Nanomechanics of type I collagen

S. Varma, J. P. R. O. Orgel, and J. D. Schieber

Abstract

Type I collagen is the predominant collagen in mature tendons and ligaments, where it gives them their load-bearing mechanical properties. Fibrils of type I collagen are formed by the packing of polypeptide triple helices. Higher-order structures like fibril-bundles and fibers are assembled from fibrils in the presence of other collagenous molecules and non-collagenous molecules. Curiously, however, experiments show that fibrils/fibril-bundles are less resistant to axial stress compared to their constituent triple helices – the Young's moduli of fibrils/fibril-bundles are an order in magnitude smaller than the Young's moduli of triple helices. Given the sensitivity of the Young's moduli of triple helices to solvation environment, a plausible explanation is that the packing of triple helices into fibrils perhaps reduces the Young's modulus of an individual triple helix, which results in fibrils having smaller Young's moduli. We find, however, from molecular dynamics and accelerated conformational sampling simulations that the Young's modulus of the buried core of the fibril is of the same order as that of a triple helix in aqueous phase. These simulations, therefore, suggest that the lower Young's moduli of fibrils/fibril-bundles cannot be attributed to the specific packing of triple helices in the fibril core. It is not the fibril core that yields initially to axial stress. Rather, it must be the portion of the fibril exposed to the solvent and/or the fibril-fibril interface that bears the initial strain. Overall, this work provides estimates of Young's moduli and persistence lengths at two levels of collagen's structural assembly, which are necessary to quantitively investigate the response of various biological factors on collagen mechanics, including congenital mutations, post-translational modifications and ligand binding, and also engineer new collagen-based materials.

Received for publication XX and in final form XX.

0006-3495/08/09/2624/12 $2.00

doi: 10.1529/biophysj.106.090944

Introduction

Collagen is the primary constituent of the extracellular matrix of animal connective tissue (1–7). To date, twenty eight different types of collagen have been identified in vertebrates and higher invertebrates (4–7). Among these collagen types, type I is the most abundant in vertebrates. It combines with other molecules in varying ratios to form a variety of tissue scaffolds, such as basal membranes, ligaments, tendons, skin and blood vessels, where it gives them their load-bearing mechanical properties.

The building block of type I collagen is a 300 nm long triple helix, which is made up of three parallel polypeptide chains wound around each other. These triple helices assemble to form fibrils. The arrangement of triple helices in a fibril is such that the N-termini of two axially adjacent triple helices are separated by $\mathrm{D} = 67$ nm and the N-termini of two collaterally adjacent triple helices are separated axially by $0.54\mathrm{D}$ (8–14). This staggered arrangement creates alternating regions of low and high protein density along the fibril axis with a repeating unit of length D (Fig. 1). These D-periods are, in fact, signature structural features of several collagen types, and are visible as alternating dark and light bands in transmission electron microscopy (TEM) and atomic force microscopy (AFM). Fibrils, in turn, combine with other collagenous and non-collagenous molecules, such as proteoglycans, to form fibril-bundles and fibers, which then assemble with each other to form tissue scaffolds.

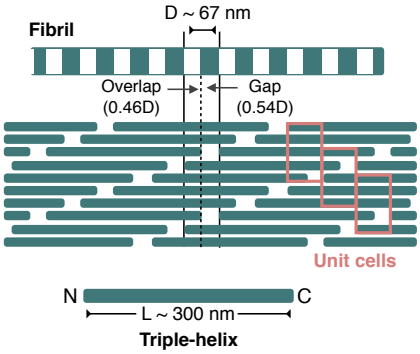

Figure 1: Structural hierarchy of type I collagen fibril. The fibril is drawn at two different spatial resolutions. While the lower resolution cartoon highlights the repeating dark and light bands visible in TEM and AFM, the higher resolution cartoon shows the 2D arrangement of the triple-helices (green rectangles) within a fibril. The higher resolution representation also outlines three representative crystallographic unit cells of the fibril. Note that a crystallographic unit cell is not a microfibril.

How do the mechanical properties of collagen vary across this structural hierarchy? Table 1 summarizes the mechanical properties of triple helices and fibrils/fibril-bundles obtained from experiment and molecular dynamics (MD) simulations (15–26). The first trend we note is that the persistence lengths of triple helices are significantly smaller than those of fibrils/fibril bundles. This trend reflects the fact that thermal fluctuations bend triple helices over much shorter lengths compared to fibrils/fibril-bundles. This is, nonetheless, expected because triple helices are three orders in magnitude thinner than fibrils/fibril-bundles. Surprisingly, however, we also note that the Young's modulus of triple helix, determined by mapping its mechanics on to an elastic rod, is an order of magnitude larger than the moduli of fibrils and fibril-bundles. This observation implies that it is easier to stretch fibrils/fibril-bundles compared to triple helices, at least in the limit of small strains. What explains this differential mechanical response?

We note first that the experiments that report the mechanical properties of fibrils/fibril-bundles, including ours (26), do not differentiate between fibrils and fibril-bundles. Consequently, the lower Young's moduli of fibrils/fibril-bundles could either be due to the packing of fibrils into fibril bundles, and/or due to the packing of triple helices into fibrils. Recent experiments show that the flexibility of type-I triple helix is sensitive to solvent conditions, and can modulate its Young's modulus by an order in magnitude (17, 27). This suggests the possibility that the packing of triple helices into fibrils, a process accompanied by partial dehydration of triple helices, perhaps reduces the Young's modulus of individual triple helix, which results in a fibril having smaller Young's moduli. It is also conceivable that the interaction between triple helices in a fibril is sufficiently weak that straining a fibril alters the relative arrangements between triple helices, but does not affect the structures of the individual triple helices. This could lead to a fibril being less resistant to axial deformation compared to a triple helix.

To explore these possibilities, we first determine using accelerated MD the Young's modulus of an isolated triple helix in salt solution. This serves as a control and also assesses the performance of MD against experiment. This is particularly important because we have demonstrated that the structural organization of triple helices in the fiber diffraction unit cell is sensitive to the employed MD protocol (28), and the question remains as to how well our protocol predicts the mechanics of

Table 1: Mechanical properties of type I collagen at different structural levels. The values of Young's moduli (Y) marked by asterisks (*) are computed from their reported respective persistence lengths (l_p) using the relationship given by Eq. 1. Similarly, the values of l_p marked by asterisks are computed from their reported respective Y using Eq. 1. Interconversion between Y and l_p through Eq. 1 requires values of temperature (T) and radius (r). In cases where temperature was not reported, it is assumed to be 298 K, and in cases where the radius of the triple helix was not reported, it is assumed to be 0.36 nm.

Structural level	r (nm)	Y (GPa)	l_p (nm)	Method	T (K)	Reference
Triple Helix	0.62	2.9 ± 0.12	$^*83.2$	X-ray	293	(15)
	–	$^*4.5$	14.5	Optical tweezers	298	(16)
	–	$^*3.7$–7.8	12–25	AFM	298	(17)
	0.78	1.3–2.4	–	Energy minimization	N/A	(18)
	0.36	$^*4.2$	12.9	Accelerated MD	310	Present work
Collagen-like Triple Helix	–	4.8 ± 1.0	$^*15.3$	Steered MD	300	(19)
	0.36	7.0	$^*22.3$	Steered MD	300	(20)
	0.36	4.0	$^*12.7$	Steered MD	310	(21)
	0.35	1.8-2.3	5.1-6.5*	MD (Force-extension)	300	(22)
Fibril Core	1.75	2.34	$^*4.0 \times 10^3$	MD (Constant-strain)	310	Present work
Fibril / Fibril-bundle	–	0.43	–	SAXS	293	(23)
	280–426	0.07-0.17	$^*0.8$–10.7×10^{11}	AFM	–	(24)
	220–570	0.123 ± 0.046	$^*0.6$–24.7×10^{11}	MEMS	–	(25)
	116–200	0.1-0.36	$^*0.03$–1.1×10^{11}	Optical tweezers	298	(26)

isolated triple helices. We then determine the Young's modulus of the buried core of a fibril. The underlying idea is that if the Young's modulus of the buried core is found to be similar to that of fibrils/fibril-bundles obtained from experiment, then we will conclude that the relatively lower Young's moduli of fibrils/fibril-bundles can be attributed to the specific packing of triple helices in the core of a fibril.

Methods

Potential of Mean Force

The potential of mean force, $U(R)$, where R is the end-to-end distance of a triple helix, is determined using well-tempered metadynamics (29–31). Gaussian hills of height $h_0 = 2$ kJ/mol and width $w = 0.2$ nm are deposited over R using a scaling factor $s = 6$ at regular intervals of 200 integration time steps. Geometrically, R is defined as the distance between the centers of masses of the backbone atoms of the N- and C-termini of the triple helix fragment.

The N- and C-termini of all three peptides in a triple helix are capped individually, and the peptides are described using the all atom Amber99sb-ildn force field (28, 32–34). The triple helix is placed in a cubic box containing explicit water molecules ($\sim$136K), and water is described using SPC/E parameters (35). NaCl salt concentration is set at 10 mM, and there are 4 extra Cl^- ions compared to Na^+ to balance the net charge on the triple helix. Electrostatic interactions are computed using the particle mesh Ewald scheme (36) with a Fourier grid spacing of 0.1 nm, a fourth-order interpolation, and a direct space cutoff of 10 Å. van der Waals interactions are computed explicitly for interatomic distance up to 10 Å. Temperature is regulated at 310 K and pressure at 0.1 MPa using extended ensemble approaches (37, 38) and with coupling constants of 1 ps. Bonds in the peptides are constrained using the P-LINCS algorithm (39), and the geometries of the water molecules are constrained using SETTLE (40). These constraints permit use of an integration time step of 2 fs. Prior to subjecting the triple helix to well-tempered metadynamics, the triple helix is subjected to 10 ns of standard MD.

Molecular Dynamics Simulations of Fibril Core

The details of our MD simulation protocol are provided elsewhere (28), and here we provide only the salient points. Since the resolution of the crystal structure (PDB ID: 3HR2) was insufficient to reveal atomic level details, including side chain orientations and inter-peptide hydrogen bonds (12, 13), an atomically detailed model was constructed by incorporating the high resolution crystallographic data of collagen-related peptides (12, 28, 41). The model includes all known hydroxylated forms of prolines and lysines, but not any glycosyl groups, as their specific sites remain unknown. We leave out all cross links between triple helices, and this is to allow the packed triple helices sufficient flexibility to rearrange under applied strains. It should be noted that the computed value of Young's modulus in the absence of cross links serves as a lower limit. The overall system contains 70,606 particles, which includes the peptide, explicit water ($\sim$11 K molecules) and 31 Cl^- ions to counter the positive charge of the peptide. The resulting ratio of Cl^- ions and water in the unit cell corresponds, roughly, to a salt concentration of 150 mM. With the exception of one simulation parameter, namely the pressure coupling scheme, all the simulation parameters are the same as those used in the MD simulation of the triple helix. In this case, we employ an anisotropic pressure coupling scheme wherein the components of the virial pressure tensor are regulated separately. While the diagonal components, σ_{ii}, are regulated at a pressure of 0.1 MPa, the non-diagonal components, $\sigma_{i\neq j}$, are regulated at zero pressure to emulate a zero-shear condition.

This MD simulation protocol preserves the dimensions of the triclinic unit cell, the gap/overlap ratio of the fibril, and also the pitch of the constituent triple helix (28). In addition, the average D-band length is $\mathrm{D}_0 = 66.28 \pm 0.08$ nm, which is in close correspondence with the values obtained from AFM, TEM and X-ray diffraction studies (8–14).

We use Gromacs v4.5 for all MD simulations (42).

Results and Discussion

Young's modulus of triple helix

To determine the Young's modulus of an isolated triple helix, we assume that its thermal fluctuations can be described by the wormlike chain model. The Young's modulus (Y) of the triple helix can then be obtained from the persistence length (l_p) using the relationship (43)

$$Y = 4k_bTl_p/\pi r^4, \tag{1}$$

where, T is the temperature, k_b is the Boltzmann constant and r is the radius of the triple helix. The persistence length l_p is obtained by solving numerically the relationship (44),

$$\langle R\rangle^2 + \langle \Delta R^2\rangle = 2l_pL_c[1 - \frac{l_p}{L_c}(1 - e^{-L_c/l_p})], \tag{2}$$

where R and L_c are, respectively, the end-to-end distance and the contour length of the triple helix.

To solve Eq. 2 for a triple helix of contour length L_c, we need to determine the averages $\langle R\rangle^2$ and $\langle \Delta R^2\rangle$. From a theoretical standpoint, these averages can be obtained using any conformational sampling method that is coupled to a temperature bath. This can, however, be practically challenging when the solvent is described explicitly and when free energy barriers are large. To circumvent these issues, we compute these averages from the potential of mean force, $U(R)$, a thermodynamic quantity whose convergence with respect to conformational sampling is tractable. We use $U(R)$ as Boltzmann weights and determine

$$\langle R\rangle = \frac{\int Re^{-U(R)/k_bT}dR}{\int e^{-U(R)/k_bT}dR} \tag{3}$$

and

$$\langle \Delta R^2\rangle = \langle R^2\rangle - \langle R\rangle^2. \tag{4}$$

We determine $U(R)$ not for the entire collagen triple helix, whose contour length exceeds 300 nm, but for a fragment of the triple helix that has a counter length of $L_c = 15.3$ nm (Fig. 2a). For even such a small fragment, the simulation cell containing explicit solvent has close to 1/2 million particles. Nevertheless, the length of the fragment is chosen such that it extends beyond two true pitch lengths of the triple helix – the true pitch of a $7/2$ helix spans 21 amino acid triplets, and our fragment contains 45 amino acid triplets. While the choice of the primary sequence of the fragment can be expected to affect the computed Young's modulus, other simulation studies (18, 22) suggest that the effect will not be significant enough to alter the conclusions of this study.

Fig. 2b shows the $U(R)$ determined using well-tempered metadynamics (29–31) where conformational sampling is conducted for 5×10^7 MD steps. Using this $U(R)$ profile in Eq. 3, we find that $\langle R\rangle = 12.78$ nm, which corresponds roughly to the

position of the lowest energy in $U(R)$. Using Eq. 4, we find that $\langle \Delta R^2 \rangle = 0.04$ nm^2. Convergence is estimated in three ways. First, we track the dynamic hill height in metadynamics (Fig. 2b inset), which does not take up a value greater than 0.001 kJ/mol during the entire second half of the trajectory, and it continues to decease asymptotically, as expected theoretically (45). Second, we compute $\langle R \rangle$ and $\langle \Delta R^2 \rangle$ using $U(R)$ profiles from shorter and longer metadynamics trajectories (29). The $U(R)$ profiles from these trajectories are also shown in Fig. 2b. The shorter trajectory is comprised of 10% fewer MD steps (4.5×10^7 MD steps) and yields $\langle R \rangle = 12.69$ nm and $\langle \Delta R^2 \rangle = 0.04$ nm^2. The longer trajectory, which is comprised of 10% more MD steps (5.5×10^7 MD steps), yields $\langle R \rangle = 12.74$ nm and $\langle \Delta R^2 \rangle = 0.04$ nm^2. $\langle R \rangle$ computed from the three trajectory lengths are within 1% of each other, with no correlation between the averages and trajectory length. Finally, we compute the averages following a time-independent reconstruction of $U(R)$ (46), and find $\langle R \rangle = 12.61$ nm, which is also within 1% of the values obtained from the unweighted $U(R)$ profiles.

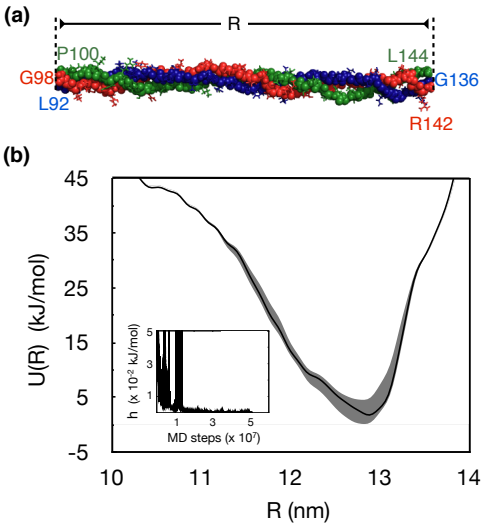

Figure 2: (***a***) Representative 45 residue fragment of the collagen molecule. The backbone atoms are drawn as spheres and the side chains are drawn as sticks. The primary sequences of the two $\alpha 1$ chains are $P^h_{100}GMK^h...PRGL_{144}$ and $G_{98}LP^hG...MGPR_{142}$, and the primary sequence of the $\alpha 2$ chain is $L_{92}P^hGF...GAKG_{136}$. The 'h' superscript over selected residues refers to their post-translationally modified hydroxylated forms, and the subscript on the terminal residues of the chains matches their specific numbers in the primary sequence taken from the fiber diffraction structure of the fibril (PDB ID: 3HR2). Note that the residue numbers in the three chains are different, and this is because the three chains in the triple helix are staggered. (***b***) Potential of mean force ($U(R)$) of this fragment evaluated as a function of its end-to-end distance (R) from well-tempered metadynamics. The solid line is an estimate from a trajectory comprising of 5×10^7 MD steps (100 ns of metadynamics time), and the shaded area indicates a range bounded by two estimates, one from a smaller trajectory comprising of 4.5×10^7 MD steps and the other from a longer trajectory comprising of 5.5×10^7 MD steps. The inset shows the time evolution of the dynamic hill height (h).

Numerical solution of Eq. 2 using $\langle R \rangle = 12.78$ nm and $\langle \Delta R^2 \rangle = 0.04$ nm^2 yields $l_p = 12.85$ nm.Plugging this value of l_p in Eq. 1, and using $r = 0.36 \pm 0.01$, which is one half of the average width of the triple helix (28) in the well-tempered metadynamics simulation, yields a corresponding $Y = 4.2$ GPa. Decreasing $\langle R \rangle$ by 1% yields $l_p = 12.09$ nm and $Y = 3.95$ GPa, and increasing $\langle R \rangle$ by 1% yields $l_p = 13.70$ nm and $Y = 4.50$ GPa. These values of persistence lengths and Young's moduli fall within the range of the most recent estimates from AFM and optical tweezer experiments (Table 1). We also note that while there are other simulation-based studies of the mechanical properties of collagen-like triple helices (Table 1), the Young's moduli in all, but one, of those studies are obtained using steered MD, where decoupling the effects of kinetics from static thermodynamic properties is challenging (47, 48).

Young's modulus of fibril core

We determine the Young's modulus of the fibril core from its stress-strain relationship. To emulate the fibril core, we simulate the crystallographic unit cell of the fibril (12, 13) under periodic boundary conditions. The crystallographic unit cell, which

is the smallest repeating unit in a fibril, is a triclinic cell whose longest lattice constant is equal to the D-band length of the fibril. The unit cell contains exactly one triple helix wrapped 4.46 times about the fibril axis, that is, the length of the triple helix $L = 4.46D$ (Fig. 1). Such a periodic system represents an infinitely wide and infinitely long fibril that lacks an interface with bulk solvent and, therefore it represents the buried core of a fibril. We note that simulating a single crystallographic unit cell under periodic conditions is an approximate representation of the buried core in that it does not incorporate aperiodic variations across multiple unit cells. Nevertheless, since the unit cell we employ is, in fact, the smallest periodic repeating unit observed in fiber diffraction studies (12, 13), we consider it as an appropriate initial model of the fibril core.

The stress-strain relationship is obtained by simulating the unit cell after straining it along the fibril axis ($\epsilon_{\parallel} > 0$) and recording the resulting stress $\sigma_{\parallel}$ (Fig. 3a). Specifically, we carry out five separate MD simulations of the unit cell, and in each simulation, the lattice constant parallel to the fibril axis (or the D-band length) is held fixed at a length $\mathrm{D} = \mathrm{D}_0(1 + \epsilon_{\parallel})$. The details of our MD simulation protocol are provided elsewhere (28), and are also described briefly in the Methods section. The evolution of the stress as a function of simulation time is shown in Fig. 3b. We find that in all cases stress-equilibration requires > 150 ns of equilibration time. Assuming that these strains are in the linear regime, we obtain the Young's modulus using Hooke's law, $Y = (\langle\sigma_{\parallel}\rangle - \sigma_0)/\epsilon_{\parallel}$, where $\sigma_0 = 0.1$ MPa (Fig. 3c). The values of $\langle\sigma_{\parallel}\rangle$ are obtained by averaging $\sigma_{\parallel}$ over the final 20 ns of each trajectory. A linear fit between $(\langle\sigma_{\parallel}\rangle - \sigma_0)$ and $\epsilon_{\parallel}$ yields $Y = 2.34$ GPa. Inserting this value in Eq. 1 and taking $r = \sqrt{\mathbf{b} \times \mathbf{c}/\pi} = 1.75$ nm, where $\mathbf{b}$ and $\mathbf{c}$ are the two smallest lattice constants of the unit cell under zero-stress condition, yields a persistence length $l_p = 4.0 \times 10^3$ nm. The values of the unit cell vectors $\mathbf{b}$ and $\mathbf{c}$ are taken from our earlier work (28).

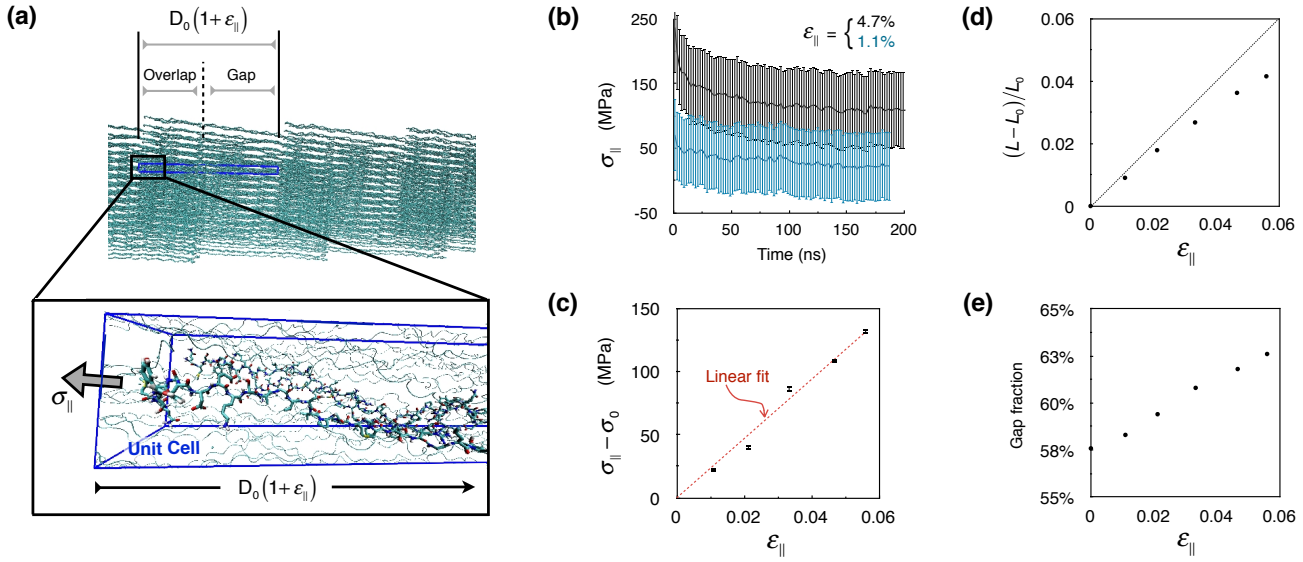

Figure 3: Constant-strain MD simulations of the crystallographic unit cell under periodic boundary conditions. (***a***) Snapshot of the MD simulation illustrating the unit cell (blue box), the strained lattice constant $|\mathbf{a}| = D_0(1+\epsilon_{\parallel})$, and the component of the virial stress tensor parallel to the fibril axis, $\sigma_{\parallel}$. (***b***) Time evolution of the $\sigma_{\parallel}$ in two different constant-strain simulations, one in which $\epsilon_{\parallel} = 0.011$, and the other in which $\epsilon_{\parallel} = 0.047$. The vertical lines represent fluctuations in $\sigma_{\parallel}$ observed over 2 ns time intervals. (***c***) Stress-strain relationship estimated for five different strain values. $\langle\sigma_{\parallel}\rangle$ are averages over the final 20 ns of the constant-strain simulations, and the standard errors are obtained via block averaging. (***d***) Relationship between the fractional change in the length of the triple-helix (L), and the applied strain, estimated from the final 20 ns of constant-strain simulations. L_0 is the length of the triple helix in the zero-strain simulation. (***e***) Relationship between the gap fraction and the applied strain, estimated from the final 20 ns of constant-strain simulations.

Analysis of the equilibrated portions of the MD simulations reveals that the fractional increase in triple helix length is not correlated linearly with the applied strain (Fig. 3d), implying that straining also induces slippage between triple helices. For the maximum applied strain, the gap fraction, estimated from axial mass densities (28), increases by 5% from 0.58 to 0.63 (Fig. 3e). Such slippage is less likely to occur under physiological conditions when the triple-helices are cross-linked. Including cross-linking in the model will only increase the Young's modulus of the fibril core, and therefore, the computed Young's modulus serves as a lower limit. Nevertheless, despite the slippage, the computed Young's modulus of the fibril

core is comparable to the Young's modulus of a triple helix and is an order in magnitude larger than the Young's moduli of fibrils/fibril-bundles (Table 1).

Conclusions

Results from our simulations suggest that the smaller Young's moduli of fibrils/fibril-bundles relative to that of their constituent triple helices cannot be attributed to the specific packing of triple helices in the fibril core. Therefore, it is not the fibril core that yields initially to axial stress. Rather, it must be the portions of the fibril exposed to the solvent and/or the fibril-fibril interface that get strained in response to the initial stress. It is conceivable that the fibril portion exposed to the solvent is packed loosely compared to the fibril core and may yield to relatively smaller axial stresses. It is also plausible that the fibril-fibril interface slips first in response to applied axial stresses. As such, the packing of fibrils into fibril-bundles is facilitated by the binding of non-collagenous molecules, such as proteoglycans, which are expected to interact with fibrils weakly compared to the interaction that holds triple helices together within fibrils (5). This explanation is supported by our recent experiments (26), which show broad distributions in both the Young's moduli (0.1–0.36 GPa) and the diameters (116-200 nm) of type I collagen fibril/fibril-bundles extracted from a single tissue-type. It is unlikely that the organization of the fibril core varies within a tissue-type (5), and it may be that the distribution in Young's moduli is systematically related to the distribution in the sizes fibrils/fibril-bundles, or the packing of the fibrils within fibril bundles. Indeed, statistical modeling of inter-fibrillar interactions suggests that the arrangement of inter-fibrillar cross-link-forming sites contributes to the broadening of the denaturation transition in fibrillar collagen (49). Additionally, studies based on a wormlike bundle model suggest that competition between the elastic properties of the filaments and those of the crosslinks leads to renormalized effective bend and twist rigidities, at least for microtubule-like or actin-like filament bundle geometries (50). Further study is, however, required to establish causality, and toward that end, this work lays the foundation by providing converged estimates of Young's moduli and persistence lengths at two levels of collagen's structural hierarchy. In general, this work provides the necessary baseline (or control) to quantitively investigate the effect of various biological factors on the mechanical properties of collagen, such as congenital mutations, post-translational modifications and ligand binding (1–7), and also engineer new collagen-based materials (51).

Author Contributions

SV, JPROO and JDS designed research; SV performed research; SV analyzed data; SV, JPROO and JDS wrote the paper.

Acknowledgements

The authors acknowledge support from the research computing center of the University of South Florida. This research was funded by DARPA grant W911NF-09-1-0378 and NSF grant MRI CHE-1531590.

References

1. Mecham, R. P. 2011. The extracellular matrix: an overview. Springer, Berlin.
2. Brinckmann, J. 2005. Collagens at a glance. Top. Curr. Chem. 247:1-6
3. Kadler, K. E., C. Baldock, J. Bella, and R.P. Boot-Handford. 2007. Collagens at a glance. J. Cell Sci. 120:1955-1958.
4. Shoulders, M. D., and R.T. Raines. 2009. Collagen Structure and Stability. Annu. Rev. Biochem. 78:929-958.
5. Orgel, J. P., J.D. San Antonio, and O. Antipova. 2011. Molecular and structural mapping of collagen fibril interactions. Conn. Tissue Res. 52:2-17
6. Ricard-Blum, S. 2011. The Collagen Family. Cold Spring Harb Perspect. Biol. 3:1-19.
7. Mienaltowski, M. J., and D. E. Birk. 2014. Structure, physiology, and biochemistry of collagens. Adv. Exp. Med. Biol. 802: 5-29.
8. Hodge, A. J., and J. A. Petruska. 1963. Recent studies with the electron microscope on ordered aggregates of the tropocollagen macromolecule. *In* Aspects of Protein Structure, G. N. Ramachandran, editor. Academic Press, London and New York, pp. 289-300.
9. Hulmes, D. J. S., and A. Miller. 1979. Quasi-hexagonal molecular packing in collagen fibrils. Nature. 282:878-880.
10. Trus, B. L., and K. A. Piez. 1980. Compressed microfibril models of the native collagen fibril. Nature. 286:300-301.
11. Hulmes, D.J., J. C. Jesior, A. Miller, C. Berthet-Colominas, and C. Wolff. 1981. Electron microscopy shows periodic structure in collagen fibril cross sections. Proc. Natl. Acad. Sci. U. S. A. 78:3567-3571.
12. Orgel, J. P., A. Miller, T. C. Irving, R. F. Fischetti, A. P. Hammersley, and T. J. Wess. 2001. The in situ supermolecular structure of type I collagen. Structure. 9:1061-1069.
13. Orgel, J. P., T. C. Irving, A. Miller, and T. J. Wess. 2006. Microfibrillar structure of type I collagen in situ. Proc. Natl. Acad. Sci. U. S.

A. 103:9001-9005.
14. Bozec, L., G. van der Heijden, and M. Horton. 2007. Collagen fibrils: nanoscale ropes. Biophys. J. 92:70-75.
15. Sasaki, N., and S. Odajima. 1996. Stress-strain curve and young's modulus of a collagen molecule as determined by the X-ray diffraction technique. J. Biomechanics. 29:655-658.
16. Sun, Y. L., Z. P. Luo, A. Fertala, and K. N. An. 2002. Direct quantification of the flexibility of type I collagen monomer. Biochem. Biophys. Res. Comm. 295:382-386.
17. Lovelady, H. H., S. Shashidhara, and W. G. Matthews. 2014. Solvent Specific Persistence Length of Molecular Type I Collagen. Biopolymers. 101:329-335.
18. Vesentini, S., C. F. C. Fitie, F. M. Montevecchi, and A. Redaelli, 2005. Molecular assessment of the elastic properties of collagen-like homotrimer sequences. Biomech. Model Mechanobiol. 3:224-234.
19. Lorenzo, A. C., and E. R. Caffarena. 2005. Elastic properties, Youngs modulus determination and structural stability of the tropocollagen molecule: a computational study by steered molecular dynamics. J. Biomechanics. 38:1527-1533.
20. Buehler, M. J. 2006. Atomistic and continuum modeling of mechanical properties of collagen: Elasticity, fracture, and self-assembly. J. Mater. Res. 21:1947-1961.
21. Gautieria, A., M. J. Buehler, and A. Redaellia. 2009. Deformation rate controls elasticity and unfolding pathway of single tropocollagen molecules. J Mech. Behav. Biomed. Mat. 2:130-137.
22. Teng, X., and W. Hwang. 2014. Chain Registry and Load-Dependent Conformational Dynamics of Collagen. Biomacromolecules. 15, 3019-3029.
23. Sasaki, N., and S. Odajima. 1996. Elongation mechanism of collagen fibrils and force-strain relations of tendon at each level of structural hierarchy. J Biomech. 29:1131-1136.
24. Yang, L., K. O. van der Werf, C. F. C. Fitie, M. L. Bennink, P. J. Dijkstra, and J. Feijen. 2008. Mechanical Properties of Native and Cross-linked Type I Collagen Fibrils. Biophys. J. 94:2204-2211.
25. Shen, Z. L., H. Kahn, R. Ballarini, and S. J. Eppell. 2011. Viscoelastic Properties of Isolated Collagen Fibrils. Biophys. J. 100:3008-3015.
26. Dutov, P.M., O.A. Antipova, S. Varma, J.P.R.O. Orgel, and J. D. Schieber. 2016. Measurement of Elastic Modulus of Collagen Type I Single Fiber. PLoS ONE 11:e0145711.
27. Grant, C. A., D. J. Brockwell, S. E. Radford, and N. H. Thomson. 2009. Tuning the Elastic Modulus of Hydrated Collagen Fibrils. Biophy. J. 97:2985-2992.
28. Varma, S., M. Botlani, J. R. Hammond, H. L. Scott, J. P. R. O. Orgel, and J. D. Schieber. 2015. Effect of intrinsic and extrinsic constraints on the simulated D-band length of Type I collagen. Proteins. 83:1800-1812.
29. Barducci, A., G. Bussi, and M. Parrinello. 2008. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. Phys. Rev. Lett. 100:020603
30. Laio, A., and F. L. Gervasio. 2008. Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry and materials sciences. Rep. Prog. Phys. 71:126601.
31. Bonomi, M., D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello. 2009. Plumed: A portable plugin for free-energy calculations with molecular dynamics. Comput. Phys. Commun. 180:1961-1972.
32. Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. 2006. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. Proteins. 65:712-725.
33. Lindorff-Larsen, K., S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins. 78:1950-1958.
34. Park, S., R. J. Radmer, T. E. Klein, and V. S. Pande. 2005. A New Set of Molecular Mechanics Parameters for Hydroxyproline and Its Use in Molecular Dynamics Simulations of Collagen-Like Peptides. J. Comput. Chem. 26:1612-1616.
35. Berendsen, H. J. C., J. R. Grigera, and T. P. Straatsma. 1987. The missing term in effective pair potentials. J. Phys. Chem. 91:6269-6271.
36. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: An NLog(N) method for Ewald sums in large systems. J. Chem. Phys. 98:10089-10092.
37. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity-rescaling. J. Chem. Phys. 126: 014101.
38. Parrinello, M., and A. Rahman. 1981. Polymorphic transition in single crystals: A new molecular dynamics method. J. Appl. Phys. 52:7128-7190.
39. Hess, B. 2008. P-LINCS: A parallel linear constraint solver for molecular simulation. J. Chem. Theory Comput. 4:116-122.
40. Miyamoto, S., and P. Kollman. 1992. SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water molecules. J. Comput. Chem. 13:952-962.
41. Okuyama, K.; Xu, X.; Iguchi, M.; Noguchi, K. Revision of collagen molecular structure. Biopolymers 2006, 84, 181-191.
42. Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl. 2008. GROMACS 4 Algorithms for highly efficient load-balanced and scalable molecular simulation. J. Chem. Theory Comput. 4:435-447.
43. de Pablo, J.J., and J. D. Schieber. 2014. Molecular Engineering Thermodynamics. Cambridge University Press.
44. Bustamante, C., J.F. Marko, E.D. Siggia, and S. Smith. 1994. Entropic elasticity of lambda-phage DNA. Science. 265:1599-1600.
45. Dama, J. F., M. Parrinello, and G. A. Voth. Well-tempered Metadynamics Converges Asymptotically. 2014. Phys. Rev. Lett. 112: 240602.

46. Tiwary, P. and M. Parrinello. A Time-Independent Free Energy Estimator for Metadynamics. 2015. J. Phys. Chem. B. 119, 736-742.
47. Jarzynski, C. 1997. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. Phys. Rev. E. 56:5018.
48. Park, S., F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten. 2003. Free Energy Calculation from Steered Molecular Dynamics Simulations Using Jarzynskis Equality. J. Chem. Phys. 119: 3559.
49. Bennett, C. B., J. Kruczek, D. A. Rabson, W. G. Matthews, and S. A. Pandit. 2013. The effect of cross-link distributions in axially-ordered, cross-linked networks. J Phys. Condens. Matter. 25: 285101.
50. Heussinger, C., F. Schuller, and E. Frey. 2010. Statics and dynamics of the wormlike bundle model. Phys. Rev. E. 81: 021904.
51. Drzewiecki, K. E., A. S. Parmar, I. D. Gaudet, J. R. Branch, D. H. Pike, V. Nanda, and D. I. Shreiber. 2014. Methacrylation Induces Rapid, Temperature-Dependent, Reversible Self-Assembly of Type-I Collagen. Langmuir. 30: 11204-11211

List of Figures

List of Tables