Contents lists available at ScienceDirect

Progress in Materials Science

Theoretical and computational hierarchical nanomechanics of protein materials: Deformation and fracture

Markus J. Buehler *, Sinan Keten, Theodor Ackbarow

Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 1-235A&B, Cambridge, MA 02139, USA

ARTICLE INFO

ABSTRACT

Proteins constitute the building blocks of biological materials such as tendon, bone, skin, spider silk or cells. An important trait of these materials is that they display highly characteristic hierarchical structures, across multiple scales, from nano to macro. Protein materials are intriguing examples of materials that balance multiple tasks, representing some of the most sustainable material solutions that integrate structure and function. Here we review progress in understanding the deformation and fracture mechanisms of hierarchical protein materials by using a materials science approach to develop structure-process-property relations, an effort defined as materiomics. Deformation processes begin with an erratic motion of individual atoms around flaws or defects that quickly evolve into formation of macroscopic fractures as chemical bonds rupture rapidly, eventually compromising the integrity of the structure or the biological system leading to failure. The combination of large-scale atomistic simulation, multi-scale modeling methods, theoretical analyses combined with experimental validation provides a powerful approach in studying deformation and failure phenomena in protein materials. Here we review studies focused on the molecular origin of deformation and fracture processes of three types of protein materials. The review includes studies of collagen – Nature's super-glue; beta-sheet rich protein structures as found in spider silk – a natural fiber that can reach the strength of a steel cable; as well as intermediate filaments – a class of alpha-helix based structural proteins responsible for the mechanical integrity of eukaryotic cells. The article concludes with a discussion of the significance of universally found structural

* Corresponding author. Tel.: +1 617 452 2750; fax: +1 617 258 6775.
 E-mail address: mbuehler@MIT.EDU (M.J. Buehler).
 URL: http://web.mit.edu/mbuehler/www/ (M.J. Buehler).

patterns such as the staggered collagen fibril architecture or the alpha-helical protein motif.

Contents

1. Introduction

Proteins constitute critical building blocks of life, forming biological materials such as hair, bone, skin, spider silk or cells, which play an important role in providing key mechanical functions in biological systems [1–9]. Failure of these materials due to flaws or extreme chemical or mechanical conditions can cause diseases and malfunctions in biological organisms. This occurs for instance in genetic disorders (e.g. rapid aging disease progeria, brittle bone disease asteogenesis imperfecta, Alport's syndrome, etc.), disease due to foreign material buildup such as in Alzheimer's disease, or during injuries and trauma. However, the fundamental deformation, fracture general failure mechanisms of biological protein materials remain largely unknown, partly due to a lack of understanding of how individual protein building blocks respond to mechanical load and how they participate in the function of the overall biological system at the mesoscale.

Such understanding is vital to advance models of diseases, the understanding of biological processes such as mechanotransduction, or and bioinspired the development of biomimetic materials. Recent progress provides us with insight into such mechanisms and clarifies for the first time how biology "works" at the ultimate, molecular scale, and how this relates to macroscopic phenomena such as cell mechanics or tissue behavior, across multiple hierarchical scales. This type of effort, the linking of mechanisms across multiple scales by using a materials science approach to provide structure-process-property links forms the emerging field of materiomics. This review article exemplifies theoretical and computational hierarchical nanomechanics approaches in the analysis of three representative protein structures, including collagenous tissues, beta-sheet rich protein materials and alpha-helix rich protein materials, illustrating how materiomics contributes to develop a *de novo* understanding of biological processes and to the potential of exploiting novel concepts in technological innovation.

1.1. Hierarchical protein materials

The field of mechanical properties of hierarchical biological materials [1–8] underwent an exciting development over the past several years, partly due to the emergence of physical science based approaches in the biological sciences, leading to cross-disciplinary investigations of materials, structures, diseases as well as the development of new treatment and diagnostics methods [10–13].

Historically, commonly used materials have been instrumental in classifying stages of civilizations, starting with the Stone Age thousands of years ago and ranging to the so-called "silicon age" in the late 20th and early 21st century. Despite the extensive and very successful utilization of materials in addressing societal needs, a systematic analysis of materials in the context of linking chemical and physical concepts with engineering applications in a systematic paradigm to link structure, processes and properties and fundamental mechanisms has not been achieved until very recently.

For instance, 50 year ago, E. Orowan, M. Polanyi and G.I. Taylor have discovered dislocations, a concept proposed theoretically in 1905 by V. Volterra. It was discovered that dislocations represent the fundamental mechanism of plastic deformation of metals [14,15]. Theoretical and physical understanding of dislocations and other nano- and microscopic deformation mechanisms was a prerequisite for major breakthroughs that utilized this knowledge for systematic material design. Advanced jet planes, cars, space shuttles and more recently, nanodevices, through synthesis of ultra-strong and heat resistant materials, could be a few examples of novel inventions that have benefited from discoveries regarding fundamental deformation mechanisms in metals. It is noted, however, that even without the knowledge of microscopic deformation mechanisms, the development of cars, airplanes and other technological achievements was possible.

Perhaps similar opportunities can be created today for the analysis and engineering of complex biological systems, based on quantitative insight into their fundamental physical and chemical features. A rigorous understanding may enable us eventually to integrate concepts from living systems into engineering materials design, seamlessly. Optical, mechanical and electrical properties at ultra-small material scales, their control, synthesis and analysis as well as their theoretical description represent major scientific and engineering opportunities. However, as in the case of conventional "engineered"
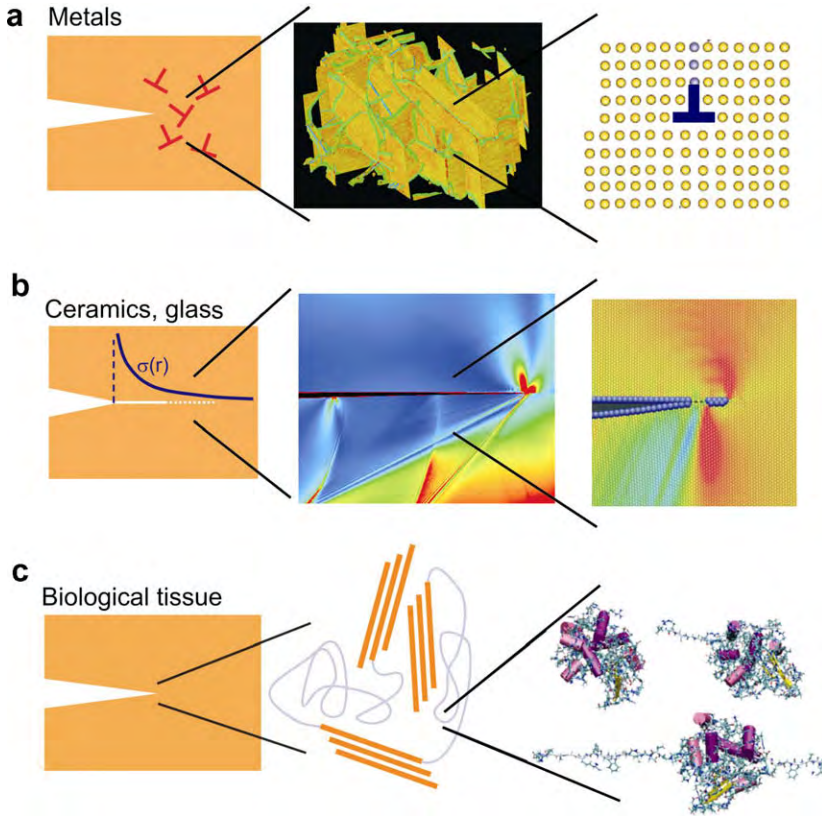
Fig. 1. Overview over the deformation and fracture behavior of different classes of materials, including (a) ductile materials [15], (b) brittle materials [16] and (c) biological protein materials (sometimes abbreviated as BPMs), as reported originally in Ref. [18]. Each subplot shows a multi-scale view of associated deformation mechanisms. In ductile materials, deformation is mediated by creation of dislocation networks; each dislocation represents localized shear of an atomic lattice. In brittle materials, fracture occurs by spreading of cracks, which is mediated by continuous breaking of atomic bonds. In BPMs, a complex interplay of different protein structures and associated mechanisms controls the mechanical response. At the ultra-scale, unfolding of individual protein molecules by rupture of hydrogen bonds (HBs) and intermolecular sliding represent some of the most fundamental deformation mechanisms.

materials, these breakthroughs may only be possible provided that their fundamental concepts are understood very well. Characterization of materials found in biology within a rigorous materials science approach is aimed towards the elucidation of these fundamental principles of assembly, deformation and fracture of these materials.

Deformation and fracture properties are intimately linked to the atomic microstructure of a material. Whereas crystalline materials show mechanisms such as dislocation spreading or crack extension, biological materials feature molecular unfolding or sliding, with a particular significance of rupture of chemical bonds such as hydrogen bonds, covalent cross-links or intermolecular entanglement. Fig. 1 displays an overview over the deformation and fracture behavior of different classes of materials, including ductile materials [15], brittle materials [16], as well as biological protein materials. Each subplot in Fig. 1 shows a multi-scale view of associated deformation mechanisms [17,18]. Different mechanisms may operate at larger length- and time-scales, where the interaction with cells and of cells with one another, different tissue types and the influence of tissue remodeling become more evident. The dominance of specific mechanisms is controlled by geometrical parameters as well as the structural arrangement of the protein elementary building blocks, across many hierarchical scales, from nano to macro (Fig. 2).
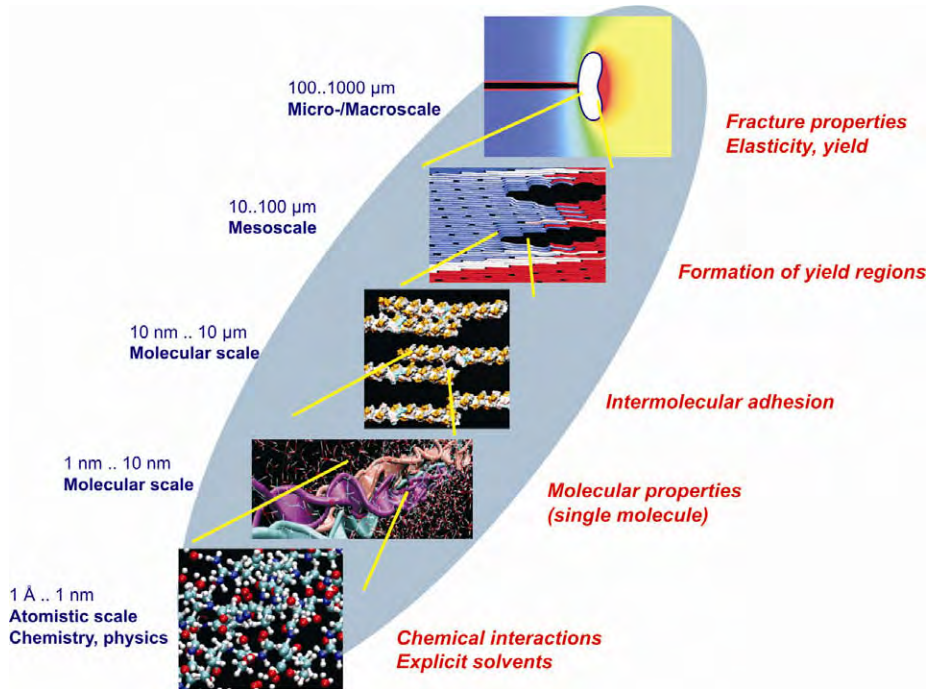
Fig. 2. Overview of different material scales, from nano to macro, here exemplified for collagenous tissue [3,4,8,26–30]. Biological protein materials such as collagen, skin, bone, spider silk or cytoskeletal networks in cells feature complex, hierarchical structures. The macroscopic mechanical material behavior is controlled by the interplay of properties throughout various scales. In order to understand deformation and fracture mechanisms, it is crucial to elucidate atomistic and molecular mechanisms at each scale and how they interact.

It is known from other fields in materials science that nano- or microscopic structures and defects control the macroscopic material behavior: for example, grain size reduction or confinement leads to an increase of the strength of crystalline metals [19–22]. Deformation maps have been proposed to characterize material properties for engineering applications [23]. Discovering similar insight for biological structures and materials represents a current frontier of research. A particularly challenging subject is the elucidation of the significance and role of nanostructures for macroscopic properties. Sensitivity analyses that show how small-scale features influence larger scale properties may effectively illustrate nanoscopic size effects in biological materials.

A major trait of biological materials is the occurrence of hierarchies and, at the molecular scale, the abundance of weak interatomic or intermolecular interactions (e.g. hydrogen bonds). The presence of hierarchies in biological materials may be vital to take advantage of molecular and sub-molecular features. For example, weak bonds in protein materials exist in hierarchical assemblies that work cooperatively to have a measurable influence for properties at larger scales. Although insignificant as individual bonds (in terms of their utilization as mechanical elements, e.g. as a glue), assemblies of weak non-covalent bonds governs structural organization and function [1,2,12,24] of these materials. Utilization of weak interactions also makes it possible to produce materials at moderate temperatures and thus with limited energy use, as well with an intrinsic self-healing ability since such bonds can reform *in situ*. An important distinction between structural control in traditional and biological materials is the geometrical occurrence of defects. While defects are often distributed randomly over the volume in crystalline and other engineered materials, biological materials consist of an ordered structure that reaches down to the nano-scale. In many biological materials, defects are placed with atomistic or molecular precision, and may play a major role in the material behavior observed at larger
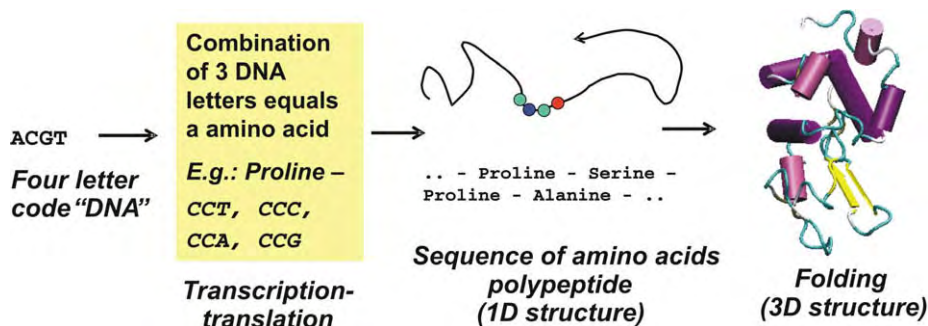
Fig. 3. Schematic to illustrate how proteins are synthesized in biological systems based on sequence information encoded in the DNA. A series of letter codes (nucleotides) is translated into the polypeptide sequence, which folds into complex three-dimensional geometries. Three DNA "letters" correspond to a specific amino acid building block (further information, see e.g. [1]).

scales. These features have been observed in bone, nacre, collagenous tissue or cellular protein networks, where some defects are highly conserved features that occur across multiple species.

Proteins are synthesized in biological systems based on sequence information encoded in the DNA (short for deoxyribonucleic acid). A series of DNA letter codes (nucleotides) is translated into the polypeptide sequence, which folds into complex three-dimensional geometries (for further details, please see [1]). This process is illustrated in schematically Fig. 3. A notable feature of this mechanism is the ability to control the structure of these materials at very small, atomistic scales, via the modification of the DNA sequence in the genetic code. Biological protein materials are typically formed by self-assembly processes, thereby forming complex hierarchical structures out of individual protein building blocks. The synthesis of a particular protein is directed via the activation of specific genes. This activation can be caused by biochemical or mechanical signals that lead to processes inside the cell's nucleus.

The mechanical properties of biological protein materials have wide ranging implications for biology. In cells for instance, mechanical sensation is used to transmit signals from the environment to the cell nucleus in order to control tissue formation and regeneration [1,25]. The structural integrity and shape of cells is controlled by the cell's cytoskeleton, which resembles an interplay of complex protein structures and signaling cascades arranged in a hierarchical fashion. Bone and collagen, providing structure to our body, or spider silk, used for prey procurement, are examples of materials that have incredible elasticity, strength and robustness unmatched by many synthetic materials, which has been attributed to its structural formation with molecular precision [1,8,12,4,26–33]. The translation of concepts observed in biology into technological applications and new materials design remains a big challenge. In particular, the combination of nanostructural and hierarchical features into materials developments could lead to significant breakthroughs.

What are the most promising strategies in order to analyze biological protein materials? Perhaps, an integrated approach that uses experiment and simulation concurrently could evolve into a new paradigm of materials research. Experimental techniques have gained unparalleled accuracy in both length- and time-scales (see Fig. 4), as reflected in development and utilization of Atomic Force Microscope (AFM) [34,35], optical tweezers [10,13,29,36] or nanoindentation [11,12,37,38] to analyze biological materials. At the same time, modeling and simulation as well as theoretical approaches have evolved into predictive tools that complement experimental analyses (see Fig. 4).

In the field of atomistic-based multi-scale stimulation, it is now possible to begin from the smallest scales (considering electrons and atoms), to reach all the way up to macroscopic scales of entire tissues, by explicitly considering the characteristic structural features at each scale. Such approaches are possible with the advent of first principles based multi-scale simulation techniques (see, for instance a review article for a broad introduction into this field [39]). The basic principle underlying these multi-scale simulation methods is "*finer scales train coarser scales*". Even though there are still major challenges ahead of us, the progress that has been reported thus far is encouraging and provides one with seemingly infinite possibilities, transforming materials science as a discipline through increased integration of computational approaches in scientific research.
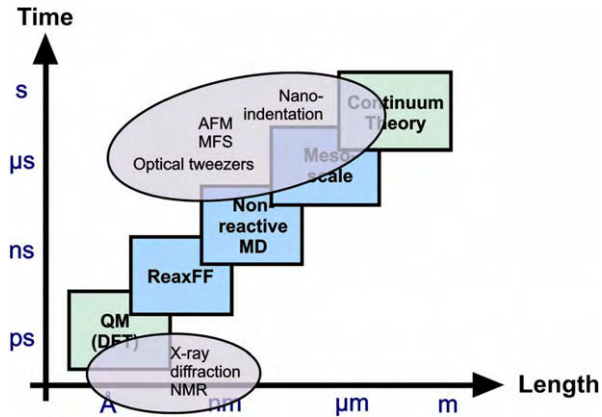
Fig. 4. Schematic that illustrates the concept of hierarchical multi-scale modeling (adapted from [18]). Hierarchical coupling of different computational tools can be used to traverse throughout a wide range of length- and time-scales. Such methods enable to provide a fundamental insight into deformation and fracture phenomena, across various time- and length-scales. Handshaking between different methods enables one to transport information from one scale to another. Eventually, results of atomistic, molecular or mesoscale simulation may feed into constitutive equations or continuum models. While continuum mechanical theories have been very successful for crystalline materials, BPMs require statistical theories, e.g. the Extended Bell Theory [40]. Experimental techniques such as atomic force microscope (AFM), molecular force spectroscopy (MFS), nanoindentation or optical tweezers now overlap into atomistic and molecular approaches, enabling the direct comparison of experiment and simulation [11,12,18,38,138].

1.2. Cross-scale interactions in biological protein materials through materiomics

A central theme of the efforts in developing materiomics is to appreciate the structure–property or structure–processing–property paradigm. This paradigm has guided materials science for many decades. For biological materials, there remain several challenges that make developing these rigorous links rather difficult and that cause for new analysis paradigms.

For example, bond energies in biological materials are often comparable to the thermal energy, as for instance in the case of hydrogen bonding, the most abundant chemical bond in biology. Biological materials often show highly viscoelastic behavior, since their response to mechanical deformation is intrinsically time-dependent. In many cases, biological structures contain extremely compliant filaments, in which entropic contributions to free energy are important and can even control the deformation behavior (e.g. elasticity and strength of hydrogen bonded protein domains [40]). Many material properties are further length scale dependent and can vary significantly. This poses grand challenges for the characterization of biological materials, as well as the comparison between different analysis approaches (e.g. simulation vs. experiment), since measuring different volumes of material lead to different material properties (e.g. elastic modulus, strength, as well as associated deformation mechanisms). Size dependent material properties are possibly crucial to understand the role of materials in physiological processes, and thus this area of research poses an exciting frontier of research. The presence of hierarchical structures calls for new paradigms in thinking about the structure–property paradigm, since corresponding concepts must include an explicit notion of the cross-scale and inter-scale interactions [40,41].

It has become evident that the atomistic scale, and in particular the notion of a chemical bond, provides a very fundamental, universal platform of materiomics at which a variety of scientific disciplines can interact: chemists, through the molecular structure of proteins, physicists, through the statistical mechanics of a large number of atoms, and materials scientists through analysis of phenomena such as elasticity, optical properties, electrical properties or thermodynamics, linking structure and function. Fig. 5 illustrates this concept. A noteworthy aspect of the materials science of biological materials is that it is interdisciplinary, by nature. Performing research in this field thus often implies to overcome barriers between scientific disciplines and to develop strategies that enable us to communicate each
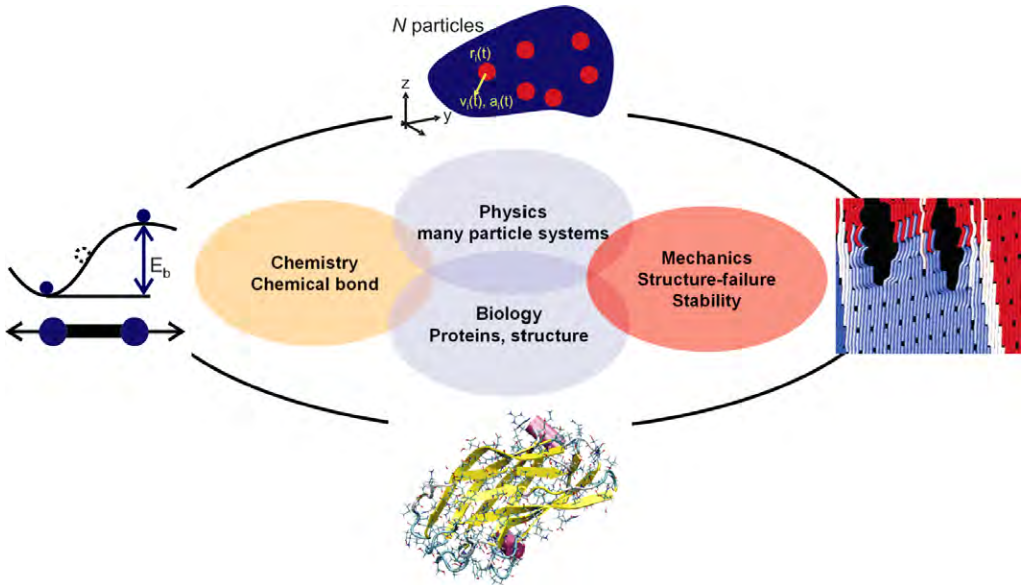
Fig. 5. Chemistry is the most fundamental "language" of materials science of protein materials. A variety of disciplines can link up with the notion of a chemical bond that defines the structure and eventually the properties of materials.

other's concepts more clearly. Structures in universities and research institutions may have to be modified to facilitate these new emerging frontiers of materials research.

It is vital to overcome the barrier that currently separates the understanding at different length- and time-scales, through the development of new experimental synthesis and characterization methods, novel model systems and an enhanced appreciation for a multi-scale view of materials in general, in order to fully understand multi-scale or cross-scale interactions in materiomics. To facilitate these developments, we must also develop a proper nomenclature to capture the various scales involved in a material. Current terminologies referring to atomistic, meso, micro and macro are insufficient to capture the subtleties of the various scales and structures observed in biological protein materials. Research should address the opportunities in integrating nanoscience and nanotechnology into biological research. What could our impact be, in the long-term perspective, in understanding some of the fundamental material concepts of biology?

For instance, is the nanomechanics of protein materials significant for biology, and may biologists have missed out on important effects due to a lack of consideration of the nanomechanics? Can we find inspiration from Nature that could guide us in the design of materials that are environmentally friendly, lightweight and yet tough and robust and can serve multiple objectives? How is robustness achieved in biology? How do universality and diversity of structural building blocks and mechanisms of material assembly or breakdown integrate into biological structures? From a theoretical viewpoint, major challenges include the development of new materials theories that include atomistic and statistical effects into an effective description, while retaining a system theoretical perspective of the overall material behavior [5,32,33,40–43]. This may eventually lead to a fruitful interaction between systems biology and materiomics.

Deformation mechanisms are a crucial element in understanding the response of materials to external stimuli. Similar to dislocation mechanics for metal plasticity, what is the theoretical framework for describing deformation and fracture of biological materials and elementary structures? It is possible that statistical theories may play a crucial role in the theoretical development of nanomechanics. Full-atomistic simulations of complex protein structures or large hierarchical protein assemblies with explicit solvents are often prohibitive due to computational limitations, and coarse-graining techniques must be used. In coarse-graining approaches, the full atomistic representation of protein
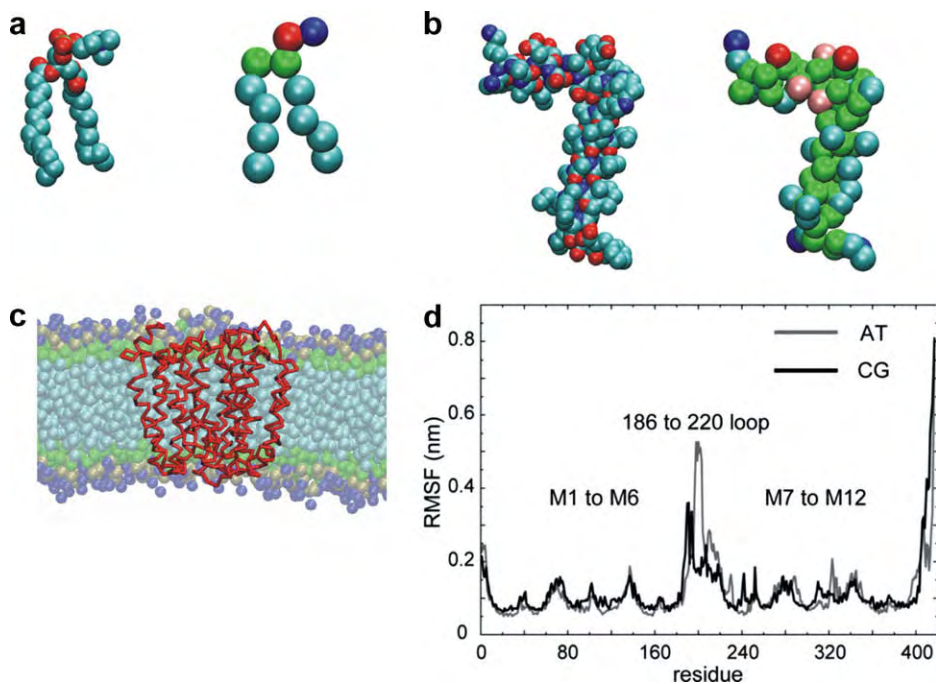
Fig. 6. Examples of a coarse-graining approach that replaces the full atomistic representation of protein structures. Thereby, small groups of atoms are treated as single particles. Subplots (a) [phosphatidyl choline] and (b) [coat protein from fd phage] show coarse-graining approaches for membrane lipid molecule structures [44]; the left part shows the full atomistic representation and the right part the corresponding coarse-grained model. In the atomistic representation, atoms are colored using the CPK convention. In the coarse-grained models the particles are colored according to the properties of the specific bead types (for details see [44]). Subplot (c) shows the results of a coarse-grained simulation, showing the protein (red backbone trace) in the lipid bilayer at the end of the 200 ns simulation. Subplot (d) shows the backbone particle root mean square fluctuation (RMSF) as a function of residue number, for both the atomistic and the coarse-grained simulation. Reprinted with permission from Ref. [44]. Copyright © 2007 Elsevier Ltd. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

structures is replaced by a model in which small groups of atoms are treated as individual (single) particles. Such methods represent a promising approach in analyzing key materials phenomena in biology focused on the mesoscale. Fig. 6 shows an example application of such a coarse-graining technique (see also Fig. 4, reference to "Mesoscale"). Fig. 6a and b shows coarse-graining approaches for membrane lipid molecule structures as reported in [44], where the left part shows the full atomistic representation and the right part the corresponding coarse-grained model. Fig. 6c shows the results of a coarse-grained simulation, showing the protein (red backbone trace) in the lipid bilayer at the end of the 200 ns simulation. Fig. 6d shows the backbone particle root mean square fluctuation (RMSF) as a function of residue number, for both the atomistic and the coarse-grained simulation, illustrating that the two descriptions provide similar results. Whereas coarse-graining works well for some systems, the method is not generally applicable and must typically be developed carefully for specific cases. Based on these challenges, it has been discussed how effective are coarse-graining techniques, overall? In general, can we indeed average out over atomistic or mesoscale structures to find equivalent representations for entire groups of atoms or are the atomistic details crucial for phenomena observed at larger length- and time-scales? How important are atomistic features at macroscale? What are the most appropriate numerical strategies to simulate the role of solvent (e.g. water) in very small confinement, where it can no longer be represented by continuum-type theories? How does confined water in small pores within and between protein constituents influence the mechanical response of natural and biological materials?

Progress in these various challenging fields will probably first occur specific to problems and applications, perhaps in those have most impact in medical or economic fields. Eventually, we must generalize our insight into the formulation of a holistic theoretical framework that transcends the current nomenclature, theoretical and experimental thinking. These efforts may provide the scientific and engineering fundamentals to develop and maintain the infrastructures to enable and evolve modern civilization. Through materiomics, materials – and materials science – could play a seminal role in these developments.

1.3. Outline

This review article provides an overview over advances in the field of deformation and fracture mechanics of biological protein materials, focusing on three model systems. The three model systems encompass the most abundant, universal building blocks of all protein materials: alpha-helices, beta-sheets and tropocollagen molecules. In each section, the article contains a broad review of works in this field. However, there is a particular focus on a summary of results reported by our group.

Section 2 provides an introduction into the numerical modeling techniques, in particular atomistic simulation methods. Section 3 provides a review of studies of collagenous tissues, including bone. Particular focus is on the mechanical properties and the properties of individual molecules and collagen fibrils. Section 4 is dedicated to a review of analyses of alpha-helical protein structures, focusing on vimentin intermediate filament protein structures. Section 5 is focused on the analysis of the mechanical properties of beta-sheet protein structures. The article concludes in Section 6 with a discussion and an outlook.

2. Numerical simulation techniques and theoretical framework

2.1. Atomistic and molecular modeling – introduction

Atomistic molecular dynamics (MD) is a tool for elucidating the atomistic mechanisms that control deformation and rupture of chemical bonds at nano-scale, and to relate this information to macroscopic material deformation and failure phenomena (see, e.g. references [18,26,27,40,44–46,93], and recent articles from our group that describes large-scale MD simulation of brittle fracture mechanisms [47–51]). The basic concept behind atomistic simulation via MD is to calculate the dynamical trajectory of each atom in the material, by considering their atomic interaction potentials, by solving each atom's equation of motion according to $F = ma$ (force equals mass times acceleration). Numerical integration of this equation by considering appropriate interatomic potentials enables one to simulate a large ensemble of atoms that represents a larger material volume, albeit typically limited to several nanoseconds of time-scale. The availability of such potentials for a specific material is often a limiting factor for the applicability of this method.

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 as an alternative approach to methods such as Monte Carlo, with the difference that MD provides full dynamical information and deterministic trajectories, a crucial element in determining dynamical details of material deformation and failure phenomena. It is emphasized that Monte Carlo schemes provide certain advantages as well; however, this point will not be discussed further here as all simulation studies reviewed here are carried out with a MD approach (for a textbook introduction into MD simulation, please see [109]). The total energy of the system is written as the sum of kinetic energy (K) and potential energy (U),

$$E = K + U, \tag{1}$$

where the kinetic energy is

$$K = \frac{1}{2} m \sum_{j=1}^{N} v_j^2, \tag{2}$$

and the potential energy is a function of the atomic coordinates r_j (where $j = 1, \ldots, N$ refers to the collection of all particles in the system),

$$U = U(r_j), \tag{3}$$

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

$$m\frac{\mathrm{d}^2 r_j}{\mathrm{d}t^2} = -\nabla_{r_j} U(r_j), \quad j = 1, \ldots, N, \tag{4}$$

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 (the equation below is the basic Verlet algorithm [109]):

$$r_j(t_0 + \Delta t) = -r_j(t_0 - \Delta t) + 2r_j(t_0) + a_j(t_0)(\Delta t)^2 + \cdots \tag{5}$$

The forces and accelerations are related by $a_j = f_j/m$. The forces are obtained from the potential energy surface – typically also called force field or interatomic potential– as

$$F_j = m\frac{\mathrm{d}^2 r_j}{\mathrm{d}t^2} = -\nabla_{r_j} U(r_j), \quad j = 1, \ldots, N. \tag{6}$$

This technique can also be used for not only single atoms but also groups of atoms as in the case of coarse-grained mesoscale approaches. Due to its fundamental formulation MD is capable of simulating a great variety of materials phenomena, for instance the response of an atomic crystal lattice to applied loading under the presence of a crack-like defect, or the deformation mechanisms of biomolecules including nucleic acids and proteins. Unlike many continuum mechanics approaches that are based on a top-down approach, atomistic techniques require no *a priori* assumption on the deformation or defect dynamics and thus provides a bottom-up description of materials. Once the atomic interactions are chosen through definition of the potential energy surface, the complete material behavior is determined. Choosing appropriate models for interatomic interactions provides a rather challenging and crucial step that remains subject of a very active discussion in the scientific community. A variety of interatomic potentials are used in the studies of biological materials at different scales, and different types of protein structures require the use of different atomistic models, some of which will be discussed below. A drawback of atomistic simulations is the challenge of analyzing results (which consist of only positions and velocities of particles as a function of time) and the large computational resources necessary to perform such simulations. Due to computational limitations, MD simulations are restricted with respect to the time-scales that can be reached, limiting overall time spans in such studies to tens of nanoseconds, or in very long simulation studies to fractions of microseconds. Therefore, many MD simulation results have been carried out at large deformation rates, exceeding several meter/second. Efforts have been reported to extend the time-scales accessible to MD to make quantitative links with experimental scales.

In this article, we review the application of several atomistic based modeling methods for three types of materials presented: collagenous tissues, alpha-helices and beta-sheets, in order to illustrate recent advancements and approaches of modeling that can be used to answer key questions regarding deformation and fracture mechanics of hierarchical biological materials.

2.2. Simulation and modeling approaches

Here we provide a brief review of interatomic force fields and modeling approaches suitable for simulating the behavior of protein structures. We refer the reader to more extensive review articles for additional information, in particular regarding force field models [52,53].

2.2.1. CHARMM force field
The basis for most studies reviewed in this article is the classical force field CHARMM [52–55], implemented for example in the MD program NAMD [56]. The CHARMM force field [56] (and

extensions, e.g. to model tropocollagen molecules [57]) is widely used in the protein and biophysics community, and provides a reasonable description of the behavior of proteins. This force field is 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 bonds (H-bonds). Since the bonds between atoms are modeled by harmonic springs or its variations, bonds (other than H-bonds) between atoms cannot be broken, and new bonds cannot be formed. Also, the charges are fixed and cannot change, and the equilibrium angles do not change depending on stretch. The CHARMM force field belongs to a class of models with similar descriptions of the interatomic forces; other models include the DREIDING force field [58], the UFF force field [59], or the AMBER model [52,53,60].

In the CHARMM model, the mathematical formulation for the empirical energy function that contains terms for both internal and external interactions has the form:

$$U(\vec{R}) = \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{UB}} K_{\text{UB}}(S - S_0)^2 + \sum_{\text{angle}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\chi(1 + \cos(n\chi - \delta))$$
$$+ \sum_{\text{impropers}} K_{\text{imp}}(\phi - \phi_0)^2 + \sum_{\text{nonbond}} \varepsilon \left[\left(\frac{R_{\min(i,j)}}{r_{ij}} \right)^{12} - \left(\frac{R_{\min(i,j)}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\varepsilon_1 r_{ij}}, \tag{7}$$

where K_b, K_{UB}, K_θ, K_χ and K_{imp} are the bond, Urey–Bradley, angle, dihedral angle and improper dihedral angle force constants, respectively; b, S, θ, χ and ϕ are the bond length, Urey–Bradley 1,3-distance, bond angle, dihedral angle and improper torsion angle, respectively, with the subscript zero representing the equilibrium values for the individual terms. Fig. 7 shows a schematic of some of the individual energy contributions listed in Eq. (7).

The Coulomb and Lennard–Jones 6–12 terms contribute to the external or non-bonded interactions; ε is the Lennard–Jones well depth and $R_{\min(i,j)}$ is the distance at the Lennard–Jones minimum, q_i is the partial atomic charge, ε_1 is the effective dielectric constant and r_{ij} is the distance between atoms i and j. The parameters in such force fields are often determined from more accurate, quantum chemical simulation models by using the concept of force field training [39] (see the hierarchical coupling schematic shown in Fig. 4).

Force fields for protein structures typically also include simulation models to describe solvents, in particular water molecules (e.g. TIP3 water), an essential part of any simulation of biological protein structures [52,53].

2.2.2. Reactive force fields: a new bridge to integrate chemistry and mechanics

Reactive force fields represent an important milestone in overcoming the limitations of classical force fields: the lack of the ability to describe rupture and formation of covalent bonds. This is because
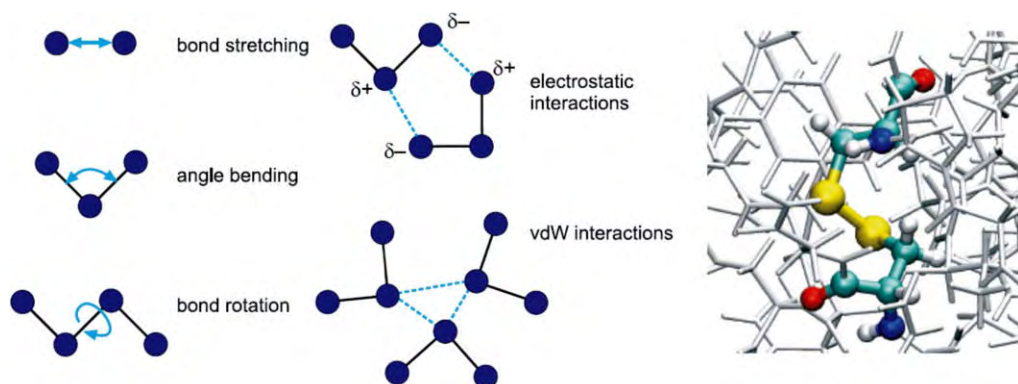
Fig. 7. Model of the individual energy contributions due to bond stretching, bond bending, bond rotation as well as electrostatic and vdW interactions. The combination of these terms constitutes the entire energy landscape of interatomic and intermolecular interactions, as given in Eq. (7).
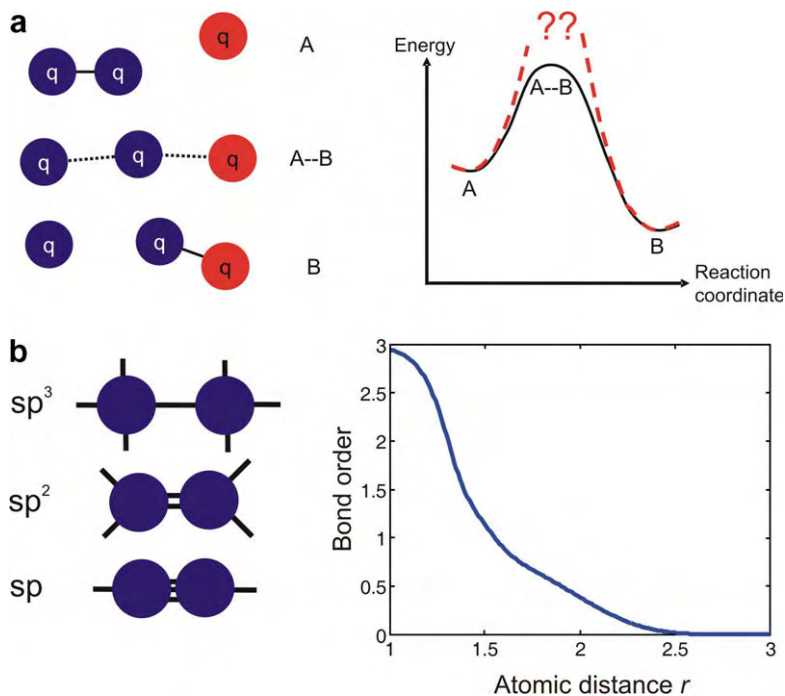
Fig. 8. Basic concept of reactive versus nonreactive force fields (figures adapted from Refs. [61,72]). Subplot (a): The reactive force fields enable describing the transition state energies between two ground states, as it is important for example during chemical reactions, including bond rupture. Nonreactive force fields (dashed line in the right subplot) are only capable of describing the ground states A and B, but not the transition state A–B. Subplot (b) illustrates the concept of bond orders, here in an example for a C–C bond. Depending on the distance between atoms, different bond orders are obtained, through a bond order mapping function. This enables to distinguish different quantum chemical states such as sp^3 (single bond), sp^2 (double bond) and sp (triple bond). The continuous change of bond orders as a function of distance ensures that reactive force fields are energy continuous, which is critical to carry our constant energy simulations. At large distances, the bond order vanishes, indicating breaking of the covalent bond. In ReaxFF, the spring constant that characterizes the strength of atomic bonding is modulated by the bond order, leading to vanishing bond strength or dissociation at large stretch.

the covalent bond terms (see Eq. (7), first term "bonds") are described using harmonic terms, (see Eq. (7), first term "bonds") which do not provide an accurate description of the bond energetics at large bond stretch. For mechanical properties of materials (that is, the large deformation and rupture mechanisms), this translates into the properties of molecules at large strain, a phenomenon referred to as hyperelasticity. Fig. 8a illustrates this effect, explaining how nonreactive force fields are not capable of describing transition state energies during bond formation and rupture correctly.

Several flavors of reactive potentials have been proposed in recent years [61–63]. 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 [5,61–63], have been extended recently to cover a wide range of materials, including metals, semiconductors [50,51,64–72] and organic chemistry in biological systems such as proteins [61,72,73]. Here we focus in particular on the ReaxFF formulation [61,69,72,73]. In some of the studies reviewed in this article, a particular flavor of the ReaxFF potentials as suggested in [73], with slight modifications to include additional QM data suitable for protein modeling is used.

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, as shown in Fig. 8b. All connectivity-dependent interactions (that means, valence and torsion angles) are formulated to be bond order dependent. This
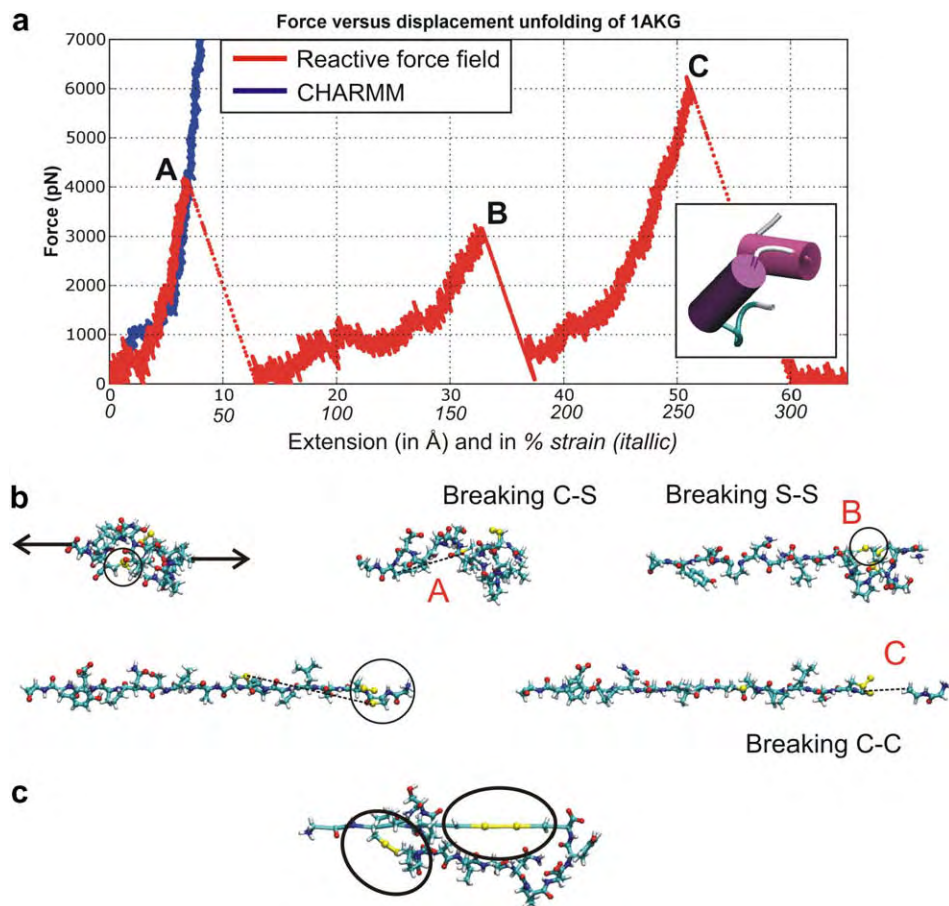
Fig. 9. Force–extension profile (unfolding of a small protein α-conotoxin PnIB from *conus pennaceus* (PDB ID 1AKG)), comparing a nonreactive CHARMM model (blue curve) with the reactive ReaxFF model (red curve) (figure adapted from Ref. [72]. It is apparent that although the ReaxFF and CHARMM descriptions agree for small deformation (below ≈7Å), 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. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

ensures that their energy contributions disappear upon bond dissociation so that no energy disconti-
nuities 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 (QEq) scheme that accounts for polarization effects and modeling of charge flow, assigning
a partial charge to each atom at each integration step. This is considered a critical advance leading to a
new bridge between QM and empirical force fields. All interactions feature a finite cutoff distance, and
all interactions are tapered off smoothly at the cutoff distance.

Further discussion regarding reactive force fields for proteins can be found elsewhere [72], includ-
ing examples that illustrate the differences between nonreactive and reactive models in describing
protein unfolding. Here we only review a few basic results. For instance, conventional nonreactive
models are not able to describe rupture of disulfide cross-links between different protein domains
(exemplified for a lysozyme protein structure). In contrast, the reactive model is capable of describing
these processes. This has major implications on the resulting force–extension curves, as shown in
Fig. 9. This figure shows the force–extension profile during unfolding of a small protein α-conotoxin

PnIB from *conus pennaceus* (PDB ID 1AKG), comparing a nonreactive CHARMM model (blue curve) with the reactive ReaxFF model (red curve) [72]. Clearly, the CHARMM model cannot capture the bond breaking events at large deformation, and deviates significantly from the ReaxFF description. Similar studies have been carried out for unfolding of lysozyme, as presented in Ref. [72]. The ReaxFF force field has been integrated with other simulation methods and force fields in the Computational Materials Design Facility (CMDF) [50,51,74] to provide a computationally efficient approach in integrating multiple simulation paradigms.

2.3. Steered molecular dynamics

To apply the forces to the molecule that induce deformation, steered molecular dynamics (SMD) has evolved into a useful tool [75]. Steered MD is based on the concept of adding a harmonic moving restraint to the center of mass of a group of atoms in order to apply external forces to induce deformation of molecular structures. This leads to the addition of the following potential to the Hamiltonian of the system:

$$U(r_1, r_2, \ldots, t) = \frac{1}{2}k[vt - (\vec{X}(t) - \vec{X}_0) \cdot \vec{n}]^2, \tag{8}$$

where $\vec{X}(t)$ is the position of restrained atoms at time t, $\vec{X}_0$ denotes original coordinates and v and $\vec{n}$ denote pulling velocity and pulling direction, respectively. The net force applied on the pulled atoms is $F(x_1, x_2, \ldots, t) = k(vt - (\vec{X}(t) - \vec{X}_0)) \cdot \vec{n}$. By monitoring the applied force (F) and the position of the atoms that are pulled over the simulation time, it is possible to obtain force versus displacement data that can be used to derive the mechanical properties such as bending stiffness or the Young's modulus (or other mechanical properties). SMD studies reviewed here are typically carried out with a spring constant $k = 10$ kcal/mol/Å^2. The SMD method mimics an AFM nanomechanical loading experiment, as illustrated in Fig. 10.
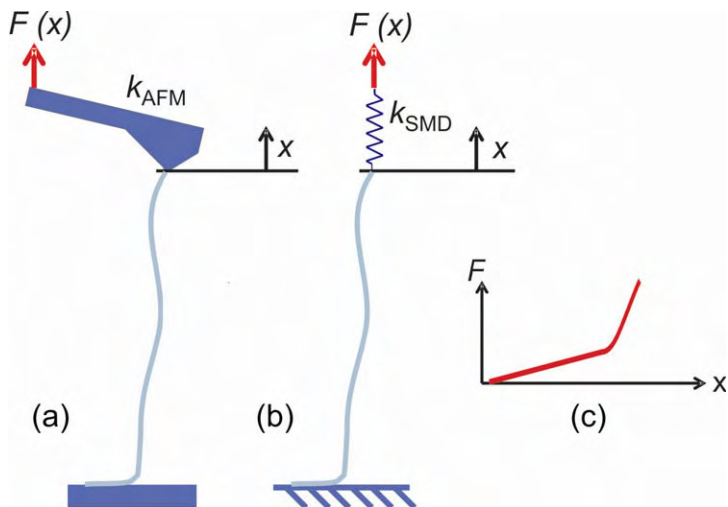
Fig. 10. Single molecule pulling experiments, carried out on a single protein molecule [18] (schematic based on Ref. [216]). Subplot (a) depicts an experimental setup based on AFM, and subplot (b) depicts a steered molecular dynamics (SMD) analogue. In the SMD approach, the end of the molecule is slowly pulled. This leads to a slowly increasing force over the displacement, as schematically shown in subplot (c). Both approaches, AFM and SMD lead to force–displacement information. In addition to the $F(x)$ curve, SMD provides detailed information about associated atomistic deformation mechanisms. Due to the time-scale limitations of MD to several nanoseconds, there is typically a large difference between simulation and experiment with respect to pulling rates. Whereas MD simulations are limited to pulling rates larger than 0.01 m/s, experimental rates are six to eight magnitudes smaller than those. This requires additional consideration in order to interpret MD results in light of experimental findings.

2.4. Large-scale parallelized computing

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 composed of hundreds of individual computers or processors that combined form an entity referred to as supercomputer. Supercomputers exceed the capabilities of ordinary PCs or laptops by several orders of magnitudes.

Whereas computing power was estimated to plateau at the gigaflop level, the broad availability of teraflop computers is now expected by the middle or end of the current decade. 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. Recent examples have demonstrated the applicability of such large-scale computing methods to solve complex problems with billions of particles [77]. Information between these small computers is shared by communicating message "packages", which is achieved by message-passing procedures, enabled via software libraries such as the "Message Passing Interface" (MPI) [76,77]. Implemented based on spatial domain decomposition in which each computational entity (=processor) solves only part of the entire computational problem, 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) [76,77]. Each processor works on a part of the computational domain referred to as a "computational cell". Fig. 11 depicts the historical development of computational power over several decades, illustrating the emergence of petaflop computers in the next few years.

With a parallel computer whose number of processors increases with the number of computational cells (the number of atoms, or more general, particles per computational cell does not change), the computational burden per processor remains constant. To achieve this, the computational space is divided up into computational cells such that in searching for neighbors interacting with a given atom or particle, only the computational 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 and more atoms or particles [77]. Alternative approaches, such as computing on Graphical Processing Units (GPUs), provide additional opportunities for extremely high performance. The concept of such approaches is to take advantage of particularly designed processing units for great efficiency in carrying out floating point
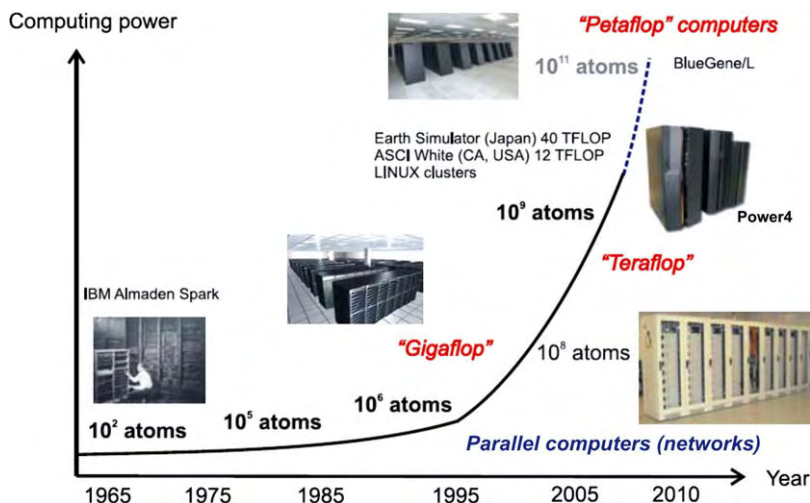
Fig. 11. Development of computing power over the past decades. The development illustrates the emergence of petaflop computers in the next few years. The plot also summarizes the number of atoms that can be treated with these computing systems; these numbers are developed for simple interatomic potentials with short cutoffs. For CHARMM or ReaxFF potentials, the number of atoms is significantly smaller.

operations to enable ultra-fast integration of the equations of motion of particles, tailored to the needs of MD.

2.5. Data analysis and visualization methods

Data analysis and visualization of complex data sets is a crucial element in a successful MD study, and its value cannot be underestimated. This is particularly true for ultra-large system sizes of millions to billions of particles, which cannot be visualized and understood by conventional approaches. An important quantity in the analysis of mechanical properties of materials is the Cauchy stress tensor. The virial stress tensor can be used to calculate the Cauchy stress tensor directly from atomistic data [79,80,93]. The reby, atomistic data is averaged over all particles (that is, to obtain a spatial average), and over several snapshots (that is, to provide a temporal average). This approach is used in Section 3, collagenous tissues, for instance, to follow the increase in stress as lateral strain is applied to collagen fibrils. For details regarding the calculation of the virial stress tensor we refer the reader to the literature [79,80]. An alternative method to calculate the stress for filamentous structures is to divide the applied force by the cross-sectional area, $\sigma = F/A$ (this could be used, for instance, in conjunction with the SMD approach that provides force-extension data). This approach is analogous to the engineering definition of stress as it is used in the study of tensile tests, for instance.

Visualization plays a crucial rule in the analysis of MD simulation results, as the raw data resulting from such numerical simulations represents merely a collection of positions, velocities and accelerations as a function of time. In particular, structural features and patterns of proteins (and how they change under deformation) are difficult to analyze. To address this point, many visualization tools have been developed that are capable of displaying biological protein molecules and clusters. A versatile, powerful and widely used visualization tool is the visual molecular dynamics (VMD) program [78], a tool closely linked to the NAMD molecular simulation code [54]. This software enables one to render complex molecular geometries using particular coloring schemes. It also enables us to highlight important structural features of proteins by using a simple graphical representation, such as alpha-helices, or the protein's backbone. The simple graphical representation of proteins is often referred to as cartoon model. These visualizations are often the key to understand complex dynamical processes and mechanisms in analyzing the motion of protein structures and protein domains, and they represent a filter to make useful information visible and accessible for interpretation.
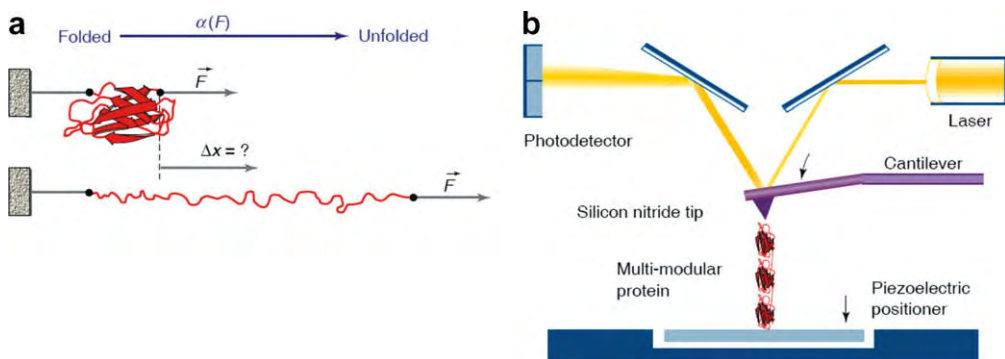
Fig. 12. The unfolding of protein domains by an external force. Subplot (a): When axial stress is applied to a folded domain the protein will unravel. Subplot (b): The force–extension mode of the atomic force microscope (AFM). When pressed against a layer of protein attached to a substrate, the silicon nitride tip can adsorb a single protein molecule. Extension of a molecule by retraction of the piezoelectric positioner results in deflection of the AFM cantilever. This deflection changes the angle of reflection of a laser beam striking the cantilever, which is measured as the change in output from a photodetector. Reprinted with permission from Ref. [302]. Copyright © 1999 Elsevier Ltd.
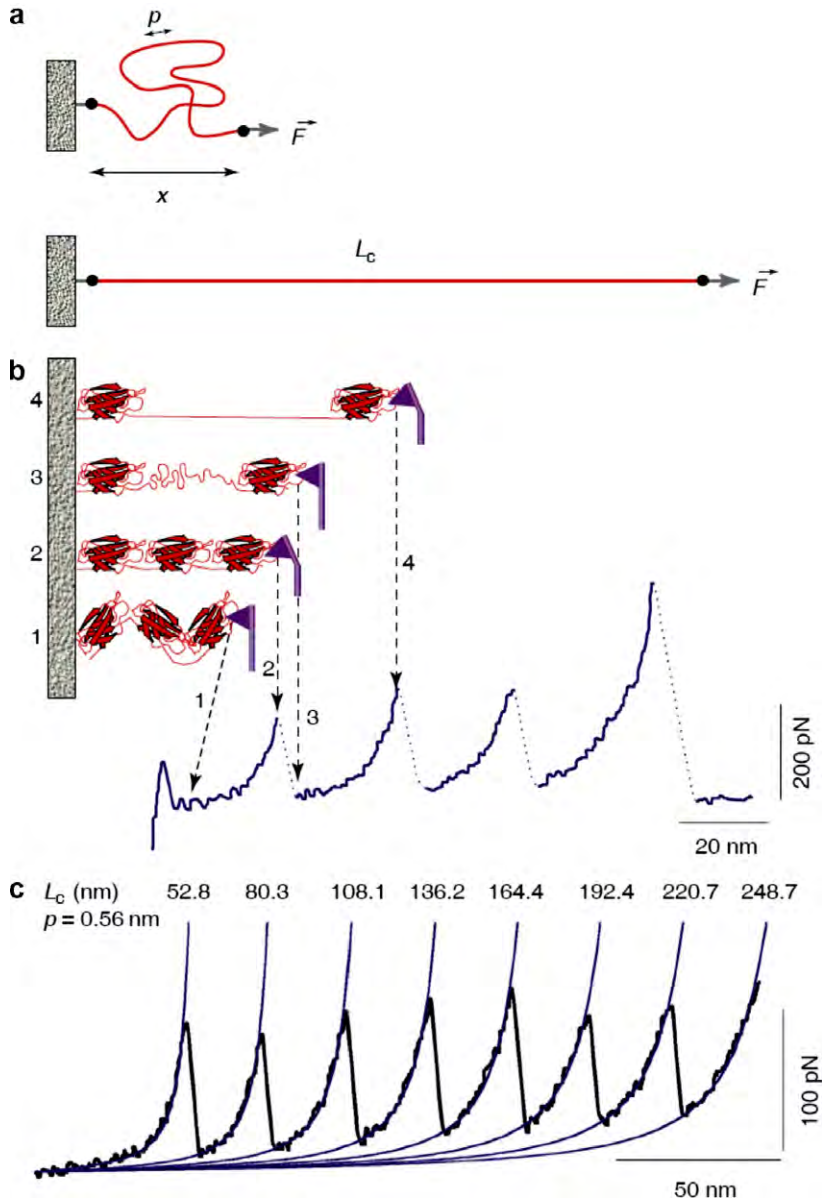
Fig. 13. The entropic elasticity of proteins and protein domain unfolding. Subplot (a): Schematic to show the geometry of the protein stretching experiment. Subplot (b): The characteristic saw-tooth pattern of peaks that is observed when force is applied to extend the protein corresponds to sequential unraveling of individual domains of a modular protein. As the distance between substrate and cantilever increases (from state 1 to state 2) the protein elongates, generating a restoring force that bends the cantilever. When a domain unfolds (state 3) the free length of the protein increases, returning the force on the cantilever to near zero. Further extension again results in force on the cantilever (state 4). The last peak represents the final extension of the unfolded protein prior to detachment from the AFM tip. Subplot (c): Consecutive unfolding peaks of recombinant human tenascin-C, fitted using the WLC model. The contour length for each of the fits is shown; the persistence length was fixed at 0.56 nm. Reprinted with permission from Ref. [302]. Copyright © 1999 Elsevier Ltd.

2.6. Experimental methods to probe nanomechanical properties

Recent advances in experimental techniques further facilitate analyses of ultra-small-scale material behavior. For instance, techniques such as nanoindentation, optical tweezers, or atomic force microscopy (AFM) can provide valuable insight to analyze the molecular mechanisms in a variety of materials, including metals, ceramics and proteins. The mechanical signature of proteins and other single biomolecules can be obtained by AFM, where the biomolecule is attached to a surface and manipulated by a cantilever that pulls the molecule at constant force (see Fig. 10). Fig. 12 depicts AFM experiments on protein structures. A saw-tooth shaped force–displacement profile is commonly observed and linked to sequential unfolding of certain domains in the protein. Worm-like chain model (WLC) [43,81] is frequently used to describe the entropic elasticity of these domains, as shown in Fig. 13. We refer the reader to other articles regarding details of these experimental approaches (see, e.g. [10,35,36,38,82–85]). A selection of experimental techniques is summarized in Fig. 4, illustrating the overlap with multi-scale simulation methods.

An important experimental technique in conjunction with atomistic modeling of protein materials is X-ray diffraction; results of such experiments provide the initial atomistic and molecular structure, the starting point for all atomistic simulations. The structure of many proteins, elucidated using such experiments, has been deposited in the Protein Data Bank [86].

Since recent advances in experimental methods now enable one to probe time- and length-scales that are also directly accessible to large-scale atomistic based simulation, the combination of experiment and simulation might lead to a particularly fruitful interaction. This is particularly promising since the kind of information obtained from experiment and simulation might be complementary.

3. Collagenous tissue

This section describes hierarchical multi-scale modeling of collagenous tissues, with a particular focus on the mechanical properties. Studies focus on elastic behavior, plastic behavior and fracture. Starting at the atomistic scale, we review the development and application of a hierarchical multi-scale model that is capable of describing the dynamical behavior of a large number of tropocollagen molecules, reaching length-scales of several micrometers and time-scales of tens of microseconds.

A particular emphasis is given on elucidating the deformation mechanisms that operate at various scales, the scale dependent properties, the effect of specific hierarchical features and length-scales (for example, cross-link densities, intermolecular adhesion). Through mineralization of a collagen template, the collagen fibril structure also forms the basis for bone, representing a composite structure of primarily protein and mineral. To investigate the role of these processes and the influence on the material properties, the effect of addition of mineral platelets during the early stages of formation of new bone will be discussed as well. We conclude with a discussion of how insight of nanomechanical behavior at the smallest scales relate with the physiological role of collagen.

3.1. Introduction

Collagen, the most abundant protein on earth, is a fibrous structural protein with superior mechanical properties, and provides an intriguing example of a hierarchical biological nanomaterial. The hierarchical structure of collagen is summarized in Fig. 14 (see also the schematic shown in Fig. 2). Collagen plays an important role in many biological tissues, including tendon, bone, teeth, cartilage or in the eye's cornea. Severe mechanical tensile loading of collagen is significant under many physiological conditions, as in joints and in bone [7,8,87–89,91,108].

In this section we review the elastic, plastic and fracture behavior of collagen fibrils, linking atomistic scale studies with the mesoscopic level of collagen fibrils. In particular the focus is on the large deformation behavior of collagen based tissues, which is particularly important under physiological conditions and during injuries. The studies reviewed here explain the limiting factors in strength of collagen fibrils, as well as the origins of its toughness. These investigations complement experimental efforts focused on the deformation mechanics of collagen fibril at nano-scale, including
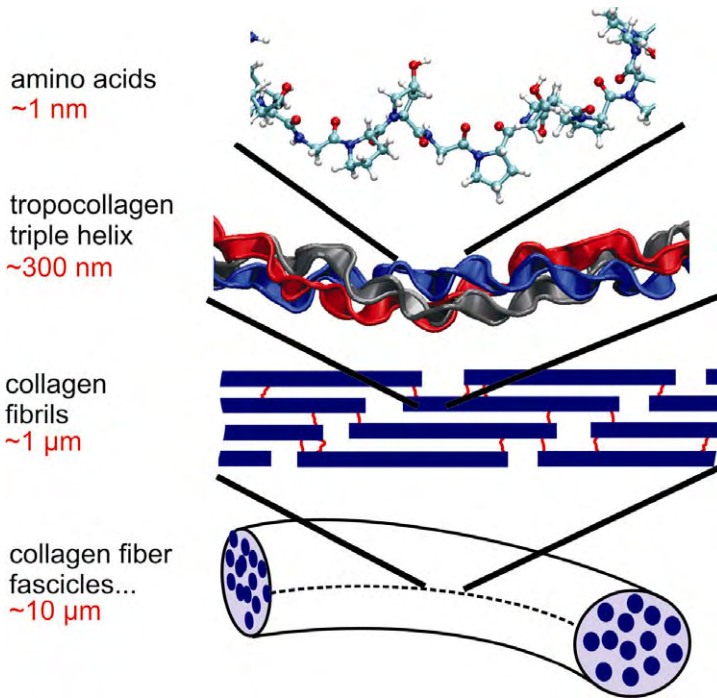
Fig. 14. Schematic view of some of the hierarchical features of collagen, ranging from the amino acid sequence level at nanoscale up to the scale of collagen fibers with lengths on the order of 10 μm (figure adapted from Refs. [26,27]). The studies reviewed in this article are focused on the mechanical properties of individual tropocollagen molecules and collagen fibrils, consisting of a staggered array of TC molecules. The red lines in the graph indicate intermolecular cross-links that are primarily developed at the ends of tropocollagen molecules. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

characterization of changes of *D*-spacing (which defines the particular staggered geometry of collagen fibrils) and fibril orientation [89], analyses that featured X-ray diffraction [92] and synchrotron radiation experiments [92] (Fig. 15).

3.2. Atomistic and molecular structure of collagenous tissues

Collagen consists of tropocollagen molecules that have lengths of $L \approx 300$ nm with approximately 1.5 nm in diameter, leading to an aspect ratio of close to 200 [1,8,27,30,100]. Staggered arrays of tropocollagen molecules form fibrils, which arrange to form collagen fibers (Fig. 14 and Fig. 15 for experimental studies of the assembly process).

Each tropocollagen molecule consists of a spatial arrangement of three polypeptides. These three molecules or polypeptides are arranged in a triple helical structure, stabilized primarily by H-bonding between amino acid residues. Every third residue in each of these molecules is a GLY amino acid, and about one fourth of the tropocollagen molecule consists of proline (PRO) and hydroxyproline (HYP) [30]. The structure of collagen has been known since classical works focusing on theoretical understanding of how tropocollagen molecules are stabilized [30]. This triple helical protein structure, also known as the "Madras triple helix", was originally proposed by G.N. Ramachandran in 1954 [30]. Fig. 16 shows a ball and stick diagrams showing two projections of the currently accepted triple helical structure for collagen with one inter-chain hydrogen bond per tripeptide.

Recently, various types of tropocollagen molecules have been crystallized and analyzed using X-ray diffraction techniques to determine their precise atomic configuration [94]. TEM experiments have
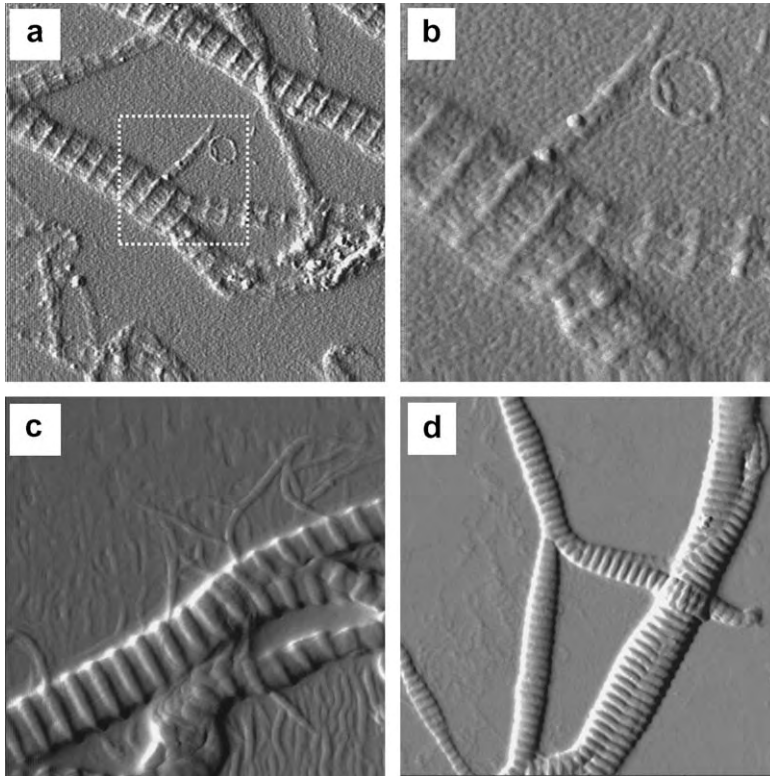
Fig. 15. AFM analyses of the assembly process of collagen fibrils, showing different stages of the assembly process (subplots (a)–(d)). The plots show the characteristic D-banding structure due to the staggered arrangement of individual tropocollagen molecules into collagen fibrils, visible by the regularly spaced lines. Reprinted with permission from Ref. [337]. Copyright © 2002 Elsevier Ltd.

also been used to study the structure of collagen in various environments, including in bone [3,4], in particular focusing on larger length scale features and its three-dimensional arrangement.

Collagen is also the most fundamental building block of bone, providing additional evidence for the great significance of collagen as a structural protein material. Bone has evolved to provide structural support to organisms, and therefore, its mechanical properties are of great physiological relevance. A total of seven hierarchical levels are found in bone [3,4]. The smallest scale hierarchical features of bone include the protein phase composed of tropocollagen molecules, collagen fibrils (CF) as well as mineralized collagen fibrils (MCFs). Tropocollagen molecules assemble into collagen fibrils in a hydrated environment, which mineralize by formation of hydroxyapatite (HA) crystals in the gap regions that exist due to the staggered geometry (see Fig. 14).

3.3. Deformation and fracture of single tropocollagen molecules

This section is focused on the nanomechanical properties of single tropocollagen molecules, as originally reported in [26,29,36,95–99]. This approach in describing collagenous tissues based on a multi-scale atomistic-based model represents a bottom-up approach. The development begins by focusing on the finest, atomistic scales governed by quantum mechanics (QM) as the starting point, reaching up to larger, mesoscopic and macroscopic scales, using hierarchical multi-scale modeling methods. The first step in achieving this goal is the careful analysis of the properties of a single molecule.
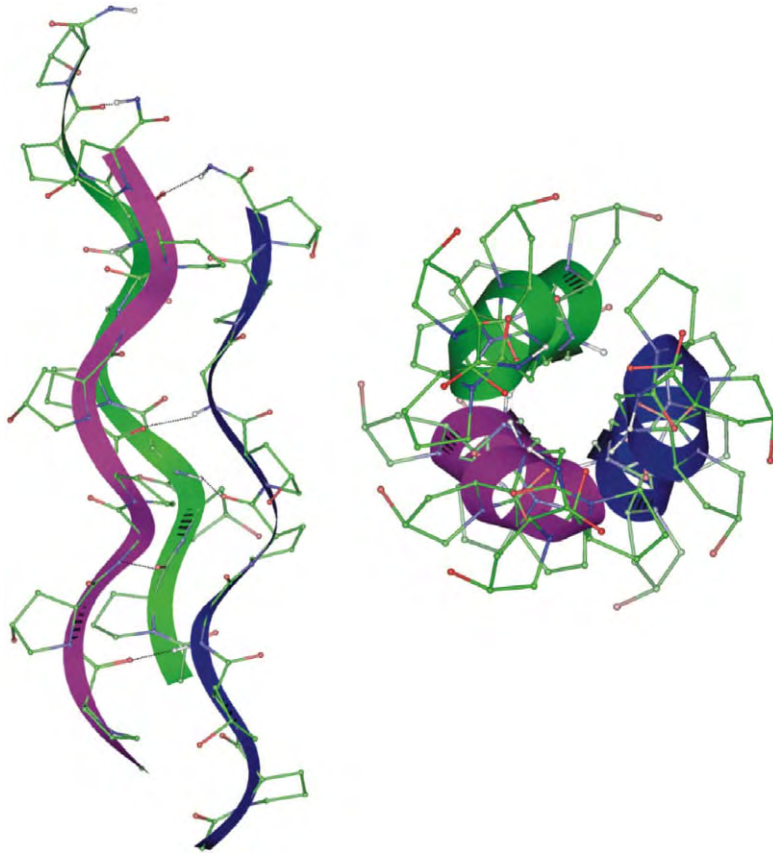
Fig. 16. Ball and stick diagrams showing two projections of the currently accepted triple helical structure for collagen with one inter-chain hydrogen bond per tripeptide, the "Madras triple helix" originally proposed by G.N. Ramachandran [30]. The sequence shown is (Gly-Pro-Hyp)$_3$ and each chain in the triple helix has a different color ribbon drawn through the backbone. Reprinted with permission from Ref. [130].

There are several reports of experimental studies focused on the mechanics of single collagen molecules [29,36,85], providing us with a basis for validation of MD simulation results. Despite its relatively simple structure, few studies of single tropocollagen molecules using molecular dynamics (MD) studies have been reported in the literature. In one of the first articles published, Lorenzo and co-workers [101] have reported investigations of the mechanical properties of tropocollagen molecules, using MD studies, focusing on their Young's modulus. Other studies focused on the stability of collagen molecules [102,103] and other structural investigations [104–106], or the effect of point mutations on the stability of tropocollagen molecules [103]. Some researchers modeled collagen at the continuum scale, using techniques such as FEM [107]. In a recent study, molecular models of collagen have been linked to collagen mechanics [56]. Other studies focused on the small- and large-deformation and fracture mechanics of collagen molecules, as reported in [26,110,111].

Questions of particular interest include: how does a tropocollagen molecule respond to mechanical stretching force, in particular at large deformation? How does it fracture, and what are the force levels at the molecular scale that lead to fracture? How can these properties be linked to the folded triple-helical structure of collagen molecules? How do ultra long collagen molecules with realistic lengths of several hundred nanometers behave in solvents and under applied forces? Such insight is important to understand the role of individual tropocollagen molecules in the context of tissue mechanics, disease and injury.

Table 1 provides an overview over important mathematical symbols used throughout this section.

Table 1

Description of the main parameters and material or molecular properties used in the studies of collagenous tissues (Section 3)

F_{tens}	Tensile strength of a bimolecular fibril
A_c	Cross-sectional area of a TC molecule
L_0	Length of an individual TC molecule
L	Length of a TC molecule in a collagen fibril
L_C	Contact length between different TC molecules (e.g. in a bimolecular assembly or in a collagen fibril)
α	Overlap parameter in an assembly of TC molecules, note that $\alpha = L_C/L$
τ_{shear}	Shear strength between two TC molecules (units: force/length)
σ_R	Critical molecular tensile stress to nucleate slip pulse
E	Young's modulus, e.g. of an individual TC molecule or a collagen fibril
σ_{tens}	Tensile stress in a TC molecule (note that $\sigma_{tens} = F_{tens}/A_c$)
F_{max}	Maximum tensile force a single TC molecule can sustain
F_F	Maximum tensile force a BM collagen fibril can sustain
χ_S	Critical molecular length scale beyond which slip pulse nucleation occurs
χ_R	Critical molecular length scale beyond which fracture occurs
L_χ	Critical molecular length scale at which maximum strength is reached
γ	Energetic barrier to nucleation of a slip pulse (units: energy per length^2)
E_{diss}	Energy dissipation during deformation (units: energy)
$E_{diss,V}$	Energy density dissipation during deformation (units: energy per volume)

Units are provided for some of the variables for which the definition deviates from the standard definition.

3.3.1. Computational atomistic model

The tool used at the finest scale of individual tropocollagen molecules is MD modeling, as widely utilized to model the behavior of materials and molecules [109] (as described in Section 2.1). The atomistic model discussed here is built based on the crystal unit cells according to X-ray diffraction data obtained by experiment [100]; a short tropocollagen segment is solvated in a water skin. These structures are taken directly from the Protein Data Bank (PDB). We review studies that are based on the crystal structure PDB ID 1QSU, with 1.75 Å resolution, as reported by Kramer et al. [100]. The 1QSU is a triple-helical collagen-like molecule with the sequence (Pro-Hyp-Gly)$_4$-Glu-Lys-Gly-(Pro-Hyp-Gly)$_5$, representing a model system to enable MD simulation studies.

The charges of each atom are assigned according to the CHARMM rules. Hydrogen atoms are added according to pH 7. The CHARMM input files (structure and topology files) are then used to perform MD stimulations. For ReaxFF calculations, no atom typing is necessary since only element types are assigned, and charges are determined dynamically during the simulation through the QEq method. Hydrogen atoms are added using the NAMD/CHARMM procedure, according to the same conditions as outlined above.

Before finite temperature dynamical calculations are performed, an energy minimization is carried out, thereby making sure that convergence is achieved, thus relieving any potential overlap in vdW interactions after adding hydrogen atoms. In the second step, the molecule is annealed after heating it up to a temperature $T = 300$ K, annealing it. The heating rate is 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 ensemble [54,109]). It is also ensured 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, mechanical forces are then applied using various types of constraints, and the response of the molecule due to the applied loading is investigated. Typically, force versus displacement data is obtained from such simulations, 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. Unless indicated otherwise, we use a steered molecular dynamics (SMD) scheme with spring constant $k = 10$ kcal/mol/Å^2. It was shown in earlier reports that this is a reasonable choice leading to independence of the measured molecular mechanical properties from the choice of the SMD spring constant [101].
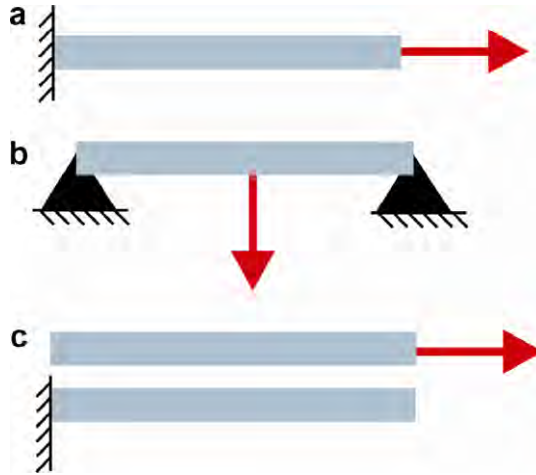
Fig. 17. Overview over various load cases studied, as reported in [26]. The load cases considered here include (a) tensile loading in axial direction of the molecule (also including compressive test), (b) bending test, (c) and shear test.

3.3.2. Tensile and bending deformation of single tropocollagen molecules

Here we review a series of studies that encompass the mechanical deformation of single tropocollagen molecules under different boundary conditions (for an overview over all cases considered, see Fig. 17). The load cases considered here include (i) tensile loading in axial direction of the molecule (also including compressive test), (ii) a bending test, (iii) and a shear test between pairs of tropocollagen molecules. The goal of these studies is to provide a fundamental understanding of the nanomechanical properties of individual molecules (and assemblies of two molecules). These results from full atomistic modeling form a so-called training set for a hierarchical multi-scale modeling approach (see schematic overview depicted in Fig. 3, where the concept of "finer scales train larger scales" is explained).

3.3.2.1. Small tensile deformation. First we review tensile testing of single tropocollagen molecules using the nonreactive CHARMM force field. After careful equilibrium (as far as the time-scales are accessible to the MD method, which is typically limited to a few nanoseconds) of the structure of the molecule, a force is applied at one end, while the other end of the molecule is kept fixed. The loading case is shown in Fig. 17a. By slowly increasing the load applied to the collagen molecule, while measuring the displacement d, we compute force–displacement curves. The force versus displacement curve $F(d)$ can be used to determine a stress versus strain curve, by proper normalization:

$$\sigma(d) = \frac{F(d)}{A_C}, \tag{9}$$

where A_C denotes an equivalent area of the cross-section of a collagen molecule (see below for further details how this quantity is estimated), and $A_C = \pi \cdot R^2 \approx 214.34 \text{ Å}^2$, assuming that $R = 8.26$ Å (this has been obtained from studies of an assembly of two tropocollagen molecules, as described below). Note that the stress is typically dependent on the stretch d. The local (in terms of strain) Young's modulus $E(d)$ is given by

$$E(d) = \frac{d_0}{A_C} \frac{\partial F(d)}{\partial d}, \tag{10}$$

where d_0 is the initial, undeformed length of the collagen fiber, and $d_0 = 84$ Å. Note that Young's modulus is independent of the length of the molecule. The definition in Eq. (10) is a consequence of the fact that the stretching force is expressed as a function of stretch d rather than strain ($\sigma = E\varepsilon$).
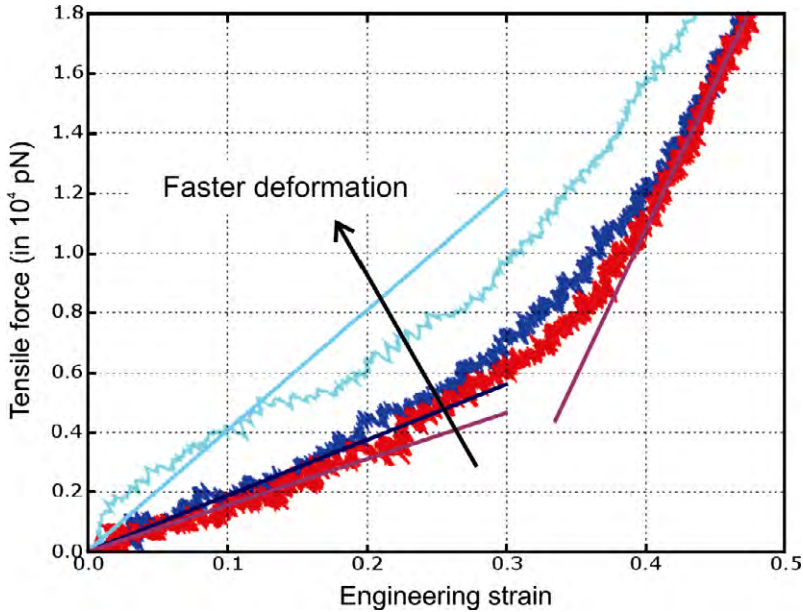
Fig. 18. Force vs. strain, pulling of a single tropocollagen molecule, for three different pulling velocities, as reported in [26]. The results indicate that faster pulling velocities lead to a stiffening effect (the loading rates in the three cases are $\dot{r}_\lambda = 0.0001$, 0.0002 and 0.001 Å/step).

Fig. 18 shows force versus displacement plots for tensile loading, for three different loading rates. The loading rates in the three cases are $\dot{r}_\lambda = 0.0001$, 0.0002 and 0.001 Å/step. Young's modulus is determined as the tangential slope corresponding to 10% tensile strain. Young's modulus is obtained to be $E_{tens} = 6.99$, 8.71 and 18.82 GPa, for increasing loading rates as specified above. An increase in stiffness for higher loading rates is observed, where a linear extrapolation to vanishing rates leads to a modulus of 5.9 Gpa. These results indicate that tropocollagen molecules show a rate dependent elastic response, at least at higher strain rates. Fig. 19 depicts snapshots of the tropocollagen molecule under increasing stretch. At large strains, the helical structure is lost and the three polypeptides appear as individual strands, then defining its elasticity by the behavior of covalent bonds. This is also confirmed in the force–strain plot (Fig. 18). The local Young's modulus associated with large strains is given by $E_{tens}^{large} = 46.7$ GPa. More recent atomistic simulation studies have shown that the tensile Young's modulus of single tropocollagen molecules converges to a constant value close to 4 GPa for loading rates below 0.1 m/s [110]. These results will not be discussed further, and the reader is referred to the literature for additional information.

3.3.2.2. Large deformation and fracture mechanics of individual tropocollagen molecules. In the next step, the focus is on the mechanical behavior of collagen fibers using a reactive force field, which is capable of describing the elastic behavior (and rupture mechanics) of covalent bonds under large deformation. The dominating forces in tropocollagen molecules occur typically in the axial direction of the molecule, so this type of loading is most critical for the mechanical integrity.

The key questions addressed here include: under which conditions do classical, nonreactive force fields fail to describe the molecular mechanics appropriately, and what are the limitations of these methods? Do the results agree for small deformation? How are the mechanical properties different once mechanical deformation is large and formation and breaking of bonds are allowed? We will pull the tropocollagen molecule until fracture occurs, and study the details of the fracture mechanisms. A central question we would like to address is, to which strain and deformation levels – or, equivalently, applied force level – can we rely on the assumption of nonreactive force fields. A solid understanding
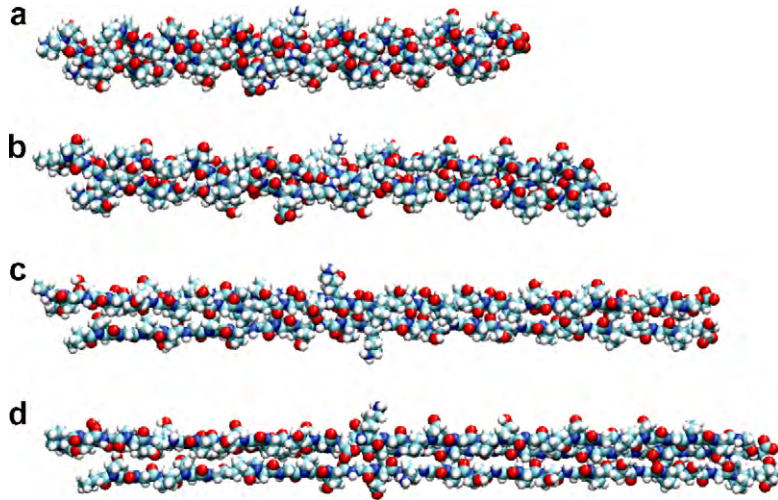
Fig. 19. Stretching of a single tropocollagen molecule, using a nonreactive CHARMM force field (water molecules are not shown for clarity) [26]. The helical structure unfolds with increasing strain and vanishes at large deformation. The three strands become independent, and the covalent bonds between atoms govern the elasticity. This loss in tertiary helical structure is represented by a change in tangent elasticity, as seen in Fig. 9, at strains beyond 35%.

of the large deformation, nonlinear and fracture properties of tropocollagen molecules becomes important when the mesoscale model is introduced at the next hierarchical level.

Fig. 20 depicts snapshots of results obtained from ReaxFF simulations of tensile deformation of tropocollagen molecules. Fig. 21 shows plots of force versus strain, comparing the CHARMM model (curve (a)) with the ReaxFF model (curve (b)). The strength of the tropocollagen molecule is determined to be 23,000 pN, reached at approximately 50% tensile strain. It is found that the CHARMM description and the ReaxFF model agree for small strains up to approximately 10% strain. The tensile
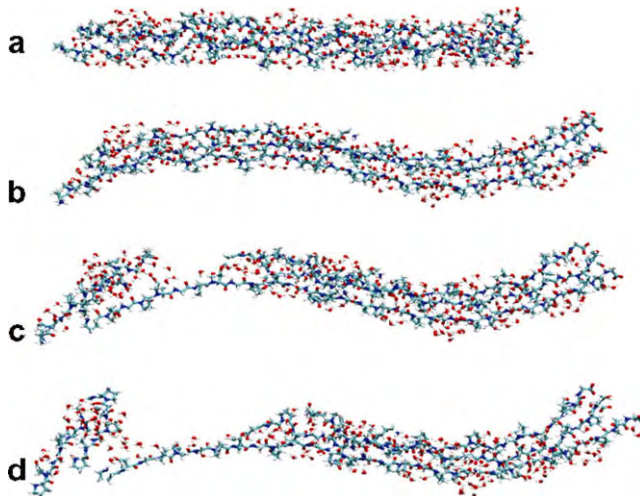
Fig. 20. Fracture mechanics of a single tropocollagen molecule, as reported in [26]. We observe a transformation from the initial straight shape to an S-shaped structure at large strain, leading to fracture of an individual polypeptide. This transformation is found consistently for a variety of loading conditions.
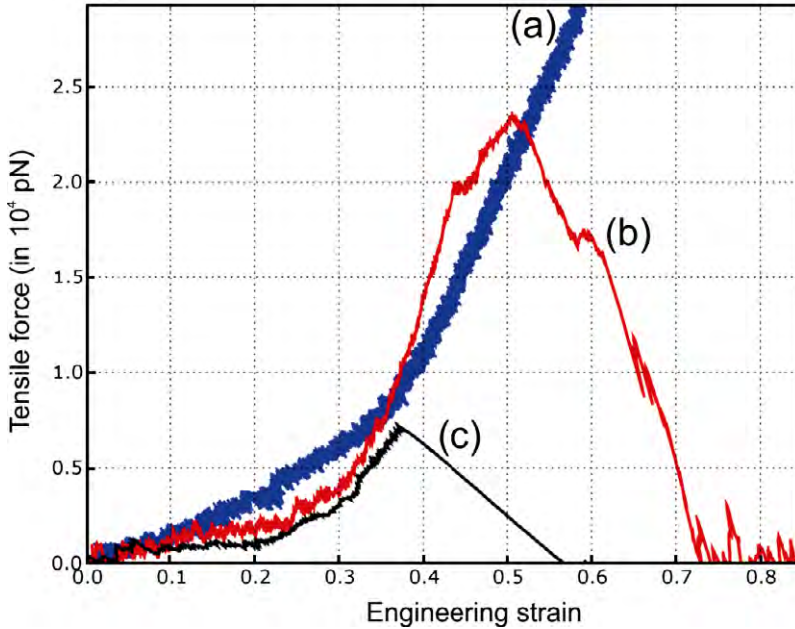
Fig. 21. Force vs. strain, pulling of a single tropocollagen molecule, as reported in [27]. An initial regime of flat, almost linear-elastic extension is followed by onset of nonlinear, stiffening behavior at larger strains beyond approximately 30–35% strain. This behavior can be explained by the fact that the tertiary, helical structure of the tropocollagen molecule begins to disappear and the elasticity of each covalent bond in the single strand governs the elastic response. This represents a significant distinction to the results obtained with reactive force fields, suggesting failure due to strain/force localization.

stress for onset of permanent deformation under rupture of covalent bonds (at around 43% strain) occurs at approximately 9.3 GPa, and the fracture stress is determined to be 11.2 GPa [26].

3.3.2.3. Bending a single tropocollagen molecule. A computational experiment is performed to describe the bending of a single tropocollagen molecule by clamping it at the boundaries and applying a force in the center of the molecule, as shown in Fig. 17b. This is equivalent to a three-point bending test that is widely used in engineering. From the force–displacement data obtained by atomistic modeling, the bending stiffness EI is given by

$$\text{EI} = \frac{dL^3}{48F_{\text{appl}}}, \tag{11}$$

where F_{appl} is the applied force, and d is the bending displacement. Fig. 22 depicts load versus displacement curves for various deformation speeds. Fig. 23 depicts the resulting bending stiffness as a function of loading speed (these results have been reported in [111]).

3.3.3. Shearing two tropocollagen molecules

We continue with a review of shearing experiments of an assembly of two tropocollagen molecules. The objective is to gain insight into the deformation mechanisms and type of interactions between two tropocollagen molecules in aqueous solution. A geometry as depicted schematically in Fig. 17c serves as the starting point for the investigations. The system is first equilibrated without application of any mechanical shear load. It is found that the equilibrium distance between two molecules depends on the presence of solvent; being reduced if solvent is present. With solvent present, it is found that the equilibrium distance between two tropocollagen molecules is $r_{\text{EQ}} \approx 16.52$ Å. All studies reported here are carried out with water molecules present. It is noted that this equilibrium dis-
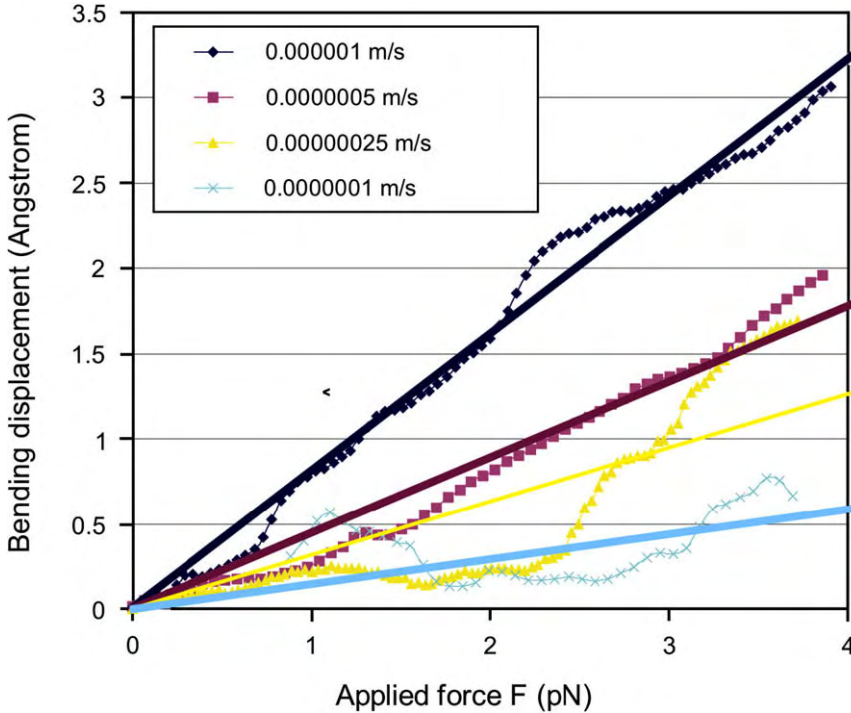
Fig. 22. Bending displacement over bending force, for the three-point bending test of a single tropocollagen molecule, as reported in [111]. Results are shown for various loading rates. The linear curves are linear fits to the MD simulation results.
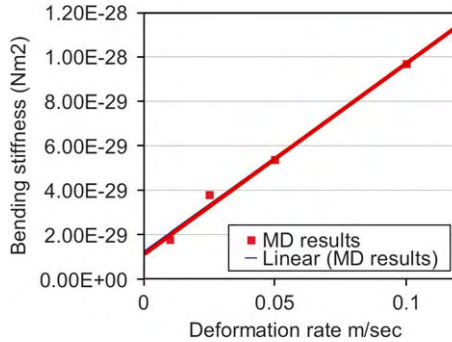
Fig. 23. Dependence of bending stiffness EI of a single TC molecule on the deformation rate (based on results shown in Fig. 15) [111]. The results indicate that EI decreases linearly with decreasing loading rate. A linear fit to the data enables us to extrapolate to smaller deformation speeds.

tance leads to a molecular radius of 8.26 Å. Fig. 24 depicts snapshots of the system as it undergoes shear deformation. Fig. 25 depicts the maximum shear force versus pulling velocity as obtained during the shearing experiment.

These shear experiments have been performed at three different loading rates, $\dot{r}_\lambda = 0.0002, 0.00005$ and 0.000025 Å/step. The resulting values are strain rate dependent, and the maximum force decreases with decrease in loading rate, assuming values $F_{\text{max,shear}} \approx 2900, 1100$ and 750 pN,
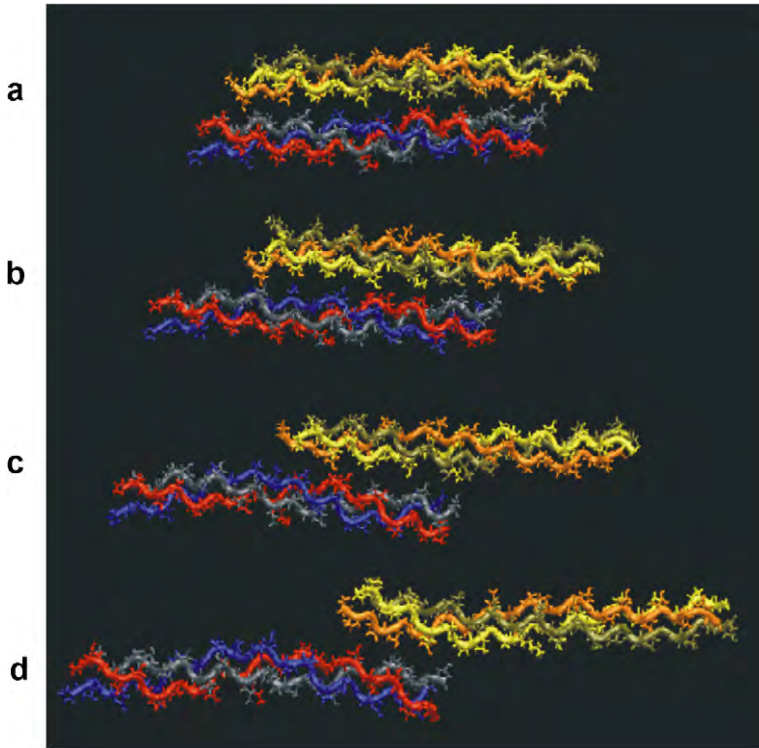
Fig. 24. Snapshots of shearing of two tropocollagen molecules [26]. The subplots (a)–(d) show the behavior as the shear strain is increased.
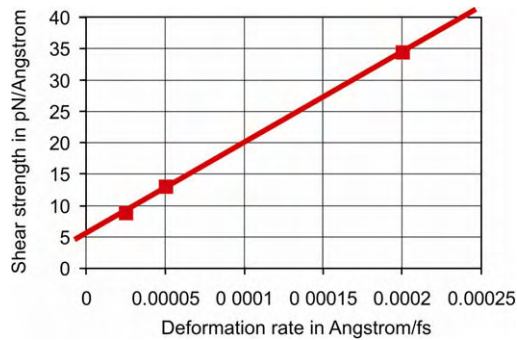
Fig. 25. Atomistic modeling of shear experiments between two TC molecules (shear load applied using the SMD method) [26]. The plot shows the shear resistance as a function of loading rate, obtained by fully-atomistic modeling using the CHARMM potential.

corresponding to the strain rates provided above. Using a linear extrapolation to vanishing loading rate, the maximum shear force is estimated to be $F_{max,shear} \approx 466$ pN, corresponding to an adhesion strength of $\tau_{max} = 5.55$ pN/Å (see Fig. 25, point where the linear curve intersects the y-axis). It is noted that the units of τ_{shear} are force/length (this is not a "stress" but rather adhesion force per unit length).
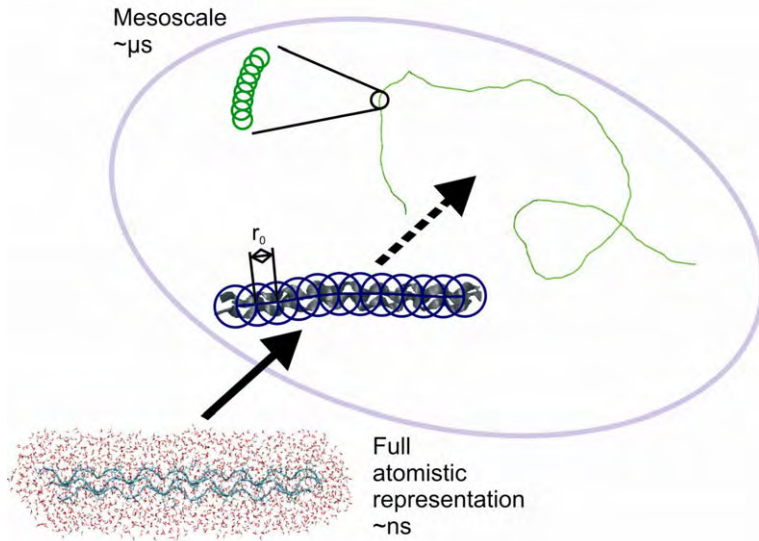
Fig. 26. Schematic showing developing of the coarse-grained molecular model from a full atomistic description, illustrating the procedure utilized in various studies reviewed in this chapter. The full atomistic representation of the triple helical TC structure is replaced by a collection of beads. The mesoscale model enables the treatment of ultra long TC molecules at time- and length-scales not in reach by full atomistic models.

3.3.4. Development of a mesoscopic, molecular model

The atomistic modeling results, carried out at the level of individual polypeptide chains, tropocollagen molecules and assemblies of those helped to develop a better qualitative and quantitative understanding of the competing mechanisms and forces during deformation of collagen, at a microscopic level. This information is now used to develop a mesoscopic model, in which beads connected by different types of springs represent collagen molecules, whereas all parameters are completely derived from atomistic calculations. The motivation for these studies is the desire to model larger length- and time-scales. This approach is similar to training empirical potentials from quantum mechanical data using the force matching approach (see, for instance [112]). The geometrical approach of coarse-graining is visualized in Fig. 26.

The reduction of degrees freedom in the mesoscale model enables one to study very long tropocollagen molecules with lengths on the order of several hundred nanometers, as well as bundles of tropocollagen molecules. This approach enables reaching the mesoscale, and makes the overall mechanics of the material accessible to atomic and molecular scale modeling (see also Fig. 6 for a schematic of how other coarse-graining methods are applied).

3.3.4.1. Model development: training from pure atomistic results using energy and force matching. The goal is to develop the simplest suitable model possible to perform large-scale studies of the mechanics of collagen molecules, eventually leading to understanding of the behavior of assemblies of such fibers. We assume that we can write the total energy of the system as

$$E = E_{\text{T}} + E_{\text{B}} + E_{\text{weak}}. \tag{12}$$

The bending energy is given by

$$\phi_{\text{B}}(\varphi) = \frac{1}{2} k_{\text{B}} (\varphi - \varphi_0)^2, \tag{13}$$

with k_{B} as the spring constant relating to the bending stiffness. The resistance to tensile load is characterized by

$$\phi_T(r) = \frac{1}{2} k_T (r - r_0)^2, \tag{14}$$

where k_T refers to the resistance of the molecule to deform under tensile load. To account for the non-linear stress–strain behavior of a single molecule under tensile loading, we replace the harmonic approximation with a bilinear model. The force between two particles is

$$F_T(r) = -\frac{\partial \phi_T(r)}{\partial r}, \tag{15}$$

where

$$\frac{\partial \phi_T(r)}{\partial r}(r) = H(r_{\text{break}} - r) \begin{cases} k_T^{(0)}(r - r_0) & \text{if } r < r_1 \\ k_T^{(1)}(r - \tilde{r}_1) & \text{if } r \geqslant r_1 \end{cases}. \tag{16}$$

In Eq. (16), $H(r - r_{\text{break}})$ is the Heaviside function $H(a)$, which is defined to be zero for $a < 0$, and one for $a \geqslant 0$, and $k_T^{(0)}$ as well as $k_T^{(1)}$ for the small and large deformation spring constants. The parameter $\tilde{r}_1 = r_1 - k_T^{(0)}/k_T^{(1)}(r_1 - r_0)$ is determined from force continuity conditions. The function E_T is given by integrating $F_T(r)$ over the radial distance. In addition, weak, dispersive interactions between either different parts of each molecule or different molecules are assumed, defined by a Lennard–Jones (LJ) function

$$\phi_{\text{weak}}(r) = 4\varepsilon \left(\left[\frac{\sigma}{r} \right]^{12} - \left[\frac{\sigma}{r} \right]^{6} \right), \tag{17}$$

with σ as the distance parameter, and ε describing the energy well depth at equilibrium. It is noted that the total energy contribution of each part is given by the sum over all pair-wise and triple (angular) interactions in the system, therefore,

$$E_I = \sum_{\text{pairs}} \phi_I(r) \quad \text{and} \quad E_B = \sum_{\text{angles}} \phi_B(\varphi). \tag{18}$$

3.3.4.2. Equilibrium distances of beads and corresponding masses. The mass of each bead is determined by assuming a homogeneous distribution of mass in the molecular model. The total mass of the tropocollagen molecule used in our studies is given by 8152.2 amu. The total length of the tropocollagen molecule used in the full atomistic MD studies is divided into $N_{\text{MD}} = 6$ pieces, thus each bead containing five amino acid residues. Each bead then has a weight of 1358.7 amu. Since the total length of the molecule is $L_0 = 84$ Å, the beads are separated by a distance $r_0 = 14$ Å (for a finer discretization of beads, say 7 Å, the mass will be half of this value). The beads represent different sequences in tropocollagen that, when added together, make up the entire sequence.

3.3.4.3. Dispersive and nonbonding interactions. The LJ parameters are determined from the calculations of shearing two collagen fibers. In all these considerations, we assume that a pair-wise interaction between different tropocollagen molecules is sufficient, and that there are no multi-body contributions. Based on these assumptions, we model the interactions between different molecules using a LJ 12:6 potential. The distance parameter σ is given by

$$\sigma = \frac{D}{\sqrt[6]{2}} \approx 14.72 \, \text{Å}, \tag{19}$$

where D is the equilibrium distance as measured in the MD simulations, $D = 16.52$ Å. The shear strength can be used to extract the LJ parameters for the weak, dispersive interactions between two fibers. Note that this interaction includes the effect of solvation water molecules and other intermolecular bonding (e.g. H-bonds, and other "nonbonding" interactions). The maximum force in a LJ potential, assuming a single LJ "bond" is given by

$$F_{\text{max,LJ}} = \frac{\Lambda \cdot \varepsilon}{\sigma}, \tag{20}$$

while noting that $\Lambda \approx 2.3964$ for the LJ 12:6 potential. The parameter σ is already determined, leaving only ε to be trained using a force matching approach. The parameter ε can be obtained by requiring a force balance at the point of rupture:

$$F_{\text{max,LJ}} N_{\text{MD}} = \tau_{\text{max}} \cdot L = F_{\text{max}}. \tag{21}$$

This expression can be used to determine ε as

$$\varepsilon = \frac{F_{\text{max}} \cdot \sigma}{\Lambda \cdot N_{\text{MD}}}. \tag{22}$$

From atomistic modeling, we calculate F_{max} which allows for estimating numerical values for ε. We find that $\varepsilon \approx 11.06$ kcal/mol predicted from Eq. (22). Based on the extrapolation of shear force $F_{\text{max,shear}} \approx 466$ pN corresponding to vanishing strain rate we finally arrive at a value $\varepsilon \approx 6.87$ kcal/mol. The presence of intermolecular cross-links effectively lead to an increased intermolecular adhesion in the region where cross-links are formed, as discussed in [113–115].

To model the effect of cross-links, the adhesion parameter ε_{LJ} is modified to account for the stronger interaction between molecules. Variation of the parameter ε_{LJ} along the molecular axis enables one to account for specific spatial distributions of cross-links. Experimental analyses of the molecular geometry suggests that intermolecular aldol cross-links between lysine or hydroxylysine residues [1,8,113] primarily develop at the ends of tropocollagen molecules [114,115]. The aldol cross-link is a C–C bond that forms between side chains of residues of two tropocollagen molecules.

The presence of cross-links is modeled by increased adhesion at the ends of each molecule, in segments of 60 Å to the left and right end of each tropocollagen molecule. According to this idea, the LJ potential parameter ε_{LJ} is increased by a factor $\beta \geqslant 1$ compared with the rest of the molecule in regions where cross-links are formed, and therefore,

$$\varepsilon_{\text{LJ,XL}} = \beta \varepsilon_{\text{LJ}}. \tag{23}$$

For a choice $\beta = 12.5$, the additional shear force exerted at the end of the molecule corresponds to ≈ 4.2 nN, which is on the order of the bond strength of covalent cross-link bonds [26,116,117]. The parameter $\beta = 12.5$ therefore corresponds to the case when approximately one cross-link is present at each end of a tropocollagen molecule, leading to a cross-link density of 2.2×10^{24} m^{-3} (the cross-link density is defined as the number of cross-links per unit volume). Similarly, doubling the value $\beta = 25$ corresponds to two covalent cross-links.

Fig. 27 shows a schematic that illustrates the modeling of intermolecular cross-links via increased adhesion.
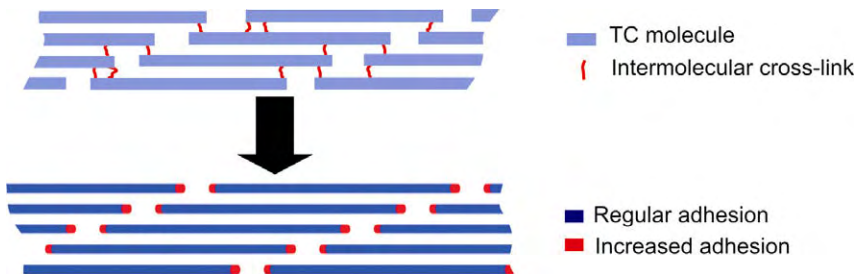
Fig. 27. Schematic showing how the presence of cross-links is modeled by increased adhesion at the ends of each molecule, in segments of 60 Å to the left and right of each tropocollagen molecule. Implementing a variation of the amplification of the adhesion strength constitutes a simplistic model for varying cross-link densities. A parameter β is introduced that describes the increase of adhesion at the ends of each TC molecule, so that $\tau_{\text{XL}} = \beta \cdot \tau$ (τ is the adhesion force/length between two TC molecules). The parameter $\beta = 15$ corresponds to the case when approximately one cross-link is present at each end of a tropocollagen molecule.

3.3.4.4. Tensile spring parameter. The tensile spring constant is determined from various calculations of stretch versus deformation, while being constrained to the regime of small loads and consequently small displacements. The spring constant k_T is then defined as

$$k_T = \frac{N_{MD}F_{appl}}{\Delta d} = \frac{A_c}{L_0}E, \tag{24}$$

with $\Delta d = L - L_0$ being the displacement of the atomistic model due to applied force F_{appl}. Based on the low-strain rate tensile testing data discussed earlier, we find that $k_T^{(0)} \approx 15.41$ kcal/mol/Å^2. Similar considerations can be used to determine a value for $k_T^{(1)}$, thereby considering the large deformation elastic behavior.

The parameters r_1 and r_{break} (unit: length) are related to the critical strains at which the tangent slopes in the stress–strain curve changes (denoted by ε_1, which is approximately 30%), and from the breaking strain (denoted by ε_{break}, which is approximately 50%):

$$r_1 = (1 + \varepsilon_1)r_0, \tag{25}$$

as well as

$$r_{break} = (1 + \varepsilon_{break})r_0. \tag{26}$$

3.3.4.5. Bending spring parameter. Using an argument of energy conservation between the atomistic and the mesoscale model, an expression for the bending stiffness parameter k_B is derived,

$$k_B = \frac{3}{2}\frac{EI}{r_0}. \tag{27}$$

Alternatively to using the energy matching approach, the force matching approach could be used as well. It is found that $k_B \approx 14.98$ kcal/mol/rad^2 as reported in [111]. These expressions are only valid for small deformations.

3.3.4.6. Computational validation of mesoscale model in tensile deformation. Fig. 28 shows the force–displacement curve of a tensile stretching experiments of tropocollagen molecules, comparing results
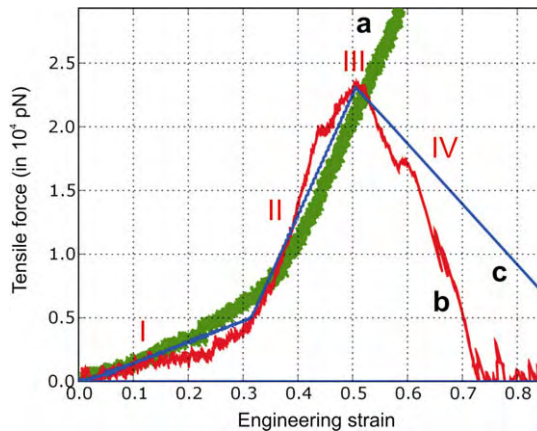
Fig. 28. Modeling of a tensile stretching experiment of a TC molecule, comparing atomistic and mesoscale model for validation (using the SMD method) (results as reported in [111]). The plot shows the force vs. displacement of a TC molecule, using the CHARMM force field (curve "a") and the reactive ReaxFF force field (curves "b" and "c"). Curve "b" shows the force–displacement response of a single TC molecule, displaying three regimes. Regime I is characterized by uncoiling of the TC molecule, regime II is associated with a larger modulus due to stretching of covalent bonds, and in regime III we observe fracture of the molecule, followed by a rapid decay of the force in regime IV. Curve "c" shows the results obtained using the reactive mesoscale model, illustrating the agreement between the different methods.

obtained with the CHARMM method, ReaxFF and the mesoscale model (for a short TC segment of approximately 8 nm length). The results confirm that the mesoscale model indeed approximates the results of reactive MD.

In particular, the mesoscale model enables us to model the characteristic four regimes of the deformation mechanics of a single TC molecule: regime I is characterized by uncoiling of the TC molecule (please, see also [110] for a detailed analysis), regime II is associated with a larger modulus due to stretching of covalent bonds, and in regime III fracture of the molecule is observed, followed by a rapid decay of the force in regime IV.

3.3.5. Stretching an ultra-long tropocollagen molecule: mesoscale modeling

The mesoscale model (development described in the sections above) now enables one to carry out studies aimed at a direct comparison with experimental results, validating the major predictions of the simulation model. The present section reviews molecular simulation results reported in [111] (for details regarding the modeling procedure see this reference).

Fig. 29 shows a computational stretching experiment obtained by using the mesoscale model, as reported in [111]. Loading of a single 300 nm long tropocollagen molecule embedded in solvent starts from a coiled entangled configuration of the molecule with end-to-end distance of approximately 100 nm. During the initial regime (I), the molecule loses its entangled structure, while the applied forces remain relatively low [110,111]. The elastic response of the molecule in this regime is due to entropic forces, and the behavior can be modeled for instance using the worm-like-chain (WLC) model [43,81]. According to the worm-like-chain (WLC) theory, the force is a function of end-to-end chain length, x, the persistence length ξ_p, and the contour length λ:

$$F(x) = \frac{k_B T}{4\xi_p} \left[\left(1 - \frac{x}{\lambda} \right)^{-2} + 4\frac{x}{\lambda} - 1 \right].$$

(28)

where k_B is the Boltzmann constant, and T is the temperature. Once the contour length is reached ($x = \lambda$), (II) uncurling of the triple helix, (III) stretching of covalent bonds in the individual polypeptides, and (IV) rupture of the tropocollagen molecule occurs, followed by a sharp drop of the forces to zero. The qualitative behavior of tropocollagen molecules under stretch is similar to recent experimental studies carried out on collagen fibrils that reach forces on the order of μN, also showing a significant hyperelastic stiffening effect [118].

Fig. 30 shows a zoom into the small force regime, providing a quantitative comparison of the MD results [111] with optical tweezers experiments [29,36,97]. The comparison is enabled by showing a direct overlay of experimental results and computational results. The plot reveals very good
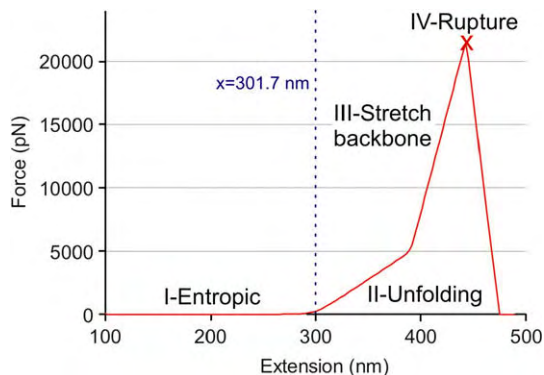
Fig. 29. Force–displacement ($F(x)$) curves of stretching a single TC molecule, $L = 301.7$ nm, at 300 K, as reported in [111]. The plot depicts the force–displacement curve over the entire deformation range, covering four stages: (I) uncoiling of the entangled configuration, (II) uncurling of the triple helix, (III) stretching of covalent bonds in the individual polypeptides, and (IV) rupture of the TC molecule. The dashed line of indicates the contour length of the molecule.
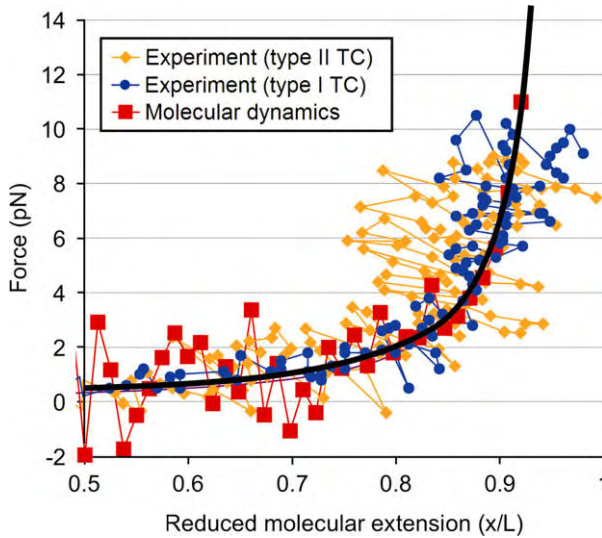
Fig. 30. This plot depicts a subset of the results depicted in the previous figure, focusing on the small force, entropic response (σ < 4 pN) (original results reported in [111]). This plot also depicts experimental results obtained for TC molecules with similar contour lengths (23,24), as well as the prediction of the WLC model with persistence length approximately 16 nm. Type I and type II collagen refer to the building blocks of different collagen tissues, displaying a different sequence of the tropocollagen molecules, leading to distinct structural at larger hierarchical scales of tissues. For instance, type I is predominantly found in skin, bone and tendon, whereas type II collagen forms cartilage tissues.

agreement, even though the deformation rate in MD is still higher. Fig. 31 shows experimental results of stretching of a single tropocollagen molecule using optical tweezers, as reported in Ref. [36]. For both cases, experiment and simulation, the WLC model is a reasonable description of the observed elastic response, with a similar persistence length on the order of approximately 15 nm [111].

3.3.6. Discussion and conclusion

We have reviewed atomistic modeling to calculate the elastic, plastic and fracture properties of tropocollagen molecules. Using full atomistic calculations, we presented a suite of calculations of different mechanical loading types and at different rates to gain insight into the deformation behavior of tropocollagen molecules. Our results suggest that it is critical to include a correct description of the bond behavior and breaking processes at large bond stretch, information which stems from the quantum chemical details of bonding.

An outcome of these studies is the observation that tropocollagen molecules undergo a transition from straight molecules to an S-shaped structure at increasingly large tensile stretch. As a consequence, we find that rupture of a single fiber does not occur homogeneously and thus at random locations, but instead, a local stress concentration develops leading to rupture of the molecule. Such 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 tropocollagen molecules at large deformation may play a critical role in the mechanical properties of bone during crack propagation [119,120,127], and elucidation of its mechanical response in particular at large strains is of critical importance during crack bridging in bone-like hard tissues [27,87,88].

Reactive modeling that takes into account the complexity of chemical bonding may be critical to understand the fracture and deformation behavior of many other biological and protein-based materials. Further, a strong rate dependence of the mechanical properties is observed, including Young's modulus and the shear interaction between molecules. This is in agreement with the fact that collagen is known to be a viscoelastic material and suggests that this behavior may at least partly originate from processes and mechanisms at the molecular scale. Further studies with respect to larger range
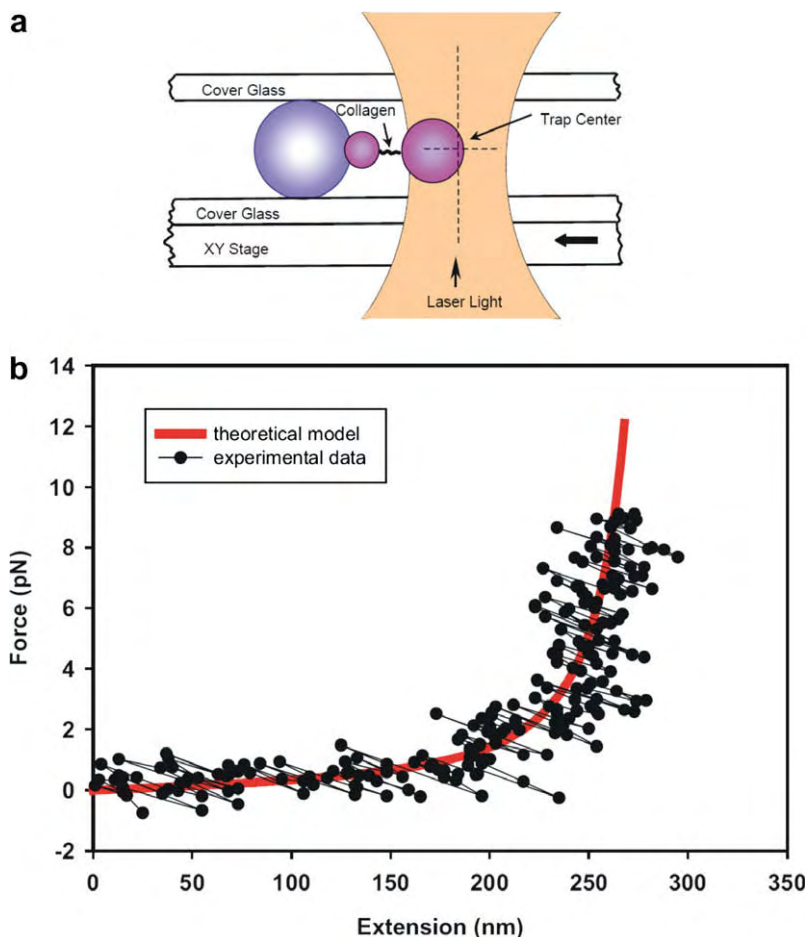
Fig. 31. Experimental stretching of a single tropocollagen molecule using optical tweezers, as reported in Ref. [36]. Subplot (a) depicts the experimental optical tweezers setup, showing the fixed bead and the laser trap that is used to apply forces to the molecules. Subplot (b) shows the force–extension curve obtained from this experiment (this data is also plotted in Fig. 30). Reprinted with permission from Ref. [36]. Copyright © 2004 Elsevier Ltd.

of time-scales provided important insight into the convergence of elastic properties of TC molecules at smaller deformation rates [110].

It is observed that the properties of collagen are scale dependent. For example, the fracture strength of individual polypeptides is different from the fracture strength of a tropocollagen molecule. The results further provide estimates of the fracture and deformation strength for different types of loading, enabling a comparison of different relative strengths.

The mesoscale model was used to predict the force–extension curve of a long tropocollagen molecule, including a direct and quantitative comparison with experimental results (as shown in Fig. 30 [111]). This result is remarkable since the molecular model does not involve empirical parameters, but instead, it is based solely on atomistic simulation results obtained from small tropocollagen structures. A full atomistic model would not have been able to describe the microsecond dynamics of stretching the molecule as shown in Fig. 30, illustrating the significance of mesoscale models to study the behavior of protein materials at these biologically critical scales.

3.4. Deformation and fracture of collagen fibrils

Now we move up in the hierarchical levels to study the mechanics of assemblies of many tropocollagen molecules into a collagen fibril (see Fig. 14, "Collagen fibrils"). Particular focus of this section is on studies of effects of the molecular length and cross-link densities. The present section reviews the molecular simulation results reported in [27,121,132] (for additional details regarding the modeling procedure see these original references).

3.4.1. Model geometry and molecular simulation approach

A two-dimensional plane stress model of collagen fibrils with periodic boundary conditions in the in-plane direction orthogonal to the pulling orientation is considered here, with a periodic array of 2×5 tropocollagen molecules (thus the total number of tropocollagen molecules is 10). The collagen fibrils show the characteristic staggered arrangement as observed in experiment. The plane stress condition is used to mimic the fact that the system is not periodic in the out of plane direction. Fully three-dimensional models are computationally very expensive. However, the model could treat such cases as well since there appears to be no intrinsic limitation of a two-dimensional case. No additional constraints other than the molecular interactions are applied to the system.

The simulations are carried out in two steps: (i) relaxation, followed by (ii) loading. Relaxation is achieved by slowly heating up the system, then annealing the structure at constant temperature, followed by energy minimization. Finite temperature calculations enable the structure to reassemble more easily, whereas energy minimization ensures finding the energetically optimal configuration of the molecules. If the initial relaxation is not carried out, pulling may be applied to a structure that is not in equilibrium and yield may be observed that is actually not due to the applied load but due to rearrangements towards the equilibrium structure. After relaxation, the structure displays the characteristics of collagen fibrils in agreement with experiment [8,82,91,337].

To model tensile deformation of collagen fibrils, displacement boundary conditions are implemented by continuously displacing a set of particles in the boundary regions (in a region of 40 Å to the left and right of the end of the collagen fibril).

The simulations are carried out by constantly minimizing the potential energy as the external strain is applied, where a displacement rate of 0.4 m/s is used for all simulations. Such rather high strain rates are a consequence of the time-scale limitation of the molecular model; total times spans of several microseconds are the most that can be simulated since the time step must be on the order of femtoseconds.

As indicated above, the virial stress is used to calculate the stress tensor for analyses of the stress–strain behavior. The yield stress σ_Y is defined as the stress at which permanent deformation of the collagen fibril begins. This is characterized either by intermolecular shear or by molecular fracture, leading to permanent deformation. The yield strain ε_Y is defined as the critical strain at which these mechanisms begin. The fracture stress σ_F is defined as the largest stress in the stress–strain curve, corresponding to the maximum load the collagen fibril can sustain. The fracture strain ε_F is the corresponding strain at which the largest stress occurs.

The strain is defined as $\varepsilon = (x - x_0)/x_0$, where x_0 is the initial, undeformed length of the collagen fibril, and x is the current, deformed length. It is noted that the extension ratio or stretch λ is related to the strain via $\lambda = 1 + \varepsilon$.

3.4.2. Size-dependent material properties: effects of molecular length

Here we focus on atomistic and molecular modeling of the mechanical properties of collagen under large stretch, leading to permanent deformation or fracture. We show that the key to understand the mechanics of collagen is to consider the interplay between the mechanics of individual tropocollagen molecules, the intermolecular chemical interactions, and the mesoscopic properties arising from hundreds of molecules arranged in fibrils. We explore the mechanics of collagen by considering different nanostructural designs, and pay specific attention to the details of molecular and intermolecular properties and its impact on the mechanical properties.

Under macroscopic tensile loading of collagen fibrils, the forces are distributed predominantly as tensile load carried by individual, and as shear forces between different tropocollagen molecules

[27]. Energetic effects rather than entropic contributions govern the elastic and fracture properties of collagen fibrils and fibers. The fracture strength of individual tropocollagen molecules is largely controlled by covalent polypeptide chemistry. The shear strength between two tropocollagen molecules is controlled by weak dispersive and hydrogen bond interactions, and by some intermolecular covalent cross-links.

3.4.2.1. Theoretical considerations: homogeneous shear. A simplistic model of a collagen fibril is considered first, by focusing on a staggered assembly of two tropocollagen molecules [27], as illustrated in Fig. 32. The shear resistance between two tropocollagen molecules, denoted as τ_{shear}, leads to a contact length dependent force

$$F_{tens} = \tau_{shear}L_C = \alpha\tau_{shear}L, \tag{29}$$

where L_C is the contact length, and F_{tens} is the applied force in the axial molecular direction, which can alternatively be expressed as tensile stress $\sigma_{tens} = F_{tens}/A_c$ by considering the molecular cross-sectional area A_c. The parameter α describes the fraction of contact length relative to the molecular length, $\alpha = L_C/L$. Due to the staggered geometry, the shear resistance increases linearly with L, thus $F_{tens} \sim \tau_{shear}L$. This model holds only if shear deformation between the molecules is homogeneous along the axial direction.

3.4.2.2. Theoretical considerations: nucleation of slip pulses. An alternative to homogeneous intermolecular shear is propagation of slip pulses due to localized breaking of intermolecular "bonds". This analysis is based on a one-dimensional model of fracture initially proposed by Hellan [122,123], which is based on Griffith's analysis of energy release rate applied to one-dimensional structures. The model describes a one-dimensional strip of material attached on a substrate, which is under tensile loading
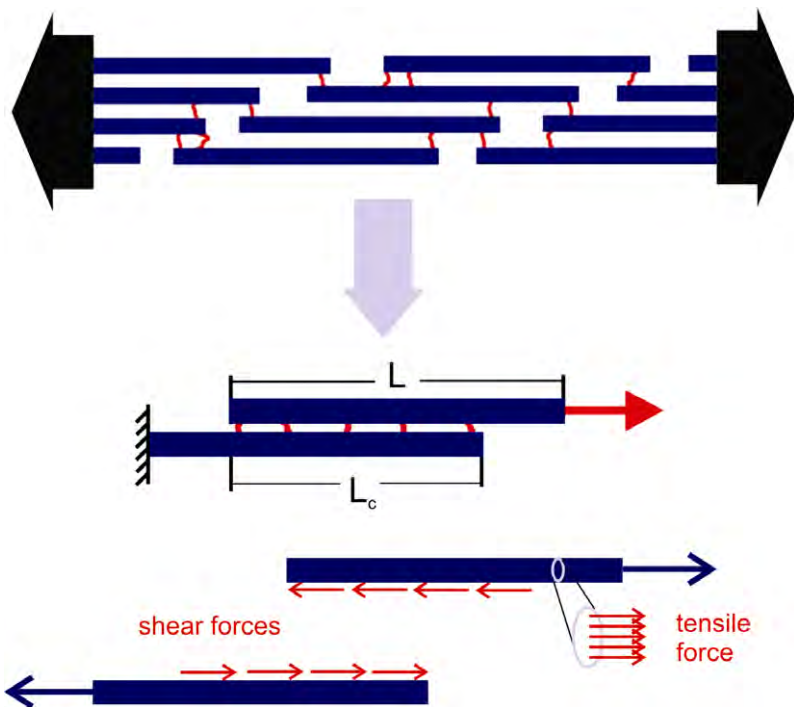
Fig. 32. Simplistic representation of the staggered collagen fibril geometry as a simple bimolecular assembly. The lower part of the figure shows the distribution of intermolecular shear forces and tensile forces within each TC molecule.

in the axial direction. The energy released per unit rupture advance of this setup is used to derive a critical condition for fracture initiation. This one-dimensional model is analogous to dislocation models or brittle fracture models in higher dimensions, based on identical concepts.

At a critical load, the energy released per advancement length of the adhesion front is equal to the energy required to break the bonding between the material strip and the substrate, leading to initiation of failure front. The failure front – corresponding to a dynamic crack tip – propagates at a fraction of the sound velocity, eventually displacing the material permanently in the direction of the applied load. This model is now applied to intermolecular deformation in collagen fibrils. The energy release rate is given by

$$G_0 = \frac{\sigma_R^2}{2E}, \tag{30}$$

where E is Young's modulus of the tropocollagen molecule, and σ_R the applied stress. With γ as the energy necessary to nucleate this defect, at the onset of nucleation the condition:

$$G_0 = \gamma, \tag{31}$$

needs to be satisfied (similar to the Griffith condition). The detachment front corresponds to the front of decohesion. Bonds behind the fracture front reform, thus forming a "slip pulse". The slip pulse is a region with increased tensile strain in the tropocollagen molecule, which is several nanometers wide. The existence of slip pulses is not a consequence of the discretization of the mesoscale model. Instead, this theoretical framework is developed based on continuum mechanics, assuming a homogeneous distribution of adhesive interactions along the molecular surface. In the spirit of Griffith's energy argument describing the onset of fracture, nucleation of slip pulses is controlled by the applied tensile stress σ_R, where

$$\sigma_R = \sqrt{2E\gamma}, \tag{32}$$

where E is Young's modulus of an individual tropocollagen molecule, and γ relates to the energy required to nucleate a slip pulse.

When $\sigma_{tens} < \sigma_R$, deformation is controlled by homogeneous shear between tropocollagen molecules. However, when $\sigma_{tens} \geqslant \sigma_R$ intermolecular slip pulses are nucleated. This leads to a critical molecular length:

$$\chi_S = \frac{\sqrt{2\gamma E}}{\tau_{shear}\alpha} A_C. \tag{33}$$

For fibrils in which $L < \chi_S$, the predominant deformation mode is homogeneous shear. When $L > \chi_S$, propagation of slip pulses dominates. The strength of the fibril is then independent of L (Eq. (33)), approaching $\tau_{shear}\alpha\chi_S$. This concept is similar to the flaw tolerance length scale proposed for mineral platelets in bone [153] (numerical studies of this phenomenon see [154]).

The length scale χ_S depends on the material parameters and interaction between molecules. If γ assume very large values – for instance due to high cross-linking density, or due to the effects of solvents (e.g. low water concentration) – the tensile forces in each tropocollagen molecule (Eq. (29), or $F_{tens} \sim L$) reaches the tensile strength of tropocollagen molecules, denoted by F_{max}, before homogeneous shear or slip pulses are nucleated (F_{max} is a material constant that ultimately depends on the molecular structure of the tropocollagen molecule).

Considering $F_{tens} = F_{max}$ leads to a second critical molecular length scale,

$$\chi_R = \frac{F_{max}}{\tau_{shear}\alpha}. \tag{34}$$

This molecular length χ_R characterizes when the transition from molecular shear to brittle-like rupture of individual tropocollagen molecules occurs. The response of collagen fibrils to mechanical load changes from shear or glide between tropocollagen molecules, to molecular fracture as L increases. For $L > \chi_R$, tropocollagen molecules break during deformation, whereas for $L \leqslant \chi_R$ deformation is characterized by homogeneous intermolecular shear.

The integrity of a complete collagen fibril is controlled by the strength of the weakest link. Thus, the interplay of the critical length-scales χ_S/χ_R controls the deformation mechanism.

When $\chi_S/\chi_R < 1$, slip pulse nucleation governs at large molecular lengths, whereas when $\chi_S/\chi_R > 1$, fracture of individual tropocollagen molecules occurs. For $L/L_\chi < 1$ homogeneous intermolecular slip dominates deformation. In both cases, the strength does not increase by making L larger. The maximum strength of the fibril is reached at $L \approx L_\chi = \min(\chi_R, \chi_S)$. This is true for any arbitrary length L of a tropocollagen molecule. Homogeneous shear deformation dominates below the critical molecular length L_χ. For molecules with $L > L_\chi$, either slip pulses or fracture set in, depending on which of the two length-scales χ_S or χ_R is smaller. For short tropocollagen molecules, the strength of collagen fibrils tends to be small and depends on L_C. When $L \approx L_\chi$, the maximal tensile strength of fibrils is reached.

Further, choosing $L \approx L_\chi$ leads to maximized energy dissipation during deformation. The work necessary to separate two fibers in contact along a length L_C under macroscopic tensile deformation is

$$E_{\mathrm{diss}} = \int_{l=0}^{l=L_C} l\tau_{\mathrm{shear}}\, \mathrm{d}l = \frac{1}{2} L_C^2 \tau_{\mathrm{shear}}. \tag{35}$$

Eq. (35) predicts an increase of the dissipated energy with increasing molecule length, therefore favoring long molecules. If $\chi_R < \chi_S$, the critical length L_χ constitutes an upper bound for L_C, since molecules rupture before shear deformation sets in. After bond rupture and formation of shorter molecules, E_{diss} decreases significantly, suggesting that $L > L_\chi$ is not favored. Energy dissipation is at a maximum for $L \approx L_\chi$. If $\chi_S < \chi_R$, the dissipated energy can be approximated by (assuming $L_C > \chi_S$)

$$E_{\mathrm{diss}} \approx \left(\frac{1}{2} \alpha^2 \chi_S^2 \tau_{\mathrm{shear}} + (L_C - \alpha\chi_S) \cdot F_{\mathrm{max}} \right), \tag{36}$$

suggesting that after a quadratic increase for small molecular lengths, the dissipated energy increases linear with L_C.

3.4.2.3. Molecular modeling of mechanical properties of collagen fibrils.
We now model the deformation behavior of a more realistic fibril geometry as shown in Fig. 21 (upper part), studying the change in mechanical properties due to variations in molecule length L.

Due to the staggered design of collagen fibrils with an axial displacement of about 25% of the molecular length [124–126], the contact length between tropocollagen molecules in a fibril is proportional to L. The length-scales suggested in the previous section therefore have major implications on the deformation mechanics of collagen fibrils.

We consider fully hydrated cross-link free collagen fibrils serving as a model for cross-link deficient collagen. Fig. 33 shows the stress versus strain response of a collagen fibril for different molecular lengths L. The results suggest that the onset of plastic deformation, the maximum strength and large strain mechanics of collagen fibrils depends on the molecular length.

Fig. 34 shows the normalized elastic strength of the fibril as a function of molecular length L. The results suggest an increase up to about 200 nm, then reaching a plateau value of around 0.3 GPa (results normalized by this value). The elastic uniaxial strains of collagen fibrils reach up to approximately 5%. The maximum stress reaches up to 0.5 GPa, during plastic deformation. The molecular length at which the saturation occurs corresponds to a change in deformation mechanism, from homogeneous shear ($L \to 0$) to nucleation of slip pulses ($L \to \infty$). The corresponding molecular length provides an estimate for the critical molecular length scale $\chi_S \approx 200$ nm.

It is noted that $\chi_R \approx 436$ nm, as described in the previous section (it is a material property of the reference system). Therefore, the ratio $\chi_S/\chi_R < 1$, suggesting a competition between slip pulses and homogeneous shear as the molecular length is varied. This suggests that cross-link deficient collagen may predominantly undergo intermolecular shear deformation.

Fig. 35 depicts the energy dissipated during deformation, per unit volume. A continuous increase with molecular length L is observed, reaching a maximum at a critical molecular length L_χ, then a slight decrease. Energy dissipation increases further at ultra large molecular lengths beyond 400 nm, due to longer shear paths during slip pulse propagation. The modest increase in energy dissipation for ultra-long molecules may be an inefficient material solution, in particular since
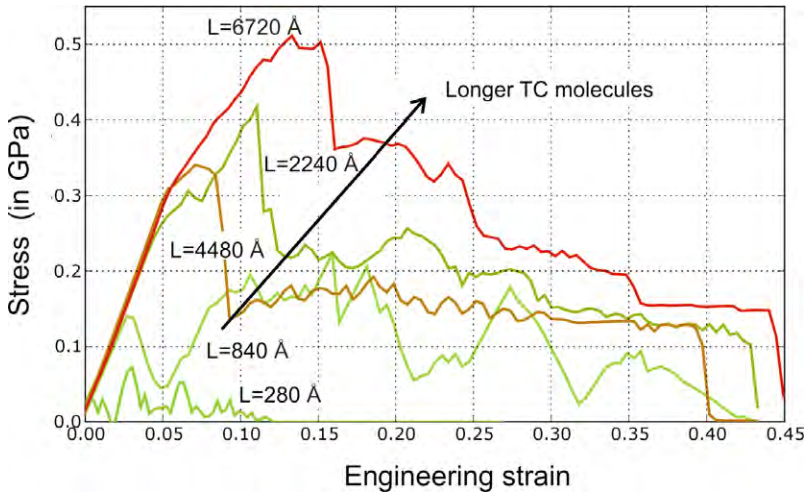
Fig. 33. Stress vs. strain of a collagen fibril, for different molecular lengths (model for cross-link deficient collagen, as no covalent cross-links are present in the collagen fibril) (results as reported in [27]). The results suggest that the longer the molecular length, the stronger the fibril. The maximum elastic strength achieved by collagen fibrils approaches approximately 0.3 GPa, with largest stress around 0.5 GPa. The onset of intermolecular shear can be recognized by the deviation of the stress–strain behavior from a linear-elastic relationship.
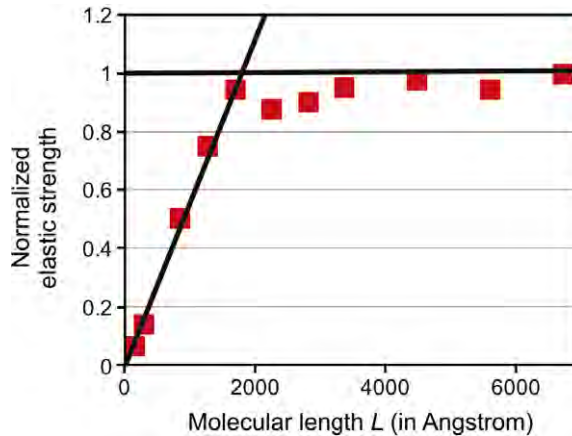
Fig. 34. This plot shows the critical stress at the onset of plastic shear between TC molecules [27]. An initial regime of linear increase of strength with molecular length is followed by a regime of finite strength, at a plateau value.

assembling such ultra-long molecules into regular fibrils might be increasingly challenging from a material synthesis viewpoint. This is because tropocollagen molecules must be aligned properly to facilitate the formation of collagen fibrils, with cross-links and other adhesion mechanisms placed with high precision. The likelihood of introducing disorder or additional defects increases with longer molecular lengths.

3.4.2.4. Discussion. The results of the simulation studies reviewed in the previous sections suggest that the length of tropocollagen molecules plays a significant role in determining the deformation mechanics, possibly explaining some of the structural features of collagen found in biological tissues.
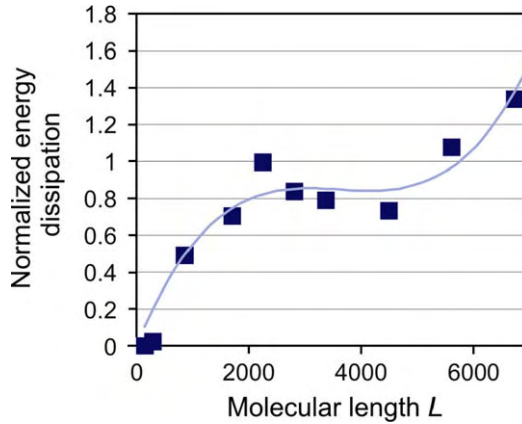
Fig. 35. The plot depicts the dissipated energy during deformation, per unit volume, in a collagen fibril, as a function of molecular length, normalized by the maximum value [27]. An initial steep increase is followed by a plateau regime, with a local maximum around 220 nm. The smooth curve is a fit of a third order expansion to the simulation data.

The two length-scales χ_S and χ_R provide a quantitative description of the three different deformation mechanisms in collagen fibrils: (i) intermolecular shear, (ii) slip pulse propagation and (iii) fracture of individual tropocollagen molecules.

The governing deformation mechanism is controlled by the ratio χ_S/χ_R: whether molecular fracture ($\chi_S/\chi_R > 1$) or slip pulses ($\chi_S/\chi_R < 1$) dominate deformation, the strength of the fibril approaches a maximum that cannot be overcome by increasing L. When $L_\chi = \min(\chi_R, \chi_S)$, tensile forces due to shear are in balance with either the fracture strength of tropocollagen molecules ($\chi_S/\chi_R > 1$) or with the critical load to nucleate slip pulses ($\chi_S/\chi_R < 1$). In either case, the maximum strength of the fibril is reached when $L \approx L_\chi$, including maximum energy dissipation.

When the length of collagen molecules is close to the critical length scale L_χ, two objectives are satisfied: (i) under large deformation, tropocollagen molecules reach their maximum strength without leading to brittle fracture, and (ii) energy dissipation during deformation is maximized. This concept may explain the typical staggered geometry of collagen fibrils found in experiment, with extremely long molecules – leading to large energy dissipation during deformation.
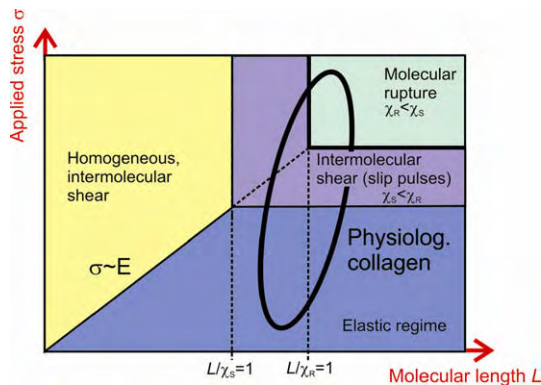
Fig. 36. Deformation map of collagen fibrils, as reported in [27]. The mechanical response of is controlled by two length-scales χ_S and χ_R. Intermolecular shear governs deformation for small molecular lengths, leading to relatively small strength of the collagen fibril. For large molecular lengths, either intermolecular slip pulses ($\chi_S/\chi_R < 1$) or rupture of individual TC molecules ($\chi_S/\chi_R > 1$) dominate. This regime refers to the case of strong intermolecular interactions (e.g. increased cross-link densities). Physiological collagen typically features long molecules, with variations in molecular interaction, so that either intermolecular shear (e.g. slip pulses) or molecular fracture are predicted to dominate.

The mechanisms of deformation and their dependence on the molecular architecture are summarized in a deformation map, shown in Fig. 36.

Slip pulses are nucleated by localized larger shear stresses at the end of the tropocollagen molecules. Thus, cross-links at these locations provide a molecular scale mechanism to prevent slip pulse nucleation, as this leads to an increase of the energy required to nucleate slip pulses, thus to a larger value of γ. This results in an increase of χ_S, due to the scaling law:

$$\chi_S \sim \sqrt{\gamma}. \tag{37}$$

As a consequence, the ratio χ_S/χ_R increases, making collagen fibrils stronger. Remarkably, this nano-scale distribution of cross-links agrees with the natural collagen design seen in experiment, often showing cross-links at the ends of the tropocollagen molecules (reminiscent of crack bridging).

Cross-links provide additional strength to the fibrils, in agreement with experiment [114]. However, extremely large cross-link densities lead to adverse effects as the material is not capable to dissipate much energy during deformation – leading to a brittle collagen that is strong, but not tough. Such behavior is observed in dehydrated collagen, or in aged collagen featuring higher cross-link density [128,129]. In contrast, decreased cross-linking as it occurs in the Ehlers-Danlos V disease [104,105] leads to significantly reduced tensile strength of collagen, as $\chi_S/\chi_R < 1$. The ratio L/L_χ decreases, resulting in skin and joint hyperextensibility due to extremely weak collagen tissue, incapable of dissipating significant energy.

Our model can be used to study different design scenarios. A design with many cross-links and short molecules would lead to a very brittle collagen, even in the hydrated state. Such behavior would be highly disadvantageous under physiological conditions. In contrast, long molecules provide robust material behavior with significant dissipation of energy (Fig. 36).

Both elastic strength and energy dissipation approach a finite value for large molecular lengths, making it inefficient to create collagen fibrils with tropocollagen molecules much longer than L_χ, which is on the order of a few hundred nanometers (Fig. 34). This length scale agrees with many experimental results of structural analyses of collagenous tissues that have focused on analyzing the typical molecular length of tropocollagen molecules [4,27,29,36,89,92,97,125,131]. These studies show that tropocollagen molecules show lengths predominantly at around 300 nm, suggesting a universal design feature common to a broad range of collagenous tissue types. The computational experiments reported in [27] that enabled a systematic variation of the molecular length provided perhaps the first insight into the reasons for these particular length-scales from a mechanical perspective (that is, the role in maximizing strength and energy dissipation).

Large deformation is a critical physiological condition for collagen-rich tissue. The risk of catastrophic brittle-like failure needs to be minimized in order to sustain optimal biological function. The nano-scale ultrastructure of collagen may be designed to provide robust material behavior under large deformation by choosing long tropocollagen molecules. Robustness is achieved by the design for maximum strength and maximized energy dissipation by shear-like mechanisms. The requirement for maximum energy dissipation plays a crucial role in determining the optimal molecular length L_χ. The layered design of collagen fibrils plays a critical role in enabling long deformation paths with large dissipative stresses. This is reminiscent of the "sacrificial bond" concept.

The properties of collagen are scale-dependent [90]. The fracture strength of an individual tropocollagen molecule (11.2 GPa) differs from the fracture strength of a collagen fibril (0.5 GPa). Similarly, Young's modulus of an individual tropocollagen molecule is approximately 7 GPa, while Young's modulus of a collagen fibrils is smaller, approaching 5 GPa (for $L \approx 224$ nm). This is in qualitative agreement with experiment [90] (it is noted that if a lower value for the molecular modulus is used in the mesoscale formulation, an even smaller fibril modulus would be expected; this could be investigated in future studies.

Quantitative theories of the mechanics of collagen have many applications, ranging from the development of new biopolymers to studies in tissue engineering, where collagen is used as a scaffolding material. In addition to optimization for mechanical properties, other design objectives such as biological function, chemical properties or functional constraints may be responsible for the structure of col-

lagen. However, the physiological significance of large mechanical deformation of collagen fibers suggests that mechanical properties could indeed be an important design objective.

3.4.3. Effect of cross-link densities

To understand the influence of cross-links on the deformation mechanics of collagen fibrils, a series of computational experiments of pulling individual collagen fibrils with increasing density of cross-links are carried out [121]. All results are compared with a control system of a cross-link free collagen fibril. Systematic increases of the density of cross-links enables one to observe the difference in mechanical behavior. Particular attention is paid to the small and large deformation behavior and the effect of intermolecular cross-links on the mechanical properties and deformation mechanisms.

In particular, studies are carried out that focus on the changes in the elastic and fracture behavior of the collagen fibril as the parameters are varied. An analysis of the molecular mechanisms allows one to develop a mechanistic understanding of the deformation behavior of collagen fibrils.

Fig. 27 visualizes how the molecular model describes the presence of intermolecular cross-links (see also definition of parameter β in Section 3.3.4.3).

3.4.3.1. Tensile deformation: stress–strain curves. Fig. 37 depicts the stress–strain curve for various cross-link densities, expressed in terms of the parameter β. For small values of cross-link densities ($\beta < 10$), the fibril starts to yield at strain in the range of 5–10%, and shows rather long dissipative deformation paths, leading to fracture at strains between 50% and 100%.

It is found that larger cross-link densities lead to larger yield strains, larger yield stresses as well as larger fracture stresses. At a critical cross-link density corresponding to one cross-link per molecule ($\beta \approx 15$), the second, steeper elastic regime is activated. This strong increase in tangent modulus corresponds to stretching of the protein backbone. This molecular deformation mode dominates after the uncoiling of the tropocollagen molecule under breaking of H-bonds [26,110]. The results clearly confirm the significance of the presence and density of cross-links on the deformation behavior.

Large elastic tensile strains of up to 50% are possible since each tropocollagen molecule itself can sustain strains of up to 50% tensile deformation (this is shown in Fig. 28, curve for a single tropocollagen molecule).

In the collagen fibril, such large strains at the molecular scale are only possible if strong links exist which prevent molecular slip and therefore enable transfer of large loads to the individual
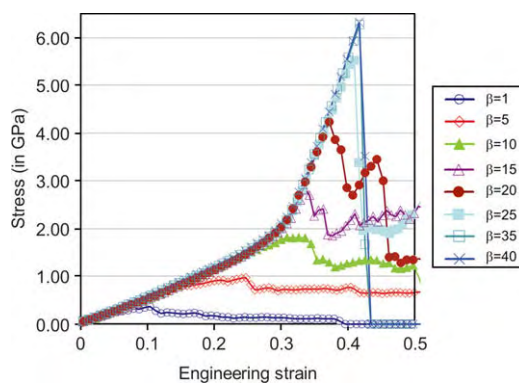
Fig. 37. Stress vs. strain of a collagen fibril, for different cross-link densities, as reported in [121]. The results clearly show that larger cross-link densities lead to larger yield strains, larger yield stresses as well as larger fracture stresses. For larger cross-link densities, the second elastic regime (seen as much steeper, second slope) is activated. As the cross-link density increases, the collagen fibril shows a more 'brittle-like' deformation behavior. For values of $\beta > 25$, the deformation mechanisms is characterized by molecular fracture, and as a consequence, the maximum fracture stress of the collagen fibril does not increase with increasing cross-link densities. This cross-link density corresponds to the case when two cross-links per molecule are present.

tropocollagen molecules. As shown in the present work, developing cross-links between molecules is a possible means of achieving this situation.

3.4.3.2. Comparison: single tropocollagen molecule and collagen fibril. Fig. 38 shows a comparison between the stress–strain curves of a collagen fibril ($\beta = 25$) and a single tropocollagen molecule, for tensile strains below 40%. Both structures are completely in the elastic regime (the tropocollagen molecule fractures at approximately 50% tensile strain and the collagen fibril starts to yield at slightly above 40% strain). These results show that the stresses in the single tropocollagen molecule are larger than in the collagen fibril.

Fig. 39 plots the tangent modulus of the stress–strain curve depicted in Fig. 38. The results clearly indicate that the tangent modulus of the single tropocollagen molecule is larger than that of a collagen fibril. The results suggest that the modulus of a single tropocollagen molecule is approximately 40% larger throughout deformation. This agrees well with experimental results [90], suggesting an increase of the stiffness from fibril to molecule close to 40%. Fig. 40 plots the results of experiments, carried out at three distinct levels of collagen organization, including tropocollagen molecules, collagen fibrils and tendon fibers. This figure is based on the results reported in Ref. [90].

Even though cross-links are stiffer than the tropocollagen molecule itself, the overall density of cross-links is rather small so that the stiffening effect is negligible. The origin of the softening is the combination of rather weak intermolecular interactions with the single molecule elasticity along most of the axial length of the tropocollagen molecules. This leads to an effective softening of the fibrillar structure even when cross-links are present. This may change for extremely large cross-link densities, for example when cross-links form along the entire axial dimension of the chain. Section 3.5 provides additional comparison with experimental results.

3.4.3.3. Yield stress and fracture stress analysis. Fig. 41 depicts the yield stress of a collagen fibril as a function of the cross-link density (curve "relative strength"). The plot shows that for larger cross-link densities, the material becomes stronger. However, when $\beta > 25$, the yield stress and fracture stress do not depend on the cross-link density any more as the yield stress reaches a plateau value. The plateau can be explained by a change in molecular deformation mechanism from predominantly intermolecular shear (for $\beta < 25$) to molecular fracture (for $\beta > 25$). Whereas the strength of the fibril is controlled
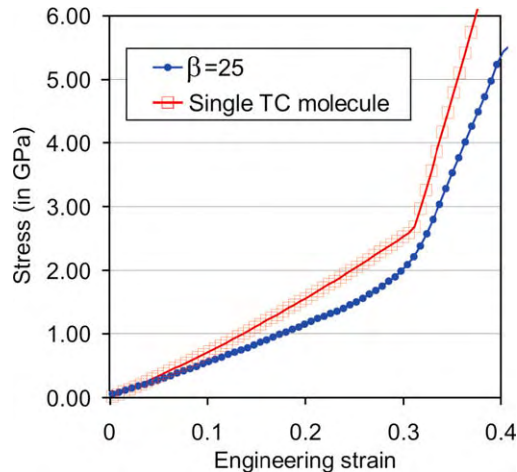
Fig. 38. Stress vs. strain, comparing a collagen fibril ($\beta = 25$) with a single TC molecule [121]. Both structures are completely in the elastic regime (the TC molecule fractures at approximately 50% tensile strain and the collagen fibril starts to yield at slightly above 40% strain). This plot shows that the stresses in the single TC molecule are larger than in the collagen fibril, and that the tangent modulus is larger throughout deformation. This finding agrees well with experimental results [90].
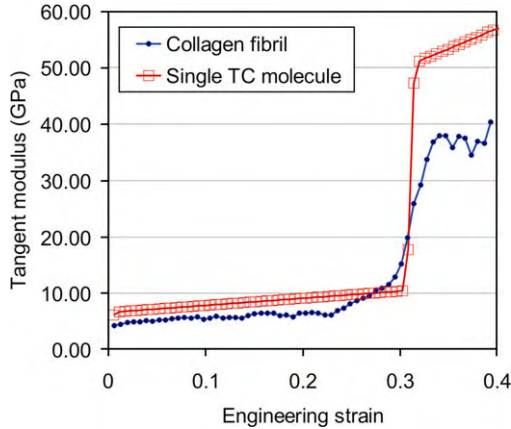
Fig. 39. Tangent modulus vs. strain, comparing a single TC molecule and a collagen fibril (cross-link parameter β = 25) [121]. The results show that the tangent modulus of the single TC molecule is approximately 40% larger, except for the transition region during which the modulus of the fibril is larger (between 20% and 30% fibril strain).
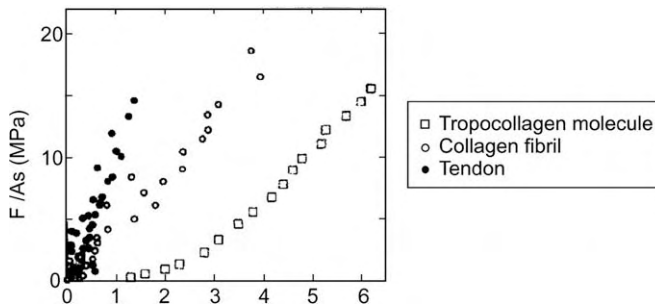
Fig. 40. Experimental results, stress–strain experiments of tropocollagen molecules, collagen fibrils and tendon. These experimental results clearly reveal the reduction of modulus when going from the pure tropocollagen molecule to the collagen fibril to the microscale tendon structure (compare the change in slope from tropocollagen molecule to collagen fibril with the results shown in Figs. 38 and 39 for the modulus). Reprinted with permission from Ref. [90]. Copyright © 1996 Elsevier Ltd.

by intermolecular adhesion for β < 25, the strength is dominated by the molecular fracture properties. This observation confirms a change in mechanisms as suggested in an earlier study [27].

As the cross-link density increases, the collagen fibril becomes more 'brittle-like'. The increasingly brittle character is clearly illustrated by the ratio of fracture stress versus yield stress. For smaller values of β < 15, the stress–strain curves show a stiffening effect after onset of yield, similar to work-hardening as known in metal plasticity. However, this stress increase decreases with increasing cross-link density. The data shows that as the cross-link parameter exceeds 15, the material becomes 'brittle-like', characterized by immediate drop of the stress after onset of yield without dissipative deformation.

An analysis of the stress–strain behavior provides further insight into the elastic and plastic deformation modes. The analysis of the stress–strain curves for varying cross-link densities corroborates the notion that for increasing cross-link densities, the material becomes increasingly 'brittle-like'. Fig. 42 further depicts the relative strength ($\sigma_{frac,rel} = \sigma_{frac}/\sigma_{frac,max}$), relative amount of plasticity (calculated with yield strain ε_{yield} and yield stress ε_{frac}, $P = (\varepsilon_{frac}/\varepsilon_{yield} - 1)/(\max(\varepsilon_{frac}/\varepsilon_{yield}) - 1)$), as well as the relative toughness of a collagen fibril ($T = \sqrt{P \cdot \sigma_{frac,rel}}$) as a function of the cross-link parameter β, as reported in [132]. Hereby the 'toughness' of a material is defined as the property of being both
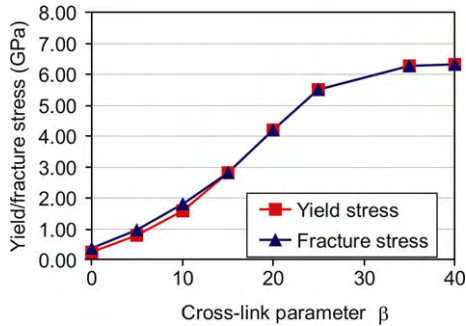
Fig. 41. Yield stress and fracture stress of a collagen fibril as a function of the cross-link density, expressed by the adhesion parameter β [157]. For large cross-link densities, the material behavior becomes increasingly brittle and the failure strength (or yield strength, equivalently) does not depend on the cross-link density any more (saturation of yield stress) since failure is controlled by fracture of individual TC molecules.
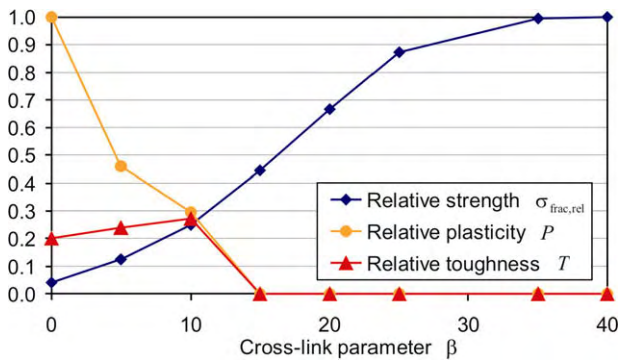
Fig. 42. Relative strength $\sigma_{\text{frac,rel}}$, relative amount of plasticity P as well as relative toughness of a collagen fibril (referred to as T) as a function of the cross-link parameter β, as reported in [132,133]. For large cross-link densities $\beta > 15$, the material behavior becomes increasingly brittle and the failure strength (or yield strength, equivalently) saturates, as failure is controlled by rupture of individual TC molecules. The variation of the toughness T suggests that a maximum relative toughness is reached for a cross-link parameter of approximately $\beta \approx 10$.

strong and requiring an increasing energy to break. The results suggest that an optimal relative toughness is reached for a cross-link parameter of approximately $\beta \approx 10$. This value density corresponds to an approximate spacing of cross-links in the molecular axial direction of $\approx$420 nm.

3.4.3.4. Discussion. Molecular modeling has been employed to predict the small and large deformation mechanics of collagen fibrils, as a function of varying cross-link densities. The results suggest that the cross-link density governs the large deformation and in particular the yield or fracture mechanics. However, it influences the small deformation mechanics only marginally (see, e.g. Fig. 41).

The model predicts that collagen fibrils are capable of undergoing extremely large deformation without fracturing; how much of this is elastic or dissipative depends on the cross-link densities. It is found that two prominent molecular mechanisms of permanent deformation dominate: molecular glide and molecular rupture.

The formation of covalent cross-links is essential to reach the elastically stiffer, second regime in the stress–strain curve of collagen, which corresponds to backbone stretching in the tropocollagen molecule. This phenomenon can be understood based on the mechanisms and the effect of the presence of cross-links: the increased traction at the end of the molecule allows for larger molecular

strains to be reached. The larger strains give rise to larger overall yield and fracture stress. However, collagen fibrils become more 'brittle-like' under these conditions as their ability to undergo dissipative, plastic deformation is reduced. These findings confirm some of the key hypotheses put forward in [27], including effect of cross-links in making the material appear more 'brittle', observed deformation mechanics and reduction of elastic modulus. This is confirmed by several analyses shown in Fig. 42, for instance.

The results improve the understanding of how molecular changes during ageing contribute to modifications of tissue properties. Ageing of organisms is primarily controlled by changes in the protein structure of elastin and collagen, when increased cross-linking between molecules develops due to non-enzymatic processes. These changes in the molecular architecture may lead to diseases that are induced by the modification of the mechanical properties of tissues [21,121,132]. The analysis confirms that cross-linking indeed leads to stiffening and increasing 'brittleness' of collagen based tissues. It is noted that the results shown in Fig. 41 are in good qualitative agreement with the results of stress–strain responses of collagen during ageing (see Fig. 3 in Ref. [90]). Both the present model and experiment predict a stronger and less dissipative behavior with the development of additional cross-links.

It is found that the material properties of collagen are scale dependent. A softening of the modulus is observed when tropocollagen molecules are assembled into a collagen fibril. The modeling suggests a reduction of modulus on the order of 40%, which is close to experimental results [90] of similar comparisons between the mechanics of collagen fibrils and tropocollagen molecules (see Fig. 42). This can also be found by taking a simple average value of all values for tropocollagen molecules reported in the literature (5.1 GPa, average of Table 2) divided by the average value of moduli for the collagen fibril (2.8 GPa, average of Table 3), which suggests an increase of modulus by approximately 80%.

The results show several features of the stress–strain behavior also found in experiment [89,92,118], notably the two regimes of moduli with a strong progressive stiffening with increasing strains. However, the magnitude of the stress is different, as MD modeling predicts larger stresses and larger moduli than seen in experiment.

Table 2
Comparison of Young's modulus (stiffness) of single tropocollagen molecules, experiment and computation

Study/case and approach	Young's modulus (GPa)
Single molecule stretching [101], atomistic modeling	4.8 ± 1
Single molecule stretching [26], reactive atomistic modeling	$\approx$7 (4 at vanishing rates [110])
Single molecule stretching [330], atomistic modeling	2.4
X-ray diffraction [90]	$\approx$3
Brillouin light scattering [331]	9
Brillouin light scattering [332]	5.1
Estimate based on persistence length [333]	3
Estimate based on persistence length [36,330]	0.35–12

Table 3
Comparison of Young's modulus of collagen fibrils, experiment and computation

Study, case and approach	Young's modulus (GPa)
MEMS stretching of collagen fibrils [118]	$\approx$0.4–0.5 (small strain modulus) $\approx$12 (large strain modulus)
X-ray diffraction [149]	1
AFM testing [334]	2–7 (ambient conditions) 0.2–0.8 (aqueous media)
Molecular multi-scale modeling [27]	4.36 (small strain modulus) $\approx$38 (large strain modulus)

A limitation of the present study is that spatial inhomogeneities of cross-link distributions are not considered. In principle, this can be implemented straightforwardly. Also, changes of molecular properties along the molecular length have not been considered, an important characteristic feature of many collagen based tissues. This aspect is particularly significant to account for entropic effects that stem from more floppy labile regions of the tropocollagen molecules [135]. These important aspects will be addressed in future work.

An improved understanding of the nanomechanics of collagen may help in the development of biomimetic materials, or for improved scaffolding materials for tissue engineering applications [128]. Diseases such as Ehlers-Danlos [128,129], *Osteogenesis Imperfecta*, Scurvy or the Caffey disease [3,4,7] are caused by defects in the molecular structure of collagen altering the intermolecular and molecular properties due to genetic mutations, modifying the mechanical behavior of collagen fibrils.

3.5. Comparison with experimental results: nanomechanics of single TC molecules and collagen fibrils

This section is dedicated to a brief discussion of the computational results reviewed above, in light of recent experimental reports of stretching experiments of tropocollagen molecules and individual collagen fibrils. Table 2 provides an overview of moduli obtained for single tropocollagen molecules. The comparison shows that our predictions for the moduli are close to experimental results, albeit fall into the higher end of the range of values reported. Table 3 summarizes results for elastic moduli of collagen fibrils from various sources. Unlike as for the single molecule case the agreement between experiment and simulation is not as good. A few important observations are discussed in more detail. Recently, MEMS devices were used to carry out tensile studies of single collagen fibrils [118]. The authors obtained a small deformation modulus of approximately 0.4 GPa, and a large deformation modulus of 12 GPa. The absolute values of the small strain moduli are approximately 10 times smaller than in the simulation results [27,121].

One possible explanation for this disagreement could be entropic effects that may make the fibril softer in particular in the small deformation regime. Such entropic effects are not considered in the studiers of collagen fibrils reviewed here, since all molecules are completely stretched out to their contour length at the beginning of the simulation and thus enter the energetic stretching regime instantaneously.

Another possible reason may be the larger deformation rates used in atomistic modeling than those used in many experiments [40,138], which might lead to overestimation of forces during mechanical deformation. Considering smaller deformation rates, for instance, may lead to smaller values for Young's modulus, as typically unfolding forces are larger for larger deformation rates (e.g. based on concepts related to Bell Theory [118]). Since the molecular model used in this study is based solely on atomistic input data, overestimation of the modulus value from the MD simulations will be transported throughout the multi-scale modeling scheme. This may explain why the values reported in the present work are close to the upper end of the range of experimental measurements. However, the ratio of large strain modulus to small strain modulus is on the same order of magnitude, being between 24 and 30 in experiment and approximately 8.4 in simulation. The transition from the small deformation to large deformation elastic modulus occurs at strains of approximately 30%, which is found consistently in both experiment and simulation.

It has been suggested in [90,118] that the tensile strength may be greater than 1 GPa, which is corroborated by our results that suggest strengths ranging from 300 MPa (cross-link deficient fibrils) to 6 GPa (highly cross-linked collagen fibrils). These values agree with the strengths predicted in our simulation (see Fig. 37, for example). On the other hand, other results show much lower failure stresses on the order of several MPa. A possible explanation for this discrepancy could be molecular defects, high loading rates or different geometries in those experiments [27] that do not resemble the perfect patterns as considered in our study.

The various results discussed in this section in comparison with experimental studies may form the basis for future investigation, for example with slower loading rates, more realistic microstructures of collagen fibrils and the considerations of higher order hierarchical levels. Such studies could provide the basis for more detailed comparisons.

3.6. Nanomechanics of mineralized collagen fibrils: molecular mechanics of nascent bone

One of the most intriguing protein materials found in Nature is bone, a material composed out of assemblies of tropocollagen molecules and tiny hydroxyapatite crystals, forming an extremely tough, yet lightweight material [3,4]. Bone has evolved to provide structural support to organisms, and therefore, its mechanical properties are of great physiological relevance. Since collagen is the most fundamental building block of bone, here we review some insight into bone's smallest scale during bone formation (nascent bone), mostly based on the study reported in [139].

Mineralized collagen fibrils are highly conserved nanostructural building blocks of bone. By a combination of molecular dynamics simulation and theoretical analysis it is shown that the characteristic nanostructure of mineralized collagen fibrils is vital for its high strength and its ability to sustain large deformation, as relevant to the physiological role of bone, creating a strong and tough material. An analysis of the molecular mechanisms of protein and mineral phases under large deformation of mineralized collagen fibrils reveals a fibrillar toughening mechanism that leads to a manifold increase of energy dissipation compared to fibrils without mineral phase. This fibrillar toughening mechanism increases the resistance to fracture by forming large local yield regions around crack-like defects, a mechanism that protects the integrity of the entire structure by allowing for localized failure.

As a consequence, mineralized collagen fibrils are able to tolerate micro-cracks on the order of several hundred micrometers size without causing any macroscopic failure of the tissue, which may be essential to enable bone remodeling. The analysis proves that adding nanoscopic small platelets to collagen fibrils increases their Young's modulus, yield strength as well as their fracture strength. We find that mineralized collagen fibrils have a Young's modulus of 6.23 GPa (versus 4.59 GPa for the collagen fibril), yield at a tensile strain of 6.7% (versus 5% for the collagen fibril) and feature a fracture stress of 0.6 GPa (versus 0.3 GPa for the collagen fibril).

The work reviewed here [139] (additional details regarding the numerical procedure can be found in this reference) is limited to the scale of mineralized fibrils, with the objective to provide insight into the most fundamental scales of bone and its deformation mechanics under tensile loading [3,28,125,140,141].
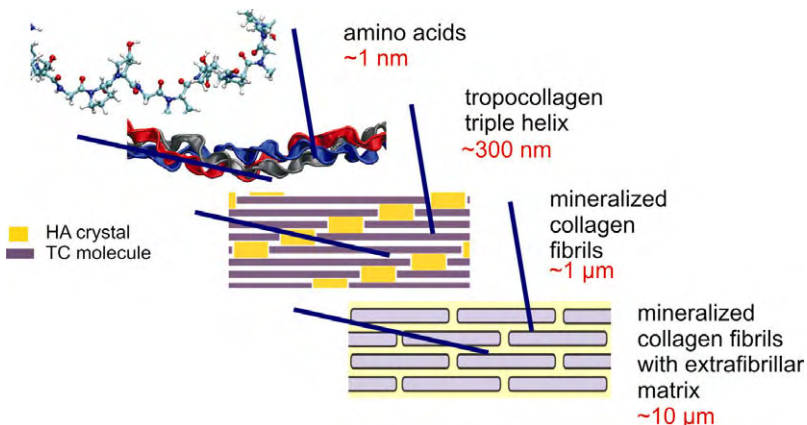
Fig. 43. Geometry of the nanostructure of bone, showing several hierarchical features from atomic to microscale. Simple schematic of the hierarchical design of mineralized collagen fibrils, forming the most basic building block of bone [3,4]. Three polypeptide strands arrange to form a triple helical tropocollagen molecule. Tropocollagen molecules assemble into collagen fibrils in a hydrated environment, which mineralize by formation of hydroxyapatite (HA) crystals in the gap regions that exist due to the staggered geometry. Mineralized collagen fibrils combine with the extrafibrillar matrix to fibril arrays, which form fibril array patters [3,4,124]. Typically, a total of seven hierarchical levels are found in bone. The present work is limited to the scale of mineralized fibrils, with the objective to provide insight into the most fundamental scales of bone and its deformation mechanics under tensile loading.

3.6.1. Introduction

Fig. 43 depicts the geometry of the nanostructure of bone, showing several hierarchical features from atomic to microscale. The smallest scale hierarchical features of bone include the protein phase composed of tropocollagen molecules, collagen fibrils (CF) as well as mineralized collagen fibrils (MCFs). Tropocollagen molecules assemble into collagen fibrils in a hydrated environment, which mineralize by formation of hydroxyapatite (HA) crystals in the gap regions that exist due to the staggered geometry. Fig. 44 shows a schematic of nascent bone formation by proposed by Landis and coworkers [142], based on TEM analyses of structural changes during bone formation. These experimental studies provided strong evidence for the nucleation of mineralization in the gap regions in the collagen fibril.

MCFs arrange together with an extrafibrillar matrix (EFM) to form the next hierarchical layer of bone. While the structures at scales larger then MCFs vary for different bone types, mineralized collagen fibrils are highly conserved, nanostructural primary building blocks of bone that are found universally [9,28,37,124,141,143–145]. Each MCF consists of tropocollagen molecules with approximately 300 nm length, arranged in a characteristic staggered pattern. Gap regions in this arrangement are filled with tiny hydroxyapatite (HA) crystals. The work reviewed in this section is limited to the scale of mineralized fibrils, with the objective to provide insight into the most fundamental scale of bone and its deformation mechanics under tensile loading.
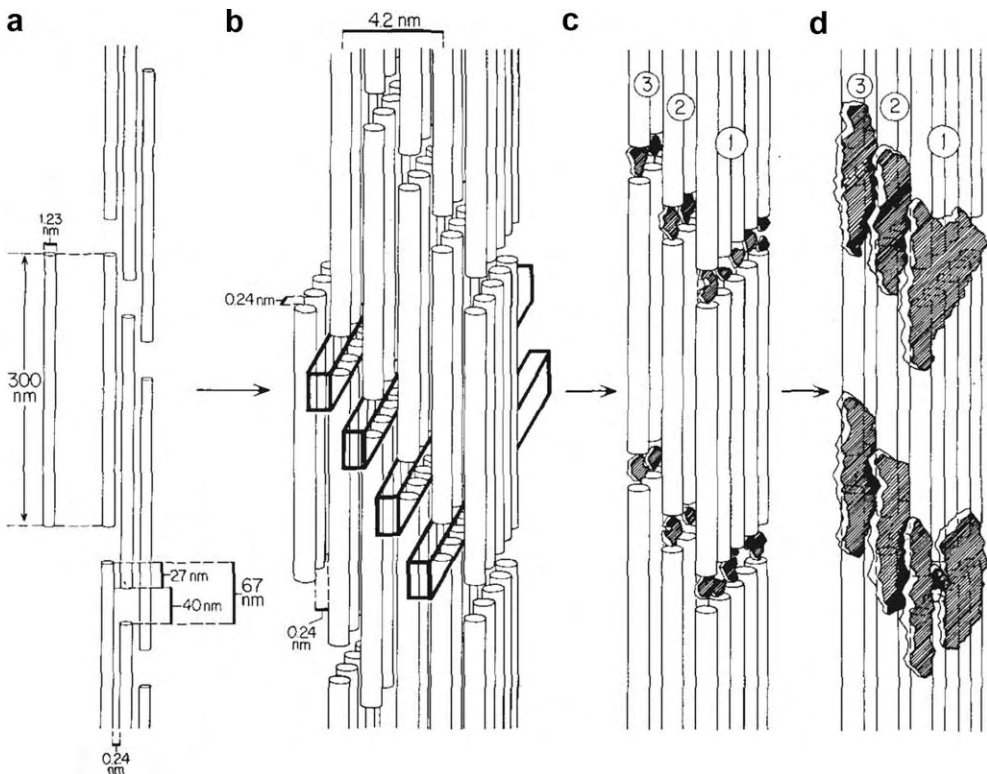
Fig. 44. Schematic of nascent bone formation by proposed by Landis et al. [142], based on TEM analyses of structural changes during bone formation. The plot illustrates the formation of nascent bone from unmineralized collagen fibrils. The study led to the conclusion that upon mineralization, the more electron dense mineral can be seen as striations within the collagen fibrils. Large and small mineral deposits grow in the hole zones. Reprinted with permission from Ref. [142]. Copyright © 1993 Elsevier Ltd.

The mechanical properties of bone have received significant attention. Particular effort has been devoted to understanding the mechanisms that make bone tough. Whereas some experimental evidence suggests that the sub-micrometer structure is critical for the mechanical properties of macroscopic bone [88,119,120,146], other results indicate that macroscopic mechanisms such as crack bridging or micro-cracking [146] contribute to the toughness of bone. Concepts such as sacrificial bonds and hidden length suggest toughening mechanisms that occur between different mineralized collagen fibrils [143,144,147].

However, due to the structural complexity of bone, the analysis and quantification of the deformation mechanisms at the ultra-scale of individual mineralized collagen fibrils (MCFs) remains an area that is not well understood. Limited knowledge exists whether, and if yes, how molecular scale mechanisms within single MCFs contribute to the toughness and stiffness of bone, as well as for its ability to repair itself. The effect of precipitating mineral crystals during bone formation remains unknown.

Most theoretical and computational analyses of bone have been carried out at continuum scales, neglecting the particular complexities of molecular interactions and chemistry. To date, there exists no molecular model of the nanostructure of bone that enables a rigorous linking between molecular and tissue scales.

In the studies reviewed here, a simple molecular model of MCFs is utilized that provides a fundamental description of its nanomechanical properties. The small and large deformation mechanics of a pure collagen fibril and a mineralized collagen fibril are systematically compared. Please see Fig. 45 for a schematic comparison of the two model systems. Corresponding experimental pictures are shown in Fig. 46, showing high resolution images of the primary components of bone, comparing a collagen fibril and a mineralized collagen fibril, as reported in [139]. Both structures are subject to identical tensile loading in the direction of the molecular axis of the tropocollagen molecules. Comparing the deformation mechanisms, stress–strain behavior and energy dissipation reveals insight into the effect of mineralization. It is found that mineralization leads to an increase in stiffness, yield stress, fracture stress and energy dissipation. The study reveals how a highly dissipative, yet strong material can be formed out of a soft polymeric collagen phase and hard, brittle HA by arranging molecules and crystals at characteristic nanostructured length-scales.

3.6.2. Molecular model

The computational model is developed with the aim to elucidate generic behavior and deformation mechanisms of MCFs. The model is a simple 2D system of a mineralized collagen fibril, based on the model of pure collagen fibrils discussed above, here extended to describe an additional mineral phase. The CF consists of a staggered array of tropocollagen molecules. The gap zones in the CF [3] are filled with a single crystal that has a planar size of approximately 28 nm × 1.4 nm, filling the entire open space, resembling the presence of the HA phase, leading to a MCF.
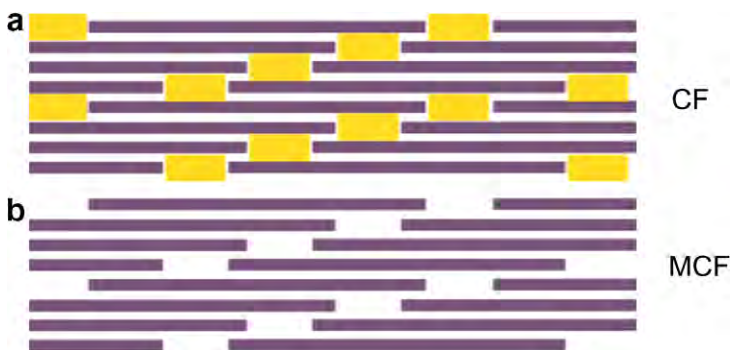
Fig. 45. Overview of the two structures considered here: subplot (a) shows a CF and subplot (b) shows a MCF (original study reported in [139]). The structures are loading in uniaxial tension along the axis of the TC molecules.
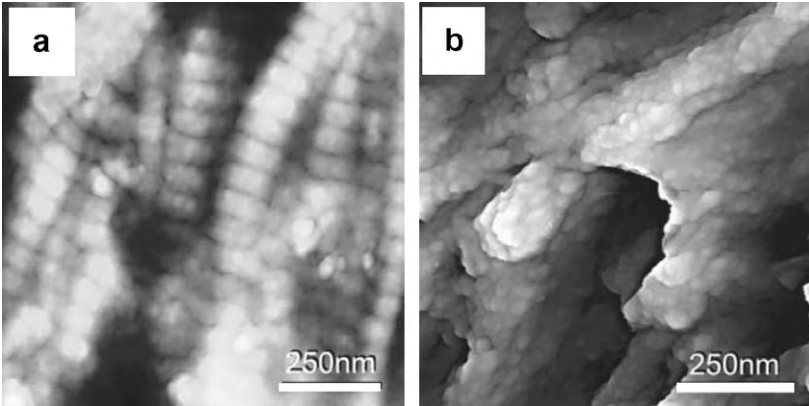
Fig. 46. High resolution images of the primary components of bone, comparing a collagen fibril and a mineralized collagen fibril. Subplot (a) shows an AFM image of the outer surface of human trabecular bone. The outer surface consists mainly of collagen fibrils showing the characteristic 67 nm D-banding. Subplot (b) depicts an AFM image of mineralized fibrils on a fracture surface of bovine trabecular bone. Here, in the interior of a trabecula, the collagen fibrils are mineralized with hydroxyapatite particles to form the primary building block of bone. Reprinted with permission from Ref. [148]. Copyright © 2006 Elsevier Ltd.

It is noted that as bone is formed, mineral crystals exceed the size of the gap region and penetrate into the collagen phase. The present study does not include these effects, and is thus limited to the early stages of bone formation ("nascent bone"). The term "hydroxyapatite" (HA) is used to refer to the mineral phase in bone, although this component is also referred to as "dahllite" or carbonated apatite [149].

The study reviewed in this section is the first molecular scale model of the nanostructure of bone. The experimental paper by Gupta et al. [149] considers a geometry that is closest to the one reviewed here. However, to the best of our knowledge no other molecular scale model of the nanostructure of bone has been reported thus far. The studies are carried out using a reactive mesoscopic model describing tropocollagen molecules as a collection of beads interacting according to interparticle multi-body potentials; the pure collagen model reviewed earlier is extended here to describe the HA phase and HA–TC interactions. The equations of motion are solved according to a classical molecular dynamics (MD) scheme implemented in the LAMMPS simulation code [150].

The total potential energy of the model is

$$E = E_T + E_B + E_{TC} + E_{HA} + E_{HA-TC}. \tag{38}$$

The difference to the model for pure collagen fibrils is extended here by two terms, E_{HA} (potential function for HA phase) and E_{HA-TC} (interaction potential HA–TC). The terms E_T and E_B are only applied within the tropocollagen molecules, as described above. Intermolecular interactions between tropocollagen particles, HA particles and between HA and tropocollagen particles are described by a Lennard–Jones 12:6 (LJ) potential

$$\phi_{TC/HA/HA-TC}(r) = 4\varepsilon\left(\left[\frac{\sigma}{r}\right]^{12} - \left[\frac{\sigma}{r}\right]^{6}\right) \tag{39}$$

with σ as the distance, and ε as energy parameter, defined separately for different materials.

The model parameters for tropocollagen properties and interactions are identical as in the studies reviewed above. The parameter $\varepsilon_{MF} = 26.72$ kcal/mol, with $\sigma_{MF} = 3.118$ Å, which leads to a Young's modulus of approximately $E_{HA} \approx 135$ GPa. These parameters are determined by fitting against the experimentally determined elastic modulus of HA. This 2D LJ model leads to extremely brittle material behavior. Since the objective in developing a model for the HA phase was to (i) provide a realistic description of the stiffness and (ii) of the fracture properties, the above described model provides a good approximation for the relevant physical and mechanical properties of the HA phase.

Interactions between the HA crystal and tropocollagen molecules are described by a LJ potential with σ_{MF-TC} = 7 Å. The adhesion strength in this potential is chosen to be ε_{MF-TC} = 25 kcal/mol for all HA–TC interactions, while the beginning and end of each tropocollagen molecule interacts with ε_{MF-TC} = 15 kcal/mol. The distinction of interaction mimics weaker adhesion between HA and tropocollagen at the head of each tropocollagen molecule due to smaller contact area. The choice of these parameters corresponds to interface surface energies of $\gamma_{MF-TC} \approx$ 0.375 and 0.225 J/m^2.

Classical MD is used to solve the equations of motion by performing a continuous energy minimization of the system as the loading is increased, with a time step of 55 fs. After energy minimization and relaxation of the initial structure, loading is applied by displacing a thin layer of particles at the ends of the system with a strain rate 7.558E−8 per integration step. Periodic boundary conditions are applied in the direction orthogonal to pulling mimicking an infinitely large fibril, subject to uniaxial tensile loading.

It is emphasized that this simple model of the molecular and physical behavior of the nanocomposite is designed to deliberately avoid modeling the atomistic details of bonding within the HA crystal or across the HA–TC interface. However, it enables one to model the inhomogeneous stress and strain fields as well as the fracture behavior. Thus, the model system enables some first fundamental insight into the nanomechanics of mineralized fibrils.

3.6.3. Computational results: elastic, plastic regime and fracture

Fig. 47 plots the stress–strain response of a pure collagen fibril (CF) compared with that of a MCF, under tensile loading, for tensile strains up to 50%. The stress–strain response for CF and MCF are qualitatively and quantitatively different, indicating that precipitation of HA crystals during bone formation significantly alters the material response. The MCF features a larger strength and much increased energy dissipation under deformation. Plastic deformation starts at approximately 6.7% tensile strain for the mineralized fibril, whereas it occurs at approximately 5% tissue strain in the case of a pure CF. Further, the MFC shows significant softening at larger strains, with a characteristic saw-tooth shaped stress–strain curve due to repeated intermolecular slip. The mineralized fibril features a higher stiffness than the pure collagen fibril.

Fig. 48 shows snapshots of the molecular geometry under increasing tensile load, clearly showing the deformation mechanism of intermolecular slip. Fig. 48a shows snapshots of the deformation mechanisms of pure CF. Fibrillar yield is characterized by intermolecular slip (see red circle highlighting a local area of repeated molecular slip). Slip leads to formation of regions with lower material
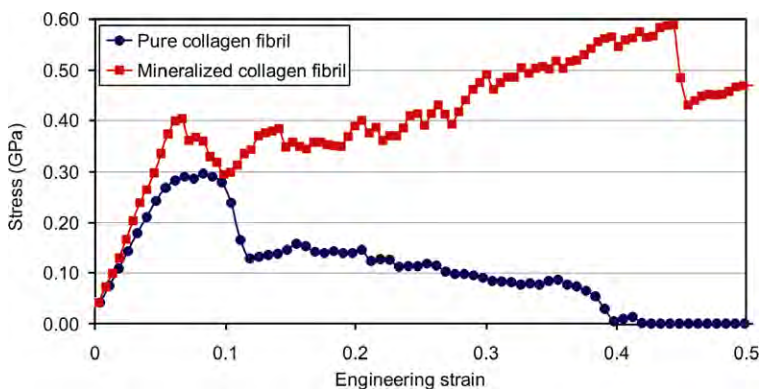
Fig. 47. Stress–strain response of a mineralized collagen fibril (MCF) and a non-mineralized, pure collagen fibril (CF), as reported in [139]. The plot shows the stress–strain curve, for the entire deformation up to tensile tissue strains of 50%. It is apparent that the MCF features a larger strength and much increased energy dissipation under deformation. Plastic deformation starts at approximately 6.7% tensile strain for the mineralized fibril, whereas it occurs at approximately 5% tissue strain in the case of a pure CF. The MFC shows significant softening at larger strains, with a characteristic saw-tooth shaped stress–strain curve due to repeated intermolecular slip. The mineralized fibril is stiffer than the pure collagen fibril.
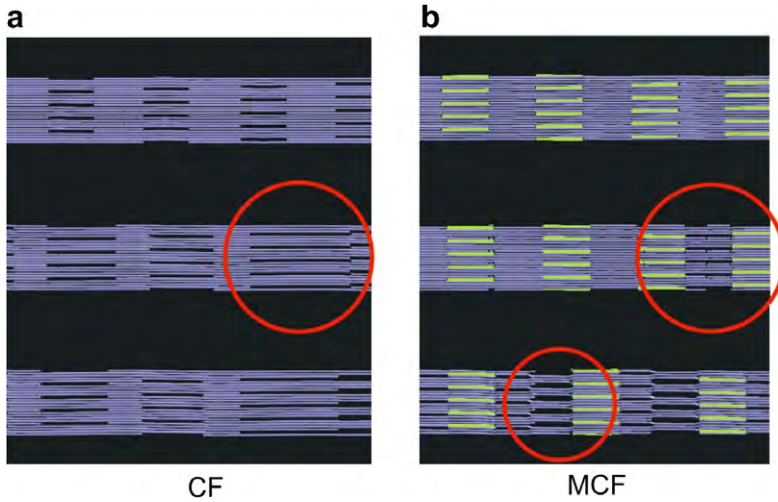
Fig. 48. Molecular geometry of plastic deformation in a pure collagen fibril and a mineralized collagen fibril, as reported in [139]. Subplot (a): Snapshots of the deformation mechanisms, pure CF, for increasing strain. Fibrillar yield is characterized by intermolecular slip (see red circle highlighting a local area of repeated molecular slip). Slip leads to formation of regions with lower material density. Subplot (b): Snapshots of the deformation mechanisms, MCF, for increasing strain. Slip initiates at the interface between HA particles and TC molecules. Slip reduces the density, leading to formation of nano-scale voids. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

density. Fig. 48b displays snapshots of the deformation mechanisms of MCFs. Slip initiates at the interface between HA particles and tropocollagen molecules. Slip reduces the density, leading to formation of nano-scale voids. The details of the differences between the MCF and CF and associated deformation mechanisms will be discussed in the following sections.

Up to the onset of yield, the mechanical response of the MCF is elastic (within the range of normal physiological function). In this regime, the increase of tissue strain leads to continuous increase of the strain in each tropocollagen molecule. Both MCF and CF display a linear-elastic regime for small deformation. The MCF is 36% stiffer than the CF (Young's modulus of a CF is 4.59 GPa versus 6.23 GPa for a MCF). The presence of HA crystals further changes the onset of plastic deformation, characterized by a sudden drop in the stress–strain response. Whereas the CF begins to yield at approximately 5%, the MCF yields at 6.7%. This represents a 34% increase in yield strain. These results are summarized in Table 4. The data shown in Table 4 was generated based on the stress–strain curve shown in Fig. 47.

Table 4
Quantitative comparison of the deformation and fracture properties of CFs and MCFs (tensile strains up to 50%)

Property	CF	MCF	Ratio value, MCF/CF
Young's modulus (small deformation)	4.59 GPa	6.23 GPa	1.36
Yield strain	5%	6.7%	1.34
Maximum stress	0.3 GPa	0.6 GPa	2
Failure mode	Molecular slip	Molecular slip and slip along HA–TC interface	N/A
Energy dissipation	3.83 GJ/m^3	19.48 GJ/m^3	≈5
Ratio of TC strain vs. tissue strain	Approaches 87% at yield	Approaches 100% at yield	1.15
Ratio of HA strain vs. tissue strain	N/A	11%	N/A
Size of fracture process zone	≈150 μm	≈200 μm	1.33

After onset of yield, the stress does not drop to zero rather quickly, but instead remains at levels of 0.4 GPa, with a slight increase with strain, approaching 0.6 GPa. After onset of yield, the MCF becomes softer, that is, less force is required for identical extension. The reduction in slope is due to the fact the strain in some of the tropocollagen molecules does not increase with tissue strain, since an increasing number of bonds to HA crystals and other tropocollagen molecules are broken. When fracture occurs, all molecular bonds inside the MCF are broken, and the strains inside each component drops to zero.

Both CF and MCF yield by intermolecular slip. Repeated glide between tropocollagen molecules and between HA particles and tropocollagen molecules initiating by slip at the HA–TC interface enables a large regime of dissipative deformation after beginning of yield. In the case of the MCF, larger stresses can be maintained after initiation of slip due to additional resistance to slip at the interface between the tropocollagen molecules and HA particles. Mineralization of the CF leads to a factor of two increase strength. Most importantly, mineralization leads to a five-fold increase in energy dissipation.

An analysis of the strain field within tropocollagen molecules and HA platelets reveals that the observations discussed in the previous section can be explained by molecular nanomechanical mechanisms, since mineralization significantly changes the strain distribution. In pure CF, the tissue strain (applied strain) is always larger than the strain within tropocollagen molecules, reaching approximately 87% immediately before yield begins. In MCF, the tissue strain and tropocollagen strain remain much closer during deformation, approaching similar strain levels at the onset of plastic deformation. This is due to the good adhesion between HA platelets and tropocollagen molecules, which hinders initiation of intermolecular slip. The HA phase carries up to 11% of the tissue strain. Such large tensile strains correspond to a stress of several GPa. Evidence for the molecular failure mechanisms of intermolecular slip is also found in experiment, as for instance shown in [124,149].

3.6.3.1. Comparison with experimental results. The most direct comparison of our molecular simulation results can be done with a recent experimental study reported in [149]. We briefly summarize the main findings. By carrying out tensile tests of MCFs obtained from mineralized turkey leg tendon, it was shown that the stiffness increases continuously with increasing mineral content. It was shown that different mineralization stages correspond to stiffness values from 500 MPa (low mineral content) to 3 GPa (high mineral content). Further, the experiments revealed that the stress–strain behavior shows a characteristic softening behavior: an initial, rather stiff regime persists up to strains of approximately 3%, which is followed by a significant softening. The observed stress–strain response is reminiscent of a bilinear softening stress–strain behavior.

The mechanical behavior calculated based on the molecular model of MCF reviewed here agrees with several observations made in experiment [124,149]. For example, the MCF yield strain is in somewhat close to experimental results (3% in experiment and 6.7% in simulation). Further, the reduced slope at large strains agrees qualitatively with experiment. The finding that Young's modulus increases is in qualitative accordance with experiment comparing mineralized and non-mineralized tendon CFs. Experimental results suggest a continuous increase in Young's modulus under mineralization, ranging up to a factor of 3 for high mineral content [134,141]. Further, the finding that component strains are smaller than tissue strains are consistent with experiment in bone [151] and tendon [140], albeit these studies were carried out at larger scales. The results prove that this is also true at the smallest hierarchical scale of bone.

3.6.3.2. Local yield protects the integrity of the entire structure. The fracture process zone describes the geometric extension of the region around a crack-like flaw that undergoes plastic deformation when the specimen is loaded. For brittle materials, the fracture process zone is extremely small, limited to a few atomic distances. In ductile materials, the fracture process zone can become very large, approaching the specimen dimensions.

The size of the plastic, dissipative zone for a crack oriented orthogonal to the alignment direction of tropocollagen molecules can be approximated as

$$\xi_{cr} \approx \frac{2\gamma E}{\sigma_{max}^2}, \tag{40}$$

where $\sigma_{\max}$ is the maximum fracture stress, γ is the energy necessary to create a new surface, and E is Young's modulus. Eq. (40) shows that the size of the fracture process zone is proportional to the fracture surface energy. Thus, increases in the dissipative work required to create two new surfaces lead to much larger plastic zones.

Based on the parameters for extracted from molecular simulation (numerical values for E and $\sigma_{\max}$ are given in Table 4; γ = 11,460 J/m^2) $\xi_{cr} \approx 400$ μm for MCFs.

This length scale has another important implication: for any defect smaller than ξ_{cr}, fracture will not be controlled by the presence of this flaw. The material is insensitive to the presence of crack-like flaws below this characteristic defect dimension.

Notably, this length scale is on the same order of magnitude as small micro-cracks typically found in bone, with characteristic dimensions of several hundred micrometer diameter [140,152]. It may also play a significant role in bone remodeling. Bone is remodeled in so-called basic multi-cellular units – BMUs – a combination of osteoclasts and osteoblasts forming small cavities inside the tissue. It has been shown that BMUs represent defects with dimensions of approximately 200 μm, thus on similar orders of magnitude as ξ_{cr} (see also Refs.[153,154]). Thus, the particular properties of MCF could be a vital component in allowing the presence of BMUs inside the tissue without compromising its strength. Further, by limiting the dimensions of individual MCF in the hierarchical structure, failure will occur homogeneously within each MCF, with plastic strains distributed over the entire geometry [155].

This analysis provides insight into how the particular MCF structure contributes to toughness by comparing a pure HA crystal. The surface energy of pure HA ranges from 0.3 to 1.6 J/m^2 (Density Functional Theory calculations [155]), leading to a rather small fracture process zone on the order several nanometers. Thus, any larger crack-like defect will lead to catastrophic failure. The estimate for γ of a MCF obtained from the MD studies of a fibril is several orders of magnitudes higher. Even though the modulus is much reduced in comparing a pure HA crystal with the MCF, the significant increase in γ outruns the reduction in modulus. Further, the fact that $\sigma_{\max}$ is smaller further leads to increase in the length scale. In comparison with a CF, the MCF has a larger fracture process zone due to an increase in γ as well as an increase in E. The effect of these two parameters outruns the effect of a smaller $\sigma_{\max}$.

3.6.4. Discussion

The work reviewed in the previous sections overcomes the limitations of the existing models of bone by explicitly considering tropocollagen molecules interacting with HA phases, providing a physics based material description that enables one to make direct links between molecular structure, topology and fracture behavior. It is found that the nanostructural arrangement of the MCF is the key to its mechanical properties, notably by allowing molecular slip as a major toughening mechanism. In the following sections, implications of the findings for the understanding of bone formation and bone mechanics are discussed.

3.6.4.1. Hierarchical toughening mechanisms. Past research has revealed that toughening occurs at different scales. Our studies and results from investigations at other length-scales suggest that each level in the hierarchy of bone may be designed to provide optimal toughness, thus being capable of taking advantage of nano-scale molecular and crystal properties, at larger scales.

The behavior discussed in this article is qualitatively similar to that suggest by the sacrificial bond model (see, e.g. [143]). However, the mechanism described here operates at a smaller length scale in bone's structural hierarchy, and has a different nanostructural origin; it is closely linked to the particular staggered molecular structure of the collagen fibrils and does not involve presence of metal ions. It is found that at the level of individual MCF, intermolecular slip is a major mechanism of dissipation (see, for instance in Figs. 47 and 48).

To enable this dissipation mechanism, the adhesion energy between HA crystals and tropocollagen molecules must be in a critical regime. This regime is characterized by the following condition: it must allow strengthening by making it more difficult to initiate molecular slip, but it must be small enough so that covalent bonds inside the tropocollagen molecules are not broken. Interface energies on the order of magnitude that allow for these deformation mechanisms correspond to ionic interactions

across the TC–HA interface. Indeed, ionic interactions have recently been suggested based on NMR studies of a TC–HA interface in physiological bone.

Pure vdW or H-bond interactions would lead to adhesion energies of approximately 0.01 J/m^2. This would be insufficient to make MCFs stronger or increase its toughness, thus rendering the presence of minerals in the gap regions insignificant. In the other extreme case, increasing γ_{MF-TC} to values corresponding to covalent bonds ($\gamma_{MF-TC} > 1$ J/m^2), the deformation mechanics changes so that plastic yield does not set in until tropocollagen molecule rupture occurs, leading to a shutdown of the toughening mechanism. The large aspect ratio of the mineral platelets leads to large shear forces between the tropocollagen molecule and the HA crystal, since $F_{shear} \sim A_c$ ($A_c \sim L_c$ is the contact area between tropocollagen and HA).

The analysis of the strain distribution inside the MCF shows that the stress in HA platelets approaches several GPa. However, macroscopic HA crystals that break at 0.1% tensile strain and stresses as low as 65 MPa [145,153]. It was shown in earlier studies that by reducing the size of a HA crystal to dimensions below 30 nm, the strength of the crystal approaches the theoretical value, even under presence of cracks or other defects [145,153,154]. Under flaw-tolerant conditions, the material does not sense the existence of defects and is thus capable of reaching its theoretical strength. Thus, the flaw tolerance concept could be a possible explanation for the fact that mineral platelets can sustain large stresses that approach one GPa, without fracturing.

3.6.4.2. Molecular design scenarios. The understanding facilitated by the model enables one to develop different design scenarios. As reported in another study of collagenous tissue [27], high cross-link densities in a pure CF without HA phase make the material stronger, but lead to a brittle polymer with low toughness and low stiffness. Such behavior is undesirable for the physiological role of bone.

As shown in the molecular analyses reviewed here, adding very stiff ceramic platelets inside the collagen fibril represents a strategy to insert high densities of covalent chemistry in order to make the material stiffer and stronger without compromising toughness. The addition of mineral platelets allows the material to yield under large load in order to protect the entire structure. The molecular role of HA platelets in MCFs thus appears to be related to the increase of the strength by providing a larger energy barrier against intermolecular slip. At the same time, presence of HA platelets increases the dissipative nature of large strain deformation of MCFs. Also, the molecular arrangement of MCFs allows to achieve a good weight–strength efficiency, since the dominating protein phase is lighter than the HA phase.

The length of tropocollagen molecules controls the mechanical behavior of CFs [27], and we expect a similar behavior for MCFs. Short tropocollagen molecules lead to reduced strength and MCFs may become rather brittle. Long tropocollagen molecules are vital to yield large toughness, as they provide a means to enable long deformation paths with large slipping inside the material. The physiological significance of toughness may explain why extremely long tropocollagen molecules in MCFs is a highly conserved molecular feature. However, if tropocollagen molecules become too long, utilization of the intermolecular glue becomes inefficient [27]. As shown in another study (see the discussion above), molecular lengths at approximately 200 nm provide an optimal basis for CFs [27]. Further analysis of the dependence of mechanical properties of MCFs on the tropocollagen molecule length is left to future studies.

The molecular toughening mechanism reviewed here unifies controversial attempts of explaining sources of toughness of bone, as it illustrates that both crack tip mechanisms [14–16] and flaw tolerance concepts [153] play a key role in the mechanical response of bone under extreme load.

The results reviewed here show the significance of the nano-scale TC–HA patterning as a toughening mechanism at the nano- and micro-scale. However, we would like to point out that the role of MCFs for the deformation and fracture mechanics of bone is only part of the larger range of deformation mechanisms that operate at larger length-scales. In particular, structural features at larger length-scales of bone, the dependence of material properties on time (e.g. via osteoblast and osteoclast cells [140,159]), and extrafibrillar matrix properties [143] are crucial for the macroscopic mechanical properties of bone.

The molecular analysis reviewed here shows that the particular properties of MCF allow to tolerate cracks at dimensions of several hundred micrometers; this may be critical to enable operation of basic molecular units (BMUs) in repair of bone, which require the presence of

small cavities inside the tissue [140]. The mechanical properties of a scaffolding material can influence the growth rate and quality of the bone tissue, providing evidence that not only chemical growth factors, but also the nano- and micro-mechanical material properties play a role in tissue development [25]. Further, it has been shown that the presence of a stiff matrix directs stem cell differentiation towards osteoblasts [25,159]. Thus, the increase in stiffness due to mineralization – as shown in Fig. 47 and Table 4 – could be a critical aspect during formation of nascent bone.

3.6.5. Conclusion: molecular nanomechanics of bone

The studies suggest that the mechanical properties of CF change significantly after mineralization. Whereas pure tropocollagen fibrils are soft and the HA minerals are stiff and extremely fragile, the stiffness of mineralized fibrils assumes intermediate values, but with much increased energy dissipation during deformation (Table 4 summarizes the main effects of mineralization). Important structural features in MCFs and their effects on the mechanical behavior are:

- Presence of HA crystals, to provide additional resistance against plastic deformation, to increase Young's modulus and the fracture strength.
- Adhesion forces between HA and tropocollagen remain weak enough, to allow for slip under large load instead of inducing fracture inside the tropocollagen molecules, but strong enough to provide significant strengthening.
- Characteristic nanoscopic dimensions, to utilize the intermolecular adhesion forces most efficiently.
- Further, the presence of long tropocollagen molecules in order to provide the basis for long deformation paths for high energy dissipation.

Modifications of the mechanical properties of CF under mineralization control the fracture properties of MCFs. Our analysis reveals that the particular constitutive behavior of MCFs induces a crack tip mechanism known as plastic shielding, effectively increasing the toughness of the tissue. The concept behind this mechanism is to sacrifice a small part of the structure in order to rescue the integrity of the entire structure. Presence of large yield regions on the order of several hundred micrometers leads to a more equal stress distribution under loading and enables operation of BMUs. In contrast, cracks in pure HA crystals lead to potentially dangerous large stress concentrations around flaws. The concept of equal stress distribution is known as a driving force for topology and shape evolution of natural structures such as bone and trees. The results suggest that this appears to be a universal principle that also holds at nano-scale.

Development of the yield region represents a toughening mechanism whose origin is intimately linked with molecular geometry and mechanisms, underlining the significance of nanostructure for bone properties. The results provide molecular scale explanations of experiments that show an increase in yield stress, maximum fracture stress, as well as failure strength with increasing mineral content [124,125,149,160,161].

4. Alpha-helical proteins: intermediate filaments

The next protein structure considered is the class of alpha-helical (AH) proteins. Together with beta-sheets, AH structures are the most abundant secondary structures found in proteins. The AH motif is commonly found in many structural protein networks, forming the basis for biological protein materials such as hair, wool and the cell's cytoskeleton. In particular, the AH protein motif plays an important role in biophysical cellular processes that involve mechanical signals, including mechanosensation and mechanotransduction, as well as in providing mechanical stability to cells [1,10,11,13,162–164]. For instance, AH rich intermediate filament networks forward signals from the cellular environment to the DNA [163,164], aspects that are critical for cell mitosis or apoptosis.

A particularly stable configuration of protein domains based on the AH motif, discovered for the first time in keratin intermediate filaments (IFs) approximately 50 years ago, are AH coiled-coils (CCs), protein structures that consist of a superhelical arrangement of two AHs. The difference in the molecular architecture of AHs and CCs is visualized in Fig. 49.
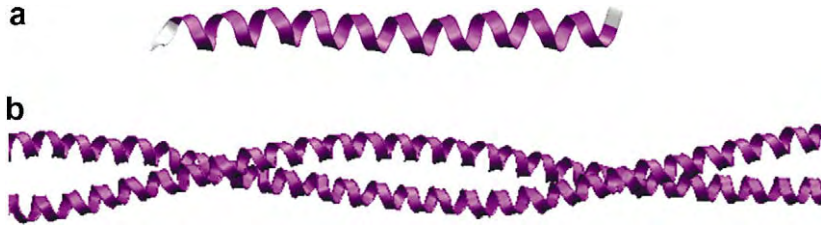
Fig. 49. Geometry of a single AH (subplot (a)) and an alpha-helical CC molecule (subplot (b)). The CC geometry represents a superhelical structure created by two AHs.

Table 5
Different types of intermediate filaments, its location in the cell and its assembly group

#	Type	Location	Assembly group
I	Acidic keratins	Cytoskeleton	AG I (heterodimer)
II	Basic keratins	Cytoskeleton	AG I (heterodimer)
III	Viementin, desmin	Cytoskeleton	AG I (heterodimer)
IV	Neurofilaments	Cytoskeleton	AG II (heterodimer)
V	Lamins	Nucleoskeleton	AG III (heterodimer)
Unknown	Phakinin, filensin	Eye lens cells	

Intermediate filaments are classified into five different types. Most of them appear in the cytoskeleton, except lamins that are found in the nucleoskeleton. Keratins, in contrast to the other IF types, assemble into heterodimers, consisting of one acidic and one basic keratin [16,26]. This table is adapted from [138].

The focus of the discussion reported here is on a particular type of a AH based protein material, the vimentin intermediate filament (IF) network in eukaryotic cells. These intermediate filaments belong to a larger class of protein materials found in cells, hair, the nuclear envelope, for instance. Table 5 provides a summary of the different classes of intermediate filaments.

Table 6 summarizes several important biological concepts used throughout this section.

Table 6
Summary of different biological terms and concepts related to IFs and AH based protein motifs

Cytoskeleton	A composite protein material inside the cell consisting of three different networks: actin filaments, microtubules and intermediate filaments. The intermediate filament network is in the focus of the review presented here. These networks connect the nucleus (nuclear membrane) with the plasma membrane and are furthermore responsible for the organization inside the cell
Intermediate filaments (IFs)	One of the three components of the cytoskeleton; mainly responsible for the large deformation behavior of the cell
Cross-bridging proteins	Cross-bridging proteins form connections inside each cytoskeletal network as well as connections between different networks (e.g. plectin proteins)
Dimer	A dimer is the elementary building block of an IF fiber. This protein consists of a head domain, a tail domain and an extremely elongated coiled-coil rod. A coiled-coil is a superhelix that consists of two alpha-helices that twist around each other (see Fig. 52)
Assembly	Individual IF dimers assemble systematically and hierarchically into filaments (see, e.g. Fig. 52). Two dimers build a tetramer, two tetramers build an octamer and four octamers build a unit length filament (ULF). Once this level of assembly is reached, ULFs ally longitudinally into long fibers
Residue	The primary structure of a protein consists of a sequence of amino acids. One residue is thus one amino acid in the polypeptide polymer
Conserved structure	A structure is conserved when parts of the residue sequence are similar or do not vary at all between the different species (e.g. human and fish). For example, certain parts of the IF sequence are very similar between different species as well as inside the IF protein family (vimentin, desmin, keratin, etc.). Conserved structures often signify a particular amino acid sequence that has proven to be particularly suitable for a specific biological task, and has thus been kept identical, or conserved, during the evolutionary process

This table is adapted from Ref. [138]. For further reference to biological concepts, we refer the reader to [1].

4.1. Background information: alpha-helical protein domains

An AH is generated when a single polypeptide chain twists around on itself, stabilized by hydrogen bonds (H-bond or HBs) made between every fourth residue, linking the O backbone atom of peptide i to the N backbone atom of peptide $i + 4$ in the polypeptide chain. Consequently, at each convolution, approximately 3.6 H-bonds are found in a parallel arrangement that stabilize the helical configuration [1].

CCs are formed by creation of a superhelical structure; this protein motif appears in approximately 10% of all proteins [165]. The CC consist of an assembly of two or more AHs in which the primary structure reveals a pronounced seven residue periodicity $(abcdefg)_n$, called a heptad repeat. Within this repeat, positions "a" and "d" are preferably occupied with nonpolar residues [166,167] such as LEU, ALA, VAL or ILE. The hydrophobic residues, concentrated on one side of the helix, are the reason why the AH proteins assemble into a CC structure: in order to avoid contact with surrounding water molecules, AHs assemble into CCs by wrapping around each other, clustering the hydrophobic side chains inside [1]. Additionally, interhelical and intrahelical salt bridges contribute to CC thermodynamic stability [168].

Even though the heptad repeat (7/2 – seven residue repeat, whereof two residues are hydrophobic) is the most common pattern for CC and is thus often referred to as the canonical CC, there exist other molecular structures, such as the 11/3, 15/4 or the 18/5 pattern. These structures result in two-, three-, four- or five-stranded CCs [162].

4.2. Vimentin intermediate filaments: structure and biological role

Here studies of the large deformation behavior of vimentin IFs, the most widely distributed type of all IFs, are reviewed. Vimentin proteins are typically expressed in leukocytes, blood vessel endothelial cells, some epithelial cells, and mesenchymal cells such as fibroblasts [1].

Together with the globular proteins microtubules (MTs) and microfilaments (MFs), IFs are one of the three major components of the cytoskeleton in eukaryotic cells [169]. The cytoskeleton plays a critical role in determining the shape and the mechanical properties of the cell, and is vital for many additional functions including protein synthesis, cell motility as well as cell division or wound healing [169–171].

Like many other biological materials, IFs are hierarchical structures with highly specific features at nano-scale. Fig. 50 depicts details of the hierarchical structure of the IF network, summarizing the structural features observed throughout several length-scales. Under tensile deformation of tissues, individual cells undergo tensile deformation, leading to a stretching of the IF network. Experimental results reported in [172] are shown in Fig. 51, clearly illustrating that the IF network undergoes large tensile deformation. Vimentin IF dimers, as shown in Fig. 52 (detailed view that also depicts the chemical structure), are the elementary building blocks of IFs. Through carefully following the various steps of assembly [173,174] it was shown how dimers associate to fibrils. Fibrils build the second level of the hierarchy. *In vivo*, these fibrils can reach a length of up to several μm and consist of 16 dimers in cross-section. The third level of hierarchy consists of three-dimensional IF-networks inside the cell, reinforcing the plasma membrane [170,175,176]. Inside the network, IF associated proteins such as plectins generate the connection between individual filaments (see Fig. 50, "Cytoskeleton").

The IF networks are connected with other cellular networks, as well as with the extracellular matrix at the plasma membrane [1]. This architecture guarantees that tensile and shear loads applied to the tissue can be carried by IF networks.

4.2.1. Biological significance of the mechanical properties of intermediate filaments

Biologically, vimentin IFs are primarily associated with mechanotransduction and with carrying passive loads applied to cells. In particular, IF networks have been shown to provide resistance to large deformation of cells [170,175,176]. It has been hypothesized that IFs are critical to provide strength to the cell under large deformation, and to absorb large amounts of energy upon a certain load by unfolding [177,178].
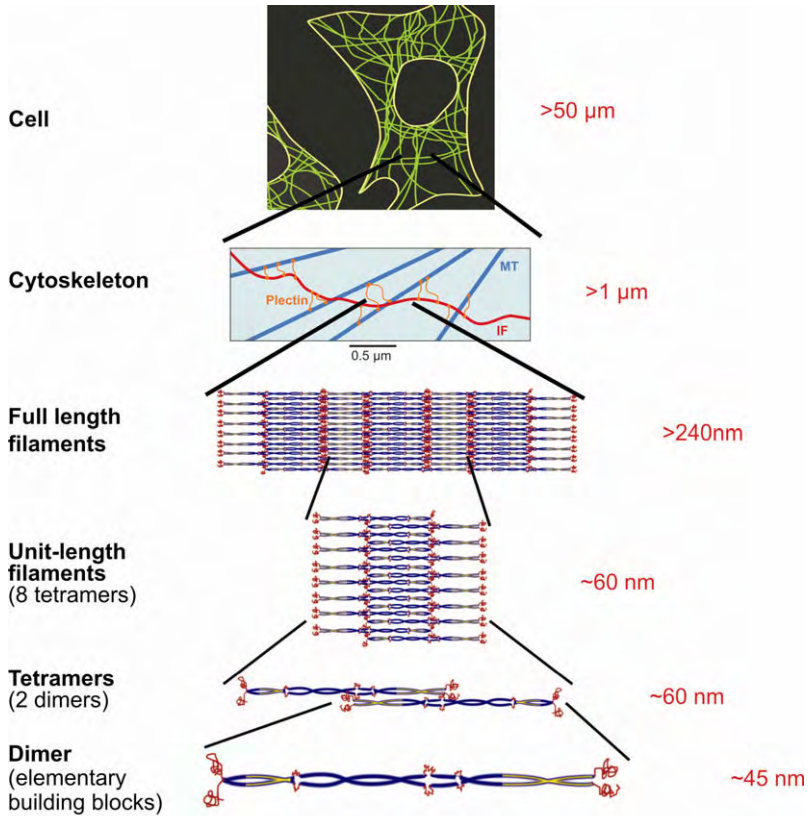
Fig. 50. Molecular hierarchical structure of vimentin IFs. Dimers, the elementary building blocks, assemble through a half-staggered, anti-parallel overlap into tetramers, which associate laterally into 13 nm thick unit length filaments (ULFs). In the next step, ULFs assembly longitudinally into more compact filaments (10 nm in diameter) with a length of more than 240 nm [1,166,175,182,188].

Under deformation of the entire IF network in a cell, each vimentin filament undergoes tensile deformation. The tensile load of filaments is carried by individual dimers at the smallest hierarchical scale (see Fig. 50). Therefore, a detailed understanding of coiled-coil dimers and their mechanical properties under small and large tensile deformation is important to provide insight into the function and mechanisms of vimentin filaments and networks. This analysis provides the starting point of the studies reviewed here.

Further, since IF filaments span from the cell's nucleus to the cell membrane and therefore interact with IF networks of other cells suggests that IFs play an important role in transmitting mechanical signals from the plasma membrane to the nucleus, where a specific response can be triggered by mechanical stimulation [163,164]. Thus, it has been suggested that IF networks may play a vital role in mechanotransduction [170,175,176].

Plakin-type cross-bridging proteins, also known as cytolinkers (e.g. plectins or desmoplakins) link all three cytoskeletal networks (MTs, MFs and IFs). These proteins attach the IFs to MTs, MFs or adhesion complexes of membranes (e.g. the cell membrane or the nuclear membrane [179]). In contrast to MTs and MFs, IFs do not participate in the dynamic functions of the cytoskeleton. Further, they do not support active transport of motor proteins such as myosin and kinesin, due to the missing polarity in
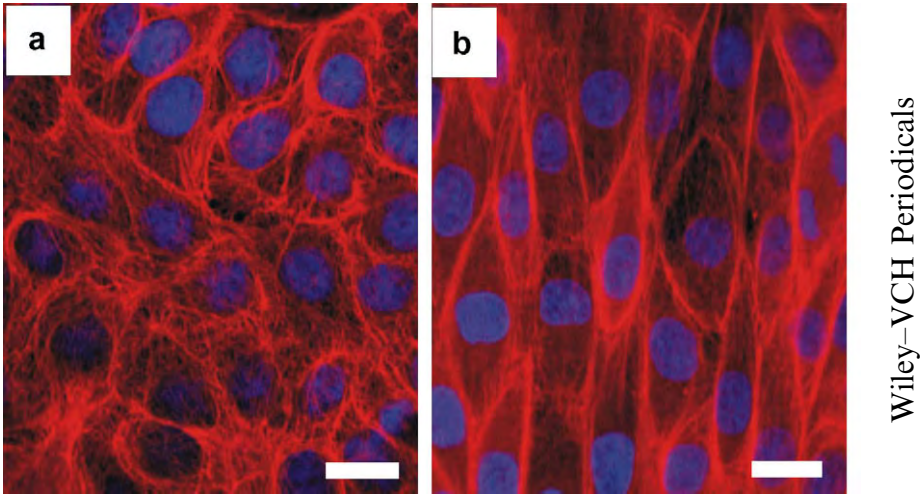
Fig. 51. The effect of large-scale uniaxial stretch on the IF network in MDCK cells. The cells were grown on collagen-coated silastic membranes and stretched using a custom cell stretcher that was mounted on a confocal microscope. Cells were fixed and stained for immunofluorescence (red = keratin IFs, blue = DNA). Subplot (a): Control cells were processed on a relaxed silastic membrane. Subplot (b): Stretched cells were fixed, stained and imaged on membranes that were held in the stretched state. Approximate uniaxial strain in stretched cells was 75%. Scale bar is approximately 25 μm. Reprinted with permission from Ref. [172].

the protein structure [166]. They do not participate in any cell movement [1]. These examples further underline the specific static-mechanical role of IFs.

Experiments have shown that IFs exhibit a highly nonlinear stress–strain relationship with a high resistance against rupture at large deformation [180]. Fig. 53 depicts experimental results of the stress–strain response of gels composed of the three cytoskeletal proteins. In these rheological experiments, gels with equal weight concentrations were sheared and the deformation response (strain) was measured. It is apparent that vimentin gels are capable of sustaining large deformation at large forces. In contrast, actin filaments rupture at low-strains but large forces, and MTs break at moderately large strains, but small forces.

Further, via shear tests of vimentin deficient cells it has been shown that they are 40% less stiff at large strains compared with wild-type cells, while their small deformation elastic properties do not change much (see plot shown in Fig. 54) [169]. These experiments strongly support the notion that the biomechanical significance of vimentin IFs lies in the *large* deformation regime, while the flexibility of IF at small strains and loading rates (compared to the properties of MFs) enables a lower mechanical resistance during cell movement, underlining the mechanical multi-functionality of IF networks.

4.2.2. Nanomechanical experimental studies

Observations in rupture experiments of single IFs have shown a dramatic change in filament diameter, which remained unchanged for hours after rupture has appeared [181]. This is an indication for a profound change in the molecular architecture under large deformation. This provided some evidence that the mechanical properties of IFs mainly depend on the nanomechanical, molecular properties of the CC dimer [181].

Finally, the mechanical role of intermediate filaments is particularly evident in diseases in which the loss of mechanical function and integrity of various tissues is associated with intermediate-filament-protein mutations [182,183]. It was shown that mutations in keratin IFs reduce the ability of these IF networks to bundle and to resist large deformation [169]. Furthermore, it has been suggested that point mutations lead to the aggregation of the cytoskeleton and extensive cell fragility in epider-
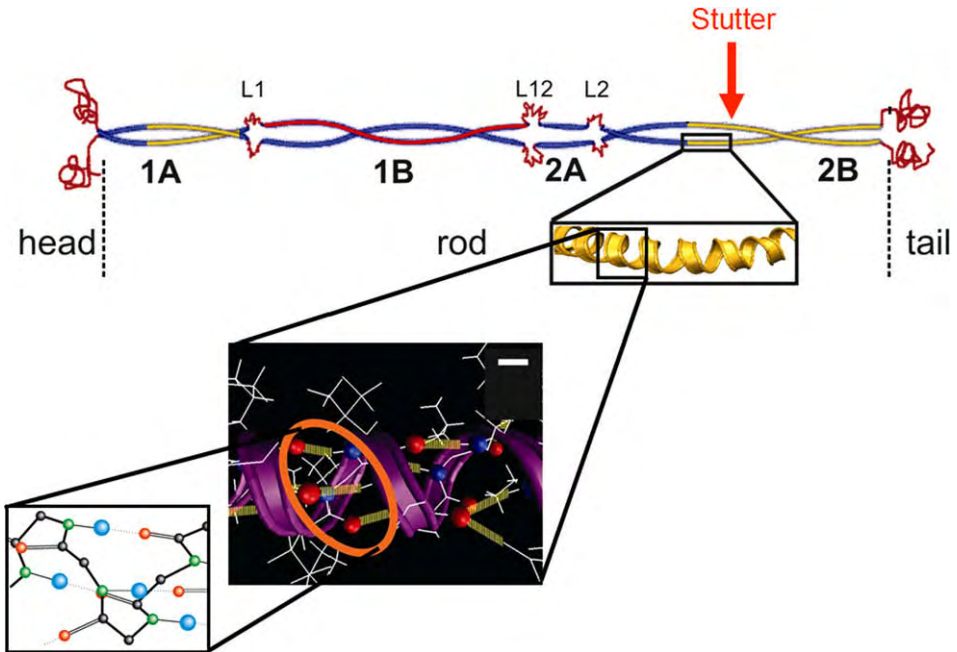
Fig. 52. Geometry of the dimeric building block in vimentin, showing the lower level hierarchical features (AH, H-bonds, chemical structure). The dimers, approximately 45 nm long, are the elementary building blocks of vimentin intermediate filaments. A dimer consists of a head, tail (plotted in red) and an elongated rod domain which is divided into four alpha-helical coiled-coils (1A, 1B, 2A, 2B) connected through linkers L1, L12, L2 (also red) [166]. The segments plotted in yellow (part of segment 1A and 2B) have been crystallized so that the atomic structure is known. The stutter is located approximately in the middle of the 2B segment (indicated with a red arrow). The molecular dynamics simulations described in this paper are performed on alpha-helices placed in the 2B segment (yellow). (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)
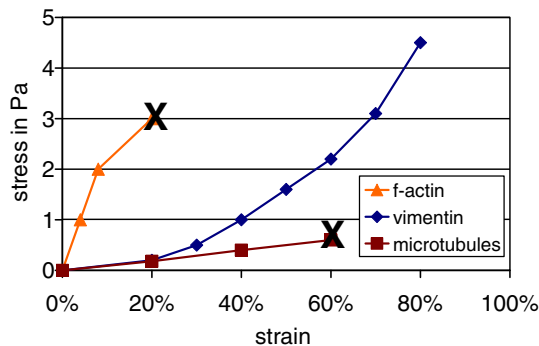
Fig. 53. Shearing experiments carried out with gels of equal weight concentration demonstrate the differences in the mechanical properties of various cytoskeletal networks (results adapted from [180]). In contrast to vimentin that sustains strains much larger than 80%, MT break at 60% strain, and MF break at 20% strain. Additionally, vimentin gels exhibits continuous significant strain hardening. We note that this stress–strain curve is different from the behavior of a single protein as reported in this paper due to the different level of assembly.

mis, heart and skeletal muscle after they are exposed to mechanical strain [184]. These examples clearly illustrate the significance of the mechanical properties of IF proteins for biological processes.
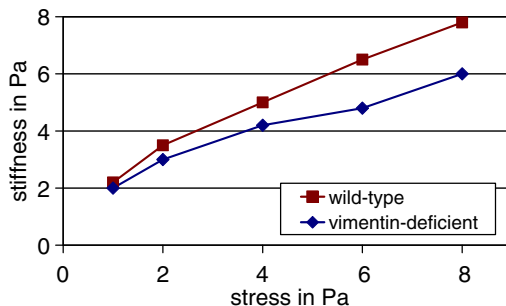
Fig. 54. Stiffness of cells as a function of stress state, comparing wild-type and vimentin deficient cells (graph based on results reported in Ref. [169]). It was shown in experiments that vimentin deficient cells are much less stiff at higher stresses than wild-type cells. These results suggest that vimentin proteins play a critical role in particular for the large deformation elastic properties of cells. It also corroborates the notion that due to the progressive strain hardening at large strains, IFs can be understood as "security belts" of the cell that operate after MTs and MFs have ruptured.

4.2.3. Structure of vimentin intermediate filament dimers

CCs are the primary building blocks of vimentin intermediate filament (IF) dimers, which are composed of a head, a tail, and an extremely elongated central rod-domain. A schematic of the vimentin dimer structure is shown in Fig. 52. The rod-like structure is 310 residues long and consists of four CCs (1A, 1B, 2A, 2B), divided by linkers (L1, L12, L2) [166,180,185].

Interestingly, all helices in the IF rod domain have different lengths (the length of the 1A segment is 26 Å, 139 Å of the 1B segment, 26 Å of the 2A segment and 158 Å of the 2B segment). However, it is notable that the lengths of each of the components are absolutely conserved for all types of IFs, even in different types of cells such as leukocytes, blood vessel endothelial cells, muscle cells or neurons [1].

A variety of discontinuities or 'defects' exist that interrupt the CC periodicity locally without destroying the overall molecular structure:

- '*Skips*' are insertions of one residue into the heptad pattern [186].
- '*Stammers*' result through an insertion of three additional residues [187–189].
- '*Stutters*' appear if four additional residues interrupt the heptad sequence [190]. Presence of a stutter results in an almost parallel run of both AHs without interrupting the CC geometry.

As of to date, little is known about the biological, mechanical or physical reasons for the presence of these defects. Thus far, only stutters have been observed in experimental analyses of the protein structure of vimentin dimers. Notably, the position of the stutter in the 2B segment is a highly conserved molecular feature (see Fig. 52, marked by an arrow) [188]. For all known types of IFs, the stutter is spaced *precisely* six heptads away from the C-terminal end of coil 2B [188].

4.3. Atomistic and molecular model

The focus of this review are atomistic simulation studies, carried out to investigate the deformation and fracture mechanisms of vimentin IF structures. All studies are based on a classical (nonreactive) molecular dynamics (MD) approach, implemented in NAMD [55] using the CHARMM22 force field [56]. All simulations were performed at a temperature of 300 K (NVT ensemble, Berendsen thermostat), with a time step 1 fs. Careful energy minimization and finite temperature equilibration of all structures are carried out before the protein is loaded. The structure obtained from the PDB is solved completely in a TIP3 water skin. In all cases, the entire protein is embedded in water, before and during deformation of the protein. This is essential to capture the correct HB rupture dynamics.

To apply forces to the molecule in order to induce deformation, the SMD technique is used [75], with a SMD spring constant is $k = 10$ kcal/mol/Å^2 are as used in the studies of collagenous proteins materials. Force versus displacement data is obtained by monitoring the time-averaged applied force

(*f*) and the position of the atom that is pulled at (*x*) over the simulation time. To apply load, C_α atoms at one end are fixed and the force is applied on the C_α atom at the other end in the AH structure, with a pulling speed *v*. The tensile boundary conditions chosen for the AH domain are closest to the physiological conditions. Several other boundary conditions have been used (changing fixed and pulled atoms, pulling at different convolutions). No changes in the AP forces have been observed, suggesting that the measured nanomechanical properties are robust with respect to changes in the boundary conditions.

Visual molecular dynamics (VMD) is used for visualization of protein structures [78]. The VMD program is also employed to carry out the analysis of the length of HBs. The rupture length of a HB is defined as 5 Å (the equilibrium length of HBs is ≈3 Å [191]). The distance from the equilibrium was chosen higher than the transition state as known from theory and experiment in order not to capture the dynamics of bond breaking and rebinding due to thermal fluctuations.

4.4. Extended Hierarchical Bell Theory: strength model for protein structures

The analysis of atomistic simulation is facilitated by the development of an appropriate theoretical framework that enables the analysis of the strength properties of protein structures.

Several theories exist that describe competing processes due to mechanically induced instabilities of protein structures. Most of them are derived from a phenomenological theory originally postulated by Bell [42], or Kramer's diffusion model [192]. Many contributions in this field have been made by Evans and co-workers [193–200], by developing a theory that describes the binding behavior of proteins in dependence of the force rate *r* (increase in force over time, here referred as the "loading rate theory"). The experimental and simulation data published by Evans led to suggest that the energy landscape is governed by multiple transition states, represented by a spline curve consisting of straight lines in the force–ln(*r*) graph. However, in most of the published work no link was made between the calculated parameters E_b and x_b, which describe the free energy landscape and the according rupture mechanisms.

The loading rate r_f is defined macroscopically as the increase in force over time, and can be seen microscopically as the rupturing force divided by the time for bond breaking (reciprocal of the off rate) [193]:

$$r_f = \frac{\Delta f}{\Delta t} = \frac{k_B \cdot T}{x_B} \cdot \omega_0 \cdot \exp\left(-\frac{(E_b - f \cdot x_b \cdot \cos(\theta))}{k_B \cdot T}\right). \tag{41}$$

The relation between the loading rate and the pulling speed is as follows:

$$r_f = K_0 \cdot v, \tag{42}$$

Here, K_0 is the pulling cantilever constant and is mostly fixed. The pulling speed is the ratio of the displacement until the angular point is reached and the time until this happens: $v = x_{AP}/t_{AP}$. By multiplying the velocity *v* with K_0 we receive, as indicated in Eq. (42) the loading rate r_f, which is consequently the force until the AP is reached (beginning of bond breaking) divided by the necessary time to generate this force. Most importantly, in SMD simulation as well as most single protein pulling experiments, the pulling speed is the controlled parameter.

In contrast to the phenomenological model, Szabo, Hummer, Dudko and co-workers [201–205] follow a slightly different approach in their "microscopic theory". They build their theory based on the assumption of only one transition state. This transition state is not only lowered with increasing external force *f* applied to the molecule (similar to the phenomenological theory), but also moved along the reaction coordinate towards the equilibrium and eventually vanishes, when the barrier disappears [202]. This results in a curved instead of a straight line in the force–log(*v*) space.

In order to derive the most appropriate information about the free energy landscape of the equilibrated system, the AFM pulling experiment or the SMD simulation (noting that both are non-equilibrium processes) need to be repeated several times. In particular for simulation studies, this might be extremely computationally expensive. The data need to be averaged out afterwards, by applying the

Jarzynski identity, which postulates that the thermodynamic free energy difference of two states equals to the work along the non-equilibrium trajectory [206,207].

The microscopic theory postulates that the simple, straight forward phenomenological regime is a very good approximation, but only over a certain magnitude of pulling velocities and lead to overestimations of the off rate at small pulling speeds. In contrast to that, the more accurate microscopic approach is more complicated, as an additional parameter (the free energy of activation) is necessary. Some of the key features of both approaches were recently summarized in [208–210]. We refer the reader to these references for further information. There exist several other theories related to this topic, which were published recently in a series of papers [211–214]. Here we focus on extensions of Bell's approach, applied to develop a strength model for hierarchical protein structures. A summary of all mathematical symbols used here is provided in Table 7.

4.4.1. Conventional, phenomenological Bell Model

In Bell's theory, the off rate χ is the product of a natural vibration frequency $\omega_0 \approx 1 \times E13 \ s^{-1}$ [42], and the quasi-equilibrium likelihood of reaching the transition state with an energy barrier E_b; x_b is the distance between the equilibrated state and the transition state. The off rate, also known from chemical reaction kinetics, is given by

$$\chi = \omega_0 \cdot \exp\left(-\frac{(E_b - f \cdot x_b)}{k_b \cdot T}\right). \tag{43}$$

The off rate describes how often a bond is broken per unit time and equals to the reciprocal of the lifetime of a bond. From this equation, it is possible to calculate the strength of a bond for given timescales of bond dissociation, τ:

Table 7
The table provides an overview over the most important mathematical symbols used in the studies focused on AH based protein materials (see Section 4)

Mathematical symbol	Unit	Description
E_b	kcal/mol	Height of the effective energy barrier at the transition state
E_b^0	kcal/mol	Energy barrier of one HB (fundamental chemical bond in the protein structures considered here)
x_b	Å	Location of the energy barrier
f	pN	Applied force
f_{AP}	pN	Force at the angular point, unfolding force
θ	° (degrees)	Angle between pulling direction and reaction coordinate
k_B	J/K	Boltzmann constant
T	K	Absolute temperature
ω_0	s^{-1}	Natural bond vibration frequency (1E13 s^{-1})
χ	s^{-1}	Off rate: bond dissociation per second
v	m/s	Macroscopically: pulling speed, microscopically bond breaking speed
v_0	m/s	Natural bond breaking speed, when no load is applied
a_1, a_2	pN	Parameter received from fitting the simulation results. Necessary for estimating E_b and x_b
ε	%	Engineering/molecular strain: displacement of the pulled atom normalized by the length of the molecule
F_{cr}, v_{cr}	pN, m/s	Force and velocity, which separate the FMD from the SDM
b	–	Number of parallel HBs
d	–	Number of HBs rupturing simultaneously
n	–	Number of parallel AHs
k	–	Number of AHs unfold simultaneously
f_v	pN	Force contribution as a consequence of the pulling speed
f_{h1}	pN	Force contribution as a consequence of the first hierarchy (parallel HBs)
f_{h0}	pN	Force contribution as a consequence of the basic hierarchy (individual HB)
r	%	Robustness, ranging from 0% to 100%
r_f	pN/s	Loading rate: increase in force over time
K_m	pN/Å	Cantilever spring constant

$$f(\tau) = \frac{1}{x_b}\left[k_B T \ln\left(\frac{1}{\omega\tau}\right) + E_b\right]. \tag{44}$$

A schematic that illustrates the main concepts (and an explanation of the variables) behind the Bell Model is shown in Fig. 55.

4.4.2. Modified Bell Model

Here we review an extension of the Bell Model to explicitly include the geometry of the protein structure, in particular the orientation of the HBs with respect to the direction of load application (see Fig. 55, lower part). This extension was originally reported in [40,138]. The goal of this model is to provide a predictive strength model for hierarchical protein structures, by modifying the Bell Model in such a way that it only includes geometry parameters and fundamental energetic properties of H-bonds.

By considering the angle θ, the angle between the direction of the reaction pathway of bond breaking (x-direction) and the direction of applied load, the energy barrier is reduced by the mechanical energy $f \cdot x_b \cdot \cos(\theta)$ due to the externally applied force (see Fig. 55, lower part). The angle can be determined by analyzing the protein geometry. Considering this angular term, the off rate is given by

$$\chi = \omega_0 \cdot \exp\left(-\frac{(E_b - f \cdot x_b \cdot \cos(\theta))}{k_B \cdot T}\right), \tag{45}$$

and describes how often a bond is broken per unit time (the reciprocal of the bond lifetime).

However, Eq. (45) does not describe the dependence of the pulling speed v (the controlled parameter in experiment and MD simulation) at which a bond breaks due to the pulling force f. Thus, Eq. (44) is modified based on the following concept: the speed v at which a bond is broken equals to the distance that needs to be overcome in order to break the bond (x_b), divided by the time for the bond breaking. Consequently, v is the product of $\chi \cdot x_b$, thus $v = \chi \cdot x_b = \Delta x/\Delta t$. Macroscopically, the pulling speed equals to the displacement Δx of the pulled atom divided by the pulling time Δt. This equation can be rewritten, leading to
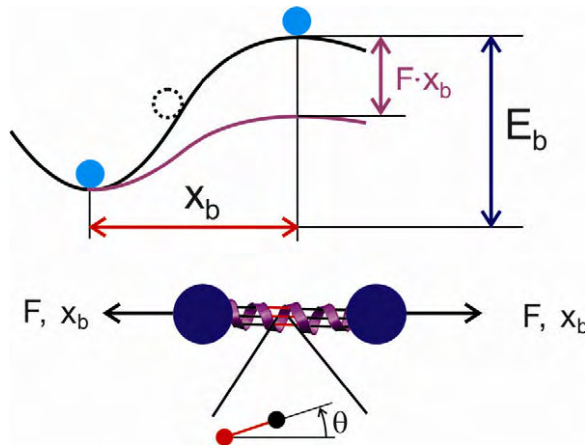
Fig. 55. The Bell Model, a statistical theory to predict the bond rupture mechanics. The graph depicts the energy as a function of deformation along a deformation variable (upper plot), along a particular pathway that leads to bond rupture; where F is the applied force, and where x_b is the displacement in the direction of the applied force. In the schematic, three HBs (indicated by the red color, lower plot) break simultaneously. Thus, x_b corresponds to the lateral displacement that is necessary to overcome the bond breaking distance of a HB, in the particular atomistic geometry present in the CC. Given that x_b^* is the distance to break a single HB, the distance $x_b^* = x_b \cos\theta$ denotes the lateral displacement at bond breaking, with the angle θ as the angle between pulling direction and orientation of the HB inside the molecule. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

$$v = v_0 \cdot \exp\left(\frac{f \cdot x_b \cdot \cos(\theta)}{k_B \cdot T}\right), \tag{46}$$

with v_0 as the natural bond breaking speed (speed of bond dissociation when no load is applied), defined as

$$v_0 = \omega_0 \cdot x_b \cdot \exp\left(-\frac{E_b}{k_B \cdot T}\right). \tag{47}$$

This modified framework enables one to calculate the force at which a bond breaks, at a given pulling velocity v:

$$f(v) = \frac{k_B \cdot T}{x_b \cdot \cos(\theta)} \cdot \ln v - \frac{k_B \cdot T}{x_b \cdot \cos(\theta)} \cdot \ln v_0 = a_1 \ln v + a_2, \tag{48}$$

where $a_1 = k_B \cdot T/(x_b \cdot \cos(\theta))$ and $a_2 = -k_B \cdot T/(x_b \cdot \cos(\theta)) \cdot \ln v_0$.

Eq. (48) predicts that the bond breaking force depends logarithmically on the pulling speed in a non-equilibrated system. The parameters a_1 and a_2 can be calculated from the parameters x_b and E_b for a certain temperature and angle. It is noted that if the free energy landscape is dominated by several transition states, each of them is characterized by a combination of E_b and x_b. This results in segments of multiple straight lines in the f–$\ln(v)$ plane. The model reduces to a phenomenological model when the $\cos(\theta)$ term is removed; the phenomenological model only contains the energy barrier E_b and x_b without structural information (the physical meaning of E_b is then an *effective* energy barrier of a protein domain). We note that a similar expression as Eq. (48) has been reported in [215].

4.4.3. Hierarchical Bell Model

Even though the modified Bell Model (Eq. (48)) explicitly considers chemical "bonds", it does not distinguish between a single chemical bond and protein architectures that include several fundamental bonds. For instance, whether a single HB ruptures or if several HBs rupture simultaneously is captured in an effective value of E_b; however, this change in mechanism is not explicitly noted in the theory given in Eq. (48).

In order to estimate the strength and the energy landscape of a protein without performing any simulations or experiments (and thus to make the model predictive), the theory is extended to explicitly consider the structural hierarchies of the protein structure with the only input parameters being the energy of a HB and the rupture distance. The AH represents a hierarchical structure, reaching from individual HBs at the lowest, atomistic level to a collection of HBs at the next higher, molecular protein scale.

The lowest hierarchy is represented by individual HBs with an E_b^0 and x_b, and the higher hierarchy consists of parallel HBs. Here we assume that b bonds in a structure are in parallel and d bonds out of these b bonds break simultaneously. Thus, $_bC_d$ (the binomial coefficient is defined as $_bC_d$) possible combinations for this rupture mechanism exists. The probability that one of these combinations constitutes a particular rupture event is one divided by $_bC_d$. Also, if d bonds break simultaneously, the total energy barrier increases by a factor d, to $d \cdot E_B^0$. This leads to the following expression for the off rate:

$$\chi_H = \omega_0 \cdot \binom{b}{d}^{-1} \cdot \exp\left(-\frac{(d \cdot E_B^0 - f \cdot x_b \cdot \cos(\theta))}{k_B \cdot T}\right). \tag{49}$$

We rewrite Eq. (49) so that the binomial coefficient appears in the exponential, which enables us to compare Eq. (49) with Eq. (44),

$$\chi_H = \omega_0 \cdot \exp\left(-\frac{\left(d \cdot E_b^0 + k_B \cdot T \cdot \ln\binom{b}{d}\right) - f \cdot x_b \cdot \cos(\theta)}{k_B \cdot T}\right). \tag{50}$$

The parameter E_b in Eq. (45) can thus be split up as follows:

$$E_b = d \cdot E_b^0 + k_B \cdot T \cdot \ln \binom{b}{d}, \tag{51}$$

where E_b^0 is the energy of a single bond and the term $k_B \cdot T \cdot \ln \binom{b}{d}$ is the contribution to the energy barrier due to the hierarchical structure. The unfolding force is

$$f(v, b, d; E_b^0, x_b, \theta) = \frac{k_B \cdot T}{x_b \cdot \cos\theta} \ln \left(\frac{v}{x_b \cdot \omega_0} \right) + \frac{k_B \cdot T}{x_b \cdot \cos\theta} \cdot \ln \binom{b}{d} + \frac{d \cdot E_b^0}{x_b \cdot \cos\theta}. \tag{52}$$

It is noted that $f = f_v + f_{h1} + f_{h0}$ where the f_v, f_{h1} and f_{h0} are the contributions to the force as a consequence of the pulling speed, the first hierarchy (number of parallel bonds, d), and the basic hierarchy (strength of bonds, E_b and x_b). This expression quantifies how the hierarchical design influences the rupture strength. In the following, we refer to this model as the Hierarchical Bell Model. This approach can be extended to three hierarchies, which enables one to predict the rupture force of a tertiary structure consisting of $2, 3, \ldots, n$ AHs, of which k unfold simultaneously. In this model, the system breaking force f^{hn} and the energy barrier E_b^{hn} of a system consisting of n hierarchies is given by

$$F^{hn} = \frac{k_B \cdot T}{x_b \cdot \cos\theta} \cdot \left[\ln \left(\frac{v}{x_b \cdot \omega_0} \right) + \ln \binom{b_n}{k_n} + \sum_{i=2}^{n} \binom{b_i}{k_i} \ln \binom{b_{i-1}}{k_{i-1}} + \frac{k_1 \cdot E_b^0}{k_B \cdot T} \right] = F_v + \sum_{i=0}^{n} F_{hi} \tag{53}$$

and

$$E_b^{hn} = k_i \cdot E_b^0 + k_B \cdot T \cdot \left[\sum_{i=2}^{n} \binom{b_i}{k_i} \cdot \ln \binom{b_{i-1}}{k_{i-1}} + \ln \binom{b_n}{k_n} \right] \tag{54}$$

In Eqs. (53) and (54) the variable b and d from previous Eqs. (49)–(51) were replaced by b_1 and k_1. These equations now enable us to predict the unfolding force at any pulling speed, once the structural geometry and the energy landscape of a single H-bond is known.

Additionally, hierarchical cross-scale interactions appear in this expression, represented by the variables k_i and b_i. In other words, f_{h0} (hierarchical level zero) depends on k_i from higher hierarchical scales ($i \geqslant 1$), that is, $f_{h0} = f(k_1, k_2, \ldots, k_n)$. Further, the coefficient b_i does not only appear in but also in h_i, so that $f_{hi-1} = f(b_i)$. Fig. 56 illustrates the meaning of the variables that characterize the hierarchical system of H-bonds in alpha-helical protein structures.

Due to the generic approach, these equations are valid for any protein structure that consists of several parallel bonds on the basic hierarchical scale, up to systems with several sub-elements on much higher hierarchical scales (for example beta-sheets or beta-helices). However, the model has thus far only been applied to alpha-helical protein structures. This could be explored in future works.

4.5. Atomistic modeling of deformation of AH and CC protein domains

In this section, we review a series of MD simulations to probe the nanomechanical behavior of CC structures from the vimentin dimer. These results are based on the studies reported in [138].

We begin with a review of the literature in this field. In recent years, a variety of AH structures and CCs have been studied in experiment as well as in simulation [216–231]. AFM experiments on single molecules of double-headed myosin, single-headed myosin as well as CC tail fragments were reported [216]. It was found that the transition to unfolding of the protein structure (in the following referred to as 'angular point' (AP) or 'unfolding force') begins at strains of about 20% of stretched protein length. Furthermore, it was shown that myosin is a very elastic protein, with almost no hysteresis at small pulling rates. Even after the CC structure has been unfolded completely under application of mechanical forces, it refolds again to its initial configuration in less than 1 s. Some of the characteristics of myosin deformation during tensile tests were also observed in MD simulations that were carried out on parts of the CC structure [219]. In addition to the CC simulation reviewed in the previous paragraph, MD simulations have been carried out on single AHs [217,219–221]. By applying tensile loads, unfolding of the helical structure has been observed after a short steep increase in force. Thereby, the
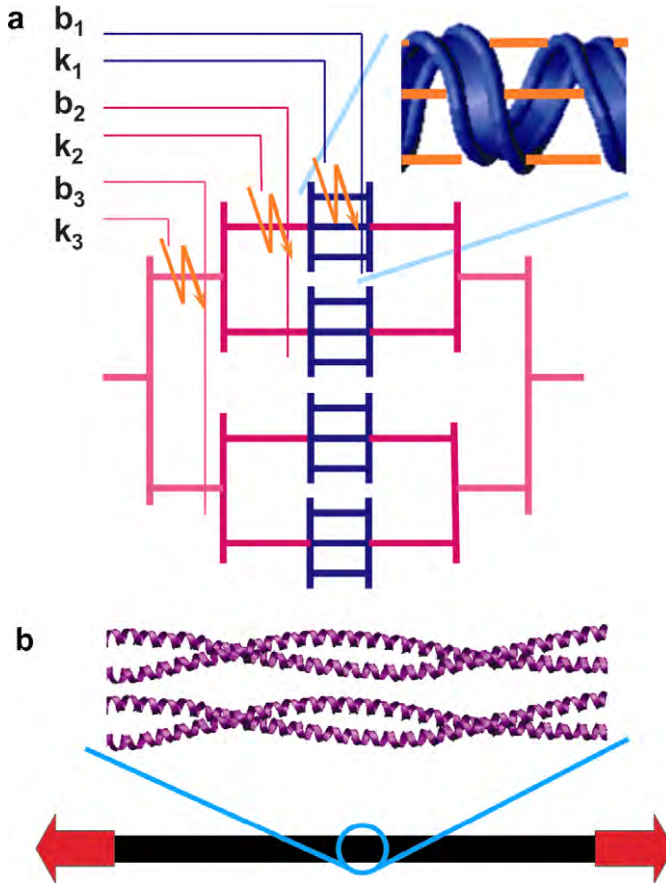
Fig. 56. Illustration of different hierarchies and their representation in the Hierarchical Bell Theory (subplot (a)), as well as representation of the corresponding physical system (subplot (b)). The inlay in the upper part of subplot (a) shows a single AH structure with ≈3 HBs per convolution. The Hierarchical Bell Theory reviewed here enables one to predict the strength of such hierarchical bond arrangements as a function of the deformation speed. Subplot (b) shows the physical system that is represented in the hierarchical model in subplot (a).

16 residues long helix in [220] started to unfold at both ends simultaneously, in contrast to the 20-residue long amino acid chain modeled in [221], which unfolded systematically from the side where force was applied. Furthermore, a multi-scale model for human hair that mainly consists of keratin IF CCs was developed, as reported in [217].

However, aside from these studies, to the best of our knowledge few MD results related to the nanomechanical behavior specifically of IFs have been published thus far. No systematic studies on the difference of mechanical properties between the behavior of single AHs and CCs were carried out. Virtually no structure–property relationships exist for these structures, and therefore theories are missing that link the molecular architecture to their mechanical properties. The nanomechanical properties, in particular their unfolding and large deformation behavior and a systematic analysis of its time-dependent fracture properties remain unknown. Here we review a series of additional classical molecular dynamics (MD) simulations reported from our group that were aimed at addressing these open questions. The goal of these investigations was a systematic analysis of the unfolding behavior of the protein domains at varying pulling rates. Fig. 57 depicts the force versus strain response of a single AH and the coiled-coil structure, both carried out at identical loading rates of
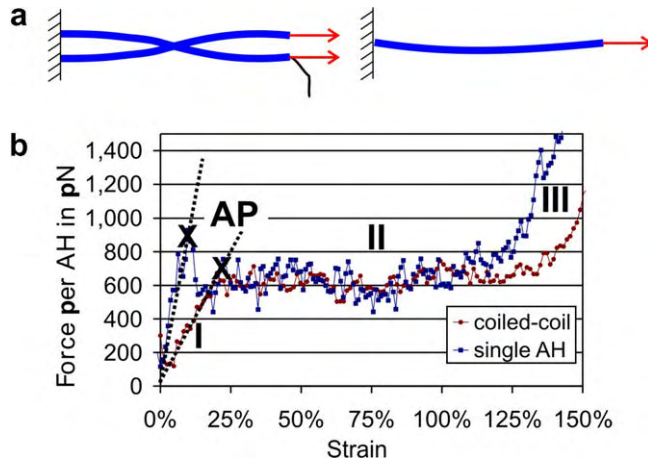
Fig. 57. Tensile experiments of vimentin proteins, comparing a AH structure with a CC structure. Subplot (a) depicts a schematic of the applied load (left: coiled-coil, right: single AH). Subplot (b): The force–strain curves of a single AH-structure and an alpha-helical coiled-coil, both at a pulling rate of 5 m/s. To enable better comparison of both curves, the force of the coiled-coil is divided by the number of helices. The first regime (I) consists of a steep increase in force until a strain of approximately 25% (referred to as angular point (AP)) for the coiled-coil and 13% for the single AH. The first regime is followed by the second regime (II) during which unfolding of the alpha-helices occurs. The forces at the AP are much higher for a single AH than for the coiled-coil. In the third regime (III), a non-linear increase in strain by stretching the backbone is observed. Thereby the single AH has a much steeper and earlier increase in force than the coiled-coil structure. The differences between the single AH and the coiled-coil structure are summarized in Table 8.

5 m/s (geometries shown in Fig. 57a). It is noted that the forces plotted in Fig. 57b are normalized by the number of AHs in the structure (that is, one for the single AH, and two for the coiled-coil).

For both structures, three distinct deformation regimes (I–III) are observed, characterizing the stretching dynamics of the protein. These three regimes are now characterized first for the coiled-coil structure before the differences between the single AH and the coiled-coil are described. In the first regime (I), the helical structure is stretched homogeneously. At the angular point (indicated as "AP" in Fig. 57, and marked with "**x**" in the plot), the structure begins to unfold by rupture of H-bonds, characterized by a significant change in slope in the force–strain plot, leading to regime (II). The AP thus represents the critical force necessary to induce a structural change in the protein, corresponding to the unfolding force. It is observed that unfolding is initiated at the end where the load is applied. In the case of the coiled-coil, each of the two alpha-helices unfolds individually, while the secondary superhelical arrangement remains intact, until strains reach more than 100%. Unfolding of each AH structure is characterized by sequential breaking of H-bonds. During this regime, the force remains approximately constant while the entire protein is unfolded at strains approaching 150%. Once the complete helix is unfolded, the slope increases continuously while the secondary superhelical configuration is lost at strains larger than approximately 150%, eventually leading to stretching of covalent bonds in the protein backbone, giving rise to rapidly increasing forces at large deformation in regime (III). This unfolding sequence (first unfolding of individual AHs, then unfolding of the superhelical structure) suggests that in the coiled-coil case, the individual AHs represent the weakest link in the structure.

Even though the unfolding curves of single AHs and coiled-coil structures are qualitatively similar, there are several significant quantitative differences. First, the slope of the single AH in the first regime is almost two times steeper and leads to much higher forces at the angular point. The angular point appears at strains of 13% in the single AH structure, compared with 20% for the coiled-coil. Second, unfolding of the single AH begins at the residue where the pulling force is applied, but is followed by immediate additional rupture of H-bonds at residue 369 and 383 inside the protein. Third, strain hardening of the single AH sets in at 20% lower strains; 125% in the single AH case versus 145% in the coiled-coil. The strain hardening increases much steeper in the case of a single AH. The reason

Table 8
Comparison of the mechanics of single AH with the coiled-coil in different dimensions

Dimension	Single AH	Coiled-coil
Equilibrated H-bond length in Å	3.08 ± 0.29	2.95 ± 0.16
Slope of the first refime in pN/Å	186.1	94.9
Force per AH at angular point in pN	930	670
Strain at angular point	13%	20%
Breaking of H-bonds	Simultaneous	Sequential
Beginning strain hardening at	125%	145
Strain hardening	Very steep	Continuous

The data have been derived from simulations with a pulling rate of 5 m/s. The results clearly indicate that the single AH has much more irregularities and instabilities compared with the coiled-coil structure, suggesting a higher mechanical stability.

for this difference is likely due to the uncoiling process of the superhelical structure that is missing in the pure AH structure. The differences between the two structures are summarized in Table 8.

To investigate the pulling rate dependence of vimentin mechanics, we have carried out pulling experiments with systematically varying pulling rates. Fig. 58 depicts force–strain curves for a wide range of different pulling velocities, clearly illustrating a strong pulling velocity (or, equivalently strain rate) dependence of the mechanics of CCs. The results reveal larger unfolding forces for larger pulling rates. Increases in the pulling velocity lead to increased slopes in particular in regimes (I) and (II).

The availability of atomistic simulation results enables the development of a mesoscopic bead-type model of a AH structure, as originally reported in [72]. This goal can be achieved with the same approach as the model development reviewed for collagenous tissues (see Section 3), and also enables one to model the deformation behavior of very long protein molecules, reaching several hundred nanometers. Fig. 59a shows the schematic of replacing the full atomistic representation with a bead model. Fig. 59b shows the force–extension curve, obtained using the bead model, for a molecule with length 249.5 nm, at 300 K. The entire force–displacement curve features different regimes of deformation, including regime I, entropic elasticity, regime II, energetic elasticity, regime III uncoiling of the AH structure, regime IV stretching of the backbone bonds, and regime V, molecular fracture. Whereas the WLC model predicts divergence of forces at the contour length, molecular modeling predicts a smooth transition from entropic to energetic elasticity. As in the case of tropocollagen molecules, entropic elasticity becomes more significant, and the force–extension behavior can be modeled using the WLC model. Fig. 60a shows a close-up view into the small force regime for which the WLC model is
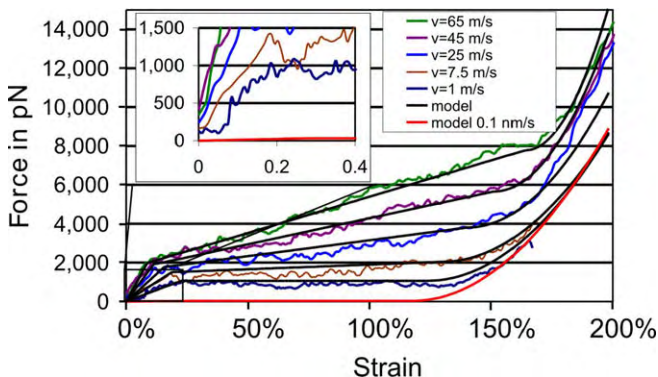
Fig. 58. Force–strain curves of a coiled-coil alpha-helical structure at different pulling rates. We describe the behavior with the developed continuum model and extrapolate to pulling rates, used AFM experiments. For pulling velocities of smaller than 5 m/s the strains, where the regime changes take place are very similar to those found in experiments and are 20% for the angular point and 120% for the beginning of the strain hardening. The straight lines refer to the prediction of our continuum model, for various pulling rates.
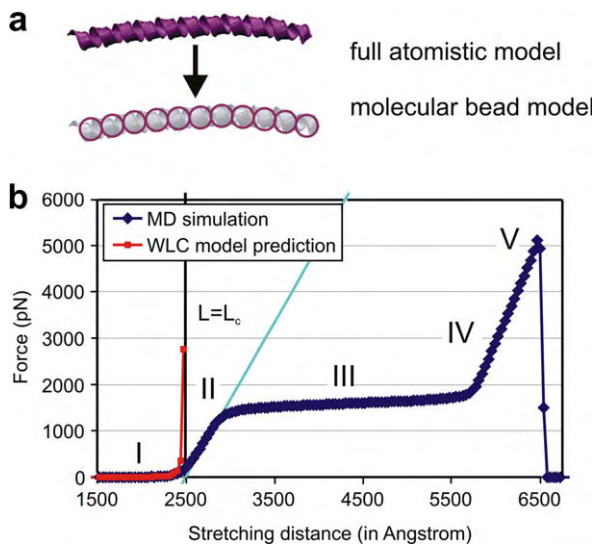
Fig. 59. Development of a mesoscopic bead-type model of a AH structure, as reported in [72]. Subplot (a) shows the schematic of replacing the full atomistic representation with a bead model. Subplot (b) shows the force–extension curve, obtained using the bead model, for a molecule with length 249.5 nm, at 300 K. The entire force–displacement curve features different regimes of deformation, including regime I, entropic elasticity, regime II, energetic elasticity, regime III uncoiling of the AH structure, IV stretching of the backbone bonds, and regime V, molecular fracture. Whereas the WLC model predicts divergence of forces at the contour length, molecular modeling predicts a smooth transition from entropic to energetic elasticity.
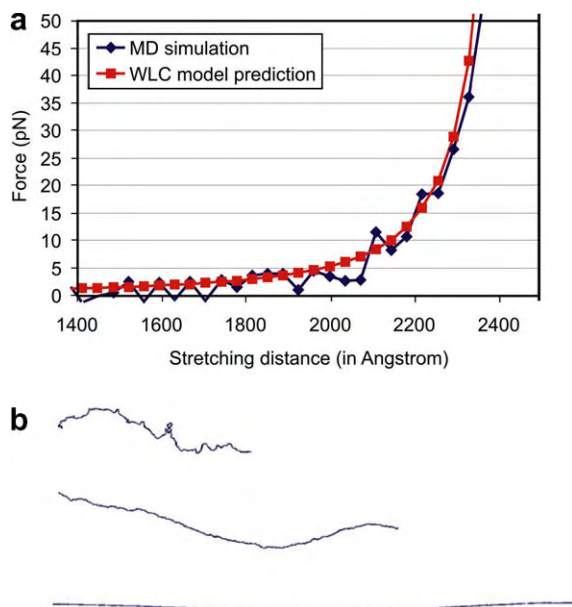
Fig. 60. Subplot (a) shows a zoom into entropic elasticity before the molecule is completely stretched to its contour length, including a quantitative comparison with the WLC model [72]. The contour length is indicated in the plot as a straight line. Subplot (b) shows snapshots of the molecular geometry as the applied force is increased.
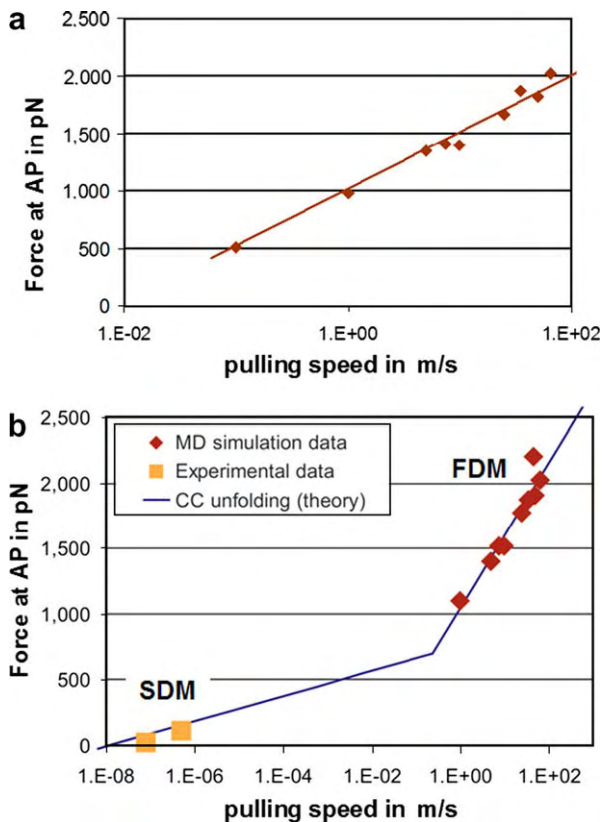
Fig. 61. Subplot (a): Unfolding force as a function of the pulling speed, for the vimentin CC structure. Subplot (b): Theoretical predictions for the unfolding force at the angular point (AP) compared with simulation and experiment, considering the transition from FDM to SDM. Experimental results [182,216] agree well with the theoretical prediction.

a good approximation. Fig. 60b depicts snapshots of the molecule as it is being stretched, losing the initial random configuration.

Fig. 61a plots the unfolding force at the angular point over the pulling velocity. By directly fitting the results for a_1 and a_2 from atomistic simulation to Eq. (48), we obtain $E_b = 5.6$ kcal/mol and $x_b = 0.185$ Å, with $\theta = 23°$ (angle determined from geometric analysis of the protein structure). These values indicate that at the angular point, approximately one individual H-bond is broken, as it is known that E_b for H-bonds is between 4 and 5 kcal/mol.

4.5.1. Hierarchical multi-scale approach: development of analytical model

Based on the simulation data for varying pulling rates and the analysis based on Bell's Theory, it is possible to generate an analytical model, consisting of linear functions for the first and the second regime and a harmonic function for the third regime. Such a fully atomistically informed, closed form model could be useful in applications of mesoscale modeling of IF structures (mesoscale modeling of such structures is not explored further here). During the first two regimes, the slope increases linearly with increasing strain rates, while in the third regime the curvature is proportional to the square of the pulling rate. As mentioned above, we find that the force at the angular points depends logarithmically on the pulling rate, following the predictions by the theoretical model.

A set of empirical equations is used to describe the rate dependent force–extension curves based on a continuum model. All parameters in the model are determined rigorously from atomistic simulation results. The form of the equation (e.g. the dependence of AP force on pulling speed according to

Eq. (48)) is based on the Extended Bell Theory described above. The forces in the three regimes are given by

$$f(x, v) = \begin{cases} F_1 = (F_a \cdot x)/x_a & \text{if } x < x_a \\ F_2 = s_2 \cdot (x - x_a) + F_a & \text{if } x_a \leqslant x < x_s, \\ F_3 = s_3 \cdot (x - x_s)^2 + F_s & \text{if } x \geqslant x_s \end{cases} \tag{55}$$

where $F_a = \max(c_1 \cdot \ln(v) + c_2; c_{12} \cdot \ln(v) + c_{13})$, $x_a = c_5 \cdot v + c_6$ and $s_2 = \max(c_7 \cdot v + c_8; 0)$ are additional parameters, and v is the pulling speed (in units of Å/fs). Further, $F_s = c_3 \cdot x_s + c_4$, $s_3 = c_9 \cdot v^2 + c_{10} \cdot v + c_{11}$ and $x_s = \left(F_a \cdot \left(1 - \frac{s_2}{s_1}\right) - c_4\right) / (c_3 - s_2)$.

It is noted that F_a and x_a are the forces and the displacements at the angular point, F_s and x_s are the force and the displacement at the beginning of the third regime. The parameters c_i (extracted purely from fitting to atomistic simulation results) are summarized in Table 9. The results of this continuum model and comparison with MD simulation results are shown in Fig. 58 (the continuous lines are the predictions by the continuum model).

4.5.2. Modeling cyclic loading of IF CC structures

The analytical model given in Eq. (54) enables us to describe the behavior of this structure even if the strain rate is varying during the stress application or cyclic loading is applied, for instance. Fig. 62a shows a prediction of the force–strain curve of an experiment of cycling loading of a coiled-coil vimentin structure.

In modeling cyclic loading, it is assumed that the relaxation curve for the coiled-coil structure equals to the one at pulling rates in the order of nm/s (rate at which re-folding into AH and coiled-coil structures appears, chosen to be much smaller than the pulling rate, in accordance with experimental observations that show refolding processes at time-scales of seconds). Similar results were shown in experiments of cyclic loading for the coiled-coil structure of myosin, even though different levels of assemblies were considered [216]. Our resulting force–extension curves (see Fig. 62) show qualitative agreement with these experimental studies.

Fig. 62b depicts the dissipation energy per loading cycle as a function of the pulling rate. This plot clearly indicates, as predicted earlier [175,232] that vimentin intermediate filaments are a stress buffering protein for higher deformation rates.

4.6. Atomistic modeling results of deformation and fracture mechanisms of individual AH domains

This section is focused on a more detailed analysis of the deformation and rupture mechanism of the vimentin IF dimer structure, here focused on individual AH domains. The chief goal of these studies is to reduce the deformation rate as much as possible to obtain insight into potential mechanism changes as extremely slow pulling rates. This is motivated by the fact that most MD simulations could not be linked quantitatively with experimental observations. For instance, considering the results shown in Fig. 61a reveals that an extrapolation of the slope to experimental pulling rates below 10E−8 m/s would result in negative force values, suggesting that the molecule would not be stable under equilibrium conditions without forces applied. Clearly, this is in contrast to experimental observations and would not enable IFs to satisfy their physiological role.

Here we review studies of two AH model protein domains (molecular geometries see Fig. 63); the choice of two model protein domains is dictated by the desire to show a universally valid strain rate

Table 9
Parameters derived from the curve fitting to match atomistic results; F_i are the forces for the different regimes, s_i are the slopes as function of the pulling rate v (in Å/fs)

c_1 (pN)	c_2 (pN)	c_3 (pN/Å)	c_4 (pN)	c_5 (pN fs/Å²)	c_6 (pN/Å)	c_7 (pN fs/Å²)	c_8 (pN/Å)	c_9 (pN fs²/Å³)	c_{10} (pN fs/Å²)	c_{11} (pN/Å)
220	3604	237	−19.65	723,894	59.6	173,519	−6.3	2.4E7	−4.55	2.7
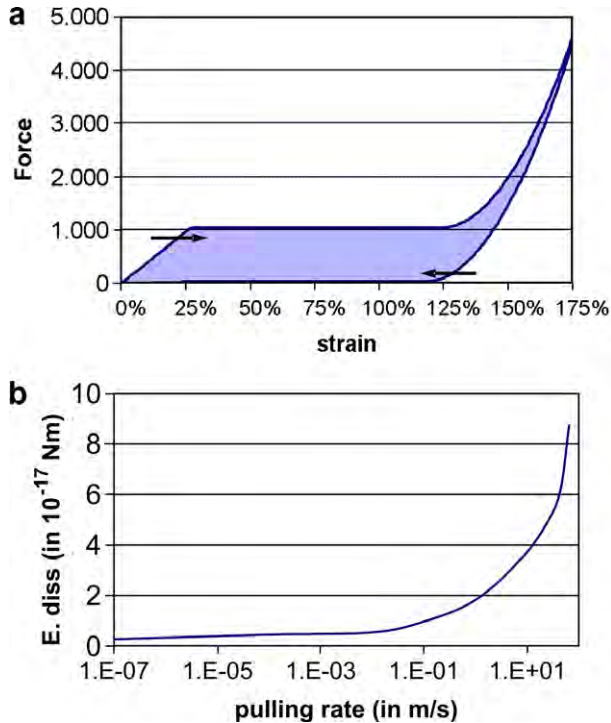
Fig. 62. Subplot (a): Using our analytical continuum model, we predict the following behavior for a cyclic load experiment. In this example, we are pulling with a speed of 1 m/s. For relaxation, we assume the same behavior as for pulling with a speed of 10 nm/s. The blue area equals to the dissipated energy. Subplot (b): Our continuum model enables us to model the dissipation energy for the coiled-coil structure as a function of the pulling rate. The dissipation energy grows rather slowly until pulling rates of 0.1 m/s, and increases almost linearly with increasing pulling velocities afterwards (here exponential, due to the log-plot). This indicates that vimentin has strain buffering properties in cells, as it features small energy dissipation rates for small pulling rates (e.g. during cell motility), but extremely large dissipation for large pulling rates. This figure is adapted from [138].
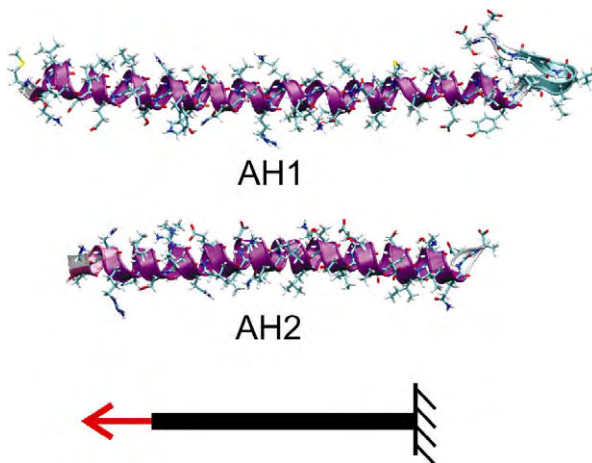
Fig. 63. Atomistic geometries of the three protein domains studied here (AH1, AH2). Surrounding water molecules are not shown for clarity. The lower part of the plot indicates the boundary conditions (tensile loading). Figure adapted from [40].
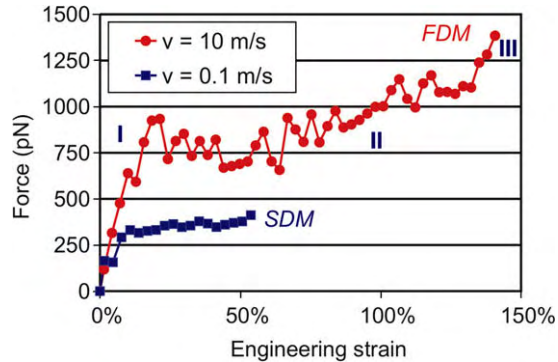
Fig. 64. Representative examples of force–extension curves, as reported in [40]. The fast deformation mode (FDM) is represented by a curve taken at a pulling speed of 10 m/s. The slow deformation mode (SDM) is represented by a pulling experiment at 0.1 m/s. The force–extension behavior consists of three regimes. (I) Linear increase in strain until the angular point is reached when the first HBs rupture, leading to unfolding of one helical convolution. (II) Plateau of approximately constant force, during which unfolding of the entire protein occurs. (III) Strain hardening (only partly shown for the FDM).

dependence effect AH1 is a domain from the 2B segment of the vimentin intermediate filament dimer (please see Fig. 52) [169–171], and AH2 is a domain from bacteriophage T4 fibritin [233].

For the vimentin AH protein domain (AH1), two characteristic force–strain curves are shown in Fig. 64 for two pulling speeds. The simulations reveal existence of three distinct deformation regimes. The first regime shows a linear increase in strain until the angular point is reached. The second regime is a plateau of approximately constant force, during which unfolding of the entire protein occurs. The last regime displays a significant strain hardening, due to pulling of the protein's backbone (only partly visible in the FDM plot). A similar behavior is observed for the AH2 structure. The change from the first to the second regime is referred as the angular point (AP), denoting the protein unfolding force (this is a qualitatively identical behavior as shown in Figs. 57 and 58). Unfolding of the protein is characterized by rupture of HBs that destroy the original protein structure (that is, the folded structure that is stable without forces applied) as the displacement is increased. In the remainder of this section, we focus on the force at the AP as a function of the pulling speed.

Computational experiments have been carried out by systematically varying the pulling velocity over four orders of magnitude, ranging from 0.05 to 100 m/s. The unfolding force is plotted as a function of the pulling speed in Fig. 65 for AH1 and AH2. Notably, in both cases (AH1 and AH2) we observe
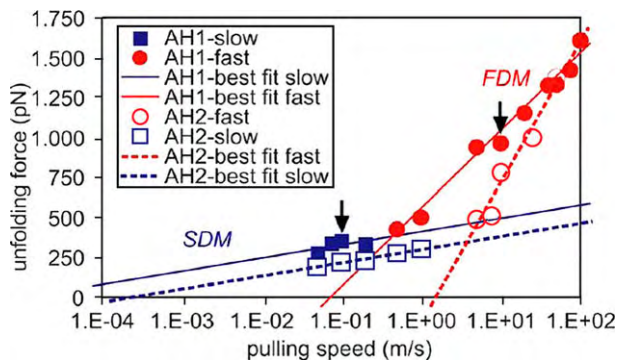
Fig. 65. Unfolding force of single AHs from the vimentin coiled-coil dimers, as a function of varying pulling speed over four orders of magnitude, ranging from 0.05 to 100 m/s (original study reported in [40]). The results clearly reveal a change in protein unfolding mechanism from the FDM to the SDM. The arrows indicate the representative pulling speeds used for the analysis reported in Fig. 64.

two distinct regimes, each of which follows a logarithmic dependence of the unfolding force with respect to the pulling rate. The existence of two discrete slopes indicates two different energy barriers and thus two different unfolding mechanisms over the simulated pulling velocity regime. The results clearly suggest a free energy landscape that consists of two transition states. In the following we refer to these two regimes as the slow deformation mode (SDM) and the fast deformation mode (FDM). The change in mechanism from the FDM the SDM occurs at $v = 0.4$ m/s (AH1) and $v = 4$ m/s (AH2), and at a force of ≈ 350 pN (AH1) and ≈ 400 pN (AH2).

In other studies, this change in mechanism has only been suggested or inferred [234,235]. For example, a comparison between MD simulation and experimental results revealed that force-pulling speed dependence must lay on two different curves in the f–ln(v) plane [234,235], suggesting a change in unfolding mechanism. Fig. 66 shows the difference in behavior, as observed from MD simulation and experimental studies. The plot shows experimentally measured [236,237] and atomistically calculated [238] peak unfolding forces of titin I27 domain as a function of the loading rate. The plot reveals a transition of the curves, reflecting a change in mechanism; however, MD simulation and experiment have not yet been linked for this protein structure. This behavior may be similar as the one shown in Fig. 65.

4.6.1. Analysis of simulation results with Bell Theory

By fitting the modified Bell Theory to the MD results of the AH1 structure, we obtain for the FDM $E_b = 4.7$ kcal/mol and $x_b = 0.20$ Å. In the SDM (with $\theta \approx 16°$), $E_b = 11.1$ kcal/mol and $x_b = 1.2$ Å. Similar values have been found for the AH2 structure, albeit the results for E_b are slightly lower (see Table 10). Considering that the bond breaking energy E_b of a HB in water ranges from 3 to 6 kcal/mol [239], the results indicate that in the FDM, individual HB rupture sequentially. In contrast, in the
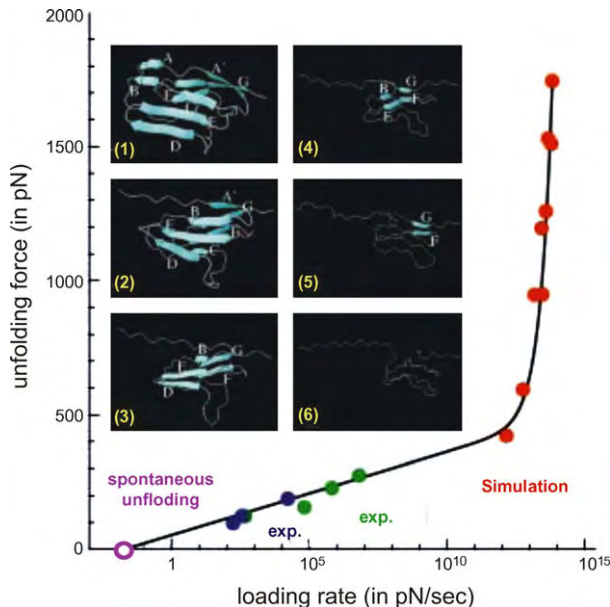
Fig. 66. Rate dependent mechanical behavior, as observed from MD simulation and experimental studies. The plot shows experimentally measured (blue [236], and green [237]) and atomistically calculated (SMD, red [238]) peak unfolding forces of titin I27 domain as a function of the loading rate. The plot reveals a transition of the curves, reflecting a change in mechanism, however, MD simulation and experiment have not yet been linked for this protein structure. This behavior might be similar as the one shown in Fig. 46. Inset figures show the deformation trajectory of titin I27 in MD simulations. Reprinted with permission from Ref. [238]. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

Table 10
Summary of the differences between the SDM and FDM, for AH1 and AH2 [40]

	AH1 (AH2) domain	
Deformation mode	SDM	FDM
Pulling speeds (m/s)	$v < 0.4$ (4)	$v > 0.4$ (4)
Unfolding forces (pN)	$F < 350$ (400)	$F > 350$ (400)
E_b (kcal/mol)	11.1 (9.11)	4.87 (3.08)
x_b (Å)	1.2 (1.19)	0.2 (0.11)
HB breaking mechanism	Simultaneous	Sequential

The values in parentheses in the AH columns represent the results for AH2.

SDM approximately three HBs rupture at once. The studies of both AH structures clearly support this observation. The details of the atomistic rupture mechanisms are summarized in Table 10.

An analysis of the atomistic structure during the rupture event is shown in Fig. 67a for the SDM in the vimentin AH1 domain. In the SDM, three HBs rupture simultaneously, within less than 20 ps timescale. It was reported that the time for HB breaking is approximately 20–40 ps [239], clearly supporting the notion that these HBs rupture at once.
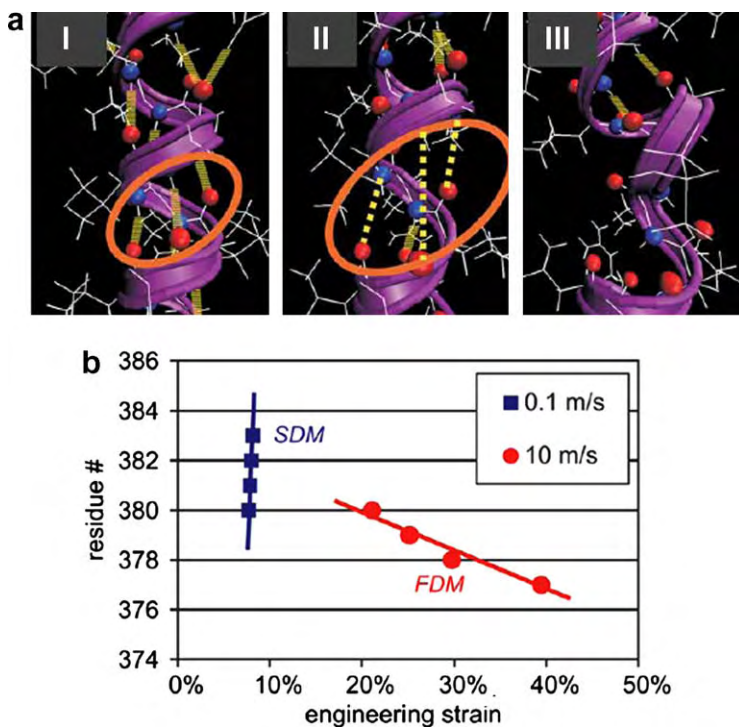
Fig. 67. Atomistic details of the unfolding mechanism of AH1 in the SDM (angular point in Fig. 1 for $v = 0.1$ m/s). Subplot (a): Atomistic representation of the rupture dynamics. The time interval between these snapshots is 20 ps (between I and II) and 40 ps (between II and III). After 20 ps (I–II), all three HBs have ruptured simultaneously, leading to local unfolding of the protein in the following 40 ps (II–III). These snapshots strongly support the concept of cooperative bond rupture in the SDM. Surrounding water molecules are not shown for clarity. Subplot (b): Rupture sequence of the first four HBs as a function of the applied strain (residue number represents the amino acid of the O-atom (H acceptor)). In the FDM, HBs rupture one by one, whereas in the SDM several HB rupture virtually simultaneously, within 20 ps. In the FDM the unfolding wave runs from the pulled residue in the direction of the fixed residue, whereas in the SDM the unfolding 'wave' runs in the opposite direction, nucleating at a random residue within the protein sequence.

Further evidence for the change in mechanism is obtained by an analysis of the HB rupture dynamics. In Fig. 67b, we plot the HB rupture as a function of the molecular strain for the vimentin AH1 domain. This provides a strategy to normalize the different time-scales by the pulling velocity (here 0.1 and 10 m/s). In agreement with the results shown in Fig. 64, the unfolding of the protein in the SDM starts at approximately 10% strain, in contrast to 20% strain in the FDM regime. This is indicated in Fig. 67b by the rupture of the first HB. The data shown in Fig. 67b clearly shows that in the FDM, HBs rupture sequentially as the lateral load is increased from 20% to 40% tensile strain. In contrast, in the SDM several HBs rupture virtually simultaneously, within less than 20 ps, at a tensile strain of ≈10%. Even though the pulling speed is several orders of magnitude slower in the SDM, the HBs in the SDM rupture significantly faster.

4.6.2. Correlation with experimental results

Notably, force spectroscopy results of individual AH domains are rare, even though the first studies with AFM have been reported almost 10 years ago. This is partly due to experimental challenges to manipulate the small physical size of the protein probes, typically on the order of tens of nanometers. For instance, it is difficult in experiment to stretch individual proteins or protein domains rather than bundles [222,240,241]. Samples of this small size cannot be visually displayed, and methods such as AFM or other imaging techniques must be employed.

The unfolding forces measured in experiments are typically between 100 and 200 pN for pulling velocities in the SDM [222,240,241]. This is close to the values obtained in our simulations, which predict forces below 350 pN in the SDM. Pulling experiments of coiled-coils, which consist of two AHs arranged in a helical geometry, can be carried in a more controlled fashion. In these systems, the unfolding force measured in experiment ranges between 25 and 110 pN, also for deformation speeds in the SDM [182,216]. These force values are also in proximity to the unfolding forces predicted by the simulations and the theoretical model.

Fig. 61b depicts a comparison of the theoretical predictions for the unfolding force at the angular point (AP) compared with simulation and experiment, considering the transition from FDM to SDM. The experimental results [182,216] agree well with the theoretical prediction. This clearly illustrates that the FDM-SDM transition is vital to understand and interpret experimental results in the low deformation rate regime.

4.7. Structure–property analysis of the AH structure: robustness against mechanical deformation

Protein folding and thus the generation of hierarchical structures are essential for biological function. First, it allows for distant parts of the amino acid chain to come physically closer together, creating local sites with specific chemical properties that derive from the collection of particular residues. Second, folding permits collective, localized motion of different regions [242]. The AH pattern is the most simple folding motif out of a one dimensional strand [162], forming a spring-like protein structure with high elasticity and large deformation capacity.

But why does an AH fold in such a way that 3.6 parallel HBs, instead of 2, 5 or 6 appear in parallel, per convolution? Notably, all AHs universally show this particular molecular architecture. To the best of our knowledge, there has been no explanation for this particular molecular feature, despite the fact that the AH is such an abundant protein structure. Maybe the structural features can be explained by considering the robustness of the AH structure against mechanical and thermodynamical unfolding.

4.7.1. Robustness of AH protein domains

The robustness can be calculated based on the definition of robustness as parameter insensitivity, postulated by Kitano [32,33]. This definition applied to the case of a AH structure corresponds to the sensitivity of the protein strength in regards to missing HBs. Starting with the Hierarchical Bell Model (Eq. (52)), the robustness is calculated as the ratio of strength of a defected system and an intact system. The intact system is defined as a system where all HBs contribute to strength, whereas in the defected system all except one HB contribute to the strength. Only the contributions due to the hierarchy in Eq. (52) are considered; the pulling speed part of this equation is not taken into account since we compare systems at identical pulling speeds. The consideration of pulling speeds would

lead to the addition of a constant value to the forces. Since this is identical for different hierarchical designs, it does not influence the comparison of the behavior of these structures. The robustness is defined as

$$r(b) = \frac{f_{h1}(d=b-1) + f_{h0}(d=b-1)}{f_{h1}(d=b) + f_{h0}(d=b)} = 1 + \frac{k_B \cdot T \cdot \ln(b) - E_b^0}{E_b^0 \cdot b}. \tag{56}$$

The robustness converges towards fault tolerance when $b \to \infty$.

Fig. 68 depicts the robustness of an AH as a function of parallel HBs per convolution. The analysis shows that for a AH structure with 3–4 HBs per convolution, 80% robustness is achieved (0% robustness means that the system is highly fragile, and 100% represent complete fault tolerance). This level of robustness in a biological structure enables to minimize waste of resources (that is, amino acids), the weight and volume, and thus making the structure overall efficient and able to sustain extreme mechanical conditions such as high loading rates and deformations. This finding is important since the only input parameters in this model is the dissociation energy of a HB, E_b^0, as well as the dissociation distance, x_b, which are fundamental, 'first-principles' properties of protein structures. These parameters can be determined reliably from either experiment or atomistic simulation (both approaches lead to similar values). The remainder of the parameters required to predict the robustness properties can be derived from the geometry of the protein structure.

Synthetic materials often do not have such high levels of intrinsic robustness. This makes it necessary to introduce safety factors that guarantee a structure's functionality even under extreme conditions. For instance, an engineering structure like a tall building must be able to withstand loads that are ten times higher than the usual load, even if this load will never appear globally. This safety factor is necessary since these structures are very fragile due to their extremely high sensitivity to material instabilities such as cracks, which might lead to such high local stresses. However, if a crack does not appear during the operation, a large fraction of the material is wasted. This illustrates the potential of engineering bio-inspired robust and efficient structures. The key may be to include multiple hierarchies and an optimal degree of redundancies, as illustrated here for the AH structure.

Forming 3–4 HBs in parallel instead of forming a single, much stronger bond is also energetically favorable, in particular in light of the moderate assembly temperatures *in vivo*. However, this only makes sense if three HBs rupture simultaneously so that they can provide considerable mechanical and thermodynamical resistance. This has indeed been shown to be the case at physiological strain rates [240]. The intimate connection of structural properties, assembly and functional processes is an overarching trait of protein materials.
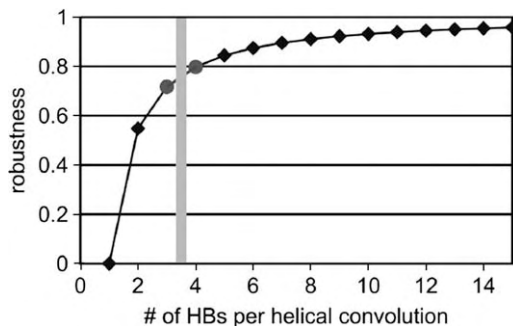
Fig. 68. Robustness of a AH as a function of parallel HBs per convolution, *b*, predicted by the Hierarchical Bell Model [40]. Robustness is defined as the ratio of strength of a failed system and an intact system. In the intact system, all HBs contribute to strength, whereas in the failed system all except one HB contribute to the strength. The dotted line indicates the number of parallel HBs per convolution (3.6 HBs) as observed in Nature. This particular molecular geometry corresponds to a robustness value of ≈80%, indicating that the AH is efficient in Pareto's sense [243,244].

4.7.2. Discussion in light of Pareto efficiency

It has been found that 3–4 parallel HBs are the most favorable bond arrangement in light of the mechanical and thermodynamical stability, leading to a robustness of approximately 80%. This result provides an indication that AHs are efficient following Pareto's principle [243,244], also known as the 80/20 rule. This rule is an empirical law that has found broad application in explaining social, economic, political and natural phenomena. Our results indicate that this concept may be also applicable to explain the nanoscopic architecture of the AH protein motif.

Pareto's efficiency rule found a broad acceptance and application in many social, economic, political, and natural phenomena, albeit clearly the Pareto rule is a rule of thumb and does not provide a quantitative measure. For example 80% of the wealth are concentrated on 20% of the entire population [245]. Similarly, 80% of the revenues, software usage or published papers typically stem from 20% of all customers, programmed code or authors [245].

The 80/20 rule is an empirical law that is found in many natural phenomena. The theoretical foundation of this principle was reported by Chen et al. [244,246]. They state that this particular distribution is the result of the probability of a new entry (that is, the addition of an additional strengthening element), which quantifies the height of the entry barrier. Most importantly, it was shown that the probability of a new entry has an inverse relation with the level of usage concentration. This leads to a concurrence between existing entities and the addition of new entities, leading to the characteristic 80/20 distribution. The more robust the structure becomes with each additional HB (see Fig. 68), the higher is the barrier to implement an additional HB, since each HB introduces an additional 'cost' due to increased material use, that is, the additionally generated weight and additionally required volume. However, the increase of robustness decreases rapidly, so that an optimal balance is found at the 80% mark that is characteristic for Pareto efficiency.

This is the first time that substantial evidence has been found for the applicability of the Pareto principle to explain the molecular structure of proteins. In Pareto's sense, the more robust the structure becomes with each additional HB, the higher the barrier to implement an additional HB, since each HB introduces an additional 'cost' due to increased material use, that is, the additional weight and volume. In light of these considerations, it is not surprising that a robustness value of 80% is found in AHs, which equals to the optimal state due to these competing mechanisms. It is noted that other reasons, such as steric effects, may also be important in explaining the particular geometry of the AH protein motif.

4.8. Stutter defects in coiled-coils

All intermediate filament dimers feature a highly conserved "stutter" region, a sequence of amino acids that interrupts the superhelical coiled-coil arrangement of the two alpha-helices, leading to a parallel arrangement of the alpha-helices in this region. Earlier studies have suggested that the stutter plays an important role in filament assembly. By utilizing predictions made by the Hierarchical Bell Model, we show that the stutter also has a significant effect on the mechanical behavior of vimentin dimers.

Here we review studies that suggest that the presence of the stutter leads to a softer structure with more homogeneous plastic strain distribution under deformation, as reported in [247]. The predictions by the Extended Bell Model are confirmed by large-scale MD simulations of three model systems: two parallel alpha-helices, a coiled-coil dimer, as well as a coiled-coil dimer with a stutter. The simulations suggest that in agreement with the prediction based on our Extended Bell Model, the stutter represents the locations at which the protein structure has the least resistance to unfolding. We discuss the implications of this molecular architecture in terms of its biological function.

4.8.1. Research design

This study involves a systematic analysis of three molecular geometries: model A is a "perfect" CC without a stutter (PDB ID 1gk6, residues 355–406). Model B is a system that contains two single AHs (extracted from model A, entire structure embedded in explicit solvent), arranged in parallel with a distance of 4 nm, mimicking the molecular geometry in the stutter region. Model C is a CC structure with a stutter (PDB ID 1gk4, residues 328–379) and thus a 'combination' of model A and B. All struc-
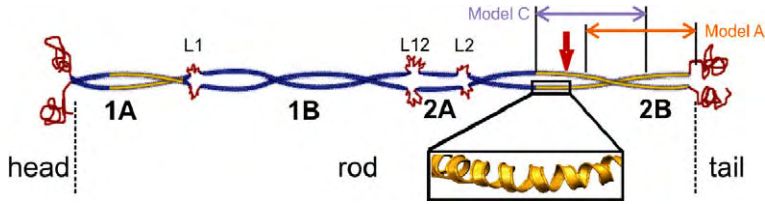
Fig. 69. Geometry of the dimeric building block in vimentin. The dimers, approximately 45 nm long, are the elementary building blocks of vimentin intermediate filaments. A dimer consists of a head, tail (plotted in red) and an elongated rod domain which is divided into four alpha-helical coiled-coils (1A, 1B, 2A, 2B) connected through linkers L1, L12, L2 (also red) [166]. The segments plotted in yellow (part of segment 1A and 2B) have been crystallized so that the atomic structure is known. The stutter is located approximately in the middle of the 2B segment (indicated with a red arrow). The molecular dynamics simulations described in this paper are performed on alpha-helices placed in the 2B segment (yellow). The locations of the model A and model C are shown. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

tures have the same length of 70 Å, to exclude any effects of the molecular length on the results. Fig. 69 depicts the location of models A and B in the vimentin dimer. The strategy of investigation is as follows. We begin with a review of the analysis of models A and B. For both systems, we then present an analysis using the Extended Bell Model and confirm the theoretical predictions by MD simulation of both systems. Then, we consider model C and carry out MD simulations to analyze the unfolding dynamics and unfolding resistance.

4.8.2. Theoretical analysis

Geometrically, the difference of models A and B/C is the angle θ. In model A, the angle $\theta_A = 23° \pm 10.2°$, whereas in models B and C, the angle $\theta_B = 16° \pm 8.5°$ and $\theta_C = 16° \pm 6.9°$. The angle fluctuates randomly over the different H-bonds. Note that in model B, the angle $\theta_B \neq 0°$ since in an AH each H-bond connects residue i with residue $i + 4$, whereas a convolution consists of only ≈ 3.5 residues. Consequently, the H-bonds of a straight AH are slightly tilted in the direction of the twisting backbone. Thus, the difference in the angle in model A and model B are due to the molecular twist in CC systems that appear in addition to the tilt of the H-bonds.

Provided that x_b and E_b are equal – that is, assuming the same unfolding mechanism – the Extended Bell Model predicts that the structure with the larger angle is stronger (for $\theta_2 > \theta_1$, $\cos\theta_2 < \cos\theta_1$ or $\cos\theta_1 / \cos\theta_2 > 1$):

$$f_2(v) = \frac{\cos\theta_1}{\cos\theta_2} \cdot f_1(v). \tag{57}$$

Therefore, in a system that consists of a hybrid combination of models A and B, the location that corresponds to model B represents the weakest segment, which will unfold most easily under increasing applied tensile load. This situation is resembled in vimentin IF, as represented in model C (see Fig. 69). In this sense, introducing the stutter corresponds to the deliberate addition of a weakening 'defect' into the molecular structure. This suggests that the role of the stutter may be to provide a predefined and controlled location where unfolding occurs under large deformation. For an angle $\theta_A \approx 23°$ of the CC (model A) and the angle $\theta_B \approx 16°$, the difference in unfolding force is approximately 5%.

As pointed out in the review section, vimentin IFs have their most significant role in the large deformation behavior. Predicting the force at the angular point and the level of force in regime II is crucial, since these regimes dominate the large deformation behavior under physiological conditions.

4.8.3. MD simulation results

In order to gain more insight into the molecular unfolding mechanism as well as the rate dependence of the results, simulations over a wide range of pulling velocities have been performed. The unfolding force is shown in Fig. 70 as a function of pulling velocity. In our analysis, the unfolding force was defined as the maximum force in the transition from the first to the second regime. One can
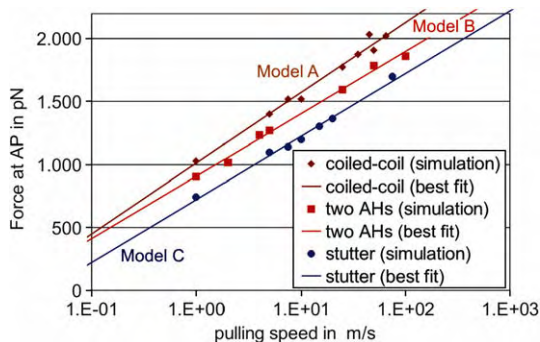
Fig. 70. Unfolding force (at angular point), for the three models, as a function of pulling velocity (results from [247]). All three models display a strong dependence of the unfolding force on the pulling velocity. Model A (perfect CC) shows the largest unfolding resistance, model B (two parallel AHs) has a lower strength. Model C (CC with stutter) has the smallest unfolding resistance. The straight lines represent best fits to the MD results.

Table 11
The parameters E_b, x_b and θ for the three models studied

	Model A, perfect CC	Model B, two aligned AHs	Model C, CC with stutter
MD results			
E_b (in kcal/mol)	5.59	5.67	5.01
x_b (in Å)	0.185	0.200	0.187
θ (in degrees)	23 ± 10.2	16 ± 8.5	16 ± 6.9

We note that while E_b and x_b direct results of MD simulation, the angle θ is determined from the molecular geometry.

clearly see that – in agreement with the predictions by the theory – model B features a reduced unfolding force compared with model A. Thus, the molecular simulations clearly confirm the hypotheses described in the previous section.

Further, fitting the parameters x_b and E_b directly from MD simulation results, following the approach given in Eq. (51), reveals that they are almost identical, suggesting the same unfolding mechanism (see Table 11 for a summary). This provides strong evidence that the strengthening effect is due to the variations in the angle θ. The particular value of $E_b \approx 6$ kcal/mol suggests that the deformation mechanism is characterized by rupture of a single H-bond, in agreement with earlier studies.

For model C, the unfolding force is even lower than in models A and B. In models A and B, unfolding begins only at the ends of the molecule, most probably due to boundary effects. Most importantly, in model C, unfolding occurs also at the stutter, in agreement with the theoretical prediction. This is confirmed by the snapshots of the unfolding dynamics for both model A (Fig. 71) and model C (Fig. 72). It has also been confirmed that the unfolding behavior is independent of the direction of pulling (results not shown).

4.8.4. Extrapolation to physiological pulling rates

The deformation rates accessible to MD simulation are many order of magnitudes higher compared to those that appear under physiological conditions or in experimental investigations. The combination of the Extended Bell Model and MD simulation results provides a possible strategy to extrapolate to much lower pulling rates and thus to overcome the time-scale limitations of MD.

The key to extrapolate towards slower pulling rates is an understanding of the change of unfolding mechanism, as discussed above. The deformation mechanism changes from rupture of single H-bonds at large strain rates (as discussed in the previous section) to simultaneous rupture of three H-bonds.
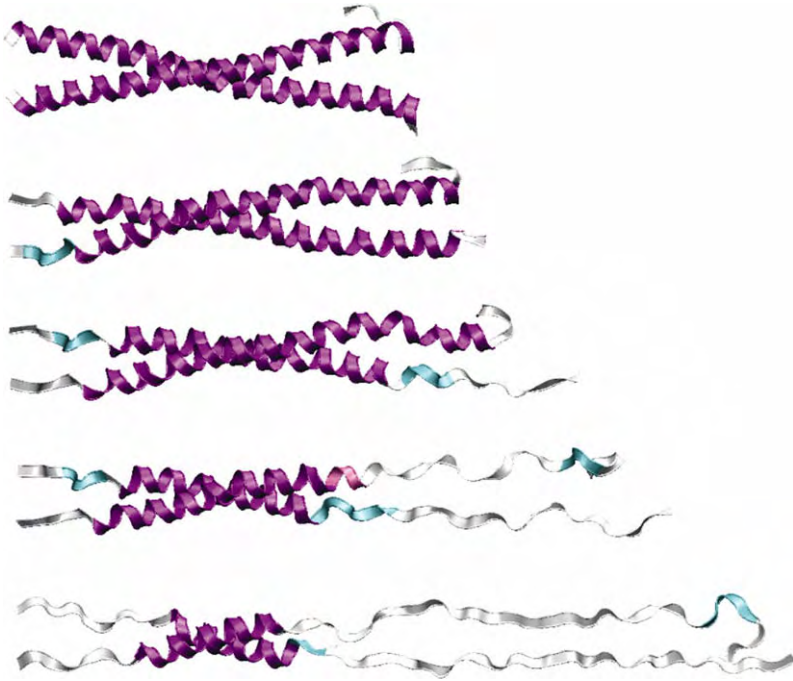
Fig. 71. Snapshots of the unfolding dynamics, model A (perfect CC). Unfolding proceeds sequentially, beginning from the end where load is applied. The structure is rendered using the VMD ribbons method; color is assigned according to the secondary protein structure. The initial length of the structure is 70 Å (results from [247]). (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

For the CC structure, the change for this particular CC structure occurs at a pulling velocity of approximately 0.161 m/s. Since Eq. (52) is generally valid for any deformation mechanism, the Extended Bell Model predicts that the weakening effect of the presence of the stutter will also persist at small deformation rates. For instance, at a pulling velocity of 8E−8 m/s, the unfolding force of model A is predicted to be 220 pN, whereas it is 84 pN for model C. This clearly confirms that the weakening effect of the modified geometry also persists at slower pulling rates. This result is significant because it shows that the discussion reported in this article also holds for physiological conditions. We note that the applicability of the theory to model C is only valid provided that the change in mechanism occurs also in this model. Thus, far we have not explicitly confirmed this using the simulation model and we leave further investigation of these aspects to future work.

We note that these unfolding force levels are somewhat in agreement with experimental results [182,216] that led to unfolding forces ranging from 25 to 110 pN in a range of pulling velocities between 8E−8 and 5E−7 m/s, whereas the predictions from modeling are in the range of 100–200 pN. We emphasize that these experimental studies were carried out on different types of IFs, suggesting that a quantitative comparison may not be immediately possible.

4.8.5. Discussion and conclusion

Vimentin IF dimers play a crucial role in determining the large deformation behavior of eukaryotic cells. In this section, we have compared the mechanical properties of vimentin coiled-coils with and without the stutter molecular defect. Further, we have reviewed systematic studies of the pulling velocity dependence and extrapolated our results to physiologically relevant deformation rates. Earlier studies have suggested that the stutter plays an important role in filament assembly [187–189]. Our work proves that the stutter also has a significant effect on the mechanical behavior of vimentin dimers under tensile loading. In summary, the stutter:
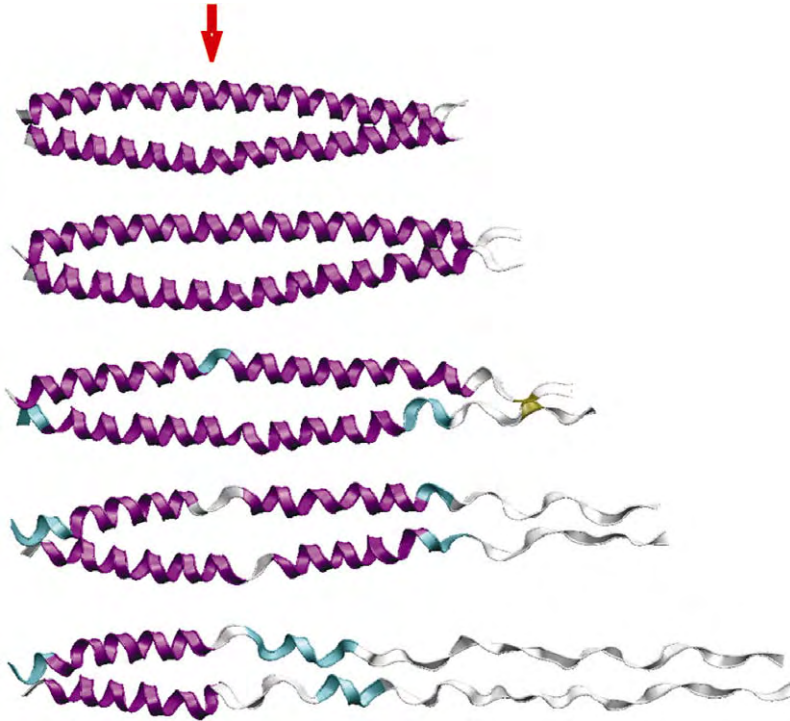
Fig. 72. Snapshots of the unfolding dynamics, model C (CC with stutter). It is apparent that the stutter represents a location where unfolding starts as well (see the second and third snapshot from the top, where it is evident that the AH motif begins to disappear at the location of the stutter). The red arrow indicates the location of the stutter. The structure is rendered using the VMD ribbons method; color is assigned according to the secondary protein structure. The initial length of the structure is 70 Å (results from [247]). (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

- Renders the molecular structures softer, that is, unfolding occurs at lower tensile forces.
- Introduces predefined locations of unfolding.
- Thus leads to a more homogeneous distribution of plastic strains throughout the molecular geometry.

Interestingly, the stutter is a defect that can only appear in the coiled-coil geometry. Our model suggests that single alpha-helices with and without the stutter sequence will have similar unfolding behavior. This is because the weakening effect is not due to the amino acid sequence, but rather due to the variation of the angle θ. However, the variation in the angle θ is not 'visible' at the scale of a single AH.

The stutter only appears in the 2B segment of the vimentin dimer. Notably, this segment is the longest of all segments in the structure (the length of the 2A segment is 158, 139 Å of the 1B segment, 48 Å of the 1A segment and 26 Å of the 2A segment). This corroborates the notion that its significance is related to control the redistribution of plastic strains during unfolding. Furthermore, based on the observed unfolding behavior – namely the occurrence of molecular unfolding at the ends – it is believed that the linkers between the CC segments may have a similar role.

The concept of division of larger structures into small, nano-sized segments with material properties of complementary nature is observed almost universally, such as in bone (hard and brittle mineral phase combined with a soft collagen phase) [4], nacre (hard mineral phase combined with a soft polymer glue) [248] spider silk (strong beta-sheet crystals embedded in a soft, amorphous matrix) [249]

and many other structural biological materials [1,7,12,250]. It is noted that the analysis of the detailed atomistic scale unfolding mechanisms have not been in the focus of this particular study of the stutter mechanics. We leave such investigations to future work.

The review of these studies illustrates the significance and role of utilizing hierarchies in the design of biological materials. Modifications in the lower hierarchy (amino acid sequence) do not influence the immediately following hierarchical level (AHs), but show effects just one hierarchical level higher, that is, on the CC level, where it has a profound impact on the mechanical properties. This illustrates that the impact of a small-scale feature can be silenced at an intermediate scale and become active at a larger scale.

Fig. 73 summarizes the effect of different molecular features. Formation of a CC superhelix out of individual AHs leads to a strengthening effect. Addition of the stutter provides a molecular defect that lead to predefined unfolding locations, effectively leading to a slight weakening of the structure while leading to a more homogeneous distribution of plastic strains. Weakening only persists in force–extension regimes I and II; the strain hardening regime III is not affected by the presence of the stutter. Biologically, this appears to make sense since IF networks are thought to be 'invisible' at small deformation and are only activated at large strains. Thus, the stutter helps to realize this trait.

There exists a delicate balance between material strengthening and weakening providing robustness through geometrical changes, patterned at nano-scale. Our results support the hypothesis that Nature seeks to provide robust mechanical function in biological materials. The assembly into CCs does not only strengthen the material, but it also allows to create structural features that direct towards a more controlled unfolding and uncoiling behavior compared to single AHs (the stutter is a feature that cannot exist based on a single AH protein; it is a property that emerges at the level of a dimer). A slight reduction in mechanical strength is sacrificed in order to obtain a robust and controlled unfolding behavior, independent of the loading rate, maybe following the rule 'safety first'.

4.9. Hierarchical effects: hierarchies of AHs and CCs

Here we review an application of the Hierarchical Bell Model (see Section 4.4.3) in an analysis of strength and robustness properties of AH based protein structures. The main question of the following discussion is: how are protein materials capable of unifying strength and robustness? We present two analyses: the first one illustrates the effect of increasing the number of elements at the same hierarchical scales. The second analysis illustrates the effect of introducing multiple hierarchies on the performance of the elements.
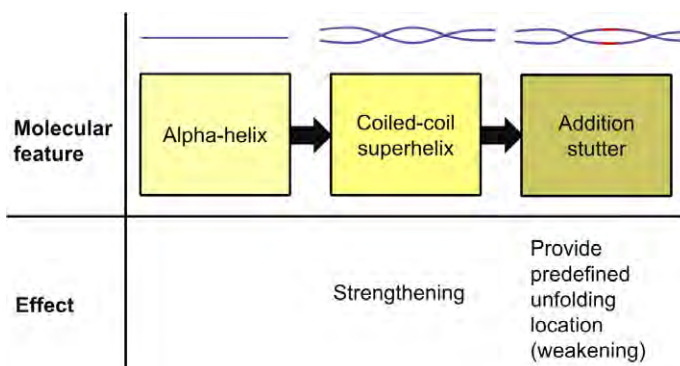
Fig. 73. Summary of the effect of different molecular features (results adapted from [247]). Formation of a CC superhelix out of individual AHs leads to a strengthening effect. Addition of the stutter provides molecular defect that lead to predefined unfolding locations, effectively leading to a weakening of the structure. Presence of the stutter leads to a 'composite' nanostructure that combines a perfect CC with two parallel aligned AH molecules. We note that the weakening only persists in force–extension regimes I and II; the strain hardening regime III is not affected by the presence of the stutter (for a definition of regimes see Fig. 4). Biologically, this makes sense since IF networks are 'invisible' at small deformation and are only activated at large strains. Thus, the stutter helps to achieve this feature.

Fig. 74. Overview over three model systems, including (a) single AH, (b) coiled-coil with two AHs (CC2) and (c) a coiled-coil protein with four AHs (CC4). The schematic on the right shows the abstract representation of each physical system. Each blue line represents a single H-bond (in the atomistic simulations, all structures are embedded entirely in a water skin during deformation; howe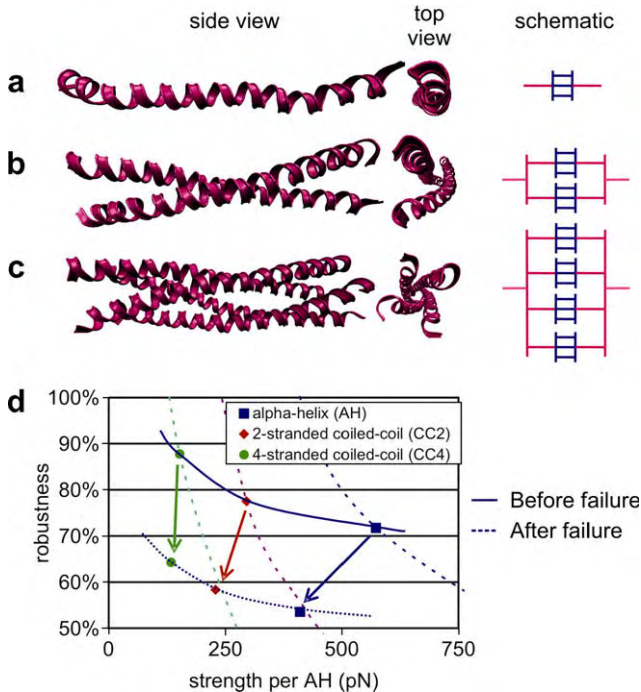ver, water molecules are not shown here for clarity). Subplot (d): Performance of the analyzed structures in the strength per AH-robustness space, before and after failure. The coarsely dashed lines represent levels of equal strength (s)–robustness (r) potential (that is, the product of both values is equal on these lines, $r \cdot s = $ const.). Robustness and strength compete on these lines. Different combinations of these two parameters can be achieved with the same elements, but under different geometrical arrangements. The figure illustrates that vimentin coiled-coils have a robustness degree of 80%. The first data point for each structure represents the intact system, whereas the second data point shows the system after failure. Thereby, robustness is defined as parameter insensitivity. In this case robustness equals to the force from hierarchical strengthening of a defect system (two instead of three HBs rupture simultaneously) divided by the force of an intact system (all three HBs rupture at once). As we can see, a defect in a system moves the system to another potential line. For example, due to the high level of robustness, the CC4 structure hardly changes its strength, whereas the strength of a single AH is significantly reduced. This illustrates how robustness and strength can be combined in dependence of the requirements.

We begin with a review of an analysis of three AH-based structures, as shown in Fig. 74. Fig. 74a–c depicts the architecture of the considered structures, and Fig. 74d shows the performance of the analyzed structures in the strength per AH-robustness space before and after failure of one basic element. Strength and robustness are not completely independent parameters; with increasing number of hierarchies, the robustness increases due to the increasing number of redundancies present in the structure. However, this goes hand in hand with a loss of strength per element (that is, per AH), since the overall strength of a structure is not directly proportional to the number of parallel elements. The strength of the failed system would be highest for an infinite number of elements on each scale – leading to a robustness value of 100%. However, increasing the number of elements at a specific hierarchical scale is inefficient, as it leads to extensive material use and a decrease in strength.

Therefore, the introduction of multiple hierarchies becomes significant. To illustrate this, eight single AHs are arranged in different hierarchical structures – asking the question: how can one arrange eight AHs to obtain various levels of robustness and strength? As shown in Fig. 75, the systems consist of two, three and four hierarchies. The differences in robustness and strength are remarkable, since they are not achieved through additional use of materials, but purely through different hierarchical
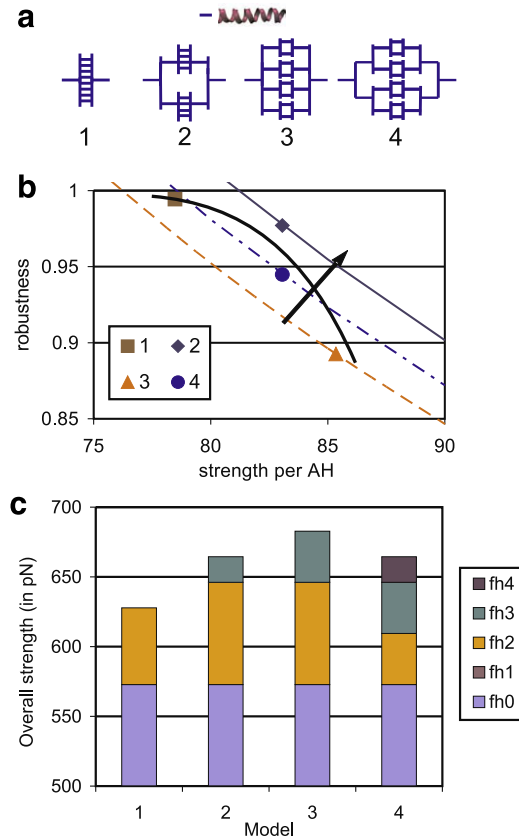
Fig. 75. This figure shows an example of four different structures with the same amount of sub elements (that is, eight AHs) but in different hierarchical arrangements. Subplot (a) shows the four different architectures. For simplicity, individual HBs on the lowest hierarchical scale are not shown; instead one line represents three HBs as one AH. Subplot (b) shows the concurrence between strength and robustness, which depends on the degree of redundancies on the different hierarchical levels. The level of robustness increases with increasing redundancies on a particular level. It seems that in dependence of the hierarchical arrangement of the elements different potentials of strength and robustness can be achieved. The arrow illustrates the direction of maximal strength robustness ratio (that is, high strength at high robustness). Subplot (c) shows the contributions of each hierarchy to the overall strength (not strength per AH, as shown in subplot (b)). As we assume that in each AH three HBs break simultaneously, each structure (having an AH as the smallest sub element) has the same contribution from h_0. This is also the highest amount of strength contribution and shows the significance of the strength of HBs, which depends on the solvent and the environment. This also shows the link between chemistry and mechanics, where variations in pH or addition of salts can significantly change the mechanical properties of a structure. The other contributions are of 'hierarchical' origin. The force contribution from h_1 is zero, as we assume that 3 of 3 HBs will break, which lowers the logarithmic multiplicity term to zero.

arrangements. To the best of our knowledge, this tuning of properties in the strength-robustness domain as reviewed here has been shown for the first time. In the robustness-strength map, the 'best' material behavior is the one in which high robustness is achieved at large strength – referred to as a 'high potential' in the following discussion. It can be seen that, remarkably, system 2 has the highest potential. It is notably not the system with the highest hierarchical level (system 4), nor the system with the highest level of redundancies (system 1). In other words, system 2 features the best combination of redundancies on the different levels.

With different structural arrangements, different combinations of strength and robustness can be achieved. This finding is a key result of this analysis: it illustrates that the conflict between strength and robustness can be resolved by introducing hierarchies as an additional design variable. These

results further suggest that the level of hierarchical depth and strength may be balanced in biological protein materials, since the robustness and strength are not completely inversely proportional, allowing biological materials to maximize the mechanical performance while minimizing the use of materials.

The analysis shown in Fig. 75 may also explain the abundance of CC2 protein motifs in biology: it is a structure that unifies robustness and mechanical resistance. Overall our analysis illustrates that the introduction of hierarchies is the key to unify disparate properties. Applying this insight will allow an extended use of hierarchies in bio-inspired or biomimetic synthetic materials at nano-scale, such as hierarchically organized CNT-bundles, nanowires or polymer–protein composites [251–253]. Understanding the fundamental design laws and processes of hierarchical protein materials enables us to link the structural protein organization to the appropriate biological functions. This is vital for the development and application of new organic and organic–inorganic composites, which will consist of elements that are available in virtually unlimited amount (C, H, N, O, S). Recently it was reported that the first bundles of coiled-coil proteins were generated synthetically [254]. Further, other peptide generation techniques are progressing rapidly [254–260]. This theory combined with these new manufacturing techniques may be the first step towards a *de novo* bottom-up structural design.

4.10. Deformation mechanics of vimentin tetramers

This section concludes with an analysis of the deformation mode at the next higher hierarchical scale in the vimentin network, the mechanical properties of tetramers (see Figs. 50 and 75 for a detailed view) [261]. This analysis is facilitated by the improved understanding of the mechanics of individual CC domains as reviewed in the sections above. The central point of the analysis is an estimate of the interaction forces between two CC dimers to better understand the dominant deformation mechanism under tensile loading of IF filaments. The rod-like structure is 310 residues long and consists of four CCs (1A, 1B, 2A, 2B), divided by several linkers [166,180,185,188].

Fig. 75 depicts how interdimer adhesion is facilitated. Extensive experimental analyses of the assembly process revealed that the adhesion between two CCs is dominated by the interaction of the head domain with the other dimer (the head domain is indicated with the red color) [166,174]. The head domain is primarily positively charged at pH 7 (due to approximately 15% Arg residues), and folds back into the CC 1A domain, primarily negatively charged at pH 7, and also connects to the 2B domain of the other dimer, which is also negatively charged (the 1A and 2B domains contain approximately 11% Glu and Asp residues). Details of the interdimer bonding are schematically shown in the blow-up (lower part of Fig. 57). It is noted that it is believed that the two CCs itself do not directly interact, which has been confirmed by analyses of the intermolecular distance between the two AHs [174]. Fig. 76 illustrates how lateral tensile load, applied at the level of a tetramer, is transmitted via interdimer shear traction.

4.10.1. Theoretical model

The description of the rupture mechanics of protein structures requires incorporation of statistical models that explicitly consider the energy landscape of the interatomic or intermolecular bonding. The unfolding force of a CC structure in physiological conditions varies between roughly 50 and 200 pN for deformation speeds varying from 5E−8 to 1E−6 m/s. These unfolding forces are somewhat in agreement with experimental results [182,216] that led to unfolding forces ranging from 25 to 110 pN in a range of pulling velocities between 8E−8 and 5E−7 m/s.

We emphasize that these experimental studies were carried out on different types of IFs, so that a quantitative comparison is not possible. However, both experiments and theory predict unfolding forces in the range of around slightly below 100 pN to lower hundreds of pN at physiologically relevant deformation speeds. In particular, F_{AP} = 90 pN for a deformation speed of 1E−7 m/s, which serves as an example for the subsequent analysis. Under tensile loading of a vimentin tetramer, the forces are distributed predominantly as tensile load carried by individual CCs, and as shear forces between different CC dimers (see Fig. 76). The key to arrive at insight into this question is therefore to consider the interplay of two competing mechanisms:
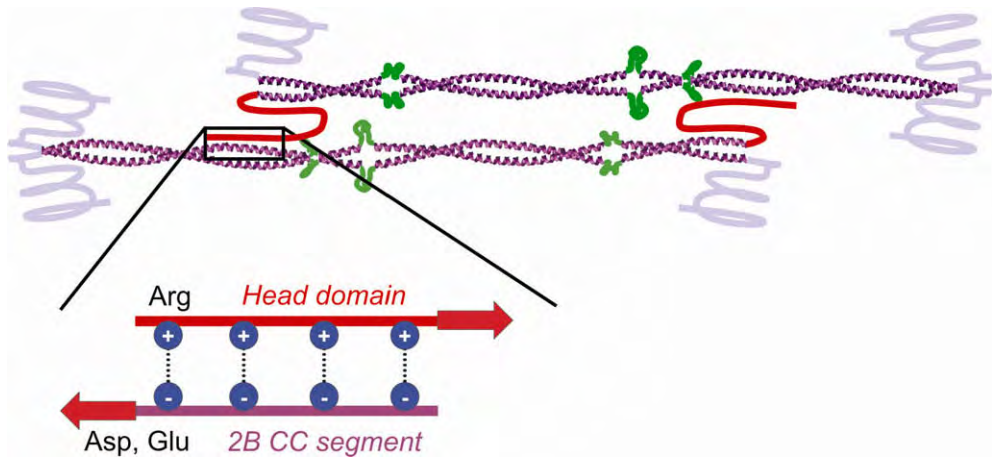
Fig. 76. Schematic of the tetramer, formed by interaction of two dimers (see Ref. [261]). Two dimers are assembled in the A_{11} assembly pattern antiparallel, approximately half-staggered, as suggested in experimental studies [174]. Interdimer adhesion is facilitated by the interaction of the head domain with the other dimer (indicated with the red color). The head domain is primarily positively charged (approximately 15% Arg residues), and folds back into the CC 1A domain, negatively charged, and also connects to the 2B domain of the other dimer, which is also negatively charged (the 1A and 2B domains contain approximately 11% Glu and Asp residues). The details of the interdimer bonding is schematically shown in the blow-up (lower part of the figure). (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

- Molecular unfolding of the CC dimer, mediated by rupture of H-bonds (as discussed in Section 2.2).
- Interdimer sliding, mediated by rupture of the head–CC interaction bonds.

The key question that must be quantified is whether or not the head–CC interactions break first, or the CC unfolding begins first. Which of the two has a lower critical force controls the dominant failure mode. Generally, for strong interdimer adhesion, CC unfolding will occur first. For weak adhesion, interdimer sliding dominates. The first step in this analysis is the determination of the critical adhesion properties of the dimer-interaction that distinguishes these two deformation regimes.

4.10.2. Critical interdimer adhesion strength

The discussion reviewed here is focused on an estimate of the critical interdimer adhesion to induce interdimer sliding rather than CC unfolding. In other words, the question is how weak should the interdimer interaction be, in order to allow for sliding rather than unfolding, and vice versa? A first simple model provides some insight into the nanomechanics of this problem. It is assumed that the shear force transmitted between two CC dimers is directly proportional to the contact length L_C (note that L_C corresponds to the total contact length, accounting for the fact that there are two contact regions as shown in Figs. 75 and 76),

$$F_{\text{tens}} = \tau_{\text{shear}} L_C, \tag{58}$$

where F_{tens} is the applied force in the axial molecular direction. This equation can be rewritten to solve for the critical adhesion strength $\tau_{\text{shear,crit}}$ in order to induce unfolding of the CC structure *before* interdimer sliding occurs:

$$\tau_{\text{shear,crit}} = \frac{F_{\text{AP}}}{L_C}, \tag{59}$$

where F_{AP} corresponds to the critical unfolding force of each dimer. The analysis reviewed in above provides estimates for the unfolding force of a dimer. For instance, $F_{\text{AP}} = 90$ pN for a deformation speed of 1E−7 m/s. Assuming the contact length ($L_C = 20$ nm) (counting the two head–CC interactions, each

of which is 10 nm long as illustrated in experiment [174]), the critical shear resistance is $\tau_{\text{shear,crit}} = 0.45$ pN/Å.

This result suggests that for $\tau_{\text{shear}} > \tau_{\text{shear,crit}}$, unfolding of the CC structure rather than interdimer sliding would occur. In contrast, for $\tau_{\text{shear}} \leqslant \tau_{\text{shear,crit}}$, interdimer sliding will be the predominant deformation mode.

In comparison, the intermolecular adhesion of tropocollagen (TC) molecules, mediated by a combination of electrostatic and H-bond interactions, was determined to be $\tau_{\text{shear,TC}} = 5.5$ pN/Å by a series of full atomistic MD simulations [26,27]. While it is rather difficult to relate this result to the specific geometry present in the head–CC interaction, this value suggests that the critical interdimer adhesion strength could be close to the interactions created by electrostatic and H-bond interactions as in the case of TC.

While the adhesion strength provides a useful measure to obtain an overall estimate of the interdimer adhesion, more details of the chemical interactions must be considered in order to arrive at a more quantitative understanding.

Experimental analyses [174] suggest that the head–CC interaction is dominated by Coulomb bonds or electrostatic interactions, formed by the interaction of partially charged Arg (positive, in the head) and Glu and Asp residues (negative, in the CC). These results suggest that the interaction between the head and CC domain is formed by formation of a cluster of interdimer weak bonds. In light of these interdimer bonds, it is useful to approximate the critical interdimer adhesion in terms of the critical shear force that must be transmitted in each cluster of bonds. The critical force exerted at each interdimer traction site is $F_{\text{shear,crit}} = 45$ pN (assuming that each adhesion region contributes the same to the shear traction). In the following sections, we will consider a statistical model to estimate the shear strength for different bond breaking scenarios.

4.10.3. Estimate of deformation mode from molecular structure

The elucidation of the specific, quantitative interdimer adhesion is a rather difficult endeavor, which could for instance be undertaken by carrying out extensive reactive MD simulations that explicitly considers the atomistic geometry and charge transfer.

Here we limit the consideration to a simple theoretical analysis that does not explicitly consider the particular atomistic geometry of the IF structure, but aims at a fundamental investigation of the force levels and rupture mechanisms. The focal point of the analysis is the consideration of two distinct bond breaking mechanism at the head–CC interface. Two cases are considered, (i) sequential tensile breaking of interdimer bonds as shown in Fig. 77a, and case (ii), concurrent breaking of interdimer bonds due to homogeneous shear, as shown in Fig. 77b. The reason why these two cases are considered is that the precise mechanism of loading of the interdimer shear interactions remains unclear.

Both cases are treated by considering a statistical model based on the Extended Bell Theory, by considering a discrete assembly of weak interaction bonds as shown in Figs. 75 and 76. It is assumed that the equilibrium spacing of the interdimer bonds is $r_0 \approx 3$ Å, which is the characteristic equilibrium distance for a non-covalent bond in organic systems [262]. We assume that the transition state for bond rupture occurs at a critical bond separation of $x_b^* = 1$ Å (that is, at a total bond length of $r_0 + x_b^* = 4$ Å). These values can be estimated from experimental analyses of weak intermolecular bonding, as re-
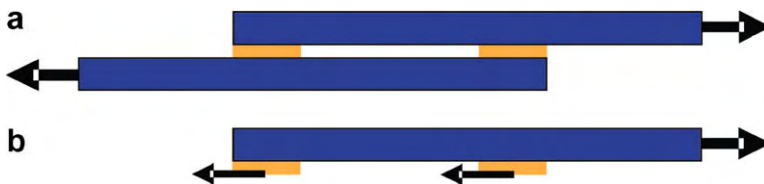
Fig. 77. Schematic representation of the dimer geometry and the interdimer traction. Subplot (a) depicts the dimer structure under lateral tensile loading. Subplot (b) depicts the upper half of the structure, showing the internal tractions, transmitted via the head–CC interaction as illustrated in the blow-up in Fig. 76 (see Ref. [261]).

ported in [262]. The two cases are different in how the internal bond breaking mechanisms depends on the laterally applied load, and how the energy barrier changes for different mechanisms.

4.10.4. Sequential bond breaking

Under sequential bond breaking (Fig. 78a), the critical breaking force is given by

$$f(v) = \frac{k_b \cdot T}{x_b} \cdot \ln v - \frac{k_b \cdot T}{x_b} \cdot \ln v_0 = a \cdot \ln v + b \tag{60}$$

with

$$v_0 = \omega_0 \cdot x_b \cdot \exp\left(-\frac{E_b}{k_b \cdot T}\right), \tag{61}$$

as pointed out above (Eqs. (59) and (60) resemble Eqs. (46) and (47) but are repeated here for convenience). It is noted that x_b relates to the global displacement necessary to achieve a local displacement of x_b^*. Therefore, the parameter $x_b = 1$ Å for the case of sequential breaking, since the force is applied directly into the loading direction.

These equations can now be used to estimate the critical energy barrier to provide sufficient resistance so that the interdimer interactions is stronger than the CC itself. For a deformation speed of 1E−7 m/s, the critical energy barrier is approximately 15 kcal/mol (this estimate is obtained by assuming that the local stretching speed sensed at the head–CC interface is the same as the laterally applied deformation speed). In other words, if each interdimer bond has an energy barrier of 15 kcal/mol or larger, the interdimer adhesion is stronger than the CC structure, so that unfolding of the CC structure is the predominant mechanism. It is noted that the energy barrier of 15 kcal/mol corresponds to the strength of several H-bonds breaking simultaneously (we refer the reader to the discussion of the mechanics of individual AHs, as reviewed above).

4.10.5. Concurrent bond breaking

The case of concurrent bond breaking (Fig. 78b) can be treated similarly as the previous case. With $v_{0,N}$ as the as the natural bond breaking speed of a collection of N bonds (each of which has an energy barrier E_b) defined as

Fig. 78. Schematic to illustrate possible interdimer bond breaking mechanism. Subplot (a) depicts sequential breaking, which would occur if the load is applied perpendicular to the molecule axis. Subplot (b) depicts concurrent bond breaking, which would occur if shear load in the direction of the molecular axis is applied (see Ref. [261]).

$$v_{0,N} = \omega_0 \cdot x_b \cdot \exp\left(-\frac{N \cdot E_b}{k_b \cdot T}\right), \tag{62}$$

the rupture force at a given deformation speed v is given by

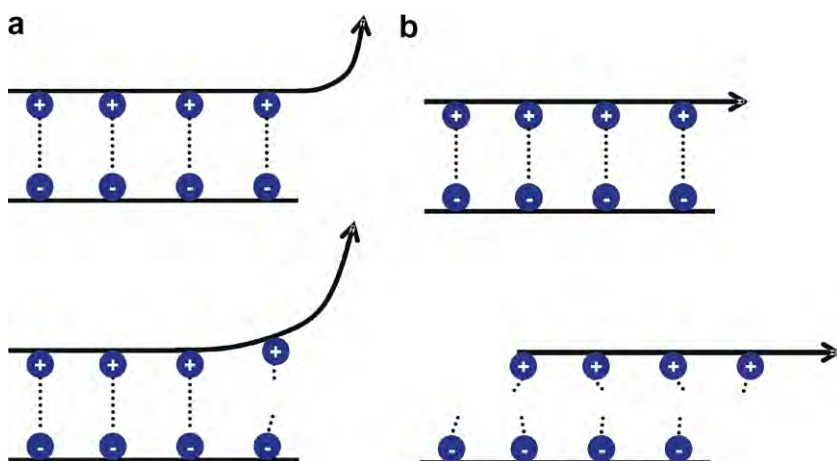$$f(v,N) = \frac{k_b \cdot T}{x_b} \cdot \ln v - \frac{k_b \cdot T}{x_b} \cdot \ln v_{0,N} = a \cdot \ln v + b, \tag{63}$$

or by considering the expression for the natural bound breaking speed $v_{0,N}$, as

$$f(v,N) = \frac{k_b \cdot T}{x_b} \cdot \ln v - \frac{k_b \cdot T}{x_b} \cdot \ln \left[\omega_0 \cdot x_b \cdot \exp\left(-\frac{N \cdot E_b}{k_b \cdot T}\right)\right] = a \cdot \ln v + b. \tag{64}$$

The underlying assumption in this model is that the total barrier against rupture is given by the total number of bond that break simultaneously, N, times the energy per bond, E_b.

Also, the parameter x_b is larger than in the previous section, since the load application in shear requires a larger shear displacement in order to reach the local transition state for the individual bond (x_b^*). Thus in this case, the parameter x_b = 2.65 Å (this value is estimated by assuming that sliding occurs without a change in the interdimer distance, and that the bond must be stretched by 1 Å in order to rupture; this leads to a lateral displacement of 2.65 Å). In this case the total energy barrier $E_{b,tot} = N \cdot E_b$ = 17.7 kcal/mol. The energy per bond is given by $E_b = E_{b,tot}/N$.

The number of possible bonds that can be formed between the head and CC domain can be estimated by considering the details of the head–CC interactions. Experiment suggests that each 'bond' consists of an interaction of an Arg residue in the head domain, with either an Asp or a Glu residue in the CC segment (see blow-up in Fig. 76). The analysis of the specific sequence of amino acids in the head and CC domain reveals that approximately 11–14% of all residues are polar, charged residues [188]. There are 80 residues in the head domain, half of which fold back to the CC, and half of which fold onto the other dimer, so that there are approximately four Arg residues in contact with the CC, due to the 11–14% of all residues participating in the bonding. That is, approximately four interdimer 'bonds' can form between the head domain and the 2B segment of the CC in the overlap region.

For this case, the critical energy barrier for the sequential breaking case is estimated to be 4.425 kcal/mol. This value is on the order of the energy of a single H-bond [262]. It is observed that the shear deformation mode requires a lower energy per bond, since multiple bonds contribute to the shear resistance.

4.10.6. Estimate of the interdimer bond energy

As indicated above and suggested from experimental analyses [174], the head–CC interaction is dominated by Coulomb 'bonds' or electrostatic interactions, as modeled in many empirical force fields [53,56,58]. The energy landscape of electrostatic interactions can be approximated using the Coulomb law,

$$E(r_0) = \frac{1}{4\pi\varepsilon_0} \cdot \frac{|q_1||q_2|}{r_0}, \tag{65}$$

where ε_0 = 8.85E−12 C^2 N^{-1} m^{-2} (permittivity of free space), and $|q_i|$ describes the two charges (modeled as point charges). Note that r_0 refers to the equilibrium spacing of the particular Coulomb 'bond' considered. As discussed above, we assume that $r_0 \approx 3$ Å, the characteristic equilibrium distance for a non-covalent bond in organic systems. It is assumed that the interdimer bonds are formed between the side chains of the participating amino acid residues without presence of water molecules between the bond. An overview over the molecular geometries of the participating amino acid residues (Arg, Glu, Asp) is provided in Fig. 79.

The charges $|q_i|$ can be determined from an analysis of the atomistic charge distribution of the participating residues. All considerations are carried out at pH 7. The head domain contains strongly positive charges due to Arginine (Arg) residues. Since the Arg side chain has a pK_a = 12.48, at pH 7 the side chain is deprotonated and charged positively. According to the charge distribution predicted by the CHARMM force field [56], the resulting partial charge of the end of the side chain is $|q_1|$ = +0.12 elementary charges (see Fig. 79).
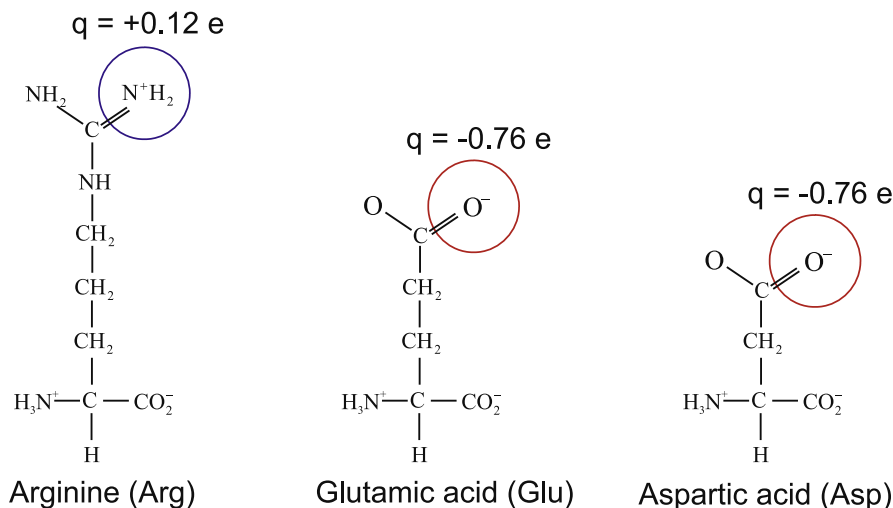
Fig. 79. Interdimer interaction, atomistic details. Interdimer interaction is facilitated by electrostatic interactions between the head and 2A domain of the CC structure. Based on sequence analyses of the vimentin structure, it has been confirmed that the head domain is rich in Arg amino acids (partially positive charged), whereas the 2A domain of the CC is rich in Asp and Glu amino acids, which are partially negative charged at pH 7 (the symbol "e" refers to elementary charge).

The 1A/2B segment of the CC domain contains negative charges due to presence of glutamic acid (Glu) residues with $pK_a = 4.07$, as well as aspartic acid (Asp) residues with $pK_a = 3.86$. In Asp and Glu, the side chain is thus protonated, with a total partial charge of the end of the side chain of $-|q_2| = -0.76$ elementary charge [56].

This information now enables us to estimate the bond energy of head–CC interactions. The energy barrier to break such a bond is given by $E_b = E(r_0)$ (since $E(r \to \infty) = 0$), with the estimates for $|q_i|$ as discussed above. The resulting energy stored in each bond is approximately 10 kcal/mol. It is emphasized that the analysis reviewed in this section is simplistic, but it provides one with a first estimate of the interdimer adhesion strength. It is noted that the value reported here is in the same order of magnitude as the bond energy reported for electrostatic dominated ligand–receptor binding in streptavidin–biotin complexes [84], suggesting that the values are reasonable.

4.10.7. Prediction of predominant deformation mode

The analysis in the previous section reveals the predominant deformation mode depends critically on the details of the head–CC interaction. The analysis reported above led to critical energy barriers under two different deformation modes, being 15 kcal/mol (sequential bond breaking) and 4.25 kcal/mol (homogeneous bond breaking under shear). The analysis of electrostatic interactions led to a quantitative estimate for the bond energy in the range of 10 kcal/mol. This value lies in between the two extreme cases considered.

It is possible that the particular interactions are indeed in the range of the critical adhesion, so that under tensile deformation the 'optimal' shear force is transduced to reach the limiting tensile strength within each CC. In other words, the interdimer glue (formed by the head–CC assembly) and the CC itself are both close to their strength limit. The analysis reviewed here is not able to clearly rule out one mechanism over the other. Since the estimate of the electrostatic bonding is rather crude, additional analysis is required to clarify the details of the interaction.

4.10.8. Discussion: rupture mechanics of tetramers

The analysis provides insight into critical adhesion forces and critical adhesion energies that lead to either CC dimer unfolding or interdimer sliding. Such models are crucial to advance the understanding

of biological process like mechanosensation, which critically depend on the nanomechanical properties of intermediate filaments.

The simple analysis put forward here suggests that the interdimer adhesion provided by the particular interaction of Arg residues with Glu and Asp residues may be at the borderline between CC unfolding and interdimer sliding. This suggests a possible balance between the two mechanisms. If this were indeed the case, the system would be structured so that the "optimal" shear force is transduced to reach the limiting tensile strength within each CC. In other words, the structure may be adapted to make optimal use of the interdimer "glue" material.

We emphasize that while this conclusion is somewhat speculative at this point, it relates well with earlier observations in the structural analysis of the collagen fibril structure [27]. It was found that the critical length scale of tropocollagen molecules may be controlled by the driving force to maximize the tensile forces in each molecule and the maximum shear that can be transmitted between molecules. Future investigation are required to clarify these issues in more detail. We refer the reader to the discussion in Section 4.9 for a comparison with experimental observations of plastic deformation of IF filaments.

The study discussed here has several limitations, and clearly, it only represents a first approximative model of the complex mechanics of tetramer structures. For instance, the quantification of the interdimer adhesion was achieved using a simplistic model based on Coulomb interactions. The atomistic details of the interdimer bonding remain elusive, and could be addressed by extensive MD simulations with reactive force fields, for instance. Further, the details of the interdimer bond rupture processes are entirely unknown. This could, however, significantly influence whether or not interdimer sliding or CC unfolding is the domination mechanism of deformation. The use of the Bell Model to estimate the forces provides another limitation, as this method does not describe the intermolecular adhesion accurately, for instance due to the neglect of the change of the transition point x_b as a function of applied force.

4.11. Discussion and conclusion

This discussion is based on the text published in [138]. In this section, we have focused on a series of atomistic modeling of tensile deformation of single and coiled-coil AH structures in the rod domain of a vimentin dimer. We have observed the existence of strain hardening and viscoelasticity, phenomena that have previously been known to exist at the filament level [6]. Our work confirms that these also exist at the molecular, dimer level.

Through systematic comparison of a single AH with a coiled-coil structure, we explored differences in the nanomechanical behavior of AH structures due to the first level of hierarchical assembly. We find that the less steep increase in force during the first and third regime, and the lower strain at the angular point are consequences of the superhelical structure. We find that the third regime of the coiled-coil features a smoother and less steep strain hardening, since uncoiling of the coiled-coil structure appears in addition to the stretching of the backbone. Some of the differences between AH structures and coiled-coils are summarized in Table 8.

It has also been observed that the moleuclar-scale unfolding mechanisms of the single AH structure at large pulling velocities is much less controlled, leading to simultaneous rupture of H-bonds at several residues at the angular point, in contrast to the force–strain curve of the coiled-coil structure, where a systematic unfolding appears by propagation of an unfolding wave. We thus conclude that coiled-coils are mechanically more stable than single AH structures. This is also exemplified by the smooth change from the first to the second regime at the AP, the continuous unfolding during the second regime and the less steep increase in force during the third regime. This could also explain why coiled-coil structures appear universally in cells (e.g. in various types of IFs, inluding lamin, vimentin, desmin etc.) and tissues (e.g. muscle myosin), where large and continuously changing mechanical loads occur.

We have reviewed an analysis of the pulling rate dependence of the coiled-coil protein domain, and developed an analytical model that enables us to predict the force–strain curves for variations in strain rates. This model appears to provide a suitable representation of the physical mechanisms for pulling rates that range over eight orders of magnitudes, and can thus be used to calculate the behavior of

vimentin proteins at pulling rates used in experiments (see Figs. 58, 64 and 65) as well as its behavior due to cyclic loading (Fig. 62). To the best of our knowledge, this is the first atomistically informed continuum model capable of describing the nanomechanical properties of vimentin coiled-coil proteins over several magnitudes of pulling rates. Based on the atomistic simulation results, we propose a change in mechanism that occurs at of approximately 1 m/sec (Fig. 65).

It was observed that the boundary conditions can play a key role in the unfolding behavior under large pulling velocities, as the H-bonds ruptures first at the fixed and the end at which the pulling force is applied. This is observed even if the fixed and pulled atoms are several convolutions away from the protein ends. This suggests that the linkers between the coiled-coils in dimers, being the "flexible" boundaries of the coiled-coils, might change the unfolding behavior and enable the coiled-coils to withstand higher forces before unfolding appears. This would give the linker a specific mechanical role in the dimer, and may change the energy landscape in particular at large pulling velocities. We may explore these aspects in future studies. The studies illustrate how Nature realizes superelasticity in vimentin coiled-coils and demonstrate that stress buffering – already known from larger scale IF networks – is also implemented at the molecular level of individual proteins.

Until now, few experimental results about the mechanical properties of IFs at the single filament level are available. AFM studies were carried out recently for pulling desmin dimers from the surface of a filament [182], and for performing bending tests on 300 nm long filaments [83]. Additionally, AFM cantilevers were moved perpendicular to the filament axis in order to analyze the stretching and the rupture behavior of IFs [181]. Results of these experimental AFM studies are shown in Fig. 80, illustrating how a IF structure undergoes large tensile deformation under applied force. Glass micro-beam force transducer apparatus were used to measure tensile properties of single keratin IF from hagfish threads [172,177,178].

In recent experiments [170], individual filaments were manipulated with the AFM cantilever perpendicular to the filament axis with forces between 30 and 40 nN and velocities of 0.4 μm/s. It was found that the breaking extension at the rupture point is up to 3.6 times of the initial length and the average extension over all experiments was 2.6 (160% strain). As there are 16 dimers in one filament, the applied force per dimer in this experiment was of about 1800–2500 pN.

By fitting these forces and velocities in our continuum model, protein strains of 150–160% are estimated. At this strain, the coiled-coil structure is in the third regime (the regime of strain hardening), but still far away from rupturing. The strength of covalent peptide bonds was reported to be in the order of 8 nN (and 16 nN for the coiled-coil [72]). Therefore, in addition to the protein extension, slip between proteins may occur, which may lead to additional strains of up to 50%. This number has been derived by taking the length of a tetramer (approximately 60 nm; two dimer, half staggered overlap) and by assuming that after slip has taken place, the dimers are arranged sequentially, leading to a
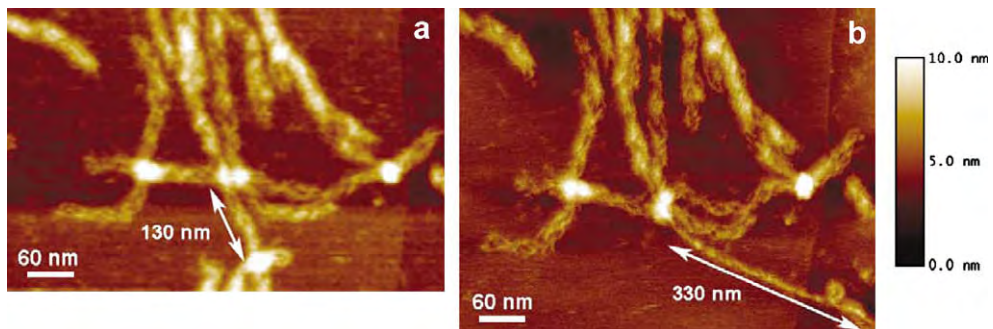
Fig. 80. Human keratin IFs, stretched using an AFM tip. Subplot (a) shows the control tapping mode AFM image, scanning speed 4000 nm/s. Subfilamentous nanostructures can be seen along the filament axis. Subplot (b) shows the results after applying a force of 30–40 nN. A piece of filament was stretched from 130 to 330 nm length, that is, approximately 2.5-fold. It can be seen that the diameter of the filament is significantly reduced after stretching. Reprinted with permission from Ref. [181]. Copyright © 2005 Elsevier Ltd.

maximal length of 90 nm (two times 45 nm). Relating the difference in length (30 nm) to the initial length (60 nm) we receive a maximum strain due to slippage of 50%. Consequently, the breaking strain being the sum of stretching and slip, can be estimated to be on the order of 200%. This is in rather good agreement with experiment (160% average, and maximum of up to 260%).

These observations let us conclude that yield in vimentin filaments features both slip between proteins, in addition to unfolding of the dimer structure. In particular, this concept explains the highly reduced diameter of the ruptured filaments, being a result of the unfolded proteins; notably, the observed plastic strains can only be explained by considering both mechanisms. The analysis reviewed in Section 4.8 supports this hypothesis. Up until now this was only suggested [177,181]. Future AFM studies could be used to determine the forces at which slip appears, which will help to understand the mechanical properties of vimentin on the filament level.

Tensile tests were recently performed on wet keratin IFs in hagfish threads [177,178]. The curves observed in these experiments contained four regimes, whereof the first three regimes appear very similar to the regimes that we observed in our simulations. Not only the shape of the curves, but also the strains where the changes between different regimes take place fit rather closely (see, for instance the results shown in Fig. 58 [138]). For example, the angular point appeared in experiment at a strain of 34% (in simulation 25%), and the beginning of the third regime was in experiment at 100% of strain (in simulation 120%).

Additionally, if we assume that each thread (diameter of 10 nm) has in average 16 dimers (a reasonable assumption based on its diameter), we can roughly calculate the average area per dimer to be 480 $Å^2$, and use this ratio for comparing Young's Modulus between the model and experiment. The simulation model predicts a modulus of approximately 25 MPa, which is in good agreement with experiment (10 MPa). Additionally, the second regime in experiment was demonstrated to begin with the opening of the AH (also known as the alpha-to-beta transition). This was also observed in the simulation at the beginning of the second regime [138].

Comparing the simulation results with those from experiments, it has been shown that the specific mechanical properties of vimentin IFs at least partly originate from the protein level [138]. Among others, the particular molecular nanomechanics characteristic to the IF protein structures explains how the large extensibility known from filament rupture experiments is possible and determined the slippage between dimers as the weaker link in filaments.

4.11.1. Linking results to other coiled-coil structures and interpretation of results in light of biological function

The force–strain curves observed in the simulations reviewed here are similar to earlier MD simulations carried out on different types of AHs and coiled-coils [217–220], as well as to previous experimental studies on other alpha-helical coiled-coil structures such as myosin [216] or desmin intermediate filaments [182], even though the force levels are typically much larger due to the very high pulling velocities.

Due to the similarity in the curve shape, the similar strain levels in experiment and simulation, as well as the similar force levels in simulations for different coiled-coil structures, we conclude that the secondary structure, reinforced by H-bonds plays a critical role in controlling the mechanical behavior during tensile deformation.

The mechanical properties of metazoan cells, its rigidity at high stresses and its integrity are mainly due to vimentin IFs. However, until now it was not known exactly from which level of hierarchy the specific mechanical properties arise. Through the molecular modeling work reviewed here it has been shown which mechanical properties appear on the protein level.

We hypothesize that coiled-coils play an important role in biology as a simple construct in order to realize superelasticity in biological materials. The deformation behavior of coiled-coils is not only elastic but also contains a long plateau at constant force, reaching mechanical strains of more than 150% (Figs. 58 and 62 for the FDM and Fig. 64 for the SDM). Thereby, the very high and reversible deformation of coiled-coils is realized by a second, stress induced phase, which equals to the rupture of the H-bonds followed by an immediate unfolding. Even though rupture of H-bonds represents a breaking of a chemical bond, it is completely reversible and the structure can self-heal. Thus once the tensile load

is reduced, the second phase becomes unstable, or in other word, the structure folds back to the more stable helical shape by reforming the H-bonds (the resulting mechanical behavior is shown in Fig. 62).

There is another interesting aspect to this observation the second phase may be used as the "security belt" of the cell, since covalent bonds rupture at forces of three magnitudes higher than H-bonds and thus the ability to undergo such large deformation without causing permanent damage could be a crucial element in this ability. Additionally, the transition from the plateau regime to the stretching of the covalent bonds is very smooth so that no shock waves appear that could lead to uncontrolled failure of the protein network. Therefore, coiled-coils may be considered as the elementary building blocks of IFs that enable cells to withstand dramatic loads, large deformation and very high deformation rates. In coiled-coil structures, H-bonds are apparently not used to generate mechanically robust structures, as can be seen by the existence of regime (II) with low unfolding forces (see, for instance Fig. 64). Instead, they provide a means of enabling low-force unfolding and simple refolding into helical structures. Perhaps strength originates from the superhelical structure and from the covalent chemistry present in the protein backbone, and further research could be used to provide a deeper understanding of these aspects.

The hysteresis (Fig. 62) and the strong strain rate dependence (Fig. 61, for instance) provide further evidence that coiled-coils represent strain buffering elements in cells, with the possibility to dissipate great amounts of energy as it undergoes repeated stretching and relaxation cycles. At the same time, at small deformations (smaller than approximately 20%) and small pulling velocities that appear during cell movement, the resistance of the coiled-coil is completely elastic, and thus do not dissipate energy.

The comparison of the simulation and theoretical results with those from simulations and experiments on other coiled-coils, underline the important role of H-bonds as a key in the realization of the specific mechanical properties of coiled-coils. The formation of triplets of H-bonds (as in AHs) is energetically much easier to realize than single bonds with strength equal to three H-bonds. The calculations predict that during pulling velocities applied in experiments, and most likely during those that appears *in vivo*, where three parallel H-bonds break simultaneously, as discussed in Section 4.4 and shown in Fig. 67 [40]. This clearly supports the idea of creating parallel bonds as a means to overcome the intrinsic weakness of H-bonds (this will be discussed further in Section 5 below). The coiled-coil configuration reinforces the protein additionally and makes it much more stable.

5. Beta-structured protein materials

This section focuses on beta-sheet and beta-helical secondary structures; key building blocks of muscle, viral, spider silk and amyloid fibers as well as many other biological materials with intriguing mechanical properties. We review computational and experimental results on the nanomechanics of protein structures that employ beta-sheets, illustrating the key link between this topological configuration and mechanically resistant protein domains found in biology. H-bond rupture mechanisms and rate dependence are given particular emphasis in this discussion as they are central to the protein unfolding problem; and thus control the strength of beta-sheet domains.

Here we briefly review experimental and computational findings on deformation mechanisms and pulling rate and geometry-dependent mechanical properties of beta-sheet structures. This is followed by a summary of recent theoretical and atomistic simulation studies on a simplistic beta-sheet model that aims on elucidating basic key deformation mechanisms and size-dependent strength of these fundamental building blocks. Results from a systematic rate dependence study employing two different pulling geometries on the model system are used here in conjunction with the modified Bell Theory reviewed in Section 4 [42,263] to illustrate the correlation between rupture strength of beta-sheets and the number of hydrogen bonds that participate in the rupture process. We review a fracture mechanics based theoretical framework that considers the free energy competition between entropic elasticity of polypeptide chains and rupture of peptide hydrogen bonds, which has been used to provide an explanation for the intrinsic strength limit of protein domains by providing a physical basis for rupture mechanisms observed in simulation studies. We discuss influence of strand length on mechanical resistance through theoretical and computational approaches, and conclude based on

earlier work that hydrogen bond assemblies in beta-sheets break in clusters of 3–4 HBs and that the strength of these assemblies are maximized when strand lengths are confined to 3–4 residues. We discuss the universal implications of this finding based on proteomics data that give characteristic dimensions of hydrogen bonded assemblies as they appear in beta-sheet structures and other protein motifs.

In the second part of this chapter, we review recent findings on mechanical properties of beta-helical structures; and present results on atomistic and continuum modeling of a model triple-sheet beta-helix structure and illustrate tensile and compressive deformation mechanisms common to beta-helical topologies. We present a deformation map for the length-dependent strength of this motif and discuss possible application of the computational technique to study this structure as it appears in cell-puncture device of the virus bacteriophage T4 (this is a virus that infects bacteria).

5.1. Significance of beta-sheet rich materials: biological, materials science and disease aspects

The beta-sheet protein motif is a very common building block of a variety of biological materials, as it is (along with alpha-helices and tropocollagen molecules) one of the basic protein structures that make up most of all functional biological matter. Many protein materials with mechanical function or extraordinary resistance to mechanical manipulation contain predominantly beta-sheets. Muscle fibers, amyloid filaments as well as commercial and spider silk are well known examples of materials that exhibit high beta-sheet content and interesting mechanical properties.

One of the most intriguing protein materials found in Nature is spider silk, a tough, elastic fiber that can exceed the strength of steel. Dragline silk is composed out of polyalanine regions that form strong beta-sheet nanocrystals that are assembled by weak inter-stand H-bonds (HBs). These tiny crystals are dispersed in a less-ordered, glycine-rich matrix that consists primarily of random coils and helices and hence has lower rigidity [264–266]. This combination of regions with distinct secondary protein structures forms an exceptionally tough fiber that blends extensibility with strength [267–269]. Fig. 81 depicts a schematic that shows the hierarchical scales observed in spider silk.

A simple molecular model was proposed by Termonia [270] for elastic behavior of spider dragline silk. Termonia's molecular model is one of the first mechanical modeling attempts of spider silk based on structural evidence from X-ray diffraction studies. Although this is a simple model, and specifically applicable to dragline silk only, this continuum-type molecular approach was instrumental in illustrating how the size of beta-sheet nanocrystals in spider silk influences mechanical behavior.
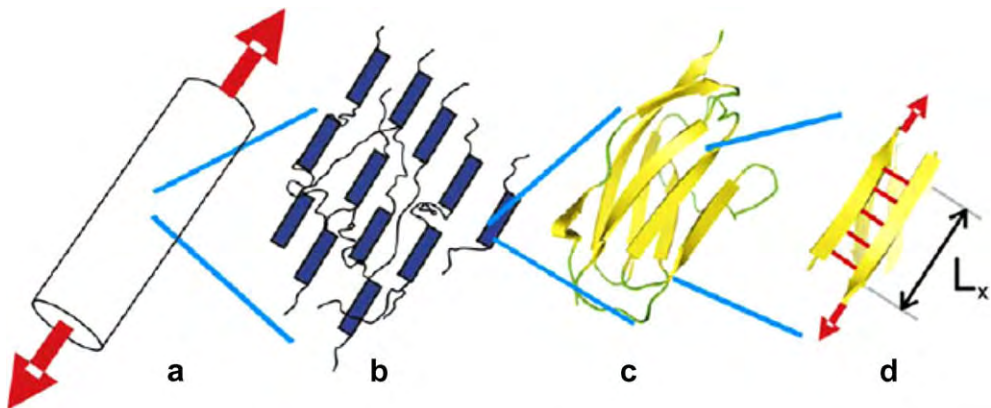
Fig. 81. Schematic views of the hierarchical structure of spider silk and the beta-sheet building blocks. The figure shows a spider silk fiber (subplot (a)), the microstructure of a spider silk fiber (subplot (b)), and a detailed view of beta-sheet crystals (subplot (c)) as well as individual beta-strands that make up beta-sheet crystals (subplot (d)). The blow-up shown in subplot (d) depicts a schematic of the predominant shear loading condition of a single beta-strand with length L_x.

American Chemical Society
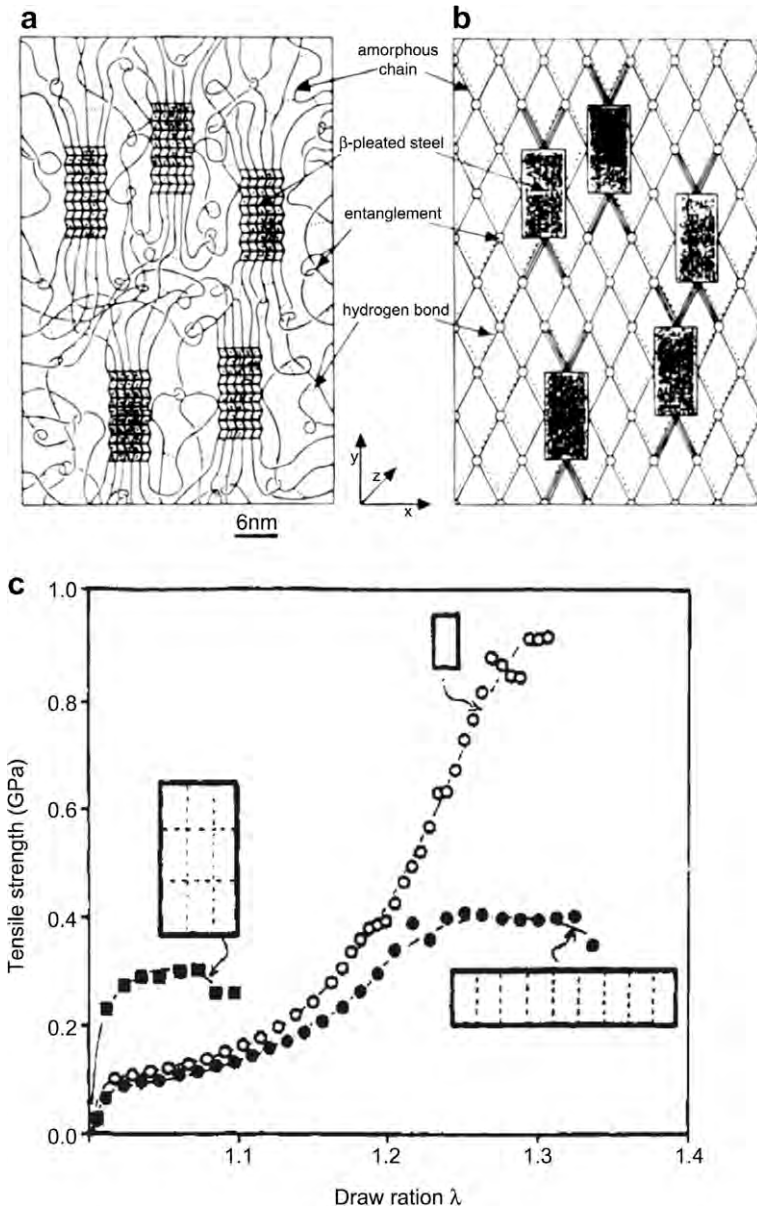
Fig. 82. Simple model proposed by Termonia [270] for elastic behavior of spider silk. Termonia's molecular model is one of the first mechanical modeling attempts of spider silk based on structural evidence from X-ray diffraction studies. Subplots (a) and (b): The complex microstructure of the material is reduced to a simple continuum model with elastic constants for H-bonds and covalent chains. A key finding from this paper is that the size of beta-sheet nanocrystals in spider silk influences elastic behavior and the observed structures in silk are optimal from a mechanical perspective, as shown in Subplot (c). The results show that smaller crystallites lead to an increased tensile strength (curve with open circles), as opposed to the curves with nine times larger crystallites (curves with closed square and closed circles). Reprinted with permission from Ref. [270].

Fig. 82 reviews this molecular model of spider silk and presents some of the resulting stress–strain behavior.

Association of beta-sheet structures in materials with superior mechanical properties is not limited to spider silk proteins. The resistance of fibronectin and immunoglobulin domains of muscle protein titin and extra-cellular matrix protein ubiquitin to unfolding is controlled by large beta-sheet dominated regions, as has been confirmed by atomistic simulation and single-molecule force microscopy studies [271–273].

Recently, interest in beta-structured materials has grown tremendously due to disease [274,275] and materials science aspects of these materials. Beta-sheet structures are very promising for self-assembling peptide applications to create biological nanotubes and nanowires [276]. Such an autocatalytic material growth phenomenon is also linked to pathological conditions such as amyloidosis [274], when amyloid fibers aggregate in tissue as a result of formation of exceptionally stable hydrogen bonded cross-beta-sheet structures. Recent studies suggest that amyloid fibrils may also have a mechanical role in some natural adhesives [277] and may be useful for materials science applications [278]. These examples illustrate the significance of the beta-sheet motif for designer protein-based structural materials.

5.2. Mechanical characterization of beta-sheet protein structures

Here we review results from theoretical and simulation studies on characterization of beta-sheet structures. We first review earlier work on mechanical manipulation of beta-sheet proteins by experimental and simulation techniques. Following this, we present recent MD simulation results on a simplistic beta-sheet model. In conjunction with theoretical approaches, we use this simple model to provide an overview of pulling rate and geometry-dependent mechanical resistance of beta-sheets. The main purpose of these simulations are to illustrate the influence of strand length, or equivalently the number of hydrogen bonds in the assembly to the strength.

We review a theoretical and computational framework that we have recently developed to predict size-dependent strength of these domains. A key observation from this section is that despite the presence of more than a dozen H-bonds in the system – all loaded in uniform shear, H-bond rupture occurs in clusters of 3–4. This phenomenon is explained by a fracture mechanics based theoretical framework which also predicts the asymptotical strength limit of beta-sheets at vanishing rates (vanishing rates refers to quasi-static, near-equilibrium rates as typically investigated by experimental techniques, often on the order of 1E−8 m/s or less). This finding furthermore shows that this strength limit also corresponds to a very prevalent, characteristic length scale at which shear strength of beta-sheets are maximum.

5.2.1. Review of computational and experimental studies on beta-sheet proteins

Natural fibers can match the strength of steel, as in the case of spider silk [279], yet the chemical bonds that give rise to these properties are almost one thousand times weaker than the interatomic metallic bonding. The source of the macroscopic strength remains a mystery: the fact such strong materials can be created using weak HBs is counter-intuitive and has been discussed controversially. For instance, mechanical resistance of proteins has been linked to their orientation [220,280–282]. However, how the topology, the number of parallel strands and their lengths influence the strength remains divisive. Most theoretical concepts that attempt to explain the strength of spider silk and similar materials are based on phenomenological continuum-type theories that are not directly linked to specific atomistic and molecular processes. As a consequence, nano-scale deformation and fracture mechanisms in silk and other beta-sheet dominated protein materials remain unknown.

Steered molecular dynamics (SMD) simulations [273,283] have confirmed that the mechanical resistance of beta-sheet rich proteins derive from hydrogen bonding between beta-strands of beta-sheet structures. Mechanical resistance is highly dependent on pulling geometry; the key beta-sheets must be loaded in shear, parallel to the strand direction [220,280–282] such that parallel hydrogen bond assemblies in between the beta-strands cooperatively resist deformation. This loading geometry also corresponds to physiological conditions in for instance muscle tissue and it justifies the orientation of beta-sheets predominantly along the fiber direction in spider silk [265] (see also Fig. 81).

The shear loading leads to stronger structures since it induces concurrent as opposed to sequential breaking of hydrogen bonds (HBs). Based on this view, longer strands would have higher mechanical

strength, since more HBs contribute to the mechanical resistance. In contrast, it is observed in experimental analyses of protein structures that strand lengths do not exceed several residues (one residue equals roughly one HB), or equivalently a few nanometers in virtually all beta-sheet structures currently known [284]. These short strands are typically arranged between short random coil segments, forming a composite of HB networks and entropy-dominated elasticity. These structural features appear to be a universal phenomenon.

Moreover, this particular geometry is astonishing since other protein structures such as those found in collagenous tissue [3,8,27,28,90–92,131,132,285] or cytoskeleton networks [1,163–177,286] display much larger structural features, reaching 300 nm and more. However, in contrast to beta-sheets, the intermolecular interactions in these protein materials are not primarily based on HBs, but include much stronger electrostatic and covalent bonding. This suggests that the ultra-small features in beta-sheets may be related to the specific properties of the HB networks.

Despite significant advancements in our understanding of the nanomechanics of beta-structures in materials [220,236,272,280–282,287,288], several key fundamental questions remains unanswered: what is the strength limit of H-bond assemblies, and how is it possible that protein materials such as spider silk or amyloids reach strengths that exceed those of steel, despite the weakness of H-bond interactions? Is there a physical explanation for the characteristic structure of beta-sheet protein domains, typically displaying ultra-short strand lengths with less than five residues, embedded in random coil protein domains? Solving these key problems based on studies on complex protein structures is extremely difficult, due to the difficulty of isolating rupture events in the unfolding process. In the next section, we present atomistic results based simple beta-sheet model system that represents a mechanical clamp as found in beta-structured proteins. The model captures intricate key details that govern force-induced unfolding of such systems, in particular entropic elasticity of polypeptide chains at low force levels and the energetics of hydrogen bonding. These two aspects of the protein unfolding problem had been commonly addressed separately in experimental and computational studies; whereas the simplistic theoretical framework that will be reviewed next illustrates that coupling between the elasticity of protein backbone and H-bond energy release due to rupture governs the unfolding of protein domains.

5.2.2. Molecular structure of beta-sheets and simplified atomistic modeling approach

Here we briefly review molecular structure of beta-sheets and the model system studied in this section. We also provide a brief explanation of the atomistic modeling approach. The beta-sheet, along with the alpha-helix, is one of the two most common types of secondary structures that make up protein materials [289]. Typically, a sheet consists of several parallel (or anti-parallel) peptide chain segments with multiple interstrand H-bonds that stabilize the structure. Regular structure of the sheet is extended onto less regular loop sections and random coils that exhibit entropic elasticity under low external force levels. Unlike alpha-helices, which consist generally of a single polypeptide chain folded onto itself in helix form by very localized H-bond interactions, beta-sheets are formed not only by intramolecular attraction between two regions of the same polypeptide chain but also by intermolecular forces between individual polypeptides, serving as Nature's glue for assembling very complex nanostructures.

The complex nature of full protein domains makes it extremely difficult to find generic, quantitative predictions of the strength of beta-sheet structures. Here we attempt to overcome this challenge by selecting a simple structural model and well-defined loading conditions to provide crucial physical insight into the common observations made by both computational and experimental approaches in the earlier works on beta-structures.

As a simple model representation of larger beta-sheet rich protein structures (Fig. 83a), we consider a small, full atomistic protein domain embedded in explicit solvent (TIP3 water), as shown in Fig. 83b. The model contains three beta-strands with inter-strand H-bonds, representing the beta-sheet protein motif that forms larger protein structures as shown in Fig. 83a. The strands in our model system have free chains at their extremities, representing already unfolded domains of a macromolecule.

All-atom MD simulations were performed using segments from the crystal structure of a beta-barrel protein obtained from Protein Data Bank [86] (PDB ID 1i78). The model three-strand beta-sheet
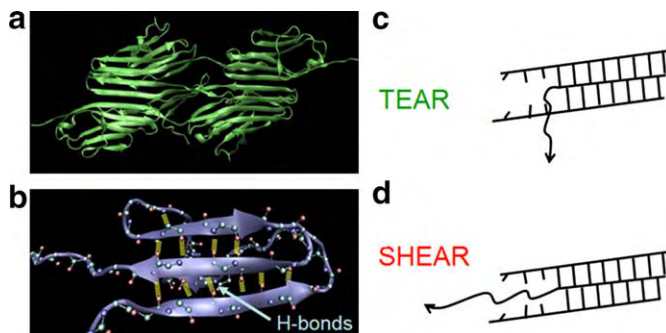
Fig. 83. Molecular structure and loading conditions, used for molecular dynamics simulations [299,300]. Beta-sandwich structures are a characteristic feature of mechanical proteins, which employ networks of parallel H-bonds that work cooperatively in shear, like the Z1Z2–telethonin complex in titin (shown in subplot (a)) [220,280–282]. Inter-strand hydrogen bonds (thick yellow lines) act as mechanical clamps that resist unfolding, as shown in subplot (b) for the I27 domain of titin. We consider a three-strand system with free chains at the ends, as shown in subplot (b) as a model to study the strength of beta-strands. We study two deformation modes, in-plane shear and out-of-plane shear, which we refer to as SHEAR and TEAR modes, respectively (subplots (c) and (d)). In the TEAR mode, the middle strand is pulled out in the direction orthogonal to the strand direction. In the SHEAR mode, the middle strand is pulled in the direction of the beta-strand, leading to uniform loading of H-bonds.

system studied corresponds to residues 134–144 (lower strand), 161–176 (middle strand and free chain) and 203–211 (upper strand) in the PDB file.

In MD simulations, energy minimization is carried out for 20,000 steps, which allows for a favorable conformation to be achieved. This is followed by an equilibration procedure, during which the molecular assembly is heated up to 300 K with a rate of 25 K every 25 steps. An NVT ensemble is employed to hold the temperature constant at the final temperature of 300 K. Using a time step of 1 fs, each molecular assembly is subjected to equilibration for 2 ns. The system is equilibrated while all C_α atoms on top and bottom strands were fixed. This constitutes the boundary conditions adopted for uniformly shearing two strands with minimal translational and rotational motion of the molecular assembly.

To simulate forced rupture of H-bonds between the strands, the steered molecular dynamics procedure (SMD) with a constant velocity pulling scheme is used, in line with the general description in Section 2. An SMD spring constant of 10 kcal/mol/Å^2 is used. We employed pulling rates ranging from 20 to 0.1 m/s in our studies to estimate the energy barriers and location of the transition state for each deformation regime. These pulling velocities are quite reasonable in light of previous studies [273,283,290] on beta-sheets. To subject the assembly to shear, the terminal C_α atom of the middle strand free chain is pulled, while the other strands are subject to fixed boundary conditions described in the equilibration process.

Bond dynamics analysis is done based on NAMD results and visualization schemes using VMD program [291]. We define that the bonds investigated are ruptured at an extension of 4 Å. We use a large sampling of trajectories (10 bond extension data points per ps) to estimate rupture dynamics accurately in each deformation mode.

Maximum force for rupture is obtained from MD simulations. We average the simulation results such that force–displacement plots employ 1 bin/Å extension. In tear mode, since the bond rupture events occur sequentially and a series of peaks are observed, we take the mean force over all rupture events as the representative maximum force as it is statistically more significant than the absolute maximum for this deformation mechanism.

5.2.3. MD simulations in tear and shear modes for a simple beta-sheet model

As a simple model representation of larger beta-sheet rich protein structures (Fig. 83a), we consider a small, full atomistic protein domain embedded in explicit water, as shown in Fig. 83b. The model contains three beta-strands with inter-strand H-bonds, representing the beta-sheet protein motif that

forms larger protein structures as shown in Fig. 83a. The strands in our model system have free chains at their extremities, representing unraveled protein domains. We model the behavior of this protein assembly by large-scale molecular dynamics (see Chapter 2 for details regarding the numerical procedure). In the analyses of the mechanical response of this system, we pull the center strand in different directions of loading, while the outermost strands remain fixed.

The goal of the computational experiments is to probe the strength and rupture mechanism of the model system in two extreme modes of deformation. In the out-of-plane shear mode (TEAR, Fig. 83c), the middle strand is pulled perpendicular to the plane of the sheet. The computational experiment in the TEAR mode is designed so that the H-bonds break sequentially, one-by-one. This case enables us to assess the rupture behavior and energy landscape of individual H-bonds. We consider a second case, the in-plane shear mode (SHEAR, Fig. 83d). Here, the middle strand is pulled out in the axial direction of the three strands. The SHEAR case is designed to asses the rupture behavior and the energy landscape of a large number of H-bonds under uniform shear loading. For both deformation modes, we carry out a systematic study of the rupture force F against pulling rate v, and plot the results on a logarithmic scale over three orders of magnitude (Fig. 84a). The curves obtained for each deformation
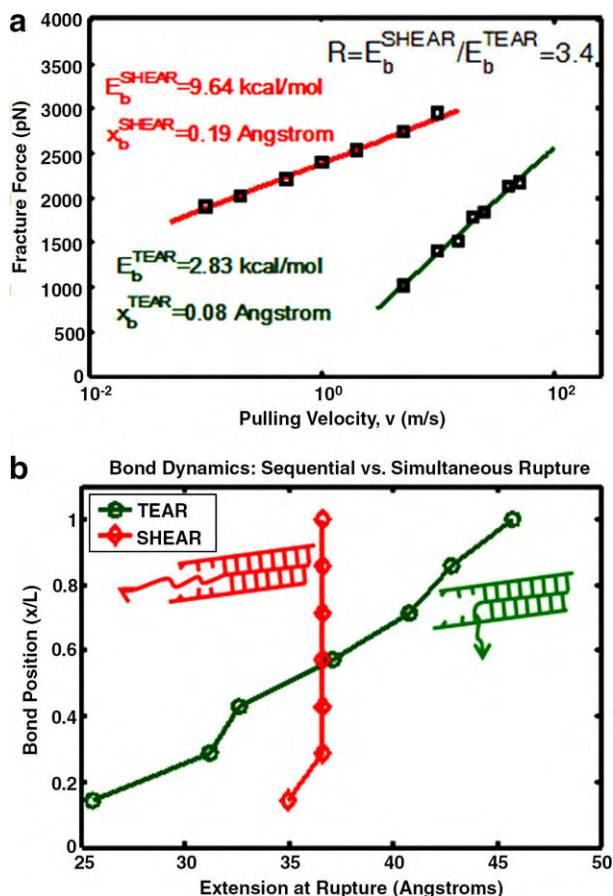
Fig. 84. Unfolding force vs. deformation speed, TEAR vs. SHEAR deformation modes, and associated H-bond rupture mechanism (results of full atomistic molecular dynamics simulations in explicit water solvent) [300]. Subplot (a): Logarithmic relationship between the pulling speed and rupture force, for both deformation modes. The parameters of the linear fit for each regime are used in conjunction with the modified Bell Model to determine the energy landscape for each deformation mode. Subplot (b) displays the H-bond rupture dynamics, which reveals that the localization of strain in the TEAR mode leads to sequential rupture of H-bonds. In contrast, several H-bonds rupture concurrently in the SHEAR mode.

mode fall on two distinct lines in the F versus $\ln(v)$ plane, which indicates according to modified Bell Model (see Section 4.4.2) [263] that the unfolding event can be characterized by a two state system with a discrete and unique energy barrier, E_b and location of the transition state, x_b [42].

The energy barrier in the TEAR deformation mode is $E_b^{TEAR} = 2.83$ kcal/mol. An analysis of the bond rupture dynamics reveals that the H-bonds break sequentially (Fig. 84b), indicating that the energy barrier to break a single H-bond in this protein is $E_{HB}^0 = 2.83$ kcal/mol. The energy barrier in the SHEAR deformation mode is $E_b^{SHEAR} = 9.64$ kcal/mol, significantly larger than in the TEAR mode. The analysis of the rupture mechanism in the SHEAR mode reveals that H-bonds break in clusters of multiple bonds (Fig. 84b).

The comparison of the energy barriers in the TEAR and SHEAR mode enables us to calculate how many H-bonds participate in the rupture mechanism in SHEAR loading. This is possible because the energy barrier in the TEAR mode corresponds to rupture of individual H-bonds. The number of H-bonds that break concertedly in the SHEAR mode is given by the ratio of the E_b values for SHEAR and TEAR; it is $E_b^{SHEAR}/E_b^{TEAR} \approx 3.4$. This suggests that H-bonds break in clusters of 3–4, but not more, even when the loading conditions (uniform shear) allow them to respond cooperatively.

Key finding from this study is that despite the presence of more than a dozen H-bonds in the system – all loaded in uniform shear, H-bond rupture occurs in clusters of 3–4. In terms of the energy barrier, it appears to be limited to a finite value, despite the possibility for a manifold increase of the resulting energy barrier due to the large number of H-bonds. We explain this phenomenon within the theoretical framework of the asymptotical strength limit prediction of hydrogen bond assemblies, which is treated in the next section.

5.2.4. Asymptotical strength of hydrogen bond assemblies in beta-sheets: experimental observations

Here we focus on the strength of uniform loaded multiple parallel hydrogen bonds, as they appear in beta-sheets and other protein structures. We review theoretical and experimental studies on multiple biological bonds. Non-covalent bonds such as H-bonds are indispensable to biological function as they play a key role in cell adhesion and motility, formation and stability of protein structures and nucleic acids, and receptor–ligand interactions. The weak nature of these bonds makes it a necessity for them to work cooperatively and function in assemblies. Strength of multiple parallel bonds has received notable interest and has been studied both theoretically, using statistical mechanics approaches as well as experimentally with the atomic force microscope (AFM) [200,292–297].

This fundamental question of 'strength' is also linked to the protein unfolding problem where the rupture of parallel inter-strand hydrogen bonds controls the mechanical unfolding pathway. Atomistic simulation [273,283] and single-molecule force microscopy studies [271,272] have shown that beta-sheet rich proteins exhibit higher rupture forces, since they employ parallel strands with numerous hydrogen bonds that act as mechanical clamps under shear loading [220,280–282]. In both experiment and simulation of mechanically resistant proteins, maximum force peak observed is linked to an individual event corresponding to the unraveling of a single beta-sheet in the protein. The dependence on pulling velocity has also been widely discussed in the literature [234,263] and explains the discrepancy of force values between MD simulations and experimental methods, the former employing fast and the latter near equilibrium pulling rates. It is astonishing to see that regardless of the variation in topology and size of structures studied, almost all so called 'mechanical' proteins (e.g. fibronectin and titin domains) examined so far exhibit a rupture force of a few hundred picoNewtons (pN) at experimental pulling rates [234,298].

Experimental and simulation results are presented in Fig. 85 where rupture force is plotted against pulling velocity on $\log(v)$ scale, as reported in [299]. The asymptotical limit at vanishingly slow (quasi-static) rates can be inferred from the overall behavior as well as from the power law fit to data based on unfolding force data corresponding to the Ig27 domain in titin (data obtained from [234,298]). This observation suggests that without the presence of salt bridges or covalent links, the strength of individual protein domains asymptotically approaches a limiting strength. Thus far, no theoretical basis or prediction has been proposed for such an intrinsic strength limit; most earlier analysis have been focused on the rate dependent behavior in order to explain the increase in unfolding forces with increasing pulling speed (Fig. 85). We have recently provided an explanation to this phenomenon by applying basic concepts from fracture mechanics to the protein unfolding problem. We focus on a simple model system that consists of a single beta-sheet strand stabilized by H-bonds, which has been identified as
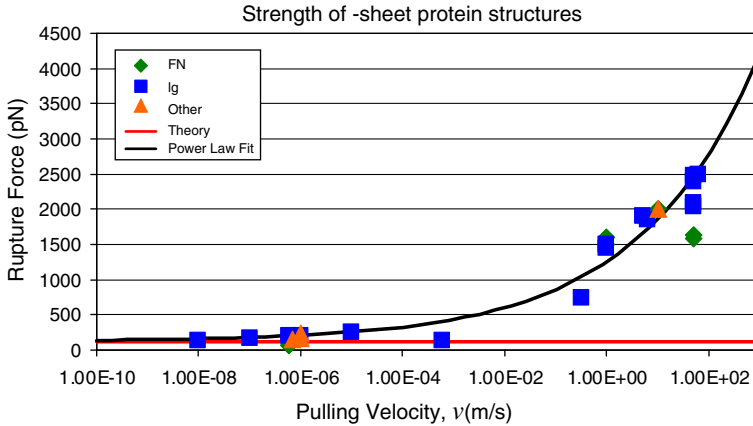
Fig. 85. Summary of earlier studies on rupture force of beta-domains as a function of pulling rate as reported in [299]. We summarize findings (adapted from [298]) of the strength of beta-sheets rich proteins (focus on fibronectin, immunoglobulin domains in ECM and titin). The overall behavior suggests that the rupture force asymptotically approaches a limiting value for vanishing pulling rates (continuous line is a power law fit to data for Ig27 [234]).

the mechanical clamp that governs strength during unraveling of the protein [220,280–282] (see Fig. 83). We describe the theoretical framework in the following section.

5.2.5. Strength prediction based on Griffith–Irwin energy balance concept

Here we review a theoretical model based on the structure presented in Fig. 83, consisting of a single strand with multiple peptide hydrogen bonds and a free chain representing unraveled protein domains [300]. The idealized model is shown in Fig. 86a, where a polypeptide chain of arbitrary length is attached to a substrate and strained in one end with force F. This setup serves as the model for which we develop our theory. As a maximum strength value that is not hindered by too few hydrogen bonds is sought after, we consider an infinitely long beta-strand (that is, we assume that the length of the strand, $L_x \rightarrow \infty$).

We adopt Griffith's theory of fracture mechanics [301] to predict the critical rupture force required to break the bonds in our theoretical model system (see Fig. 86a and b). The concept is based on the idea that onset of fracture is characterized by the condition that the change in potential energy W_P of a system due to extension of a crack balances the energy necessary to create new surfaces.

In the Griffith theory, the negative of the expression for the potential energy change with respect to a crack advance of one unit distance δa is called the energy release rate or crack driving force:

$$G = -\frac{\delta W_P}{\delta a}. \tag{66}$$

The total potential energy in the system can be expressed as the sum of the elastic energy stored in the material, denoted by Φ, and the work done by the external load, denoted by W_F:

$$W_P = \Phi + W_F. \tag{67}$$

The onset of fracture is characterized by the condition that the potential energy released due to a unit crack advance δa must equal the energy required per unit length to create a new fracture surface. The idea that fracture surface energy, γ_s must equal the energy release rate at the onset of fracture leads to

$$G = -\frac{\delta W_P}{\delta a} = \gamma_s. \tag{68}$$

Despite being originally proposed as a continuum fracture theory, the energy balance notion relies on purely thermodynamical concepts and has been used in the past to describe fracture phenomena at various length scales, including atomistic phenomena in collagenous tissue (see Section 3 and [285]).
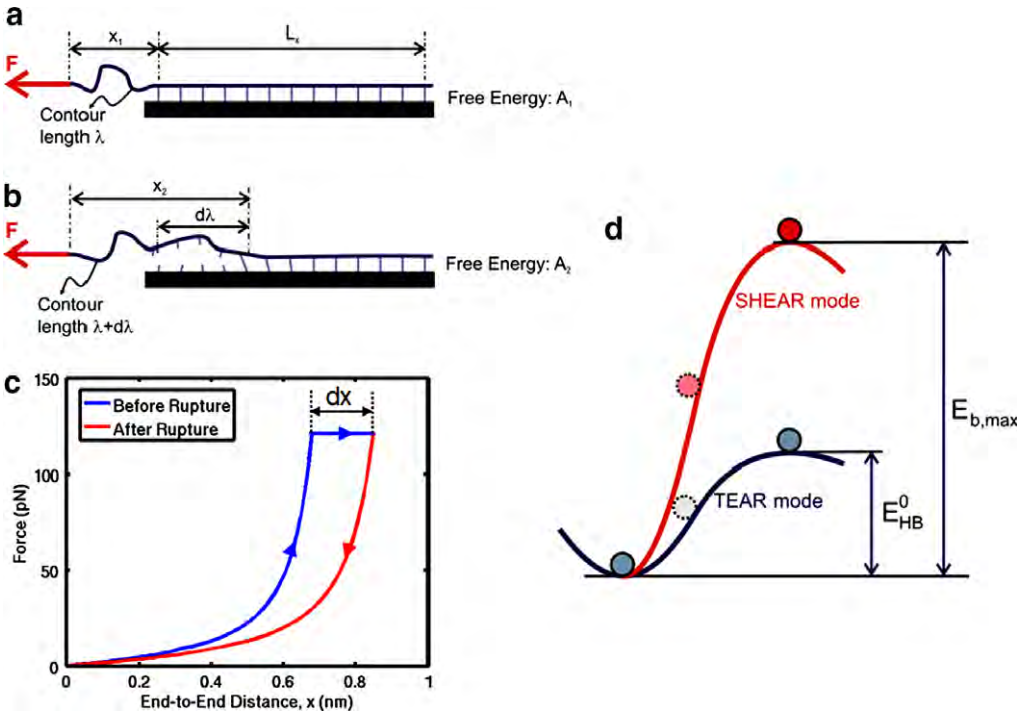
Fig. 86. Schematics to explain the main concepts of the WLC-based fracture theory, as well as the implications on the energy landscape [300]. Subplot (a): We consider a single double-strand system as a polypeptide chain of infinite length ($L_x \to \infty$) with a free end of length a. The polypeptide chain is stabilized by an array of parallel H-bonds, and strained due to the external force F. At the onset of rupture (subplot (b)), several H-bonds break, and the contour length increases due to the detachment of a piece of the chain of length $d\lambda$. Subplot (c): The change in the contour length and the end-to-end distance at the onset of rupture yields two distinct WLC curves, relating to the states before and after rupture; the area enclosed between the WLC force extension curves is equivalent to the change in free energy before and after rupture. This dissipated (released) energy must equal the adhesion energy per unit length at the onset of rupture. Subplot (d) shows a schematic of how the energy barrier changes from rupture of a single H-bond up until the limiting energy barrier $E_{b,max}$, when 3–4 H-bonds rupture simultaneously. This energy barrier is the highest energy barrier of an assembly of H-bonds under uniform loading.

The model system can be broken down into three components, the external force, the peptide hydrogen bonds and the elasticity of the free chain that represents the backbone of unfolded protein domains. As we probe the strength at vanishing pulling rates (quasi-static deformation), we assume that the system is in equilibrium and the force is constant over infinitesimal time-scales. We characterize the strength of hydrogen bonds by their dissociation energy, E_b.

To understand the fracture behavior of polypeptides bonded by HBs, we extend the theoretical framework laid out by Griffith and Irwin by taking into consideration that at the force levels of interest, the elasticity of the protein backbone is primarily due to entropic rather than energetic effects. The Marko–Siggia worm-like chain (WLC) model [81] (see Eq. (28)) is one of the most widely used expressions to predict the entropic elasticity of polypeptide chains, and has been adopted widely as the elastic description of the backbone (a wider range of earlier studies provide substantial evidence that this model is an excellent model for the behavior of individual, unconstrained polypeptide chains [236,288,302–304]).

While the Griffith theory considers the change in internal energy, U of the system, in our case we seek an expression of how the free energy – rather than the internal energy, changes as a function of crack extension. We emphasize that this generalization from the internal energy to the free energy does not change the concept behind the Griffith model and thus we can immediately apply the Griffith model without any further modifications.

The approach can be summarized as follows: (i) we calculate how much free energy is released per unit polypeptide rupture distance (free energy release rate, it increases with larger applied force); (ii) we determine the critical free energy release rate so that it is equal to the cohesion energy of a HB system, and then (iii) finally, we calculate the critical force required to initiate rupture based on this critical free energy release rate.

The integration of the force–extension equation of the WLC model (see Eq. (28), Section 3.3.5) yields the free energy $A = U - TS$ at a specific deformation state. This is because the WLC equation describes forces that arise due to changes in the entropy of the system and is derived by calculating the change of free energy as a function of the deformation variable x. The potential free energy stored in the molecule of end-to-end distance $\hat{x}$ and contour length $\hat{\lambda}$, subject to force F is given by the integral of the force F, integrated from the reference point taken at zero to the end-to-end distance $\hat{x}$ (this is a general expression valid for any contour length $\hat{\lambda}$ and any end-to-end distance $\hat{x}$):

$$A(\hat{x}, \hat{\lambda}) = \int_0^{\hat{x}} F(\tilde{x})\mathrm{d}\tilde{x} = \frac{k_\mathrm{B}T}{\xi_\mathrm{p}} \left[\frac{1}{4}\left(1 - \frac{\tilde{x}}{\hat{\lambda}}\right)^{-1}\hat{\lambda} - \frac{\tilde{x}}{4} + \frac{\tilde{x}^2}{2\hat{\lambda}} \right]_0^{\hat{x}}. \tag{69}$$

Before the onset of rupture, the molecule has a free end-to-end distance of length x_1 and an initial contour length of λ_1 and is strained by an external force, F which remains constant at the onset of rupture. When hydrogen bonds rupture the end-to-end distance as well as the contour length both increase. We now assume that a piece of the molecule of length $\delta\lambda$ detaches at the moment of rupture. The contour length then changes from λ_1 (before rupture) to $\lambda_2 = \lambda_1 + \delta\lambda$ (after rupture). The change of the end-to-end distance from x_1 (before rupture) to x_2 (after rupture) can be calculated by considering that the force before and after rupture is identical. Before fracture,

$$F_1 = \frac{k_\mathrm{B}T}{4\xi_\mathrm{p}}\left[\left(1 - \frac{x_1}{\lambda_1}\right)^{-2} + 4\frac{x_1}{\lambda_1} - 1\right], \tag{70}$$

and after rupture,

$$F_2 = \frac{k_\mathrm{B}T}{4\xi_\mathrm{p}}\left[\left(1 - \frac{x_2}{\lambda_1 + \delta\lambda}\right)^{-2} + 4\frac{x_2}{\lambda_1 + \delta\lambda} - 1\right], \tag{71}$$

Assuming that the system is equilibrated at all times, the force remains constant before and after fracture ($F = F_1 = F_2$). Based on this assumption, we equate Eqs. (70) and (71) and obtain an expression for end-to-end distance after rupture, x_2 as a function of the initial end-to-end distance, x_1:

$$x_2 = \frac{\lambda_1 + \delta\lambda}{\lambda_1}x_1. \tag{72}$$

The change in end-to-end distance during rupture is

$$\delta x = \left(\frac{\lambda_1 + \delta\lambda}{\lambda} - 1\right)x_1. \tag{73}$$

Knowledge of these parameters now enables us now to calculate the change in free energy ($\delta W_\mathrm{P} = \delta\Phi + \delta W_\mathrm{F}$) due to rupture of a piece of length $\delta\lambda$. Recall that the quantity $G = -\delta W_\mathrm{P}/\delta\lambda$ is the energy release rate. Since

$$\delta\Phi = A_2 - A_1, \tag{74}$$

and

$$\delta W_\mathrm{F} = -F\delta x, \tag{75}$$

the net change in free energy is given by

$$\delta W_\mathrm{P} = A_2 - A_1 - F\delta x. \tag{76}$$

From the expression for the free energy state given in Eq. (69), free energies before and after onset of fracture are given as

$$A_1(x_1, \lambda_1) = \frac{k_B T}{\xi_p} \left[\frac{1}{4} \left(\frac{\lambda_1^2}{\lambda_1 - x_1} \right) - \frac{x_1}{4} + \frac{x_1^2}{2\lambda_1} - \frac{\lambda_1}{4} \right], \tag{77}$$

before rupture, and after rupture,

$$A_2(x_2, \lambda_2) = \frac{k_B T}{\xi_p} \left[\frac{1}{4} \left(\frac{\lambda_2^2}{\lambda_2 - x_2} \right) - \frac{x_2}{4} + \frac{x_2^2}{2\lambda_2} - \frac{\lambda_2}{4} \right]. \tag{78}$$

The variables x_2 and λ_2 can both be expressed in terms of x_1 and λ_1 (that is, $x_2 = (\lambda_1 + d\lambda)\lambda_1^{-1}x_1$ and $\lambda_2 = \lambda_1 + d\lambda$ as shown above). Therefore,

$$A_2(x_1, \lambda_1 + d\lambda) = -\frac{k_B T}{\xi_p} \left[\frac{1}{4} \left(\frac{(\lambda_1 + d\lambda)^2}{\lambda_1 + d\lambda - (\lambda_1 + d\lambda)\lambda_1^{-1}x_1} \right) - \frac{(\lambda_1 + d\lambda)\lambda_1^{-1}x_1}{4} + \frac{((\lambda_1 + d\lambda)\lambda_1^{-1}x_1)^2}{2(\lambda_1 + d\lambda)} - \frac{\lambda_1 + d\lambda}{4} \right]. \tag{79}$$

To express the free energies A_1 and A_2 in a simpler form, we introduce the variable α given as

$$\alpha = \frac{x_1}{\lambda_1}. \tag{80}$$

The physical meaning of α is the ratio of end-to-end length of the free chain to its contour length, and is defined in the range from 0 to 1 in context of the WLC theory. Further simplification and expressing Eqs. (78) and (79) in terms of α leads to

$$A_1 = \frac{k_B T}{4\xi_p} (\lambda_1)[(1 - \alpha)^{-1} - \alpha + 2\alpha^2 - 1], \tag{81}$$

and

$$A_2 = \frac{k_B T}{4\xi_p} (\lambda_1 + d\lambda)[(1 - \alpha)^{-1} - \alpha + 2\alpha^2 - 1]. \tag{82}$$

The energy release rate expressed in differentials is

$$G(\lambda_1, x_1) = -\frac{dW_P}{d\lambda}, \tag{83}$$

or

$$G = (A_1 - A_2 + F \, dx)/dx \tag{84}$$

where $dx = ((\lambda + d\lambda)/\lambda - 1)x_1$, or equivalently,

$$dx = \frac{d\lambda}{\lambda_1} x_1. \tag{85}$$

The force F is given as

$$F(\alpha) = \frac{k_B T}{4\xi_p} [(1 - \alpha)^{-2} + 4\alpha - 1]. \tag{86}$$

Substituting all terms into Eq. (84) and further simplification eliminates differential terms and yields

$$G(\alpha) = \frac{k_B T}{4\xi_p} [\alpha(1 - \alpha)^{-2} - (1 - \alpha)^{-1} + 2\alpha^2 + 1]. \tag{87}$$

The value of G depends only on the ratio α but not on the individual values of x_1 and λ_1. This equation can immediately be used to calculate a critical condition for initiation of failure by assuming $G = \gamma_s$ and obtaining the value α_{cr} that satisfies this condition. The parameter γ_s is the 1D analogy to the fracture surface energy, and is defined as

$$\gamma_s = \frac{E_{HB}^0}{L_{x,0}}, \tag{88}$$

where E_{HB}^0 is the energy released by a single bond and $L_{x,0}$ is the average distance between two bonds. Once the critical value, α_{cr} is solved for, the rupture force can be obtained as

$$F_{break}(\alpha_{cr}) = \frac{k_B T}{4\xi_p}[(1 - \alpha_{cr})^{-2} + 4\alpha_{cr} - 1]. \tag{89}$$

Fig. 86c displays the process of loading the structure, the rupture event and the return to the initial configuration. The area between the two curves is the dissipated free energy that is employed in the breaking of the hydrogen bonds.

The sole input parameters in this model are the persistence length ξ_p of the polypeptide chain, as well as the dissociation energy of a H-bond, E_{HB}^0 (γ_s is calculated directly from E_{HB}^0). Both parameters can be determined reliably from either experiment or atomistic simulation. E_{HB}^0 for peptide H-bonds is known to generally lie in the range of 2–9 kcal/mol, whereas the persistence length has been found to be approximately 0.4 nm for many different polypeptides and also for DNA strands [236,262,265,288,304].

5.2.6. Numerical estimate for strength

To obtain a numerical prediction for N_{cr}, we choose $T = 300$ K and $\gamma_s = 0.94$ kcal/mol/Å (for $E_{HB}^0 = 2.83$ kcal/mol from the TEAR mode MD simulation), and $\xi_p = 0.4$ nm from previous experimental studies [236,262,288]. These values are based on earlier experimental observations [236,262,288]. We solve this equation numerically to find the parameter $\alpha = x/\lambda$, the ratio of the end-to-end distance to contour length. We find that $\alpha = 0.85$ for the particular value of γ_s selected. This means that rupture occurs when the free chain is stretched to 85% of its contour length. The predicted asymptotic force limit is 127 pN. We emphasize that this value does not depend on the number of HBs that participate; rather, it is an intrinsic shear strength limitation.

This is the most important result of this analysis: the force F_{break} is the absolute maximum strength that can be attained by any parallel bond arrangement, at near equilibrium pulling rates. The rupture force F_{break} is independent of the contour length λ_1, since the solution to Eq. (67) only depends on α.

The significance of this result is evident by the comparison with experimental and computational results of the unfolding force as a function of pulling speed, as shown in Fig. 85. It is apparent that the pulling speed dependence suggests that a limiting force of approximately 100–300 pN is reached for vanishing pulling rates. This observation has been noted previously in the literature; however, it has never been explained. This work closes this gap and provides – to the best of our knowledge – the first rigorous explanation.

These findings indicate that the key to understand this strength limitation is the interplay between entropic elasticity and the energy of HB adhesion; since both components are 'material intrinsic properties' of protein structures, the force limit represents a universally valid value.

5.2.7. Scaling relationship and size-dependence: size of fracture process zone

MD simulations on the simplistic system suggest that hydrogen bonds exhibit limited cooperation under external force. The equilibrium theory presented in the previous section complements this idea by showing that the rupture force of a beta-sheet with numerous hydrogen bonds is finite. Theoretical derivation illustrates an extreme case of infinitely many hydrogen bonds loaded extremely slowly and uniformly; this is essentially a scenario that allows maximum possible cooperation of the H-bonds.

The finite strength limit indicates that indeed a limited number of hydrogen bonds actually participate in the rupture process and contribute to strength. Calculation of the critical number of hydrogen bonds that break simultaneously from the strength prediction would be of great interest as this would yield a critical length scale for beta-strands beyond which further mechanical stability can't be achieved by additional bonds. Here we attempt to link the strength limit to the number of bonds broken at the initiation of the rupture by using the conventional, phenomenological Bell Model (see Section 4.4.1) for providing the link between external rupture force and energy barrier to unfolding. This idea is similar to the concept of fracture process zone estimation in cracks propagating in a specimen.

We recall that the rupture force of the protein assembly is given as

$$F_{\text{break}} = \frac{k_B T}{4\xi_p}[(1 - \alpha_{\text{cr}})^{-2} + 4\alpha_{\text{cr}} - 1].$$

(90)

Now we calculate the number of H-bonds that participate in a unit fracture event. We consider the process at the rupture front, at the moment when an *unknown* number of N_{cr} H-bonds break. We model the strength of this cluster of H-bonds by a statistical theory based on Bell's original model [42] (reviewed in Section 4.4.1 as the conventional, phenomenological Bell Model). The key parameter is the energy barrier: the energy barrier rises N_{cr}-fold when N_{cr} hydrogen bonds break simultaneously. Thus, the rupture force of this system is

$$F_{\text{break}}^{\text{local}}(N_{\text{cr}}) = \frac{1}{x_B}\left[k_B T \ln\left(\frac{1}{\omega\tau}\right) + E_{\text{HB}}^0 N_{\text{cr}} \right],$$

(91)

where $\omega = 1 \times 10^{13}$ s^{-1} (natural frequency of bond vibration), τ is the characteristic time-scale for H-bond dissociation, and x_B is the applied pulling distance at the moment of bond rupture (this is a slightly modified version of the equation presented in Eq. (43)).

The parameter $x_B = 4$ Å, corresponding to the distance required to pull the strand to initiate permanent rupture of bonds (based on geometric considerations), and the characteristic time-scale $\tau = 20$ ps (τ corresponds to the time-scale at which H-bond rupture occurs; this characteristic time-scale of H-bond rupture is chosen according to experimental [262] and computational [263] results, and characterizes the dynamics of H-bond rupture, which is much faster than the applied loading in the asymptotic limit).

At the moment of rupture, the local rupture force given by Eq. (91) must equal the critical rupture load F_{break} given by Eq. (90). Setting $F_{\text{break}}^{\text{local}}(N_{\text{cr}}) = F_{\text{break}}$ thus enables us to calculate the number of H-bonds that break simultaneously in the unit rupture event,

$$N_{\text{cr}} = \frac{k_B T}{E_{\text{HB}}^0}\left[\frac{x_b}{4\xi_p}[(1 - \alpha_{\text{cr}})^{-2} + 4\alpha_{\text{cr}} - 1] - \ln\left(\frac{1}{\omega\tau}\right) \right].$$

(92)

To obtain a numerical prediction for N_{cr}, we choose $T = 300$ K and $\gamma_s = 0.94$ kcal/mol/Å (for $E_{\text{HB}}^0 = 2.83$ kcal/mol from the TEAR mode MD simulation), and $\xi_p = 0.4$ nm from previous experimental studies [236,262,288]. The prediction of the critical number of H-bonds that break concurrently is $N_{\text{cr}} = 3.1$. The corresponding predicted maximum energy barrier (see Fig. 86d) for this case is $E_{\text{b,max}} = N_{\text{cr}} E_{\text{HB}}^0 = 8.77$ kcal/mol, with an asymptotic strength limit (at vanishing pulling rates) of 127 pN.

A schematic summary of the entire derivation is shown in Fig. 87. Further, Table 12 provides an overview over all variables used in this section.

This is the most important result of this derivation: the value $E_{\text{b,max}} = N_{\text{cr}} E_{\text{HB}}^0$ represents the highest energy barrier that can be achieved by a uniformly loaded H-bond assembly, providing an upper limit for its strength. Moreover, the prediction for the energy barrier is very close to what is observed in the SHEAR mode ($E_b^{\text{SHEAR}} = 9.64$ kcal/mol $\approx E_{\text{b,max}} = 8.77$ kcal/mol) in MD simulations. There exists an intrinsic upper limit of 3–4 H-bonds that can break concurrently, illustrating that the conventional assumption of uniform shear loading of H-bonds in beta-strand fails once the number of H-bonds in an assembly exceeds this critical value.

The shear strength of beta-sheets can be calculated as a function of the strand length, in the asymptotic limit (the analysis of the geometry of beta-sheets reveals that 1 H-bond $\approx$ 1 residue is a good approximation). We assume that the maximum rupture force of the sheet is given by as

$$F_{\text{ult}}(N) = \max(F_{\text{break}}^{\text{local}}(N), F_{\text{break}}),$$

(93)

where N is the number of H-bonds in the assembly.

Eq. (93) expresses the fact that any beta-sheet that employs less than the critical number of H-bonds N_{cr}, features a rupture force that is described by Eq. (91), that is, the strength of increases with a increasing number of participating H-bonds. If $N \geqslant N_{\text{cr}}$, the rupture force equals F_{break}, the asymptotic strength limit of an assembly of H-bonds.
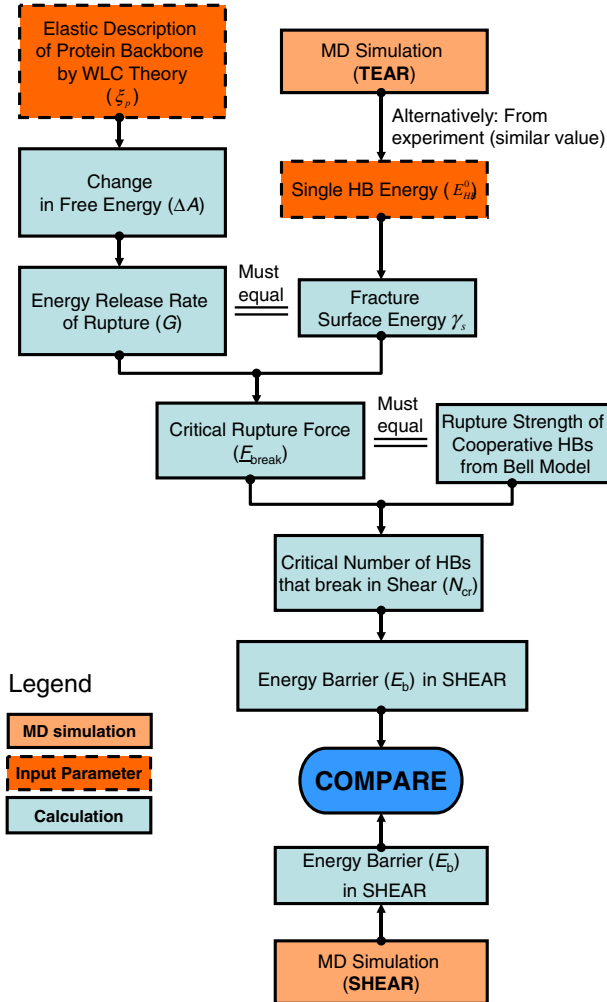
Fig. 87. Flowchart that summarizes the theoretical derivation. This figure shows a flowchart that explains each step in the theoretical derivation. The energy balance concept is used to derive energy release rate of fracture of a polypeptide chain attached to a surface. We estimate strength of an individual hydrogen bond from our MD simulations in TEAR mode. This value is used to calculate the fracture surface energy per unit length, which must equal energy release rate for fracture to be initiated. This equality can be used in conjunction with WLC expression for force to calculate the critical rupture force. This rupture force is linked to a critical number of hydrogen bonds that break simultaneously in shear. The increase in the energy barrier due to simultaneous rupture of bonds is estimated from the critical number, and compared to our SHEAR simulation results. Results from both theory and simulation show close agreement.

The shear strength equals the force required to break all of the bonds in the strand, $F_{ult}(N)$, divided by the contact area between the strands. We assume that the thickness T, of a beta-strand equals 5 Å, and its length is given by $N \times L_{x,0}$. The rupture force divided by the area $NL_{x,0}T$ gives the theoretical shear strength estimate for a beta-sheet,

$$\sigma_{max}(N) = \frac{F_{ult}(N)}{NL_{x,0}T}. \tag{94}$$

For short strand lengths, since the rupture force does not increase with strands longer than the critical length scale, the shear strength of longer strands is much lower than short ones, scaling as

Table 12
Important variables and symbols used in the manuscript for explaining theoretical considerations for beta-sheet structures (Section 5)

Variable/ symbol	Description
E_b	Variable used to describe the energy barrier in general
E_{HB}^0	Energy barrier of rupture of a single H-bond
E_b^{TEAR}	Energy barrier for TEAR loading, as measured from MD simulations
E_b^{SHEAR}	Energy barrier for SHEAR loading, as measured from MD simulations
$E_{b,max}$	Maximum energy barrier attainable by an arbitrary assembly of H-bonds
x_b	Distance of transition point
E_b	Energy barrier
L_x	Length of beta-strand
F_{break}	Rupture force
F_{ult}	Ultimate strength of a beta-strand or cluster of H-bonds
F	Force
σ_{max}	Shear strength of a beta-strand
G	Energy release rate; this corresponds to the free energy dissipated due to breaking of a cluster of H-bonds
γ_s	Adhesion energy, H-bond assembly
L_{cr}	Critical strand length
N_{cr}	Critical number of H-bonds that rupture concurrently
τ	Characteristic time-scale for H-bond rupture
ξ_p	Persistence length of polypeptide

$$\sigma_{max}(N) \sim \frac{1}{N}, \qquad (95)$$

for large $N > N_{cr}$. We discuss the implications of this results in the following section.

5.2.8. Structure–property relationship for beta-sheet systems

Since the number of H-bonds in a beta-strand is proportional to the number of residues ($L_{cr} \sim N_{cr}$, assuming that the H-bonds are arranged in a linear geometry as shown in Fig. 88a), numerical prediction for N_{cr} leads to a critical geometric strand length L_{cr}: strands beyond L_{cr} are prone to localization
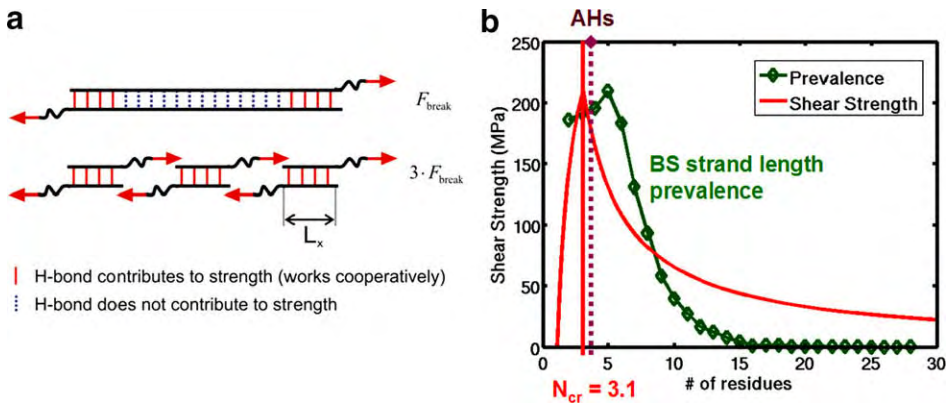
Fig. 88. Size effects, shear strength and prevalence of strand length of beta-sheets (adapted from [300]). Subplot (a): This plot illustrates the difference of the strength of a single, long beta strand vs. a combination of multiple small strands. In the upper plot, only H-bonds at the boundary are being stretched and contribute to the strength. In the lower plot, all H-bonds throughout the entire structure contribute to the strength, making the overall structure three times stronger. Subplot (b): Shear strength of beta-sheets as a function of strand length, and prevalence of beta-sheet strand length as reported in [284]. The highest shear resistance is found at a characteristic length scale of 3.1 residues. Beyond this length scale, the shear strength drops rapidly. The plot of the prevalence over the strand length illustrates that shorter beta-strands are more prevalent [284]; in particular, strands that employ less than five residues are most common, and the prevalence decays sharply after this point. The dotted line marks the number of H-bonds in alpha-helical protein domains, where each convolution features 3.6 H-bonds per turn.

of deformation and rupture of H-bonds that reduces their efficacy, since not all of the H-bonds partic-ipate in the rupture process and thus contribute to the strength (Fig. 88a). The physical basis for this limitation is the entropic elasticity of the protein backbone (an intrinsic property of protein structures) as well as the characteristic energy of H-bonds (also an intrinsic property of protein structures).

Notably, this finding explains recent proteomics data [284] obtained by analyzing the strand lengths of a wide range of beta-sheet structures. Fig. 88b illustrates that shorter beta-strands are more prevalent [284]; strands that employ less than five residues are most common, and the prevalence de-cays sharply after this point.

Further, it is found that the shear strength correlates closely to the prevalence of the strand length (see Fig. 88b). The key conclusion of this analysis is that the evolutionary driving force for the selection of strand lengths in beta-sheet structures may be the maximization of the mechanical (and hence, thermodynamical) stability. This hypothesis is strongly supported by the fact that both the prevalence curve and the shear strength show a similar behavior (Fig. 88b).

The intrinsic strength limit presented here applies to individual protein domains and can only be overcome by using structural hierarchies. Indeed, such hierarchical structures are commonly found in many mechanically strong biological materials such as spider silk and muscle tissue.

Moreover, the theoretical development reported here is not limited to beta-strands. As long as the system of interest contains an assembly of H-bonds that are loaded uniformly, the theoretical deriva-tion reported here is valid. Thus, our finding also explains the characteristic structure of alpha-helical proteins, which feature 3.6 H-bonds per turn, which are loaded in parallel and that break concertedly (as recently shown in [263]). The characteristic number of H-bonds per turn in AHs closely resembles the strength limit predicted by our theory. In other words, increasing the number of H-bonds per turn does not have an effect in increasing the mechanical stability of the protein. Similarly, beta-helix pro-tein structures also feature a helical arrangement of ultra-short beta-strand segments with each less than 5 H-bonds. Fig. 89 compares characteristic length-scales observed in common protein secondary structures and the theoretical prediction presented here.

The fact that the results of this model only depend on fundamental, 'first-principles' properties of protein structures underlines the significance of the findings. The only input parameters in this model are the persistence length ξ_p of the polypeptide chain, as well as the dissociation energy of a H-bond, E_{HB}^0 (γ_s is calculated directly from E_{HB}^0). Both parameters can be determined reliably from either exper-iment or atomistic simulation. Fig. 90 further illustrates the robustness of the value of N_{cr} due to vari-ations in the parameters ξ_p, E_{HB}^0 and τ.
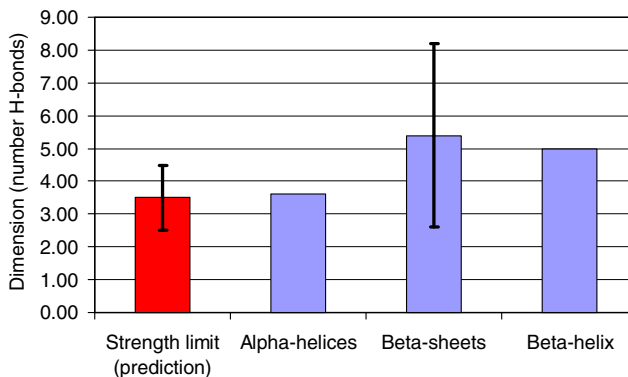
Fig. 89. Characteristic dimensions of common protein secondary structures. We compare the characteristic dimensions of alpha-helices, beta-sheets and beta-helices to our strength limit. Since the theoretical derivation considers uniform deformation of hydrogen bonds with no particular specificity to geometry, it may also apply to other protein structures where Nature utilizes geometric confinement to achieve higher mechanical stability. The fact that 3.6 HBs per convolution exists on alpha-helices and beta-sheets on the sides of helices occur in clusters of 4 may be indicative of such a biological concept.
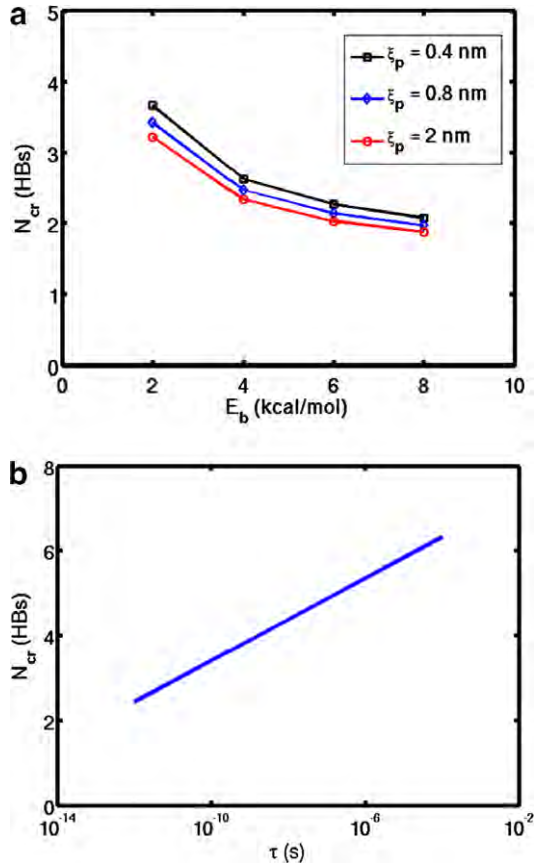
Fig. 90. Sensitivity (error) analysis for the critical number of H-bonds, N_{cr} (adapted from [300]). This plot illustrates the sensitivity of N_{cr} due to variations of the key physical parameters used in the length-scale derivation. This analysis illustrates the effect of uncertainties in these parameters on N_{cr}. Subplot (a): We plot N_{cr} as a function of E_b and ξ_p and show that $N_{cr} \approx 3 \pm 1$ H-bonds over the domain of a range of measured values for variations in the H-bond energy and for variations in the persistence length. Subplot (b): There is also a weak logarithmic dependence of N_{cr} on the characteristic time of H-bond rupture, τ. The critical number of H-bonds $N_{cr} \approx 4 \pm 2$ H-bonds over eight orders of magnitude of time-scale (here we use $E_b = 2.83$ and $\xi_p = 0.4$ nm). These results illustrate the significance and universality of our finding and support that the results obtained are very robust.

5.2.9. Summary and conclusions

A variety of models for the fracture mechanics of ceramics and metals have been reported over the past decades, involving detailed descriptions of dislocation and crack extension mechanisms. However, similar advances for biological protein materials have thus far remained elusive. To the best of our knowledge, the results reviewed in this section for the first time describe a rigorous fracture mechanics approach to describe the fundamental bond rupture events in protein materials. In analogy to dislocation nucleation and propagation in ductile metals, the breaking of H-bonds represents a fundamental unit mechanism of materials failure.

It has been shown that the rupture strength of H-bond assemblies is governed by geometric confinement effects, suggesting that clusters of at most 3–4 H-bonds break concurrently, even under uniform shear loading of a much larger number of H-bonds. This universally valid result leads to an intrinsic strength limitation that suggests that shorter strands with less H-bonds achieve the highest shear strength.

As reported in Fig. 85, our prediction close to 127 pN is in excellent agreement to a wide range of experimental data, and may explain the convergence to a finite limiting value at small deformation speeds. Our size-dependence analysis becomes of particular interest when compared with recent experimental proteomics data, suggesting a correlation between the shear strength and the prevalence of beta-strand lengths in biology (see Figs. 88b and 89). Shorter strands are more common in nature, and they exhibit higher mechanical (and hence thermodynamical) stability according to our prediction. This hypothesis is confirmed by direct large-scale full atomistic MD simulation studies of beta-sheet structures in explicit solvent. It is also in agreement with experimental results that show a non-linear, diminishing increase of strength with increasing H-bond cluster size [237]. The finding suggests that the intrinsic strength limitation of H-bonds is overcome by the formation of a nanocomposite structure of H-bond clusters, thereby enabling the formation larger, much stronger beta-sheet structures. Such hierarchical features are commonly observed in strong materials such as spider silk and muscle fibers. The results could have interesting implications for the design mechanically robust, strong biomimetic or peptide based self-assembled materials. The model further confirms that fracture mechanics concepts, previously primarily applied to macroscale fracture phenomena, can also be directly applied at nano-scale, and can be used to describe failure mechanisms in protein materials. Similar approaches can be used to lay the foundation for the development of plasticity and strength models of biological protein materials.

5.3. Mechanical characterization of beta-helical protein structures

This part is focused on the nanomechanical properties of the β-helix, reviewing a molecular simulation study reported in [305]. A snapshot of the molecular structure of this protein material is shown in Fig. 91. The β-helix is a protein structure that has been discovered relatively recently [306]. Essen-
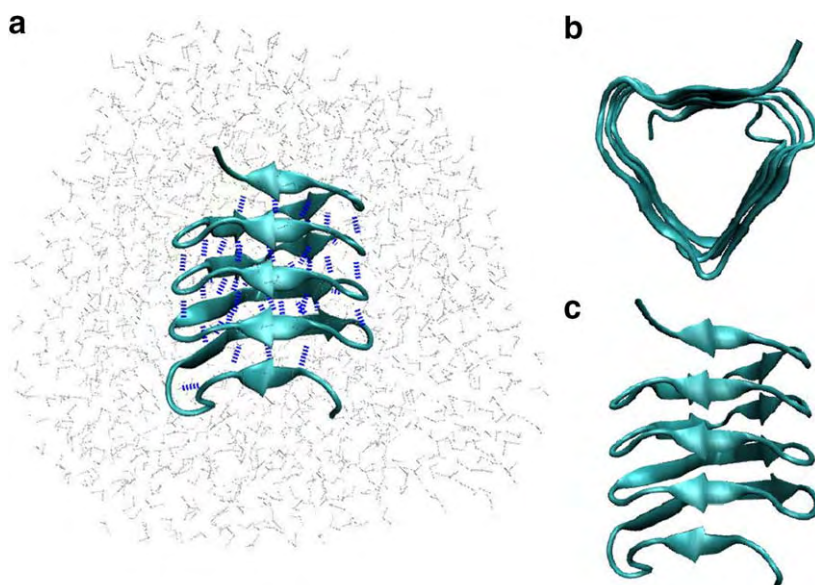
Fig. 91. Full atomistic model (subplot a): The structural coordinates of the beta-helix protein are obtained from the Protein Data Bank (PDB ID 1fwy) and solvated with a skin of water. Hydrogen bonds that are oriented parallel to the helix axis and perpendicular to the strands enable this system to be stable in this conformation. The beta-helix is a nanotube-like structure (subplot b) with a triangular cross-section formed by three strands on the edges. This particular protein structure is classified as a single-stranded left handed beta-helix, consisting of a single polypeptide chain that creates a helix with approximately four to five beta-strands per side (subplot c). It is noted that in most snapshots shown in this paper, the water molecules are not plotted to improve the clarity of the figures.

tially, this protein structure represents a helical nano-tube that consists of beta-sheet walls. The beta-helix is generally considered as a separate type of secondary structure. Beta-sheets and beta-helices make up all beta-structured protein motifs discovered so far, and will be treated in the same section here since the chemical and mechanical properties of these two structure types are linked.

The purpose of the analysis reviewed in this section is to illustrate the applicability of the simulation approaches to a recently discovered structure, for which little or no experimental data exists. Predictive modeling approaches such as the one presented here may provide valuable insight into structure–function–property relationships of protein structures and may motivate experimental studies associated with disease aspects or materials science applications of proteins.

Due to its recent discovery, this protein structure has not been studied as extensively as α-helices or β-sheets or many other protein structures, and therefore, there exists virtually no understanding of its mechanical behavior. To the best of our knowledge, no experimental or computational studies of the mechanics of this protein structure have been reported up to date. Here we summarize results from molecular dynamics simulation and atomistically informed continuum modeling of the nanomechanics of the beta-helix protein motif, a key building block of amyloids associated with Alzheimer's, Huntington's and prion diseases, forming a nanotube-like protein structure.

As reported in [305], we carry out MD simulations that employ a fully atomistic model with an explicit solvent and very slow pulling rates over extraordinarily long time-scales reaching the fraction of a microsecond (250 ns) to simulate the tensile deformation of this motif. Inspired by the beta-helical domain of the needle-like cell puncture device of bacteriophage T4 (see Fig. 92), we review a series of compressive loading molecular dynamics simulations of the beta-helix, and show that this protein motif can withstand extremely large compressive loads, far exceeding the tensile strength. Different mechanical deformation modes under compressive loading are summarized in a deformation map. The findings illustrate the potential of the beta-helix protein motif as an inspiration for nano-scale materials applications, ranging from stiff nanotubes to self-assembling peptide based fibers inspired by amyloids.

The studies reviewed here elucidate the intriguing mechanical properties of this important self-assembling protein structure that may be eventually be used for developing exceptionally strong and elastic fibers that can surpass the strength of steel, as in the case of spider silk [279].

5.3.1. Significance of beta-helical structures: biological, materials science and disease aspects

Despite being discovered only recently as a protein structure, the β-helix has already gained tremendous interest in the scientific community. This is largely attributed to the belief that the β-helix is the fundamental structural unit of several different kinds of pathological amyloid fibers [307–309]. Amyloids are caused by uncontrolled aggregation of proteins in tissue and organs during the course of many different fatal diseases such as Alzheimer's, Huntington's, prion diseases and type 2 diabetes [278]. It is not clearly known whether the amyloid fibrils form as the result of a mechanism to convert more toxic oligomers to larger, less harmful aggregates, or if they are the pathological agents themselves [274]. However, it is now generally accepted that large amounts of amyloid deposition in tissue may be a cause of organ malfunction. There is also growing evidence suggesting that amyloids play a central role in pathogenesis of at least some of the aforementioned diseases [310]. Investigation of this new protein structure is therefore of utmost importance for the medical and biophysical community as it may reveal crucial insight into the formation and assembly of these fibers and may provide ideas for novel remedies to diseases affecting millions. Due to recent X-ray diffraction analyses and elucidation of the atomistic structure proteins that contain this motif, the protein structure is now accessible to all-atom simulations [306].

Amyloids are most widely known for the pathological conditions that they are directly or indirectly associated with. From a materials scientist's perspective, however, amyloid formation is a fascinating materials growth phenomenon that can be put in use for beneficial applications [278]. Engineered self-assembling amyloid fibrils have been used as nano-wire templates, taking advantage of their patterned nanotubular structure [311,312]. Amyloids are also found in some natural adhesives [277], and in spider silk in silk glands [313,314], which has attracted some interest into possible structural applications of these materials and their interesting mechanical properties [315,316]. Notably, amyloids are found to be extremely sturdy protein structures that can often not be disintegrated by the organ-
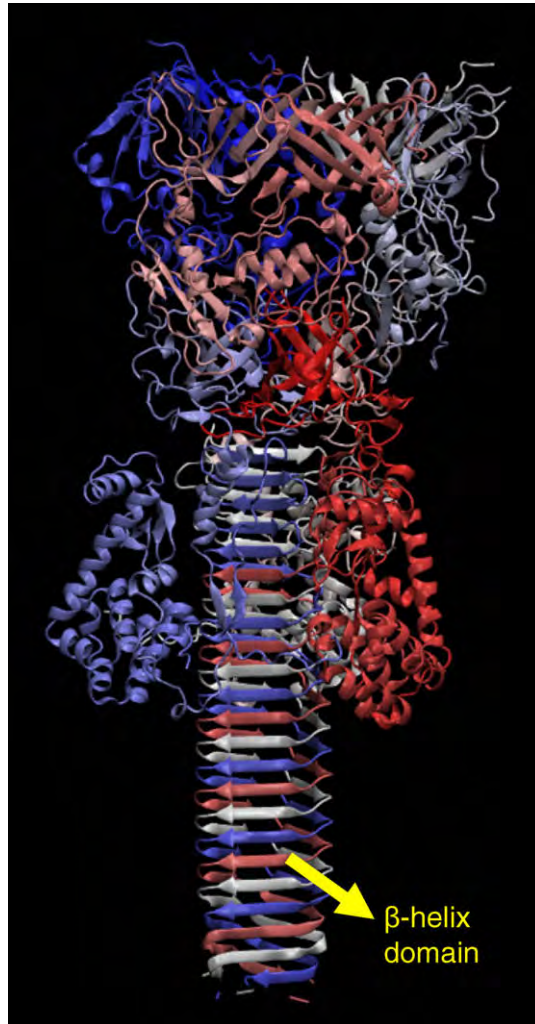
Fig. 92. Molecular structure of the triple beta-helix domain of bacteriophage T4 in its biological context (PDB ID 1K28, visualization with VMD [291]). The tail structure of bacteriophage T4 is used as a cell-puncture device and serves key function in its host infection mechanism. Our atomistic compression studies with SMD approach on our model beta-helix system is inspired by the cell-puncture device of T4.

ism itself, suggesting a particularly stable structural arrangement and as a consequence, possibly large resistance against mechanical unfolding. These aspects motivate the studies of the mechanical properties of this protein structure.

Although a single structural motif such as the β-helix may not be an accurate molecular level description for all the different kinds of amyloids discovered so far, it is a model that captures all essential structural and biochemical features common to amyloids and is therefore an excellent representative system to begin a rigorous, nanomechanical characterization of amyloids and similar fibrous materials.

### 5.3.2. Molecular structure of beta-helices and the simple atomistic model for MD studies

Beta-helices form when a single polypeptide chain is coiled into a wide helix, formed by beta-strands separated by loop regions of variable length. Each side of the helix is an anti-parallel

beta-sheet structure and helices generally contain two or three-sheets (sides). For the triple-sheet beta-helices investigated here, no specific sequence pattern has been identified so far [317].

All-atom MD simulations are performed using segments from the crystal structure of a protein obtained from Protein Data Bank [86] (PDB ID 1fwy). The segments used for the simulations form the characteristic β-helix structure. For investigating mechanics of β-helix structures, we utilize the C-terminal domain of this protein, which adopts a left-handed parallel β-helix motif (residues 250–328).

5.3.3. Atomistic simulation results: tension versus compression

Here we review the force–displacement profiles obtained from simulations of the β-helix model. The results reveal that the mechanical behavior of this molecule is different from a Hookean spring. An important feature is that the mechanical response is quite asymmetric with respect to tension and compression, as shown in Fig. 93a and b. The maximum force observed in compression happens at approximately 17% engineering strain and is six times larger than the peak force observed in tension at approximately 800% strain (Fig. 93c). Up until 17% strain, both systems behave more or less linearly, allowing comparison of stiffness values in small deformation regime. The elastic moduli computed using both hollow and solid cross-section assumptions show that the β-helix in compression is approximately 18 times stiffer than in tension, as shown in Fig. 93d.
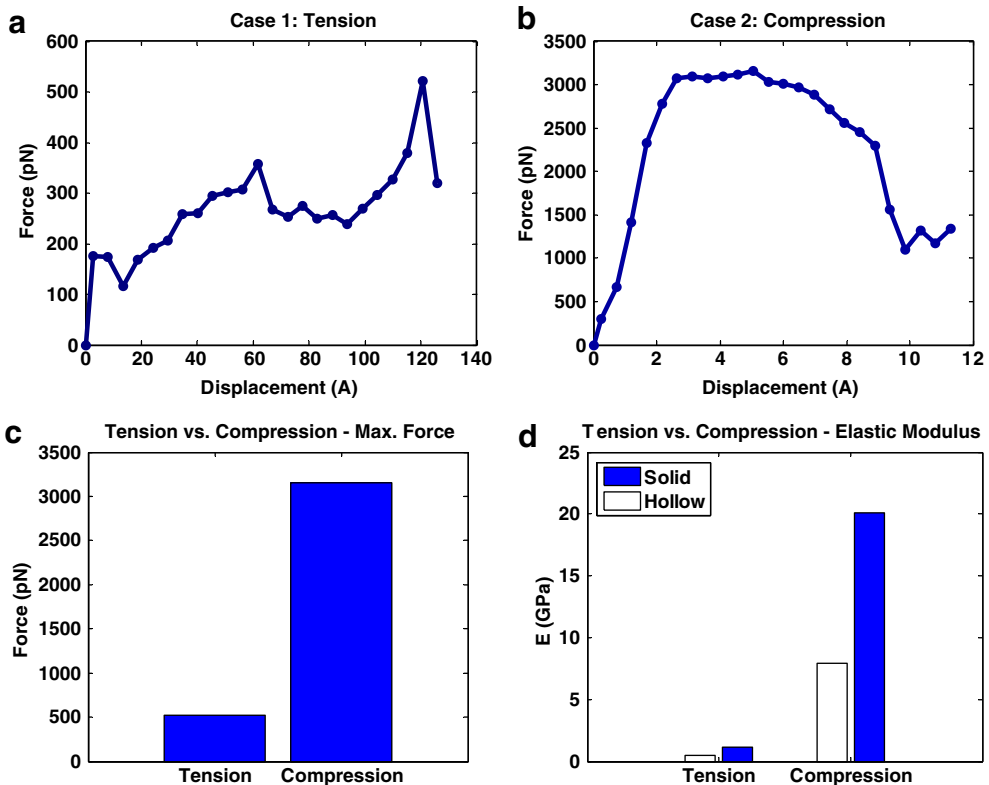
Fig. 93. Tensile vs. compressive strength (subplots a and b). The mechanical behavior of the beta-helix is rather different from a Hookean spring, since the response is strongly asymmetric in tension and compression. The maximum force observed in compression is at ≈17% engineering strain, and is six times larger than the peak force observed in tension at ≈800% strain (subplot c). Up to ≈17% strain, both systems behave more or less linearly, allowing comparison of stiffness values in small deformation regime. The elastic moduli computed using both hollow and solid cross-section assumptions show that the β-helix is approximately 18 times stiffer in compression than in tension (subplot d).
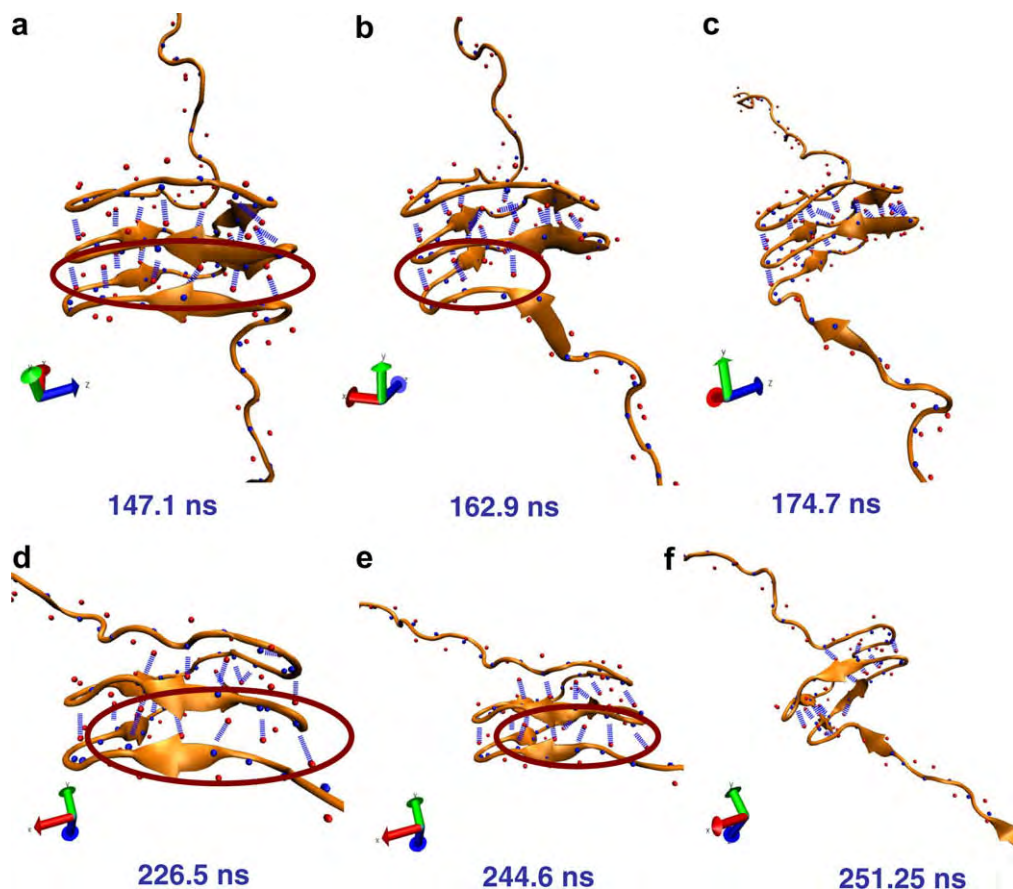
**Fig. 94.** Tensile deformation – sequential vs. simultaneous rupture. The peaks, rises and valleys in the complex, non-linear force response curve can be better interpreted by studying individual bond rupture events. While most of the helix loops uncoil with hydrogen bonds breaking more or less sequentially (subplots a–c), comparatively long time-scales of the simulation and the coupling between tension and torsion in the helix allow the molecule to reorient its strands during the course of the simulation. At this point, five hydrogen bonds in a sheet deform rather homogeneously under shear, cooperatively resisting further stretching like a strong traction point (subplots d–f). Simultaneous rupture of these bonds corresponds to the largest force peak obtained over the course of the 250 ns time span of the simulation.

The peaks, rises and valleys in the complex non-linear force response curve can only be understood by studying the details of the individual bond rupture events. Most of the helix loops uncoil with hydrogen bonds breaking more or less sequentially. Comparatively long time-scales of the simulation and the coupling between tension and torsion in the helix allow the molecule to reorient its strands during the course of the simulation. In the later stages of unfolding, a sheet with five hydrogen bonds aligns itself parallel to the loading direction, as opposed to the general perpendicular orientation of the strands. In this case, five hydrogen bonds in this sheet deform homogeneously under shear, cooperatively resisting further stretching like a hook-up point. Simultaneous rupture of these bonds corresponds to the force peak observed at approximately 120 Å pulling displacement. Simultaneous versus sequential rupture events are illustrated in Fig. 94.

The force versus displacement profile of the β-helix reveals distinct deformation mechanisms under tensile loading, which are marked in Fig. 95a. Here, the same averaging scheme is used, with a different number of averaged points (bins) to illustrate different levels of detail in force response curve (different curves in Fig. 95a) to enable a more careful analysis of the force–extension curve. The first
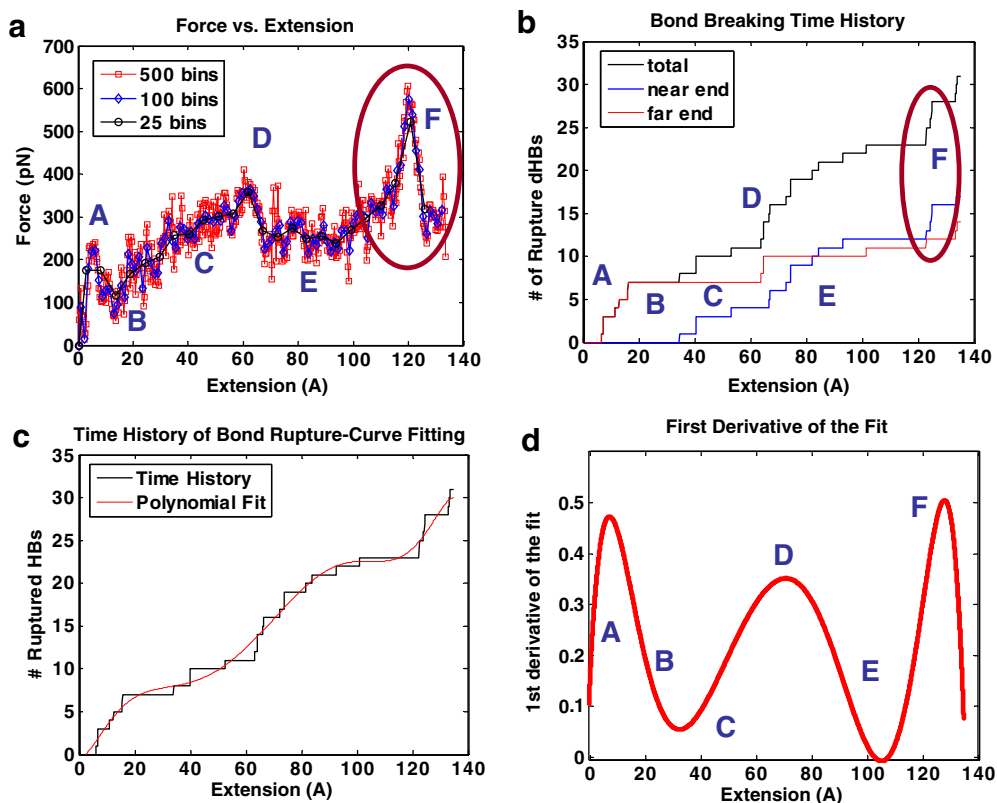
Fig. 95. Relationship between bond rupture events and mechanical signature. The force vs. displacement profile of the β-helix reveals distinct deformation mechanisms under tensile loading (subplot a). Here, we use the same averaging scheme with different number of bins (see methods) to illustrate the differences in level of detail associated with this process. The first peak in force–displacement curve (A) corresponds to the unfolding of the first loop at the fixed end. Unfolding of this loop relaxes the structure, resulting in the low force point (B). The regime that follows this is characterized by steady uncoiling of loops at the far end, with no more than 1–2 hydrogen bonds breaking at the same time (C). Peak force at ∼60 Å extension corresponds to the unfolding of the first near end loop (D). Following this, the molecule relaxes again and steadily uncoils with lower resistance (E). The force begins to increase again when the helix is reoriented and hydrogen bonds are loaded in parallel, and reaches a crest with a force value of ∼600 pN (F) when five hydrogen bonds on a single side break simultaneously. The direct correlation between the number of bonds that break and the mechanical response can be quantified by plotting the number of ruptured bonds as a function of extension (subplot b). To further analyze the bond rupture events as a function of time, a ninth order polynomial function is fitted onto the bond data (subplot c). The slope of this function relates to the speed of bond breaking events. The troughs correspond to periods where the backbone is being stretched out without any hydrogen bonds breaking, and the peaks indicate simultaneous rupture of a large number of bonds (subplot d).

peak (A) corresponds to the unfolding of the first loop at the fixed end. Unfolding of this loop relaxes the structure, resulting in a decrease in force (B). The regime that follows this is characterized by steady uncoiling of loops at the far end (C). The peak at approximately 60 Å pulling extension corresponds to the unfolding of the first loop at the near end, namely the N-terminal of the protein (D). Following this, the molecule relaxes again and steadily uncoils at a lower force value (E). The force begins to rise again due to reorientation of the helix to load hydrogen bonds in parallel and reaches a peak force value of approximately 600 pN (F), when approximately five hydrogen bonds on a single side break simultaneously.

The direct correlation between the number of bonds that break and the mechanical response can be quantified by plotting the number of ruptured bonds as a function of extension ($N_{HB}(x)$), as shown in Fig. 95b. To more rigorously analyze the bond rupture events as a function of deformation, a ninth or-

der polynomial function is fitted to the bond break history data. The fit is depicted in Fig. 95c. The first derivative of this function, plotted in Fig. 95d relates linearly to the rate of bond rupture events since the protein is pulled at a constant pulling velocity (thus $N_{HB}(t) \sim N_{HB}(x)$ since $x \sim t$ at constant $\dot{x}$).

The valleys correspond to periods where the backbone is being stretched out without any or with few hydrogen bonds breaking, and the peaks indicate simultaneous rupture of a larger number of hydrogen bonds. This can be confirmed by comparing Fig. 95d with Fig. 95a. It is apparent that each peak in the force–extension curve corresponds to a peak in the slope of the H-bond rupture history $N_{HB}(t)$, suggesting that the instantaneous strength of the β-helix is proportional to the H-bond breaking rate,

$$F(t) \sim \frac{\partial N_{HB}(t)}{\partial t}. \tag{96}$$

This result may be useful for the development of constitutive relations or strength theories of protein structures.

In compression simulations, it is observed that failure occurs by sliding out of an entire single strand, corresponding to a local fracture scenario at a critical compressive force $P_{fr} = 3154.40$ pN.

5.3.4. Theoretical model: size-dependence of compressive strength

In order to gain understanding of the mechanical behavior at larger length-scales, we propose a simplistic continuum model for the compressive strength of the beta-helix. Schematic views, boundary conditions and the dimensions of the simplified hollow triangular section are shown in Fig. 96. We believe that localized failure as observed in MD simulations competes with Euler buckling of the structure. Using the simplified model illustrated in Fig. 96, we calculate the Euler buckling load [318] for this specific geometry as

$$P_{cr,0} = \frac{\pi^2 EI}{4L_0^2}, \tag{97}$$

where L_0 is the original length of the helix and the factor of four in the denominator derives from fixed-free boundary conditions. For the length L_0 used in the simulations, the Euler buckling load is $P_{cr,0} = 6474.11$ pN (see also Table 13). This value is larger than the maximum compressive force value observed from simulation, denoted $P_{fr} = 3154.40$ pN, which is observed for a localized shear failure of the nanotube. Thus, it is expected that the dominating failure mode for the geometry used in the MD simulations is not Euler buckling. This prediction is in good agreement with the MD simulation results that show that local failure is the governing mechanism. However, as the length of the molecule is increased we expect that Euler buckling rather than localized failure will be the dominating deformation mechanism.

Using the peak force from SMD as an estimate for the maximum strength of the helix for a local fracture scenario, the critical molecular length beyond which Euler buckling [318] governs is calculated using this value as the critical buckling load:

$$L_{cr} = \pi \sqrt{\frac{EI}{P_{fr}}} \approx 2.15 \text{ nm}. \tag{98}$$

Based on these results, the compressive strength of the β-helix as a function of its length can be plotted, as shown in Fig. 97a. The analysis reveals that buckling governs at length-scales exceeding only a few nanometers. The resulting compressive strength of the helix is determined by the governing mechanism, given by

$$P_{max}(L) = \min[P_{fr}, P_{cr}(L)], \tag{99}$$

where P_{fr} is the maximum force observed in local fracture deformation mode and $P_{cr}(L)$ is the buckling load as a function of length. As expected, strength diminishes with increasing length. We note that this simple calculation is based on the assumption that the local failure strength P_{fr} is independent of the molecular length. The persistence length [319] of this structure, ξ_p, can also be estimated as
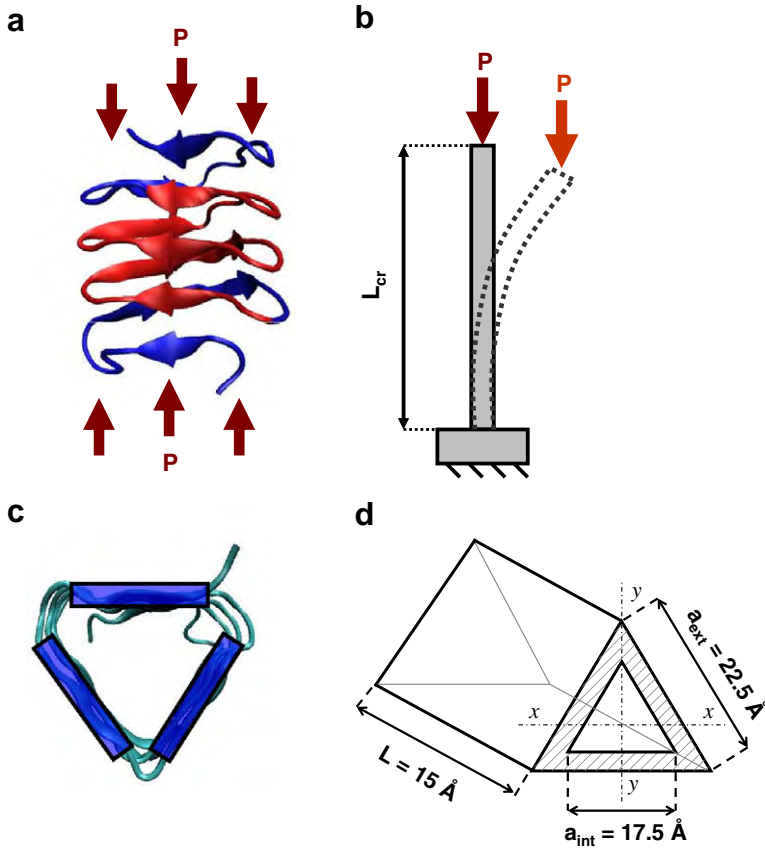
Fig. 96. Schematics of the compression studies in an engineering analysis. The molecular structure is subjected to uniform compression load on each Cα atom on the first loop at the N-terminal (subplot a). All of the Cα atoms are fixed in the last loop, namely C-terminal of the domain. The boundary conditions (subplot b) are idealized based on the SMD setup and are used for buckling analysis. The cross-section of the β-helix is triangular, with a strand on each side (subplot c). The dimensions of the helix can be approximated from VMD to build a simplistic model. In this case, we model our helix as a hollow triangular section. All calculations are repeated for a solid section assumption as well (subplot d).

$$\xi_p = \frac{EI}{k_B T}, \tag{100}$$

where k_B is Boltzmann's constant, and T is temperature, given as 300 K, and EI is determined from material and structural properties determined from all-atom model and simulation results. At length-scales exceeding ξ_p, even thermal vibrations could cause significant deformation in the molecule, so the behavior is like an ultra-thin wire. For the β-helix structure, we find this length to be in the order of a few micrometers (ξ_p = 1.425 μm, see Table 13). The length-scales estimated here can be utilized to construct a simple deformation map, indicating regions of different mechanical behavior as a function of the protein nanotube's length, as presented in Fig. 97b.

From SMD simulations and continuum studies, one can conclude that in compression; very short helices below a few nanometer lengths are sturdy and fail locally. Longer helices are governed by buckling effects, and extremely long helices beyond a molecular length on the order of a micrometer behave like wires with negligible capacity for sustaining compressive loads. This suggests that the nanotube length must limited to a critical length scale on the order of a few nanometers in order to provide most optimal resistance to compressive deformation.

Table 13
Summary of nanomechanical properties of the β-helical protein model

Maximum force (tension)	522.04	pN
Maximum force (compression)	3154.40	pN
Stiffness, k (tension)	0.65	N/m
Stiffness, k (compression)	11.63	N/m
E (hollow – tension)	1.13	GPa
E (hollow – compression)	20.14	GPa
E (solid – tension)	0.45	GPa
E (solid – compression)	7.96	GPa
I_{xx} (hollow)	2.93E−37	m^4
I_{xx} (solid)	4.62E−37	m^4
Buckling load (hollow)	6474.11	pN
Buckling load (solid)	4033.87	pN
L_{cr}, buckling length (hollow)	2.15	nm
L_{cr}, buckling length (solid)	1.70	nm
χ_p, persistence length (hollow)	1425.34	nm
χ_p, persistence length (solid)	888.10	nm
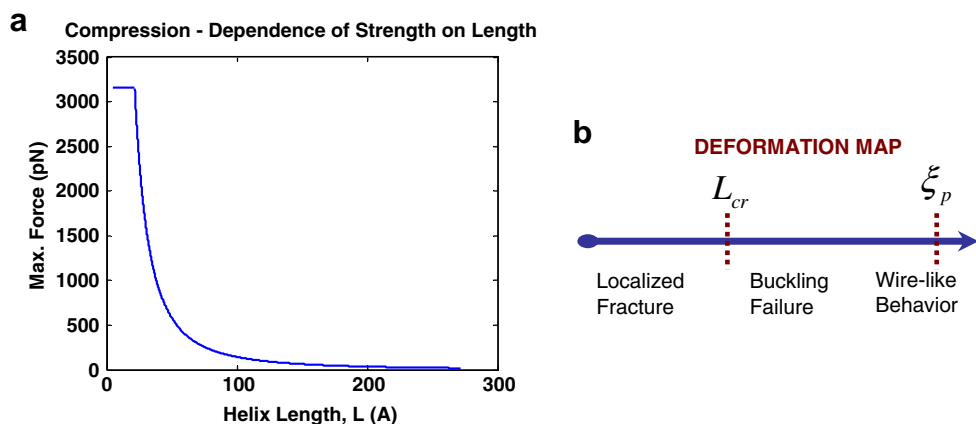
Fig. 97. Compression – analysis of results: in SMD simulations, failure occurs by sliding out of a single strand, corresponding to a local fracture scenario. The max force value observed for this case is lower than the Euler buckling load, further supporting that local failure is the governing mechanism. The maximum force observed from SMD is used as an approximation for the fracture strength for the 'short column', and the critical length beyond which buckling governs is calculated based on this value. Using these results, the strength of the helix as a function of its length is plotted (subplot a). The persistence ξ_p length can also be determined, identifying the length scale at which thermal vibrations can cause significant deformation in the molecule. These results are summarized in a deformation map (subplot b), indicating different regions of mechanical behavior as a function of length. Very short helices are sturdy and fail locally, longer helices are governed by buckling effects, and extremely long helices behave like wires with negligible capacity for sustaining compressive loads.

It is noted that a similar transition of the behavior has been observed for carbon nanotubes, albeit at different values of the critical length-scales [320]. A summary of all numerical results for tension–compression analysis of the beta-helix are provided in Table 13.

5.3.5. Structure–property relationships for beta-helical protein domains under compressive loading

MD simulations reveal that the β-helix nanotube structure is extremely extensible and can sustain tensile deformation up to ≈800% without rupture of the covalently bonded protein backbone. As the β-helix structures are similar to β-spirals found in extensible domains found in elastin and spider silk [321,322], our modeling attempts illustrate that extension over tens of nanometers and rupture of dozens of hydrogen bonds during deformation may be the molecular basis of the extraordinary tough-

ness of these materials, in line with previous hypotheses [34], as the protein structure deforms while dissipating a significant amount of energy.

Recent experimental and MD studies have revealed that proteins with a shear topology have been found to provide exceptional resistance to forced unfolding due concurrent breaking of multiple hydrogen bonds upon loading parallel to strand orientation [220,273,280–283,323]. With the analysis shown in Fig. 95d, we have further validated this claim. We have also showed that during the course of an ultra-long time-scale MD simulation that extends to fractions of microseconds, significant reorientation in the molecule can take place as a result of external force. This observation may provide insight into how beta-sheet rich domains in amyloid-like spider silk in the silk gland may be transformed to a shear topology due to forces experienced in spinning and elongational flow and may illustrate the impact of this reorientation on mechanical strength. The processing of protein fibers to generate such shear topologies could provide a means to enhance the mechanical resistance.

We also reviewed a study that is focused into the compressive strength of a protein structure. Usually, there is little motivation to study proteins under uniform compressive loads, as most proteins studied mechanically so far have found applications in fibers, materials that are predominantly loaded in tension. Thus, there are very few studies in the literature that focus on this mode of deformation. The inspiration for the compressive study reported here derives from the β-helix structure in the gp5 domain in the needle-like cell-puncture device of bacteriophage T4 [324]. Extreme responsiveness under compression shown here is a truly unique feature of the beta-helical topology among protein structures identified so far. The hollow triangular cross section of the beta-helix is thereby very efficient for serving under compression as it can sustain significant loads without buckling, while at the same time providing a highly dissipative behavior under tensile stretching. The methods presented here can perhaps be used to develop a critical length scale for the optimal resistance under compressive deformation for bacteriophage T4 cell puncture device. We leave a detailed investigation of this aspect to future work.

The persistence length for the β-helix nanotube estimated here is much higher than those calculated for single polypeptide chains, but it is lower than that of a (5,5) single wall carbon nanotube that consists of all covalent bonds, as shown in the comparison depicted in Table 14. These findings are in excellent agreement with experimental results obtained from amyloid based nano-tubes used for growing silver nanowires [311], and illustrates the suitability of this protein motif for nano-wire and nano-tube applications where stiffness is a prerequisite for robust design.

The deformation map depicted in Fig. 97b represents a first, simplistic attempt to develop a structure–function–property relationship for this particular protein topology.

5.3.6. Conclusions

Here we summarized findings from molecular dynamics simulations and atomistically informed continuum modeling of nanomechanics of the β-helix protein motif, a key building block of amyloids and self-assembling peptide based fibers. It was shown that this nano-tube like protein motif is highly extensible under tensile deformation; sustaining strains up to 800%. Under compression, the structure is very robust and exhibits high resistance before failure occurs. Through continuum approaches, the length-dependent deformation map of beta-helices under compression was illustrated.

Table 14
Comparison of the persistence length for different nanostructures and molecules

Type	Structure	Length (nm)	References
Single molecules	Tropocollagen	11–16	[97,335]
	Ig domains in titin	0.4	[236]
	Silk protein	0.4	[265]
	Tenascin	0.4	[288]
Nanotube structures	CNT	16,000	[336]
	Peptide nanotubes	>1000	[311]
	β-Helix	1430	Present study

The approach reviewed here illustrates how molecular dynamics simulations can reveal crucial information about biological structures when experimental data is not readily available. Such information can be used to build predictive models that explain materials behavior and structure–property–function relationships for proteins and protein-based materials.

The findings point out the potential of using the β-helix protein motif for a variety of nano-scale materials science applications, ranging from stiff peptide nano-tubes to self-assembling peptide based fibers inspired by amyloids. These results illustrate that beta-helix structures appear as a promising candidate for many functional nanomaterials that may be used for structural applications, tissue scaffolds and non-invasive drug delivery systems inspired by the bacteriophage T4.

5.4. Discussion and conclusions

Improving the current understanding of the nanomechanics of beta-structured protein materials – key building blocks of some of the most impressive biological fibers such as silks and muscle fibers – is of crucial importance for materials science applications of peptide systems. Such simple building blocks are the starting point for designing self-assembling peptide materials. Their thermodynamical and mechanical characterization is a necessary step to accurately create materials that will serve a particular function. The relatively easy edge-to-edge assembly and aggregation of beta-sheet structures quickly into fibrils and fibers makes these simple systems extremely promising for new materials and material synthesis techniques.

So far nanomechanical characterization of protein materials has focused on individual proteins, in an attempt to explain their specific biological functions or structural features from a biophysical perspective. Difficulties in interpretation and generalization of results from unfolding experiments or simulations of complex large-scale individual proteins are evident. For this reason, molecular modeling of particularly simplistic protein structures that play a universal and crucial role in nano-scale deformation mechanisms of many protein materials is currently the only reliable option for shedding light on deformation mechanisms, size effects and other important aspects that cannot easily be generically defined for more complex systems. Experimental studies at these length-scales are extremely challenging with current technologies, therefore predictive models that can generalize MD simulation results at these length-scales is indispensable for explaining experimental findings at larger scales.

From a materials scientist's perspective, generic properties of chemical interactions, bond rupture mechanisms and deformation modes are key input parameters that one needs to build multi-scale models that can predict properties of bulk materials. In light of this, the fundamental, simplistic approach we reviewed here for beta-sheets and beta-helices is extremely important for explaining universal deformation mechanisms, size effects and chemical/thermodynamical properties of biological materials and structures. Simulation methods coupled with theoretical frameworks such as those reviewed here illustrate the state-of-the-art in nano-scale/multi-scale modeling techniques for materials applications.

6. Discussion and conclusion

The review of results in the previous chapters has provided an overview over developments and capabilities of bottom-up computational models in studies of deformation and fracture of biological protein materials. In this final section we provide a discussion and an outlook to possible future research directions.

This review article was focused on three protein materials, collagenous tissues, alpha-helical rich protein materials, as well as beta-sheet rich protein materials. These represent some of the most abundant protein materials found in biology.

First, we focused on atomistically informed multi-scale modeling of collagenous materials, in search for an accurate micro-scale description of composite materials that employ collagen networks, including pure collagenous tissue and mineralized tissue such as bone. The multi-scaling technique reviewed here (see schematic in Fig. 2) enables a bottom-up approach for modeling complex compos-

ite systems that have components of different mechanical properties (tropocollagen molecules, as shown in Fig. 14, cross-linking between tropocollagen molecules forming collagen fibrils as shown in Fig. 27, mineral platelets dispersed in collagen networks, as shown in Figs. 45 and 47, and other structures). Similar approaches may be used for other protein materials. A condition for applicability of these models is the clear separation of scales, that is, that structural features can be clearly associated with specific length-scales. This separation of scales enables one to adapt coarse-graining approaches, as reviewed throughout the manuscript, for instance in Figs. 6, 26, 59 and 60. The discussion then continued with a discussion the rupture mechanics of alpha-helices, with a focus on atomistic and molecular processes. We reviewed a modified Bell formulation, which allows for predicting the properties of hierarchical protein structures directly from knowing the geometry and basic properties of the H-bonds. This was illustrated in several case studies, for instance the example depicted in Fig. 74 that showed the approach for different hierarchical arrangements of alpha-helices. Although this model may be directly adapted for other structures without significant changes, this approach has been discussed only for alpha-helical protein structures. Future work could focus on an extension of these models to other protein structures, in particular to more complex geometries. Further simulation studies to validate the predictions by the theoretical model could be carried out.

For the beta-sheet section, we focused on the shear topology of these protein materials and attempted to specifically address the size- and geometry-dependent strength properties of these protein domains. For this, the WLC model is used in conjunction with the Griffith–Irwin energy balance concept to describe the physical basis of protein unfolding and rupture of H-bonds (see, for instance, Fig. 86). This model can also be adapted to other protein structures for similar purposes; however, the discussion of this method is focused solely on the application to beta-sheet rich structures. A preliminary analysis of these concepts is shown in Fig. 89.

The examples illustrate the range of mechanical properties (elasticity, e.g. Figs. 28 and 30, plasticity, Figs. 33 and 58, and fracture, Figs. 37 and 59) and studies at different length- and time-scales (picoseconds, as shown in Fig. 67, or microseconds, as shown in Fig. 30) that can be carried out with atomistic and molecular simulation methods. The studies reveal fundamental deformation mechanisms of protein materials, most notably:

- Molecular unwinding, intermolecular molecular slip and molecular rupture for tropocollagen molecules (e.g. Figs. 18 and 20).
- Rupture of clusters of H-bonds, for both alpha-helices and beta-sheet structures (e.g. Figs. 67, 83 and 84).
- Intermolecular slip and rupture between protein domains, as illustrated for tetramer structures of vimentin IFs (e.g. Fig. 76) and for the collagen fibril (e.g. Figs. 24 and 48).

Many of these mechanisms can ultimately be linked to the breaking of H-bonds, which may be considered an elementary unit event of deforming many biological materials as introduced schematically in Fig. 1c. The models reviewed in Sections 4 and 5, focused on atomistic description of breaking of clusters of H-bonds, provide quantitative predictions of the mechanisms and mechanical resistance of such processes. For instance, the studies shown in Fig. 67 (rupture of H-bonds in alpha-helices) and the results reviewed in Figs. 83 and 84 (rupture of H-bonds in beta-sheets) provide important insight into these mechanisms [40,299,300]. The strength analysis over a range of experimental and computational studies as shown in Fig. 65 reveals that the strength limit may be a universal property of H-bond clusters in proteins (see discussion in Section 5), as suggested in Refs. [299,300].

6.1. Structural complexity and simplicity in biological protein materials: Universality-diversity paradigm

Even though protein materials lead to vastly complex structures such as cells, organs or organisms, an analysis of their composition reveals simple underlying mechanisms that can be classified into two major categories, *universal* and *diverse* features [41]: some of the structural features materials are commonly found in different tissues, that is, they are highly conserved. Examples of such universal building blocks include alpha-helices, beta-sheets or tropocollagen molecules. In contrast, other features are highly specific to tissue types, such as particular intermediate filament assemblies

[181,188,189,285,325], beta-sheet nanocrystals in spider silk [326] or tendon fascicles [8,89,90,92]. These examples illustrate that the coexistence of universality and diversity is an overarching feature in protein structures. This represents an important concept of materiomics. Future studies might be aimed at further investigating these aspects, as they could play an important role in the understanding of the architecture and evolution of biological materials and structures.

Such analyses may be of vital importance from a scientific, technological as well as sociological perspective. The collaboration between materials scientists and structural biologists is vital and will be mutually beneficial: materials scientists have extensive of experience in treating structures, processes and properties of materials systematically and with rigorous mathematical methods. On the other hand, biologists have gained a detailed understanding of biological systems and structures and related functions. This will lead to a better understanding and new theories, which will build the foundation for the design of synthetic hierarchical structures and systems. A detailed analysis of hierarchical biological materials may contribute to a variety of scientific disciplines, such as the science of fracture, materials theory, genetic research (e.g. the hierarchical three-dimensional folding of the DNA). In these examples, the link between structural organization and function [24] is a vital component that might further contribute to the understanding of which driving forces in Nature create hierarchical biological materials.

6.2. Theoretical, computational and experimental challenges and future directions

In order to realize the promising opportunities that arise from an improved understanding of hierarchial protein materials several critical challenges must be overcome. Up until now, theories describing hierarchical biological materials are still lacking. Virtually no understanding exists about how specific features at distinct scales interact, and for example, participate in mechanical deformation. However, such models are vital to arrive at a solution for the universality-diversity question and Nature's hierarchical material concepts. The path to success is to develop cross-scale relationships and constitutive equations for different hierarchical scales within the structure–property paradigm of materials science (see Fig. 98), that is, to understand if and how nano-/meso-/micro-changes affect properties at larger scales. To achieve this goal, structural architecture will have to be considered across the scales, possibly combined with fractal theory [327].

Furthermore, the nomenclature for hierarchical biological materials is still missing. Definitions and measures for material properties such as hierarchical degree, level of robustness, degree of universality, and others, are crucial. Appropriate terminology for cross-scale relations such as scale-separation,
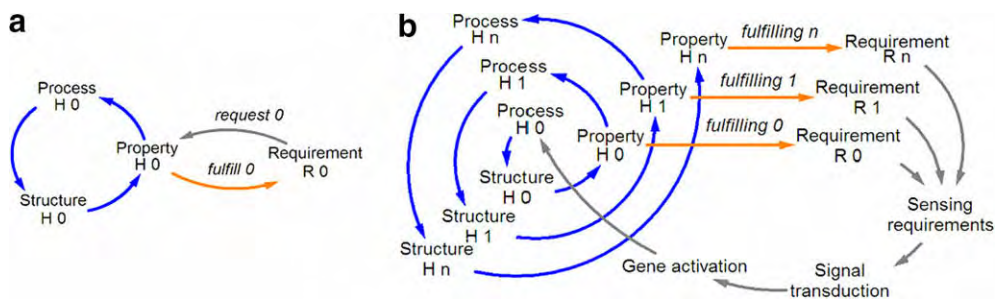
Fig. 98. Structure–process–property paradigm in biological protein materials, the basis of materiomics. In biological materials hierarchical structures, decentralized processes, material properties and environmental requirements, are brought together in mutual completion [41]. Subplot (a) illustrates the traditional paradigm in materials science where process, structure and property build the "magic" triangle on a single hierarchical level. Subplot (b) illustrates the paradigm for hierarchical (biological) materials. In contrast to the traditional paradigm, relations between "external" functions/requirements and "internal" properties exist on several scales resulting in multi-functionality. Further, as requirements are consistently changing over time (e.g. changing loads, changing environment), continuous adaptation is necessary. In addition to multi-functionality, robust feedback loops that result in smart signaling chains allow decentralized self-organization. Consequently, in HBM level-specific properties (H_i) do not only fulfill the required functions, but also initiate the decentralized processes on the next hierarchical level (H_{i+1}), and thus generate the structures on this level (H_{i+1}).

-integration and -interaction must be defined. Computational modeling techniques have progressed enormously during the last few years, and simulation techniques like molecular dynamics find broad application and increasing acceptance. But these simulation approaches are still limited to samples of a few nanometers in size and modeling techniques, linking atomistic to continuum scale in biological materials, which lack a regular atomic lattice, are in a very early stage of development. To overcome these limitations, new numerical models will be necessary, followed by new approaches of data analysis and visualization methods.

In addition to the computational techniques, experimental techniques on the level of individual molecules progressed immensely during the last decade (AFM, optical or magnetic tweezers, nanoindentation, micromanipulators). Beyond experimental challenges, there exist many manufacturing challenges that must be overcome, such as the application of recombinant DNA techniques to sustain industrial volumes and large-scale mass production, or the construction of macro-materials from nano-devices.

6.3. Concluding remarks

Overcoming the challenges associated with developing *de novo* materials will require a convergence of scientific disciplines in two regards. First, experimental, theoretical and computational approaches will need to be combined quantitatively, in order to understand and explain observed phenomena that are found in the biological nano-world and in order to successfully apply the knowledge gained to technological developments. Assessment and minimization of the multi-faceted risks associated with nano- and bio-technologies will be a vital part of this process. Second, different disciplines like Physics, Chemistry, Biology, Engineering, Computer Science and Medicine will have to be integrated, as each of these fields is indispensable for successful pursuit of the translation of biological nanoscience into new technologies (see also Fig. 5).

Humans have first exploited natural materials, such as stone, wood and clay. Later, with the advent of Bronze and Iron Ages, metallurgy and synthetic materials have become more dominant. Following the Industrial Revolution, mass production of synthetic materials through energy intensive manufacturing processes became the norm. However, developing public awareness of the adverse impact of many large-scale industrial processes on the environment and the depletion of natural energy resources in line with ever-growing global energy demand caused a paradigm shift in the field, setting the trend back to exploring natural methods in materials design and synthesis a develop sustainable and energy efficient structural materials [328].

The fundamental interest in biological materials and structures is intimately related to recent advances in nanoscience and nanotechnology; and these fields blend together in an interdisciplinary framework to become a conspicuous starting point for new directions of materials research towards the development of new materials and structures that can meet future technical and social challenges. Over the past few centuries, scientists and engineers have developed a good understanding of how to create complex synthetic structures out of a diverse range of constituents, at various scales (machines, buildings, airplanes, nuclear reactors and many others). Increased development and research funding into these areas of research will lead to breakthroughs not only on the fundamental sciences, but also in technological applications. Research in the area of mechanics of biological materials will extend our ability to carry out structural engineering, as used for buildings or bridges today, to the ultimate scale – nano-scale, and may be a vital component of the realization of nanotechnology.

A better understanding of the mechanics of biological and natural materials, integrated within complex technological systems could make it possible to combine living and non-living environments to develop sustainable technologies. New materials technologies such as protein-based materials produced by recombinant DNA techniques represent new frontiers in materials design and synthesis [255,258]. These questions have potentially very high impact in the understanding and design of environmentally friendly technologies and may enhance the quality of life of millions of people, through advances in the medical sciences as well as through improvements of the living environment. A currently pressing question is the development of new technologies to address challenges associated with providing energy at an ever increasing demand. The development of new materials and processes could be a key element in solving some of these issues (e.g. for fuel cells, biofuels, advanced biotech-

nologies for fuel processing). Technological advances may also be possible by utilization of bacteria to produce and process fuel from crops, or by enabling the synthesis of materials at reduced processing temperature.

Nanoscience and nanotechnology enable us to make structures at the ultimate scale (self-assembly, recombinant DNA, utilization of motor proteins for nano-machines and others). This will perhaps lead to novel complex structural materials, designed from nano to macro. The theoretical progress in understanding hierarchical biological materials will facilitate to use an extended physical space, through the use of multiple hierarchies, in an efficient and controlled manner, that is, lead to a bottom-up structural design on the sub-macroscopic scale, instead of empirical approaches. For example, the extended design space might serve as mean to realize new physical realities that are not accessible to a single scale, such as material synthesis at moderate temperatures, or fault tolerant hierarchical assembly pathways [329], which enable biological systems to overcome the limitations to particular chemical bonds (soft) and chemical elements (organic) present under natural conditions.

The improved understanding of hierarchical design laws through materiomics might further enable the development and application of new organic and organic–inorganic multi-featured composites (such as assemblies of carbon nanotubes and proteins or polymer–protein composites [251–253]), which will mainly consist of chemical elements that appear in our environment in a virtually unlimited amount. In the future, these materials might consequently help us to solve human's energy and resource problems (e.g. to overcome dependencies on fossil resources, iron, etc.), and allow us to manufacture nano-materials, which will be produced in the future by techniques like recombinant DNA [255–257] or peptide self-assembly [258–260], techniques where the borders between materials, structures and machines vanish.

Applications of these new materials and structures are new biomaterials, new polymers, new composites, engineered spider silk, new scaffolding tissues, improved understanding of cell–ECM interactions, cell mechanics, hierarchical structures and self-assembly. In addition to the long-term impact in biology, bioengineering and medicine, this research may eventually contribute to our theoretical understanding of how structural features at different scales interact with one another. In light of the extended physical design space that is available to hierarchical protein materials, materiomics may transform engineering approaches not only for materials applications, but also in manufacturing, transportation, or in the designs of communication networks.

Acknowledgements

References

[1] Alberts B et al. Molecular biology of the cell. Taylor & Francis; 2002.
[2] Astbury WT, Street A. X-ray studies of the structures of hair, wool and related fibres. I. General. Trans R Soc Lond A 1931;230:75–101.
[3] Weiner S, Wagner HD. The material bone: structure mechanical function relations. Annu Rev Mater Sci 1998;28:271–98.
[4] Currey JD. Bones: structure and mechanics. Princeton, NJ: Princeton University Press; 2002.
[5] Lakes R. Materials with structural hierarchy. Nature 1993;361(6412):511–5.
[6] Wegst UGK, Ashby MF. The mechanical efficiency of natural materials. Philos Mag 2004;84(21):2167–81.
[7] Vincent JFV. In: Structural biomaterials. Princeton, NJ: Princeton University Press; 1990. p. 244.
[8] Fratzl P, Weinkamer R. Nature's hierarchical materials. Prog Mater Sci 2007;52(8):1263–334.

[9] Aizenberg J et al. Skeleton of *Euplectella* sp.: structural hierarchy from the nanoscale to the macroscale. Science 2005;309(5732):275–8.
[10] Dao M, Lim CT, Suresh S. Mechanics of the human red blood cell deformed by optical tweezers. J Mech Phys Solids 2003;51(11–12):2259–80.
[11] Suresh S et al. Connections between single-cell biomechanics and human disease states: gastrointestinal cancer and malaria. Acta Biomater 2005;1(1):15–30.
[12] Bao G, Suresh S. Cell and molecular mechanics of biological materials. Nat Mater 2003;2(11):715–25.
[13] Lim CT et al. Large deformation of living cells using laser traps. Acta Mater 2004;52(7):1837–45.
[14] Taylor GI. Mechanism of plastic deformation in crystals. Proc R Soc A 1934;145:362.
[15] Hirth JP, Lothe J. Theory of dislocations. Wiley–Interscience; 1982.
[16] Broberg KB. Cracks and fracture. Academic Press; 1990.
[17] Buehler MJ. Nano- and micromechanical properties of hierarchical biological materials and tissues. J Mater Sci 2007;42(21):8765–70.
[18] Buehler MJ, Ackbarow T. Fracture mechanics of protein materials. Mater Today 2007;10(9):46–58.
[19] Wolf D et al. Deformation mechanism and Inverse Hall–Petch behavior in nanocrystalline materials. Z Metall 2003;94:1052–61.
[20] Yip S. The strongest size. Nature 1998;391:532–3.
[21] Blanckenhagen Bv, Gumbsch P, Arzt E. Dislocation sources in discrete dislocation simulations of thin film plasticity and the Hall–Petch relation. Model Simul Mater Sci Eng 2001;9:157–69.
[22] Nieh TG, Wadsworth J. Hall–Petch relation in nanocrystalline solids. Scripta Met 1991;25(4).
[23] Frost HJ, Ashby MF. Deformation–mechanism maps. Oxford: Pergamon Press; 1982.
[24] Fraser P, Bickmore W. Nuclear organization of the genome and the potential for gene regulation. Nature 2007;447(7143):413–7.
[25] Engler AJ et al. Matrix elasticity directs stem cell lineage specification. Cell 2006;126(4):677–89.
[26] Buehler MJ. Atomistic and continuum modeling of mechanical properties of collagen: elasticity, fracture and self-assembly. J Mater Res 2006;21(8):1947–61.
[27] Buehler MJ. Nature designs tough collagen: explaining the nanostructure of collagen fibrils. Proc Natl Acad Sci USA 2006;103(33):12285–90.
[28] Fratzl P et al. Structure and mechanical quality of the collagen-mineral nano-composite in bone. J Mater Chem 2004;14(14):2115–23.
[29] An KN, Sun YL, Luo ZP. Flexibility of type I collagen and mechanical property of connective tissue. Biorheology 2004;41(3–4):239–46.
[30] Ramachandran GN, Kartha G. Structure of collagen. Nature 1955;176:593–5.
[31] Doyle J. Rules of engagement. Nature 2007;446:860.
[32] Kitano H. Computational systems biology. Nature 2002;420(6912):206–10.
[33] Kitano H. Systems biology: a brief overview. Science 2002;295(5560):1662–4.
[34] Smith BL et al. Molecular mechanistic origin of the toughness of natural adhesives, fibres and composites. Nature 1999;399(6738):761–3.
[35] Prater CB, Butt HJ, Hansma PK. Atomic force microscopy. Nature 1990;345(6278):839–40.
[36] Sun YL et al. Stretching type II collagen with optical tweezers. J Biomech 2004;37(11):1665–9.
[37] Tai K, Ulm FJ, Ortiz C. Nanogranular origins of the strength of bone. NanoLetters 2006;11:2520–5.
[38] Lim CT et al. Experimental techniques for single cell and single molecule biomechanics. Mater Sci Eng C: Biomim Supramol Syst 2006;26(8):1278–88.
[39] Goddard WA. A perspective of materials modeling. In: Yip S, editor. Handbook of materials modeling. Springer; 2006.
[40] Ackbarow T et al. Hierarchies, multiple energy barriers and robustness govern the fracture mechanics of alpha-helical and beta-sheet protein domains. Proc Natl Acad Sci USA 2007;104:16410–5.
[41] Ackbarow T, Buehler MJ. Hierarchical coexistence of universality and diversity controls robustness and multi-functionality in protein materials. J Comput Theor Nanos 2008;5(7):1193–204.
[42] Bell GI. Models for specific adhesion of cells to cells. Science 1978;200(4342):618–27.
[43] Bustamante C et al. Entropic elasticity of lambda-phage DNA. Science 1994;265(5178):1599–600.
[44] Bond PJ et al. Coarse-grained molecular dynamics simulations of membrane proteins and peptides. J Struct Biol 2007;157(3):593–605.
[45] Vashishta P, Kalia RK, Nakano A. Large-scale atomistic simulations of dynamic fracture. Comp Sci Eng 1999:56–65.
[46] Rountree CL et al. Atomistic aspects of crack propagation in brittle materials: multimillion atom molecular dynamics simulations. Annu Rev Mater Res 2002;32:377–400.
[47] Buehler MJ, Gao H. Dynamical fracture instabilities due to local hyperelasticity at crack tips. Nature 2006;439(7074):307–10.
[48] Buehler MJ, Abraham FF, Gao H. Hyperelasticity governs dynamic fracture at a critical length scale. Nature 2003;426:141–6.
[49] Buehler MJ, Gao H. Ultra large scale atomistic simulations of dynamic fracture. In: Schommers W, Rieth A, editors. Handbook of theoretical and computational nanotechnology. American Scientific Publishers (ASP); 2006.
[50] Buehler MJ, van Duin ACT, Goddard WA. Multi-paradigm modeling of dynamical crack propagation in silicon using the ReaxFF reactive force field. Phys Rev Lett 2006;96(9):095505.
[51] Buehler MJ et al. Threshold crack speed controls dynamical fracture of silicon single crystals. Phys Rev Lett 2007;99:165502.
[52] Wang W et al. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein–ligand, protein–protein, and protein–nucleic acid noncovalent interactions. Annu Rev Biophys Biomol Struct 2001;30:211–43.
[53] Mackerell AD. Empirical force fields for biological macromolecules: overview and issues. J Comput Chem 2004;25(13):1584–604.

[54] Phillips JC et al. Scalable molecular dynamics with NAMD. J Comput Chem 2005;26(16):1781–802.
[55] Nelson MT et al. NAMD: a parallel, object oriented molecular dynamics program. Int J Supercomput Appl High Perform Comput 1996;10(4):251–68.
[56] MacKerell AD et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102(18):3586–616.
[57] Anderson D. Collagen self-assembly: a complementary experimental and theoretical perspective. Toronto, Canada: University of Toronto; 2005.
[58] Mayo SL, Olafson BD, Goddard WA. Dreiding – a generic force-field for molecular simulations. J Phys Chem 1990;94(26):8897–909.
[59] Rappe AK et al. UFF, a full periodic-table force-field for molecular mechanics and molecular-dynamics simulations. J Am Chem Soc 1992;114(25):10024–35.
[60] Pearlman DA et al. AMBER, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Commun 1995;91(1–3):1–41.
[61] van Duin ACT et al. ReaxFF: a reactive force field for hydrocarbons. J Phys Chem A 2001;105:9396–409.
[62] Brenner DW et al. A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons. J Phys: Condens Matter 2002;14(4):783–802.
[63] Stuart SJ, Tutein AB, Harrison JA. A reactive potential for hydrocarbons with intermolecular interactions. J Chem Phys 2000;112(14):6472–86.
[64] Strachan A et al. Shock waves in high-energy materials: the initial chemical events in nitramine RDX. Phys Rev Lett 2003;91(9).
[65] Nielson KD et al. Development of the ReaxFF reactive force field for describing transition metal catalyzed reactions, with application to the initial stages of the catalytic formation of carbon nanotubes. J Phys Chem A 2005;109:49.
[66] van Duin ACT et al. ReaxFF SiO: reactive force field for silicon and silicon oxide systems. J Phys Chem A 2003;107:3803–11.
[67] Han SS et al. Optimization and application of lithium parameters for the reactive force field, ReaxFF. J Phys Chem A 2005;109(20):4575–82.
[68] Chenoweth K et al. Simulations on the thermal decomposition of a poly(dimethylsiloxane) polymer using the ReaxFF reactive force field. J Am Chem Soc 2005;127(19):7192–202.
[69] Strachan A et al. Thermal decomposition of RDX from reactive molecular dynamics. J Chem Phys 2005;122(5).
[70] Cheung S et al. ReaxFF(MgH) reactive force field for magnesium hydride systems. J Phys Chem A 2005;109(5):851–9.
[71] van Duin ACT et al. Application of ReaxFF reactive force fields to transition metal catalyzed nanotube formation. Abstr Papers Am Chem Soc 2004;227:U1031.
[72] Buehler MJ. Hierarchical chemo-nanomechanics of stretching protein molecules: entropic elasticity, protein unfolding and molecular fracture. J Mech Mater Struct 2007;2(6):1019–57.
[73] Datta D, van Duin ACT, Goddard WA. Extending ReaxFF to biomacromolecules. Unpublished; 2005.
[74] Buehler MJ et al. The computational materials design facility (CMDF): a powerful framework for multiparadigm multi-scale simulations. Mater Res Soc Proc 2006;894:LL3.8.
[75] Lu H et al. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. Biophys J 1998;75(2):662–71.
[76] Gropp W, Lusk W, Skjellum A. Using MPI. MIT Press; 1999.
[77] Kadau K, Germann TC, Lomdahl PS. Large-scale molecular-dynamics simulation of 19 billion particles. Int J Mod Phys C 2004;15:193.
[78] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14(1):33.
[79] Tsai DH. Virial theorem and stress calculation in molecular-dynamics. J Chem Phys 1979;70(3):1375–82.
[80] Zimmerman JA et al. Calculation of stress in atomistic simulation. Model Simul Mater Sci Eng 2004;12:S319–32.
[81] Marko JF, Siggia ED. Stretching DNA. Macromolecules 1995;28(26):8759–70.
[82] Bozec L et al. Atomic force microscopy of collagen structure in bone and dentine revealed by osteoclastic resorption. Ultramicroscopy 2005;105(1–4):79–89.
[83] Guzman C et al. Exploring the mechanical properties of single vimentin intermediate filaments by atomic force microscopy. J Mol Biol 2006;360(3):623–30.
[84] Yuan CB et al. Energy landscape of streptavidin–biotin complexes measured by atomic force microscopy. Biochemistry 2000;39(33):10219–23.
[85] Sun YL, Luo ZP, An KN. Stretching short biopolymers using optical tweezers. Biochem Biophys Res Commun 2001;286(4):826–30.
[86] Bernstein FC et al. Protein data bank – computer-based archival file for macromolecular structures. J Mol Biol 1977;112(3):535–42.
[87] Nalla RK et al. Mechanistic aspects of fracture and R-curve behavior in human cortical bone. Biomaterials 2005;26(2):217–31.
[88] Ritchie RO et al. Characteristic dimensions and the micro-mechanisms of fracture and fatigue in 'nano' and 'bio' materials. Int J Fract 2004;128(1–4):1–15.
[89] Hulmes DJS et al. Radial packing, order, and disorder in collagen fibrils. Biophys J 1995;68(5):1661–70.
[90] Sasaki N, Odajima S. Elongation mechanism of collagen fibrils and force–strain relations of tendon at each level of structural hierarchy. J Biomech 1996;29(9):1131–6.
[91] Orgel JPRO et al. Microfibrillar structure of type I collagen in situ. Proc Natl Acad Sci USA 1995;103(24):9001–5.
[92] Puxkandl R et al. Viscoelastic properties of collagen: synchrotron radiation investigations and structural model. Philos Trans R Soc Lond Ser B: Biol Sci 2002;357(1418):191–7.
[93] Buehler MJ. Atomistic Modeling of Materials Failure. Springer; 2008.
[94] Zervakis M, Gkoumplias V, Tzaphlidou M. Analysis of fibrous proteins from electron microscopy images. Med Eng Phys 2005;27(8):655–67.

[95] Layton BE et al. Nanomanipulation and aggregation limitations of self-assembling structural proteins. Microelectron J 2005;36(7):644–9.
[96] Bozec L, Horton M. Topography and mechanical properties of single molecules of type I collagen using atomic force microscopy. Biophys J 2005;88(6):4223–31.
[97] Sun YL et al. Direct quantification of the flexibility of type I collagen monomer. Biochem Biophys Res Commun 2002;295(2):382–6.
[98] Arnoux PJ et al. Numerical damage models using a structural approach: application in bones and ligaments. Eur Phys J: Appl Phys 2002;17(1):65–73.
[99] Waite JH, Qin XX, Coyne KJ. The peculiar collagens of mussel byssus. Matrix Biol 1998;17(2):93–106.
[100] Kramer RZ et al. Staggered molecular packing in crystals of a collagen-like peptide with a single charged pair. J Mol Biol 2000;301(5):1191–205.
[101] Lorenzo AC, Caffarena ER. Elastic properties, Young's modulus determination and structural stability of the tropocollagen molecule: a computational study by steered molecular dynamics. J Biomech 2005;38(7):1527–33.
[102] Persikov AV et al. Electrostatic interactions involving lysine make major contributions to collagen triple-helix stability. Biochemistry 2005;44(5):1414–22.
[103] Israelowitz M et al. Computational modeling of type I collagen fibers to determine the extracellular matrix structure of connective tissues. Protein Eng Des Select 2005;18(7):329–35.
[104] Mooney SD, Klein TE. Structural models of osteogenesis imperfecta-associated variants in the COL1A1 gene. Mol Cell Proteom 2002;1(11):868–75.
[105] Mooney SD, Kollman PA, Klein TE. Conformational preferences of substituted prolines in the collagen triple helix. Biopolymers 2002;64(2):63–71.
[106] Mooney SD et al. Computed free energy differences between point mutations in a collagen-like peptide. Biopolymers 2001;58(3):347–53.
[107] Bischoff JE, Arruda EM, Grosh K. Finite element modeling of human skin using an isotropic, nonlinear elastic constitutive model. J Biomech 2000;33(6):645–52.
[108] Freeman JW, Silver FH. Elastic energy storage in unmineralized and mineralized extracellular matrices (ECMs): a comparison between molecular modeling and experimental measurements. J Theor Biol 2004;229(3):371–81.
[109] Allen MP, Tildesley DJ. Computer simulation of liquids. Oxford University Press; 1989.
[110] Gautieri A, Buehler MJ, Redaelli A. Deformation rate controls elasticity and unfolding pathway of single tropocollagen molecules. J Mech Behav Biomed Mater 2008. doi:10.1016/j.jmbbm.2008.03.001.
[111] Buehler MJ, Wong SY. Entropic elasticity controls nanomechanics of single tropocollagen molecules. Biophys J 2007;93(1):37–43.
[112] Ercolessi F, Adams JB. Interatomic potentials from 1st principle-calculations – the force matching method. Europhys Lett 1994;28(8):583–8.
[113] Lodish HB Arnold, Zipursky S Lawrence, Matsudaira Paul, Baltimore David, Darnell James E. Molecular cell biology. New York: WH Freeman & Co.; 1999.
[114] Bailey AJ. Molecular mechanisms of ageing in connective tissues. Mech Ageing Dev 2001;122(7):735–55.
[115] Robins SP, Bailey AJ. The chemistry of the collagen cross-links. Biochem J 1973;135:657–65.
[116] Lantz MA et al. Quantitative measurement of short-range chemical bonding forces. Science 2001;291(5513):2580–3.
[117] Grandbois M et al. How strong is a covalent bond? Science 1999;283(5408):1727–30.
[118] Eppell SJ et al. Nano measurements with micro-devices: mechanical properties of hydrated collagen fibrils. J R Soc Interf 2006;3(6):117–21.
[119] Ritchie RO et al. Fracture and ageing in bone: toughness and structural characterization. Strain 2006;42(4):225–32.
[120] Nalla RK et al. Fracture in human cortical bone: local fracture criteria and toughening mechanisms. J Biomech 2005;38(7):1517–25.
[121] Buehler MJ. Nanomechanics of collagen fibrils under varying cross-link densities: atomistic and continuum studies. J Mech Behav Biomed Mater 2008;1(1):1–2. doi:10.1016/j.jmbbm.2007.04.001.
[122] Hellan K. Introduction to fracture mechanics. McGraw-Hill Inc.; 1984.
[123] Griffith AA. The phenomenon of rupture and flows in solids. Philos Trans R Soc A 1920;221:163–98.
[124] Gupta HS et al. Nanoscale deformation mechanisms in bone. NanoLetters 2005;5(10):2108–11.
[125] Jager I, Fratzl P. Mineralized collagen fibrils: a mechanical model with a staggered arrangement of mineral particles. Biophys J 2000;79(4):1737–46.
[126] Peterlik H et al. From brittle to ductile fracture of bone. Nat Mater 2006;5(1):52–5.
[127] Nalla RK, Kinney JH, Ritchie RO. Effect of orientation on the in vitro fracture toughness of dentin: the role of toughening mechanisms. Biomaterials 2003;24(22):3955–68.
[128] Lichtens Jr et al. Defect in conversion of procollagen to collagen in a form of Ehlers-Danlos syndrome. Science 1973;182(4109):298–300.
[129] Glorieux FH. Caffey disease: an unlikely collagenopathy. J Clin Invest 2005;115(5):1142–4.
[130] Bhattacharjee A, Bansal M. Collagen structure: the Madras triple helix and the current scenario. IUBMB Life 2005;57(3):161–72.
[131] Borel JP, Monboisse JC. Collagens – why such a complicated structure. Comptes Rendus Des Seances De La Societe De Biologie Et De Ses Filiales 1993;187(2):124–42.
[132] Buehler MJ. Molecular architecture of collagen fibrils: a critical length scale for tough fibrils. Curr Appl Phys 2008;8(3–4):440–2. doi:10.1016/j.cap.2007.10.058.
[133] Buehler MJ. Hierarchical Nanomechanics of Collagen Fibrils: Atomistic and Molecular Modeling. In: Fratzl P, editor. Collagen: Structure and Mechanics. Springer; 2008.
[134] Borsato KS, Sasaki N. Measurement of partition of stress between mineral and collagen phases in bone using X-ray diffraction techniques. J Biomech 1997;30(9):955–7.
[135] Miles CA, Bailey AJ. Thermally labile domains in the collagen molecule. Micron 2001;32(3):325–32.

[138] Ackbarow T, Buehler MJ. Superelasticity, energy dissipation and strain hardening of vimentin coiled-coil intermediate filaments: atomistic and continuum studies. J Mater Sci 2007;42(21):8771–87.
[139] Buehler MJ. Molecular nanomechanics of nascent bone: fibrillar toughening by mineralization. Nanotechnology 2007;18:295102.
[140] Taylor D, Hazenberg JG, Lee TC. Living with cracks: damage and repair in human bone. Nat Mater 2007;6(4):263–6.
[141] Gupta HS et al. Cooperative deformation of mineral and collagen in bone at the nanoscale. Proc Natl Acad Sci USA 2006;103:17741–6.
[142] Landis WJ et al. Mineral and organic matrix interaction in normally calcifying tendon visualized in 3 dimensions by high-voltage electron-microscopic tomography and graphic image reconstruction. J Struct Biol 1993;110:39–54.
[143] Fantner GE et al. Sacrificial bonds and hidden length dissipate energy as mineralized fibrils separate during bone fracture. Nat Mater 2005;4(8):612–6.
[144] Thompson JB et al. Bone indentation recovery time correlates with bond reforming time. Nature 2001;414(6865):773–6.
[145] Gao HJ. Application of fracture mechanics concepts to hierarchical biomechanics of bone and bone-like materials. Int J Fract 2006;138(1–4):101–37.
[146] Nalla RK, Kinney JH, Ritchie RO. Mechanistic fracture criteria for the failure of human cortical bone. Nat Mater 2003;2(3):164–8.
[147] Hansma PK, Turner PJ, Ruoff RS. Optimized adhesives for strong, lightweight, damage-resistant, nanocomposite materials: new insights from natural materials. Nanotechnology 2007;18(4).
[148] Fantner GE et al. Hierarchical interconnections in the nano-composite material bone: fibrillar cross-links resist fracture on several length scales. Compos Sci Technol 2006;66(9):1205–11.
[149] Gupta HS et al. Synchrotron diffraction study of deformation mechanisms in mineralized tendon. Phys Rev Lett 2004;93(15).
[150] Plimpton S. Fast parallel algorithms for short-range molecular-dynamics. J Comput Phys 1995;117:1–19.
[151] Screen HRC et al. Local strain measurement within tendon. Strain 2004;40(4):157–63.
[152] Turner CH. Bone strength: current concepts. Skeletal development and remodeling in health, disease, and aging; 2006. p. 429–46.
[153] Gao H et al. Materials become insensitive to flaws at nanoscale: lessons from nature. Proc Natl Acad Sci USA 2003;100(10):5597–600.
[154] Buehler MJ et al. Cracking and adhesion at small scales: atomistic and continuum studies of flaw tolerant nanostructures. Model Simul Mater Sci Eng 2006;14:799–816.
[155] Zhu WH, Wu P. Surface energetics of hydroxyapatite: a DFT study. Chem Phys Lett 2004;396(1–3):38–42.
[157] Buehler MJ. Nanomechanics of collagen fibrils under varying cross-link densities: atomistic and continuum studies. J Mech Behav Biomed Mater 2007. doi:10.1016/j.jmbbm.2007.04.001.
[159] Alsberg E et al. Regulating bone formation via controlled scaffold degradation. J Dental Res 2003;82(11):903–8.
[160] Louis O et al. Cortical mineral-content of the radius assessed by peripheral qct predicts compressive strength on biomechanical testing. Bone 1995;16(3):375–9.
[161] Lotz JC, Gerhart TN, Hayes WC. Mechanical-properties of trabecular bone from the proximal femur – a quantitative CT study. J Comput Assist Tomogr 1990;14(1):107–14.
[162] Gruber M, Lupas AN. Historical review: another 50th anniversary – new periodicities in coiled coils. Trends Biochem Sci 2003;28(12):679–85.
[163] Moir RD, Spann TP. The structure and function of nuclear lamins: implications for disease. Cell Mol Life Sci 2001;58(12–13):1748–57.
[164] Wilson KL, Zastrow MS, Lee KK. Lamins and disease: insights into nuclear infrastructure. Cell 2001;104(5):647–50.
[165] Rose A, Meier I. Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. Cell Mol Life Sci 2004;61(16):1996–2009.
[166] Strelkov SV, Herrmann H, Aebi U. Molecular architecture of intermediate filaments. Bioessays 2003;25(3):243–51.
[167] Maccallum JL et al. Hydrophobic association of {alpha}-helices, steric dewetting, and enthalpic barriers to protein folding. Proc Natl Acad Sci USA 2007;104(15):6206–10.
[168] Burkhard P et al. The coiled-coil trigger site of the rod domain of cortexillin I unveils a distinct network of interhelical and intrahelical salt bridges. Struct Fold Des 2000;8(3):223–30.
[169] Wang N, Stamenovic D. Mechanics of vimentin intermediate filaments. J Muscle Res Cell Motil 2002;23(5–6):535–40.
[170] Mucke N et al. Assessing the flexibility of intermediate filaments by atomic force microscopy. J Mol Biol 2004;335(5):1241–50.
[171] Helfand BT, Chang L, Goldman RD. Intermediate filaments are dynamic and motile elements of cellular architecture. J Cell Sci 2004;117(2):133–41.
[172] Kreplak L, Fudge D. Biomechanical properties of intermediate filaments: from tissues to single filaments and back. Bioessays 2007;29(1):26–35.
[173] Smith TA et al. Modeling alpha-helical coiled-coil interactions: the axial and azimuthal alignment of 1B segments from vimentin intermediate filaments. Proteins Struct Funct Genet 2003;50(2):207–12.
[174] Kreplak L, Aebi U, Herrmann H. Molecular mechanisms underlying the assembly of intermediate filaments. Exp Cell Res 2004;301(1):77–83.
[175] Herrmann H, Aebi U. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular scaffolds. Annu Rev Biochem 2004;73:749–89.
[176] Strelkov SV et al. Crystal structure of the human lamin a coil 2B dimer: implications for the head-to-tail association of nuclear lamins. J Mol Biol 2004;343(4):1067–80.
[177] Fudge DS et al. The mechanical properties of hydrated intermediate filaments: insights from hagfish slime threads. Biophys J 2003;85(3):2015–27.
[178] Fudge DS, Gosline JM. Molecular design of the alpha-keratin composite: insights from a matrix-free model, hagfish slime threads. Proc R Soc Lond Ser B: Biol Sci 2004;271(1536):291–9.
[179] Coulombe PA et al. The 'ins' and 'outs' of intermediate filament organization. Trends Cell Biol 2000;10(10):420–8.

[180] Janmey PA et al. Viscoelastic properties of vimentin compared with other filamentous biopolymer networks. J Cell Biol 1991;113(1):155–60.
[181] Kreplak L et al. Exploring the mechanical behavior of single intermediate filaments. J Mol Biol 2005;354(3):569–77.
[182] Kiss B, Karsai A, Kellermayer MSZ. Nanomechanical properties of desmin intermediate filaments. J Struct Biol 2006;155(2):327–39.
[183] Omary MB, Coulombe PA, McLean WHI. Mechanisms of disease: intermediate filament proteins and their associated diseases. New Engl J Med 2004;351(20):2087–100.
[184] Schietke R et al. Mutations in vimentin disrupt the cytoskeleton in fibroblasts and delay execution of apoptosis. Eur J Cell Biol 2006;85(1):1–10.
[185] Mucke N et al. Molecular and biophysical characterization of assembly-starter units of human vimentin. J Mol Biol 2004;340(1):97–114.
[186] McLachlan AD, Karn J. Periodic features in the amino-acid-sequence of nematode myosin rod. J Mol Biol 1983;164(4):605–26.
[187] Brown JH, Cohen C, Parry DAD. Heptad breaks in alpha-helical coiled coils: stutters and stammers. Proteins Struct Funct Genet 1996;26(2):134–45.
[188] Strelkov SV et al. Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. EMBO J 2002;21(6):1255–66.
[189] Herrmann H, Aebi U. Intermediate filament assembly: temperature sensitivity and polymorphism. Cell Mol Life Sci 1999;55(11):1416–31.
[190] Parry DAD. Fibrinogen – preliminary-analysis of amino-acid sequences of portions of alpha-chains, beta-chains and gamma-chains postulated to form interdomainal link between globular regions of molecule. J Mol Biol 1978;120(4):545–51.
[191] Warshel A, Papazyan A. Energy considerations show that low-barrier hydrogen bonds do not offer a catalytic advantage over ordinary hydrogen bonds. PNAS 1996;93(24):13665–70.
[192] Kramers HA. Brownian motion in a field of force and the diffusion model of chemical reactions. Physica 1940;7:10.
[193] Evans EA, Calderwood DA. Forces and bond dynamics in cell adhesion. Science 2007;316(5828):1148–53.
[194] Bayas MV et al. Lifetime measurements reveal kinetic differences between homophilic cadherin bonds. Biophys J 2006;90(4):1385–95.
[195] Evans E. Mechanical switching and cooperative coupling of unbinding pathways in bioadhesion bonds. Abstr Papers Am Chem Soc 2004;227:U469–70.
[196] Evans E. Probing the relation between force – lifetime – and chemistry in single molecular bonds. Annu Rev Biophys Biomol Struct 2001;30:105–28.
[197] Evans E et al. Chemically distinct transition states govern rapid dissociation of single L-selectin bonds under force. Proc Natl Acad Sci USA 2001;98(7):3784–9.
[198] Evans EB. Looking inside molecular bonds at biological interfaces with dynamic force spectroscopy. Biophys Chem 1999;82(2–3):83–97.
[199] Merkel R et al. Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. Nature (London) 1999;379(6714):50–3.
[200] Evans E, Ritchie K. Dynamic strength of molecular adhesion bonds. Biophys 1997;72(4):1541–55.
[201] Dudko OK, Hummer G, Szabo A. Intrinsic rates and activation free energies from single-molecule pulling experiments. Phys Rev Lett 2006;96(10).
[202] Dudko OK et al. Extracting kinetics from single-molecule force spectroscopy: nanopore unzipping of DNA hairpins. Biophys J 2007;92(12):4188–95.
[203] Hummer G, Szabo A. Kinetics from nonequilibrium single-molecule pulling experiments. Biophys J 2003;85(1):5–15.
[204] Hummer G, Szabo A. Free energy surfaces from single-molecule force spectroscopy. Acc Chem Res 2005;38(7):504–13.
[205] Hummer G, Szabo A. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. Proc Natl Acad Sci USA 2001;98(7):3658–61.
[206] Jarzynski C. Equilibrium free-energy differences from nonequilibrium measurements: a master-equation approach. Phys Rev E 1997;56(5):5018–35.
[207] Jarzynski C. Nonequilibrium equality for free energy differences. Phys Rev Lett 1997;78(14):2690–3.
[208] Makarov DE. Unraveling individual molecules by mechanical forces: theory meets experiment. Biophys J 2007;92(12):4135–6.
[209] Li PC, Makarov DE. Theoretical studies of the mechanical unfolding of the muscle protein titin: bridging the time-scale gap between simulation and experiment. J Chem Phys 2003;119(17):9260–8.
[210] Schlierf M, Rief M. Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape. Biophys J 2006;90(4):L33–5.
[211] Gilli P et al. Covalent versus electrostatic nature of the strong hydrogen bond: discrimination among single, double, and asymmetric single-well hydrogen bonds by variable-temperature X-ray crystallographic methods in beta-diketone enol RAHB systems. J Am Chem Soc 2004;126(12):3845–55.
[212] Wiita AP et al. Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques. Proc Natl Acad Sci USA 2006;103(19):7222–7.
[213] Maloney CE, Lacks DJ. Energy barrier scalings in driven systems. Phys Rev E 2006;73(6).
[214] West DK, Olmsted PD, Paci E. Mechanical unfolding revisited through a simple but realistic model. J Chem Phys 2006;124(15).
[215] Bagni BCMA, Romano G, Cecchi G. Characterization of actomyosin bond properties in intact skeletal muscle by force spectroscopy. Proc Natl Acad Sci USA 2007;104(22):9284–9.
[216] Schwaiger I et al. The myosin coiled-coil is a truly elastic protein structure. Nat Mater 2002;1(4):232–5.
[217] Akkermans RLC, Warren PB. Multiscale modelling of human hair. Philos Trans R Soc Lond Ser A: Math Phys Eng Sci 2004;362(1821):1783–93.

[218] Root DD et al. Coiled-coil nanomechanics and uncoiling and unfolding of the superhelix and alpha-helices of myosin. Biophys J 2006;90(8):2852–66.
[219] Cieplak M, Hoang TX, Robbins MO. Thermal folding and mechanical unfolding pathways of protein secondary structures. Proteins Struct Funct Genet 2002;49(1):104–13.
[220] Rohs R, Etchebest C, Lavery R. Unraveling proteins: a molecular mechanics study. Biophys J 1999;76(5):2760–8.
[221] Bornschlogl T, Rief M. Single molecule unzipping of coiled coils: sequence resolved stability profiles. Phys Rev Lett 2006;96(11).
[222] Mitsui K et al. Dynamic measurement of single protein's mechanical properties. Biochem Biophys Res Commun 2000;272(1):55–63.
[223] Hanke F, Kreuzer HJ. Breaking bonds in the atomic force microscope: theory and analysis. Phys Rev E 2006;74(3).
[224] Wolgemuth CW, Sun X. Elasticity of alpha-helical coiled coils. Phys Rev Lett 2006;97(24):248101.
[225] Forman JR, Clarke J. Mechanical unfolding of proteins: insights into biology, structure and folding. Curr Opin Struct Biol 2007;17(1):58–66.
[226] Brockwell DJ. Force denaturation of proteins – an unfolding story. Curr Nanosci 2007;3(1):3–15.
[227] Randles LG, Rounsevell RWS, Clarke J. Spectrin domains lose cooperativity in forced unfolding. Biophys J 2007;92(2):571–7.
[228] Paramore S, Voth GA. Examining the influence of linkers and tertiary structure in the forced unfolding of multiple-repeat spectrin molecules. Biophys J 2006;91(9):3436–45.
[229] Ortiz V et al. Unfolding a linker between helical repeats. J Mol Biol 2005;349(3):638–47.
[230] Finke JM et al. Equilibrium unfolding of the poly(glutamic acid)(20) helix. Biopolymers 2007;86(3):193–211.
[231] Day R, Daggett V. Direct observation of microscopic reversibility in single-molecule protein folding. J Mol Biol 2007;366(2):677–86.
[232] Storm C et al. Nonlinear elasticity in biological gels. Nature 2005;435:191–4.
[233] Boudko SP et al. Design and crystal structure of bacteriophage T4 mini-fibritin NCCF. J Mol Biol 2004;339(4):927–35.
[234] Sotomayor M, Schulten K. Single-molecule experiments in vitro and in silico. Science 2007;316(5828):1144–8.
[235] Gao M, Lu H, Schulten K. Unfolding of titin domains studied by molecular dynamics simulations. J Muscle Res Cell Motil 2002;23(5–6):513–21.
[236] Rief M et al. Reversible unfolding of individual titin immunoglobulin domains by AFM. Science 1997;276(5315):1109–12.
[237] Marszalek PE et al. Mechanical unfolding intermediates in titin modules. Nature 1999;402(6757):100–3.
[238] Matthias Rief HG. Force spectroscopy of single biomolecules. ChemPhysChem 2002;3(3):255–61.
[239] Sheu S-Y et al. Energetics of hydrogen bonds in peptides. PNAS 2003;100(22):12683–7.
[240] Kageshima M et al. Insight into conformational changes of a single alpha-helix peptide molecule through stiffness measurements. Chem Phys Lett 2001;343(1–2):77–82.
[241] Lantz MA et al. Stretching the alpha-helix: a direct measure of the hydrogen-bond energy of a single-peptide molecule. Chem Phys Lett 1999;315(1–2):61–8.
[242] Lezon TR, Banavar JR, Maritan A. The origami of life. J Phys Condens Matter 2006;18(3):847–88.
[243] Pareto V. Manual of political economy. New York; 1909.
[244] Chen YS, Chong PP, Tong YG. Theoretical foundation of the 80/20 rule. Scientometrics 1993;28(2):183–204.
[245] Chen JCH, Chong PP, Chen YS. Decision criteria consolidation: a theoretical foundation of Pareto principle to Porter's competitive forces. J Org Comput Electron Commer 2001;11(1):1–14.
[246] Chen YS, Chong PP, Tong MY. Mathematical and computer modeling of the pareto principle. Math Comput Model 1994;19(9):61–80.
[247] Ackbarow T, Buehler MJ. Molecular mechanics of stutter defects in vimentin intermediate filaments. Exp Mech, in press. doi:10.1007/s11340-007-9100-6.
[248] Barthelat F et al. On the mechanics of mother-of-pearl: a key feature in the material hierarchical structure. J Mech Phys Solids 2007;55(2):306–37.
[249] Heslot H. Artificial fibrous proteins: a review. Biochimie 1998;80(1):19–31.
[250] Ball P. Made to measure: new materials for the 21st century. Princeton, NJ, USA: Princeton University Press; 1997.
[251] Cui XQ et al. Biocatalytic generation of ppy-enzyme-CNT nanocomposite: from network assembly to film growth. J Phys Chem C 2007;111(5):2025–31.
[252] Hule RA, Pochan DJ. Polymer nanocomposites for biomedical application. MRS Bull 2007;32(4):5.
[253] Winey KI, Vaia RA. Polymer nanocomposites. MRS Bull 2007;32(4):5.
[254] Papapostolou D et al. Engineering nanoscale order into a designed protein fiber. PNAS 2007:0700801104.
[255] Langer R, Tirrell DA. Designing materials for biology and medicine. Nature 2004;428(6982):487–92.
[256] Petka WA et al. Reversible hydrogels from self-assembling artificial proteins. Science 1998;281(5375):389–92.
[257] Smeenk JM et al. Controlled assembly of macromolecular beta-sheet fibrils. Angew Chem Int Ed 2005;44(13):1968–71.
[258] Zhao XJ, Zhang SG. Designer self-assembling peptide materials. Macromol Biosci 2007;7(1):13–22.
[259] Zhao XJ, Zhang SG. Molecular designer self-assembling peptides. Chem Soc Rev 2006;35(11):1105–10.
[260] Mershin A et al. A classic assembly of nanobiomaterials. Nat Biotechnol 2005;23(11):1379–80.
[261] Buehler MJ. Rupture mechanics of vimentin intermediate filament tetramers. J Eng Mech (ASCE), in press.
[262] Sheu SY et al. Energetics of hydrogen bonds in peptides. Proc Natl Acad Sci USA 2003;100(22):12683–7.
[263] Ackbarow T et al. Hierarchies, multiple energy barriers, and robustness govern the fracture mechanics of {alpha}-helical and beta-sheet protein domains. Proc Natl Acad Sci 2007:0705759104.
[264] Du N et al. Design of superior spider silk: from nanostructure to mechanical properties. Biophys J 2006;91(12):4528–35.
[265] Oroudjev E et al. Segmented nanofibers of spider dragline silk: atomic force microscopy and single-molecule force spectroscopy. Proc Natl Acad Sci USA 2002;99:6460–5.
[266] van Beek JD et al. The molecular structure of spider dragline silk: folding and orientation of the protein backbone. Proc Natl Acad Sci USA 2002;99(16):10266–71.
[267] Parkhe AD et al. Structural studies of spider silk proteins in the fiber. J Mol Recognit 1997;10(1):1–6.

[268] Vollrath F, Porter D. Spider silk as archetypal protein elastomer. Soft Matter 2006;2(5):377–85.
[269] Vollrath F, Porter D. Spider silk as a model biomaterial. Appl Phys A: Mater Sci Process 2006;82(2):205–12.
[270] Termonia Y. Molecular modeling of spider silk elasticity. Macromolecules 1994;27(25):7378–81.
[271] Carrion-Vazquez M et al. The mechanical stability of ubiquitin is linkage dependent. Nat Struct Biol 2003;10(9):738–43.
[272] Rief M et al. The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy. Biophys J 1998;75(6):3008–14.
[273] Lee EH et al. Mechanical strength of the titin Z1Z2–telethonin complex. Structure (London, England: 1993) 2006;14(3):497–509.
[274] Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem 2006;75:333–66.
[275] Dobson CM. Protein folding and misfolding. Nature 2003;426(6968):884–90.
[276] Zhang SG. Fabrication of novel biomaterials through molecular self-assembly. Nat Biotechnol 2003;21(10):1171–8.
[277] Mostaert AS et al. Nanoscale mechanical characterisation of amyloid fibrils discovered in a natural adhesive. J Biol Phys 2006;32(5):393–401.
[278] Cox DL et al. The materials science of protein aggregation. MRS Bull 2005;30(6):452–7.
[279] Vollrath F, Knight DP. Liquid crystalline spinning of spider silk. Nature 2001;410(6828):541–8.
[280] Brockwell DJ et al. Pulling geometry defines the mechanical resistance of a beta-sheet protein. Nat Struct Biol 2003;10(9):731–7.
[281] Grater F et al. Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations. Biophys J 2005;88(2):790–804.
[282] West DK et al. Mechanical resistance of proteins explained using simple molecular models. Biophys J 2006;90(1):287–97.
[283] Lu H, Schulten K. The key event in force-induced unfolding of titin's immunoglobulin domains. Biophys J 2000;79(1):51–65.
[284] Penel S et al. Length preferences and periodicity in beta-strands. Antiparallel edge beta-sheets are more likely to finish in non-hydrogen bonded rings. Protein Eng 2003;16(12):957–61.
[285] Buehler MJ, Keten S. Elasticity, strength and resilience: a comparative study on mechanical signatures of α-helix, β-sheet and tropocollagen domains. Nano Research 2008;1(1):63–71.
[286] Alberts B. Molecular biology of the cell, fourth ed., vol. xxxiv. New York: Garland Science; 2002. 1463, 86 pp..
[287] Rief M et al. Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles. J Mol Biol 1999;286(2):553–61.
[288] Oberhauser AF et al. The molecular elasticity of the extracellular matrix protein tenascin. Nature 1998;393(6681):181–5.
[289] Voet D, Voet JG. In: Biochemistry. New York: John Wiley and Sons; 2004. p. 1591.
[290] Gao M et al. Structure and functional significance of mechanically unfolded fibronectin type III1 intermediates. Proc Natl Acad Sci USA 2003;100(25):14784–9.
[291] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14(1):33.
[292] Merkel R et al. Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. Nature 1999;397(6714):50–3.
[293] Erdmann T, Schwarz US. Stability of adhesion clusters under constant force. Phys Rev Lett 2004;92(10).
[294] Hanke F, Kreuzer HJ. Breaking bonds in the atomic force microscope: theory and analysis. Phys Rev E 2006;74(3).
[295] Heymann B, Grubmuller H. Dynamic force spectroscopy of molecular adhesion bonds. Phys Rev Lett 2000;84(26):6126–9.
[296] Seifert U. Rupture of multiple parallel molecular bonds under dynamic loading. Phys Rev Lett 2000;84(12):2750–3.
[297] Sulchek T, Friddle RW, Noy A. Strength of multiple parallel biological bonds. Biophys J 2006;90(12):4686–91.
[298] Sulkowska JI, Cieplak M. Mechanical stretching of proteins – a theoretical survey of the Protein Data Bank. J Phys Condens Matter 2007;19(28).
[299] Keten S, Buehler MJ. Asymptotic strength limit of hydrogen bond assemblies in proteins at vanishing pulling rates. Phys Rev Lett 2008;100(19):198301.
[300] Keten S, Buehler MJ. Geometric confinement governs the rupture strength of H-bond assemblies at a critical length scale. Nano Letters 2008;8(2):743–8.
[301] Griffith AA. Philos Trans R Soc Lond Ser A 1921(221):163–8.
[302] Fisher TE et al. The study of protein mechanics with the atomic force microscope. Trends Biochem Sci 1999;24(10):379–84.
[303] Rief M, Fernandez JM, Gaub HE. Elastically coupled two-level systems as a model for biopolymer extensibility. Phys Rev Lett 1998;81(21):4764–7.
[304] Bustamante C et al. Single-molecule studies of DNA mechanics. Curr Opin Struct Biol 2000;10(3):279–85.
[305] Keten S, Buehler MJ. Large deformation and fracture mechanics of a beta-helical protein nanotube: atomistic and continuum modeling. Comput Method Appl Mech Eng, 2008. doi:10.1016/j.cma.2007.11.028.
[306] Yoder MD, Lietzke SE, Jurnak F. Unusual structural features in the parallel beta-helix in pectate lyases. Structure 1993;1(4):241–51.
[307] Govaerts C et al. Evidence for assembly of prions with left-handed beta 3-helices into trimers. Proc Natl Acad Sci USA 2004;101(22):8342–7.
[308] Kishimoto A et al. beta-Helix is a likely core structure of yeast prion Sup35 amyloid fibers. Biochem Biophys Res Commun 2004;315(3):739–45.
[309] Perutz MF et al. Amyloid fibers are water-filled nanotubes. Proc Natl Acad Sci USA 2002;99(8):5591–5.
[310] Kisilevsky R. Review: amyloidogenesis-unquestioned answers and unanswered questions. J Struct Biol 2000;130(2–3):99–108.
[311] Reches M, Gazit E. Casting metal nanowires within discrete self-assembled peptide nanotubes. Science 2003;300(5619):625–7.
[312] Scheibel T et al. Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. Proc Natl Acad Sci USA 2003;100(8):4527–32.
[313] Kenney JM et al. Amyloidogenic nature of spider silk. Eur J Biochem 2002;269(16):4159–63.

[314] Slotta U et al. Spider silk and amyloid fibrils: a structural comparison. Macromol Biosci 2007;7(2):183–8.
[315] Fukuma T, Mostaert AS, Jarvis SP. Explanation for the mechanical strength of amyloid fibrils. Tribol Lett 2006;22(3):233–7.
[316] Kellermayer MSZ et al. Reversible mechanical unzipping of amyloid beta-fibrils. J Biol Chem 2005;280(9):8464–70.
[317] Brändén C-I, Tooze J. Introduction to protein structure, second ed., vol. xiv. New York: Garland Pub.; 1999. 410 pp..
[318] Timoshenko S, Gere J. Theory of elastic stability. New York, NY, USA: McGraw-Hill; 1988.
[319] Boal D. Mechanics of the cell. Cambridge, UK: Cambridge University Press; 2001.
[320] Buehler MJ, Kong Y, Gao HJ. Deformation mechanisms of very long single-wall carbon nanotubes subject to compressive loading. J Eng Mater Technol 2004;126(3):245–9.
[321] Urry DW, Parker TM. Mechanics of elastin: molecular mechanism of biological elasticity and its relationship to contraction. J Muscle Res Cell Motil 2002;23(5–6):543–59.
[322] Hayashi CY, Lewis RV. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. J Mol Biol 1998;275(5):773–84.
[323] Li PC, Makarov DE. Simulation of the mechanical unfolding of ubiquitin: probing different unfolding reaction coordinates by changing the pulling geometry. J Chem Phys 2004;121(10):4826–32.
[324] Kanamaru S et al. Structure of the cell-puncturing device of bacteriophage T4. Nature 2002;415(6871):553–7.
[325] Janmey PA, Leterrier JF, Herrmann H. Assembly and structure of neurofilaments. Curr Opin Colloid Interf Sci 2003;8(1):40–7.
[326] Kerkam K et al. Liquid crystallinity of natural silk secretions. Nature 1991;349(6310):596–8.
[327] Mandelbrot BB. Fractal geometry – what is it, and what does it do. Proc R Soc Lond Ser A: Math Phys Eng Sci 1989;423(1864):3–16.
[328] Bruck HA, Evans JJ, Peterson ML. The role of mechanics in biological and biologically inspired materials. Exp Mech 2002;42(4):361–71.
[329] Holland JH. Hidden order – how adaptation builds complexity. Reading, MA: Helix Books; 1995.
[330] Vesentini S et al. Molecular assessment of the elastic properties of collagen-like homotrimer sequences. Biomech Model Mechanobiol 2005;3(4):224–34.
[331] Harley R et al. Phonons and elastic-moduli of collagen and muscle. Nature 1977;267(5608):285–7.
[332] Cusack S, Miller A. Determination of the elastic-constants of collagen by brillouin light-scattering. J Mol Biol 1979;135(1):39–51.
[333] Hofmann H et al. Localization of flexible sites in thread-like molecules from electron-micrographs – comparison of interstitial, basement-membrane and intima collagens. J Mol Biol 1984;172(3):325–43.
[334] van der Rijt JAJ et al. Micromechanical testing of individual collagen fibrils. Macromol Biosci 2006;6(9):697–702.
[335] Buehler MJ. Atomistic and continuum modeling of mechanical properties of collagen: elasticity, fracture, and self-assembly. J Mater Res 2006;21(8):1947–61.
[336] Buehler MJ. Mesoscale modeling of mechanics of carbon nanotubes: self-assembly, self-folding, and fracture. J Mater Res 2006;21(11):2855–69.
[337] Rainey JK, Wen CK, Goh MC. Hierarchical assembly and the onset of banding in fibrous long spacing collagennext term revealed by atomic force microscopy. Matrix Biol 2002;21(8):647–60.