UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A COMPUTATIONAL STUDY OF AMYLOID FIBIRLS AND THEIR STRUCTURAL PROPERTIES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ERIK ALRED
Norman, Oklahoma
2017

A COMPUTATIONAL STUDY OF AMYLOID FIBIRLS AND THEIR STRUCTURAL PROPERTIES

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

BY

_____
Dr. U.H.E. Hansmann, Chair

_____
Dr. Ronald Barnes

_____
Dr. Chuanbin Mao

_____
Dr. Susan J. Schroder

_____
Dr. Rakhi Rajan

I dedicate this text to my family and lab mates, without whose continued support I would not be where I am

Table of Contents

List of Tables

List of Figures

Abstract

The term amyloid describes misfolded protein aggregates in which a highly ordered cross β-sheet pattern is adopted. While there exist functional amyloids, the majority of known amyloids are associated with diseases in multicellular organisms. One example is the association is that between Amyloid β (Aβ) and Alzheimer's disease, a neurodegenerative disorder in humans. Several mechanisms of toxicity have been proposed, yet a lack of dynamic data prevents a full molecular explanation for the toxicity of Aβ and other amyloid systems. Mutational effects often increase the degree of polymorphism in observable structures, compounding the issues with a molecular level examination. In this thesis, Molecular Dynamic (MD) simulations of wild-type and mutant sequences of both Aβ and Prion proteins are performed to explore the structural dynamics of amyloids and amyloid-like systems. The data generated will provide physics-based explanations of the traits of amyloids on a molecular level which may guide further physical experimentation into the mechanism of amyloid toxicity and formation.

Chapter 1 - Introduction

Organic life requires a complicated and balanced system of chemical machinery in order to exist and function. One primary component of this biochemical machinery is proteins. The sequence of amino acids in a peptide chain denotes the most rudimentary level of protein structure. These chains can then adopt specific spatial arrangements necessary for function via a process known as protein folding. The apparent disorder of the starting structure, coupled with the rapidity at which it adopts an ordered structure, leads to the conclusion that the energy landscape must have a funnel shaped topology.[1,2] While the funnel shaped energy landscape certainly provides an elegant explanation, it fails to fully describe observed dynamics of protein folding. Mutational experiments performed by Obran, Bryan *et. al.* on *Streptococcus* protein- G domains allowed for the creation of two distinct protein structures, with unique activity, that shared over 90% sequence homology.[3-6] These studies underscore the need for a new paradigm for protein folding, in which the energy landscape possesses multiple funnels with pathways leading to multiple structures that exist in a dynamic equilibrium. This equilibrium may be shifted by mutation in the protein sequence and interaction with small molecules that change the structural dynamic. Due to the nature of the multi-funnel landscape, multiple different folds may exist of the same sequence in a polymorphic system of structures, with the energy barriers between structures preventing rapid coversion.[7] While methods of analyzing and fully mapping the energy landscape of protein folding are still in their infancy, computational probing using

molecular dynamics (MD) allows for the quick and robust analysis of small structural changes and their effects on the greater protein system.

Often times shifts in a protein's dynamic equilibrium is necessary for the function of a protein, such as in allosteric transitions. However, alternative pathways for protein folding can potentially lead to nonfunctioning or deleterious structures. Mutations or changes in environment that create these pathways are often associated with disease in humans. The formation and growth of highly structured aggregates of Tau and Amyloid β proteins has long been shown to be associated with Alzheimer's disease.[8,9] The aggregate formations, called amyloids, posses high structural stability largely due to the large number of hydrogen bonds, hydrophobic packing, and steric-zipper-like Van der Waals forces allowed by the cross β sheet pattern.[10]

Due to the nature of the fibrilar aggregates, X-ray diffraction crystallography can only be used to determine a two dimensional structure.[10-12] To enhance structural elucidation, solid state Nuclear Magnetic Resonance (ssNMR) may be employed to provide additional constraints. However, this still creates an undetermined system with more degrees of freedom than constraints.[13] To bypass this limitation, molecular modeling is employed to generate an ensemble of structures that fit an energy function defined by experimental constraints. This allows for an ensemble of possible structures to be derived, the lowest energy of which may be deposited in the protein databank (PDB).[14] While these structures are invaluable to research, they provide little information on the dynamics of the protein systems. Many types of protein interactions occur on time or

size scales difficult to assess in a controlled environment and impossible to observe in a biological setting.[10-12] To enhance the scope of experimentally derived structural data, computational methods may be employed to predict dynamics details on sub microsecond timescales with atomistic resolution.

Common methods of computational prediction are molecular dynamics[15-17] and Monte Carlo simultations.[18,19] Monte Carlo simulations are valuable in that they may be used to assess the statistical weights of structures in an ensemble of states.[20] While useful for assessing the shifts in population for protein structures, its field of use is beyond the scope of the research performed. Molecular dynamics may be used to generate realistic trajectories for proteins using Newtonian physics to calculate the forces.[21] This allows for the assessment of structural qualities of proteins and the extrapolation of these qualities to their function in the protein folding and aggreagation[22-24] The generated data may be used to identify key features to guide further experimentation.[15,25] A more in-depth exploration of computational methods and algorithms will be provided in

Chapter 2: Molecular Dynamics Overview

The consistency of molecular dynamics simulations is dependent on the accuracy of the empirical energy function that makes up the forcefield.[26-28] Forcefields are designed and parameterized to model specific protein interactions, and the selection of improper forcefields will impact the reliability of the generated data. This makes the selection of the forcefield a critical step in the simulation protocals.[29] For the experiments performed, the selection of forcefield was determined through root mean square

deviation (RMSD) analysis of structures in short simulations. The rationale behind this selection is that the deposited structures had to be highly energetically stable to be observable experimentally, and as such, without external perturbation structural characteristics should be preserved in the simulated trajectories. The forcefield that routinely generated the lowest RMSD for the pilot trajectories was CHARMM 27 with cmap corrections and CHARMM 36.[26,30] Forcefields selection will be further explained in the method sections of the individual chapters.

One of the questions where simulations can compliment experimentation is the role of polymorphism. While the cross β patterns are highly regular, there are multiple ways amyloid aggregates may be arranged into fibrilar super-structures. Many of the fibrilar morphologies may be observed simultaneously in experimental procedures.[31] The ability of amyloid aggregates to exist as a system of polymorphic structures is of particular importance since amyloid fibrils can differ in growth rates and cytotoxicity based on their structure.[32-34]

The ensemble of polymorphs is shifted by mutations in the peptide sequence. The Iowa mutation of the Amyloid β protein is a key example of such a phenomenon. The arrangement of the β sheets is exclusively parallel in wild-type aggreagtes.[35] However, when the Iowa mutation (D23N) is introduced in the sequence, aggregates can be observed forming with both parallel and antiparallel β sheet aragments.[36] The antipallel Iowa mutant structures are observed slowly converting to parallel structures *in vitro* as illustrated in **Figure 1**.[36] This behavior indicates that the Iowa mutation increases the

energy barrier of transition between antiparallel and parallel aggregates, creating observable metastable antiparallel aggregates. This change in equilibrium becomes important when considering the fact that the Iowa mutation is associated with more aggressive and early onset variants of Alzheimer's disease.[37,38] Thus an understanding of the transitional pathways between parallel and antiparallel aggregates and the factors changing the barrier of transition are both key to the development of chemical interventions to lessen the effects of Alzheimer's Disease. A computational exploration of the structural properties of Iowa mutant and wild-type structures will be presented in

Chapter 3: Stability of Iowa Mutant and Wild-Type Aβ-peptide Aggregates

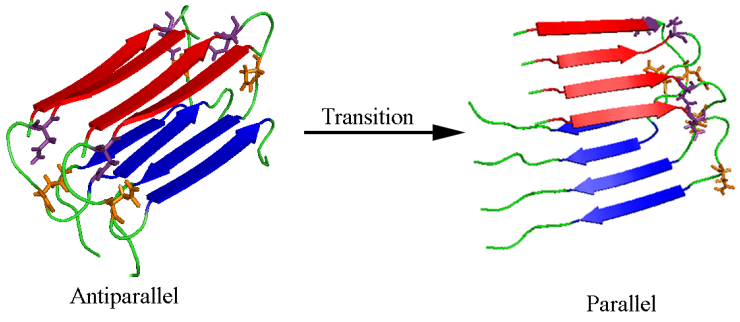

Figure 1: Antiparallel to parallel transition for Aβ Iowa mutant. Red marks the N-terminal β –strand, blue the C-terminal β strand, and the salt bridge between reside 23 and 27 is shown in orange (23) and purple (27).

Iowa mutation illustrate the effect of substitution mutations on the structural dynamics of amyloid proteins. Deletion mutations can also cause unique changes in the observable fibrilar structures and their mechanisms of growth.[39] A clear illustration of these effects can be seen in the Osaka mutation (ΔE22) of Aβ.[40] This mutation, much like the Iowa mutation, is associated with both a unique structure and elevated

cytotoxicity.[40] The structure adopted by Osaka mutant aggregates resemble a "Greek key" as opposed to the traditional hairpin structure as observed in **Figure 2**. However, unlike the Iowa mutation and observed structures, Osaka mutant aggregates may be used to seed wild-type peptides to adopt an Osaka mutant structure, but wild-type aggregates do not induce Osaka mutant peptides to form wild-type aggregate structures.[41,42] This differs from the observations of the Iowa mutation in which all sequences eventually convert to a parallel hairpin structure.[43]

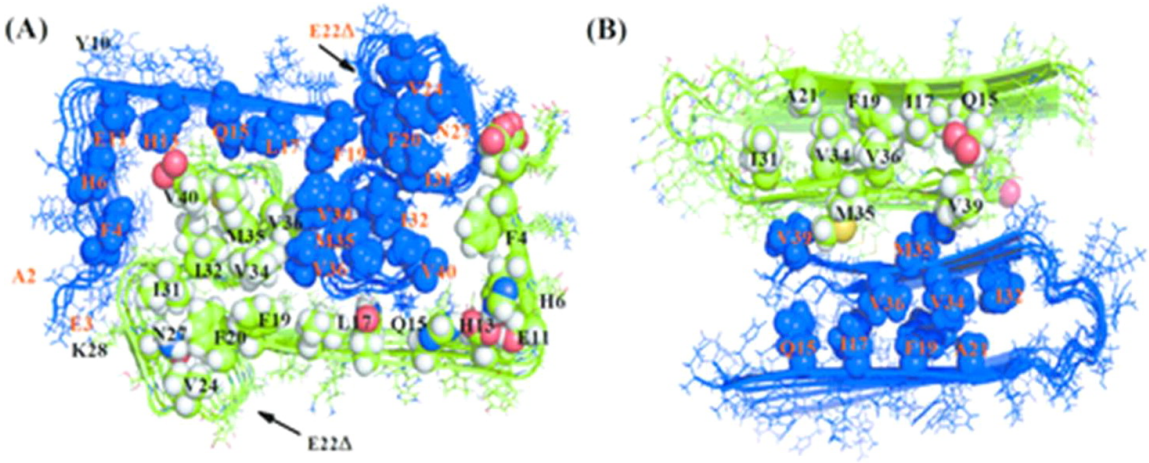

Figure 2: A comparison of the greek key structure of the Osaka mutant (A) and the parallel hairpin structure of wildtype sequence Aβ. Important intersheet contacts are indicated by their Vander wall radius spheres and sequence number. Reprinted with permission from Stability of Osaka Mutant and Wild-Type Fibril Models. J Phys Chem B 119:13063-70. Copyright 2015 American Chemical Society.

Up until this point, the discussion of amyloid structures has focused on systems derived from *in vitro* sources. This allows for a controlled setting in which the environment may be changed to observe the effects on nucleation.[7] These studies all observe polymorphic systems of amyloid aggregates with predominantly either two-fold or 3-fold symmetry on the fibrilar axis. However, recent data provided by ssNMR on structures isolated

from patients' brains post mortem indicate the presence of only a single structure in all regions of the brain with no observable polymorphism.[44] Furthermore, these structures are patient-specific, with the two patients of the study showing aggregates of different morphologies. Since the nucleation of aggregates is relatively quicker than the initial aggregation, there must exist some mechanism for structural selection during the initial stage of aggregation. However, currently this mechanism is unclear, indicating the need for further understanding of the structural properties of amyloids. This question will be explored in **Chapter 4: Lack of Polymorphism in Aβ Aggregates from Patient Brains**.

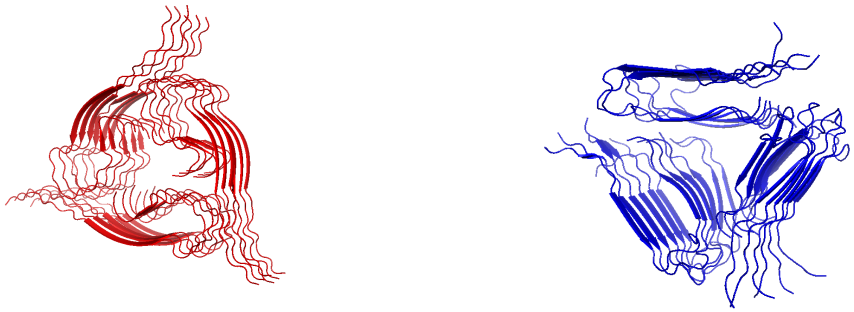

Figure 3: Shows one of the struture of the patient derived system and the predomiate 3 fold structure of *in vitro* derived system

As previously discussed, structural data on fibrilar aggregates is typically generated as an ensemble of structures derived from experimental constraints. These structures are deposited in the Protein Data Bank (PDB) in order of increasing energy.[14] This presents a problem for computational modeling of aggregate systems. Usually, the lowest energy may be considered the most statistically probable and a reasonable initial structure for simulation.[34] However, since there are typically more intra-chain contacts in amyloids

than inter-chain in the same layer,[55] the intra-chain contacts will be weighted higher in the determination of the lowest energy structure. The structural stability of the aggregate super-structure is largely dependent on inter-chain contacts,[55] thus the "lowest-energy" structure may not be the most physiologically relevant. Care must be taken in determining what structures from a fibrilar ensemble will be used as the basis for simulation as erroneous selection may bias the results. Previous research into the lack of polymorphism of patient-derived structures would be impacted by the potential bias introduced through exclusive use of low energy structures. As such, a follow up study was performed to investigate the structural variance observable in MD simulations using previously excluded structures from the ensemble. A comparison of structural dynamic differences of various NMR ensemble entries will be provided in **Chapter 5: Stability Differences in the NMR ensembles of Aβ**.

While much of the current literature on amyloid aggregation focuses on the associated neurodegenerative diseases,[45-47] mounting evidence suggests that amyloid fibrils can also be observed in a diverse set of tissue. An example of this range can be seen in systemic amyloidosis in which multiple organs can be observed concurrently forming cytotoxic amyloid plaques.[48,49] The wide array of tissue that is affected by the amyloid plaques makes diagnosis difficult, with multiple disparate illnesses arising including congestive renal and cardiac system failure.[50] Various different peptide chains are involved in the formation of systemic amyloidosis plaques, with the most common being the light-chain immunoglobulin κ protein.[51] As previously discussed, the aggregation rate and structure is largely determined by the peptide sequence. The R61N,

G68D, and A84T mutations in the κ protein sequence are shown to enhance the aggregation rate resulting in structured, polymorphic aggregates while D82I is associated with the formation of largely unstructured, amorphous aggreagates.[52-54] Mutations are again shown to either enhance or reduce the number of structures observable in the polymorphic ensemble much like in the Iowa and Osaka mutation of amyloid β. This again illustrates the importance of understanding the physical interactions governing the aggregation of amyloid forming proteins. However, consideration of sequence mutations fails to explain the lack of polymorphism observed in amyloid β aggregate structures derived from patients suffering from Alzheimer's as discussed in chapters 4 and 5. One possible explanation is that amyloid structures can self-replicate and propagate, i.e. be infectious, similar to the observed mechanisms of growth in misfolded prion structures.

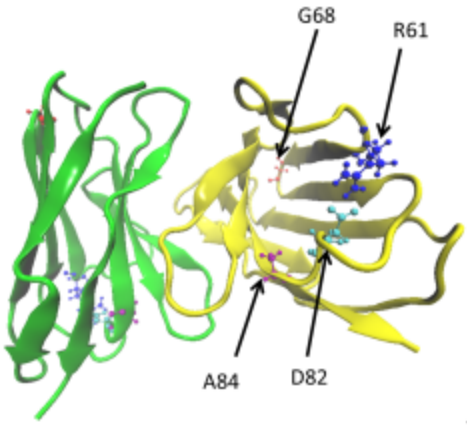

Figure 4: Dimer aggregate of light-chain κ protein showing the location of the various mutations in the sequence

One possible explanation can be seen when considering the method of pseudo-infectious method of proliferation employed by misfolded prion structures. Prions are found in a diverse set of organisms and are responsible for the development of neuronal

synapses.[56-58] An important property of prion proteins is the induction of fold that occurs when misfolded prion protein, denoted as Scrapie prion protein (PrP^{SC}), interact with healthy cellular prion protein (PrP^{C}) leading to the conversion of PrP^{C} into PrP^{SC}.[58] This phenomenon creates a system in which misfolded prions are communicable upon ingestion of neuronal tissue from an "infected" organism.[58,59] Diseases resulting from the misfolded state can be observed in humans (Creutzfeldt-Jacob) as well as other animals (Mad Cow Disease) and can often times jump the species barrier.[58,59] While limited experimental evidence exists as to the exact structure of the misfolded prions, there is an associated drop in helicity (43% in healthy prion compared to 30% in misfolded) and a dramatic increase in β sheets (3% in healthy prions compared to 43% in misfolded).[60-67] The elevated β sheet content in misfolded prion proteins and subsequent aggregation is similar to properties observed in amyloid aggregates. What's more, prions can be induced to fully form amyloid aggregates via removal of the glycosylphosphatidylinositol (GPI) anchor at the C terminus.[58,59] These findings suggest a common mechanism of aggregation and stabilization of aggregates which could be identified partially by an enhanced understanding of the physical properties of protein aggregate structures. Furthermore, the resulting data on how one pseudo-infectious predominate structure is generated using external cofactors could be used in refining the explanation of lack of polymorphism observed in certain amyloid systems. A discussion of this shared mechanism will be presented in **Chapter 6: MD Simulations of RNA-Mediated Prion Conversion.**

Much of this text will focus on the structural dynamics and degree of polymorphism observable for amyloids and how they are affected upon distinct mutational and environmental changes. The effects of RNA interaction with prion protein will also be explored to provide an example of amyloid like mechanisms that exist throughout many proteins in the body that leads to one dominate structure. Note that the dynamics of this protein-RNA interaction can also be dramatically shifted with small mutations in the peptide sequence. The D178N mutation can be shown to be associated with an elevated risk of the formation of cytotoxic prion structures, while the M129V mutate prions possess a resistance to the misfolding mechanism.[68,69] While the initial stages of prion conversion to infectious structure, alpha helical breakdown, can be observed in timescales accessible using computational methods, the formation of ordered β sheet structures occur far slower, and are difficult to observe without the usage of specialized expensive hardware.[56,70,71] This puts these and other properties of protein structural dynamics outside the scope of computational methods available to traditional academic institutions, underscoring the need for novel methods to efficiently sample the structural landscape. The results of pilot studies investigating mutational effects on prion structural dynamics as well as a discussion on potential future techniques that may be used to expand the scope of the provided research will be presented in **Chapter 7: Further Mutational Effects on Aggregation and Outlook on Computational Methods.**

The contained research seeks to utilize existing methods in molecular dynamics to explore the structural properties of amyloid and amyloid-like aggregates. Comparison

of trajectories generated from Iowa mutant sequence structures to those of wild-type sequence is used to determine the effects of the D23N mutation and how it shifts the stability of antiparallel aggregates to be metastable increasing the number of polymorphic structures. Subsequently, an evaluation of *in vitro* and patient-derived structures is performed to determine potential reasons for the selection of a single structure observed in patients from the polymorphic ensemble of structures *in-vitro*. A comparison of structures from the ensemble deposited for *in-vitro* and patient-derived is then used to show the importance of considering structures outside the lowest energy state. Finally, an analysis of prion proteins interacting with RNA snippets is shown to illustrate the initial stages of misfolding and aggregation observed following an amyloid-like mechanism and presents a mechanism for the selection of one aggregate structure from an ensemble of potential structures. These predictive studies can then be used as guides for future experimentation to focus the studies on key structural regions for the disruption of the aggregation process.

Chapter 2 - Molecular Dynamics Overview

Protein structural determination is an arduous task due to the need for structures of relatively high rigidity for examinations via experimental methods.[72-74] This rigidity is induced by various methods to enhance the sum of forces locking the protein in its folded state and eliminating entropic motion.[75] While this process is useful for pairing down the structural ensemble to a manageable level, it is not without drawbacks. Evidence suggests that evolutionary pressures in organisms favor optimization of activity, not stability, for protein selection.[76-80] Thus, experimental methods that favor the selection of stable structures may not provide a complete picture of a protein's structural ensemble. Physics based computational predictions can therefore be employed to expand the understanding of the dynamic nature of proteins from initially rigid structures.

One common method of structural prediction is through the use Molecular Dynamics (MD) simulations. MD simulations involve a process by which iterative numerical calculations are used to calculate the motion of a molecular system from the instantaneous forces generated from classical Newtonian mechanics. Typically MD predictions are made assuming the Born-Oppenheimer approximation of nuclear motion, that is to say quantum mechanical effects are ignored and atoms, or particles, are considered to be a single point mass.[81] With this assumption, the relationship between mass of particle *I,* denoted by m_i, and velocity $\left(\frac{dr_1}{dt}\right)$, where r_1 is a three dimensional vector is given by equation 1.[15]

$$\frac{dr_1}{dt} = \frac{p_i}{m_i} \quad (1)$$

with p_i being the momentum of particle *i*. The net force (F_i) exerted on particle *i* can be described as the negative gradient of a potential energy function (U) with respect to the position of the particle as denoted by equation 2.[15]

$$F_i = -\frac{dU}{dr_i} \quad (2)$$

This allows for the Newtonian equation of motion to be expressed for the particle as equation 3.[15]

$$\frac{dp_i}{dt} = F_i \quad (3)$$

Positional changes are calculated on a finite timestep Δt and can be expressed as the following Taylor series expansion (for a simplistic single dimensional model along the x axis) as shown in equation 4.[15]

$$x(t + \Delta t) = x(t) + \frac{dx(t)}{dt}\Delta t + \frac{d^2x(t)}{dt^2}\frac{\Delta t^2}{2} + \cdots \quad (4)$$

where $\frac{dx(t)}{dt}$ is the particle's velocity and $\frac{d^2x(t)}{dt^2}$ is the particle's acceleration. Typically molecular mechanics numerical solutions are only calculated for velocity and acceleration, with other higher order terms approximated empirically.[15] From this

simplification, the force of the simple single dimension system may be expressed as equation 5.[15]

$$\frac{d^2x(t)}{dt^2} = \frac{F_x}{m} \quad (5)$$

Higher order terms in the Taylor series may be approximated as being equal to 0, however this leads to a lack of conservation of energy with significant drift of total energy for long time simulations.[15]

Since integration is a relatively costly computationally process, there exists numerous methods to algorithmically integrate the equation for motion. The Verlet algorithm at timestep t uses positional data from the previous step, $x(t - \Delta t)$, and acceleration data at time t, $\frac{d^2x(t)}{dt^2}$, to calculate the new positions for the subsequent step, $x(t + \Delta t)$ as illustrated in equation 6.[82]

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + \frac{d^2x(t)}{dt^2}\Delta t^2 \quad (6)$$

This method of integration is employed in GROMACS and heavily optimized in version 5.0 onward due to enhanced accuracy at the expense of computational performance.[83] A variation of this integrator is the leap-frog algorithm, in which positional data at time (t), $x(t)$, and velocities data from time $\left(t - \frac{\Delta t}{2}\right)$, $\frac{dx(t)}{dt}\left(t - \frac{\Delta t}{s}\right)$, are used to

simultaneously update positional and velocity data lists.[84] For positional data the following algorithm is used

$$x(t + \Delta t) = x(t) + \frac{dx(t)}{dt}\left(t + \frac{\Delta t}{2}\right)\Delta t \quad (7)$$

For velocity data is calculated using equation 8.[84]

$$\frac{dx(t)}{dt}\left(t + \frac{\Delta t}{2}\right) = \frac{dx(t)}{dt}\left(t - \frac{\Delta t}{2}\right) + \frac{d^2x(t)}{dt^2}\Delta t \quad (8)$$

For the performed simulations of this work, a leap-frog algorithm was selected for the integrator as the increased accuracy of a Verlet integrator was not shown to affect the results in a discernable manner, thus selection was made based on performance optimization.

Methods of calculating potential energy functions are also an important aspect of molecular dynamic simulations. These functions are parameterized to approximate energy landscape while minimizing the needed degrees of freedom, allowing for both accurate and rapid predictions.[85] One notable popular potential function is the CHARMM forcfield.[26,86,87] While previous iterations of CHARMM used a different method of algorithmically predicting the potential-energy landscape, iterations after CHARMM27 use the scheme outlined in equation 9.[26]

$$U_{CHARMM} = U_{bonded} + U_{nonbonded} \quad (9)$$

in which U_{bonded} is defined as equation 10.[26]

$$U_{bonded} = U_{bond} + U_{angle} + U_{UB} + U_{dihedral} + U_{improper} + U_{CMAP} \quad (10)$$

U_{bond} is a function of positional data bonded pairs b and b_0 as well as parameterized force constant K_b shown in equation 11.[26]

$$U_{bond} = \sum_{bonds} K_b(b - b_0)^2 \quad (11)$$

U_{angle} is a function of angular data for the bonded pairs θ and θ_0 as well as parameterized force constant K_θ shown in equation 12.[26]

$$U_{angle} = \sum_{angle} K_\theta(\theta - \theta_0)^2 \quad (12)$$

U_{UB} is a function of positional data b^{1-3} and b_0^{1-3} as well as parameterized force constant K_{UB} shown in equation 13.[26]

$$U_{UB} = \sum_{Urey-Bradley} K_{UB}(b^{1-3} - b_0^{1-3})^2 \quad (13)$$

$U_{dihedral}$ is a function of dihedral angular data φ, phase shift data δ, and periodicity n as well as parameterized force constant K_φ shown in equation 14.[26]

$$U_{dihedral} = \sum_{dihedral} K_{\varphi}((1 + \cos(n\varphi - \delta)) \quad (14)$$

$U_{improper}$ is a function of out of plane angular data ω and ω_0 as well as parameterized force constant K_{ω} shown in equation 15.[26]

$$U_{improper} = \sum_{improper} K_{\omega}(\omega - \omega_0)^2 \quad (15)$$

U_{CMAP} is a correction term derived from backbone dihedral angular data ϕ and ψ for each residue. This is employed to more closely model experimental data using equation 16.[26]

$$U_{CMAP} = \sum_{residues} u_{CMAP}(\phi, \psi) \quad (16)$$

$U_{nonbonded}$ is defined as the summation of two terms shown in equation 17.[26]

$$U_{nonbonded} = U_{LJ} + U_{elec} \quad (17)$$

where U_{LJ} is defined as the Lennard-Jones potential which is a function of the distance between mass centers i and j, r_{ij}, as well as forcefield specific parameters ε_{ij} and r_{ij}^{min} shown in equation 18.[26]

$$U_{LJ} = \sum_{nonbond\ pairs} \varepsilon_{ij} \left[\left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{6} \right] \quad (18)$$

U_{elec} is the electrostatic potential which is a functiom of distance between mass centers i and j, r_{ij}, the charges of these mass centers, q_i and q_j, and forcefield parameter ϵ shown in equation 19.[26]

$$U_{elec} = \sum_{nonbond\ pair} \frac{q_i q_j}{\epsilon r_{ij}} \quad (19)$$

While many of the definitions of the potential function appear mathematically simple, each atom pair in a simulation theoretically represents multiple mathematical operations. This means performance decreases exponentially with system size, necessitating the need for a drastic reduction of operations. Typically cutoff schemes may be employed in which atoms outside of a predefined cutoff region are not considered for potential calculation. A "Simple-Cutoff" scheme involves the use of a circle of radius r. If $r_{ij} > r$ the potential for a given interaction is set to zero.[83] Note that due to differences between nonbonded and bonded interaction distance, two separate circles must be generated for each type of interaction.[83] While more sophisticated methods of cutoff determination exist, they showed no change in accuracy for the system of study and a drastic decrease in performance. Thus they will not be discussed in this text.

Chapter 3 - Stability of Iowa Mutant and Wild-Type Aβ-peptide Aggregates

The following chapter was published in similar form in the Journal of Chemical Physics by the author of this dissertation as the following article: Stability of Iowa Mutant and Wild Type Aβ-Peptide Aggregates. J. Chem. Phys. 141:175101. All text and figures are taken with the permission of the publisher.

Introduction

Deposits of amyloids are associated with a growing number of human illnesses. An example is Alzheimer's disease[8], which is correlated with the appearance of fibrils in patient brains that are formed by β-amyloid (Aβ) peptides[88]. The amyloid deposits consist of elongated spines made of many β-sheets strands[8] held together by a dense hydrogen-bonded network and steric-zipper-like van der Waals and hydrophobic forces[89], which in turn depends on shape complementarity and the organization of β-sheets into either a parallel or antiparallel structure[90]. Amyloid-forming peptides, such as Aβ, can simultaneously assemble into fibrils with different morphologies[91]. Such fibril polymorphism arises from differences in packing of the peptides into parallel and antiparallel β-sheets, proto-filaments, filaments, and fibrils[92]; and this polymorphism is important because the various fibril morphologies differ in growth rate and toxic potential[93-95]. Insight into this relationship is therefore crucial for understanding the disease mechanism, which in turn may open the way to new therapeutic strategies.

One possible avenue to probe this relationship is by comparing the wild-type with the various pathogenic mutations that are known to modify the physicochemical properties

of the peptide[96] and to cause early onset of Alzheimer's disease. One example of these pathological Aβ mutants is the so-called Iowa mutant (D_{23}N Aβ), which has higher neurotoxicity than wild-type Aβ[43]. While the existing experimental structures of wild-type Aβ are built out of in-register parallel β-sheets[35], recent *in vitro* studies indicate Aβ (1–40) Iowa mutant (D_{23}N) fibrils can contain either parallel or antiparallel β-sheets[97]. The antiparallel D_{23}N-Aβ_{1-40} fibrils propagate less efficiently in seeded fibril growth and were found to be thermodynamically meta-stable, transient, intermediates that convert over time into fibrils with parallel structure. Evidence for this conversion results from measurements of intermolecular dipole-dipole couplings among ^{13}C labels at A21 methyl carbons, however, there are no experimental structures of the parallel fibril Iowa mutant deposited in the Protein Data Bank[98]. The larger stability of parallel configurations has been related to more ordered residues, longer β-strand segments, and interactions between cross-β units in parallel D_{23}NAβ_{1-40} fibrils than found in antiparallel structures[98]. Additional factors that lead to a predominance of parallel structures are the more efficient packing of hydrophobic side chains at the C-terminal interface in the two fold and/or polar zipper interactions involving Q15, N23, or N27 side chains[98].

The solid state NMR data of the Iowa mutant fibrils suggests that antiparallel cross-β motifs could also exist in other cases[98]. As antiparallel cross-β motifs also exist and are believed to be thermodynamically stable in polyglutamine[99], this NMR data casts further doubt on the widely accepted assumption that amyloid fibrils are built out of in-register parallel β-sheets[100]. Instead, the data suggest coexistence of parallel and

antiparallel conformations in Aβ. Even for wild-type Aβ where the existing experimental structures are built only out of in-register parallel β-sheets, there is evidence that the wild-type could exist to some extent in an antiparallel manner[101]. Hence, while both forms are cytotoxic to neuronal cell cultures[98], it is reasonable to conjecture that the different pathologies are related to the different probabilities that both forms are observed in wild-type and Iowa mutant.

Assuming such a scenario, it becomes important to understand the polymorphism and conformational stability of Aβ peptides. For this purpose, we have performed a series of molecular dynamic simulations of antiparallel and parallel two-folds (stacks) of $A\beta_{15-40}$ and $D_{23}N$ $A\beta_{15-40}$ fibril-like oligomers. Such molecular dynamics stability studies suffer from the problem that the energy landscape of amyloid aggregation is rugged. The achievable scope of protein aggregation[102] computer simulations is limited as aggregation and conversion between forms of aggregates happen on time scales not accessible by molecular dynamics. One way to circumvent this problem is through use of stability studies, which do not directly model amyloid assembly but provide indirect input on the various factors that modulate fibril formation. An evaluation of the stability of the pre-formed initial structures requires that the system evolves with its natural kinetics. This excludes the use of enhanced sampling techniques such as generalized ensemble sampling and replica exchange molecular dynamics. Since these techniques rely on artificial dynamics[103-105], they make it difficult to interpret the changes in stability observed during the simulation. Instead, we test in the present work a different approach where the computational costs are lowered by using reduced solvent masses

that lower friction and enhance conformational sampling[106,107]. Previous studies have also shown that reducing the masses of peptide side chain atoms in combination with solvent mass can enhance peptide conformation sampling[107,108] and, in the present paper, we evaluate the efficiency of this approach in the context of protein aggregation. The approach is similar to another method for increasing sampling efficiency, namely raising the temperature of the system. However, unlike that method it avoids the problem that hydrophobicity is strongly temperature dependent which, in the case of high temperature simulations, may lead to significantly altered dynamics[109]. Mass-scaling also will lead to artificial dynamics, but we conjecture that it leads to a smaller disturbance of the system and therefore smaller deviations from the natural dynamics. Using this improved sampling technique we then explore how the differences in stability between parallel and antiparallel forms are modulated by the sequence of amino acids in both wild-type and Iowa mutants.

Materials and Methods

In our simulations, we investigate the stability of aggregates of the wild-type and the Iowa-mutant of Aβ, both in configurations with parallel β-sheets and such with anti-parallel β-sheets. The four start configurations are decamers built from two U-shaped penta-peptides with C terminal to C terminal interfaces whose structures were derived from the NMR amyloid fibrils (PDB codes, 2LNQ and 2LMO). The fibril-like oligomers used as starting configurations are made out of residues 15-40 for both $A\beta_{15-40}$ and $D_{23}N$ $A\beta_{15-40}$ peptides, which are capped with acetyl and amide groups in order to have equal length in the simulated molecules. This helps to avoid systemic error when comparing differences in the stability of the preformed oligomers of Aβ and

its Iowa mutant. We assume that the parallel $D_{23}NA\beta_{1-40}$ oligomer structures closely resemble that of the parallel WT-$A\beta_{1-40}$ structures, and the anti-parallel $D_{23}NA\beta_{1-40}$ oligomers closely resemble the anti-parallel WT-$A\beta_{1-40}$. We thus replace the residue N23 of the experimentally known $D_{23}NA\beta_{15-40}$ fibril (PDB id of 2LNQ) with D to generate the anti-parallel wild-type $A\beta_{15-40}$ fibril-like oligomer. The mutation is done by replacing the side chains of the targeted residues and keeping its original backbone conformations. The parallel $D_{23}N$ $A\beta_{15-40}$ fibril model is generated in the same manner by replacing residue N23 with D in the experimentally determined wild-type $A\beta_{10-40}$ fibril (PDB id of 2LMO) structure and removing residues 10-14. The double-layered model of $D_{23}N$-$A\beta_{15}$-40 is constructed from the experimentally derived single layer by setting the interlayer distance between two pentamers to values between 9 Å, which is in the range of values observed via x-ray[8]. The two β-strands belonging to each peptide are offset from one another along the oligomer axis by roughly 5 Å as in the experimental structure of the parallel wild-type Aβ10-40 double-fold. This allows for a more meaningful comparison between the two different structural models. The interface in the antiparallel model was then staggered to create a steric zipper that minimized clashes between residues. This was done by maximizing the distance between the initial side chain distances and orienting the stagger such that the bulky phenylalanine was positioned between glycine and leucine[110].

We ran several long all-atom explicit water molecular dynamic simulations in order to explain the structural and energetic differences between the parallel and antiparallel arrangement of the wild-type and Iowa mutant aggregates. Our molecular dynamics

simulations use a combination of the CHARMM27 force field with CMAP corrections[86,87,111] with explicit water (TIP3P)[112,113], a common choice for exploring amyloid peptide aggregation[114,115], as implemented in the GROMACS program version 4.6.2[116]. Hydrogen atoms are added with the *pdb2gmx* module of the GROMACS suite. The start configurations for all proteins are put in the center of a cubic box, with at least 12 Å between the solute and the edge of the box[99]. We enforce periodic boundary conditions in all three directions to simulate a pseudo-infinite amyloid. In order to account for the periodicity, electrostatic interactions are calculated with the PME algorithm[117-118]. We use a time step of 2 fs. Hydrogen atoms are constrained with the LINCS[119] algorithm, while the Settle algorithm is used for water[120]. The temperature of 310 K is kept constant by the Parrinello-Donadio-Bussi algorithm[121] (τ = 0.1 fs) which is similar to Berendsen coupling but adds a stochastic term that ensures a proper canonical ensemble[121,122]. In a similar way, the pressure is kept constant at 1 bar by the Parrinello-Rahman algorithm[123] (τ = 1 fs). After energy-minimizing the solvated start configuration using first the steepest descent method, followed by conjugate gradient, the system is equilibrated in two steps of 500 ps, first in an NVT ensemble and secondly in an NPT ensemble at 1 bar. After equilibration, 300 ns trajectories are analyzed for each system to monitor how the oligomer structures evolve with time; however, averages are only calculated over the last 100 ns to ensure equilibrium conditions. Data are saved at 4.0 ps intervals for further analysis. For each system, we run three distinct simulations with different initial velocity distributions. Since they start from the same initial structure, these three trajectories are correlated, and error estimates from standard deviations therefore have to be taken with a grain of salt. However, comparing the three

trajectories gives some indication for reaching equilibrium and guarantees three maximal independent sets of measurements.

Two sets of simulations were run for each of the four cases: one with the physical mass of the atoms in the molecules, and one where these masses are scaled by a factor of 0.5. Since the viscosity is proportional to mass, the mass-scaling leads to a reduced viscosity which in turn allows faster sampling of configurations. Comparing the two sets of simulations allows us to quantify the improvement in sampling efficiency.

The molecular dynamics trajectories are analyzed with the tool set of the GROMACS package. Specifically, we monitor conformational changes and stability of the oligomer models through the time evolution of root means square deviations of the Cα atoms (RMSD), root mean square fluctuation (RMSF), hydrophobic contacts distances and hydrogen bonds, measured with the *g_hbond* and *g_dist* modules in GROMACS. Hydrogen bonds are defined by a distance cutoff between donor and acceptor of 0.36 nm and an angle cutoff of 30°. Configurations are visualized using PyMOL[124].

Result and Discussion

Sampling Efficiency of Protein Aggregate Simulations

We begin our analysis by comparing molecular dynamics simulations of the four systems (wild-type and Iowa mutant, with either parallel or anti-parallel β-sheet arrangement) that use physical masses with simulations. The physical masses are scaled by a factor of one half, corresponding to a reduced viscosity of the system. As a metric to evaluate the efficiency of the two methods, we have calculated the root-mean-square

deviation (RMSD) to the start configurations as reference. The expectation is that the RMSD increases faster in the systems with scaled mass than in the control system, with both eventually approaching a similar steady state. Such behavior is indeed observed for the antiparallel Iowa mutant protein system, where both the control system and the scaled mass simulations approach a final RMSD value of approximately 6 Å (**Figure 5A).** This value is reached in the scaled mass simulations after only 7 ns while the control, run with full physical mass, requires 104 ns to reach this value (i.e. in this example we find an approximately 15 fold increase in efficiency). The improvement is even more dramatic in the case of the wild-type antiparallel model where the control system within 300 ns of simulation time never reaches the RMSD value of about 6.1 Å, whereas the scaled mass simulation approaches this value after only 30 ns (**Figure 5B)**. Assuming again an increase in efficiency by a factor of 15, the control system would only approach this value after 450 ns, much longer than the simulation time of 300 ns for our systems. A qualitatively similar picture is also observed for the two systems with parallel sheet organization (**Figure 5 C and D**). However, the gain in efficiency in both systems (wild-type and mutant) is not large enough that the systems would approach a region of constant RMSD. This is not surprising, as the parallel configurations are expected to be much more stable.

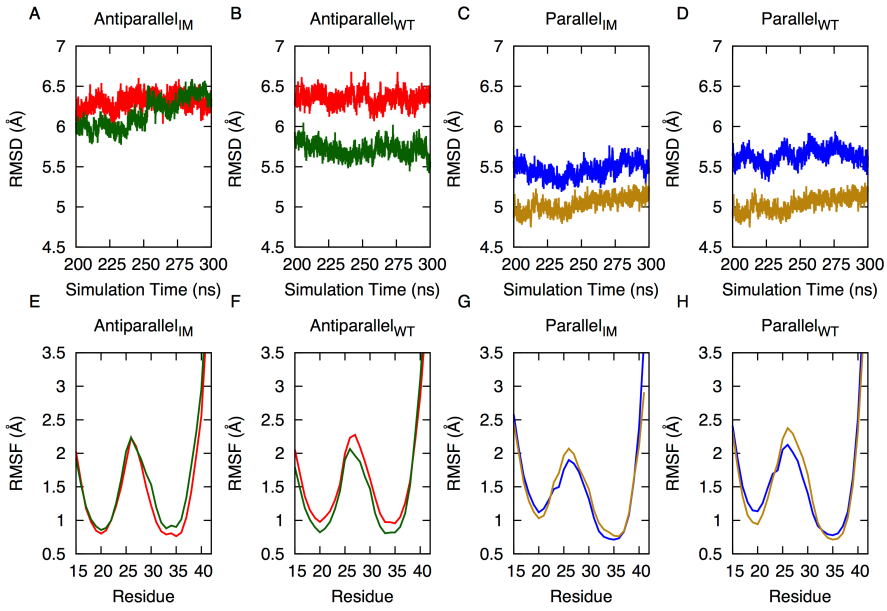

Figure 5: A comparison of the RMSD and RMSF values for Iowa mutant and wild-type with physical mass (Green/Tan) and with scaled mass (Red/Blue). There is an observable increase in the models variance with time in the case of scaled mass simulations, and thus an increase in the overall sampling efficiency. (A) and (E) show the RMSD and RMSF respectively for antiparallel Iowa mutant (AP_{IM}). (B) and (F) show the RMSD and RMSF respectively for antiparallel wild-type (AP_{WT}). (C) and (G) show the RMSD and RMSF respectively for parallel Iowa mutant (P_{IM}). (D) and (H) show the RMSD and RMSF respectively for parallel wild-type (P_{WT})

Note that in all cases similar root-mean-square-deviation values are approached. Comparison of the root-mean-square fluctuation (RMSF) values for the last 100 ns of simulations (**Figure 5 E-H**) between the antiparallel and parallel full mass (green and tan respectively) and the half mass (red and blue respectively) shows little difference in the residue behavior of the equilibrated structure. From this, we conclude that the mass scaling does not alter the behavior of the overall structure to a large degree. We conclude that mass scaling can indeed enhance the sampling efficiency in simulations of protein aggregates, albeit more work needs to be done to optimize the mass tuning. This

will be explored in future studies. In the present work, we use only our reduced mass simulations for further analysis, as the sampling efficiency in all four cases is higher than in the regular molecular dynamics runs.

Structural Stability of Wild-Type and Iowa Mutant Aggregates

Following the trajectories of the four systems, we find that the $A\beta_{15-40}$ and $D_{23}N$ $A\beta_{15-40}$ decamers have very similar dynamic behavior. Both parallel β-sheet variants (**Figure 5C and D**) change less in their root-mean-square deviation values than the anti-parallel β-sheet variants (**Figure 5A** and **B**), but the RMSD values indicate that there are no significant differences in stability between the fibril-like oligomers of $A\beta_{15-40}$ and $D_{23}N$ $A\beta_{15-40}$. A more sensitive quantity is the root-mean-square fluctuation (RMSF), which allows one to distinguish between flexible and stable residues. The average RMSF is calculated for 6 of the 10 peptide-chains for the last 100 ns in each system, with the 1^{st}, 5^{th}, 6^{th}, and 10^{th} chains removed due to aberrant flexibility caused by increased surface exposure. Both conformations of mutant and wild-type are highly flexible at the ends and in the loop regions of the protein strands, and much less so in the β-strand regions. The C-terminal region has a high RMSF value that can be explained by its C-C terminal bilayer hydrophobic interactions resulting from an increased solvent exposure. While less pronounced, high RMSF values are observed also for the three N-terminal residues. Face-to-face interactions observed in β1 (residues 18-22) and C-terminal to C-terminal hydrophobic interactions between adjacent β2 layers (residues 30-38) appear to lead to similar relative stability in both antiparallel systems (**Table 1**) as these residues have the the same average RMSF value of 1 Å in both antiparallel systems. On the other

hand, the C-terminal to C-terminal interactions are more stable in the parallel form of both mutant and wild-type, leading for the residues 30-38 to a RMSF value of less than 1 Å (**Figure 5)**. This suggests that this C-terminal to C-terminal interface is responsible for the increased stability that is observed experimentally for the parallel system of Aβ when compared to antiparallel Aβ. Note that the β1 region of both parallel systems (residues 18-22 in **Figure 5**) has an average RMSF of around 1.1 angstroms, which is more than the values observed in the antiparallel system. This is likely an artifact resulting from our truncation of the four ordered residues per peptide strand in the parallel model causing increased solvent exposure to normally buried residues.

In all four cases, the oligomer models for the Aβ wild-type and Iowa mutant keep the general characteristics and topologies of their initial conformation (**Figure 6**). The β-sheet-loop-β-sheet topology is stable in all chains of all oligomers and the hydrophobic interface between the U-shaped stacks stays intact throughout the simulation. However, the outer chains and the turn regions have an enhanced flexibility as shown in the above RMSF analysis. Hence, both parallel and antiparallel organizations are stable and therefore can contribute to the polymorphism during amyloid fibril formation. Previous simulations of the U-turn polymorphism of $A\beta_{17-42}$ by Miller *et al* [94] also showed that both parallel and antiparallel arrangements are stable and can contribute to a polymorphic population.

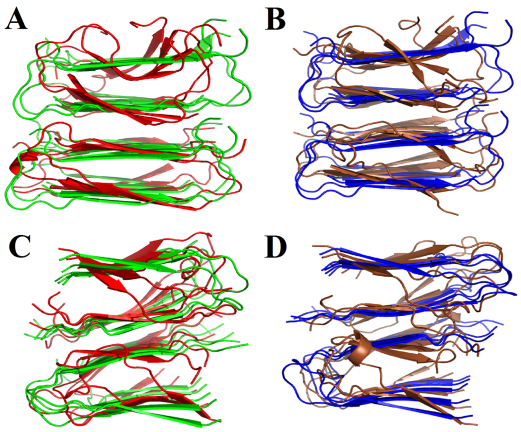

Figure 6: The structures of Aβ wild-type and Iowa mutants at end of 300 ns simulation compared with the corresponding minimized start configurations of the three independent trajectories over of (A) D23N Aβ15–40 antiparallel β-sheet where green is starting configuration and red is the structure after 300ns of simulation; (B) wild-type Aβ15–40 antiparallel β-sheet where blue is starting configuration and tan is the structure after 300ns of simulation; (C) D23N Aβ15–40 parallel β-sheet where green is starting configuration and red is the structure after 300ns of simulation and (D) wild-type Aβ15–40 parallel β-sheet where blue is starting configuration and tan is the structure after 300ns of simulation.

C to C Interface	Dist	Dist			
Antiparallel Iowa Mutant	**Start**	**Average**	**Run 1**	**Run 2**	**Run 3**
G29-E22'	10	13.4 (0.6)	13.5 (0.5)	14.4 (0.7)	12.5 (0.7)
I31-F20'	9.6	8.9 (0.4)	8.8 (0.3)	9.2 (0.4)	8.7 (0.4)
G33-V18'	9.9	8.31 (0.4)	8.6 (0.4)	8.1 (0.4)	8.3 (0.4)
M35-K16'	9.5	10.5 (0.5)	11.7 (0.5)	10.4 (0.4)	9.4 (0.6)
Antiparallel Wild Type	**Start**	**Average**	**Run 1**	**Run 2**	**Run 3**
G29-E22'	10.3	11.7 (0.6)	12.1 (0.5)	10.7 (0.5)	12.4 (0.7)
I31-F20'	10	8.6 (0.3)	9.2 (0.4)	7.8 (0.3)	8.7 (0.3)
G33-V18'	9.6	8.2 (0.4)	8.7 (0.4)	7.8 (0.4)	8.1 (0.4)
M35-K16'	10.6	9.9 (0.4)	11.2 (0.4)	9.2 (0.4)	9.2 (0.5)
Parallel Iowa Mutant	**Start**	**Average**	**Run 1**	**Run 2**	**Run 3**
I31-M35'	8.4	9.8 (0.3)	9.7 (0.4)	9.7 (0.3)	10.1 (0.3)
G33-G33'	7.2	8.1 (0.4)	8.2 (0.4)	8.0 (0.3)	8.3 (0.4)
M35-I31'	8.2	9 (0.3)	8.9 (0.4)	9.0 (0.3)	9.0 (0.3)
Parallel Wild Type	**Start**	**Average**	**Run 1**	**Run 2**	**Run 3**
I31-M35'	8.4	10 (0.4)	9.9 (0.3)	9.8 (0.3)	10.3 (0.5)
G33-G33'	7.4	8.4 (0.4)	8.0 (0.3)	8.2 (0.4)	9.1 (0.5)
M35-I31'	8.2	9.2 (0.4)	9 (0.4)	9.1 (0.4)	9.6 (0.4)

Table 1: Average hydrophobic residues distance of the C-terminal to C-terminal interactions of β-sheets in adjacent layers measured in angstroms. The standard deviation (calculated from averaging over three trajectories for the last 100 ns of the simulation) is shown in parenthesis.

Energetics of Wild-Type and Mutant Aggregates

The above findings from visual and RMSF analysis are also confirmed by an analysis of side chain contacts, hydrogen bonds, and the secondary structure content. The fibril-like oligomers of both the Aβ wild-type and Iowa mutant have tightly packed β-sheets with complementary side chains acting as a steric zipper[89,125] along the C-terminal interfaces of the double layer interface (see **Figure 7A and B**). We have monitored the distance between the Cα atoms complementary side-chains of the first β-sheet to the second β-sheet for the last 100 ns of the simulation in order to get further insight into their role in stabilizing the four different simulated oligomers. Our analysis of the data is shown in **Table 2.** The distances between G_{33}−G_{33}, and M_{35}−I_{31} residues in the parallel β-sheets are smaller than 9.5 Å, while the distance for the I_{31}−M_{35} pair is slightly larger than 10.0 Å. These distances are close to the experimental values of 8–11 Å[89], indicating a good fit between the interacting amino acid side chains, which keeps the oligomer intact during the simulation. The contacts between the residues G_{29}−E_{22} in the antiparallel β-sheet model is larger than the experimental value of 8–11[89] and increases by about 2 Å during the simulation, indicating poor packing between these residues. However the remaining pairs of residues that are involved in the C-terminal β-sheets interface have a tight inter-digitation, keeping the structure in both wild-type and Iowa mutant stable.

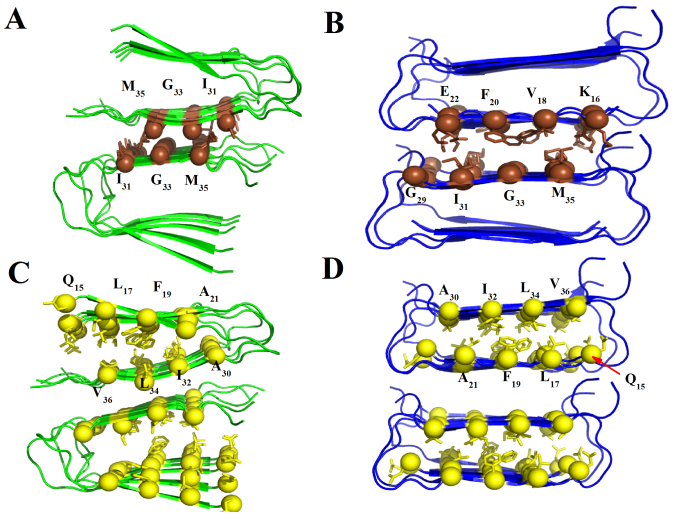

Figure 7: Face-to-face interactions between β-sheets of a single fold and side-chain interactions in the double fold along the interface between C-terminal β-sheets for decamers of Aβ wild-type and Iowa mutant**. (A)** Side chain interactions of the double fold along the C-terminal β-sheets interface residues I31/M35, G33/G33 and M35/I31 of parallel β–sheet double fold and **(B)** Side chain interactions of the double fold along the C-terminal β-sheets interface residues G29/E22, I31/F20, G33/V18 and M35/K16 of antiparallel β–sheet double fold. The side-chains involved in the complementary interactions are shown in sphere with yellow spheres representing the face-to-face interactions between β-sheets of the single fold residues and brown spheres representing C-C terminal sidechain interactions between protein layers. Parallel systems are shown in green and antiparallel are shown in blue **(C)** Face-to-face interactions between β-sheets of the single fold residues Q15/V36, L17/L34, F19/I32 and A21/A30 of Aβ15–40 parallel β–sheet, **(D)** Face-to-face interactions between β-sheets of the single fold residues Q15/V36, L17/L34, F19/I32 and A21/A30 of D23NAβ15–40 antiparallel β–sheet

Another type of interaction that is important for stabilizing the fibril-like oligomers is the face-to-face interactions between β-sheets in each single fold (see **Figure 7C** and **D). Table 2** displays the results of the face-to-face interaction dynamics for the last 100 ns of the simulation. Among selected pairs of residues for both the parallel and antiparallel experimental oligomer models, the distances increase during the simulation for the antiparallel conformation more than for the parallel. The face-to-face contacts in the parallel structure tighten during the simulation by about 1 angstrom. This may be

due to a better steric fitting of adjacent side chains in parallel conformations than in antiparallel ones that have been proposed by Antzutkin *et al*[90]. Additional stability results from the ordering of residues 10-14 in the parallel structure while the same residues are disordered in the antiparallel structures[35,98]. As the face-to-face distance indicates how well the side chain packing is, our simulation points out differences between the two systems. These differences are not enough by themselves to establish differences in thermodynamic stability; however, they point to face-to-face interactions as an important factor contributing towards differences in stability. The looser fit between the side chains, as determined from the β-sheet to β-sheet distances at the interface (Table I) and the face-to-face interactions (Table II) leads to a creation of a cavity in the anti-parallel aggregates that does not exist in the parallel form. We conjecture that the more efficient packing between β-sheets makes the parallel aggregates more stable than the antiparallel forms.

Face to Face Interface	Dist	Dist			
Antiparallel Iowa Mutant	**Start**	**Average**	**Run 1**	**Run 2**	**Run 3**
Q15-V36	8.4	9.3 (0.7)	8.8 (0.7)	9.4 (0.6)	9.6 (0.7)
L17-L34	9.5	10.1 (0.5)	10.1 (0.4)	9.6 (0.5)	10.8 (0.5)
F19-I32	9.5	10.4 (0.5)	9.9 (0.4)	10.2 (0.4)	11.0 (0.6)
A21-A30	9.9	10.6 (0.5)	10.0 (0.5)	11.1 (0.5)	10.8 (0.6)
Antiparallel Wild Type	**Start**	**Average**	**Run 1**	**Run 2**	**Run 3**
Q15-V36	8.6	9.1 (0.6)	8.8 (0.6)	9.5 (0.6)	9.1 (0.7)
L17-L34	9.6	9.8 (0.5)	10.1 (0.5	9.7 (0.5)	9.7 (0.4)
F19-I32	9.6	9.9 (0.4)	10.3 (0.5)	9.8 (0.4)	9.7 (0.4)
A21-A30	10	10.6 (0.5)	10.3 (0.5)	11.1 (0.5)	10.3 (0.5)
Parallel Iowa Mutant	**Start**	**Average**	**Run 1**	**Run 2**	**Run 3**
Q15-V36	14.7	12.9 (0.7)	12.7 (0.8)	13.9 (0.7)	12 (0.6)
L17-L34	13.9	12.0 (0.5)	12.9 (0.6)	11.2 (0.5)	11.9 (0.4)
F19-I32	12.3	12.0 (0.5)	12.0 (0.6)	11.7 (0.4)	12.4 (0.5)
A21-A30	11.5	11.1 (0.6)	10.4 (0.6)	11.7 (0.5)	11.3 (0.5)
Parallel Wild Type	**Start**	**Average**	**Run 1**	**Run 2**	**Run 3**
Q15-V36	14.5	13.5 (0.8)	13.8 (0.7)	14.2 (0.9)	12.5 (0.9)
L17-L34	13.3	12.1 (0.6)	12.9 (0.5)	12.5 (0.7)	11 (0.5)
F19-I32	12.3	11.9 (0.5)	12.1 (0.5)	11.6 (0.5)	12.1 (0.5)
A21-A30	11.5	10.5 (0.6)	11.2 (0.6)	9.2 (0.6)	11.2 (0.8)

Table 2: Averaged sheet to sheet interactions measured in angstroms for the last 100ns of the simulation. The standard deviation (calculated from averaging over three trajectories) is shown in parenthesis.

Root-mean-square deviation and root-mean-square fluctuations differ by approximately 1Å between the parallel and the antiparallel conformations. This difference and the one in packing indicate that the parallel conformation is only slightly more stable than the antiparallel conformation. However, our set-up underestimates the stability of parallel aggregates as we have removed residues 9-14 of the parallel structure in order to have the same size in all our structures (simplifying our simulation set-up). These residues are part of the ordered β-sheet in the parallel-structure fibril, but are disordered in and do not contribute to the stability of the antiparallel experimental structure. The presence of these additional five residues in the N-terminal region of the full-sized

parallel fibril increases further the differences in stability between the two forms but is not accounted for in our simulations.

Another important factor in stabilizing the supra-molecular organization of amyloid peptides[126] is the arrangement of hydrogen bonds. We have analyzed the extent of hydrogen bonds in order to determine whether there is a difference between parallel or anti-parallel β-sheet organizations of the wild-type and Iowa mutant oligomers. We find that the total number of hydrogen bonds is comparable for Iowa mutant and wild-type in both parallel and anti-parallel systems. However, the number of backbone, side chain and side chain-backbone (i.e. protein-protein) hydrogen-bonds are slightly larger (by 18(8) bonds) in the antiparallel systems than the parallel systems (see **Figure 8)** of $A\beta_{15-40}$ and $D_{23}N$ $A\beta_{15-40}$. On the other hand, the number of hydrogen bonds between the aggregate and the surrounding solvent is larger in the antiparallel configuration than in the parallel one by about the same amount (21.9(19) bonds). However, the error in the latter number is large and makes an interpretation difficult. No signal for a preference of one over the other configuration was found in the solvent-solvent hydrogen bonding. From these results, we conjecture that, in the experimentally determined structures, the four additional ordered residues in the parallel arrangement of $A\beta_{10-40}$ increase the number of hydrogen bonds and therefore stabilize this structure over that of the anti-parallel $D_{23}N$ $A\beta_{15-40}$ oligomer. This is also in agreement with the experimental data and will make the parallel form more stable than the antiparallel form[35,98].

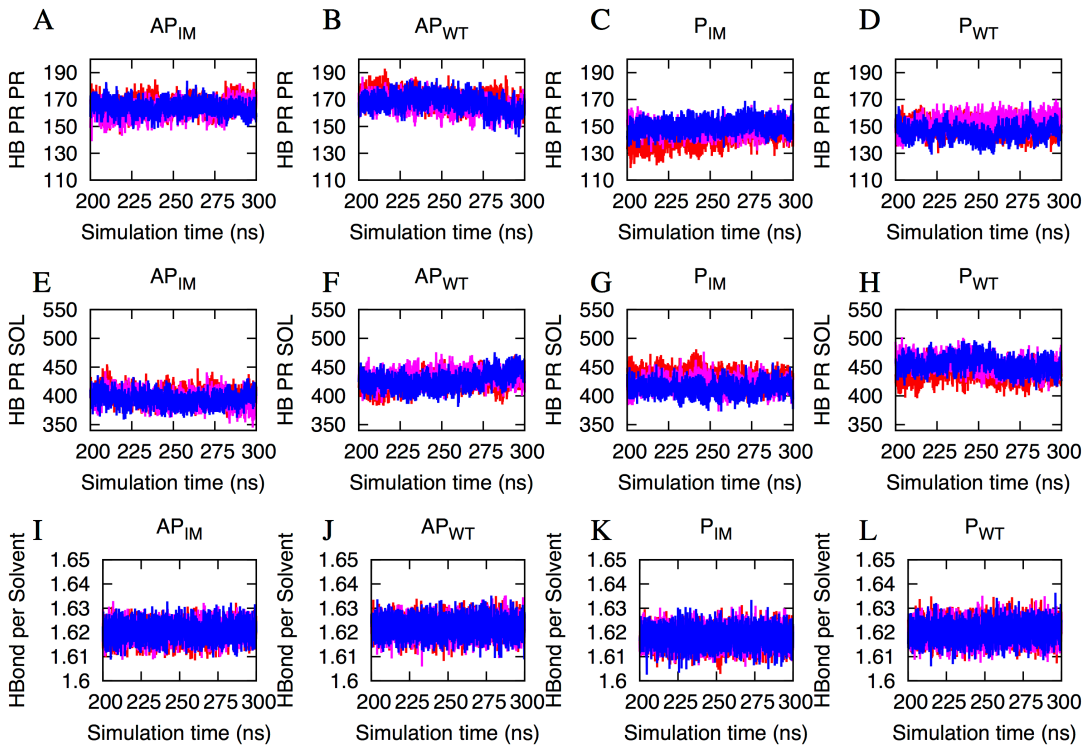

Figure 8: Hydrogen bonding of Aβ wild-type and Iowa mutant involving residues 15-40 and plotted versus time. This shows the main chain to main chain hydrogen bonding in the first row for (A) antiparallel Iowa mutant (AP_{IM}), (B) antiparallel wild-type (AP_{WT}), (C) parallel Iowa mutant (P_{IM}), (D) parallel wild-type (P_{WT}). The second row show side chain to side chain hydrogen bonding for (E) antiparallel Iowa mutant, (F) antiparallel wild-type, (G) parallel Iowa mutant, (H) parallel wild-type. The third row shows the hydrogen bonding for (I) antiparallel Iowa mutant, (J) antiparallel wild-type, (K) parallel Iowa mutant, (L) parallel wild-type. The three independent runs are shown in red, blue, and pink.

Evidence for this conjecture is also found in our monitoring of the secondary structure during the simulation. Given that amyloid fibrils are composed mainly of β-sheets, this quantity provides information on the relation between interactions that involve the β-strand motif and the stability of the aggregates[127]. **Table 3** lists the average secondary structure content from the DSSP analysis for the initial structure and the last 100 ns of

each system. As a general tendency, we find lower β-sheet content in the decamers with parallel organization than in such with antiparallel structure. The increased β-sheet content in the latter case could contribute to the stability of the antiparallel organizations and could explain how this structural organization can compete with the parallel conformation[98]. Wei *et al*[98] have proposed that the difference in the number of ordered residues (50% of antiparallel versus 75% in parallel oligomers) is responsible for the thermodynamic preference of parallel over antiparallel structures. However, our simulation of $A\beta_{15\text{-}40}$ fibril-likes oligomer of Aβ wild-type and Iowa mutants indicates that the antiparallel structure has a slightly larger percentage of β-sheet secondary structure. This difference may be due to limitations of our model, which considers only residues 15-40. While residues 1-15 are disordered in the experimentally determined antiparallel $D_{23}N$ $A\beta_{15–40}$ fibril, in the Aβ wild-type fibril only residues 1-10 are disordered. We therefore would expect more β-sheet secondary structure in simulations than seen in the experimental $A\beta_{10\text{-}40}$ parallel β-sheet oligomer. Hence, our results suggest that the relative stability of parallel β-sheet over antiparallel β-sheet conformations dependents on the larger number of ordered residues in the parallel β-sheet conformation.

System	Percentage of secondary structure of the initial sample				Percentage of secondary structure averaged over 200-300ns			
	Helix	Beta Sheet	Turns	Random	Helix	Beta Sheet	Turns	Random
Antiparallel Iowa mutant	0	49.59	13.11	37.30	1.1 (0.49)	51.65 (0.59)	16.96 (0.45)	30.29 (1.47)
Antiparallel Wild type	0	47.48	15.87	36.65	0.56 (0.86)	50.01 (0.8)	18.41 (1.91)	31.01 (1.18)
Parallel Iowa mutant	0	46.33	14.26	39.41	1.3 (1.11)	44.43 (1.71)	18.41 (2.89)	35.87 (2.68)
Parallel Wild type	0	46.28	14.96	38.76	2.82 (2.37)	44.68 (1.75)	19.90 (0.8)	32.59 (1.1)

Table 3: Average secondary structure of the wild-type Aβ and Iowa mutant antiparallel and parallel sheet structure structures during the initial structure and last 100 ns. The standard deviation between the three runs is shown in parenthesis.

Free Energies of Wild-Type and Iowa Aggregates

Common paradigm is that the protein-aggregation landscape resembles a rugged valley with numerous close local minima corresponding to different polymorphic forms of fibrils. Multiple polymorphic aggregates can emerge during fibril formation, with different environmental conditions favoring different polymorphic forms over the others[128,129]. In order to gain further insight into the relative stability of our four oligomers we have done MM-PBSA calculations of the oligomers allowing us to monitor their interactions[130] through calculating approximate binding free energies from molecular simulations[131,132]. While the MM-PBSA approach in general does not replicate the absolute binding free energy values, [133] it was chosen by us because it allows one to calculate quickly an estimate for differences in the free energy of binding, and because it usually exhibits a good correlation with experimental data[134]. In the present study the binding energy between the two β-sheets (that is between the pentamers that form the decamer) are estimated with the MM-PBSA methodology as implemented in AMBER12 using the using the all atom AMBER99SB force field and TIP3P water model. Before starting the MM-PBSA analysis, all water molecules and ions were excluded from the trajectory. The dielectric constants used for the solute and

surrounding solvent are 1 and 80, respectively. Multiple molecular dynamics runs (three trajectories of 40 ns that unlike our production runs did not utilize mass-scaling) were performed for each of the four double layer aggregate models and carefully equilibrated for 5ns before the production run. Such an approach is preferable over a single long time run as it leads to a more efficient sampling of phase space. The MM-PBSA single trajectory approach is used to calculate the binding energy. Snapshots are gathered at intervals of 40 ps during the 40 nanoseconds of simulation leading to 1000 equally spaced snapshots for each single trajectory from the three 40 ns MD trajectories. For our analysis we use the Python implementation of MM-PBSA as provided with AmberTools 12[135]. The results of the binding energy in **Table 4** are the average of the three calculations. Note that we did not account for entropic contributions to binding since we compare only systems that are very similar where these contributions could even raise the overall uncertainty in the calculated binding energies[136]. For a detailed discussion, see [137,138].

System	ΔE_{vdw}	ΔE_{ele}	ΔE_{PB}	ΔE_{SA}	$\Delta E_{non\text{-}polar}$	ΔE_{polar}	$\Delta G_{binding}$	$\Delta\Delta G^{\S}$
Parallel Wild-type	-325.0(12.3)	-96.2(111.6)	187.6(89.2)	181.4(3.6)	-143.6(8.7)	91.4(22.6)	-52.2(26.5)	-21.6
Antiparallel Wild-type	-351.1(18.1)	-307.23(36.7)	431.8(20.7)	195.7(7.9)	-155.3(10.2)	124.7(18.4)	-30.6(9.2)	
Parallel Iowa mutant	-344.8(9.2)	-457. 7(31.6)	546.1(42.8)	192.3(0.8)	-152.6(8.7)	88.3(17.3)	-64.3(24.2)	-23.4
Antiparallel Iowa mutant	-351.1(22.1)	-670.3(64.2)	782.8(60.3)	197.7 (10.0)	-153.4 (12.4)	112.5(24.4)	-40.9(14.4)	

Table 4: Binding energies and standard deviations as calculated with the MM-PBSA approach, including its components (Kcal/mol) for the four fibril models. The standard deviation (calculated from averaging over three trajectories) is shown in parenthesis.

§ $\Delta\Delta G$ is the difference in binding free energy $\Delta G_{binding}$ between the parallel and antiparallel conformation of Aβ wild-type and Iowa mutant. The negative values indicate that parallel conformations are more stable than the corresponding antiparallel conformations. $\Delta G_{binding} = \Delta E_{vdw} + \Delta E_{ele} + \Delta G_{sol}$; $\Delta G_{sol} = \Delta E_{PB} + \Delta E_{SA}$. Here, ΔE_{vdw} is van der Waals energy and ΔE_{ele} is the non-solvent electrostatic energy. The contributions to the solvation free energy are split into a nonpolar and polar part. E_{SA} is a nonpolar contribution to solvation free energy; ΔG_{PB} is electrostatic contributions to the solvation free energy calculated by the Poisson-Boltzmann equation. The nonpolar

term ($\Delta E_{non\text{-}polar}$) consists of the van der Waals interaction energies (E_{vdW}) and the nonpolar contribution to the solvation free energy (E_{SA}). The polar term (ΔE_{polar}) is the sum of Coulomb interaction energy (E_{elec}) and polar contribution to the solvation free energy (E_{PB}). Various entropic contributions are neglected in our MM-PBSA approximation.

The obtained binding energies listed in **Table 4** are in agreement with the experimental observation that the parallel Iowa mutant is more stable than the transient antiparallel structure[98], and that while there is experimental evidence for the possibility of wild-type Aβ existing in an in antiparallel form[99] all experimental determined models of Aβ are build out of of parallel β-sheets[35]. This high-energy difference predicted by our MM-PBSA calculation suggests that the conversion between parallel and antiparallel forms cannot be achieved by simple rearrangement (since it is a topological difference), but is most likely due to individual strands detaching from the less stable antiparallel form and assuming the more stable parallel β-sheet structure[98,134,135] .

We have further analyzed the various components in order to identify the dominant factors in the binding affinity. The van der Waals (ΔE_{vdw}), electrostatic term ($\Delta \mathbf{E_{ele}}$) and nonpolar terms ($\Delta E_{non\text{-}polar}$) favor complex formation in all cases. The nonpolar solvation ($\Delta \mathbf{E_{SA}}$) which describes the process of transferring a nonpolar molecule from vacuum to water including the creation of a cavity in water, is unfavorable in all cases. Similarly, the polar solvation term opposes the protein-protein binding due to polarization of the solvent environment by the solute. The favorable electrostatic contribution (ΔE_{ele}) between the proto-filament pairs, which due to the destabilizing electrostatic repulsions by charged amino acids is larger in an in-register parallel β-sheet

structure than in the anti-parallel β-sheet structure with its pairing of negatively and positively charged groups, is cancelled by the electrostatic solvation term ($\Delta \mathbf{E_{PB}}$).

For both the Iowa mutant and wild-type Aβ does the binding energy of the two parallel and antiparallel conformations suggest a slightly higher thermodynamic stability of the parallel organization over that of the antiparallel one. While free energy perturbation and other free energy calculations methods are more accurate than the MM-PBSA approach we used[139], this ranking of the stability could explain the experimental observation of the conversion of the less stable antiparallel Iowa mutant into the more stable parallel conformation fibril.[98,140] While our data are noisy with only small differences, they are supported by a recent theoretical study by Okamato et al[141] using the *ab initio* MP2/6-31G method that also found the parallel conformation in both the wild-type and Iowa mutants more stable than the antiparallel ones.

Conclusion

We have investigated in silico the stability of decamers of Aβ wild-type and its Iowa mutant that are reported to exhibit either an antiparallel or parallel β-sheet organization. Simulations with scaled mass are compared with such simulations where the physical masses are unchanged. Our data demonstrates that mass-scaling leads indeed to increased sampling efficiency by reducing the viscosity of the system. Similar to the more common raising of system temperature, this allows for an easier escape from local minima and therefore enhanced sampling while at the same time keeping deviations from the natural dynamics small. In some cases we found with mass scaling

improvements in sampling efficiency by factors 15. There is potential for further refinement of selective scaling to provide even greater yields in efficiency. Even with this increase in efficiency our simulations clearly do not cover the time scales on which the conversion of antiparallel to parallel forms is observed in experiments. However, signals for these transitions can be found by comparing the stability of wild-type and mutant in the two forms. An analysis of the free energy of binding by MM-PBSA of our data as derived from configurations that originated from these enhanced molecular dynamics simulations, indicates that the parallel forms of both wild-type and Iowa mutant aggregates are the most stable, while the antiparallel aggregates is less stable for the Iowa mutant and least stable for the wild-type. This ranking of stabilities is consistent with previous experimental results and is explained by us with the dependency of the structurally important sheet-to-sheet interface interactions on the side chain complementarity. The direct alignment of hydrophobic interactions in the in-register parallel oligomers makes them more stable than the antiparallel aggregates. The parallel supra-molecular organization could be due to favorable residues stacking while each adjacent residue overcome potential repulsive charge interactions making it energetically more favorable than the antiparallel aggregates[90]. However, our data demonstrate that both parallel and anti-parallel β-sheet fibrils can exist under physiological conditions even though they have differences in thermodynamic and structural stability. Hence, both forms may contribute to the polymorphism of Aβ aggregates. David Eisenberg and co-workers [142] utilizing the parallel β-sheet fibril models of an Alzheimer's peptide in combination with computer modeling, have found several compounds that reduce amyloid toxicity. Our data suggest that these results

could be improved by using for Aβ aggregation inhibitor design as templates not only parallel, but also anti-parallel β-sheet fibrils.

The slightly higher thermodynamic stability of the Iowa mutant oligomers in its parallel organization over that of the mutant in antiparallel form is supported by previous experimental measurements showing slow inter-conversion of antiparallel aggregates into parallel ones. Our calculations indicate that this conversion is energetically costly. This suggests that the conversion is not a simple re-arrangement but rather involves the detachment of monomers from the less stable antiparallel form and re-aggregation into the more stable parallel structure. Future plans include computational studies that rely on novel sampling techniques currently under development in our lab.

Acknowledgement

This work is supported by the National Institutes of Health under Grant No. GM62838 and the National Science Foundation under Grant CHE-1266256, and used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH1123. Other parts of the simulations were done on the BOOMER cluster of the University of Oklahoma. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Institutes of Health, the National Science Foundation, the Department of Energy, or the University of Oklahoma.

Chapter 4 - On the Lack of Polymorphism in Aβ-peptide Aggregates Derived from Patient Brains

The following chapter was published in similar form in Protein Science by the author of this dissertation as the following article: On the Lack of Polymorphism in a-Peptide Aggregates Derived from Patient Brains. Protein Science 24:923-935. All text and figures are taken with the permission of the publisher.

Introduction

Amyloid Beta (Aβ) oligomers and fibrils found in the extracellular space of the brain and on the walls of cerebral blood vessels in Alzheimer's patients [143] cause, or at least contribute to, the outbreak of Alzheimer's disease. Hence, characterizing the organization of these aggregates may help to combat or reverse the effects of the disease [55]. Multiple *in vitro* studies have shown that Aβ fibrils are polymorphic, and that their structures depend on the conditions encountered during their nucleation [7]. However, solid-state NMR data of fibrils acquired post mortem indicate lack of such polymorphism in the brains of Alzheimer's patients [44]. Only a single, patient-specific, form of aggregates was found in samples of several regions in the brains of two deceased patients [44]. The form of this patient-specific structure is correlated with the aggressiveness of the disease in the patient [44]. The question arises why only a single (patient specific) form is found in patient brains, and how these patient-derived forms differ from the ones observed *in vitro*.

Amyloid formation is difficult to observe directly in experiments and computer simulations. The later suffer from the problem that aggregation, and conversions between different forms of aggregates, takes hours or longer, i.e. happens on time scales that are not accessible in molecular dynamic simulations [17,144]. For this reason, we address the above question by comparing the stability of structurally distinct Aβ amyloid fibrils in atomistic molecular dynamics simulations with explicit solvent. Specifically, we study a two-fold *in vitro* model of Aβ fibrils (PDB code: **2LMO**), a threefold *in vitro* model (PDB code: **2LMP**), and the three-fold *in vivo* (patient-derived) model (PDB code: **2M4J**). While faster than direct aggregation simulations, such stability investigations are still computationally taxing, and often the temperature is raised to increase sampling speed [144,145]. In the present paper, we use a different approach, reducing the viscosity of the system through mass scaling. Long-time all-atom explicit solvent molecular simulations are performed of *in vitro* and *vivo* Aβ fibril models of various arrangements and sizes. The so-generated data are used to evaluate the conformational stability of the assemblies in order to provide insights into (a) structural stability, (b) hydration in the cavities, and (c) the role of intra and inter-sheets packing. Our goal is to determine through these enhanced molecular dynamic simulations the factors that modulate the stability of Aβ-fibrils and lead to the lack of polymorphism in the fibrils collected from patient brains. This is important, as knowledge of the specific forms that (Aβ) oligomers assume in the environment of patient brains could lead to new approaches in development of drugs targeting Alzheimer's disease. The difference between the *in vivo* and *in vitro* fibril models also questions the relevance of previous computational studies that were based on the *in*

vitro models and explored the effect of mutations, alternate fibril morphologies or nucleation, in cross-seeding or for *in silico* structure-based search of inhibitors or imaging agents [93,146-149].

Our simulations indicate that the brain-derived model is less stable than the two *in vitro* models. The three-fold *in vivo* and *in vitro* fibril models are characterized by a central hydrated channel that may explain the toxicity of the amyloids as such channels interfere with nerve cells communication by changing membrane potential, and also may cause cell death due to loss of solutes.

Structural Characterization of the Experimentally Derived *in vivo* and *in vitro* Aβ Fibril Models

Incubation of Aβ peptides in buffer under shaking conditions leads to formation of fibrils with two-fold symmetry while under quiescent conditions the resulting fibrils have a twisted morphology and a three-fold symmetry [35,150]. These *in vitro* fibrils, with either two-fold or three-fold symmetry with respect to the axis of fibril growth, differ in the arrangement of the ordered *strand-loop-strand*, inter-peptide packing, and in the inter-residue interactions. For *in vitro* fibrils with three-fold symmetry, an ssNMR-derived molecular model is only available for $A\beta_{9-40}$. Both two-fold and three-fold *in vitro* Aβ fibril models display a similar *strand-loop-strand* unit, in which residues 10-22 and 30-40 form two β-strands connected by a loop. However, fibril symmetry and number of β cross units per layer differ in the two models. In the two-fold fibrils the side chains of D_{23} and K_{28} form salt bridges in the loop region that are absent in the *in*

vitro fibril with three-fold symmetry (**Figure 9B-C**). On the other hand, the *in vivo* fibril structure (seeded with an extract from an Alzheimer's patient's brain tissue) has a three-fold symmetry, with the entire $A\beta_{40}$ sequence participating in the ordered structure built out of *strand-loop-strand units*.

Both the *in vivo* and the three-folded *in vitro* (**Figure 9**) models are stabilized by side chain hydrophobic interactions between the two β-strands (for example F_{19}-L_{34} contacts). The N-termini of the chains (with a majority of charged and polar side chains, residues G_9-E_{22} for *in vitro* and residues D_1–E_{22} for *in vivo*) are exposed to the exterior bulk solvent while the C-termini (including the majority of hydrophobic amino acid side chains, residues A_{30}-V_{40}) face the internal fibril surface (**Figure 9A-C**) and form interior channels in the center of the fibrils. Both termini are connected by a turn region consisting of residues 25-29. The channels in the center of the fibrils contain rows of water when hydrated [150]. The three-fold *in vitro* and *in vivo* models differ in how the three β-layers are packed through a hydrophobic center, different contacts between residues. For instance, in the *in vivo* model there are salt bridges between residues D_{23}-K_{28}, residues 1–8 are ordered, and residues 30–40 lack a single β-strand (**Figure 9 A-B**), with several kinks and bends in the β_2 strand that results in non-parallel contact of β_1 and β_2 (**Figure 9A**). However, the *in vitro* fibril forms parallel (non-kinked) contacts between β_1 and β_2 strands (**Figure 9 B-C**). The side chains of I_{31} and V_{39} are in close contacts between *strand-loop-strand* units in the central cavity and contribute to the stabilization of the *in vitro* model three fold. The *in vivo* models, on the other hand,

have the close contact between *strand-loop-strand* units in the central cavity, between I_{32} and V_{40} (**Figure 9 A-B**).

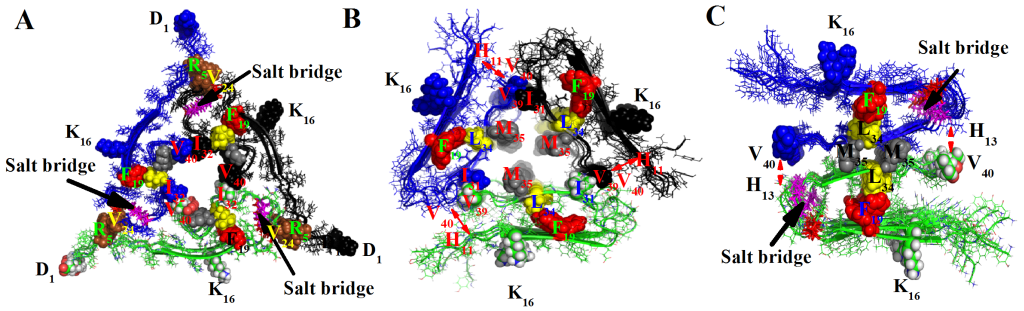

Figure 9: I*n vivo* (three fold symmetry) and *in vitro* (three-fold and two-fold symmetry) SSNMR structural models of Aβ40 fibrils. (a) The three-fold *in vivo* Aβ40 fibrils model (PDB code 2M4J) with the entire residues 1–40 structurally ordered. Six strands within the *strand-loop-strand units* are shown. The three *strand-loop-strand* units are colored blue, green, or black for clarification. The side chains of D_{23} (red) and K_{28} (magenta) form a salt bridge (indicated with an arrow). (b) The three-fold *in vitro* model, consisting of three *strand-loop-strand* structures (PDB codes 2LMP) with six strands (repeats) within the *strand-loop-strand* motifs shown. The three *strand-loop-strand* units are colored blue, green, or black for clarification. Only residues 9–40 are shown. (c) The two-fold *in vitro* Aβ40 fibril model, consisting of two *strand-loop-strand* structures with nine strands (2LMO). The side chains of D_{23} (red) and K_{28} (magenta) form a salt bridge (indicated with an arrow). Residues 9–40 are shown in which the β-sheets are associated through C-terminal-to-C-terminal interfaces. The two *strand-loop-strand* units are colored blue and green. In all models, we display hydrophobic contacts between F_{19} (red color) and L_{34} (yellow color) in each *strand-loop-strand* units as spheres. The I_{32} and V_{40} form contacts between different cross-β units in the *in vivo* three fold aggregate; similar contacts are formed in the *in vitro* three-fold conformer between I_{31} and V_{39} (shown as spheres). The *in vivo* three fold also forms contacts between the side chain of R_5 and V_{24} adjacent layers; while in the *in vitro* three fold it involves H_{14} and V_{40} (shown as spheres in brown color). The hydrophobic cavity is formed by residues M_{35} and V_{40} for the three-fold *in vivo* model and M_{35} for the *in vitro* model (shown gray spheres)

All three fibril models share common hydrophobic contacts (i.e. the hydrophobic between F_{19} and L_{34}) between the two layers in the *strand-loop-strand* units, with a parallel β-sheet in each layer. The *in vivo* three-fold model also forms contacts between the side chains of R_5 of one layer with V_{24} of the adjacent layer, while the *in vitro* three-

fold forms similar contacts involving H_{14} and V_{40}. The hydrophobic core of the interior channel in the *in vivo* model consists of M_{35} and V_{40} residues, however in the *in vitro* model with three-fold symmetry the core consists of only M35 residues.

Material and Methods

Construction of the Fibril-Like Oligomer Models

In order to compare *in vivo* and *in vitro* structures of Aβ aggregates we chose three NMR-based models derived by the same laboratory (the Tycko group at NIH, http://www.niddk.nih.gov/about-niddk/staff-directory/intramural/robert-tycko/) as this minimizes errors resulting from different experimental protocols for purification and characterization of the aggregates. These three models include two *in vitro* fibril models, one two-fold (PDB-ID: **2LMO**) and one three-fold (PDB-ID: **2LMP)**; and the *in vivo* fibril model which has three-fold symmetry (PDB-ID: **2M4J**). The positive- and negative-stagger 2-fold-symmetric A β 40 fibril models were deposited separately by Petkova et al [35] with PDB entries 2LMN and 2LMO, respectively. Recent computational studies[151] of models of A β 40 fibrils suggest that only the negative stagger model (PDB code: 2LMO) can form left-handed helical superstructures, the twist that has been observed in scanning electron microscopy studies of amyloid superstructures [152]. In contrast to the negative stagger model, the positive-stagger A β 40 fibril filaments cannot adopt the super-structural helical twist. The negative stagger is suggested as A β 40 fibril that is physiologically relevant[151]. The free energy surface [153] for negative and positive stagger structures based on extensive molecular dynamics simulations coupled with umbrella sampling is found to be consistent with prior data

that suggest that positive-stagger A β 40 fibril cannot adopt the super-structural helical twist. Therefore in our simulation we selected negative stagger model (PDB code: 2LMO) because such periodic alignment of β -hairpins would lead to the experimentally detected left-handedness of fibrils for the two-fold A β fibril model[151]. In order to probe the stability of the three forms, we have generated a cascade of models of increasing size, with three, six and twelve layers for oligomers with three-fold symmetry (**Figure 10-11**) (each layer made out of three Aβ chains), and nine and eighteen layers of two Aβ chains for oligomers with two-fold symmetry (**Figure 12**).

The *in vivo* model **2M4J** is made out of three layers, while the *in vitro* model **2LMP** has six layers. Each layer is made out of three Aβ chains. Models with equal number of layers are derived by, for instance, removing surplus chains in the **2LMP** aggregate, or forming a six-layer *in vivo* model by aligning two copies of the three-layered **2M4J** such that the inter-strand distances are ~4.7 angstroms. In the same way, we generate twelve layer systems by aligning two six-layer systems, requiring again inter-strand distances to be ~4.7 angstroms. In the same way, we take the two-fold **2LMO** model which is made of six layers of two *strand-loop-strand units*, and built out of it models with either nine or eighteen layers. In this way, we arrive at models with two-fold symmetry that have the same number of chains as the models with three-fold symmetry. The oligomer with nine layers is constructed using the six layers coordinates of the **2LMO** structure, and three additional layers are added by aligning the strands in such a way that the contacts mirror the ones in the NMR model, setting the inter-strand distances again to ~4.7 angstroms. This process was repeated to generate the eighteen

layers system. Since N-terminal residues 1−8 residues are structurally disordered in the two the 2-fold and 3-fold in vitro ssNMR based structures A β 40 fibrils (PDB entries 2LMN and 2LMO) respectively [35], we did not included residues 1–8 in the MD simulations on in vitro structures simulations.

Simulation Protocol

In order to probe the stability of the various aggregates we run a number of long all-atom explicit water molecular dynamic simulations relying on a combination of the CHARMM27 force field (with CMAP corrections [154,155]) with explicit water (TIP3P)[156,157], a common choice for exploring amyloid peptide aggregation,[158,159] as implemented in the GROMACS program version 4.6.2[160]. MD simulations were performed using as the starting structure the first structure of the 10 and 20-member lowest energy NMR structural ensembles for both the in vitro and in vivo fibril models respectively [35,44]. We use the *pdb2gmx* module of the GROMACS suite to add hydrogen atoms. The molecules are put in the center of a cubic box, with at least 12 Å between the solute and the edge of the box. Periodic boundary conditions are enforced, and electrostatic interactions are calculated with the PME algorithm[161,162]. We use a time step of 2 fs. Hydrogen atoms are constrained with the LINCS[163]algorithm, while for water the Settle algorithm is used[164]. The temperature of 310 K is kept constant by the Parrinello-Donadio-Bussi algorithm [165] (τ = 0.1 fs) which is similar to Berendsen coupling but adds a stochastic term that ensures a proper canonical ensemble [166]. In a similar way, the pressure is kept constant at 1 bar by the Parrinello-Rahman algorithm [123] (τ = 1 fs). In order to increase sampling, we use mass-scaling of all solvent atoms (masses are scaled by a factor of 0.5), which effectively decreases the viscosity of the

solvent. In previous work[167] , we found that such mass-scaling can lead to modest improvement of factors 2-8 in sampling efficiency without disturbing the dynamics of the system.

After energy-minimizing the solvated start configuration, using first steepest descent followed by conjugate gradient, the system is equilibrated in two steps of 500 ps, first in an NVT ensemble and secondly in an NPT ensemble at 1 bar. After equilibrization, we perform 300 ns of molecular dynamics simulations of the three-fold fibril models with three and six layers, and the two-fold model with nine layers. Due to limitations in computational resources are the three-fold systems with twelve layers and the two-fold system with eighteen layers only simulated for 100ns.

For each system, data are saved at 10 ps intervals in the so-obtained trajectories, and are analyzed to monitor the evolution of oligomer structures with time. For each system, we run three distinct simulations with different initial velocity distributions. Note, however, that these three trajectories are correlated since they start from the same initial structure. Hence, error estimates calculated from standard deviations over the three runs have to be taken with a grain of salt. The trajectories are analyzed with the tool set available in the GROMACS package. We monitor conformational changes and stability of the our systems by measuring the time evolution of the root-means-square deviations of the Cα atoms (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg), pore diameter, secondary structure contents, hydrophobic contacts and hydrogen bonds. PyMOL is used for visualizing structural changes and for generating images.

Results

Systems	C_α RMSD (Å)	Radius of gyration (Å)	Inner pore diameter for the three-fold	Normalized main chain hydrogen bonds	Normalized side chain hydrogen bonds
6 layer 3 fold *in vivo* model	12.5 (4.4)	66.2 (13.9)	20.2 (2.5)	0.36 (0.01)	0.15 (0.01)
6 layer 3 fold *in vitro* model	9.8 (1.7)	45.3 (10.1)	19.9 (1.4)	0.44 (0.01)	0.06 (0.01)
9 layer 2 fold *in vitro* model	9.8 (.5)	60 (17.0)	N/A	0.39 (0.02)	0.07 (0.01)
12 layer 3 fold *in vivo* model	9.6 (1.3)	42.2 (7.8)	19.4 (1.8)	0.32 (0.02)	0.16 (0.01)
12 layer 3 fold *in vitro* model	6.8 (0.5)	31.2 (4.9)	19.6 (1.6)	0.38 (0.01)	0.07 (0.01)
18 layer 2 fold *in vitro* model	7.7 (0.9)	36.4 (2.4)	N/A	0.36 (0.01)	0.07 (0.01)

Table 5: The average C_α RMSD (Å) calculated for residues 9-40 for all systems, Radius of gyration (Å), inner pore diameter for the three folds (Å), main chain and side chain hydrogen bonds. Standard deviations are calculated over three independent runs and listed in brackets.

In order to get an impression of the relative stability of the various aggregates we start our analysis by visually comparing the initial and final structures. Measurements of the root-mean-square deviation (RMSD) and Radius of Gyration (Rg) quantify these observations and are listed in **Table 5**. The three-layered *in vivo* model looses its initial structure in all three runs, as the three U-shaped subunits separate rapidly and dissociate, while the *in vitro* fibril model with three-fold symmetry is more stable but still subjects to large structural shifts. The low stability of the two models is not surprising as hydrophobic contacts at the C-terminal-C-terminal (CC) interface between the U-shaped subunits are not sufficient to stabilize the three-fold in small oligomers (i.e. three layers in each *strand-loop-strand* subunits) (**Figure 10**). Increasing the size to six layers increases the stability of both models. The U-shaped subunits now stay in contact throughout the simulation via the hydrophobic interface, but show some drift along this interface. The Radius of gyration is larger in the *in vivo* model than in the *in vitro* model indicating that the latter is more compact. The *strand-loop-strand*

conformation of the three subunits, which constitutes a common motif for the Aβ peptides,[168] is maintained in most of the aggregates; with the N-terminal residues tightly packed and therefore more stable than the C-terminal (C) residues. However, the contacts between the subunits are only partially retained. Hence, while with increased size the hydrophobic CC interface contacts in the six-layer systems become stronger, they are still not sufficient to stabilize its three-fold geometry (**Figure 11**). On the other hand, both twelve-layer systems keep the topologies and general characteristics of their initial conformations. While the exposed outer chains and turn regions are highly flexible, the sheet-turn-sheet topology is stable in both models. The hydrophobic contacts are more pronounced than in the smaller-sized systems, and stable throughout the simulations (**Figure 10** and **Figure 11**). The smaller Radius-of-gyration values show that both the *in vivo* and *in vitro* twelve-layer models with three-fold symmetry are more compact than the smaller aggregates of the same topology. At the same time, the differences between *in vivo* and *in vitro* models become smaller as the system size increases.

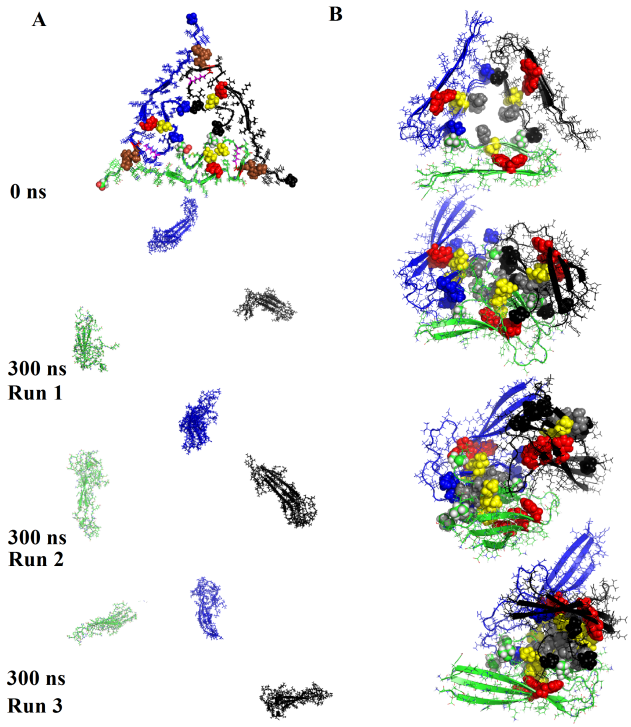

Figure 10: Snapshots of the three-fold *in vivo* and *in vivo* fibril models (consisting of 3 layers), at the start and end of a 300 ns molecular dynamics simulation. The three *strand-loop-strand* units are colored blue, green, or black. F_{19}, L_{34} and M_{35} are shown as spheres colored red, yellow or gray.

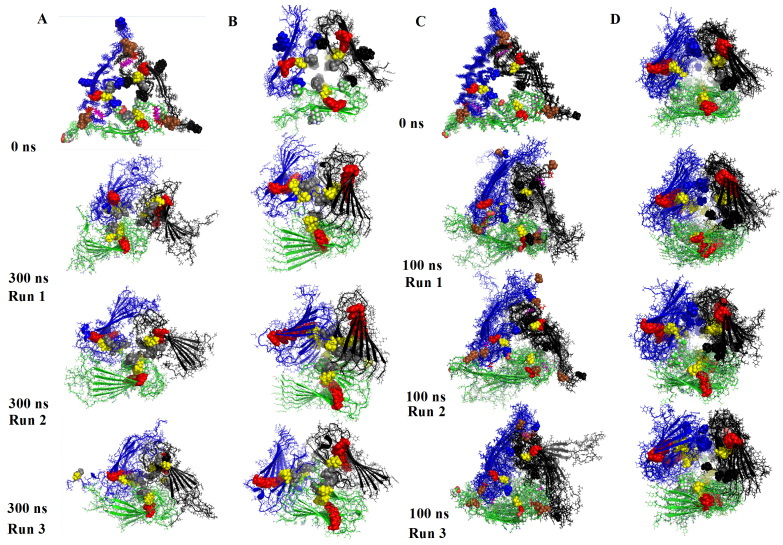

Figure 11: Snapshots of three-fold *in vivo* (A and B) and *in vitro* (C and D) fibril models consisting of either six or twelve layers, at the start and end of the respective molecular dynamics simulations. The three *strand-loop-strand* units are colored blue, green, or black. F_{19}, L_{34} and M_{35} are shown as spheres colored in red, yellow and gray.

A similar size dependence is also observed in the simulations of the *in vitro* models with two-fold symmetry. The size-dependent stability of all our oligomers is in agreement with previous numerical studies of amyloid micro-fibrils where the fibril geometry no longer changes with system size once a critical fibril length (about 20 layers) is approached [169]. Because the number of contacts along the CC interface is in the two-fold models larger than in the ones with three-fold symmetry, the two-fold models are more stable and the strands separate less than equally-sized models with three-fold symmetry (**Figure 12**). For instance, we find that the six-layered *in vivo* model with its three-fold symmetry has a larger root-mean-square deviation (12.5 Å) between final and initial configuration than the *in vitro* model of same size and topology (~9.8 Å), which in turn is larger than the equal-sized model with two-fold symmetry (~7.1 Å). Note that as size increases, the average root-mean-square-deviations became smaller, and the difference between the three systems decreases. Our simulations also indicate that the outermost strands are in all models (both such with two-fold symmetry and such with three-fold symmetry) more flexible than the strands located in the center (**Figure 10-12**). This instability at the edges of a growing fibril may provide a flexible docking point for incoming Aβ peptide while at the same time maintaining the underlying structure of the fibril [143,170]. As more Aβ peptides are added and the fibril grows, the previously exposed molecules will adopt a more stable structure (**Figure 10-12**).

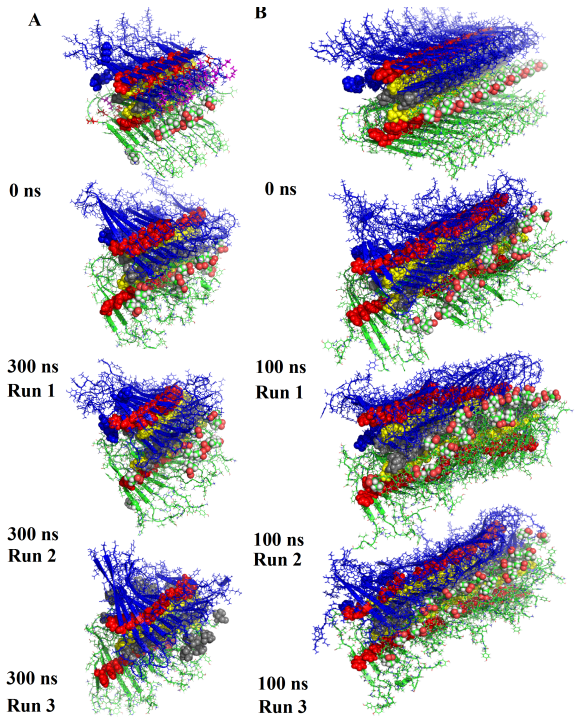

Figure 12: Snapshots of *in vitro* two-fold A*β* fibril model consisting of 9 layers (A) or 18 layers (B) at the start and end of molecular dynamics simulations of 300 ns and 100ns respectively. The two *strand-loop-strand* units are colored blue, and green. F_{19}, L_{34} and M_{35} are shown as spheres colored in red, yellow and gray.

Amyloids consist of extended intermolecular β-sheets, and the more stable the β-sheets are in an amyloid, the more likely the β-sheet and β-strands stack together tightly in an aggregate. Hence, high β-sheet content correlates with the stability of amyloids. For this reason, we have compared the changes in secondary structure between the energy-equilibrated structure (0 ns) and the last 100 ns or, for larger systems sizes, 50 ns, of the molecular dynamics trajectories. Our data, calculated with the DSSP [125] software, are listed in **Table 6** and indicate that the original parallel, in-register, β-sheet is conserved in all cases. The *in vivo* model has a twist in the β_2 strand that results in non-parallel contact of β_1 and β_2. As a result, it has about 10 % less residues in a β-sheet than the *in vitro* aggregates with three-fold symmetry of similar size. However, in both

the six and twelve-layer *in vivo* models the original β-sheet is maintained, and even increases slightly by ~4% and ~3%, respectively, in the last 100 ns (50 ns for the twelve-layer system). The twelve-layer *in vitro* model with three-fold symmetry also maintains its overall architecture, however, its β-sheet content decreases by about 6 %. The slightly higher level of regular β-sheet in the *in vitro* model results from the kinked and non-parallel sheets that differ in the location and length of residues. At small aggregate size, *in vitro* oligomers with two-fold symmetry have less β-sheet content than those with three-fold symmetry. This difference is because in oligomers of *in vitro* systems with three-fold symmetry, more of the β_2 region is buried inside the hydrophobic inner core. However, this difference becomes smaller as the systems grow, and is negligible for our largest models. This is because even in the oligomers with two-fold symmetry, the solvent exposure gets smaller and side chains are better packed as the number of chains grows, thus allowing the aggregates to maintain most of their initial β–sheet structure. The size dependence of β-sheet structure is in agreement with the root mean square fluctuations (RMSF) values shown in **Figure 13** that indicate significantly higher flexibility for the smaller size aggregates than observed for larger size aggregates.

System	Percentage of β-Sheet secondary structure of the energy minimized structure	Percentage of β-Sheet secondary structure averaged over last 50 ns
6 layer 3 fold in vivo model	29%	34% (1.7)
6 layer 3 fold in vitro model	45%	44% (1.4)
9 layer 2 fold in vitro model	36%	39% (0.9)
12 layer 3 fold in vivo model	28%	31% (1.2)
12 layer 3 fold in vitro model	45%	39% (0.7)
18 layer 2 fold in vitro model	38%	39% (0.5)

Table 6: Frequency ofβ-Sheet secondary structure in the three-fold and two-fold models. The values are presented as percentages for the energy-equilibrated structure (0 ns) and the last 50 ns of the molecular dynamics trajectory. Standard deviations are calculated over three independent runs and listed in brackets.

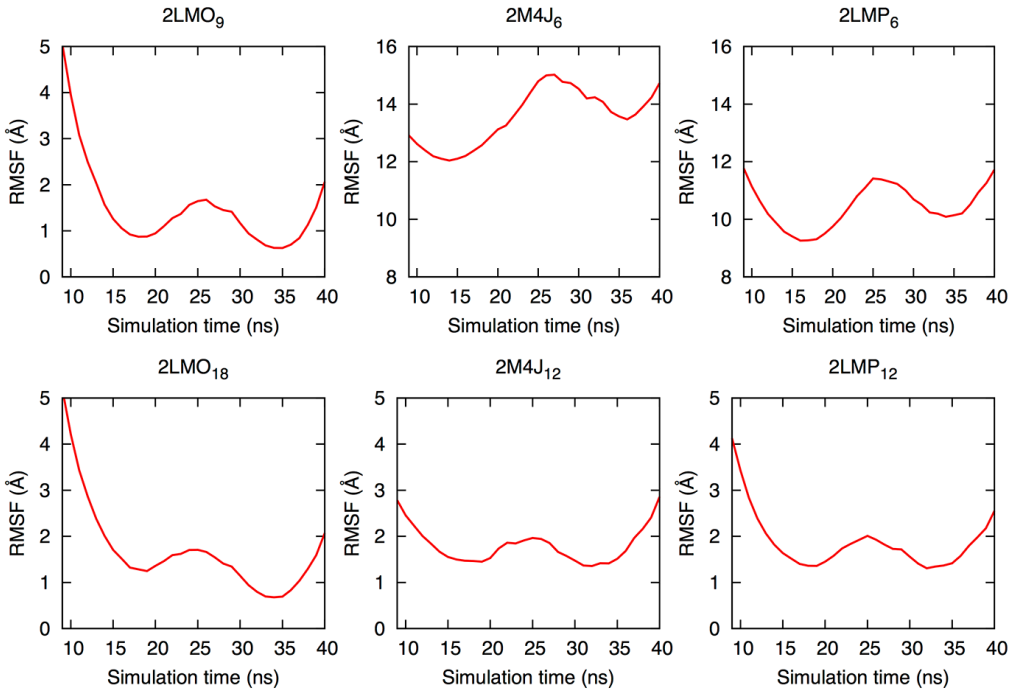

Figure 13: Average root-mean-square-fluctuation (RMSF) for the *in vivo* three-fold **(A and B),** *in the vitro* three-fold (**C and D**), and the *in vitro* two-fold systems **(E and F).**

The thermodynamic stability and the dynamical changes of the amyloid fibrils are modulated by local interactions such as hydrogen bonds or salt bridge, and hydrophobic side chain packing [8,55,171,172]. Polymorphism implies different arrangements of β-sheets in amyloid fibrils and therefore differences in the network of hydrogen bonds that determines the structural characteristics and properties of amyloids [55,115]. Previous computational studies have demonstrated that due to differences in the contribution of hydrophobic interaction and hydrogen bonds [169] oligomers with double fold are more stable than such with the three-fold symmetry. For this reason, we have also analyzed the hydrogen bond networks observed in our models (**Table 5**). The number of hydrogen bonds is normalized per amino acids to take into account the differences in size, and to enable comparisons between the *in vivo* models (consisting of residues 1 to 40) and *in vitro models* (consisting of only residues 9 to 40).

The *in vitro* fibrils have a larger number of backbone hydrogen bonds (**Table 5**) than the *in vivo* model. This difference highlights the reduced stability of this fibril and is in agreemnt with our visual inspection and the data for RMSD, Rg, RMSF and secondary structure discussed above. Similar to the other quantities, as the size of the aggregates increases, the difference in the number of hydrogen bonds between the various models becomes smaller. However, the *in vivo* model has an increased content of side chain hydrogen bonds compared to *in vitro* models. This is because of the higher flexibility of the N-terminal residues, and several kinks and bends (**Figure 13**) with a twisting movement (**Figure 9**), that bring side-chain hydrogen-bond donor and acceptor groups closer together, and increase in this way the number of side chain hydrogen bonds

(**Table 5**) in the N–terminal region. As a consequence, residues that are far apart are brought together in the *in vivo* fibril, thereby breaking the backbone hydrogen-bonding. Unlike in the folded state, where most interactions between residues are intra-molecular, inter-molecular interactions (steric zippers contacts) between amino acid residues[173] dominate in amyloids. These interactions include inter-chain backbone hydrogen bonding and such between side chains in adjacent β-sheets, or inter-peptide (i.e. lateral packing) hydrophobic interactions of hydrophobic residues [174].

We first measure the intra-molecular interaction through hydrophobic contacts between phenylalanine (F_{19}) and leucine (L_{34}) in adjacent β-sheets of the *strand-loop-strand* units (**Figure 9**). This allows us to explore the effect of this particular contact on the structure of the various aggregates (**see Table 7-8**). In the *in vitro* two-fold system of nine and eighteen layers the distance between F_{19} and L_{34} is about 10 Å (see **Table 8**). This distance is comparable to experimental values in which the β-sheets are typically apart by about 9-10 Å when two β-sheets pack together in amyloid fibrils. On the other hand, the F_{19}–L_{34} contact distance for models with three-fold symmetry is within the range of 11.5 Å (*in vivo* model*)* or 10 Å in the *in vitro* model (**Table 4**). Both values are again close to the experimentally measured ones, but slightly larger than the ones seen in the models with two fold symmetry. This indicates that side chains of amino acids present in the two β-sheets' intra-digitation of the hydrophobic steric zipper are responsible for retaining the *strand-loop-strand* motif. Visual inspection (**Figure 10-12**) also indicates that the difference in stability arises from the differences in the *strand-loop-strand* units interface contacts, and not from the instability of the *strand-loop-*

strand subunit themselves. Hence, the difference in inter-*strand-loop-strand units* contacts (as opposed to intra-*strand-loop-strand* units contacts) is responsible for the structural variation in the three systems.

Timeframe	6 layer 3 fold in vivo model	6 layer 3 fold in vitro model	12 layer 3 fold *in vivo* model	12 layer 3 fold *in vitro* model
▪	**<$F_{19}L_{34}$>**	**<F_{19}/I_{34}>**	**<F_{19}/L_{34}>**	**<F_{19}/L_{34}>**
Initial	11.6 (1.1)	10.9 (1.7)	11.2 (1.2)	10.3 (1.5)
Final 50ns	11.5 (2.9)	10.7 (2.1)	11.3 (1.1)	10.0 (1.2)
	<V_{40}/I_{32}>	**<V_{39}/I_{31}>**	**<V_{40}/I_{32}>**	**<V_{40}/I_{32}>**
Initial	18.4 (2.1)	10.3 (3.1)	17.1 (1.5)	10.2 (1.3)
Final 50ns	19.9 (1.5)	20.0 (1.2)	17.9 (1.1)	19.7 (0.8)
	<M_{35}/M_{35}>	**<M_{35}/M_{35}>**	**<M_{35}/M_{35}>**	**<M_{35}/M_{35}>**
Initial	23.3 (1.3)	18.8 (2.2)	23.5 (1.8)	18.2 (1.4)
Final 50ns	20.7 (2.7)	18.3 (3.0)	19.4 (1.5)	19.6 (1.2)
	<R_5/V_{24}>	**<H_{11}/V_{40}>**	**<<R_5/V_{24}>**	**<R_5/V_{24}>**
Initial	10.4 (2.4)	19.8 (3.1)	10.2 (1.5)	19.4 (1.5)
Final 50ns	26.1 (4.2)	17.6 (2.9)	12.9 (2.6)	17.6 (0.8)

Table 7: Hydrophobic contacts distances of the *in vivo* and *in vitro* three-fold systems. Measured hydrophobic contacts include such between F_{19} and L_{34} in each β-hairpin; the I_{32} and V_{40} form contacts between different cross-β units in the *in vivo* three-fold, and the corresponding contacts in the *in vitro* three-fold conformer formed by I_{31} and V_{39}. We also monitor in the *in vivo* three-fold contacts between the side chains of R_5 and V_{24} of adjacent layers, and the corresponding contacts in the *in vitro* three-fold between H_{14} and V_{40}. In three fold models, the hydrophobic cavity is formed by residues M_{35} and V_{40} for *in vivo* model and M_{35} for *in vitro* model. Standard deviations are calculated over three independent runs and listed in brackets.

Other stabilizing inter-molecular interactions are contacts between the side chains of residues located at the edge of the *strand-loop-strand* units (between the C-terminal and N terminal edges of adjacent *strand-loop-strand units*). Examples are the contacts between residues R_5 and V_{24} of the adjacent of *strand-loop-strand* units in *in vivo* three-fold, H_{13} and V_{40} of the *in vitro* three-fold, and H_{13} and V_{40} within the *strand-loop-strand* units of the *in vitro* two-fold. The contact distances in **Table 3-4** show that in the six-layer *in vivo* model with three-fold the Cα-Cα distance measured between residues R_5 and V_{24} grows from an initial 10 Å to 26 Å averaged over the last 100ns. Hence, side chain packing between R_5 and V_{24} is not sufficient to conserve the U-shaped

motifs. However, in the twelve-layer *in vivo* structure, the R_5 and V_{24} contact distance changes by only about 2.5 Å. The H_{14} and V_{40} contact distances at the edge of the *strand-loop-strand units* between the N-terminal and C-terminal decrease by approximately 1.5 Å to about 17.5 Å, irrespective of size (**Table 7**). Hence, similar combinations of favorable hydrophobic interactions stabilize the *strand-loop-strand* units in both the six and twelve layer *in vitro* models with three-fold symmetry.

<F_{19}/L_{34}>		**<F_{19}/L_{34}>**	
Initial	10.2 (0.8)		10.1 (0.5)
Final 50ns	9.9 (0.6)		9.4 (0.3)
<I_{31}/G_{37}>	**9 layer 2 fold**	**<I_{31}/G_{37}>**	**18 layer 2 fold**
Initial	10.1 (0.6)		10.1 (0.4)
Final 50ns	10.6 (0.4)		9.9 (0.3)
<H_{13}/V_{40}>		**<H_{13}/V_{40}>**	
Initial	15.8 (2.4)		15.9 (1.4)
Final 50ns	18.0 (0.4)		15.5 (1.2)

Table 8: Hydrophobic contacts distances in the in vitro two-fold systems. Measured hydrophobic contacts include these between F_{19} and L_{34} in each β-hairpin, and the G_{33} and M_{35} contacts between different adjacent β-sheet layers across the hydrophobic core. We also monitor in *in vitro* two-fold contacts between the side-chains of H_{13} and V_{40} in adjacent layers.

The interior channel in the three-fold has few hydrophobic contacts between CC termini-interfaces, while the experimental fibril model with two-fold symmetry mostly consists of a longer hydrophobic inter-molecular steric zipper. Thus contacts among subunits are obviously different in models with two-fold and three-fold symmetry. We monitor the distances between selected hydrophobic amino acids interactions that are involved in connecting the *strand-loop-strand* units at the points where they meet in the inner region of the amyloid fibril models. The side chains of V_{40} and I_{32} in the *in vivo* fibril model (V_{39} and I_{31} for *in vitro* three-fold) form hydrophobic contacts at the corners

of adjacent *strand-loop-strand* units (**Figure 9A-B**). The contacts between the V_{40} to I_{32} of adjacent *strand-loop-strand* units of the *in vivo* models and that of V_{39} to I_{31} between adjacent *strand-loop-strand* units in three-fold *in vitro* models are listed in **Table 7-8**. These distances (**Table 7**) grow from their initial values of ≈17.5 Å and 10 Å, respectively. A previous shorter simulation of the eight layers *in vitro* three-fibril model indicated values of ≈6-7 Å for the distance of V_{39}-I_{31} residues, a result suggesting that proximity of these residues stabilizes the cavity along the fibril axis. However, our much longer simulations indicate that such side-chain (V_{40}-I_{32} and I_{31}-V_{39}) packing is not sufficient to hold the adjacent *strand-loop-strand* units together. We rather believe that in the three-fold fibril models the stiff three *strand-loop-strand* units prevent internal hydrophobic segments from getting closer due to larger twist [175].

The hydrophobic contacts (**Table 8**) measured between F_{19} and L_{34} in each *strand-loop-strand units* for the two-fold *in vitro* models are around 10 Å, close to experimental values[92]. The side-chain contacts in the hydrophobic core of the double layer include those between I_{31} to G_{37}, G_{33} to M_{35} and G_{37} to I_{31} across the adjacent β-sheet. Their values in Table 4 indicate that Cα-Cα distance are within the range of 9.5 to 10.0 Å, which is characteristic for lateral association of two β-sheet. Visual inspection of the two-fold structures indicates that the larger steric zipper interface stabilizes such contacts much better than in the three-fold fibril models characterized by small steric zipper (with Cα-Cα at interface being ≈19.0 Å) and less efficient packing between U-shaped units [34,176]. However, the Cα-Cα distance between H_{13} and V_{40} in the termini of the *in vitro* two-fold model is larger, indicating flexibility of residues at edge. Hence,

our simulations imply that larger number of interface contacts in the two-fold systems leads to their higher stability over the three-fold systems, with the *in vivo* three-fold being the least stable (**Figures 10-12**). We conjecture that in aqueous environment nucleation of the *in vivo* model would be the slowest.

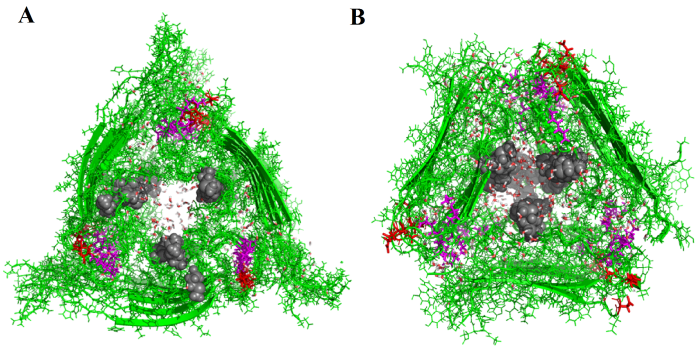

Figure 14: The amyloid aggregates of Aβ with *in vivo* (A) and the *in vitro* (B) three-fold accommodate water molecules in their hydrophobic core of the *in vivo* (A). Note that the corresponding experimental fibril structures of Aβ (PDB ID: 2M4J, and 2LMP) do not contain water.

As it is already known that in *in vitro* models with two-fold symmetry the hydrophilic cavities within *strand-loop-strand* units form a water channel [125,147,169], we limit here our examination to the systems with three-fold symmetry. Here, we find a water channel within the central hydrophobic cavity as the interaction between the *strand-loop-strand* subunits at their edge in the C terminal region of the three-fold creates a pore in which hydrophobic side chains of M_{35} are lining the inner surface. We estimate the pore diameter by measuring the distance between the central M_{35} residues of the adjacent U-shaped sub-units and averaging these values [177]. The observed pore diameter range of 19.5-20 Å for the 6 and 12 layer Aβ oligomers (**Table 5)** is comparable to the average pore diameter of 15□20 Å that was observed in experiments of amyloids

formed by $^{E22G}A\beta_{1-40}$ [170]. The resulting interior channels are water-accessible [150,178]. This can be seen in a snapshot from an equilibrated part of the simulation (5 ns of simulation) where water molecules enter into the interior channel (**Figure 14**) [147]. The average number of water molecules around each side-chain is shown in **Figure 15** for both the *in vivo* model and the *in vitro* systems with three-fold symmetry. The presence of water in the central pore [92,115] observed in our molecular dynamics simulations suggests cell membrane leakage as a possible mechanism for oligomer-mediated toxicity [170]. The possibility that oligomers of Aβ could be formed outside the membrane, bind to the cell surface and then span the membrane for the formation of active channels has been proposed by Shafrir *et al* [179]. Further computational studies study of Aβ aggregation with three-fold symmetry in the presence of lipid membranes may help to identify which of the amyloid peptide models are stable and remain sufficiently open for ion diffusion. Note that a similar pore with side chains of M_{35} lining the inner surface was also observed by Stroud *et al* [168] in the toxic Aβ42 fibrillar oligomer model.

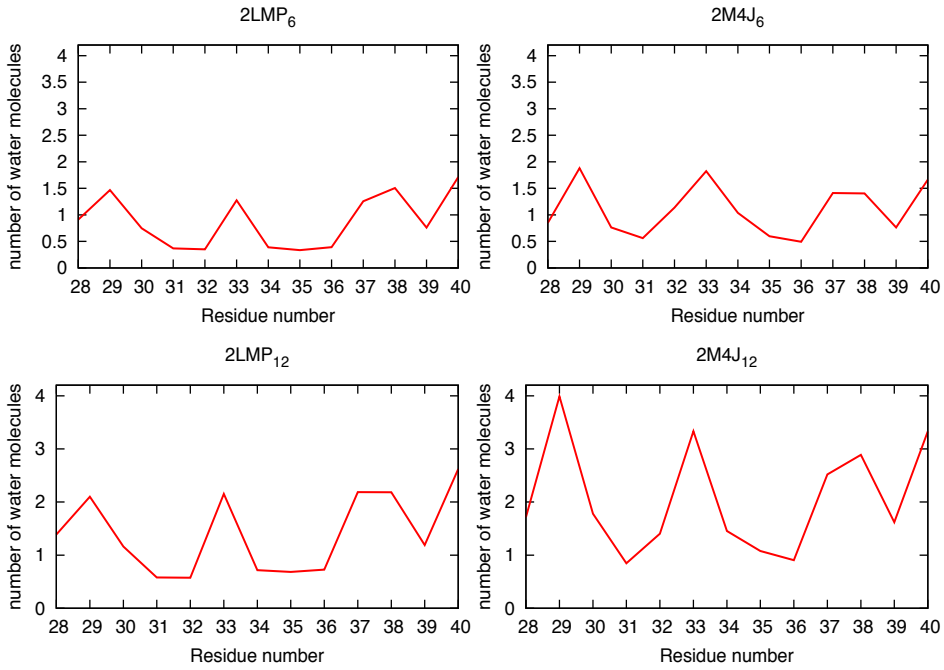

Figure 15: Average number of water molecules within 4 Å of each side-chain C_β carbon as measured for both *in vivo* and *in vitro* three-fold systems. Solvation exposure of side-chain hydrophobic core in the three-fold models and hydrophilic cavities in the two-fold system indicates the existence of water in the hollow. **(A)** *In vivo* three-fold with six layers ($2M4J_6$), **(B)** *in vitro* three-fold with six layers ($2LMP_6$), **(C)** *In vivo* three-fold with twelve layers ($2M4J_{12}$) and **(D)** *in vitro* three-fold with twelve layers ($2LMP_{12}$).

Discussion

We have compared the stability of various in *vivo* and *in vitro* Aβ fibril models using all-atom explicit solvent molecular dynamics simulations. We find water molecules flowing in the hydrophobic cavity of the models with three-fold symmetry (both *in vivo* and *in vitro*). This observation suggests water-leakage as a potential mechanism for the toxicity of amyloids. While *in vitro* grown Aβ fibrils are polymorphic, and their structures depend on the conditions encountered during their nucleation, fibrils acquired

post mortem from the Alzheimer's patient's brain lack such polymorphism and appear in a single, patient-specific form. Lu *et al* [44] discussed three possible reasons for the formation of a sole structure in patient brains: 1) that there is a mechanism for selection of a single oligomer via the brain's environment; 2) multiple oligomers form but only one is stable, either by internal or external means; 3) only one structure is stable enough to exist after death. Our simulations exclude the third possibility as they indicate that the three-fold brain-derived model is less stable than the *in vitro* Aβ amyloid fibrils, both such with three-fold symmetry and such with a two-fold structure. The weak stability we observed in our simulation contradicts the Lu et al [4] argument of the stability as criteria for the lack of polymorphism. While the current evidence we have is indirect, further evidence could be provided in future studies on quantitative information about thermodynamic stability using Free Energy Perturbation (FEP) or other accurate methods of free energy calculation. The difference in stability is due to the different frequencies of hydrophobic contacts between inter- *strand-loop-strand* sub units: the two-fold system is more stable due to a longer association of the β-sheets of the two adjacent *strand-loop-strand* units through tight side chain complementary. The difference in stability between the various forms depends on the size of the aggregates and decreases as the aggregates grow.

While our simulations indicate that the initial morphology is well maintained for all three models of the largest-sized aggregates, the significant lower stability of the *in vivo* model when compared to both forms of *in vitro* models (two-fold symmetry and three-fold symmetry) indicates that the brain-derived *in vivo* structures cannot be generated

easily under *in vitro* conditions, i.e. in solution. Note however, that our simulations are limited by the solvent model and require counter-ions to neutralize the system; whereas the fibrils formed under more physiologically relevant conditions include different ionic strengths and cofactor molecules such as membrane and proteins [173]. As the brain-derived aggregates strongly resemble fibril structures of Aβ formed in the presence of the lipid vesicles [180], it seems likely that the lack of polymorphism in the brain-derive fibrils is due to the different environmental condition by that the *in vitro* and *in vivo* models are grown. We anticipate that the results from our simulation will provide a basis for further studies into structural differences and differences in propagation mechanism between the brain-derived A β -fibril structure and the *in vitro* models, knowledge that is important for rational design of aggregation inhibitors and imaging agents.

Acknowledgment

This work is supported by the National Institutes of Health under Grant No. GM62838, and the National Science Foundation under Grant No CHE-1266256, and used XSEDE resources by the National Science Foundations and computational resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH1123. Other parts of the simulations were done on the BOOMER cluster of the University of Oklahoma. M.P. acknowledges support by the STEP-UP program administrated by OMHRC in NIH/NIDDK. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not

necessarily reflect the views of the National Institutes of Health, the National Science Foundation, the Department of Energy, or the University of Oklahoma. The authors have no conflict of interests to declare.

Chapter 5 - Stability Differences in the NMR Ensembles of Amyloid β Fibrils

The following chapter was published in similar form in the Journal of Theoretical Computational Chemistry by the author of this dissertation as the following article: Stability Differences in the NMR Ensembles of Amyloid β Fibrils. J. Theor. Comput. Chem. 15: 1650059. All text and figures are taken with the permission of the publisher.

Introduction

Alzheimer's, as well as other amyloid diseases, is characterized by the presence of amyloid fibrils that appear green under ultraviolet light after staining with Congo Red.[8,143] These fibrils contain highly ordered cross β arrangements held together by steric zipper van der Waals, hydrophobic packing, and hydrogen bonding networks.[9,10] The cytotoxicity and growth rates of the fibril aggregates are largely dependent on the structural arrangement and side chain packing.[93,115,181] In order to understand the pathogenesis of these diseases, one needs to know at a molecular level the structure of the fibrils and their precursors.[55] For instance, aggregates of the amyloid beta (Aβ) peptide are implicated in Alzheimer's disease, and differences in their molecular structure correlate with the disease progress.[182-184] Interestingly, fibrils derived post mortem from the brains of Alzheimer patients appear in a single, patient-specific form[44] while *in vitro* grown Aβ-fibrils are polymorphic.[7] In a recent study[185] we excluded the possibility that the lack of polymorphism in patient-derived fibrils is due to larger stability of the brain-derived model over that of the *in vitro* models. We arrived at this

conclusion from an analysis of the stability of various fibril fragments followed over long molecular dynamics trajectories. The molecular structure of these fibrils were resolved experimentally by solid state NMR (ssNMR). However, determining the structure of a protein or aggregate from measured NMR signals is an underdetermined problem.[13] Hence, molecular modeling with a given energy function is used to generate an ensemble of configurations that are compatible with the NMR signals. The subset of configurations with the lowest energies is then deposited in the Protein Data Bank (PDB), with the entries ordered according to their energy[14]. Following common practice[181], we have considered in our earlier study only the first entry (i.e., the one with the lowest energy). This approach is justified if one can assume that the NMR-derived-configurations form a good approximation of the canonical ensemble at physiologically relevant temperature. This approximation is less valid for solid-state NMR than for NMR of solvated molecules.[186,187] Considering only the first entry therefore may skew a computational study if the NMR entries differ significantly and the simulation is not sufficiently long to yield equilibrium configurations (which would be independent on the start configuration). In order to ensure that our previous results, comparing patient-derived and *in vitro* generated fibrils, are not biased by considering only the first NMR entry we study in the present paper how stability investigations of amyloids depend on the choice of a specific NMR entry as start configuration. In addition, we consider the role of staggering on the stability of fibrils.

For this purpose, we investigate the differences in stability between the members of NMR ensembles for the three-fold patient-derived fibril model (PDB-ID: **2M4J**)[44] and

the three-fold *in-vitro* fibril (PDB-ID: **2LMP**)[150], that were the focus of our earlier study. In a β hairpin loop, the relative position of the β sheets is used to determine the stagger of the overall structure. If there is no displacement such that the β sheets are aligned in the same plane on the z-axis, the structure does not posses staggering. However, displacement of the second β sheet of peptide chain *n* along the z-axis such that it interacts with the first β sheet chain *n+1* is positive staggering. Conversely, displacement such that the second β sheet of chain *n* interacts with the first β sheet chain *n-1* is negative staggering. This creates structures with unique sets of contacts meaning they may be functionally distinct.[35] The *in-vitro* model 2LMP has positive staggering, and in order to test the role of staggering on fibril stability we now consider additionally its counterpart with negative staggering (PDB-ID: **2LMQ**).[150] Note that no staggering has been observed in the patient-derived structure. While monitoring long molecular dynamics trajectories we measure quantities such as root-mean-square deviation, radius of gyration and secondary structure content for two members of each system: the first entry (i.e. the lowest energy configuration) and the entry in the NMR ensemble that has the largest Cα-Cα root-mean-square deviation to the first entry, i.e. differs most from it. These configurations are entry #20 (9.3 Å) for the patient-derived fibril model **2M4J**, entry #2 (3.7Å) for the positively staggered *in vitro* fibril model **2LMP**, and entry #6 (3.2 Å) for its negatively staggered counterpart **2LMQ**. The numbers in brackets are the root-mean-square-deviations of these entries to the corresponding first NMR entry. Comparing molecular dynamics simulations of these systems, we observe significant variations in molecular flexibility, compactness, and secondary structure between the entries. Of particular interest is the presence of a water-soluble central channel.

Experimentally, water accessible pores have been observed with a diameter of 15-20Å from $^{E22G}A\beta_{1-40}$ aggregates.[150,170,178] This data suggests a membrane leakage mechanism of cytotoxicity for Aβ. We previously observed the presence of channels of similar dimensions between residues 28-40 of both *in vivo* and *in vitro* models with 6 or more peptide layers with a pore diameter around 20Å. In this study we confirm that these channels are present in all simulated trajectories through comparison of the Inner pore diameter. To maintain parity with the original study, the inner pore diameter is estimated in the same manner, by averaging the distances between residues M_{35} of the same peptide layer. In addition, we find for the *in vitro* model that the negatively staggered models are less stable than the positively staggered one that we previously studied. On the other hand, the earlier focus on the first entry let to an underestimation of the stability of the patient-derived fibril. While our new results do not change the overall conclusions of our earlier investigation, they demonstrate that the full NMR ensemble and different staggering patterns should be considered for a physiologically relevant description of Aβ fibrils.

Materials and Methods

Model Generation

As in our previous work we consider fibril fragments of six and twelve layers in order to test how the stability of fibril fragments changes with size of the fragments. The patient-derived model **2M4J,** as deposited in the Protein Data Bank, contains three protein layers, while the *in-vitro* models **2LMQ** and **2LMP** has six protein layers, with each layer assembled from three coordinated Aβ chains and having three-fold

symmetry. In order to create a six-layer patient-derived model we duplicate the three-layer 2M4J structure, align them in the xy plane, and then moved them along the z-axis until the inter-strand distances are ~4.7 angstroms. In the same way, we generate twelve layer systems by aligning two six-layer systems, requiring again the inter-strand distances to be ~4.7 angstroms. As the first eight N-terminal residues are structurally disordered and not present in the *in vitro* **2LMP** and **2LMQ** ensemble entries, we exclude them in our simulations. However, these residues are present in the NMR ensemble data for the patient-derived **2M4J** model where they form junctures between adjacent strands in the same layer. For this reason, these residues are accounted for in our simulation of the patient-derived model, but in order to compare the different models these residues are excluded in the measurement of the analyzed quantities.

Simulation Protocol

In order to compare the stability of the various fibril fragments selected from respective the NMR ensembles, we run three long molecular dynamic simulations for each model, using as all-atom force field CHARMM27 (with CMAP corrections[188,189] and modeling solvent interactions with TIP3P water[190,191], a common computational setup for simulations of amyloid β peptide aggregation.[192,193] We use the GROMACS program (version 4.6.2)[194] adding hydrogen atoms with the pdb2gmx module. The peptide chains are put in the center of a box filled with water molecules such that there is at least a 12 Å distance between protein and the edges of the cubic box. Electrostatic interactions are calculated with the PME algorithm[115,195] to account for the periodic boundary conditions. The equations of motions are integrated with a time step of 2 fs. Hydrogen

atoms are constrained with the LINCS algorithm,[196] while for water the Settle algorithm is used.[197] The temperature of 310 K is kept constant by the Parrinello-Donadio-Bussi algorithm[198] (τ = 0.1 fs) which is similar to Berendsen coupling but adds a stochastic term that ensures a proper canonical ensemble.[199] In a similar way, the pressure is kept constant at 1 bar by the Parrinello-Rahman algorithm [200] (τ = 1 fs).

After solvation, the protein is allowed to relax to an energy-minimized state, which is subsequently equilibrated in a 2 ns NVT-ensemble simulation followed for another 2ns by a pressure coupled NPT ensemble simulation. After equilibration, three 300 ns molecular dynamic trajectories in a NPT ensemble are followed for the selected Aβ fibril structures. The trajectories differ in their randomly generated initial velocities and are considered by us as independent when calculating error estimates, i.e., we neglect correlations resulting from the three trajectories sharing the same initial fibril structure. Trajectory data are saved every 4 ps and examined using the analysis tools in the GROMACS software package. Specifically, we measure root-means-square deviations of the Cα atoms (RMSD), pore diameter of the oligomer cavity, secondary structure contents, solvent accessible surface area (SASA), radius of gyration (Rg) and hydrogen bonding. PYMOL is used for imaging and visualization of our data. As the *in vitro* fibril fragments do not include the first eight residues, we either exclude the residues when calculating quantities for the patient-derived fibril models or where appropriate by normalizing them according to the different number of residues.

Results

Due to the creation of artificial contacts between patient derived (2M4J) models in the construction of six and twelve layer systems from the original three layer structure, the validity of the created contact must first be confirmed before further analysis is performed. All artificial contacts are at the point of contact between the three layer structures, therefore the only contacts to be compared are interlayer. To this end, we compared the hydrogen bond occupancy of the created contacts to interlayer contacts present in the original structure presented in **Table 9**. This comparison indicates that there is little difference between the artificial contacts occupancy and the contacts present in the original structure, as well as the interlayer contacts of both *in vitro* structures. This shows that the hydrogen-bonding network at the artificial contact appears to have formed properly. This is also shown in the visual analysis of the system as many of the observed changes involve changes in intralayer contacts, which were not artificially generated.

H-Bond Occupancy	0-100ns artificial	0-100ns original	200-300ns artificial
2M4J(Entry 1) 6 Layer	70.3% (3.1)	69.9% (3.4)	75.1% (3.1)
2M4J(Entry 1) 20 Layer	72.8% (4.0)	70.1% (3.3)	79.2% (2.6)
2M4J(Entry 20) 6 Layer	71.6% (3.6)	70.5% (3.7)	76.5% (2.2)
2M4J(Entry 20) 20 Layer	73.0% (3.4)	73.2% (2.9)	80.1% (1.8)
2LMP(Entry 1)	N/A	72.7% (3.2)	N/A
2LMQ(Entry 1)	N/A	73.1% (3.9)	N/A

Table 9: A comparison of the hydrogen bonding occupancies of the artificial interlayer contacts to those present in the original structure. Note that since 2LMP and 2LMQ were deposit as six layer constructs, there is no artificial contacts

In a recent comparison of Aβ fibril models[185] we have seen that the stability of both *in vitro* and the patient-derived fibrils increases with the size of the fragments, but that the patient-derived fibrils are less stable than the *in vitro* ones. We arrived at this conclusion by visual inspection of 300 ns-long molecular dynamics trajectories which started with structures build from the first entries (i.e. the lowest energy configurations) of the respective NMR ensembles. Comparing these trajectories we found also a larger radius of gyration, lower β-sheet content and a lower number of backbone hydrogen bonds for the patient-derived fibrils. One goal of the present investigation is to test whether these previous results are biased by our choice of the start configuration. For this purpose, we have now added new simulations which start from these entries in the various NMR ensembles that differ maximally from the respective first entries.

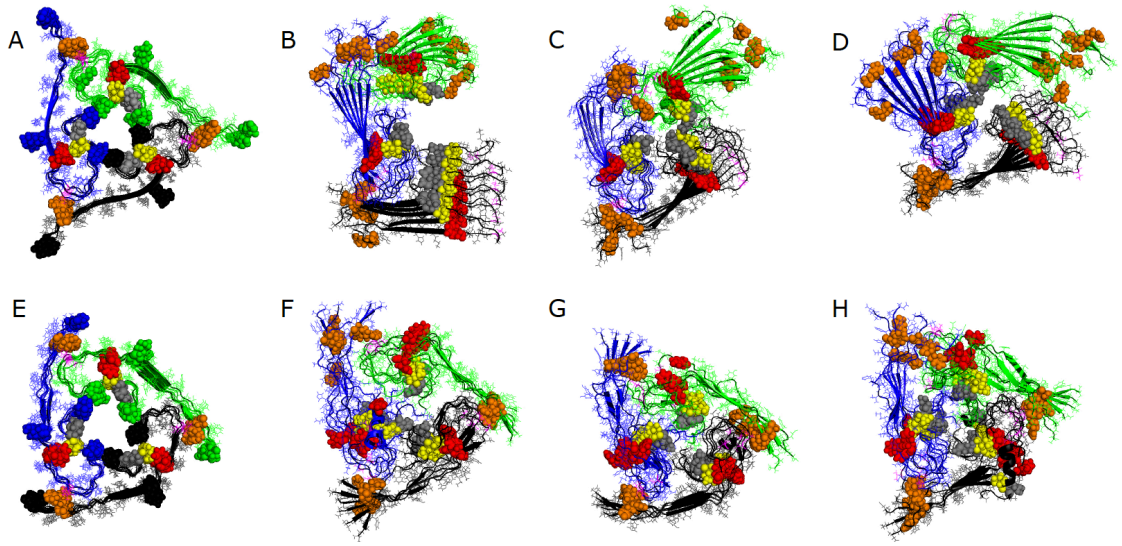

Figure 16: Start (A) and final configurations (B-D) of three 300 ns long trajectories from simulations of the six-layer model derived from entry #20 of the patient-derived fibril 2M4J. For comparison we show also the start (E) and final configurations (F-H) of three trajectories that start from models derived from lowest-energy configuration (entry #1) as used in our previous study.[185]

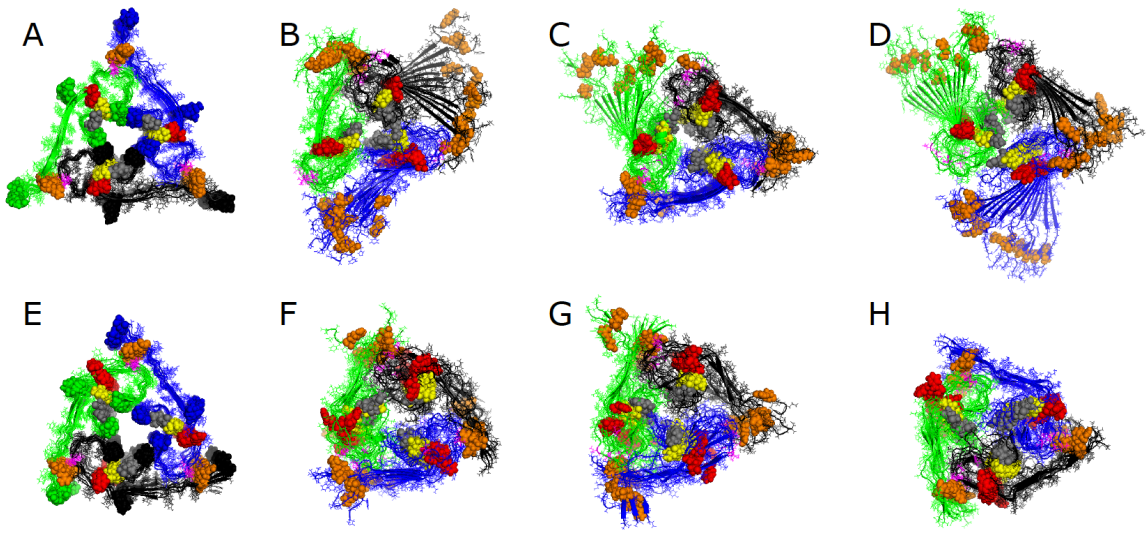

Figure 17: Start (A) and final configurations (B-D) of three 300 ns long trajectories from simulations of the twelve-layer model derived from entry #20 of the patient-derived fibril 2M4J. For comparison we show also the start (E) and final configurations (F-H) of three trajectories that start from models derived from lowest-energy configuration (entry #1) as used in our previous study.[185]

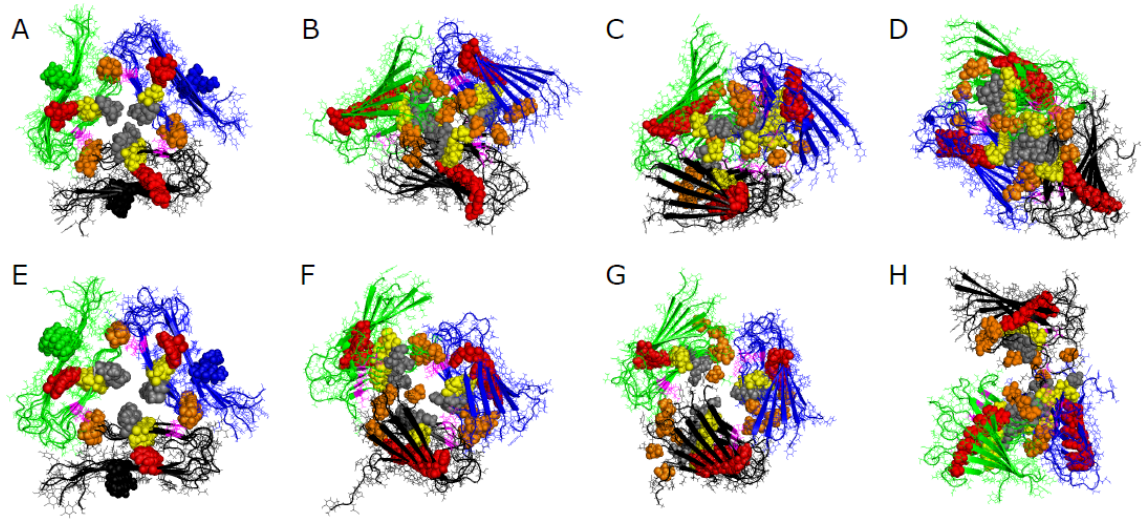

Figure 18: Start (A) and final configurations (B-D) of three 300 ns long trajectories from simulations of the six-layer model derived from entry #2 of the positive staggered *in vitro* fibril 2LMP. For comparison we show also the start (E) and final configurations (F-H) of three trajectories that start from models derived from lowest-energy configuration (entry #1) as used in our previous study.[185]

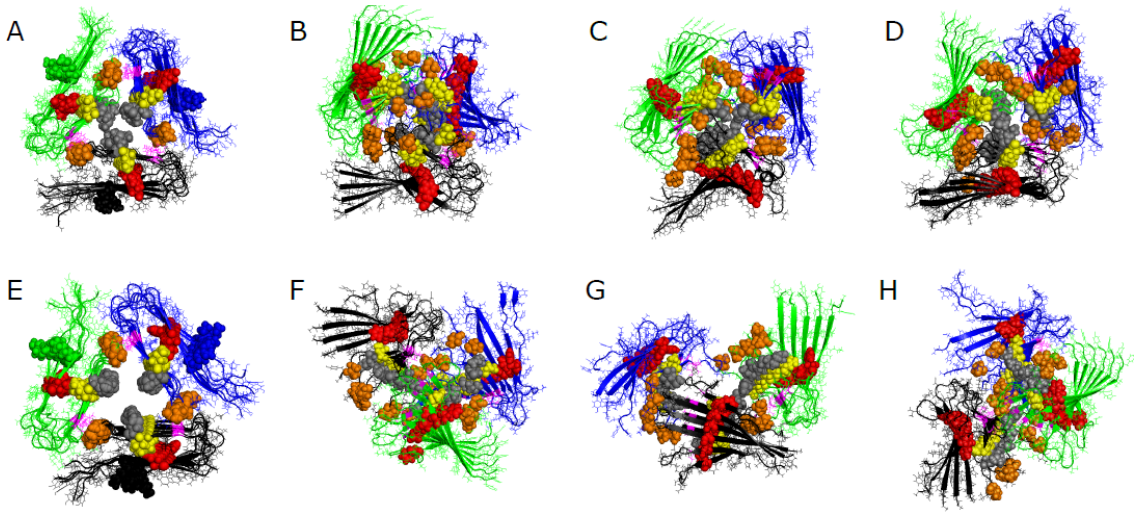

Figure 19: Start (A) and final configurations (B-D) of three 300 ns long trajectories from simulations of the six-layer model derived from entry #6 of the negative staggered *in vitro* fibril 2LMQ. In the second row we show also the start (E) and final configurations (F-H) of three trajectories that start from models derived from lowest-energy configuration (entry #1).

In same way as in our previous work, we start our analysis with a visual inspection of the various trajectories. As an example, we show in **Figure 16** the start and final configuration of the six-layer model derived from entry #20 of the patient-derived fibril **2M4J.** This entry differs from the previously used lowest-energy configuration (entry #1) by a root-mean-square-deviation of 9.3 Å. For comparison we show also the corresponding structures for the first entry that was used in our previous study. Similar figures for the same system with twelve layers are shown in **Figure 17**, and for the two *in vitro* systems in **Figures 18 and 19**. Visual inspection of these structures is complemented by measurements of the root-mean-square deviation (RMSD) and Radius of Gyration (Rg) as listed in **Table 10.**

For the patient-derived fibril (**Figure 16**) the lowest-energy configuration loses its central column over the course of the simulation, and the repeating chains become

elongated or stretched. On the other hand, we find that the fibril build out of the entry #20 remains intact over the course of the simulation and keeps its central column despite some 'shearing' along the fibril growth axis. This difference in stability is also seen in the time evolution of the root-mean-square deviation shown in **Table 10**: in the simulation that started from the first entry of the NMR ensemble, the final configuration differs from the start configuration by 14.5 (2.9) Å and has a radius of gyration of 55.6 (8.4) Å, while the final configuration in the simulation started that started from entry #20 has a radius of gyration of 37.9 (6.7) Å and differs by only 7.9 (0.7) Å from the start configuration. Comparison of the inner pore indicates that the channel is better maintained in trajectories started from entry #20, with an average diameter of 20.2 (1.1) Å compared to that of entry #1, 24.2 (3.2) Å. The difference between NMR models becomes smaller, but for the most part persists when going to larger systems. The inner pore appears to be of similar stability between the two entries with a pore diameter value of 20.1 (2.2) Å and 19.7 (1.2) Å for entry # 1 and #20 respectively. Any difference in the measure is likely due to noise. For the fibril with twelve layers is the root-mean-square deviation between final and start configurations 10.2 (1.9) Å when starting from the first entry, and 7.6 (0.7) Å when starting from entry #20. The radius of gyration at the end of the simulation again larger for the first entry than for entry #20: 48.2 (3.5) Å versus 43.4 (2.7) Å, but the difference between both values is also smaller than in the case of the six-layer system. This smaller difference can be also seen by visual inspection of the structures in **Figure 17**. Hence, we find that the stability of the patient-derived fibril in molecular dynamics simulations seems to depend strongly on the choice of the NMR entry that is chosen. However, this difference in stability

between the various entries decreases when going to larger systems. Note that the difference in stability between the two entries is not due to variations in their hydrogen bonding pattern. The less stable first entry has in both the six-layer and the twelve layer systems a larger number of main-chain - main-chain hydrogen bonds while the number of main-chain - side-chain and side-chain side -chain hydrogen bonds differs little, see **Table 10**. Together with the decreasing deviation in stability with increasing fibril size, this suggest that the d stability of the two entries depends not so much on the hydrogen bonding pattern but on the packing of hydrophobic residues. This can also be seen from the solvent accessible surface area (SASA) measurements that are listed separately for hydrophobic and hydrophilic residues in **Table 10**. For both the six-layer and the twelve-layer systems we find that the hydrophobic residues have a larger surface area exposed to the solvent for the less stable first entry than the corresponding residues have in the more stable entry #20.

System	C_α RMSD	Radius of Gyration	Inner Pore Diameter	Hydrogen Bonds			SASA		
				Main Chain	Side Chain	Mixed	Total	Non-polar	polar
Six-layer patient-derived model									
#1	14.5 (2.9)	55.6 (8.4)	24.2 (3.2)	0.40 (0.02)	0.12 (0.1)	0.13 (0.01)	538.0	180.9 (5.7)	357.0 (4.8)
#20	7.9 (0.7)	37.9 (6.7)	20.2 (1.1)	0.30 (0.02)	0.14 (0.02)	0.14 (0.01)	530.4	166.5 (3.2)	363.9 (6.0)
Twelve-layer patient-derived model									
#1	10.1 (1.9)	43.4 (2.7)	20.1 (2.2)	0.41 (0.02)	0.11 (0.01)	0.14 (0.01)	527.1	178.6 (6.1)	348.5 (3.2)
#20	7.6 (0.7)	48.2 (3.5)	19.7 (1.2)	0.28 (0.01)	0.15 (0.02)	0.13 (0.01)	506.2	158.8 (2.7)	347.4 (2.9)
Six-layer in-vitro model (positive staggering)									
#1	7.5 (0.3)	33.9 (7.8)	19.6 (1.4)	0.42 (0.01)	0.12 (0.02)	0.07 (0.01)	636.3	227.2 (5.5)	409.2 (5.8)
#2	7.0 (0.4)	33.5 (6.0)	24.1 (4.6)	0.43 (0.03)	0.11 (0.01)	0.07 (0.01)	651.7	242.3 (4.6)	409.4 (5.4)
Six-layer in-vitro model (negative staggering)									
#1	9.4 (0.8)	38.9 (7.7)	19.5 (1.8)	0.43 (0.02)	0.13 (0.01)	0.07 (0.01)	623.1	223.1 (5.1)	400.0 (4.1)
#6	8.6 (0.5)	40.8 (5.9)	24.9 (5.1)	0.41 (0.02)	0.12 (0.02)	0.06 (0.02)	640.4	240.7 (4.8)	399.7 (4.6)

Table 10: Average root-mean-square deviation between C_α-atoms (C_α RMSD) in the final and the start configuration (Å), calculated for Residues 9–40, Radius of Gyration (Å), Inner Pore Diameter (Å), number of Main-Chain – Main Chain, Side- Chain –

Side- Chain and Main-Chain - Side-Chain Hydrogen Bonds, solvent accessible surface area (SASA) for all residues, for polar residues only and for non-polar residues only. The various NMR ensemble entries are marked by their identifier

The variation in stability between NMR ensemble entries is also much smaller for the *in vitro* structures **2LMP** (positive staggering) and **2LMQ** (negative staggering) for which we only have studied six-layer systems. For the positive staggered system we find final root-mean-square-deviations of 9.4 (0.8) Å and 8.6 (0.5) Å for the two NMR entries, and corresponding values of 7.5 (0.3) Å and 7.0 (0.4) Å for the negatively staggered system, i.e. the staggering has a larger influence on the stability than the choice of NMR-entry. Note, however, that for both the positively and negatively staggered systems we find a stronger hydrophobic packing of the first entry in the respective NMR ensembles, see the variation in solvent accessible surface area for hydrophobic residues in **Table 10,** which leads to a higher stability of the central column and a about 5 Å smaller inner pore diameter**.** Visual inspection of the structures in **Figures 18** and **19** also shows that despite these differences the NMR entries of the *in vitro* structures vary less than the corresponding configurations of patient-derived fibrils. This smaller variability is likely due to the fact that the start configurations also differ less: ~3.1 Å for both *in vitro models* compared to a difference of 9.3 Å for the patient-derived model.

It is interesting to compare of the positive and the negative staggered *in vitro* structures. Focusing on the first entries of the respective NMR ensembles (the top rows in **Figures 18** and **19**) we see that both staggering arrangements maintain the overall three-fold structure throughout the 300 ns trajectory. This can be also seen from a comparison of radius of gyration and inner pore diameter, also shown in **Table 10**. However, the

negatively staggered 2LMQ is less stable than the positive staggered system, 2LMP, with final root-mean-square-deviations of 9.4 (0.8) Å versus 7.5 (0.3) Å, and while both systems started from similar start structures, the trajectories deviate over the 300 ns of simulation. In the positively staggered 2LMP the internal column is preserved as indicated by the triagonal arrangement of the gray M_{35} residues (**Figure 19)**. Furthermore, the yellow and red contacts (L_{34} and F_{12}) show that the hairpin units are well maintained.

We also see well ordered hairpin loops on the negatively staggered 2LMQ, however, the central column appears to be unstable and sometimes even re-assembles into a fibril with two-fold symmetry. The slightly smaller stability of the positive stagger model may be due to a lower β sheet content. Using the DSSP[201] implementation in GROMACS we have calculated the changes in secondary structure between the energy-equilibrated structure (0 ns) and the last 100 ns of the molecular dynamics trajectories. In the 2LMP structure are at the start 45% of residues part of a beta-sheet, and at the end only 42 (1) %, while in the 2LMQ structure initially 47 % of residues are part of a beta-sheet, and this number stays constant: 47 (2) %.

Discussion and Conclusion

We have compared in a recent article the differences in stability between an *in vitro* fibril with three-fold symmetry and a patient-derived fibril with the same symmetry. We extend these investigations in this article by addressing two potential shortcomings of the earlier work: first, we studied previously only the stability of the first entry (the

lowest-energy configuration) of the respective ssNMR ensembles; and secondly, we had ignored the existence of fibrils with different staggering for the *in vitro* fibril structure (the patient-derived structure does not have staggering). Hence, our previous conclusion, that the lack of polymorphism in patient-derived fibrils is not due to larger stability of the brain-derived model over that of *in vitro* models, depends on the assumption that fibril stability differs little between the ssNMR entries and does not depend strongly on the staggering. Evaluating now also the stability of different entries in the respective ssNMR ensembles, and of structures with other staggering, our new simulations lead to two observations.

First, the variation in stability between the NMR entries depends on the specific system and, not surprisingly, seems to reflect the diversity of the ensemble. For the patient-derived structure the difference in the entries is about 9 Å, and we find large differences in stability. On the other hand, for the in vitro fibril structures the differences between the NMR entries are only between 3-4 Å, leading to much smaller differences in stability. However, while this seems to imply that we have previously underestimated the stability of six-layered patient-derived fibrils, the new data do not change our conclusion. In addition, differences between NMR entries disappear with increasing fibril size. Already for our twelve-layer patient-derived fibril fragment are the deviations much smaller.

Our second observation is that for the *in vitro* model the negatively staggered models are less stable than the positively staggered one. The negative staggering disrupts stabilizing bonds between the three-fold motif leading to a rapid conversion to a two fold like structure. Hence, our new data show the importance of considering staggering

when evaluating the stability of certain fibril arrangements, however, both arrangements are more stable than the patient-derived fibril structure.

We conclude that while our new results do not undermine the overall conclusions of our earlier investigation, they demonstrate that the complete NMR ensemble and different staggering patterns must be considered for a physiologically relevant description of Aβ fibrils.

Acknowledgment

The simulations in this work were done on the BOOMER cluster of the University of Oklahoma. We acknowledge financial support from NSF CHE-1266256 and OCAST HR14-129

Chapter 6 - Molecular Dynamics Simulations of Early Steps in RNA-Mediated Conversion of Prions

The following chapter was published in similar form in Protein Science by the author of this dissertation as the following article: Molecular Dynamics Simulations of Early Steps in Rna-Mediated Conversion of Prions. Protein Sci 26:1524-1534. All text and figures are taken with the permission of the publisher.

Introduction

Prions are proteins commonly found in a wide range of species including bacteria and fungi. In mammals and birds they play a role in the growth and maintenance of neuronal synapses [56-58]; and mis-folded prions (usually addressed in this form as scrapie form or PrP^{SC}, while prions in their functional form are called PrP^{C}) assemble into cell-toxic aggregates that are the cause for a number of neurodegenerative diseases in humans (Creutzfeldt-Jacob, Kuru, Fatal Familial Insomnia) and animals (Scrapie, Bovine Spongiform Encephalopathy).[58,59] While the structure of the infectious PrP^{SC} state is not known, circular dichroism measurements indicate a lower helicity (43% in PrP^{C} compared to 30% in PrP^{SC}) and a larger β sheet content (3% in PrP^{C} compared to 43% in PrP^{SC}) than seen in the native PrP^{C} structures that are resolved and deposited in the Protein Data Bank.[66] In the often studied mouse prion (total length of 254 residues) the loss of helicity results from the helix proximal to the N-terminal domain referred to as helix A^{60-67} (residues 143 to 161) converted into a beta sheet, while the central helix

B (172-196) and the C-terminal helix C (200-229) stay intact. The now generally accepted protein-only hypothesis of Stanley Prusiner states that mis-folded prions are infectious, i.e., a prion protein in its disease-causing PrP^{SC} form can convert an unfolded prion, or one that is in its native PrP^{C} state, into its own form.[202] Hence, prion diseases progress by a nucleation mechanism where initial mis-folded prions seed the spread of the cell-toxic aggregates. There is evidence[183,203,204] that such "infectiousness'' of aggregates plays also a role in Alzheimer's and other amyloid diseases. Hence, the rate-limiting factor in all these diseases is the formation of the initial seed that nucleates the growth of the toxic aggregates. Data on the initial stages of aggregation can be immensely valuable as the list of diseases associated with fibrillar aggregates is numerous and continually growing.[205] Recent evidence[206,207] strongly indicates the pregnancy specific disorder preeclampsia to be associated with protein aggregations, which has previously been shown to be associated with elevated levels of expression for prion protein.[208] This knowledge underscores the potential impact of structural data for the initial stages of prion aggregation for a wide array of disorders affecting multiple areas of the body.

There is experimental evidence that polyanions, such as polyadenosine RNA (poly-A-RNA), can catalyze conversion of the native PrP^{C} form into PrP^{SC} seeds through interacting with the N-terminal of the prion at either a segment made from residues 21 - 31 or at the segment made from residues 111-121.[56,209] For instance, the deletion of polybasic domain 21-31 creates prion proteins that do not undergo conversion to the infectious state in the presence of poly-A-RNA.[210] However, it is not clear how

interactions between the RNA and these two polybasic segments lead to unfolding of helix A and its re-folding into a beta-sheet.

In the present paper we use molecular dynamics simulations to probe this mechanism and to identify early steps in the conversion of PrP^{C} into PrP^{SC} upon binding to RNA. The mouse prion, for which the wealth of experimental data is available, is compared with the medically more relevant human prion protein allowing us to assess the effect of sequence on the mechanism of conversion. As the N-terminal domain (residues 1 - 121) has not been resolved in the PrP^{C} structures of mouse (254 residues) and human (253 residues) prion protein (deposited in the Protein Data Bank (PDB) under identifiers 2L40 and 2LSB), it is not possible to study directly in *silico* how docking of RNA to the PrP^{C} structure modulates the stability of helix A and its conversion to a beta-sheet. For this reason, we have augmented the experimental structures with structure predictions of the N-terminal domain. Comparing proposals from the well-established MODELLER, which is based largely on homology modeling and refinement,[211,212] and ITASSER, which used Replica Exchange Monte Carlo simulations to construct models from known structural templates,[213-215] programs we generate initial protein structures for the docking of the prions in their PrP^{C} form with poly-A-RNA fragments. The quality of the predicted structures are inferred prior to docking by comparing the scores from three separate assessments: RAMPAGE[216], ERRAT[217] and ProQ[218]. The stability of such complexes, predicted by the docking software Autodock,[219] is then evaluated in 300 ns long all-atom molecular dynamics simulations with an explicit solvent. Through monitoring the protein's secondary structure, contacts formed with the RNA fragment and the protein (specifically helix A

and polybasic domains of the N-terminal region), and root mean square fluctuation of residues in both the bound and unbound state of prion proteins, we characterize how the interaction between RNA and prion protein modulates the transition from the cellular prion protein to its infectious scrapie form.

Materials and Methods

Model Generation

In mammals, prion proteins are anchored to the cellular membrane via a glycosylphosphatidylinositol (GPI) anchor at the C terminus, added to the protein after cleavage of the last 24 C-terminal residues.[57,58] Simulations of the protein anchored to the membrane are computationally expensive, and in order to explore efficiently the dynamics of RNA-prion interactions, we chose an unanchored prion model for our study. This requires to use in our simulations the non-cleaved protein (of length 253 residues for human prions or 254 residues for mouse prions) to prevent the presence of an erroneously charged C terminus. In order to study the conversion of the PrP^{C} form into the disease-causing PrP^{SC} state by molecular dynamic simulations we need a suitable model of the mouse or human prion protein in its native PrP^{C} structure. These structures have been only partially resolved by X-ray diffraction, with the N-terminal first 121 residues missing in the structures deposited in the Protein Data Bank (PDB) under PDB-ID 2LSB (human) and 2L40 (mouse). As binding of the RNA to the protein is supposed to involve these N-terminal residues, we propose to extend the PDB models by adding the N-terminal first 121 residues, taking as their structure one predicted by the well-established MODELLER or ITASSER programs.[211-215] While both software packages rely on aligning a protein sequence to existing structures, they

differ in the prediction protocols and protein models. Comparisons of the C terminal residues and critical polybasic residues between prediction algorithms showed an RMSD less than 5 Å showing key features were maintained regardless of method of structural prediction. Hence, by comparing two independent sets of predictions we hedge our study against potential biases. For both packages we select the two structures that have the highest ranking, and allow them to relax in a short, 5 ns long molecular dynamic run at T=310 K and 1 bar pressure. From each trajectory we then collect ten evenly spaced configurations for quality assessment. Model quality was calculated by averaging the scores of RAMPAGE[216], ERRAT[217] and ProQ[218] and comparing them to a cutoff value. We then find the RNA complexed structures using the docking software Autodock 25, dock these configurations to a five-nucleotide snippet of poly-A-RNA, allowing free rotation around all single bonds in the poly-A-RNA. Autodock 25 was selected as the docking software as it has been used extensively to model docking to both large proteins and RNA.[220,221] We choose a fragment size of five nucleotides because this was in recent experiments the minimal size where photo-degradation does not change the rate of conversion.[222,223] The predictions by either MODELLER or ITASSER lead to a total of two times ten docked structures for each software package (MODELLER/ITASSER) and protein (human/mouse), that are used as start point for short molecular dynamics simulations of 10 ns length at T=310 K. In 29 (25) of the 40 runs for the human (mouse) prion protein, the RNA got detached from the protein, and we discarded the corresponding docked structures. Analyzing the remaining 12 docked structures for the human prion, we identified three stable binding sites for the human protein. Site 1 is found five times and corresponds to the binding

between the RNA and the prion protein at the polybasic domain of residues 21-31. Site 2 involves the interaction with polybasic domain of residues 111-121 and is found four times, and site 3 (found three times) involves the residues 135-145 located in the N-terminal helix A. For the mouse prion protein we find four binding sites in the 15 docked structures that did not dissolve in the molecular dynamics runs. These include the three binding sites seen for the human prion protein and a fourth one, involving the residues 1-5 of the positively charged N-terminal. Site 1 is found five times, site 2 three times, site 3 four-times, and site 4 three-times. Out of the docked structures that share the same binding site we then choose for each protein (human/mouse) and prediction algorithm (ITASSER/MODELLER) a structure for further analysis with the smallest root-mean-square-deviation over the 10ns run, i.e, the structure that appeared to be most stable. Note that in the case of the human prion we also repeated the above protocol for the post-translated protein where the C-terminal 23 residues are cleaved. This test led to the same predicted binding sites. The complexes with full-length protein the residues 230-253 do not form contacts with either the poly-A-RNA fragment or helix A, and the structure of this segment changes by approximately 10 Å in the simulations described below. This indicated that residues 230-253 do not appear to interact with key regions for prion conversion, thus we conclude that the post-translational cleavage does not affect the conversion between PrP^{C} and PrP^{SC} form of human and mouse prion in our model.

Simulation Protocol

The stability of the so-derived 2 x 3 (2 x 4) docked structures predicted for the human (mouse) prion protein by either MODELLER or ITASSER was studied in longer molecular dynamics simulations, and compared with the outcome of control simulations of the two proteins (in their predicted configurations) without being docked to the RNA. These simulations rely on the GROMACS software package version 4.6.5[83] and utilize the CHARMM36 force field with associated nucleic acids parameters[30,224-226] and TIP3P water molecules[227,228], a common choice for simulation of amyloid-β systems which aggregate via pathways similar to that of prions[112,229,230]. A cubic box with a side length of 12 Å is placed around the center of mass of the docked and undocked peptide systems. As this solvent box has periodic boundaries, electrostatic interactions are calculated using the PME algorithm[231,232] (36,37). The various RNA-prion complexes are first minimized by steepest descent, before being equilibrated in a succession of a 2 ns molecular dynamics run in a constant number volume temperature (NVT) ensemble and a 2 ns run in a constant number pressure temperature (NPT) ensemble.

The six human prion and eight mouse prion protein structures, generated with the protocol described above, are the start point for long molecular dynamics simulations of 300 ns length in a NPT ensemble at 310 K and 1 bar pressure that allow us to probe the stability of the various systems. The equations of motions are integrated with a 2 fs time step, where hydrogen atoms are constrained by the LINCS algorithm[119] and water using the Settle algorithm (39). The temperature is held constant at a physiological

temperature of 310K by a Parrinello-Donadio-Bussi thermostat[120,233] (τ = 0.1 fs), and pressure is similarly held constant at 1bar by the Parrinello-Rahman algoritht (τ = 1 fs).[234] As the simulation was not being performed on GPUs, a group cutoff scheme was used, as we would not gain noticeable performance increases with a Verlet scheme. Due to the system size, neighbor searching was set to grid with a cutoff of 1.5 nm. Long-range electrostatics was handles by Particle Mesh Ewald with cubic interpolation and grid dimension set to 0.15 nm.

Generating different velocity distributions we follow each system in three trajectories in order to get a simple estimate for the statistical fluctuations. Data are saved every 4 ps to allow for analysis with the tools available in GROMACS. Primarily, we measure the following quantities: root-means-square deviations of the Cα atoms (RMSD), secondary structure contents, contact distances and hydrogen bonding footprint. Configurations are visualized using PYMOL.[43]

Results and Discussion

Model Confirmation

Quality assessment of predicted structures was preformed by comparing the scores of three separate methods of validation: RAMPAGE[216], ERRAT[217] and ProQ[218]. The cutoff used for the RAMPAGE score required the relaxed trajectory to have an aggregate of more than 85% of residues in favored regions and less than 0.8% (2 residues or less) in disfavored regions. Validation by ERRAT required the trajectories to have an average quality factor greater than 90%, as this suggests an acceptable

model[216]. ProQ predicts two scores, LGscore and MaxSub with cutoffs of 2.5 and 0.1 respectively.[235] Averages for all analyze structures are presented in **Table 11**.

Sequence	Rampage Favored	Rampage Allowed	RAMPAGE Disallowed	ERRAT	ProQ LgScore	ProQ MaxSub
Human Modeller	88.6% (3.5)	11.0% (2.1)	0.6% (0.2)	90.32% (2.1)	0.265 (0.05)	3.4 (0.4)
Human ITASSER	90.7% (2.8)	8.8% (2.8)	0.4% (0.2)	92.02% (3.2)	0.350 (0.10)	3.8 (0.6)
Mouse Modeller	87.2% (4.0)	12.4% (4.2)	0.5% (0.3)	90.03% (2.2)	0.27 (0.20)	3.5 (0.5)
Mouse ITASSER	89.3% (3.0)	10.3% (2.6)	0.4% (0.1)	91.91%(2.4)	0.330 (0.08)	3.7 (0.4)

Table 11: The average quality scores of trajectories used for docking provided by three separate methods of validation. Standard deviation is presented in parenthesis.

Visual Inspection

Eight docked structures (one for each of the four binding sites as predicted by either ITASSER or MODELLER) of the mouse prion – RNA complex, and six of the human prion – RNA complex, were followed in three independent trajectories over 300 ns to evaluate the stability of the complexes. We start our analysis of these 42 trajectories with a visual inspection of the final configurations. As a control, we also simulated the four unbound mouse, and three unbound human, prion protein structures, predicted by MODELLER and ITASSER, in three trajectories having the same length and using the same protocols as for the RNA-protein complexes. From the docked structures we were able to identify three binding sites for the human prion, and both docked and unbound structures are shown in **Figure 20**. The corresponding structures for the mouse prion, where we found four binding sites, are shown in **Figure 21**.

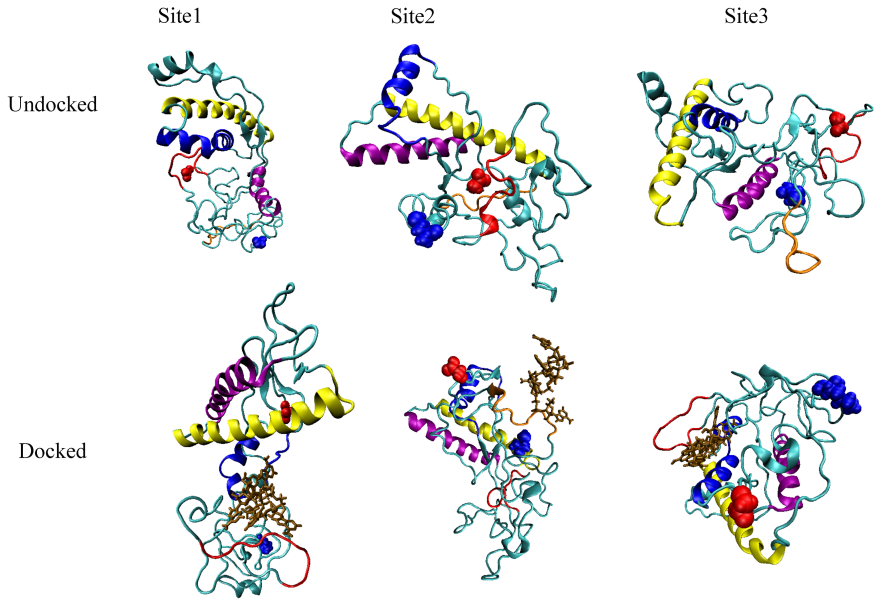

Figure 20: Binding motifs observed for the human prion protein. The upper row displays the unbound systems and the lower row the bound ones. Helix A is drawn in blue, helix B in yellow, the polybasic domain of residues 21-31 in red, and polybasic domain of residues 121-131 in orange (binding site 2 only). Blue spheres denote the N terminal domain while red spheres denote the C terminal domain.

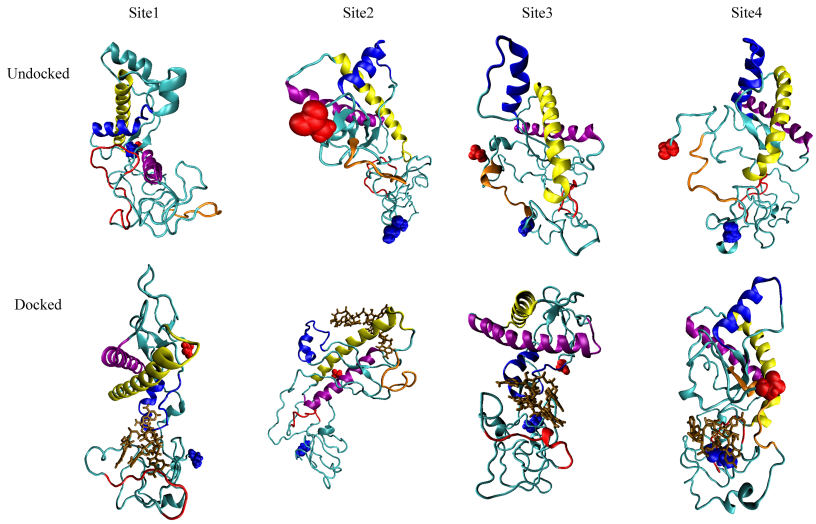

Figure 21: Binding motifs observed for the mouse prion protein. The upper row displays the unbound systems and the lower row the bound ones. Helix A is drawn in blue, helix B in yellow, the polybasic domain of residues 21-31 in red, and polybasic domain of residues 121-131 in orange (binding site 2 and 4 only). Blue spheres denote the N terminal domain while red spheres denote the C terminal domain.

In the six trajectories that followed the time evolution of the RNA docked to the mouse prion protein at site 1 (the polybasic segment of residues 21-31), a pincer-like structure between helix A (shown in blue) and the polybasic domain (shown in red) is quickly formed and encapsulates the RNA (shown in brown). This pincer motif is also seen in the corresponding six trajectories for the human prion protein bound with RNA at the polybasic segment of residues 21-31. This motif is characterized by distinct contacts between RNA and Protein which we list in **Table 12**. Note that this motif does not differ between structures generated by MODELLER or ITASSER, but it is not observed in the control simulations of the unbound proteins. In the trajectories of the prion-RNA complex where this motif appears, helix A dissolves over the course of the simulations, but not in the control simulations of the unbound proteins. Since this loss of helicity in the bound protein occurs in the same region where it is seen in the conversion to the infectious state, we conclude that this polybasic-helical pincer motif initiates the conversion to the infectious state. This direct interaction with helix A does not affect the secondary structure of the other helixes B and C, again in line with experimental results.[62] In the region of alpha helix A, and the surrounding coils, we see the generation of transient β structures indicating the potential conversion to the PrP^{SC} state as shown in **Figure 22**. However, these structures are not consistently observed across all trajectories and when observed oscillate between coil and sheet conformation, indicating that our 300ns trajectories are not sufficiently long to sample the conversion to the infectious state.

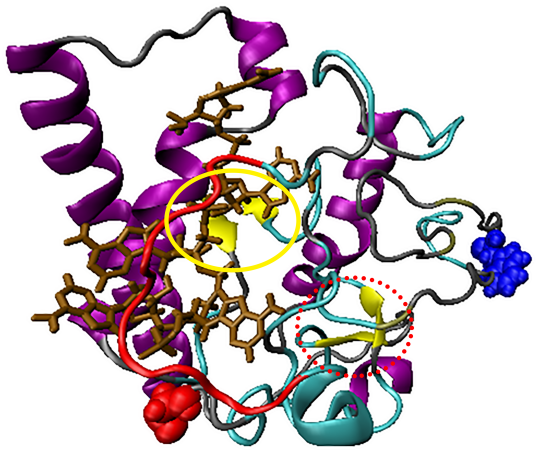

Figure 22: Secondary structure of the prion protein for system of RNA docked to site 1. Residues with β sheet secondary structure are indicated by yellow; α helices are indicated by purple; RNA is indicated by brown; coils are indicated by gray, and the polybasic domain 21-31 is indicated by red. Transient β strands around docking site are circled in a red dotted line while sheets that are from the PDB are circled in a yellow solid line. Blue spheres denote the N terminal domain while red spheres denote the C terminal domain.

0-100ns	**200-300ns**
Contacts with Helix A	
144 ASP	144 ASP
145 TRP	145 TRP
147 ASP	147 ASP
148 ARG	148 ARG
	139 HIS
	140 PHE
	146 GLU
	149 TYR
	150 TYR
Contacts in the Neighborhood of Polybasic Domain of Residues 21 - 31	
25 ARG	25 ARG
27 LYS	27 LYS
34 GLY	34 GLY
35 GLY	35 GLY
33 THR	33 THR
41 GLN	41 GLN

Table 12: Contacts between RNA and prion that define the helix pincer motif.

In the preliminary short trajectories, a strong binding of the RNA molecule to the prion proteins was also observed for binding site 2, defined as the region around the polybasic

domain of residues 111-121. In the 300ns long trajectories of complexes formed with this binding site we do not see the formation of the pincer-motif seen in the simulation of complexes with binding site 1. This is likely because helix A (residues 140-161) is too close to the binding site and a pincer arrangement would lead to steric clashes. Consequently, the secondary structure changes not more in the simulation of the bound prion protein than in that of the unbound protein. Again we find little difference between human and mouse proteins, and models generated with either ITASSER or MODELLER.

The RNA fragment can also interact with mouse and human prion proteins at binding site 3, made of residues 135-145, which partially overlaps with the helix A (residues 140-161). As with the complexes involving binding site 1, we observe in multiple trajectories the formation of the helical-polybasic pincer motif, and the dissolution of helix A happens even faster when the RNA binds to site 3 than when it binds to site 1. This may be because when binding at site 3, the RNA binds directly to the helix A, and only later forms the helical-polybasic pincer, which then traps the RNA in a region close to the helix A. This observation is further supported by data from two trajectories for RNA bound to the mouse ITASSER structure at site 3. While in two of the three trajectories the RNA separated from the protein (at ~75ns and ~95ns, respectively), the complex persisted in one trajectory where, unlike in the other two trajectories, the pincer between the polybasic domain and helix A is formed and helix A dissolves.

Binding site 4 involves the RNA interacting with residues 1-5 of the N terminal domain and occurs only in mouse prion, despite that mouse and prion protein have the same sequence in this region. However, in the three structures where the RNA binds to the protein at site 4, the N-terminal residues 1-5 are within 5 Å to the polybasic domain 111-121 with which they form a pocket around the RNA. As the unbound systems, RNA-bound systems at binding site 4 do not change secondary structure. Because of the close proximity to the polybasic domain of residues 111-121, one may consider binding site 4, which is only observed for a single structure predicted by MODELLER, as a variant of binding site 2. As most of the structures, predicted by MODELLER, have the first five N-terminal residues within 5 Å to the polybasic domain 111-121, binding site 4 may be in artifact of the protein structure prediction method.

Our visual inspection of the 42 trajectories suggest that RNA can initiate conversion into the infectious scrapie form when binding to the prion protein at sites 1 and 3. Hence, we focus in our further analysis on these two binding sites. As we find little differences between ITASSER and MODELLER generated structures, we will no longer distinguish between them, but combine them in our analysis. Hence, in the following analysis our statistics are calculated for prion proteins of respective human or mouse sequence from the total of 6 trajectories (3 from ITASSER and 3 from MODELLER) for both binding sites 1 and 3.

RMSF Analysis

One way to quantify the effects of docking RNA to specific regions of the prion protein is by comparing the root mean squared fluctuation (RMSF) of residues in the various RNA-prion complexes. These are shown in **Figure 23** for all residues of human or mouse prion. The figure displays for each residue the ratio of RMSF measure for the bound protein divided by the corresponding value measured for the unbound prion. A value larger than one therefore implies that a given residue is more flexible when RNA binds to the prion than in the unbound protein. In agreement with the visual inspection of the various trajectories, we find for both human and mouse prion docked to either site 1 or 3 a characteristic spike in region 150-180, which includes many residues of helix A (140-161), that is not seen for the other binding sites (data not shown). Binding to these two pockets leads for the human prion protein also to increased flexibility in the region of residues 35 – 90. This higher flexibility results from a loss of contacts between polar residues in this segment and helix A after binding of the RNA to the protein. The frequency of such contacts falls, with the standard deviations between trajectories presented in parenthesis, from 32.1% (7.2) in undocked structures to 12.1% (2.9) in docked structures. A similar decrease of frequency for such contacts is also observed for the mouse protein when the RNA is binding to site 1 and site 3, declining from 31.1% (4.4) in undocked to 10.1 (3.3) in docked structures, and consequently an increase in flexibility is observed for the segment of residues 35-90.

The distribution of the residue RMSF corroborates our previous observation that formation of the pincer-like structure leads to dissolution of helix A. This dissolution is

also seen in **Table 13** where we show how the probability to find 1+4 - backbone hydrogen bonds, characteristic for α- helices, changes along the trajectories. The helix starts dissolving by losing hydrogen bonds involving residues 144-148, which are also the ones that first form contacts with the RNA. As the simulations evolve, further helical back bone hydrogen bonds break, especially the bond between Tyr150 and Glu146, and TYR149 and ANS153; all residues that now also form contacts with the RNA fragment.

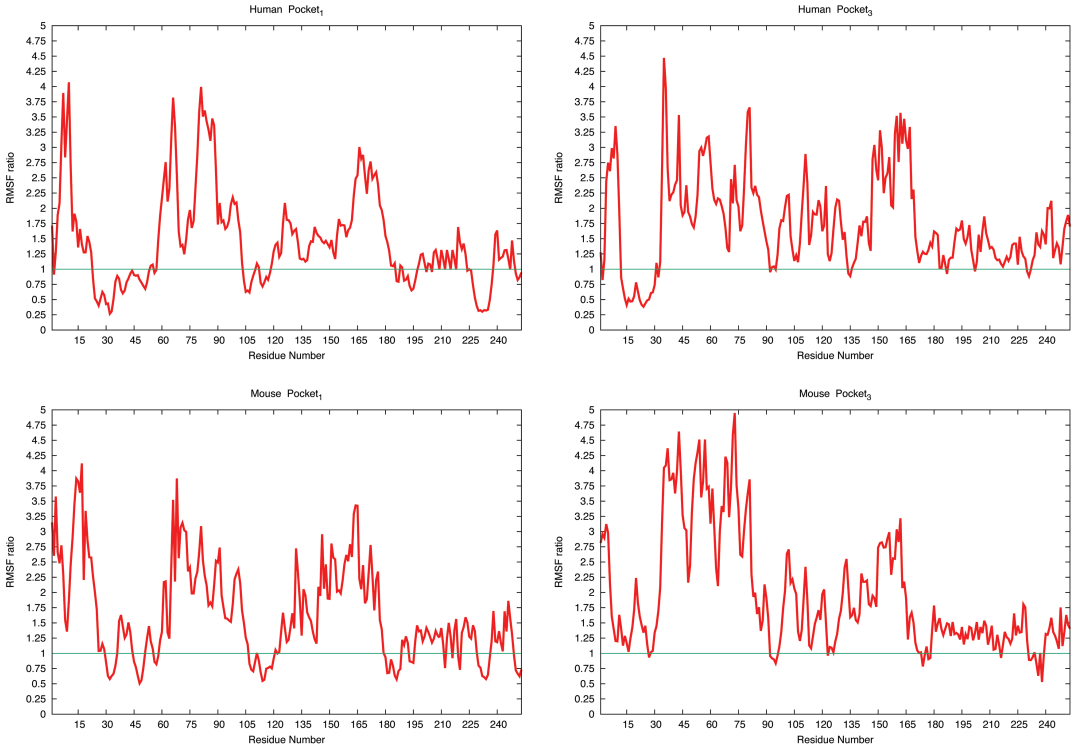

Figure 23: RMSF values measured for residues in docked structures divided by that measured in undocked structures. Ratios for the human prion protein are shown in the upper row, and ratios for the mouse protein in the lower row. The left column shows the ratios for structures docked to binding site 1, and the right column that for binding site 3. The ratio is calculated by dividing the average RMSF value for a given residues in the docked protein-RNA structure to the average RMSF of the undocked protein. The green line indicates a ratio of one.

Backbone Hydrogen Bond	Control Undocked	RNA Docked 0-100ns	RNA Docked 200-300ns
VAL161-PRO157	50.0% (3.7)	45.1% (4.6)	10.1% (2.5)
VAL160-HIS155	75.4% (1.2)	55.4 (3.1)	15.6% (2.5)
GLN159-MET154	88.6% (3.6)	61.5% (3.7)	20.4% (4.1)
ASN158-ASN153	81.7% (3.5)	65% (4.4)	19.5% (3.5)
PRO157-ASN153	45.5% (4.1)	43.4 (3.8)	20.4 (4.0)
ARG156-GLU152	78.7% (2.3)	65.5% (7.7)	20.9% (3.7)
HIS155-ARG151	97.9% (4.2)	81.0% (8.3)	51.0% (3.4)
MET154-TYR150	74.3% (3.1)	65.6% (7.4)	30.5% (4.4)
ASN153-TYR149	86.1% (3.8)	50.1% (4.2)	0% (0)
GLU152-ARG148	97.1% (2.1)	71.70% (9.2)	21.70% (4.2)
ARG151-ASP147	94.7% (3.0)	40.6% (5.3)	20.6% (3.5)
TYR150-GLU146	98.8% (1.3)	60.1% (4.9)	10% (2.1)
TYR149-TRP145	91.5% (4.5)	41.2% (4.1)	15.0% (4.8)
ARG148-ASP144	75.6% (5.1)	42.5% (6.2)	22.5% (4.9)
147ASP-143ARG	96.9% (1.9)	62.1% (4.8)	25.6% (3.2)

Table 13: The probability of a given helical contact averaged over the trajectories where the polybasic-helix pincer motif is seen.

The net- effect of the loss of backbone hydrogen bonds in helix A and the newly formed contacts of residues in this segment with the RNA is an increase in the total number of hydrogen bonds; i.e., the binding of the RNA is despite the loss of helix-stabilizing hydrogen bonds energetically favorable. This can be seen from **Table 14** where this relative increase is shown for both human and mouse prion protein, at both binding sites. The control simulations show no noticeable growth in the number of hydrogen bonds while the bound forms do show an increase, with a larger growth observed for binding to site 1. In order to emphasize this point we show in **Table 14** also the differences in the total number of hydrogen bonds. Average over the first 100ns there is only a weak signal, in that there is only an increase of about six hydrogen bonds; however, this number increases to about ten or more hydrogen bonds over the last 100 ns of the simulations. Hence, despite the dissolution of helix A additional backbone

hydrogen bonds appear, possibly indicating the start of formation of another ordered structure.

	0-100ns			200-300ns		
	Ctrl	**Docked**		**Ctrl**	**Docked**	
Name	Average Number MC-MC H Bonds	Average Number MC-MC H Bonds	Δ Bonds 0-100ns	Average Number MC-MC H Bonds	Average Number MC-MC H Bonds	Δ Bonds 100-300ns
Pocket 1 Human	78.3 (2.9)	84.2 (4.8)	5.9 (3.5)	78.9 (2.3)	91.4 (3.0)	12.5 (2.7)
Pocket 3 Human	71.9 (3.4)	78.2 (5.4)	6.3 (4.2)	84.6 (2.9)	94.9 (3.0)	10.3 (2.9)
Pocket 1 Mouse	72.5 (4.0)	77.6 (4.8)	5.1 (4.1)	71.4 (4.3)	84.1 (3.7)	12.7 (3.9)
Pocket 3 Mouse	75.1 (4.5)	79.1 (3.3)	4.0 (3.5)	74.7 (4.5)	75.5 (3.9)	10.8 (4.2)

Table 14: Number of main chain - main chain hydrogen bonds averaged over three trajectories for each pocket of the human sequence. Δ Bonds is calculated as the difference between the docked and undocked average main chain – main chain hydrogen bonds.

The decrease in helicity resulting from the loss of stabilizing backbone hydrogen bonds in helix A is also seen in **Table 14**. While there no change in the control, the average helicity decreases from about 43% to 32% for both proteins and binding sites. However, binding of the RNA with the protein leads not only to a loss of helical structures but also to a gain in β arrangements. This effect is small but significant if one considers in addition the life times that such sheet-like element exist. For this reason, we list in **Table 15** also the β strand occupancy, defined as the average amount of time a β-strand can be observed along the trajectory. While there is only an increase of ~4.5% in total β strand propensity, the average life time of observed transient β strands grows by approximate 25%, which may indicate the beginning transition to the β sheet rich PrP^{SC} structure.

Time	Human		Mouse	
	Control	Docked	Control	Docked
Helicity				
0-100 ns	43.1% (1.8)	38.9% (2.7)	39.2% (0.9)	36.6% (2.2)
200-300 ns	41.9% (1.6)	32.0% (2.3)	42.7% (1.3)	31.5% (2.4
β Strands				
0-100ns	4.3% (1.3)	6.9% (1.6)	4.2% (1.2)	6.5% (1.7)
200-300ns	4.5% (1.2)	8.8% (1.8)	4.4% (1.5)	9.1% (1.4)
β Strand Occupancy				
0-100 ns	26.7% (1.4)	38.4% (4.6)	26.2% (1.3)	37.8% (4.0)
200-300 ns	27.9% (1.3)	57.9% (4.2)	28.2% (1.2)	57.3% (4.5)

Table 15: Average secondary structure content and β-strand occupancy of the C terminal domain (residues 121 to 253) of for all trajectories of structures with polybasic-helix pincer, i.e., binding sites 1 and 3. The averages are calculated for the first and for the last 100ns in order to show the evolution of the structures. Occupancy is defined as the percentage of time a given residue exists in a beta strand confirmation.

In order to emphasize the shift in hydrogen bonding and the corresponding structural rearrangements we show in **Figure 24** a contact map of the human prion (upper triangle) and mouse prion (lower triangle). For this figure, we measured the frequency of backbone-backbone hydrogen bonds in both docked and undocked structures, and the color coding marks the difference between these two frequencies. A reddish coloring indicates that a given hydrogen bond is more frequently seen in the bound structure than in the unbound one, and a greenish color that this hydrogen bond is more common in the unbound structure than in the bound one. We find the strongest signal for contacts involving residues 140-161, which are part of a helix in the native structure. In this region, there is a decline of helical contacts (see the greenish coloring of points parallel to the diagonal that mark 1-4 hydrogen bonds) and a corresponding increase in contacts that suggest β arrangements (the reddish colored data points orthogonal to the diagonal).

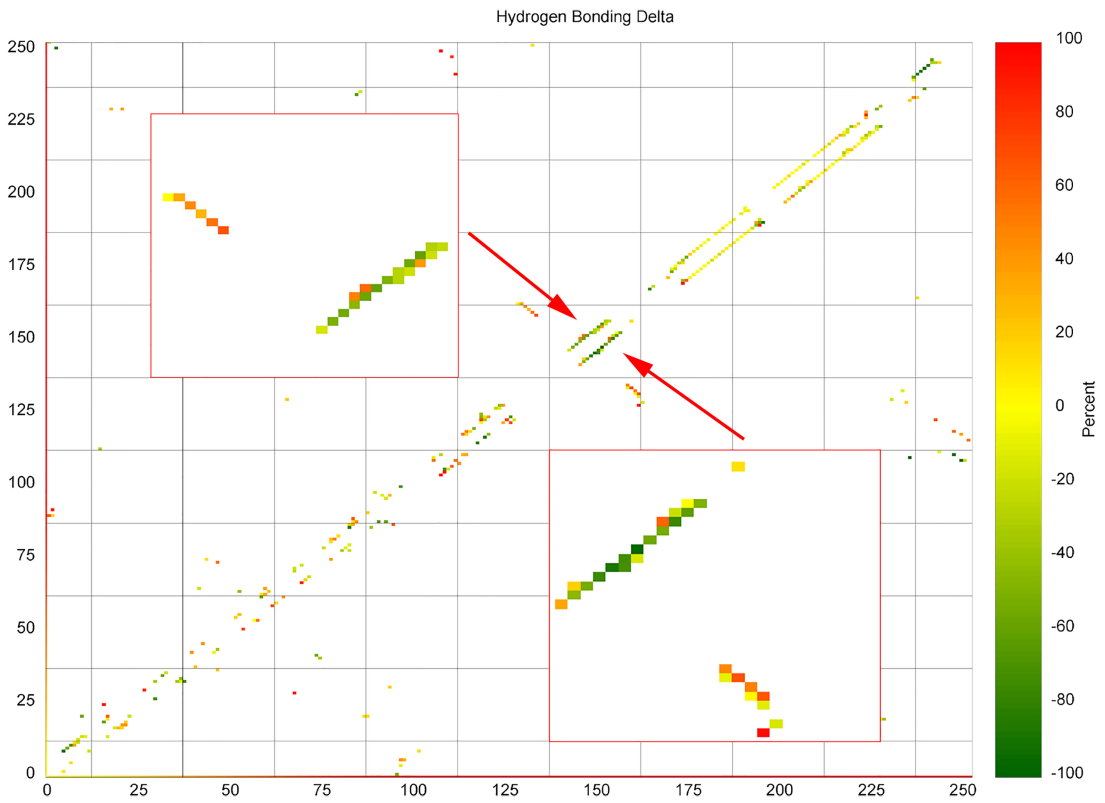

Figure 24: Difference between the frequency of backbone H-bond contacts found in the bound prion protein and the frequency of such contacts found in the unbound protein. Diagonal elements correspond to helical contacts and linear clusters orthogonal to the diagonal correspond to β strand contacts. The upper triangle shows the contact map for the human prion protein and lower triangle the map for the mouse prion protein. The helix A region is shown enlarged in the inset.

Conclusion

We have simulated the effect of polyadenosine RNA bound to human and mouse prions on the stability and secondary structure of these proteins. Potential binding sites were predicted with Autodock where unresolved parts of the experimental prion structures (in their Pr^C form) where assumed to take structures predicted by standard software packages ITASSER and MODELLER. The RNA-prion complexes, generated in this way, where followed in long molecular dynamics simulations. In cases were the

RNA binds with the N-terminal polybasic segment 21-31, or directly with the N-terminal helix A, we observe that binding of the RNA leads to formation of a pincer-like structure between helix A and the polybasic domain that encapsulates the RNA. Because of steric clashes, the pincer cannot be formed when the RNA binds to the polybasic region 121-131, the other predicted binding site. Formation of the pincer seems necessary to recruit and trap the RNA, and it precedes dissolution of helix A, which starts with the N-terminus of the helix and the helix subsequently unraveling toward its C-terminal end. As the molecular dynamic trajectories proceed the helical contacts are replaced by short β-strand arrangements that eventually will lead to the characteristic high β-sheet content of disease-causing mis-folded PrP^{SC} prion structure. This picture differs little between mouse and human prion protein, i.e. is independent of the sequence differences between the two proteins. Hence, our results suggest a mechanism by that RNA binding to the prion protein at the segment 21-31 can trigger the conversion of the cellular prion protein structure PrP^{C} to its infectious scrapie PrP^{SC} form that then becomes the seed for formation of toxic amyloids. The shift in structure upon interaction with RNA from helix to coil could indicate the presence of intermediate states in the process of prion conversion. Recent evidence suggests that these intermediates can affect aggregation and conversion rates due to the energy barriers between the intermediates and mis-folded proteins.[236] Furthermore, increasing the flexibility of an aggregate intermediate structure has been shown to reduce the time for conversion to a mis-folded state for amyloid β.[237] Amyloid β aggregates have also been shown to have α helical intermediate structures, which go through a similar breakdown and conversion to β sheets in the aggregation process.[238] These studies

coupled with our observations of similar behavior suggest a possible relationship between the initial stages of amyloid aggregation and prion conversion. While the importance of this mechanism for prion diseases is not clear, our results put an interesting twist on the "protein-only" hypothesis in which prions are converted solely through interaction with other prion proteins.

Acknowledgment

The simulations in this work were done sing XSEDE resources funded under project MCB160005 and on the BOOMER cluster of the University of Oklahoma. We acknowledge financial support from NSF CHE-1266256 and OCAST HR14-129.

Chapter 7 – Further Mutational Effects on Aggregation and Outlook on Computational Methods

Deletion Mutation Effects on Amyloid Structure

The Osaka mutation (ΔE22) of amyloid peptides leads to the formation of a unique "Greek key" structure for the aggregate system.[239] This structure possesses solvent permeable pores with a mean diameter of 20.8Å compared to the wild-type sequence structure's mean pore diameter of 11.1Å.[183] This allows for far more solvent molecules to be observed in the pore during computational simulations as illustrated in **Figure 25**.

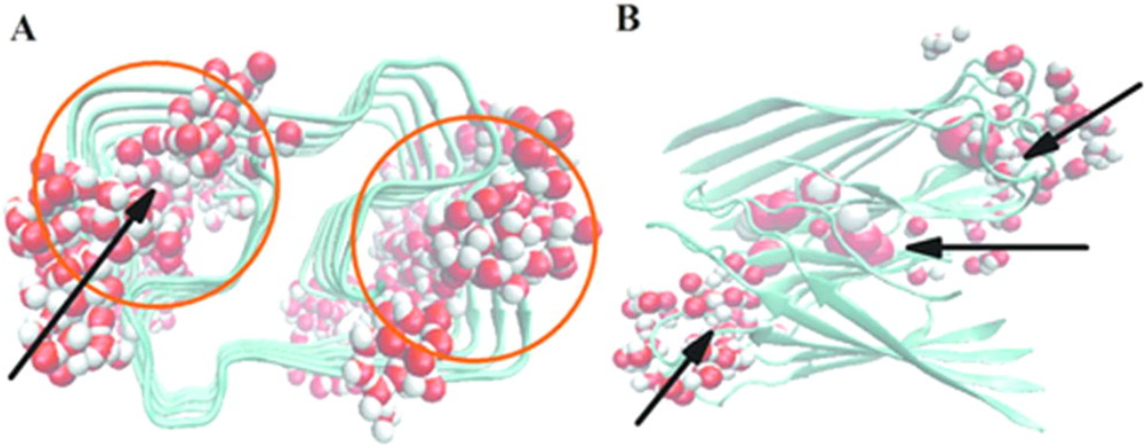

Figure 25: A comparison of solvent molecules found in solvent accessible pores between Osaka mutant **(A)** and wild-type **(B)** Aβ fibrils Reprinted with permission from Stability of Osaka Mutant and Wild-Type Fibril Models. J Phys Chem B 119:13063-70. Copyright 2015 American Chemical Society.

Using the same experimental methods discussed in **Chapter** 4, Osaka mutant structures can be shown to posses an average flow rate of 30 solvent molecules ns^{-1} as compared to 12.3 solvent molecules ns^{-1} observed in the wild-type structure.[183] This data is important when considering a proposed mechanism for amyloid toxicity of membrane leakage. In this mechanism, amyloid fibrils form ion permeable pores that allow for the unregulated flow of ions into and out of the neurons, depolarizing the membrane

potential and leading to an eventual loss of function and cellular death.[240] The elevated flow rate of solvent molecules[183] can be used to infer increased ion permeability of the Osaka mutant which can then be associated with its higher toxicity compared to the wild-type[239] to provide evidence for the membrane leakage mechanism of cytotoxicity.

However, ions traveling through the pore of amyloids is difficult to observe computationally using existing methods. Assuming a standard density for water of 1 g/mL and an extracellular ion concentration on the mM scale, derived from cerebrospinal concentrations of potassium and calcium[241], an approximate ratio of 1 ion per 5,000 to 10,000 water molecules can be derived. Using the previous mean flow rate of 30 solvent molecules ns^{-1} [183], an approximate time-step of 150 to 300ns can be calculated for each ion permeation event. This data suggests that to observe statistically relevant data on the flow of ions through the pores of an amyloid formation using physiochemical conditions requires microseconds of data to be generated. This timecale is far outside what is accessible using traditional molecular dynamics methods. Furthermore, even statistical methods may not be able to adequately map the transition of such an infrequent and quick event as they traditionally provide population data and do not relate to one specific trajectory.[103-105] This underscores the need for methods that drastically enhance the rate of sampling, while maintaining the accuracy of traditional molecular dynamics.

Mutational Effects on Prion Proteins

Further illustration of the need for novel computational methods can be shown in an expansion of the data presented for the interaction of prions with RNA snippets. The

structural dynamics of prions can be drastically altered by single point mutations, much like amyloid proteins. The D178N mutation is associated with elevated risk of formation of scrapie structures, as well as a suggested increase in the polymorphism of observed scrapie structures.[68] The M129V mutation conversely can be shown to be associated with a higher resistance to the formation of scrapie structures as well as increased polymorphism in the healthy cellular prion proteins.[69] Using the experimental setup described in Chapter 5, the following changes can be observed in the hydrogen bonding map upon introduction of the D178N and M129V mutation illustrated in **Figure 26.** The D178N mutation (**Figure 26 B)** enhances the rate of helical breakdown when compared to the wild-type **(Figure 26 A)**, with both helix 1 and 3 of the C terminal region breaking down after 300ns. Furthermore, the N terminal residues are observed forming a helical formation, which restricts the motion of polybasic domain 1 (residues 21-31) and enhances the frequency of contacts between the RNA and the polybasic domain. The M129V mutation **(Figure 26 C)** is observed with minimal helical breakdown in helices 1 and 3 of the C terminal region.

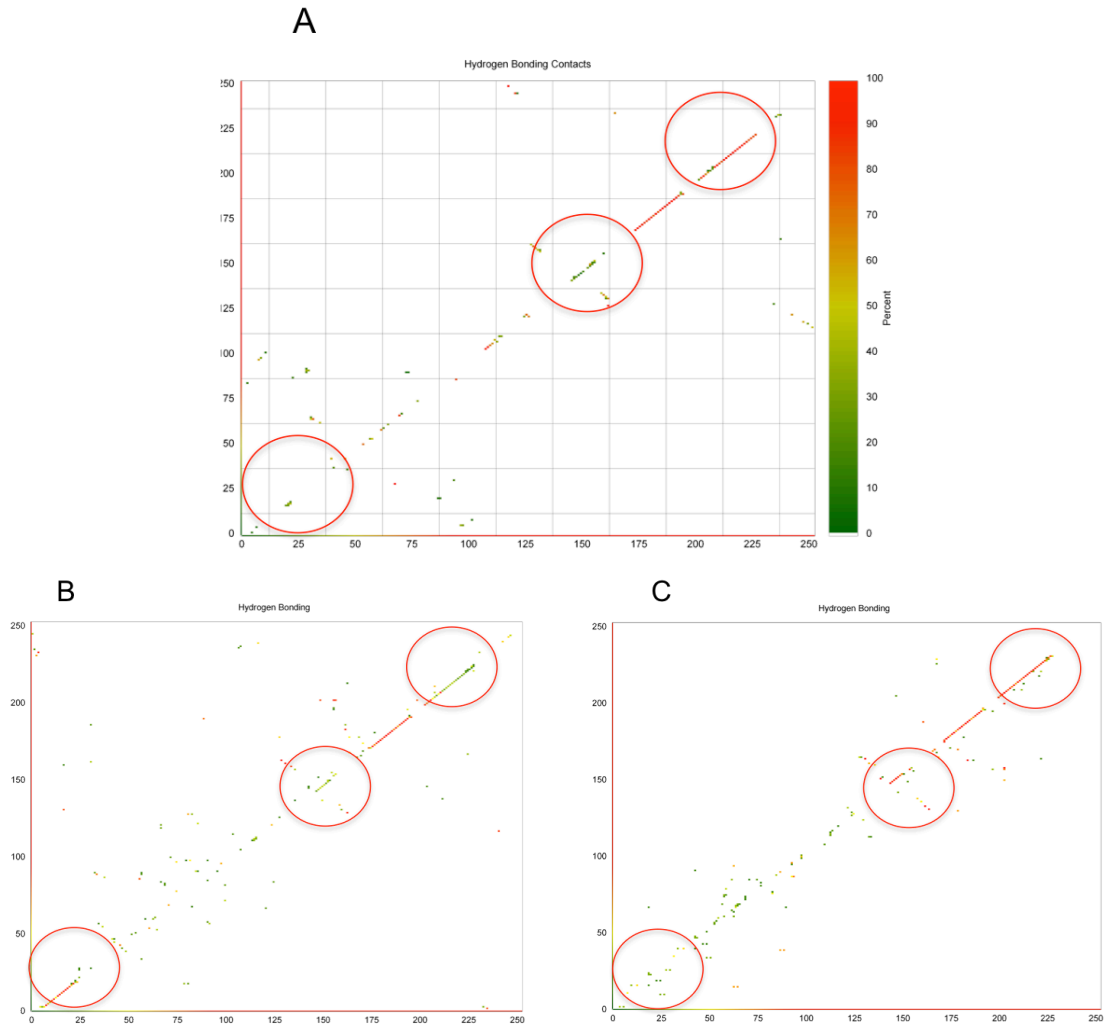

Figure 26: The mutational effects on the hydrogen bonding map of prion proteins after a 300ns computational simulation upon interaction with polyadenosine RNA. (A) indicates the original sequence, (B) indicates the D178N mutation and (C) indicates the M129V mutation. Red circles show the region of critical polybasic domain (residues 21-31) as well as helix A (residues 143 to 161) and helix C (residues 200 to 229)

The changes in the prion-RNA complex dynamics can be used to infer a potential cause for the increase in polymorphism of the scrapie form of the D178N mutant. The data suggests the elevated levels of helical breakdown would allow an increase in degrees of freedom during the formation of the misfolded structures due to loss of helical contacts

in helix C in addition to helix A **(Figure 26 B)**. However, as the formation of the β sheets rich C terminal domain of the scrapie form is not observed in any computational simulation, the pathway between cellular and scrapie prion forms is not fully explored in the trajectories. There may be unforeseen transitional structures that are not accounted for in the data. The observation of the exact dynamics of the formation of both folded and misfolded structures occurs on a timescale inaccessible to computational methods outside of employing folding biased potential models[242] and expensive specialized equipment.[70,71] This hurdle is especially cumbersome as computational costs traditionally increase with system size following an exponential function.[15] The scaling function can be converted to a more manageable power law function through the use of sophisticated statistical techniques[103-105], yet this change is often not enough to sample the full structural landscape of medium to large size proteins or protein systems. To increase the purview of the scale and types of interactions observable using computation there is a constant need for novel methods which enhance sampling while maintaining physiological relevance.

Outlook

The knowledge gained from the experiments presented in this manuscript may be used to provide a valuable roadmap for further physical experimentation into the dynamics of amyloid and amyloid-like mechanisms of formation, stabilization, and proliferation. The studies had to be restricted to comparisons of relative stability between systems and inference of possible mechanisms due to limitations in the methodology available at the time. As the field continues to evolve and develop novel methods of system sampling,

including those listed here[243], there will be a potential shift in the system sizes and types of interactions accessible using computational methods.

References

1. Leopold PE, Montal M, Onuchic JN (1992) Protein Folding Funnels: A Kinetic Approach to the Sequence-Structure Relationship. Proc Natl Acad Sci U S A 89:8721-5.
2. Onuchic JN, Wolynes PG (2004) Theory of Protein Folding. Curr Opin Struct Biol 14:70-5.
3. Alexander PA, He Y, Chen Y, Orban J, Bryan PN (2007) The Design and Characterization of Two Proteins with 88% Sequence Identity but Different Structure and Function. Proc Natl Acad Sci U S A 104:11963-8.
4. Alexander PA, He Y, Chen Y, Orban J, Bryan PN (2009) A Minimal Sequence Code for Switching Protein Structure and Function. Proc Natl Acad Sci U S A 106:21149-54.
5. He Y, Chen Y, Alexander PA, Bryan PN, Orban J (2012) Mutational Tipping Points for Switching Protein Folds and Functions. Structure 20:283-91.
6. Bryan PN, Orban J (2010) Proteins That Switch Folds. Curr Opin Struct Biol 20:482-8.
7. Tycko R (2014) Physical and Structural Basis for Polymorphism in Amyloid Fibrils. Protein Sci 23:1528-39.
8. Eisenberg D, Jucker M (2012) The Amyloid State of Proteins in Human Diseases. Cell 148:1188-1203.
9. Ahmed M, Davis J, Aucoin D, Sato T, Ahuja S, Aimoto S, Elliott JI, Van Nostrand WE, Smith SO (2010) Structural Conversion of Neurotoxic Amyloid-Beta(1-42) Oligomers to Fibrils. Nat Struct Mol Biol 17:561-7.
10. Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJ, McFarlane HT and others (2007) Atomic Structures of Amyloid Cross-Beta Spines Reveal Varied Steric Zippers. Nature 447:453-7.
11. Comellas G, Rienstra CM (2013) Protein Structure Determination by Magic-Angle Spinning Solid-State Nmr, and Insights into the Formation, Structure, and Stability of Amyloid Fibrils. Annu Rev Biophys 42:515-36.
12. Middleton CT, Marek P, Cao P, Chiu CC, Singh S, Woys AM, de Pablo JJ, Raleigh DP, Zanni MT (2012) Two-Dimensional Infrared Spectroscopy Reveals the Complex Behaviour of an Amyloid Fibril Inhibitor. Nat Chem 4:355-60.
13. Tycko R (2011) Solid-State Nmr Studies of Amyloid Fibril Structure. Annu Rev Phys Chem 62:279-99.
14. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S and others (2002) The Protein Data Bank. Acta Crystallogr D Biol Crystallogr 58:899-907.

15. Adcock SA, McCammon JA (2006) Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. Chem Rev 106:1589-615.
16. Karplus M, McCammon JA (2002) Molecular Dynamics Simulations of Biomolecules. Nat Struct Biol 9:646-52.
17. Morriss-Andrews A, Shea JE (2014) Simulations of Protein Aggregation: Insights from Atomistic and Coarse-Grained Models. Journal of Physical Chemistry Letters 5:1899-1908.
18. Jonsson SA, Mohanty S, Irback A (2012) Distinct Phases of Free Alpha-Synuclein--a Monte Carlo Study. Proteins 80:2169-77.
19. Li DW, Mohanty S, Irback A, Huo S (2008) Formation and Growth of Oligomers: A Monte Carlo Study of an Amyloid Tau Fragment. PLoS Comput Biol 4:e1000238.
20. Rizzuti B, Daggett V (2013) Using Simulations to Provide the Framework for Experimental Protein Folding Studies. Arch Biochem Biophys 531:128-35.
21. Dror RO, Dirks RM, Grossman JP, Xu H, Shaw DE (2012) Biomolecular Simulation: A Computational Microscope for Molecular Biology. Annu Rev Biophys 41:429-52.
22. Piana S, Klepeis JL, Shaw DE (2014) Assessing the Accuracy of Physical Models Used in Protein-Folding Simulations: Quantitative Evidence from Long Molecular Dynamics Simulations. Curr Opin Struct Biol 24:98-105.
23. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y and others (2010) Atomic-Level Characterization of the Structural Dynamics of Proteins. Science 330:341-6.
24. Straub JE, Thirumalai D (2011) Toward a Molecular Theory of Early and Late Events in Monomer to Amyloid Fibril Formation. Annu Rev Phys Chem 62:437-63.
25. Mousseau N, Derreumaux P (2005) Exploring the Early Steps of Amyloid Peptide Aggregation by Computers. Acc Chem Res 38:885-91.
26. Huang J, MacKerell AD, Jr. (2013) Charmm36 All-Atom Additive Protein Force Field: Validation Based on Comparison to Nmr Data. J Comput Chem 34:2135-45.
27. Lopes PE, Huang J, Shim J, Luo Y, Li H, Roux B, Mackerell AD, Jr. (2013) Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. J Chem Theory Comput 9:5430-5449.
28. Vymetal J, Vondrasek J (2013) Critical Assessment of Current Force Fields. Short Peptide Test Case. J Chem Theory Comput 9:441-51.
29. Berhanu WM, Hansmann UH (2012) Side-Chain Hydrophobicity and the Stability of Abeta(1)(6)(-)(2)(2) Aggregates. Protein Sci 21:1837-48.
30. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S and others (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. J Phys Chem B 102:3586-616.
31. Petkova AT, Leapman RD, Guo Z, Yau WM, Mattson MP, Tycko R (2005) Self-Propagating, Molecular-Level Polymorphism in Alzheimer's Beta-Amyloid Fibrils. Science 307:262-5.

32. Smaoui MR, Poitevin F, Delarue M, Koehl P, Orland H, Waldispuhl J (2013) Computational Assembly of Polymorphic Amyloid Fibrils Reveals Stable Aggregates. Biophys J 104:683-93.
33. Miller Y, Ma B, Nussinov R (2009) Polymorphism of Alzheimer's Abeta17-42 (P3) Oligomers: The Importance of the Turn Location and Its Conformation. Biophys J 97:1168-77.
34. Berhanu WM, Hansmann UHE (2012) Structure and Dynamics of Amyloid-Beta Segmental Polymorphisms. Plos One 7:12.
35. Petkova AT, Yau WM, Tycko R (2006) Experimental Constraints on Quaternary Structure in Alzheimer's Beta-Amyloid Fibrils. Biochemistry 45:498-512.
36. Tycko R, Sciarretta KL, Orgel JP, Meredith SC (2009) Evidence for Novel Beta-Sheet Structures in Iowa Mutant Beta-Amyloid Fibrils. Biochemistry 48:6072-84.
37. Murakami K, Irie K, Morimoto A, Ohigashi H, Shindo M, Nagao M, Shimizu T, Shirasawa T (2003) Neurotoxicity and Physicochemical Properties of Abeta Mutant Peptides from Cerebral Amyloid Angiopathy: Implication for the Pathogenesis of Cerebral Amyloid Angiopathy and Alzheimer's Disease. J Biol Chem 278:46179-87.
38. Fuse S, Matsumura K, Fujita Y, Sugimoto H, Takahashi T (2014) Development of Dual Targeting Inhibitors against Aggregations of Amyloid-Beta and Tau Protein. Eur J Med Chem 85:228-34.
39. Tomiyama T, Matsuyama S, Iso H, Umeda T, Takuma H, Ohnishi K, Ishibashi K, Teraoka R, Sakama N, Yamashita T and others (2010) A Mouse Model of Amyloid Beta Oligomers: Their Contribution to Synaptic Alteration, Abnormal Tau Phosphorylation, Glial Activation, and Neuronal Loss in Vivo. Journal of Neuroscience 30:4845-4856.
40. Cloe AL, Orgel J, Sachleben JR, Tycko R, Meredith SC (2011) The Japanese Mutant a Beta (Delta E22-a Beta(1-39)) Forms Fibrils Instantaneously, with Low-Thioflavin T Fluorescence: Seeding of Wild-Type a Beta(1-40) into Atypical Fibrils by Delta E22-a Beta(1-39). Biochemistry 50:2026-2039.
41. Berhanu WM, Hansmann UHE. Stability of Amyloid Oligomers. In: KarabenchevaChristova T, editor. Biomolecular Modelling and Simulations. Volume 96, Advances in Protein Chemistry and Structural Biology. San Diego: Elsevier Academic Press Inc; 2014. p 113-141.
42. Ma BY, Nussinov R (2006) Simulations as Analytical Tools to Understand Protein Aggregation and Predict Amyloid Conformation. Curr. Opin. Chem. Biol. 10:445-452.
43. Murakami K, Irie K, Morimoto A, Ohigashi H, Shindo M, Nagao M, Shimizu T, Shirasawa T (2003) Neurotoxicity and Physicochemical Properties of a Beta Mutant Peptides from Cerebral Amyloid Angiopathy - Implication for the Pathogenesis of Cerebral Amyloid Angiopathy and Alzheimer's Disease. Journal of Biological Chemistry 278:46179-46187.
44. Lu JX, Qiang W, Yau WM, Schwieters CD, Meredith SC, Tycko R (2013) Molecular Structure of Beta-Amyloid Fibrils in Alzheimer's Disease Brain Tissue. Cell 154:1257-1268.

45. Sisodia SS, Price DL (1995) Role of the Beta-Amyloid Protein in Alzheimer's Disease. FASEB J 9:366-70.
46. Winner B, Jappelli R, Maji SK, Desplats PA, Boyer L, Aigner S, Hetzer C, Loher T, Vilar M, Campioni S and others (2011) In Vivo Demonstration That Alpha-Synuclein Oligomers Are Toxic. Proc Natl Acad Sci U S A 108:4194-9.
47. Ross CA, Tabrizi SJ (2011) Huntington's Disease: From Molecular Pathogenesis to Clinical Treatment. Lancet Neurol 10:83-98.
48. Kyle RA, Gertz MA (1995) Primary Systemic Amyloidosis: Clinical and Laboratory Features in 474 Cases. Semin Hematol 32:45-59.
49. Sanchorawala V (2006) Light-Chain (Al) Amyloidosis: Diagnosis and Treatment. Clin J Am Soc Nephrol 1:1331-41.
50. Murtagh B, Hammill SC, Gertz MA, Kyle RA, Tajik AJ, Grogan M (2005) Electrocardiographic Findings in Primary Systemic Amyloidosis and Biopsy-Proven Cardiac Involvement. Am J Cardiol 95:535-7.
51. Rosenzweig M, Giralt S, Landau H (2013) Light-Chain Amyloidosis: Sct, Novel Agents and Beyond. Bone Marrow Transplant 48:1022-7.
52. Hurle MR, Helms LR, Li L, Chan W, Wetzel R (1994) A Role for Destabilizing Amino Acid Replacements in Light-Chain Amyloidosis. Proc Natl Acad Sci U S A 91:5446-50.
53. Nowak M (2004) Immunoglobulin Kappa Light Chain and Its Amyloidogenic Mutants: A Molecular Dynamics Study. Proteins 55:11-21.
54. Kobayashi Y, Tsutsumi H, Abe T, Ikeda K, Tashiro Y, Unzai S, Kamikubo H, Kataoka M, Hiroaki H, Hamada D (2014) Decreased Amyloidogenicity Caused by Mutational Modulation of Surface Properties of the Immunoglobulin Light Chain Bre Variable Domain. Biochemistry 53:5162-73.
55. Knowles TPJ, Vendruscolo M, Dobson CM (2014) The Amyloid State and Its Association with Protein Misfolding Diseases. Nature Reviews Molecular Cell Biology 15:384-396.
56. Deleault NR, Harris BT, Rees JR, Supattapone S (2007) Formation of Native Prions from Minimal Components in Vitro. Proc Natl Acad Sci U S A 104:9741-6.
57. Westergard L, Christensen HM, Harris DA (2007) The Cellular Prion Protein (Prp(C)): Its Physiological Function and Role in Disease. Biochim Biophys Acta 1772:629-44.
58. Bremer J, Baumann F, Tiberi C, Wessig C, Fischer H, Schwarz P, Steele AD, Toyka KV, Nave KA, Weis J and others (2010) Axonal Prion Protein Is Required for Peripheral Myelin Maintenance. Nat Neurosci 13:310-8.
59. Caiati MD, Safiulina VF, Fattorini G, Sivakumaran S, Legname G, Cherubini E (2013) Prpc Controls Via Protein Kinase a the Direction of Synaptic Plasticity in the Immature Hippocampus. J Neurosci 33:2973-83.
60. Pan KM, Baldwin M, Nguyen J, Gasset M, Serban A, Groth D, Mehlhorn I, Huang Z, Fletterick RJ, Cohen FE and others (1993) Conversion of Alpha-Helices into Beta-Sheets Features in the Formation of the Scrapie Prion Proteins. Proc Natl Acad Sci U S A 90:10962-6.

61. Norstrom EM, Mastrianni JA (2006) The Charge Structure of Helix 1 in the Prion Protein Regulates Conversion to Pathogenic Prpsc. J Virol 80:8521-9.
62. Singh J, Kumar H, Sabareesan AT, Udgaonkar JB (2014) Rational Stabilization of Helix 2 of the Prion Protein Prevents Its Misfolding and Oligomerization. J Am Chem Soc 136:16704-7.
63. Singh J, Udgaonkar JB (2013) Dissection of Conformational Conversion Events During Prion Amyloid Fibril Formation Using Hydrogen Exchange and Mass Spectrometry. J Mol Biol 425:3510-21.
64. Prusiner SB (1997) Prion Diseases and the Bse Crisis. Science 278:245-51.
65. Schlepckow K, Schwalbe H (2013) Molecular Mechanism of Prion Protein Oligomerization at Atomic Resolution. Angew Chem Int Ed Engl 52:10002-5.
66. Tycko R, Savtchenko R, Ostapchenko VG, Makarava N, Baskakov IV (2010) The Alpha-Helical C-Terminal Domain of Full-Length Recombinant Prp Converts to an in-Register Parallel Beta-Sheet Structure in Prp Fibrils: Evidence from Solid State Nuclear Magnetic Resonance. Biochemistry 49:9488-97.
67. Salamat K, Moudjou M, Chapuis J, Herzog L, Jaumain E, Beringue V, Rezaei H, Pastore A, Laude H, Dron M (2012) Integrity of Helix 2-Helix 3 Domain of the Prp Protein Is Not Mandatory for Prion Replication. J Biol Chem 287:18953-64.
68. Synofzik M, Bauer P, Schols L (2009) Prion Mutation D178n with Highly Variable Disease Onset and Phenotype. J Neurol Neurosurg Psychiatry 80:345-6.
69. Golanska E, Sieruta M, Corder E, Gresner SM, Pfeffer A, Chodakowska-Zebrowska M, Sobow TM, Klich I, Mossakowska M, Szybinska A and others (2013) The Prion Protein M129v Polymorphism: Longevity and Cognitive Impairment among Polish Centenarians. Prion 7:244-7.
70. Shaw DE (2009) Anton: A Specialized Machine for Millisecond-Scale Molecular Dynamics Simulations of Proteins. Abstracts of Papers of the American Chemical Society 238.
71. Grossman JP, Kuskin JS, Bank JA, Theobald M, Dror RO, Ierardi DJ, Larson RH, Ben Schafer U, Towles B, Young C and others (2013) Hardware Support for Fine-Grained Event-Driven Computation in Anton 2. Acm Sigplan Notices 48:549-560.
72. Haki GD, Rakshit SK (2003) Developments in Industrially Important Thermostable Enzymes: A Review. Bioresour Technol 89:17-34.
73. Choi JM, Han SS, Kim HS (2015) Industrial Applications of Enzyme Biocatalysis: Current Status and Future Aspects. Biotechnol Adv 33:1443-54.
74. Wilson CJ (2015) Rational Protein Design: Developing Next-Generation Biological Therapeutics and Nanobiotechnological Tools. Wiley Interdiscip Rev Nanomed Nanobiotechnol 7:330-41.
75. Pace CN, Shirley BA, McNutt M, Gajiwala K (1996) Forces Contributing to the Conformational Stability of Proteins. FASEB J 10:75-83.
76. Yokota A, Takahashi H, Takenawa T, Arai M (2010) Probing the Roles of Conserved Arginine-44 of Escherichia Coli Dihydrofolate Reductase in Its

Function and Stability by Systematic Sequence Perturbation Analysis. Biochem Biophys Res Commun 391:1703-7.
77. Torrado M, Revuelta J, Gonzalez C, Corzana F, Bastida A, Asensio JL (2009) Role of Conserved Salt Bridges in Homeodomain Stability and DNA Binding. J Biol Chem 284:23765-79.
78. Jomain JB, Tallet E, Broutin I, Hoos S, van Agthoven J, Ducruix A, Kelly PA, Kragelund BB, England P, Goffin V (2007) Structural and Thermodynamic Bases for the Design of Pure Prolactin Receptor Antagonists: X-Ray Structure of Del1-9-G129r-Hprl. J Biol Chem 282:33118-31.
79. Fredricksen RS, Swenson CA (1996) Relationship between Stability and Function for Isolated Domains of Troponin C. Biochemistry 35:14012-26.
80. Stouffer AL, Nanda V, Lear JD, DeGrado WF (2005) Sequence Determinants of a Transmembrane Proton Channel: An Inverse Relationship between Stability and Function. J Mol Biol 347:169-79.
81. Sutcliffe BT, Woolley RG (2012) On the Quantum Theory of Molecules. J Chem Phys 137:22A544.
82. Verlet L (1967) Computer Experiments on Classical Fluids .I. Thermodynamical Properties of Lennard-Jones Molecules. Physical Review 159:98-+.
83. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D and others (2013) Gromacs 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. Bioinformatics 29:845-854.
84. Hockney RW (1968) Potential Calculation. Bulletin of the American Physical Society 13:1747-&.
85. Cheatham TE (2004) Simulation and Modeling of Nucleic Acid Structure, Dynamics and Interactions. Curr Opin Struct Biol 14:360-7.
86. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S and others (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. Journal of Physical Chemistry B 102:3586-3616.
87. MacKerell AD, Feig M, Brooks III CL (2004) Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. Journal of Computational Chemistry 25:1400-1415.
88. Ahmed M, Davis J, Aucoin D, Sato T, Ahuja S, Aimoto S, Elliott JI, Van Nostrand WE, Smith SO (2010) Structural Conversion of Neurotoxic Amyloid-Beta(1-42) Oligomers to Fibrils. Nature Structural & Molecular Biology 17:561-556.
89. Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJW, McFarlane HT and others (2007) Atomic Structures of Amyloid Cross-Beta Spines Reveal Varied Steric Zippers. Nature 447:453-457.
90. Antzutkin ON, Balbach JJ, Leapman RD, Rizzo NW, Reed J, Tycko R (2000) Multiple Quantum Solid-State Nmr Indicates a Parallel, Not Antiparallel,

Organization of Beta-Sheets in Alzheimer's Beta-Amyloid Fibrils. Proceedings of the National Academy of Sciences of the United States of America 97:13045-13050.
91. Petkova AT, Leapman RD, Guo ZH, Yau WM, Mattson MP, Tycko R (2005) Self-Propagating, Molecular-Level Polymorphism in Alzheimer's Beta-Amyloid Fibrils. Science 307:262-265.
92. Fitzpatrick AWP, Debelouchina GT, Bayro MJ, Clare DK, Caporini MA, Bajaj VS, Jaroniec CP, Wang LC, Ladizhansky V, Muller SA and others (2013) Atomic Structure and Hierarchical Assembly of a Cross-Beta Amyloid Fibril. Proceedings of the National Academy of Sciences of the United States of America 110:5468-5473.
93. Smaoui MR, Poitevin F, Delarue M, Koehl P, Orland H, Waldispuhl J (2013) Computational Assembly of Polymorphic Amyloid Fibrils Reveals Stable Aggregates. Biophysical Journal 104:683-693.
94. Miller Y, Ma BY, Nussinov R (2009) Polymorphism of Alzheimer's a Beta(17-42) (P3) Oligomers: The Importance of the Turn Location and Its Conformation. Biophysical Journal 97:1168-1177.
95. Berhanu WM, Hansmann UHE (2012) Structure and Dynamics of Amyloid-Beta Segmental Polymorphisms. Plos One 7:e0041479.
96. Nilsberth C, Westlind-Danielsson A, Eckman CB, Condron MM, Axelman K, Forsell C, Stenh C, Luthman J, Teplow DB, Younkin SG and others (2001) The 'Arctic' App Mutation (E693g) Causes Alzheimer's Disease by Enhanced a Beta Protofibril Formation. Nature Neuroscience 4:887-893.
97. Tycko R, Sciarretta KL, Orgel J, Meredith SC (2009) Evidence for Novel Beta-Sheet Structures in Iowa Mutant Beta-Amyloid Fibrils. Biochemistry 48:6072-6084.
98. Qiang W, Yau WM, Luo YQ, Mattson MP, Tycko R (2012) Antiparallel Beta-Sheet Architecture in Iowa-Mutant Beta-Amyloid Fibrils. Proceedings of the National Academy of Sciences of the United States of America 109:4443-4448.
99. Buchanan LE, Carr JK, Fluitt AM, Hoganson AJ, Moran SD, Pablo JJ, Skinner JL, Zanni MT (2014) Structural Motif of Polyglutamine Amyloid Fibrils Discerned with Mixed-Isotope Infrared Spectroscopy. PNAS 111:5796-5801.
100. Tycko R (2011) Solid State Nmre Studies of Amyloid Fibril Structure. Annu Rev Phys Chem. 62:279-299.
101. Schmechel A, Zentgraf H, Scheuermann S, Fritz G, Pipkorn RD, Reed J, Beyreuther K, Bayer TA, Multhaup G (2003) Alzheimer Beta-Amyloid Homodimers Facilitate a Beta Fibrillization and the Generation of Conformational Antibodies. Journal of Biological Chemistry 278:35317-35324.
102. Piana S, Lindorff-Larsen K, Shaw DE (2012) Protein Folding Kinetics and Thermodynamics from Atomistic Simulation. Proceedings of the National Academy of Sciences of the United States of America 109:17845-17850.

103. Hansmann UHE, Okamoto Y (1993) Prediction of Peptide Conformation by Multicanonical Algorithm: A New Approach to the Multiple-Minima Problem. J. Comp. Chem. 14:1333-1338.
104. Hansmann UHE (1997) Parallel Tempering Algorithm for Conformational Studies of Biological Molecules. Chem. Phys. Lett. 281:140-150.
105. Sugita Y, Okamoto Y (1999) Replica-Exchange Molecular Dynamics Method for Protein Folding. Chem. Phys. Lett. 314:141-151.
106. Schulz JCF, Schmidt L, Best RB, Dzubiella J, Netz RR (2012) Peptide Chain Dynamics in Light and Heavy Water: Zooming in on Internal Friction. Journal of the American Chemical Society 134:6273-6279.
107. Lin IC, Tuckerman ME (2010) Enhanced Conformational Sampling of Peptides Via Reduced Side-Chain and Solvent Masses. Journal of Physical Chemistry B 114:15935-15940.
108. Wright LB, Walsh TR (2013) Efficient Conformational Sampling of Peptides Adsorbed onto Inorganic Surfaces: Insights from a Quartz Binding Peptide. Physical Chemistry Chemical Physics 15:4715-4726.
109. Huang DM, Chandler D (2000) Temperature and Length Scale Dependence of Hydrophobic Effects and Their Possible Implications for Protein Folding. Proc Natl Acad Sci U S A. 97:8324-8327.
110. Zhao J, Yu X, Liang G, Zheng J (2011) Structural Polymorphism of Human Islet Amyloid Polypeptide (Hiapp) Oligomers Highlights the Importance of Interfacial Residue Interactions. Biomacromolecules 12:210-220.
111. Foloppe N, MacKerell AD (2000) All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. Journal of Computational Chemistry 21:86-104.
112. Zachariae U, Schneider R, Briones R, Gattin Z, Demers JP, Giller K, Maier E, Zweckstetter M, Griesinger C, Becker S and others (2012) Beta-Barrel Mobility Underlies Closure of the Voltage-Dependent Anion Channel. Structure 20:1540-9.
113. Kutzner C, Grubmuller H, de Groot BL, Zachariae U (2011) Computational Electrophysiology: The Molecular Dynamics of Ion Channel Permeation and Selectivity in Atomistic Detail. Biophysical Journal 101:809-817.
114. Jang H, Connelly L, Arce FT, Ramachandran S, Kagan BL, Lal R, Nussinov R (2013) Mechanisms for the Insertion of Toxic, Fibril-Like Beta-Amyloid Oligomers into the Membrane. Journal of Chemical Theory and Computation 9:822-833.
115. Miller Y, Ma BY, Nussinov R (2011) The Unique Alzheimer's Beta-Amyloid Triangular Fibril Has a Cavity Along the Fibril Axis under Physiological Conditions. Journal of the American Chemical Society 133:2742-2748.
116. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D and others (2013) Gromacs 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. Structural bioinformatics 29 845–854.

117. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald - an N. Log(N) Method for Ewald Sums in Large Systems. Journal of Chemical Physics 98:10089-10092.
118. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A Smooth Particle Mesh Ewald Method. Journal of Chemical Physics 103:8577-8593.
119. Hess B (2008) P-Lincs: A Parallel Linear Constraint Solver for Molecular Simulation. Journal of Chemical Theory and Computation 4:116-122.
120. Miyamoto S, Kollman PA (1992) Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. Journal of Computational Chemistry 13:952-962.
121. Bussi G, Donadio D, Parrinello M (2007) Canonical Sampling through Velocity Rescaling. Journal of Chemical Physics 126:014101.
122. Bussi G, Zykova-Timan T, Parrinello M (2009) Isothermal-Isobaric Molecular Dynamics Using Stochastic Velocity Rescaling. Journal of Chemical Physics 130:074101.
123. Parrinello M, Rahman A (1981) Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. Journal of Applied Physics 52:7182-7190.
124. DeLano WL (2002) Pymol Molecular Graphics System. Version 1.3.0.4 Schrödinger, Llc. PyMOL molecular graphics system. Version 1.3.0.4 Schrödinger, LLC.
125. Bernhardt NA, Berhanu WM, Hansmann UHE (2013) Mutations and Seeding of Amylin Fibril-Like Oligomers. Journal of Physical Chemistry B 117:16076-16085.
126. Antzutkin ON, Iuga D, Filippov AV, Kelly RT, Becker-Baldus J, Brown SP, Dupree R (2012) Hydrogen Bonding in Alzheimer's Amyloid-Beta; Fibrils Probed by N-15{O-17} Reapdor Solid-State Nmr Spectroscopy. Angewandte Chemie-International Edition 51:10289-10292.
127. Berhanu WM, Masunov AE (2012) Controlling the Aggregation and Rate of Release in Order to Improve Insulin Formulation: Molecular Dynamics Study of Full-Length Insulin Amyloid Oligomer Models. Journal of Molecular Modeling 18:1129-1142.
128. Guo JL, Lee VMY (2014) Cell-to-Cell Transmission of Pathogenic Proteins in Neurodegenerative Diseases. Nature Medicine 20:130-138.
129. Miller Y, Ma BY, Nussinov R (2010) Polymorphism in Alzheimer a Beta Amyloid Organization Reflects Conformational Selection in a Rugged Energy Landscape. Chemical Reviews 110:4820-4838.
130. Kahler A, Sticht H, Horn AHC (2013) Conformational Stability of Fibrillar Amyloid-Beta Oligomers Via Protofilament Pair Formation - a Systematic Computational Study. Plos One 8:e70521.
131. Wright DW, Hall BA, Kenway OA, Jha S, Coveney PV (2014) Computing Clinically Relevant Binding Free Energies of Hiv-1 Protease Inhibitors. J. Chem. Theory Comput. 10:1228−1241.
132. Berhanu WM, Yasar F, Hansmann UHE (2013) In Silico Cross Seeding of a Beta and Amylin Fibril-Like Oligomers. ACS Chem Neurosci 4:1488-1500.

133. Berhanu WM, Hansmann UHE (2013) The Stability of Cylindrin B-Barrel Amyloid Oligomer Models—a Molecular Dynamics Study. Proteins: Structure, Function, and Bioinformatics 81:1542-1555.
134. Spiliotopoulos D, Spitaleri A, Musco G (2012) Exploring Phd Fingers and H3k4me0 Interactions with Molecular Dynamics Simulations and Binding Free Energy Calculations: Aire-Phd1, a Comparative Study. Plos One 7:e46902.
135. Miller BR, McGee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE (2012) Mmpbsa.Py: An Efficient Program for End-State Free Energy Calculations. Journal of Chemical Theory and Computation 8:3314-3321.
136. Homeyer N, Gohlke H (2012) Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. Molecular Informatics 31:114-122.
137. Park J, Kahng B, Hwang W (2009) Thermodynamic Selection of Steric Zipper Patterns in the Amyloid Cross-Beta Spine. PLoS Comput Biol 5:e10000492.
138. Laganowsky A, Liu C, Sawaya MR, Whitelegge JP, Park J, Zhao ML, Pensalfini A, Soriaga AB, Landau M, Teng PK and others (2012) Atomic View of a Toxic Amyloid Small Oligomer. Science 335:1228-1231.
139. Wereszczynski J, McCammon JA (2012) Statistical Mechanics and Molecular Dynamics in Evaluating Thermodynamic Properties of Biomolecular Recognition. Quaterly Reviews of Biophysics 45:1-25.
140. Tycko R (2014) Physical and Structural Basis for Polymorphism in Amyloid Fibrils. Protein Science.
141. Okamoto A, Yano A, Nomura K, Higai S, Kurita N (2014) Effect of D23n Mutation on the Dimer Conformation of Amyloid B-Proteins: Ab Initio Molecular Simulations in Water. J Mol Graph Model 50:113-124.
142. Jiang L, Liu C, Leibly D, Landau M, Zhao ML, Hughes MP, Eisenberg DS (2013) Structure-Based Discovery of Fiber-Binding Compounds That Reduce the Cytotoxicity of Amyloid Beta. Elife 2:e00857.
143. Rochet JC, Lansbury PT (2000) Amyloid Fibrillogenesis: Themes and Variations. Current Opinion in Structural Biology 10:60-68.
144. Berhanu WM, Hansmann UHE (2014) Stability of Amyloid Oligomers. Advances in Protein Chemistry and Structural Biology. 96:113 - 141.
145. Domert J, Rao SB, Agholme L, Brorsson AC, Marcusson J, Hallbeck M, Nath S (2014) Spreading of Amyloid-P Peptides Via Neuritic Cell-to-Cell Transfer Is Dependent on Insufficient Cellular Clearance. Neurobiology of Disease 65:82-92.
146. Jiang L, Liu C, Leibly D, Landau M, Zhao ML, Hughes MP, Eisenberg DS (2013) Structure-Based Discovery of Fiber-Binding Compounds That Reduce the Cytotoxicity of Amyloid Beta. Elife 2.
147. Berhanu WM, Yasar F, Hansmann UHE (2013) In Silico Cross Seeding of a Beta and Amylin Fibril-Like Oligomers. Acs Chemical Neuroscience 4:1488-1500.

148. Kahler A, Sticht H, Horn AHC (2013) Conformational Stability of Fibrillar Amyloid-Beta Oligomers Via Protofilament Pair Formation - a Systematic Computational Study. Plos One 8.
149. Berhanu WM, Hansmann UHE (2012) Structure and Dynamics of Amyloid-Beta Segmental Polymorphisms. Plos One 7.
150. Paravastu AK, Leapman RD, Yau WM, Tycko R (2008) Molecular Structural Basis for Polymorphism in Alzheimer's Beta-Amyloid Fibrils. Proceedings of the National Academy of Sciences of the United States of America 105:18349-18354.
151. GhattyVenkataKrishna PK, Uberbacher EC, Cheng XL (2013) Effect of the Amyloid Beta Hairpin's Structure on the Handedness of Helices Formed by Its Aggregates. Febs Letters 587:2649-2655.
152. Rubin N, Perugia E, Goldschmidt M, Fridkin M, Addadi L (2008) Chirality of Amyloid Suprastructures. Journal of the American Chemical Society 130:4602-4615.
153. Gurry T, Stultz CM (2014) Mechanism of Amyloid-Beta Fibril Elongation. Biochemistry 53:6981-6991.
154. Foloppe N MA (2000) All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. Journal of Computational Chemistry 21:86-104.
155. MacKerell AD, Bashford M, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhorn B, Relher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. . Journal of Physical Chemistry B 102:3586-3616.
156. Kutzner C GH, de Groot BL, Zachariae U (2011) The Molecular Dynamics of Ion Channel Permeation and Selectivity in Atomistic Detail. Biophysical Journal 101:809-817.
157. Zachariae U SR, Briones R, Gattin Z, Demers JP, Giller K, Maier E, Zweckstetter M, Griesinger C, Becker S, Benz R, de Groot BL, Lange A (2012) Beta-Barrel Mobility Underlies Closure of the Voltage-Dependent Anion Channel. Structure 20:1540-1549.
158. Jang H, Connelly L, Arce FT, Ramachandran S, Kagan BL, Lal R, Nussinov R. (2013) Mechanisms for the Insertion of Toxic, Fibril-Like Beta-Amyloid Oligomers into the Membrane. Journal of Chemical Theory and Computation 9:822-833.
159. Miller Y, Ma B, Nussinov R (2011) The Unique Alzheimer's Beta-Amyloid Triangular Fibril Has a Cavity Along the Fibril Axis under Physiological Conditions. Journal of the American Chemical Society 133:2742-2748.
160. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B Lindahl E (2013) Gromacs 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. Structural bioinformatics 29:845-854.

161. Darden T, Darrin Y, Pedersen L (1993) Particle Mesh Ewald - an N. Log(N) Method for Ewald Sums in Large Systems. Journal of Chemical Physics 98:10089-10092.
162. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A Smooth Particle Mesh Ewald Method. Journal of Chemical Physics 103:8577-8593.
163. Hess B (2008) P-Lincs: A Parallel Linear Constraint Solver for Molecular Simulation. Journal of Chemical Theory and Computation 4:116-122.
164. Miyamoto S, Kollman PA (1992) Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. Journal of Computational Chemistry 13:952-962.
165. Bussi G, Donadia D, Parrinello M (2007) Canonical Sampling through Velocity Rescaling. Journal of Chemical Physics 126.
166. Bussi G, Zykova-Timan T, Parrinello M (2009) Isothermal-Isobaric Molecular Dynamics Using Stochastic Velocity Rescaling. Journal of Chemical Physics 130.
167. Alred EJ, Scheele EG, Berhanu WM, Hansmann UHE (2014) Stability of Iowa Mutant and Wild Type Aβ-Peptide Aggregates. J. Chem. Phys. 141:175101.
168. Stroud JC, Liu C, Teng PK, Eisenberg D (2012) Toxic Fibrillar Oligomers of Amyloid-Beta Have Cross-Beta Structure. Proceedings of the National Academy of Sciences of the United States of America 109:7717-7722.
169. Buchete NV, Hummer G (2007) Structure and Dynamics of Parallel Beta-Sheets, Hydrophobic Core, and Loops in Alzheimer's a Beta Fibrils. Biophysical Journal 92:3032-3039.
170. Lashuel HA, Hartley D, Petre BM, Walz T, Lansbury PT (2002) Neurodegenerative Disease - Amyloid Pores from Pathogenic Mutations. Nature 418:291-291.
171. Adler J, Scheidt HA, Kruger M, Thomas L, Huster D (2014) Local Interactions Influence the Fibrillation Kinetics, Structure and Dynamics of a Beta(1-40) but Leave the General Fibril Structure Unchanged. Physical Chemistry Chemical Physics 16:7461-7471.
172. Wetzel R, Shivaprasad S, Williams AD (2007) Plasticity of Amyloid Fibrils. Biochemistry 46:1-10.
173. Ellis RJ, Minton AP (2006) Protein Aggregation in Crowded Environments. Biological Chemistry 387:485-497.
174. Gorkovskiy A, Thurber KR, Tycko R, Wickner RB (2014) Locating Folds of the in-Register Parallel B-Sheet of the Sup35p Prion Domain Infectious Amyloid. PNAS 111(43):E4615-4622.
175. Paparcone R, Sanchez J, Buehler MJ (2010) Comparative Study of Polymorphous Alzheimer's a Beta(1-40) Amyloid Nanofibrils and Microfibers. Journal of Computational and Theoretical Nanoscience 7:1279-1286.
176. Berhanu WM, Hansmann UHE (2014) Inter-Species Cross-Seeding: Stability and Assembly of Rat - Human Amylin Aggregates. Plos One 9.

177. Berhanu WM, Hansmann UHE (2013) The Stability of Cylindrin -Barrel Amyloid Oligomer Modelsa Molecular Dynamics Study. Proteins-Structure Function and Bioinformatics 81:1542-1555.
178. McDonald M, Box H, Bian W, Kendall A, Tycko R, Stubbs G (2012) Fiber Diffraction Data Indicate a Hollow Core for the Alzheimer's a Beta 3-Fold Symmetric Fibril. Journal of Molecular Biology 423:454-461.
179. Shafrir Y, Durell S, Arispe N, Guy HR (2010) Models of Membrane-Bound Alzheimer's Abeta Peptide Assemblies. Proteins-Structure Function and Bioinformatics 78:3473-3487.
180. Niu A, Zhao W, Zhang Z, Xiao F, Tang F, Yang J (2014) The Molecular Structure of Alzheimer B-Amyloid Fibrils Formed in the Presence of Phospholipid Vesicles. Angew. Chem. Int. Ed. 53:9294 -9297.
181. Berhanu WM, Hansmann UHE (2012) Structure and Dynamics of Amyloid-Beta Segmental Polymorphisms. Plos One 7:e41479.
182. Tycko R (2015) Amyloid Polymorphism: Structural Basis and Neurobiological Relevance. Neuron 86:632-645.
183. Berhanu WM, Alred EJ, Hansmann UHE (2015) Stability of Osaka Mutant and Wild-Type Fibril Models. J Phys Chem B 119:13063-70.
184. Alred E, Scheele EG, Berhanu WM, Hansmann UHE (2014) Stability of Iowa Mutant and Wild Type Aβ-Peptide Aggregates. J. Chem. Phys. 141:175101.
185. Alred E, Phillips M, Berhanu WM, Hansmann UHE (2015) On the Lack of Polymorphism in a-Peptide Aggregates Derived from Patient Brains. Protein Science 24:923-935.
186. Heise H (2008) Solid-State Nmr Spectroscopy of Amyloid Proteins. Chembiochem 9:179-89.
187. Dybowski C, Bai S (2008) Solid-State Nmr Spectroscopy. Anal Chem 80:4295-300.
188. Kajava AV, Baxa U, Steven AC (2010) Beta Arcades: Recurring Motifs in Naturally Occurring and Disease-Related Amyloid Fibrils. Faseb Journal 24:1311-1319.
189. Kajava AV, Steven AC (2006) Beta-Rolls, Beta-Helices, and Other Beta-Solenoid Proteins. Fibrous Proteins: Amyloids, Prions and Beta Proteins 73:55-96.
190. Foloppe N, MacKerell AD (2000) All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. Journal of Computational Chemistry 21:86-104.
191. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S and others (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. . Journal of Physical Chemistry B 102:3586-3616.
192. Kutzner C, Grubmuller H, de Groot BL, Zachariae U (2011) The Molecular Dynamics of Ion Channel Permeation and Selectivity in Atomistic Detail. Biophysical Journal 101:809-817.
193. Zachariae U, Schneider R, Briones R, Gattin Z, Demers JP, Giller K, Maier E, Zweckstetter M, Griesinger C, Becker S and others (2012) B-Barrel Mobility

Underlies Closure of the Voltage-Dependent Anion Channel. Structure 20:1540-1549.
194. Jang H, Connelly L, Arce FT, Ramachandran S, Kagan BL, Lal R, Nussinov R (2013) Mechanisms for the Insertion of Toxic, Fibril-Like Beta-Amyloid Oligomers into the Membrane. Journal of Chemical Theory and Computation 9:822-833.
195. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E(2013) Gromacs 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. Structural bioinformatics 29:845-854.
196. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald - an N. Log(N) Method for Ewald Sums in Large Systems. . Journal of Chemical Physics 98:10089-10092.
197. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A Smooth Particle Mesh Ewald Method. Journal of Chemical Physics 103:8577-8593.
198. Hess B (2008) P-Lincs: A Parallel Linear Constraint Solver for Molecular Simulation. Journal of Chemical Theory and Computation 4:116-122.
199. Miyamoto S, Kollman PA (1992) Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. Journal of Computational Chemistry 13:952-962.
200. Bussi G, Donadio D, Parrinello M (2007) Canonical Sampling through Velocity Rescaling. Journal of Chemical Physics 126:014101.
201. Kabsch W, Sander C (1983) Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. Biopolymers 22:2577-637.
202. Prusiner SB (1998) Prions. Proc Natl Acad Sci U S A 95:13363-83.
203. Makarava N, Savtchenko R, Alexeeva I, Rohwer RG, Baskakov IV (2016) New Molecular Insight into Mechanism of Evolution of Mammalian Synthetic Prions. Am J Pathol 186:1006-14.
204. Berhanu WM, Alred EJ, Bernhardt NA, Hansmann UHE (2015) All-Atom Simulation of Amyloid Aggregates. Physics Procdia 68C:61.
205. Chiti F, Dobson CM (2006) Protein Misfolding, Functional Amyloid, and Human Disease. Annu Rev Biochem 75:333-66.
206. Buhimschi IA, Nayeri UA, Zhao G, Shook LL, Pensalfini A, Funai EF, Bernstein IM, Glabe CG, Buhimschi CS (2014) Protein Misfolding, Congophilia, Oligomerization, and Defective Amyloid Processing in Preeclampsia. Sci Transl Med 6:245ra92.
207. Kouza M, Banerji A, Kolinski A, Buhimschi IA, Kloczkowski A (2017) Oligomerization of Fvflm Peptides and Their Ability to Inhibit Beta Amyloid Peptides Aggregation: Consideration as a Possible Model. Phys Chem Chem Phys 19:2990-2999.
208. Hwang HS, Park SH, Park YW, Kwon HS, Sohn IS (2010) Expression of Cellular Prion Protein in the Placentas of Women with Normal and Preeclamptic Pregnancies. Acta Obstet Gynecol Scand 89:1155-61.

209. Miller MB, Geoghegan JC, Supattapone S (2011) Dissociation of Infectivity from Seeding Ability in Prions with Alternate Docking Mechanism. PLoS Pathog 7:e1002128.
210. Zurawel AA, Walsh DJ, Fortier SM, Chidawanyika T, Sengupta S, Zilm K, Supattapone S (2014) Prion Nucleation Site Unmasked by Transient Interaction with Phospholipid Cofactor. Biochemistry 53:68-76.
211. Webb B, Sali A (2014) Comparative Protein Structure Modeling Using Modeller. Curr Protoc Bioinformatics 47:5 6 1-32.
212. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A (2000) Comparative Protein Structure Modeling of Genes and Genomes. Annu Rev Biophys Biomol Struct 29:291-325.
213. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-Tasser Suite: Protein Structure and Function Prediction. Nat Methods 12:7-8.
214. Roy A, Kucukural A, Zhang Y (2010) I-Tasser: A Unified Platform for Automated Protein Structure and Function Prediction. Nat Protoc 5:725-38.
215. Zhang Y (2008) I-Tasser Server for Protein 3d Structure Prediction. BMC Bioinformatics 9:40.
216. Colovos C, Yeates TO (1993) Verification of Protein Structures: Patterns of Nonbonded Atomic Interactions. Protein Sci 2:1511-9.
217. Lovell SC, Davis IW, Arendall WB, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure Validation by Calpha Geometry: Phi,Psi and Cbeta Deviation. Proteins 50:437-50.
218. Wallner B, Elofsson A (2003) Can Correct Protein Models Be Identified? Protein Sci 12:1073-86.
219. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) Autodock4 and Autodocktools4: Automated Docking with Selective Receptor Flexibility. J Comput Chem 30:2785-91.
220. Peltonen K, Colis L, Liu H, Trivedi R, Moubarek MS, Moore HM, Bai B, Rudek MA, Bieberich CJ, Laiho M (2014) A Targeting Modality for Destruction of Rna Polymerase I That Possesses Anticancer Activity. Cancer Cell 25:77-90.
221. Hosen MJ, Zubaer A, Thapa S, Khadka B, De Paepe A, Vanakker OM (2014) Molecular Docking Simulations Provide Insights in the Substrate Binding Sites and Possible Substrates of the Abcc6 Transporter. PLoS One 9:e102779.
222. Piro JR, Supattapone S (2011) Photodegradation Illuminates the Role of Polyanions in Prion Infectivity. Prion 5:49-51.
223. Piro JR, Harris BT, Supattapone S (2011) In Situ Photodegradation of Incorporated Polyanion Does Not Alter Prion Infectivity. PLoS Pathog 7:e1002001.
224. Mackerell AD, Feig M, Brooks CL (2004) Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. J Comput Chem 25:1400-15.
225. Denning EJ, Priyakumar UD, Nilsson L, Mackerell AD (2011) Impact of 2'-Hydroxyl Sampling on the Conformational Properties of Rna: Update of the Charmm All-Atom Additive Force Field for Rna. J Comput Chem 32:1929-43.

226. Best RB, Zhu X, Shim J, Lopes PE, Mittal J, Feig M, Mackerell AD (2012) Optimization of the Additive Charmm All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Phi, Psi and Side-Chain Chi(1) and Chi(2) Dihedral Angles. J Chem Theory Comput 8:3257-3273.
227. Mahoney MW, Jorgensen WL (2000) A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. The Journal of Chemical Physics 112:8910-8922.
228. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. Journal of Chemical Physics 79:926-935.
229. Kutzner C, Grubmuller H, de Groot BL, Zachariae U (2011) Computational Electrophysiology: The Molecular Dynamics of Ion Channel Permeation and Selectivity in Atomistic Detail. Biophys J 101:809-17.
230. Jang H, Connelly L, Arce FT, Ramachandran S, Kagan BL, Lal R, Nussinov R (2013) Mechanisms for the Insertion of Toxic, Fibril-Like Beta-Amyloid Oligomers into the Membrane. J Chem Theory Comput 9:822-833.
231. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald: An N·Log(N) Method for Ewald Sums in Large Systems. The Journal of Chemical Physics 98:10089-10092.
232. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A Smooth Particle Mesh Ewald Method. The Journal of Chemical Physics 103:8577-8593.
233. Bussi G, Donadio D, Parrinello M (2007) Canonical Sampling through Velocity Rescaling. Journal of Chemical Physics 126:7.
234. Parrinello M, Rahman A (1981) Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. Journal of Applied Physics 52:7182-7190.
235. Maiti R, Van Domselaar GH, Zhang H, Wishart DS (2004) Superpose: A Simple Server for Sophisticated Structural Superposition. Nucleic Acids Res 32:W590-4.
236. Buchanan LE, Dunkelberger EB, Tran HQ, Cheng PN, Chiu CC, Cao P, Raleigh DP, de Pablo JJ, Nowick JS, Zanni MT (2013) Mechanism of Iapp Amyloid Fibril Formation Involves an Intermediate with a Transient Beta-Sheet. Proc Natl Acad Sci U S A 110:19285-90.
237. Kouza M, Co NT, Nguyen PH, Kolinski A, Li MS (2015) Preformed Template Fluctuations Promote Fibril Formation: Insights from Lattice and All-Atom Models. J Chem Phys 142:145104.
238. Rojas AV, Liwo A, Scheraga HA (2011) A Study of the Alpha-Helical Intermediate Preceding the Aggregation of the Amino-Terminal Fragment of the Beta Amyloid Peptide (Abeta(1-28)). J Phys Chem B 115:12978-83.
239. Ovchinnikova OY, Finder VH, Vodopivec I, Nitsch RM, Glockshuber R (2011) The Osaka Fad Mutation E22 Delta Leads to the Formation of a Previously Unknown Type of Amyloid Beta Fibrils and Modulates a Beta Neurotoxicity. Journal of Molecular Biology 408:780-791.
240. Quist A, Doudevski L, Lin H, Azimova R, Ng D, Frangione B, Kagan B, Ghiso J, Lal R (2005) Amyloid Ion Channels: A Common Structural Link for Protein-

Misfolding Disease. Proceedings of the National Academy of Sciences of the United States of America 102:10427-10432.
241. Andrews SB, Leapman RD, Landis DM, Reese TS (1987) Distribution of Calcium and Potassium in Presynaptic Nerve Terminals from Cerebellar Cortex. Proc Natl Acad Sci U S A 84:1713-7.
242. Tozzini V (2005) Coarse-Grained Models for Proteins. Curr Opin Struct Biol 15:144-50.
243. Bernhardt NA, Xi W, Wang W, Hansmann UHE (2016) Simulating Protein Fold Switching by Replica Exchange with Tunneling. J Chem Theory Comput 12:5656-5666.