Mutable Mechanics in Biomaterials: a Study of Trichocyte Keratin and Nvjp-1 Protein

by

Chia-Ching Chou

B.S. Civil Engineering,
National Taiwan University, 2006
M.S. Civil Engineering,
National Taiwan University, 2008

Submitted to the Department of Civil and Environmental Engineering in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the
Massachusetts Institute of Technology

June 2015

Signature of Author: Signature redacted
Department of Civil and Environmental Engineering, 2015

Certified by: Signature redacted
Markus J. Buehler
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by: Signature redacted
Heidi M. Nepf
Donald and Martha Harleman Professor of Civil and Environmental Engineering
Chair, Departmental Committee for Graduate Students

Mutable Mechanics in Biomaterials: a Study of Trichocyte Keratin and Nvjp-1 Protein

by

Chia-Ching Chou

Submitted to the Department of Civil and Environmental Engineering on May 4, 2015, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Field of Structures and Materials in Civil and Environmental Engineering

ABSTRACT

Disulfide bonds and metal-coordination bonds serve as flexible connections on account of their variable bond strengthes in different chemical environments. Both bonds can act as an effective switch to stabilize or weaken a protein's geometry. These mechanisms are responsible for protein mutability, enabling distinct biological functions. Understanding the bonds' chemistry and mechanics is important for our knowledge of molecular, cellular and tissue level properties of biological materials that contain high ratios of disulfide bonds or metal-coordination bonds, such as hair and Nereis jaws.

In this thesis, we develop the first multiscale framework to study the hierarchical structure of trichocyte keratin proteins with disulfide bonds from a bottom-up approach. We build the first full atomistic structure of keratin k35/k85 heterodimer and upscale to the keratin macrofibril level. Using molecular dynamics simulations, we provide insights into the rupture mechanisms of disulfide bonds in keratin protein and illustrate the importance of the redox environment which governs the rupture mechanisms and locations. Disulfide bonds result in a higher strength and toughness of keratin proteins, but the system loses helical structures under loading, suggesting that disulfide bonds play a significant role in achieving the characteristic mechanical properties of trichocyte keratin. In keratin macrofibrils, disulfide crosslinks contribute to the initial modulus and enhance the robustness of macrofibrils by facilitating cooperative deformation of microfibrils at larger deformation.

We also study the mutability of Nvjp-1 protein with metal-coordination bonds in varied chemical environments from the single molecule level to larger length scales (~μm). Nvjp-1 forms a more compact structure in the presence of Zn ions with more stable intra-molecular metal coordination complexes at higher ion concentrations. As pH increases, deprotonation of histidine amino acids weakens the electrostatic repulsion, allowing a greater intramolecular interaction and a shift to a folded conformation. We find that pH also controls the stiffness and modulus of Nvjp-1. This study suggests that the metal-coordination crosslinks and pH induce significant Nvjp-1 aggregation and achieve the contraction of the Nvjp-1 stripe observed in experiments. The methodology illustrated here elucidates the trigger mechanism of material mutability for the design of biomaterials with varied properties through a bottom-up computational approach.

Thesis Supervisor: Markus J. Buehler

Title: Professor of Civil and Environmental Engineering

Author's Declaration

I am the author of all the work presented in this thesis. Research was conducted in the Department of Civil and Environmental Engineering at the Massachusetts Institute of Technology. Part of the work presented here has been published:

Chou, C.C., Lepore, E., Antonaci, P., Pugno, N. and Buehler, M. J. (2015). "Mechanics of trichocyte alpha-keratin fibers: Experiment, theory, and simulation," Journal of Materials Research **30**(01): 26-35.

Chou, C.C. and Buehler, M. J. (2013). "Molecular mechanics of disulfide bonded alpha-helical protein filaments," BioNanoScience, **3**.

Chou, C.C. and Buehler, M. J. (2012) "Structure and mechanical properties of human trichocyte keratin intermediate filament protein," Biomacromolecules, **13**(11).

Keten, S., Chou, C.C., van Duin, A. C.T. and Buehler, M. J. (2012). "Tunable Nanomechanics of Protein Disulfide Bonds in Redox Microenvironments," Journal of the Mechanical Behavior of Biomedical Materials, **5**,.

Chou, C.C. and Buehler, M. J. (2011). "Breaking out of the cage," Nature Chemistry, **3**(11), pp. 837-839.

Qin, Z., Chou, C., Kreplak, L. & Buehler, M. J. (2011). "Structural mechanics and functional properties of intermediate filaments from the atomistic to the cellular scales." Chapter 4 in Advances in Cell Mechanics, edited by Shaofan Li and Bohua Sun, published by Springer-Verlag

Acknowledgements

I take this opportunity to express my appreciation to many people who have helped, supported and encouraged me during this PhD journey. I would never have been able to finish my dissertation without the guidance of my committee members, help from friends, and support from my family and husband.

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Markus Buehler, for giving me this opportunities to work on this study and spending countless time to discuss with me and encourage me. His advice was essential to the completion of this dissertation. This thesis would not have been possible without his patience, support and insightful comments.

I would like to thank my thesis committee members, Professor Oral Buyukozturk, Professor Mark Bathe, Professor Niels Holten-Andersen, for spending time meeting with me and providing many valuable comments to make this thesis more complete. I would also like to thank many of my collaborators, Professor Keten, Professor Pugno, Dr. Dennis and Dr. Naik, for their help.

At MIT, I am so grateful to meet a group of friends and colleagues in LAMM. I would like to thank them for their friendship, never-ending support and encouragement in the past years. I really enjoy their company.

I would like to thank my parents, who have supported and encouraged me in all my academic endeavors.

This research was grants by the NSF, DOD-MURI, IRG and AFOSR. Their support is greatly appreciated. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number TG-MSS090007.

To my husband, for his support and encouragement.

Contents

List of figures

List of tables

1 Introduction

1.1 Background and objectives

Chemical bonds vary greatly in strength as exemplified by the covalent bonds in diamond compared to the extremely weak van der Waals interactions between layers of graphite in the two most common allotropes of carbon. These different bonds play a key role in regulating the properties of functionally diverse protein materials in nature. Strong covalent bonds form a permanent connection between atoms in the backbone of polypeptide chains. Weaker non-covalent bonds, such as hydrogen bonds, define a protein's three-dimensional structure. Intermediate between these strong and weak interactions are the disulfide bonds and metal-coordination bonds, which serve as flexible connections on account of their variable bond strengths in different chemical environments. Disulfide bonds weaken in the presence of reducing agents (approaching the strength of hydrogen bonds) and are strengthened in an oxidizing environment (approaching the strength of other strong covalent bonds). On the other hand, the strength of metal-coordination bonds changes as pH changes. Therefore, both bonds can act as an effective switch to stabilize or weaken a folded protein structure, or even to rearrange a protein's geometry altogether.

These mutable mechanisms allow a protein to provide distinct biological functions, a phenomenon called mutability, and are also a critical feature in molecular sensing and signalling. Understanding the chemistry and mechanics of these bonds is important to our knowledge of molecular, cellular and tissue level properties of biological materials that contain high ratios of disulfide bonds or metal-coordination bonds, such as hair and *Nereis* jaws. In this research, using a bottom-up approach, we study the fundamental mechanical properties of biological materials with disulfide bonds and metal-coordination bonds, and investigate the interplay of bond energy, chemical microenvironment and structure to achieve the desired mechanical properties at multiple length scales. The insights of this study could be applied to designing nanoscale

structures of peptide-based and other biomaterials consisting of different bond crosslinks to achieve desired mechanical properties. Diversity of mechanical properties can arise from the same building block with mutable bonds in different chemical environments, presenting a new paradigm in materials design.

1.1.1 Mutability of the disulfide bonded protein

Protein disulfide bonds are typically formed through the oxidization of thiol groups in two cysteine amino acids. Disulfide bonds play a fundamental role in the structure, flexibility and mechanical properties of biological materials (Brändén and Tooze 1999, Voet and Voet 2004, Cheek, Krishna et al. 2006, Fratzl and Weinkamer 2007, Buehler and Yung 2009, Garcia-Manyes, Liang et al. 2009). This is because disulfide bonds are stronger and have larger bond energy in oxidizing environments, while under the presence of a reducing agent, disulfide bonds are weakened and can be broken easily. Therefore, disulfide bonds provide mutability to the folded structures of proteins, enabling specific mechanical functions such as molecular sensing, switching and signalling in biology (Aslund and Beckwith 1999, Mayans, Wuerges et al. 2001, Hogg 2003).

Intermolecular disulfide bonds control the mechanical properties of proteins and polymers such as vulcanized rubber (Flory 1953), hair, feather, beak and wool (Parbhu, Bryson et al. 1999, Mayans, Wuerges et al. 2001, Mucke, Kreplak et al. 2004). The change in the strength of disulfide bonds is similarly important in the context of disease where a failure to form the correct disulfide bond or excessive strengthening under oxidative stress can lead to malfunction through protein misfolding or aggregation, as seen in inflammation, cardiovascular disease, amyotrophic lateral sclerosis or cataract formation (Wang, Xu et al. 2006, Buehler and Yung 2009, Grützner, Garcia-Manyes et al. 2009, Pande, Gillot et al. 2009). Overall, disulfide bond chemistry and mechanics are fundamental to our understanding of molecular, cellular and tissue level properties of biological materials in both physiological and disease conditions.

One of the disulfide bonded-rich proteins, keratin, is an intermediate filament (IF) protein which is a key component of hair, nail and skin in vertebrates. The human intermediate

filament database has shown that at least 54 of the 70 intermediate filament genes are keratins, making this a very prominent form of IFs that are widely found (Szeverenyi, Cassidy et al. 2008). According to their mechanical properties, keratin protein materials are divided into trichocyte keratin and epithelial keratin. Trichocyte keratin has a larger elastic modulus but low extensibility, and epithelial keratin has a much lower elastic modulus but larger extensibility. For different trichocyte keratin proteins, the elastic modulus ranges from approximately 1-4 GPa. For example, the modulus is 1.6-4.5 GPa (Oxenham 1989, Chou and Overfelt 2011) in wool, 2.0-3.7 GPa (Fudge and Gosline 2004, Guthold, Liu et al. 2007, Seshadri and Bhushan 2008) in human hair, 1.0-3.5 GPa (Chou and Overfelt 2011) in porcupine quill, and 0.4 GPa (Bertram and Gosline 1987, Chou and Overfelt 2011) in hoof. Epithelial keratin, on the other hand, is usually found in epithelial cells of skin and its elastic modulus is significantly lower, and typically about 3 MPa (Park and Baddiel 1972).

From an engineering perspective, trichocyte keratin is an extremely durable material that has been widely used in commercial products for over 2,000 years, e.g., wool in the textile industry. The marvelous durability of the material is exemplified by extremely old artifacts that exist to this day. Apparently, trichocyte keratin features a very long degradation time, perhaps due to its high resistance against mechanical loading and other environmental influences such as chemical degradation that break down many other biological tissues. It is thought that disulfide bonds play a crucial role in defining the stability of this protein and provide a critical feature in mechanical properties (Yu, Yu et al. 1993, Naito and Arai 1996, Parbhu, Bryson et al. 1999, Wang, Parry et al. 2000, Fudge and Gosline 2004, Strnad, Usachov et al. 2012).

Due the outstanding durable features of trichocyte keratin, several researchers have studied this class of proteins from the 1930s onwards (Speakman 1927, Astbury and Street 1932, Astbury and Woods 1934, Kreplak, Doucet et al. 2001, Kreplak, Doucet et al. 2004), including analysis of the stress-strain curve of wool, examination of the trichocyte keratin protein structure, and measurement of the modulus of trichocyte keratin IFs in oxidizing/reducing environments. As a result of experimental work, some deformation mechanism models were proposed to explain the mechanical behaviors of

keratin proteins (Bendit 1960, Chapman and Feughelman 1967, Wortmann and Zahn 1994, Hearle 2000, Feughelman 2002, Kreplak, Franbourg et al. 2002). However, due to the lack of a structural model of keratins with atomistic resolution, relatively little is known of the physical material concepts. In this thesis, we aim to develop an atomistic model of this protein to open the possibility of studying the mechanical properties of hair fibrils and other fibers from a bottom-up perspective. Moreover, we provide general insights into the effect of disulfide crosslinks on mechanical properties and present opportunities of synthesizing peptide materials with higher durability and resistibility against external loading.

1.1.2 Mutability of the metal coordination binding protein

We also study the Nvjp-1 protein, another mutable material, with metal-coordination bonds which occur when a species (ligand) donates its lone pair of electrons to a metal ion which is usually a transition metal, such as Fe^{2+}, Co^{2+}, Cu^{2+}, and Zn^{2+}. The metal-coordination bond is a flexible connection due to its variable bond strength in different chemical microenvironments (Lee, Scherer et al. 2006, Holten-Andersen, Harrington et al. 2011, Fullenkamp, He et al. 2013). The strength of a metal-coordination bond decreases in low pH conditions and the bond strength is increased with higher pH. Moreover, unlike most covalent bonds whose rupture is usually irreversible, metal-coordination bonds are able to reform after rupture. The amino acid, histidine, often forms metal-coordination bonds with surrounding metal ions such as Zn^{2+} in the Nvjp-1 protein, a key component of *Nereis* jaw (Lichtenegger, Schoberl et al. 2003) (Figure 1-1(A)). The jaws are composed of roughly 90% protein, halogens (~8%) and Zn ions (~2%). However, the hardness and stiffness of the jaws are comparable to the human dentin (~1-2 GPa of hardness and 10-20 GPa of stiffness (Cribb, Lin et al. 2010)). Nvjp-1 protein contains over 25 mol% histidine amino acid and it is believed that the formation of metal-coordination crosslinks between histidine-rich Nvjp-1 protein and Zn ions dominates the structural stability and superb mechanical properties of *Nereis* jaw. It has been shown that pH affects the aggregation and secondary structure of Nvjp-1 protein (Broomell, Chase et al. 2008), whereby an increase in pH causes the proteins to form more regular secondary structures such as alpha-helix and beta-sheet as shown in Figure

1-1(B). In this research, we use atomistic simulation to study the effect of pH and ions on the structural and mechanical properties of Nvjp-1 protein and to investigate the role of metal-coordination crosslinks in Nvjp-1 protein. The results provide fundamental insights into mutable mechanical properties of Nvjp-1 protein and other proteins with metal-coordination bonds from a bottom-up perspective, yielding insights into synthesizing peptide materials with mutable properties.

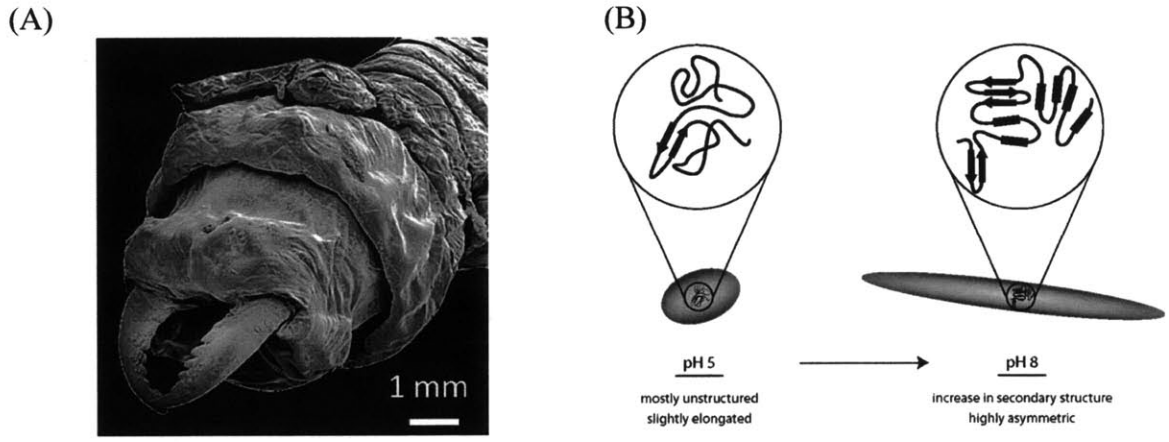

Figure 1-1 (A) *Nereis* jaw. (B) pH effect on secondary structure content of Nvjp-1 protein. Reprinted with permission from (Broomell, Chase et al. 2008). Copyright (2008) American Chemical Society

1.2 Research approach and area

The goal of this research is to study the fundamental mutable mechanics of biological materials with disulfide bonds and metal coordination bonds and to provide insights into the interplay of bond energy, chemical microenvironment and structure to achieve desired mechanical properties at multiple scales. The development of computational models for specific material systems and the analysis follow four steps:

I. **Mutable nanomechanics of protein disulfide bonds in redox microenvironments.**

We use a molecular simulation approach to show that the fracture strength of disulfide bonds is decreased under the presence of reducing agents, due to a loss of crosslink stability controlled by the chemical microenvironment. We provide a

mechanistic insight into the fracture mechanism of disulfide bonds in proteins and illustrate the importance of the redox microenvironment, where factors such as accessibility, mechanical strain and local redox potential govern the dominating rupture mechanism and location. The method used here provides a general computational protocol for studying mechanochemical fracture of large-scale protein materials concurrently with experimental efforts.

II. **Molecular dynamics simulations of disulfide bonded-rich proteins.**
Keratin, a type of intermediate filament protein and a key component of hair, nail and skin in vertebrates, is studied as a model system. Keratin plays the role of maintaining the stability of the cell and contributes to the stability of adjacent cells. Fundamental work focuses on building a bottom-up molecular based model of keratin, starting at the atomistic scale with the complete amino acid sequence. The geometric and mechanical properties of keratin proteins are investigated using full atomistic simulation and validated against experimental data from AFM, SEM and x-ray diffraction.

III. **Coarse-grained modeling of disulfide bond-rich proteins.**
A coarse-grained multi-scale model is developed to understand the outstanding mechanical properties at larger scales of filaments and filament networks in cells and biomaterials. In this work, we study how disulfide bonds play a significant role in achieving the characteristic mechanical properties of varied protein materials. A particular interest is on the interplay of weak and strong bonds and the role of hierarchical structures, and the elucidation of deformation mechanisms at different scales.

IV. **Investigation of Nvjp-1 protein structure and mechanical properties in various microenvironments.**
Nvjp-1 protein, the key component of Nereis jaw, is studied as a model system. Fundamental work focuses on building an atomistic model of Nvjp-1 based on the full-length amino acid sequence. The geometric and mechanical properties of Nvjp-1 proteins in various environments are investigated by full atomistic simulation and validated against experimental data.

1.3 Organization of the thesis

The organization of the thesis is as follows: Chapter 2 provides details on the methodologies used in this study, including conventional molecular dynamics simulations, replica-exchange molecular dynamics, and pH molecular dynamics. In Chapter 3, we describe the different fracture mechanisms of disulfide bond in redox environments. Chapter 4 focuses on atomistic simulations of the keratin heterodimer based on the full amino acids sequence of keratin proteins (k35/k85) and investigates the geometric and mechanical properties of resulting protein fibers. We show the structural analysis of the equilibrated geometries of both heterodimer and tetramer of keratin k35/k85 protein and compare with experimental data. Then, we present the effect of disulfide crosslinks on mechanical properties of keratin protein under mechanical testing. We scale up to the keratin macrofibril level in Chapter 5. The results on deformation mechanisms of keratin macrofibrils with disulfide bonds are discussed. In Chapter 6, we focus on full atomistic simulations of nanostructure of Nvjp-1 protein in varied ion concentrations. In Chapter 7, we study the effect of pH on the structure and mechanical properties of Nvjp-1 proteins. Chapter 8 summarizes the major findings of the thesis and discusses possible future research directions.

2 Methodology

2.1 Atomistic modeling

2.1.1 Classical molecular dynamics

Molecular Dynamics (MD) is a numerical method to study the behavior of a set of atoms or molecules based on the classical equations of motion (Newton's Equations). In MD, the interatomic force fields which are obtained from the empirical approximation or quantum mechanics calculation are used to describe the interaction between atoms. MD counts the position $\mathbf{r}_i$, velocity $\mathbf{v}_i$ and acceleration $\mathbf{a}_i$ of each atom in a system at every short step of a time period under a specified statistical ensemble, and then the properties of system are calculated based on the above information of time. Figure 2-1 shows a flow chart of performing a MD simulation.

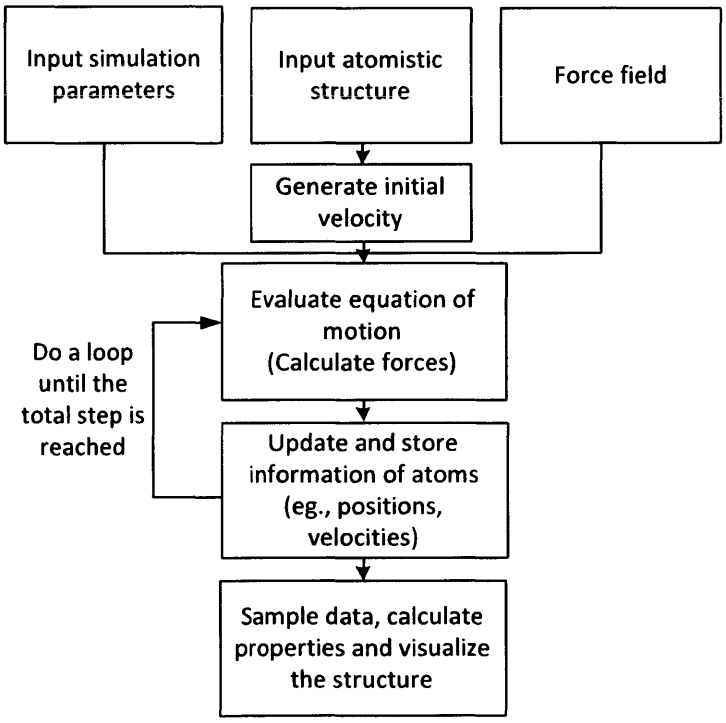

Figure 2-1 Overview of the procedure for molecular dynamics simulation (Buehler 2008).

In MD, the movement of each atom follows Newton's second law of motion, in which the motion of an atom i is caused by a force $\mathbf{f}_i$:

$$\mathbf{f}_i = m\ddot{\boldsymbol{r}}_i \tag{2-1}$$

where m is the mass of the atom and the acceleration $\ddot{\boldsymbol{r}}_i$ is given by:

$$\ddot{\boldsymbol{r}}_i = \frac{d^2\mathbf{r}_i}{dt^2} \tag{2-2}$$

In the classical molecular dynamics, the total energy consists of the kinetic energy and potential energy of the N atoms:

$$E = K + U \tag{2-3}$$

where K is the kinetic energy and U is the potential energy:

$$K = \frac{1}{2}\sum_i m\dot{\mathbf{r}}_i^2 \tag{2-4}$$

$$U = U(\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_N) \tag{2-5}$$

If we take the time derivative of the total energy for a system with conservative energy, we have:

$$\sum_i m\dot{\mathbf{r}}_i \cdot \ddot{\boldsymbol{r}}_i + \sum_i \frac{\partial U}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i = 0 \tag{2-6}$$

Since the velocities are all independent of one another, Equation (2-6) can be satisfied only if, for each atom i:

$$\frac{\partial U}{\partial \mathbf{r}_i} = -m\ddot{\mathbf{r}}_i = -\mathbf{f}_i \tag{2-7}$$

Therefore the force is calculated as the negative gradient of a potential function $U(\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_N)$. With a well-defined potential energy, the trajectory of each atom can be determined by applying a finite difference method to evaluate equation of motion.

2.1.2 Force fields for biomaterials

The interatomic force fields are used to describe the interaction between atoms in molecular dynamics simulations for biological materials. For biomaterials, a wide range of force fields, for example, AMBER (Pearlman, Case et al. 1995), CHARMM force field, OPLS force field (Jorgensen, Maxwell et al. 1996, MacKerell, Bashford et al. 1998), and GROMACS force field, have been developed.

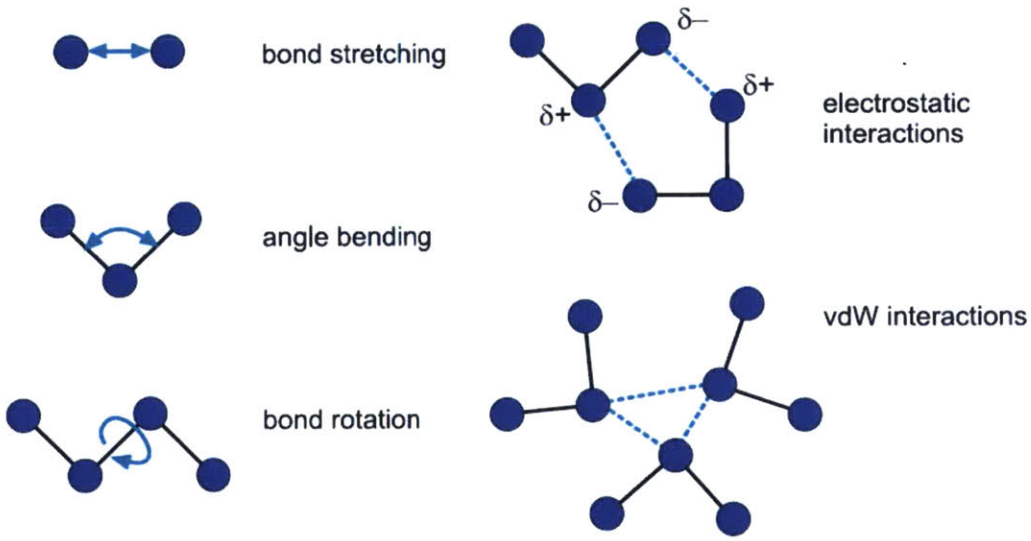

Figure 2-2 Illustrations of the contributions of the different terms in the potential expressions given in Eq. (2-8). Figure reprinted from (Buehler 2010).

For a conventional force field for organic materials, such as CHARMM force field, the potential energy includes the bonded and non-bonded interactions among atoms (Figure

2-2). The bonded interatomic force field describes the bond, angle and dihedral interaction between bonded atoms. The non-bonded force field includes the Coulomb interaction (electrostatic force) among charge particles and van der Waals interaction. The potential energy, U, is expressed as

$$U = U_{\text{bond}} + U_{\text{angle}} + U_{\text{torsion}} + U_{\text{coulomb}} + U_{\text{vdW}} + \cdots \tag{2-8}$$

where U_{bond} is the bonded interaction describing the stretching behavior, U_{angle} is the bending energy, U_{torsion} is the torsion energy, U_{coulomb} and U_{vdW}, are the coulomb energies and the van der Waals interactions between atoms in the system. U_{bond} is usually written as

$$U_{\text{bond}} = \sum_{\text{bonds}} K_i(b_i - b_0)^2 \tag{2-9}$$

where K_i is the force constant of a given bond, b_0 is the equilibrium bond length, and b_i is the bond length between the two atoms. b_0 and K_i are the parameters to describe the type of the bond.

The harmonic angle term is given by

$$U_{\text{angle}} = \sum_{\text{angles}} K_\theta(\theta_i - \theta_0)^2 \tag{2-10}$$

where K_θ is bending stiffness and the θ is the angle between three bonded beads relative to the equilibrium angle, θ_0. And the cosine form of the torsion potential is expressed as

$$U_{\text{torsion}} = \sum_{\text{torsions}} K_\chi(1 + \cos(n\chi - \delta)) \tag{2-11}$$

where K_χ is the force constant, δ is the equilibrium angle where the potential passes through its minimum value, and χ is the torsion angle. The parameters for each type of angle and torsion interaction between bonded atoms are well-defined in the force field. For non-bonded interaction, the most common potential to model the van der Waals interactions is known as the Lennard-Jones potential:

$$\phi_{LJ} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{2-12}$$

where ε is the energy at equilibrium and σ is the distance as $\phi_{LJ} = 0$. The Coulomb interaction between atoms is expressed by using Coulombs law:

$$U_{\text{coulomb}} = \sum_{\text{coulomb}} \frac{q_i q_j}{4\pi\epsilon_{ij} r_{ij}} \tag{2-13}$$

where ϵ_{ij} is a constant, q_i and q_j are charges of atoms, and r_{ij} is the distance between atoms.

2.2 Trichocyte keratin k35/k85 and Nvjp-1 protein sequences

The real sequences of the human trichocyte keratin k35/k85 and the Nvjp-1 protein of the Nereis jaw are used to generate the keratin and Nvjp-1 molecules. The sequences are adapted from Human Intermediate Filament Database (http://www.interfil.org/) and (Broomell, Chase et al. 2008).

The sequence of the human trichocyte keratin k35 is:

MASKCLKAGFSSGSLKSPGGASGGSTRVSAMYSSSSCKLPSLSPVARSFSACSVG
LGRSSYRATSCLPALCLPAGGFATSYSGGGGWFGEGILTGNEKETMQSLNDRLA
GYLEKVRQLEQENASLESRIREWCEQQVPYMCPDYQSYFRTIEELQKKTLCSKAE
NARLVVEIDNAKLAADDFRTKYETEVSLRQLVESDINGLRRILDDLTLCKSDLEA
QVESLKEELLCLKKNHEEEVNSLRCQLGDRLNVEVDAAPPVDLNRVLEEMRCQ

YETLVENNRRDAEDWLDTQSEELNQQVVSSSEQLQSCQAEIIELRRTVNALEIEL
QAQHSMRDALESTLAETEARYSSQLAQMQCMITNVEAQLAEIRADLERQNQEY
QVLLDVRARLECEINTYRGLLESEDSKLPCNPCAPDYSPSKSCLPCLPAASCGPSA
ARTNCSPRPICVPCPGGRF

The sequence of the human trichocyte keratin k85 is:

MSCRSYRISSGCGVTRNFSSCSAVAPKTGNRCCISAAPYRGVSCYRGLTGFGSRS
LCNLGSCGPRIAVGGFRAGSCGRSFGYRSGGVCGPSPPCITTVSVNESLLTPLNLEI
DPNAQCVKQEEKEQIKSLNSRFAAFIDKVRFLEQQNKLLETKWQFYQNQRCCES
NLEPLFSGYIETLRREAECVEADSGRLASELNHVQEVLEGYKKKYEEEVALRATA
ENEFVVLKKDVDCAYLRKSDLEANVEALVEESSFLRRLYEEEIRVLQAHISDTSVI
VKMDNSRDLNMDCIIAEIKAQYDDVASRSRAEAESWYRSKCEEMKATVIRHGET
LRRTKEEINELNRMIQRLTAEIENAKCQRAKLEAAVAEAEQQGEAALSDARCKL
AELEGALQKAKQDMACLLKEYQEVMNSKLGLDIEIATYRRLLEGEEHRLCEGVG
SVNVCVSSSRGGVSCGGLSYSTTPGRQITSGPSAIGGSITVVAPDSCAPCQPRSSSF
SCGSSRSVRFA

The sequence of the Nvjp-1 protein is:

HNDGYGHDDHHGHGHGGYGGHGHGDYGGHGHGGYGGHGHGHGHGHGHFDD
HPFYTIPAFGHGYGHGHGGHGHGYGGHDGYGGHGGYGGHGGYGDHGHGGHG
YGSHGGHGQDYGGDYGGHGHGGHHHGGHDHDDFGHDFGHHGGDHGHHGGG
HHGHHDGYGHDQGHGHGHGDYDYSHGNQHDNGHREHYGGHQAAGHHGHGE
SGSHYGGQGGGGGSHHSGGHDGSHNHGKFQTQYSYRGNDKYGGHNQYQGHG
AYNEYSKSKGTGKYVAHGTEEGHHQLDGYEKDNHQGKYEGHGEHHSHQHGD
GHHKRKEHGDHAGISGFHGQGHVQGNHALHGHHGGGHGHHGGEHGHSARGH
HGGGGYGGGGHGHRGHHGGHQGHHGHH

2.3 Steered molecular dynamics

In this study, we apply steered molecular dynamics (SMD) which mimics the AFM experiments to probe the mechanics of proteins in molecular dynamics simulations. The

basic concept is to pull the center of mass of a collection of chosen atoms via a harmonic spring along the pulling direction, while keeping the center of mass of another group of atoms fixed. The force applied on the pulled atoms can be computed by

$$F(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t) = k[vt - (\mathbf{r}(t) - \mathbf{r}_0) \cdot \mathbf{n}] \tag{2-14}$$

in which $\mathbf{r}(t)$ is the position of the restrained atoms at time t, $\mathbf{r}_0$ is the original coordinates of the restrained atom, v is the pulling velocity and $\mathbf{n}$ is the pulling direction. The force-displacement curve can be obtained by monitoring the time-averaged applied force (2-14) and the position of the pulled atom over the simulation time.

2.4 Replica-exchange molecular dynamics

The basic idea of the replica exchange molecular dynamics (REMD) method is to perform independent conventional MD simulations at different temperature (replicas) at the same time (Sugita and Okamoto 1999) (Figure 2-3 (A)). After a certain period of time, the neighboring replicas are attempted to exchange their temperature. The probability of successful exchange is determined by the potential energy and temperature difference between replicas based on the Metropolis criterion:

$$P(x \to x') = \begin{cases} 1, & \Delta < 0 \\ \exp\left[-\left(\frac{1}{k_b T_1} - \frac{1}{k_b T_2}\right)(E_2 - E_1)\right], & \Delta \geq 0 \end{cases} \tag{2-15}$$

in which $\Delta = \left(\frac{1}{k_b T_1} - \frac{1}{k_b T_2}\right)(E_2 - E_1)$, E_1 and E_2 are the potential energies at T_1 and T_2. Figure 2-3 (B) illustrates the overlap of potential energy between replicas. Note that in the REMD, the replica at lower temperature has the possibility to exchange to a higher temperature thus has the possibility to overcome high energy barrier. Similarly, the replica at high temperature has the possibility to reduce its temperature to low temperature to explore different conformations at low temperature. Therefore, the REMD enhances the sampling of different conformations in the conventional MD simulations.

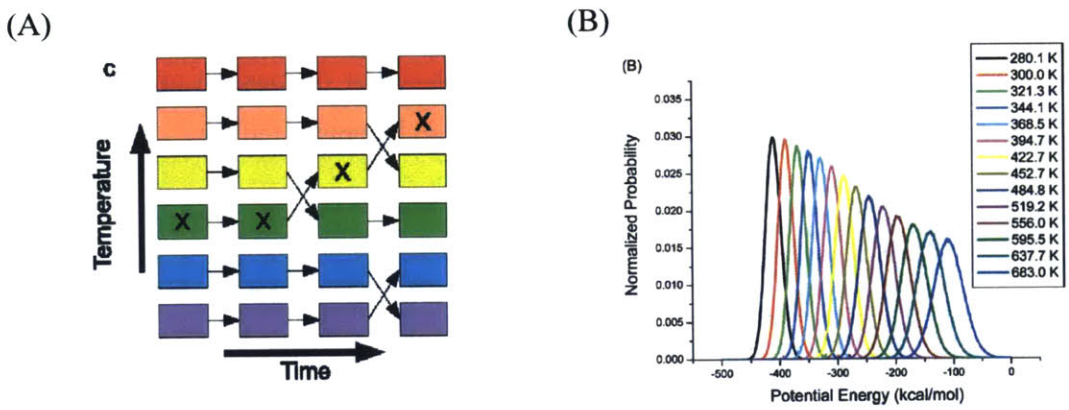

Figure 2-3 (A) Illustration of REMD. Reprinted figure with permission from van der Spoel, D. and M. M. Seibert (2006). "Protein folding kinetics and thermodynamics from atomistic simulations." Physical Review Letters 96(23). Copyright (2006) by the American Physical Society. (B) The overlap of potential energy between replicas. Reprinted with permission from Sindhikara, D., Y. L. Meng and A. E. Roitberg (2008). "Exchange frequency in replica exchange molecular dynamics." Journal of Chemical Physics 128(2). Copyright (2008), AIP Publishing LLC.

2.5 Constant pH molecular dynamics

Solution pH plays a critical role in controlling chemical and biological processes. The stability and function of proteins are highly dependent on the environmental pH. However, it is not possible to understand the mechanism of these pH-dependent biological processes, including ATP synthesis (Yoshida, Muneyuki et al. 2001), enzyme catalysis (Benkovic and Hammes-Schiffer 2003), and protein folding (Khandogin and Brooks 2007), using conventional MD simulations which are performed at preset and fixed protonation states. Many effort have been reported to develop constant pH MD algorithms to investigate the pH-dependent phenomena. The algorithms allow us to determinate the protonation states of titratable molecules during the MD simulations at a specific pH. In this thesis, we apply the constant continuous pH MD method to study the pH-dependent protein mutability. The basic idea of the continuous pH MD (CpHMD) method (Lee, Salsbury et al. 2004, Khandogin and Brooks 2005, Khandogin and Brooks 2006) is to introduce a virtual coordinate, titration coordinate, to the total Hamiltonian to

dictate the protonation states of titratable sites in the system. The total Hamiltonian can be written as,

$$H(\{\boldsymbol{r}_i\},\{\boldsymbol{\theta}_j\}) = \sum_{1}^{N_{atom}} \frac{m_i}{2}\dot{\boldsymbol{r}}_i^2 + U^{\text{int}}(\{\boldsymbol{r}_i\}) + U^{\text{hybrid}}(\{\boldsymbol{r}_i\},\{\boldsymbol{\theta}_j\}) + \sum_{1}^{N_{titr}} \frac{m_j}{2}\dot{\boldsymbol{\theta}}_j^2 + U^*(\{\boldsymbol{\theta}_j\}) \tag{2-16}$$

where the $i = 1, N_{atom}$ is the index for atomic spatial coordinates, and $j = 1, N_{titr}$ is the index for a continuous variable, θ, which is connected to titration coordinate, λ, through a sine function as shown in the following,

$$\lambda_j = \sin^2 \theta_j \tag{2-17}$$

where $\lambda_j = 0$ corresponds to the protonated state and $\lambda_j = 1$ corresponds to the deprotonated state.

In the Equation (2-16), $U^{\text{int}}(\{\boldsymbol{r}_i\})$ is the configurational energy of the system independent to protonation states, $U^{\text{hybrid}}(\{\boldsymbol{r}_i\},\{\boldsymbol{\theta}_j\})$ is the energy coupling the configurational and protonation degrees of freedom, and $U^*(\{\boldsymbol{\theta}_j\})$ represents the biasing energy along the titration coordinates. In this study, a hybrid solvent system is applied in the CpHMD (Wallace and Shen 2011), which takes advantage of the efficiency of the Generalized-Born model to derive the force on titration coordinates and the accuracy of explicit solvent to compute the conformational dynamics. $U^{\text{hybrid}}(\{\boldsymbol{r}_i\},\{\boldsymbol{\theta}_j\})$ is written as a sum of the van der Waals, Coulombic, and GB solvation energy shown as below.

$$U^{\text{hybrid}}(\{\boldsymbol{r}_i\},\{\boldsymbol{\theta}_j\}) = U^{\text{vdW}}(\{\boldsymbol{r}_i\},\{\boldsymbol{\theta}_j\}) + U^{\text{Coul}}(\{\boldsymbol{r}_i\},\{\boldsymbol{\theta}_j\}) + U^{\text{GB}}(\{\boldsymbol{r}_i\},\{\boldsymbol{\theta}_j\}) \tag{2-18}$$

For the biasing potential, $U^*(\{\boldsymbol{\theta}_j\})$, it is written as

$$U^{*}(\boldsymbol{\theta}_i) = \sum_i -U^{\mathrm{mod}}(\boldsymbol{\theta}_i) + U^{\mathrm{pH}}(\boldsymbol{\theta}_i) + U^{\mathrm{barr}}(\boldsymbol{\theta}_i) \tag{2-19}$$

where $U^{\mathrm{mod}}(\boldsymbol{\theta}_i)$ is the potential of mean force along the titration coordinate for the model compound, $U^{\mathrm{pH}}(\boldsymbol{\theta}_i)$ is the pH-dependent protonation free energy, and $U^{\mathrm{barr}}(\boldsymbol{\theta}_i)$ is a Harmonic potential for suppressing the population of mixed states where $0.1 < \lambda_j < 0.9$:

$$U^{\mathrm{mod}}(\boldsymbol{\theta}_i) = \boldsymbol{A}_i(\boldsymbol{\lambda}_i - \boldsymbol{B}_i)^2 \tag{2-20}$$

$$U^{\mathrm{pH}}(\boldsymbol{\theta}_i) = 2.3k_B T\lambda_i(\boldsymbol{\theta}_i)(\mathrm{pK_a}(i) - \mathrm{pH}) \tag{2-21}$$

$$U^{\mathrm{barr}}(\boldsymbol{\theta}_i) = 4\beta_i\left(\lambda_i(\boldsymbol{\theta}_i) - \frac{1}{2}\right)^2 \tag{2-22}$$

3 Tunable nanomechanics of protein disulfide bonds in redox microenvironments

(This chapter is adapted from ((Keten, Chou et al. 2012)))

Disulfide bonds are important chemical crosslinks that control the elasticity of fibrous protein materials such as hair, feather and wool. Here we present a novel computational approach using the first principles based ReaxFF reactive force field and apply this approach to show the various fracture strength of disulfide bonds under the presence of oxidizing/reducing agents We provide insight into the rupture mechanism of protein disulfide bonds and illustrate the importance of the redox environment, where factors such as accessibility, mechanical strain and local redox potential govern the dominating rupture mechanism and location. The method used here provides a general computational protocol for studying mechanochemical fracture of large-scale protein materials concurrently with experimental efforts.

3.1 Introduction

Disulfide bond (S-S), a weak covalent bond, can be formed reversibly between sulphur side-chain atoms in two cysteine (abbreviation CYS) amino acids in biology. Disulfide bond plays a fundamental role on structure, flexibility and mechanical properties of biological materials (Brändén and Tooze 1999, Voet and Voet 2004, Cheek, Krishna et al. 2006, Fratzl and Weinkamer 2007, Buehler and Yung 2009, Garcia-Manyes, Liang et al. 2009). This is because disulfide bonds are stronger and have larger bond energy in oxidizing environments, and under the presence of a reducing agent, disulfide bonds are weaken and can be broken easily. Thereby, disulfide bonds provide tunable stability to folded structures of proteins, enabling specific mechanical functions such as molecular

sensing, switching and signaling in biology (Aslund and Beckwith 1999, Mayans, Wuerges et al. 2001, Hogg 2003).

Facilitated by recent advancements in experimental techniques, pioneering studies of thiol-disulfide exchange reactions under external force application and presence of reducing agents, have demonstrated a rate-dependent bond rupture mechanism and the kinetic of disulfide bond isomerization (Bell 1978, Wiita, Ainavarapu et al. 2006, Alegre-Cebollada, Kosuri et al. 2011, Chou and Buehler 2011). However, it is difficult to derive an explanation of reaction path and reaction mechanism from experimental data only, pointing out a need for further investigation. One approach to this challenging problem is to perform atomistic simulation under the experimental condition and external loading. A major limitation to computational studies of disulfide bonds is that that the rupture of covalent bonds cannot be simulated via the classical force fields such as DREIDING, CHARMM or AMBER (MacKerell, Bashford et al. 1998, Ponder, Case et al. 2003), since covalent bonds are typically treated as unbreakable harmonic springs. Previous computational studies on disulfide bond properties have been limited to small systems with often only tens to hundreds of atoms, where Density Functional Theory (DFT) or Car-Parrinello Molecular Dynamics (CPMD) and QM/MM methods have been employed to investigate potential reaction pathways (Fernandes and Ramos 2004, Hofbauer and Frank 2010, Li and Grater 2010). Studying disulfide bond dynamics beyond the single molecule level remains a computational challenge, due to the prohibitive cost of quantum calculations. As a result, efficient interatomic force fields are in demand to extend the length and time-scale of simulations involving chemical reactions. Here we propose a novel approach in addressing this challenge for proteins using molecular dynamics simulations with the ReaxFF reactive force field (van Duin, Dasgupta et al. 2001). The key objective of this paper is to answer a fundamental question: What is reaction mechanism of disulfide bond in varied chemical environments under external loading?

3.2 Simulation setup

3.2.1 ReaxFF reactive force field and simulation model

ReaxFF is derived from first-principles calculations and is capable of modelling chemical reactions (including transition states during reactions, changes in bond order, charge equilibration) while retaining computational efficiency, which has been previously demonstrated for a variety of reactive systems including oxidization of hydrocarbons, silicon fracture and catalysis of carbon nanotube formation (Nielson, van Duin et al. 2005, Buehler, van Duin et al. 2006, Chenoweth, van Duin et al. 2008). We consider a simple model system consisting of two cysteine amino acids covalently linked by a single disulfide bond to gain fundamental insight into the nanomechanics of disulfide bonds in proteins under varied chemical and physical conditions. Figure 3-1(A&B) shows the general molecular setup simulation protocol. As depicted in Figure 3-1(C), reactions are explored by using a variable biasing potential that tethers the hydrogen atoms to the sulfur atoms to enhance the sampling of reactions (Laioa and Parrinello 2002, Bonomi, Branduardi et al. 2009, Bonomi and Parrinello 2010). All simulations are carried out in a periodic box of water with explicit solvent, as shown in Figure 3-1(D) in a system containing about 1,000 fully reactive protein and water solvent atoms, to capture solvent effects (hydrophobicity, viscosity, screening, possible reactions etc.) accurately.

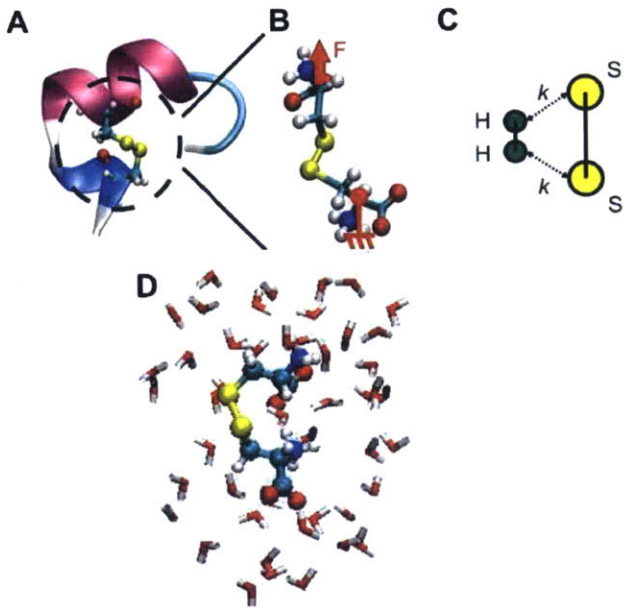

Figure 3-1 Schematic views illustrating the simulation setup. The disulfide bond model is taken from a small protein (Protein Data Bank identification code 1AKG) shown in (A). Steered Molecular Dynamics simulations are carried out, where one of the amino acids is fixed and the other is pulled at the alpha-carbon atom as depicted in (B). (C) shows the method for inducing reactions. To induce a reducing potential, hydrogen atoms (in the schematic shown as a hydrogen molecule) are tethered towards the S-S bond with a harmonic bias potential, thereby sampling the reaction at a shorter time scale accessible to Molecular Dynamics. (D) shows the overall system solvated in an explicit water box with periodic boundary conditions. Reprinted from Keten, S., C. C. Chou, A. C. T. van Duin and M. J. Buehler. "Tunable nanomechanics of protein disulfide bonds in redox microenvironments." Journal of the Mechanical Behavior of Biomedical Materials 5(1): 32-40, Copyright 2012, with permission from Elsevier.

3.2.2 Molecular Dynamics simulations setup

Molecular Dynamics simulations are carried out using the LAMMPS software (Plimpton 1995) with ReaxFF (van Duin, Dasgupta et al. 2001) and Steered Molecular Dynamics (Israelewitz, Gao et al. 2001) packages. Simulations are carried out for a simple protein system consisting of two cysteine amino acids linked via a disulfide bond. The initial

atomic coordinates of the model are obtained from the alpha-connotoxin protein (Protein Data Bank identification code 1AKG) in the Protein Data Bank (Bernstein, Koetzle et al. 1977). The system is solvated within a box of explicit reactive water using Visual Molecular Dynamics (Humphrey 1996). Each simulation is carried out in explicit solvent within a periodic simulation box (approximately containing 1,000 atoms), using the recently developed ReaxFF parameter set for proteins (Rahaman, van Duin et al. 2010). A Steered Molecular Dynamics approach is used to provide external mechanical perturbation to system, to induce the breaking of the peptide chain. The Steered Molecular Dynamics spring constant is set as 10 kcal/mol/Å^2 and the pulling velocity is 0.05 Å/ps. All visualizations are done using the VMD. For the simulations with bias potentials, reactions are driven varying the equilibrium length (sliding restraints) or force constant of the springs systematically over several simulations.

3.2.3 Research design

We study several cases designed to assess the rupture mechanisms of disulfide bonds under combined mechanical forces and reducing agents and using different harmonic bias potentials to sample varied time-scales. The bias potentials used here are implemented as harmonic functions defined as where is the distance between the pair of atoms, and denotes the target distance between the pair of atoms. The parameter is a constant that defines the strength of the bias potential (referred to as "bias potential constant"). The goal of these simulations is to quantify the mechanical rupture strength and rupture point of the cysteine molecule in conditions that mimic a variety of redox chemical environments. The first case involves stretching simulations of the cysteine molecule in the close presence of molecular hydrogen (reducing agent) and molecular oxygen (oxidizing agent). The O_2 or H_2 molecule is biased to the sulfur atoms in the protein with harmonic bias potentials that have gradually increasing bias potential constants. The purpose of this investigation is to probe the reduction of the disulfide bond with hydrogen and oxygen molecules, and show that in the false test of oxygen, the reduction of the S-S bond cannot occur regardless of the biasing function. Our simulations employ an *NVT* ensemble with a temperature set at 300 K, with an integration time step of 0.25 femtoseconds. The simulations with typical simulation time-scales are hundreds of

picoseconds to a few nanoseconds, depending on the bond rupture time-scale under mechanical strain and biasing potentials.

3.2.4 Mechanical strength analysis

We study the maximum force of disulfide bond breaking under mechanical force applied with hydrogen and oxygen molecules, respectively. To sample the reaction at longer time-scales, two bias potentials are applied between pairs of each of the sulfur atoms (of the disulfide bond) and corresponding hydrogen/oxygen atoms in the system. To achieve this bias potentials are added between S-O and S-H for the two cases investigated. The bias potential is implemented as described above. In order to assess the mechanical force needed to break disulfide bond and to examine associated cleavage mechanisms under H_2 or O_2 presence, we vary effect of the bias potential and compute the maximum force of the bond breaking with H_2 and O_2. We systematically vary the bias potential constant k from 1 to 50 kcal/mol/Å^2 while keeping all other parameters constant. The alpha-carbon atoms at the cysteine dimer are pulled on by using Steered Molecular Dynamics (Isralewitz, Gao et al. 2001) with a spring constant of 10 kcal/mol/Å^2 and the pulling force F is recorded versus the position. All simulations are carried out at a pulling velocity of 0.05 Å/ps.

3.2.5 Metadynamics analysis

We use the metadynamics method to investigate the energy barrier of our system along collective variables (CVs). The metadynamics algorithm is implemented in a portable plug-in, the PLUMED package (Bonomi, Branduardi et al. 2009), which can be included in the LAMMPS software. During the simulation, an external bias potential, $V(s,t)$, is developed as a sum of Gaussian functions centered on visited configurations $s^{(0)}$ in the collective variables space, s, where

$$V(s,t) = \sum_{\substack{t'=\tau,2\tau,3\tau\ldots \\ t'<t}} W \exp\left(-\sum_{i=1}^{d} \frac{\left(s_i - s_i^{(0)}(t')\right)^2}{2\sigma_i^2} \right), \tag{3-1}$$

in which s_i is the ith CV, σ_i is the Gaussian width corresponding to the i-th CV, W is the Gaussian height, d is the number of collective variables and τ is the deposit frequency of the Gaussian. Eventually, the free energy of the system is obtained by

$$V(s,t) = -F(s,t), \tag{3-2}$$

To investigate the effect of the bias potential to the energy barrier of cysteine with different agents, we perform metadynamics calculations using the S–S distance of the two cysteine residues as the only CV with different the spring constants of bias potential for oxygen and hydrogen molecules. All calculations are performed with a Gaussian height, W = 0.1 kcal/mol, width of 0.35 Å and a simulation time of 375 ps.

3.3 Results and Discussion

We start with an analysis of molecular hydrogen and oxygen attacking the disulfide bond, to assess bond rupture strengths and locations in representative oxidizing and reducing conditions. Figure 3-2(A) depicts snapshots from the simulation of disulfide rupture under reducing conditions, illustrating the reduction mechanism of the disulfide bond in the presence of a hydrogen molecule. Mechanical stretching of the system leads to exposure and straining of the disulfide bond. Before the reaction occurs, the bond is slightly elongated due to the presence of hydrogen atoms, which weakens the S-S bond. This is followed by the reduction of the sulfur atoms and fracture of the protein at the S-S bond. Next we carry out an identical simulation setup, but this time we replace the hydrogen molecule with an oxygen molecule (Figure 3-2(B)). In this case, regardless of the proximity of the oxygen molecule to the S-S bond, rupture of the S-S bond does not occur, which is in direct contrast to the hydrogen molecule case presented previously (Figure 3-2(A)). Indeed, this is expected from basic redox chemistry, since further oxidization of a disulfide bond via a reaction with oxygen is unlikely. An interesting observation is that in this scenario, the large mechanical forces induce breaking of the adjacent C-S bond, rather than the S-S bond, a result that we also observe consistently in pure mechanical stretching of the dimer in solvent. These findings agree with previous

theoretical predictions based on quantum mechanical methods that have also suggested the comparable bond energy and rupture likelihood of the C-S bond under external forces (Hofbauer and Frank 2010).

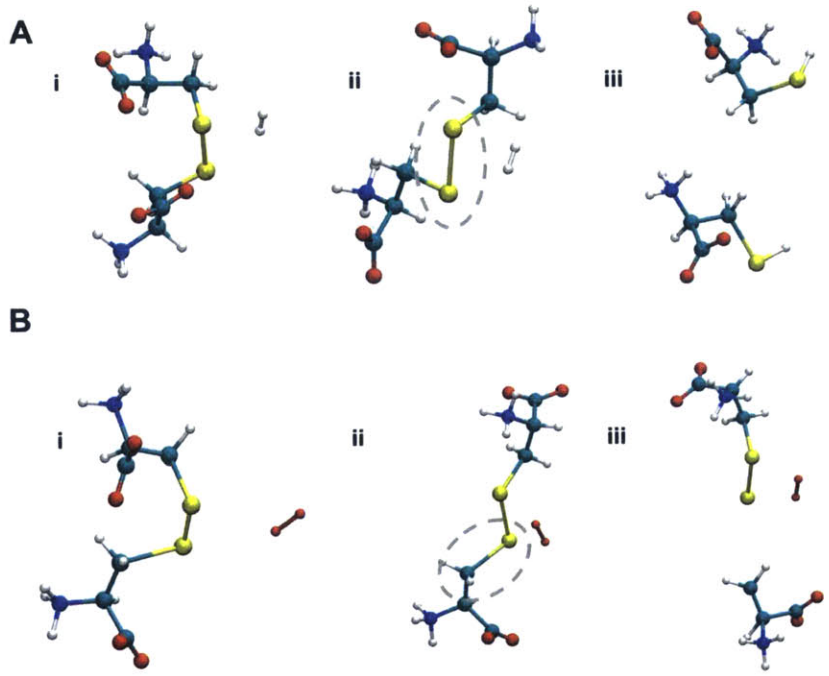

Figure 3-2 Reaction mechanisms with molecular hydrogen and oxygen. Rupture mechanism of the disulfide bond depends on the local redox environment, mimicked here by the presence of a tethered molecular oxygen (oxidizing agent) or molecular hydrogen (reducing agent) in the vicinity of the disulfide bond. Dissociation of the hydrogen atom occurs simultaneously with the rupture of the S-S bond as shown in panel A (i-iii), forming stable reduced monomers of cysteine. We carry out an identical simulation, this time replacing hydrogen with an oxygen molecule. Since oxygen is an oxidizing agent, a reaction with the disulfide bond does not take place. Instead, the adjacent C-S bond ruptures at very high mechanical forces as shown in panel B (i-iii). For clarity, surrounding water atoms are not shown. Reprinted from Keten, S., C. C. Chou, A. C. T. van Duin and M. J. Buehler. "Tunable nanomechanics of protein disulfide bonds in redox microenvironments." Journal of the Mechanical Behavior of Biomedical Materials 5(1): 32-40, Copyright 2012, with permission from Elsevier.

To directly demonstrate the effect of reducing agent on mechanics, we designed a setup in which we pull the ends of the protein using Steered Molecular Dynamics at a constant rate, while adding a bias potential between the sulfur atoms of the disulfide bond and an attacking molecule (here, H_2 and O_2 molecules are used as simplified models for reducing and oxidizing agents). The bias potential has a varied strength to model sampling at different effective time-scales, arising from modified effective collision frequency in analogy to the Arrhenius description. Figure 3-3(A) shows the breaking force as a function of the bias potential constant and reflecting varied time-scales of observation. In the system with O_2, the C-S bond always breaks first and the strength does not seem to be influenced by the chemical environment significantly, since the cleavage of the S-S bond cannot be activated. On the other hand, the mechanism and force level of bond failure is quite strongly affected in the system with H_2 as a model reducing agent. As the bias potential constant increases—representing the behavior at longer time-scales—the force level of bond failure drops significantly, and the S-S bond always breaks first instead of the C-S bond. Figure 3-3(B) shows representative force-displacement curves comparing the influence of H_2 and O_2. Because the accessible time-scales in molecular dynamics simulations is limited to nanoseconds and we successfully observe the thiol/disulfide exchange reaction in our simulations, the reaction would likely occur violently in the real time-scales once the molecules are in close proximity to each other. However, the difference of the bond breaking time-scale is hardly observed in experimental studies because the cleavage of a single bond usually happens in a very short time, whereas the diffusion of the molecules occurs at much longer time scales. Hence, the molecular dynamics simulations provide us with an opportunity to explore the detailed configuration changes during the reaction. It is worth noting here that the temperature plays a key role during a chemical reaction. From Bell's theory (Bell 1978), for example, raising the temperature will likely decrease the bond life time, as it increases the vibrational excitation of the S-S bond and also favors the high-entropy broken-bond state. Therefore, in a system at a higher temperature, the breaking force for a disulfide bond would be smaller than the breaking force shown in Figure 3-3.

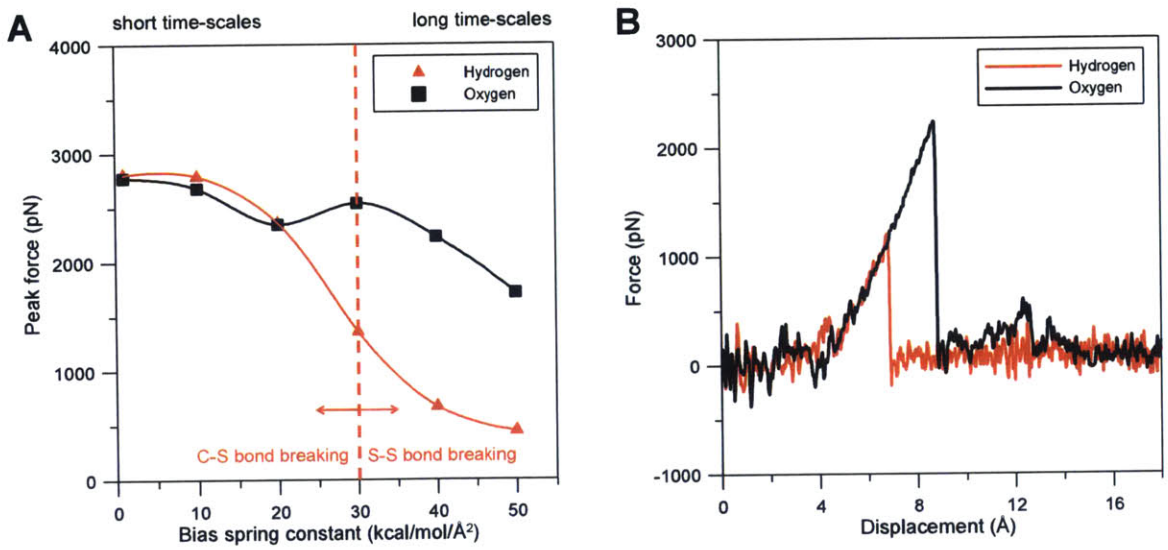

Figure 3-3 Changes of nanomechanical strength properties of disulfide bond under presence of a reducing and oxidizing agent. Panel A: Breaking force as a function of the bias potential spring constant. This plot shows the bond breaking force is strongly dependent on the presence of a reducing or oxidizing environment. In the system with O_2, an oxidizing environment causes a small reduction in strength, but the C-S bond always breaks first. In stark contrast, the system with H_2 shows that the mechanism and force required to induce bond breaking is strongly affected in a reducing environment. As the bias potential constant increases (representing the behavior at longer time-scales) the peak forces drop to significantly smaller levels than in the oxidizing environment and the S-S breaks first rather than C-S bond. The fitting function of the trend lines is a cubic spline function. Panel B: Force-displacement curve showing a comparison of systems with H_2 and O_2 at a fixed bias potential constant of 30 kcal/mol/A^2. The result shows a much lower peak force due to the presence of H_2, very different from the O_2 case, with a reduction by 45 %). (Each data point is an average of every ten points of data.). Reprinted from Keten, S., C. C. Chou, A. C. T. van Duin and M. J. Buehler. "Tunable nanomechanics of protein disulfide bonds in redox microenvironments." Journal of the Mechanical Behavior of Biomedical Materials 5(1): 32-40, Copyright 2012, with permission from Elsevier.

The curves show a much lower peak force and earlier rupture due to the presence of H_2. To solidify our findings, we also perform metadynamics simulations to obtain the free

energy landscape of each system with a bias potential between the sulfur atoms of the disulfide bond and an attacking molecule (H_2 and O_2 molecules). The results of the resulting energy barriers are listed in Table 3-1. In the systems with O_2, the disulfide bond remains stable with an increasing bias potential. On the other hand, as the bias potential constant increases in the system with H_2, the energy barrier drops drastically, in agreement with the results shown in Figure 3-3 where a smaller force is shown to be adequate to overcome the effective energy barrier. These results have two important implications. First, they directly show that the presence of a reducing agent leads to a significant reduction in the force required to break the disulfide bond. Second, these results suggest that the behavior of disulfide bonds is ultimately time-scale dependent, where the weakening effect of the reducing environment is only seen at longer effective sampling times, as shown here qualitatively through the use of bias potentials. This indicates that the mechanics of disulfide crosslinks are rate-dependent; similar to the observations made on weaker protein interactions such as hydrogen bonds (Brockwell, Paci et al. 2003, Ackbarow, Chen et al. 2007, Dietz and Rief 2008, Keten and Buehler 2008).

Table 3-1 The energy barrier of disulfide bond rupture under the hydrogen and oxygen molecules. This table shows the energy barrier is strongly dependent on the presence of a reducing or oxidizing environment. In the system with O_2, the energy barriers fluctuate in the range of 39 to 68 kcal/mol, but the larger spring constant does not cause a smaller energy barrier. Conversely, the system with H_2 shows that the reducing environment weakens the stability of disulfide bond. The energy barriers drop dramatically while the bias potential constant increases.

Bias potential spring constant (kcal/mol/Å^2) Case considered	1	10	20	30	40
Protein with molecular oxygen, O_2	43.0	39.0	68.0	46.1	51.8
Protein with molecular hydrogen, H_2	48.4	55.0	30.6	26.2	*

* As the spring constant equals to ≈40 kcal/mol/Å^2, the energy minimum is located at a S-S distance of 15 Å, and the S-S bond breaks spontaneously.

3.4 Conclusions

In this chapter, we demonstrate that the stability of the disulfide bond is strongly influenced by the redox potential of the chemical environment using the first principles based ReaxFF reactive force field. A direct simulation of the fracture strength showed that the presence of a reducing agent drastically changes the strength of a protein disulfide bond, and controls the rupture location of a bond. The results of metadynamics simulations confirm our Steered Molecular Dynamics results and show that the reducing agent decreases the energy barrier of disulfide bond rupture. The first principles based ReaxFF reactive force field provides a computational approach for studying the link between chemistry and mechanics of protein materials in various physical and chemical environmental conditions. Future computational studies of the structural dynamics of disulfide bonds subject to chemical cues will hopefully provide insight into the mutable mechanics of protein crosslinks, a biological mechanism that could potentially be exploited in bio-inspired polymers or targeted drug delivery systems (Bernkop-Schnürch, Schwarz et al. 1999, Leitner, Walker et al. 2003, Napoli, Valentini et al. 2004, Albrecht and Bernkop-Schnurch 2007, Cerritelli, Velluto et al. 2007, Lv, Xu et al. 2010).

4 Structure and mechanical properties of trichocyte human keratin intermediate filament protein

(This chapter is adapted from (Chou and Buehler 2012))

Keratin is a protein in the intermediate filament family, and the key component of hair, nail and skin. Here we report a bottom-up atomistic model of the keratin dimer, using the complete human keratin type k35 and k85 amino acid sequence. A detailed analysis of geometric and mechanical properties through full-atomistic simulation with validation against experimental results is presented. We introduce disulfide crosslinks in a keratin tetramer and compare the mechanical behavior of the disulfide bonded systems with a system without disulfide bonds. Disulfide bond results in a higher strength (20% increase) and toughness (49% increase), but the system loses alpha-helical structures under loading, suggesting that disulfide bonds play a significant role in achieving the characteristic mechanical properties of trichocyte alpha-keratin. Our study provides general insight into the effect of disulfide crosslink on mechanical properties. Moreover, the availability of an atomistic model of this protein opens the possibility to study the mechanical properties of hair fibrils and other fibers from a bottom-up perspective.

4.1 Introduction

Keratin is an intermediate filament (IF) protein and the key component of hair, nail and skin in vertebrates. The human intermediate filament database has shown that at least 54 of the 70 intermediate filament genes are keratins, making this a very prominent form of IFs that are very widely found (Szeverenyi, Cassidy et al. 2008). According to their mechanical properties, keratin protein materials are divided into trichocyte keratin and epithelial keratin. Trichocyte keratin has a larger elastic modulus but low extensibility, and epithelial keratin has a much lower elastic modulus but larger extensibility. For

different trichocyte keratin proteins, the elastic modulus ranges from approximately 1-4 GPa. For example, the modulus is 1.6-4.5 GPa (Oxenham 1989, Chou and Overfelt 2011) in wool, 2.0-3.7 GPa (Fudge and Gosline 2004, Guthold, Liu et al. 2007, Seshadri and Bhushan 2008) in human hair, 1.0-3.5 GPa(Chou and Overfelt 2011) in porcupine quill, and 0.4 GPa (Bertram and Gosline 1987, Chou and Overfelt 2011) in hoof. Epithelial keratin, on the other hand, is usually found in epithelia cells of skin and its elastic modulus is significantly lower, and typically about 3 MPa (Park and Baddiel 1972).

What's interesting from an engineering perspective is that trichocyte keratin is an extremely durable material that has been widely used in commercial products over 2,000 years, for example as wool in the textile industry. The marvelous durability of the material is exemplified by extremely old artifacts that exist to date. For example, a wool carpet from 500 B.C. is exhibited in "*The State Hermitage Museum*" in St Petersburg (Popescu and Wortmann 2010), and wigs made of human hair originating from ancient Egypt have been found (Kitahara and Ogawa 1991, Robins 1999), all structurally intact. Apparently, trichocyte keratin features a very long degradation time, perhaps due to its high resistance against mechanical loading and other environmental influences such as chemical degradation that break down many other biological tissues. Due the outstanding durable features of trichocyte keratin, several researchers have examined this class of proteins from the 1930s onwards (Speakman 1927, Astbury and Street 1932, Astbury and Woods 1934, Kreplak, Doucet et al. 2001, Kreplak, Doucet et al. 2004), and analyzed the stress-strain curve of wool (Speakman 1927), examined the trichocyte keratin protein structure, and the alpha helix-beta sheet transition that was observed early on during the stretching of hair and wool (Astbury and Street 1932, Astbury and Woods 1934, Kreplak, Doucet et al. 2001, Kreplak, Doucet et al. 2004).

As a result of experimental work, a deformation mechanism model was proposed aimed at explaining the mechanical behavior of keratin proteins (Bendit 1960, Chapman and Feughelman 1967, Wortmann and Zahn 1994, Hearle 2000, Feughelman 2002, Kreplak, Franbourg et al. 2002). However, due to the lack of a structural model of keratins with atomistic resolution, relatively little is known of the physical material concepts. Previous computational studies of keratin proteins (or other alpha-helix intermediate filament

proteins) used different methods to construct atomistic models, where a homology approach using the crystal structure of vimentin and GCN4 leucine zipper has been applied to build models of keratin proteins (Liovic, Stojan et al. 2001, Akkermans and Warren 2004, Smith and Parry 2008), and several works focused on vimentin intermediate filaments have been reported (Qin, Kreplak et al. 2009, Arslan, Qin et al. 2011, Qin and Buehler 2011, Qin, Chou et al. 2012). However, although it is the key molecular constituent of trichocyte keratin, no full-atomistic structural model of the entire keratin heterodimer molecule exists to date. An atomistic model is critical in order to probe the properties of keratin proteins and to exploit the insight for bioengineering and other applications, and specifically to address the origin of this material's durability by assessing a direct link between the molecule's chemical structure (including the abundance of disulfide crosslinks) and its larger-scale properties.

In this study, we focus on the study of structure and mechanical properties of the human hair keratin protein. It is well established that human hair (and many related biomaterials) exhibits a hierarchical structure (Feughelman 1997, Feughelman 2002) as shown in Figure 4-1(A), ranging from the alpha-helix, dimer (coiled-coil) consisting of two alpha-helical chains, microfibril consisting of 16 dimers, and microfibrils and matrix assembling into macrofibrils, cell to entire hair fibers. Previous studies suggest that disulfide bonds, formed between dimmers as shown schematically in Figure 4-1(B), play a crucial role in defining the stability of this protein and provide a critical feature in mechanical properties (Yu, Yu et al. 1993, Naito and Arai 1996, Parbhu, Bryson et al. 1999, Wang, Parry et al. 2000, Fudge and Gosline 2004, Strnad, Usachov et al. 2012). The disulfide bond can only be formed by two cysteine residues and typically occurs in oxidizing environments.

The study reported in this chapter is aimed to build the keratin heterodimer based on the full amino acids sequence of keratin proteins (k35/k85) and investigate the geometric and mechanical properties of resulting protein fibers. We also consider the deformation and rupture behavior of a keratin tetramer intermolecular linked by disulfide bonds or H-bonds to gain insight into the nanomechanics of disulfide crosslinks in keratin proteins under external loading.

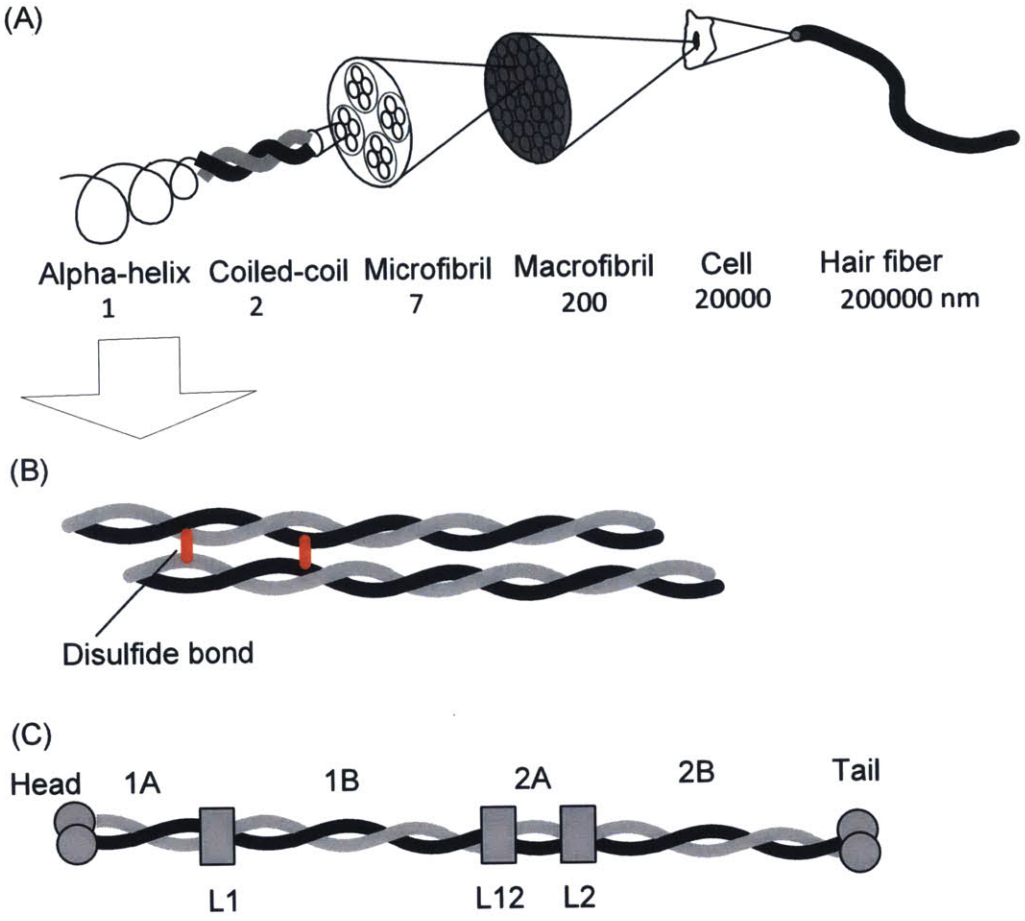

Figure 4-1 Hierarchical structure of human hair. (A) Human hair features a hierarchical structure, ranging from alpha-helix, dimers with a coiled-coil structure, microfibrils, macrofibrils to the cellular and eventually entire hair fiber level. (B) Schematic visualization that shows how two dimers are connected by intramolecular disulfide bonds. (C) Schematic representation of the configuration of a keratin dimer with key domains labeled. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

4.2 Materials and Methods

4.2.1 Atomistic keratin dimer and tetramer structures

Trichocyte keratin is generally divided into type I (k31–k40) and type II (k81–k86) proteins according to their physical and chemical properties. For example, keratin type I is neutral or basic and keratin type II is acidic. In contrast to other intermediate filaments, keratin homodimers are very unstable, and the basic structural unit of intermediate filaments is a heterodimer of one type I keratin, where its partner is type II keratin (Hatzfeld and Weber 1990, Steinert 1990). In this study, we build an atomistic model of the keratin heterodimer composed of k35 and k85 (these sequences are selected based on the availability of experimental data for comparison). Like all intermediate filament proteins, the characteristic of keratin is the central alpha-helical rod domain that is composed of four helices (1A, 1B, 2A, and 2B), three linker domains (L1, L12, and L2), and terminal groups formed by globular head and tail domains as shown in Figure 4-1(C). Based on the sequence analysis in experiments (Szeverenyi, Cassidy et al. 2008), the amino acids in the alpha-helical domains (1A, 1B, 2A, and 2B) are connected by dihedral angles of $\varphi = -58°$ and $\psi = -47°$, and others in the linker (L1, L12 and L2), head and tail domains are largely unstructured, and connected by dihedral angles $\varphi = 180°$ and $\psi = 180°$ (Branden and Tooze 1999, Alberts, Johnson et al. 2002) in the initial configuration.

We use two coiled-coil structure prediction algorithms, COILS and Paircoil2, to determine the geometric characteristic of coiled structure (most importantly the twisting angle) and the position of each amino acid in the heptad-repeated coiled coil structure (Lupas, Van Dyke et al. 1991, Lupas 1996, McDonnell, Jiang et al. 2006). The coiled-coil prediction algorithms can identify whether or not an unknown sequence contained coiled-coil characteristics and determine its features in the heptad repeat. In COILS, the algorithm compares the unknown sequence with known coiled-coil structures and utilizes the frequencies of amino acids at different positions of the heptad repeats in their database to predict the coiled-coil structure. In Parcoil2, correlations of amino acid occurrence and the position of each in the heptad repeat is used to predict whether or not

a coiled-coil is formed. A comprehensive comparison of performance of these two algorithms was reported in the literature (Gruber, Soding et al. 2006).

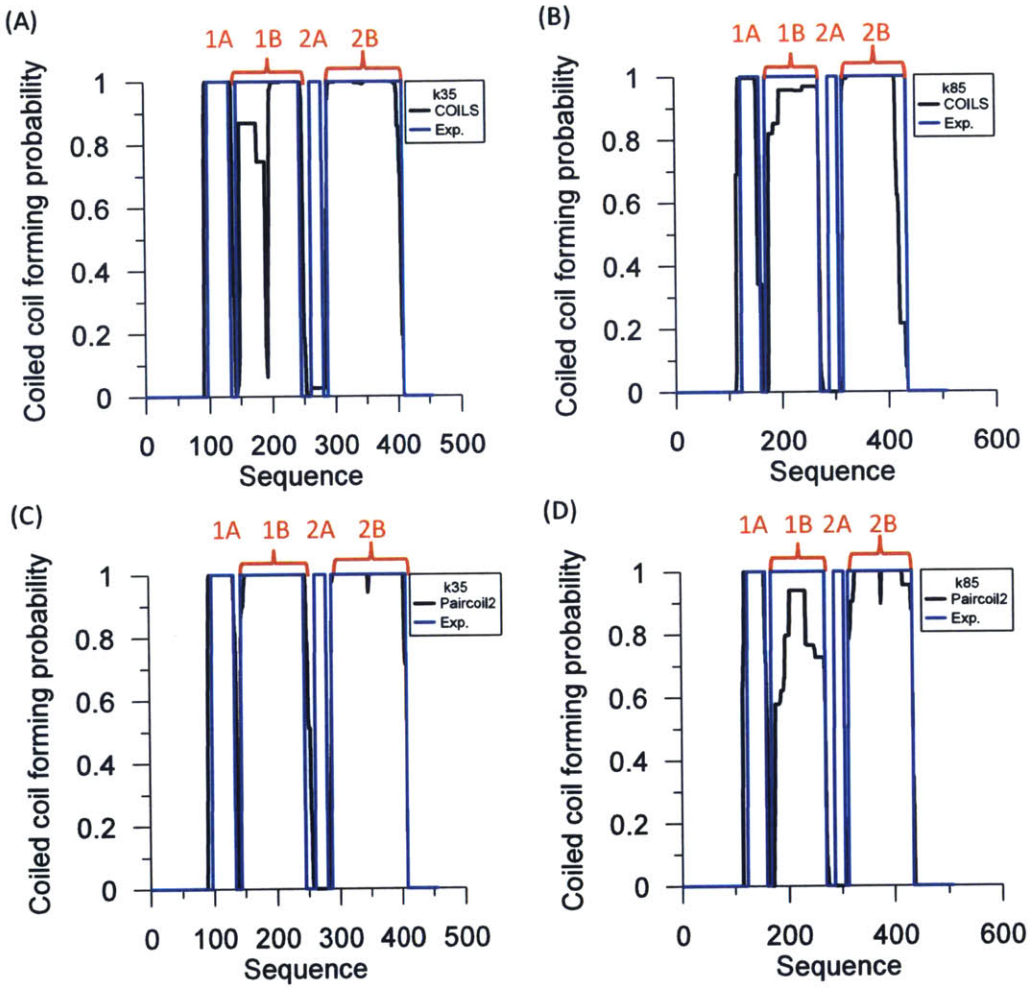

Figure 4-2 Coiled-coil-forming probability of k35 and k85 based on the COILS (A and B) and Paircoil2 (C and D) algorithms. Comparison of coiled-coil-forming probability of k35 and k85 sequence determined from different prediction algorithms, and experimental results(Szeverenyi, Cassidy et al. 2008). The probability shows whether the sequence will adopt a coiled coil structure or not. The high probability means the sequence has the high potential to form a coiled coil structure and the coiled coil structure is unlikely if the sequence has a low coiled-coil-forming probability. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

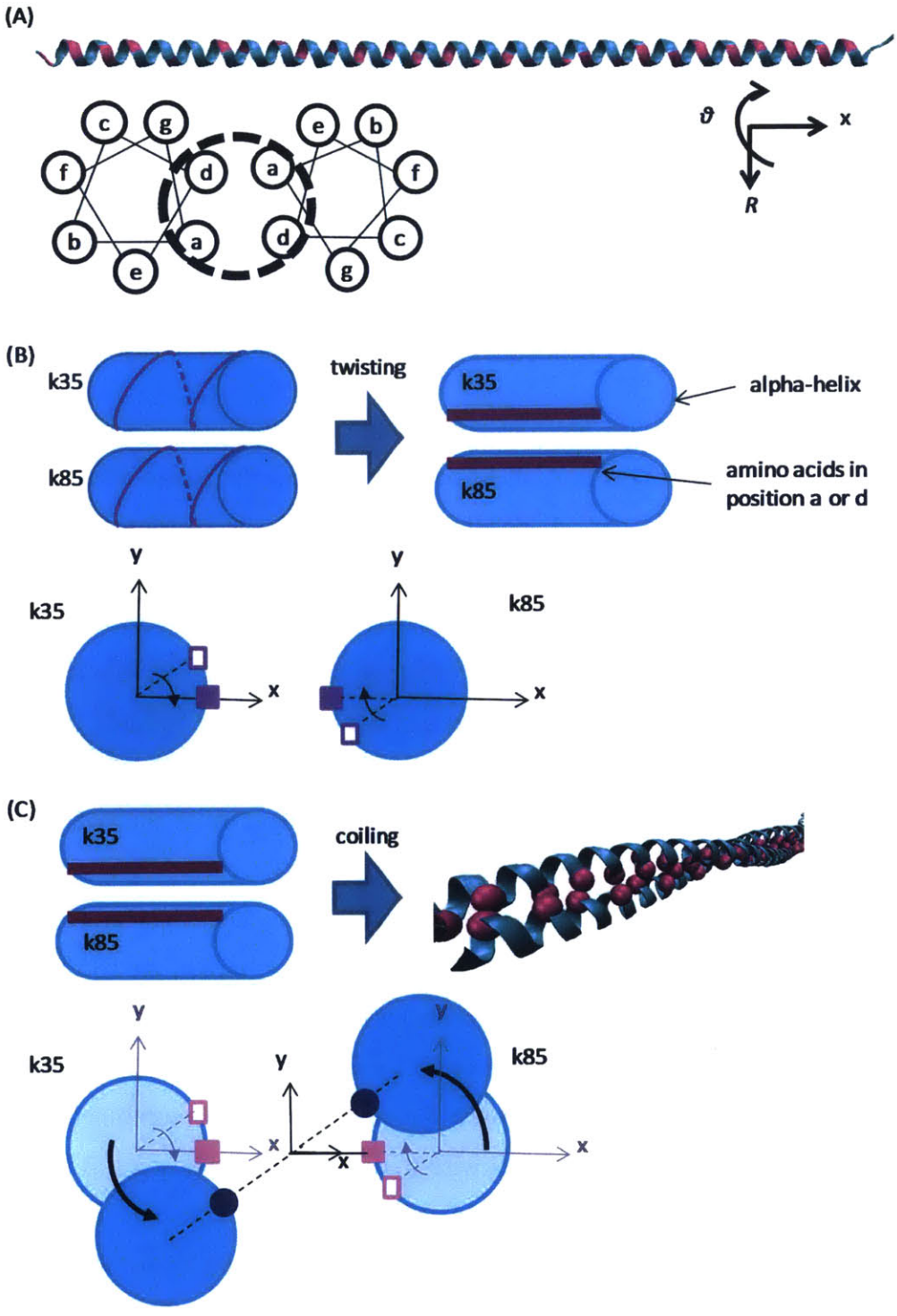

Figure 4-3 Geometric approach to construct the initial dimer model. (A) The amino acids situated in the core of the dimer (the 'a' and 'd' positions in the heptad repeat shown below) in k35 1B domain predicted by COILS generate a purple stripe winding around its helical axis with a left-handed heptad repeat. Panel (B) and (C) display the steps and the

mechanism of building the initial structure of keratin heterodimer. The alpha-helical domain is plotted as a blue cylinder with the purple stripe. The open square (□), the solid square (■) and the solid circle (●) mark the positions of an amino acid on the purple stripe before twisting, after twisting and after coiling in a dimer, respectively. This process is applied systematically to all domains in the protein. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

Figure 4-2 shows the coiled-coil forming probability of each amino acid predicted by COILS and Paircoil2. The results from these predictions show the amino acids sitting in the core of the dimer (the 'a' and 'd' positions in heptad repeat (abcdefg) depicted in Figure 4-3(A)) in 1A, 1B and 2B domains generate a stripe winding around its helical axis with a left-handed heptad repeat (the stripe colored in purple shown in Figure 4-3(A)). About 70% of amino acids in the stripe are hydrophobic and the structure is energetic favored while these amino acids form a stable hydrophobic bonding in the core of dimer. We build the initial structure of dimer by twisting the k35 and k85 chains into a coiled coil structure so as to shield amino acids in two stripes from the polar solvent environment (Apgar, Gutwin et al. 2008). First, in Figure 4-3(B), each purple stripe in the alpha-helical domain of k35 and k85 is twisted clockwise along the helical axis to line the amino acids in 'a' and 'd' in a straight line. Then we align the stripes in the k35/k85 face-to-face, letting the stripes form a stable hydrophobic bonding. The lower part in Figure 4-3(B) depicts the geometry transformation of k35 and k85 with respect to their helical axis. Next, due to the torsion applied to twist the alpha-helices, the combined k35 and k85 chains coil back *together* counter-clockwise to reach a stress-free state. Because the amino acids in purple stripes form a stable hydrophobic bonding in the core of dimer, the alpha-helices coil with respect to the axis of the dimer to shield the hydrophobic amino acids inside the dimer shown in Figure 4-3(C). In our algorithm we manually coiled two chains. The right part in Figure 4-3(C) summarizes the two steps of the geometry transformation of k35 and k85 chains in constructing the initial model of keratin heterodimer. The twisting angle is determined from the linear fitting of the angular coordinate θ of the hydrophobic amino acids with respect to their axial coordinate x in the

R–θ–x coordinate system of C_α atoms shown in Figure 4-4(A-F) and Table 4-1. The stripes are almost parallel to the axis of the alpha-helix in the 2A domain and form a stable hydrophobic bonding in the heterodimer without twisting. Due to varied sequence in each domain, the twisting angles are distinct in different segments, and the angle for coiling each domain is the average of twisting angles in k35 and k85 chains. The initial configuration of the k35/k85 heterodimer based on the twisting angles predicted by COILS is shown in Figure 4-5(A). All visualizations for proteins are done using the Visual Molecular Dynamics (Humphrey 1996).

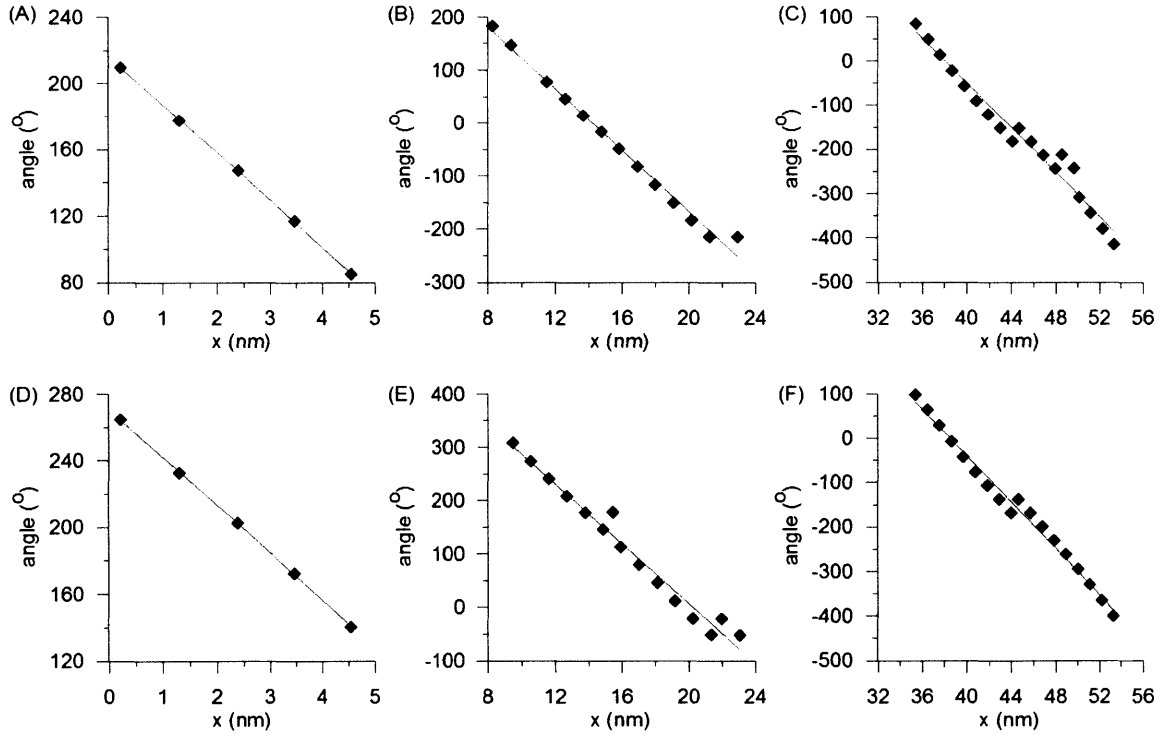

Figure 4-4 Twisting angles of the k35 and k85 sequences based on the predictions based on the COILS algorithm. The geometric character of each coiled-coil domain in a keratin dimer is shown: (A) k35-1A, (B) k35-1B, (C) k35-2B, (D) k85-1A, (E) k85-1B and (F) k85-2B. The twisting angle is obtained by using the slope of the trend line in each domain. The values of twisting angles in each domain are listed in **Table 1**. Reprinted with permission from Chou, C. C. and M. J. Buehler (2012). "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

Table 4-1 Twisting angles (°/Å) of each coiled coil domain in k35/k85 heterodimer based on the predictions of the COILS and Paircoil2 algorithms, respectively.

Sequence/ Algorithm \ Domains		1A	1B	2B
k35	COILS	2.875	2.90	2.52
	Paircoil2	2.875	3.022	3.149
k85	COILS	2.86	2.80	2.61
	Paircoil2	2.86	3.02	3.2

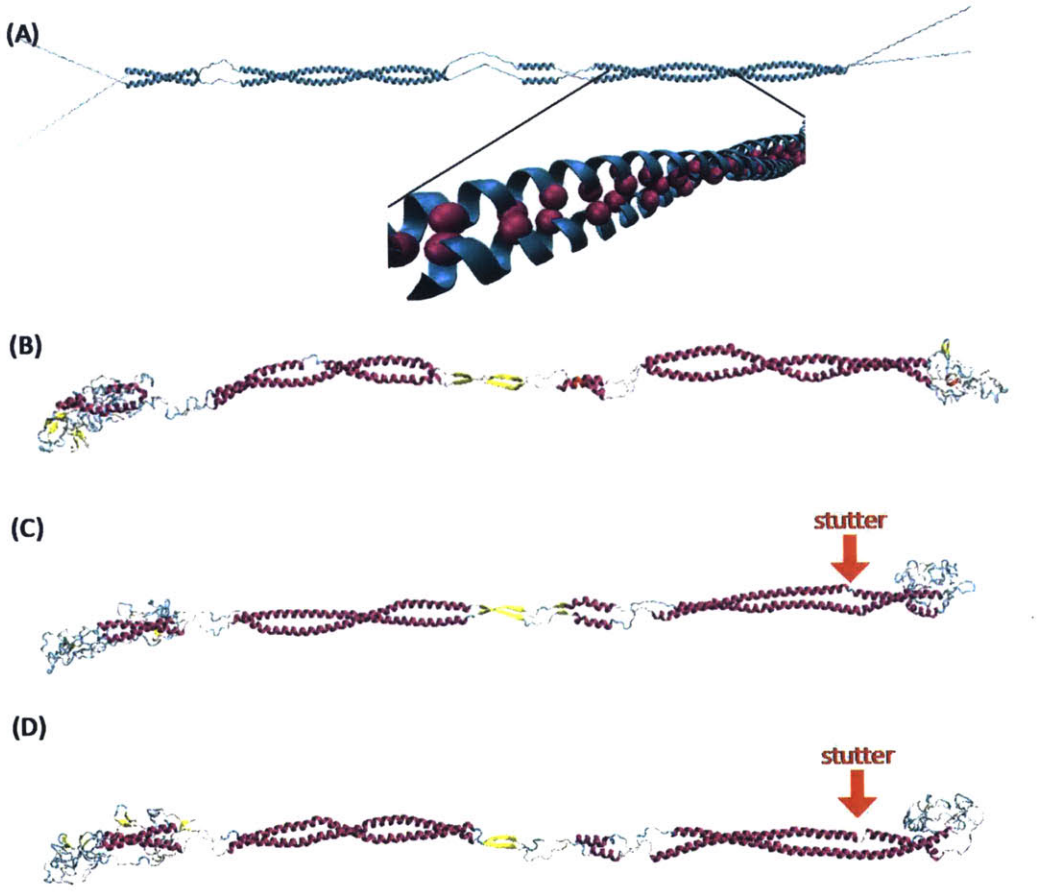

Figure 4-5 Visualizations of the keratin heterodimer initial and equilibrium configurations in implicit and explicit solvent. (A) Snapshots of initial configuration of k35/k85 heterodimer in implicit solvent predicted by the COILS algorithm. The amino acids in the core of heterodimer are highlighted in purple shown in the close-up. K35 and

k85 keratin proteins are assembled into a heterodimer based on the predicted coiling angles from two different algorithms. Assembled keratin dimers are equilibrated using the CHARMM force field with implicit solvent. (B) and (C) show the snapshots of k35/k85 heterodimers in implicit solvent predicted by the Paircoil2 and COILS, respectively. The model based on COILS is used in the simulation with explicit solvent. Panel (D) displays a snapshot of the k35/k85 heterodimer after equilibration in explicit solvent. For clarity, surrounding explicit water molecules are not shown. In panel (B), (C) and (D), the alpha-helix, the beta-sheet structure and random coiled structure are colored in purple, yellow and blue, respectively. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

At a higher hierarchical level there exists a keratin tetramer that consists of two anti-parallel keratin heterodimer (Feughelman 1997, Feughelman 2002). In earlier work it was proposed that there are three possible alignment models of two heterodimers assembling into a tetramer. However, close inspection of the sequence and structure suggests that disulfide crosslinks can only exist in one of the alignment models, the A_{11} model(Wang, Parry et al. 2000, Parry, Marekov et al. 2001). Previous study shows the connectivity of the disulfide bonds in the keratin tetramer including the sequence number of cysteine amino acid and the relative orientation of cysteine in coiled-coil structure(Fraser and Parry 2007). Although the keratin sequences studied in the previous literature(Fraser and Parry 2007) are slightly different from this study, the trichocyte keratin sequences share a high level of similarity(Marchuk, McCrohon et al. 1985, Rogers, Winter et al. 1997), which allows us to adapt the disulfide crosslink model for the case considered here. Furthermore, due to the geometric constraints of disulfide bond connectivity (specifically, $\approx$0.2 nm disulfide bond length and $\approx$0.4 nm(Fraser, MacRae et al. 1988) β-carbon-β-carbon separation in a disulfide bond between two α-helices), there is only one feasible disulfide bond connectivity. This refers to the formation of a disulfide bond between cysteine residue160 of k35 and cysteine residue 234 of k85.

4.2.2 Atomistic simulation

We use both implicit and explicit solvent models for equilibration. Molecular simulations using implicit solvent are carried out using the CHARMM19 all-atom energy function with an effective Gaussian model for the water solvent (Lazaridis and Karplus 1997, Lazaridis and Karplus 1999). An energy minimization using the steepest descent method is performed. The Effective Energy Function 1 (EEF1) is used to compute solvation free energy and nonbonding interactions. After energy minimization, all simulations are performed at a temperature of 300 K at *NVT* ensemble, with a 1 fs time step. For the keratin heterodimer, the molecular dynamics equilibration is performed with total simulation time of 50 ns. The first 10 ns of the simulation are used to equilibrate the system and are excluded from structural analyses due to observed structural changes. During the last 40 ns simulations, the structure remains stable and configurations of keratin are recorded every 10 ps, resulting in 4,000 frames. Data collected every 5 frames of the 4,000 frames are used in the analyses. Molecular simulation using implicit solvent reduces the computational cost to obtain a first estimate of the structure in the equilibrium state. Although the implicit solvent has limitations to describe the ionic solvent and overestimate the some energy terms in particular structural domain of protein (Zhou and Berne 2002), it adequately describes the equilibrated structure and dynamics behavior of alpha-helical proteins in water solvent (Qin, Kreplak et al. 2009, Qin and Buehler 2011). However, further equilibration in explicit water solvent is needed to examine the structural stability and to refine the molecular model.

As such, for structure refinement, simulations with explicit water molecules are carried out with the molecular simulation software LAMMPS (Plimpton 1995) (the entire atomistic structure is embedded completely in a periodic water box composed of TIP3 explicit water molecules). The CHARMM force field is applied (MacKerell, Bashford et al. 1998). Energy minimization and finite temperature equilibration of all structures are performed before the protein is subjected to mechanical loading. An energy minimization using a conjugate gradient algorithm is performed before finite temperature equilibration. Nonbonding interactions are computed using a switching function between 8 and 10 Å for van der Waals interactions. Electrostatic interactions are modeled by the

particle–particle particle–mesh (PPPM) method. After energy minimization, all simulations were performed at a temperature of 300 K at *NVT* ensemble, with a 1 fs time step. For the keratin heterodimer, the explicit water molecular dynamics simulation is performed with total simulation time of 20 ns. The first 10 ns simulations are used to equilibrate the system and are excluded from analyses. During the last 10 ns simulations, the configurations of keratin are recorded every 2 ps which results in 5,000 frames. Every 5 frames of the 5,000 frames are used in the analyses. For the keratin tetramer, the molecular dynamics simulation is performed with total simulation time of 10 ns. The first 2 ns simulations are used to equilibrate the system and are excluded from analyses. During the last 8 ns simulations, the configurations of keratin are recorded every 2 ps which results in 4,000 frames. The systems of heterodimer and tetramer embedded in explicit solvent have ≈250,000 atoms with a box size of 56 nm × 6.5 nm × 6.0 nm and ≈500,000 atoms with a box size of 72 nm × 10 nm × 8 nm, including explicit water molecules, respectively. The computational cost is about 2,500 CPU hours in heterodimer case and 6,000 CPU hours in tetramer case for structure equilibrium in explicit solvent. In order to reduce the computational cost, the truncated tetramer model which is composed of 1B domain of k35/k85 and two disulfide crosslinks is used in the study of mechanical testing. The Supporting Information includes atomistic structures of the keratin heterodimer in explicit and implicit solvent in the Protein Data Bank (PDB) format.

In this study, one of our goals is to study the mechanical properties and rupture behavior of the disulfide bonded keratin proteins. The disulfide bond serves as a strong but flexible connection on account of its variable bond strength in different chemical environments. Here we introduce an additional cutoff in CHARMM force field to model the breakable behavior of disulfide bond and show the effect of disulfide bond with varied bond strength on keratin proteins under an external force. We consider disulfide crosslink with bond energy of 7.17 kcal/mol in the disulfide bonded system.

4.2.3 Steered Molecular Dynamics

To apply forces to the molecule to induce deformation, we use Steered Molecule Dynamics (SMD), with an SMD spring constant k_{SMD} =10 kcal/mol. For the keratin

k35/k85 heterodimer, the two C_α atoms at the end of 1A segment are fixed and the other C_α atoms at the end of 2B segment are pulled by using SMD. In the truncated tetramer simulation, the force is applied to the two C_α atoms at one end of one dimer and the two C_α atoms of the other dimer at the other end are fixed in the alpha-helix structure, with pulling velocities v = 0.1 Å/ps. The force-displacement curve is obtained by monitoring the time-averaged applied force and the position of the atom that is pulled at fiber direction over the simulation time is recorded. Mechanical tensile testing is performed using LAMMPS (Plimpton 1995) with explicit water solvent, properly accounting for hydrophobicity, viscosity, and other effects.

4.3 Results and Discussion

We first show the structural analysis of the equilibrated geometries of both heterodimer and tetramer of keratin k35/k85 protein and compare with experimental data. Then we proceed to study the mechanical properties of the keratin heterodimer and tetramer and present the effect of disulfide crosslink on mechanical properties of keratin protein in the results of mechanical testing.

The initial configuration of k35/k85 heterodimer is shown in Figure 4-5(A), and reveals that the heterodimer is composed of two alpha-helix chains (k35 and k85) that are assembled based on the predicted twisting angles (Table 4-1) obtained from two different algorithms (Lupas, Van Dyke et al. 1991, McDonnell, Jiang et al. 2006). Assembled keratin heterodimers are equilibrated using CHARMM with implicit solvent. Figure 4-5(B) and (C) show the snapshots of k35/k85 heterodimers after equilibration in implicit solvent predicted by the Paircoil2 and COILS. First, we observe that the 1A domain tangle with the head domain and due to the interaction between the hydrophobic amino acids in 1A and head domains, the hydrophobic amino acids in 1A domain might not shield in the core of dimer. The 1A domain loses the well coiled-coil structure, in agreement with the suggestions in the previous study (Parry, Marekov et al. 2002). In Figure 4-5(C), the two chains in 1B domain form a stable coiled-coil structure, in agreement with the literatures(Parry, Strelkov et al. 2007, Smith and Parry 2008) in which the 1B domain is observed as a stable coiled form in absence of discontinuity and

plays a key role in providing the stability of self-assembly of keratin proteins. Figure 4-5(B) and (C) show the linker L12 domain as a beta-sheet structure in the heterodimer and our results and the earlier study (Parry and Steinert 1999) are in the well agreement. The 2B domain is the longest coiled-coil domain in the heterodimer and has less stability than 1B domain due to the presence of a stutter (the discontinuity shown in the Figure 4-5(C)), and the amino acids on the right end of the 2B domains unwind. These observations in the simulation are in the agreement with earlier experimental studies (North, Steinert et al. 1994, Brown, Cohen et al. 1996, Smith and Parry 2008). It is worth noting here that the equilibrated structure based on the COILS prediction captures more structural features, such as the beta-sheet structure in L12 domain and stutter in 2B domain shown in the Figure 4-5(C), compared with the structure based on the Paircoil2 prediction depicted in Figure 4-5(B) after implicit solvent equilibration.

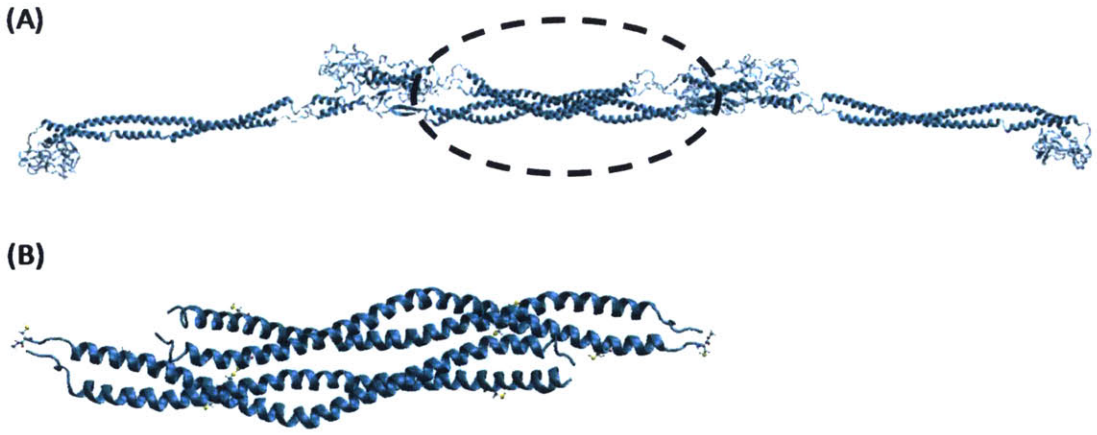

Figure 4-6 Keratin full-length and truncated tetramer model for tensile testing. (A) The full-length keratin tetramer model is composed of two anti-parallel heterodimers connected by two disulfide crosslink. (B) Visualization of the truncated keratin tetramer model. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

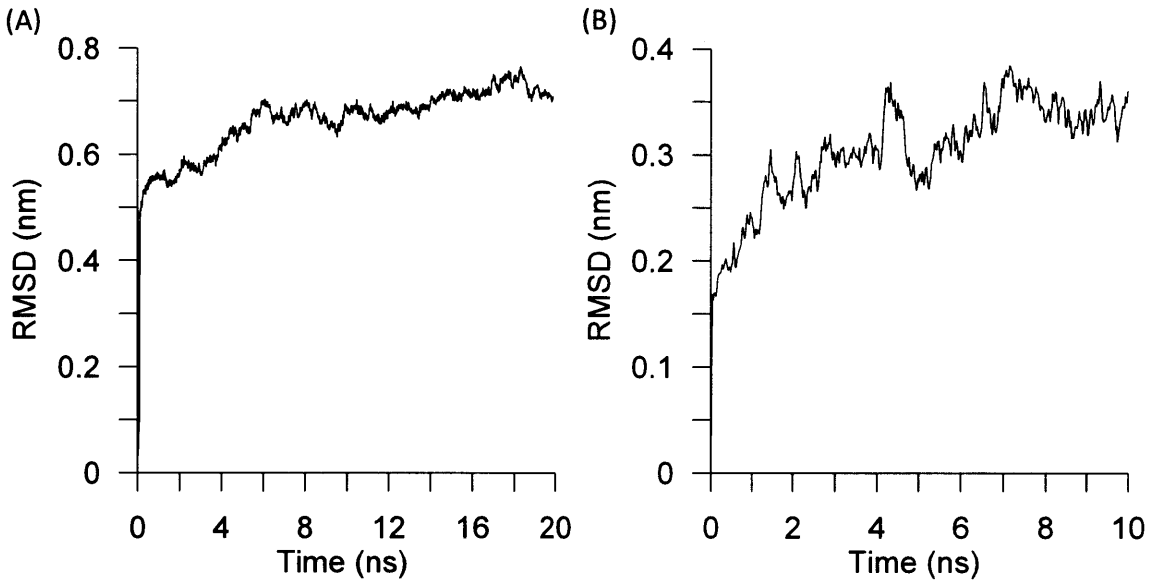

Figure 4-7 The root mean square displacement (RMSD) versus simulation time, for (A) keratin heterodimer and (B) tetramer systems in explicit solvent. After several nanoseconds simulation time, the RMSDs of heterodimer and tetramer systems fluctuate around a constant value, suggesting that the systems has reached a stable configuration. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

The keratin tetramer is composed of two anti-parallel heterodimers with the A_{11} alignment model shown in Figure 4-6(A). Referring to the disulfide bonded tetramer model based on experimental data(Fraser and Parry 2007), we introduce two disulfide bonds in our model, which intermolecular connect residue 160 of k35 and residue 234 of k85 in the two heterodimers. The average diameter of tetramer is ≈2.2 nm, which is close to the experimental data of 2 nm (Wortmann and Zahn 1994). To reduce the system size for mechanical testing, the head, tail, 1A, 2A, 2B, L1, L12 and L2 domains of the keratin protein distant from the disulfide bond are removed, and only residues in 1B domains are kept as depicted in Figure 4-6(B). Figure 4-7 shows the time history of the root mean square deviation (RMSD) for the dimer (Figure 4-7(A)) and the truncated tetramer (Figure 4-7(B)) in explicit solvent. The results show that after 10 ns and 2 ns, the RMSD values for the heterodimer and tetramer converge to a constant value, showing structural

stability of the predicted structure in the explicit solvent model. Images of the final structures are shown in Figure 4-5(D) and Figure 4-6(B), respectively.

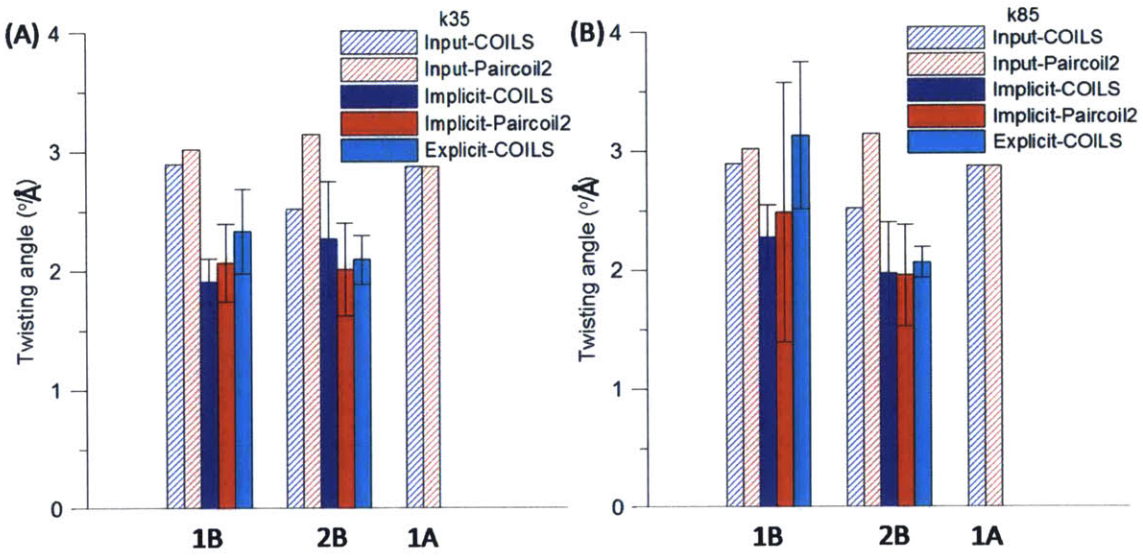

Figure 4-8 Twisting angles of k35/k85 keratin heterodimer after equilibration in implicit and explicit solvent. Comparison between twisting angles predicted by COILS and Paricoil2 in 1B and 2B domains of k35 (A) and k85 (B) before and after equilibration in implicit and explicit solvent. The main bar is the average values of twisting angles in the last 10% of the total simulation time while the simulation reaches the equilibrium state, and the error bar corresponds to the standard deviation. The twisting angle of 1B domain predicted by the Paircoil2 algorithm has a relatively large standard deviation after equilibration in implicit solvent. It shows that the structure predicted by Paircoil2 has less stability so the model based on the COILS prediction is used in the simulation with explicit solvent and following structural comparison, as well as mechanical testing. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

Table 4-2 Structural properties of trichocyte alpha-keratin. Comparison of the length of the rod and inter-rod domains (1A, 1B, 2A and 2B) in the keratin heterodimer determined from experimental study and simulation in both implicit and explicit solvent. The

simulation model is built based on the prediction of COILS. In the reference, the length of each inter-rod domain is the product of number of amino acids and the axial rise per residue in the coiled coil structure, 0.1485 nm (Fraser and Parry 2005).

Properties	Experimental result (nm) (Steinert, Marekov et al. 1993, Fraser and Parry 2005)	Simulation/implicit solvent (nm)	Simulation/explicit solvent (nm)
Length of rod domain	44-47	54.2±0.1	50.2±0.13
1A	5.2	7.03±0.12	7.03±0.1
1B	15	16.55±0.19	16.09±0.16
2A	2.822	4.63±0.11	4.72±0.1
2B	17.97	19.67±0.16	19.45±0.26

We evaluate the twisting angles of each alpha-helical domain structural properties of equilibrated structures in implicit solvent predicted by COILS and Paircoil2. Figure 4-8 shows the comparison between the twisting angles of 1B and 2B domains in implicit and explicit solvent predicted by COILS and Paircoil2. After equilibration in the implicit solvent, the twisting angles are slightly decreased but the majority of the structures are stable and well coiled. However, the twisting angle of 1B domain predicted by the Paircoil2 algorithm has a large standard deviation after equilibration in implicit solvent, suggesting that the structure predicted by Paircoil2 has less stability. Therefore, the model predicted by COILS is used in the simulation with explicit solvent and the following structural comparison and mechanical testing. We further calculate the length of central alpha-helical rod domain and four helical domains (1A, 1B, 2A and 2B) in the heterodimer after equilibration in implicit and explicit solvent shown in Table 4-2. The

simulation results are in the well agreement with the experimental data (Steinert, Marekov et al. 1993, Fraser and Parry 2005).

We calculate the distribution of dihedral angles (φ, ψ) of each amino acid in the full-length heterodimer after equilibration in explicit solvent and 2B domain after implicit and explicit solvent equilibration and plot as the Ramachandran plot (http://en.wikipedia.org/wiki/Ramachandran_plot) with the labels showing the approximate locations of common structure features in biomaterials, such as alpha-helix (α), beta-sheet (β) in Figure 4-9. The plots clearly show the most of dihedral angles of amino acids in alpha-helix domains are located near the α region (where $-90° < \varphi < -15°$, $-70° < \psi < -35°$) and the dihedral angles of amino acids in unstructured domains are dispersed in the Ramachandran plot. Two blue points are the dihedral angles of residues 392 and 393 of k85, showing the local discontinuity that resembles the stutter defect in 2B domain. The model used in explicit solvent is based on the equilibrated structure in implicit solvent as its initial configuration. It is important to note that the Ramachandran plot of the equilibrated structure in explicit solvent is similar to the one in implicit solvent shown in Figure 4-9(C), providing additional evidence for the structural stability. Figure 4-9(D) shows the time history of dihedral angle change of the stutter during the implicit solvent equilibration. In the initial configuration, dihedral angles of residue 392 and 393 are defined as the angles of $\varphi = -58°$ and $\psi = -47°$. During equilibration, the local coiled structure unwinds gradually and reaches a steady state where $-83° < \varphi < -75°$, $45° < \psi < 75°$ during the last 10 ns of the simulation.

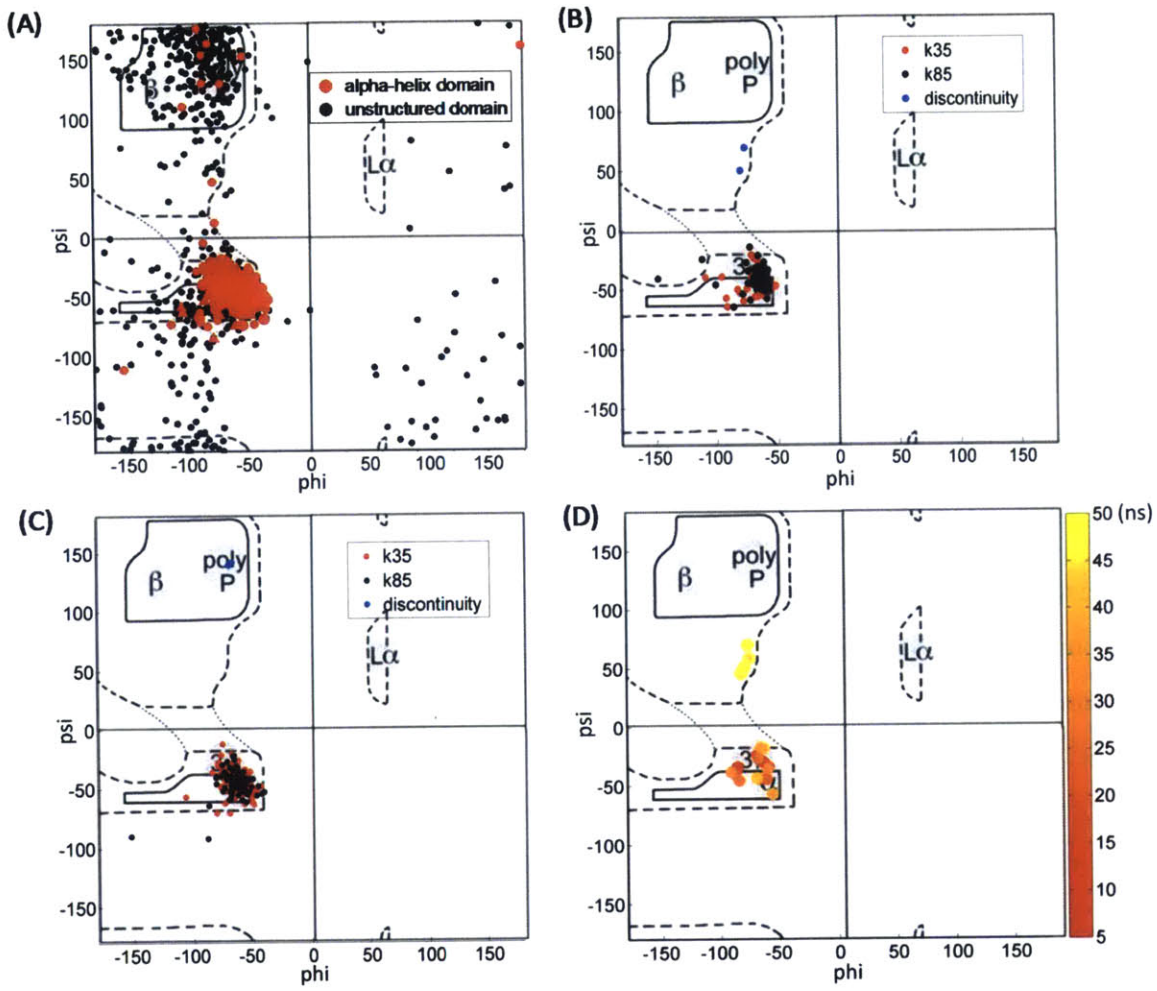

Figure 4-9 Ramachandran plot for (A) the full-length equilibrium structure of heterodimer in explicit solvent and the equilibrium structure of 2B domain in (B) implicit solvent, (C) explicit solvent and (D) time history of discontinuity in 2B domain. Panel (A) shows the dihedral angles (φ, ϕ) of each amino acid in full-length heterodimer at equilibrium state in explicit solvent on the Ramachandran plot (http://en.wikipedia.org/wiki/Ramachandran_plot) (with the labels showing the approximate locations of structural features, such as α-helix (α), β-sheet (β)). Panels (B) and (C) show the dihedral angles of each amino acid in 2B domain at the equilibrium state in implicit and explicit solvent, respectively. The blue points depict the discontinuity in 2B domain. Panel (D) displays the dihedral angle change of the discontinuity in panel (B) during the equilibration in implicit solvent. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin

Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

We proceed to study the mechanical properties of the keratin heterodimer and tetramer structure by direct tensile testing. The goal of these simulations is to study the mechanical properties and rupture behavior of the disulfide bonded keratin proteins. The mechanical testing involves stretching simulation of keratin heterodimer and shearing simulations of the keratin tetramer in the presence of disulfide crosslink and in the absence of disulfide crosslink. The purpose of this investigation is to show the strength of the keratin protein is enhanced by disulfide bond and in the false test, the two heterodimer only stabilized by hydrogen bonds (H-bonds) have less strength under an external loading.

We begin with an analysis of the mechanical behavior of the keratin heterodimer. The force-displacement curve consists of three distinct regimes of deformation as shown in Figure 4-10(A). In the first regime (I), the pulling force increases linearly with displacement until the angular point is reached. In the second regime (II), the pulling force fluctuates around a constant value where alpha-helical domains start unfolding. In the third regime (III), stretching the covalent bonds in the polypeptide backbone of the heterodimer leads to a rapid increase of force as displacement increases. The heterodimer is stretched until all helical domains are fully extended. The characteristics of force-displacement curves of the keratin heterodimer are similar to other alpha-helical or coiled-coil proteins, such as vimentin and myosin proteins (Schwaiger, Sattler et al. 2002, Ackbarow and Buehler 2007, Ackbarow, Chen et al. 2007, Qin, Kreplak et al. 2009). Notably, the magnitude of force value might depend on the pulling rate. In previous studies (Ackbarow, Chen et al. 2007, Ackbarow, Keten et al. 2009) , the pulling rate effect on the force-displacement relationship of alpha-helix protein is shown. For alpha-helical and coiled-coil systems, the higher pulling rate results a greater strength and unfolding force. At 0.001 Å/ps pulling rate, the unfolding force is ≈350 pN and for higher pulling speed which is 0.1Å /ps, the unfolding force is ≈1,000 pN (Ackbarow, Chen et al. 2007), in agreement with the force level in the second regime in the force-displacement curve of keratin heterodimer.

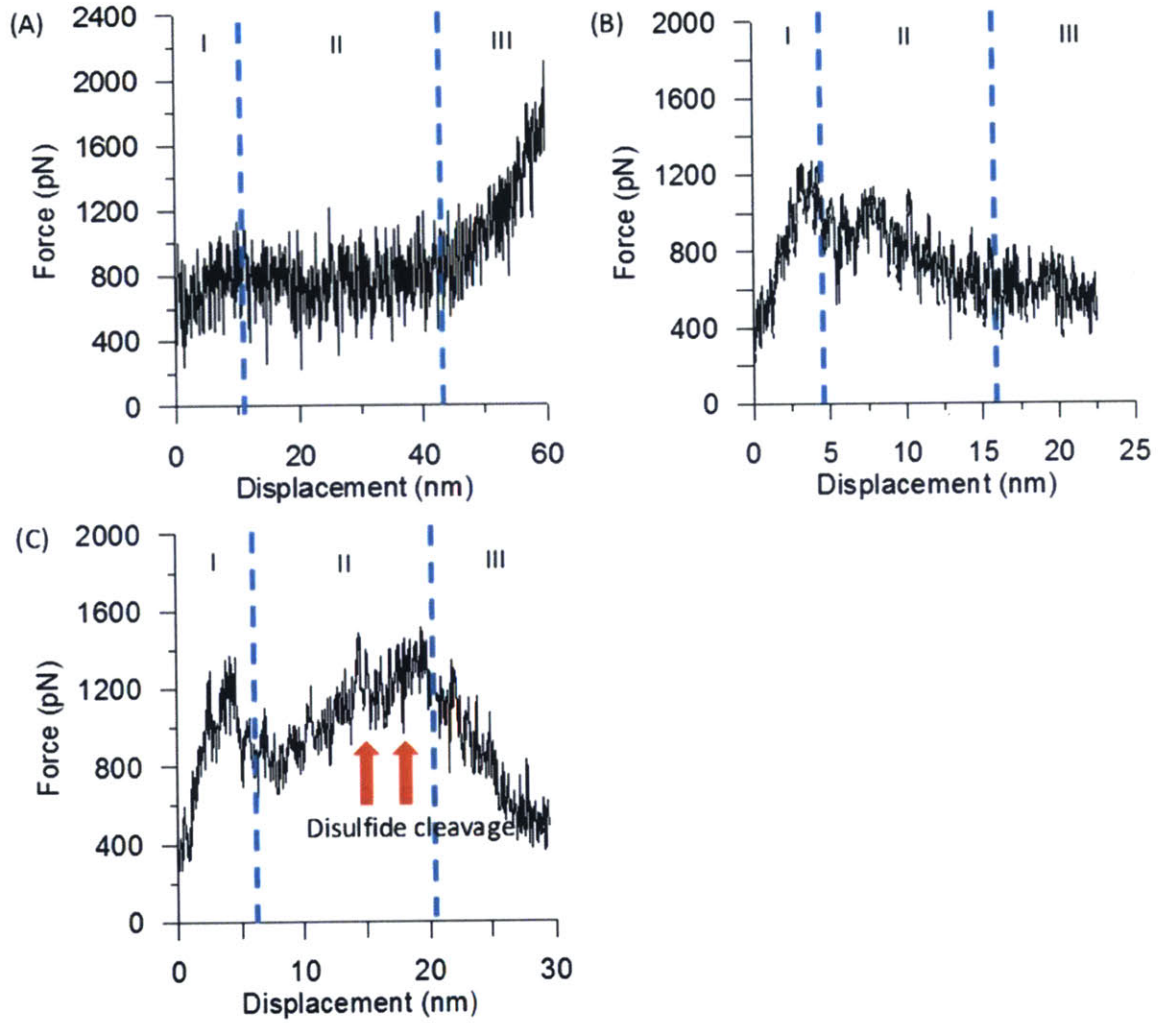

Figure 4-10 Force-displacement relations for tensile testing of (A) heterodimer, truncated tetramer model (B) without and (C) with disulfide crosslink. (A) The force-displacement curve features three regimes: (I), the force increases with displacement until the rupture of the H-bonds and uncoiling of alpha-helices takes place. (II), the dimer starts unfolding in a plateau of a constant force. (III), the force increase dramatically due to stretching the backbone of coiled coils. (B) Force-displacement relations of the model without disulfide bond. The force-displacement curve features three regimes: (I), the force increases with displacement until the rupture of the H-bonds and uncoiling of alpha-helices takes place. (II), dimer sliding and cleavage of H-bond between dimers. (III), failure. (C) In the presence of disulfide crosslink, force-displacement curve features three regimes: (I), the force increases with displacement where the rupture of the H-bonds and uncoiling of

alpha-helices takes place. (II), stretching of backbone of dimer and breaking of two disulfide crosslinks (marked by red arrows) at the displacements of 14.41 nm and 19.44 nm. (III), sliding of the dimers. The maximum forces of the truncated tetramer models without and with disulfide crosslink are 1,250 pN at the displacement of 2.96 nm and 1,510 pN at the displacement of 19.44 nm, respectively. The increment of maximum force due to the disulfide crosslink is greater than 20%. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

Figure 4-10(B) shows the force-displacement behavior of truncated tetramer without disulfide crosslinks. The force-displacement plot shows three distinct regimes. In the first regime (I), the force increases linearly with displacement until the rupture of the H-bonds between coils and uncoiling of alpha-helices takes place. Beyond the peak value, the force decreases and the upper strand starts sliding in the second regime (II). In the third regime (III), the upper strand is pulled out the system and the system fails. In the Figure 4-11(A) and (B), it shows the change of number of H-bonds over the displacement. The number of H-bonds between the heterodimers drops dramatically as displacement increases, showing the inter-dimer H-bonds mediate the tensile deformation and the cleavage of inter-strand H-bonds mitigate the uncoiling of the alpha-helical structure. During the loading, the system remains the most of alpha-helical structure. In the Figure 4-11(B), the number of H-bonds in the heterodimer slightly decreases by $\approx$20%, suggesting that the inter-dimer H-bonds facilitate sliding behavior under stretching and localize damage in the protein fiber.

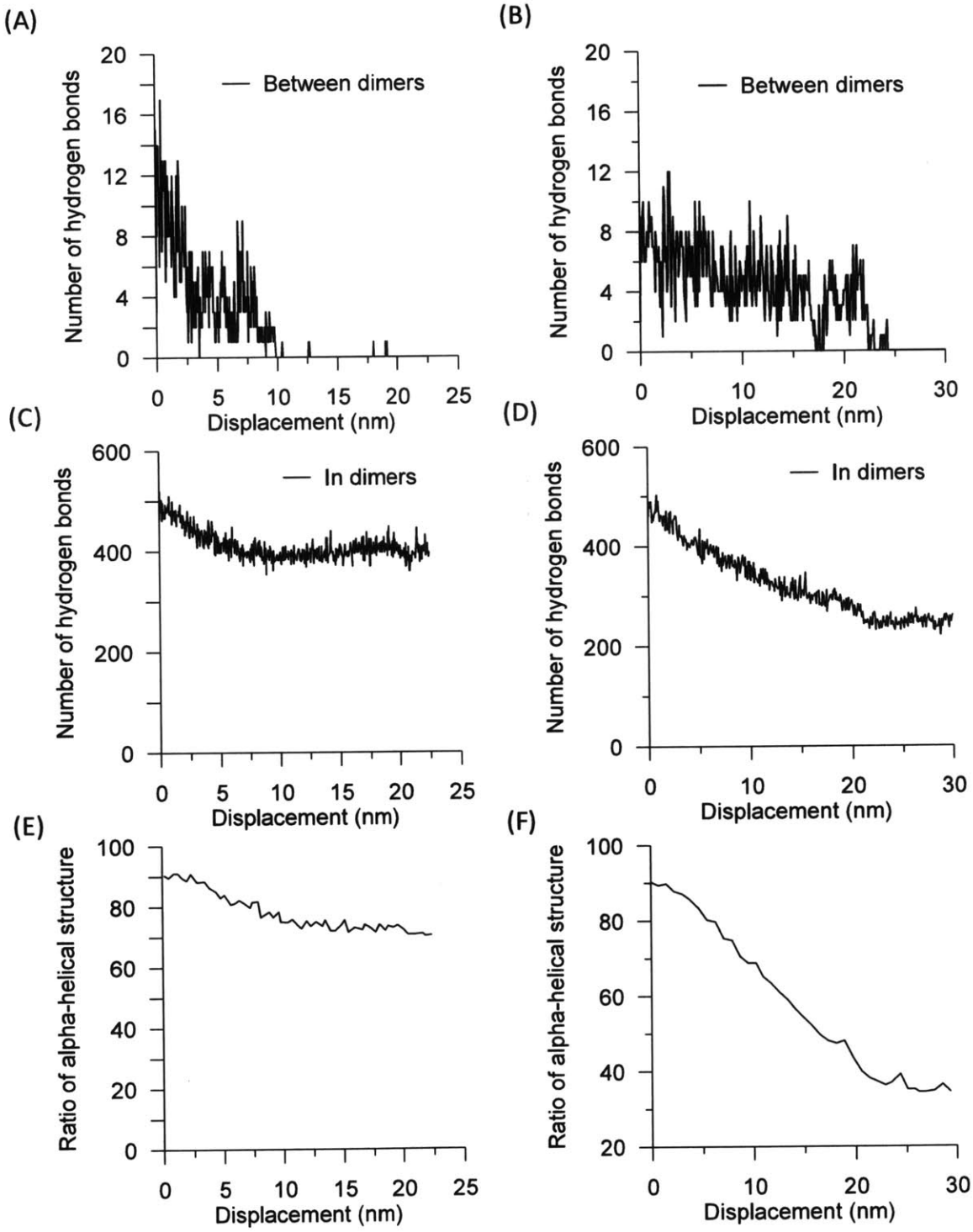

Figure 4-11 Change of the number of H-bonds (hydrogen bonds) and the ratio of alpha-helical structure over displacement of truncated tetramer models (A, C and E) without and (B, D and F) with disulfide crosslinks. Panels (A) and (B) show the number of inter-

molecular H-bonds between dimers, panels (C) and (D) show the number of intra-molecular H-bonds in the dimers and (E) and (F) show the ratio of alpha-helical structure of the dimer. These results imply that in the system without disulfide crosslink, the interdimer H-bonds facilitate sliding under stretching and thereby localize the damage in the system. In the system with disulfide bonds, during loading, the system loses the more than 50% of the alpha-helical secondary structure, and the eventual failure is dominated by the breaking of the disulfide bond. Reprinted with permission from Chou, C. C. and M. J. Buehler. "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules 13(11): 3522-3532. Copyright 2012 American Chemical Society.

In the system with disulfide crosslinks, the simulation also reveals three distinct regimes of deformation as shown in Figure 4-10(C). In the first regime (I), the force increases with displacement where the rupture of the hydrogen bonds and uncoiling of alpha-helices takes place. In the second regime (II), we observe that the pulling force increases from a small load drop to a peak value, marking the onset of the final regime (III). Below the peak value, the stretching of the tetramer leads to expose the disulfide crosslink between two dimers and the breaking of two disulfide crosslink occurs individually in the regime (II) marked by the two arrows at the displacements of 14.41 nm and 19.44 nm. The force drops beyond the peak value in regime (III). The dimers are detached at the displacement larger than 30 nm. The data depicted in Figure 4-11(C) and (D) shows that the number of H-bonds in the heterodimer drops as the displacement increases, suggesting that more than 50% of alpha-helical structure is lost during the loading (Kreplak, Doucet et al. 2001) . The number of inter-dimer H-bonds slightly decreases with the increase of displacement and reaches zero after the rupture of disulfide bonds, implying that the failure behavior is dominated by the disulfide bond. In Figure 4-11(E) and (F), we show the change of the ratio of alpha-helix in the system under external loading for the systems without and with disulfide crosslink. In the absence of the disulfide crosslink, the system loses its alpha-helical structure in the small displacement regime. After the peak value of force, the dimer starts sliding and most of alpha-helix remains. In the system with disulfide crosslink, the loss of secondary structure corresponds to a linear increase of the displacement during stretching. After disulfide

bond breaking, interdimer sliding takes place and the ratio of alpha-helical structure remains at a constant value. The observation in the simulation is in good agreement with earlier experimental data (Kreplak, Doucet et al. 2001).

We continue with a direct comparison of the mechanical properties in these two systems. The maximum forces of the truncated tetramer models without and with disulfide crosslink are 1,250 pN at the displacement of 2.96 nm and 1,510 pN at the displacement of 19.44 nm, respectively. The increment of maximum force due to the disulfide crosslink is more than 20%. We also calculate the toughness modulus by dividing the area under the force-displacement curve until breaking occurs by initial volume of the tetramer. The initial volume is equal to the product of the initial length of truncated tetramer, $L_0 = 17.94$ nm, and the cross-sectional area of the tetramer, $A_0 = 3.80$ nm^2. The toughness moduli are 131.4 pN/nm^3 for the system without disulfide crosslink, and 195.3 pN/nm^3 for the disulfide bonded system, marking a 49% increase. The result shows the disulfide bonded system is stronger and tougher than the one without, where the underlying mechanism is mediated by its ability to resist a greater force and dissipating more energy by breaking the intra-molecular H-bonds under the application of an external load.

We note that the alpha-beta transition is not observed during the mechanical testing in this study (Astbury and Street 1932, Astbury and Woods 1934, Kreplak, Doucet et al. 2001, Kreplak, Doucet et al. 2004). There are two possible explanations: First, a previous fully atomistic study on vimentin protein suggested that there is a strong interaction between the head-coiled-coil overlap regions in the tetramer (Qin, Kreplak et al. 2009), leading to unfolding of the coiled-coil domains, accompanied by beta-sheet formation rather than inter-dimer sliding behavior. However, for the truncated keratin tetramer model, due to the lack of the head-coiled-coil overlap regions, once the inter-dimer sliding begins, the H-bonds between dimers facilitate the sliding behavior and unfolding of the coiled-coil domains is no longer observed (as depicted in the Figure 4-10(B), Figure 4-11(A) and (B)). Secondly, previous computational studies showed that H-bonds in beta-sheets have less stability in solvent due to stronger hydrogen bonding with explicit water molecules in atomistic simulation (Sheu, Yang et al. 2003), implying that perhaps the formation of the beta-sheet structure might not be observed under the

competition of hydrogen bonding with water molecules. For the truncated keratin tetramer with disulfide bonds, because the tetramer is embedded in the explicit water molecules after the coiled-coil domains unfold, the amino acids in the unfolded regions form stable hydrogen bonding with the water molecules rather than other residues. Further studies to address these observations are needed.

4.4 Conclusion

The study reported here presents a first atomistic configuration of the entire keratin heterodimer based on the full amino acid sequence of k35/k85, including the analysis of geometric and mechanical properties of the keratin heterodimer and tetramer by full atomistic simulation. We reported the detailed steps of constructing the keratin heterodimer based on two different algorithms (Lupas, Van Dyke et al. 1991, Lupas 1996, McDonnell, Jiang et al. 2006), used to determine twisting angle of the coiled-coil structure. The geometric properties were validated against experimental data. The force-displacement curve shows that the disulfide bonds are important crosslinks that strengthen and toughen the tetramer. In the systems without disulfide crosslinks, the H-bond between dimers is sacrificed to and the alpha-helical structures are largely maintained, but the system has a lower strength and toughness. The stronger disulfide bond resists higher external load, which results in a loss of the alpha-helical structures and allowing for greater energy dissipation and hence greater toughness. We observe that the loss of secondary structure is increases linearly with the displacement during stretching, in good agreement with experimental data (Kreplak, Franbourg et al. 2002, Kreplak, Doucet et al. 2004). Our analysis suggests the trichocyte alpha-keratin has a higher durability and resistibility against the external loading in the presence of disulfide crosslink. It is interesting to note that the stability of disulfide crosslinks is enhanced under oxidizing conditions, perhaps explaining why keratin materials are so durable under atmospheric conditions with relatively minor degradation over thousands of years. Still, future work is needed to further understand the precise mechanisms by which this material enhances its durability. The availability of an atomistic model allows us to study the mechanical properties at larger scales of keratin fibrils, and perhaps whole fibers using a coarse-grained multi-scale model. The systematic approach of building hair

keratin k35/k85 coiled-coil structure applied here could be directly extended to other disulfide bonded alpha-helical proteins, such as wools and porcupine quills.

5 Mechanics of trichocyte alpha-keratin fibers: Experiment, theory and simulation

(This chapter is adapted from (Chou, Lepore et al. 2015))

The mechanical behavior of human hair is determined by the interaction of trichocyte alpha keratin protein, matrix, and disulfide bonds crosslinking. Much effort has been spent to understand the link between the microscopic structure and the macroscopic fiber properties. Here we apply a mesoscopic coarse-grained model of the keratin macrofilament fibril combined with an analytical solution based on the concept of entropic hyperelasticity of the protein helix to investigate the link between the microscopic structure and the macroscopic properties of keratin fibers. The mesoscopic model provides good agreement with a wide range of experimental results. Based on the mesoscopic model, the predicted stress–strain curve of hair fibers agrees well with our own experimental measurements. The disulfide crosslink between the microfibril–matrix and matrix–matrix contributes to the initial modulus and provides stiffening at larger deformation of the trichocyte keratin fibers. The results show that the disulfide bonds reinforce the macrofilament and enhance the robustness of the macrofilament by facilitating the microfilaments to deform cooperatively. The availability of a mesoscopic model of this protein opens the possibility to further explore the relationship between microscopic chemical structure and macroscopic performance for a bottom-up description of soft materials.

5.1 Introduction

Hair fiber is a hierarchical structure ranging from alpha-helix, coiled-coils, microfibrils, macrofibrils to fibers, as sketched in Figure 5-1(A). The primary structural molecules in wool and hair fibers are keratin intermediate filament (IF) proteins, forming a microfibril

which is built from the assembly of coiled-coils (heterodimers). The IF is embedded in a sulfur-rich protein matrix and assembles into macrofibrils. For keratinized materials, the elastic modulus ranges from approximately 1-4 GPa. Specifically, the modulus is 1.6-4.5 GPa (Oxenham 1989, Chou and Overfelt 2011) in wool, 2.0-3.7 GPa (Fudge and Gosline 2004, Guthold, Liu et al. 2007, Seshadri and Bhushan 2008) in human hair, 1.0-3.5 GPa (Chou and Overfelt 2011) in porcupine quill, and 0.4 GPa (Bertram and Gosline 1987, Chou and Overfelt 2011) in hoof.

Fueled by demand from the textile industry, intensive studies have examined this class of proteins from the 1930s onwards (Speakman 1927, Astbury and Street 1932, Astbury and Woods 1934, Kreplak, Doucet et al. 2001, Kreplak, Doucet et al. 2004), aiming at explaining the mechanical behavior of keratin-based fibers and the link between the structural change and the mechanical properties. Several deformation models were proposed to interpret the shape of stress-strain curves, and to correlate with the fiber structure to explain the mechanical behavior of keratin fibers (Feughelman and Haly 1959, Chapman and Feughelman 1967, Hearle 1967, Hearle 1969, Feughelman 1979, Feughelman 1994, Wortmann and Zahn 1994, Feughelman 1997, Hearle 2000, Feughelman 2002). Earlier deformation mechanism models were proposed by Hearle (Hearle 1967, Hearle 1969, Hearle 2000) and Feughelman (Feughelman and Haly 1959, Feughelman 1979, Feughelman 1994, Feughelman 1997, Feughelman 2002). The deformation models used to describe the mechanical behavior of keratin fiber focus primarily on the geometric change of microfibril/matrix structure. In this earlier work the authors attempted to interpret the shape of curve in relation to the keratin IF protein (coiled coil structure) and an amorphous matrix structure. The major difference between these two models is that the Feughelman model does not take into account any interaction between the coiled coil region and the matrix, while the model proposed by Chapman and Hearle does. Besides these deformation models, previous computational studies of keratin proteins and other α-helical proteins, such as vimentin IF, focusing on the atomistic configuration and nanomechanical properties at atomistic scale have been reported (Akkermans and Warren 2004, Qin, Kreplak et al. 2009, Qin and Buehler 2011, Azoia, Fernandes et al. 2012, Chou and Buehler 2012). However, relatively little is known of the physical mechanisms that drive its deformation behavior, thus presenting an opportunity

to generate a new approach that considers the structure-property paradigm from the atomistic level to the macroscopic scale. In order to provide a bottom-up description of materials behavior from a fundamental perspective, here we apply a multi-scale simulation approach that links the molecule's chemical structure (including the abundance of disulfide crosslinks) and its larger-scale properties. Using the multi-scale simulation, we are able to upscale the atomistic simulations based on the architecture of materials to exploit insights that are relevant for bioengineering and other applications.

In this chapter, we aim to build a mesoscopic simulation model of keratin macrofilaments and to link the micro-scale structural changes with the stress-strain curves of large-scale keratin fibers. We demonstrate that the deformation prediction from the mesoscopic model is in good agreement with the stress-strain curve in experiments, and it provides insights and a general methodology to assess the mechanics of other biological systems.

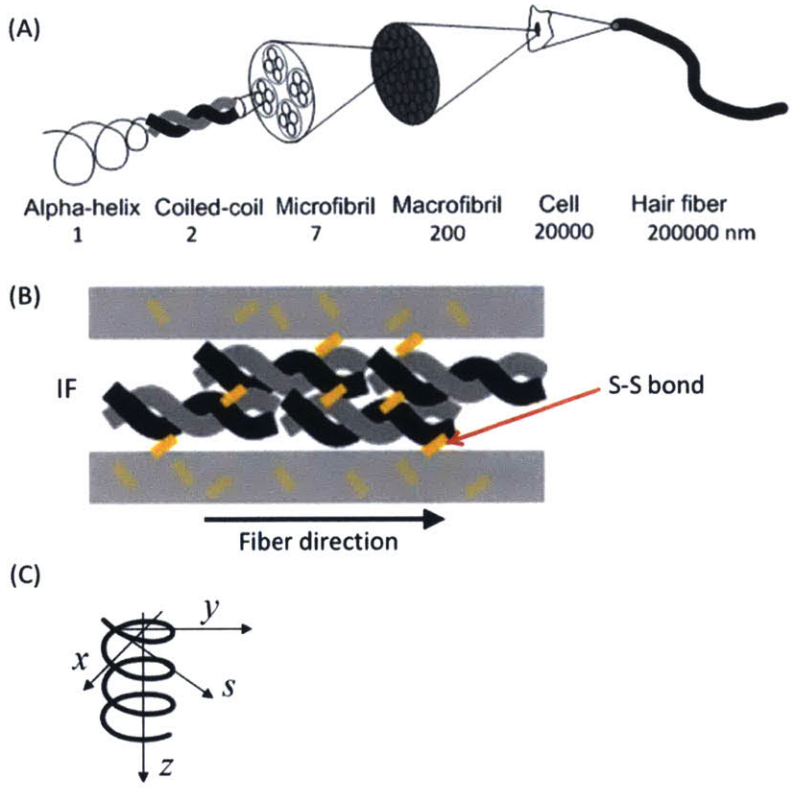

Figure 5-1 (A) Human hair features a hierarchical structure, ranging from alpha-helix, dimers with a coiled-coil structure, microfibrils, macrofibrils to the cellular structure and eventually entire hair fiber level. (B) Schematic visualization that shows IFs are

embedded in the matrix connected by intramolecular disulfide bonds (S-S bond in the figure). (C) Geometry of the coil and reference system. Reproduced with permission Chou, C.-C., E. Lepore, P. Antonaci, N. Pugno and M. J. Buehler (2015). "Mechanics of trichocyte alpha-keratin fibers: Experiment, theory, and simulation." Journal of Materials Research 30(01): 26-35.

5.2 Simulation details

5.2.1 Mesoscale modeling

In this study, a mesoscopic "bead-spring" method is used to investigate the mechanics of trichocyte keratin macrofilament with disulfide crosslinks. The method is a coarse-grained description of alpha-helical protein and matrix structure. This has proven to be a suitable approach to simulate certain aspects of alpha-helix based proteins, such as intermediate filaments in the cell's cytoskeleton (Ackbarow, Sen et al. 2009, Qin and Buehler 2012) and other fibril protein systems, such as collagen fibril with mineral and various crosslink densities (Buehler 2006, Buehler 2007, Buehler 2008). In the mesoscale model, each bead represents clusters of amino acids in explicit solvent, and all of beads interact according to a specific intermolecular multi-body potential to reflect the physical behavior of keratin proteins. The total energy of the interaction of beads in the system, E_{tot}, is expressed as,

$$E_{tot} = E_{bond} + E_{angle} + E_{non-bonded} \tag{5-1}$$

where E_{bond} is the bonded interaction describing the stretching behavior, E_{angle} is the bending energy, and $E_{non-bonded}$ is the nonbonded energy due to the van der Waals interactions. Within each intermediate filament, a multi-linear model provides a nonlinear force-extension behavior of intermediate filament under stretching as

$$F(r) = -\frac{\partial E_{bond}(r)}{\partial r}$$

$$= \begin{cases} k_1(r - r_0) & r < r_1 \\ k_1(r_1 - r_0) + k_2(r - r_1) & r_1 < r < r_2 \\ k_1(r_1 - r_0) + k_2(r_2 - r_1) + k_3(r - r_2) & r_2 < r < r_3 \\ k_1(r_1 - r_0) + k_2(r_2 - r_1) + k_3(r_3 - r_2) + k_4(r - r_4) & r_3 < r < r_4 \\ 0, & r_4 < r \end{cases} \quad (5\text{-}2)$$

where $F(r)$ is the force between two beads, and k_1, k_2, k_3, k_4 are four spring constants which represent the stiffness of molecular stretching at different regimes as shown in Figure 5-2 (A). The force-displacement relationship in eq. (2) is fitted to reproduce the nanomechanical behavior obtained using the full atomistic model of keratin tetramer (Chou and Buehler 2012) which has been well validated against experimental measurements. $E_{angle} = \frac{1}{2}k_\theta(\theta - \theta_0)^2$ with $k_\theta = \frac{3EI}{r_0}$, where EI is the bending stiffness, and θ is the angle between three beads relative to the equilibrium angle, θ_0. The bending stiffness, EI, can be obtained through the relationship of $EI = L_p k_B T$, where L_p is the intermediate filament persistence length, k_B is the Boltzmann constant and T is the temperature. The bending stiffness, EI, is obtained from bending deformation calculations of alpha-helical molecules in the previous publications (Ackbarow and Buehler 2007, Buehler 2007). The nonbonded interaction, $E_{non-bonded}$, is described by Lennard–Jones potential to simulate the van der Waals interactions between the filament chains and matrix-filament, $\phi_{LJ} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$, where ε is the energy at equilibrium and σ is the distance as $\phi_{LJ} = 0$. A harmonic/shift/cut function in LAMMPS (Plimpton 1995) with spring stiffness of 7.35 kcal/mol/Å^2 is applied to model the disulfide crosslinking. When the bond deformation is larger than 4 Å, the disulfide crosslinking breaks and the force drops to zero. The parameters are based on the full atomistic simulation results using ReaxFF reactive force field (Keten, Chou et al. 2012). The time step used in our coarse-grained model is 20 fs. The simulations are performed using the large-scale atomic/molecular massively parallel simulator (LAMMPS)

(Plimpton 1995). The mesoscopic parameters of the coarse-grained trichocyte keratin macrofilament model are summarized in Table 5-1.

Table 5-1 Summary of the parameters used in the mesoscopic model

Parameter and units	Numerical value
Equilibrium bead distance r_0 (Å)	28
Critical distances r_1, r_2, r_3, r_4 (Å)	30.257, 45.96, 57.93, 75.6
Tensile stiffness parameters k_1 , k_2 , k_3 , k_4 (kcal/mol/Å^2)	4.49, 0.0, 0.46, -0.86
Equilibrium angle θ_0 (°)	180
Bending stiffness parameter k_θ (kcal/mol/rad^2)	3.44
Equilibrium energy ε (kcal/mol)	6.8
Equilibrium distance σ (Å)	25

In the coarse-grained model, the full length sequence of tetramer is replaced by a chains of beads, and eight chains of the tetramer model are placed in a cylinder with diameter 7.0 nm to form a microfilament according the measurement using X-ray scattering on human hair (Kreplak, Franbourg et al. 2002, Kajiura, Watanabe et al. 2006). The filaments are placed on a hexagonal lattice packing based on the observation in experiments shown in Figure 5-2 (B) (Bruce Fraser and Parry 2003). The distance between the center of mass of one filament and its nearest neighbor is determined by the experimental mean filament-filament distance 9.0 nm (Kreplak, Franbourg et al. 2002,

Kajiura, Watanabe et al. 2006). Assuming the eight chains are close packed on the cylinder, the diameter of one filament particle is roughly equal to the one-eighth of the circumference of the cylinder. Since a keratin tetramer is ~50 nm long and has a molecular weight of 200 kg/mol (around 2000 amino acids), one filament bead represents about 60 amino acids with a molecular weight of 12 kg/mol. The diameter of the matrix is determined using the separation between two filaments in order to place at least one particle between the filaments. Based on all of the structural parameters and the molecular weight fraction of microfibrils (43 %) and matrix (28 %) in experiments (Zahn 2002), we calculate the ratio of the number of filament and matrix beads equal to 12. Finally, the coarse-grained model containing the microfilaments embedded in the matrix is shown in Figure 5-2 (C, D).

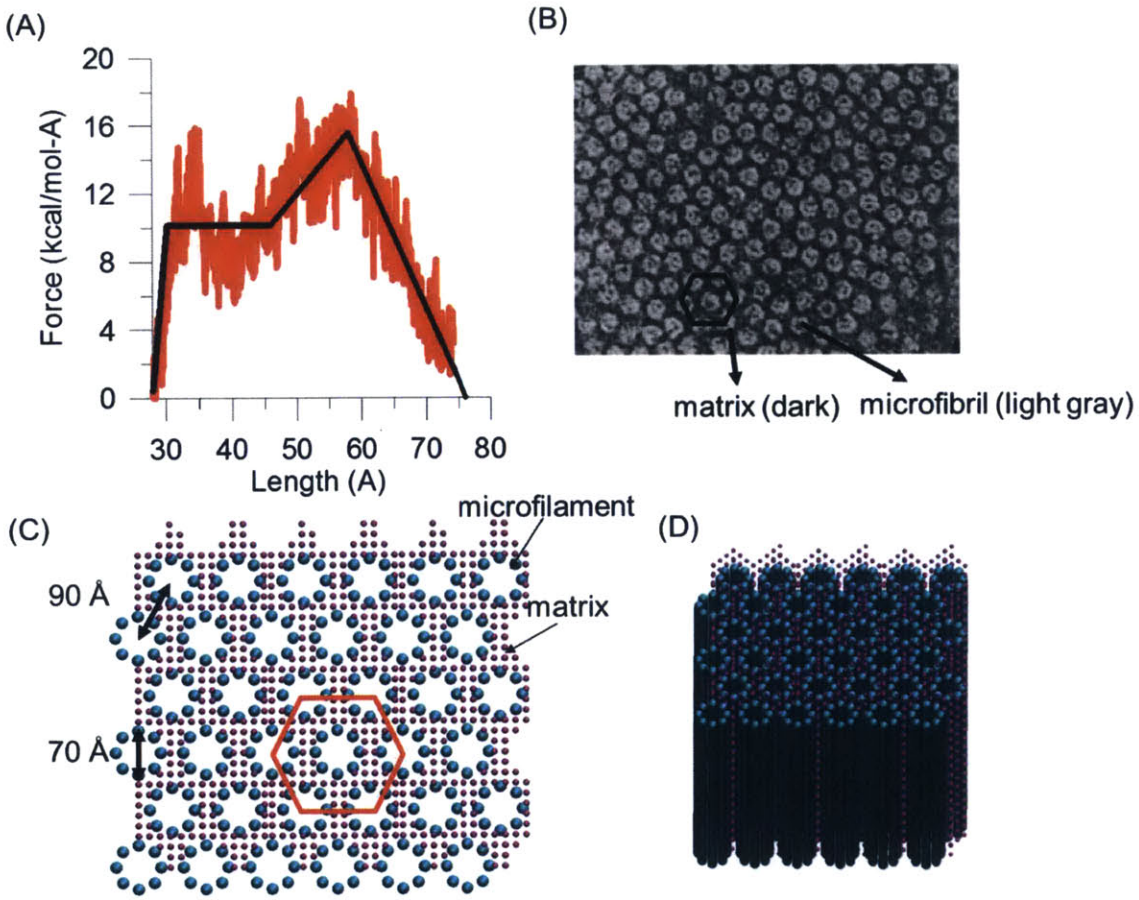

Figure 5-2 (A) The force-strain relationship of individual keratin microfilament in our bead-spring model is derived from full-atomistic simulations (Chou and Buehler 2012)

and has been validated against experimental studies. (B) Electron micrograph of the central portion of macrofibril in the cortex of fine wool depicts the hexagonal microfibril packing in macrofibril (Bruce Fraser and Parry 2003). Reprinted from Journal of Structural Biology, Vol. 142(2), R.D. Bruce Fraser and D.A.D. Parry, Macrofibril assembly in trichocyte (hard α-) keratins, Copyright (2003), with permission from Elsevier. (C, D) Coarse-grained model of keratin macrofibril. The filaments are placed on a hexagonal lattice. Reproduced with permission Chou, C.-C., E. Lepore, P. Antonaci, N. Pugno and M. J. Buehler (2015). "Mechanics of trichocyte alpha-keratin fibers: Experiment, theory, and simulation." Journal of Materials Research 30(01): 26-35.

5.2.2 Mechanical testing

Pulling is applied in fiber-direction to model a tensile loading, as indicated in Figure 5-1(B). We stretch the systems by deforming the simulation box length in the fiber-direction. In this study, we choose a small strain rate of 0.01 m/s which is computationally feasible, and no major rate effect on the system is observed. The simulations are carried out at 300 K in a NVT ensemble (constant temperature, constant volume, and constant number of particles). During the simulation, the pressure in fiber-direction and the box length are monitored to measure the stress-strain curve. The strain is defined by $\varepsilon = \Delta L/L$ (= engineering strain), where ΔL is the applied elongation and L is the length of the system in the pulling direction. The initial elastic modulus (Young's modulus) is calculated as the steepest slope of the stress-strain curve in the initial linear region.

5.2.3 Entropic hyperelasticity of helix

Hyperelasticity

Here, we use an entropic hyper-elasticity of helix model to describe the behaviour of alpha-helix domain in hair. The coil geometry shown in Figure 5-1(C) is described by the following equations:

$$x = r\cos 2\pi\frac{s}{l} \tag{5-3}$$

$$y = r\sin 2\pi\frac{s}{l} \tag{5-4}$$

$$z = \lambda\frac{s}{l} \tag{5-5}$$

where r is the radius of the coil, s is the curvilinear coordinate, l is the length of a ring and λ is the distance between two adjacent rings. The coil angle is defined by $\tan\alpha = \frac{\lambda}{2\pi r}$.

If two self-equilibrated forces F are applied along z at the ends of the coil, its elastic energy per unit length will be:

$$\frac{dE_L}{ds} = \frac{1}{2}\left(\frac{N^2}{EA} + \frac{M^2}{GI_p}\right) \tag{5-6}$$

where N is the axial load, M is the twisting moment, E and G are the material Young and shear modulus, A is the cross-section area and I_p is the torsional (e.g. polar, for circular cross- sections) moment of inertia of the fiber. It is worth noting that $M = Fr$, $N = F\sin\alpha$, we can easily derive the elastic energy in the coil composed by m rings, as:

$$E_L = \frac{m}{2}\frac{F^2}{EA}\left(\sin^2\alpha + 2(1+\nu)\frac{Ar^2}{I_p}\right)l \tag{5-7}$$

where v is the Poisson's ratio. The equivalent (in term of stored energy) elastic stiffness can be derived by comparison with:

$$E_L = \frac{1}{2}\frac{F^2}{K_E} \tag{5-8}$$

yielding

$$K_E = \frac{EA/ml}{\sin^2\alpha + 2(1+\nu)\frac{Ar^2}{I_p}} \tag{5-9}$$

Note that $\frac{Ar^2}{I_p}$ describes the slenderness of a ring and EA/l its axial stiffness.

Now we consider the actual values of r and α as imposed by a nominal strain $\varepsilon = \frac{\lambda-\lambda_0}{\lambda_0}$, where subscript 0 refers to the unstrained initial configuration. The strained coil geometry, described by the two functions $\alpha(\varepsilon)$ and $r(\varepsilon)$, can be deduced by imposing the inextensibility of the coil (i.e. neglecting the axial compliance with respect to the torsional one), i.e. imposing $l = l_0$, from which we obtain:

$$\sin\alpha(\varepsilon) = \sin\alpha_0\,(1+\varepsilon) \tag{5-10}$$

then consequently

$$r(\varepsilon) = r_0\frac{\cos\alpha(\varepsilon)}{\cos\alpha_0} = r_0\sqrt{1 - \frac{\varepsilon^2 + 2\varepsilon}{\frac{1}{\sin^2\alpha_0} - 1}} \tag{5-11}$$

The coil will reach the straight configuration ($\sin\alpha(\varepsilon^*) = 1$) under a critical strain:

$$\varepsilon^* = \frac{l-\lambda_0}{\lambda_0} = \frac{1}{\sin\alpha_0} - 1 \tag{5-12}$$

Thus, in general, the hyper-elastic stiffness of a coil of contour length ml per cross-section area can be described by:

$$K_{HE}(\varepsilon) = K_E A\lambda_0 = \begin{cases} \dfrac{E\lambda_0/ml}{\sin^2\alpha_0\,(1+\varepsilon)^2 + 2(1+\nu)\dfrac{A}{I_p}r_0^2\left(1-\dfrac{\varepsilon^2+2\varepsilon}{\dfrac{1}{\sin^2\alpha_0}-1}\right)}, & \varepsilon < \varepsilon^* \\ \dfrac{E\lambda_0}{ml}, & \varepsilon \geq \varepsilon^* \end{cases} \quad (5\text{-}13)$$

Hydrogen bonds (H-bonds)

We describe a stiffness of a linear elastic spring with a bundle of H-bonds between two adjacent coils as K_{HB}. We estimate the value of K_{HB} from the MD simulation results (Ackbarow and Buehler 2008, Ackbarow, Keten et al. 2009), where the 3.6 H-bonds in one convolution break simultaneously, and the corresponding energy barrier and the distance between the equilibrated state and the transition state are $E_{HB} = 11.1$ kcal/mol and $\Delta b_{HB} = 1.2$ Å. Thus K_{HB} can be obtained by

$$K_{HB} = \frac{(E_{HB}/\Delta b_{HB})N}{A} \quad (5\text{-}14)$$

where N is the number alpha-helices in an unit area A. We apply the cross-section area ~10 nm^2 (Sokolova, Kreplak et al. 2006) of a tetramer that consists of 4 alpha-helices, and we obtain $K_{HB} = 254.8$ MPa.

It is worth noting here that for alpha-helical structure, the previous studies have shown that the hydrophobic interactions between helices have relatively small contribution (Qin, Fabre et al. 2013). Also, in trichocyte keratins, the full atomistic model showed that the disulfide bonds hinder the sliding behavior between coils (Chou and Buehler 2012) so there is little contribution of interaction between helices to the force. Thus, the interaction

between helices including hydrophobic interaction is not considered in our theoretical model.

5.3 Results and Discussion

In this study, we consider two different approaches to identify how the trichocyte keratin fiber with disulfide crosslink responds to mechanical deformation: a mesoscopic model of keratin macrofilament with disulfide crosslink and a force-displacement relationship based on the helix-like theoretical model.

5.3.1 Mesoscopic model of keratin macrofilament with disulfide crosslink

In order to understand how the trichocyte keratin materials response under an external force, it is important to develop a model which reflects the composite structure of keratin materials and allows us to probe the effect of disulfide crosslinks. Since the experiments suggests that macrofilament serves the major component of trichocyte keratin (Zahn 2002), we build a mesoscopic model of the macrofilament consisting of microfilaments and matrix with disulfide crosslinks inside the matrix and in between the matrix and the microfilament as shown in Figure 5-2(C, D). Details about the model formulation are included in the ''Materials and Methods'' section.

We proceed with stretching the systems by deform simulation box length in fiber-direction and measure the stress-strain response of this material, until failure occurs. Figure 5-3**(A)** depicts stress-strain curves of the mesoscopic macrofilament model. We observe three major regimes in the stress-strain response. In the first regime (I), the stress increases linearly under a small deformation (strain < 7%). In the second regime (II), a relative flat increase in stress from 7 to 17 % strain, followed by an increasing stiffness of the stress, which lasts up to the stress close to 80 MPa (III). Eventually, strong bonds in the microfilaments chains break, and the entire system fails at 30 % strain. The phenomenon of the increase of the stress in the regime (III) is referred to as strain hardening.

The stress-strain curves of the mesoscopic simulation features the similar characteristic shape as the experimental measurements. From the analysis of the stress-strain behavior, the Young's modulus = 1.38 GPa is calculated as the steepest slope of the stress-strain curve in the region (I). This modulus is close to experimental results for human hair, where a range of 1.1 to 6 GPa (our experiments) was reported (Greenberg and Fudge 2013). The maximum strain for the mesoscpic filament model is predicted to be 30% (Figure 5-3(A)) which is close to experimental measurements of ~33 %. We note that the maximum stress, 80MPa, of our mesoscopic model is almost twice smaller than the experimental data. The discrepancies between the experiments and the simulation might be caused by the fact that whole hair samples are considered in experiment, whereas a part of macrofilament system without any structural flaws is modeled in our simulations.

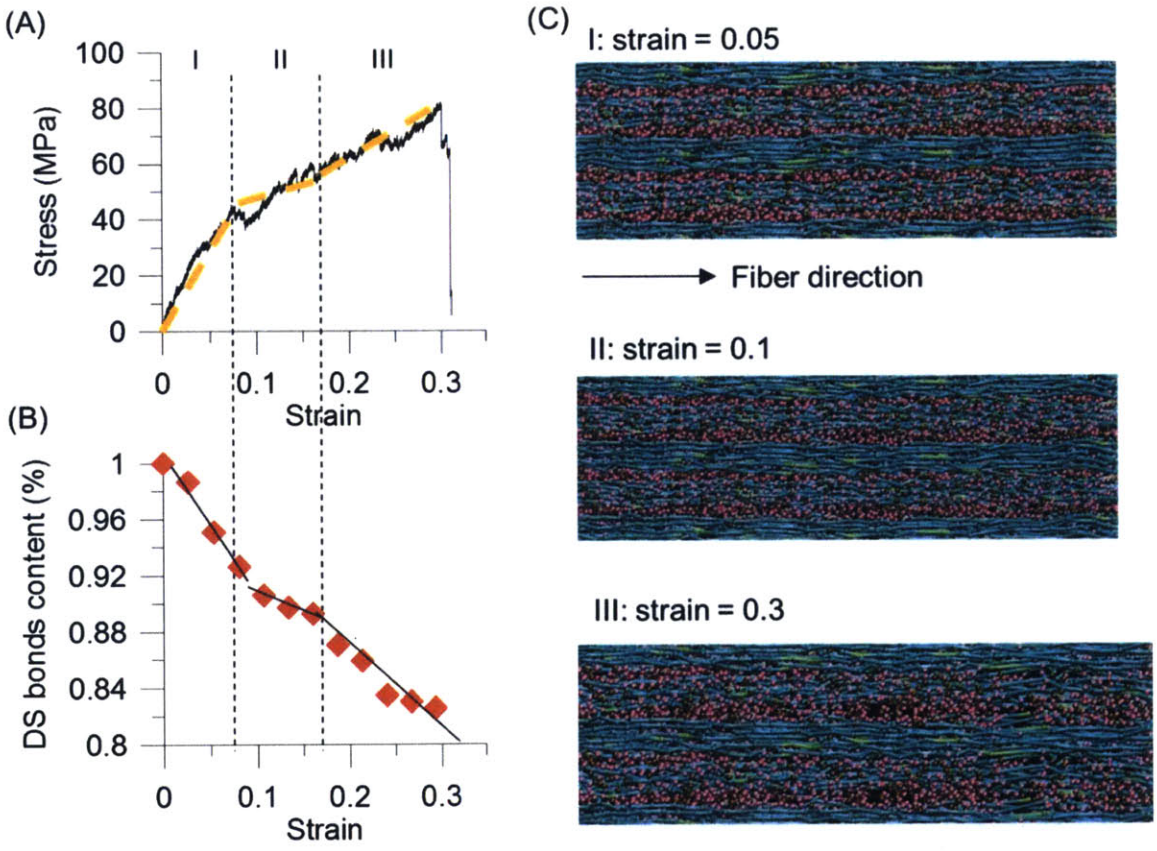

Figure 5-3 (A) The stress-strain curves of coarse-grained simulation. (B) Decrease of disulfide crosslink content, normalized by the number of disulfide bond in the initial model. (C) Simulation snapshots at different pulling strains. Reproduced with permission

Chou, C.-C., E. Lepore, P. Antonaci, N. Pugno and M. J. Buehler (2015). "Mechanics of trichocyte alpha-keratin fibers: Experiment, theory, and simulation." Journal of Materials Research 30(01): 26-35.

We further analyze the disulfide crosslink responses during the pulling. Figure 5-3(B) depicts the decrease of disulfide crosslink content which is normalized by the number of disulfide bond in the initial model plotted as a function of increasing applied strain. This plot also exhibits three different regimes during mechanical strain. The rupture of disulfide bonds occurs from the start of straining (0 to 7%). This shows t the rupture of the linkages between the matrix and the IFs, and the results further support Chapman and Hearle's model that incorporates a link between the matrix and the IFs. In the second regime, the curve shows a small decrease in the crosslink content until ~17% strain, followed by a rapid decrease again until the system fractures. This phenomenon is consistent to the experimental measurement using Raman spectra (Paquin and Colomban 2007). Figure 5-3(C) shows the simulation snapshots of the macrofilament at different applied strains. Under strain, the microfilaments deform cooperatively, and the chains start aligning along the fiber direction. At the larger strain, once a bundle of the microfilaments breaks, then the system fails. The coarse-grained model is developed from a bottom-up approach and parameterized from the previous full atomistic study (Chou and Buehler 2012) so we expect that the stress-strain curve of the coarse-grained simulation and the experiments might not perfectly match. However, despite its simplicity, our model captures the essential physical properties of alpha-helical based keratin macrofilaments as identified in earlier theoretical and experimental studies. Through simulation of a larger-scale filament and matrix structure, our model enables us to provide an important link between single molecule properties and mechanisms and the overall material behavior at much larger length-scales.

5.3.2 Theoretical model: helix-like hyper-elastic stiffness model

We proceed with the study of mathematical equation of hyper-elastic stiffness coil model linking the structural change and stress-strain curve of keratin fiber through a spring model. Under small deformation, the stretch of H-bonds in convolutions of alpha-helical

structure in keratin proteins plays the major role. The stiffness of an alpha-helix can be calculated as (H-bonds in parallel to the elasticity of a coil as plotted in Figure 5-4(A))

$$K = K_{HB} + K_{HE} \tag{5-15}$$

Once the deformation of H-bonds elongates to a critical points Δb_{HB} which is the critical displacement for H-bond fracture, the unfolding of alpha-helix occurs. We consider a chain with N units of folded structure involved in H-bond fracture and assume the ratio of units of folded structure is $\frac{Nx_\alpha}{L\,(N)} = \frac{Nx_\alpha}{(L_0+Nx_\alpha)} = 0.15$, where $L\,(N)$ is the total length of the chain, L_0 is the length of chains not involved in H-bond fracture and x_α is the unit length of helix. Thus due to $\frac{\Delta b_{HB}}{x_\alpha} = \frac{1.2}{5.4} = 0.22$, the critical strain for the rupture of H-bonds $\varepsilon_1 = \frac{N\Delta b_{HB}}{(L_0+Nx_\alpha)} = 0.03$ can be obtained. Beyond the critical point ε_1, the helical structures start to unfold, and the regime is dominated by the hyper-elasticity mechanics until the coil reaches the straight configuration where the backbone of the protein starts to be stretched at the strain ε^* in (E13). It is known that the ratio of the unit length of helix (x_α) to its contour length (λ) is about 0.5, i.e. $\varepsilon^* = \frac{\lambda - x_\alpha}{x_\alpha} \approx 1.0$; based on the previous atomistic simulation(Qin, Kreplak et al. 2009), assuming the ratio of the length of helix structure (Mx_α) involved in the unfolding $\frac{Mx_\alpha}{L\,(N)} = \frac{Mx_\alpha}{(L_1+Mx_\alpha)} \approx 0.4$, where L_1 is the length of chains not involved in unfolding, ~~thus~~ the critical strain for the alpha-helix unfolding is $\varepsilon_2 = \frac{M(\lambda - x_\alpha)}{(L_1+Mx_\alpha)} = 0.4$. After the hyper-elasticity regime where $\varepsilon < \varepsilon_2$, the intrinsic axial stiffness of the fiber, $K_{HE} = \frac{E\lambda_0}{ml}$, dominates up to fracture.

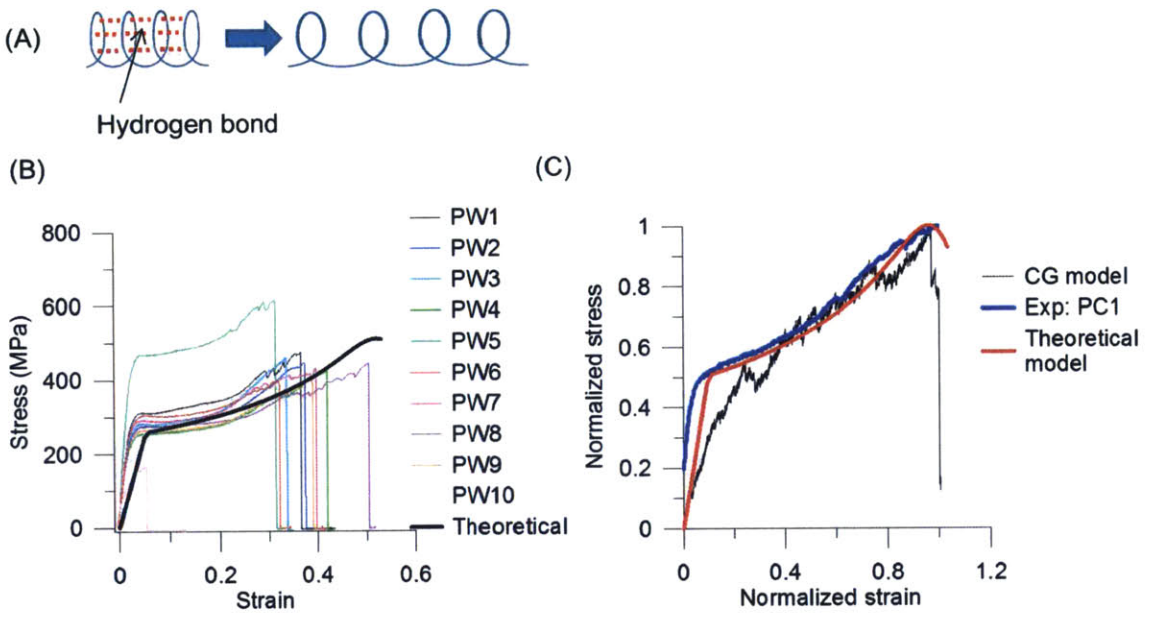

Figure 5-4 (A) Illustration of H-bonds in parallel in coils and unfolding of the helical structures (B) Comparison of stress-strain curves of experimental data and theoretical prediction based on the stiffness in Equation (5-18), where the parameters in binomial distribution are obtained by fitting the failure stress, failure strain and the critical point at onset of third regime in stress-strain curves of keratin fibers. (C) Comparison of normalized stress-strain curves of coarse-grained simulation, theoretical prediction and experimental data. The stress and strain are normalized to the maximum values in each data set, individually, to compare the shape of stress-strain curves and the mechanical response in tensile testing. Reproduced with permission Chou, C.-C., E. Lepore, P. Antonaci, N. Pugno and M. J. Buehler (2015). "Mechanics of trichocyte alpha-keratin fibers: Experiment, theory, and simulation." Journal of Materials Research 30(01): 26-35.

Here we further apply a binomial distribution to consider the effect of disulfide crosslinks in the system. We assume that the number of breaking disulfide bonds is k, each with a bond breaking probability p. The total number of disulfide bonds is n so that the probability of k bonds breaking in a pulling event is given by a binomial distribution:

$$b = \binom{n}{k} p^k (1-p)^{n-k} \tag{5-16}$$

Moreover, we assume that after a certain number x of disulfide bond are broken, the system fails. The probability that the total number X of disulfide bond breaking, with a given probability distribution b, during the entire pulling process will exceed the tolerance x is given by

$$B = 1 - \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i} \tag{5-17}$$

The parameters used in the binomial distribution are obtained by fitting the failure stress, failure strain and the critical point at the onset of the third regime in the experimental stress-strain curves of keratin fibers. The fitted binomial distribution mainly affects the third regime of the stress-strain curve but has little effect below the critical point at the onset of the third regime. Thus, considering the breaking of disulfide bond under stretching, the stiffness of an alpha-helical chain of can be described by

$$K = B(K_{HE} + K_{HB}) \tag{5-18}$$

Based on the stiffness in Equation (18), the stress-strain relationship is given by $\sigma = K\varepsilon$, where σ is the stress and ε is the engineering strain. Figure 5-4(B) depicts the comparison of the stress-strain curve of the theoretical model and experiments. The result shows that the theoretical data agrees well with experiments as the value of the stress–strain curves of the theoretical data is within the range of the experimental data set. The availability of a theoretical model of the helix provides a mathematical equation as a basis to explore the relationship between structure, mechanical properties and performance from a theoretical perspective. The good agreement between the results of experimental data and theoretical formulation suggests that the behavior of helix domain plays a dominant role in hair as stated in literatures(Feughelman 1997, Hearle 2000).We also compare the mesoscale simulation results with experimental data and theoretical prediction plotted in Figure 5-4(C). The stress and strain are normalized to the maximum values in each data set, individually, to compare the shape of stress-strain curves and the

mechanical response in tensile testing. The comparison in Figure 5-4(C) shows that the stress-strain curves have similar shape which features three regimes. In the first regime (I), both stresses increase linearly with strain. In the second (II), a stress plateau relates to the unfolding of the coiled-coil segments and alpha-helical structure. For the third (III) regime, the stretching of the backbone of the alpha-helix with the breaking of disulfide crosslinks causes the strain stiffening of the material, and the system fails.

5.4 Summary

The study reported here presented a mesoscopic coarse-grained model to link the microscopic structure of trichocyte keratin fibers to the mechanical properties at larger scales. The composite structure of keratin fiber at the mesoscale is addressed using a bead-spring model for a keratin macrofilament. Mesoscale beads represent coil-coiled keratin tretramer proteins which are assembled to filaments and embedded in a matrix of soft particles with disulfide crosslinks. The mesoscopic model gives good agreement with a wide range of experimental results. The stress-strain curve of hair fibers predicted by the mesoscopic model agrees well with our experimental data. The disulfide crosslink the microfibril-matrix and matrix-matrix contributes the initial modulus and provides stiffening at larger deformation of trichocyte keratins. The results show that the disulfide bonds reinforce the macrofilament and enhance its robustness by facilitating the microfilaments to deform cooperatively. The availability of the coarse-grained multi-scale model now allows us to study the mechanical properties at larger scales of keratin fibrils, and perhaps entire hair fibers. The mesoscopic model of keratin fiber built here could be directly extended to other soft fibrous-matrix materials, such as porcupine quills.

6 Ionic effect on metal-coordination and nanostructure of Nvjp-1 protein in the Nereis jaw

Nvjp-1 protein is the key component of the Nereis jaw and contains over 25 mol% histidine, which is believed to play a role in the metal-coordination crosslinks between the Nvjp-1 and metal ions, resulting in the structural stability and the exceptional mechanical properties of the Nereis jaw. In order to investigate this intriguing mechanism of metal-coordination crosslinks in Nvjp-1 protein, a molecular dynamics simulation is used to build a bottom-up molecular based model of Nvjp-1 protein. A detailed analysis of metal coordination and nanostructure properties of the Nvjp-1 at different ion concentrations using replica exchange molecular dynamics simulations is presented. The simulations show that the solvent accessible surface area and the radius gyration of Nvjp-1 decrease at higher ionic concentrations. Nvjp-1 forms a more compact structure in the presence of Zn ions, and more stable intra-molecular metal coordination complexes are formed at higher ion concentrations. The results are in good agreement with experimental observations. The study suggests that the metal-coordination crosslinks induce significant Nvjp-1 aggregation and achieve the contraction of the Nvjp-1 stripe observed in experiments.

6.1 Introduction

Nvjp-1 protein is the key component of Nereis jaw which is composed of roughly 90% protein, halogens (~8%) and Zn ions (~2%) (Bryan and Gibbs 1979, Lichtenegger, Schoberl et al. 2003, Waite and Broomell 2012). Surprisingly, although most of the jaw is composed of soft materials, the hardness and stiffness of the jaws are comparable to that of human dentin (~1-2 GPa of hardness and 10-20 GPa of stiffness). The Nvjp-1 protein contains over 25 mol% histidine, which is a kind of amino acid that often forms the metal-coordination bonds with surrounding metal ions such as Cu^{2+} or Zn^{2+}. Studies show

that the formation of the metal-coordination crosslinks between histidine-rich Nvjp-1 protein and Zn ions dominates the structural stability and the marvelous mechanical properties of Nereis jaw (Bryan and Gibbs 1979, Lichtenegger, Schoberl et al. 2003, Birkedal, Khan et al. 2006, Broomell, Mattoni et al. 2006, Broomell, Chase et al. 2008, Broomell, Zok et al. 2008, Waite and Broomell 2012).

The metal-coordination bond occurs when a species (ligand) donates its lone pair of electrons to a metal ion that is usually a transition metal. The metal-coordination bond serves as a flexible connection due to its variable bond strength in different chemical microenvironments (Holten-Andersen, Harrington et al. 2011, Fullenkamp, He et al. 2013). The strength of the metal-coordination bond decreases in low pH conditions, and the bond strength increases as the pH value increases. Moreover, unlike most covalent bonds whose rupture is usually irreversible, the metal-coordination bonds are able to reform after rupture. With the mutable strength of the metal-coordination bond in different chemical microenvironments, the Nvjp-1 protein has also shown intriguing mutability such as the ionic effect on the aggregation and hydrodynamics properties of the Nvjp-1 protein (Broomell, Chase et al. 2008). In addition, recent studies show the effect of the metal-coordination bonds on optical properties of materials, for example, the material's visual appearance (Fullenkamp, He et al. 2013). The mechanism of color change of the materials could be applied to identify different metal-coordination bonds in nature.

However, although prominent in the Nereis jaw, relatively little is known about the mechanism of mutability, particularly how the proteins behave in various environments. An atomistic model is critical in order to probe the properties of the Nvjp-1 proteins and to exploit the insight for bioengineering and other applications. Such a model is needed to address the origin of this material's mutability by assessing a direct link between the molecule's chemical structure (including the abundance of metal-coordination crosslinks) and its larger-scale properties. It will also benefit the current studies on jaw-inspired materials to improve their design (Srivastava, Holten-Andersen et al. 2008, Holten-Andersen, Harrington et al. 2011).

In this study, we focus on the study of the effect of the Zn ions on the structure and mutability of the Nvjp-1 protein and investigate the role of the metal-coordination crosslinks in Nvjp-1 proteins using atomistic simulation. This study provides fundamental insight into mutable mechanical properties of the Nvjp-1 protein and other proteins with metal-coordination bonds from a bottom-up perspective. It also provides insight into synthesizing peptide materials with mutable properties.

6.2 Materials and Methods

6.2.1 Replica-exchange molecular dynamics simulation

In this study, replica-exchange molecular dynamics (REMD) simulation (Sugita and Okamoto 1999) is used to investigate the formation of protein structures in various environments. REMD is an effective tool that allows us to identify a native protein structure by overcoming kinetic trapping in local energy minima in the simulation. The simulation approach is illustrated in Figure 6-1.

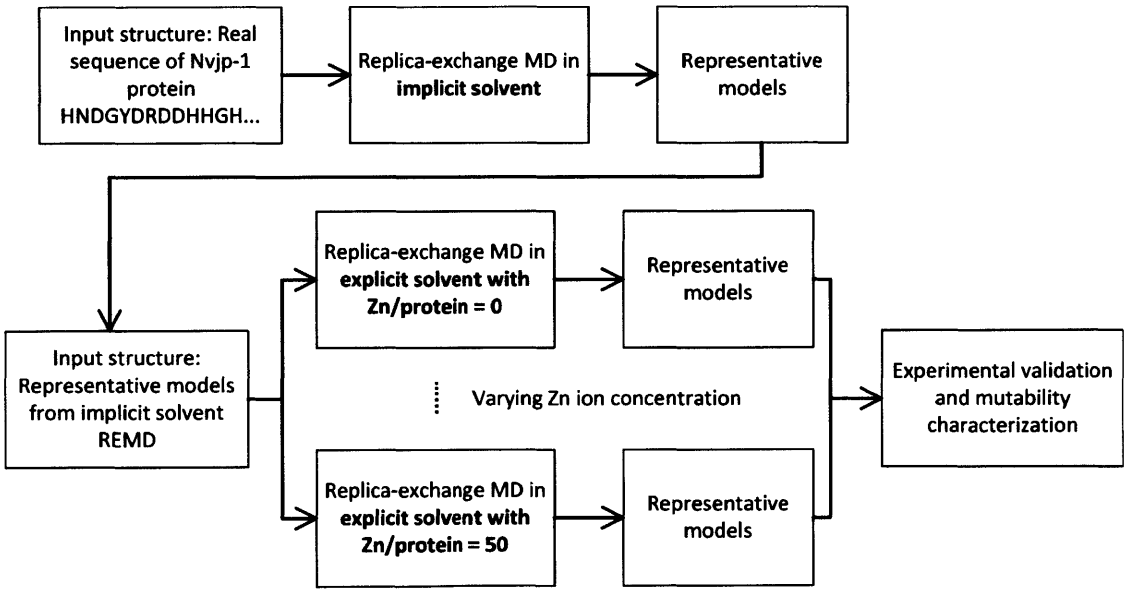

Figure 6-1 Illustration of simulation protocol for identifying the structure of Nvjp-1 in this study.

Primary protein structures are created with the real amino acid sequence of the Nvjp-1 protein (Broomell, Chase et al. 2008). To simulate the system at pH 8, the protonation states of Nε2H for histidine (His), Oδ2(−) for asparatic acid (Asp) residues, and Oε2(−) for glutamic acid (Glu) are considered, respectively. We first carry out simulations with Langevin dynamics using the CHARMM program with implicit solvent (EEF1), which is based on the CHARMM19 all-atom energy function with an effective Gaussian model to compute solvation free energy and nonbonding interactions (Lazaridis and Karplus 1999, Brooks, Brooks et al. 2009). The implicit-solvent REMD is performed using the Multiscale Modeling Tools for Structural Biology (MMTSB) toolset (Feig, Karanicolas et al. 2004) with a time step of 2 fs. The simulations are performed on 24 replicas within the 300–480 K temperature range with an exchange time step of 2 ps to allow the relaxation of the system. Each replica is simulated for a total of 40 ns, corresponding to a total simulation time of 0.96 μs.

We further analyzed the structure from the last 1,000 exchanges of the production run in the trajectory at 300K. The K-means clustering algorithm in the MMTSB toolset is used to cluster the structures based on mutual conformational similarity according to a root mean square deviation (RMSD) within 2 Å. Four representative structures with the lowest energy and close to the center of clusters are identified, as shown in Figure **6-2**.

In order to study the ion effect, the REMD with explicit solvent is applied to explore the ensemble of the Nvjp-1 protein structure in the presence of varied concentrations of Zn ion. The REMD simulations with explicit solvent and various ionic environments are carried out in the NAMD 2.8 (Phillips, Braun et al. 2005). The CHARMM force field is applied (MacKerell, Bashford et al. 1998) to describe the interaction between atoms. The initial models used in the explicit-solvent REMD simulations are obtained from the representative structures of the implicit solvent REMD. The entire atomistic structure is embedded completely in a TIP3 explicit water box with different ionic concentrations (Zn/protein = 0, 10, 20, 50). A simulation time step of 2 fs is used, and the particle mesh ewald function with a grid width of 1 Å is used to calculate electronic interactions.

Conventional MD simulations were first performed on these systems in the NPT ensemble with a constant pressure of 1 atmosphere at 300 K for 10 ns to equilibrate the structure in explicit solvent. The explicit-solvent REMD simulations were performed with the NVT ensemble at a temperature range of 300-480 K to simulate the ionic effect on the protein structure. The computational cost is 768 CPUs for 96 replicas (8 CPUs for each replica) in XSEDE. The equilibrium time within exchange is 2 ps with an average acceptance ratio of 36 %. For each ion concentration, every replica is simulated for a total of 2,000 cycles until equilibrium is established. We then analyze the conformations from the trajectory at 300 K through the K-means clustering algorithm with 2 Å radius based on mutual similarity according to RMSD. The structures of selected clusters with the lowest energy are identified as the representative conformations.

6.3 Results and Discussion

Four representative structures obtained from the REMD simulations in implicit solvent are shown in Figure 2, and the secondary structure of each amino acid is identified using the STRIDE algorithm in Visual Molecular Dynamics (VMD) software (Humphrey 1996). The implicit-solvent REMD simulation results show that the Nvjp-1 protein has primary turn and random coil configurations, and small clusters of beta-sheet and helix secondary structures. Different secondary structure contents are illustrated in Figure 2(A)-(D). Both N- (residue 1-174) and C-terminals (residue 175-381) contain roughly the same number of clusters of beta-sheet and helical structures, but C-terminal exhibits relatively longer clusters of helix. This difference might be due to the lower content of glycine amino acids in the C-terminal domain (30%) than in the N-terminal domain (42%). The representative structures predicted here are used as initial structures in the explicit-solvent MD simulations that follow.

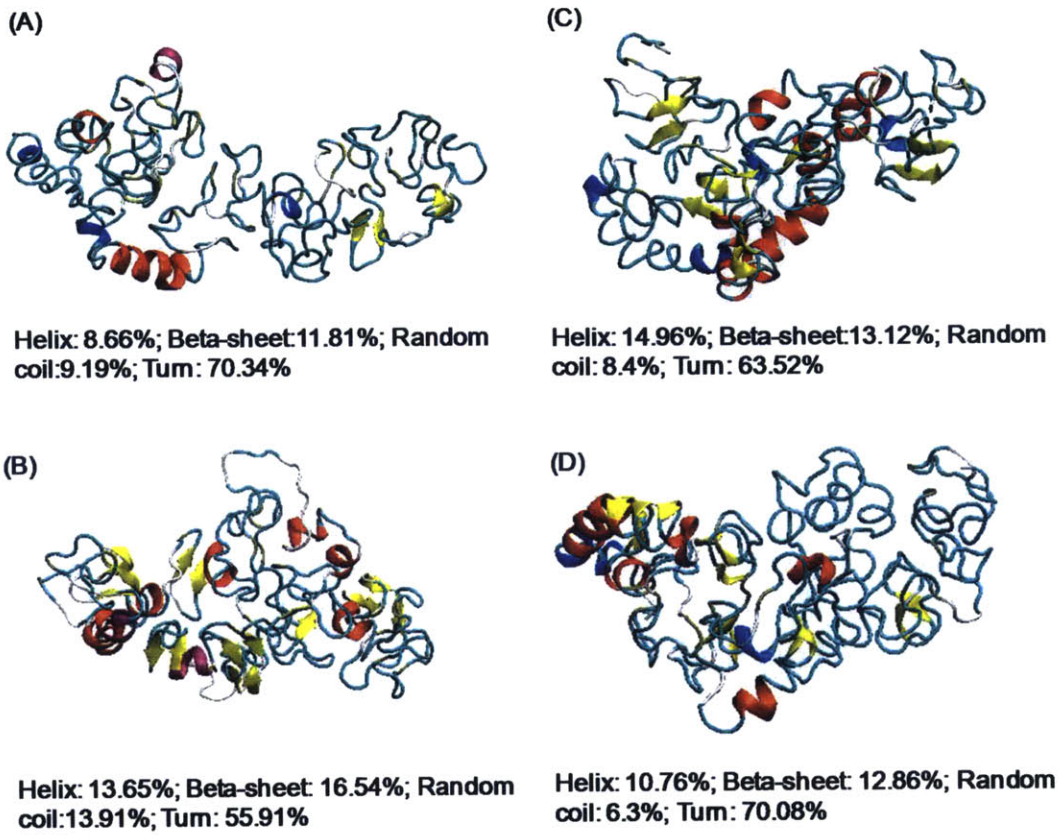

Figure 6-2 Four representative structures from implicit solvent replica-exchange MD simulation. The coloring is based on the secondary structure of each amino acid using the STRIDE algorithm in VMD.

In order to understand the mechanism of mutability of the Nvjp-1, the explicit-solvent REMD simulations with various ionic concentrations are employed to probe the nanostructure of the Nvjp-1 and mutable properties at an atomistic scale. In experiments, mutability of the Nvjp-1 protein in various ionic concentrations has been observed as shown in Figure **6-3**. Figure **6-3**(A) shows that the significant Nvjp-1 aggregation started at low levels of Zn/protein molar ratio, and the aggregation slightly increased at a Zn/protein molar ratio of 50. Contraction behavior of the Nvjp-1 crosslinked strips was also achieved through a high ionic concentration (Figure 3(B)).

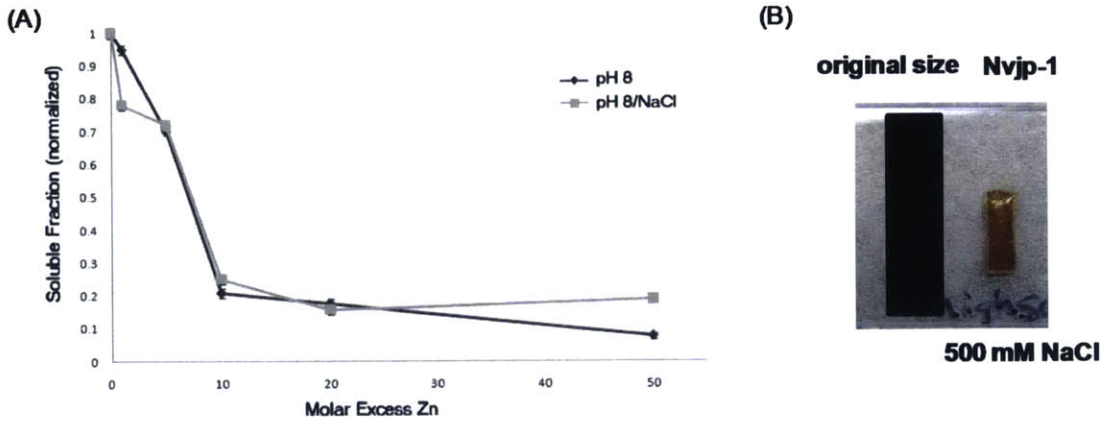

Figure 6-3 Ionic effect on Nvjp-1 protein observed in experiments. (A) Solubility: Zn2+ induced significant rNvjp-1 aggregation even at low levels; about 80% of the protein was precipitated at a Zn/protein molar ratio of 10. Image reproduced from (Broomell, Chase et al. 2008). (B) Contraction behavior of the Nvjp-1 crosslinked strips is achieved by ionic concentration.

Here, we first calculate the solvent accessible surface area (SASA) from the REMD simulations at different ionic concentrations (Zn/protein = 0, 20, 50) as shown in Figure **6-4**(A). SASA is calculated surface area that is accessible to a water molecule with a radius of 1.4 Å using VMD. The SASA of the Nvjp-1 protein decreases as the Zn/protein ratio increases, showing the same trend in the solubility of the Nvjp-1 protein measured in experiments in Figure **6-2**(A). From Zn/protein = 0 to 20, the SASA value decreases from 28000 $Å^2$ to approximately 26250 $Å^2$, indicating less solvent exposure of buried residues. While Zn/protein ratio increases from 20 to 50, the SASA decreases only slightly at the Zn/protein ratio of 50. We further calculate the radius of gyration to measure the size of Nvjp-1. Figure **6-4**(B) shows the radius of gyration of Nvjp-1 as a function of Zn/protein ratio. The radius of gyration of the Nvjp-1 protein decreases as the concentration of Zn ions increases. The results suggest that the more compact structures are formed in the presence of Zn ions, consistent with the results of the SASA calculations.

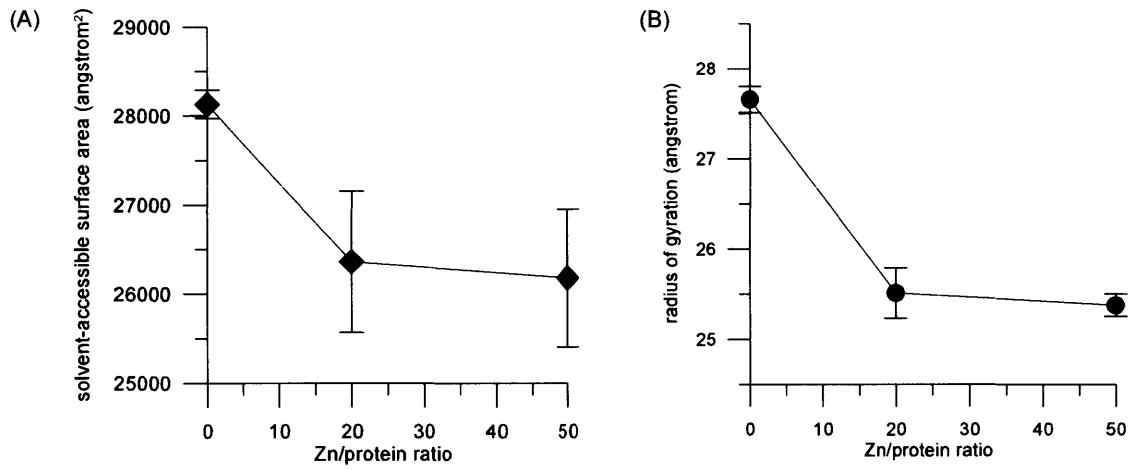

Figure 6-4 Solvent accessible surface area (SASA) and radius of gyration of the Nvjp-1 as a function of Zn/protein ratio. (A) The SASA of the Nvjp-1 protein decreases as the Zn/protein ratio increases, showing the same trend in the solubility of the Nvjp-1 protein measured in experiments. (B) The radius of gyration of the Nvjp-1 protein decreases in the presence of Zn ions.

We look deeper into the ionic effect on metal-coordination crosslinks of the Nvjp-1 protein. Figure 5 shows that the number of histidines and Zn ions forming metal-coordination crosslinks increases as the Zn/protein increases. The formation of metal-coordination crosslinks is almost saturated at lower Zn/protein ratios, and both the number of histidines and the Zn ions rise only slightly at a Zn/protein of 50. This pattern implies that more intra-molecular crosslinks are formed as the Zn/protein ratio increases, and the Zn ions stay in the center of histidines to form stable crosslinks in a chain. However, the number of crosslinks might not increase much in the presence of high Zn/protein molar ratios. Combined with the SASA and the radius of gyration calculations, the results of the formation of the compact structure and intra-chain crosslinks provide an explanation for the contraction behavior observed in the experiment. It is notable here that, because there is a single protein in the simulation, only the intra-molecular metal-coordination crosslink is considered.

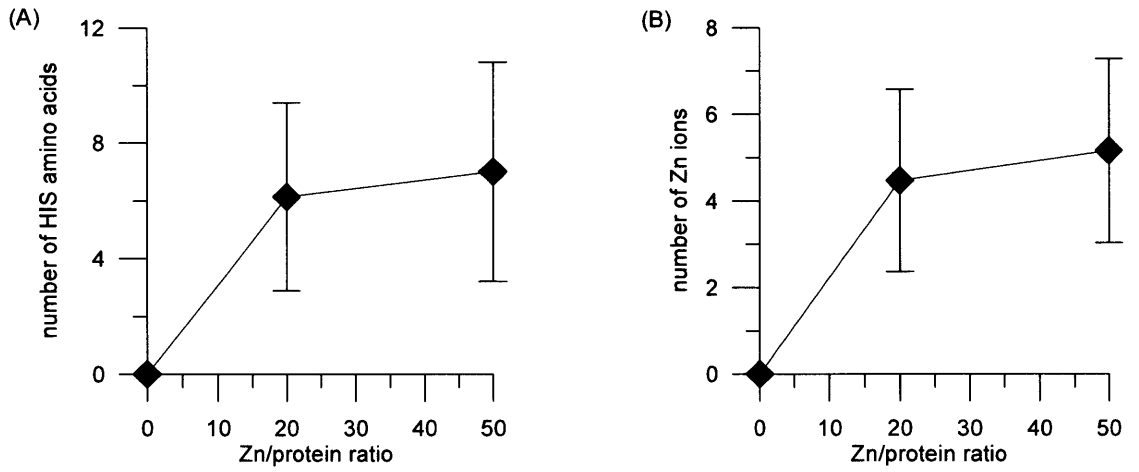

Figure 6-5 The number of (A) histidine amino acids and (B) zinc ions forming in metal coordination complexes.

Based on the formation of the intra-chain crosslinks in the Nvjp-1 protein, we further analyze the relationship between the sequence of the Nvjp-1 protein and the location of metal-coordination crosslinks. The analysis indicates that the crosslinks mainly occur in the two glycine- and histidine-rich segments, GHHG and GHGH. The segments are also identified as repetitive motifs of the Nvjp-1 protein using Statistical Analysis of Protein Sequence (SAPS) program (Brendel, Bucher et al. 1992). Some histidine-rich repetitive peptides in proteins have been characterized to form metal-coordination bonds (Gupta, Dobritsa et al. 2002, Pappalardo, Impellizzeri et al. 2002, Hara, Fujinaga et al. 2005, Zhao and Waite 2006, La Mendola, Magrì et al. 2012). Since glycine amino acids add local structural flexibility to the protein, histidine amino acids nearby are allowed to form a crosslink without causing a steric effect. The similar feature of glycine- and histidine-rich peptides binding metal ion has also been studied (Gupta, Dobritsa et al. 2002, Srivastava, Holten-Andersen et al. 2008).

We further compare the simulation snapshot with the model of proposed crosslinks in the Nereis jaws [1] in Figure **6-6**. The simulation snapshot clearly shows that Zn-mediated crosslink connects segments in the histidine-rich protein, implying the metal-coordination bond is responsible for stabilizing the material and further enhancing the mechanical properties of the Nereis jaws.

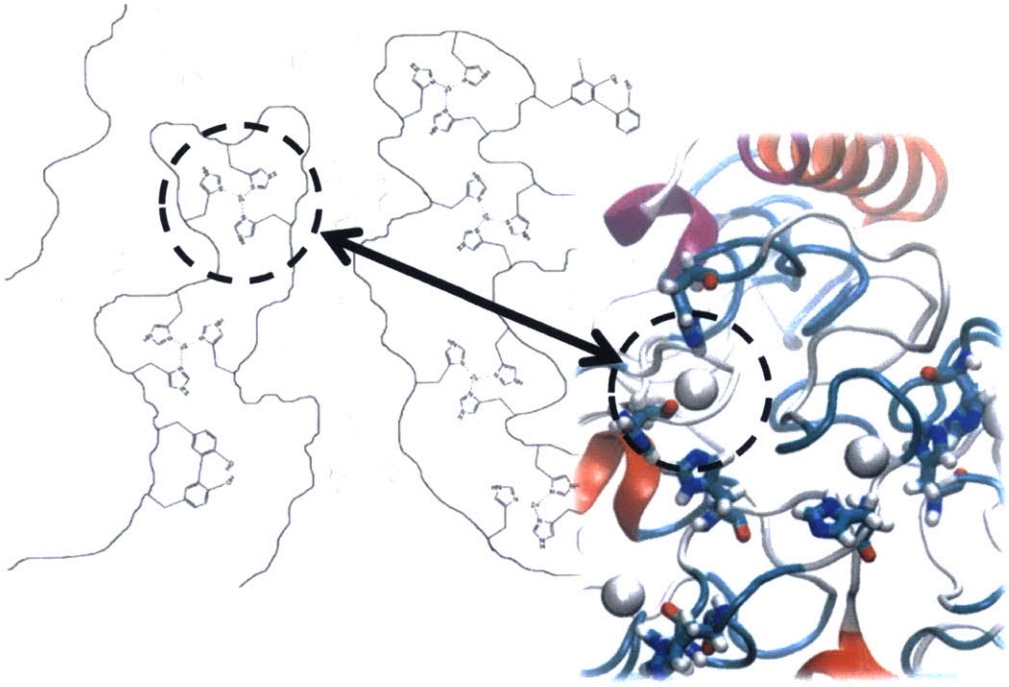

Figure 6-6 Simulation snapshot of intra-molecular crosslink forming in a metal-coordination complex. Zn ions and histidine amino acids are highlighted, and the protein is colored based on the secondary structure. Water molecules are not showed. The left image is reproduced from (Broomell, Mattoni et al. 2006).

6.4 Conclusions

We developed an atomistic model for the Nvjp-1 protein to predict response in solvents with various ion concentrations. Structurally tunable properties of the Nvjp-1 protein are demonstrated and validated against experiments. The SASA of the Nvjp-1 protein is calculated, and it shows the same trend in the solubility of the Nvjp-1 protein measured in experiments. The SASA of the Nvjp-1 protein decreases as the Zn/protein ratio increases. We also calculate the radius of gyration to measure the size of the Nvjp-1 protein. The result shows that the more compact structure is formed in the presence of Zn ions. Our results suggest that metal-coordination crosslinks play a significant role in achieving the characteristic properties of this protein material. This study provides fundamental insight into mutable mechanical properties of the Nvjp-1 protein and other proteins with metal-coordination bonds from a bottom-up perspective. It also provides an opportunity to

synthesize biomaterials with mutable properties in the presence of metal-coordination crosslinks.

7 Role of pH in structural changes of Nvjp-1 protein in the Nereis jaw: Molecular to mesoscale insights

Nvjp-1 protein is the predominant component of the Nereis jaw which contains >90% protein and no mineral content but has exceptional mechanical properties. In this study, we seek to understand how the solution pH affects the conformation of Nvjp-1 protein. Constant pH molecular dynamics simulations of the full-length Nvjp-1 and its truncations are performed to explore the role of pH in the conformation changes of Nvjp-1 proteins. Structurally mutable properties of the Nvjp-1 protein are demonstrated and validated against experiments. The simulations show that the solvent accessible surface area and the radius gyration of Nvjp-1 decrease as the pH level increases. The result reveals that the protein become more buried and the more compact structure formed at a higher pH. Our results indicate that at low pH, the protonation of histidine amino acids leads to the strong intramolecular electrostatic repulsion and results in the unfolding and disordered conformation of Nvjp-1. In contrast, at high pH, the deprotonation of histidine amino acids weakens the electrostatic repulsion, allowing a greater intramolecular interaction and the formation of the folded and structured conformation. We further develop coarse-grained models to investigate the intermolecular interaction and mechanical properties of Nvjp-1 protein assemblies at a larger scale. We demonstrate that the effect of pH on Nvjp-1 proteins leads to the formation of curvature structure. Using nanoindentation, the mutable stiffness/modulus of Nvjp-1 can be achieved through varied pH. The study suggests that the pH induces significant Nvjp-1 aggregation and the formation of metal-coordination crosslink to achieve the contraction of the Nvjp-1 stripe observed in experiments. This study provides fundamental insight into pH-dependent mutable mechanical properties of the Nvjp-1 protein and other proteins from a bottom-up perspective. The computational methodology illustrated here provides an ability to elucidate various roles of pH in material's mutability and to design biomaterials with mutable properties triggered by the various pH.

7.1 Introduction

Recently, many efforts have been reported to study biological materials with the goal of creating materials that are light weight but still have robust mechanical properties. One such biological material of great interest is Nereis Virens jaw (Bryan and Gibbs 1979, Lichtenegger, Schoberl et al. 2003, Birkedal, Khan et al. 2006, Broomell, Mattoni et al. 2006, Broomell, Chase et al. 2008, Broomell, Zok et al. 2008, Waite and Broomell 2012), a type of the marine polychaete jaws (Lichtenegger, Schoberl et al. 2003, Lichtenegger, Birkedal et al. 2005). This material is of particular relevance because it has magnificent strength that enables the Nereis Virens to grasp and tear prey. Remarkably, this jaw contains >90% protein and no mineral content. In comparison, dentin contains ~20% organic material and ~70% mineral (by mass). Yet, the two materials, dentin and jaw, have comparable hardness and stiffness (~1-2 GPa of hardness and 10-20 GPa of stiffness respectively) (Lichtenegger, Schoberl et al. 2003, Broomell, Khan et al. 2007, Pontin, Moses et al. 2007). Therefore it is intriguing how this material gains such strength having Nvjp-1 as its predominant protein (Broomell, Chase et al. 2008).

The Nereis jaw gains its strength from the ability of the Nvjp-1 to form metal-coordination bonds. Other proteins containing similar bonding schemes are the byssal thread in mussles whose strength and self-healing ability can be modulated with pH (Harrington and Waite 2007, Harrington, Gupta et al. 2009, Degtyar, Harrington et al. 2014). An increase in pH has shown to stiffen byssal threads in mussels. Moreover, experimental investigations have identified that the Nereis virens jaw protein Nvjp-1 exhibits a dramatic structural transition at different pH levels, suggesting that pH triggers a protein self-assembly mechanism (Broomell, Chase et al. 2008). Therefore, here we study the role that pH plays in molecular structural transition and the mechanical performance of Nvjp-1 protein in atomic and mesoscale. Understanding such biological control, for example determined through pH, can provide important insights for the field of bioinspired materials. This understanding allows replication in other material systems which is the basic premise of bio-inspired design.

The aforementioned phenomenon has prompted us to investigate the effect of pH on the structure of Nvjp-1 using continuous constant pH molecular dynamics (CpHMD) (Lee, Salsbury et al. 2004, Khandogin and Brooks 2005, Wallace and Shen 2011) in explicit solvent, which allows us to determine the pKa values of titratable amino acids and probe the pH-dependent conformational dynamics in atomistic detail. However, although prominent in the Nereis jaw, relatively little is known about the mechanism of its mutability, particularly the relationship to pH. An atomistic model is crucial to probe these properties of the Nvjp-1 proteins and to extrapolate this mechanism for bioengineering and other applications. The availability of an atomistic model of this protein will open the possibility to study the mechanical properties of the sea worm's jaw from a bottom-up perspective. It will also benefit the current studies on jaw-inspired materials to improve their design (Srivastava, Holten-Andersen et al. 2008, Holten-Andersen, Harrington et al. 2011).

Furthermore, a coarse-grained molecular dynamics (CGMD) model is utilized to study the intermolecular interaction and the mechanical behavior of Nvjp-1 proteins. Indeed, CGMD simulations are advantageous to handle large systems (more than 1,000,000 atoms) and long-time scales (~ ms) with reasonable accuracy and predictive power, when compared to full atomistic MD simulations (Marrink, Risselada et al. 2007, Monticelli, Kandasamy et al. 2008). Thus, CGMD simulations will enable us to simulate large scale phenomena using information obtained from molecular dynamics at the atomistic level. Through CGMD, we are able to incorporate the intermolecular interaction and metal-coordination crosslink to demonstrate the effect of pH on the mutability of Nvjp-1 proteins at a larger length scale.

The outline of this chapter is as follows. We begin with a description of the atomistic model constructed and system arrangement for the Nvjp-1 protein in the Materials and Methods. In the Results and Discussion, we then describe the atomistic simulation results, with a focus on the pH effect on the nanostructure of the Nvjp-1 protein. The results of the Nvjp-1 coarse-grained model is also presented. The conclusions and implications are presented in the Conclusion.

7.2 Methods

7.2.1 Continuous constant pH Molecular dynamics simulation

In order to study the pH effect, a set of constant pH molecular dynamics simulations (CpHMD) with explicit solvent is applied to explore the ensemble of the Nvjp-1 protein structure at the various pH values (pH=2, 5, 8 and 10). We also proceed with the CpHMD to study the pH effect on the conformation of two truncated Nvjp-1 proteins mutants. The mutants are two unequal halves, the N-terminal half (residue 1-174, Δ174) and the C-terminal half (residue 175-381, Δ207) of the full-length sequence of Nvjp-1 protein. Initial models in CpHMD are adopted from the representative structures in implicit solvent replica-exchange simulations (See Chapter 6).

The CpHMD simulations with explicit solvent and various pH environments are carried out in the CHARMM 37b1 (Brooks, Brooks et al. 2009) with PHMD module (Lee, Salsbury et al. 2004, Khandogin and Brooks 2005, Wallace and Shen 2011). The CHARMM force field is applied (Mackerell, Feig et al. 2004) to describe the interaction between atoms. The initial models used in the explicit-solvent REMD simulations are obtained from the representative structures of the implicit solvent REMD. The entire atomistic structure is embedded completely in a TIP3 (Jorgensen, Chandrasekhar et al. 1983) explicit water box with various pH values (pH = 2, 5, 8 and 10). The generalized Born (GB) implicit model (Im, Lee et al. 2003) was applied to calculate solvent-modulated electrostatic energies for the propagation of titration coordinates (Wallace and Shen 2011). The GB calculation and titration coordinates are updated every 10 dynamic steps. A simulation time step of 2 fs is used, and all simulations run at 300 K and 1 bar. The coordinates and titration states are monitored during the simulations. The data shown in the *Results* are the average values and standard deviations calculated from the last 1 ns CpHMD simulations at each pH value.

pK_a values are calculated by fitting the unprotonated fractions (S) at each pH (pH = 2, 5, 8, 10) to the Hill equation:

$$S = \frac{1}{10^{n(\mathrm{p}K_a - \mathrm{pH})} + 1}, \tag{7-1}$$

where S is the fraction of the total simulation that the titratable amino acids spend in an unprotonated state, and n is the Hill coefficients.

7.2.2 Coarse-Grained Modeling of Nvjp-1 proteins

In study, we develop coarse-grained models of full-length Nvjp-1 and two Nvjp-1 truncations to investigate their intermolecular interactions and mechanical behavior at the larger scale. The Nvjp-1 protein is coarse-grained from the representative structures predicted from the CpHMD at pH 5 and pH 8 following the philosophy of the MARTINI force field (Marrink, Risselada et al. 2007, Monticelli, Kandasamy et al. 2008). The MARTINI force field is used to describe the inter- and intra-molecular interaction of the Nvjp-1 proteins in explicit water. Four water molecules are modeled using one particle in the MARTINI force field.

pH-induced curvature simulations

First, we build three composite models to simulate the behavior of full-length Nvjp-1 and two Nvjp-1 truncations. Each model consists of two layers with Nvjp-1 proteins that are coarse-grained based on the atomistic structure and protonation stated from the full atomistic modeling at two different pH, pH 5 and 8. $ZnCl_2$ is added at 100 mM concentration to the system, which allows the formation of metal-coordination crosslinks in the system. The Nvjp-1 protein composite simulations are performed at a temperature of 300 K at NVT ensemble, with a 10 fs time step. It is worth noting that the system consists of around 1,600,000 CG particles (consisting of 170,000 CG protein particles and 1,430,000 CG solvent particles), which is equivalent to around 7 million atoms. The equilibration of the simulation is confirmed by monitoring the time history of root mean square displacements (RMSDs) of the system particles' movement. The coarse-grained MD simulations of Nvjp-1 composites are carried out using the GROMACS 4.5 software package (Hess, Kutzner et al. 2008).

A shape-fitting approach is applied to capture the curved composite. The 85 nm by 20 nm by 40 nm composite is first divided into layers with averaged 0.3 nm thickness parallel to the bending axis (z axis). The shape of the each layer is determined by the center of mass of each box that is used to compute a best-fit shape through least-squares fitting. The mathematical expression for the best-fit surface is a three-dimensional quadratic equation to describe the curvature:

$$z(x) = a + bx + cx^2 + dy + exy + fy^2, \tag{7-2}$$

where $\boldsymbol{z}$ axis is the direction perpendicular to the plane of the interface between upper and bottom layer of the Nvjp-1 composite, the $\boldsymbol{x}$ axis is the long axis of the composite, and the $\boldsymbol{y}$ axis is the short axis. $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$, $\boldsymbol{e}$, and $\boldsymbol{f}$ are coefficients determined during the fitting. Once a best-fit plane in the form of Equation (2) is obtained, the radius of curvature of each layer along the long axis ($\boldsymbol{x}$ axis) in the composite can be determined by averaging $\boldsymbol{r_x}$ where $\boldsymbol{x}$ is in the same layer. The radius of curvature along the $\boldsymbol{y}$ axis, $\boldsymbol{r_y}$, can be computed by the same approach. Given a plane defined by Equation (2), $\boldsymbol{r_x}$ and $\boldsymbol{r_y}$ can be evaluated via the analytical expressions (Hyde 1997):

$$r_x = \frac{\left[1+\left(\frac{dz}{dx}\Big|_y\right)^2\right]^{3/2}}{\frac{d^2z}{dx^2}\Big|_y} = \frac{[1+(b+ey+2cx)^2]^{3/2}}{2c}, \qquad r_y = \frac{\left[1+\left(\frac{dz}{dy}\Big|_x\right)^2\right]^{3/2}}{\frac{d^2z}{dy^2}\Big|_y} = \frac{[1+(d+ex+2fy)^2]^{3/2}}{2f}, \tag{7-3}$$

The curvatures κ_x and κ_y can be determined by $\kappa_x = {}^1/_{r_x}$ and $\kappa_y = {}^1/_{r_y}$.

Nanoindentation study of the mechanical behavior of Nvjp-1 proteins

Secondly, we build two 20 nm by 20 nm by 10 nm film structures that consist of coarse-grained full-length Nvjp-1 proteins at pH 5 and 8, individually to investigate the mechanical property of Nvjp-1 using the nanoindentation method. The coarse-grained MD simulations of Nvjp-1 film models are carried out using the LAMMPS software package (Plimpton 1995). After the system is equilibrated at 300K using an NPT

ensemble, the model subjects to indentation by a prescribed repulsive spring force acting on the system at a constant velocity and a set radius. A spherical indenter of radius R = 5 nm is used in this study and all samples are indented with a velocity of 10 m/s. As the nanoindentation proceeds, we record the force as a function of the depth of indentation. The force on each atom due to a spherical indenter is given as (Plimpton 1995)

$$F = -k(r - R)^2, \tag{7-4}$$

where $\boldsymbol{k}$ is the specified force constant, $\boldsymbol{r}$ is the distance from the atom to the centre of the indenter and $\boldsymbol{R}$ is the radius of the indenter. Note that the indenter is frictionless. When the indentation depth is smaller than the radius of indenter, the contact area can be approximated as (Fischer-Cripps 2011)

$$A = 2\pi R h_{\max}, \tag{7-5}$$

where $\boldsymbol{R}$ is the radius of the indenter at the maximum indentation depth $\boldsymbol{h_{max}}$.

Force-displacement data is analyzed assuming elastic contact (Hertzian), and the material indentation modulus (E^*) can be determined by (Pharr, Oliver et al. 1992, Fischer-Cripps 2011)

$$E^* = \frac{1}{2}\frac{\sqrt{\pi}}{\sqrt{A}}\frac{\Delta F}{\Delta h}, \tag{7-6}$$

The coarse-grained MD simulations of Nvjp-1 composite were carried out using the GROMACS 4.5 software package (Hess, Kutzner et al. 2008). The MARTINI force field (Marrink, Risselada et al. 2007) was used to describe the inter- and intra-molecular interaction of the Nvjp-1 composite in explicit water. Four water molecules were modeled using one particle in the MARTINI force field. The Nvjp-1 protein was coarse-grained from the representative structures predicted from the full atomistic simulation

(see Chapter 7) at different pH values following the philosophy of the MARTINI force field (Monticelli, Kandasamy et al. 2008).

7.3 Results and Discussion

7.3.1 pKa and protonation state of Nvjp-1 protein

In order to understand the mechanism of mutability of the Nvjp-1 protein, explicit-solvent CpHMD simulations with various pH conditions are employed. This is done to probe the nanostructure of the Nvjp-1 and mutable properties at the atomistic scale. Experimentally, the mutability of the Nvjp-1 protein in various pH is studied and are shown in Figure 7-1. Figure 7-1(A) shows significant Nvjp-1 aggregation at high pH. The swelling behavior of the Nvjp-1 crosslinked strips is observed at low pH as shown in Figure 7-1(B). As stated before, pH titration simulations are performed using the CpHMD method in explicit solvent for the full-length Nvjp-1 and its truncations. Molecular dynamics simulations at pH conditions of 2, 5, 8 and 10 are performed while simultaneously titrating all acidic and histidine amino acids. The simulations allow us to determine protonation states of each titratable amino acid and obtain the details of the conformational dynamics at each pH condition.

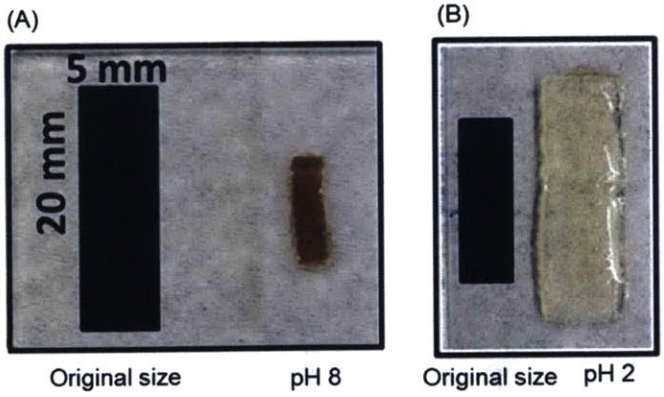

Figure 7-1. pH effect on Nvjp-1 protein observed in experiments. Contraction behavior of the Nvjp-1 crosslinked strips is achieved by pH.

From the CpHMD simulations, we calculate the pK_a values of titratable amino acids based on their protonation state at different pHs. Figure 3 shows the titration curves for a single amino acid of the Nvjp-1, His-199, in the explicit-solvent CPHMD simulations. The time history of the calculated pK_a for His-199 of the Nvjp-1 from the explicit-solvent CpHMD simulation shows that good convergence for the Nvjp-1 protein titrations and a stable calculated pK_a for the His-199 in the last 1 ns simulations. The calculation of pK_a is determined by fitting the unprotonated fractions S to the Hill equation shown in the Figure 7-2(B). From the CpHMD simulations, we observe highly protonated histidines at a lower pH and deprotonated histidines at a higher pH as illustrated in Figure 7-3(C). We compare the calculated pK_a values of titratable amino acids in CpHMD against the experimental measurements (Chapman and Feughelman 1967) listed in the Table 7-1. The calculated data are averaged all of the histidine, aspartic acid, glutamic acid and lysine amino acids in the full-length, Δ174, and Δ207 Nvjp-1, individually. The CpHMD simulations results and experiments are in close agreement.

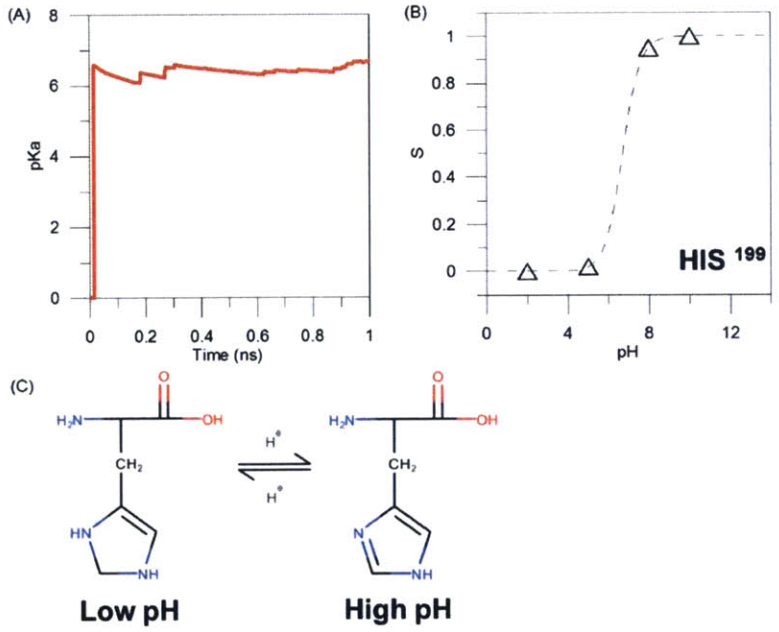

Figure 7-2 Titration curves for histidine-199 of the Nvjp-1 in the explicit-solvent CPHMD simulations. (A) Time history of the calculated pKa for histidine-199 of the Nvjp-1 from the explicit-solvent CpHMD simulation. The data shows that the

convergence of Nvjp-1 protein titrations and the stable calculated pKa of the histidine-199 in the last 1 ns simulations. (B) The unprotonated fractions S are averaged based on the last 1 ns simulations and fit to the Hill equation. (C) The illustration of highly protonated histidine at low pH and deprotonated histidine at high pH observed in the CpHMD simulations.

Table 7-1 Calculated and experimental $\mathbf{p}K_a$ values of histidine, aspartic acid, glutamic acid and lysine amino acids. The calculated data are averaged in the full-length, Δ174, and Δ207 Nvjp-1. The $\mathbf{p}K_a$ values measured in experiments are based on the blocked single amino acids from (Nozaki and Tanford 1967).

Amino acid	Full-length	Δ174	Δ207	Experiment
Histidine	6.78 ± 1.02	7.18 ± 0.76	6.74 ± 0.74	6.6/7.0
Aspartic acid	2.91 ± 1.06	3.00 ± 0.84	2.23 ± 1.52	4.0
Glutamic acid	3.07 ± 0.60	-	2.96 ± 0.75	4.4
Lysine	10.54 ± 1.29	-	11.70 ± 1.25	10.4

7.3.2 pH-dependent solvent exposure and radius of gyration

We further calculate the solvent accessible surface area (SASA) using the equilibrium conformations of full-length Nvjp-1 and two truncations (Δ174 and Δ207) from the CpHMD simulations at different pH conditions (pH = 2, 5, 8 and 10) as shown in Figure 7-3(A). SASA is the calculated surface area that is accessible to a water molecule with a radius of 1.4 Å using Visual Molecular Dynamics (Humphrey 1996). The SASA of the full-length Nvjp-1 decreases as pH increases, indicating that the protein becomes less exposed to water at a higher pH which is a stabilizing effect. Figure 7-3(A) shows, for both Nvjp-1 truncations, that the proteins also become more buried as pH increases. We

also calculate the radius of gyration to measure the size of full-length Nvjp-1 and its truncations. Figure 7-3(B) shows the radius of gyration of Nvjp-1 as a function of pH. The radius of gyration of the Nvjp-1 protein decreases as the pH increases, suggesting that the more compact structures are formed at a higher pH which is consistent with the results of the SASA calculations.

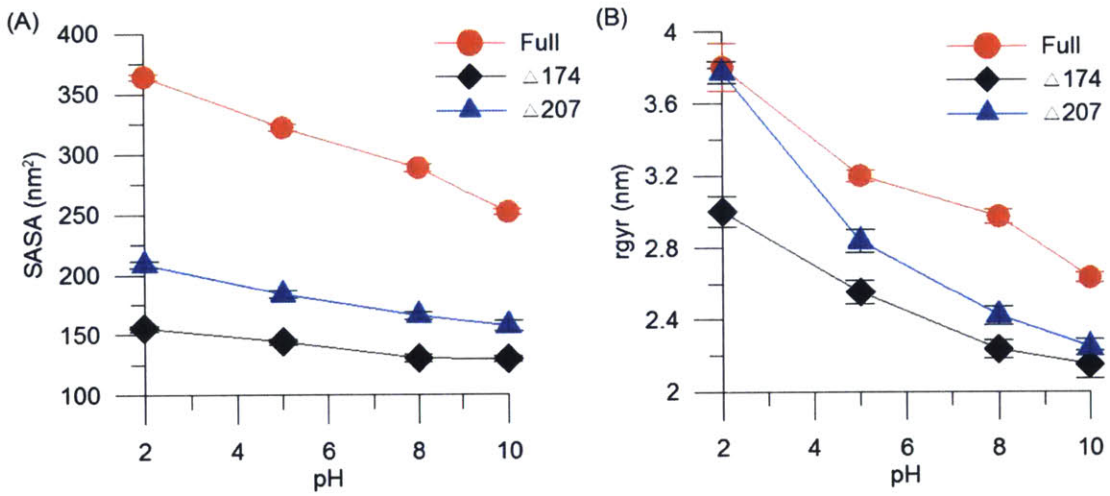

Figure 7-3 Solvent accessible surface area (SASA) and radius of gyration (rgyr) of the Nvjp-1 as a function of pH. (A) The SASA of the Nvjp-1 protein decreases as pH increases, indicating the protein becomes less exposed at a higher pH. (B) The radius of gyration of the Nvjp-1 protein decreases as pH increases, showing the same trend in the size of the Nvjp-1 stripe measured in experiments.

We next examine how the pH affects the hydrogen bond distribution in the Nvjp-1 systems. Figure 7-4 shows the number of hydrogen bonds within the full-length, Δ174 and Δ207 Nvjp-1 as a function of pH. As pH increases, the number of hydrogen bonds of the systems increases, indicating a greater concentration of intramolecular interactions within the system at a higher pH. The results demonstrate a morphology change that is consistent with the SASA and radius of gyration calculations from pH 2 to 10. Again, the pH-dependent H-bond distribution suggests a more compact structure for the Nvjp-1 proteins at elevated pH conditions.

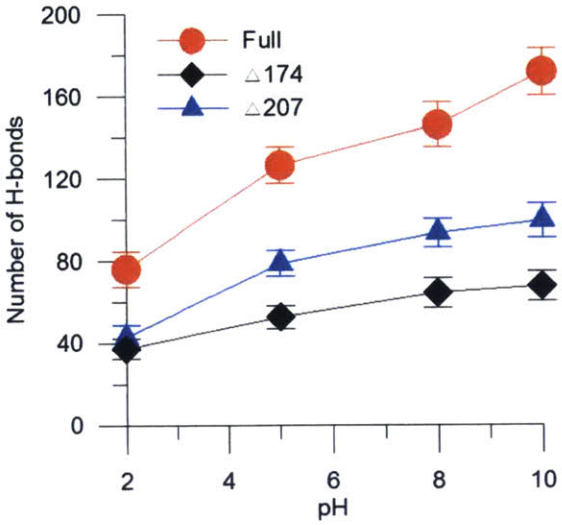

Figure 7-4 Hydrogen bonds distribution of full-length Nvjp-1 and two Nvjp-1 truncations (Δ174 and Δ207) as a function of pH.

7.3.3 pH-dependent secondary structure propensity

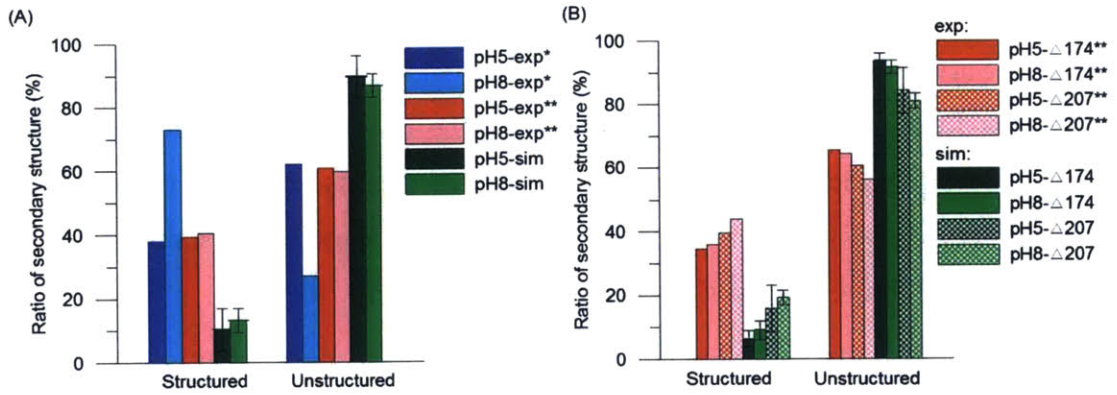

Figure 7-5 pH-dependent secondary structure transition of (A) full-length Nvjp-1 and (B) truncated Nvjp-1 (Δ174 and Δ207), simulation and experiment. Experimental data are obtained from (Broomell, Chase et al. 2008) (*) and from this study (**). Structured includes α-helix and β-sheet, and unstructured includes coil and turn.

Analysis of the equilibrium data reveals that the preferred secondary structure of the Nvjp-1 protein depends on pH. Figure 7-5(A) shows that Nvjp-1 protein forms more structured conformation at a higher pH. This trend implies strong influence of pH on the aggregation and assembly of Nvjp-1 protein. The simulation results clearly show the

structural transition from pH 5 to pH 8 which is the same trend seen in the literature (Broomell, Chase et al. 2008) (ATR-FTIR Spectroscopy) as well as our experimental data measured using circular dichroism spectroscopy. We now study with CpHMD the effect of pH on the conformation of two truncated Nvjp-1 proteins mutants. Both truncated segments form a more structured conformation at the higher pH value and have a more unstructured configuration at the lower pH as shown in Figure 7-5(B). Truncations of Nvjp-1 demonstrate the same structural transition from pH 5 to pH 8 compared to the full-length protein. The analysis also reveals the Δ174 truncation forms less structured conformation than the Δ207 truncation which is in good agreement with experimental measurements, showing the various pH sensitivities of two mutants.

We further analyze how the secondary structure changes through the sequence at various pH. Figure 7-6(A-B) shows the time history of the secondary structure of full-length Nvjp-1 at low pH and high pH along the sequence. We clearly see that the system exhibits more structured conformation at the higher pH, such as helix which is colored in blue, than at the low pH. For instance, the residue 210-265 forms more than 60% helix at the high pH but only forms less than 15% helix at the low pH. Furthermore, we compare the secondary structure of full-length Nvjp-1 and two Nvjp-1 truncations (Δ174 and Δ207) at the high pH, which is shown in Figure 7-6(C). The secondary structure is calculated from the final conformation of the CpHMD simulations. The result indicates that the sequence of the full-length and truncations all share similarities in the secondary structure. For example, the sequence of full-length and truncations with the helical structure are at approximately the same locations, meaning residue 60-70 and 210-285. Also, the secondary structure content of full-length and truncations are close to each other: the residue 1-174 and residue 175-381 in the full-length have roughly 4.54 and 14.97 percent of helix content which are similar to the helix contents, 3.27 % and 14.05 %, in the Δ174 and Δ207 truncations. The analysis also reveals the various pH sensitivities of different regions. In this regard, the residue 1-174 in the full-length Nvjp-1 forms less structured conformation than the residue 175-381, a result that is consistent to the different structured content of Δ174 and Δ207.

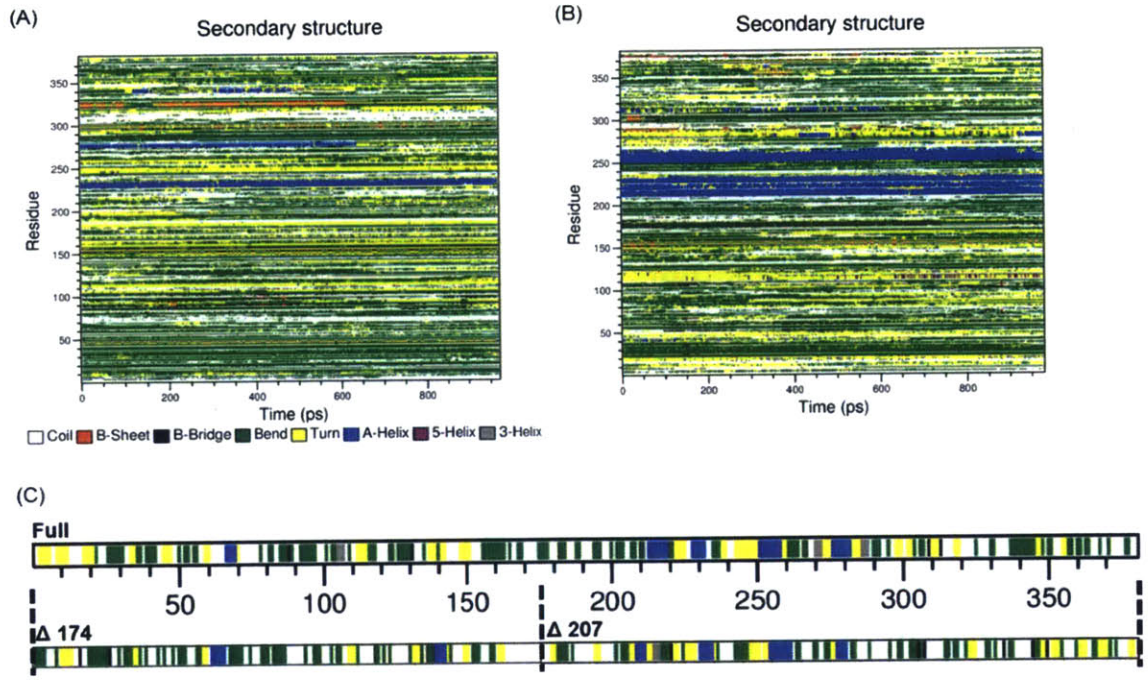

Figure 7-6 Time history of the secondary structure of full-length Nvjp-1 at (A) low pH and (B) high pH. At the higher pH, the system exhibits more structural conformation, such as helix which is colored in blue, than at the low pH. For instance, the residue 210-265 forms more than 60% helix at the high pH but only forms less than 15% helix at the low pH environments. (C) The comparison of the secondary structure of full-length Nvjp-1 and two Nvjp-1 truncations (Δ174 and Δ207) at high pH condition. The secondary structure is calculate from the final conformation in the CpHMD simulations. The result shows that the sequence of the full-length and truncations share similarities in the secondary structure. For example, the sequence of full-length and truncations with the helical structure are at the approximate same locations. Also, the secondary structure content of full-length and truncations is closed to each other: the residue 1-174 and residue 175-381 in the full-length have roughly 4.54 and 14.97 percent of helix content which are similar to the helix contents, 3.27 and 14.05 %, in the Δ174 and Δ207 truncations. The helix content is computed as the percentage of the number of helical residues.

Our data suggest that at the lower pH, the disruption of ordered formation is due to the electrostatic repulsion interaction between histidines. We use a simulation snapshot of

Nvjp-1 Δ207 as an example to show that structure transition of the unfolded and folded structures of the histidine-rich Nvjp-1 protein from the low to high pH in Figure 7-7. At the low pH, the protonation of histidine amino acids leads to the strong intramolecular electrostatic repulsion and the unfolding conformation of the Nvjp-1 protein. However, at the high pH, the deprotonation of histidine amino acids weakens the electrostatic repulsion, which allows the more intramolecular interaction and hydrogen bonds formation and results in the formation of the folded structure and more ordered conformation. Combined with the titration curve calculations of Nvjp-1 proteins, the results of the protonation states of histidine amino acids and the formation of the compact structure provide an explanation for the shrinkage behavior observed in the experiment. Our result also imply that the structural change of the single molecule contributes to the shrinkage observed in the experiments.

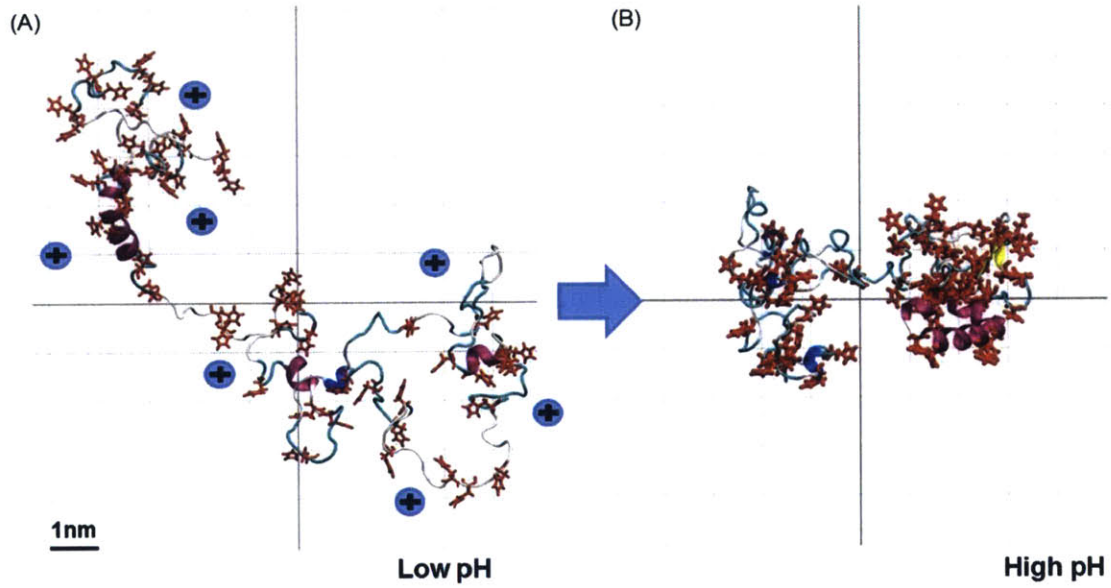

Figure 7-7 pH-dependent conformational change of Nvjp-1 truncation (Δ207). Simulation snapshots are taken from the simulations at (A) low (pH=2) and (B) high (pH=10) pH. (A) The protonation of histidine amino acids leads to the strong intramolecular electrostatic repulsion and the unfolding conformation of Nvjp-1 protein. (B) At the high pH, the deprotonation of histidine amino acids weakens the electrostatic repulsion, resulting the more compact structure is formed. Histidine amino acids are highlighted,

and the protein is colored based on the secondary structure. Water molecules are not shown.

7.3.4 Coarse-Grained Modeling of Nvjp-1 proteins

In order to investigate the intermolecular interactions of Nvjp-1 proteins at the larger scale, three coarse-grained composite structures of full-length Nvjp-1 and two Nvjp-1 truncations are built using the MARTINI force field. Each composite consists of two layers with coarse-grained Nvjp-1 proteins at pH 5 and 8. One of the initial composite models, full-length Nvjp-1, is shown in Figure 7-8(A). After more than 50 nanoseconds simulation time, the RMSDs of the composite systems fluctuate around a constant value, suggesting that the systems has reached a stable configuration. For the equilibrated model of full-length Nvjp-1, the molecules in the upper layer farther away from each other lead the entire structure curved (Figure 7-8(B)). Similar curved configurations for two Nvjp-1 truncations are also observed.

The curvature for each composite structure is analyzed as described in *Materials and Methods*, and radius of curvature, r_x, is computed using Equation (7-2). Note that the simulations are carried out in 3D, but our primary interest is understanding the radial (2D) behavior, perpendicular to the bending axis. For the full-length Nvjp-1, the initial radius of curvature along the x axis (r_x) is measured to be ~1400 nm, and the radius is reduced to ~190 nm at the end of the simulation. The initial curvature is significantly smaller ($\kappa_x = {}^{1}/_{r_x}$) than the equilibrated curvature, confirming quantitatively the curvature present at equilibrium. For the two Nvjp-1 truncations (Δ174 and Δ207), the initial radii are in accordance with the radius of full-length Nvjp-1 since the models are built by the same approach. In reaching equilibrium, the radius is reduced to ~270 nm for Δ174 and ~196 nm for Δ207.

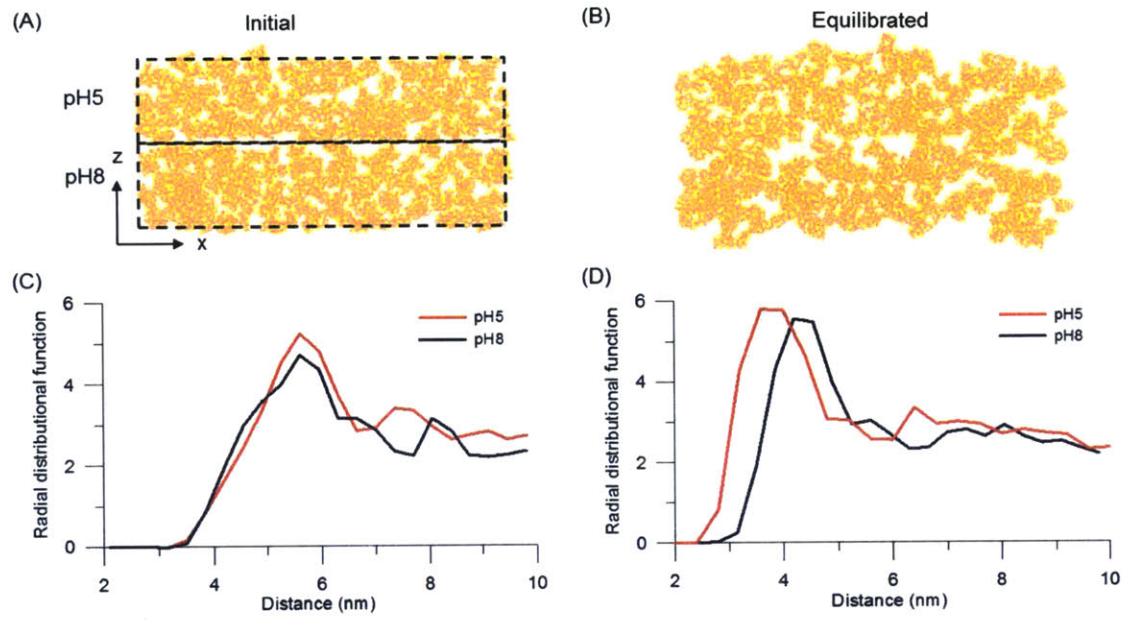

Figure 7-8 (A) Initial coarse-grained composite structures of Nvjp-1 proteins built using MARTINI force field. The composite is consisted of two layers with coarse-grained Nvjp-1 proteins at pH 5 and 8. (B) Equilibrated configuration of Nvjp-1 composite. (C) Radial distribution function of the upper (pH 5) and bottom (pH 8) layer in the (C) initial and (D) equilibrated composite structure.

We further analyze the radial distribution function for Nvjp-1 proteins in the upper and bottom layers of the composites to investigate the intermolecular interaction between the Nvjp-1 proteins. As can be seen in Figure 7-8(C), we initially place the molecules with equal intermolecular distance for both of upper and bottom layers in the initial model. After equilibrium (Figure 7-8(D)), we find that at the higher pH the intermolecular distance is shorter, suggesting the molecules pack tightly. This is because of the stronger electron repulsion force between highly charged histidine molecules at the lower pH, which leads to weaker intermolecular interactions and a larger intermolecular distance. Moreover, due to the molecular protonation states, the formation of the metal-coordination crosslink between Zn ion and histidines are more energetically favorable at the higher pH, inducing a shorter intermolecular distance. The finding of the varied intermolecular distance at lower and higher pH is also consistent with the experimental results.

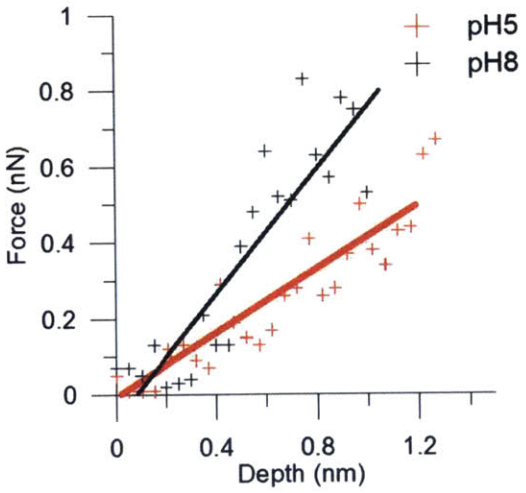

Figure 7-9 Force–displacement curves of the indentation tests. An increase of pH in the system leads to a steeper slope of the elastic region of the force–displacement curve. The linear segment of the force–displacement curve is fitted to determine the contact stiffness of the indenter on the film.

The force–displacement curves of nanoindentation are shown in Figure 7-9. As observed in Figure 7-9, an increase of pH in the system leads to a steeper slope of the elastic region of the force–displacement curve. The linear segment of the force–displacement curve is fitted to determine the contact stiffness of the indenter on the film. This linear fitting is shown in Figure 7-9. Force-displacement data is analyzed assuming elastic contact (Hertzian), and the material indentation modulus (E^*) can be determined by (Pharr, Oliver et al. 1992, Fischer-Cripps 2011)

$$E^* = \frac{1}{2}\frac{\sqrt{\pi}}{\sqrt{A}}\frac{\Delta F}{\Delta h}, \tag{7-7}$$

where A is the contact area, F is the force and h is the indentation depth. When the indentation depth is smaller than the radius of indenter, the contact area can be approximated as (Fischer-Cripps 2011)

$$A = 2\pi R h_{\max}, \tag{7-8}$$

where R is the radius of the indenter at the maximum indentation depth $h_{\max}$.

From the analysis, we see that the system at a higher pH condition has four times larger modulus than the system at lower pH ($E^*_{pH=8}/E^*_{pH=5} \cong 4$). As discussed above, since the ability of histidines to bind metal ions is pH-dependent, at the higher pH, we also observe a great concentration of metal-coordination crosslink in the system. This indicates that the metal-coordination crosslinks strengthens the system and contributes to the mechanical response. This result is consistent with the experiment study (Broomell, Mattoni et al. 2006), which shows the removal of Zn crosslinks by chelation in the Nereis jaw tip decreases both hardness and modulus by over 65%. Similar studies on byssal thread (Harrington and Waite 2007, Degtyar, Harrington et al. 2014) also show that pH-dependent metal-ion-binding ability of histidine influences thread mechanics. Our study depicts that the mutable stiffness/modulus of Nvjp-1 proteins can be achieved by changing pH conditions. These coarse-grained models enable us to simulate large scale phenomena using chemical structures obtained from molecular dynamics at the atomistic level. Through the models, we are able to incorporate the intermolecular interaction and metal-coordination crosslink to demonstrate the effect of pH on Nvjp-1 proteins that leads to the curvature structure and mutable mechanical properties.

7.4 Conclusions

We develop an atomistic model for the Nvjp-1 protein to predict response in solvents with various pH conditions. Structurally tunable properties of the Nvjp-1 protein are demonstrated and validated against experiments. We calculate the solvent accessible surface area and radius of gyration to measure solvent exposure and the size of the Nvjp-1 protein at different pH conditions. The result shows that the protein becomes more buried with a more compact structure formed at a higher pH condition. Our results reveal that at the low pH, the protonation of histidine amino acids leads to the strong intramolecular electrostatic repulsion and the unfolding and disordered conformation of Nvjp-1 protein. In contrast, at the high pH, the deprotonation of histidine amino acids weakens the electrostatic repulsion, which allows the more intramolecular interaction and results in the formation of the folded and structured conformation. The study suggests that the pH induces significant Nvjp-1 aggregation and achieves the contraction of the Nvjp-1 stripe observed in experiments (Figure **7-1**). This study provides fundamental

insight into pH-dependent mutable mechanical properties of the Nvjp-1 protein, and perhaps other similar proteins, for instance, mussel byssal threads and sandcastle worm glue (Shao, Bachus et al. 2009, Stewart, Wang et al. 2011, Degtyar, Harrington et al. 2014), from a bottom-up perspective. Based on the structural and chemical information at atomistic level, we develop coarse-grained models to investigate the intermolecular interaction and mechanical property of Nvjp-1 proteins at a larger scale. We demonstrate that the effect of pH on Nvjp-1 proteins leads to the formation of curvature in the structure. Using nanoindentation, the mutable stiffness/modulus of Nvjp-1 can be achieved through varied chemical conditions. Our work represents an effort toward understanding the environmental effects on protein folding/structural transition and the mechanical properties. The pH-dependent intramolecular/intermolecular mechanism unveiled by our study may underlie the mutable phenomena of other pH-sensitive biological materials. The computational methodology illustrated here provides an ability to elucidate various roles of pH in a material's mutability and to design and synthesize biomaterials with mutable properties triggered by various pH.

8 Conclusions and opportunities for future research

8.1 Summary of key findings and significances

In this thesis, we develop the first multiscale framework to study the hierarchical structure of keratin proteins with disulfide bond crosslink from a bottom-up molecular approach. We build the first full atomistic structure of keratin (heterodimer) and upscale to the keratin microfibril level. We also study the mutability of Nvjp-1 protein with metal-coordination bonds in the varied chemical environments from the single molecule level to larger length scales (~μm). Investigations of fundamental mutability of biological materials with disulfide bonds and metal coordination bonds from atomistic scale provide essential insights into the interplay of bond energy, chemical microenvironment and structure to achieve the desired mechanical properties at multiple scales.

By modeling keratin molecules, we show that full atomistic simulation can predict the structure from real sequences of keratin heterodimer k35/k85 molecules. We provide insights into the rupture mechanism of disulfide crosslinks in keratin protein and illustrate the importance of the redox environment, where factors such as accessibility, mechanical strain and local redox potential govern the dominating rupture mechanism and location. Disulfide bonds result in a higher strength and toughness of keratin proteins, but the system loses alpha-helical structures under loading, suggesting that disulfide bonds play a significant role in achieving the characteristic mechanical properties of trichocyte alpha-keratin. We provide a general computational protocol for studying mechanochemical fracture of large-scale protein materials concurrently with experimental efforts.

Further, we study the keratin macrofilaments using coarse-grained modeling. In the macrofilaments, disulfide crosslinks between the microfibril–matrix and matrix–matrix contribute to the initial modulus and provide stiffening behavior at larger deformation of the trichocyte keratin fibers. The results show that disulfide bonds reinforce the macrofilament and enhance the robustness of the macrofilament by facilitating the

microfilaments to deform cooperatively. The mesoscopic model of this protein, proposed in this study, opens the possibility to further explore the relationship between microscopic chemical structure and macroscopic performance for a bottom-up description of soft materials.

Using molecular dynamics simulation, we investigate the mutability of Nvjp-1 triggered by ion concentration and pH. Nvjp-1 forms a more compact structure in the presence of Zn ions, and more stable intra-molecular metal coordination complexes are formed at higher ion concentrations. Moreover, the protein also becomes more buried and a more compact structure is formed at a higher pH. At a high pH, the deprotonation of histidine amino acids weakens the electrostatic repulsion, allowing a greater intramolecular interaction and the formation of the folded and structured conformation. A coarse-grained model is developed to investigate the intermolecular interactions and mechanical properties of Nvjp-1 proteins at a larger scale. We demonstrate that the effect of pH on Nvjp-1 proteins leads to varied intermolecular distances of Nvjp-1. Using nanoindentation, the mutable stiffness/modulus of Nvjp-1 can be achieved through varied pH. This study suggests that metal-coordination crosslinks and pH effects induce significant Nvjp-1 aggregation and achieve the contraction of the Nvjp-1 stripe observed in experiments. This study provides fundamental insights into ionic/pH-dependent mutable mechanical properties of the Nvjp-1 protein and other proteins from a bottom-up perspective. The computational methodology illustrated here provides an ability to elucidate various roles of ion/pH in material's mutability and to design biomaterials with mutable properties triggered by the various chemical conditions.

8.2 Opportunities for future research

The multiscale framework developed in this thesis provides a useful tool to study many aspects of mutable biological materials under various chemical conditions. Possible opportunities for future research are discussed in this section.

One opportunity for future research is to extend our trichocyte keratin model to study other keratin mechanical properties in epidermis or endothelial cells. Recent studies have

shown that keratin intermediate filament network plays an important role in a cell's response to tensile and shear stresses (Yoon, Yoon et al. 2001, Yano, Komine et al. 2004, Fudge, Russell et al. 2008, Sivaramakrishnan, Schneider et al. 2009, Obarzanek-Fojt, Favre et al. 2011, Yoon, Na et al. 2012). Also, several mutation locations along the keratin molecules which cause severe disease conditions have been identified (Ma, Yamada et al. 2001, Godsel, Hobbs et al. 2008, Bolling, Lemmink et al. 2011). Our approach can be used to build the other keratin proteins and to investigate how the mutation affects the structure and mechanical property using full atomistic models. This multiscale framework can be used to study the mechanisms of keratin network mechanotransduction in a cell and the effects of various single-point mutations, which would help understanding the effect of each mutation and developing treatments.

Our models provide molecular details on how metal-coordination crosslinks and pH affect the structure and mechanics of Nvjp-1 protein. These molecular understandings and the approach can be applied to simulate other mutable materials, for example, the mussel byssus and other synthetic polymers (Harrington and Waite 2007, Hwang, Zeng et al. 2010, Holten-Andersen, Harrington et al. 2011, Fullenkamp, He et al. 2013, Degtyar, Harrington et al. 2014). Full atomistic modeling would allow us to study the mechanism of mutability of these materials. Furthermore, the results of atomistic modelling may feed into coarse-grained modelling to study the mechanical response at larger scales, including a dynamic response, from a microscale molecular level to a macroscale level.

References

Ackbarow, T. and M. Buehler (2007). "Superelasticity, energy dissipation and strain hardening of vimentin coiled-coil intermediate filaments: atomistic and continuum studies." Journal of Materials Science **42**(21): 8771-8787.

Ackbarow, T. and M. J. Buehler (2008). "Hierarchical Coexistence of Universality and Diversity Controls Robustness and Multi-Functionality in Protein Materials." Journal of Computational and Theoretical Nanoscience **5**(7): 1193-1204.

Ackbarow, T., X. Chen, S. Keten and M. J. Buehler (2007). "Hierarchies, multiple energy barriers, and robustness govern the fracture mechanics of α-helical and β-sheet protein domains." Proceedings of the National Academy of Sciences **104**(42): 16410-16415.

Ackbarow, T., S. Keten and M. J. Buehler (2009). "A multi-timescale strength model of alpha-helical protein domains." Journal of Physics-Condensed Matter **21**(3): 035111.

Ackbarow, T., D. Sen, C. Thaulow and M. J. Buehler (2009). "Alpha-Helical Protein Networks Are Self-Protective and Flaw-Tolerant." Plos One **4**(6): e6015.

Akkermans, R. L. C. and P. B. Warren (2004). "Multiscale modelling of human hair." Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **362**(1821): 1783-1793.

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter (2002). Molecular Biology of the Cell. Ner York, Taylor & Francis.

Albrecht, K. and A. Bernkop-Schnurch (2007). "Thiomers: forms, functions and applications to nanomedicine." Nanomedicine **2**(1): 41-50.

Alegre-Cebollada, J., P. Kosuri, J. A. Rivas-Pardo and J. M. Fernández (2011). "Direct observation of disulfide isomerization in a single protein." Nat Chem **3**(11): 882-887.

Apgar, J. R., K. N. Gutwin and A. E. Keating (2008). "Predicting helix orientation for coiled-coil dimers." Proteins: Structure, Function, and Bioinformatics **72**(3): 1048-1065.

Arslan, M., Z. Qin and M. J. Buehler (2011). "Coiled-coil intermediate filament stutter instability and molecular unfolding." Computer Methods in Biomechanics and Biomedical Engineering **14**(5): 483-489.

Aslund, F. and J. Beckwith (1999). "Bridge over troubled waters: Sensing stress by disulfide bond formation." Cell **96**(6): 751-753.

Astbury, W. T. and A. Street (1932). "X-ray studies of the structure of hair, wool, and related fibres I - General." Philosophical Transactions of the Royal Society of London Series a-Containing Papers of a Mathematical or Physical Character **230**: 75-101.

Astbury, W. T. and H. J. Woods (1934). "X-ray studies of the structure of hair, wool, and related fabrics II - The molecular structure and elastic properties of hair keratin." Philosophical Transactions of the Royal Society of London Series a-Containing Papers of a Mathematical or Physical Character **232**: 333-U367.

Azoia, N. G., M. M. Fernandes, N. M. Micaêlo, C. M. Soares and A. Cavaco-Paulo (2012). "Molecular modeling of hair keratin/peptide complex: Using MM-PBSA calculations to describe experimental binding results." Proteins: Structure, Function, and Bioinformatics **80**(5): 1409-1417.

Bell, G. I. (1978). "Models for Specific Adhesion of Cells to Cells." Science **200**(4342): 618-627.

Bendit, E. G. (1960). "A Quantitative X-Ray Diffraction Study of the Alpha-Beta Transformation in Wool Keratin." Textile Research Journal **30**(8): 547-555.

Benkovic, S. J. and S. Hammes-Schiffer (2003). "A Perspective on Enzyme Catalysis." Science **301**(5637): 1196-1202.

Bernkop-Schnürch, A., V. Schwarz and S. Steininger (1999). "Polymers with Thiol Groups: A New Generation of Mucoadhesive Polymers?" Pharmaceutical Research **16**(6): 876-881-881.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi (1977). "Protein Data Bank - Computer-based archival file for macromolecular structures." Journal of Molecular Biology **112**(3): 535-542.

Bertram, J. E. and J. M. Gosline (1987). "Functional design of horse hoof keratin: the modulation of mechanical properties through hydration effects." Journal of Experimental Biology **130**(1): 121-136.

Birkedal, H., R. K. Khan, N. Slack, C. Broomell, H. C. Lichtenegger, F. Zok, G. D. Stucky and J. H. Waite (2006). "Halogenated veneers: Protein cross-linking and halogenation in the jaws of Nereis, a marine polychaete worm." Chembiochem **7**(9): 1392-1399.

Bolling, M. C., H. H. Lemmink, G. H. L. Jansen and M. F. Jonkman (2011). "Mutations in KRT5 and KRT14 cause epidermolysis bullosa simplex in 75% of the patients." British Journal of Dermatology **164**(3): 637-644.

Bonomi, M., D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia and M. Parrinello (2009). "PLUMED: A portable plugin for free-energy calculations with molecular dynamics." Computer Physics Communications **180**(10): 1961-1972.

Bonomi, M. and M. Parrinello (2010). "Enhanced sampling in the well-tempered ensemble." Phys Rev Lett **104**(19): 190601.

Brändén, C.-I. and J. Tooze (1999). Introduction to protein structure. New York, Garland Pub.

Branden, C. I. and J. Tooze (1999). Introduction to Protein Structure. New York, Garland Pub

Brendel, V., P. Bucher, I. R. Nourbakhsh, B. E. Blaisdell and S. Karlin (1992). "Methods and algorithms for statistical analysis of protein sequences." Proceedings of the National Academy of Sciences **89**(6): 2002-2006.

Brockwell, D. J., E. Paci, R. C. Zinober, G. S. Beddard, P. D. Olmsted, D. A. Smith, R. N. Perham and S. E. Radford (2003). "Pulling geometry defines the mechanical resistance of a beta-sheet protein." Nature structural biology **10**(9): 731-737.

Brooks, B. R., C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus (2009). "CHARMM: The Biomolecular Simulation Program." Journal of Computational Chemistry **30**(10): 1545-1614.

Broomell, C. C., S. F. Chase, T. Laue and J. H. Waite (2008). "Cutting edge structural protein from the jaws of Nereis virens." Biomacromolecules **9**(6): 1669-1677.

Broomell, C. C., R. K. Khan, D. N. Moses, A. Miserez, M. G. Pontin, G. D. Stucky, F. W. Zok and J. H. Waite (2007). Mineral minimization in nature's alternative teeth.

Broomell, C. C., M. A. Mattoni, F. W. Zok and J. H. Waite (2006). "Critical role of zinc in hardening of Nereis jaws." Journal of Experimental Biology **209**(16): 3219-3225.

Broomell, C. C., F. W. Zok and J. H. Waite (2008). "Role of transition metals in sclerotization of biological tissue." Acta Biomaterialia **4**(6): 2045-2051.

Brown, J. H., C. Cohen and D. A. D. Parry (1996). "Heptad breaks in α-helical coiled coils: Stutters and stammers." Proteins: Structure, Function, and Bioinformatics **26**(2): 134-145.

Bruce Fraser, R. D. and D. A. D. Parry (2003). "Macrofibril assembly in trichocyte (hard α-) keratins." Journal of Structural Biology **142**(2): 319-325.

Bryan, G. W. and P. E. Gibbs (1979). "Zinc - a major inorganic component of nereid polychaete jaws." Journal of the Marine Biological Association of the United Kingdom **59**(04): 969-973.

Buehler, M. J. (2006). "Nature designs tough collagen: Explaining the nanostructure of collagen fibrils." Proceedings of the National Academy of Sciences **103**(33): 12285-12290.

Buehler, M. J. (2007). "Hierarchical chemo-nanomechanics of proteins: Entropic elasticity, protein unfolding and molecular fracture." Journal of Mechanics of Materials and Structures **2**(6): 1019-1057.

Buehler, M. J. (2007). "Molecular nanomechanics of nascent bone: fibrillar toughening by mineralization." Nanotechnology **18**(29): 295102.

Buehler, M. J. (2008). Atomistic Modeling of Materials Failure, Springer.

Buehler, M. J. (2008). "Nanomechanics of collagen fibrils under varying cross-link densities: Atomistic and continuum studies." Journal of the Mechanical Behavior of Biomedical Materials **1**(1): 59-67.

Buehler, M. J. (2010). Atomistic modeling of materials failure / Markus J. Buehler, New York : Springer, 2010.

Buehler, M. J., A. C. T. van Duin and W. A. Goddard (2006). "Multiparadigm Modeling of Dynamical Crack Propagation in Silicon Using a Reactive Force Field." Physical Review Letters **96**(9): 095505-095504.

Buehler, M. J. and Y. C. Yung (2009). "Deformation and failure of protein materials in physiologically extreme conditions and disease." Nat Mater **8**(3): 175-188.

Buehler, M. J. and Y. C. Yung (2009). "Deformation and failure of protein materials in physiologically extreme conditions and disease." Nature Materials **8**(3): 175-188.

Cerritelli, S., D. Velluto and J. A. Hubbell (2007). "PEG-SS-PPS: Reduction-sensitive disulfide block copolymer vesicles for intracellular drug delivery." Biomacromolecules **8**(6): 1966-1972.

Chapman, B. M. and M. Feughelman (1967). "Aspects of the structure of α-Keratin derived from mechanical properties." Journal of Polymer Science Part C: Polymer Symposia **20**(1): 189-199.

Cheek, S., S. S. Krishna and N. V. Grishin (2006). "Structural classification of small, disulfide-rich protein domains." J Mol Biol **359**(1): 215-237.

Chenoweth, K., A. C. T. van Duin and W. A. Goddard (2008). "ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation." Journal of Physical Chemistry A **112**(5): 1040-1053.

Chou, C.-C. and M. J. Buehler (2011). "Biophysics: Breaking out of the cage." Nat Chem **3**(11): 837-839.

Chou, C.-C., E. Lepore, P. Antonaci, N. Pugno and M. J. Buehler (2015). "Mechanics of trichocyte alpha-keratin fibers: Experiment, theory, and simulation." Journal of Materials Research **30**(01): 26-35.

Chou, C. C. and M. J. Buehler (2012). "Structure and Mechanical Properties of Human Trichocyte Keratin Intermediate Filament Protein." Biomacromolecules **13**(11): 3522-3532.

Chou, S. F. and R. A. Overfelt (2011). "Tensile deformation and failure of North American porcupine quills." Materials Science & Engineering C-Materials for Biological Applications **31**(8): 1729-1736.

Cribb, B. W., C. L. Lin, L. Rintoul, R. Rasch, J. Hasenpusch and H. Huang (2010). "Hardness in arthropod exoskeletons in the absence of transition metals." Acta Biomaterialia **6**(8): 3152-3156.

Degtyar, E., M. J. Harrington, Y. Politi and P. Fratzl (2014). "The Mechanical Role of Metal Ions in Biogenic Protein-Based Materials." Angewandte Chemie International Edition **53**(45): 12026-12044.

Dietz, H. and M. Rief (2008). "Elastic Bond Network Model for Protein Unfolding Mechanics." Physical Review Letters **100**(9): 098101.

Feig, M., J. Karanicolas and C. L. Brooks (2004). "MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology." Journal of Molecular Graphics & Modelling **22**(5): 377-395.

Fernandes, P. A. and M. J. Ramos (2004). "Theoretical insights into the mechanism for thiol/disulfide exchange." Chemistry-a European Journal **10**(1): 257-266.

Feughelman, M. (1979). "Role of the Microfibrils in the Mechanical-Properties of Alpha-Keratins." Journal of Macromolecular Science-Physics **B16**(1): 155-162.

Feughelman, M. (1994). "A Model for the Mechanical-Properties of the Alpha-Keratin Cortex." Textile Research Journal **64**(4): 236-239.

Feughelman, M. (1997). Mechanical Properties and Structure of Alpha-keratin Fibres: Wool, Human Hair and Related Fibres. Sydney, UNSW Press.

Feughelman, M. (2002). "Natural protein fibers." Journal of Applied Polymer Science **83**(3): 489-507.

Feughelman, M. and A. R. Haly (1959). "Structural Features of Keratin Suggested by Its Mechanical Properties." Biochimica Et Biophysica Acta **32**(2): 596-597.

Fischer-Cripps, A. C. (2011). Nanoindentation., New York : Springer ; c2011. 3rd ed.

Flory, P. J. (1953). Principles of polymer chemistry. Ithaca,, Cornell University Press.

Fraser, R. D. B., T. P. MacRae, L. G. Sparrow and D. A. D. Parry (1988). "Disulphide bonding in α-keratin." International Journal of Biological Macromolecules **10**(2): 106-112.

Fraser, R. D. B. and D. A. D. Parry (2005). "The three-dimensional structure of trichocyte (hard α-) keratin intermediate filaments: Features of the molecular packing deduced from the sites of induced crosslinks." Journal of Structural Biology **151**(2): 171-181.

Fraser, R. D. B. and D. A. D. Parry (2007). "Structural changes in the trichocyte intermediate filaments accompanying the transition from the reduced to the oxidized form." Journal of Structural Biology **159**(1): 36-45.

Fratzl, P. and R. Weinkamer (2007). "Nature's hierarchical materials." Progress in Materials Science **52**(8): 1263-1334.

Fudge, D., D. Russell, D. Beriault, W. Moore, E. B. Lane and A. W. Vogl (2008). "The Intermediate Filament Network in Cultured Human Keratinocytes Is Remarkably Extensible and Resilient." PLoS ONE **3**(6): e2327.

Fudge, D. S. and J. M. Gosline (2004). "Molecular design of the α–keratin composite: insights from a matrix–free model, hagfish slime threads." Proceedings of the Royal Society of London. Series B: Biological Sciences **271**(1536): 291-299.

Fullenkamp, D. E., L. He, D. G. Barrett, W. R. Burghardt and P. B. Messersmith (2013). "Mussel-Inspired Histidine-Based Transient Network Metal Coordination Hydrogels." Macromolecules **46**(3): 1167-1174.

Garcia-Manyes, S., J. Liang, R. Szoszkiewicz, T. L. Kuo and J. M. Fernandez (2009). "Force-activated reactivity switch in a bimolecular chemical reaction." Nature Chemistry **1**(3): 236-242.

Godsel, L. M., R. P. Hobbs and K. J. Green (2008). "Intermediate filament assembly: dynamics to disease." Trends in Cell Biology **18**(1): 28-37.

Greenberg, D. A. and D. S. Fudge (2013). "Regulation of hard α-keratin mechanics via control of intermediate filament hydration: matrix squeeze revisited." Proceedings of the Royal Society B: Biological Sciences **280**(1750): 20122158.

Gruber, M., J. Soding and A. N. Lupas (2006). "Comparative analysis of coiled-coil prediction methods." Journal of Structural Biology **155**(2): 140-145.

Grützner, A., S. Garcia-Manyes, S. Kötter, C. L. Badilla, J. M. Fernandez and W. A. Linke (2009). "Modulation of Titin-Based Stiffness by Disulfide Bonding in the Cardiac Titin N2-B Unique Sequence." Biophysical Journal **97**(3): 825-834.

Gupta, R., S. Dobritsa, C. Stiles, M. Essington, Z. Liu, C.-H. Chen, E. Serpersu and B. Mullin (2002). "Metallohistins: A New Class of Plant Metal-Binding Proteins." Journal of Protein Chemistry **21**(8): 529-536.

Guthold, M., W. Liu, E. Sparks, L. Jawerth, L. Peng, M. Falvo, R. Superfine, R. Hantgan and S. Lord (2007). "A Comparison of the Mechanical and Structural Properties of Fibrin Fibers with Other Protein Fibers." Cell Biochemistry and Biophysics **49**(3): 165-181.

Hara, M., M. Fujinaga and T. Kuboi (2005). "Metal binding by citrus dehydrin with histidine-rich domains." Journal of Experimental Botany **56**(420): 2695-2703.

Harrington, M. J., H. S. Gupta, P. Fratzl and J. H. Waite (2009). "Collagen insulated from tensile damage by domains that unfold reversibly: In situ X-ray investigation of mechanical yield and damage repair in the mussel byssus." Journal of Structural Biology **167**(1): 47-54.

Harrington, M. J. and J. H. Waite (2007). "Holdfast heroics: comparing the molecular and mechanical properties of Mytilus californianus byssal threads." Journal of Experimental Biology **210**(24): 4307-4318.

Hatzfeld, M. and K. Weber (1990). "The coiled coil of in vitro assembled keratin filaments is a heterodimer of type I and II keratins: use of site-specific mutagenesis and recombinant protein expression." The Journal of Cell Biology **110**(4): 1199-1210.

Hearle, J. W. S. (1967). "The structural mechanics of fibers." Journal of Polymer Science Part C: Polymer Symposia **20**(1): 215-251.

Hearle, J. W. S. (1969). "Chapman Mechanical Model for Wool and Other Keratin Fibers." Textile Research Journal **39**(12): 1109-&.

Hearle, J. W. S. (2000). "A critical review of the structural mechanics of wool and hair fibres." International Journal of Biological Macromolecules **27**(2): 123-138.

Hess, B., C. Kutzner, D. van der Spoel and E. Lindahl (2008). "GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation." Journal of Chemical Theory and Computation **4**(3): 435-447.

Hofbauer, F. and I. Frank (2010). "Disulfide Bond Cleavage: A Redox Reaction Without Electron Transfer." Chemistry-a European Journal **16**(17): 5097-5101.

Hogg, P. J. (2003). "Disulfide bonds as switches for protein function." Trends in Biochemical Sciences **28**(4): 210-214.

Holten-Andersen, N., M. J. Harrington, H. Birkedal, B. P. Lee, P. B. Messersmith, K. Y. C. Lee and J. H. Waite (2011). "pH-induced metal-ligand cross-links inspired by mussel yield self-healing polymer networks with near-covalent elastic moduli." Proceedings of the National Academy of Sciences of the United States of America **108**(7): 2651-2655.

http://en.wikipedia.org/wiki/Ramachandran_plot.

Humphrey, W., Dalke, A. and Schulten, K. (1996). "VMD - Visual Molecular Dynamics." J. Molec. Graphics **14**: 6.

Hwang, D. S., H. Zeng, A. Masic, M. J. Harrington, J. N. Israelachvili and J. H. Waite (2010). "Protein- and Metal-dependent Interactions of a Prominent Protein in Mussel Adhesive Plaques." Journal of Biological Chemistry **285**(33): 25850-25858.

Hyde, S. (1997). The Language of shape : the role of curvature in condensed matter--physics, chemistry, and biology, Amsterdam [Netherlands] ; New York : Elsevier, 1997.

Im, W. P., M. S. Lee and C. L. Brooks (2003). "Generalized born model with a simple smoothing function." Journal of Computational Chemistry **24**(14): 1691-1702.

Isralewitz, B., M. Gao and K. Schulten (2001). "Steered molecular dynamics and mechanical functions of proteins." Current opinion in structural biology **11**(2): 224-230.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein (1983). "Comparison of simple potential functions for simulating liquid water." The Journal of Chemical Physics **79**(2): 926-935.

Jorgensen, W. L., D. S. Maxwell and J. TiradoRives (1996). "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids." Journal of the American Chemical Society **118**(45): 11225-11236.

Kajiura, Y., S. Watanabe, T. Itou, K. Nakamura, A. Iida, K. Inoue, N. Yagi, Y. Shinohara and Y. Amemiya (2006). "Structural analysis of human hair single fibres by scanning microbeam SAXS." Journal of Structural Biology **155**(3): 438-444.

Keten, S. and M. J. Buehler (2008). "Geometric Confinement Governs the Rupture Strength of H-bond Assemblies at a Critical Length Scale." Nano Lett. **8**(2): 743-748.

Keten, S., C. C. Chou, A. C. T. van Duin and M. J. Buehler (2012). "Tunable nanomechanics of protein disulfide bonds in redox microenvironments." Journal of the Mechanical Behavior of Biomedical Materials **5**(1): 32-40.

Khandogin, J. and C. L. Brooks (2005). "Constant pH molecular dynamics with proton tautomerism." Biophysical Journal **89**(1): 141-157.

Khandogin, J. and C. L. Brooks (2006). "Toward the accurate first-principles prediction of ionization equilibria in proteins." Biochemistry **45**(31): 9363-9373.

Khandogin, J. and C. L. Brooks (2007). "Linking folding with aggregation in Alzheimer's β-amyloid peptides." Proceedings of the National Academy of Sciences **104**(43): 16880-16885.

Kitahara, T. and H. Ogawa (1991). "The extraction and characterization of human nail keratin." Journal of Dermatological Science **2**(6): 402-406.

Kreplak, L., J. Doucet and F. Briki (2001). "Unraveling double stranded alpha-helical coiled coils: An x-ray diffraction study on hard alpha-keratin fibers." Biopolymers **58**(5): 526-533.

Kreplak, L., J. Doucet, P. Dumas and F. Briki (2004). "New Aspects of the α-Helix to β-Sheet Transition in Stretched Hard α-Keratin Fibers." Biophysical Journal **87**(1): 640-647.

Kreplak, L., A. Franbourg, F. Briki, F. Leroy, D. Dallé and J. Doucet (2002). "A New Deformation Model of Hard α-Keratin Fibers at the Nanometer Scale: Implications for Hard α-Keratin Intermediate Filament Mechanical Properties." Biophysical Journal **82**(4): 2265-2274.

La Mendola, D., A. Magrì, A. M. Santoro, V. G. Nicoletti and E. Rizzarelli (2012). "Copper(II) interaction with peptide fragments of histidine–proline-rich glycoprotein: Speciation, stability and binding details." Journal of Inorganic Biochemistry **111**(0): 59-69.

Laioa, A. and M. Parrinello (2002). "Escaping free-energy minima." P. Natl. Acad. Sci. USA **99**: 12562-12566.

Lazaridis, T. and M. Karplus (1997). ""New View" of Protein Folding Reconciled with the Old Through Multiple Unfolding Simulations." Science **278**(5345): 1928-1931.

Lazaridis, T. and M. Karplus (1999). "Effective energy function for proteins in solution." Proteins: Structure, Function, and Bioinformatics **35**(2): 133-152.

Lee, H., N. F. Scherer and P. B. Messersmith (2006). "Single-molecule mechanics of mussel adhesion." Proceedings of the National Academy of Sciences **103**(35): 12999-13003.

Lee, M. S., F. R. Salsbury and C. L. Brooks (2004). "Constant-pH molecular dynamics using continuous titration coordinates." Proteins-Structure Function and Bioinformatics **56**(4): 738-752.

Leitner, V. M., G. F. Walker and A. Bernkop-Schnurch (2003). "Thiolated polymers: evidence for the formation of disulphide bonds with mucus glycoproteins." European Journal of Pharmaceutics and Biopharmaceutics **56**(2): 207-214.

Li, W. and F. Grater (2010). "Atomistic evidence of how force dynamically regulates thiol/disulfide exchange." J Am Chem Soc **132**(47): 16790-16795.

Lichtenegger, H. C., H. Birkedal, D. M. Casa, J. O. Cross, S. M. Heald, J. H. Waite and G. D. Stucky (2005). "Distribution and Role of Trace Transition Metals in Glycera Worm Jaws Studied with Synchrotron Microbeam Techniques." Chemistry of Materials **17**(11): 2927-2931.

Lichtenegger, H. C., T. Schoberl, J. T. Ruokolainen, J. O. Cross, S. M. Heald, H. Birkedal, J. H. Waite and G. D. Stucky (2003). "Zinc and mechanical prowess in the jaws of Nereis, a marine worm." Proceedings of the National Academy of Sciences of the United States of America **100**(16): 9144-9149.

Liovic, M., J. Stojan, P. E. Bowden, D. Gibbs, A. Vahlquist, E. B. Lane and R. Komel (2001). "A novel keratin 5 mutation (K5V186L) in a family with EBS-K: a conservative substitution can lead to development of different disease phenotypes." Journal of Investigative Dermatology **116**(6): 964-969.

Lupas, A. (1996). "Prediction and analysis of coiled-coil structures." Methods in enzymology **266**: 513-525.

Lupas, A., M. Van Dyke and J. Stock (1991). "Predicting coiled coils from protein sequences." Science (New York, N.Y.) **252**(5010): 1162-1164.

Lv, L.-P., J.-P. Xu, X.-S. Liu, G.-Y. Liu, X. Yang and J. Ji (2010). "Disulfide-Crosslinked Biomimetic Micelles: Formation, Thiol Reactivity and Cytotoxicity Behavior." Macromolecular Chemistry and Physics **211**(21): 2292-2300.

Ma, L., S. Yamada, D. Wirtz and P. A. Coulombe (2001). "A 'hot-spot' mutation alters the mechanical properties of keratin filament networks." Nat Cell Biol **3**(5): 503-506.

MacKerell, A. D., D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus (1998). "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins." The Journal of Physical Chemistry B **102**(18): 3586-3616.

Mackerell, A. D., M. Feig and C. L. Brooks (2004). "Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations." Journal of Computational Chemistry **25**(11): 1400-1415.

Marchuk, D., S. McCrohon and E. Fuchs (1985). "Complete sequence of a gene encoding a human type I keratin: sequences homologous to enhancer elements in the regulatory region of the gene." Proceedings of the National Academy of Sciences **82**(6): 1609-1613.

Marrink, S. J., H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries (2007). "The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations." The Journal of Physical Chemistry B **111**(27): 7812-7824.

Mayans, O., J. Wuerges, S. Canela, M. Gautel and M. Wilmanns (2001). "Structural evidence for a possible role of reversible disulphide bridge formation in the elasticity of the muscle protein titin." Structure **9**(4): 331-340.

McDonnell, A. V., T. Jiang, A. E. Keating and B. Berger (2006). "Paircoil2: improved prediction of coiled coils from sequence." Bioinformatics **22**(3): 356-358.

Monticelli, L., S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman and S.-J. Marrink (2008). "The MARTINI Coarse-Grained Force Field: Extension to Proteins." Journal of Chemical Theory and Computation **4**(5): 819-834.

Mucke, N., L. Kreplak, R. Kirmse, T. Wedig, H. Herrmann, U. Aebi and J. Langowski (2004). "Assessing the flexibility of intermediate filaments by atomic force microscopy." Journal of Molecular Biology **335**(5): 1241-1250.

Naito, S. and K. Arai (1996). "Type and location of SS linkages in human hair and their relation to fiber properties in water." Journal of Applied Polymer Science **61**(12): 2113-2118.

Napoli, A., M. Valentini, N. Tirelli, M. Muller and J. A. Hubbell (2004). "Oxidation-responsive polymeric vesicles." Nature Materials **3**(3): 183-189.

Nielson, K. D., A. C. T. van Duin, J. Oxgaard, W. Q. Deng and W. A. Goddard (2005). "Development of the ReaxFF reactive force field for describing transition metal catalyzed reactions, with application to the initial stages of the catalytic formation of carbon nanotubes." Journal of Physical Chemistry A **109**(3): 493-499.

North, A. C. T., P. M. Steinert and D. A. D. Parry (1994). "Coiled-coil stutter and link segments in keratin and other intermediate filament molecules: A computer modeling study." Proteins: Structure, Function, and Bioinformatics **20**(2): 174-184.

Nozaki, Y. and C. Tanford (1967). [84] Examination of titration behavior. Methods in Enzymology. C. H. W. Hirs, Academic Press. **Volume 11:** 715-734.

Obarzanek-Fojt, M., B. Favre, M. Huber, S. Ryser, A. M. Moodycliffe, P. J. Wipff, B. Hinz and D. Hohl (2011). "Induction of p38, tumour necrosis factor-α and RANTES by mechanical stretching of keratinocytes expressing mutant keratin 10R156H." British Journal of Dermatology **164**(1): 125-134.

Oxenham, W. (1989). "The mechanics of wool structures R. Postle, G. A. Carnaby and S. de Jong, Ellis Horwood, Chichester, 1988. pp. 462, price £59.50. ISBN 0-7458-0322-9." British Polymer Journal **21**(3): 279-279.

Pande, A., D. Gillot and J. Pande (2009). "The Cataract-Associated R14C Mutant of Human γD-Crystallin Shows a Variety of Intermolecular Disulfide Cross-Links: A Raman Spectroscopic Study." Biochemistry **48**(22): 4937-4945.

Pappalardo, G., G. Impellizzeri, R. P. Bonomo, T. Campagna, G. Grasso and M. G. Saita (2002). "Copper(ii) and nickel(ii) binding modes in a histidine-containing model dodecapeptide." New Journal of Chemistry **26**(5): 593-600.

Paquin, R. and P. Colomban (2007). "Nanomechanics of single keratin fibres: A Raman study of the alpha-helix ->beta-sheet transition and the effect of water." Journal of Raman Spectroscopy **38**(5): 504-514.

Parbhu, A. N., W. G. Bryson and R. Lal (1999). "Disulfide Bonds in the Outer Layer of Keratin Fibers Confer Higher Mechanical Rigidity: Correlative Nano-Indentation and Elasticity Measurement with an AFM." Biochemistry **38**(36): 11755-11761.

Park, A. C. and C. B. Baddiel (1972). "Rheology of Stratum Corneum .1. Molecular Interpretation of Stress-Strain Curve." Journal of the Society of Cosmetic Chemists **23**(1): 3-12.

Parry, D. A. D., L. N. Marekov and P. M. Steinert (2001). "Subfilamentous Protofibril Structures in Fibrous Proteins." Journal of Biological Chemistry **276**(42): 39253-39258.

Parry, D. A. D., L. N. Marekov, P. M. Steinert and T. A. Smith (2002). "A Role for the 1A and L1 Rod Domain Segments in Head Domain Organization and Function of Intermediate Filaments: Structural Analysis of Trichocyte Keratin." Journal of Structural Biology **137**(1–2): 97-108.

Parry, D. A. D. and P. M. Steinert (1999). "Intermediate filaments: molecular architecture, assembly, dynamics and polymorphism." Quarterly Reviews of Biophysics **32**(02): 99-187.

Parry, D. A. D., S. V. Strelkov, P. Burkhard, U. Aebi and H. Herrmann (2007). "Towards a molecular description of intermediate filament structure and assembly." Experimental Cell Research **313**(10): 2204-2216.

Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel and P. Kollman (1995). "Amber, a Package of Computer-Programs for Applying Molecular Mechanics, Normal-Mode Analysis, Molecular-Dynamics and Free-Energy Calculations to Simulate the Structural and Energetic Properties of Molecules." Computer Physics Communications **91**(1-3): 1-41.

Pharr, G. M., W. C. Oliver and F. R. Brotzen (1992). "On the generality of the relationship among contact stiffness, contact area, and elastic modulus during indentation." Journal of Materials Research **7**(03): 613-617.

Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten (2005). "Scalable molecular dynamics with NAMD." Journal of Computational Chemistry **26**(16): 1781-1802.

Plimpton, S. (1995). "Fast Parallel Algorithms for Short-Range Molecular Dynamics." Journal of Computational Physics **117**(1): 1-19.

Ponder, J. W., D. A. Case and D. Valerie (2003). Force Fields for Protein Simulations. Advances in Protein Chemistry, Academic Press. **66:** 27-85.

Pontin, M. G., D. N. Moses, J. H. Waite and F. W. Zok (2007). "A nonmineralized approach to abrasion-resistant biomaterials." Proceedings of the National Academy of Sciences **104**(34): 13559-13564.

Popescu, C. and F.-J. Wortmann (2010). Wool – Structure, Mechanical Properties and Technical Products based on Animal Fibres. Industrial Applications of Natural Fibres, John Wiley & Sons, Ltd: 255-266.

Qin, Z. and M. Buehler (2011). "Structure and dynamics of human vimentin intermediate filament dimer and tetramer in explicit and implicit solvent models." Journal of Molecular Modeling **17**(1): 37-48.

Qin, Z. and M. J. Buehler (2012). "Mechanical properties of crosslinks controls failure mechanism of hierarchical intermediate filament networks." Theoretical and Applied Mechanics Letters **2**(1): 13-014005.

Qin, Z., C.-C. Chou, L. Kreplak and M. J. Buehler (2012). Structural, Mechanical and Functional Properties of Intermediate Filaments from the Atomistic to the Cellular Scales. Advances in Cell Mechanics. S. Li and B. Sun, Springer Berlin Heidelberg: 117-166.

Qin, Z., A. Fabre and M. Buehler (2013). "Structure and mechanism of maximum stability of isolated alpha-helical protein domains at a critical length scale." The European Physical Journal E **36**(5): 1-12.

Qin, Z., L. Kreplak and M. J. Buehler (2009). "Hierarchical Structure Controls Nanomechanical Properties of Vimentin Intermediate Filaments." Plos One **4**(10): e7294.

Rahaman, O., A. C. T. van Duin, W. A. Goddard and D. J. Doren (2010). "Development of a ReaxFF Reactive Force Field for Glycine and Application to Solvent Effect and Tautomerization." The Journal of Physical Chemistry B **115**(2): 249-261.

Robins, G. (1999). "Hair and the Construction of Identity in Ancient Egypt, c. 1480-1350 B.C." Journal of the American Research Center in Egypt **36**: 55-69.

Rogers, M. A., H. Winter, J. Schweizer, L. Langbein, S. Praetzel, I. Moll and T. Krieg (1997). "Sequences and differential expression of three novel human type-II hair keratins." Differentiation **61**(3): 187-194.

Schwaiger, I., C. Sattler, D. R. Hostetter and M. Rief (2002). "The myosin coiled-coil is a truly elastic protein structure." Nat Mater **1**(4): 232-235.

Seshadri, I. P. and B. Bhushan (2008). "In situ tensile deformation characterization of human hair with atomic force microscopy." Acta Materialia **56**(4): 774-781.

Shao, H., K. N. Bachus and R. J. Stewart (2009). "A Water-Borne Adhesive Modeled after the Sandcastle Glue of P. californica." Macromolecular Bioscience **9**(5): 464-471.

Sheu, S.-Y., D.-Y. Yang, H. L. Selzle and E. W. Schlag (2003). "Energetics of hydrogen bonds in peptides." Proceedings of the National Academy of Sciences **100**(22): 12683-12687.

Sindhikara, D., Y. L. Meng and A. E. Roitberg (2008). "Exchange frequency in replica exchange molecular dynamics." Journal of Chemical Physics **128**(2).

Sivaramakrishnan, S., J. L. Schneider, A. Sitikov, R. D. Goldman and K. M. Ridge (2009). "Shear Stress Induced Reorganization of the Keratin Intermediate Filament Network Requires Phosphorylation by Protein Kinase C ζ." Molecular Biology of the Cell **20**(11): 2755-2765.

Smith, T. A. and D. A. D. Parry (2008). "Three-dimensional modelling of interchain sequence similarities and differences in the coiled-coil segments of keratin intermediate filament heterodimers highlight features important in assembly." Journal of Structural Biology **162**(1): 139-151.

Sokolova, A. V., L. Kreplak, T. Wedig, N. Mücke, D. I. Svergun, H. Herrmann, U. Aebi and S. V. Strelkov (2006). "Monitoring intermediate filament assembly by small-angle x-ray scattering reveals the molecular architecture of assembly intermediates." Proceedings of the National Academy of Sciences **103**(44): 16206-16211.

Speakman, J. B. (1927). "38—The intracellular strcuture of the wool fiber." Journal of the Textile Institute Transactions **18**(10): T431-T453.

Srivastava, A., N. Holten-Andersen, G. D. Stucky and J. H. Waite (2008). "Ragworm Jaw-Inspired Metal Ion Cross-Linking for Improved Mechanical Properties of Polymer Blends." Biomacromolecules **9**(10): 2873-2880.

Steinert, P. M. (1990). "The two-chain coiled-coil molecule of native epidermal keratin intermediate filaments is a type I-type II heterodimer." Journal of Biological Chemistry **265**(15): 8766-8774.

Steinert, P. M., L. N. Marekov, R. D. B. Fraser and D. A. D. Parry (1993). "Keratin Intermediate Filament Structure: Crosslinking Studies Yield Quantitative Information on Molecular Dimensions and Mechanism of Assembly." Journal of Molecular Biology **230**(2): 436-452.

Stewart, R. J., C. S. Wang and H. Shao (2011). "Complex coacervates as a foundation for synthetic underwater adhesives." Advances in Colloid and Interface Science **167**(1–2): 85-93.

Strnad, P., V. Usachov, C. Debes, F. Grater, D. A. D. Parry and M. B. Omary (2012). "Unique amino acid signatures that are evolutionarily conserved distinguish simple-type, epidermal and hair keratins." Journal of Cell Science **124**(24): 4221-4232.

Sugita, Y. and Y. Okamoto (1999). "Replica-exchange molecular dynamics method for protein folding." Chemical Physics Letters **314**(1-2): 141-151.

Szeverenyi, I., A. J. Cassidy, C. W. Chung, B. T. Lee, J. E. Common, S. C. Ogg, H. Chen, S. Y. Sim, W. L. Goh, K. W. Ng, J. A. Simpson, L. L. Chee, G. H. Eng, B. Li, D. P. Lunny, D. Chuon, A. Venkatesh, K. H. Khoo, W. I. McLean, Y. P. Lim and E. B. Lane

(2008). "The Human Intermediate Filament Database: comprehensive information on a gene family involved in many human diseases." Human Mutation **29**(3): 351-360.

van der Spoel, D. and M. M. Seibert (2006). "Protein folding kinetics and thermodynamics from atomistic simulations." Physical Review Letters **96**(23).

van Duin, A. C. T., S. Dasgupta, F. Lorant and W. A. Goddard (2001). "ReaxFF: A reactive force field for hydrocarbons." Journal of Physical Chemistry A **105**(41): 9396-9409.

Voet, D. and J. G. Voet (2004). Biochemistry. New York, J. Wiley & Sons.

Waite, J. H. and C. C. Broomell (2012). "Changing environments and structure–property relationships in marine biomaterials." The Journal of Experimental Biology **215**(6): 873-883.

Wallace, J. A. and J. K. Shen (2011). "Continuous Constant pH Molecular Dynamics in Explicit Solvent with pH-Based Replica Exchange." Journal of Chemical Theory and Computation **7**(8): 2617-2629.

Wang, H., D. A. D. Parry, L. N. Jones, W. W. Idler, L. N. Marekov and P. M. Steinert (2000). "In Vitro Assembly and Structure of Trichocyte Keratin Intermediate Filaments." The Journal of Cell Biology **151**(7): 1459-1468.

Wang, J., G. Xu and D. R. Borchelt (2006). "Mapping superoxide dismutase 1 domains of non-native interaction: roles of intra- and intermolecular disulfide bonding in aggregation." Journal of Neurochemistry **96**(5): 1277-1288.

Wiita, A. P., S. R. K. Ainavarapu, H. H. Huang and J. M. Fernandez (2006). "Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques." Proceedings of the National Academy of Sciences of the United States of America **103**(19): 7222-7227.

Wortmann, F.-J. and H. Zahn (1994). "The Stress/Strain Curve of α-Keratin Fibers and the Structure of the Intermediate Filament." Textile Research Journal **64**(12): 737-743.

Yano, S., M. Komine, M. Fujimoto, H. Okochi and K. Tamaki (2004). "Mechanical Stretching In Vitro Regulates Signal Transduction Pathways and Cellular Proliferation in Human Epidermal Keratinocytes." J Investig Dermatol **122**(3): 783-790.

Yoon, G., S. Na and K. Eom (2012). "Loading device effect on protein unfolding mechanics." Journal of Chemical Physics **137**(2).

Yoon, K. H., M. Yoon, R. D. Moir, S. Khuon, F. W. Flitney and R. D. Goldman (2001). "Insights into the Dynamic Properties of Keratin Intermediate Filaments in Living Epithelial Cells." The Journal of Cell Biology **153**(3): 503-516.

Yoshida, M., E. Muneyuki and T. Hisabori (2001). "ATP synthase — a marvellous rotary engine of the cell." Nat Rev Mol Cell Biol **2**(9): 669-677.

Yu, J., D.-W. Yu, D. M. Checkla, I. M. Freedberg and A. P. Bertolino (1993). "Human Hair Keratins." J Investig Dermatol **101**(s1): 56S-59S.

Zahn, H. (2002). "Progress report on hair keratin research." International journal of cosmetic science **24**(3): 163-169.

Zhao, H. and J. H. Waite (2006). "Proteins in Load-Bearing Junctions: The Histidine-Rich Metal-Binding Protein of Mussel Byssus." Biochemistry **45**(47): 14223-14231.

Zhou, R. and B. J. Berne (2002). "Can a continuum solvent model reproduce the free energy landscape of a β-hairpin folding in water?" Proceedings of the National Academy of Sciences **99**(20): 12777-12782.