Probing the Mechanics of Protein Materials with Molecular Dynamics Simulations

Dissertation submitted to
the Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of Doctor of Natural Sciences

Presented by: Senbo Xiao, *M.Sc.*
Born in: Guangdong, China

Oral-examination: July 26^{th}, 2011

Referees:

Prof. Dr. Roland Eils

Prof. Dr. Jörg Langowski

Abstract

Mechanical force is one important type of biomechanical signal, the specific role of which, however, remains to be explored. The mechanism of force signaling within a biomolecule is lacking a comprehensive understanding. More importantly, how biomolecules such as proteins withstand high mechanical forces, thereby demonstrating their delicate mechanical properties, is a missing page, yet is the basis for designing biomaterials.

Silk is a mysterious natural protein material that out-performs any industrious counterparts. Its molecular combination of highly stiff as well as tremendously elastic subunits gives rise to an invincible toughness, which inspires extensive research efforts on silk mechanics. Here, multiscale modelling and simulations using all-atom molecular dynamics and finite element methods are used to study silk protein mechanics. Combining these two different methods allows to maintain a certain accuracy of the model, while raising the computational limit of the modelled system to the micrometer scale. Using this hybrid approach, we could suggest a silk fiber structure with serially arranged stiff crystalline subunits in a soft amorphous peptide matrix, which is stronger than the commonly assumed random arrangement. Protein β-sheet crystals with smaller cross section area and longer β-strand length are found to better reinforce the fiber structure. These findings are likely to apply to similar semi-crystal materials like polyamides, preliminary studies of which have been performed in this thesis.

Muscle protein fibrils form networks for carrying the mechanical load present in living cells. The complex muscle protein network consists of molecules with well-tailored properties for mechanical adaptivity. The M-band protein myomesin is one such protein fibril. In collaboration with experimentalists using X-ray analysis and force spectroscopy, we analyzed myomesin's

response to force using force-probe Molecular Dynamics simulations. Myomesin comprises long α-helices between rigid domains, which we found to act as strain absorbers in muscle contraction cycles in this study. It also includes a highly force-resistant dimerization interface, which maintains the stability of the fibril network in the M-band of the force generating unit of muscle.

This thesis focused on the two biological systems, silk and myomesin, to dissect the determinants of mechanical function. Using advanced techniques, Molecular Dynamics simulations, Force Distribution Analysis and Finite Element Methods, we could show how the two proteins have been optimized by nature in terms of their mechanical properties. As a common theme, these two proteins utilize a specific assembly of building blocks, most importantly β-sheet structures oriented along the pulling direction, to achieve their mechanical function. This study, as a theoretical approach, opens the road for further efforts to dissect the role of forces in biosystems on a range of scales. It can serve as a guideline for experiments on these and analogous protein systems. As demonstrated for silk fibers, the computational approach employed in this thesis can be expected to aid the design of natural or artificial nano-structured materials.

Zusammenfassung

Mechanische Kraft ist eine wichtige Form des biomechanischen Signals. Ihre Rolle und Wirkung jedoch sind noch unerforscht und der Mechanismus der meisten Kraftsensoren wirft Rätsel auf. Wie widerstehen Biomoleküle wie Proteine hohen mechanischen Kräften? Wie können sie so empfindlich auf eine Änderung der mechanischen Umgebung reagieren? Antworten auf diese Fragen würden neue Türen für die Herstellung von Biomaterialien öffnen.

Seide ist ein natürliches Protein. Es erscheint uns als geheimnisvolles Material, ist es doch all seinen industriell hergestellten Pendants mechanisch überlegen. Die Kombination von extrem steifen mit enorm elastischen Untereinheiten führt zu einer ungewöhnlich hohen Zähigkeit, die zu umfangreichen Forschungsaufwand auf dem Gebiet der Seidenmechanik inspiriert. In dieser Arbeit wurde die Mechanik von Seidenproteinen mithilfe von Multiskalen-Modelierung untersucht. Hierzu wurden molekular-dynamische Simulationen, die alle Atome betrachten sowie Finite-Elemente-Methoden verwendet.

Durch die Kombination dieser beiden Methoden konnten wir einerseits eine hohe Genauigkeit des Modells erhalten, und andererseits das modellierte System trotz des hohen Rechenaufwands ebenso auf der Mikrometerskala betrachten. Mit diesem hybriden Ansatz liefern wir heute den Vorschlag für eine seidene Faserstruktur, die stärker ist, als was uns die Natur liefern kann: Eine serielle Anordnung der steifen kristallinen Untereinheiten in einer weichen, amorphen Peptid-Matrix ist der gemeinhin angenommen zufälligen Anordnung, die als die natürliche gilt, überlegen. Protein β-Faltblatt-Kristalle mit kleiner Querschnittsfläche und erhöhter β-Strang Länge stärken die Faserstruktur. Wir gehen davon aus, dass diese Erkenntnis ebenso auf ähnliche semi-kristalline Materialien wie Polyamide zutrifft.

Vorläufige Studien zu dieser Fragestellung wurden im Rahmen dieser Arbeit durchgeführt.

Auch für Muskeln spielt Kraft eine entscheidende Rolle. Muskel Protein-Fibrillen bilden Netzwerke, welche die mechanische Belastung in lebenden Zellen tragen. Das komplexe Netzwerk aus Muskelproteinen besteht aus Molekülen, die für ihre mechanische Adaptivität magermaßgeschneiderten sind.

Das M-Band Protein Myomesin ist eine solche Proteinfaser. In Zusammenarbeit mit Experimentalisten, die uns mit Röntgenstrukturanalysen und Daten aus ihrer Kraft-Spektroskopie unterstützten, untersuchten wir, wie Myomesin in molekular-dynamischen Simulationen auf Krafteinwirkungen reagiert. Myomesin besteht aus langen α-Helices die zwischen starren Domänen lokalisiert sind. Unsere Analysen zeigen, dass diese als Stamm-Absorber im Muskelkontraktionszyklus fungieren. Des weiteren enthält Myomesin eine Kraft resistente Dimerisierungs-Schnittstelle. Diese erhält die Stabilität des Fibrillen-Netzwerks im M-Band der Kraft generierenden Einheit im Muskel.

Der Fokus dieser Arbeit liegt auf zwei biologischen Systemen: Seide und Myomesin. Ziel war, deren mechanische Funktion zu verstehen. Mit Hilfe modernster Techniken wie molekular-dynamischen Simulationen, Kraftverteilungsanalyse und Finite Elemente Methoden konnten wir zeigen, dass die Natur diese beiden Proteine in Hinsicht auf ihre mechanischen Eigenschaften optimiert hat. Diese beiden Proteine teilen ihr Aufbauschema: Beide Strukturen werden vom β-Faltblatt dominiert, das sich entlang der Zugrichtung ausrichtet, um seine mechanische Stabilität zu erreichen. Unser theoretischer Ansatz bildet einen Grundstein für weitere Untersuchungen, die die Rolle von Kraft in Biosystemen analysieren. Er kann als Leitlinie für neue Experimente auf diese und ähnliche Protein-Systeme dienen. Wie wir für Seidenfasern zeigen konnten, lässt sich der rechnerische Ansatz bei der Gestaltung natürlicher oder künstlicher nanostrukturierter Materialien nutzen.

Acknowledgements

Upon finishing this project, I would like to take this chance to express my special thanks to Dr. Frauke Gräter. Her ideas and insightful suggestions have been key to the whole project. I owe many important results to inspiring discussions with Dr. Gräter. As a student working with Dr. Gräter, what I have learned is more than science but how to be a researcher. I believe I can benefit from these lessons learned throughout my life time.

I would like to sincerely thank Dr. Wolfram Stacklies. As a colleague and close friend, he brought great help and joy both in science and life. I also would like to thank Dr. Fei Xia, Dr. Murat Cetinkaya, Dr. Agnieszka Bronowska, Martin Brown, Dr. Carsten Baldauf, Shijun Xiao, Dr. Scott Edwards, the whole Molecular Biomechanics group working at the Heidelberg institute for Theoretical Studies and the whole Klaus Tschira Lab at MPG-CAS Partner Institute for Computational Biology.

I would like to thank both my supervisors, Prof. Roland Eils and Prof. Jörg Langowski, for continuously monitoring my progress. I also thank MPG-CAS Partner Institute for Computational Biology and Heidelberg Institute for Theoretical Studies for firm support of my work.

I also owe thanks to all collaboration partners throughout the project, especially Prof. Dr. Thorsten Hugel at Technische Universität Müchen, Prof. Dr. Matthias Rief at Technische Universität Müchen, Dr. Matthias Wilmanns at EMBL-Hamburg and Dr. Nikos Pinotsis, University of Vienna.

Finally, I would love to thank my family for all.

Contents

Abbreviations

AFM	atomic force microscopy
DNA	deoxyribonucleic acid
FCMD	force-clamp molecular dynamics
FDA	force distribution analysis
FE	Finite Element
FEA	finite element analysis
FEM	finite element method
FPMD	force-probe molecular dynamics
Ig	immunoglobulin
MD	molecular dynamics
MM	molecular mechanics
NMR	nuclear magnetic resonance
PA	polyamide
PBC	periodic boundary condition
PDB	protein data base
PDE	partial differential equation
PME	particle mesh Ewald
QM	quantum mechanics
RNA	ribonucleic acid
WLC	worm like chain

ABBREVIATIONS

1

Introduction

1.1 Protein materials

Proteins are outstanding materials. As major constituents of fibers, such as muscle fibers in vivo or silk fibers in vitro, protein materials demonstrate great adaptability as well as extraordinary mechanical properties. These protein materials normally show a high elasticity, and at the same time exhibit significant rupture strain and stresses, that easily outperform man-made counterparts. Up to today, mankind can still not reproduce materials with such properties at industrial scales. The secret lies in the building blocks of protein and their intricate assembly mechanism, most of which has remained largely unknown and awaits systematic studies.

Proteins are probably the most versatile biomolecules in nature. They are highly optimized to fulfill the requirements of their environment. Not only do they perform enzymatic functions to support life, they also serve as the most fundamental materials to build the cell. The biggest and toughest protein found in a living cell is titin, (1) which is found in different types of muscle and involved in withstanding the large forces generated by a muscle cell. Silk (Fig. 1.1) is mysterious in the ability of balancing an outstanding resistance against rupture and a tremendous elasticity to reach an extraordinary toughness for hunting and protection. (2–6) While it has been utilized by mankind for thousands of years, silk nowadays still outperforms any similar industrial materials produced by cutting edge modern technology.

The key determinants of the amazing features of protein materials lie in their molecular architecture. Proteins, like other materials, use basic physical interactions to build

Figure 1.1: Spider web. Spider silk is one of the strongest fiber in nature. Being made of protein, silk has been optimized by evolution to provide the spider with an extremely tough fiber. (This figure is provided by *Dr. Frauke Gräter*.)

up their molecular structures. There are common building blocks of structural motifs that can be further assembled to higher architectures. Understanding the physics behind protein structures is the key to unraveling the robustness of protein materials.

This thesis focuses on theoretical studies of protein materials. Protein structures of silk and the muscle protein myomesin are chosen as intriguing examples. Using molecular structures either from molecular modeling or experiments, computer simulations are performed to study their structural mechanics. Continuum modeling is also used to link different study scales to clarify the relationship between material molecular architectures and macro properties. The goal of this study is to understand how protein building blocks affect structural mechanics and thereby influence the material's behavior, aiming at establishing a guideline for protein material design.

1.2 Protein mechanics

How and how much can proteins withstand force? Modern technologies such as atomic force microscopy (AFM) or optical tweesers have been widely used to probe protein mechanics. The principle of these experiments is to apply mechanical load to protein structures (see Section 2.1.2). By monitoring the unfolding process and recording the force and extension until rupture, and thus the overall mechanical response of protein

structures can be obtained. Such experiments have been carried out extensively to study ubiquitin and immunoglobulin (Ig) domains. (7; 8) They indicate that proteins can withstand force as high as a few hundred piconewtons, although protein structures in equilibrium are held together by non-covalent interactions, each of which is marginally stable and subjected to thermal fluctuations. The collection of these interactions maintain the integrity of protein structures. They are described in the following basis of molecular mechanics.

Protein structures

As this thesis will illustrate the structure of a protein is the major determinant of its mechanical response. For this reason, the structures a protein can adopt are shortly described in this section.

Protein structures are the folded states of unbranched polypeptides of amino acid residues. Much effort has been dedicated to research on protein structures which have been classified into four main levels.

The protein primary structure is the sequence of amino acids in a peptide (Fig. 1.2). There are normally 20 different amino acids found in proteins. These amino acids are attached to each other in a chain using their carboxyl group and amino group of the next amino acid by dehydration reactions. The length of these unbranched peptide ranges from a dozen to a couple of thousand residues.

Protein structures feature some common building blocks by assembly or local folding of peptides. These building blocks are called protein secondary structures. The most abundant protein secondary structures are the α-helix and the β-sheet, as shown in Fig. 1.2. An α-helix is a right-handed spiral. Amino acid residues in an α-helix interact strongly via hydrogen bonds with residues which are four residue positions away in sequence, which stabilizes the helix structure. Peptide strands in β-sheets are fully extended and positioned in line with each other. Interactions in between create strands through hydrogen bonds hold the structure together.

The relative arrangement of secondary structures is the tertiary structure (Fig. 1.2), i.e. the three-dimensional spatial structure, of a protein. Different proteins take up very different tertiary structures. Some protein structures are similar and classified into the same structure category called family. Spatially independent folds of peptides are defined as domains. The keys for structural stability of protein tertiary structures

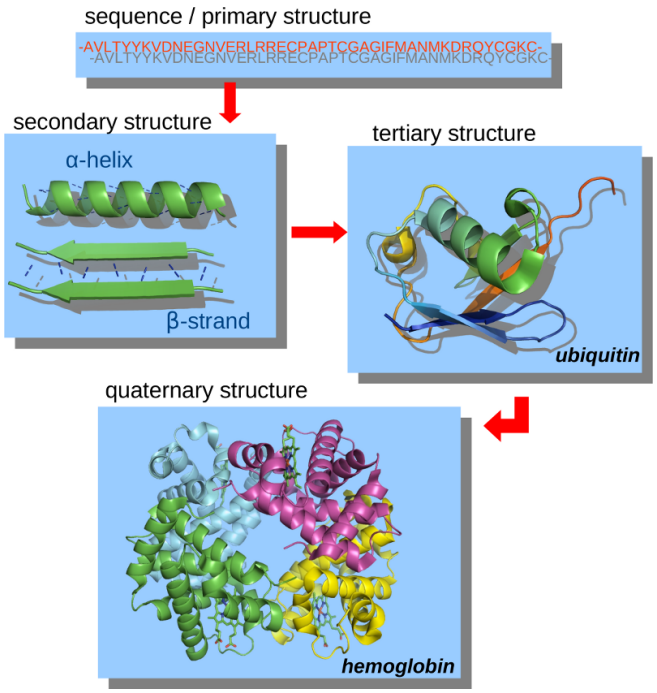

Figure 1.2: Protein structural hierarchy as a determinant of protein mechanical function. Protein primary structures fold into secondary structures which in turn serve as building blocks for tertiary structures. α-helix and β-strands are shown as typical protein secondary structures. Hydrogen bonds are depicted as blue dashed lines. Ubiquitin (PDB code: 3N32) and hemoglobin (PDB code: 3QJD) are shown as protein tertiary and quaternary structures, respectively. Different colors are used for different subunits in hemoglobin; heme groups are shown in sticks.

is the complicated pattern of both hydrophobic and hydrophilic interactions, which is complemented by special bonds like disulfide bonds.

Some protein structures are further assembled from more than one peptide chain. The assembly of protein tertiary structures subunits into a larger complex results in the so-called quaternary structure. Covalent bonds between subunits are absent in a protein quaternary structure. Proteins use quaternary structures to achieve special molecular functions based on structural or chemical interactions between subunits, a prime example of which is hemoglobin. (9) While silk protein fibers are dominated by β-sheets as their secondary structure (Chapter 3), myomesin shows an interesting alternation of α-helices and β-sheets (Chapter 3). This thesis gives numerous evidence for that unstructured or α-helical structures are soft and β-sheets are comparatively

rigid and force-resistant.

Driving forces

A linear peptide is a polymer chain. It shares the same mechanical properties of a polymer strand. A linear polymer chain tends to collapse into unorganized structures to increase its entropy.

A protein folded structure is defined by its primary sequence. In many cases, a given peptide sequence has a specific tertiary structure when properly folded. This fact suggests that in a protein folding process from a linear peptide entropy first increases during chain collapse and then decreases till the final well-defined low entropy structure is reached. The decrease of entropy is unfavorable, and is compensated by the forming of favorable interactions between different parts of the peptide. These interactions optimize protein structures to lower energy states and stabilize the architecture.

These non-bonded interactions within a folded protein rupture when the structure is unfolded by a mechanical force, the process primarily studied in this thesis. In a classical mechanical description, interactions in proteins are classified into two types: Coulombic interactions and van der Waals interactions. Two point charges can either attract or repel each other with the energy given by Coulomb's law:

$$V_{coul} = \frac{1}{4\pi\epsilon_o\epsilon_r}\frac{q_1 q_2}{r} \ , \tag{1.1}$$

where q_1 and q_2 are point charges, r is the distance between them, ϵ_o is electric constant, and ϵ_r is the permittivity of the medium which equals to 1 in vacuum. An atom consists of a positively charged nucleus and a negative electron cloud. The electron distribution around the nuclei of a molecule hence is approximated by partial charges on the atoms.

Van der Waals interactions are due to dispersion interactions between atoms. The dynamics of electron clouds causes random fluctuations of the electrostatic field. Random dipoles in adjacent atoms resulting from these fluctuations give rise to a short range attractive potential. An approximate and computationally efficient description of this interaction is the Lennard-Jones potential, given by

$$V_{LJ} = 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}] \ , \tag{1.2}$$

where r is the distance between two atoms, ϵ is the depth of interaction energy, and σ is the interaction radius. The van der Waals interaction has two contributions: (i)

London's dispersion, which is negative and given by $(\frac{\sigma}{r})^6$, is the attraction energy between two dipoles, and dominant for two distant dipoles; (ii) Pauli exclusion is positive and given by $(\frac{\sigma}{r})^{12}$; it is the repulsion energy between two atoms, and dominant if their electron clouds overlap. The potential has an energy minimum, where the two atoms are at the most favorable distance. This distance is the sum of the van der Waals radii of two atoms.

Hydrogen bonds

Hydrogen bonds are maybe the most determinant of protein structures. Importantly, they are generally the interactions stabilizing protein against force-induced rupture and thus are of primary importance of this thesis (see e.g. Section 3.1.1 and Section 3.1.5 in Chapter 3). They involve both Coulombic and van der Waals interactions. In contrast to chemical bonds, hydrogen bonds are non-covalent, and thus reversible interactions between polar group. When a hydrogen atom is covalently linked to an electronegative atom such as oxygen, the electron cloud on the hydrogen atom shifts towards the adjacent atom, leaving the hydrogen atom partially positively charged. This type of hydrogen atoms then are likely to attract other electronically negative atoms such as nitrogen atoms or oxygen atoms to form strong interactions, the hydrogen bonds (Fig. 1.3).

Hydrogen bonds are far more weaker than real covalent bonds, with a lifetime of some picoseconds, and an interaction energy of 2$\sim$10 kJ/mol. (10) Although an individual hydrogen bond is weak and constantly breaks and reforms, several hydrogen bonds ca n be strong and reach lifetimes of nanoseconds. There can be hundreds of hydrogen bonds in a protein structure. These hydrogen bonds behave in clusters to resist shearing or tearing forces, which impellingly stabilizes protein structures.

Amino acid residues are capable of forming hydrogen bonds because of their carbonyl and amino groups. These two groups are partially charged and ideal for forming hydrogen bonds between each other. The protein backbone thereupon can form hydrogen bonds between amide groups close or distant in sequence. Hydrogen bonds between different parts of the peptide backbone largely create and define secondary structures. α-helices have hydrogen bonds between every fourth residues, resulting in a helical shape. β-sheets exhibit clusters of hydrogen bonds, which can be either arranged or in zig-zag patterns depending on the relative arrangement of the peptide

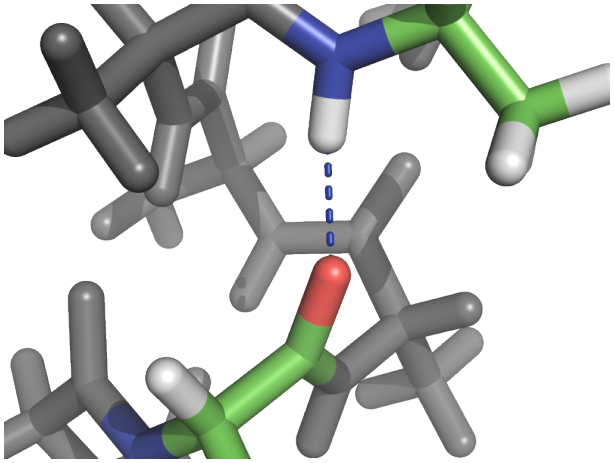

Figure 1.3: A hydrogen bond formed between polar groups in the protein backbone. Blue, white, red and green atoms are nitrogen, hydrogen, oxygen, and carbon atoms, respectively. A hydrogen bond results from both the combination of Coulombic and van der Waals interactions between the four atoms.

strands (Fig. 1.2). These hydrogen can cooperatively strengthen a protein structure such that they give rise to rupture forces up to a few nanonewton (Chapter 3).

Hydrophobicity and hydrophilicity

We so far focused on the structural rol of backbone hydroegn bonds, amino acid side-chains also massively contribute to protein structural stability. One can classify favourable contacts between sidechains into two types, hydrophobic and hydrophilic interactions.

Some of the amino acid side-chains comprise ethyl groups, methyl groups or benzene rings (such as leucine and phenylalanine). These chemical groups are largely apolar, being in a neutral charge state with virtually no charge center displacement. Water molecules, the main solvent in nature, are polar in contrast. There is a dipole in each water molecule, rendering interactions with other dipoles of the solvent or solute favorably. Introducing an apolar solute into water causes an ordering of water dipoles at the solute surface, which entails an entropy penalty. Reducing the contact area of waters with apolar groups therefore is entropically favorable, leading to the so-called hydrophobic effect. (11) Thus, apolar side-chains in proteins are excluded from water molecules in general. Amino acids with hydrophobic (apolar) side-chains are likely

to form the core of protein structure where they are far away from water molecules. Again, a hydrophobic protein core can be a very efficiently stabilized structure against mechanical forces (e.g. Chapter 3, Section 3.2).

The hydrophilic effect can be considered as the opposite of the hydrophobic effect. Some amino acids have polar or charged side-chains (such as serine and threonine). These amino acid side-chains are likely to favorably interact with water molecules and are mostly found on the surface of protein structures.

Both hydrophobicity and hydrophilicity are important for organizing protein secondary structures to a higher structural hierarchy, and thus for th overall stability. α-helices and β-sheets may have polar side-chains at one side to face solvents and apolar side-chains at the other side to construct the core of a protein domain. These effects are important driving forces of protein folding into their native structures, and therefore also crucial molecular forces for resisting external pulling forces..

Salt bridges and disulfide bonds

Salt bridges and disulfide bonds are special bonds constructed by pairs of amino acid side-chains. They are less commonly observed but are also important to protein structural stability. When negatively charged side-chains of either aspartic acid or glutamic acid are close to positively charged side-chains of lysine or arginine, they form a non-covalent interaction, termed salt bridges. Disulfide bonds are covalent chemical bonds. They can only form between two cysteine residue side-chains. Disulfide bonds can strongly reinforce protein structure from rupturing. The rupture of disulfide bonds normally requires special enzymes, since covalent bond rupture in the absence of a catalyst requires several nanonewton of force. (12)

Protein mechanics from simulations

Experimental methods for single molecule studies have been revolutionized by advances in technology over the last few decades. In particular, atomic force microscopy (AFM) and optical tweezers have greatly improved the experimental capabilities of single molecule studies in molecular biology. Single molecule techniques reveal processes which would remain obscured by statistical ensemble measurements such as Nuclear

Magnetic Resonance (NMR) or Fluorescence Resonance Energy Transfer (FRET) experiments.

Both AFM and optical tweezers are capable of applying mechanical force to a single molecule. They are frequently used in probing dynamics on molecular transition pathways, such as protein unfolding or ligand-protein interactions. Due to its ability to work in liquids and at room temperature and pressure, AFM became a valuable tool in structural molecular biology. (13) The AFM uses an extremely sharp tip on a cantilever to measure deflection at the surface of the tip. This can be used to apply mechanical forces directly to the objects underneath the tip, and thus to manipulate the proteins under dynamic load (see Section 2.1.2). AFM has been successfully used to open protein domains to probe molecular mechanisms. (14–16) It is also used to explore protein refolding energy landscapes (17–19). Optical tweezers, on the other hand, use optically trapped beads to apply force indirectly onto macromolecules. Optical tweezers are used in studies of folding pathways, (20) as well as DNA-protein interactions and recognition events. (21; 22)

NMR spectroscopy, although it is not a single-molecule technique, is frequently used for simultaneous determination of structure and dynamics of macromolecules. (23) The principle of NMR spectroscopy is the non-zero nuclear magnetic moment of many elements, such as ^{1}H, ^{13}C, ^{15}N, or ^{19}F. (24) When put into an external static magnetic field ($\mathbf{B}$), the different nuclear spin states of these elements become quantized with energies proportional to their projections onto vector $\mathbf{B}$. The energy differences are also proportional to the field strength and dependent on the chemical environment of the element, which makes NMR an ideal technique to study 3-dimension structural and dynamical properties of the systems. (25; 26) Development of techniques for incorporation of stable ^{13}C and ^{15}N isotopes into expressed proteins allows for application of modern multidimensional heteronuclear NMR techniques. As a result, the maximum size of the macromolecule studied using these techniques rose from about 10 kDa (when ^{1}H homonuclear NMR is used) to 50 kDa and beyond (using ^{13}C and ^{15}N heteronuclear NMR with fractional ^{2}H enrichment). Application of modern TROSY (transverse relaxation optimized spectroscopy) techniques further expanded the size limitations of NMR, reaching up to 900 kDa. (27; 28)

One of the major challenges in all single molecule techniques is the size of the system studied, therefore, the requirement for macro-scale measurements of nano-scale objects.

Another challenge lies within the physical variables measured by these techniques. In case of both AFM and optical tweezers, the measurable quantities are force applied and end-to-end distance of the investigated molecule; internal dynamics and transient phenomena are not measurable. On the other hand, NMR, which is capable of probing transient phenomena and internal dynamics on various time scales, lacks the ability to apply mechanical force.

Molecular biology experiments have been traditionally limited by both the size and time scales of single molecule events. Computer simulations helped to overcome these limitations by providing insights into events at short time scales and high resolutions. Dynamical behaviors of macromolecules can be studied using extensive molecular dynamics (MD) simulation. Computer simulations take great advantage of huge amount of molecular structural data available nowadays. In a situation when experimental structures of the protein target or its part is missing, computational techniques also allow for generating such structure (e.g. by homology modeling) and use it for studies which can be validated experimentally. (29) Computational approaches allow also for investigations of transient phenomena, e.g. short-living conformers from an ensemble that contribute to the binding event but which cannot be readily observed. Therefore, usage of theoretical methods is indispensable – not only for the interpretation of the existing experimental data, but also to direct and design new experiments.

Among computational methods suitable for macromolecular studies, high-level quantum mechanical (QM) methods are the most accurate. QM methods, which rely on solving the Schrödinger equation, are e.g. used for charge derivation for molecular dynamics simulations, for description of intramolecular interactions, for calculations of pKa, protonation, redox states, and for studying solvation effects, such as computing the free solvation energies.(30) QM methods, on various levels of theory, have been extensively used to study mechanisms of chemical reactions in biochemistry, and macromolecular association events. (31–33) The application area of QM methods is not limited to equilibrium conditions. Recently, QM methods have been employed for studies of disulfide bond breaking under mechanical force with the great success. (34; 35) Two main disadvantages of QM methodology are the computational cost of the calculations and the size of the system suitable for studies, which is fairly small (up to 1000 atoms). These make QM calculations, despite their great accuracy, impractical

for studies of protein structural transitions such as large conformational changes or protein folding.

In order to improve computational efficiency, hybrid Quantum Mechanics/Molecular Mechanics (QM/MM) methods are put forward. These calculations only treat the most interesting part of molecule, such as the reaction center, with quantum mechanics while other parts with molecular mechanics. (36–38) Hybrid QM/MM methods take advantage of accuracy of QM and efficiency of MM, which are successfully applied in structure-based drug design, enzymatic reaction, covalent bond breaking and forming and others. Although these calculations have largely improved the computational performance, QM potential calculations are still the time-limiting steps. Calculations of bigger systems need a further approximation.

Molecular dynamics (MD) simulations introduce a level of approximation to study the dynamics of atomic-level phenomena of larger systems. MD simulations are based on the numerical solution of the Newton's equations of motion of a macromolecular system. The potential energy of the particle system is described by the so-called force field (see Section 2.1.2), which is described as a sum of energy terms for covalent bonds, angles, dihedral angles, van der Waals non-bonded terms, and non-bonded electrostatic terms. (39; 40) All-atom force fields provide parameters for every type of atom in a system, united-atom force fields treat the functional groups such as methyl as single units, while coarse-grained force fields, such as the Go-Model (41), provide even more crude representations for increased computational performance and simulations at larger time scales.

MD simulations (see Section 2.1.2) are successful owing its efficiency when compared to QM calculations. Since the kinetic energy is also taken into account, MD simulation systems are able to move across the energy barriers on the potential energy surface, which implies substantial changes (e.g. conformational) during the simulation. This makes MD simulations suitable for studying dynamic events on atomic level. It has been proven to be extremely potential in molecular biology research such as DNA, RNA and protein mechanics. The wide usage of MD simulations is also reinforced by dozens of effective and convenient simulation software packages. Different simulation techniques also enable scientists to explore an even wider range in molecular science on nanometer and microsecond scale. The combination of MD simulations with single molecule experiments such as AFM or optical tweezers is even more significant. (42)

Force-probe MD and force-clamp MD, in which mechanical force is applied to molecular structures, allow mimicking single-molecule experiments to gain insight into structural and dynamical details of molecular processes such as protein unfolding. (43–45) The results obtained not only help to explain experimental observations (e.g. AFM, optical tweezers), but also shed light on the nature of the molecular processes under investigation.

In summary, MD simulations are an efficient and relatively straightforward technique suitable for studying structural, dynamical, and mechanical properties of macromolecules. They offer a good combination of atomic resolution, accuracy, and computer performance, and hence are the primary tool used in this dissertation (Chapter 3).

1.3 Continuum mechanics

MD simulations are an efficient approach to study objects on the nanometer scale. They become impracticable when the investigated system and its relevant properties has significantly larger dimensions, i.e. on the meso or macroscales.. Large scale simulations of protein materials are not feasible using MD simulation, neither can macro scale structural and mechanical material properties be straightforwardly inferred from MD simulations of subsystems due to the complexity of biological systems. Continuum structural mechanics, by introducing a higher level of approximation, can significantly extend the length and time scale of the system under investigation.

Continuum structural mechanics is a classical method applicable at all scales. Here, as its name suggests, atomic details of the object are disregarded to focus on the effect of the overall structure on mechanics only. In certain cases, such as determining an elastic modulus of a rubber band, or understanding the stress bearing pattern of a building, atomisation is normally not the determinant of the question asked. In such cases, the object or parts thereof are treated as a continuum sharing the same physical properties. This approach has invincible computational efficiency by greatly reducing the complexity, while at the same time largely maintaining the accuracy, at least as long as the continuum approximation is adequate.

Using different levels of simplifications, continuum mechanics is able to tackle problems for structure size that can not be reached by MD simulations. In spite of the conceptual differences between continuum mechanics and all-atom simulations, the combi-

nation of these two different methods is promising. Parameters for continuum mechanics can be obtained from all atom simulations that ensure high accuracy. The efficiency of continuum mechanics allows to infer mechanical properties on the macroscopic scale from the molecular scale, which is the key to understand material's interplay of the atomic structure and its overall performance. We have employed continuum mechanics to accompany all atom MD simulations in silk protein modeling. In collaboration with Murat Cetinkaya, within this thesis, we have followed this approach (Chapter 3, Sections 3.1.3 and 3.1.4). (46; 47)

Finite element method

Finite element method (FEM), also call finite element analysis (FEA), is a well established continuum mechanics method. It is also the most common and efficient technique as compared to other similar continuum methods. FEM was invented in the middle of the 20th century, and has since then greatly developed to become an indispensable tool in engineering science.

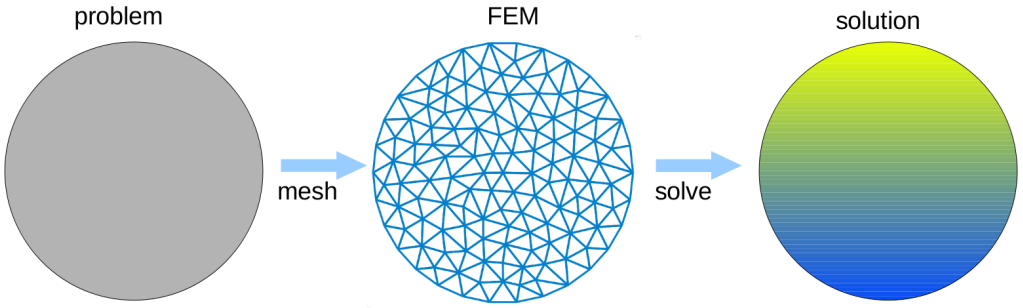

Figure 1.4: A scheme of problem solving using finite element method. A complicated problem is divided into well understood smaller elements. Solutions in each element are assembled to yield a final answer.

In structural mechanics, especially for odd structures, determining the partial differential equations (PDE) that govern physical variables such as stress, strain and displacement between each part is critically challenging. FEM is an approximate numeric way to find the solution of the PDE. The first step in FEM involves dividing the structure into different regular domains representing different material properties. Domains are then further meshed down to smaller units the properties of which are well understood. These units are the elements in FEM. Physical problems are solved in

each element based on the submitted conditions and elemental connection properties. The results from every element are finally assembled to find the mechanical response of the whole structure.

Because of its simplicity, FEM is widely applied in all kinds of continuum mechanics analysis such as architecture engineering, force distribution, damage simulation, thermodynamics, fluid dynamics studies and others. It is employed in wide aspects of industry. FEM is used in this project to work with all atom MD simulation to study silk fiber mechanics (Chapter 3, Section 3.1). Continuum mechanics models built in FEM use parameters from all atom simulations in a bottom-up approach. Fiber mechanics are then predicted in FEM models for different types of fiber assembly. Overall, combining MD with FEM results in a predictive scale-bridging technique far beyond the capabilities of one of the two methods.

More details on FEM theory are given in Section 2.2. Technical details on FEM as used in this thesis are provided in relevant sections in Chapter 3.

2

Methods

2.1 Molecular Dynamics Simulations

The mechanical properties of a protein are determined by its dynamical behavior under force. Protein dynamics cover a huge range of time scales. Bonds vibrate at a time scale as small as few femtoseconds, yet folding of a medium sized protein takes several microseconds. While dynamic processes determine the molecular and more importantly the mechanical properties, their determination often is beyond the limitation of modern technologies. With the fast growing computing power, computational approaches based on well established physical laws to predict molecular dynamics of proteins on the microsecond time scale have become possible.

Theoretically, the time-dependent Schrödiger equation can accurately predict any molecular process at any size. It considers the electron distributions in molecules, and deduces how the electron distribution governs molecular transitions. Such quantum mechanics calculations have proven useful for various systems and problems.

As the major limitation of quantum mechanics, it can only deal with small systems of typically less than 100 atoms by nowadays algorithms. The cost of quantum mechanics calculations rises with the square of the number of electrons in the system, which renders computational studies on large molecular systems impossible. Other computational methods with reasonable approximations are needed in order to overcome the limitations of quantum mechanics. Molecular mechanics that ignore quantum effects is such approximation; one most successful representative is Molecular Dynamics (MD).

The molecular mechanics framework describes atomic motion using classical physics

laws. Atoms are modeled as spherical particles to each of which specific properties, namely a charge, mass and van de Waals parameters are assigned. Different parameters are assigned to different atom types. Different interactions between atoms are modeled by bonded or non-bonded interactions. All parameters to describe the interaction potentials are termed as force field. They are point based, and obtained by quantum mechanics calculation and more importantly by experimental measurements. Comparing to quantum mechanics, the molecular mechanics description is not only orders of magnitude faster, but can also produce reliable results on the nano to micro second scale, when electron transitions are not relevant.

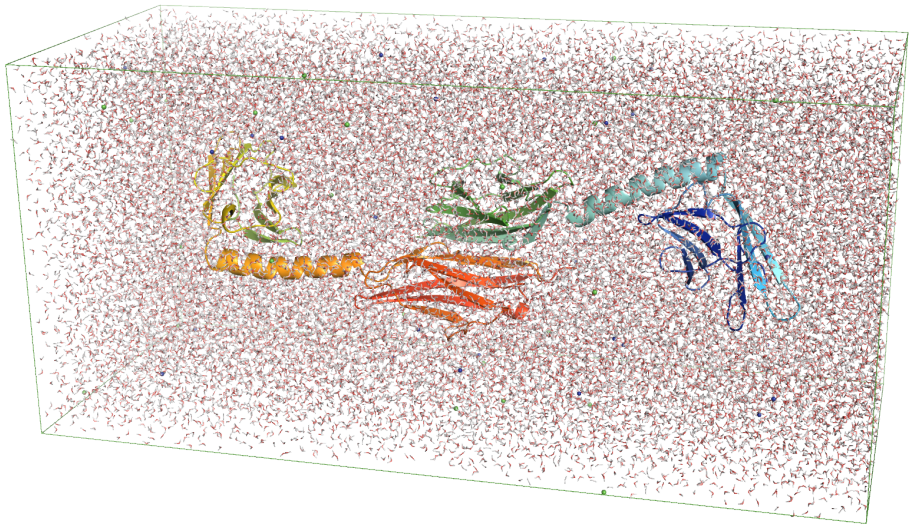

Figure 2.1: A molecular dynamics simulation system. Protein structure (colored cartoon representation) in this system is solvated in water solvent with ions. Water molecules are shown as lines, with oxygen atoms in red and hydrogen atoms in white; sodium and chlorine ions are shown in blue and green spheres, respectively. The protein shown is the dimerization interface of myomesin, the mechanics of which have been studied in this thesis (Chapter 3, Section 3.2).

Using the molecular mechanics description detailed above, MD simulations calculate the motion of every single atom in the studied system in time, termed a trajectory. Molecular structures, like proteins, RNA or DNA, contain coordinates of all atoms, and served as the starting structure for a simulation. They are obtained by experiments

such as X-ray crystallography, Nuclear Magnetic Resonance (NMR), or other methods like computer modeling. The molecular topology, the second requisite for the MD simulation, contains the atomic and inter-atomic properties for the molecular structure, as given by the force field (see below). It is a collection of parameters need to describe the simulation system and does not change during the whole simulation process. Initial velocities of all atoms are applied randomly, taken e.g. from a Poisson distribution. Forces acting on every atom are calculated using the potential functions defined in the force field. The dynamics of all atoms by calculated by integrating Newton's equation of motion:

$$\begin{aligned} F_i &= -\frac{\partial V\left(r_1, \cdots, r_N\right)}{\partial r_i} \\ &= m_i \alpha_i \ , \end{aligned} \tag{2.1}$$

where F_i is the force acting on atom i by all other atoms; m_i and α_i are the atomic mass and its acceleration under force, respectively. The coordinates and velocities of all atoms are updated accordingly for next cycle of force calculation. This process is repeated until the end of the simulation.

MD simulations are carried out for a certain simulation system such as the box shown in Fig. 2.1. The size and shape of the simulation system are chosen such that it can accommodate the studied molecule which is normally incubated by water molecules and ions. In order to remove artifacts from the box boundary, period boundary conditions (PBC) are commonly used. When using PBC, images of the simulation system are generated at each side of the box and repeated infinitely. The most obvious effect of PBC is that molecules reappear at the opposite side of the simulation box if they cross box borders. Because of PBC, artifacts due to molecular interaction with its image can arise. Thus, the simulation box should be large enough to prevent these artifacts by ensuring one single molecule not to interact with its own image.

The most time consuming step in MD simulations is the evaluation of non-bonded potentials. These potentials are atom-wise interactions between non-bonded atoms. A cut-off treatment, which only evaluates interactions between one atom and its neighbors within the cut-off distance, is normally used to accelerate the computations. Introducing a cut-off into a simulation does not significantly affect the Lennard-Jones potentials which decay rapidly with increasing distance. In contrast, a cut-off treatment ignores

long range electrostatic interactions which can be large. For also taking long range electrostatic into account, other techniques like Particle Mesh Ewald summation (PME) are used as approximations beyond the cut-off distance. PME uses Fourier transforms to estimate long-range Coulombic interactions. (48) It provides a mostly sufficient accuracy and smooth energetic transitions, and thus stabilizes the simulation system.

An additional option to accelerate a simulation is to constrain covalent bond vibrations. Covalent bond vibrations have a time scale of femtoseconds, and are the fastest motion in the system. This high frequency motion is normally not of interest for MD simulations, but is the most important determinant of the maximal length of the time steps in the simulation. By constraining covalent bonds to rigid beams, the simulation process can be speeded up by increasing the simulation time step. Covalent bond constraint algorithms such as SHAKE (49) and LINCS (50) are typically available in most MD software packages. For most MD simulations, a time step length of 1 or 2 femtoseconds is used.

By delivering a trajectory of the molecular system, i.e. a conformational ensemble, MD simulations can in principle yield thermodynamic quantities for comparison with bulk experiments. The reason is that, theoretically, the time average of one system in equilibrium should be the same with the ensemble average, given a long enough observation time. Thus, the results that are provided by MD simulations should be the same observed in a reaction test tube. However, while MD simulations can successfully explain why sodium ions can pass through their specific transporter in the cell membrane, they can not provide the electron voltage between the two sides of the cell membrane. The reason is the limited nano to microsecond time scale of MD simulations. Theoretical methods are therefore needed to infer such relationships.

Nowadays, many MD simulation software packages are available. Some of these packages are commercial such as AMBER (51), others are free and open-source, such as NAMD (52) and GROMACS (53). GROMACS is known for its computational efficiency and Linux command style, which enable it to be well embedded into a Linux operating system. Being open-source, the source code is available for all users for debugging and the implementation of new features. New simulation methods such as force distribution analysis (FDA) (54) as used for this thesis (see 2.1.2) have been implemented into GROMACS. For these reasons, in this study, the MD simulations have been carried out using GROMACS.

2.1.1 Force Fields

A force field is a collection of particle properties (atom mass, charge and so on) and interaction potential functions, which are used to describe the molecular energy in MD simulations. Both the parameters and potential functions are based on classical mechanics and developed from quantum calculations and experimental results. The potential energy is a function of the coordinates of particles that represent atoms. Electrons of the atoms are ignored.

There is a variety of classical force fields nowadays available for MD simulations. The most wide-spread ones are the OPLS-AA force field, (40; 55) the AMBER force filed, (39) and the GROMOS force field (56), which vary in their strength and weaknesses. (57) Some of these force fields are all-atom force field like OPLS-AA and AMBER, in which all atoms are precisely described. Others, like the GROMOS force field, are united-atom force fields that fuse non-polar atoms (hydrogen atoms) into its bonded heavy atom to treat them as a single bead. Each force field has their own characteristics: e.g. the AMBER force field is famous for DNA and RNA simulations, and the OPLS-AA force field is famous for protein simulations. These force fields are developed on the basis of different concepts put forward by different groups, and thus have different particle properties to treat the same atoms. Nevertheless, they make use of similar potential functions for energy calculation.

A force field classifies all interactions into two types, namely bonded and non-bonded interactions:

$$E_{\text{total}} = E_{\text{bonded}} + E_{\text{non-bonded}} \tag{2.2}$$

Bonded interactions, E_{bonded}, in molecules include covalent bond stretching, angle bending and dihedral angle rotations (Fig. 2.2). Non-bonded interactions, $E_{\text{non-bonded}}$, include Lennard-Jones interactions and Coulombic interactions.

Covalent bonds are normally modeled as harmonic springs in force fields. The potential equation of bond stretching follows the Hooke's Law, which has a form of:

$$E_{\text{bonds}} = \sum_{\text{bonds}} K_r (r - r_{eq})^2 \ , \tag{2.3}$$

with K_r as the spring constant that specifies the bond strength, and r_{eq} the equilibrium bond length that gives the potential minimum. Because of this potential form, a chemical bond in a MD simulation can not break but only be extended (Fig. 2.3).

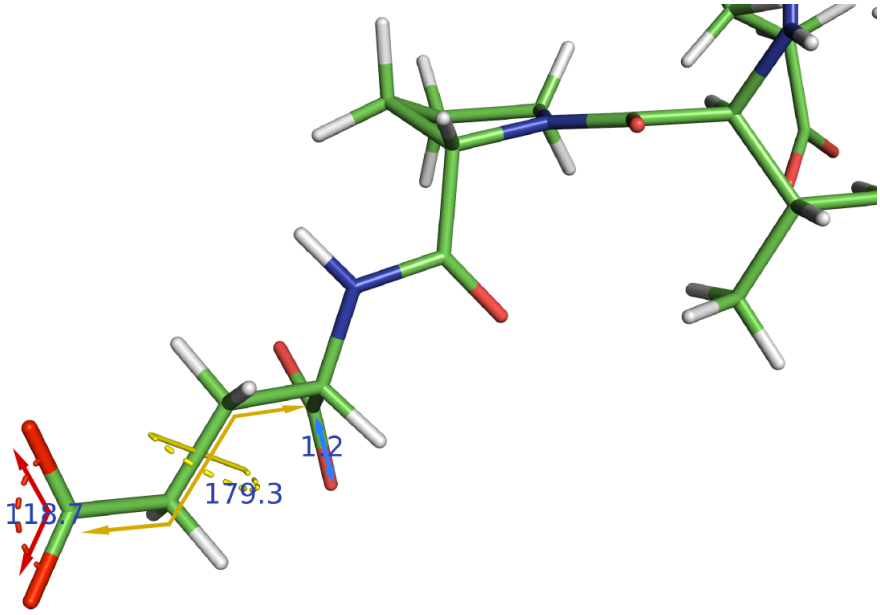

Figure 2.2: Bonded interactions in molecules. A bond (blue), an angle (red) and a dihedral angle (yellow) are three types of bonded interactions in molecules. The represented bond has a length of 1.2 Å, the angle and the dihedral angle are of 118.7° and 179.3°, respectively.

A bonded angle is formed by three connected atoms. Similar to bond potentials, angle potentials have functional form of:

$$E_{\text{angles}} = \sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2 \ , \tag{2.4}$$

with an angle force constant K_θ and an equilibrium angle θ_{eq} (Fig. 2.3).

A dihedral angle is formed by four bonded atoms. The energy function has periodicity for the rotation around the central bond of the four atoms. Its form (here for OPLA-AA) is more complicated than two other bonded potentials above, as:

$$\begin{aligned} E_{\text{dihedral}} = & \frac{V_1}{2}[1 + \cos(\phi)] + \frac{V_2}{2}[1 - \cos(2\phi)] \\ & + \frac{V_3}{2}[1 + \cos(3\phi)] + \frac{V_4}{2}[1 - \cos(4\phi)] \ , \end{aligned} \tag{2.5}$$

with V_1, V_2, V_3 and V_4 the coefficients in the Fourier series and ϕ the corresponding phase angle. The cosine function produces the periodicity (Fig. 2.3).

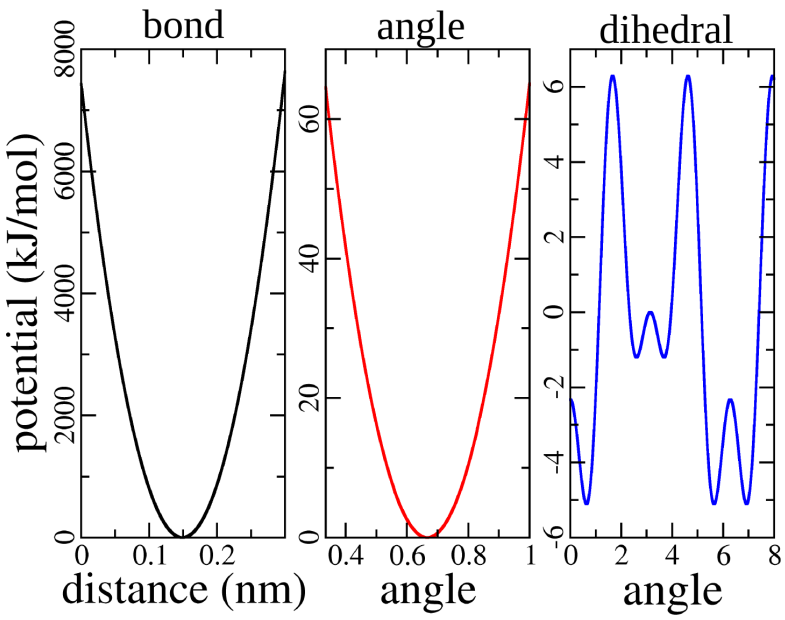

Figure 2.3: Bonded potentials in the OPLS-AA force field. Examples are shown for a bond potential (black), an angle potential (red), and a dihedral angle potential (blue).

Atoms that are not bonded to each other and stay at least three bonds away also obviously interact with each other if their positions are close enough. The interaction between these atoms are classified as non-bonded interactions. The non-bonded interactions include Van der Waals and Coulombic interactions, namely:

$$E_{\text{non-bonded}} = \sum_i \sum_j \left\{ 4\epsilon_{ij}[(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^{6}] + \frac{q_i q_j e^2}{4\pi\epsilon_o r_{ij}} \right\} , \tag{2.6}$$

where the first two terms together describe the Lennard-Jones potential (Fig. 2.4), with ϵ as the energy minimum, σ as the distance between two atoms that gives the energy minimum; and r_{ij} as the distance between atom i and j. The first term of the Lennard-Jones potential is exclusively repulsive and dominates when two atoms are so close that the electron densities overlap. The second term is London's dispersion that dominates and defines attraction when two atoms are far away. The third term in the equation is the electrostatic potential (Fig. 2.4), q_i and q_j are the charges assigned to the atoms i and j. It is attractive (repulsive) if the atoms have opposite (the same) charges.

Hydrogen bonds and hydrophobic effects are produced by non-bonded interactions. As one of the most important interaction in proteins or nuclear acid molecules, hydro-

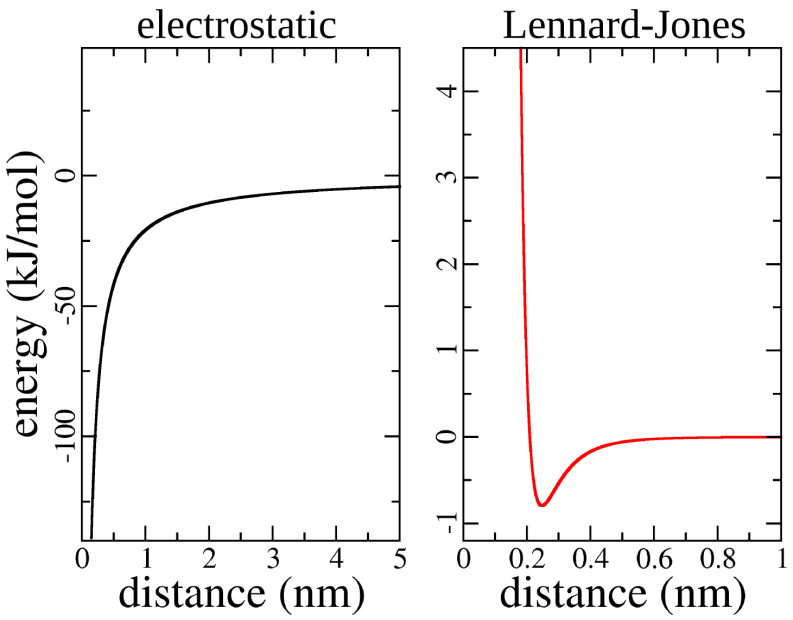

Figure 2.4: Nonbonded potential examples in OPLS-AA force field. The electrostatic potential is shown in black; the Lennard-Jones potential is shown in red.

gen bonds are formed, without that specific parameters are used. The partial positive charge of hydrogen atoms and the partial negative charge of oxygen atoms have strong Coulombic interactions which contribute to the strength of the hydrogen bond. Lennard-Jones interactions of all the hydrogen bond forming atoms contribute as well. Some important features of hydrogen bonds, such as a slight bonding angle dependence or polarization effects, are not fully captured by force fields. However, the energetics of hydrogen bonds are well produced in MD simulations. Hydrophobic attractions are also generated between non-bonded interactions, even though they are not directly implemented in the force field, but an indirect entropic effect. Atoms in hydrophobic groups such as fatty chains or benzene rings of proteins are modeled as apolar atoms which have a small or no charge. By doing so, interactions with polar groups, especially water molecules, are reduced. The hydrophobic effect has been further described in Chapter 1, Section 1.2.

Water

Water is the solvent in biological systems. Almost all proteins, or parts of a protein (like membrane proteins), perform their function in the presence of water. Not only do water molecules provide the reaction environment, they often also take part in molecular reactions. As a consequence, the water model is of key importance for MD simulations. Water models can determine the quality of MD simulations and even simulation results.

Water is a typical molecule with many-body effects. It has two partially positive hydrogen atoms and also two partially negative 2s-2p hybridization electron orbitals that are not coupled. The molecular structure enables a water molecule to be both hydrogen bond acceptor and donator at the same time. Water molecules are capable of forming hydrogen bonds between themselves which gives water a high melting and boiling temperature. Water molecules also form hydrogen bonds with proteins, nuclear acids and other biomolecules. Interactions between protein and water are the driving forces of protein dynamics.

A handful of water models have been developed and are available nowadays. Some of the water models use three particles to represent the atoms in water molecule, with charge and mass at the same position as the atoms. The SPC water model (58; 59) and TIP3P (60) water model are both well known three-point water models, as shown in Fig. 2.5. They are similar with slight differences in their molecular geometry and charge distribution. Many previous studies show the power of these two water models when used with force fields such as GROMOS and AMBER. Four-point and five-point water models introduce more particles into the model in order to capture the water molecule characteristics more precisely and thus better bulk mimic water behavior. The TIP4P water model (60) is a well-known four-point water model. Combining the TIP4P water model with OPLS-AA force field has been proven to give very accurate structures and energetics in protein simulations.

Hydrogen bonding in water is directional because of the geometry of hybridized electron orbitals of oxygen atoms. Three-point water models can not produce the correct hydrogen bonding direction caused by the electron distribution in water molecules. Four-point water models, such as TIP4P (60), separate the mass and charge of the oxygen atom by introducing one extra dummy particle into the molecular structure, trying to better represent the electron density distribution. The dummy particle of TIP4P sits

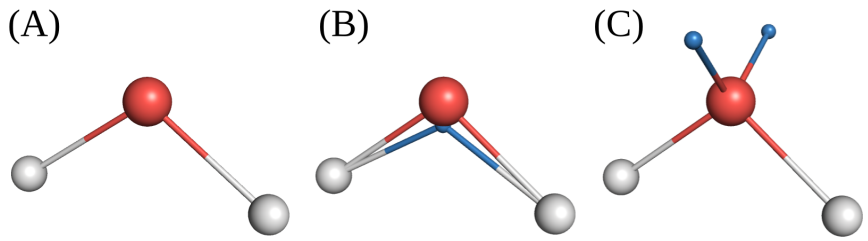

Figure 2.5: Water models. Three-point (A), four-point (B) and five-point (C) water models used in computational simulations are shown. Oxygen atoms are shown in red, hydrogen atoms are shown in white and dummy particles, which only carry charge points, are shown in blue.

close to the oxygen atom on the angular bisector line of the hydrogen-oxygen-hydrogen angle, as shown in Fig. 2.5. It only contains the charge of the oxygen atom while the oxygen particle contains only its weight. Five-point water models, such as TIP5P (61), further separate the negative charge on the water molecule. Two massless dummy particles located at two empty 2s-2p hybridization positions are introduced into the water molecular structure. These two dummy particles can greatly improve the hydrogen bonding geometry.

Choosing a water model in MD simulation is a compromising task which should balance accuracy and simulation costs. As water molecules normally take up more than 80% of the simulation system, introducing one more particle in a water molecule can result in a significant increase of computational cost. Although there are even six-point water models available, the usage of multiple particle water models is quite limited. MD simulations in this project use TIP4P water model with OPLS-AA force field, a combination which has been specifically developed and validated in numerous studies.

2.1.2 MD simulations with Force

Non-equilibrium processes driven by mechanical force are widely observed in biology. The generation of mechanical force by muscle is one good example to demonstrate how molecules response to these mechanical forces. Muscle fibrils in the sarcomere, the force generating unit in muscle, withstand high forces to maintain the integration of the muscle network. Molecules such as titin and myomesin, the latter of which has been studied in this thesis (see Chapter 3, Section 3.2), mediate force propagation and

restore inter-connection structures. Understanding these mechanical processes is still a challenge in biology, yet it is often the only way to uncover biological function.

The biggest challenge in understanding biological mechanical forces is the magnitude of these forces. Common biological forces, such as those generated by molecular motors, are as small as couple of piconewtons. Force signals of such small scale are easily lost in the noise of thermodynamic vibrations, which renders studies on these forces even more difficult. A couple of modern experimental technologies have been developed to study such extremely diminutive interactions. These experiments, such as atomic force microscopy (AFM) and optical tweezers, provide extreme sensitivity combined with a high resolution to allow to probe of mechanical forces in biological systems. Even so, atomic details of molecular transitions like protein unfolding or ligand dissociation pathways under tensile forces are still out of range. For a higher resolution in our understanding of these molecular processes, MD simulations serve as a complementary tool for experiments.

Applying mechanical force onto molecular structures in MD simulations to manipulate a biological process is a relative simple task when having the molecular structure and the force field of the system at hand. MD simulations of processes under forces are of particular value, if validated by single molecule experiments such as those using an atomic force microscope (AFM). MD simulations can complement AFM experiments by explaining the experimental observations. Unique insights have been gained previously and within this thesis (Chapter 3, Sections 3.1.5 and 3.2.1) using these two techniques, (14; 45; 62) and more are likely to be under way.

Atomic Force Microscopy

Atomic force microscopy (AFM) is one of the most important and exciting single molecule experimental methods developed in the last two decades. (13; 63–65) The major component of an AFM is a super sharp tip attached to a soft arm, called cantilever, to probe and record atomic interaction forces. It has a nanometer resolution and a sensitivity at the piconewton scale, which falls into the range of biological interactions.

AFM is an advanced experimental technology. An AFM setup mainly includes three parts, a sample stage, a cantilever and a force detector, as shown in Fig. 2.6. The sample stage is a movable supporting platform for the samples under investigation. It can be

very precisely positioned by a piezomotor. The cantilever is the force probing part of the system. It contains a laser reflection point and a sharp tip. The force detector contains a laser source and a signal detector. Laser is focused onto the cantilever which reflects the laser beam into the signal detector. The atomic force can be measured by the displacement of laser reflection point.

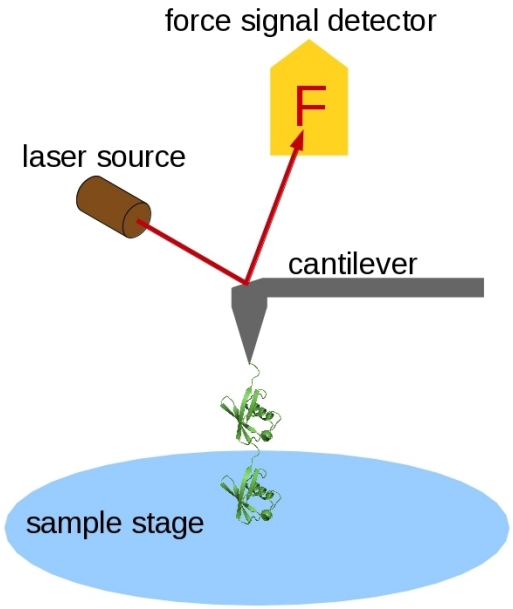

Figure 2.6: Schematic AFM setup. The main components of the AFM are indicated. The cantilever is shown in grey. Its bending causes changes in the laser beam reflection that can be used as force measurements. The sample stage is shown in light blue, and a protein sample is shown in green. The laser beam is shown in red.

AFM is widely used in surface imaging that utilizes atomic interaction between the cantilever tip and the surface atoms on the sample. More importantly for the topic of this thesis, a molecule or molecular system can be attached to the sample stage and the cantilever tip to detect molecular interactions. This allows, among others, to unfold single protein structures using the AFM. (13; 63)

Force-probe MD simulations

Manipulations of molecular structures are in principle easy and straightforward in MD simulations. MD simulations keep track of coordinates of all atoms in the system during the whole simulation time. Applying a mechanical force onto the macromolecule is one

of the standard features of MD simulation software packages. Force-probe MD (FPMD) simulations is one of such simulation techniques to mimic AFM experiments. (66) Mechanical force can be applied to specific parts of the simulated molecule that are of interests, such as a single atom or atom groups. By applying a pulling force, protein unfolding or DNA unwrapping can be monitored uand the required force recorded . FPMD may be the most important way to interpret AFM experiments, as illustrated by many previous studies.

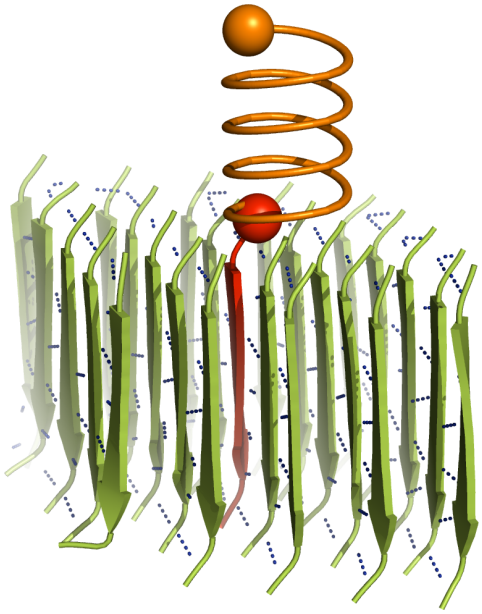

Figure 2.7: Force-probe MD simulation setup. The molecular structure in this setup are protein β-strands (for more details, see Chapter 3, section 3.1.1). A harmonic potential, shown as an orange spring, is applied to the pulled strand which is shown in red. The force application point is shown as a red sphere.

In FPMD simulations, a harmonic potential, represented as a virtual spring as indicated in Fig. 2.7, is assigned to the force application point. During the simulation, the virtual spring steadily moves away from the force application point with constant speed along the specified direction. With the increasing displacement between the virtual spring and the force application point, force is generated according to the Hooke's law. Parameters in FPMD simulations thus include pulling speed, the virtual spring force constant, and the pulling direction, which need to be predefined in the simulation setup.

The most significant difference between FPMD simulations and AFM experiments is the pulling speed. MD simulations are still limited by the computational power available today. The largest time scale that MD simulations can achieve on general-purpose high-performance computers is normally some microseconds. In order to observe molecular structural events such as protein unfolding or DNA double strand unzipping, the pulling speed in MD simulation should be high enough to induce the structural transition within the achievable time scale. This results in a pulling speed in MD simulations in the range of $nm/\mu s$, which is some orders higher than those used in AFM experiments. This difference in pulling speed is reflected by the forces measured by the two methods. More precisely, MD simulations with faster pulling speed tend to measure higher forces.

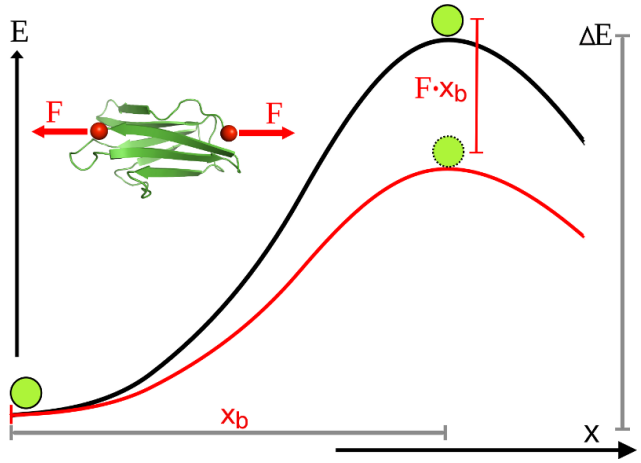

Figure 2.8: Bell model of how mechanical force manipulates the energy landscape along the reaction coordinate. An example of a protein structure under a pulling force is shown in green, with force indicated as red arrows and force application points as red spheres. The system energetic states are indicated as green spheres. Mechanical force lowers the energy barrier by a value of the applied force times the transition distance x_b, assuming the pulling force acts only along the reaction coordinate.

A statistical model, the Bell model, postulated by G.I. Bell, has been successfully used to connect FPMD simulations and AFM experiments, thereby explaining the different results. (67; 68) Both in FPMD simulations and AFM experiments, unfolding of molecular structures is induced by lowering the transition energy barrier using mechanical force, as depicted in Fig. 2.8. The mechanical force needed to drive the system to

its transition state depends on the pulling speed, as follows:

$$F(v) = \frac{k_b \cdot T}{x_b} \cdot \ln v - \frac{k_b \cdot T}{x_b} \cdot \ln v_o \ , \tag{2.7}$$

where k_b is the Boltzmann constant, T is the system absolute temperature, x_b is the distance between initial and the transition state along the reaction coordinate, which is assumed to be the pulling direction in both FPMD simulations and AFM experiments. v_o is the system's intrinsic speed to reach the transition state, it depends on the height of the energy barrier and also x_b.

Bell's model established a connection between simulations and experiments and enables theoretical comparison. Because rupture forces of protein or DNA unfolding depend logarithmically on the pulling speed, x_b can be obtained from the linear dependency of the force on $\ln v$, and the experimental and computed x_b can be compared. Force profiles from these two methods also can be qualitatively compared and were found to agree with each other, which allows to explain AFM experimental observations with the details observed in MD simulations. (45; 69; 70) This connection is key to our understanding of biomolecular mechanics, which is the general aim of this thesis. On one hand, FPMD simulations can be verified by AFM experiments; on the other hand, FPMD simulations can serve as a high resolution approach to interpret AFM experiments. FPMD simulations, in combination with AFM experiments carried out by collaborators, have been performed for silk protein and myomesin (see Chapter 3, Sections 3.1.5 and 3.2.1).

Force-clamp MD simulations

Force-clamp MD (FCMD) simulations are also designed to mimic AFM experiments, namely those carried out with constant loading force. Practically, the only difference between FPMD and FCMD is the way of force application. FCMD uses a constant pulling force during the simulation, while the pulling force varies in FPMD simulations, depending on and reflecting the events of molecular transitions. They have other simulation parameters in common.

Quenching macromolecules such as unfolded proteins to a constant pulling force brings their structures to other equilibrium states, which allows monitoring their refolding. (18) This simulation method is important for deducing the kinetics and thermodynamics of biomolecules. It also is highly relevant considering that molecules can

be under quasi-constant biological stress in the cell. It is also an important way to understand how molecular structures cope with external forces, either by rearranging structural building blocks or by carrying it by intramolecular interactions. Applying a constant force to a structure also allows to understand how mechanical stress is distributed and propagated (see next paragraph). It therefore is another method of choice for studying both silk and myomesion in this thesis.

Force Distribution Analysis

Mechanical properties play a crucial role for structural molecules such as muscle proteins or fibrous proteins like silk. How these protein distribute mechanical stress is an important measure for understanding their mechanical robustness. Force distribution analysis (FDA) is the technique of choice to discover stress distribution in molecular structures. (54) FDA is also an ideal method for investigating molecular stress triggered by ligand binding and DNA-protein interactions. The pattern of stress distribution is the pathway of mechanical signal transduction.

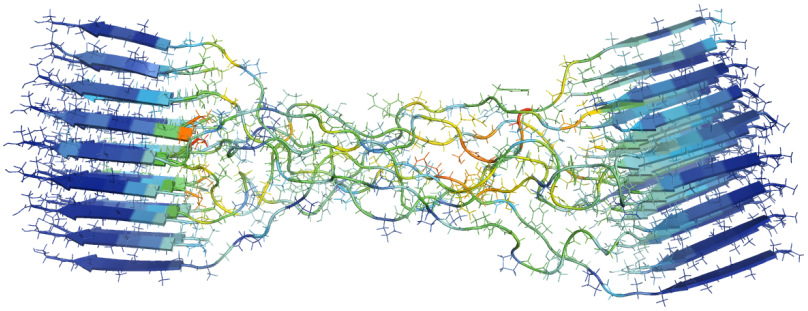

Figure 2.9: Force distribution analysis on a protein structure. The silk composite unit model consisting of two β-sheet crystals and a random-structure amorphous phase is shown as a example. Red color represents high mechanical stress in the figure while blue represents lower stress. For details see Chapter 3, Section 3.1.3.

Force distribution analysis (FDA) is a new method developed in the group of Dr. F. Gräter. It has been already applied to different systems and could yield insightful results. (54; 71; 72) It is reminiscent of the mechanical analysis used in civil engineering. A result from FDA is shown in Fig. 2.9.

FDA is based on FCMD (see above). The principle of FDA is to equilibrate the molecular structure in two mechanical states by applying different constant mechanical

pulling forces which are low enough to prevent the structure from rupturing. Force between each atom pair is collected at every simulation step and averaged during the whole simulation time. The absolute value of the force difference between two states of all the atom pairs are summed up in the whole simulation. Finally, the sum of absolute value of force differences as given by the force field (Equation 2.2) is mapped onto the corresponding atoms to give rise to the force distribution pattern. Details of the FDA algorithm is give in the section where FDA has been used, namely Chapter 3, Sections 3.1.1, 3.1.2 and 3.1.3.

2.2 Finite Element Method

2.2.1 Introduction

The classical molecular mechanics description (Section 2.1) is successful in solving simple structural problems but becomes very complex and computationally costly for large systems. This limitation of classical mechanics for structures can be overcome by an approximation, the finite element method (FEM).

FEM, also called finite element analysis (FEA), is the most successful continuum modeling method. Since first introduced in the mid twentieth century (73), FEM has been applied to all possible aspects of physical modeling, including vibrational analysis, structural analysis, fatigue analysis, fluid dynamics dynamics, heat transfer analysis, and others. (74–76) It is a steadily revised method that is adjusted to the growing computational capabilities.

The concept of FEM is to divide a complex continuum mechanics problem into simple units whose properties are already well understood. Not only can FEM divide structures made of different materials into different domains, but also divide odd structure to regular or symmetrical units. By doing so, FEM transforms the solving of a mechanical problem with complex components into assembling the solutions of regular and relatively simple questions. One example of solving the stress distribution in a car wheel is demonstrated in Fig. 2.10. This thesis uses FEM to study the structural mechanics in silk fibers.

The basic units of the structure after discretization are the elements in FEM. One element in FEM is constructed by nodes which are also the connections to other elements, as shown in Fig. 2.11. A stiffness matrix is used to describe the behavior of an

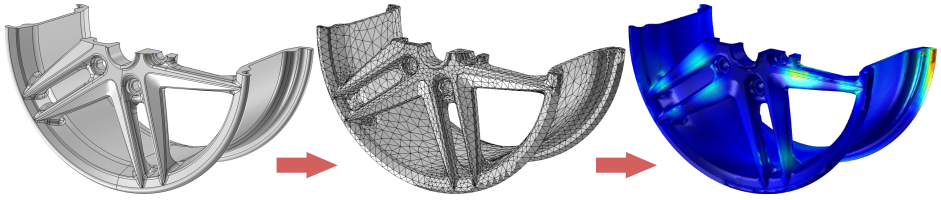

Figure 2.10: Stress distribution analysis in a car wheel. Structures like a car wheel are not symmetrical in shape and consist of different materials with complicated connections between the parts, which makes it impossible to use pure classical methods to solve the stress distribution. FEM discretizes a car wheel into parts and solves the stress in each part using given conditions and connections to finally obtain an approximate solution. As shown in the Figure, high stress concentrates in red areas in the structure, while blue areas have lower stress. Figure is generated with the COMSOL FEM commercial package (http://www.comsol.com).

element by vectorizing its nodes' respond under load. This matrix contains the properties of the elements including elastic modulus, shearing modules and others. Commonly, there are different kinds of elements used in applications, such as beam elements, rod elements, mass elements. By correctly choosing element stiffness matrices, both linear and non-linear systems can be modeled and studied.

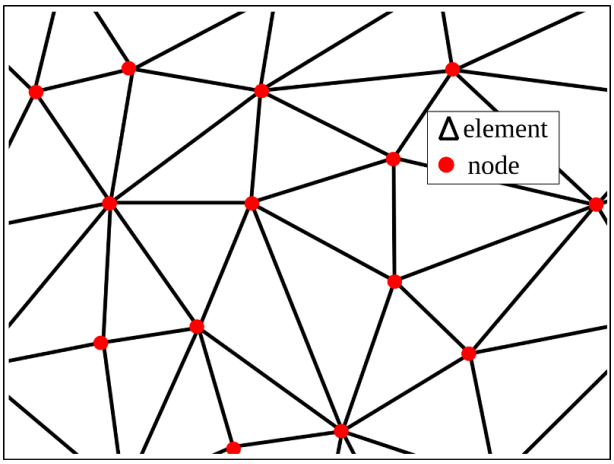

Figure 2.11: Example of elements and nodes in FEM. Triangle 2D plain elements are shown in this Figure. Nodes are highlighted as red circles.

FEM creates a mesh map in the studied system by dividing the structure into a combination of elements, as shown in Fig. 2.10. The sizes of elements depends on the

studied area and problem. A higher density of nodes and thus smaller size of elements gives higher accuracy of solutions in FEM, though at higher computational cost. If some area in the investigated system is more sensitive to mechanical loads, such as the area subjected to massive changes in stress or initial points of structural facture, smaller elements are normally used to achieve requirements of accuracy. It is also normal to re-mesh the structure using different elements of smaller size to recalculate the solution.

Predefined conditions of physical problems in FEM are the boundary conditions. These conditions include how loads are applied to the system and how the system is subjected to some special restraints or constraints etc. They also include known quantities such as some physical laws of equilibrium or constancy in mass or energy.

2.2.2 Theory

The core concepts of FEM are the elements, their types and their properties. The properties of FEM elements, such as elastic modulus, shearing modulus and etc., are the basis of mechanical properties of FEM models. The response of an element to external load are described by a so-called stiffness matrix which contains the vectorized elements' behavior under mechanical load, namely the direction and amplitude of nodes displacement, physical inflow and outflow, and so forth.

For most FEM software package, all kinds of elements for commonly found materials such as metals (e.g. steel or cooper) are provided by experimental measurements. New elements for constructing new materials should be implemented and parameterized by the material scientists. In this thesis, the silk fiber is studied by FEM. Thus, new elements needed for building a silk fiber are needed, as these are not provided by the FEM commercial software. In order to create these new elements, more precisely to determine element stiffness matrices, MD simulations are performed on all-atom models of silk protein subunits. Mechanical properties of silk protein subunits are then incorporated into stiffness matrices of new FEM elements for building and simulating continuum silk fibers. For more details, see Chapter 3 (Sections 3.1.3 and 3.1.4).

Determining a stiffness matrix for the whole system is a process of constructing polynomial functions. Because all the connections between elements are located at their nodes, the behavior of elements depends on the states of these nodes. A stiffness matrix of an element is developed from the equilibrium states equations of all the nodes concerned. Its complexity grows dramatically with the increase of the node number

and dimensions. A stiffness matrix can be developed using a simple model, as shown in Fig. 2.12. The principle of developing stiffness matrices for more complex elements is the same.

Fig. 2.12 shows a simple two-dimensional FEM element. The element is a harmonic spring with one node at each end. The nodes can only move in x and y directions. Given a certain tension in the spring, the force at each node in both x and y directions depending of both nodes' displacements in each direction can be summed up by the equations below:

$$\begin{aligned}
F_{1x} &= k_{1x-1x}r_{1x} + k_{1x-1y}r_{1y} + k_{1x-2x}r_{2x} + k_{1x-2y}r_{2y} \\
F_{1y} &= k_{1y-1x}r_{1x} + k_{1y-1y}r_{1y} + k_{1y-2x}r_{2x} + k_{1y-2y}r_{2y} \\
F_{2x} &= k_{2x-1x}r_{1x} + k_{2x-1y}r_{1y} + k_{2x-2x}r_{2x} + k_{2x-2y}r_{2y} \\
F_{2y} &= k_{2y-1x}r_{1x} + k_{2y-1y}r_{1y} + k_{2y-2x}r_{2x} + k_{2x-2y}r_{2y}
\end{aligned} \tag{2.8}$$

where F_{1x} and F_{1y} are forces on node 1 and F_{2x} and F_{2y} are forces on node 2 in both x and y directions, respectively; r_{1x}, r_{1y}, r_{2x} and r_{2y} are the displacements of both nodes in x and y directions; k_{i-j} is the stiffness coefficient of force because of the r_j displacement with respect to r_i (x or y positions). For example, k_{1x-1y} denotes the spring force constant in y direction of node 1 with respect to its displacement in x directions.

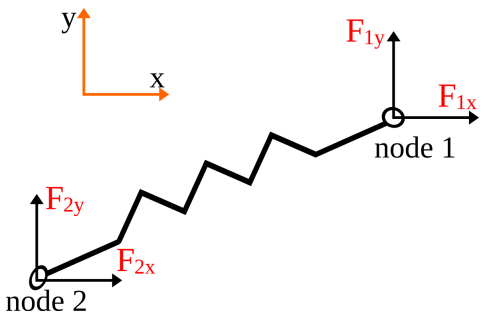

Figure 2.12: An example of a 2-dimensional FEM spring element. This element consists of only one harmonic spring with a node at each end. The spring has a force constant of k The nodes are allowed to move in the XY plane only. Force acting on each node in each direction can be calculated by the Hooke's law based on the displacement of the node in each direction, which can be summed up into polynomial equations.

The equations of the forces acting on each node of the example element shown above can be written in the following way:

$$\begin{pmatrix} F_{1x} \\ F_{1y} \\ F_{2x} \\ F_{2y} \end{pmatrix} = \begin{bmatrix} k_{1x-1x} & k_{1x-1y} & k_{1x-2x} & k_{1x-2y} \\ k_{1y-1x} & k_{1y-1y} & k_{1y-2x} & k_{1y-2y} \\ k_{2x-1x} & k_{2x-1y} & k_{2x-2x} & k_{2x-2y} \\ k_{2y-1x} & k_{2y-1y} & k_{2y-2x} & k_{2x-2y} \end{bmatrix} \cdot \begin{pmatrix} r_{1x} \\ r_{1y} \\ r_{2x} \\ r_{2y} \end{pmatrix}$$
$$= K_{i-j} \cdot \begin{pmatrix} r_{1x} \\ r_{1y} \\ r_{2x} \\ r_{2y} \end{pmatrix} \tag{2.9}$$

where the matrix consisting of all force constants, k_{i-j}, is the stiffness matrix of the example element, simplified as K_{i-j}. If the harmonic spring has a force constant of k and has an angle to x axis of θ, K_{i-j} has a known form (74):

$$K_{i-j} = k \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta & -\cos^2\theta & -\cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta & -\sin\theta\cos\theta & -\sin^2\theta \\ -\cos^2\theta & -\cos\theta\sin\theta & \cos^2\theta & \cos\theta\sin\theta \\ -\sin\theta\cos\theta & -\sin^2\theta & \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} \tag{2.10}$$

The stiffness matrix, K_{i-j}, contains the properties of the spring element. Not only it contains the spring force constant, but also the geometrical response of the spring under mechanical load. It emerges that the stiffness matrix becomes very complex, when more nodes and dimensions are taken into account. Modern FEM computers and software greatly simplified the development of stiffness matrices for different materials. Modern commercial and open source FEM software packages contain a variety of different stiffness matrices for different elements. Materials, such as iron, copper or glass, have been successfully modeled using experimental data as a guide.

As mentioned in the earlier sections, in this thesis, FEM was combined with MD simulations. Stiffness matrices for protein materials were developed based on the parameters obtained from MD simulations. Using MD data, which were validated by experimental results, allows for avoiding the usage of low-accuracy and scarce experimental parameters. The combination of MD simulations and FEM enable us to tackle the relationship between the nanoscale architecture and macroscopic properties of silk proteins (see Chapter 3, Section 3.1). This is the first such attempt in material engineering, to our knowledge.

3

Protein Materials

3.1 Silk

3.1.1 Mechanical Response of Silk Crystalline Units from Force Distribution Analysis

The outstanding mechanical toughness of silk fibers is thought to be mainly brought about by embedded crystalline units acting as crosslinks of silk proteins in the fiber. Here, we examine the robustness of these highly ordered β-sheet structures by molecular dynamics simulations and finite element analysis. Structural parameters and stress-strain relationships of four different models, from spider and *bombyx mori* silk peptides, in anti-parallel and parallel arrangement, were determined and found to be in good agreement with X-ray diffraction data. Rupture forces exceed those of any previously examined globular protein many times over, with spider silk (poly-alanine) slightly outperforming *bombyx mori* silk ((Gly-Ala)$_{\mathrm{n}}$). All-atom force distribution analysis reveals both intra-sheet hydrogen bonding and inter-sheet side-chain interactions to contribute to stability to similar extent. In combination with finite element analysis of simplified β-sheet skeletons, we could ascribe the distinct force distribution pattern of the anti-parallel and parallel silk crystalline units to the difference in hydrogen bond geometry, featuring a in-line or zigzag arrangement, respectively. Hydrogen bond strength was higher in anti-parallel models, and ultimately resulted in higher stiffness of the crystal, compensating the effect of the mechanically disadvantageous in-line hydrogen bond geometry. Atomistic and coarse-grained force distribution patterns can thus explain

differences in mechanical response of silk crystals, opening up the road to predict full fiber mechanics. (72)

Introduction

Silk proteins build up the most tough yet elastic fibers known (77; 78). Relating the extraordinary fiber mechanics to the underlying molecular architecture is a requisite for rationally altering properties of natural silk fibers and for designing artificial analogues. Understanding the intricate correlation of the elastic response with the complex nano-scale protein structure of silk fibers, however, has remained a challenge.

Silk proteins produced from different insect species, the most commonly studied representatives of which are spider drag line silk and cocoon silk from the silk worm *bombyx mori*, share a common protein sequence and fiber architecture (Fig. 3.1A and B). Repeat units of six to nine amino acids in length, from alanine or from alternating alanine and glycine residues, for spider and *bombyx mori* silk, respectively, build up highly ordered β-sheet rich crystalline units (Fig. 3.1C) (79–81). These crystals are connected by and embedded into an amorphous matrix of disordered proteins from non-repetitive sequence motifs. The ratio of β-sheet versus matrix forming motifs in the silk block copolymer sequence as well as the spinning process define the relative amount of β-sheet. The transition from highly ordered β-sheet crystals to the disordered region appears to be blurt and to involve semi-crystalline regions (82). Crystals of a few nanometers in size with highly ordered β-strands oriented along the fiber axis have been found by X-ray analysis to constitute 10-15 % of silk fibers (83), with the overall β-sheet content amounting to 40–50 % for spider and silkworm silk, respectively (84). While the ratio of anti-parallel and parallel β-sheets remains largely unknown for most silk types, solid-state NMR experiments suggested roughly a 2:1 ratio of anti-parallel to parallel conformations in wild silkworm fibers, independent of fiber stretch (85).

The crystalline units crosslink the protein chains in the fiber via hydrogen bonding. In a stretched fiber, the external force propagates along the fiber axis by straightening the disordered protein chains and subjecting the cystalline β-sheet regions to a tensile force along the β-strand axis. Their elastic modulus in silkworm silk has been recently determined by X-ray diffraction experiments (86). The extraordinary toughness of silk fibers is assumed to be encompassed by the strong and stiff crystalline units, taking up the mechanical load in stretched fibers as stiffness attracts force and thereby protecting

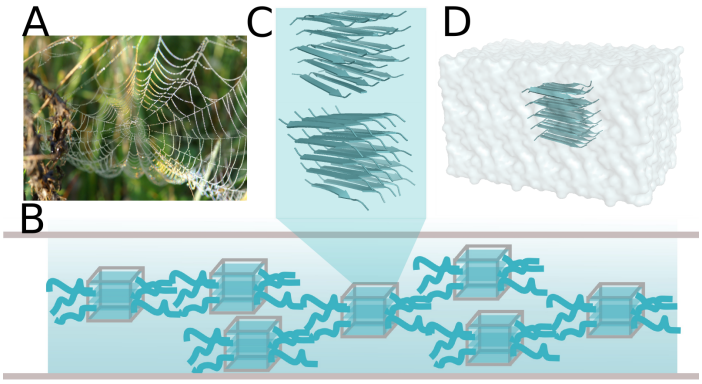

Figure 3.1: Silk structure and simulation systems. (A) Silk is one of the toughest materials known, and has evolved in nature for manifold purposes, from predation to protection. In courtesy of Michael Goduscheit. (B) Silk proteins form an amorphous matrix of disordered segments, into which crystalline units are embedded and serve as crosslinks, depicted as boxes. (C) All atom models of silk crystalline units in cartoon representation, in both anti-parallel (upper) and parallel (lower) arrangement. (D) Representative MD simulation system of crystalline unit models, water is shown in transparent surface representation.

against failure. Theoretical studies that focus on mechanical properties of silk have been restricted to simple models to date (87). Recent studies on the related amyloid fibers and small β-strand topologies have shed light onto the mechanics of β-sheet dominated fibers (88; 89). The molecular basis of the rupture strength and stiffness of β-sheet stacks as they occur in silk is currently unknown. What are the forces necessary to fracture silk protein crystals? How does the force distribute through such a structure and what are the force-bearing molecular interactions?

We here present atomic-detail models of the crystalline units of *bombyx mori* and spider drag line silk, in both parallel and anti-parallel arrangement. We determine and compare the rupture forces, stiffness in terms of a backbone pull-out resistance and internal force distribution from molecular dynamics simulations. We then develop simplified model structures to dissect the contributions of hydrogen bond geometry and strength to overall strength of the β-sheet. We find the weaker hydrogen bonding in parallel β-sheets to be compensated for by the stiffer geometry with inclined hydrogen bonds with respect to the anti-parallel analogue. Focusing onto idealized models of the crystalline units as the major stabilizing building block of silk allows us to determine the mechanics of silk-like crystals detached from the complex multi-layer organization

of a full fiber. This study thereby presents a first step towards a comprehensive understanding of the molecular ingredients of silk fiber mechanics.

Methods

Modeling and equilibration

We here focus on the mechanical response of the β-sheet rich units in silk fibers. In the absence of a high-resolution structure, we modelled highly ordered crystalline units on the basis of the available substantial experimental data. We built all-atom models composed of the repeat units found to be present in spider drag-line silk and cocoon silk (77; 79; 90), AAAAAAAA (denoted AA model) and GAGAGAGAAS (denoted GA model), respectively. Since silk fibers presumedly consist of a mixture of possible β-sheet arrangements (85), both parallel and anti-parallel models were constructed, denoted here as AA_p, AA_{ap}, GA_p, and GA_{ap}. We arranged five layers of β-sheets, each consisting of five β-strands of the respective sequence, such that the model exhibits optimal hydrogen bonding in the absence of steric repulsion. We found 0.55 nm and 0.47 nm as inter-strand distances for the AA and GA models, respectively, to be a reasonable choice. We obtained models approximately $2.5 \times 2.5 \times 3.0\,\mathrm{nm}^3$ in size, in agreement with X-ray experiments that found crystals to be a few nanometers in size in each direction (83). Since the detailed number of strands in a silk crystalline unit is currently unknown and might vary within a fiber and between different silks, we performed additional simulations of larger crystals. Uncharged peptide termini were chosen to mimic the situation in a silk fiber, in which the β-strands do not terminate but reach out into the amorphous region. We did not include the disordered parts of the silk protein into the models, allowing us to focus on the mechanical properties of the force-bearing crystalline units exclusively.

We used the Gromacs 3.3.1 package (53) for all subsequent molecular dynamics (MD) simulations, and the OPLS-AA force field (40) for the protein. Simulation boxes of approximately $6.4 \times 6.8 \times 6.4\,\mathrm{nm}^3$ were used. Periodic boundary conditions were employed to remove artificial boundary effects. We chose a cutoff of 1.0 nm for non-bonded interactions, and the Particle-Mesh Ewald method (91) to account for long-range electrostatics interactions. In order to increase the simulation time step, we used LINCS (92) to constrain all bond vibrations. The time step was 0.002 ps. Simulations were performed in the NpT ensemble with a temperature of $T = 300\,\mathrm{K}$ and a pressure

of $p = 1$ bar in all the simulations. We used Nosé-Hoover (93; 94) temperature coupling with a coupling time constant $\tau_T = 0.1$ ps, and Berendsen (95) pressure coupling with a coupling time constant of $\tau_p = 1$ ps.

We relaxed the modelled crystalline units by energy minimization and short MD simulations in vacuum. The models were subsequently solvated in TIP4P water (60). In a silk fiber, cyrstalline units are surrounded by amorphous peptide chains as well as water molecules. Solvation with water was chosen to mimic this environment, a condensed and polar phase, in an efficient and more realistic way than vacuum. The solvent included Na and Cl ions with a concentration of 0.1 mol/liter, resulting in a system size of ∼35,000 atoms. After energy minimization using the steepest descent method, we performed 500 ps position restrained simulations to further relax our simulation systems, subjecting each protein atom to a harmonic potential with a force constant of 1000 kJ mol^{-1} nm^{-2}. Each model was then fully equilibrated for 10 ns. Energy and coordinates of the simulation systems were collected every 1000 time steps. The resulting equilibrated simulation systems served as starting points for Force probe and Force clamp MD simulations (see below).

Force-probe MD simulations

To assess the mechanical resistance of the four different silk models, we performed force-probe molecular dynamics simulations (96). Final equilibrated structures obtained from the free MD simulations of the four systems were exposed to an external stress to monitor rupture. More precisely, the terminal residue of the central strand was subjected to a pulling force along the strand direction by moving a spring with constant velocity away from the silk block. A counter force was applied to the center of mass of the protein to prevent translation of the protein by the pulling force. We alternatively also considered applying the counter force to all strands except the pulled one, and obtained the same mechanical response. In a silk fiber, the complex mechanical stress pattern acting onto the silk crystal is determined by how the individual silk peptide chains are embedded and connected within the amorphous matrix. As the simplest scenario, we chose the central strand out of the unit of 5 x 5 strands to be pulled. This maximally reduces the effect of the protein-water interface, which does not exist as pronounced in the more densely packed natural silk fiber. However, we do not expect the rupture forces and force distribution patterns to largely depend on this choice. The pulling velocity of 0.2 nm ns^{-1}, and a spring constant of 500 kJ mol^{-1} nm^{-2} was used.

To accomodate the protein also after rupture, we increased the box dimension along the pulling direction to 12.0 nm, resulting in a system size of ∼55,000 atoms. The simulated time of all the models to monitor their full rupture was 20–25 ns, depending on the resistance against rupture. These non-equilibrium MD simulations used the same simulation parameters as the equilibrium simulations (see above). A representative simulation system is shown in Fig. 3.1D. Mechanical response is characterized by the rupture force, which is the maximal force observed for rupturing the cystal, F. Stiffness was measured by a quantity we here denote backbone pull-out resistance, R_{bb}. In analogy to the elastic modulus typically given to measure stiffness of a material, R_{bb} is defined as

$$R_{\mathrm{bb}} = \frac{\text{stress}}{\text{strain}} = \frac{F/A}{\Delta\, l/l}, \tag{3.1}$$

where F is the force acting on the strand, and A is the cross section area of the interacting strands. We defined the cross-section as the area covered between adjacent strands, which gives $A = 1\,\mathrm{nm}^2$. In a full silk fiber, the external stress applied to the fiber's area distributes highly inhomogeneously through the fiber. Depending on the entanglement of the chains in the amorphous matrix, tensile stresses in the chains, and thus the forces acting on the individual strands in the crystalline units, differ from each other. Thereby, the applied stretching force acting on only one or a few peptides in the silk crystals, are effectively translated into a shear force within the β-sheet arrangement. However, in analogy to previous experimental measurements of stiffness of crystalline units in terms of the Young's modulus E (86), we here defined a backbone pull-out resistance to measure stiffness. How exactly the amorphous matrix structural properties determine force distribution onto the β-sheet rich phase will be subject of future investigations, and will help to find a remedy for this definition.

Force-clamp MD (FCMD) simulations and force distribution analysis

To determine the internal strain of the crystalline units prior to rupture, we employed a newly developed force distribution analysis. We here shortly outline the basic concept. Details have been published elsewhere (54).

Force distribution analysis is based on Gromacs-3.3.1 modified to write out forces F_{ij} between each atom pair i, j. Forces include individual bonded (bond, angle, dihedral) and non-bonded (electrostatic, van der Waals) terms below the cutoff distance of 1 nm. The force between an atom pair is represented as the norm of the force vector and thus is scalar, attractive and repulsive forces are distinguished by oppoite signs. As we

consider the direct force between each atom pair, the equilibrium force can be different from zero, even for the theoretical case of a system without motion. Atomic forces, i.e. the sum over all force vectors acting on a single atom would instead average out. We hereby obtain the advantage to be able to observe strain propagation even through stiff materials, such as the silk crystal, where forces propagate without causing major atomic displacement.

Forces were monitored in the relaxed state during equilibrium simulation (EQ) and in the strained state during FCMD simulation (FC). In the FCMD simulations, a constant external force of 1660pN was applied to the terminal amino acid along the strand direction as described above. The four silk models were equilibrated for 20 ns in total in the strained and relaxed state in two independent FCMD / equilibrium simulations each. Average forces were written every 10 ps. To obtain converged averages, forces were afterwards averaged over the complete simulation time.

A change in pairwise force reflects internal strain and thus is considered as a measure for load-bearing interaction. Consequently, the force propagation pattern becomes visible when observing the differences in forces F_{ij} between strained and relaxed state, defined as:

$$\Delta F_{ij} = F_{ij}^{\mathrm{FC}} - F_{ij}^{\mathrm{EQ}} \tag{3.2}$$

where F_{ij}^{FC} is the force between atom i and j in the strained state and F_{ij}^{EQ} is the force in the relaxed state. The mechanical coupling of a single atom with respect to all other atoms is then defined as the absolute sum of changes in force ΔF_j:

$$\Delta F_j = \sum_i |\Delta F_{ij}| \tag{3.3}$$

Individual hydrogen bond forces were obtained from summing up over pair-wise Lennard Jones and Coulombic forces between all atom pairs of the C=O and N-H groups.

Finite element analysis

To predict the effect of hydrogen bond geometry on the mechanical response, we developed a simplified β-skeleton model for the anti-parallel and parallel β-sheets on the basis of the all-atom models (Fig. 3.5). Distances and hydrogen bond geometry were directly taken from the all-atom models. The geometries were imported into a common beam frame analysis finite element software where the backbone, the hydrogen bonds, and their short connections to the backbone, i.e. the C=O and N-H groups,

were modelled by rigid-jointed, linear-elastic beam elements with circular cross-sections. For each of the elements an individual bending stiffness EI and a tensile stiffness EA were defined where A is the cross-sectional area and I denotes the geometrical moment of inertia (second moment of area). The bending stiffness for the backbone element was calculated as

$$EI = p \cdot k_B T \tag{3.4}$$

where p is the persistence length of a peptide, k_B is the Boltzmann constant and T is the temperature, here 300 K. From the worm-like chain model (97) of a peptide that only takes conformational flexibility into account, a persistence length of 1.2 nm was previously obtained and used here (98). A bending modulus of $EI = 58\,\mathrm{pN\,nm^2}$ was obtained.

The tensile stiffness of the backbone element is defined as

$$EA = \frac{F}{\Delta l / l} \tag{3.5}$$

In FCMD simulation, we obtained a strain of $\Delta l/l = 0.06$ at a force of 1660 pN, resulting in a stretching modulus of $EA = 23300\,\mathrm{pN}$.

EA of the hydrogen bond elements, was calculated on the basis of the hydrogen bond potential of the OPLS-AA force field, namely as the second derivative of the energy summed up over all Lennard-Jones and electrostatic interactions between the atoms in the C=O and H-N groups (Fig. 3.2). As a first approximation, the curvature at the potential energy minimum, i. e. at an O-H distance of 0.195 nm, was taken as the tensile stiffness EA of both the parallel and anti-parallel skeletons, β_p and β_ap, resulting in $EA = 1797\,\mathrm{pN}$. A bending modulus of hydrogen bonds cannot be straightforwardly obtained from the all-atom models. We used an effective radius of 0.06 nm for hydrogen bonds to obtain EI from EA, using the relation $r = 2\sqrt{EI/EA}$. This radius is a measure for the relative bending versus tensile stiffness and was chosen similarly to the one obtained for the backbone. We obtained $EI = 1.62\,\mathrm{pN\,nm^2}$ for hydrogen bond elements. The elastic modulus of the hydrogen bonds then is $E = 159\,\mathrm{GPa}$. However, in fact, when calculating the hydrogen bond length distribution in the relaxed silk crystals, we found that hydrogen bonds in all-atom parallel units were longer (0.203 nm on average) than those in anti-parallel models (0.195 nm on average), resulting in a lower hydrogen bonding energy in parallel models.

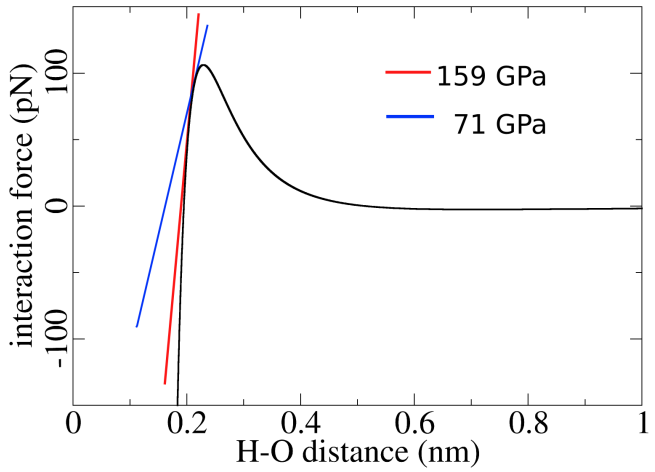

Figure 3.2: Hydrogen bond force profile. The first derivative of the potential energy between C=O and N-H atom groups in OPLS, including electrostatics and Lennard-Jones interactions, is shown. The slope of the curve, indicated for the potential minimum at an equilibrium length of 0.195 nm, and for extended average hydrogen bond length of 0.203 nm, gives the hydrogen bond stretching modulus.

To account for the lower strength of hydrogen bonds in the parallel model an additional parallel β-skeleton was considered, $\beta_{\mathrm{p,weak}}$. At a distance of 0.203 nm, the interaction potential from the OPLS-AA force field gives a stretching modulus of EA=802 pN (Fig. 3.2). The bending modulus then reduces to $EI = 0.72\,\mathrm{pN\,nm^2}$. The elastic modulus of hydrogen bonds in $\beta_{\mathrm{p,weak}}$ then is $E = 71\,\mathrm{GPa}$.

For the C=O and N-H elements, parameters similar to the backbone element were applied. These elements had an only minor effect on the mechanics of the skeleton. Using the parameters described above and the structures shown in Fig. 3.5, a finite element analysis was performed to calculate the dislocation of the central strand in each skeleton upon application of a force of 1660 pN, as in previous MD simulations. This analysis gave effective Young's moduli E of the whole β-skeletons which were directly compared to MD results. First, the anti-parallel and parallel β-skeletons with an identical hydrogen bond strength, $EA = 1797\,\mathrm{pN}$ for β_{ap} and $\beta_{\mathrm{p,strong}}$, were compared to focus on effects based on geometry only. Secondly, different hydrogen bond strength, $EA = 1797$ for β_{ap} and $EA = 802$ for $\beta_{\mathrm{p,weak}}$ were applied (see above). We note that the

simple finite element models proceeded from geometrical as well as material linearity. Non-linear effects, in particular for the hydrogen bonds, might be incorporated in future analysis. Using finite element analysis, the backbone pull-out resistance, as defined in Eq. 3.1 above for the all-atom models, was determined and compared to the MD results.

Results and Discussion

Structural validation of models for silk crystalline units

The toughness of silk fibers is brought about by the β-sheet rich crystalline units which crosslink the protein chains. They consist of a poly-alanine or a GAGAGAGAAS sequence, in spider and silkmoth silk, in an anti-parallel or parallel arrangement of the strands. While the three-dimensional structure of any of these crystalline units remains largely unknown, the inter-strand spacing within the crystal has been measured by X-ray diffraction of silk fibers (83). We constructed four different models, denoted AA_p, AA_{ap}, GA_p, and GA_{ap} for the parallel and anti-parallel spider and silkmoth silk, respectively. We here only considered purely parallel and anti-parallel arrangements within and between β-sheets, and expect mixtures thereof to show intermediate behavior. Arranging the strands such that hydrogen bonding and sidechain packing is optimized does not leave any other degrees of freedom. We equilibrated these models in water in molecular dynamics (MD) simulations. The models show remarkable agreement with the experimental β-strand spacing in the crystal, as shown for GA in Table 3.1. In addition, they show high conformational stability during the 10 ns equilibration, with a root-mean-square deviation from the initial model not higher than 0.15 nm in all models. Our idealized crystal units can be considered representatives of the most regular β-sheet rich regions which occur with various degrees of regularity in real silk fibers.

Characterization of mechanical response

We characterize the mechanical response of the four different models by determining their stress-strain relationship and backbone pull-out resistance (Fig. 3.3). The way the force acts onto a crystalline unit in a stretched silk fiber is complex and varies from unit to unit. We here simplify the stress application by subjecting the terminus of the central strand to a pulling force arising from a virtual spring moved along the strand axis with constant velocity, as schematically shown in Fig. 3.3C.

The stress-strain curves of the AA models are shown in Fig. 3.3A. The elastic response is mostly linear up to the rupture of the pulled strands. The slope of the

Table 3.1: Structural and mechanical properties of crystalline units compared to experimental data. d_1 , distance between two neigbouring C_α along one peptide; d_2 , distance between two β-sheet layers; d_3 , distance between neigbouring strands in the same β-sheet layer. R_{bb}, backbone pull-out resistance from force-probe MD simulations and experiment (86). R_{FEM}, backbone pull-out resistance of β-skeletons from finite element model, with the same hydrogen bond strength, $\beta_{\mathrm{p,strong}}$ and β_{ap}, or with the weaker strength found for parallel crystals, $\beta_{\mathrm{p,weak}}$.

	AA_{ap}	AA_p	
d_1 (nm)	0.351 ± 0.005	0.331 ± 0.008	
d_2 (nm)	1.043 ± 0.023	1.082 ± 0.026	
d_3 (nm)	0.477 ± 0.011	0.485 ± 0.016	
R_{bb} (GPa)	68.1 ± 1.5	28.9 ± 1.4	
R_{FEM} (GPa)	$\beta_{\mathrm{ap}} : 26.3$	$\beta_{\mathrm{p,weak}} : 17.1$	$\beta_{\mathrm{p,strong}} : 24.3$
	GA_{ap}	GA_p	GA experiment
d_1 (nm)	0.351 ± 0.005	0.330 ± 0.008	0.348
d_2 (nm)	0.905 ± 0.023	1.025 ± 0.025	0.970
d_3 (nm)	0.477 ± 0.013	0.480 ± 0.016	0.466
R_{bb} (GPa)	86.8 ± 2.5	27.3 ± 1.8	26.5 ± 0.8

stress-strain relationship of AA_{ap} gives an backbone pull-out resistance of 67.2 ± 3.0 GPa (Table 3.1), which is more than twice of the corresponding parallel crystalline unit AA_p(28.6 ± 2.2 GPa). The stiffness as measured by the backbone pull-out resistance correlates for the two β-sheet arrangements of AA with the forces upon which rupture occurs. With a rupture force of 4074 pN, AA_{ap} clearly outperforms AA_p (2988 pN, Fig. 3.3D). To assess the effect of the loading rate onto the obtained stress-strain relation, we performed additional simulations with a 10-fold higher loading rate, using a spring constant of $5000\,\mathrm{kJ\,mol^{-1}\,nm^{-2}}$. We obtained the same backbone pull-out resistance. We can therefore assume the stress-strain response, involving only slight sub-nanometer structural rearrangements, to be independent from the loading rate or magnitude of force applied, in contrast to the load-dependent forces inducing complete rupture.

Similarly, linear stress-strain relations were found for the analogous GA crystalline units, with a higher stiffness for the anti-parallel (red curve in Fig. 3.3B) over the

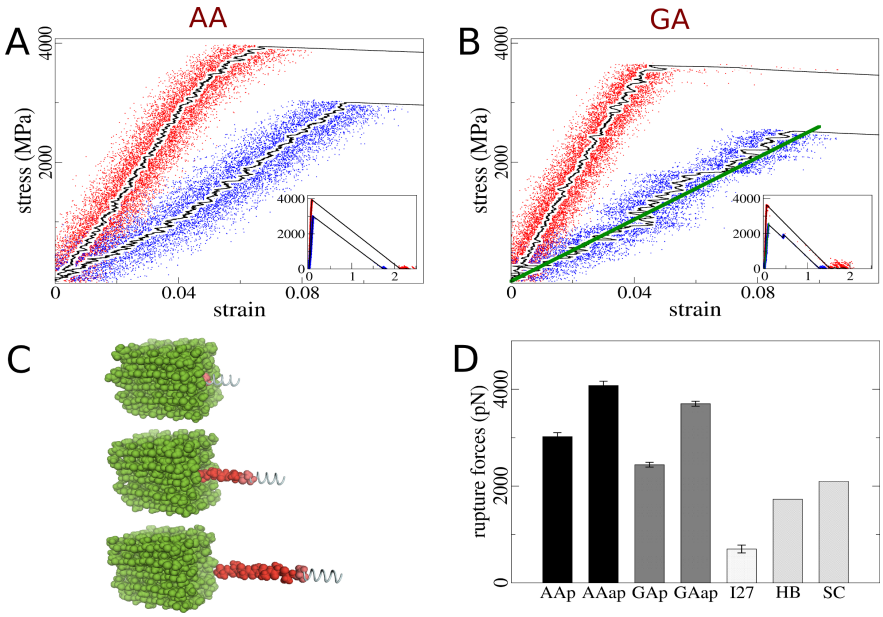

Figure 3.3: Elasticity and stability of crystalline units from force-probe MD simulations. (A) Stress-strain curves of AA_p (blue) and AA_{ap} (red), with solid black curves for averages. The inset shows the complete extension and rupture process. (B) Stress-strain curves of the GA_p (blue) and GA_{ap} (red). (C) Simulation snapshots of the rupture process. Crystalline unit is in green with the pulled strand in red. The pulling force is depicted as a spring. (D) Rupture forces of different units, of titin I27 domain as comparison, and of AA_{ap} after hydrogen bond or side-chain interaction of the central strand are switched off.

parallel structure (blue) (backbone pull-out resistance given in Table 3.1). Assuming a mixture of 2:1 (GA_{ap}:GA_p) (85), our simulations predict a modulus of GA crystals in the range of 27–87 GPa, comparing to 26.5 GPa as the experimental value (86). With regard to the putative role of crystalline units as the major force-bearing units in silk fibers, the total net area of crystalline units onto which force is primarily applied is effectively smaller than the fiber cross-section used to calculate stress, suggesting the experimental value to serve as a lower bound (personal communication, M. Müller). We note that an elastic modulus as given in the previous experimental work is not straightforwardly defined for crystalline units due to the translation of tensile to primarily shear stress within the nano-scale structure (see Methods). Nevertheless, in analogy to macroscopic fiber stretch experiments and to previous experiments on the shear-deformation within a silk crystal (86), we here define the mechanical response in terms of stress and strain as well, resulting in a modulus to quantify the backbone pull-out resistance of the

crystal. We conclude that the calculated backbone pull-out resistance are in good agreement with the modulus recently obtained from X-ray diffraction measurements, further validating our models for the crystalline silk units. Again, a higher stiffness coincides with a higher rupture force for GA_{ap} (3628 pN) comparing to GA_{p} (2435 pN, Fig. 3.3D). The correlation of the stiffness with fracture resistance, as found for both silk and spider crystalline units, was expected, since load-bearing interactions are of the same nature and range of attraction. In all five independent simulated rupture events of GA_{p}, the first rupture leads to the formation of intermediate states with newly formed hydrogen bonds, reflected by jumps in the stress-strain curve (Fig. 3.3B, inset). GA_{p} as the softest out of the four models considered here therefore shows a non-linear elastic behaviour with an effectively even lower modulus by this sliding-snapping mechanism.

As shown in Fig. 3.3D, the rupture forces of the AA models are higher than the GA models. Comparing to a change of ∼1000 pN when converting a parallel to an anti-parallel conformation, replacing glycine by alanine only increases the rupture force by ∼400 pN. The additional methyl sidechain in alanine residues thus adds mechanical resistance, but only marginally. Interestingly, all of the crystalline units (forces between 2 and 4 nN) have several times higher rupture forces than titin I27 (700 pN at a very similar pulling velocity of 0.4 nm/ns), one of the most stable protein domains known to date (99; 100). Thus, the periodically arranged β-sheets in silk fibers have an outstanding toughness, higher than any globular protein examined to date.

Force distribution from molecular dynamics

The elastic response and rupture forces we observe suggest silk crystalline units to largely outperform other previously investigated globular proteins of high mechanical toughness such as immunoglobulin-like domains (101). This high toughness is further fine-tuned by differences in the arrangement of strands into parallel or anti-parallel sheets. What are the determinants of the robustness of the β-sheet stacks in silk? To reveal the force-bearing motifs in crystalline units, and to thereby rationalize the high stiffness and differences in stiffness due to strand orientation, we performed a force distribution analysis for AA_{ap} and AA_{p}, as shown in Fig. 3.4A and B, respectively. In this analysis, atomic pairwise forces were obtained from the strained structure, held at constant force of 1660 pN in force clamp (FCMD) simulations. These forces, F_{ij}^{FC}, were compared to forces in the relaxed state, F_{ij}^{EQ}, obtained from equilibrium simulation, i. e. in the absence of force (see Methods). Force averages over time converged well within

the total simulation time of 20 ns, with an average statistical error of ∼3.6 pN. For both models, strain is maximal (red) at the point of force application at the central strand, and decays horizontally along the β-sheet involving hydrogen bonding, and vertically along the layers of alanine sidechain packing. Thus, both inter-strand hydrogen bonding and inter-sheet sidechain interaction are similarly involved in force propagation. Force distribution therefore predicts that eliminating either the hydrogen bonds or sidechain interactions will lead to a decrease in stability. We tested this by determining the rupture force after selectively switching off (i) the electrostatic backbone interactions of the central strand, or (ii) the Lennard Jones interactions of the alanine sidechain methyl group of the central strand. Indeed the rupture force is lowered from 2988 pN to 2096 pN (Fig. 3.3D, SC) (i) and 1726 pN (Fig. 3.3D, HB) (ii), respectively, confirming the notion from force distribution analysis that both types of non-covalent interactions contribute to silk crystal toughness to similar extent.

In AA_{ap} (Fig. 3.4B), most of the applied stress is already taken up by the strands adjacent to the central strand, and the stress shows a fast non-linear decay with distance. Outer strands are merely strained (blue). The subset of a few central strands suffices to sustain the external load, apparently due to strong non-bonded interactions, rendering the crystal stiff and robust (compare Fig. 3.3A and D). In sharp contrast, in AA_p (Fig. 3.4A), force is more widely distributed, along hydrogen bonding layers as well as inter β-sheet layers. Along the central strand, strain decays in a linear way and thus significantly more slowly than in AA_{ap}. Each individual sidechain or hydrogen bond interaction can take up less of the external strain, resulting in a softer structure comparing to the anti-parallel model. Consequently, both the stiffness and rupture force of parallel arrangements are generally lower. The same tendency is found for the force distribution in the respective parallel and anti-parallel GA models of *bombyx mori*.

We find the difference in hydrogen bond geometry between AA_p and AA_{ap} to be the major determinant for the difference in force distribution within one β-sheet (Fig. 3.4 C and D). By nature, parallel β-sheets feature an inclined zigzag geometry of hydrogen bonds, whereas anti-parallel counterparts show an in-line geometry. The hydrogen bonds in AA_{ap}, being oriented in-line, are responding to the external load in a homogeneous way. Fig. 3.4F shows the force differences between strained and relaxed state, ΔF, for the two sets of hydrogen bonds formed by the central strand with two adjacent

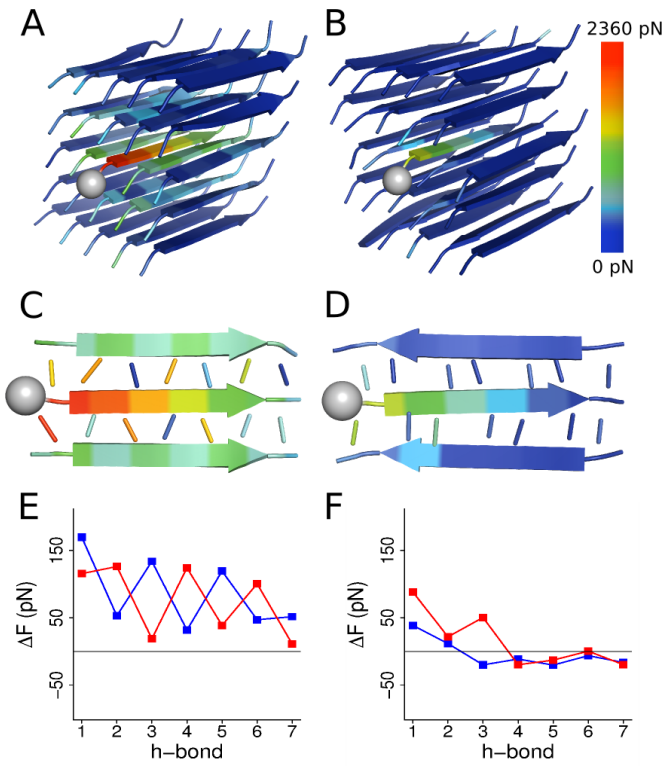

Figure 3.4: Force distribution analysis of crystalline units from MD simulations. Coloring indicates internal strain, averaged over residues, from low (blue) to high (red) levels of ΔF. Protein are shown as cartoon, force application sites as spheres. (A-B) Force distribution in AA_p (A) and AA_{ap} (B). (C-D) Force distribution within inter-strand hydrogen bonding between the central and the two adjacent strands for AA_p (C) and AA_{ap} (D), hydrogen bonds between central and adjacent strands are shown as sticks. Coloring indicates ΔF in hydrogen bonds (sticks) and residues (cartoon), using the same color code as (A-B). (E-F) ΔF for inter-strand hydrogen bonds along the strands. The upper and lower hydrogen bonds in (C-D) are shown in red and blue, respectively, starting from the point of force application in AA_p (E) and AA_{ap} (F).

strands. All hydrogen bonds generally become strained by the externally applied force. ΔF varies from ∼50 pN to ∼10 pN and decays along the strand to zero (Fig. 3.4F and color code in Fig. 3.4D).

The zigzag geometry of hydrogen bonds in parallel β-sheets instead entails an analogous zigzag pattern in the force distribution (Fig. 3.4C and E). Hydrogen bonds oriented along the pulling direction are significantly strained, with ΔF up to ∼160 pN, while oppositely oriented hydrogen bonds merely respond to the external pulling force.

The overall larger ΔF and the slower decay along the strand indicate that hydrogen bonds in the parallel β-sheet geometry are less capable of taking up load, rendering the structure softer.

However, the impact of hydrogen bond geometry onto the overall elastic response and rupture forces cannot be directly inferred from force distribution of the hydrogen bonds when being part of the whole crystal unit, since other structural components may have similar influence. To dissect the hydrogen bonding properties from other potential determinants of mechanical stability, we next examined the impact of bonding geometry only, using simplified β-sheet skeletons.

Force distribution of β-sheet skeletons

To examine how hydrogen bonding geometry affects the mechanical properties of the silk crystalline units, we built β-sheet skeletons, simplified models of one β-sheet layer, which are shown in Fig. 3.5A and B for anti-parallel and parallel β-sheets models, respectively. Structural and elastic parameters were adopted from the MD simulations, to mimic hydrogen bonding geometry and strength and backbone elastic properties of a β-sheet layer in silk. The backbone of each strand, hydrogen bonds, and the connecting C=O and N-H groups are each treated as one element. The parallel β-sheet skeleton,$\beta_{\mathrm{p,strong}}$, features a zigzag hydrogen bond geometry (Fig. 3.5A), while hydrogen bonds are oriented in-line in the anti-parallel β-sheet skeleton, β_{ap} (Fig. 3.5B). The stretching modulus for hydrogen bonds was defined from the second derivative of the all-atom hydrogen bond potential at the potential minimum, and thus was the same for both geometries. The squares in Fig. 3.5A and B represent connections between elements. Details are given in the Methods section.

Distortion of the structure upon application of a pulling force of 1660 pN (arrow) was determined by finite element analysis and is shown as solid black line (Fig. 3.5A and B). The resulting effective (macroscopic) backbone pull-out resistance are 24.3 GPa ($\beta_{\mathrm{p,strong}}$) and 26.3 GPa ($\beta_{\mathrm{ap}}$). The zigzag geometry is generally the more stable structure from a mechanical point of view, as reflected by the wide spread use of inclined cross beams as e.g. in trusses exploiting the structural stability of triangular shapes. Here, the hydrogen bonds are taking up the role of such cross beams, but due to their bending stiffness, are not able to add considerable mechanical stability. Consequently, we find nearly the same stiffness, and thus structural stability, for the two β-sheet skeletons.

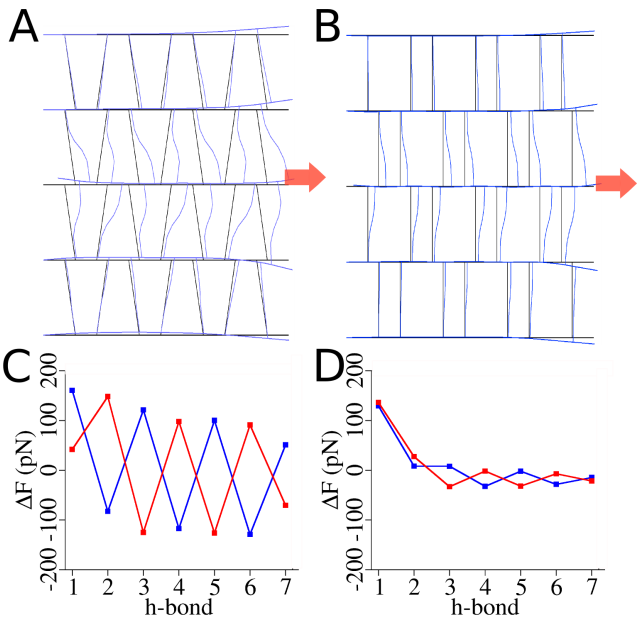

Figure 3.5: Force distribution in simple β-sheet skeletons. (A-B) Skeletons for parallel (A) and anti-parallel (B) β-sheets, consisting of connected elements for the backbone, hydrogen bonds and C=O/N-H groups (black). Dislocation at an external force of 1660 pN is obtained from finite element analysis and shown in blue. (C-D) ΔF for interstrand hydrogen bonds along the strand, as in Fig. 3.4E-F, for parallel (C) and anti-parallel (D) β-skeletons.

The β_{p} model would more clearly outperform the β_{ap} model, if hydrogen bonds would act as weaker cross beams in terms of bending, as depicted in Fig. 3.6.

However, in contrast to this prediction, the all-atom simulations indicate $\mathrm{AA}_{\mathrm{ap}}$ and $\mathrm{GA}_{\mathrm{ap}}$ to be stiffer and more robust than their parallel counterparts. A possible explanation might lie in the fact that the hydrogen bond strength is found to be the second major difference between the two alternative β-sheet arrangements. Namely, parallel β-sheets showed extended hydrogen bonds stretched out of the potential energy minimum, apparently due to steric restraints, which resulted in an effectively lower stretching modulus of the hydrogen bond element of 71 GPa comparing to 159 GPa (Fig. 3.2). By taking this difference of hydrogen bond strength between parallel and anti-parallel β-sheet skeletons into account, we found the anti-parallel β-sheet skeleton (26.3 GPa, as above) to clearly outperform the parallel, comparably weakly hydrogen bonded, analogue with a backbone pull-out resistance reduced to 17.1 GPa. Not surprisingly, the backbone pull-out resistance of the full β-sheet stacks at atomic detail

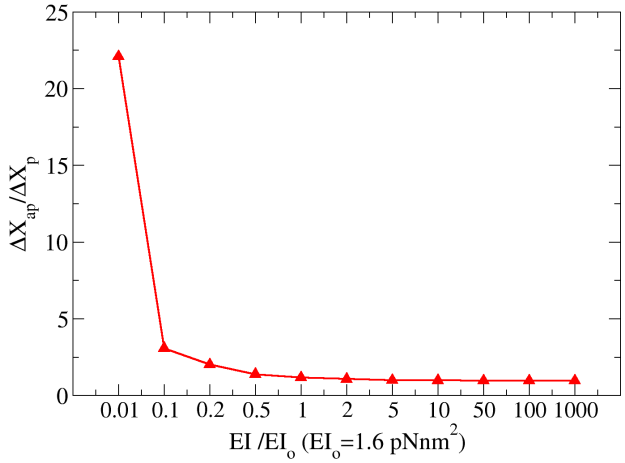

Figure 3.6: Dislocation ratio of parallel (ΔX_{ap}) to anti-parallel (ΔX_p) β-skeletons on different hydrogen bond bending moduli, EI. The X-axis has a unit of $EI_o = 1.6\,\mathrm{pN\,nm^2}$. The lower the bending stiffness, the more pronounced is the stabilizing effect of a zigzag geometry (parallel) relative to the in-line geometry (anti-parallel). Caculations have been done for a given hydrogen bond stretching modulus ($EA = 1797\,\mathrm{pN}$). For $EI \geq EI_o$, the geometry og hydrogen bonds (anti-parallel vs. parallel) has a negligible effect on the overall β-skeleton stiffness.

are overall higher (68 GPa and 21 GPa for AA_{ap} and AA_p) than those of the one-layer skeletons, since the former are additionally stabilized by inter-layer sidechain packing.

The force distribution within the β-sheet skeletons can be directly compared to the force distribution from molecular dynamics simulations. We find good quantatitive agreement of the changes in hydrogen bond forces ΔF between the simplified skeletons (Fig. 3.5C and D) and the all-atom crystalline units (Fig. 3.4E and F). ΔF decays continuously and quickly along the strand in the anti-parallel β-sheet skeleton (Fig. 3.5C) from 170 pN to zero. Thus, as in the all-atom force distribution, the pulling force propagates along the pulled strand and is taken up quickly. In the parallel skeleton, again, the zigzag pattern is clearly recovered in the ΔF pattern along the strand (Fig. 3.5D). Bonds are alternatingly compressed and stretched, depending on their relative orientation towards the pulling force, as reflected by the change in sign of ΔF. In contrast, even bonds oriented against the pulling force were stretched and weakened in the all-atom parallel silk units. Apparently, due to sidechain imposed steric restrictions, the hydrogen bond force changes are shifted upwards comparably (Fig. 3.4F).

The good agreement of the finite element analysis of simplified silk single β-sheet

layers with the full silk unit analysis suggests that hydrogen bond force distribution is largely determined by geometry and strength of the involved bonds. Thus, the presented skeletons are good approximations for a silk crystalline β-sheet layer. We can conclude that the difference in overall elasticity and structural stability between parallel and anti-parallel crystal units can be largely explained by the differential geometry and strength of hydrogen bonds, the latter compensating for the former.

Conclusion

We here suggested three-dimensional structural models of spider and silkmoth silk crystals in two feasible conformations, namely anti-parallel and parallel arrangements. We find the crystalline units to agree well with X-ray diffraction data, suggesting our idealized highly ordered models to closely resemble the structure of crystalline regions in silk fibers.

We examined the mechanical response of the silk crystalline units by three complementary approaches. First, we performed force-probe molecular dynamics simulations to determine relative stabilities from rupture forces. We find spider silk crystals (poly(A)) to outperform silkmoth crystals (poly(GA)), and anti-parallel to outperform parallel arrangements in terms of rupture forces and stiffness. Estimated backbone pull-out resistance compare well to experimental data. Second, force distribution analysis was used to reveal how external forces propagate through the crystal. The analysis determined inter-sheet alanine sidechain packing and inter-strand hydrogen bonding as the major force-bearing elements. Third, simple β-sheet skeletons were developed on the basis of all-atom MD simulations to focus on one layer of hydrogen bonds allowing to investigate how the hydrogen bond geometry affected silk crystal stability. While the zigzag motif as found in parallel β-sheets renders a structure generally stiffer than the in-line geometry of the anti-parallel analogues, the lower hydrogen bond strength in parallel sheets compensates for this effect and renders parallel conformations overall less stiff and stable. The simple finite element analysis of β-sheet skeleton semi-quantitatively reproduced the backbone pull-out resistance obtained from all-atom MD simulations. The analysis is computationally highly efficient, and opens the door towards simulations of full silk fibers. Incorporating silk crystal skeletons into the amorphous matrix will lead to a model of disordered protein chains crosslinked in

crystal units, a model that aims at accurately describing the complex force propagation on the nano-scale. Studies in this direction are under way.

The idealized silk crystalline units studied here could withstand surprisingly high forces up to 4nN. Interestingly, covalent bonds have been shown to rupture in this high force regime (12; 102; 103). For example, siloxane bonds were shown to rupture at 4.4nN in ab initio molecular dynamics simulations (102). Similarly, for a peptide bond, the bond under tensile stress in the silk crystal, we find a rupture force of roughly 4nN on a picosecond time scale (unpublished results). However, in contrast to the elongated chain molecules investigated in these studies, silk crystals very effectively distribute the tensile stress throughout the crystal. In fact, forces as high as the externally applied force are only found at the point of force application and then decrease rapidly, according to our force distribution analysis. We also note that the rupture forces up to 4nN found in our simulations are a result of the high loading rates, which are typically around six orders of magnitude higher than experimental rates. Therefore, rupture forces for silk crystals can be expected to be significantly lower than those required to induce covalent bond rupture in the backbone.

We note that our understanding of force propagation in silk crystal units obtained from MD and finite element analysis is based on a molecular mechanics model. Force fields are known to not accurately predict the angle dependency of hydrogen bonds (104), which questions our detailed force distribution analysis on inter-strand hydrogen bonds (Fig. 3.4). Quantum mechanical calculations at sufficient level of theory might allow to assess the error involved in the classical mechanical approximation.

Our study presents a first step towards a physical structure-based model of silk fiber mechanics. The ultimate aim of the ongoing effort in the field of silk mechanics is to design new silk-inspired high-performance materials. As a first conclusion, we find spider silk crystals to be slightly stiffer and structurally more stable than silkmoth silk due to the additional Ala sidechain. How other alterations of the primary sequence in the amorphous and crystalline regions will affect overall silk mechanical response remains to be analyzed.

3.1.2 Force Distribution determines Optimal Length of β-sheet Crystals for Mechanical Robustness

Atomistic models of β-sheet crystals with varying number of alanines were mechanically tested in Molecular Dynamics simulations. The gain in mechanical robustness per residue is maximal for strands that are eight residues long, because the external force is efficiently deflected into neighbouring strands on the distance they span, rendering eight residues a mechanical optimum for a semi-crystalline material like silk. (105)

Introduction

Among the secondary structure types a protein can adopt, β-sheets are generally those with the highest resistance to mechanical force. Consequently, evolution has selected β-sheets as the major building blocks of the most resilient proteins, such as titin immunoglobulin and fibronectin domains, (106; 107) amyloid fibers, (108) or the crystalline phase in natural silk fibers. (109; 110) As the hallmark of β-sheets and similarly found in the crystalline phases of some polymers like polyamides or polyurethanes, hydrogen bonds are periodically formed between adjacent β-strands in the sheet, the concurrent rupture of which can resist high forces.

The mechanical stability of a molecular architecture like a β-sheet depends on a number of important factors, such as the chemistry and relative arrangement of the chains. Given a certain composition and geometry of a β-sheet crystal with respect to the applied force, is there an optimal length of the strands in the crystal so that the material maximizes its macroscopic mechanical toughness? Optimizing the resistance against rupture gained by each unit of the protein or polymer would allow the efficient usage of the material, a competitive advantage in a world of limited resources.

Calculation and Results

We first determined the preferred length of strands in naturally occurring β-sheets, i.e. the strand length dominantly selected by evolution. We measured the distribution of β-strand length in the SCOP(Structural Classification of Proteins) (111) database by determining the secondary structure of the SCOP domains using STRIDE(STRuctural IDEntification). (112) STRIDE is known to give a more accurate assignment for 70% of cases as compared to DSSP. (113) The length distribution for all SCOP domains

features a maximum at 4-5 residues (Fig. 3.7), which reflects a positive selection for foldable and functional proteins in general and only partly an optimization for mechanical robustness. Interestingly, when restricting the analysis to a protein fold known for its mechanical function in the biological context and for its extra-ordinary stability against mechanical force, namely the immunoglobulin-like β-sandwich fold, the overall distribution shifts to longer strands, with a maximum at 6 residues per β-strand. A similar distribution with a peak for 6 residues is found when only considering domains of titin, tenascin, and fibronectin from this family, for which clear evidence for their force-carrying function is at hand (data not shown).

A prime example for a natural β-sheet structure optimized for the ability of bearing high mechanical load are the β-sheets formed by silk proteins. Spider silk shows strand lengths of 8-10 residues, (77) and thus another shift in the distribution towards longer β-strands as compared to other protein domains (Fig. 3.7, green bar). Silk is a material with outstanding toughness not matched by any other synthetic material. The crystalline units are formed from stacked β-sheets, each of which comprises several strands of mostly eight alanine residues (or alternating alanine-glycine, depending on individual type of silk). These crystals act as force-bearing units, cross-linking the amorphous phase they are embedded in (Fig. 3.8A). Since the primary evolutionary constraint for silk crystalline regions is their mechanical stability, we conclude that β-strand lengths of 8-10 residues are an optimal choice from a material design point of view. This length is roughly two times the length generally dominating proteins under multiple evolutionary constraints.

In contrast, in a recent computational study, 2D models of β-sheets consisting of three β-strands were mechanically analyzed and found to reach a mechanical optimum in terms of the pull-out resistance against force for a strand length of 4 residues. (88) However, the validity of this number for the naturally occurring 3D structures, which are stabilized by both intra-sheet hydrogen bonding and inter-sheet sidechain interactions, remains to be investigated.

To reconcile these observations, we here ask how the rupture force of a β-sheet structure varies with strand length to determine the mechanically optimal number of residues in a β-strand. We have recently built and mechanically characterized atomistic models of spider dragline silk protein crystals. (72) Crystals differing in their protein

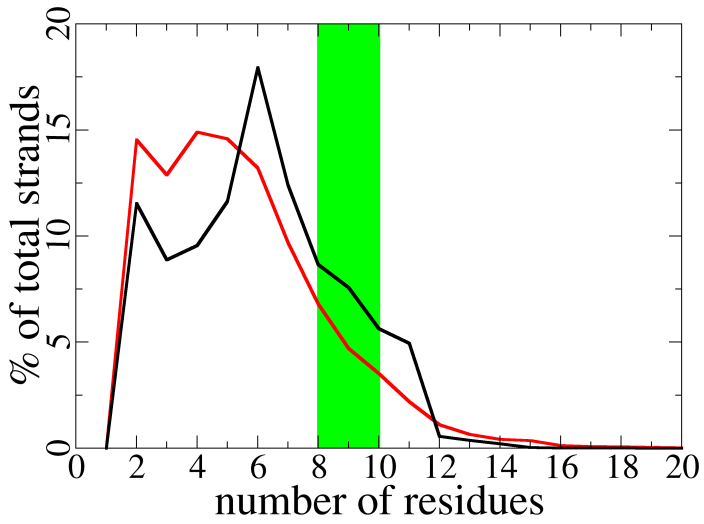

Figure 3.7: β-strand lengths distribution in proteins. The distribution for all protein structures in SCOP shows a peak at 4-5 residues (red), and for proteins of the SCOP immonoglobulin-like fold a peak at 6 residues (black). β-strand lengths found in silk protein sequences, shown as green bar, vary between 8 and 10 residues.

sequence and strand arrangement (parallel versus anti-parallel) were compared and validated by experiments. (86) Their extra-ordinary rupture forces and stiffness explained their function of reinforcing silk fibers. One thing that specifically interests us is that, how β-strand length affects pulled out rupture force? Is there a correlation between β-strand lengths and rupture force, that governs nature selection of protein structures? Here, we built silk β-sheet protein crystals which differed in their number of alanines in each strand and therefore the number of inter-strand hydrogen bonds from 4 to 16 (Fig. 3.8B). Of which are the most nature occuring β-strand lengths. All of them consisted of five strands per β-sheet layer and five of these layers. We chose an anti-parallel arrangement of strands in a β-sheet. (72) Crystals of parallel sheets are softer, but can be expected to show a similar scaling. We fully equilibrated the five all-atom structures in explicit water for 10 ns, before subjecting them to force-probe molecular dynamics (FPMD) simulations, (66) in which mechanical force was applied to the C-terminus of the central strand as indicated in Fig. 3.8B.

We used the simulation package GROMACS 4.0.5 (114) to perform all simulations.

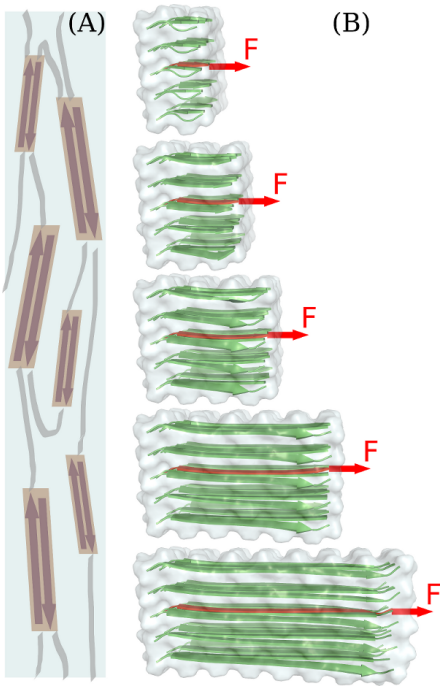

Figure 3.8: Molecular structure of silk. (A) Schematic drawing of a silk fiber comprising two different phases, namely the β-sheet crystals simplified as brown anti-parallel arrows, cross-linking the amorphous peptides shown in grey. (B) Molecular models of silk protein crystals, varying in their strand length from 4 (top), 6, 8 (corresponding to natural dragline silk), and 12 to 16 residues (bottom). Central strands are shown in red with arrows indicating the external pulling force applied in FPMD simulations.

In FPMD simulations, we moved a virtual harmonic spring with a force constant of 500kJ/mol/nm^2 away from the pulled group, the terminal C-α atom of the central strand, with a pulling rate of 0.2 nm/ns. More details have been published previously. (72) The resulting force profiles are shown in Fig. 3.9A. A sudden drop of the force reflects failure of the crystalline structures. In all cases, prior to failure, we observed the same linear increase in force, indicating a high constant elastic modulus steadily increasing with the length of the β-strands. However, rupture forces strongly varied with strand length, with larger rupture forces for longer strands, as expected (Fig. 3.9A).

The interesting quantity for optimizing a material is the rupture force gained per residue in the crystal. Thus, we calculated the force per residue, i.e. per hydrogen bond and sidechain interaction, by normalizing the rupture force by the number of alanine

groups in the strand (shown as black squares in Fig. 3.9B). We found a maximal force load of 427 pN/residue for β-sheets of 8 residues. The rupture force per residue in 4-residue-long crystals is 169 pN/residue, and thus less than half of the mechanical optimum. For longer β-sheets comprising 12 or 16 residues, the load carried per residue decreases again. Apparently, adding additional residues to an 8-residue long crystal does not lead to a significant gain in mechanical resistance. At even longer scales, we expect the total rupture force to reach a plateau.

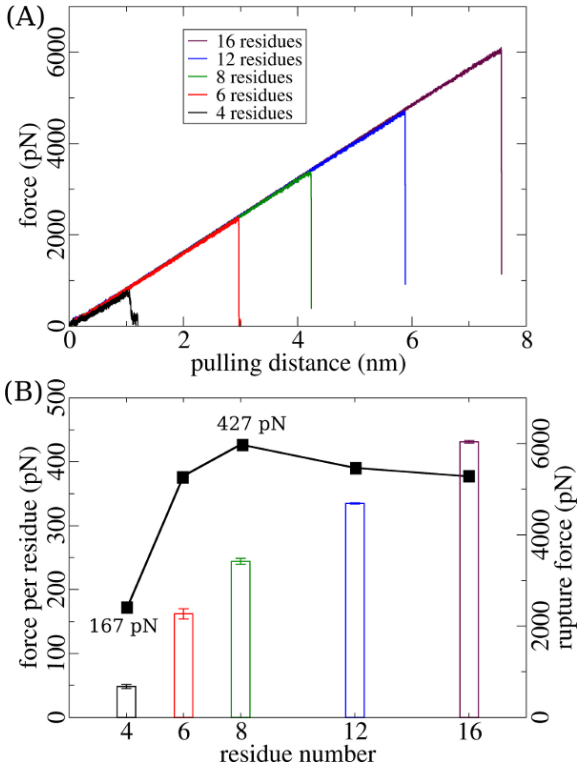

Figure 3.9: Force for rupturing β-sheet crystals of varying strand length. (A) Force profiles for models with strand lengths varying from 4 to 16 residues as indicated. We observe a linear increase in force prior to failure in all cases. (B) Rupture forces, i.e. maximal forces observed in FPMD simulations as shown in (A), again for models with varying strand lengths (color bar). The longer the crystal, the higher the rupture force required to pull out the central strand. By dividing the rupture force by the number of residues in the strand, we obtained a normalized rupture force per residue (black curve), which shows a maximum of 427 pN for strands of 8 residues.

What is a microscopic interpretation of the maximum force per residue for $(Ala)_8$ β-sheet crystals? Apparently, the external force, propagating along the pulled strand, is

deflected vertically into adjacent strands via hydrogen bonds and sidechain interactions such that the residues at a distance of 8 residues and more from the pulled terminus do not take part in carrying any significant strain. To test this interpretation, we performed force distribution analysis (FDA) to detect the internal propagation of stress in the structure prior to rupture. (54)

The main idea of the newly developed FDA is to analyze the inter-atomic forces as the structure is subjected to an external constant force in force-clamp molecular dynamics (FCMD) simulations. The externally applied force is low enough to monitor the force distribution in a mechanically equilibrated state prior to rupture. The atomic pair-wise forces are averaged over time and substracted from those forces observed in the relaxed state of the structure, i.e. in the absence of an external force. FDA is reminiscent of finite element analysis performed for revealing the stress distribution in macroscopic objects. The investigated molecular structure was equilibrated on two states, namely a perturbed force loaded state and a non-force relax state. All possible interacting atomic pairwise forces were collected and averaged for a long simulation time. Corresponding pairwise forces are then subtracted each other to get the absolute force differences which were then remapped onto the molecular structure. For more technical details, please consult to the original FDA publication. (54) Here, we subjected the $(Ala)_{16}$ crystal to MD simulations in the absence of an external force, and in presence of a constant force of 1660 pN. This external force is low enough to keep the structure intact within the nanosecond time scale of the simulations. We then calculated the internal force distribution in the $(Ala)_{16}$ crystal from the difference in the atomic pair-wise force between the resulting strained and relaxed state. The atomic forces have been averaged over 10 ns simulations for each state. The details of this method are published elsewhere. (54)

The forces carried by individual hydrogen bonds formed between the pulled central strand and its two adjacent strands in the central β-sheet layer are shown in Fig. 3.10A. The hydrogen bond forces, also shown as color codes on the β-sheet structure, are maximal at the point of force application, and decay along the strands. They reach almost zero at the eighth residue in the central strand. Residues more than 8 residues away from the point of force application are virtually not involved in the force distribution. Similarly, the next rows of hydrogen bonds in the structure do not carry any significant load (Fig. 3.10B). On the basis of this central result of the FDA, we predict the addition

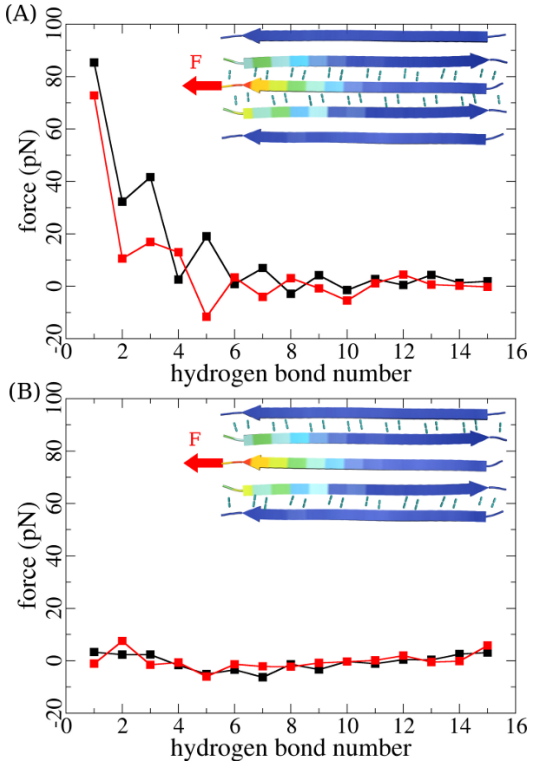

Figure 3.10: Force distribution in the hydrogen bonds of a β-sheet layer. (A) Forces taken up by the hydrogen bonds between the central pulled strand and the two adjacent strands, as indicated as dashed lines in the structure. The upper and lower strand hydrogen bonds are shown in red and black, respectively. The same data is shown as a color code mapped onto the β-sheet structure, where red indicates high and blue indicates low strain. (B) Force distribution of hydrogen bonds between the outer strands of the β-sheet layer. Otherwise same as (A).

of alanine residues to strands of eight residues or longer not to add to the β-sheet crystal's mechanical stability. This is in agreement with and gives a molecular explanation for our results from force-probe MD simulations above. We conclude that β-strands of eight residues in length are the most efficient building block of mechanically resistant structures.

We note that we obtained the consistent result of 8 residues as the optimal β-sheet length from two complementary, yet very distinct simulations techniques, namely FPMD and FDA. While in FPMD simulations, a constant loading rate leading to forces as high as 6 nN is employed, FDA makes use of a constant and significantly lower force

(1.66 nN). This underlines that our finding is independent from the details of force application, but instead is largely determined by the nature of the molecular structure.

On the basis of the above results, we can explain the length of β-strands found in spider dragline silk. The segment of the sequence forming β-sheet crystals is poly$(\text{Ala})_8$. Apparently, during evolution, spiders optimized the sequence such that the resource of the natural material silk is most efficiently used, resulting in the maximal resilience of the crystalline subunits at a given sequence length. Similarly, intra and extra-cellular proteins with mechanical function often feature β-sheet sandwich domains with β-strands in the range of 8 residues. Among others, the immunoglobulin domains of the muscle protein titin are among the mechanically most resistant domains known and consists of strands typically showing a length of 6 ± 1 residues as shown in Fig. 3.7. We note that the optimum of eight residues is an approximate result. Differences in the rupture forces conferred per residue for 6 and 12 residues, though significant, are not large and might depend on the particular sequence and packing geometry. However, in contrast to earlier results, (88) we clearly show that β-sheets formed from four residues or even less perform poorly in comparison and as a consequence have not been the length of choice during evolution.

Conclusion

In conclusion, we compared the mechanical efficiency of β-sheet crystals of different strand lengths in terms of rupture force per residue and found 8-residue long β-strands to be the optimum. Further analysis by force distribution confirmed this view. Our results explain the prevalence of β-strands 8 residues in length in nature and also have important implications for designing artificial semi-crystalline materials. For synthetic silk-mimetic polymers, we predict eight residues, i.e. hydrogen bonds, in the crystalline phase an optimal choice for the cross-linking and force-bearing properties of the crystals.

3.1.3 Silk Fiber Mechanics from Multiscale Force Distribution Analysis

Here we decipher the molecular determinants for the extreme toughness of spider silk fibers. Our bottom-up computational approach incorporates molecular dynamics and finite element simulations. Therefore, the approach allows the analysis of the internal strain distribution and load-carrying motifs in silk fibers on scales of both molecular and continuum mechanics. We thereby dissect the contributions from the nano-scale building blocks, the soft amorphous and the strong crystalline subunits, to silk fiber mechanics. We identify the amorphous subunits not only to give rise to high elasticity, but to also ensure efficient stress homogenization through the friction between entangled chains, which also allows the crystals to withstand stresses as high as 2 GPa in the context of the amorphous matrix. We show that the maximal toughness of silk is achieved at 10-40% crystallinity depending on the distribution of crystals in the fiber. We also determined a serial arrangement of the crystalline and amorphous subunits in lamellae to outperform a random or a parallel arrangement, putting forward a new structural model for silk and other semi-crystalline materials. The multi-scale approach, not requiring any empirical parameters, is applicable to other partially ordered polymeric systems. Hence, it is an efficient tool for the design of artificial silk fibers. (47)

This Study was performed in collaboration with my colleague Dr. Murat Cetinkaya, who performed the finite element analysis.

Introduction

Silk fibers constitute an intriguing class of natural materials. Through a flawless assembly of strong and soft building blocks, they exhibit astonishing mechanical properties. Silk fibers may have an ultimate strength comparable to steel, toughness greater than that of kevlar, and a density lower than that of cotton or nylon. (77) Furthermore, many natural silk fibers exhibit high rupture strain ($\sim$30%), which is one of the major reasons for their energy-absorbent behavior upon impact. (77; 115) Even today, natural silk fibers outperform their artificial counterparts in terms of mechanical performance. Therefore, many experimental (77; 86; 87; 116–119) and theoretical studies (72; 86; 87; 116; 120; 121) have tried to understand the process of silk fiber formation and the origins of the mechanical characteristics of silk fibers.

Natural silk fibers share a common structural architecture consisting of two types of major components, namely the crystalline and amorphous subunits (Fig. 3.11). (79; 109; 110) Crystalline subunits of spider silk involve short peptides of six to ten amino acids containing alanine or glycine-alanine residues. These short peptides organize themselves into mechanically strong crystal blocks measuring 2∼5 nm on a side. (83) They are made of parallel or anti-parallel layers of β-strands (85) interacting via hydrogen bonds and via non-covalent bonds between amino acid side chains. (72) Crystalline subunits constitute 10-25% of the fiber volume in spider silk. (77; 84) They are axially oriented along the fiber and reinforce the soft amorphous subunits by acting as comparably stiff cross-linking sites. (117) The amorphous subunit of a silk fiber is composed of longer, glycine-rich peptide sequences. (77; 116; 117) Even though the amorphous subunit may include some semi-ordered peptide chains with partial secondary structure (84; 120; 122; 123), it is predominantly disordered (118; 124; 125), rendering itself as a soft matrix in stretching experiments. (77; 86; 87; 117)

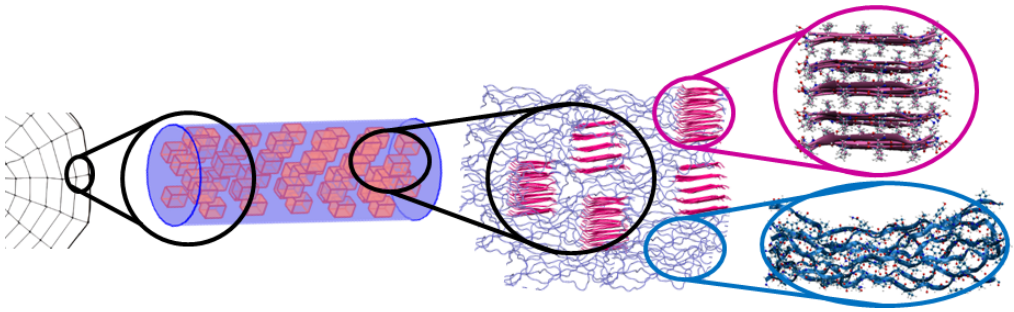

Figure 3.11: Schematic of the spider silk architecture from macro- to nanoscales. Spider silk fibers are composed of crystalline subunits that are made of β-sheets (pink), and semi-extended disordered peptide chains (blue).

Assessing the determinants of the mechanics of natural silk fibers is challenging not only due to the possible variety in the peptide chemistry, but also due to the influence of the assembly conditions, such as the shear rate and ion concentration, on the fiber mechanics. (119; 126) The macroscopic mechanical characteristics of silk fibers originate from the nano-scale morphology and the interactions within the fiber subunits. However, technical difficulties in experimental and computational studies (e.g. isolation and structure determination of the subunits in experiments, simulations of larger systems for longer time scales) prevent a comprehensive analysis of spider silk

fibers and provide only partial information. In particular, computational studies on the atomistic scale have so far considered only the crystalline subunits and, therefore, were limited in the understanding of silk fiber mechanics. (72; 120; 127) An approximate yet elegant nanoscale mean approach by Vollrath, Porter, and co-worker (87; 128) successfully predicted some features of macroscopic fiber mechanics and its dependence of order.

In this study, our aim was to understand the mechanical properties of spider silk fibers using a bottomup computational approach that incorporates both the crystalline and amorphous subunits. Our approach bridges two discrete methods and scales: atomistic molecular dynamics (MD) simulations for individual and coupled subunits, and finite element (FE) simulations for a comprehensive fiber model on the continuum scale. The contributions from the subunits to the mechanical properties of spider silk were investigated. Different structural architectures regarding fiber subunits were tested for optimal mechanical performance. The calculated values for the elastic moduli, rupture stress and strain, and toughness were in good agreement with experimental values. Our results make testable suggestions for a synthetic silk-mimicking polymer. Since our approach is applicable to similar polymeric systems, it is a useful tool for the design of artificial silk fibers.

Materials and Methods

All-atom Molecular Dynamics (MD) simulations

All MD simulations were performed with Gromacs 4.0.5 (53). For force distribution analyses, an extension of Gromacs (54) was utilized to write out atomic pair-wise forces. The simulation parameters were the same as in the previous study (72). Shortly, the OPLS-AA force field (40) for the protein and the TIP4P model (60) for water were employed, a time step of 2 fs was used, and simulations were performed at constant temperature (300 K) and pressure (1 bar), with periodic boundary conditions and with Particle-Mesh Ewald summations for long-range (>1 nm) electrostatics (91). All simulations were run without any bias regarding force field parameters or boundary conditions.

Crystalline subunit

The all-atom model is composed of five layers, each layer containing five β-strands with each β-strand having eight alanine repeats. Details on the modeling, equilibration,

and force distribution analysis of the crystalline subunits have been published (72). The crystalline subunit was totally intact throughout the equilibration and up to the rupture point in the force-probe simulations. Fig. 3.12 shows the RMSD of the β-sheets in a crystalline subunit during an equilibrium MD simulation. The overall RMSD is as low as ~0.11 nm after 10 ns.

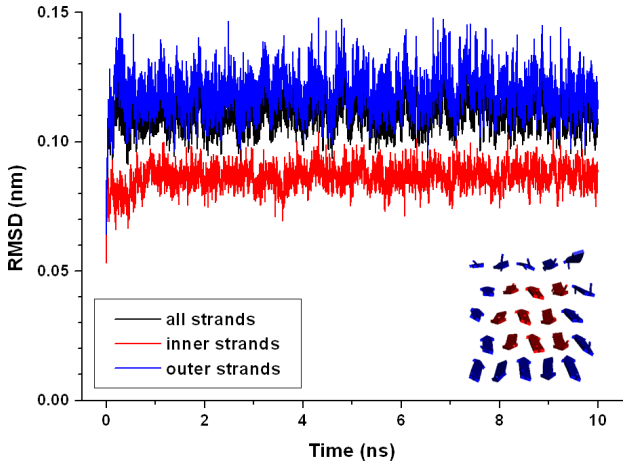

Figure 3.12: RMSD of a crystalline subunit during an equilibrium MD simulation of 10 ns. The overall RMSD of all 25 β-sheets in the crystalline subunit (black) is shown along with the RMSD regarding the inner 9 (red) and outer 16 β-sheets (blue).

Amorphous subunit

The all-atom model comprises the 24 residue sequence known to form the amorphous subunit in Araneus diadematus spider silk (GPGGYGPGSQGPSG PGGYGPGGPG, where G, P, Y, S, Q are glycine, proline, tyrosine, serine, and glutamine, respectively). A single peptide chain was generated in a partially extended conformation (end-to-end distance of 6.2 nm), solvated in water with physiological ion concentration (100 mM), resulting in a system of $6\times6\times10$ nm^3 and ~47000 atoms. During 20 ns of equilibrium MD simulations, the chain collapsed to an average distance of 0.38 nm. To obtain a force-extension curve with minimal nonequilibrium effects, we used umbrella sampling, in which harmonic potentials acted on the peptide termini along z-direction. We sampled in 34 windows for 20 ns each, starting from snapshots with varying the end-to-end distances (0.12 to 8.87 nm). The harmonic potential force constant was 100 kJ/mol/nm^2. All other simulation parameters were kept the same as in the all-atom simulations of the crystalline subunit. Weighted histogram analysis method (129)

was applied to check the overlap between the umbrella samples and to calculate the stretching force of the peptide as a function of its end-to-end distance.

Composite unit

Two crystalline subunits were serially coupled with an amorphous subunit and pulled away from each other. The amorphous subunit was composed of eight entangling peptide chains having 50% initial extension relative to the chain contour length. This initial extension was chosen as an intermediate value of a polymer in shear flow (see e.g. Rammensee, et al. (119)), as it is present in the spider gland during silk fiber spinning. In natural spider silk fibers, cross linking of the crystalline units via disordered chains occurs randomly to an extent that depends on the alignment of the chains during fibrillogenesis due to shear flow. Our model corresponds to a situation where approximately one third of the disordered peptide chains (8 out of 25) connected the same crystalline subunit to another one. Conformations for the disordered peptide chains at an extension of 50% of the contour length were taken from umbrella sampling MD simulations (see above). The strands from the two crystals were connected by the eight peptide chains as shown in Fig. 3.17b in the main text. For the skeleton model, we find the connectivity not to influence the effective elastic response of the composite unit. Probing a composite unit with different connectivity (e.g. with all straight chains) is computationally demanding, and can be expected to also give a similar elastic response, given the efficient cross-chain force distribution (see Results in the main text).

The composite unit was solvated with TIP4P water in a $7\times7\times45$ nm^3 box with 100 mM ion concentration. The system had a total number of 300,000 atoms. After energy minimization using steepest descent method, 500 ps position restrained simulations were followed by 10 ns equilibrium simulations. As a further validation of the all-atom model, the density of the composite unit was calculated as 1.12 g/cm^3, in good agreement with $\sim$1.14 g/cm^3, calculated by Fossey (120) and in reasonable agreement with the experimental results for similar silk fibers with $\sim$1.35 g/cm^3 (130) and for the amorphous regions with $\sim$1.14 g/cm^3 (131). To obtain a force-elongation curve, crystalline subunits were pulled using virtual springs with harmonic force constants (500 kJ/mol/nm^2) acting at the center of mass of the alanines at the C and N-termini of the composite unit. The springs were moved away from each other with a velocity of 0.2 m/s. The force-probe MD simulations were stopped after rupture occurred at $\sim$60 ns. For the force distribution analysis, constant forces of 1600 kJ/mol/nm and

100 kJ/mol/nm were applied to the same pulling groups to sample the perturbed and relaxed states, respectively. The system was equilibrated under this constant force and the pairwise forces in the composite unit were averaged and written out over 10 ps intervals. For both states, five independent MD simulations (10 ns each) were performed and the pairwise forces averaged over the aggregated simulation time. All other parameters were identical to those described above.

Finite Element (FE) simulations of the skeleton and comprehensive fiber models

Crystalline subunit

We constructed a skeleton model of the crystalline subunit (5×5 β-strands) consisting of four types of isotropic Euler beams. Each protein backbone was represented by a single elastic member. We defined a linear elastic modulus, E, of ∼743.0 GPa as obtained from stretching a single $(Ala)_8$ chain in force-probe MD simulations. For other members, representing single covalent bonds (i.e. C=O, N-H, and C_α-C_β), the elastic modulus was directly obtained from the force constant k of these bonds in the OPLS/AA force field (40), using $E = kL_0A_o^{-1}$, resulting in ∼297.0 GPa. The nominal cross sectional area Ao was calculated by assuming a bond radius of 0.1 nm for all covalently bonded members. Similar to the covalent bonds, the nonbonded interactions between alanine side-chain (CH_3) groups in the all-atom model were also represented with linear elastic members in the skeleton model. These members connected backbone members between β-sheet layers. The elastic response was estimated from a Lennard-Jones (LJ) potential fitted to the total interaction energy between alanine CH_3 groups in equilibrium MD simulations. Nominal length of the corresponding members in the skeleton model, L_0, was the equilibrium CH_3-CH_3 distance observed in MD simulations.

Hydrogen bonds in the crystalline subunit, connecting the backbone members within a β-sheet layer, were represented with nonlinear elastic members in order to consider the weakening of hydrogen bonds with increasing strain. The strain-dependent elastic modulus of the hydrogen bonds was calculated from the first derivative of the hydrogen bond potential, defined as the sum of the Lennard-Jones and Coulomb interactions between the four atoms taking part in hydrogen bonding (i.e., N-H and C=O) as defined by the OPLS-AA force field (40). The cross sectional radius of the members representing the nonbonded interactions (intrasheet hydrogen bonds and intersheet side-chain interactions) was taken as 0.06 nm (details provided in Xiao, et al. (72)). Nominal

length, L_0, of the members representing the hydrogen bonds was defined as the equilibrium hydrogen bond (O$\cdots$H) distance observed in MD simulations.

Amorphous subunit

The amorphous subunit was modeled with nonlinear elastic members, each of which corresponds to a 24-mer peptide chain as in the all-atom model. The force elongation curve of the amorphous chain from all-atom MD simulations (see above) was used as an input for the skeleton model. The elastic modulus increases with chain length L, as a result of the entropic stiffening characteristic for polymer chains (Fig. 3.13a). This mechanical behavior was observed at different portions of umbrella sampling MD simulations, indicating that the simulation time scale was not a major limitation (Fig. 3.13b). At low forces the convergence between force-extension curves is reasonable, and at higher forces the curves overlap. The cross sectional radius of the corresponding members in the skeleton model was 0.1 nm as for other covalent members, and nominal member length, L_0, was taken as 0.38 nm, which was the average end-to-end distance of the coiled peptide chain during equilibration MD simulations.

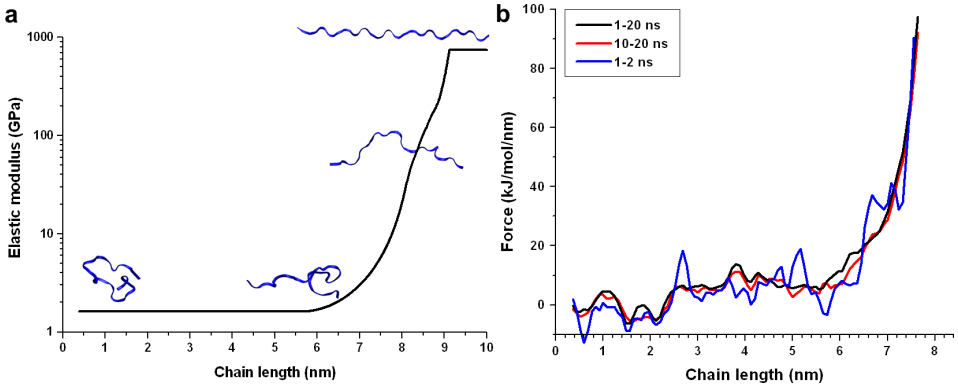

Figure 3.13: Mechanical behavior of the disordered peptides in the amorphous subunits. (a) Elastic modulus of the members representing the disordered chains of the amorphous subunit in the skeleton model. The nonlinear behavior of these chains resembles a neo-Hookean solid. Up to a certain chain length, a constant value of 1.6 GPa was used for convenience. The upper limit was set as 743.0 GPa, which is the elastic modulus value for the fully extended backbone members. Schematics represent the state of chain extension. (b) Force- extension curves of a disordered peptide chain obtained from different portions of umbrella sampling MD simulations. At low forces the convergence between force-extension curves is reasonable, and at high forces the curves overlap indicating full convergence.

Composite unit

Analogous to the all-atom composite unit, we serially coupled two crystalline subunits with an amorphous subunit consisting eight diagonally oriented chains. External tensile stress was applied in a distributed manner at the termini of each crystal in opposite directions. Parameters of the members in the composite unit were same as in the skeleton models of the individual subunits.

Comprehensive fiber model

This model was only tested with FE simulations, since MD simulations of the same system at atomistic scale (∼1.2 million atoms) were computationally unfeasible. A three dimensional solid stress-strain model was used for the fiber FE simulations. Amorphous subunits were assumed to be completely isotropic (E_a = 2.7 GPa) while the crystalline subunits were modeled as transversely isotropic (E_c = 80.0 GPa in the pulling direction). The elastic moduli of the subunits were calculated from the all-atom simulations of the composite unit. Poissons ratio for both types of subunits was taken as ∼0.33 in the pulling direction, as suggested by Fossey for poly(Gly-Ala) crystals (120). These values were used as input for the comprehensive fiber model. The fiber was represented by a linear elastic cylindrical solid member pulled by distributed loads acting on both ends.

The fiber was 7 nm in radius and 39 nm in length. The crystalline subunits were represented with linear elastic cubic members 12 nm^3 in size, corresponding to the dimensions of the all-atom crystalline subunit, and embedded into a continuous amorphous phase without any slip at the interfaces. We compared a random distribution state of the crystals in the fiber with two other extreme states, named as serial (lamellar) and parallel (longitudinal) distribution of the crystals. We note that the parallel distribution, even though insightful for comparison, cannot be realized due to the block-copolymer like silk protein sequence.

The elastic modulus of the fiber, E_f, was determined with varying crystallinity, which is the volume percentage of the crystals in the fiber (Fig. 3.18a). Rupture strength and strain of the fiber were estimated from the stress distribution analysis across the fiber (Fig. 3.18b). More precisely, the fibers rupture stress and strain were the values at the instant where the average tensile stress in the crystalline subunits reached 2.0 GPa. This was the value observed in the all-atom model of the composite unit prior to rupture. Toughness, the integral of the stress-strain curve, of the fiber was calculated as half the product of the rupture stress and the strain based on our

definition. We here do not take yielding of the fiber into account, which allows for a further increase in (plastic) strain without the failure of the fiber. We thus consider the estimated rupture strain and toughness values to present the lower limits of the actual rupture strain and toughness, respectively.

Results and Discussion

Mechanical characteristics of individual subunits

Silk mechanics on macroscale originates from the molecular characteristics of crystalline and amorphous subunits on nano-scale. Therefore, the first step of our approach was to model the individual subunits with simplified skeleton models and validating them by comparing to the all-atom models. We developed a threedimensional skeleton model for the crystalline subunits formed by five layers of fivestranded anti-parallel β-sheets, in which the side-chain interactions link the twodimensional skeleton layers, which were previously developed. (72) The stress-strain curve and the resulting pull-out resistance of the skeleton model obtained from FE simulations (54.1 GPa, Fig. 3.14 and Methods) compared well to that obtained from the more accurate all-atom model in MD simulations (66.3 GPa).

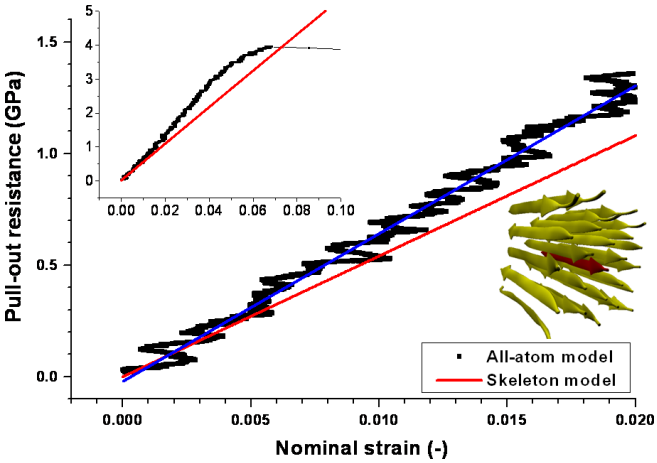

Figure 3.14: Pull-out resistance of the crystalline subunit. Results from the all-atom (black) and skeleton models (red) as the middle strand is pulled out from the crystalline subunit (indicated by the red β-strand in the lower inset). A linear fit for the all-atom model is shown with blue line. The upper inset shows the pull-out resistance-strain curves for a higher range of values.

In addition, the skeleton model was able to reproduce the force distribution pattern as observed in the all-atom model, namely a steady decay of the load along the pulled strand and in lateral and vertical directions perpendicular to the plane of pulling. We concluded that the skeleton model for FE simulations, being parameterized on the basis of the all-atom description, captures the features of the crystalline subunit mechanics.

We straightforwardly modeled the second constituent of the spider silk fibers, the disordered peptide chains of the amorphous subunit, by inputting the force-extension data obtained from long-time-scale all-atom MD simulations. The non-linear elastic response, approximately following a worm-like chain behavior up to intermediate extensions, was the basis for the non-linear elastic modulus of the amorphous subunits (Fig. 3.13).

Mechanical characteristics of the composite unit

We next serially coupled the crystalline and amorphous subunits to dissect their relative contributions to the strength and toughness of spider silk fibers. The connectivity between the disordered chains in the amorphous subunit and the β-sheet crystal in a silk fiber is unknown, but is likely to show a high variation within the fiber. We here chose a symmetric connectivity, details of which are given in the Methods. The composite unit comprised 300,000 atoms (including solvent) in the all-atom model, which served as a reference, and 3,000 members in the simplified skeleton model. Fig. 3.15a shows the force-elongation curves for the skeleton and all-models of the composite unit. In both models, the strength of the composite unit increased with higher elongation. The skeleton structure showed a force-elongation relationship that one would expect if the subunits, most importantly the individual disordered protein chains, behaved additively.

More specifically, the non-linear elasticity of the chains in the amorphous subunit dominated the overall mechanical response, and the crystalline subunits were strained only at high stresses. This mechanical behavior was also independent of the way the peptide chains connected the crystalline subunits, as entangled and straight peptide chains gave indistinguishable force-elongation curves (data not shown). The force-elongation behavior of the all-atom model shares the main features with the skeleton model. However, the all-atom model yielded a significantly lower elasticity for the amorphous chains, resulting in a smaller strain at rupture. Apparently, the more realistic all-atom description of the amorphous subunit took internal friction into account

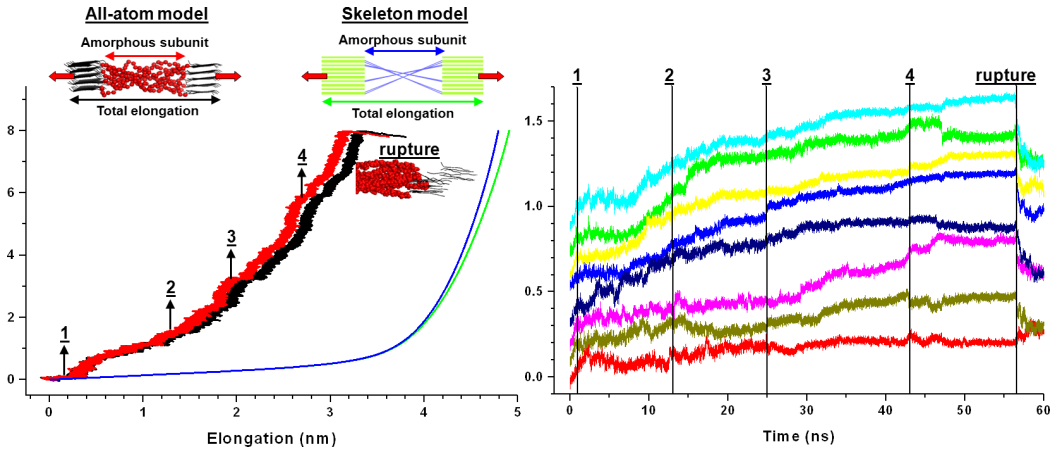

Figure 3.15: Mechanical response of the composite unit. (a) Force-elongation curves from the skeleton (green, whole unit; blue, amorphous subunit) and all-atom models (black, whole unit; red, amorphous subunit) (Insets at the top). Schematics for the composite unit from all-atom (left) and skeleton models (right) (Inset next to the rupture point). Representative structure of the all-atom model right after the rupture event, at which peptide chains partially disengage from the crystal. Numbers 1∼4 in panels a and b indicate the events at which peptide chains rearrange themselves. (b) Root-mean-square displacement (RMSD) of the peptide chains in the all-atom model during force-probe MD simulations. Individual traces are vertically shifted for clarity.

that led to higher energy absorption than what was expected from the mere sum of the individual chains.

Indeed, there were jumps in the force-elongation curve of the all-atom model (Fig. 3.15a) correlating with structural rearrangements of individual disordered chains relative to each other (Fig. 3.15b). Such short-range diffusional motion is characteristic for rubbery-elastic mechanics of amorphous polymers. (132) We interpret the difference between the force-elongation curves of the all-atom and skeleton models as the additional mechanical work taken up by the entangled chains within the amorphous subunit. The skeleton model failed to take this effect into account since inter-chain interactions were neglected. We suggest that this non-additive behavior of the chains in the amorphous subunit crucially enhances the toughness of spider silk fibers.

The all-atom model of the composite unit ruptured at a stress of 2.0 GPa with 33% strain (black curve in Fig. 3.15a). Gosline et al. reported an experimental rupture stress of 1.1 GPa with 27% strain for the same spider silk type. (77) Experiments by

Liu et al. resulted in a rupture stress of 1.7 GPa with 24% strain. (133) Considering the simplicity and the limited size of the all-atom system, our values are in good agreement with the experimental results in literature. In fact, given the high crystallinity of our composite model (65% as opposed to 25% for natural spider silk), the overestimation of the rupture stress in the all-atom model was expected.

Varying the loading rate by one order of magnitude resulted in a change in rupture stress of smaller than 10% and in toughness of only ∼3%, due to the high extension of amorphous chains at the rupture point. Complementary MD simulations with a bare amorphous subunit showed that in the absence of crystalline subunits the peptide chains would slide from each other and rupture at a low stress of 0.4 GPa (Fig. 3.16). Therefore, the crystalline subunits are the determinant factors in the rupture of the composite unit, in agreement with the previous notion of crystalline subunits acting as cross-links in silk fibers. (84; 117; 121) For further validation we also estimated the toughness of the all-atom model and obtained 147 MJ/m3. Previous experimental studies reported values between 141-225 MJ/m3. (77; 133–135) The agreement is striking regarding the tremendous difference between the length scales.

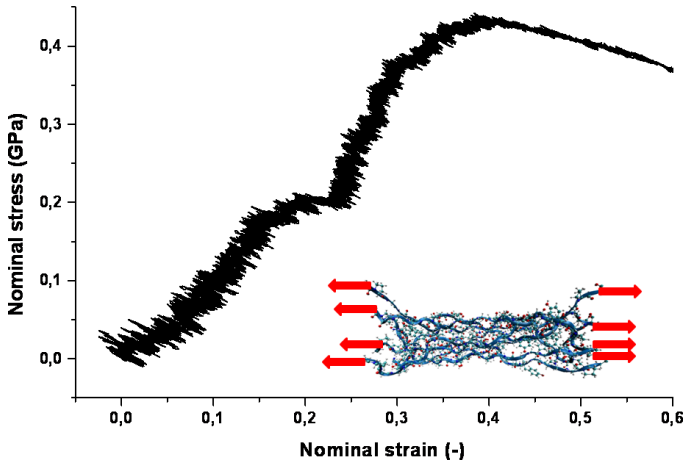

Figure 3.16: Stress-strain curve of the amorphous subunits in the absence of the crystalline subunits in an all-atom model. Inset shows the schematic of the pulling process.

The distribution of the pulling force within the composite unit gives additional insight into the molecular contributions to the toughness of silk fibers. As a general finding, the axial force was high in disordered peptide chains but decayed within the crystalline subunits due to an efficient distribution of the pulling force along and across

β-strands (Fig. 3.17a and b). This was in particular the case in the skeleton model, where the peptide chains in the amorphous subunit acted as individual strings and thus carried axial loads of very high and strongly varying magnitude (up to 400 pN) depending on their linkage to the crystalline subunit (Fig. 3.17a). As a consequence, the distribution of the pulling force within the crystalline subunit was heterogeneous (Fig. 3.17c). In contrast, the amorphous chains in the all-atom model carried forces only slightly higher than the ones in the crystalline subunit.

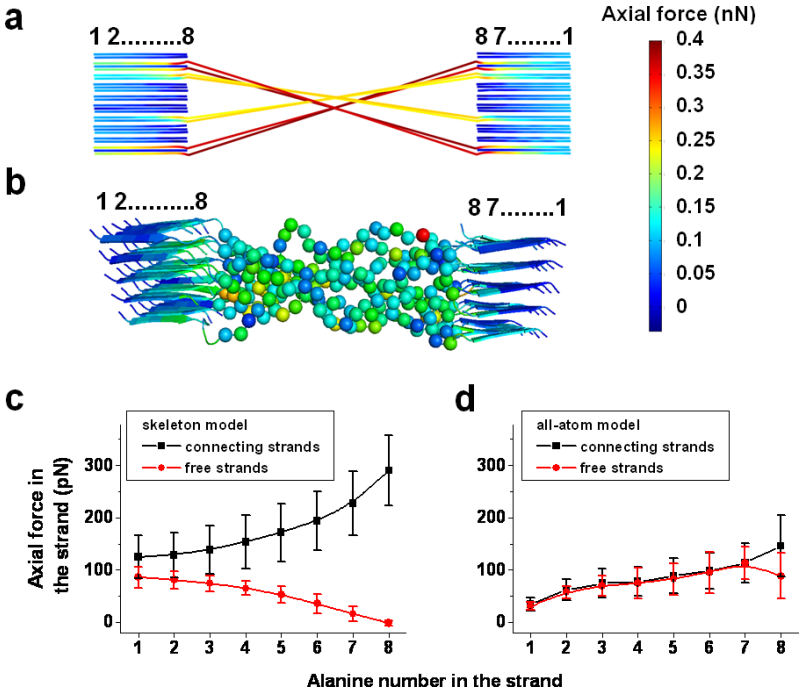

Figure 3.17: Distribution of the external pulling force (∼2.5 nN) within the composite unit. (a and b) The force distribution pattern in the skeleton (a) and all-atom models (b) of the composite unit. (c and d) Average axial force along the poly(Ala) strands in the crystalline subunits in the skeleton (c) and all-atom (d) models. The order of alanines are indicated in panels a and b. The term connecting strands' in panels c and d refers to the poly(Ala) strands bonded to the disordered chains in the amorphous subunits. Error bars represent the standard deviation for 34 free strands and 16 connecting strands in a 55 crystalline subunit. Smooth lines in panels c and d are spline fits for the data points.

More importantly, these forces (∼200 pN) were independent of the linkage and entanglement of the chains (Fig. 3.17b). The distribution of the pulling force in the crystalline subunit was also more homogeneous (Fig. 3.17d). Overall, the skeleton model

overestimated the absolute and relative axial loads in the composite unit due to the heterogeneous distribution of the pulling force across amorphous subunits. The force distribution analysis corroborates our earlier finding that the inter-chain interactions enable efficient stress propagation and homogenization, as an important contribution to toughness. Overall, the pronounced differences between the all-atom MD simulations and the FE calculations of the skeleton model helped us to dissect the contributions of the structure to mechanical resilience. Improving the skeleton model further by incorporating interchain interactions might be the subject of further investigations, but was not of interest here, as the further upscaling was based on only the all-atom simulations.

Elastic moduli of the crystalline (E_c) and amorphous (E_a) subunits in the all-atom model were calculated as 80.0 GPa and 2.7 GPa, respectively. The elastic modulus of the poly(Ala) spider silk crystals is currently unknown, but Krasnov et al. reported the E_c for silkworm silk poly(AlaGly) crystals as 26.5 GPa. (86) In these experiments, the stress in the crystalline subunits in a full fiber was inferred by assuming a homogeneous stress distribution. However, we below show that in the range of experimental crystallinity values (10∼25%), crystalline subunits carry stresses 2∼3 times larger than the macroscopic fiber stress (see the comprehensive fiber model).

Based on this information, the corrected experimental E_c is ∼50.0∼80.0 GPa, thus close to our calculated E_c value. Regarding the amorphous subunits, Krasnov et al. reported an E_a value of 6.3 GPa as an indirect estimation based on their experiments with silkworm silk. (86) In Termonia's computational study (121), the chains in the amorphous subunits were assumed to be entropic springs, with E_a calculated as 70.0 MPa. Apparently, similar to our skeleton model of the individual peptide chains, Termonia's study underestimated the stiffness of the amorphous subunit by two orders of magnitude because it overlooked the effects of inter-chain force distribution, suggests to be a hallmark of rubbery behavior.

Mechanical characteristics of the comprehensive fiber model

The skeleton and all-atom models of the composite unit, albeit representing minimal models of a spider silk fiber, proved useful in assessing the fiber mechanics. The important question arises how the relative amount and arrangement of the two subunits influence the macroscopic mechanical properties. We built a comprehensive fiber model with crystals embedded into an amorphous matrix and parameterized their elastic properties from the all-atom simulations. The comprehensive fiber model enables

us to reach higher length scales with a smaller computational cost (1∼5 CPU hours as compared to ∼1 Mio CPU hours for a several million atom system and microsecond time scales).

Fig. 3.18a shows the dependence of the fiber elastic modulus, E_f, on the fiber's crystallinity. We considered three possible distributions of the crystalline subunits in the fiber:

1. a serial (lamellar-like) arrangement of the crystalline and amorphous subunits.
2. a parallel (longitudinal) arrangement.
3. a random arrangement between these two extremes.

As shown in Fig. 3.18a, for any given crystallinity, a serial assembly resulted in lower fiber elastic modulus E_f dominated by the mechanical characteristics of the amorphous subunits. On the other hand, a parallel assembly, being reinforced by the crystalline subunits throughout the fiber axis, resulted in the highest values of E_f. Unsurprisingly, these two extremes follow the behavior of simple two-phase isotropic models (*blue* and *red curves* in Fig. 3.18a), validating our FE simulations. A random distribution of the crystalline subunits in the fiber yielded intermediate E_f values, as expected. At natural (i.e. 25%) crystallinity, a fiber with randomly arranged crystals, as suggested by Gosline et al. (77), has an E_f value of 6.4 GPa. This value is in excellent agreement with the reported experimental data ranging between 4.0∼10.0 GPa (77; 133–135) despite the approximations in our model such as a two-phase system with linear elastic components.

While the serial arrangement at 25% crystallinity (E_f = 4.2 GPa) is another feasible model for natural spider silk, the parallel arrangement (E_f = 22.1 GPa) can be excluded on the basis of our calculations. We note that our comprehensive fiber model neglects the details of connectivity between the two phases and of the entanglement within the amorphous matrix. Interestingly, given that the experimental results can be reproduced, the amorphous and crystalline phases apparently act largely as a continuum, presumably due to the stress homogenization described above, and render this approximation feasible.

Apart from high elasticity, spider silk was presumably also optimized for high rupture strength. The comprehensive fiber model, being based on conventional FE calculations, is not capable of showing fiber rupture. However, the rupture can be indirectly

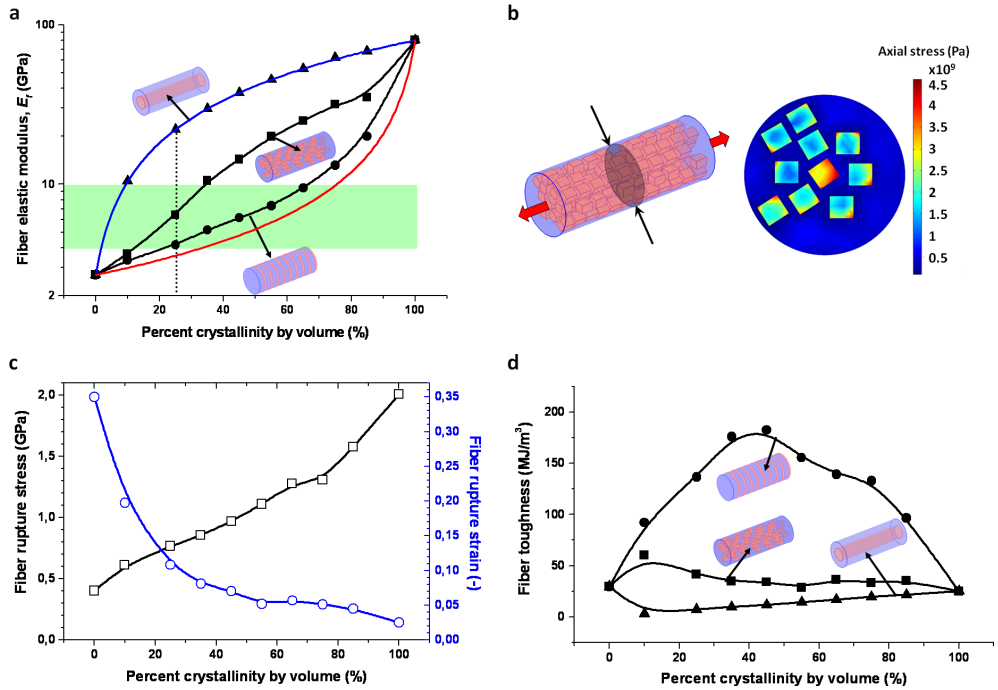

Figure 3.18: Variation of the mechanical properties with respect to the silk fiber crystallinity. (a) The effects on the fiber elastic modulus, E_f, are shown. Results for three different crystalline subunit distributions are shown as indicated by the schematics. The parallel (triangles) and serial (circles) distributions obey simple two-phase isotropic models, namely, the Voigt composite model (blue line) for the parallel distribution: $E_f = E_c u_c + E_a u_a$ and the Reuss composite model (red line) for the serial distribution: $1/E_f = u_c/E_c + u_a/E_a$, where E_c, u_c, and E_a, u_a are the elastic moduli and volume percentage of the crystalline and amorphous subunits, respectively. The small discrepancy with the Reuss model is due to the geometric effects arising from nonzero Poissons ratios used in three-dimensional FE simulations. The random distribution (squares) yields intermediate values for E_f. (Dashed line) Percent crystallinity of dry spider silk in nature. Green band spans the interval of experimental values from literature. (77; 133–135) (b) Schematic of the comprehensive fiber model with random distribution of crystals (left) and the stress distribution at an arbitrary cross section of a 25% crystalline fiber with an external stress of 0.8 GPa (right). (c) Variation of the rupture stress (squares) and rupture strain (circles) with respect to the fiber crystallinity, again for a fiber with random distribution of crystals. Values for 0% crystallinity were calculated with the all-atom model. (d) Variation of the toughness with respect to the silk fiber crystallinity. Results for the parallel (triangles), serial (circles), and random (squares) distribution states of the crystalline subunits are shown as indicated by the schematics. Smooth lines in panels a, c, and d are spline fits for the data points.

inferred from the average tensile stress in the crystalline subunits, as they failed at a tensile stress of ∼2.0 GPa in the all-atom model (Fig. 3.15a). The stress distribution across the fiber cross section showed that the crystalline subunits carried higher stresses (by a factor of ∼3 at 25% crystallinity with random distribution) compared to the amorphous subunit (Fig. 3.18a).

In other words, the tensile stress was concentrated in the crystalline subunits. However, the distribution of tensile stress became more homogeneous within the fiber with increasing crystallinity and it approached the upper limit of 2.0 GPa, the rupture strength of the crystalline subunits (Fig. 3.18b). Our comprehensive model predicted a fiber with 25% crystallinity to rupture at 0.8 GPa, which is close in magnitude to the experimental value of 1.1 GPa reported by Gosline et al. (77). The corresponding rupture strain was calculated as 11%, a clear underestimation of the experimental value of 27%. (77)

Our model reached 27% rupture strain only with mechanically weak fibers having 5% crystallinity. A major reason for this discrepancy lies in the neglect of any plastic deformation, as constant elastic moduli (i.e. E_a and E_c) were employed for the fiber model. In contrast, in silk stretching experiments and for many semi-crystalline polymers, the linear elastic regime is followed by yielding that leads to a further increase in the strain prior to rupture. (77; 132) Despite these discrepancies, Fig. 3.18 reveals interesting insight to the structural constraints of natural spider silk fibers. As expected, a fiber purely composed of amorphous subunits would be highly elastic, but also very weak, while a 100% crystalline fiber would be brittle. Thus, the crystalline and amorphous subunits improve fiber strength and elasticity, respectively, resulting in a balanced tradeoff between the two.

The question arises if this trade-off results in a maximal toughness, as another consequence of the structural architecture in spider silk fibers. Fig. 3.18d shows that the parallel distribution of crystalline subunits resulted in the lowest toughness, while the random distribution state yielded slightly higher values. Fibers with serial distribution largely outperformed the other alternatives in terms of toughness. An optimal toughness of silk fibers is achieved in the range of 10% (random distribution) and 40% (serial distribution), which overlaps with the spider silk's natural crystallinity level of 10-25%. Naturally, the calculated toughness (41 MJ/m3 at 25% crystallinity for the random distribution) underestimates the experimental results (141-225 MJ/m3 (77; 133–135))

due to neglecting plastic deformations prior to rupture in this linear elastic model. Furthermore, a certain degree of serial arrangement in the fiber might further increase the toughness, as suggested by our model. Overall, the qualitative comparison between the possible geometries clearly reveals that the serial distribution is superior over other alternatives in terms of toughness.

Conclusions

Our bottom-up computational approach incorporated both the crystalline and amorphous subunits in the fiber and provided a deeper understanding for the mechanical nature of spider silk fibers. Mechanical properties such as strength, elasticity, rupture, and toughness were investigated and we provided their calculated values (Table 3.2). We compared the capabilities of the skeleton and all-atom models of spider silk components. The skeleton models were in qualitative agreement with the all-atom models, and the quantitative discrepancies between them allowed us to identify the important determinants of spider silk mechanics – in particular, the friction between the entangled chains causing higher stiffness and energy absorbance, and allowing an efficient homogenization of stress. We show that the crystalline and amorphous subunits are the sources of strength and elasticity, respectively. In particular, the crystalline subunits are crucial elements as they act as cross-linking sites at which the stress is higher compared to the amorphous subunits in the fiber. The rupture in spider silk fibers originates from the failure of the crystalline subunits.

We show that the crystalline and amorphous subunits are the sources of strength and elasticity, respectively. In particular, the crystalline subunits are crucial elements as they act as cross-linking sites at which the stress is higher compared to the amorphous subunits in the fiber. The rupture in spider silk fibers originates from the failure of the crystalline subunits.

We showed that Nature makes a balanced trade-off between elasticity, strength, and toughness in spider silk fibers by choosing a moderate crystallinity level of 10-25%. Different possibilities for the distribution of crystalline subunits provide advantages in specific mechanical characteristics. We find that 10% crystallinity with random distribution yields a moderate elastic modulus, rupture strength and rupture strain. Interestingly, a significantly higher toughness can be achieved in fibers with lamellar structures at 40% crystallinity. While being directly inferable from simple composite

Table 3.2: Comparison of the calculations mechanical properties with the experimental data in literature

	all-atom	FE*	Experiment
Pull-out resistance			
crystalline subunit	66.3 GPa	54.1 GPa$^{\circledast}$	N/A
Rupture stress			
crystalline subunit	2.0 GPa		N/A
Amorphous subunit	0.4 GPs		N/A
Fiber	2.0 GPa	0.8 GPa†	1.1∼1.7 GPa‡,¶
Rupture strain			
crystalline subunit	25%		N/A
Amorphous subunit	35%		N/A
Fiber	33%	11%†	24∼27%‡,¶
Elastic modulus			
crystalline subunit	80.0 GPa		26.5 GPa$^{\S,\|}$
Amorphous subunit	2.7 GPa		6.3 GPa§,**
Fiber		4.2∼6.4 GPa*,†,††	4.0∼10 GPa‡,¶,‡‡,¶¶
Toughness			
Fiber	147 MJ/m^3	41∼137 MJ/m^{3*,†,††}	141∼225 MJ/m^{3‡,¶,‡‡,¶¶}

*The term FE stands for finite element calculations regarding the comprehensive fiber model.

$^{\circledast}$obtained by skeleton model.

†The result was obtained for 25% fiber crystallinity with transversely isotropic crystalline members.

‡From Gosline et al. (77). ¶From Liu et al. (133).

§From Krasnov et al. (86). The result was reported for poly(GLY-ALA) crystals.

$^{\|}$The corrected value we propose takes the higher stress in crystals into account (see text), that should be 50∼80 GPa.

**The result was reported for silkworm silk and it is an indirect estimation from the overall fiber elastic modulus.

††The results were obtained for serial and random distribution of the crystalline subunits.

‡‡From Swanson et al. (135). ¶¶From Koehler and Vollrath. (134)

models, to our knowledge this has not yet been suggested for silk, and the current picture for silk fibers is the one with randomly distributed crystals. However, a certain tendency of silk proteins to assemble into lamellae is plausible, as:

1. We find that the mechanical properties of the lamellar structures fall into the experimental range.

2. Multimerization domains in silk might enhance the alignment of the poly(Ala) repeats to form lamellae.

3. Similar structures have been observed for synthetic polymer blends (see e.g. (136)).

In conclusion, in order to reach the toughness of natural silk fibers with synthetic analogues, our simulations emphasize the importance of well-defined segment lengths in crystalline-amorphous copolymers in order to form lamellae (137) and strongly interacting, yet highly disordered polymer types for the amorphous subunit.

Using a bottom-up computational approach, our conclusions are not based on empirical parameters. A major limitation, however, lies in the need of high loading rates in the MD simulations, which entail an overestimation of the predicted rupture forces and toughness, as discussed above. Together with the restricted system size of all-atom models,here for the composite model several 100,000 atoms, moving to the continuum scale for the full fiber, could relieve these limitations. The assumption of a simple two-phase fiber with isotropic (amorphous) and transversely isotropic (crystalline) subunits is a drastic, but apparently a reasonable approximation in the elastic deformation regime, as we could validate our model by a number of quantitative comparisons to the experiments. The incorporation of viscous and plastic behavior will help to further improve the accuracy of the model. The explicit treatment of water at the nanoscale allows us to address the pivotal role of wetting in silk mechanics (138), which will be the subject of future investigations. Our computational approach is a helpful tool in artificial fiber design, inasmuch as it does not require any empirical parameters and therefore is straightforwardly applicable to other partially or highly ordered polymeric and composite systems.

3.1.4 The Effects of Crystalline Subunit Size on Silk Fiber Mechanics

Here, we utilize a computational bottom-up approach for deciphering the size effects of crystalline subunits in silk fiber mechanics. We varied the crystal size in terms of their cross-sectional area, i.e. the number of layers of b-strands, S, in the crystal, and the backbone length along the fiber axis, N, while constraining other major parameters such as chemical composition, fiber crystallinity and their relative orientation in the fiber. The computational approach incorporates Molecular Dynamics and Finite Element simulations of the crystalline subunits along with Finite Element simulations of a two phase silk fiber in order to determine their stress-strain behavior, elastic moduli and toughness. Overall, the fiber elastic modulus and toughness increase with the length of the crystals as given by the number of residues in the beta-strands, and decreases with the crystal cross-section area, i.e. the number of strands per crystal. The minimal cross-sectional area investigated, a 3×3 crystal, shows the highest sensitivity of the mechanical properties towards the crystal length. The presented approach is a versatile tool in artificial fiber design since it does not require any empirical parameters and it is similarly applicable to other semicrystalline polymeric systems or composite materials. (139)

As in Section 3.1.3, finite element analysis have been performed by Dr. Murat Cetinkaya in this project.

Introduction

There has been a long-standing and growing interest in natural silk fibers due to their intriguing mechanical properties. Mimicking the mechanical properties of these biological materials has remained a challenge due to their structural complexity. Consequently, natural silk fibers continue to outperform their artificial counterparts. Through a hierarchical arrangement of the strong crystalline and elastic amorphous subunits at the atomistic scale, silk fibers exhibit superior mechanical properties such as high toughness, elasticity, and rupture strength. (77; 140) These properties can be further fine-tuned by the amino acid sequences of the constituent proteins (77; 133; 135; 141) and the conditions during fiber formation. (119; 134; 142; 143) For instance, the reeling speed of spider silk fibers and their interaction with water affects the size of the crystalline

subunits (142), and the pretension (i.e. initial extension) of the disordered peptide chains in the amorphous subunits (143), respectively.

Experimental studies on silk fibers have provided great detail about their chemical and physical nature such as the mechanical response at the atomistic and macroscopic levels (77; 116; 134; 141–144), the crucial features of the silk protein sequences (133; 135; 145), and the size and architecture of the building blocks. (83; 85; 123; 142; 143; 146) There has also been considerable effort regarding the artificial synthesis of spider silk fibers. (119) On the other hand, many computational studies at different length scales aimed at understanding the mechanical characteristics of the nanoscale building blocks (72; 120; 127) and the macroscopic properties of the spider webs as an assembly. (121; 144; 147) The size effects of the building blocks has been a recent point of interest (105; 148) as it is a requisite for designing artificial silk fibers with tailored mechanical properties. Due to technical limitations in the experimental studies, computational methods have been the method of choice for understanding the relationship between the subunit size and macroscopic fiber properties.

In this study, we investigate the size effects of the crystalline subunits on silk fiber mechanics by utilizing a bottom-up computational approach that bridges atomistic Molecular Dynamics (MD) and continuum scale Finite Element (FE) calculations. (149) First, the effects of the crystal length and the cross-sectional area on crystal elasticity and crystal toughness as obtained from MD simulations are presented. The optimum size for a crystalline subunit is discussed. On this basis, the relationship between the size of the crystalline subunits and the fiber mechanical properties such as elastic modulus, rupture strength and stress, and toughness are analyzed with FE models. In comparison to other computational studies in literature (120; 121; 127; 144; 147; 148), we consider the contributions both from the crystalline and amorphous subunits of the silk fibers, and combine atomistic and continuum scale methods. In our analysis, we utilize a random distribution of crystalline subunits in the fiber and include the effects of both tensile and shear forces on the crystals by a new pulling setup. The results from our comprehensive fiber model are in good agreement with previous experiments, and the model demands considerably less computational power compared to the purely atomistic models.

Methods

The spider silk fibers are built of fibrils composed of highly ordered crystalline subunits being surrounded by the amorphous subunits which are made of semi-extended, disordered peptide chains (Fig. 3.19a). As in our previous study, (149) we consider the crystalline and amorphous subunits of the spider silk fiber from the major ampullate gland of Araneus diadematus. (77) The crystalline subunit is composed of layers of *beta*-strands from poly(Ala) peptides. The size of the crystalline subunits is expressed in terms of the backbone length (i.e. number of residues), N, and the number of layers, S, that is proportional to the cross-sectional area of the crystals (Fig. 3.19b). Therefore, crystalline subunits are composed of S×S β-strands each of which are made up of N amino acids (i.e. S×S×N amino acids).

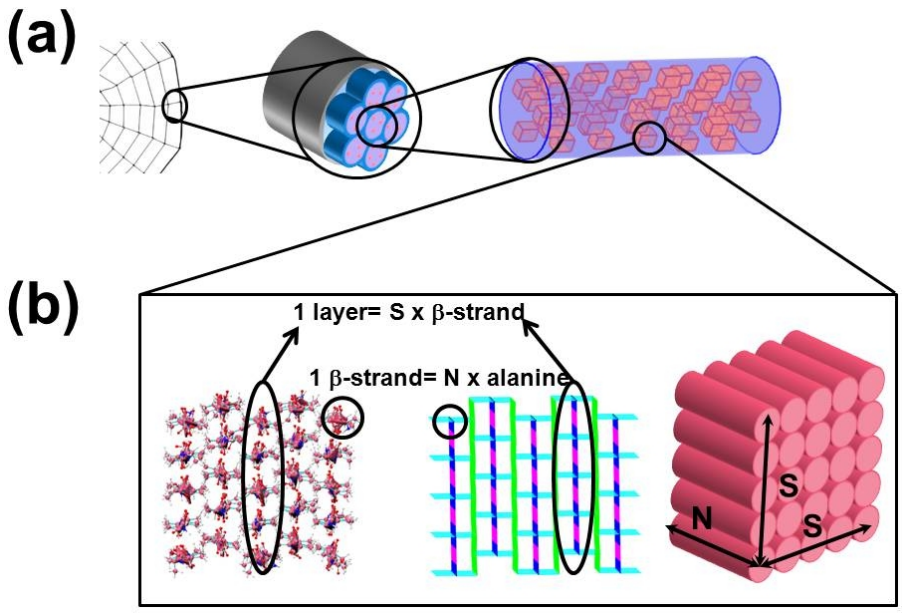

Figure 3.19: The structural hierarchy in spider silk fibers. (a) Fibers of spider dragline silk are composed of fibrils which are made up of strong crystalline subunits (pink) segregated in a soft amorphous matrix (blue). (b) The crystalline subunits (left: all-atom model, middle: skeleton model) are composed of layers containing poly(alanine) β-strands with their backbone axis oriented along the fiber axis. The size of the crystalline subunits is quantified by the number of β-strands in one layer, S, and the backbone length (i.e. the number of alanines in a β-strand), N.

Molecular Dynamics (MD) simulations

All MD simulations were performed with Gromacs 4.0.5. (53) The parameters for equilibrating the simulation systems were assigned the same as in our previous studies. (72; 149) In summary, the OPLS-AA force field (40) for the peptides and the TIP4P model (60) for water were used. A time step of 2 fs was chosen, and simulations were performed at constant temperature (300 K) and pressure (1 bar), with periodic boundary conditions and with Particle-Mesh Ewald summations for long-range (>1 nm) electrostatics. (91)

The all-atom models of the crystalline subunits were composed of anti-parallel β-strands of alanine. Tensile stretching of the crystals was accomplished by symmetrically pulling the β-strands in opposite directions in parallel to the backbone axis (see Fig. 3.20 inset). In this pulling setup, neighboring strands were pulled in opposite directions, while no force was applied to the central strand. Compared to similar studies, (127; 148) this way of pulling involves both the axial and shear effects on the peptides in the crystal, and as such is more close to the situation in a fiber, where most of the strands are subjected to forces of some magnitude depending on the connectivity between crystals. The pulling setup described here is rather an idealized case that results in a homogeneous distribution of pulling stress across the crystalline subunits, and therefore, can be expected to give an upper limit of the rupture stress. The toughness of a crystalline subunit was calculated by integrating its stress-strain curve up to the rupture point, at which the β-strands in the crystals disengage from each other.

Finite element (FE) simulations

FE simulations were performed for (i) skeleton models of the crystalline subunits and (ii) the fiber model. All FE simulations were carried out with the COMSOL Multiphysics package. Nominal values of stress and strain were used in calculations. We considered all systems as purely elastic, thereby neglecting rate-dependent viscous and dissipative plastic effects.

The three-dimensional 'skeleton model' of the crystalline subunit was described previously. (149) In short, the skeleton models are built of isotropic Euler beams representing peptide backbones and side-chain elements with linear elasticity, and hydrogen bonds with a strain dependent, non-linear elasticity. The starting geometry of the N=8, S=5 structure, the nominal lengths of the included members and elastic (Young's) moduli of these members were obtained from the equilibrium MD simulations with the corresponding all-atom model. For other crystal sizes, the skeleton models were built

proportional to the N=8, S=5 structure. The pulling setup for the skeleton models was the same as the one described for the all-atom models.

The fiber model was a three dimensional solid stress-strain model composed of transversely isotropic crystalline subunits embedded into a matrix of isotropic amorphous subunits without any slip at the interfaces. The fiber was represented by a cylindrical solid member pulled by distributed loads acting on both ends. The fiber was 7 nm in radius and 39 nm in length. The crystalline subunits were randomly distributed in the fiber, but their longitudinal axes were kept aligned with the fiber axis. (123) Both types of subunits were treated as linear elastic solids. All fiber models were 25% crystalline by volume. The elastic modulus of the amorphous subunit (2.7 GPa) was calculated from the all-atom model in our previous study. (149) For the crystalline subunits, the elastic moduli in axial and transverse directions and the shear moduli were calculated with the skeleton models of the corresponding geometries. Poisson's ratio for both types of subunits was assigned as ∼0.33 along the axial direction, as suggested by Fossey for poly(Gly-Ala) crystals. (120) Here, we did not consider the yielding of the fiber, which allows for a further increase in the strain without the failure of the fiber. Therefore, the estimated rupture and toughness values present the lower limits of the actual values, respectively. Fiber rupture was assumed to happen at the instant where the average von Mises stress in the crystalline subunits reached their rupture stress in the corresponding all-atom models, which naturally depends on the crystal size. Fiber toughness was calculated as half the product of the fiber rupture stress and strain.

Results

Mechanical properties of artificial fibers could be tailored by controlling (i) the crystallinity level, (ii) the structural architecture (e.g. lamellar or random distribution), or (iii) the size of the subunits in the fiber. Recently, we have investigated the significance of the former two factors and confirmed that natural spider silks choose a moderate crystallinity level of 10∼25% in order to achieve a balanced-trade off between strength and elasticity. (149) Furthermore, the crystalline subunits were suggested to be distributed in a random or lamellar fashion due to the agreement between our calculations and the experimental results. Considering the fiber toughness, lamellar distribution was found to be a more advantageous option for the structural architecture when designing artificial fibers. Here, we constrain the parameters regarding fiber crystallinity and

structural architecture (i.e. random distribution with 25% crystallinity), and investigate the size effects of the crystalline subunits by varying the length of the β-strands, N, and the number of layers that build up a crystalline subunit, S (Fig. 3.19b).

Mechanical properties of the crystalline subunits

When pulling alternating β-strands from one end, crystalline subunits mostly exhibited a linear stress-strain relationship (Fig. 3.20a). We have observed a similar stress-strain profile also with other pulling setups (e.g. pulling out one strand out of a crystal, or pulling all β-strands from both ends). (72; 149) Unlike the stick-slip mechanism of hydrogen bonds as suggested by Keten et al. (127) and Nova et al., (148) the crystalline subunits in our all-atom simulations did not show such a plastic deformation prior to rupture. Such a discrepancy might arise from the boundary conditions applied on the crystalline subunit. In our simulations the top and bottom layers of a crystalline subunit were not constrained. Therefore, the hydrogen bond network between the β-strands in the crystal did not lead to such a deformation regime.

Both the all-atom and skeleton models showed that the elastic modulus of the crystalline subunits decreases with increasing number of layers, S (Fig. 3.20b). Hence, a higher number of *beta*-strands results in softer crystals. For instance, the all-atom models with S=3 and S=9 (with constant N=8) have elastic moduli of 96.0 GPa and 49.8 GPa, respectively. This observation qualitatively agrees with the one reported by Nova et al. although a different pulling setup was used. (148) The backbone length of the *beta*-strands, N, on the other hand, has an opposite effect on crystal elastic modulus as shown in Fig. 3.21. Longer crystalline subunits from the skeleton and all-atom models have higher elastic modulus. The all-atom models with N=4 and N=16 (with constant S=5) had 19.8 GPa and 137.0 GPa elastic moduli, respectively. Despite their simplicity, the skeleton models quantitatively agree with the all-atom models and thus allowed us to investigate a larger set of values for N and S. Fig. 3.20c summarizes the effects of both geometrical parameters, S and N, on crystal elastic modulus. It is clear that higher N and lower S values result in higher elastic modulus for the crystalline subunits. Longer β-strands with more intra-layer hydrogen bonds and smaller layers with less axial strain are the key determinants for a stiffer crystal, with a maximum of the elastic modulus of 134.6 GPa for N=16 and S=3. However, when considering the average contribution of each amino acid to the stiffness by normalizing by the total number of alanines in the crystal, a different maximum emerges. As shown in Figure

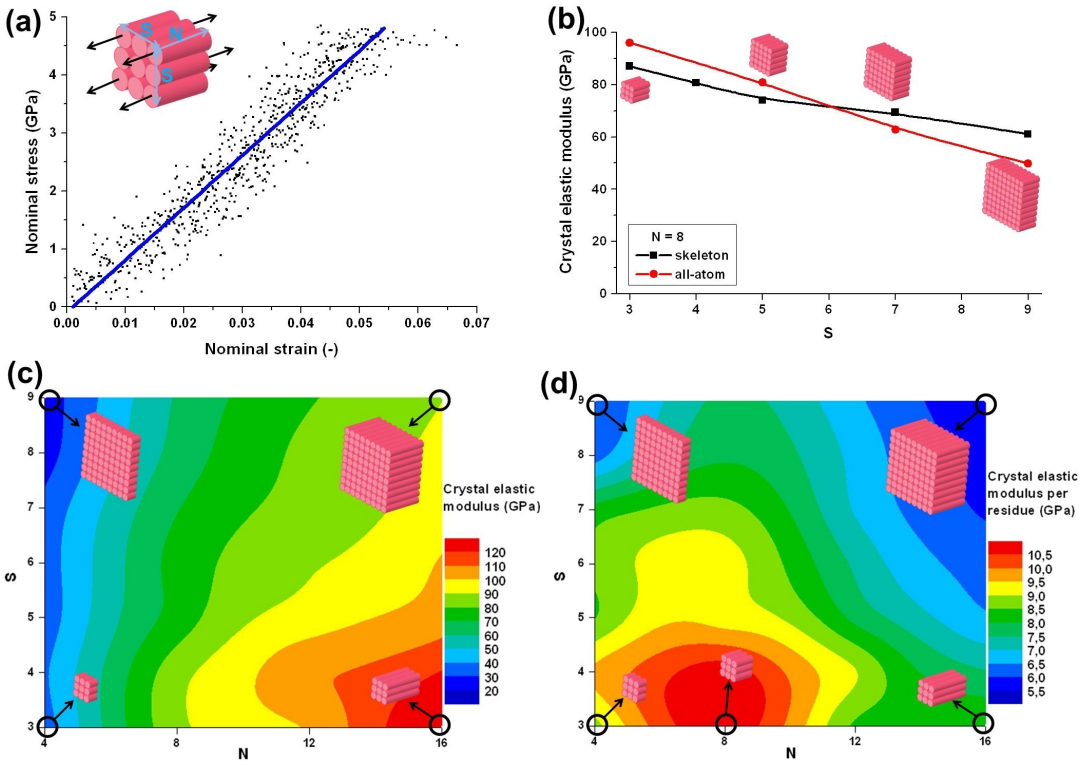

Figure 3.20: Mechanical properties of the crystalline subunits. (a) The stress-strain curve from the all-atom simulations of a crystalline subunit (N=8, S=3). Blue line shows a linear fit up to the rupture point. Inset shows the pulling setup. (b) A comparison of the all-atom and skeleton models showing the decrease in the axial elastic modulus with the number of layers, S. (c) A contour map constructed for the skeleton models showing the change in axial elastic modulus with respect to N and S. (d) The normalized version of (c) showing the per residue contribution to the elastic modulus. Smooth lines in (b) are spline fits to the data points.

2d, we found a small cross-section (S=3) and an intermediate β-strand length (N=8) to be the mechanically optimal configuration, with each of its residues providing an average stiffness of 10.9 GPa.

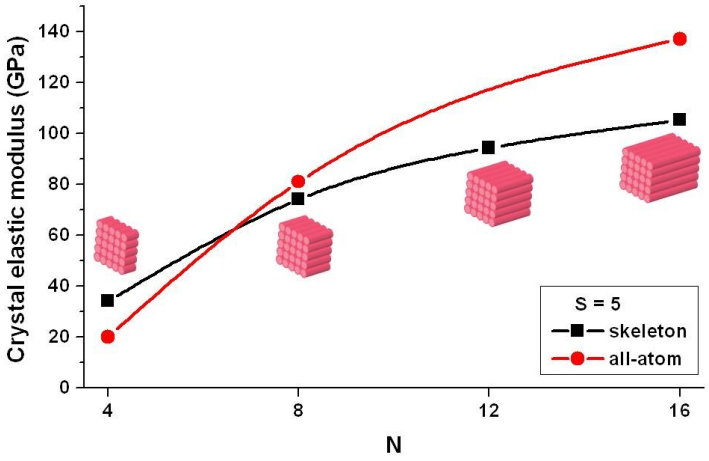

Figure 3.21: A comparison of the all-atom and skeleton models. The increase in the axial elastic modulus with respect to the backbone length, N is shown. Smooth lines are spline fits to the data points.

Toughness is a measure for the energy required to rupture these structures and can be inferred from the stress-strain profile until rupture as observed for the all-atom models in Molecular Dynamics simulations. The toughness follows a trend similar to that observed for the crystal elastic modulus. Crystal toughness increases with higher N (Fig. 3.22a) and decreases with higher S values (Fig. 3.22b). For instance, the S=5, N=8 structure has 48.8 MJ/m^3 toughness, whereas the S=5, N=4 and S=7, N=8 structures have 26.5 and 35.4 MJ/m^3 toughness, respectively. Unlike the crystal elastic modulus, crystal toughness does not yield an optimum geometry in terms of the contribution per residue within the probed range of crystal size, but continues to increase (decrease) with N (S). Considering the same crystal sizes, our toughness values overlap with the resilience values reported by Keten et al. (127) Resilience is calculated by considering only the elastic deformation region, while toughness involves the total energy input until the rupture point. In our simulations we have observed that the crystalline subunits showed little or no plastic deformation as suggested by a linear slope in the stress-strain curves (Fig. 3.20a). This renders toughness and resilience to

be highly similar for silk crystals, in contrast to previous observations which, however, might be driven by boundary effects. (127)

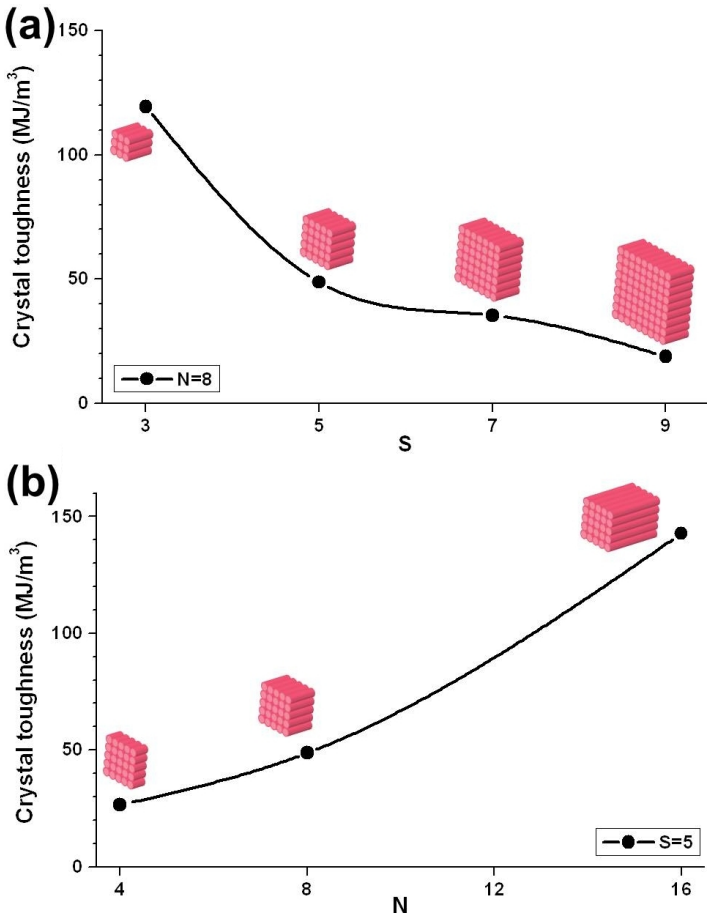

Figure 3.22: Toughness of the crystalline subunits. All-atom models show a decrease (a) and increase (b) with respect to the number of β-strands in the crystal, S, and the backbone length given by N, respectively. Smooth lines in (a) and (b) are spline fits to the data points.

The toughness and rupture properties of the crystalline subunits show size dependencies parallel to the crystal elastic modulus, as compared in Fig. 3.23. Taken together, they suggest that smaller S and larger N values yield stiffer and stronger crystals. In contrast to the previously suggested geometries (127; 148), we find that crystals with lower aspect ratios, namely with several nm *beta*-strand length and 1 nm cross-sectional diameter, exhibit the highest mechanical performance. To the best of our knowledge, there have been no experimental studies to compare our findings, except from the ob-

servation that the crystal length along the fiber axis is indeed of several nanometers 8-16 alanine residues. (77; 83) Hence, we consider our results as valuable predictions to be tested by future experimental work.

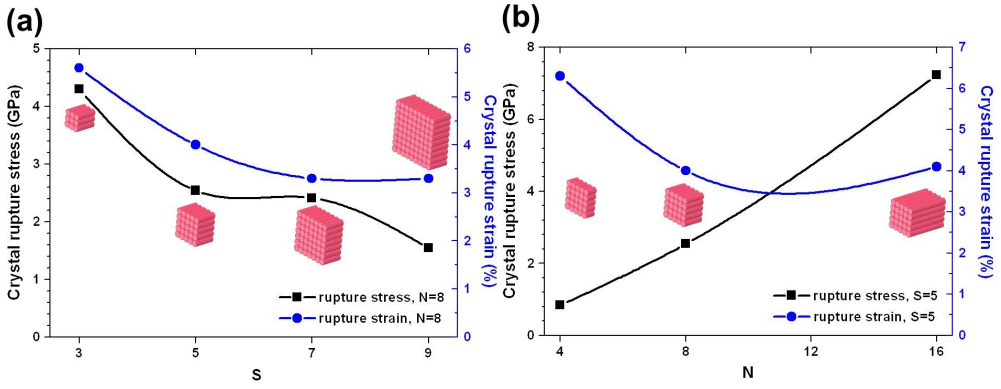

Figure 3.23: Rupture properties of the crystalline subunits obtained from MD simulations. (a) The variation of the rupture stress and strain with respect to (a) the number of layers in a crystal, S, and (b) the backbone length, N. Smooth lines in (a) and (b) are spline fits to the data points.

Mechanical properties of the fibers

Crystalline subunits are one of the main constituents of natural silk fibers, but their contribution to the fiber mechanical properties can not be straightforwardly predicted solely on the basis of the crystals' mechanical properties. Here, we employ FE analysis to determine the fiber elasticity and toughness for varying silk crystalline sizes. We limit our analysis by keeping the degree of ordering (25% crystallinity) and the structural architecture of the fiber (randomly distributed crystals) as fixed parameters. Although the total volume occupied by the crystalline subunits is the same in all fibers, on the fiber elastic modulus changes with the crystal geometry. Fig. 3.24 shows that a smaller number of layers (i.e. smaller S) yields stiffer fibers, and this effect is more pronounced as N increases. As an example, fibers having crystals with S=3 and S=7 have similar elastic moduli at N=4 (6.2 GPa vs 5.2 GPa) while the difference is larger at N=16 (13.0 GPa vs 7.5 GPa). In a coarse grained fiber model with a higher crystallinity rate (45%), Termonia has observed the same qualitative behavior for the fiber elastic modulus. (121) In an experimental study by Du et al., spider silk fibers are reported to become stiffer with crystals having smaller S values. (142) In contrast to our finding,

they also reported that smaller N leads to stiffer fibers, which, however, can be ascribed to the higher crystallinity of their fibers with shorter crystals.

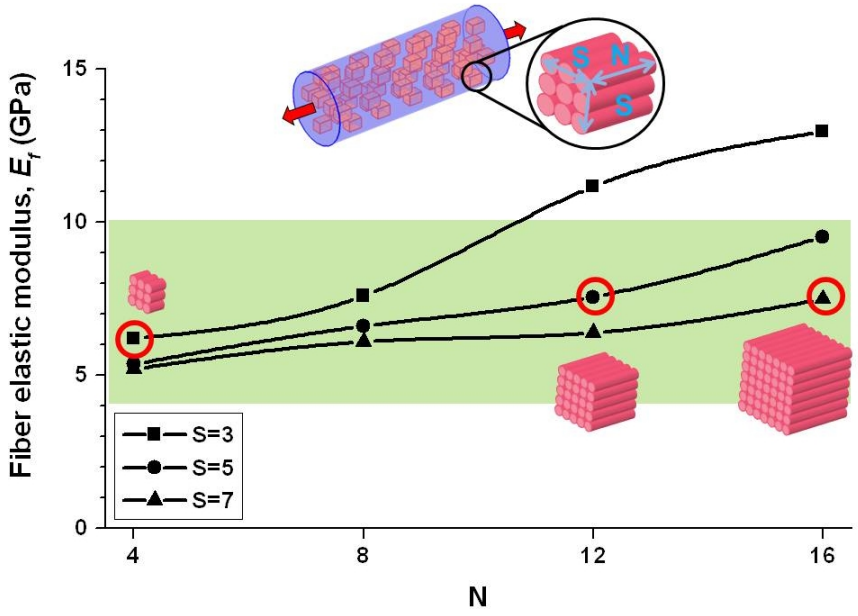

Figure 3.24: The variation of the fiber elastic modulus with respect to the crystalline subunit geometry obtained from FE calculations. Upper inset shows the pulling setup for the fiber and the green band indicates the range of experimental results in literature. (77; 133–135) Representative crystal sizes are shown as schematics and highlighted by red circles. Smooth lines are spline fits to the data points.

Our simulations with the all-atom and skeleton models showed that the elastic characteristics of the crystalline subunits resemble transversely isotropic materials. Consequently, the fiber model shows a non-isotropic behavior. Contrary to the general assumption of isotropic crystals, (143) the crystalline subunits are stiff in the axial direction and relatively weaker in the transverse directions while having low shear moduli. For instance, the skeleton model of the S=5, N=8 structure has an axial, transverse, and shear modulus of ∼74.0, ∼36.0, and ∼6.2 GPa, respectively. Fibers made of these crystals have 6.6 GPa axial tensile modulus, which falls well into the range of experimental values regarding spider silk, 4∼10 GPa. (77; 133–135) Moreover, the compressive modulus of these fibers in the transverse direction, 0.7 GPa, agrees well with the experimental results for spider silk, 0.6 GPa, (144) and silkworm silk, 0.7 GPa. (150)

We next assessed the second parameter crucial for mechanical performance, fiber rupture and resulting toughness, Despite the large number of experimental work addressing the toughness and rupture of natural silk fibers, (77; 133–135) there is limited information regarding size effects of the crystalline subunits. Although Du et al. reported the significance of the size effects on the silk fiber mechanics, the contribution of fiber crystallinity was not explicitly considered in their study. (142) We here systematically investigated size effect by varying the crystalline subunit size at constant crystallinity. Since our fiber model was based on conventional FE calculations, fiber rupture could not be directly simulated, but instead was indirectly estimated from the failure of crystalline subunits observed in the corresponding all-atom models (see Methods).

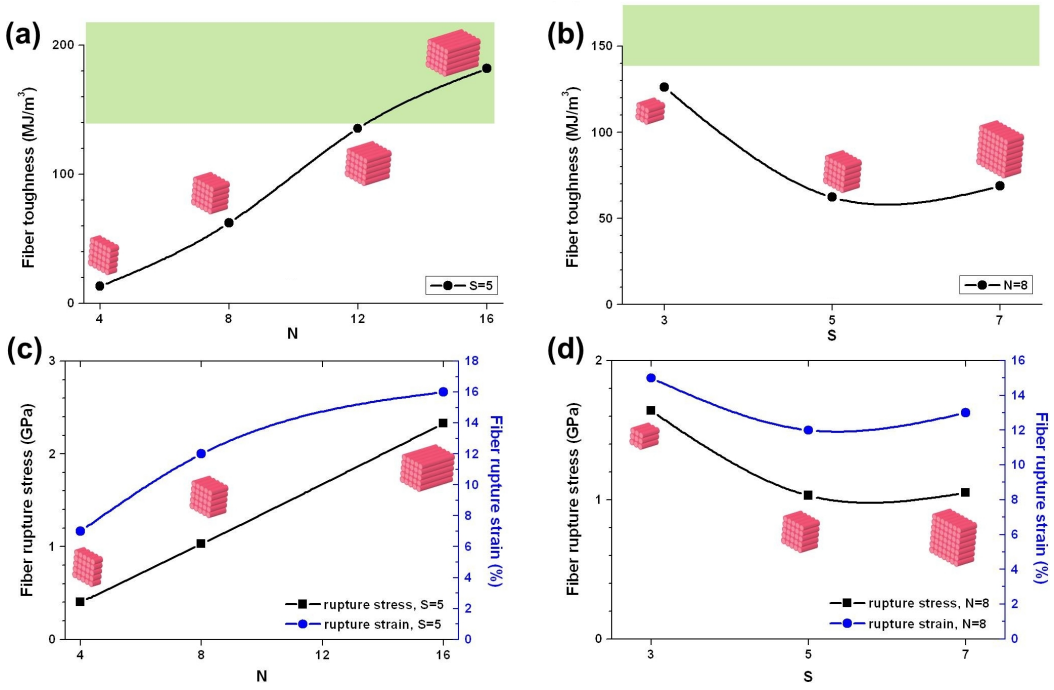

Figure 3.25: Fiber toughness and rupture properties obtained from FE calculations. (a-b) The variation of the fiber toughness with respect to the crystal size parameters, N (a) and S (b), is shown. Green bands in (a) and (b) indicate the range of experimental results in literature. (77; 133–135) Fiber rupture stress and strain curves with respect to the crystal size parameters, N (c) and S (d). Smooth lines in (a-d) are spline fits to the data points.

As the crystals within the fibers are composed of β-strand of increasing length

N, the fiber toughness increases as much as an order of magnitude (Fig. 3.25a). A similar behavior was observed for the all-atom models of the bare crystalline subunits (Fig. 3.22b). On the other hand, the fibers with wider crystals (i.e. crystals with higher S) tend to have lower toughness (Fig. 3.25b), again showing a similar trend as the bare crystalline subunits (Fig 3.22a). However, the size dependency is less pronounced for the parameter S, the cross-sectional area. The overall increase in toughness with increasing N and decreasing S originates from two parallel tendencies, namely higher ruptures strains and stresses (Fig. 3.25c and d). Overall, at moderate crystallinity, the highest fiber toughness is achieved with crystals composed of small layers of long β-strands. This observation is also valid for a serial (lamellar) arrangement of crystalline and amorphous subunits, but the toughness is of higher magnitude. For instance, fibers with N=8, S=5 crystals have a toughness of 215.0 MJ/m^3 with a serial distribution of crystals, as compared to 62.4 MJ/m^3 for the random arrangement conventionally assumed and considered in Fig. 3.25. In fact, we have already previously reported that a serial arrangement of subunits is superior over a random arrangement in terms of high fiber toughness. (149) As we now show here, nanometer sized crystalline blocks of small cross-sectional area are the components of choice within these lamellae, (86; 148; 151) the detailed arrangement of which on the molecular scale, however, is still open for discussion.

Conclusions

Here, we have quantified the crystalline size effects on silk fiber mechanics with a bottom-up computational approach. All-atom models of crystalline subunits have shown that the stiffness, toughness, and strength of the crystalline subunits increase with respect to the backbone length of the β-strands, N, in the crystal. These quantities,however, are inversely dependent on the number of the β-strand layers, S, in the crystals. All-atom models also showed that the crystalline subunits do not undergo plastic deformations, but instead feature a purely elastic regime up to the final rupture.

Finite element simulations of the skeletons models of the crystalline subunits were in quantitative agreement with the all-atom models at tremendously decreased computational cost. The skeleton models enabled us to determine the mechanically optimal size for the crystalline subunits. Based on the elastic modulus calculations, the N=8,

S=3 structure shows the most efficient usage of protein crystalline material to maximize crystal stiffness, in close agreement with the prevalence of roughly eight alanine residues in silk proteins and the experimentally determined crystal dimensions in silk fibers. (83)

Finite element simulations of the comprehensive fiber model allowed us to reach mesoscopic length scales at feasible computational cost in order to determine the contribution of the crystalline subunits on fiber mechanical properties. Being composed of transversely isotropic crystal blocks, the simulated silk fibers showed a similar material behavior. The calculated tensile and transverse elastic moduli were in good agreement with the experimental results in literature. Fiber stiffness is found to be strongly dependent on the crystalline subunit stiffness when S=3, and this effect is less pronounced for larger cross-sections. Fiber toughness substantially increases with the length of the strands in the crystal, but again is largely unaffected for higher number of layers (i.e. S>3).

In contrast to previous studies, other fiber parameters such as fiber crystallinity and structural architecture were constrained, allowing us to systematically scan silk fiber mechanics in function of crystal dimensions. Given the challenge to determine the dimensions of crystalline phases in fibers experimentally, our study provides a guide for artificial fiber design.

Our results on fiber toughness are based on three major approximations. First, the fiber rupture stress was directly estimated from the crystal rupture stress so that plastic deformations in the fiber following crystal rupture are neglected. Secondly, plasticity within crystals was not taken into account as well. Thirdly, the non-linear elasticity of the amorphous phase was approximated by a linear model. Therefore, toughness values provided here have to be considered as lower bounds of results to be expected in experiments (Fig. 3.25a and b). Including plastic effects in the crystals and the fiber, and employing model for the amorphous chains that takes connectivity and non-linearity into account will be important future steps. Despite the above drawbacks, the finite element models allow us to sample a large set of design possibilities with acceptable accuracy and affordable computational cost. Our computational approach combining MD simulations and FE analysis does not make use of any empirical parameters and is easily applicable to similar semi-crystalline systems such as non-biological block copolymers or composite systems.

3.1.5 Controlling the Structure of Proteins at Surface

With the help of single molecule force spectroscopy and molecular dynamics simulations, we determine the surface-induced structure of a single engineered spider silk protein. An amyloid like structure is induced in the vicinity of a surface with high surface energy and can be prohibited in the presence of a hydrophobic surface. The derived molecular energy landscapes highlight the role of single silk protein structure for the macroscopic toughness of spider silk. (69)

The experiments presented in this section have been performed in the lab of Prof. Thorsten Hugel, TU Munich.

Introduction

In many applications, proteins or polypeptides have to be immobilized at surfaces. (152; 153) To keep their functionality, conformational changes in polypeptides or even denaturation of proteins induced by substrates have to be avoided. Evidence has accumulated that already soft surfaces like lipid membranes can cause polypeptides to lose their functionality by forming amyloid-like secondary structures (154; 155) related to recognized clinical disorders, including type II diabetes, and Alzheimer disease. (156) Yet is it also possible to use this obstacle to our advantage by creating functionality through surface contact? Key to answer this question are molecular forces involved in the structure formation and how these forces are affected by thermodynamic parameters. Here, we use force spectroscopy to monitor structural changes with one and the same single polypeptide. We show that structure formation within a single protein, a genetically engineered variant of spiders dragline silk, eADF4, which consists of 16 identical repeats, (119) is induced in the vicinity of surfaces with high surface energy and can be prohibited in the presence of hydrophobic surfaces. In addition, we investigated how salts and temperature modulate surface-induced folding. The experimentally determined molecular structure agrees well with force probe molecular dynamics (FPMD) simulations. This surface-induced structure bears striking similarities to the conformation of amyloidogenic peptides in their fibrilar assemblies and gives insight into the molecular origin of the enormous energy absorption and strength-bearing properties of silk threads.

Materials and Methods

Experimental setup is shown in Fig. 3.26. The experimental material, eADF4, used in this study comprises 16 repeats of the sequence GSSAAAAAAAASGPGGYGPENQGPSGPGGYGPGGP and has a molecular mass of 47.7 kDa. (157) The preparation of this material and corresponding experiments are performed by cooperators of this study, with details are published. (158)

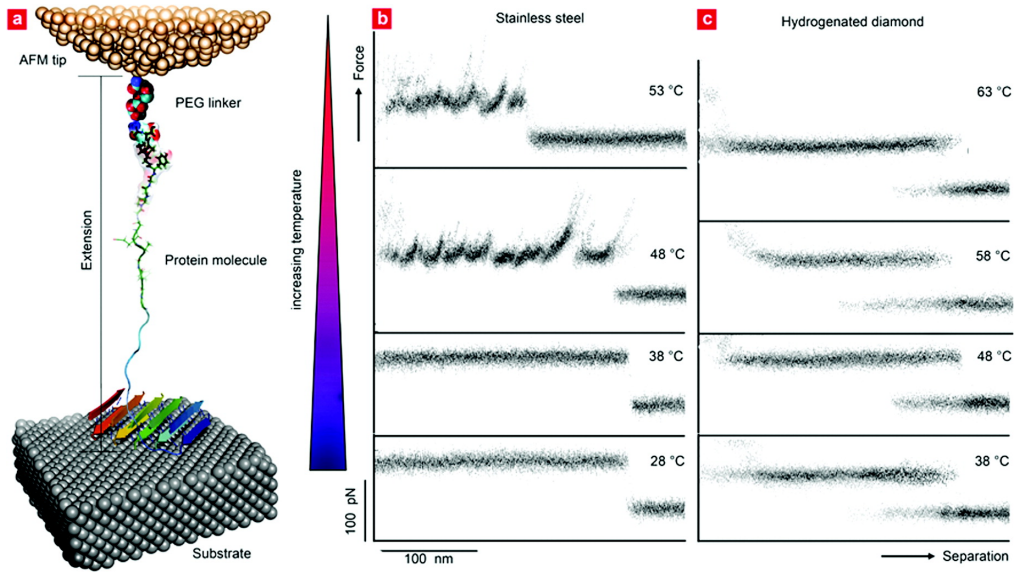

Figure 3.26: Monitoring surface-induced protein folding in eADF4. (a) Sketch of the single molecule desorption experiment of eADF4 attached to the AFM tip via a PEG linker (not to scale). Here, the molecule exhibits β-sheet structure on the steel surface. (b) Force-extension traces of a single eADF4 moleculein water at varying temperatures desorbed from stainless steel and (c) from hydrogenated diamond. Each graph shows the superposition of 20 consecutive force-separation curves

Modeling and equilibration of eADF4 β-sheets

Silk proteins are known to form beta-sheet patches via their poly(Ala) sequences. Multi-layer β-sheet stacks from spider silk $(Ala)_8$ have been modeled previously at atomic detail (72). The models were found to be both structurally highly stable and extraordinarily force-resistant, as well as to agree with X-ray diffraction data (83). The structural details of these models were published elsewhere (72).

Constructs consisting of two layers are the minimal building blocks to account for the experimental force distributions. As expected for a hydrophilic surface, the

polar backbone functions, namely the amide groups, were oriented towards the surface, resulting in a beta-sheet orientation perpendicular to the surface. Five of these two-stranded beta-sheets were packed against each other along the surface plane, allowing alanine sidechain packing parallel to the surface plane. Here, we are interested in the force-induced rupture of the poly(Ala) repeats in the eADF4.

Three different β-sheet patterns, namely an all-parallel structure (pp), a mixed structure of anti-parallel sheets with parallel orientations between adjacent β-sheets (ap), and all anti-parallel (aa), each consisting of two layers of β-strands, were modeled (Fig. 3.29). We used the OPLS-AA force field18 for the protein. A surface of 5×5 nm^2 in size consisting of Lennard-Jones spheres was placed below the two-layer systems. For this generic surface, ϵ=0.276 kJ/mol and σ=0.35 nm were used as Lennard-Jones parameters. The initial distance between protein and surface was 0.8 nm to ensure that atoms do not overlap, and to allow water molecules to fill the gap between protein and surface at the beginning of the equilibration. Because detailed knowledge of the molecular details of the steel surface (highly oxidized and thus polar under experimental conditions) is lacking, we kept the surface overall uncharged. Even though the modeled surface is apolar, i.e. does not provide favorable electrostatic interactions partners for the protein amide groups, the modeled structures remained highly stable within the simulation time at the nanosecond scale. We conclude that the surface-induced alignment was sufficient to keep the silk protein beta-sheet assembly in this energy minimum.

We used the GROMACS 3.3.3 package (53) for all subsequent molecular dynamics (MD) simulations. Simulation boxes of ∼7.5×7.5×6 nm^3 were used. The models were solvated in TIP4P water20 mimicking the experimental environment. The solvent included Na+ and Cl- ions with a concentration of 0.1 mol·l^{-1}, resulting in a total system size of ∼40,000 atoms. Periodic boundary conditions were employed to remove artificial boundary effects. We chose a cutoff of 1.0 nm for non-bonded interactions, and the particle mesh Ewald method (91) to account for long range electrostatics interactions. To increase the simulation time step, we used LINCS (92) to constrain all bond vibrations. The time step was 0.002 ps. Simulations were performed in the NpT ensemble with a temperature of T = 300 K and a pressure of p = 1 bar in all the simulations. We used Nosé-Hoover (93; 94) temperature coupling with a coupling time constant τ_T = 0.1 ps, and Berendsen (95) pressure coupling with a coupling time

constant of $\tau_P = 0.1$ ps. After energy minimization using the steepest-descent method, we performed 500 ps position-restrained simulations to further relax our simulation systems, subjecting each protein atom to a harmonic potential with a force constant of 1000 kJ·mol^{-1}·nm^{-2}. Each model was then fully equilibrated for 10 ns to allow the protein to bind readily to the surface. Energy and coordinates of the simulation systems were collected every 1000 time steps. The resulting equilibrated simulation systems served as starting points for force-probe simulations.

Force-probe MD simulations

To monitor force-induced unfolding and to determine unbinding forces for the different self-assembled silk protein structures, we performed force-probe molecular dynamics (FPMD) simulations (66). Final equilibrated structures obtained from the free MD simulations of the three systems were exposed to an external stress to monitor rupture. To accommodate the pulled peptide also after the rupture occurred, we increased the initial box dimension along the pulling direction to 12.0 nm, resulting in a system size of ∼90,000 atoms. The terminal residue of different strands in the 5×2 β-strand structure was subjected to a pulling force acting perpendicular to the surface by moving a spring with constant velocity away from the protein fold, i.e. the central strand in the upper and lower layer were pulled in the direction orthogonal to the surface plane. Linear and angular motions of the center of mass of the surface were removed during the simulation to serve as a reference. A pulling velocity of 0.2 nm·ns^{-1}, and a spring constant of 500 kJ·mol^{-1}·nm^{-2} were used. The simulated time of all the models was 15∼20 ns each, depending on the resistance against unfolding. These non-equilibrium MD simulations used the same simulation parameters as the equilibrium simulations (see above). Representative simulation systems are shown in Fig. 3.29.

For each configuration, pp, ap and aa, respectively, three independent pulling simulations were performed and the rupture forces and transition state distances were averaged. We determined the distance to the transition state Δx as the difference in distance along the pulling direction (reaction coordinate) between the pulled group and the surface as it changed from the equilibrium conformation to the onset of the unfolding event.

Amorphous chain segments

We did not include the amorphous chain segments connecting the $(Ala)_8$ β-strands into the simulation, because we focused on the force-induced unfolding of the ordered

structural elements giving rise to the distinct pattern of force peaks in the force-extension traces of our AFM experiments, rather than the underlying plateau stemming from equilibrium desorption. However, as a reference, we also performed FPMD simulations of a single disordered peptide comprising both the $(Ala)_8$ motif and the amorphous chain sequence. The resulting force profile is shown in Fig. 3.27.

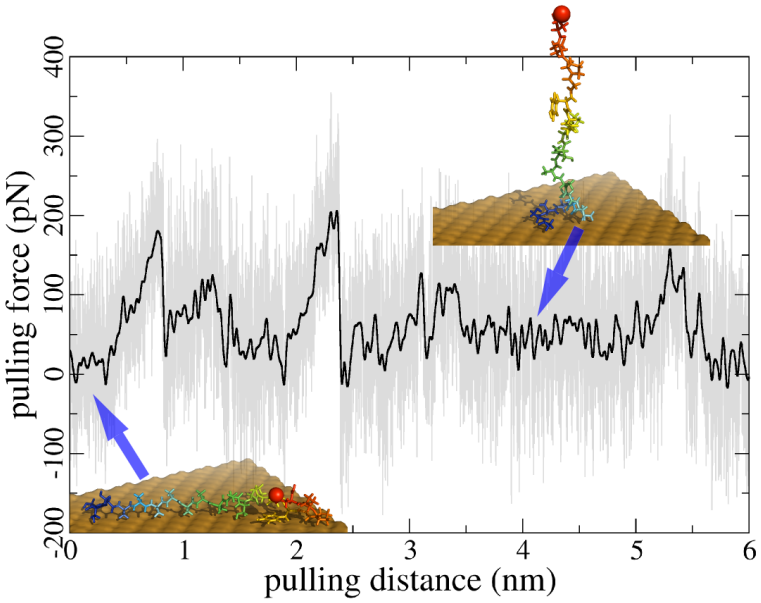

Figure 3.27: Force profile for pulling a single silk peptide from the substrate surface in the absence of secondary structure formation such as beta-sheets. The rupture forces are overall lower and correspond to the plateau force observed experimentally.

Results and Discussion

For the following experiments a single eADF4 molecule has been covalently attached with its N-terminus via a long flexible linker of poly(ethylene glycol) (PEG) to an atomic force microscopy (AFM) cantilever tip (Fig. 3.26a). Covalent bonding provided the necessary long-term stability to allow measurements for hours with one and the same protein on different substrate materials in solution. The tip with the protein was brought in contact with the surface for 1 s to adsorb and possibly to form its fold before being retracted at a velocity of 1 μm/s.

Up to 48 °C, we obtain velocity independent force plateaus (Fig. 3.26b) with a length equal to the contour length of the protein. Accordingly, the desorption of eADF4 from

its N-terminus to its C-terminus occurs in equilibrium and stable structural features are absent (159). As can be seen from Fig. 3.26b, the traces above 48 °C on stainless steel 316 L, which is one of the most common alloys in medicine, are considerably different. Here, a regular pattern of peaks with amplitudes ranging from 70 to 200 pN appears on top of the desorption plateaus. Such regular peaks are the characteristic fingerprint for unfolding secondary structures (8; 160). We still probe the full contour length of eADF4 at 48 °C, and hence the unfolding of all structural elements. The observation of a desorption plateau underlying the peaks is a fingerprint for a very high mobility in the surface plane and excludes the presence of distinct adhesion sites in this plane. In other words, the peptide is removed (perpendicular to the surface plane) in steps, which are defined by the stability of its secondary structure, while it is highly mobile in the surface plane.

All data in Fig. 3.26b were recorded with one and the same eADF4 molecule as indicated by the single step drop of the force plateau to the zero line. (158) The only parameter that was varied was the solution temperature. We identified the temperature around 48 °C as the transition temperature for structure formation on hydrophilic stainless steel. A second set of experiments in Fig. 3.26c substantiates that the structure is not formed in solution but is surface-induced, because the same experiments on hydrophobic, hydrogenated diamond show no conformational transition (even up to 63 °C). Not only temperature, but also a higher ionic strength of phosphate of about 100 mM can facilitate the transition in line with previous bulk experiments, (161) but again only on the hydrophilic metal substrate. The reason probably is the polar character and thus the high interfacial energy of steel of 38 mJ/m^2 as compared to 18 mJ/m^2 for the diamond substrate (determined by contact angle measurements).

What are the characteristic features underlying the surface induced structure of eADF4? Following previously described methods, (70) we transformed the 20 force-extension traces from Fig. 3.28a into contour length space and obtained histograms that directly reflect particular locations along the polymer contour that constitute an energy barrier for unfolding (Fig. 3.28b). The histogram shows 16 equidistant unfolding barriers corresponding to the 16 repeats of eADF4 plus one terminal barrier originating from the detachment of the C-terminus from the surface. For a detailed analysis of the force distributions, well-defined loading rate conditions (constant spacer lengths) are essential. (162) Therefore, all 16 rupture peaks have to be analyzed separately.

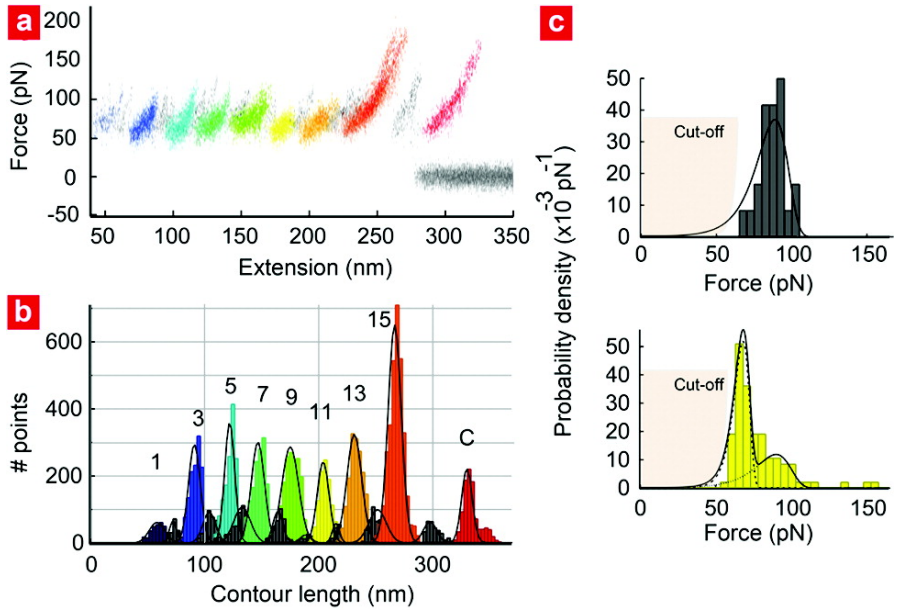

Figure 3.28: Forced unfolding of the eADF4 structure. (a) Superposition of 20 force-extension traces of eADF4 with odd peaks being marked in color. The traces were recorded in salt solution on steel (316 L) at room temperature. (b) Averaged barrier position histogram obtained by transformation of each of the 15 unfolding peaks with a persistence length of 5.1 Å into contour length space and averaging the 20 resulting barrier position histograms for each peak. (c) Unfolding force distributions and normalized unfolding probability density for two sequential unfolding barriers (same color coding). The shaded area in the lower force regime illustrates the resolution limit of the instrumentation (cut-off). The best fit to the data (solid black) reveals two discrete barrier binding potentials: a low force unbinding around 75 pN (dashed) and high force unbinding around 90 pN (dotted).

Fig 3.28c displays two consecutive force histograms out of the 16 together with their probability functions fitted to the data based on the assumption of a one-barrier binding potential (163). These histograms differ from odd to even unfolding events. The former have to be fitted with two separate potentials, while the even barrier numbers can be fitted with one. The potential widths derived from these probability functions are all summarized in Table 3.3. We find a short potential width of 4.7 Å for high force unfolding and a longer width of 8.9 Å for all low range forces. Only the last two values are different due to the influence of the C-terminus interacting with the substrate. Our study of the conformational transition caused by increased salt concentration on steel yields similar values and suggests an equal folding mechanism. The high mobility in the

plane of the surface and the two different barriers cannot be explained by the strong interaction of a single amino acid in a repeat unit of eADF4. The slight broadening of the rupture force peaks and the analogy of the transition barrier distances to diffraction patterns of amyloidogenic structures make the $(Ala)_8$ repeats very likely candidates for the observed structure.

Table 3.3: Potential Width (Δx_{pp} and Δx_{ap} (Å)) and Unfolding Rate Constants (k_{pp}^{off} and k_{ap}^{off} ($10^{-2}s^{-1}$)) at Zero Force (Water, 48 °C) for Parallel and Anti-parallel Arrangements, Respectively[a].

	Event No.												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Δx_{pp}	8.5		8.6		8.7		9.0		8.3	9.9	9.3	3.0	4.6
k_{pp}^{off}	1.3		1.5		1.4		1.4		1.5	1.6	1.5	1×10^3	2×10^2
Δx_{ap}	4.7	4.0	5.1	4.8	4.9	4.9	4.5	4.9	4.5	4.6	4.8	1.4	1.7
k_{ap}^{off}	3.0	3.0	3.5	3.6	3.0	3.6	3.0	3.9	3.2	3.5	3.0	55	55

[a] The unbinding events are numbered sequentially starting at the N-terminus.

To relate these data to three-dimensional surface-adhered structures of eADF4, we therefore modeled structures formed by the amyloidogenic repeat units of eADF4, that is, $(Ala)_8$, to assess their mechanical response using FPMD simulations (see methods). Force fields for steel-protein interactions are not available yet,16 but previous studies of amino acids on noble metals have found polar interactions to prevail. (164; 165) We here assume steel to mainly serve as a template, inducing order in the poly(Ala) segments similar to the role of shear flow in silk fibrillogenesis. The backbone amide groups in (Ala)8 are polar and therefore likely candidates to orient toward the hydrophilic metal surface. We assembled a two-layer β-sheet, consisting of $(Ala)_8$ β-strand motifs connected by interstrand hydrogen bonding, that are stacked against each other via side-chain interactions parallel to the surface plane (Fig. 3.29a). Coil domains of eADF4 were omitted in the simulations. The various relative orientations in a double layer can be grouped into antiparallel, parallel, and mixed arrangements (Fig. 3.29a).

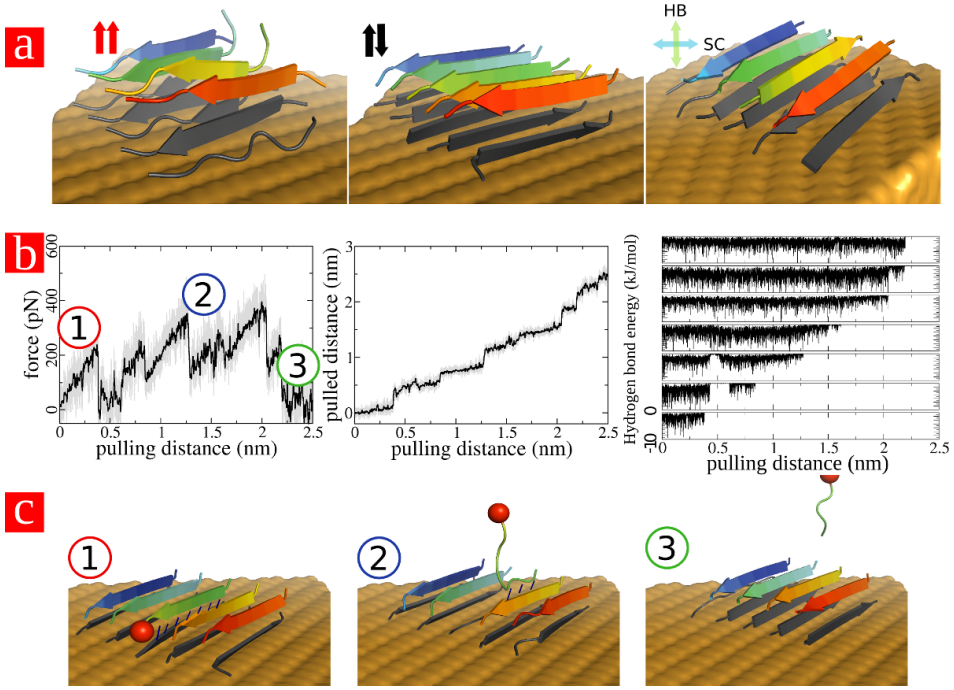

Figure 3.29: Surface-induced structures probed by force-probe MD simulations. (a) Modeled two-layer $(Ala)_8$ arrangements: all parallel on the left, mixed in the middle, and all antiparallel on the right. Red and black arrows make the connection to the energy landscapes in Fig 3.31a. Directions of hydrogen bonds and side-chains are indicated. (b) Representative FPMD simulation of rupturing one strand from the upper layer of the antiparallel structure. Shown are the force profile (top), the displacement of the pulled group from which Δx is estimated (middle), and potential energies of hydrogen bonds between the pulled strand and its neighbors (bottom). (c) Simulation snapshots, with hydrogen bonds shown as blue sticks, are numbered to indicate the corresponding position on the force profile.

From FPMD simulations, we find independently from the detailed arrangement strands in the upper layer of the crystalline structure that detach with forces of 250–450 pN. These forces are due to the rupture of strong hydrogen bonds to the adjacent bottom strand (Fig. 3.29b). The subsequent rupture of the bottom strand, not stabilized by interstrand hydrogen bonds anymore but only adsorbed to the surface, does not give rise to a force peak. This also implies that a one-layer arrangement would lead to a mere force plateau and thus cannot explain the experimental force pattern. We note that the experimental distances to the transition state are wider than those from simulations by a factor of about 3, most probably due to temperature soften-

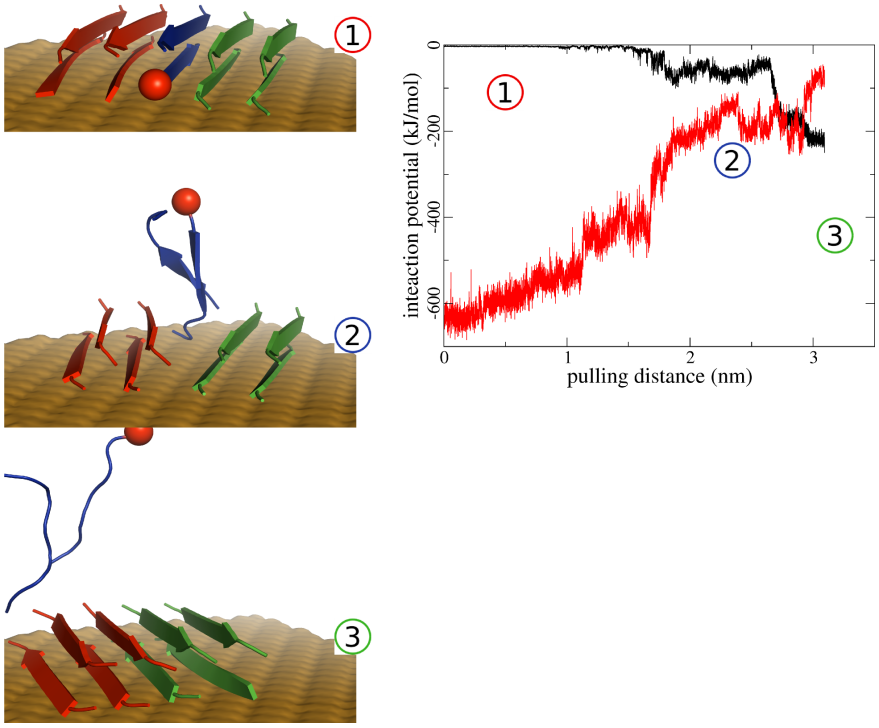

Figure 3.30: Self-healing of ruptured silk protein assemblies during FPMD simulations. Snapshots show how the beta-sheets (red and green) adjacent to the pulled strands (blue) form new direct interactions with each other after rupture, leading to a selfhealed structure. The plot shows the changes in potential energy during rupture for the interaction of the pulled strands with protein and surface (red) and for the newly forming interaction of the two adjacent sheets with each other (black).

ing, (166) but the relative changes match very well. Strikingly, strands in antiparallel arrangements show higher forces and smaller transition distances than those in parallel arrangements, as previously found for the structurally related silk crystalline units in bulk. (72) These differences qualitatively agree with the two discrete unfolding barriers for odd numbered peaks observed in our AFM experiments (Fig. 3.28c and Table 3.3). One also has to consider strand-turn-strand motifs to be surface-assembled such that the pulling force acts onto the strand in the lower layer first, resulting in a concurrent detachment of the whole motif and therefore causing higher forces, as observed for the rare peaks with even numbers in Fig. 3.28b and in corresponding FPMD simulations. Finally, the plateau of constant force underlying the unfolding pattern stems from the desorption of the glycine-rich coiled segments between the poly(Ala) patches,

which behave like a nonstructured flexible polymer with high in-plane mobility. When detaching a β-strand-coil-β-strand motif from within the two-layer β-sheet system as described above, we observe the structure to self-heal by sealing the introduced flaw on the nanosecond time scale of the simulations, as depicted in Fig. 3.30.

Taken together, our structural model of a single eADF4 molecule derived from force spectroscopy matches the MD simulations. It also shows remarkable similarities with eADF4 assembled into nanofibrils and into microspheres (167) and with typical protein structures within amyloid aggregates. (168) Our combined experimental and theoretical data suggest parallel β-sheets to outnumber antiparallel counterparts in the surface induced assembly. This is in contrast to the observed preference of antiparallel β-sheets of many amyloidogenic peptides, (168) but consistent with the parallel/antiparallel mixture in silk fibres determined by Asakura et al. (85)

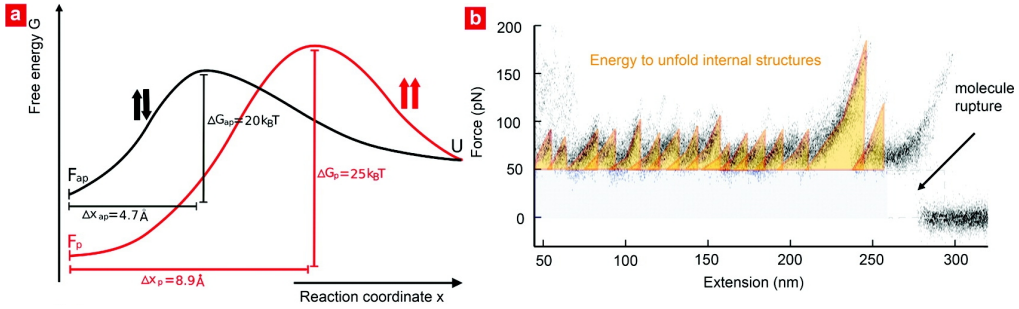

Figure 3.31: Strength-bearing properties and energy absorption of eADF4. (a) Schematic energy landscapes with potential widths (Δx) and energy barrier heights (ΔG for unfolding parallel (red) and antiparallel (black) sheets from the folded (F) to the unfolded (U) state. (b) The area under the superposition of 20 typical force-extension traces of eADF4 is divided into the free energy contribution due to the equilibrium desorption (gray) at negligible energy dissipation (friction) and the energy that is needed to break individual elements of the secondary structure (orange) on steel 316L at 48 °C.

In the following, we quantify the stability of the molecular structure and its origin. The parallel structure exhibits a very low mean transition rate of $1.5\times10^{-4}\mathrm{s}^{-1}$ and therefore high stability, which compares, for example, to that of the strong immunoglobulin fold of the muscle protein titin, (8) to ubiquitin, (169) or to the extracellular matrix protein tenascin. (170) In contrast, the mean transition rate of the antiparallel β-sheet formed by eADF4 ($3.3\times10^{-2}\mathrm{s}^{-1}$) is 2 orders of magnitude higher and is similar to those of calmodulin (19) and F-actin cross-linker filamin. (166) Assuming an Arrhenius

prefactor of $10^7 s^{-1}$, the transition energy can be calculated (Fig 3.31a). Remarkably, the parallel structure, which breaks first under load, is stronger and hence more stable at zero force and will therefore rapidly reform the moment the molecule relaxes. (116) The wider potential renders it more susceptible to force-induced destabilization. (171)

Conclusion

We have shown that single engineered silk proteins can form secondary structures in a solid substrate, salt, and temperature-dependent manner. On hydrogenated diamond no secondary structure was observed, while the stainless steel induced an amyloid like secondary structure. A control of such surface-induced structures should not only help to sustain the functionality of protein coated devices, but also to create new ones. Finally, our combined approach of single molecule force spectroscopy and MD simulations allowed us to determine the response of the surface-induced structure to an external force and provides an explanation for the toughness of spider silk threads.

3.1.6 Mechanical Comparison of Polyamide Crystals by Molecular Dynamics Simulations

Crystals of nanometer sizes are found in all industrial polyamide materials similar to those in silk. These crystals are highly ordered polyamide strands serving as force bearing subunits in the material. The molecular structures of these crystals feature inter-strand hydrogen bonds and hydrophobic interactions, which are reminiscent of crystals found in silk proteins. Despite of their importance for the material, the nano-mechanics of these crystals have remained unclear. This study uses all-atom modeling and Molecular Dynamics simulations to investigate different types of crystals in different polyamides, namely polyamide-2, polyamide-4, polyamide-6 and polyamide-6,6, aiming at understanding their structural mechanics and stability. By mechanical comparison of these polyamide crystals with similar crystalline structures as in silk, their crystallinity, robustness and elastic moduli are elucidated. This study can be regarded as a basis for understanding and mechanically improving semi-crystalline materials.

Introduction

Polyamides are among most important materials in the society nowadays. Different types of polyamides, such as nylon (polyamide-6), polyamide-6,6 and others, are used in almost all aspects of daily life that make them indispensable to mankind. The multiformity of polyamide materials originates from their molecular structures which are comprehensively crystal-embedded amorphous peptide matrices.(172–174) Crystals in polyamides serve as reinforcing subunits that maintain the integrity of the random peptide structures under load. Different polyamides show diverse morphologies of crystallinity, and thus mechanical properties. Despite this general understanding of polyamides, molecular details of the peptide chain organization in crystals and the interactions that bring about the structural stability are still not clear. As increasing attention is paid to similar natural semi-crystalline materials such as silk (46; 47; 72; 105), studies on polyamides, from atomic subunits to overall material performance, are required for a better understanding and more importantly mechanical improvement of polyamide materials.

Polyamide chains are unbranched and consist of thousands of residues. Polyamide residues which have no side-chain are covalently bonded via peptide bonds (-[C=O–N-

H]-). (175; 176) The number of carbon atoms in the backbone of a polyamide residue defines the type of polyamide, such as polyamide-6 has six carbon atoms in its residue while polyamide-4 has four carbon atoms, as shown in Fig. 3.32. All carbon atoms that are not involved in the peptide bonds are ethenyl groups and hydrophobic in nature. Some of the polyamides are co-polymer with two types of residues like polyamide-6,6. In polyamide-6,6, one residue has amide groups at both sides and the other has carboxyl groups (Fig. 3.32).

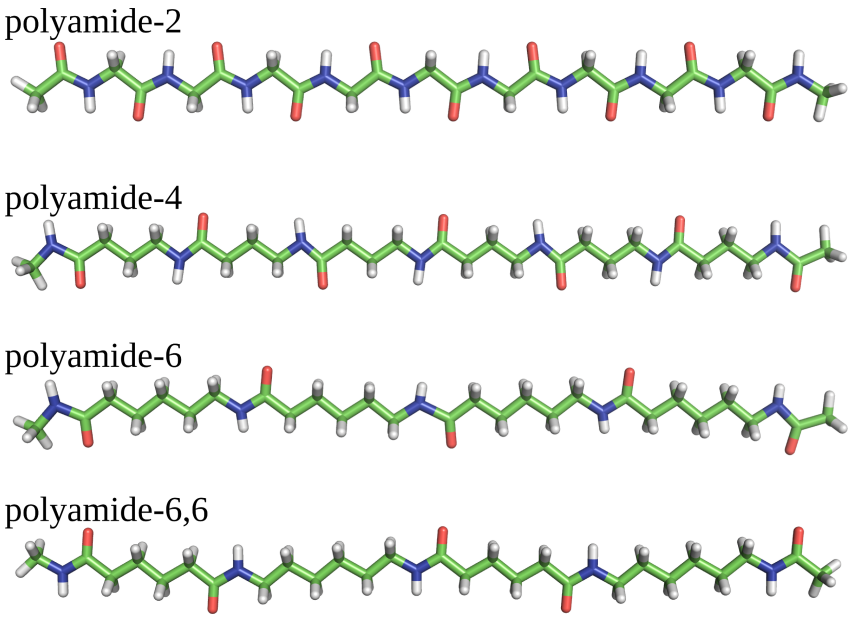

Figure 3.32: Four all-atom chain models of different polyamides. Carbon atoms are shown in green, oxygen atoms are in red, nitrogen atoms are in blue, and hydrogen atoms are in white.

Polyamide materials are generally semi-crystalline. (172–177) More specifically, a large part of the polyamide chains form unstructured random coils, while some parts of the chains are ordered in arrays via a high number of hydrogen bonds. These highly ordered peptide chain arrays are the crystalline subunits. These crystalline units are stiff because of hydrogen bonds and hydrophobic packing. Former studies of similar molecular organizations of silk proteins (see Chapter 3, Sections 3.1.3) have suggested that both percentage and arrangement of crystals in amorphous peptides can highly affect material performance. (47) Polyamide materials share a similar molecular architecture with silk protein, which makes the morphology of polyamide crystals, both

strand arrangement and mechanics, particularly interesting concerning the possibility of improving polyamides' mechanical performance.

This study focuses on molecular modeling and mechanical simulations of crystals of four different types of polyamides (Fig. 3.32). All-atom polyamide chain models are built and parameterized based on a universal force field. Crystalline models are then built by arranging copies of chain models into ordered arrays to capture the features of crystals in polyamide materials. All the crystalline models are relaxed by structural energy minimization and equilibrated by Molecular Dynamics (MD) simulations. With validation from experimental results, all the equilibrated crystalline models are submitted to force-probe MD (FPMD) simulations to determine their mechanical properties. The simulation results are compared with similar studies in crystals of silk protein. Patterns in mechanical properties of different polyamide crystals in this study can be used to achieve a better understanding of relevant materials and to lay the foundation for further studies.

Methods

All-atom modeling and parameterizing of polyamide chains

Polyamide (PA) chains are modeled by modifying similar protein peptide chains. A fully extended poly-glycine chain of 10 residues is taken as a model of the polyamide-2 chain. Other polyamide chain models are built simply by inserting different numbers of ethenyl groups into the backbone of the poly-glycine chain depending on the chemistry of the target polyamide chain model, namely polyamide-4, polyamide-6 and polyamide-6,6 (two types of residues). In order to allow systematic comparisons, the lengths of all models are kept within ∼30 atoms. Both ends of all chains are capped with acetyl or methlamide groups to keep them neutral, thereby avoiding artificially high Coulombic interactions. This resulted in final lengths of all PA chain models of ∼3.6 nm, as shown in Fig. 3.32.

Parameters for new polyamide residues are based on the OPLS-AA force field (40), which allows direct comparison to the respective protein simulations. Bonded interaction parameters of the peptide-bond atoms are kept the same with those in the protein, which feature the planarity of the four peptide bond atoms, namely oxygen, carbon, nitrogen and hydrogen. New parameters, atom types for carbon and hydrogen atoms, angle bending potentials as well as dihedral rotation potentials, are introduced into the

new polyamide residues accordingly. Details in force field parameters such as dihedral angles between ethenyl groups in the backbone also feature other quantum mechanics study results (178). Finally, five new residues (polyamide-6,6 has two types of residues) are introduced into the OPLA-AA force field, and are used in all subsequent MD simulations.

Building of polyamide crystals

Models of polyamide crystals are built by assembling copies of the four chains models, respectively. Same copies of each four chains are either arranged in a parallel or anti-parallel pattern to reach ordered structures. Thus, eight crystalline structures are built from four chain models, namely parallel polyamide-2 (pPA-2), polyamide-4 (pPA-4), polyamide-6 (pPA-6), polyamide-6,6 (pPA-6,6) and anti-parallel polyamide-2 (apPA-2), polyamide-4 (apPA-4), polyamide-6 (apPA-6), polyamide-6,6 (apPA-6,6). Each model has five layers of strands with five strands in each layer. One sample of the built structure of pPA-4 is shown in Fig. 3.33. The atoms in all these models are arranged such that they do not overlap with each other. These initial structures are then subjected to structural energy relaxations and subsequent MD simulations.

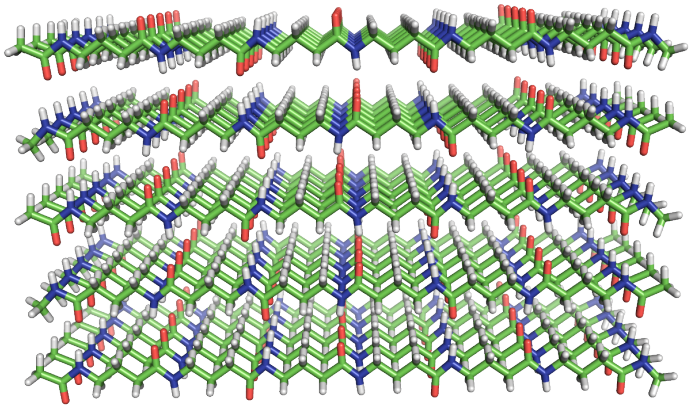

Figure 3.33: Initial all atom crystalline model of pPA-4. The strands are positioned without any atomic overlap.

Crystalline structure optimization

The Gromacs 4.0.5 package (53) with the modified OPLS-AA force field (40) containing newly built polyamide residue parameters are used for all simulations. All crystalline models are accommodated in simulation boxes treated with periodical boundary

conditions. All simulation boxes have additional space between boxes edges and the polyamide models to avoid artificial interactions between molecular images, which is 1.5 times of a non-bonded interaction cutoff distance of 1.0 nm. The Particle-Mesh Ewald method (91) is used to account for the long-range electrostatic interactions. In order to increase the simulation time step, LINCS (92) algorithm is used to constrain all bond vibrations. The time step was 0.002 ps. All the simulations were performed in the NpT ensemble with a temperature of $T = 300$ K and a pressure of $p = 1$ bar. Nosé-Hoover (93; 94) temperature coupling method and Berendsen (95) pressure coupling method are used with a coupling time constant of $\tau_T = 0.1$ ps and $\tau_p = 1$ ps, respectively.

All crystalline models are first relaxed by energy minimization and short MD simulations in vacuum. The models are subsequently solvated in TIP4P water (60). The solvent includes Na and Cl ions with a concentration of 0.1 mol/liter, resulting in a system size of ∼30,000 atoms. After energy minimization using the steepest descent method to remove possible atomic close contacts, position restrained simulations to further relax the simulation systems are performed for another 500 ps, subjecting each heavy atom in the polyamide chains to a harmonic potential with a force constant of $1000\,\mathrm{kJ\,mol^{-1}\,nm^{-2}}$. Each model is then fully equilibrated using Molecular Dynamics (MD) for 10 ns. Energy and coordinates of the simulation systems are collected every 1000 time steps. The resulting equilibrated simulation systems are then selected as starting structures for Force-probe MD (FPMD) simulations (see below).

All molecular structures are stable during MD structural equilibration as shown in Fig. 3.34. The highest structural root-mean-square deviation of ∼0.4 nm among the models is observed during simulations. The structural fluctuations are mainly because of the vibration of chain terminals. The equilibrated structures of all the models after a simulation time of 10 ns are found to well reproduce the main characteristics of polyamide crystals. Distances between peptide stands and between layers of sheets are found to be 0.48 nm and 0.37 nm, respectively, which are in good agreement with previous experimental studies (179; 180).

Force-probe MD simulations

All equilibrated crystalline models are further subjected to a pulling force in FPMD simulations (96). The terminal residue of the central strand in each model is attached to a harmonic pulling potential with a force constant of $500\,\mathrm{kJ\,mol^{-1}\,nm^{-2}}$. The harmonic potential moves along the strand direction at a pulling speed of $0.2\,\mathrm{nm\,ns^{-1}}$, generating

mechanical force by the increasing displacement from the force application point. A counter force is applied to the center of mass of the other stands to prevent translation of the model because of the pulling force. The simulated time for monitoring the rupture of the pulled strand is 20∼25 ns, depending on the PA. These FPMD simulations used the same simulation parameters as those in the equilibrium simulations (see above).

The backbone pull-out resistance, R_{bb}, is used to quantify the stiffness of the PA crystals. (72) R_{bb} is in analogy to the elastic modulus typically given to measure the stiffness of a material, as being defined as

$$R_{\mathrm{bb}} = \frac{\mathrm{stress}}{\mathrm{strain}} = \frac{F/A}{\Delta\, l/l}, \tag{3.6}$$

where F is the force acting on the strand, and A is the cross section area of the pulled strands. Here, the cross-section area is defined as the area covered between adjacent strands, which gives $A = 0.69\,\mathrm{nm}^2$. Measuring the elastic modulus of crystals in polyamides is challenging in experiments because of their nanometer size. Only the Young's modulus of whole semi-crystalline fiber is available.

Results and Discussion

Crystallinity of polyamides

All eight crystalline models are stable during structural energy minimization and subsequent MD simulation as described in the Methods section. The structures show obvious fluctuations when compared to all-atom crystals from silk (72), with an RMSD reaching as high as 0.4 nm. In general, anti-parallel models are more ordered than parallel models in general. Final equilibrated structures of the four parallel crystalline models are shown in Fig. 3.34. The most extended strands are found in pPA-6,6, with more ordered hydrogen bonds and strand arrangements, and thus higher crystallinity than other models. The crystalline structure is not as stable as their silk protein counter partner (72), as strand bending, twisting and tilting are observed during the simulations, especially in pPA-6. Despite of the structural flexibilities, hydrogen bonds are firmly established in all models.

All-atom silk protein crystalline models are much more ordered and stable, as previously reported (72). A major difference between silk and polyamide crystalline models are the side-chains, because silk proteins consist of protein peptide sequences of polyalanine or alternative alanine-glycine. When comparing the structure between pPA-2

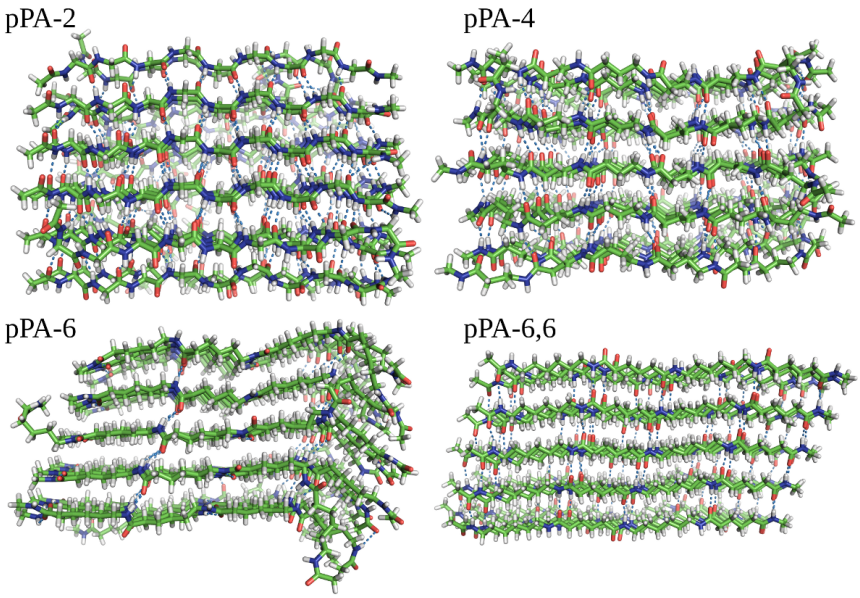

Figure 3.34: Equilibrated parallel polyamide crystalline models. Hydrogen bonds are shown as blue dashed lines to indicate their bonding patterns. The disorder of pPA-6 as compared to the others is obvious.

and parallel crystals of alternative alanine-glycine in silk, it is evident that the methyl group side-chains of the protein peptide are the key maintainers of the crystal (72). Alanine has a side-chain of one methyl group, which can fill gaps between strands (77; 122), thereby effectively decreasing the rotational freedom of peptides and stabilizing them.

Crystalline phase and hydrogen bonds

Hydrogen bonds are the mechanical determinants of crystals in silk, which also most likely applies to crystals in polyamides. It is known that different crystal forms are observed in polyamides materials. Two most commonly observed of these forms are the stable α-form and unstable γ-form (179; 180). The α-form of polyamide crystals is characterized by an in-line hydrogen pattern in which all hydrogen bonds are mostly orientated vertical to the strands. The γ-form of polyamide crystals features hydrogen bonds in a zigzag pattern. The α-form is found to be more stable than the γ-form because of the favorite hydrogen bonding positions and distances. Polyamides, such as polyamide-6, can perform transitions between the two crystalline forms under different conditions (181).

Hydrogen bonds are monitored in the PA crystals during all the whole MD simula-

tion trajectories. Conformational examples are shown in Fig. 3.35. In the eight models, especially in polyamide-6 and polyamide-6,6, the α-form is well produced. Fig. 3.35(A) and (B) show two hydrogen bonding patterns found in polyamide-6 starting from the α-form and the γ-form asinitial structures, respectively. After equilibration, the α-form is preserved in the sheet structure (Fig. 3.35(A)), while the γ-form has partly transformed (Fig. 3.35(B)). The transition in the γ-form of polyamide-6 confirms former studies that the γ-form is not a stable crystalline form. (179–181) A sample of the α-form of polyamide-6,6 is also shown in Fig. 3.35(A). Polyamide-6,6 favors the α-form for most of the simulation time because of its different copolymer peptide chain structure. This allows strong hydrogen bonding and also results in high resisting pull-out force (see below).

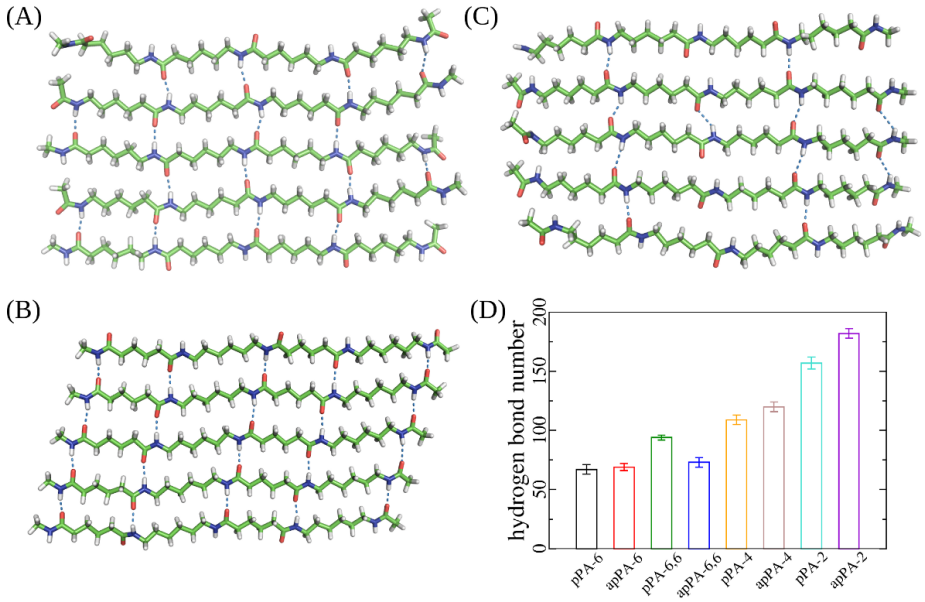

Figure 3.35: hydrogen bonds in crystalline models. (A) α-form crystal pattern in polyamide-6. (B)Partly α and γ-form of crystal pattern in polyamide-6. (C) α-form of crystal pattern in polyamide-6,6. (D) hydrogen numbers from all 8 models in equilibrium.

Hydrogen bond numbers differ in the eight models as expected (Fig. 3.35(D)). With more ethenyl groups separating neighboring peptide bonds, hydrogen bond densities decrease. Unsurprisingly, two polyamide-2 models, pPA-2 and apPA-2, show higher hydrogen bond numbers than the other models. Anti-parallel models, starting with the α-form crystalline pattern, feature higher hydrogen bond numbers than their par-

allel counter parts. Surprisingly, pPA-6,6 shows a higher hydrogen bond number than polyamide-6 models and its anti-parallel counter part, for which same hydrogen bond densities would be expected. The pPA-6,6 model maintains the α-form during the whole simulation. Thus the α-form is most likely the hydrogen bonding pattern in a polyamide-6,6 material. This stable hydrogen bonding pattern also explains, why polyamide-6,6 is more robust than polyamide-6 in terms of high melting point and toughness (see below).

Strand pull-out force and resistance

All eight equilibrated crystals are subjected to a pulling force in FPMD simulations. Rupture forces of the pulled strands are the highest force peak observed in the pulling force profiles and shown in Fig. 3.36(A). As a definition of the elastic modulus of the PA crystals, strand pull-out resistance is measured during the FPMD simulation (Fig. 3.36(B)) (72).

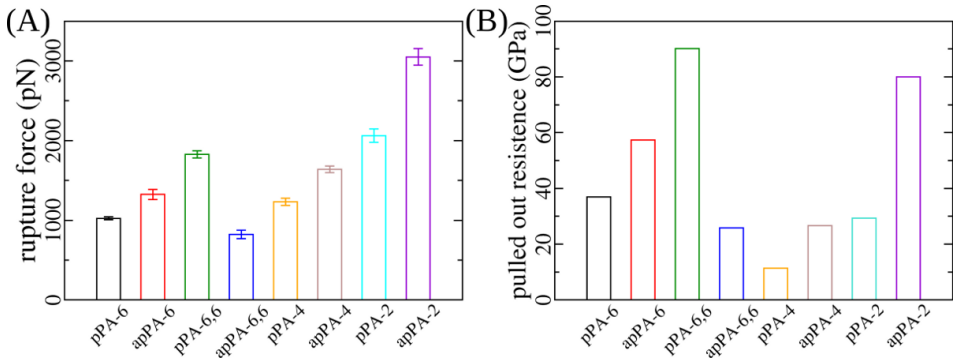

Figure 3.36: Strand pull-out forces and resistances of different PA-crystals.

The strand pull-out force correlates with the hydrogen bond number density in the models. A higher hydrogen bond density results in a higher rupture force, as shown in Fig. 3.36(A). pPA-6,6 again shows a higher strand pull-out force among the polyamide-6 and polyamide-6 models, apparently directly reflecting its higher hydrogen bond density (Fig. 3.35).

More efforts are needed to understand strand pull-out resistance in the different PA crystals. Surprisingly, pPA-6,6 shows the highest pull-out resistance of 90.1 GPa, even higher than apPA-2, i.e. poly-glycine, which has a higher hydrogen bond density, and already outperforms the stiffest silk crystal reported (~86 GPa) (72). This fact mostly results from the strand extension when the models are subjected to pulling force, as

shown in Fig. 3.34. Strands are extended in the polyamide-6,6 model more strongly than those in polyamide-4 and polyamide-2 models. The extension is limited in an extended peptide strand, which gives rise to a much lower strain and thus higher pull-out resistance. This result suggests pPA-6,6 to be a better choice for higher stiffness, due to its straightened conformation in crystals.

Conclusion

This study uses all-atom molecular simulations to probe the mechanics of different polyamide materials. Different crystalline morphologies in the models during the simulations are characterized. Two important crystalline forms, α and γ-forms, are identified and compared. The α-form crystals represent the favored state in terms of hydrogen bonding between polyamide peptides, especially for polyamide-6,6. The hydrogen bond density determines the strand pull-out rupture force, with higher hydrogen bond density resulting in a higher stability against strand rupture. The α-form polyamide-6,6 is found to be extremely stiff, featuring the highest pull-out resistance of all models studied. Both, favorite hydrogen bonding and fully extended strands, give rise to the outstanding performance of the α-form polyamide-6,6 crystals. This study demonstrates for the first time the mechanical properties of all-atom crystalline models of different polyamides, which is a first step to analyze polyamide semi-crystalline fiber mechanics for comparison with (silk) protein fibers.

3.2 Myomesin

3.2.1 Fast Folding α-helices as Reversible Strain Absorbers in the Muscle Protein Myomesin

The highly oriented filamentous proteins of muscle constantly experience significant mechanical load during muscle operation. The dimeric protein myomesin has been identified as an important M-band component to support the mechanical integrity of the entire sarcomere. Recent structural studies have revealed long α-helical linkers between the individual Immunoglobulin (Ig) domains in the C-terminal part of myomesin. In this paper we have used single molecule force-spectroscopy in combination with molecular dynamics simulations to characterize the mechanics of the myomesin dimer comprinsing Ig domains 9-13. We find that at forces of approximately 30 pN the α-helical linkers elongate reversibly and allow the molecule to extend to 150% of its folded length. High-resolution measurements directly reveal the equilibrium folding/unfolding kinetics of individual helices. We show that α-helix unfolding mechanically protects the molecule homo-dimerization that would otherwise break before the myomesin molecule can elongate. As fast and reversible molecular springs, myomesin α-helices are essential components for the structural integrity of the M-band. (182)

Introduction

Filamentous modular proteins play a prominent role in the force-bearing structures of the sarcomere (183; 184). The most prominent example is the giant muscle protein titin. For titin, a detailed mechanical hierarchy ranging from entropic stretching of unstructured segments to mechanical kinase activation and unfolding of individual domains has been described (62; 185). While in the sarcomeric I-band titin provides the muscle with its passive tension (186), within the M-band, the 185 kDa protein myomesin (187) as well as other filamentous proteins such as metabolic enzymes and kinase domains form a large network to constitute a well organized compartment that has both structural and metabolic properties (188). Myomesin comprises 13 domains, with the first one (My1) being unique and the others (My2-My13) either of the immunoglobulin type II (Ig) or fibronectin type III fold (189). It is part of a complex network that involves interactions with myosin, titin, obscurin and obscurin-like 1 (190; 191). Through its

N-terminal myosin binding domain (My1) and the ability to form anti-parallel homodimers via an interface residing in its C-terminal domain (My13) (192), myomesin acts as a cross-linker of myosin in the M-band. Ehler et al. have shown that together with the C-terminal part of titin, myomesin is a requirement for the integration of myosin into the sarcomere; they further suggest that myomesin in the M-band, α-actinin in the Z-disk, and titin in between form the basic stabilizing structure of the sarcomere ((193), cf. also (194)). This implicates that myomesin is one of the key factors in maintaining the structural integrity of the M-Band under load.

During normal muscle operation, the M-band and consequently myomesin will constantly be subject to stress and strain imposed by muscle contraction and relaxation (195; 196). Hence its elastic properties are crucial (196). A special isoform of myomesin predominantly expressed in embryonic heart muscle (EH isoform) contains a long unstructured repeat of amino acids. It has been shown that this insert acts as an entropic spring providing significant elasticity to the EH isoform of myomesin (197; 198). However, other isoforms of myomesin lack the specific EH insert. The extensibility of a rigid rod structure consisting of sequential poly Ig and FnIII domains is limited and would therefore directly transmit stress to the dimerization bond of myomesin thus jeopardizing its structural integrity.

Recent structural studies have revealed long α-helical freestanding linkers connecting the C-terminal Ig domains of myomesin (199). Such a linker motif is novel in Ig-repeat proteins. Pinotsis et al. have therefore speculated that the α-helical linker segments of myomesin may provide the necessary elasticity for the C-terminal part of this molecule (199). In this study we use high-resolution single molecule AFM force spectroscopy accompanied by Molecular Dynamics simulations to reveal the unique mechanical characteristics of these linkers for the elastic properties of the myomesin molecule.

Materials and Methods

Preparation of materials as well as all experiments are carried out by our project collaborator, Prof. M. Rief's lab in TU Munich. Details of these experiments have been published elsewhere. (182) Here, only the computational methods are given.

Molecular Dynamics simulations

The X-ray structure of the myomesin C-terminal dimer (PDB code 2R15) (199) containing two immunoglobulin-like domains, My12 and My13, was subjected to molecular dynamics (MD) simulations. We used the WHAT IF (200) package to determine the protonation states of all histidine residues. On solvating the molecular structure with TIP4P (60) water in a rectangular box, we ensured the distance between the protein and the box edge to be 1.5 times of the non-bonded interactions cut-off distance , which is 1 nm. We used an ion concentration of 0.1 mM to mimic the physiological environment, which finally resulted in a simulation system $\sim$7.5$\times$8.0$\times$18.0 nm^3 in size, containing $\sim$135,000 atoms.

We used the GROMACS 3.3.1 package (201) for all the subsequent MD simulations, and the OPLS-AA force field (40) for the protein. Artificial boundary effects were removed by employing periodic boundary conditions. We chose a cut-off of 1 nm for non-bonded interactions, and the Particle-Mesh Ewald method (48) to account for long-range electrostatics. In order to increase the simulation time step, we used LINCS (50) to constrain all bond vibrations. The time step was 0.002 ps. Simulations were performed in the NpT ensemble with a temperature of T = 300 K and a pressure of p = 1 bar in all the simulations. Nosé-Hoover (93; 94) temperature coupling with a coupling time constant τ_T = 0.1 ps, and Berendsen (95) pressure coupling with a coupling time constant of τ_P = 1 ps were employed.

After we relaxed the simulation system by energy minimization using the steepest descent method, we performed 500 ps position restrained simulations to further relax our simulation system, subjecting each heavy atom in the protein to a harmonic potential with a force constant of 1000 kJ$\cdot$mol^{-1} $\cdot$ nm^{-2}. The protein was then fully equilibrated for 20 ns. Energy and coordinates of the simulation systems were collected every 1000 time steps.

The resulting equilibrated simulation system served as starting point for force-probe MD simulations (66) to mimic the atomic force microscopy experiments. It was subjected to an external pulling force. More precisely, we applied the pulling force onto both N-terminal residues of the My12 domains by moving two virtual springs with constant velocity away from the center of the protein. We chose a pulling velocity of 0.5 nm$\cdot$ns^{-1}, and a spring constant of 500 kJ$\cdot$mol^{-1} $\cdot$ nm^{-2}. To accomodate the extending protein, we increased the box dimension along the pulling direction to 30.0 nm, resulting

in a system size of ∼240,000 atoms. Other simulation parameters were the same as listed above.

The resulting Gaussian-smoothed force profile and representative structures are shown in Fig. 3.38. We monitored the initial helix unwinding by determining the protein secondary structure using DSSP (113). For simulations longer than 20 ns, the protein reached the simulation box borders. The helices were fully extended, and the force started to rise to cause immunoglobulin domain unfolding.

Results

In order to focus on my contribution to this project within the thesis, the molecular simulations, only those experimental results most relevant to the theoretical part are presented here. For further details on experiments, we refer to the original publication. (182)

Mechanical stability of wild-type My9-My13 dimers

AFM force spectroscopy was performed on a construct consisting of the 5 C-terminal Ig domains My9-My13 that formed a homo-dimer via the Ig domain My13 (Fig. 3.37A). Typical force-extension traces exhibit saw-tooth like patterns as given in Fig. 3.37B and C. The contour length increase of ΔL = 29.7±0.04 nm (n=965) is in good agreement with the expected unfolding length of an Ig domain containing 86∼90 residues and exhibiting average unfolding forces of 87.9±0.4 pN (n=965) (Fig. 3.37D). The force distribution is broad, likely reflecting the different stabilities of the individual Ig domains within the construct. It is important to note that the adsorption to the cantilever is of non-specific nature, and hence can occur anywhere along the protein chain. We observe up to 6 Ig unfolding events in individual stretching curves (see section Mechanical stability of the My13 dimer interface). At the beginning of the unfolding traces, deviating from the typically observed force-extension behavior due to entropic polymer elasticity (8) (black lines in force-extension traces are worm-like chain fits), regions of apparently constant force can be detected (see arrows in Fig. 3.37B). These plateau regions occur within a relatively narrow force range around 29.6±0.6 pN (n=95, dashed line and black histogram in Fig. 3.37D). In contrast to the typical mechanical unfolding of protein domains far from equilibrium, this force plateau reappears between peaks if the force drops below the plateau value after the unfolding of an Ig domain (2nd arrow in Fig. 3.37B). To test the reversibility of the plateau behavior, we performed a series

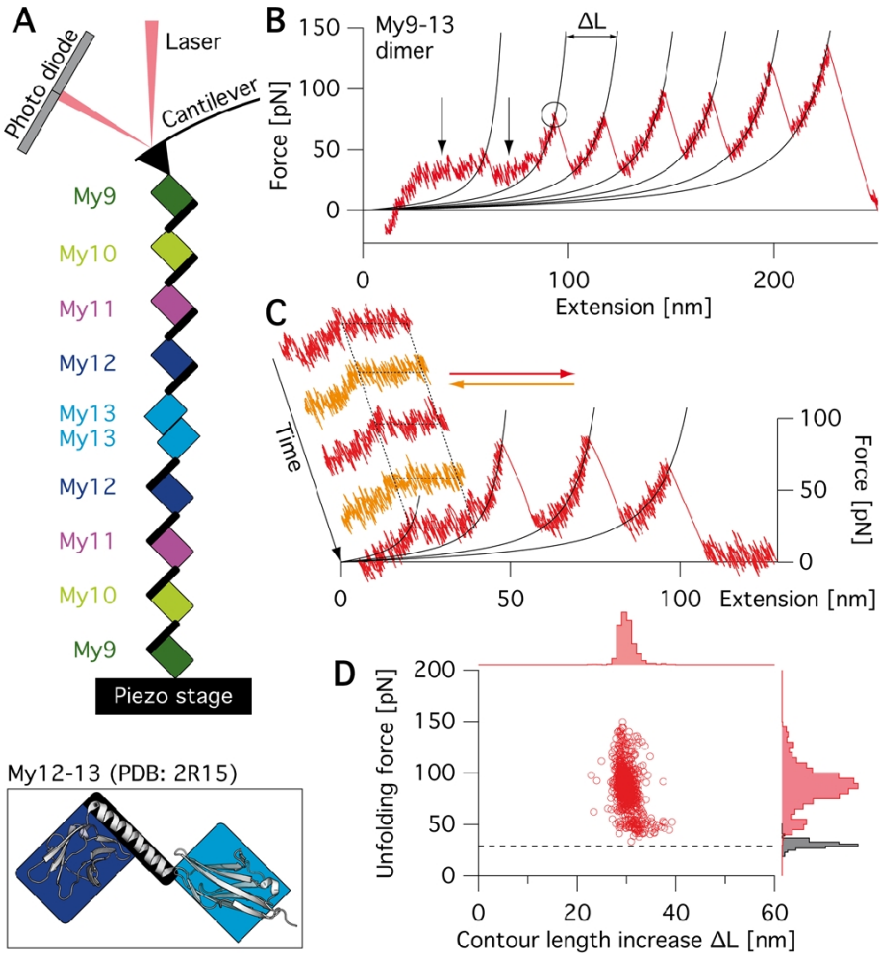

Figure 3.37: (A) Schematic representation of a My9-My13 dimer in the AFM experimental setup (not to scale). The colored squares and the black bars represent the Ig domains My9 to My13 and their linker helices, respectively (see bottom). (B) Exemplary force-extension trace of My9-My13 unfolding. The circle marks a single unfolding event, black traces represent worm-like chain fits providing the contour length increases $\Delta L$of a single unfolding. The arrows indicate plateau force regions. (C) Force-extension trace containing stretch (red) and relax (orange) cycles to test the reversibility of the plateau. In the final stretching cycles two unfoldings can be observed before the sample detaches. (D) Scatter plot of unfolding forces and corresponding contour length increases with respective distributions (red). The black histogram gives the plateau force distribution and the dashed line its mean value.

of subsequent stretching and relaxation cycles on one individual myomesin molecule. The plateau exhibits no observable hysteresis between stretching and relaxation cycles up to the highest velocity measured of 1 μm/s (Fig. 3.37C). From these observations we conclude that the elongation/shortening transitions underlying this force plateau occur fast compared to the experimental time scale.

The α-helix linker is an extensible element

To understand the nature of the plateau in the force-extension traces, we performed force-probe molecular dynamics (MD) simulations. A force-extension curve of a myomesin My12-My13 dimer is shown in Fig. 3.38 (red curve). Above extensions of 4 nm a plateau region appears in which the molecule extends at constant forces. Beyond 13 nm the force again continues to rise. A series of snapshots of the associated molecular conformations reveals the nature of the force plateau: While all four Ig domains remain intact, the α-helical linkers start to unfold in the plateau region. In our simulations α-helix unfolding starts at those segments of the helix that are solvent exposed (C-terminal part; cf. Fig 3.38, top). Part of this region has already unfolded when the plateau force is reached. The other segments of the α-helix forming a hydrophobic interface with the Ig domains are more stable and unfold later in the plateau. It is important to note that the timescales are much smaller in the simulations than in the experiments and hence, forces from these two methods cannot be directly compared. However, the mechanical hierarchy within the molecule is likely conserved (45; 70) as also suggested by the qualitative agreement between the experimental and calculated force profiles, which both feature a plateau followed by a steep increase in force.

Discussion

Again, the discussion here mainly covers results from MD simulations. Details can be found in the joint publication with the experimental partners. (182)

For its role as a cross-linker of myosin filaments, the elastic properties of myomesin as well as the stability of its dimerization interface are important. In the absence of detailed structural insight, early mechanical measurements of myomesin have naturally focused on the stability of the fold of Ig and Fn domains within the myomesin rod (197; 198). In those earlier studies, in addition to the domain unfolding forces, the entropic elasticity of the EH segment, a putatively unstructured element containing ca. 100 amino acid residues, was investigated. However this segment only occurs in an isoform

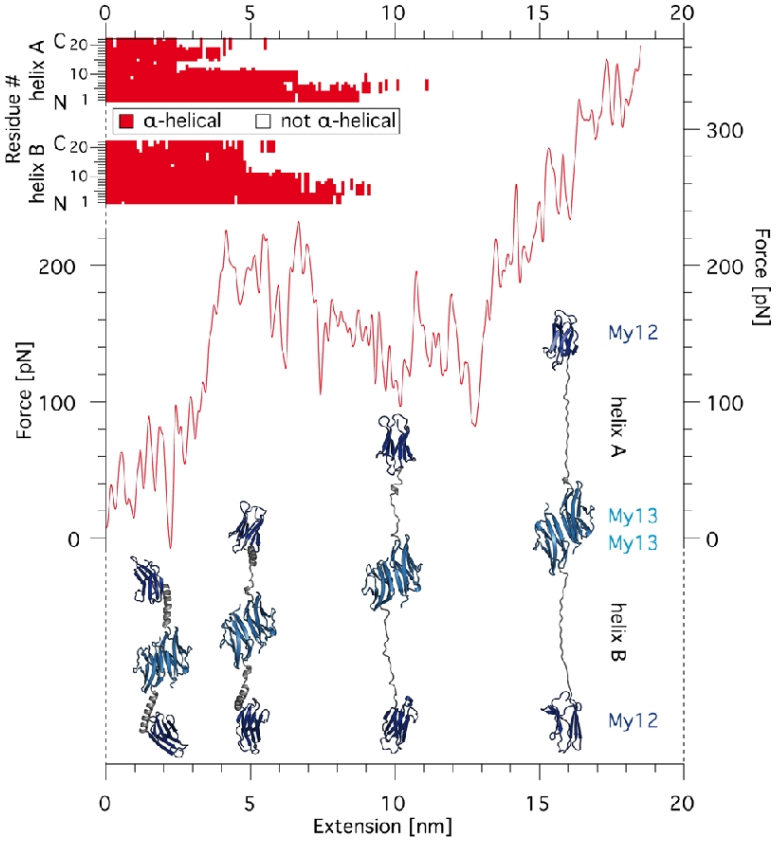

Figure 3.38: Force-probe molecular dynamics simulation of My12-My13 dimer unfolding. The red force-extension trace clearly exhibits a plateau. Corresponding structural snapshots at 2 nm, 5 nm, 10 nm, and 16 nm are given below in cartoon representation and show that the plateau corresponds to α-helix unfolding. The individual conformational state of both helices is mapped for each residue against extension (top).

predominantly expressed in embryonic heart muscle. The sources of elasticity for the most commonly expressed myomesin isoforms lacking the unstructured segment have so far remained unclear. The recent discovery of long α-helical linker segments as well as the detailed structure of the dimerization complex has raised the possibility of a novel elastic element within myomesin (199).

Generally, an elastic mechanism based on non-equilibrium processes such as Ig domain unfolding, exhibits broad force distributions and can therefore not supply sharply defined force values. Equilibrium unfolding of the α-helical linkers combines two essential features for a reliable elastic mechanism: the possibility for considerable elongation

at force values far below dimer dissociation at the level of 30 pN while still providing stability and rigidity at forces below 20 pN. At first sight, the possible elongation upon α-helix unfolding may appear small. Nevertheless, full extension of a total of 8 helices in a dimer can elongate the molecule by approximately 50 nm, which corresponds to 50% of the total folded length of the myomesin dimer. This will provide the myomesin with enough adaptability to react to misalignment or changes in spacing of thick filaments (196). One of the most important roles we assign to unfolding of the α-helical segments is the protection of the homo-dimerization against deleterious loads. We can estimate a lifetime for the dimeric interface of 12 s at the plateau force of 30 pN. (182) Hence, even under those extreme stretching conditions the molecule dimerization can survive a significant time.

Intriguingly, according to current models of the M-line structure, myomesin is oriented parallel to the sarcomere filaments only with its C-terminal part (My7-My13) (196). This C-terminal portion of the myomesin dimer then spans the region between the M4 and M4' line being a constituent of the M-filament as seen in EM micrographs (202; 203). Hence, in this region the whole mechanical stress will act on the dimer interface. Towards the N-terminus, the myomesin will be oriented rather perpendicular to the sarcomere axis and the axial mechanical stress will then be likely distributed over many more supporting structures. Such binding partners interacting with myomesin in the N-terminal region include myosin, titin, and obscurin (190; 191).

In summary, we show that the unique α-helical linkers found between the C-terminal Ig domains in myomesin have mechanical properties that make them ideally suited to act as strain absorbers in the M-band. The fast and reversible two-state-folding kinetics protect the stability of the myomesin dimer under load up to strains of 150%. The mechanical properties of myomesin elasticity presented here form an important building block for the emerging mechanical and structural understanding of the M-band.

3.2.2 Myomesin is an Elastic Structural Maintainer of the Sarcomere

The M-band, consisting of a complex assembly of fibrils, has long been speculated is one of the most important structural maintainers of the sarcomere, the force generating unit in striated muscle. Although much efforts has been devoted to studying the M-band, its organization, and thus mechanism are still under debate. This study employs Molecular Dynamics simulations to study the most crucial structural molecules, myomesin, of the M-band. By quantifying the mechanical robustness of immunoglobulin domains and the flexible domain-connecting helices, we are able to show that myomesin acts both as a stress absorber and structural restoring element in the M-band during a force generating cycle.

Introduction

The M-band is located in the middle of the muscle force generating unit, the sarcomere. It consists of molecular fibrils shown as dark lines in microscopic images. These fibrils interconnect tails of myosin thick fibrils and anchor titin C-termini in the M-band. (196) As an integrating molecular network of myosin tails, the M-band acts as a safe guard of the sarcomere. (195) Although a reconstructed three-dimensional structure of the M-band is available nowadays from single-particle analysis, (204) the molecular details of the M-band protein network are still missing. Thus, how the M-band molecular network balances mechanical stress in a force generating cycle is unknown.

Three most important molecules have been identified up-to-date which are responsible for the M-band lines, namely myomesin, M-protein and myomesin-3. (205) These three molecules are genetically related. As the most important molecule, being found in all types of striated muscles (187), myomesin is the most promising candidate for deciphering the secrets of the M-band mechanics.

Myomesin consists of 13 domains, as shown in Fig. 3.39. The first domain of myomesin, my1, has a special molecular structure and binds to a myosin tail. The other domains of myomesin are either immunoglobulin (Ig) domains or fibronection type-III (FNIII) domains. The 13th domain, my13, is able to form a homodimer, and thereby enabling two myomesin molecules to expand from the M4 to M4' lines of the M-band. (196) One iso-form of myomesin, identified in early development of heart and termed as EH-myomesin, has an unstructured peptide segment between the 6th and

7th domains, my6 and my7. (206) This unstructured molecular segment in this iso-form of myomesin features a similar function as the PEVK segment in titin, providing the molecule with elasticity. (197; 198; 207)

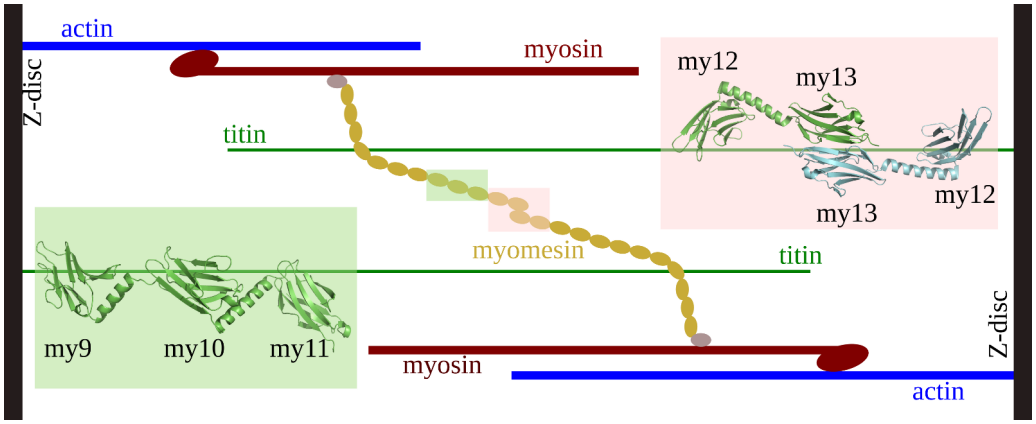

Figure 3.39: Schematic figure of the molecular connections established by myomesin. Myomesin domains are depicted as elliptic beads. Main components of the sarcomere are labeled in the Figure. Two segments of detailed molecular structures are shown in cartoon representation, with corresponding positions in the myomesin molecule highlighted with relevant colors.

The interactions between myomesin and other muscle molecules are the key to the integration of other fibrils in the M-band of sarcomere. It is already known that myomesin's 4th domain, my4, interacts with titin and helps titin to anchor its C-terminus in the M-band.(192) More importantly, the connection between two anti-parallel myosin tails via two myomesin molecules not only serves as the structural organization in M-band, but also as force transduction bridges to balance mechanical force in force generating cycles of the sarcomere. (195) Although being critically important to the M-band, the mechanical function of myomesin is still not very clear because of the lack of molecular structures and adequate experimental studies.

Recently, the molecular structure of a homodimer of two C-terminal domains of myomesin, my12 and my13, has been reported. (199) This molecular structure contains two peptide chains dimerized by the two my13 domains, as depicted in Fig. 3.39 (cartoon representation). More interestingly, there is a long α-helix between my12 and my13. Studies both by experiments and Molecular Dynamics (MD) simulations showed that this helix functions as strain absorber and elongation units upon force applications on

the molecule (see Section 3.2.1). (182) A mechanical plateau force of 30 pN is needed to unfold this helix in Atomic Force Microscopy (AFM) experiment, which results in an elongation of 150% of its initial helical length to an extended peptide (see Chapter 3, Section 3.2.1). In order to fully explore the mechanical properties of myomesin, here additional MD simulations are performed to studies both reported and unreported molecular segments of myomesin, my9-my11 and my12-my13 dimer. The my9-my11 fragment is currently unpublished, but was provided to us by experimental partners (Dr. M. Wilmanns). Force-probe MD (FPMD) simulations are carried out to apply force to 5 Ig domains, my9-13, as well as the dimer formed by two m13 domains. Force-clamp MD (FCMD) simulations are employed to quench selected helices to a low constant force, namely the helix between my9 and my10 (helix910), the helix between my10 and my11 (helix1011), and the helix between my12 and my13 (helix1213).

Methods

Structural equilibration

The molecular structure of a segment of the myomesin molecule, my9-13, is submitted to Molecular Dynamics (MD) simulations for structural equilibration. We use the WHAT IF (200) package to determine the protonation states of all histidine residues. On solvating the molecular structure with TIP4P (60) water in a rectangular box, we ensure the distance between the protein and the box edge to be 1.5 times of the non-bonded interactions cut-off distance of 1 nm. We use an ion concentration of 0.1 mM to mimic the physiological environment in muscle. We choose the GROMACS 4.0.7 package (201) for all subsequent MD simulations, the OPLS-AA force field (40) for the protein and TIP4P (60) to describe water. Artificial boundary effects are removed by employing periodic boundary conditions. We use the Particle-Mesh Ewald method (91) to account for long-range electrostatics. In order to increase the simulation time step, we use LINCS (50) to constrain all bond vibrations. The time step is 0.002 ps. All simulations are performed in the NpT ensemble with a temperature of T = 300 K and a pressure of p = 1 bar. Nosé-Hoover (93; 94) temperature coupling with a coupling time constant $\tau_T = 0.1$ ps, and Berendsen (95) pressure coupling with a coupling time constant of $\tau_P = 1$ ps are employed, respectively.

The protein structure is energetically minimized and fully equilibrated for 20 ns. The resulting equilibrated molecular structure is cut into different parts, namely the

dimer of two my13 domains, my9-12 domains, and connecting helices. Each part of the molecular structure is then solvated into simulation boxes with the same conditions as listed above for subsequent MD simulations. More precisely, systems of the my13 dimer and the of my9-12 domains are set up and equilibrated for force-probe MD simulations, and the helix1213 for force-clamp MD simulations.

Force-probe MD (FPMD) simulations

The dimer and Ig domains of myomesin are subjected to FPMD simulations (66) to mimic their unfolding process in the AFM experiments (see Section 3.2.1). Pulling force is applied onto the molecular structures by moving two virtual springs with constant velocity away from the force application points which are the two termini of each Ig domain or two different peptide chains of the dimer, respectively. A pulling velocity of 0.5 nm·ns^{-1}, and a spring constant of 500 kJ·mol^{-1} · nm^{-2} are chosen for all FPMD simulations. In order to accommodate the extending protein during unfolding, the box dimensions along the pulling direction are increased to ∼25.0 nm, resulting in system sizes of ∼200,000 atoms. The pulling force and trajectories are constantly recorded during the simulations until the domains are fully unfolded. Other simulation parameters are the same as for the structural equilibration listed above.

Force-clamp MD (FCMD) simulations

Helices are subjected to FCMD simulations. Different constant forces are used to hold the helices ranging from a high value of 100 pN to a low value of 6 pN. Secondary structures of the helices are monitored during the simulations using DSSP (113). FCMD simulation parameters are the same with those in equilibration simulation.

Results and Discussion

Mechanical robustness of Ig domain my9-13

All five Ig domains, my9-13, are independently subjected to pulling forces. Using a pulling speed of 0.5 nm·ns^{-1} and a spring constant of 500 kJ·mol^{-1} · nm^{-2}. The resulting loading rate is orders of magnitude higher than experiments, which results in higher rupture forces in simulations than AFM experiments as predicted by Bell's model. (67) Five independent simulations are performed to rupture each Ig domain, namely my9-13. The average rupture forces of each domain range from 443∼716 pN, as shown in Fig. 3.40. The my9-13 Ig domains thus show a similar robustness as the Ig-domains in titin, such as I1 and I27, studied by similar simulations. (43; 44) As expected

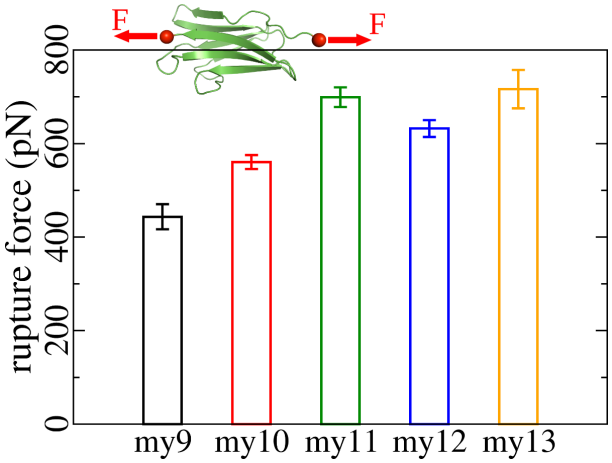

Figure 3.40: Rupture forces of the five Ig domains in myomesin. Each rupture force is the average of five independent FPMD simulations. Top snapshot shows my10, with red spheres indicating the force application points and the force arrows indicating the pulling directions in the FPMD simulations.

due to the higher loading rate, all rupture forces observed in MD simulations are higher than corresponding those in AFM experiments, see also Section 3.2.1. The key patterns of the unfolding processes and the critical features of the molecular mechanics obtained by two methods, AFM and MD simulation, are likely to conserved, as being confirmed by former studies. (45; 70) Interestingly, the mechanical stability of the Ig domains differ by nearly 300 pN, in spite of their highly similar topologies. The sequence variations causing these differences remains to be elucidated.

Stability of my13 homodimer

Force is applied to N-termini of two my13 domains of the dimer in FPMD simulations as depicted in Fig. 3.41. The N-terminal β-strands of the two my13 domains in the dimer are establishing the interaction between the two domains. Force induces the detachment of the N-terminal β-strands from each other. This dissociation of the dimer does not significantly affect the molecular structures of the two my13 domains. They instead remain in their folded states after rupture. We note that the FPMD simulations of my13 homodimer use the same loading rate as those of other Ig domains.

The detachment force of the myomesin dimer obtained from MD simulations is 818$\pm$51 pN, which is higher than the rupturing forces of all Ig domains investigated above. A higher force needed to detach the my13 homodimer is confirmed by AFM

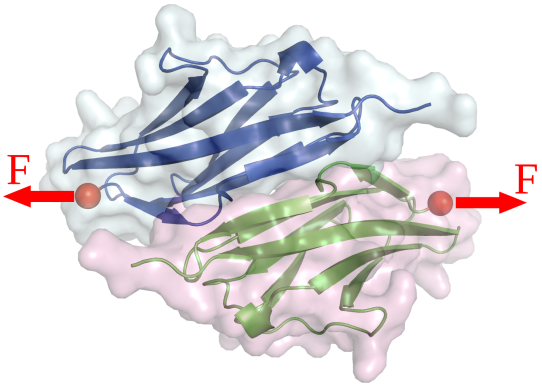

Figure 3.41: Molecular structure of a myomesin dimer and pulling setup in FPMD simulations. The two peptide chains are shown in cartoon and transparent surface representation and colored in blue and red, respectively. Pulling forces are shown as red arrows, and force application points as red spheres.

experiments (see Section 3.2.1) (182), which suggests that the linkage between two myomesin molecules can survive over other Ig domains when over-stretching myomesin fibrils. The elongation of domain-connecting helices and the unfolding of Ig domains progresses under mechanical load are ensured by the integrity of the homodimer. As a consequence, myomesin can elongate to an extreme extra yet maintains a bridge between two anti-parallel myosin fibrils.

Helices in equilibrium

Myomesin is the only muscle molecule that possesses the unique structural architecture of helix-connected Ig domains, which is likely to present a molecular model for elasticity. (182) Former studies (Section 3.2.1) have already indicated that these α-helices can serve as strain absorbers in the molecule that provide viscous extensibility. Further investigations are needed to explore the mechanical properties of these helices, such as their refolding abilities and their contribution to the molecular adaptability in sarcomere.

To this end, each of the three helices, with the adjacent Ig domains being removed, are first equilibrated separately for a simulation time ranging from 310 ns to 450 ns, as shown in Fig. 3.42. None of the three helices is fully stable in solvent. As depicted in Fig. 3.42, all three helices lose helix secondary structures gradually with the increase of simulation time. In the end of the simulations, helix1011 and helix1213 have lost nearly all of their helical secondary structure. When compared to helices in the full

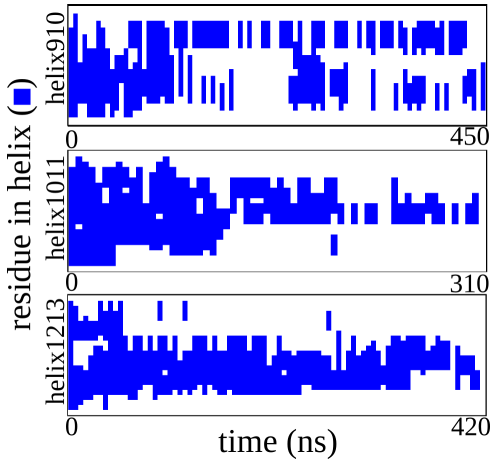

Figure 3.42: Secondary structure plots of three helices in myomesin in equilibrium. Vertical axis is residue number (not shown in the Figure). Simulation times are indicated at the end of horizontal axis of each plot. Blue color blocks indicate α-helical configuration of the corresponding residues in the peptide sequence.

myomesin structure, the simulation environment of equilibration is artificial because of the missing link to and interactions with the Ig domains. The destabilization and unfolding of the α-helices in solvent, similar to other helices of similar sizes, can be explained by the change in the environment, from a mainly hydrophobic Ig interface to the polar water solvent.

Interestingly, all these helices, especially helix910, show refolding events of helical turn on a short time scale during the simulations. Residues in these helices seem to have a high tendency of α-helix refolding, as shown by secondary structure plots in Fig. 3.42. This fast refolding is likely to be a key function of myomesin helices, considering the fast force generating circles in muscle operation. In order to probe this feature of myomesin, FCMD simulation of these helices are carried out (see next section).

Helices under constant pulling force

We further perform simulations to mimic the situation of myomesin α-helices under constant pulling force and assess their mechanics. We choose helix1213 as a representative helix for further investigations because of its complete unfolding in the end of the equilibrium MD simulation (Fig. 3.42). Using constant force applied at each end of helix1213, the freedom of rotation and bending of helix1213 is restrained. This simulation setup also features the condition of how a helix in myomesin experiences external

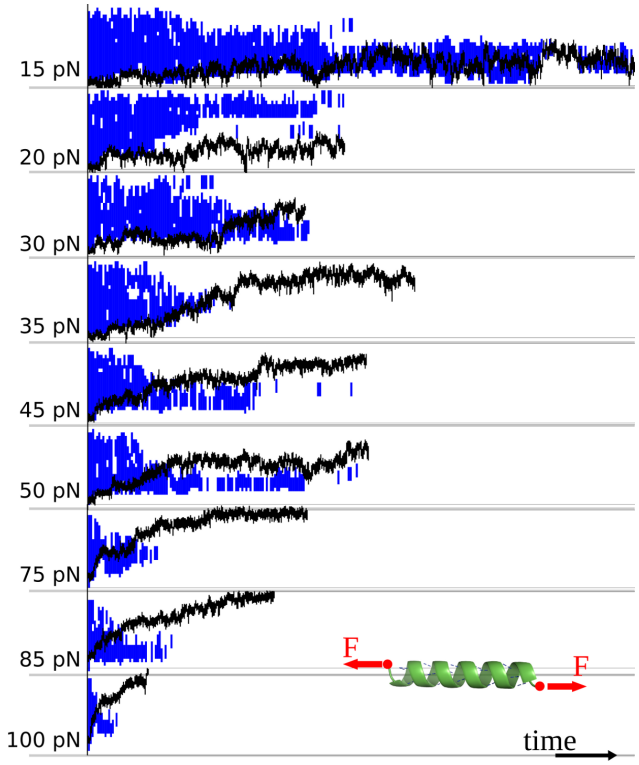

Figure 3.43: Secondary structure plot (blue) and end-to-end distance (black) of helix1213 under different constant forces. Secondary plots are obtained using DSSP, and shown along helix sequence from bottom (N-terminal) to top (C-terminal) in each plot, with blue area representing α-helical conformation. The inset shows the pulling setup in all simulations. Black curves in each data set are the end-to-end distances recorded in the simulation. Scales of the vertical axis are not shown, which are residue numbers for DSSP plots and nm for distance curves.

force in the sarcomere. A pulling setup is schematically shown as inset in Fig. 3.43. Different constant forces ranging from 15 pN up to 100 pN are used in the FPMD simulations (Fig. 3.43). Simulation times are adopted to the pulling forces. Simulations are stopped when the whole helix has unfolded. The secondary structure of helix1213 and its end-to-end distance are monitored during the simulations, as shown in Fig. 3.43

As expected, helix1213 loses its helicity quickly under high constant forces of 75$\sim$100 pN. Correspondingly, its end-to-end distance increases extremely fast with the loss of helicity. AFM experiments of complete myomesin molecules indicated that the resisting unfolding force of myomesin helices is $\sim$30 pN (see Section 3.2.1). (182) The isolated

helix1213 dose show a longer life time in this force range, though only on a nanosecond scale, presumably due to the lack of Ig domain interations. When the pulling force further decreases to 15 pN, helix1213 keeps its helicity even longer and shows obvious helical turn refolding events as indicated by the secondary structure shown in Fig. 3.43. This result indicates that helices in myomesin can partially refold on a very short time scale of nanoseconds, much shorter than the refolding time of seconds of an Ig domain. The reversibility of helix refolding under constant force is expected to apply to other helices in myomesin, especially helix910 which showed refolding within nanoseconds already in the absence of force (Fig 3.42). This result also confirms that the decrease in bending flexibility of myomesin helices by applying of constant holding force can enhance their helix propensity, similar to their incorporation in the myomesin molecule.

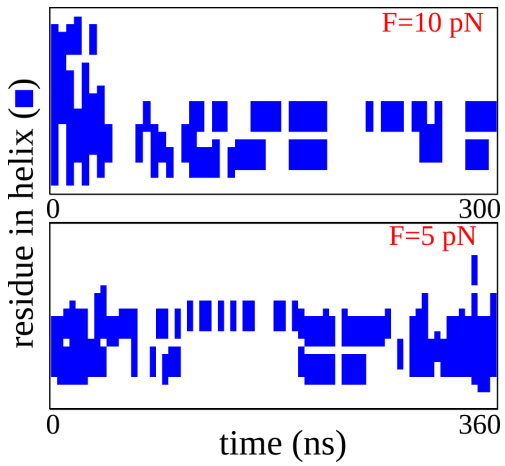

Figure 3.44: Secondary structure plot of helix1213 under constant forces of 10 and 5 pN in FPMD simulations. FPMD Simulation using 5 pN (bottom plot) continues at 300 ns of the FPMD simulation using 10 pN (top plot).

A simulation with low constant force of 10 pN with a simulation time of 300 ns confirms the reversible helix unfolding described above (Fig. 3.44). Again, the helix partly loses its helicity during the simulation. But pronounced helix refolding events on this submicrosecond scale. The constant pulling force is then switched to 5 pN at 300 ns and maintained till the end of the FPMD simulation.

As being shown in Fig. 3.44, helix1213 under a pulling force of only 5 pN continues to randomly refold or lose its turns. Remarkably, even longer sections of the helix also

refold, as indicated by secondary structure plot in Fig. 3.44. Shortening of the end-to-end distance of helix1213 again accompanies the partial refolding of the helix (Fig. 3.45). Helix refolding is expected to be more pronounced in the presence of Ig domains and in experiments at larger timescales, given the strong hydrophobic interactions between the helix and the my12 Ig domain. Indeed, AFM experiments give unfolding and refolding traces with low hysteresis for myomesin with intact Ig domains.

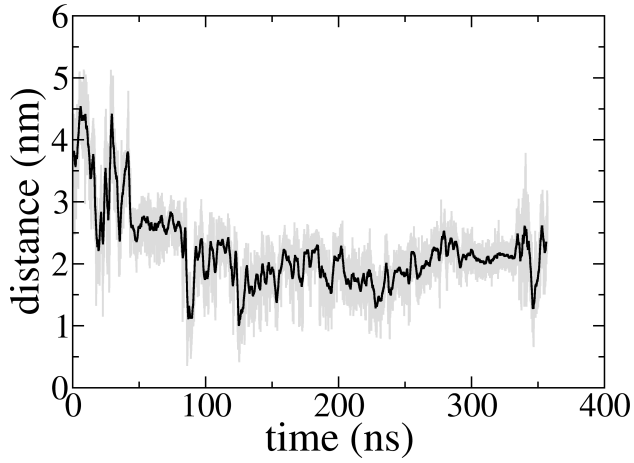

Figure 3.45: End-to-end distance of helix1213 after switching of constant holding force from 10 to 5 pN. Raw data is shown in gray curve with average shown in black curve.

Except from its iso-form, EH-myomesin, myomesin does not contain a disordered elastic module equivalent to the PEVK segment in titin. However, the helices in myomesin here have been demonstrated to be extensible elements that provide elasticity and serve as shock absorbers (see Section 3.2.1). These helices are thus able to serve as elastic modules that a muscle molecule commonly features. Unfolding of Ig domains are rare physiological events in muscle. Only extreme force lead to their unfolding, with refolding only occurring on the second timescale. This represents a risk for the integrity of the sarcomere given the fact that the main function of myomesin is to anchor the titin filament and to interconnect anti-parallel myosin molecules. By integrating helices into its molecular structure, myomesin is able to enormously elongate which lowers the risk of unfolding its Ig-domains.

In an activated sarcomere, two anti-parallel myosin molecules consume energy to move in opposite directions thereby exerting mechanical force onto the linking molecules,

mostly myomesin, in the M-band. This cause an elongation of two dimerized myomesin molecules in the sarcomere axis direction due to the displacement of two myosin molecules. Because myomesin molecules also interact with other fibrils in the M-band, such as titin, myomesins are able to transduce mechanical force in the M-band network and to balance stress distribution. Most importantly, fast helical refolding is likely to play a crucial role in the relaxation period of the force generating cycle in the sarcomere. The involved helix shortening produces a counter force to restore the unactivated states of myosin filaments. The time scale of myomesin helix refolding is short enough, as shown by our simulation results, to restore the sarcomere structure for the next force generating cycle. The results confirm that myomesin is one of the most important M-band, and thus the sarcomere safeguards.

4

Conclusions

This thesis focuses on two types of force-withstanding proteins, namely silk and the muscle protein myomesin, with the aim to understand the determinants of their mechanical properties. Two simulation methods, MD simulations and FEM, are used, and results thereof validated by experimental data of various sources. New insights into self-assembly, rupture mechanics, force bearing units and refolding were gained, which will be outlined in the next paragraphs.

Proteins are materials tailored for bearing force

Proteins can achieve a toughness outperforming other materials like steel or Kevlar; they can also show a high elasticity and extensibility which allow them to dynamically and reversibly adapt to their environment. The determinants of protein mechanics lie in the assembly of the protein structural building blocks.

Silk

Subunits of silk proteins, crystals and amorphous peptides, enable a silk fiber to carry a mechanical load with outstanding robustness and tremendous elasticity, which together result in an immense toughness.

Crystals in the silk protein are extremely stiff. These results are presented in Chapter 3, Section 3.1.1. A force of ∼4 nN is needed to rupture one peptide strand from a crystal along the axial direction. The elastic moduli measured during the MD simulations are also extremely high, reaching ∼86 GPa. The peptide arrangements

in silk protein crystals affect the mechanical properties of the molecular structure. An anti-parallel arrangement of β-strand is more mechanically robust than a parallel counterpart, with a higher strand rupture force and a narrower distribution of internal force as indicated by FDA.

The β-strand lengths found in silk protein crystals are about eight residues, which also applies to other force bearing protein structures as investigated in Chapter 3, Section 3.1.2. As the evolution of protein materials in nature suggests, the mechanical performance for a given material volume is maximized for 8-residue long sheets. Higher robustness, in term of β-strand rupture force, can be achieved by longer strand lengths, but β-strands of 8-residue length show the best compromise between mechanical properties and material usage, especially given the fact that most of the mechanical load is taken up by the first eight residues linked to the force application point.

Amorphous peptides in silk fibers are softer than crystals and can undergo a significant elongation under force. The topology of amorphous peptides is likely to involve random entanglements with various interactions, which is captured by the composite model of the crystal-amorphous-crystal structure in Chapter 3, Section 3.1.3. Typically, the elongation of amorphous peptides features friction and reversible rupture of random interactions between peptides. These resist pulling forces and consume substantial mechanical energy.

The function of subunits in silk protein, namely crystals and amorphous peptides, are obvious. Amorphous peptides are mechanical buffers used to consume mechanical energy, while crystals are structural enhancers which prevent fiber disintegration. The outstanding mechanical properties of silk fibers come from the synergy between the two subunits of silk. By assembling two type of subunits with different mechanical properties, a silk fiber is optimized for both load carrying and load-induced stretching.

The assembly mechanism of silk in vivo is still mysterious. By controlling the percentage of crystals in silk protein, different types of silk are produced by spiders. Although natural silk fibers include a volume percentage of 10%$\sim$25% of protein crystals which is assumed to be randomly distributed in the amorphous phase, this study shows that a better structural fiber hierarchy can be achieved. This is discussed in Chapter 3, Sections 3.1.3 and 3.1.4. A silk fiber model with a serial arrangement of crystals in the fiber shows a higher toughness than other fiber models either with parallel or random arrangements of crystals at any percentage of crystallinity. Our fiber model

also suggests that in the serial arrangements, 40% of crystallinity yields the best fiber performance. Fiber performance is also sensitive to the sizes of crystals. Crystals with longer strand length and smaller cross-section area can better reinforce the fiber hierarchy. All these results are critical for material engineering and also shed light on the refinement of semi-crystalline industrial materials like polyamides.

Polyamide materials such as nylon and polyamide-6,6 are similar to silk on the atomistic scale, as they form highly similar hydrogen bonding patterns. Crystals of different polyamide materials are modeled and investigated in Chapter 3, Section 3.1.6. Hydrogen bond densities are found to differ for different polyamides crystals, which results in different crystallinities, strand rupture forces, and elastic moduli. The α-form polyamide-6,6 crystal is found to show the highest ordered in its atomic structure. It also shows the highest rupture force and elastic modulus among all polyamide crystal models. The results explain the outstanding mechanical properties of polyamide-6,6.

Myomesin

Myomesin is the most important structural molecule in the M-band of the sarcomere. While largely differing from silk in term of structure and function, myomesin also consists of rigid Ig domains rich in β-sheets which are oriented along the pulled direction. Thus, both proteins have a primary determinant of mechanical resilience in common. It is studied in Chapter 3, Sections 3.2.1 and 3.2.2.

The construct of rigid Ig domain with soft helical connections render myomesin a new model for molecular elasticity. The helices in myomesin behave as mechanical absorbers when the molecule is subjected to external load. The unfolding of these helices feature large elongations with a constant resistance against the pulling force. This effect provides myomesin with viscous elasticity to adapt to the mechanical environment in the M-band in the middle of sarcomere. These α-helices compensate for the lacking of an unstructured and entropically elastic peptide segment in myomesin. The helices investigated also show a remarkable ability of fast refolding with a refolding time scale of nanoseconds. The fast refoldability is crucial for myomesin considering the short contraction-and-relaxation cycles of force generating of sarcomere.

The dimer formed by two myomesin via my13 domains is found to be stronger than other Ig domains in the molecule. This result suggests the key contribution of myomesin to the integrity of fibrils in the M-band. Two dimerized myomesin molecules connect

two anti-parallel myosin fibrils which constantly apply force to the myomesin dimer in opposite directions during the activation of the sarcomere. Elongation of myomesin and rare unfolding events of Ig domains can take up considerable mechanical work. As the last and strongest safeguard, the dimer of two myomesin molecules establishes the connection and thus the integrity of the molecular network in the M-band, which serves as a foundation for structure restoration of the M-band with the refolding and shortening of myomesin helices occurring within the relaxation period of sarcomere.

Protein aggregation induced by a surface

Controlling protein folding has remained a highly challenging problem. Although it is well accepted that a protein spatial structure is determined by its amino acid sequence, amyloid-type β-sheets form from various peptides under certain conditions. Both chemical or mechanical conditions such as ion concentration and shearing flow strength can modify protein folding pathways. Controlling protein secondary structure formation on a single molecule level opens the door for understanding protein material assembly mechanisms.

As presented in Section 3.1.5, silk proteins form secondary structures on a high energy metal surface, namely stacks of β-sheets. Apparently, surfaces can create an environment for controlling protein folding. MD simulation results confirm that β-sheets can explain the force curves obtained from desorbing the protein from the metal surface. A parallel arrangement of strands is preferred to form on the surface. It absorb more mechanical energy upon detachment, and thus possesses a higher mechanical stability. The studies by both AFM experiments and MD simulations open a road for guiding protein self-assembly into new materials.

Combination of MD simulations and FEM

Combining of different methods for bridging length on time scales has been proven to be very effective, a successful example of which are hybrid QM/MM methods used in studying molecular transitions. Combining MD simulations with FEM modeling successfully pioneered for studying protein mechanics on multiple scale has been demonstrated in this thesis (Chapter 3, Sections 3.1.3 and 3.1.4), which can be easily applied

to other similar problems.

MD simulations are fairly accurate yet limited to systems of nanometer size. In order to unravel the mechanical relationship between the atomic structure and the overall performance of materials using a bottom-up approach, a coarser description that yet maintains the atomic characteristics of the investigated material is required. Continuum mechanics based on finite elements is well established for structural mechanics studies and can be applied to all scale. The efficiency of FEM largely outperforms MD simulations and allows macroscale modeling. Accurate parameters needed for the FEM modeling can be obtained from MD simulations, allowing the bridging of the two methods. Silk fiber modeling using both MD simulations and FEM conserves the accuracy of MD simulations and benefits from the advantages of FEM. More successful studies combining these two methods can be expected.

Outlook

This dissertation probes the mechanics of protein materials. All-atom simulations are applied to model and quantify molecular mechanics. Continuum modeling method is used to further increase the simulation scale. Our results can serve as guidelines for artificial material design and provide a basis for further studies.

Our approach can be readily improved by refining the models. Bridging of MD simulations and FEM entails open questions, such as the modeling of the amorphous peptide chain topology in FEM and their mechanical properties. The FEM model used in this study is in elastic regime of material deformation, while it shows visco-plastic effects in real fiber. As for the study of myomesin molecular mechanics, a larger all-atom molecular segment structure should be studied for better mechanical comparisons. Improvements in this direction will make our methods predictive enough for systematic comparisons to and interpretations of experiments. All these questions require more research strength and time. This thesis is a starting point for more exciting studies on protein mechanics.

References

[1] Marie-Louise Bang, Thomas Centner, Friderike Fornoff, Adam J. Geach, Michael Gotthardt, Mark McNabb, Christian C. Witt, Dietmar Labeit, Carol C. Gregorio, Henk Gr anzier, and Siegfried Labeit. **The Complete Gene Sequence of Titin, Expression of an Unusual approx700-kDa Titin Isoform, and Its Interaction With Obscurin Identify a Novel Z-Line to I-Band Linking System**. *Circ Res*, **89**(11):1065–1072, 2001.

[2] Markus Heim, David Keerl, and Thomas Scheibel. **Spider Silk: From Soluble Protein to Extraordinary Fiber**. *Angew Chem Int Ed Engl*, **48**(20):3584–3596, 2009.

[3] X. Hu, K. Vasanthavada, K. Kohler, S. McNary, A. M. Moore, and C. A. Vierra. **Molecular mechanisms of spider silk.** *Cell Mol Life Sci*, **63**(17):1986–1999, 2006.

[4] Fiorenzo G. Omenetto and David L. Kaplan. **New Opportunities for an Ancient Material**. *Science*, **329**(5991):528–531, 2010.

[5] F. Vollrath. **Strength and structure of spiders' silks**. *J Biotechnol*, **74**(2):67–83, 2000.

[6] S. Winkler. **Molecular biology of spider silk**. *J Biotechnol*, **74**(2):85–93, 2000.

[7] Chia-Lin L. Chyan, Fan-Chi C. Lin, Haibo Peng, Jian-Min M. Yuan, Chung-Hung H. Chang, Sheng-Hsien H. Lin, and Guoliang Yang. **Reversible mechanical unfolding of single ubiquitin molecules.** *Biophys J*, **87**(6):3995–4006, 2004.

[8] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. **Reversible unfolding of individual titin immunoglobulin domains by AFM.** *Science (New York, N.Y.)*, **276**(5315):1109–1112, 1997.

[9] M. F. Perutz. **Structure of hemoglobin.** *Brookhaven Symp Biol*, **13**:165–183, 1960.

[10] S. Woutersen. **Hydrogen-bond lifetime measured by time-resolved 2D-IR spectroscopy: N-methylacetamide in methanol**. *Chemical Physics*, **266**(2-3):137–147, 2001.

[11] Alenka Luzar and David Chandler. **Hydrogen-bond kinetics in liquid water**. *Nature*, **379**(6560):55–57, 1996.

[12] M. Grandbois, M. Beyer, M. Rief, H. Clausen-Schaumann, and H. E. Gaub. **How strong is a covalent bond?** . *Science (New York, N.Y.)*, **283**(5408):1727–1730, 1999.

[13] M. Ludwig, M. Rief, L. Schmidt, H. Li, F. Oesterhelt, M. Gautel, and H. E. Gaub. **AFM, a tool for single-molecule experiments**. *Appl Phys A Mater Sci Proc*, **68**(2):173–176, 1999.

[14] Jorge Alegre-Cebollada, Raul Perez-Jimenez, Pallav Kosuri, and Julio M. Fernandez. **Single-molecule force spectroscopy approach to enzyme catalysis.** *J Biol Chem*, **285**(25):18961–18966, 2010.

[15] Raul Perez-Jimenez, Alvaro Inglés-Prieto, Zi-Ming Zhao, Inmaculada Sanchez-Romero, Jorge Alegre-Cebollada, Pallav Kosuri, Sergi Garcia-Manyes, T. Joseph Kappoc k, Masaru Tanokura, Arne Holmgren, Jose M. Sanchez-Ruiz, Eric A. Gaucher, and Julio M. Fernandez. **Single-molecule paleoenzymology probes the chemistry of resurrected enzymes**. *Nat Struct Mol Biol*, **advance online publication**, 2011.

[16] Arun P. Wiita, Raul Perez-Jimenez, Kirstin A. Walther, Frauke Grater, B. J. Berne, Arne Holmgren, Jose M. Sanchez-Ruiz, and Julio M. Fernandez. **Probing the chemistry of thioredoxin catalysis with force**. *Nature*, **450**(7166):124–127, 2007.

[17] Ronen Berkovich, Sergi Garcia-Manyes, Michael Urbakh, Joseph Klafter, and Julio M. Fernandez. **Collapse Dynamics of Single Proteins Extended by Force**. *Biophys J*, **98**(11):2692–2701, 2010.

[18] J. M. Fernandez and H. B. Li. **Force-clamp spectroscopy monitors the folding trajectory of a single protein**. *Science*, **303**:1674–1678, 2004.

[19] Jan P. Junker, Fabian Ziegler, and Matthias Rief. **Ligand-Dependent Equilibrium Fluctuations of Single Calmodulin Molecules**. *Science*, **323**(5914):633–637, 2009.

[20] E. A. Shank, C. Cecconi, J. W. Dill, S. Marqusee, and C. Bustamante. **The folding cooperativity of a protein is controlled by its chain topology**. *Nature*, **465**(7298):637–640, 2010.

[21] Josep M. Huguet, Cristiano V. Bizarro, Núria Forns, Steven B. Smith, and Felix Bustamante, Carlos nad Ritort. **Single-molecule derivation of salt dependent base-pair free energies in DNA**. *Proc Natl Acad Sci USA*, **107**:15431–6, 2010.

[22] Hagar Zohar, Craig L. Hetherington, Carlos Bustamante, and Susan J. Muller. **Peptide nucleic acids as tools for single-molecule sequence detection and manipulation.** *Nano Lett*, **10**(11):4697–4701, 2010.

[23] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch. **A New Method of Measuring Nuclear Magnetic Moment**. *Physical Review*, **53**:318–318, 1938.

[24] Vincent Breukels, Albert Konijnenberg, Sanne M. Nabuurs, Jurgen F. Doreleijers, Nadezda V. Kovalevskaya, and Geerten W. Vuister. *Overview on the Use of NMR to Examine Protein Structure*. John Wiley & Sons, Inc., 2001.

[25] James Keeler. *Understanding NMR Spectroscopy*. University of Cambridge, Department of Chemistry, 2002.

REFERENCES

[26] G.E Martin and A.S. Zekter. *Two-Dimensional NMR Methods for Establishing Molecular Connectivity: A Chemist's Guide to Experiment Selection, Performance, and Interpretation.* VCH Publishers, Inc:New York, 1988.

[27] César Fernández and Gerhard Wider. **TROSY in NMR studies of the structure and function of large biological macromolecules.** *Curr Opin Struct Biol*, **13**(5):570–580, 2003.

[28] Jocelyne Fiaux, Eric B. Bertelsen, Arthur L. Horwich, and Kurt Wüthrich. **NMR analysis of a 900K GroEL GroES complex.** *Nature*, **418**(6894):207–211, 2002.

[29] M. A. Martí-Renom, A. C. Stuart, A. Fiser, R. Sánchez, F. Melo, and A. Sali. **Comparative protein structure modeling of genes and genomes.** *Annual review of biophysics and biomolecular structure*, **29**(1):291–325, 2000.

[30] A. K. Bronowska. *Thermodynamics of ligand-protein interactions: implications for molecular design. Thermodynamics (Book I)*. InTech Publ, 2011.

[31] Jiali Gao, Shuhua Ma, Dan T. Major, Kwangho Nam, Jingzhi Pu, and Donald G. Truhlar. **Mechanisms and free energies of enzymatic reactions.** *Chem Rev*, **106**(8):3188–3209, 2006.

[32] Hao Hu and Weitao Yang. **Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods.** *Annu Rev Phys Chem*, **59**(1):573–601, 2008.

[33] Hans Martin M. Senn and Walter Thiel. **QM/MM methods for biomolecular systems.** *Angew Chem Int Ed Engl*, **48**(7):1198–1229, 2009.

[34] Wenjin Li and Frauke Gräter. **Atomistic Evidence of How Force Dynamically Regulates Thiol/Disulfide Exchange**. *J Am Chem Soc*, **132**(47):16790–16795, 2010.

[35] Raul Perez-Jimenez, Jingyuan Li, Pallav Kosuri, Inmaculada Sanchez-Romero, Arun P. Wiita, David Rodriguez-Larrea, Ana Chueca, Arne Holmgren, Ant onio Miranda-Vizuete, Katja Becker, Seung-Hyun H. Cho, Jon Beckwith, Eric Gelhaye, Jean P. Jacquot, Eric A. Gaucher, Eric Gaucher, Jose M. Sanchez-Ruiz, Bruce J. Berne, and Juli o M. Fernandez. **Diversity of chemical mechanisms in thioredoxin catalysis revealed by single-molecule force spectroscopy.** *Nat Struct Mol Biol*, **16**(8):890–896, 2009.

[36] Marco De Vivo. **Bridging quantum mechanics and structure-based drug design.** *Front Biosci*, **16**:1619–1633, 2011.

[37] Victor Guallar and Frank H. Wallrapp. **QM/MM methods: looking inside heme proteins biochemistry.** *Biophys Chem*, **149**(1-2):1–11, 2010.

[38] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertrán. **Theoretical QM/MM studies of enzymatic pericyclic reactions.** *Interdiscip Sci*, **2**(1):115–131, 2010.

[39] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. **A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules**. *J Am Chem Soc*, **117**(19):5179–5197, 1995.

[40] W. L. Jorgensen and J. Tirado-Rives. **The OPLS potential functions for proteins - energy minimizations for crystals of cyclic-peptides and crambin.** *J Am Chem Soc*, **110**:1657–1666, 1988.

[41] D. K. Klimov and D. Thirumalai. **Mechanisms and kinetics of beta-hairpin formation.** *Proc Natl Acad Sci USA*, **97**(6):2544–2549, 2000.

[42] Albert Galera-Prat, Angel Gómez-Sicilia, Andres F. Oberhauser, Marek Cieplak, and Mariano Carrión-Vázquez. **Understanding biology by stretching proteins: recent progress.** *Curr Opin Struct Biol*, **20**(1):63–69, 2010.

[43] M Gao, H Lu, and K Schulten. **Unfolding of titin domains studied by molecular dynamics simulations**. *J Muscle Res Cell Motil*, **23**(5-6):513–521, 2002.

[44] M Gao, M Wilmanns, and K Schulten. **Steered molecular dynamics studies of titin I1 domain unfolding**. *Biophys J*, **83**(6):3435–3445, 2002.

[45] P. E. Marszalek, H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten, and J. M. Fernandez. **Mechanical unfolding intermediates in titin modules.** *Nature*, **402**(6757):100–103, 1999.

[46] Murat Cetinkaya, Senbo Xiao, and Frauke Gräter. **Bottom-up computational modelling of semi-crystalline fibers: From atomistic to continuum scale**. *Phys Chem Chem Phys*, **13(22)**:10426–10429, 2011.

[47] Murat Cetinkaya, Senbo Xiao, Bernd Markert, Wolfram Stacklies, and Frauke Gräter. **Silk fiber mechanics from multiscale force distribution analysis.** *Biophys J*, **100**(5):1298–1305, 2011.

[48] T. Darden, D. York, and L. Pedersen. **Particle mesh Ewald: An N . log(N) method for Ewald sums in large systems**. *J Chem Phys*, **98**(12):10089–10092, 1993.

[49] J. Ryckaert, G. Ciccotti, and H. Berendsen. **Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes**. *J Comput Phys*, **23**(3):327–341, 1977.

[50] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. **LINCS: A linear constraint solver for molecular simulations**. *J Comput Chem*, **18**(12):1463–1472, 1997.

[51] D. Pearlman. **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**. *Computer Physics Communications*, **91**(1-3):1–41, 1995.

[52] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kalé, and Klaus Schulten. **Scalable molecular dynamics with NAMD.** *J Comput Chem*, **26**(16):1781–1802, 2005.

[53] David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E. Mark, and Herman J. C. Berendsen. **GROMACS: Fast, flexible, and free**. *J Comp Chem*, **26**:1701–1718, 2005.

[54] Wolfram Stacklies, M. Cristina Vega, Matthias Wilmanns, and Frauke Gräter. **Mechanical Network in Titin Immunoglobulin from Force Distribution Analysis**. *PLoS Comput Biol*, **5**(3):e1000306, 2009.

[55] William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives. **Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids**. *J Am Chem Soc*, **118**(45):11225–11236, 1996.

[56] Chris Oostenbrink, Alessandra Villa, Alan E. Mark, and Wilfred F. van Gunsteren. **A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6.** *J Comput Chem*, **25**(13):1656–1676, 2004.

[57] Jay W. Ponder and David A. Case. **Force fields for protein simulations.** *Adv Protein Chem*, **66**:27–85, 2003.

[58] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. **The missing term in effective pair potentials**. *J Phys Chem*, **91**(24):6269–6271, 1987.

[59] Kahled Toukan and Aneesur Rahman. **Molecular-dynamics study of atomic motions in water**. *Physical Review B*, **31**(5):2643–2648, 1985.

[60] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R.W. Impey, and M. L. Klein. **Comparison of simple potential functions for simulating liquid water.** *J Chem Phys*, **79**:926–935, 1983.

[61] Michael W. Mahoney and William L. Jorgensen. **A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions**. *J Chem Phys*, **112**(20):8910–8922, 2000.

[62] Elias M. Puchner, Alexander Alexandrovich, Ay L. Kho, Ulf Hensen, Lars V. Schäfer, Birgit Brandmeier, Frauke Gräter, Helmut Grubmüller, Hermann E. Gaub, and Mathias Gautel. **Mechanoenzymatics of titin kinase**. *Proc Natl Acad Sci USA*, **105**(36):13385–13390, 2008.

[63] T E Fisher, P E Marszalek, and J M Fernandez. **Stretching single molecules into novel conformations using the atomic force microscope**. *Nat Struct Biol*, **7**(9):719–724, 2000.

[64] A. Janshoff, M. Neitzert, Y. Oberdorfer, and H. Fuchs. **Force spectroscopy of molecular systems - Single molecule spectroscopy of polymers and biomolecules**. *Angew Chem Int Ed Engl*, **39**:3213–3237, 2000.

[65] Keir C. Neuman and Attila Nagy. **Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy**. *Nat Methods.*, **5**(6):491–505, 2008.

[66] H Grubmüller, B Heymann, and P Tavan. **Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force**. *Science*, **271**(5251):997–999, 1996.

[67] G. I. Bell. **Models for the specific adhesion of cells to cells**. *Science*, **200**(4342):618–627, 1978.

[68] Emily B. Walton, Sunyoung Lee, and Krystyn J. Van Vliet. **Extending Bell's model: how force transducer stiffness alters measured unbinding forces and kinetics of molecular complexes.** *Biophys J*, **94**(7):2621–2630, 2008.

[69] Michael Geisler, Senbo Xiao, Elias M. Puchner, Frauke Grater, and Thorsten Hugel. **Controlling the Structure of Proteins at Surfaces.** *Journal of the American Chemical Society*, **132**(48):17277–17281, 2010.

[70] E. Puchner. **Comparing Proteins by Their Unfolding Pattern**. *Biophys J*, **95**(1):426–434, 2008.

[71] Wolfram Stacklies, Fei Xia, and Frauke Gräter. **Dynamic allostery in the methionine repressor revealed by force distribution analysis.** *PLoS Comput Biol*, **5**(11):e1000574+, 2009.

[72] S Xiao, W Stacklies, M Cetinkaya, B Markert, and F Gräter. **Mechanical response of silk crystalline units from force-distribution analysis**. *Biophys J*, **96**(10):3997–4005, 2009.

[73] G. Pelosi. **The finite-element method, Part I: R. L. Courant [Historical Corner]**. *IEEE Antennas and Propagation Magazine*, **49**(2):180–182, 2007.

[74] Jacob Fish and Ted Belytschko. *A First Course in Finite Elements.* John Wiley & Sons, Ltd., 2007.

[75] Roger W Pryor. *Multiphysics Modeling Using COMSOL–A First Principles Approach.* Jones and Bartlett Publishers, LLC, 2011.

[76] O.C. Zienkiewicz and R.L. Taylor. *The Finite Element Method for Solid and Structural Mechanics (Sixth edition).* Elsevier Ltd., 2005.

[77] J M Gosline, P A Guerette, C S Ortlepp, and K N Savage. **The mechanical design of spider silks: from fibroin sequence to mechanical function**. *J Exp Biol*, **202**(Pt 23):3295–3303, 1999.

[78] Fritz Vollrath and David P. Knight. **Liquid crystalline spinning of spider silk**. *Nature*, **401**:541–548, 2001.

[79] John Gatesy, Cheryl Hayashi, Dagmara Motriuk, Justin Woods, and Randolph Lewis. **Extreme Diversity, Conservation, and Convergence of Spider Silk Fibroin Sequences**. *Science*, **291**:2603–2605, 2001.

[80] Paul A. Guerette, David G. Ginzinger, Bernhard H. F. Weber, and John M. Gosline. **Silk Properties Determined by Gland-Specific Expression of a Spider Fibroin Gene Family**. *Science*, **272**:112–115, 1996.

REFERENCES

[81] Yasuhiro Takahashi, Mikio Gehoh, and Kimio Yuzuriha. **Structure refinement and diffuse streak scattering of silk (Bombyx mori)**. *Int J Biol Macromol*, **24**:127–138, 1999.

[82] ME Rousseau, Cruz D Hernandez, MM West, AP Hitchcock, and M. Pézolet. **Nephila clavipes spider dragline silk microstructure studied by scanning transmission X-ray microscopy**. *J Am Chem Soc*, **129**:3897–905, 2007.

[83] David T. Grubb and Lynn W. Jelinski. **Fiber Morphology of Spider Silk: The Effects of Tensile Deformation**. *Macromolecules*, **30**:2860–2867, 1997.

[84] T. Lefevre, M. Rousseau, and M. Pezolet. **Protein Secondary Structure and Orientation in Silk as Revealed by Raman Spectromicroscopy**. *Biophys J*, **92**(8):2885–2895, 2007.

[85] Tetsuo Asakura, Michi Okonogi, Yasumoto Nakazawa, and Kazuo Yamauchi. **Structural analysis of alanine tripeptide with antiparallel and parallel beta-sheet structures in relation to the analysis of mixed beta-sheet structures in Samia cynthia ricini silk protein fiber using solid-state NMR spectroscopy.** *J Am Chem Soc*, **128**(18):6231–6238, 2006.

[86] I Krasnov, I Diddens, N Hauptmann, G Helms, M Ogurreck, T Seydel, S S Funari, and M Müller. **Mechanical properties of silk: interplay of deformation on macroscopic and molecular length scales**. *Phys Rev Lett*, **100**(4):048104–048104, 2008.

[87] D. Porter, F. Vollrath, and Z. Shao. **Predicting the mechanical properties of spider silk as a model nan ostructured polymer.** *Eur Phys J E Soft Matter*, **16**(2):199–206, 2005.

[88] S Keten and M J Buehler. **Geometric confinement governs the rupture strength of H-bond assemblies at a critical length scale**. *Nano Lett*, **8**:743–748, 2008.

[89] T P Knowles, A W Fitzpatrick, S Meehan, H R Mott, M Vendruscolo, C M Dobson, and M E Welland. **Role of intermolecular forces in defining material properties of protein nanofibrils**. *Science*, **318**:1900–1903, 2007.

[90] Tetsuo Asakura, Rena Sugino, Tatsushi Okumura, and Yasumoto Nakazawa. **The role of irregular unit, GAAS, on the secondary structure of Bombyx mori silk fibroin studied with ^{13}C CP/MAS NMR and wide-angle X-ray scattering**. *Protein Sci*, **11**:1873–1877, 2002.

[91] T. Darden, D. York, and L. Pedersen. **Particle Mesh Ewald — An N log(N)method for Ewald sums in large systems.** *J Chem Phys*, **98**:10089–10092, 1993.

[92] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. **LINCS: A linear constraint solver for molecular simulations.** *J Comp Chem*, **18**:1463–1472, 1997.

[93] W. G. Hoover. **Canonical dynamics: equilibrium phase-space distributions**. *Phys Rev A*, **31**:1695–1697, 1985.

[94] S. Nosé. **A molecular dynamics method for simulations in the canonical ensemble**. *Mol Phys.*, **52**:255–268, 1984.

[95] H. J. C. Berendsen. **Transport properties computed by linear response through weak coupling to a bath.** In *Computer Simulations in Material Science*, pages 139–155, 1991.

[96] H Grubmüller, B Heymann, and P Tavan. **Ligand binding: Molecular mechanics calculation of the streptavidin biotin rupture force**. *Science*, **271**:997–999, 1996.

[97] J. Skolnick and M. Fixman. **Electrostatic persistence length of a wormlike polyelectrolyte**. *Macromolecules*, **10**:944, 1977.

[98] Frauke Gräter, Pascal Heider, Ronen Zangi, and B. J. Berne. **Dissecting entropic coiling and poor solvent effects in protein collapse**. *J Am Chem Soc*, **130**:11578+, 2008.

[99] Hongbin Li, Andres F. Oberhauser, Susan B. Fowler, Jane Clarke, and Julio M. Fernandez. **Atomic force microscopy reveals the mechanical design of a modular protein**. *Proc Natl Acad Sci USA*, **97**:6527–6531, 2000.

[100] Hui Lu and Klaus Schulten. **Steered molecular dynamics simulation of conformational changes of immunoglobulin domain I27 interprete atomic force microscopy observations**. *J Chem Phys*, **247**:141–153, 1999.

[101] Philip M. Williams, Susan B. Fowler, Robert B. Best, Josè Luis Toca-Herrera, Kathryn A. Scott, Annette Steward, and Jane Clarke. **Hidden complexity in the mechanical properties of titin**. *Nature*, **422**:446–449, 2003.

[102] E M Lupton, F Achenbach, J Weis, C Bräuchle, and I Frank. **Modified chemistry of siloxanes under tensile stress: interaction with environment**. *J Phys Chem B*, **110**(30):14557–14563, 2006.

[103] P Schwaderer, E Funk, F Achenbach, J Weis, C Bräuchle, and J Michaelis. **Single-molecule measurement of the strength of a siloxane bond**. *Langmuir*, **24**(4):1343–1349, 2008.

[104] Alexandre V. Morozov, Tanja Kortemme, Kiril Tsemekhman, and David Baker. **Close agreement between the orientation dependence of hydrogen bonds observed in proteinstructures and quantum mechanical calculations**. *Proc Natl Acad Sci USA*, **101**:6946–6951, 2004.

[105] S Xiao, M Stacklies, C Debes, and F Gräter. **Force distribution determines optimal length of β-sheet crystals for mechanical robustness**. *Soft Matter*, **4(7)**:1308–1311, 2010.

[106] D Bisig, P Weber, L Vaughan, K H Winterhalter, and K Piontek. **Purification, crystallization and preliminary crystallographic studies of a two fibronectin type-III domain segment from chicken tenascin encompassing the heparin- and contactin-binding regions**. *Acta Crystallogr D Biol Crystallogr*, **55**(Pt 5):1069–1073, 1999.

[107] M Pfuhl and A Pastore. **Tertiary structure of an immunoglobulin-like domain from the giant muscle protein titin: a new member of the I set**. *Structure*, **3**(4):391–401, 1995.

[108] J T Nielsen, M Bjerring, M D Jeppesen, R O Pedersen, J M Pedersen, K L Hein, T Vosegaard, T Skrydstrup, D E Otzen, and N C Nielsen. **Unique identification of supramolecular structures in amyloid fibrils by solid-state NMR spectroscopy**. *Angew Chem Int Ed Engl*, **48**(12):2118–2121, 2009.

[109] P A Guerette, D G Ginzinger, B H Weber, and J M Gosline. **Silk properties determined by gland-specific expression of a spider fibroin gene family**. *Science*, **272**(5258):112–115, 1996.

[110] Y Takahashi, M Gehoh, and K Yuzuriha. **Structure refinement and diffuse streak scattering of silk (Bombyx mori)**. *Int J Biol Macromol*, **24**(2-3):127–138, 1999.

[111] A G Murzin, S E Brenner, T Hubbard, and C Chothia. **SCOP: a structural classification of proteins database for the investigation of sequences and structures**. *J Mol Biol*, **247**(4):536–540, 1995.

[112] D Frishman and P Argos. **Knowledge-based protein secondary structure assignment**. *Proteins*, **23**(4):566–579, 1995.

[113] W Kabsch and C Sander. **Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features**. *Biopolymers*, **22**(12):2577–2637, 1983.

[114] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. **GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation**. *J Chem Theory Comp*, **4**(3):435–447, 2008.

[115] Fritz Vollrath and Donald T. Edmonds. **Modulation of the mechanical properties of spider silk by coating with water**. *Nature*, **340**(6231):305–307, 1989.

[116] Nathan Becker, Emin Oroudjev, Stephanie Mutz, Jason P. Cleveland, Paul K. Hansma, Cheryl Y. Hayashi, Dmitrii E. Makarov, and Helen G. Hansma. **Molecular nanosprings in spider capture-silk threads**. *Nat Mater*, **2**(4):278–283, 2003.

[117] Victoria Brookes, Robert Young, and Fritz Vollrath. **Deformation micromechanics of spider silk**. *J Mate Sci*, **43**(10):3728–3732, 2008.

[118] E. Oroudjev, J. Soares, S. Arcidiacono, J. B. Thompson, S. A. Fossey, and H. G. Hansma. **Segmented nanofibers of spider dragline silk: Atomic force microscopy and single-molecule force spectroscopy**. *Proc Natl Acad Sci USA*, **99**(Suppl 2):6460–6465, 2002.

[119] S. Rammensee, U. Slotta, T. Scheibel, and A. R. Bausch. **Assembly mechanism of recombinant spider silk proteins**. *Proc Natl Acad Sci USA*, 2008.

[120] S. Fossey. **Atomistic modeling of interphases in spider silk fibers**. *Int J Biol Macromol*, **24**(2-3):119–125, 1999.

[121] Yves Termonia. **Molecular Modeling of Spider Silk Elasticity**. *Macromolecules*, **27**(25):7378–7381, 1994.

[122] C Y Hayashi, N H Shipley, and R V Lewis. **Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins**. *Int J Biol Macromol*, **24**(2-3):271–275, 1999.

[123] J. D. van Beek, S. Hess, F. Vollrath, and B. H. Meier. **The molecular structure of spider dragline silk: Folding and orientation of the protein backbone**. *Proc Natl Acad Sci USA*, **99**(16):10266–10271, 2002.

[124] Michael B. Hinman and Randolph V. Lewis. **Isolation of a Clone Encoding a Second Dragline Silk Fibroin:** *Nephila Clavipes* **Silk is a Two-Protein Fiber**. *J Biol Chem*, **267**:19320–19324, 1992.

[125] Alexandra Simmons, Ed Ray, and Lynn W. Jelinski. **Solid-State 13C NMR of Nephila clavipes Dragline Silk Establishes Structure and Identity of Crystalline Regions**. *Macromolecules*, **27**(18):5235–5237, 1994.

[126] Zhengzhong Shao, Fritz Vollrath, Yong Yang, and Hans C. Thøgersen. **Structure and Behavior of Regenerated Spider Silk**. *Macromolecules*, **36**(4):1157–1161, 2003.

[127] Sinan Keten, Zhiping Xu, Britni Ihle, and Markus J. Buehler. **Nanoconfinement controls stiffness, strength and mechanical toughness of [beta]-sheet crystals in silk**. *Nat Mater*, **9**(4):359–367, 2010.

[128] F. Vollrath and D. Porter. **Spider silk as a model biomaterial**. *Appl Phys A Mater Sci Proc*, **82**(2):205–212, 2006.

[129] Shankar Kumar, John M. Rosenberg, Djamal Bouzida, Robert H. Swendsen, and Peter A. Kollman. **THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method**. *J Comput Chem*, **13**(8):1011–1021, 1992.

[130] K. Guess. **Thermal analysis of major ampullate (drag line) spider silk: the effect of spinning rate on tensile modulus**. *Thermochim Acta*, **315**(1):61–66, 1998.

[131] D. Saravanan. **Spider Silk Structure, Properties and Spinning.** *Journal of Textile and Apparel, Technology and Management*, **5**:1–14, 2006.

[132] D.W. van Krevelen and Klaas te Nijenhuis. *Properties of polymers : their correlation with chemical structure : their numerical estimation and prediction from additive group contributions*. Amsterdam : Elsevier, 2009.

[133] Yi Liu, Alexander Sponner, David Porter, and Fritz Vollrath. **Proline and Processing of Spider Silks**. *Biomacromolecules*, **9**(1):116–121, 2008.

[134] Tamara Köhler and Fritz Vollrath. **Thread biomechanics in the two orb-weaving spiders Araneus diadematus (Araneae, Araneidae) and Uloborus walckenaerius (Araneae, Uloboridae)**. *J Exp Zool*, **271**(1):1–17, 1995.

[135] B. O. Swanson, T. A. Blackledge, J. Beltrán, and C. Y. Hayashi. **Variation in the material properties of spider dragline silk across species**. *Appl Phys A Mater Sci Proc*, **82**(2):213–218, 2006.

REFERENCES

[136] C. Basire and D. A. Ivanov. **Evolution of the lamellar structure during crystallization of a semicrystalline-amorphous polymer blend: time-resolved hot-stage SPM study.** *Phys Rev Lett*, **85**(26 Pt 1):5587–5590, 2000.

[137] O. Rathore and D. Y. Sogah. **Self-assembly of β-sheets into nanostructures by poly(alanine) segments incorporated in multiblock copolymers inspired by spider silk.** *J Am Chem Soc*, **123**(22):5231–5239, 2001.

[138] Chengjie Fu, David Porter, and Zhengzhong Shao. **Moisture Effects on Antheraea pernyi Silk's Mechanical Property**. *Macromolecules*, **42**(20):7877–7880, 2009.

[139] Murat Cetinkaya, Senbo Xiao, and Frauke Gräter. **The Effects of Crystalline Subunit Size on Silk Fiber Mechanics**. *submitted*, 2011.

[140] C. Vepari and D. Kaplan. **Silk as a biomaterial**. *Prog Polym Sci*, **32**(8-9):991–1007, 2007.

[141] Ken N. Savage and John M. Gosline. **The effect of proline on the network structure of major ampullate silks as inferred from their mechanical and optical properties**. *J Exp Biol*, **211**(12):1937–1947, 2008.

[142] Ning Du, Xiang Y. Liu, Janaky Narayanan, Lian Li, Matthew L. Lim, and Daiqin Li. **Design of Superior Spider Silk: From Nanostructure to Mechanical Properties**. *Biophys J*, **91**(12):4528–4535, 2006.

[143] Roxana Ene, Periklis Papadopoulos, and Friedrich Kremer. **Combined structural model of spider dragline silk**. *Soft Matter*, **5**(22):4568–4574, 2009.

[144] Frank K. Ko and Jovan Jovicic. **Modeling of mechanical properties and structural design of spider web.** *Biomacromolecules*, **5**(3):780–785, 2004.

[145] F Hagn, L Eisoldt, J G Hardy, C Vendrely, M Coles, T Scheibel, and H Kessler. **A conserved spider silk domain acts as a molecular switch that controls fibre assembly**. *Nature*, **465**(7295):239–242, 2010.

[146] H. Shulha, C. Pofoo, D. Kaplan, and V. Tsukruk. **Unfolding the multi-length scale domain structure of silk fibroin protein**. *Polymer*, **47**(16):5821–5830, 2006.

[147] Yuko Aoyanagi and Ko Okumura. **Simple Model for the Mechanics of Spider Webs**. *Phys Rev Lett*, **104**(3):038102+, 2010.

[148] Andrea Nova, Sinan Keten, Nicola M. Pugno, Alberto Redaelli, and Markus J. Buehler. **Molecular and Nanostructural Mechanisms of Deformation, Strength and Toughness of Spider Silk Fibrils**. *Nano Lett*, **10**(7):2626–2634, 2010.

[149] Murat Cetinkaya, Senbo Xiao, Bernd Markert, Wolfram Stacklies, and Frauke Gräter. **Silk Fiber Mechanics From Multiscale Froce Distribution Analysis**. *Biophys J*, **,**:*in press*, 2010.

[150] K. Zhang, F. W. Si, H. L. Duan, and J. Wang. **Microstructures and mechanical properties of silks of silkworm and honeybee.** *Acta Biomater*, **6**(6):2165–2171, 2010.

[151] Periklis Papadopoulos, Jan Sölter, and Friedrich Kremer. **Hierarchies in the structural organization of spider silk – a quantitative model**. *Colloid Polym Sci*, **287**(2):231–236, 2009.

[152] Anita Goel and Viola Vogel. **Harnessing biological motors to engineer systems for nanoscale transport and assembly**. *Nat Nanotechnol*, **3**(8):465–475, 2008.

[153] Mehmet Sarikaya, Candan Tamerler, Alex K. Jen, Klaus Schulten, and Francois Baneyx. **Molecular biomimetics: nanotechnology through biology**. *Nat Mater*, **2**(9):577–585, 2003.

[154] Lisa A. Capriotti, Thomas P. Beebe, and Joel P. Schneider. **Hydroxyapatite Surface-Induced Peptide Folding**. *J Am Chem Soc*, **129**(16):5281–5287, 2007.

[155] Yegor A. Domanov and Paavo K. Kinnunen. **Islet amyloid polypeptide forms rigid lipid-protein amyloid fibrils on supported phospholipid bilayers.** *J Cell Biol*, **376**(1):42–54, 2008.

[156] Christopher M. Dobson. **Protein folding and misfolding**. *Nature*, **426**(6968):884–890, 2003.

[157] Thomas Scheibel. **Spider silks: recombinant synthesis, assembly, spinning, and engineering of synthetic proteins.** *Microbial cell factories*, **3**(1):14+, 2004.

[158] Michael Geisler, Tobias Pirzer, Christian Ackerschott, Simon Lud, Jose Garrido, Thomas Scheibel, and Thorsten Hugel. **Hydrophobic and Hofmeister effects on the adhesion of spider silk proteins onto solid substrates: an AFM-based single-molecule study.** *Langmuir*, **24**(4):1350–1355, 2008.

[159] D. Horinek, A. Serr, M. Geisler, T. Pirzer, U. Slotta, S. Q. Lud, J. A. Garrido, T. Scheibel, T. Hugel, and R. R. Netz. **Peptide adsorption on a hydrophobic surface results from an interplay of solvation, surface, and intrapeptide forces**. *Proc Natl Acad Sci USA*, **105**(8):2842–2847, 2008.

[160] Yi Cao and Hongbin Li. **Engineered elastomeric proteins with dual elasticity can be controlled by a molecular regulator.** *Nat Nanotechnol*, **3**(8):512–516, 2008.

[161] Andreas Lammel, Martin Schwab, Ute Slotta, Gerhard Winter, and Thomas Scheibel. **Processing conditions for the formation of spider silk microspheres.** *ChemSusChem*, **1**(5):413–416, 2008.

[162] Claudia Friedsam, Angelika K. Wehle, Ferdinand Kühner, and Hermann E. Gaub. **Dynamic single-molecule force spectroscopy: bond rupture analysis with variable spacer length**. *J Phys Condens Matter*, **15**(18):S1709+, 2003.

[163] Mariano Carrion-Vazquez, Piotr E. Marszalek, Andres F. Oberhauser, and Julio M. Fernandez. **Atomic force microscopy captures length phenotypes in single proteins**. *Proc Natl Acad Sci USA*, **96**(20):11288–11292, 1999.

[164] Luca M. Ghiringhelli, Berk Hess, Nico F. van der Vegt, and Luigi Delle Site. **Competing adsorption between hydrated peptides and water onto metal surfaces: from electronic to conformational properties.** *J Am Chem Soc*, **130**(40):13460–13464, 2008.

[165] Pim Schravendijk, Luca M. Ghiringhelli, Luigi D. Site, and Van D. Vegt. **Interaction of Hydrated Amino Acids with Metal Surfaces: A Multiscale Modeling Description**. *J Phys Chem C*, **111**(6):2631–2642, 2007.

[166] M. Schlierf and M. Rief. **Temperature Softening of a Protein in Single-molecule Experiments**. *J Mol Biol*, **354**(2):497–503, 2005.

[167] Ute K Slotta, Sebastian Rammensee, Stanislav Gorb, and Thomas Scheibel. **An Engineered Spider Silk Protein Forms Microspheres**. *Angew Chem Int Ed Engl*, **47**(24):4592–4594, 2008.

[168] O. Sumner Makin and Louise C. Serpell. **Structures for amyloid fibrils.** *FEBS J*, **272**(23):5950–5961, 2005.

[169] Mariano Carrion-Vazquez, Hongbin Li, Hui Lu, Piotr E. Marszalek, Andres F. Oberhauser, and Julio M. Fernandez. **The mechanical stability of ubiquitin is linkage dependent.** *Nat Struct Biol*, **10**(9):738–743, 2003.

[170] Andres F. Oberhauser, Piotr E. Marszalek, Harold P. Erickson, and Julio M. Fernandez. **The molecular elasticity of the extracellular matrix protein tenascin**. *Nature*, **393**(6681):181–185, 1998.

[171] Felix Hanke and Hans Jürgen J. Kreuzer. **Breaking bonds in the atomic force microscope: theory and analysis.** *Phys Rev E Stat Nonlin Soft Matter Phys*, **74**(3 Pt 1), 2006.

[172] S. León, C. Alemán, and S. Muñoz-Guerra. **Structure of the Extended Crystal Forms of Nylon-6**. *Macromolecules*, **33**:5754–5756, 2000.

[173] Youyong Li and William A. Goddard. **Nylon 6 Crystal Structures, Folds, and Lamellae from Theory**. *Macromolecules*, **35**(22):8440–8455, 2002.

[174] Xiaohui Liu, Qiuju Wu, and Lars A. Berglund. **Polymorphism in polyamide 66/clay nanocomposites**. *Polymer*, **43**(18):4967–4972, 2002.

[175] N. Jones. **Chain-folded lamellar crystals of aliphatic polyamides. Investigation of nylons 4 8, 4 10, 4 12, 6 10, 6 12, 6 18 and 8 12**. *Polymer*, **38**(11):2689–2699, 1997.

[176] N. A. Jones, E. D. T. Atkins, M. J. Hill, S. J. Cooper, and L. Franco. **Polyamides with a Choice of Structure and Crystal Surface Chemistry. Studies of Chain-Folded Lamellae of Nylons 8 10 and 10 12 and Comparison with the Other 2N 2(N + 1) Nylons 4 6 and 6 8**. *Macromolecules*, **30**(12):3569–3578, 1997.

[177] T. D. Fornes and D. R. Paul. **Crystallization behavior of nylon 6 nanocomposites**. *Polymer*, **44**(14):3945–3961, 2003.

[178] Carsten Baldauf, Robert Günther, and Hans-Jörg Hofmann. **Helix Formation and Folding in γ-Peptides and Their Vinylogues**. *HCA*, **86**(7):2573–2588, 2003.

[179] B. Dalo, G. Coppola, and B. Pallesi. **Studies of crystalline forms of nylon-6 by X-ray and i.r. spectrophotometry**. *Polymer*, **15**(3):130–132, 1974.

[180] D. R. Holmes, C. W. Bunn, and D. J. Smith. **The crystal structure of polycaproamide: Nylon 6**. *J Polym Sci*, **17**(84):159–177, 1955.

[181] H. Arimoto. **α–γ Transition of nylon 6**. *J Polym Sci A Gen Pap*, **2**(5):2283–2295, 1964.

[182] Felix Berkremeier, Morten Bertz, Senbo Xiao, Nikos Poinotsis, Matthias Wilmanns, Frauke Gräter, and Matthias Rief. **Fast Folding α-Helices as Reversible Strain Absorbers in the Muscle Protein Myomesin**. *submitted*, 2011.

[183] Y. Au. **The muscle ultrastructure: a structural perspective of the sarcomere.** *Cell Mol Life Sci*, **61**(24):3016–3033, 2004.

[184] C. Gregorio. **Muscle assembly: a titanic achievement?** *Curr Opin Cell Biol*, **11**(1):18–25, 1999.

[185] Hongbin Li, Wolfgang A. Linke, Andres F. Oberhauser, Mariano Carrion-Vazquez, Jason G. Kerkvliet, Hui Lu, Piotr E. Marszalek, and Julio M. Fernandez. **Reverse engineering of the giant muscle protein titin**. *Nature*, **418**(6901):998–1002, 2002.

[186] H. Granzier and S. Labeit. *Titin and Its associated proteins: the third myofilament system of the sarcomere*, **71** of *Advances in Protein Chemistry*, pages 89–119. Elsevier, 2005.

[187] B. K. Grove, V. Kurer, C. Lehner, T. C. Doetschman, J. C. Perriard, and H. M. Eppenberger. **A new 185,000-dalton skeletal muscle protein detected by monoclonal antibodies.** *J Cell Biol.*, **98**(2):518–524, 1984.

[188] J. Squire, H. Alkhayat, C. Knupp, and P. Luther. *Molecular Architecture in Muscle Contractile Assemblies*, **71** of *Advances in Protein Chemistry*, pages 17–87. Elsevier, 2005.

[189] W. M. Obermann, U. Plessmann, K. Weber, and D. O. Fürst. **Purification and biochemical characterization of myomesin, a myosin-binding and titin-binding protein, from bovine skeletal muscle.** *Eur J Biochem*, **233**(1):110–115, 1995.

[190] Atsushi Fukuzawa, Stephan Lange, Mark Holt, Anna Vihola, Virginie Carmignac, Ana Ferreiro, Bjarne Udd, and Mathias Gautel. **Interactions with titin and myomesin target obscurin and obscurin-like 1 to the M-band - implications for hereditary myopathies**. *J Cell Sci*, pages jcs.028019+, 2008.

[191] Wolfgang M. J. Obermann, Mathias Gautel, Klaus Weber, and Dieter O. Furst. **Molecular structure of the sarcomeric M band: mapping of titin and myosin binding domains in myomesin and the identification of a potential regulatory phosphorylation site in myomesin**. *EMBO J*, **16**(2):211–220, 1997.

REFERENCES

[192] Stephan Lange, Mirko Himmel, Daniel Auerbach, Irina Agarkova, Katrin Hayess, Dieter O. Fürst, Jean-Claude C. Perriard, and Elisabeth Ehler. **Dimerisation of myomesin: implications for the structure of the sarcomeric M-band.** *J Mol Biol*, **345**(2):289–298, 2005.

[193] E. Ehler, B. M. Rothen, S. P. Hämmerle, M. Komiyama, and J. C. Perriard. **Myofibrillogenesis in the developing chicken heart: assembly of Z-disk, M-line and the thick filaments.** *J Cell Sci*, **112 (Pt 10)**:1529–1539, 1999.

[194] Matthew J. Potthoff, Michael A. Arnold, John McAnally, James A. Richardson, Rhonda Bassel-Duby, and Eric N. Olson. **Regulation of skeletal muscle sarcomere integrity and postnatal muscle function by Mef2c.** *Mol Cell Biol.*, **27**(23):8143–8151, 2007.

[195] Irina Agarkova, Elisabeth Ehler, Stephan Lange, Roman Schoenauer, and Jean-Claude Perriard. **M-band: a safeguard for sarcomere stability?** *J Muscle Res Cell Motil*, **24**(2):191–203, 2003.

[196] Irina Agarkova and Jean-Claude C. Perriard. **The M-band: an elastic web that crosslinks thick filaments in the center of the sarcomere.** *Trends Cell Biol*, **15**(9):477–485, 2005.

[197] P. Bertoncini, R. Schoenauer, I. Agarkova, M. Hegner, J. C. Perriard, and H. J. Güntherodt. **Study of the Mechanical Properties of Myomesin Proteins Using Dynamic Force Spectroscopy**. *J Cell Biol*, **348**(5):1127–1137, 2005.

[198] Roman Schoenauer, Patricia Bertoncini, Gia Machaidze, Ueli Aebi, Jean-Claude C. Perriard, Martin Hegner, and Irina Agarkova. **Myomesin is a molecular spring with adaptable elasticity.** *J Mol Biol*, **349**(2):367–379, 2005.

[199] N Pinotsis, S Lange, J C Perriard, D I Svergun, and M Wilmanns. **Molecular basis of the C-terminal tail-to-tail assembly of the sarcomeric filament protein myomesin**. *EMBO J*, **27**(1):253–264, 2008.

[200] G. Vriend. **WHAT IF: a molecular modeling and drug design program.** *J Mol Graph*, **8**(1), 1990.

[201] Carsten Kutzner, David Van der Spoel, Martin Fechner, Erik Lindahl, Udo W. Schmitt, Bert L. De Groot, and Helmut Grubmuller. **Software news and update - Speeding up parallel GROMACS on high-latency networks**. *J Comp Chem*, **28**:2075–2084, 2007.

[202] Gustav G. Knappeis and Frits Carlsen. **The Ultrastructure of The M Line in Skeletal Muscle**. *J Cell Biol*, **38**(1):202–211, 1968.

[203] P. Luther. **Three-dimensional structure of the vertebrate muscle M-region**. *J Cell Biol*, **125**(3):313–324, 1978.

[204] Hind A. Al-Khayat, Robert W. Kensler, Edward P. Morris, and John M. Squire. **Three-dimensional structure of the M-region (bare zone) of vertebrate striated muscle myosin filaments by single-particle analysis.** *J Mol Biol*, **403**(5):763–776, 2010.

[205] Roman Schoenauer, Stephan Lange, Alain Hirschy, Elisabeth Ehler, Jean-Claude C. Perriard, and Irina Agarkova. **Myomesin 3, a novel structural component of the M-band in striated muscle.** *J Mol Biol*, **376**(2):338–351, 2008.

[206] I. Agarkova, D. Auerbach, E. Ehler, and J. C. Perriard. **A novel marker for vertebrate embryonic heart, the EH-myomesin isoform.** *J Biol Chem*, **275**(14):10256–10264, 2000.

[207] S. Labeit and B. Kolmerer. **Titins: giant proteins in charge of muscle ultrastructure and elasticity.** *Science (New York, N.Y.)*, **270**(5234):293–296, 1995.

Declaration

I herewith declare that I have produced this thesis without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This thesis has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted from August 2007 to April 2011 under the supervision of Dr. Frauke Gräter at Heidelberg University.

Heidelberg, Germany