

Large-scale hierarchical molecular modeling of nano-structured biological materials

Markus J. Buehler

*Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
77 Massachusetts Ave. Room 1-272
Cambridge, MA, 02139, USA*

Email: mbuehler@MIT.EDU

Phone: 617-452-2750

Fax: 617-258-6775

Abstract:

The atomistic and molecular mechanisms that occur during mechanical deformation of natural and biological materials remain largely unknown. In recent years, development of new quantitative experimental, analytical and computational methods have led to advances in understanding of some details of these deformation mechanisms. Particular progress has been made in how to relate the molecular-scale chemistry to mesoscopic and macroscopic material properties. Here we review large-scale atomistic and molecular modeling methods to investigate the mechanical properties of natural and biological materials with nanostructured hierarchical designs. We discuss basic concepts of hierarchical multi-scale modeling capable of providing a bottom-up description of chemically complex materials. We compare the deformation mechanisms of biological materials with crystalline materials such as metals or ceramics. We emphasize on the importance of entropic contributions to elasticity, and the interplay of chemical bonding of different strengths, at different length- and time scales. We exemplify some of the techniques in studies of the mechanics of polypeptides, tropocollagen molecules and collagen fibrils.

Keywords:

Mechanics, protein, collagen, molecule, elasticity, fracture, atomistic modeling, self-assembly

This article appeared as:

Buehler, M. J. Large-scale hierarchical molecular modeling of nano-structured biological materials. *Journal of Computational and Theoretical Nanoscience* Vol. 3, pp. 603-623 (2006).

Contents

1. Introduction

- 1.1 Cracking the secrets of Nature's material design concepts
- 1.2 Outline

2. Classes of materials, modeling challenges and cross-scale interactions

- 2.1 Ductile versus brittle materials behavior: Crystalline solids
- 2.2 Small-scale materials: Thin metal films, nanowires and nanocrystalline solids
- 2.3 Biological materials: Nanostructured, heterogeneous hierarchies
- 2.4 From nano to macro: Interactions across the scales

3. Physical foundations of elasticity: Chemistry and mechanics

- 3.1 Entropic elasticity
- 3.2 Energetic elasticity

4. Atomistic and molecular modeling: Algorithms and methods

- 4.1 Basic classical molecular dynamics
- 4.2 Massively parallelized molecular modeling
- 4.3 Software integration and multi-paradigm modeling: Multi-paradigm modeling

5. Interatomic potentials and force fields

- 5.1 Nonreactive force fields for organic systems
- 5.2 Reactive force fields
- 5.3 Example: Application of reactive force fields to describe unfolding of a small protein

6. Application: Mechanics of tropocollagen molecules and collagen fibrils

- 6.1 The mechanics of collagen
- 6.2 Atomistic and mesoscale modeling of collagen
- 6.3 Full atomistic calculations
- 6.4 Reactive mesoscale model development
- 6.5 Mesoscale parameter training
- 6.6 Deformation of ultra-long tropocollagen molecules: Mechanics and entropy, fracture
- 6.7 Mechanics of collagen fibrils

7. Discussion and conclusion

8. Acknowledgements

References

1. Introduction

When materials are deformed, they display a small regime in which deformation is reversible or elastic ¹. Once the forces on the material are increased, deformation becomes irreversible, and the deformation remains after the stress is removed. This is referred to as the plastic regime ¹. Plastic deformation is typically followed by fracture. Many materials, ranging from metals, ceramics, amorphous materials, and also biological materials can be characterized in this way ².

Currently, the understanding of how materials deform and break is often limited to phenomenological engineering theories that describe the macroscale materials behavior, neglecting the underlying atomistic microstructure ¹. However, deformation and fracture of materials is controlled by atom-by-atom processes governed by quantum mechanics ³. These quantum mechanical effects that control chemical bonds are neglected in most existing theories. To include these effects, atomistic models are developed that are capable to simulate the motion of all atoms in the material, with systems comprising up to several billion particles, thus reaching macroscopic scales of material behavior that can be directly observed in experiment ⁴⁻⁹. The goal of such atomistic models is to understand the macroscopic response of materials, for example to mechanical stimulation, based on their fundamental, atomistic ultrastructure.

Modern multi-scale modeling techniques use a sequence of overlapping hierarchies encompassing various simulation tools to bridge the scales from nano to macro (see Figure 1). These modeling techniques allow a rigorous linking of material properties from quantum mechanics to mesoscale and macroscale. This fundamental viewpoint could revolutionize the engineering approach to use and create materials, by incorporating the atomistic scale into materials analysis and synthesis.

Further, understanding how materials break enables to decode the blueprint and building blocks of materials. This article is focused on large-scale atomistic and molecular modeling techniques capable of describing deformation and fracture mechanisms, with a particular focus on materials that show high chemical complexity and multiple hierarchies of structures across many length scales.

1.1 Cracking the secrets of Nature's material design concepts

Theoretical, numerical and experimental methods now enable to investigate the nanoscale mechanics of materials using quantitative analysis techniques, an area referred to as “nanomechanics”. For example, development and application of nanoindentation, AFM, and other tools enables scientists to probe the origins of elastic and plastic deformation, with forces in the range of pN to μ N, and at scales approaching that of individual atoms and molecules (e.g., see ^{10, 11}). At the same time, computational methods, computational power and theoretical approaches have led to significant advances in addressing nanomechanics from a first principles viewpoint (see Figure 1). This combination of experiment, theory and computation has proven to be very fruitful, and could lead to major advances in materials theories in the coming decades.

Computational methods can help to crack the secrets of how Nature designs materials with unique properties such as extreme toughness, hardness or compliance, as well as to shed light on how diseases inducing minute changes of the molecular structure of proteins can make such big impact on the behavior at larger scales. Development of physics based atomistic models of biological materials will have significant impact. It may change the way biologists and medical doctors think of materials, help to cure diseases, and may change how engineers design novel materials, for example by using recombinant DNA technologies.

The combination of high-level structural control of matter as achieved in nanoscience and nanotechnology, and integration of living and non-living systems into technologies and their interfaces may lead to the development of new technologies that utilize the advantages of both micro- and nanotechnology with the principles of biology¹². With increasing complexity, materials start to resemble systems or machines, so that the borderlines between concepts such as “machine” and “material” start to disappear. These concepts have been used systematically by Nature for millions of years. However, their systematic exploitation for technological applications has so far been severely hindered due to lack of understanding of how to link the chemistry scale with material and device properties and function.

Among the many building blocks of biological materials, the structure and behavior of proteins, as well as the materials that are based on proteins, stand out. In biology, these materials play an overarching role in determining the function and properties. New technologies could be based on such principles. Fundamental research that develops the links between chemistry, biology, physics and material properties could lay the foundation for these developments. Computational methods as those reviewed here may be a starting point towards realization of this vision.

1.2 Outline

In the following sections, we review basic concepts of a computational approach to address the materials science of biological materials, by linking structure, function and property.

This review article is particularly focused on:

- Outlining challenges and opportunities in the field of atomistic and molecular modeling of hierarchical materials,
- Summarizing hierarchical multi-scale modeling strategies at the atomistic and molecular level, suitable for chemically complex materials, in particular suitable models for atomic interactions,
- Reviewing an application of such a multi-scale modeling approach to tropocollagen molecules and collagen fibrils, making links between molecular chemistry and mechanical properties.

We proceed as follows. In Section 2, we review basic concepts of mechanics of materials, emphasizing on classical “bulk” materials as well as on nanoscale materials such as nanowires. We also introduce the particular complexities of biological materials, and review recent progress in the analysis of such materials, at different scales. Section 3 is dedicated to a discussion of the atomistic and molecular origins of elasticity, introducing the competition between entropic and energetic effects, which is crucial for many biological materials. We provide two hands-on examples of how to link molecular properties to the effective elastic response. Section 4 introduces basic molecular dynamics concepts, including a brief review on parallel computation and software architectures. In Section 5, we discuss strategies to model the atomic interactions, with a particular emphasis on those models that enable to treat the multitude of chemical interactions present in biological materials. We discuss a recent study of unfolding of a small protein that illustrates the interaction of chemical interactions of different magnitude in biological systems. Section 6 is focused on a hierarchical approach to describe the mechanics of collagen, including first principles atomistic, molecular and reactive mesoscale modeling of individual tropocollagen molecules and entire collagen fibrils. We conclude with a discussion in Section 7.

2. Classes of materials, modeling challenges and cross-scale interactions

Material behavior is typically characterized into brittle and ductile, as depicted schematically in Figure 2. With the advent of new nanoscale materials, the field of small-scale materials or interfacial materials have gained importance.

2.1 Ductile versus brittle materials behavior: Crystalline solids

The mechanics of brittle and ductile materials is relatively well understood, and many molecular-scale models have been proposed over the last decade. Results of large-scale molecular dynamics simulations of brittle and ductile materials are shown in Figure 3¹³⁻¹⁸. For crystalline materials, the understanding of plastic deformation and fracture behavior has advanced significantly¹⁹.

In the brittle case, atomic bonds are broken as material separates along a crack front. The type of failure of such materials is often characterized by the simultaneous motion of thousands of small cracks; as observed when glass shatters¹.

This type of failure usually happens rapidly, as following a large impacts cracks propagate at velocities close the speed of sound. An enormous amount of research has been carried out over the last hundred years or so and has been summarized in recent books. The origin of fracture research dates back to the early 20th century in studies by Inglis¹, Griffith²⁰ and Irwin²¹. The Griffith criterion provides a quantitative estimate of the condition under which material fails, and it is based on simple energetic and thermodynamic arguments. The Griffith criterion states that materials fail when the mechanical elastic energy released by crack propagation (denoted by G) equals the fracture surface energy (denoted by γ_s):

$$G = 2\gamma_s. \quad (1)$$

This thermodynamic view of fracture was the foundation for the field of fracture mechanics.

The continuum mechanical theory of fracture is a relatively well-established framework²². In this theory, the stress field in the vicinity of the crack tip is given by the asymptotic solution and exhibits a universal character independent of the details of the applied loading^{1, 23}. Further details can be found in the literature^{1, 23, 24}.

The response of ductile materials to severe loading can be analyzed with similar methods, involving energy parameters such as the unstable stacking fault energy (denoted by γ_{us}). This parameter describes how difficult it is to generate dislocations at the tip of a crack. The energetic competition between dislocation nucleation and creation of surfaces determines whether a material is ductile or brittle.

2.2 Small-scale materials: Thin metal films, nanowires and nanocrystalline solids

Recent research results indicate that dislocation-based processes and cleavage are not the only mechanisms for deformation of materials. Materials under geometrical confinement are also referred to as materials in small dimensions. The behavior of these materials is characterized by the interplay of interfaces (for example, grain boundaries), constraints (for example, substrates), and free surfaces²⁵⁻³². Examples for such materials are nanocrystalline materials³³⁻⁴⁴ or ultra thin submicron films⁴⁵⁻⁵².

It was shown by computer simulation that in such materials, with grain sizes of tens of nanometers and below, deformation could be largely dominated by grain boundary processes, including diffusive mechanisms. Even though such material behavior is ductile, as these materials can be bent without cleavage, no dislocation motion is required, in contrast to what is known from larger-scale microcrystalline materials. Because of the small sizes of the grains, dislocations cannot be generated, because, for instance, Frank–Read sources are too large to fit within a grain, or because dislocations are energetically very expensive under very small geometrical confinement, as shown in earlier work⁵³. This leads to unexpected mechanisms of deformation, such as motion of partial dislocations^{36, 37}.

The study of the mechanical properties nanowires, both experimentally and with molecular dynamics modeling, has been rather fruitful and led to significant insight about the behavior of such one-dimensional materials⁵⁴⁻⁶³. It is an excellent example for an area where classical continuum mechanical concepts cannot be used any more, and where molecular dynamics modeling is rather readily applicable⁶³. Such nanowires, ultra small high aspect ratio wire-like structures, have been synthesized out of a variety of distinct materials, including metals, ceramics, and even biological materials^{56, 58, 64, 65}.

2.3 Biological materials: Nanostructured, heterogeneous hierarchies

Whereas metals and ceramics often show a regular structure of polycrystalline nature, many biological materials show a much more complex hierarchical and rather heterogeneous structure.

Sometimes, as for example in the case of elastin⁶⁶, the microstructure at the molecular or atomistic level is not completely known – as it consists of random domains that have – thus far – largely thwarted any attempt to decipher their structure. This also applies to some natural composite materials such as cement^{67, 68}, for which not even the basic building blocks or their structure is known. Many of these natural materials also belong to the class of chemically complex materials, and share critical features with many biological composites. Many of the modeling techniques discussed here may also be applicable to these materials.

In this case, the deformation behavior is much more complicated. Also, due to the heterogeneous and hierarchical design, deformation mechanics may be different at different scales. For example, bone is composed of two phases: Single crystal platelets of hydroxyapatite, and a matrix of collagen⁶⁹⁻⁷⁴. Whereas hydroxyapatite itself is extremely fragile (as classroom chalk), collagen is extremely compliant. However, the combination of these two phases into a nanostructural composite makes a material that is tough⁶⁹⁻⁷⁴.

Thus, in order to understand the mechanics of hierarchical materials, it is critical to investigate the responses at different scales, including a treatment of the individual phases and the interaction of different phases. Figure 4(a) exemplifies such a complex materials subject to a macroscopic crack. Figure 4(b) illustrates schematically the different interactions that need to be considered.

Protein based materials hold great potential as biomaterials or new synthetic polymers with unique mechanical, optical and electrical properties. In recent years, proteins have also been proposed as basis of new materials for technological applications^{75, 76}. Genetically engineered biopolymers based on recombinant DNA technologies have been developed by Tirrell and coworkers⁷⁵⁻⁷⁹ for various applications, including pH sensitive hydrogels⁷⁶. This is an example for the vanishing borderline between technology and biology, enabling new critical breakthroughs for novel material concepts, allowing to translate nature's structural concepts into engineered materials⁷⁵⁻⁸⁰.

Much effort has recently gone into understanding the nano- and micromechanics of living organisms, as for example cells^{10, 11, 81, 82}. Earlier work focused on simple network models to describe the cellular filaments (rods, ropes)⁸³⁻⁸⁶. More recently, Suresh and coworkers analyzed the mechanics of red blood cells within the context of malaria, and discovered that the disease has a dramatic effect of making the cell much stiffer, leading to spreading of the disease⁸⁷.

The quantitative, precise analysis revealed that this stiffening effect due to Malaria is much more severe than previous assumed, leading to stiffening up to ten times compared

to the original stiffness of the cell. The analysis of cells also included molecular modeling at the scale of spectrin networks⁸⁸.

2.4 From nano to macro: Interactions across the scales

Modifications of ultra-small or nanoscale features in materials can lead to significant changes in the macroscopic behavior. Similar mechanisms have been known for a long time in the physics or mathematics communities, who have been studying the dynamical properties of nonlinear systems. In fact, many optical or magnetic properties are based on such microscopic collective mechanisms, leading to dramatic changes at the macroscopic scale.

For mechanics of materials, such concepts have thus far not systematically been exploited, and are often poorly understood. One reason is that in the past, structural changes at the nanoscale have not been possible, which made it difficult to confirm theoretical explanations and concepts that link nanoscale and macroscale quantitatively. However, such atomic-scale manipulation of matter is now reality.

For biological materials, and in particular those based on proteins, genetic engineering has become an important tool to address precisely these questions. These techniques enable scientists and engineers to change specific genes, leading to mutations in the sequence of proteins, at the scale of several Angstrom. This has significant impact in understanding diseases, which often origin from such mutations. On the other hand, it also allows to create new materials. Bacteria can be used to produce new biological materials, derived from known structures – modified to feature improved properties. This particular case is another intriguing example for a systematic microscopic changes – which can actually be performed, with extremely high accuracy – and whose result can be measured. Understanding such interactions across the scales remains elusive. Yet, it is the key to decipher the origins of diseases, mechanisms of cellular and tissue processes, and the development of new biologically inspired materials⁸⁹⁻⁹⁴.

What are theoretical and computational frameworks to describe the deformation and failure of such complex materials? Historically, the continuum-based theories have been the basis for most theoretical and computational tools of engineers. Such approaches were extremely successful, for example to describe plasticity of crystal materials. In early stages of computational plasticity, dislocations and cracks were often treated using linear continuum mechanics theory, relying on numerous phenomenological assumptions. Over the last decades, there has been a new realization that understanding nanoscale behavior is required for understanding how materials fail. This is partly because of the increasing trend to miniaturization as relevant length scales of materials approach several nanometers in modern technology. Once the dimensions of materials reach submicron length scales, the continuum description of materials is questionable and the full atomistic information is necessary to study materials phenomena⁹⁵.

The applicability of continuum theories to biological materials remains subject to discussion. Certainly, at a rather macroscopic scale these methods can still be used.

However, recent research indicates that the governing processes often occur at the nanoscale, where properties change rapidly and often irregularly across variations in space – thus the basic assumptions of continuum theories are violated. This particularly applies to natural and biological materials, which often have structural features across many length scales, with smallest features at the nanoscale. The fact that these materials are used at dimensions comparable to their structural heterogeneity – for example, enzymes operate on substrates with sizes comparable to their diameter – complicates the situation and certainly questions the applicability of continuum-based methods.

3. Physical foundations of elasticity: Chemistry and mechanics

Elasticity stems from the interactions of atoms, and thus it is intimately linked to chemistry. Depending on what material is considered, these chemical interactions are more or less complex. In metals, for example, bonding is primarily non-directional, and can be characterized by positive ions embedded in a gas of electrons. Other materials show much greater chemical complexity – often featuring many different chemical bonds with varying strength. In biological materials, it is vital to consider various types of chemical interactions, including:

- Ionic interactions (columbic interactions),
- Covalent bonds (due to overlap of electron orbitals),
- Hydrogen-bonds, and
- Weak or dispersive van der Waals (vdW) interactions.

In general, the elastic properties of materials can be expressed as the partial derivative of the free energy density with respect to the strain tensor that characterizes the deformation. Based on the free energy

$$A = U - TS \tag{2}$$

we can define the free energy density,

$$\Phi = \frac{A}{V}. \tag{3}$$

The (scalar) stress is then is given by

$$\sigma = \frac{\partial \Phi}{\partial \varepsilon}, \tag{4}$$

and the elastic modulus can be expressed as

$$E = \frac{\partial^2 \Phi}{\partial \varepsilon^2}. \tag{5}$$

Note that stress and strain are related by Hooke's law, $\sigma = E\varepsilon$.

Whereas the elasticity of crystalline materials is primarily controlled by energetic changes of the internal energy, in natural and biological materials, elasticity is also significantly influenced by entropic contributions.

This is because in many crystalline materials, the entropic term can be neglected, so that Φ in eqs. (3) and (4) can be directly substituted by U , the internal energy. However, in biopolymers, entropic contributions can dominate the elasticity, in particular for small deformation, and therefore Φ in eqs. (3) and (4) can be substituted by TS . Theoretical concepts such as the Neo-Hookean constitutive equation are derived based directly from these concepts.

In the following two sections, we review strategies how to link molecular or atomistic properties with overall elastic moduli, for two extreme cases: Pure entropic elasticity ($dU/d\varepsilon \approx 0$), and pure energetic elasticity ($d(TS)/d\varepsilon \approx 0$).

3.1 Entropic elasticity

Dominance of entropic behavior is a well-known phenomenon from many polymers. The contributions to the entropic term to elasticity can be described in several ways, including classical descriptions such as the worm-like chain model or the freely jointed Gaussian chain model. Such descriptions are similar to constitutive models in continuum elasticity, and require input parameters that are typically determined empirically. In contrast to these models, molecular dynamics modeling can provide a first principles based description of entropic elasticity – without any additional fitting parameters beyond the atomic interactions.

The persistence length is defined as the molecular length at which entropic contributions to elasticity become important, as the molecule shows significant bending purely due to its thermal energy. A molecule with length far beyond the persistence length will bend, even without application of forces, and assume a conglomerated, wiggly shape. With the bending stiffness of a molecule denoted as EI , the persistence length is defined as

$$\xi_p = \frac{EI}{k_B T}. \quad (6)$$

When the length of molecules, denoted by L , is beyond the persistence length, that is, $L \gg \xi_p$, thermal energy can bend the molecule, and entropic elasticity typically plays a role. On the other hand, when $L \ll \xi_p$, entropic effects play a minor role, and energetic elasticity governs.

A schematic of a convoluted molecule is shown in Figure 5(a). A schematic response of a molecule due to stretching is shown in Figure 5(b). Entropic effects become important and appear in measurements, for example when one stretches a convoluted molecule⁹⁶.

Assuming that the initial point-to point distance is $x < L$ (expressing the fact that the molecule is convoluted) the force that resists stretching is given by

$$F = \frac{kT}{\xi_p} \left(\frac{1}{4} \frac{1}{(1-x/L)^2} - \frac{1}{4} + x/L \right). \quad (7)$$

This model is called the Worm-Like-Chain Model (WLC) or Marko–Siggia equation. The molecular properties enter this equation in form of the persistence length, which is a function of the bending stiffness. If these properties are known from atomistic calculations, the WLC model provides a quantitative estimate of entropic elasticity.

Similar considerations have been used to understand the macroscopic elastic response of rubber – a class of material whose elasticity is largely dominated by entropic effects. Figure 6 depicts a sample unit cell of an entropic polymer that undergoes deformation, characterized by the extension ratios $\lambda_i = r_{i,new} / r_i$ (see Figure 6).

We briefly review how to link the molecular properties to the overall elasticity. Considering a freely jointed Gaussian chain with n links of length l each, the entropy of such a chain is

$$S = c - kb^2 r^2, \quad (8)$$

where

$$b^2 = \frac{3}{2nl^2}. \quad (9)$$

The change in entropy due to deformation is given by

$$\Delta S = -kb^2 \sum_{N_b} (\lambda_1^2 - 1)x^2 + (\lambda_2^2 - 1)y^2 + (\lambda_3^2 - 1)z^2. \quad (10)$$

By averaging over all chains in the system,

$$\Delta S = -kb^2 N_b [(\lambda_1^2 - 1)\langle x^2 \rangle + (\lambda_2^2 - 1)\langle y^2 \rangle + (\lambda_3^2 - 1)\langle z^2 \rangle]. \quad (11)$$

Assuming an isotropic solid, i.e. the end-to-end distances of the N_b chains are directed equally in all directions, we find that

$$\langle x^2 \rangle = \langle y^2 \rangle = \langle z^2 \rangle = \frac{1}{3} \langle r_b^2 \rangle. \quad (12)$$

Since

$$r_{RMS} = \sqrt{n} \cdot l = \sqrt{\langle r_b^2 \rangle}, \quad (13)$$

and $\langle r_b^2 \rangle = n \cdot l^2$, we arrive at

$$\langle x^2 \rangle = \langle y^2 \rangle = \langle z^2 \rangle = \frac{1}{3} n \cdot l^2 = \frac{1}{2b^2}, \quad (14)$$

and finally at

$$\Delta S = -kN_b / 2 [(\lambda_1^2 - 1) + (\lambda_2^2 - 1) + (\lambda_3^2 - 1)]. \quad (15)$$

The change in free energy is then given by

$$A = -T\Delta S = \frac{1}{2} NkT (\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3). \quad (16)$$

The free energy density is given by

$$\Phi = \frac{A}{V}. \quad (17)$$

This expression of the free energy density function can be used to derive expression for Young's modulus (using eqs. (4) and (5)),

$$E = 3N^* kT, \quad (18)$$

linking uniaxial stress σ and uniaxial extension ratio λ through $\sigma = (E/3)(\lambda^2 - 1/\lambda)$.

This provides a link of microscopic parameters ($N^* = N/V$, the density of molecular chains per unit volume) with a macroscopic measurable, E . From this equation, we also note that the stiffness is proportional to the temperature, $E \sim T$.

Despite the significant progresses made in advancing the understanding of brittle or ductile materials, the atomistic and molecular mechanisms of deformation of natural and biological materials in which entropy governs are often poorly understood. The study of elastic and plastic deformation of such materials using large-scale computer simulations poses great challenges and thus great opportunities.

3.2 Energetic elasticity

Energetic elasticity is characterized by the stretching of atomic bonds, metallic, covalent or ionic bonds, that lead to a change in potential energy in the material volume. Many techniques have been proposed to link micro-structural deformation with macroscopic concepts of elasticity.

A prominent example is the Cauchy-Born (CB) rule, which relates microscopic deformation fields with macroscopic stresses and strains⁹⁷. The CB rule or different implementations is the basis for several multi-scale modeling algorithms, including the quasicontinuum method⁹⁸⁻¹⁰⁰ or the virtual internal bond method¹⁰¹⁻¹⁰⁴. A limitation of

these methods is that they are largely limited to zero temperature studies, and that the technique is only applicable far from defects or large stress and strain gradients.

Recently, the CB has been extended to capture entropic effects as well – thus overcoming some of the limitations of the classical CB rule, and making such multi-scale techniques applicable to finite temperature studies^{100,105}.

We review a brief example how to calculate the effective elastic properties of a one-dimensional string of atoms, to demonstrate how the CB rule can be used to make a rigorous link between interatomic potential and continuum elasticity. This is schematically visualized in Figure 7.

The central assumption of the Cauchy-Born rule is to express the energy of an atomic system as a function of an applied strain tensor ε_{ij} , and it is assumed that the continuum fields can immediately be mapped to the atomic scale. The goal is to express the strain energy density (SED) of an atomic unit cell as a function of the macroscopic strain field. As indicated above, first derivatives of the SED with respect to the strain yields the stress, the second derivatives of the SED with respect to the strain give the modulus.

In crystal lattices with pair-wise interactions, the sum over all bonds in a unit cell yields the SED:

$$\Phi(\varepsilon) = \frac{1}{\Omega} \sum_i \phi_i(r_i). \quad (19)$$

As a first example, we consider a 1D string of atoms, corresponding to a biological molecule. In its undeformed configuration, the distance between pairs of atoms is r_0 . This 1D crystal is defined by a unit cell that contains one bond and one atom. The volume of the unit cell is $\Omega = r_0 \times A_c$ where A_c is the cross-sectional unit area of the molecule, orthogonal to the axial direction. Assuming existence of a function describing the relation of bond energy and bond length, the SED is given by

$$\Phi(\varepsilon) = \frac{1}{r_0 \cdot A_c} \phi(r). \quad (20)$$

The function $\phi(r)$ could, for example, be a Lennard-Jones (LJ) 12:6 potential that is defined as

$$\phi(r) = 4\varepsilon^* \left[\left(\frac{\sigma^*}{r} \right)^{12} - \left(\frac{\sigma^*}{r} \right)^6 \right], \quad (21)$$

where ε^* and σ^* are the LJ potential parameters. The bond length r in the deformed state is given by $r = (1 + \varepsilon)r_0$. Therefore,

$$\Phi(\varepsilon) = \frac{1}{r_0 \cdot A_c} \phi(r) = \frac{1}{r_0 \cdot A_c} \phi((1 + \varepsilon) \cdot r_0). \quad (22)$$

Since we are interested in obtaining elastic coefficients within the concept of linear elasticity, the second derivatives of $\Phi(\varepsilon)$ with respect to ε should be constant. Thus we consider a second order Taylor expansion of the interatomic potential around the equilibrium bond length r_0 :

$$\phi(r) = a_0 + a_1(r - r_0) + a_2/2(r - r_0)^2 + \dots, \quad (23)$$

Since the force versus bond stretch is zero at the equilibrium bond length, $a_1 = \phi'(r_0) = 0$. The parameter $a_2 = \phi''(r_0)$ corresponds to the force-bond distance slope around r_0 and is now denoted as k . The parameter a_0 is a constant, which is set to zero since we are only interested in the derivatives of $\Phi(\varepsilon)$. We approximate

$$\phi(r) \approx k/2(r - r_0)^2. \quad (24)$$

We can then rewrite eq. (5) as

$$\Phi(\varepsilon) = \frac{1}{r_0 \cdot A_c} \frac{1}{2} k ((1 + \varepsilon - 1)r_0)^2 = \frac{r_0}{2A_c} k \varepsilon^2. \quad (25)$$

The first derivative of $\Phi(\varepsilon)$ yields the stress for a given strain ε ,

$$\frac{\partial \Phi(\varepsilon)}{\partial \varepsilon} = \sigma = \frac{r_0}{A_c} k \varepsilon. \quad (26)$$

The second derivative of yields the modulus,

$$\frac{\partial^2 \Phi(\varepsilon)}{\partial \varepsilon^2} = E = \frac{r_0}{A_c} k. \quad (27)$$

Eqns. (9) and (10) provide relationships between stress and strain and define Young's modulus for this system. This expression can be used to calculate other material properties, such as the wave speed. For the 1D system, the (only) wave speed is given by

$$c = \sqrt{\frac{E}{\rho}}, \quad (28)$$

where ρ is the atomic density ($\rho = m/r_0$ with m denoting the atomic mass). These relationships have been established by only considering (i) the atomic microstructure, (ii) the interatomic potential, and (iii) by assuming that the Cauchy-Born rule can be applied (*i.e.* that we can map the continuum displacement field to the atomic lattice). No other phenomenological assumptions have been made.

We note that it is possible that condition (iii) is not satisfied, for example when dealing with molecular crystals, including crystals of proteins. This is because in such systems, atomic displacements within the unit cell do not necessarily correspond to the continuum displacement field. Other systems where this is relevant are amorphous materials where the Cauchy-Born rule can only be applied in a statistical sense.

4. Atomistic and molecular modeling: Algorithms and methods

Atomistic simulations have proved to be a unique and powerful way to investigate the complex behavior of dislocations, cracks, and grain boundary processes at a very fundamental level. Atomistic methods are often the core in modern materials modeling. One of the strengths and the reason for the great success of atomistic methods is its very fundamental viewpoint of materials phenomena. The only physical law that is put into the simulations is Newton's law and a definition of how atoms interact with each other. Despite this very simple basis, very complex phenomena can be simulated. Unlike many continuum mechanics approaches, atomistic techniques require no *a priori* assumption on the defect dynamics.

Drawbacks of atomistic simulations are the difficulty of analyzing results and the large computational resources necessary to perform the simulations. This becomes more evident as the simulation sizes increase to systems with billions of atoms. One of the largest simulations reported to date treated almost 20 billion particles⁸.

The atomistic or molecular approach brings great promises, but also holds many big challenges. Once the atomic interactions are chosen, the complete material behavior is determined. Although in some cases it is difficult to find the correct potential for a specific material, atomic interactions can often be chosen such that generic properties common to a large class of materials are incorporated (e.g., ductile materials). This allows us to design "model materials" to study specific materials phenomena. Despite the fact that model building has been in practice in fluid mechanics for many years, the concept of "model materials" in materials science is relatively new.

However, atomic interactions can be calculated very accurately for a specific atomic interaction, using quantum mechanics methods such as the density functional theory³. Also, the development of better interatomic potentials – such as reactive force fields^{13, 106-109} capable of describing chemistry in systems comprising of several million particles – major breakthroughs are expected in the years to come.

Recently, an increasing number of researchers have grown to consider the computer as a tool to do science, engineering – including the design of new materials, new structures or

development of drugs, without ever synthesizing them. Along such lines of thought, computer simulations have been sometimes referred to as “computational experiments.” Designing smart computer experiments is the key to a successful simulation. Computational modeling allows a bottom-up approach in understanding biological materials, focusing on the finest, atomistic scales of detail governed by quantum mechanics (QM) as starting point, reaching up to large, macroscopic scales, using hierarchical multi-scale modeling. This could be viewed as an alternative, supplemental method with respect to many experimental efforts starting at the macroscopic level reaching down to ever finer scales.

4.1 Basic classical molecular dynamics

Classical molecular dynamics generates the trajectories of a large number of particles, interacting with a specific interatomic potential, leading to positions $r_i(t)$, velocities $v_i(t)$ and accelerations $a_i(t)$. It can be considered an alternative to methods like Monte-Carlo, with the difference that MD actually provides full dynamical information. – and deterministic trajectories.

The total energy of the system is written as the sum of kinetic energy (K) and potential energy (U),

$$E = K + U \quad (29)$$

where the kinetic energy is

$$K = \frac{1}{2} m \sum_{j=1}^N v_j^2, \quad (30)$$

and the potential energy is a function of the atomic coordinates r_j ,

$$U = U(r_j), \quad (31)$$

with a properly defined potential energy surface $U(r_j)$. The numerical problem to be solved is a system of coupled 2nd order nonlinear differential equations:

$$m \frac{d^2 r_j}{dt^2} = -\nabla_{r_j} U(r_j) \quad j = 1..N, \quad (32)$$

which can only be solved numerically for more than two particles, $N > 2$. Typically, MD is based on updating schemes that yield new positions from the old positions, velocities and the current accelerations of particles:

$$r_i(t_0 + \Delta t) = -r_i(t_0 - \Delta t) + 2r_i(t_0)\Delta t + a_i(t_0)(\Delta t)^2 + \dots \quad (33)$$

The forces and accelerations are related by $a_i = f_i / m$. The forces are obtained from the potential energy surface – sometimes also called force field – as

$$F = m \frac{d^2 r_j}{dt^2} = -\nabla_{r_j} U(r_j) \quad j = 1..N. \quad (34)$$

This technique cannot only be used for particles that are atoms; it can also be applied for particles that represent groups of atoms, such as in bead models (see Sections 6.4 and 6.5).

4.2 Massively parallelized molecular modeling

Large-scale molecular dynamics simulations often require a significant amount of computing resources. Classical molecular dynamics can be quite efficiently implemented on modern supercomputers using parallelized computing strategies. Such supercomputers are made out of hundreds of individual computers.

Whereas computing power was estimated to plateau at the gigaflop level, we now expect petaflop computers by the middle or end of the current decade¹¹⁰. Based on the concept of concurrent computing, modern parallel computers are made out of hundreds or thousands of small computers working simultaneously on different parts of the same problem. Information between these small computers is shared by communicating, which is achieved by message-passing procedures, enabled via software libraries such as the “Message Passing Interface” (MPI)¹¹¹.

Implemented based on spatial domain decomposition, parallel MD reaches linear scaling, that is the total execution time scales linear with the number of particles $\sim N$, and scales inversely proportional with the number of processors used to solve the numerical problem, $\sim 1/P$ (where P is the number of processors)⁸.

With a parallel computer whose number of processors increases with the number of cells (the number of particles per cell does not change), the computational burden remains constant. To achieve this, the computational space is divided up into cells such that in searching for neighbors interacting with a given particle, only the cell in which it is located and the next-nearest neighbors have to be considered. This scheme allows to treat huge systems with several billion particles⁸.

4.3 Software integration and multi-paradigm modeling: Multi-paradigm modeling

Predicting the properties and behavior of materials by computer simulation from a fundamental, *ab initio* perspective has long been a vision of computational material scientists. The key to achieving this goal is utilizing hierarchies of paradigms and scales that connect macro-systems to first principles quantum mechanics (QM). Such approaches have been suggested by various groups over the last years.

The need for new simulation methods that enable integration of disparate codes and paradigms has resulted in the development of software environments, including simulation environments such as “CAMPOS”¹¹², Pyre¹¹³ or the “Computational Materials Design Facility” (CMDf)¹¹⁴⁻¹¹⁷. They have a common goal make various standalone codes – representing distinct simulation paradigms – available to users, so that they can combine them arbitrarily, often within easy to learn scripting languages. This notion of “paradigm complexity” could help to build complex simulation algorithms whose capabilities emerge far beyond those of each individual code.

For example, the CMDf environment^{13, 114-116} is capable of simulations of complex materials studies using a variety of simulation paradigms, by using a Python scripting layer to integrate different computational tools to build up multi-scale simulation applications. A central data structure “Extended OpenBabel (XOB)” plays a critical role as the glue between individual applications. Separate computational engines are wrapped using automated wrapping tools (e.g. SWIG¹¹⁸). This strategy allows rapid integration of low-level C or FORTRAN codes with scripting languages. Scripting languages provide interfaces to compiled code, and rely significantly on external libraries for important subsystems.

5. Interatomic potentials and force fields

Definition of the atomic interactions by force fields is at the heart of MD methods, as it defines the complete materials behavior. The expressions for the potential energy landscape are used to determine the forces and accelerations used in a MD updating scheme.

Many flavors and types of force fields exist. Here we briefly review some very basic concepts of force fields for metals, and then describe some methods to model organic systems including proteins. The methods developed within the realm of empirical force fields are widely applicable, and are also the basis for mesoscopic descriptions of molecular behavior – treating collection of atoms as “bead” like objects, thus significantly reducing the computational cost, both in time and space.

Many force fields have been developed to describe crystalline systems, including metals and ceramics. These are typically separated into pair potentials and multi-body potentials. In pair potentials, the total potential energy of the system is expressed as the sum over all pairs of bonds. In multi-body potentials, the strength of the pair wise interactions are modulated dependent on the atomic environment. Examples for pair potentials are the Lennard-Jones potential¹¹⁹ and the Morse function¹²⁰. Examples for popular multi-body potentials are the Tersoff potential¹²¹ and the Stillinger-Weber¹²² potential to model covalent interactions, or the EAM and MEAM potentials suitable for metals¹²³⁻¹²⁷.

5.1 Nonreactive force fields for organic systems

A popular classical force field CHARMM¹²⁸, implemented in the MD program NAMD^{129, 130}. The CHARMM force field¹²⁸ is widely used in the protein and biophysics community, and provides a basic description of proteins. Other popular examples are the AMBER force field¹³¹ and the DREIDING¹³² or the all-atom UFF force field¹³³.

Force fields of this type are typically based on harmonic and anharmonic terms describing covalent interactions, in addition to long-range contributions describing van der Waals (vdW) interactions, ionic (Coulomb) interactions, as well as hydrogen bonding. Since the bonds between atoms are modeled by harmonic springs or its variations, bonds between atoms can not be broken, and new bonds can not be formed. Also, the charges are fixed and can not change, and the equilibrium angles do not change depending on stretch.

5.2 Reactive force fields

Reactive force fields represent a major milestone in overcoming the limitations classical force fields in not being able to describe chemical reactions. For mechanical properties, this translates into the properties of molecules at large-strain, a phenomenon referred to as hyperelasticity.

Several flavors of reactive potentials have been proposed in recent years^{106, 108, 134, 135}. Reactive potentials can overcome the limitations of empirical force fields and enable large-scale simulations of thousands of atoms with quantum mechanics accuracy. The reactive potentials, originally only developed for hydrocarbons^{106, 134, 135}, have been extended recently to cover a wide range of materials, including metals, semiconductors and organic chemistry in biological systems such as proteins^{108, 109, 136-142}. Here we use the ReaxFF formulation¹⁰⁶. We employ a particular flavor of the ReaxFF potentials as suggested in^{106, 140}, with slight modifications to include additional QM data suitable for protein modeling¹⁴³.

Reactive potentials are based on a more sophisticated formulation than most nonreactive potentials. A bond length/bond order relationship is used to obtain smooth transition from non-bonded to single, double, and triple bonded systems. All connectivity-dependent interactions (that means, valence and torsion angles) are formulated to be bond-order dependent. This ensures that their energy contributions disappear upon bond dissociation so that no energy discontinuities appear during reactions. The reactive potential also features non-bonded interactions (shielded van der Waals and shielded Coulomb).

The reactive formulation uses a geometry-dependent charge calculation scheme similar to QEq¹⁴⁴ that accounts for polarization effects and modeling of charge flow. This is a critical breakthrough leading to a new bridge between QM and empirical force fields. All interactions feature a finite cutoff.

In ReaxFF, the total energy of a system is expressed as the sum of different contributions that account for specific chemical interactions. The total energy is given by

$$E_{system} = E_{bond} + E_{vdWaals} + E_{Coulomb} + E_{val,angle} + E_{tors} + E_{over} + E_{under}. \quad (35)$$

The terms $E_{bond} + E_{vdWaals} + E_{Coulomb}$ are two-body contributions, the terms $E_{val,angle} + E_{tors}$ are 3-body and 4-body terms, and $E_{over} + E_{under}$ correspond to multi-body contributions due to the local chemical environment.

ReaxFF is based on the concept of bond orders, going back to early work by Abell, Tersoff, Brenner and others^{121, 134, 135, 145}. The basic concept of bond orders is simple to explain. The key idea is to modulate the bond strength based on the atomic environment, taking advantage of some theoretical chemistry principles.

Consider a pair potential in which the total energy of the system is given by the sum over all pairs of atoms (note the factor 1/2 to avoid double counting):

$$U_{total} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1, i \neq j}^N \varphi(r_{ij}). \quad (36)$$

Instead of writing $\varphi(r_{ij})$ as a LJ function (see above), in the Abell-Tersoff approach the interaction between two atoms is expressed as

$$\varphi(r_{ij}) = \varphi_R(r_{ij}) - M_{ij} \varphi_A(r_{ij}), \quad (37)$$

where $\varphi_R(r_{ij})$ and $\varphi_A(r_{ij})$ are pair repulsive (subscript ‘‘R’’) and attractive (subscript ‘‘A’’) interactions. The parameter M_{ij} that multiplies the attractive interactions represents a many-body interaction parameter. This parameter describes how strong the attraction is for a particular bond, from atom i to atom j . Most importantly, the parameter M_{ij} can range from zero to one, and describes how strong this particular bond is, depending on the environment of atom i . It can thus be considered a normalized bond order, following the concept of the Pauling relationship between bond length and bond order. Abell suggested that

$$M_{ij} \sim Z^{-\delta}, \quad (38)$$

where δ depends on the particular system, and Z is the coordination number of atom i .

Eqs. (37) and (38) immediately lead to a relationship between bond length, binding energy and coordination, through the parameter M_{ij} . This method has been very successful to describe the interatomic bonding in several materials, for example the C-C bonds in diamond, graphite and even hydrocarbon molecules. It is also the basis for the ReaxFF force field.

5.3 Example: Application of reactive force fields to describe unfolding of a small protein

Here we demonstrate the difference between the reactive potential and a classical CHARMM potential in a study of unfolding of a small protein PDB ID 1AKG.

Figure 8 shows the force-displacement curve during mechanical unfolding of a small protein, comparing the predictions made by the CHARMM force field and ReaxFF. Individual rupture events of covalent bonds can be associated with peaks in the force-displacement curve. Figure 9 shows a more detailed view of this system, focusing in on the small displacements and the first bond rupture event. Figure 10 depicts several snapshots of the atomic structure as the protein undergoes deformation.

It is apparent that classical CHARMM-like descriptions are not capable to address the various chemical events properly, and lead to incorrect behavior at large deformation.

6. Application: Mechanics of tropocollagen molecules and collagen fibrils

Here we review a multi-scale approach applied to model the mechanics of collagen fibrils^{146, 147}. The hierarchical nanostructure of collagen is shown schematically in Figure 11. The plot visualizes the different design hierarchies, ranging from the nano-scale of polypeptides to microscopic collagen fibers¹⁴⁸.

Each tropocollagen molecule consists of a spatial arrangement of three polypeptides. These three molecules or polypeptides are arranged in a helical structure, stabilized mainly by H-bonding between different residues. Every third residue in each of these molecules is a GLY amino acid, and about one fourth of the molecule consists of proline (PRO) and hydroxyproline (HYP). The structure of collagen has been known since classical works by Ramachandran in 1951, focusing on theoretical understanding of how tropocollagen molecules are stabilized¹⁴⁹. Recently, various types of tropocollagen molecules have been crystallized and analyzed using X-ray diffraction techniques to determine their exact atomic configuration¹⁵⁰. TEM experiment has also been used to study the structure of collagen in various environments, including in bone¹⁵¹, in particular focusing on larger length scale features and its three-dimensional arrangement.

Collagen is one of the most abundant protein-based structural materials, for which the atomic structure is relatively well-known^{148, 149, 152}. Together with such new experimental techniques for synthesis and characterization⁷⁵⁻⁸⁰, a sound theoretical understanding, and the availability of accurate models is essential for developing engineering applications of such complex materials. There are several reports of experimental studies focused on the mechanics of single tropocollagen fibers^{93, 96, 152-156}. However, despite its relatively simple structure¹⁵⁰, collagen molecules have rarely been studied using molecular dynamics (MD) studies. In one of the few reports we found in the literature, Lorenzo and coworkers¹⁵⁷ have recently reported investigations of the mechanical properties of collagen fibers, using MD studies, focusing on their Young's modulus. Other studies focused on the stability of collagen molecules¹⁵⁸ and other

structural investigations¹⁵⁹, or the effect of point mutations on the stability¹⁶⁰⁻¹⁶². Some researchers modeled collagen at the continuum scale, using techniques such as FEM^{159, 163}. In a recent study, molecular models of collagen have been linked to collagen mechanics¹⁶⁴.

6.1 The mechanics of collagen

The mechanical properties of collagen are determined from complex interaction of individual scale-specific properties and their interactions. Up to date it remains unclear how different properties interact, and which are most important, but it is believed that the hierarchical design plays a vital role. Fundamental fracture and deformation theories of collagen are missing, in particular those that include the effects of the hierarchical design.

6.2 Atomistic and mesoscale modeling of collagen

Our modeling approach includes studies of:

- Tropocollagen molecules (elasticity, fracture),
- Assemblies of tropocollagen molecules (shear deformation, assembly mechanics), and
- Collagen fibrils (elasticity, fracture, and response to large deformation).

The different scales associated with each structure become accessible through hierarchical multi-scale modeling, where the model parameter at each scale are rigorously informed from more accurate calculations. This method is following the idea of “finer scales train coarser scales” (see Figure 1 for schematic).

These multi-scale studies enable to understand the fundamental properties of collagen, such as bending stiffness, response to tensile stretching, and the interaction of several collagen fibers with each other, in particular under shear loading. Furthermore, none of the previously reported studies have investigated the large-strain, permanent or plastic or fracture behavior of single fibers of tropocollagen, and the self-assembly processes associated with those molecules.

6.3 Full atomistic calculations

The atomic structure of a single tropocollagen molecule is determined according to X-ray diffraction data obtained by experiment, and the structures are taken directly from the Protein Data Bank (PDB). For the studies reviewed here, the crystal structure PDB ID 1QSU, with 1.75 Å resolution, as reported by Kramer and coworkers¹⁵⁰ is used. The 1QSU is a triple-helical collagen-like molecule with sequence (Pro-Hyp-Gly)₄-Glu-Lys-Gly-(Pro-Hyp-Gly)₅.

The charges of each atom are assigned according to the CHARMM rules implemented in the tLEaP program. Hydrogen atoms are added according to pH 7, and the protonation states of individual amino acids are assigned according to pH 7 as well. The CHARMM input files (structure and topology files) are then used to perform NAMD calculations.

For ReaxFF calculations, no typing is necessary, and charges are determined dynamically during the simulation. Hydrogen atoms are added using the NAMD/CHARMM procedure, according to the same conditions as outlined above. We include crystallographic water in the simulations, and also add a skin of water of a few Angstrom surrounding each tropocollagen molecule, or, in the case of two molecules, surrounding the overall geometry. We ensure that the whole molecule assembly is embedded in water.

Before finite temperature, dynamical calculations are performed, we carry out an energy minimization for several thousand steps, making sure that convergence is achieved, thus relieving any potential overlap in vdW interactions after adding hydrogen atoms. In the second step, we anneal the molecule after heating it up to a temperature $T=300$ K, annealing it. The heat up rate is $\Delta T = 25$ K every 25 steps, and we keep the temperature fixed after the final temperature $T = 300$ K is achieved (then we apply a temperature control in an NVT ensemble¹¹⁹). We also ensure that the energy remains constant after the annealing procedure. The relaxed initial length of each molecule (consisting of 30 residues in each of the three chains) is $L_0 = 84$ Å.

Depending on the details of the loading case, we then apply mechanical forces using varies types of constraint, and investigate the response of the molecule due to the applied loading. Typically, we obtain force-versus-displacement data, which is then used to extract mechanical quantities such as stress and strain, using continuum mechanical concepts by drawing analogies between the molecular level and continuum mechanical theories. Steered MD is based on the concept of adding restraint force to groups of atoms by extending the Hamiltonian by an additional restraint potential of the form $\frac{1}{2}k_{\text{SMD}}(r - r_\lambda)^2$. Unless indicated otherwise, we use a Steered Molecular Dynamics (SMD) scheme¹²⁹ with spring constant $k_{\text{SMD}} = 10$ kcal/mol/Å². It was shown in previous studies that this is a reasonable choice leading to independence of the measured molecular mechanical properties from the choice of the SMD spring constant¹⁵⁷.

Figure 12 shows force-displacement curves for a single tropocollagen molecule, comparing the CHARMM and the ReaxFF force field. This simulation provides insight into the failure strength of the molecule, Young's modulus and the critical strains for onset of different modes of deformation. Figure 13 shows a few snapshot as the tropocollagen molecule undergoes significant deformation followed by fracture.

The force-versus displacement curve $F(d)$ can be used to determine a stress versus strain curve, by proper normalization:

$$\sigma(d) = \frac{F(d)}{A_c}, \quad (39)$$

where A_c denotes an equivalent area of the cross-section of a collagen molecule, and $A_c = \pi \cdot R^2 \approx 214.34$ Å², assuming that $R = 8.26$ Å (obtained from studies of an assembly of two tropocollagen molecules). Young's modulus is obtained by taking the partial derivative of the stress with respect to strain. Young's modulus of a single

tropocollagen molecule is obtained to be $E_{tens} = 6.99$ GPa. The strength of the tropocollagen molecule is determined to be $F_{max,tens} \approx 2.4 \times 10^4$ pN, reached at about 50% tensile strain.

Similar computational experiments that model the response of a tropocollagen molecule to three-point bending are carried out to calculate the bending stiffness EI , and from this, the persistence length¹⁴⁷. At $T = 300$ K, we find the persistence length $\xi \approx 23.4$ nm, which is close to results obtained by experimental investigations suggesting a persistence length of 14.5 ± 7.3 nm¹⁵⁴.

Similar studies have been carried out to model the interaction of two tropocollagen molecules, in particular to determine the axial shear strength. The shear strength is shown in Figure 14, as a function of shear deformation rate. Figure 15 depicts some snapshots during deformation. Using a linear extrapolation to vanishing shear rate, we estimate a maximum shear force of $F_{max,shear} \approx 466$ pN, corresponding to adhesion strength $\tau_{max} = 5.55$ pN/Å.

6.4 Reactive mesoscale model development

The goal is to develop a bead-type model that is also capable of describing the fracture properties, associated with bond breaking in the tropocollagen molecule, and the change in structure during deformation leading to changes in tangent elastic moduli. In the bead model, each tropocollagen molecule is represented as a linear collection of connected particles (see inlay in Figure 16).

We express the total energy of the system as

$$E = E_T + E_B + E_{weak}, \quad (40)$$

where E_T is the energy stored in chemical bonds due to stretching along the axial direction, E_B is the energy due to bending of the CNT, and E_{weak} constitute weak (vdW) interactions. The total energy contribution of each part is given by the sum over all pair-wise and triple (angular) interactions in the system, thus

$$E_{T,weak} = \sum_{pairs} \phi_I(r) \quad (41)$$

for the tensile and weak interactions (both summed pair-wise), and

$$E_B = \sum_{angles} \phi_B(\varphi), \quad (42)$$

summed over all triples of particles in the systems. The bending energy is given by

$$\phi_B(\varphi) = \frac{1}{2}k_B(\varphi - \varphi_0)^2, \quad (43)$$

with k_B as the spring constant relating to the bending stiffness, and φ as the bending angle between three particles. Calculation of the angle thus requires consideration of the position of three atoms. The molecular potential is thus a three-body potential.

The tensile force between two particles is given by

$$F_T(r) = \partial\phi_T(r)/\partial r = k_T(r)(r - r_0), \quad (44)$$

where

$$k_T = H(r_{break} - r) \begin{cases} k_T^{(0)} & r < r_1 \\ k_T^{(1)} & r \geq r_1 \end{cases}. \quad (45)$$

In eq. (45), $H(r - r_{break})$ is the Heaviside function $H(a)$, which is defined to be zero for $a < 0$, and one for $a \geq 0$, and $k_T^{(0)}$ as well as $k_T^{(1)}$ for the small and large-deformation spring constants. This particular bilinear shape of force law is chosen since it provides a reasonable approximation to the two regimes observed in full atomistic calculations (Fig. 12), corresponding to the effective elastic regime due to uncoiling ($k_T^{(0)}$) and that of stretching of covalent bonds ($k_T^{(1)} > k_T^{(0)}$).

In addition, we assume weak, dispersive interactions between either different parts of each molecule or different molecules, defined by a Lennard-Jones 12:6 (LJ) function as introduced in eq. (21), with σ^* as the distance parameter, and ε^* describing the energy well depth at equilibrium.

6.5 Mesoscale parameter training

The mass of each bead is determined by assuming a homogeneous distribution of mass in the molecular model (which is an excellent assumption). The total mass of the tropocollagen molecule used in our studies is given by 8,152.2 amu. We divide the total length of the tropocollagen molecule used in the MD studies into $N_{MD}=6$ pieces, each bead containing 5 amino acid residues. Each bead then has a weight of 1358.7 amu. Since the total length of the molecule in full atomistic studies is $L_0 = 84 \text{ \AA}$, the beads are separated by a distance $r_0 = 14 \text{ \AA}$. The beads represent different sequences in tropocollagen that when added together make the entire sequence.

The LJ potential has been shown to be a good model for such dispersive interactions¹³². The mesoscale model parameters are all determined from an extensive set of full

atomistic simulation, part of which has been briefly reviewed in Section 6.2. For example, the distance parameter σ^* is given by

$$\sigma^* = \frac{D}{\sqrt[6]{2}} \approx 14.72 \text{ \AA}, \quad (46)$$

where D is the equilibrium distance as measured in the MD simulations, $D=16.52 \text{ \AA}$. The parameter ε^* is obtained so that the mesoscale and atomistic models yield the identical shear strength between two tropocollagen molecules (Figs. 14 and 15).

The spring constant $k_T^{(0)}$ for small deformation is calculated from Young's modulus for small deformation is given by the expressions derived within the framework of the Cauchy-Born model (see eq. (27) solved for the spring constant k):

$$k_T^{(0)} = \frac{A_c}{r_0} E. \quad (47)$$

An equivalent expression can be found for the large-strain (tangent) modulus. The parameters r_1 and r_{break} are obtained from the critical strains for switch between the two tangent slopes (denoted by ε_1 , which is approximately 30%), and from the breaking strain (denoted by ε_{break} , which is approximately 50%):

$$r_1 = (1 + \varepsilon_1)r_0. \quad (48a)$$

$$r_{break} = (1 + \varepsilon_{break})r_0. \quad (48b)$$

The bending stiffness parameter k_B is obtained by the condition that the mesoscale model and the full atomistic model yield the same bending stiffness. Similar expressions can be derived for all other parameters involved in the mesoscale model (for further details, see ¹⁴⁶).

Table 1 summarizes all parameters used in the mesoscale model. The mesoscale model has been validated to reproduce the tensile response of the pure atomistic calculations as shown in Figure 12, for instance. This result also illustrates that the bilinear approximation for the force law is a reasonable approximation.

6.6 Deformation of ultra-long tropocollagen molecules: Entropy and fracture

Figure 16 shows results stretching a highly coiled tropocollagen molecule (see schematic in Figure 5). Figure 17 shows a zoom into the regime of elastic deformation during which entropy controls. Figure 17 also includes a quantitative comparison with the WLC model discussed earlier. The total simulation time for these tensile tests is approximately 2 μ sec, with an end-displacement rate on order of approximately 0.5 m/sec. This is

clearly not in reach with pure atomistic methods, but appropriate for the reactive mesoscale model.

The agreement between the molecular studies and the WLC model is reasonably good. Compared to the elastic deformation modes (Figure 16), the entropic regime is characterized by rather low tensile forces, below 100 pN.

Figure 18 depicts several snapshots as the tropocollagen molecule undergoes tensile deformation from a coiled structure to a linear molecule that eventually ruptures.

6.7 Mechanics of collagen fibrils

The mesoscale model is capable of modeling complete collagen fibrils, at scales that are not reachable using pure atomistic calculation. The reactive mesoscale model enables a new method to treat the various complexities of chemical interactions while considering entropic effects naturally.

Figure 19 depicts results of large-scale molecular modeling for the deformation of a single collagen fibril, composed out of 200 individual tropocollagen molecules with each 335 nm length. Figure 20 shows the stress-strain plot of this collagen fibril as it undergoes tensile deformation. We observe several deformation regimes, including linear elasticity (A), onset of plastic (permanent) deformation (B), fracture (C) and continuous shearing of molecules (D).

7. Discussion and conclusion

We have reviewed recent advances in atomistic and molecular modeling to describe the elastic, plastic and fracture properties of biological materials, from an atomistic perspective, linking the molecular scale of chemistry with the mechanics of materials.

We have exemplified some of the techniques in models of tropocollagen molecules and collagen fibrils. Using full atomistic calculations, we reviewed a suite of calculations of different mechanical loading types to gain insight into the deformation behavior of tropocollagen molecules. The results suggest that it is critical to include a correct description of the bond behavior and breaking processes at large bond stretch, information stemming from the quantum chemical details of bonding. We have used the full atomistic simulation results within a hierarchical modeling approach to describe the properties of ultra long tropocollagen molecules (Figures 16-18). Our results provide estimates of the fracture and deformation strength of tropocollagen molecules, for different types of loading, enabling a comparison of different relative strengths. We find that the tensile strength of tropocollagen molecules is largest, approaching 2.4×10^4 pN (Figure 12), while the shear deformation constitutes the weakest link in the system, on the order of 466 pN (Figure 14). This further suggests the importance of shear deformation on collagen mechanics, and opens the possibility to improve the material behavior by tuning the shear resistance, for example by introducing cross-linking.

Further, we find a strong rate dependence of the mechanical properties, including Young's modulus and shear strength (see, for example Figure 14). This is in agreement with the fact that collagen is known to be a viscoelastic material. We also observe that the properties of collagen are scale dependent. For example, the fracture strength of an individual tropocollagen molecule (maximum stress of approximately 11.2 GPa) is different from the fracture strength of a collagen fibril (maximum stress of approximately 0.66 GPa). This agrees with the general notion that the properties of biological hierarchical materials are often scale-dependent.

Information about the fracture behavior of collagen may be essential to understand the role of collagen components in biological materials. For example, the mechanics of collagen fibers at large stretch may play a critical role in the mechanical properties of bone during crack propagation, and elucidation of its mechanical response in particular at large strains is of critical importance during crack bridging in bone-like hard tissues¹⁶⁵⁻¹⁶⁷.

We believe that reactive modeling that takes into account the complexity of chemical bonding may be critical to understand the fracture and deformation behavior of other biological and protein-based materials.

We have further reviewed an application of the new reactive mesoscale model, applied to model the tensile response of collagen fibrils (Figures 19 and 20). We have also demonstrated that the mesoscopic model can be used to describe tensile experiments of pulling long molecules, naturally including the effects of entropy. Such results could be immediately compared with experiment¹⁵².

The mesoscopic model can also be used to probe how changes in the sequence would influence macroscopic properties, including the possibility of cross-links. Such effects could be captured using the reactive potential. Our numerical methods could eventually be developed into engineering tools for recombinant DNA technologies.

Similar methods as those discussed here could be used for collagen molecules with different sequences leading to different adhesion parameters. The importance of chemistry could also be investigated in studies related to enzymatic activity in the proximity of collagen. Future studies could thus focus on addressing some of those aspects, including degradation of collagen due to enzymes, or other chemicals in the environment, similar to stress assisted cracking. We believe that the new reactive force fields are a promising approach to address some of those aspects of collagen mechanics.

8. Acknowledgements

MJB gratefully acknowledges support from the Department of Civil and Environmental Engineering at the Massachusetts Institute of Technology. MJB is grateful for helpful discussions with Yu Ching Yung (Harvard University) on the general topic of biological materials.

References

1. Broberg, K. B., *Cracks and Fracture*. Academic Press: 1990.
2. Courtney, T. H., *Mechanical behavior of materials*. McGraw-Hill: 1990.
3. Springborg, M., *Density-functional methods in chemistry and materials science*. Wiley research series in Theoretical Chemistry: 1997.
4. Vashishta, P.; Kalia, R. K.; Nakano, A. *Comp. in Science and Engrg.* **1999**, 56-65.
5. Zhou, S. J.; Beazly, D. M.; Lomdahl, P. S.; Holian, B. L. *Phys. Rev. Lett.* **1997**, 78, 479-482.
6. Zhou, S. J.; Preston, D. L.; Lomdahl, P. S.; Beazley, D. M. *Science* **1998**, 279, 1525-1527.
7. Buehler, M. J.; Gao, H., *Large-scale atomistic modeling of dynamic fracture*. CRC Press: 2006.
8. Kadau, K.; Germann, T. C.; Lomdahl, P. S. *Int. J. Mod. Phys. C* **2004**, 15, 193.
9. Buehler, M. J. *Journal of Computational and Theoretical Nanoscience* **2006**, in press.
10. Bao, G.; Suresh, S. *Nature Materials* **2003**, 2, (11), 715-725.
11. Van Vliet, K. J.; Bao, G.; Suresh, S. *Acta Materialia* **2003**, 51, (19), 5881-5905.
12. Drexler, K. E. *Journal Of Computational And Theoretical Nanoscience* **2006**, 3, (1), 1-10.
13. Buehler, M. J.; Duin, A. C. T. v.; Goddard, W. A. *Phys. Rev. Lett.* **2006**, 96, (9), 095505.
14. Buehler, M. J.; Gao, H., *Ultra large scale atomistic simulations of dynamic fracture*. American Scientific Publishers (ASP): 2006.
15. Buehler, M. J.; Abraham, F. F.; Gao, H. *Nature* **2003**, 426, 141-146.
16. Buehler, M. J.; Gao, H. *Physik in unserer Zeit* **2004**, 35, (1).
17. Buehler, M. J.; Hartmaier, A.; Gao, H. *Modelling And Simulation In Materials Science And Engineering* **2004**, 12, (4), S391-S413.
18. Buehler, M. J.; Hartmaier, A.; Duchaineau, M. A.; Abraham, F. R.; Gao, H. J. *Acta Mechanica Sinica* **2005**, 21, (2), 103-111.
19. Hirth, J. P.; Lothe, J., *Theory of Dislocations*. Wiley-Interscience: 1982.
20. Griffith, A. A. *Phil. Trans. Roy. Soc. A* **1920**, 221, 163-198.
21. Irwin, G. R. in: *Fracturing of Metals* **1948**, 147-166.
22. Freund, L. B. *J. Mech. Phys. Solids* **1977**, 25, 69-79.
23. Freund, L. B., *Dynamic Fracture Mechanics*. Cambridge University Press, ISBN 0-521-30330-3: 1990.
24. Anderson, T. L., *Fracture mechanics: Fundamentals and applications*. CRC Press: 1991.
25. Buehler, M. J. *J. Mater. Res.* **Under submission**.
26. Blanckenhagen, B. v.; Gumbsch, P.; Arzt, E. *Phil. Mag. Lett.* **2003**, 83, 1-8.
27. Nix, W. D. *Scripta Materialia* **1998**, 39, 545-554.
28. Gao, H.; Zhang, L.; Nix, W. D.; Thompson, C. V.; Arzt, E. *Acta Mater.* **1999**, 47, 2865-2878.
29. Hartmaier, A.; Buehler, M. J.; Gao, H. J. *Advanced Engineering Materials* **2005**, 7, (3), 165-169.

30. Buehler, M. J.; Hartmaier, A.; Gao, H. *Modelling And Simulation In Materials Science And Engineering* **2004**, 12, (4), S391-S413.
31. Buehler, M. J.; Kong, Y.; Yao, H.; Gao, H. *Journal Of Engineering Materials And Technology* **2006**, 128, (1), 3-10.
32. Buehler, M. J.; Yao, H.; Gao, H.; Ji, B. *Model. Sim. Mat. Science and Engr.* **In press**.
33. Haslam, A. J.; Moldovan, D.; Phillpot, S. R.; Wolf, D.; Gleiter, H. *Comput. Mat. Science* **2003**, 23, 15-32.
34. Nieh, T. G.; Wadsworth, J. *Scripta Met.* **1991**, 25, (4).
35. Yamakov, V.; Wolf, D.; Salazar, M.; Phillpot, S. R.; Gleiter, H. *Acta mater.* **2001**, 49, 2713-2722.
36. Yamakov, V.; Wolf, D.; Phillpot, S. R.; Gleiter, H. *Acta mater.* **2002**, 50, 61-73.
37. Yamakov, V.; Wolf, D.; Phillpot, S. R.; Mukherjee, A. K.; Gleiter, H. *Nature Materials* **2002**, 1, 1-4.
38. Schiotz, J.; Vegge, T.; DiTolla, D.; Jacobsen, K. W. *Phys. Rev. B* **1999**, 60, 11971-11983.
39. Yamakov, V.; Wolf, D.; Phillpot, S. R.; Gleiter, H. *Phil. Mag. Lett.* **2003**, 83, 385-393.
40. Hasnaoui, A.; Swygenhoven, H. v.; Derlet, P. M. *Science* **2003**, 300, (5625), 1550-1552.
41. Farkas, D.; Swygenhoven, H. v.; Derlet, P. *Phys. Rev. B* **2002**, 66, 184112.
42. Swygenhoven, H. v.; Derlet, P. M. *Phys. Rev. B* **2001**, 64, 224105.
43. Van Swygenhoven, H.; Derlet, P. M.; Froseth, A. G. *Nature Materials* **2004**, 3, (6), 399-403.
44. Yamakov, V.; Wolf, D.; Phillpot, S. R.; Mukherjee, A. K.; Gleiter, H. *Nature Materials* **2004**, 3, (1), 43-47.
45. Vinci, R. P.; Zielinski, E. M.; Bravman, J. C. *Thin Solid Films* **1995**, 262, 142-153.
46. Weiss, D.; Gao, H.; Arzt, E. *Acta mater.* **2001**, 49, 2395-2403.
47. Dehm, G.; Balk, T. J.; Blanckenhagen, B. v.; Gumbsch, P.; Arzt, E. *Z. Metallk.* **2002**, 93, (5), 383-391.
48. Buehler, M. J.; Hartmaier, A.; Gao, H. *J. Mech. Phys. Solids* **2003**, 51, 2105-2125.
49. Nix, W. D. *Metall. Trans. A* **1989**, 20, 2217-2245.
50. Nicola, L.; Giessen, E. V. d.; Needleman, A. *J. Appl. Phys.* **2003**, 93, 5920-5928.
51. Hartmaier, A.; Buehler, M. J.; Gao, H., A discrete dislocation plasticity model of creep in polycrystalline thin films. In *Defects And Diffusion In Metals: An Annual Retrospective Vi*, 2004; Vol. 224-2, pp 107-125.
52. Buehler, M. J.; Hartmaier, A.; Gao, H. *Journal Of The Mechanics And Physics Of Solids* **2003**, 51, (11-12), 2105-2125.
53. Arzt, E. *Acta Mater.* **1998**, 46, 5611-5626.
54. Diao, J. K.; Gall, K.; Dunn, M. L.; Zimmerman, J. A. *Acta Materialia* **2006**, 54, (3), 643-653.
55. Gall, K.; Diao, J. K.; Dunn, M. L.; Haftel, M.; Bernstein, N.; Mehl, M. J. *Journal Of Engineering Materials And Technology-Transactions Of The Asme* **2005**, 127, (4), 417-422.

56. Gall, K.; Diao, J. K.; Dunn, M. L. *Nano Letters* **2004**, 4, (12), 2431-2436.
57. Lin, J. S.; Ju, S. P.; Lee, W. J. *Physical Review B* **2005**, 72, (8).
58. Nam, C. Y.; Jaroenapibal, P.; Tham, D.; Luzzi, D. E.; Evoy, S.; Fischer, J. E. *Nano Letters* **2006**, 6, (2), 153-158.
59. Wang, B. L.; Shi, D. N.; Jia, J. M.; Wang, G. H.; Chen, X. S.; Zhao, J. J. *Physica E-Low-Dimensional Systems & Nanostructures* **2005**, 30, (1-2), 45-50.
60. Mylvaganam, K.; Zhang, L. C. *Journal Of Computational And Theoretical Nanoscience* **2005**, 2, (3), 385-388.
61. Diao, J. K.; Gall, K.; Dunn, M. L. *Journal Of The Mechanics And Physics Of Solids* **2004**, 52, (9), 1935-1962.
62. Liang, W.; Zhou, M. *Proceedings Of The Institution Of Mechanical Engineers Part C-Journal Of Mechanical Engineering Science* **2004**, 218, (6), 599-606.
63. Park, H. S.; Zimmerman, J. A. *Physical Review B* **2005**, 72, (5).
64. Gu, Q.; Cheng, C. D.; Gonela, R.; Suryanarayanan, S.; Anabathula, S.; Dai, K.; Haynie, D. T. *Nanotechnology* **2006**, 17, (1), R14-R25.
65. Gudiksen, M. S.; Lauhon, L. J.; Wang, J.; Smith, D. C.; Lieber, C. M. *Nature* **2002**, 415, (6872), 617-620.
66. Li, B.; Daggett, V. *Journal Of Muscle Research And Cell Motility* **2002**, 23, (5-6), 561-573.
67. Ulm, F. J. *Materials And Structures* **2003**, 36, (261), 426-438.
68. Constantinides, G.; Ulm, F.; Van Vliet, K. *Materials And Structures* **2003**, 36, (257), 191-196.
69. Peterlik, H.; Roschger, P.; Klaushofer, K.; Fratzl, P. *Nature Materials* **2006**, 5, (1), 52-55.
70. Gupta, H. S.; Wagermaier, W.; Zickler, G. A.; Aroush, D. R. B.; Funari, S. S.; Roschger, P.; Wagner, H. D.; Fratzl, P. *Nano Letters* **2005**, 5, (10), 2108-2111.
71. Jaeger, C.; Groom, N. S.; Bowe, E. A.; Horner, A.; Davies, M. E.; Murray, R. C.; Duer, M. J. *Chemistry Of Materials* **2005**, 17, (12), 3059-3061.
72. Fratzl, P.; Gupta, H. S.; Paschalis, E. P.; Roschger, P. *Journal Of Materials Chemistry* **2004**, 14, (14), 2115-2123.
73. Weiner, S.; Wagner, H. D. *Annual Review Of Materials Science* **1998**, 28, 271-298.
74. Gao, H.; Ji, B.; Jäger, I. L.; Arzt, E.; Fratzl, P. *P. Natl. Acad. Sci. USA* **2003**, 100, (10), 5597-5600.
75. Langer, R.; Tirrell, D. A. *Nature* **2004**, 428, (6982), 487-492.
76. Petka, W. A.; Harden, J. L.; McGrath, K. P.; Wirtz, D.; Tirrell, D. A. *Science* **1998**, 281, (5375), 389-392.
77. Maskarinec, S. A.; Tirrell, D. A. *Current Opinion In Biotechnology* **2005**, 16, (4), 422-426.
78. Smeenk, J. M.; Otten, M. B. J.; Thies, J.; Tirrell, D. A.; Stunnenberg, H. G.; van Hest, J. C. M. *Angewandte Chemie-International Edition* **2005**, 44, (13), 1968-1971.
79. Michon, T.; Tirrell, D. A. *Biofutur* **2000**, 2000, (197), 34-38.
80. Tirrell, D. A. *Abstracts Of Papers Of The American Chemical Society* **2002**, 224, U408-U408.
81. Lim, C. T.; Dao, M.; Suresh, S.; Sow, C. H.; Chew, K. T. *Acta Materialia* **2004**, 52, (7), 1837-1845.

82. Dao, M.; Lim, C. T.; Suresh, S. *Journal Of The Mechanics And Physics Of Solids* **2003**, 51, (11-12), 2259-2280.
83. Boey, S. K.; Boal, D. H.; Discher, D. E. *Biophysical Journal* **1998**, 75, (3), 1573-1583.
84. Discher, D. E.; Boal, D. H.; Boey, S. K. *Biophysical Journal* **1998**, 75, (3), 1584-1597.
85. Boal, D. H. *Biophysical Journal* **1994**, 67, (2), 521-529.
86. Boal, D. H.; Seifert, U.; Zilker, A. *Physical Review Letters* **1992**, 69, (23), 3405-3408.
87. Suresh, S.; Spatz, J.; Mills, J. P.; Micoulet, A.; Dao, M.; Lim, C. T.; Beil, M.; Seufferlein, T. *Acta Biomaterialia* **2005**, 1, (1), 15-30.
88. Li, J.; Dao, M.; Lim, C. T.; Suresh, S. *Biophysical Journal* **2005**, 88, (5), 3707-3719.
89. Yung, Y. C., Mooney, D.J., Engineering Smooth Muscle. In *CRC Biomedical Engineering Handbook*, Fisher, J. P., Ed. CRC Press: Boca Raton, FL, In Press.
90. Martin, R. B.; Burr, D. B.; Sharkey, N. A., *Skeletal tissue mechanics*. Springer: New York, 2001.
91. Hellmich, C.; Ulm, F. J. *Journal Of Biomechanics* **2002**, 35, (9), 1199-1212.
92. Hellmich, C.; Ulm, F. J. *Journal Of Engineering Mechanics-Asce* **2002**, 128, (8), 898-908.
93. An, K. N.; Sun, Y. L.; Luo, Z. P. *Biorheology* **2004**, 41, (3-4), 239-246.
94. de Carmejane, O.; Morris, M. D.; Davis, M. K.; Stixrude, L.; Tecklenburg, M.; Rajachar, R. M.; Kohn, D. H. *Calcified Tissue International* **2005**, 76, (3), 207-213.
95. Gates, T. S.; Odegard, G. M.; Frankland, S. J. V.; Clancy, T. C. *Composites Science And Technology* **2005**, 65, (15-16), 2416-2434.
96. Sun, Y. L.; Luo, Z. P.; Fertala, A.; An, K. N. *Journal Of Biomechanics* **2004**, 37, (11), 1665-1669.
97. Weiner, J. J. *Phys. Rev. B* **1983**, 24, 845-848.
98. Tadmor, E. B.; Ortiz, M.; Phillips, R. *Phil. Mag. A* **1996**, 73, 1529.
99. Knap, J.; Ortiz, M. *J. Mech. Phys. Sol.* **2001**, 49, (9), 1899-1923.
100. Dupuy, L. M.; Tadmor, E. B.; Miller, R. E.; Phillips, R. *Physical Review Letters* **2005**, 95, (6).
101. Gao, H.; Ji, B. *Engineering Fracture Mechanics* **2003**, 70, 1777-1791.
102. Gao, H.; Klein, P. *J. Mech. Phys. Solids* **2001**, 46, (2), 187-218.
103. Klein, P.; Gao, H. *Engineering Fracture Mechanics* **1998**, 61, 21-48.
104. Guo, X.; Wang, J. B.; Zhang, H. W. *International Journal Of Solids And Structures* **2006**, 43, (5), 1276-1290.
105. Jiang, H.; Huang, Y.; Hwang, K. C. *Journal Of Engineering Materials And Technology-Transactions Of The Asme* **2005**, 127, (4), 408-416.
106. Duin, A. C. T. v.; Dasgupta, S.; Lorant, F.; Goddard, W. A. *J. Phys. Chem. A* **2001**, 105, 9396-9409.
107. Cheung, S.; Deng, W.; Duin, A. C. T. v.; Goddard, W. A. *J. Phys. Chem. A* **2005**, 109, 851.
108. Duin, A. C. T. v.; Strachan, A.; Stewman, S.; Zhang, Q.; Xu, X.; Goddard, W. A. *J. Phys. Chem. A* **2003**, 107, 3803-3811.

109. van Duin, A. C. T.; Nielson, K.; Deng, W. Q.; Oxgaard, J.; Goddard, W. A. *Abstracts Of Papers Of The American Chemical Society* **2004**, 227, U1031-U1031.
110. <http://www.top500.org/>, TOP 500 Supercomputer Sites. In.
111. Gropp, W.; Lusk, W.; Skjellum, A., *Using MPI*. MIT Press: 1999.
112. Norskov, J. K.; Schiotz, J.; Jacobsen, K. W. https://wiki.fysik.dtu.dk/ase/Atomic_Simulation_Environment.
113. Cummings, J.; Aivazis, M.; Samtaney, R.; Radovitzky, R.; Mauch, S.; Meiron, D. *Journal Of Supercomputing* **2002**, 23, (1), 39-50.
114. Buehler, M. J. CMDF website: <http://www.wag.caltech.edu/home/mbuehler/cmdf/>. 2005.
115. Buehler, M. J.; Goddard, W. A. http://www.wag.caltech.edu/home/mbuehler/cmdf/CMDF_Proceedings.pdf **2005**.
116. Buehler, M. J.; Dodson, J.; Meulbroek, P.; Duin, A.; Goddard, W. A. *Mat. Res. Soc. Proceedings* **2006**, 894, LL3.8.
117. H. Tang, R. R., M.J. Buehler, A.C.T. v. Duin, W.A. Goddard. *Mat. Res. Soc. Proceedings* **2006**.
118. Hauch, J. A.; Holland, D.; Marder, M.; Swinney, H. L. *Future Generation Computer Systems* **2003**, 19, 599-609.
119. Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*. Oxford University Press: 1989.
120. Morse, P. M. *Phys. Rev.* **1929**, 34, 57-64.
121. Tersoff, J. *Phys. Rev. Lett.* **1988**, 61, (25), 2879-2883.
122. Stillinger, F.; Weber, T. A. *Phys. Rev. B* **1985**, 31, (8), 5262-5271.
123. Oh, D. J.; Johnson, R. A. *J. Mater. Res.* **1988**, 3, 471-478.
124. Baskes, M. I. *Phys. Rev. B* **1984**, 29, (12), 6443-6543.
125. Mishin, Y.; Mehl, M. J.; Papaconstantopoulos, D. A.; Voter, A. F.; Kress, J. D. *Phys. Rev. B* **2001**, 63, 224106.
126. Baskes, M. I. *Journal Of Metals* **1988**, 40, (11), 123-123.
127. Baskes, M. I. *Materials Chemistry and Physics* **1997**, 50, (2), 152-158.
128. 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. *Journal Of Physical Chemistry B* **1998**, 102, (18), 3586-3616.
129. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *Journal Of Computational Chemistry* **2005**, 26, (16), 1781-1802.
130. Nelson, M. T.; Humphrey, W.; Gursoy, A.; Dalke, A.; Kale, L. V.; Skeel, R. D.; Schulten, K. *International Journal Of Supercomputer Applications And High Performance Computing* **1996**, 10, (4), 251-268.
131. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Computer Physics Communications* **1995**, 91, (1-3), 1-41.
132. Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *Journal Of Physical Chemistry* **1990**, 94, (26), 8897-8909.

133. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *Journal Of The American Chemical Society* **1992**, 114, (25), 10024-10035.
134. Brenner, D. W.; Shenderova, O. A.; Harrison, J. A.; Stuart, S. J.; Ni, B.; Sinnott, S. B. *Journal Of Physics-Condensed Matter* **2002**, 14, (4), 783-802.
135. Stuart, S. J.; Tutein, A. B.; Harrison, J. A. *Journal Of Chemical Physics* **2000**, 112, (14), 6472-6486.
136. Strachan, A.; van Duin, A. C. T.; Chakraborty, D.; Dasgupta, S.; Goddard, W. A. *Physical Review Letters* **2003**, 91, (9).
137. Nielson, K. D.; Duin, A. C. T. v.; Oxgaard, J.; Deng, W.; Goddard, W. A. *J. Phys. Chem. A* **2005**, 109, 49.
138. Han, S. S.; van Duin, A. C. T.; Goddard, W. A.; Lee, H. M. *Journal Of Physical Chemistry A* **2005**, 109, (20), 4575-4582.
139. Chenoweth, K.; Cheung, S.; van Duin, A. C. T.; Goddard, W. A.; Kober, E. M. *Journal Of The American Chemical Society* **2005**, 127, (19), 7192-7202.
140. Strachan, A.; Kober, E. M.; van Duin, A. C. T.; Oxgaard, J.; Goddard, W. A. *Journal Of Chemical Physics* **2005**, 122, (5).
141. Cheung, S.; Deng, W. Q.; van Duin, A. C. T.; Goddard, W. A. *Journal Of Physical Chemistry A* **2005**, 109, (5), 851-859.
142. Tao, L.; Buehler, M. J.; Duin, A. C. T. v.; Goddard, W. A. *Under submission* **2005**.
143. Datta, D.; Duin, A. C. T. v.; Goddard, W. A. *Unpublished* **2005**.
144. Rappé, A. K.; Goddard, W. A. *J. of Physical Chemistry* **1991**, 95, (8), 3358-3363.
145. Brenner, D. W. *Physical Review B* **1990**, 42, (15), 9458-9471.
146. Buehler, M. J. *J. Mater. Res.* **In press**.
147. Buehler, M. J. *Under submission*.
148. Borel, J. P.; Monboisse, J. C. *Comptes Rendus Des Seances De La Societe De Biologie Et De Ses Filiales* **1993**, 187, (2), 124-142.
149. Bhattacharjee, A.; Bansal, M. *Iubmb Life* **2005**, 57, (3), 161-172.
150. Kramer, R. Z.; Venugopal, M. G.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *Journal Of Molecular Biology* **2000**, 301, (5), 1191-1205.
151. Zervakis, M.; Gkoumplias, V.; Tzaphlidou, M. *Medical Engineering & Physics* **2005**, 27, (8), 655-667.
152. Bozec, L.; Horton, M. *Biophysical Journal* **2005**, 88, (6), 4223-4231.
153. Layton, B. E.; Sullivan, S. M.; Palermo, J. J.; Buzby, G. J.; Gupta, R.; Stallcup, R. E. *Microelectronics Journal* **2005**, 36, (7), 644-649.
154. Sun, Y. L.; Luo, Z. P.; Fertala, A.; An, K. N. *Biochemical And Biophysical Research Communications* **2002**, 295, (2), 382-386.
155. Arnoux, P. J.; Bonnoit, J.; Chabrand, P.; Jean, M.; Pithioux, M. *European Physical Journal-Applied Physics* **2002**, 17, (1), 65-73.
156. Waite, J. H.; Qin, X. X.; Coyne, K. J. *Matrix Biology* **1998**, 17, (2), 93-106.
157. Lorenzo, A. C.; Caffarena, E. R. *Journal Of Biomechanics* **2005**, 38, (7), 1527-1533.
158. Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2005**, 44, (5), 1414-1422.
159. Israelowitz, M.; Rizvi, S. W. H.; Kramer, J.; von Schroeder, H. P. *Protein Engineering Design & Selection* **2005**, 18, (7), 329-335.

160. Mooney, S. D.; Klein, T. E. *Molecular & Cellular Proteomics* **2002**, 1, (11), 868-875.
161. Mooney, S. D.; Kollman, P. A.; Klein, T. E. *Biopolymers* **2002**, 64, (2), 63-71.
162. Mooney, S. D.; Huang, C. C.; Kollman, P. A.; Klein, T. E. *Biopolymers* **2001**, 58, (3), 347-353.
163. Bischoff, J. E.; Arruda, E. M.; Grosh, K. *Journal Of Biomechanics* **2000**, 33, (6), 645-652.
164. Freeman, J. W.; Silver, F. H. *Journal Of Theoretical Biology* **2004**, 229, (3), 371-381.
165. Ritchie, R. O.; Kruzic, J. J.; Muhlstein, C. L.; Nalla, R. K.; Stach, E. A. *International Journal Of Fracture* **2004**, 128, (1-4), 1-15.
166. Nalla, R. K.; Kruzic, J. J.; Kinney, J. H.; Ritchie, R. O. *Biomaterials* **2005**, 26, (2), 217-231.
167. Nalla, R. K.; Kinney, J. H.; Ritchie, R. O. *Biomaterials* **2003**, 24, (22), 3955-3968.
168. Buehler, M. J.; Hartmaier, A.; Gao, H. J.; Duchaineau, M.; Abraham, F. F. *Computer Methods In Applied Mechanics And Engineering* **2004**, 193, (48-51), 5257-5282.

Tables

Equilibrium bead distance r_0 (in Å)	14.00
Critical hyperelastic strain r_1 (in Å)	18.20
Bond breaking distance r_{break} (in Å)	21.00
Tensile stiffness parameter $k_T^{(0)}$ (in kcal/mol/Å ²)	17.13
Tensile stiffness parameter $k_T^{(1)}$ (in kcal/mol/Å ²)	97.66
Equilibrium angle φ_0 (in degrees)	180.00
Bending stiffness parameter k_B (in kcal/mol/Å ²)	14.98
Dispersive parameter ε^* (in kcal/mol)	6.87
Dispersive parameter σ^* (in Å)	14.72
Mass of each bead (in amu)	1358.7

Table 1: Summary of the parameters used in the reactive mesoscopic model, chosen based on full atomistic modeling of solvated tropocollagen molecules.

Figures

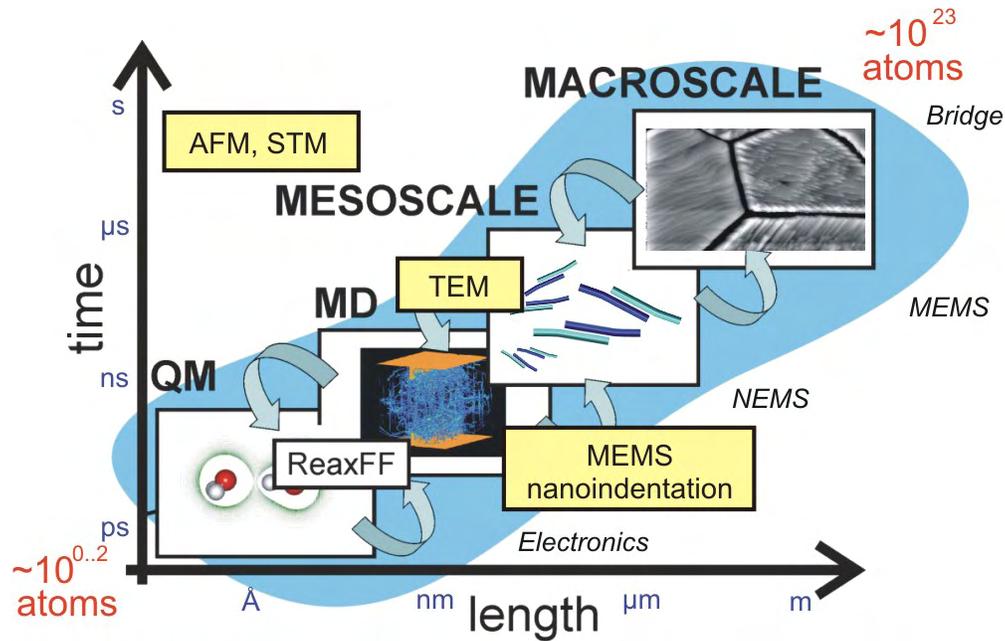


Figure 1: Multi-scale view of materials, summarizing different modeling techniques. Hierarchical multi-scale modeling is based on the idea to train coarser scales with finer scale calculations, enabling to link DFT, ReaxFF, empirical potentials to mesoscale and continuum models. This scheme of overlapping methods has been quite successful in bridging the small to the large. We also include a selection experimental techniques in the plot that operate on time- and length-scales comparable to the modeling and simulation tools. Techniques such as AFM, STM, TEM and nanoindentation allow to probe the nanomechanics of materials at extremely fine length scales, and opens up new avenues for complementing theoretical-experimental studies.

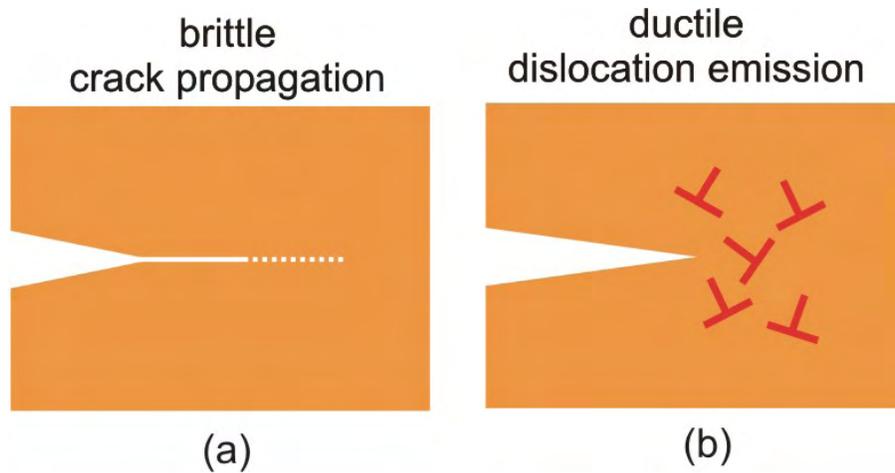


Figure 2: Brittle versus ductile material behavior ^{14, 18, 168}. Cracks or similar defects exist in any material, and the tip of the crack represents an area of extremely large stresses that can induce failure of the material due to breaking or reformation of atomic bonds. How the material responds to these large stresses determines whether it is brittle or ductile. Brittle fracture is characterized by formation of new surfaces due to breaking of atomic bonds, often leading to catastrophic and sudden failure (example for brittle materials: glass, many ceramics, some polymers). Ductile deformation is characterized by nucleation of dislocations, characteristic line defects in crystals, which are indicated by the \top -symbol (example for ductile materials: many metals such as copper, nickel).

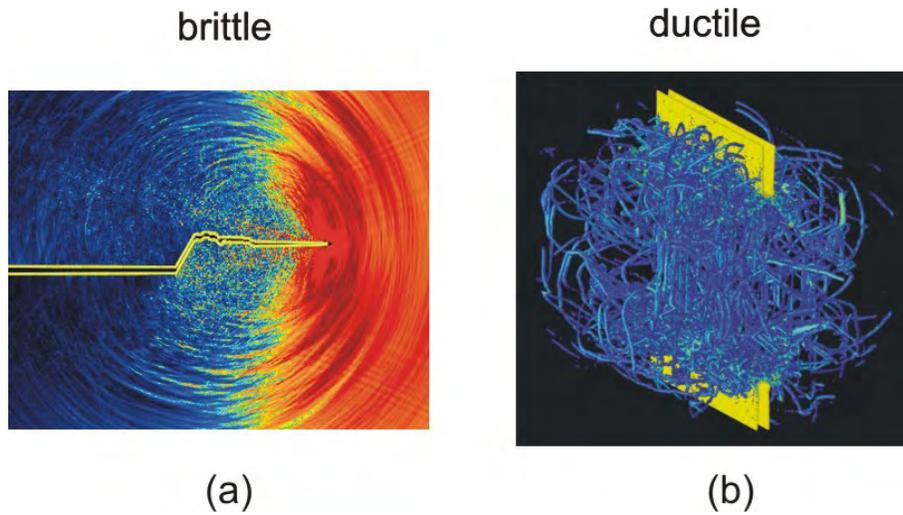


Figure 3: Results of large-scale molecular dynamics modeling of brittle (subplot (a)) and ductile (subplot (b)) materials behavior^{14, 18, 168}. It is clearly visible that brittle fracture involves breaking of atomic bonds leading to creation of new fracture surfaces (yellow color). Ductile materials deform by creation of many dislocations, which appear as the wiggly bluish lines in the plot.

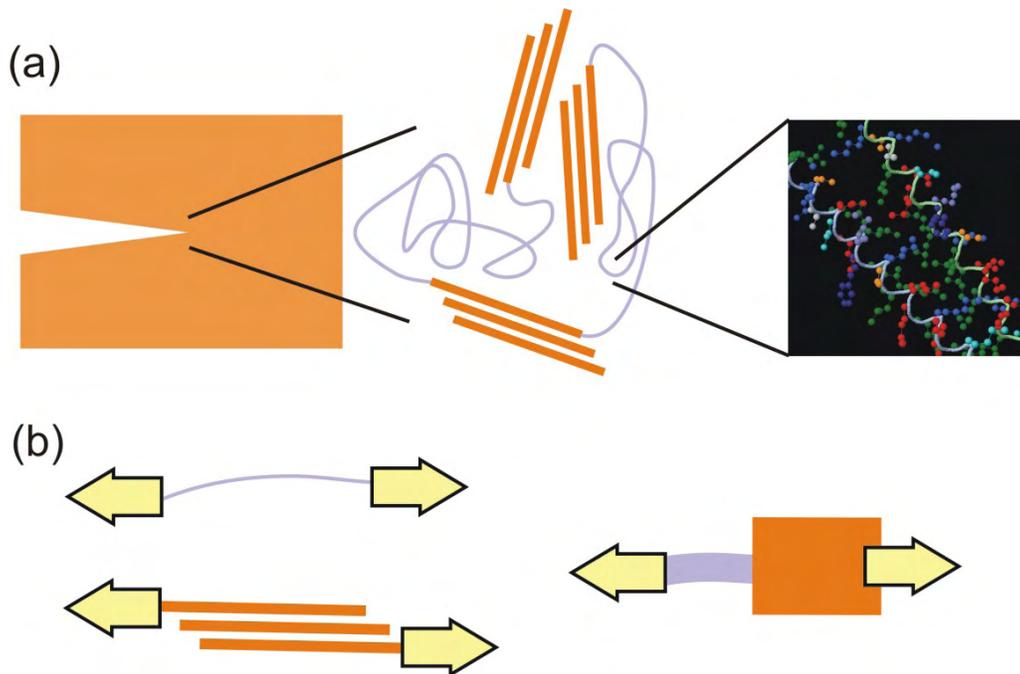


Figure 4: Response of a hierarchical material to a macroscopic defect, e.g. a crack (subplot (a)). The plot shows a “macro-view”, the molecular scale, and the atomistic building blocks. Due to the heterogeneous structure of this material, the response of the individual building blocks needs to be considered (subplot (b)). The individual building blocks may respond in various ways, including energetic elasticity, fracture, or entropic elasticity, or by a mixture of different types of elastic and plastic behavior. The size of the defect, here the crack, relative to the size of the building blocks is important to characterize the response of the material.

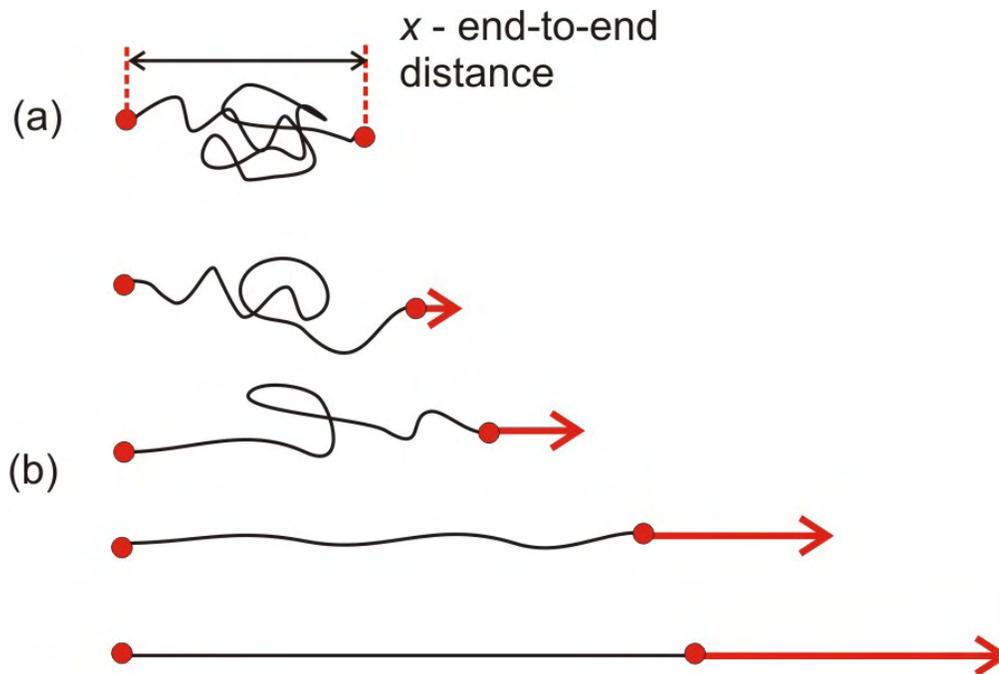


Figure 5: Entropy controlled molecular elasticity. Subplot (a): Coiled, entangled state of a molecule with contour length much larger than the persistence length. The end-to-end distance is measured by the variable x . Subplot (b): Response of the molecule to mechanical loading. As the applied force is increased, the end-to-end distance x increases until the molecule is fully extended. Clearly, the continued disentanglement leads to a reduction of entropy in the system, which induces a force that can be measured as an elastic spring. Once the molecule is fully extended, the change in entropy due to increased force approaches zero, and the elastic response is controlled by changes in the internal potential energy of the system, corresponding to the energetic elasticity.

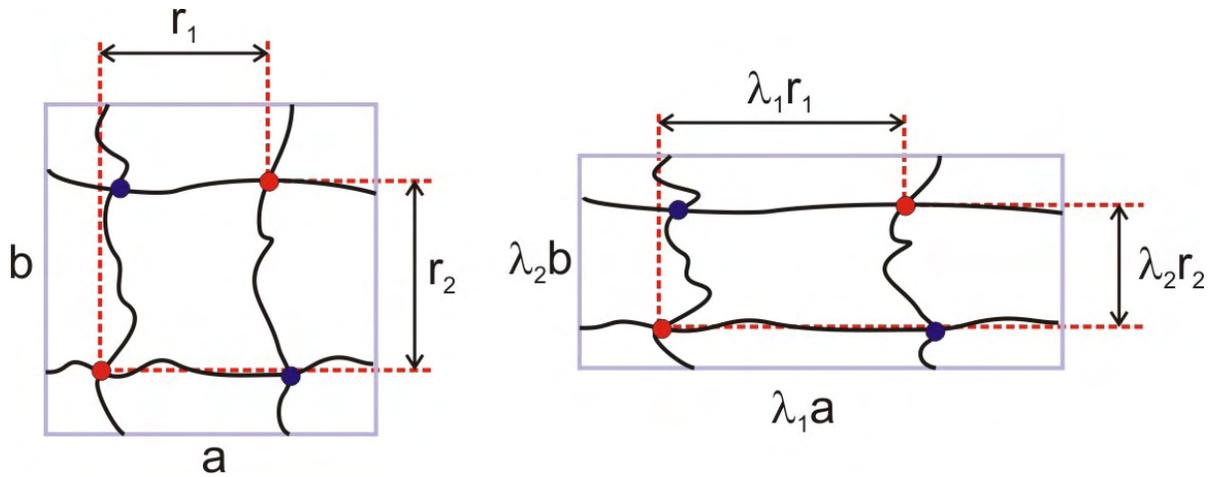


Figure 6: The concepts of entropic elasticity of single molecules can be immediately applied to understand 2D and 3D networks of molecules in a polymer. This figure demonstrates how a change in state of deformation poses constraints on the end-to-end distances of molecules, influencing the entropy of the system. Such considerations enable to link the properties of single molecules (their entropy) with the overall macroscopic elastic behavior of the material.

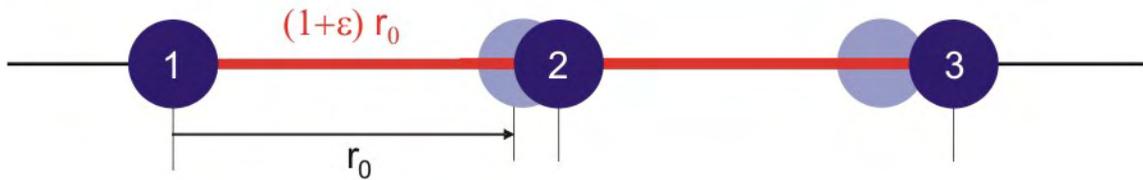


Figure 7: Application of Cauchy-Born rule to express the strain energy density as a function of the macroscopic deformation field, taking into consideration the specific atomic microstructure. This technique, here outlined for a 1D molecule, represented by a collection of connected particles, can be used to link molecular interactions with the overall mechanics, under conditions where energetic elasticity governs over entropic contributions.

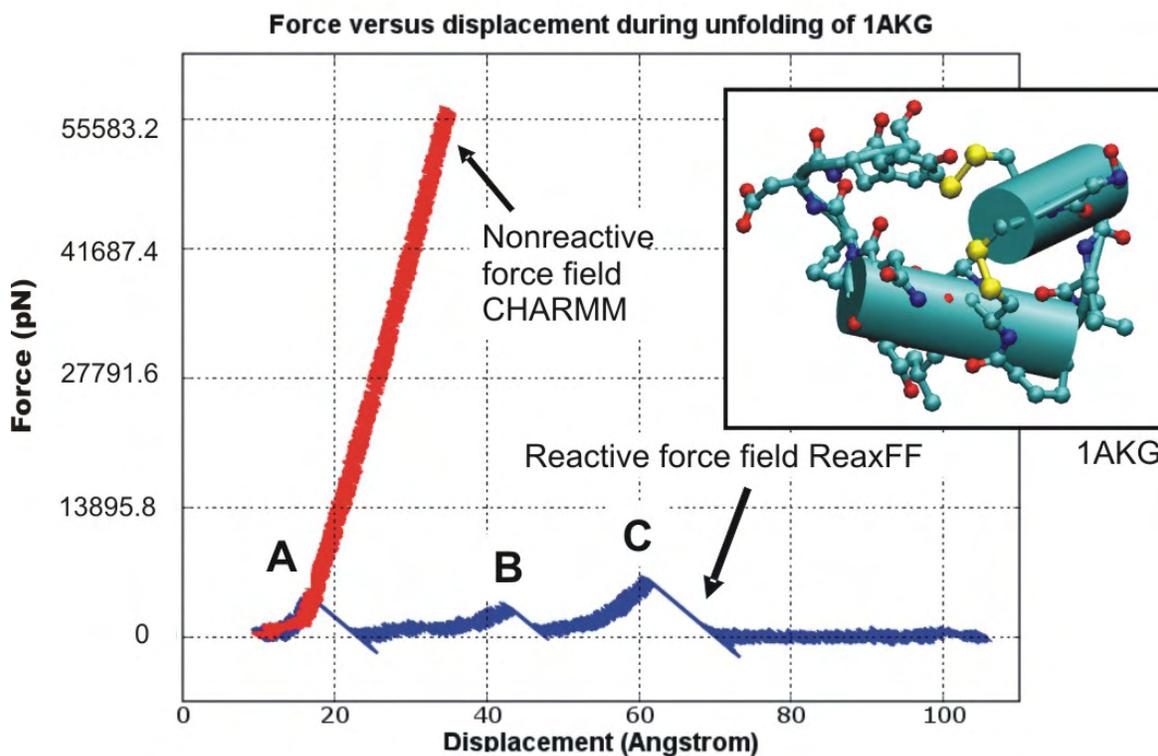


Figure 8: Mechanically stimulated unfolding of a small protein PnIB 1AKG, modeling results obtained using CHARMM (red curve) and the reactive force field ReaxFF (blue curve). We apply a slowly increasing force at the N- and C-terminus of the polypeptide. It is apparent that although the ReaxFF and CHARMM descriptions agree for small deformation (below $\approx 17\text{\AA}$), they disagree strongly for larger deformation. The difference can be explained based on the fact that the CHARMM potential is incapable of describing breaking of the disulfide bonds. The points “A”-“C” characterize different stages of deformation; “A” and “B” correspond to breaking of the two disulfide bonds, with lower force for the second one. Point “C” characterizes breaking of the backbone chain. The strength of the backbone chain is slightly larger than the disulfide bonds.

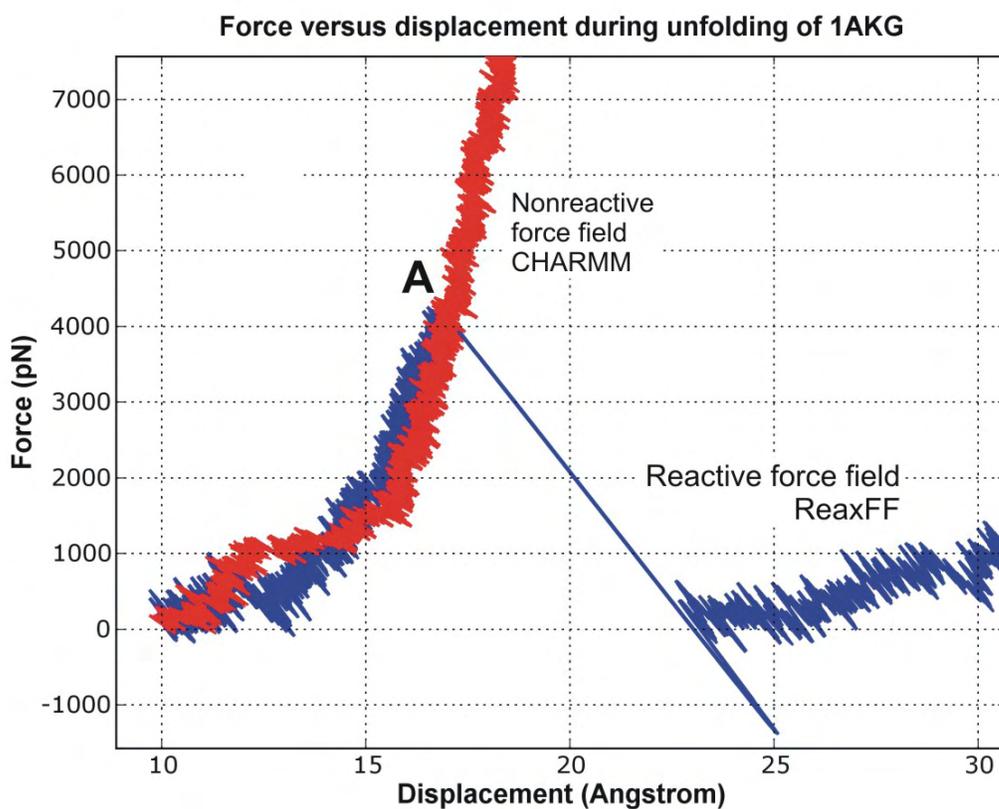


Figure 9: Mechanically stimulated unfolding of a small protein PnIB 1AKG. This plot shows a zoom into smaller displacements (same data as shown in the previous figure). It is apparent that although the ReaxFF and CHARMM descriptions agree for small deformation (below $\approx 17\text{\AA}$), they disagree strongly for larger deformation, until the first covalent disulfide bond ruptures (“A”).

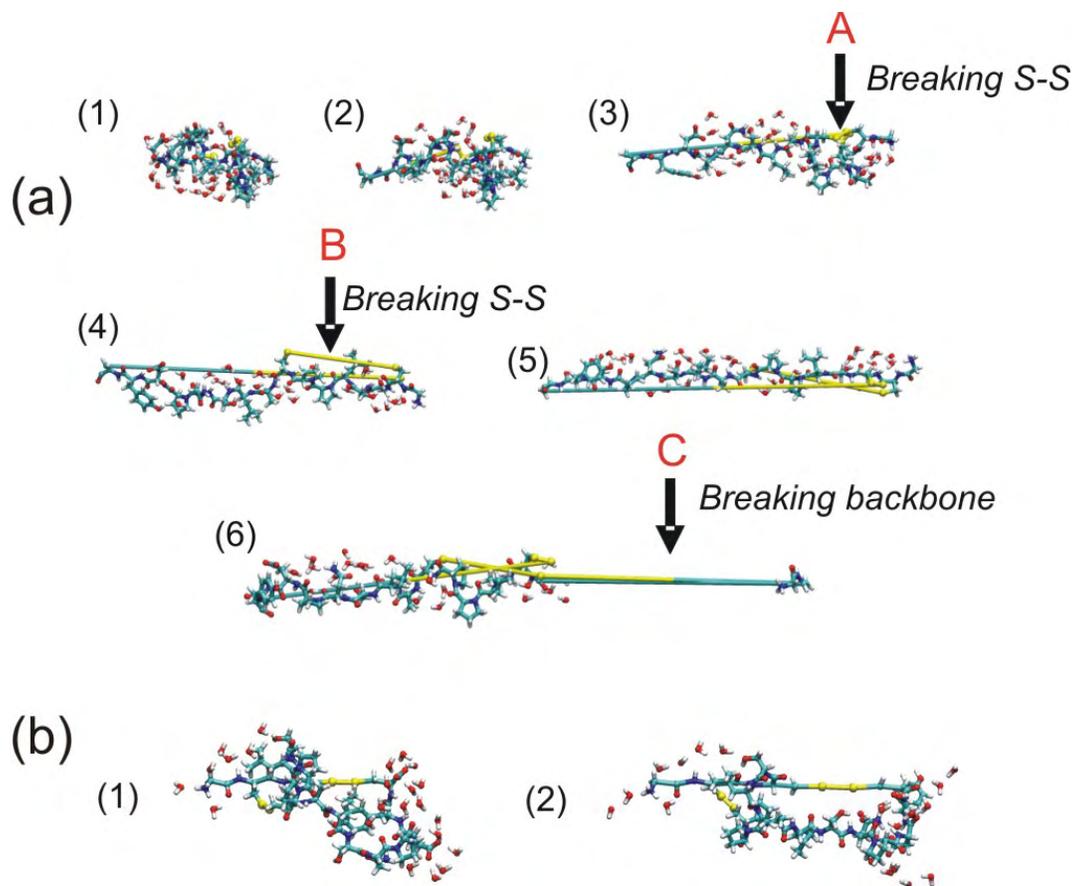


Figure 10: Snapshots of mechanically stimulated unfolding of a small protein. The plot depicts results obtained using the reactive force field ReaxFF (subplot (a)), and CHARMM (subplot (b)). The points “A”-“C” correspond to those shown in Figure 8.

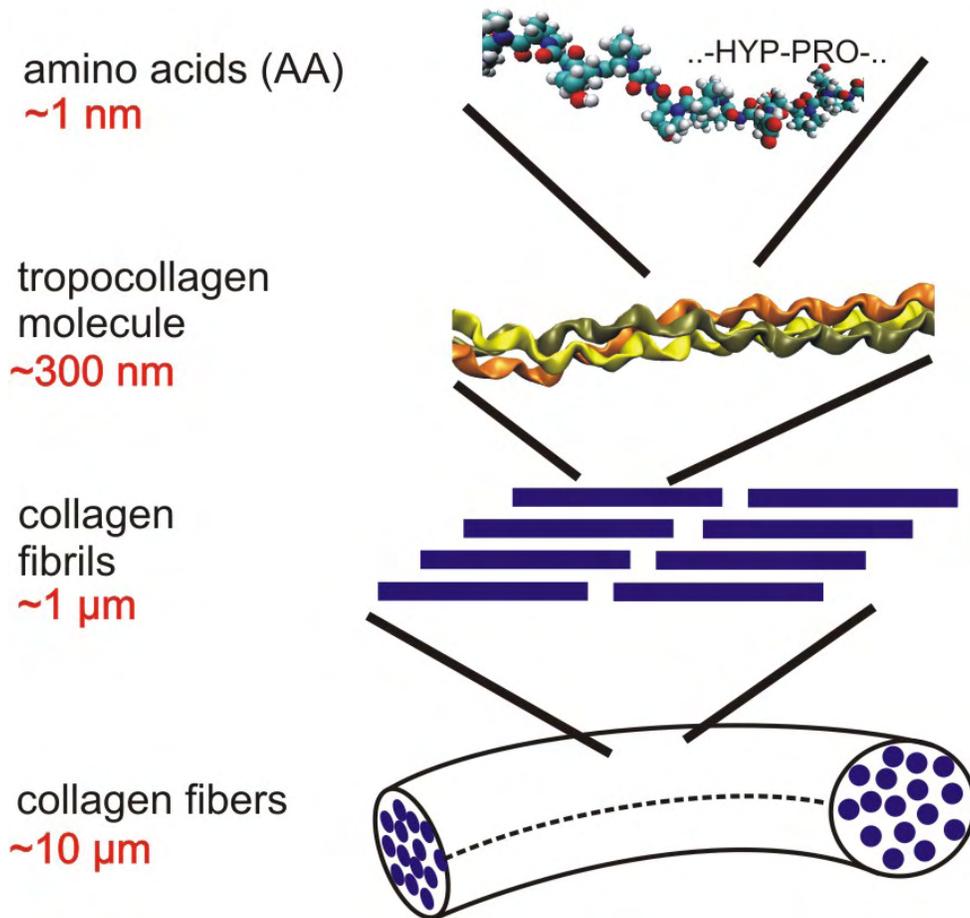


Figure 11: The hierarchical structure of collagen. The figure shows a schematic view on some of the hierarchical features of collagen, ranging from the amino acid or polypeptide level at the nanoscale up to the scale of collagen fibers with lengths on the order of 10 μm . In this article, we focus on the mechanical properties of collagen fibrils, consisting of a staggered array of tropocollagen molecules.

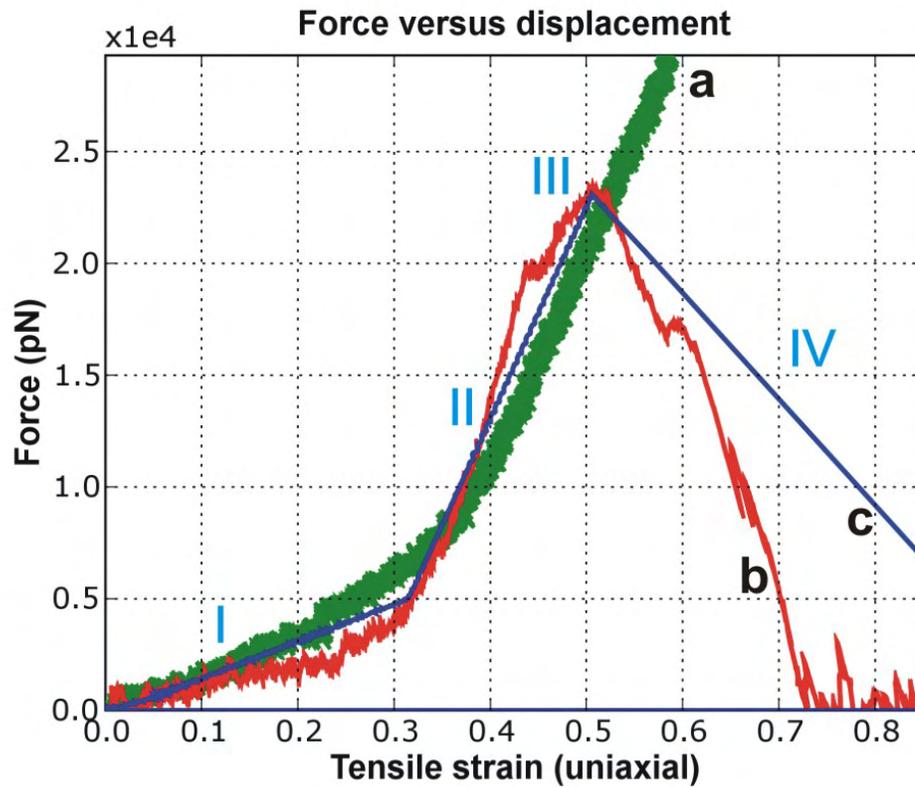


Figure 12: Force-displacement response of a single tropocollagen molecule. Curve (a) shows the results obtained using the nonreactive CHARMM force field model, curve (b) depicts results of ReaxFF modeling, and curve (c) the results obtained using the reactive mesoscale model. We observe four different regimes and states during 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 after the triple helices have been uncoiled. In regime III, we observe onset of fracture followed by a rapid decay of the force in regime IV.

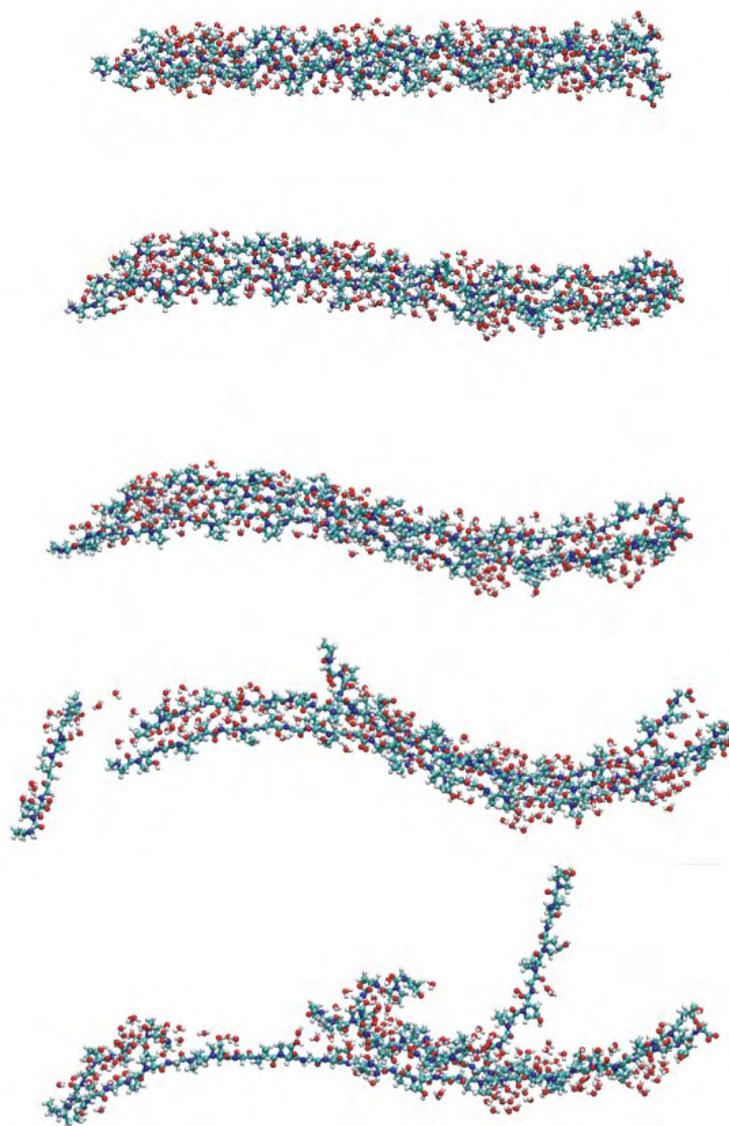


Figure 13: Fracture dynamics of a single tropocollagen molecule. Fracture initiates at the location with large radius of curvature, and starts with breaking of an individual polypeptide. The two other polypeptides break soon after.

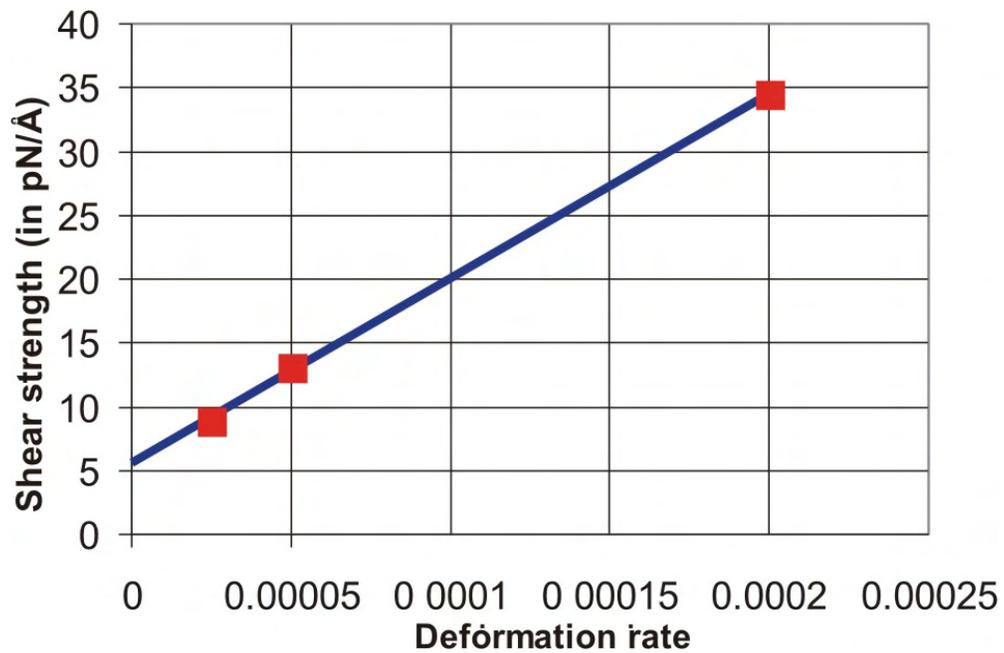


Figure 14: Shear strength (adhesion strength) between two tropocollagen molecules, as a function of shear deformation rate. The results indicate a strong rate dependence, possibly a manifestation of the viscoelastic properties of collagen at the molecular scale.

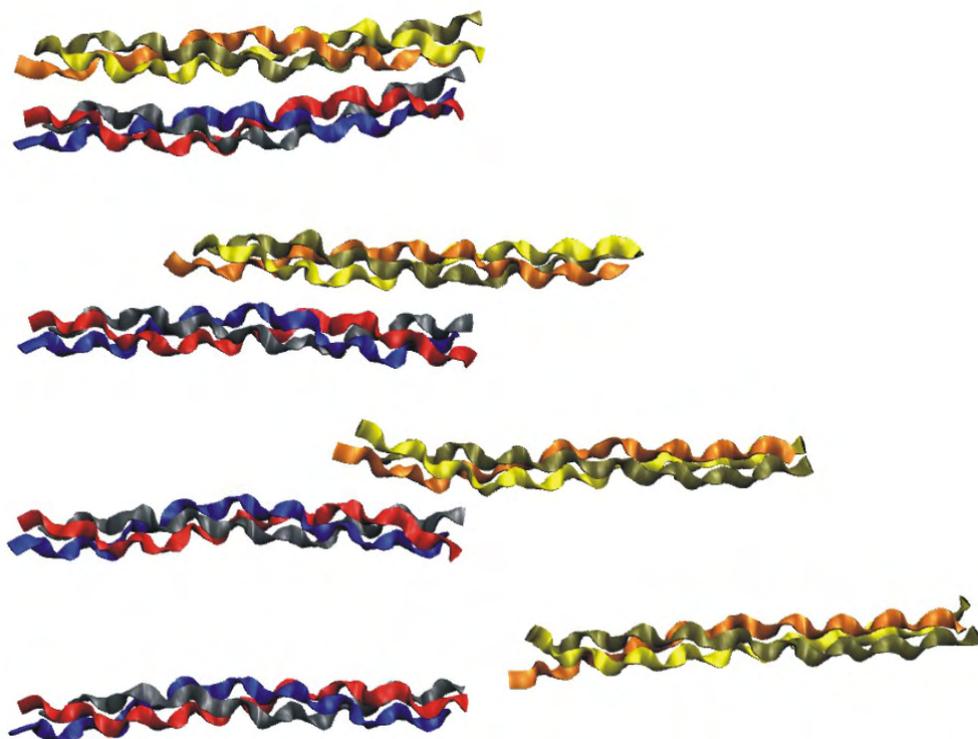


Figure 15: Snapshots of shearing of two tropocollagen molecules. The polypeptides are colored by molecule number, showing the characteristic triple helical structure.

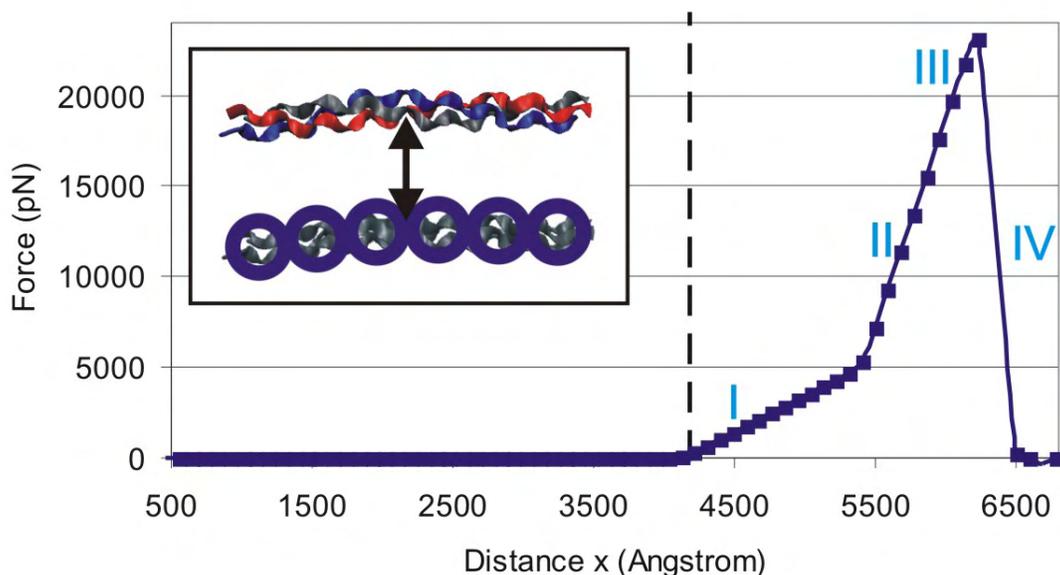


Figure 16: Force-displacement curve of stretching a single tropocollagen molecule, with molecular length of approximately 418 nm, obtained using the reactive mesoscale model, at 300K. The inlay depicts the scheme of replacing the molecule by a linear assembly of beads. The loading of the long molecule starts from a coiled entangled configuration of the molecule with end-to-end distance of approximately 500Å (entangled configuration is shown in Figure 18). During the initial regime, the molecule uncoils, while the applied forces are relatively low. This regime characterized by entropic elasticity appears to be basically flat is followed by the four stages discussed earlier, during which energetic elasticity governs. The four stages are (I) uncurling of the triple helix, and (II) stretching of covalent bonds in the individual polypeptides, and (III/IV) rupture of the tropocollagen molecule. The dashed line of indicates the contour length of the molecule ($x = L$).

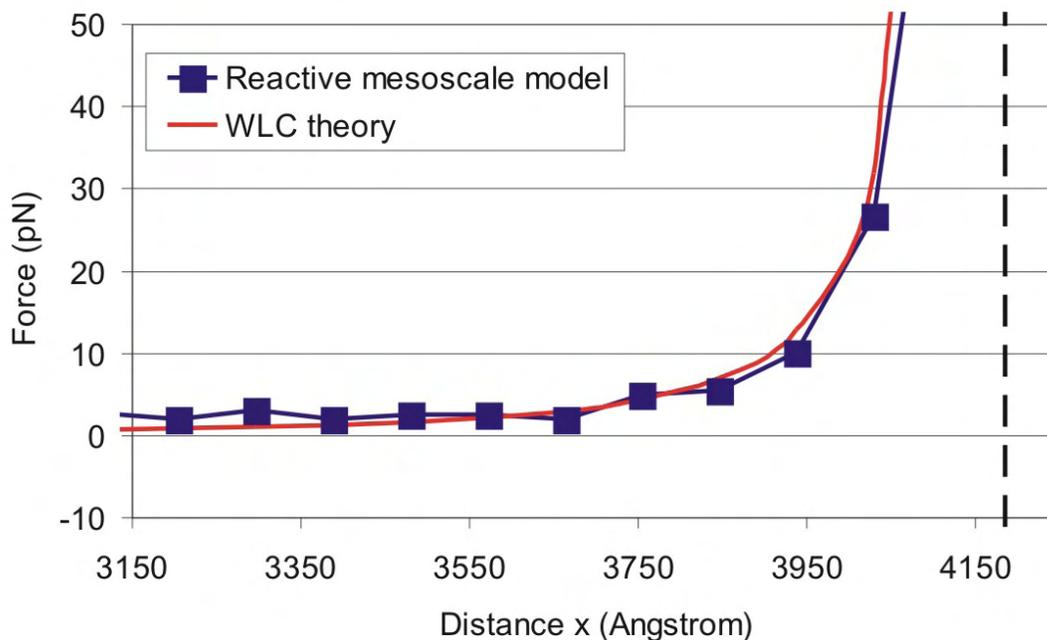


Figure 17: Entropic elasticity of a single tropocollagen molecule, length approximately 418 nm, at 300 K. The data shown here is a subset of the results depicted in the previous figure, focusing on the uncoiling response that is controlled by reduction of entropy. The dashed line indicates the contour length of the tropocollagen molecule. Once the contour length is reached, the elasticity transforms into energetic elasticity with significantly increased modulus. The forces due to entropic contributions are relatively low, rising up to approximately 100 pN until energetic elasticity takes over. Note that the WLC model diverges when $x \rightarrow L$, which explains the larger slope of the WLC prediction for large distances x . The dashed line indicates the contour length of the molecule ($x = L$).

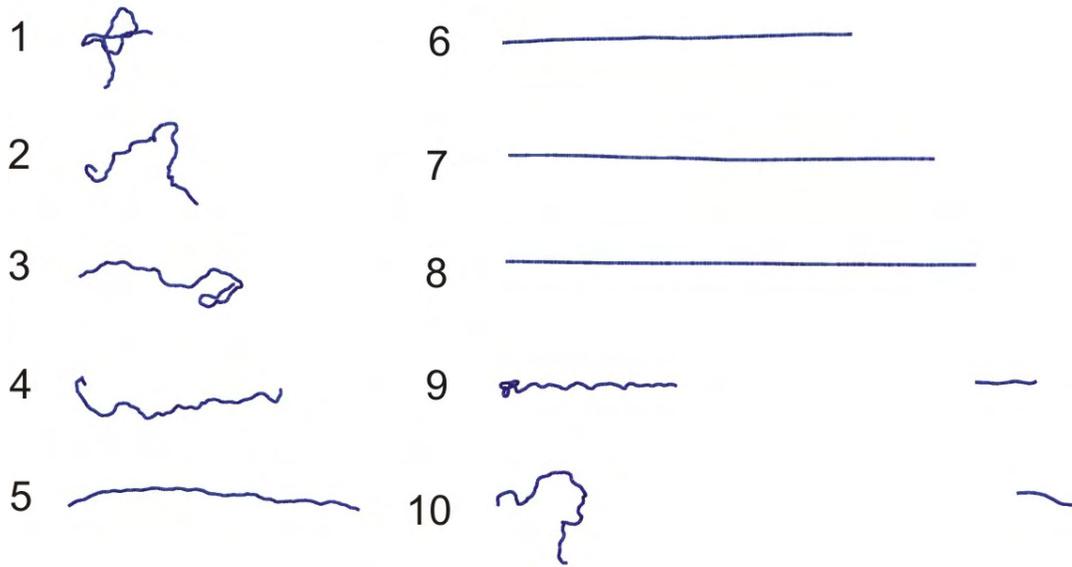


Figure 18: Snapshots of deformation of a tropocollagen molecule under tensile force, at 300K. The left column shows early stages of deformation, which is controlled by entropy (change of elastic energy small, entangled molecule, snapshots 1-5). At a critical applied force, the molecule becomes straight (snapshot 6), when energetic elasticity starts to dominate (change of entropy small, snapshots 7-8). This deformation continues until the molecule ruptures (right column, lower part, snapshots 9 and 10). After the molecule has ruptured, the free molecule starts to entangle again, due to lack of applied force. The smaller piece on the right however stays relatively straight, since it is close to the persistence length.

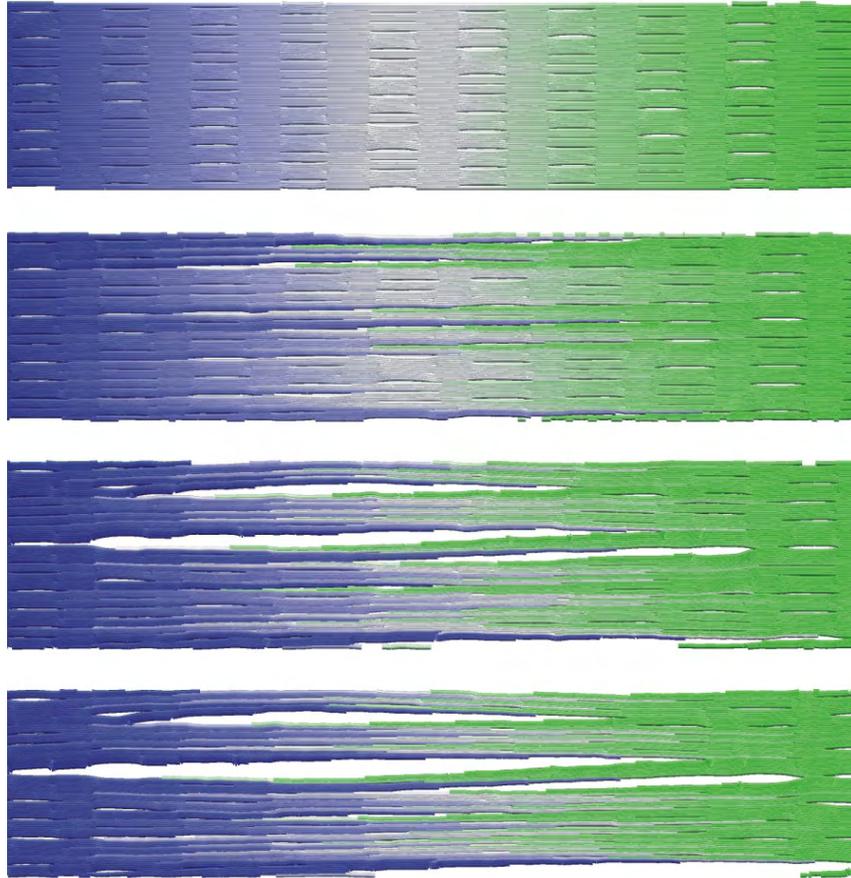


Figure 19: Large-scale molecular modeling of the mechanics of a collagen fibril that consists of 200 tropocollagen molecules, with approximately 335 nm length each. The simulation is 2D with periodic boundary conditions applied in the direction orthogonal to the applied load. The system comprises of 48,000 mesoscale beads and was carried out on a LINUX computing cluster. The results indicate that the initial elastic regime is followed by a regime of intermolecular shear that induces void formation at later stages of deformation. (Color scheme: Displacement relative to initial configuration.)

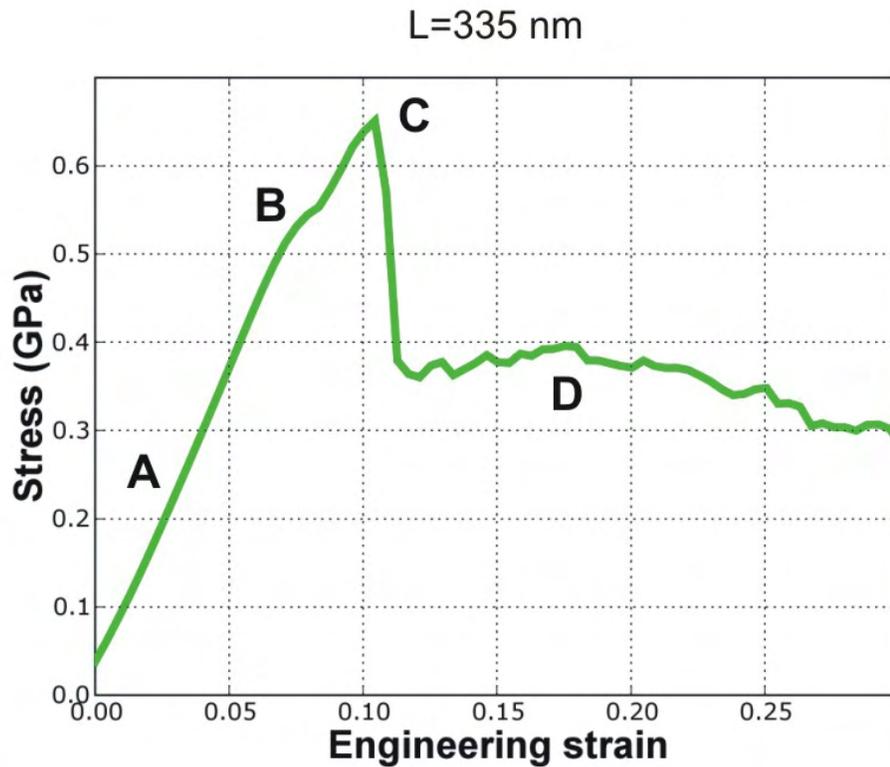


Figure 20: Stress versus strain plot, for a solvated collagen fibril without cross-links, with molecular length of approximately 335 nm. The tensile elastic modulus of the fibril is around $E \approx 6.6$ GPa, with maximum stress of approximately 0.66 GPa. In comparison many metals have a modulus on the order of 100.200 GPa. In regime (A), linear elastic deformation dominates, in regime (B) we observe onset of plastic (permanent) deformation at engineering strain of approximately 7%. In regime (C), fracture is observed, indicated by a sudden drop of the stress in the fibril, at 10.5% tensile strain. Regime (D) is continuous sliding of tropocollagen molecules above each other. This shear deformation continues until the complete fibril is broken, and constitutes a regime of much energy dissipation.