not dehydrated like those of peptides on the slab. In the results for Simulation B1, there are two distinct populations, one with a high level of side chain hydration (about 5 waters around each Cγ) and high backbone hydration (about 16 waters per chain), and another with low levels of side chain hydration (centered at 1) and low backbone hydration (12 waters per chain). The hydrated population corresponds to the monomers in water, and suggests that slab binding causes a decrease in the number of bound waters, as well as dehydration of valine side chains. Results for Simulation B2 are similar, with more dehydration observed due to the longer timescale. On average, the sidechains in simulation B1 and B2 appear slightly less dehydrated than those in simulation A. This is in part due to the shorter equilibration times after slab binding in B1 and B2 compared to A, and dissipates over time, as can be seen from the results of the longer simulations in C1 and C2. Peptides in Simulation C1 and C2 resemble those

in Simulation B2, but are more dehydrated due to the longer simulation timescale and greater concentrations.

iii) Peptide binding to the slab alters the conformational equilibrium of GV4

We used the end-to-end distance distribution for the GV4 peptide as a representation of its conformational equilibrium. *Figure 5* shows the distributions of end-to-end distance for Simulations A through D. All simulations on the slab have a strong peak at around 0.66 nm, with the peak being most prominent for Simulation A, and decreasing in prominence for C1 and C2. At the same time, a plateau becomes prominent for conformations of 2 nm or longer, especially in simulations C1 and C2, as well as in B2, albeit to a lesser extent. This suggests that there is a slow exchange of closed conformations into extended conformations that persists over longer timescales. In contrast, the simulations in water have few conformations of 1.3 nm or longer, and have many conformers in the mid-range, likely corresponding to hydrogen-bonded turns. The conformations in simulation D differ slightly from the monomers in water, but, unlike the simulations of aggregates on the slab, do not exhibit a plateau in the 2.0 nm and over range. The additional bumps in the conformational distribution are likely due to the comparatively smaller amount of sampling in Simulation D in comparison to the other experiments.

To examine the evolution of the distance distribution over time we binned the end-to-end distances over periods of 500 ps. The results of this analysis for Simulations B1 and B2, shown in *Figure 6*, reveal that the extension is a rapid process, and the bulk of the extension occurs within the first 10 ns, by transitioning of conformers from the intermediate range into extended conformers. Although conversion from hairpins to extended forms can also occur, this is most likely to occur in two stages, and may involve initial conversion of the hairpin form into an

intermediate form. The conversion of the hairpin into more extended conformations likely takes longer, which is why it is difficult to see at short timescales like those explored in B1 and B2.

To investigate the possibility of conformational exchange more thoroughly, we quantified the proportions of conformers in the closed/hairpin, intermediate, and extended conformations for all simulations (Figure 7). For simulations on the slab, the proportion of peptides in the closed/hairpin conformation decreases as the concentration and time spent on the slab increases, and it is most prominent in simulation C2. In contrast, the simulations in water maintain a level of the closed conformation similar to that of the monomers on the slab, regardless of their aggregation state. Another important distinction between the slab and the water systems can be seen in *Figure 7b*, as the monomers on the slab show distinctly more extended conformations than peptides in water, regardless of their aggregation state. This effect increases along with time and concentration, and is highest in Simulation C2. This all suggests that conversion of intermediate forms into extended forms is a characteristic of GV4 binding to the slab that is observed in neither monomers nor aggregates in water. Furthermore, the frequency of hairpin conformations decreases over time, suggesting that an conversion of hairpins into more extended conformations is indeed occurring as aggregation proceeds. It is also interesting to note that there is no significant difference in the proportion of hairpin-like conformations observed in aggregates in water as compared to monomers in water and monomers on the slab.

iv) Time-evolution of hydrogen bonding

In addition to affecting the conformational equilibrium of the peptides, binding to the slab also affects the way that they form hydrogen bonds. In simulation B1 and B2, there is a linear increase in the number of hydrogen bonds formed by the peptides, resulting in almost 1.5 intermolecular hydrogen bonds formed by each chain by the end of the 30 ns. In contrast, the

number of intramolecular hydrogen bonds stays firm at about 0.6 bonds per chain, suggesting that slab binding has little overall effect on the quantity of intramolecular hydrogen bonds in the ensemble. Simulation C2 (Figure 8b) represents a longer-scale and higher-concentration system, and exhibits a far higher level of hydrogen bonding, with up to 8 hydrogen bonds per chain by the end of the 219 ns simulation. Even after 219 ns, more hydrogen bonds still appear to be forming, at a relatively constant rate. Furthermore, the number of intramolecular hydrogen bonds per chain seems to dip slightly over time, suggesting that as the peptides form ordered aggregates, the level of intramolecular hydrogen bonding gradually drops from its original state. Aggregation in water (Figure 8c) occurs with a more rapid formation of hydrogen bonds than observed for Simulation B1 and B2. One possible reason for this discrepancy is the fact that slab binding effectively dilutes the peptides, by distributing approximately four peptides on each slab surface. Another possibility is that at the concentrations in water that were tested, there are numerous possible geometries through which peptides can interact and form hydrogen bonds, many of which may become less favourable once the peptides are bound to the slab surface. The initial aggregation stage in water tapers off after about 15 ns, after which hydrogen bonds rise at a much lower rate, potentially due to rearrangements within existing aggregates rather than formation of new ones.

We also investigated whether certain conformations were more likely to participate in particular hydrogen-bonding interactions than others. To this end, we separated the end-to-end distance distribution into four mutually exclusive categories: peptides with only intermolecular hydrogen bonds, those with only intramolecular hydrogen bonds, those participating in both types of hydrogen bonds, and those with no hydrogen bonding at all. The results of this analysis show that the four classes of conformations tend to form two clusters. The peptides with either

no hydrogen bonding or only intermolecular hydrogen bonds tend to be distributed in the more extended portion of the end-to-end distance range and show similar distributions. However, the peptides with only intermolecular bonds form a pronounced plateau at conformations > 2 nm, which is especially noticeable in the results from Simulation C2. In contrast, the peptides with either both types of hydrogen bonds or only intramolecular bonds exhibit very similar distributions, likely due to the significant contribution of hairpin-like conformations to these two populations. Another notable distinction is that the divergence between the two distributions appears the greatest in Simulation C2, with several peaks appearing in the distribution for peptides with both types of hydrogen bonds, suggesting that the states with multiple types of hydrogen bonds may act as intermediates.

v) Aggregation occurs while on the slab

Because of the large number of intermolecular hydrogen bonds formed, we decided to examine the rate at which peptides were aggregating. To do this we counted the number of peptides that had intermolecular hydrogen bonds, and then calculated these as the fraction aggregated. In Simulation B1 and B2, about 50% of all peptides had aggregated by the end of Simulation B1, and aggregation continued at a slower rate once all peptides were on the slab (Figure 10a). After a longer period of time on the slab, as in Simulation C2, the fraction of peptides aggregated was close to 100%, that is, all peptides were participating in intermolecular bonds.

To further investigate the nature of hydrogen bonding and examine the hydration of the aggregates, we compared the backbone hydration of the peptides to the number of hydrogen bonds they formed, using the same criteria as in Figure 9. There is an anticorrelation between the number of hydrogen bonds formed and the number of backbone bound waters, suggesting that

the bound waters are being replaced by hydrogen bonds formed within the aggregate. The number of hydrogen bonds increases as the timescale and concentration of the peptides increases – note the gradual increase in the maximum from simulation B1, to simulation C2. Another interesting trend occurs in the peptides with both intramolecular and intermolecular hydrogen bonds. The relative frequency of peptides that fall into this category seems to increase from simulation B1, to B2, and onwards to C1. However, it drops off in simulation C2, even though the peptides in this category do show the largest number of hydrogen bonds in this particular simulation. This suggests that the peptides with both types of hydrogen bonds are acting as an intermediate. Their frequency increases early on in the aggregation process, and decreases gradually as aggregates mature and intramolecular hydrogen bonds are dissociated. The overall trend when these two types of peptides are combined is a gradual increase in the number and distribution of hydrogen bonds, with most chains having between 4 and 12 hydrogen bonds in simulation C2, as compared to 0 to 4 hydrogen bonds in simulation B1. A similar analysis performed for simulation D shows results similar to those for simulation B, except with the base (0 hydrogen bonds) occurring at a higher number of bound waters, suggesting that the mechanism occurring is similar, but it is occurring from a different baseline level of dehydration (data not shown).

v) Structure of aggregates

The large number of intermolecular hydrogen bonds seen in simulation B1, and even more so C1 and C2 suggested that the aggregates may have converted into a more ordered form. To investigate the ordering of the aggregates, we examined the secondary structure of the systems. We found that there was a notable increase in the amount of extended β-sheet structure that was formed both in Simulations B1 and B2, as well as in C1 and C2 (Figure 13). In

simulations B1 and B2, the levels of β-sheet increased from none initially to 4% of all resides, and up to 6% after the end of simulation B2. The levels of β-structure observed in Simulation C1 and C2 were even more striking, with the levels of β-sheet going from 10% at the 32 ns mark all the way up to almost 30% of all residues after 219 ns of simulation. Notably, the levels of β-bridges remained low, at around 5% of all residues, in all simulations. Because extended β-sheets are essentially β-bridges with more than two involved residues, the constant levels of β-bridges suggest that these structures may act as potential substrates for further formation of extended β-sheets. These were the main contributors to hydrogen-bonded structure in these peptides, as the total contribution of all other hydrogen-bonded structural types (e.g. turns, helices) in Simulation B2, C1 and C2 was less than 5%.

Discussion

Implications of Slab Binding

Binding of the GV4 peptide to the slab is an interesting result, and the driving forces behind this are easy to determine. For example, hydrophobic interactions between the alkyl groups in the side chains of amino acids like valine and the methyl groups in octane are likely to drive association of the side chains and the octane clump. Although we can look at octane as a model of what happens at a lipid-water interface, binding of peptides to clumps of saturated hydrocarbons in water also has merit. A recent study by Lee et al treated α-synuclein with hexane granules, and found that this led to rapid formation of amyloid fibrils[45]. The present study suggests a theoretical explanation for the effects observed in these experiments, as the hexane may act as a scaffold and induce some conformational change in the α-synuclein, just as the octane slab does for the GV4 peptide. Although the effects of binding to a water-nonpolar interface likely depend on the character of the side chains involved, we believe that it is likely

that such binding may induce local conformational changes in certain portions of the peptide, priming them for the formation of β-sheets.

Furthermore, the binding of peptides to an interface such as the slab acts to simultaneously dilute and concentrate them. For example, in Simulation B2, we have about 4 peptides on each side rather than 8 in one block of water – effectively diluting the concentration of peptide by 2. Yet, as peptides are titrated onto the surface, and due to the decreased number of degrees of freedom available to the peptide while it is on the surface, the local concentration of peptide increases. So in this manner, bidning to an interface performs two roles. At first, it dilutes the peptide in solution by taking peptides away from the water phase. As peptides are titrated into the solution, however, the concentration increases, and a very high local concentration of peptide can be reached in this way. Because of these concentration effects alone, the binding to the slab may act to promote the formation of β-sheets by increasing the local concentration of peptide.

The Slab as a Model

Although the slab directly models the effects of binding of hydrophobic side chains to alkanes it is also of value as a model. There are two main types of biologically relevant polar-nonpolar interfaces: the lipid bilayer, and the surface of a molten globule. Lipid bilayers differ from the slab in that they have polar and charged headgroups, and have low density and relatively high fluidity in the middle. TALK ABOUT LIPIDS

→ charge repulsion – bad – decreases activity

→ rugged surface – bad – decreases kinetics

Another biologically relevant situation that the slab may help to model is that of a molten globule. Globular oligomers of peptides have been found in many amyloid-forming peptides, and

in some instances have been thought to be a precursor to the mature fibril stage, although that is debatable. Because it is hydrophobic, behaviour on the slab may represent the behaviour of peptides at the surface of large proteinaceous aggregates. Furthermore, many models of protein folding consider hydrophobic collapse an important early step in the process. Once a protein has collapsed into a molten globule, then its exterior will in fact be in an environment rather similar to a polar-nonpolar interface, as the inside will be free of water. Furthermore, as many β-sheets, in particular antiparallel ones, in functional protein are formed near or at the water-nonpolar surface, this would provide another possible rationale for the ability of a water-nonpolar interface to drive their formation.

Relevance of GV4 peptide

The process of nucleation and propagation commonly assumed to be the in vitro mechanism of amyloid formation may occur differently when the monomers are at least partly adsorbed to a hydrophobic interface. There are two main additional entropic considerations that come into play when dealing with interfaces that act to increase rates of ordered β-sheet formation: loss of conformational space in the chain leading to increased prevalence of extended conformations, and the replacement of 3D diffusion processes with a 2D diffusion process on the interface.

As monomers on the interface, the GV4 peptides initially tend to adopt one of two main classes of conformations, a hairpin-like conformation, which makes up approximately 30% of monomers on the slab, or one of many extended conformations, contributing about 45% in the case of monomers, and up to 70% in simulation C2. The ratio of the two appears to depend on concentration, as when the slab is more densely populated, forming many hydrogen bonds to neighbouring extended peptides is more favourable than forming an internal hairpin. Along with

the trend of increasing extended conformations with time and concentration, these data suggest that the slab indeed does lead to a large increase in levels of extended conformations, which could contribute to β-sheet formation.

Extended forms can directly form ordered dimers by joining together, but our results suggest that the hairpin-like structures are capable of forming both intramolecular and intermolecular peptide-peptide hydrogen bonds simultaneously. Peptides exhibiting this class of bonding are surprisingly common, consisting of 18.9% of all peptides in Simulation C2. All of these processes are effectively equilibria, as hydrogen bonds are not permanent and interactions held together by only a few hydrogen bonds do not always last—the levels of fluctuation in the secondary structure and aggregate size metrics for GV4 support this view. Nevertheless, dimers do elongate, and eventually can form elongated sheet-like structures, like those that were observed in Simulation C2.

Diffusion along the slab as opposed to in water also contributes to the catalysis of β-sheets. When within a bilayer environment, peptides can remain monomers and be guided to each other through 2D diffusion on the slab. This has several benefits: it reduces the range of motion of the monomers, and discourages formation of amorphous aggregates. Alignment also increases the local concentration of the peptide, as the same number of units are aligned on a two-dimensional plane rather than a three-dimensional volume, reducing their entropy and increasing the likelihood of collision. In effect, the octane slab acts as a giant scaffold, and behaves as if it was the surface of a large amorphous aggregate, only unable to hydrogen bond. Because of this, the behaviour of peptides at the surface of the octane may be similar to their behaviour at the surface of large, amorphous, hydrophobic aggregates. The only appreciable difference would be that the octane surface is flatter, and may be a better scaffold for β-

sheet formation. Interestingly, in systems where concentration is higher than the maximum amount of peptide that can fit on that region of the slab, there are more amorphous aggregates and less ordered dimer formation. This may suggest that density on the slab (or concentration in solvent) determines the preferred path to formation of extended β sheets. **Brief outline for discussion:**

→ what about effects of membrane recycling and vesicular transport on all of this?

→ possible theories of slab binding

→ burial of hydrophobic residues

→ behaviour as the surface of a large (key word LARGE) globular aggregate (of something, not necessarily proteinaceous)

→ discuss ratios of peptide in each conformation, comparative ratios in water/slab/diff conc

→ more complex sequences would have slower kinetics (predicted) but fewer off-pathway forms because of the bulkiness of side chains and decreased number of possible stable aggregated sarrangements

→ this would support why the sub-micellar concentrations of SDS etc. would cause changes rather than micellar, because of the exposure of the hydrophobic portions of the surface

→ "like likes like"

→ emphasis on DISORDERED nature of peptide in these instances (that is, a majority of the peptides are in a "coil" conformation in all cases, even when there is additional b-structure)

→ possible mechanics of aggregation on the slab

→ diffusion

→ joining of hairpins + extended ➔ 3 extended (does that even occur?)

→ relate to known papers/theories of hairpin folding and sheet folding and important role of hairpins in folding and so on and so forth → significance of hairpins (putative intermediates)

[Also: general implications to amyloid formation and possibly toxicity; relevance to beta sheet formation in membrane proteins (beta barrels) and in soluble proteins; discussion of entropy: interface reduces entropy penalty of beta sheet formation in two ways: (1) by displacing the conformational ensemble of disordered states in water to fewer, more ordered conformations at the interface; (2) by replacing a 3D conformational search by a 2D one.

How do our results relate to two different regimes: (a) the low concentration limit, where addition of monomer to preformed fibril is the RLS; (b) the high concentration limit, where reorganization of aggregates is the rate-limiting step (see the recent Vendruscolo papers in which they used their "tube" model)?

Discuss limitations of our molecular model.

How do these results relate to the nucleation-propagation model?]

→ comparison to known experimental/theoretical work

Conclusion

Acknowledgments

1. Chiti, F. & Dobson, C.M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem* **75**, 333—366(2006).
2. Mandel-Gutfreund, Y. & Gregoret, L.M. On the Significance of Alternating Patterns of Polar and Non-polar Residues in Beta-strands. *Journal of Molecular Biology* **323**, 453-461(2002).
3. Broome, B.M. & Hecht, M.H. Nature disfavors sequences of alternating polar and non-polar amino acids: implications for amyloidogenesis. *Journal of Molecular Biology* **296**, 961-968(2000).
4. Jiménez, J.L. et al. Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *The EMBO Journal* **18**, 815—821(1999).
5. Serpell, L.C., Blake, C.C. & Fraser, P.E. Molecular structure of a fibrillar Alzheimer's Aβ fragment. *Biochemistry* **39**, 13269—13275(2000).
6. Nelson, R. et al. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **435**, 773—778(2005).
7. Sawaya, M.R. et al. Atomic structures of amyloid cross-bold beta spines reveal varied steric zippers. *Nature* **447**, 453—457(2007).
8. Pastor, M.T. et al. Amyloid toxicity is independent of polypeptide sequence, length and chirality. *J Mol Biol* **375**, 695—707(2008).
9. Reches, M., Porat, Y. & Gazit, E. Amyloid Fibril Formation by Pentapeptide and Tetrapeptide Fragments of Human Calcitonin. *Journal of Biological Chemistry* **277**, 35475—35480(2002).
10. Wiltzius, J.J.W. et al. Atomic structure of the cross-{beta} spine of Islet Amyloid Polypeptide (Amylin). *Protein Science* (2008).
11. Vilar, M. et al. The fold of {alpha}-synuclein fibrils. *Proceedings of the National Academy of Sciences* **105**, 8637(2008).
12. Zhao, H., Tuominen, E.K. & Kinnunen, P.K. Formation of amyloid fibers triggered by phosphatidylserine-containing membranes. *Biochemistry* **43**, 10302—10307(2004).
13. Dobson, C.M. Protein folding and misfolding. *Nature* **426**, 884—890(2003).
14. Fandrich, M. The behaviour of polyamino acids reveals an inverse side chain e□ect in amyloid structure formation. *The EMBO Journal* **21**, 5682—5690(2002).
15. Stohr, J. et al. Mechanisms of prion protein assembly into amyloid. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 2409-14(2008).
16. McLaurin, J.A. & Chakrabartty, A. Membrane Disruption by Alzheimer beta-Amyloid Peptides Mediated through Specific Binding to Either Phospholipids or Gangliosides. *Journal of Biological Chemistry* **280**, 30001—30008(1996).
17. Volles, M.J. et al. Vesicle permeabilization by protofibrillar α-synuclein: implications for the pathogenesis and treatment of Parkinson's disease. *Biochemistry* **40**, 7812—7819(2001).
18. Valincius, G. et al. Soluble amyloid {beta} oligomers affect dielectric membrane properties by bilayer insertion and domain formation: Implications for cell toxicity. *Biophysical Journal* (2008).
19. Canale, C. et al. Natively Folded HypF-N and Its Early Amyloid Aggregates Interact with Phospholipid Monolayers and Destabilize Supported Phospholipid Bilayers. *Biophysical Journal* **91**, 4575(2006).

20. Zhu, M. & Fink, A.L. Lipid Binding Inhibits α-Synuclein Fibril Formation. *Journal of Biological Chemistry* **278**, 16873—16877(2003).
21. Hartley, D.M. et al. Protofibrillar Intermediates of Amyloid beta-Protein Induce Acute Electrophysiological Changes and Progressive Neurotoxicity in Cortical Neurons. *Journal of Neuroscience* **19**, 8876(1999).
22. Come, J.H., Fraser, P.E. & Lansbury Jr, P.T. 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—5963(1993).
23. Cannon, M.J. et al. Kinetic analysis of beta-amyloid fibril elongation. *Analytical Biochemistry* **328**, 67—75(2004).
24. Xue, W., Homans, S.W. & Radford, S.E. Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proceedings of the National Academy of Sciences* **105**, 8926-8931(2008).
25. Collins, S.R. et al. Mechanism of prion propagation: amyloid growth occurs by monomer addition. *PLoS Biol* **2**, e321(2004).
26. Sharp, J.S., Forrest, J.A. & Jones, R.A.L. Surface denaturation and amyloid fibril formation of insulin at model lipid-water interfaces. *Biochemistry* **41**, 15810—15819(2002).
27. Nguyen, H.D. & Hall, C.K. Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proc Natl Acad Sci US A* **101**, 16180—16185(2004).
28. Jang, H., Zheng, J. & Nussinov, R. Models of ${$beta$}$-Amyloid Ion Channels in the Membrane Suggest That Channel Formation in the Bilayer Is a Dynamic Process. *Biophysical Journal* **93**, 1938(2007).
29. Crowet, J.M. et al. In Silico tilted properties of the 67-78 fragment of α-synuclein are responsible for membrane destabilization and neurotoxicity. *Proteins* **68**, 936—947(2007).
30. Miao, M. et al. Sequence and Structure Determinants for the Self-aggregation of Recombinant Polypeptides Modeled after Human Elastin*. *Journal of Biological Chemistry* **278**, 48553—48562(2003).
31. Kenney, J.M. et al. Amyloidogenic nature of spider silk. *FEBS Journal* **269**, 4159—4163(2002).
32. Rauscher, S. et al. Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. *Structure* **14**, 1667—1676(2006).
33. Lindahl, E., Hess, B. & van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling* **7**, 306—317(2001).
34. Berendsen, H.J.C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43—56(1995).
35. Jorgensen, W.L., Maxwell, D.S. & Tiradorives, J. Development and Testing of the OPLS All-Atom Force-Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society* **118**, 11225—11236(1996).
36. Jorgensen, W.L. et al. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926(1983).
37. Essmann, U. & others A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577(1995).
38. Berendsen, H.J.C. et al. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **81**, 3684(1984).

39. Hess, B. et al. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463—1472(1997).
40. Miyamoto, S. & Kollman, P.A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* **13**, 952-962(1992).
41. Pandit, S.A., Bostick, D. & Berkowitz, M.L. An algorithm to describe molecular scale rugged surfaces and its application to the study of a water/lipid bilayer interface. *The Journal of Chemical Physics* **119**, 2199(2003).
42. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577—2637(1983).
43. Pettersen, E.F. et al. UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605-1612(2004).
44. Allen, M.P. & Tildesley, D.J. *Computer Simulation of Liquids*. (Clarendon Press: 1987).
45. Lee, J. et al. Instantaneous Amyloid Fibril Formation of ${$alpha$}$-Synuclein from the Oligomeric Granular Structures in the Presence of Hexane. *Biophysical Journal* (2008).