ABSTRACT

Title of Document: TYPE I COLLAGEN HOMOTRIMERS; THEIR ROLE IN COLLAGEN FIBRIL FORMATION AND TISSUE REMODELING

Sejin Han, Doctor of Philosophy, 2009

Directed By: Professor Wolfgang Losert
Department of Physics

Formation and remodeling of type I collagen fibril networks are paradigms of biopolymer self-assembly, yet many of their aspects remain poorly understood. Type I collagen is the most abundant vertebrate protein which self assembles into fibrils and hierarchical fibril network structures, forming scaffolds of bone, skin, tendons and other tissues. The normal isoform of type I collagen is a heterotrimer of two α1(I) and one α2(I) chains, but homotrimers of three α1(I) chains have been reported, e.g., in cancer and fibrosis. Despite their importance in various disorders, very little is known about potential effects of the type I collagen homotrimers on self-assembly, physical properties, and remodeling of collagen fibrils and fibril networks. Thus, we selected characterization of these effects and understanding the underlying physical mechanisms as the topic of the present thesis. Some of our most important findings were: (i) different nucleation mechanism and morphology in homotrimer fibrils compared to the normal heterotrimers fibrils; (ii) segregation of the homo- and heterotrimers within fibrils; (iii) increased bending rigidity of homotrimer fibrils; and (iv) homotrimer resistance to cleavage by enzymes responsible for fibril degradation and remodeling due to increased triple helix stability at the cleavage site. The corresponding *in vitro* experiments and theoretical analysis of the results suggested drastically different physics of the fibril networks composed of the homo/heterotrimer mixtures and pointed to a potential role of these physics in various disorders, e.g., in cancer and fibrosis pathology.

TYPE I COLLAGEN HOMOTRIMERS; THEIR ROLE IN COLLAGEN FIBRIL FORMATION AND TISSUE REMODELING

By

Sejin Han

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:
Associate Professor Wolfgang Losert , Chair / Advisor
Special member Dr. Sergey Leikin
Assistant Professor Arpita Upadhyaya
Assistant Professor Arthur La Porta
Professor Sergei Sukharev / Dean's representative

Acknowledgements

First, I would like to thank Our Lord Jesus Christ who led and blessed me and my study.

I am indebted to Dr. Sergey Leikin my supervisor at the National Institute of Health in many ways. He provided me with the set of problems addressed in this dissertation and also with relentless support over the last five years. He was always encouraging and supportive throughout while allowing me a great amount of independence and creativity. I am deeply grateful to Dr. Leikin for his intellectual insights, unending encouragements and constructive advices, without which I do not believe I was able to finish this study.

I am thankful to Professor Wolfgang Losert, my academic advisor at the University of Maryland, for his support, mentoring, and continued advice and encouragement over the course of my education.

All of my group members in the Section on Physical Biochemistry at NICHD have always been to me wonderful colleagues, collaborators and friends. In retrospective, I find public discussions and private conversations with them invaluable and I really appreciate all their direct and indirect contributions to my research, even including their most severe criticisms.

I am also grateful to my friends and colleagues at University of Maryland who have supported me in many ways during these years, and especially I want to thank KBS (Korean Bible Study) members for their love and friendship. Their prayer provided me with a great spiritual power upon which I could regroup myself even when I was most tired and discouraged.

Most of all, I would like to thank my husband, my parents and other family members for their constant support, love, and encouragement.

During the course of this work, I was supported by an intramural fellowship (a pre-doctoral visiting fellow) from NICHD. I am grateful to this support as well.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The strength of bone and cartilage, and the elastic properties of tendon, ligament, and skin all arise from the physical properties of collagen scaffolds [1]. The most important building blocks in these scaffolds are type I collagen molecules, which self-assemble into complex hierarchical structures. Each collagen molecule consists of three polypeptide chains woven together into a 300 nm long triple helix. In normal adult tissues, the type I collagen molecule is a heterotrimer of two α1(I) chains and one α2(I) chain [2]. However, type I collagen also has a second isoform, which is a homotrimer of three α1(I) chains. This homotrimeric isoform has been found in fetal tissues [3] as well as in cancer [4-9], fibrosis [10], and other diseases [11, 12]; but its role in development and pathology has not been understood. In this thesis, we investigated how α1(I) homotrimers affect the self-assembly and physical properties of type I collagen fiber networks *in vitro*.

Collagen fiber networks are in many respects similar to polymer networks extensively studied in soft matter physics [13, 14]. Based on the accumulated knowledge of the physical properties critical for the formation and function of such networks, we focused on the effects of α1(I) homotrimers on: (i) collagen fibril self-assembly, (ii) network formation and its connectivity, (iii) mechanical properties of individual fibrils, and (iv) remodeling of individual fibrils and fibril networks. In these studies, we had to account for distinct features of collagen, distinguishing it from typical synthetic polymers. As a result, this study was guided by the logic of

polymer network physics, but it also often relied on the methods more common to protein biochemistry.

Since the methods utilized for understanding different aspects of the collagen fiber network physics are fairly distinct, we introduce them in more detail in the corresponding chapters. The content of each chapter is based on a separate article; resulting in some recurrence and overlap in the text of this thesis. Below, we briefly summarize the ideas behind each step of the study.

In Chapter 2 we review the literature on the structure and biosynthesis of heterotrimeric type I collagen as well as formation, morphology, mechanical properties, degradation, and remodeling of the corresponding collagen fibril networks. The chapter concludes with a brief review of much more limited reports on the properties of the homotrimeric type I collagen.

Chapter 3 describes our study of the effects of homotrimers on fibril network self-assembly in different mixtures with the heterotrimers.

Self-assembly is a ubiquitous phenomenon. Much of the information determining the structure and properties of resulting systems is contained within this process. The physical principles and practical applications of self-assembly have been extensively studied in different areas of physics, chemistry, material and tissue engineering, biology, and pharmaceutical technology [15-17].

One of the common features of these processes is S-shape kinetics, e.g., observed in the growth of salt crystals, formation of biopolymer networks, and

fibrillogenesis of heterotrimeric type I collagen. This kinetic pattern is characterized by an initial delay (lag phase) followed by rapid growth and saturation. The lag-phase is typically associated with nucleation, whereby a critical number of molecules must interact stochastically. Once the critical molecular number is reached, the accretion of additional molecules becomes more favorable than the removal of molecules, resulting in a rapid growth.

An abnormal kinetics without the lag-phase was reported for *in vitro* self-assembly of type I collagen homotrimers [18], pointing to an altered nucleation mechanism that might affect fibril network architecture and individual fibril properties. Therefore, we investigated the physical principles underlying the lag-less self-assembly of the homotrimers and their implications for fibril formation in homo/heterotrimer mixtures. Our observations suggested that the apparent lack of the lag phase was caused by rapid formation of nonspecific homotrimer aggregates, which was detected as an assembly process in turbidity measurements. Confocal microscopy imaging revealed that these aggregates served as centers from which the homotrimer fibrils grew in all directions. We observed no nonspecific aggregation of the heterotrimers.

Different fibril nucleation mechanisms resulted in kinetic segregation of the homo- and heterotrimers within fibrils in mixtures. However, a more detailed analysis also revealed equilibrium sub-fibrillar segregation of the molecules. Such segregation is often observed in binary mixtures of different polymers[19, 20]. In most cases, it results from major differences in the size, shape, or surface chemistry of the molecules [21]. In our case of closely similar molecules, the segregation appeared to

be associated with more subtle differences in the homotrimer-homotrimer, homotrimer-heterotrimer, and heterotrimer-heterotrimer interactions.

Chapter 4 describes a subsequent study in which we focused on the effects of homotrimers on mechanical properties of individual fibrils. Previous studies found that type I collagen homotrimer fibers were more loosely packed, more hydrated [22], and had lower tensile strength [23]. Based on these observations, we expected the homotrimer fibril network to be more flexible. However, confocal imaging revealed spear-like homotrimer fibrils, which appeared to be more rigid than thread-like heterotrimer fibrils, as described in Chapter III. Furthermore, the homotrimers assembled into mostly isolated clusters with multiple spear-like fibrils protruding from a common center, while heterotrimer fibrils formed an entangled meshwork. This difference in the network architecture also appeared to be at least partially related to higher bending rigidity of individual homotrimer fibrils.

To characterize the resistance of individual fibrils to bending, we measured their persistence length. The persistence length is commonly used as a measure of the bending rigidity in long, semi-flexible polymers in solution [21]. However, the relationship between the persistence length, bending rigidity, and shape of collagen fibrils might be less straightforward due to the more complex, hierarchical structure of the fibrils. To distinguish different modes of bending associated with thermal motions and static defects in fibrils, we measured different types of correlations in the fibril shape and determined persistence lengths corresponding to these correlations. Normal mode analysis suggested that the observed bending fluctuations were mostly

thermal. At the same time, analysis of the curvature distribution indicated that the apparent bending rigidity of the fibrils might be different at different length scales. **In Chapter 5, we describe a study of the homotrimer effect on collagen fiber network degradation and remodeling.** The molecular mechanism of degradation and remodeling clearly distinguishes collagen fibers from most synthetic and many natural self-assembled systems. For instance, polymer networks usually self-assemble and disassemble through reversible association and dissociation in response to changes in temperature, pressure, concentration, solvent, etc. [24, 25]. In contrast, collagen fibers, which are stabilized by intermolecular covalent cross-links, are disassembled through the action of specialized enzymes (collagenases).

The homotrimers artificially refolded from isolated α(I) chains were earlier found to be resistant to the cleavage by fibroblast and neutrophil collagenases [26]. We established that naturally produced homotrimers were also resistant to these enzymes as well as to all other collagenases [27], but the molecular mechanism of this resistance remained unknown.

To clarify this mechanism, we analyzed different models for collagen cleavage by collagenases and experimentally investigated the interaction of the most common human collagenase (MMP-1) with homo- and heterotrimeric type I collagen. By combining the theoretical analysis with experiments, we were able to decouple different steps involved in the cleavage process and demonstrate that the enzyme binding to the homo- and heterotrimers as well as the cleavage rates of the α1(I) and α2(I) chains were similar. The inefficient cleavage was caused by the homotrimer resistance to local opening (unwinding) of collagen triple helix at the cleavage site

required for placement of individual chains in the catalytic cleft of the enzyme. Confocal imaging of the degradation of fibrils prepared from different mixtures of the homo- and heterotrimers revealed that the combination of their very different susceptibility to collagenases and tendency to segregate fibrils drastically affected the remodeling of the mixed fibrils.

Chapter 6 concludes the thesis with the discussion of potential implications of the collagen fiber self-assembly physics uncovered in this work for cancer invasion, fibrosis, and other disorders. In particular, we propose how cancer cells may utilize the tendency of the homo- and heterotrimers to segregate, the higher rigidity of the homotrimeric type I collagen fibers, and the resistance of these fibrils to collagenases for promoting their own proliferation and migration while destroying surrounding tissues.

Chapter 2: Type I collagen (Literature review)

This literature review provides a general background for our study by describing current understanding of collagen biosynthesis, fibrillogenesis, interactions, and degradation with particular attention to type I collagen.

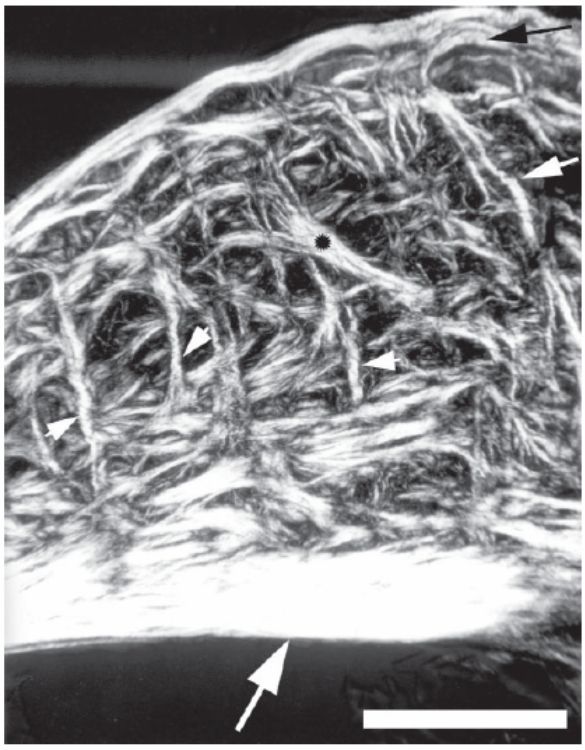

Fig 2.1. Polarizing microscopy image of collagen fibers in human jaw joint disk. Type I collagen is the predominant collagen by volume in this tissue. Also Type II collagen is present pericellularly and between Type I collagen lattice. Branched and joined fibers to each other have been believed to serve a stress distribution function. White scale bar: 1mm. (reproduced from Ref [28]).

Collagens are the most abundant proteins in the animal kingdom [29]. They are the major structural proteins of extracellular matrix (ECM), which provides scaffolding for formation and maintenance of multicellular systems [Fig. 2.1]. Their distinguishing feature is the presence of one or several triple helical domains formed by three polypeptide chains with the characteristic Gly-X-Y repeat in each chain, where Gly represents glycine and X and Y are other amino acids.

There are over 25 types of collagen. The first to be purified and characterized were vertebrate fibril forming collagens from connective tissues such as skin, bone, tendon, ligament, blood vessel walls and cornea [30]. Traditionally, distinct types of collagen molecules were numbered following the order in which they were discovered, using Roman numerals for each collagen type and Arabic numerals for individual α-chains.

Collagens can be crudely separated into three large functional groups [31]:

Fibril forming collagens (types I, II, III, V, XI) form fibrous scaffolds of connective tissues. They are characterized by a large uninterrupted triple helical domain of approximately 300 nm. They are synthesized as procollagens composed of three pro α-chains, which contain large, mostly globular N- and C- terminal propeptides cleaved by specialized extracellular enzymes prior to fibril assembly. After the cleavage, mature types I-III collagens are essentially triple helices with only short non-helical terminal peptides (telopeptides). In contrast, types V and XI collagens retain a large N-terminal domain, that is not cleaved prior to their

incorporation into fibrils and that may play an important role in regulating fibril diameter[32, 33].

Type I collagen is arguably the most important structural protein in vertebrates responsible for the mechanical strength of skin, tendon, bone, cornea, sclera, dentine, and other tissues. Type II collagen is the major structural protein of cartilage. Type III collagen coassembles into fibrils with type I collagen in many tissues, particularly blood vessel walls and skin. Types V and XI are "minor" collagens, found in small quantities within type I and II collagen fibrils. [34]

Basement membrane collagens (types IV, VII, XV, and XVIII) are found mostly within thin layers of specialized ECM supporting epithelial or endothelial cells, muscle fibers, and peripheral nerves. They have multiple interruptions in the Gly-X-Y sequence, serving as flexible hinges between more rigid triple helical domains. The more abundant type IV collagen forms polygonal 2D networks, which constitute an insoluble scaffold of basement membranes. It has large triple-helical domains with the total length of >350 nm (>160 kDa) separated by more than 20 short interruptions. Type VII collagen is one of the largest of the known mammalian collagens, which anchors the basement membrane to collagen fibers in the supporting connective tissue. Types XV and XVIII collagens may also serve as basement membrane anchors in some tissues.

Other collagens are expressed in multiple tissues and have a variety of different functions. Types VI, VIII, and X are short-chain collagens with continuous triple

helical domains. They show significant homology but different tissue distribution; type VI collagen is believed to be particularly important in muscles, type VIII in cornea, and type X in cartilage. Types XIII and XVII are transmembrane collagens, which contain transmembrane domains. Types IX, XII, XIV, XIX and XXI are often referred to as FACIT collagens (Fibril Associated Collagens with Interrupted Triple helices) based on their primary structure and function [2, 35-37]. In subsequent sections of this chapter we focus on biosynthesis, formation, properties, and remodeling of type I collagen fibrils, which is the subject of this thesis. Hereafter, "collagen" refers to "type I collagen", unless specifically stated otherwise.

2.1. Structure of collagen triple helix

The triple helices of all collagens are formed by three parallel, left-handed, helical α-chains [38]. They are connected by hydrogen bonds, staggered by one residue and wound together into a right-handed supercoil [Fig2.2.]. These three α-chains are either identical (homotrimers) or a mixture of two or three genetically distinct chains (heterotrimers).

Every third residue of each α-chain must be occupied by Glycine (Gly). It is the smallest amino acid and the only one which can pack tightly in the center of the triple helix. The –NH groups of Gly form hydrogen bonds with the –C=O groups of the X-position residues on the adjacent chain. These hydrogen bonds (one per Gly-X-Y triplet) are aligned perpendicular to the helical axis. They tie the α-chains together and provide the main stabilizing force for the triple helix structure [39, 40] [Fig 2.3.(a)]. In contrast to the α-helix or β-sheet conformations, the remaining two

backbone C=O groups in each triplet and the backbone -NH groups of X and Y residues are not involved in peptide hydrogen bonding [38].

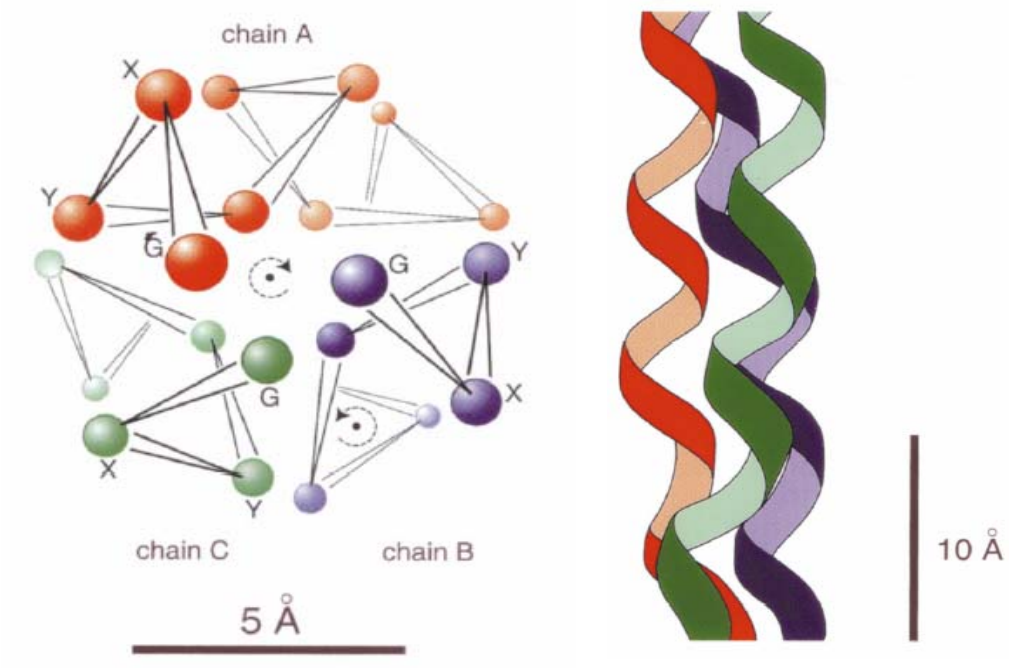

Fig 2.2. Cross-and longitudinal section of a triple-helix

The left picture shows a collagen triple-helix cross-section, viewed from the amino terminus. The positions of Cα atoms are marked by dots. The right picture shows a longitudinal projection of the three α-chains, each forming a left-handed helix with 9.5 Å pitch. The chains are wound together into a right handed supercoil.[38].

Collagen Triple Helix	
Sequence	(Gly-X-Y)n
Helical Parameters of individual chains	
Handedness	Left
Residues per helix turn, n	3.33
Rise/Residue, d (Ao)	2.86
Helix Radius, Cα (Ao)	1.5

[Table 2.1.]. Structural parameters of the Collagen Triple-Helix, based on fiber diffraction data and the crystal structure of a collagenous peptide.[39, 41]

In the repeating Gly-X-Y sequence, the X residue is frequently proline (Pro) and the Y residue is frequently hydroxyproline (Hyp). High Hyp content in the triple helical domain indicates that these residues play a crucial role in stabilizing the triple helical conformation [42]. The helical structure of individual α-chains results largely from the preferential conformation of proline residues in the X position and 4-hydroxyproline residues in the Y position, since the pentagonal pyrrolidine rings of these imino acids are rigid and limit rotation about the peptide N-C bond.

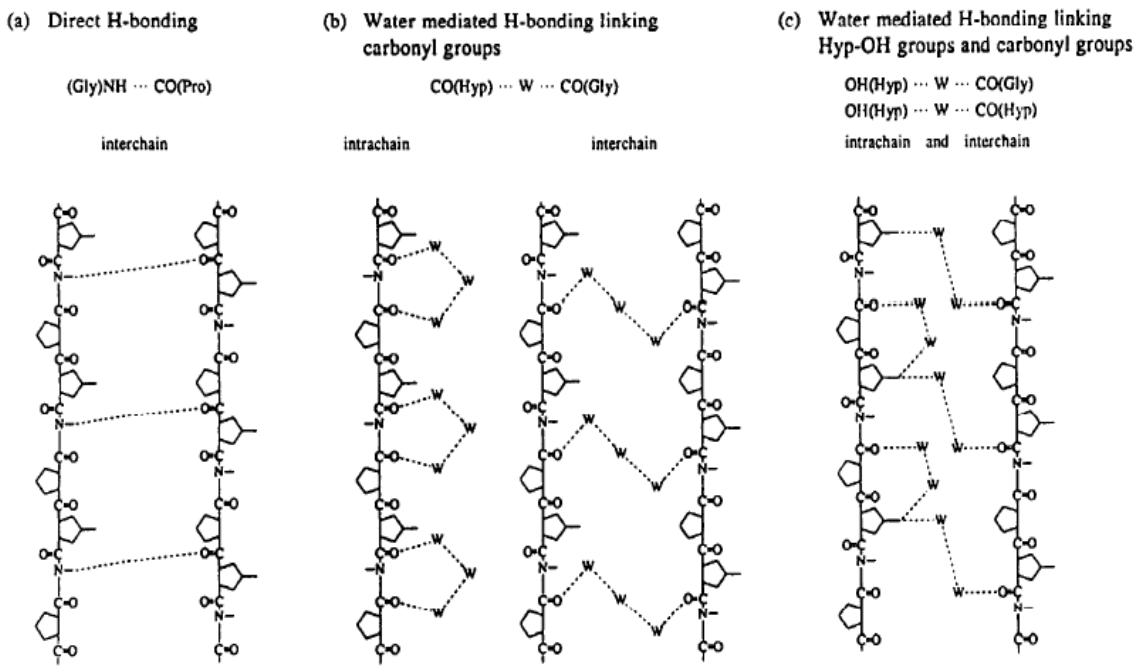

Fig 2.3.Hydrogen bonding patterns found in the triple-helix (reproduced from Ref [40]): (a) direct peptide group hydrogen bonding; (b) water mediated hydrogen bonding linking carbonyl groups; and (c) water mediated hydrogen bonding linking hydroxyproline OH groups and carbonyl groups.

It was also proposed that hydroxyproline contributes to the stability of the triple helix by forming specific hydrogen bonds with water molecules [43]. A highly ordered hydration network (Figure 2.3.(b), (c)) was observed in crystals of collagen-like triple helical peptides formed by three $(Gly\text{-}X\text{-}Y)_{10}$ chains [44, 45].

However, more recent studies of triple helical peptides with fluorinated Pro derivatives suggested that subtle differences in preferential conformations of the

pyrrolidine rings of Pro and Hyp might provide even more important contribution to the triple helix stability than the difference in hydrogen bonding with water [46, 47]. Hydroxylation of the proline residues before Gly, at Y positions, increases the triple helix thermal stability while residues after Gly, at X positions, destabilize the triple helix [46, 47].

The side chains of the X and Y amino acids face outward and are available for intermolecular interactions with surrounding solvent and other molecules [48, 49]. This arrangement allows various amino acid residues (acidic, basic etc) to be accommodated in the molecule. Collagen electrostatics are believed to be determined exclusively by the side chains since the triple helix bears no intrinsic dipole moment [50]. The side chains of the X and Y residues contribute to the hydrophilicity, ionization, hydrophobicity, and steric roughness of the triple helix surface, including the size of the helical groove.

The folding of the three α-chains into the triple helix is nucleated at the C-terminus and proceeds toward the N-terminus [51]. The folding rate depends on the sequence, particularly on the identity of the Y position residues. The highest triple helix propensity and folding rate were observed for Gly-Pro-Hyp triplets [52].

2.2. Collagen biosynthesis

Collagen synthesis, folding, and maturation proceed through a number of intracellular and extracellular steps, which are briefly discussed below [29, 53, 54].

2.2.1. Intracellular steps

Individual collagen chains are synthesized by ribosomes and injected into the lumen of the endoplasmic reticulum (ER) as precursor pro-α-chains. These precursor chains have a short amino-terminal signal peptide required to direct the nascent chain to the ER as well as propeptides at both N- and C- terminal ends. In contrast to most other proteins, the folding of procollagen begins only after completion of the synthesis of all three chains, when the C-terminal ends of the chains exit the ribosomes. The chain synthesis and folding in ER are accompanied by posttranslational modifications of over ~10 % of all collagen residues, including hydroxylation of some proline and lysine residues and glycosylation of some of the resulting hydroxylysine residues[54].

In vertebrate collagens, most proline residues in the Y positions are hydroxylated to 4R-hydroxyproline (4R-Hyp). The presence of an appropriate number of 4-hydroxyproline residues in the -Y- position of the Gly-X-Y triplet is crucial for the stability of the triple helix. For instance, the thermal denaturation temperature (melting temperature) of triple helices with nonhydroxylated proline residues in the Y-position is about 15 ^{o}C lower than in the fully hydroxylated collagen.

Some X and Y prolines are hydroxylated to 3-hydroxyproline, particularly in basement membrane collagens. The 3-Hyp residues both in X and Y positions destabilize the triple helix [55]. Nevertheless, these residues appear to be very important. Their formation requires specialized enzymes. The locations of 3-Hyp residues (e.g., in type I collagen) are highly conserved. However, the specific roles of these residues in collagen structure and formation are not known.

In addition to proline, some of the lysine residues are also hydroxylated. Hydroxyl groups of hydroxylysine residues have at least two important functions: they serve as attachment sites for carbohydrate units and they are crucial for the stability of intramolecular and intermolecular collagen cross-links [29, 56, 57].

In vertebrate collagens, some hydroxylysine residues within triple-helical domains are further modified by covalent attachment of galactose or glucosyl-galactose. This glycosylation is unique to collagens. Its extent varies between collagen types and within the same collagen in different tissues and at different ages. A potential biological role of the specific carbohydrate units would be regulation of lateral assembly of the triple-helical domains. For instance, an inverse relationship between the carbohydrate content and collagen fibril diameter suggests that the bulky glycosylated hydroxylysine residues may sterically hinder formation of highly ordered fibrils [2, 29, 58]

Notably, the hydroxylation of prolines and lysines and glycosylation of hydroxylysines occur only within unfolded regions of collagen chains so that the extent of the posttranslational modification depends on the folding rate. Once the folding is completed, procollagen is transported through the Golgi stack and secreted from cells. [59-62]

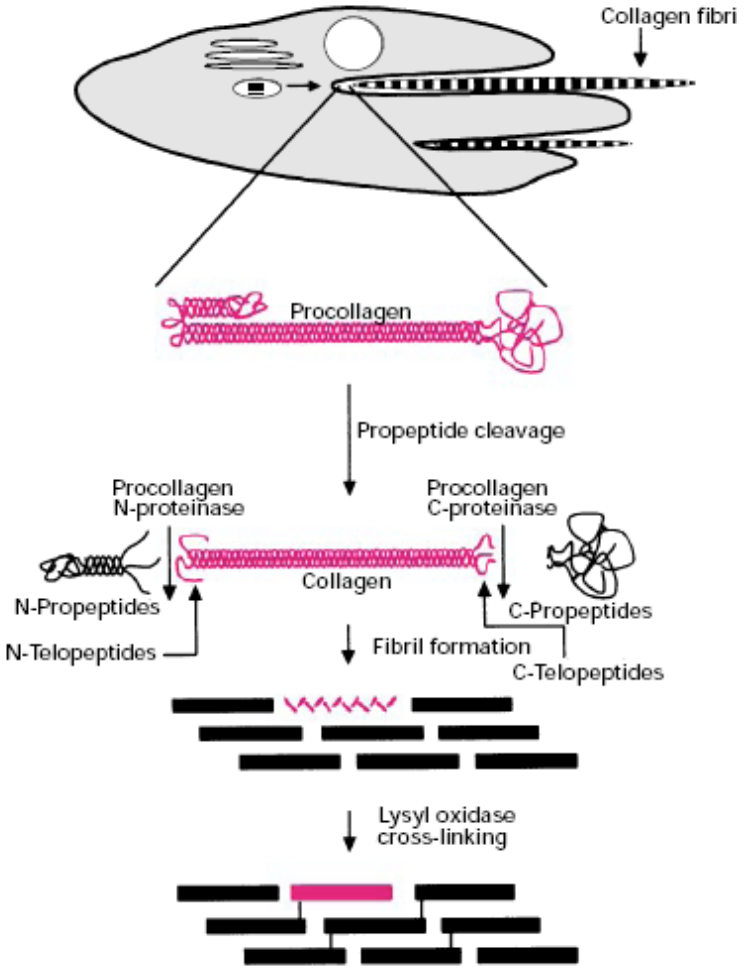

Fig 2.4. Biosynthesis of collagen fibrils (reproduced from Ref [60]).

2.2.2. Extracellular steps

Procollagen secreted from cells is converted into collagen by specialized N- and C-proteinases, which remove the N- and C-propeptides respectively. At the same time, oxidative deamination by lysyl oxidase converts specific lysine and hydroxylysine residues into allysine and hydroxyallysine, respectively [63-65] [66-

68]. The resulting collagen spontaneously assembles into short cross-striated fibril segments approximately 25 to 50 μm in length, which fuse linearly and laterally to form longer and thicker fibrils [69]. Subsequently, allysine and hydroxyallysine aldehyde groups react with amine groups of lysine and hydroxylysine forming intra- and inter-molecular cross-links. The locations of these cross-links depend on the amino acid sequence and quaternary structural arrangements, but may also be quite flexible , e.g., dependent on the type of tissue [63-65]. The lysine and hydroxylysine-derived cross-links stabilize fibrils and provide the tensile strength and mechanical stability to connective tissues [70, 71].

The assembly of collagen molecules into fibrils, known as fibrillogenesis [60, 72], is affected by pH, ionic strength, ion valence, and other micro- and macromolecular additives [73]. The self-assembly of fibrils is commonly considered to be an entropy driven process [69], similar to that occurring in other protein self-assembly systems, such as microtubules[74] and actin filaments [75]. These processes are presumed to be driven by the loss of structural solvent molecules from hydrophobic surfaces. They result in reducing the energetically and entropically unfavorable interactions of hydrophobic surfaces with water [69]. However, it was suggested that the role of the hydrophobic effect appears to be negligible in collagen-collagen interactions. Measurements of forces between triple helices of type 1 collagen at different temperatures, pH, and solute concentrations suggested that the temperature favorable to collagen fibril assembly might be driven by hydrophilic rather than hydrophobic interactions [76, 77].

2.3. In vitro fibrillogenesis

Although cells are known to play an important role in guiding collagen fibril assembly *in vivo*, it is apparent that much of the information that determines the structure of the Extracellular Matrix is contained within the macromolecules which form these structures [78]. Therefore, collagen fibrillogenesis is often studied in a better defined environment *in vitro* with purified solutions of collagen molecules of interest.

In particular, type 1 collagen fibrillogenesis occurs spontaneously at neutral pH and physiological temperature. It can be initiated either by a temperature jump or by cleavage of globular propeptides (C- and N-proteinases). Other important extracellular matrix components, such as other collagens and various noncollagenous proteins, may be added in well-defined amounts in such experiments. The corresponding studies with mixed-component solutions have shown that many matrix molecules affect the fibrillogenesis and, therefore, may play important roles during fibril assembly *in vivo* [79].

In vitro fibril formation is typically monitored by light scattering (turbidity) measurements. Turbidity studies showed collagen fiber assembly to be similar to the growth of salt crystals. Fibrillogenesis starts with a lag period, followed by rapid fibril growth and saturation, showing the characteristic sigmoidal growth pattern [Fig 2.5.]. Similar behavior was described for many other biological self-assembly processes as well, e.g., insulin polymerization and the G-F actin polymerization [80, 81].

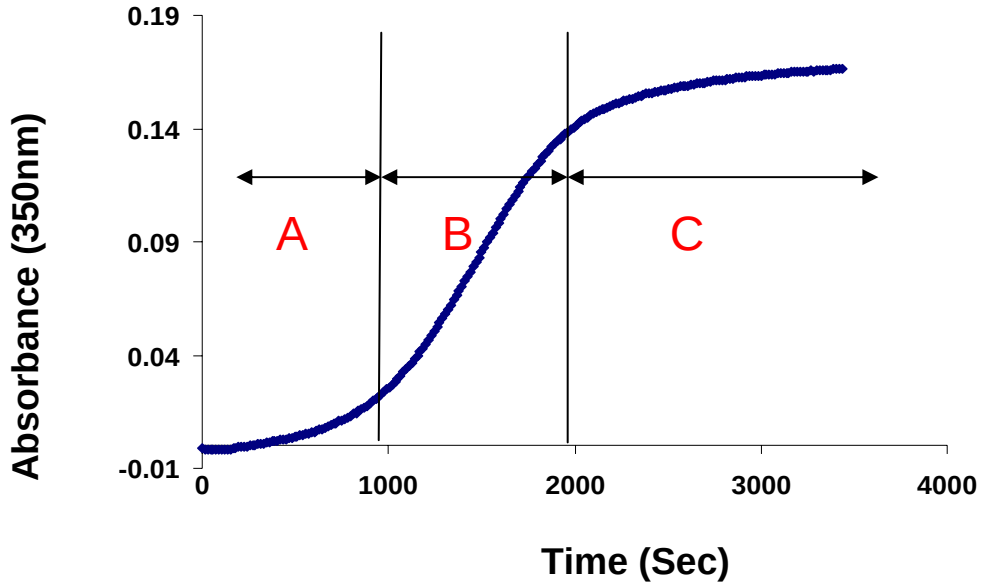

Fig 2.5. Kinetics of *in vitro* fibrillogenesis of type I collagen in 0.13 M NaCl, 30 mM sodium phosphate (pH 7.4) at 32 ^{o}C

(A) Initiation: Soluble collagen aggregates form and grow after a temperature jump from 4 ^{o}C to 32 ^{o}C. It was suggested that a collagen fibril precursor forms and grows linearly until a critical length is reached[82]. It might be short microfibrils or small bundles of short microfibrils[83].

(B) Growth: The fibrils grow linearly and laterally increasing the solution turbidity.

(C) Plateau: Native banded fibrils are obtained[84].

By analogy, the lag phase in collagen fibrillogenesis was initially interpreted as fibril nucleation, that is, the period during which fibril precursors form [85]. This hypothesis was disputed in later studies of the lag phase by viscometry, sedimentation equilibrium, and light scattering, which failed to reveal the formation of aggregates [86]. However electron microscopic observations and laser light scattering studies did

confirm the initial hypothesis[87-89]. Based on measurements of the translational diffusion coefficient and the intensity of scattered light, it was proposed that single collagen molecules initially form staggered dimers and trimers. It was suggested that very long, thin, linear aggregates form during the lag phase, which laterally aggregate during the growth or secondary phase of turbidity-time curves [82, 84].

2.4. Collagen fiber structure and morphology

Packing of the triple helices in fibrils is controlled to a large extent by the amino acid sequence of collagen, specifically by the distribution of polar and hydrophobic residues that are exposed on the triple helix surface. As a result of interactions between these residues, adjacent molecules in fibrils are staggered by integrals of D with respect to one another (D = 67 nm). This D-periodicity can be observed through fibril banding patterns in electron microscopy [Fig 2.6.] [60]. It was suggested that this periodicity arises from the five most favorable staggered conformations for collagen-collagen interaction [90, 91] [Fig 2.7.].

Lateral assembly of axially staggered collagen molecules in fibrils may be viewed as pentameric microfibrils tightly packed together. [92-94]. This hierarchical organization of collagen fibrils was recently confirmed by an X-ray structure of crystalline tendon fibrils [Fig 2.8 (b,c)], which also revealed an overall quasi-hexagonal [95, 96] organization of the molecules in fibril cross-sections [Fig. 2.8 (a)][97].

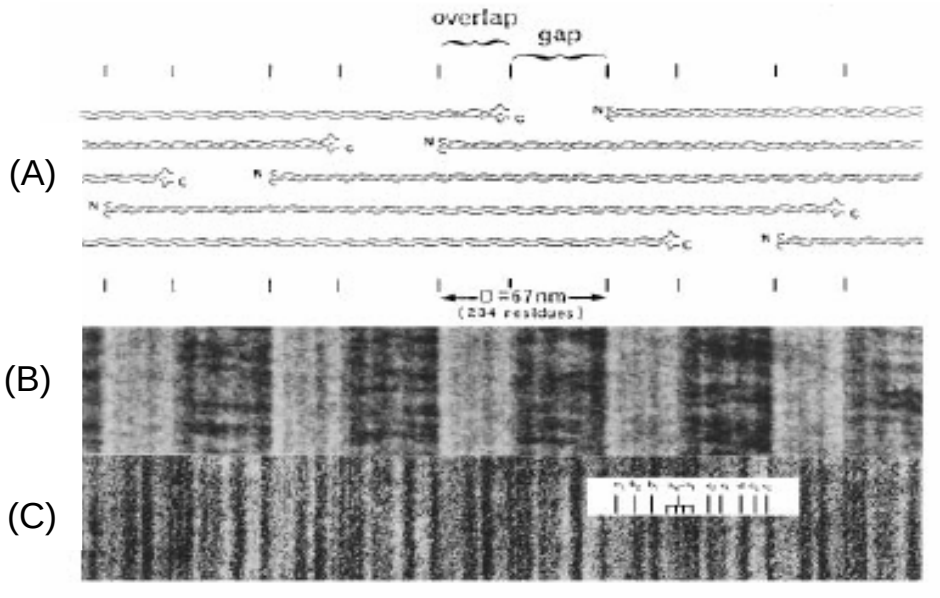

Fig 2.6. Electron microscopy image of a collagen fibril (reproduced from [60]). (A) Schematic drawing of the axial packing arrangement of the triple-helical collagen molecules in a fibril, (B) Negatively stained collagen fibril, (C) Positively stained collagen fibril.

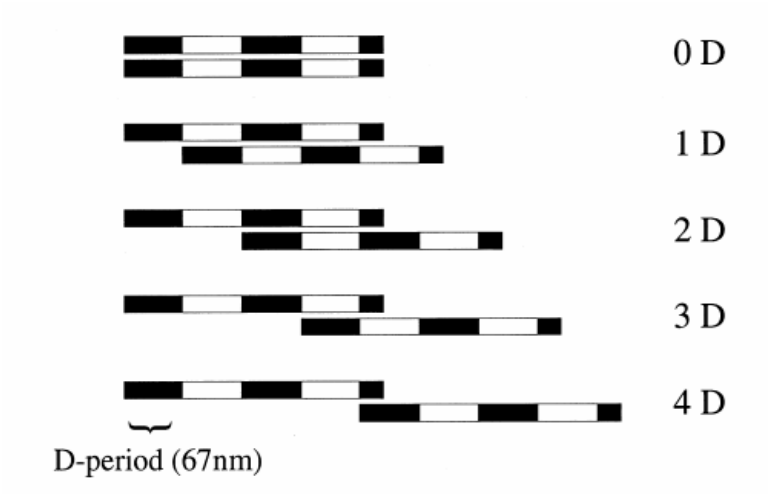

Fig 2.7. The collagen molecule is approximately 300nm with 4.5*D*-periods in length (D = 67nm). There are 5 specific ways for possible binding conformation of 2 collagen molecules, defined as the 0*D*, 1*D*, 2*D*, 3*D* and 4*D* staggers (reproduced from [98]).

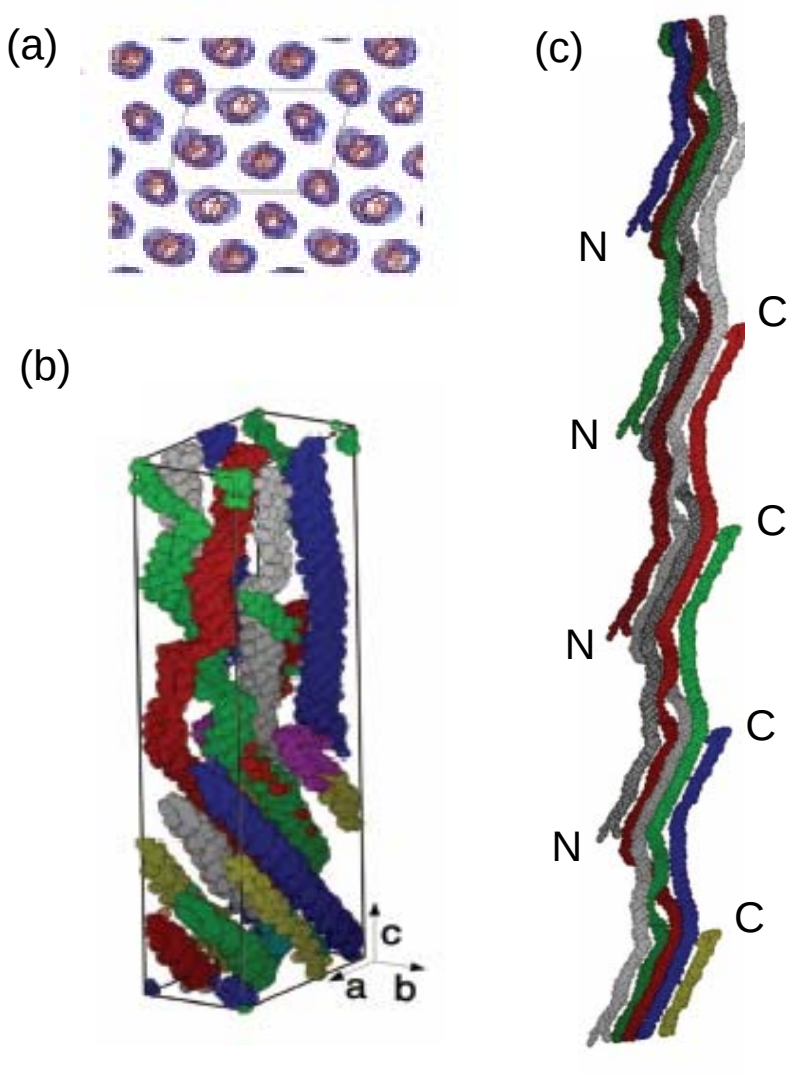

Fig 2.8. X-ray structure of collagen of quasi-crystalline collagen fibrils in tendon (reproduced from Ref [97]). (a) Quasi-hexagonal packing of collagen of fibril cross-section visualized by simulation of electron density. The lines indicate the unit cell containing five molecules. (b) Arrangement of collagen molecules in a unit cell, and (c) a possible collagen microfibril structure. The collagen molecules progress from bottom (N –terminal) to top (C-terminal) showing the conformation of the D-staggered packing and right-handed twist. Each color indicates different molecules. N and C denote N-terminal and C-terminal of a collagen molecule, respectively.

In vivo, collagen fibrils are usually arranged in large bundles usually referred to as fibers [Fig. 2.9]. The arrangement and size of collagen fibers vary with the function of tissues in different regions of the body. For instance, tendon and ligament fibers are thick, long, and highly oriented [Fig 2.9.(a)]. In blood vessels, intestine, and glandular ducts, collagen fibers are helically arranged around the long axes of the tubular structures. In cornea, collagen fibers are layered similarly to plywood, with their packing determining the transparency and refractive ability of the tissue. In skin, cartilage, and bone, collagen fibers are less well oriented and their large scale organization is more reminiscent of a network. Highly interconnected 3D networks are formed by collagen fibers, e.g., in jaw joint discs [Fig. 2.1].

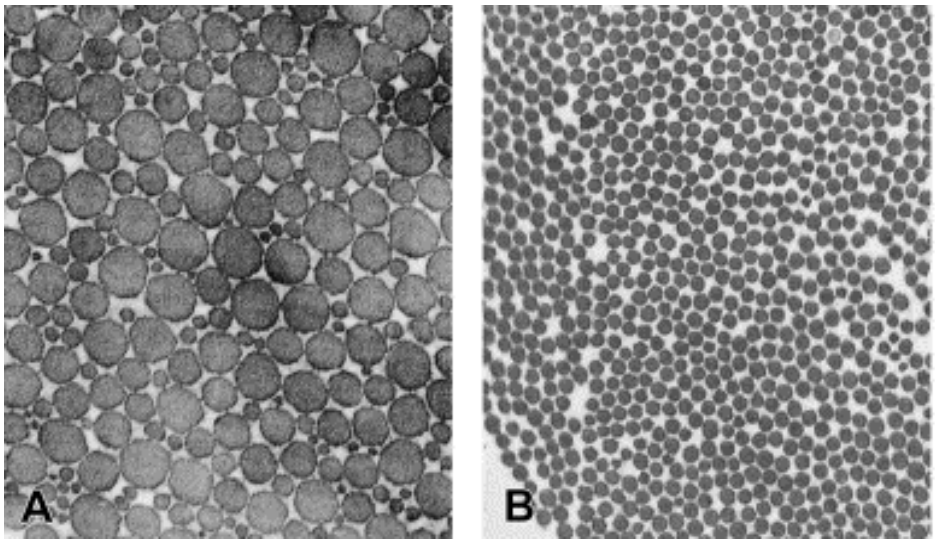

Fig 2.9. Transmission electron micrographs of transected fibril bundles from rat tendon (A) and ox nerve sheath (B) (reproduced from Ref [99]). It shows lateral packing of collagen. The horizontal field of view spans approximately 10 μm (A) and 3 μm (B).

In vitro fibrillogenesis results mostly in the formation of individual collagen fibrils. With increasing collagen concentration, these fibrils form entangled and interconnected 3D networks [Fig. 2.10] rather than fiber bundles, probably reflecting the lack of other matrix molecules that may be responsible for lateral aggregation of the fibrils into fibers. However, each fibril has the same basic organization of the molecules, the same banding pattern in electron microscopy, and the same range of diameters as observed *in vivo*.

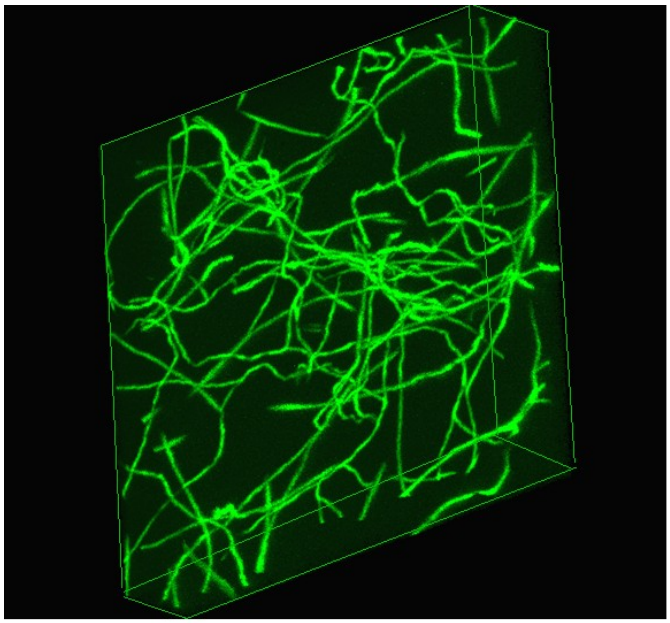

Fig 2.10. Fibrils reconstituted from rat-tail-tendon type I collagen *in vitro*. Collagen type I was polymerized in 8-well chambered cover glass (Schleicher & Schuell). The stained 3-dimensional matrices were viewed with an inverted confocal laser-scanning microscope (Leica TCS SP2) using a 100X, oil immersion lens. Laser light of 488 nm which illuminated the hydrated samples was detected with a photomultiplier tube. Samples were maintained at room temperature.

2.5. Collagen associated proteoglycans

Proteoglycans are the most abundant noncollageneous matrix molecules that are known to associate with collagen fibrils and are likely to affect assembly and lateral association of the fibrils *in vivo*. It was suggested that the proteoglycans may inhibit lateral growth of fibrils, reducing their diameter [100].

Proteoglycans are glycosylated proteins containing long chains of covalently linked glycosaminoglycans, i. e., chondroitin sulfate, dermatan sulfate, heparan sulfate, heparin, and keratan sulfate. The core protein directs proteoglycan biosynthesis, structure, and function [101]. Proteoglycans regulate collagen fibrillogenesis, influence cell growth, and act as tissue organizers. For instance, they affect the tensile strength of skin, corneal transparency, neurite outgrowth [102], and maturation of other specialized tissues [101].

The two largest groups of proteoglycans are small leucine-rich proteoglycans and large, modular proteoglycans [101]. Modular proteoglycans are generally multidomain proteins containing motifs shared with other proteins. Small leucine-rich proteoglycans (SLRPs) have a small core containing a characteristic leucine-rich repeat (LRR) flanked by cystein-rich clusters, which stabilize the protein structure by forming disulfide bonds. There are four cysteine residues at the amino terminal region and two cysteine residues at the carboxyl terminal side [103]. SLRPs include at least nine different molecules. In particular, decorin is a ubiquitous extracellular matrix SLRP found in association with collagen fibrils [104, 105]. Type I collagen has a specific binding site for the decorin core protein [106-108]. Decorin delays lateral

assembly of collagen molecules into fibrils and decreases the diameter of the fibrils by binding to their surfaces [100, 109].

2.6. Mechanical properties of collagen fibrils

The outstanding mechanical properties of connective tissues containing collagen, e.g., tendons, ligaments, and bones, are largely determined by collagen fibrils. Collagen fibrils transmit forces, dissipate energy, and prevent premature mechanical failure of tissues [89, 110]. Tendons are among the most important stress-carrying structures in mammals. They are composed of parallel fibrils from 20 to several hundred nanometers in diameter and fibril bundles, which can be as large as several hundred micrometers [111]. The stress-strain relationship in tendons is illustrated in Fig 2.11 [112, 113].

The structure of collagen fibrils is clearly important for the tensile strength of tendon and mechanical properties of other connective tissues [114-116], but the relationship between changes in this structure, applied stress, and resultant strain are still not completely clear. For instance, it was recently proposed that collagen fibrils may have a relatively hard shell and less dense core, providing for additional flexibility and elasticity [117]. It was also speculated that hierarchical folding of α-chains into left-handed helices wound together in a right-handed supercoils might contribute to the unique mechanical properties of collagen fibrils. As the molecule is stretched, the right handed coil might unwind, resulting in a small extension of the molecular length [118]. The stress induced by further stretching might be applied to the hydrogen bonds that link the polypeptide chains together. A very large stress may

induce breakage of these bonds and also prevent the molecule from regaining its original conformation. Finally, micro-unfolding of the triple helix in regions lacking proline and hydroxyproline residues may allow for reversible, elastic stretching of individual collagen molecules [92, 95, 118, 119]. Note that proteoglycans and other collagen associated proteins may also play a important role in the mechanical properties of tendons, especially for fibril sliding [120, 121].

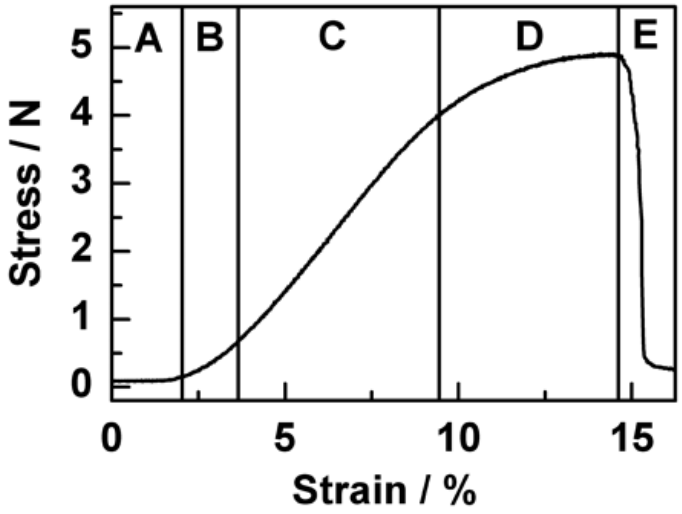

Fig 2.11. The stress-strain curve of rat tail tendon (reproduced from [113]). (A) Toe region: The tendon can be extended with very little force by removing a macroscopic crimp of the fibrils [122]. (B) Heel region: Higher strain removes fibril kinks and reduces disorder in the lateral packing of molecules [23, 123]. (C) Linear region: 3-9 % strain is accommodated by stretching of collagen triple helices and/or cross-links between the helices, e.g., increasing the length of the gap region with respect to the overlap region. Some slippage of molecules within fibrils may also play a role because not all of the elongation could be explained by the stretching of the molecules. (D) Plateau region: This region is pronounced in cross-link deficient collagen and indicates slippage within the fibrils [121]. (E) Rupture: The strain above 15% usually leads to fibril failure.

2.7. Degradation of collagen

Controlled degradation of the extracellular matrix is an integral feature in a variety of biological processes, such as embryonic development, tissue remodeling, and tissue repair. Matrix metalloproteinases (MMPs) are a large family of zinc-dependent endopeptidases that play the most important role in normal matrix degradation during tissue remodeling and repair as well as within pathologies such as cancer invasion [124-126]. Interstitial (soluble) enzymes primarily responsible for cleavage of native fibrillar collagens are known as collagenases, specifically: MMP-1, MMP-8, and MMP-13. The catalytic activities of collagenases towards fibrillar collagens differ to some extent. For example, MMP-1 preferentially cleaves type III and type I collagens while MMP-13 cleaves type II collagen and gelatin more efficiently [126, 127]. Gelatinase-A (MMP-2) is primarily responsible for cleavage of denatured collagen chains, yet it also displays weak activity toward native, triple-helical collagens [128]. Membrane collagenase MT1-MMP (MMP-14) is particularly important for pericellular cleavage of collagen fibrils [129].

These collagenolytic enzymes are active at neutral pH and are produced by many different cells such as fibroblasts, epithelial cells, and leukocytes. However, collagen degradation during bone resorption by osteoclasts, which occurs at acidic pH, also involves an enzyme from a different family known as cathepsin K [130].

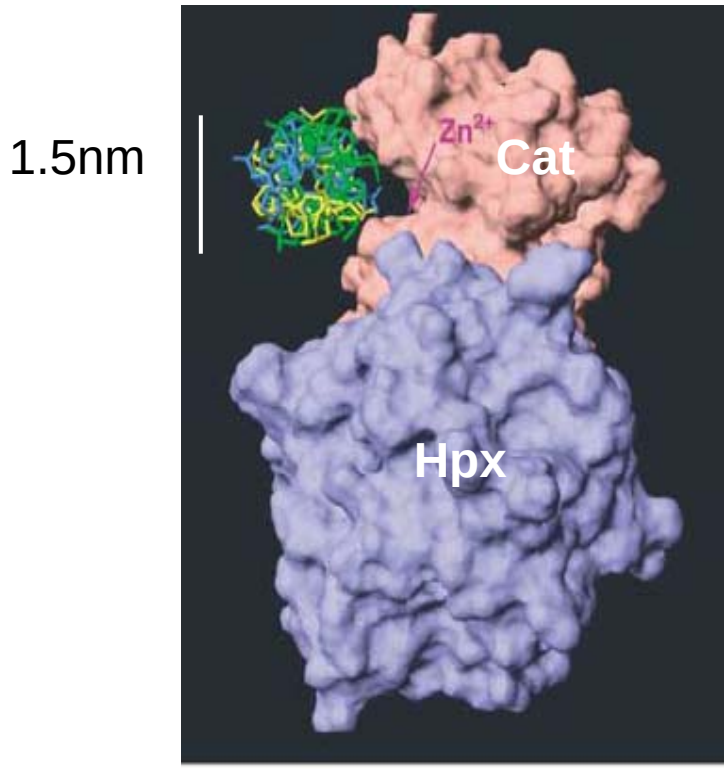

Fig 2.12. Model of collagen triple helix alignment with the active site of MMP-1 (reproduced from Ref [131]). Pink, catalytic domain (Cat); purple, hemopexin domain (Hpx).

All interstitial collagenases and ectopic regions of membrane collagenases are composed of a propeptide, catalytic domain, hinge domain, and hemopexin domain. Their collagenolytic activity is triggered upon propeptide cleavage by other ECM proteinases [124, 125]. The proteolysis of collagen chains occurs within the catalytic domain, but the hemopexin (Hpx) domain, which has no proteolytic activity by itself, is required for cleavage of collagen triple helices [132]. It is believed that the Hpx domain is essential for collagen binding [131].

Recent studies of MMP-1 revealed that proMMP-1 activation is accompanied by

a change in the conformation of the hinge connecting the catalytic and Hpx domains, switching the enzyme from a "closed" to "open" state [133, 134]. However, even in the open conformation, the catalytic site of MMP-1 is located within a narrow catalytic cleft that is only 5 Å wide. Type I collagen triple helix, which is approximately 12 Å in diameter cannot fit within this cleft [Fig 2.11]. The cleft width is sufficient to accommodate only a single polypeptide chain. Thus, unwinding of the triple-helix is required for placement of collagen chains at the active site and peptide bond hydrolysis [131, 135].

The cleavage of type I, II and III collagens occurs between Gly and Leu or Ile residues, generating 3/4 N-terminal and 1/4 C-terminal triple-helical fragments. The cleaved fragments denature spontaneously at body temperature and are further degraded by gelatinases (MMP-2 and MMP-9) and nonspecific tissue proteinases. In type I collagen, the cleavage occurs between Gly775-Ile776 residues of the α1 chains and Gly775-Leu776 residues of the α2 chain. It is believed that other Gly-Ile or Gly Leu bonds are not cleaved because of tighter triple helix packing (higher helix stability) at these sites, e.g., due to Y-position a hydroxyprolines residue in adjacent Gly-X-Y the following repeat triplets [136-138]. Computer modeling and fiber X-ray diffraction suggested that low triple helix stability and partial dissociation of the chains at the cleavage site may be crucial [138]. However, the exact sequence and structural requirements for the substrate recognition by collagenases are still poorly understood. For instance, analysis of type I-III collagens from different species suggested that an Arg residue located on the C-terminal side, 5 or 8 residues away from the cleavable bond may be important for the enzyme binding or chain cleavage

[139]. However, analysis of collagens from OI patients with Leu and Cys substitutions for Arg-780 in the α1(I) chain, which was performed in our laboratory, did not confirm this hypothesis.

2.7. Type I collagen homotrimer

Type I, II, and III collagens evolved from the same evolutionary predecessor. As their common predecessor, type II and III collagens are homotrimers composed of three identical chains. In contrast, type I collagen is normally a heterotrimer of two α1(I) and one α2(I) chains. The α2(I) chain in type I collagen evolved in early vertebrates and was retained through the evolution of all higher vertebrates [140, 141]. Homotrimers of three α1(I) chains [Fig 2.13.] have been found only in fetal tissues [3, 142] and pathological conditions [4-9, 143].

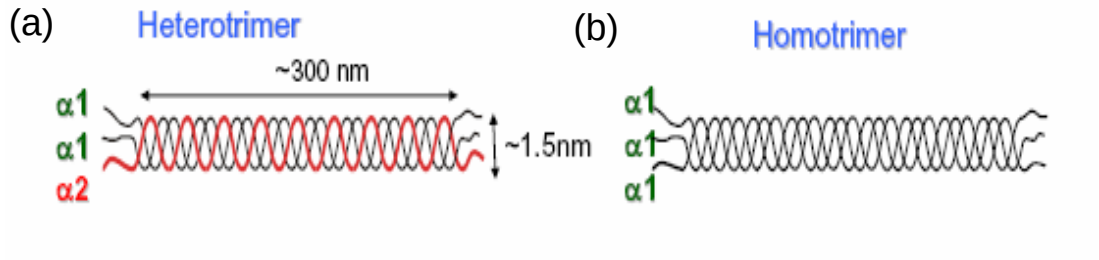

Fig 2.13. Schematic representation of type I collagen triple helix. (a) heterotrimeric triple helix and (b) homotrimeric triple helix.

Total absence of the α2(I) chain in homozygous or compound heterozygous patients was found to result in formation of type I collagen homotrimers, causing rare recessive forms of Osteogenesis Imperfecta (OI) and Ehlers-Danlos syndrome (EDS) [12]. Whereas, heterozygous parents of these patients did not have any pronounced symptoms, even though up to 50% of mature type I collagen in their tissues were α1(I) homotrimers. Synthesis of a fraction of the homotrimers was observed in cancer [6, 7], and some forms of fibrosis [10, 144-146]. Interestingly, a major genetic marker of predisposition to age-related osteoporosis was proposed to be a polymorphism in the α1(I) promoter region, resulting in synthesis of larger than normal amount of α1(I) chains [147-150]. It was speculated that formation of α1(I) homotrimers in these patients might be involved, but the homotrimer presence in patient tissues was not established[150].

The absence of the α2(I) chain decreases the mechanical strength of tissues [23] and leads to a decrease in the order of axial packing [22] and loss of crystalline lateral packing of collagen in mouse tendons [22]. The interactions between homotrimeric triple helices of type I collagen were shown to be different from interactions between the heterotrimers [151]. The homotrimers still assemble into fibrils with normal D-periodicity, but the concentration of the homotrimers necessary for the assembly is higher than for the heterotrimers [22].

The kinetics of homotrimer fibril formation were also found to be different from those of heterotrimers. Virtually, no lag phase was observed upon fibrillogenesis of pepsin treated homotrimers isolated from an oim mouse model of Osteogenesis Imperfecta and homotrimers refolded from α1(I) chains of rat-tail-tendon collagen

[Fig 2.13.]. The unusual fibrillogenesis kinetics of the homotrimers was not a simple consequence of pre-existing fiber nuclei formed by covalently cross-linked molecules. Apparently, the nucleation barrier for the homotrimers is much lower than for the heterotrimers, resulting in much shorter fiber nucleation time. However, the exact origin of this phenomenon remains unknown [18].

Thus, while the homotrimers of type I collagen clearly play an important and special role in embryonic development and pathology, the molecular mechanisms involved in their normal function and pathophysiology remain largely unknown. Better understanding of the physical mechanisms involved in the effects of the homotrimers on formation, structure, mechanical properties, and degradation of collagen fibrils may provide a key to solving this very important puzzle. Investigation of these mechanisms is the subject of the present thesis.

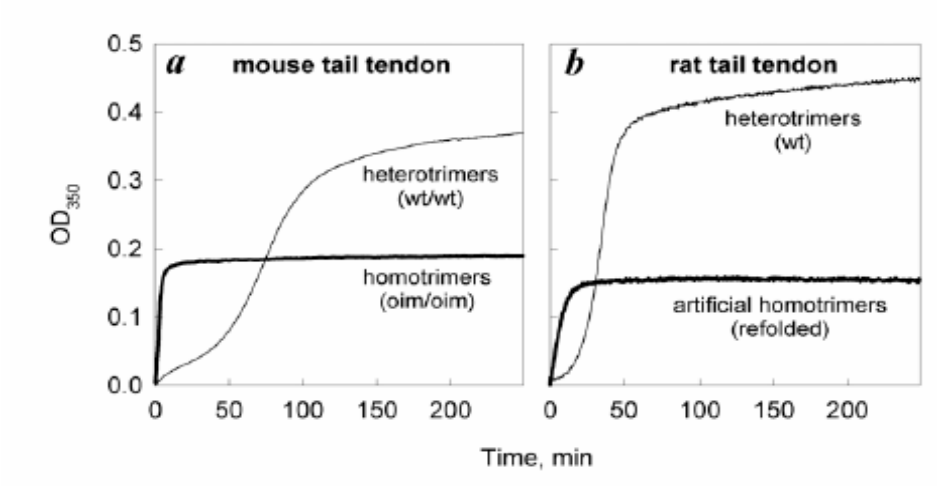

Fig 2.14. Kinetics of *in vitro* fibrillogenesis of type I homo- and heterotrimers from mouse-tail-tendon in 0.13 M NaCl, 30 mM sodium phosphate (pH 7.4) at 32 ^{o}C (a) Pepsin treated

oim /oim and wild-type mouse (b) artificial rat-tail-tendon and rat-tail-tendon heterotrimers (reproduced from[18]).

Chapter 3: Segregation of type I collagen homo- and heterotrimers in fibrils.

This chapter was adapted from a published paper. ***Journal of Molecular Biology*** *Volume 383, Issue 1, 31 October 2008, Pages 122-132*

Normal type I collagen is a heterotrimer of two α1(I) and one α2(I) chains, but various genetic and environmental factors result in synthesis of homotrimers which consist of three α1(I) chains. The homotrimers completely replace the heterotrimers only in rare recessive disorders. In the general population, they may comprise just a small fraction of type I collagen. Nevertheless, they may play a significant role in pathology, e.g., synthesis of 10-15% homotrimers due to a polymorphism in the α1(I) gene may contribute to osteoporosis. Homotrimer triple helices have different stability and less efficient fibrillogenesis than heterotrimers. Their fibrils have different mechanical properties. However, very little is known about their molecular interactions and fibrillogenesis in mixtures with normal heterotrimers. Here we studied the kinetics and thermodynamics of fibril formation in such mixtures by combining traditional approaches with 3D confocal imaging of fibrils, in which homo- and heterotrimers were labeled by different fluorescent colors. Following a temperature jump from 4 to 32 ^{o}C, in a mixture we first observed rapid formation of homotrimer aggregates. The aggregates promoted nucleation of homotrimer fibrils which served as seeds for mixed and heterotrimer fibrils. The separation of colors in confocal images indicated segregation of homo- and heterotrimers at a subfibrillar level throughout the process. The fibril color patterns continued to change slowly after the fibrillogenesis appeared to be complete, due to dissociation and reassociation of the pepsin-treated homo- and heterotrimers, but this remixing did not significantly reduce the segregation even after several days. Independent homo- and heterotrimer solubility measurements in mixtures confirmed that the subfibrillar segregation was an equilibrium property of intermolecular interactions

and not just a kinetic phenomenon. We argue that the subfibrillar segregation may exacerbate effects of a small fraction of α1(I) homotrimers on formation, properties, and remodeling of collagen fibers.

3.1. Introduction

Type I collagen is the most abundant structural protein found in the extracellular matrix of vertebrates. It assembles into fibrils forming the structural scaffold of bone, skin and other connective tissues. It consists of three polypeptide α-chains folded into a 300-nm-long triple helix with short nonhelical terminal peptides (telopeptides). The triple helical region of each chain has the distinctive amino acids repeat (Gly-X-Y) with an obligatory Gly in every third position. The amino acids in the X and Y positions vary along each chain. Frequently, X is a proline and Y is a hydroxyproline [59].

The normal form of type I collagen is a heterotrimer of two α1(I) and one α2(I) chain. The chains are encoded by COL1A1 and COL1A2 genes located on different chromosomes. The chains have significant sequence homology, but α2(I) is more hydrophobic and has fewer proline and hydroxyproline residues. The α2(I) chain appeared in early vertebrates and was retained through the evolution in all higher vertebrates [140, 141]. Nevertheless, homotrimers of three α1(I) chains form, e.g., in embryonic tissues [3] and in a variety of pathological conditions [5-9, 12, 143, 152-154].

Effects of α1(I) homotrimers on tissue properties are poorly understood. The evidence is conflicting. For instance, a severe bone pathology (Osteogenesis

Imperfecta, OI) was reported in a patient whose mutant α2(I) chains were synthesized but not incorporated into collagen molecules [152]. In contrast, no such pathology was found in few patients without any α2(I) chain synthesis [12, 153]. Type I collagen in all these patients consisted entirely of α1(I) homotrimers, but the resulting disease phenotype was dramatically different.

Several studies of α1(I) homotrimer properties and interactions have been reported. It has been established that homotrimer triple helices have higher overall stability [18, 155] and an altered pattern of local stability [18]. Homotrimer helices exhibit less efficient mutual recognition and weaker intermolecular attraction [18, 151]. As a result, fibrillogenesis requires a higher concentration of homotrimers [156]. Homotrimer fibrils are morphologically similar to heterotrimer fibrils [22, 156], but they have lower tensile strength [23] and sparser and less precise packing of the helices [22].

Most of the knowledge accumulated so far pertains to fibrils containing only heterotrimers or only homotrimers, although just a handful of recessive cases with no synthesis or incorporation of α2(I) chains into collagen have been reported. Much less is known about interactions of α1(I) homotrimers with the heterotrimers in mixed fibrils, formation of which may be quite common. The homotrimers were found, e.g., in cultures of some kidney [8, 9] and cancer cells [5, 157-159]. Moreover, cultured osteoblasts with a recently discovered COL1A1 polymorphism were found to secrete up to 10-15% of α1(I) homotrimers [148]. Multiple studies linked this COL1A1 polymorphism to increased bone fragility in age-related osteoporosis in the general population, sparking a renewed interest in α1(I) homotrimers [147-150]. However, it

is still not clear whether and how a small fraction of the homotrimers can affect bone and whether the synthesis of the homotrimers is responsible for bone fragility.

To understand how the presence of a fraction of homotrimers may affect formation and properties of fibrils, we investigated different mixtures of heterotrimeric type I collagen from wild type mouse tail tendon and homotrimeric collagen from homozygous oim mouse tail tendon. The latter mice make a nonfunctional α2(I) chain that is not incorporated into collagen, similar to the OI patient discussed above [160]. We measured the *in vitro* fibrillogenesis kinetics after a temperature jump from 4 to 32 ^{o}C and the equilibrium solubility of the homo- and heterotrimers at 32 ^{o}C as a function of the mixture composition. We used confocal microscopy to examine the morphology of the resulting fibrils and partitioning of homo- and heterotrimers labeled by different fluorescent dyes.

To achieve equilibration of fibrils with the surrounding solution, we pretreated the homo- and heterotrimers by pepsin at 4 ^{o}C. By partially degrading the telopeptides, pepsin prevents irreversible covalent crosslinking between the molecules [82, 161]. While such a treatment alters the fibrillogenesis kinetics [82, 162], it does not affect the integrity of the triple helices and it does not alter their packing and interactions in fibrils [163].

We found that homotrimers affected the fibril bending rigidity and nucleation mechanism. More importantly, we observed equilibrium segregation of homo- and heterotrimers at a subfibrillar level. We argue that this segregation may dramatically exacerbate effects of even a small fraction of the homotrimers on tissues.

3.2. Results

3.2.1. Fibrillogenesis Kinetics

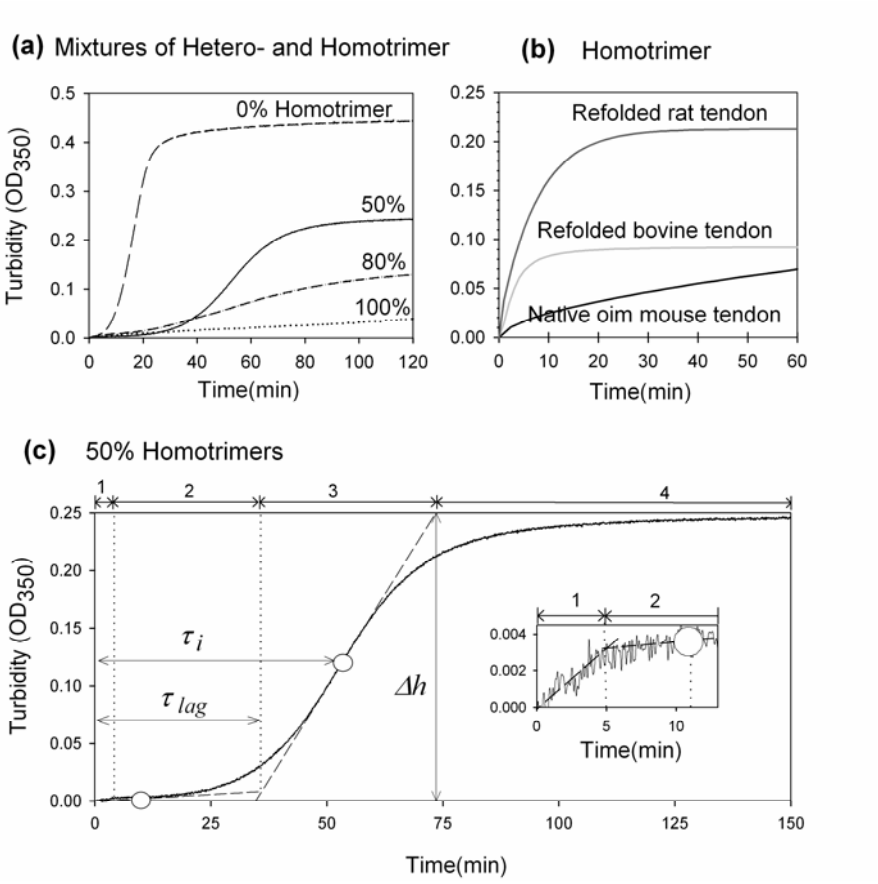

Fig 3.1. Fibrillogenesis kinetics in hetero/homotrimer mixtures of type I collagen: (a) 0.3 mg/ml mixtures of type I collagen from wild type (heterotrimers) and homozygous oim (homotrimers) mice; (b) homotrimers from homozygous oim mouse tendons (native, 0.2 mg/ml, 20mM sodium-phosphate), bovine Achilles tendons (refolded, 0.36 mg/ml, 30mM sodium-phosphate),and rat tail tendons (refolded, 0.25 mg/ml, 20mM sodium-phosphate); (c) 1:1 mixture of mouse heterotrimers and homotrimers (50% curve in (a)). In (c), the white circles show two inflection points on the turbidity curve; the dashed tangential lines indicate turbidity trends at different fibrillogenesis steps; the vertical dotted lines separate five fibrillogenesis steps: 1-early growth, 2- lag phase, 3- delayed growth, and 4- high plateau.

After a temperature jump from 4 to 32°C, the turbidity of pepsin-treated heterotrimer solutions in 0.13 M NaCl, 30 mM sodium phosphate, pH 7.4 followed the expected [164] time course with lag, growth, and plateau phases (Fig.3.1. (a), dashed line). The turbidity, which is caused by light scattering from forming collagen fibrils, was monitored by measurement of the optical density at 350 nm (OD_{350}). In contrast, we observed no lag phase for natural homotrimers purified from oim mouse tail tendons and reconstituted homotrimers refolded from rat tail and bovine Achilles tendon collagen (Fig 3.1.(b)), consistent with a previous report [151]. Slightly different sodium phosphate and collagen concentrations were used for collagen from different species to optimize the fibrillogenesis rate and OD_{350}. To avoid potential counterion-induced aggregation, no divalent or multivalent ions were present in the fibrillogenesis buffer, and collagen was extensively dialyzed against 2 mM HCl before mixing with the fibrillogenesis buffer.

In mixtures of heterotrimers and homotrimers, we observed four fibrillogenesis steps: 1 – early growth, 2 – lag-phase, 3 – delayed growth, and 4 – plateau (Fig. 3.1. (c)). The early growth was probably associated with formation of homotrimer nuclei and fibrils, while mixed and/or heterotrimer fibrils presumably formed primarily during the lag-phase and delayed growth. Table 3.1 shows turbidity parameters extracted from the experiments in which constant, 0.3 mg/ml total concentration of collagen and various molar fractions of homotrimers were used. The final turbidity at the plateau gradually decreased with increasing homotrimer fraction, but the lag time, maximum-growth-rate delay, and half-maximum-turbidity time, defined in Fig.3.1, appeared to change only between 0 and 50% homotrimers. Note

that the plateau turbidity value may be difficult to interpret since it depends both on the number and size of the fibrils. Also, the lag time and maximum-growth-rate delay are not applicable for 100% homotrimers.

Table 3.1. Kinetic parameters of fibrillogenesis.

Homotrimers (%)	Lag time (τ_{lag}, min)	Maximal-growth-rate delay (τ_i, min)	Half-maximal-turbidity time ($\tau_{1/2}$, min)	Final turbidity (Δh)
0	9	17	16	0.46
50	36	52	54	0.25
80	26	53	55	0.13
90	34	52	74	0.09
100			55	0.04

All parameters were determined from turbidity curves as shown in Fig. 1 (c): τ_{lag} is the time at the end of the lag phase, τ_i is the time at the inflection point during the delayed growth step, $\tau_{1/2}$, is the time at 50% of maximal turbidity, Δh is the maximal turbidity reached on the high plateau (measured at 150 min).

3.2.2. Fibril morphology and molecular segregation

To test whether homo- and heterotrimers coassembled into the same fibrils, we pre-labeled them with Alexa fluor 488 (green color, G) and Alexa fluor 568 (red color, R). At less than one Alexa fluor dye per 4-5 collagen triple helices, we observed normal turbidity curves, critical fibrillogenesis concentrations and fibril morphology (see below). In contrast, Cy dyes altered the fibrillogenesis kinetics

curve and induced formation of non-fibrillar aggregates even at low labeling fractions.

We prepared 0.3 mg/ml binary mixtures of G- and R-labeled homo- and heterotrimers in 0.13 M NaCl, 30 mM sodium-phosphate, transferred them into sealed microscopy cells at 4 ^{o}C followed by equilibration at 32 ^{o}C for several hours. Subsequent examination by confocal microscopy in a controlled environment chamber set at 32 ^{o}C revealed three dimensional (3D) networks of flexible, long heterotrimer fibrils (Fig. 3.2 (a)). The homotrimers formed fewer fibrils, which looked like straight spears emanating in several directions from a clump-like common center (Fig. 3.2 (b)). In homo/heterotrimer mixtures, we observed an intermediate morphology (Fig. 3.2 (c)). We observed similar morphologies of unlabeled collagen fibrils by light reflection confocal microscopy.

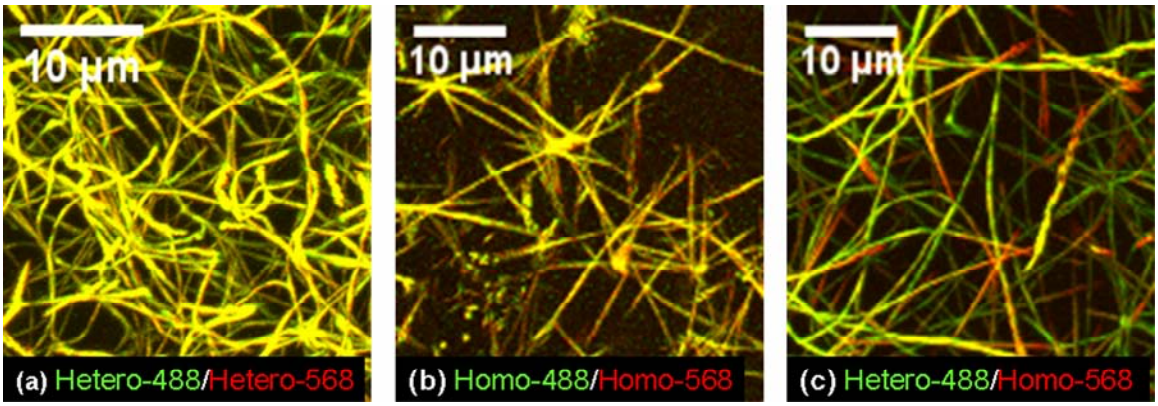

Fig 3.2. 3D reconstruction of confocal images of collagen fibril networks in different 0.3 mg/ml mixtures of fluorescently labeled wild type heterotrimers (Het) and oim homotrimers (Homo) from mouse tail tendons. Molecules labeled by AlexaFluor-488 (G) are

shown in green color. Molecules labeled by AlexaFluor-568 (R) are shown in red color. The yellow overlay color appears in fibrils containing approximately the same amount of molecules labeled by each dye.

In mixtures, both heterotrimers and homotrimers appeared to be present within each fibril. However, highly heterogeneous color distribution, particularly in thinner fibrils and at fibril tips, suggested some segregation of the homo- and heterotrimers at a sub-fibrilarl level, e.g. into separate micro-fibrils. We quantified segregation by calculating the relative excess of red color in each pixel (I_R-I_G/I_R+I_G) in the range of -1.0 (completely green) to 1.0 (completely red). Here I_R and I_G represent red and green intensities normalized as described in Methods. The resulting color separation fraction is shown by relative frequency histograms in Fig 3.3. In controls where the same kind of collagen was labeled with both colors (hetero-G/hetero-R and homo-G/homo-R mixtures) we observed a single peak at I_R-I_G/I_R+I_G=0. In hetero-G/homo-R mixtures, however, we found two distinct peaks, indicating segregation of the molecules into homo- and heterotrimer rich regions.

Note that the broadening of the homo-G/homo-R mixture histogram compared to the histogram for hetero-G/hetero-R may be caused by thinner homotrimer fibrils. For thinner fibrils, statistical variation in the number *N* of molecules labeled by each color within the focal volume may cause an appreciable peak broadening. In a confocal microscope, the fluorescence is collected from a small focal volume (<1 μm^3), which may contain *N*~1000 or fewer labeled molecules. A smaller number of molecules in thinner fibrils will result in a larger statistical variation in I_R-I_G/I_R+I_G

(which is proportional to $1/\sqrt{N}$). Chromatic aberrations, producing displacement of one color with respect to the other in the image, may cause further broadening of the peak in the histogram. The same chromatic aberration will cause a larger change in I_R-I_G/I_R+I_G for thinner fibrils.

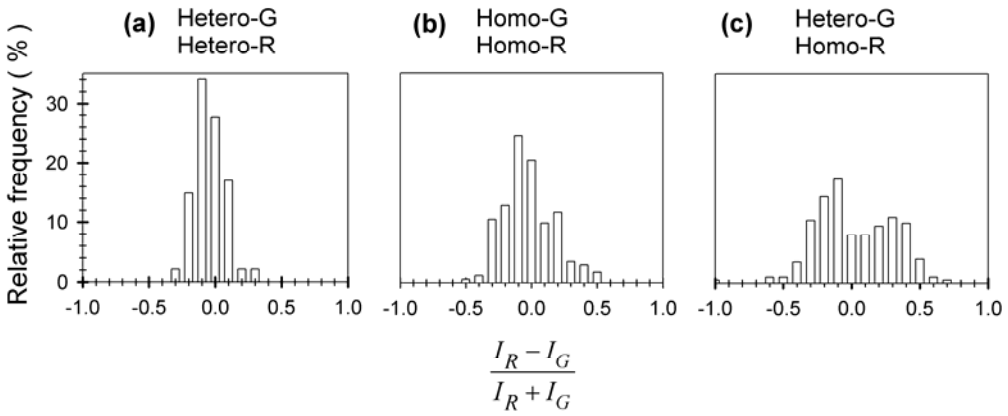

Fig 3.3. Collagen segregation in mixed fibrils. Each histogram of I_R-I_G/I_R+I_G fluorescence intensity ratio shows distribution of R and G labeled molecules. Note that each voxel of the confocal image may contain *N*~400 or more labeled molecules.

Coassembly accompanied by subfibrillar segregation also occurred when a homotrimer solution was added to preformed heterotrimer fibers. A 0.3 mg/ml solution of G-labeled heterotrimers in 30mM sodium-phosphate, 0.13 M NaCl was equilibrated for 2 hours at 32^{o}C. Subsequent examination by confocal microscopy revealed a well-formed fibril network (Fig. 3.4 (a)). A cold (4 ^{o}C) 0.3 mg/ml solution

of R-labeled homotrimers in the same buffer was then injected into the heterotrimer fibril network by a micropipette, to achieve ~1:3 homo-: heterotrimer ratio in the final mixture. Following the injection, the sample was re-equilibrated at 32 ^{o}C and repeatedly re-examined by confocal microscopy in several different areas. After ~1 h, homotrimers were observed predominantly at the tips of preformed heterotrimer fibrils, creating an appearance of longitudinal fibril growth (Fig. 3.4 (b)). After, ~ 2h, more homotrimers were observed in other fibril areas (Fig.3. 4 (c)). Their distribution was more reminiscent of that observed after fibrillogenesis of premixed samples (Fig. 3.2 (c)). No independent homotrimer fibrils were observed, suggesting that the interaction between a homotrimer and a heterotrimer was more favorable than between homotrimers. In contrast, injection of a cold solution of G-labeled heterotrimers to preformed R-labeled homotrimer fibrils resulted in independent formation of separate heterotrimer fibrils in addition to heterotrimer binding to preexisting fibrils (not shown).

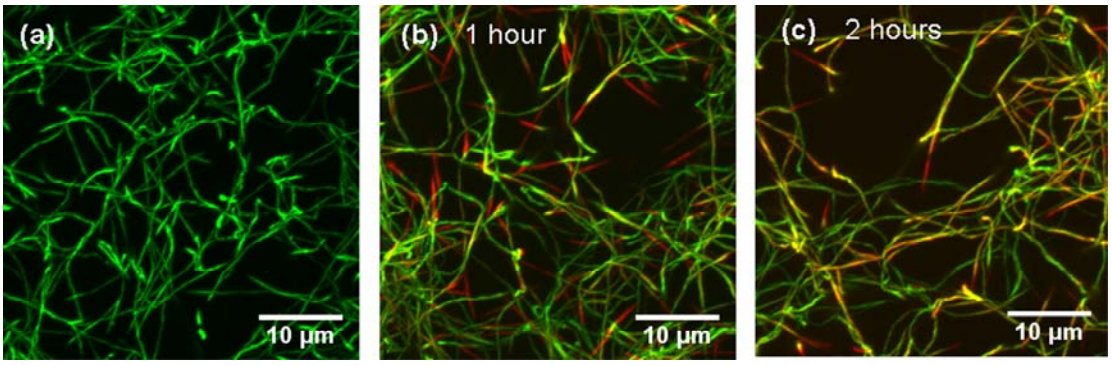

Fig 3.4. 3D reconstruction of confocal images illustrating interaction of homotrimer solution with preassembled heterotrimer fibrils. (a) Heterotrimer-G fibrils before addition of

homotrimer-R solution. (b) Mixed fibrils after 1 hour (b) and 2 hour (c) equilibration with added homotrimer solution at 32^{o}C.

3.2.3. Collagen solubility

To further investigate the segregation of homo- and heterotrimers, we systematically examined their solubility in equilibrium with mixed fibrils (sometimes referred to as a critical fibrillogenesis concentration). Different mixtures of homo- with heterotrimers in 30mM Na-phosphate, 0.13M NaCl, pH 7.4 (0.3 - 1.6mg/ml total collagen concentration) were equilibrated for one day at 32 ^{o}C. Fibrillar and soluble collagens were then separated by centrifugation and fluorescently labeled with Cy5. The composition of the fibrils (pellet) and the solubilities of the heterotrimers and homotrimers (their concentrations in the supernatant) were analyzed by gel electrophoresis, relying on linear dependence of the fluorescent intensities of the gel bands on the concentration of the corresponding chains in the sample (Fig. 3.5 (a)).

The first lane of the gel in Fig. 3.5 (a) shows electrophoretic bands of the α1(I) chain, α2(I) chain, and covalent dimers of $\alpha1\alpha1$ and $\alpha1\alpha2$ in a 100% heterotrimer sample. The relative intensities of the α2 and $\alpha1\alpha2$ bands decreased with increasing homotrimer fraction in the sample. Only α1 and $\alpha1\alpha1$ bands were present in a 100% homotrimer sample (last lane in Fig. 3.5(a)). Based on the calibration curves for the $\alpha1/\alpha2$ and of $\alpha1\alpha1/\alpha1\alpha2$ intensity ratios shown in Fig 3.5 (b), we determined the homotrimer fraction in each pellet and supernatant. By comparing the intensities of the α1 bands with those of standards labeled at the same time and identical conditions, we measured the concentration of the α1 chains in each

supernatant, from which we determined the solubilities of both homo- and heterotrimers.

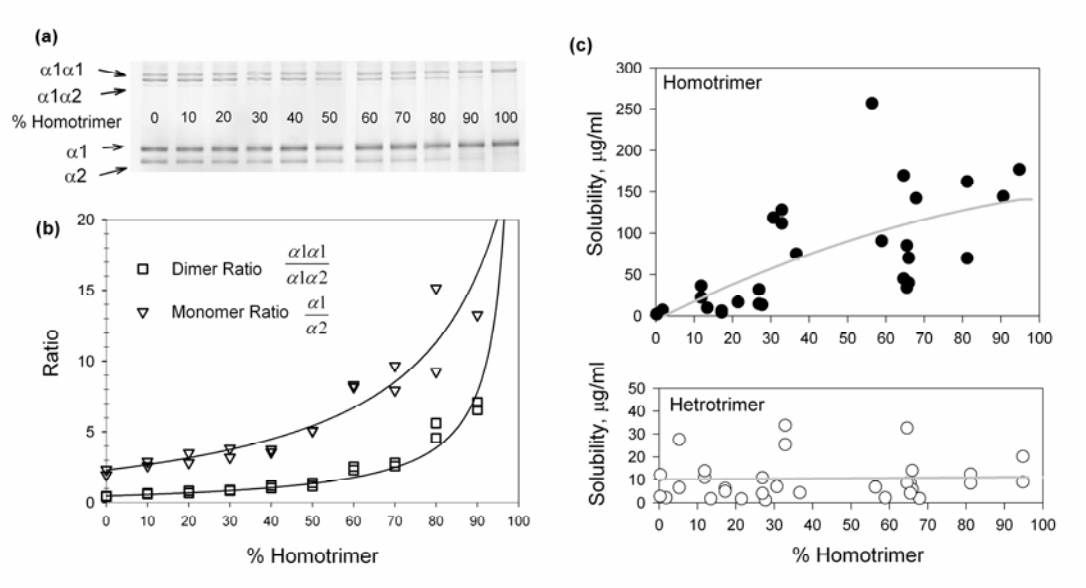

Fig 3.5. Measurement of hetero- and homotrimer solubility by gel electrophoresis. (a) Monomer and dimer gel bands of type I collagen chains in different initial mixtures. (b) Calibration curves for monomer (α1(I)/ α2(I)) and dimer (α1(I)-α1(I)/ α1(I)-α2(I)) fluorescent intensity ratios of the corresponding bands shown on the gel. (c) Concentrations of soluble hetero- and homotrimer in equilibrium with fibrils after overnight equilibration at 32 ^{o}C plotted vs. fibril composition. The soluble and fibrillar collagen fractions were separated by centrifugation and analyzed by gel electrophoresis.

Fig. 3.5(c) shows the solubility of the hetero- and homotrimers in equilibrium with fibrils containing different homotrimer fractions. As in previous reports [156], we found the solubility of the homotrimers to be 15-20 times higher than that of the

heterotrimers in pure samples. The solubility of the homotrimers decreased with their decreasing fraction in fibrils, consistent with a simple ideal mixture model for the molecules within fibrils (Supplementary Material). In contrast, the solubility of the heterotrimers was independent of the overall fibril composition, inconsistent with the ideal mixture model. Such constant solubility, however, may be expected when a fraction of the heterotrimers is completely segregated from the homotrimers within some fibril regions (see Discussion).

3.3. Discussions

3.3.1.The spear-shape of homo trimer fibrils indicates increased bending stiffness.

We found that pepsin-treated homotrimers formed more rigid, spear-like fibrils compared to flexible, thread-like heterotrimer fibrils (Fig. 3.2). Previously, the spear-like morphology was reported for both homo- [156] and heterotrimers [165]. In our experiments, all fibrils were thinner than in these previous studies, making them more susceptible to bending by thermal fluctuations and weak convective flows, but only heterotrimer fibrils exhibited significant bending.

The apparent higher bending rigidity of homotrimer fibrils cannot be explained by their thickness, packing or intermolecular cross-linking. Homotrimer fibrils have the same D-periodic axial alignment of molecules [22] and lower lateral density in the cross-section [22, 151]. In addition, our homotrimer fibrils were thinner than heterotrimer fibrils, which should make them easier rather than more difficult to bend. Formation of covalent cross-links, which might affect the bending rigidity, was precluded by the pepsin treatment.

Thus, the increased homotrimer fibril stiffness is likely to be related to the rigidity of their component triple helices. Since bending distorts helical structure, higher stability of homotrimer helices [18, 155] may contribute to their increased rigidity. The reported difference in the pattern of flexible and rigid regions within homo- and heterotrimer helices [18] may also contribute by altering the alignment of these regions on adjacent molecules. The increased rigidity of homotrimer fibrils may

play an important role, e.g., in the reduced tensile strength of tendon [23] and increased fragility of bone [155] in oim mice.

3.3.2.Non-fibrillar aggregates nucleate fibrillogenesis of pepsin-treated homotrimers

Although homotrimers have higher solubility[156]and exhibit weaker intermolecular attraction in fibrils [151], they assemble into fibrils without the lag phase [18], i.e. fibril nucleation occurs much faster in homotrimers than heterotrimers. Our new findings not only confirm the latter counterintuitive observation but also suggest why this observation may not conflict with less efficient recognition between homotrimers.

The appearance of homotrimer fibrils as multiple spears growing from a common center indicates that heating from 4^{o}C to 32^{o}C induces non-specific, rapid aggregation of homotrimers into non-fibrillar clumps without proper mutual alignment of molecules. The fibrils are nucleated within these aggregates. They grow outward in several directions resulting in the characteristic morphology of multiple spears emanating from a common center (Fig. 3.2). The interactions causing the initial, non-specific aggregation may be weaker than the recognition and attraction at the proper alignment, but they seem to have no energy barrier. Thus, the aggregation proceeds rapidly even though the energy gain from it may be lower than from proper fibril assembly. It is detected by light scattering as the initial phase of fibrillogenesis, so that the solution turbidity increases without a lag phase. The kinetics of formation of homotrimer fibrils with properly aligned molecules is masked by this aggregation.

Presently, we do not know whether it is slower or faster than formation of heterotrimer fibrils.

Fibrillogenesis of pepsin-treated heterotrimers does not involve formation of non-fibrillar aggregates: Heterotrimers remain undetected by light scattering or confocal microscopy during the lag phase until they overcome a nucleation barrier and assemble into well-defined fibrils. Note that the heterotrimer nucleation barrier may be higher even though their mutual interactions in fibrils at equilibrium are more favorable than between homotrimers. The observed initial kinetics of heterotrimer fibril formation was slower than for homotrimers because it involves higher energy intermediates. This does not contradict more favorable equilibrium interactions between heterotrimers within fibrils.

Interestingly, rapid formation of non-fibrillar aggregates after the temperature jump was also observed in homo- and heterotrimer mixtures (Fig. 3.1) but not when the homotrimers were added to preformed heterotrimer fibrils (Fig. 3.4). The latter observation suggests that homotrimer binding at tips of heterotrimer fibrils occurs without a significant energy barrier and that this binding is more energetically favorable than the non-fibrillar aggregation.

The non-fibrillar aggregation seems to be a property of only pepsin-treated α1(I) homotrimers with partially degraded telopeptides, i.e., it is prevented by the α2(I) chain and by α1(I) telopeptides. Indeed, pepsin-treated heterotrimers, containing the α2(I) chain, do not form such non-fibrillar aggregates. Full-length homotrimer molecules with intact α1(I) telopeptides, which were not treated by pepsin, were also reported to form isolated spears after a normal lag phase rather than

hedgehog-like spear clusters [156]. By preventing the non-specific aggregation, the α2(I) chain and the telopeptides may play an important role in initial reorientation and alignment of triple helices required for fibril formation.

3.3.3. Homo- and heterotrimers co-assemble within the same fibril but segregate at a sub-fibrillar level

The observed multistep kinetics of fibril formation (Fig. 3.1 (c)) suggests the following sequence of events in mixtures of pepsin-treated homo- and heterotrimers. Based on the properties of pure homo- and heterotrimers, the first step of early growth likely involves fast non-fibrillar aggregation of homotrimers followed by the outward growth of homotrimer fibrils. The next step (lag phase) involves nucleation of mixed homo/heterotrimer fibrils or predominantly heterotrimer fibrils, which may be promoted by binding of heterotrimers to the homotrimer aggregates and fibrils. Provided that the binding/diffusion time is shorter than the nucleation time, such binding may explain why the lag time at this step appears to be independent of the heterotrimer fraction in the mixture. The growth of heterotrimer enriched fibril regions from these nuclei occurs at the delayed growth step of the turbidity curve. The sequential formation of homo- and heterotrimer enriched fibril regions may contribute to the sub-fibrillar segregation observed in our experiments. The appearance of homotrimers at fibril tips may be related, e.g., to their higher solubility and, therefore, lower binding energy. Once formation of heterotrimer enriched fibrils is nucleated, it proceeds faster than homotrimer binding or assembly, depleting the surrounding solution of heterotrimers. The remaining homotrimers bind last, thereby

appearing at fibril tips. It is also possible, however, that homotrimers preferentially bind at fibril tips due to different geometry of intermolecular contacts at the growing tip vs. the side surface. We presently have no data that could be unequivocally interpreted in terms of the influence of fibril geometry and characteristic distance scales for segregation.

In addition to kinetics, sub-fibrillar segregation of homo- and heterotrimers may result from equilibrium interactions, as indicated by the persistence of the segregation after overnight or longer equilibration. Pepsin-treated collagen molecules dissociate from and re-associate with fibrils since they do not form intermolecular cross-links. The equilibrium between fibril surfaces and surrounding solution was previously found to be established after several hours [166]. Consistently, here we observed significant fibril remodeling several hours after addition of homotrimers to preformed heterotrimer fibrils. After initial association at tips of heterotrimer fibrils, the homotrimers gradually appeared along the whole length of the fibrils and the heterotrimers appeared in the places where the fibril tips were elongated (Fig. 3.4). We should note, however, that complete re-equilibration of composition in fibril cores may take much longer than at fibril surfaces and it may not be practically attainable.

Additional evidence for equilibrium segregation of homo- and heterotrimers is the dependence of their equilibrium solubility on the mixture composition (Fig. 3.5). At equilibrium between a homogeneously mixed binary aggregate and a surrounding solution, the solubility of the minor component should be approximately proportional to its fraction within the aggregate provided that this fraction is sufficiently small

(Supplementary Material). We observed such proportionality from 0 up to ~30% of homotrimers in aggregates, suggesting that a relatively small amount of homotrimers can be well mixed with heterotrimers. In contrast, the solubility of heterotrimers did not change at all when their fraction was varied, suggesting poor mixing of the molecules in fibrils containing a small fraction of heterotrimers.

Much lower solubility of the heterotrimers [156] (Fig. 3.5) and their segregation suggest that heterotrimer-heterotrimer interactions in fibrils are much more energetically favorable than homotrimer-homotrimer and heterotrimer-homotrimer interactions. Presently, we do not know which of the latter two interactions is more favorable than the other. The asymmetric segregation, which occurs at a low fraction of heterotrimers but not at a low fraction of homotrimers, significantly complicates evaluation of the heterotrimer-homotrimer interaction energy. It indicates, e.g., that the heterotrimer-homotrimer interaction may be different for a homotrimer surrounded by heterotrimers and for a heterotrimer surrounded by homotrimers.

3.3.4. Biomedical implications

The subfibrillar segregation may significantly enhance the effect of even a small homotrimer fraction on the collagen fibril integrity and mechanical properties because the homotrimers have different triple helix stability, stretching elasticity, and bending rigidity. The segregation certainly occurs *in vitro* in fibrils assembled from pepsin-treated collagen molecules with partially degraded telopeptides. *In vivo*, the fibrils assemble from molecules with intact telopeptides in a different environment

and under different conditions. Not only it is a non-equilibrium process, but it also may be actively regulated by cells [78]. The resulting fibrils are stabilized by irreversible covalent cross-links. Their remodeling involves proteolytic degradation of old collagen and its replacement with newly synthesized molecules rather than equilibrium dissociation and re-association of the molecules. Unfortunately, differential fluorescent labeling and direct observation of the subfibrillar segregation *in vivo* or even in cell culture is presently unrealistic.

Still, our findings suggest that the tendency of homo- and heterotrimers to segregate, particularly the kinetic segregation, is likely to be preserved *in vivo* as well. The dramatic difference between pepsin-treated homotrimers and pepsin-treated heterotrimers indicates that the α2(I) chain plays a major role both in kinetic and equilibrium segregation, even though telopeptides may inhibit the formation of non-fibrillar aggregates *in vivo*. Involvement of tissue collagenases in fibril remodeling may further exaggerate the segregation because α1(I) homotrimers are less susceptible to cleavage by these enzymes than the heterotrimers [26].

If this hypothesis is correct, even a small fraction of α1(I) homotrimers may have a large effect on tissues. For instance, accumulation of homotrimers with multiple bone remodeling cycles may exacerbate gradual degradation of bone quality with age in individuals with the overproducing allele of COL1A1, increasing osteoporotic fractures. Synthesis and accumulation of collagenases-resistant homotrimers inside kidney glomeruli may result in glomerular sclerosis contributing to renal failure [167]. Similarly, accumulation of collagenases-resistant homotrimer-

rich fibrils or subfibrillar regions may play a detrimental role in a broader spectrum of fibrotic tissues and disorders.

3.4. Materials and Methods

3.4.1. Native collagen

Pepsin-soluble collagen was purified from tail tendons of wild-type (normal heterotrimers) and homozygous *oim* mice (α1(I) homotrimers) as previously described [151]. Briefly, frozen mouse tails were thawed in ice-cold 3.5M NaCl, 10 mM Tris, 20 mM EDTA, 2 mM *N*-ethylmaleimide, and 1 mM phenylmethylsulfonyl fluoride (pH 7.5). Tendons were excised from tails, washed in the same enzyme-inhibiting buffer for several days at 4 °C, dissolved in 0.5 M acetic acid (pH 2.8), and digested by pepsin (100 mg /1 g collagen in two doses for 24 h at 4 °C each). After removal of insoluble material by centrifugation, collagen was precipitated overnight by 0.9 M NaCl, separated by centrifugation, redissolved in 0.5M acetic acid, and reprecipitated by 0.7M NaCl. The final precipitate was re-suspended and extensively dialyzed against excess 2 mM HCl. The resulting collagen was characterized by gel electrophoresis using the same procedure as described in *In Vitro* Fibrillogenesis section. The concentration of collagen was evaluated by circular dichroism (CD) on a J810 spectrometer (Jasco Inc.).

3.4.2. Reconstituted α1(I) homotrimers

Type I collagen from rat tails and fetal calf tendons (both from Pel-Freez Biologicals) was prepared by the same method. The pepsin-soluble collagen was transferred into 0.2M sodium phosphate, 0.5M glycerol (pH 7.4), diluted to 0.8 mg/ml, denatured at 55^{o}C for 10min, and refolded at 26^{o}C for 10 days. The CD spectra of solution aliquots at room temperature were measured to control the extent of denaturation and refolding. The refolded collagen was dialyzed into 0.5 M acetic acid at 4 ^{o}C, digested by pepsin (0.2 - 0.9 mg/ml) for 40 min at 32 ^{o}C followed by 2-3 hours at room temperature (to remove partially folded molecules and residual heterotrimers), and precipitated by 1 M NaCl overnight at 4 ^{o}C. The precipitate was resuspended in and dialyzed against 2 mM HCl, mixed 1:1 (volume ratio) with 0.26 M NaCl, 60 mM Na-phosphate, pH 7.4 (after mixing) and reprecipitated by overnight fibrillogenesis at 32 ^{o}C (to remove residual impurities and fibrillogenesis incapable molecules). Fibrillar collagen was separated by centrifugation, resuspended in and dialyzed against 2 mM HCl and characterized by gel electrophoresis, CD, and differential scanning calorimetry at 1 ^{o}C/min in N-DSC II (Calorimetry Sciences Corp.) [168].

3.4.3. In vitro fibrillogenesis

For kinetic measurements, a pepsin-treated collagen solution in 2 mM HCl was mixed 1:1 with 40 mM or 60 mM sodium phosphate, 0.26 M NaCl (pH 7.4 after mixing) at 4 °C. The solution was degassed, transferred into a pre-cooled at 4 °C quartz cuvette (1 cm optical path) and placed into a V-560 spectrophotometer (Jasco

Inc.) equipped with a thermoelectric temperature controller set at 4 °C. Fibrillogenesis was initiated by a temperature jump to 32 °C. Once the temperature reached 32 °C, the time evolution of the solution turbidity was monitored by measuring the optical density at 350 nm.

For solubility measurements, different mixtures of mouse homo- and heterotrimer in 2mM HCl were mixed 1:1 with 60mM sodium phosphate, 0.26M NaCl, 0.03% Brij 35 (pH 7.4 after mixing) and equilibrated for one day at 32 ^{o}C. The fibrils and surrounding solutions were separated by 30 min centrifugation at 20,000 G. The supernatants were collected with pipette tips rinsed with 0.03% Brij 35 into tubes also rinsed with 0.03 % Brij 35. Pellets resuspended in 2mM HCl and supernatants were mixed 4:1 with 2.5M NaCl, 0.5 M sodium carbonate, pH 9.3. Each sample was then rapidly mixed 5:1 with freshly prepared mono-reactive Cy5 NHS-ester (GE Healthcare #PA 35001) solution in anhydrous dimethylformamide (1 vial/ml) and immediately placed on a shaker for 30 min at room temperature. The same stock solution of the dye was used for all labeling reactions to ensure the same labeling efficiency. The accuracy of the labeling was tested by adding bovine serum albumin to some of the samples as an internal standard [166]. After 30 min labeling, the samples were mixed with 4x NuPAGE LDS Sample Buffer (Invitrogen), denatured at 60^{o}C for 5min, and the alpha chains were separated on pre-cast 3–8% Tris–acetate mini-gels (Invitrogen). The gels were scanned on a FLA5000 fluorescence scanner (Fuji Medical Systems) and analyzed by Multi Gauge 3.0 software supplied with the scanner.

3.4.4. Confocal imaging

Collagen solution in 2 mM HCl was mixed 4:1 with 2.5 M NaCl, 0.5 M sodium carbonate (pH 9.3) and fluorescently labeled by 2-10 μg/ml Alexa Fluor 488 or Alexa Fluor 568 carboxylic acid succinimidyl ester (Invitrogen). The amount of the dye was adjusted to result in 1 attached dye molecule per 4-5 collagen triple helices. After 30 min labeling at room temperature, the reaction was stopped and collagen was precipitated by adding acetic acid to 0.5 M and NaCl to 0.9 M. After overnight equilibration, precipitated collagen was separated by centrifugation, resuspended in 2 mM HCl and extensively dialyzed against 2 mM HCl to remove residual free dye molecules. To evaluate the labeling efficiency, the collagen concentration was determined by CD at 221 nm. The concentration of the conjugated dye was determined from visible light absorption in a Lambda 900 spectrometer (Perkin Elmer) as recommended by the dye manufacturer.

Binary mixtures of collagens labeled by Alexa Fluor 488 and Alexa Fluor 568 were mixed 1:1 with 60mM sodium phosphate, 0.26M NaCl (pH 7.4 after mixing), transferred into SecureSeal hybridization chambers (Grace Bio-Labs) and sealed. The cover glass used with the chambers was pretreated by Sigmacote (Sigma) to prevent attachment of collagen fibrils. The hybridization chambers were attached to microscope slides and preequilibrated at 32 ^{o}C or directly mounted on a microscope stage in a controlled environment chamber set at 32 ^{o}C. The fibrils were imaged in a Zeiss 510 Inverted Meta confocal, laser-scanning microscope (Carl Zeiss) with a 488 nm Argon ion and a 543 nm Helium-Neon lasers. The emission was collected by independent photomultipliers in two channels with 505-530 nm band pass and 560

nm long pass filters. The sample was illuminated by only one of the lasers at any given time, utilizing a line scanning mode with automatic laser switching. Virtually no Alexa Fluor 568 signal was detected in the Alexa Flour 488 channel and visa versa. The digital zoom, z-oversampling, scanning rate, and laser power were optimized to eliminate photobleaching.

3.4.5. Image processing and color separation analysis

A non-uniform background caused by interference and reflections was subtracted with Image J software (http://rsb.info.nih.gov/ij/). The residual noise was eliminated with an intensity threshold that did not affect colors within fibrils. Three-dimensional images were reconstructed with MIPAV software (http://mipav.cit.nih.gov/). For color separation analysis, both channels were normalized to the same total intensity within each 3D stack. The color separation ratio

$$\frac{I_R - I_G}{I_R + I_G}$$

was calculated for each pixel, where I_R and I_G were normalized intensities of the red and green channels, respectively. The color separation within fibrils was characterized by histograms of the relative frequency of $(I_R - I_G)/(I_R + I_G)$ from -1.0 to 1.0 with 0.1 step.

3.5. Supplementary material

The dependence of the equilibrium solubility of the homo- and heterotrimers on their fraction in fibrils can be deduced from the following thermodynamic arguments.

At equilibrium, the chemical potential $\mu_{i,s}$ of each molecule *i* (homo- or heterotrimers) in solution must be the same as the chemical potential $\mu_{i,f}$ of the same molecule in fibers,

$$\mu_{i,s} = \mu_{i,f} . \quad (1)$$

At fibrillogenesis conditions, the equilibrium concentrations c_i of the homo- and heterotrimers in solution are small enough for collagen-collagen interactions to be negligible. Their solutions are nearly ideal and $\mu_{i,s}$ can be approximated by

$$\mu_{i,s} \approx \mu_{i,s}^{0} + k_B T \ln\left(c_i\right) \quad (2)$$

where k_B is the Boltzman constant, T is the absolute temperature, and $\mu_{i,s}^{0}$ is the standard chemical potential.

In fibrils, the homo- and heterotrimers may also form an ideal binary mixture. Provided that the fraction ϕ_i of one of the components is sufficiently small, the chemical potential of this minor component *i* can be approximated by [e.g., Landau&Lifshitz]

$$\mu_{i,f} \approx \mu_{i,f}^{0} + k_B T \ln\left(\phi_i\right) \tag{3}$$

From Eqs. (1)-(3), we then find

$$c_i \approx \phi_i \exp\left(\frac{\mu_{i,f}^{0} - \mu_{i,s}^{0}}{k_B T}\right), \tag{4}$$

i.e., the equilibrium concentration (solubility) of the minor component in the surrounding solution is proportional to the fraction of this component in the aggregate.

Note, however, that Eq. (4) may be expected to work well only when interactions of the molecules *i* with each other in fibrils can be neglected, i.e., when the probability of these molecules being next to each other is negligible. Segregation within fibrils will always lead to a violation of this condition, resulting in significant deviations from the proportional dependence even at small ϕ_i. The lack of this proportionality at small concentrations of either component may thus be used as an indicator of segregation.

Chapter 4: Bending rigidity of reconstituted type I collagen hetero- and homotrimers

4.1. Introduction

Type I collagen is the most abundant protein in all vertebrates. It is a 300 nm long triple helix which self-assembles into elastic fibrils. Tightly packed fibril bundles form fibers that provide structural and mechanical scaffolds for tissues. For example, type I collagen fibers define the mechanical properties of tendons and ligaments and provide crucial contribution to the strength and elasticity of skin and bone.

The mechanical properties of collagen scaffolds are determined by the fiber network organization and by the mechanics of individual collagen fibrils [169, 170]. In particular, tendons are more rigid axially than laterally due to collagen fiber orientation parallel to the tendon axis [1]. In the jaw joint disk, collagen fibers are not aligned, but form a branched, connected network with a wide range of fiber orientations [28]. This organization improves energy dissipation, allowing collagen fibers to absorb impacts [28].

Numerous biomechanical measurement of different fibrous tissues had been reported [171, 172], but interpretation of molecular interactions responsible for the observed mechanical properties was difficult due to the complicated organization of tissues. Even tendons, which are essentially bundles of collagen fibrils, still contain several different types of collagen and proteoglycan molecules that may affect tendon mechanics.

To gain a better insight into the underlying physics, various simpler model studies of individual molecules and fibril networks reconstituted *in vitro* from purified type I collagen were also carried out, for instance, mechanical properties of individual collagen triple helices were studied in [118, 173-175]. Connectivity, elasticity, and rheology of type I collagen fibril networks were studied in [176-178]. However, less is known about the bending rigidity of collagen fibrils and its role in the mechanical and structural properties of fibril networks [179-181].

Here we compare bending of reconstituted fibrils composed of heterotrimeric and homotrimeric type I collagen. The normal isoform of type I collagen is a heterotrimer of two α1(I) and one α2(I) chains, but homotrimers of three α1(I) chains form in fetal tissues [3], carcinomas[6, 7], and some fibrotic tissues [10, 144-146]. We suggested that the appearance of the homotrimers in cancer might be related to resistance of the homotrimeric molecules to collagenases[26, 182, 183], which are massively secreted by invasive cancer cells and tumor-related fibroblasts [184]. However, we also observed enhanced proliferation and migration of cancer cells on matrix reconstituted from homotrimeric type I collagen. The latter observation was more surprising because homotrimers present the cells with the same α1(I) chains as heterotrimers do. Since cells are known to respond differently to substrates with different mechanical properties [185-187], we speculated that a difference in the rigidity between the homo- and heterotrimer fibrils may be the underlying cause for the observed differences in cell proliferation and migration, additionally motivating the present study.

Like normal heterotrimers, homotrimers form fibrils with the characteristic D-periodicity[156], but homotrimer fibers are more hydrated [22], have less regular lateral packing [22], and lower tensile strength [23]. Hence, we expected reconstituted homotrimer fibrils to have lower bending rigidity and to form entangled networks. Instead, examination by confocal microscopy revealed clusters of thin, straight, spear-like homotrimer fibrils vs. thicker yet curved heterotrimer fibrils forming entangled three dimensional (3D) networks [188]. It seemed that the homotrimer fibrils were intrinsically more rigid.

In the present study, we quantified these differences in the homo- and heterotrimer fibril shape. Visualization of fibril dynamics by time-lapse confocal microscopy suggested that fluctuations in the curvature of isolated fibrils were likely driven by thermal motions. Thus, we compared the observed fibril shapes with predictions for thermal fluctuations of long, thin filaments with constant bending rigidity (worm-like chains). As expected within the worm-like chain model [21], we obtained positive agreement between fibril persistence lengths extracted from the autocorrelation function [189] and normal mode analysis of shape fluctuations [190]. Despite somewhat smaller fibril diameter, the persistence lengths and bending rigidities of reconstituted homotrimer fibrils derived from this analysis were about two times larger than for the heterotrimers. In addition, analysis of the curvature by distribution functions suggested larger variations in the bending rigidity within an apparently more heterogeneous homotrimer fibril population.

4.2. Results

4.2.1. Fibril morphology

Figure 4.1 shows fluorescently labeled fibrils reconstituted from purified mouse-tail-tendon type I collagen homo- and heterotrimers by *in vitro* fibrillogenesis visualized through laser scanning confocal microscopy. The heterotrimers formed an entangled network of flexible, long fibrils (Fig 4.1, (a)), while the homotrimers formed clusters of thinner, straight, spear-like fibrils (Fig 4.1, (b)). Both samples were reconstituted from collagen solutions with the same concentrations at the same conditions. The lower density of the homotrimer fibrils resulted from higher equilibrium solubility of the homotrimer molecules, so that a smaller fraction of these molecules self-assembled into fibrils [188].

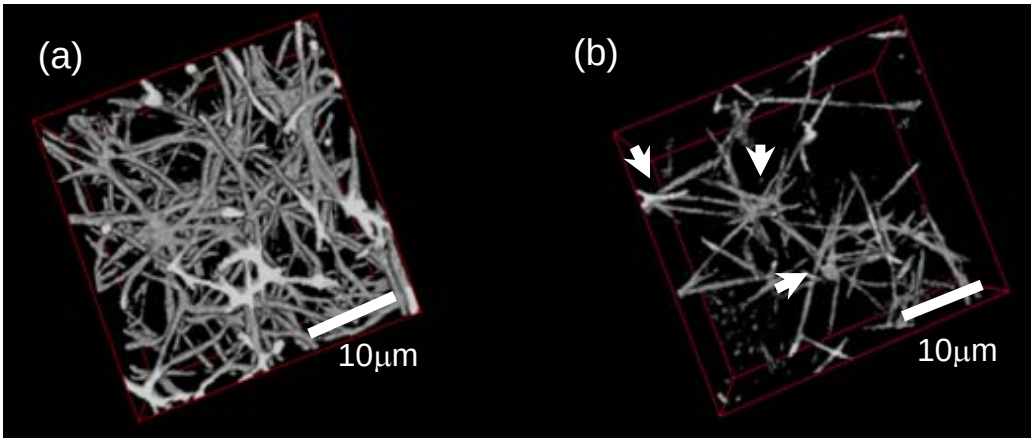

Fig 4.1. 3D reconstruction of confocal images of collagen fibril networks in 0.3 mg/ml of fluorescently labeled wild type heterotrimers (Het) and oim homotrimers (Homo) from mouse tail tendons. Spear-like homotrimer fibrils appeared to emanate from common centers shown by arrows.

4.2.2. Fibril growth, fluctuations, and formation

To gain a better insight into fibril network formation and dynamics, we captured time-lapse series of 3D images. Figure 4.2 shows a time sequence of confocal 3D projections from the same sample area, in which two representative fibrils are marked by arrows in the first snapshot. After fibril formation was initiated by a temperature jump from 4 to 32 ^{o}C, freely moving, short, isolated fibrils were observed. With time, fibrils elongated, grew in diameter, and exhibited large motions and shape fluctuations. However, once they formed an entangled network, their motions and shape fluctuations became limited and difficult to detect. These observations suggested that fluctuations in fibril shape were driven by thermal forces rather than defects in fibril formation. However, because of technical limitations of laser scanning confocal microscopy, the latter possibility could not be completely excluded.

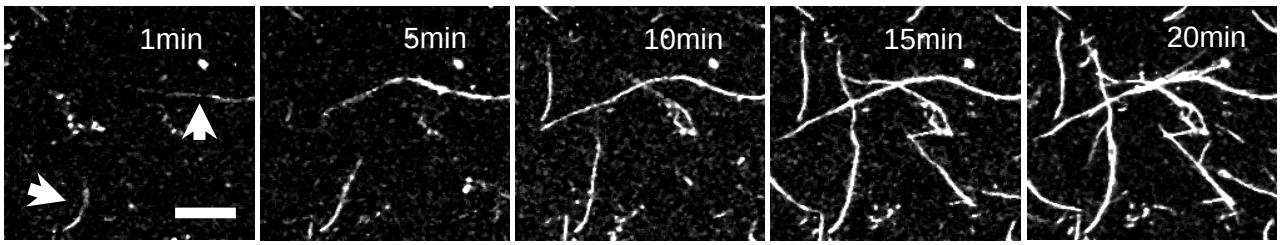

Fig 4.2. Projections of 3D confocal images illustrating growth and motions of heterotrimer fibrils after a temperature jump from 4 to 32 ^{o}C at 0 min; white bar: 10 μm

4.2.3. Fibril curvature fluctuations

The observed fluctuations in fibril curvature should decrease with increasing fibril rigidity regardless of their thermal or non-thermal origin. To characterize these fluctuations, we traced individual fibrils in 3D stacks and measured their persistence lengths by three different methods. Short and isolated fibrils, observed during very early fibrillogenesis, drifted too fast for reliable detection by 3D confocal imaging (rapidly moving and fluctuating fibrils could result in spurious contributions to the curvature). Therefore, we analyzed larger (> 15 μm) segments of individual fibrils between (or without) network connections after the fibril formation was complete. At this stage, significant changes in the fibril shape occurred at the time scale of tens of minutes or longer, while the scanning time for each 3D stack was ~1 min. We also looked a curvature in the xy plane vs. the yz or zx planes. The scanning speed is much faster in the xy plane than along the z direction. If the curvature resulted from an artifact of the actual fibril position moving between scans, we would expect higher curvature for fibrils aligned in the z direction than for fibrils in the x or y direction. However, no noticeable difference of curvature depending on the scanning direction was observed.

First, we determined an autocorrelation function

$$R_{\vec{T}}(s) = \left\langle \vec{T}(s')\vec{T}(s'+s) \right\rangle \tag{4.1}$$

for the tangential vector of unit length $\vec{T}(s)$ along the fibrils, where < > denotes statistical averaging over multiple filaments and time, and *s* is the distance along the

contour of the chain, measured from an arbitrary starting point s_0. At $s < 3\mu m$, the autocorrelation function (shown within circles in Fig. 4.3.) might be affected by data smoothing used in tracing fibril centerlines. At distances longer than the average length over which the fibrils remained in the field of view and could be clearly identified as single fibrils (> 20 μm for heterotrimer fibrils and > 30 μm for homotrimer fibrils), the results might be affected by network entanglement and visualization artifacts. Therefore, both regions were excluded from further analysis.

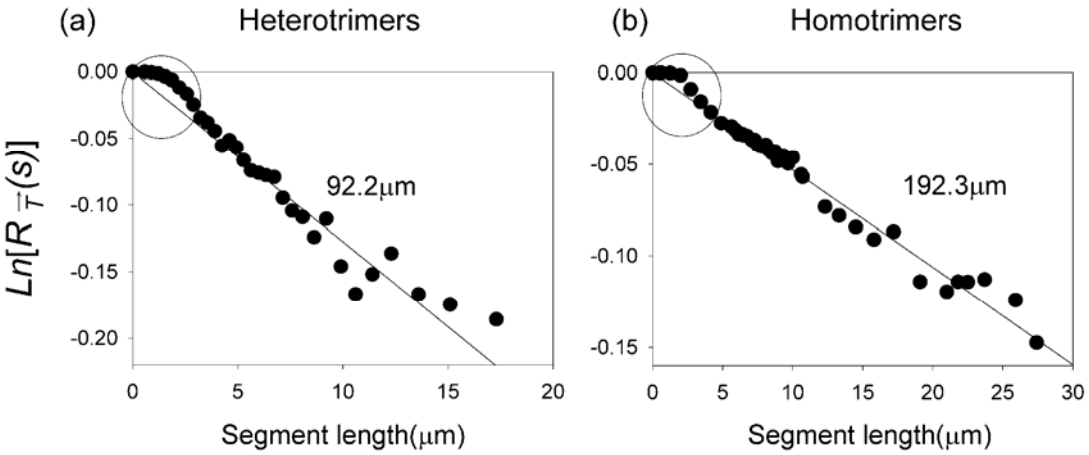

Fig 4.3. Estimation of the persistence length from the tangential vector autocorrelation function. (a) heterotrimer fibrils (b) homotrimer fibrils.

Within these limits, the observed autocorrelation function decayed exponentially with the characteristic length 92±4 μm for the heterotrimers and 192±4 μm for the homotrimers, as shown by the plots of $\ln\left[R_{\bar{T}}(s)\right]$ in Fig. 4.3. The decay length of this autocorrelation function for fluctuating chains is usually referred to as the chain persistence length L_p [189].

Next, we decomposed the shape of each individual collagen fibril into a sum of normal modes by Fourier transformation of the angle $\theta(s)$ of the tangential vector, defined in the plane of the fibril curvature at each point s (see Methods).

$$\theta(s) = \sqrt{\frac{2}{L}} \sum_{n=0}^{\infty} a_n \cos\left(\frac{n\pi s}{L}\right) \tag{4.2}$$

where L is the contour length of the fibril selected for the analysis (based on the same criteria as for $R_{\bar{T}}(s)$). For a worm-like chain with constant bending rigidity, the variance for thermal fluctuations in the amplitude of each normal mode is given by

$$\left\langle \left(a_n - \langle a_n \rangle\right)^2 \right\rangle = \frac{L^2}{n^2 \pi^2 L_p}. \tag{4.3}$$

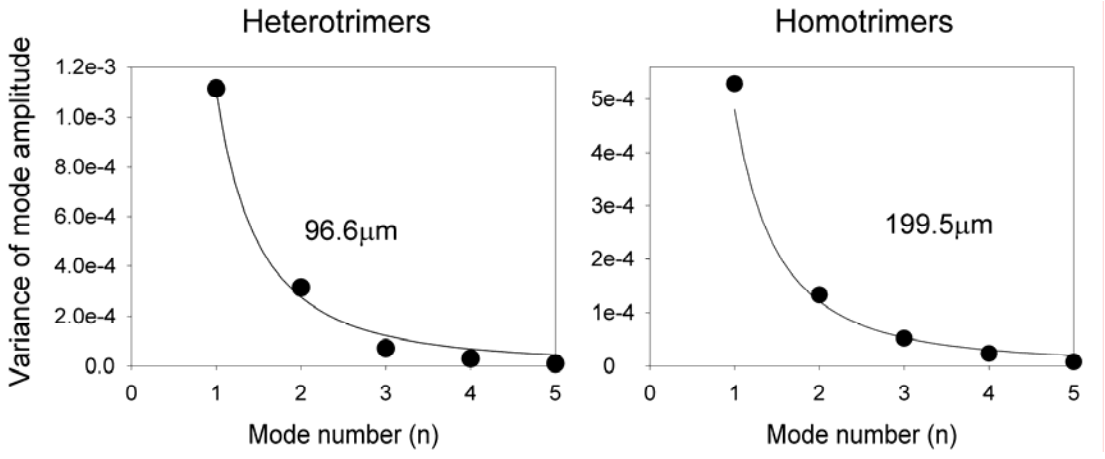

Fig 4.4. Normal mode analysis of fluctuations in the tangential vector angle. Fitted curves depict the expected $1/n^2$ behavior.

The observed variance was well fitted by this equation at L_p=97±5 μm for the heterotrimers and L_p=200±6 μm for the homotrimers (Fig. 4.4), in agreement with the persistence length values obtained from the tangential vector autocorrelation. Because of potential artifacts associated with fitting of fibril centerlines, modes with $n > 5$ were not included into the analysis (analogous to the exclusion of $s < 3$ μm from the autocorrelation function analysis).

Both measurements reported above produce an average persistence length for a potentially heterogeneous population of fibrils. To gain a better insight into the range of individual fibril persistence lengths in the populations of homotrimeric and heterotrimeric fibrils, we determined the distribution function for the local curvature along fibrils. Note that the local curvature probability in a fluctuating worm-like chain depends on the length scale at which it is measured (see Appendix and Eq. (4.4)). Thus, we defined the curvature C at each point s along fibril centerline as the curvature of a fibril segment of fixed contour length ΔS centered at this point. To obtain good statistics, we calculated the curvature distribution within over 200 individual fibrils longer than 10 μm (in each sample set).

The curvature distribution was characterized by histograms of the relative frequency $P(C)$ with 0.01 μm^{-1} bin size for the heterotrimers and 0.003 μm^{-1} bin size for the homotrimers (Fig 4.5). As expected from simple visual inspection of fibril shapes, the maximum of the curvature distribution histogram for the homotrimers was

shifted toward smaller curvatures compared to the heterotrimers. However, the homotrimer histogram appeared to have a longer tail at high curvatures.

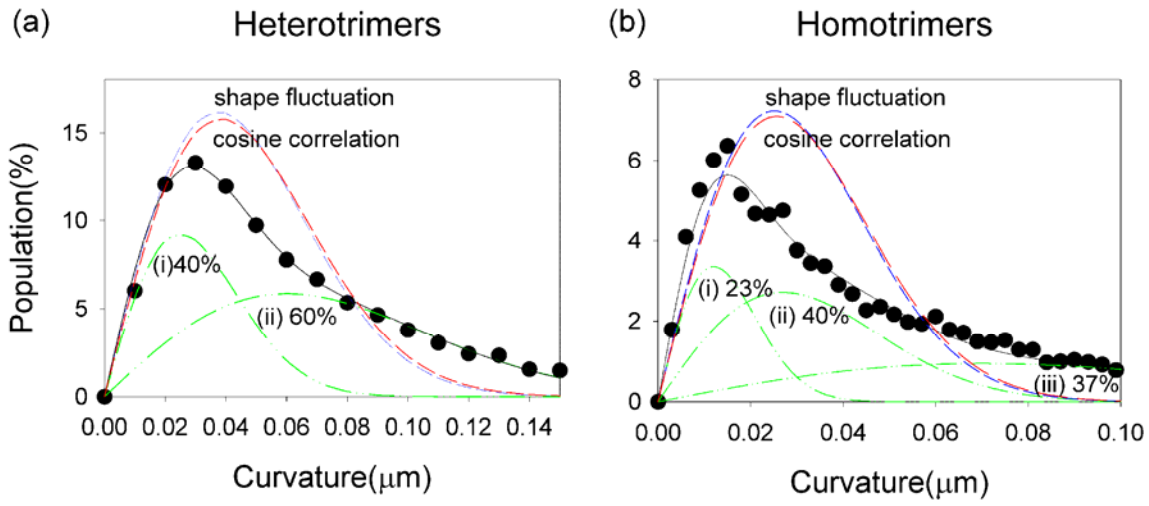

Fig 4.5. Local curvature distributions ($P(C)$ 100%) in (a) hetero- and (b) homotrimer fibrils. Blue and red dashed lines represent curves described by Eq. (4.4) with the persistence length calculated from the normal mode analysis and autocorrelation function for fibril shape fluctuations, respectively. Green dash-dotted lines (i, ii, and iii) represents deconvoluted subpopulations of fibrils with different persistence lengths. Solid black lines show the corresponding fits with several subpopulations of fibrils. At least three different subpopulations were required for the homotrimers.

For thermal fluctuations of a worm-like chain, the relative curvature frequency is described by (see Appendix)

$$P(C) = L_p \Delta S C e^{-(\frac{1}{2} L_p \Delta S C^2)}. \quad (4.4)$$

The corresponding functions calculated for the hetero- and homotrimers, with the L_p values deduced above, did not fit the observed histograms (red and blue curves in Fig. 4.5). In fact, the histograms could not be fitted with Eq. (4.4) at any single L_p. Long tails of the distributions suggested that the histograms were superpositions of several or multiple curvature distributions characterized by different values of L_p, as expected for heterogeneous populations of fibrils with different diameters.

Since a continuous distribution of L_p could not be uniquely deconvoluted from the observed histograms, we analyzed the histograms with a minimal discrete set of L_p values providing a reasonable fit. For heterotrimer fibrils, we found a good fit assuming ~40% subpopulation with $L_p \approx 207$ μm and ~60% with $L_p \approx 36$ μm (Fig. 4.5 (a)). For homotrimers, the apparent range of L_p variation was wider, requiring at least three subpopulations for a reasonable fit; ~23% with $L_p \approx 877$ μm, ~40% with $L_p \approx$ 173 μm, and ~37% with $L_p \approx 26$μm (Fig. 4.5 (b)). Average L_p based on these subpopulations (~99 μm for the heterotrimers and ~270 μm for the homotrimers) were somewhat larger but consistent with the average L_p deduced above. On a cautionary note, these fits indicate an approximate range for L_p variation within fibril populations rather than actual discrete subpopulations of heterotrimer or homotrimer fibrils.

4.3. Discussions

Thus, interpretation of the observed fibril shape fluctuations as thermal motions of a worm-like chain results in the persistence length L_p from 35 to 200 μm (average $L_p \sim 100$ μm) for reconstituted heterotrimer fibrils and from 25 to 900 μm (average $L_p \sim 200$ μm) for thinner homotrimer fibrils. To compare these values with prior measurements of heterotrimer fibril mechanics, consider two limiting cases for fibril bending.

(i) During fibril bending, individual molecules may freely slip along each other. If that is the case, the persistence length of a fibril L_p is simply proportional to the number of molecules N in the fibril cross-section; $L_p = N\, l_p$, where l_p is the persistence length of an individual molecule. Assuming that each molecule in reconstituted fibrils may be modeled as an elastic rod with the Young modulus $E \approx$ 400-800 MPa [191-193]and the moment of inertia $I = \pi r^4/4$ [194], where $r \approx 0.6$ nm, we find $l_p = EI/kT =$ 10-20 nm, consistent with 15 nm reported in single molecule measurements [173, 174, 195]. For reconstituted fibrils with 50 to 300 nm diameter[94, 196], containing 10^3 to 10^4 molecules in cross-section, we then find the expected range of L_p from 10 to 700 μm, in a good agreement with 35 to 200 μm we observed for heterotrimeric type I collagen fibrils (Fig. 4.5 (a)).

(ii) In the opposite limit of no slippage, the whole fibril may be viewed as an elastic rod with the moment of inertia $I = \pi R^4/4$, where R is the fibril radius. For fibrils with 50 to 300 nm diameter with 400 to 800 MPa Young modulus, we then expect L_p from 3 cm to 80 m, clearly inconsistent with our observations. While we cannot exclude a contribution of fibril defects (resulting in intrinsic fibril curvature)

to the observed persistent length, the observed range of thermal fibril motions was at least a couple of orders of magnitude larger than expected for 3 cm to 80 m thermal persistence length. Unless bending involves drastic changes in fibril profile, e.g., as discussed in [117], such large discrepancy is difficult to explain.

In other words, the most consistent interpretation of our observations for fibrils reconstituted from type I collagen heterotrimers appears to be as follows. The consistent fibril persistence lengths (Figs. 4.3-4.5) estimated by three different methods are associated with thermal fluctuations, in agreement with visual observations of fibril dynamics at the early stages of fibrillogenesis (Fig. 4.2). The low bending modulus indicates that fibril bending involves free slippage of molecules along each other so that the bending rigidity of a fibril is simply a sum of bending rigidities of individual molecules in the fibril cross-section. Crosslinked molecules would provide a significantly larger rigidity.

The tendency of collagen to forming highly regular, D-periodic alignment suggests that the slippage of adjacent molecules along each other is likely to involve significant energy barriers and occur very slow. Slow slippage may explain the nearly frozen network appearance in time lapse confocal images. One might expect much faster motions without slippage, but fibril resistance to such motions should be much higher, dramatically reducing their amplitude.

Note that the free slippage of the molecules along each other in our fibrils could possibly be due to the pepsin treatment we utilized during collagen purification. Such a treatment partially degrades non-helical telopeptides. While this treatment does not affect the axial and lateral packing or interactions between the molecules in fibrils

[151, 163], it prevents formation of intermolecular covalent cross-links, which would inhibit slippage. Consistently, fibrils reconstituted from collagen with full-length telopeptides always appear to be much more straight (data not shown) [60, 156], indicating significantly higher bending rigidity.

Although the observed fibril shapes appear to be consistent with thermal fluctuations, other factors may also contribute. In particular, network connections and entanglement may limit possible fibril conformations, violating the assumptions implicit in our analysis. Defects in collagen packing within fibrils may result in intrinsic (static) curvature of fibrils that does not reflect thermal motions. However, these factors should affect the dependence of curvature fluctuations on the wave number, e.g., network entrapment should primarily affect long wave length fluctuations (at distance scales comparable or larger than the characteristic distance between network connections). More precise evaluation of these contributions would require additional measurements, e.g, of the persistence length of fibrils with free ends and time evolution of the apparent persistence length during network formation. These measurements were beyond the scope of our study, which was focused on a more qualitative evaluation of bending rigidities and on contrasting properties of homo- and heterotrimers.

Because intermolecular interactions between homotrimers in fibrils are weaker than between heterotrimers [151], the slippage of homotrimers should occur more readily than that of heterotrimers. Therefore, the most likely explanation for longer persistence lengths of reconstituted homotrimer fibrils is larger bending rigidity of individual homotrimer molecules. Note that the homotrimer fibrils were thinner (Fig.

4.1). Thus, two times longer average persistence length of the homotrimer fibrils indicates an even larger difference in the persistence length of individual molecules. Since stretching and bending of molecules may involve transient, local unwinding (unfolding) of triple helices [18, 197], higher stability of the homotrimeric triple helix [18, 155] may be the source of higher bending rigidity.

Regardless of its molecular interpretation, higher bending rigidity of homotrimer fibrils may have very important implications for cancer and other disorders. For instance, cancer cells are known to proliferate and migrate faster on stiffer substrates [185, 198, 199]. They secrete type I collagen, of which 20-40% is the homotrimeric isoform [183]. Apparently, the homotrimers may promote cancer invasion not only because they resist cleavage by collagenases secreted by cancer cells [182, 183][see also Chapter 5 of this thesis], but also because they form a stiffer matrix. Indeed, our preliminary measurements suggested faster proliferation and migration of cancer cells on matrix reconstituted from the homotrimers [183].

4.4. Materials and Methods

4.4.1. Collagen preparation and purification

Pepsin-soluble collagen was purified from tail tendons of wild-type (normal heterotrimers) and homozygous *oim* mice (α1(I) homotrimers) as previously described [151]. Briefly, frozen mouse tails were thawed in ice-cold 3.5M NaCl, 10 mM Tris, 20 mM EDTA, 2 mM *N*-ethylmaleimide, and 1 mM phenylmethylsulfonyl fluoride (pH 7.5). Tendons were excised from tails, washed in the same enzyme-inhibiting buffer for several days at 4 °C, dissolved in 0.5 M acetic acid (pH 2.8), and

digested by pepsin (100 mg /1 g collagen in two doses for 24 h at 4 °C each). After removal of insoluble material by centrifugation, collagen was precipitated overnight by 0.9 M NaCl and subsequently separated by centrifugation, redissolved in 0.5 M acetic acid, and reprecipitated by 0.7 M NaCl. The final precipitate was re-suspended and extensively dialyzed against excess 2 mM HCl. The resulting collagen was characterized by gel electrophoresis using the same procedure as described in the *In Vitro* Fibrillogenesis section of chapter 3. The concentration of collagen was evaluated by circular dichroism (CD) on a J810 spectrometer (Jasco Inc.).

4.4.2. In vitro fibrillogenesis and confocal imaging

Collagen solution in 2 mM HCl was mixed 4:1 with 2.5 M NaCl, 0.5 M sodium carbonate (pH 9.3) and fluorescently labeled by 2-10 μg/ml Alexa Fluor 488, 546, 568, or 647 carboxylic acid succinimidyl ester (Invitrogen). The amount of the dye was adjusted to result in 1 attached dye molecule per 4-5 collagen triple helices. After 30 minute labeling at room temperature, the reaction was stopped and collagen was precipitated by adding acetic acid to 0.5 M and NaCl to 0.9 M. After overnight equilibration, precipitated collagen was separated by centrifugation, resuspended in 2 mM HCl and extensively dialyzed against 2 mM HCl to remove residual free dye molecules. To evaluate the labeling efficiency, the collagen concentration was determined by CD at 221 nm. The concentration of the conjugated dye was determined from visible light absorption in a Lambda 900 spectrometer (Perkin Elmer) as recommended by the dye manufacturer.

Solutions of Alexa Fluor labeled homo- or heterotrimers in 2 mM HCl (0.4-0.6 mg/ml) were mixed 1:1 with 60mM sodium phosphate, 0.26M NaCl (pH 7.4 after mixing), transferred into SecureSeal hybridization chambers (Grace Bio-Labs) and sealed. The cover glass used with the chambers was pretreated by Sigmacote (Sigma) or bovine serum albumin to prevent attachment of collagen fibrils. The hybridization chambers were attached to microscope slides and preequilibrated at 32 ^{o}C or directly mounted on a microscope stage in a controlled environment chamber set at 32 ^{o}C. The fibrils were imaged in a Zeiss 510 Inverted Meta confocal, laser-scanning microscope (Carl Zeiss) with a 488 nm Argon ion laser and 543 nm and 647 nm Helium-Neon lasers. The emissions were collected by independent photomultipliers in two channels, with band pass and long pass filters optimized for the selected laser lines. The sample was illuminated by only one of the lasers at any given time, utilizing a line scanning mode with automatic laser switching. The digital zoom, z-oversampling, scanning rate, and laser power were optimized to eliminate photobleaching.

4.4.3. Image processing and analysis

In order to characterize the shape of each fibril, we separated fibril segments from the network by a particle tracking algorithm written by a colleague in our laboratory. This algorithm tracks fibrils vertically oriented to the analysis plane. Since most fibrils tended to align mostly parallel to the scanning plane (xy plane), we rotated 3D image stacks about the y- or x-axis, as illustrated in Fig 4.6.(a),(b). In re-sliced, rotated stacks most fibrils appeared as dots in each plane of the stack (Fig 4.6.(c)). At the next step, we reduced noise by applying an intensity threshold (d).

Each fibril segment was then numbered and tracked along a vertical direction in a 3D stack (e). A set of centroid points along each fibril was extracted and used for further analysis as the coordinates of fibril centerline (f).

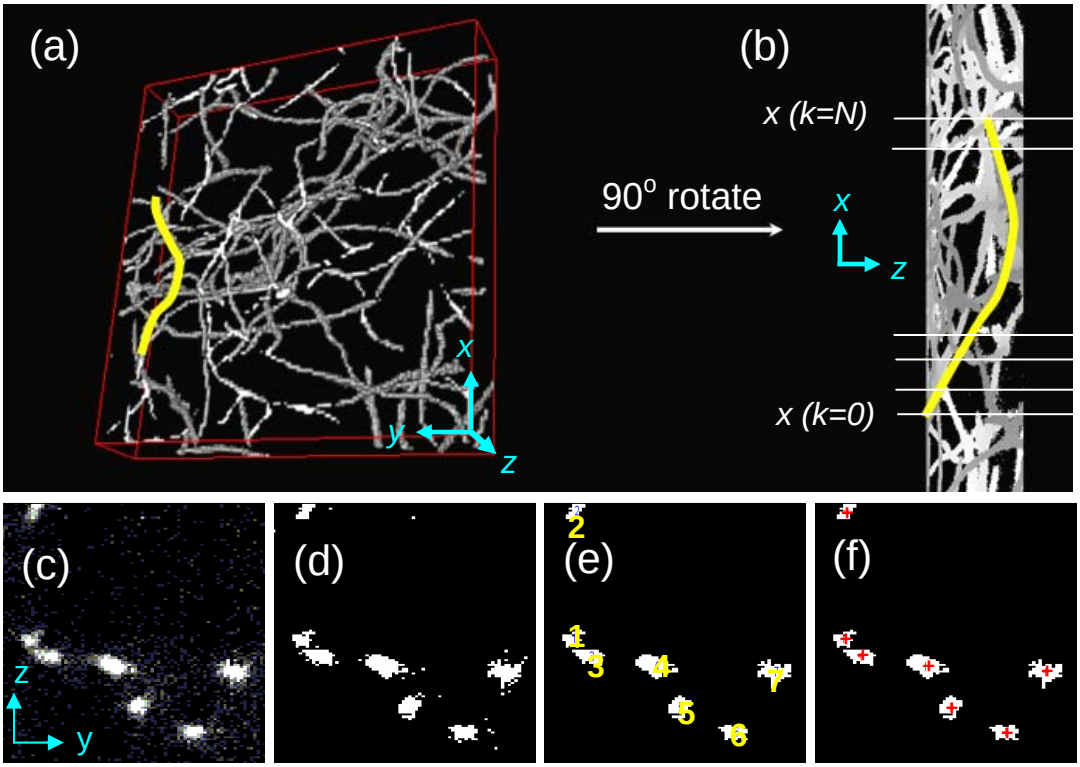

Fig 4.6. Image processing steps. (a) 3D rendered image of heterotrimer fibrils. Fibril segment which is tracked is colored with yellow. (b) Image stack rotation about the x-axis; (c) re-slicing of the rotated image stack; (d) noise removal; (e) fibril labeling and tracking; (f) extraction of centroid coordinates.

The extracted centroid points were not spaced equally and they exhibited residual noise not eliminated by threshold application. Therefore, we generated a smooth centerline for each fibril by fitting contiguous, overlapping stretches of centroid

points with a 2nd order polynomial. Specifically, the fitting was performed at each centroid point with 9 other centroid nodes included on each side of the fitting point (19 nodes total in each fit). Fibril ends containing the first and the last 9 centroid nodes were discarded. The optimal number of nodes for the fitting was selected to provide robust and consistent results for the curvature. Reaching the optimal number is important because if the number is too small, imaging noise may result in artificially high, rapidly fluctuating curvatures. When the number is too large, the fitting may artificially straighten fibrils, underestimating real curvature fluctuations.

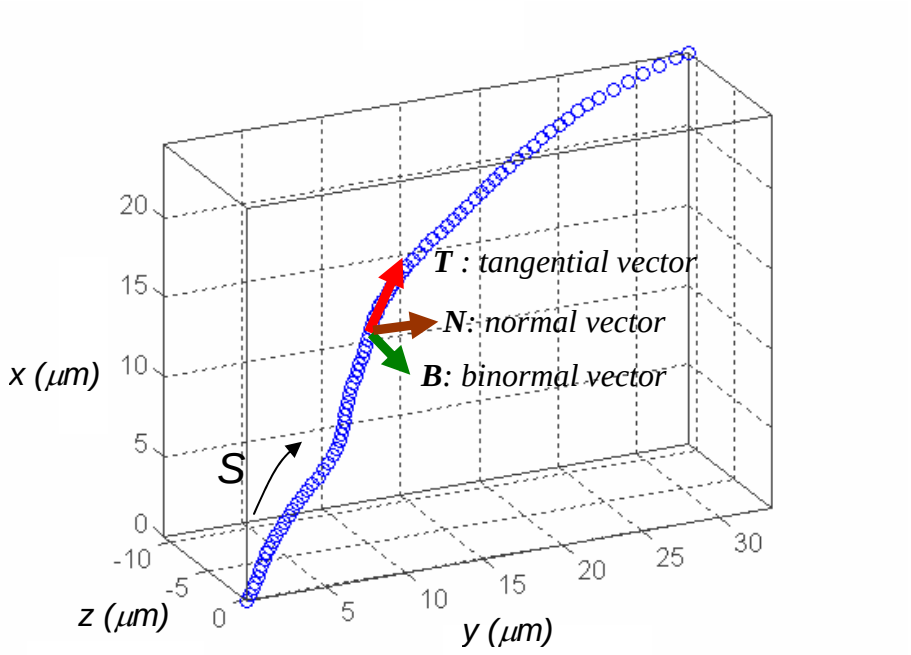

Fig 4.7. Fibril centerline extracted from a 3D confocal stack. Blue circles represent centroid points extracted as illustrated in Fig. 4.6. $\vec{T}$ is a tangent vector, pointing in the fibril direction. $\vec{N}$ is the derivative of $\vec{T}$ with respect to the centerline contour length coordinate *s*.

$\vec{T}$ and $\vec{N}$ vectors form the plane of fibril curvature at each at each point *s*. $\vec{B}$ is the cross product of $\vec{T}$ and $\vec{N}$.

From the fitted centerline, we generated a new set of equally spaced centerline points $\{X_k,Y_k,Z_k\}$ with the running index *k* increasing sequentially from 0 at one end to *N* at the other end; and we defined the tangential vectors of unit length $\vec{T}_k$ ($\left|\vec{T}_k\right|=1$) at each point (Fig. 4.7). We then used these sets of centerline coordinates and tangential vectors for calculating the tangential vector autocorrelation and curvature distribution as well as for analysis of normal modes of curvature fluctuations.

Specifically, we defined s_0=0 and calculated the contour length s_k at k>0 as

$$s_k \approx \sum_{j=1}^{k} \sqrt{\left(X_k - X_{k-1}\right)^2 + \left(Y_k - Y_{k-1}\right)^2 + \left(Z_k - Z_{k-1}\right)^2}\,. \tag{4.5}$$

We calculated the tangential vector autocorrelation function from Eq. (4.1), in which s_i and s_j were used in place of *s'* and *s'*+*s*, respectively (s=s_j-s_i). We defined $\theta(s) \equiv \theta(s_k)$ for the normal mode analysis as $\theta(s_0)$=0 at k=0, and

$$\theta(s_k) = \sum_{i=1}^{k} \cos^{-1}\left(\vec{T}_{k-1} \cdot \vec{T}_k\right). \tag{4.6}$$

at k>0. The amplitudes of the normal modes for $\theta(s)$ fluctuations are then given by

$$a_n = \sqrt{\frac{2}{s_N}} \int_0^{s_N} \theta(s) \cos\left(\frac{n\pi s}{s_N}\right) ds$$
$$\approx \sqrt{\frac{2}{s_N}} \sum_{k=1}^{N} \frac{1}{2} \left[\theta(s_{k-1}) \cos\left(\frac{n\pi s_{k-1}}{s_N}\right) + \theta(s_k) \cos\left(\frac{n\pi s_k}{s_N}\right) \right] (s_k - s_{k-1}) \quad . \tag{4.7}$$

Finally, we calculated the curvature at each point as

$$C(s_k) = \frac{d\theta(s_k)}{ds_k} \approx \frac{\cos^{-1}\left(\vec{T}_{k-1} \cdot \vec{T}_k\right)}{s_k - s_{k-1}} . \tag{4.8}$$

Note that for two dimensional (2D) curves, the angle $\theta(s)$ is simply an angle between the tangential vectors $\vec{T}(s)$ and $\vec{T}(0)$. In 3D, the meaning of this angle is more complicated. Mathematically, it may be defined by Eq. (4.6) or as

$$\theta(s) = \int_0^s C(s') ds' \tag{4.9}$$

Geometrically, it may be understood as an angle between $\vec{T}(s)$ and a projection of $\vec{T}(0)$ onto the plane of fibril curvature at the point *s* (which is the plane formed by the vectors $\vec{T}(s)$ and $\vec{N}(s)$ (Fig. 4.7)).

4.5. Appendix: Worm-like chain curvature distribution function

Consider a worm-like chain without an intrinsic curvature. The bending energy of such a chain is described by

$$E_C = \frac{kT}{2}\int L_p\left[C(s)\right]^2 ds \tag{4.10}$$

where L_p is the persistence length of the chain.

The probability of the mean curvature of a chain segment with a contour length ΔS to be between C and $C+dC$ is given by [189].

$$p(C) \equiv P(C)\,dC = \frac{\exp\left(-\frac{W_{\min}(C)}{kT}\right)CdC}{\int_0^\infty \exp\left(-\frac{W_{\min}(C)}{kT}\right)CdC} \tag{4.11}$$

where $P(C)$ is the curvature distribution function and $W_{\min}(C)$ is the minimal work required of achieving the mean curvature C from a minimum energy state of the chain.

From Eq. (4.10), we find that the minimal work of generating the mean curvature C in a uniform chain (constant L_p) is given by

$$W_{\min}(C) = \frac{kT}{2} L_p \Delta S C^2, \tag{4.12}$$

since the work is minimal when the chain curvature is constant. After substitution of Eq. (4.12) into Eq. (4.11), we then arrive at Eq. (4.4).

Note that the proportionality of the minimal work to the length of the segment imposes an important restriction on using the worm-like chain model. Indeed, the minimal work of generating a curvature C in a single point (infinitely small segment) within this model is zero and the curvature probability distribution becomes meaningless. This is an artifact of the model, which can be used only at length scales larger than some minimal length and at curvatures smaller than some maximum curvature. Outside these boundaries, Eq. (4.10) breaks down. This minimal length and the maximum curvature are defined by the physics and molecular structure of the actual object that is being modeled. For instance, in applying the model to a collagen fibril, we must take into account that Eq. (4.10) may break down at curvatures comparable to the fibril radius R, so that the model can be used only at $\Delta S >> R^2/L_p$.

Chapter 5: Molecular mechanism of the resistance of a homotrimeric type I collagen isoform to mammalian collagenases

The normal isoform of type I collagen is a heterotrimeric triple helix formed by two α1(I) and one α2(I) chains, but homotrimers with three α1(I) chains have been reported, e.g., in fetal tissues, carcinomas and fibrosis. In a separate study, we found the $\alpha1(I)_3$-homotrimers to be resistant to all major mammalian collagenases (MMP-1,2,8,13, and 14). Homotrimers were secreted by cancer cells but not by normal fibroblasts. Since MMP production is elevated in fibrosis and cancer, the synthesis of MMP-resistant collagen isoform in these tissues may be physiologically important. Moreover, distinct physical and chemical properties of the homotrimers may make them a potential diagnostic and treatment target. Nevertheless, the mechanism and physiological role of the homotrimer resistance to MMPs have remained unclear. In this chapter, we focused on the homotrimer interaction with MMP-1. We found that binding of MMP-1 to heterotrimeric and homotrimeric isoforms of type I collagen was identical. The cleavage rate of unfolded α1(I) and α2(I) chains by MMP-1 was similar as well. Instead, the homotrimer resistance to MMP-1 was caused by less efficient triple helix unwinding at the cleavage site. Increased triple helix stability at this site may explain the universal resistance of the homotrimers to different MMPs with collagenolytic activity. We also observed segregation and selective degradation of heterotrimer fibrils in reconstituted, mixed matrix, indicating that the homotrimer resistance to MMPs may have significant consequences for tissue remodeling.

1.1. Introduction

Type I collagen is the most abundant vertebrate protein. The mature protein is a triple helix flanked by short terminal peptides. It is the main component of collagen fibrils, forming the structural scaffold of many tissues. The most common isoform of type I collagen is a heterotrimer of two α1(I) chains and one α2(I) chain. However, a homotrimer of three α1(I) chains has also been found in fetal tissues[3] and various pathological conditions, e.g., fibrosis [10, 144-146] and cancer [6, 7].

Initial studies of triple helices refolded *in vitro* from isolated α1(I) chains suggested that reconstituted α1(I) homotrimers may be resistant to fibroblast (MMP-1) and neutrophil (MMP-8) collagenases [26]. We found that naturally produced human and mouse α1(I) homotrimers were also over 5-10 times more resistant to these collagenases as well as to collagenase-3 (MMP-13), gelatinase (MMP-2), and membrane collagenases (most importantly MMP-14 also referred to as MT1-MMP)[183].

Collagenases are enzymes of the matrix metalloproteinase (MMP) family, which are believed to be primarily responsible for extracellular and pericellular cleavage of type I collagen triple helix. The cleavage occurs at a specific site between Gly-775 and Ilu-776 in the α1(I) chain and between Gly-775 and Leu-776 in the α2(I) chain, approximately three quarters away from the N-terminus [200]. (We follow the convention in which the helical residues are numbered starting from the first Gly in the triple helical domain). The cleavage is a stepwise process involving MMP binding, local triple helix unwinding at the cleavage site, binding of a locally

unfolded chain inside the catalytic cleft in the catalytic domain of the enzyme, and peptide bond hydrolysis[131].

While much has been learned about the molecular mechanism of collagen cleavage by MMPs [131, 139, 201, 202], it has remained unclear which of the cleavage steps might be affected in α1(I) homotrimers and why. The universal resistance of the homotrimers to all collagenolytic MMPs is particularly puzzling since each MMPs interacts with collagen in a distinct manner. For instance, MMP-2 is believed to bind and unwind collagen in a different manner from MMP-1[135].

In the present study, we focus on understanding the molecular mechanism behind the dramatically different susceptibility of the two isoforms of type I collagen to collagenases and on potential effects of this difference on tissue remodeling and disease. Understanding this mechanism may be important not only from the fundamental knowledge perspective, but also for practical applications, e.g., utilization of α1(I) homotrimers as diagnostic and/or therapeutic targets.

Indeed, fibrotic and cancer tissues are characterized by increased expression of collagenases, but the homotrimers do not appear to be simple byproducts of collagen synthesis and selective heterotrimer degradation. Our observations indicate that α1(I) homotrimers are not produced in normal, non-fetal tissues. Although a small fraction of the homotrimeric isoform was reported in normal human skin[142], we were not able to confirm this observation. We tested skin biopsy from several patients as well as collagen secreted in cell culture by several lines of normal dermal fibroblasts and found no evidence of the homotrimers by using an assay[183] with better than 1% detection limit. We believe that the homotrimers may be physiologically important

and selectively produced in fibrosis, and cancer. For instance, analysis of homotrimer synthesis in human cancer cell cultures and xenografts of these cells in mice suggests that tumor homotrimers are made selectively by cancer cells rather than fibroblasts recruited into the tumors[183]. Cancer cells may utilize homotrimer fibrils as tracks for invading and degrading surrounding tissue, taking advantage of their collagenase resistance and ability of the homotrimer fibrils to promote faster cell proliferation and migration[183].

To understand the mechanism of α1(I) homotrimer resistance to collagenases, in the present study we focused on the interaction of the homotrimers with MMP-1, which is one of the most common and best studied collagenases. Because of the previously described preferential interaction of MMP-1 with the α2(I) chain[131], we expected the MMP-1 resistance of α1(I) homotrimers to be associated with reduced enzyme binding, but our measurements revealed a different answer. The data clearly pointed to the resistance of α1(I) homotrimers to local unwinding at the MMP cleavage site. To better understand the results, we analyzed different kinetic models of collagen cleavage, as described in the supplementary online material. This analysis confirmed the qualitative interpretation of the data and provided further evidence for the crucial role of the unwinding step in the cleavage of normal and mutant collagens in various disorders.

5.2.Results

5.2.1. Cleavage kinetics

To compare the cleavage kinetics of type I collagen homo- and heterotrimers we utilized binary mixtures in which the homo- and heterotrimers were labeled by different fluorescent dyes, mixed together and co-processed with activated MMP-1 in the same sample tube at different temperatures. Fig 5.1. (a) illustrates gel electrophoresis patterns of sample aliquots collected at different times after the addition of MMP-1. The homotrimer and heterotrimer patterns are images of the same area of one gel in different fluorescence colors.

Comparison of the electrophoretic patterns revealed the same ¾ and ¼-length fragments after cleavage of the homotrimers and heterotrimers. The initial cleavage rate V_0 (Fig. 1b) was ~17 times slower for the homotrimers at 25 ^{o}C and ~ 6 times slower at 35 ^{o}C (Fig. 5.1 c). The smaller difference at higher temperatures indicates higher apparent cleavage activation energy for the homotrimers. Aside from the slightly lower overall enzyme efficiency, the results were essentially the same for the homo- and heterotrimer cleavage by MMP-13 (Suppl. Fig. 5.7).

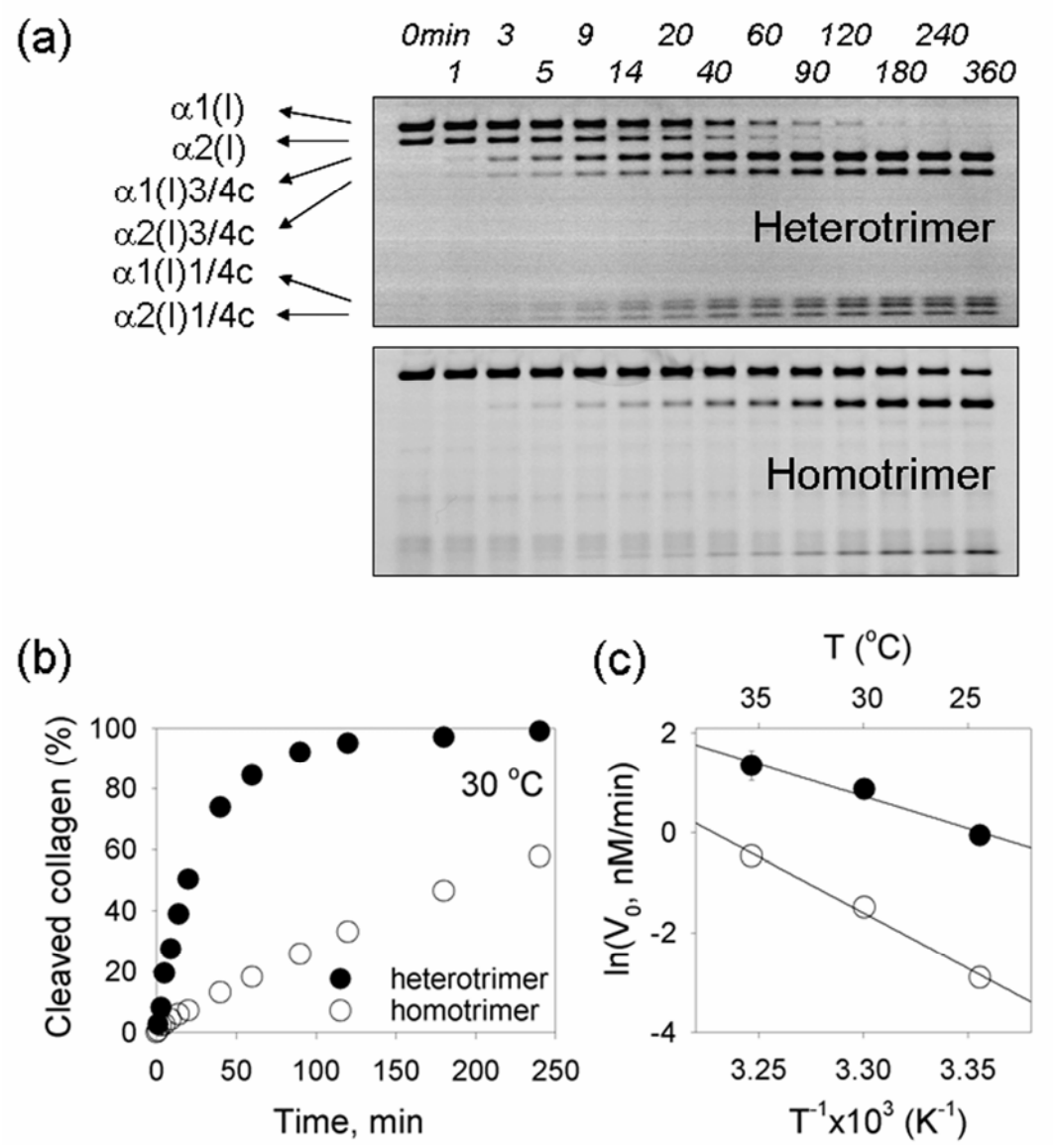

Fig. 5.1. Cleavage kinetics of human type I collagen homo- and heterotrimers in a 1:1 mixture (50 μg/ml) by recombinant human (rh) MMP-1 (2 nM). The homo- and heterotrimers were labeled with Cy5 and AlexaFluor 488, respectively. Inverse labeling produced the same results. (a) Gel electrophoresis of sample aliquots at different reaction times. The same gel area is shown in AlexaFluor 488 fluorescence (top) and Cy5 fluorescence (bottom). α1(I)3/4 and α1(I)1/4c are cleaved fragments at Gly(775)-Ile(776) of α1(I) chain. α2(I)3/4 and α2(I)1/4c are cleaved fragments at Gly(775)-Leu(776) of α2(I) chain. (b) The cleaved

fraction calculated from the intensities of the α1(I) band and its cleavage fragments. The initial cleavage rate V_0 was determined from the slope of the cleavage curve at zero time. (c) Temperature dependence of V_0 for rh-MMP-1. The lines show Arrhenius plots with the apparent activation energies of 44 (homo-) and 26 kcal/mol (heterotrimers).

The dependence of the initial cleavage rate on collagen concentration was consistent with the simple Michaelis-Menten kinetics (Fig. 5.2.).

$$V_0 = \frac{V_{\max}[C]}{K_m + [C]} \qquad (1)$$

where $V_{\max}$ is the maximum cleavage rate, [C] is the total collagen (substrate) concentration in the sample, and K_m is the Michaelis constant.

The apparent K_m= 0.8±0.08 μM for heterotrimeric human fibroblast collagen was the same as previously determined by Welgus et. al. [203]. To our surprise, however, the apparent K_m=0.9±0.2 μM for human homotrimers was also the same; the slower cleavage of the homotrimers was entirely attributable to the difference in $V_{\max}$. (We later found that similar K_m value was previously reported for the homotrimers artificially refolded from purified α1(I) chains [26]).

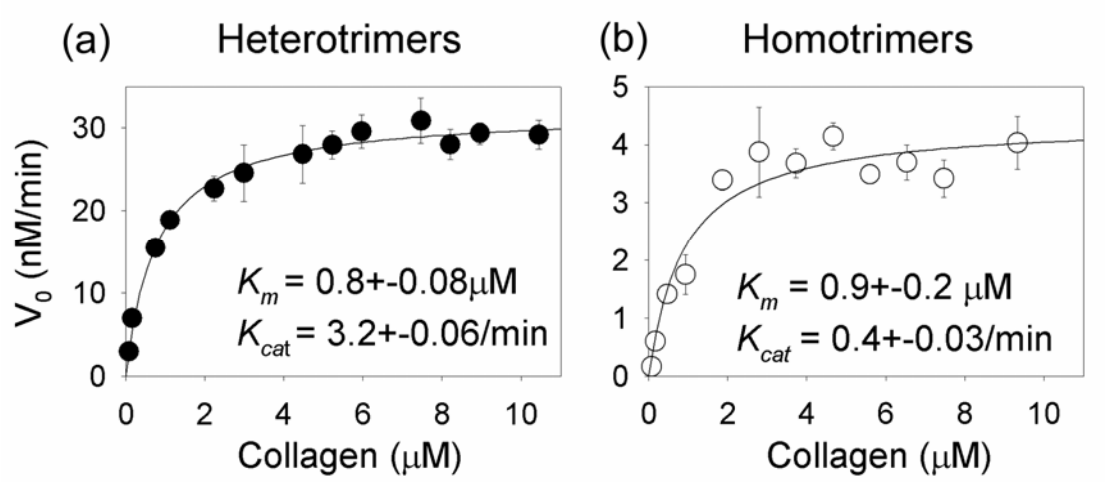

Fig. 5.2. Dependence of the initial cleavage rate by rh-MMP-1 on human collagen concentration at 25 °C. Solid lines represent fitting of the experimental data to Michaelis-Menton equation (Eq. (1)). The catalytic constant $k_{cat}=V_{max}/[\text{MMP-1}]$ was calculated based on the nominal enzyme concentration ([MMP-1]) in the solution, although we cannot exclude that only a fraction of enzyme molecules were active. To inhibit fibrillogenesis at high collagen concentrations, particularly at 35 °C, the reaction buffer was supplemented with 1 M glycerol, which had no effect on the cleavage rate.

5.2.2. Enzyme binding

Since K_m is usually close to the enzyme dissociation constant [50], the kinetic data indicated that MMP-1 binding to homo- and heterotrimers might be similar. This observation was contrary to our expectation based on the previous report [131] of a preferential interaction of MMP-1 with the α2(I) chain. Therefore, we measured the equilibrium dissociation constant by utilizing a catalytically disabled MMP-1(E200A)

[131] with a Glu-200 → Ala substitution at the catalytic site.

Because of the limited supply of human fibroblast α1(I) homotrimers, we performed the binding and other subsequent measurements with mouse tail-tendon collagen. The cleavage rate of mouse collagen by recombinant human MMP-1 was slightly slower than for human collagen. However, the relative cleavage rate of human homotrimers vs. human heterotrimers was identical to the relative cleavage rate of mouse homotrimers vs. mouse heterotrimers.

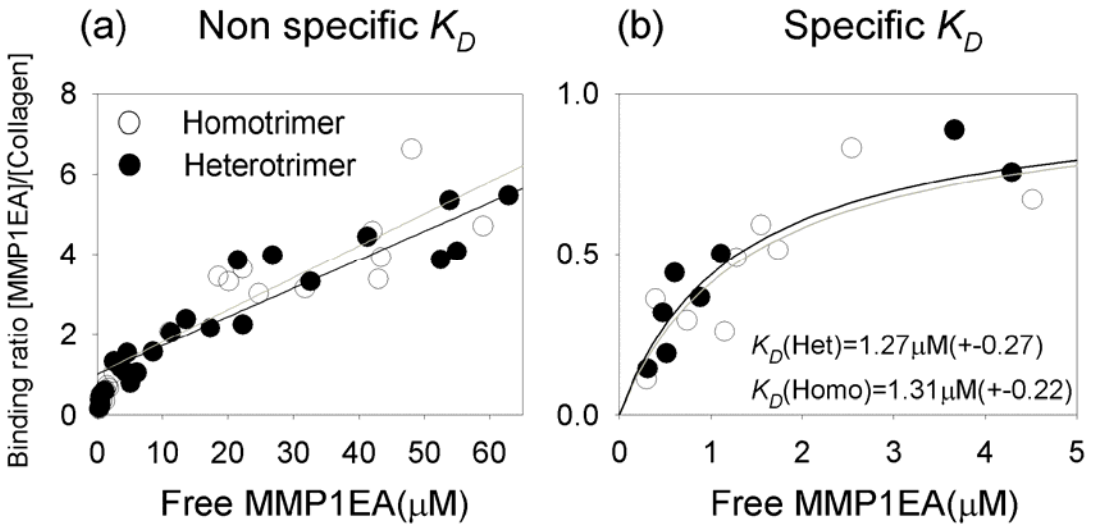

Fig. 5.3. Equilibrium binding of MMP-1(E200A) to homo- and heterotrimeric mouse-tail-tendon type I collagen. (a) The number *N* of bound MMP-1(E200A) molecules at 10 to 65 μM concentration of free MMP-1(E200A) was well fitted by a straight line with the intercept at *N*=1. (b) At concentrations below 5 μM, the binding curve was consistent with the Langmuir isotherm with one binding site and K_d =1.3±0.3 μM (solid lines). This curve was corrected for the nonspecific binding by subtraction of the fitted straight line from the data shown in panel (a).

Equilibrium micro-dialysis experiments revealed nonspecific, low-affinity binding (K_d > 50 μM) of at least five MMP-1(E200A) molecules per mouse-tail-tendon collagen triple helix (Fig. 5.3. (a)) as well as specific, high-affinity binding (K_d = 1.3±0.3 μM) of a single MMP-1(E200A) molecule, apparently at the cleavage site (Fig. 5.3. (b)). Both non-specific and specific binding to the homo- and heterotrimeric collagen were undistinguishable. The specific K_d was consistent with that reported in [131] for type I collagen heterotrimers and with the value of K_m measured in our kinetic experiments. These measurements confirmed that the homotrimer resistance to MMP-1 could not be attributed to weaker enzyme binding and had to be related to triple helix unwinding and/or cleavage rate of individual chains.

5.2.3. Triple helix unwinding and chain cleavage

To understand which of these steps might be affected by the lack of the α2(I) chain, the triple helix unwinding and cleavage of individual chains had to be decoupled. It is believed that MMP-1 may be released either from intact collagen or only after it sequentially cleaves all three chains [131]. To remove this coupling, we investigated homo- and heterotrimeric collagen cleavage by a combination of MMP-1(E200A) and the catalytic domain of MMP-1 (MMP-1(ΔC)). In such a combination, MMP-1(E200A) binds to the triple helix and promotes local unwinding of the chains, exposing them to cleavage by MMP-1(ΔC) or any other enzyme capable of cutting at the exposed sequences. [131]. The cleavage of locally unfolded individual chains then occurs independently.

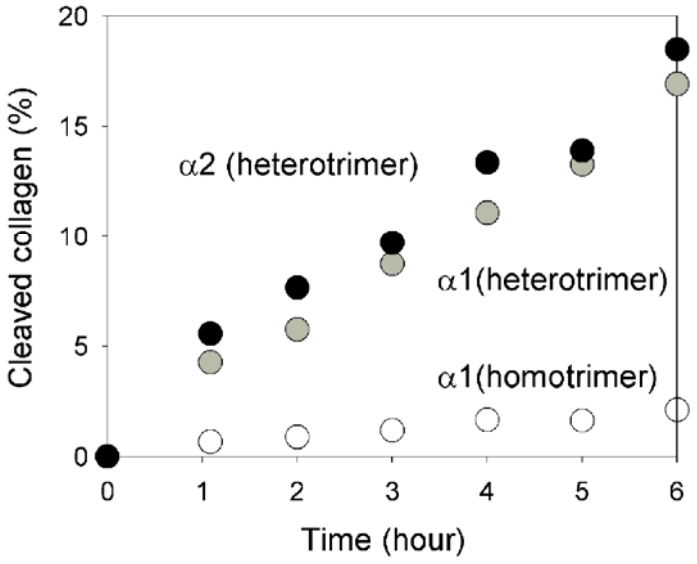

Fig. 5.4. Cleavage of a1(I) and a2(I) chains by MMP-1(E200A)/MMP-1(ΔC) mixture. A binary mixture of homo- and heterotrimeric mouse type I collagens (0.2 μM each) was incubated with 11μM MMP-1(E200A) and 5.8μM MMP-1(ΔC) at 25^{o}C.

Consistent with previous reports [131] , we did not observe noticeable collagen cleavage within the first 6-8 hours by either MMP-1(E200A) alone or MMP-1(ΔC) alone. In mixture of MMP-1(E200A) with MMP-1(ΔC), we observed approximately 10 times faster cleavage of the heterotrimers than homotrimers (Fig. 5.4), as in the case of MMP-1 (Fig.5.1). However, the cleavage rate of the α2(I) and α1(I) chains within the heterotrimers was approximately the same, suggesting that the much slower α1(I) chain cleavage in the homotrimers was related to the resistance of the latter to triple helix unwinding by MMP-1(E200A).

5.2.4. Fibril cleavage

Our studies of reconstituted collagen matrix cleavage by cells [183] and digestion of wild type and oim tendon fibrils with MMP-1 revealed that fibrils of homotrimeric type I collagen were also resistant to collagenases. However, fibrils composed exclusively of the homotrimers are formed only in rare cases of hereditary α2(I) chain deficiency.[12]. Without a genetic defect affecting the α2(I) chain synthesis, the homotrimers comprise just a fraction of all type I collagen. They coassemble with the heterotrimers into the same fibrils[188]. The mixed fibrils may then be susceptible to degradation by collagenases due to the cleavage of one of their main components.

To test whether this were the case, we reconstituted collagen fibrils by *in vitro* fibrillogenesis from a mixture of AlexaFluor-546-labeled mouse heterotrimers and AlexaFluor-647-labeled homotrimers. We observed complete disintegration of most fibrils after incubation with 150 nM MMP-1 at 37 ^{o}C for several hours (Fig. 5.5. a, b). However, we also observed small, residual fibrils, which appeared to be comprised entirely or almost entirely of the homotrimers and were MMP-1-resistant (shown by arrows in Fig. 5.5. b). Since the homo- and heterotrimeric type I collagen molecules tend to segregate at a subfibrillar level[188], these residual homotrimer fibrils likely originated from segregated fibril regions composed of the homotrimers with little or no heterotrimers.

To test what might happen with such fibrils after several tissue remodeling cycles, we induced fibrillogenesis of the heterotrimers in the presence of preformed homotrimer fibrils. Examination in a confocal microscope revealed a fraction of the heterotrimers bound on surfaces of homotrimer fibrils and a fraction that formed new

heterotrimer fibrils (Fig. 5.5. c). After incubation with 150 nM MMP-1 at 37 ^{o}C, we observed complete degradation of the heterotrimer fibrils and complete removal of the heterotrimeric collagen from homotrimer fibril surfaces (Fig. 5.5. d). At the same time, most homotrimer fibrils remained intact.

When we added homotrimers to preformed heterotrimer fibrils, we observed no formation of new homotrimer fibrils. Instead, all homotrimers were bound on surfaces of some but not all preexisting heterotrimer fibrils (Fig. 5.5. (c)). Subsequent incubation with MMP-1 resulted in rapid degradation of fibrils that had little or no bound homotrimers, while the homotrimer-covered fibrils were protected from the cleavage. (Fig. 5.5. (d)).

Thus, homotrimer fibrils formed, e.g., due to homo/heterotrimer segregation, are likely to survive multiple tissue remodeling cycles and grow due to further build up of the collagenase resistant molecules.

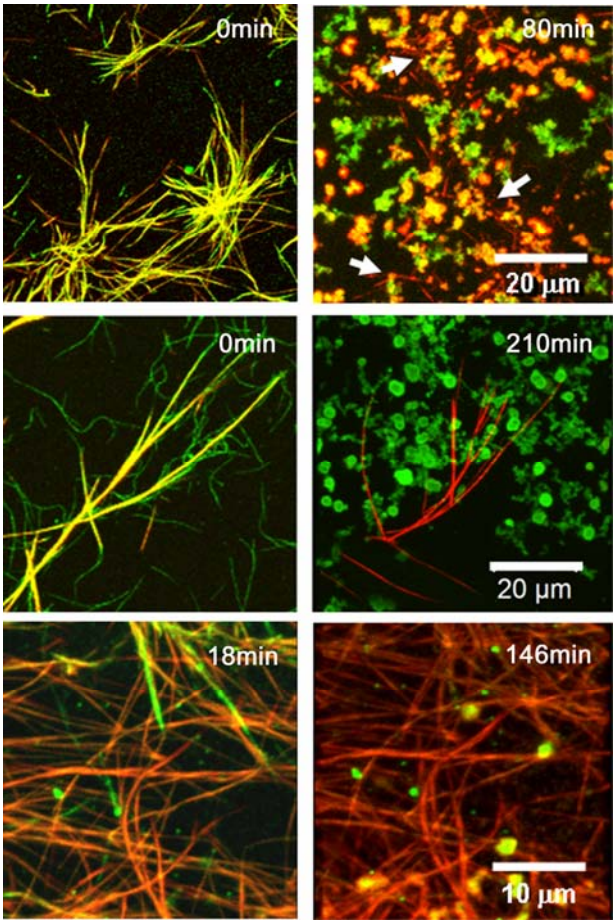

Fig. 5.5. Confocal projections of fibrils reconstituted from heterotrimeric (green) and homotrimeric (red) mouse-tail-tendon collagen before and after digestion with rh-MMP-1. (a) Fibrils reconstituted from a 1:1 mixture of AlexaFluor-546-labeled homotrimers and AlexaFluor-647-labeled-heterotrimers (0.2 mg/ml). [188] (b) A different area of the sample was imaged after replacing the buffer with 150 nM MMP-1 in TNC and incubating for 1.5 h at 37 ^{o}C. (c) An ice-cold solution of heterotrimers (0.2 mg/ml) was added to preformed homotrimer fibrils (0.2 mg/ml), incubated at 32 ^{o}C for 6days, and imaged. (d) A different area of the same sample was imaged 3.5 h after replacing the buffer with 150 nM MMP-1 in TNC and incubating at 37 ^{o}C. (e,f) An ice-cold solution of homotrimers (0.2 mg/ml) was added to preformed heterotrimer fibrils (0.2 mg/ml) and incubated at 32 ^{o}C for 6days. The same area of the sample was imaged after incubating with 150 nM MMP-1 at 37 ^{o}C for 18 min (e) and 2.5 h (f).

5.3. Discussions

5.3.1.MMP-1 binding is essential for efficient triple helix unwinding at the cleavage site

The need for triple helix unwinding prior to chain binding and cleavage in the catalytic cleft of MMPs is generally accepted, but two different mechanisms that account for this observation have been proposed. Within one mechanism, the unwinding is promoted by MMP binding to collagen, as summarized by the following scheme

$$C_N + E \leftrightarrow EC_N \leftrightarrow EC_U \rightarrow E + P \qquad (2)$$

where C_N and C_U indicate collagen in which the MMP cleavage site is in native (helical) state and in locally unwound state, respectively; E is the enzyme, and EC_N and EC_U are enzyme-collagen complexes [131]. Alternatively, the local unwinding of the triple helix is postulated to precede enzyme binding and the EC_N complex is considered to be nonproductive [131],

$$EC_N \leftrightarrow C_N + E \leftrightarrow C_U + E \leftrightarrow EC_U \rightarrow E + P \qquad (3)$$

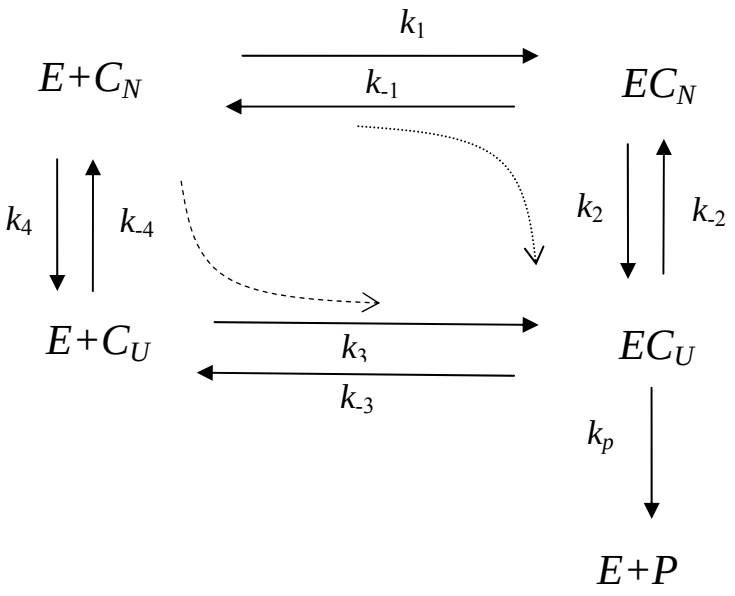

Fig 5.6. Schematic representation of possible collagen cleavage pathways. Each reaction step is characterized by the corresponding rate constant k; E denotes free enzyme. C_N denotes collagen in its native, triple-helical state; C_U denotes collagen with a locally unwound (unfolded) collagenase cleavage site; EC_N and EC_U are enzyme–collagen complexes. P denotes triple helix cleavage products. For simplicity, sequential cleavage of three collagen chains is represented with a single irreversible step. A more detailed model with three irreversible cleavage steps produced substantively similar predictions.

Since the only irreversible step in the cleavage process is peptide bond hydrolysis, both of these mechanisms may be combined within a single kinetic scheme (Fig. 5.6). Within the latter scheme, the first mechanism corresponds to postulating $k_3 = k_{-3} = k_4 = k_{-4} = 0$; the second corresponds to postulating $k_2 = k_{-2} = 0$.

A detailed analysis of previously published experimental data within the full model, without *a priori* assumptions, clearly suggests that the upper pathway proposed in [131] is dominant ($k_2/k_{-2} >> k_4/k_{-4}$). The triple helix unwinding at the cleavage site is greatly enhanced by the enzyme binding (Supplementary Material).

5.3.2. MMP-1 facilitates stochastic triple helix opening rather than actively unwinds it

Our present observations may be summarized as follows: (a) The equilibrium dissociation constants (K_d) for MMP-1(E200A) and the Michaelis constants (K_m) for MMP-1 are the same for the homo- and heterotrimers (Figs. 5.2, 5.3); and they are also close to each other (K_d=1.3±0.3 μM and K_m=0.9±0.2 μM). (b) MMP-1 hydrolyses the α1(I) and α2(I) chains at approximately the same rate (Fig. 5.4), i.e., k_p should be similar for the homo- and heterotrimeric collagens. (c) Nevertheless, the maximum cleavage rate (V_{max}) of the homotrimers is more than 5-10 times slower than heterotrimers (Figs. 5.1, 5.2).

These observations can be reconciled with the cleavage scheme shown in Fig. 5.6 only if $K_d \approx K_m \approx k_{-1}/k_1$, $V_{max} \propto k_2/k_{-2}$, and $k_2/k_{-2} << 1$ (Supplementary Material). The small value of k_2/k_{-2} indicates that the unwinding step is a rare, stochastic event even in the presence of bound MMP-1. In other words, MMP-1 does not *actively* unwind the cleavage site; and most of the triple helices still remain in the native state after MMP-1 binding. Instead, MMP-1 binding facilitates the unwinding by destabilizing the native helical state (increasing k_2 compared with k_4) and/or stabilizing the unwound state (decreasing k_{-2} compared with k_{-4}).

In addition, our study of a mutant type I collagen with a Gly-Ala-Hyp insertion, in which the collagenase cleavage site is completely unwound [204], showed dramatically faster cleavage of the molecules triple helices by MMP1 ΔC compared with normal type I collagen. In this mutant, the α1(I) and α2(I) chains were cleaved independently with the same rate. These observations confirmed that the unwinding is the rate limiting step in the collagenase cleavage ($k_2/k_{-2} << 1$), further supporting our model of collagenase resistance of type I collagen homotrimers.

5.3.3. Type I homotrimers resist triple helix unwinding

Since neither MMP-1 binding (k_{-1}/k_1) nor the cleavage of individual collagen chains (k_p) are affected by the α2(I) chain loss, much lower V_{max} for the homotrimers may be explained only by the homotrimer resistance to the local triple helix unwinding at the cleavage site (reduced k_2/k_{-2}). This conclusion derived from the formal analysis of our results within the Fig. 5.6 cleavage scheme coincides with the simple qualitative interpretation of the observations shown in Fig. 5.4. The resistance of α1(I) homotrimers to this unwinding may explain their universal resistance to different collagenolytic MMPs [183], because the unwinding is required for the cleavage by all of these enzymes.

The heterotrimers may be less resistant to this unwinding because preferential MMP-1 binding to the α2(I) chain stabilizes the enzyme interaction with unwound collagen chains (decreases k_{-2}). A preferential binding of the unwound α2(I) chain in the catalytic cleft of MMP-1(E200A) was indeed reported [131]. However, nearly equal cleavage rate of the α2(I) and α1(I) chains by MMP-1(ΔC) in the presence of

MMP-1(E200A) (Fig. 5.4) suggest that this preferential binding is relatively weak and unlikely to be sufficient for more than 5-10 fold decrease in k_{-2}. Note that we also observed similar cleavage rate of the α2(I) and α1(I) chains by MMP-1(ΔC) in a heterotrimeric type I collagen with a mutation resulting in thermal unwinding of the cleavage site even without MMP-1(E200A) (to be reported elsewhere).

More likely, the resistance of α1(I) homotrimers to the local unwinding at the cleavage site originates from inherently higher stability of the homotrimer triple helix at this site (lower k_2). The latter interpretation is consistent with higher overall stability of the homotrimer helix [18, 155], apparent partial dissociation of the α2(I) chain from the triple helix at the MMP-1 cleavage site [138, 201], and higher apparent activation energy for the homotrimer cleavage (Fig. 1c).

5.3.4. Type I homotrimers may alter tissue remodeling

Since type I homotrimers were found in a variety of pathological conditions [6, 7, 10, 144-146], their resistance to MMPs may have important biological implications. For instance, the homotrimers may promote cancer invasion by resisting collagenases massively secreted by invasive cancer cells and supporting proliferation and migration of these cells[183]. In fibrosis, the MMP resistance of the homotrimers may hinder degradation of sclerotic collagen deposits.

In addition, our observations suggest that the homotrimers may alter tissue remodeling even when they comprise a minor component of type I collagen. Indeed, MMP-1 digestion of mixed fibrils suggested that segregation of the homo- and heterotrimers[188] may result in formation of collagenase-resistant fibril regions

composed primarily of the homotrimers (Fig. 5.5. a,b). These regions may not only survive collagenase degradation but may nucleate homotrimer fibrils that will grow instead of being disassembled during multiple tissue remodeling cycles, eventually causing tissue disorganization (Fig. 5.5. c,d). This effect may be further amplified by the presence of other collagens (e.g., type III and V) in heterotypic fibrils, resulting in pathogenic accumulation of collagenase-resistant fibrils. We are presently testing this hypothesis.

One crucial unanswered question on the potential role of collagenase-resistant type I collagen in fibrosis is why only localized and relatively mild fibrosis is observed in mice in which all type I collagen is collagenase-resistant. In particular, homozygous oim mice, in which all type I collagen is homotrimeric, develop glomerular sclerosis[205]. Homozygous Col1a1$^{r/r}$ mice with an altered α1(I) chain sequence at the primary MMP cleavage site develop thickening of uterine wall, skin, and bones [206-209]. In contrast, MMP-14 knock-out mice develop severe, generalized fibrosis and have short life span, even though their type I collagen has normal collagenase susceptibility and they lack only one membrane collagenase[210]. In our opinion, these observations point to alternative, still poorly understood mechanism(s) of collagen degradation and remodeling as well as to other antifibrotic functions of MMP-14, in addition to type I collagen cleavage. In other words, the homotrimeric type I collagen resistance to collagenases appears to be an essential piece of the fibrosis puzzle, but there are clearly other pieces that are still insufficiently understood or completely unknown.

5.4. Materials and methods

5.4.1. Collagen purification and fluorescent labeling

Human type I collagen heterotrimers were purified from cell culture media of CRL-2127 fibroblasts (American Type Culture Collection, Manassas, VA). Human type I collagen homotrimers were purified from cell culture media of fibroblasts from a homozygous, α2(I)-chain-deficient patient described in [12]. Fibroblasts were grown to confluence at 37 °C in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum and 2 mM glutamax in the presence of 5% CO_2. Collagen secretion was stimulated with 50 μg/ml ascorbate in 0.1% fetal bovine serum. After 24 h stimulation, type I collagen was purified from the media by ammonium sulfate precipitation, pepsin treatment, and several rounds of selective precipitation with NaCl at acidic pH [211]. Type III collagen content was reduced below 1% for heterotrimers and 4% for homotrimers by selective precipitation with NaCl at neutral pH [212]. .Mouse type I collagen heterotrimers and homotrimers were purified by pepsin extraction and several rounds of selective salt precipitation from tail tendons of wild type and homozygous oim mice, respectively. Fluorescent labeling with amino-reactive carboxylic acid succinimidyl ester AlexaFluor 488, 546, or 647 dyes (Invitrogen) or Cy5 NHS-etser dye (GE Healthcare) was performed as described in [188]. The dye concentration was adjusted to 1 dye per 3-5 collagen triple helices labeling efficiency. The unreacted dye was removed by collagen precipitation in 0.9M NaCl and 0.5M acetic acid, dialysis, or size-exclusion chromatography on G-50 spin mini-columns (GE Healthcare). After purification and/or labeling collagen was transferred into and dialyzed against 2 mM HCl. Collagen concentration in each

preparation was measured by circular dichroism at 221 nm with a J810 spectrometer (Jasco Inc.). The purity of each preparation was evaluated by electrophoresis on pre-cast 3–8% Tris–acetate mini-gels (Invitrogen). The gels were scanned on a FLA5000 fluorescence scanner (Fuji Medical Systems) and analyzed by Multi Gauge 3.0 software supplied with the scanner.

5.4.2. Collagen cleavage

Recombinant human proMMP-1, MMP-13, MMP-1(E200A), MMP-1(ΔC), and MMP-3(ΔC) were prepared and purified as described in [131]. ProMMP-1 was activated with 1mM aminophenylmercuric acetate (APMA) and one-fifteenth molar amount MMP-3ΔC at 37^{o}C for 1-2h. The length of enzyme activation was optimized by testing sample aliquots with gel electrophoresis. The activated enzyme was separated from cleaved propeptides, APMA, and MMP-3(ΔC) on G-50 spin mini-columns.

Binary mixtures of type I collagen homotrimers and heterotrimers, in which one component was labeled with AlexaFluor 488 and the other with Cy5, were transferred into 50mM Tris–HCl (pH 7.5), 0.15M NaCl, 10mM $CaCl_2$, 0.05% Brij 35 (TNC buffer) and incubated with activated MMP-1, MMP-13, MMP-1(E200A), MMP-1(ΔC), or combination of MMP-1(E200A) and MMP-1(ΔC). To stop the reaction, sample aliquots were mixed with an LDS gel-loading buffer (Invitrogen) supplemented with 20 mM EDTA. The aliquots were then denatured at 60^{o}C for 5min, and analyzed on 3-8% Tris-Acetate mini gels. To prevent fibrillogenesis at high collagen concentrations, 1 M glycerol was added to the TNC buffer. Testing revealed

no detectable effects of 1 M glycerol or fluorescent labeling on the collagen cleavage rate.

5.4.3. Equilibrium micro-dialysis

MMP-1(E200A) binding to homo- and heterotrimeric type I collagen was measured by equilibrium micro-dialysis in a dialyzer with 50 μL chambers separated by a 100,000 MWCO cellulose acetate membrane (Harvard Apparatus, Holliston, MA). One chamber was filled with a mixture of 0–65 μM MMP-1(E200A) and 0.8-11 μM ALexaFluor-488-labeled-mouse-tail-tendon collagen in TNC buffer supplemented with 1 M glycerol. The other chamber was filled with TNC/glycerol buffer or with MMP-1(E200A) solution in this buffer at the concentration slightly lower than in the first chamber. The dialyzer was incubated at room temperature on a rotator for 2-3 days, until the equilibrium was reached. After the incubation, light absorption spectra from each solution were measured from 210 to 650 nm in a Lambda-900 spectrophotometer (Perkin-Elmer). The collagen concentration was determined from the 495 nm absorption peak of AlexaFluor 488. The concentration of MMP-1(E200A) was determined from the absorption at 280 nm corrected for the contribution of AlexaFluor 488 labeled collagen. No collagen leakage through the dialysis membrane was observed. The collagen cleavage by MMP-1(E200A) during the equilibration time was negligible, as verified by gel electrophoresis. The number of bound MMP-1(E200A) molecules per triple helix was determined based on the difference in total MMP-1(E200A) concentration in the two chambers.

5.4.4. Confocal microscopy imaging

Type I homo- and heterotrimers were labeled by Alexa Fluor 546 and Alexa Fluor 647 carboxylic acid succinimidyl ester (Invitrogen), respectively. Collagen solution in 2 mM HCl was mixed 1:1 with 60mM sodium phosphate, 0.26M NaCl (pH 7.4 after mixing), transferred into SecureSeal hybridization chambers (Grace Bio-Labs), sealed, and incubated at 32 ^{o}C for 2 days, until fibril formation was completed. In some experiments, the two collagen isoforms were mixed before the fibrillogenesis. In other experiments, the fibrillogenesis was initially performed with one isoform and a cold solution of the other isoform was added to the fibrils thermostated at 32 ^{o}C, followed by additional incubation at 32 ^{o}C for XX h.

Fibril degradation by collagenase was visualized in an LSM 510 Inverted Meta confocal laser scanning microscope (Carl Zeiss) with a 60X, 1.4 NA oil-immersion objective. The hybridization chamber was mounted on a microscope slide and placed inside a controlled environment chamber set at 37 ^{o}C. After initial fibril imaging, the fibrillogenesis buffer solution was replaced with MMP-1 solution in TNC and the sample was quickly remounted in the microscope. The samples were imaged with a 543 nm and 633 nm Helium-Neon lasers. The emission was collected by independent photomultipliers in two channels with 560-615 nm band pass and 650 nm long pass filters as described in [188]. The digital zoom, z-oversampling, scanning rate, and laser power were optimized to eliminate photo damage and photobleaching. Very low laser power, longer laser wavelength (≥ 543 nm), and short laser exposure times were required to avoid the damage of MMP-1. Otherwise, irreversible accumulation of

MMP-1 at fibril surfaces and no fibril cleavage were observed within areas imaged before the degradation was completed.

5.4. Supplementary material

5.4.1. Homo- and heterotrimer cleavage by MMP-13

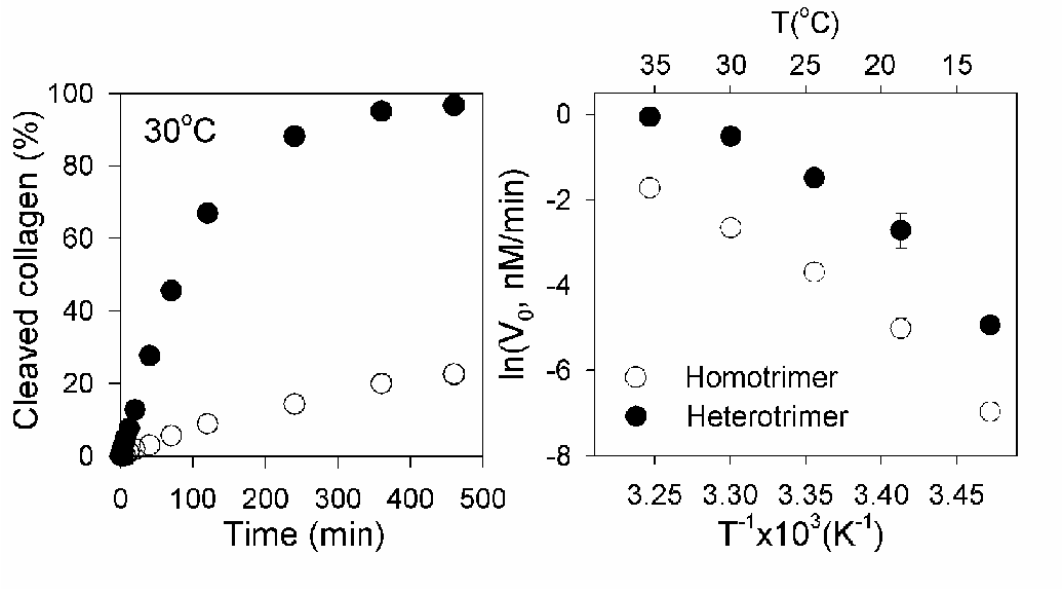

Fig 5.7. Cleavage kinetics of human type I collagen homo- and heterotrimers (0.05 mg/ml) by rh MMP-13 (5nM). The experiment was performed as described in Fig. 5.1.

5.4.2. General steady-state cleavage kinetics by MMP-1

Consider the kinetic scheme of collagen cleavage by MMP-1 represented in Fig. 6 in more detail. Provided that the total enzyme concentration $[E]_{total}$ is sufficiently small compared to the collagen concentration, the initial cleavage rate of collagen V_0 within this scheme can be calculated by using the following steady-state approximation.

$$\frac{d[EC_n]}{dt} = k_1[E][C_n] + k_{-2}[EC_u] - k_{-1}[EC_n] - k_2[EC_n] \approx 0 \quad \text{(S1)}$$

$$\frac{d[EC_u]}{dt} = k_3[E][C_u] + k_2[EC_n] - k_{-2}[EC_u] - k_{-3}[EC_u] - k_5[EC_u] \approx 0 \quad \text{(S2)}$$

$$[E]_{total} = [E] + [EC_n] + [EC_u] \quad \text{(S3)}$$

Here C_n and C_u indicate native, triple helical (n) and unwound (u) states of collagen chains at the MMP-1 cleavage site, respectively, and [E] is the concentration of free enzyme.

By solving these equations, we find

$$V_0 = k_p[EC_u] \approx \frac{k_p\{k_1k_2[C_n] + k_3(k_2 + k_{-1})[C_u]\}[E]_{total}}{k_1(k_2 + k_{-2} + k_{-3} + k_p)[C_n] + k_3(k_{-2} + k_2 + k_{-1})[C_u] + k_2(k_{-3} + k_p) + k_{-1}(k_{-2} + k_{-3} + k_p)} \quad \text{(S4)}$$

where

$$[C_u] \approx \frac{k_4}{k_{-4}}[C_n] \ . \quad \text{(S5)}$$

These expressions may be significantly simplified by taking into account the results of experiments described in, as discussed below.

5.4.3. Collagen cleavage by MMP-1E200A/cutter-enzyme mixtures

Indeed, it has been well documented that the catalytic domain of MMP-1 (MMP-1(ΔC)) alone cannot efficiently cleave native collagen even at high concentration, although it readily cleaves denatured collagen [132, 213]. However, MMP-1(ΔC) and even serine proteases cleave collagen at the MMP-1 cleavage site in the presence of MMP-1(E200A), a catalytically disabled form of MMP-1 that binds to collagen and promotes unwinding of the cleavage site with the same affinity as MMP-1 [131]. The latter experiment may be represented by the kinetic scheme shown in Fig 5.8, which is constructed based on the scheme for MMP-1 (Fig. 5.6), assuming that MMP-1(ΔC) and serine proteases are responsible only for cleaving the collagen with unwound triple helix at the MMP-1 cleavage site (C_u).

Since no detectable cleavage of collagen by MMP-1(ΔC) in the absence of MMP-1(E200A) is observed for hours, we can conclude from Fig 5.8 that

$$\frac{k_4}{k_8} << 1 \qquad \text{(S6)}$$

In contrast, rapid collagen cleavage is observed at the same MMP-1(ΔC) concentrations upon addition of MMP-1(E200A) (E_A). Since MMP-1(E200A) has almost no catalytic activity, it mostly facilitates the unwinding of the MMP-1 cleavage site by forming the E_AC_u complex [131]. This complex (E_AC_u) and the collagen with spontaneously unwound cleavage site (C_u) serve as the substrates for

MMP-1(ΔC). The cleavage rate should then be proportional to the sum of the concentration of C_u and E_AC_u:

$$V_0 \propto [E_AC_u] + [C_u] \quad \text{(S7)}$$

Since

$$[C_u] \approx \frac{k_4}{k_{-4}}[C_n] \leq \frac{k_4}{k_{-4}}[C]_{total} \quad \text{(S8)}$$

and

$$[E_AC_u] \approx \frac{k_2}{k_{-2}}[E_AC_n] \leq \frac{k_2}{k_{-2}}\frac{k_1}{k_{-1}}[E_A][C]_{total} \quad \text{(S9)}$$

we find

$$\frac{V_0([E_A])}{V_0([E_A]=0)} \leq 1 + \frac{k_1}{k_{-1}}[E_A]\cdot\frac{k_2/k_{-2}}{k_4/k_{-4}} \quad \text{(S10)}$$

Because collagen cleavage at $[E_A]\sim k_{-1}/k_1$ (≈ 1 μM) is orders of magnitude more efficient than at $[E_A]=0$, i.e.,

$$V_0(k_{-1}/k_1) >> V_0(0) \quad \text{(S11)}$$

(i.e., $V_0(k_{-1}/k_1) >> V_0(0)$) [1], we also find

$$\frac{k_2}{k_{-2}} >> \frac{k_4}{k_{-4}} \qquad \text{(S12)}$$

From the equilibrium in the absence of MMP-1(ΔC), we find

$$\frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}} = \frac{k_4}{k_{-4}}\frac{k_3}{k_{-3}} \qquad \text{(S13)}$$

Since the on-rates k_1 and k_3 are determined mostly by diffusion, one may expect them to be similar. Then, from Eqs. (S12), (S13) we obtain

$$\frac{k_{-3}}{k_{-1}} << 1 \qquad \text{(S14)}$$

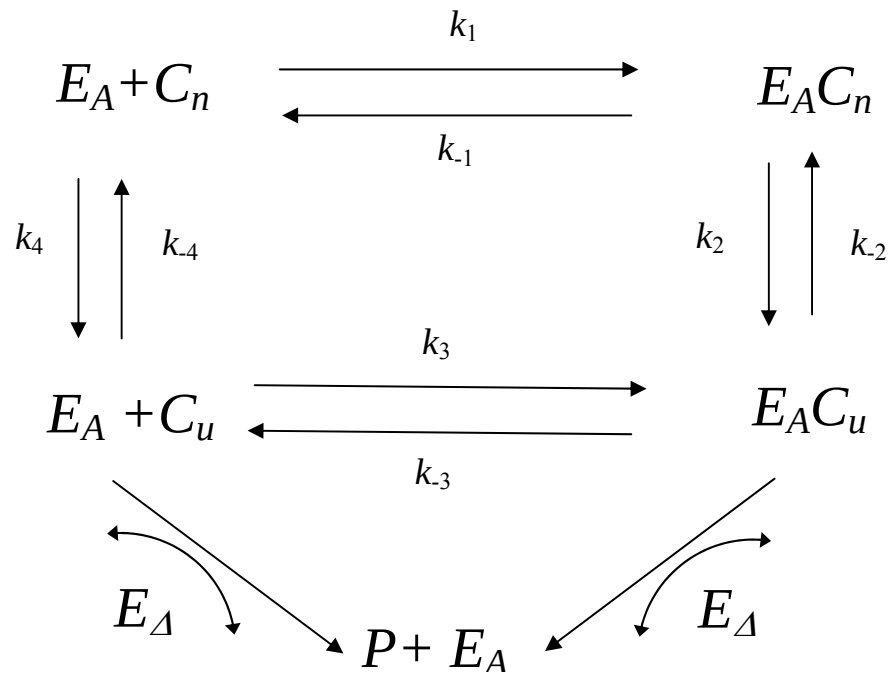

Fig 5.8. Schematic representation of collagen cleavage by MMP-1(ΔC) in the presence of MMP-1(E200A). E_A denotes MMP-1(E200A) and E_Δ denotes MMP-1(ΔC). Following Ref. [131], we assume that MMP-1(E200A) interact with collagen in exactly the same way as MMP-1, but it cannot hydrolyze peptide bonds. Therefore, MMP-1(E200A) interaction with collagen is represented by the same scheme as in Fig. 5.6 in the main text, but with k_p=0. MMP-1ΔC can cleave the vulnerable form of collagen with unwound triple helix at the MMP-1 cleavage site (C_u), which may be present free in solution and in complex with MMP-1(E200A). *P* indicates collagen cleavage products.

5.4.4. Cleavage kinetics of type I homotrimers

Using Eqs. (S5), (S6), and (S12)-(S14), we may reduce Eq. (S4) to the Michaelis-Menton equation (Eq. (1)), in which K_m and V_{max} are given by

$$K_m \approx \frac{k_{-1}}{k_1}\left\{\frac{1}{\left(1+\frac{k_2}{k_{-2}+k_{-3}+k_p}\right)}\right\} \qquad \text{(S15)}$$

and

$$V_{\max} \approx \frac{k_p k_2 [E]_{total}}{k_2 + k_{-2} + k_{-3} + k_p} \qquad \text{(S16)}$$

This means that the upper pathway in Fig. 5.6 is dominant and the full kinetic scheme may essentially be reduced to the one proposed in [131] with separate steps of enzyme binding and local triple helix unwinding.

Since MMP-1(E200A) binding to type I homo and heterotrimers is identical, the value of K_m for MMP-1 cleavage is also the same, and the value of $V_{\max}$ is much smaller for homotrimers, we can conclude that

$$K_m \approx \frac{k_{-1}}{k_1} \quad \left(\frac{k_2}{k_{-2}+k_{-3}+k_p} << 1\right) \qquad \text{(S17)}$$

The lack of the α2(I) chain in type I collagen homotrimers appears to reduce the value of $k_2/(k_2 + k_{-2} + k_{-3} + k_p)$ by inhibiting the local unwinding of the triple helix (reducing k_2) and/or promoting the helix rewinding (increasing k_{-2}).

Chapter 6: Conclusions

At the onset of this study, we were not aware of the potential importance of homotrimeric type I collagen in cancer, fibrosis, osteoporosis, and other common disorders. We were motivated mostly by studies of rare recessive forms of Osteogenesis Imperfecta (OI) and Ehlers-Danlos syndrome (EDS) in which only homotrimeric, but no normal heterotrimeric type I collagen was found [12]. Our initial goal was to understand the physics of type I collagen fibril nucleation, specifically to uncover the reasons why no lag phase was observed during the self-assembly of homotrimers[18]. The explanation turned out to be mundane: pepsin-treated homotrimers were prone to fast nonspecific aggregation in fibrillogenesis conditions[188]. Nevertheless, this work led us in an exciting foray into more complex physics of collagen fibril network self-assembly, mechanics, and remodeling as well as into studying the implications of physics in the biology of cancer and other disorders. The main conclusions of the resulting study and hypotheses based on these conclusions are summarized below.

Homo- and heterotrimeric type I collagen molecules co-assemble within the same fibrils but segregate into homotrimer and heterotrimer rich regions. In Chapter 3, we established that this segregation had both kinetic and equilibrium components. The kinetic segregation appears to be associated with different rates of homo- and heterotrimer fibril assembly, e.g., more pronounced tendency of the homotrimers to nonspecific aggregation. The equilibrium segregation is associated with the hierarchy

of interaction energies, the heterotrimer-heterotrimer interactions being most favorable and homotrimer-homotrimer interactions being least favorable. We also observed how homotrimers dramatically affect fibril network architecture, which is associated with the homotrimer tendency to nonspecifically aggregate. Spear-like homotrimer fibrils grew in multiple directions outward from relatively large nonspecific aggregates. The resulting homotrimer fibril network was more reminiscent of a convention of long-needled sea urchins (Fig. 4.1 (b)) rather than a normal meshwork of entangled threads (Fig. 4.1 (a)).

Homotrimer fibrils form stiffer networks; the spear-like fibril morphology results from longer persistence length. More hydrated, and less well organized homotrimer fibrils were less willing to bend. Therefore, we carefully measured the persistence lengths of homo- and heterotrimer fibrils (Chapter 4). A comparison of the results with previously reported mechanical measurements suggested that fibril bending was accompanied by free slippage of the molecules along each other. Thus, we concluded that increased fibril stiffness was related to higher bending rigidity of individual homotrimer molecules. Most likely, this higher bending rigidity originates from the increased stability of the homotrimer triple helix [18, 155]. The "sea urchin convention" architecture of homotrimer fibril network is likely to increase the stiffness and fragility of networks, although we do not know yet whether such architecture occurs *in vivo*.

Higher homotrimer triple helix stability results in resistance to cleavage by collagenases, enhancing detrimental effects of replacing a stronger, more interconnected, and more flexible heterotrimer network with a stiffer and more fragile homotrimer one. Type I collagen triple helix cleavage by collagenases is crucial for extracellular matrix degradation and remodeling. The cleavage requires local unwinding of the triple helix at the cleavage site[131]. In Chapter 5, we demonstrated that significantly less efficient unwinding of homotrimeric type I collagen triple helices was responsible for their collagenase resistance. We found the unwinding to be a stochastic rather than active process, although promoted by collagenase binding. Our data were most consistent with a simple interpretation of inefficient unwinding as resulting from an increased stability of the homotrimer triple helix at the cleavage site.

More rigid, abnormally organized homotrimer fibril network may accumulate after multiple tissue remodeling cycles, resulting in tissue pathology. Combining our findings uncovered throughout all steps in this project, we realized how synthesis of even a small fraction of homotrimeric type I collagen molecules in various disorders might result in very serious pathological consequences. Indeed, type I collagen fibrils are constantly renewed; the old ones are degraded by collagenases and replaced with newly synthesized ones. In humans, under normal conditions, complete fibril remodeling takes from months (in soft tissues) to several years (in bones)[214]. Under inflammatory conditions, however, this process may occur much faster. We found that due to subfibrillar segregation and collagenase resistance of the

homotrimers, homotrimer-rich fibril regions might not be degraded. Instead of being disassembled, these regions may nucleate growth of predominantly homotrimeric fibrils, which may then accumulate with each tissue remodeling cycle (Chapter 5). After multiple remodeling cycles, these fibrils may form an extended, abnormal network (stiff, fragile, and collagenase resistant), resulting in significant tissue pathology. Our laboratory is currently testing this hypothesis. Preliminary evidence suggests that this process may contribute, e.g., to pathological collagen accumulation resulting in kidney failure.

Distinctive physical properties of individual homotrimeric type I collagen molecules and their networks may play an important role in cancer invasion. Massive production of collagenases is a hallmark of invasive cancers[183], often associated with poor prognosis. These enzymes, produced by cancer cells and tumor-associated fibroblasts, degrade the surrounding collagen matrix, allowing cancer cells to penetrate through blood vessel walls and invade different tissues[126]. However, cancer cells also utilize the type I collagen matrix as an essential substrate for their own proliferation and migration. The ability of cancer cells to produce collagenase-resistant homotrimeric molecules may enable them to degrade the host tissue collagen matrix at the leading front of the invasion while laying their own matrix resistant to the degradation. The tendency of homotrimeric molecules to segregate may help in separating the host matrix from the matrix produced by fibroblasts recruited into tumors. The latter matrix, unlike the matrix made by the cancer cells, needs to be degraded during outward tumor growth. Stiffness and the peculiar architecture of the

homotrimer matrix may also be beneficial for cancer cells, which are known to proliferate and migrate faster on stiffer substrates.

Over the last several months, my colleagues in our laboratory tested 9 different invasive cancer cell lines from five different types of cancer (melanoma, mammary adenocarcinoma, prostate adenocarcinoma, fibrosarcoma, and pediatric neuroblastoma). All of these cells produced type I collagen, of which 20-40% were homotrimeric molecules[183]. Preliminary studies suggested that at least some of these cells did proliferate and migrate faster on the matrix reconstituted from homotrimeric fibrils compared to heterotrimeric fibrils. Analysis of xenograft tumors produced by human cancer cells in immunodeficient mice confirmed that mouse fibroblasts recruited into the tumors produced only heterotrimeric type I collagen while over 50% of human collagen molecules made by the cancer cells were homotrimers.

In other words, so far all key elements of our hypothesis appear to be supported by the data, although much work remains to be done before the testing is complete. We hope that some of the physical properties of homotrimeric type I collagen we discovered may prove to be useful in developing new approaches for the diagnostic and therapeutic targeting of cancer through the homotrimers made by cancer cells.

Bibliography

1. Fratzl, P. and SpringerLink (Online service). *Collagen structure and mechanics*. [SpringerLink] 2008 [cited; xviii, 506 p.]. Available from: http://www2.lib.purdue.edu:2048/login?url=http://dx.doi.org/10.1007/978-0-387-73906-9 An electronic book accessible through the World Wide Web; click for information
2. Keith, C.M., Hopkinson,I $Grant,M.E, *Collagen :The collagen Family:Structure,Assembly and Organizasion in the Extracellular Matrix in Connective tissue and its heritable disorders molecular,genetic,and medical aspects*, ed. B.S. Peter M.Royce. 2002: Wiley-liss,Inc.,NY. 103-147.
3. Jimenez, S.A., et al., *Identification of collagen alpha1(I) trimer in embryonic chick tendons and calvaria.* Biochem Biophys Res Commun, 1977. **78**(4): p. 1354-61.
4. Pucci-Minafra, I., et al., *A new form of tumor and fetal collagen that binds laminin.* Biochemistry, 1993. **32**(29): p. 7421-7.
5. Moro, L. and B.D. Smith, *Identification of collagen alpha1(I) trimer and normal type I collagen in a polyoma virus-induced mouse tumor.* Arch Biochem Biophys, 1977. **182**(1): p. 33-41.
6. Ghersi, G., A.M. La Fiura, and S. Minafra, *Direct adhesion to type I and homotrimer collagens by breast carcinoma and embryonic epithelial cells in culture: a comparative study.* Eur J Cell Biol, 1989. **50**(2): p. 279-84.
7. Rupard, J.H., et al., *Synthesis of type I homotrimer collagen molecules by cultured human lung adenocarcinoma cells.* Am J Pathol, 1988. **133**(2): p. 316-26.
8. Haralson, M.A., H.R. Jacobson, and R.L. Hoover, *Collagen polymorphism in cultured rat kidney mesangial cells.* Lab Invest, 1987. **57**(5): p. 513-23.
9. Gibbs, S.R., et al., *Characterization of the collagen phenotype of rabbit proximal tubule cells in culture.* Connect Tissue Res, 1999. **40**(3): p. 173-88.
10. Kuo, M.Y., et al., *Collagen biosynthesis in human oral submucous fibrosis fibroblast cultures.* J Dent Res, 1995. **74**(11): p. 1783-8.
11. Prockop, D.J., *Mutations in collagen genes as a cause of rare and perhaps common diseases of connective tissue.* Acta Paediatr Scand Suppl, 1991. **379**: p. 55-7; discussion 58.
12. Schwarze, U., et al., *Rare autosomal recessive cardiac valvular form of Ehlers-Danlos syndrome results from mutations in the COL1A2 gene that activate the nonsense-mediated RNA decay pathway.* Am J Hum Genet, 2004. **74**(5): p. 917-30.
13. Kabla, A. and L. Mahadevan, *Nonlinear mechanics of soft fibrous networks.* J R Soc Interface, 2007. **4**(12): p. 99-106.
14. BREGG, R.K., *Current topics in polymer research*. illustrated ed. 2006: Nova Publishers.
15. Rosler, A., G.W. Vandermeulen, and H.A. Klok, *Advanced drug delivery devices via self-assembly of amphiphilic block copolymers.* Adv Drug Deliv Rev, 2001. **53**(1): p. 95-108.

16. Whitesides, G.M. and M. Boncheva, *Beyond molecules: self-assembly of mesoscopic and macroscopic components.* Proc Natl Acad Sci U S A, 2002. **99**(8): p. 4769-74.
17. Lehn, J.M., *Toward self-organization and complex matter.* Science, 2002. **295**(5564): p. 2400-3.
18. Kuznetsova, N.V., D.J. McBride, and S. Leikin, *Changes in thermal stability and microunfolding pattern of collagen helix resulting from the loss of alpha2(I) chain in osteogenesis imperfecta murine.* J Mol Biol, 2003. **331**(1): p. 191-200.
19. Arêas, E.P.G. and J.A.G. Arêas, *Mesoscopic segregation in binary liquid solutions: application of a model for block copolymers to non-macromolecular systems.* Journal of Molecular Structure: THEOCHEM, 1999. **464**(1-3): p. 199-209.
20. Bates, F.S., *Polymer-Polymer Phase Behavior.* Science, 1991. **251**(4996): p. 898-905.
21. Teraoka, I., *Polymer solutions : an introduction to physical properties*. 2002, New York: Wiley. xv, 338.
22. McBride, D.J., Jr., et al., *Altered collagen structure in mouse tail tendon lacking the alpha 2(I) chain.* J Mol Biol, 1997. **270**(2): p. 275-84.
23. Misof, K., et al., *Collagen from the osteogenesis imperfecta mouse model (oim) shows reduced resistance against tensile stress.* J Clin Invest, 1997. **100**(1): p. 40-5.
24. Collins, C.A. and R.B. Vallee, *Temperature-dependent reversible assembly of taxol-treated microtubules.* J Cell Biol, 1987. **105**(6 Pt 1): p. 2847-54.
25. Dou, H., et al., *pH-dependent self-assembly: micellization and micelle-hollow-sphere transition of cellulose-based copolymers.* Angew Chem Int Ed Engl, 2003. **42**(13): p. 1516-9.
26. Narayanan, A.S., et al., *Action of mammalian collagenases on type I trimer collagen.* Coll Relat Res, 1984. **4**(4): p. 289-96.
27. S. Han, E.M., N.V. Kuznetsova, A.M. DeRidder, M.B. Sutter, D.J. McBride, C.L. Phillips, U. Schwarze, J.M. Pace, P.H. Byers, R. Visse, H. Nagase, and S. Leikin, *Molecular mechanism of the resistance of a homotrimeric type I collagen isoform to mammalian collagenases.* PNAS (in preparation), 2009.
28. Scapino, R.P., A. Obrez, and D. Greising, *Organization and function of the collagen fiber system in the human temporomandibular joint disk and its attachments.* Cells Tissues Organs, 2006. **182**(3-4): p. 201-25.
29. Kadler, K., *Protein profile,Extracelluelar Matrix 1: fibril-forming collagens*. Vol. 1. 1994. 519-522
30. Kielty, C.M., Hopkinson,I Grant,M.E, *Collagen :The collagen Family:Structure,Assembly and Organizasion in the Extracellular Matrix in Connective tissue and its heritable disorders molecular,genetic,and medical aspects*, ed. B.S. Peter M.Royce. 2002: Wiley-liss,Inc.,NY. 159-221.
31. Miller, E.J., *The structure of fibril-forming collagens.* Ann N Y Acad Sci, 1985. **460**: p. 1-13.

32. Linsenmayer, T.F., et al., *Type V collagen: molecular structure and fibrillar organization of the chicken alpha 1(V) NH2-terminal domain, a putative regulator of corneal fibrillogenesis.* J Cell Biol, 1993. **121**(5): p. 1181-9.
33. Morris, N.P. and H.P. Bachinger, *Type XI collagen is a heterotrimer with the composition (1 alpha, 2 alpha, 3 alpha) retaining non-triple-helical domains.* J Biol Chem, 1987. **262**(23): p. 11345-50.
34. Boot-Handford, R.P. and D.S. Tuckwell, *Fibrillar collagen: the key to vertebrate evolution? A tale of molecular incest.* Bioessays, 2003. **25**(2): p. 142-51.
35. Wu, J.J., D.R. Eyre, and H.S. Slayter, *Type VI collagen of the intervertebral disc. Biochemical and electron-microscopic characterization of the native protein.* Biochem J, 1987. **248**(2): p. 373-81.
36. Fitzgerald, J. and J.F. Bateman, *A new FACIT of the collagen family: COL21A1.* FEBS Lett, 2001. **505**(2): p. 275-80.
37. Tuckwell, D., *Identification and analysis of collagen alpha 1(XXI), a novel member of the FACIT collagen family.* Matrix Biol, 2002. **21**(1): p. 63-6.
38. Beck, K. and B. Brodsky, *Supercoiled protein motifs: The collagen triple-helix and the alpha-helical coiled coil.* Journal of Structural Biology, 1998. **122**(1-2): p. 17-29.
39. Crick, A.R.a.F.H.C., *The molecular structure of collagen.* Journal of Molecular biology, 1961. **3** p. 483–506.
40. Brodsky, B. and J.A.M. Ramshaw, *The collagen triple-helix structure.* Matrix Biology, 1997. **15**(8-9): p. 545-554.
41. Bella, J. and H.M. Berman, *Crystallographic evidence for C-alpha-H center dot center dot center dot O=C hydrogen bonds in a collagen triple helix.* Journal of Molecular Biology, 1996. **264**(4): p. 734-742.
42. Sakakibara, S., et al., *Synthesis of (Pro-Hyp-Gly) n of defined molecular weights. Evidence for the stabilization of collagen triple helix by hydroxypyroline.* Biochim Biophys Acta, 1973. **303**(1): p. 198-202.
43. Gustavson, K.H., *The function of hydroxyproline in collagens.*
. Nature 1955. **175**: p. 70-74.
44. Bella, J., B. Brodsky, and H.M. Berman, *Hydration Structure of a Collagen Peptide.* Structure, 1995. **3**(9): p. 893-906.
45. Bella, J., et al., *Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution.* Science, 1994. **266**(5182): p. 75-81.
46. Berisio, R., et al., *Recent progress on collagen triple helix structure, stability and assembly.* Protein Pept Lett, 2002. **9**(2): p. 107-16.
47. Jenkins, C.L. and R.T. Raines, *Insights on the conformational stability of collagen.* Nat Prod Rep, 2002. **19**(1): p. 49-59.
48. Chan, V.C., et al., *Positional preferences of ionizable residues in Gly-X-Y triplets of the collagen triple-helix.* J Biol Chem, 1997. **272**(50): p. 31441-6.
49. Ramshaw, J.A., N.K. Shah, and B. Brodsky, *Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides.* J Struct Biol, 1998. **122**(1-2): p. 86-91.
50. Jeremy M.Berg, J.L.T., Lubert Stryer, *biochemistry.* 2002: p. 41-76.

51. Engel, J. and D.J. Prockop, *The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper.* Annu Rev Biophys Biophys Chem, 1991. **20**: p. 137-52.
52. Ackerman, M.S., et al., *Sequence dependence of the folding of collagen-like peptides. Single amino acids affect the rate of triple-helix nucleation.* J Biol Chem, 1999. **274**(12): p. 7668-73.
53. Brownell, A.G. and A. Veis, *Intracellular location of triple helix formation of collagen. Enzyme probe studies.* J Biol Chem, 1976. **251**(22): p. 7137-43.
54. Alberts, B.J., Alexander; Lewis, Julian,Raff, Martin; Roberts, Keith; Walter, Peter. , *Molecular Biology of the Cell*. 4 ed. 2002: Garland. 1096-1102.
55. Jenkins, C.L., et al., *Effect of 3-Hydroxyproline Residues on Collagen Stability.* J. Am. Chem. Soc., 2003. **125**(21): p. 6422-6427.
56. Kivirikko, K.I. and T. Pihlajaniemi, *Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases.* Adv Enzymol Relat Areas Mol Biol, 1998. **72**: p. 325-98.
57. Guzman, N.A., *Prolyl Hydroxylase, Protein Disulfide Isomerase and Other Proteins*. 1997, New York, USA: Marcel Decker, Inc.
58. Valli, M., et al., *"In vitro" fibril formation of type I collagen from different sources: biochemical and morphological aspects.* Connect Tissue Res, 1986. **15**(4): p. 235-44.
59. Prockop, D.J. and K.I. Kivirikko, *Collagens: molecular biology, diseases, and potentials for therapy.* Annu Rev Biochem, 1995. **64**: p. 403-34.
60. Kadler, K.E., et al., *Collagen fibril formation.* Biochem J, 1996. **316 (Pt 1)**: p. 1-11.
61. Myllyharju, J. and K.I. Kivirikko, *Collagens and collagen-related diseases.* Ann Med, 2001. **33**(1): p. 7-21.
62. Schofield, J.D. and D.J. Prockop, *Procollagen-a precursor form of collagen.* Clin Orthop Relat Res, 1973(97): p. 175-95.
63. Siegel, R.C., *Biosynthesis of collagen crosslinks: increased activity of purified lysyl oxidase with reconstituted collagen fibrils.* Proc Natl Acad Sci U S A, 1974. **71**(12): p. 4826-30.
64. Kagan, H.M., et al., *Influence of sequence and charge on the specificity of lysyl oxidase toward protein and synthetic peptide substrates.* J Biol Chem, 1984. **259**(18): p. 11203-7.
65. Bozec, L., et al., *Atomic force microscopy of collagen structure in bone and dentine revealed by osteoclastic resorption.* Ultramicroscopy, 2005. **105**(1-4): p. 79-89.
66. Kagan, H.M., et al., *Ultrastructural immunolocalization of lysyl oxidase in vascular connective tissue.* J Cell Biol, 1986. **103**(3): p. 1121-8.
67. Kagan, H.M., *Lysyl oxidase: mechanism, regulation and relationship to liver fibrosis.* Pathol Res Pract, 1994. **190**(9-10): p. 910-9.
68. Pinnell, S.R. and G.R. Martin, *The cross-linking of collagen and elastin: enzymatic conversion of lysine in peptide linkage to alpha-aminoadipic-delta-semialdehyde (allysine) by an extract from bone.* Proc Natl Acad Sci U S A, 1968. **61**(2): p. 708-16.

69. Kadler, K.E., Y. Hojima, and D.J. Prockop, *Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process.* J Biol Chem, 1987. **262**(32): p. 15696-701.
70. Lightfoot, S.J., et al., *Type I procollagens containing substitutions of aspartate, arginine, and cysteine for glycine in the pro alpha 1 (I) chain are cleaved slowly by N-proteinase, but only the cysteine substitution introduces a kink in the molecule.* J Biol Chem, 1992. **267**(35): p. 25521-8.
71. Bailey, A.J., *Molecular mechanisms of ageing in connective tissues.* Mech Ageing Dev, 2001. **122**(7): p. 735-55.
72. Prockop, D.J., et al., *The biosynthesis of collagen and its disorders (first of two parts).* N Engl J Med, 1979. **301**(1): p. 13-23.
73. Wood, G.C., *The precipitation of collagen fibers from solution.* Int Rev Connect Tissue Res, 1964. **2**: p. 1-31.
74. Lambeir, A. and Y. Engelborghs, *A quantitative description of microtubule formation in the presence of tubulin-colchicine.* Eur J Biochem, 1983. **132**(2): p. 369-73.
75. Oosawa, F., and Asakura,S., *thermodynamics of the polymerization of protein*. 1975: Academic press,New york.
76. Leikin, S., D.C. Rau, and V.A. Parsegian, *Temperature-favoured assembly of collagen is driven by hydrophilic not hydrophobic interactions.* Nat Struct Biol, 1995. **2**(3): p. 205-10.
77. Leikin, S. and V.A. Parsegian, *Temperature-induced complementarity as a mechanism for biomolecular assembly.* Proteins, 1994. **19**(1): p. 73-6.
78. Canty, E.G. and K.E. Kadler, *Procollagen trafficking, processing and fibrillogenesis.* J Cell Sci, 2005. **118**(Pt 7): p. 1341-53.
79. payne, A.V.a.K., *collagen fibrillogenesis.* **1**: p. 113-137.
80. Kasai, M., S. Asakura, and F. Oosawa, *The cooperative nature of G-F transformation of actin.* Biochim Biophys Acta, 1962. **57**: p. 22-31.
81. Kasai, M., S. Asakura, and F. Oosawa, *The G-F equilibrium in actin solutions under various conditions.* Biochim Biophys Acta, 1962. **57**: p. 13-21.
82. Gelman, R.A., B.R. Williams, and K.A. Piez, *Collagen fibril formation. Evidence for a multistep process.* J Biol Chem, 1979. **254**(1): p. 180-6.
83. Birk, D.E. and R.L. Trelstad, *Extracellular compartments in matrix morphogenesis: collagen fibril, bundle, and lamellar formation by corneal fibroblasts.* J Cell Biol, 1984. **99**(6): p. 2024-33.
84. Williams, E.M., et al., *Phenotypical features of an unique Irish family with severe autosomal recessive osteogenesis imperfecta.* Clin Genet, 1989. **35**(3): p. 181-90.
85. Wood, G.C., *The formation of fibrils from collagen solutions. 2. A mechanism of collagen-fibril formation.* Biochem J, 1960. **75**: p. 598-605.
86. Bernengo, J.C., B. Roux, and D. Herbage, *Electrical birefringence study of monodisperse collagen solutions.* Biopolymers, 1974. **13**(3): p. 641-7.
87. Comper, W.D. and A. Veis, *The mechanism of nucleation for in vitro collagen fibril formation.* Biopolymers, 1977. **16**(10): p. 2113-2131.

88. Comper, W.D. and A. Veis, *Characterization of nuclei in in vitro collagen fibril formation.* Biopolymers, 1977. **16**(10): p. 2133-42.
89. Silver, D., et al., *Helical model of nucleation and propagation to account for the growth of type I collagen fibrils from symmetrical pointed tips: a special example of self-assembly of rod-like monomers.* Proc Natl Acad Sci U S A, 1992. **89**(20): p. 9860-4.
90. Kadler, K.E., Y. Hojima, and D.J. Prockop, *Collagen fibrils in vitro grow from pointed tips in the C- to N-terminal direction.* Biochem J, 1990. **268**(2): p. 339-43.
91. Holmes, D.F., et al., *Growing tips of type I collagen fibrils formed in vitro are near-paraboloidal in shape, implying a reciprocal relationship between accretion and diameter.* Proc Natl Acad Sci U S A, 1992. **89**(20): p. 9855-9.
92. George, A. and A. Veis, *FTIRS in H2O demonstrates that collagen monomers undergo a conformational transition prior to thermal self-assembly in vitro.* Biochemistry, 1991. **30**(9): p. 2372-7.
93. Trus, B.L. and K.A. Piez, *Compressed microfibril models of the native collagen fibril.* Nature, 1980. **286**(5770): p. 300-1.
94. Prockop, D.J. and A. Fertala, *The collagen fibril: the almost crystalline structure.* J Struct Biol, 1998. **122**(1-2): p. 111-8.
95. Perret, S., et al., *Unhydroxylated triple helical collagen I produced in transgenic plants provides new clues on the role of hydroxyproline in collagen folding and fibril formation.* J Biol Chem, 2001. **276**(47): p. 43693-8.
96. Fraser, R.D., T.P. MacRae, and A. Miller, *Molecular packing in type I collagen fibrils.* J Mol Biol, 1987. **193**(1): p. 115-25.
97. Orgel, J.P., et al., *Microfibrillar structure of type I collagen in situ.* Proc Natl Acad Sci U S A, 2006. **103**(24): p. 9001-5.
98. Parkinson, J., K.E. Kadler, and A. Brass, *Simple physical model of collagen fibrillogenesis based on diffusion limited aggregation.* J Mol Biol, 1995. **247**(4): p. 823-31.
99. Ottani, V., et al., *Hierarchical structures in fibrillar collagens.* Micron, 2002. **33**(7-8): p. 587-96.
100. Vogel KG, T.J., *The effect of proteoglycans on the morphology of collagen fibrils formed in vitro.* Coll Relat Res. , 1987. **7**(2): p. 105-14.
101. Iozzo, R.V. and A.D. Murdoch, *Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function.* Faseb J, 1996. **10**(5): p. 598-614.
102. Iozzo, R.V., *Matrix proteoglycans: from molecular design to cellular function.* Annu Rev Biochem, 1998. **67**: p. 609-52.
103. Kobe, B. and J. Deisenhofer, *The leucine-rich repeat: a versatile binding motif.* Trends Biochem Sci, 1994. **19**(10): p. 415-21.
104. Cawston, T.E., et al., *The levels of collagenase, tissue inhibitor of metalloproteinases-1 (TIMP-1), collagenase approximately TIMP-1 complexes and glycosaminoglycan (GAG) in sequential samples of synovial fluid aspirated from patients with osteoarthritis.* Clin Exp Rheumatol, 1995. **13**(4): p. 431-7.

105. Krusius, T. and E. Ruoslahti, *Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA.* Proc Natl Acad Sci U S A, 1986. **83**(20): p. 7683-7.
106. Scott, J.E., *Proteoglycan-fibrillar collagen interactions.* Biochem J, 1988. **252**(2): p. 313-23.
107. Scott, J.E. and C.R. Orford, *Dermatan sulphate-rich proteoglycan associates with rat tail-tendon collagen at the d band in the gap region.* Biochem J, 1981. **197**(1): p. 213-6.
108. Uldbjerg, N. and C.C. Danielsen, *A study of the interaction in vitro between type I collagen and a small dermatan sulphate proteoglycan.* Biochem J, 1988. **251**(3): p. 643-8.
109. Rosenberg, A., K.W. Gratz, and H.F. Sailer, *Should titanium miniplates be removed after bone healing is complete?* Int J Oral Maxillofac Surg, 1993. **22**(3): p. 185-8.
110. Gosline, J.M. and R.E. Shadwick, *The mechanical properties of fin whale arteries are explained by novel connective tissue designs.* J Exp Biol, 1996. **199**(Pt 4): p. 985-97.
111. Goldoni, S., et al., *Biologically Active Decorin Is a Monomer in Solution.* J. Biol. Chem., 2004. **279**(8): p. 6606-6612.
112. Hulmes, D.J., et al., *Radial packing, order, and disorder in collagen fibrils.* Biophys J, 1995. **68**(5): p. 1661-70.
113. Gutsmann, T., et al., *Force spectroscopy of collagen fibers to investigate their mechanical properties and structural organization.* Biophys J, 2004. **86**(5): p. 3186-93.
114. Doillon, C.J., et al., *Collagen fiber formation in repair tissue: development of strength and toughness.* Coll Relat Res, 1985. **5**(6): p. 481-92.
115. Doillon, C.J., et al., *Collagen deposition during wound repair.* Scan Electron Microsc, 1985(Pt 2): p. 897-903.
116. Parry, D.A., *The molecular and fibrillar structure of collagen and its relationship to the mechanical properties of connective tissue.* Biophys Chem, 1988. **29**(1-2): p. 195-209.
117. Gutsmann, T., et al., *Evidence that collagen fibrils in tendons are inhomogeneously structured in a tubelike manner.* Biophys J, 2003. **84**(4): p. 2593-8.
118. Bozec, L. and M. Horton, *Topography and mechanical properties of single molecules of type I collagen using atomic force microscopy.* Biophys J, 2005. **88**(6): p. 4223-31.
119. John C. Thomas, G.C.F., *Dynamic light scattering from collagen solutions. II. Photon correlation study of the depolarized light.* Biopolymers, 1979. **18**(6): p. 1333-1352.
120. Cribb, A.M. and J.E. Scott, *Tendon response to tensile stress: an ultrastructural investigation of collagen:proteoglycan interactions in stressed tendon.* J Anat, 1995. **187 (Pt 2)**: p. 423-8.
121. Puxkandl, R., et al., *Viscoelastic properties of collagen: synchrotron radiation investigations and structural model.* Philos Trans R Soc Lond B Biol Sci, 2002. **357**(1418): p. 191-7.

122. Diamant, J., et al., *Collagen; ultrastructure and its relation to mechanical properties as a function of ageing.* Proc R Soc Lond B Biol Sci, 1972. **180**(60): p. 293-315.
123. Misof, K., G. Rapp, and P. Fratzl, *A new molecular model for collagen elasticity based on synchrotron X-ray scattering evidence.* Biophys J, 1997. **72**(3): p. 1376-81.
124. Nagase, H., R. Visse, and G. Murphy, *Structure and function of matrix metalloproteinases and TIMPs.* Cardiovasc Res, 2006. **69**(3): p. 562-73.
125. Visse, R. and H. Nagase, *Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry.* Circ Res, 2003. **92**(8): p. 827-39.
126. Ala-aho, R. and V.M. Kahari, *Collagenases in cancer.* Biochimie, 2005. **87**(3-4): p. 273-86.
127. Knauper, V., et al., *Biochemical characterization of human collagenase-3.* J Biol Chem, 1996. **271**(3): p. 1544-50.
128. Patterson, M.L., et al., *Specific collagenolysis by gelatinase A, MMP-2, is determined by the hemopexin domain and not the fibronectin-like domain.* FEBS Lett, 2001. **503**(2-3): p. 158-62.
129. Ohuchi, E., et al., *Membrane type 1 matrix metalloproteinase digests interstitial collagens and other extracellular matrix macromolecules.* J Biol Chem, 1997. **272**(4): p. 2446-51.
130. Garnero, P., et al., *The collagenolytic activity of cathepsin K is unique among mammalian proteinases.* J Biol Chem, 1998. **273**(48): p. 32347-52.
131. Chung, L., et al., *Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis.* Embo J, 2004. **23**(15): p. 3020-30.
132. Clark, I.M. and T.E. Cawston, *Fragments of human fibroblast collagenase. Purification and characterization.* Biochem J, 1989. **263**(1): p. 201-6.
133. Li, J., et al., *Structure of full-length porcine synovial collagenase reveals a C-terminal domain containing a calcium-linked, four-bladed beta-propeller.* Structure, 1995. **3**(6): p. 541-9.
134. Jozic, D., et al., *X-ray structure of human proMMP-1: new insights into procollagenase activation and collagen binding.* J Biol Chem, 2005. **280**(10): p. 9578-85.
135. Tam, E.M., et al., *Characterization of the distinct collagen binding, helicase and cleavage mechanisms of matrix metalloproteinase 2 and 14 (gelatinase A and MT1-MMP): the differential roles of the MMP hemopexin c domains and the MMP-2 fibronectin type II modules in collagen triple helicase activities.* J Biol Chem, 2004. **279**(41): p. 43336-44.
136. Miller, E.J., et al., *Specific cleavage of the native type III collagen molecule with trypsin. Similarity of the cleavage products to collagenase-produced fragments and primary structure at the cleavage site.* Arch Biochem Biophys, 1976. **173**(2): p. 631-7.
137. Miller, E.J., et al., *Cleavage of Type II and III collagens with mammalian collagenase: site of cleavage and primary structure at the NH2-terminal portion of the smaller fragment released from both collagens.* Biochemistry, 1976. **15**(4): p. 787-92.

138. Perumal, S., O. Antipova, and J.P. Orgel, *Collagen fibril architecture, domain organization, and triple-helical conformation govern its proteolysis.* Proc Natl Acad Sci U S A, 2008. **105**(8): p. 2824-9.
139. Fields, G.B., *A model for interstitial collagen catabolism by mammalian collagenases.* J Theor Biol, 1991. **153**(4): p. 585-602.
140. Brodsky, B. and E.F. Eikenberry, *Characterization of fibrous forms of collagen.* Methods Enzymol, 1982. **82 Pt A**: p. 127-74.
141. Kelly, J., et al., *Fibril-forming collagens in lamprey.* J Biol Chem, 1988. **263**(2): p. 980-7.
142. Uitto, J., *Collagen polymorphism: isolation and partial characterization of alpha 1(I)-trimer molecules in normal human skin.* Arch Biochem Biophys, 1979. **192**(2): p. 371-9.
143. Rojkind, M., M.A. Giambrone, and L. Biempica, *Collagen types in normal and cirrhotic liver.* Gastroenterology, 1979. **76**(4): p. 710-9.
144. Schillaci, R., C. Luparello, and S. Minafra, *Type I and I-trimer collagens as substrates for breast carcinoma cells in culture. Effect on growth rate, morphological appearance and actin organization.* Eur J Cell Biol, 1989. **48**(1): p. 135-41.
145. Pucci Minafra, I., et al., *Collagen changes in the ductal infiltrating (scirrhous) carcinoma of the human breast. A possible role played by type I trimer collagen on the invasive growth.* J Submicrosc Cytol, 1986. **18**(4): p. 795-805.
146. Rojkind, M. and R. Perez-Tamayo, *Liver fibrosis.* Int Rev Connect Tissue Res, 1983. **10**: p. 333-93.
147. Grant, S.F., et al., *Reduced bone density and osteoporosis associated with a polymorphic Sp1 binding site in the collagen type I alpha 1 gene.* Nat Genet, 1996. **14**(2): p. 203-5.
148. Mann, V., et al., *A COL1A1 Sp1 binding site polymorphism predisposes to osteoporotic fracture by affecting bone density and quality.* J Clin Invest, 2001. **107**(7): p. 899-907.
149. Mann, V. and S.H. Ralston, *Meta-analysis of COL1A1 Sp1 polymorphism in relation to bone mineral density and osteoporotic fracture.* Bone, 2003. **32**(6): p. 711-7.
150. Ralston, S.H., et al., *Large-scale evidence for the effect of the COLIA1 Sp1 polymorphism on osteoporosis outcomes: the GENOMOS study.* PLoS Med, 2006. **3**(4): p. e90.
151. Kuznetsova, N., D.J. McBride, Jr., and S. Leikin, *Osteogenesis imperfecta murine: interaction between type I collagen homotrimers.* J Mol Biol, 2001. **309**(3): p. 807-15.
152. Nicholls, A.C., et al., *The clinical features of homozygous alpha 2(I) collagen deficient osteogenesis imperfecta.* J Med Genet, 1984. **21**(4): p. 257-62.
153. Sasaki, N. and S. Odajima, *Stress-strain curve and Young's modulus of a collagen molecule as determined by the X-ray diffraction technique.* J Biomech, 1996. **29**(5): p. 655-8.
154. Pace, J.M., et al., *Defective C-propeptides of the Pro{alpha}2(I) Chain of Type I Procollagen Impede Molecular Assembly and Result in Osteogenesis Imperfecta.* J Biol Chem, 2008. **283**(23): p. 16061-7.

155. Miles, C.A., et al., *The role of the alpha2 chain in the stabilization of the collagen type I heterotrimer: a study of the type I homotrimer in oim mouse tissues.* J Mol Biol, 2002. **321**(5): p. 797-805.
156. McBride, D.J., Jr., et al., *Self-assembly into fibrils of a homotrimer of type I collagen.* Matrix, 1992. **12**(4): p. 256-63.
157. Minafra, S., et al., *Collagen biosynthesis by a breast carcinoma cell strain and biopsy fragments of the primary tumour.* Cell Biol Int Rep, 1988. **12**(10): p. 895-905.
158. Minafra, S., et al., *Collagen composition in the ductal infiltrating carcinoma of human breast.* Cell Biol Int Rep, 1984. **8**(1): p. 79-85.
159. Yamagata, S. and T. Yamagata, *FBJ virus-induced osteosarcoma contains type I, type I trimer, type III as well as type V collagens.* J Biochem, 1984. **96**(1): p. 17-26.
160. Chipman, S.D., et al., *Defective pro alpha 2(I) collagen synthesis in a recessive mutation in mice: a model of human osteogenesis imperfecta.* Proc Natl Acad Sci U S A, 1993. **90**(5): p. 1701-5.
161. Helseth, D.L., Jr. and A. Veis, *Collagen self-assembly in vitro. Differentiating specific telopeptide-dependent interactions using selective enzyme modification and the addition of free amino telopeptide.* J Biol Chem, 1981. **256**(14): p. 7118-28.
162. Comper, W.D. and A. Veis, *The mechanism of nucleation for in vitro collagen fibril formation.* Biopolymers, 1977. **16**: p. 2113-2131.
163. Kuznetsova, N. and S. Leikin, *Does the triple helical domain of type I collagen encode molecular recognition and fiber assembly while telopeptides serve as catalytic domains? Effect of proteolytic cleavage on fibrillogenesis and on collagen-collagen interaction in fibers.* J Biol Chem, 1999. **274**(51): p. 36083-8.
164. Veis, A. and K. Payne, *Collagen fibrillogenesis*, in *Collagen*, M.E. Nimni, Editor. 1988, CRC Press: Boca Raton, Florida. p. 113-137.
165. Kadler, K.E., et al., *Assembly of type I collagen fibrils de novo by the specific enzymic cleavage of pC collagen. The fibrils formed at about 37 degrees C are similar in diameter, roundness, and apparent flexibility to the collagen fibrils seen in connective tissue.* Ann N Y Acad Sci, 1990. **580**: p. 214-24.
166. Mertz, E.L. and S. Leikin, *Interactions of inorganic phosphate and sulfate anions with collagen.* Biochemistry, 2004. **43**(47): p. 14901-12.
167. Phillips, C.L., et al., *Novel collagen glomerulopathy in a homotrimeric type I collagen mouse (oim).* Kidney Int, 2002. **62**(2): p. 383-91.
168. Leikina, E., et al., *Type I collagen is thermally unstable at body temperature.* Proc Natl Acad Sci U S A, 2002. **99**(3): p. 1314-8.
169. Hiltner, A., J.J. Cassidy, and E. Baer, *Mechanical Properties of Biological Polymers.* Annual Review of Materials Science, 1985. **15**(1): p. 455-482.
170. Comninou, M. and I.V. Yannas, *Dependence of stress-strain nonlinearity of connective tissues on the geometry of collagen fibers.* J Biomech, 1976. **9**(7): p. 427-33.

171. Barrett, S.R., et al., *Experimental measurement of the mechanical properties of carotid atherothrombotic plaque fibrous cap.* J Biomech, 2009. **42**(11): p. 1650-5.
172. Martin, R.B., D.B. Burr, and N.A. Sharkey, *Skeletal tissue mechanics*. 1998, New York: Springer. xiv, 392 p.
173. in t Veld, P.J. and M.J. Stevens, *Simulation of the Mechanical Strength of a Single Collagen Molecule.* 2008. **95**(1): p. 33-39.
174. Sun, Y.L., et al., *Direct quantification of the flexibility of type I collagen monomer.* Biochem Biophys Res Commun, 2002. **295**(2): p. 382-6.
175. Yang, L., et al., *Micromechanical bending of single collagen fibrils using atomic force microscopy.* J Biomed Mater Res A, 2007. **82**(1): p. 160-8.
176. Mickel, W., et al., *Robust Pore Size Analysis of Filamentous Networks from Three-Dimensional Confocal Microscopy.* 2008. **95**(12): p. 6072-6080.
177. Knapp, D.M., et al., *Rheology of reconstituted type I collagen gel in confined compression.* Journal of Rheology, 1997. **41**(5): p. 971-993.
178. Velegol, D. and F. Lanni, *Cell Traction Forces on Soft Biomaterials. I. Microrheology of Type I Collagen Gels.* 2001. **81**(3): p. 1786-1792.
179. Wenger, M.P., et al., *Mechanical properties of collagen fibrils.* Biophys J, 2007. **93**(4): p. 1255-63.
180. Parkinson, J., et al., *The mechanical properties of simulated collagen fibrils.* J Biomech, 1997. **30**(6): p. 549-54.
181. Fratzl, P., et al., *Fibrillar Structure and Mechanical Properties of Collagen.* Journal of Structural Biology, 1998. **122**(1-2): p. 119-122.
182. Han, E.M., N.V. Kuznetsova, A.M. DeRidder, M.B. Sutter, D.J. McBride, C.L. Phillips, U. Schwarze, J.M. Pace, P.H. Byers, R. Visse, H. Nagase, and S. Leikin, *Molecular mechanism of the resistance of a homotrimeric type I collagen isoform to mammalian collagenases.* PNAS (in preparation), 2009.
183. Makareeva, S.H., J.C. Vera, D.L. Sackett, K. Holmbeck, D.J. McBride, C.L. Phillips, U. Schwarze, J.M. Pace, P.H. Byers, R. Visse, H. Nagase, and S. Leikin, *Cancer cells secrete MMP-resistant isoform of type I collagen.* 2009.
184. Bauer, E.A., et al., *Enhanced collagenase production by fibroblasts derived from human basal cell carcinomas.* Cancer Res, 1979. **39**(11): p. 4594-9.
185. Zaman, M.H., et al., *Migration of tumor cells in 3D matrices is governed by matrix stiffness along with cell-matrix adhesion and proteolysis.* Proc Natl Acad Sci U S A, 2006. **103**(29): p. 10889-94.
186. Tony, Y., et al., *Effects of substrate stiffness on cell morphology, cytoskeletal structure, and adhesion.* Cell Motility and the Cytoskeleton, 2005. **60**(1): p. 24-34.
187. Pelham, R.J. and Y.-l. Wang, *Cell locomotion and focal adhesions are regulated by substrateâ€‰flexibility.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(25): p. 13661-13665.
188. Han, S., et al., *Segregation of type I collagen homo- and heterotrimers in fibrils.* J Mol Biol, 2008. **383**(1): p. 122-32.
189. Landau, L.D., E.M. Lifshiëtìs, and L.P. Pitaevski¸, *Statistical physics*. 3rd ed. Course of theoretical physics ; v. 5, 9. 1980, Oxford [England]: Butterworth-Heinemann.

190. Gittes, F., et al., *Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape.* J Cell Biol, 1993. **120**(4): p. 923-34.
191. Lewis, G. and K.M. Shaw, *Tensile properties of human tendo Achillis: effect of donor age and strain rate.* J Foot Ankle Surg, 1997. **36**(6): p. 435-45.
192. Eppell, S.J., et al., *Nano measurements with micro-devices: mechanical properties of hydrated collagen fibrils.* J R Soc Interface, 2006. **3**(6): p. 117-21.
193. Wren, T.A., et al., *Mechanical properties of the human achilles tendon.* Clin Biomech (Bristol, Avon), 2001. **16**(3): p. 245-51.
194. Howard, J., *Mechanics of motor proteins and the cytoskeleton*. 2001, Sunderland, Mass.: Sinauer Associates. xvi, 367 p.
195. Sun, Y.-L., Z.-P. Luo, and K.-N. An, *Stretching Short Biopolymers Using Optical Tweezers.* Biochemical and Biophysical Research Communications, 2001. **286**(4): p. 826-830.
196. Fratzl, P. and SpringerLink (Online service). *Collagen Structure and Mechanics*. 2008 [cited; Available from: http://alias.libraries.psu.edu/eresources/proxy/login?url=http://dx.doi.org/10.1007/978-0-387-73906-9.
197. Makareeva, E., et al., *Structural heterogeneity of type I collagen triple helix and its role in osteogenesis imperfecta.* J Biol Chem, 2008. **283**(8): p. 4787-98.
198. Ulrich, T.A., E.M. de Juan Pardo, and S. Kumar, *The mechanical rigidity of the extracellular matrix regulates the structure, motility, and proliferation of glioma cells.* Cancer Res, 2009. **69**(10): p. 4167-74.
199. Alexander, N.R., et al., *Extracellular matrix rigidity promotes invadopodia activity.* Curr Biol, 2008. **18**(17): p. 1295-9.
200. Birkedal-Hansen, H., et al., *Matrix metalloproteinases: a review.* Crit Rev Oral Biol Med, 1993. **4**(2): p. 197-250.
201. Paul, S.N., S.-E. Ramon, and M.S. Collin, *Do collagenases unwind triple-helical collagen before peptide bond hydrolysis? Reinterpreting experimental observations with mathematical models.* Proteins: Structure, Function, and Bioinformatics, 2008. **70**(4): p. 1154-1161.
202. Overall, C.M., *Molecular determinants of metalloproteinase substrate specificity: matrix metalloproteinase substrate binding domains, modules, and exosites.* Mol Biotechnol, 2002. **22**(1): p. 51-86.
203. Welgus, H.G., J.J. Jeffrey, and A.Z. Eisen, *The collagen substrate specificity of human skin fibroblast collagenase.* J Biol Chem, 1981. **256**(18): p. 9511-5.
204. Cabral, W.A., et al., *Type I collagen triplet duplication mutation in lethal osteogenesis imperfecta shifts register of alpha chains throughout the helix and disrupts incorporation of mutant helices into fibrils and extracellular matrix.* J Biol Chem, 2003. **278**(12): p. 10006-12.
205. Phillips, C.L., et al., *Type I collagen glomerulopathy: oim (COL1A2 deficient) mice.* Matrix Biology, 2006. **25**(Supplement 1): p. S66-S66.

206. Liu, X., et al., *A targeted mutation at the known collagenase cleavage site in mouse type I collagen impairs tissue remodeling.* J. Cell Biol., 1995. **130**(1): p. 227-237.
207. Beare, A.H.M., et al., *Severely Impaired Wound Healing in the Collagenase-Resistant Mouse.* 2003. **120**(1): p. 153-163.
208. Chiusaroli, R., et al., *Collagenase Cleavage of Type I Collagen Is Essential for Both Basal and Parathyroid Hormone (PTH)/PTH-Related Peptide Receptor-Induced Osteoclast Activation and Has Differential Effects on Discrete Bone Compartments.* Endocrinology, 2003. **144**(9): p. 4106-4116.
209. Zhao, W., et al., *Bone resorption induced by parathyroid hormone is strikingly diminished in collagenase-resistant mutant mice.* The Journal of Clinical Investigation, 1999. **103**(4): p. 517-524.
210. Holmbeck, K., et al., *MT1-MMP-deficient mice develop dwarfism, osteopenia, arthritis, and connective tissue disease due to inadequate collagen turnover.* Cell, 1999. **99**(1): p. 81-92.
211. Makareeva, E., et al., *Molecular Mechanism of {alpha}1(I)-Osteogenesis Imperfecta/Ehlers-Danlos Syndrome: UNFOLDING OF AN N-ANCHOR DOMAIN AT THE N-TERMINAL END OF THE TYPE I COLLAGEN TRIPLE HELIX.* J. Biol. Chem., 2006. **281**(10): p. 6463-6470.
212. Cunningham, L.W. and D.W. Frederiksen, *Structural and contractile proteins*. 1982, New York: Academic Press. v.
213. Murphy, G., et al., *The role of the C-terminal domain in collagenase and stromelysin specificity.* J Biol Chem, 1992. **267**(14): p. 9612-8.
214. Comper, W.D., *Extracellular matrix*. 1996, Amsterdam: Harwood Academic Publishers.