

# Entropic Elasticity Controls Nanomechanics of Single Tropocollagen Molecules

Markus J. Buehler\* and Sophie Y. Wong†

\*Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, and †Division of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

**ABSTRACT** We report molecular modeling of stretching single molecules of tropocollagen, the building block of collagen fibrils and fibers that provide mechanical support in connective tissues. For small deformation, we observe a dominance of entropic elasticity. At larger deformation, we find a transition to energetic elasticity, which is characterized by first stretching and breaking of hydrogen bonds, followed by deformation of covalent bonds in the protein backbone, eventually leading to molecular fracture. Our force-displacement curves at small forces show excellent quantitative agreement with optical tweezer experiments. Our model predicts a persistence length  $\xi_p \approx 16$  nm, confirming experimental results suggesting that tropocollagen molecules are very flexible elastic entities. We demonstrate that assembly of single tropocollagen molecules into fibrils significantly decreases their bending flexibility, leading to decreased contributions of entropic effects during deformation. The molecular simulation results are used to develop a simple continuum model capable of describing an entire deformation range of tropocollagen molecules. Our molecular model is capable of describing different regimes of elastic and permanent deformation, without relying on empirical parameters, including a transition from entropic to energetic elasticity.

## INTRODUCTION

Collagen, the most abundant protein on earth, is a fibrous structural protein with superior mechanical properties, and provides an intriguing example of a hierarchical biological nanomaterial. Collagen consists of triple helical tropocollagen (TC) molecules that have highly conserved lengths of  $L \approx 300$  nm, roughly 1.5 nm in diameter (1–7). Staggered arrays of TC molecules form fibrils, which arrange to form collagen fibers (1,2,8).

Collagen plays an important role in many biological tissues, including tendon, bone, teeth, cartilage, and in the cardiovascular system. Collagen is typically subjected to mechanical tensile load under physiological conditions (8–10).

Due to lack of understanding of nanomechanical properties of collagen, the source of the elasticity of individual TC molecules, fibrils, and collagen fibers remains controversial.

There are several reports of experimental studies focused on the mechanics of single tropocollagen fibers (7,15–20). However, tropocollagen molecules have rarely been studied using molecular dynamics (MD) studies. In a recent investigation, the mechanical properties of collagen fibers, using MD studies, focus on their Young's modulus (11) for tensile strains up to 4%. Another recent study was carried out using molecular mechanics, considering the stretching mechanics of tropocollagen molecules up to 10% deformation (5). Other studies focused on the stability of collagen molecules (12) or the effect of point mutations on the stability (13,14). Some

researchers modeled collagen at the continuum scale, using techniques such as finite element modeling (15).

However, no studies of the large deformation behavior of tropocollagen molecules, including elastic and fracture regimes, have been reported. Further, most existing atomistic studies were limited to short sequences of tropocollagen proteins that did not exceed 10 nm, thus significantly below the length of molecules extracted from tissue, which typically features molecular lengths  $\approx 300$  nm (1–7). Atomistic investigations of the bending mechanics and the mechanics of long tropocollagen molecules at finite temperature are still missing. Such studies are vital to understand the effect of entropic contributions to elasticity and its competition with energetic terms.

Therefore, limited understanding exists about the transition from entropic to energetic elasticity, under which conditions either one dominates, and how properties change when individual TC molecules assemble into collagen fibrils (8). Understanding the nanomechanical properties of single TC molecules is essential to develop theories that link molecular and tissue properties of collagen.

Early experiments suggested that TC molecules are rigid, rod-shaped structures (16,17). However, based on hydrodynamic methods and transmission electron microscopy (TEM), it was found that TC molecules feature some flexibility, which can be measured by the persistence length. Hydrodynamic methods suggested a persistence length of 130–180 nm (18–20); TEM-based estimates (21) range between 40 nm and 60 nm. Recent experiments using optical tweezers suggested a much lower persistence length, with values ranging between 11 nm and 15 nm (22–24). Thus far, no satisfying explanation for these discrepancies has been

Submitted December 6, 2006, and accepted for publication February 22, 2007.

Address reprint requests to M. J. Buehler, Tel.: 626-628-4087; E-mail: mbuehler@MIT.EDU.

Editor: Ivet Bahar.

© 2007 by the Biophysical Society

0006-3495/07/07/37/07 \$2.00

doi: 10.1529/biophysj.106.102616

provided. No molecular model of TC has been developed capable of describing different regimes of deformation.

Here we report hierarchical modeling consisting of sequential atomistic and molecular simulation that confirms optical tweezer results quantitatively, without relying on empirical parameters. We employ a mesoscale molecular model as a strategy to upscale the mechanical properties of single TC molecules to large timescales to enable capturing entropic effects. Our model is capable of carrying out molecular stretching experiments, with high precision, at time-scales of tens of microseconds, approaching experimental deformation rates.

We find that when the TC molecule is stretched below its contour length ( $x < L$ ), entropic elasticity governs elasticity. When the TC molecule is stretched close to and beyond its contour length ( $x > L$ ), we observe a transition to energetic elasticity, featuring a two-regime hyperelastic or nonlinear elastic behavior due to unfolding and breaking of hydrogen bonds and stretching of the polypeptide backbone, followed by molecular fracture. We find that assembling TC molecules into fibrillar structures leads to rapid increase of the persistence length, rendering entropic elasticity less significant.

## MATERIALS AND METHODS

Our molecular, mesoscale model is derived from a series of full atomistic simulations, similar to those described in the literature (25,26). We note that the focus of this article is not development of the mesoscale model. Instead, here we focus on applying the mesoscale model to describe deformation of ultra-long TC molecules.

However, extending earlier work (25), here we pay particular attention to the bending stiffness of TC molecules and its dependence on the loading rate. Therefore, we carry out additional molecular dynamics simulations of bending a short sequence of a TC molecule.

### Atomistic model

For the full atomistic simulations, we use the classical force field CHARMM (27,28), implemented in the molecular dynamics (MD) program NAMD (29). The CHARMM force field (27,28) is widely used in the protein and biophysics community, and provides a basic description of the atomistic interactions in proteins. It is based on harmonic and anharmonic terms describing covalent interactions, in addition to long-range contributions describing van der Waals interactions and ionic (Coulomb) interactions as well as hydrogen-bonding.

We visualize the molecular structures using the visual molecular dynamics (30) program. We consider the atomic coordinates of a single TC molecule based on the crystal structure PDB ID 1QSU (31) with 1.75 Å resolution, a short sequence of a TC molecule with length  $L \approx 8.4$  nm. The 1QSU is a triple-helical type II collagen-like molecule with sequence (Pro-Hyp-Gly)<sub>4</sub>-Glu-Lys-Gly-(Pro-Hyp-Gly)<sub>5</sub>.

This model serves as a simple representation for the amino-acid sequence that makes up the sequence of longer tropocollagen molecules extracted from live tissue.

We use the psfgen program of NAMD (29) to add hydrogen atoms and to assign charges. The molecular model is set up according to pH 7, resembling conditions in experiment (24). No electrolytes are present in our atomistic setup.

We include crystallographic water in the simulations, and also add a skin of water of a few Ångströms surrounding the tropocollagen molecule. We

note that electrostatic interactions can be significantly weakened by screening due to solvent molecules, which can lead to interactions that are weaker than van der Waals interactions and that decay over distances of several Ångströms (32).

Simulations with forces applied are carried out after energy minimization and finite temperature relaxation at  $T = 300$  K. All simulations reported in this article are carried out at  $T = 300$  K with temperature control (NVT ensemble), using a time step  $\Delta t = 1$  fs. We apply forces to the molecule using an SMD scheme (33). Forces and fixed constraints are applied at the C $_{\alpha}$  atom in each chain. Forces are applied on each of the three  $\alpha$ -chains equally. In the case of a fixed end, each C $_{\alpha}$  at the end of the molecule is fixed. No torsional loads are applied since the end of the tropocollagen molecule is free to rotate. This is essential to allow for proper unfolding of the molecular structure.

Forces in the SMD scheme are applied by adding a potential energy term to the Hamiltonian of the system,

$$\phi_{\text{SMD}}(\vec{r}, t) = \frac{1}{2} k_{\text{SMD}} (\dot{r}t - (\vec{r}(t) - \vec{r}_0) \cdot \vec{n}), \quad (1)$$

where  $k_{\text{SMD}}$  is the SMD spring constant,  $\dot{r}$  the pulling rate,  $\vec{r}(t)$  the current position vector of the atom being pulled,  $\vec{r}_0$  the position vector of the atom at  $t = t_0$  (beginning of simulation), and  $\vec{n}$  the normalized direction vector in which the atom is being pulled.

To carry out the tensile experiments, we use a SMD spring constant  $k_{\text{SMD}} = 10$  kcal/mol/Å<sup>2</sup>, with  $\dot{r} = 0.0001$  Å/step.

Postprocessing of the simulations is done using the CMDf framework (34), a computational Python-based simulation environment capable of seamless integration of various file formats and computational engines, including molecular and crystal building tools. CMDf Python scripts are used to analyze and postprocess the simulation data, as required, for example, to compute statistical averages of force-displacement curves.

Fig. 1 depicts results of stretching calculations of a single TC molecule, based on the CHARMM model and the ReaxFF model (for details and simulation conditions, see (25)). Fig. 2 shows snapshots of the atomistic geometry of the tropocollagen molecule as it undergoes tensile deformation.

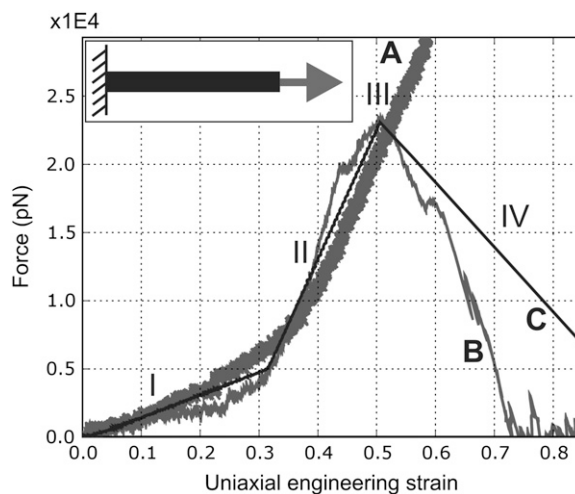


FIGURE 1 Force-displacement response of a single TC molecule with  $L \approx 8.4$  nm under tensile loading, showing results obtained using different methods (nonreactive CHARMM force field (A) (27), ReaxFF (B), and the reactive mesoscale model (C) (25)). We distinguish four regimes as the molecule undergoes deformation: Regime I is characterized by uncoiling of the triple helical structure, and Regime II with a larger tangent modulus corresponds to stretching of covalent bonds. In Regime III, we observe molecular fracture followed by rapid decay of the force in Regime IV.

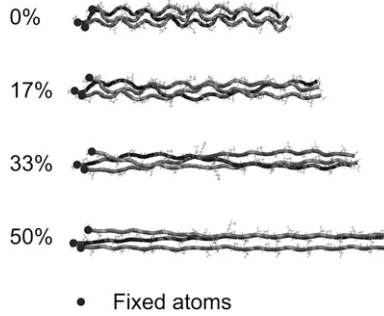


FIGURE 2 Snapshots during deformation of a single tropocollagen molecule, for different levels of tensile strain ranging from 0% to 50%.

The curves reveal the highly nonlinear or hyperelastic stress-strain behavior of a single TC molecule as the tensile strain increases up to 50%. We observe four regimes as the molecule undergoes large tensile deformation: Regime I is characterized by uncoiling of the triple helical structure; Regime II with a larger tangent modulus corresponds to stretching of covalent bonds in the protein backbone. In Regime III, we observe molecular fracture due to breaking of peptide bonds, followed by rapid decay of the force in Regime IV. The stiffness ratio of the second to first regime is  $\sim 8$ , underlying the significant stiffening effect at tensile deformation beyond  $\sim 30\%$ .

To determine the bending stiffness, we fix both ends of the TC molecules. This is achieved by fixing the terminal  $C_\alpha$  atom in each of the three polypeptides. We apply a force in the center of mass of the TC molecule using steered molecular dynamics (using  $k_{\text{SMD}} = 0.01$  kcal/mol/Å<sup>2</sup>). To study the variation of bending stiffness with deformation rate, we simulate this three-point bending experiment at varying displacement rates, where  $\dot{\epsilon} = 0.01$  to 0.1 m/s. The applied force is orthogonal to the axial direction of the TC molecule, leading to bending of the molecule upon increasing force. From the force-displacement data obtained by atomistic modeling, the bending stiffness  $EI$  is given by

$$EI = \frac{F_{\text{appl}} L^3}{48d}, \quad (2)$$

where  $F_{\text{appl}}$  is the applied force, and  $d$  is the bending displacement. This equation can be derived by solving a simple beam problem with proper boundary conditions ( $u(0) = u(L) = 0$  and no moment applied at the boundaries), and provides a link between atomistic and nanocontinuum mechanics.

Fig. 3 depicts the bending stiffness as a function of displacement rate. The results suggest a linear relationship between the bending stiffness and the deformation rate. Extrapolation to zero deformation rate yields an estimate for the quasi-static bending stiffness,  $EI \approx 1.247 \times 10^{-29}$  Nm<sup>2</sup>.

We note that the TC molecule is several thousand times stiffer than a carbon nanotube of comparable diameter (e.g., a (5,5) carbon nanotube).

## Molecular model

The full atomistic calculation results in feed parameters for the molecular model, representing TC molecules at a mesoscale, as shown schematically in Fig. 4. Our molecular model is implemented in LAMMPS (35), and parameters are identical to the model described in the literature (25,26). Here we chose a smaller bead-to-bead distance of 7 Å and adapt the bending stiffness to the extrapolated value for zero deformation rate (see Fig. 3). The effects of the solvent are treated implicitly in the model.

Since the parameters in our molecular model are fitted directly to match the nanomechanical properties of the single TC molecule (PDB ID 1QSU), our model intrinsically assumes that the sequence along the length of the tropocollagen molecule does not vary. This is an approximation since col-

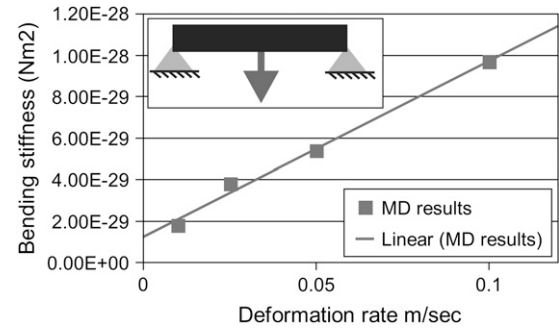


FIGURE 3 Dependence of bending stiffness  $EI$  of a single TC molecule on the deformation rate,  $L \approx 8.4$  nm, at  $T = 300$  K. The results indicate that  $EI$  decreases linearly with decreasing loading rate. A linear fit to the data enables us to extrapolate to  $\dot{\epsilon} \rightarrow 0$ . The resulting bending stiffness for vanishing deformation rate is used as input parameter for the mesoscale molecular model.

lagen extracted from live tissue displays sequence variations along the axial directions. These sequence variations may induce modifications of the mechanical properties, including changes to the bending stiffness. These effects are neglected in the present study and left to future work.

The bottom subplot in Fig. 4 depicts how the load is applied to the long tropocollagen molecule using the SMD scheme (33). The outermost particle of the molecule is kept fixed during the simulation, and a slowly increasing force is applied on the right end. The SMD spring constant is chosen to be  $k_{\text{SMD}} = 1$  kcal/mol, and we simulate deformation at a pulling velocity of 0.025 m/s (for the SMD formulation, see Eq. 1; the SMD model in the atomistic and molecular model have the same formulation). The continuously increasing force slowly stretched the molecule from a convoluted initial shape to a straight configuration, followed by further extension and fracture. During the simulation, we record snapshots of magnitude of force applied and corresponding end-to-end distance of the molecule. This force-extension data is averaged into bins after the simulation is over to obtain a

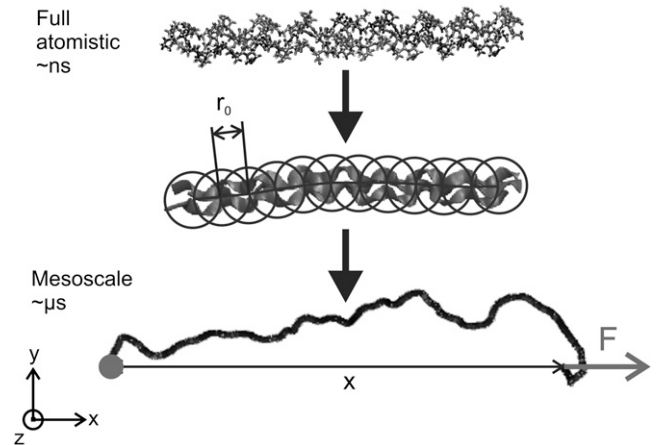


FIGURE 4 Schematic showing development of the coarse-grained molecular model from a full atomistic description. The full atomistic representation of the triple-helical TC structure is replaced by a collection of beads, each of which represents  $\approx 10$  protein atoms plus water molecules (25), with  $r_0 = 7$  Å. The bottom subplot depicts how the load is applied to the long tropocollagen molecule. The outermost particle of the molecule is kept fixed during the simulation, and a slowly increasing force is applied on the right end.

smooth curve. We find that further reduction of loading rate shows negligible changes in the force-extension curve.

We consider a stretching experiment of a TC molecule with  $L = 301.7$  nm. The time step used in molecular modeling is  $\Delta t = 20 \times 10^{-15}$  s, and we reach a total simulation time of  $t_{\text{tot}} \approx 20 \times 10^{-6}$  s. Each mesoscale particle has a mass of  $m = 1358.7$  amu, accounting for the mass of the TC segment it represents plus a skin of water molecules in a radius of 3 Å.

All calculations are carried out in three dimensions.

## RESULTS AND DISCUSSION

Fig. 5 *a* shows the entire force-displacement curve of stretching a single TC molecule ( $x$  denotes the end-to-end distance of the TC molecule).

Loading of the TC molecule starts from a coiled entangled configuration of the molecule with end-to-end distance of

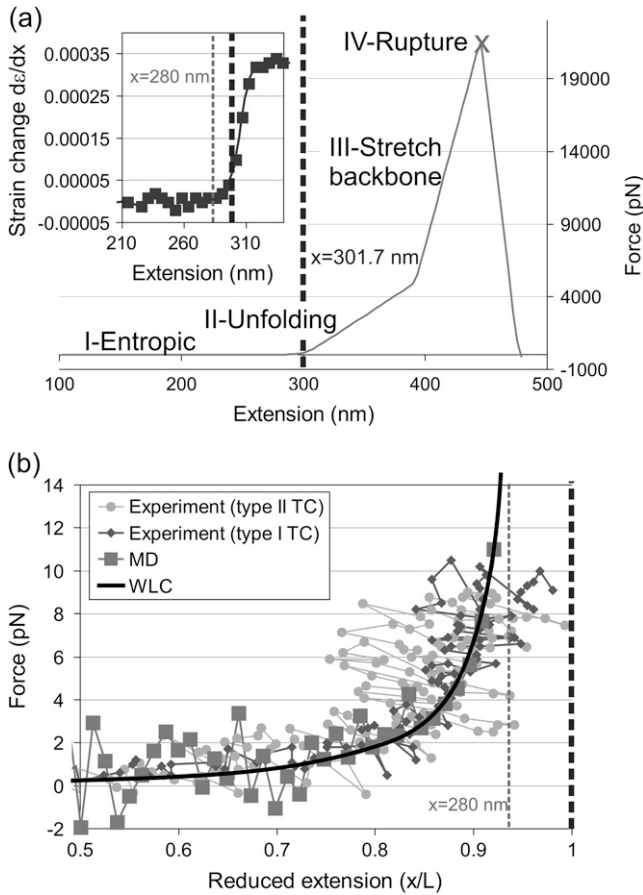


FIGURE 5 Force-displacement ( $F(x)$ ) curves of stretching a single TC molecule,  $L = 301.7$  nm, at 300 K. Subplot (a): Force-displacement curve over the entire deformation range, covering four stages: (I) uncoiling of the entangled configuration, (II) uncurling of the triple helix, (III) stretching of covalent bonds in the individual polypeptides, and (IV) rupture of the TC molecule. The dashed line indicates the contour length of the molecule. The inset depicts  $d\epsilon/dx(x)$ , showing that entropic contributions dominate for  $x < 280$  nm. Subplot *b* depicts a subset of the results depicted in the previous figure, focusing on the small force, entropic response ( $F < 14$  pN). This plot also depicts experimental results obtained for TC molecules with similar contour lengths (23,24), as well as the prediction of the WLC model with persistence length  $\xi_p \approx 16$  nm.

$\approx 100$  nm. During the initial regime (Regime I), the molecule loses its entangled structure, while the applied forces remain relatively low. Once the contour length is reached ( $x = L$ ), (Regime II) uncurling of the triple helix, (Regime III) stretching of covalent bonds in the individual polypeptides, and (Regime IV) rupture of the TC molecule occurs, followed by a sharp drop of the forces to zero.

The qualitative behavior of TC molecules under stretch is similar to recent experimental studies carried out on collagen fibrils that reach forces on the order of several  $\mu\text{N}$  (36), also showing a significant hyperelastic stiffening effect.

Fig. 5 *b* depicts a detailed view into Regime I, including a quantitative comparison with experiment (23,24). The plot reveals very good agreement, even though the deformation rate in MD is still much higher.

The elastic properties of materials can be expressed as partial derivatives of the free energy density with respect to strain  $\epsilon$ . Based on the free energy density  $A_V = (U - TS)/V$  (where  $V$  stands for the system volume), the scalar stress  $\sigma = \partial A_V / \partial \epsilon$ . At constant temperature, the entropic regime is characterized by  $\partial U / \partial \epsilon \approx 0$ . In contrast, during the energetic regime,  $\partial S / \partial \epsilon \approx 0$ .

Changes in the internal energy  $U$  can be correlated with changes in the average internal tensile strain in the TC molecule ( $d\epsilon(x)/dx$ ), where

$$\epsilon = \frac{1}{N-1} \sum_{i=1, \dots, N-1} \epsilon_i \quad (3)$$

and

$$\epsilon_i = \frac{|\vec{x}_{i+1} - \vec{x}_i|}{r_0}. \quad (4)$$

We note that  $U = f(\epsilon)$ . The result of  $d\epsilon/dx(x)$  shown in the inset of Fig. 5 *a* confirms that the regime  $x < 280$  nm is dominated by entropic elasticity since  $d\epsilon/dx \approx 0$ .

For  $x \geq 280$  nm,  $\epsilon$  increases slowly, leading to a change of internal energy that contributes to the elastic response. We find that  $\epsilon \approx 1.2\%$  when  $x \rightarrow L$ . These changes indicate the onset of a smooth transition from entropic to energetic elasticity at  $x = L$ .

Energetic elasticity dominates when  $x > L$  and leads to a significantly stiffer response. Entropic contributions to the elastic response decay quickly and approach zero at  $x \approx 317$  nm, suggesting that energetic terms dominate for  $x > 317$  nm. This is confirmed by considering that  $d\epsilon/dx$  approaches a constant value (see inset of Fig. 5 *a*).

By fitting the Marko-Siggia estimate of the worm-like chain (WLC) model (37,38) to the MD results (for  $F < 11$  pN), we determine an effective persistence length  $\xi_p \approx 16$  nm, with  $L = 301.7$  nm.

Even though the WLC model is a reasonable approximation for small forces, it ceases to hold when significant deformation occurs and the molecule is stretched toward its contour length (see Fig. 7 *a*, curve  $f_{\text{WLC}} - f_1$ ).

The significance of entropic effects for a given end-to-end distance  $x$  is characterized by the ratio  $\alpha = L/\xi_p$ , noting that  $\xi_p \sim EI$ .

Assembly of TC molecules into fibrils leads to strongly increased bending stiffness, thus reducing the significance of entropic elasticity. Fig. 6 depicts the bending stiffness as a function of the number of molecules in a fibril, suggesting a power-law scaling  $EI \sim n^p$ , where  $p \approx 2.5$ . Elastic beam theory predicts  $p = 3$ , confirming the results of our numerical model.

We note that, for simplicity, in this study we only consider a simple stacking without axial displacement, as shown in the inlay of Fig. 6. We note that collagen fibrils show a staggered arrangement of molecules (8). This effect is not accounted for in the present article. To implement the three-point bending test, we fix two outermost left and right particles in the lowest molecule. We apply a slowly increasing force in the center particle of the top molecule using the SMD scheme.

Collagen fibrils have diameters of several tens of nanometers and consist of hundreds of aligned TC molecules. Since  $\alpha \sim EI^{-p}$ ,  $\alpha < 1$  for  $n > 10$ —suggesting that the elasticity in collagen fibrils includes significant energetic contributions, particularly in the regions of dense molecular packing. We have confirmed the effect of increasing  $EI$  by direct MD calculations that clearly show the reduction in entropic elasticity, leading to much lower forces in the stretching regime below the contour length ( $x < L$ ). We note that while the bending stiffness increases significantly after assembly, our model predicts that the tensile stiffness or Young's modulus decreases continuously with increasing number of molecules.

However, a possible source of entropic elasticity in collagen fibrils are hydroxyproline-deficient sequences (labile domains) that represent more floppy parts of a collagen fibril with reduced molecular density (6–8,39), as they occur, for instance, in the gap regions. The mechanism studied in this article may thus relate to the transition from less ordered to

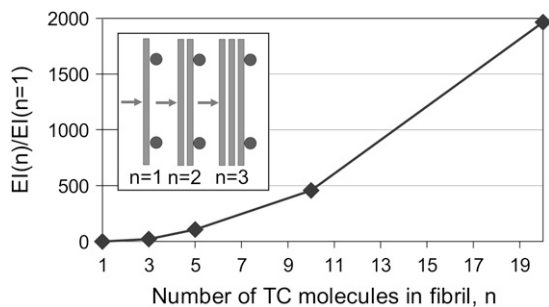


FIGURE 6 Increase in bending stiffness of a collagen fibril, as a function of number of TC molecules in the radial direction. The plot shows the bending stiffness normalized by the value for a single TC molecule. The results suggest a strong increase of the bending stiffness with number of molecules. A power-law fit suggests a scaling of  $EI \sim n^{2.5}$ .

more ordered molecule states, as discussed in the literature (6,7), where it was suggested that disappearance of “thermally activated kinks in the collagen molecules” may account for the entropic elasticity in fibrils. Thus, the entropic regime found in this article may account for the initial regime of stress-strain behavior for fibril strains up to 8%, shown in Fig. 6 in Misof et al. (6).

In summary, the toe region may relate to Regime I (entropic elasticity), and the heel region may relate to Regime II (energetic elasticity) (6,7). The qualitative shape of the force-extension curves showing the strong progressive increase of the tangent slope with strain is found in both our simulations and experiments.

However, since the length of labile domains in collagen fibrils is typically limited to 60 residues (or  $\sim 20$  nm), quantifying their relative contribution to elasticity of a collagen fibril remains unclear. We leave further investigations of this topic to future research.

Further, variations of the molecule properties along the molecule axis may change the stretching mechanics of single tropocollagen molecules. This is because the variation of sequence along the axial direction leads to locally varying bending stiffnesses and therefore modifications in the local persistence length. Such effects could be analyzed by carrying out studies similar to the one presented here, by considering different amino-acid sequences and then mapping them onto a axial variation of nanomechanical properties. We emphasize that our molecular modeling approach can, in principle, easily be adapted to account for such effects.

We use the MD results to develop a continuum model for the elasticity of TC molecules. The force as a function of distance ( $F(x)$ ) is given by

$$F(x) = \begin{cases} f_{\text{WLC}}(x) - f_1(x) & \text{if } x < L, \\ f_{\text{ENG}}(x) + f_2(x) & \text{if } x \geq L, \end{cases} \quad (5)$$

where (37,38)

$$f_{\text{WLC}}(x) = \frac{k_B T}{\xi_p} \left( \frac{1}{4(1-x/L)^2} - \frac{1}{4} + \frac{x}{L} \right), \quad (6)$$

and

$$f_{\text{ENG}}(x) = H(x_{\text{break}} - x) \begin{cases} k^{(0)}(x - L) & \text{if } x < x_1, \\ k^{(1)}(x - \tilde{x}_1) & \text{if } x \geq x_1. \end{cases} \quad (7)$$

The functions  $f_1$  and  $f_2$  are correction functions that describe the transition from entropic to energetic elasticity. We note that  $f_1(x) = 0$  for  $x \rightarrow 0$ ,  $f_1(x) \rightarrow \infty$  for  $x \rightarrow L$  (due to the divergence of the WLC model), and  $f_2(x) = 0$  for  $x \gg L$ , but  $f_2$  reaches a finite value for  $x \rightarrow L$ . Note that  $\tilde{x}_1 = x_1 - k^{(0)}/k^{(1)}(x_1 - L)$ .

Both functions are determined numerically by MD (see Fig. 7). For  $L = 301.7$  nm,  $x_{\text{break}} = 452.6$  nm,  $x_1 = 39.0$  nm,  $k^{(0)} = 51.7$  pN/nm,  $k^{(1)} = 294.8$  pN/nm, and  $\xi_p = 16$  nm. The predictions of our model agree with the numerical results shown in Fig. 5.

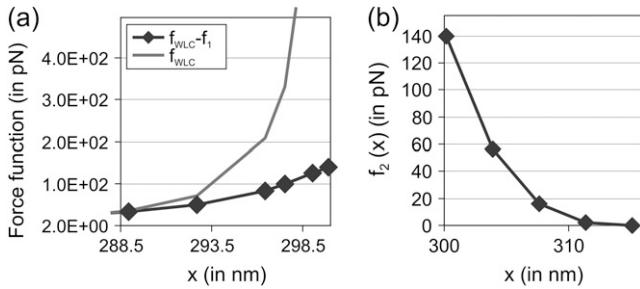


FIGURE 7 Numerical estimates of the functions  $f_{WLC} - f_1(x)$  and  $f_2(x)$ . These functions describe the smooth transition from entropic elasticity to energetic elasticity. The results show that entropic forces reach a maximum when the molecule is stretched to its contour length, approaching 140 pN.

## CONCLUSIONS

Our model includes contributions from entropic elasticity, energetic elasticity, and molecular fracture, within a single formulation, and does not contain empirical fitting parameters. All physical properties of TC molecules are determined from full atomistic calculations. Based on our force-stretching data, we find a persistence length of  $\xi_p \approx 16$  nm. Our force-stretching curves are in quantitative agreement with results obtained by optical tweezer experiments (see Fig. 5 b) (23,24), which suggested persistence lengths between 11 and 15 nm.

We note that our estimate of  $\xi_p$  from the force-stretching curve is larger than the value obtained by considering

$$\xi_p^{EI} = \frac{EI}{k_B T}, \quad (8)$$

where  $EI$  is the bending stiffness of the molecule. This equation leads to  $\xi_p^{EI} \approx 3$  nm. This difference could be attributed to coupling between bending and tension deformation modes that are not captured in the WLC theory, but included in our dynamical calculations, leading to an increased persistence length.

This suggests that  $\xi_p$  determined from force-extension curves in both molecular dynamics and experiment (23,24) represents an effective persistence length. Calculation of the persistence length from the tangent correlation function confirms that  $\xi_p > \xi_p^{EI}$ , suggesting that  $\xi_p \approx 30 \dots 45$  nm. Interestingly, this is close to experimental TEM results obtained by optical methods (21) that are based on a similar geometrical analysis scheme. These observations may help to explain the discrepancies noted in earlier studies (23,24).

An alternative reason for the large variations in persistence lengths in experiment (18–20) is based on the idea that not a single molecule, but an assembly of several molecules has been stretched, effectively increasing  $EI$  that leads to an increased persistence length (24). This possibility is not inconsistent with our model (see Fig. 6).

The observed transition from an entropic to an energetic elastic regime found here may correspond to experimental observations of progressive increase in the tangent slope of the

stress-strain curve during stretching of collagen fibrils (6,7). Our results confirm that the transition from less-ordered to more-ordered molecule states as discussed in the literature (6,7) can lead to the characteristic mechanical behavior.

Our results illustrate a possible strategy to develop molecular or mesoscale models of complex biological systems, considering various mechanisms of deformation, ranging from entropic to energetic elasticity, considering fracture mechanics of individual molecules. Similar methods could be used for other biopolymers with similar geometry. Future studies of mechanics of individual and ensembles of TC molecules could help to shed light on assembly processes or help to clarify the mechanisms during cellular production of collagen molecules. Further, the effect of sequence changes along the molecular axis is an important subject of study.

Part of this research was supported by the Army Research Office, grant No. W911NF-06-1-0291 (program officer Dr. Bruce LaMattina). S.W. was supported by MIT's Undergraduate Research Opportunities Program. Support from both sources is most greatly appreciated.

## REFERENCES

1. Ramachandran, G. N., and G. Kartha. 1955. Structure of collagen. *Nature*. 176:593–595.
2. Rich, A., and F. Crick. 1955. The structure of collagen. *Nature*. 176: 915–916.
3. Bhattacharjee, A., and M. Bansal. 2005. Collagen structure: the madras triple helix and the current scenario. *IUBMB Life*. 57:161–172.
4. Jager, I., and P. Fratzl. 2000. Mineralized collagen fibrils: a mechanical model with a staggered arrangement of mineral particles. *Biophys. J.* 79:1737–1746.
5. Vesentini, S., C. Fitie, F. Montecocchi, and A. Redaelli. 2002. Molecular assessment of the elastic properties of collagen-like homotrimer sequences. *Biomech. Model. Mechanobiol.* 3:224–234.
6. Misof, K., G. Rapp, and P. Fratzl. 1997. A new molecular model for collagen elasticity based on synchrotron x-ray scattering evidence. *Biophys. J.* 72:1376–1381.
7. Puxkandl, R., I. Zizak, O. Paris, J. Keckes, W. Tesch, S. Bernstorff, P. Purslow, and P. Fratzl. 2002. Viscoelastic properties of collagen: synchrotron radiation investigations and structural model. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357:191–197.
8. Hulmes, D. J. S., T. J. Wess, D. J. Prockop, and P. Fratzl. 1995. Radial packing, order, and disorder in collagen fibrils. *Biophys. J.* 68:1661–1670.
9. Weiner, S., and H. D. Wagner. 1998. The material bone: Structure mechanical function relations. *Annu. Rev. Mater. Sci.* 28:271–298.
10. Screen, H. R. C., D. L. Bader, D. A. Lee, and J. C. Shelton. 2004. Local strain measurement within tendon. *Strain*. 40:157–163.
11. Lorenzo, A., and E. Caffarena. 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.
12. Persikov, A., J. Ramshaw, A. Kirkpatrick, and B. Brodsky. 2005. Electrostatic interactions involving lysine make major contributions to collagen triple-helix stability. *Biochemistry*. 44:1414–1422.
13. Mooney, S., and T. Klein. 2002. Structural models of *Osteogenesis imperfecta*-associated variants in the COL1A1 gene. *Mol. Cell. Proteomics*. 1:868–875.
14. Mooney, S., P. Kollman, and T. Klein. 2002. Conformational preferences of substituted prolines in the collagen triple helix. *Biopolymers*. 64:63–71.

15. Israelowitz, M., S. Rizvi, J. Kramer, and H. von Schroeder. 2005. Computational modeling of type I collagen fibers to determine the extracellular matrix structure of connective tissues. *Protein Eng. Des. Sel.* 18:645–652.
16. Boedke, H., and P. Doty. 1956. The native and denatured states of soluble collagen. *J. Am. Chem. Soc.* 78:4267–4280.
17. Oebrink, B. 1972. Non-aggregated tropocollagen at physiological pH and ionic strength. a chemical and physio-chemical characterization of tropocollagen isolated from the skin of lathyritic rats. *Eur. J. Biochem.* 25:563–572.
18. Utiyama, H., K. Sakato, K. Ikehara, T. Setsuiye, and M. Kurata. 1973. Flexibility of tropocollagen from sedimentation and viscosity. *Biopolymers.* 12:53–64.
19. Nestler, F. H. M., S. Hvidt, J. D. Ferry, and A. Veis. 1983. Flexibility of collagen determined from dilute-solution viscoelastic measurements. *Biopolymers.* 22:1747–1758.
20. Saito, T., N. Iso, H. Mizuno, N. Onda, H. Yamato, and H. Odashima. 1982. Semi-flexibility of collagens in solution. *Biopolymers.* 21:715–728.
21. Hofmann, H., T. Voss, K. Kuhn, and J. Engel. 1984. Localization of flexible sites in thread-like molecules from electron-micrographs—comparison of interstitial, basement-membrane and intima collagens. *J. Mol. Biol.* 172:325–343.
22. Sun, Y. L., Z. P. Luo, and K. N. An. 2001. Stretching short biopolymers using optical tweezers. *Biochem. Biophys. Res. Commun.* 286: 826–830.
23. 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. Commun.* 295:382–386.
24. Sun, Y. L., Z. P. Luo, A. Fertala, and K. N. An. 2004. Stretching type II collagen with optical tweezers. *J. Biomech.* 37:1665–1669.
25. Buehler, M. 2006. Atomistic and continuum modeling of mechanical properties of collagen: elasticity, fracture and self-assembly. *J. Mater. Res.* 21:1947–1961.
26. Buehler, M. 2006. Nature designs tough collagen: explaining the nanostructure of collagen fibrils. *Proc. Natl. Acad. Sci. USA.* 103: 12285–12290.
27. MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
28. Anderson, D. 2005. Collagen self-assembly: a complementary experimental and theoretical perspective. PhD thesis. University of Toronto, Toronto, Ontario.
29. Nelson, M. T., W. Humphrey, A. Gursoy, A. Dalke, L. V. Kale, R. D. Skeel, and K. Schulten. 1996. NAMD: a parallel, object oriented molecular dynamics program. *Intl. J. Supercomputer Appl. High Perform. Comput.* 10:251–268.
30. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
31. Kramer, R. Z., M. G. Venugopal, J. Bella, P. Mayville, B. Brodsky, and H. M. Berman. 2000. Staggered molecular packing in crystals of a collagen-like peptide with a single charged pair. *J. Mol. Biol.* 301: 1191–1205.
32. Feig, M., and C. Brooks. 2004. Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Biochem. Biophys. Res. Commun.* 14:217–224.
33. Lu, H., B. Israelowitz, A. Krammer, V. Vogel, and K. Schulten. 1998. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.* 75:662–671.
34. Buehler, M., J. Dodson, A. v. Duin, and W. Goddard. 2006. The computational materials design facility (CMDf): a powerful framework for multiparadigm multi-scale simulations. *Mat. Res. Soc. Proc. (Combinat. Methods Inform. Mat. Sci.)*. 894:LL3.8.
35. Plimpton, S. 1995. Fast parallel algorithms for short-range molecular-dynamics. *J. Comput. Phys.* 117:1–19.
36. Eppell, S. J., B. N. Smith, H. Kahn, and R. Ballarini. 2006. Nano measurements with micro-devices: mechanical properties of hydrated collagen fibrils. *J. Roy. Soc. Interface.* 3:117–121.
37. Bustamante, C., J. F. Marko, E. D. Siggia, and S. Smith. 1994. Entropic elasticity of lambda-phage DNA. *Science.* 265:1599–1600.
38. Marko, J. F., and E. D. Siggia. 1995. Stretching DNA. *Macromolecules.* 28:8759–8770.
39. Miles, C. A., and A. J. Bailey. 2001. Thermally labile domains in the collagen molecule. *Micron.* 32:325–332.