

UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

IN SILICO STUDIES OF AMYLOID FORMATION AND AMYLOID STABILITY

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the
Degree of

DOCTOR OF PHILOSOPHY

By
WENHUA WANG
Norman, Oklahoma
2020

IN SILICO STUDIES OF AMYLOID FORMATION AND AMYLOID STABILITY

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

By THE COMMITTEE CONSISTING OF

Dr. Ulrich Hansmann, Chair

Dr. Tohren Kibbey

Dr. Si Wu

Dr. Rakhi Rajan

© Copyright by WENHUA WANG 2020
All Rights Reserved.

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

Table of Contents

LIST OF TABLES	VIII
LIST OF FIGURES	VIII
ABSTRACT	X
CHAPTER 1. SCOPE AND CONTRIBUTION OF THE THESIS.....	1
1.1. A SHORT OVERVIEW	1
1.2. STABILITY STUDY OF RECENTLY FOUND AMYLOID B (AB) MOTIF	2
1.3. N-TERMINAL REGION OF SERUM AMYLOID A (SAA) AND ITS ROLE IN AMYLOID FORMATION	5
1.4. DOWNREGULATION AND AMYLOID FORMATION MECHANISM OF SAA	8
1.5. STABILITY STUDY OF RECENTLY FOUND SAA FIBRIL.....	11
1.6. SUMMARY	13
CHAPTER 2. OVERVIEW OF PROTEIN AND PROTEIN FOLDING BIOCHEMISTRY	14
2.1. AMINO ACIDS AND PROTEINS	14
2.1.1 Levels of Structures	17
2.1.2. Protein Folding and its Driving Forces.....	19
2.2 PROTEIN MISFOLDING DISEASE: AMYLOIDOSIS	21
2.2.1. Common Features of Amyloidosis	21
2.2.2. Amyloid β Amyloidosis.....	23
2.2.3. Amyloid A Amyloidosis.....	26
2.3. SIMULATION METHODS.....	29
2.3.1. Molecular Dynamics (MD).....	29
2.3.2. Force Field Potentials	32
2.3.3. Replica Exchange Molecular Dynamics (REMD).....	34
2.3.4. Summary	36
CHAPTER 3. STABILITY OF A RECENTLY FOUND TRIPLE-B-STRANDED AB ₁₋₄₂ FIBRIL MOTIF	37
3.1. INTRODUCTION.....	37
3.2. MATERIALS AND METHODS	40
3.2.1. Fibril Conformations.....	40
3.2.2. Simulation Protocol and Data Analysis	41
3.3. RESULTS AND DISCUSSION.....	43
3.3.1. Stability of A β ₁₁₋₄₂ Fibrils with the Triple- β Motif.....	43
3.3.2. Temperature Dependence of Fibril Stability.....	47

3.3.3. Role of the K28-A42 Salt-bridge.....	49
3.3.4. Packing Surfaces of A β ₁₋₄₂ Fibrils with Two-fold-packing Symmetry	52
3.4. CONCLUSIONS.....	60
3.5. ACKNOWLEDGMENTS.....	62
CHAPTER 4. STABILITY OF THE N-TERMINAL HELIX AND ITS ROLE IN AMYLOID FORMATION OF SERUM AMYLOID A	63
4.1. INTRODUCTION.....	63
4.2. MATERIALS AND METHODS	66
4.2.1. Models.....	66
4.2.2. Simulation Protocols.....	66
4.3. RESULTS AND DISCUSSION.....	67
4.3.1. System Set-up	67
4.3.2. Comparing the Efficiency of REMD and ME-REMD Sampling.....	68
4.3.3. Configurational Ensemble of SAA (1-27) Monomers.....	72
4.4. CONCLUSIONS.....	78
4.5. ACKNOWLEDGMENTS.....	80
CHAPTER 5. CLEAVAGE, DOWNREGULATION, AND AGGREGATION OF SERUM AMYLOID A.....	81
5.1. INTRODUCTION.....	81
5.2. MATERIALS AND METHODS	84
5.2.1. Initial Conformations	84
5.2.2. Simulation Protocols.....	85
5.2.3. Observables.....	86
5.3. RESULTS AND DISCUSSION.....	89
5.3.1. Hexamer	89
5.3.2. Monomer.....	94
5.4. CONCLUSION.....	108
5.5. ACKNOWLEDGMENTS.....	111
CHAPTER 6. STABILITY STUDY OF A RECENTLY FOUND HUMAN SAA FIBRIL STRUCTURE	112
6.1. INTRODUCTION.....	112
6.2. MATERIALS AND METHODS	115
6.3. RESULTS AND DISCUSSIONS.....	117
6.3.1. Critical size of SAA fibril and potential amyloid formation mechanism	117
6.3.2. Acidic conditions enhance early stage aggregation	120
6.3.3. Roles of N-terminal amyloidogenic region and C-terminal disordered region	124
6.4. CONCLUSIONS.....	128

CHAPTER 7. SUMMARY AND OUTLOOK.....	130
REFERENCES	133
APPENDIX I: CHAPTER 3 SUPPLEMENTAL.....	146
APPENDIX II: CHAPTER 4 SUPPLEMENTAL.....	148
APPENDIX III: CHAPTER 5 SUPPLEMENTAL	151
APPENDIX IV: COPYRIGHTS	156

List of Tables

TABLE 3.1.....	58
TABLE 4.1.....	72
TABLE 4.2.....	76
TABLE 5.1.....	96
TABLE 6.1.....	116
TABLE S5.1.....	151

List of Figures

FIGURE 1.1.....	2
FIGURE 1.2.....	5
FIGURE 1.3.....	8
FIGURE 1.4.....	12
FIGURE 2.1.....	15
FIGURE 2.2.....	19
FIGURE 2.3.....	21
FIGURE 2.4.....	27
FIGURE 3.1.....	44
FIGURE 3.2.....	46
FIGURE 3.3.....	48
FIGURE 3.4.....	50
FIGURE 3.5.....	52
FIGURE 3.6.....	54
FIGURE 3.7.....	55
FIGURE 3.8.....	56
FIGURE 3.9.....	59
FIGURE 4.1.....	65
FIGURE 4.2.....	71
FIGURE 4.3.....	74
FIGURE 4.4.....	75
FIGURE 4.5.....	77
FIGURE 5.1.....	83
FIGURE 5.2.....	90
FIGURE 5.3.....	92
FIGURE 5.4.....	94
FIGURE 5.5.....	98
FIGURE 5.6.....	99
FIGURE 5.7.....	101

FIGURE 5.8.....	103
FIGURE 5.9.....	104
FIGURE 5.10.....	110
FIGURE 6.1.....	118
FIGURE 6.2.....	119
FIGURE 6.3.....	120
FIGURE 6.4.....	121
FIGURE 6.5.....	123
FIGURE 6.6.....	124
FIGURE 6.7.....	125
FIGURE 6.8.....	126
FIGURE 6.10.....	128
FIGURE S3.1.....	146
FIGURE S3.2.....	147
FIGURE S4.1.....	148
FIGURE S4.2.....	149
FIGURE S4.3.....	150
FIGURE S4.4.....	150
FIGURE S5.1.....	152
FIGURE S5.2.....	153
FIGURE S5.3.....	154
FIGURE S5.4.....	155

Abstract

The term “Amyloid” describes the precursor proteins misfolded and aggregated into fibril-like structures that are built by cross β -sheet subunit. The disease caused by depositing amyloid fibril in tissues and organs is called amyloidosis. Amyloidosis can not only damage tissues and organs but also could lead to death. More than thirty amyloidoses have been observed in human body, these amyloidoses can be classified to mainly two types, local and systemic amyloidoses. One well-known local amyloidosis, amyloid β amyloidosis ($A\beta$ amyloidosis) is caused by its precursor protein amyloid β ($A\beta$) misfolding and aggregation, and $A\beta$ amyloidosis leads to Alzheimer’s Disease (AD). Patients who developed AD suffer from the neurodegenerative disorder and losing memory. On the other hand, precursor protein serum amyloid A (SAA) can cause so-called systemic amyloidosis, specifically amyloid A amyloidosis (AA amyloidosis). AA amyloidosis can cause damage in tissues and organs non-localized in the human body, such as heart, spleen, kidney, and liver. The mechanisms for amyloid fibril formation and the difference in toxicity among species for both $A\beta$ and SAA remain unclear, especially missing is the atomic detail of structural and dynamic information. In this thesis, molecular dynamics (MD) and other enhanced sampling methods are applied to probe the structural dynamics for $A\beta$ and SAA systems. Specifically, the study in **chapter 3** for $A\beta$ derives 1) the critical size, 2) the important role of last two hydrophobic residues on C-terminus to stabilize structures and 3) potential packing patterns to form two-fold structures, for the newly found S-shaped $A\beta$ fibril structure. As for SAA, our study of the N-terminal fragments revealed that the key salt-bridge interaction between residues 1R and 9E controls the misfolding and aggregation of the amyloidogenic region, the dissolving of the salt-

bridge can initialize the amyloid formation process, see details in **chapter 4**. Furthermore, our simulations of hexamer and monomer fragments and full-sized SAA protein suggested that SAA amyloid formation happens after the failure of a downregulation mechanism. The downregulation mechanism is proposed based on the following two observations. First, the difference in stability between full-sized hexamer and the hexamer built from shorter fragments, second, different structural properties between variant monomer motifs, third, different pH conditions, details can be seen in **chapter 5**. Recently, a high resolution SAA fibril structure has been resolved via Cryo-EM. In order to understand of SAA fibril formation better, we study the thermodynamic stability of the fibril via molecular dynamics simulations. Our preliminary results in **chapter 6** reveal that SAA fibril formation starts from monomers stacking into meta-stable one-fold subunit and then packing into stable fibril with two-fold symmetry. The two-fold two-layers system is the minimum size to maintain the fibril stable. The meta-stable structure can be stabilized under acidic conditions, it is consisted with our previous observation that low pH is critical to initialize SAA mis-folding and aggregation. We also discussed the roles of N-terminal amyloidogenic region and C-terminal disordered region play in SAA fibril formation. The data and results generated from our studies (chapters 3 - 6) reveal the posterities of amyloids and provide physical explanations at the atomic level, also this information of precursor protein downregulation, amyloid formation, and toxicity can be generalized to understand different type of amyloidoses and provide insights into future studies.

Chapter 1. Scope and Contribution of the Thesis

1.1. A Short Overview

Amyloidosis is a disease caused by the buildup and deposition of abnormal amyloid fibrils in human tissues which can further cause tissue damage, dysfunction of organs, and eventually death.¹ Amyloid fibrils formation follows a sigmoidal growth kinetic pattern, a pattern that is typically observed for nucleated self-assembly reactions. Figure 1.1 represents the stages of amyloid formation. Three significant phases are observed: the lag phase, the growth/transition phase, and the saturation/steady phase.^{2, 3} The lag phase is the period during which the amyloid becomes nucleated, the saturation phase represents the concentration required to reach equilibrium, the transition/growth phase describes the highest conversion rate from monomers or oligomers to the fibril state. Understanding the dynamics of each stage and the transition between different phases can primarily benefit human health by offering guidance in developing a new drug or treatment against amyloidosis. Our work in this thesis provides the protein structural dynamic data which reveals the molecular underpinning of different phases (steady phase of A β , lag phase and steady phase of SAA) in the amyloid formation process by taking advantage of molecular dynamics (MD) and other advanced computational approaches. Specifically, chapter 2 provides backgrounds of protein folding and related simulation methods. Work in chapter 3 reveals the critical size and potential packing of recently found amyloid β (A β) fibril structure. Work in chapter 4-6 answers what role does N-terminal region, cleavage or acidic condition play in down-regulation and amyloid formation of SAA.

In summary, the scope of this thesis covered our work in studying the post / steady phase of amyloid fibril formation of A β and the lag and steady phase for SAA. Many previous studies focused on the elongation/transition phase for A β , however the elongation phase for SAA was rarely reported. Studying the transition between native structure and fibril is critical to understand the mechanism of SAA fibril formation and provide the potential guidance for drug design and treatment.

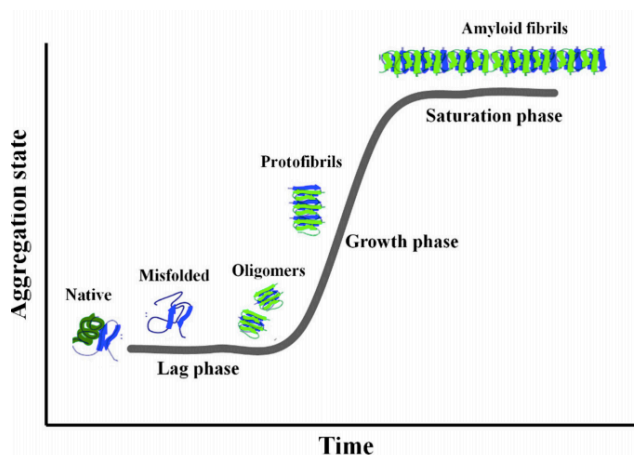


Figure 1.1. Aggregation states over time ⁴ (Iannuzzi, C.; Irace, G.; Sirangelo, I., *Molecules* **2015**, *20* (2), 2510-2528. Open access journal and no copyright is required)

1.2. Stability Study of Recently Found Amyloid β (A β) motif

Amyloidosis starts with the precursor protein misfolding and aggregating. For example, Amyloid β (A β) is the precursor protein that triggers A β amyloidosis and leads to Alzheimer's Disease (AD).⁵⁻⁷ The deposition of A β fibril in the human brain can cause senile dementia which is a result of the brain degrading and is commonly seen in AD patients.⁷ A β fibril structures have been

reported as fibril structures that are assembled from peptides of lengths from 37 to 42 residues. Two most common fragments are A β ₁₋₄₀ and A β ₁₋₄₂, and the latter one reported to have greater toxicity. Also, compared to the A β mature fibril, the soluble oligomers are more toxic to cell.⁷⁻⁹ The in vitro observations of oligomers assembled by peptides with the same sequence are not necessarily to end up with the same structure; the polymorphic property of A β oligomers and fibrils could lead to different pathologies.^{10, 11} Several oligomer structures have been reported before. There are two significant types of fibrils known today, U-shaped and S-shaped fibril structures. More than five variations of U-shaped A β ₁₋₄₀ structures have been reported. The A β ₁₋₄₂ U-shaped structure contains two β -strand. They cross regions of residues 18-26 and 31-42 and a loop region connects two β -strand (see figure 1.2a).¹¹⁻¹³

Until recently only a U-shaped low-resolution structure, discovered in 2005, was available for A β ₁₋₄₂. In 2015 a high-resolution A β ₁₋₄₂ was reported by Xiao et al. and produced by using the ssNMR approach.¹⁴ Interestingly, this structure was different than the U-shaped structure mentioned before; this new model adopted an S-shaped structure. Each chain of the model is built by three-stranded β -sheets, β 1 (residues 12-18), β 2 (residues 24-33), and β 3 (residues 36-40). Later, more similar S-shaped structures were also reported (see figure 1.2b). The unique three-stranded β -sheet structure drew the interest of researchers immediately. Our previous work over molecular dynamics (MD) simulation, covered in chapter 3, was applied to study the stability of A β ₁₋₄₂ S-shaped structure motif. This study concluded that the critical size (6 chains) is the driving force to maintain the motif's stability. Specifically the hydrophobic interaction between β 2 and β 3 is the driving force keeping the structure stable. Truncation of hydrophobic residues on C-terminal (I41 and A42) can significantly disrupt the conformation due to the damage to β 2-turn- β 3 hydrophobic core. Based

on this observation, residues 41-42 are essential to structural stability. This observation could explain that why S-shaped $A\beta_{1-42}$ cannot cross-seed the $A\beta_{1-40}$, as seen in the previous study, is due to the fact that $A\beta_{1-40}$ does not contain the last two residues. Lacking the last two residues leads to a high energy barrier for cross-seeding. This also explains why the observation of S-shaped structure came late.

Furthermore, two different 2-fold packing symmetries are proposed. The first, PSA, is packing between $\beta 1$ sheet from each fold. Residue K16 on one fold stabilizes the structure by forming salt-bridge with residue D22/E23 on another fold. The second, PSB, is packing between the $\beta 3$ sheet from each fold. Contact between V40 and G38 stabilized the 2-fold packing symmetry.¹⁵ More specific details can be found in chapter 3 of this thesis. The work described above not only gives a better understanding of the novel S-shaped $A\beta_{1-42}$ but is also crucial for later work. The stability study introduces the idea that there is a difference in neurotoxicity between the U-shaped and S-shaped structures, bringing up the idea of “in-register” and “out-of-register” fibril. These two types of fibril differ by the hydrogen bond interaction between different chains, which is important for the structure transferring to a β barrel structure. The β barrel is responsible for damaging cell membranes and generates neurotoxicity. S-shaped $A\beta_{1-42}$ can go through a structural rearrangement from “in-register” to “out-of-register” fibril and eventually transfer into the more toxic β barrel confirmation. The U-shaped structure lacks this ability.¹⁶⁻¹⁸ The proposed packing symmetry PSA is also further extended from two folds to N folds to propose the ring-like models with a larger pore in the center and offered a better understanding of oligomerization and toxicity.

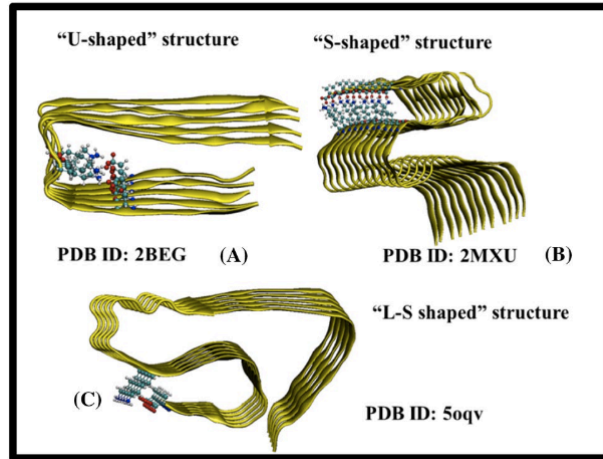


Figure 1.2. U-shaped, S-shaped, and L-S shaped amyloid β .¹⁹ (Agrawal, N.; Skelton, A. A., *The protein journal* **2019**, 38 (4), 425-434. Copyright 2019 by The Protein Journal, see details in Appendix IV)

1.3. N-terminal Region of Serum Amyloid A (SAA) and its role in Amyloid Formation

A β amyloidosis is only one kind of amyloidosis. Thirty-seven peptides or proteins have been reported that can cause different forms of amyloidosis. Seven of these proteins are related to amyloidosis in the central nervous system, and these amyloidosis can cause neurodegeneration, and lead to diseases, such as Alzheimer's diseases and Parkinson's diseases. Fifteen of the remaining thirty proteins are highly related to the amyloidosis that occur in tissues which include, but are not limited to, the liver, spleen, kidney, and heart. Amyloidosis in tissues is classified as systemic amyloidosis.¹ Recently, research is mainly focused on neurodegenerative amyloids. The number of patients with neurodegeneration disorder is higher than systemic amyloidosis; however, many more cases and various tissues are involved with the latter group. ¹ Understanding the pathogenesis of systemic amyloidosis is becoming more and more critical. Although many other

forms of systemic amyloidosis are mentioned in this thesis, the focus of the thesis is AA amyloidosis which is a systemic amyloidosis triggered by the precursor protein serum amyloid A (SAA).

At room temperature SAA adopts a disordered structure. As the temperature cools to 4 °C, and salt begins precipitating, the ratio of α -helical increase significantly.²⁰⁻²² All-atom structures of human SAA1 has been resolved by X-ray experiment. The monomer is made up of four helices together as a bundle and wrapped by a long C-terminal tail. Helix 1 to 4 and the C-terminal correspond to residues 1-27, 32-47, 50-69, 73-88, and 89-104 respectively (see fig 1.3a).²² Before the SAA monomer structure was resolved the N-terminus of SAA was previously reported as a critical region for misfolding and aggregation because the losing of the first 11 residues prevents misfolding and aggregation from happening.²³ Nordling et al built a model of N-terminal helix with 13 residues. The structure switching between α -helix and β -hairpin was observed in Nordling's simulation, which consisted of previous Raman spectra measurement for SAA residues 1-12 fragment aggregation.^{24,25} Later, more simulations were conducted showing that single-point mutations on the Nordling's model can cause the population shift between the helix and hairpin conformations.²⁶ However, in the recently resolved structure, the N-terminal region is a long straight helix (residues 1-27). Compared with previous studies that all focused on the first 11-13 residues, the neighbor helices can affect the molecular underpinning of helix-hairpin transition due to the local constraint difference.

To address what role the N-terminal helix plays in SAA amyloid formation, **our work in chapter 4**, a stability study of SAA N-terminal helix I (residue 1-27) was conducted using replica exchange molecular dynamics (REMD) and in-house multi-exchange replica exchange molecular dynamics

(ME-REMD). REMD is a widely used enhanced sampling algorithm that can solve the problem of slow sampling caused by the significant barrier in the energy landscape compared to MD. REMD alleviates slow sampling by exchanging replicas at different temperatures. High temperatures can easily overcome the energy barrier. In contrast, the low temperature can explore low energy conformations. REMD works well for relatively small systems; it is cumbersome for the proteins in explicit water. ME-REMD is an in-house modified approach that can improve the acceptance vanishing limitation of REMD caused by the increasing system size. If the number of replicas stays the same, the exchange rate will significantly decrease. REMD and ME-REMD will be discussed in detail in **chapter 2**. Comparison of simulation data from REMD and ME-REMD reveals the physical properties of the system, such as secondary structure probability, the outcome of ME-REMD consists of REMD. ME-REMD can maintain the same exchange rate between replicas by reducing one-fourth of the replicas comparing in REMD, due to this property of ME-REMD, it has the potential to be applied on more extensive systems to save computational resources. Using ME-REMD allowed us to explore three motifs of SAA N-terminal helix: 1) the straight helix, 2) the helix hairpin, and 3) the N-terminus dynamic. The straight helix maintains a similar helix structure as crystal structure. The helix hairpin more than 50 % of the straight helix break in the region of residues 11-13, and the two sub helices form a hairpin shaped structure that is maintained through the hydrophobic interaction between two helices. The N-terminus dynamic is the first helix of the helix hairpin motif to loses its structure. This is due to disruption of the salt-bridge between R1 and E9 which normally maintains the first helix in a stable state. The N-terminus dynamic motif could be the starting point for the first 11 residues to misfold into a β -hairpin and the potentially forming aggregates. Mutation E9A can disrupt the salt-bridge between R1 and E9 causing the equilibrium to be shifted to a motif with a higher β -strand content. This

observation is consistent with previous reported SAA shorter fragment simulation studies.²⁶ Although the mutation E9A is not observed in nature, the mutation changed the isoelectric point of the SAA fragment and mimicked an acidic condition. The increasing probability of misfolding and aggregating is consistent with experimental observations under low pH conditions. Our work on the N-terminal SAA resulted in the proposed mechanism that the key residues maintain the stability of the critical region and prevent it from misfolding and aggregating.

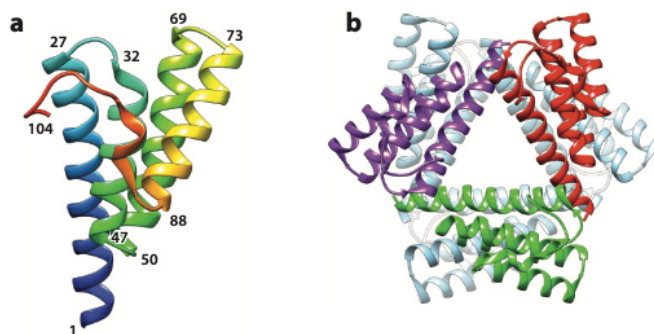


Figure 1.3. a). Monomer, and b). hexamer of serum amyloid A.¹ (Westermarck, G. T.; Fändrich, M.; Westermarck, P., *Annual Review of Pathology: Mechanisms of Disease* **2015**, *10*, 321-344. Copyright 2015 by Annual Review of Pathology, see details in Appendix IV)

1.4. Downregulation and Amyloid Formation Mechanism of SAA

While misfolding and aggregation are a trigger, they do not absolutely lead to the diseased condition. The human body has protective mechanisms against the misfolding and aggregation of proteins, when these mechanisms fail the diseased conditions happen.²⁷ It is not only essential to understand the misfolding and aggregation process, but also the protective mechanisms; however, for amyloidoses, these mechanisms remain mostly unclear. Although the whole picture of the

mechanisms is not clear, the factors that are known to play essential roles during the process, are precursor protein concentration, genetic mutation, cleavage, binding to other molecules, and seeding. As mentioned before, more than 30 proteins have been found related to amyloidosis, and while the amyloid formation mechanisms can vary for different proteins, several features are common among all proteins: the nucleus formation and fibril elongation fibril.²⁸ On the other hand, the protection mechanisms might also include standard components among different proteins; understanding these components can offer more precise treatments. Individual proteins need to be researched to recognize these crucial features and eventually the profiles can help us find the pattern. In this thesis, SAA monomer and hexamer fragments are studied to answer this question.

Previous studies have shown that SAA oligomeric states are not limited to one structure; hexamers, octamers, and monomers have been identified in murine SAA2.2. A recent study of the SAA atomic structure has also established the full length of human SAA1.1. It was found to be a 104 residues monomer that is assembled as a hexamer and functions as the biologically active state (see fig 1.3b).²² The biological function of the SAA is rarely understood. Still, some important functions have been identified before. For example, under certain health condition, SAA will bind to high density lipoprotein (HDL) and transport cholesterol away from the injured site for the human body to recycle later.¹ Under the condition that the human body has an infection, such as rheumatoid arthritis (RA), the concentration of SAA can increase to 1 mg/mL and reach to 1000 fold higher than would otherwise be if no infection were present. With the help of glycosaminoglycans (GAGs), such as heparin/HS, SAA will no longer be bound to HDL, from which point the misfolding and aggregation can be triggered.

The mechanism of SAA fibril formation was proposed to happen in the following way. SAA first dissociates from HDL, enters the cell, and resides in the acidic lysosomal environment. Structural rearrangement of SAA happens in the lysosomal environment. During the rearrangement oligomers forms in the cell, and then the oligomer breaks the lysosomal and cell membranes reaching the extracellular environment. Finally, SAA oligomers extend to mature amyloid fibrils.

²⁹ The concentration of SAA increases up to 1000 folds during acute inflammation. However, for different individuals, the duration and exact levels of SSA concentrations vary. The difference in concentrations among patients may explain why not all patients with acute inflammation suffer from SAA amyloidosis. However, more importantly, the protein quality control system plays an critical role in the degradation of SAA and its fibril deposits. The degradation process is an efficient way to release the pressure of SAA overexpression. It can downregulate the concentration of SAA via cleavage. Failure of degradation/ downregulation can lead to the diseased condition. The mechanism of SAA degradation in fibril formation is not precise.

Our work in chapter 5 proposes a down-regulation mechanism for SAA under disease condition. Down-regulation failure causes SAA fibrils formation. The down-regulation mechanism is a two step-process. First, under the high concentration diseased condition, SAA activity has been reduced via the cleavage, which can shift the equilibrium from full-length hexamer to shorter monomer fragments. Secondly, the shorter fragment SAA₁₋₇₆ can switch between two structural forms. The first form (helix-weaken) is easy to proteolyze, however, it could also trigger aggregation and the second form (helix-broken) is hard to proteolyze but prevents aggregation. These two forms are highly related to the environmental conditions such as pH. Under natural conditions the helix weaken-structure dominates and can lead to further cleavage of the N-terminal

region and fast degradation. Inversely, acidic conditions can increase the chance of amyloid formation. The population shifts to helix-broken structure to protect the N-terminal region from aggregation. This is accomplished primarily by the salt-bridge interaction between residues R1 and E9 being disrupted at acidic condition, which makes N-terminal region shift to a disordered state. However, the failure of this switching mechanism could lead to aggregation and disease conditions. Our work revealed SAA downregulation and aggregation mechanism. This mechanism can be generated as a model to study other amyloidosis.

1.5. Stability Study of Recently Found SAA fibril

To better understand the SAA fibril formation, mechanism we extended our work to study the stability of recently reported all-atomic SAA fibril structures. The structure is built by a short fragment (residues 2-69). The region that contains residues 2-55 have been identified as ordered cross- β structure, and the region containing residues 56-69 is classified as a disordered structure (see figure 1.4).³⁰ The fibril adopts a two-fold packing symmetry.

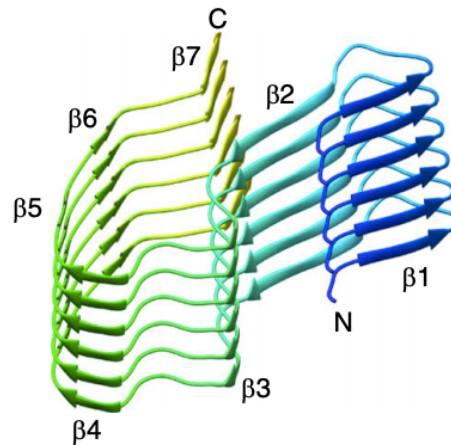


Figure 1.4. Fibril structure of SAA residues 2-55.³⁰ (Liberta, F.; Loerch, S.; Rennegarbe, M.; Schierhorn, A.; Westermark, P.; Westermark, G. T.; Hazenberg, B. P.; Grigorieff, N.; Fändrich, M.; Schmidt, M., *Nature communications* **2019**, *10* (1), 1-10. Open access journal and no copyright is required)

Understanding the fibril formation is difficult regardless of whether experimental or computational approaches are used. **Our preliminary data in chapter 6** provides a potential mechanism of SAA fibril formation through studying the thermodynamic stability of different sizes of fibril systems. The subunit of fibrils with variant stability could reveal interactions that are critical for fibril formation. First, we identified that the crucial dimension of SAA fibril is a two-fold two layers system. The stacking between monomers happens first to form a meta-stable one-fold two layer structure which is stabilized via packing into a two-fold system. The interactions on the N-terminal region maintain the two-fold symmetry. Second, acidic pH does not affect the stability of the two-fold systems but low pH conditions can stabilize the one-fold two-layer meta-stable structure via altering the sidechain interaction pattern in the C-terminal cavity (forming sidechain hydrogen bonds). This observation indicates that acidic conditions are critical for the early stage of SAA fibril formation. Third, we also discussed that the N-terminal amyloidogenic region (residues 2-

11) is essential to maintain the two-fold packing and stabilize SAA fibril. The C-terminal disordered region (residues 56-76) does not reduce the stability of the two-fold system. Our work proposes a potential mechanism for SAA fibril formation. Monomers stack to form meta-stable structures.

Under acidic conditions, the meta-stable states are stabilized. Then, the meta-stable structures pack via their N-terminal region to develop a two-fold symmetry. The C-terminal disordered regions did not interrupt this process. This potential mechanism can offer atomic details and insight into early-stage SAA misfolding and aggregation.

1.6. Summary

In summary, the scope of this thesis covered our work in studying the fibril formation of A β and SAA. Many previous studies focused on the elongation/transition phase for A β , however the elongation phase for SAA was rarely reported. Understanding the transition between native structure and fibril is critical for revealing mechanism of SAA amyloid fibril formation and provides potential guidance for drug design and treatment. The work can be continued by applying state-of-art computational approaches to explore the landscape and transition pathway of SAA fibril formation, this is discussed in detail in **chapter 7 summary and outlook**.

Chapter 2. Overview of Protein and Protein Folding Biochemistry

2.1. Amino Acids and Proteins

The amino acid is the basic unit of proteins. Around 500 different amino acids have been identified in nature by humans however, only 22 amino acids are related to protein formation. Of these 20 amino acids are classified as proteinogenic amino acids.³¹⁻³³ Amino acid has a carboxyl and an amine group, along with a unique side chain. Based on the side-chains properties, proteinogenic amino acids can be classified as charged (R, K, H, D, E), polar (S, T, N, Q, C, G, P) or non-polar (A, I, L, M, F, W, V, Y). The remained 2 amino acids are O and U. More details about amino acids can be seen in figure 2.1. Combinations of these 22 kinds form the sequence of peptide chains and assemble into to a different proteins. Proteinogenic amino acids are all "left-handed" isomers or L-amino acids (L-stereoisomers), and a few D-amino acids are identified in bacteria cell walls but not in bacteria protein and antibiotics.³⁴

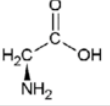
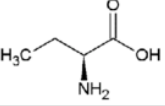
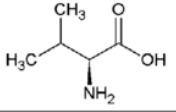
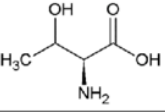
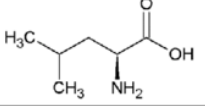
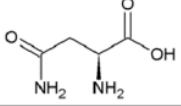
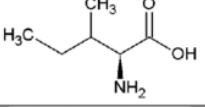
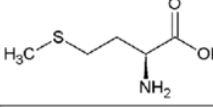
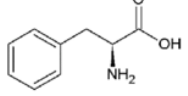
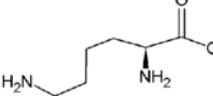
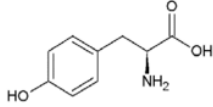
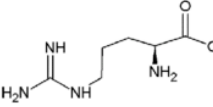
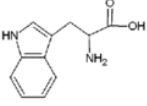
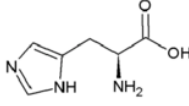
Amino Acid	Side chain properties	Chemical structure	Amino Acid	Side chain properties	Chemical structure
Alanine	Hydrophobic side chain		Serine	Polar, uncharged side chain	
Valine	Hydrophobic side chain		Threonine	Polar, uncharged side chain	
Leucine	Hydrophobic side chain		Asparagine	Polar, uncharged side chain	
Isoleucine	Hydrophobic side chain		Methionine	Polar, uncharged side chain	
Phenylalanine	Hydrophobic side chain		Lysine	Positively charged side chain	
Tyrosine	Hydrophobic side chain		Arginine	Positively charged side chain	
Tryptophan	Hydrophobic side chain		Histidine	Positively charged side chain	

Figure 2.1. Examples of proteinogenic amino acids under physiological pH³⁵ (Aerts, J. W., Röling, W. F., Elsaesser, A., & Ehrenfreund, P. (2014) *Life*, 4(4), 535-565. Open access journal and no copyright is required)

The biochemical properties of different amino acids depend on their different side-chains. Side-chains protonation states are highly related to the pH of the environment. Isoelectric point (pI) is a pH value that maintains the average net charge of zero. When environmental pH is lower than the pI, the charged amino acids side-chains are protonated, on the contrary, a higher value of pH can deprotonate them. The change of side chains protonation states can affect the biological functions of amino acid; for example, the protonated / deprotonated of side chains can disrupt the electrostatic interaction between amino acids. ³⁶

The initiation of protein formation begins with the peptide bond formation. The peptide bond formation is a type of condensation reaction. During the condensation, the α -carbon atom of one amino acid moves close to the α -nitrogen atom of the other. Two amino acids lose one oxygen and two hydrogen atoms. A water molecule is produced during this reaction and the peptide bond connects the two amino acids.³⁷ To break a peptide bond, it requires overcoming a high energy barrier and releases energy of ~ 2 -4 kcal/mol. This energetically unfavorable process takes an exceptionally long time; at room temperature the half-life of the peptide bond is 350-600 years.³⁸ This process can be accelerated by the participation of enzymes in living organisms. Peptide bond hydrolysis with enzymes is known as protease.³⁹

The peptide bond is a double covalent bond. The amide group forms a planar bond and exists as cis or trans-isomers. The two isomers can both exist in the unfolded protein structure; however, in the folded structures only one isomer can be adopted in each position. Compared to cis the trans isomers are adopted in most of the cases (99.9 %). One exception is Proline. In the X-Pro peptide bond (X means any amino acid in front of Proline) the unique side chain of Proline causes the energy difference between cis and trans isomer to be much less causing the ratio of cis over trans isomer to increase to 3-10 %.

2.1.1 Levels of Structures

Proteins are polypeptides and chains under 30 amino acids are considered to be peptides rather than proteins by convention.⁴⁰ Proteins fold into specific 3D structures via driving forces including hydrogen bonding, hydrophobic interaction, Van der Waals and electrostatic. These 3D structures allow the proteins to perform their biological functions. The adopted structures have a tight relationship with their function. Many experimental techniques, X-ray crystallography, cryogenic electron microscopy, and solid-state NMR (ssNMR) spectroscopy are employed to characterize the structures of proteins. The size of proteins range from about ten to thousands amino acids with the physical size being between 1 to 100 nm. Massive aggregates can be formed by subunits, such as oligomers.⁴¹

Protein structure has four distinct levels (primary, secondary, tertiary, and quaternary). Amino acids sequence determine the primary structure.⁴² In between N- and C- terminus are residues connected by a peptide bond and this connection forms the backbone of proteins. The sequence of primary protein structures can be determined directly or from DNA sequences. Their sequences can often determine the primary structures of the protein. However, variant modifications can be involved which alter the structures such as disulfide bond cross-linking, transition between L- and D- amino acid, phosphorylation, and cleavage.

The amino acid sequence is highly related to protein 3D structure. The local features of the 3D structure, such as the secondary structure segmentation, can be predicted from the sequence (primary structure). 3D structures of proteins cannot be determined only by its sequence. However, the structural information of a homologous protein increases the accuracy of prediction through homology modeling.⁴³

The secondary structure of the protein is 3D local segments of protein.^{44,45} Secondary structure is strictly defined by the backbone hydrogen bond pattern or by using the backbone dihedral angles pattern. A Ramachandran plot is a method to classify the secondary structure by using backbone dihedral angles directly.⁴⁶ α -helices and β -sheets are most commonly observed secondary structures. In an α -helix structure, backbone atoms spiral around a virtual axis. There are average 3.6 residues in each helical turn, and a hydrogen bond ring contains 13 atoms. Compared to other secondary structure motifs, the α -helix is most prevalent and most natural to predict from the sequence.⁴⁷ Other standard helix structures, such as π -helix, is less observed due to the unfavorable packing of the backbone atoms. The backbone of β -strands adopted a zig-zag pattern, differing from the α -helix. Adjacent strands can be formed by the sequences that are far away from each other. β -sheets can be classified as parallel/antiparallel by the direction of adjacent strands (see fig 2.2). β -sheets can lead to amyloid formation in human diseases. Turns and loops link other secondary structures together. The coil is a disordered structure that represents an undefined secondary structure.

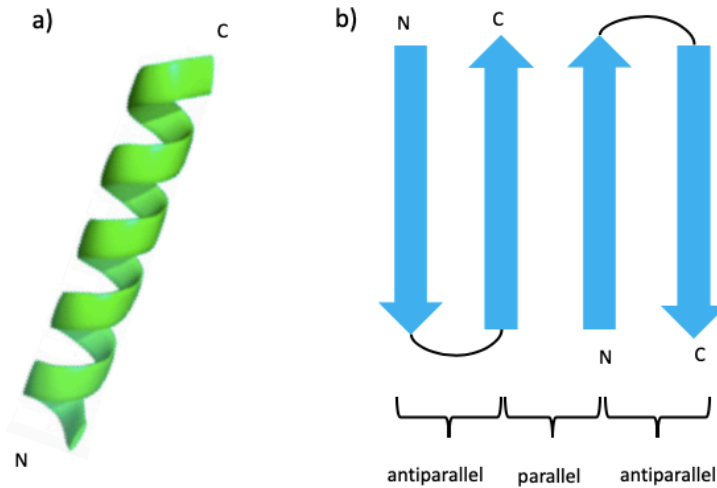


Figure 2.2. Secondary structure motifs. a). α -helix. b). β -sheets. N and C stand for terminus.

The tertiary structure describes the protein 3D structural information.⁴⁸ The backbone of the tertiary structure connects one or more secondary structure motifs while the side chains of protein interact with other side chains. Hydrophobic interactions drive folding of the proteins. Hydrophobic residues are buried inside the globular structure, but further stabilization of tertiary structure may require stronger interaction, such as salt-bridge interactions to lock up the structure. Quaternary structure is a 3D structure assembled by more than two polypeptide chains; the mechanism of quaternary structure stabilization is similar to tertiary structure.

2.1.2. Protein Folding and its Driving Forces

From the primary structure to 3D structure, proteins go through a folding process that happens spontaneously and is driven by van der Waals interactions, hydrophobic interactions, and hydrogen

bond contacts.⁴⁹ The local environment, such as pH conditions, salt concentration, and temperature effects can profoundly affect protein folding.

The fundamental driving force behind protein folding is hydrophobic interactions among side chains. Minimizing the side chains exposition to water can effectively increase hydrophobic interactions. The hydrophobic regions of protein collapse to form the hydrophobic core during folding process. When the protein surrounds by water, water molecules aggregate outside of the hydrophobic core and form a water shell which decreases the entropy. This makes the folding process favorable to entropy. Inside the hydrophobic core there are van der Waals interactions among these hydrophobic groups and hydrogen bonds which significantly increase the stability of protein after folding.^{50, 51} The distribution of hydrophobic residues can determine if a primary sequence can fold into a protein with globular structure, random distribution, or cluster of hydrophobic residues because they can disrupt the folding process causing the protein to adopt an intrinsically disordered structure.⁵²⁻⁵⁴

Protein function is determined by its structure and structure highly depends on folding. The theoretical understanding of protein folding remains incomplete. An energy landscape can describe the conformational space of protein folding. Bryngelson and Wolynes proposed that due to the principle of minimal frustration proteins, nature already has optimized energy landscapes for folding.⁵⁵ The natural evolution determines the specific primary sequence of the protein making the folding process faster and native structure is sufficiently stable. Although the frustration has been minimized for proteins, some left-over cases can still be observed as local minima. The energy landscape adopts a "funneled" shape, this funneled landscape allows folding to occur

through more than one pathway towards the native structure at the bottom (see fig 2.3).⁵⁶ The model has been proved very powerful to explain the experimental and simulation data.

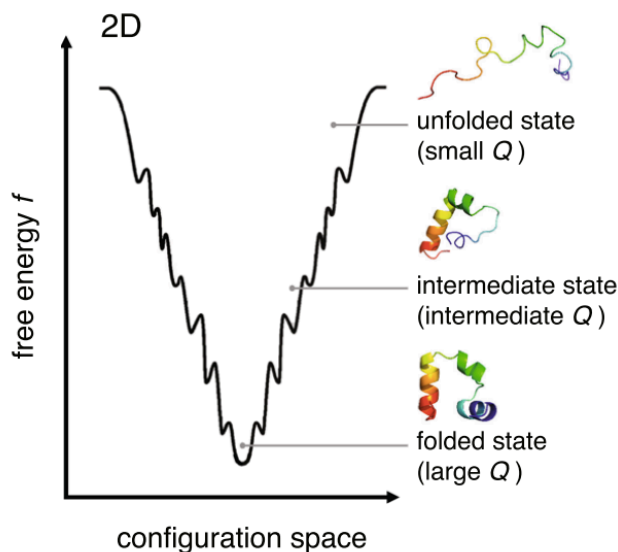


Figure 2.3. Sketch of the funnel-shaped free energy landscape.⁵⁷ (Chong, S. H., & Ham, S. (2019). Scientific reports, 9(1), 1-9. Open access journal and no copyright is required)

Proteins are not guaranteed to fold into their native structures. Misfolding can be triggered by amino acid sequence mutations or external factors that can disrupt the folding process.^{58, 59} β -sheets are often involved in protein misfolding, and β -sheets can be clustered and arranged into a stable cross- β structure via the backbone hydrogen bonds. The cross- β structure is highly stable and insoluble which allows the protein to protect itself from proteolysis. The misfolded protein can further aggregate into oligomers and amyloid fibril structure and eventually cause cell death.⁶⁰

2.2 Protein Misfolding Disease: Amyloidosis

2.2.1. Common Features of Amyloidosis

Amyloidosis are diseases caused by the protein misfolding and aggregation into abnormal fibril deposition. The abnormal fibril depositions are triggered by different factors, such as mutation or cleavage of the precursor protein. ¹

Formation of amyloid fibrils is characteristic of amyloidosis; examples of this include amyloid β fibrils which originate from the precursor protein of A β amyloidosis, or SAA fibrils which originate from the precursor protein of AA amyloidosis. The two proteins have significantly different primary, secondary, and tertiary structures; however, they eventually can form amyloid fibrils that have uniformed structure. The structure can be observed by the electron micrograph of tissues or in vitro samples. The fibril structure is typically made up by a single or several protofilaments. It appears straight and unbranched. The lateral width (~ 10 nm) of the single fibril is highly consistent among different precursor proteins. ^{61, 62} The cross- β structure is the fundamental structural component of the amyloid fibril. ⁶³ The structure is formed by the interchain interaction between β strands. ⁶¹ The Cross- β structure can be determined by different experimental techniques. For example, X-ray diffraction can measure a pattern of cross- β structure which contains two reflection signals at ~ 4.7 angstroms and ~ 8 -12 angstroms. The first one is relatively sharp compare to the second. ⁶¹ The structure also can be stained by Congo red and thioflavin S. Dyes can be used to diagnose amyloidosis.

Amyloid fibrils can adopt a similar cross- β structure regardless of the difference in precursor proteins that they are derived from. This cross- β structure suggests that a general mechanism is possible for amyloid fibril formation. A nucleation-dependent aggregation mechanism has been proposed based on many experimental and computational results. ⁶⁴⁻⁷¹ This mechanism contains

two essential phases, the nucleate formation and fibril elongation. During the nucleate formation oligomers are assembled by unstructured soluble monomers. High energy barrier needs to be overcome for oligomers forming. This forming process is rate-limiting and the oligomers at this time are in a so-called "fluid-like" disordered structure. When it approaches to a critical size, the disordered oligomer rearranges its structure into more organized oligomer that is rich in β -sheets.⁷²
⁷³ This structural rearrangement produces protofilaments. Unlike the nucleation phase, the elongation phase is a more energetically favorable process. During this process more and more monomers are extended to the protofilaments via a "dock and lock" mechanism, the elongated protofilaments form protofibrils and that then develop into mature fibrils.⁶⁸

2.2.2. Amyloid β Amyloidosis

In this thesis two different amyloidosis ($A\beta$ and AA amyloidosis) are discussed in detail and many other amyloidosis are referenced. $A\beta$ amyloidosis is a result of amyloid β ($A\beta$) peptides aggregation, it can lead to Alzheimer's disease (AD). AD affects more than ten millions of people across the world and causes substantial social and economic burden with its dramatically fast-growing numbers.⁷⁴⁻⁷⁶ $A\beta$ peptides are segments of amyloid precursor protein (APP) caused by proteolytic cleavage.⁷⁷ APP can be sequentially cleaved by α -secretases and γ -secretases, which generates the P3 peptide that is non-pathogenic, or by β -secretases and γ -secretases, which can generate a pathogenic $A\beta$ peptide. $A\beta$ peptides have different length because γ -secretases do not cleave at the same location absolutely. The most common variants are 40 ($A\beta_{1-40}$) and 42 ($A\beta_{1-42}$) residues long peptides. $A\beta_{1-42}$ contains two more hydrophobic residues at the C-terminal region, and this makes $A\beta_{1-42}$ more toxic and aggregation-prone.⁷⁸⁻⁸⁰

An A β monomer is termed as an intrinsically disordered protein (IDP) due to its structural flexibility in an aqueous environment. A β monomers have the ability to adopt multiple conformations.^{81, 82} The structures of A β monomer are mainly determined using ssNMR and computational simulation, such as molecular dynamics (MD). Under a membrane-like environment, A β ₁₋₄₀ and A β ₁₋₄₂ monomers adopted structures that contain high α -helix content. For example, the A β ₁₋₄₂ monomer has two helix regions (residues 8-24 and residues 28-38), and the two helices are linked by a turn region (residues 25-28).^{83, 84} However, a significant loss of helix content in the aqueous environment has been observed. The A β monomer in water solvent has the lowest helix content. Due to the IDP property of A β monomer peptides, MD simulations are employed to measure the structural information of A β monomer systems in both aqueous and membrane environments.^{85, 86} For example, the replica-exchange molecular dynamics study of A β ₁₋₄₀ and A β ₁₋₄₂ monomers have shown that the N-terminal region of monomers is famous for misfolding and aggregation while C-terminal regions favor membrane-like environment.⁸⁷

A β oligomers are intermediate states that form before the mature fibril. Among different A β structures A β oligomers are the most toxic.⁸⁸⁻⁹² Few structures have been determined through experimental approaches due to its unstable nature.⁹³ The mechanism of A β oligomers neurotoxicity is not fully understood. Three primary mechanisms have been proposed: the creation of pores that disrupt cell membranes, membrane structure perturbation, binding with the receptors on the membrane surface.^{93, 94}

Unlike the highly insoluble structures of A β oligomers, A β fibrils adopt a uniformed cross- β structure.^{61, 95, 96} Early ssNMR studies have shown that each chain of fibril adopts a turn region that connects to a "U-shaped" structure (both A β ₁₋₄₀ and A β ₁₋₄₂), which is built by two β -sheets

and a hydrophobic core is formed in between the β -sheets.^{97, 98} Multiple driving forces are involved in keeping the U-shaped structure stable. The first is an interchain backbone hydrogen bond between adjacent chains. Secondly, van der Waals interaction are observed in the hydrophobic core. And thirdly there is salt bridge interaction in between residues Lys28 and Asp23.^{99, 100} In 2015, an S-shaped fibril structure of $A\beta_{1-42}$ was revealed using ssNMR. This structure was different from the U-shaped structure which contains three β -sheets. Instead of relying on the salt bridge interaction between residues Lys28 and Asp23 for stabilizing the structure, S-shaped fibril relies on the hydrophobic packing between the last two residues and nearby β -sheet. Interestingly, unlike the U-shaped structure, only $A\beta_{1-42}$ S-shaped structure was observed. Simulation results showed that the S-shaped structure can shift between "in-register" and "out-of-register" fibril and eventually transforms to a β -barrel structure which can disrupt membranes; this may explain why $A\beta_{1-42}$ is more toxic. In a more recent cryogenic electron microscopy (Cryo-EM) study, an "L-S" shaped $A\beta_{1-42}$ structure has been reported. In the "L-S" shaped structure, L-shaped at N-terminus and S-shaped at C-terminus. The structure is stabilized through hydrophobic interactions among three hydrophobic clusters.¹⁰¹ More and more structures will be resolved with the improvement of experimental characterization. The same sequence of amino acids of $A\beta$ protein can lead to different fibril shapes, the polymorphism cause different interaction sites, and these sites give different properties of fibril surface, which may affect their cellular toxicity. Regardless of different topologies of structures, understanding the factors that stabilize the structure can provide insight into treatment for amyloidosis.

In recent years, more knowledge of A β structures has been accumulated through different experimental techniques. However, more works are still needed in understanding the early stage of misfolding and the mechanism of neurotoxicity caused by A β oligomers/fibrils.

2.2.3. Amyloid A Amyloidosis

The typical systemic amyloidosis is called AA amyloidosis. AA amyloidosis is a byproduct of the primary disease, such as neoplastic, chronic inflammatory, and infectious diseases. Under the primary disease conditions, the precursor protein SAA gets overexpressed and approaches a concentration that is 1000 times what the average concentration is. This crowded environment provides the SAA with the opportunity to misfold and aggregate.

SAA is highly conserved among different species, but we will use the SAA nomenclature here to represent proteins in the human system. There are three individual members, including SAA1, 2, and 4. Different genes encode them.¹⁰² SAA 1 and 2 are mainly expressed in the liver, and SAA 4 is expressed in tissues. SAA 1 and 2 are classified as acute-phase proteins that can bind with high-density lipoprotein (HDL) in plasma. The production of SAA is under the control of cytokines, of which the most important is IL-6.¹⁰³ The structure information can be seen in fig 1.3. This arrangement of the structure may also help to keep N-terminal stable due to N-terminal has been proved critical for amyloidosis. However, the structure of SAA binding HDL is still not known.

Not only SAA, but many amyloid fibrils are built by shorter fragments rather than full-length precursor protein. Cleavage of the parent protein can expose the hidden regions in the native structure, and the hidden regions of different chains can interact with each other, especially under

high local concentration. The full-length SAA in humans has never been observed to form amyloids in humans. However, shorter fragments of SAA that miss the C-terminal region can form fibril in vivo. The length of the fragments start from a few residues to more than 50 residues. With the cleavage on C-terminus, the N-terminal region is free to transfer to β -sheet structure.²²

The mechanism of SAA released from HDL is unclear. Increasing levels of HSPG may be critical for SAA and HDL dissociation. This is concluded from the observation that heparan sulfate (HS) can dissociate SAA from HDL.¹⁰⁴ Strong evidence have shown that macrophages in the reticuloendothelial system are critical to the SAA dissociation process.¹⁰⁵ Previous studies also reveals that the early stages of fibril formation happens in lysosomes when SAA is endocytosed. The C-terminus is truncated and the N-terminus is misfolded by the lysosome/cell causing intracellular oligomers to appear and disrupt the lysosome and cell membrane. This releases the protein into the extracellular environment where it is assembled into fibril structure (see figure 2.4).²⁹

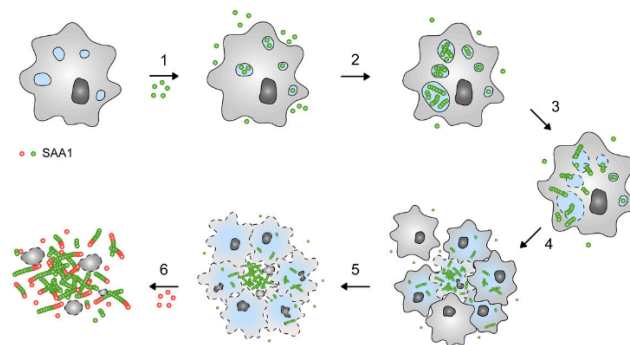


Figure 2.4. Formation of SAA fibril. 1). SAAs are endocytosed into cells. 2). Intracellular nucleation. 3). Intracellular leakage. 4). Cellular clusters formation. 5). Release to extracellular space. 6). Extracellular growth.²⁹ (Claus, S.; Meinhardt, K.; Aumüller, T.; Puscalau-Girtu, I.; Linder, J.; Haupt, C.; Walther, P.; Syrovets, T.; Simmet, T.; Fändrich, M., *EMBO reports* **2017**, *18* (8), 1352-1366. Copyright 2017 by EMBO Report, see details in Appendix IV)

Enzymes play an role in degradation and SAA fibril formation. The SAA amyloid formation can be accomplished in a short period with cleaved monomer. This observation indicates that cleavage is critical for SAA fibril formation. A previous study indicates that the lysosomal pathway degrades SAA. This degradation may be overloaded when local SAA concentration is too high.¹⁰⁶ Cathepsin B, a protease in the lysosome, can cleavage SAA to the most common 76 residues long fragment for amyloid formation.¹⁰⁷ In the extracellular environment there are metalloproteinases, MMP1, MMP2, and MMP3 that can interact with SAA protein. MMP3 can cleave SAA into shorter fragments for aggregation.^{108, 109}

As mentioned before, full-length SAA is not a notable species in SAA amyloid formation. In this case understanding the importance of cleavage seems critical to SAA fibril formation. Until recently the role of cleavage for SAA was unclear. Our work has shown that the cleavage is critical for shifting the equilibrium of SAA hexamer to a monomer which reduces the pressure caused by overexpression during the acute phase reaction in responding to the primary diseases; the monomer fragments, can later be degraded by Cathepsin D in the lysosome. The Cathepsin D can cleave the SAA fragment on the N-terminus and this cleavage can protect shorter fragments from aggregation. However, this degradation may be overloaded and causing it to eventually fail and this in turn causes intracellular aggregation.¹¹⁰

Recently, murine and human SAA fibril all-atom structures have been resolved via cryo-EM, (the human AA fibril was collected from a patient's kidney). The human fibril is made from the SAA fragment (residues 2-55) and the disordered region (residues 56-69); this disordered region is recognized by the diffuse density of cryo-EM.³⁰ This fibril structure is a well-packed two fold

structure. The packing interface between two folds is large and contains ionic, polar, and hydrophobic interactions as well as two layers to maintain a stable fibril structure. Studying the structure-property of fibril is critical to understand the fibril formation mechanism needed to provide insight into prevention/treatment development.

2.3. Simulation Methods

The majority of determined protein structures are collected in a protein data bank. They are determined by using a different methods such as X-ray diffraction crystallography, cryo-EM, and solid-state Nuclear Magnetic Resonance (ssNMR).¹¹¹⁻¹¹⁴ The structures are valuable and essential, but it is hard to get dynamics information from them. Interactions between proteins depend on the types of interactions. They happen on time and size scales, which are hard for the experimental methods to measure and observe. In order to capture the protein interactions in microsecond timescales and under atomistic resolution computational methods seem to be an excellent tool to be employed here.

2.3.1. Molecular Dynamics (MD)

Molecular Dynamics (MD) is an often used approach for protein structural prediction. MD can output realistic trajectories by using the Newtonian equation. The Born-Oppenheimer approximation of nuclear motion are assumed in MD, quantum effects are ignored, and each atom is defined as a point of mass.¹¹⁵ The relationship between the mass of atom i , and its position r_i , can be described by equation 2.1, where p_i stands for the momentum, m_i stands for the mass of atom i , dt stands for a short time interval.¹¹⁶

$$\frac{dr_i}{dt} = \frac{p_i}{m_i} \quad (2.1)$$

The net force F_i which is applied on atom i is calculated by using a potential energy function (dV) shown in equation 2.2. ¹¹⁶

$$F_i = - \frac{dV}{dr_i} \quad (2.2)$$

The expression for atom i of Newtonian equation of motion can be seen in equation 2.3. ¹¹⁶

$$\frac{dp_i}{dt} = F_i \quad (2.3)$$

Where r_i is a three-dimensional vector stands for atom i position in x-, y-, and z-direction. For example, $x(t)$ stands for at time t the x-coordinate of atom i . Using the standard Taylor series the position along x-direction in next concise time interval can be calculated by equation 2.4. ¹¹⁶

$$x(t + \delta t) = x(t) + \frac{dx(t)}{dt} \delta t + \frac{d^2x(t)}{dt^2} \frac{\delta t^2}{2} + \dots \quad (2.4)$$

$x(t)$ stands for a position, $\frac{dx(t)}{dt}$ for velocity and $\frac{d^2x(t)}{dt^2}$ for acceleration. Higher terms of the Taylor series expansion can be considered as 0 and the acceleration can be connected to mass and force by Newton's second law, see equation 2.5. F_x stands for the force to atom i in the x -direction. ¹¹⁶

$$\frac{d^2x(t)}{dt^2} = \frac{F_x}{m} \quad (2.5)$$

If the atom i 's position in direction x at time t and before t have already been understood, integration algorithms can be applied to get the position after the next δt time interval. One famous integration algorithms is the Verlet algorithm. Using Verlet, with the positions at time t , $t - \delta t$, and $t + \delta t$ is calculated using equation 2.6. ¹¹⁶

$$x(t + \delta t) = 2x(t) - x(t - \delta t) + \frac{d^2x(t)}{dt^2} \delta t^2 \quad (2.6)$$

The Verlet algorithm gives the position data using equation 2.7 and velocity data using equation 2.8. ¹¹⁶

$$x(t + \delta t) = x(t) + v_x \delta t + \frac{\delta t^2}{2m} F_x(t) \quad (2.7)$$

$$v_x(t + \delta t) = v_x(t) + \frac{\delta t}{2m} [F_x(t) + F_x(t + \delta t)] \quad (2.8)$$

2.3.2. Force Field Potentials

From the Verlet algorithm we can see that the position and velocity of atom at a particular time depend on the force applied on an atom. The force is determined by potential energy calculation. In MD the potential energy calculation is done using a force field. Many force fields are proposed and refined for accurate prediction of proteins. Here, CHARMM27, one of the most studied force fields, is utilized. To illustrate how potential energy is calculated using force fields, the potential energy equation can be seen in equation 2.9. Both bonded energy and non-bonded energy contribute to the potential energy.^{116, 117}

$$U_{CHARMM} = U_{bonded} + U_{nonbonded} \quad (2.9)$$

The bonded energy term includes bond stretching, bond angle bending, a "Urey-Bradley" term, dihedral energy term, improper dihedral energy term, and cmap term. They can be seen in equation 2.10.¹¹⁷

$$U_{bonded} = U_{bond} + U_{angle} + U_{UB} + U_{dihedral} + U_{improper} + U_{CMAP} \quad (2.10)$$

$$\begin{aligned}
U_{bonded} = & \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 \\
& + \sum_{Urey-Bradleys} k_{UB}(b - b_0)^2 + \sum_{dihedrals} k_\phi(1 + \cos(n\phi - \delta)) \\
& + \sum_{impropers} k_\omega(\omega - \omega_0)^2 + \sum_{residues} u_{CMAP}(\Phi, \Psi) \quad (2.11)
\end{aligned}$$

U_{bonds} is the term that defines the bond stretching, the force field constant k_b and b_0 (naught values) are taken from force field parameters. U_{angle} is the term that defines the angle bending, the force field constant k_θ and naught values are taken from force field parameters. U_{UB} is the so-called "Urey-Bradley" term which is an addition to U_{bonds} and U_{angle} . The term is introduced to angle term in some cases during the vibrational spectra optimization; the force field constant k_{UB} and naught values are taken from force field parameters. $U_{dihedral}$ is the term that describes the dihedral angle, periodicity parameters n , phase shift parameters δ , and the force field constant k are taken from force field parameters. $U_{improper}$ is the term that describes the out of plane dihedral angle, the force field constant k_ω and naught values are taken from force field parameters. U_{CMAP} term is a correction for dihedral angle values on the backbone. It is used to improve protein backbones' conformational properties. ¹¹⁷

The $U_{nonbonded}$ term contains the electrostatic potential and Lennard-Jones potential, see equation 2.12.

$$U_{nonbonded} = U_{LJ} + U_{elec} \quad (2.12)$$

$$U_{nonbonded} = \sum_{pairs} \epsilon_{ij} \left[\left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{min}}{r_{ij}} \right)^6 \right] + \sum_{pairs} \frac{q_i q_j}{\epsilon r_{ij}} \quad (2.13)$$

The Lennard-Jones potential term is defined using the first term of equation 2.13. It describes the center of the mass distance between atoms i and j. ϵ_{ij} is force field constant and is taken from force field parameters. $\left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{12}$ stands for the short-range repulsive term, $\left(\frac{r_{ij}^{min}}{r_{ij}} \right)^6$ stands for the attractive long-range term between atom i and j.¹¹⁷

The electrostatic potential is a function of electrostatic interactions between pairs of charged atoms. Due to the non-uniform distribution of charges for some molecules, one way to solve this in MD is to assign partial charges to the center of mass for both atoms.

2.3.3. Replica Exchange Molecular Dynamics (REMD)

According to the funnel-shaped energy landscape, while a protein is searching for its native structure it could become trapped in local minima. Although MD can provide the physical trajectory the conformations only follow one pathway, and it is easy for the protein to become

trapped in the free energy local minima. In order to solve this problem, replica exchange molecular dynamics (REMD) was proposed to increase sampling efficiency. REMD alleviates slow sampling by exchanging replicas at different temperatures. High temperature can easily overcome the energy barrier while the low temperatures can explore the low energy conformations. In REMD, under different temperatures, multiple MD simulations run simultaneously for one same system; the distribution of temperature is from low to high. MD simulations under different temperature are replicas. Replicas get their configurations swapped with the neighbor replicas by velocities reassignment if a Metropolis criterion is satisfied. π_{T_i} stands for the Boltzmann distribution of replica under temperature T. Temperatures distribution optimization is critical for REMD simulation to be successful.

$$\pi_T(q, p) = \frac{e^{-\frac{E(q,p)}{k_B T}}}{\sum_{(q,p)} e^{-\frac{E(q,p)}{k_B T}}} \quad (2.14)$$

The Metropolis criterion is satisfied. The metropolis criterion is given below

$$\min\left(1, \frac{\pi_{T_i}(q_j, p_j) \pi_{T_j}(q_i, p_i)}{\pi_{T_j}(q_j, p_j) \pi_{T_i}(q_i, p_i)}\right) \quad (2.15)$$

2.3.4. Summary

To summarize, this chapter provides the biochemical background of proteins and protein folding and presents the protein folding process from a theoretical aspect. Under the situation of misfolding, the misfolded peptides can aggregate into amyloid fibrils. The fibril is further deposited in tissues and organs and it is this abnormal deposition that can cause diseases related to amyloidosis. Amyloidosis can become either local or systemic depending on the position of fibril deposition. Different forms of amyloidosis is generally triggered by different precursor proteins misfolding and then aggregating. Understanding the misfolding or amyloid formation is essential as it can offer guidance for treatment development for related diseases. Studies always start with the structures of precursor proteins. Experimental approaches have provided crucial high-resolution information for these structures; however, the experimental approaches cannot reveal the dynamics or critical interaction of the structure as the time scale of these interactions is smaller than experimental approaches can measure. Computational approaches are more suitable to answer these questions. An overview of selected modern simulation methods is given, and exactly how the protein systems and interactions in the systems are extracted to models and potentials.

Chapter 3. Stability of a Recently Found Triple- β -Stranded A β ₁₋₄₂ Fibril Motif

The material in chapter 3 is adapted from Xi, W., Wang, W., Abbott, G., & Hansmann, U. H. (2016). Stability of a recently found triple- β -stranded A β ₁₋₄₂ fibril motif. *The Journal of Physical Chemistry B*, 120(20), 4548-4557. The copyright permission is obtained from, see details in Appendix IV.

Author Contributions: Work in this chapter is majorly credited to Dr. Wenhui Xi. However, Wenhua Wang spent three month of his rotation working with Wenhui Xi and contributed his fair share to this work. Wenhua Wang run the productive run of one-fold four-layer system, one-fold five-layer system, and test runs for two-fold systems. Wenhua also participated in all RMSD, RMSF and cluster analyses. Wenhua wrote part of the draft of this paper (materials and methods, part of results and discussion, specifically, results for one-fold systems) and contributed to revisions of the paper. Above all, it's legitimate to include this work as a chapter of Wenhua's dissertation.

3.1. Introduction

A growing number of human diseases are correlated with the presence of amyloid fibrillar plaques¹¹⁸⁻¹²⁰, visible under ultraviolet light after staining with Congo Red.¹²¹ A prominent example is Alzheimer's Disease¹²², a neurodegenerative disorder that is the leading cause of senile dementia^{123, 124} and characterized by plaques made mainly from 37-43 residue-long amyloid- β (A β)

peptides.¹²⁴ Most common are the $A\beta_{1-40}$ and $A\beta_{1-42}$ species, with the $A\beta_{1-42}$ more toxicity than $A\beta_{1-40}$.¹²⁵ While the insoluble amyloid plaques are neurotoxic, there is evidence that solvable oligomers (which could be either on the fibril formation pathway or off-pathway) are more toxic than the mature fibrils.^{126, 127} The self-assembly of $A\beta$ peptides leads *in vitro* to polymorphic oligomers and fibrils that share a cross- β sheet packing.¹²⁸⁻¹³² The differences in molecular structure between these amyloid polymorphs is correlated with the speed of disease progress.^{132, 133} Hence, for the purpose of deriving and evaluating treatment options it is important to understand the details of amyloid formation, especially of the process that selects specific polymorphs. An example is the seeding and acceleration of the prolongation of new fibrils by pre-formed fibril fragments.¹³⁴ Such fragments can not only seed proteins and peptides of the same kind but also different amyloidogenic molecules. For instance, $A\beta_{1-40}$ fibrils can cross-seed $A\beta_{1-42}$ fibrils and vice versa.¹³⁵ This effect is known as cross-seeding,¹³⁶ and may explain the statistically observed correlations between occurrence of certain amyloid diseases.¹³⁷ For example, $A\beta_{1-40}$ and $A\beta_{1-42}$ can seed *in vitro* other amyloid peptides such as hIAPP (amylin) implicated in type II diabetes.^{135, 138-141} Similarly, there is evidence that some of the $A\beta$ -amyloid polymorphs induce conversion of amyloids with different structure to their own form, and in this way seed their own growth. This gives rise to the idea that some polymorphs can act as “infectious strains” in patients, which may also explain why no polymorphism is found in fibril taken from Alzheimer patient brains.¹⁴²⁻

144

An understanding of the role polymorphism in the pathology of Alzheimer disease requires therefore a characterization and comparison of such forms. A number of fibril structures have been resolved for wild-type $A\beta_{1-40}$ -fibrils. Most resemble each other and are made out of chains with

two β -strands connected by a loop region.¹⁴⁵ For instance, in one model are the two β -strands formed by residues 9-23 and 31-30 and connect without a without salt-bridge holding them together. The fibril has two-fold packing symmetry with the chains in contact through the β 2-strands.¹³⁰ Another model of $A\beta_{1-40}$ has three-fold packing symmetry of the chains, with each chain made of two β -strands of residues 12-19 and 31-38 connected by a loop.¹³² For a long time, only low-resolution $A\beta_{1-42}$ fibril structures had been known, which indicate the presence of the same U-shape motif, with two β -strand: residues 18-26 and residues 31-42 connected by a loop, and the arrangement stabilized by salt bridges between residues D23-K28.¹²⁹ However, using ss-NMR Y. Xiao et al¹⁴⁶ have determined recently a high-resolution structure for wild type homogenous $A\beta_{1-42}$ fibrils that has a different morphology. This $A\beta_{1-42}$ fibril structure is not build out of U-shaped chains, but each peptide forms a three-stranded β -sheet with strands β 1 made of residues 12-18, β 2 of residues 24-33, and β 3 of residues 36-40. An intra-chain salt bridge between side chain of residue K28 and the main chain residue A42 that is not seen in previously found structures connects the β 2 and β 3 strands. While the newly found fibril motif can seed $A\beta_{1-42}$ chains, it cannot seed $A\beta_{1-40}$ peptides. Xiao et al have speculated that this is because the salt bridge K28-A42 cannot be formed in $A\beta_{1-40}$ peptides, i.e. that the stability of the newly found structure depends on formation of this salt bridge.¹⁴⁶

The striking difference in structure and seeding properties between the newly found form and previously known $A\beta$ fibril structures opens up a new opportunity to study the mechanisms that lead to formation of distinct polymorphs. As the process by that fibrils form is difficult to follow in either experiment or simulation, such investigations best start by investigating the thermodynamic stability of the respective polymorphs. Extending the many computational studies

have looked into the stability of the previously found fibril forms¹⁴⁷⁻¹⁴⁹ to the newly found triple- β -strand motif of $A\beta_{1-42}$ fibrils, we hope to identify the mechanism that leads to the differences between the various forms. For this purpose, we focus on the following questions: first, what is the driving force of stability in this triple- β -strand conformation? Second, what is the relation between the truncation of last two residues and the mechanism of cross-propagation barriers? Third, do protofibrils made out of $A\beta_{1-42}$ chains with this new motif have a single-fold or two-fold packing symmetry. Our molecular dynamic simulations allow us to identify a critical minimal size for of six chains such fibril fragment. We find a surprising stability for the $\beta 2$ -turn- $\beta 3$ motif even in simulations with substantially raised temperatures. The triple- β -strand motif of $A\beta_{1-42}$ fibrils remains stable even in mutants where the K28-A42 salt bridge cannot form. However, truncation of last two residues leads to the exposition of the hydrophobic core in the U-shape motif and results the disruption of the conformation. Hence, cross-seeding of $A\beta_{1-40}$ fibrils is not so much inhibited because the K28-A42 salt bridge cannot be formed but because of the difficulty to form in $A\beta_{1-40}$ peptides the well-packed hydrophobic core of the $\beta 2$ -turn- $\beta 3$ motif that is seen in the $A\beta_{1-42}$ triple- β -strand motif. Our simulations also suggest that the $A\beta_{1-42}$ with the newly found structure will likely arrange in fibrils with a two-fold packing symmetry, with the chains packing at their $\beta 1$ chains.

3.2. Materials and Methods

3.2.1. Fibril Conformations

In order to test the stability of the newly found triple- β -strand motif of $A\beta_{1-42}$ fibrils and the forces that drive their assembly, we have built a number of systems that are derived from the first model

in Protein Data Bank ensemble (PDB identifier: 2MXU). As the first ten residues are flexible, this model describes only fibrils of A β ₁₁₋₄₂ fragments. In order to exclude the possibility of a bias in our data resulting from excluding these first ten residues, we perform also as a control a simulation of the fibril structure for the full-length peptide A β ₁₋₄₂ with the first ten residues forming a β -strand. Fibrils of different size, ranging from four to seven chains are simulated at the physiological relevant temperature of 310 K, and these simulations suggest a critical size of five to six chains. Hence, in most of our simulations we considered fibrils formed by five chains. Besides the wild type we also study fibril fragments of this size with the mutation K28A, or such where the last two residue have been truncated (i.e. fibrils of A β ₁₁₋₄₀). In order to study the possible conformations of two-fold-symmetric fibrils, we have constructed two assemblies differing in the packing surfaces between the two parts. As the β 1-strand and the β 3-strand are potentially exposed to the solvent, we assume that the chains will associate at the surfaces formed by these strands, and that the packing surface are either between hydrophobic residues or between oppositely charged residues. Ten different initial states with β 1- β 1 association, and ten states with β 3- β 3 packing, have been constructed by optimizing the inter-chain contacts. The ten initial states of each arrangement are followed over short molecular dynamics simulations 10 ns length and tested for their stability. The most stable configurations are used for as start point of our simulations and shown in **Fig. 3.6** for both arrangements.

3.2.2. Simulation Protocol and Data Analysis

Our molecular dynamics simulations are carried out in an isothermal–isobaric (NPT) ensemble using the GROMACS 4.5.7 software package¹⁵⁰ with the amber ff99sb-ilde force field¹⁵¹ and TIP3P water molecules¹⁵². Bond lengths are constrained in the solute with the LINCS algorithm¹⁵³,

while the water geometry is constrained with the SETTLE algorithm¹⁵⁴, allowing for an integration time step of 2 fs. Electrostatic interactions are calculated with the particle mesh Ewald (PME) method and a real space cutoff of 1.0 nm.¹⁵⁵ The cutoff for van der Waals (vdW) interactions is 1.4 nm. Solute and solvent are coupled separately to external temperature and pressure baths using, respectively, v-rescale¹⁵⁶ thermostat and Parrinello–Rahman barostat¹⁵⁷. Most of the simulations are at a temperature of 310 K and a pressure of 1 bar, but we also study some of our systems at temperatures of 330K, 350K, 375K and 400K, Each system us followed over at least four trajectories that start from slightly different initial states. Each trajectory is followed for at least 200 ns, and for fibril fragments with either five or six chains is one of the trajectory extended up to 500ns. In order to allow the systems to thermalize, we omit the first 50 ns in each trajectory. Cluster analysis of the remaining trajectory is used to identify representative conformations. The binding energy are analysis with the MMGBSA tools¹⁵⁸ in the Amber 12 software package¹⁵⁹ where the generalized Born (GB) model developed by H. Nguyen et al is employed to estimate the solvent effect.¹⁶⁰ This procedure leads to an estimate for the binding free energy given by:

$$\Delta G_{bind} = \Delta E_{vdw} + \Delta E_{electrostatic} + \Delta E_{GB} + \Delta E_{surf} - T\Delta S$$

where E_{elec} is the electrostatic and E_{vdw} van der Waals interaction energy in gas phase. E_{GB} is the Generalized-Born approximation of the polar solvation energy, and E_{surf} is non-polar solvation energy. Note that configurational entropy contributions are neglected in this approximation.

3.3. Results and Discussion

3.3.1. Stability of A β_{11-42} Fibrils with the Triple- β Motif

In order to analyze the stability of the newly-found A β_{11-42} fibril structure with its triple- β -motif, we determine first the minimal size of A β_{11-42} fibril fragments with this structure. For this purpose, we have simulated suitable fragments with either four, five, six or seven chains, with the start configuration taken from the Protein Data Bank (PDB id: 2MXU) as shown in **Fig. 3.1a**. The root mean square distances (RMSD) to the corresponding initial conformation (build from the first entry of the ssNMR ensemble as deposited in the Protein Data Bank) are shown in **Fig. 3.1b** as function of time. For each of the systems we have generated four trajectories, and **Fig. 3.1b** shows these trajectories where the root-mean-square deviation after 200 ns was maximal. Over the whole 200 ns are the root-mean-square deviations of the fragments with four chains (black) or five chains (red) larger than the corresponding values of the systems with six (blue) or seven (pink) chains. This size depends is due to the larger relative weight of end chains in the smaller fragments, as end chains are more flexible and lose their secondary structure more easily, see **Fig. 3.1b**. We also note that the β 1-strands (12-18) are more flexible than other parts of the chains, especially for the fragments with four or five chains (**Fig 3.1c**). In order to check the robustness of our results we have for the systems with five and six chains the simulations extended up to 500ns. Representative final conformations are shown in **Fig. 3.1c**. Note, that between 200 ns and 500 ns the root-mean-square-deviation of the five-chain-system (red line) continues to rise while the system with six chains (blue line) appears to approach a plateau, Some of the β -strands have dissolved in the final conformation of the five-chain system while all are preserved in the six-chain system (**Fig. 3.1c**). From these results we conclude that the critical size A β_{11-42} fibrils with triple-

β -motif is about six chains, with the fibril fragment size of five chains right on the boundary between stable and unstable. Hence, most of our simulations are based on systems with either five or six chains. Note that with five to six chains the critical size of the $A\beta_{11-42}$ fibrils with triple- β -motif is larger than that of earlier found $A\beta$ wild type fibril structures. For instance, the critical size for the previously found low-resolution $A\beta_{11-42}$ fibril structure 2BEG is three to four chains.¹⁶¹ The larger critical size may indicate a higher free energy barrier for forming fibrils with the triple- β -motif than seen in the previously found fibril structures with U-shaped chains, which could explain why this motif has not been observed earlier.

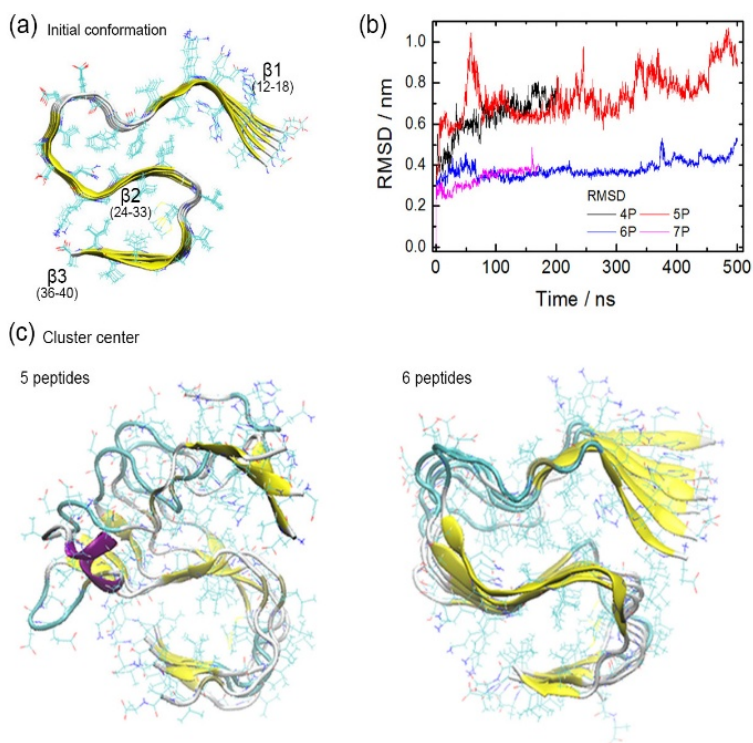


Figure 3.1. (a) The initial state of a system with five chains as used in our simulations. (b) The rmsd of fibrils, followed over 500ns; (c) representative conformation of the final configurations of fibrils with five and six chains.

In the initial configurations as build from the PDB configurations, all chains are in a periodic arrangement without twist in the β -strands. Such lack of twist has been seen also in other models of $A\beta_{1-40}$ fibrils^{129, 130, 132}, however, for these models did the β strands always acquire a twist over the course of molecular dynamics simulations.^{149, 161, 162} This is different for the $\beta 2$ and $\beta 3$ strand of the new $A\beta_{1-42}$ fibril model which both do not develop a twist during 200 ns of simulation. At the end of the respective trajectories these two strands have a twist of less than 1 degree per chain, unlike the $\beta 1$ -strand which behaves more as expected and develops a twist of about 3 degree per chain. As far as we know, this is the first time that such lack of twist as here in the $\beta 2$ and $\beta 3$ strands has been seen in atomistic molecular dynamics simulations of $A\beta$ fibrils. As the energy of β -strands is more favorable if the strands have a slight twist of about 3-5 degree per chain, the lack of twist implies that strong packing interactions must compensate for the unfavorable backbone interactions that result from lack of a twist. Note that the secondary structure (as measured by the `do_dssp` tools¹⁶³ in GROMACS) is maintained during the 200 ns albeit the edge chains lose their $\beta 1$ strand (data not shown). The lack of twist implied that single-fold $A\beta_{1-42}$ fibrils with the triple- β -motif would vary over 45 nm between five and six nm, while with twisted strands one would expect that the diameter variations happens on a length scale of 15 nm.

Previous experiments have shown that the first 15 residues of $A\beta_{1-42}$ are coiled without taking an unique conformation.¹²⁹ Other work showed that the first 8 residues of $A\beta_{1-40}$ are flexible.¹³⁰ In the new $A\beta_{1-42}$ model, the first 10 residues are also considered disordered, and are therefore omitted in our simulations. In order to exclude the possibility that this omission skews our data we have performed simulations with a model where we have added these first ten residues, initially forced into a β -strand to have a defined start structure. As expected¹⁶⁴, this N-terminal

strand dissolves over the length of the simulation and the first ten residues line up as a disordered flexible segment. However, the addition of these additional N-terminal residues does not change the stability of fibril where in each chain the residues 11-42 stay in their triple- β -strand motif, see **Fig. 3.2**. Hence, we feel that our choice of omitting the first ten residues in our simulations is justified

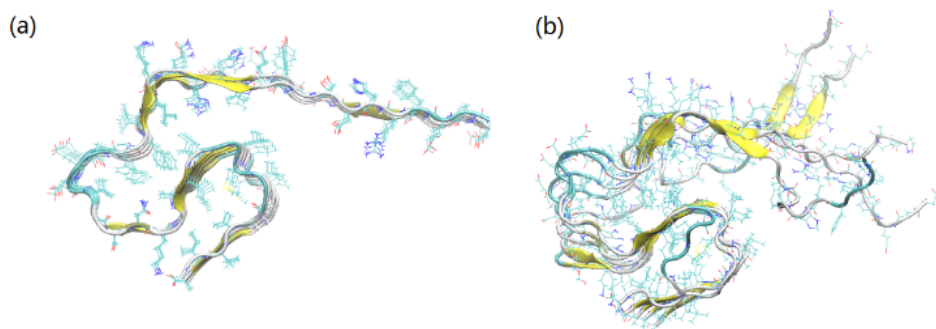


Figure 3.2. A β_{1-42} (a) initial conformation. (b) representative conformation after 200ns.

We remark that we have also simulated the five-chain fragment with a different force field, CHARMM27¹⁶⁵, to exclude the possibility of artifacts introduced by choice of a specific force field (Amber ff99sb-ild¹⁵¹). Final configurations, and the evolution of secondary structure and root-mean-square deviation along the trajectories are similar for both force-fields (data not shown) suggesting that our results are not biased by the choice of force-field. In order to exclude yet another possible source of bias, we have also repeated our simulations taking this time as start configuration not the first entry of the NMR ensemble but the one that has the largest root-mean-square deviation to the first entry. This is the ninth model which differs from the first model in the side chain packing in the β_2 -turn- β_3 motif. The results after 200 ns are again for both types of

simulation similar in terms of stability (data not shown), and we conclude that our results do also not rely on the choice of a specific NMR entry.

3.3.2. Temperature Dependence of Fibril Stability

The speed with that fibrils evolve over time is a function of temperature. In general, fibrils will dissolve faster at elevated temperature. Hence, in order to assess the stability of our fibril fragments in another way we have simulated the five-chain-fragment at elevated temperatures 330K, 350K, 375K and 400K. We display in **Fig. 3.3** the root-mean-square fluctuation (RMSF) of the chain residues as function of temperature together with representative configurations sampled at the various temperatures. As the temperature rises, the root-mean-square-fluctuations of residues 12-18 in the β_1 strand grow far more than the corresponding values of residues in the β_2/β_3 regions. This indicates a higher flexibility of the β_1 strand compared to the more rigid β_2 -turn- β_3 motif of residues 24-40. The difference in stability persist up to 375K, and only at 400 K have both the β_1 (12-18) and β_3 (36-40) lost their initial conformation. The root-mean-square-fluctuations of the residue 17-20 (forming a turn), and of the residues 30-42 of the β_2 -strand are lower than that of the other segments, indicating the higher stability of these parts. The residues K16 to E22 are known to form the hydrophobic core of A β peptides.¹⁶⁶ In the triple- β -motif, the phenylalanine residues at position 19 and 20 have their side chains engulfed by the turn 1 region. This is different from the previous models of A β_{1-42} ¹²⁹ where the residue F19 is exposed to solvent. In the β_2 region, the hydrophobic residues I31 and I32 are also packed inside the turn region instead of being exposed to solvent. The side chain of residue I31 points to the turn 2 region while the residue I32 is located between turn 1 and the β_1 strand, and is in contact with residue F19. Hence,

because of this packing of hydrophobic residues it appears that the turn 1 and $\beta 2$ region in the new model are energetically more favorable than they are in the previous models of $A\beta_{1-42}$.

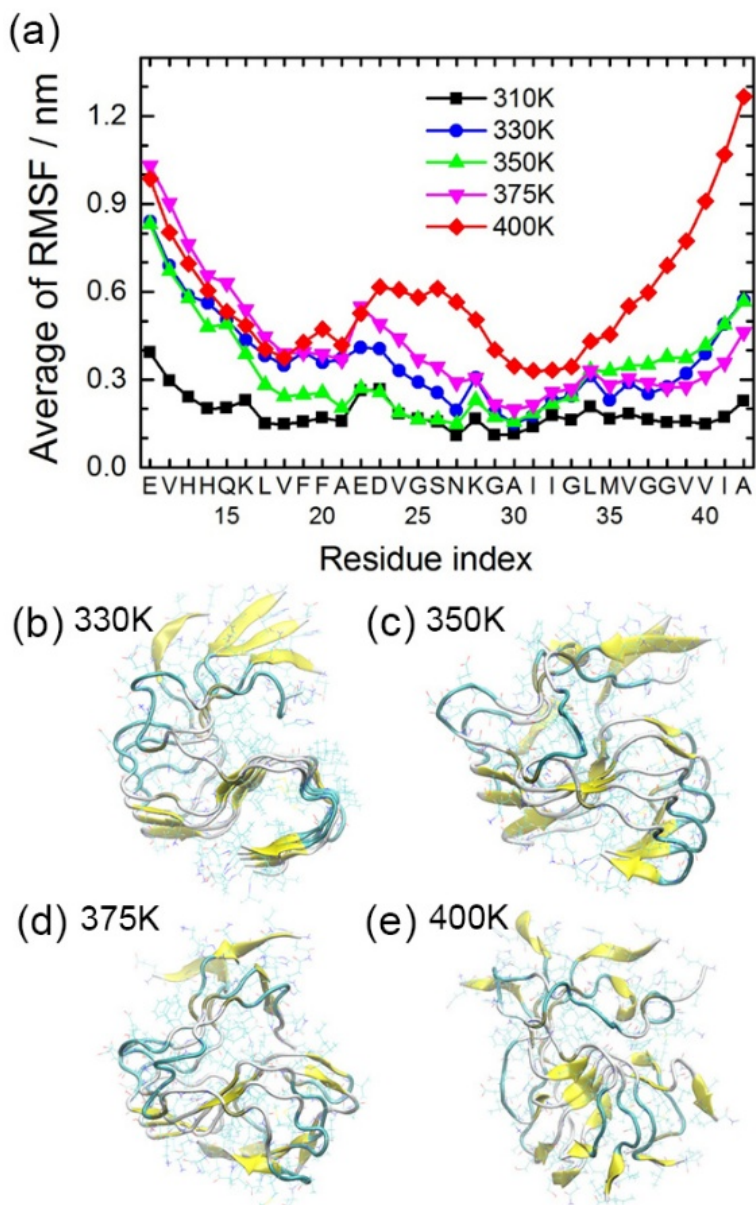


Figure 3.3. (a) Variation of the residue-specific RMSF with temperature averaged over all five chains of the fibril. Shown are also representative conformations of the chains at temperatures 330 K(b), 350K (c), 375K (d) and 400K (e).

3.3.3. Role of the K28-A42 Salt-bridge

The previous model of $A\beta_{1-42}$ has a salt bridge between residues D23-K28¹²⁹ that is also found in $A\beta_{1-40}$ ¹³⁰ fibrils with two-fold-symmetry, and that is known to be important for the stability of these fibrils.¹⁶⁷ Note, however, that no salt-bridge is formed by residue K28 in $A\beta_{1-40}$ fibrils with three-fold-symmetry. In the new model, the side chain of K28 form an intra-chain salt-bridge with the main chain of the C-terminal residue A₁₋₄₂.¹⁴⁶ Xiao et al have speculated that the stability of the newly found structure depends on formation of this salt bridge,¹⁴⁶ and that the impossibility to form this salt bridge in $A\beta_{1-40}$ peptides is the reason why the triple- β -strand motif is not seen in $A\beta_{1-40}$ fibrils. In order to test this hypothesis, we have compared simulations of the wild type $A\beta_{1-42}$ with that of mutants K28A. However, while the C-terminal segment became more flexible and the β -strand propensity in the $\beta 3$ region decreased, see **Fig. 3.4a and c**, did the $\beta 2$ -turn- $\beta 3$ motif in the mutant not dissolve over 200 ns of simulation. The stability of this motif suggests that, unlike expected, this salt bridge is not necessary for the stability of the new structure.

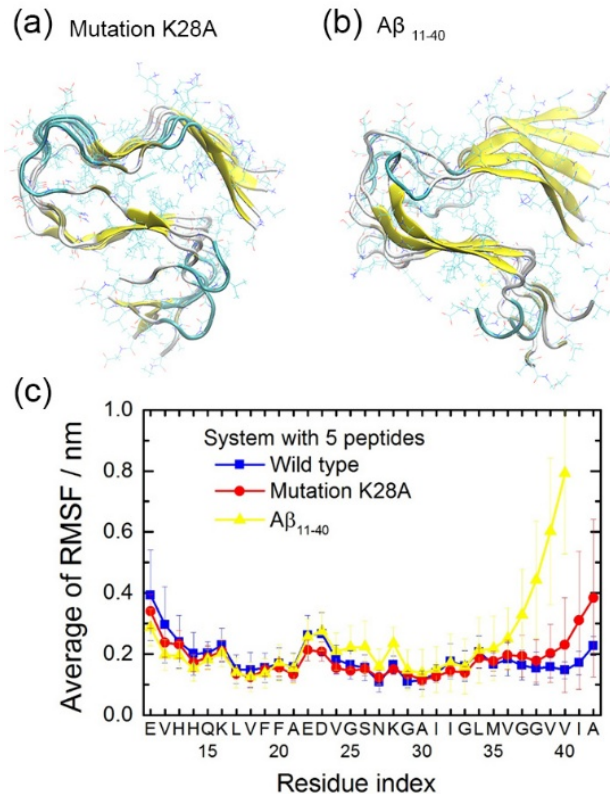


Figure 3.4. Representative conformation of mutation K28A(a) and Aβ₁₁₋₄₀(b). (c)The RMSF of the K28A mutant Aβ₁₁₋₄₂ (red) and Aβ₁₁₋₄₀(yellow) compare to the wild type Aβ₁₁₋₄₂ (blue).

One way to test this conjecture is by simulating the fibril with each of the five chains truncated to residues 11-40. If this system which also does not have the K28-42 salt bridge has a different behavior, it would indicate that not the K28-A42 salt bridge is responsible for the stability of the new form but rather other interactions involving the missing two C-terminal residues I41 and A42. Indeed we find in simulations of the truncated system the stability of the β₂-turn-β₃ motif (residues 24-40) greatly reduced. (**Fig. 3.4b**). Without the last two residues, the β₃ strand (residues 36-40) becomes too short to form a β-strand and the chains rather assume an U-shape form similar to the one seen for Aβ₁₋₄₀ in previous ss-NMR experiments.^{129, 130} The C-terminal residues 35-40 turn into a random coil, see also Fig. 3.4c, with the hydrophobic residues I31 and M35 now exposed to

the solvent. Besides excluding the possibility to form the K28-A42 salt bridge, the lack of the terminal two residues I41 and A42 also removes the hydrophobic contacts of residues I31-I41 and residues V39-I41. The side chains of residues I31 and V39 are now exposed to solvent and the three hydrophobic residues I31, M35 and V39 can no longer maintain the β 1-turn- β 2 motif. On the other hand, the root-mean-square-fluctuations of the β 1-turn- β 2 motif (residues 12-33) did not change in the truncated version (**Fig 3.4c**), i.e. the β 2-strand is still preserved. Hence, comparing the truncated system with the mutant system, which both lack the salt-bridge K28-A42, it appears that the stability of the newly found $A\beta_{1-42}$ fibril structure depends more on the contacts and steric constraints induced by the C-terminal residues I41 and A42 than on the salt bridge K28-A42.

While our above result suggest that the salt bridge K28-A42 is not the leading factor in the stability of the new motif seen in $A\beta_{1-42}$ fibrils, formation of the salt bridge may be a key factor in the kinetic process that leads to folding of the triple β -strand structure. In order to see whether the salt-bridge is a guiding factor, we have further simulated wild type systems that started from configurations in which the β 2- β 3 sheet is broken, but where the K28-A42 salt-bridge is preserved. If this salt bridge guides the formation of the new motif one would expect that this motif (i.e. the β 3-strand and its attachment to the β 2 strand) would re-form over the course of our simulation. The initial conformation is derived from high-temperature simulations with the K28-A42 salt-bridge and residues 11-28 restrained, and is shown in supplemental **Fig. S3.5a**. As for the other systems, we follow our system in four independent trajectories. In none of these runs does the β 3 strand re-form. Instead, we observe a gradual disruption of the amyloid conformation as shown in **Fig. 3.5b**. During the 200 ns of simulation time, the K28-A42 salt bridge remains stable, but the residues 29-42 turn into a disordered loop instead of re-forming the β 2- turn - β 3 motif. This lack of repair of

the C-terminal β -sheet raises further doubts on the importance of the K28-A42 salt-bridge for forming and stabilizing the new structure. However, we cannot exclude the possibility that the β sheet would need longer than the 200 ns of our simulation time to form again. Previous experiments and molecular dynamics simulations indicate that a strand-loop-strand motif requires microsecond to milliseconds to form,¹⁶⁸ much longer than our simulations, albeit we would expect a faster process for our system as the K28-A42 salt bridge restricts the conformation space.

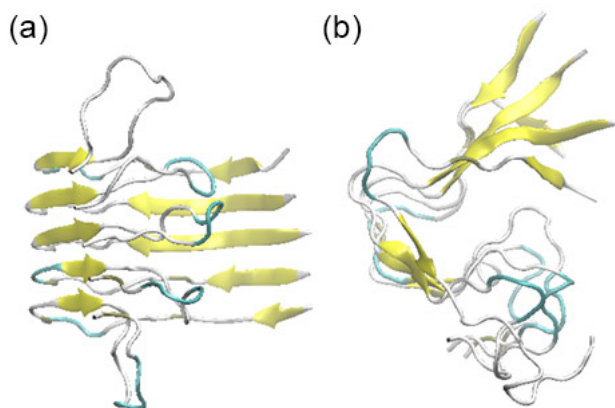


Figure 3.5. A β_{11-42} fibril with the $\beta 2$ - $\beta 3$ strand broken but intact K28-A42 salt-bridge. (a) initial states. (b) representative final conformation with disrupted $\beta 2$ -turn- $\beta 3$ motif but most of the K28-A42 salt-bridge is still preserved.

3.3.4. Packing Surfaces of A β_{1-42} Fibrils with Two-fold-packing Symmetry

While it was initially assumed that the A β_{1-42} fibril existed as single-fold conformation¹²⁹ build out of a single protofibrils, later ss-NMR experiments have shown that the A β_{1-40} fibril can exhibit two-fold¹³⁰ or three-fold packing symmetry^{131, 132}, with, for instance, in the two-fold arrangement the two chains in each unit associated by their $\beta 2$ -strand (segment 31-37) and with an intra-chain

salt-bridge formed between D23 and K28.¹³⁰ In the three-fold arrangement there is no D23-K28 salt-bridge and the packing surfaces are made out by segments of residues 28-31 and the C-terminal of the chains.¹³¹ In the patient brain-derived fibril the chains also arrange with three-fold symmetry, packing through residues 22-31 and the C-terminal.¹³² In all the above cases, the A β chains take the previously known U-shape form. Whether the recently found triple- β -stranded form is part of fibrils with a single fold, or such with two-fold or three-fold symmetry, is yet not resolved experimentally. In order to explore the possibility of fibrils with two-fold packing symmetry, made out of chains with the triple- β -strand motif, we have constructed various arrangements that differ in their packing surface, and have investigated their stability. The single chain has two strands that can be exposed to the solvent, the β 1-strand and the β 3-strand. Hence, one can expect that the two chains in each unit will associate through the surfaces formed by these strands. This suggests the two arrangements shown in **Fig. 3.6**. **Fig. 3.6a** displays the case where the two chains assemble by way of their β 1-strands in such a way that the charged residues K17, E22, and D23 form inter-chain salt-bridges. **Fig. 3.6b** presents another possible arrangement where the two chains assemble by way of much shorter (five-residue long) β 3-strands. Here, the two chains are arranged in such a way that the number of contacts involving the hydrophobic residues G38-V40 is maximized. We call the arrangement in **Fig 3.6a** where the two chains assemble by way of their β 1-strands PSA, and the corresponding arrangement by way of their β 3-strands in **Fig. 3.6b** PSB.

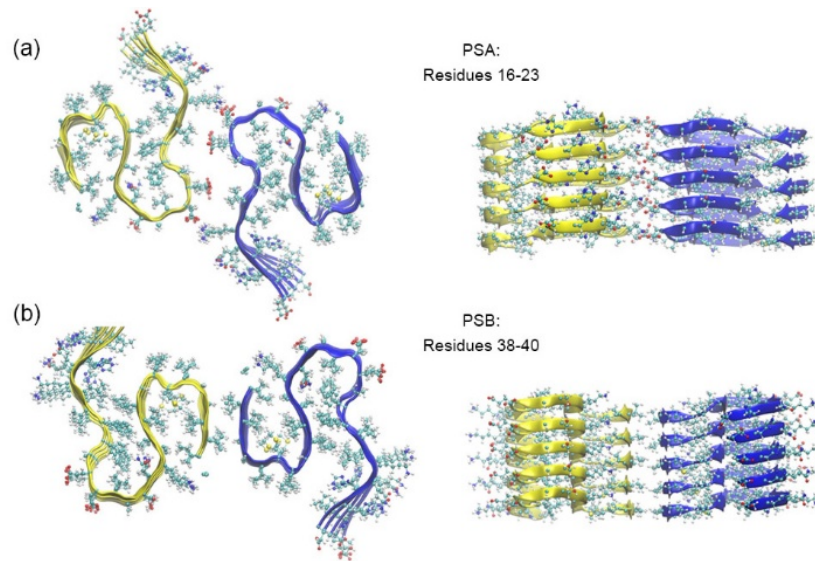


Figure 3.6. The initial states of two proposed chain arrangements in fibrils with two-fold symmetry (a) PSA, packing between $\beta 1 - \beta 1$. (b) PSB, packing between $\beta 3 - \beta 3$.

Even after 200 ns of simulation time did neither the PSA nor the PSB fibrils separate. Representative conformations are shown in **Fig 3.7**. In the PSA fibril does the residue K16 form an inter-chain salt bridge with residues E22 or D23 that stabilizes the two-fold (**Fig. 3.7a**). The residues segment 16-23 bent so that a water pore is formed between the surfaces in contact. Since the $\beta 1$ strands are twisted, the whole fibril has a twist. On the other hand, the chains in the PSB fibril interact mainly through contacts between the side chains of residue V40 (**Fig. 3.7b**). The $\beta 2$ -turn- $\beta 3$ motif stays stable in both chains, and as the $\beta 3$ -strand does not develop a twist, the whole fibril is also not twisted.

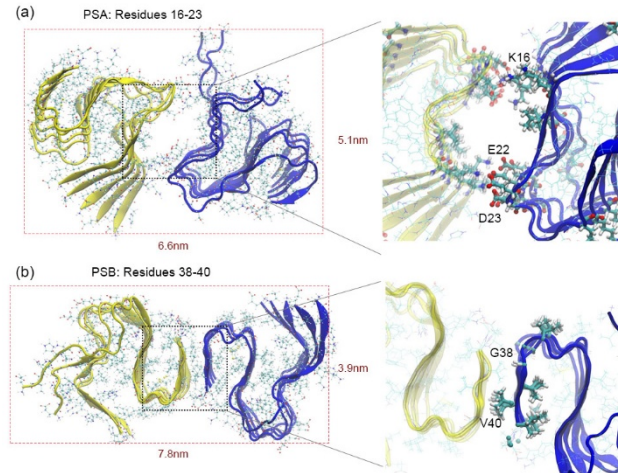


Figure 3.7. Representative conformations of fibrils with two-fold symmetry. (a) In the PSA arrangement, residue K16 forms an inter-chain salt bridge with residues E22/D23. The cross-section of the fibril is about $6.6 \times 5.1 \text{ nm}^2$. (b) the PSB arrangement is stabilize by contacts between the residues V40, and the cross-section is about $7.8 \times 3.9 \text{ nm}^2$.

We show the time evolution of the root-mean-square-deviation of each chain shown for both the PSA and PSB fibrils is in **Fig. 3.8**. The values for both arrangements are with about 0.7nm comparable with the values measured for the single-fold fibrils shown in **Fig 3.1b** In agreement with the above discussed visual inspection of the final states these values indicate that on the time scale of our simulations (200ns) fibrils with two-fold packing symmetry are stable. However, the PSA and PSB arrangements differ in the diameter of the resulting fibrils. For the PSA fibril we find a cross-section of $6.6 \times 5.1 \text{ nm}^2$, and a corresponding value of $7.8 \times 3.9 \text{ nm}^2$ for the PSB fibril. Hence, fibrils with a PSB arrangement of chains are more oblate than such within a PSA arrangement of chains. Transmission electron microscopy measurements lead to experimental values for the fibril diameters in the range of 4.5-6.0 nm for thinner filaments, and 6.0-14.0 nm for wider filaments. As the diameters of single fold protofibrils are of order 4.0-5.0 nm, the wider

filaments cannot be formed by single-fold fibrils, and both the PSA and PSB arrangements are in better agreement with the experimentally measured diameters.

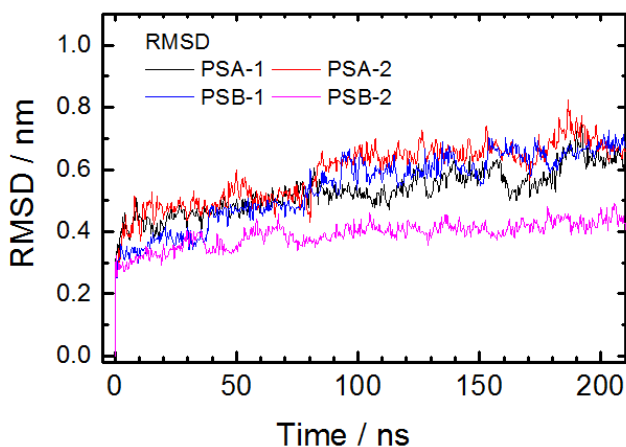


Figure 3.8. Root-mean-square-deviation of each layer in fibrils with two-fold packing -symmetry.

We have further analyzed the two chain arrangements in fibrils with two-fold symmetry by estimating the binding energy of the two chains within the MM-GBSA approximation. In fibrils with PSA arrangement of chains we find as total binding energy -56.3 ± 1.2 kcal/mol and a slightly higher value of -46.6 ± 0.5 kcal/mol for fibrils with PSB arrangement of chains. The various terms that make up the binding energy are listed in **Table 3.1**. The contributions of each residue to the binding energy are shown in **Fig. 3.9a**. The two chain arrangements are shown again **Fig. 3.9b** and **3.9c**, but now with the residues colored according to their binding energy contribution, with red marking attractive interaction and blue repulsive terms. In PSA fibrils the strongest contribution come from interactions involving residues K16 and E22 or D23, i.e. the residues that form the inter-chain salt bridge. This observation is consistent with the large contributions from electrostatic and solvent term for this fibril arrangement, see **Table 3.1**. Similarly, the most important contribution in the PSB fibrils come from the hydrophobic residue V40, and

correspondingly, the binding energy estimate is dominated by the VDW interactions and solvation energy. While the absolute values indicate a slighter preference for a PSA arrangement of chains, these numbers have to be taken with a grain of salt as conformational entropy contributions are not accounted for in our MM-GBSA analysis. However, the calculated binding energy estimates support our earlier observation that both chain arrangement will lead to stable fibrils. As the diameters of such fibrils with two-fold packing symmetry is in better agreement with the Transmission Electron Microscopy measurements than that of single-fold fibrils, we believe that A β ₁₋₄₂ chains with triple- β -strand structure likely build fibrils with two-fold packing of the chains. Both the PSA and the PSB arrangements appear to be valid candidates. While the PSB fibril model is in better agreement with the experimentally measured fibril diameters, the PSA model leads to a lower MM-GBSA binding estimate. As the chains in this arrangement are hold together by contacts involving the charged residues K17, E22 and D23 between the two chains, a PSA arrangement in fibrils with two-fold packing symmetry may also explain why the in other fibril structures often found salt bridge between residues K28 and E22 or D23 is not seen.

Table 3.1. Binding energy of fibrils with either PSA or PSB arrangement

	Binding energy	
	Kcal/mol with Std. Err.	
	PSA	PSB
E_{vdw}	-12.7 (0.9)	-73.1 (0.6)
E_{elect}	-1800 (30)	173(3)
E_{GB}	1764 (29)	-139 (3)
E_{surf}	-8.5 (0.2)	-7.9 (0.1)
E_{total}	-56.3 (1.2)	-46.6 (0.5)

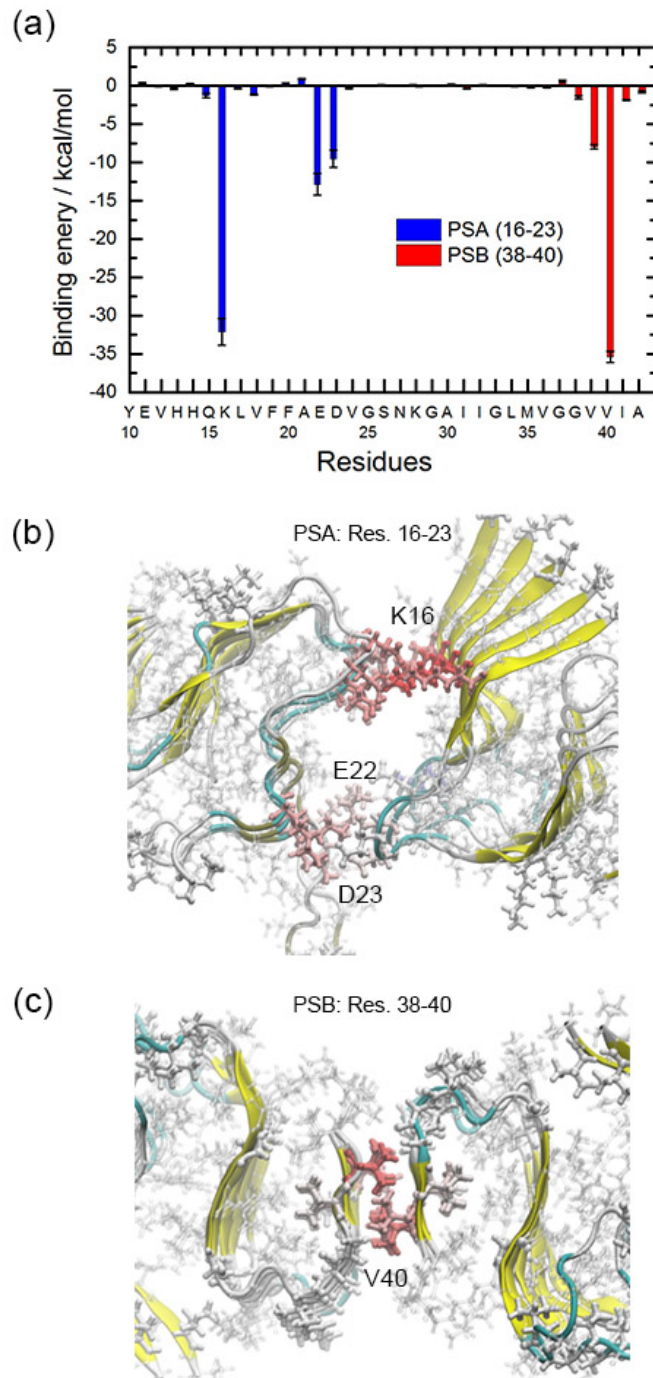


Figure 3.9. Binding energy between protofibrils in fibrils with two-fold packing. (a) Binding energy distributions for each residues. The coloring for corresponding residues in $\beta 1$ (b) and $\beta 3$ (c) corresponds to the one in (a), with “red” indicating attractive and “blue” contributions.

3.4. Conclusions

Using molecular dynamics simulations we have studied the stability of a recently resolved high-resolution structure for wild type homogenous $A\beta_{1-42}$ fibrils¹⁴⁶ that differs significantly from previously found $A\beta_{1-40}$ fibril¹³⁰⁻¹³² and low-resolution $A\beta_{1-42}$ fibril structures¹²⁹: the individual chains form a triple- β -strand motif instead of a U-shape. It is tempting to assume that the difference in structure is correlated with the higher neurotoxicity of $A\beta_{1-42}$ amyloids but the relevance of the new motif is not clear. We find a critical fibril size of five to six chains. Fibril fragments with less than five chains dissolve while such with six or more chains change little over the course of our simulations. Hence, the critical size is larger for the newly found $A\beta_{1-42}$ fibrils than for the earlier found low-resolution $A\beta_{1-42}$ fibril structures¹⁶¹ where it is three to four chains. This larger critical size may indicate a higher free energy barrier for forming the $A\beta_{1-42}$ fibrils with the triple- β -strand motif than for forming $A\beta_{1-42}$ fibrils with U-shaped chains. This could explain why this motif has not been observed earlier and would also suggest a longer nucleation phase for such fibrils. On the other hand, the new structure is surprisingly stable for temperatures up to 400 K. At a physiological temperature of 310 K all three β -strands persist during molecular dynamics simulations of up to 500 ns, with the $\beta 2$ -turn- $\beta 3$ motif more stable than the $\beta 1$ strand. Unlike the $\beta 1$ -strand the strands in this $\beta 2$ -turn- $\beta 3$ motif surprisingly do not become twisted even after 500ns. The lack of an energetically more favorable twist in the middle and C-terminal strand suggest strong packing interactions that were missing in previous molecular dynamics simulations of other $A\beta$ -fibril models where initially untwisted strands always quickly became twisted. These strong packing interactions suggest that $A\beta_{1-42}$ fibril with the triple- β -strand motif are energetically more favorable than the previously found low-resolution $A\beta_{1-42}$ fibril models with U-shaped chains. One can speculate that once formed such amyloids are more stable and more capable of seeding

aggregates. Simulating the mutant K28A and wild type systems where the residues I41 and A42 are truncated, we find that the salt bridge between the side chain of K28 and C-terminal of A42 adds to the stability of the structure, but appears to be neither for kinetic nor thermodynamics reasons necessary for its formation and stability. Instead, the stability of the β 2-turn- β 3 motif depends on the hydrophobic packing of side chains and steric constraints induced by the C-terminal residues I41 and A42. This explains why $A\beta_{1-42}$ fibril with the triple- β -strand motif can seed fibril growth of $A\beta_{1-42}$ chains but not of $A\beta_{1-40}$ chains, and we expect that mutations of these two C-terminal residues will disrupt the triple- β -strand motif, i.e that these mutants could not be seeded by wild type $A\beta_{1-42}$ fibril with the triple- β -strand motif. Such mutation studies may also help to identify the mechanisms that separate formation of wild type $A\beta_{1-42}$ fibril with the triple- β -strand motif from such with U-shaped chains. In order to study this question we are now also preparing simulations that will probe transitions between these two states. We have finally considered different packing arrangements of protofibrils build out of $A\beta$ chains with the triple-strand motif. In these fibril models, two chains either interact by contacts between residues in their β 1 strands, or by residues in their β 3 strands. Approximating the binding energies between the two protofibrils by MM-GBSA estimates, we find that both arrangements are more favorable than single fold fibrils, but our analysis does not allow us to select one of the two arrangements as the more likely one. As the geometries of both of our proposed arrangements are more consistent with Transmission Electron Microscopy measurements than single-fold fibrils¹⁴⁶, we conclude that both arrangements are valid candidates for fibrils with two-fold packing symmetry. This conjecture could be tested by mutation experiments involving the residues K16, E22, E23 that are not critical for the stability of fibril but for the packing chains. We would expect that such mutations lead to an reduction of the fibril diameters. We remark that our results do not exclude the possibility of

fibril packing with three-fold or higher fold organization. While the PSB arrangement cannot be extended to a three-fold model, this would be possible for a PSA like packing involving residues K16 to E22 or D23. Packing through contacts involving these residues allows for formation of an equilateral triangle leading to fibrils with a diameter of approximately 11 nm, also consistent with the values of 6.0 nm to 14.0 nm measured for thicker filaments by Transmission Electron Microscopy.

3.5. Acknowledgments

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 4. Stability of the N-terminal Helix and its Role in Amyloid Formation of Serum Amyloid A

The material in chapter 4 is adapted from Wang, W., Xi, W., & Hansmann, U. H. (2018). Stability of the N-Terminal Helix and Its Role in Amyloid Formation of Serum Amyloid A. *ACS omega*, 3(11), 16184-16190. Open access journal and no copyright is required.

Author Contributions: This work in this chapter is majorly credited to Wenhua Wang, Dr. Wenhui Xi participated in discussion and model preparation.

4.1. Introduction

A large spectrum of diseases is connected with the presence of amyloid fibrils that after staining with Congo Red are visible under ultraviolet light. Depending on the specific disease, these deposits are either systemic or localized (for instance, in the patient's brain in the case of Alzheimer's disease), and appear either spontaneous or as the consequence or byproduct of another disease.¹⁶⁹⁻¹⁷³ An example for the latter, so-called secondary, amyloidosis is Colonic Amyloidosis where in response to inflammation the precursor protein Serum Amyloid A (SAA) is overexpressed.¹⁷⁴ Normally, SAA is found with a concentration of 1-3 µg/ml in human blood, but in patients with colon cancer or inflammatory bowel disease the concentration of SAA can approach more than 1 mg/ml. This extremely high concentration then encourages mis-folding and aggregation of SAA, leading to the outbreak of colonic amyloidosis as a secondary disease that adds to the symptoms of the primary disease.¹⁷⁵⁻¹⁷⁷

Serum Amyloid A is built out of 104 amino acids, however, in extracellular environment, or in amyloids, one does not find the complete protein but rather the truncated fragment SAA(1-76). The structure of the full-size protein has been resolved by X-Ray crystallography²² and deposited in the Protein Data Bank (PDB) under identifier 4IP8. While the crystal structure, shown in **Figure 4.1**, is for a hexamer, it is stable for the monomer, and characterized by a long C-terminal tail that wraps four helices corresponding to residues 1-27, 32-47, 50-69 and 73-88.²² From mutation experiments it is known that the first eleven residues are crucial for the mis-folding and aggregation of SAA in colonic amyloidosis as their lack prevents the aggregation of SAA.²³ In the crystal structure, these residues are part of the N-terminal α -helix, but it was proposed by Nordling and Nordling²⁴ that residues 1-13 can mis-fold into a β -hairpin structure whose presence will trigger aggregation.

This idea is consistent with Jannone's observation of Raman spectra seen during SAA(1-12) aggregation²⁵, and our own simulations which show an equilibrium between α -helix to β -hairpin configurations for the N-terminal segment SAA(1-13) that can be shifted by various single point mutations.²⁶

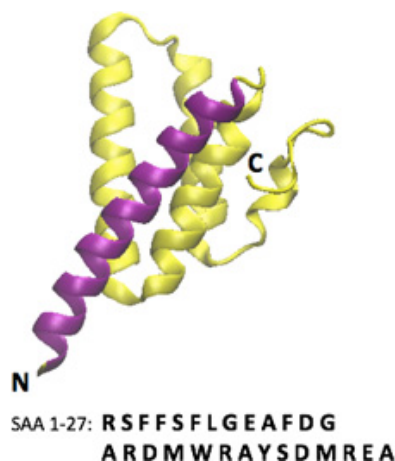


Figure 4.1. The crystal structure (PDB ID: 4IP8) of the full-size SAA monomer. The purple segment is helix I (residues 1-27). Helix II, III, and IV and the C-terminal tail, consisting of residues 28-104, are drawn in yellow. Labels mark the N- and C-terminals.

A drawback of our previous study is that it considered only an isolated segment formed by residues 1-13. However, in the crystal structure these residues part of an α -helix spanning from residue 1-27. The local environment formed by this α -helix will likely alter the landscape of the 1-13 segment and the mechanism of the helix-hairpin transition. For this reason, we extend here our previous study to the segment formed by the first 27 residues of SAA, investigating structural transitions in it, and the potential role of these transitions in the formation of SAA amyloids. For this purpose, we consider not only the wild type, but probe also how landscape and transition mechanism are altered by the mutation of residue 9 from a Glutamic Acid into an Alanine (E9A), a mutation that disrupts a potentially helix-stabilizing salt bridge seen in the wild type at neutral pH. Our results rely on large-scale simulations using a variant replica exchange molecular dynamics (REMD), multi-exchange replica exchange molecular dynamics (ME-REMD), which raises the efficiency of regular REMD by attempting multiple exchanges of neighboring replicas between the molecular dynamic segments. An additional goal of this chapter is to evaluate in a systematic way the gain

in efficiency by this approach.

4.2. Materials and Methods

4.2.1. Models

The initial configuration is derived from the corresponding fragment (residue 1-27) of the crystal structures of full-length serum amyloid A protein 1.1 (PDB ID: 4IP8). The mutation-type SAA (1-27) E9A were built from the wild-type by replacing the negatively charged side chain of Glutamic acid with the hydrogen atom of the alanine side chain. The so-obtained configurations of wild-type and mutant are together with 4700 water molecules placed in a box with box size 5.4 nm and periodic boundary conditions, and melted by molecular dynamics simulation at 500 K to generate random configurations as starting point for the molecular dynamics simulations in this study.

4.2.2. Simulation Protocols

These simulations are done with the GROMACS 4.5.6 software package.¹⁷⁸ The CHARMM 27 force field with CMAP and TIP3P water model^{152, 179} CHARMM TIP3P are employed in replica exchange molecular exchange (REMD) and multi-exchange replica-exchange molecular dynamics (ME-REMD) of the SSA(1-27) fragment in the isothermal-isobaric (NPT) ensemble.¹⁷⁸ Each replica is followed over 300ns. For REMD, the system contains 36 replicas, spread between 300 K and 420K, for ME_REMD only 28 replica are distributed over this temperature interval. The temperatures are controlled by velocity rescaling^{2,180} and the pressure is kept at 1 bar by the isotropic Parrinello-Rahman's method. Constraining peptide bonds with LINCS¹⁵³ allows us to integrate the equations of motions with a time step of 2 fs. Replica exchange are attempted every 1000 steps. , method is employed to maintain a constant temperature¹⁸⁰ and is involved to

keep a constant pressure at 1 bar. Because we use periodic boundary conditions are long-range electrostatic interactions calculated with Particle Mesh Ewald¹⁸¹ algorithm using a 1 nm cutoff. The same cutoff is employed for calculation of van der Waals interactions.

For the data analysis we omit the first 200 ns to allow the system to reach equilibrium. Only the 310 K replica is considered in REMD simulations, and the 308 K replica in ME-REMD simulations. The secondary structure of configurations is calculated with the PROSS algorithm¹⁸² which relies on measurements of dihedral angles only. Configurations were cluster with an in house tool defining clusters by certain geometric motifs (such as extended helices, helix-hairpin) of the configurations. Sidechain contact maps were derived using the gromacs tool `g_mdmat` which is based on the mean distance between each residue. The solvent accessible surface area analysis is conducted by using gromacs tool `g_sas`.

4.3. Results and Discussion

4.3.1. System Set-up

In the present study we use molecular dynamics simulations to investigate structural transitions in the N-terminal segment SAA (1-27), a protein implicated in colonic amyloidosis. In order to ensure validity of our results we have to make certain that the set-up of our simulation does not introduce a bias to our data. Such bias could come, for instance, from a too small simulation box that would cage the molecule and therefore restrict its configurational space. In supplemental Figure S4.1 we present results from short constant temperature (400 K) simulations of the wild type fragment SAA (1-27) for cubic boxes of length 4.8 nm, 5.4 nm, and 5.6 nm. Shown is the probability of a residue to be part of a helix, strand or turn. Comparing the results, we find that the results from box sizes

5.4 nm and 5.6 nm are consistent, but differ from that of box size 4.8 nm. Hence, we conclude that a box size of at least 5.4 nm is necessary in simulations to minimize the finite size effects.

In a similar way, we need to make sure that the temperature distribution allows a walk of replicas through temperature space, and that the highest temperature is sufficiently large to enable crossing of all relevant barriers in the system. The lowest temperature is given by our target temperature, $T=300$ K. Short constant temperature runs indicate that a temperature of $T=420$ K is sufficient to dissolve the helix, and for this reason we choose $T=420$ K as our maximal temperature. The distribution of temperatures between the two extremes can be determined for a given number of replicas by the algorithm of van der Spoel.¹⁸³ Especially, we find that 36 replicas lead to an average acceptance probability of 18% which we consider acceptable. Simulating the wild type with this box size and temperature distribution for 300ns, and evaluating various quantities, such as the secondary structure probability of certain segments, for different time intervals, we find that a regular REMD simulation converges within 200 ns. Hence, this set-up for regular REMD (see supplemental Figure S4.2) is used as our standard against which we compare all considered variants when testing the efficiency of multi-exchange replica-exchange molecular dynamics (ME-REMD).

4.3.2. Comparing the Efficiency of REMD and ME-REMD Sampling

Replica exchange molecular dynamics (REMD) is a method to enhance sampling of configurations in systems where large barriers in the energy landscape lead to slow sampling. REMD alleviates this problem by introducing a walk through temperature space. At high temperature, barriers can be easily crossed, while at low temperatures more relevant low-energy configurations are explored. A disadvantage of this technique is that the temperature spacing decreases with increasing system

size (or otherwise the acceptance for exchange moves becomes vanishingly small). This makes simulations of proteins in explicit water (i.e., where the system consists of the protein and a much larger number of water molecules) cumbersome. One way to raise the effective transition rate is to attempt multiple exchanges for two neighboring replicas.¹⁸⁴ As a consequence, a given replica can cover a large range of temperatures between two molecular dynamics segments if a series of exchange attempts between pairs of neighboring replicas are made, resulting in a faster walk in temperature space. While it appears reasonable to assume that such an approach will lead to improved sampling, it is important to explore systematically the gain in efficiency that can be obtained, and how this gain depends on the number of such exchange attempts within one exchange cycle.

Hence, in a first step we took our system of 36 replicas, simulated with regular REMD, and we varied the number of exchange attempts from $N=1$ to $N=5$, 10 , 20 and $N=100$. The resulting effective exchange rates for the various temperature pairs are shown in Figure 2a, and suggest a fast approach to an optimum, with increasing the number of attempts beyond $N=10$ not raising the effective exchange rate noticeable. Here, we define the effective exchange rate by the number of times that, after the series of N exchanges, the final and initial configurations of a replica differ. As a consequence of the higher exchange rate, replicas walk faster in temperature space. This can be seen in table 1 where we list the number of tunneling events and average tunneling times for various N . Here, we define a tunneling events as the walk of a replica through the temperature space, from lowest temperature to highest temperature and back, see Figure 4.2c and 4.2d for illustration. Note in table 1 the much faster walk in temperature space with ME-REMD than seen for regular REMD.

One possible application of the increased effective exchange rates and faster walk in temperature space, is the possibility to use ME-REMD with a smaller number of replicas than needed for regular REMD. Varying the number of replicas in the same temperature interval of $T=300\text{ K}$ to $T=420\text{ K}$ from $M=36$ to $M=28$ we measured again for various numbers N of exchange attempts the resulting effective exchange rates, and plot these in Figure 4.2b. As in the case of $M=36$ replicas, the effective exchange rate approaches an optimum for $N \approx 10$. For $M=28$ replicas corresponds this optimal effective exchange rate to the one seen when the original system ($M=36$) is simulated with regular REMD. In a similar way agree the number of tunneling events and average tunneling times, see table 1. For illustration, we show in Figure 4.2c and 4.2d for both the 36 REMD and the 28 ME-REMD simulations example walks of replicas through temperature space. Hence, by using ME-REMD instead of regular REMD we reduce the number of replicas in our system from $M=36$ to $M=28$ replicas, a reduction of computational resources by more than 20%!

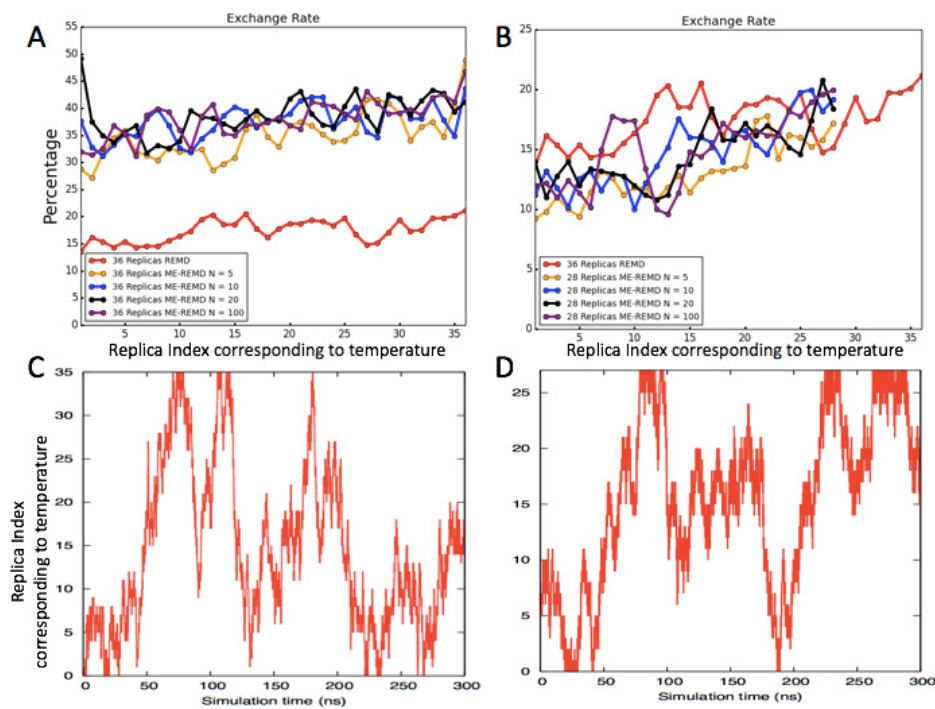


Figure 4.2. a) Observed exchange rate in simulations of the SAA (1-27) fragment, the red line is the exchange rate for 36 replica REMD, orange, blue, black and purple lines stand for the accumulated exchange rate in ME-REMD with $N=5, 10, 20,$ and 100 . B) Similar to a), the orange, blue, black and purple lines stand for the accumulated exchange rate in ME-REMD with $N=5, 10, 20,$ and 100 , but now for a simulation with only 28 replicas distributed over the same temperature interval of 200K to 420 K; for comparison is also shown the regular REMD simulation of the 36 replica system.

Various approaches have been proposed in the past that result into faster walks through temperature space. In some cases, these approaches lead to biased sampling and non-reliable results. In order to exclude this possibility, we compare in Figure 4.3 the secondary structure propensity of the residues as measured with regular REMD for our original system of $M=36$ replicas, with the data measured in our ME-REMD simulation of the system with $M=28$ replicas. In both cases, we find a propensity for formation of α -helices of about 50%, 51 % for ME-REMD and 55 % for REMD, and substantial propensity for turn and β -strand formation, see the left and

middle panel of Figure 4.3. In both REMD and ME-REMD results, we observed that the wild type SAA(1-27) fragment has a tendency for the helix to break around residue 11-13. The qualitative and quantitative similarity between the two runs shows that the enhanced sampling in ME-REMD does not introduce any biases into our simulation; and for our further analysis we therefore rely on the data obtained with ME-REMD sampling.

Table 4.1. Number of tunneling events, average tunneling time and average exchange rate over 300ns for regular REMD and ME-REMD. The various quantities are defined in the text.

Systems	Time Interval (ns)	Tunneling Event Number	Average Tunneling Time (ns)	Average Exchange Rate (%)
36 Replicas REMD WT	0-300	46(1)	47.7(3.8)	17.9(0.2)
28 Replicas ME-REMD WT	0-300	40(5)	54.6(7.1)	14.7(0.1)
28 Replicas ME-REMD Mutation E9A	0-300	30(3)	59.4(4.1)	11.3(0.1)

4.3.3. Configurational Ensemble of SAA (1-27) Monomers

While in the crystal structure of Ref. ¹¹ all 27 residues form an α -helix, the helix propensity is reduced in our simulations, with a break of the helix around residues 11-13. These results are consistent with previous experiments, showing that the truncation of the first eleven residues can

prevent aggregation. Our data indicate three regions: helix Ia (residue 1-11), turn (residue 12-13), helix Ib (residue 14-27). As the first segment, residues 1-11, has a lower helicity than the segment formed by residues 14-27, it follows that the helix Ia is less stable than helix Ib, and its presence or lack of therefore may well be the key for aggregation to happen.

The secondary structure propensity distribution is consistent with our clustering analysis where we group configurations according to the secondary structure propensity of residues 1-11 and 14-21 (residues 22-27 are always helical and therefore ignored for clustering). We find three main clusters. The first cluster is made of configurations with a helix-turn-helix hairpin, with at least five consecutive residues in each of the two segments identified as helical, but separated by at least two residues that are not helical. The second cluster is made of configurations with a stable helix Ib (again defined by the requirement of at least five helical residues in the segment 14-21), but where the N-terminal residues 1-11 form an elongated, disordered and dynamically changing segment (with no more than three consecutive helical residues). Finally, the configurations in the third cluster are characterized by single long helices, resembling the crystal structure for this fragment, and are defined by the requirement of a single helix of at least 14 residues covering both segments. Example configurations are shown in the upper level of the left and middle panel of Figure 4.3. The helix-turn-helix cluster has the largest population and is contains about 33% of all configurations. The second largest group is the disordered N-terminal cluster consisting of about 25% of configurations. The lowest frequency (5%) is seen for configurations with the native-structure-like single extended helix. The various frequencies are listed in Table 2.

The above results raise the question on what breaks the extended helix seen in the crystal structure, what leads to the dominance of the helix-turn-helix hairpin motif, and how is this all connected to

the role of the first eleven residues in amyloid formation of SAA? Residue 13 is a glycine and residue 14 an alanine, and both residues make it likely to break the helix in this segment. Residue 11 is a phenylalanine whose aromatic sidechain may form a weak hydrophobic contact with residue 21Y to break the extended helix motif and stabilize the helix-turn-helix motif, shown in Figure 4.5B. Residue 12 is an aspartic acid, and Figure 4.5 A and B indicate that there is weak contact between residue 1R and residue 12D sidechain in the extended helix motif that is not observed in the helix-turn-helix motif. Hence, the two residues may form a transient salt bridge which stabilizes the extended helix motif, while on the contrary the helix-turn-helix motif is preferred when the salt bridge dissolves. Hence, while helix 1 (residues 1-27) is stable in the crystal structure of the full protein as a hexamer, the higher flexibility of the much smaller segment SAA(1-27) allows this helix to break up around residues 12 and 13 into two segments, the N-terminal helix Ia (residues 1-11) and Ib (residues 14-27).

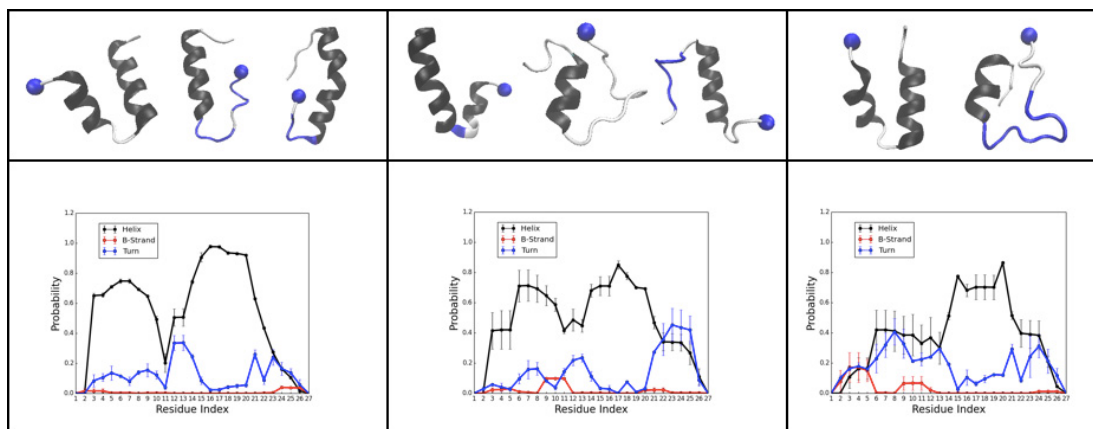


Figure 4.3. Secondary structure frequency (lower row) and representative configurations for the dominant motifs (upper row) as found in REMD simulations with 36 replica of the wild type SAA (1-27) fragment (left panel), in ME-REM simulations with 28 replicas of the wild type SAA(1-27) fragment (middle panel), and in ME-REM simulations with 28 replicas of the E9A mutant SAA(1-27) fragment (right panel). The frequency of α -helices is drawn in black, that of turns in blue, and of β -strands in red. The same color coding is used in the figures in the upper panel, with blue balls marking the N terminal of the segments.

In order to understand the appearance of the helix-hairpin motif we calculate the contributions of each residue to the solvent accessible surface area (SASA) and compare the obtained values for the different motifs. The corresponding values are shown in Figure 4.4. In helix-hairpin configurations, the hydrophobic residues 3F, 21Y and 24 are less exposed to water than in the straight helix of the crystal structure. Hence, we conclude that the helix hairpin is stabilized by hydrophobic contacts involving these residues as shown in Figure 4.5B, that do not exist in the elongated helix of the crystal structure (see Figure 4.5A).

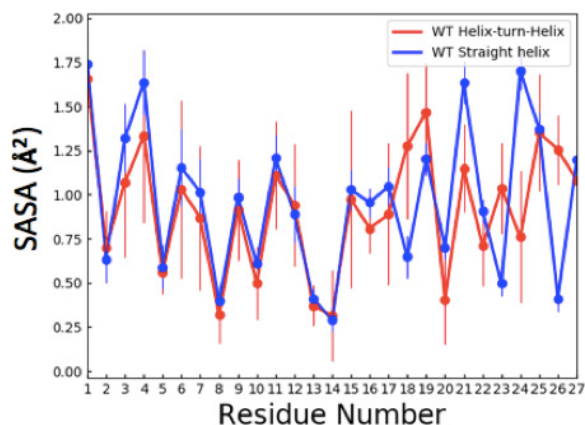


Figure 4.4. Solvent accessible surface area (SASA) of each residue as measured in our ME-REMD simulations when the SAA (1-27) fragment is in either a helix-turn-helix configuration (red) or in the elongated helix also seen in the PDB structure (blue).

In earlier work¹⁰ we have studied the segment SAA (1-13) which includes the helix Ia forming residues. These previous results established a conformational transition between an α -helix and a β hairpin that is associated with the dissolving and forming of salt bridges involving residues 1R, 5S, 9E, and 12D. Specifically, by comparing the wild type with suitable mutants we found that the salt bridge between 1R and 9E stabilizes the helix structure, as the helix propensity for this

fragment vanished in the mutant E9A. A corresponding analysis of the contact map of the wild type SAA (1-27) in Figure 4.5B also indicates that the N-terminal helix Ia is stabilized in the helix-turn-helix configuration by electrostatic interactions between the charged residues 1R and 9E, which are within 0.45 nm, and therefore have a high probability to form a salt bridge. We conjecture that the hydrophobic contacts between the helices Ia and Ib favor a helix hairpin; however, the propensities of residues 1-11 to form helix Ia are weak, and require an alignment of residues enforced by the salt bridge between residue 1R and 9E. Loss of this salt bridge leads to a transition of the helix-hairpin structure to configurations with helix Ib intact but the N-terminal residues forming a dynamically changing disordered elongated segment (see Figure 4.3c) with transient β -strands that is prone to aggregation.

Table 4.2. Percentage of population of different type of cluster for different systems

Systems	Helix-turn-helix	N-terminus dynamic	Straight Helix
ME-REMD Wild Type	33 (2) %	25 (9) %	5 (2) %
ME-REMD Mutation E9A	13 (1) %	49.7 (0.3) %	0

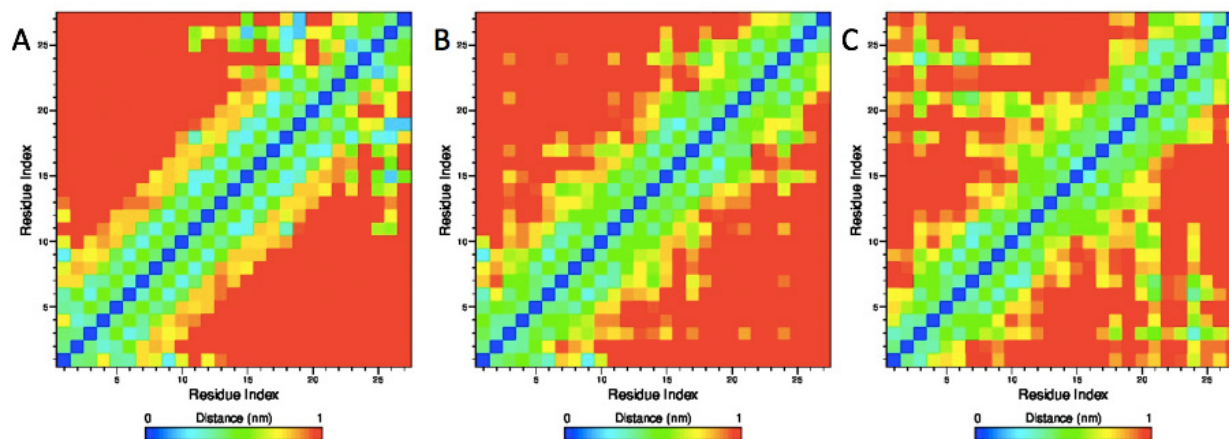


Figure 4.5. Sidechain-sidechain contact map as calculated from our ME-REMD simulations of the SAA (1-27) fragment. Shown are in A) the results obtained for the case that the wild type peptide assumes the extended helix motif seen in the PDB structure. Correspondingly, we show in B) this map for the case of the wild type peptide in the helix-turn-helix configuration. Note the strong signal for a salt bridge between residues 1R and 9E sidechain and the hydrophobic contacts between residues 3F, 21Y and 24M, missing in the PDB structure. By design is the salt bridge between residues 1 and 9 missing in the contact map of the E9A mutant in mutant in a the helix-turn-helix configuration, shown in C). Instead we find here a signal for a salt bridge between residue 1R and 26E, and for hydrophobic interaction between residue 3F and 21Y or 24M.

In order to test this hypothesis, we have performed additional ME-REMD simulations of the E9A mutant, using the same temperature distribution and number of replicas as for the wild type. Acceptance rates and tunneling times are also listed in table 1. Comparing the mutant with the wild type, we find indeed a lower propensity for helix formation in the first 11 residues, and an increase of turn formation to 40%, see the middle and right panel of figure 4.3. As a consequence, we observe only two dominant clusters for the E9A mutant, with the frequency of helix-hairpin configurations decreasing to 13 % compared with the 33 % in wild type, and on the contrary, the content of aggregation-prone configurations lacking helix Ia growing to 50 % compared with the 25 % in wild type, see Table 2. Hence, the replacement of a charged lysine by an alanine as residue 9, makes the N-terminal of SAA more flexible, increasing the population of aggregation-prone

configurations. However, the loss of the salt bridge between residues 1 and 9 does not completely destabilizes helix Ia. Instead this helix is still weakly stabilized by a salt bridge between the first residue 1R and residue 26E, and a larger number of hydrophobic contacts between residue 3F and 21Y or 24M, that together restrict fluctuations of helix hairpin configurations. These contacts can be seen in Figure 4.6C, and also the solvent accessible surface area(SASA) contributions of residues 3F,21Y and 24M. Specifically, the residue 3F has in the E9A mutant about 0.3 \AA^2 less surface exposed to solvent than in the wild type. The difference is with about 0.2 \AA^2 similar for residues 21Y and 24M.

4.4. Conclusions

In this study, we have used a variant of the replica exchange molecular dynamic, multi-exchange replica exchange molecular dynamic (ME-REMD) to investigate the conformational ensemble of the isolated N-terminal segment (1-27) of serum amyloid A, In the folded structure these residues form a single extended helix, but especially residues 1-11 are implicated in the amyloid formation of the overexpressed protein. Our first result is that ME-REMD is robust and depends little on the number of exchange attempts; i.e. the improved sampling efficiency is not bought by the need to adjust an additional parameter. The rather trivial modification of REMD leads to immediate gains of about 20% in sampling efficiency or reduction of required computational resources.

While a noteworthy result, our main interest is in the dynamics of the long helix formed by residue 1 to 27 in the crystal structure. This helix is not stable in our simulations of the isolated fragment SAA (1-27). Instead, the helix breaks up around residues 12 and 13 in a more than 50% of configurations. The resulting helix-hairpin is hold together by hydrophobic contacts between helix Ia (residues 1-11) and helix Ib (residues 14-27); however, this motif is itself not stable as helix Ia

is only weakly hold together by a salt bridge between residues 1R and 9E. Hence, in about 25% of configurations this helix Ia dissolves and the residues 1-11 form a dynamically changing elongated segment, with transient β -strand content that in a more dense environment would become the starting point for aggregation. Mutation of residue 9 from a glutamic acid to an alanine destroys the salt bridge, shifting the equilibrium away from the helix-hairpin toward a motif with increased β -strand content as also seen in our earlier simulations of the short fragment SAA(1-13).

We conjecture that a similar mechanism also applies for the full-length protein. While here the N-terminal helix I is likely stabilized by residues 28-76, we expect that it also exist in a dynamical equilibrium with an unstable helix-hairpin configuration where the first eleven residues may form transient β -strands that become the nucleus for the aggregation observed *in vivo* for overexpressed SAA. This scenario is supported by experimental observations showing a reduction of helicity at pH=2²¹, and an increase in fibril formation when the SAA(1-27) segment binds with acidic lysophospholipids instead of neutral lysophospholipids.¹⁸⁵ Since the wild type SAA(1-11) fragment has an isoelectric point of 6.0, while that of the E9A mutant is 9.8 (as calculated by the ExPASy Server¹⁸⁶⁻¹⁸⁸), the loss of charge at residue 9 due to the mutation E9A is similar to going in the wild type from neutral pH to acidic conditions. Hence, the experimental measurements are consistent with the mechanism that we have derived from molecular dynamics studies of a wild type and mutant SAA fragment.

4.5. Acknowledgments

Simulations were done on the SCHOONER cluster of the University of Oklahoma and the Extreme Science and Engineering Discovery Environment (XSEDE) which is supported by the National Science Foundation (NSF). Computing time was awarded under grant MCB160005. We acknowledge financial support from National Institutes of Health (NIH) under research grant R01GM120634.

Chapter 5. Cleavage, Downregulation, and Aggregation of Serum Amyloid A

The material in chapter 5 is adapted from Wang, W., Khatua, P., & Hansmann, U. H. (2020). Cleavage, Downregulation, and Aggregation of Serum Amyloid A. *The Journal of Physical Chemistry B*, 124(6), 1009-1019. The copyright permission is obtained from ACS, see details in Appendix IV.

Author Contributions: Wenhua Wang and Dr. Prabir Khatua contributed equally in this work.

5.1. Introduction

In order to function properly, a protein has to take a specific structure either by itself or in complex with other molecules. Misfolded proteins lose their function and are usually degraded in cells by proteolytic cleavage.¹⁸⁹ In general, much longer time scales are required for a competing process by which unfolded or misfolded proteins aggregate instead into assemblies characterized by a cross-beta structure termed as “amyloid”. While these amyloids sometimes have specific functions in organisms (for instance as storage of hormones¹⁹⁰ or as a matrix in bacterial biofilms¹⁹¹), their presence in humans and other mammals is more often the hallmark of neurodegenerative, metabolic, and other diseases.^{169, 192, 193}

These amyloid diseases are in some cases secondary illnesses. For instance, the 104-residue long serum amyloid A (SAA) protein plays a role in the transport of cholesterol in high-density lipoprotein (HDL) particles and is thought to be involved also in a role in the regulation of inflammation. In its active form, the protein assembles as a hexamer, built from two layers of trimers. The structure of the monomer has been resolved by X-ray crystallography and deposited

in the Protein Data Bank (PDBID: 4IP9). The four-helix bundle is shown in Figure 5.1a and consists of N-terminal helix-I formed by residues 1–27, helix-II by residues 32–47, helix-III by residues 50–69, and the C-terminal helix-IV by residues 73–88. The structure of SAA changes little when part of the hexamer (PDB-ID: 4IP8),²² where the chains in each trimer are packed together by the N-terminal helices, see Figure 5.1b, and cholesterol binds at the interface between the two trimers. Diseases such as colon cancer, inflammatory bowel disease, or rheumatoid arthritis can cause overexpression of SAA. The resulting serum concentration of about 1 mg/mL¹⁷⁵ is more than 1000 times higher than that in healthy persons and gives rise to the outbreak of SAA amyloidosis in some patients, which is characterized by the appearance of amyloid deposits, most commonly in liver, spleen, and kidney. As the deposits may interfere with the function of the affected organs, they add to the symptoms of the primary disease.¹⁹⁴ A drastic example is the failure of renal function and subsequent death in captive cheetahs caused by amyloid deposits, which themselves are because of overexpression of SAA as a consequence of stress-related inflammatory diseases.^{195, 196}

Because SAA amyloidosis is not observed in all patients with the primary disease, it cannot be caused solely by crowding resulting from the high concentration. More likely, it is a failure of a mechanism to downregulate the HDL transport and other functionalities after the SAA overexpression and/or to regulate the immune response to the primary disease.¹⁹⁷ However, this regulation mechanism is not fully understood. Binding with glycosaminoglycans (GAGs), such as heparin/heparan sulfate (HS), leads to dissociation of the hexamer.^{104, 198} This process should by itself foster downregulation of the SAA activity; however, the released 104-residue SAA_{1–104} monomers are in a second step and are cleaved into smaller fragments of 45 to 94 residues.¹⁹⁹ Most commonly found in amyloid deposits is the truncated fragment SAA_{1–76},¹⁰⁷ but the only resolved

fibril model is for the even smaller fragment SAA_{2–55} (PDB-ID: 6MST).³⁰ It is known from mutation experiments that the first eleven N-terminal residues are crucial for SAA misfolding and aggregation.²³ As part of a hexamer, in the native structure, this segment is cached in helix-I (residues 1–27) and stabilized by interactions with the neighboring chains; however, in isolated monomers and fragments, it may be flexible enough to form strand-like segments.²⁰⁰ This was observed in our lab during molecular dynamics simulations of the fragment SAA_{1–13}, where the segment alternated between an α -helix and a β -hairpin.²⁶ Nordling and Abraham-Nordling²⁴ have proposed that by taking such strand-like configurations, the N-terminal segment could nucleate fibril formation. Given the raised risk for amyloid formation (and subsequent amyloidosis), the question arises for what reason the full-length SAA proteins are cleaved into smaller fragments.

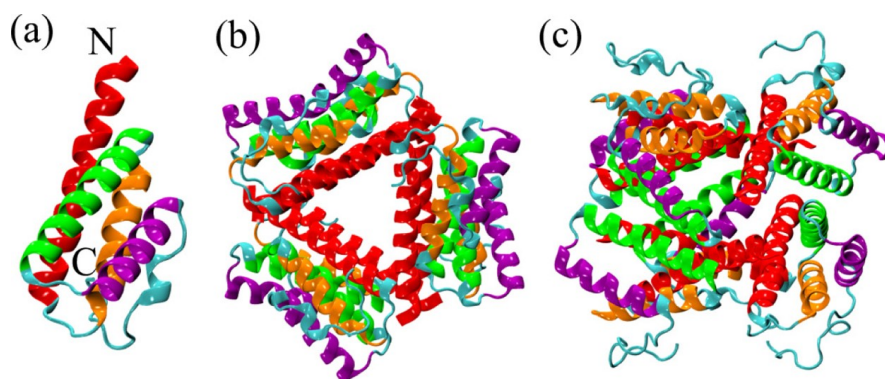


Figure 5.1. Crystal structure model of the full-sized SAA (SAA_{1–104}) (a) monomer (PDB-ID: 4IP9) and the hexamer (PDB-ID: 4IP8) as deposited in the Protein Data Bank. The hexamer is shown both in a top-down view (b) and a side view (c). The four helices are colored as follows: helix-I (red), helix-II (orange), helix-III (green), and helix-IV (magenta).

In the present chapter, we have explored this question using molecular dynamics simulations of the full-length SAA protein and various fragments, both as monomers and assembled into a hexamer. We establish that the cleavage contributes to the downregulation of the SAA activity by

shifting the equilibrium from hexamers toward monomers, thereby reducing the SAA concentration. We hypothesize that SAA₁₋₇₆ is the most commonly found fragment because unlike smaller or larger fragments, it allows switching between two distinct structures, enabling a fine-tuning of the response to the SAA overexpression. Dominant at neutral pH is a structure (coined by us helix-weakened) that allows for easy degradation; a process that helps to lower the SAA concentration quickly. However, the first eleven N-terminal residues also have an increased risk in helix-weakened configurations to unfold from helix-I and to form strand-like segments which in turn may nucleate amyloid growth. If acidic conditions raise the risk for aggregation and amyloid formation, the equilibrium shifts toward an alternative configuration (termed by us helix-broken), where the N-terminus is more stable but is more difficult to degrade. We speculate that, in most patients, where colon cancer, inflammatory bowel disease, or rheumatoid arthritis lead to overexpression of SAA, the described mechanism downregulates the activity and concentration of SAA, but if the switching mechanism is impeded or overwhelmed, an overabundance of the more aggregation-prone helix-weakened configurations gives rise to SAA amyloidosis.

5.2. Materials and Methods

5.2.1. Initial Conformations

For the full-sized SAA protein SAA₁₋₁₀₄, we use the crystal structures as start configurations in our simulations, derived by X-ray crystallography and deposited in the Protein Data Bank (PDB) under identifiers 4IP8 (monomer) and 4IP9 (hexamer). The initial configuration of the fragment SAA₁₋₇₆ monomer is derived by deleting the residues 77–104 from the crystal structure of the full-length SAA monomer. Similarly, the residues 77–104 are deleted for each of the six chains in the

crystal structure of the full-length hexamer to construct the start configuration of the truncated SAA₁₋₇₆ hexamer. Each of the two monomer models is placed in the center of a cubic box of edge length 6.8 nm and is filled with ~10,000 water molecules; while for two hexamer systems, the cubic box has an edge size of 9.8 nm and is filled with ~28,000 water molecules. In a similar way, we also generate from the SAA monomer two shorter fragments SAA₁₋₂₇ and SAA₁₋₄₇ that are put again into cubic boxes with edge sizes of 6.2 nm (SAA₁₋₂₇) and 6.8 nm (SAA₁₋₄₇), filled with ~7800 and ~10,000 water molecules, respectively. The above obtained configurations are then minimized by steepest descent (as implemented in GROMACS²⁰¹) before simulated by molecular dynamics for 20 ps, with positional restraints on the protein atoms. The relaxed structures obtained in this way are the start configurations for our main molecular dynamics runs.

5.2.2. Simulation Protocols

Our molecular dynamics simulations are performed by using the GROMACS 2018.2 software package.²⁰¹ Protein–protein and protein–water interactions are modeled with the CHARMM 36m force field²⁰² and the TIP3P¹⁵² water model, a combination that is known to perform well in simulations of the amyloid assembly.²⁰³ We use the LINCS algorithm¹⁵³ to constrain the bond length and the SETTLE algorithm¹⁵⁴ to maintain the water geometry. As periodic boundary conditions are selected, we use the particle mesh Ewald algorithm¹⁸¹ to calculate the electrostatic interaction with a cut-off distance of 1.2 nm, the default value in GROMACS for electrostatic and van der Waals interactions.

The molecular dynamics simulations are performed in an isothermal-isobaric (NPT) ensemble, with temperature set to 310 K using a v-rescale thermostat¹⁸⁰ and pressure set to 1 bar using a Parrinello–Rahman barostat.²⁰⁴ The equations of motions are integrated with a time step of 2 fs

using the leapfrog algorithm as implemented in GROMACS.²⁰¹ Each system is followed over three independent trajectories of either 1 μ s (monomers) or 500 ns (hexamers) duration, starting from distinct configurations derived from the above generated models by introducing a small random movement to the coordinates and randomizing the velocity distribution. Measurements are taken every 50 ps and stored for further analysis.

5.2.3. Observables

Time evolution of structures is followed by calculating the root-mean-square-deviation (rmsd) to the start configuration using our in-house code. Similarly, we measured the solvent accessible surface area (SASA) and the cavity diameter $\langle d_{\text{cavity}} \rangle$ with our in-house codes. The later quantity is approximated by averaging over the center-of-mass distances between the N-terminal helix-I segments of adjacent units of both layers. This approximation is justified as each of the two trimer layers (see Figure 5.1b) is formed from the helices in the respective chains through a circular head-to-tail packing, where helix-I remains within the interior cavity. Another measure for the similarity of a given configuration to the start configuration is the fraction of native contacts, which is defined as²⁰⁵

$$Q(X) = \frac{1}{N} \sum_{(i,j)} \frac{1}{1 + \exp[\beta(r_{ij}(X) - \lambda r_{ij}^0)]} \quad (5.1)$$

The sum runs over the N pairs of (i,j) nonhydrogen atoms i and j belonging to residues θ_i and θ_j that form a contact in the start configuration, that is, their distance is less than 4.5 Å in the start configuration. Note that $|\theta_i - \theta_j| > 3$. $r_{ij}(X)$ denotes the distance between the atoms i and j in

conformation X, while r_{ij}^0 represents this distance in the native state. β is a smoothing parameter taken to be 5 \AA^{-1} , and the factor λ accounts for the fluctuation when contact is formed, taken to be 1.8.

Correlations between contacts, defined here by the condition that the distance between two residues i and j (with $|i - j| > 3$) is less than 4.5 \AA , are quantified by the intermittent time correlation function (TCF), $C_{\text{contact}}(t)$, which is defined as^{206, 207}

$$C_{\text{contact}}(t) = \frac{\langle h(0)h(t) \rangle}{\langle h(0)h(0) \rangle} \quad (5.2)$$

Here, $h(t)$, a population variable, is set to one if a pair of residues is in contact at a particular time t and zero otherwise. We have also calculated similar TCFs for the helicity, $C_H(t)$, where the population variable $h(t)$ is now set to one if a residue is in helix at a particular time t and zero otherwise.

We define the cross correlation function between residues i and j , $C(i,j)$, as²⁰⁸⁻²¹⁰

$$C(i,j) = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{\langle \Delta r_i^2 \rangle^{\frac{1}{2}} \langle \Delta r_j^2 \rangle^{\frac{1}{2}}} \quad (5.3)$$

where angular brackets mark ensemble averages, Δr_i and Δr_j are the displacement vectors of the i -th and the j -th C_α atoms of the protein, respectively. $C(i,j)$ can vary by definition between +1 (complete correlated motion) and -1 (complete anti-correlated motion). Correlated residues move in the same direction, and anti-correlated residues move in the opposite direction.

The motion of secondary structure elements is quantified by measuring the dipole moments of various helices and comparing them with the ones observed in the start configurations. For this purpose, we define the dipole moment of a helix, μ , as

$$\mu = \sum_{i=1}^N (r_i - r_0) q_i \quad (5.4)$$

where r_i and r_0 represent the position vectors of the i -th backbone atom and the center of mass of the helix, respectively, while q_i is the partial charge of the respective backbone atom. Because the three helices differ in their number of residues, we have normalized the magnitude of the dipole moment vector by dividing it by the respective number of residues.

The stability of configurations is also evaluated by calculating the NMR N–H bond order parameter. Following Zhang and Brüschweiler,²¹¹ we define the order parameter for the i -th residue S_i^2 as

$$S_i^2 = \tanh \left(0.8 \sum_k [\exp(-r_{i-1,k}^O) + 0.8 \exp(-r_{i,k}^H)] \right) + b \quad (5.5)$$

where k runs over all the nonhydrogen atoms except those from residues i and $i-1$. r^H and r^O denote the distance from the nonhydrogen atom k to the amide hydrogen in residue i and carbonyl oxygen in residue $i-1$, respectively. The parameter value $b = -0.1$ is motivated by the observation that one finds usually for rigid protein regions, an order parameter value of around 0.9. Note that we have used the corrected version of S_i^2 ²¹¹ according to which the distances $r_{i-1,k}^O$ and $r_{i,k}^H$ in eq 5 should be shortened by 1.2 Å.

5.3. Results and Discussion

5.3.1. Hexamer

Recent experiments have established that the 104-residue serum amyloid A SAA₁₋₁₀₄ assembles as a hexamer in its biologically active state and forms a complex with HDL in blood serum. Dissociation of the hexamer, which is not amyloidogenic,²¹² is assisted by binding to GAGs, such as heparin/HS. The so-formed SAA monomers are in the second step cleaved by enzymes into shorter fragments with the 76residue fragment SAA₁₋₇₆, which is the most commonly found species. In principle, one can think of two reasons for the cleavage. First, cleavage may aid downregulation of the SAA activity by shifting the equilibrium toward monomers making a reassembly toward the biologically active hexamers less likely for SAA₁₋₇₆ than for the full-sized SAA₁₋₁₀₄. A second possibility is that the shorter fragments allow for a faster degradation, allowing in this way for rapid reduction of the SAA concentration.

In the present section, we focus on the role of the cleavage for the equilibrium between the hexamer and monomer exploring how SAA₁₋₁₀₄ and SAA₁₋₇₆ hexamers differ in their stability and decay dynamics. In order to ensure convergence, we have monitored the time evolution of the rmsd of the hexamer SAA₁₋₇₆ and SAA₁₋₁₀₄ with respect to their start configurations, taking into account all the nonhydrogen atoms in residues 1–76. This choice allows us to compare rmsd values for the two systems despite their unequal length. Our results are shown in Figure 5.2 and demonstrate that both systems equilibrate after approximately 200 ns. For this reason, we consider only the last 300 ns of the 500 ns long trajectories for further analysis. Snapshots of the final configurations of each run are shown in Supporting Information Figure S5.1 and help to visualize the structural changes seen in each run. In order to test whether the elevated rmsd values of the SAA₁₋₇₆ hexamer in

relation to the SAA₁₋₁₀₄ hexamer are indeed markers for differing thermodynamic stability, we have performed additional molecular dynamics simulations of the two hexamers at elevated temperatures of 325, 350, 375, and 400 K. Set-up and simulation protocol are analog to the ones described in Section 2 and trajectories are followed over 350 ns. The corresponding time evolutions of rmsd are shown in Figure 5.2c,d. As expected, the rmsd values rise in both systems faster and higher with increasing temperature. However, while in the case of SAA₁₋₁₀₄, the final rmsd values vary within 3–4.5 Å for temperatures up to 375 K and a value of 5 Å is already reached at 325 K for SAA₁₋₇₆. Both this difference in final rmsd values and their growth rate with time indicate a lower thermal stability of the SAA₁₋₇₆ hexamer compared to the SAA₁₋₁₀₄ hexamer.

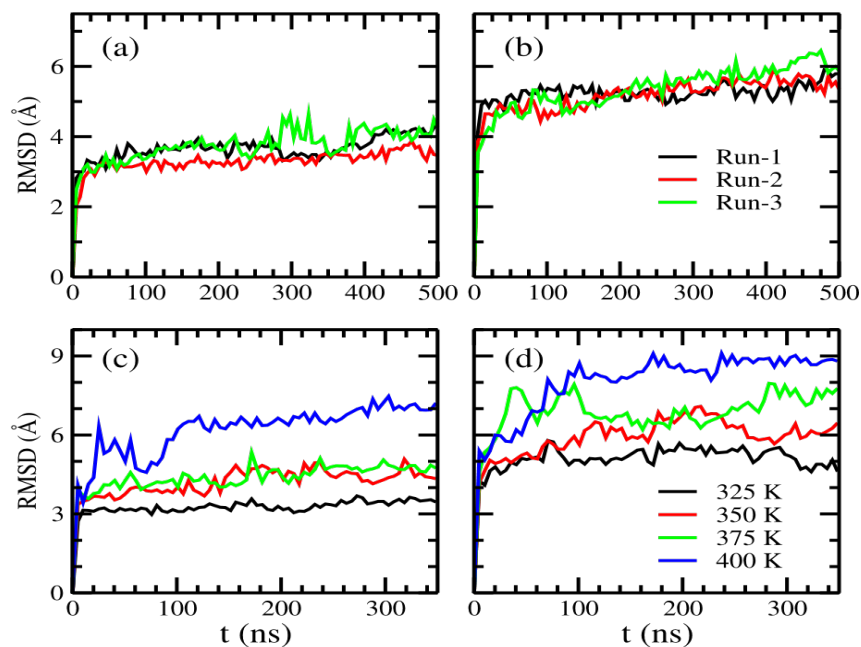


Figure 5.2. Time evolutions of rmsd for all three trajectories of (a) SAA₁₋₁₀₄ and (b) SAA₁₋₇₆ hexamers. The rmsd values are calculated with respect to the start hexamer configuration considering all the nonhydrogen atoms in residues 1–76 in each of the six chains. The average rmsd values as function of time for different temperatures are shown in (c) for SAA₁₋₁₀₄ and (d) SAA₁₋₇₆ hexamers.

In order to find the origin of the lower stability of the SAA₁₋₇₆ hexamer, we have calculated the average cavity diameter $\langle d_{\text{cavity}} \rangle$ and the SASA of both hexamers. While the cavity diameter is similar ($24.5 \pm 0.2 \text{ \AA}$ for SAA₁₋₇₆ and $24.3 \pm 0.7 \text{ \AA}$ for SAA₁₋₁₀₄), individual residues are more exposed to the solvent for SAA₁₋₇₆, leading to the larger SASA values ($55.5 (0.3) \text{ \AA}^2$ compared to $51.2 (0.4) \text{ \AA}^2$ for SAA₁₋₁₀₄, see Supporting Information Table S5.1). This higher solvent exposure suggests that the lower SAA₁₋₇₆ hexamer stability, leading to the higher rmsd values seen in Figure 5.2, is caused by increased flexibility of individual residues in the six chains.

This higher flexibility of residues in the SAA₁₋₇₆ hexamer results from the missing favorable interchain hydrogen bonds, salt-bridges, and hydrophobic interactions that stabilize the SAA₁₋₁₀₄ crystal structure.⁷ This can be seen in Figure 5.3, where we show the (a) fraction of all native contacts and the (b) same quantity restricted to intrachain native contacts. Corresponding to the increase in rmsd, the fraction of native contacts decreases with time for both hexamers, with the loss of native contacts more pronounced for the SAA₁₋₇₆ hexamer than for the SAA₁₋₁₀₄ hexamer. As in both cases, only contacts formed by the first 76 residues are considered; it follows that the higher stability of the SAA₁₋₁₀₄ hexamer is not caused by the additional contacts formed by the C-terminal tail.

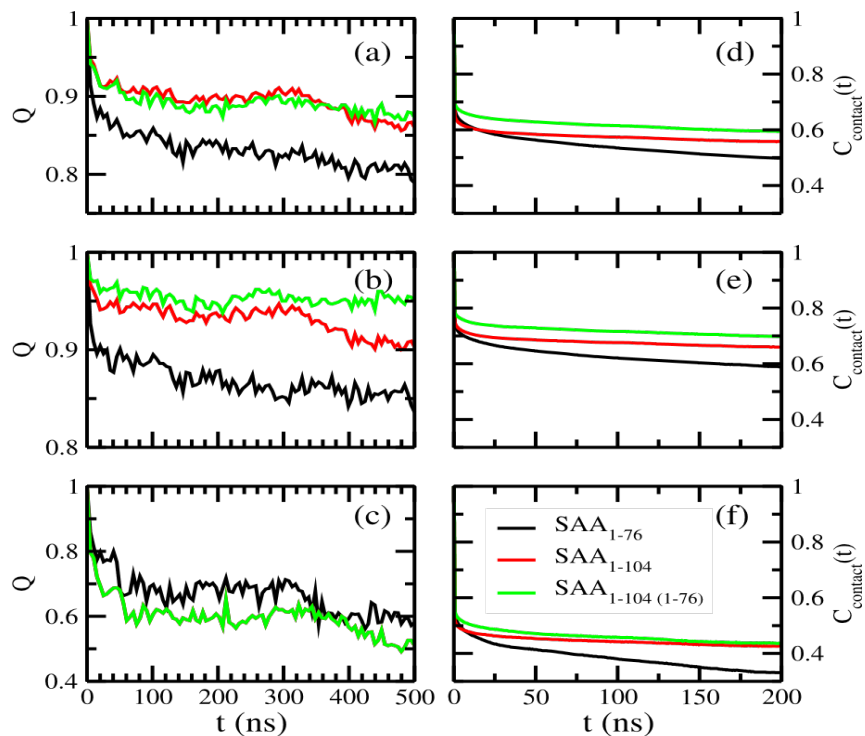


Figure 5.3. Time evolutions of the fraction of native contacts (Q) are shown in the left column, considering either (a) all native contacts, (b) only intrachain native contacts, or (c) only interchain native contacts. Data are from the first trajectory of either the SAA_{1-76} or the SAA_{1-104} hexamer simulation. The right column presents plots of the intermittent contact TCF, $C_{\text{contact}}(t)$, considering either (d) all contacts, (e) only intrachain contacts, or (f) only interchain contacts, formed in SAA_{1-76} and SAA_{1-104} hexamers. Data are from all three trajectories for each system, but only the last 300 ns is considered for calculating $C_{\text{contact}}(t)$. For comparison, we show also for the full-sized SAA_{1-104} hexamer the corresponding values for the case when only contacts between residues 1 and 76 are taken into account.

Surprisingly, the fraction of interchain contacts, as shown in Figure 5.3c, is higher for SAA_{1-76} . This on first look unexpected result is likely caused by the higher flexibility of the SAA_{1-76} chains, which allow them to form interchain contacts more easily. However, these interchain contacts are transient and do not contribute to the stability of the hexamer. This can be seen from Figure 5.3d–f, where we show TCFs of the contacts, taking into account either all contacts or considering only either interchain or intrachain contacts. Irrespective of the type of contacts, TCFs for the SAA_{1-76} hexamer decay fast and monotonically, whereas for the SAA_{1-104} hexamer the decay is slower and

quickly approaches a plateau. This is also the case when for the later only residues 1–76 are considered. The decay time is especially short for the interchain contacts in the SAA_{1–76} hexamer, demonstrating the short life times and the transient nature of these contacts. Hence, the additional interchain contacts can only partially compensate for the loss of stability resulting from the reduced number of intrachain contacts as their lifetimes are short, and overall, the total fraction of native contacts is lower for SAA_{1–76} hexamers than for SAA_{1–104} hexamers. Hence, the intrachain contacts rather than the interchain contacts determine the overall stability of the SAA hexamers.

The above discussion implies that when part of the hexamer, the intrinsic stability of the folded SAA_{1–104} chains is higher than for the SAA_{1–76} chains. This is supported by Figure 5.4a, where we show the residue-wise NMR order parameter (S_2) of the first 76 residues of both proteins, averaged over all six chains in a hexamer and all three trajectories. The lower the value of S_2 , the higher will be the flexibility of the respective N–H bond and hence the corresponding residue and vice versa. Comparing the two proteins, we find a signal for increased flexibility of SAA_{1–76} chains in two regions: one made of residues 30–43, and the other of residues 63–76. These residues have a higher probability to interact with the solvent, starting in this way the dissociation process of the hexamer. Figure 5.4b confirms indeed that the SASA of these residues, belonging to either helix-II or helix-III, is higher for SAA_{1–76} than for the full-sized SAA_{1–104}, where the C-terminal tail (including helix-IV) protects these residues from being exposed to the solvent.

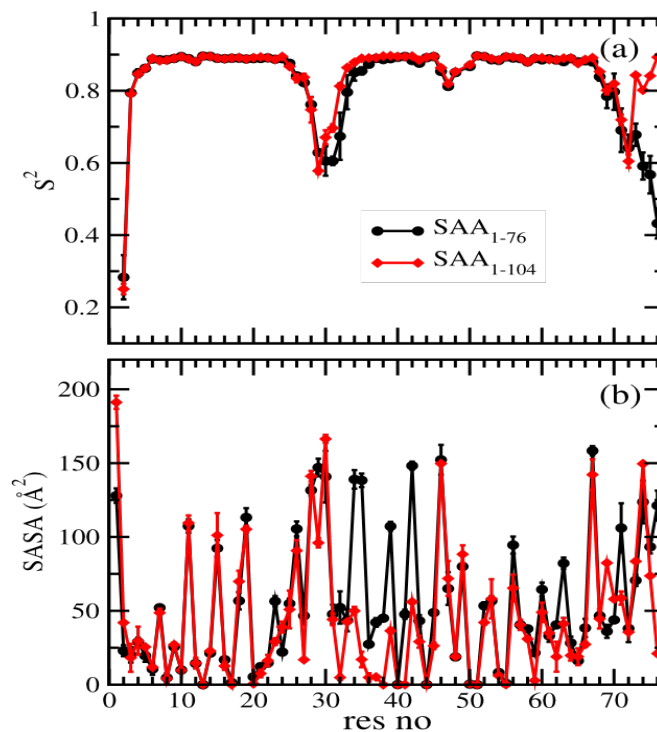


Figure 5.4. Residue-wise (a) backbone N–H order parameter (S^2) and (b) SASA for the SAA₁₋₇₆ and SAA₁₋₁₀₄ hexamers as calculated over the last 300 ns of all three trajectories. Data are shown only for the first 76 residues.

5.3.2. Monomer

In the abovementioned section, we have demonstrated that hexamers formed by SAA1–104 have a higher stability than those formed from the fragment SAA1–76. Hence, it is unlikely that SAA1–76 and similar fragments, once generated by enzymatic cleavage from the full sized protein, will reassemble into a hexamer and if formed by chance, such a hexamer would decay quickly. However, it is not clear whether the sole purpose of the cleavage is inhibiting reassembly into the biologically active hexamer. Another reason for the cleavage could be to decrease the SAA concentration by easing degradation of the protein. We have already shown in the previous section that within the hexamer, the SAA1–76 chains have lower internal stability than full-length SAA1–104 chains. A similar increased flexibility of the isolated chains may allow for the

structural changes of fragments that could encourage proteolysis. However, the higher flexibility would also increase the risk of aggregation as it could lead to release of the first eleven residues from helix-I. These N-terminal residues are considered to be crucial for amyloid formation. Hence, in order to probe the effect of the cleavage on stability and potential amyloid-forming tendencies of SAA monomers, we have also studied the isolated monomers of the full-length SAA1–104 and the SAA1–76 fragment in another set of molecular dynamic simulations.

Similar to the hexamer, we first establish what part of the simulated trajectories can be used for analysis. For this purpose, we monitor for the two monomers the time evolution of rmsd with respect to their start configurations considering again only the nonhydrogen atoms in the first 76 residues. As for the hexamers, we show the last snapshots of each runs in Supporting Information, Figure SF2, to highlight the structural changes of the monomers. Our results, as shown in Supporting Information, Figure S5.3, are similar to the hexamers in which the changes of rmsd are smaller for SAA1–104 monomers (around 2 Å) than for the SAA1–76 monomers (around 6–12 Å), which again have more pronounced fluctuations. However, as a plateau is approached for both molecules and the simulations converge in about 500 ns, 500 ns of the 1 μ s-long trajectories remain for our analysis.

The stability differences between SAA1–76 and SAA1–104 monomers, seen in the time evolution of rmsd, are further investigated by comparing the average number of native contacts ($\langle N_{nat} \rangle$), residue–residue contacts ($\langle N_c \rangle$), and helicity ($\langle h \rangle$). In Table 5.1, we list the differences of these values to the ones measured in the respective start configurations. For the full-sized protein, we have also calculated these values restricted to the first 76 residues and have added the values in the table to allow for a better comparison with the fragment SAA1–76. The twice smaller values

of $\langle N_{\text{nat}} \rangle$ for SAA1–76 monomers (in relation to full-length monomers) confirm the larger deviations, and in turn, lower stability of the truncated fragment. This difference is also seen when only the first 76 residues of the full-sized protein are considered. Similarly, the loss of helicity $\langle h \rangle$ is larger in the fragment than in the full-sized protein. On the other hand, the total number of contacts $\langle N_{\text{c}} \rangle$ is approximately the same in the truncated fragment and between residues 1–76 of the full-sized protein, indicating that the fragment not only loses helical contacts but the structure of the fragment changes and new contacts are formed.

Table 5.1. Difference of Various Quantities with Respect to the Start Configuration, Measured in All Three Trajectories of SAA1–76 a and SAA1–104 Chains in the Hexamer and in Isolated Monomers

system		hexamer				monomer			
		$\langle \Delta \text{SASA} \rangle$	$\langle \Delta N_{\text{nat}} \rangle$	$\langle \Delta N_{\text{non}} \rangle$	$\langle \Delta h \rangle$	$\langle \Delta \text{SASA} \rangle$	$\langle \Delta N_{\text{nat}} \rangle$	$\langle \Delta N_{\text{non}} \rangle$	$\langle \Delta h \rangle$
SAA1–76	run-1	-3.3	-369	295	-3	-0.6	-391	324	-10
	run-2	-3.8	-366	266	-3	-5.0	-385	424	-4
	run-3	-3.3	-378	258	-4	0.9	-461	380	-7
	average	-3.5 (0.3)	-371 (6)	273 (19)	-3 (1)	-2.3 (2.8)	-412 (42)	376 (50)	-7 (3)
SAA1–104 (1–76)	run-1	-1.9	-285	212	-3	-1.5	-291	248	-3
	run-2	-1.4	-296	210	-2	-1.6	-258	261	-3
	run-3	-1.4	-275	221	-3	-1.8	-298	267	-3
	average	-1.6 (0.3)	-285 (11)	214 (6)	-3 (1)	-1.6 (0.1)	-282 (21)	259 (10)	-3 (0)
SAA1–104	run-1	1.5	-553	403	-2	4.1	-615	417	-3
	run-2	1.1	-546	397	-2	3.4	-589	453	-3
	run-3	1.9	-545	410	-2	4.6	-643	425	-4
	average	1.5 (0.4)	-548 (4)	403 (7)	-2 (1)	4.1 (0.6)	-616 (27)	432 (19)	-3 (1)

^aDifferences calculated for the solvent accessible surface area $\langle \Delta \text{SASA} \rangle$ per residue (in \AA^2), number of native contacts ($\langle \Delta N_{\text{nat}} \rangle$), number of nonnative contacts ($\langle \Delta N_{\text{non}} \rangle$), and helicity ($\langle \Delta h \rangle$). For comparison, we also present data for SAA1–104 chains considering only the first 76 residues. Averages are taken over the three trajectories with the standard deviations listed in parenthesis.

For the full-sized protein, the loss of native contacts is mainly in the C-terminal region, with only few (about 10%) of native contacts within residues 1–76 lost in the first 1 μs . This demonstrates the protective role of the C-terminal segment of residues 77–104 and indicates that after cleavage, the fragment SAA1–76 quickly loses its tertiary structure. This picture is supported by the side-

chain–side-chain contact maps, as shown in Figure 5.5, which shows that absence of the C-terminal tail leads SAA1–76 to lose the contacts between residues 70–76 and 60–69, that is, the cleavage increases the flexibility of helix-III. In turn, contacts between residues 32–38 and 65–67, located on helix-II and helix-III, are also dissolved. A similar loss of interhelix contacts is seen between helix-I (residues 26–28) and helix-II (residues 32–34). This reduction in the number of contacts increases the residue-wise SASA in the fragment, see also Supporting Information Figure S5.4. Similar to that of the hexamer, most of the residues present in helix-II of the truncated SAA monomer become more exposed toward the solvent. The difference in SASA between the fragment and the full-sized monomer is the largest for the residues 30–40, belonging to helix-II, and about 10% larger in the monomer than in the hexamer. This SASA difference results from the weaker contacts of helix-II with the adjunct helices I and III in the fragment. This is consistent with the fact that, as shown in Supporting Information Figure S5.2, helix-II loses the contacts with both helix-I and helix-III that are present in the native structure. However, the SASA values of the residues 57, 58, 61, 64, 65, 68, and 69 on helix-III, which in the native structure form contacts with helix-II, do not increase after losing the contacts. Instead, the SASA values even decrease, indicating that these hydrophobic residues have formed alternative contacts, leading to a structural rearrangement of the remaining three helices.

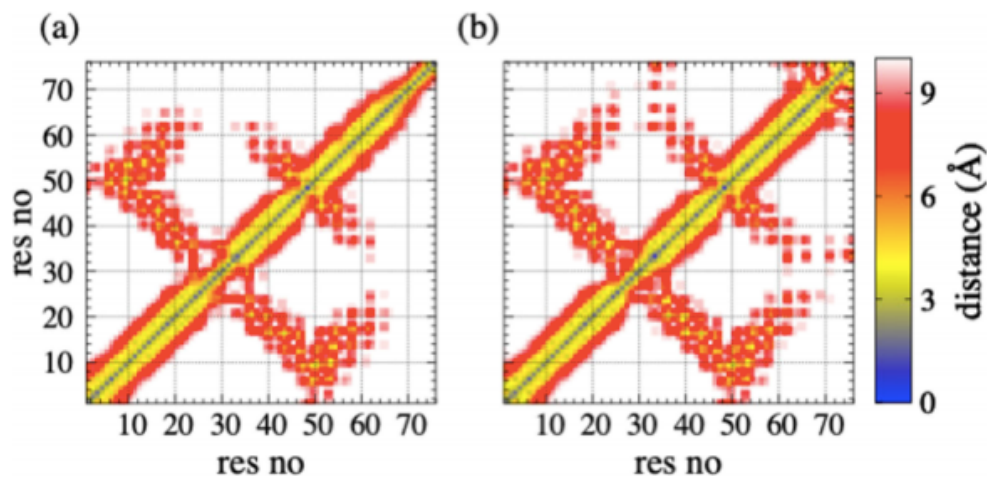


Figure 5.5. Side-chain–side-chain contact map of (a) SAA₁₋₇₆ and (b) SAA₁₋₁₀₄ monomers as calculated over the last 500 ns of all three trajectories of each system. Data are shown only for the first 76 residues. Residue pairs whose average contact distance is more than 10 Å are excluded.

Hence, cleavage of the C-terminal tail destabilizes the native structure of SAA monomers through reducing interhelix contacts and increased solvent exposure. As soon as (after release from the hexamer) the full-length SAA monomers are cleaved and lose the C-terminal tail (including helix-IV), their native structure decays, and the chains are unlikely to re-associate into a hexamer. On the other hand, unfolding of the full-length monomer structure (SAA₁₋₁₀₄) happens, if at all, only on much longer time scales. These longer lifetimes allow the full-length protein, unlike the fragment, to reassemble into a hexamer.

However, the reduced stability of the fragment SAA₁₋₇₆ affects also the N-terminal eleven residues, which are known to be critical for SAA amyloid formation. The strength of this effect

depends on the length of the cleaved fragment. This can be seen in Figure 5.6, where we show for SAA fragments of various lengths the residue-wise NMR order parameter (S^2) for the backbone N–H bond vectors (eq 5). The flexibility of the first eleven N-terminal residues increases drastically with subsequent cleavage of the helices, facilitating more misfolding into an aggregation prone configuration. A special role here seems to fall to the segment SAA1–76, which is in between the two extreme cases of the N-terminal segment fully caged in helix-I (the full-sized protein made of all four helices) or of the N-terminal segment being highly flexible and stabilized only by the environment of helix-I (SAA1–27) and helix-II (SAA1–47). This intermediate position of the SAA1–76 fragment, therefore, indicates that helix-III takes a special role in the misfolding and aggregation of SAA.

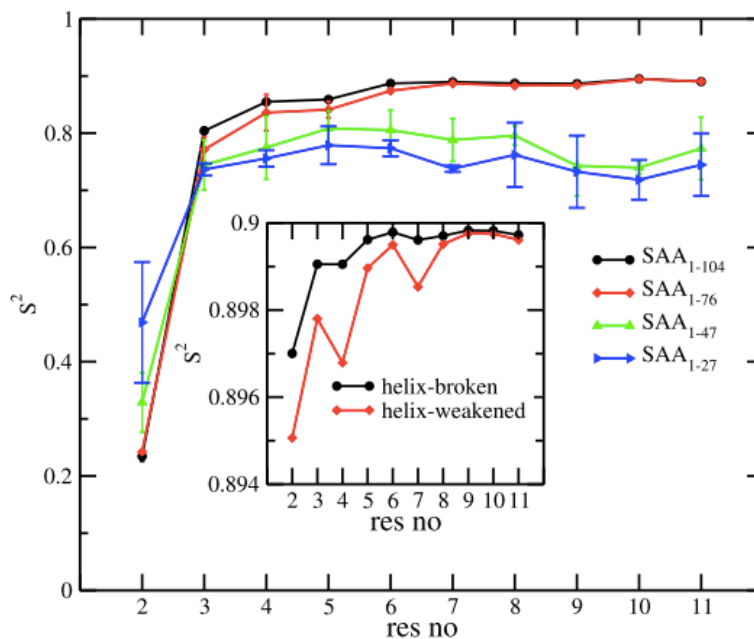


Figure 5.6. Residue-wise backbone N–H order parameter (S^2) for the first eleven residues of different SAA protein monomer fragments, (1– 104), (1–76), (1–47), and (1–27) as calculated over the last 500 ns of all three trajectory of each system. For the fragment SAA_{1–76}, we distinguish further between helix-broken and helix-weakened structures (see text). Residue-wise backbone N–H order parameter (S^2) for the first eleven residues corresponding to these two structures of SAA_{1–76} monomers is shown in the inset.

In order to understand this special role of helix-III, we have calculated for each helix the corresponding dipole moment. The time evolutions of the per residue dipole moment (μ) for all three helices are displayed in Figure 5.7a–c, where we depict for this quantity, the magnitude of the difference between the actual value and the one measured in the start configuration. While the dipole moments of helix-I and helix-II change little with time, that is, when not affected by the cleavage of the C-terminal residues, there is a clear signal for helix-III. The change observed here in the dipole moment could indicate either breakage of helix-III into two shorter segments connected by a kink (named by us a helix-broken configuration) or a shortening of helix-III with the released residues taken random orientations (called by us a helix-weakened configuration). Visual inspection shows that both possibilities happen: the helix-broken case in run 2 and the helix-weakened one in run 1 and run 3. The relative orientation of helix-III with respect to the other two helices changes in both motifs, as seen from the relative orientation between pairs of helices in Figure 5.7d–f and the center-of mass distances between these pairs in Figure 5.7g–i. This is particularly noticeable for the helix-broken case (Figure 5.7e,h).

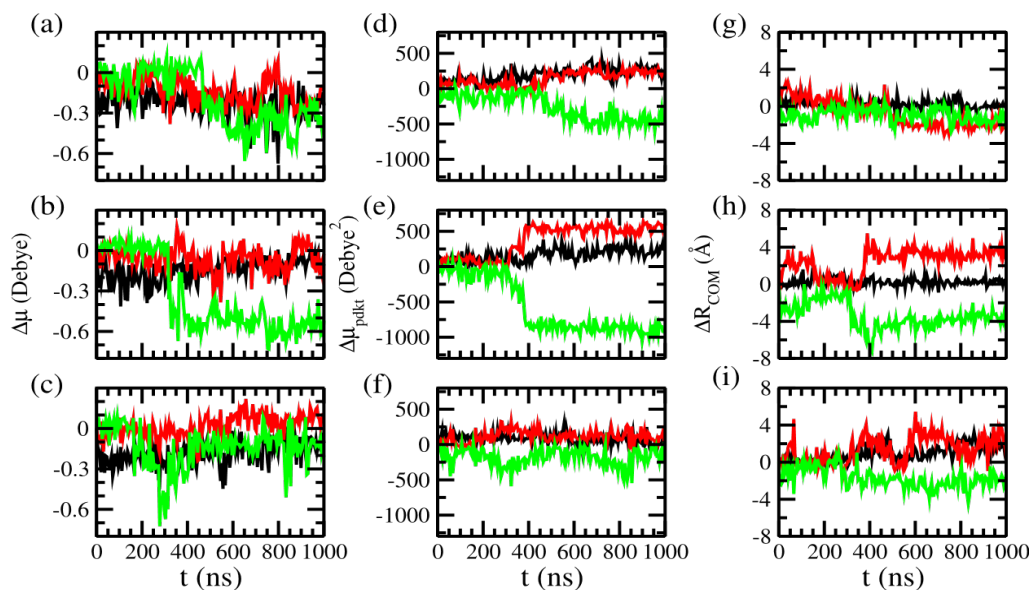


Figure 5.7. (a–c) Time evolution of the difference in dipole moments (μ) of a helix at a given time minus the corresponding value in the start configuration. Shown in the figure are the magnitudes of this difference divided by the number of residues for helix-I, helix-II, and helix-III in the SAA1–76 fragment. The color code is as follows: helix-I is shown in black, helix-II in red, and helix-III in green. In (d–f) we show the corresponding time evolution of the scalar product of dipole moment vectors and in (g–i) center of mass distances between pairs of helices. The color coding is as follows: values for helix pair (I,II) are displayed in black, for helix pair (II,III) in red, and for helix pair (I,III) in green. The results for run-1 is shown in the first row, for run-2 in the middle row, and for run-3 in the last row.

In order to go beyond visual inspection, we introduce the following definition for the two motifs.

In helix-weakened configurations, helix-III is still preserved for residues 50–62, but at most three residues are still helical in the C-terminus of helix III (residues 63–69). On the other hand, a kink is formed within residues 50–62 in case of helix-broken configuration (i.e., the helicity of this segment is less than 13) and the C-terminus of helix-III is preserved (at least five residues are helical in the segment 63–69). In both cases, the orientation of helix-III changes toward helix-I with the $C\alpha$ residue distance between residue 1 and 69 (${}^dC\alpha$) less than 20 Å in case of helix-broken configuration and more than 20 Å in the helix-weakened case.

Our above results show that upon cleavage of the C-terminal residues 77–104, the position of helix-III is no longer restrained by contacts with these residues but can now form contacts with helix-I. If helix-III moves toward the C-terminal end of helix I, the lack of helix-stabilizing contacts will lead to a shortening of the helix and a disordered segment at the C-terminus becomes oriented toward the C-terminus of helix-I. On the other hand, if helix-III moves toward the N-terminus of helix-I, helix-III breaks up into two shorter pieces connected by a kink around the residues 55–58, with the C-terminal end of helix-III now pointing toward the first eleven N-terminal residues of helix-I. We believe that the contrasting realignment of helix-I and helix III allows for the helix-weakened motif release the N-terminal residues 1–11 from helix I (which then may refold into a β hairpin as needed for attachment of other chains and starting amyloid formation), but the helix-broken motif protects these residues from being released from helix-I. This assumption is supported by Figure 5.8, where we show how rapidly the contact number and helicity measured for the first eleven N-terminal residues change with time. This speed of change is quantified by the intermittent TCF of the two quantities, calculated for the helix-weakened case from the last 500 ns of run 1 and 3, and the helix-broken case from the last 500 ns of run 2. For comparison, we show values obtained for the full-sized SAA1–104 monomer. While for the helix-weakened configuration both the helicity and the number of contacts of the N-terminal segment decay faster than in the more stable full-sized SAA1–104 monomer, there is no qualitative difference seen between the helix-broken configuration and the full-sized protein. Hence, our results suggest a higher chance for the release of the first eleven N-terminal residues from helix-I in helix-weakened configurations than in helix-broken configurations, where the segment is stabilized by contacts with the C-terminus of helix-III.

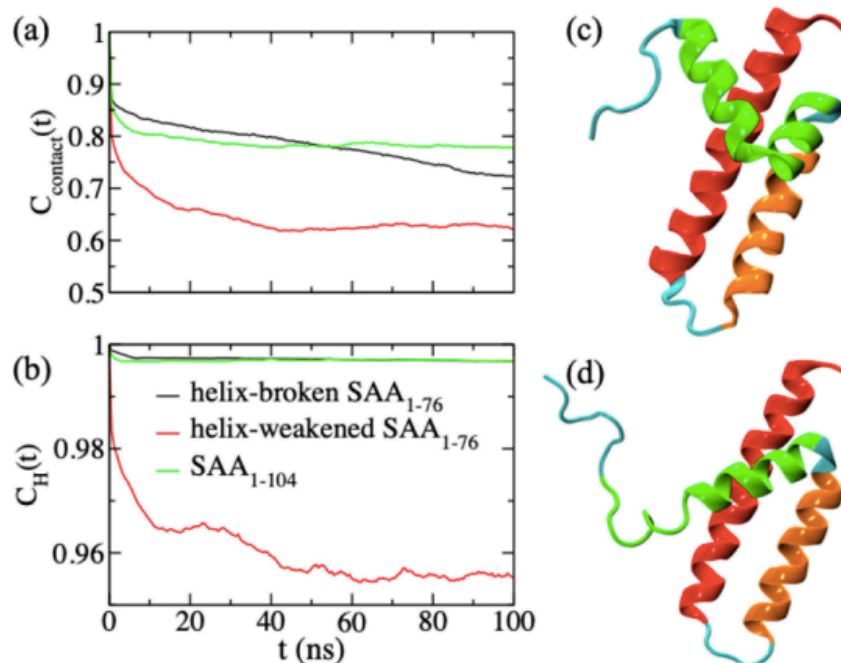


Figure 5.8. Intermittent TCF for (a) contact $C_{\text{contact}}(t)$ and (b) helicity $C_H(t)$, where contacts and helicity are calculated for the first eleven N-terminal residues of helix-I of either helix-broken or helix-weakened SAA1–76 conformations. As a reference, data for SAA1–104 monomers are also shown. Representative configurations for the (c) helix-broken and (d) helix-weakened SAA1–76 structures are presented in the right subfigures. The color scheme of the three helices, helix-I to III, is same as in Figure 1.

In order to probe in more detail how the relative movement of helix-III in the two motifs affects residues in other parts of the molecule, especially, the first eleven N-terminal residues, we show in Figure 5.9a–d, respectively, for both motifs the contact map and cross correlation function, $C(i,j)$, between the residues. In the helix-weakened structure the C-terminal residues of helix III form contacts with the C-terminal residues of helix-I. Formation of these contacts is by a correlated motion of the C-terminal of helix-III (residues 58–69) with the C-terminal of helix-I (residues 19–27), and a consequently anticorrelated motion with N-terminal of helix-I (residues 1–11), driving them in opposite directions. This motion is reversed in the helix-broken structure, where these contacts are absent and instead the C-terminus of helix-III forms contacts with the N-terminal

residues of helix-I (here primarily the first eleven residues). These additional contacts, which do not exist in the helix-weakened structure, stabilize helix-I in the helix-broken structure, that is, prevent release of the first eleven residues and their misfolding into an amyloid-prone configuration. The motif is further stabilized by a reorientation of the hydrophilic residues on helix-I, which in the native structure point toward helix-II but in the helix-broken motif face outward to the solvent. On the other hand, in the helix-weakened motif, the hydrophilic residues stay oriented toward helix-II, but the hydrophobic residues now face toward the solvent, which in the native structure is the case for only 30% of the hydrophobic residues on helix-I. Note that exposure of such large hydrophobic patches as seen on helix-I in the helix-weakened structure often serves as a signal for activating degradation by the proteasome.

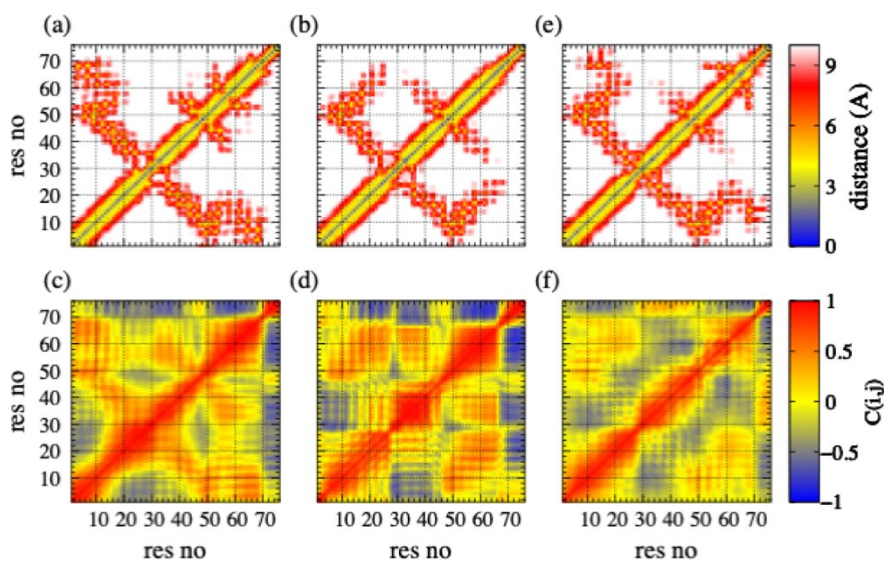


Figure 5.9. Side-chain–side-chain contact map of (a) helix-broken and (b) helix-weakened SAA1–76 monomer configurations at neutral pH, showing the distance between pairs of residues. The corresponding two-dimensional dynamic cross-correlation map $C(i,j)$ of these pairs are shown for helix-broken configurations in (c) and for helix-weakened configurations in (d). For the later motif, we also show in (e,f) corresponding figures obtained from simulations under acidic conditions (mimicking a $\text{pH} = 4$).

The already existing stable contacts between the C-terminal end of helix-I and helix-II in the helix-weakened structure restrict the possibility of formation of new contacts between the C-terminal end of helix-III and the C-terminal end of helix-I. This also simultaneously restricts the formation of salt-bridges between the opposite charged residues or hydrogen bonds between the side chains. On the contrary, repulsion between the similar charged residues on the C-terminal ends of helix-III and helix-I will further destabilize the overall structure. The only reasonably stable contacts that we found in helix-weakened structures are Y21–R62, M24–I65, and R25–I65. However, these contacts are neither sufficient to compensate the charge–charge repulsion, and the helix-weakened structure is flexible, with frequent transitions toward the original arrangement of helix-I and helix-III. The net-effect is a loss of the forces that stabilize helix-I in the native structure of the full-sized SAA1–104 monomer but are missing in the shorter segments. The resulting higher entropy of helix-I leads to its de-stabilization, especially, at the N-terminus as the C-terminus is still supported by contacts with helix-III. As a consequence, there is an increased chance for the first eleven residues to be released from helix-I and to misfold into strand-like configuration (see inset of Figure 5.6) that may start the aggregation process.

On the other hand, in the helix-broken structures, strong contacts are formed between residues F4–F68, L7–I65, and F11–F69. The phenyl rings of the pair of Phe residues that is, (F4, F68), and (F11, F69) orient in such a way that they remain one upon another with an interplanar angle of $\sim 0^\circ$, resulting in strong π – π stacking interactions of the phenyl rings. The result is a helix–helix linkage between the C-terminal end of helix-III and the N-terminal end of helix-I, which not only

stabilizes the overall structure but, especially, the N-terminus, keeping the first eleven residues cached in helix-I and therefore reducing the probability of their misfolding and subsequent aggregation.

Hence, based on our observations, we propose the following mechanism for tuning the SAA concentration. After the release of SAA1–104 chains from the hexamer enzymatic cleavage, it leads to SAA1–76 monomers that are unstable as they lack the helix–helix linkage interaction between helix-III and helix-IV. The C-terminal of helix-III becomes exposed, and helix-III can now move relative to helix-I. This process is facilitated by formation of transient contacts between residues 20–21 (in helix-I) and 61–62 (in helix-III), which now permit movement of the C-terminus of helix-III toward either the C-terminus or N-terminus of helix-I, that is, enabling transitions between the two resulting structures, coined by us helix-weakened and helix-broken structures. With only three runs, we do not have sufficient statistics to quantify the relative frequency of the two motifs, but our simulations indicate that after cleavage of the full-length protein, most SAA1–76 monomers evolve into a helix-weakened configuration.

We conjecture that in helix-weakened configurations the large exposed hydrophobic patches on helix-I will activate further proteolytic degradation, likely by protease cathepsin D41 after binding with the heat shock protein Hsp70²¹³⁻²¹⁵ recruited in response to the primary disease. The net-effect would be a reduction in the SAA concentration. However, the larger flexibility of the N-terminal segment of the first eleven residues also increases the risk that these residues unfold from helix-I and take strand-like configurations which could nucleate amyloid formation. Under neutral pH, the helical conformation of the segment is stabilized through a transient salt bridge between

residues 1R and 9E. This salt bridge (defined by us by the condition that the center-of-mass distance between the ammonium or carboxylate groups of residues 1R and E9 is below 4 Å) is formed in about 70% of the full-length SAA1–104 monomers but only in about 25% of the helix-weakened SAA1–76 configurations. On the other hand, in helix-broken configurations, this salt bridge is found with similar frequency as in the full-length protein, and the N-terminal segment is stabilized additionally by interactions with the C-terminal of helix-III. Hence, the helix-broken configurations are less aggregation prone than helix-weakened configurations. We hypothesize that this difference is minor under neutral conditions making helix-weakened configurations more desirable than helix-broken configurations, which do not have large hydrophobic patches exposed to the solvent and therefore are more difficult to degrade.

The situation may be different under acidic conditions such as seen in conjunction with cancer or inflammatory diseases,^{216, 217} where this salt bridge between residues 1R and 9E can no longer be formed. In preliminary simulations, where we approximate acidic conditions by fixing the protonation of charged residues to states given by the known pK_a value of the free residue and the selected pH-value, the salt bridge is found at pH = 5 in only about 40% for helix-broken and about 15% for helix-weakened configurations and at pH = 4 it is less than 5% of the SAA1–76 configurations. The disappearance of the salt bridge will likely increase the chance for aggregation and amyloid formation with the risk mitigated in helix-broken configurations by the additional contacts between the N-terminal segment and the C-terminus of helix-III. Interestingly, we also observe in the same set of simulations a shift from helix-weakened to helix-broken configurations, when going to acidic conditions, see Figure 5.9e,f. This transition between the two forms, therefore, counteracts the increased danger of unfolding of the N-terminal segment by a structural conversion

that stabilizes this segment through additional contacts with helix-III. Hence, this pH-driven transition between helix-weakened and helix-broken configurations appears to be a mechanism to counteract the increased chance for aggregation and amyloid formation, otherwise often seen under acidic conditions.^{216, 217} Unfortunately, it would not make sense to quantify the free energy differences between the two forms as function of pH, as our approximate treatment of pH limits the reliability of our measured data. For a more quantitative treatment, one would need to utilize the constant pH techniques^{218, 219} as applied to amyloids in ref ²²⁰.

5.4. Conclusion

Various diseases cause overexpression of the SAA protein leading in some, but not all, cases to amyloidosis as a secondary disease. The response to overexpression involves dissociation of the SAA hexamer and subsequent enzymatic cleavage of the full-length SAA₁₋₁₀₄ chains into shorter fragments most commonly the 76-residue fragment SAA₁₋₇₆. Analyzing extensive molecular dynamics simulations, we propose a mechanism to downregulate the SAA activity and concentration that relies on this cleavage and is sketched in Figure 5.10. As SAA in its functional form assembles into hexamers, we have first tested the hypothesis that the cleavage shifts the equilibrium for SAA₁₋₇₆ fragments from the biologically active hexamers to potentially amyloidogenic monomers. Our molecular dynamics simulations confirm that hexamers built from full-length SAA₁₋₁₀₄ chains are indeed more stable than such formed from SAA₁₋₇₆ fragments. We explain this lower stability with the larger exposure of helix-II and helix-III in the SAA₁₋₇₆ hexamer chains. This lower stability reduces the chance for the SAA₁₋₇₆ fragments to reassemble into a hexamer, and if formed by chance, such a hexamer would decay quickly.

Our hexamer simulations also suggested a higher internal stability of the full-length SAA₁₋₁₀₄ chains compared with the SAA₁₋₇₆ fragments. As the reduced stability of the fragment likely triggers further degradation, reducing SAA concentration, we have added molecular dynamics simulations of isolated monomers to probe also the lower stability of the SAA₁₋₇₆ fragments. As the first eleven residues are crucial for amyloid formation,²³ we were especially interested in the question of how the cleavage affects the stability of this N-terminal segment which is part of helix-I (residues 1–27) in the native structure and protected by the C-terminal tail (including helix-IV). On the other hand, in SAA fragments too short to contain helix-III (residues 50–69), helix-I is less protected, and only a transient salt bridge between residues 1R and 9E stabilizes the helical configuration of the N-terminal segment.²⁰⁰ Especially under acidic conditions, where this salt bridge cannot be formed, the N-terminal residues may unfold to form strand-like segments^{26, 200} and to nucleate aggregation.

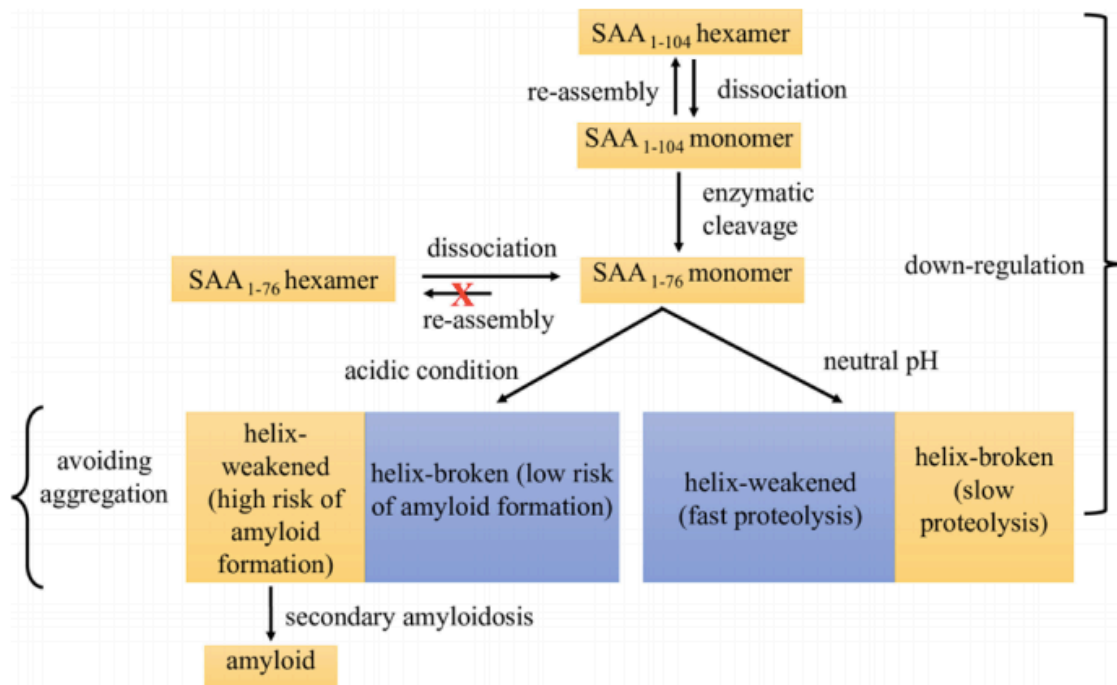


Figure 5.10. Sketch of the proposed mechanism that balances the desire for the downregulating SAA activity and concentration with the need to avoid harmful aggregation.

SAA₁₋₇₆ fragments are on the cusp between structures, where the N-terminal residues are firmly integrated in helix-I (as in the full-length protein) and such where they are flexible enough to be released easily from helix-I, see Figure 5.6. This is because interactions with helix-III may either stabilize or de-stabilize helix-I. Because the fragment SAA₁₋₇₆ lacks the helix-helix linkage interactions between helix-III and helix-IV, seen in the full-length protein, the C-terminus of helix-III is able to move toward either the C-terminus of helix-I or toward the N-terminus of helix-I. The first and more common case leads to helix-weakened structures characterized by a weakened helix-III causing large exposed hydrophobic patches on helix-I that will trigger further degradation by the proteasome, therefore, reducing the SAA concentration and downregulating the SAA activity. However, this motif is also characterized by a reduced stability of helix-I which raises the probability for a release of the aggregation-prone first eleven residues, providing a potential start

point for subsequent amyloid formation. This is especially the case under acidic conditions, where the transient salt bridge 1R–9E cannot be formed that stabilizes the helical conformation of the N-terminal segment under neutral conditions. On the other hand, in the helix-broken structures, newly formed contacts between the C-terminus of helix-III and the N-terminus of helix-I stabilize the latter helix, reducing the probability that the aggregation-prone first eleven residues are released from helix I. Hence, the increased risk for aggregation and amyloid formation associated with acidic conditions is mitigated by switching from helix-weakened to helix-broken configurations, a process that we have observed in preliminary simulations designed to mimic an acidic environment. The possibility for such transitions may be the reason why cleavage of the full-length SAA protein leads most often to SAA_{1–76} fragments, which by switching between helix-weakened and helix-broken configurations can optimize the chance for degradation while minimizing the risk of aggregation and amyloid formation. In most patients where colon cancer, inflammatory bowel disease, or rheumatoid arthritis leads to overexpression of SAA, the described mechanism enables downregulation of the activity and concentration of SAA. We speculate that SAA amyloidosis indicates failure of this switching mechanism.

5.5. Acknowledgments

The simulations in this work were done using the SCHOONER cluster of the University of Oklahoma and XSEDE resources allocated under grant MCB160005 (National Science Foundation). We acknowledge financial support from the National Institutes of Health under grant GM120634.

Chapter 6. Stability Study of A Recently Found Human SAA Fibril Structure

The material in chapter 6 is work that was unpublished at the time this dissertation was written, only preliminary data are included in this chapter. A manuscript is in preparation.

Author Contributions: This work is fully credited to Wenhua Wang.

6.1. Introduction

Most of my doctoral research has focused on identifying the factors that accelerate SAA fibril formation, such as enzyme cleavage, low pH condition, presenting of glycosaminoglycans (GAGs). Our previous results showed that enzyme cleavage can shift the equilibrium to monomer state. The cleavage always happens in the C-terminal region of SAA and causes significantly increasing the flexibility of monomer. The most common fragment of SAA cleavage product is SAA₁₋₇₆ (contains only the first 76 residues). The fragments can downregulate and protect from aggregation by switching the conformation between two identified motifs. Many previous studies have revealed that amyloid fibrils are more stable under acidic condition, especially for SAA fibril, the early stage of fibril formation happens in lysosomes, the pH condition in lysosome is acidic (pH = ~ 4.3), SAA is endocytosed. Then its C-terminus is truncated, N-terminus is misfolded by the lysosome/cell, intracellular oligomers appear, and disrupt the lysosome and cell membrane to get into the extracellular environment and assemble into fibril structure.²²¹ This proposed SAA fibril formation mechanism indicated that low pH is critical for early-stage fibril formation. Our previous studies for SAA shorter fragments and full-length protein also shown that acidic

conditions can shift the equilibrium to a fibril formation favorable state.^{110, 200} Although key interactions stabilizing the amyloidogenic region in native structure have been identified and a potential mechanism for SAA cleavage, downregulation and aggregation has been proposed, many questions still remain unclear, such as, what interactions are critical in SAA fibril formation, what are the driving forces to stabilize SAA fibril.

The kinetics of fibril formation is difficult to probe by both experimental or computational studies. Computationally more accessible are thermodynamic stability measurements of fibrils. Such a stability analysis requires a structural model of the SAA fibrils, which only recently was resolved for human and mouse SAA by electron cryo-microscopy.³⁰ Because of experimental difficulties is the human SAA fibril model only available for the fragment SAA₂₋₆₉, and here only residues 2-55 have been resolved while residues 56-69 are disordered. However, as SAA₁₋₇₆ (and not SAA₁₋₅₅) is the most commonly found fragment in human SAA fibrils, it appears to be likely that the disordered region, likely extending to residue 76, plays also a role in fibril formation.³⁰

The ordered part of the human SAA fibril adopts a right-hand twist, with each chain forming seven β -sheets, β 1- β 7. The N-terminal region is made of about 20 residues and takes a similar structure in both human and mouse SAA fibrils. It is stabilized by hydrophobic interaction involving six hydrophobic residues (F3, F5, F10, M16, Y17, and A19). The C-terminal region of human SAA fibril is stabilized by the ionic interactions among the charged residues (D23, R25, E26, K34, D43, and R47) buried inside a C-terminal cavity. Two proto-fibrils are packed together in human SAA fibrils, with the packing interface stabilized by ionic, hydrophobic, and polar cross-stack interactions. The interface is a so-called “steric zipper”, formed by the packing of β 3-strands (residues Y29 and I30) of chains on opposite proto-fibrils. A salt-bridge interaction between the

N-terminus of chains in one layer of a protofibril with residue D33 of a chain of the same layer in the opposite proto-fibril is also stabilizing the packing interface.³⁰ Compared with the mouse SAA fibril is the human SAA fibril packing much tighter, making us wonder whether packing interactions between two chains may lead to a stable nucleus, i.e., whether fibril formation starts with nucleation and elongation of single protofibrils, or whether both protofibrils grow out of dimer kept together by packing interactions. Hence, the first question that we pose is that for the smallest stable fibril fragment, i.e., the nucleus which starts fibril growth, and whether this fragment is stabilized by stacking or packing. As it is known that acid conditions further SAA fibril formation, we ask as a second question how acidic conditions affect the fibril stability. Finally, in our third line of research we study the role of the first eleven residues (amyloidogenic region) and residues 56-76 (disordered region) for fibril stability.

From our molecular dynamics simulations we conclude that the critical size for fibril stability is a two-fold two-layer tetramer. We conjecture that stacking within the same fold precedes packing between two folds. Acidic conditions are most important at early stages of fibril formation. By altering the salt-bridge network for the C-terminal cavity, they encourage the stacking of chains. Our results further suggests that SAA fibril formation is initially dominated by a kinetically favorable process where N-terminal regions of different chains interact with each other and form anchors which allow the chains to stack together. On the other hand, we do not find indications that the disordered C-terminal tail alters the stability of the SAA fibril geometry. This observation is consisted with C-terminal region (residues 65-76) turned into disorder state to release N-terminal region (residues 1-11) for further proteolysis/aggregation in our previous study of SAA monomer fragment SAA₁₋₇₆.

6.2. Materials and Methods

Models used in our simulation are generated from the cryo-EM structure as deposited in the Protein Data Bank under identifier (PDB ID: 6MST). This model is for the fragment SAA(2-55) i.e., without the first residue which we also did not consider in our investigation. The model is a two-fold six-layer system, from which we successively deleted chains in the bottom and top end layer to obtain a sequence of models with two-fold three-layers (F2L3), two-fold two-layers (F2L2), and two-fold one-layer(F2L1). By deleting one of the folds, corresponding single-fold models are generated, i.e., one-fold three-layers (F1L3), one-fold two-layers (F1L2), and one-fold one-layer (F1L1) systems. In order to test the role of the first eleven N-terminal residues we generated also a series of truncated models (F1L1-,F1L2-,F1L3-,F2L1- and F2L2-) where these residues were deleted from the chains. In another series of models (F1L1A, F1L2A, F1L3A, F2L1A, F2L2A, F2L3A) we accounted for the effect of acid conditions in an approximated way by altering the protonation states of residues E9, E26, H37 to the one expected at pH=4 while it corresponds in the parent models to pH=7. Each of the single-fold models is placed in a cubic box of edge size ~ 9.6 nm and filled with ~ 28000 water molecules, while for two-fold models the box has an edge size of ~ 10.4 nm and is filled with ~ 37000 water molecules. Since most fibrils taken from patients are made from the larger SAA(1-76) fragments, we have in addition generated one model where the residues 56-76 were added. These residues could not be resolved in the cryo-EM structure indicating that that this segment is unstructured. We generated this specific model only after we had determined the two-fold two-layer systems as the smallest stable system. Hence, our extended model F2L2+ is also a two-fold two-layer system, with the additional residues 56-76 added in a coil configuration. The extended model is again placed in a cubic box of edge size of

~ 17.5 nm that is filled with ~ 74000 water molecules. All considered models are listed in Table 6.1.

Table 6.1. List of considered models.

System index	pH value	N-terminal truncation	C-terminal extension	Folds	Layers
1	7	No	No	One	One
2	7	No	No	One	Two
3	7	No	No	Two	One
4	7	No	No	Two	Two
5	7	No	No	One	Three
6	7	No	No	Two	Three
7	4	No	No	One	One
8	4	No	No	One	Two
9	4	No	No	Two	One
10	4	No	No	Two	Two
11	4	No	No	One	Three
12	4	No	No	Two	Three
13	7	Yes	No	One	One
14	7	Yes	No	One	Two
15	7	Yes	No	Two	One
16	7	Yes	No	Two	Two
17	7	Yes	No	One	Three
18	7	No	Yes	Two	Two

Our simulations rely on the GROMACS 2018.2 software package.²⁰¹ We use the CHARMM36m force field to describe interactions between atoms, and chose TIP3P for modeling water.^{152, 202} The SETTLE and LINCS algorithms are used for water geometry and bond length constraint.^{153,}¹⁵⁴ Particle mesh Ewald algorithm (cut off 1.2 nm) is applied to calculate electrostatic interactions for periodic boundary conditions. The same cutoff is used in calculations of van der Waals interactions. Molecular dynamics simulations are done in an isothermal-isobaric (NPT) ensemble,¹⁸¹ with the temperature set to 310 K by a v-rescale thermostat and pressure to 1 Bar by a Parrinello–Rahman barostat.^{180, 204} Three independent trajectories of 100ns length are generated for each system. Each of the three trajectories starts with an initial configurations derived from the above generated models by small random variations of coordinates and the velocity distribution. Data are collected and stored every 50 ps for later analysis. Only the last 70 ns of the trajectories are used for our stability analysis which relies on GROMACS built-in tools and in-house code for measuring quantities such as RMSD, RMSF, Contact map, etc.²⁰¹

6.3. Results and Discussions

6.3.1. Critical size of SAA fibril and potential amyloid formation mechanism

In order to test the stability of the recently resolved SAA fibril structure, we start by determining the critical size of the fibril, that is the size above which the fibril is stable over extended periods of time while the fibril will be decaying on the time scale of our simulation below this size. We show in Figure 6.1 root-mean-square deviations of backbone atoms to the start configuration, and root-mean-square fluctuations of the residues, with the values derived by averaging over all chains, and the highest value of all three trajectories shown. We find that at neutral pH, the one-fold one-

layer F1L1, i.e., the SAA₂₋₅₅ monomer, and the two-fold one-layer model F2L1 are unstable, which is consistent with visual inspection of the final configurations, see for instance the representative structures in fig 6.3. b and e. Hence, our data exclude not only SAA₂₋₅₅ monomers as possible nucleus for fibril formation, but also dimers made of two chains packed together. On the other hand, the stability of the models increases while the number of layers increase. Note especially that the one-fold two-layer model F1L2, i.e., a dimer made of two chains stacked on each other, has higher stability than the two-fold one-layer model F2L1, i.e., the dimer where the two chains are in the same plane stabilized by packing interactions. This can be seen from both the variation in RMSD in fig 6.2 and visual inspection, see the representative structure in fig 6.3. f. Hence, while this dimer is still not stable enough to serve as nucleus for fibril growth, it may be a metastable state on the way to forming the two-fold two layers system F2L2, the smallest assembly seen in our simulation that is stable enough to serve as a nucleus for fibril elongation.

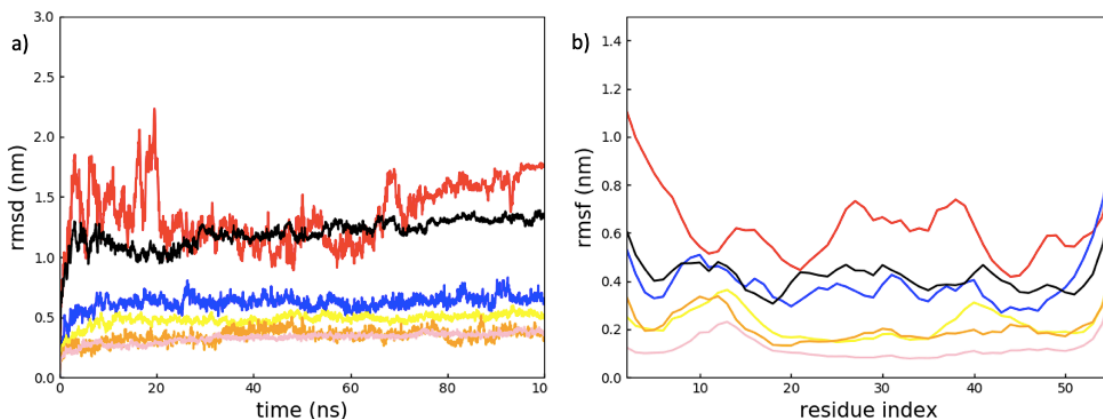


Figure 6.1. Under pH = 7 condition. a). Average RMSD per chain (overall lowest value among three trails are chosen to plot) for different systems, red, F1L1 system, blue, F1L2 system, black, F2L1 system, yellow, F2L2 system, orange, F1L3 system, pink, F2L3 system. b). Average RMSF value for different systems, the color scheme is the same as a).

Hence, comparing the stability of the various studied models suggest that at neutral pH the formation of the critical nucleus starts with two monomers forming mainchain hydrogen bonds and stacking into a metastable dimer, with each chain assuming a fibril-like configurations. Packing of two such dimers with their N-terminal region leads to a tetramer where the chain configurations are sufficiently stabilized by the interactions between the chains, and this two-fold two-layers tetramer then serves as a nucleus for further fibril growth, see fig 6.3. c.

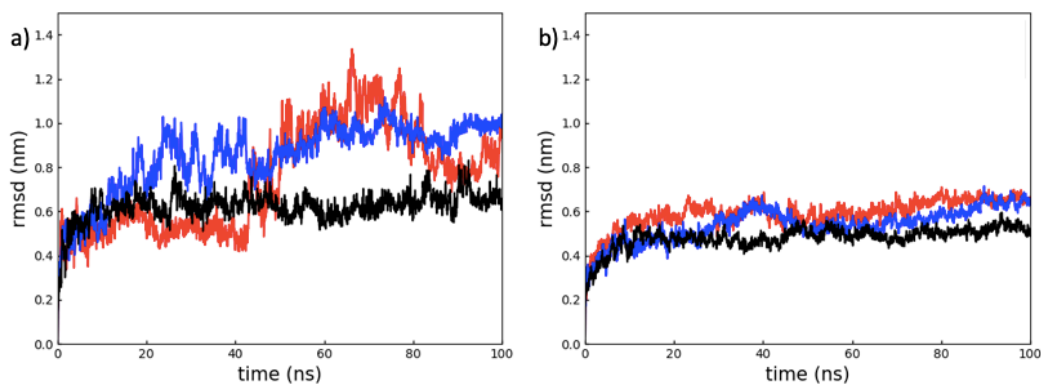


Figure 6.2. Average RMSD per chain for three trails of pH = 7 systems, a). F1L2 system. b). F2L2 system. Red, blue and black stand for trial 1 to 3.

The critical size determined by our simulation is also consist with previous measurements of the Gibbs free energy needed for dissociation of two protofibrils. Using PDBePISA it was found that a protofibril has to have at least two layers in order for this dissociation energy to be positive.³⁰ With growing number of layers decrease the RMSD and RMSF values of the various systems studied in our simulations. The differences between the observed stabilities are small, and systems larger than the critical size preserve their fibril structure on the time scale of our simulations, i.e., can be considered to be stable., see also the representative final configurations in fig 6.3. d and g.

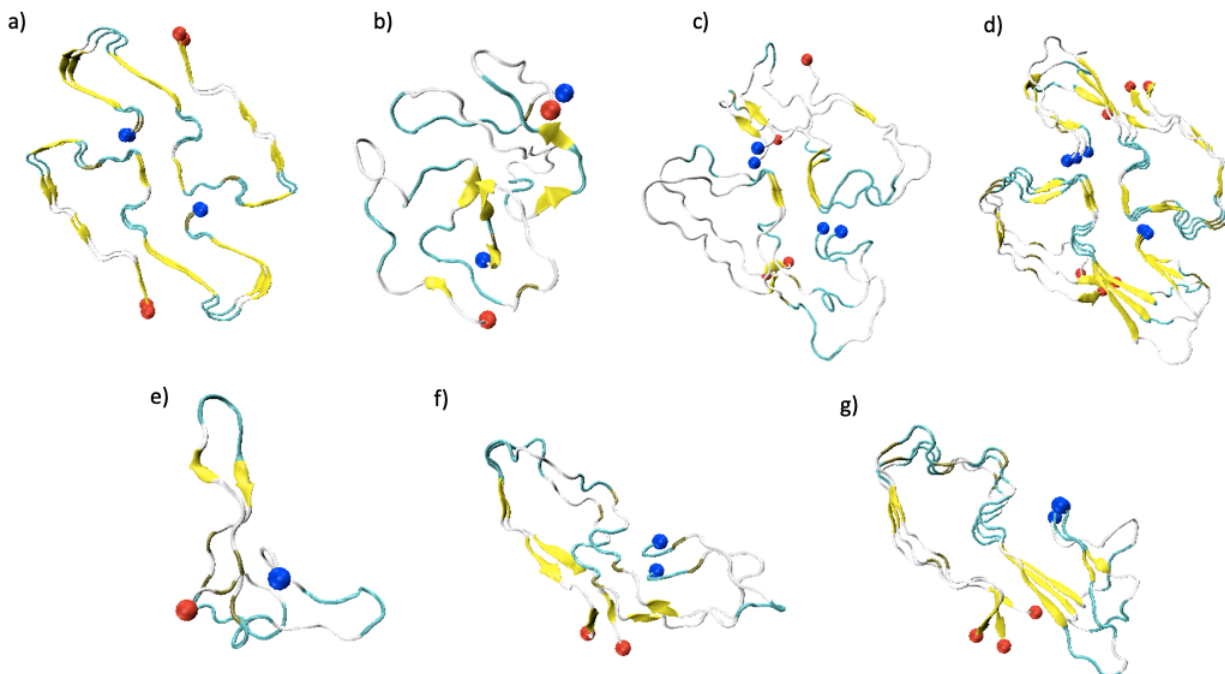


Figure 6.3. At pH 7, a). Initial structure (F2L2 system). Representative structure of b). F2L1, c). F2L2, d). F2L3, e). F1L1, f). F1L2, g). F1L3 systems. Red and Blue balls are N-terminus and C-terminus. Yellow, cyan and gray stand for β -strand, turn and coil.

6.3.2. Acidic conditions enhance early stage aggregation

The situation is different under acidic conditions at pH 4, which we have simulated in a simplistic way by protonating the sidechains of Glutamic acid and Histidine. While the general patterns resemble that seen in the simulations done under neutral pH conditions, the RMSF for the one-fold two layers system F1L2A (fig 6.4a) is smaller than the one for the corresponding system at pH 7, namely F1L2. Differences are especially seen for residues 12 – 37 and for the C-terminal cavity region former by residues 45 – 55. The lower root-mean-square fluctuations imply a higher stability of this assembly than seen under neutral condition. These differences in RMSF and stability are not seen when comparing the two-fold two-layer F2L2A tetramer with the corresponding system at neutral pH, F2L2, see (fig 6.4b). Hence, while acidic conditions do not

raise the stability of fully-formed fibrils, see fig 6.4 c-d, we conjecture that they increase the probability for fibril formation by making it easier to form a nucleus. Under acidic conditions leads the stacking of two chains to a dimer that is more stable than it is at pH=7. The resulting longer lifetime makes it more likely for such a dimer to find a partner and assemble to the F2L2 tetramer that is the critical nucleus.

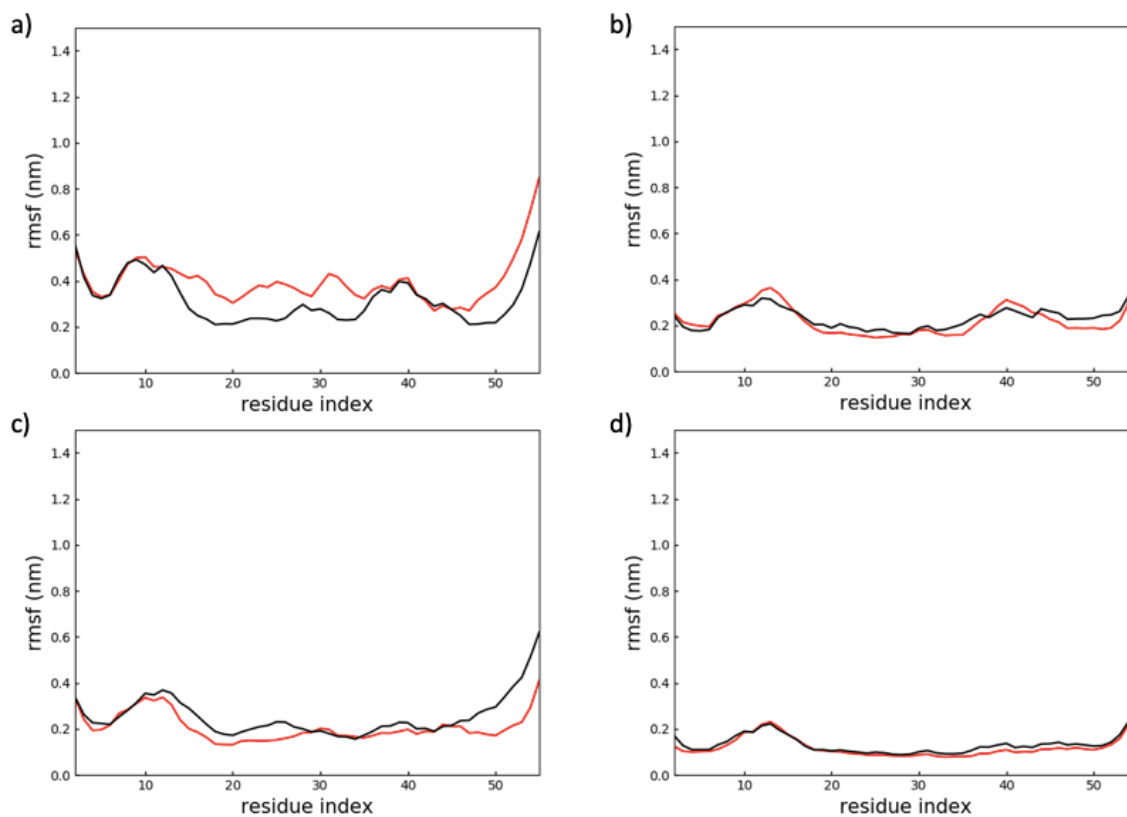


Figure 6.4. Average RMSF value comparison for same system size under different pH conditions, red, pH = 7, black, pH = 4. a). F1L2 systems. b). F2L2 systems. c). F1L3 systems. d). F2L3 systems

What causes the additional stabilization of the F1L2 dimer at acid conditions? From fig 6.5 a)-b) one sees that at neutral pH only a few long-range side-chain contacts (between F3 and M24)

contribute to stability, while under acidic conditions additional contacts are found inside the C-terminal cavity, such as between residue E26 and group K46-R47-G48, and between group R25-E26, K34 and D43. Note especially the role of residue E26, which becomes protonated at pH=4, allowing it to form multiple sidechain hydrogen bonds and salt-bridges in the C-terminal cavity (residues 23 - 51) which further stabilize the meta-stable dimer. With the changing of protonation state, instead of salt-bridge interactions between E26 and K34 or E26 and K46 at pH 7 condition, E26 forms sidechain interaction with D43 at pH=4 condition. The shift of E26 hydrogen bond makes the interaction between R25 and D43 possible at pH 4 condition. Furthermore, and both D23 and R25, R47 and D43 can form strong sidechain interactions with occupancy of 49.6 % and 27.2 % respectively under pH 4 condition. At pH 4 condition, interactions between R25 and D43, E26 and D43 can connect group D23-R25-E26 and group R47-K46-D43, and along with the connection with group Y35-K34-D33-Y29, a stronger salt-bridge/hydrogen bond network forms in C-terminal cavity. At pH 7 condition, only and group R47-K46-D43 and group Y35-K34-D33-Y29 are connected, and E26 is only weakly (7.1 % occupancy) connected to group R47-K46-D43, see fig 6.6. for details. Hence, the different pH conditions alter the intrachain sidechain hydrogen bond networks in C-terminal cavity increasing the structural stability of the chain folds.

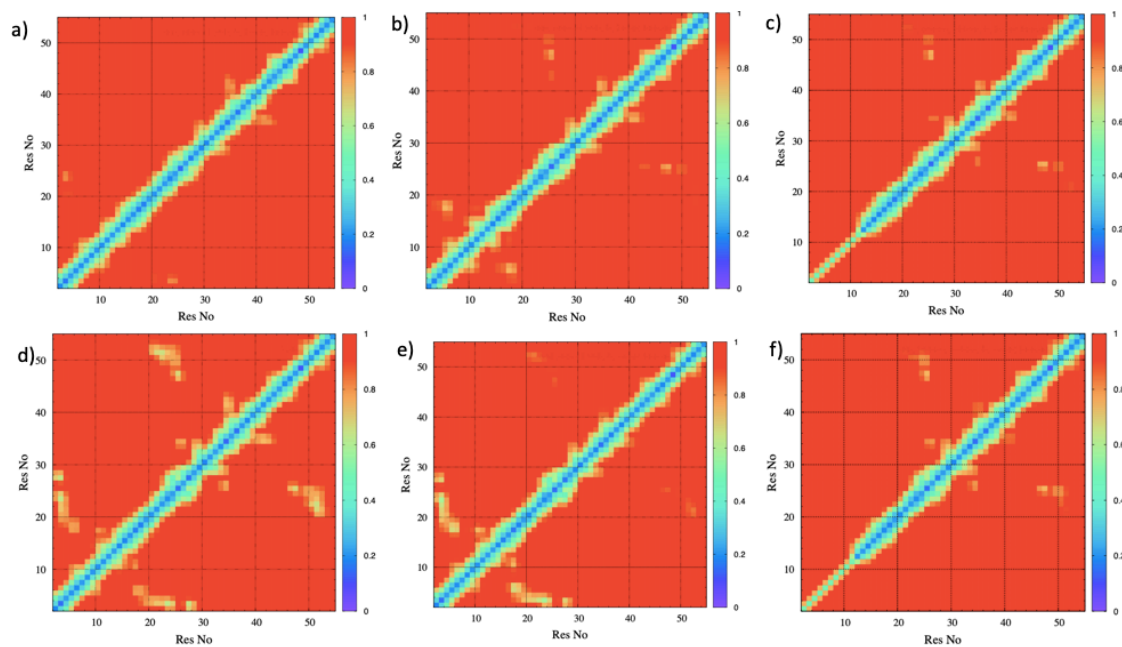


Figure 6.5. Intrachain sidechain contact map, a) – c), one-fold two-layers systems, a), F1L2, pH = 7, b). F1L2, pH = 4, c). F1L2-, pH = 7. d) – f), two-fold two layers systems, d), F2L2, pH = 7, e). F2L2, pH = 4, f). F2L2-, pH = 7. Color bar stands for distance in nm unit.

This hydrogen bond network shifting observation is also consistent with the sidechain contact we captured in fig 6.5. b and could explain the low RMSF value in C-terminal cavity under acidic conditions, see figure 6.4. a.

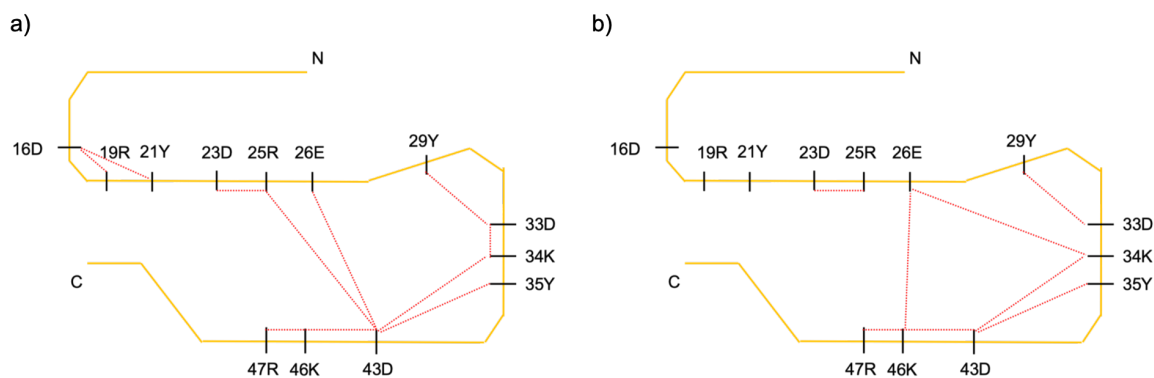


Figure 6.6. Intrachain Sidechain salt-bridge/hydrogen bond network at a). pH 4 and b). pH 7. Yellow, black and red lines stand for protein chain, residues and side-chain interactions (average occupancy per chain larger than 4 %). N and C stand for terminus.

On the other hand, see fig 6.5 d)-e), for the two-fold two-layer fibril fragment F2L2 exist already at pH=7 a multitude of contacts involving both N- and C-terminal region of SAA chains, and the overall stability is similar for neutral and acidic conditions. This indicates that unlike for the meta-stable F1L2 dimer, N-terminal packing can maintain the stability of the fibril by both inter- and intra- chain interactions once a critical size is reached. Hence, acidic conditions do not change the stability of fibrils but further their formation by stabilizing early-stage on-pathway meta-stable states.

6.3.3. Roles of N-terminal amyloidogenic region and C-terminal disordered region

The critical role of the N-terminal residues for maintaining both inter- and intra- molecular interactions is comparable with previous experimental studies, which showed that the truncation of the first eleven residues prevents SAA fibril formation. Hence, to understand the role of these residues in more detail, we have also considered truncated fragments without these residues. These

simulations were done at neutral pH. A RMSD and RMSF values comparison for the two class of models is shown in figure 6.7.

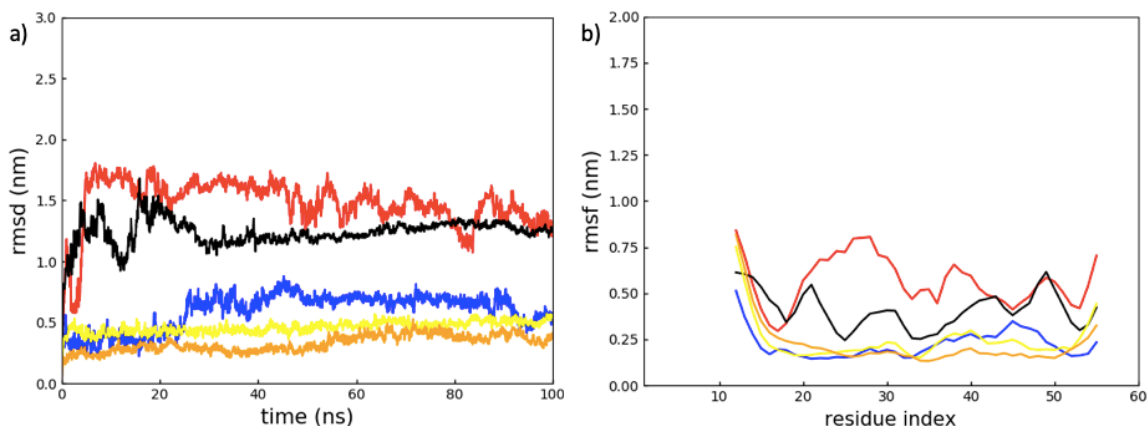


Figure 6.7. Under pH = 7 condition without residues 2-11 for each chain. a). Average RMSD per chain (overall lowest value among three trails are chosen to plot) for different systems, red, F1L1- system, blue, F1L2- system, black, F2L1- system, yellow, F2L2- system, orange, F1L3- system, pink, F2L3- system. b). Average RMSF value for different systems, the color scheme is the same as a).

For one-fold models, F1L1-, F1L2-, F1L3-, the lack of the N-terminal region leads to a more rigid C-terminal region, see fig 6.8. a-b, that implies a higher stability of these fibril models. These results appear to contradict experimental observations that SAA cannot form fibrils when the first eleven residues are truncated. However, our results can be explained if one assumes that the SAA fibril form by a process where the N-terminal regions serve as anchors, that ease the stacking of the rest of chains by enabling interchain interactions. Once the stacking of two chains is completed, they will be stabilized not only by N-terminal residues but also in addition by C-terminal residues. Our hypothesis implies that existing fibrils will not dissolve if the first eleven residues are cleaved, but that fragments derived from such fibrils with truncated chains cannot nucleate fibril growths and elongation. While going beyond the scope of this work, this could in principle be tested in suitable seeding experiments. We remark that in ~ 15 % of conformations

of the one-fold two-layer N-terminal-truncated F1L2- dimer simulations, the region spanned by residues 12-21 in a chain folds back into a helix, see fig 6.9a. This refolding is not observed for the full-sized F1L2 dimer. Hence, while the more rigid C-terminal region is stabilizing the fibril geometry, without the N-terminal, stacking contacts involving residues 12-21 are easily lost, leading likely to further degradation of the SAA fibril. Note that also the N-terminal truncated two-fold models, F2L1- and F2L2-, are less stable than the corresponding models F2L1 and F2L2 built from full-sized chains. This is despite the fact that the intrachain contacts in the C-terminal cavity are increased for each fold, see figure 6.5. f. Hence, the truncation encourages break-up of the two protofibrils which, without additional packing contacts will also further decay, see the representative structure in fig 6.9b.

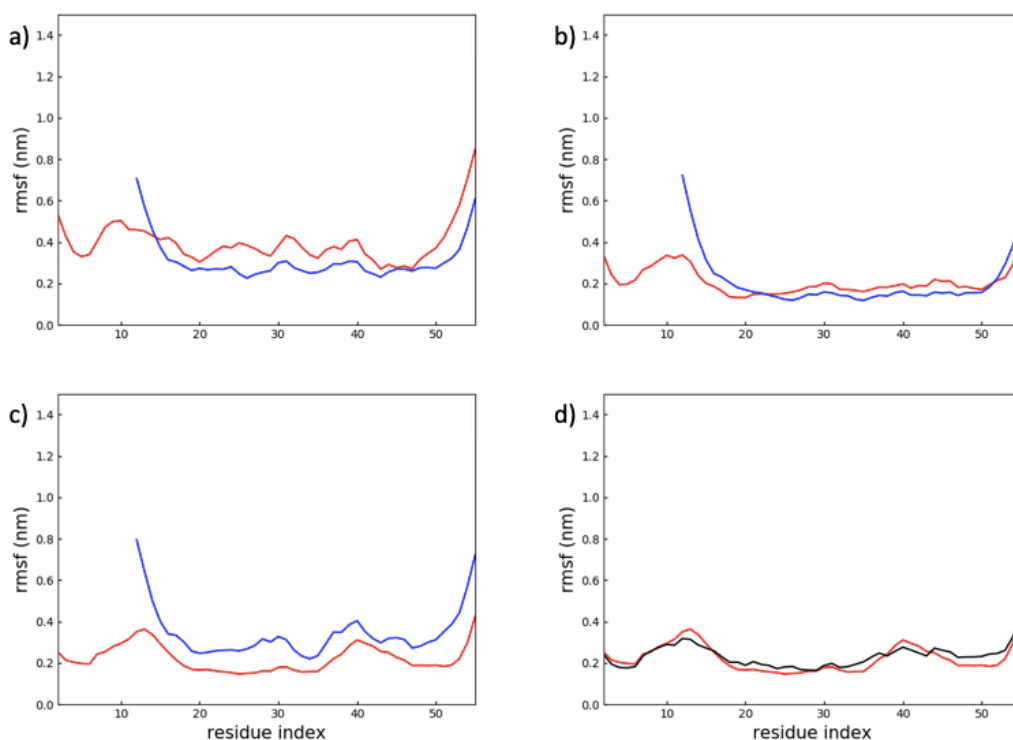


Figure 6.8. Average RMSF value comparison for the same system size under different conditions. a). red (pH = 7, F1L2 system), blue (pH = 7, F1L2- system). b). red (pH = 7, F1L3 system), blue (pH = 7, F1L3- system). c). red (pH = 7, F2L2 system), blue (pH = 7, F2L2- system). d). red (pH = 7, F2L2 system), black (pH = 7, F2L2+ system).

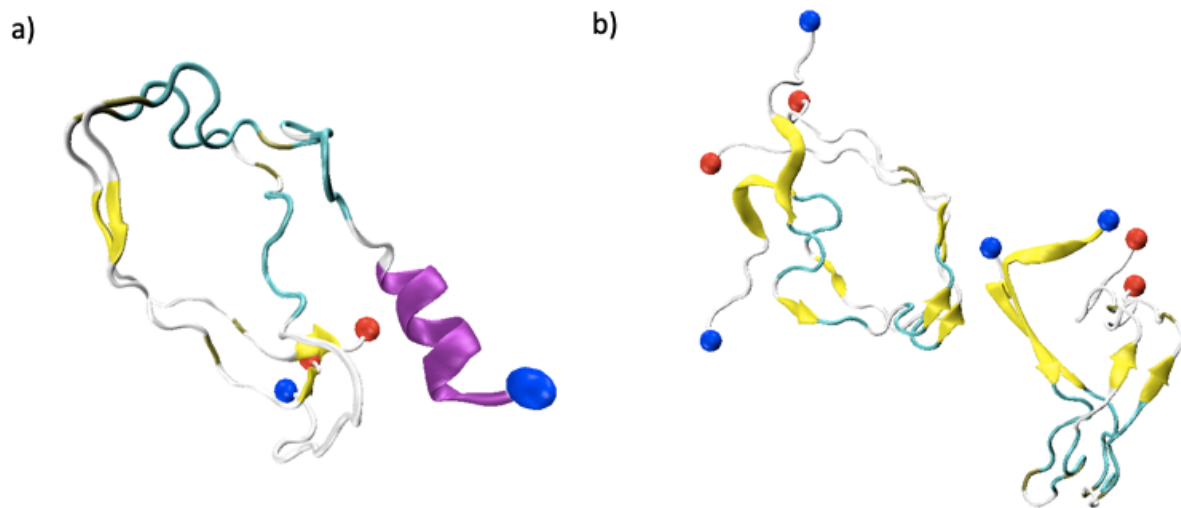


Figure 6.9. Representative structure of a). F1L2- and b). F2L2- systems at pH 7. Color scheme is consist with fig 6.3. (Purple stands for helix)

The experimentally resolved SAA fibril structure contains a disordered C-terminus formed by residues 56 - 69. This region likely extends to residue 70-76 in the most common SAA₁₋₇₆ fibrils. Wondering about the role of this disordered region for stability of SAA fibrils, we have also performed fibril simulations where we have extended the chains in the resolved PDB-structure, assuming a disordered, random, conformation for the added residues 56-76. Our data reveal that for two-fold stable systems, F2L2+ and F2L3+, the stability of the fibril region (residues 2 - 55) is similar to that of the non-extended systems, see fig 6.10 and 6.8d. Hence, the disordered region does not alter the stability of two-fold systems.

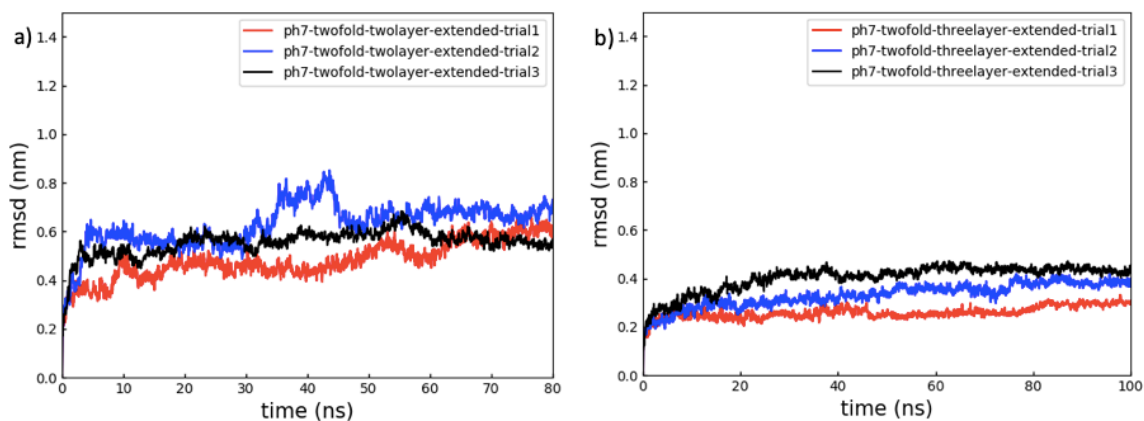


Figure 6.10. Average RMSD per chain for three trails of pH = 7, systems with extended residues 56-76 region a). F2L2+ system. b). F2L3+ system. Only the residues 2-55 region is considered.

6.4. Conclusions

SAA in its native structure is stabilized by intra-chain interactions, especially the salt-bridge interaction between residues 9E and 1R, which maintains the stability of N-terminal amyloidogenic region and prevents misfolding and aggregation. Also, in the fibril structure studied in this work, both N-terminal packing and a network of ionic interaction and hydrogen bond inside the C-terminal cavity are critical for stabilizing the fibril structure. However, we find the disordered C-terminal tail has very limited effects on fibril stability. We have identified as the critical size of fibrils a tetramer with two-folds and two layers, which we conjecture is formed by first stacking two chains on each other, before two such dimers pack together into a two-fold structure anchored by their N-terminal regions. The critical role of the N-terminus explains the experimental observation that without the first eleven residues SAA does not aggregate into fibrils. We speculate that amyloid formation requires these residues to unravel from the N-terminal helix seen in the native structure. This allows the now extended segment to interact with another one

from a second chain, forming a β -anchor that is the starting point of fibril formation. While our stability analysis gives already some hints on the underlying dynamics of SAA fibril formation, more elaborated simulations are needed to establish the mechanism of SAA fibrilization.

Chapter 7. Summary and Outlook

In this thesis, two proteins, A β and SAA, lead to amyloidoses are discussed. Despite having different sequence, A β and SAA both exhibit highly organized cross β fibril structures. Previous experimental studies were fruitful in revealing high resolution fibril structures, however, the understanding of soluble oligomers and intermediates is not as deep as amyloid fibril structure despite of these soluble oligomers to contain more toxicity than mature fibrils.

Due to the soluble and unstable nature of oligomers, it could be challenging for experimental approach to capture them, but the knowledge can provide better understanding of the early stage of aggregation. In order to achieve this goal, computational approaches are more suitable to be applied. Although molecular dynamic (MD) simulation can provide physic-based atomic structural information in nanosecond to microsecond time scale, MD is very likely to be trapped in local minima, leading to inefficient sampling. Replica exchange molecular dynamics (REMD) introduced in chapter 2 can enhance the sampling efficiency compared to MD, it works very well for smaller system, once the system increase to above ~ 200 residues with explicit solvent, the computational cost will dramatically increase if the temperature interval remains same, In chapter 4, an approach called multi-exchange replica exchange molecular dynamics (ME-REMD) was implemented and it can decrease one fourth of the computational cost, ME-REMD could be a potential solution for medium size system.

However, to continue based on the results of this thesis, since we have already studied the N-terminal region of SAA and understood the initiation of misfolding and aggregation on the critical

region, and further on our study of hexamer and monomer SAA (both fragments and full-sized protein) proposed a mechanism of SAA downregulation under natural and acidic pH, the failure of the downregulation eventually leads to aggregation and amyloid formation. From fragments to full-sized protein, this bottom-up investigation provide the structure information of SAA as non-fibril structure from pieces to a relative whole picture, together this two studies provided the information of early stage of SAA misfolding and aggregation. Our study of SAA fibril stability revealed the key interaction to stabilize the structure and also provide the potential treatment/prevention for SAA fibril formation, this study provided information about the post fibril formation, a gap of knowledge between early stage and post stage amyloid formation needs still to be filled.

To fulfill the gap, we need to employ appropriate computational methods. Based on previous experimental results and our simulations, the SAA native structure is a hexamer and three layers and two-fold is the critical size to stabilize the fibril. As discussed above, neither REMD and ME-REMD can sufficiently explore the transition/intermediates landscape in between native structure with high helix content and β -rich fibril structure due to their limitation of system size and simulation time scale. Targeted molecular dynamics (TMD) could be one solution which can provide some information of transition pathway, however, the output pathway could be biased due to the nature of TMD.

Our lab has developed an unbiased approach that can fast explore the landscape of the possible intermediates, this replica-exchange-with-tunneling (RET) approach can possibly answer questions, such as, the transition pathway from native to fibril structure, the key components controlling this transition, and provide insights into drug design to interrupt the transition pathway.

In future work a combination of multi-scale simulation and RET will be used to solve the conversion mechanism between SAA in its native form and its fibril structure.

References

1. Westermark, G. T.; Fändrich, M.; Westermark, P., AA amyloidosis: pathogenesis and targeted therapy. *Annual Review of Pathology: Mechanisms of Disease* **2015**, *10*, 321-344.
2. Ferrone, F. A.; Hofrichter, J.; Eaton, W. A., Kinetics of sickle hemoglobin polymerization: II. A double nucleation mechanism. *Journal of molecular biology* **1985**, *183* (4), 611-631.
3. Ferrone, F., [17] Analysis of protein aggregation kinetics. In *Methods in enzymology*, Elsevier: 1999; Vol. 309, pp 256-274.
4. Iannuzzi, C.; Irace, G.; Sirangelo, I., The effect of glycosaminoglycans (GAGs) on amyloid aggregation and toxicity. *Molecules* **2015**, *20* (2), 2510-2528.
5. Querfurth, H. W.; LaFerla, F. M., Mechanisms of disease. *N Engl J Med* **2010**, *362* (4), 329-344.
6. DeToma, A. S.; Salamekh, S.; Ramamoorthy, A.; Lim, M. H., Misfolded proteins in Alzheimer's disease and type II diabetes. *Chemical Society Reviews* **2012**, *41* (2), 608-621.
7. Nasica-Labouze, J.; Nguyen, P. H.; Sterpone, F.; Berthoumieu, O.; Buchete, N.-V.; Cote, S.; De Simone, A.; Doig, A. J.; Faller, P.; Garcia, A., Amyloid β protein and Alzheimer's disease: When computer simulations complement experimental studies. *Chemical reviews* **2015**, *115* (9), 3518-3563.
8. Walsh, D. M.; Selkoe, D. J., A β oligomers—a decade of discovery. *Journal of neurochemistry* **2007**, *101* (5), 1172-1184.
9. Sakono, M.; Zako, T., Amyloid oligomers: formation and toxicity of A β oligomers. *The FEBS journal* **2010**, *277* (6), 1348-1358.
10. Schütz, A. K.; Vagt, T.; Huber, M.; Ovchinnikova, O. Y.; Cadalbert, R.; Wall, J.; Güntert, P.; Böckmann, A.; Glockshuber, R.; Meier, B. H., Atomic-resolution three-dimensional structure of amyloid β fibrils bearing the Osaka mutation. *Angewandte Chemie International Edition* **2015**, *54* (1), 331-335.
11. Lu, J.-X.; Qiang, W.; Yau, W.-M.; Schwieters, C. D.; Meredith, S. C.; Tycko, R., Molecular structure of β -amyloid fibrils in Alzheimer's disease brain tissue. *Cell* **2013**, *154* (6), 1257-1268.
12. Petkova, A. T.; Yau, W.-M.; Tycko, R., Experimental constraints on quaternary structure in Alzheimer's β -amyloid fibrils. *Biochemistry* **2006**, *45* (2), 498-512.
13. Lührs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Döbeli, H.; Schubert, D.; Riek, R., 3D structure of Alzheimer's amyloid- β (1–42) fibrils. *Proceedings of the National Academy of Sciences* **2005**, *102* (48), 17342-17347.
14. Xiao, Y.; Ma, B.; McElheny, D.; Parthasarathy, S.; Long, F.; Hoshi, M.; Nussinov, R.; Ishii, Y., A β (1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nature structural & molecular biology* **2015**, *22* (6), 499-505.
15. Xi, W.; Wang, W.; Abbott, G.; Hansmann, U. H., Stability of a recently found triple- β -stranded A β 1–42 fibril motif. *The Journal of Physical Chemistry B* **2016**, *120* (20), 4548-4557.
16. Xi, W.; Vanderford, E. K.; Hansmann, U. H., Out-of-Register A β 42 Assemblies as Models for Neurotoxic Oligomers and Fibrils. *Journal of chemical theory and computation* **2018**, *14* (2), 1099-1110.
17. Xi, W.; Hansmann, U. H., Ring-like N-fold Models of A β 42 fibrils. *Scientific reports* **2017**, *7* (1), 1-14.

18. Xi, W.; Dean, D. N.; Stockmal, K. A.; Morgan, S. E.; Hansmann, U. H.; Rangachari, V., Large fatty acid-derived A β 42 oligomers form ring-like assemblies. *The Journal of chemical physics* **2019**, *150* (7), 075101.
19. Agrawal, N.; Skelton, A. A., Structure and Function of Alzheimer's Amyloid β Proteins from Monomer to Fibrils: A Mini Review. *The protein journal* **2019**, *38* (4), 425-434.
20. Srinivasan, S.; Patke, S.; Wang, Y.; Ye, Z.; Litt, J.; Srivastava, S. K.; Lopez, M. M.; Kurouski, D.; Lednev, I. K.; Kane, R. S., Pathogenic serum amyloid A 1.1 shows a long oligomer-rich fibrillation lag phase contrary to the highly amyloidogenic non-pathogenic SAA2. 2. *Journal of Biological Chemistry* **2013**, *288* (4), 2744-2755.
21. McCubbin, W. D.; Kay, C.; Narindrasorasak, S.; Kisilevsky, R., Circular-dichroism studies on two murine serum amyloid A proteins. *Biochemical Journal* **1988**, *256* (3), 775-783.
22. Lu, J.; Yu, Y.; Zhu, I.; Cheng, Y.; Sun, P. D., Structural mechanism of serum amyloid A-mediated inflammatory amyloidosis. *Proceedings of the National Academy of Sciences* **2014**, *111* (14), 5189-5194.
23. Egashira, M.; Takase, H.; Yamamoto, I.; Tanaka, M.; Saito, H., Identification of regions responsible for heparin-induced amyloidogenesis of human serum amyloid A using its fragment peptides. *Archives of biochemistry and biophysics* **2011**, *511* (1-2), 101-106.
24. Nordling, E.; Abraham-Nordling, M., Colonic amyloidosis, computational analysis of the major amyloidogenic species, Serum Amyloid A. *Computational biology and chemistry* **2012**, *39*, 29-34.
25. Jannone, J. M.; Grigg, J. I.; Aguirre, L. M.; Jones, E. M., Electrostatic Interactions at N- and C-Termini Determine Fibril Polymorphism in Serum Amyloid A Fragments. *The Journal of Physical Chemistry B* **2016**, *120* (39), 10258-10268.
26. Bernhardt, N. A.; Xi, W.; Wang, W.; Hansmann, U. H., Simulating protein fold switching by replica exchange with tunneling. *Journal of chemical theory and computation* **2016**, *12* (11), 5656-5666.
27. Colón, W.; Aguilera, J. J.; Srinivasan, S., Intrinsic stability, oligomerization, and amyloidogenicity of HDL-free serum amyloid A. In *Lipids in Protein Misfolding*, Springer: 2015; pp 117-134.
28. Esler, W. P.; Stimson, E. R.; Jennings, J. M.; Vinters, H. V.; Ghilardi, J. R.; Lee, J. P.; Mantyh, P. W.; Maggio, J. E., Alzheimer's disease amyloid propagation by a template-dependent dock-lock mechanism. *Biochemistry* **2000**, *39* (21), 6288-6295.
29. Claus, S.; Meinhardt, K.; Aumüller, T.; Puscalau-Girtu, I.; Linder, J.; Haupt, C.; Walther, P.; Syrovets, T.; Simmet, T.; Fändrich, M., Cellular mechanism of fibril formation from serum amyloid A1 protein. *EMBO reports* **2017**, *18* (8), 1352-1366.
30. Liberta, F.; Loerch, S.; Renegarbe, M.; Schierhorn, A.; Westermark, P.; Westermark, G. T.; Hazenberg, B. P.; Grigorieff, N.; Fändrich, M.; Schmidt, M., Cryo-EM fibril structures from systemic AA amyloidosis reveal the species complementarity of pathological amyloids. *Nature communications* **2019**, *10* (1), 1-10.
31. Wagner, I.; Musso, H., New naturally occurring amino acids. *Angewandte Chemie International Edition in English* **1983**, *22* (11), 816-828.
32. Nelson, D. L.; Cox, M. M.; Lehninger, A. L., *Principles of biochemistry*. Freeman New York:: 2008.
33. Latham, M., Body composition, the functions of food, metabolism and energy. *Human nutrition in the developing world. Food and Nutrition Series* **1997**, (29).

34. Michal, G.; Schomburg, D., *Biochemical pathways: an atlas of biochemistry and molecular biology*. John Wiley & Sons: 2012.
35. Aerts, J. W.; Röling, W. F.; Elsaesser, A.; Ehrenfreund, P., Biota and biomolecules in extreme environments on Earth: implications for life detection on Mars. *Life* **2014**, *4* (4), 535-565.
36. Meisenberg, G.; Simmons, W. H., *Principles of Medical Biochemistry E-Book*. Elsevier Health Sciences: 2016.
37. Muller, P., Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994). *Pure and Applied Chemistry* **1994**, *66* (5), 1077-1184.
38. Martin, R. B., Free energies and equilibria of peptide bond hydrolysis and formation. *Biopolymers: Original Research on Biomolecules* **1998**, *45* (5), 351-353.
39. Radzicka, A.; Wolfenden, R., Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases. *Journal of the American Chemical Society* **1996**, *118* (26), 6105-6109.
40. Stoker, H. S., *General, organic, and biological chemistry*. Nelson Education: 2012.
41. Brocchieri, L.; Karlin, S., Protein length in eukaryotic and prokaryotic proteomes. *Nucleic acids research* **2005**, *33* (10), 3390-3400.
42. Sanger, F., The arrangement of amino acids in proteins. In *Advances in protein chemistry*, Elsevier: 1952; Vol. 7, pp 1-67.
43. Chothia, C.; Lesk, A. M., The relation between the divergence of sequence and structure in proteins. *The EMBO journal* **1986**, *5* (4), 823-826.
44. Schellman, J. A.; Schellman, C. G., Kaj Ulrik Linderstrøm-Lang (1896–1959). *Protein science* **1997**, *6* (5), 1092-1100.
45. Linderstrøm-Lang, K. U., *Lane medical lectures: proteins and enzymes*. Stanford University Press: 1952; Vol. 6.
46. Ramachandran, G. N., Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **1963**, *7*, 95-99.
47. Dunitz, J. D., Pauling's left-handed α -helix. *Angewandte Chemie International Edition* **2001**, *40* (22), 4167-4173.
48. Branden, C. I.; Tooze, J., *Introduction to protein structure*. Garland Science: 2012.
49. Anfinsen, C. B., The formation and stabilization of protein structure. *Biochemical Journal* **1972**, *128* (4), 737.
50. Tanford, C., The hydrophobic effect and the organization of living matter. *Science* **1978**, *200* (4345), 1012-1018.
51. Deechongkit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W., Context-dependent contributions of backbone hydrogen bonding to β -sheet folding energetics. *Nature* **2004**, *430* (6995), 101-105.
52. Irbäck, A.; Sandelin, E., On hydrophobicity correlations in protein chains. *Biophysical Journal* **2000**, *79* (5), 2252-2258.
53. Irbäck, A.; Peterson, C.; Potthast, F., Evidence for nonrandom hydrophobicity structures in protein chains. *proceedings of the National Academy of Sciences* **1996**, *93* (18), 9533-9538.
54. Foy, S. G.; Wilson, B. A.; Bertram, J.; Cordes, M. H.; Masel, J., A shift in aggregation avoidance strategy marks a long-term direction to protein evolution. *Genetics* **2019**, *211* (4), 1345-1355.

55. Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G., Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Structure, Function, and Bioinformatics* **1995**, *21* (3), 167-195.
56. Leopold, P. E.; Montal, M.; Onuchic, J. N., Protein folding funnels: a kinetic approach to the sequence-structure relationship. *Proceedings of the National Academy of Sciences* **1992**, *89* (18), 8721-8725.
57. Chong, S.-H.; Ham, S., folding free energy Landscape of ordered and intrinsically Disordered proteins. *Scientific reports* **2019**, *9* (1), 1-9.
58. Dobson, C. M., Protein folding and misfolding. *Nature* **2003**, *426* (6968), 884-890.
59. Fersht, A. R., Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism. *Proceedings of the National Academy of Sciences* **2000**, *97* (4), 1525-1529.
60. De Felice, F. G.; Vieira, M. N.; Meirelles, M. N. L.; Morozova-Roche, L. A.; Dobson, C. M.; Ferreira, S. T., Formation of amyloid aggregates from human lysozyme and its disease-associated variants using hydrostatic pressure. *The FASEB journal* **2004**, *18* (10), 1099-1101.
61. Sunde, M.; Blake, C. C., From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quarterly reviews of biophysics* **1998**, *31* (1), 1-39.
62. Shirahama, T.; Cohen, A. S., High-resolution electron microscopic analysis of the amyloid fibril. *The Journal of cell biology* **1967**, *33* (3), 679-708.
63. Eanes, E.; Glenner, G. G., X-ray diffraction studies on amyloid filaments. *Journal of Histochemistry & Cytochemistry* **1968**, *16* (11), 673-677.
64. Meinhardt, J.; Tartaglia, G. G.; Pawar, A.; Christopeit, T.; Hortschansky, P.; Schroeckh, V.; Dobson, C. M.; Vendruscolo, M.; Fändrich, M., Similarities in the thermodynamics and kinetics of aggregation of disease-related A β (1-40) peptides. *Protein science* **2007**, *16* (6), 1214-1222.
65. Jarrett, J. T.; Lansbury Jr, P. T., Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **1993**, *73* (6), 1055-1058.
66. Caughey, B.; Lansbury Jr, P. T., Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annual review of neuroscience* **2003**, *26* (1), 267-298.
67. Teplow, D. B., Structural and kinetic features of amyloid β -protein fibrillogenesis. *Amyloid* **1998**, *5* (2), 121-142.
68. Straub, J. E.; Thirumalai, D., Toward a molecular theory of early and late events in monomer to amyloid fibril formation. *Annual review of physical chemistry* **2011**, *62*, 437-463.
69. Pellarin, R.; Caflich, A., Interpreting the aggregation kinetics of amyloid peptides. *Journal of molecular biology* **2006**, *360* (4), 882-892.
70. Pallitto, M. M.; Murphy, R. M., A mathematical model of the kinetics of β -amyloid fibril growth from the denatured state. *Biophysical journal* **2001**, *81* (3), 1805-1822.
71. Harper, J. D.; Lansbury Jr, P. T., Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annual review of biochemistry* **1997**, *66* (1), 385-407.
72. Sharma, R.; Chakraborty, S. N.; Chakravarty, C., Entropy, diffusivity, and structural order in liquids with waterlike anomalies. *The Journal of chemical physics* **2006**, *125* (20), 204501.

73. Agarwal, M.; Kushwaha, H. R.; Chakravarty, C., Local order, energy, and mobility of water molecules in the hydration shell of small peptides. *The Journal of Physical Chemistry B* **2010**, *114* (1), 651-659.
74. Association, A. s., 2016 Alzheimer's disease facts and figures. *Alzheimer's & Dementia* **2016**, *12* (4), 459-509.
75. Wimo, A.; Guerchet, M.; Ali, G.-C.; Wu, Y.-T.; Prina, A. M.; Winblad, B.; Jönsson, L.; Liu, Z.; Prince, M., The worldwide costs of dementia 2015 and comparisons with 2010. *Alzheimer's & Dementia* **2017**, *13* (1), 1-7.
76. Uflacker, A.; Doraiswamy, P. M., Alzheimer's disease: an overview of recent developments and a look to the future. *Focus* **2017**, *15* (1), 13-17.
77. O'Brien, R. J.; Wong, P. C., Amyloid precursor protein processing and Alzheimer's disease. *Annual review of neuroscience* **2011**, *34*, 185-204.
78. Del Prete, D.; Checler, F.; Chami, M., Ryanodine receptors: physiological function and deregulation in Alzheimer disease. *Molecular neurodegeneration* **2014**, *9* (1), 21.
79. Bergström, P.; Agholme, L.; Nazir, F. H.; Satir, T. M.; Toombs, J.; Wellington, H.; Strandberg, J.; Bontell, T. O.; Kvartsberg, H.; Holmström, M., Amyloid precursor protein expression and processing are differentially regulated during cortical neuron differentiation. *Scientific reports* **2016**, *6* (1), 1-14.
80. Spies, P. E.; Verbeek, M. M.; van Groen, T.; Claassen, J., Reviewing reasons for the decreased CSF Aβ42 concentration in Alzheimer disease. *Front Biosci* **2012**, *17*, 2024-34.
81. Ball, K. A.; Phillips, A. H.; Nerenberg, P. S.; Fawzi, N. L.; Wemmer, D. E.; Head-Gordon, T., Homogeneous and heterogeneous tertiary structure ensembles of amyloid-β peptides. *Biochemistry* **2011**, *50* (35), 7612-7628.
82. Uversky, V. N., Intrinsic disorder in proteins associated with neurodegenerative diseases. In *Protein folding and misfolding: neurodegenerative diseases*, Springer: 2009; pp 21-75.
83. Janek, K.; Rothmund, S.; Gast, K.; Beyermann, M.; Zipper, J.; Fabian, H.; Bienert, M.; Krause, E., Study of the conformational transition of Aβ (1–42) using d-amino acid replacement analogues. *Biochemistry* **2001**, *40* (18), 5457-5463.
84. Crescenzi, O.; Tomaselli, S.; Guerrini, R.; Salvadori, S.; D'Ursi, A. M.; Temussi, P. A.; Picone, D., Solution structure of the Alzheimer amyloid β-peptide (1–42) in an apolar microenvironment: Similarity with a virus fusion domain. *European Journal of Biochemistry* **2002**, *269* (22), 5642-5648.
85. Agrawal, N.; Skelton, A. A., Binding of 12-crown-4 with Alzheimer's Aβ40 and Aβ42 monomers and its effect on their conformation: insight from molecular dynamics simulations. *Molecular pharmaceutics* **2018**, *15* (1), 289-299.
86. Luttmann, E.; Fels, G., All-atom molecular dynamics studies of the full-length β-amyloid peptides. *Chemical physics* **2006**, *323* (1), 138-147.
87. Miyashita, N.; Straub, J. E.; Thirumalai, D., Structures of β-Amyloid Peptide 1–40, 1–42, and 1–55 in the 672–726 Fragment of APP in a Membrane Environment with Implications for Interactions with γ-Secretase. *Journal of the American Chemical Society* **2009**, *131* (49), 17843-17852.
88. Nimmrich, V.; Grimm, C.; Draguhn, A.; Barghorn, S.; Lehmann, A.; Schoemaker, H.; Hillen, H.; Gross, G.; Ebert, U.; Bruehl, C., Amyloid β oligomers (Aβ1–42 globulomer) suppress spontaneous synaptic activity by inhibition of P/Q-type calcium currents. *Journal of Neuroscience* **2008**, *28* (4), 788-797.

89. Narayan, P.; Meehan, S.; Carver, J. A.; Wilson, M. R.; Dobson, C. M.; Klenerman, D., Amyloid- β oligomers are sequestered by both intracellular and extracellular chaperones. *Biochemistry* **2012**, *51* (46), 9270-9276.
90. Kaye, R.; Lasagna-Reeves, C. A., Molecular mechanisms of amyloid oligomers toxicity. *Journal of Alzheimer's Disease* **2013**, *33* (s1), S67-S78.
91. He, Y.; Zheng, M.-M.; Ma, Y.; Han, X.-J.; Ma, X.-Q.; Qu, C.-Q.; Du, Y.-F., Soluble oligomers and fibrillar species of amyloid β -peptide differentially affect cognitive functions and hippocampal inflammatory response. *Biochemical and biophysical research communications* **2012**, *429* (3-4), 125-130.
92. Baglioni, S.; Casamenti, F.; Bucciantini, M.; Lheshi, L. M.; Taddei, N.; Chiti, F.; Dobson, C. M.; Stefani, M., Prefibrillar amyloid aggregates could be generic toxins in higher organisms. *Journal of Neuroscience* **2006**, *26* (31), 8160-8167.
93. Pham, J. D.; Chim, N.; Goulding, C. W.; Nowick, J. S., Structures of oligomers of a peptide from β -amyloid. *Journal of the American Chemical Society* **2013**, *135* (33), 12460-12467.
94. Benilova, I.; Karran, E.; De Strooper, B., The toxic A β oligomer and Alzheimer's disease: an emperor in need of clothes. *Nature neuroscience* **2012**, *15* (3), 349.
95. Kirschner, D. A.; Abraham, C.; Selkoe, D. J., X-ray diffraction from intraneuronal paired helical filaments and extraneuronal amyloid fibers in Alzheimer disease indicates cross-beta conformation. *Proceedings of the National Academy of Sciences* **1986**, *83* (2), 503-507.
96. Inouye, H.; Fraser, P. E.; Kirschner, D. A., Structure of beta-crystallite assemblies formed by Alzheimer beta-amyloid protein analogues: analysis by x-ray diffraction. *Biophysical journal* **1993**, *64* (2), 502.
97. Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R., A structural model for Alzheimer's β -amyloid fibrils based on experimental constraints from solid state NMR. *Proceedings of the National Academy of Sciences* **2002**, *99* (26), 16742-16747.
98. Paravastu, A. K.; Leapman, R. D.; Yau, W.-M.; Tycko, R., Molecular structural basis for polymorphism in Alzheimer's β -amyloid fibrils. *Proceedings of the National Academy of Sciences* **2008**, *105* (47), 18349-18354.
99. Riek, R.; Eisenberg, D. S., The activities of amyloids from a structural perspective. *Nature* **2016**, *539* (7628), 227-235.
100. Agrawal, N.; Skelton, A. A., 12-crown-4 ether disrupts the patient brain-derived amyloid- β -fibril trimer: Insight from all-atom molecular dynamics simulations. *ACS chemical neuroscience* **2016**, *7* (10), 1433-1441.
101. Gremer, L.; Schölzel, D.; Schenk, C.; Reinartz, E.; Labahn, J.; Ravelli, R. B.; Tusche, M.; Lopez-Iglesias, C.; Hoyer, W.; Heise, H., Fibril structure of amyloid- β (1-42) by cryo-electron microscopy. *Science* **2017**, *358* (6359), 116-119.
102. Husby, G.; Marhaug, G.; Dowtor, B.; Sletten, K.; Sipe, J. D., Serum amyloid A (SAA): biochemistry, genetics and the pathogenesis of AA amyloidosis. *Amyloid* **1994**, *1* (2), 119-137.
103. Kisilevsky, R.; Manley, P. N., Acute-phase serum amyloid A: perspectives on its physiological and pathological roles. *Amyloid* **2012**, *19* (1), 5-14.
104. Snow, A. D.; Bramson, R.; Mar, H.; Wight, T. N.; Kisilevsky, R., A temporal and ultrastructural relationship between heparan sulfate proteoglycans and AA amyloid in experimental amyloidosis. *Journal of Histochemistry & Cytochemistry* **1991**, *39* (10), 1321-1330.

105. Noborn, F.; Ancsin, J. B.; Ubhayasekera, W.; Kisilevsky, R.; Li, J.-P., Heparan sulfate dissociates serum amyloid A (SAA) from acute-phase high-density lipoprotein, promoting SAA aggregation. *Journal of Biological Chemistry* **2012**, *287* (30), 25669-25677.
106. Kluge-Beckerman, B.; Manaloor, J.; Liepnieks, J., Binding, trafficking and accumulation of serum amyloid A in peritoneal macrophages. *Scandinavian journal of immunology* **2001**, *53* (4), 393-400.
107. Yamada, T.; Liepnieks, J.; Kluge-Beckerman, B.; Benson, M., Cathepsin B generates the most common form of amyloid A (76 residues) as a degradation product from serum amyloid A. *Scandinavian journal of immunology* **1995**, *41* (1), 94-97.
108. Stix, B.; Kähne, T.; Sletten, K.; Raynes, J.; Roessner, A.; Röcken, C., Proteolysis of AA amyloid fibril proteins by matrix metalloproteinases-1,-2, and-3. *The American journal of pathology* **2001**, *159* (2), 561-570.
109. van der Hilst, J.; Yamada, T.; Op den Camp, H.; van der Meer, J.; Drenth, J.; Simon, A., Increased susceptibility of serum amyloid A 1.1 to degradation by MMP-1: potential explanation for higher risk of type AA amyloidosis. *Rheumatology* **2008**, *47* (11), 1651-1654.
110. Wang, W.; Khatua, P.; Hansmann, U. H., Cleavage, Downregulation, and Aggregation of Serum Amyloid A. *The Journal of Physical Chemistry B* **2020**, *124* (6), 1009-1019.
111. Roychoudhuri, R.; Yang, M.; Hoshi, M. M.; Teplow, D. B., Amyloid β -protein assembly and Alzheimer disease. *Journal of Biological Chemistry* **2009**, *284* (8), 4749-4753.
112. Morris, K. L.; Serpell, L. C., X-ray fibre diffraction studies of amyloid fibrils. In *Amyloid Proteins*, Springer: 2012; pp 121-135.
113. Schmidt, M.; Rohou, A.; Lasker, K.; Yadav, J. K.; Schiene-Fischer, C.; Fändrich, M.; Grigorieff, N., Peptide dimer structure in an A β (1-42) fibril visualized with cryo-EM. *Proceedings of the National Academy of Sciences* **2015**, *112* (38), 11858-11863.
114. Scheidt, H. A.; Morgado, I.; Rothemund, S.; Huster, D., Dynamics of amyloid β fibrils revealed by solid-state NMR. *Journal of Biological Chemistry* **2012**, *287* (3), 2017-2021.
115. Sutcliffe, B. T.; Woolley, R. G., On the quantum theory of molecules. *The Journal of chemical physics* **2012**, *137* (22), 22A544.
116. Adcock, S. A.; McCammon, J. A., Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical reviews* **2006**, *106* (5), 1589-1615.
117. Huang, J.; MacKerell Jr, A. D., CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of computational chemistry* **2013**, *34* (25), 2135-2145.
118. Chiti, F.; Dobson, C. M., Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* **2006**, *75*, 333-66.
119. Eisenberg, D.; Jucker, M., The amyloid state of proteins in human diseases. *Cell* **2012**, *148* (6), 1188-203.
120. Knowles, T. P.; Vendruscolo, M.; Dobson, C. M., The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol* **2014**, *15* (6), 384-96.
121. Sipe, J. D.; Cohen, A. S., Review: History of the amyloid fibril. *J Struct Biol* **2000**, *130* (2-3), 88-98.
122. Querfurth, H. W.; LaFerla, F. M., Alzheimer's disease. *N Engl J Med* **2010**, *362* (4), 329-44.
123. DeToma, A. S.; Salamekh, S.; Ramamoorthy, A.; Lim, M. H., Misfolded proteins in Alzheimer's disease and type II diabetes. *Chem Soc Rev* **2012**, *41* (2), 608-21.

124. Nasica-Labouze, J.; Nguyen, P. H.; Sterpone, F.; Berthoumieu, O.; Buchete, N. V.; Cote, S.; De Simone, A.; Doig, A. J.; Faller, P.; Garcia, A.; Laio, A.; Li, M. S.; Melchionna, S.; Mousseau, N.; Mu, Y. G.; Paravastu, A.; Pasquali, S.; Rosenman, D. J.; Strodel, B.; Tarus, B.; Viles, J. H.; Zhang, T.; Wang, C. Y.; Derreumaux, P., Amyloid beta Protein and Alzheimer's Disease: When Computer Simulations Complement Experimental Studies. *Chem Rev* **2015**, *115* (9), 3518-3563.
125. Hardy, J.; Selkoe, D. J., Medicine - The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. *Science* **2002**, *297* (5580), 353-356.
126. Walsh, D. M.; Selkoe, D. J., A beta Oligomers - a decade of discovery. *J Neurochem* **2007**, *101* (5), 1172-1184.
127. Sakono, M.; Zako, T., Amyloid oligomers: formation and toxicity of A beta oligomers. *Febs J* **2010**, *277* (6), 1348-1358.
128. Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R., A structural model for Alzheimer's beta -amyloid fibrils based on experimental constraints from solid state NMR. *Proc Natl Acad Sci U S A* **2002**, *99* (26), 16742-7.
129. Luhrs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dobeli, H.; Schubert, D.; Riek, R., 3D structure of Alzheimer's amyloid-beta(1-42) fibrils. *Proc Natl Acad Sci U S A* **2005**, *102* (48), 17342-7.
130. Petkova, A. T.; Yau, W. M.; Tycko, R., Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry-Us* **2006**, *45* (2), 498-512.
131. Paravastu, A. K.; Leapman, R. D.; Yau, W. M.; Tycko, R., Molecular structural basis for polymorphism in Alzheimer's beta-amyloid fibrils. *Proc Natl Acad Sci U S A* **2008**, *105* (47), 18349-54.
132. Lu, J. X.; Qiang, W.; Yau, W. M.; Schwieters, C. D.; Meredith, S. C.; Tycko, R., Molecular structure of beta-amyloid fibrils in Alzheimer's disease brain tissue. *Cell* **2013**, *154* (6), 1257-68.
133. Schutz, A. K.; Vagt, T.; Huber, M.; Ovchinnikova, O. Y.; Cadalbert, R.; Wall, J.; Guntert, P.; Bockmann, A.; Glockshuber, R.; Meier, B. H., Atomic-resolution three-dimensional structure of amyloid beta fibrils bearing the Osaka mutation. *Angew Chem Int Ed Engl* **2015**, *54* (1), 331-5.
134. Andersen, C. B.; Yagi, H.; Manno, M.; Martorana, V.; Ban, T.; Christiansen, G.; Otzen, D. E.; Goto, Y.; Rischel, C., Branching in Amyloid Fibril Growth. *Biophys J* **2009**, *96* (4), 1529-1536.
135. O'Nuallain, B.; Williams, A. D.; Westermarck, P.; Wetzel, R., Seeding specificity in amyloid growth induced by heterologous fibrils. *J Biol Chem* **2004**, *279* (17), 17490-9.
136. Vishveshwara, N.; Liebman, S. W., Heterologous cross-seeding mimics cross-species prion conversion in a yeast model. *Bmc Biol* **2009**, *7*, 26.
137. Haan, M. N., Therapy Insight: type 2 diabetes mellitus and the risk of late-onset Alzheimer's disease. *Nat Clin Pract Neuro* **2006**, *2* (3), 159-166.
138. Do, T. D.; Economou, N. J.; Chamas, A.; Buratto, S. K.; Shea, J. E.; Bowers, M. T., Interactions between amyloid-beta and Tau fragments promote aberrant aggregates: implications for amyloid toxicity. *J Phys Chem B* **2014**, *118* (38), 11220-30.
139. Morales, R.; Moreno-Gonzalez, I.; Soto, C., Cross-seeding of misfolded proteins: implications for etiology and pathogenesis of protein misfolding diseases. *Plos Pathog* **2013**, *9* (9), e1003537.

140. Workalemahu, M.; Berhanu, U. H. E. H., *Proteins* **2013**, *81*, 1542.
141. Berhanu, W. M.; Hansmann, U. H., Inter-species cross-seeding: stability and assembly of rat-human amylin aggregates. *PLoS One* **2014**, *9* (5), e97051.
142. Berhanu, W. M.; Alred, E. J.; Hansmann, U. H., Stability of Osaka Mutant and Wild-Type Fibril Models. *J Phys Chem B* **2015**, *119* (41), 13063-70.
143. Spirig, T.; Ovchinnikova, O.; Vagt, T.; Glockshuber, R., Direct evidence for self-propagation of different amyloid-beta fibril conformations. *Neurodegener Dis* **2014**, *14* (3), 151-9.
144. Stohr, J.; Watts, J. C.; Mensinger, Z. L.; Oehler, A.; Grillo, S. K.; DeArmond, S. J.; Prusiner, S. B.; Giles, K., Purified and synthetic Alzheimer's amyloid beta (A β) prions. *Proc Natl Acad Sci U S A* **2012**, *109* (27), 11025-30.
145. Tycko, R., Solid-state NMR studies of amyloid fibril structure. *Annu Rev Phys Chem* **2011**, *62*, 279-99.
146. Xiao, Y.; Ma, B.; McElheny, D.; Parthasarathy, S.; Long, F.; Hoshi, M.; Nussinov, R.; Ishii, Y., A β (1-42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nat Struct Mol Biol* **2015**, *22* (6), 499-505.
147. Buchete, N. V.; Tycko, R.; Hummer, G., Molecular dynamics simulations of Alzheimer's beta-amyloid protofilaments. *J Mol Biol* **2005**, *353* (4), 804-21.
148. Berhanu, W. M.; Hansmann, U. H., Structure and dynamics of amyloid-beta segmental polymorphisms. *PLoS One* **2012**, *7* (7), e41479.
149. Alred, E. J.; Scheele, E. G.; Berhanu, W. M.; Hansmann, U. H., Stability of Iowa mutant and wild type A β -peptide aggregates. *J Chem Phys* **2014**, *141* (17), 175101.
150. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* **2008**, *4* (3), 435-447.
151. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E., Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78* (8), 1950-8.
152. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics* **1983**, *79* (2), 926-935.
153. Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G., LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry* **1997**, *18* (12), 1463-1472.
154. Miyamoto, S.; Kollman, P. A., Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of computational chemistry* **1992**, *13* (8), 952-962.
155. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N \cdot log (N) method for Ewald sums in large systems. *The Journal of chemical physics* **1993**, *98*, 10089.
156. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J Chem Phys* **2007**, *126* (1).
157. Berendsen, H.; Postma, J.; Van Gunsteren, W.; Hermans, J., Interaction models for water in relation to protein hydration. *Intermolecular forces* **1981**, *11* (1), 331-342.
158. Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* **2012**, *8* (9), 3314-3321.

159. D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman, AMBER 12. *University of California, San Francisco*. **2012**.
160. Nguyen, H.; Roe, D. R.; Simmerling, C., Improved Generalized Born Solvent Model Parameters for Protein Simulations. *J Chem Theory Comput* **2013**, *9* (4), 2020-2034.
161. Sun, Y.; Xi, W.; Wei, G., Atomic-level study of the effects of O4 molecules on the structural properties of protofibrillar Abeta trimer: beta-sheet stabilization, salt bridge protection, and binding mechanism. *J Phys Chem B* **2015**, *119* (7), 2786-94.
162. Wu, C.; Scott, J.; Shea, J. E., Binding of Congo red to amyloid protofibrils of the Alzheimer Abeta(9-40) peptide probed by molecular dynamics simulations. *Biophys J* **2012**, *103* (3), 550-7.
163. Kabsch, W.; Sander, C., Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22* (12), 2577-637.
164. Zhou, X.; Xi, W.; Luo, Y.; Cao, S.; Wei, G., Interactions of a water-soluble fullerene derivative with amyloid-beta protofibrils: dynamics, binding mechanism, and the resulting salt-bridge disruption. *J Phys Chem B* **2014**, *118* (24), 6733-41.
165. Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd, 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* **2004**, *25* (11), 1400-15.
166. Hilbich, C.; Kisters-Woike, B.; Reed, J.; Masters, C. L.; Beyreuther, K., Substitutions of hydrophobic amino acids reduce the amyloidogenicity of Alzheimer's disease beta A4 peptides. *J Mol Biol* **1992**, *228* (2), 460-73.
167. Reddy, G.; Straub, J. E.; Thirumalai, D., Influence of preformed Asp23-Lys28 salt bridge on the conformational fluctuations of monomers and dimers of Abeta peptides with implications for rates of fibril formation. *J Phys Chem B* **2009**, *113* (4), 1162-72.
168. Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E., How fast-folding proteins fold. *Science* **2011**, *334* (6055), 517-20.
169. Chiti, F.; Dobson, C. M., Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **2006**, *75*, 333-366.
170. Eisenberg, D.; Jucker, M., The amyloid state of proteins in human diseases. *Cell* **2012**, *148* (6), 1188-1203.
171. Fändrich, M.; Schmidt, M.; Grigorieff, N., Recent progress in understanding Alzheimer's β -amyloid structures. *Trends in biochemical sciences* **2011**, *36* (6), 338-345.
172. Hazenberg, B. P., Amyloidosis: a clinical overview. *Rheumatic Disease Clinics* **2013**, *39* (2), 323-345.
173. Sipe, J. D.; Benson, M. D.; Buxbaum, J. N.; Ikeda, S.-i.; Merlini, G.; Saraiva, M. J.; Westermarck, P., Nomenclature 2014: amyloid fibril proteins and clinical classification of the amyloidosis. Taylor & Francis: 2014.
174. Pettersson, T.; Kontinen, Y.; Maury, C., Treatment strategies for amyloid A amyloidosis. *Expert opinion on pharmacotherapy* **2008**, *9* (12), 2117-2128.

175. Gillmore, J. D.; Lovat, L. B.; Persey, M. R.; Pepys, M. B.; Hawkins, P. N., Amyloid load and clinical outcome in AA amyloidosis in relation to circulating concentration of serum amyloid A protein. *The Lancet* **2001**, *358* (9275), 24-29.
176. Lachmann, H. J.; Goodman, H. J.; Gilbertson, J. A.; Gallimore, J. R.; Sabin, C. A.; Gillmore, J. D.; Hawkins, P. N., Natural history and outcome in systemic AA amyloidosis. *New England Journal of Medicine* **2007**, *356* (23), 2361-2371.
177. Westermark, G. T.; Westermark, P., Prion-like aggregates: infectious agents in human disease. *Trends in molecular medicine* **2010**, *16* (11), 501-507.
178. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29* (7), 845-854.
179. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S. a.; Karplus, M., CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry* **1983**, *4* (2), 187-217.
180. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *The Journal of chemical physics* **2007**, *126* (1), 014101.
181. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of chemical physics* **1993**, *98* (12), 10089-10092.
182. Fleming, P. J.; Gong, H.; Rose, G. D., Secondary structure determines protein topology. *Protein Science* **2006**, *15* (8), 1829-1834.
183. Patriksson, A.; van der Spoel, D., A temperature predictor for parallel tempering simulations. *Physical Chemistry Chemical Physics* **2008**, *10* (15), 2073-2077.
184. Rhee, Y. M.; Pande, V. S., Multiplexed-replica exchange molecular dynamics method for protein folding simulation. *Biophysical journal* **2003**, *84* (2), 775-786.
185. Tanaka, M.; Nishimura, A.; Takeshita, H.; Takase, H.; Yamada, T.; Mukai, T., Effect of lipid environment on amyloid fibril formation of human serum amyloid A. *Chemistry and physics of lipids* **2017**, *202*, 6-12.
186. Bjellqvist, B.; Basse, B.; Olsen, E.; Celis, J. E., Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis* **1994**, *15* (1), 529-539.
187. Bjellqvist, B.; Hughes, G. J.; Pasquali, C.; Paquet, N.; Ravier, F.; Sanchez, J. C.; Frutiger, S.; Hochstrasser, D., The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* **1993**, *14* (1), 1023-1031.
188. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M. R.; Appel, R. D.; Bairoch, A., Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook*, Springer: 2005; pp 571-607.
189. Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J., Molecular cell biology 4th edition. *National Center for Biotechnology Information, Bookshelf* **2000**.
190. Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P. R.; Simon, R.; Schubert, D., Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **2009**, *325* (5938), 328-332.
191. Taglialegna, A.; Lasa, I.; Valle, J., Amyloid structures as biofilm matrix scaffolds. *Journal of bacteriology* **2016**, *198* (19), 2579-2588.

192. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K., Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* **2008**, *37*, 215-246.
193. Murphy, M. P.; LeVine III, H., Alzheimer's disease and the amyloid- β peptide. *Journal of Alzheimer's disease* **2010**, *19* (1), 311-323.
194. de Asúa, D. R.; Costa, R.; Galván, J. M.; Filigheddu, M. T.; Trujillo, D.; Cadiñanos, J., Systemic AA amyloidosis: epidemiology, diagnosis, and management. *Clinical epidemiology* **2014**, *6*, 369.
195. Papendick, R.; Munson, L.; O'Brien, T.; Johnson, K., Systemic AA amyloidosis in captive cheetahs (*Acinonyx jubatus*). *Veterinary pathology* **1997**, *34* (6), 549-556.
196. Caughey, B.; Baron, G. S., Are cheetahs on the run from prion-like amyloidosis? *Proceedings of the National Academy of Sciences* **2008**, *105* (20), 7113-7114.
197. Cai, L.; de Beer, M. C.; de Beer, F. C.; van der Westhuyzen, D. R., Serum amyloid A is a ligand for scavenger receptor class B type I and inhibits high density lipoprotein binding and selective lipid uptake. *Journal of Biological Chemistry* **2005**, *280* (4), 2954-2961.
198. Snow, A. D.; Kisilevsky, R.; Stephens, C.; Anastassiades, T., Characterization of tissue and plasma glycosaminoglycans during experimental AA amyloidosis and acute inflammation. Qualitative and quantitative analysis. *Laboratory investigation; a journal of technical methods and pathology* **1987**, *56* (6), 665-675.
199. Westermark, G.; Sletten, K.; Westermark, P., Massive vascular AA-amyloidosis: a histologically and biochemically distinctive subtype of reactive systemic amyloidosis. *Scandinavian journal of immunology* **1989**, *30* (5), 605-613.
200. Wang, W.; Xi, W.; Hansmann, U. H., Stability of the N-Terminal Helix and Its Role in Amyloid Formation of Serum Amyloid A. *ACS omega* **2018**, *3* (11), 16184-16190.
201. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19-25.
202. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **2017**, *14* (1), 71-73.
203. Man, V. H.; He, X.; Derreumaux, P.; Ji, B.; Xie, X.-Q.; Nguyen, P. H.; Wang, J., Effects of all-atom molecular mechanics force fields on amyloid peptide assembly: the case of $\alpha\beta$ 16-22 dimer. *Journal of chemical theory and computation* **2019**, *15* (2), 1440-1452.
204. Berendsen, H. J.; Postma, J. P.; van Gunsteren, W. F.; Hermans, J., Interaction models for water in relation to protein hydration. In *Intermolecular forces*, Springer: 1981; pp 331-342.
205. Best, R. B.; Hummer, G.; Eaton, W. A., Native contacts determine protein folding mechanisms in atomistic simulations. *Proceedings of the National Academy of Sciences* **2013**, *110* (44), 17874-17879.
206. Stillinger, F. H., Water revisited. *Science* **1980**, *209* (4455), 451-457.
207. Rapaport, D., Hydrogen bonds in water: Network organization and lifetimes. *Molecular Physics* **1983**, *50* (5), 1151-1162.
208. Swaminathan, S.; Harte Jr, W.; Beveridge, D. L., Investigation of domain structure in proteins via molecular dynamics simulation: application to HIV-1 protease dimer. *Journal of the American Chemical Society* **1991**, *113* (7), 2717-2721.

209. Luo, J.; Bruice, T. C., Ten-nanosecond molecular dynamics simulation of the motions of the horse liver alcohol dehydrogenase· PhCH₂O⁻ complex. *Proceedings of the National Academy of Sciences* **2002**, *99* (26), 16597-16600.
210. Ghosh, A.; Vishveshwara, S., A study of communication pathways in methionyl-tRNA synthetase by molecular dynamics simulations and structure network analysis. *Proceedings of the National Academy of Sciences* **2007**, *104* (40), 15711-15716.
211. Zhang, F.; Brüschweiler, R., Contact model for the prediction of NMR N– H order parameters in globular proteins. *Journal of the American Chemical Society* **2002**, *124* (43), 12654-12655.
212. Wang, L.; Colón, W., The interaction between apolipoprotein serum amyloid A and high-density lipoprotein. *Biochemical and biophysical research communications* **2004**, *317* (1), 157-161.
213. Van der Hilst, J., Recent insights into the pathogenesis of type AA amyloidosis. *The Scientific World Journal* **2011**, *11*, 641-650.
214. Kundel, F.; De, S.; Flagmeier, P.; Horrocks, M. H.; Kjaergaard, M.; Shamma, S. L.; Jackson, S. E.; Dobson, C. M.; Klenerman, D., Hsp70 inhibits the nucleation and elongation of tau and sequesters tau aggregates with high affinity. *ACS chemical biology* **2018**, *13* (3), 636-646.
215. Fernandez-Funez, P.; Sanchez-Garcia, J.; De Mena, L.; Zhang, Y.; Levites, Y.; Khare, S.; Golde, T. E.; Rincon-Limas, D. E., Holdase activity of secreted Hsp70 masks amyloid-β42 neurotoxicity in Drosophila. *Proceedings of the National Academy of Sciences* **2016**, *113* (35), E5212-E5221.
216. Rajamäki, K.; Nordström, T.; Nurmi, K.; Åkerman, K. E.; Kovanen, P. T.; Öörni, K.; Eklund, K. K., Extracellular acidosis is a novel danger signal alerting innate immunity via the NLRP3 inflammasome. *Journal of Biological Chemistry* **2013**, *288* (19), 13410-13419.
217. Kato, Y.; Ozawa, S.; Miyamoto, C.; Maehata, Y.; Suzuki, A.; Maeda, T.; Baba, Y., Acidic extracellular microenvironment and cancer. *Cancer cell international* **2013**, *13* (1), 89.
218. Mongan, J.; Case, D. A.; McCammon, J. A., Constant pH molecular dynamics in generalized Born implicit solvent. *Journal of computational chemistry* **2004**, *25* (16), 2038-2048.
219. Lee, M. S.; Salsbury Jr, F. R.; Brooks III, C. L., Constant-pH molecular dynamics using continuous titration coordinates. *Proteins: Structure, Function, and Bioinformatics* **2004**, *56* (4), 738-752.
220. Khandogin, J.; Brooks, C. L., Linking folding with aggregation in Alzheimer's β-amyloid peptides. *Proceedings of the National Academy of Sciences* **2007**, *104* (43), 16880-16885.
221. Jayaraman, S.; Gantz, D. L.; Haupt, C.; Gursky, O., Serum amyloid A forms stable oligomers that disrupt vesicles at lysosomal pH and contribute to the pathogenesis of reactive amyloidosis. *Proceedings of the National Academy of Sciences* **2017**, *114* (32), E6507-E6515.

Appendix I: Chapter 3 Supplemental

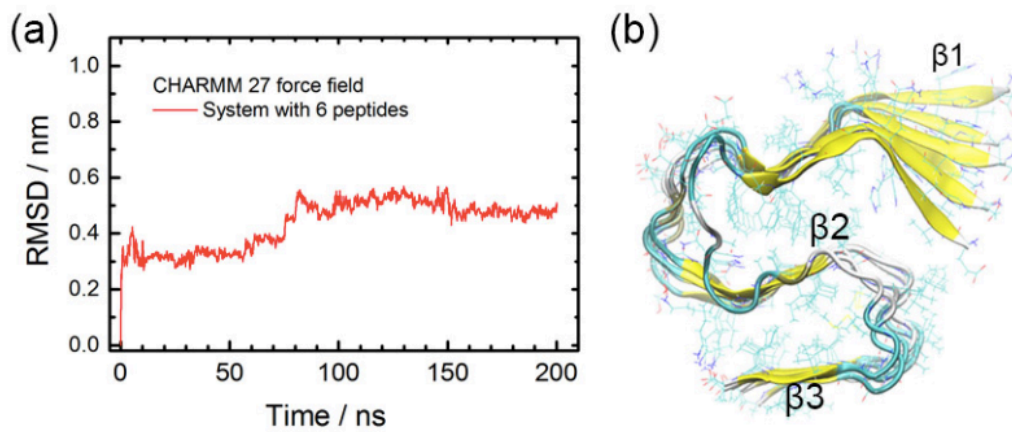


Figure S3.1. (a) The rmsd of six-chain fragment with force field CHARMM27. (b) representative conformation after 200ns.

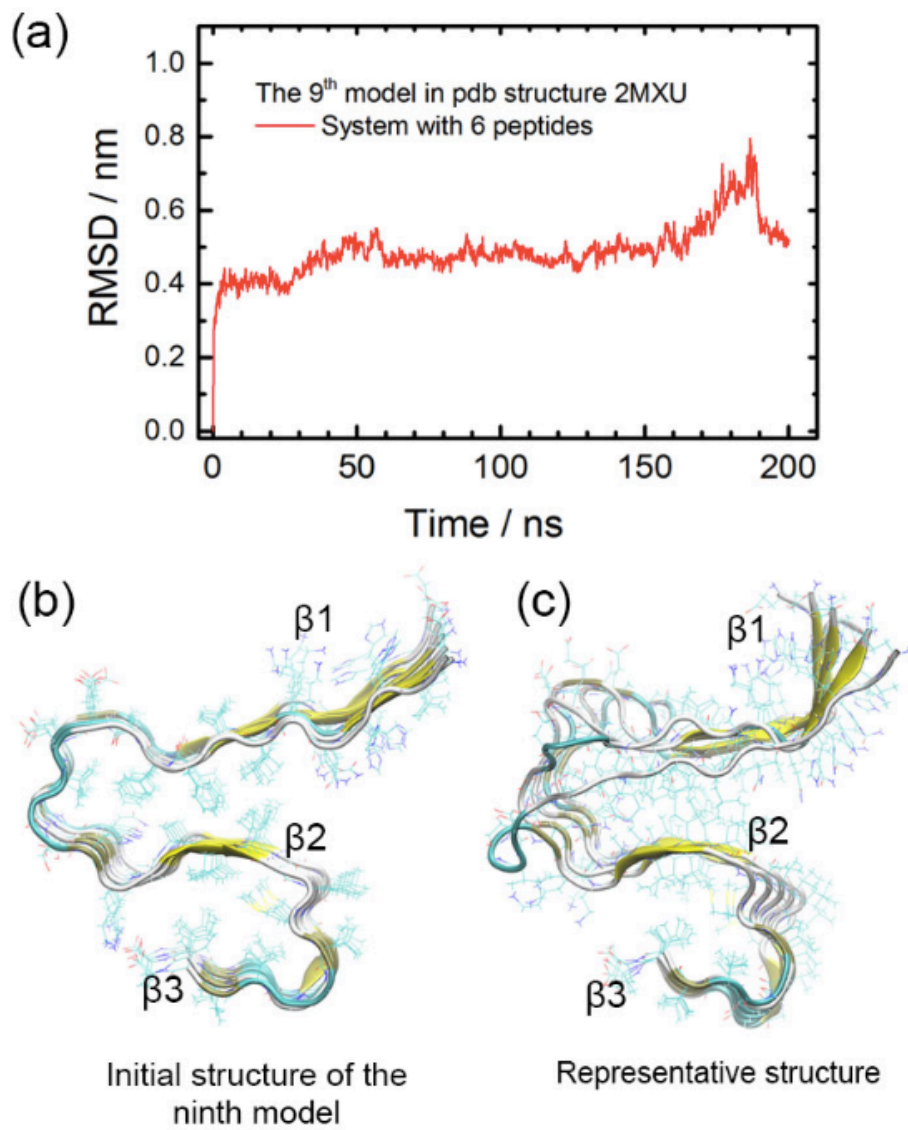


Figure S3.2. (a) The rmsd of six-chain system starting from the ninth model of pdb structure 2MXU. (b) initial conformation of the ninth model. (c) representative conformation after 200ns.

Appendix II: Chapter 4 Supplemental

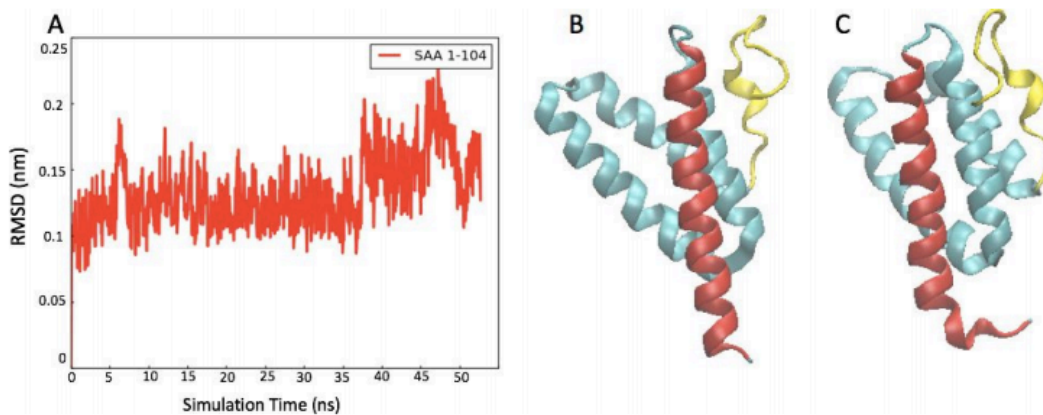


Figure S4.1. Root-mean-square-deviation (RMSD) to the crystal structure for the full-sized SAA protein as function of time (A). The initial configuration is shown in (B) and the final one in (C), where helix I is marked in red, helix II-IV in cyan, and the C terminal tail in yellow. Data are from a 40ns molecular dynamic simulation at 310 K, using the set-up described in the manuscript.

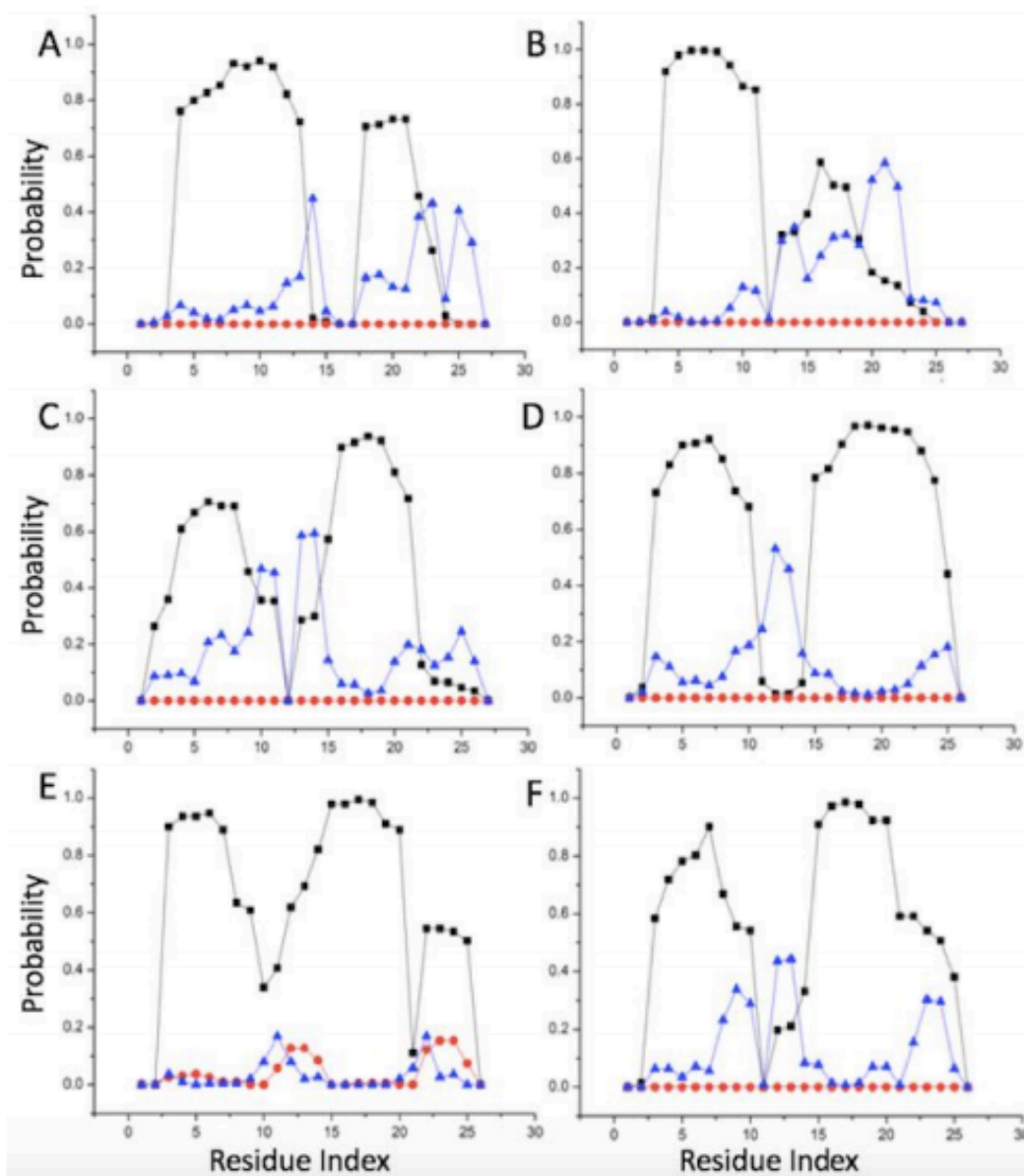


Figure S4.2. A and B, Secondary structure probability for the box size of 4.8 nm. C and D, Secondary structure probability for the box size of 5.4 nm. E and F, Secondary structure probability for the box size of 5.6 nm. The frequency of α -helices is drawn in black, that of turns in blue, and of β -strands in red.

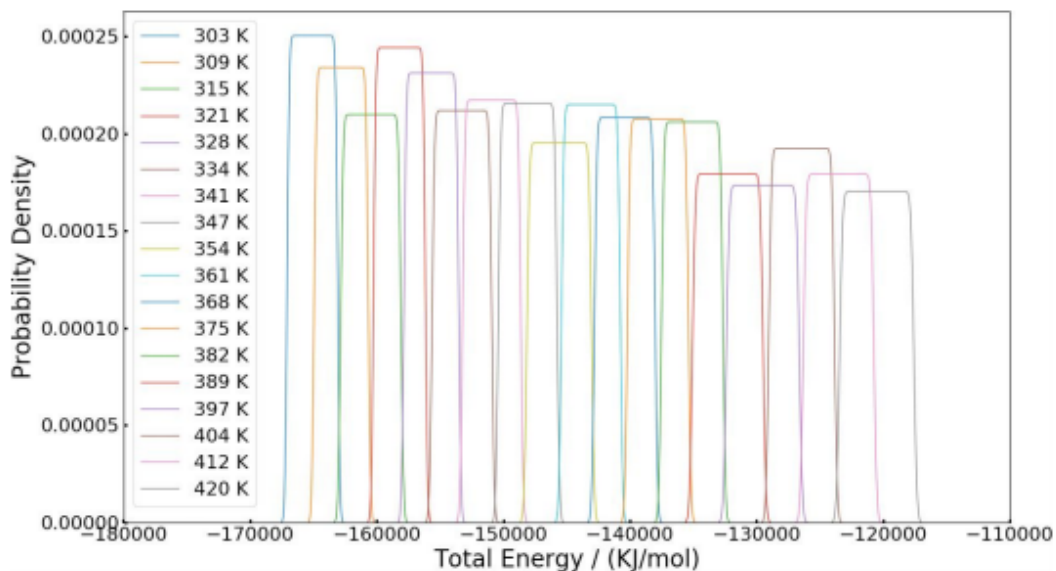


Figure S4.3. Distribution of potential energies for the 36 replicas distributed in a temperature interval between 300K and 420 K. Only every second replica is shown.

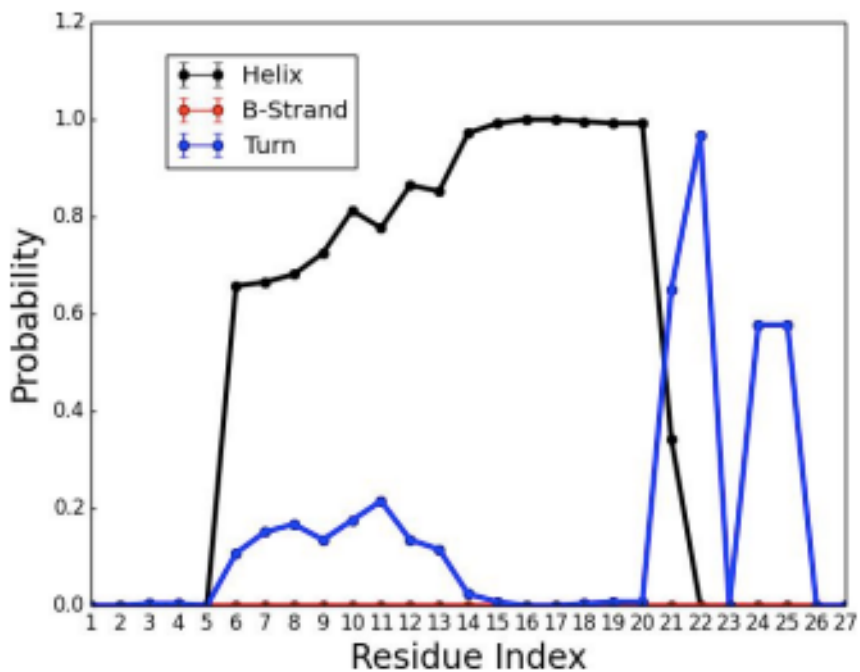


Figure S4.4. Secondary structure probability analysis for molecular dynamics simulation of the mutation E9A based on the straight structure from wild type.

Appendix III: Chapter 5 Supplemental

Table S5.1. Solvent Accessible Surface Area (SASA) per residue averaged over all chains in the SAA₁₋₇₆ and SAA₁₋₁₀₄ hexamers. We show both the values for the individual trajectories and for the resulting averages (with standard deviation listed in parenthesis). For comparison we list also the respective values for the free monomers.

system	Hexamer	Monomer
SAA ₁₋₇₆	Run-1	77.6
	Run-2	72.7
	Run-3	79.1
	Average	76 (3)
SAA ₁₋₁₀₄	Run-1	69.0
	Run-2	68.7
	Run-3	69.1
	Average	69.2 (0.4)

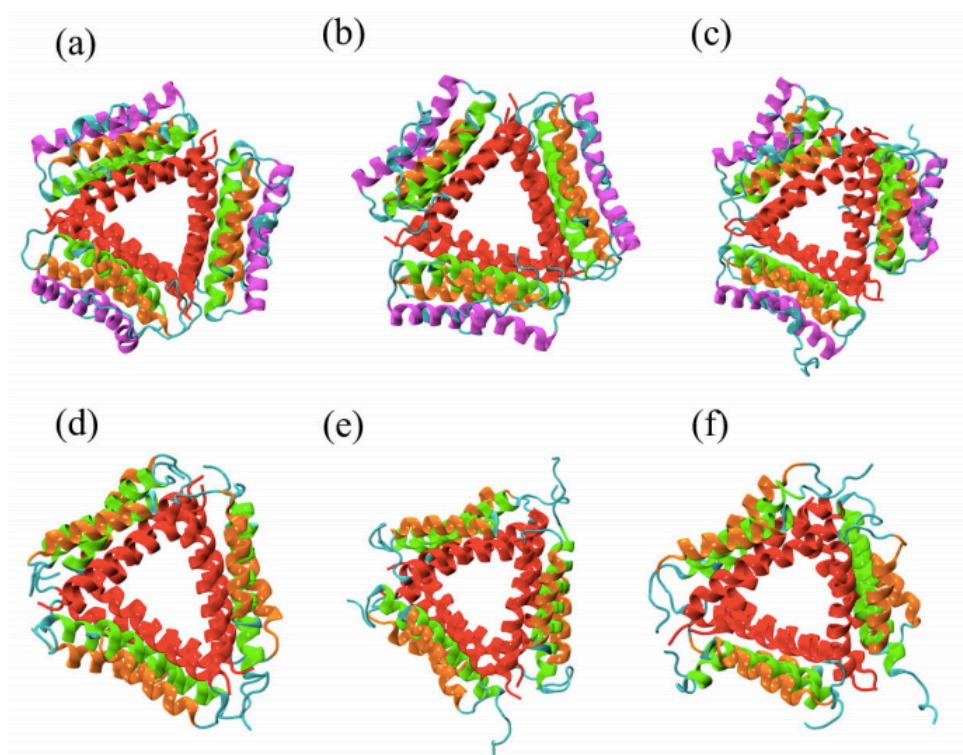


Figure S5.1. The snapshots of (a-c) SAA_{1-104} and (d-f) SAA_{1-76} hexamers as obtained at the end of each of the three trajectories. The trajectories are shown as follows: first column (Run-1), second column (Run-2), and third column (Run-3). Color coding of the helices: helix-I (red), helix-II (orange), helix-III (green), and helix-IV (magenta).

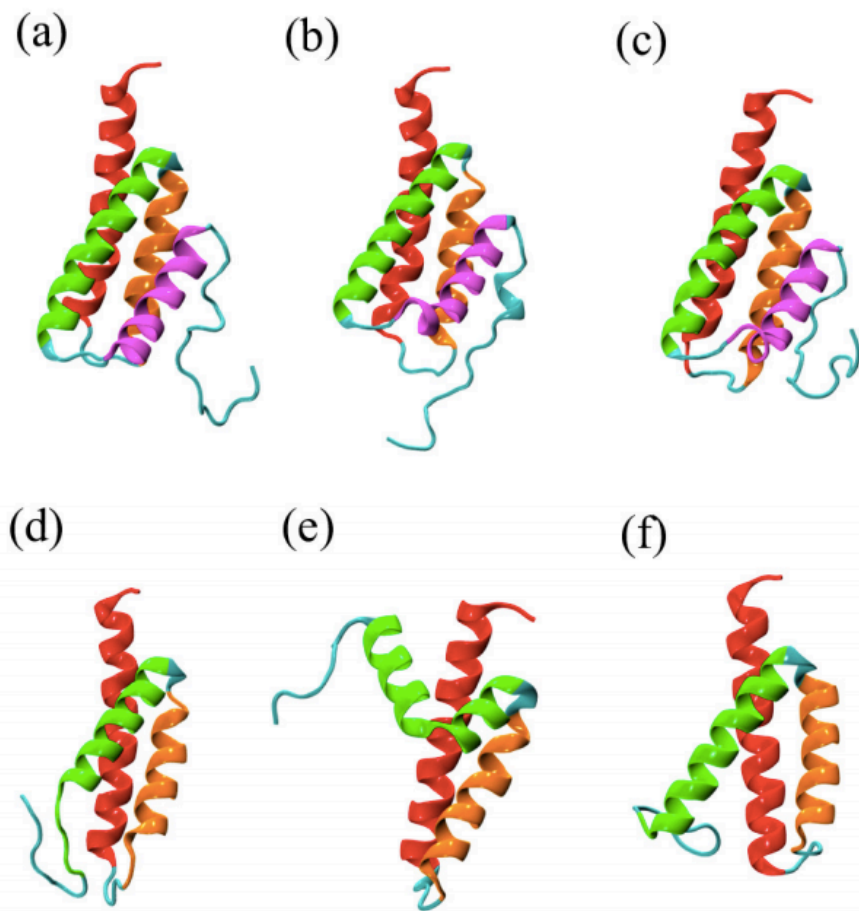


Figure S5.2. The snapshots of (a-c) SAA₁₋₁₀₄ and (d-f) SAA₁₋₇₆ monomers as obtained at the end of each of the three trajectories. The trajectories and color coding of the helices are same as that of Figure 5.1.

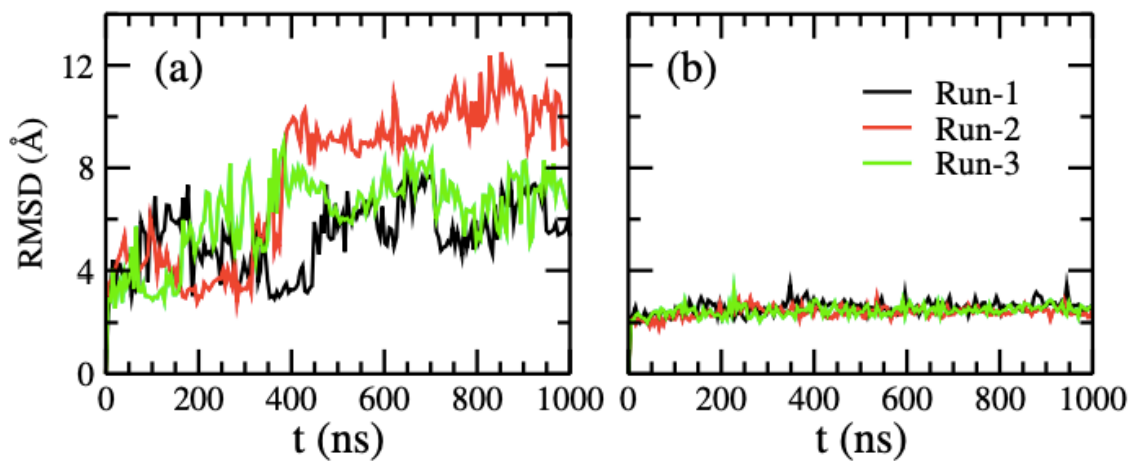


Figure S5.3. Time evolution of the root mean square deviation (RMSD) of isolated (a) SAA₁₋₇₆ and (b) SAA₁₋₁₀₄ monomers, taking into account all non-hydrogen atoms in the first 76 residues. Data are shown for all three trajectories of each system. The RMSD values are calculated with respect to the corresponding start configurations.

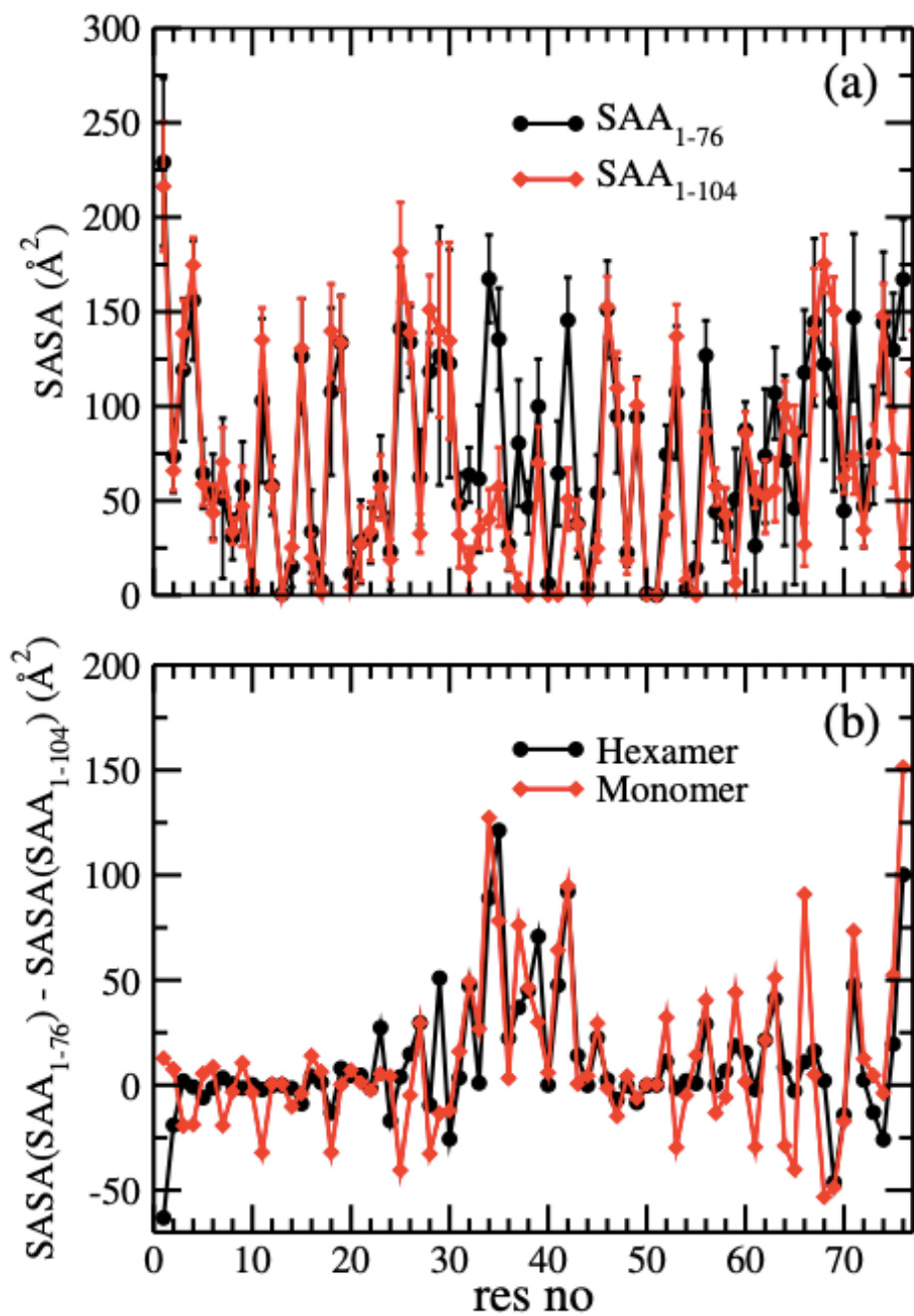


Figure S5.4. (a) Average residue-wise solvent accessible surface area (SASA) of SAA₁₋₇₆ and SAA₁₋₁₀₄ monomers as calculated over the last 500 ns of each of the three trajectories of each system. Data are shown only for the first 76 residues. In (b) we show for each residue the difference $\Delta\text{SASA} = \text{SASA}(\text{SAA}_{1-76}) - \text{SASA}(\text{SAA}_{1-104})$ calculated either for the isolated monomer or measured in the corresponding hexamers.

Appendix IV: Copyrights

6/30/2020

RightsLink Printable License

SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Jun 30, 2020

This Agreement between University of Oklahoma -- wenhua wang ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4858991326808
License date	Jun 30, 2020
Licensed Content Publisher	Springer Nature
Licensed Content Publication	The Protein Journal
Licensed Content Title	Structure and Function of Alzheimer's Amyloid β Proteins from Monomer to Fibrils: A Mini Review
Licensed Content Author	Nikhil Agrawal et al
Licensed Content Date	Jul 19, 2019
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of	1

Copyright for Figure 1.2.

Annual Reviews, Inc. - License Terms and Conditions

This is a License Agreement between Wenhua Wang ("You") and Annual Reviews, Inc. ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Annual Reviews, Inc., and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date	30-Jun-2020	Type of Use	Republish in a thesis/dissertation
Order license ID	1045388-1	Publisher Portion	ANNUAL REVIEWS
ISSN	1553-4014		Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Annual review of pathology	Rightholder	Annual Reviews, Inc.
Date	12/31/2005	Publication Type	e-Journal
Language	English	URL	http://arjournals.annualreviews.org/loi/pathmechdis
Country	United States of America		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	United States
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
Format (select all that apply)	Print, Electronic	Copies for the disabled?	No
Who will republish the content?	Academic institution	Minor editing privileges?	No
Duration of Use	Life of current and all future editions	Incidental promotional use?	No
Lifetime Unit Quantity	Up to 499	Currency	USD
Rights Requested	Main product		

Copyright for Figure 1.3.

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Jul 01, 2020

This Agreement between University of Oklahoma -- wenhua wang ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number 4860361009047

License date Jul 01, 2020

Licensed Content Publisher John Wiley and Sons

Licensed Content Publication EMBO Reports

Licensed Content Title Cellular mechanism of fibril formation from serum amyloid A1 protein

Licensed Content Author Marcus Fändrich, Thomas Simmet, Tatiana Syrovets, et al

Licensed Content Date Jun 21, 2017

Licensed Content Volume 18

Licensed Content Issue 8

Licensed Content Pages 15

Type of use Dissertation/Thesis

Copyright for Figure 2.4.



RightsLink®



Home



Help



Email Support



wenhua wang ▾

Stability of a Recently Found Triple- β -Stranded A β 1-42 Fibril Motif**Author:** Wenhui Xi, Wenhua Wang, Gabrielle Abbott, et al**Publication:** The Journal of Physical Chemistry B**Publisher:** American Chemical Society**Date:** May 1, 2016*Copyright © 2016, American Chemical Society***PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE**

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)[CLOSE WINDOW](#)

Copyright for Chapter 3



RightsLink®



Home



Help



Email Support



wenhua wang ▾

Cleavage, Downregulation, and Aggregation of Serum Amyloid A

Author: Wenhua Wang, Prabir Khatua, Ulrich H. E. Hansmann

Publication: The Journal of Physical Chemistry B

Publisher: American Chemical Society

Date: Feb 1, 2020

Copyright © 2020, American Chemical Society

ACS Publications

Most Trusted. Most Cited. Most Read.

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)[CLOSE WINDOW](#)**Copyright for Chapter 5**