Matrix Biology 34 (2014) 161–169

Contents lists available at ScienceDirect

Matrix Biology

journal homepage: www.elsevier.com/locate/matbio

Mechanical recruitment of N- and C-crosslinks in collagen type I

Albert L. Kwansa [a,b], Raffaella De Vita [b], Joseph W. Freeman [a,c,*]

[a] Virginia Polytechnic Institute and State University, Virginia Tech-Wake Forest School of Biomedical Engineering and Sciences, Musculoskeletal Tissue Regeneration Laboratory, 330 Kelly Hall, 325 Stanger St., Blacksburg, VA 24061, USA
[b] Virginia Polytechnic Institute and State University, Department of Engineering Science and Mechanics, Mechanics of Soft Biological Systems Laboratory, 202 Norris Hall, 495 Old Turner St., Blacksburg, VA 24061, USA
[c] Rutgers University, Department of Biomedical Engineering, Musculoskeletal Tissue Regeneration Laboratory, 317 BME Building, 599 Taylor Rd., Piscataway, NJ 08854, USA

ARTICLE INFO

Article history:
Received 28 June 2013
Received in revised form 25 October 2013
Accepted 25 October 2013
Available online 21 November 2013

Keywords:
Collagen type I
Crosslink
Molecular dynamics
Molecular model
Telopeptide

ABSTRACT

Collagen type I is an extracellular matrix protein found in connective tissues such as tendon, ligament, bone, skin, and the cornea of the eyes, where it functions to provide tensile strength; it also serves as a scaffold for cells and other extracellular matrix components. A single collagen type I molecule is composed of three amino acid chains that form a triple helix for most of the molecule's length; non-triple-helical extensions called N- and C-telopeptides are located at the amino/N-terminal and carboxy/C-terminal ends of the molecule, respectively. In two of the three chains, the C-telopeptide has been reported to possess a hair-pin/hook conformation, while the three N-telopeptides display a more extended structure. These telopeptides are crucial for the formation of enzymatic covalent crosslinks that form in collagens near their N- and C-ends. Such crosslinks provide structural integrity, strength, and stiffness to collagenous tissues. However, deformation mechanisms of N- and C-crosslinks and functional roles for the N- and C-telopeptide conformations are not yet well known. By performing molecular dynamics simulations, we demonstrated that two dehydro-hydroxylysino-norleucine crosslinks, positioned at the N- and C-crosslinking sites, exhibited a two-stage response to the mechanical deformation of their parent molecules. We observed that the N-crosslink served as the first responder to mechanical deformation, followed by the C-crosslink. The results of our simulations suggest a mechanical recruitment mechanism for N- and C-crosslinks. Understanding this mechanism will be crucial for the development of larger-scale predictive models of the mechanical behavior of native collagenous tissues, engineered tissues, and collagen-based materials.

1. Introduction

Collagens are extracellular matrix (ECM) proteins that are found in nearly all eukaryotic organisms except for plants and protozoa (Urich, 1994). In mammals, collagens are especially abundant and can comprise up to one-third of all protein in the body by weight (Williams, 1978). There are approximately 27 different types of collagens that have been identified (von der Mark, 2006); type I collagen is the most prevalent and is found in vertebrate connective tissues such as tendon, ligament, bone, skin, and the cornea of the eyes (von der Mark, 2006). Type I collagen functions to provide tensile strength to these connective tissues, and it serves as a structural framework for cells and other ECM components such as fibronectin, proteoglycans, and bone mineral (Sweeney et al., 2008).

Type I collagen is a rod-shaped molecule with a length of ~300 nm and a diameter of ~1.5 nm (Birk and Bruckner, 2005). It is composed of three amino acid chains (often called α-chains), each containing just over 1000 amino acids; there are two α1 chains and one α2 chain. The entire molecule is composed of an N-telopeptide domain at the beginning, a triple-helical domain, and a C-telopeptide domain at the end (The UniProt Consortium, 2012). The N-telopeptides have been reported to display an extended coil structure; and the longer α1 C-telopeptides a hair-pin/hook structure (Orgel et al., 2006).

Fibril-forming collagens (e.g., types I, II, III, V, and XI) are able to assemble into structures called fibrils, which are stabilized by covalent bonds called crosslinks. There are enzymatic crosslinks and non-enzymatic crosslinks. Enzymatic crosslink formation is regulated by lysyl hydroxylase and lysyl oxidase enzymes. Lysyl hydroxylases are intracellular enzymes that convert specific lysines to hydroxylysines, and lysyl oxidases are extracellular enzymes that convert the side chain ε-amino group of telopeptide lysines and hydroxylysines into an aldehyde group (Avery and Bailey, 2008; Kagan and Ryvkin, 2011). These aldehyde groups can react readily with other ε-amino groups to form these crosslinks (Knott and Bailey, 1998). There are several types

Abbreviations: α, alpha value; ANOVA, analysis of variance; Cα, alpha carbon; CHARMM/CHARMm, Chemistry at HARvard Molecular Mechanics; deH-HLNL, dehydro-hydroxylysino-norleucine; deH-LNL, dehydro-lysino-norleucine; ECM, extracellular matrix; GBIS, generalized Born implicit solvent; NAMD, NAnoscale Molecular Dynamics; *P*, p-value; RMS, root-mean-square; SMD, steered molecular dynamics; TCL, tool command language; VMD, Visual Molecular Dynamics.

* Corresponding author at: Rutgers University, Department of Biomedical Engineering, Musculoskeletal Tissue Regeneration Laboratory, 317 BME Building, 599 Taylor Rd., Piscataway, NJ 08854, USA. Tel.: +1 848 445 6595.
E-mail address: joseph.freeman@rutgers.edu (J.W. Freeman).

0945-053X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.matbio.2013.10.012

of enzymatic crosslinks that can form (e.g., immature, divalent crosslinks and mature, trivalent crosslinks), depending upon factors such as tissue type, age, and health (Eyre et al., 1984a, 1984b; Fujii et al., 1994; Sims and Bailey, 1992). Non-enzymatic crosslinks tend to form in old age or in certain disease states (e.g., diabetes) outside of the regulation of enzymes (Avery and Bailey, 2008). The focus of this paper is on one of the enzymatic crosslinks called dehydro-hydroxylysino-norleucine (deH-HLNL). deH-HLNL involves the reaction of a lysine aldehyde (from an N- or C-telopeptide domain) to a hydroxylysine (from the triple-helical domain) (see Fig. 1 for the chemical structure of deH-HLNL and its precursor amino acids).

Steered molecular dynamics (SMD) and moving constraints are two extensions of molecular dynamics that allow for constant force pulling (e.g., constant force SMD) or constant velocity pulling (e.g., constant velocity SMD and moving constraints). In moving constraints and constant velocity SMD, each pulled atom or the center of mass of a group of pulled atoms, respectively, is connected through a virtual spring (k = stiffness) to a virtual atom that moves at a constant velocity (v = velocity) (Bhandarkar et al., 2011). The use of virtual springs and virtual atoms is meant to provide an analog to position-controlled atomic force microscopy (AFM) (Isralewitz et al., 2001a). Constant velocity SMD and its related techniques have been used to simulate the mechanical functions of proteins and protein unfolding pathways (Isralewitz et al., 2001b), and to predict the Young's modulus of collagen-like molecules (Gautieri et al., 2009; Lorenzo and Caffarena, 2005).

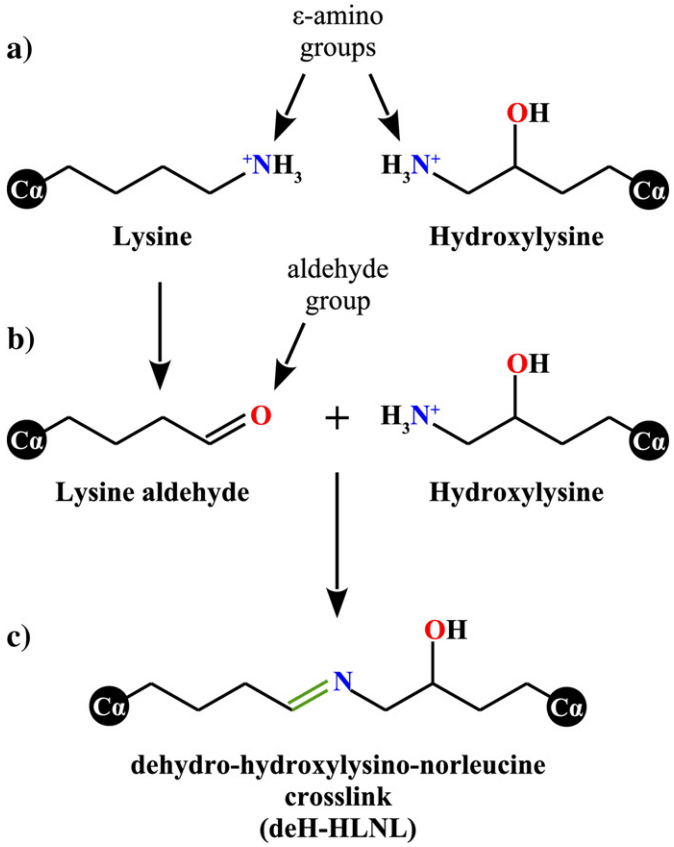

Fig. 1. A schematic of the formation of the enzymatic crosslink dehydro-hydroxylysino-norleucine (deH-HLNL) from the amino acids lysine and hydroxylysine (a modification of lysine). (a) shows the precursor amino acids lysine and hydroxylysine. (b) shows the product of the conversion of lysine to a lysine aldehyde. (c) shows the product of the reaction between this lysine aldehyde and hydroxylysine. In the case of deH-HLNL, the lysine is derived from one of the N- or C-telopeptide domains, while the hydroxylysine is derived from the main triple-helical domain. In the online color version of this figure, colors are used to further distinguish oxygen atoms (red), nitrogen atoms (blue), and the newly formed covalent bond within the deH-HLNL crosslink (green) (i.e., a carbon-nitrogen double bond).

The pulling velocity and the stiffness of the virtual springs have been shown to influence constant velocity pulling simulations. For instance, faster pulling rates can lead to over-estimated mechanical properties (Gautieri et al., 2009). Furthermore, overly compliant virtual springs can lead to under-estimated mechanical properties, and overly stiff virtual springs can result in increased numerical noise (Lorenzo and Caffarena, 2005). One, therefore, typically seeks a sufficiently reduced pulling velocity (i.e., closer to what might be employed experimentally or experienced physiologically) and sufficiently stiff virtual springs. There is a balance, however, that must be sought when choosing a pulling velocity, since a reduced pulling velocity necessitates a greater number of simulation timesteps to reach the same level of deformation, which increases the computational demands of one's simulation (e.g., time and/or computing resources).

Collagen and collagen-like molecules have been investigated through experiment and simulation; however, only a few investigations have been carried out on the mechanism through which collagen crosslinks respond to deformation. Uzel and Buehler conducted constant velocity SMD simulations (k = 4000 kJ/mol/nm^2 = ~9.56 kcal/mol/Å^2 (Gautieri et al., 2009); v = 1 m/s) with a 30-nm-long collagen type I molecular model containing an enzymatic C-terminal lysine-lysine crosslink called dehydro-lysino-norleucine (deH-LNL). They reported that this C-crosslink exhibited an initial delayed response due to the unfolding and straightening of the C-telopeptide. It was also reported that the crosslink contributed more to load-bearing at higher levels of deformation (Uzel and Buehler, 2011). In another study, Bourne and Torzilli investigated how perpendicular forces applied to a crosslink affect the conformation of its parent collagen molecule. Constant velocity SMD (k = 1 kcal/mol/Å^2; v = not reported) was used to pull on a non-enzymatic crosslink precursor amino acid (arginine) in a direction perpendicular to the long axis of a collagen-like molecule. It was reported that the collagen-like molecule offered little resistance to the perpendicular forces applied, which led to molecular bending and conformational disruption of the triple helix before covalent failure might be expected. It was thus suggested that these conformational disruptions, in response to crosslink loading, present an additional mechanism of damage within collagen proteins (Bourne and Torzilli, 2011). Interesting and important insights have been reported in these studies that have already been employed in the development of a rheological model of a crosslink's load-deformation behavior (Uzel and Buehler, 2011) and a finite-element model of unmineralized and mineralized fibrillar collagen (Hambli and Barkaoui, 2012).

In order to contribute to this area of research, we sought to investigate the influence of crosslink location (i.e., N vs. C-crosslinking site) using deH-HLNL as a model crosslink. deH-HLNL is an enzymatic lysine-hydroxylysine crosslink that has been detected in connective tissues such as skin (Saito et al., 1997; Sims and Bailey, 1992), tendon, and ligament (Fujii et al., 1994). deH-HLNL is especially abundant in developing (i.e., immature) and healing tissues. We used molecular dynamics to model the deformation of a ~23-nm-long molecular model in representation of the ends of two crosslinked collagen type I molecules with two deH-HLNL crosslinks positioned at the N- and C-crosslinking sites; these simulations were designed to model modes of molecular deformation that have been proposed through the X-ray diffraction of bovine Achilles tendons (Sasaki and Odajima, 1996).

2. Results and discussion

2.1. Higher spring constants improve the accuracy of the observed pulling velocity

In order to select an appropriate spring constant for the virtual springs, a spring constant study was conducted (k = 0.01, 0.1, 1, 10, 100, 1000, 9999 kcal/mol/Å^2 with v = 100 m/s). We found that stiffer

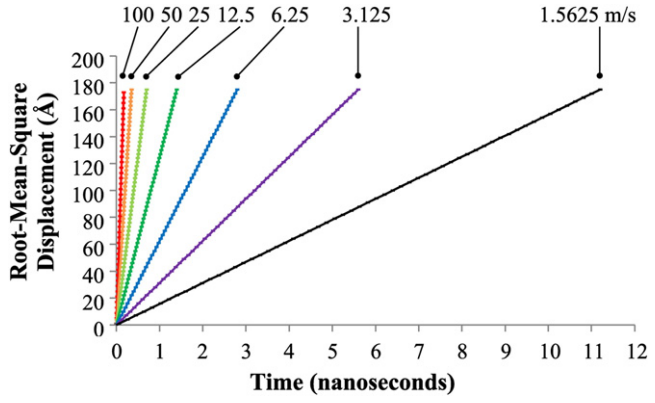

Fig. 2. Root-mean-square displacement vs. time data for different pulling velocities. This data was used to compute observed velocities (observed velocity = slope from linear fitting) and percent errors of the observed velocities (% error = (observed − expected) / expected * 100%). The x- and y-axes are plotted on log and linear scales, respectively. Each data point represents a mean ± s.d (n = 3).

virtual springs (i.e., higher spring constants) improved the accuracy of the observed pulling velocity; the value of each observed pulling velocity was based on the slope of the root-mean-square (RMS) displacement of the pulled atoms vs. time data (Fig. 2), and the error associated with each observed velocity was calculated as the difference between the observed and set velocities divided by the set velocity (i.e., $v = 100$ m/s).

The velocity errors were found to be statistically different between spring constants of 0.01, 0.1, 1, 10, and 100, but not between 100, 1000, and 9999 kcal/mol/Å^2 (Welch's ANOVA, $P < 0.0001$; Student's t-tests, $\alpha = 0.0024$) (Fig. 3). We further analyzed the numerical noise associated with the spring constants; however, no conclusive significant differences were found (Welch's ANOVA, $P = 0.0085$; Student's t-tests, all $P > \alpha = 0.0024$). Thus, 9999 kcal/mol/Å^2 was selected, given the order of magnitude reduction in error between 100, 1000, and 9999 kcal/mol/Å^2 (Fig. 3). Statistical power values for the velocity error and numerical noise analyses were found to be ~1.00 and 0.695, respectively, with sample sizes of three for each group (n = 3).

2.2. Pulling direction influences crosslink alignment, strain, and strain energy

The pulling direction (i.e., N- vs C-direction) was then investigated to determine if the selection of a particular direction affected the simulation results. The deH-HLNL crosslink, employed in our simulations, involves a link between two amino acids, leading to two alpha carbon (Cα) atoms per deH-HLNL crosslink (Fig. 1). The following output data was obtained from the simulations: crosslink alignment, crosslink engineering strain, crosslink strain energy, overall engineering strain, and

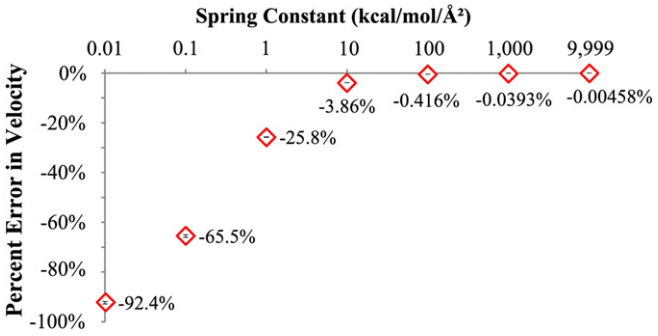

Fig. 3. Percent error in the observed pulling velocity at 100 m/s. The x- and y-axes are plotted on log and linear scales, respectively, and the value adjacent to each data point denotes the y-axis value. Negative errors, here, indicate that the observed velocity was less than the expected velocity. Each data point represents a mean ± s.d (n = 3).

overall strain energy (see Section 3.5). For the remainder of this manuscript, the phrases "crosslink strain" and "overall strain" have been used to refer to "crosslink engineering strain" and "overall engineering strain", respectively.

It was found that the pulling direction produced a bias in crosslink alignment, crosslink strain, and crosslink strain energy. When pulling in the C-direction (Fig. 4a), the crosslink closer to the pulled atoms (C-crosslink) began to align, strain, and store energy at lower overall strains relative to the crosslink located further from the pulled atoms (N-crosslink). When pulling in the N-direction (Fig. 4b), the reverse was observed. This bias was less prominent with slower pulling velocities; it was found that crosslink alignment was the least sensitive to the pulling velocity, followed by crosslink strain, and then by crosslink strain energy. At our slowest pulling velocity of 1.5625 m/s, the bias was essentially non-existent for crosslink strain energy and crosslink strain, but it was still present for crosslink alignment. Thus, data acquired from pulling in the N- or C- directions were averaged to correct for the remaining bias on crosslink alignment. These observations highlight the importance of one's chosen pulling direction and pulling velocity. Thus, it would be worthwhile to conduct a check of one's simulation to determine if such a bias exists for one's data of interest.

2.3. Pulling velocity influences crosslink alignment, crosslink strain, and strain energy

The pulling velocity ($v = 100$, 50, 25, 12.5, 6.25, 3.125, and 1.5625 m/s) was found to have a marked influence on the overall strain energy response of the ~23-nm-long molecular model and on the individual N- and C-crosslinks. Increased velocities led to increased overall strain energy (Fig. 5) and crosslink strain energy (Inline Supplementary Figs. S1 and S2). The effect of the pulling velocity became less pronounced as the velocity was reduced. For instance, it can be seen qualitatively that there are small differences for the overall strain energy and crosslink strain energy data between 6.25, 3.125, and 1.5625 m/s, but there are notable differences for the higher velocities (Fig. 5 and Inline Supplementary Figs. S1 and S2). These results show that the sensitivity to the pulling velocity was dependent upon the velocity regime (i.e., higher vs. lower velocities), which is in agreement with Gautieri et al. (2009).

Inline Supplementary Figs. S1 and S2 can be found online at http://dx.doi.org/10.1016/j.matbio.2013.10.012.

It was observed that the crosslink strains were elevated with higher velocities; however, the starting and ending strains (i.e., crosslink strains at 0% and 75% overall strain, respectively) were generally independent of the pulling velocity (Inline Supplementary Figs. S3 and S4). This shows that the crosslinks were strained or deformed more readily at higher velocities, while reduced velocities allowed for an initial lag in crosslink strain at lower overall strains (Inline Supplementary Figs. S3 and S4).

Inline Supplementary Figs. S3 and S4 can be found online at http://dx.doi.org/10.1016/j.matbio.2013.10.012.

Interestingly, lower velocities led to more alignment in the crosslinks. At higher velocities, the crosslinks displayed a three-phase behavior that resembled an inverse sigmoidal function. This three-phase behavior was comprised of initial alignment (an initial downward sloping region), a region of a nearly constant angle (a flatter region), and further alignment (another downward sloping region) (Inline Supplementary Figs. S5 and S6). The length of the constant-angle region became shorter as the velocity was reduced. These findings suggest that reduced velocities provided the crosslinks with more time to adjust (e.g., bond lengths, bond angles, bond dihedrals, and improper dihedrals) in order to achieve lower, more favorable levels of crosslink strain energy (Inline Supplementary Figs. S1 and S2) and greater alignment (Inline Supplementary Figs. S5 and S6).

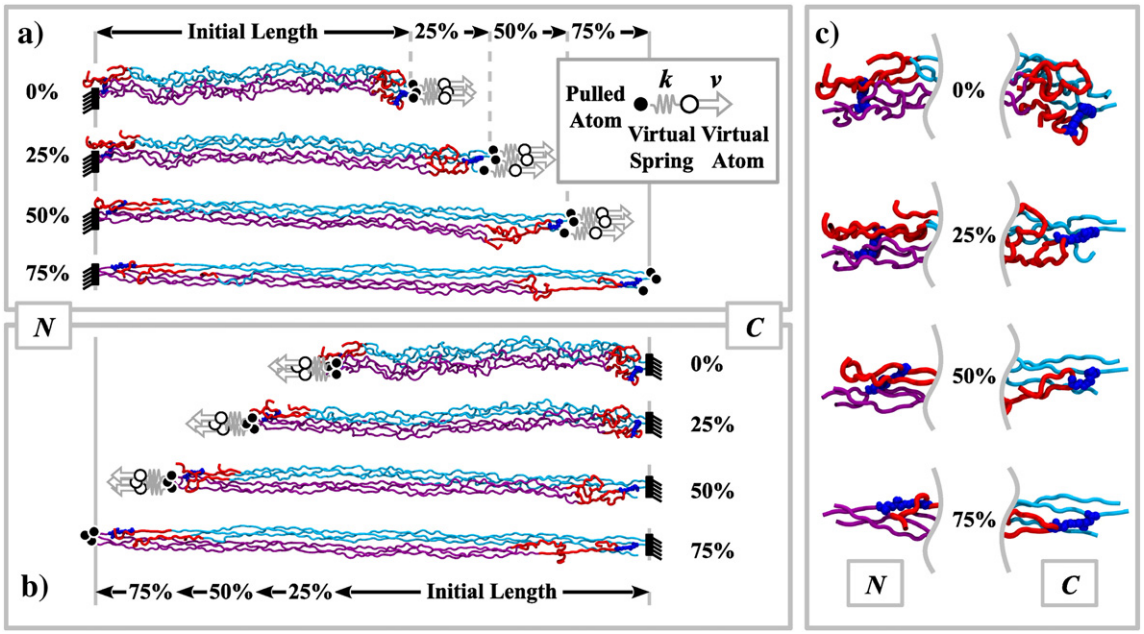

Fig. 4. Snapshot images of the model being pulled in the (a) C- or (b) N-direction at 1.5625 m/s. Triple-helical domains (cyan = upper segment; purple = lower segment), telopeptides (red), two crosslinks (blue), three fixed atoms (thick vertical line with diagonal hash lines), three pulled atoms (closed circles), and three virtual atoms (open circles). "N" and "C" represent the amino/N-terminal and carboxy/C-terminal directions. A schematic of the moving constraints technique is shown as an inset in (a) (k = virtual spring stiffness; v = virtual atom velocity). Close-up views of the N- and C-ends are shown in (c); these close-up views correspond to (a). Our molecular models were in representation of the ends of two collagen type I molecules covalently interconnected with two deH-HLNL crosslinks. Each of the two crosslinks was an "immature" or "divalent" crosslink only capable of interconnecting two amino acids each, namely, a lysine aldehyde from a telopeptide domain and a hydroxylysine from the triple-helical domain. Thus, for all of the deH-HLNL crosslinks depicted in this figure (colored in blue in the online version of the manuscript), only two amino acid chains are interconnected per crosslink. There are, however, other types of collagen crosslinks that may link more than two amino acid chains (e.g., pyridinolines, pyrroles, and histidine-based crosslinks); such crosslinks were not investigated for the work described herein.

Inline Supplementary Figs. S5 and S6 can be found online at http://dx.doi.org/10.1016/j.matbio.2013.10.012.

2.4. N- and C-crosslinks display a two-stage mechanical recruitment mechanism

The N- and C-crosslinks aligned from being initially more perpendicular to the pulling direction (~90° and ~60°, respectively) to being more parallel at 75% overall strain (~10 ± 1° and ~21 ± 2°, respectively, at 1.5625 m/s) (mean ± s.d.). It was observed qualitatively that the change in N- and C-crosslink alignment appeared to be fairly linear; thus, a linear fit of crosslink angle vs. overall strain was conducted leading to R^2 values of 0.96 and 0.86, respectively, for the N- and C-crosslinks. Slopes computed from these linear fits were used to quantify the rate of alignment with increasing overall strain. It was found that the N-crosslink aligned ~2× faster than the C-crosslink, and that by 75% overall strain, the N-crosslink displayed ~2× greater alignment (Fig. 6a).

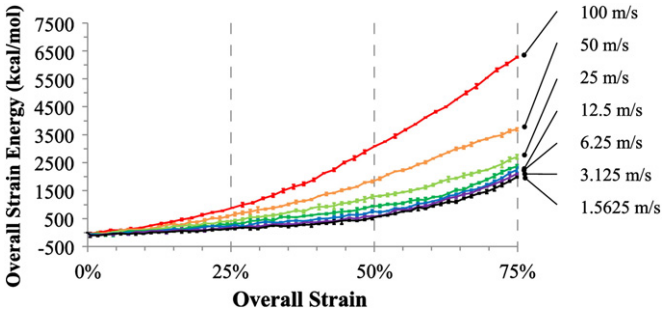

Fig. 5. The influence of the pulling velocity on overall strain energy. Overall strain represents the engineering strain of the ~23-nm-long molecular model and was defined as the change in length over the initial length. Overall strain energy represents the change in total potential energy relative to the unstrained state. Each data point represents a mean ± s.d (n = 3).

N- and C-crosslink strain was initially similar up to ~40% overall strain, after which the N-crosslink reached ~2× that of the C-crosslink by 75% overall strain (Fig. 6b). Strain energy in the N- and C-crosslinks

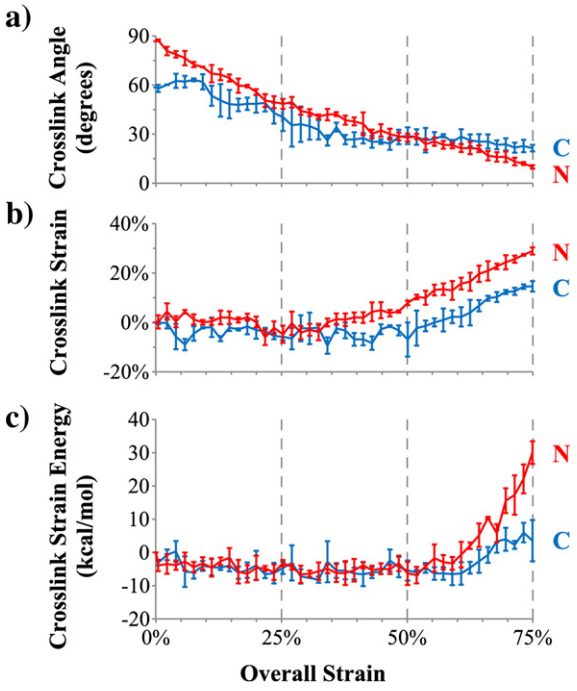

Fig. 6. Orientation angle (a), strain (b), and energy storage (c) in the N- and C-crosslinks at 1.5625 m/s averaged from both pulling directions (mean ± s.d., n = 3). (a), (b), and (c) all have the same x-axis shown in (c). Crosslink angles and crosslink strains were based on the alpha carbon atoms of the crosslinked amino acids, and crosslink angles were calculated relative to the pulling direction.

was similar up to ~60% overall strain, during which neither crosslink stored appreciable levels of strain energy; this lag or delay phase is consistent with that reported by Uzel and Buehler (2011). We observed that after this lag phase, strain energy in the N-crosslink reached ~8× that of the C-crosslink by 75% overall strain (Fig. 6c).

It is interesting to note also that the C-crosslink responded in a manner that was offset or delayed relative to the N-crosslink. This offset for crosslink strain energy was calculated as ~11 ± 3% overall strain (Fig. 6c). The crosslinks exhibited, essentially, a two-stage, recruitment-like response. These observations suggest that for a collagen fibril with many N- and C-crosslinks, N-crosslinks may serve as first responders and C-crosslinks as secondary responders to mechanical deformation. We propose that the recruitment response predicted here is attributable to the conformations of the N- and C-telopeptides and their specific points of connection to the crosslinks. The C-crosslink is reported to attach at amino acid position #16 of the C-telopeptide (Orgel et al., 2006); this position is located after the C-telopeptide hooks around at amino acid positions #13 and #14 (Orgel et al., 2000). In our investigation, this configuration of the C-telopeptide and its C-crosslink allowed the C-telopeptide to undergo a prolonged unfolding (Fig. 4c), permitting the delayed or offset response of the C-crosslink relative to the N-crosslink.

The extension of the two-stage response predicted here for N- and C-crosslinks to greater collagenous length scales, such as collagen fibrils, and for modeling applications is, however, dependent upon factors such as the site-specificity of crosslink formation (i.e., whether a crosslink forms specifically at the N- or C-crosslinking sites) and N–C directionality (i.e., whether the collagen type I amino acid chains maintain a common N–C direction within each molecule, and whether the molecules maintain a common N–C direction within fibrils). Regarding the first factor, we anticipate that this two-stage response would occur for other enzymatic crosslinks that form specifically at the telopeptides, but not for non-enzymatic crosslinks that form more haphazardly outside of enzyme involvement. Regarding the second factor, N–C directionality is typically maintained due to a multitude of intracellular and extracellular processes that drive and regulate the synthesis and folding of the collagen type I α-chains (Khoshnoodi et al., 2006), the release, modification, and assembly of molecules into fibrils (Kadler et al., 1996; Leikin et al., 1995), and the growth of these fibrils within tissues via fibril-to-fibril fusion (Kadler et al., 1996). It has also been reported that collagen fibrils can form unipolar N–C fibrils, or they can become bipolar N–N fibrils; however, even in bipolar N–N fibrils, most of their length has reported to be unipolar with a single region of C–C transition (Kadler et al., 2000, 1996) (Fig. 7). These are only two factors that may influence the extension of this recruitment mechanism to larger length scales; there are likely several other factors such as the initial alignment of the N- and C-crosslinks, the initial spacing between the molecules at the N- and C-crosslinking sites, and the crosslink type located at each site. It would be interesting to explore these and other factors in future work.

The ~23-nm-long N–C region that was modeled here (i.e., N-end of one molecule adjacent to the C-end of another molecule) is expected to be a prevalent and important region of interest within collagen fibrils (both unipolar and bipolar), which can span hundreds of micrometers in length. A mechanical recruitment mechanism of N- and C-crosslinks, as reported here, could be accounted for in the development of larger-scale predictive models for the mechanical behavior of native collagenous tissues, engineered tissues, and collagen-based materials.

3. Experimental procedures

3.1. Molecular modeling programs and parameters

The molecular model was constructed using the computer programs Discovery Studio Visualizer (Accelrys Software, Inc., 2011) and Visual Molecular Dynamics (VMD) (Humphrey et al., 1996). Simulations were carried out using the NAnoscale Molecular Dynamics (NAMD) code (Kale et al., 2011; Phillips et al., 2005), and the Tachyon ray tracing code was used with VMD to generate molecular renderings for visualization (Stone, 1998).

The Chemistry at HARvard Molecular Mechanics version 19 (CHARMM19) united-atom force field (Neria et al., 1996) was used. This force field consists of a potential energy expression containing various covalent and non-covalent energy terms (Eq. (1)), a parameter file (param19.inp), and a topology file (toph19.inp). These parameter and topology files were obtained from http://mackerell.umaryland.edu/CHARMM_ff_params.html. The parameter file contains parameters for these energy terms (Table 1), while the topology file contains information about different amino acids and solvents being modeled (e.g., their constituent atoms, how these atoms are connected, and the partial atomic charges and masses of these atoms).

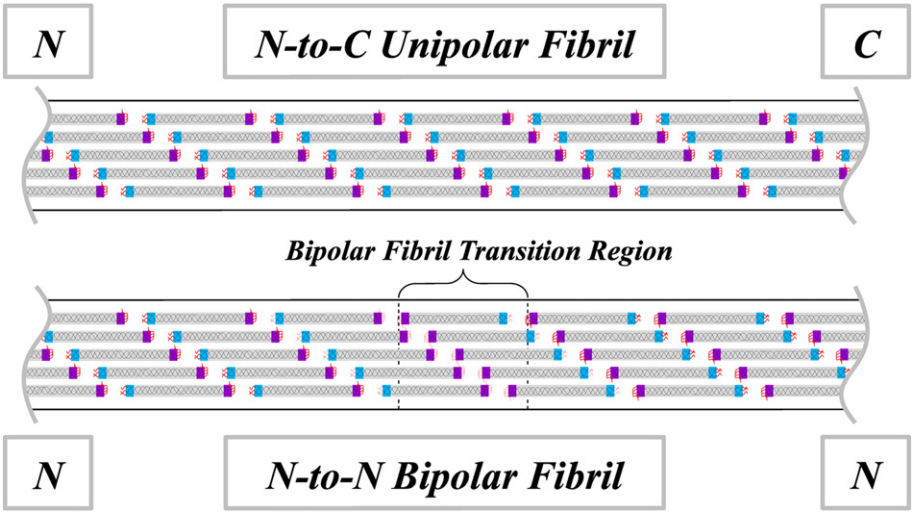

Fig. 7. A schematic of longitudinal slices through unipolar and bipolar fibrils with their internal arrangement of collagen type I molecules. "N" and "C" represent the amino/N-terminal and carboxy/C-terminal directions. The wavy lines at the left and right represent truncations of the fibril lengths. The number of molecules depicted in these fibril cross-sections is smaller than the actual number; this was done for visual clarity. The color scheme is identical to that described for Fig. 4. The transition region represents a region where the C-ends of two molecules are adjacent rather than the C-end of one molecule and the N-end of another molecule.

Table 1
CHARMM19 force field parameters (Brunner et al., 2008; Neria et al., 1996).

Term	Parameter	Description	Unit
Bonds	k_b	Force constant	kcal/mol/Å^2
	b	Bond length	Å
	b_0	Equilibrium bond length	Å
Bond angles	k_θ	Force constant	kcal/mol/radian^2
	θ	Bond angle	Degrees
	θ_0	Equilibrium bond angle	Degrees
Dihedral angles	k_φ	Force constant	kcal/mol
	n	Periodicity/symmetry parameter	Unitless
	φ	Torsion angle	Degrees
	δ	Phase angle	Degrees
Improper angles	k_ω	Force constant	kcal/mol/radian^2
	ω	Out-of-plane bending angle	Degrees
	ω_0	Equilibrium out-of-plane bending angle	Degrees
vdW (van der Waals)	r_{ij}	Distance between atoms i and j (center-to-center)	Å
	r_i^{min} and r_j^{min}	van der Waals radii for atoms i and j	Å
	ε_i and ε_j ε_{ij}	vdW energy minimum for atoms i and j ε_{ij} = sqrt(ε_i*ε_j)	kcal/mol
	σ_{ij}	Distance between atoms i and j at ε_{ij} $\sigma_{ij} = r_i^{min} + r_j^{min}$	Å
ES (electrostatics)	C	Coulomb's constant = 332.0636	(kcal · Å)/(mol · e^2)
	ϵ_0	Dielectric constant; "dielectric" NAMD parameter	Unitless
	ε_{ij}^{14}	1–4 scaling factor; "1–4scaling" NAMD parameter	Unitless
	q_i and q_j	Partial atomic charges of atoms i and j	Electron charge, e
	r_{ij}	Distance between atoms i and j (center-to-center)	Å

CHARMM19 parameters for two non-standard amino acids, hydroxyproline and hydroxylysine, were obtained from the CHARMm polar-hydrogen force field (Accelrys Software, Inc., 2011). CHARMM19 parameters for the deH-HLNL crosslink were derived from existing CHARMM19 parameters (Neria et al., 1996), following a review of a CHARMM22 combined topology-parameter file for "retinal/retinol and related model compounds" (stream/toppar_all22_prot_retinol.str) obtained from http://mackerell.umaryland.edu/CHARMM_ff_params.html (MacKerell et al., 1998). This CHARMM22 topology-parameter file contains parameters for a deprotonated Schiff's base structure resulting from the covalent linking of retinal to lysine; this structure is referred to as "SCH1" or "Schiff's base model compound 1, deprotonated" within this topology-parameter file. A deprotonated Schiff's base is the same type of structure that is present within the deH-HLNL crosslink (Fig. 1), motivating our review of the CHARMM22 SCH1 structure to aid our derivation of CHARMM19 parameters for deH-HLNL. The additional CHARMM19 parameters derived for the deH-HLNL crosslink are tabulated in Supplementary Tables 1–4.

In the CHARMM19 force field (Eq. (1)), total potential energy is expressed as a summation of a "bonds" term (bond stretching between two covalently bonded atoms), "bond angles" term (angle bending amongst three covalently bonded atoms), "dihedral angles" term (bond twisting amongst four covalently bonded atoms), "improper angles" term (out-of-plane bending for four covalently bonded atoms), "vdW" term (van der Waals; induced dipoles between two non-covalently interacting atoms), and "ES" term (electrostatics; permanent dipoles between two non-covalently interacting atoms). Each of these potential energy terms is summed over various atom combinations within one's simulated system. Table 1 provides a listing of parameters employed within these different CHARMM19 potential energy terms. S_{vdw} and S_{elec} are switching and shifting functions, respectively. R_{on} is the switching distance, and R_{off} is the cutoff distance. $S_{vdw} = 1$ when $r_{ij} \le R_{on}$, and $S_{vdw} = 0$ when $r_{ij} \ge R_{off}$. $S_{vdw} = (R_{off}^2 - r_{ij}^2)^2 * (R_{off}^2 + 2r_{ij}^2 - 3R_{on}^2) / (R_{off}^2 - R_{on}^2)^3$ when $R_{on} < r_{ij} < R_{off}$. $S_{elec} = (1 - r_{ij}^2 / R_{off}^2)^2$ when $r_{ij} < R_{off}$, and $S_{elec} = 0$ when $r_{ij} \ge R_{off}$ (Brunner et al., 2008).

$$V = \tfrac{1}{2}\sum_{\text{bonds}} k_b(b-b_0)^2 + \tfrac{1}{2}\sum_{\text{bond angles}} k_\theta(\theta-\theta_0)^2 + \tfrac{1}{2}\sum_{\text{dihedral angles}} k_\varphi(1+\cos(n\varphi-\delta)) + \tfrac{1}{2}\sum_{\text{improper dihedrals}} k_\omega(\omega-\omega_0)^2 + \sum_{\text{vdW}} \varepsilon_{ij}\left[\left(\sigma_{ij}/r_{ij}\right)^{12} - 2*\left(\sigma_{ij}/r_{ij}\right)^{6}\right] * S_{vdw} + \sum_{\text{ES}} \left(\varepsilon_{ij}^{14} * C * q_i * q_j\right) / \left(\epsilon_0 * r_{ij}\right) * S_{elec} \quad (1)$$

There are several groups of force fields that have been developed for molecular simulations of biological macromolecules such as proteins, nucleic acids, lipids, and/or carbohydrates (e.g., CHARMM, CHARMm, AMBER, GROMOS, and OPLS). Our use of the CHARMM19 force field, with additional parameters derived from the CHARMm polar-hydrogen and CHARMM22 force fields, was based on the need for hydroxyproline and hydroxylysine parameters and deH-HLNL crosslink parameters for our system of interest (i.e., collagen type I molecular segments covalently linked with deH-HLNL crosslinks).

The non-standard amino acid hydroxyproline does not have parameters within the CHARMM19, CHARMM22/27, and OPLS force fields; however, hydroxyproline is parameterized in the AMBER (94, 96, 99, 03, and 99SB), CHARMm (all-atom and polar-hydrogen), and GROMOS (43a1 and later) force fields. The non-standard amino acid hydroxylysine is parameterized in the CHARMm (all-atom and polar-hydrogen) force fields, but not in the CHARMM, AMBER, GROMOS, nor OPLS force fields. Schiff's bases, such as that present in our crosslink of interest, are parameterized in the CHARMM all-atom force fields (22 and 27), but not in the CHARMM19, CHARMm, AMBER, GROMOS, nor OPLS force fields. For these reasons, we decided to utilize parameters from the CHARMM/m force fields for our simulations. Furthermore, we used the CHARMM19 united-atom force field, rather than the CHARMM22 all-atom force field, to reduce the number of atoms in our simulations and to thus reduce the computational demand and time required for our simulations at the sacrifice of the structural accuracy available with an all-atom force field.

The determination of whether hydroxyproline, hydroxylysine, and Schiff's base compounds were parameterized in existing force fields was based on the following resources: AMBER, GROMOS, and OPLS force fields ("aminoacids.rtp" files within the "share/top/" directory of the "gromacs-4.6.3.tar.gz" source code archive available from http://www.gromacs.org/Downloads). CHARMM force fields ("toph19.inp" and "top_all22_prot.inp" files within the "toppar/" directory and the "toppar_all22_prot_retinol.str" file within the "toppar/stream/" directory of the "toppar_c35b2_c36a2.tgz" archive available from http://mackerell.umaryland.edu/CHARMM_ff_params.html), and CHARMm force fields (CHARMm "AMINO.RTF" and "AMINOH.RTF" topology files available after the installation of Discovery Studio Visualizer by Accelrys).

3.2. Molecular model building steps

The triple-helical structure in our molecular model was based upon a 29-amino acid-long collagen-like (proline-proline-glycine)$_n$ crystal

structure acquired from the Protein Data Bank (1K6F.pdb) (Berisio et al., 2002). Crystallographic water molecules were removed, and the structure was replicated and spliced together to form a single collagen-like α1-α2-α1 triple helix containing 1054 and 1026 amino acids for the α1 and α2 chains, respectively (Inline Supplementary Fig. S7a). These amino acid-lengths of 1054 and 1026 included the N- and C-telopeptides; however, at this stage of the model's construction, the amino acids that would later comprise the telopeptides were still in a triple-helical conformation. This triple-helical collagen-like molecule was then divided into five segments (segments 1, 2, 3, 4, and 5) (Inline Supplementary Fig. S7b), and the segments were assembled into a quasi-hexagonal compressed microfibril representing an entire collagen type I D-period (Inline Supplementary Fig. S7c) (Fraser et al., 1983). The overlap region of the two main crosslinking segments (segments 1 and 5) was then isolated for further use, leading to a length of ~30 nm (Inline Supplementary Fig. S7d); this 30-nm-long collagen-like model contained 106 ± 8 amino acids per chain including the telopeptides. The collagen-like amino acid sequences were then replaced with sequences for rat collagen type I; these sequences were obtained from a Protein Data Bank file (1Y0F.pdb) of Orgel et al. (Chapman and Hulmes, 1984; Orgel et al., 2006). The conformations of the N- and C-telopeptides were then later adjusted based upon the revised Protein Data Bank 3HR2.pdb structure of Orgel et al. (2006), allowing for favorable alignments between four potential crosslinking sites (Eyre and Wu, 2005). These latter two steps are depicted in Inline Supplementary Fig. S7e.

Inline Supplementary Fig. S7 can be found online at http://dx.doi.org/10.1016/j.matbio.2013.10.012.

The sequences in our model were as follows from 1Y0F.pdb (Orgel et al., 2006):

α1 chain EMSYGYDEKSTGISVP | GPM … GPP | SGGYDLSFLPQPPQQ KAHDKGRYY
α2 chain EFDAKG | GGP … GPP | SGGYEF

These sequences contained 1054 amino acids for each α1 chain and 1026 amino acids for the α2 chain based on 1Y0F.pdb (Orgel et al., 2006). The pipe symbols "|" denote separations of the N-telopeptide, triple-helical, and C-telopeptide domains assigned to our molecular model; this domain assignment was based on retaining 1014 amino acids for all three chains within the triple-helical domain and aligning the sequences of the three chains near the C-terminal end of the triple-helical domain such that, for all three chains, the glycine-proline-proline | serine-glycine-glycine-tyrosine (GPP | SGGY) sequences represented the end of the triple-helical domain (Weiss and Jayson, 1982) and the start of the C-telopeptide domain (Orgel et al., 2000). This approach resulted in the following domain assignments for the α1 chains (1054 amino acids total with 16 for the N-telopeptide, 1014 for the triple-helical domain, and 24 for the C-telopeptide) and the α2 chain (1026 amino acids total with 6 for the N-telopeptide, 1014 for the triple-helical domain, and 6 for the C-telopeptide domain). The number of amino acids in our model, based on the X-ray fiber diffraction structure of Orgel et al. (1Y0F.pdb), was employed so that experimentally based conformational adjustments could be applied to our modeled N- and C-telopeptides.

3.3. Energy minimization, heating, & equilibration

Energy minimization via the conjugate gradients method was performed until an NAMD gradient tolerance of 0.01 kcal/mol/Å was reached based on a review of convergence criteria that have been employed in the literature. For heating, temperature reassignment was used to increase the temperature from 0 to 293.15 K over 10 ps. During subsequent equilibration steps, the Langevin dynamics method was used for temperature control at 293.15 K. The generalized Born implicit solvent (GBIS) approach of Onufriev, Bashford, and Case, namely, GB^{OBC} II, was used to indirectly model the effects of solvent and ions (Onufriev et al., 2000). The solvent dielectric, ion concentration, and GBIS alpha cutoff parameters were set to 80.0, 0.150 mol/L, and 14 Å, respectively. The simulation timestep was initially set to 1 fs during energy minimization, heating (10 ps), and the initial stages of equilibration (1 ns); it was later switched to 2 fs during the remaining equilibration steps (6 ns, with rigid bonds and the settle algorithm for covalent bonds involving hydrogens). Furthermore, during energy minimization, heating, and the first 0.5 ns of equilibration, harmonic positional restraints were placed on all alpha carbon atoms (k = 0.5 kcal/mol/Å^2); during the remaining 6.5 ns of equilibration, all positional restraints were removed. Non-bonded interactions were treated with a switching distance of 15 Å, a cutoff distance of 16 Å, and a pair list distance of 17.5 Å. The solvent dielectric of 80.0 was selected in representation of water at 293.15 K and 1 atm, and the ion concentration of 0.150 mol/L was selected to match physiological saline solution (i.e., 9% w/v sodium chloride in water). The non-bonded distances (15–16–17.5 Å) and the GBIS alpha cutoff (14 Å) were based upon values suggested in the NAMD 2.8 user's guide for use with GBIS (Bhandarkar et al., 2011).

3.4. Constant velocity pulling simulations

We employed the moving constraints technique implemented in NAMD to conduct constant velocity pulling (Bhandarkar et al., 2011). Each pulled atom was connected through a virtual spring of stiffness, k, to a virtual atom (moving constraint) that was set to move at a constant velocity, v. A schematic of the moving constraints technique is shown as an inset in Fig. 4a. In our simulations, three N- or C-terminal Cα atoms of one segment were fixed, while three C- or N-terminal Cα atoms of the other segment were pulled in the opposite direction (Fig. 4).

Three parameters were investigated, namely, the spring constant, the pulling direction, and the pulling velocity. First, with a set velocity of 100 m/s and while pulling in the C-terminal direction, the spring constant parameter was varied from 0.01 to 9999 kcal/mol/Å^2 in factors of ~10 (0.01, 0.10, 1.00, 10.0, 100, 1000, and 9999). 9999 was used as an approximation of 10,000, due to the number of characters permitted for the occupancy (O) column of a PDB file; the occupancy column was used to set the spring constant parameter for each pulled atom (e.g., replacing 0 with the spring constant value), and the beta (B) column was used to identify the fixed atoms (e.g., replacing 0 with 1). The observed velocity of the pulled atoms was calculated as the slope of the RMS displacement of the pulled atoms versus time. Percent error was then calculated as the difference between this observed velocity and the set velocity divided by the set velocity all multiplied by 100%. Numerical noise in the potential energy versus time data was calculated as the error sum of squares (SS_{error}). The potential energy versus time data for each spring constant were fitted with a quadratic polynomial function, and SS_{error} values were calculated using the JMP statistical software as $SS_{error} = \sum (a_i - f_i)^2$ where i increases from 1 to n x-axis data points (e.g., time), a_i represents the actual data value, and f_i represents the fitted value (SAS Institute, Inc., 2010). Second, with a velocity of 100 m/s and a spring constant of 9999 kcal/mol/Å^2, the pulling direction was varied in either the C-direction or N-direction. Lastly, with a spring constant of 9999 kcal/mol/Å^2 and while pulling in the C-terminal direction, the pulling velocity was varied from 100 down to 1.5625 m/s in factors of two (100, 50, 25, 12.5, 6.25, 3.125, and 1.5625 m/s). See Fig. 4 for representative snapshots of the simulations.

3.5. Analysis of deformation and strain energy of the crosslinks and the overall model

Crosslink deformation was characterized by two measures, namely, crosslink alignment and crosslink engineering strain. Crosslink alignment was defined as the orientation angle of a line connecting the two Cα atoms (Cα-to-Cα line) of each crosslink relative to the pulling direction. Crosslink engineering strain was calculated as the change in

length of this Cα-to-Cα line divided by its initial length. Crosslink strain energy was defined as the change in total potential energy (Eq. (1)) of the atoms comprising each crosslink, excluding all backbone atoms except for the Cα atoms.

Overall deformation was measured from overall engineering strain, which was defined as the RMS displacement of the pulled atoms divided by the initial length of the molecular model (22.55 nm). This initial length was measured as the RMS difference between the coordinates of the three fixed Cα atoms and the three pulled Cα atoms, using the "measure rmsd" tool command language (TCL) command in VMD. The RMS displacement of the pulled atoms was calculated using the RMSD Trajectory Tool in VMD. All data were reported out to 75% overall strain, which was selected based upon average failure strains that have been reported for single collagen type I fibrils (Shen et al., 2010; van der Rijt, 2004). Overall strain energy was defined as the change in total potential energy (Eq. (1)) of all of the atoms in the model.

It was observed that the length of our molecular model after equilibration and just prior to the pulling simulations was shorter than values typically reported for the overlap region (e.g., (0.4 to 0.48) *67 nm = 26.8 to 32.2 nm) (Orgel et al., 2000). Our molecular model experienced a contraction of its length during the equilibration steps, in particular, once the harmonic positional restraints were removed from the alpha carbon atoms. The length of our molecular model was 29.75 nm before energy minimization, 29.91 nm after energy minimization, 29.93 nm after heating, 29.89 nm after the initial restrained equilibration step, and 22.55 nm after the subsequent unrestrained equilibration steps. This length of 22.55 nm was the length just prior to the pulling simulations. The discrepancy of the length of our molecular model (~23 nm) compared to values typically reported for the overlap region of collagen type I's D-period (26.8 to 32.2 nm) (Orgel et al., 2000) may be attributed to the fact that our molecular model represented two collagen type I molecular segments isolated from their parent collagen molecules and isolated from a microfibrillar arrangement. Thus, the length of our modeled collagen type I segments was less restricted in our model than if they had been modeled as full-length collagen molecules within a microfibrillar array.

3.6. Statistical analysis

All molecular dynamics simulations were conducted in triplicate ($n = 3$), and all statistical analyses were performed via the JMP statistical software (SAS Institute, Inc., 2010). Unless otherwise stated, α (also known as the alpha value, significance level, or type I error rate) was set to 0.05. The Brown–Forsythe and Levene tests were used to assess if variances were homogeneous. If so, a one-way analysis of variance (ANOVA) employing a pooled standard deviation was used, followed by a Tukey–Kramer Honestly Significant Difference test, when appropriate. If not, then a Welch's one-way ANOVA and Student's t-tests with Bonferroni correction were used. For the Bonferroni correction, α was set to 0.05 divided by the number of t-tests (e.g., number of t-tests $= g * (g - 1) / 2$, where g denotes the number of groups). It should be noted, however, that with sample sizes of only three ($n = 3$), one is not able to verify the distribution of one's data (i.e., we assumed that our data was normally distributed, but could not verify this assumption with our sample sizes). This assumption of normality underlies our reporting of means and standard deviations and the statistical analyses that we employed.

3.7. Computing resources

These molecular dynamics simulations were conducted using the following two computing systems: (1) an Intel 64-bit Linux computing cluster (34 compute nodes, ~20 to 30 GB RAM per node, 8 or 16 processor cores per node, and 2.21 to 2.60 GHz per core) and (2) an HP desktop computer (2 GB RAM and a Core2Duo processor with 3.16 GHz per core).

Author contributions

A.L. Kwansa and J.W. Freeman designed the molecular modeling methods; A.L. Kwansa prepared and submitted the simulation input files, performed the analyses of the simulation output files, and conducted the statistical analyses; A.L. Kwansa, J.W. Freeman, and R. De Vita were involved in the interpretation and discussion of the results and the writing of the manuscript.

Acknowledgments

We thank the Virginia Tech Department of Engineering Science and Mechanics for granting access to their Linux computing cluster and Prof. D. Bevan for providing feedback on the manuscript. Graduate student support was received through the Harriett G. Jenkins Predoctoral Fellowship Project (National Aeronautics and Space Administration and the UNCF Special Programs Corporation) and the Virginia Tech-Initiative for Maximizing Student Development fellowship program (National Institutes of Health Biomedical and Behavioral Sciences Research Training Grant # R25 GM072767). This work is a part of a project funded by the National Science Foundation's Chemical, Bioengineering, Environmental, and Transport Systems Division (Award # 0932024 granted to R. De Vita, J.W. Freeman, and J. Barrett).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.matbio.2013.10.012.

References

Accelrys Software, Inc., 2011. Discovery Studio Modeling Environment. Accelrys Software, Inc., San Diego, CA.
Avery, N.C., Bailey, A.J., 2008. Restraining cross-links responsible for the mechanical properties of collagen fibers: natural and artificial. In: Fratzl, P. (Ed.), Collagen: Structure and Mechanics. Springer US, New York, NY, pp. 81–91.
Berisio, R., Vitagliano, L., Mazzarella, L., Zagari, A., 2002. Crystal structure of the collagen triple helix model [(Pro-Pro-Gly)(10)](3). Protein Sci. 11, 262–270.
Bhandarkar, M., Bhatele, A., Bohm, E., Brunner, R., Buelens, F., Chipot, C., Dalke, A., Dixit, S., Fiorin, G., Freddolino, P., Grayson, P., Gullingsrud, J., Gursoy, A., Hardy, D., Harrison, C., Hénin, J., Humphrey, W., Hurwitz, D., Krawetz, N., Kumar, S., Kunzman, D., Lee, C., McGreevy, R., Mei, C., Nelson, M., Phillips, J., Sarood, O., Shinozaki, A., Tanner, D., Zheng, G., Zhu, F., 2011. NAMD User's Guide, Version 2.8.
Birk, D.E., Bruckner, P., 2005. Collagen suprastructures. In: Brinckmann, J., Notbohm, H., Muller, P.K. (Eds.), Collagen: Primer in Structure, Processing, and Assembly, Topics in Current Chemistry. Springer Berlin Heidelberg, New York, NY, pp. 185–205.
Bourne, J.W., Torzilli, P.A., 2011. Molecular simulations predict novel collagen conformations during cross-link loading. Matrix Biol. 30, 356–360.
Brunner, R., Dalke, A., Gursoy, A., Humphrey, W., Nelson, M., 2008. NAMD User Guide, Version 1.5.
Chapman, J.A., Hulmes, D.J.S., 1984. Electron microscopy of the collagen fibril. In: Ruggeri, A., Motta, P.M. (Eds.), Ultrastructure of the Connective Tissue Matrix, Electron Microscopy in Biology and Medicine. Martinus Nijhoff Publishers, Boston, MA, pp. 1–33.
Eyre, D.R., Wu, J.-J., 2005. Collagen cross-links. In: Brinckmann, J., Notbohm, H., Muller, P.K. (Eds.), Collagen: Primer in Structure, Processing, and Assembly, Topics in Current Chemistry. Springer Berlin Heidelberg, New York, NY, pp. 207–229.
Eyre, D.R., Koob, T.J., Van Ness, K.P., 1984a. Quantitation of hydroxypyridinium crosslinks in collagen by high-performance liquid chromatography. Anal. Biochem. 137, 380–388.
Eyre, D.R., Paz, M.A., Gallop, P.M., 1984b. Cross-linking in collagen and elastin. Annu. Rev. Biochem. 53, 717–748.
Fraser, R.D., MacRae, T.P., Miller, A., Suzuki, E., 1983. Molecular conformation and packing in collagen fibrils. J. Mol. Biol. 167, 497–521.
Fujii, K., Yamagishi, T., Nagafuchi, T., Tsuji, M., Kuboki, Y., 1994. Biochemical properties of collagen from ligaments and periarticular tendons of the human knee. Knee Surg. Sports Traumatol. Arthrosc. 2, 229–233.
Gautieri, A., Buehler, M.J., Redaelli, A., 2009. Deformation rate controls elasticity and unfolding pathway of single tropocollagen molecules. J. Mech. Behav. Biomed. 2, 130–137.
Hambli, R., Barkaoui, A., 2012. Physically based 3D finite element model of a single mineralized collagen microfibril. J. Theor. Biol. 301, 28–41.
Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33–38.
Israelewitz, B., Baudry, J., Gullingsrud, J., Kosztin, D., Schulten, K., 2001a. Steered molecular dynamics investigations of protein function. J. Mol. Graph. Model. 19, 13–25.
Israelewitz, B., Gao, M., Schulten, K., 2001b. Steered molecular dynamics and mechanical functions of proteins. Curr. Opin. Struct. Biol. 11, 224–230.

Kadler, K.E., Holmes, D.F., Trotter, J.A., Chapman, J.A., 1996. Collagen fibril formation. Biochem. J. 316 (Pt 1), 1–11.
Kadler, K.E., Holmes, D.F., Graham, H., Starborg, T., 2000. Tip-mediated fusion involving unipolar collagen fibrils accounts for rapid fibril elongation, the occurrence of fibrillar branched networks in skin and the paucity of collagen fibril ends in vertebrates. Matrix Biol. 19, 359–365.
Kagan, H.M., Ryvkin, F., 2011. Lysyl oxidase and lysyl oxidase-like enzymes. In: Mecham, R.P. (Ed.), The Extracellular Matrix: An Overview. Springer, New York, NY, pp. 303–306.
Kale, L.V., Bhatele, A., Bohm, E.J., Phillips, J.C., 2011. NAnoscale Molecular Dynamics (NAMD). In: Padua, D. (Ed.), Encyclopedia of Parallel Computing. Springer US, New York, NY, pp. 1249–1254.
Khoshnoodi, J., Cartailler, J.P., Alvares, K., Veis, A., Hudson, B.G., 2006. Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers. J. Biol. Chem. 281, 38117–38121.
Knott, L., Bailey, A.J., 1998. Collagen cross-links in mineralizing tissues: a review of their chemistry, function, and clinical relevance. Bone 22, 181–187.
Leikin, S., Rau, D.C., Parsegian, V.A., 1995. Temperature-favored assembly of collagen is driven by hydrophilic not hydrophobic interactions. Nat. Struct. Biol. 2, 205–210.
Lorenzo, A.C., Caffarena, E.R., 2005. Elastic properties, Young's modulus determination and structural stability of the tropocollagen molecule: a computational study by steered molecular dynamics. J. Biomech. 38, 1527–1533.
MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., Karplus, M., 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B 102, 3586–3616.
Neria, E., Fischer, S., Karplus, M., 1996. Simulation of activation free energies in molecular systems. J. Chem. Phys. 105, 1902–1921.
Onufriev, A., Bashford, D., Case, D.A., 2000. Modification of the generalized Born model suitable for macromolecules. J. Phys. Chem. B 104, 3712–3720.
Orgel, J.P., Wess, T.J., Miller, A., 2000. The in situ conformation and axial location of the intermolecular cross-linked non-helical telopeptides of type I collagen. Structure 8, 137–142.
Orgel, J.P.R.O., Miller, A., Irving, T.C., Fischetti, R.F., Hammersley, A.P., Wess, T.J., 2001. The in situ supermolecular structure of type I collagen. Structure 9, 1061–1069.
Orgel, J.P.R.O., Irving, T.C., Miller, A., Wess, T.J., 2006. Microfibrillar structure of type I collagen in situ. Proc. Natl. Acad. Sci. U. S. A. 103, 9001–9005.
Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., Schulten, K., 2005. Scalable molecular dynamics with NAMD. J. Comput. Chem. 26, 1781–1802.
Saito, M., Marumo, K., Fujii, K., Ishioka, N., 1997. Single-column high-performance liquid chromatographic fluorescence detection of immature, mature, and senescent cross-links of collagen. Anal. Biochem. 253, 26–32.
SAS Institute, Inc., 2010. JMP. SAS Institute, Inc., Cary, NC.
Sasaki, N., Odajima, S., 1996. Elongation mechanism of collagen fibrils and force-strain relations of tendon at each level of structural hierarchy. J. Biomech. 29, 1131–1136.
Shen, Z.L., Dodge, M.R., Kahn, H., Ballarini, R., Eppell, S.J., 2010. In vitro fracture testing of submicron diameter collagen fibril specimens. Biophys. J. 99, 1986–1995.
Sims, T.J., Bailey, A.J., 1992. Quantitative analysis of collagen and elastin cross-links using a single-column system. J. Chromatogr. 582, 49–55.
Stone, J.E., 1998. An Efficient Library for Parallel Ray Tracing and Animation. (M.S. Thesis) University of Missouri-Rolla, Rolla, MO.
Sweeney, S.M., Orgel, J.P., Fertala, A., McAuliffe, J.D., Turner, K.R., Di Lullo, G.A., Chen, S., Antipova, O., Perumal, S., Ala-Kokko, L., Forlino, A., Cabral, W.A., Barnes, A.M., Marini, J.C., Antonio, J.D.S., 2008. Candidate cell and matrix interaction domains on the collagen fibril, the predominant protein of vertebrates. J. Biol. Chem. 283, 21187–21197.
Tarini, M., Cignoni, P., Montani, C., 2006. Ambient occlusion and edge cueing for enhancing real time molecular visualization. IEEE Trans. Vis. Comput. Graph. 12, 1237–1244.
The UniProt Consortium, 2012. Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res. 40, D71–D75.
Urich, K., 1994. Extracellular structural and secretory proteins. In: King, P.J. (Tran.) (Ed.), Comparative Animal Biochemistry. Springer-Verlag, New York, NY, pp. 378–383.
Uzel, S.G., Buehler, M.J., 2011. Molecular structure, mechanical behavior and failure mechanism of the C-terminal cross-link domain in type I collagen. J. Mech. Behav. Biomed. Mater. 4, 153–161.
Van der Rijt, J.A.J., 2004. Micromechanical Testing of Single Collagen Type I Fibrils. (Ph.D. Dissertation) University of Twente, Enschede, The Netherlands.
Von der Mark, K., 2006. Structure, biosynthesis and gene regulation of collagens in cartilage and bone. In: Seibel, M.J., Robins, S.P., Bilezikian, J.P. (Eds.), Dynamics of Bone and Cartilage Metabolism. Principles and Clinical Applications. Academic Press/Elsevier, Burlington, MA, pp. 3–40.
Weiss, J.B., Jayson, M.I.V., 1982. Collagen in Health and Disease. Longman Group Limited, New York, NY.
Williams, A.P., 1978. The amino acid, collagen and mineral composition of preruminant calves. J. Agr. Sci. 90, 617–624.