The Role of Type I Collagen Heterotrimers and Homotrimers in Mechanical Strength and Collagen Cleavage

by

Shu-Wei Chang

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

September 2014

© 2014 Massachusetts Institute of Technology. All Rights Reserved.

Signature redacted

Signature redacted

Certified by:

Markus J. Buehler
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by:

Heidi M. Nepf
Chair, Departmental Committee for Graduate Students
The Role of Type I Collagen Heterotrimers and Homotrimers in Mechanical Strength and Collagen Cleavage

by

Shu-Wei Chang

Submitted to the Department of Civil and Environmental Engineering on August 7, 2014 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

Collagen is a crucial structural protein, formed through a hierarchical assembly of molecules, arranged in collagen fibrils, which constitutes the basis for larger-scale fibers. Normal type I collagen is a heterotrimer triple-helical molecule consisting of two alpha-1 chains and one alpha-2 chain. A mouse model of the genetic brittle bone disease, osteogenesis imperfecta, oim, is characterized by a replacement of the alpha-2 chain by an alpha-1 chain, resulting in a homotrimer collagen molecule. Experimental studies of oim mice tendon and bone have shown reduced mechanical strength compared to normal mice. The relationship between the molecular content and the decrease in strength is, however, still unknown. In this thesis, we use a bottom-up molecular simulation approach to examine the role of type I normal collagen and oim collagen from a single collagen molecule to collagen microfibril to mineralized collagen microfibril. At the molecular level, we find that the replacement of the alpha-2 chain results in a collagen molecule with more kinks and a more thermally stable cleavage site. The higher thermal stability of the cleavage site of the homotrimer explains the enzyme resistances of homotrimers. Furthermore, we reveal a molecular mechanism of force induced stabilization of collagen against enzymatic breakdown for the heterotrimer. At the fibril level, we find that the kinks affect the packing of collagen molecules. The homotrimer microfibril has a less dense packing of collagen molecules which leads to a reduced modulus. The alterations on the assembly of collagen molecules further alter the space for mineral deposition at the mineralized collagen fibril level. We find that the mineralized homotrimer collagen fibril has more space for mineral deposition but the mineral size is smaller because the kinks at the molecular level result in a more discontinuous space for mineral deposition. The mineralized oim collagen microfibril has a reduced modulus due to the alterations on the collagen assembly and mineral deposition. Our results provide fundamental insight into the effect of the loss of alpha-2 chain at the molecular level and help understanding the molecular origin of many diseases such as the brittle bone at much larger length-scales.

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:


Acknowledgements

I would like to express my gratitude for many who have supported me during my studies at MIT. 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 wife.

First and foremost, I would like to express my appreciation to my advisor, Professor Markus J. Buehler, for his support, patience, and encouragement throughout my graduate studies. His advice was essential to the completion of this dissertation. He has taught me countless lessons and insights on the workings of academic research in general. Without his supervision and constant help, this dissertation would not have been possible.

I would like to thank my committee members, Professor Kausel, Professor Grodzinsky and Professor Shefelbine, for providing many valuable comments that improved the presentation and contents of this dissertation. Especially, I would like to thank Professor Shefelbine for many valuable and inspiring discussions in the past few years. I would also like to thank many of my collaborators, Professor Ruberti, Dr. Gautieri, Dr. Flynn and Dr. Mašić, for their help.

In my daily work I have been blessed with a friendly and cheerful group of my peers in LAMM. I would like to thank them for their friendship, never-ending support and encouragement in the past years.

Finally, I would like to express my deepest gratitude to my family, who have supported and encouraged me in all my academic endeavors. I would like to thank my wife Chia-Ching who has been a constant source of love, support and encouragement. I am thankful that we were able to experience this together.

This research was funded by an MIT Presidential Fellowship, and grants by the NSF, DOD-PECASE, ONR, AFOSR, and NIH U01. Their support is greatly appreciated.
To my wife, for inspiration and loving support
Contents

1 Introduction ............................................................................................................. 18
   1.1 Background and objectives ........................................................................... 18
   1.2 Research approach and area ......................................................................... 21
   1.3 Organization of the thesis ........................................................................... 23

2 Literature review .................................................................................................... 25
   2.1 Overview of collagen materials ..................................................................... 25
      2.1.1 Structure of collagen molecules ......................................................... 27
      2.1.2 Hierarchical structure of collagen ...................................................... 29
   2.2 Collagen mechanics ....................................................................................... 30
      2.2.1 Mechanical properties of a single collagen molecule ......................... 30
      2.2.2 Mechanical properties of collagen fibrils ......................................... 32
   2.3 Mutations and diseases in collagenous tissues ............................................ 34
      2.3.1 Single residue mutations in brittle bone disease ............................... 34
      2.3.2 The osteogenesis imperfecta mouse model ...................................... 35
   2.4 Collagen degradation ..................................................................................... 37
      2.4.1 Collagen degradation mechanisms ..................................................... 37
      2.4.2 Mechanical force effects on the collagen degradation rate ............... 38
      2.4.3 Homotrimer resistance of collagen degradation ............................... 40

3 Methodology ......................................................................................................... 42
   3.1 Atomistic modeling ....................................................................................... 42
      3.1.1 Atomic description of classical molecular dynamics ......................... 42
      3.1.2 Newtonian dynamics ......................................................................... 43
      3.1.3 Force fields for biomaterials .............................................................. 45
   3.2 Collagen sequences ....................................................................................... 47
   3.3 Steered molecular dynamics ........................................................................ 48
   3.4 Persistence length calculation ...................................................................... 50
   3.5 Structure of collagen molecules ................................................................... 52
4 Structural and mechanical differences between wild type and *oim* collagen molecules ................................................. 55

4.1 Introduction .................................................................................................. 55

4.2 Model construction ....................................................................................... 55

4.2.1 Collagen molecule generation ................................................................. 55

4.2.2 All-atom equilibration and mechanical test ............................................ 57

4.2.3 Analysis of the structures of collagen molecules .................................... 58

4.2.4 Analysis of the freely rotating behavior .................................................. 60

4.3 Results .......................................................................................................... 60

4.3.1 Persistence length calculations ................................................................. 60

4.3.2 Structural differences between the heterotrimer and homotrimer ............ 63

4.3.3 Analysis of the kinks of the heterotrimer and homotrimer ....................... 67

4.3.4 Mechanical tests of the heterotrimer and the homotrimer ....................... 69

4.3.5 Repeated simulations ............................................................................ 72

4.4 Summary ....................................................................................................... 74

5 The role of alpha-2 chain and mechanical force effects in collagen cleavage ................................................................. 78

5.1 Introduction ................................................................................................... 78

5.2 Simulation details .......................................................................................... 81

5.2.1 Collagen molecule generation ................................................................. 81

5.2.2 All-atom modeling and equilibration ....................................................... 82

5.2.3 In silico mechanical tests ........................................................................ 83

5.2.4 Analysis of the structures of collagen molecules .................................... 83

5.2.5 Analysis of H-bonds at the cleavage site ................................................. 83

5.3 Results .......................................................................................................... 84

5.3.1 Molecular structures at the cleavage site ............................................... 84

5.3.2 Mechanical testing of the cleavage sites ............................................... 88

5.3.3 Analysis of H-bonds .............................................................................. 91

5.3.4 Analysis of the cleavage kinetics ............................................................. 92
6 Structural and mechanical differences between wild type and oim collagen microfibrils

6.1 Introduction

6.2 Materials and methods
   6.2.1 Atomistic collagen microfibril
   6.2.2 All-atom equilibration
   6.2.3 Mechanical testing

6.3 Results
   6.3.1 Mechanical properties of hydrated collagen microfibrils
   6.3.2 The oim collagen microfibrils have larger lateral spacing
   6.3.3 Mechanical properties of dry collagen microfibrils
   6.3.4 Structural analysis of the normal and oim collagen microfibrils

6.4 Summary

7 Mineralized collagen microfibrils of wild type and oim collagen

7.1 Introduction

7.2 Materials and methods
   7.2.1 Model construction
   7.2.2 Mineral density in mineralized collagen fibrils
   7.2.3 Full atomistic simulations
   7.2.4 Analysis of the mineral distribution and size

7.3 Results
   7.3.1 Mineral distributions in the normal and oim collagen microfibrils
   7.3.2 Analysis of the mineral sizes
   7.3.3 Mechanical properties of mineralized collagen microfibrils
   7.3.4 Analysis of the deformation mechanisms

7.4 Summary

8 Conclusions and opportunities for future research
8.1 Summary of key findings and significances ..................................................... 140
8.2 Opportunities for future research ................................................................. 142

References ......................................................................................................... 144
List of figures

Figure 1.1 Hierarchical structure of collagenous tissues and the proposed research...... 20
Figure 2.1 Structure of collagenous tissues and a schematic of biomechanics of collagen materials........................................................................................................................................... 26
Figure 2.2 Molecular structure of a collagen molecule, which consists of three chains forming a triple helical structure............................................................................................................................ 27
Figure 2.3 Calculated relative thermal stability of type I and type II collagen along the twisting axis ........................................................................................................................................... 28
Figure 2.4 A periodic cell of a collagen fibril model................................................... 29
Figure 2.5 Mechanical response of a single collagen molecule................................. 31
Figure 2.6 Mechanical properties of collagen molecules and fibrils.......................... 33
Figure 2.7 Mutations throughout the entire type I collagen molecule which have been identified to cause severe, moderate and mild OI................................................................. 34
Figure 2.8 Experimental findings on the difference between wild-type and oim collagen. ........................................................................................................................................... 36
Figure 3.1 Point representations of a material in molecular dynamics simulations........ 42
Figure 3.2 Illustrations of the contributions of the different terms in the potential expressions given in Eq. (3.11)....................................................................................... 45
Figure 3.3 Illustration of the use of steered molecular dynamics to apply mechanical load ........................................................................................................................................... 49
Figure 3.4 Illustration of the geometric parameters in the freely rotating chain model... 50
Figure 3.5 Structure of collagen protein materials....................................................... 53
Figure 4.1 Initial structures of the heterotrimer and homotrimer collagen molecules..... 57
Figure 4.2 The relation between the separated distance and the loss of the direction of the tangent........................................................................................................................................... 61
Figure 4.3 Statistics of structural parameters of collagen protein. ............................... 62
Figure 4.4 Unit height distributions of heterotrimer and homotrimer at 0, 20, 30, 40 and 50 ns........................................................................................................................................... 64
Figure 4.5 Time history of unit heights of heterotrimer and homotrimer for different regions of the segment analyzed ........................................................................................................................................... 65
Figure 4.6 Snapshots of collagen protein. Each chain is plotted in ribbon with the same color. ................................................................. 66
Figure 4.7 The comparison of the kink and freely rotating behaviors between heterotrimer and homotrimer collagen molecules ......................................................... 68
Figure 4.8 Mechanical tests of the heterotrimer and the homotrimer .................... 70
Figure 4.9 Direct comparisons of the unit heights of heterotrimer and homotrimer between the repeated and the original simulations ................................. 71
Figure 4.10 Direct comparisons of the distances between Cα atoms of heterotrimer and homotrimer between the repeated and the original simulations ............. 72
Figure 4.11 Direct comparisons of the radii of heterotrimer and homotrimer between the repeated and the original simulations .................................................. 73
Figure 4.12 The projection of the rotating vectors on the y-z plane for (3,6,9) of the repeated and the original simulations .................................................... 74
Figure 4.13 The role of the kink of collagen molecule in the microfibril ................. 75
Figure 5.1 Comparisons of the cleavage rates between the heterotrimer and the homotrimer without and with applying forces ............................................ 80
Figure 5.2 Snapshots of collagen molecules for heterotrimer (top) and homotrimer (bottom) with and without force ......................................................... 85
Figure 5.3 Structural analyses of the vicinity of cleavage site of a type I mouse heterotrimer and homotrimer collagen molecule without applied force ............... 86
Figure 5.4 Changes in the lengths of the cleavage sites with applying force .......... 87
Figure 5.5 Structural analyses of the vicinity of cleavage site of type I mouse heterotrimer and homotrimer with constant force applied ............................ 89
Figure 5.6 Analyses of the N-O distance at the vicinity of the cleavage site of heterotrimer and homotrimer with and without forces ....................................... 90
Figure 5.7 Possible cleavage mechanisms of collagen molecules ....................... 94
Figure 5.8 Analysis of the Michaelis constant and degradation rate ................... 96
Figure 6.1 Stress-strain curves of normal and oim collagen microfibrils along the axial direction ................................................................. 102
Figure 6.2 Analysis of the a-axis lengths of the normal and oim collagen microfibrils at different stress levels ................................................................. 103
Figure 6.3 Analysis of the b-axis lengths of the normal and oim collagen microfibrils at different stress levels........................................................................................................... 104
Figure 6.4 Analysis of the end-to-end distances of the normal and oim collagen microfibrils at different stress levels........................................................................................................... 105
Figure 6.5 Analysis of the average unit height of the normal and oim collagen microfibrils at different stress levels........................................................................................................... 106
Figure 6.6 Stress-strain curves of the normal and oim collagen microfibrils in hydrated and dry conditions........................................................................................................... 108
Figure 6.7 Analysis of the a-axis lengths of the dehydrated normal and oim collagen microfibrils at different stress levels........................................................................................................... 110
Figure 6.8 Analysis of the b-axis lengths of the dehydrated normal and oim collagen microfibrils at different stress levels........................................................................................................... 111
Figure 6.9 Stress-strain curves of compression tests on the microfibril models.............. 112
Figure 6.10 Direct comparison of the structures of normal and oim collagen molecules................................................................. 114
Figure 6.11 Structures of normal collagen in x-y plane, several kinks are observed while the collagn molecules mostly align with the end-to-end direction (shown by solid line) of the collagen molecule. ................................................................. 114
Figure 6.12 Structures of normal collagen in x-z plane, several kinks are observed while the collagn molecules mostly align with the end-to-end direction (shown by solid line) of the collagen molecule. ................................................................. 115
Figure 6.13 Structures of oim collagen in x-y plane, oim collagen molecule has more kinks (mostly in the gap region), which lead to the increase of lateral distance and the reduction of modulus. ........................................................................................................... 115
Figure 6.14 Structures of oim collagen in x-z plane, oim collagen molecule has more kinks (mostly in the gap region), which lead to the increase of lateral distance and the reduction of modulus. ........................................................................................................... 116
Figure 6.15 Direct comparison between the structures of normal and oim collagen microfibrils................................................................................................................................. 117
Figure 6.16 Analysis the twisting of the normal and oim collagen molecule by analyzing the 3D structures of the Gly residues........................................................................................................... 117
Figure 6.17 Analysis of the projections of the centers of masses of Glycine on the plan perpendicular to the end-to-end direction of collagen molecule. ........................................ 118
Figure 7.1 A bright field TEM image of longitudinal section of human bone. .......... 121
Figure 7.2 Structure of normal and oim collagen microfibrils and the hydroxyapatite crystal................................................................. 123
Figure 7.3 Constructing mineralized collagen microfibril models based on the equilibrium structures obtained from full atomistic simulations. ...................... 125
Figure 7.4 Relation between the mineral density of bone ($p_{\text{bone}}$) and the mineral density of a mineralized collagen fibril ($p_{\text{mineral_coll}}$), under the assumption that 70% of the minerals in bone are external to collagen fibrils .................................................. 126
Figure 7.5 Mineralized normal and oim collagen microfibrils constructed by using the same criterion (N=0.74). ................................................................. 129
Figure 7.6 Equilibrium structures of mineralized normal and oim collagen microfibrils at mineral densities of 25% and 35%. The blue regions indicate the minerals. ........ 130
Figure 7.7 Normalized mineral density distribution along the collagen axis for mineralized normal collagen microfibrils with 25% and 35% mineral densities. ...... 131
Figure 7.8 Normalized mineral density distribution along the collagen axis for mineralized oim collagen microfibrils with 25% and 35% mineral densities. ............ 132
Figure 7.9 Analysis of the largest mineral in mineralized normal and oim collagen microfibril at 25% mineral density: ......................................................... 133
Figure 7.10 Analysis of the largest mineral in mineralized normal and oim collagen microfibril at 35% mineral density ......................................................... 134
Figure 7.11 Stress-strain curves of mineralized collagen microfibrils. The results show that the mineralized oim collagen microfibril has reduced modulus. ............... 136
Figure 7.12 Analysis of the gap/overlap ratio under different tensile stress levels for non-mineralized and mineralized collagen microfibrils................................. 138
Figure 8.1 Summary of the multiscale framework developed in this thesis and the major findings. ............................................................................... 141
List of tables

Table 2-1 Summary of recent experimental studies on mechanical force effects on collagen degradation rate. (Adapted from (Chang and Buehler 2014))........................ 41
Table 4-1 Comparison of kink angles at different regions of heterotrimer and homotrimer collagen molecules. (Chang, Shefelbine et al. 2012).......................................................... 69
1 Introduction

1.1 Background and objectives

Collagen, which is a basic construction material in human body, constitutes one third of the human proteome, providing mechanical stability, elasticity and strength to organisms (Rainey, Wen et al. 2002; Fratzl 2008; Buehler and Yung 2009). Collagen is a crucial structural protein, formed through a hierarchical assembly of tropocollagen molecules, arranged in collagen fibrils, which constitutes the basis for larger-scale fibers (Figure 1.1). Normal type I collagen is a heterotrimer consisting of two alpha-1 chains and one alpha-2 chain. Deviations from the natural structure of collagen in the form of mutations or substitutions of entire chains cause severe changes in the mechanical properties and result in diseases that affect connective tissues such as tendon or bone. A variation of the natural type I collagen molecule is the type I homotrimer, which consists of three alpha-1 chains, and has been found in fetal tissues (Jimenez, Bashey et al. 1977), fibrotic tissues (Ehrlich, Brown et al. 1982), carcinomas (Makareeva, Han et al. 2010), and fetal and cancer cells (Minafra, Luparello et al. 1988; Makareeva, Han et al. 2010) in human. It is also found in a mouse model of the genetic brittle bone disease, osteogenesis imperfecta (OI), the oim mutation of type I collagen. Experimental studies have shown that the mechanical strength of oim bone and tail tendon is significantly less than that of the normal mice (McBride, Choe et al. 1997; Misof, Landis et al. 1997). The oim phenotype has extreme skeletal fragility, reduced stature, and bone deformities, mimicking moderate to severe OI in humans.

Much research has investigated the oim model, not only to understand and develop treatments for osteogenesis imperfecta, but also to probe the function of the alpha-2 chain in collagen. Indeed, the oim model provides an interesting platform for investigating the ramifications of alterations in molecular structure on the mechanics and structural behavior of bone throughout its hierarchy. At the organ level, oim bones are extremely fragile, with smaller cross sections, and little or no plastic deformation resulting in brittle-
type fractures (Saban, Zussman et al. 1996; McBride, Choe et al. 1997; Misof, Landis et al. 1997; Camacho, Hou et al. 1999; Miller, Delos et al. 2007; Vanleene, Saldanha et al. 2011). At the tissue level, *oim* has more disorganized woven bone than organized lamellar bone (Chipman, Sweet et al. 1993). At the matrix level, *oim* bones have higher mineral density, with smaller apatite crystals closely packed together (Camacho, Landis et al. 1996; Fratzl, Paris et al. 1996). All of these changes are caused fundamentally by alterations at the molecular level. Understandings how the mutation at the molecular level results in the brittle behaviour at the tissue level would help developing treatments for a variety of diseases such as brittle bone disease. However, how the molecular changes at single collagen molecular level affects the properties throughout different levels of its hierarchy remains unclear.

Another important aspect of collagen mechanics is its ability to "remodel" — *i.e.*, to grow and break down "on demand" to form tissues where needed. In this context collagen cleavage is crucial for many biological and pathological processes such as wound healing, tissue remodeling, and cancer invasion (Baragi, Qiu et al. 1997; Nagase and Visse 2003; Helary, Ovtracht et al. 2006). Normal physiological remodeling processes involve precisely regulated collagen degradation, where excessive or deficient degradation has been associated with numerous diseases. Accelerated breakdown of collagen has been associated with arthritis, atherosclerotic heart disease, tumor cell invasion, glomerulonephritis, and cell metastasis (Riley, Harrall et al. 1995; Barnes and Farndale 1999; Bode, Mosorin et al. 1999; McDonnell, Morgan et al. 1999). Deficient degradation of collagen has been shown to result in spontaneous abnormal growth plate and increased trabecular bone in mice (Stickens, Behonick et al. 2004).

The mutation at the collagen molecular level alters not only the mechanical properties but also the biological processes of collagenous tissues. Previous studies have shown that the type I heterotrimer and homotrimer have distinct degradation behaviors. Type I homotrimers are found to be resistant to all mammalian collagenases (Narayanan, Meyers et al. 1984; Han, Makareeva et al. 2010; Makareeva, Han et al. 2010), with a cleavage rate much slower for homotrimers than for heterotrimers. The enzyme resistance of homotrimers may play an important role in homotrimer-related diseases or in early
development, during which necessary collagen degradation may be hindered with detrimental results. For example, it has been shown in fibers reconstituted from mouse tail tendon collagen that a minor fraction of homotrimer-based fibers may grow instead of being disassembled during tissue remodeling cycles, which may eventually result in tissue disorganization (Han, Makareeva et al. 2010). Moreover, it has been shown experimentally that the mechanical force applied to collagen molecules controls the cleavage rate (Ellsmere, Khanna et al. 1999; Flynn, Bhole et al. 2010; Adhikari, Chai et al. 2011). Therefore, how to modulate the activity is of great interest and is crucial for developing treatments for a variety of diseases. However, it remains unclear how the loss of the alpha-2 chain in the homotrimer leads to the enzyme resistance at the molecular level and how the mechanical forces mediate the degradation rate of collagen molecules.

Figure 1.1 Hierarchical structure of collagenous tissues and the proposed research. Collagen is a crucial structural protein, formed through a hierarchical assembly of tropocollagen molecules, arranged in collagen fibrils, which constitutes the basis for collagenous tissues. Normal type I collagen (wild type) is a heterotrimer consisting of two alpha-1 chains and one alpha-2 chain. A variation of the natural type I collagen molecule is the type I homotrimer, which consists of three alpha-1 chains, and has been
found in a mouse model of the genetic brittle bone disease, *osteogenesis imperfecta* (OI), the *oim* mutation of type I collagen. In this research, we use a bottom up molecular simulation approach to investigate the mechanisms on how the *oim* mutation affects the behavior at the molecular level and how those effects could lead to altered effects such as brittle bone disease at macroscale. This study covers the understanding of the behavior from a single collagen molecule (I & II) to collagen microfibrils (III) to mineralized collagen microfibrils (IV), helping to define the basic structural and mechanical differences responsible for the changes throughout different levels of its hierarchy.

We are motivated to understand how the mutations at the microscale molecular level alter the macroscale properties including the mechanical properties and the biological functions such as collagen degradation (Figure 1.1). We use a bottom up molecular simulation approach to investigate the mechanism on how the *oim* mutation affects the behavior at the molecular level and how those effects could lead to altered effects such as brittle bone disease at macroscale. We also study the molecular origin of the enzyme resistance of the homotrimer collagen molecular and how the mechanical forces modulate the activity of collagen cleavage. These studies cover the understanding of the behavior from a single collagen molecule to collagen microfibrils to mineralized collagen microfibrils, helping to define the basic structural and mechanical differences responsible for the changes throughout different levels of its hierarchy.

1.2 Research approach and area

In this research, full atomistic simulation, which provides a powerful tool for studying the material properties with molecular details, is used to study the following aspects of collagen mechanics under physiological and diseased conditions:

1. **The role of alpha-2 chain in type I collagen at single collagen molecular level.** We use a molecular simulation approach to study the mechanical and structural differences between small segments of type I heterotrimer and homotrimer collagen molecules. In this research, a comprehensive understanding of the molecular and structural differences between the heterotrimer and homotrimer
collagen molecules are studied to understand the molecular differences at the single collagen molecular level.

II. **The molecular mechanisms of collagen cleavage in the heterotrimer and homotrimer.**

Earlier work has shown that the homotrimer collagen molecule is resistant to collagen cleavage and mechanical force can mediate the cleavage rate. We use atomistic simulations to study the structures of the cleavage sites from both wild type and oim collagen molecules. We focus on understanding the molecular mechanism of the enzyme resistance of the homotrimer and how the mechanical force could mediate the collagen cleavage rate.

III. **Investigation of the structural and mechanical differences between full atomistic microfibril collagen models from wild type and oim mice models.**

It is important to understand how the molecular changes at single collagen molecular level affects the properties throughout different levels of its hierarchy. Therefore, we construct full atomistic collagen microfibril models from wild type and oim mice models to study the alterations of the packing of collagen molecules at the microfibril level resulting from the oim mutation. The mechanical properties of the wild type and the oim microfibrils are computed and the structural differences are compared to provide fundamental insights into how the oim mutations alter the properties of collagen fibrils.

IV. **Constructions of full atomistic models of mineralized collagen microfibrils from wild type and oim mice models to study the molecular origins of brittle bone disease.**

Bone is a composite material consisting of collagen molecules and mineral. Mineral depositions occur after the formation of collagen fibrils. We construct full atomistic models of mineralized collagen microfibrils from wild type and oim mice models based on the equilibrium structures of both wild type and oim collagen microfibrils obtained from III to study the relationship between the change of the building block at the molecular level, oim mutation, and the brittle bone disease at large scale.
The real sequences of type I alpha-1 and type I alpha-2 chains of mus musculus (wild type mouse) are used to generate the collagen molecules. The heterotrimer collagen molecule is built of two alpha-1 chains and one alpha-2 chain while the homotrimer collagen molecule is built of three alpha-1 chains. The sequences of the alpha-1 and alpha-2 chains are adapted from NCBI protein database (http://www.ncbi.nlm.nih.gov/protein): AAH50014.1 for alpha-1 chain and NP_031769.2 for alpha-2 chain. The entire alpha-1 and alpha-2 chains consist of 1,014 residues with repeated G-X-Y triplets, excluding the C-terminal and N-terminal sequences.

The collagen molecules are created by inputting the sequences of three chains into the software THeBuScr (Rainey and Goh 2004), which enables a user to build a triple-helical molecule based on any specified amino acid sequence. The code uses derived conformations from statistical analyses of high-resolution x-ray crystal structures of triple-helical peptides to build collagen molecule structures. For the collagen microfibril models, the models are generated based on the in situ structure of full length collagen type I molecule (Orgel, Irving et al. 2006) (Protein Data Bank identification code 3HR2), which has a triclinic unit cell (~40.0 Å, ~27.0 Å, ~678 Å, α~89.2°, β~94.6°, γ~105.6°). Full atomistic simulations are performed using NAMD (Nelson, Humphrey et al. 1996) and LAMMPS (Plimpton 1995) with the CHARMM force field (MacKerell, Bashford et al. 1998) that includes parameters for hydroxyproline amino acids, based on a model put forth by Anderson (Anderson 2005). This force field has been widely validated for a variety of biochemical models of proteins including collagen (Gautieri, Buehler et al. 2009; Srinivasan, Uzel et al. 2009). More details on the methodology are provided in Chapter 3.

1.3 Organization of the thesis

This study examine the behavior from a single collagen molecule to collagen microfibril to mineralized collagen microfibrils, helping to define the basic structural and mechanical differences responsible for the changes throughout different levels of its hierarchy. We provide understandings on the molecular differences between wild type and oim collagen molecules not only at the single collagen molecule level but the structural and mechanical
differences at the collagen fibril level. This study provides fundamental insights into the effect of the loss of alpha-2 chain at the molecular level and helps understanding the molecular origin of many diseases such as the brittle bone at much larger length-scales. These understandings on the biological functions and mechanical properties of collagen would help developing treatments for a variety of diseases.

The organization of the thesis is as follows: Chapter 2 gives a literature review on the collagen materials, including the hierarchical structure of collagenous tissues, the mechanical properties of collagen, the mutations and diseases in collagen and a review on the collagen degradation and current understandings on how the mechanical forces mediate the collagen degradation rate. Chapter 3 provides details on the methodologies used in this study. In Chapter 4, we describe the differences on the mechanical and structural differences between the wild type and oim collagen at the molecular level. Chapter 5 focuses on the differences on the cleavage sites between the wild type and oim collagen molecules. We provide the molecular origin of the enzyme resistance of the homotrimer and a molecular mechanism of force induced stabilization of collagen against enzymatic breakdown. We scale up to the collagen microfibril level in Chapter 6. The results on the differences between wild type and oim collagen at the microfibril level are discussed. In Chapter 7, we focus on the mineralized collagen microfibril level. We study how the mutation at single collagen molecular level affects the space for mineral deposition. The differences on the mineral distributions and mechanical properties between the mineralized wild type and oim collagen microfibrils are provided. Chapter 8 summarizes the major findings of the thesis and discusses possible future research directions.
2 Literature review

2.1 Overview of collagen materials

Collagen molecules are the basic component of collagenous tissues, such as tendon and bone, which provide mechanical stability, elasticity and strength to organisms (Rainey, Wen et al. 2002; Eppell, Smith et al. 2006; Fratzl 2008; Shen, Dodge et al. 2008; Buehler and Yung 2009; Tang, Ballarini et al. 2010; Nair, Gautieri et al. 2013). Unlike many engineering materials, collagen materials are 'smart' materials that have the ability to adapt their properties in response to mechanical forces through altering their structures from the molecular level up (KJÆR 2004; Wang 2006; Magnusson, Langberg et al. 2010). They are able to convert mechanical forces into biochemical signals that control many biological and pathological processes such as wound healing and tissue remodeling (Figure 2.1). For example, appropriate physical training increases the cross-sectional area and the tensile strength of tendons (Tipton, Matthes et al. 1975; Suominen, Kiiskinen et al. 1980; Langberg, Rosendal et al. 2001), while inappropriate physical training can lead to tendon injuries (Khan and Maffulli 1998; Maffulli, Khan et al. 1998).

Mechanical loading is important in collagenous tissue formation and remodeling (Prajapati, Chavally-Mis et al. 2000). It is understood that physical activity influences both collagen synthesis and degradation (Magnusson, Langberg et al. 2010). Previous study has shown that both collagen formation and degradation increase initially after exercise in humans (Huang and Yannas 1977). A net collagen synthesis is found 36-72 hours after exercise. These results show that collagen degradation is a fundamental event in connective tissue growth and remodeling (Huang and Yannas 1977). At single molecule level, collagen molecules alter their structures in response to mechanical forces, providing signals to mediate the collagen degradation rate.

About 28 types of collagen have been identified while type I collagen, which is found in tendon, skin, teeth, cornea and bone, is the most abundant collagen in the human body. Collagen is a crucial structural protein, formed through a hierarchical assembly of tropocollagen molecules, arranged in collagen fibrils, which constitutes the basis for
larger-scale fibers (Figure 2.1). Normal type I collagen is a heterotrimer consisting of two alpha-1 chains and one alpha-2 chain. As collagen molecules are the basic building blocks, the properties of collagenous tissues at macroscale are known to be affected by the properties of collagen molecules, governed by their sequences, at microscales. Changes on the sequences at molecular level would alter not only the mechanical properties but also the biological processes of collagen materials. Mutations at the molecular level have been associated with various diseases, such as brittle bone disease (Beck, Chan et al. 2000).

Figure 2.1 Structure of collagenous tissues and a schematic of biomechanics of collagen materials. Collagenous tissues are hierarchical structures composed of self-assembled collagen molecules. Adequate loading enables the adaptation to strengthen the collagenous tissue while inadequate loading may lead to injuries. (Adapted from (Chang and Buehler 2014); Hierarchical structure of connective tissues reprinted with permission from Gautieri, A., et al., Hierarchical Structure and Nanomechanics of Collagen Microfibrils from the Atomistic Scale Up. Nano Letters, 2011. Copyright 2011 American Chemical Society).
2.1.1 Structure of collagen molecules

Collagen is the most abundant protein in the vertebrate and is the basic component of connective tissues. It provides mechanical strength and biological functions for connective tissues (Rainey, Wen et al. 2002; Eppell, Smith et al. 2006; Fratzl 2008; Buehler and Yung 2009). A collagen molecule consists of three chains stacked alongside forming a triple helical structure as shown in Figure 2.2. Each chain of a collagen molecule consists of amino acids and has a characteristic repeating sequence of (Gly-X-Y)ₙ. The spheres in Figure 2.2 represent the alpha-carbon atoms of Gly residues, which are mostly located within the center of a collagen molecule. The triple helical structure is stabilized by hydrogen bonds between each chain. A type I collagen molecule has a length of about 300 nm with a radius of approximately 1.6 nm.

![Figure 2.2 Molecular structure of a collagen molecule, which consists of three chains forming a triple helical structure. Each chain is plotted with one color and the spheres represent the alpha-carbon atoms of glycine residues. The collagen molecule has a characteristic sequence of (Gly-X-Y)ₙ. The triple helical structure is stabilized by inter-chain hydrogen bonds. (Tarakanova, Chang et al. 2014)](image)

The collagen molecule is a heterogeneous structure along its twisting axis due to the variation of amino acid sequence. The variation of sequence is crucial for varied biological functions along each segment of the collagen molecule. A collagen molecule has a varying unit height of ~0.853 nm for iminorich regions and ~0.865 nm for aminorich regions (Rainey and Goh 2002). Triple helix builders, such as
GENCOLLAGEN (Huang, Couch et al. 1998) and TTheBuScr (Rainey and Goh 2004), have been developed recently to create idealized atomic coordinates by using sequence information of collagen. Variation of sequence also affects the thermal stability of local conformation of a collagen molecule. Persikov et al. have developed a thermal stability calculator for collagen molecules based on experimental measurements (Persikov, Ramshaw et al. 2005). Figure 2.3 shows examples of relative thermal stability profiles for full length type I and type II collagen molecules. It identifies a non-uniform distribution of thermal stability along the entire length of the molecule. The highest thermal stability are found at the N- and C-terminals, which impacts the mechanical and biological properties. For example, the two regions with low stability in the type I collagen molecule (indicated by arrows in the figure) are identified to be the cross-linking sites.

![Thermal Stability Profiles](image)

Figure 2.3 Calculated relative thermal stability of type I and type II collagen along the twisting axis. Collagen molecule is a heterogeneous material along its length. The variation of sequence impacts its material and biological properties. (Reprinted from (Persikov, Ramshaw et al. 2005) with permission)
2.1.2 Hierarchical structure of collagen

Collagen molecules are the basic building blocks of connective tissue, having a hierarchical structure as illustrated in Figure 2.1. Collagen molecules, produced by cells, are stacked together in a characteristic \( D \)-period to form collagen fibrils which have diameters of \(~100\) nm. Collagen fibrils are the basic components of collagen fiber which forms connective tissues. The orientation of collagen fibrils varies in different connective tissues to provide particular mechanical and biological functions. In tendon and bone, collagen fibrils align mostly parallel to each other to provide mechanical strength in the axial direction of the tissue. In the cornea, collagen fibrils align radially to form a membrane structure. Alignment also varies across animal species to provide specific biomechanical properties (Boote, Hayes et al. 2008).

![Figure 2.4](image)

Figure 2.4 A periodic cell of a collagen fibril model. Collagen molecules packing in a specific arrangement form a \( D \)-period of \(~67\) nm, which contains an overlap region and a gap region. The \textit{in situ} structure of full length type I collagen fibril is revealed by Orgel \textit{et al.} (Orgel, Irving et al. 2006). The figure shows a full atomistic model of human collagen fibril which is reprinted with permission from Gautieri, A., \textit{et al.}, \textit{Hierarchical structure

Recently, the in situ structure of the full length type I collagen fibril (Protein Data Bank identification code 3HR2) has been revealed (Orgel, Irving et al. 2006) as shown in Figure 2.4. The collagen fibril has a triclinic unit cell with dimensions $a$-40.0 Å, $b$-27.0 Å, $c$-678 Å, $\alpha$=89.2°, $\beta$=94.6°, $\gamma$=105.6°. A fibril has a gap region with a length of 0.54 $D$ and an overlap region with a length of 0.46 $D$ (Figure 2.4). Here $D$ (-67 nm) denotes the length of the $D$-period of collagen fibril. Overlap and gap regions have different biological properties. The cell interaction domain of collagen has been linked to the overlap region while tissue mineralization occurs in the gap region (Sweeney, Orgel et al. 2008). The length of the $D$-period of the collagen fibril has slightly different values for different tissues. The $D$-period is about 67 nm for tendon and bone which contains primarily type I collagen. Skin, which contains about 15% type III and 85% type I collagen, has been shown to have a slightly shorter $D$-period of ~65 nm (Brodsky, Eikenberry et al. 1980; Stinson and Sweeny 1980; Gathercole, Shah et al. 1987).

2.2 Collagen mechanics

2.2.1 Mechanical properties of a single collagen molecule

A typical force-displacement curve of a single collagen molecule is shown in Figure 2.5. For mechanical forces below 14 pN, a collagen molecule is flexible and behaves a worm-like chain behavior (Sun, Luo et al. 2001; Sun, Luo et al. 2004; Buehler and Wong 2007). This is the entropic elasticity regime of a collagen molecule, where the molecule exhibits large strain which may play a role in cell signaling. Collagen molecules have persistence lengths in the range of 10-15 nm, depending on collagen type. Mechanical tests using optical tweezers have shown a persistence length of 14.5 ± 0.73 nm for type I collagen (Sun, Luo et al. 2002) and a persistence length of 11.2 ± 8.4 nm for type II collagen (Sun, Luo et al. 2004). Atomistic simulations have predicted a similar range of the persistence length of collagen (Buehler and Wong 2007). Although the worm-like chain model can describe the force-displacement curve of a collagen molecule quite well overall, the
collagen molecule is known to feature non-uniform deformation throughout its length due to the variations of sequence.

Figure 2.5 Mechanical response of a single collagen molecule. When exposed to a mechanical force below \(-14\) pN, the collagen molecule behaves like a flexible worm-like chain with a persistence length of \(10\) to \(15\) nm. This is the entropic elasticity regime of a collagen molecule. When mechanical force is larger than \(-14\) pN, there are three regimes in a force-displacement curve of a collagen molecule. The collagen molecule is uncurling first followed by stretching backbone covalent bonds and then rupture. (Reprinted from Biophysical Journal, Buehler, M.J. and S.Y. Wong, *Entropic Elasticity Controls Nanomechanics of Single Tropocollagen Molecules.*, 2007. 93(1): p. 37-43., Copyright 2007, with permission from Elsevier (Buehler and Wong 2007))

The mechanical properties of a collagen molecule vary along its twisting axis, and the local conformations are known to change and have different biological functions (Sweeney, Orgel et al. 2008). There exist micro-unfolding regions in a collagen molecule
Micro-unfolding regions are thought to be important for biological functions such as collagen degradation. When a collagen molecule is stretched in the entropic elasticity regime, the mechanical force induces larger deformations in the micro-unfolding regions (as they are softer), while only inducing small deformation for regions that have higher thermal stability. The mechanical response of a collagen molecule in the entropic elasticity regime is likely relevant for its biological functions. For example, recent studies have shown that low mechanical force on the order of pN is sufficient to alter the collagen degradation rate greatly (Adhikari, Chai et al. 2011; Camp, Liles et al. 2011; Adhikari, Glassey et al. 2012).

Once a collagen molecule is pulled out of the entropic elasticity regime, there are three other regimes (Buehler and Wong 2007). Firstly, the collagen molecule undergoes uncurling through its entire length. Earlier experimental and computational studies have shown that the Young’s modulus of a collagen molecule is in the range of 3-9 GPa (Harley, James et al. 1977; Cusack and Miller 1979; Hofmann, Voss et al. 1984; Sasaki and Odajima 1996; Sun, Luo et al. 2002; Gautieri, Vesentini et al. 2011; Chang, Flynn et al. 2012). Further stretching a collagen molecule will induce the stretching of the backbone of each chain resulting in a markedly stiffer response. In the last regime, a collagen molecule is ruptured if it is deformed beyond its strength.

2.2.2 Mechanical properties of collagen fibrils

The mechanical response of collagen fibrils is distinct from collagen molecules. The mechanical features of a collagen microfibril obtained from a molecular model are shown in Figure 2.6. For a hydrated collagen fibril, two regimes have been identified in the stress-strain curve. In the first regime (strain below ~10% and stress below ~50 MPa), the collagen fibril has a non-linear and softer mechanical response. The end-to-end distance of a collagen molecule within the fibril is increased in this region, which suggests that the micro-unfolding domains are stretched. The extensibility of a collagen fibril in this regime is important for cell-matrix interactions and for many biological functions. Once the stress in a collagen fibril reaches ~50 MPa, the end-to-end distance of a collagen molecule reaches its contour length indicating that it has been straightened. Atomistic
simulations have revealed that the straightening of a collagen molecule happens primarily in the gap region (Gautieri, Vesentini et al. 2011). Beyond this point, the stress-strain curve of a collagen fibril enters the second regime which has a linear behavior. Because the collagen molecule has been straightened, the collagen fibril becomes stiffer in this regime (Figure 2.6).

A hydrated collagen fibril has a Young's modulus of \(-300\) MPa at small strain and a modulus of \(-1.2\) GPa at large strain (Gautieri, Vesentini et al. 2011), while a dehydrated fibril has a larger modulus of \(-2\) GPa independent of the applied stress. Figure 2.6 (b) shows values of the Young's modulus of collagen molecules and fibrils from various experimental studies and molecular simulations. A collagen fibril has been found to have a smaller modulus \(-0.4\) to \(0.9\) GPa compared with a single collagen molecule, which features a modulus \(-3\) to \(9\) GPa. These data suggests a strong scale and environment dependence of collagen properties.

Figure 2.6 Mechanical properties of collagen molecules and fibrils. a) Mechanical response of a hydrated and dehydrated collagen fibril. A dehydrated collagen fibril has a linear stress-strain curve while a hydrated collagen fibril has two regimes. In the first regime of a hydrated collagen, the collagen fibril exhibits non-linear relations, while in the second regime, it is stiffer and has a linear stress-strain relation. b) Young's modulus of single collagen molecules and collagen fibrils obtained from various experimental and computational studies (Reprinted with permission from Gautieri, A., et al., Hierarchical structure and nanomechanics of collagen microfibrils from the atomistic scale up. Nano Letters, 2011. 11(2): p. 757-66. Copyright 2011 American Chemical Society.).
2.3 Mutations and diseases in collagenous tissues

2.3.1 Single residue mutations in brittle bone disease

Single residue mutations in collagen molecules have been identified and associated with various diseases. For example, osteogenesis imperfecta (OI), which is known as brittle bone disease, is a rare genetic disorder of collagenous tissues. The OI mutation is caused primarily by a replacement of the Gly residue in the repeating (Gly-X-Y)n triplets. In brittle bone disease, mutations at a single molecule level alter the material properties of collagenous tissue at macro scale. As of now, several mutation locations and types along the entire collagen molecule have been identified and classified into severe, moderate and mild disease conditions as shown in Figure 2.7 (Beck, Chan et al. 2000).

![Figure 2.7 Mutations throughout the entire type I collagen molecule which have been identified to cause severe, moderate and mild OI. (Reprinted from (Beck, Chan et al. 2000), Copyright (2000) National Academy of Sciences, U.S.A.)](image)

Although it remains unclear how a mutation of a collagen molecule could lead to a change in the material property of collagenous tissues, recent studies have revealed that mutations in the molecule cause changes in its structure and mechanical properties (Gautieri, Vesentini et al. 2012) at the molecular level. The OI mutations are found to disrupt the triple helical structure of the collagen molecule in the vicinity of the mutation.
The unfolding of the triple helical structure at the mutation site results from the disruption of inter-chain hydrogen bonds. The severity of structure disruption is found to depend on mutation phenotype (Gautieri, Vesentini et al. 2012). There is a strong correlation with the severity of the phenotype and the inter-chain distance at energy minimum. The OI mutations have been shown to cause reductions in the Young's modulus of collagen molecules (Gautieri, Uzel et al. 2009).

2.3.2 The osteogenesis imperfecta mouse model

The osteogenesis imperfecta mouse model, oim, is also caused by mutations in the collagen molecule. In the case of oim mutations, the alpha-2 chain of the collagen molecule is replaced by an alpha-1 chain, resulting in a homotrimer molecule. Experimental studies of oim mice bone and tendon have shown reduced mechanical strength compared to normal mice. As shown in Figure 2.8 (a), material properties of 1-year-old mice have been measured to examine the severity of phenotype (Camacho, Hou et al. 1999). The oim/oim mice are found to have significant reductions of their failure torque and torsional stiffness compared with normal bone.

At the molecular level it has been proposed that the alpha-2 chain has a critical role in the integrity of the triple helix of collagen (McBride, Choe et al. 1997). X-ray diffraction patterns of oim tail tendon demonstrates a loss of lateral packing in the collagen fibrils (McBride, Choe et al. 1997). Interestingly, the longitudinal structure of the collagen molecule is unaltered in oim tissue and has the normal D-periodic banding (Chipman, Sweet et al. 1993; McBride, Choe et al. 1997). Using differential scanning calorimetry, studies have shown that the denaturation temperature of oim homotrimeric collagen is higher than wild-type heterotrimeric collagen, which implies oim collagen is more thermally stable (Figure 2.8 (b)) (Miles, Sims et al. 2002; Kuznetsova, McBride et al. 2003) and requires more energy to disrupt the integrity of the triple helix. At the same constant temperature, homotrimeric collagen denaturates 100 times slower than heterotrimeric collagen (Kuznetsova, McBride et al. 2003). Using differential scanning calorimetry in both acetic acid and water, Miles et al. calculate that oim collagen has approximately 5% greater volume fracture of water, which they propose increased the
lateral distance between collagen molecules by 1.4 Å in oim fibers (Miles, Sims et al. 2002). The increase in water may be due in part to the lower hydrophobic content of the alpha-1 chain compared to the alpha-2 chain (Kuhn 1982; Kuznetsova, McBride et al. 2001). The increased water content and therefore lateral distance between collagen molecules may be responsible for the loss of lateral packing and reduced cross-linking (Miles, Sims et al. 2002). As cross-linking of molecules are critical for mechanical integrity of collagen (Buehler 2008; Uzel and Buehler 2011), this could account for the weaker mechanical properties of the tissue at larger length scales. These studies provide insight into the role of the alpha-2 chain at the fibril level, however, it remains unclear how the loss of the alpha-2 chain in the homotrimer affects the behavior at the molecular level, which may lead to its altered fibril behaviour and eventually the brittle bone disease at tissue level.

Figure 2.8 Experimental findings on the difference between wild-type and oim collagen. a) Failure torque and torsional stiffness of bone from oim and normal mice (Reprinted from (Camacho, Hou et al. 1999) with permission). b) Thermograms of tail tendon of wild-type and oim mice in water. The oim mice fibers have higher denaturation temperature, indicating that the mutation alters the packing of collagen molecules. (Reprinted from J. Mol. Biol., Miles, C.A., et al., The role of alpha2 chain in the stabilization of the collagen type I heterotrimer: a study of the type I homotrimer in oim
2.4 Collagen degradation

2.4.1 Collagen degradation mechanisms

Collagenases of the matrix metalloproteinase (MMP) family, including MMP-1, MMP-8, MMP-13 and membrane-bound MMP-14 (Nagase, Visse et al. 2006), are major mammalian proteases involved in the physiological cleavage of collagen. MMPs consist of propeptide, catalytic and hemopexin domains. They play an important role in cleaving collagen into characteristic ¼ and ½ fragments. For type I collagen, the specific cleavage site is after the 775th residue (Gly), in the sequences of G-IA for alpha-1 chain and G-LL for alpha-2 chain.

MMPs can only cleave one chain at a time because the binding site of a stable triple helical structure is too narrow. Therefore it is widely accepted that MMPs are not able to cleave a stable triple helical structure. The cleavage site of collagen molecules must be in a vulnerable state to be cleaved. Two possible cleavage mechanisms have been proposed earlier. The first suggests that the collagen does not unwind by itself and MMPs unwind the collagen after binding (Chung, Dinakarpandian et al. 2004). In the second, a collagen molecule is believed to thermally unwind locally at the vicinity of the cleavage site before MMPs bind to it (Stultz 2002).

There is evidence that both mechanisms exist and whether the cleavage site unwinds by itself depends on the thermal stability of the collagen molecule. Atomistic simulations, which serve as a tool that allows us to study the behavior at the vicinity of cleavage site with molecular details, have shown that there exists a vulnerable state of collagen at the cleavage site in the absence of MMPs (Stultz 2006; Nerenberg and Stultz 2008; Stultz and Salsas-Escat 2009; Stultz and Salsas-Escat 2010), suggesting that the cleavage site could be thermally unfolded. On the other hand, recent experimental work has revealed that, for a stable triple helix, the hemopexin domain of MMP-1 binds to the cleavage site first, then a back-rotation of the catalytic domain leads to a "closed" conformation of.
MMP-1 and thus releases one chain out of the triple-helix, suggesting that the enzyme enables the unwinding of the cleavage site of the collagen molecules (Bertini, Fragai et al. 2011).

The MMPs are able to cleave the covalent bond between G-I/L while only one of the several other sites in the collagens that contain the same G-I/L bonds is hydrolyzed (Fields 1991), suggesting that the local conformation at the vicinity of the cleavage site plays an important role in providing a recognition signal for MMPs since amino acid sequence alone is not sufficient for the high specificity of collagen recognition by MMPs (Welgus, Jeffrey et al. 1981; Xiao, Addabbo et al. 2010). Atomistic simulations of all G-I/L sites in type III collagen molecules have provided further evidences that local conformations of all sites have different vulnerability scores (Stultz and Salsas-Escat 2010), indicating that the local conformation provides a recognition signal for enzymes.

The degradation varies in different types of collagen due to varied thermal stability and local conformation of the cleavage site. Experimental studies of human skin fibroblast collagenase have found large differences in the degradation rates, from 1.0 to 565 h⁻¹, for different types of collagen, including collagen type I, type II and type III (Welgus, Jeffrey et al. 1981). Remarkably, the enzyme-substrate affinity is similar for all types of collagen. Han et al. also find similar enzyme-substrate affinity for type I heterotrimer and type I homotrimer which have very different degradation rates (Han, Makareeva et al. 2010). These suggest that the variations of the sequences of collagen molecules do not alter the binding affinity of enzyme but affect the proteolysis rate after enzyme binding.

2.4.2 Mechanical force effects on the collagen degradation rate

Degradation of collagen molecules is a crucial step for many biological and pathological processes such as wound healing, tissue remodeling, cancer invasion and organ morphogenesis (Baragi, Qiu et al. 1997; Cawston, Billington et al. 1999; Matrisian and Brinckerhoff 2002; Nagase and Visse 2003). Precisely regulated collagen degradation is required for normal physiological remodeling and repairing processes. Excessive or deficient degradations have been associated with many diseases. For example, accelerated breakdown of collagen may result in arthritis, atherosclerotic heart disease,
tumor cell invasion, glomerulonephritis, and cell metastasis (Riley, Harrall et al. 1995; Celentano and Frishman 1997; Barnes and Farndale 1999; Bode, Mosorin et al. 1999; McDonnell, Morgan et al. 1999; Sternlicht and Werb 2001; Aigner, Gelse et al. 2003; Nerenberg, Salsas-Escat et al. 2007). Deficient degradation of collagen has been shown to result in increased trabecular bone in mice (Stickens, Behonick et al. 2004).

The chemical composition of a collagen molecule defines its material properties and how it alters its conformation in response to mechanical force. It is clear that mechanical force is able to alter the conformation of collagen molecule and thus mediates the collagen degradation rate. However, it remains a challenging question to understand whether mechanical force speeds up or slows down the degradation rate at different magnitudes of force and in different types of collagen molecules.

We summarize recent studies on mechanical effects on the collagen degradation rate in Table 2-1. Bhole et al. have shown that mechanical strain enhances survivability of collagen micronetworks in the presence of collagenase (Bhole, Flynn et al. 2009). The same mechanism has also been found for the reconstituted collagen fibrils in the presence of MMP-8 (Flynn, Bhole et al. 2010). The fact that mechanical forces are able to slow down the enzymatic cleavage has also been found in different conditions, including native tissue with dynamic loading, uniaxial tension on tendon in vitro, and at low force levels (Nabeshima, Grood et al. 1996; Ruberti and Hallab 2005; Zareian, Church et al. 2010; Camp, Liles et al. 2011). There is also evidence that mechanical force accelerates enzymatic degradation (Ellsmere, Khanna et al. 1999; Adhikari, Chai et al. 2011; Adhikari, Glassey et al. 2012). Experimental studies on single collagen molecule have shown that mechanical force is able to speed up the collagen degradation rate even with a low mechanical force level on the order of pN (Adhikari, Chai et al. 2011; Adhikari, Glassey et al. 2012).

Two mechanisms have been proposed in the literature to explain how mechanical force speeds up and slows down the degradation rate. Mechanical forces can slow down the degradation rate by enhancing the thermal stability of the cleavage site of the collagen molecule (Flynn, Bhole et al. 2010; Camp, Liles et al. 2011). On the other hand, Adhikari
et al. have proposed a molecular mechanism that mechanical force pulls the collagen molecules to another vulnerable state which is more accessible to enzymatic breakdown [69, 70].

These two mechanisms suggest that there exists two vulnerable states of the cleavage sites of collagen molecules. One is the micro-unfolding conformation of the cleavage site, which has a lower thermal stability. The other is the unwinding conformation of the cleavage site of a collagen molecule.

2.4.3 Homotrimer resistance of collagen degradation

Previous studies have shown that the type I heterotrimer and homotrimer have distinct degradation behaviors. Type I homotrimers are found to be resistant to all mammalian collagenases (Narayanan, Meyers et al. 1984; Han, Makareeva et al. 2010; Makareeva, Han et al. 2010), with a cleavage rate much slower for homotrimers than for heterotrimers. The MMP resistance of homotrimers may play an important role in homotrimer-related diseases or in early development, during which necessary collagen degradation may be hindered with detrimental results. For example, it has been shown in fibers reconstituted from mouse tail tendon collagen that a minor fraction of homotrimer-based fibers may grow instead of being disassembled during tissue remodeling cycles, which may eventually result in tissue disorganization (Han, Makareeva et al. 2010).
Table 2-1 Summary of recent experimental studies on mechanical force effects on collagen degradation rate. (Adapted from (Chang and Buehler 2014))

<table>
<thead>
<tr>
<th>Collagen Type</th>
<th>Collagenase</th>
<th>Results</th>
<th>Effect on degradation rate (increase ↑ or decrease ↓)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single collagen trimer peptide (GPQGIAGQRGVVGL)</td>
<td>MMP-1</td>
<td>~10 pN induces a 100-Fold increase in collagen degradation rate (Adhikari, Chai et al. 2011)</td>
<td>↑</td>
</tr>
<tr>
<td>Type I recombinant human collagen molecule</td>
<td>Bacterial collagenase</td>
<td>3-10 pN force slows enzymatic cleavage (Camp, Liles et al. 2011)</td>
<td>↓</td>
</tr>
<tr>
<td>Recombinant, post-translationally modified human collagen I</td>
<td>MMP-1</td>
<td>16 pN causes an 8-fold increase in collagen proteolysis rates (Adhikari, Glassey et al. 2012)</td>
<td>↑</td>
</tr>
<tr>
<td>Recombinant, post-translationally modified human collagen I</td>
<td>Bacterial collagenase</td>
<td>16 pN force does not affect cleavage rates (Adhikari, Glassey et al. 2012)</td>
<td>_</td>
</tr>
<tr>
<td>Reconstituted collagen fibrils</td>
<td>MMP-8</td>
<td>mechanical strain stabilizes enzymatic degradation (Flynn, Bhole et al. 2010)</td>
<td>↓</td>
</tr>
<tr>
<td>Cornea and dissected from mature whole bovine eyes</td>
<td>Bacterial collagenase</td>
<td>collagen degradation corresponds inversely to the tensile stress (Ruberti and Hallab 2005)</td>
<td>↓</td>
</tr>
<tr>
<td>Pepsin-extracted, bovine, type I, atelo-collagen monomers</td>
<td>Bacterial collagenase</td>
<td>mechanical strain enhances survivability of collagen micronetworks (Bhole, Flynn et al. 2009)</td>
<td>↓</td>
</tr>
</tbody>
</table>
3 Methodology

3.1 Atomistic modeling

3.1.1 Atomic description of classical molecular dynamics

Molecular dynamics (MD) is a computational modeling method to simulate the structures and the dynamics of a system with atom by atom precision. In a molecular dynamics simulation, each atom is treated as an individual particle which has its mass $m_i$, atomic position $\mathbf{r}_i$, velocity $\mathbf{v}_i$ and acceleration $\mathbf{a}_i$ as shown in Figure 3.1. The basic idea behind MD is to compute the dynamical trajectory of each atom in the system by considering their atomic interaction potentials and solving each atom’s motion by following Newton’s equations of motion.

![Diagram of molecular dynamics](image)

Figure 3.1 Point representations of a material in molecular dynamics simulations. In MD, each atom is treated as an individual particle which has its mass, atomic position, velocity and acceleration. All the atoms in MD follow Newton’s equations of motion, with suitable initial and boundary conditions, and proper interatomic potentials, while satisfying macroscopic thermodynamical (ensemble-averaged) constraints, to compute the dynamical trajectories of a system (Buehler 2010). Figure reprinted from (Buehler 2010).
3.1.2 Newtonian dynamics

In the Newtonian interpretation of dynamics, the translational motion of an atom \( i \) is caused by a force \( f_i \) exerted by some external agents. The motion and the applied force are explicitly related through Newton's second law:

\[
f_i = m\ddot{r}_i
\]  

(3.1)

in which \( m \) is the mass of the atom. The acceleration \( \ddot{r}_i \) is given by:

\[
\ddot{r}_i = \frac{d^2 r_i}{dt^2}
\]  

(3.2)

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

\[
E = K + U
\]  

(3.3)

where the kinetic energy is

\[
K = \frac{1}{2m} \sum_i p_i^2
\]  

(3.4)

in which the momentum \( p_i \) of an atom \( i \) is defined in terms of its velocity by:

\[
p_i = m\dot{r}_i
\]  

(3.5)

and the potential energy is a function of all the atomic positions:
\[ U = U(r_1, r_2, \ldots, r_N) \]  

(3.6)

For a system with conservative energy, by taking total time derivative of the total energy, we have:

\[ \frac{1}{m} \sum_i p_i \cdot \dot{r}_i + \sum_i \frac{\partial U}{\partial r_i} \cdot \dot{r}_i = 0 \]  

(3.7)

Substituting \( p_i = m \dot{r}_i \) into (3.7), we get:

\[ \sum_i \left( \dot{p}_i + \frac{\partial U}{\partial r_i} \right) \cdot \dot{r}_i = 0 \]  

(3.8)

Since the velocities are all independent of one another, Eq. (3.8) can be satisfied only if, for each atom \( i \):

\[ \frac{\partial U}{\partial r_i} = -\dot{p}_i \]  

(3.9)

or

\[ \frac{\partial U}{\partial r_i} = -f_i \]  

(3.10)

This is the usual expression for a conservative force; that is, any conservative (non-dissipative) force can be written as the negative gradient of a potential function \( U(r_1, r_2, \ldots, r_N) \). Therefore, the basis of a MD simulation is a physical definition of the potential function \( U(r_1, r_2, \ldots, r_N) \) of the system (also called force field), which describes
how atoms interact with each other. With a well-defined potential energy, the time trajectory of each atom in the system can be solved by numerical integrating the equations of motion of each atom.

3.1.3 Force fields for biomaterials

As discussed in the previous section, the fundamental of the MD simulation is a force field to describe the interactions between atoms. A wide range of force fields have been developed for the simulations of biomaterials, such as collagen molecule. AMBER (Pearlman, Case et al. 1995), CHARMM force field (MacKerell, Bashford et al. 1998), OPLS force field (Jorgensen, Maxwell et al. 1996), and GROMACS force field are some of the well-known available force fields for organic materials.

Figure 3.2 Illustrations of the contributions of the different terms in the potential expressions given in Eq. (3.11). Figure reprinted from (Buehler 2010).

In a typical force field for organic materials, such as CHARMM force field, the potential energy of a system (or a molecule) has the form:
\[ U = U_{\text{bond}} + U_{\text{angle}} + U_{\text{torsion}} + U_{\text{coulomb}} + U_{\text{vdW}} + \ldots \]  

(3.11)

in which an assumption that the total energy can be split into the contributions from different types of interactions has been made.

The bond energy of a system is usually written as

\[ U_{\text{bond}} = \sum_{bonds} K_i (b_i - b_0) \]  

(3.12)

where \( b_0 \) denotes the equilibrium bond length, \( b_i \) denotes the bond length between the two atoms, and \( K_i \) denotes the bending stiffness of a given bond. Note that these parameters (\( b_0 \) and \( K_i \)) depend on the type of the bond, i.e. the types of atoms which form the bond.

Similarly, the angle and torsion terms are given by

\[ U_{\text{angle}} = \sum_{\text{angles}} K_\theta (\theta - \theta_0) \]  

(3.13)

and

\[ U_{\text{torsion}} = \sum_{\text{torsions}} K_\chi (1 + \cos(n\chi - \delta)) \]  

(3.14)

where the summations are over all the angles and torsions in the system. The parameters for each type of chemical angle and torsion are well-defined in the force field, such as CHARMM force field. \( U_{\text{coulomb}} \) and \( U_{\text{vdW}} \) are the coulomb energies and the van der Waals interactions between each pair of atoms in the system.
It is worth noting that the molecular simulations is a fundamental approach to model a material, since it considers fundamental interactions between atoms, including bond, angle, torsion, coulomb and vdW interactions. Therefore, MD is a powerful tool to solve structures and dynamics of materials with molecular details. However, it features several limitations mainly due to its extreme cost on the demand of computational resource. Typically, in a MD simulation, the size of the systems is limited to several hundred nanometers and the timescale is limited to hundred nanoseconds (Buehler 2010).

### 3.2 Collagen sequences

The real sequences of type I alpha-1 and type I alpha-2 chains of mus musculus (wild type mouse) are used to generate the collagen molecules. The sequences of the alpha-1 and alpha-2 chains are adapted from NCBI protein database (http://www.ncbi.nlm.nih.gov/protein): AAH50014.1 for alpha-1 chain and NP_031769.2 for alpha-2 chain.

The sequence of the alpha-1 chain is:

```
MFSFVDRL LLLLGLATALLTHGQEDIFEVSCIHNGLRVFVNGETWKPEVCLICICHNTAVCDDVQCNELLD
CPNPQRGEGECAFCPEEYVSPNSEDVGVEPGKDGFPGFQQGRGFVPGRDGIPQQPLGPLPGPFPGPFPGP
PGFLGGMNFASQMSYGYDEKASAVSVFGPRGFGPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
PGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPGFPG
MGFPFPKGTAGEPKGAKERLGPLPGPAVPGPAGDKGDEGAQAPAGFAPAGERGEGPAQPSPFQPQGLGPL
PGPEAGKPGEQVFPGDLAPGSGARGERFGPGEVQCGFPFPAGPQGNGAAPGNGAGKDTQAPAGAPSQ
GAPGLQZMPGZPGZAAGLPKGGDAGPGADGPDKGDAGGLTGPQIPQPAGAPGDKGEAGPSGPPGP
TQRGAIRPSGDRGAEAGPPAGFAPGUGDPDQAGKPGAEPTQTVKGEGAGPAGAPGASGQ
GAPGAAPGAGATCIPGGAARVGPPGPSPGNSAPPPFPGFVGKEGKPGPGETGAPRGEVGPPGPPGAGEK
GSPPGDAGPAPSGPGTPQPGIAQARGJQVGLPGQRGERGFPGLPGPSGEPEPGKSQSGSGERPGPPMPGPL
AGPGESREPAGPGEOSGPRDQAPGAKGERGTPAPEAPGAPAGGAPAGAPEAP PangKNGDRGETGAPAGPAG
PIQAPAGARSPAPQPGPQRDGEGTEQQDRGIEKHKRGFSGLQQPGPSGPQGEQGSGPAGPAAPRPGPASA
GSPPGDQLNLPGPIPGPPGPRGDSQPDAPGGPGFQPGPGFPGFPGFPGFPGFQGDDYDFSFLQPPQQEGSSDGGRYYRA
```

47
The sequence of the alpha-2 chain is:

MLSFVDTRTLLLLAVTSCLATCQLQSGSVRKGPTGDRGPRGRGQPAGPRGRGDPGAPPRGPRAGSPGAPPGLTGNFAQYSDKGVSSGPMPGLMGPRPPAGVGAPFQGQFQGPAEPGE腾讯QTPQAPPRGPAGSPKAGEDGHPKPRGPRGERGVGPGQARFFGPQGGLPGFKGVKHSMDGLKQPGAQQVKEPGAGEN
GTPGQAGARGLPGREGRVPGAPAGARGDSDGSGVPFGAPIGSAAGPPGAPFPAPKLGELPQNGAPGAPAP
APGRGEPVGLPLGSPGGPNGTNGLTGAKATGPLLNGVAGMGAPGPRGIEIPFGAGAGATAGARGVLHPEG
PAGSKGEGSNKGEPSGSGAQGPGGPSEEKRGSPEAGSAGPAGGLRGPSLPGSLAGDGRAGVMGPG
NGRSSTGPAGIRPNGDAGRPRGEPGLMRPRPGPSPGQNVSPKSGKQETPVGLPCIDGRPPIPAGPAPRGEAGN
IGPFPKPKGSDDPGKPRGERHPLAGARAPGDNGNAGAQPPGPQGQGKGEQQPAGPPFPQGQLPGPS
TTGEGRPKPRGELPGEFLPGAPPRGPRGERGPQGESAAGPSGPGIRGPRGPPDPGGEGAGAVAGAAGSA
GASGPGLPGERAGAIGPPGKGKEGGETLQPGTDGNTGRDARGIPAVAGAPGAGASGDRGEAAAGPSGP
AGPRGSPGERGEVPGAGPNFAGPAAGAAGCPAGKEGKTGPRGKENGIVGTGGSVAAGPSPNPGPGVG
SRDGDPGMTGFPAGAAGTPGPGPSGIAPFGPAGAEGIRFPRQDQFVGTGETASAGPSGPFGVEK
GPSGEPGTAGATGAPAGPQLLAPGLGPGLSGRERGLGPIAGALKEGPFLGISGPPARGPPPAGVSGP
NGAPEGAGDRGNPSGGPPGPPGHGPERFGPSICPTGAAGAPPPHSGVPAKHCNRIEGPAPSSGVC
PVGAVGPRSPGQPQRGIRDGKEPDKHRGLFQGSKLYQGLQLGLPQLAGLHGOLDQAGAPGVPGAAPPPAPAPS
GPVGDGRSRQPGPPVGPAVGVRGSQGSQPAPGGPGPPGGPAPVGSGYDPIFRGDFYRADQPSQPSLRP
KDYEVATKLNSNQIETLTLPGERRKSPFARCRTCDRLRSLPFEWSDYIMDPQCTMDIAKVYCDSTGE
TCIAQRPWNTPAKNSYSRQAANKHWWLGETINGGSFQFEYNVNGVSSKEMATQLMRLLANRASQNYTHC
KNSIAYLDEETGSNLAVLQLGSNSDVELVAEGNSRFTYSVLVDCSKKTNEWGTKIIIEYTKNFPSRLFLD
IAPLDIGADDQFVRVEVPVCFK

in which the collagen domain are highlighted as red and bold.

3.3 Steered molecular dynamics

Steered molecular dynamics (SMD) is a useful tool to probe the mechanical property of proteins in molecular dynamics simulations. It is used to probe differences of the mechanical properties between the type I collagen heterotrimer and homotrimer. Here we briefly explain the SMD method.
This method is based on the concept of pulling the center of mass of a collection of chosen atoms via a spring along the pulling direction, while keeping the center of mass of another group of atoms fixed. To achieve this, a harmonic moving restraint to the center of mass of a group of atoms is added into the system. This leads to the addition of the following potential to the total energy of the system:

$$U(r_1, r_2, \ldots, r_n, t) = \frac{1}{2} k \left[ v t - (r(t) - r_0) \cdot n \right]$$

in which $r(t)$ is the position of the restrained atoms at time $t$, $r_0$ is the original coordinates of the restrained atom, $v$ the pulling velocity and $n$ is the pulling direction. The force applied on the restrained atoms can be computed by

$$F(r_1, r_2, \ldots, r_n, t) = k \left[ v t - (r(t) - r_0) \cdot n \right]$$

Figure 3.3 Illustration of the use of steered molecular dynamics to apply mechanical load (a) AFM experiment (b) steered molecular dynamics model (c) an illustrations on the obtained force-displacement curve. Figure reprinted from (Buehler 2010)
By monitoring the applied force (3.16) and the displacement of the restrained atoms, we can thus obtain the force-displacement curve from a full atomistic simulation. It is worth noting that the SMD mimics the AFM nano-mechanic tests as shown in Figure 3.3.

3.4 Persistence length calculation

The persistence length is a measure of the mechanical properties of a molecule, characterizing the stiffness of a polymer. Here we briefly explain the concept of the persistence length. Assume a polymer is made of \( n \) repeating segments and each segment is straight with length \( l \). For a freely rotating chain model, the bonds (segments) are linked by a fixed valence angle and are thought to freely rotate along a cone:

![Diagram of the geometric parameters in the freely rotating chain model.](Chang, Shefelbine et al. 2012)

Let \( r \) denote the end-to-end vector of a polymer, then we have

\[
r = \sum_{i=1}^{n} l_i \tag{3.17}
\]

and

\[
r^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} l_i l_j = n l^2 + 2 \sum_{i<j} l_i l_j \tag{3.18}
\]

The expected value of end-to-end distance is thus
\[ \langle r^2 \rangle = n l^2 + 2 l^2 \sum_{i<j} \langle \cos \theta_{ij} \rangle \]

From the assumption of freely rotating model that angles between two segments are the same, we have

\[ \langle \cos \theta_{ij} \rangle = \langle \cos \gamma \rangle^{j-i} \]

(3.20)

Here \( \gamma \) is the fixed angle between two connected segments. For a sufficiently large number of bonds \(|j - i| \gg 1\), this relation has the following property

\[ \langle \cos \theta_{ij} \rangle = e^{-|j-i|/L_p} \]

(3.21)

Here \( L_p \) is the persistence length. Using this fact and the relation that \( L_c = n l \), integrate Eq. (3.19) for a polymer with infinite number of segments, we obtain the relation that

\[ \langle r^2 \rangle = 2 \langle L_c \rangle L_p - 2 l^2 \left( 1 - e^{-\langle L_c \rangle / L_p} \right) \]

(3.22)

The bending stiffness \( B_s \) of a polymer can be calculated from its persistence length by the relation \( L_p = B_s / k_B T \), where \( k_B \) is Boltzmann's constant and \( T \) is the absolute temperature. The Young's modulus \( E \) of a polymer can also be calculated by \( E = L_p k_B T / l \) noting that \( B_s = E l \), in which \( l \) is the moment of inertia of the cross section of the polymer. We have shown the derivations of the persistence length for a freely rotating model. For an infinitely long polymer satisfying the assumptions made for the freely rotating model (\( n \) repeating segments of constant length \( l \)), two approaches (Eq. (3.21) and Eq. (3.22)) for calculating the persistence length should give the same results. However, violation of any assumption will result in different values of persistence length. It is worth noting that for collagen molecules, (1) the molecule does not have repeating segments, (2) it does not have fixed angles between segments, and (3) the molecule is not
infinitely long. There are also errors resulting from the estimations we use in our calculations, such as (1) the estimations of contour length and end-to-end distance, (2) the error in approximating the tangent vector of collagen molecule by the vector of nearest center of masses, and (3) the approximation of the molecule shapes by a spline fit.

3.5 Structure of collagen molecules

In collagen molecules, each chain is a repeating primary sequence of $(\text{Gly-X-Y})_n$ and the three chains are staggered with respect to each other. The positions of the repeated Gly therefore capture the characteristic structures of collagen molecules. We use the $C_a$ atom of each Gly residue as a measurement point to characterize molecular structure (Figure 3.5, a and b).

In our analysis, $r_i$ denotes the position of $C_a$ atom of $i$-th Gly in the collagen molecule. To estimate the contour length and the end-to-end distance of the collagen molecules, we compute the center of mass of all the nearest three $C_a$ atoms (one from each chain because three chains in triple helix are staggered) by $b_i = \frac{1}{3}(r_i + r_{i+1} + r_{i+2})$. Here $b_i$ is the center of mass, then the contour length is calculated by summing the distances between adjacent $b_i$, i.e.

$$L_c = \sum_{i=1}^{n-1} ||b_{i+1} - b_i||,$$

(3.23)

where $n$ is the number of the center of masses and $L_c$ is the contour length. The end-to-end distance is estimated by the distance between the first and the last $b_i$, i.e. $||b_n - b_1||$.

The distance between the $i$-th $C_a$ atom to its nearest $C_a$ atom can be simply computed by $d_i = ||r_{i+1} - r_i||$.

Furthermore, we use $r_i^A$, $r_i^B$, and $r_i^C$ to be the positions of the $C_a$ atoms of the $i$-th Gly in chain A, B, and C respectively and $b_i$ to be the center of mass of the $C_a$ atoms of the $i$-th Gly in three chains ($b_i = \frac{1}{3}(r_i^A + r_i^B + r_i^C)$). Note that because three chains in triple helix are staggered, we have the relations: $r_i^A = r_{3i-2}$, $r_i^B = r_{3i-1}$ and $r_i^C = r_{3i}$. The unit
height is calculated by $h_i = \|\mathbf{r}_{i+1} - \mathbf{r}_i\|$ (Figure 3.5 d). We calculate the radius as the distance between each $C_\alpha$ to the twisting axis of the collagen molecule, which is approximated by constructing a 3d natural interpolating cubic spline curve to $\mathbf{r}_i$ (Figure 3.5 c). Once the spline curve is obtained, the $i$-th radius, i.e. the radius at $\mathbf{r}_i$, is defined by averaging the radius at $\mathbf{r}_i^A$, $\mathbf{r}_i^B$, and $\mathbf{r}_i^C$. The radii at $\mathbf{r}_i^A$, $\mathbf{r}_i^B$, and $\mathbf{r}_i^C$ are obtained by compute the distances between them to the twisting axis.

![Image](image.png)

Figure 3.5 Structure of collagen protein materials. (a) The cross-section of the collagen molecule. Each chain is plotted by ribbon in the same color. The spheres are the $C_\alpha$ atoms of Gly residues. The arrow indicates the radius which is defined by the distance from the $C_\alpha$ atom to the twisting axis. (b) The side view of the collagen molecule. (c) The spline fit. The blue (red, and cyan respectively) line is the spline fit passing through the $C_\alpha$ atoms of chain A (B and C respectively). x represents the positions of $C_\alpha$ atoms. The black line is the approximated twisting axis of collagen molecule obtained by the spline fit. (d) Illustration on the unit height ($h_2$), separated distance ($l_{13}$) between two points on collagen molecule and the angle ($\theta_{13}$) between the tangent vectors of these two points. (Chang, Shefelbine et al. 2012)
The structural analysis approach described here is applied in the following chapters to analyze and compare the structural differences between the normal and *oim* collagen molecules.
4 Structural and mechanical differences between wild type and *oim* collagen molecules

(This chapter is adapted from (Chang, Shefelbine et al. 2012))

4.1 Introduction

Normal type I collagen is a heterotrimer triple helical molecule consisting of two alpha-1 chains and one alpha-2 chain. A mouse model of the genetic brittle bone disease, osteogenesis imperfecta, *oim*, is characterized by a replacement of the alpha-2 chain by an alpha-1 chain, resulting also in a homotrimer collagen molecule. Experimental studies of *oim* mice tendon and bone have shown reduced mechanical strength compared to normal mice. The relationship between the molecular content and the decrease in strength is, however, still unknown.

In this chapter, full atomistic simulations of a section of mouse type I heterotrimer and homotrimer collagen molecules are developed to explore the effect of the substitution of the alpha-2 chain at the molecular level. We calculate the persistence length and carry out a detailed analysis of the structure to determine differences in structural behavior between hetero- and homotrimers.

4.2 Model construction

4.2.1 Collagen molecule generation

The real sequences of type I alpha-1 and type I alpha-2 chains of *mus musculus* (wild type mouse) are used to generate the collagen molecules. The heterotrimer collagen molecule is built of two alpha-1 chains and one alpha-2 chain while the homotrimer collagen molecule is built of three alpha-1 chains. The sequences are adapted from NCBI
protein database (http://www.ncbi.nlm.nih.gov/protein): AAH50014.1 for alpha-1 chain and NP_031769.2 for alpha-2 chain. The entire alpha-1 and alpha-2 chains consist of 1014 residues with repeated G-X-Y triplets, excluding the C-terminal and N-terminal sequences. A specific section with a length of 57 residues of the 1014 total residues (from the 403rd to 459th residues) is selected to generate the heterotrimer and homotrimer collagen molecules. The chosen sequences are

alpha-1:
GFPGPKGTAGEPGKAGERGLPGPPGAVGPAGKDGEAGAAGAPGAPPGAPAGPAGERGE

and

alpha-2:
GFPGPKGPSGDGKPGERGHGPLAGARGAPGDGNNGAQGPPGPQGVQQGKGE

Note that these sequences of the alpha-1 chain and alpha-2 chain are chosen particularly such that the six residues of the two ends are the same for both alpha-1 and alpha-2 chains in order to avoid the possible boundary effects of the heterotrimer and homotrimer collagen molecules. We have verified that the amino acid composition (by %) in the segment is similar to the composition of the complete collagen molecule.

The collagen molecules are created by inputting the sequences of three chains into the software THeBuScr (an interactive triple-helical collagen building script) (Rainey and Goh 2004), which enables a user to build a triple-helical molecule based on any specified amino acid sequence. The code uses derived conformations from statistical analyses of high-resolution x-ray crystal structures of triple helical peptides to build collagen molecule structures. The snapshots of initial structures of the heterotrimer and homotrimer are shown in Figure 4.1. To neutralize the terminals, two ends of the heterotrimer and homotrimer are capped by assigning the first residue to ACE, the acetylated N-terminus, and the last residue to CT3, the N-Methylamide C-terminus. The length of the collagen molecule is about 160 Å and is solvated in a periodic water box composed of TIP3 water molecules. The VMD software (Humphrey, Dalke et al. 1996) is used to solvate the collagen and also to neutralize the system by adding ions. The final
solvated all-atom system contains ~37,000 atoms, with a box size of 280 Å × 40 Å × 40 Å.

Figure 4.1 Initial structures of the heterotrimer and homotrimer collagen molecules. (Chang, Shefelbine et al. 2012)

4.2.2 All-atom equilibration and mechanical test

Full atomistic simulations are performed using NAMD (Nelson, Humphrey et al. 1996) and the CHARMM force field (MacKerell, Bashford et al. 1998) that includes parameters for hydroxyproline amino acids based on a suggestion by Anderson (Anderson 2005). This force field has been widely validated for a variety of biochemical models of proteins including collagen (Gautieri, Vesentini et al. 2008; Gautieri, Vesentini et al. 2008; Gautieri, Buehler et al. 2009; Gautieri, Uzel et al. 2009; Srinivasan, Uzel et al. 2009). An energy minimization using a conjugate gradient scheme is performed before molecular dynamics simulations. Rigid bonds are used to constrain covalent bond lengths, thus allowing an integration time step of 2 fs. Nonbonding interactions are computed using a cut-off for neighbor list at 13.5 Å, with a switching function between 10 and 12 Å for van der Waals interactions. The electrostatic interactions are modeled by the particle mesh Ewald summation (PME) method.

After energy minimization, the collagen molecule is simulated through 50 ns at a constant temperature of 310 K and 1.013 bar pressure in molecular dynamics simulations. The first 10 ns simulations are used to equilibrate the system and are excluded from analyses. The last 40 ns simulations are used for analyses. During the last 40 ns simulations, the
configurations of collagen molecule are recorded every 20 ps which results in 2,000 frames. Each frame of the 2000 frames is used in the analyses shown in this paper. The total simulation takes about 10,000 CPU hours. We also perform repeated simulations for both the heterotrimer and homotrimer, from a different starting point, to confirm the reliability of our results.

Steered molecular dynamics (SMD) is used to probe the mechanical properties of the heterotrimer and the homotrimer. This method is based on the concept of pulling the center of mass of a collection of chosen atoms via a spring along the pulling direction, while keeping the center of mass of another group of atoms fixed. More details on the SMD is provided in Chapter 3.3. After 40 ns NPT simulations, we fix the left ends of the collagen molecules and apply constant velocity pulling method to stretch both the heterotrimer and the homotrimer. In both simulations, the spring constants are 7 kcal/mol/Å² and the pulling velocities are set to 5 Å/ns.

4.2.3 Analysis of the structures of collagen molecules

Five structural parameters of collagen molecule are analyzed for the heterotrimer and homotrimer collagen molecules: contour length, end-to-end distance, the distance between nearest Cα atoms, unit height, and radius (the distance from Cα atom to twisting axis). The details on how we compute the structural parameters are provided in Chapter 3.5.

We also calculate and compare the persistence lengths of the heterotrimer and the homotrimer. The persistence length characterizes the stiffness of a long protein. Using θ to denote the angle between the tangent of two points of the protein separated by a distance L and $L_p$ to denote the persistence length of the protein, the expectation value of the cosine angle exponentially decays with the distance (Flory 1969), that is,

$$\langle \cos \theta \rangle = e^{-L/L_p}. \quad (4.1)$$
Alternatively, using \( r \) to denote the end-to-end distance of a long protein, the following relation can also be derived under the assumption that the protein is infinite long (Schäfer and Elsner 2004):

\[
\langle r^2 \rangle = 2\langle L_e \rangle L_p - 2L_p^2 \left(1 - e^{-\langle L_e \rangle / L_p}\right).
\] (4.2)

Both Eq. (4.1) and Eq. (4.2) are used to calculate the persistence length in this paper. A brief derivation of Eq. (4.1) and Eq. (4.2) is provided in Chapter 3.4. To use Eq. (4.1), we use relation between \( \theta \) and \( L \). For each frame, we calculate \( l_{ij}(t) \), the distance between \( \vec{r}_j \) and \( \vec{r}_i \), and \( \theta_{ij}(t) \), the angle between the tangent of these two positions at time \( t \) (Figure 3.5d), thus we have

\[
l_{ij}(t) = \|\vec{r}_j(t) - \vec{r}_i(t)\|, \quad \text{and} \quad (4.3)
\]

\[
cos \theta_{ij}(t) = (\vec{r}_{j+1}(t) - \vec{r}_j(t)) \cdot (\vec{r}_{i+1}(t) - \vec{r}_i(t)). \quad (4.4)
\]

Note that here we approximate the tangent at \( \vec{r}_i(t) \) by \( \vec{r}_{i+1}(t) - \vec{r}_i(t) \). Then, the expectation value of the distance \( l_m \) and the cosine angle \( \cos \theta_m \) between \( m \)-th neighboring center of masses are obtained by averaging through simulation time:

\[
\langle l_m \rangle = \langle l_{ij} | j - i = m \rangle, \quad \text{and} \quad (4.5)
\]

\[
\langle \cos \theta_m \rangle = \langle \cos \theta_{ij} | j - i = m \rangle. \quad (4.6)
\]

The expectation value of the distance and the cosine angle are calculated by averaging the last 40 ns simulations. Then, an exponential fitting, i.e. Eq. (4.1), is applied to obtain the persistence length of collagen molecules. Note that the exponential decay in Eq. (4.1) holds only for long proteins and the length of the protein in the simulation is only 160 Å. Therefore, we only use the expected values up to fifth neighboring center of masses.
An alternative way of calculating the persistence length is by using Eq. (4.2). The expectation value of the contour length and the end-to-end distance of collagen molecule are calculated by averaging the last 40 ns of the simulations. Then the persistence length is obtained directly by solving Eq. (4.2).

4.2.4 Analysis of the freely rotating behavior

We use \( \vec{f}_i, \vec{f}_j \) and \( \vec{f}_k \) to denote three center of masses on the collagen molecule. The rotating angle \( \theta \) is the supplementary angle formed by these three center of masses (Figure 4.7 a). The expectation value of the rotating angle is obtained by averaging the rotating angles for last 40 ns simulations. In order to further reveal the freely rotating behavior, for each configuration during the last 40 ns simulations, a rotation is performed such that \( \vec{f}'_j - \vec{f}_i \) is parallel to the x-axis. Here \( \vec{f}'_j \) is the new coordinate of \( \vec{f}_j \) after rotation. Finally, we use \( \vec{n}'_{jk} \) to denote the unit vector of \( \vec{f}_k' - \vec{f}_j' \). The projection of \( \vec{n}'_{jk} \) on the y-z plane is calculated for each recorded frame.

We analyze the freely rotating behavior at the center of masses of C\(_\alpha\) atoms of the 2\(^{\text{nd}}\), 3\(^{\text{rd}}\), 4\(^{\text{th}}\), 5\(^{\text{th}}\), 6\(^{\text{th}}\), 9\(^{\text{th}}\), 13\(^{\text{th}}\), and 17\(^{\text{th}}\) Gly residues. We chose these regions to demonstrate differences between local regions based on the structural parameters calculated above.

4.3 Results

4.3.1 Persistence length calculations

We calculate the contour lengths and the end-to-end distances of heterotrimer and homotrimer collagen molecules of mus musculus by averaging the last 40 ns of the total 50 ns full atomistic simulations. The contour length of heterotrimer collagen is 166.6 ± 1.7 Å and homotrimer collagen is 162.5 ± 1 Å, a slightly smaller value. Interestingly, there is a striking difference between the end-to-end distance of the heterotrimer (147.6 ± 6.8 Å) and homotrimer (127.1 ± 5.5 Å) collagen molecules. The persistence length, which characterizes the mechanical stiffness, can be obtained through the relation between the contour length and the end-to-end distance of the collagen molecule. The calculated persistence length is 215 Å for the heterotrimer collagen molecule and 96 Å.
for homotrimer collagen molecule, suggesting that the homotrimer collagen molecule is more flexible (smaller overall bending stiffness).

Figure 4.2 The relation between the separated distance and the loss of the direction of the tangent. The relation between averaged distance $\langle l_m \rangle$ and cosine angles $\langle \cos \theta_m \rangle$ of heterotrimer and homotrimer collagen molecules. The line represents the exponential fit (Eq. (4.1)) with persistence length of 209 Å for heterotrimer collagen molecule and 89 Å for homotrimer collagen molecule. (Chang, Shefelbine et al. 2012)

Our calculation results of the persistent lengths are within the range of previously reported values. By fitting the force-extension curve from mechanical test using optical tweezers to a worm-like chain elasticity model, the persistence length of type I heterotrimer collagen obtained by Sun et al. (Sun, Luo et al. 2002) is found to be $145 \pm 73$ Å and the persistence length of type II homotrimer collagen is found to be $112 \pm 84$ Å (Sun, Luo et al. 2004). These studies also suggest that the homotrimer has a smaller persistence length than the heterotrimer, but it should be noted that the homotrimer in the study cited above is type II collagen. The persistence lengths obtained from earlier simulations are in the range from approximately 150-250 Å (Buehler and Wong 2007), and coarse-grained modeling of collagen molecules using a MARTINI force field...
suggests a value of 515 ± 67 Å (likely higher due to the generally stiffer nature of protein structures in the MARTINI approach) (Gautieri, Russo et al. 2010).

Figure 4.3 Statistics of structural parameters of collagen protein. (a) Averaged unit height of heterotrimer collagen molecule. (b) Averaged unit height of homotrimer collagen molecule. (c) Averaged nearest Cα distance of heterotrimer and homotrimer collagen molecules. (d) Averaged radius of heterotrimer and homotrimer collagen molecules. (Chang, Shefelbine et al. 2012)
As an alternative approach we also calculate the persistence length as a measure of the direction of the tangent loss of a protein over a distance. The relation between the separated distance and the cosine of the angle between two tangent vectors at two points on the heterotrimer and homotrimer collagen molecules for simulations from 10 ns to 50 ns are shown in Figure 4.2. The decrease of the cosine angles of homotrimer collagen molecule is larger than the heterotrimer collagen molecule as the distances increases. That is, the homotrimer collagen molecule has a shorter persistence length and is more flexible to change its direction. Exponential fits (see Eq. (4.1)) are fit to the results of heterotrimer and homotrimer collagen molecules shown in Figure 4.2. The persistence lengths from the exponential fitting are 209 Å for heterotrimer collagen molecule and 89 Å for homotrimer collagen molecule. The results are consistent to the values obtained by the contour length and end-to-end distance approach used above and again show that the homotrimer collagen molecule has a shorter persistence length and is more flexible.

4.3.2 Structural differences between the heterotrimer and homotrimer

The persistence length calculations assume the collagen molecules to be homogeneous, which is an over-simplification of a heterogeneous molecular structure. We thus perform additional analyses to identify structural mechanisms that governed the apparent change in molecular stiffness.

Figure 4.3, a and b show the averaged values of unit heights of heterotrimer and homotrimer collagen molecules, respectively. The results show that both the heterotrimer and homotrimer collagen molecules are not homogeneous along the length of the collagen section analyzed. To validate our results with experimental data, the average of all the unit heights (excluding two at each end) are calculated. The averaged unit height is 8.57 Å for the heterotrimer and 8.59 Å for the homotrimer collagen molecule. The results are similar to that obtained from a statistical analysis of high-resolution X-ray crystal structures of triple-helical peptides, which shows a varied unit height of 8.53 Å for iminorich regions and 8.65 Å for aminorich regions (Rainey and Goh 2002).
Figure 4.4 Unit height distributions of heterotrimer and homotrimer at 0, 20, 30, 40 and 50 ns. The results show that the unit heights at both ends are very stable during the simulation. The unit height distributions converge during the simulation. The micro-unfolding regions of heterotrimer and homotrimer have decreased unit heights. One more region with decreased unit height is found at the 5th unit height of homotrimer and the value of decreased unit height is very stable. (Chang, Shefelbine et al. 2012)

The results show that both the heterotrimer and homotrimer collagen molecules have abnormal shorter unit heights at the region around the 12th triplet (corresponding to the 436th residues in terms of the entire type I collagen molecule). The standard deviation of unit heights at this region of heterotrimer collagen molecule is larger than other regions of the molecule, which indicates that thermal energy easily varies the conformations of the heterotrimer. This can also be confirmed by unit height distributions through simulation time (Figure 4.4), and time history of unit heights (Figure 4.5). This finding is in accordance with experimental studies that have shown lower denaturation temperature of heterotrimers (Miles, Sims et al. 2002; Kuznetsova, McBride et al. 2003). There is an additional region with a shorter unit height for homotrimer collagen molecule around the 5th triplet (corresponding to the 418th residues of type I collagen molecules).
Figure 4.5 Time history of unit heights of heterotrimer and homotrimer for different regions of the segment analyzed. For the heterotrimer, decreased unit heights are found only at the micro-unfolding region ($h_{11}$ and $h_{12}$). For the homotrimer, unit heights
decreased at two kink regions \((h_5, h_6, h_{11} \text{ and } h_{12})\) and are very stable during simulation. (Chang, Shefelbine et al. 2012)

Figure 4.6 Snapshots of collagen protein. Each chain is plotted in ribbon with the same color. (a) 20 ns, heterotrimer collagen molecule. (b) 40 ns, heterotrimer collagen molecule. (c) 20 ns, homotrimer collagen molecule. (d) 40 ns, homotrimer collagen molecule. (Chang, Shefelbine et al. 2012)

The distances between the nearest \(\text{Ca}\) atoms of the Gly residues are shown in Figure 4.3. The results show that at the region around the 34\(^{th}\) to 39\(^{th}\) carbon atoms in Figure 4.3, the distance between the nearest \(\text{Ca}\) atoms of both heterotrimer and homotrimer collagen are larger than other regions. Note that each triplet has three Gly residues (one from each chain), and therefore the 12\(^{th}\) triplet corresponds to the 34\(^{th}\) to 36\(^{th}\) Gly residues. Similarly, the radii (Figure 4.3) at this region are larger than other regions; that is, the distances from the \(\text{Ca}\) atoms of Gly to the twisting axis are larger. Therefore, we refer to this region as the “micro-unfolding” region (Figure 4.6). This finding is consistent with previous simulation results that show micro-unfolding at 436\(^{th}\) Gly residue of type I
collagen (as can be confirmed in Table 1 in Ref. (Bodian, Radmer et al. 2011)). In a region of micro-unfolding, the H-bond between Gly on adjacent chains might not form easily since the distances between Gly are large. There are only two nearest Cα distances (d_{30} and d_{33}) abnormally larger in homotrimer collagen molecules, while many abnormally larger nearest Cα distances are found in heterotrimer collagen molecule. This also supports our finding that the heterotrimer collagen molecule is more thermally unstable and the micro-unfolding is more severe. In contrast to the micro-unfolding region, at the region near the 5th triplet of the homotrimer, where unit height decreases, there is no increase in the radius nor the distances between nearest Cα atoms. These decreases of the unit height do not result in thermal instability in this region. It is also shown directly in the snapshot in Figure 4.6 that this region does not feature micro-unfolding.

4.3.3 Analysis of the kinks of the heterotrimer and homotrimer

We investigate how regions with a shorter unit height may contribute to a decrease in persistence length. Table 4-1 shows the rotating angles at different regions for both heterotrimer and homotrimer collagen molecules. In the 3rd and 4th triplets (2,3,4 and 3,4,5), the angles of rotation for both heterotrimer and homotrimer collagen molecules are small. In regions where the unit height is short (12th residue in heterotrimer and 5th and 12th in the homotrimer), there is an increase in the angle of rotation. When a large rotation (larger than 15°) is present, we call this 'kink' indicating a local bend in the molecule (Figure 4.6, c and d). Both heterotrimer and homotrimer have kinks at the micro-unfolding region but the kink angle of the homotrimer is about three times larger than it of the heterotrimer (22.4° vs. 63.9°). Moreover, one more kink region (the 5th triplet) is found in homotrimer. Therefore, we conclude that the loss of alpha-2 chain leads to the ability to kink with larger angles at specific regions.
Figure 4.7 The comparison of the kink and freely rotating behaviors between heterotrimer and homotrimer collagen molecules. (a) Illustration of the analysis method. (b) Illustration of several center of masses of collagen molecules. (c) The projection of the rotating vectors on the $y$-$z$ plane for $(i,j,k) = (3,6,9)$. The number indicates the simulation time in ns unit. The results show that at the 5th triplet region, a specific kink is found in homotrimer collagen molecule and the homotrimer rotates first counterclockwise (from 10 ns to 28 ns), then clockwise (28 ns to 48 ns) at this region. (Chang, Shefelbine et al. 2012)

The projections of the rotating vectors on the cross-section of collagen molecule at the 5th residue are shown in Figure 4.7 c. The inner circle represents the line with a rotating angle of 20° and the outer circle represents the line with a rotating angle of 50°. It is seen that the projections of the rotating vectors of heterotrimer collagen molecule lie within the inner circle and are closed to the origin, with an average of $10.1 \pm 5.3°$ (Table 4-1).
Projections of the rotating vectors of the homotrimer collagen molecule are much larger with an average value of 37.3 ± 7.6°. The numbers shown in the figure indicate the simulation time in ns unit, and indicate that the homotrimer rotates first counterclockwise, then clockwise at this location. From the results, we observe a very specific rotating angle for the homotrimer at this region, which is reflected in the overall shorter persistence length. The specific rotating behavior identified here provides the molecular origin of the result that the homotrimer is more flexible and has an overall smaller bending stiffness.

Table 4-1 Comparison of kink angles at different regions of heterotrimer and homotrimer collagen molecules. (Chang, Shefelbine et al. 2012)

<table>
<thead>
<tr>
<th>(i,j,k)</th>
<th>Heterotrimer</th>
<th>Homotrimer</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2,3,4)</td>
<td>7.6° ± 4.5°</td>
<td>7.7° ± 4.2°</td>
</tr>
<tr>
<td>(3,4,5)</td>
<td>6.3° ± 3.3°</td>
<td>7.2° ± 3.7°</td>
</tr>
<tr>
<td>(3,6,9)</td>
<td>10.1° ± 5.3°</td>
<td>37.3° ± 7.6°</td>
</tr>
<tr>
<td>(9,13,17)</td>
<td>22.4° ± 9.6°</td>
<td>63.9° ± 10.3°</td>
</tr>
</tbody>
</table>

4.3.4 Mechanical tests of the heterotrimer and the homotrimer

We use steered molecular dynamics to probe the mechanical responses of the heterotrimer and the homotrimer. Figure 4.8 a shows the illustration of the mechanical simulations and the structures (after 40 ns NPT equilibrium) of the heterotrimer and the homotrimer before applying force. The end-to-end distances of the heterotrimer and the homotrimer are 149 Å and 128 Å respectively before applying force.

The stress-strain curves of the heterotrimer and the homotrimer are shown in Figure 4.8 b. We assume that the heterotrimer and the homotrimer have the same diameters equal to 15 Å in the calculations of the stress. The stress-strain curves of both collagen molecules can be divided into two regimes: entropic elasticity and the energetic elasticity. For the first regime where the stress level is small (< 100 MPa), the collagen molecule is in the entropic elasticity region where the applying force decreases the entropic energy of the
collage molecule and the collagen molecule is flexible and straightening out. The stress required is fairly constant as the molecule straightens. In the second regime, the collagen molecule is in the energetic elasticity region where the collagen becomes stiffer and larger stresses (~ 1 GPa) are required to stretch the collagen molecule. In this regime, stress increases as strain increases.

Figure 4.8 Mechanical tests of the heterotrimer and the homotrimer. (a) Illustration of the simulation condition and the equilibrium structures of the heterotrimer and the homotrimer before applying forces. (b) Stress-strain curves of the heterotrimer and the homotrimer. (c) The structures of the heterotrimer and the homotrimer when stretched to 160, 170 and 180 Å. (Chang, Shefelbine et al. 2012)

It is worth noting that the homotrimer has a larger entropic elasticity region than the heterotrimer. The heterotrimer transitions to the energetic elasticity region when the strain is larger than 0.15 while the homotrimer stays in the entropic elasticity region until the strain exceeds 0.3. The results show that the homotrimer is more flexible than the heterotrimer and has an overall smaller stiffness. The specific kinks of the homotrimer
result in the shorter end-to-end distance and thus the homotrimer has larger extension before stretching into the energetic elasticity region. The mechanical tests confirm the predictions given by the calculations of the persistence lengths that the homotrimer is more flexible.

![Diagram](image)

Figure 4.9 Direct comparisons of the unit heights of heterotrimer and homotrimer between the repeated and the original simulations. (a), Heterotrimer. (b), Homotrimer. These results show that the unit heights of heterotrimer and homotrimer in the repeated simulations are in accordance with the results of the original simulations. (Chang, Shefelbine et al. 2012)

Interestingly, we find that the stresses required to stretch the heterotrimer and the homotrimer to the same end-to-end distances are the same. Figure 4.8 c shows the stretched structures of the heterotrimer and the homotrimer when the end-to-end distances are stretched to 160, 170 and 180 Å. The corresponding stress-strain relations of every structure are indicated by blue and red numbers for the heterotrimer and the homotrimer respectively in Figure 4.8 b. It indicates that the differences in the stress-strain behaviors are the consequences of the different equilibrium structures of the heterotrimer and the homotrimer. The homotrimer has kinks which result to shorter end-to-end distance at equilibrium and thus has softer stress-strain responses compared to the heterotrimer when stretched. Therefore, the differences in the equilibrium structures of
the heterotrimer and the homotrimer are crucial and dominate the different mechanical responses between them.

4.3.5 Repeated simulations

We repeat our simulations for both the heterotrimer and homotrimer in order to confirm the reliability of our results from different starting configurations. To do this we assign new sets of random velocities to the structures of heterotrimer and homotrimer after 10 ns \textit{NPT} simulation and then 20 ns \textit{NPT} simulations are performed for both collagen molecules. The results of the repeated simulations are compared to the results from 10 to 30 ns of the original simulations and we find that the results are in agreement with the original simulations.

Figure 4.10 Direct comparisons of the distances between Ca atoms of heterotrimer and homotrimer between the repeated and the original simulations. (a), Heterotrimer. (b), Homotrimer. These results coincide with the original simulations that at the region around the 34\textsuperscript{th} to 39\textsuperscript{th} carbon atoms, the distances between the nearest Ca atoms of both heterotrimer and homotrimer collagen are larger than other regions. That is, the same micro-unfolding regions as reported in the manuscript are confirmed by the repeated simulations. (Chang, Shefelbine et al. 2012)
The direct comparisons of all structural parameters of heterotrimer and homotrimer between the original and repeated simulations are shown in Figure 4.9 to Figure 4.11 and the analysis of the freely rotating behavior is shown in Figure 4.12. The results confirm that there is a specific kink at the 5th triplet region of the homotrimer but no kink is found at this region of the heterotrimer. The kink angles at this region calculated from the repeated simulations are $8.1^\circ \pm 4.3^\circ$ and $38.9^\circ \pm 7.7^\circ$ for heterotrimer and homotrimer, which are in agreement with the values shown in Table 4-1 ($10.1^\circ \pm 5.3^\circ$ for the heterotrimer and $37.3^\circ \pm 7.6^\circ$ for the homotrimer). It is seen from the data that the homotrimer rotates first counterclockwise, then clockwise at this location. It is in accordance with our observation in the set of first simulation that the homotrimer can freely rotate at this specific region. Therefore, the results of the repeated simulations confirm our findings and suggest that the results are generally valid.

Figure 4.11 Direct comparisons of the radii of heterotrimer and homotrimer between the repeated and the original simulations. (a), Heterotrimer. (b), Homotrimer. These results show that larger radii are found at the micro-unfolding regions and the radii calculated from the repeated simulations coincide with the radii calculated from the original simulations. (Chang, Shefelbine et al. 2012)
Figure 4.12 The projection of the rotating vectors on the $y$-$z$ plane for $(3,6,9)$ of the repeated and the original simulations. The number indicates the simulation time of the repeated simulation of the homotrimer in ns unit. The kink angles at this region calculated from the repeated simulations are $8.1^\circ \pm 4.3^\circ$ and $38.9^\circ \pm 7.7^\circ$ for heterotrimer and homotrimer, which are in agreement with the values obtained from the original simulations ($10.1^\circ \pm 5.3^\circ$ for heterotrimer and $37.3^\circ \pm 7.6^\circ$ for homotrimer). These results confirm the finding reported in the manuscript that the replacement of the alpha-2 chain results the ability of kinking and freely rotating behavior of homotrimer. The kinking and freely rotating behavior of homotrimer collagen is likely to affect the packing of the molecules into fibrils and provides possible insights and relevance for the molecular origin of brittle bone disease. (Chang, Shefelbine et al. 2012)

4.4 Summary

In this chapter, we consider a segment of collagen that is 57 amino acids ($\approx 160 \text{ Å}$) long, which is only a small portion of the 1014 amino acids ($\approx 3,000 \text{ Å}$) of the entire molecule, but computationally feasible to model at the atomic level. Despite modeling only a small portion (which can be limiting), we are able to clearly characterize differences in mechanical flexibility and local structural behavior. There is a possibility that throughout the molecule there are similar local regions of micro-unfolding in both the heterotrimer...
and homotrimer cases. We anticipate that kinking also occurs at other regions, predominately in the homotrimer.

![Diagram of collagen microfibril with heterotrimer and homotrimer structures showing kinking and freely rotating behavior.]

Figure 4.13 The role of the kink of collagen molecule in the microfibril. The kinking and freely rotating behavior of homotrimer collagen is likely to affect the packing of the molecules into fibrils. The lateral space used by the homotrimer collagen molecule is larger since it has larger kink angles. The larger kink angles are likely to lead to a larger distance between homotrimer molecules in the microfibril. (Chang, Shefelbine et al. 2012)

The behavior of the collagen molecule can affect the mechanical properties of many tissues including bone, tendon, and ligament. In these tissues, collagen molecules are aligned in a quarter staggered arrangements to form fibrils (Figure 4.13). The kinking and freely rotating behavior of homotrimer collagen is likely to affect the packing of the molecules into fibrils. The lateral space used by the homotrimer collagen molecule is larger since it has larger kink angles. Therefore, the larger kink angles are likely to lead to a larger distance between homotrimer molecules in the fibril. The larger distance between collagen molecules could lead to the reduction of cross-linking between collagen
molecules and the loss of mechanical integrity of the tissue. Our simulation results provide evidence of the increase in possible lateral configurations for the homotrimer collagen molecule. As shown directly in Figure 4.6 and Figure 4.13, the explored lateral distance is increased in the kink regions. Indeed, from computations of the volume fraction of the water, Miles et al. (Miles, Sims et al. 2002) suggested an increase in the lateral packing of 1.4 Å between collagen molecules in homotrimer collagen compared to the heterotrimer collagen, which is consistent with our findings. Our studies on the structural and mechanical behavior of a single collagen molecule provide possible insights and relevance for the loss of mechanical integrity of the molecule at the atomistic level.

In summary, we carried out fully atomistic simulations of mus type I heterotrimer and homotrimer in explicit water to explore the role of the loss of alpha-2 chain in collagen molecules. The simulations range over 50 ns to systematically explore the different structural and mechanical behaviors between heterotrimer and homotrimer collagen molecules. We found that the persistence length of the heterotrimer is twice that of the homotrimer, indicating the homotrimer collagen molecule is significantly more flexible and hence softer. Specifically, structural analyses of collagen molecules show that the greater flexibility of homotrimer collagen molecule, smaller bending stiffness and smaller persistence length, results from the emergence of kinking at two specific regions. The observation of these freely rotating behaviors at specific regions confirms the experimental and computational findings that the collagen molecules are not homogeneous along their length and the behavior is dominated by the local variation of amino acid sequences. Local kinking is likely to disrupt the lateral distance between collagen molecules in a fibril. Our results provide a possible mechanism that accounts for the increase in the lateral inter-molecular distance that has been suggested by experiments. We believe that local kinking would not affect the axial spacing of the collagen molecules, which is why the D-period remains relatively unchanged in homotrimeric collagen.

Micro-unfolding regions are known to be important for protein digestion; in triple helical conformation, the protein is protected, but unfolded regions are susceptible to cleavage
by proteolytic enzymes. Our atomistic model suggests that we can predict such cleavage sites from fundamental structural analyses of the molecule. Experimental studies have revealed that homotrimeric \textit{oim} collagen is more resistant to digestion than heterotrimers (Chipman, Sweet et al. 1993), which is in accordance with the higher denaturation temperature of \textit{oim} collagen (Leikina, Mertts et al. 2002; Miles, Sims et al. 2002; Kuznetsova, McBride et al. 2003).
5 The role of alpha-2 chain and mechanical force effects in collagen cleavage

(This chapter is adapted from (Chang, Flynn et al. 2012))

5.1 Introduction

Collagen cleavage is crucial for many biological and pathological processes such as tissue modeling and remodeling, growth, wound healing, cancer invasion and organ morphogenesis (Baragi, Qiu et al. 1997; Nagase and Visse 2003; Helary, Ovtracht et al. 2006). Normal physiological remodeling processes involve precisely regulated collagen degradation, where excessive or deficient degradation has been associated with numerous diseases. Accelerated breakdown of collagen has been associated with arthritis, atherosclerotic heart disease, tumor cell invasion, glomerulonephritis, and cell metastasis (Riley, Harrall et al. 1995; Barnes and Farndale 1999; Bode, Mosorin et al. 1999; McDonnell, Morgan et al. 1999). On the other hand, deficient degradation of collagen has been shown to result in spontaneous abnormal growth plate and increased trabecular bone in mice (Stickens, Behonick et al. 2004). Therefore, understanding the mechanism of collagenolysis and how it is possible to modulate its activity is crucial for developing treatments for a variety of diseases.

Collagenases of the matrix metalloproteinase (MMP) family (Nagase, Visse et al. 2006) are mammalian proteases involved in the physiological cleavage of collagen. The most prevalent collagenases, including MMP-1, MMP-8 and MMP-13, consist of propeptide, catalytic and hemopexin domains. They play an important role in cleaving collagen in the extracellular matrix, resulting in characteristic %- and %-fragments. The specific cleavage site is after the 775th residue (a Gly amino acid), in the sequences of G-IA for alpha-1 chain and G-LL for alpha-2 chain. Interestingly, there are several other sites in the
collagens that contain the same G-I/L bonds but they are not hydrolyzed. Since the amino acid sequence alone is not sufficient to explain the high specificity of collagen recognition by MMPs (Xiao, Addabbo et al. 2010), the local conformation at the vicinity of the cleavage site might play an important role in providing a recognition signal for MMPs.

The type I heterotrimer and homotrimer have distinct degradation behaviors (Figure 5.1). Type I homotrimers are found to be resistant to all mammalian collagenases (Narayanan, Meyers et al. 1984; Han, Makareeva et al. 2010; Makareeva, Han et al. 2010), with a cleavage rate much slower for homotrimers than for heterotrimers (Figure 5.1 (a)). The MMP resistance of homotrimers may play an important role in homotrimer-related diseases or in early development, during which necessary collagen degradation may be hindered with detrimental results. For example, it has been shown in fibers reconstituted from mouse tail tendon collagen that a minor fraction of homotrimer-based fibers may grow instead of being disassembled during tissue remodeling cycles, which may eventually result in tissue disorganization (Han, Makareeva et al. 2010).

Moreover, it has been shown experimentally that mechanical force applied to collagen molecules alters the cleavage rate (Ellsmere, Khanna et al. 1999; Flynn, Bhole et al. 2010; Adhikari, Chai et al. 2011). Notably, conflicting results between homo- and heterotrimer collagen have been reported. Experiments using a single, collagen homotrimer snippet (14 kDa) have shown that mechanical load induces an 81±3-fold increase in the rate of collagen proteolysis (Figure 5.1 (b)) (Adhikari, Chai et al. 2011). On the contrary, it has been found that mechanical load stabilizes heterotrimer (arranged in reconstituted fibrils) against enzymatic degradation (Figure 5.1 (c)) (Flynn, Bhole et al. 2010). The lower specificities in Figure 5.1 (a and b) can be at least partially explained by the lower reaction temperatures.

In this chapter, we perform molecular dynamics simulations of the cleavage sites from both the heterotrimer and homotrimer to examine the molecular origins of these discrepancies and to investigate possible molecular mechanisms by which mechanical force could mediate the rate of collagen cleavage.
Figure 5.1 Comparisons of the cleavage rates between the heterotrimer and the homotrimer without and with applying forces. (a), Cleavage rates of type I human heterotrimer and homotrimer (data from (Han, Makareeva et al. 2010)). The homotrimer is resistant to MMP cleavage (MMP-1, 25°C). (b), Force enhances the cleavage rate of the type I collagen homotrimer (data from (Adhikari, Chai et al. 2011), MMP-1, Room Temperature). (c), Force slows down the cleavage rate of type I bovine heterotrimer.
exposed to MMP-8 (data from (Flynn, Bhole et al. 2010), 37 °C). (Chang, Flynn et al. 2012)

5.2 Simulation details

5.2.1 Collagen molecule generation

We use full atomistic simulations to study the structure in the vicinity of the cleavage site of the type I mouse heterotrimer and homotrimer. The real sequences of type I alpha-1 and type I alpha-2 chains of *mus musculus* (wild type mouse) are used to generate the collagen molecules. The heterotrimer collagen molecule is built of two alpha-1 chains and one alpha-2 chain while the homotrimer collagen molecule is built of three alpha-1 chains. The sequences are adapted from NCBI protein database (http://www.ncbi.nlm.nih.gov/protein): AAH50014.1 for the alpha-1 chain and NP_031769.2 for the alpha-2 chain. The entire alpha-1 and alpha-2 chains consist of 1014 residues with repeated G-X-Y triplets, excluding the C-terminal and N-terminal sequences. Segments of real sequences with 63 residues long centered at the MMP-1 cleavage site (748th to 810th residues) of alpha-1 and alpha-2 chains are chosen to construct the heterotrimer and homotrimer collagen molecules. The chosen sequences are

**alpha-1:**

GPPGPAGEKGSAGSPGADGPAGSPGTPGPQG-
IAGQRGVVGLPGQRGERGFPGLPGPSGEPGKQGPS

**alpha-2:**

GPPGFVGEGKSAGSPGTPGATAPGTAGPQG-
 LLGAPGILGLPGSRGERGLPGIAGALGEPGPGIS.

Note that "-" indicates the scissile bond (the bond cut by MMP) at the cleavage site (G-IA for alpha-1 chain and G-LL for alpha-2 chain) of MMP-1.
The collagen molecules are created by inputting the sequences of three chains into the software THeBuScr (an interactive triple-helical collagen building script) (Rainey and Goh 2004), which enables us to build triple-helical molecules based on any specified amino acid sequence using conformations derived from statistical analyses of high-resolution X-ray crystal structures of triple-helical peptides. To neutralize the terminals, two ends of the heterotrimer and homotrimer are capped by assigning the first residues to ACE, the acetylated N-terminus, and the last residues to CT3, the N-Methylamide C-terminus. The length of the collagen molecule is about 170 Å and is solvated in a periodic water box composed of TIP3 water molecules. Visual Molecular Dynamics (VMD) (Humphrey, Dalke et al. 1996) is used to solvate the molecule and to neutralize the system by adding ions. The final solvated all-atom system contains ≈70,000 atoms, with a periodic box size of 230 Å × 60 Å × 60 Å.

5.2.2 All-atom modeling and equilibration

Full atomistic simulations are performed using NAMD (Nelson, Humphrey et al. 1996) and the CHARMM force field (MacKerell, Bashford et al. 1998) that includes parameters for hydroxyproline amino acids (Anderson 2005). This force field has been broadly validated for a variety of biochemical models of proteins including collagen (Gautieri, Buehler et al. 2009; Gautieri, Uzel et al. 2009; Srinivasan, Uzel et al. 2009). An energy minimization using a conjugate gradient scheme is performed before molecular dynamics simulations. Rigid bonds are used to constrain hydrogen atoms, thus allowing an integration time step of 2 fs. Nonbonding interactions are computed using a cut-off for neighbor list at 13.5 Å, with a switching function between 10 and 12 Å for van der Waals interactions. The electrostatic interactions are modeled by the particle mesh Ewald summation (PME) method. After energy minimization, the collagen molecule is fixed and the system is simulated with NPT at a constant temperature of 310 K and 1.013 bar pressure to reach an equilibrium box size with water molecules. Thermal fluctuations are thought to play an important role in the cleavage site of collagen molecule (Stultz, Nerenberg et al. 2008). To ensure a sufficient exploration on the configurations of collagen molecules, after reaching an equilibrium box size, both heterotrimer and homotrimer are equilibrated with a NVT ensemble at 310 K for 80 ns. During the
simulations, the configurations of collagen molecule are recorded every 20 ps, resulting in 4,000 frames (all used in the analyses). The total simulation takes about 12,800 CPU hours for each collagen molecule.

5.2.3 In silico mechanical tests

In order to examine the effects of applied force on the structural changes of heterotrimer and homotrimer collagen and to explore how force affects the MMP cleavage, we use steered molecular dynamics to apply a constant 100 pN force to each chain of heterotrimer and homotrimer after 80 ns equilibrium. The Cα atom of the 13th residue (the 5th Gly) of each chain is fixed and a constant force is applied at the Cα atom of the 49th residue (the 17th Gly) of each chain. With the constant force applied, the simulations of heterotrimer and homotrimer are performed for 25 ns. During the simulation the configurations of collagen molecule are recorded every 20 ps.

5.2.4 Analysis of the structures of collagen molecules

We focus on the structural analyses on the cleavage site, i.e. the 18 residues at the center of the collagen molecules used in our simulations. The unit heights and radii are calculated every 0.02 ns and an averaged of the data from 60 ns to 80 ns are taken to study the structural behavior of heterotrimer and homotrimer without applying force. For the analyses of the structural behavior of heterotrimer and homotrimer with applying force, an average of the data from 15 ns to 25 ns with force is used. The details on how we compute the structural parameters are provided in Chapter 3.5. The length of the cleavage site under applying force is calculated by summing the unit heights from the 5th height to 16th height.

5.2.5 Analysis of H-bonds at the cleavage site

Averaged N-O distances for heterotrimer and homotrimer with and without force at the cleavage site are calculated to study the H-bond forming. We calculate the distance from N of Gly residue to O of the residue at the X position of (GXY)n triplets. The bond distances with and without force are calculated by averaging from 60 ns to 80 ns for NVT simulation and from 15 ns to 25 ns for steered molecular dynamics simulation
respectively. There are three H-bonds can be formed at the Gly at the scissible bonds: from Gly at chain A to X position at chain B, from Gly at chain B to X position at chain C and from Gly at chain C to X position at chain A. If the N-O distance is smaller than 0.4 nm, it is assumed that there exists an H-bond. The life time of an H-bond is calculated by the ratio of the time an H-bond exists to the total time of observation. We calculate the life time of H-bond with and without force by observing from 60 ns to 80 ns for NVT simulation and from 15 ns to 25 ns for

5.3 Results

5.3.1 Molecular structures at the cleavage site

We find that the heterotrimer is thermally unfolded locally in the vicinity of the cleavage site at body temperature. In contrast, the homotrimer is thermally more stable and retains the triple helical structure characteristic of collagen molecules, suggesting stark differences in the structure of hetero- versus homotrimer molecules (Figure 5.2). Further quantitative geometric analysis supports this notion. Specifically, the unit heights and radii in the vicinity of the cleavage sites of the heterotrimer and homotrimer are shown in Figure 5.3. Comparing to a statistical analysis of high-resolution X-ray crystal structures of triple-helical peptides (Rainey and Goh 2002) (a study which reports that the unit height of the triple helical structure of collagen molecule is around 9 Å and the radius is about 2 Å), we find that the structure of the homotrimer in the vicinity of cleavage site is a rather stable triple-helical structure (Figure 5.2). However, decreased unit heights and increased radii are found in the vicinity of the cleavage site of the heterotrimer (Figure 5.3), supporting the visual analysis of the two structures. Notably, a large radius, 5.7 ± 0.4 Å, is found at the scissile bond of the heterotrimer, indicating that each chain is separated and exposed (Figure 5.2) and may play a role in the recognition of MMPs.
Figure 5.2 Snapshots of collagen molecules for heterotrimer (top) and homotrimer (bottom) with and without force. The vicinity of the cleavage site of the heterotrimer is thermally unfolded while the homotrimer behaves stable triple helical structure, indicating it is thermally stable. The applying force stabilizes the cleavage site of the
heterotrimer but does not affect the triple helical structure of the homotrimer. (Chang, Flynn et al. 2012)

Figure 5.3 Structural analyses of the vicinity of cleavage site of a type I mouse heterotrimer and homotrimer collagen molecule without applied force. (a) The unit heights of heterotrimer at the vicinity of the cleavage site. (b) The radii of heterotrimer at the vicinity of the cleavage site. (c) The unit heights of homotrimer at the vicinity of the cleavage site. (d) The radii of homotrimer at the vicinity of the cleavage site. This figure shows that the cleavage site of the homotrimer is a stable triple helical structure while the
heterotrimer is thermally unfolded at the cleavage site. The radius of the heterotrimer at the cleavage site is three times larger than it of the homotrimer indicating that each chain is exposed. (Chang, Flynn et al. 2012)

Figure 5.4 Changes in the lengths of the cleavage sites with applying force. (a), The lengths of the cleave sites (summation from the 5th height to 16th height) with and without force. (b), The time history of the increase of the length for heterotrimer and homotrimer. (c), The strain of each unit height at the vicinity of cleavage site. This figure shows that force induces both entropic and energetic strain to the heterotrimer but only induces energetic strain to the homotrimer. In the heterotrimer, the applied force reduces the entropy of the unfolded region and stabilizes the cleavage site by refolding it into the
triple helical structure. In contrast, the force only induces uniform strain in the homotrimer. (Chang, Flynn et al. 2012)

5.3.2 Mechanical testing of the cleavage sites

Next we explore the response of the molecular structures to the application of mechanical force. We find that the heterotrimer and homotrimer also show very different mechanical behaviors, as revealed in Figure 5.4. A detailed analysis of the strain distribution along the twisting axis of the homotrimer reveals that it features a uniform strain distribution with a total strain close to 0.03 (Figure 5.4 (c)). By assuming the diameter of collagen molecule is 15 Å, we obtain a 5.7 GPa Young's modulus, which is within reasonable bounds established in earlier studies (Sasaki and Odajima 1996; Gautieri, Vesentini et al. 2011) and suggests that the strain is entirely elastic strain. In contrast, for the heterotrimer, elastic strains similar to the homotrimer are only found at the regions outside of the vicinity of the cleavage site. Interestingly, large strains beyond the elastic strains are found highly localized in the vicinity of the cleavage site. Indeed, the maximum local strain is about 0.7 found at the scissile bond, which results in a small, 0.24 GPa, local Young's modulus indicating the applying force not only induces elastic strain but also stretches the collagen by reducing the entropy of the unfolded region.

The time history of the length of the cleavage site (Figure 5.4 (b)) provides further evidence that there are two regimes of stretching for the heterotrimer. This confirms that the applied force contributes to both entropic and energetic strain. In the first regime, the applied force reduces the entropy of the unfolded region and stabilizes the cleavage site by refolding it into the triple helical structure, which explains why a rather large local strain is observed in the vicinity of the cleavage site. In the second regime, the deformation of collagen molecules is in the energetic region, where the applied force induces elastic strain. Remarkably, after just 15 ns of application of force, the heterotrimer and the homotrimer feature the same lengths. This indicates that the applying the force pulls the heterotrimer out of the entropic elasticity region (Figure 5.4 (a)).
Figure 5.5 Structural analyses of the vicinity of cleavage site of type I mouse heterotrimer and homotrimer with constant force applied. (a) The unit heights of heterotrimer at the vicinity of the cleavage site. (b) The radii of heterotrimer at the vicinity of the cleavage site. (c) The unit heights of homotrimer at the vicinity of the cleavage site. (d) The radii of homotrimer at the vicinity of the cleavage site. (Chang, Flynn et al. 2012)
### Figure 5.6 Analyses of the N-O distance at the vicinity of the cleavage site of heterotrimer and homotrimer with and without forces.

(a) Average N-O distances for heterotrimer and homotrimer with and without force. The arrow indicates the distance point from N of Gly residue to O of the residue at the X position of (GXY)n triplets. Distances larger than 0.4 nm are marked in red.

(b) Bond life time of N-O distance point from the Gly at the scissile bond for heterotrimer.

(c) Bond life time for homotrimer. This figure shows that the force increases the H-bond life time of the heterotrimer but does not alter the H-bond life time of the homotrimer at the cleavage site, indicating that...
the force induces stabilization in the heterotrimer but not in the homotrimer. (Chang, Flynn et al. 2012)

Structural analyses show increases in the unit heights and decreases in the radii in the vicinity of the cleavage site (Figure 5.5), which provides additional direct evidence that applied force stabilizes (i.e., refolds) the cleavage site. This is consistent with experimental studies which show that applied force enhances thermal stability of collagen (Humphrey, Wells et al. 2005). It is worth mentioning that although a large force (~100 pN) is applied to speed up the force induced stabilization in our atomistic simulations, we anticipate a low force (~10 pN) within the entropic elastic region would result in the same stabilization mechanism in the heterotrimer since the conformational changes at the cleavage site of the heterotrimer are primarily due to entropic contributions (Figure 5.4 (c)).

5.3.3 Analysis of H-bonds

An analysis of averaged N-O distances and the H-bond life time in the vicinity of the cleavage site also shows that the heterotrimer thermally unfolded without applying force and that the thermal stability is increased by applying force at the cleavage site (Figure 5.6). Without applying force, all H-bonds close to the scissile bond are stably formed in the homotrimer, but none of them are formed in the heterotrimer (Figure 5.6). When force is applied, no significant change is observed in the homotrimer, i.e. the applied force does not alter the life time of stably formed H-bonds. Interestingly, a recovery of H-bonds close to the scissile bond is found in the heterotrimer, providing direct evidence for the stabilizing effect of mechanical force. The simulations show that H-bonds are broken 87% (no force) versus 17% (with force) of the time (Figure 5.6). Thus, we conclude that the applying force stabilizes the cleavage site of the heterotrimer. Most importantly, the same effect is not found for the homotrimer since it is thermally stable without applying force.
5.3.4 Analysis of the cleavage kinetics

It is widely accepted that collagen cleavage by MMPs involves three steps (Stultz 2002; Chung, Dinakarpandian et al. 2004; Han, Makareeva et al. 2010): enzyme binding, helix unwinding at the cleavage site and sequential hydrolysis of the chains. However, two different models have been proposed to explain the cleavage mechanism. In the first model, due to the thermal instability of the cleavage site, collagen is thermally unwound locally at the vicinity of the cleavage site before MMPs bind (Stultz 2002). In the second model, collagen is unwound after MMPs bind to the cleavage site (Chung, Dinakarpandian et al. 2004). Without a priori assumption about the effect of MMPs on the local triple helix unwinding, these two models have been integrated into a more general mechanism (Han, Makareeva et al. 2010) as shown in Figure 5.7. By setting \( k_2 = k_-2 = 0 \), the degradation scheme reduces to the first model and by setting \( k_3 = k_4 = k_-3 = k_-4 = 0 \), the scheme reduces to the second model.

Based on the cleavage kinetic scheme (Figure 5.7) and steady-state approximations, Han et al. (Han, Makareeva et al. 2010) have derived the Michaelis constant \( K_m \) and maximum cleavage rate \( V_{\text{max}} \) as follows:

\[
K_m = \frac{(k_2(k_3 + k_p) + k_-1(k_-2 + k_-3 + k_p))(1 + \frac{k_4}{k_-4})}{k_1(k_2 + k_-2 + k_-3 + k_p) + k_3 \frac{k_4}{k_-4}(k_-2 + k_2 + k_-1)}
\]

(5.1)

and

\[
V_{\text{max}} \approx \frac{k_p(k_1k_2 + \frac{k_4}{k_-4}k_3(k_-2 + k_-1)) [E]_{\text{total}}}{k_1(k_2 + k_-2 + k_-3 + k_p) + k_3 \frac{k_4}{k_-4}(k_-2 + k_2 + k_-1)}
\]

(5.2)

Here \([E]_{\text{total}}\) is the total enzyme concentration. Experimental studies have shown that for heterotrimer, \( K_m = 0.8 \pm 0.08 \mu \). Surprisingly, the homotrimer has a similar Michaelis
constant, $K_m = 0.9 \pm 0.2 \mu m$. Therefore, the smaller cleavage rate of homotrimer results entirely from the smaller maximum cleavage rate $V_{\text{max}}$.

We use $k_{-1} = 5 \times 10^{-3} \text{ s}^{-1}$ and $k_p = 1 \text{ s}^{-1}$ in our analysis because experimental results show that $k_{-1} \approx 5 \times 10^{-3} \text{ s}^{-1}$ (Ottl, Gabriel et al. 2000; Stultz, Nerenberg et al. 2008; Salsas-Escat, Nerenberg et al. 2010) and $k_p = 0.11 \sim 11 \text{ s}^{-1}$ (Fields, Van Wart et al. 1987; Stultz, Nerenberg et al. 2008; Salsas-Escat, Nerenberg et al. 2010). Note that $k_{-3}$ should be within the same order of $k_{-1}$, we have $k_p \gg k_{-3}$. Because atomistic simulations show that the homotrimer cleavage site is thermally stable, we assume $\frac{k_4}{k_{-4}} \ll 1$ for homotrimers. These assumptions lead to

$$K_m(\text{homotrimer}) = \frac{k_{-1}}{k_1} \left( \frac{1 + \frac{k_p K_U}{K_{-3} + K_p}}{1 + K_U} \right)$$

(5.3)

and

$$k_{\text{cat}}(\text{homotrimer}) = \frac{V_{\text{max}}(\text{homotrimer})}{[E]_{\text{total}}} \approx \frac{k_p K_U}{1 + K_U}$$

(5.4)

Here we have used $K_U = \frac{k_2}{k_{-2} + k_{-3} + k_p}$. We use experimental results by Han et al. (Han, Makareeva et al. 2010) ($K_m(\text{homotrimer}) = 0.9 \mu m$, and $k_{\text{cat}}(\text{homotrimer}) = 0.007 \text{ s}^{-1}$) to solve for unknown variables in Eq. (5.3) and Eq. (5.4).

The heterotrimer is thermally unfolded within nanoseconds and previous studies observe $\frac{k_4}{k_{-4}} \approx 1$ at room temperature (Stultz 2002), thus we are motivated to study the rate equations without assuming $\frac{k_4}{k_{-4}} \ll 1$ and neglecting all terms proportional to $\frac{k_4}{k_{-4}}$ at the first step. We assume that $k_{-2}$ is relatively small compared to $k_p$ and $k_{-1}$, which is reasonable since we do not expect the MMP would be likely to fold the unfolded collagen
into triple helical structure and we expect once MMP binds a vulnerable conformation of collagen it would likely cleave it. Based on these assumptions, we rearrange Eq. (5.1) and Eq. (5.2) into

$$K_m = \frac{k_{-1}}{k_1} \left( \frac{1 + \frac{k_p}{k_{-1}} K_U \left( 1 + \frac{k_4}{k_{-4}} \right)}{1 + \left( 1 + \frac{k_3}{k_1} \right) K_U + \frac{k_3}{k_1} K_U + \frac{k_4}{k_{-4}} \frac{k_{-1}}{k_p}} \right)$$

(5.5)

and

$$k_{cat} = \frac{V_{max}}{[E]_{total}} \approx \frac{k_p K_U \left( 1 + \frac{k_3}{k_1} \frac{k_4}{k_{-4}} \right)}{1 + \left( 1 + \frac{k_3}{k_1} \right) K_U + \frac{k_3}{k_1} K_U + \frac{k_4}{k_{-4}} \frac{k_{-1}}{k_p}}$$

(5.6)

Figure 5.7 Possible cleavage mechanisms of collagen molecules. There are two widely used models to explain the cleavage mechanism of collagen. In the first one, path I, collagen is thermally unfolded to a vulnerable state before the MMP binds to the collagen molecule (Stultz 2002). The other one, path II, MMP binds to and unwinds the collagen.
molecule before cutting it (Chung, Dinakarpandian et al. 2004). Our results show that the heterotrimer is thermally unfolded at the cleavage site within 80 ns, indicating that it primarily goes path I. Mechanical tests reveal that the force stabilizes the cleavage site of the heterotrimer by refolding it into a triple helical structure and thus protects against enzymatic breakdown. (Chang, Flynn et al. 2012)

Note that the remaining unknowns are $k_3$ and $k_4/k_1$. All other parameters, $K_U$, $k_{-1}$, and $k_p$ have been determined experimentally for homotrimers. It is worth noting that $k_3/k_1$ represents the ratio of the rate enzyme binds to the vulnerable state of collagen compared to the rate enzyme binds to the native state of collagen. Previous studies have shown that the local conformation of the cleavage site may provide a recognition signal for MMPs (Ottl, Gabriel et al. 2000; Salsas-Escat and Stultz 2010; Xiao, Addabbo et al. 2010). Thus, we expect that $k_3 > k_1$. Because there are no experimental results for the values of $k_3/k_1$ or $k_4/k_{-4}$, we study the values of $0 < k_3/k_1 < 100$ and $10^{-6} < k_4/k_{-4} < 1$. We consider an extreme case that $k_3/k_1$ is the only different parameter between heterotrimer and homotrimer to study how the thermal stability of collagen affects the degradation rate.

The results (Figure 5.8) show that within the range of the values we studied, there is not much change of $K_m$ which is consistent with the experimental finding that $K_m$ (heterotrimer) $\approx$ $K_m$ (homotrimer). When $k_4/k_{-4}$ is small enough, the $K_m$ remains constant. If both $k_3/k_{-4}$ and $k_2/k_1$ are large, a slightly smaller value of $K_m$ is obtained which is likely the case found in experiments that $K_m$ (heterotrimer) = 0.8 $\mu$m which is slightly smaller compared to $K_m$ (homotrimer) = 0.9 $\mu$m. On the other hand, large $k_4/k_{-4}$ and small $k_2/k_1$ would result in a larger value of $K_m$. 

95
Interestingly, the results of degradation rate reveal that increasing $\frac{k_4}{k_{-4}}$, i.e. decreasing the thermal stability, would result in a larger degradation rate. If $\frac{k_4}{k_{-4}} \ll 10^{-4}$, which is the case of homotrimer, the degradation rate is slow and is not sensitive to the thermal stability $\frac{k_4}{k_{-4}}$. While when $\frac{k_4}{k_{-4}}$ is large ($\frac{k_4}{k_{-4}} > 10^{-4}$) as the case of the heterotrimer, the degradation rate is fast and is dependent on the thermal stability. We find that an increase in thermal stability slows down the degradation rate. Moreover, when the rate MMP binds to vulnerable state is larger than it binds to the naive state $k_3 > k_1$, the degradation rate is more sensitive to the thermal stability. In summary, together with the atomistic simulations results, we find that the MMP resistance of homotrimer could be explained by the difference in thermal stability. The MMP resistance of the homotrimer results from its high thermal stability, which is in good accordance with the experimental results (Han, Makareeva et al. 2010).

We have shown that applied force stabilizes the heterotrimer. Therefore, it is clear from Figure 5.8 (b) that by decreasing $\frac{k_4}{k_{-4}}$, the applied force decreases the degradation rate. In contrast, for the homotrimer, applied force does not slow down the degradation by improving the thermal stability of homotrimer because the degradation rate of the
homotrimer is not sensitive to the thermal stability when $\frac{k_4}{k_{-4}} \ll 1$ as shown in Figure 5.8 (b). It can be also seen from Eq. (5.4) that the degradation rate of the homotrimer is not dependent to the thermal stability.

In summary, from the results of mechanical simulations, we see that applied force stabilizes the cleavage site of the heterotrimer, indicating that $K_4(\text{heterotrimer, force}) \gg K_4(\text{heterotrimer, no force})$. Because the cleavage rate of the heterotrimer depends on the thermal stability and the increase of thermal stability slows down the cleavage rate, we postulate that the applying force slows down the rate of thermally unfolding and should also slow down the cleavage rate of the heterotrimer. This suggests a molecular mechanism in which force induces stabilization of collagen and against enzymatic breakdown.

The applied force is less likely to slow down the cleavage rate by the same mechanism for the homotrimer because it is thermally stable without applying force. However, the remaining question is why the force speeds up the cleavage rate of the homotrimer (Adhikari, Chai et al. 2011). One possible explanation is that in the homotrimer, force results in homogeneously distributed elastic strain that eventually leads to uncurling of the molecule, suggesting that force leads to destabilization of the molecular structure. This explanation would not affect the fact that applying force slows down the cleavage of the heterotrimer since it primarily follows path I and the degradation rate is mainly governed by the thermal stability.

5.4 Summary

Molecular modelling shows that while the vicinity of the cleavage site of the heterotrimer is thermally unfolded at body temperature, the vicinity of the cleavage site of the homotrimer remains triple helical structure, as it is thermally stable. We conclude that the higher thermal stability of the homotrimer is also responsible for its greater resistance to MMP degradation in free solution. Moreover, we find that in heterotrimers, the application of force alters the structure of the cleavage domain such that the molecule exhibits enhanced stability and behaves in a manner similar to the homotrimer. Thus,
force appears to "switch" the heterotrimer from unstable to stable by analogy. Direct quantitative evidence was found to support this finding, as established by the observation that the molecular radii decreases and that the H-bond lifetime increases, by forming a triple helical geometry at the cleavage site. The reformation of a triple helical structure should protect against enzymatic breakdown if unwinding is necessary for alpha chains to enter the cleavage site on MMPs. If, however, thermal fluctuations at the cleavage site make catalysis more difficult, then the applied force with its associated triple helix reformation could accelerate the cleavage rate. Such behavior may have been observed in the work of Adhikari et al. (Adhikari, Chai et al. 2011), but the nearly 100 fold acceleration in MMP-1 cleavage is difficult to fully explain. In the homotrimer, force results in homogeneously distributed elastic strain that likely leads to frank unfolding of the molecule, suggesting that force leads to destabilization of the entire molecular structure.

Our study suggests a molecular mechanism by which thermal stability plays a role in the cleavage mechanism and the potential explanation of the force-stabilization results found in experimental studies as well as the distinct behaviours of the heterotrimer and homotrimer collagen. In addition to providing a possible explanation for the seemingly conflicting experimental results reported in earlier papers based on a mechanistic model, we anticipate that our study is crucial for developing new biomaterials that serve as platforms for treatments for a variety of diseases. For example, through tuning the mechanical force, one can precisely regulate collagen degradation. This concept can also find application in tunable collagen-based biomaterials, and may be important to understand the development of tissue.
6 Structural and mechanical differences between wild type and oim collagen microfibrils

6.1 Introduction

Collagen molecules, produced by cells, are stacked together in a characteristic D-period to form collagen fibrils which have diameters of ~100 nm. Collagen fibrils are the basic components of collagen fiber which forms connective tissues. Recently, the in situ structure of a full length type I collagen fibril (Protein Data Bank identification code 3HR2) has been revealed (Orgel, Irving et al. 2006). It has been shown that the collagen fibril has a triclinic unit cell with dimensions $a=40.0\ \text{Å}$, $b=27.0\ \text{Å}$, $c=678\ \text{Å}$, $\alpha=89.2^\circ$, $\beta=94.6^\circ$, and $\gamma=105.6^\circ$. The fibril has a gap region with a length of 0.54 $D$ and an overlap region with a length of 0.46 $D$. Here $D$ (~67 nm) denotes the length of the D-period of collagen fibril. This model allows us to examine the structure of type I collagen fibril with molecular details.

In this chapter, we scale up our modelling from single collagen molecular level (as presented in Chapter 4) to the collagen fibril level. We use a molecular simulation approach to study the mechanical and structural differences between type I heterotrimer and homotrimer microfibril models using real sequence mouse collagen molecules (Collagen, type I, alpha 1 and alpha 2 chain precursor [Mus musculus]). Different stress levels along both the axial and lateral directions of the collagen microfibril are performed to investigate the differences on the mechanical properties between the normal and the oim collagen microfibril. We also analyze the structural differences between the microfibril models to explore the effects of the replacement of the alpha-2 chain with alpha-1 chain at the molecular level. The results presented in this chapter provide insights into how the oim mutations alter the structures and properties of collagen fibrils.
6.2 Materials and methods

6.2.1 Atomistic collagen microfibril

The collagen microfibril models are generated based on the *in situ* structure of full length collagen type I molecule (Protein Data Bank identification code 3HR2), which has a triclinic unit cell \((a \approx 40.0 \text{ Å}, \ b \approx 27.0 \text{ Å}, \ c \approx 678 \text{ Å}, \ \alpha \approx 89.2^\circ, \ \beta \approx 94.6^\circ, \ \text{and} \ \gamma \approx 105.6^\circ)\). Note that the structure reported in (Orgel, Irving et al. 2006) includes only backbone alpha carbons and the primary sequence of *rattus norvegicus*, we therefore use homology modeling as described in (Gautieri, Vesentini et al. 2011) to obtain a full-atomistic structure with the mus collagen sequence. The real sequences of type I alpha-1 and type I alpha-2 chains of mus musculus (wild type mouse) are used. The heterotrimer collagen microfibril model is built of two alpha-1 chains and one alpha-2 chain while the homotrimer collagen microfibril model is built of three alpha-1 chains. The sequences are adapted from NCBI protein database (http://www.ncbi.nlm.nih.gov/protein): AAH50014.1 for alpha-1 chain and NP_031769.2 for alpha-2 chain. Full sequences are provided in Chapter 3.2. In both collagen microfibril models, ions are added to neutralize the system.

6.2.2 All-atom equilibration

Full atomistic simulations are performed using modeling code LAMMPS (Plimpton 1995) (http://lammps.sandia.gov/) and the CHARMM force field (MacKerell, Bashford et al. 1998) that includes parameters for hydroxyproline amino acids based on a model put forth by Anderson (Anderson 2005). An energy minimization using a conjugate gradient scheme is performed before molecular dynamics simulations. Rigid bonds are used to constrain covalent bond lengths and an integration time step of 1 fs is used. Nonbonding interactions are computed using a cut-off for neighbor list at 13.5 Å, with a switching function between 10 and 12 Å for van der Waals interactions. The electrostatic interactions are modeled by the ewald/n style, which performs standard coulombic Ewald summations in a more efficient manner (in't Veld, Ismail et al. 2007). After energy
minimization, the collagen molecule is simulated through 5 ns at a constant temperature of 310 K in molecular dynamics simulations to obtain an initial equilibrium structure.

6.2.3 Mechanical testing

In order to investigate the mechanical properties of both the hydrated and dehydrated atomistic microfibril models, we perform NPT molecular dynamics simulations with different constant mechanical stress levels (in the range from 0 to 150 MPa) in compression (along the fibril \(a\)-axis) or in tension (along the fibril \(c\)-axis) for each model. The pressures on the other axes are kept constant at 1.013 bar pressure. In each simulation, we keep monitoring the convergence of the lattice parameters. Every simulation is performed until the lattice parameters on all directions are converged. Most of the simulations are performed up to 50 ns and the longest simulation is performed for 90 ns. After reaching the equilibrium, we use the last 1 ns of molecular dynamics simulations for computing the unit cell lattice parameters corresponding to each load applied.

6.3 Results

6.3.1 Mechanical properties of hydrated collagen microfibrils

Tensile tests of hydrated full atomistic microfibrils along the collagen \(c\)-axis reveal higher modulus of the normal (heterotrimer) compared with the \(oim\) (homotrimer) microfibril. Full atomistic modelling reveals nonlinear stress-strain relations for both normal and \(oim\) collagen microfibril as shown in Figure 6.1. Each point in the figure is obtained from an individual long time NPT simulation to ensure the equilibrium of the collagen microfibril structures at the given stress level. Here we use third order polynomial equations to fit the stress-strain curves for both collagen microfibrils as shown in the figure (solid lines). The fitting result of the normal collagen microfibril is \(\sigma = -0.9 + 143.6\varepsilon + 480.6\varepsilon^2 + 7136.3\varepsilon^3\) while the fitting result of the \(oim\) collagen microfibril is \(\sigma = -0.98 - 38.0\varepsilon + 1740.4\varepsilon^2 + 2543.4\varepsilon^3\). These results reveal the modulus-strain relations as \(E_{\text{normal}} = 143.6 + 961.2\varepsilon + 21408.9\varepsilon^2\) and \(E_{\text{oim}} = -38 + 3480.8\varepsilon + 7630.2\varepsilon^2\).
Figure 6.1 Stress-strain curves of normal and oim collagen microfibrils along the axial direction. The results reveal nonlinear stress-strain relations for both normal and oim microfibrils. At low stress level (0–50 MPa), both normal and oim collagen microfibrils are flexible and sensitive to the mechanical forces. While when the applied stress larger than 50 MPa, both collagen microfibrils enter linear regions where the modulus becomes much higher.

As shown in the figure, at low stress level (0–50 MPa), both normal and oim collagen microfibrils are flexible and sensitive to the mechanical forces. When the applied stress reaches ~50 MPa, both collagen microfibrils enter linear regions where the modulus becomes much higher. The nonlinear stress-strain behaviour is consistent with previous computational study (Gautieri, Vesentini et al. 2011). We find that at low strain level ($\varepsilon = 0.06$), the normal collagen microfibril has a modulus of 278.3 MPa while the oim collagen microfibril has a reduced modulus of 198.3 MPa. The difference between the normal and oim fibril is 29%. It is worth noting that the oim collagen microfibril also has a reduced modulus at higher strain region. At a strain of 0.15, the normal collagen microfibril has a modulus of 769.5 MPa while the oim collagen microfibril has a reduced modulus of 655.8 MPa. The difference between the normal and oim at larger strain ($\varepsilon = 0.15$) is 14.8%, indicating the reduction of the modulus is more severe at the lower strain region for the oim collagen microfibril. These results indicate that the replacement of the alpha-2 chain with an alpha-1 chain in the collagen molecular level results in the
formation of a collagen microfibril with reduced tensile modules. However, the mutation does not affect the deformation mechanism of collagen microfibrils. As shown in the figure, both normal and oim collagen microfibrils have the nonlinear stress-strain relations and transitions from a toe region to a linear region are found in both collagen microfibrils.

6.3.2 The oim collagen microfibrils have larger lateral spacing

In order to understand the molecular differences between the normal and oim collagen microfibrils and the molecular mechanism of the reduced modulus of oim collagen microfibrils, we further analyze the structures of the normal and oim collagen microfibril models at different stress levels.

![Graph showing a-axis lengths of normal and oim collagen microfibrils](image)

Figure 6.2 Analysis of the a-axis lengths of the normal and oim collagen microfibrils at different stress levels. The results show that as the stress along the collagen axis direction increases, the lengths of the a-axis of both the normal and oim collagen microfibrils decrease. At all stress levels studied in this work, the oim collagen microfibril has a larger a-axis length (≈1-2 Å larger) compared with the normal collagen microfibril.

Firstly, we analyze the a-axis and b-axis lengths of the normal and oim collagen microfibrils at different stress levels. Figure 6.2 shows the results of the a-axis lengths of both collagen microfibrils at different stress levels. The result reveals that as the tensile stress along the collagen axis increases, the lengths of the a-axis of both the normal and
oim collagen microfibrils decrease. More importantly, at all stress levels, the oim collagen microfibril has a larger a-axis length (~1-2 Å larger) compared with the normal collagen microfibril. Figure 6.3 further shows the results of the b-axis lengths of both collagen microfibrils at different stress levels. Similar to the results of the a-axis lengths, we find that as the tensile stress along the collagen axis increases, the lengths of the b-axis of both the normal and oim collagen microfibrils decrease, and, at all stress levels, the oim collagen microfibril has a larger b-axis length (~2 Å larger) compared with the normal collagen microfibril.

![Graph showing b-axis lengths vs stress](image)

Figure 6.3 Analysis of the b-axis lengths of the normal and oim collagen microfibrils at different stress levels. The results show that as the stress along the collagen axis direction increases, the lengths of the b-axis of both the normal and oim collagen microfibrils decrease. At all stress levels studied in this work, the oim collagen microfibril has a larger b-axis length (~2 Å larger) compared with the normal collagen microfibril.

Both Figure 6.2 and Figure 6.3 reveal that the oim collagen molecules have larger lateral spacing when forming into collagen microfibrils. This finding is consistent with previous experimental studies. From computations of the volume fraction of the water, Miles et al. (Miles, Sims et al. 2002) suggested an increase in the lateral packing of 1.4 Å between collagen molecules in homotrimer collagen compared with the heterotrimer collagen. Our studies provide the first full atomistic computational models with full molecular information for both normal and oim collagen microfibrils. The increases of the lateral
spacing between collagen molecules in the oim collagen microfibrils are resulted from the formation of kinks of the homotrimer as we found at the single molecular level (Chapter 4). Direct comparisons on the structural and packing differences between the normal and the oim collagen microfibrils are provided in section 6.3.4.

It is worth noting that our results show that a tensile stress (~40 MPa) is required for the microfibril unit cell to maintain the dimension of type I collagen fibrils in situ (a~40.0 Å, b~27.0 Å, c~678 Å). This might suggest that there exists pre-stress or pre-strain within type I collagen fibrils in situ. However, other factors such as cross-links and solvent might also affect the structure of type I collagen fibrils. The cross-links are believed to play an important role in the structure of type I collagen fibrils. No cross-links is included in the model because of the limitation of the model that is made of "one" molecule. We anticipate that the cross-links can at least contribute to a part of the pre-stress.

![Graph](image)

Figure 6.4 Analysis of the end-to-end distances of the normal and oim collagen microfibrils at different stress levels. Two regimes of the relations between the end-to-end distance and the stress are found. At small deformation, the end-to-end distance of the collagen molecule increase linearly until the stress reaches ~50 MPa. Beyond this point, the relations between the end-to-end distance of the collagen molecules and the
stress enter the second regime where the end-to-end distance increases linearly with the stress while the slope is distinctly smaller.

Figure 6.4 shows the analysis of the end-to-end distance of the normal and oim collagen microfibrils at different stress levels. The results provide more insights into the molecular deformation mechanisms during stretching the hydrated collagen microfibrils. The molecular origins of the nonlinear stress-stress curves can be explained by analysing the molecular structures. Two regimes of the relations between the end-to-end distance and the stress are found as shown in Figure 6.4, which are in accordance with the stress-strain curves. At small deformation, the end-to-end distance of the collagen molecule increase linearly until the stress reaches ~50 MPa. Beyond this point, the relations between the end-to-end distance of the collagen molecules and the stress enter the second regime where the end-to-end distance increases linearly with the stress while the slope is distinctly smaller. This deformation mechanism is consistent with previous computational modelling in which human type I collagen sequences are used (Gautieri, Vesentini et al. 2011).

Figure 6.5 Analysis of the average unit height of the normal and oim collagen microfibrils at different stress levels.
It is worth noting that our results reveal an end-to-end distance of ~250 nm at zero stress for the normal mouse collagen microfibril and an end-to-end distance of ~300 nm when stress approaching 200 MPa. These values match very well with the results obtained from human collagen microfibrils (Gautieri, Vesentini et al. 2011), suggesting that wild-type mouse collagen microfibrils have similar properties as the human type I collagen microfibrils.

The results of the relations between the end-to-end distance and stress for the $oim$ collagen microfibrils show similar behaviour as the normal collagen microfibrils. Two regimes of the curves are also found for the case of $oim$ collagen microfibrils. As shown in Figure 6.4, the slopes of the curves in both regions are close when comparing the normal collagen microfibrils with the $oim$ collagen microfibrils. These results suggest that the replacement of the alpha-2 chain with an alpha-1 chain does not alter the deformation mechanism of the collagen microfibrils. However, the end-to-end distances of the $oim$ collagen microfibrils are smaller at all stress levels compared with the normal collagen microfibrils. The smaller end-to-end distances can be explained by the formation of kinks in the $oim$ collagen molecules. These results are consistent with the simulations of the heterotrimer and the homotrimer at the single molecular level (Chapter 4).

We further analyze the average unit height of the normal and $oim$ collagen microfibrils at different stress levels as shown in Figure 6.5. The results reveal similar trend as the end-to-end distances. Two regimes are found in the relations between the average unit height of the collagen molecules and the applied stress. In both regimes, the average unit height of the collagen molecules increases linearly as the stress increases while the slope of the second regime is distinctly smaller. The results of the $oim$ collagen microfibrils show the same trend as the normal collagen microfibrils while the $oim$ collagen microfibrils have a slightly smaller (~0.2 Å) average unit height compared with the normal collagen microfibrils. These results further confirm that the replacement of the alpha-2 chain with an alpha-1 chain affects the structures of the collagen molecules while does not alter the deformation mechanisms of the collagen microfibrils.
6.3.3 Mechanical properties of dry collagen microfibrils

We also perform mechanical tests of the dehydrated (dry) collagen microfibrils to study the effect of water solvent on the mechanical properties of the normal and oim collagen at fibril level. Figure 6.6 shows the stress-strain curves of the normal and oim collagen microfibrils in hydrated and dry conditions. We perform NPT simulations for both the normal and oim microfibril models under five stress levels (0, 30, 50, 75 and 150 MPa) as shown in the figure. The solid lines in the figure represent the stress-strain curves obtained from curve fitting. We find a stress-strain relation of $\sigma = 1.4 + 1048.0\varepsilon + 5313.3\varepsilon^2$ for the normal collagen microfibril and a stress-strain relation of $\sigma = 1.5 + 161.2\varepsilon + 7347.0\varepsilon^2$ for the oim collagen microfibril. These results reveal the modulus-strain relations as $E_{\text{normal}} = 1048.0 + 10626.6\varepsilon$ and $E_{\text{oim}} = 161.2 + 14692.0\varepsilon$.

![Stress-strain curves of the normal and oim collagen microfibrils in hydrated and dry conditions.](image)

Figure 6.6 Stress-strain curves of the normal and oim collagen microfibrils in hydrated and dry conditions.

We find that at low strain level ($\varepsilon = 0.06$), the normal collagen microfibril has a modulus of 1.7 GPa, while the modulus approaches ~2.6 GPa for larger strain ($\varepsilon = 0.15$) in the dehydrated condition. The modulus of the dehydrated collagen microfibril is much larger than the modulus we found for the normal collagen microfibril in the hydrated condition (278.3 MPa when $\varepsilon = 0.06$; 769.5 MPa when $\varepsilon = 0.15$). The results are in good agreement with previous computational study on human collagen microfibril (Gautieri, Vesentini et al. 2011) and experimental studies using Atomic Force Microscopy (AFM) (Aladin,
Cheung et al.; van der Rijt, van der Werf et al. 2006). These results are consistent with previous studies which have shown that the water plays an important role in the mechanics of collagen microfibrils.

In the dehydrated condition, the oim collagen microfibril has reduced modulus when compared with the normal collagen microfibril as shown in Figure 6.6. At low strain level, \( \varepsilon = 0.06 \), the oim collagen microfibril has a modulus of 1.0 GPa, while the modulus approaches \(-2.4\) GPa for larger strain \( \varepsilon = 0.15 \) in the dehydrated condition. The difference between the normal and oim fibril is 41\% at low strain level \( \varepsilon = 0.06 \) while the difference between the normal and oim reduces to a significant smaller value of \(-7.7\%\) at larger strain \( \varepsilon = 0.15 \), indicating the reduction of the modulus is more severe at lower strain region for the oim collagen microfibrils, which is similar with our findings in the hydrated condition. These results suggest that the replacement of the alpha-2 chain with an alpha-1 chain in the collagen molecular level results in the formation of a less dense packing of collagen microfibril which leads to the reduced tensile modulus in both hydrated and dehydrated conditions. However, at larger strain level, the reduction on the modulus of the oim collagen microfibril becomes insignificant. We anticipate it is because the force stretches the kink regions and results a more dense packing of the oim collagen microfibril at larger strain region.

We further analyze the a-axis and b-axis lengths of the dehydrated normal and oim collagen microfibrils at different stress levels. Figure 6.7 shows the results of the a-axis lengths of both collagen microfibrils at different stress levels. The results show significant reductions on the the length of the a-axis of both the normal and oim collagen microfibril when the collagen microfibril are dehydrated. The a-axis of the normal collagen microfibril reduces from \( \sim 43\) Å in the hydrated condition to \( \sim 34\) Å in the dehydrated condition. This indicates that when the collagen microfibril is dehydrated, the collagen molecules rearranged into a more dense packing due to the removal of the water molecules. The more dense packing of the collagen microfibril thus results a much larger modulus of the collagen microfibril in the dehydrated condition.
Similar to the results of the hydrated collagen microfibrils, we find that as the tensile stress along the collagen axis increases, the lengths of the \( a \)-axis of both the normal and \( oim \) collagen microfibrils decrease in the dehydrated condition. More importantly, the \( oim \) collagen microfibril has a larger \( a \)-axis length (~2 Å larger) compared with the normal collagen microfibril even in the dehydrated condition. These results suggest that the larger lateral spacing between the \( oim \) collagen molecules are resulted from the collagen molecular structures (for example, kink formations) while cannot be solely explained by the higher hydrophobicity of the alpha-2 chain as suggested by Miles et al. (Miles, Sims et al. 2002).

Figure 6.8 shows the results of the analysis of the \( b \)-axis lengths of the dehydrated normal and \( oim \) collagen microfibrils at different stress levels. The results are similar to the findings on the lengths of the \( a \)-axis. A significant reduction on the lengths of the \( b \)-axis is found (from ~28 Å to ~24.8 Å) when the normal collagen microfibril is dehydrated. At all stress levels studied in this work, the \( oim \) collagen microfibril has larger \( b \)-axis lengths (~1 Å larger) compared with the normal collagen microfibril, indicating a less dense packing of the \( oim \) collagen molecules at the fibril level in the dehydrated condition.

![Figure 6.7 Analysis of the \( a \)-axis lengths of the dehydrated normal and \( oim \) collagen microfibrils at different stress levels. The results show that as the stress along the collagen axial direction increases, the lengths of the \( a \)-axis of both the normal and \( oim \) collagen microfibrils decrease. At all stress levels studied in this work, the \( oim \) collagen microfibril has larger \( a \)-axis lengths (~2 Å larger) compared with the normal collagen microfibril, indicating a less dense packing of the \( oim \) collagen molecules at the fibril level in the dehydrated condition.](image-url)
microfibril has a larger $a$-axis length (~2 Å larger) compared with the normal collagen microfibril.

We have shown that the oim collagen microfibrils have larger lateral spacing between collagen molecules and therefore have reduced modulus along the fibril axial direction. We further study the transverse modulus of both the normal and oim collagen microfibrils to investigate how the larger lateral spacing affects the mechanical property of the collagen microfibrils in the transverse direction.

![Figure 6.8](image)

**Figure 6.8** Analysis of the $b$-axis lengths of the dehydrated normal and oim collagen microfibrils at different stress levels. The results show that as the stress along the collagen axial direction increases, the lengths of the $b$-axis of both the normal and oim collagen microfibrils decrease. At all stress levels studied in this work, the oim collagen microfibril has a larger $b$-axis length (~1 Å larger) compared with the normal collagen microfibril.

We find that the oim collagen microfibril also has a decreased transverse modulus. In silico compression tests along the transverse direction of the collagen microfibril model suggest higher transverse modulus of the normal compared with the oim microfibril. The results reveal nonlinear stress-strain relations for both the normal and oim microfibrils as shown in Figure 6.9 ($E_{\text{normal}} = 1.05$ GPa and $E_{\text{oim}} = 0.79$ GPa, when the applied stress is 150 MPa). The difference between the normal and oim fibril is 25%. The results suggest that the increase of lateral distance in oim fibril, due to local kink formations, not only...
causes a reduction of the modulus along the axial direction but also causes a reduction of the modulus along the transverse direction. Figure 6.9 shows the stress-strain curve as determined by a polynomial fitting ($\sigma = 2.7 - 402.7\varepsilon + 1629.2\varepsilon^2$ for the normal and $\sigma = 0.7 - 431.9\varepsilon + 738.4\varepsilon^2$ for the oim microfibril) for four different applied stress levels (1 atm, 30 MPa, 75 MPa and 150 MPa).

Figure 6.9 Stress-strain curves of compression tests on the microfibril models. The transverse modulus (along $a$-axis direction of the microfibril model) is determined from the tangent of the stress strain curve at each applied stress and is found to be 25% less for oim at applied stress of 150 MPa when compared to normal (heterotrimer) collagen fibrils.

6.3.4 Structural analysis of the normal and oim collagen microfibrils

Full atomistic models allow us to investigate the role of alpha-2 chain in type I collagen microfibril with molecular details. In this section, we provide further structural analysis to investigate the molecular origins on the increased larger lateral spacing.

By modelling a single collagen molecule we have shown that the replacement of alpha-2 chain in type I collagen molecule results the kink formation (see Chapter 4 for more detail). We anticipate that throughout the full length of the type I collagen molecule, there
are more kinks which are the molecular origins of the increased lateral spacing of the oim collagen microfibrils. Indeed, by directly comparing the structures of the full length normal and oim collagen molecules, we find that the replacement of the alpha-2 chain in type I collagen molecule results in the formation of more local kinks as shown in Figure 6.10. This confirms that the local kinks in oim collagen molecules affect the packing of collagen molecules into fibrils and result in a less dense structure. It is clearly shown in the figures that the oim collagen molecule has more kinks with larger angles that is the molecule mechanism of the increased lateral spacing. The larger lateral spacing between oim collagen molecules results in looser packing of the microfibril in the lateral directions which leads to a decreased modulus of the oim microfibril.

Figure 6.11, Figure 6.12, Figure 6.13 and Figure 6.14 further shows how the kinks affect the packing of the collagen microfibrils. In these figures, the blue molecules represent the collagen molecules in the overlap region and the red molecules represent the collagen molecules in the gap region. Both normal and oim collagen molecules form local kinks which are more clearly observed in the gap region compared with the overlap region of the fibrils. Local kinks are larger in the oim compared with the normal molecule. As indicates in the figures, it clearly shows that these kinks lead to an increased distances between the collagen molecules to its end-to-end direction (shown by solid lines). Interestingly, although local kinks alter the lateral packing of collagen molecules, we find that the replacement of the alpha-2 chain does not affect the specific D-period of the oim collagen microfibril, which can be seen from Figure 6.15. As shown in the figure, both collagen microfibrils have the specific D-period and the overlap and gap regions.
Figure 6.10 Direct comparison of the structures of normal and oim collagen molecules. The blue regions indicate molecules within the overlap region while the red regions indicate molecules within the gap region. The oim collagen molecule has more kinks (mostly in the gap region), which lead to the increase of lateral distance and the reduction of modulus.

Figure 6.11 Structures of normal collagen in x-y plane, several kinks are observed while the collagen molecules mostly align with the end-to-end direction (shown by solid line) of the collagen molecule.
Figure 6.12 Structures of normal collagen in x-z plane, several kinks are observed while the collagen molecules mostly align with the end-to-end direction (shown by solid line) of the collagen molecule.

Figure 6.13 Structures of oim collagen in x-y plane, oim collagen molecule has more kinks (mostly in the gap region), which lead to the increase of lateral distance and the reduction of modulus.
Figure 6.14 Structures of oim collagen in x-z plane, oim collagen molecule has more kinks (mostly in the gap region), which lead to the increase of lateral distance and the reduction of modulus.

We also analyze the twisting of the normal and oim collagen molecules by analyzing the 3D structures of the Gly residues. Glycine (Gly) is found at every third residue along alpha chains of collagen (Kadler, Holmes et al. 1996). The analysis of the centers of masses of Gly residues provides quantitative information on how the local kink formations affect the twisting of collagen molecules (Figure 6.16). The figure shows that local kinks are more clearly observed in the oim collagen molecules. It is not surprising that these kinks would affect the packing of the collagen fibrils.

We further analyze the projections of the positions of centers of masses of Glycine on the plan perpendicular to the end-to-end direction of the collagen molecule. The results are shown in the Figure 6.17. The color of each point represents the order of index of (Gly-X-Y)_n triplet. Note that there are total 1014 amino acids in each chain of the collagen molecule with 338 (Gly-X-Y)_n triplet. Therefore, there are 338 points for each collagen molecule. Gly residues are more scattered in the oim collagen molecule compared with the normal collagen molecule. This supports our finding that there are more kink formations in the oim collagen molecule within the microfibril. Due to the more local...
kink formations in the \textit{oim} collagen molecule, the inter-molecular lateral distance is increased in the \textit{oim} collagen microfibril.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6_15}
\caption{Figure 6.15 Direct comparison between the structures of normal and \textit{oim} collagen microfibrils.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6_16}
\caption{Figure 6.16 Analysis the twisting of the normal and \textit{oim} collagen molecule by analyzing the 3D structures of the Gly residues.}
\end{figure}
6.4 Summary

In this chapter, we use molecular simulations to study the mechanical and structural differences between the heterotrimer and homotrimer collagen microfibrils. The collagen microfibril models are generated based on the *in situ* structure of full length collagen type I molecule with the actual amino acid sequence of real mouse collagen sequence (Collagen, type I, alpha 1 and alpha 2 chain precursor [Mus musculus]). Our studies provide fundamental insight into the effect of the loss of alpha-2 chain at the molecular level and help understanding the molecular origin of the bone brittle disease at much larger length-scales.

In summary, atomistic models suggest that the oim collagen fibril has reduced modulus in both lateral and axial directions and reveals that the absence of the alpha-2 chain alters the inter-molecular packing because of local kink formations. In silico mechanical tests and structural analyses of collagen microfibril models reveal nonlinear stress-strain relations for both normal and oim microfibril models, which are consistent with previous experiments that have shown the toe, heel and linear regions in the stress-strain curve of a
single type I collagen fibril. Our results suggest that the oim microfibril is less dense compared with the normal microfibril as a result of local kink formations which explains the reduced modulus in the oim microfibril. Furthermore, we find that the increased lateral spacing between collagen molecules of oim microfibril occurs in both dehydrated and hydrated conditions, which coincides with experiments and suggests that the increased lateral spacing is not solely due to the higher hydrophobicity of the alpha-2 chain.

The increase in the lateral distance will further affect the organization of inter-fibrillar cross-links, non-collagenous proteins and proteoglycans in tendon. This may be the reason why tendons from oim mice exhibit lower ultimate stress compared to the wild type. A decrease in the inter-molecular cross linking in oim collagen fibrils from the tail tendon has been reported by Sims et al. (Sims, Miles et al. 2003). These cross-links are significantly reduced in the collagen from oim bone tissue by 27 % (Sims, Miles et al. 2003). Insufficient cross-links will affect the stabilization of the fibril and thus alter its mechanical properties. We anticipate that these kink formations along the collagen molecules and the reduction of immature cross-links in oim can explain the lower elastic moduli measured from the oim collagen fibril.

The individual collagen fibril is recognized to be the building block of many important biological tissues that cannot actually be defined as materials or as structures but as a consequence of this combination during which structural elements merge from the nanoscale to the macroscale through a hierarchical formation. Since the properties of a structure at a given hierarchical level controls the assembly at the next level, a significant relationship between the mechanics of collagen fibrils, increased mineralization and failure process of oim bones is expected. The proposed full atomistic normal and oim microfibrils, which provide molecular details on the packing of collagen molecules, allow us to study the alterations on the space for mineral depositions and their effects on the mechanical properties of mineralized collagen fibrils, which are discussed in the next chapter.
7 Mineralized collagen microfibrils of wild type and oim collagen

7.1 Introduction

Bone is a remarkable hierarchical biomaterial consisting of two major constituents: collagen molecules and inorganic mineral (Glimcher, Krane et al. 1967; Fratzl, Gupta et al. 2004; Fratzl 2008). Approximately, 35% of bone tissue by weight is organic matrix (90-90% are collagen), and 65% of bone tissue is inorganic minerals (Glimcher, Krane et al. 1967). The mineralized collagen fibrils are highly conserved across species and are the fundamental building block of bone (Jager and Fratzl 2000; Gupta 2006; Fratzl and Weinkamer 2007). The inorganic minerals in bone are primary hydroxyapatite (Ca_{10}(PO_4)_6(OH)_2) or crystals that have similar structure as hydroxyapatite (Bogert and Hastings 1931; Carlstrom 1955; Posner, Stutman et al. 1960; Herman and Dallemagne 1961; Jager and Fratzl 2000), although the exact compositions of the minerals in bone is still not fully understood (Glimcher, Krane et al. 1967).

In bone, the collagen fibril is the substrate which acts as the nucleation catalyst in the formation of the mineral phase (Glimcher, Krane et al. 1967). The initiation of the mineralization in bone is governed by the structural components of the organic matrix. It has been proposed that the packing of collagen molecules is a necessary structural feature which provides the space, mainly in the gap region, within which the mineral-phase transformation occurs (Glimcher, Krane et al. 1967). Therefore, the assembly of collagen molecules at the collagen fibril level is likely to play a significant role in the mineralization of mineralized collagen fibril.

Recently, experimental studies have revealed that hydroxyapatite occurs largely as elongated plate-like structures external to and oriented parallel to the collagen fibrils (McNally, Schwarcz et al. 2012). Figure 7.1 shows a bright field TEM image of longitudinal section of human bone. McNally et al. have suggested that about 70% of the hydroxyapatite occurs as mineral structures external to the fibrils while 30% of the
hydroxyapatite occurs within the collagen fibrils and is found constrained to the gap regions as shown in the figure (McNally, Schwarcz et al. 2012).

Figure 7.1 A bright field TEM image of longitudinal section of human bone. The figure clearly shows the high-contrast gap regions and the low-contrast overlap regions, indicating most intra-fibril minerals are within the gap regions. (Arrows point to possible boundaries of constituent crystals. Figure reprinted from (McNally, Schwarcz et al. 2012))

Experimental studies have shown that, at the matrix level, oim bones have higher mineral density with smaller apatite crystals closely packed together (Camacho, Landis et al. 1996; Fratzl, Paris et al. 1996). Since the mechanical properties of mineralized collagen
fibrils are governed by the mineral density, distribution, shape and size, it is believed that the alterations on the structure of mineralized collagen fibril results from the mutation in the organic collagen molecule account for the brittle-type fractures of oim bones (Saban, Zussman et al. 1996; McBride, Choe et al. 1997; Misof, Landis et al. 1997; Camacho, Hou et al. 1999; Miller, Delos et al. 2007; Vanleene, Saldanha et al. 2011). However, how the replacement of the alpha-2 chain affects the structure of mineralized collagen fibril remains unclear.

In this chapter, we focus on the mechanical and structural differences between the mineralized normal and oim collagen microfibrils. In our full atomistic investigations of the normal and oim collagen microfibrils, we find that the oim microfibril is less dense compared with the normal microfibril as a result of local kink formations (Chapter 6). We construct the full atomistic mineralized collagen microfibrils for both normal and oim collagen by depositing minerals based on the assembly of collagen molecules discussed in Chapter 6 to investigate how the alterations on the assembly affect the mineral distributions, sizes and mechanical properties of mineralized collagen fibrils.

7.2 Materials and methods

7.2.1 Model construction

We construct mineralized collagen microfibrils for normal and oim collagen based on the equilibrium structures obtained from full atomistic simulations (see Chapter 6 for more detail). As we discussed in Chapter 6, the sequences of the collagen molecules define the structures of collagen molecules and thus affect how the collagen molecules pack into a fibril. We find that the oim collagen microfibril has larger lateral spacing due to the local kink formations and thus has different arrangements when packing into fibrils (Figure 7.2 (a)). It has been proposed that the space between collagen molecules is the place where the mineralization occurs. Therefore, we deposit the minerals based on the space within collagen microfibrils obtained in Chapter 6. Hydroxyapatite (Ca_{10}(PO_4)_6(OH)_2) crystal (Figure 7.2 (b)) which has a unit cell with dimensions 0.94240 x 0.81614 x 0.68790 nm^3 is used to construct the mineralized collagen microfibrils.
We use the following procedures to construct mineralized collagen microfibril models for both normal and oim collagen microfibrils:

I. Create a super cell of collagen microfibril from the equilibrium structure (as shown in Figure 7.3 (a))

II. Create a mineral plate (hydroxyapatite) with a dimension larger than the size of the collagen microfibril unit cell. The c-axis of the mineral are placed to align with the collagen axis.

III. Remove the hydroxyapatite crystals within N Å of collagen molecules. Here by changing the value of N, we can create mineralized collagen microfibrils with various mineral densities.

IV. Considering the periodicity of the model, remove the hydroxyapatite crystals which overlap across the boundaries.

Figure 7.2 Structure of normal and oim collagen microfibrils and the hydroxyapatite crystal. (a) The normal and oim collagen microfibrils obtained from full atomistic simulations (Chapter 6). As shown in the figure, the space between collagen molecules in oim collagen microfibril is different from the normal collagen microfibril due to the alterations on the assembly of collagen molecules. (b) hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$) unit cell.

Figure 7.3 (a) shows an illustration of the procedure of creating mineralized collagen microfibrils, and Figure 7.3 (b) shows examples on the mineral distributions obtained based on the packing of collagen molecules. This method has been used to construct
mineralized human collagen microfibril. It has been shown that the mineral distribution obtained from the method is in good agreement with experimental data (Nair, Gautieri et al. 2013).

7.2.2 Mineral density in mineralized collagen fibrils

At the tissue (bone) level, the bone consists of minerals external to the collagen fibrils and minerals within the collagen fibrils (McNally, Schwarcz et al. 2012). Let $W$ denote the total weight of a bone tissue, we have:

\[ W = W_{\text{collagen}} + W_{\text{hap\_int}} + W_{\text{hap\_ext}} \]  

(7.1)

where $W_{\text{collagen}}$ is the weight of collagen molecules, $W_{\text{hap\_int}}$ is the weight of the minerals within collagen fibrils and $W_{\text{hap\_ext}}$ is the weight of the minerals external to the collagen fibrils. Let $W_{\text{mineralized\_coll}}$ denote the total weight of mineralized collagen fibril and $W_{\text{mineral}}$ denote the total weight of the minerals of bone, we have

\[ W_{\text{mineralized\_coll}} = W_{\text{collagen}} + W_{\text{hap\_int}} \]  

(7.2)

\[ W_{\text{mineral}} = W_{\text{hap\_int}} + W_{\text{hap\_ext}} \]  

(7.3)

Using equations (7.1), (7.2) and (7.3), we can rewrite the mineral density of bone ($\rho_{\text{bone}}$) and the mineral density of a mineralized collagen fibril ($\rho_{\text{mineral\_coll}}$) as:

\[ \rho_{\text{bone}} = \frac{W_{\text{mineral}}}{W} = \frac{W_{\text{hap\_int}} + W_{\text{hap\_ext}}}{W} \]  

(7.4)

\[ \rho_{\text{mineral\_coll}} = \frac{W_{\text{hap\_int}}}{W_{\text{mineralized\_coll}}} = \frac{W_{\text{hap\_int}}}{W_{\text{collagen}} + W_{\text{hap\_int}}} \]  

(7.5)
Experimental studies have revealed that about 70% of the minerals are external minerals in bone (McNally, Schwarcz et al. 2012), we have the following relation:

\[
\frac{W_{\text{hap, int}}}{W_{\text{hap, ext}}} = \frac{0.3}{0.7} \tag{7.6}
\]

By using equation (7.6), we obtain a relation between the mineral density of bone \( (\rho_{\text{bone}}) \) and the mineral density of a mineralized collagen fibril \( (\rho_{\text{mineral, coll}}) \):

\[
\rho_{\text{mineral, coll}} = \frac{0.3 \rho_{\text{bone}}}{1 - 0.7 \rho_{\text{bone}}} \tag{7.7}
\]

Figure 7.4 shows the results of equation (7.7). The equation suggests that for a bone with 40% ~ 60% mineral density, the mineralized collagen fibril has a mineral density of 20% ~ 35%.

Figure 7.3 Constructing mineralized collagen microfibril models based on the equilibrium structures obtained from full atomistic simulations. (a) Illustration of a super cell of collagen microfibril and a mineral plate (hydroxyapatite). The c-axis of the mineral is placed to align with the collagen axis. (b) Two examples on the mineral
distributions obtained based on the packing of collagen molecules with different values of $N$.

![Graph showing the relation between mineral density of bone ($\rho_{bone}$) and mineral density of a mineralized collagen fibril ($\rho_{mineral_coll}$) under the assumption that 70% of minerals are external to collagen fibrils.]

Figure 7.4 Relation between the mineral density of bone ($\rho_{bone}$) and the mineral density of a mineralized collagen fibril ($\rho_{mineral_coll}$), under the assumption that 70% of the minerals in bone are external to collagen fibrils.

### 7.2.3 Full atomistic simulations

Full atomistic simulations are performed using MD simulation software, LAMMPS (Plimpton 1995), ([http://lammps.sandia.gov/](http://lammps.sandia.gov/)). The CHARMM force field (MacKerell, Bashford et al. 1998) that includes parameters for hydroxyproline amino acids based on a model put forth by Anderson (Anderson 2005) is used to model the collagen molecule. Force fields for biomaterials such as CHARMM do not include parameters for minerals. Therefore, we use extended CHARMM force field with non-bonded, bond, angle and dihedral parameters for hydroxyapatite and the collagen-hydroxyapatite interactions reported earlier (Hauptmann, Dufner et al. 2003; Bhowmik, Katti et al. 2007; Nair, Gautieri et al. 2013).
An energy minimization using a conjugate gradient scheme is performed before molecular dynamics simulations. Rigid bonds are used to constrain covalent bond lengths and an integration time step of 1 fs is used. Nonbonding interactions are computed using a cut-off for neighbor list at 13.5 Å, with a switching function between 10 and 12 Å for van der Waals interactions. The electrostatic interactions are modeled by the ewald/n style, that performs standard coulombic Ewald summations in a more efficient manner (in't Veld, Ismail et al. 2007). After energy minimization, we perform MD simulations for 5 ns at a constant temperature of 310 K to obtain an initial equilibrium structure.

In order to assess the mechanical properties of the mineralized atomistic microfibril models, we perform $NPT$ molecular dynamics simulations with different constant mechanical stress levels (in the range from 0 to 150 MPa) in tension (along the fibril $c$-axis) for models with different mineral density. The pressures on the other axes are kept constant at 1.013 bar pressure. In each simulation, we keep monitoring the convergence of the lattice parameters. Every simulation is performed until the lattice parameters on all directions are converged. After reaching the equilibrium, we use the last 1 ns of molecular dynamics simulations for computing the unit cell lattice parameters corresponding to each load applied.

7.2.4 Analysis of the mineral distribution and size

We develop a method to analyze the mineral distribution and size. For analyzing the mineral distribution, we divide the mineralized collagen microfibrils into finite cross-sections along the $c$-axis of the collagen microfibril model. By counting the number of calcium ions in each cross-section, we can obtain the mineral distribution along the $c$-axis of collagen microfibril.

It is nontrivial to analyze the continuity and the size of each mineral crystal in a full atomistic mineralized collagen microfibril because the atoms of hydroxyapatites are actually ‘discrete’ and mineral crystals could aggregate to form a larger mineral crystal during a full atomistic simulation. In order to analyze and compare the mineral size of the normal and oim mineralized collagen microfibrils, we develop an algorithm to find out all
the hydroxyapatite atoms within the same mineral crystal for a given atom \(i\). The procedure is as following:

For each atom \(i\) of hydroxyapatite:

Step 1: Find all the hydroxyapatite atoms within \(x\) Å of atom \(i\), denoted by \(S\)

Step 2: Find all the hydroxyapatite atoms within \(x\) Å of \(S\), denoted by \(T\)

Step 3: If \(T\) is empty then go to step 4

else

set \(S = S \cup T\)

repeat step 2

endif

Step 4: \(S\) is the group of all hydroxyapatite atoms connected with atom \(i\), analyze the shape and size

Here \(x\) is a threshold value on the order of atomic distance (\(-1\) Å) chosen to define the continuity of the mineral. The above procedure allows us to analyze the continuity of the hydroxyapatite crystals and the size of the largest mineral within the mineralized collagen microfibril.

7.3 Results

7.3.1 Mineral distributions in the normal and \(oim\) collagen microfibrils

Based on the equilibrium normal and \(oim\) collagen microfibril structures, we firstly use the same criterion (\(N=0.74\)) to construct the normal and \(oim\) mineralized collagen microfibrils. Note that \(N=0.74\) is chosen for constructing a mineralized normal collagen microfibril with a mineral density of \(-25\%\), which is corresponding to a bone with a mineral density of \(-52\%\) (computed from equation (7.7)). Figure 7.5 shows the results of the mineralized normal and \(oim\) collagen microfibrils. We find that by using the same mineralization criterion, there are much more spaces between collagen molecules for mineral deposition in the \(oim\) collagen microfibril. As \(N=0.74\), the mineralized normal collagen microfibrils has a mineral density of 25\% while the mineralized \(oim\) collagen microfibril has a much higher mineral density (\(-35\%\)). This is because \(oim\) collagen
microfibril has larger lateral spacing between collagen molecules as we observed in the full atomistic simulations of non-mineralized collagen microfibrils (see chapter 6 for more detail). This finding is consistent with previous experimental study (Vanleene, Porter et al. 2012; Carriero, Zimmermann et al. 2014).

These results suggest that mutations at the molecular level cause the alterations throughout the multiple hierarchical levels of bone. For the case of oim, the replacement of the alpha-2 chains with an alpha-1 chain at the molecular level results in a collagen molecule with more local kink regions. These changes on the structures at the molecular level alters the packing of the collagen molecules at the fibril level (as discussed in Chapter 6) and the alterations on the packing further modify the space for the deposition of minerals at the mineralized collagen microfibril level.

\[ \text{Normal (~25\%)} \]

\[ \text{oim (~35\%)} \]

Figure 7.5 Mineralized normal and oim collagen microfibrils constructed by using the same criterion (N=0.74). By using the same criterion to deposit minerals, we obtain a mineralized collagen microfibril with much higher mineral density for the case of oim collagen.

We further construct two more mineralized collagen microfibrils to investigate the differences between mineralized normal and oim collagen microfibrils at the same mineral density. One is the mineralized normal microfibril with a mineral density of ~35\% and the other is the mineralized oim microfibril with a mineral density of ~25\%.

We find that in order to construct a mineralized collagen microfibril with higher mineral
density (~35%), a much smaller value of N (0.66) is required for the normal collagen. On the other hand, a larger value of N (0.87) results in a mineralized oim collagen microfibril with a mineral density of ~25%. These suggest that at the same mineral density, the mineralized normal collagen microfibril has more closed interactions at the interface between the minerals and the collagen molecules.

![Diagram](image.png)

Figure 7.6 Equilibrium structures of mineralized normal and oim collagen microfibrils at mineral densities of 25% and 35%. The blue regions indicate the minerals.

Figure 7.6 shows the equilibrium structures of mineralized normal and oim collagen microfibrils at mineral densities of 25% and 35%. The blue region indicates the minerals within collagen microfibrils. As shown in the figure, most of the minerals are within the gap region of the collagen microfibrils for both normal and oim collagen. This is in good agreement with the findings in experiments (Weiner and Wagner 1998; Jager and Fratzl 2000; Gupta 2006; Nudelman 2010; Alexander 2012).

We further analyze the mineral distributions along the collagen axis as shown in Figure 7.7 and Figure 7.8. Figure 7.7 shows the mineral distributions for the mineralized normal collagen microfibrils with mineral densities of 25% and 35%. The y-axis represents the normalized mineral density. The mineral density is normalized by the cross-section area of the unit cell of a hydroxyapatite crystal. The unit is set to be the cross-section area of
the unit cell of a hydroxyapatite crystal ($0.94240 \times 0.81614$ nm$^2$). The results confirm that most of the minerals are within the gap region for both mineral densities. For the collagen microfibril with lower mineral density (25%), all the minerals are within the gap region, while for the higher mineral density (35%), small minerals are found in the overlap region. The minerals have larger cross-section area in the case of higher mineral density, as shown by the larger normalized mineral density in the figures. It is worth noting that the results show that the highest mineral densities for both models are found at the gap region close to the C-telopeptide domain. This is in accordance with the experiment finding that the mineral nucleation point is close to the carboxy terminus (in the gap region, immediately after the gap/overlap transition) (Nudelman 2010).

![Figure 7.7 Normalized mineral density distribution along the collagen axis for mineralized normal collagen microfibrils with 25% and 35% mineral densities.](image)

Figure 7.7 Normalized mineral density distribution along the collagen axis for mineralized normal collagen microfibrils with 25% and 35% mineral densities.

Figure 7.8 shows the mineral distributions for the mineralized $oim$ collagen microfibrils with mineral densities of 25% and 35%. Similar to the mineralized normal collagen microfibrils, most of the minerals are within the gap region of the collagen microfibrils. The minerals have larger cross-section areas in the case of higher mineral density (35%), which is the same as the results of the mineralized normal collagen microfibrils. However, the results reveal that different to the mineralized normal collagen microfibril, small minerals are found in the overlap region in the mineralized $oim$ collagen microfibril even at lower mineral density (25%). This suggests that the $oim$ collagen microfibril has a
less dense packing even in the overlap region compared with the normal collagen microfibril.

Figure 7.8 Normalized mineral density distribution along the collagen axis for mineralized oim collagen microfibrils with 25% and 35% mineral densities.

7.3.2 Analysis of the mineral sizes

The mechanical properties of the mineralized collagen microfibrils not only depend on the mineral density of the fibrils but also depend on the mineral distributions and mineral sizes. Experimental studies have shown that the oim bone has higher mineral density but smaller mineral crystals (Camacho, Landis et al. 1996; Fratzl, Paris et al. 1996). By constructing mineralized collagen microfibrils based on the space between collagen molecules from full atomistic simulations, we have shown that the higher mineral density of the oim can be explained by the larger lateral spacing between collagen molecules, i.e. there are more spaces for mineral deposition. In this section, we use the algorithm described in section 7.2.4 to investigate the mineral sizes of both mineralized normal and oim collagen microfibrils at 25% and 35% mineral densities.

Figure 7.9 shows the results of the analysis of the mineral size for both mineralized normal and oim collagen microfibrils at 25% mineral density. The shaded area represents the mineral distribution of the largest mineral crystal found in the microfibril. The inset shows the snapshot of the largest mineral crystal, and the grid serves as a scale bar which
has a size of 1 x 1 nm². The results show that at 25% mineral density, the largest mineral crystals in both mineralized normal and oim collagen microfibrils have similar dimensions (~12-13 nm long). It is worth noting that the largest mineral crystal of the mineralized normal collagen microfibril is found to be at the region close to the C-telopeptide domain of the full length type I collagen while the largest mineral crystal of the mineralized oim collagen microfibril is found at the center of the gap regions. This suggests that the replacement of the alpha-2 chain in the oim collagen microfibril affects the space within fibrils for mineral deposition in the gap regions due to the disruptions on the packing of the collagen molecules.

![Normal 25%](image1)

![oim 25%](image2)

Figure 7.9 Analysis of the largest mineral in mineralized normal and oim collagen microfibril at 25% mineral density.

We further analyze the largest mineral crystal size of both mineralized normal and oim collagen microfibrils with 35% mineral density as shown in Figure 7.10. Interestingly, our results reveal that at mineral density of 35%, which corresponds to a bone with up to 60% mineral (equation (7.7)), the largest mineral crystal in the mineralized normal collagen microfibril has a length of ~24 nm, while the largest mineral crystal in the mineralized oim collagen microfibril has a shorter length of ~19 nm. These results
suggest that although the oim collagen microfibril has more space for mineral deposition (due to the increased lateral spacing between collagen molecules), the kinks of the oim collagen molecules are likely disrupting the space for mineral deposition. That is, the normal collagen microfibril has less space for mineral deposition but the space is more continuous. Our findings on the differences of the mineral size for mineralized normal and oim collagen microfibrils are in good agreement with experiments (Vanleene, Porter et al. 2012).

Figure 7.10 Analysis of the largest mineral in mineralized normal and oim collagen microfibril at 35% mineral density

7.3.3 Mechanical properties of mineralized collagen microfibrils

We perform mechanical tests of the mineralized normal and oim collagen microfibrils to investigate their mechanical properties and deformation mechanisms. Mechanical tests for mineralized normal collagen microfibril with 25% mineral density and oim collagen microfibril with 35% mineral density are performed to study the mechanical differences between the mineralized normal and oim collagen microfibril by using the same mineral deposition criterion. An additional mechanical test on the mineralized normal collagen microfibril with 35% mineral density is performed to understand how the mineral density
affects the mechanical properties and the differences between the mineralized normal and oim collagen microfibril at the same mineral density.

Figure 7.11 shows the stress-strain curves of the three mineralized normal and oim collagen microfibrils. We perform NPT simulations for each microfibril model under four stress levels (0, 30, 75 and 150 MPa) as shown in the figure. The solid lines in the figure represent the stress-strain curves obtained from curve fitting. We find a stress-strain relation of $\sigma = 2.6 + 1044.6\varepsilon + 16700.5\varepsilon^2$ for the mineralized normal collagen microfibril with 25% mineral density, a stress-strain relation of $\sigma = -1.2 + 663.9\varepsilon + 19102.9\varepsilon^2$ for the mineralized oim collagen microfibril with 35% mineral density, and a stress-strain relation of $\sigma = 2.3 + 2794.5\varepsilon + 2622.3\varepsilon^2$ for the mineralized normal collagen microfibril with 35% mineral density. These results reveal the modulus-strain relations as $E_{\text{normal},25\%} = 1044.6 + 33400\varepsilon$, and $E_{\text{oim},35\%} = 663.9 + 38206\varepsilon$, and $E_{\text{normal},35\%} = 2794.5 + 5244.6\varepsilon$.

As shown in Figure 7.11, when the mineral density of the mineralized normal collagen microfibril increases from 25% to 35%, the modulus increases. At low strain ($\varepsilon = 0.01$), the modulus of the mineralized collagen microfibril increases from 1.38 GPa to 2.8 GPa when the mineral density increases from 25% to 35%. It is worth noting that at lower mineral density (25%), the stress-strain exhibits a non-linear relationship while as the mineral density increases to a higher mineral density (35%), the mineralized collagen microfibril becomes tougher and exhibits a linear stress-strain curve. These results are consistent with previous computational studies on the mineralized human collagen microfibrils (Nair, Gautieri et al. 2013).

We find that the mineralized oim collagen microfibril has reduced modulus compared with the mineralized normal collagen microfibril. As shown in Figure 7.11, the 35% mineralized oim collagen microfibril has a modulus of ~1.05GPa which is much smaller than the modulus of the mineralized normal collagen microfibril with the same mineral density (35%). More importantly, we find that the modulus of the 35% mineralized oim collagen microfibril is even smaller than the modulus of mineralized normal collagen microfibril with much lower mineral density (25%). These indicate that the replacement of the alpha-2 chain at the molecular level leads to an altered packing of collagen.
molecules. These alterations on the packing of the collagen molecules affect the spaces for mineral deposition at the mineralized collagen fibril level. Although the oim collagen microfibril forms a structure with more spaces for mineralization, the disruption on the packing of the molecules makes it more difficult to reinforce its mechanical property through the mineral deposition process.

Figure 7.11 Stress-strain curves of mineralized collagen microfibrils. The results show that the mineralized oim collagen microfibril has reduced modulus.

7.3.4 Analysis of the deformation mechanisms

To understand the deformation mechanisms of the mineralized collagen microfibrils, we analyse the gap/overlap ratio under different stress level (Figure 7.12). As shown in Figure 7.12, for the non-mineralized collagen microfibril, there are two regimes of the curve. In the low stress regime, the gap/overlap ratio increases linearly as the stress increases with a larger slope, while in the high stress regime, the gap/overlap ratio increases linearly as the stress with a distinct smaller slope (see Chapter 6 for more detail on the property of non-mineralized collagen microfibril.). For the mineralized normal
collagen microfibril with 25% mineral density, the result shows that the gap/overlap remains almost constant as the stress increases. This is because the minerals, mostly within the gap region, reinforce the modulus of the gap region, making it as stiff as the overlap region. Interestingly, as the mineral density increases to 35%, an inverse trend is found. For the 35% mineralized normal collagen microfibril, we find that the gap/overlap ratio slightly decreases as the stress increases. This indicates the overlap region is softer compared with the gap region and therefore stretches more as stress increases. This suggests that the higher mineral density, i.e. the more mineral within the gap region, makes the gap region stiffer. These findings are also consistent with previous computational studies on the mineralized human collagen microfibrils (Nair, Gautieri et al. 2013).

For the mineralized oim collagen microfibril with 35% mineral density, the result shows that the gap/overlap ratio firstly increases as the applied stress increases from 0 to 30 MPa and then decreases when the stress increases further. The results suggest that at the equilibrium (zero tensile stress), the gap region is softer even with 35% mineral density. One reason is that the gap region of the oim collagen microfibril is much softer compared with the normal collagen microfibril due to the kink formations in the gap region (see Chapter 6 for more detail on the structural differences between normal and oim collagen microfibrils). It is also likely that because the mineral crystals are smaller and discrete in the mineralized oim collagen microfibril, the modulus of the gap region does not increase as much as the mineralized normal collagen microfibril. As the applied stress larger than 30 MPa, because the kink regions are stretched, the gap region of the mineralized oim collagen microfibril becomes stiffer and thus has similar trend as the mineralized normal collagen microfibril.
Figure 7.12 Analysis of the gap/overlap ratio under different tensile stress levels for non-mineralized and mineralized collagen microfibrils.

### 7.4 Summary

In this chapter, we construct mineralized collagen microfibrils for normal and oim collagen based on the equilibrium structures obtained from full atomistic simulations. The minerals are deposited based on the space within collagen molecules defined by the assembly of collagen molecules. Two mineral densities (25% and 35%) for both normal and oim mineralized collagen microfibrils are constructed. By analyzing the space within collagen molecules, we find that there is more space for mineral deposition for the oim collagen fibril. This finding is consistent with experiments and provides the molecular insight into how the replacement of alpha-2 chain at molecular level affects the structure of mineralized collagen fibrils.
By analyzing the mineral distributions and mineral sizes of the normal and oim mineralized collagen microfibrils, we find that the normal collagen microfibril has less space for mineral deposition, but the space is more continuous. On the other hand, although the oim collagen microfibril has more space for mineral deposition, the space is disrupted by collagen molecules and thus less continuous. This finding could explain why the experiments have revealed the oim collagen has higher mineral density, with smaller apatite crystals closely packed together (Camacho, Landis et al. 1996; Fratzl, Paris et al. 1996).

The mechanical tests on the mineralized normal and oim collagen microfibrils reveal that the oim mineralized collagen has reduced modulus even with higher mineral density. This suggests that the mineral distribution and mineral size also play important roles in the mechanical properties of mineralized collagen fibrils. Our results suggest that the reduction on the modulus of oim mineralized collagen microfibril is resulted from at least: (1) the less dense packing of collagen molecules; (2) the disrupted mineral distribution which leads to (3) the smaller mineral size.

In summary, we have developed the first full atomistic structures of the mineralized normal and oim collagen microfibrils. Because the mineralized collagen fibrils are highly conserved across species and are the fundamental building block of bone, our model provide important molecular information which help us understand the loss of mechanical integrity at larger length scales. Our model also provides a possible connection between the molecular content and the brittle bones in the oim model of osteogenesis imperfecta.
8 Conclusions and opportunities for future research

8.1 Summary of key findings and significances

In this thesis, we use the oim model to develop the first multiscale framework to study the hierarchical structure of bone from a bottom-up molecular approach. Our work is the first to compare the full atomistic structures between the normal (heterotrimer) and oim (homotrimer) collagen from a single molecular level to collagen microfibril level and to mineralized collagen microfibril level. Figure 8.1 illustrates the multiscale framework and the major findings of this thesis. Investigations of collagenous tissues from atomistic scale provide fundamental insights into the molecular origins of many diseases such as the brittle bone and enzyme resistance.

By modeling single collagen molecules, we show that full atomistic simulation can predict the structures of real sequences type I collagen molecules. We find that the replacement of alpha-2 chains alters the molecular structure of collagen molecules. The homotrimer forms larger kinks, which further affect the properties in the next hierarchical levels. Full atomistic simulations also reveal that the homotrimer has higher thermal stability at the cleavage site. Our study suggests a molecular mechanism by which thermal stability plays a role in the cleavage mechanism and the potential explanation of the force-stabilization results found in experimental studies as well as the distinct behaviours of the heterotrimer and homotrimer collagen. In addition to providing a possible explanation for the seemingly conflicting experimental results reported in earlier papers based on a mechanistic model, we anticipate that our study is crucial for developing new biomaterials that serve as platforms for treatments for a variety of diseases. For example, through tuning the mechanical force, one can precisely regulate collagen degradation. This concept can also find application in tunable collagen-based biomaterials, and may be important to understand the development of tissue.
Figure 8.1 Summary of the multiscale framework developed in this thesis and the major findings.

On the other hand, multiscale modeling allows us to examine the structure and mechanics of collagen, helping to define the basic structural and mechanical differences responsible for the changes throughout the hierarchy:

I. At the molecular level, we find that the \( oim \) collagen molecule forms kinks which alters the assembly of collagen molecules at the fibril level.

II. At the microfibril level, our models provide direct evidence that the \( oim \) microfibrils have larger lateral spacing due to the alteration on the single molecular structure. The less dense packing of the collagen molecules results in the reduced mechanical properties of \( oim \) collagen fibrils.

III. At the mineralized microfibril level, we find that the alteration on the assembly of collagen molecules indeed alters the space for mineral deposition. The \( oim \) collagen microfibril has more space for mineral deposition while the space is less continuous and therefore the \( oim \) mineralized collagen microfibril has smaller mineral size. We find that the mineralized \( oim \) collagen microfibril has reduced
modulus due to the less dense assembly of oim collagen molecules and the alterations on the mineral distributions and sizes caused by the mutations at the molecular level.

Our results help in defining the structural and mechanical differences between the heterotrimer and homotrimer collagen responsible for each hierarchy, which underlie the loss of mechanical integrity at larger length scales. Our model also provides a possible connection between the molecular content and the brittle bones in the oim model of osteogenesis imperfecta. This work provides important insights into mechanical alterations at the fibril level due to pathologies such as OI, in this case allowing insights into the mechanisms behind altered macromechanical behaviour of oim tissue from a full atomistic investigation. Given the extreme importance of collagen as the major construction material in biology—and the associated use of collagen in many fields ranging from regenerative medicine, pathology to bioengineering—the impact of our work is highly significant and illustrates the critical role of the alpha-2 chain.

8.2 Opportunities for future research

The multiscale framework developed in this thesis provides a useful tool to study many aspects of collagenous tissue under various physiological and diseased conditions. Possible opportunities for future research are discussed in this section.

One opportunity for future research is to extend our model to study the osteogenesis imperfecta (OI) in human. The OI mutation is caused primarily by a replacement of the Gly residue in the repeating (Gly-X-Y)₈ triplets. As of now, several mutation locations and types along the entire collagen molecule have been identified and classified into severe, moderate and mild disease conditions (Figure 2.7) (Beck, Chan et al. 2000). This multiscale framework can be used to study the effects of various single point mutations, which would help understanding the effect of each mutation and developing treatments for severe, moderate and mild OI.

Another opportunity is to use this model to investigate the effect on the modification of the mineral in bone. It is known that the inorganic mineral phase in bone is not simply
pure hydroxyapatite crystals. However, it remains unclear how the modification of the mineral phase might alter the mechanical property of bone. Our model can be used to investigate how the modification on the mineral compositions affects the mechanical property by modifying the mineral structures deposited in our full atomistic models. For example, future work can simulate the carbonated calcium-deficient hydroxyapatite and investigate its effect on the mechanical property of mineralized collagen fibrils.

More importantly, our models provide molecular details on the assembly of collagen molecules and the mineral deposition. These molecular understandings can be feed into a coarse-grained modelling to simulate real size collagen fibrils. The coarse-grained modelling would allow us to study the failure of single collagen fibril. Furthermore, the results of the coarse-grained modelling may be further feed into a continuum modelling to complete the multiscale framework for modelling bone from microscale molecular level to macroscale tissue level.
References


148


152


experimental observations with mathematical models." Proteins-Structure Function and Bioinformatics 70(4): 1154-1161.


