Investigation on the conformational conversion of human prion protein by molecular dynamics simulation

Yang, Hong

2017

Yang, H. (2017). Investigation on the conformational conversion of human prion protein by molecular dynamics simulation. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/70177>

<https://doi.org/10.32657/10356/70177>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**INVESTIGATION ON THE CONFORMATIONAL
CONVERSION OF HUMAN PRION PROTEIN BY
MOLECULAR DYNAMICS SIMULATION**

YANG HONG

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

2017

**INVESTIGATION ON THE CONFORMATIONAL
CONVERSION OF HUMAN PRION PROTEIN BY
MOLECULAR DYNAMICS SIMULATION**

YANG HONG

YANG HONG

School of Physical and Mathematical Sciences

A thesis submitted to the Nanyang Technological University

in fulfillment of the requirement for the degree of

Doctor of Philosophy of Science

2017

Acknowledgements

I would like to express my earnest appreciation to all the people whose guidance, support and encouragement have made this thesis possible.

Firstly, I would like to sincerely and gratefully thank my supervisor, Asst. Prof. Zhang Dawei, for his guidance, motivation and support during my Ph.D studies. This thesis would unlikely to be successfully completed without his patient guidance and valuable advice. His expertise and sharing are of great benefit in helping me grow as a computational chemist.

My sincere gratitude also goes to Assoc. Prof. Li Tianhu for his support and kindness on supervising me in the final year.

I would also like to thank my lab mates in Dr. Zhang's group, Cui Jinglan, Yip Yew Mum, Dr Wei Caiyi, Dr. Sun Tiedong, Dr. Raudah Lazim, and Dr. Duan Rui. Thanks to Caiyi for helping me perform my first simulation with Amber Program. Thanks to Tiedong for his patience in helping me solve computer problems. Thanks for all the inspiring discussions, kind assistance, joys and friendships have ever had with them. Heartfelt thanks to all my friends in Singapore for sharing their living and studying experiences with me.

Last but not least, I would like to express a special thanks to my parents and my brother. No words can express my grateful feelings for their constant unconditional love, encouragement and support throughout my life. Thank you!

Contents

Acknowledgements	1
Contents	2
Abstract	5
List of abbreviations	6
List of Figures	8
Chapter 1 Introduction	13
1.1 A brief introduction of prion diseases	13
1.2 The cellular prion protein, PrP ^C	14
1.2.1 The C-terminal domain of PrP ^C	16
1.2.2 The N-terminal domain of PrP ^C	19
1.3 The scrapie prion protein, PrP ^{Sc}	24
1.4 Molecular Dynamics Studies on PrP ^C to PrP ^{Sc} Conversion	29
1.5 Organization of this thesis	34
Chapter 2 pH Effect on the Conformational Conversion of Human Prion Protein	35
2.1 Introduction	35
2.2 Methods.....	42

2.2.1	Structure Preparation for MD simulation.....	42
2.2.2	Molecular Dynamics Simulation	43
2.2.3	Analysis Methods.....	44
2.3	Results and Discussions	44
2.3.1	Stability of Helix 1 (H1)	47
2.3.2	Stability of Helix 2 (H2)	52
2.3.3	Stability of Helix 3 (H3)	55
2.3.4	Stability of β -sheets S1S2	57
2.4	Conclusion	59
 Chapter 3 The C-terminal of H2 Helix in the Prion Protein Induced α-helix to β-sheet Conformational Changes		
		60
3.1	Introduction	60
3.2	METHODS	64
3.2.1	Structure Preparation for MD simulation.....	64
3.2.2	Molecular Dynamics Simulation	64
3.2.3	Simulated annealing simulation	65
3.2.4	Analysis Methods.....	66
3.3	Results and Discussions.....	67
3.3.1	Speeding up the unfolding by high temperature simulation	67
3.3.2	Formation of new β -strand S3.....	73
3.3.3	Probing the role of disulfide bond in the conformational transition	

.....	77
3.4 Conclusion	80
Chapter 4 Probing the β constitution at the N-terminus of prion protein..	82
4.1 Introduction	82
4.2 Methods	86
4.2.1 Polarizable Structure-Specific Backbone Charge Model.....	86
4.2.2 Replica Exchange Molecular Dynamics methodology	88
4.2.3 Conventional MD simulations	90
4.2.4 Analysis Methods.....	91
4.3 Results and Discussions.....	91
4.3.1 Formation of four-stranded β -sheet structure.....	92
4.3.2 Formation of five-stranded β -sheet structure	101
4.3.3 Formation of six-stranded β -sheet structure.....	108
4.4 Conclusion	114
Chapter 5 Summary and future directions	117
Publications.....	122
Appendix.....	123
References	124

Abstract

Conformational conversion of normal α -rich cellular prion protein PrP^{C} to its oligomeric β -rich isoform PrP^{Sc} can lead to fatal prion diseases. Despite numerous studies, the underlying mechanism for the β -enrichment during the conversion still remains elusive. In this molecular dynamics (MD) study, firstly, we examined the influence of pH on the PrP C-terminal domain and suggested that acidic pH can facilitate the denaturation of PrP which starts from H2 helix. After further conformational changes in the elevated temperature simulation and simulated annealing simulation, a new β -strand, named as S3, is formed at the denatured region of H2 and aligns antiparallel with native S2 β -strand, leading to a stable three-stranded β -sheet structure in the PrP C-terminal domain. With a combination of polarizable structure-specific backbone charge (PSBC) model and replica exchange molecular dynamics (REMD) simulation, we simulated the folding of three N-terminal fragment peptides. Three new β -strands are found. By incrementally modeling the newfound β -strand onto the previous three-stranded β -sheet structure, we obtained a novel β -rich conformer with six-stranded antiparallel β -sheet spanning both the C-terminal domain and the N-terminal amyloidogenic region of PrP, which might serve as the β -core of the amyloid fibrils and can provide important insights into the mechanism of the conformational conversion of PrP^{C} to PrP^{Sc} .

List of abbreviations

AD Alzheimer's disease

BSE Bovine spongiform encephalopathy

CD Circular dichroism

CJD Creutzfeldt-Jakob disease

CWD Chronic wasting disease

DSSP Dictionary of secondary structure of proteins

FFI Fatal familial insomnia

GPI Glycosylphosphatidylinositol

GSS Gerstmann-Straussler-Scheinker syndrome

HD Huntington's disease

MD Molecular Dynamics

MM Molecular Mechanics

NMR Nuclear magnetic resonance

PCA Principal Component Analysis

PDB Protein Data Bank

PPC Polarized Protein-specific Charge

PSBC Polarizable Structure-Specific Backbone Charge

PrP^C Cellular Prion Protein

PrP^{Sc} Scrapie Prion Protein

QM Quantum Mechanics

REMD Replica Exchange Molecular Dynamics

RESP Restrained Electrostatic Potential

RMSD Root mean square deviation

RMSF Root mean square fluctuation

TSE Transmissible Spongiform Encephalopathies

List of Figures

Figure 1.1: Structural illustration of prion protein.....	16
Figure 1.2: Model of prion protein showing the binding sites of metal ions.....	20
Figure 1.3: Structure models of prion aggregates.....	28
Figure 2.1: NMR structure of C-terminal domain of human prion protein.....	39
Figure 2.2: RMSD evolutions of human PrP (residues 128-228) in simulations of different pH levels.....	45
Figure 2.3: RMSF evolutions of human PrP (residues 125-228) in simulations of different pH levels.....	46
Figure 2.4: Secondary structure evolution of H1 helix by DSSP analysis of simulations at different pH levels and the final structures in each simulation...	48
Figure 2.5: Fractional α -helical content evolutions of H1 helix in simulations at different pH levels.....	49
Figure 2.6: Long-range salt-bridges identified connecting H1 and H2H3 hydrophobic domain in the simulations at different pH levels.....	51
Figure 2.7: Secondary structure evolution of H2 helix by DSSP analysis of simulations at different pH levels and the final structures in each simulation...	53
Figure 2.8: Fractional α -helical content evolutions of H2 helix in simulations at different pH levels.....	54

Figure 2.9: Variations in radius of gyration of H2 helix in simulations at different pH levels.....	55
Figure 2.10: Secondary structure evolution of H3 helix by DSSP analysis of simulations at different pH levels and the final structures in each simulation...	56
Figure 2.11: Fractional α -helical content evolutions of H3 helix in simulations at different pH levels.....	57
Figure 2.12: Secondary structure evolution of S1 β -strand by DSSP analysis of simulations at different pH levels and the final structures in each simulation...	58
Figure 3.1: Simulated annealing protocol.....	66
Figure 3.2: (A) Secondary structure evolution of huPrP by DSSP analysis of simulation at high temperature and acidic pH conditions. (B) Starting and final structures in the simulation.....	69
Figure 3.3: Essential dynamics of prion protein identified by the second eigenvector of PCA analysis.....	71
Figure 3.4: Representative structures given by the clusters analysis of the last 100ns of high temperature simulation.....	72
Figure 3.5: (A) Secondary structure evolution of huPrP by DSSP analysis of the simulated annealing simulation. (B) Hydrogen bonds detected between the newly formed S3 and native S2 β -strands.....	74
Figure 3.6: Secondary structure evolution of three-stranded β -sheet structure by DSSP analysis of the conventional simulation.....	75

Figure 3.7: (A) H-bonds detected between S2 and S3 in the conventional MD simulation of three-stranded β -sheet structure. (B) Distances of H-bonds over simulation time.....	76
Figure 3.8: Secondary structure evolution of reduced huPrP by DSSP analysis of the conventional simulation.....	79
Figure 4.1: Secondary structure evolution of PrP(116-131) by DSSP analysis under different conditions.....	95
Figure 4.2: Backbone hydrogen bonds between S0 and S1 β -strands in the crystal structure by Jan Steyaert et al.....	96
Figure 4.3: Distances variations over simulation time of the five hydrogen bonds calculated based on the 262.00 K trajectory of PrP(116-131) using PSBC model.....	96
Figure 4.4: (A): Representative structure of PrP(116-131). (B): Backbone hydrogen bonds detected in the representative structure. (C): Illustration of the hydrogen bonds in the representative structure.....	98
Figure 4.5: (A) Illustration of the three-stranded β -sheet structure of PrP118-228 in the crystal structure. (B) Illustration of the constructed four-stranded β -sheet structure of PrP118-228 in our simulation.....	100
Figure 4.6: Secondary structure evolution of the four-stranded β -sheet structure of PrP118-228 by DSSP analysis of the conventional MD simulation.....	100
Figure 4.7: Secondary structure evolution of PrP(105-122) by DSSP analysis of	

the trajectories with PSBC model at 264.00 K and 298.17K	103
Figure 4.8: Representative structures of three most populated clusters of PrP(105-122) by the cluster analysis of 264.00 K trajectory using PSBC model.....	104
Figure 4.9: Backbone hydrogen bonds detected in the representative structure in Cluster C1 of PrP(105-122).....	105
Figure 4.10: Backbone hydrogen bonds detected in the representative structure in Cluster C2 of PrP(105-122).....	105
Figure 4.11: Upper panel: Secondary structure evolution of the five-stranded β -sheet structure of PrP104-228 by DSSP analysis of the conventional MD simulation. Lower panel: The snapshot of final structure in the simulation.....	107
Figure 4.12: Secondary structure evolution of PrP(95-112) by DSSP analysis of the trajectories with PSBC model at 265.00K and 295.66K	109
Figure 4.13: Representative structures of three most populated clusters of PrP(95-112) by the cluster analysis of the trajectory with PSBC model at 265.00 K.....	110
Figure 4.14: Backbone hydrogen bonds detected in the representative structure in Cluster C1 of PrP(95-112).....	110
Figure 4.15: Upper panel: Secondary structure evolution of the six-stranded β -sheet structure of PrP93-228 by DSSP analysis of the conventional MD simulation. Lower panel: The snapshot of final structure in the simulation.....	112

Figure 5.1: (A) NMR structure of PrP125-228. (B) Unfolding started from the H2 C-ter (C) Further unfolding at H2 C-ter and H3 N-ter. (D) H2H3 loop aligns closely to native S1S2. (E) Three-stranded antiparallel β -sheet structure of PrP125-228. (F) Four-stranded antiparallel β -sheet structure of PrP118-228. (G) Five-stranded antiparallel β -sheet structure of PrP104-228. (H) Six-stranded antiparallel β -sheet structure of PrP93-228..... 120

Chapter 1 Introduction

1.1 A brief introduction of prion diseases

Prion diseases, also known as transmissible spongiform encephalopathies (TSEs)¹, are a class of lethal neurodegenerative disorders in human beings and a wide range of mammalian animals. Human prion diseases include Creutzfeldt-Jakob disease (CJD), Gerstmann-Straussler-Scheinker syndrome (GSS), fatal familial insomnia (FFI) and kuru. In animals, there are bovine spongiform encephalopathy (BSE, also known as mad cow disease) in cattle, scrapie in sheep, chronic wasting disease (CWD) in deer and elk.²⁻⁴ Prion diseases share some similar pathophysiological and morphological characteristics with other progressive encephalopathies, such as Alzheimer's disease (AD), Parkinson's disease (PD) and Huntington's disease (HD).⁵ They are commonly associated with the accumulation of pathogenic misfolded isoform of prion protein, leading to the formation of intracellular or extracellular protein amyloid fibrils or aggregates. However, prion diseases are considered unique in that there seems no informational nucleic acid genome which encodes for the progeny, and the infectious isoform of prion protein can propagate itself.⁶ The unique biology, along with the public health risks posed by prion diseases has attracted

considerable attention on prion protein. Prions have also been adopted as models for providing a comprehensive evaluation of other neurodegenerative diseases involving accumulation of misfolded host proteins.

The infectious agent responsible for prion diseases was recognized in the 1960's.⁷ The agent was termed as “prion” by Stanley Prusiner in 1982, which stands for proteinaceous infectious particle.⁸ In diseased humans and animals, prion protein exists in two forms, the physiological, normal cellular form named as PrP^C, and the pathological scrapie isoform indicated as PrP^{Sc}. Since then, enormous efforts have been devoted to investigating the biophysical and biochemical properties of the prion protein and unravelling the mechanism of prion pathogenesis. Numerous research studies have provided evidence that the misfolding and aggregation of healthy PrP^C to pathological PrP^{Sc} is the key event for the development of prion diseases.⁹ According to the widely accepted protein-only hypothesis, PrP^{Sc} is the major component of the infectious agent and it can “replicate” in the absence of nucleic acid by templating the misfolding of other healthy PrP^C in an autocatalytic manner.¹⁰ Detailed structural studies on both the two isoforms are of great significance in the prion biology since they may clarify the molecular mechanisms responsible for the prion pathologies.

1.2 The cellular prion protein, PrP^C

The cellular prion protein (PrP^C) is highly present in the central nervous system

and various tissues of humans as well as animals¹¹, and has around 90% similarity across mammals.^{12,13} It is a glycosylphosphatidylinositol anchored (GPI-anchored) glycoprotein which is located on the outer surface of the neurons cells.¹ In human prion protein, the immature PrP^C is a long polypeptide composed of 253 amide acids. During the protein maturation and translocation, the N-terminal endoplasmic reticulum signal peptide and the C-terminal GPI-anchoring peptide are cleaved off. The remained mature functional protein is 209 amid acids (residue 23-231) in length and is attached to the extracellular cell membrane by a GPI anchor.¹⁴⁻¹⁶

Structural studies at the atomic level have been hampered by the difficulties in obtaining enough quantities of highly purified materials and relatively low-expression levels determination method. So far, most structural information on PrP^C has come from recombinant prion protein (rec PrP^C) in bacteria or chemical synthesis. Nuclear magnetic resonance¹⁷⁻²¹ (NMR) and X-ray crystallography^{22,23} studies on more than 20 different species²⁴ and mutants^{25,26} revealed that PrP^C shares a highly conserved fold of two distinct structural regions, including a largely disordered N-terminal and a well-structured C-terminal globular domain (Figure 1.1)²⁷.

Figure 1.1: Structural illustration of prion protein

1.2.1 The C-terminal domain of PrP^C

Various structures of the globular C-terminal domain of prion protein have been solved at atomic level by solution NMR and X-ray crystallographic techniques. These structures are from a range of mammalian animals, such as mouse¹⁹, sheep²³, bovine¹⁷, elk²¹, horse²⁶, humans²⁷, and so on. Collectively, these structures share a consistent fold of three α -helices (H1, H2 and H3) and a short two-stranded antiparallel β -sheets (S1 and S2). H2 and H3 form the hydrophobic bulk and are linked by a characteristic disulfide bond. Taking human prion protein as an example, the C-terminal domain (residue 125 to 228) consists of three α -helices (H1: residue 144-156, H2: residue 174-194, H3: residue 200-228) and a short two-stranded antiparallel β -sheets (S1: residue 128-131, S2: residue

161-164).²⁷ H2 and H3 are linked by a disulfide bond between Cys179 and Cys214. Defining the particular role of each secondary structure domain in the conversion to a pathological form or partially unfolded intermediates is one of major research interests in unraveling the conformational conversion of PrP^C to PrP^{Sc}.

The first helix H1 is characterized by a highly unusual sequence. Unlike normal α -helices and β -sheets consisting of most hydrophobic residues²⁸, it is composed of hydrophilic residues (DYEDR YYREN MHR, residues 144-156 in human prion protein), so it lacks the ability to form stabilizing hydrophobic contacts with the protein core²⁹. The protonation state of His155 at different pH conditions may affect the structure of H1 C-terminus³⁰. Even though a number of studies implicated that H1 may be the initial site of the protein misfolding³¹⁻³³, NMR and CD spectroscopy studies with synthetic peptides of H1 demonstrated that H1 has remarkably high intrinsic α -helical propensity³⁴, and shows high stability under a wide range of environmental conditions³⁵. More and more studies indicated that H1 may not take part in the early conformational changes, but may refold in the late stage as involved in global structure rearrangements.

The H2 and H3 helix form the hydrophobic bulk of the protein and are linked by a characteristic disulfide bond. H2 and H3 are often studied as an integrity domain. It is reported that the isolated H2H3 domain can independently undergo glycosylations and glycosylphosphatidylinositol anchoring in a way similar to that

of full-length prion protein³⁶. The H2H3 fragment can form proteinase K-resistant aggregates and reproduce the oligomerization pathways of full-length prion protein³⁷. H2H3 undergoes structural transformation early during the aggregation of recombinant PrP, indicating to be the seeding sites of the pathogenic conversion³⁸. It is hypothesized that H2H3 might be the minimal region involved in the conformation of PrP³⁶.

The amino acid sequence at H2 C-terminus(TVTTTT, residues 188–193 in human prion protein) is very unusual since it has a high propensity for β -sheet formation. The H2 C-terminus also exhibits the most structural perturbation in a number of experimental³⁹⁻⁴² and theoretical^{41,43-46} studies. This region has been shown to form β -conformation or amyloid fibrils^{41,42,47} and to be neurotoxic in vitro⁴⁸. Thus, the H2 C-terminus may play an early role in the conformational conversion, but this role is still not elucidated.

Most of the pathogenic mutations⁴⁹ that have been associated with the prion diseases cluster in the C-terminal domain, mainly in H2, H3 and the loop between them. To date, available solved structures of mutants by NMR⁵⁰⁻⁵² and X-ray crystallography⁵³ show minor structural differences. None of these mutations causes dramatic changes in the globular conformation. Some mutations⁵⁴ may affect the side chain interaction of the intermolecular salt-bridges or modify the structural contacts between H2 and H3, leading to the disturbance of the normal folding pathway of prion protein.

1.2.2 The N-terminal domain of PrP^C

The highly flexible N-terminus of PrP features the glycine-rich octarepeat region (OR), which is evolutionarily conserved among different mammalian species.⁵⁵ In human prion protein, there are four perfect OR repeats of PHGGGWGQ (residues 60-91) and one homologous “pseudorepeat” of PQGGGGWGQ (residues 51-59). Within the PHGGGWGQ repeats, the His and Trp residues were found to be possible binding sites to a number of metal ions, including manganese⁵⁶, zinc⁵⁷⁻⁵⁹, and copper ions^{58,60-62} (see the copper binding sites in Figure 1.2).

Its ability to bind copper ions has attracted many research interests.⁶³ Numerous studies have been done to identify the role of OR regions in protein function and the disease pathogenesis, which is complicated and controversial. An X-ray study⁶⁴ with the HGGGW peptide fragment showed that this region can be wrapped into a β -turn configuration by the coordination to Cu²⁺ ion. Zahn⁶⁵ reported that the HGGGW residues adopt a loop conformation and the GWGQ residues can form a β -turn-like structure in a pH-dependent aggregation. However, the PrP isolated from normal and pathological brain cells is not necessarily the full-length form, but a truncated form^{66,67}. The truncated molecule contains the C-terminus and the amyloidogenic portion of N-terminal, but lacks the OR regions, indicating that the OR regions is not necessarily essential for the prion infectivity. Although the OR region might not be necessarily required for the β -rich core of the infectious pathogen PrP^{Sc}, it is still hypothesized to play a role in the

pathological process of the prion protein.

Figure 1.2 Model of human prion protein showing the binding sites of metal ions.

Another region which contains functional subdomains associated with copper coordination is a positively charged segment (residues 96-111) in human prion protein) at N-terminus of prion protein (see the copper binding sites in Figure 1.2).⁶⁸ The low-occupancy binding form with Cu^{2+} bridging to His96 and His111 is regarded as the most likely route to induce local β -sheet formation. A Circular dichroism (CD) spectroscopy study⁶⁹ on the PrP(91–115) synthetic peptide observed an increase in β -sheet formation in the low-occupancy coordination of Cu^{2+} . Younan et al.⁷⁰ postulated that the β -sheet increase could be the result of the Cu^{2+} coordination to His96 and His111, which can cause main-chain conformation changes, induce the formation of anti-parallel β -hairpin and cause intermolecular self-association and aggregation. Pushie et al.⁷¹ proposed that Cu-induced PrP conversion might occur in the lysosome compartments which have a lower pH

and favor low-occupancy copper coordination form, as compared to extracellular environments.⁷¹ MD simulation on PrP(96-111) fragment with low-occupancy Cu²⁺ binding form demonstrated that Cu²⁺ bridging to His96 and His111 can promote localized β -sheet structure formation.⁷² However, on the contrary, a combined spectroscopic and MD simulations by Migliorini et al.⁷³ observed β -sheet structure occurrence in apo peptide of huPrP(91-127) without copper binding, while the metal bound complex showed a lower β -sheet content, suggesting that copper binding to His111 imposes precise backbone bending and weakens β -sheet formation propensity.

The relationship between copper binding and prion infectivity is complicated and controversial. At this point, it hasn't reached agreement on the role of copper in prion pathogenesis. It has been reported that copper inhibited the formation of amyloid fibrils in vitro experiment of recombinant PrP.⁷⁴ The presence of copper increased the PrP^C accumulation in normal hamster but decrease the PrP^{Sc} accumulation in the scrapie infected hamster.⁷⁵ In another study, however, it showed that copper induced the PrP^C to PrP^{Sc} conformational changes in recombinant and cellular PrP.⁷⁶ Addition of copper are found to be capable of recover the protease resistance and infectivity of PrP in vitro.⁷⁷ Recently, a single-molecule fluorescence assay demonstrated that copper could induce monomers of PrP to misfold before oligomerization and the misfolded PrP could promote subsequent aggregation. Therefore, it is still unclear whether copper

inhibits or promotes the prion disease.

Although copper may not have a causative effect in prion diseases, the ability of copper binding suggests PrP might play a role in copper homeostasis^{60,78,79}. In addition, there are also studies indicate that PrP might be involved in neuroprotection⁸⁰⁻⁸³, signal transduction⁸⁴⁻⁸⁶ or peripheral myelin maintenance⁸⁷. However, a unifying definition of PrP function has not been uncovered. Therefore, PrP is often regarded to exert pleiotropic effects in diverse physiological functions of neuronal and glial cells⁸⁸.

The downstream region to the primary copper-binding sites at N-terminal is a highly hydrophobic sequence containing a high percentage of Ala amino acids (residues 106-126 in human prion sequence). This region has also received special research interest because the corresponding peptide has shown highly amyloidogenic and neurotoxic properties in vivo and in vitro studies.⁸⁹⁻⁹³ PrP106-126 peptide of human PrP exhibited a predominantly β -sheet structure and showed amyloidogenicity and neurotoxicity in vitro at pH 5.⁹⁴ Synthesized peptides corresponding to the hydrophobic sequence in hamster PrP (residues 109-122) spontaneously formed amyloids composed largely of β -sheet structure.⁹⁵ The fibrillogenicity and neurotoxicity of PrP106-126 peptide can induce neuronal cell death in vitro by an apoptotic mechanism.⁹⁶ A long-time-scale (1.1 μ s in total) all-atoms explicit-solvent MD simulation showed that the peptide fragment of PrP109–122 region can spontaneously form into a β -hairpin.⁹⁷ All these studies

confirm that this hydrophobic region is a β -rich region and may play a role in the amyloid fibril formation.

The PrP106-126 sequence contains a highly hydrophobic and amyloidogenic palindrome AGAAAAGA (residues 113 to 120 in human prion sequence). This hydrophobic core plays an important role in the fibrillation and toxicity of the PrP106-126 sequence.⁹⁸ This palindromic AGAAAAGA region is indicated to be solvent exposed in PrP^{Sc}, the strong hydrophobic interaction of this region may play a role in the infectivity of prion.^{99,100} Tryptophan fluorescence spectrum and ¹H-NMR studies¹⁰¹ showed that the palindromic region was associated with the oligomer fibrils and was engaged in intra- and/or inter-molecular interactions. Some studies¹⁰²⁻¹⁰⁴ has suggested that this region has a strong propensity to form steric zipper amyloid fibrils, with different β -sheet organizations and the neighboring β -sheet layers packing modes due to various local environments of the sequence. Mastrianni et al.¹⁰⁵ reported that the deletion of AGAAAAGA palindrome can prevent the formation of protease-resistant PrP^{Sc}-like protein in yeast, suggesting that the palindrome segment is essential for the prion conversion. Brown¹⁰⁶ proposed that AGAAAAGA was necessary but insufficient for the toxicity of PrP.

1.3 The scrapie prion protein, PrP^{Sc}

Although the atomic structure of PrP^C has been determined, the detailed structure of PrP^{Sc} remains elusive. This is due to the polymeric and heterogeneous nature of PrP^{Sc}, which makes it difficult to get enough quantities of purified materials for conventional structural characterization. Detail information in structural details of PrP^{Sc} are of vital importance for understanding the molecular mechanism of the transmission and for potential pharmacological and clinical intervention.

So far, considerable effort has been devoted to define its structure. PrP^{Sc} shares a high degree of similarity in the amino acid sequence with PrP^C, but has huge differences in secondary and tertiary structure. Unlike the α -rich PrP^C, it is demonstrated that PrP^{Sc} contains significant β -sheet and few α -helical structure. Low-resolution optical measurements reveal that the PrP^{Sc} is β -sheet structure enriched.^{107,108} The secondary structural component of PrP^{Sc} remains elusive. Studies with Frouier-ransform infrared (FTIR) spectroscopy¹⁰⁹ shows that PrP^{Sc} has a less proportion of α -helical structure (30%) and a pronouncedly higher β -sheets structure content (43%), in comparison to PrP^C which contains three α -helices (42%) and a very short two-strand β -sheets (3%). These low-resolution optical measurements are of high uncertainty. However, all these studies confirm the hypothesis that part of the original α -helices in healthy PrP^C would refold into β -sheets during the pathogenic conversion to PrP^{Sc}.

Owing to the structural features of PrP^{Sc}, it reveals different physicochemical

properties with PrP^C. PrP^C is monomeric, soluble, and can easily be degraded by proteinase K. PrP^{Sc}, on the other hand, is insoluble, highly resistant to protease K, and prone to aggregates.^{8,109,110} The insolubility and aggregation tendency of PrP^{Sc} have hampered many classic methods, such as NMR and X-ray, to identify the detailed structure at the atomic level, because it is difficult to obtain sufficient amounts of highly purified samples for experiments. In-depth structural analysis at the atomic level remains limited. Thus, to unravel the atomic structure and biophysical properties of PrP^{Sc} remains one of the most critical challenges in prion biology.

It is well-known that PrP^{Sc} can form aggregates¹¹¹ with characteristics similar to amyloid fibrils. Amyloid fibrils are β -rich filamentous peptide or protein aggregates, in which β -strands are perpendicular to the long axis of the filaments and inter-strand hydrogen bonds are approximately parallel to the long axis.¹¹² The formation of amyloid fibrils appears to initiate from partially structured conformations of proteins, which can provide vital intermolecular interaction for aggregation. There are four types of amyloid fibril architecture: antiparallel, parallel in-register, parallel out-of-register and β -helix.

Diverse types of fibrils have been found in studies with full-length recombinant PrP and various PrP fragments in vitro. In-register parallel β -sheets fibrils have been reported by solid state NMR studies with full-length Syrian hamster prion protein¹¹³, as well as PrP fragment containing residues 106-126¹¹⁴ and residues

127-147¹¹⁵. Antiparallel β -sheets have also been found in fibrils formed by residues 109-122.¹⁰³ Baskakov et al.¹¹⁶ proposed that PrP has an intrinsic propensity to form conformational variational self-propagating structures, and identified two morphologies of fibrils (R- and S-fibrils) formed by full length PrP (residues 23-231).¹¹⁷ In vitro kinetic studies on the aggregation pathway of prion protein suggested multiple misfolding pathways and that the forms of β -sheet-rich amyloid fibrils are dependent on the experimental conditions.

The pathogenic process can be divided into two steps. The first step is the misfolding of normal prion protein to a partially unfolded form which can provide the necessary flexibility and potential intermolecular interaction for the subsequent aggregation. The second step is that the infectious intermediates propagate in an autocatalytic manner to form amyloid fibrils. According to the “protein only”, PrP^{Sc} is the major component of the infectious agent and it can “replicate” in the absence of nucleic acid by templating the misfolding of other healthy PrP^C in an autocatalytic manner. Despite numerous studies, little is known about the mechanism of conversion of PrP^C to PrP^{Sc}. So far, multiple mechanisms of prion replication have been put forward. The most accepted and general model is that conversion between PrP^C and PrP^{Sc} is reversible.¹¹⁸⁻¹²⁰ But the PrP^{Sc} is much less stable than PrP^C, which means that the equilibrium would be strongly shifted to PrP^C. PrP^{Sc} can self-associate into a stable oligomeric nucleus in an ordered manner and this step is the rate-limiting in the mechanism. Upon the

formation of the stable oligomeric species, additional PrP monomers or oligomers are attached. In this manner, new monomeric chains are recruited and the aggregates grow rapidly. Recent advances suggest that prions amyloid fibrils are not the toxic form¹²¹, but rather their propagation process involves production of toxic species. In this way, it is possible that toxic intermediate oligomers or protofibrils accumulate during the process of PrP^C to PrP^{Sc} conversion.

Although it has not been definitively determined that the toxic or infectious form of prion protein is its amyloid fibril form or the intermediate oligomers, the self-propagating mechanism of the infectious PrP can be readily uncovered with reliable hypothetical molecular structure of PrP^{Sc}, which has characteristics similar to amyloid fibrils. Several models have been proposed to represent the structure of PrP^{Sc}. Based on the electron crystallography studies, Govaerts et al.¹²² suggested a trimeric model (Figure 1.3 a) in which residues 89-174 form a parallel left-handed β -helical region and the C terminus, residues 176–227, retain its original globular α -helical domain linked by a disulfide-bond. DeMarco et al.³² built a 3-fold protofibrils model (Figure 1.3 b) by molecular dynamics simulation with a scrapie prion protein. In their model, the 3-strand β -sheets consist of residues 116–119, 119–132 and 160–164 and an isolated strand containing residues 135–140. The three α -helices of the native protein retain their structure, and do not involve in the conversion. The propagation of fibrils appears to occur by connecting the isolated fourth strand to the three-stranded sheet, leading to the

formation of a continuous four-stranded β -sheet. Cobb et al.¹²³ investigated the molecular architecture of the recombinant PrP amyloid fibrils using site directed spin labeling and EPR spectroscopy. Their studies show that, in contrast to the previous models, the PrP^C to PrP^{Sc} conversion majorly involve the refolding of the α -helices of the C-terminus (residues 160-220) including helix 2, a major part of helix 3 and the loop between them. This region form stacking single-molecule layers with parallel, in-register alignment of β -strands (Figure 1.3 c), which is corresponding to the studies carried by Tycko et al.¹¹³ on full-length PrP with NMR measurements.

Figure 1.3: Structure models of prion aggregates. a) Trimeric model with left-handed β -helical formation by Govaerts et al.¹²². b) Three-fold protofibrils model by DeMarco et al.³². c) Stacking layers with parallel, in-register alignment of β -strands by Tycko et al.¹¹³.

1.4 Molecular Dynamics Studies on PrP^C to PrP^{Sc} Conversion

Despite the continuing research, there are still many questions remaining unsolved, including the biological function of PrP^C, the detailed structure of PrP^{Sc}, the mechanism of PrP^C to PrP^{Sc} conversion, the mechanism of propagation and infectivity and so on. A wide variety of experimental measurements have been employed in the past decades, such as X-ray crystallography, NMR spectroscopy and Dual-polarization interferometry. However, many research efforts have been hampered due to the fibrillogenicity, amyloidogenicity and neurotoxicity of prion protein. Computational simulation serves as an important alternative and complementary approach to figure out some of the problems.

Molecular dynamics (MD) simulation is a well-established technique which has been widely used to theoretically simulate the evolvement of biomolecular systems. To model efficiently, electrons are generally ignored and the calculation is based on the nuclear positions only following the classical mechanics. The physical movement of particles and the interaction between them are described using forcefield potential energy function, including the chemical bonding, van der Waals interactions and electrostatic interaction. The trajectories derived from numerically solving the motion equation for each particle can provide detailed information on the time evolution of the simulated molecule, giving an insight into the stability, local dynamics, conformational changes and energy barrier.

To date, molecular dynamics simulation has been widely used in exploring and

understanding the prion biology^{124,125}, for example, mimicking mutations, ions or environmental effect on the conformational conversion, constructing 3D atomic models for amyloid fibril or unstructured regions.

The accuracy of protein folding largely relies on the accuracy of the force fields.¹²⁶ Generally speaking, referring to a set of energy function forms and semi-empirical derived parameters, force fields describe the contributions of various forces acting on the system or atoms. The hydrogen bond energy is usually regarded as an electrostatic term in the force fields, which is dependent on the atomic charges and distances. Hydrogen bonds are recognized as one of the most important structural elements of protein, which have a strong polarization effect.¹²⁷ Most of the conventional force fields use fixed charge and lack polarization effect. The strength of hydrogen bonds has been demonstrated to be underestimated in simulations with standard force fields without polarization effect.¹²⁸ Thus, including polarization effect in the force field is critical for protein dynamics simulation.

Many studies have demonstrated that electrostatic polarization plays an important role in stabilizing the intra-protein hydrogen bonds and protein structure.^{129,130} In previous polarizable charge models, such as polarized protein-specific charge (PPC) scheme¹³¹, adaptive hydrogen bond specific charge (AHBC) model¹³², semifluctuating charge (SFC) model¹³³, and the polarizable hydrogen bond (PHB) model¹³⁰, the electrostatic polarization effect of hydrogen

bonds is included in the MD simulation by periodically updating atomic charges.

Recently, our group developed a polarizable structure-specific backbone charge (PSBC) model based on the existing AMBER force field. The hydrogen bonds formed by the backbone atoms have been discovered to be one of the main forces driving the formation of secondary protein structures, which are namely the α -helix and the β -sheet. Polarized charge distribution between the hydrogen bond donor and acceptor will be distorted when the hydrogen bond's microenvironment changes, for instance, the hydrogen bond forms or breaks, or the hydrogen bond distance varies. This charge transfer will affect the electrostatic interaction energy, which in turn affect the stability of the protein structure. It is important to ensure the accuracy of the backbone atomic charges at each updating step so that the hydrogen bond energy is properly accounted for during the simulation.

In the PSBC model, the polarization effect is accounted by periodically updating the atomic charges (N, H, C, O) involving in backbone hydrogen bonds according to precomputed equations. These equations are derived from quantum-mechanical calculations of the charge transfer between the hydrogen bond donor and acceptor depending on the bond distance in specific secondary structure of α -helix or β -sheet. While most polarizable force field models (e.g. induced dipole models¹³⁴, classic Drude oscillators, fluctuating charges) require computationally expensive iterative and/or *ab initio* methods at every time step of a MD simulation, the PSBC method offers a simpler solution via polynomial

functions to account for atomic polarizability between the backbone atoms.

In the development of this model, firstly, we constructed a two-alanine system to calculate the atomic charges for backbone atoms during the hydrogen bond formation. A hydrogen bond was formed by constraining two alanine residues in a specific secondary structure (α -helix/ β -sheet) with ACE and NME capping at the N- and C-terminus of each residue. The hydrogen bond distance between the two residues was varied from 2.5 to 6.5 Å in increments of 0.1 Å. Meanwhile, the electrostatic potential (ESP) of the two-alanine system was calculated using Gaussian03 at the HF/6-31g* level along the hydrogen bond length. The atomic charges were fitted based on the calculated ESP. Only the charges of atoms involving in the backbone hydrogen bond formation (N, H, C, O) were updated, but all the other atoms were fixed at AMBER94 charges. Then a single-alanine system was constructed by treating the partnering alanine residue in the two-alanine system as ghost atoms. The atomic charges of the single alanine residue were calculated using the same way. By subtracting the charges obtained from the two-alanine system from that of the single-alanine residue, we obtained the charge transfer between the backbone atoms involving in the backbone hydrogen bond formation. Finally, equations were fitted based on the calculated charge transfer and hydrogen bond distance. These equations are implemented in the SANDER algorithm in AMBER simulation program to update the atomic charges of the backbone atoms according the hydrogen bond distance. During the

simulation, a check is performed periodically to inspect if charge updating is needed. If yes, the atomic charge of the atoms involving in the backbone hydrogen bond will be updated according to the precalculated equations. The equations derived from β -sheet structure are as follows:

$$\Delta q_N = 0.000587r^5 - 0.01576r^4 + 0.1701r^3 - 0.9286r^2 + 2.591r - 3.029 \quad (1.1)$$

$$\Delta q_O = 0.0008088r^3 - 0.01742r^2 + 0.1247r - 0.3051 \quad (1.2)$$

where r is the distance between donor and acceptor in angstroms. The values of transferred charges for H and C atoms were the negative of that for N and O atoms.

The PSBC model has been successfully applied to the folding of peptides with distinct secondary structural fold, such as β -hairpin tryptophan zipper, Beta3s peptide (three-stranded β -sheet), and the mini protein 1FME with mixed α/β structure.

Here, in this prion study, we employed the PSBC scheme in the replica exchange molecular dynamics (REMD) simulations to study the folding of peptide fragment corresponding to the hydrophobic region at the PrP N-terminus, aiming to provide possible atomic-level conformation for this undetermined region. The hydrophobic region at PrP N-terminus (residues 90-126) has been demonstrated to show PrP^{Sc}-like amyloidogenic and neurotoxic features and exhibit β -structure in a range of experimental and theoretical studies.

1.5 Organization of this thesis

In this thesis, we use MD simulations to investigate the conformational conversion of human prion protein. The outline of this thesis is as follow:

Chapter 1 is a literature review of prion protein, which mainly introduces the prion diseases, the structural information of PrPC and PrPSc, and MD simulations in prion studies.

Chapter 2 investigates the pH effects on the early stages of conformational conversion of C-terminal domain of human prion protein. The conformational changes of each secondary structural element have been discussed.

Chapter 3 further explores the conformational conversion of huPrP C-terminus by high-temperature simulation and simulated annealing simulation. A new β strand, named as S3, forms at the unfolding part of H2 and folds into a three-stranded β -sheet with native S1S2.

Chapter 4 probes the role of huPrP N-terminus in the conformational conversion. Finally, a six-stranded β -sheet encompassing the C-terminus and the hydrophobic region of N-terminus is achieved.

Chapter 5 summarizes the thesis and proposes possible future directions.

Chapter 2 pH Effect on the Conformational Conversion of Human Prion Protein

2.1 Introduction

The misfolding and aggregation of a normal cellular prion protein (PrP^{C}) protein to a misfolded, oligomeric form prion protein (PrP^{Sc}) can cause a group of fatal prion diseases also known as transmissible spongiform encephalopathies (TSEs).¹³⁵ Prion diseases are a class of inevitably lethal neurodegenerative disorders which can be acquired by sporadic, inherited, or infectious means, and for which there is no treatment yet. Prion diseases in humans include Gerstmann-Sträussler-Scheinker disease (GSS), Creutzfeldt-Jakob disease (CJD)², fatal familial insomnia (FFI)³ and kuru¹³⁶. In animals, relevant prion diseases are scrapie in cattle and sheep¹³⁷, bovine spongiform encephalopathy (BSE, or mad cow disease), transmissible mink encephalopathy (TME) and chronic wasting disease (CWD) in elk and deer.⁴

Owing to the high free energy barriers of PrP^{C} to PrP^{Sc} conversion, prion disease has a very long incubation period, which can exceed several years¹³⁸. However, reduced or lacking efficiency of physicochemical conditions can strongly accelerate the conformational change of PrP^{C107} , such as mutations,

chemical modification, or changes in the environmental conditions¹³⁹.

In neuron cells, PrP^C is attached to outer cellular membrane of neuron cells by a GPI anchor. Generally, PrP^C cycles between the cell surface and endosomal compartments and its cellular trafficking is thought to play a critical role in its pathogenic conversion.¹⁴⁰ PrP will experience very different pH environments during its trafficking process. As internalized PrP molecules progress through the endocytic pathway, they are exposed to an increasingly acidic pH. The physiological pH of neuron cell surface is 7.4, early endosomes is around 6, late endosomes around 5, and lysosomes lower still.^{141,142} To date, it is still unclear where the pathological conversion occurs. Nevertheless, PrP^{Sc} has been found to accumulate in endosomal compartments of the infected neuronal cell.¹⁴³ Hence, it was reasonable to hypothesize that the increasingly acidic environment may play a significant role in the pathogenic conversion.

Many experimental findings have suggested that acidic pH may have significant implications for the presumed formation of PrP^{Sc} during the endocytosis. Oligomers formed at acidic pH exerted cytotoxicity in vitro.¹⁴⁴ Borchelt et al.¹⁴⁵ reported that PrP^{Sc} synthesis took place after the PrP^C transits from the membrane and the acidic pH of the endosomal compartment induce the conformational conversion. DeMarco and Daggett³² have also conducted studies in vitro which indicates a relationship between acidic pH and the misfolding of prion protein. Hornemann and Glockshuber¹⁴⁶ identified a scrapie-like unfolding

intermediate of PrP (121–231) which shows spectral characteristics of β -sheet structure and is exclusively populated at acidic pH conditions. Swietnicki et al.¹⁴⁷ reported that recombinant human prion protein (residues 90-231) can convert into β -sheet-rich oligomer in the presence of 1 M GdnHCl and acidic pH conditions (pH 3.6-5) without disrupting the disulfide bond. Zou and Cashman¹⁴⁸ also observed the effect of acidic pH on the conformational change of human prion protein in vitro studies. All these experimental findings suggest that acidic pH condition may facilitate the presumed formation of PrP^{Sc}.

The cellular prion protein PrP^C is composed a largely disordered N-terminal and a well-structured C-terminal globular domain.²⁷ NMR¹⁷⁻²¹ and X-ray^{22,23} studies have revealed that PrP^C is an α -rich structure with three α -helices and two small β -strands. Unlike PrP^C, PrP^{Sc} has a significant β -sheet content and few α -helical content. It is insoluble, partially resistant to protease K, and prone to aggregates.^{8,109,110} The insolubility and aggregation tendency of PrP^{Sc} have hampered classic methods, such as NMR and X-ray, to identify the detailed structural information. In-depth structural analysis of PrP^{Sc} at the atomic level is of vital importance for understanding the molecular mechanism of the prion conversion.

Due to the structural features, the N- and C-terminal domains of PrP are often studied individually in the exploration of the pathogenic conformational transition.¹⁴⁹ The C-terminal is well-structured and resolved by NMR and X-ray

crystal techniques. The reduction in α -helical content of PrP^{Sc} does imply that the C-terminal helical domain is necessarily involved in the conformational changes. In addition, most diseases-associated mutations¹⁵⁰ leading to the spontaneous transitions are clustered in the C-terminal domain region of PrP^C, hence suggesting the importance of this region. On the other hand, the N-terminal is highly flexible and disordered, making it difficult to have a common viewpoint across studies. Some fragments of N-terminus are reported show β -propensity in specified conditions, e.g. residues 106–126^{94,151}. We will discuss the β -constitution of PrP N-terminus in Chapter 3. Here, in this chapter we will focus on the C-terminus of human prion protein.

In the human prion protein, the globular C-terminal domain (residue 125 to 228, Figure 2.1) consists of three α -helices (H1: residue 144-156, H2: residue 174-194, H3: residue 200-228) and a short two-stranded antiparallel β -sheets (S1: residue 128-131, S2: residue 161-164).²⁷ H2 and H3, which are linked by form a disulfide bond between Cys179 and Cys214, form the hydrophobic bulk of globular domain.

Figure 2.1: NMR structure of C-terminal domain of human prion protein (PDB ID 1QM0, residues 125-228). Three helices H1, H2 and H3 are colored in violet, and the two β -strands S1, S2 are colored in yellow.

Defining the particular role of each secondary structure domain in the conformational conversion can provide clues about the early misfolding events. The stability of isolated peptide fragment encompassing particular secondary structure domains have been investigated by NMR, X-ray Crystal, or related structural studies. Earlier research emphasizes the role of S1H1S2 region³¹⁻³³ which is thought to be the initial site of the protein misfolding. It is implicated that hydrophilic H1 helix is responsible for the α -helix-to- β -sheet conformational change^{29,122,152}, whereas H2 and H3 tend to maintain their helical structure in the PrP^c to PrP^{Sc} transition¹⁵²⁻¹⁵⁴. Norstrom et al.¹⁵² also suggested that H1 is the site associated with PrP^c-PrP^{Sc} complex with emphasis on the importance of specific charge orientations. Conversely, it has also been shown by some NMR and CD

studies that H1 has remarkably high intrinsic α -helical propensity³⁴, and reveals high structural stability under a wide range of environmental conditions³⁵. Some researchers¹⁵⁵ hypothesized that H1 was not as believed previously to be the seeding region of the conformational change, but may refold in the late stage as involved in global structure rearrangements of the protein¹⁵⁶.

In addition, more and more studies support the hypothesis that the H2H3 domain rather than H1 is more associated with the conformational change. NMR study with isolated H2H3 fragment reported that H2H3 can form PK-resistant amyloid fibers similar to that of full-length protein.³⁷ The H2H3 domain undergoes structural transformation early during the aggregation of recombinant PrP, indicating to be the seeding sites of the pathogenic conversion³⁸. Thompson et al.⁴⁸ reported that both H2 and H3 have a high propensity to form β -sheet structure in pH-dependent solutions, but electron microscopy measurement show that only H2 can form amyloid aggregates. Yamaguchi's⁴⁷ study shows that H2 exhibited a high propensity to form long, straight and ordered β -sheet which could induce full length α -to- β conformational change, while H3 contains only short, flexible and random β -sheet. Although many studies confirm that the H2H3 region may play a role in the α -helix to β -sheet conversion, detailed information on the behavior of H2H3 during the conformational transition is still not elucidated.

Due to these limited low-resolution experimental studies on PrP^{Sc}, molecular dynamics (MD) simulation has served as an invaluable and complementary

approach to explore the complicated mechanism of prions, which may give the microscopic information and provide a better understanding of the conversion mechanism.¹⁵⁷ So far, MD simulations have been widely used to explore the pH effect on the PrP^C to PrP^{Sc} conversion. A Constant-pH simulation by Campos et al.¹⁵⁸ suggested a strong pH-dependence in the conformational change, indicating a pH-dependent trend that the α -helices decreased and the β -sheets enriched towards acidic pH. DeMarco and Daggett¹⁵⁹ reported the destabilization of PrP^C via disruption of critical long-range salt-bridges and the disconnection of H1 helix under low pH conditions. Study of Langella et al.¹⁶⁰ revealed a significant loss of α -helix content under mildly acidic conditions, especially in the C-terminal part of H2. Massih et al.¹⁶¹ reported a new β -strands formation between the S1H1 loop (residues 138–141) and C-terminal of H2 (residues 183–192) by restrained MD simulations. Gu et al.¹⁶² reported that at low pH and high temperature, H2 helix is the most unstable region in human PrP^C. This is further supported by a 1 μ s MD simulation at high temperature and acidic pH conditions.¹⁶³ The folded C-terminal showed a drastic conformational conversion by the partial unfolding of the H2H3 and persistent β -sheets between H2 and H3.

In this chapter, MD simulations were performed with C-terminal domain of human PrP^C (residues 125-228). Simulations under room temperature (300K) and different pH levels (neutral: pH ~7, middle: pH~5, low: pH~3) were carried out to investigate the pH effect on the early stages of prion conformational conversion.

The structural and dynamical behavior of each secondary structure (three α -helices: H1, H2, H3; two β -sheets: S1, S2) was discussed separately to define its particular role in the early conversion. We found that the C-terminus of H2 was the most unstable region under acidic pH conditions and seemed to be the initial region for the conformational conversion.

2.2 Methods

2.2.1 Structure Preparation for MD simulation

All the starting structures for our MD simulation were taken from the NMR structure of the human prion protein C-terminal globular domain (PDB ID: 1QM0, residues 124-228). In order to model the three different pH environment, the protonation state of Asp (pKa ~ 4.5), Glu (pKa ~ 4.6) and His (pKa ~ 6.2) residues were manually changed. A low pH environment at an approximate value of pH 3, which falls within the pH range of both endosomal pH and in vitro conversion conditions, was realized by protonating all the Asp, Glu and His residues. A middle pH condition of approximately 5 was attained by protonating only His residues. In neutral (pH~7) simulations, all the Asp, Glu and His residues were left unprotonated. All the simulation systems were neutralized with counterions (Na^+ or Cl^-). The disulfide bond between Cys179 and Cys214 was kept intact in our simulations as there is evidence indicating it is retained during the misfolding and aggregation of PrP¹⁶⁴.

2.2.2 Molecular Dynamics Simulation

MD simulations were carried out with AMBER10 software package and AMBER 99SB force field. All structures were solvated in octahedral periodic box by using TIP3P water molecules and counterions (Na^+ or Cl^-) were added to adjust the system charge to zero. The minimum distance between the surfaces of the waterbox and the protein was set as 10 Å. The systems were minimized in two steps. In the first step, only the solvent molecules are optimized by 1000 steps of steepest descent minimization, and by a 3000 steps of conjugate gradient method. In the second step, the whole system was optimized without restraint, with a 10000 steps of steepest descent minimization, followed by 10000 steps of conjugate gradient minimization. After minimization, the temperature for the system was gradually increased using the canonical ensemble (NVT, constant number of particles, volume, and temperature). For the three simulations at room temperature simulations under different pH conditions, the system was heated up from 0 to 300 K in 100 ps. In all the simulations, Langevin thermostat was used to regulate the temperature of the system and SHAKE algorithm was used to fix the covalent bonds involving hydrogen atoms. The following equilibration stage, as well as MD production run, was performed without any restrains in the isothermal-isobaric ensemble (NPT, constant number of particles, pressure, and temperature). After equilibration, the MD simulations were carried out for 50 ns with a time step of 1 fs. 50 ns MD simulation was performed for each simulation

and the trajectory was saved every 5 ps for further analysis.

2.2.3 Analysis Methods

After the production run, the trajectories were analyzed with Amber Tools and the Visual Molecular Dynamics (VMD) software. The VMD Salt Bridges plugin was used to identify the presence of salt-bridges, with bond lengths not exceeding 3.2 Å. Dictionary of secondary structure of proteins (DSSP) was used to assign secondary structures to the amide acids of proteins. Root mean square fluctuation (RMSF) was used to measure the average deviation between the position of an atom and the reference position. Ca Root mean square deviation (Cα RMSD) was calculated to measure the scalar distance between a conformation and a reference structure as a function of time. And the native structure determined by NMR was used as the reference structure in our calculations. Radius of gyration is the root mean square distance of the protein about its center of mass. It is used to measure the compactness of its folding. The α-helical fraction of each helix domain was examined by the backbone dihedral angle (ϕ, ψ) of the residues in the region of $[(\phi, \psi) = (-60 \pm 20, -40 \pm 20)]$.¹⁶⁵

2.3 Results and Discussions

Here, MD simulations at room temperature (300K) and different pH levels (neutral: pH~7, middle: pH~5, low: pH~3) were performed to study the effect of pH on the conformational behavior of early misfolding. The protonation states of

His, Glu and Asp residues are altered to mimic different pH conditions in the simulation.

With a reference of the native NMR structure, the C α RMSD of the PrP globular domain (residues 128–228) was calculated to monitor the structure stability of conformational changes in all simulations (Figure 2.2). A slight increase in RMSD values with some fluctuations is observed for all the three simulations, indicating that there are some unfolding events in all the three simulations under different pH conditions.

Figure 2.2: RMSD evolutions of human PrP (residues 128-228) in simulations of different pH levels

Subsequently, RMSF analysis was carried out to identify residues that possess greater atomic fluctuations, since any unfolding process should result in a greater flexibility within the secondary structure. As seen in Figure 2.3, it is observed that

the loops connecting different secondary structural domains show high RMSF values, especially the loop connecting S2 and H2. As for the secondary structural elements of β -sheets and α -helices, the C-terminal part of H2 has a much higher rise in the fluctuation when lowering the pH levels, suggesting this region is unstable at acidic pH conditions.

Figure 2.3: RMSF evolutions of human PrP (residues 125-228) in simulations of different pH levels.

In the following discussion, the particular behavior with respect to each secondary structural element including three α -helices (H1, H2, H3) and two β -strands (S1S2) will be discussed separately.

2.3.1 Stability of Helix 1 (H1)

Unlike normal α -helices consisting of most hydrophobic residues, the H1 helix is characterized by a highly unusual sequence of hydrophilic residues (DYEDR YYREN MHR, residues 144-156). So, it lies away from the hydrophobic core formed by H2 and H3²⁹. The DSSP criterion is utilized to assign the secondary structure of H1 (Figure 2.4). In comparison with the healthy cellular PrP^c (NMR structure), the major part of H1 (residues 140-153) retains its native α -helix structure and only a few residues at the C-terminus (residues 154-156) fluctuate between 3_{10} helix structure and turn structure. The fluctuation is caused by the protonation state of His155 at different pH conditions³⁰. At acidic pH, the protonation of His155 introduced a positive charge, which is unfavorable to the neighboring residue Arg156. This can be further supported by the helical content calculation (Figure 2.5), which shows slight differences in the extent of denaturation under different pH changes.

These observations of our computations agree with experimental results with H1 fragment. NMR and CD studies of H1 fragment peptide reported that H1 has strong intrinsic helical propensity³⁴, and this significant helicity remains stable under a wide range of conditions (pH variation, high salt and organic cosolvents)³⁵.

Figure 2.4: Secondary structure evolution of H1 helix by DSSP analysis of simulations at different pH levels and the final structures in each simulation.

Figure 2.5: Fractional α -helical content evolutions of H1 helix in simulations at different pH levels.

Although no secondary structure change occurred in H1, visual inspection of the MD trajectory with VMD program revealed large movements of H1 in all simulations. H1 moved away from the globular H2H3 domain toward different directions, into solutions. The biggest movement is observed in the simulation at low pH. We quantified the movement of H1 by the distance between the disulfide bond and the center of mass of H1. The largest outward movement of H1 takes place in the simulation at low pH with 1.8 Å away relative to its starting position, following by that of middle pH (1.5 Å), and the smallest movement at neutral pH (1.4 Å). Since the movements took place in all simulations, this structural flexibility of H1 might be an intrinsic property of PrP^C.

Previous MD studies of PrP^C in various conditions^{159,166,167} also observed the dislocation of H1 with different types of movements. Different types of H1 movements were observed. H1 slides on the H2H3 plane and moves away from the H2H3 core in diverse directions, downward or toward the N-terminus of H3. However, the movement of H1 does not mean complete dissociation and solvation of H1.¹⁶⁶ It has been suggested that the disruption of salt-bridges is the key event under the disconnection of H1 in acidic conditions. The side chains of Asp and Glu residues are protonated under low pH, which prevent the formation of salt-bridges. DeMarco and Daggett¹⁵⁹ ascribe the large movements of H1 under acidic conditions to the disruption of two long-range salt-bridges linking S1H1S2 region to the H2H3 globular core: Arg156-Glu196 and Arg164-Asp178.

Here, in our study, some highly occupied salt-bridges were identified by a detailed analysis of salt-bridges using VMD Salt Bridges plugin. In the neutral pH simulation, a total of four persistent salt-bridges have been identified, in which three of them are found in connection with the H1 region (Figure 2.6). When pH decreases, these salt-bridges got disrupted gradually. Two of the three salt-bridges remained in the middle pH simulation, while all the three disappeared due to the protonation of Asp and Glu residues in low pH simulation. The disruption of salt-bridges allows more conformational freedom of H1. However, the disruption cannot lead to the unfolding of this region. Actually, due to the high intrinsic α -propensity of H1, there may be on secondary structural changes in the

early misfolding steps. But we cannot preclude the possibilities that H1 may refold in the late stage as involved in global structural rearrangements.

Salt-bridges		neutral	Middle	low
GLU146	ARG208	✓	×	×
ASP202	ARG156	✓	✓	×
GLU196	ARG156	✓	✓	×

Figure 2.6: Long-range salt-bridges identified connecting H1 and H2H3 hydrophobic domain in the simulations at different pH levels.

Experimental studies have reported that the disruption of salt-bridges connecting H1 cannot lead to the unfolding of this region. Actually, H1 has very high intrinsic α -helical propensity³⁴, and retain its secondary structure under a wide range of environmental conditions³⁵, which is in agreement with our analysis.

2.3.2 Stability of Helix 2 (H2)

H2 is the most unstable region in our simulations. The denaturation began from its C-terminus. And it seems that the unfolding of H2 is correlated with the decrease in pH. From the DSSP analysis (Figure 2.7), we can see that the denaturation occurred much earlier under low pH conditions as opposed to middle and neutral pH conditions. Under low pH conditions, there appear to be some evidence of unfolding after approximately 17 ns. Under middle and neutral pH conditions, this evidence of denaturation occurred later and fewer residues were affected approximately 25 ns and 30 ns respectively.

We calculate the helical content of H2 to show the helical variation over simulation time (Figure 2.8). It can be seen that the simulation at low pH resulted in a more pronounced reduction of helical content, following by that at middle pH. The helical content at neutral pH shows the smallest reduction. The extent of denaturation also shows a relationship with pH. Under neutral pH, the decrease in helical content was approximately 10%, with this decrease increasing to 20% and 40% under middle and low pH conditions respectively (Figure 2.8). Thus, it is evident that acidic pH conditions can facilitate the unfolding process. The instability of H2 under low pH simulations was further confirmed by the radius of gyration analysis (Figure 2.9). It can be seen that radius of gyration of H2 under low pH shows a gradual increase over time, suggesting that the H2 helix tends towards unfolding.

Figure 2.7: Secondary structure evolution of H2 helix by DSSP analysis of simulations at different pH levels and the final structures in each simulation.

Figure 2.8: Fractional α -helical content evolutions of H2 helix in simulations at different pH levels.

The instability of H2 observed in our simulation is in consistent with experimental results. NMR studies of human PrP at pH 4.5 reveal that the C-terminus of H2 (residues 187-194) can exist in a disordered conformation²⁷. The amino acid sequence at H2 C-terminus (TTTT, residues 190–193) is rare. It is usually found in a β -strand or loop conformation^{40,168} and makes H2 C-ter an inherently unstable region. The H2 C-terminus is reported to exhibit the most structural perturbation in a number of experimental studies.³⁹⁻⁴² This region has been shown to form β -conformation or amyloid-like fibrils^{41,42,47} and to be neurotoxic in vitro⁴⁸. Thus, the H2 C-terminus may play an early role in the conformational conversion, but this role is still not elucidated.

Figure 2.9: Variations in radius of gyration of H2 helix in simulations at different pH levels.

2.3.3 Stability of Helix 3 (H3)

H3 is the longest and most rigid one among the three helices. The DSSP diagram of H3 (Figure 2.10) shows that the major part of H3 helix (residues 200-224) is conserved well throughout the three simulation with only a few residues at the C-terminal unfolded. Many studies indicated that H3 would retain a large part of its native secondary structure in the PrP^C to PrP^{Sc} conversion. Gallo et al.¹⁶⁹ reported that a “capping box” at the top of H3 may help to keep the helical structure of H3 stable. The disulfide bond, connecting H2 and H3 helices, may also stabilize these two helices to some extent. The H3 helix starts from approximately residue 200 and is almost 30 residues long. However, the C-terminal region of H3 shows some flexibility and is poorly defined in several

species of prion protein, such as mouse prion protein.¹⁷⁰ Thus, the fluctuation in our simulation was in expectation.

Figure 2.10: Secondary structure evolution of H3 helix by DSSP analysis of simulations at different pH levels and the final structures in each simulation.

Figure 2.11: Fractional α -helical content evolutions of H3 helix in simulations at different pH levels

2.3.4 Stability of β -sheets S1S2

The relationship between pH conditions with the two native β -sheets is not so obvious as compared to the above studies of α -helices. From the DSSP analysis in Figure 2.12, we note that the elongation of the native antiparallel β -sheet occurred sporadically in the middle and low pH conditions. However, under neutral conditions, there was an unexpected elongation that occurred fairly stable. This may have something to do with the limitations of the Amber 99SB forcefield used in our simulations. In previous work¹²⁶ of our group on native β -sheet of PrP^C, stable elongation of β -sheet at acidic conditions is seen in the simulation with a charge updated force model (adaptive hydrogen bond-specific charge, AHBC),

while weakening of β -sheet can be observed in the simulation with traditional Amber03 forcefield, indicating that the polarization effect may affect the depicting of secondary structure. Some other studies have also observed the elongation of native β -sheet.

Figure 2.12: Secondary structure evolution of S1 β -strand by DSSP analysis of simulations at different pH levels and the final structures in each simulation.

In summary, no significant structural changes were observed at S1, S2 β -strands and H1, H3 α -helices during the simulations. The C-terminus of H2 exhibited the most structural perturbation in the acidic simulations, which might play an early role in the conformational conversion of prion protein. In the next chapter, we raised the simulation temperature to a mildly high temperature (350K) in order to speed up the denaturation process. With the elevated temperature simulation, we intended to explore whether there would be further changes at each secondary structure domain, and to what extent the denaturation would be.

2.4 Conclusion

In this work, we performed room temperature MD simulations under three different pH conditions to study the pH effect on the early misfolding of prion protein. While H1 retains its structure identity in the simulations, increasing movement due to the disruption of salt-bridges as lowering the pH is noted. The H3 remains stable throughout the simulations. No new β -structure is observed. The C-terminus of H2 is the most unstable region and may be the initial site of the prion protein misfolding. The extent of unfolding of H2 increases with decreasing pH conditions, suggesting that acidic pH can facilitate the unfolding of the protein. Our conclusions are not only in agreement with both experimental and theoretical studies, but also provide insight into the early steps of pH-induced misfolding.

Chapter 3 The C-terminal of H2 Helix in the Prion Protein Induced α -helix to β -sheet Conformational Changes

3.1 Introduction

According to the protein-only hypothesis, the key event in the pathogenesis is the transition of the normal α -rich cellular prion protein (PrP^{C}) to the β -rich isoform PrP^{Sc} . PrP^{C} consists of a flexible disordered N-terminal tail and a well-structured C-terminal domain including two small β -strands and three α -helices. Although the structural properties of PrP^{C} in different mammalian species have been widely investigated, detailed information of PrP^{Sc} structure is still poorly characterized.

Elucidating the molecular mechanism of how PrP^{C} convert to PrP^{Sc} has been the focus of intensive experimental and theoretical studies in the past decades. Several factors have been identified to destabilize the native structure of PrP^{C} and induce amyloid fibrillogenesis, including denaturant agents, mutations, pressure, environmental pH, or elevated temperature.^{139,171} Years of studies have shed some light on the neural pathogenesis of prion disease.

In our previous studies (Chapter 2), we studied the pH effects on the early

conformational changes of the C-terminal domain of human prion protein. MD simulations were performed with C-terminal domain of human PrP^C (residues 125-228) at room temperature and different pH levels (neutral: pH ~7, middle: pH~5, low: pH~3). The particular conformational and dynamical behavior with respect to each secondary structural region including three α -helices (H1, H2, H3) and a two-strand β -sheets (S1, S2) have been discussed separately. We found that H1 maintains its native secondary structure throughout the simulation. The movement of H2 increases with the decrease of pH. Highly unstable behavior of the H2 C-terminal part was found in acidic pH simulations. The extent of unfolding seems to be correlated with the decrease in pH. The longest and most rigid helix H3 remained stable in all the simulations. No obvious elongation of the native β -strands is observed. Our findings are in agreement with a range of both experimental^{47,48} and theoretical studies^{126,161,162}. We concluded that the conformational changes started from the C-terminus of H2 and acidic pH can trigger the conversion.

High temperature may increase the probability of overcoming the activation barrier for protein unfolding. It is demonstrated that the unfolding of proteins can be speeded up by increasing temperature without changing the unfolding pathway¹⁷², which enables us to unravel the protein unfolding pathway and explore diverse intermediate conformation at minimal computational cost. The effects of temperature and the combination of temperature and pH on the stability

and dynamics of human PrP^C have been extensively investigated. Wong et al.¹⁷³ demonstrated that elevated temperature could stimulate the incubation of the PrP conversion in cell-free conditions. Incubation of recombinant PrP at temperatures above 40 °C can lead to the unfolding and aggregation of prion protein. Bocharova et al.¹⁷⁴ reported substantial conformational rearrangements of amyloid fibrils may occur within an mature amyloid structure generated from recombinant prion protein by brief exposure to 80 °C. MD simulation of human PrP D178N mutant at 300K and 500K were conducted by Gilbert et al.¹⁷⁵ Compared to the results of simulation at 300K, they observed a more pronounced increase in β content and further breakdown of the structure in the 500K simulation. Liu et al.¹⁶² studied the unfolding of the human PrP under different temperature and pH conditions and found that the C-terminal of H2 is the most unstable region at low pH and high temperature simulations. Blinov et al.¹⁷⁶ and Colacino et al.¹⁷⁷ also reported that H2 unfolds first at low pH and high temperatures. A long-time scale (1 μ s) simulation with ovine PrP at mildly high temperature (338K) and low pH (pH 3.4) conditions was carried out by Chakroun et al.¹⁶³ They observed the unfolding of H2H3 and the formation of small β segments between H2 and H3.

In this Chapter, firstly, we performed a relatively long time-scale (200ns), explicit solvent, all-atom MD simulation of the human prion protein (PDB ID: 1QM0, residues 125 to 228) at low pH level (pH~3) and moderately high temperature (350K), aiming to speed up the unfolding process and gain more

insight into the conformational changes. Based on the long-time simulation, a principal component analysis of the molecular dynamics was conducted to identify the essential dynamics in the misfolding pathway. The loop between H2 and H3, which was elongated by unfolding of H2 C-terminus and H3 N-terminus, revealed a significant dynamic. It showed a tendency to approach the native S1S2 β -strands. This was of particular interest because this movement trend hinted at the possibility of new β -strand formation adding onto the native β -strands. We verified this hypothesis by a simulated annealing simulation with a starting structure taken from the high temperature simulation. A new β -strand, named as S3, formed at the unfolding part of H2. This new β -strand aligned antiparallel with native S2. Finally, a three-strand- β -sheet structure was achieved and this structure was proved to be stable in a subsequent 100ns regular MD simulation. This three-strand- β -sheet structure might be part of the seeding site for subsequent aggregation.

3.2 METHODS

3.2.1 Structure Preparation for MD simulation

The NMR structure of the C-terminal globular domain (residue 125-228) of the human prion protein (PDB ID: 1QM0) was used as the starting structure. The Asp (pKa ~ 3.7), Glu (pKa ~ 4.2) and His (pKa ~ 6.1) residues were protonated to mimic the low pH conditions. The disulfide bond between Cys179 and Cys214 was kept intact.

3.2.2 Molecular Dynamics Simulation

The MD simulation under high temperature (350K) and low pH (pH~3) condition was carried out with AMBER10 software package and AMBER 99SB force field. The starting structures were solvated in octahedral periodic box by using TIP3P water molecules. Counterions (Na^+ or Cl^-) were added to adjust the system charge to zero. The minimum distance between the surfaces of the waterbox and the protein was set as 10 Å. After minimization, the temperature for the system was gradually increased from 0 to 350 K in 100 ps using the canonical ensemble (NVT). Langevin thermostat was used to regulate the temperature of the system and SHAKE algorithm was used to fix the covalent bonds involving hydrogen atoms. The equilibration and MD production run were performed without any

restrains in the isothermal-isobaric ensemble (NPT). The MD production run was carried out for 200 ns with a time step of 1 fs. The trajectory was saved for further analysis.

3.2.3 Simulated annealing simulation

The starting structure for simulated annealing was taken from the cluster analysis of high temperature simulation. The simulation system was treated with minimization in order to ensure that the system is energetically favourable prior to the simulation. Then the simulated annealing protocol was as performed in a simple stepwise protocol of heating up from 250K to 450K and cooling down to 250K (Figure 3.1). The system was heated from 250K to 300K linearly for 20ps and was equilibrated at 300K for 20ps. Then the system was continuously heated from 300K to 350K for 20ps, and was equilibrated at 350K for 20ps. This intermittent heating and equilibration process went on until the temperature reached the highest temperature of 450K. Then the mirror symmetric cooling down procedure began until the temperature went back to 250K. Simulation time of one cycle is 320ps. This cycle was repeated 50 times.

Figure 3.1: Simulated annealing protocol

3.2.4 Analysis Methods

Amber Tools and the Visual Molecular Dynamics (VMD) software are used to conduct analysis of MD trajectories. The hydrogen bond analysis was conducted with Ptraj in Amber Tools. The criterions of hydrogen bond and distance cutoff were set as $3.5\text{\AA}$ and 120° , respectively. Cluster analysis was also carried out with Ptraj in Amber Tools, grouping similar structures from the simulation trajectories into distinct conformational states. Dictionary of secondary structure of proteins (DSSP) was used to assign secondary structures to the amide acids of proteins. Principal Component Analysis (PCA) was used to determine the most important motions in protein dynamics.

3.3 Results and Discussions

In this work, we use MD simulations to study the misfolding mechanism of human prion protein. We focus on the C-terminus of prion protein. A 200ns MD simulation under high temperature and acidic pH was carried out to speed up the unfolding process, aiming to gain further insights of the denaturation. Simulated annealing was used to overcome the energy barrier and achieve the formation of new β -strand.

3.3.1 Speeding up the unfolding by high temperature simulation

In Chapter 2, we performed room temperature MD simulations under different pH conditions to examine the pH effect on the early misfolding of prion protein. We found that H2 C-terminus was the initial site of unfolding, while the other regions largely retain their native secondary structure. The extent of unfolding of H2 increases with decreasing pH conditions, suggesting that acidic pH can facilitate the unfolding of the protein.

Here, in this chapter, we raised the simulation temperature to a mildly high temperature, aiming to speed up the denaturation process and reveal more structural information about the PrP^C to PrP^{Sc} conformational conversion. It has been demonstrated that increasing temperature in MD simulations can accelerate the unfolding of protein without changing its pathway of unfolding¹⁷². A 200ns simulation under elevated temperature (350 K) and low pH (pH~3) conditions was

performed, in hope to gain further insights.

DSSP analysis was carried out to show the secondary structure assignment over time (Figure 3.2). The secondary structure changes in the early stage of high temperature simulation were in consistent with that of room temperature simulation in Chapter 2. The H1, H3, S1 and S2 domains appeared to largely maintain their native structures, while partially denaturation occurred first at C-terminus of H2 helix.

As the simulation went on, let us have a look at further conformational change at each secondary structural domain. There is a large reduction in helicity of the globular domain around 100ns. In the simulation before 100ns, unfolding is only seen at C-terminal part (residues 184-193) of the H2 helix. After 100ns, the N-terminal of H3 (residues 200-205) become unstable and a further denaturation in H2 C-terminal (residues 180-183) is observed. Nevertheless, the H1 helix retains its native helical structure throughout the simulation. The dramatic denaturation at H3 N-terminal and H2 C-terminal generates a long floppy loop between H2 and H3. As for the β -sheet formation, S1S2 extended the β -sheet length around 40ns. In addition, two small β segments are also observed. One is a small antiparallel β -turn formed at the S1H1 loop region between residue 135 and 138. The other one is a small parallel β -turn segment formed at the unfolded region of H2 between residue 186 and 190.

Figure 3.2: (A) Secondary structure evolution of huPrP by DSSP analysis of simulation at high temperature and acidic pH conditions. (B) Starting and final structures in the simulation.

Principal Components Analysis (PCA) was performed to determine the essential dynamics of the misfolding of the prion protein. PCA is a powerful method used for the analysis of collective modes of fluctuations in biomolecules.¹⁷⁸ It has been found that the vast majority of protein dynamics can be described by a surprisingly low number of collective degrees of freedom. And

the major collective motions are assumed to dominate the functional dynamics of proteins. PCA is an effective way to identify the essential motions in protein dynamics and builds a relationship between dynamics and protein function.

Here, we use the Ptraj program in Ambertools to perform PCA in Cartesian space. The essential motions are visualized with VMD. We found that the essential motions located at the loop between H2 and H3 loop in the analysis. In the first eigenvector of PCA, the largest motion is the extensive movement of the long H2H3 loop. What arouses our interest is the motion of H2H3 loop in the second eigenvector (Figure 3.3). We found that the motion seems to have a certain pattern — the H2H3 loop move towards the native S1S2 β -sheet. This motion is very interesting because it indicates a possibility of additional β -strand formation onto the native β -sheet.

The residues corresponding to the elongated H2H3 loop, especially those residues unfolded from original H2 C-terminal (TVTTTT, residues 188–193) has a high intrinsic propensity for extensive β formation.¹⁷⁹ The C-terminus of H2 have been reported to exhibit the most structural perturbation in a number of studies.^{41,43-46} The H2 helix has also been shown to form β -conformation or amyloid fibrils^{41,42,47} in vitro. Thus, we hypothesize that new β conformation may come about at H2H3 loop and form connection with the the native S1S2 β -strands, achieving a more extensive β -sheet conformation.

Figure 3.3: Essential dynamics of prion protein identified by the second eigenvector of PCA analysis

A cluster analysis of the last 100ns of trajectory was conducted in order to identify the structural properties of the conformational substrates of the partially denatured prion protein. Through cluster analysis, we managed to obtain 6 distinct clusters (Figure 3.4). Compared to the native structure, all the six substrates have a dislocated H1, a largely unfolded H2, a partially unfolded H3 and elongated S1S2 β -strands. The major structural differences among the six substrates located at the long floppy loop resulted from the denaturation of H2H3. The most populated cluster is Structure F (Figure 3.4) with H2H3 loop away from the native S1S2 strands. In the second largest cluster (Structure E, Figure 3.4), the H2H3 loop is quite close to native S1S2. Moreover, the alignment of the loop is almost parallel to S2, indicating a potential of adding a new β -strand on to the native S1S2. In the rest clusters, the H2H3 loop is located at different distances with

regard to S1S2. In fact, the results of cluster analysis are in accordance to our PCA analysis. The representative structures of cluster analysis correspond to the stationary points of the dynamics at different stage.

Figure 3.4: Representative structures given by the clusters analysis of the last 100ns of high temperature simulation.

3.3.2 Formation of new β -strand S3

In the simulation at high temperature and acidic pH conditions, the loop between H2 and H3 increased from 6 residues to 23 residues in length by the unfolding of H2 C-terminal and H3 N-terminal. The movement of this long floppy H2H3 loop has been recognized the most significant dynamics of the protein during the conformational conversion process. It is interesting to observe that this loop can approach the native S1S2 β -sheets. In consideration of its intrinsic high β propensity, we hypothesize that a new β -strand would form at this region.

A simulation annealing simulation was performed to test our hypothesis of new β -strand formation. The simulated annealing was performed in a repeated protocol of heating up (heat from 250K to 450K, equilibrate every 50K in between) and cooling down (cool from 450K to 250K, equilibrate every 50K in between), see Figure 2. The heating can overcome energy barriers, and then the cooling is used to trap the folded state. The repeated heating and cooling of the molecule allows the molecule to search different regions of conformational space. The heating and cooling process was repeated 50 times, resulting in a 16ns trajectory. The starting structure was taken from the cluster analysis of high temperature simulation, Structure E (Figure 3.4). It is the second largest cluster. In this structure, H2H3 loop lies close and parallel to the native S1S2 strands, indicating high possibility for new S3 formation.

Figure 3.5: (A) Secondary structure evolution of huPrP by DSSP analysis of the simulated annealing simulation. (B) Hydrogen bonds detected between the newly formed S3 and native S2 β -strands.

DSSP analysis was conducted to monitor the secondary structure conversion during the repeated simulated annealing process (Figure 3.5 A). Formation of new β -strand, named as S3, was observed around 9 ns (28th repeat), and remained stable for a quite long simulation time till 15 ns (45th repeat). S3 formed at the unfolding part of H2 C-terminus encompassing residues 187-189, and aligned

antiparallel with S2 (residues 160-164). Two pairs of hydrogen bonds (160O-189N & 187O-162N) (Figure 3.5 B) were identified between S2 and S3 through the backbone hydrogen bonds analysis. The newly formed S3, together with native S1 and S2, led to a β -rich conformation with an antiparallel three-stranded β -sheet.

Figure 3.6: Secondary structure evolution of three-stranded β -sheet structure by DSSP analysis of the conventional simulation.

A subsequent 100-ns explicit solvent MD simulation at room temperature and acidic pH conditions was carried out to check the stability of the three-stranded β -sheet structure. From the DSSP analysis (Figure 3.6), we can see that the overall structure remains fairly stable throughout the simulation. Elongation of S2 and S3

β -sheet are observed around 60 ns.

A detailed hydrogen bond analysis (Figure 3.7) based on the conventional MD shows that, in addition to the two backbone hydrogen bonds (160O-189N & 187O-162N) detected in the simulated annealing simulation, three more hydrogen bonds emerge (189O-160N, 189O-159N, and 190O-159N), which further stabilize the newly formed S3.

Figure 3.7: (A) H-bonds detected between S2 and S3 in the conventional MD simulation of three-stranded β -sheet structure. (B) Distances of H-bonds over simulation time.

The formation of S3 in our simulation is not a matter of chance. S3 stems from the denatured loop between H2H3, which showed high flexibility and β

propensity in previous experimental and theoretical studies. The β -structure in our simulations is in great agreement with the high β -propensity regions identified by a NMR investigation¹⁷⁹ of urea-denatured state of oxidized human prion protein by Schwalbe's group. In that study, three regions have been shown to have high β -propensity, including native S2 (residues 162-164), the loop between S1 and H1 (residues 135-142), and the C-terminus of H2 (residues 182-189). In our simulations, the C-terminus of H2 was the most unstable region and converted from the α -helical structure to new β -strand S3. The newly formed β -strand S3 encompassed residues 187-189 and aligned antiparallel with S2 (residues 160-164), leading to the formation of a three-stranded β -sheet. This conversion of α -helices to β -sheets at the C-terminus of H2 could possibly be the trigger for the PrP^C to PrP^{Sc} conformational change. A small β -turn was noticed to form at the loop between S1 and H1 (residues 135-142). Both Schwalbe's group's and our studies suggest that the two native β -strands, S1H1 loop and C-terminus of H2 are responsible for the intermolecular β -structure augmentation and may be that seeding site of the subsequent aggregation.

3.3.3 Probing the role of disulfide bond in the conformational transition

Prion protein contains a disulfide bond between Cys179 and Cys214, which can stabilize the hydrophobic fold of the protein by linking the two long helices H2

and H3. In the NMR investigation by Schwalbe's group¹⁷⁹ of urea-denatured state of human prion protein, they demonstrated that the reduced state (disulfide bond reduced) of human prion protein did not form amyloid fibrils and loss of restrictions of disulfide bond could increase the dynamics around the two Cys residues. Here, we have also performed an 80 ns MD simulation at high temperature (350K) and acidic pH (pH~3) conditions with the reduced human prion protein, aiming to provide some structural information about the reduced state.

In comparison to our simulation of oxidized human PrP under high temperature and low pH (first 80 ns), the conformational changes in S1, S2, H1 and H2 structural domains were similar (Figure 3.8). Sporadic elongation of S1S2 was observed. H1 retained its helical structure and moved away from its original position. The unfolding at H2 C-terminus was also noted in both simulations. The most obvious difference is located at the H3 helix, especially its C-terminus. In the simulation of reduced form, the unfolding started from the N-terminus of H3, followed by the C-terminus around 45 ns. A large part of the H3 helix became unfolded in the simulation of reduced form, while there was only minor unfolding at the N-terminus in the simulation of oxidized form. In addition, the relative positions of H2 and H3 were also affected by loss of restrictions of the disulfide bond. Once the disulfide bond breaks, it may provide additional flexibility and affect the conformational stabilities of H2 and H3 helices.

Figure 3.8: Secondary structure evolution of reduced huPrP by DSSP analysis of the conventional simulation.

In fact, the role of disulfide bond during the pathogenic conversion process is controversial in the past decades. Some experiments suggested that the formation of PrP^{Sc} is inhibited in reducing environment and the integrity of the intramolecular disulfide bond is required during the conversion.¹⁶⁴ On the contrary, there are also studies reported that the reduction of disulfide bond can induce a transition to a β-rich PrP^{Sc}-like conversion. Studies also indicate that the reduction of disulfide bond may be a temporary event during the conversion process and the reformed disulfide bond in PrP^{Sc} can be intermolecular or intramolecular.¹⁸⁰ Our simulations suggest that the breakage of the disulfide bond

may not be necessary in the initial stage of prion misfolding, as unfolding were observed in the simulations of both oxidized form and reduced form. Based on the simulation of reduced form, the disulfide bond may have a stabilizing effect on the structure of H3, especially its C-terminus. Additional simulation on the reduced form is needed to be carried out in our future work to further explore the role of the disulfide bond.

3.4 Conclusion

In summary, firstly, we carried out a long time-scale simulation at a high temperature and acidic environment to speed up the unfolding process of prion protein, aiming to provide more insight with regards to conformational changes. During the high temperature simulation, the secondary structure of H1 helix remained stable but showed a large locational mobility. The unfolding started from H2 C-terminus. Further unfolding of H2 C-terminus and H3 N-terminus were observed as simulation went on. Stable elongation of S1S2 was also observed. What caught our attention was the movement of the extended loop resulted from H2H3 unfolding, which showed a tendency to approach the native S1S2 strands. The mobility of this loop, combined with its high β -structure propensity, indicated a possibility of new β formation in this region. We successfully verified the hypothesis with a simulated annealing simulation. A new β -strand, named as S3, formed at the unfolding part of H2 C-terminus and aligned antiparallel with native S2, leading to three-stranded β -sheet structure. This

three-stranded β -sheet structure was proved to be quite stable by a subsequent 100-ns explicit solvent MD simulation at room temperature and acidic pH conditions. Finally, we obtained a three-stranded β -sheet conformation of the prion protein with α helix content decrease and β -sheet content increase.

In our simulations, the conformational conversion of α -helices to β -sheets occurred at the C-terminus of H2. It is consistent with recent experimental and computational studies, which suggest that C-terminus of the H2 helix is intrinsically unstable and has high β -structure formation propensity. The difference between previous computational studies and our studies is that, in our studies, the S3 refolded from denatured H2 C-terminus is associated with the native S1S2 β -strands, leading to more extensive three-strand β -sheet structure. The three-strand β -sheet structure might be the seeding site for β -structure expansion and the subsequent aggregation.

Chapter 4 Probing the β constitution at the N-terminus of prion protein

4.1 Introduction

PrP^C is composed of a largely disordered N-terminal and a well-structured C-terminal globular domain.²⁷ In our previous work, we mainly focus on conformational conversion of the C-terminus of human prion protein. Our results revealed α -decrease and β -increase in the secondary structure of PrP C-ter. The C-terminus of H2 helix was the most unstable region and underwent α -to- β conformational change, leading to the form of a new β -strand named S3. The new S3 coexisted with S1S2 and formed a stable three-strand- β -sheet, which might be the β -core and the initial site for β enrichment and subsequent aggregation.

The experimentally observed increase of β content is far more than what has been observed in this work with PrP C-ter. FTIR study reveals that PrP^{Sc} has a less proportion of α -helical structure (30%) and a pronouncedly higher β -sheets structure content (43%) compared to that of the healthy form PrP^C which contains 42% α -content and 3% β -content. In this work, we attempt to probe the β ingredient in N-terminus of prion protein.

The N-terminus of prion protein contains a signal peptide that directs its intracellular trafficking to Golgi (residues 1-22 in huPrP)¹⁸¹, four octarepeat

regions associated with copper binding (residues 60-91 in huPrP), a hydrophobic region (residues 106-126 in huPrP), and the non-octarepeat copper binding site (His96 and His111 in huPrP)^{68,182}.

The hydrophobic region is highly conserved among different mammalian species and contains a high percentage of Ala amino acids.¹⁸³ It is located at the N-terminus of prion protein, and is adjacent to the well-structured C-terminal domain. This region has received special research interest because the corresponding peptide has shown many PrP^{Sc}-like features in vivo and vitro.⁸⁹⁻⁹³ PrP106-126 peptide of human PrP exhibited a predominantly β -sheet structure and showed amyloidogenicity and neurotoxicity in vitro at pH 5.⁹⁴ Synthesized peptides corresponding to the hydrophobic sequence in hamster PrP (residues 109-122) spontaneously formed amyloids composed largely of β -sheet structure.⁹⁵ The fibrillogenicity and neurotoxicity of PrP106-126 peptide can induce neuronal cell death in vitro by an apoptotic mechanism.⁹⁶ A long-time-scale (1.1 μ s in total) all-atoms explicit-solvent MD simulation showed that the peptide fragment of PrP109–122 region can spontaneously form into a β -hairpin.⁹⁷ All these studies confirm that this hydrophobic region is a β -rich region and may play a role in the amyloid fibril formation.

Within the PrP106-126, the most highly hydrophobic and amyloidogenic sequence is AGAAAAGA (residues 113-120). This hydrophobic core plays an important role in the fibrillation and toxicity of the PrP106-126 sequence.⁹⁸ This

palindromic AGAAAAGA region is indicated to be solvent exposed in PrP^{Sc}, the strong hydrophobic interaction of this region may play a role in the infectivity of prion.^{99,100} It is reported that the palindromic region was associated with the oligomer fibrils and was engaged in intra- and/or inter-molecular interactions.¹⁰¹ Deletion of this palindrome sequence¹⁰⁵ can prevent the formation of protease-resistant PrP^{Sc}-like protein in yeast, suggesting that the palindrome segment is critical for the prion conversion. Another study further proposed that AGAAAAGA was necessary but insufficient for the toxicity of PrP.¹⁰⁶

The non-octarepeat copper binding region is adjacent to the hydrophobic palindrome region, containing two His residues (His96 and His111 in HuPrP sequence).⁶⁸ The His residues were found to be possible binding sites for copper ions. Localized β -sheet formation has been observed within this region in both experimental⁶⁹ and theoretical⁷² studies, especially in the low-occupancy binding form⁷⁰ with Cu²⁺ bridging to His96 and His111 residues under acidic pH conditions⁷¹. However, the relationship between copper binding and prion conversion is complicated and controversial. Some studies reported that β -structure was more likely to form in this region with high occupancy copper binding form¹⁸⁴ (copper binding to single His residue), or even without copper coordination⁷³ rather than the low-occupancy binding form. At this point, it hasn't come to agreement on the role of copper in prion pathogenesis, whether copper inhibits^{74,75} or promotes^{76,77} the prion disease.

In this work, we use an incremental approach to model the unstructured N-terminal residues onto the globular domain of prion protein. Many studies have suggested that the unstructured residues 90-126 at PrP N-terminus are associated with the amyloidogenic and toxic nature of PrP and can form β -structure. Three overlapping N-terminal fragments have been studied, including PrP(116-131), PrP(105-122), and PrP(95-112). Each fragment was studied in two steps. In the first step, a combination of REMD simulation and PSBC updating scheme was employed to predict the secondary structure of the fragment. This combination has been successfully used in the folding of peptides with distinct secondary structural fold, such as β -hairpin tryptophan zipper, Beta3s peptide (three-stranded β -sheet), and the mini protein 1FME with mixed α/β structure. In the next step, referring to the results of REMD simulations with PrP fragments, the corresponding N-terminal residues were built up onto the globular C-terminal domain of PrP and a long time scale conventional MD simulation was carried out to check the stability of the constructed model. By the incremental build-up, we successfully obtained a β -rich conformer with a six-stranded antiparallel β -sheet encompassing both the prion C-terminus and the amyloidogenic region of the N-terminus.

4.2 Methods

4.2.1 Polarizable Structure-Specific Backbone Charge Model

MD simulation is considered as a powerful tool for elucidating the protein folding mechanism. The accuracy of protein folding largely relies on the accuracy of the force fields. Most of the conventional force fields use fixed charge, include implicit polarization in an average way, and lack polarization effect, which plays an important role in the stability of hydrogen bonds and protein structure.¹²⁷ Recently, our group developed a simple formula-based polarizable structure-specific backbone charge (PSBC) model based on the existing AMBER force field to include the electrostatic polarization in the simulation.

In the PSBC model, the polarization effect is accounted by calculating the charge transfer between the backbone atoms (N, H, C, O) during backbone hydrogen bond formation. By varying the hydrogen bond length from 2.5 to 6.5 Å in a two-alanine system of a specific secondary structure (α -helix/ β -sheet), the atomic charges of the backbone atoms (N, H, C, O) were fitted based on the calculated electrostatic potential (ESP) of the system. While all the other atoms used Amber fixed charges. Then, the atomic charges of the single-alanine system were calculated in a similar way by treating the partnering alanine residue in the two-alanine system as ghost atoms. By subtracting the charges obtained from the

two-alanine system from that of the single-alanine residue, the charge transfer between the backbone atoms (N, H, C, O) involving in the backbone hydrogen bond formation was obtained. Finally, equations were fitted by relating the charge transfer and the backbone hydrogen bond length. The equations derived from β -sheet structure system are as following:

$$\Delta q_N = 0.000587r^5 - 0.01576r^4 + 0.1701r^3 - 0.9286r^2 + 2.591r - 3.029 \quad (4.1)$$

$$\Delta q_O = 0.0008088r^3 - 0.01742r^2 + 0.1247r - 0.3051 \quad (4.2)$$

where r is the distance between donor and acceptor in angstroms. The values of transferred charges for H and C atoms were the negative of that for N and O atoms. These functions are implemented in the SANDER algorithm in AMBER simulation program to periodically update the atomic charges of the backbone atoms according the hydrogen bond distance. Details about the PSBC model can be found in Chapter 1 Section 1.4.

The PSBC model has been successfully applied to the folding of peptides with distinct secondary structural fold, such as β -hairpin tryptophan zipper, Beta3s peptide (three-stranded β -sheet), and the mini protein 1FME with mixed α/β structure.

Here, we employed the PSBC scheme in the replica exchange molecular dynamics (REMD) simulations to study the folding of peptide fragment corresponding to the hydrophobic region at the PrP N-terminus, aiming to provide possible atomic-level conformation for this undetermined region. The

hydrophobic region at PrP N-terminus (residues 90-126) has been demonstrated to show PrP^{Sc}-like amyloidogenic and neurotoxic features and exhibit β -structure in a range of experimental and theoretical studies.

4.2.2 Replica Exchange Molecular Dynamics methodology

The replica exchange molecular dynamics (REMD) method is routinely used to study the protein folding.^{185,186} REMD can enhance conformational samplings and help overcome energy barriers efficiently. It is a generalized ensemble algorithm, in which several simulations (replicas) run in parallel at different temperatures and exchanges between adjacent replicas are attempted every time interval based on the Metropolis criterion with probability:

$$P(i \leftrightarrow j) = \min \left(1, \exp \left[\left(\frac{1}{K_B T_i} - \frac{1}{K_B T_j} \right) (U_i - U_j) \right] \right) \quad (4.3)$$

where $P(i \leftrightarrow j)$ is the exchange probability, K_B is Boltzmann's constant, T_i and T_j are the temperatures for the replicas, U_i and U_j are the instantaneous potential energies. This equation is suited for the NVT ensemble, and an additional term should be introduced for the NPT ensemble.

The starting structures were linear structures built with LEaP module of Amber10 using the corresponding peptide sequences and terminal caps. AMBER99SB force field and generalized Born solvation model were employed. The systems were minimized until convergence, and then were heated up to the target temperatures in 500 ps. The time step was 1 fs. The SHAKE algorithm was

used to constrain all bonds with the hydrogen atoms and the Langevin dynamics was used to regulate the temperatures. The exchanges between adjacent replicas were attempted every 5 ps. For each REMD simulation, 20000 exchange attempts were performed, resulting in a total of 50 ns long simulation.

In our work, The AMBER 10 software package was used for all the REMD simulations and subsequent data analysis. REMD simulations were carried out for three PrP fragment peptide corresponding to different sequence at the hydrophobic region of PrP N-terminus. For the simulation of PrP(116-131), the REMD productions runs were carried out with both standard AMBER99SB force field and the PSBC model. The simulations were composed of 8 replicas with temperatures setting as 262.00, 280.67, 300.45, 321.48, 343.82, 367.50, 392.63, 419.30 K. For the rest two fragments, the simulations were carried out with only the PSBC model. The temperatures of the 8 replicas for the simulation of PrP(105-122) were set as 264.00, 280.63, 298.17, 316.42, 335.91, 356.45, 378.12, 400.96 K. The temperatures of 10 replicas for the simulation of PrP(95-112), were set as 265.00, 279.96, 295.66, 312.11, 329.38, 347.47, 366.45, 386.36, 407.22, 429.16 K. The exchange acceptance ratio is about 40%. During the simulation with the PSBC model, the atomic charges for atoms involving in backbone hydrogen bonds (N, H, C, O) were periodically updated (every 2 ps) according to the precalculated PSBC equations.

4.2.3 Conventional MD simulations

The structural stability of the constructed model is examined by explicit solvent conventional MD simulations. The missing residues at PrP N-terminus were built up onto the globular domain of PrP by Maestro program. The constructed structure was minimized by Macromodel panel in Maestro. The disulphide bond between Cys179 and Cys214 was kept intact in our simulations as there is evidence indicating it is retained during the misfolding and aggregation of PrP¹⁶⁴.

MD simulations were carried out with AMBER10 software package and AMBER 99SB force field. All structures were solvated in octahedral periodic box by using TIP3P water molecules. The minimum distance between the surfaces of the waterbox and the protein was set as 10 Å. Counterions (Na⁺ or Cl⁻) were added to adjust the system charge to zero. The systems were minimized until convergence was achieved. After minimization, the temperature for the system was gradually increased from 0 to 300 K in 300 ps. In the simulations, Langevin thermostat was used to regulate the temperature of the system and SHAKE algorithm was used to fix the covalent bonds involving hydrogen atoms. The following equilibration and the MD production run were performed in NPT ensemble with a time step of 2 fs. 50 ns MD simulation was performed for each simulation and the trajectory was saved for further analysis.

4.2.4 Analysis Methods

The analysis of MD trajectories was performed with Amber Tools and the Virtual Molecular Dynamics (VMD) software. The hydrogen bond analysis was conducted with Ptraj in Amber Tools. The criteria of hydrogen bond and distance cutoff were set as 3.5 Å and 120°, respectively. Cluster analysis was also carried out with Ptraj in Amber Tools, grouping similar structures from the simulation trajectories into distinct conformational states. Dictionary of secondary structure of proteins (DSSP) was used to assign secondary structures to the amide acids of proteins.

4.3 Results and Discussions

In our previous studies of the C-terminal domain of PrP (Chapter 3), we obtained a three-stranded β -sheet structure with α -helical content decrease and β -sheet content increase. The decrease in α -helical structure was primarily due to the unfolding of H2 and H3 helices. The increase in β -sheet structure mainly lies in the elongation of native S1S2 β -sheets and the formation of new β -strand at the unfolding region of H2, which was named as S3. Recent years, numerous studies have shown that residues 90-126 at PrP N-terminus exhibits many PrP^{Sc}-like biochemical features and forms neurotoxic amyloid-like fibrils in vitro, indicating that this region might be a β -rich region. It is important to include this region in

the conformational conversion study of prion protein.

In this work, we intend to study the secondary structure preference of N-terminal fragments and construct the missing N-terminal residues onto the globular domain of PrP, attempting to provide an atomic structure view of a larger PrP fragment spanning both the N- and C-terminal region. Our construction was based on the three-stranded β -sheet structure of the globular C-terminal domain in previous work (Chapter 3). Our study focuses on two aspects: whether the N-terminal residues of PrP can form β -sheet structure; if yes, what is the relationship of N-terminal β -sheet and C-terminal β -sheet? Our study would provide conformational information of the missing residues at the N-terminus of the prion protein.

4.3.1 Formation of four-stranded β -sheet structure

The highly hydrophobic region (residues 113-120) falls within the unstructured N-terminus (residues 23-124) and is adjacent to the well-defined C-terminal globular domain. This region has a palindromic sequence of AGAAAAGA, which is highly conserved among various mammalian species. This palindrome region has been reported to possess amyloidogenic and neurotoxic properties⁹⁸, indicating to be a critical region in prion generation.¹⁰⁵ Due to the insoluble nature and strong aggregation tendency of amyloidogenic region, detailed structural information of this region remains limited. To date, various experimental and

computational approaches have been utilized to predict the atomic structure of this region. Some studies¹⁰²⁻¹⁰⁴ have suggested that this region has a strong propensity to form steric zipper amyloid fibrils, with different β -sheet organizations and the neighboring β -sheet layers packing modes due to diverse local environments of the sequence. Zhang et al.¹⁸⁷ constructed an optimal atomic-resolution of AGAAAGA amyloid fibrils by simulated annealing method. Zanuy et al.¹⁸⁸ reported that the antiparallel arrangement for the β -sheet formed by the palindrome sequence was the most stable conformation within sheets, while the parallel packing was stable between sheets. Recently, a new β -strand S0 was solved in a nanobody assisted crystallography study of human prion protein by Jan Steyaert et al.¹⁸⁹ Half of the palindromic sequence (residues 118-122) at N-terminus adopts an extended β -configuration, named as S0. The newly found S0 aligns antiparallel to native S1 and folds into a three-stranded antiparallel β -sheet together with native S2, adopting a S0-S1-H1-S2-H2-H3 structural organization. This study provides a first crystal structural view of the palindromic region adopting a β -sheet conformation.

Here, the secondary structure of PrP N-terminal fragment of residues (116-131) is studied with REMD simulation with both standard Amber 99SB force field and the PSBC model. PrP(116-131) consists of a major part of the palindrome region and the native S1 β -strand sequence. Based on the research studies discussed above, an antiparallel β -sheet was expected in the simulations.

The REMD simulation was comprised of 8 replicas with an acceptance ratio of 40%. The temperatures were set as 262.00, 280.67, 300.45, 321.48, 343.82, 367.50, 392.63, 419.30 K. The simulation of each replica starts from an extended-linear structure, which was heated up to the target temperature and was equilibrated using the Amber Program. The exchanges between adjacent replicas were attempted every 5 ps. 25000 exchange attempts were performed, resulting in a total of 50 ns long simulation. During the simulation with the PSBC updating scheme, the charges of atoms involving the β -sheet backbone hydrogen bond donor and acceptor are updated according to the equation in the PSBC model, while all the other atoms use standard Amber99SB fixed charge. Periodic check of backbone hydrogen bonds and charge update is performed every 2 ps.

The secondary structure evolutions of PrP(116-131) are determined with DSSP analysis (Figure 4.1). In the simulation with standard Amber99SB forcefield, the secondary structure distribution is highly scattered at both 262.00K and 300.45K. Various types of secondary structure elements are observed while none of them are the dominant structure. However, in the simulation with PSBC model, the secondary structure is mainly consisted of antiparallel β conformations with a transient show-up of α -helix and 3_{10} helix structure. It can be seen that the starting extended structure quickly runs into a β -sheet conformation. Especially at 262.00K, the structure undergoes some changes on the detailed strand alignment but remains the β -conformation throughout the simulation.

Figure 4.1: Secondary structure evolution of PrP(116-131) by DSSP analysis of the trajectories under different conditions.

In the recently reported crystal structure by Jan Steyaert et al¹⁸⁹, the detailed structural features of S0 and S1 are as followings: Gly123 and Gly124 are the corner residues for the S0S1 β -strands. The irregular hydrogen bonding between Leu125 and Val122 disrupted the regular hydrogen bonding between S0 and S1, causing the S1 adopted a β -bulge motif. Five hydrogen bond pairs are found to connect S0 and S1, including 118O-130N, 120O-128N, 122O-125N, 125O-122N and 128O-120N. The details of the hydrogen bonds in the crystal structure are shown in Figure 4.2.

Figure 4.2: Backbone hydrogen bonds between S0 and S1 β -strands in the crystal structure by Jan Steyaert et al.¹⁸⁹

The distances between the hydrogen bond donor and acceptor corresponding to that in the crystal structure are calculated based on the trajectory at 262.00 K with PSBC model to see if these hydrogen bonds can form during our computational study. From Figure 4.3, it can be seen that all of the five hydrogen bonds can form with a distance around 3.0 Å during the last half of the simulation, suggesting that the S0S1 β -sheet may form during our simulation.

Figure 4.3: Distances variations over simulation time of the five hydrogen bonds calculated based on the 262.00 K trajectory of PrP(116-131) using PSBC model.

The representative structure of PrP(116-131) is shown in Figure 4.4, A. It is the most populated cluster the 262.00K trajectory which accounts for 26.2% with an average backbone rmsd of 1.76 Å. The structure is shaped in a well-folded, long β -sheet, which is very similar to the S0S1 β -sheet in the crystal structure. Hydrogen bonding analysis is performed based on the representative structure to gain more insight into the detailed structure. A total of seven hydrogen bond pairs have been detected (Figure 4.4, B&C). Four of them are in accordance with the hydrogen bonds in the crystal structure by Jan Steyaert et al¹⁸⁹, including 118O-130N, 120O-128N, 122O-125N, and 128O-120N. Another hydrogen bond acceptor 126N is competing with 125N for the hydrogen bond donor 122O. Two new hydrogen bonds are found to form between 126O and 122N, and 130O and 118N.

Figure 4.4: (A): Representative structure of PrP(116-131). (B): Backbone hydrogen bonds detected in the representative structure. (C): Illustration of the hydrogen bonds in the representative structure.

Currently, it is very challenging to perform folding simulations of proteins with β structure or mixed α/β structure with existing empirical forcefields. Most of the existing forcefields have intrinsic secondary structure bias towards overstabilizing α -helix and understabilizing β -extended structure. In this PrP N-terminal fragment study, the REMD simulation with standard Amber 99SB forcefield shows a chaotic structural distribution with different secondary structural elements. In the simulation with PSBC model, by including the electrostatic polarization of backbone hydrogen bonds, we successfully obtained a well-folded β -sheet conformation of PrP(116-131), which is in comparably good agreement with crystal structure data¹⁸⁹ of S0 and S1.

It is interesting to observe that new β -strands formed at both the outer sides of the native S1S2 β -sheet. It seems that the native S1S2 β -sheet might be a nucleus for intermolecular β -enrichment in conformational conversion of PrP. In our previous study with PrP C-terminus (Chapter 3), the unfolding part of H2 moved close to S2 and refolded into a new β -strand, named as S3, which was in antiparallel alignment with S2. This new β -strand S3, together with native S1S2, formed a stable three-strand β -sheet. In this REMD study, the palindrome region of the PrP N-terminus formed new antiparallel β -sheet with native S1, which was very similar to newly resolved S0 in the crystal structure. It is reasonable to raise the question whether the four β -strands could coexist and form a more extensive β -sheet structure.

Figure 4.5: (A) Illustration of the three-stranded β -sheet structure of PrP118-228 in the crystal structure.¹⁸⁹ (B) Illustration of the constructed four-stranded β -sheet structure of PrP118-228 in our simulation.

Figure 4.6: Secondary structure evolution of the four-stranded β -sheet structure of PrP118-228 by DSSP analysis of the conventional MD simulation.

Since it is very time-consuming and computationally expensive to do the direct folding with large extended structure, we artificially built up the N-terminal residues corresponding to S0 region onto our previous three-stranded β -sheet structure with a reference to the hydrogen bonding in the crystal structure¹⁸⁹, achieving a more elaborate S0-S1-H1-S2-H2-S3-H3 structural fold (Figure 4.5, B). Then a 50-ns MD simulation with explicit water solvation at room temperature was carried out to check the stability of this architecture.

Secondary structure determination of the architecture was performed with DSSP analysis. As illustrated in Figure 4.6, it is pleasing to find that the four β -strands (S0, S1, S2, and S3) can coexist well. The four-stranded antiparallel β -sheet remain stable throughout the simulation, except that the irregular β -bulge at S1 showed some instability. In the end, we obtained a β -rich structure with a stable antiparallel four-stranded β -sheet adopting a S0-S1-H1-S2-H2-S3-H3 structural fold.

4.3.2 Formation of five-stranded β -sheet structure

In the above section, by a combination of PSBC model and REMD simulation, we successfully conducted folding simulation of the PrP fragment peptide corresponding to S0S1 region (residues 116-131), which is in comparably good agreement with the crystal structure data¹⁸⁹. By manually built up S0 onto our previous three-stranded β -sheet structure with S3, we successfully obtained a

four-stranded antiparallel β -sheet encompassing both PrP C-terminus and the palindrome region of N-terminus, which is proved to be quite stable in a subsequent 50-ns MD simulation. Then we continued our research with extra N-terminal residues, intending to further explore the structural fold of PrP N-terminus.

A fragment of PrP(105-122) was studied. Part of this fragment peptide is overlapped with PrP(116-131), corresponding to the residues involved in S0 formation (residues 118-122) in the above section. The extra sequence (residues 105-117) is at the downstream of S0, which is also within the amyloidogenic region of PrP N-terminus and have high β -structure formation propensity. Thus, this fragment was expected to adopt a β -sheet configure.

Here, we have also run a folding simulation of PrP(105-122) with REMD and PSBC updating scheme towards the secondary structure of this region, in a similar way to PrP(116-131) in the above section. The REMD simulation was composed of 8 replicas, setting as 264.00, 280.63, 298.17, 316.42, 335.91, 356.45, 378.12, 400.96 K. The starting structure was linear structure. The PSBC updating scheme was inserted into the REMD simulation and a total of 50 ns long simulation was obtained.

DSSP analysis was conducted based on the extracted trajectories to display the secondary structure evolution of PrP(105-122). As shown in Figure 4.7, the secondary structure distribution mainly consists of β conformations with

sporadically formation of α -helix and 3_{10} -helix. PrP(105-122) formed a two-stranded β -sheet structure. The first β -strand involves residues 115-119, which is within the palindrome region and is matched with S0 β -strand in the previous section. The second β -strand is formed by the extra residues in PrP N-terminus encompassing residues 108-112 and aligned antiparallel with S0. This new β -strand named as S-1. Gly114 and Ala115 form the β -corner.

Figure 4.7: Secondary structure evolution of PrP(105-122) by DSSP analysis of the trajectories with PSBC model at 264.00 K and 298.17K.

Cluster analysis is conducted based on the trajectory of 264.00K to identify the structural properties of the conformational substrates of this peptide folding. Most of the clusters are in β -sheet conformation. The representative structures of the three most populated clusters are shown in Figure 4.8. The largest cluster C1 accounts for 35.6% with an average backbone rmsd of 1.89 Å. The second largest cluster C2 has a population of 18.7% with an average backbone rmsd of 1.14 Å. The third one is 16.3% populated with an average backbone rmsd of 1.15 Å. All

these three clusters are in very similar conformation of long β -sheet with difference in the backbone hydrogen bond zipping.

Figure 4.8: Representative structures of three most populated clusters of PrP(105-122) by the cluster analysis of 264.00 K trajectory using PSBC model.

Hydrogen bond analysis of the clusters is carried out at 264.00K to gain detailed structural data. Six hydrogen bond pairs are found in cluster C1 which can be seen in Figure 4.9. However, residues of Ala118, Ala120 and Val 122 in S0 β -strand, which are involved in the hydrogen bond connection with new S-1 β -strand, have already participated in the hydrogen bond formation between S0 and S1 in previous section. The hydrogen bonds in cluster C2 are also investigated and depicted in Figure 4.10. In this structure, residues Ala115, Ala117, and Gly119 involving in the hydrogen bond formation between S0 and S-1 are in staggered positions with those residues (Ala118, Ala120 and Val 122) in the S0S1 β -sheet hydrogen bond connecting work.

Figure 4.9: Backbone hydrogen bonds detected in the representative structure in Cluster C1 of PrP(105-122).

Figure 4.10: Backbone hydrogen bonds detected in the representative structure in Cluster C2 of PrP(105-122).

With a reference to the conformation of cluster C2, we manually built up extra N-terminal residues (residues 104-117), corresponding to the potential S-1 β -strand, onto the four-stranded β -sheet structure in the above section, leading to a five-stranded β -sheet structure. A 50-ns all-atom MD simulation with explicit water solvation at room temperature was carried out to check the stability of this five-stranded β architecture.

From the DSSP analysis (Figure 4.11), we can see that this five-stranded β structure remain fairly stable. The final structure is illustrated in Figure 4.12. S-1 disappeared around 31 ns, but reformed after 36 ns in the simulation. S0 is elongated by forming additional hydrogen bonds with new S-1. There is a “hunch-up” at the structure of S1. It is possible due to the irregular hydrogen bonding with S0 as described in previous section. S2 and S3 are almost unchanged. Thus, we obtained a stable five-stranded antiparallel β -sheet structure of prion protein with continuously growing β -structure at its N-terminus.

Figure 4.11: Upper panel: Secondary structure evolution of the five-stranded β -sheet structure of PrP104-228 by DSSP analysis of the conventional MD simulation. Lower panel: The snapshot of final structure in the simulation.

4.3.3 Formation of six-stranded β -sheet structure

Then we went on to the next N-terminal fragment of PrP encompassing residues 95-112. This fragment contains an overlapped sequence with previous PrP(105-122) fragment, which is involved in the formation of S-1. In fact, this region has received special research interest in the past decades, because it contains two His residues (His96 and His111), which were considered to be possible binding sites for copper ions. Some studies suggested that copper coordination to His96 and His111 can cause main-chain conformation changes and induce the formation of β -structure in this region.⁷⁰ It is proposed that Cu-induced PrP conversion might occur in low-occupancy copper coordination form.⁷¹ An increase in β -sheet formation was observed in a CD spectroscopy study of PrP(91-115) peptide with low-occupancy Cu^{2+} coordination.⁶⁹ MD simulation on PrP(96-111) fragment with low-occupancy Cu^{2+} binding form demonstrated that Cu^{2+} bridging to His96 and His111 can promote localized β -sheet structure formation.⁷² Controversially, copper binding to single His residue was reported to be more prone to the conformational transition by Giachin et al.¹⁸⁴ Moreover, there is also study indicating that copper binding would impose precise backbone bending and weaken β -sheet formation propensity. Structure of β -sheet was observed in PrP(91-127) peptide without copper binding.⁷³ Since there seems no direct link between copper binding and β -sheet formation, our simulation does not include the copper ions.

Same as previous two PrP fragments, REMD simulation with PSBC updating scheme was employed to predict the secondary structure of this PrP(95-112). The REMD simulation was made up of 10 replicas, setting as 265.00, 279.96, 295.66, 312.11, 329.38, 347.47, 366.45, 386.36, 407.22, 429.16 K. The starting structure was linear structure. The PSBC updating scheme was inserted into the REMD production run and a total of 50 ns long simulation was obtained.

We also checked the secondary structure distributions of PrP(116-131) with DSSP analysis. As seen in Figure 4.12, a short two-stranded antiparallel β -sheet was found to shape up between residues 96-97 and residues 110-111. A small α -helix turn was formed in the region of residues 105-109. The first β -strand at the region of residues 96-97 is corresponding to S-1 β -strand in the previous section. The other one encompassing residues 110-111 is newly emerged, named as S-2.

Figure 4.12: Secondary structure evolution of PrP(95-112) by DSSP analysis of the trajectories with PSBC model at 265.00 K and 295.66K.

Figure 4.13: Representative structures of three most populated clusters of PrP(95-112) by the cluster analysis of the trajectory with PSBC model at 265.00 K.

Figure 4.14: Backbone hydrogen bonds detected in the representative structure in Cluster C1 of PrP(95-112).

The representative structures of the three most populated clusters at 265.00K are shown in Figure 4.13. The most populated cluster is a mixed structure of a small α -helical turn and a small β segment, which accounts for 53.6% and has a backbone rmsd of 0.83 Å. The hydrogen bonds identified in the β -sheet region are shown in Figure 4.14. Hydrogen bond can be formed between the two His

residues. The result of our simulation with PrP95-122 is very similar to that of the simulation of PrP(96-111) with Cu^{2+} bridging to His96 and His111 by Vogel et al⁷², suggesting that this region can form β -sheet with or without copper ions.

Referring to the dominant structure (cluster C1) of PrP (116-131) in the REMD simulation, we manually built up potential S-2 β -strand (residues 94-104) onto the five-stranded β -sheet structure in the above section, leading to a six-stranded β -sheet structure. A 50-ns all-atom MD simulation with explicit water solvation at room temperature was carried out to check the stability of this six-stranded β -sheet architecture.

As indicated in the DSSP analysis (Figure 4.15), the six-stranded β -sheet structure remains stable throughout the simulation. The final structure in the simulation is also shown in Figure 4.9. It is noted that the irregular β -bulge at S1 disappears during this simulation. Instead of that, regular hydrogen bonds form between S0 and S1 beta-strand, leading to a more flat β -sheet structure. Stable elongation of S1, S0, and S-1 is also observed. The cooperative interaction effect between the extensive β -strands contributes to this multiple β -strand structure. Finally, we obtained a β -rich intermediate of prion protein with a six-stranded antiparallel β -sheet conformation adopting S₋₂-S₋₁-S₀-S₁-H₁-S₂-H₂-S₃-H₃ structural organization.

Figure 4.15: Upper panel: Secondary structure evolution of the six-stranded β -sheet structure of PrP93-228 by DSSP analysis of the conventional MD simulation. Lower panel: The snapshot of final structure in the simulation.

Overall, our simulations provide a β -rich model with a six-stranded- β -sheet encompassing both C-terminus and the hydrophobic region of N-terminus. Our data strongly supports the hypothesis that the N-terminus may be involved in the β -enrichment in the monomer as one of the early events in the PrP^C to PrP^{Sc} conformational conversion. Moreover, the β -sheet formed at N-terminus is linked together with the C-terminus, leading to a more extensive β -sheet. Our study may help to provide one of the possible underlying routines for the intermolecular β -sheet nucleation.

The native S1S2 β -sheets might serve as intramolecular nucleus for β -enrichment. It is suggested that the exposed edges of regular β -sheets are prone to form hydrogen bonding interaction with other β -strand available, because they are already set up in the right backbone conformation.¹⁹⁰ For the native S1S2 β -stands in PrP, the exposed edge of S2 formed hydrogen bonds with the unfolding part of H2, leading to the formation of new β -stand S3 in our previous study with C-terminal domain (Chapter 3). Residues 90–126 of PrP N-terminus are unstructured in the native form and have been proposed to possess PrP^{Sc}-like neurotoxic and amyloidogenic properties. This region is located adjacent to the C-terminal domain. The regional proximity hints at the possibility of forming extensive β -structure with the C-terminal β -sheets. Here, in this work, the edge of S1 formed hydrogen bonding connection with the palindrome sequence, leading to the formation of the first β -strand at N-terminus, which was donated as S0. The

formation of new hydrogen bonds with native S1S2 β -strands also brought about the elongation of native strands. S1 is extended by two residues at its N-terminal direction by the forming β -hairpin with S0, while S2 extended by two residues at its C-terminus by the formation of new hydrogen bonds with S3. Two more β -strands were built up in the continuous sequence of N-terminus, resulting in a six-stranded β -sheet encompassing the prion C-terminus and the hydrophobic region of the N-terminus.

4.4 Conclusion

Elucidating the molecular mechanism of the PrP^C to PrP^{Sc} conformational conversion is a key issue in prion biology. To date, the detailed structure at atomic level of PrP^C has been only partially resolved, with an unresolved N-terminal region (residues 23-124) and a well-defined C-terminal globular domain (residues 125-231). In previous chapters, we studied the C-terminal domain and found that the α -to- β conformational changes occurred at H2 helix. Residues 90–126 at PrP N-terminus are unstructured in the native form, comprising a hydrophobic core (residues 106–126) and the so-called fifth copper binding site¹⁸² involving His96 and His111. This region has been proposed as a key region for the amyloidogenic and toxic nature of PrP. Furthermore, the location of this region is just adjacent to S1 β -strand at the C-terminal domain. The regional proximity and amyloidogenicity imply a special role for this region in the β -structure enrichment process. It is important to include the N-terminal region in the prion study.

Using a combination of REMD simulation and PSBC updating scheme, we investigated the secondary structure of isolated N-terminal fragment peptide. Spontaneously β -sheet formation was observed in the simulation, confirming that the N-terminal region has β -structure formation property. Build upon the three-stranded β -sheet conformation in previous PrP C-terminal study, we incrementally modeled the unstructured N-terminal residues onto the globular C-terminal domain, with a reference to the structure obtained in REMD simulations with PrP N-terminal fragments. Finally, we successfully got a stable six-stranded antiparallel β -sheet structure of prion protein. This architecture contains four consecutive β -strands, among which three (S0, S-1, S-2) are formed at the amyloidogenic region of N-terminus (residues 90-126) and the remaining one (S1) is at the starting region of the C-terminus. For the two discontinuous β -strands, one is the native S2 β -strand, and the other one is derived from the unfolding part of H2 helix, which has been demonstrated to be intrinsically unstable and have high β -structure formation propensity. The native S1S2 β -sheets might serve as nucleus for intramolecular β -enrichment. The β -sheet grows by addition of new strand along the exposed edge side.

Our model provides a first possible atomic-level view of a six-stranded antiparallel β -sheet conformation spanning both N- and C-terminal regions of prion protein. The regions involved in the β -sheet formation in our study have all been reported to be capable of forming β -sheet structure or amyloid oligomers in a

wide range of experimental and theoretical studies with full-length or truncated form of prion protein. Moreover, the N- and C-terminal β -structure are connected together to form a β -nucleation in prion monomer. This intermolecular β -sheet nucleation can provide important insights towards the mechanism of the conformational conversion of PrP^{C} to PrP^{Sc} and may be core for subsequent oligomerization and aggregation.

Chapter 5 Summary and future directions

Prion diseases are fatal neurodegenerative disorders characterized by the misfolding of normal prion protein and the accumulation of amyloid deposits in the central nervous system. Understanding the conversion of the α -helix-rich cellular prion protein (PrP^{C}) into pathogenic β -sheet-rich isoform (PrP^{Sc}) is fundamental to unravel the early stages of the development of prion diseases. PrP^{C} consists of an unstructured N-terminus and a well-defined C-terminal globular domain. Due to the insolubility and aggregation nature of PrP^{Sc} , it is difficult to obtain its atomic-level structure by classic detection techniques. MD simulations have served as an alternative and invaluable means to study the mechanism of prion misfolding and provide microscopic information on the structure and dynamics of PrP^{Sc} . In this thesis, different strategies of MD simulations have been utilized to investigate the conformational conversion of human prion protein.

Firstly, we examined the pH effect on the conformational changes of prion protein since many studies have suggested that acidic pH may have implications for the presumed formation of PrP^{Sc} . Room temperature simulations of human PrP^{C} C-terminal domain (residues 125-228, Figure 5.1 A) at different pH levels (pH ~7, pH~5 and pH~3) were performed. Our data showed that, in the initial stage of the conversion, no obvious secondary structural changes occurred in H1, H3, S1 and

S2. Displacement of H1 was observed. The movement was due to the disruption of salt bridges in acidic pH simulations and increased with decreased pH levels. H2 exhibited the significant structural perturbation with partially unfolding at its C-terminus (Figure 5.1 B). The extent of unfolding increased toward acidic pH. Our findings are in agreement with a range of experimental and theoretical studies, suggesting that acidic pH can trigger the conformational changes and H2 may play an early role in the conversion.

Next, we raised the simulation temperature to a mildly high temperature, aiming to speed up the denaturation process and gain further insights into the conformational conversion. A 200 ns simulation at 350 K and pH~3 conditions was performed. During this elevated temperature simulation, further unfolding at H2 C-terminus and H3 N-terminus were observed (Figure 5.1 C). Elongation of the native S1S2 β -sheet was also noted. An essential dynamic regards to the loop between partially denatured H2 and H3 has been identified by PCA analysis — it showed a movement tendency toward the native S1S2 β -sheet. This movement aroused our interest because this region has been reported to have strong β -structure formation propensity. It is reasonable to hypothesize that new β -strand formed at this region by adding onto the native β -strands (Figure 5.1 D). We verified this hypothesis by a simulated annealing simulation. A new β -strand, named as S3, formed at the unfolding part of H2 and aligned antiparallel with S2 (Figure 5.1 E). This newfound S3 was elongated and further stabilized by three

additional hydrogen bonds in subsequent 100 ns room temperature MD simulation. Thus, we successfully obtained a stable three-strand β -sheet conformation of prion protein, adopting a S1-H1-S2-H2-S3-H3 structural organization.

In the third study, we attempted to investigate the β -constitution in PrP N-terminus with recently developed PSBC model. The unstructured N-terminal residues 90-126 have been reported to be associated with the amyloidogenicity and neurotoxicity of PrP^{Sc}. By implementing PSBC into REMD simulation, we simulated the folding of PrP(116-131). A well-organized two-stranded β -sheet was obtained, which was in good agreement with the S0S1 region of a recent crystal structure¹⁸⁹. By manually attaching the newfound S0 onto our previous three-stranded β -sheet structure, we obtained a four-stranded structure (Figure 5.1 F), which was proved to be quite stable in a subsequent 50-ns conventional MD simulation. Then we extended our studies to additional N-terminal fragments of PrP(105-122) and PrP(95-112) in a similar way. Two more β -strands, named as S-1 and S-2, were found. These two β -strands were modeled onto the globular domain successively. Finally, we obtained a stable six-stranded antiparallel β -sheet structure (Figure 5.1 H) spanning both the C-terminus and the amyloidogenic region of PrP N-terminus. This β -rich structure can provide important insights towards the mechanism of the conformational conversion of PrP^C to PrP^{Sc} and may be core for subsequent oligomerization and aggregation.

Figure 5.1: (A) NMR structure of PrP125-228. (B) Unfolding started from the H2 C-ter (C) Further unfolding at H2 C-ter and H3 N-ter. (D) H2H3 loop aligns closely to native S1S2. (E) Three-stranded antiparallel β -sheet structure of PrP125-228. (F) Four-stranded antiparallel β -sheet structure of PrP118-228. (G) Five-stranded antiparallel β -sheet structure of PrP104-228. (H) Six-stranded antiparallel β -sheet structure of PrP93-228

The pathogenic process of prion diseases can be divided into two steps. The first step is the conversion of α -helix-rich PrP^C into β -sheet-rich isoform PrP^{Sc}. The second step is that the infectious intermediates propagate in an autocatalytic manner to form amyloid fibrils. In this work, we have studied the conformational conversion of prion monomer and obtained a β -rich conformer with a six-stranded antiparallel β -sheet spanning both the C-terminal domain and the amyloidogenic region of N-terminus. The remaining N-terminal region contains a special sequence of four perfect OR repeats of PHGGGWGQ (residues 60-91) among different mammalian species and additional homologous “pseudorepeat” of PQGGGGWGQ (residues 51-59) in human prion protein. This region has attracted many research interests in the past decades because of its structural features and its ability to binding copper ions. To date, the role of this region in the pathological process is still elusive. We are planning to study the structure features of this region with MD simulation to thoroughly investigate the β composition in PrP^{Sc} and provide a comprehensive understanding towards the mechanism of prion misfolding. Dimer, trimer, or even oligomers of prion protein could be built up based on the β -rich structure of prion monomer to further explore the possible aggregation pathway.

Publications

1. Hong Yang, Dawei Zhang. Formation of four-stranded β -sheet towards the elucidation of the conformational conversion of prion protein. (under review)
2. Hong Yang, Dawei Zhang. Probing the β constitution at the N-terminus of prion protein. (under preparation)

Appendix

The amino sequence of human prion protein (residues 23-231)

23	K	K	R	P	K	P	G	G	W	N	T	33	G	G	S	R	Y	P	G	Q	G
43	S	P	G	G	N	R	Y	P	P	Q	G	53	G	G	G	W	G	Q	P	H	G
63	G	G	W	G	Q	P	H	G	G	G	W	73	G	Q	P	H	G	G	G	W	G
83	Q	P	H	G	G	G	W	G	Q	G	G	93	G	T	H	S	Q	W	N	K	P
103	S	K	P	K	T	N	M	K	H	M	A	113	G	A	A	A	A	G	A	V	V
123	G	G	L	G	G	Y	M	L	G	S	A	133	M	S	R	P	I	I	H	F	G
143	S	D	Y	E	D	R	Y	Y	R	E	N	153	M	H	R	Y	P	N	Q	V	Y
163	Y	R	P	M	D	E	Y	S	N	Q	N	173	N	F	V	H	D	C	V	N	I
183	T	I	K	Q	H	T	V	T	T	T	T	193	K	G	E	N	F	T	E	T	D
203	V	K	M	M	E	R	V	V	E	Q	M	213	C	I	T	Q	Y	E	R	E	S
223	Q	A	Y	Y	Q	R	G	S	S			231									

References

- (1) Prusiner, S. B. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, 95, 13363.
- (2) Head, M. W.; Ironside, J. W. *Neuropathology and Applied Neurobiology* **2012**, 38, 296.
- (3) Aguzzi, A.; Calella, A. M. *Physiological Reviews* **2009**, 89, 1105.
- (4) Imran, M.; Mahmood, S. *Virology Journal* **2011**, 8.
- (5) Aguzzi, A.; Haass, C. *Science* **2003**, 302, 814.
- (6) Meyer-Luehmann, M.; Coomaraswamy, J.; Bolmont, T.; Kaeser, S.; Schaefer, C.; Kilger, E.; Neuenschwander, A.; Abramowski, D.; Frey, P.; Jaton, A. L.; Vigouret, J. M.; Paganetti, P.; Walsh, D. M.; Mathews, P. M.; Ghiso, J.; Staufenbiel, M.; Walker, L. C.; Jucker, M. *Science* **2006**, 313, 1781.
- (7) Alper, T.; Cramp, W. A.; Haig, D. A.; Clarke, M. C. *Nature* **1967**, 214, 764.
- (8) Prusiner, S. B.; Bolton, D. C.; Groth, D. F.; Bowman, K. A.; Cochran, S. P.; McKinley, M. P. *Biochemistry* **1982**, 21, 6942.
- (9) Soto, C.; Estrada, L.; Castilla, J. *Trends in Biochemical Sciences* **2006**, 31, 150.
- (10) Griffith, J. S. *Nature* **1967**, 215, 1043.
- (11) Dohura, K.; Perryman, S.; Race, R.; Chesebro, B. *Microbial Pathogenesis* **1995**, 18, 1.
- (12) Chien, P.; Weissman, J. S.; DePace, A. H. *Annual Review of Biochemistry* **2004**, 73, 617.
- (13) Wopfner, F.; Weidenhofer, G.; Schneider, R.; von Brunn, A.; Gilch, S.; Schwarz, T. F.; Werner, T.; Schatzl, H. M. *J Mol Biol* **1999**, 289, 1163.
- (14) Benvegna, S.; Poggiolini, I.; Legname, G. *Journal of Comparative*

Neurology **2010**, 518, 1879.

(15)Poggiolini, I.; Legname, G. *Plos One* **2012**, 7.

(16)Sales, N.; Hassig, R.; Rodolfo, K.; Di Giamberardino, L.; Traiffort, E.; Ruat, M.; Fretier, P.; Moya, K. L. *European Journal of Neuroscience* **2002**, 15, 1163.

(17)Garcia, F. L.; Zahn, R.; Riek, R.; Wuthrich, K. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, 8334.

(18)Billeter, M.; Riek, R.; Wider, G.; Hornemann, S.; Glockshuber, R.; Wuthrich, K. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, 94, 7281.

(19)Riek, R.; Hornemann, S.; Wider, G.; Billeter, M.; Glockshuber, R.; Wuthrich, K. *Nature* **1996**, 382, 180.

(20)Luhrs, T.; Riek, R.; Guntert, P.; Wuthrich, K. *J Mol Biol* **2003**, 326, 1549.

(21)Gossert, A. D.; Bonjour, S.; Lysek, D. A.; Fiorito, F.; Wuthrich, K. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, 102, 646.

(22)Knaus, K. J.; Morillas, M.; Swietnicki, W.; Malone, M.; Surewicz, W. K.; Yee, V. C. *Nature Structural Biology* **2001**, 8, 770.

(23)Haire, L. F.; Whyte, S. M.; Vasisht, N.; Gill, A. C.; Verma, C.; Dodson, E. J.; Dodson, G. G.; Bayley, P. M. *J Mol Biol* **2004**, 336, 1175.

(24)Guest, W. C.; Plotkin, S. S.; Cashman, N. R. *J. Toxicol. Env. Health Part A* **2011**, 74, 154.

(25)Ilc, G.; Giachin, G.; Jaremko, M.; Jaremko, L.; Benetti, F.; Plavec, J.; Zhukov, I.; Legname, G. *Plos One* **2010**, 5.

(26)Perez, D. R.; Damberger, F. F.; Wuthrich, K. *J Mol Biol* **2010**, 400, 121.

(27)Zahn, R.; Liu, A. Z.; Luhrs, T.; Riek, R.; von Schroetter, C.; Garcia, F. L.; Billeter, M.; Calzolari, L.; Wider, G.; Wuthrich, K. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, 145.

- (28)Fujiwara, K.; Toda, H.; Ikeguchi, M. *Bmc Structural Biology* **2012**, 12.
- (29)Morrissey, M. P.; Shakhnovich, E. I. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, 96, 11293.
- (30)Calzolari, L.; Zahn, R. *Journal of Biological Chemistry* **2003**, 278, 35592.
- (31)Wille, H.; Michelitsch, M. D.; Guenebaut, V.; Supattapone, S.; Serban, A.; Cohen, F. E.; Agard, D. A.; Prusiner, S. B. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, 3563.
- (32)DeMarco, M. L.; Daggett, V. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, 101, 2293.
- (33)Eghiaian, F.; Daubenfeld, T.; Quenet, Y.; van Audenhaege, M.; Bouin, A. P.; van der Rest, G.; Grosclaude, J.; Rezaei, H. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, 104, 7414.
- (34)Liu, A. Z.; Riek, P.; Zahn, R.; Hornemann, S.; Glockshuber, R.; Wuthrich, K. *Biopolymers* **1999**, 51, 145.
- (35)Ziegler, J.; Sticht, H.; Marx, U. C.; Muller, W.; Rosch, P.; Schwarzingier, S. *Journal of Biological Chemistry* **2003**, 278, 50175.
- (36)Xu, Z.; Prigent, S.; Deslys, J. P.; Rezaei, H. *Journal of Biological Chemistry* **2011**, 286, 40060.
- (37)Adrover, M.; Pauwels, K.; Prigent, S.; de Chiara, C.; Xu, Z.; Chapuis, C.; Pastore, A.; Rezaei, H. *Journal of Biological Chemistry* **2010**, 285, 21004.
- (38)Chen, J.; Thirumalai, D. *Biochemistry* **2013**, 52, 310.
- (39)Singh, J.; Kumar, H.; Sabareesan, A. T.; Udgaonkar, J. B. *Journal of the American Chemical Society* **2014**, 136, 16704.
- (40)Dima, R. I.; Thirumalai, D. *Biophysical Journal* **2002**, 83, 1268.
- (41)Tizzano, B.; Palladino, P.; De Capua, A.; Marasco, D.; Rossi, F.; Benedetti, E.; Pedone, C.; Ragone, R.; Ruvo, M. *Proteins-Structure Function and Bioinformatics* **2005**, 59, 72.
- (42)Ronga, L.; Palladino, P.; Ragone, R.; Benedetti, E.; Rossi, F. *Journal of*

Peptide Science **2009**, *15*, 30.

(43) Yamamoto, N. *Journal of Physical Chemistry B* **2014**, *118*, 9826.

(44) Pappalardo, M.; Milardi, D.; La Rosa, C.; Zannoni, C.; Rizzarelli, E.; Grasso, D. *Chemical Physics Letters* **2004**, *390*, 511.

(45) Dima, R. I.; Thirumalai, D. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 15335.

(46) Chebaro, Y.; Derreumaux, P. *Journal of Physical Chemistry B* **2009**, *113*, 6942.

(47) Yamaguchi, K. I.; Matsumoto, T.; Kuwata, K. *Biochemistry* **2008**, *47*, 13242.

(48) Thompson, A.; White, A. R.; McLean, C.; Masters, C. L.; Cappai, R.; Barrow, C. J. *Journal of Neuroscience Research* **2000**, *62*, 293.

(49) Mead, S. *European Journal of Human Genetics* **2006**, *14*, 273.

(50) Hosszu, L. L. P.; Tattum, M. H.; Jones, S.; Trevitt, C. R.; Wells, M. A.; Waltho, J. P.; Collinge, J.; Jackson, G. S.; Clarke, A. R. *Biochemistry* **2010**, *49*, 8729.

(51) Chiesa, R.; Pestronk, A.; Schmidt, R. E.; Tourtellotte, V. G.; Ghetti, B.; Piccardo, P.; Harris, D. A. *Neurobiology of Disease* **2001**, *8*, 279.

(52) Zhang, Y. B.; Swietnicki, W.; Zagorski, M. G.; Surewicz, W. K.; Sonnichsen, F. D. *Journal of Biological Chemistry* **2000**, *275*, 33650.

(53) Lee, S.; Antony, L.; Hartmann, R.; Knaus, K. J.; Surewicz, K.; Surewicz, W. K.; Yee, V. C. *Embo Journal* **2010**, *29*, 251.

(54) van der Kamp, M. W.; Daggett, V. *Protein Engineering Design & Selection* **2009**, *22*, 461.

(55) Wuthrich, K.; Riek, R. *Advances in Protein Chemistry, Vol 57* **2001**, *57*, 55.

(56) Marino, T.; Russo, N.; Toscano, M.; Pavelka, M. *Interdisciplinary Sciences-Computational Life Sciences* **2010**, *2*, 57.

- (57)Pan, K.; Yi, C.-W.; Chen, J.; Liang, Y. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2015**, 1854, 907.
- (58)Walter, E. D.; Stevens, D. J.; Visconte, M. P.; Millhauser, G. L. *Journal of the American Chemical Society* **2007**, 129, 15440.
- (59)Shearer, J.; Rosenkoetter, K. E.; Callan, P. E.; Pham, C. *Inorganic Chemistry* **2011**, 50, 1173.
- (60)Millhauser, G. L. In *Annual Review of Physical Chemistry* 2007; Vol. 58, p 299.
- (61)Luczkowski, M.; Kozłowski, H.; Stawikowski, M.; Rolka, K.; Gaggelli, E.; Valensin, D.; Valensin, G. *Journal of the Chemical Society-Dalton Transactions* **2002**, 2269.
- (62)Singh, R. K.; Srivastava, V. C.; Singh, U. P. *Protein and Peptide Letters* **2011**, 18, 1280.
- (63)Lau, A.; McDonald, A.; Daude, N.; Mays, C. E.; Walter, E. D.; Aglietti, R.; Mercer, R. C. C.; Wohlgemuth, S.; van der Merwe, J.; Yang, J.; Gapeschina, H.; Kim, C.; Grams, J.; Shi, B.; Wille, H.; Balachandran, A.; Schmitt-Ulms, G.; Safar, J. G.; Millhauser, G. L.; Westaway, D. *Embo Molecular Medicine* **2015**, 7, 339.
- (64)Burns, C. S.; Aronoff-Spencer, E.; Dunham, C. M.; Lario, P.; Avdievich, N. I.; Antholine, W. E.; Olmstead, M. M.; Vrieling, A.; Gerfen, G. J.; Peisach, J.; Scott, W. G.; Millhauser, G. L. *Biochemistry* **2002**, 41, 3991.
- (65)Zahn, R. *J Mol Biol* **2003**, 334, 477.
- (66)Chen, S. G.; Teplow, D. B.; Parchi, P.; Teller, J. K.; Gambetti, P.; Autiliogambetti, L. *Journal of Biological Chemistry* **1995**, 270, 19173.
- (67)Jimenez-Huete, A.; Lievens, P. M. J.; Vidal, R.; Piccardo, P.; Ghetti, B.; Tagliavini, F.; Frangione, B.; Prelli, F. *American Journal of Pathology* **1998**, 153, 1561.
- (68)D'Angelo, P.; Della Longa, S.; Arcovito, A.; Mancini, G.; Zitolo, A.; Chillemi, G.; Giachin, G.; Legname, G.; Benetti, F. *Biochemistry* **2012**, 51, 6068.

- (69) Jones, C. E.; Abdelraheim, S. R.; Brown, D. R.; Viles, J. H. *Journal of Biological Chemistry* **2004**, 279, 32018.
- (70) Younan, N. D.; Klewpatinond, M.; Davies, P.; Ruban, A. V.; Brown, D. R.; Viles, J. H. *J Mol Biol* **2011**, 410, 369.
- (71) Pushie, M. J.; Vogel, H. J. *J. Toxicol. Env. Health Part A* **2009**, 72, 1040.
- (72) Pushie, M. J.; Rauk, A.; Jirik, F. R.; Vogel, H. J. *Biomaterials* **2009**, 22, 159.
- (73) Migliorini, C.; Sinicropi, A.; Kozłowski, H.; Luczkowski, M.; Valensin, D. *Journal of Biological Inorganic Chemistry* **2014**, 19, 635.
- (74) Bocharova, O. V.; Breydo, L.; Salnikov, V. V.; Baskakov, I. V. *Biochemistry* **2005**, 44, 6776.
- (75) Hijazi, N.; Shaked, Y.; Rosenmann, H.; Ben-Hur, T.; Gabizon, R. *Brain Research* **2003**, 993, 192.
- (76) Qin, K. F.; Yang, D. S.; Yang, Y.; Chishti, M. A.; Meng, L. J.; Kretzschmar, H. A.; Yip, C. M.; Fraser, P. E.; Westaway, D. *Journal of Biological Chemistry* **2000**, 275, 19121.
- (77) McKenzie, D.; Bartz, J.; Mirwald, J.; Olander, D.; Marsh, R.; Aiken, J. *Journal of Biological Chemistry* **1998**, 273, 25545.
- (78) Pauly, P. C.; Harris, D. A. *Journal of Biological Chemistry* **1998**, 273, 33107.
- (79) Herms, J.; Tings, T.; Gall, S.; Madlung, A.; Giese, A.; Siebert, H.; Schurmann, P.; Windl, O.; Brose, N.; Kretzschmar, H. *Journal of Neuroscience* **1999**, 19, 8866.
- (80) Kuwahara, C.; Takeuchi, A. M.; Nishimura, T.; Haraguchi, K.; Kubosaki, A.; Matsumoto, Y.; Saeki, K.; Yokoyama, T.; Itohara, S.; Onodera, T. *Nature* **1999**, 400, 225.
- (81) Weise, J.; Crome, O.; Sandau, R.; Schulz-Schaeffer, W.; Bahr, M.; Zerr, I. *Neuroscience Letters* **2004**, 372, 146.
- (82) Watt, N. T.; Taylor, D. R.; Gillott, A.; Thomas, D. A.; Perera, W. S. S.;

- Hooper, N. M. *Journal of Biological Chemistry* **2005**, 280, 35914.
- (83)Rangel, A.; Burgaya, F.; Gavin, R.; Soriano, E.; Aguzzi, A.; del Rio, J. A. *Journal of Neuroscience Research* **2007**, 85, 2741.
- (84)Mouillet-Richard, S.; Ermonval, M.; Chebassier, C.; Laplanche, J. L.; Lehmann, S.; Launay, J. M.; Kellermann, O. *Science* **2000**, 289, 1925.
- (85)Spielhaupter, C.; Schatzl, H. M. *Journal of Biological Chemistry* **2001**, 276, 44604.
- (86)Chiarini, L. B.; Freitas, A. R. O.; Zanata, S. M.; Brentani, R. R.; Martins, V. R.; Linden, R. *Embo Journal* **2002**, 21, 3317.
- (87)Bremer, J.; Baumann, F.; Tiberi, C.; Wessig, C.; Fischer, H.; Schwarz, P.; Steele, A. D.; Toyka, K. V.; Nave, K.-A.; Weis, J.; Aguzzi, A. *Nature Neuroscience* **2010**, 13, 310.
- (88)Aguzzi, A.; Baumann, F.; Bremer, J. In *Annual Review of Neuroscience* 2008; Vol. 31, p 439.
- (89)Li, A.; Christensen, H. M.; Stewart, L. R.; Roth, K. A.; Chiesa, R.; Harris, D. A. *Embo Journal* **2007**, 26, 548.
- (90)Brown, D. R.; Schmidt, B.; Kretzschmar, H. A. *Nature* **1996**, 380, 345.
- (91)Alier, K.; Li, Z.; MacTavish, D.; Westaway, D.; Jhamandas, J. H. *Journal of Neuroscience Research* **2010**, 88, 2217.
- (92)Walsh, P.; Yau, J.; Simonetti, K.; Sharpe, S. *Biochemistry* **2009**, 48, 5779.
- (93)Dorosz, J.; Volinsky, R.; Bazar, E.; Kolusheva, S.; Jelinek, R. *Langmuir* **2009**, 25, 12501.
- (94)Degioia, L.; Selvaggini, C.; Ghibaudi, E.; Diomede, L.; Bugiani, O.; Forloni, G.; Tagliavini, F.; Salmona, M. *Journal of Biological Chemistry* **1994**, 269, 7859.
- (95)Gasset, M.; Baldwin, M. A.; Lloyd, D. H.; Gabriel, J. M.; Holtzman, D. M.; Cohen, F.; Fletterick, R.; Prusiner, S. B. *Proceedings of the National Academy of Sciences of the United States of America* **1992**, 89, 10940.

- (96) Forloni, G.; Angeretti, N.; Chiesa, R.; Monzani, E.; Salmona, M.; Bugiani, O.; Tagliavini, F. *Nature* **1993**, 362, 543.
- (97) Daidone, I.; Amadei, A.; Di Nola, A. *Proteins-Structure Function and Bioinformatics* **2005**, 59, 510.
- (98) Jobling, M. F.; Stewart, L. R.; White, A. R.; McLean, C.; Friedhuber, A.; Maher, F.; Beyreuther, K.; Masters, C. L.; Barrow, C. J.; Collins, S. J.; Cappai, R. *Journal of Neurochemistry* **1999**, 73, 1557.
- (99) Khalili-Shirazi, A.; Summers, L.; Linehan, J.; Mallinson, G.; Anstee, D.; Hawke, S.; Jackson, G. S.; Collinge, J. *Journal of General Virology* **2005**, 86, 2635.
- (100) Kocisko, D. A.; Lansbury, P. T.; Caughey, B. *Biochemistry* **1996**, 35, 13434.
- (101) Sasaki, K.; Gaikwad, J.; Hashiguchi, S.; Kubota, T.; Sugimura, K.; Kremer, W.; Kalbitzer, H. R.; Akasaka, K. *Prion* **2008**, 2, 118.
- (102) Cheng, H.-M.; Tsai, T. W. T.; Huang, W. Y. C.; Lee, H.-K.; Lian, H.-Y.; Chou, F.-C.; Mou, Y.; Chan, J. C. C. *Biochemistry* **2011**, 50, 6815.
- (103) Lee, S. W.; Mou, Y.; Lin, S. Y.; Chou, F. C.; Tseng, W. H.; Chen, C.; Lu, C. Y. D.; Yu, S. S. F.; Chan, J. C. C. *J Mol Biol* **2008**, 378, 1142.
- (104) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. O.; Riek, C.; Eisenberg, D. *Nature* **2007**, 447, 453.
- (105) Norstrom, E. M.; Mastrianni, J. A. *Journal of Biological Chemistry* **2005**, 280, 27236.
- (106) Brown, D. R. *Molecular and Cellular Neuroscience* **2000**, 15, 66.
- (107) Temussi, P. A.; Masino, L.; Pastore, A. *Embo Journal* **2003**, 22, 355.
- (108) Huang, Z. W.; Prusiner, S. B.; Cohen, F. E. *Folding & Design* **1996**, 1, 13.

- (109) Pan, K. M.; Baldwin, M.; Nguyen, J.; Gasset, M.; Serban, A.; Groth, D.; Mehlhorn, I.; Huang, Z. W.; Fletterick, R. J.; Cohen, F. E.; Prusiner, S. B. *Proceedings of the National Academy of Sciences of the United States of America* **1993**, *90*, 10962.
- (110) Prusiner, S. B.; Groth, D. F.; Bolton, D. C.; Kent, S. B.; Hood, L. E. *Cell* **1984**, *38*, 127.
- (111) Kitamoto, T.; Tateishi, J.; Tashima, T.; Takeshita, I.; Barry, R. A.; Dearmond, S. J.; Prusiner, S. B. *Annals of Neurology* **1986**, *20*, 204.
- (112) Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. F. *J Mol Biol* **1997**, *273*, 729.
- (113) Tycko, R.; Savtchenko, R.; Ostapchenko, V. G.; Makarava, N.; Baskakov, I. V. *Biochemistry* **2010**, *49*, 9488.
- (114) Walsh, P.; Simonetti, K.; Sharpe, S. *Structure* **2009**, *17*, 417.
- (115) Lin, N. S.; Chao, J. C. H.; Cheng, H. M.; Chou, F. C.; Chang, C. F.; Chen, Y. R.; Chang, Y. J.; Huang, S. J.; Chan, J. C. C. *Chemistry-a European Journal* **2010**, *16*, 5492.
- (116) Makarava, N.; Baskakov, I. V. *Journal of Biological Chemistry* **2008**, *283*, 15988.
- (117) Lu, X. J.; Wintrode, P. L.; Surewicz, W. K. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 1510.
- (118) Jarrett, J. T.; Lansbury, P. T. *Cell* **1993**, *73*, 1055.
- (119) Collinge, J.; Clarke, A. R. *Science* **2007**, *318*, 930.
- (120) Baskakov, I. V. *Febs Journal* **2007**, *274*, 3756.
- (121) Kaye, R.; Head, E.; Thompson, J. L.; McIntire, T. M.; Milton, S. C.; Cotman, C. W.; Glabe, C. G. *Science* **2003**, *300*, 486.
- (122) Govaerts, C.; Wille, H.; Prusiner, S. B.; Cohen, F. E. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 8342.

- (123) Cobb, N. J.; Sonnichsen, F. D.; McHaourab, H.; Surewicz, W. K. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 18946.
- (124) Kocisko, D. A.; Priola, S. A.; Raymond, G. J.; Chesebro, B.; Lansbury, P. T.; Caughey, B. *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 3923.
- (125) Barducci, A.; Chelli, R.; Procacci, P.; Schettino, V. *Biophysical Journal* **2005**, *88*, 1334.
- (126) Xu, Z. J.; Lazim, R.; Mei, Y.; Zhang, D. W. *Chemical Physics Letters* **2012**, *539*, 239.
- (127) Best, R. B.; Hummer, G. *Journal of Physical Chemistry B* **2009**, *113*, 9004.
- (128) Ji, C. G.; Zhang, J. Z. H. *Journal of Physical Chemistry B* **2009**, *113*, 13898.
- (129) Wei, C.; Tung, D.; Yip, Y. M.; Mei, Y.; Zhang, D. *Journal of Chemical Physics* **2011**, *134*.
- (130) Gao, Y.; Lu, X.; Duan, L. L.; Zhang, J. Z. H.; Mei, Y. *Journal of Physical Chemistry B* **2012**, *116*, 549.
- (131) Ji, C.; Mei, Y.; Zhang, J. Z. H. *Biophysical Journal* **2008**, *95*, 1080.
- (132) Duan, L. L.; Mei, Y.; Zhang, D.; Zhang, Q. G.; Zhang, J. Z. H. *Journal of the American Chemical Society* **2010**, *132*, 11159.
- (133) Ji, C. G.; Xiao, X. D.; Zhang, J. Z. H. *Journal of Chemical Theory and Computation* **2012**, *8*, 2157.
- (134) Yu, H. B.; van Gunsteren, W. F. *Computer Physics Communications* **2005**, *172*, 69.
- (135) Lysek, D. A.; Schorn, C.; Nivon, L. G.; Esteve-Moya, V.; Christen, B.; Calzolari, L.; von Schroetter, C.; Fiorito, F.; Herrmann, T.; Guntert, P.; Wuthrich, K. *Proceedings of the National Academy of Sciences of the United*

States of America **2005**, 102, 640.

(136) Klatzo, I.; Gajdusek, D. C.; Zigas, V. *Journal of Neuropathology and Experimental Neurology* **1959**, 18, 335.

(137) Hadlow, W. J. *Laboratory Investigation* **1959**, 8, 1478.

(138) Collinge, J.; Whitfield, J.; McKintosh, E.; Beck, J.; Mead, S.; Thomas, D. J.; Alpers, M. P. *Lancet* **2006**, 367, 2068.

(139) Gianni, S.; Ivarsson, Y.; De Simone, A.; Travaglini-Allocatelli, C.; Brunori, M.; Vendruscolo, M. *Nature Structural & Molecular Biology* **2010**, 17, 1431.

(140) Sunyach, C.; Jen, A.; Deng, J.; Fitzgerald, K. T.; Frobert, Y.; Grassi, J.; McCaffrey, M. W.; Morris, R. *Embo Journal* **2003**, 22, 3591.

(141) Marijanovic, Z.; Caputo, A.; Campana, V.; Zurzolo, C. *Plos Pathogens* **2009**, 5.

(142) Arnold, J. E.; Tipler, C.; Laszlo, L.; Hope, J.; Landon, M.; Mayer, R. J. *Journal of Pathology* **1995**, 176, 403.

(143) Lee, R. J.; Wang, S.; Low, P. S. *Biochimica Et Biophysica Acta-Molecular Cell Research* **1996**, 1312, 237.

(144) Sanghera, N.; Wall, M.; Venien-Bryan, C.; Pinheiro, T. J. T. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2008**, 1784, 873.

(145) Borchelt, D. R.; Taraboulos, A.; Prusiner, S. B. *Journal of Biological Chemistry* **1992**, 267, 16188.

(146) Hornemann, S.; Glockshuber, R. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, 95, 6010.

(147) Swietnicki, W.; Morillas, M.; Chen, S. G.; Gambetti, P.; Surewicz, W. K. *Biochemistry* **2000**, 39, 424.

(148) Zou, W. Q.; Cashman, N. R. *Journal of Biological Chemistry* **2002**, 277, 43942.

(149) Prusiner, S. B. *Science* **1991**, 252, 1515.

- (150) Giachin, G.; Biljan, I.; Ilc, G.; Plavec, J.; Legname, G. *Molecules* **2013**, *18*, 9451.
- (151) Kuwata, K.; Matumoto, T.; Cheng, H.; Nagayama, K.; James, T. L.; Roder, H. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 14790.
- (152) Norstrom, E. M.; Mastrianni, J. A. *Journal of Virology* **2006**, *80*, 8521.
- (153) Viles, J. H.; Donne, D.; Kroon, G.; Prusiner, S. B.; Cohen, F. E.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2001**, *40*, 2743.
- (154) Hosszu, L. L. P.; Baxter, N. J.; Jackson, G. S.; Power, A.; Clarke, A. R.; Waltho, J. P.; Craven, C. J.; Collinge, J. *Nature Structural Biology* **1999**, *6*, 740.
- (155) Speare, J. O.; Rush, T. S.; Bloom, M. E.; Caughey, B. *Journal of Biological Chemistry* **2003**, *278*, 12522.
- (156) Lingenheil, M.; Denschlag, R.; Tavan, P. *European Biophysics Journal with Biophysics Letters* **2010**, *39*, 1177.
- (157) Ribeiro, A.; de Alencastro, R. B. *Journal of Molecular Graphics & Modelling* **2013**, *42*, 1.
- (158) Campos, S. R. R.; Machuqueiro, M.; Baptista, A. M. *Journal of Physical Chemistry B* **2010**, *114*, 12692.
- (159) DeMarco, M. L.; Daggett, V. *Biochemistry* **2007**, *46*, 3045.
- (160) Langella, E.; Improta, R.; Barone, V. *Biophysical Journal* **2004**, *87*, 3623.
- (161) Khorvash, M.; Lamour, G.; Gsponer, J. *Biochemistry* **2011**, *50*, 10192.
- (162) Gu, W.; Wang, T. T.; Zhu, J.; Shi, Y. Y.; Liu, H. Y. *Biophysical Chemistry* **2003**, *104*, 79.
- (163) Chakroun, N.; Fornili, A.; Prigent, S.; Kleinjung, J.; Dreiss, C. A.;

Rezaei, H.; Fraternali, F. *Journal of Chemical Theory and Computation* **2013**, 9, 2455.

(164) Welker, E.; Raymond, L. D.; Scheraga, H. A.; Caughey, B. *Journal of Biological Chemistry* **2002**, 277, 33477.

(165) Kemp, D. S.; Allen, T. J.; Oslick, S. L.; Boyd, J. G. *Journal of the American Chemical Society* **1996**, 118, 4240.

(166) Chen, W.; van der Kamp, M. W.; Daggett, V. *Biochemistry* **2010**, 49, 9874.

(167) van der Kamp, M. W.; Daggett, V. *Biophysical Journal* **2010**, 99, 2289.

(168) Minor, D. L.; Kim, P. S. *Nature* **1994**, 367, 660.

(169) Gallo, M.; Paludi, D.; Cicero, D. O.; Chiovitti, K.; Millo, E.; Salis, A.; Damonte, G.; Corsaro, A.; Thellung, S.; Schettini, G.; Melino, S.; Florio, T.; Paci, M.; Aceto, A. *International Journal of Immunopathology and Pharmacology* **2005**, 18, 95.

(170) Calzolari, L.; Lysek, D. A.; Guntert, P.; von Schroetter, C.; Riek, R.; Zahn, R.; Wuthrich, K. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, 8340.

(171) DeMarco, M. L.; Daggett, V. *Comptes Rendus Biologies* **2005**, 328, 847.

(172) Day, R.; Bennion, B. J.; Ham, S.; Daggett, V. *J Mol Biol* **2002**, 322, 189.

(173) Wong, C.; Xiong, L. W.; Horiuchi, M.; Raymond, L.; Wehrly, K.; Chesebro, B.; Caughey, B. *Embo Journal* **2001**, 20, 377.

(174) Bocharova, O. V.; Makarava, N.; Breydo, L.; Anderson, M.; Salnikov, V. V.; Baskakov, I. V. *Journal of Biological Chemistry* **2006**, 281, 2373.

(175) El-Bastawissy, E.; Knaggs, M. H.; Gilbert, I. H. *Journal of Molecular Graphics & Modelling* **2001**, 20, 145.

- (176) Blinov, N.; Berjanskii, M.; Wishart, D. S.; Stepanova, M. *Biochemistry* **2009**, *48*, 1488.
- (177) Colacino, S.; Tiana, G.; Broglia, R. A.; Colombo, G. *Proteins-Structure Function and Bioinformatics* **2006**, *62*, 698.
- (178) Kitao, A.; Hirata, F.; Go, N. *Chemical Physics* **1991**, *158*, 447.
- (179) Gerum, C.; Silvers, R.; Wirmer-Bartoschek, J.; Schwalbe, H. *Angewandte Chemie-International Edition* **2009**, *48*, 9452.
- (180) Turk, E.; Teplow, D. B.; Hood, L. E.; Prusiner, S. B. *European Journal of Biochemistry* **1988**, *176*, 21.
- (181) Harris, D. A. *British Medical Bulletin* **2003**, *66*, 71.
- (182) Di Natale, G.; Turi, I.; Pappalardo, G.; Sovago, I.; Rizzarelli, E. *Chemistry-a European Journal* **2015**, *21*, 4071.
- (183) Prusiner, S. B. *Science* **1982**, *216*, 136.
- (184) Giachin, G.; Phuong Thao, M.; Thanh Hoa, T.; Salzano, G.; Benetti, F.; Migliorati, V.; Arcovito, A.; Della Longa, S.; Mancini, G.; D'Angelo, P.; Legname, G. *Scientific Reports* **2015**, *5*.
- (185) Hansmann, U. H. E.; Okamoto, Y. *Current Opinion in Structural Biology* **1999**, *9*, 177.
- (186) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96.
- (187) Zhang, J. P.; Sun, J.; Wu, C. Z. *Journal of Theoretical Biology* **2011**, *279*, 17.
- (188) Zanuy, D.; Ma, B. Y.; Nussinov, R. *Biophysical Journal* **2003**, *84*, 1884.
- (189) Abskharon, R. N. N.; Giachin, G.; Wohlkonig, A.; Soror, S. H.; Pardon, E.; Legname, G.; Steyaert, J. *Journal of the American Chemical Society* **2014**, *136*, 937.
- (190) Richardson, J. S.; Richardson, D. C. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 2754.