INAUGURAL - DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich - Mathematischen

Gesamtfakultät

der Ruprecht - Karls - Universität

Heidelberg

vorgelegt von

Dipl.-Phys. Johannes Andreas WAGNER

aus Arnstadt

Tag der mündlichen Prüfung: 21.07.2017

How internal stress influences molecular properties on multiple scales

Gutachter: Prof. Dr. Dieter Heermann
Prof. Dr. Joachim Spatz

Zusammenfassung

Interner Stress kann entscheidend die funktionalen Eigenschaften molekularer Strukturen bestimmen. Die wichtige Rolle der internen Stressverteilung in gespannten Molekülen und Materialen wird zunehmend erkannt. Die hier vorgestellte Arbeit untersucht den Einfluss von Stress auf unterschiedlichen Längenskalen für zwei bestimmte molekulare Prozesse: chemische Reaktionen und das Strecken einer komplex aufgebauten Faser. Dabei wird ein sehr interdisziplinärer Ansatz unter Verwendung von quantenmechanischen Berechnungen und Finite-Elemente Simulationen verfolgt.

Der erste Teil hatte zum Ziel, den Beitrag von Stress zur chemischen Reaktivität besser zu verstehen. Ein Hauptergebnis der quantenmechanischen Berechnungen ist, dass die Kinetik von Cycloadditionen durch verschiedene stereoisometrische Zustände sehr genau kontrolliert und eingestellt werden kann, und dies im direkten Zusammenhang mit dem bei der Reaktion involvierten molekularen Stress steht. Im zweiten Teil wurde auf einer größeren Längenskala die Beiträge von Stress auf die mechanischen Eigenschaften von komplexen Materialien analysiert. Die Finite-Elemente Simulationen sagen eine stress-induzierte Ordnung in der Struktur von Spinnenseide vorher, ein bisher unbeschriebenes Phänomen, das durch Neutronenstreuungsexperimente bestätigt wurde. Jenseits des grundlegenden Verständnisses der untersuchten Systeme eröffnen die Ergebnisse der vorliegenden Arbeit neue Wege, die Reaktionskinetik und Spezifität von Cycloadditionen für das selektive Markieren von Makromolekülen passgenau einzustellen sowie die Biomechanik von Spinnenseidenfasern für das Design von neuen Kompositmaterialien zu nutzen.

Abstract

INTERNAL stress in molecular structures is increasingly appearing as a crucial factor in determining functional properties at a large range of length scales. The present work has investigated the effect of stress for two distinct molecular processes, chemical reactions and straining of a two-phase fiber, using a highly interdisciplinary approach involving quantum-mechanical and finite element simulations. First, with the aim of better understanding the contribution of stress to reactivity, we discovered that kinetics of strain-promoted cycloadditions can be finely tuned by exploiting the reactants' stereoisomeric and regioselective states and thereby the strain coupled to the reaction. Secondly, at longer length scales, we dissected the contributions of stress to the mechanical properties of complex materials. Our simulations predict a stress-induced ordering in spider silk, which was validated by scattering experiments. Beyond the basic understanding of the investigated systems, the findings presented in this work open new scenarios for the possibility to enhance strain-dependent chemical reactivity and the mechanics of composite materials.

Dedicated to my parents.

List of Publications

During the time of this Ph.D., the following papers have been published or prepared for publication:

A **Dynamic Prestress in a Globular Protein**
S.A. Edwards, J.A. Wagner, F. Gräter
PLOS Computational Biology **8** (2012) 5:6e1002509

B **Origin of Orthogonality of Strain-Promoted Click Reactions**
J.A. Wagner, D. Mercadante, I. Nikić, E.A. Lemke, F. Gräter
Chemistry: a European Journal **21** (2015) 35:12431-12435

C **Modulation of Electron Demand in benzyl-derivatives for Strain-Promoted Click Reactions**
J.A. Wagner, D. Mercadante, F. Gräter
In preparation

D **Stress-induced long-range ordering in spider silk**
J.A. Wagner, S.P. Patil, I. Greving, M. Lämmel, K. Gkagkas, T. Seydel, K. Kroy, M. Müller, B. Markert, and F. Gräter
In review at Scientific Reports

Contents

Chapter 1

Introduction

1.1 Mechanical Stress on Multiple Scales

In the strict definition of continuum mechanics, stress is the internal force per unit cross-sectional area. As a macroscopic concept it requires the force and unit area to be in a homogenous state. However, at microscopic length scales, where atomic interactions govern, area cannot be described using the macroscopic definition as forces are defined by the individual inter-particle interactions. In this case the term stress is not uniquely defined, and instead internal forces are used to measure stress (see Figures 1.1 & 1.2). Stress can occur in a large variety of forms such as tensile, contractile, shear or viscous (or frictional). For a fluid or gas at equilibrium, stress equals pressure. While stress, as a measure of internal force and strain as a measure of deformation, are inherently linked to each other, stress can exist without strain, but not *vice versa*. In that sense, stress is prevalent in our description of systems from atomistic length scales such as small molecules with strained bonds to mesoscopic and macroscopic length-scales such as structural macromolecules or load-bearing fibers and beyond. The influence and effects of stress are found in all branches of natural sciences and their study is fundamental to our understanding of many aspects of Nature. While the influence of stress on a continuum mechanics level is well described by the theories of elasticity, plasticity, (Non-) Newtonian fluids and rheology, on molecular length-scales it is much more complex and not often obvious. As an example, allostery in proteins is signal transduction through stress propagation [1, 2]. New findings continue to emerge such as how very localized stress caused by crystal defects affect rupture strength and how it can be relieved by inserting biomolecules in the regions of highest stress gradients [3, 4]. Another example reports how combined static compression and tension stresses in a tensegrity fashion, were found to be present in cells [5] and proteins [6] and defined as pre-stress. The causes of stress can be manifold, it can be internal as shown in Fig. 1.1 or such as the above mentioned tensegrity structures in proteins.

Otherwise stress can simply be imposed by applying external forces (Figure 1.3). The influence of applied stress on the behavior of macromolecules or bulk materials can be very diverse. Stress on solid bulk materials generates deformations that can fall into the plastic or elastic regime. A plastic deformation of bulk materials on a molecular scale can be either a rupture of covalent bonds or breakage of non-covalent interactions. The exploitation of bulk materials' elastic properties, however, is one of the most pursued areas of research targeting the enhancement of material properties or the creation of new, smart, materials. On the other hand, the application

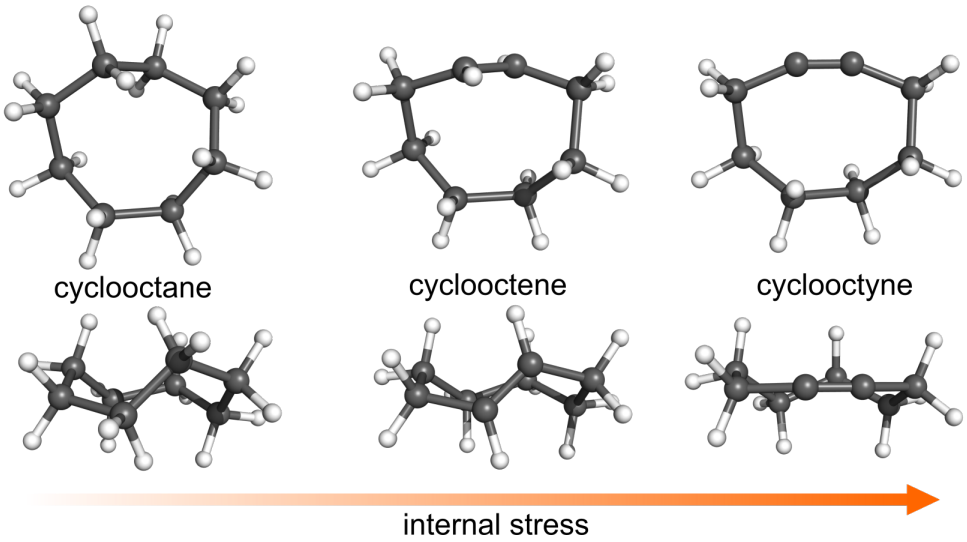

Figure 1.1: Internal stress in eight-membered cyclic compounds. Most cyclic compounds contain strain energy by non-equilibrium C-C bonds which are subject to angular, torsional and Van der Waals or also called *Pitzer* strain [7]. From left to right the internal strain increases by incrementing bond multiplicity from single (cyclooctane) to double (cyclooctene) and triple bonds (cyclooctyne). The illustration was rendered using the program PyMol [8].

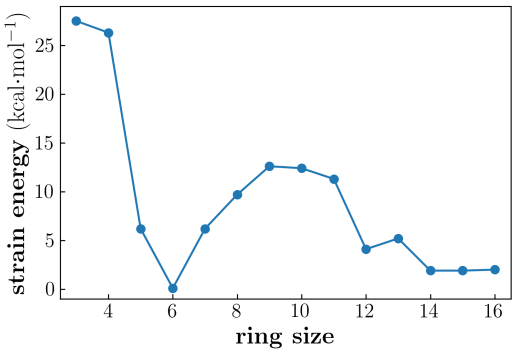

Figure 1.2: Strain energy in cycloalkanes as a function of ring size [9]. The C-C-C bonds have a preferred tetrahedral angle of 109.5° which, depending on ring size, drive cyclic compounds in stable chair, boat-chair or crown conformations. The deviations from the ideal angles and bond-lengths imposed by cyclicity create intrinsic strain and therefore and impose strain energy. The six-membered cyclohexane in chair conformation has minimal strain energy due to almost ideal C-C-C angles, whereas for large rings the increased degrees of freedom reduce strain.

of stress on a molecule or the modulation of the internal stress within a molecular structure can lead to functionally specific conformational variations. These aim to fullfill specified functions, promoting the occurence of molecular interactions or propagating a signal across the molecular structure [10, 11]. Overall, these conforma-

tional variations, either occurring in bulk materials or molecules, where individual atomic interactions govern, can be modulated within different micro-environmental conditions such as variations of temperature, chemical state or application of electromagnetic fields [12].

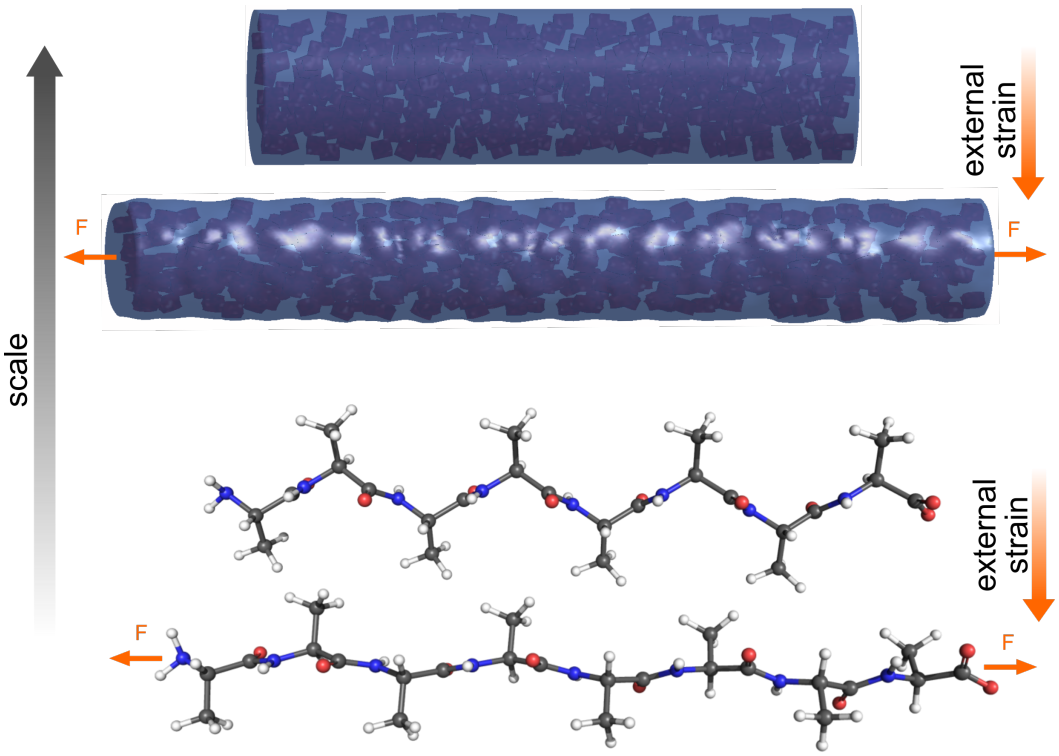

Figure 1.3: External strain at different length-scales. Top: section of a fiber in equilibrium and (below) strained using external force. Bottom: peptide composed of eight alanine amino acids in the linear β-strand conformation and (below) subjected to external force.

In most cases stress is defined such that the stressed system maintains static equilibrium, *i.e.* the system can not accelerate, only deform. At certain stress levels, materials can undergo plastic deformation in different ways. They can, for instance, fracture or rupture as seen in graphene sheets [13] or exhibit plastic flow as it has been shown for lithium fluoride crystals [14] or glass [15]. They can also alter crystal structures as it was shown for Ni_2MnGa crystals [16], reversibly induce crystal transitions for polymers [17], or cause stress-induced crystallization, as it was shown for polyurethane elastomers [18] and natural rubber [19, 20]. Lastly, external applied stress has also been shown to change chemical states by force-induced covalent-bond activation [21, 22] or biasing reaction pathways [23]. When under stress, most systems can elastically or also partially inelastically absorb and store internal energy which, upon release, can be transformed to strain or other

forms of work [24]. Most interestingly, in a similar fashion chemical reactions can directly be tuned by exploiting the intrinsic strain that exists within reactants that is then released during the reaction upon atoms rearrangement [25, 26, 27]. These examples all highlight the importance of stress in a diverse range of processes found across multiple length-scales.

1.2 Aims and Motivation

Considering the manyfold occurences of stress at a molecular level, the relationship linking causes and effects of stress can be very complex. The motivation for this work arises from the interest in understanding particular effects caused by stress, or exploiting certain properties of stressed systems and steer them towards a desired behavior. Therefore, in this work we aim to study on a theoretical level the influence of stress at two distinct length-scales, both in the microscopic regime.

We aim to show how stress can be exploited to steer a system towards desired properties. On a small-molecule scale, we investigate chemical reactions with strained cyclic compounds. They can harness their spring-loaded internal energy in a very specific way with highly tunable reaction rates. The application for this kind of reactions under the umbrella of "click-chemistry" is already successfully deployed in labeling proteins for visualization purposes in super-resolution microscopy. Our goal is to expand the scope and versatility of these applications by better understanding the tunability of strain-promoted click-reactions.

On a larger, yet microscopic scale, we aim to contribute to the understanding of the influence of stress on the structural topology of a fibrous system. A very interesting system still lacking profound understanding is the spider silk-fiber, a two-phase spider silk-fiber that exibits astounishing mechanical properties. We aim to contribute to the understanding of the influence of external stress on changes in internal configuration of the two phases of the highly ordered and unordered regions, which originate from different secondary structure parts in the spider silk-protein.

1.3 Techniques

Several very distinct simulation methods are at hand to study molecular and material properties on a theoretical basis at various length- and time scales (Figure 1.4).

Quantum Mechanical simulations, more precisely Hartree-Fock and Density Functional calculations, are used to study the electronic configuration and thus reaction kinetics of small strained cyclic molecules. By means of activation energies, molecular orbital energy levels, electron densities and conformations of the reactants, products and transition states, a comprehensive picture of the system properties can be obtained. In this thesis, QM was used to study the kinetics of strained cycloaddition reactions. The basic principles are described in Section 2.1. Additionally, experimental reaction rates are used to validate activation energies, such that a strong conclusion of the simulation results can be derived.

Finite element simulations are a well established simulation technique for continuum materials. They present a big approximation by neglecting any molecular interactions. In this thesis finite element simulations with viscoelastic and elasto-plastic material models with parameters derived from MD simulations are deployed to study the structural stress response on a two-phase spider silk fiber. Due to the availability of Small-angle Neutron Scattering experimental data, the scatter intensities are calculated based on the finite element simulation data to allow a direct validatian of the numerical simulations.

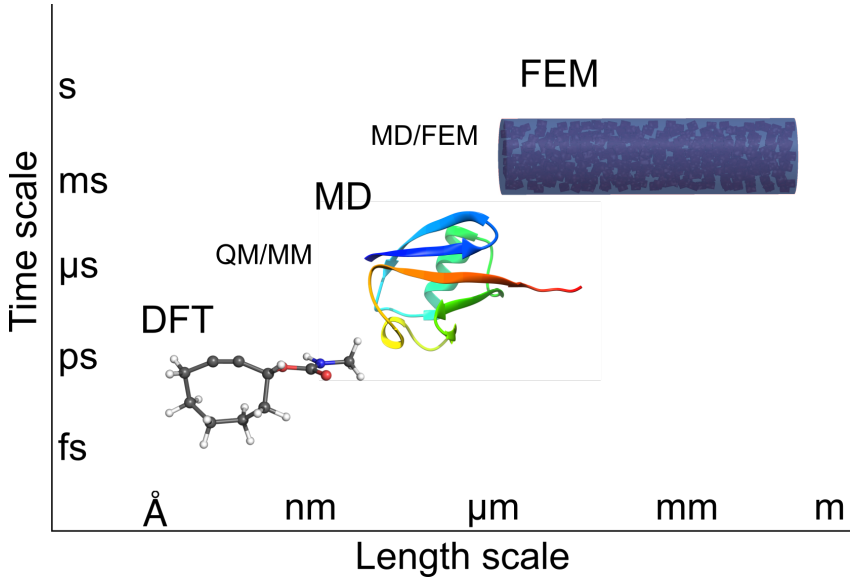

Figure 1.4: Range of simulation techniques suiteable to study phenomena from microscopic over mesoscopic to macroscopic length- and time scales. For small molecules such as the depcited strained-cyclooctene, shown in ball-and-stick representation, *ab-initio* or density functional theory (DFT) methods are used. For larger molecules such as ubiquitin (PDB code: 1UBQ), here shown in cartoon representation, molecular dynamics simulations are used. All-atom molecular dynamics simulation can be scaled up by coarse-graining multiple atoms in one coarse-grain interaction site. At the upper end, for large systems or bulk materials, continum mechanics based methods such as finite element methods (FEM) are used.

1.4 Objectives

Guided by the central aim to understand the influence of stress on molecular systems we formulate two main goals for this thesis.
Firstly, on an atomistic scale we will study the reaction kinetics of a set of strain-promoted cycloaddition reactions by quantum-chemical simulation methods. We try to answer two questions: Which properties of strained cyclic compounds and the reaction partners allow to tune the kinetics in a controlled fashion? And can we find design principles to obtain a reaction set with the desired properties?

Secondly, on a continuum mechanics scale we will study the influence of external stress on the structure and conformation of a general two-phase fiber model. Using a bottom-up approach adopting parameters from atomistic simulations on the highly ordered and disordered phases of spider silk fibers, we design and mesh an almost full-scale finite element model of the spider-silk fibril. The goal is to characterize structural changes induced by stress, which ideally can be validated experimentally. Taken together, this thesis addresses understanding stress-induced properties of molecules on two vastly different length-scales, namely the molecular sub-nanometer and the macroscopic micrometer scale.

1.5 Outline of this Thesis

This thesis is organized into the following chapters: Chapter 2 introduces the essential principles and concepts of the major computational techniques used in this work. These are primarily quantum mechanics and continuum mechanics methods. This chapter also includes a brief introduction of small-angle neutron scattering concepts, as this experimental technique was used for direct comparison of computational data. Furthermore, neccesary theoretical concepts to analyze and interpret the data are introduced.

Chapter 3 outlines the influence of different strain-promoted cycloaddition click reactions on reaction kinetics in two parts. The first part derives determinants which allow reactant pairs to undergo slow or fast kinetics. The related second part addresses the influence of different functional groups on the reactions electron demand modulated via a benzyl ring on a tetrazine reactant, which undergoes cycloaddition with a strained cyclic eight-membered ring.

Chapter 4 investigates the influence of strain and stress on the structural order in a two-phase fiber by finite element simulations. Here, experimental small angle neutron scattering served as a complementing method to assess the nature of ordering. Lastly, in Chapter 5 a summary, conclusion and outline future work is provided.

Chapter 2

Fundamentals

Contents

For the understanding of effect of stress on molecules and materials, a phletora of computational methods has been devised, which covers a wide range of time and length scales. This chapter introduces a subset of these methods, namely quantum mechanical, molecular dynamics and finite element calculations. We will introduce and develop the fundamental theoretical concepts relevant for their applications within this thesis.

The first part of this chapter is dedicated to the basics of quantum mechanics. Necessary assumptions and approximations are outlined, which allow numerical

calculations of the electronic configuration of complex molecular systems where no analytical solutions exist (Section 2.1). The higher levels of theories used in this thesis, Hartee-Fock (Section 2.1.2) and Density Functional Theory (Section 2.1.3) are described in-depth, as well as the relevant calculated properties for this thesis: energy levels of individual electrons, total system energy and electron densities. The following sections will focus on necessary concepts to understand and explain reaction kinetics and the link to experimental rate constants. Section 2.1.6 covers frontier molecular orbital theory, which as we show in Chapter 3 is one of the theories to explain differences in activation energies upon introducing functional groups. The concepts of strain (or distortion) and interaction energies are explained in Section 2.1.8, which plays an important role in bisecting energy contributions between intramolecular distortion or stress and electronic interaction.

The bridge between the length scales of a quantum mechanical descriptions of small molecules and a continuum mechanics-based description of material systems is atomistic Molecular Dynamics (MD) simulations based on the equations of motion formulated by Newton. In this second part (Section 2.2) we briefly introduce this computational technique. Molecululer Dynamics simulations have not been used directly within this thesis. However, parameters required for the FE simulations of silk have been based on MD simulations under force. In addition, MD simulations have been used in the publication A: 'Dynamic Prestress in a Globular Protein' [6], which is not further described in this thesis.

In the third part of this chapter, we introduce the continuum mechanics based finite element (FE) method together with a short introduction on meshing (Section 2.3). Finite element methods allow the exploration of length and time scales inaccessible to computationally more expensive lower level discrete particle based methods such as Molecular Dynamics simulations. FE methods are wide-spread computational techniques to analyze and predict structural mechanics of materials and objects. They here allowed us to model and describe the silk fiber and to predict its structural and mechanical properties.

Parts of this thesis made use of and also predicted structural information of silk fibers as obtained from small-angle neutron scattering experiments. The third and last section of this chapter will briefly review the fundamentals of small-angle scattering (Section 2.4). We will limit this to static, inelastic scattering of a neutron beam and its measured quantity of scattering intensity, which allows access to microscopic structural information in equilibrium. Finally, in we will introduce the related scattering quantities, the structure factor and form factor, both of which are relevant to interpret structural data jointly from experimental and simulated data.

2.1 Quantum Mechanics

The introductory section on the quantum mechanical framework relevant for this work is adapted from *Introduction to Computational Chemistry* [28] and limited to the non-relativistic descriptions. The subsection of QM solvation models (Section 2.1.5) is further derived from the excellent review by Mennucci *et al.* [29].

2.1.1 Schrödinger Equation

The theory of quantum mechanics is the description of properties of matter at small scales, where classical Newtonian physics as used for Molecular Dynamics simulations (Section 2.2) fails to correctly reproduce observations. As opposed to Newtonian mechanics, where uniquely knowing a particle's two measurable states, momentum $\mathbf{p}$ and position $\mathbf{x}$, is sufficient to describe any future state, quantum mechanics has no such unique states. In contrast, quantum mechanics uses the concept of probabilities to describe measurable properties, and thus its states. All measurable properties of a system, such as position, angular momentum or energy are assigned mathematical objects called observables. To each such observable belongs a set of states which return unique results upon measurement. These states are called Eigenstates, and the respective measurement value is an Eigenvalue. Depending on the mathematical description, a quantum mechanical state is represented either as a vector $|\psi\rangle$ in Hilbert space $\mathcal{H}$, a complex inner product space, or as a wavefunction ψ. By definition, the state ψ contains all information about the system and is a superposition of all possible states, until a measurement reduces it to its Eigenstates. The Eigenvalue equation for the observable of the energy is referred to as the stationary or also called time-independent Schrödinger equation:

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle\,. \tag{2.1}$$

The Hamilton operator $\hat{H}$ is defined as

$$\hat{H} = \frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{x}) \tag{2.2}$$

for a single particle of mass m moving inside the potential V. This allows for an analytical solution for the hydrogen atom, a single electron moving in the electrostatic potential of a proton.

Adiabatic and Born-Oppenheimer Approximation. Complex molecular systems beyond the H_2^+ molecule do not have analytical solutions of the full Schrödinger equation. To be able to solve such systems, the *Born-Oppenheimer* and *adiabatic* approximation are used. The main concept originates from the idea that the big difference in inertia of electrons and nuclei ($m_n/m_e \geq 1836$) stands in contrast to similar Coulomb forces from identical absolute charges. Max Born's adiabatic theorem states: 'A physical system remains in its instantaneous eigenstate if a given perturbation is acting on it slowly enough and if there is a gap between the eigenvalue

and the rest of the Hamiltonian's spectrum' [30]. This implies that in this approximation, the light electrons instantaneously adjust their position with respect to the heavy nuclei positions. The movement of the electrons therefore happens on a much faster time-scale, considered to be independent of the movement of the nuclei. Therefore, both movements can be described independently of each other.

To describe more complicated multiple particle systems such as molecules, the total Hamilton operator $\hat{H}_{\text{tot}}$ for a molecule can be split into

$$\hat{H}_{\text{tot}} = \left[\hat{T}_{\text{e}} + \hat{V}_{\text{e-e}} + \hat{V}_{\text{e-n}}\right] + \hat{T}_{\text{n}} + \hat{V}_{\text{n-n}} \tag{2.3}$$

where the kinetic terms $\hat{T}$ and potential terms $\hat{V}$ are split into the contributions from electrons and nuclei: electronic-electronic repulsion $\hat{V}_{\text{e-e}}$, nuclear-nuclear repulsion $\hat{V}_{\text{n-n}}$, and electronic-nuclear attraction $\hat{V}_{\text{e-n}}$. The electronic Hamilton $\hat{H}_{\text{e}}$ in the center of mass coordinate system is composed of

$$\hat{H}_{\text{e}} = \hat{T}_{\text{e}} + \hat{V}_{\text{e-e}} + \hat{V}_{\text{e-n}} + \hat{V}_{\text{n-n}} \tag{2.4}$$

and only depends on the positions of the electrons and nuclei ($\hat{V}_{\text{e-n}}$,$\hat{V}_{\text{n-n}}$), but not on the nuclear momenta. The nuclear

Therefore with electron positions $\mathbf{r}$ and nuclei positions $\mathbf{R}$, the Schrödinger equation (2.1) can be written as

$$\hat{H}_{\text{tot}}\psi_{\text{tot}}(\mathbf{R},\mathbf{r}) = E_{\text{tot}}\psi_{\text{tot}}(\mathbf{R},\mathbf{r}). \tag{2.5}$$

The total wavefunction of the electrons and nuclei can be expanded into a complete set of electron wavefunctions with nuclear coordinate coefficients $\psi_{\text{n}i}(\mathbf{R})$:

$$\psi_{\text{tot}}(\mathbf{R},\mathbf{r}) = \sum_{i}^{\infty} \psi_{\text{n}i}(\mathbf{R}) \cdot \psi_{\text{e}i}(\mathbf{R},\mathbf{r}) \ . \tag{2.6}$$

Combining equation (2.5) and (2.6), one can obtain after integration (more detailed in [28]):

$$\begin{aligned} \nabla_{\text{n}}^2\psi_{\text{n}i} + E_j\psi_{\text{n}i} + \sum_{i=1}^{\infty} \big\{ \langle\psi_j|\nabla_{\text{n}}|\psi_i\rangle\,(\nabla_{\text{n}}\psi_{\text{n}i}) + \langle\psi_j|\nabla_{\text{n}}^2|\psi_i\rangle\,\psi_{\text{n}i} \\ + \langle\psi_j|\hat{H}_{\text{mp}}|\psi_i\rangle\,\psi_{\text{n}i}\big\} = E_{\text{tot}}\psi_{\text{n}j} \end{aligned} \tag{2.7}$$

where $\hat{H}_{\text{mp}}$ is the mass-polarization Hamiltonian [28]. In mathematical terms, the *adiabatic* approximation restricts the total wave function to one electronic surface by neglecting all coupling elements $\langle\psi_j|\nabla_{\text{n}}^2|\psi_i\rangle$, $i \neq j$ in Equation (2.7) while the *Born-Oppenheimer* approximation further neglects the diagonal terms $\langle\psi_j|\nabla_{\text{n}}^2|\psi_j\rangle$ in Equation (2.7), resulting in a Schrödinger equation for the nuclear component:

$$\left[\hat{T}_{\text{n}} + E_{\text{e}j}(\mathbf{R})\right] \cdot \psi_{\text{n}j}(\mathbf{R}) = E \cdot \psi_{\text{n}j}(\mathbf{R}) \tag{2.8}$$

where E_{ej} effectively composes the potential energy surface (PES) on which the nuclei move.

To obtain the Schrödinger equation for the electronic component, the nuclei are considered stationary to the electrons allowing to neglect the kinetic term of the nuclear Hamilton operator $\hat{T}_{\mathrm{n}}(\mathbf{R}) \approx 0$. Applying the total Hamilton (Equation (2.3)) on the total wave function (Equation (2.6)) results in an electronic Schrödinger equation with an attractive ($\hat{V}_{\text{e-n}}$) and repulsive ($\hat{V}_{\text{n-n}}$) potential that has Eigenstates $\psi_{\mathrm{e}j}$ and Eigenenergies $E_{\mathrm{e}j}$:

$$\hat{H}_{\mathrm{e}}(\mathbf{r}, \mathbf{R}) \cdot \psi_{\mathrm{e}j}(\mathbf{r}, \mathbf{R}) = E_{\mathrm{e}j} \cdot \psi_{\mathrm{e}j}(\mathbf{r}, \mathbf{R}) \tag{2.9}$$

where the positions $\mathbf{R}$ of the nuclei are kept constant. It is now possible to obtain $E_{\mathrm{e}}(\mathbf{R})$ by repeatedly solving Equation (2.9) and varying $\mathbf{R}$ in small steps and succesively solving Equation (2.8).

The *Born-Oppenheimer* together with the *adiabatic* approximation are generally a good approximation for ground states, where solutions of energy Eigenvalues are in close proximity [28].

2.1.2 Hartree-Fock Method

The Hartree-Fock method, also called Self-consistent Field Theory, allows the calculation of wave functions or orbital energies based on the 'Hartree' or 'independent particles' approximation [31, 32]. It states that electrons interact via the mean field created by the other electrons. Simply stated, electrons avoid space already occupied by other electrons by means of Coulomb repulsion. This allows the separation of the wave function $\hat{\Psi}$ into the product of individual electron wave functions:

$$\hat{\Psi}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = \Phi_1(\mathbf{x}_1) \cdot \Phi_2(\mathbf{x}_2) \cdot ... \cdot \Phi_n(\mathbf{x}_N) \tag{2.10}$$

also known as 'Hartree product'. The HF method additionally accounts for the spin of electrons by one-electron wave functions that are a product of spatial and spin wave functions. To obey the Pauli principle, stating that two electrons cannot occupy the same orbital with equal spin, the total electronic wave function is required to be antisymmetric. Consequently, a single Slater determinant is introduced as an approximation to compose the one-electron trial wave function:

$$\Phi_{\mathrm{SD}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(N) & \phi_2(N) & \dots & \phi_N(N) \end{vmatrix}, \quad \langle \phi_i | \phi_j \rangle = \delta_{ij}$$

which, by swapping two columns or rows, changes the sign and thus obeys the antisymmetry principle. Without interacting electron, the total Hamiltonian can then be described by a set of one-electron Hamiltonians $\hat{F}_i$, called the 'Fock' operators. The resulting pseudo-eigenvalue Hartee-Fock equations:

$$\hat{F}_i \phi_i' = \varepsilon_i \phi_i' \tag{2.11}$$

then describe single electrons i where the corresponding ε_i are interpreted as the orbital energies. Each Fock operator

$$\hat{F}_i = \hat{h}_i + \sum_j^N (\hat{J}_j - \hat{K}_j) \tag{2.12}$$

consists of an one electron operator $\hat{h}_i$ describing the attractive term caused by the potential of all the nuclei and terms for the repulsion with other electrons: $\hat{J}$ for Coulomb self interaction and $\hat{K}$ for exchange interaction. To numerically compute Hartree-Fock equations, the wave functions for the molecular orbitals are expanded into a linear combination of atomic orbitals (LCAO) centered on the nuclei positions. These functions are commonly called basis functions. An expansion into M basis functions χ_α is given by:

$$\phi_i = \sum_\alpha^M c_{\alpha i} \chi_\alpha \tag{2.13}$$

In general, the more and better suited basis functions are used, the better the approximation. The derivation originates from the analytical solution of the wave functions of the hydrogen atom with a similar angular and radial part. The orbitals of this type are also known as Slater Type Orbitals (STO) and have the form

$$\chi_k \propto \cdot e^{-\alpha |\mathbf{r} - \mathbf{R}_k|}. \tag{2.14}$$

Even though the STO matrix elements are better suited for electronic structure calculations, they usually do not have an analytical solution, hence in practice Gaussian Type Orbitals (GTO) of the form

$$\chi_k \propto \cdot e^{-\alpha (\mathbf{r} - \mathbf{R}_k)^2} \tag{2.15}$$

are almost exclusively used.
A Hartree-Fock problem is solved with the following algorithm:

1. intial guess of density matrix of molecular orbitals
2. form fock matrix
3. diagonalize fock matrix
4. determining HF potential
5. form new density matrix
6. repeat step (2 – 5) until self-consistency is reached

The Hartree-Fock method is commonly used, also in combination with other methods. Despite the fact that even by deploying an infinite number of STOs or GTOs, they can never reach the exact solution of the Schrödinger equation, but instead reach the Hartree-Fock limit. The reason for this is the above mentioned 'independent particles approximation', neglecting direct electron-electron interactions. However, the Hartree-Fock method is well suited to obtain electronic energy levels, very accurate electron densities and dipole moments. According to the variational principle, any inital guess of an approximate one-electron wave function has an energy equal or above the ground state energy. If these initial guess 'trial' wave functions do not yield the exact energy, they can be minimized as a function of their parameters. As a consequence, the *ab intio* Hartree-Fock energy yields an upper bound to the exact energy, which is called the 'Hartree-Fock limit' that other approximative methods cannot [33, 34, 35].

2.1.3 Density Functional Theory

Density Functional Theory (DFT) is another method to obtain the ground state of a quantum mechanical many-body system. The big advantage over other methods is that DFT does not require a complete solution of the Schrödinger equation (2.1) of the whole electronic system. The name is derived from its use of functionals: functions of electron densities, which themselves again are functions. Similar to the HF theory, the system's wave function is decomposed into N single-electron wave functions, called the Kohn-Sham function ϕ_j. Here, the Hohenberg-Kohn-Theorem states that for a given electron density $n(\mathbf{x})$ the respective potential $V(\mathbf{x})$ and thus the ground-state wave function Ψ are uniqely defined. Or vice versa, the ground-state of an N-electron system is uniqely defined by an electron density $n(\mathbf{x})$. Therefore, instead of calculating the Schrödinger equation for N electrons simultaneously, the problem reduces to determine N single-electron wave functions with independent solutions of one Schrödinger equation with only 3 spatial coordinates in the electron density. These single electron wave functions are termed Kohn-Sham-equations:

$$\left[-\frac{\hbar}{2m}\nabla^2 + v_{\text{eff}}(\mathbf{x})\right]\varphi_j(\mathbf{x}) = -\varepsilon_j\varphi_j(\mathbf{x}) \tag{2.16}$$

and the density is obtained by the sum of all Kohn-Sham function densities:

$$n(\mathbf{x}) = \sum_{j=1}^{N} |\varphi_j(\mathbf{x})|^2. \tag{2.17}$$

The effective potential is then in turn obtained from the density

$$v_{\text{eff}}(\mathbf{x}) = v(\mathbf{x}) \int \frac{n(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} d^3x' + v_{\text{xc}}(\mathbf{x}). \tag{2.18}$$

DFT approximations. The inherent problem of DFT is that the exact functional for the exchange-correlation ($v_{\text{xc}}(\mathbf{x})$) is mostly unknown. The many-body problem

has only been shifted to $v_{xc}(\mathbf{x})$, but not actually been solved. Since $v_{\mathrm{xc}}(\mathbf{x})$ depends on the electron density $n(\mathbf{x})$ of all positions, which can be exactly calculated only for a few simple problems like an electron gas.

1. Local Density Approximation (LDA). To address this issue in DFT methods, a local-density approximation is used

$$E_{\mathrm{XC}}^{\mathrm{LDA}}(n) = \int \varepsilon_{\mathrm{XC}}(n) n(\mathbf{x}) \mathrm{d}^3 r. \tag{2.19}$$

   In the LDA, as the name suggests, it is assumed that the exchange correlation $v_{\mathrm{xc}}(\mathbf{x})$ only depends on the local density $n(\mathbf{x})$ at this very position. This is naturally a very good approximation where the density is relatively uniform as is the case for an electron gas. If the the density is not the same everywhere, LDA typically over-estimates the exchange-correlation energy and therefore under-estimates bond lengths.

2. Generalized Gradient Approximation (GGA): To mitigate this effect, in addition to the local density $n(\mathbf{x})$, its gradient can be considered for a better global density estimate. This type of approximation is termed generalized gradient approximation (GGA) and is still commonly used nowadays.

$$E_{\mathrm{XC}}^{\mathrm{GGA}}(n) = \int \varepsilon_{\mathrm{XC}}(n, \nabla n) n(\mathbf{x}) \mathrm{d}^3 x \tag{2.20}$$

3. Meta-GGA: Taking this approach one step further are meta-GGA functionals, which additionally use the second derivative of the electron density $n(\mathbf{x})$. Examples of this type are the TPSS [36] and Minnesota family [37] of functionals. Due to their nature of using local densities, DFT methods are generally not well suited for longer range van der Waals interactions. Opposite to the Kohn-Sham methods, electron orbital energy differences are generally underestimated using LDA and GGA approximations. Meta-GGA was the DFT method of choice in this thesis (Section 3.3.1).

4. Hybrid Methods: The idea is to partly include the exact exchange energy calculated from HF to the exchange-correlation energies from DFT methods. Such hybrid methods can be more accurate, especially for non-uniform density systems such as molecules. The most commonly used hybrid method is B3LYP.

Empirical corrections. A very efficient, empirical correction for the missing dispersion or vdW interaction in DFT calculations are the D1/2/3 correction terms [38]. They all add a Lennard-Jones potential in slightly varying ways to the total DFT energy:

$$E_{\text{DFT-D3}} = E_{\mathrm{DFT}} - E_{\mathrm{disp}} \ . \tag{2.21}$$

The latest developed D3 dispersion correction term consists of a two- and three-body correction term [38]:

$$E_{\text{disp}} = E^{(2)} + E^{(3)} \tag{2.22}$$

where $E^{(2)}$ is the major contributing atom pair term of order n with isotropic dispersion coefficients C_n^{AB}, scaling factors s_n and damping function $f_{d,n}$:

$$E^{(2)} = \sum_{AB} \sum_{n=6,8,10,\ldots} s_n \frac{C_n^{AB}}{r_{AB}^n} f_{d,n}(r_{AB}) \ . \tag{2.23}$$

The three-body term $E^{(3)}$ incorporates a geometry dependend C_{6ij} term according to the coordination number for:

$$E^{(3)} = \sum_{ABC} f_{d,3}(r'_{ABC}) E^{ABC} \ , \tag{2.24}$$

where E^{ABC} is the triple dipole term derived from third-order pertubation theory. These dispersion correction terms considerably increase the accuracy of DFT methods for longe range correlations of charge distributions at virtually no added computing cost.

2.1.4 Basis sets

In essentially all *ab initio* methods, the wavefunctions used to describe molecular orbitals are approximated by a linear combination of a finite number of basis functions, which together compose a basis set:

$$\phi_i = \sum_{\alpha} C_{i\alpha} \Psi_{\alpha}. \tag{2.25}$$

All of the described methods used to determine the electronic structure of a system share the very high demand of computing power. A major impact is the choice of function type and number of basis functions used in a basis set, which in turn is deployed in solving the Kohn-Sham equations (Equation (2.16)).

There are two main philosophies used to describe molecular orbitals. The first is to think of molecules as an assembly of distorted atoms and hence construct the whole electronic configuration as a basis of many atomic orbitals. The second is to consider molecules as an assembly of atoms that distort free electrons and thus use plane waves to describe the overall configuration.

Atomic orbital basis The basis functions use the respective nuclear positions $\mathbf{R}$ as the origin. Even with few basis functions, good results to describe atoms in a chemically accurate fashion can be achieved. Especially the inner electrons in spherical proximity of the nuclei (Muffin-tin region), which are not subject to bonds, are well described. In contrast, basis functions describing the outer electrons

of different, nearby atoms can overlap, leading to the so-called basis set superposition error (BSSE). Therefore, the interstitial space between atoms or delocalized electrons as found in periodic systems cannot be described well. Individual orbitals are constructed by a number of non-orthogonal functions following the LCAO approach (Equation (2.13)), which can include basis functions to represent polarized or diffuse orbitals. Due to their many parameters, they are easily tunable. A single basis function has the form

$$\varphi_\zeta(\mathbf{r}; R_l) = \phi_\zeta(r) Y_{l,m}(\Sigma, \phi). \tag{2.26}$$

$Y_{l,m}$ are the spherical harmonic functions, and $\varphi_\zeta(r)$ can be a Slater type orbital (STO, compare Equation (2.14))

$$\varphi_\zeta(r) = e^{-\zeta r} \tag{2.27}$$

or the Gaussian type orbital (GTO, compare Equation (2.15)) typically used for molecules

$$\varphi_\zeta(r) = e^{-\zeta r^2}. \tag{2.28}$$

Plane wave basis Here, the basis functions are of orthogonal and periodic nature and are independent of nuclear positions. Therefore they do not exibit BSSE and hence are well suited to describe delocalized electrons in systems with periodic boundary conditions such as found in continuum materials. Fast-Fourier-transformations (FFT) can be exploited to make some integral evaluations in reciprocal space less computationally intensive. However, they are ill-suited to describe the very localized wave functions of inner electrons which exibit high density gradients. Plane-waves are typically augmented by an 'effective core potential' or pseudopotential, such that only valence electrons remain to be described by plane-waves. Yet, a large number of basis functions is needed for an accurate description of the electron density. A single plane-wave basis function has the form

$$\varphi_\alpha(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{(i\mathbf{G}_\alpha \cdot \mathbf{r})}, \tag{2.29}$$

with $\mathbf{G}$ being a vector of the reciprocal lattice.

Since using an atomic orbital basis corresponds more closely to the chemistry of atoms compared to a plane-wave basis, an atomic orbital basis set was chosen for the study on small organic molecules presented in this thesis (Section 3.3.1).

2.1.5 Solvation Model

Quantum-mechanical calculations are typically performed *in vacuo* and therefore solvent effects are not intrinsically accounted for. Since they can have a large effect on the electronic configuration, several methods to account for solvent effects exist.

They can be divided in two main cateogries: describing solvent molecules explicitly or implicitly. Since explicitly describing enough solvent molecules around the compound of interest on a quantum-mechanical level is beyond currently available computing power, implicit solvation models prevail. The effects of solvation contain short-range effects such as hydrogen bonding and long-range electrostatic effects such as screening of charges. The long-range effects are typically modelled via the Self-Consistent Reaction Field (SCRF) approach: a uniform dielectric medium that is polarizable. The compound of interest (or solute) M is placed into a suitable shaped cavity C (Figure 2.1), whose creation causes an energy penalty. The added dispersion (van der Waals) interaction and polarization of the dielectric medium by the solute, on the other hand, causes an energy gain.

$$\Delta \mathbf{G}_{\text{solvation}} = \Delta \mathbf{G}_{\text{cavity}} + \Delta \mathbf{G}_{\text{dispersion}} + \Delta \mathbf{G}_{\text{electrostatic}} \tag{2.30}$$

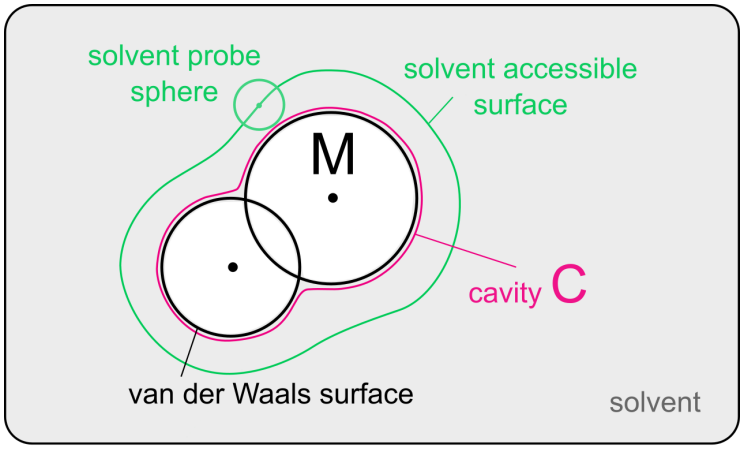

Figure 2.1: Solvation models. For the reaction field model, a solute M is placed into a suitable shaped cavity C inside a uniformly polarized solvent of permittivity ε. The shape of the cavity can be defined by van der Waals spheres. The solvent accessible surface (SAS) obtained from an additional solvent radius is typically used to calculate the dispersion term $G_{\text{dispersion}}$, whereas ∂C is used for the electrostatic term $G_{\text{electrostatic}}$.

The SCRF models differ in few aspects, such as how the size and shape of the cavity is determined (see Figure 2.1) or how the charge distribution of the solute M is defined. Several SCRF models have been developed to implicitely incorporate solvent effects. The model most widely used is the polarizable continuum model (PCM) [39, 29], which is also used for the work presented in this thesis. The basic theory behind continuum models is solving the Poisson problem:

$$-\nabla \times \left[\varepsilon(\mathbf{r})\nabla V(\mathbf{r})\right] = 4\pi\rho_{\text{M}}(\mathbf{r}) \tag{2.31}$$

where $\varepsilon(\mathbf{r})$ is the local permittivity and $\rho_{\text{M}}(\mathbf{r})$ the solutes' charge distribution contained inside the cavity C. This allows to assign ε to be homogenous inside the

cavity:

$$\varepsilon(\mathbf{r}) = \begin{cases} 1, & \mathbf{r} \in C \\ \varepsilon, & \mathbf{r} \notin C \end{cases}$$

and to define $V(\mathbf{r})$ as the sum of the solutes potential $V_{\mathrm{M}}(\mathbf{r})$ and a contribution $V_\sigma(\mathbf{r})$ from an apparent surface charge $\sigma(\mathbf{s})$ of ∂C. This now reduces the solution of finding $\varepsilon(\mathbf{r})$ of the whole solvent to only finding the apparent surface charge distribution $\sigma(\mathbf{s})$ of the solute's ∂C:

$$V_\sigma(\mathbf{r}) = \int_{\partial C} \frac{\sigma(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|} \mathrm{d}^2 s \ . \tag{2.32}$$

Subsequently, the electrostatic term of Equation (2.29) can be written as

$$G_{\text{electrostatic}} = \frac{1}{2} \int_{\partial C} \sigma(\mathbf{s}) \int_V \frac{\rho_M(\mathbf{r})}{|\mathbf{r}-\mathbf{s}|} \mathrm{d}^2 s \tag{2.33}$$

and is typically solved with the integral equation formalism [40].

The other two non-electrostatic components of Equation (2.29) are combined into a single term which is assumed to scale with the atoms' accessible surface areas S_i and the coefficients ζ_i as a measure of surface tension contribution from the same atom:

$$\Delta \mathbf{G}_{\text{cavity}} + \Delta \mathbf{G}_{\text{dispersion}} = \sum_i^{\mathrm{N}} \zeta_i S_i \tag{2.34}$$

where the coefficients ζ_i are obtained by experiments and N the total number of atoms.

2.1.6 Frontier Molecular Orbital Theory

The concept of Frontier Molecular Orbital (FMO) theory was first conceived by Kenichi Fukui [41, 42], and describes interactions between the two frontier molecular orbitals: namely the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO). Initially thought to be valid only for π electron compounds, FMO became universally accepted to consider the HOMO and the LUMO orbitals as the reactive sites of a nucleophile and electrophile reagents, respectively. The total electron density of the reactants could therefore be ignored. Key observations of the FMO theory are that the HOMO and LUMO of two interacting molecules act as the essential part and that their orbital symmetries are decisive for a chemical reaction to occur as a selection rule. HOMO-LUMO interactions can be quantified by the orbital gap $\Delta\ \mathrm{E}_{\mathrm{FMO}}$ inbetween two molecules i and j as:

$$\Delta E^{ij}_{\mathrm{FMO}} = \Delta E^{i}_{\mathrm{LUMO}} - \Delta E^{j}_{\mathrm{HOMO}} \ . \tag{2.35}$$

By swapping i and j, the lower of both gaps ΔE_{FMO} dictates the direction of charge transfer and thus defines the nucleo- and electrophile. However, due to electron delocalisation, the HOMO and LUMO orbitals are not necessarily localised on a particular site of a molecule, rendering the application of FMO theory difficult. On the other hand, FMO cannot satifactorily describe the reation of organo-metallic compounds featuring hypervalency, as in this bonding and anti-bonding orbitals interact. There are also other causes that drive reactions other than FMO interactions, such as aromatic systems where the HOMO composes the conjugated electron system and in case the reaction would cause the loss of resonance stabilization, it may not occur. Since HOMO and LUMO energies are typically calculated by the HF-method, in which the LUMO energy is generally overestimated as it is derived from a system with $n+1$ electrons and the HOMO is underestimated because of the lack of self-interaction (Section 2.1.2). Therefore ΔE_{FMO} can be greatly overestimated and should be seen as an upper boundary [43].

2.1.7 Transition State Theory

The Boltzmann distribution states the probability of finding a molecule in a given energy state by $e^{-\Delta E/k_B T}$, therefore the rate constant in which a reaction occurs can be written as

$$k = \frac{k_B T}{h} e^{\frac{-\Delta G}{k_B T}}, \tag{2.36}$$

relating the rate constant k to the Gibbs free energy difference ΔG between the transition state and reactants for a given temperature T. Furthermore, k_{B} is the Boltzmann constant and h the planck constant. The Gibbs free energy is composed of the enthalpy and entropy $G = H - TS$. In practice, experimental rates are analyzed in the form of an Arrhenius equation

$$k = A e^{\frac{-E_{\mathrm{act}}}{RT}}, \tag{2.37}$$

where the pre-exponential factor A is constant for a given chemical reaction and R the gas constant. The main interest is in comparing relative differences of rates, and therefore also relative differences in the activation energies. Theoretical methods are almost exclusively used to compare activation energies of different reactions by using the same method, but not to predict absolute rates.

Therefore, in the assumption that the enthalpy can be written as the sum of translational, rotational, vibrational and electronic terms

$$H_{\mathrm{tot}} = H_{\mathrm{trans}} + H_{\mathrm{rot}} + H_{\mathrm{vib}} + H_{\mathrm{elec}}, \tag{2.38}$$

and, by neglecting the first three, the activation energy is set as the enthalpy difference.

$$E_{\mathrm{act}} = H_{\mathrm{elec}}^{\mathrm{TS}}. \tag{2.39}$$

The activation energy E_{act} is the height of the barrier of the rate-limiting step along the reaction coordinate (see Figure 2.2 and 3.6). Calculating the barrier height requires knowledge of the Potential Energy Surface, such that the smallest barrier is known. Those smallest barriers are called Transition States (TS) and are mathematically defined as first order saddle points as shown in Figure 2.2.

Figure 2.2: Potential energy landscape along two arbitrary reaction coordinates. Originating from a local minima of the reactants conformations, the red lines depict two possible reaction pathways over two different transition states A or B ending in too different products A and B. Transition states A and B are first order saddle points, yielding only one imaginary vibrational frequency in the harmonic approximation. The local maximum is a second order saddle point. The illustration was rendered using Matplotlib [44].

Vibrational Frequency Analysis To verify a potential transition state, a vibrational frequency analysis can be carried out on the stationary point investigated. In the most simple assumption, molecular vibrations can be approximated by a harmonic oscillator. By expanding the energy $E(\mathbf{R})$ into a Taylor series around the stationary point and only keeping the first term that is non-zero yields:

$$E(\Delta\mathbf{R}) \approx \frac{1}{2}\frac{\mathrm{d}^2E}{\mathrm{d}R^2}\Delta\mathbf{R}^2 = \frac{1}{2}\Delta\mathbf{R}^2 \tag{2.40}$$

where k is the force constant. The vibrational frequency η for a one-dimensional harmonic oscillator can be derived as

$$\eta = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}, \tag{2.41}$$

where μ is the reduced mass. For a molecule consisting of N atoms, k is substituted by a $3N \times 3N$ Hessian, containing the second derivatives analogous to Equation (2.40). A transition state is confirmed if by mass weighting, transforming to vibrational normal coordinates and diagonalization, the resulting matrix con-

tains one and only one negative Eigenvalue. Correspondingly, there will be only one imaginary frequency.

In this thesis, transition state structures and energies E_{act} were calculated using two-dimensional scans along relevant reaction coordinates, namely atom distances and/or angles (Section 3.3.1). This allowed to compare to experimental rates for validation, but also to determine strain and interaction energies (Section 2.1.8).

2.1.8 Activation Strain Model

Originally developed by Bickelhaupt *et al.* [45], the activation strain model (ASM) or also commonly termed interaction/distortion energy model introduces the idea of 'activation strain', which is the energy difference within each reactant between the relaxed conformation and the deformed transition state conformation (see Figure 2.2). Together with the 'transition state interaction' between the deformed reactants in the transition state, strain and interaction energies compose the total activation energy:

$$\Delta E = \Delta E_{\text{strain}} + \Delta E_{\text{int}} \ . \tag{2.42}$$

The strain or distortion energy ΔE_{strain} is determined by the stiffness of the reactants towards deformation and is the energy required to bring the individual minimized (or optimized) reactants i into their transition state conformations:

$$\Delta E_{\text{strain}} = \sum_i E_i^{\text{TS}} - E_i^{\text{opt}} \ . \tag{2.43}$$

The interaction energy is the difference between the sum of all reactants' individual energies in the TS conformation and the transition state complex:

$$\Delta E_{\text{int}} = E_{\text{TS}} - \sum_i E_i^{\text{TS}} \tag{2.44}$$

and originates from the electronic structures. It encompasses Pauli repulsion between occupied orbitals, polarization and charge transfers and longer range electrostatic interactions.

The decomposition of the activation barrier into the destabilizing strain energy penalty and the stabilizing interaction energy gain (Figure 2.3) helps to understand in detail how competing mechanisms affect reaction efficiencies. For example, a high distortion of the reactants towards the transition state can be offset by a high interaction energy, resulting in a similar activation energy as reactions with small distortion and interaction energies. However, the interaction energy could be significantly changed by a different solvent without affecting the distortion energy much. Furthermore, the different reactants contributions can be isolated to gain useful insights in how to tune reactions to achieve the desired properties. For practical reasons it has to be mentioned that the interaction energy ΔE_{int} is very prone to basis-set superposition errors.

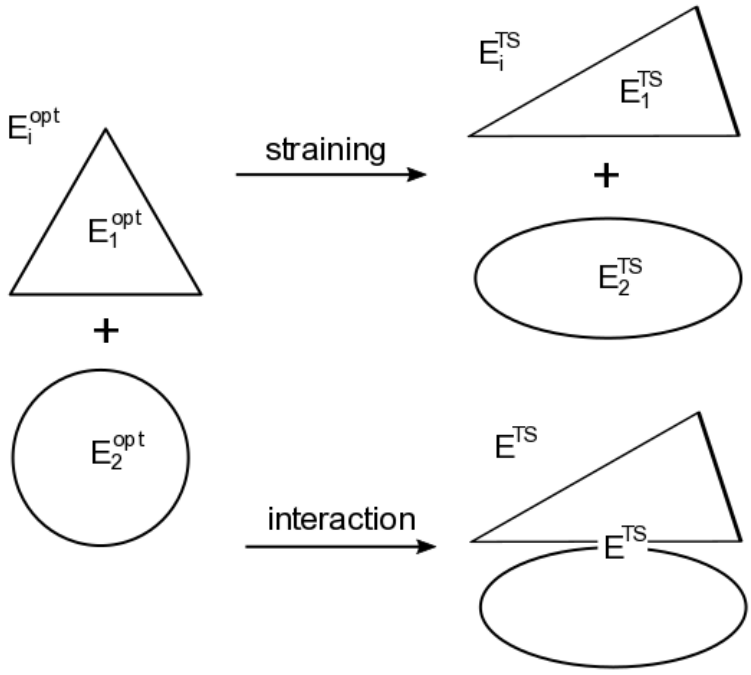

Figure 2.3: The activation strain model is used to decompose the contributions to the activation energy from activation strain between the reactants relaxed and transition state (Eq. (2.43)) and the reactants interaction energy (Eq. (2.44)).

2.2 Molecular Dynamics

Molecular dynamics (MD) simulations are an invaluable tool used to sample the dynamical behavior of molecules at an atomistic level of detail. This section outlines the most basic concepts of MD. This approach was also applied in the work that led to the publication "Dynamic prestress in a globular protein" [6], which is not included further in this thesis. It also is the method used on intermediate length scales between the quantum mechanics based methods for small molecules covered in the previous section (2.1) and the continuum mechanics based description of large systems or bulk materials covered in the following section (2.3). To bridge length and time scales, parameters or structural data can be passed between the levels. In this regard, MD simulations were pivotal for studying silk fiber mechanics, as they delivered mechanical parameters as required for the continuum mechanics description [46, 47, 48].

In case one wants to obtain a detailed describtion of a small region with a lower level method embedded in a large system described on a higher level, coupling schemes exist that link QM with MD systems, and also MD with FEM systems. A more detailed description of the principles of MD simulations can be found in various books [49, 50] or in other relevant literature like the Gromacs user manual [51].

2.2.1 Equations of Motion

MD simulations are a particle based computational method to study the time-evolution of molecular systems. The force acting on each atom of the system as a function of the atom's position and velocities can be calculated as:

$$\mathbf{F}_i(\mathbf{r}_i) = -\frac{\partial V}{\partial \mathbf{r}_i}. \tag{2.45}$$

Given initial coordinates and velocities, the Newton equations of motion

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = \frac{\mathbf{F}_i(\mathbf{r}_i)}{m_i} \tag{2.46}$$

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{v}_i \tag{2.47}$$

can therefore be numerically solved for each particle $i = 1, .., N$ composing the system and updated coordinates and velocities for the next time-step are obtained.

2.2.2 Force Fields

Atoms are categorized into different types and these categories are parametrized in their local environment. The set of parameters and relations used to describe the interactions of different types of atoms is called a force field. Due to the static local environment in which bonded atoms occur, partial charges are assigned to the atoms to approximately resemble the dipolar character of a molecule. They implicitly

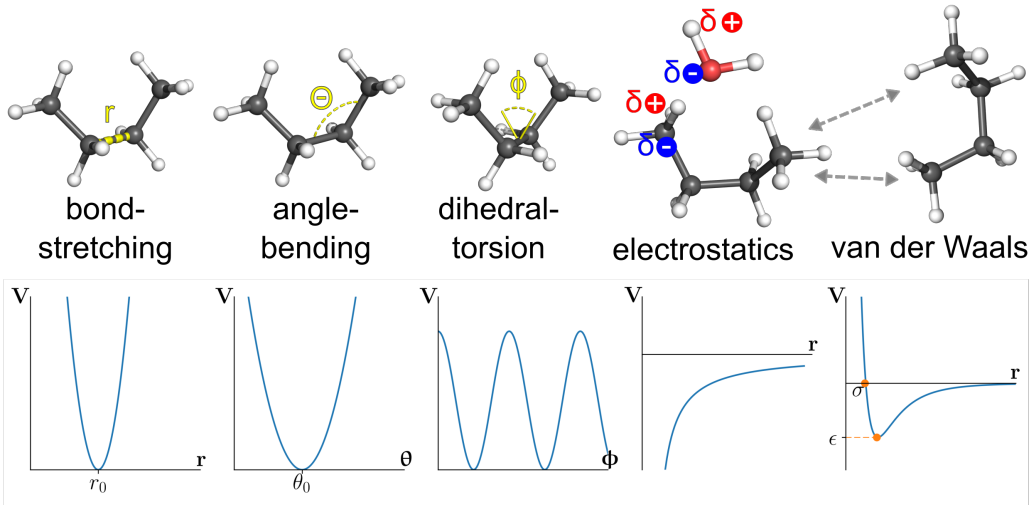

Figure 2.4: Bonded and non-bonded interaction types and shape of the potential functions in canonical force fields. From left to right: the bonded interaction potentials for bond-stretching, angle-bending and dihedral-torsion, Eq. (2.49), and the non-bonded Couloumb and Lennard-Jones potentials, Eq. (2.52).

describe a constant polarization on a bonded set of atoms. Other parameters are the Lennard-Jones parameters (ε defining the depth of the potential's well and σ defining the inter-atomic distance at which the energy is the lowest), the spring-constants k of the interaction potentials representing the stiffness of the length of bonds and the angles for bending and torsion (Figure 2.4), equilibrium values for bond lengths, angles and dihedrals, and lastly their mass.

The interaction energies of each pair of particles are described by a set of potentials, resembling the underlying physical and chemical principles described above. The potentials are divided into two main categories, namely bonded and non-bonded:

$$V(\mathbf{r}) = V_{\text{bonded}}(\mathbf{r}) + V_{\text{nonbonded}}(\mathbf{r}). \tag{2.48}$$

The bonded potentials describe interactions originating from covalently bonded atoms, and can be well approximated by just three terms for 2, 3 and 4-body interactions (Figure 2.4). The explicit bonded terms are bond-stretching (V_{distance}), angle-bending (V_{angular}), proper (V_{dihedral}) and improper (not shown) dihedrals:

$$V_{\text{distance}}(r_{ij}) = \frac{1}{2}k_{ij}^{\text{dist}}(r_{ij} - r_{ij}^0)^2 \tag{2.49}$$

$$V_{\text{angular}}(\Theta_{ijk}) = \frac{1}{2}k_{ijk}^{\text{angle}}(\Theta_{ijk} - \Theta_{ijk}^0)^2 \tag{2.50}$$

$$V_{\text{dihedral}}(\phi_{ijkl}) = k_{ijkl}^{\text{dihed}}[1 + \cos(n\phi_{ijkl} - \phi_{ijkl}^0)] \tag{2.51}$$

The equilibrium bond-lengths r_{ij}^0, angles Θ_{ijk}^0 and dihedrals ϕ_{ijkl}^0 as well as spring constants for each atom k_{ij}^{dist}, k_{ijk}^{angle} and k_{ijkl}^{dihed} (see Figure 2.4) are typically obtained from structural data, other experimental data or quantum mechanical calculations.

The non-bonded interactions have an electrostatic term (V_{Coulomb}) and a Lennard-Jones term ($V_{\text{L-J}}$) which combines short-range Pauli repulsion and long-range Van der Waals or dispersion attraction. Both typically use a cut-off between 1.0 – 1.2 nm to avoid excessive computations for negligible longer-range interaction energies. However, as coulombic interactions only decay with $1/r_{ij}$, approximations for the long-range components are commonly used. The respective non-bonded terms are:

$$V_{\text{Coulomb}}(r_{ij}) = \frac{1}{4\pi\varepsilon_0\varepsilon_r}\frac{q_i q_j}{r_{ij}} \tag{2.52}$$

$$V_{\text{L-J}}(\mathbf{r}_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{2.53}$$

with $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Once the individual forces are obtained, all forces acting on the given atom i exerted by another atom j are summed up

$$\mathbf{F}_i = \sum_j \mathbf{F}_{ij} \tag{2.54}$$

and placed into equation (2.46). The time-integration is typically calculated in the iterative leap-frog or similar Verlet scheme [52], where velocities and positions are alternatively updated in a leap-frog manner. This requires the knowledge of the positions $\mathbf{r}_i$ at time t and velocities $\mathbf{v}_i$ at time $t - \frac{\Delta t}{2}$:

$$\mathbf{v}\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}\left(t - \frac{\Delta t}{2}\right) + \frac{\mathbf{F}(t)}{m}\Delta t \tag{2.55}$$

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}\left(t + \frac{\Delta t}{2}\right)\Delta t. \tag{2.56}$$

More accurate force fields exist that better resemble bonded potentials and non-bonded interactions such as hydrophobic effects at the expense of performance and/or versatility. The two main limitations of canonical force fields are polarizability and reactivity. Induced polarizablility can be accounted for by the Drude oscillator model, introducing charged virtual sites attached to polarizable atoms by a harmonic potential [53]. Reactive force fields such as ReaxFF [54] use relationships between bond distance and order and additionally bond order and energy to deploy real Morse potentials and allow proper forming and breaking of bonds.

2.2.3 Simulation Conditions

High frequency vibrations of covalent bonds, particularly for the light hydrogen atoms on short time-scales cannot be captured accurately by classical mechanics and,

in most cases, are not essential to the understanding of the behavior of molecules on longer time-scales. One ansatz to handle this, is to constrain bonds with hydrogen atoms or even between heavy atoms [55]. Constraints on the bond-distances are introduced with algorithms such as LINCS [56]. Despite a moderate increase in computational complexity, LINCS allows to use a larger integration time step Δt of $\geq$ 0.2fs, and therefore save many iterative steps to arrive at the same total simulation time. To simulate under proper thermodynamic conditions, correct temperature and pressure ensembles can be imposed by coupling mechanisms to an external bath [57, 58]. Since the number of particles is always constant, the ensembles can be isothermal-isobaric (NPT) or canonical isochoric-isothermic (NVT) or when no coupling is used, constant volume and energy (NVE). Furthermore, periodic boundary conditions are applied to avoid finite size effects. By means of steered molecular dynamics simulations, external forces or bias potentials can be applied to drive a system in conditions that are otherwise not accessible, or not accessible in a timely manner.

2.2.4 Applications and Limitations

MD simulations as a tool are considered a computational microscope, opening access to a large variety of observables that are otherwise almost unreachable or not accessible. They are often used in conjunction with experimental methods such as X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy and widely used to study the behavior of biomolecules or polymers in the nano- to millisecond timescale [59] and from few amino-acid peptides to greater than 10^8 particles [60]. On the lower end of the accessible length and time scales, classical mechanics MD methods cannot be used to describe quantum effects such as chemical reactions or charge transfer processes. Hybrid quantum mechanics/molecular mechanics (QM/MM) schemes [61] exist and embed a QM region into a classical force-field based molecular system. The aim of QM/MM methods is to study chemical reactions within a large molecule (*e.g.*, protein) matrix that cannot be treated on a QM level as a whole due to the current limitations of available computing power and algorithms. Another way to expand the scope of MD simulations is coarse-graining: namely combining multiple interaction sites into one coarse-grain interaction site [62]. By having fewer particles, a less-rugged potential energy-landscape allowing larger integration time-steps, one gains access to greater length and time scales, however this is at the expense of detail and accuracy. Even greater scales fall well into the regime of continuum-mechanics based methods, such as those described in the following section (2.3). However, coupling schemes exist that combine particle based MD methods with continuum-based FEM methods [63], such that both, atomistic effects and long range interactions can be incoorporated into a multi-scale simulation.

2.3 Continuum Mechanics

The study of objects on the mesoscopic and macroscopic scale is out of reach for particle based models such as MD described in the previous section. A continuum mechanics based approach can be applied to study large scale effects such as stress distributions of objects sharing the same physical properties. Atomistic details are disregarded, however, collective properties of atomistic simulations such as molecular friction or rupture forces are used to parametrize continuum mechanics properties such as viscosity and yield strength. The following sections describe the necessary continuum mechanics concepts based on the book 'Continuum mechanics and theory of materials' [64], the LS-DYNA theory manual [65] and the work of Patil [66], as well as material models and the basics of the finite element method (FEM), which we deployed for the study of structural and mechanical properties of silk fibers.

2.3.1 Kinematics

The time-dependent motion of a body $\mathcal{B}$ can be expressed by a mapping function χ in which a point $\mathbf{X}_\alpha$ in the reference configuration of $\mathcal{B}$ moves to a new point $\mathbf{x}_i$ in ⌊ within the same coordinate system (Figure 2.5). In the Lagrangian description the new coordinates can be described in the reference coordinate system:

$$\mathbf{x} = \chi(\mathbf{X}, t) \tag{2.57}$$

The displacement vector $\mathbf{u}$, associating the particles reference and new position in the materials coordinate system:

$$\mathbf{u}(\mathbf{x}_0, t) = \mathbf{x}(\mathbf{x}_0, t) - \mathbf{x}_0 \tag{2.58}$$

can be a translation or deformation of the continuum body $\mathcal{B}$. The deformation gradient can than be expressed as

$$\mathbf{F}(\mathbf{x}_0, t) = \mathrm{GRAD}(\chi(\mathbf{x}_0, t)) = \frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} \tag{2.59}$$

where $\mathrm{GRAD} = \sum_i \frac{\partial \phi}{\partial x_i}\mathbf{e}_i$ and the Jacobian matrix can be written as $\mathrm{J} = \det \mathbf{F}$, where the requirement of $\mathrm{J} \neq 0$ allows the inversion of χ. The deformation gradient is subsequently used to construct the deformation tensors. The polar decomposition theorem

$$\mathbf{F} = \mathbf{R} \cdot \mathbf{U} = \mathbf{V} \cdot \mathbf{R} \tag{2.60}$$

splits the deformation into an orthogonal rotation tensor $\mathbf{R}$ and a symmetric positiv definite right or left stretch tensor $\mathbf{U}$ or $\mathbf{V}$.

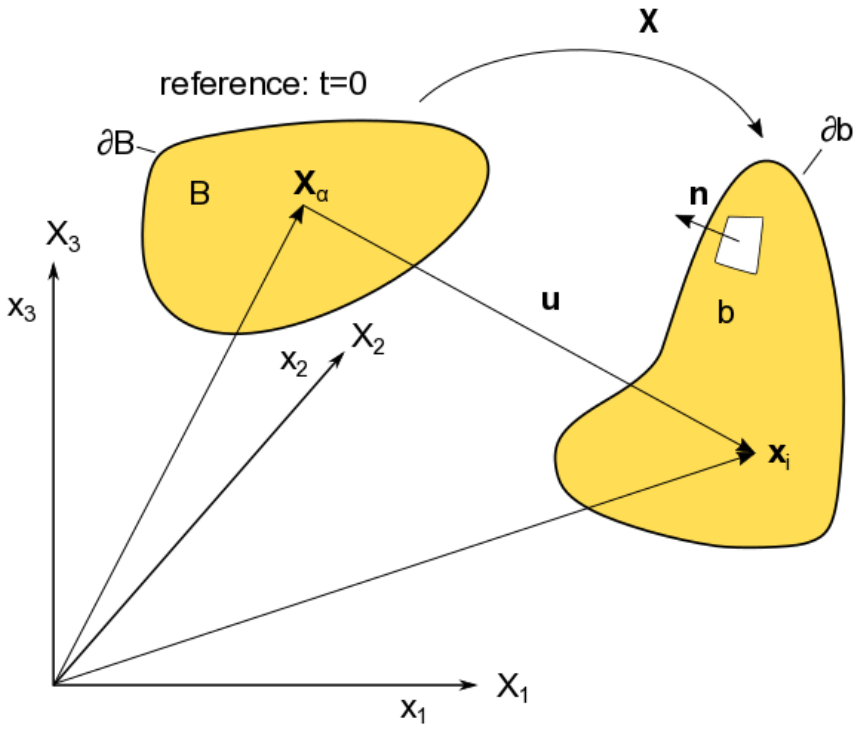

Figure 2.5: Translation and deformation of a point with position $\mathbf{X}_\alpha$ in body $\mathcal{B}$ at $t = 0$ with motion function χ and displacement vector $\mathbf{u}$ to point $\mathbf{x}_i$ in b in Lagrange notation.

2.3.2 Stress Tensor and Balance Laws

A correct description of physical quantities can only be obtained if the balance relations are obeyed, namely the conservation of mass, linear and angular momentum and energy. Since the work presented in this thesis neglects temperature effects, the balance law for energy is not further elaborated.

Cauchy Stress Tensor. Essential for the formulation of the balance laws is the definition of a stress tensor which relates the internal stress in a body to external loads. This is a similar concept to Newtons second law $\mathbf{F} = m\mathbf{a}$ applied to continuum mechanics, where the simplified picture of a force acting on the center of mass does not apply. If we take a infinite small volume from a body (Figure 2.6) and define an interface with normal vector $\mathbf{n}$, Cauchy's fundamental theorem states that the relationship between the stress vector $\mathbf{t}$ and the normal vector $\mathbf{n}$ is linear:

$$\mathbf{t} = \sigma^{\mathrm{T}} \cdot \mathbf{n} \tag{2.61}$$

such that the Cauchy stress tensor transforms the normal vector $\mathbf{n}$ of an intersection into the stress vector $\mathbf{t}$, which has the unit of force per area.

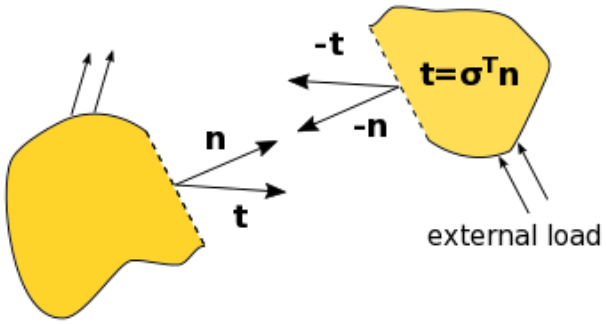

Figure 2.6: Stresses at the interface of an intersected body $\mathcal{B}$ with normal vectors $\mathbf{n}$ and stress vectors $\mathbf{t}$ resulting from the stress tensor σ caused by the applied external load.

Balance of Mass. Assuming that no sources of mass exists and $\rho(\mathbf{x}, t)$ is the density of body, its mass is defined as

$$m = \int_v \rho \mathrm{d}v. \tag{2.62}$$

The balance of mass states that mass is conserved, *i.e.* does not change over time:

$$\frac{d}{dt} \int_v \rho \mathrm{d}v = 0. \tag{2.63}$$

In the Eulerian local form the balance of mass can be written as

$$\frac{\partial}{\partial t}\rho + \mathrm{div}(\rho \dot{\mathbf{x}}) = \dot{\rho} + \rho \mathrm{div}(\dot{\mathbf{x}}) = 0 \tag{2.64}$$

Balance of Momentum. The momentum of a body is defined as

$$\mathbf{p} = \int_v \rho \dot{\mathbf{x}} \mathrm{d}v \tag{2.65}$$

and the balance of momentum states that the change of momentum of a body is equal to all forces acting on it:

$$\rho \frac{\mathrm{d}}{\mathrm{d}t} \int_v \rho \dot{\mathbf{x}} = \int_v \rho \mathbf{k} \mathrm{d}v + \int_a \mathbf{t} \mathrm{d}a \tag{2.66}$$

where a is the surface of volume v. This relation is also know as *Cauchy's first law of motion*. The local form reads

$$\rho \frac{\partial}{\partial t} \dot{\mathbf{x}} + \rho \ \mathrm{grad}(\dot{\mathbf{x}}) = \rho \mathbf{k} + \mathrm{div}\sigma \tag{2.67}$$

where σ is the stress tensor from Equation (2.61).

Balance of Angular Momentum. The balance of angular momentum or also termed *Cauchy's second law of motion* states that the change in momentum of a body equals the external applied angular momentum on the volume or surface:

$$\rho \frac{\mathrm{d}}{\mathrm{d}t} \int_v (\mathbf{x} - \mathbf{c}) \times \rho \dot{\mathbf{x}} = \int_v (\mathbf{x} - \mathbf{c}) \times \rho \mathbf{k} + \int_a (\mathbf{x} - \mathbf{c}) \times \mathbf{t} \mathrm{d}a \tag{2.68}$$

where $\mathbf{c}$ is an arbirtrary fixed vector. The local form is given by

$$\mathbf{x} \times \rho \frac{\partial}{\partial t} \dot{\mathbf{x}} + \rho \ \mathrm{grad}(\dot{\mathbf{x}}) = \mathbf{x} \times \rho \mathbf{k} + \mathrm{div} \sigma \tag{2.69}$$

Inserting Equation (2.69) into Equation (2.67) reduces the local form to just a symmetry requirement on the Cauchy stress tensor

$$\sigma = \sigma^{\mathrm{T}}. \tag{2.70}$$

2.3.3 Material Models

The material models used for simulating the silk fiber model consist of a viscoelastic model for the amorphous phase and an elastoplastic model for the crystalline phase. They state the *constitutive laws* by relating the strains and stresses for the different material models. Both models are available in the FEM solver LS-DYNA[67] and are outlined below.

Viscoelastic Material Model. Viscoelastic materials exibit a strain rate dependent stress response. In the work carried out for this thesis, simple linear viscoelasticity is assumed by an elastic and viscous component:

$$\sigma_{\mathrm{e}} = E\varepsilon \ , \tag{2.71}$$

$$\sigma_{\mathrm{v}} = \eta\dot{\varepsilon} \ , \tag{2.72}$$

where the stress σ and strain ε are related via the materials elastic modulus E and η is the coefficient of viscosity. Both components can be combined in an arbitrary fashion to a total stress σ_{tot} such as a simple addition:

$$\sigma_{\mathrm{tot}} = c_1 \cdot \sigma_{\mathrm{e}} + c_2 \cdot \sigma_{\mathrm{v}} \tag{2.73}$$

The viscous components' relaxation time constant τ is defined as

$$\tau = \eta / E \ . \tag{2.74}$$

The viscoelastic material model implemented in LS-DYNA used for describing the amorphous phase of the silk fibers is listed as Material Type 006 [67]. The model is used to simulate solids with a shear relaxation behaviour described by [68]:

$$G(t) = G_\infty + (G_0 - G_\infty) e^{-\beta t} \tag{2.75}$$

where G_0 is the short-time shear relaxation modulus, G_∞ the long-time shear relaxation modulus and β the decay constant, all as a function of temperature. The stress response under strain is formulated by the Jaumann rate [69]:

$$\sigma'_{ij} = 2 \int_0^t G(t-\tau) \frac{\partial \varepsilon'_{ij}(\tau)}{\partial \tau} d\tau \tag{2.76}$$

where σ'_{ij} is the stress rate and $\frac{\partial \varepsilon'_{ij}(\tau)}{\partial \tau}$ the strain rate. Furthermore the material density ρ and the elastic bulk modulus K

$$K = -V \frac{\mathrm{d}P}{\mathrm{d}V} \tag{2.77}$$

is required for a complete description. The elastic bulk modulus defines the compressibility of a material, relating the change in volume V to a change in applied pressure p.

Elastoplastic Material Model. The elastoplastic material model implemented in LS-DYNA is listed as Material Type 003 [67], and is used to simulate solids with isotropic or kinematic hardening plasticity or a combination of both. In the elastic regime with stress σ below the yield stress σ_y, the material is described by Hooke's law (Equation (2.71)), whereas above the yield stress, a plastic response with plastic modulus H is present:

$$\mathrm{d}\sigma = H \mathrm{d}\varepsilon^p \tag{2.78}$$

from the plastic strain component ε^p with $\varepsilon = \varepsilon^e + \varepsilon^p$.

Furthermore, dynamic rate effects can be included by using the Cowper and Symonds model [67], which scales the static yield stress σ_{sy} with the Cowper-Symonds relation to obtain a dynamic yield stress σ_{dy}:

$$\sigma_{dy} = \sigma_{sy} \cdot \left(1 + \frac{\varepsilon}{C}^{\frac{1}{P}}\right) \tag{2.79}$$

with C and P being strain rate parameters and ε the strain rate.

2.3.4 Finite Element Method

Most physical processes can be formulated via partial differential equations (PDE). The partitioning of a problem into smaller, simpler parts allows the numerical solution to the boundary value problem for PDE's by local approximations on the elements' node points. The finite element method consists of several parts: the discretization of a problem including the boundary conditions, the weak formulation of the PDE and time integration. A very detailed description of the finite-element method can be found in the books of Hutton, Zienkiewicz and Logan [70, 71, 72] or the very comprehensive LS-DYNA theory manual [65], which this section is based on.

The general steps of FEM can be summarized as follows:

- Discretization of the domain (meshing) and selection of element types (Paragraph 2.3.4).
- Approximation of the displacement field (Figure 2.5). A continuous function throughout the whole body is approximated by discretized by piecewise functions in the subdomains/elements. These piecewise functions are also commonly called shape or basis functions. This step requires the weak formulation of the PDE.
- Definition of the Stress-Strain relationship. These are derived from the material models and state the constitutive laws (Section 2.3.3).
- Derivation of the elements stiffness matrix. This generally involves either work/energy based methods such as the principle of virtual work of d'Alembert, or weighted residuals methods such as the Galerkin method.
- Assembly of element equations into global equations and introduction of boundary conditions.
- Solving for the primary unknowns: displacements of the nodes and/or rotations (degrees of freedom).
- Solving for the secondary unknowns: strains and stresses of elements.
- Postprocessing: analysis and interpretation of results.

A more detailed description is provided below for the meshing, weak formulation, volume- and time integration

Meshing. The spatial discretization of a domain, *i.e.* a line, area or volume into small 'finite elements' is termed meshing. However, no optimal solution for meshing an arbitrary volume exists. Since the numerical solutions depend on the mesh's geometry, a general understanding of the different discretization methods is required to create a mesh suitable for the problem investigated. Meshing methods can be divided into two main categories, structured and unstructured. The two main element types used for meshing consist of triangular or quadrangular faces (see Figure 2.7).

Unstructured mesh algorithms are widespread due to their versatility in the two- and three dimensional domain. However, compared to structured meshs, unstructured meshs consist of more elements for a given number of nodes involved. Therefore they are computationally more expensive, and in general not well suited for regular geometries. Since the finite element boundaries do not follow the geometries' inherent boundaries (see Figure 2.7), the solutions obtained are usually

not optimal. Commonly used unstructured mesh algorithms are the Delaunay, MeshAdapt or Frontal/NETGEN algorithm [73, 74, 75, 76]. A typical meshing strategy for complicated geometries starts with creating an unstructured triangular or tetrahedral mesh and, if possible, recombining triangular or tetrahedral elements into quadrangular or hexahedral elements. Other strategies of efficently creating unstructured, yet quadriliteral meshs exist such as the Blossom-Quad method [77].

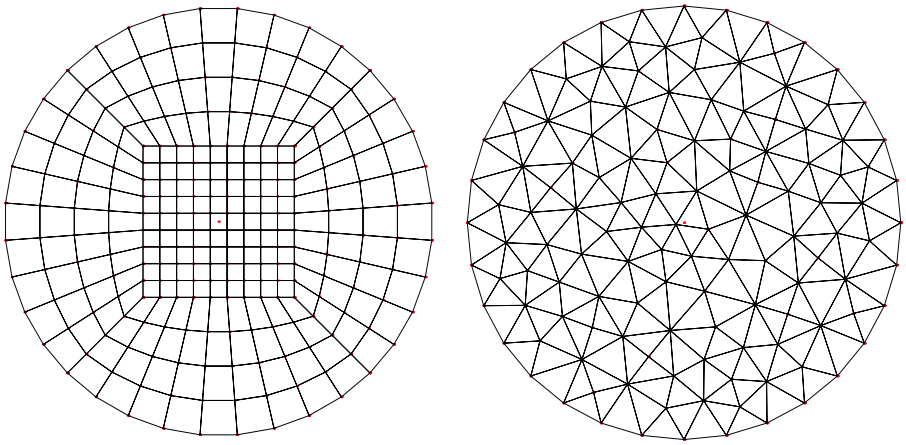

Figure 2.7: Examples for meshs of a circle. The left figure shows a structured mesh consisting of quadriliterals, the right figure shows an unstructured mesh consisting of triangles. The meshs were created and rendered using GMSH [74].

Structured mesh algorithms, on the other hand, are more efficient by mapping the geometrical domain into the computational domain, but are also prone to artefacts such as shear and volume locking (also known as hourglass effect). The main disadvantage of structured meshs, however, is that no universal algorithm exists that is able to mesh arbitrary volumes. Structured meshs are heuristically created by decomposing a complicated geometry into smaller subsets with easier to define boundaries, which in turn are manually meshed by a simple 3D extrusion of a 2D mesh, or by one of the numerous structured meshing methods such as Advancing Front, Medial Surface, superposition or grid based methods [78].

The meshing of the fiber models presented in this work consisting of a cylinder with randomly inserted cuboids was performed using the program GMSH [74], deploying the unstructured Delaunay algorithm together with the NETGEN algorithm to optimize for a more uniform distribution of angles (see Section 4.3).

In the mathematical description, the domain Ω is be approximated by a discretized domain Ω^h which in turn is composed of the union of n disjoint subdomains Ω_e:

$$\Omega \approx \Omega^h = \bigcup_e^n \Omega_e \tag{2.80}$$

Weak Formulation. The formulation of most physical problems can be expressed by PDE's, which might be difficult or impossible to obtain an analytical solution for. The weak formulation of a PDE states that it is possible to obtain an approximate, numerical solution, to an analytical solution f, where f is not required to be smooth. As a simple example can formulate the balance of momentum (Equation (2.67)) in the static case where $\dot{\mathbf{x}} = 0$:

$$0 = \rho\,\mathbf{k} + \operatorname{div}\sigma \tag{2.81}$$

The weak formulation implies that Equation (2.81) must not be well defined in all points. We can multiply both sides of Equation (2.81) by a test function $\phi \in \mathcal{H}$:

$$0 = \rho\,\mathbf{k} \cdot \phi + \operatorname{div}\sigma \cdot \phi \tag{2.82}$$

and succesively integrate over the domain Ω

$$0 = \int_\Omega \rho\,\mathbf{k} \cdot \phi \,\mathrm{d}v + \int_\Omega \operatorname{div}\sigma \cdot \phi \,\mathrm{d}v \tag{2.83}$$

and, using Green's first identity and taking into account the boundary condition on $\partial\Omega$ we obtain

$$\delta\Omega = \int_\Omega \rho\,\mathbf{k} \cdot \phi \,\mathrm{d}v + \int_\Omega \rho\,\mathbf{k} \cdot \phi \,\mathrm{d}v - \int_{\partial\Omega} \sigma\mathbf{n} \cdot \phi \,\mathrm{d}a = 0 \tag{2.84}$$

as the weak formulation of (2.81) and also states the principle of virtual work. A further requirement on the weak formulation is that Equation (2.84) has to hold for all $\phi \in \mathcal{H}$, formulating the variational principle. The test functions ϕ have a well defined norm and can easily be converted to vectors in the Hilbert space $\mathcal{H}$. The boundary condition $\delta\Omega$ can be given in the form of either a Dirichlet boundary Γ_D or a Neumann boundary Γ_N (Figure 2.5). For the Dirichlet boundary, the displacement is known (requiring ϕ) whereas for the Neumann boundary the normal stresses $\mathbf{t}$ (or derived σ) is required. In practical terms, the flow of force or stress between neighbouring elements is stated as Neumann boundary conditions on the node points.

We can then apply a finite mesh of node points on our reference configuration and express $\mathbf{x}$ as a linear combination of basis functions ψ_j:

$$\mathbf{x}(\mathbf{X}, t) = \mathbf{x}(\mathbf{X}(\xi, \eta, \zeta), t) = \sum_{j=1}^{k} \psi_j(\xi, \eta, \zeta)\,\mathbf{x}^j(t) \tag{2.85}$$

in their local, parametric coordinates ξ, η, ζ for k node points. Each basis function is required to have only a finite support. Since $\delta\Omega = 0$ has to hold for all ϕ, it also has to hold for all basis functions ψ_j. Therefore, by having Ω discretized (Equation (2.80)) we can spatially separate the integral (2.84) additively into n terms:

$$\delta\Omega = \sum_{m=1}^{n} \delta\Omega_m = 0 \tag{2.86}$$

and succesively assemble the individual element contributions into a total system of equation which can be solved numerically.

Volume integration. After having obtained an integral formulation of the governing equations, the volume integration over an element is typically carried out numerically using the Gaussian quadrature method. A function f defined over a volume v is approximated by

$$\int_v f \mathrm{d}v = \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} f \det \mathbf{J}\, \mathrm{d}\xi\, \mathrm{d}\eta\, \mathrm{d}\zeta \approx \sum_{j=1}^{n}\sum_{j=1}^{n}\sum_{j=1}^{n} f_{jkl}|J_{jkl}|w_j w_k w_l \tag{2.87}$$

where ξ, η and ζ are the local coordinates, $w_{j,k,l}$ the weight factors and n the number of integration points inside an element. For the finite element simulations carried out in this thesis, single integration point ($n = 1$) tetrahedrals have been used.

Time Integration. The time discretization used for the finite-element simulations carried out in this work was performed in an explicit manner where, the step r_{n+1} is a direct function of the previous step r_n:

$$\mathbf{r}_{n+1} = f(\mathbf{r}_n, ...) \ , \tag{2.88}$$

using the central difference scheme [65]. An equation of motion with diagonal mass matrix $\mathbf{M}$, external and body loads $\mathbf{P}$ and stress divergence vector $\mathbf{F}$ at time n in the form

$$\mathbf{M}\mathbf{a}_n = \mathbf{P}_n - \mathbf{F}_n \tag{2.89}$$

advances to time t_{n+1} by updating the nodal displacements $\mathbf{u}$, velocities $\mathbf{v}$ and accelerations $\mathbf{a}$ in the following manner:

$$\mathbf{a}_n = \mathbf{M}^{-1}(\mathbf{P}_n) \ , \tag{2.90}$$

$$\mathbf{v}_{n+1/2} = \mathbf{v}_{n-1/2} + \mathbf{a}_n \Delta t_n \ , \tag{2.91}$$

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \mathbf{v}_{n+1/2}\Delta t_{n+1/2} \ , \tag{2.92}$$

$$\Delta t_{n+1/2} = \frac{\Delta t_n + \Delta t_{n+1}}{2} \tag{2.93}$$

with initial displacements $\mathbf{u}_0$ and initial velocities $\mathbf{v}_0$. Eventually we obtain the updated nodal positions $\mathbf{r}$ by:

$$\mathbf{r}_{n+1} = \mathbf{r}_0 + \mathbf{u}_{n+1} \ . \tag{2.94}$$

2.4 Small-Angle Neutron Scattering

Due to their missing Coulomb interaction, neutrons are excellent scattering probe particles for nuclei and allow to extract structural information in a non-destructive way. They have a spin of 1/2 and belong to the group of baryons with a mean lifetime of 881.5±1.5 s [79] before decaying into an proton, electron and antineutrino:

$$n \rightarrow p^+ + e^- + \bar{\nu}_e.$$

Small-angle neutron scattering (SANS) is a widely used tool to investigate the structure of macromolecues in the length-scale of 1–100 nm. In this section we only provide a brief description of the quantities extracted from small-angle neutron scattering used for the structural investigation of spider silk-fibers presented in this thesis. A comprehensive overview of neutron scattering can be found in the textbook of Lovesey [80], or the work of Hennig [81], which this short section is derived from. The basic properties of a neutron are listed in Table 2.1.

The strength of the interaction between a neutron and a nucleus is quantified by the coherent scattering length. The large difference of the neutron scattering length between hydrogen ($-0.3741 \times 10^{-12}\,\mathrm{cm}^{-1}$) and deuterium ($0.6671 \times 10^{-12}\mathrm{cm}^{-1}$) [82] makes SANS a very powerful technique to achieve contrast between deuterated and non-deuterated parts of a molecular system.

Properties of Neutrons	
mass (m)	1.00866491597 u
electric charge (q)	0
spin (s)	1/2
magnetic moment	$-1.91304273\mu_N$
mean lifetime	881.5 s

Table 2.1: Basic properties of the neutron [79, 81].

When a neutron with wavevector $\mathbf{k}$ is scattered from a sample, the scattered wave $\mathbf{k}'$ has an amplitude of:

$$A(\mathbf{q}) = \int \rho(\mathbf{r}) e^{(\mathrm{i}\,\mathbf{q}\,\mathbf{r})\,\mathrm{d}^3\mathbf{r}}\ , \tag{2.95}$$

where ρ_b is the scattering length density and $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ is the momentum transfer vector (see Figure 4.6). In SANS experiments, the measured quantity is the flux of neutrons through the area element $d\,a$ in the plane of the impact of the solid angle element $d\mathbf{\Omega}$. This partial differential cross-section coincides with $A(\mathbf{q})$:

$$\frac{\mathrm{d}\sigma(\mathbf{q})}{\mathrm{d}\mathbf{\Omega}} = \langle A(\mathbf{q})\, A(\mathbf{q})^* \rangle\ . \tag{2.96}$$

If the scattering length density is taken as an average value $\bar{\rho}$ inside the particle

volume V, we can obtain the scattering intensity:

$$\frac{\mathrm{d}\sigma(\mathbf{q})}{\mathrm{d}\mathbf{\Omega}} = N\,V^2\,\bar{\rho}^2\,S(\mathbf{q})\,F(\mathbf{q}). \tag{2.97}$$

where the equilibrium ensemble average structure of the system composing the structure factor S is:

$$S(\mathbf{q}) = \left\langle \frac{1}{N} \sum_{j,k=1}^{N} e^{[\mathrm{i}\,\mathbf{q}\,(\mathbf{R}_j - \mathbf{R}_k)]} \right\rangle \,, \tag{2.98}$$

and the Fourier transform of the particle shape defines the form factor F:

$$F(\mathbf{q}) = \left| \frac{1}{V} \int_V e^{(\mathrm{i}\,\mathbf{q}\,\mathbf{r})\,\mathrm{d}^3\mathbf{r}} \right|^2 \,, \tag{2.99}$$

The structure factor S can also be expressed in terms of the pair-correlation function (or radial distribution function) $g(r)$:

$$S(q) = 1 + 4\,\pi\,N \int_0^\infty [g(r) - 1]\,\frac{\sin(q\,r)}{q\,r}\,r^2\,\mathrm{d}r \tag{2.100}$$

which describes the spatial arrangement of particles defined by their mutual interactions.

Chapter 3

Strain-Promoted Click-reactions

Contents

3.1 Abstract

Abstract: Site-specific labeling of biomolecules is rapidly advancing due to the discovery of novel mutually orthogonal reactions. Quantum chemistry studies have also increased our understanding of their relative rates, although these have until now been based on highly simplified reactants. Here we examine a set of strain-promoted click-type cycloaddition reactions of n-propyl azide, 3-benzyl tetrazine and 3-benzyl-6-methyl tetrazine with cyclooctenes/ynes, in which we aim to address all relevant structural details of the reactants. Our calculations have included the obligatory handles used to attach the label and biomolecule as these can critically influence the stereochemistry and electron demand of the reaction. Furthermore we investigated the influence of these handles on the diene by substituting the benzyl group with chemical moieties ranging from heavily electron withdrawing to heavily electron donating and determined how they modulate electron demand via the benzyl ring on the reactive site. We systematically computed orbital gaps, activation and strain energies using density functional theory and determined experimental rates for validation. Even though frontier molecular theory and is often considered sufficient to explain differences in the reaction rates experimentally investigated, our study points towards the contribution of other factors such as internal strain, stereo- and regioselectivity, which appear as the predominant components to explain the kinetic differences. Most importantly, the presence and isomeric configuration of handles can decisively determine the reaction speed and regioselectivity. Our findings can be directly utilized in engineering genuine cycloaddition click chemistries for biological labeling.

3.2 Introduction

Stress and strain stored in molecular systems are important to reactivity. Understanding the influence of stress on chemical reactions in detail and exploiting it the right way gives leverage to specifically steer reactivities and to design useful new reactions. One such important type of reactions that exploits stressed bonds is the class of click reactions.

3.2.1 Click Chemistry

Click chemistry is an approach introduced in 2001 by K. Barry Sharpless that essentially encompasses chemical processes which rapidly, reliably, and selectively 'click' together small units to synthesize useful new compounds and, ultimately promote new functionalities [84]. For a reaction to meet the criteria of 'click chemistry' it must be "*modular*, *wide in scope*, give *very high yields*, generate only *inoffensive by-products* and be *stereospecific* (but not necessarily enantio-selective)" [84]. Desired characteristics of such reactions are: satisfying simple reaction conditions such as insensitivity to water and oxygen, using only good solvents such as water or not needing solvents at all. Likewise it should produce stable, simple to isolate products

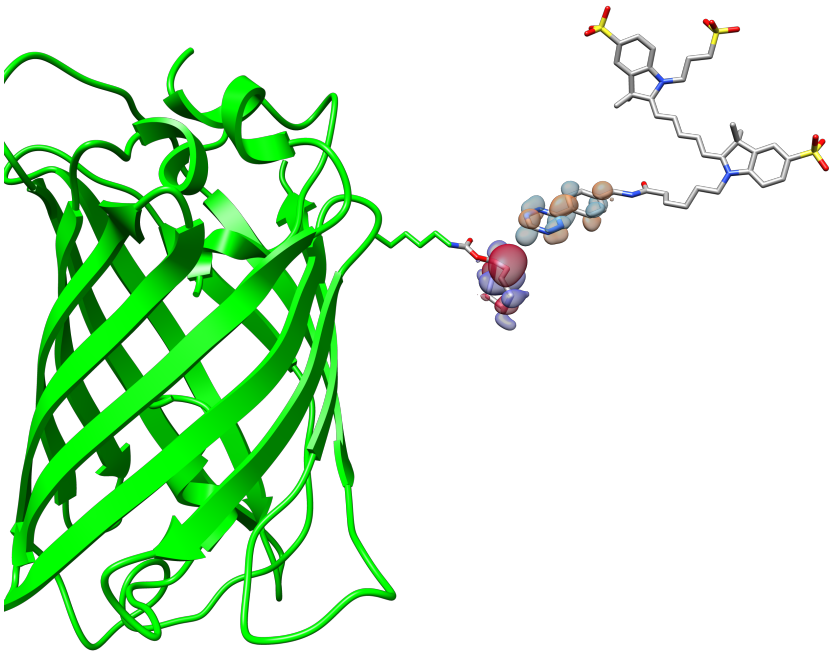

Figure 3.1: Rapid labeling via click chemistry: The Green fluorescent protein (GFP) (green, shown in cartoon representation, PDB code: 2WUR) with an unnatural amino acid (shown in green licorice) incoorporating a TCO*ax isomer reacting with a tetrazine attached to the fluorophore Cy5. TCO*ax, the tetrazine and Cy5 are shown in licorice and colored by atom type (carbon:grey, nitrogen:blue, oxygen:red, sulphur:yellow). The molecular orbitals relevant to the reaction (HOMO for the TCO*ax and LUMO+1 for the tetrazine) are shown as transparent iso-surfaces and colored red or blue wether the wave function is positive or negative. The image was rendered using the program Chimera [83].

under physiological conditions and start from readily available reagents. Moreover, the reaction should reliably work on small and large scale applications. The idea is to have only a few but "nearly perfect reactions" [84]. Reactions that obey this philosophy are typically highly exothermic ($>$ 20 kcal mol^{-1}), and therefore proceed rapidly, tend to be highly selective for a single product and are termed "spring loaded for a single trajectory" [84]. This directly implies 'stress' as a driving force, which however has not been fully quantified and explored yet.

With the aim of expanding the scope of click chemistry beyond the discovery of large numbers of new compounds, this approach is also used for multiple applications [85] ranging from drug discovery by lead-finding through combinatorial chemistry [86], to the selective labeling of biomolecules for radiochemistry [87, 88, 89, 90], the imaging of cellular components [91], or the conception of new functional materials [92, 93, 94, 95].

The high reactivity and selectivity of click reactions can be achieved in various ways, and using catalysis, e.g. by metal ions is a popular choice. However, an alternative option is to speed up a reaction by stress. The click reaction can be designed such that reactants are built into strained rings, the release of which increases reactivity. This is described in further detail in Section 3.2.3.

3.2.2 Cycloadditions and their Applications

The most prominent type of "click chemistry" reactions that fulfill the above outlined requirements are cycloadditions [96, 86].

In cycloadditions, two or more unsaturated molecules combine and form a cyclic adduct with a net reduction in bond multiplicity [97, 98]. Very often they are of pericyclic nature, by which the geometry of the transition state has a cyclic component and bonds are formed and broken through a concerted shift of electrons, omitting the formation of intermediates [99]. Key properties of cycloadditions are their high selectivity and bio-orthogonality: namely these are chemical reactions that do not perturb native biochemical processes inside living cells. Therefore they strongly serve the purpose to label biomacromolecules and to follow their activity inside cells by means of spectroscopic detection [86]. According to K. Barry Sharpless, the 'premier example' of click chemistry is the Huisgen 1,3-dipolar cycloaddition of alkynes and azides yielding triazoles [96, 100, 86, 101]. In IUPAC notation [97, 98] this corresponds to a $[_\pi 2_s +_\pi 4_s]$ cycloaddition, where [i+j] denotes the number of electrons participating in the ring formation. The preceding subscript π specifies the orbitals involved, and the appended subscript s specifies the stereochemistry as suprafacial. The high selectivity and specificity of the azide-alkyne cycloadditions allow the detection of very small quantities of reactants. Despite the slow progress of the reaction, typically in the tens of hours [102], it is often used as the labeling strategy of choice when other methods prove infeasible [91, 89]. To speed up the reaction, an *in silico* study suggested the use of metal catalysis to potentially increase the rate by up to $\sim 10^8$-fold [103], even though metal catalysts are a disadvantage for labeling biomolecules in living cells due to their toxicity. Since the application

in biology can only be meaningfully exploited if the time-scale of the reaction is comparable to the timescale of the biological process investigated, this imposes a limitation to the versatility of the Huisgen 1,3-dipolar cycloaddition. However, the rapid construction of molecules via copper(I)-catalyzation is, among others, widely used in organic synthesis or to construct dendrimers and polymers [89]. The Diels-Alder cycloaddition can occur in two manners: normal or inverse electron demand (Figure 3.2). In the latter case, the dienophil is the electron-rich reactant that donates electrons to the electron-poor diene [104, 105].

3.2.3 Strained Click Reactions

Although the use of metal catalysts has enormously improved the timescales of cycloaddition reactions, the citotoxicty of copper makes the use of ion-catalysed reactions unsuitable for biological applications and different approaches to enhance the kinetics of this reaction class have been explored. Particularly, the use of strained bonds and angles within the chemical structures of the reactants has been considered for $[_{\pi}2_{s}+_{\pi}2_{s}]$ cycloadditions (Figure 3.2), suggesting that internal strain can be considered as a strategy to substitute the adoption of metal catalysts in order to have kinetically relevant cycloadditions. This work focuses on strain-promoted cycloadditions (SPIEDAC) [106] as they constitute a class of reactions that can widen the application range of click chemistry to the investigation of biological molecules without the risk of creating cellular toxicity. One of the most prominent applications of strain-promoted cycloadditions is indeed the labeling of biological entities, such as DNA, proteins or lipid moieties, to follow their activity or interacting capabilities inside cells. The most popular labeling method for the visualization of cellular processes involves the use of the green fluorescent protein (GFP, Figure 3.1). Almost every cellular process has so far been characterized using GFP or GFP-like proteins [107]. However, GFP and related proteins are relatively large biomolecules, which can induce significant steric hindrance and thus can influence the physiological function of tagged targets. GFP-like tags are therefore poorly or not suited to tag lipids, glycans or nucleic acids. On the other hand, small-molecule probes generally have better access to intracellular molecules with less chance of interference with their natural way of functioning inside cells [108], and this is why recent advances for site-specific labelling in conjunction with super-resolution microscopy of cellular components were made possible through the use of cycloadditions [84, 109, 110, 111, 112]. Using strain promoted cycloaddition as the tool of choice for cellular imaging via molecular labeling, the study of how internal stress enhances reactivity and channels reaction kinetics into specialised ranges of applications is of enormous interest. Therefore the aim of this work is to clarify the influence of strain on the kinetics of this important and well established class of reactions.

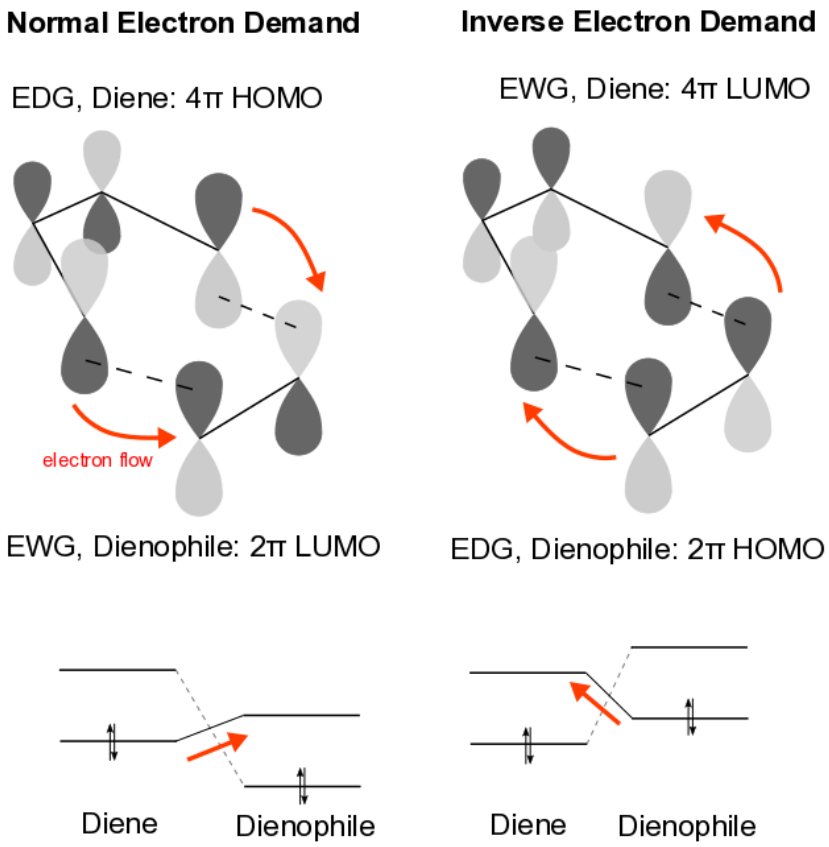

Figure 3.2: Scheme of the Diels-Alder cycloaddition in normal and inverse electron demand together with the FMO involved in the reaction. Due to no intermediate states involved, orbital symmetries (suprafacial) drive the reaction in a $[_{\pi}4_s+_{\pi}2_s]$ manner. The difference in normal and inverse electron demand is the direction of electron transfer between the diene and the dieneophile (red arrows). Normal demand implies that electrons move from the diene HOMO to the dienophile LUMO, whereas inverse demand implies electron to move from the dienophile HOMO to the diene LUMO.

The major part of this section has been published as:

Origin of Orthogonality of Strain-promoted Click Reactions

Chemistry: a European Journal **21** (2015) 35:12431-12435

Authors: J.A. Wagner[1,2], D. Mercadante[1,3], I. Nikić[4], E. Lemke[4], F. Gräter[1,3]

[1] Heidelberg Institute for Theoretical Studies,
69118 Heidelberg, Germany
[2] Institute for Theoretical Physics, Heidelberg University,
69120 Heidelberg, Germany
[3] Interdisciplinary Center for Scientific Computing,
69120 Heidelberg, Germany
[4] European Molecular Biology Laboratory,
69117 Heidelberg, Germany

Authors contribution: J. A. Wagner and E. Lemke and F. Gräter designed research. Computer simulations and data analysis were performed by J. A. Wagner. The experiments were performed by I. Nikić. The article was written by J. A. Wagner and F. Gräter. All authors contributed to the discussion.

3.3 Part I: Origin of Orthogonality

Strain-promoted cycloadditions can be highly selective or even specific and therfore allow for mutual orthogonality of several particularly slow and fast click-type reactions [113]. In this context, mutual orthogonal reactions allow the simultaneous labeling of multiple sites in a kinetically controlled fashion *in vivo* and *in vitro* [114, 115, 116].

Recently, a set of mutually orthogonal reactions has been used by Nikić *et al.* [117] all of which are promoted by molecular strain. This set has been emplyoed for the time dependent monitoring of the insulin receptor and Influenza proteins labelled through the insertion of non-canonical amino acids bearing strained eight-membered rings. The click reactions utilized in this work are n-propyl azide, 3-benzyl tetrazine (H-Tet) and 3-benzyl-6-methyl tetrazine (Me-Tet) ligations to three types of eight-member cyclic rings (Figure 3.3a and Figure 3.5, IUPAC names in Appendix 3.3.4), namely (i) racemic equatorial trans-cyclooct-2-en-1-methylcarbamate (TCO*e) and racemic axial trans-cyclooct-2-en-1-methylcarbamate (TCO*a), (ii) two enantiomers of strained cyclooct-2-yn-1-methylcarbamate (SCO), and (iii) *endo*- and *exo*-bicyclonon-5-yn-1-methylcarbamate (BCN^{endo} & BCN^{exo}). The tetrazine ligation, a [4+2]-cycloaddition, is a strain-promoted inverse electron demand Diels-Alder reaction (SPIEDAC) [118, 119], while the alkyne-azide ligation is a strain-promoted Huisgen-type 1,3-dipolar $[_{\pi}2_{s}+_{\pi}4_{s}]$-cycloaddition (SPAAC) [120, 121, 122]. The large differences in reaction speed when employing supposedly similar ligation partners, such as H-Tet versus Me-Tet, or TCO* versus SCO, render dual color labeling of biological systems possible [123, 121].

Quantitatively predicting such differences and understanding their origin by quantum chemical calculations would increase the ability to further enhance the mutual orthogonality and augment the available set of reactions for preferential labeling. A large body of previous quantum-chemical computational work has characterized and thereby significantly advanced our understanding of a number of highly related strain-promoted click reactions [124, 125, 126, 127, 128]. They mostly ascribed the differences in reactivity of these reactions, which have been shown to feature an inverse Diels-Alder electron demand, to contributions from orbital interactions, strain-distortion and Pauli repulsion [129, 130]. The higher reactivity of tetrazines towards trans-cyclooctene was attributed to their higher electrophilicity as compared to azides [125], while introducing nitrogens into the diene was proposed to decrease Pauli-repulsion, rendering tetrazines the fastest reaction partner for cycloadditions with alkenes [130].

Notwithstanding these advances, the previous studies described above have been focusing on highly simplified reactant scaffolds limiting the comparison of calculated activation energies to reaction rates determined experimentally for more complex scaffolds.

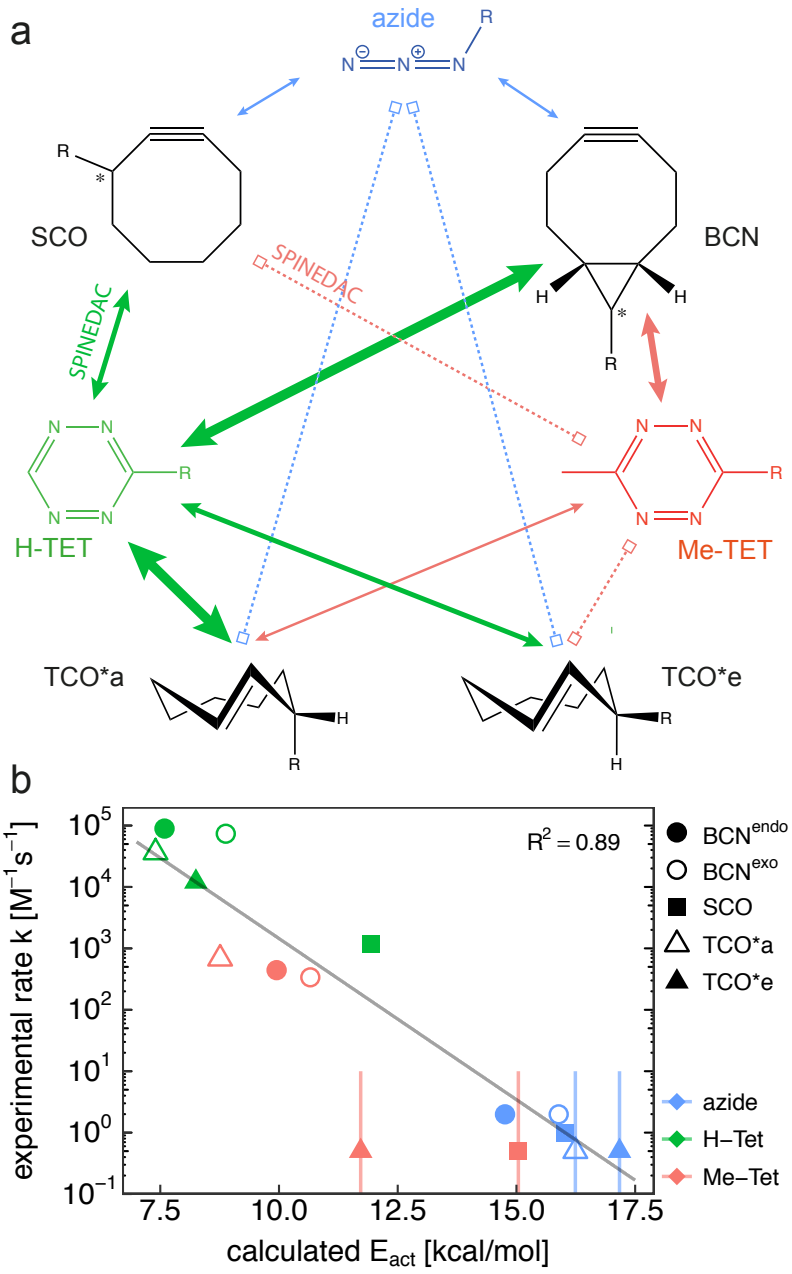

Figure 3.3: Kinetics of mutually orthogonal click reactions. a) Overview of the investigated reaction set between TCO*e/a isomers, $BCN^{endo/exo}$ isomers and SCO with azide, H-Tet and Me-Tet. Different isomers are marked as *. Differently sized arrows represent the reaction rates, where thicker means faster. Dashed lines show reactions that were too slow to be measured. Structures of all isomers are given in Figure 3.5. b) Correlation between experimental rate constants k and the calculated activation energies E_{act} for each pair of reactants. The solid line shows a linear regression. The SCO/BCN-azide rates are taken from Borrmann *et al.* [120]. For the reactions too slow to be measured, we only obtained an upper limit for k as indicated by the error lines and omitted such points for the linear fit. Experimental error bars are smaller than the symbol size and are omitted for clarity.

However, attaching handles to the functional groups of azides, tetrazines and the different eight-membered rings is the crucial requisite to label biomolecules, and may also critically affect the reactivity of the molecule, as experimental data increasingly certify that the speed of these reactions is highly susceptible to minor changes in substituents [117]. Importantly, handles on cyclooctyne rings introduce a single stereocenter whereas handles on cyclooctene rings add an additional stereocenter. This results in two SCO enantiomers and two pairs of TCO* enantiomers (see Figure 3.5). The configuration of the stereocenters emerge as a crucial determinant for this class of click reactions [131, 132]. Considering the effect of the genuine handles used for functionalization is therefore indispensible for a direct validation of insights from quantum chemical calculations and, more importantly, for the application of these findings to the design of new orthogonal reactions.

We addressed this fundamental challenge and herein present the first comprehensive set of measured rates and computed barriers for reactants with direct relevance for biological applications. We systematically computed the conformations and energies for the SPIEDAC and SPAAC reactions described above, using density functional theory (DFT, see Section 2.1.3). This resulted in a total of 48 reactions with configurationally distinct transition states and products, out of which 24 are chemically distinct enantiomers. The resulting system sizes required advanced sampling of the complex energy landscape, and the diminishing barriers of the high speed TCO* reactions rendered it necessary to incorporate van der Waals complexes into the reaction pathway. On this basis, unexpectedly, we find that SPIEDAC reactions involving SCO and tetrazines to follow normal electron demand, ascribed to the electrophilicity of the SCO. We thus term this class of reactions SPINEDAC (strain-promoted inherently normal electron demand Diels-Alder cycloaddition). We observe isomeric configurations to critically fine-tune the reactivity of the investigated click-reactions. Our study aids the engineering of a currently widely used set of cycloadditions to further enhance their reactivity and mutual orthogonality for their applicability in biology.

3.3.1 Methods

To carry out the quantum chemical calculations, DFT was used as quantum mechanical method in conjunction with the M06-2X density functional [133, 134, 135], which is classified as a Global Hybrid (GH) meta-GGA exchange correlation functional. It uses the generalized gradient approximation to the local density (Section 2.1.3), as well the second derivative of the density. In addition it incorporates a 54% Hartree-Fock (HF, see Section 2.1.2) exchange to account to some extent for longer ranging interactions and thus features a medium range dispersion interaction.

More recently developed DFT functionals show improvements in estimating ground state energies at the expense of less accurate electron densities and therefore also geometries, a category the chosen M06-2X clearly falls into [136]. However, this has only been shown for single atoms and ions. For more complicated molecules, and especially the type of reactions investigated in this work, the M06-2X functional

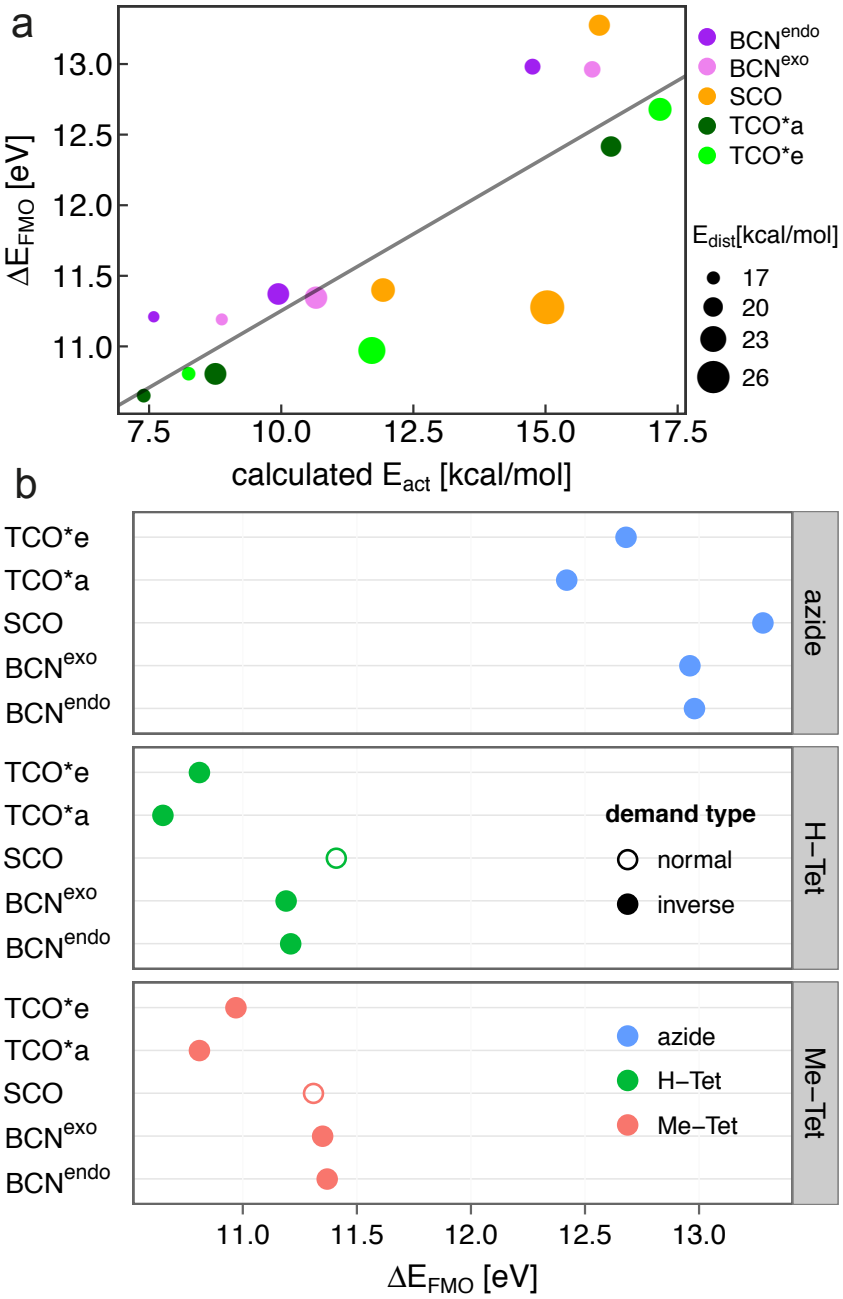

Figure 3.4: Electron demand determines cycloaddition rate. a) Correlation between FMO energy differences (ΔE_{FMO}) of the 15 distinct reactions (E_{act} is the barrier for the two enantiomers obtained in antiparallel tail orientation). The dot size represents the summed strain energies of both reactants. The solid line shows the linear fit of the data. b) Electron demand of the 15 different reactant pairs according to FMO energy gaps.

has been shown to perform exceptionally well. In the Diels-Alder (DARC) subset of the large GMNTK reaction set used to benchmark DFT functionals for kinetics

and non-covalent interactions, M06-2X is one of top performers [137, 138, 125, 135].

No empirical dispersion correction is used in conjunction with M06-2X, since due to its hybrid approach it already accounts for medium range dispersion interaction originating from the HF part. The M06-2X functional comes at moderate computational cost with a scaling of the order of $\sim O(N^4–N^6)$ [134, 139, 137, 140]. As basis set we used the 6-311+G(d,p) [141], consisting of 6 plane-wave basis functions and additional polarizable and diffuse functions on heavy as well as hydrogen atoms.

The quantum chemical calculations were performed using the program Gaussian 09 [142], which has the necessary functional and basis set incorporated.

Initial conformations of all 11 reactants (Figure 3.3a, Figure 3.5) were obtained by steepest descent and conjugate gradient energy minimization starting from a large rotamer ensemble of each reactant using the general amber force field (GAFF) [133, 143], deploying the software Avogadro [144]. The lowest energy conformations were then used as input coordinates for quantum mechanical (QM) optimization using the M06-2X Minnesota functional [133] and the Pople split valence basis set 6-311+G(d,p) [141] deploying Gaussian 09 [142]. The M06-2X functional is known to underestimate dispersion interactions such as those of benzene in Me-Tet and H-Tet [145, 146]. However, this error should largely cancel out in our calculations as the van der Waals complex that we chose as the reference state for E_{act} has dispersion interactions highly similar to the one of the TS.

Initally all calculations have been performed *in vacuo* and a polarizable conductor model (PCM) has been applied on the final geometry optimized compounds to account for solvation effects. A general stabilization by the solvent (water) of about $\sim$0.2 kcal/mol was obtained.

We report the experimentally more relevant results from calculations in solvent only. For several reactions, very different conformations and also reaction pathways were obtained in solvent calculations vs. vacuum, hence computing solvated energies on conformations obtained from vaccum calculations was not deemed suitable.

For post-processing of the QM data, Python scripts using the excellent cclib library have been used [147]. Cclib-based scripts allow to parse and analyze large amounts of QM output data in a standardized and automated manner across different QM programs.

Upon ligation of each of the eight eight-membered ring isomers to azide, Me-Tet and H-Tet, respectively, in both an antiparallel and parallel fashion with respect to the tail orientations of the two reactants, we obtained 48 products, which were again minimized using M06-2X.

Due to the multi-configurational phase space of the complex molecules investigated here, obtaining global minima and real TS structures pose a challenge. The transition state (TS) search for each of the reactions was performed through relaxed potential energy surface (PES) scans in reverse reaction coordinate direction, at the full M06-2X/6-311+G(d,p) level of theory. For SPAAC reactions, one-dimensional PES scans were performed by opening the N-N-N azide angle in steps of 2 degrees, refined to 0.5 degrees close to the saddle point. The two highest energy conformations from the PES scans were then used as starting points for a TS search employing the

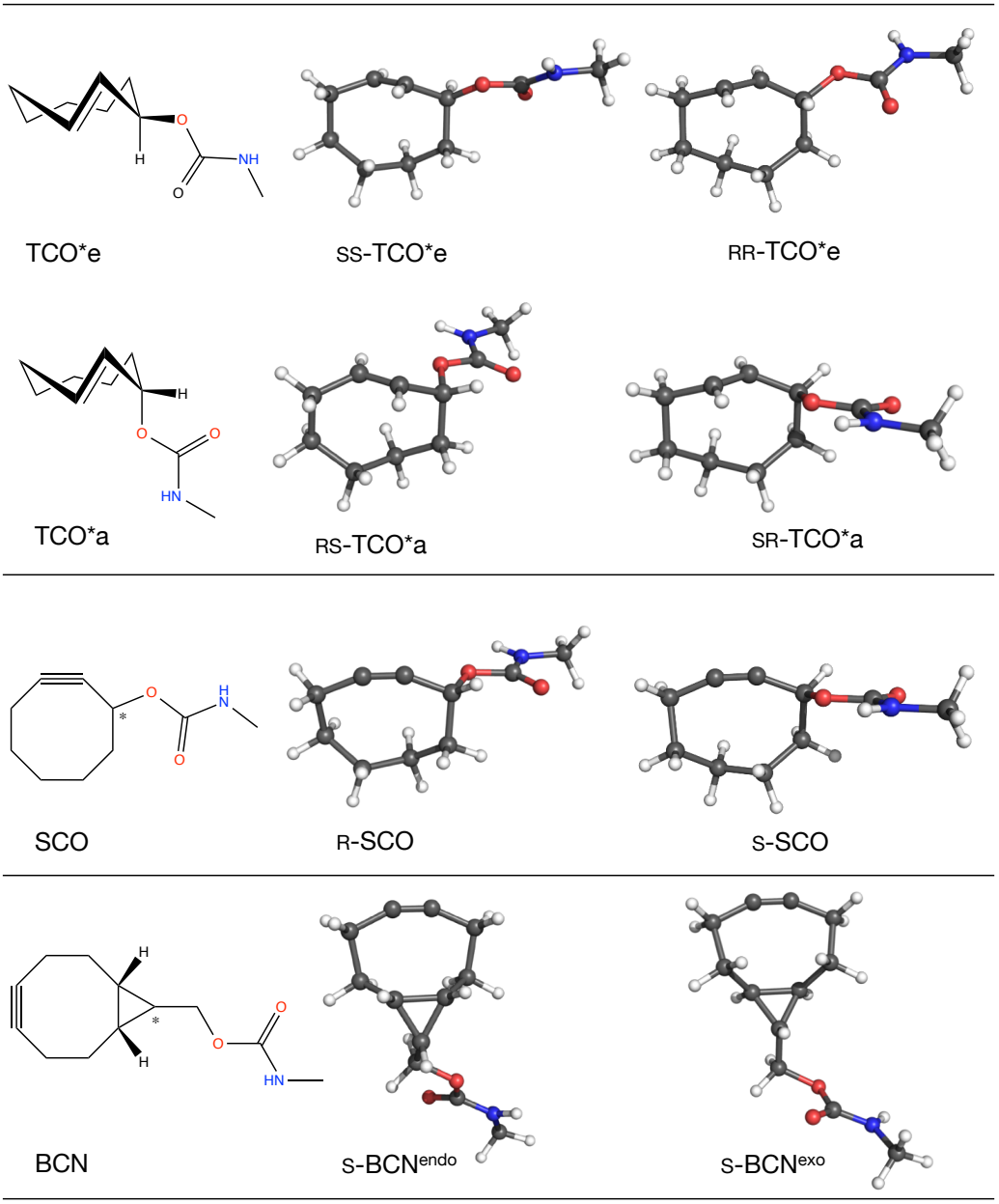

Figure 3.5: Structures and naming scheme of M06-2X/6-311+G(d,p) optimized geometries of all eight-member ring confomers.

Berny algorithm. SPIEDAC/SPINEDAC scans were performed along two dimensions by increasing the lengths of the two newly formed C-N bonds by about 0.2 Å per step. Here, three scan points in the vicinity of the saddle point were used as starting points for the TS search. While not always all three starting points lead to a TS structure as indicated by a single imaginary frequency, with SPIEDAC/SPINEDAC reactions being particularly challenging, we could finally obtain TS for all 48 reactions from at least two different starting points in the scanned PES. The imaginary

frequencies of the transition states are listed in Table 3.2, Section 3.3.4.

Van der Waals complexes of all reactions were obtained by geometry optimizing a conformation close to the TS with the PES scan coordinates slightly above the TS treshold. All of the obtained geometry optimized conformations (reactants, complexes, and products) and their energies were validated by additional optimizations starting from other initial configurations.

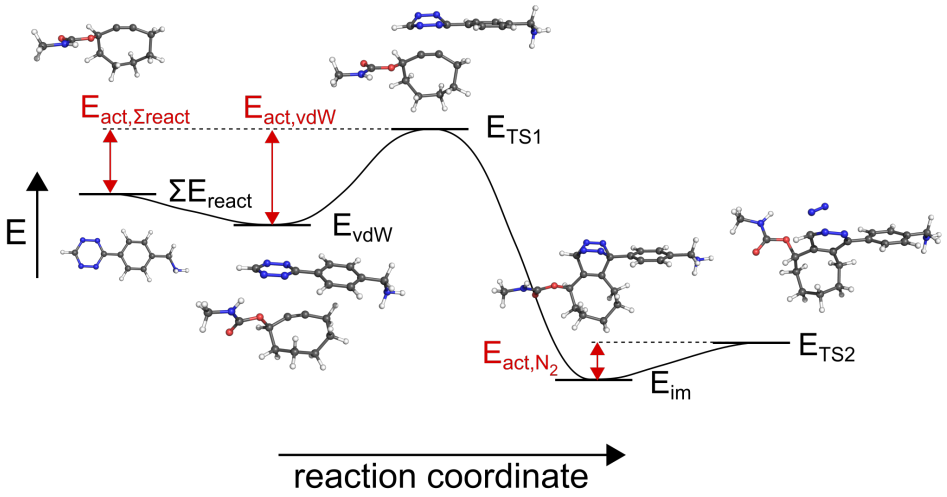

Figure 3.6: Reaction pathway of the R-SCO – H-Tet cyclocaddition. E_{act} was calculated relative to the van der Waals complex' energy (E_{vdW}) as opposed to the sum of the isolated reactants energies ($\sum E_{react}$). Also, the activation energy $E_{act,vdW}$ between the van der Waals complex (E_{vdW}) and transition state 1 (E_{TS1}) is the rate limiting step. The barrier of the N_2 exit (E_{act,N_2}) between the intermediate complex (E_{im}) and transition state 2 (E_{TS2}) is around ~1.5kcal/mol.

The ligation of Me-Tet and H-Tet to the strained rings is followed by a low barrier N_2 exit, for which we calculated barrier heights of 1.5–2 kcal/mol, so that the first cycloaddition step represents the rate-limiting step (see Figure 3.6). This is in agreement with previous studies [126].

We identified the set of possible interacting molecular orbitals by suprafacial symmetry overlaps between the relevant occupied and unoccupied orbitals for any pair of reactants. The smallest energy difference defined the interacting orbitals and thereby if a normal (diene HOMO with dienophile LUMO) or inverse demand Diels-Alder reaction (diene LUMO with dienophile HOMO) is at play (Tables 3.3 and 3.4). An example for the interacting orbitals of H-Tet and TCO*ax is shown in Figure 3.1.

As a measure for the reaction rates of any of the putative 48 reactions (3 reactants azide/Me-Tet/H-Tet times 8 different enantiomers and diastereomers of TCO*/SCO/BCN, times two different relative tail orientations, Table 3.2), we calculated the activation energy E_{act} along the reaction pathway (Figure 3.6). Con-

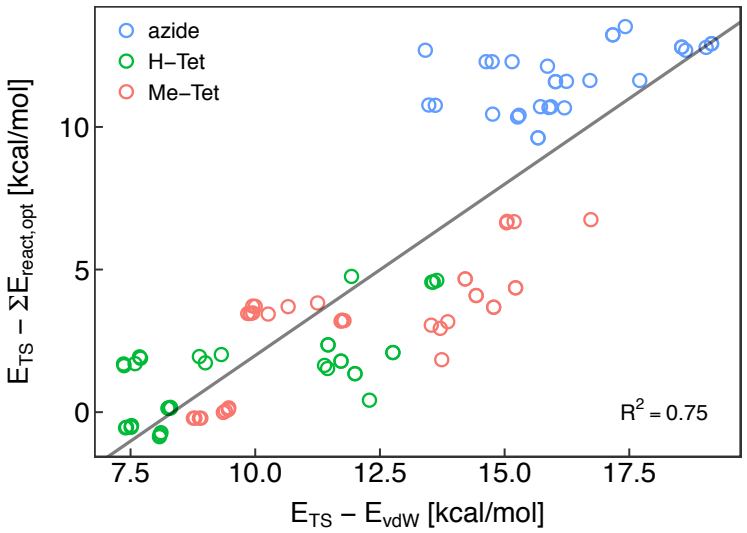

Figure 3.7: Correlation of vdW complex activation energies $E_{\mathrm{act,vdW}}$ with those using the sum of reactant energies $\sum E_{\mathrm{react,opt}}$ as a reference. The line depicts the linear regression over all energies and has a slope of ∼1.2.

ventionally, and also as previously reported for this class of reactions,[125] E_{act} is defined as the difference between the transition state (TS) energy and the sum of the energies of geometry optimized reactants, $E_{\mathrm{act},\sum\mathrm{react}} = E_{\mathrm{TS}} - \sum E_{\mathrm{react,\,opt}}$. Here, we instead use the energy of the van der Waals complex formed by the two reactants, E_{vdW}, as a reference, *i.e.* defined $E_{\mathrm{act}} = E_{\mathrm{act,vdW}} = E_{\mathrm{TS}} - E_{\mathrm{vdW}}$ (Figure 3.6), the reason for which was two-fold. First, TS conformations and van der Waals complexes are affected by basis set superposition errors (BSSE), whereas the conformations of the smaller single reactants are less affected. Thus, the BSSE is inherently included in $E_{\mathrm{act},\sum\mathrm{react}}$ but largely cancels out in $E_{\mathrm{act,vdW}}$. In addition, we found that $\sum E_{\mathrm{react,\,opt}}$ is larger than E_{TS} for some of the fastest reactions involving TCO*, which would imply the absence of any activation barrier and is an artefact from overlooking the formation of a favourable van der Waals complex prior to the reaction. Nevertheless, we found a correlation of $E_{\mathrm{act,vdW}}$ with $E_{\mathrm{act},\sum\mathrm{react}}$ (Figure 3.7), and report $E_{\mathrm{act,vdW}}$ as activation energies in the following. Orbital energies were computed using the HF/6-311+G(d,p) level of theory and we define ΔE_{FMO} as the energy difference between two relevant interacting orbitals.

For computational efficiency, smaller compounds compared to the experimental reactants [117, 132] were used, but including a significant fraction of the handles and therefore still more detailed when compared to the highly simplified scaffolds of previous computations. The tetrazines R-group features a benzylamine, the azide a N-propyl, and the eight-ring a N-methylcarbamate group.

While for the DFT calculations we considered all conformers separately, experimental rates were obtained with mixtures of the TCO*a, TCO*e and SCO enantiomers, using stopped-flow spectroscopy [132]. This resulted in a total of 15 measured reactions rates directly comparable to our calculated barriers.

3.3.2 Results and Discussion

The understanding of click-chemistry reactions at their highest level of detail is crucial to improve their applicability in biomolecule labeling at the base of single-molecule spectroscopy. Therefore, azide/tetrazine cycloadditions to cyclooctenes and cyclooctynes were herein investigated taking into account the effect of R-substituents used for protein functionalization.

We obtained M06-2X energies from optimization of all stereoisomers including enantiomeric pairs of the reactants' set (Figure 3.5). These energies suggest two energetically distinct TCO isomers, of which the equatorial TCO*e was found to be more stable than the axial TCO*a by 1.1 kcal mol^{-1} (Table 3.1). TCO*e also is the less reactive molecule among the tetrazine cycloadditions (Figure 3.3a), but undergoes more readily cis-isomerisation [132]. The kinetic measurements are taken from Hoffmann *et al.* [132] and were obtained using stopped-flow spectoscropy. All relevant experimental and calculated values are listed in Table 3.2.

We next calculated conformations and energies of van der Waals complexes and transition states for all possible reactions to estimate energy barriers E_{act}. We obtained a high correlation between measured rates and calculated energy barriers (Figure 3.3b, $R^2 = 0.89$), validating our quantum mechanical calculations. Due to the sterically demanding protein that the eight-membered rings are attached to, the azide and tetrazine substituents are prone to orient in an antiparallel fashion to the carbamate sidechain of the eight-ring. Thus the energy barriers of such antiparallel oriented linker configurations were chosen here. Although a similar correlation between experimental rates and calculated activation energies was obtained when the side-chain regioselectivity was ignored (Figure 3.10), individual barriers can vary by up to 3 kcal/mol when changing tail orientation (Table 3.2). This suggests the steric demand of the linkers including the bulky label and biomolecule to crucially determine the reaction kinetics.

To identify the origin of the differences in reactivity, we analyzed the energy differences ΔE_{FMO} between interacting frontier molecular orbitals (FMO) of all 15 reactant pairs. The relevant FMOs and energies are listed in Tables 3.3 and 3.4. Overall, ΔE_{FMO} correlates very well with E_{act} (R^2=0.72, Figure 3.4a). As expected, ΔE_{FMO} systematically underestimates the barrier for ligations with high strain energies (dot size in Figure 3.4a). We obtained an improved correlation when comparing the sum of strain energies and ΔE_{FMO} with E_{act} (R^2=0.82), implying that in a first approximation these two contributions can be considered additive and are both critically determining reactivitiy.

The main contribution to strain energies generally comes from the azide or tetrazine, respectively (Figure 3.8). Tetrazine (Me-Tet and H-Tet) ligations to SCO as well as any of the cycloadditions involving Me-Tet show significantly higher strains than the other cycloadditions (Figure 3.4a and 3.8), suggesting that the methyl group of Me-Tet as well as the carbamate right next to the triple bond of SCO creates a steric hindrance for the transition state formation.

We also analyzed the origins for the different observed reactivity between the

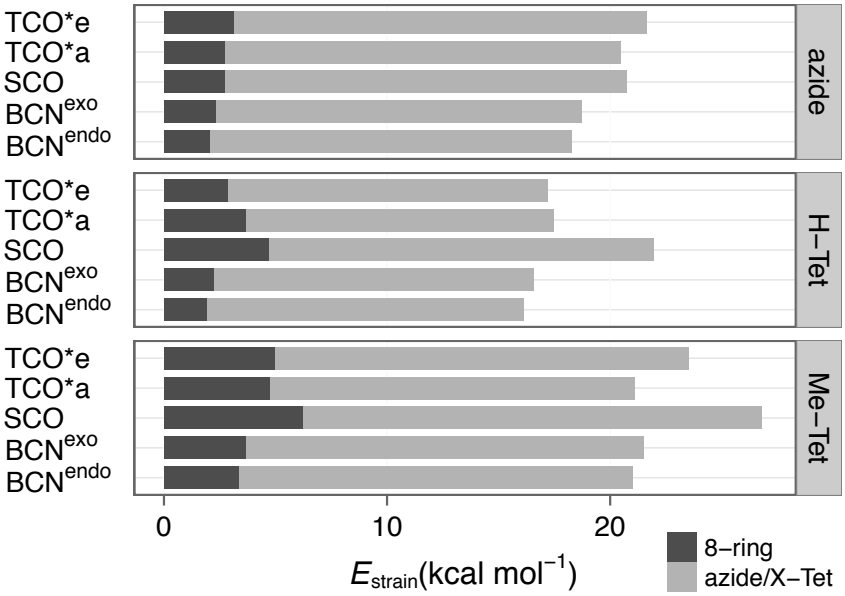

Figure 3.8: Strain energies E_{strain} from the vdW complex to the transition state conformation, decomposed for the ring and azide/tetrazine compounds (see Methods Sec. 2.1.8 for details).

investigated reactions. As expected, ΔE_{FMO} are overall higher in cycloadditions involving azides rather than tetrazines, explaining the well-known fact that n-propyl azide reacts less efficiently overall (Figure 3.4a) [121, 120]. Our calculations and measurements also confirm the previously established higher preference of this azide with $\text{BCN}^{\text{endo/exo}}$ over SCO [106, 109], for which we predict a ~3 kcal/mol difference in barrier. Our results further agree with the higher reactivity of azide with BCN^{endo} over BCN^{exo} [148], which is explained by a higher orbital overlap in *endo* reactions, while *exo* reactions are generally favoured by sterics [149]. The higher BCN^{endo} reactivity was found to hold similarly for reactions with both, H-Tet and Me-Tet (Figure 3.4a).

We can now ascribe the smaller barrier for the cycloaddition of the azide with BCN to a smaller HOMO-LUMO gap between the two reactants (Figure 3.4b). According to both calculations and measurements, TCO*a reacts faster than TCO*e with both H-Tet (~3-fold) and Me-Tet (~700-fold, Figure 3.3b), again directly in line with the smaller FMO energy gap for the TCO*a isomer (Figure 3.4b). The axial position increases the electron-withdrawing effect of the carbamate group, an effect that is further enhanced by the smaller strain deformations required for the ligation of TCO*a to a tetrazine (Figure 3.8).

Reactions involving H-Tet are generally faster than those with Me-Tet, a trend to be expected in this case of inverse electron demand of SPIEDAC reactions, as the methyl group shifts electron density into the reacting 6-ring (Figure 3.9a). Surprisingly, the so-called SPIEDAC involving SCO and a tetrazine [117] we instead predict to proceed with normal electron demand and correspondingly term this reac-

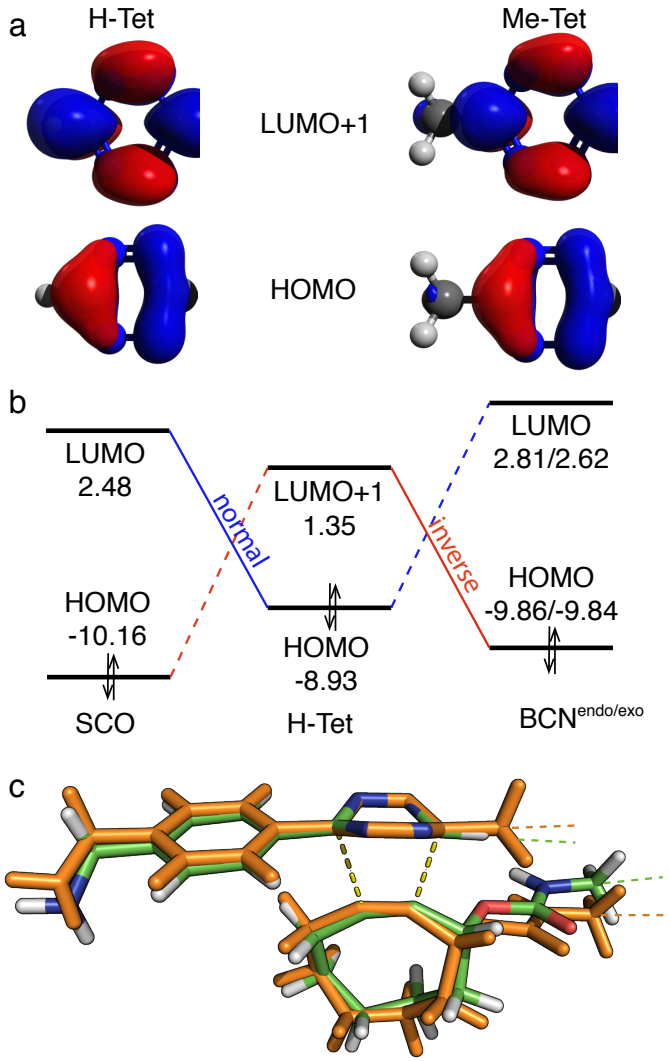

Figure 3.9: Origin of differences in electron demand and distortion strain a) The methyl-group of Me-Tet shifts electron density into the tetrazine ring, as evidenced by the differences in shape of the LUMO and HOMO between H-Tet and Me-Tet (for clarity only the substituted tetrazine ring is shown). b) Energy gaps in eV between FMOs of H-Tet reacting with SCO and BCN$^{endo/exo}$. SCO features lower FMO energies than BCN$^{endo/exo}$ and seems to undergo cycloadditions with H-Tet (and Me-Tet, see Table 3.4) with normal electron demand (solid blue line) instead of inverse electron demand as for BCN$^{endo/exo}$ (solid red line). The energy levels are not drawn to scale. c) The transition state of SCO–Me-Tet (orange) shows a significantly higher distortion than the one of SCO–H-Tet (green, compare orange and green dashed lines) resulting in a larger asymmetry with regard to the forming of C-C bonds (dashed yellow lines).

tion SPINEDAC. The carbamate group makes SCO more electrophilic than BCN, rendering the interaction of its LUMO with the tetrazine HOMO more favorable

(Figure 3.9b). SCO ligation to Me-Tet nevertheless is slower than to H-Tet, because the sterically more demanding methyl group gives rise to a 6 kcal/mol higher distortion of the transition state (Figure 3.4 and 3.9c). We therefore propose that this SPINEDAC reaction can be sped up conversely to SPIEDAC reactions, namely by more strongly electron-drawing cyclooctyne substituents and/or by further electron-donating tetrazine substituents, both with as little steric demand as possible. Our results, however, also emphasize that care must be taken as reactions can switch between inverse and normal electron demand upon allegedly minuscule chemical changes.

3.3.3 Conclusions

We here analyzed the determinants of reactivity for a set of strain-promoted cycloaddition reaction partners that is identical to those used for in vitro and in vivo labeling.

We could show that handles not only play the role to attach the reactive moieties to fluorescent labels and proteins. Instead, they also decisively influence the reaction rate of the cycloadditions, due to their electronic properties and steric demand. Our data explain how different isomeric states, methyl groups and carbamate linkers can independently and decisively alter the speed of SPAAC, SPIEDAC and the newly defined SPINEDAC reactions. Interestingly, the presence of the methyl and/or benzyl moieties of H-Tet and Me-Tet together with the electron-withdrawing properties of the carbamate handle of the SCO cyclo-octyne seems to switch the electron demand from an inverse to a normal type, with direct consequences for the rational tuning of the reactivity of these reactions. Furthermore, the two identified energetically distinct TCO* isomers, experimentally shown to be stable [132], exibit significantly different reaction rates upon ligation with both tetrazines. This finding leads to the observation that reaction kinetics can even be tuned by exploiting isomeric state differences that otherwise possess almost identical electronic configurations. The different incorporated ring-strain energies of isomeric states have a considerable impact on the selective labeling of biological processes.

A more strongly electron-donating group replacing the methyl or benzyl groups of H-Tet/Me-Tet will increase the reactivity of ligations with SCO, but not with TCO* and BCN. Even modifications of the aminogroup attached to the benzene ring of H-Tet/Me-Tet, although distant, can modulate the electron density of the benzene and thereby the reactivity of the tetrazines. Likewise, replacing the carbamate handle with more strongly electron-withdrawing groups, is predicted to have a similar effect. Benzyl and methyl moieties at Me-Tet and H-Tet are sterically highly demanding. Their substitution by smaller and/or electron-withdrawing groups should considerably enhance SPIEDAC, but not SPINEDAC reactions. These structural features should thus be taken into account when re-engineering cycloadditions in computations and experiments.

3.3.4 Appendix Part I

IUPAC nomenclature for the reactants used in this work (Figure 3.5): n-propyl azide,
(4-(1,2,4,5)-tetrazin-3-yl)phenyl methanamine
or short benzylamino tetrazine (H-Tet) and
(4-(1,2,4,5)-6-methyl-tetrazin-3-yl)phenyl methanamine
or short benzylamino methyltetrazine (Me-Tet) with
R,2E,R_P- and S,2E,S_P-cyclooct-2-en-1-methylcarbamate (TCO*e),
R,2E,S_P- and S,2E,R_P-cyclooct-2-en-1-methylcarbamate (TCO*a),
R- and S-cyclooct-2-yn-1-methylcarbamate (SCO), and
endo- and exo-bicyclonon-5-yn-1-methylcarbamate (BCN^{endo} & BCN^{exo}).

Table 3.1: M06-2X energies of the geometry optimized dienophiles and dienes.

dienophile	diene	Energy (H)
TCO*e		-596.4094
TCO*a		-596.4077
BCN^{endo}		-672.5492
BCN^{exo}		-672.5529
SCO		-595.1671
	H-Tet	-621.9683
	Me-Tet	-661.2818
	Azide	-282.6729

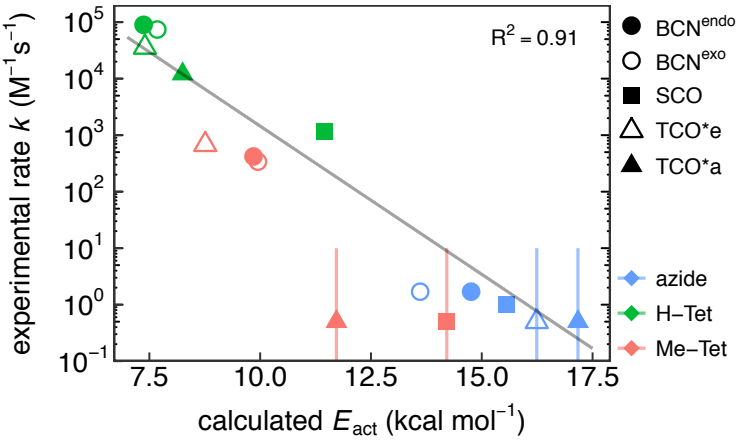

Figure 3.10: Correlation between experimental reaction rates k and the calculated activation energies E_{act} for each pair of reactants. In contrast to Figure 3.3b of the main text (only antiparallel tail orientations), we here chose the lower E_{act} among antiparallel and parallel tail orientations, ignoring regioselectivity. Even in this case we obtain an excellent correlation of 0.91 with the experimental rates. Only the lowest E_{act} among the different enantiomers (of SCO and TCO*) are shown. Experimental error bars are smaller than the symbol size and are omitted for clarity.

diene	dienophile	E_{act}	ΔE_{FMO}	measured k	error k	$E_{strain,ring}$	$E_{strain,azide/tet}$	$\sum E_{strain}$	Img freq
regioselectivity: antiparallel tail orientation									
azide	TCO*e	17.17	12.68	n.d.	n.d.	3.15	18.49	21.64	-457.3
azide	TCO*a	16.24	12.42	n.d.	n.d.	2.75	17.72	20.47	-457.0
azide	SCO	16.02	13.28	1.0	-	2.75	18.01	20.76	-454.7
azide	BCN^{endo}	14.76	12.98	1.7	0.3	2.08	16.21	18.29	-434.6
azide	BCN^{exo}	15.89	12.96	1.7	0.3	2.37	16.37	18.74	-436.9
H-Tet	TCO*e	8.25	10.81	12080	150	2.89	14.31	17.20	-307.9
H-Tet	TCO*a	7.40	10.65	35900	420	3.68	13.76	17.44	-289.9
H-Tet	SCO	11.93	11.41	1180	80	4.71	17.24	21.95	-372.7
H-Tet	BCN^{endo}	7.59	11.21	89700	3200	1.97	14.17	16.14	-297.0
H-Tet	BCN^{exo}	7.68	11.19	74090	1100	2.25	14.31	16.56	-299.8
Me-Tet	TCO*e	11.72	10.97	n.d.	n.d.	4.98	18.55	23.53	-350.7
Me-Tet	TCO*a	8.76	10.81	680	5	4.78	16.32	21.10	-326.6
Me-Tet	SCO	15.04	11.31	n.d.	n.d.	6.22	20.56	26.78	-394.4
Me-Tet	BCN^{endo}	9.95	11.37	420	10	3.36	17.63	20.99	-335.2
Me-Tet	BCN^{exo}	10.66	11.35	340	10	3.68	17.81	21.49	-337.9
regioselectivity: parallel tail orientation									
azide	TCO*e	19.04	12.68			3.72	19.91	23.63	-471.5
azide	TCO*a	18.55	12.42			2.78	18.51	21.29	-459.7
azide	SCO	15.56	13.28			3.13	17.70	20.83	-443.2
azide	BCN^{endo}	15.29	12.98			2.04	16.30	18.34	-437.0
azide	BCN^{exo}	13.61	12.96			2.31	16.27	18.58	-434.9
H-Tet	TCO*e	11.39	10.81			4.99	17.50	22.49	-340.8
H-Tet	TCO*a	8.08	10.65			3.57	13.72	17.29	-288.2
H-Tet	SCO	11.46	11.41			4.14	17.89	22.03	-375.1
H-Tet	BCN^{endo}	7.62	11.21			1.95	14.23	16.18	-295.7
H-Tet	BCN^{exo}	7.68	11.19			2.24	14.25	16.49	-296.3
Me-Tet	TCO*e	13.53	10.97			5.71	19.98	25.69	-369.9
Me-Tet	TCO*a	9.40	10.81			5.15	16.30	21.45	-325.8
Me-Tet	SCO	14.21	11.31			6.37	20.50	26.87	-389.3
Me-Tet	BCN^{endo}	9.85	11.37			3.38	17.66	21.04	-336.7
Me-Tet	BCN^{exo}	9.95	11.35			3.62	17.71	21.33	-336.3

Table 3.2: Overview of all activation energies (in kcal mol^{-1}), FMO gaps (in eV) and strain energies (in kcal mol^{-1}) for the eight-rings and the azide or tetrazine, imaginary frequencies (in cm^{-1}) of the TS for all isomers and tail orientations and measured experimental rates k (in $M^{-1}s^{-1}$). The measured reaction rates are for H-Tet-Cy5 and Me-Tet-Cy5, respectively (see [132] for more details). The reaction rates for the azide-SCO/BCN cycloadditions were taken from Borrmann *et al.* [120].

	MO and corresponding energies in (eV)			
compound	HOMO-1	HOMO	LUMO	LUMO+1
azide	-12.021	-10.260	3.116	4.063
H-Tet	-11.153	-8.929	0.765	1.346
Me-Tet	-10.989	-8.821	0.823	1.508
TCO*e	-11.807	-9.457	2.859	3.620
TCO*a	-11.812	-9.301	2.757	3.710
SCO	-10.550	-10.158	2.484	3.557
BCN^{endo}	-9.975	-9.859	2.805	3.317
BCN^{exo}	-9.966	-9.839	2.616	2.754

Table 3.3: FMO energies in eV computed at the HF/6-311+G(d,p) level of theory on geometry optimized structures.

diene	orbital 1	dienophile	orbital 2	ΔE_{FMO} (eV)
azide	LUMO	TCO*e	HOMO	12.68
azide	LUMO	TCO*a	HOMO	12.42
azide	LUMO	SCO	HOMO	13.28
azide	LUMO	$\mathrm{BCN}^{\mathrm{endo}}$	HOMO	12.98
azide	LUMO	$\mathrm{BCN}^{\mathrm{exo}}$	HOMO	12.96
H-Tet	LUMO+1	TCO*e	HOMO	10.81
H-Tet	LUMO+1	TCO*a	HOMO	10.65
H-Tet	HOMO	SCO	LUMO	11.41
H-Tet	LUMO+1	$\mathrm{BCN}^{\mathrm{exo}}$	HOMO	11.21
H-Tet	LUMO+1	$\mathrm{BCN}^{\mathrm{exo}}$	HOMO	11.19
Me-Tet	LUMO+1	TCO*e	HOMO	10.97
Me-Tet	LUMO+1	TCO*a	HOMO	10.81
Me-Tet	HOMO	SCO	LUMO	11.31
Me-Tet	LUMO+1	$\mathrm{BCN}^{\mathrm{endo}}$	HOMO	11.37
Me-Tet	LUMO+1	$\mathrm{BCN}^{\mathrm{exo}}$	HOMO	11.35

Table 3.4: ΔE_{FMO} of interacting orbitals in eV. The red rows indicate normal demand Diels-Alder reactions, all others are inverse electron demand.

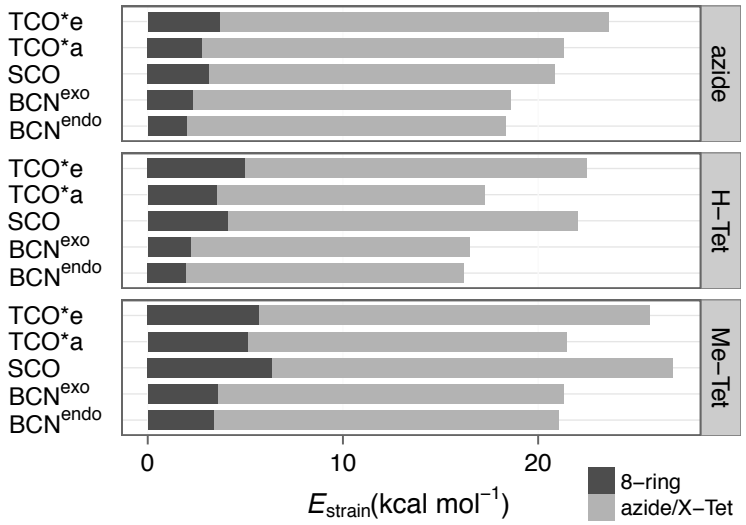

Figure 3.11: Decomposed strain energies for all reactions analogous to Figure 3.8, but for regioselectivity in parallel tail orientation. Particularly, all equatorial TCO's (TCO*e) show a higher strain energy compared to the antiparallel tail orientation.

Modulation of Electron Demand in Benzyl-Derivatives of Strain-Promoted Click Reactions

To be submitted

Authors: J.A. Wagner[1,2], D. Mercadante[1,3], F. Gräter[1,3]

[1] Heidelberg Institute for Theoretical Studies,
69118 Heidelberg, Germany
[2] Institute for Theoretical Physics, Heidelberg University,
69120 Heidelberg, Germany
[3] Interdisciplinary Center for Scientific Computing,
69120 Heidelberg, Germany

Authors contribution: J. A. Wagner and F. Gräter designed research. Computer simulations and data analysis were performed by J. A. Wagner. The article was written by J. A. Wagner under supervision of F. Gräter. All authors contributed to the discussion.

3.4 Part II: Modulation of Electron Demand in Benzyl-Derivatives

In the previous section, we have established the different reactivities for various types of strain-promoted click reactions. One finding was that the electronic properties of the two reactants jointly decide on the reactivity. The focus of the second part of this study, outlined in this section, is the impact of the properties of the diene on the reactivity, including its frontier molecular orbitals, but also very importantly, the involved molecular distortions.

As motivated in the conclusion of Part I (Sec. 3.4), the substitution of the methylamino group by smaller and/or electron-withdrawing groups should considerably enhance SPIEDAC, but not SPINEDAC reactions. We here investigate the effect of substituted benzylamino tetrazines with high electron-withdrawing or donating capabilities. By modulating the electron demand with substituents via the benzylamino tetrazines benzyl ring, the cycloaddition reaction rate can be tuned not only by modifying the strained eight-rings, but also by modifying, and thus functionalizing the tetrazines. Therefore the understanding of the impact of the modulated electron demand via the benzyl ring on the tetrazine reaction center allows to achieve greater flexibility in designing and engineering strain-promoted cycloaddition reactions with the desired properties.

Electron donating groups (EDG) and electron withdrawing groups (EWG) mainly act via resonance or inductive effects. EDGs act via the p-electron system, where delocalized electrons and its bonding cannot be expressed by one single Lewis structure, but instead are represented by resonance structures and are considered activating groups. EWGs instead act via s-electrons by means of charge transmission or electronegativity, resulting in dipoles in a bond and are considered deactivating. Both shift electrons either towards the tetrazine (donating) or draw electrons away from it (withdrawing).

While the hydrogen atom is considered the standard and regarded as neutral, meaning neither activating nor deactivating, the activating/deactivating terming originates from an increase or decrease of rates relative to hydrogen. However, if for a given reaction the FMO theory switches the electron demand type from inverse to normal, *i.e.* converting a SPIEDAC to a SPINEDAC type reaction (see Sec. 3.3), the opposite rate effect should be observed.

We here present rationally designed dienes which are comprised of benzyltetrazines with two R_1 substituents on the tetrazine ring and seven different R_2 substituents on the benzyl ring (Figure 3.12). The R_2 substituents were designed to cover a wide range of electron withdrawing and donating capabilities and therefore to functionalize the linker to the fluorophore or any other molecule. While R_1 acts activating or deactivating directly on the tetrazine, R_2's influence is modulated via the benzyl ring. To isolate the influence of both R-groups on the ΔE_{FMO} gap, only two different strained cyclooctynes where choosen as dienophile reaction partner: namely $BCN^{endo/exo}$ and SCO. As established in the previous section, despite the

similar HOMO energy levels of SCO and BCN$^{endo/exo}$ (Tab. 3.3), the cycloaddition of BCN$^{endo/exo}$ with a tetrazine has a smaller Δ E_{FMO} gap in inverse electron demand direction while the cycloaddition of SCO with a tetrazine has the smaller Δ E_{FMO} gap in normal electron demand direction (Figure 3.4).

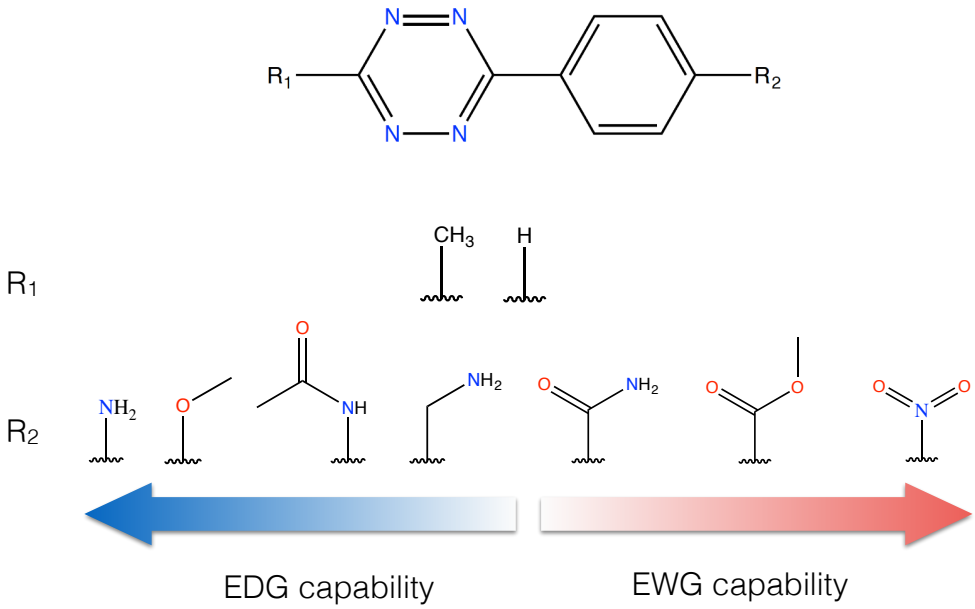

Figure 3.12: Scheme of the investigated R-groups on the benzyl-tetrazine. From left to right: amino, methylether, acetamide, methylamino, carboxyamide, acetate and nitro. The R-groups are sorted by their electron donating or withdrawing capability. The electron donating groups (EDG) are considered activating, while the electron withdrawing groups (EWG) are considered deactivating.

Figure 3.12 shows all seven R_2-groups. The four R_2 EDG's ordered from weakly to strongly activating (ortho/para-directing) are:

moderately activating

- methylamino (-CH_2NH_2)
- acetamide (-$NHCOCH_3$)

strongly activating

- methylether (-OCH_3)
- amino (-NH_2)

and the three R_2 EWG's ordered from weakly to strongly deactivating (meta-directing) are:

moderately deactivating

- carboxyamide (-$CONH_2$)
- acetate (-$COOCH_3$)

strongly deactivating

- nitro (-NO_2)

3.4.1 Methods

Analogous to Part I (Sec. 3.3), all eight-ring members and substituted tetrazine rings have been geometry optimized at the M06-2X level of theory. The intermediate complex before the N_2 exit (see Figure 3.6) has been taken as starting point for a scan of the PES in direction of the transition state. From the obtained PES, three points in the vicinity of the saddle point have been used as starting points for a transition state search. After having obtained candidates for a transition state, frequency calculations have been performed to confirm the transition state is a first order saddle point in the harmonic approximation. Originating from the transition state, a small displacement towards separating the reactants has been introduced and a geometry optimization to lower the energy was performed. The resulting structure at the local energy minima is the van der Waals complex (see Figure 3.6). Figure 3.13 shows the correlation between the activatin energy E_{act} with respect to the van der Waals complex energies and the energies of both reactants in the relaxed state summed up. The R_1 methyl substituents show a much weaker correlation in activation energies between the summed reactants and van der Waals complex reference energies then the R_1 hydrogen substituents, particularly in conjunction with the BCN compounds. Furthermore the first definition can yield negative activation energies for BCN in combination with R_1 hydrogen substituents (Figures 3.13). Therefore, it is crucial to use the van der Waals complex as a reference to the transistion state for accurate results.

3.4.2 Results and Discussion

The electron donating/withdrawing capability can be quantified by the HOMO energy: a withdrawing of electrons out of the tetrazine ring stabilizes and thus lowers its energy. Figure 3.14 shows the FMOs for all substituted tetrazine rings. Despite being separated by a benzene ring, the R_2 groups influence on the HOMO, LUMO and LUMO+1 energies of the diene is clearly evident. By showing that by going from the heavily electron donating group amino to the heavily withdrawing group nitro, a stabilization of $\sim 1.9\,eV$ can be seen. The R_1 methyl substituent destabilizes the HOMO over the hydrogen R_1-group, whereby such a destabilization consistently increases with R_2's stabilizing effect from $\sim 0.05\,eV$ in the case of amino to $\sim 0.2\,eV$ in the case of the nitro group. The opposite effect is observed for the LUMO+1 energy. The R_1 methyl substituent still destabilizes, but the destabilizing effect consistently decreases with R_2's stabilizing effect. For the investigated dienes, the absolute value of the FMO energy levels is mainly influenced by the R_2 group.

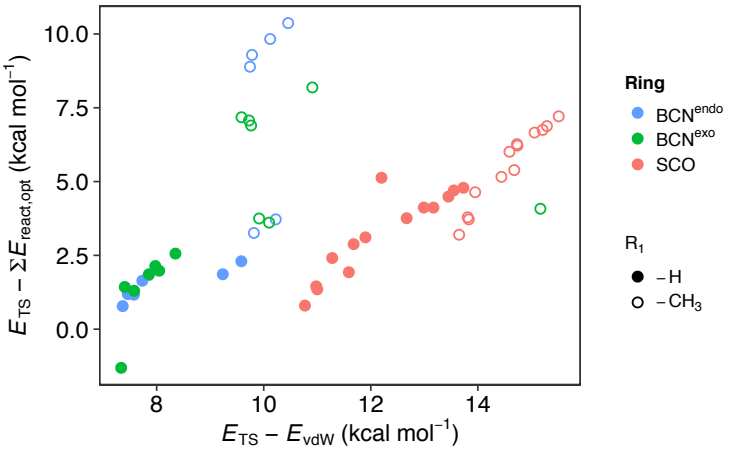

Figure 3.13: Correlation of activation energies E_{act} using as a reference the energy of the van der Waals complex E_{vdW} with those using the sum of reactant energies $\sum E_{react}$ as shown in Figure 3.6. The R_1 groups are depicted as solid (-H) and empty circles (-CH_3). The R_1 methyl group shows a much weaker correlation between the two definitions of the activation energy compared to the -H group. Data for the SCO include the barrier heights for both handle arrangements, parallel and antiparallel.

It is interesting to note that the insertion of an acyl group between the benzene ring and the amino group (R_2: methylamino) considerably lowers the HOMO energy ($\Delta E \approx 0.9$ eV). Similarly the insertion of a carbonyl group on the benzyl ring consists in a large change of the three orbitals energies monitored. With respect to these observations, the choice of functional groups immediately attached to the benzyl ring broadens the opportunity to modulate the energies of the HOMO and the relevant virtual orbitals, as well as the FMO gaps (Figure 3.15).

As shown in Sec. 3.3 of this chapter, the Diels-Alder type reactions between benzylamino tetrazine or benzylamino methyltetrazine and BCN$^{endo/exo}$ are SPIEDAC, while the reactions with SCO are suggested to be SPINEDAC undergoing cycloaddition with normal electron demand. However, the difference in the FMO gaps for the reaction with SCO in inverse (11.51 eV) and normal (11.41 eV) demand direction is very small. The wide range of electron withdrawing and donating R-groups on the tetrazine (Figure 3.12) was chosen to obtain a consistent increase in the differences between the inverse and normal electron demand FMO gaps such that a clearer distinction can be drawn. As the dienophile reaction partner, the two highly strained cyclooctynes SCO and BCN were choosen (Tab. 3.1).

Figure 3.15 shows the orbital gaps ΔE_{FMO} between the eight-rings in inverse and normal electron demand direction for all benzyl derivatives (R_2 substituents) and the R_1 substituents. Consistent for all eight-membered rings is the influence of the R_1 substituent: a lowered ΔE_{FMO} gap in inverse electron demand direction for hydrogen compared to the methyl substituent, and the opposite in normal electron demand direction: a higher ΔE_{FMO} gap for the hydrogen compared to methyl.

Due to the electron donating or withdrawing influence of the R_2 group, a clear

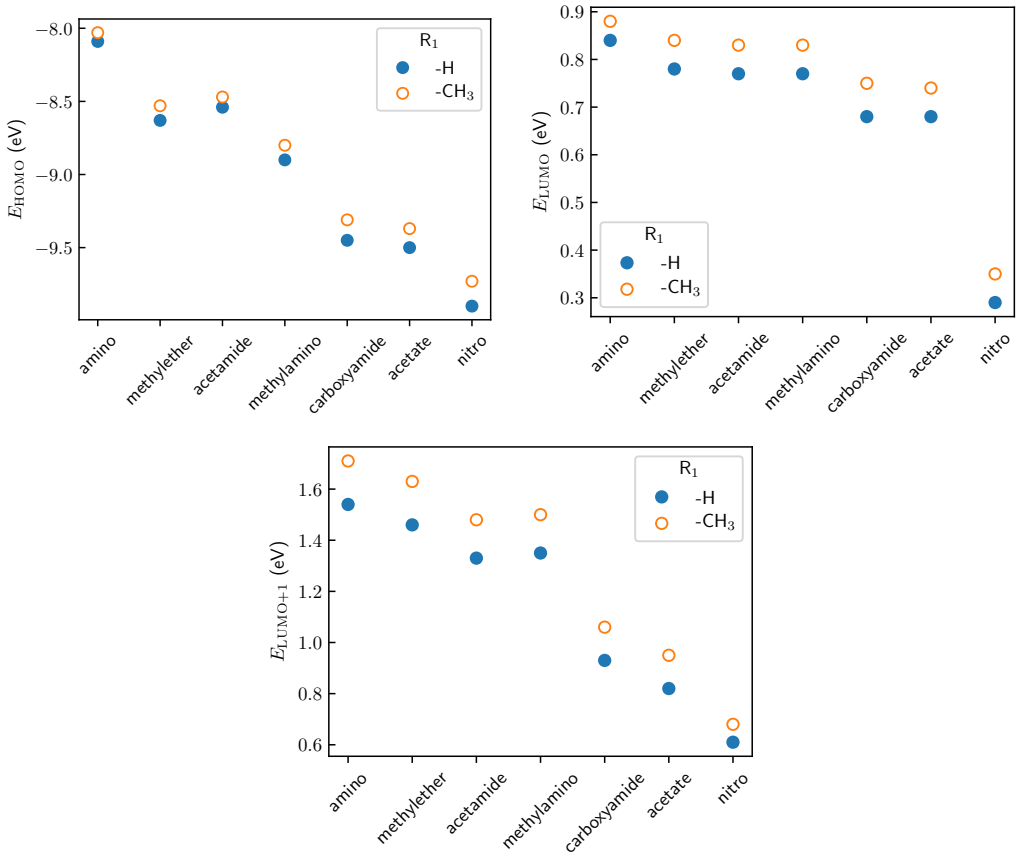

Figure 3.14: Energy levels of the HOMO (top left panel), LUMO (top right panel) and LUMO+1 (bottom panel) orbitals, calculated at the HF/6-311+G(d,p) level of theory for all R-groups of the benzyl-tetrazines. The orbital energy levels for the compounds featuring a hydrogen or methyl as R_1 group are shown as solid and empty circles, respectively. The x-axis lists the R_2 groups ordered from electron donating to withdrawing capability, compare Figure 3.12 for the respective structures.

trend towards higher ΔE_{FMO} gaps in inverse electron demand can be observed by going from the heavily electron withdrawing nitro group to the highly electron donating amino group. The energy gap in normal electron demand direction shows the opposite trend, leading to a crossing of ΔE_{FMO}.

In Part I (Sec. 3.3) we report that the regioselectivity of the eight-ring and tetrazines handles are a decisive factor for the activation energies. The handles of both, the diene and dienophile can be arranged in a parallel or antiparallel fashion. This has a direct impact on the strain and interaction energies, but also on experimentally accessible rates. Since k_{on} is typically measured by having the eight-ring incorporated into an unnatural amino acid (UAA) while the tetrazine is attached to a fluorophore, the reaction most likely measured is in the antiparallel handle arrangement due to the large steric hindrances of both, the protein and fluorophore. Therefore, we report the activation energies for reactions with SCO separately in

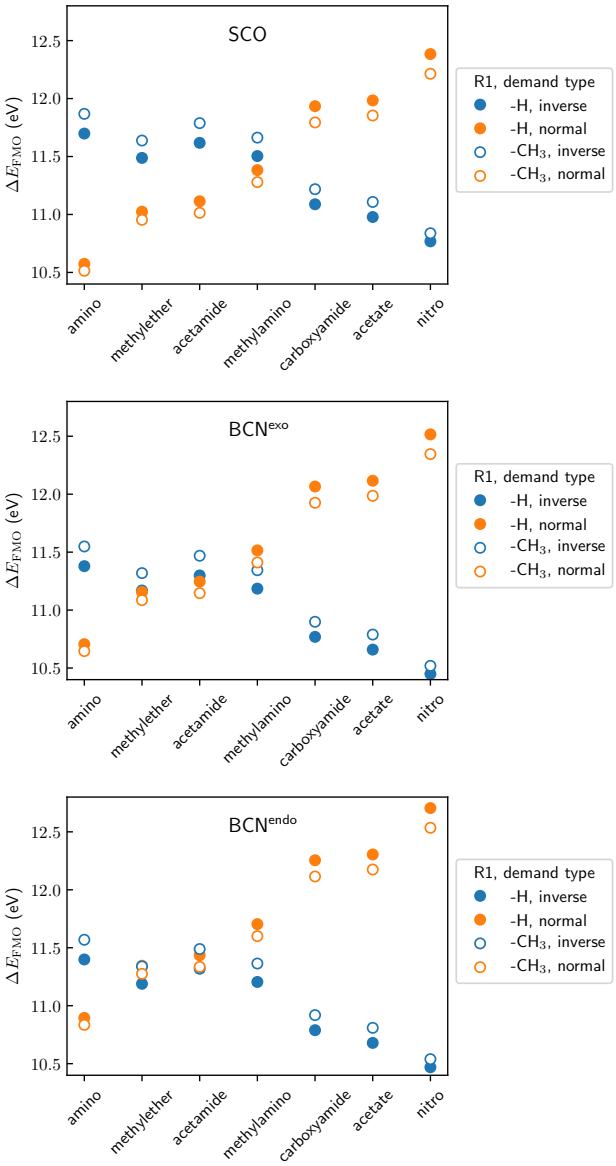

Figure 3.15: Δ E_{FMO} gap in inverse and normal electron demand direction shown for both R_1 groups, all R_2 groups (x-axis) and the three eight-rings SCO, BCN^{exo} and BCN^{endo}. The smaller Δ E_{FMO} gap should dictate the direction of electron transfer occuring in the reaction. The crossing of inverse vs. normal demand Δ E_{FMO} shifts towards more electron withdrawing dienes by going from SCO over BCN^{exo} to BCN^{endo}.

both, parallel and antiparallel handle arrangement (Figure 3.16). For $BCN^{endo/exo}$, the regioselectivity has almost no influence due to the handle being located on the opposite site of the triple bond (see Figure 3.5).

Figure 3.16 shows the activation energies with respect to both R-groups. For all

eight-rings, the R_1 substituent, directly attached to the tetrazine, has a large effect on the observed activation energy E_{act}. By substituting hydrogen with methyl, the activation energy for all R_2 substituents increases about ~ 2.5 kcal mol^{-1} for SCO, ~ 5.5 kcal mol^{-1} for BCN^{endo} and ~ 8.0 kcal mol^{-1} for BCN^{exo}. Across all seven R_2 substituents, a consistent increase in activation energy is observed by decreasing electron withdrawing or increasing electron donating capability. In a parallel arrangement of the handles of SCO and the tetrazine, the observed barrier is lower than in antiparallel arrangement. The increased strain energy contribution to the transition state is offset by a yet larger increase in interaction energy (Table 3.2).

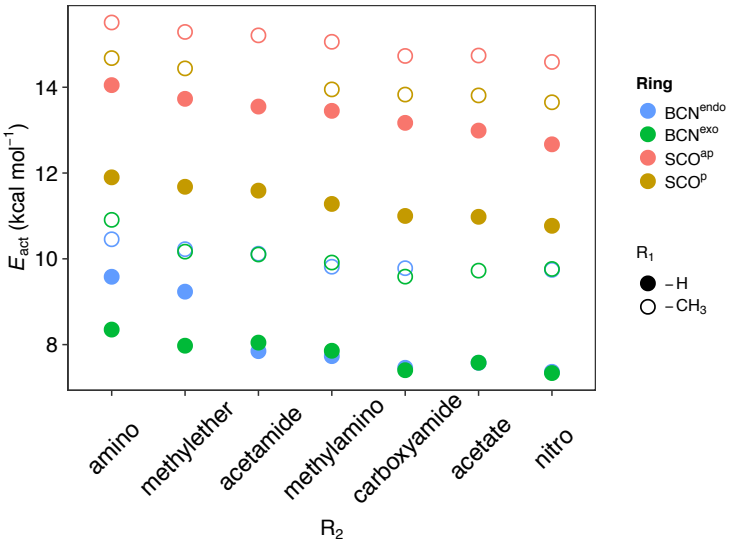

Figure 3.16: Activation energies for the cycloaddition reactions in kcal mol^{-1} for both R_1 and all seven R_2 substituents on the tetrazine together with the dienophiles SCO or BCN$^{endo/exo}$. The activation energies with SCO are shown for both handle arrangements, parallel (SCO^{p}) and antiparallel (SCO^{ap}). The higher activation energies for the R_1 methyl substituent compared to hydrogen mainly originates from the higher strain/distortion imposed by the methyl group (see Section 3.3).

According to FMO theory, the lower of the two ΔE_{FMO} gaps should direct the electron demand, switching the reaction from inverse to normal demand by going towards more electron donating R-groups. For the cycloaddition with SCO, ΔE_{FMO} in normal electron demand direction becomes lower by substituting carboxyamide with methylamino, for BCN$^{endo/exo}$, the flip occurs for groups with higher electron donating capability, namely by going from methylamino to acetamide (Figure 3.15). This lead to the assumption that the cycloaddition of SCO with a tetrazine incoorporating the methylamino R_2 group takes place in normal electron demand. If the ΔE_{FMO} values really measure reactivity, one would expect a strong correlation of ΔE_{FMO} with E_{act}, *i.e.* E_{act} should also show a turnover along the R_2 substituents as found for ΔE_{FMO}. However, as shown in Figure 3.17, the correlation of the activation energy E_{act} with ΔE_{FMO} is only positive for the inverse electron demand, *i.e.* an electron transfer from the electron rich dienophile (eight-ring) to the electron

poor diene (substituted tetrazine).

Interestingly, from Figure 3.15 it can be seen that the orbital gaps for the reaction between SCO and methylamino tetrazine fall on very similar values for both normal and inverse electron demands, whereas for the reaction with $BCN^{endo/exo}$ the ΔE_{FMO}, suggesting inverse demand, is considerably lower. This has previously led to the consideration that the SCO-methylamino tetrazine reaction could be driven by normal demand[reference the paper]. Nevertheless, in Figure 3.17 the correlation between activation and ΔE_{FMO} is negative for normal demand, signifying that the reaction is rather driven by inverse demand. This observation suggests that the FMO theory derived from HF calculations limits the full understanding of the mechanisms driving this type of reactions and that exclusively looking at the orbital energy gap ΔE_{FMO} is not sufficient. Indeed, HF theory is known to overestimate the HOMO-LUMO gap as the energies calculated for occupied and unoccupied states can only be considered as lower and upper boundaries respectively. Moreover, additional factors, such as strain energy, can significantly contribute to the modulation of the reactivity of the investigated compounds, overall adding to the knowledge provided by the analysis of HOMO-LUMO energy differences. Therefore, only by looking at the overall correlation of activation energy with the ΔE_{FMO} for a range of R_2 substituents modulating the FMO levels, it is possible to identify if the investigated reaction is responding to a normal or inverse demand mechanism.

3.4.3 Conclusions

We could show that the FMO gap as well as the activation energy consistently increases with reduced activation. While the R_1-group directly attached to the tetrazine ring has the larger impact on the activation energy of the cycloaddition of a tetrazine and an eight-ring, the R_2-group attached to the benzyl ring has the larger effect on the absolute values of the FMO energies and ΔE_{FMO} gaps. By shifting the heavily electron donating amino group one acyl group further away, rendering it a methylamino group, its electron donating capabilities are almost neglibile. Opposed to the previous finding that the reaction between a benzylamino tetrazine and a strained cyclooctyne undergoes cycloaddition in normal electron demand, we showed that this was a misleading result. By looking at the correlation of the ΔE_{FMO} gaps and activation energies for a wide range of electron donating or withdrawing R-groups, we could demonstrate that the investigated cycloaddition reactions do occur in inverse electron demand SPIEDAC type.

Considering the previously reported excellent agreement of our calculated activation energies with experimental rates, we can assume that the here presented theoretical calculation capture the relative differences in activation energies fairly accurate. The FMO theory holds well by correlating Hartree-Fock ΔE_{FMO} with the activation energies, but fails in conjunction with overestimated HF orbital gaps in explaining the irrelavance of the lower energy normal electron demand ΔE_{FMO} gap, which is not reflected in barrier heights.

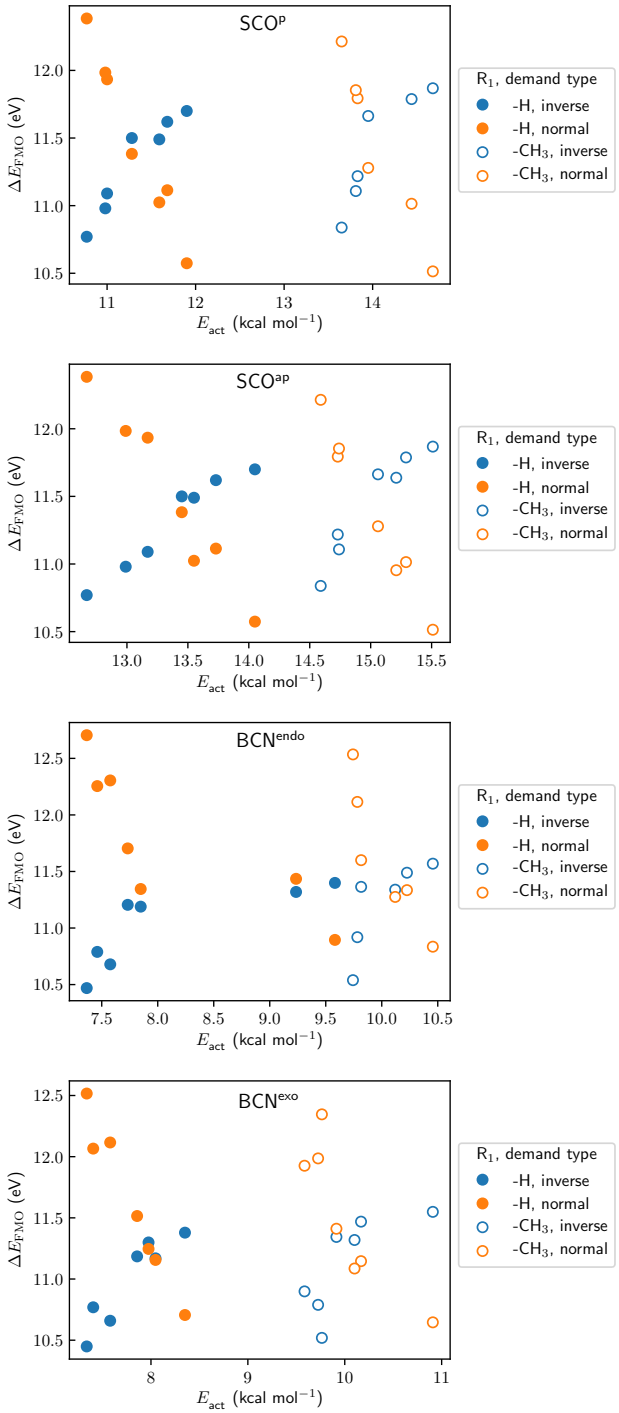

Figure 3.17: Correlation of activation energies with orbital gaps according to FMO theory for all investigated cycloaddition reactions. From top to bottom: SCO for parallel and antiparallel tail orientation, BCN^{endo} and BCN^{exo}. The FMO-gaps are shown for both, inverse and normal electron demand. It can be clearly seen that a positive correlation between the activation energy and FMO gap is only present for the inverse demand, suggesting that all reactions occur in inverse demand.

3.4.4 Appendix Part II

Table 3.5: M06-2X energies of all geometry optimized dienes.

R_2	M06-2X Energies (H)	
	R1: -H	R1: -CH_3
amino	-582.67736	-621.99042
methylether	-641.83099	-681.14429
acetamide	-735.32185	-774.63526
methylamino	-621.96833	-661.28177
carboxyamide	-696.02002	-735.33378
acetate	-755.17994	-794.49375
nitro	-731.80547	-771.11955

		MO and corresponding energies in (eV)		
R_1	R_2	HOMO	LUMO	LUMO+1
-H	amino	-8.090	0.838	1.537
-H	methylether	-8.626	0.784	1.460
-H	acetamide	-8.544	0.773	1.332
-H	methylamino	-8.899	0.771	1.346
-H	carboxyamide	-9.447	0.683	0.934
-H	acetate	-9.501	0.676	0.824
-H	nitro	-9.897	0.290	0.607
-Me	amino	-8.028	0.884	1.706
-Me	methylether	-8.533	0.838	1.632
-Me	acetamide	-8.470	0.827	1.483
-Me	methylamino	-8.795	0.829	1.505
-Me	carboxyamide	-9.314	0.747	1.059
-Me	acetate	-9.366	0.741	0.945
-Me	nitro	-9.728	0.355	0.675

Table 3.6: FMO energies in eV computed at the HF/6-311+G(d,p) level of theory on geometry optimized tetrazine substituents.

dienophile	regioselectivity	R_1	R_2	E_{act}	ΔE_{FMO} inverse	ΔE_{FMO} normal
SCO	ap	-H	amino	12.20	11.698	10.574
SCO	ap	-H	methylether	13.73	11.618	11.114
SCO	ap	-H	acetamide	13.55	11.488	11.024
SCO	ap	-H	methylamino	13.45	11.504	11.383
SCO	ap	-H	carboxyamide	13.17	11.088	11.934
SCO	ap	-H	acetate	12.99	10.978	11.984
SCO	ap	-H	nitro	12.67	10.768	12.384
SCO	ap	-Me	amino	15.51	11.868	10.514
SCO	ap	-Me	methylether	15.29	11.788	11.014
SCO	ap	-Me	acetamide	15.21	11.638	10.954
SCO	ap	-Me	methylamino	15.06	11.663	11.279
SCO	ap	-Me	carboxyamide	14.73	11.218	11.794
SCO	ap	-Me	acetate	14.74	11.108	11.854
SCO	ap	-Me	nitro	14.59	10.838	12.214
SCO	p	-H	amino	11.90	11.698	10.574
SCO	p	-H	methylether	11.68	11.618	11.114
SCO	p	-H	acetamide	11.59	11.488	11.024
SCO	p	-H	methylamino	11.28	11.504	11.383
SCO	p	-H	carboxyamide	11.00	11.088	11.934
SCO	p	-H	acetate	10.98	10.978	11.984
SCO	p	-H	nitro	10.77	10.768	12.384
SCO	p	-Me	amino	14.68	11.868	10.514
SCO	p	-Me	methylether	14.44	11.788	11.014
SCO	p	-Me	acetamide	-	11.638	10.954
SCO	p	-Me	methylamino	13.95	11.663	11.279
SCO	p	-Me	carboxyamide	13.83	11.218	11.794
SCO	p	-Me	acetate	13.81	11.108	11.854
SCO	p	-Me	nitro	13.65	10.838	12.214

Table 3.7: Overview of all activation energies E_{act} in kcal mol^{-1} with the SCO dienophile in both antiparallel and parallel handle arrangement and Δ E_{FMO} gaps in eV in inverse and normal electron demand direction.

dienophile	R_1	R_2	E_{act}	ΔE_{FMO} inverse	ΔE_{FMO} normal
BCN^{endo}	-H	amino	9.58	11.399	10.895
BCN^{endo}	-H	methylether	9.24	11.319	11.435
BCN^{endo}	-H	acetamide	7.85	11.189	11.345
BCN^{endo}	-H	methylamino	7.73	11.205	11.704
BCN^{endo}	-H	carboxyamide	7.46	10.789	12.255
BCN^{endo}	-H	acetate	7.58	10.679	12.305
BCN^{endo}	-H	nitro	7.37	10.469	12.705
BCN^{endo}	-Me	amino	10.46	11.569	10.835
BCN^{endo}	-Me	methylether	10.23	11.489	11.335
BCN^{endo}	-Me	acetamide	10.12	11.339	11.275
BCN^{endo}	-Me	methylamino	9.82	11.364	11.600
BCN^{endo}	-Me	carboxyamide	9.78	10.919	12.115
BCN^{endo}	-Me	acetate	-	10.809	12.175
BCN^{endo}	-Me	nitro	9.74	10.539	12.535
BCN^{exo}	-H	amino	8.351	11.379	10.706
BCN^{exo}	-H	methylether	7.975	11.299	11.246
BCN^{exo}	-H	acetamide	8.047	11.169	11.156
BCN^{exo}	-H	methylamino	7.857	11.185	11.515
BCN^{exo}	-H	carboxyamide	7.404	10.769	12.066
BCN^{exo}	-H	acetate	7.579	10.659	12.116
BCN^{exo}	-H	nitro	7.337	10.449	12.516
BCN^{exo}	-Me	amino	10.909	11.549	10.646
BCN^{exo}	-Me	methylether	10.168	11.469	11.146
BCN^{exo}	-Me	acetamide	10.101	11.319	11.086
BCN^{exo}	-Me	methylamino	9.914	11.344	11.411
BCN^{exo}	-Me	carboxyamide	9.585	10.899	11.926
BCN^{exo}	-Me	acetate	9.725	10.789	11.986
BCN^{exo}	-Me	nitro	9.764	10.519	12.346

Table 3.8: Overview of all activation energies E_{act} in kcal mol^{-1} with the $BCN^{endo/exo}$ dienophiles and Δ E_{FMO} gaps in eV in inverse and normal electron demand direction.

Chapter 4

Stress-induced long-range ordering in spider silk

Contents

Stress-induced long-range ordering in spider silk

In review at Scientific Reports

Authors: J.A. Wagner[1,2], S. P. Patil[3], I. Greving[4], Marc Lämmel[5], Konstantinos Gkagkas[6], T. Seydel[4], K. Kroy[5], M. Müller[4,8], Bernd Markert[3] and F. Gräter[1,9]

[1] Heidelberg Institute for Theoretical Studies,
69118 Heidelberg, Germany
[2] Institute for Theoretical Physics, Heidelberg University,
69120 Heidelberg, Germany
[3] Institute of General Mechanics, RWTH Aachen University,
52062 Aachen, Germany
[4] Institute of Materials Research, Helmholtz-Zentrum Geesthacht,
21502 Geesthacht, Germany
[5] Institute for Theoretical Physics, University of Leipzig,
04103 Leipzig, Germany
[6] Technical Center, Toyota Motor Europe NV/SA,
1930 Zaventem, Belgium
[7] Institute Max von Laue-Paul Langevin,
38042 Grenoble, France
[8] Institute for Experimental and Applied Physics, University of Kiel,
24098 Kiel, Germany
[9] Interdisciplinary Center for Scientific Computing,
69120 Heidelberg, Germany

Authors contribution: J. A. Wagner, S. P. Patil, I. Greving, M. Müller and F. Gräter designed research. Computer simulations and data analysis were performed by J. A. Wagner. The experiments were performed by I. Greving and T. Seydel. The analytical toy model was devised by J. Wagner and M. Lämmel. The article was written by J. A. Wagner, I. Greving and F. Gräter. All authors contributed to the discussion.

4.1 Abstract

The emergence of order from disorder is a topic of vital interest. We here propose that long-range order can arise from a randomly arranged two-phase material under mechanical load. Using Small-Angle Neutron Scattering (SANS) experiments and Molecular Dynamics based finite element (FE) models we show evidence for stress-induced ordering in spider dragline silk. Both methods show striking quantitative agreement of the position, shift and intensity increase of the long period upon stretching. We demonstrate that mescoscopic ordering does not originate from silk-specific processes such as strain-induced crystallization on the atomistic scale or the alignment of tilted crystallites. It instead is a general phenomenon arising from a non-affine deformation that enhances density fluctuations of the stiff and soft phases along the direction of stress. Our results suggest long-range ordering, analogously to the coalescence of defects in materials, as a wide-spread phenomenon to be exploited for tuning the mechanical properties of many hybrid stiff and soft materials.

4.2 Introduction

A material's mechanical behavior is ultimately defined by its underlying structure. Enormous progress has been made in the last decades in the knowledge and design of the molecular to macroscopic structure for even high-complexity materials, in order to tailor their mechanics and mechanical properties. However, our understanding of how the application of mechanical stress effects the material's structure on the relevant length scales, even though at least similarly critical for mechanical performance, needs expanding.

Many composite materials of mechanical interest feature a coexistence of stiff and soft phases, in which the stiff phase reinforces and thereby strengthens the soft phase. Analogously, semi-crystalline polymers, including synthetic and biological materials, consist of a highly structured stiff phase alternating with an amorphous, soft phase. It is well established that stretching semi-crystalline materials can lead to order on a short length scale, namely by stress-induced crystallization [150, 151, 19, 152]. Amorphous chains straighten and align, and thereby self-assemble into more ordered structures under force. This also leads to a reduction in the average tilt angle of the backbones within crystallites relative to the fiber axis [150, 153, 151, 154]. A completely unaddressed question however is if applied stress can also increase long-range order in the distribution of stiff components within the amorphous matrix of composite or semi-crystalline materials.

Interestingly, in a material with randomly distributed voids or defects, stress application causes these voids to nucleate. The coalescence and gradual volumetric growth of voids leads to the formation of larger defect-free regions and lowers the total internal energy of the system [155, 156, 157]. Similarly, stress can increase structural order in metals with irradiation-induced defects [158, 159, 160]. Each defect in a periodic arrangement of atoms or molecules gives rise to a local stress concentration. This stress can cause these defects to be 'dragged' towards each other,

which lowers the total internal energy of the system. This results in larger defect-free regions separated by clusters of defects along the direction of stress application.

Protein crystals show a similar behavior: Zemlin *et al.* could show that upon irradiation beyond a critical threshold, defects vanish in finite sized protein crystals, resulting in stress-induced order due to the build-up of internal stresses [161]. They observed the same effect in regular 2D-lattices of macroscopic bubbles, corroborating their conclusion that mechanical stress and not thermal energy induced the order. These seminal experiments nearly two decades ago have not yet been attempted with other systems. Stress-induced order as a general principle has yet to be uncovered.

Given that defects can be considered as a second phase in a two-phase system with inherently different mechanical properties, an obvious question is if stress-induced long-range order is a phenomenon similarly occuring in composite or semi-crystalline materials (Figure 4.1). An intriguing consequence of such ordering would be an attenuation of stress concentrations and an enhancement in toughness.

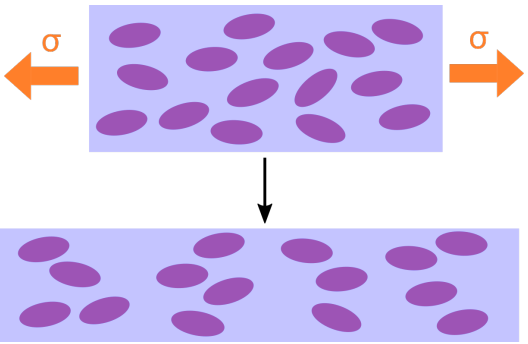

Figure 4.1: Scheme to illustrate the hypothesis of stress-induced order. Tensile loading can lead to ordering of the stiffer (purple) and softer (blue) components in a two-component system. The material builds up periodic density fluctuations featuring larger soft regions with fewer stiff particles, resulting in long-range order along the loading direction at a length scale larger than the stiffer components' dimension.

We here addressed this question for the nanoscale arrangement of crystallites within silk fibers. Silk is a protein-based high-performance semi-crystalline material. In contrast to synthetic semi-crystalline materials, crystalline regions are imprinted into the sequence, namely poly-alanine repeat units (Figure 1.3) of a length of 6-10 monomers [162, 163, 164]. They alternate with glycine-rich disordered regions of roughly three times the sequence length, which form the amorphous phase. The nanometer-sized crystalline units show a preferred orientation of the crystallites c-axis along the fiber direction [150, 165, 154], but are otherwise thought to be randomly positioned [150]. Both, Small-angle neutron (SANS) and X-ray scattering (SAXS) experiments show a mesoscopic long period in fiber direction at a length scale similar to the length scale of individual crystallites, which currently lacks a structural interpretation [166, 167, 168, 169, 170, 171]. Interestingly, we could previously show in a highly simplified finite element model that an ordered arrangement

of crystallites, with regions of high and low crystallinity alternating along the fiber axis, would increase the fiber's toughness [172].

We here present results from a purely atomistically-informed (bottom-up) finite element model of a silk fiber under stress and small-angle neutron scattering (SANS) experiments of *Nephila edulis* fibers. Both simulations and experiments unequivocally show an increase in long-range order of crystals along the fiber axis beyond the purely geometric effect of fiber elongation. Fluctuations of crystal density along the fiber axis increase upon stretching, thereby lowering mechanical energy. Our data suggests long-range order arising from mechanical work performed on the material to be vitally important to our understanding of silk fiber structure and mechanics, and potentially many other semi-crystalline or composite materials.

4.3 Methods

Finite element modelling

Model building

Fiber Geometry. The β-sheet nanocrystal size was defined as in previous FE studies of silk based on experimental data with dimensions in interchain (H-bonds) $\times$ intersheet $\times$ chain direction (fiber axis) of $1.9 \times 2.0 \times 2.7\,\mathrm{nm}^3$ [46, 172, 173] which falls in line with the length of the poly-alanine repeats present in the spider silk sequences [174]. Another range of nanocrystal sizes of $\sim 4.2 \times 2.3 \times 5.2\,\mathrm{nm}^3$ [175] was reported from X-ray diffraction (XRD) experiments and derived via the Scherrer equation (Equation 4.1) that relates peak broadening to crystallite size.

$$l = \frac{k\lambda}{\Delta(2\Theta)\cos\Theta} \tag{4.1}$$

where Θ is the Bragg-angle, $\Delta(2\Theta)$ the FWHM peak broadening, λ the X-ray wavelength, k the shape factor and l the crystallites mean size in perpendicular direction of the reflecting planes. The reported larger crystal dimensions could potentially originate from an induced densely-ordererd packing of residues in the glycine-rich sequence [174] in close vicinity to the highly ordered poly-alanine β-sheets. Due to parameters available only for the smaller crystal size, and the unknown influence of the non-β-sheet parts in the larger crystals on the strength of those, the fiber model was constructed using the small crystallites.

The geometry of the fiber defines the positions and arrangement of the crystals inside the amorphous matrix of the cylinder shaped fiber. We created fiber models with a length of 85 nm and a diameter of 22 nm, which roughly corresponds to half a nano-fibril [170]. We varied the crystallinity from 9–14 volume %, which is consistent with experimental observations [150, 176]. The crystals were randomly placed by a packing routine that only obeyed a minimum inter-crystal and crystal-to-fiber surface distance. The random placed crystals were further aligned along the fiber axis with an angle randomly chosen within a range of 0–30 ° giving an average of

∼15 °, as observed experimentally [150]. The maximum volume % of crystallinity that could be achieved with this packing approach is ∼18%. Larger crystallinities were achieved with packing algorithms such as Packmol [177], that incoorporate an optimization routine. The optimization process, however, resulted in a regular and aligned arrangement of crystals to accomodate the higher crystallinity and was thus not suitable.

In addition, as a reference, a model with a serial arrangement of crystal slices and an overall crystallinity of 24.7 % was created with alternating cross-sections, which were either fully crystalline or fully amorphous [173].

Meshing and Convergence. The generally preferred type of 3D mesh is consisting of hexahedral elements instead of tetrahedral elements. The advantages are a highly reduced number of elements and thus compute time for the same number of nodes as well as better accuracy of the finite-element calculations [178]. Disadvantages are possible volume and shear locking effects (hourglass effect) that need to be accounted for [65]. However, no general algorithm exists that can decompose and mesh arbitrary geometries with quadriliteral elements. The here presented model of a 3-dimensional fiber with randomly embedded crystallites is one such example. Successful attempts exist for a 2-dimensional fiber model [179]. We therefore had to resort to unstructured meshing algorithms.

The fiber geometry was meshed using GMSH 2.13.2 and the in-built unstructured 3D-Delaunay tetrahedral mesh algorithm [74], which is well suited for creating a coherent interface mesh between the amorphous and randomly embedded crystalline parts (Section 2.3.4). The mesh was then optimized by deploying the NETGEN algorithm [76], which increases element quality by displacing, reconnecting and/or removing nodes (Figure 4.2). The final mesh quality was evaluated by monitoring distributions of three criteria for tetrahedral element quality:

- η, the volume of the element to the power of 2/3 divided by the squared sum of the edge lengths.
- γ, the volume divided by the summed face areas and the maximal edge length.
- ρ, the minimum edge length divided by the maximum edge length of the tetrahedron.

The tetrahedral elements were defined as LS-DYNA element formulation 13, a 1 integration point (Section 2.3.4) constant stress element with nodal pressure averaging, which significantly alleviates volumetric locking. To obtain the minimal possible mesh resolution that still yields accurate results, the characteristic length of a finite element (l_c) was varied over the range of 0.45 to 1.2 nm. Fiber models with different mesh resolutions in this l_c range were pulled with a strain rate of 0.0175 ns^{-1} to a maximum strain of 1.75 % (strain velocity of 1.4875 nm/ns for a 85 nm long fiber) and the terminal stress was obtained. This terminal stress should converge with decreasing l_c. Figure 4.4 shows the mesh convergence for element

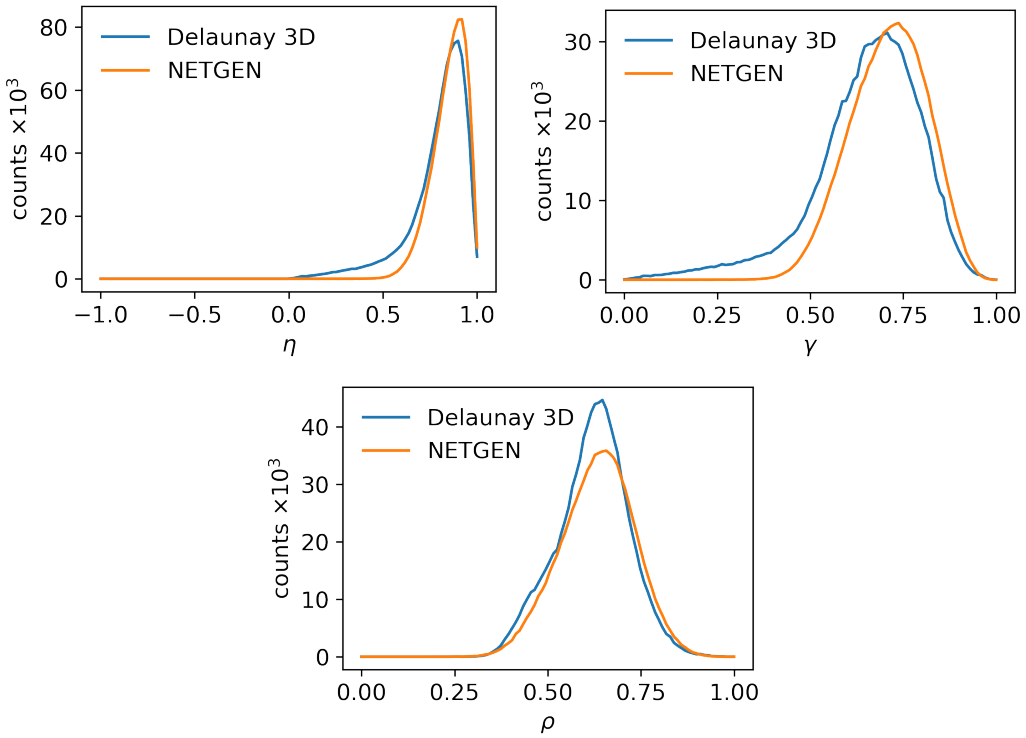

Figure 4.2: Mesh quality: distributions before (blue) & after (orange) NETGEN optimization for a representative fiber (10% crystallinity). The elements with the worst quality, *i.e.* lowest η, γ and ρ values were significantly reduced after optimization while the average values increased. Before optimization: $\gamma = 0.7847$ & $\rho = 0.6504$, after optimization: $\gamma = 0.8437$ & $\rho = 0.7085$.

formulation 13 and, for comparison, also the less well-suited LS-DYNA element formulation 4, a 1 point nodal pressure tetrahedron. To balance computational effort and accuracy, a value of $l_c = 0.55\,\text{nm}$ for the almost converged terminal stress was then used for the production simulations.

Parametrization. For the amorphous phase, a viscoelastic material model was used. Parameters for the density, elastic modulus, and viscosity were obtained by a bottom-up approach using Molecular Dynamics simulations of a representative atomistic building block of the amorphous phase as described previously [47]. The crystalline phase was modelled as a plastoelastic material, with a density, elastic modulus, and yield stress again obtained from Molecular Dynamics simulations as published earlier [180, 173, 46, 172]. All underlying Molecular Dynamics simulations were performed for the representative sequences from *Araneus diadematus* spider silk. We emphasize that this geometry and parametrization results in a highly simplified model of spider silk, as it considers only two phases: a fully crystalline and fully amorphous phase. Importantly, this model neglects, by design, partially crystalline intermediate regions [181, 150, 167, 153, 182] and other structural units

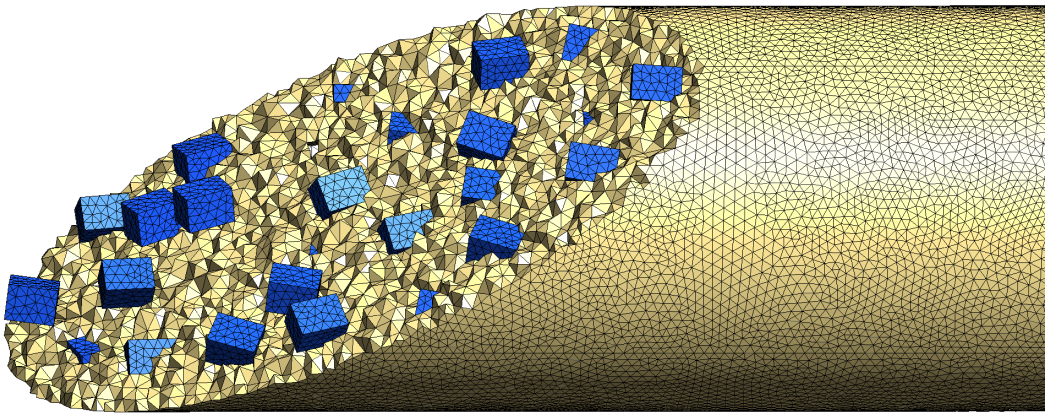

Figure 4.3: Cross section of meshed fiber with a crystallinity of 10 volume %. The crystals are randomly placed and rotated, but the tilt-angles against the fiber axis are constraint in the range of 0–30°.

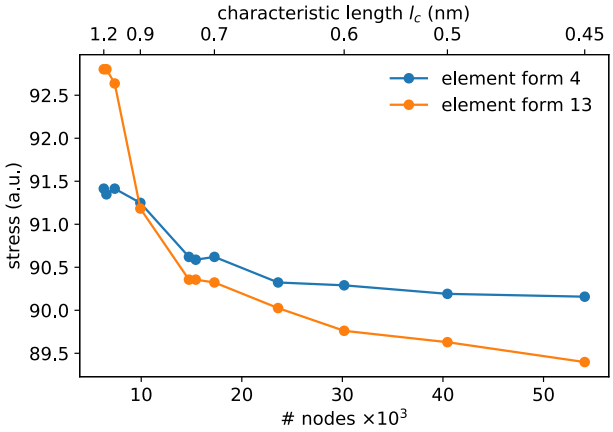

Figure 4.4: Mesh convergence for tetrahedral element formulations 4 & 13. The terminal stress after straining (1.75 %) is plotted against the number of nodes, which is determined by the characteristic length l_c of the meshing algorithm. A value of $l_c = 0.55\,\text{nm}$ and element formulation 13 were used for production runs.

such as the debated polyproline-II helices [183, 184, 185, 186, 187] that have been shown to exist. It also ignores any chain connectivity between the amorphous and crystalline phases, as we here aim to address a general phenomenon that is likely present in any two-phase material with stiff and soft components.

As another reference model that does not feature differences in the mechanical properties of the two phases, an all-amorphous model was used with a geometry and mesh identical to the fiber described above but with parameters of the amorphous phase throughout the whole fiber. In this case, the geometry is only relevant for the definition of scatterers (see below) but irrelevant for the tensile test.

Computational Tensile Tests

All finite element calculations were carried out using LS-DYNA R8.1.0 and results visualized and processed with LS-Prepost 4.3 as well as tools developed in Julia [188] and Python [44]. The fibers were subjected to a constant strain rate of $0.1\,\mathrm{ns}^{-1}$, which was comparable to the range used in experiments [189, 190], up to a maximum strain of $\varepsilon = 0.5$. The corresponding SANS experiments were done in quasi-equilibrium, but slow strain rates have been shown to not expose any significant differences in the stress-strain behaviour [191, 190]. Nodal positions and von-Mises nodal stresses were written out at a rate of $0.02\,\mathrm{ns}^{-1}$. As boundary conditions, the nodes of one cylinder surface were fixed, and the surface nodes on the other side were subjected to load. Fiber rupture was defined at the point when more than 5% of all crystal nodes exceeded the yield stress and underwent plastic deformation (Figures 4.10,4.11,4.9). We note that our model does not predict the strain-induced stiffening often observed in experiments [163], as the model does not capture the finite chain length of proteins in the amorphous phase nor the increase in crystallinity under tensile load (Figure 4.10).

Calculation of $I(\mathbf{q})$ from a FEM model

The scatter intensity is defined as (see Section 2.4)

$$I(\mathbf{q}) = F(\mathbf{q}) \cdot S(\mathbf{q}) \tag{4.2}$$

with $F(\mathbf{q})$ being the form factor and $S(\mathbf{q})$ the structure factor. These in turn are defined as

$$F(\mathbf{q}) = \frac{1}{N} \sum_{i,j}^{N} e^{-i\langle \mathbf{q}, \mathbf{r}_{ij} \rangle} \tag{4.3}$$

and

$$S(\mathbf{q}) = \frac{1}{N} \sum_{k,l}^{N} e^{-i\langle \mathbf{q}, \mathbf{r}_{kl} \rangle} \tag{4.4}$$

where $\mathbf{r}_{i,j}$ are scatterer positions within a single crystal, and $\mathbf{r}_{k,l}$ are the crystal positions as scatterers. This is similar to the WAXS pattern calculations previously done by Ulrich *et al.* [192].

Analogous to the experimental scatterers, which are the non-deuterated hydrogen atoms inside the crystals, we defined scatterer positions of the model similar to the mesh's nodes of the crystallites. Since non-structured tetrahedral meshs typically use a higher density of nodes (or smaller elements) at edges or regions of high curvatures such as our crystal surface edges, using the mesh node points as virtual scatterers (Figure 4.5b) would give an inaccurate analouge to the hydrogen atom scatterers used for the SANS experiment, where D_2O/H_2O exchange would first occur at the surface of the highly ordered and densely packed crystalline regions and replace hydrogen with deuterium. To avoid a higher number of virtual scatterers on the surface of our modelled crystals, we used the crystal geometry defined by the outer 8 edges to span a regular, equidistant set of virtual scatterers (Figure 4.5c) to calculate the scatter intensity of our modelled fiber. The distance in-between these virtual scatteres is of no importance to us, since they would be only seen at high q-values typically for wide-angle spectroscopy such as WAXS.

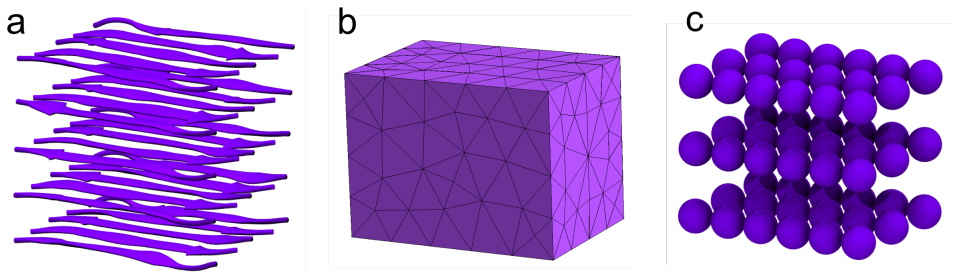

Figure 4.5: Crystalline unit in the FEM model. (a) Atomistic depiction of a spider silk crystalline subunit in cartoon representation comprised of 8 poly-alanine repeats, composing one β-strand, of which 5 in antiparallel orientation compose one β-sheet. Five β-sheets stacked in turn compose the crystal. (b) Model of the crystallite with tetrahedral volume mesh. (c) Virtual scatterer representation of one crystal unit used to calculate $I(\mathbf{q})$.

Additionally, the form factor can be calculated out of the geometric description of a single crystal of our FE model. This later allows to isolate the the structure factor and form factor independently and its effect on the overall intensity can be quantified.

Rotational averaging and angular integration were taken into account as shown in Figure 4.6 by averaging $S(\mathbf{q})$ over the rotation $\mathbf{R}(\zeta)$ of the angle around the fiber axis ζ and angualar integration over the rotation $\mathbf{R}(\chi)$ of the azimuthal angle χ:

$$I(\mathbf{q}) = \frac{1}{2\pi KN} \sum_{\zeta_k \in [0,2\pi)} \sum_{\chi_l \in [0,\pi/4]}^{K} \sum_{i,j=0}^{N} e^{-i\langle \mathbf{R}(\zeta_\mathbf{k},\chi_\mathbf{l},)\mathbf{q},\mathbf{r_{ij}}\rangle} \tag{4.5}$$

with K $= 2\pi/\Delta\zeta$. The experimental small-angle neutron scattering (SANS) data

were angular integrated with an opening angle of $\chi = \pi/2$ and $\chi = \pi/12$, likewise for the calculated data.

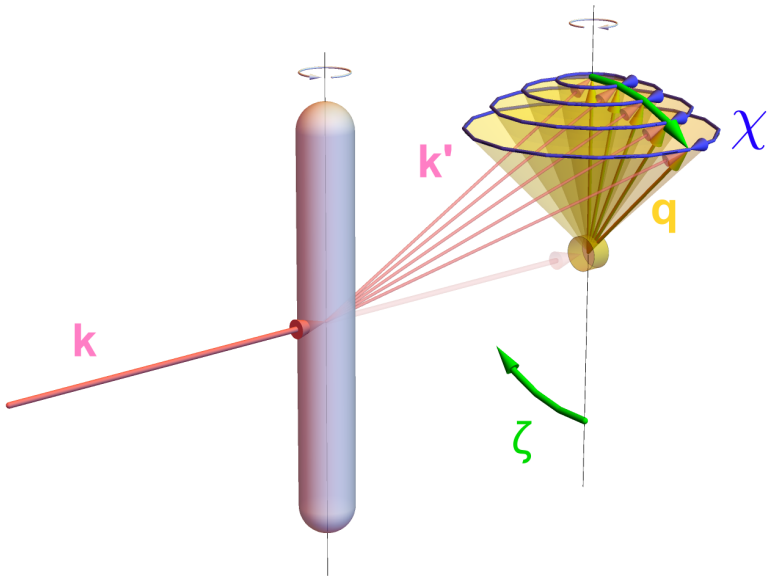

Figure 4.6: Scatter geometry of the silk fiber for an incoming neutron beam $\mathbf{k}$ and scattered beam $\mathbf{k}'$. Rotational averaging and angular integration were performed in the same range as for the experimentally recorded spectra for which the reason were to obtain a sufficient signal-to-noise ratio.

Experimental Details

The SANS experiments were performed by Imke Greving and Tilo Seydel at the Institute Laue-Langevin, Grenoble (ILL) for a fiber bundle of *Nephila edulis* spider silk at different strains. The fibers were humidified with D_2O, and the scattering contrast arose from the H/D-contrast between the water-accessible disordered (amorphous) regions and the crystallites. The latter are not accessible for H_2O and therefore do not undergo an H/D-exchange. A more detailed description of the experiment can be found in the recent publication from Krasnov *et al.* concerning strain-dependent diffusion originating from the same experiment [193].

Small-Angle Neutron Scattering Experiments. The *Nephila edulis* spider silk sample with a total mass at ambient humidity of $m_n = 104\,\mathrm{mg}$ was obtained by the forced silking technique [194] with a reeling speed of $10\,\mathrm{m/s}$. The sample consisted of a few uninterrupted dragline fibers wound around a pair of steel hooks with an initial gap between the inner edges of approximately $4\,\mathrm{cm}$. The sample thus formed a bundle with a width of approximately $1\,\mathrm{cm}$ and a thickness of $2.0 \times 2.7\,\mathrm{mm}$. It had previously been used in a neutron spectroscopy experiment (immediately prior

to the experiment), and the fibers were therefore previously strained and relieved from strain before the experiment.

The SANS experiments were carried using the beam line D11 at the high-flux research reactor of the ILL in Grenoble, France, using a wavelength of $\lambda = 6\,\text{Å}$. The beam size was defined by a rectangular slit 7 mm wide and 10 mm high, and the reported SANS patterns were recorded at a detector distance of 4 m.

The silk sample was mounted on a computer-controlled tensile machine [195, 182] inside a hermetically sealed chamber at ambient temperature with the silk fiber bundle axis vertically aligned. Within the sample chamber, a humid atmosphere saturated with heavy water (D_2O) vapor was set by filling small troughs located at the bottom of the chamber with heavy water. The saturated (100% RH) humidity level and temperature inside the chamber were monitored by adequate sensors.

SANS patterns were recorded at constant elongations, and the elongations between recordings were carried out at a constant strain rate of $0.01\,\text{mm}^{-1}$. Patterns were recorded immediately after reaching each new elongation with an integration time of 1 min. As soon as a stable SANS pattern was reached, the patterns were integrated over a 40 min recording time.

Experimental Scatter Intensity $I(\mathbf{q})$. The data were reduced using GRASP [196]. The background (empty sample cell) was subtracted from the data, and the runs were normalised for transmission in oder to account for the material, which is pulled out of the Field of View during the stretching process. Since the sample is measured within a container, the sample transmission is defined as the ratio of the transmission through the sample and container $T_{\text{s+c}}$ and the transmission of the empty container T_{c}. Similarly the empty beam transmission can be defined as ratio between the empty container T_{c} and the incident beam intensity or 'empty beam', T_{eb}.

$$T_{\text{s}} = \frac{T_{\text{s+c}}}{T_{\text{c}}} \quad T_e = \frac{T_{\text{c}}}{T_{\text{eb}}} \tag{4.6}$$

The intensity of the sample (fibres in cell) corrected for empty container scattering, background and electronic noise (I_{cd}) becomes

$$I_{\text{corr}} = \frac{1}{T_{\text{s}} T_{\text{e}}}(I_{\text{s}} - I_{\text{cd}}) - \frac{1}{T_{\text{e}}}(I_{\text{c}} - I_{\text{cd}}) \tag{4.7}$$

The fiber volume is assumed to be constant upon straining, *i.e.* the increased length coincides with radial thinning. Therefore, the fiber volume exposed to the neutron beam upon straining is decreasing, and can assumed to be proportional to the strain ε [197]. Since the scatter intensity $I(\mathbf{q})$ is proportional to the number of scatterers, $I(\mathbf{q})$ is corrected to:

$$I(\mathbf{q}, \varepsilon) = I(\mathbf{q})\,\varepsilon\,. \tag{4.8}$$

Another effect visible, against intuition, is that more water appears to be incoorporated into the fiber upon straining [182], therefore increasing the overall intensity.

Towards $q = 0$, some contrast from the amorphous phase but also atmosphere is seen additionally, which typically has a wide gaussian profile around $q = 0$ and is corrected by subtracting a linear background in our narrow regime of interest. A difference between the observations from SANS and FE calculations (see above) due to the two different spider species used for experiments (*Nephila edulis*) and for computational modelling (*Araneus diadematus*) can not be excluded. However, the two spider silks have closely related sequences, most importantly sharing the same length of poly-alanine repeats [162, 163, 164], and our simplified and rather generally valid FE model is unlikely to capture the sequence differences.

4.4 Results and Discussion

A range of composite or semi-crystalline materials consist of particles or crystallites embedded randomly in a much softer phase. Emergence of long-range order of these particles within the softer matrix could result in lowering the mechanical energy of the system upon stretching, in consistence with the well-known coalescence of defects in materials. Using small-angle neutron scattering (SANS) and finite element (FE) models we show the presence of such stress-induced ordering in spider silk fibers. Both methods show striking quantitative agreement of the position, shift and intensity increase of the long period upon stretching. We demonstrate that this mescoscopic ordering does not originate from strain-induced crystallization at the atomistic scale, and instead arises from a non-affine deformation that enhances density fluctuations of the two phases along the direction of stress. Our results suggest long-range ordering as a wide-spread phenomenon that can be exploited for tuning the mechanical properties of many hybrid materials with stiff and soft phases.

Crystals Distribute Periodically in Strained Silk Fibers.

To analyze potential stress-induced ordering in silk, we here resorted to a simplified two-phase finite element (FE) model of silk fiber comprising of the stiffer crystalline and the softer amorphous phase. We have developed this bottom-up FE model on the basis of atomistic MD simulations of *Araneus diadematus* silk protein [173, 46]. We here present a novel refined FE model, which describes the amorphous phase as a viscoelastic and the crystalline phase as a plastoelastic material, of which the viscosity and yield stress were obtained from MD simulations [180, 47, 48]. Crystals were randomly distributed into the amorphous phase with the experimentally known average tilt angle of $\sim 15°$ at crystallinities varying between 9 and 14 % (Figure 4.7a) [150, 176]. We applied constant strain rates to the fibers, and observed typical stress-strain curves with yield stress and strain, rupture stress and strain, and toughness values in agreement with experimental observations (see Table 4.1), validating our simplified model.

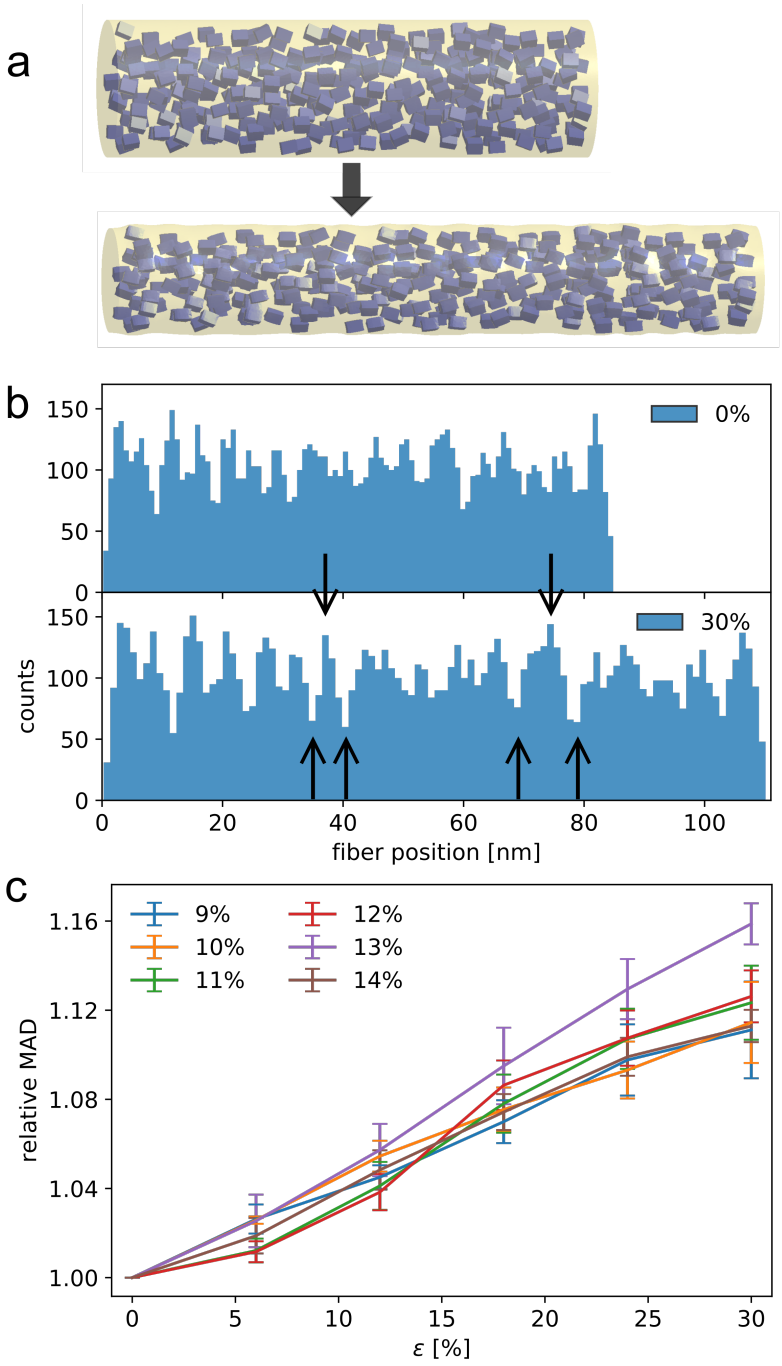

Figure 4.7: Increased long-range order in FE fiber models during loading. (a) A simplified 3D FE fiber model (see Suppl. Methods for details) with crystallites immersed into the amorphous phase was subjected to tensile load and the crystal distribution monitored. (b) Fluctuating crystallinity along the fiber axis of a fiber with 11 % overall crystallinity for two different strains. Arrows exemplify cross-sections with pronouncedly increased amplitudes in crystallinity variations during loading. (c) Mean absolute deviations (MAD) of crystallinity along the fiber axis as a function of external strain. For each overall crystallinity, values have been averaged over five individual fiber models.

Changes in structural order were also analyzed within the fiber under load. Figure 4.7 shows the crystallinity along the fiber axis for different strain values. The initial random packing transforms into a crystal distribution with pronounced peaks of higher crystallinity upon stretching (Figure 4.7a). The increase in crystallinity fluctuations was quantified by summing up the deviations from the mean crystallinity along the fiber axis (Figure 4.7b,c). We always observed a significant increase in these fluctuations upon straining. These fluctuations roughly lie in the 4–8 nm range, *i.e.* beyond the length of single crystals along the fiber axis, which here was chosen to be 2.7 nm [46]. This advocates a stress-induced ordering, resulting in more dense crystalline regions and resulting in wider amorphous patches (arrows in Figure 4.7b).

SANS Confirms Stress-Induced Order in Silk

Small-angle neutron scattering (SANS) was undertaken to test our computational prediction of stress-induced ordering in silk fibers. Deuterated fiber bundles using H/D exchange of *Nephila edulis* spider silk [170, 182] were subjected to increasing values of constant strains, for each of which a SANS pattern was recorded. An increase in scattering peak intensity on the nanometer length scale upon fiber stretch would reflect a long-range order of crystalline units within the flexible and thus more strongly deuterated amorphous phase. The two-dimensional SANS patterns (inset in Figure 4.8a) show the distinct meridional peaks as an indication of the mesoscopic long period in fiber direction previously reported [166, 167, 168, 169, 170, 171, 195] as well as a strong equatorial streak. The latter would yield additional information on the lateral (perpendicular to the fiber axis) arrangement of the nanocrystals in the spider silk fiber; however, in the context of the ordering effect in fiber direction, the respective analysis is not relevant and will be discussed elsewhere. Upon tensile load, the 2D SANS pattern exhibits two obvious changes: the position of the meridional peaks moves towards smaller scattering angles and the peaks become more pronounced. The integration of the two meridional peaks (symmetrically above and below the equator) over a 45 ° azimuthal angle each result in radial intensity distributions $I(q)$ where q is the modulus of the wave vector transfer defined as $q = \frac{4\pi}{\lambda} \sin\theta$ with the scattering angle 2θ. $I(q)$ measures correlations of distances $d = \frac{2\pi}{q}$ along the fiber axis between less deuterated regions, which here are the highly ordered β-sheet rich crystal units. The respective $I(q)$ curves are shown in Figure 4.8a for different values of the macroscopic strain ε. The second order peak (at $2q$) is extremely weak, which is consistent with SAXS experiments on wet silk fibers [167]. As already seen in the 2D pattern, the long-period peak intensity increases and shifts to smaller q upon stretch.

The increase in $I(q)$ of the long period observed in the SANS experiment is in line with the observations from the FE silk fiber model (Figure 4.7). For a more direct comparison, we calculated the SANS observable, $I(q)$, from the fiber model. More precisely, $I(q)$ was obtained from the ensemble averaged correlations of distances $d = \frac{2\pi}{q}$ between the mesh nodes within and between the crystals (Figure 4.8b, see

Suppl. Information for details). As expected, the finite size of the model causes high structure factors below $1\,\mathrm{nm}^{-1}$, and correlations within the crystalline units result in high structure factors in the $q = 10$–$13\,\mathrm{nm}^{-1}$ range (Figure 4.8b, inset). We can also recover the long period, *i.e.* the correlations of inter-crystal distances along the fiber axis, as observed in the SANS experiments, with high $I(q)$ in the range of $1\,\mathrm{nm}^{-1}$. We note that quantitative comparisons of absolute $I(q)$ values between SANS experiments and FE simulations are not possible, because the number of scatterers in the experiment are unknown. As in the SANS experiments, the maximum of $I(q)$ increases with strain and moves towards smaller q upon stretching.

For both, the SANS and FE long period, the q-position of the peak maximum was then used to calculate the corresponding d-spacing. Taking the value at no strain as reference $d_0 = d(\varepsilon = 0)$, we can define the microscopic small-angle strain as $\varepsilon_d(\varepsilon) = \frac{d(\varepsilon)-d_0}{d_0}$. The strain dependence of ε_d is plotted in Figure 4.8c and is clearly linear with a ratio of 0.86 (SANS) or 0.85–0.91 (FE) between ε_d and ε at maximal experimental ε. A value < 1 is associated with deformation by fibril slip [198], which is consistent with our simulation results assuming lateral friction within the amorphous phase of spider silk [47]. The lateral friction within the amorphous phase is connected to the increased unraveling of the random polymer chains upon strain, which is a process associated with friction and thus reflected in the spectroscopic response [47, 182].

Figure 4.8d shows the stress-induced change in the intensity maximum. In both the SANS experiments and FE calculations, independent from the initial arrangement and crystallinity of the model, we obtained a significant stress-induced increase in $I(q)$ of the long period. This finding indicates increased order of the periodic arrangement of the long period structure. To distinguish such stress-induced ordering from a purely geometric effect, we calculated the strain-dependent long periods also for a fictitious reference system, namely a homogeneous fiber with identical viscoelastic properties throughout (Figure 4.8d). In this case, the intensity of the long-period also steadily increases upon loading, but to a significantly smaller extent than in the SANS experiment or in the actual FE silk fiber models. Remarkably, the extent of ordering we observe in our silk fiber models is as high as the stress-induced ordering expected for a toy model of an already completely pre-ordered fiber. In this simple model, patches of high crystallinity move away from one another under tensile stress in an affine deformation. The toy model qualitatively confirms the strain-dependence of the position and height of the long-range peak observed in FE calculations and SANS experiments (Figure 4.8c,d and methods Section. 4.3):

$$\begin{aligned} q_{\max} &\sim 2\pi/(1+\varepsilon) \\ S(q_{\max}) &\sim 1 + N[1-(2\pi\sigma)^2/(1+\varepsilon)^2] \qquad (0 < \varepsilon \ll 1) \end{aligned} \tag{4.9}$$

Taken together, the combined results from experiments and simulations strongly support the notion of stress-induced ordering of the crystalline phase along spider silk fibers, which relies on the fact that one component is significantly stiffer than

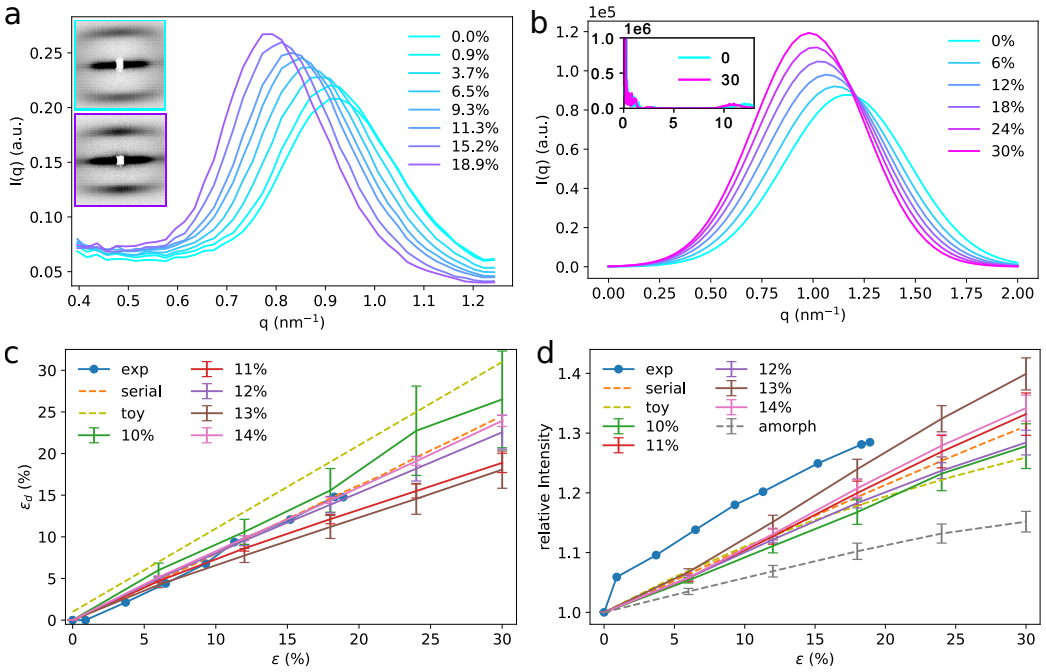

Figure 4.8: Scattering intensities from SANS and FE corroborate stress-induced long-range order. (a) Meridional long period peaks of $I(q)$ recorded by SANS on a D_2O humid spider silk sample for varying strain values ε in % elongation. Inset: 2D scattering image at no strain (top) and maximum strain (bottom), respectively. The peak position shifts towards lower q with increasing tensile load. In addition, the peak becomes slightly sharper. (b) Long period peaks of $I(q)$ calculated from FE calculations of a strained fiber model of 11 % crystallinity. Inset: $I(q)$ over a larger q-range also including peaks from intra-crystallite scattering. (c) Small-angle strain ε_d as manifested by the long period peak position versus macroscopic fiber strain ε, calculated from SANS (a) and the FE model (b). (d) Relative increase of scattering intensity, $I(q)/I_{\varepsilon=0}(q)$, as recorded by SANS or calculated for FE models of different crystallinity. As a reference, the intensity increase predicted by a simplified analytical model (see Suppl. Methods), by a fully amorphous model, in which the second phase serves as scatterer, and by a perfectly ordered serial model is also shown. (Note: the intensity increase from the analytical model was chosen to fall into the FE data by using N=60 in Equation (4.9))

the other, resulting in a phase separation. Regions rich in amorphous phase can more readily expand separating crystal-reinforced regions from one another and minimizing mechanical energy.

Mechanical Properties. Although this works focus is not on mechanical properties of spider-silk fibers, stress-strain curves were obtained from our FEM simulations. The stress-strain curves (Figure 4.10) are terminated by fiber rupture using the definition outlined above at which more than 5% of crystal elements have to

exceed their yield strength (Figure 4.9). Therefore, the rupture strain (ultimate strain), rupture stress (ultimate strength) and toughness (Figure 4.11) can be calculated and was found to be in alignment with experimental values (see Table 4.1).

Table 4.1: Mechanical properties of the investigated spider silk-fiber model as obtained from FEM simulations presented in this work and experimentally determined and reported in literature [199, 163, 200, 201].

Mechanical property	Simulations	Experiments
Ultimate Strength $\sigma_{\max}$ (GPa)	1.04 – 1.22	0.65 – 1.61
Ultimate Strain $\varepsilon_{\max}$	0.27 – 0.38	0.23 – 0.45
Toughness (MJ m^{-3})	125 – 158	120 – 225

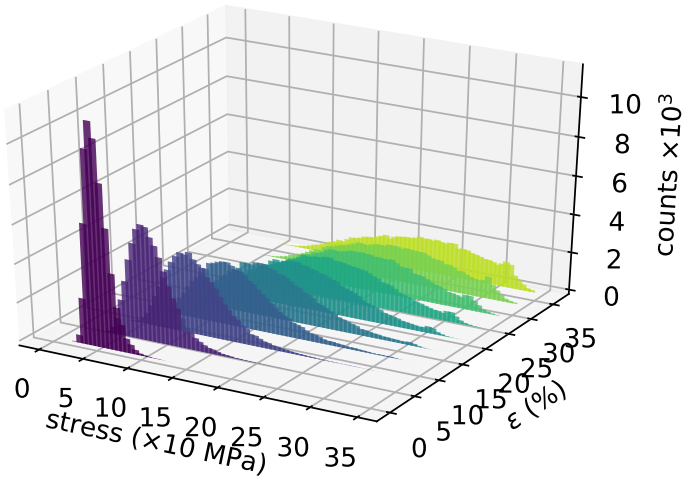

Figure 4.9: Histograms of von Mises-stresses in all crystal nodes for varying strains. Data for one fiber model with 11 % crystallinity is shown. We defined fiber rupture when more than 5 % of crystals are stressed beyond the defined yield stress of 0.3 GPa, which here results in a rupture strain of 32.8 %.

Analytical Toy Model

General Transformation of $S(\mathbf{q})$ under Deformation. Under a general deformation $\mathbf{f}$, which gives the new position $\mathbf{r}' = \mathbf{f}(\mathbf{r})$ of a particle that was initially located at $\mathbf{r}$, the structure factor $S(\mathbf{q})$ transforms as

$$S'(\mathbf{q}) = 1 + n' \int \mathrm{d}\mathbf{r}\, h'(\mathbf{r}) \mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}} = 1 + n' \int \mathrm{d}\mathbf{r}\, \det[D\mathbf{f}] h(\mathbf{r}) \mathrm{e}^{i\mathbf{q}\cdot\mathbf{f}(\mathbf{r})}\,, \tag{4.10}$$

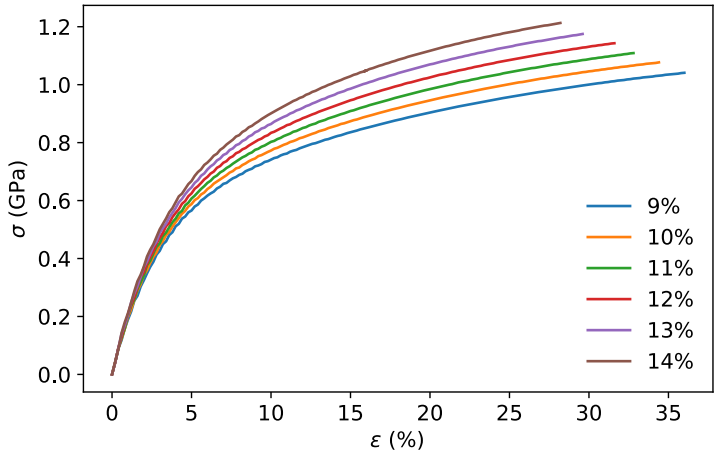

Figure 4.10: Representative stress-strain curves obtained from FE calculations of fiber models with varying crystallinity.

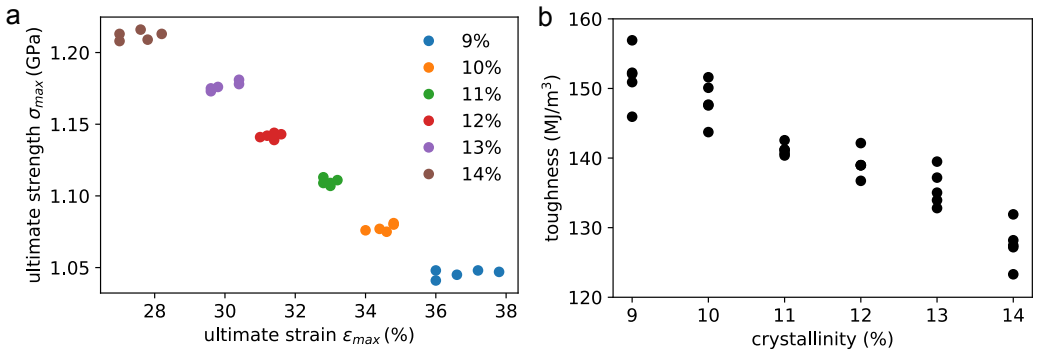

Figure 4.11: Mechanical properties for fiber models with different crystallinity: (a) ultimate strength $\sigma_{\max}$ vs. rupture (or ultimate) strain $\varepsilon_{\max}$ and (b) resulting toughness.

because the total correlation function after the deformation is given by $h'(\mathbf{r}) = h[\mathbf{f}^{-1}(\mathbf{r})]$. For an affine deformation $\mathbf{f}(\mathbf{r}) = \Lambda \cdot \mathbf{r}$, where the matrix Λ itself does not depend on the position $\mathbf{r}$, the average density n' of the deformed solution is given by $n' = n/\det\Lambda$ and Equation (4.10) therefore becomes

$$S'(\mathbf{q}) = 1 + n\int \mathrm{d}\mathbf{r}\, h(\mathbf{r}) \mathrm{e}^{i(\Lambda^{\mathrm{T}}\cdot\mathbf{q})\cdot\mathbf{r}} = S(\Lambda^{\mathrm{T}}\cdot\mathbf{q})\,. \tag{4.11}$$

This implies that the peak positions of the structure factor are changed, whereas the peak heights remain unchanged upon deformation. An extension of magnitude ε, *i.e.*, $\Lambda(\varepsilon) = \mathrm{diag}[1+\varepsilon, 1-\nu\varepsilon, 1-\nu\varepsilon]$ with Poisson ratio ν, for instance, characterized by its magnitude ε in x-direction and the Poisson ratio ν, one obtains that the strain-dependent structure factor reads $S(q,\varepsilon) = S[(1+\varepsilon)q, 0]$ if the scattering vector is parallel to the stretch direction, *i.e.*, $\mathbf{q} = q\mathbf{e}_x$.

Scattering from a Fiber. To investigate the scattering from a stretched silk fiber, we assume that the fiber is parallel to the x-axis and that the incoming beam points into the y-direction. Exploiting the systems symmetry, we use spherical and cylindrical coordinates

$$\mathbf{q} = q \begin{pmatrix} \sin\gamma \\ \cos\gamma \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{r} = \begin{pmatrix} x \\ \rho\cos\varphi \\ \rho\sin\varphi \end{pmatrix} , \tag{4.12}$$

for the scattering vector and the particle-particle distance vector, respectively, where is is assumed that the scattering profile is recorded in the x-y-plane. In this way the azimuthal angle of $\mathbf{q}$ can be set to 0. The polar angle γ between $\mathbf{q}$ and the incoming beam is related to $q \equiv |\mathbf{q}|$ and the wave number k of the beam as

$$q = 2k\cos\gamma\,, \tag{4.13}$$

which can be derived as follows. The wave vectors $\mathbf{k}_\mathrm{i}$ and $\mathbf{k}_\mathrm{s}$ of the incoming and the scattered wave have both the same absolute value k and differ in their direction by the scattering angle θ, *i.e.*, $\mathbf{k}_\mathrm{i} \cdot \mathbf{k}_\mathrm{s} \equiv k^2\cos\theta$. The length q of the scattering vector $\mathbf{q} = \mathbf{k}_\mathrm{s} - \mathbf{k}_\mathrm{i}$ is thus related to θ as $q^2 = 2k^2(1-\cos\theta)$. Inserting these relations into the defining equation $\cos\gamma \equiv -\mathbf{q}\cdot\mathbf{k}_\mathrm{i}/(qk) = (k/q)(1-\cos\theta)$ of the polar angle γ, we obtain Equation (4.13).

From the parametrization in Equation (4.12) follows $\mathbf{q}\cdot\mathbf{r} = q(x\sin\gamma + \rho\cos\gamma\cos\varphi)$ and with $h(\mathbf{r}) = h(x,\rho)$, which is independent of φ due to the rotational symmetry of the system around the fiber axis, we obtain

$$\begin{aligned} S'(\mathbf{q}) \equiv S(\mathbf{q},\varepsilon) &= 1 + n\int \mathrm{d}\mathbf{r}\, h(x,\rho,\varepsilon)\mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}} \\ &= 1 + 2\pi n \int \mathrm{d}x \int \mathrm{d}\rho\, \rho h(x,\rho,\varepsilon) J_0\left[q\rho^2/(2k)\right] \cos\left[qx\sqrt{1-q^2/(4k^2)}\right] \\ &\sim 1 + 2\pi n \int \mathrm{d}x \int \mathrm{d}\rho\, \rho h(x,\rho,\varepsilon)\cos(qx)\,, \end{aligned} \tag{4.14}$$

where the asymptotic scaling in the third line applies for $q \ll k$. In the second line, we performed the integral over φ, used Equation (4.13) to substitute q for γ, and removed the imaginary contributions exploiting that the result is invariant under the transition $x \mapsto -x$. For the silk experiments, the neutron beam has a wavelength of $2\pi/k = 0.6\,\mathrm{nm}$, the considered peak is located around $q \approx 0.1\,\mathrm{nm}^{-1}$, yielding a sufficiently small ratio $q/k \approx 0.01$ (corresponding to $\theta \approx 0.8\,°$), thereby justifying the use of the asymptotic relation in Equation (4.14).

Stretch-dependent Scattering from a Pre-ordered Fiber. To investigate how stretching affects the scattering from the fiber, an integrated total correlation function is introduced:

$$h(x,\varepsilon) \equiv (2\pi nL/N) \int \mathrm{d}\rho\, \rho h(x,\rho,\varepsilon)\,, \tag{4.15}$$

so that the structure factor in Equation (4.14) takes the more simple form

$$S(q,\varepsilon) = 1 + (N/L) \int \mathrm{d}x\, h(x,\varepsilon)\cos(qx)\,, \tag{4.16}$$

which suggests the toy model approach outlined below.

We describe the silk fiber as a one-dimensional chain of scatterers, obtained by projecting the three-dimensional scatterer positions onto the center line of the fiber. Their coordinates x along the fiber axis represent the interfaces between the matrix and the embedded crystallites. The function $h(x,\varepsilon)/L$ in Equation (4.16) then the distribution of the scatterer distances x. The structure of the silk material is characterized by spatial domains of high crystallite density.

The mean distance between neighbouring domains is set to 1 and denotes their size (*i.e.*, the length along the fiber axis) by 2σ. Coherent scattering from objects placed in two adjacent domains thus generates a peak of the structure factor around $q = 2\pi$. In the following, we address the influence of fiber stretching on the position and shape of this (first-order) domain-domain peak. The more densely packed conglomerates within the crystallite-rich domains are stiffer, while the matrix between the domains is more easily stretchable. In a first approximation, one can therefore take the relative positions of the scatterers within the domains to be conserved and the inter-domain distances to follow an externally imposed strain affinely.

Within our toy model approach, we represent the two neighboring domains by two Gaussian peaks in the scatterer density, which yields the following distribution of the scatterer distances:

$$h(x,\varepsilon)/L = \frac{1}{4\sigma\sqrt{2\pi}} \left[2\mathrm{e}^{-\frac{x^2}{2\sigma^2}} + \mathrm{e}^{-\frac{(x+1+\varepsilon)^2}{2\sigma^2}} + e^{-\frac{(x-1-\varepsilon)^2}{2\sigma^2}} \right]. \tag{4.17}$$

The ε-dependent peak positions account for the affine elongation of large distances between scatterers located in different domains, while the conservation of the scattering geometry within the same domain is represented by the ε-independent peak width σ. Inserting this distribution into Equation (4.16), the peak of the structure factor for the stretched chain is as follows:

$$S(q,\varepsilon) = 1 + N\cos^2\left[\frac{(1+\varepsilon)q}{2}\right] \mathrm{e}^{\frac{-\sigma^2 q^2}{2}}\,. \tag{4.18}$$

The toy model qualitatively confirms the strain-dependence of the position and height of the long-range peak observed in FE calculations and SANS experiments (Figure 4.8):

$$q_{\max} \sim 2\pi/(1+\varepsilon)\,, \quad S(q_{\max}) \sim 1+N[1-(2\pi\sigma)^2/(1+\varepsilon)^2] \qquad (0 < \varepsilon \ll 1) \tag{4.19}$$

4.5 Conclusions

By combining simulations and experiments, we here show that tensile forces acting on systems with two mechanically distinct phases can lead to an increase in long-range order. The ordering goes significantly beyond a purely geometric effect, which solely arises from the lengthening and thinning of the specimen. Remarkably, we obtained a striking quantitative agreement between the finite element simulations, the SANS experiments and a simple analytical model. Together, this strongly advocates stretch-induced ordering as a general principle of any two-phase system, being it semi-crystalline or a composite, and at varying length scales of the two phases, from the nanometer to micrometer scale or even beyond. Whenever one of the two phases is much softer than the other (or actually a defect [161]), the softer phase will more readily extend in regions less stiffened by the other phase, resulting in a mechanically induced phase segregation. This has analogies to the recent observations of stress-induced reorientation in hybrid soft and stiff liquid crystal elastomers [202]. We also show that stress-induced long-range ordering is largely independent from the extent of crystallinity fluctuations at zero stress. It does rely, however, on some pre-existing fluctuations in the distribution of crystals, such that order should not arise from a fiber with a perfectly homogenous crystal distribution. The observed ordering effect is in line with the very recent observation by quasi-elastic neutron scattering experiments of spider silk that the diffusive mobility of the amorphous phase increases with tensile strain [182]. We propose that tensile forces present during the spinning process of silk or other synthetic fibers enhance long-range pre-order, which in turn boosts further ordering during tensile loading. The parameter range of volume ratios, pre-order, and mechanical properties of the two phases within which stress-induced ordering is at play remains to be systematically quantified. We envision a wide range of materials, from nano-composites to block-copolymers, for which stress-induced long-range order can be exploited as a mean to tune mechanical properties.

Chapter 5

Summary & Conclusions

The aim of this thesis is to understand the influence of internal stress on molecular properties across multiple length scales. On a small molecular scale we investigated the kinetics of cycloaddition reactions for differently strained cyclic compounds reacting with tetrazines or azides and how it changes as a function of electron demand from tetrazine substitutents. On a larger, mesoscopic scale we studied the influence of stress on the structural topology for a two-phase fiber model system. This work was inspired by the many-fold occurences in which stress is the driving force for natural processes ranging from chemical reactions over biomechanical features of supramolecular components to biomolecular functions. The motivation to understand the complex and often hidden interplay between internal stress and its effects drove the study of two examples at different scales by exploiting knowledge and methods established in quantum mechanical and structural mechanical research.

In the following two paragraphs, we will discuss the findings of our investigation on the effects of internal stress in strained cyclic rings with respect to the reaction kinetics in cycloadditions and on the external stress applied to spider silk fiber models with respect to the rearrangement of the crystalline phase within the amorphous matrix.

Strain-promoted Click-reactions Strain can promote chemical reactions, a feature that has been very successfully explored in the field of click chemistry. Click-chemistry reactions in general and more specifically strain-promoted cycloadditions in the form of SPAAC or SPIEDAC are nowadays an essential and widely used tool for super resolution microscopy. Particularly, in the recently achieved triple-color labelling of biological entities [117], an ever more precise understanding and control over the reactions is needed to achieve desired properties such as mutual orthogonality of a set of reactions. Subtle differences in conformations and electronic properties of the reaction partners, such as caused by different isomeric states or substituents, can be quantified with computational methods, and by understanding them, the gained knowledge can be applied towards a better control of experimentally relevant reaction rates.

In the *first* study, we therefore investigated the kinetics of a large set of strain-promoted cyclocaddition reaction partners comprising of strained eight-rings, azides and tetrazines as determined by quantum mechanical calculation methods. In order to resemble experimental conditions, we included the handles of the strained

eight-rings attached to the protein and of the tetrazine attached to the fluorophore. We determined the DFT optimized geometries and energies of all strained eight-rings, tetrazine and azide reactants individually, obtained their transition state and van der Waals complex conformations and energies, as well as the frontier molecular orbitals and their energy levels. Our prediction of two energetically distinct groups of trans-cyclooctene stereoisomers, namely axial (TCO*a) and equatorial (TCO*e) TCO* isomers, was subsequently experimentally verified [132]. Furthermore the different isomeric states exibit significantly different reaction rates upon ligation with both tetrazines, leading to the conclusion that reaction kinetics can be tuned even by exploiting isomeric state differences. This becomes apparent not only in the electronic configuration, but also in the inherent different ring-strain energies of isomeric states with a considerable impact on the selective labelling of biological processes. Moreover, by taking into account regioselectivity of the handles we could show that their role is not only to attach to the fluorescent labels and proteins, but to decisively influence the reaction rate by their steric demand and electronic properties.

The *second* study expands upon the first part by systematically investigating the effect of electron demand on the reactive site of the diene modulated by the insertion of benzyl derivatives. We showed the impact of seven different benzyl derivatives ranging from heavily electron donating to heavily electron withdrawing groups on the reaction kinetics of cycloadditions with strained cyclooctynes. By correlating the FMO gap to the calculated activation energies we could demonstrate that FMO theory is not sufficient to explain differences in reaction kinetics of cycloaddition reactions. Moreover, small differences in the orbital gaps by substituting a hydrogen with a methyl group can result in drastic changes in the activation energies as well as differences in reaction rate by several orders of magnitude. A switch from inverse to normal electron demand as determined by the lowest of the two FMO gaps could not be observed in the activation energy. However, a consistent modulation of both, the inverse electron demand FMO gap as well as activation energy of the cycloaddition reaction, was found. We further identifed how the distance of an electron donating group from the reactive site of the diene influences the orbital energy levels and activation energies. Both showed to be considerably attenuated by increasing the spacing between the electron donating group and the benzyl ring through the insertion of a single methylene group.

This study ultimately demonstrates that reaction kinetics are defined not only by chemical effects, but also by geometric effects in terms of stereo- and regioselectiviy as well as strain energies. Such a combination of effects enlarges the parameter space available to re-engineer cycloaddition reactions to optimize them according to specfic experimental needs.

The success of our approach to calculate and predict accurate activation energies for a large set of cycloaddition reactions is promising for further investigations, such as identifying substituents that could modify the electronic configuration of the reactive sites to switch electron demand. An important step in further understanding strain promotion would be to quantify the ring strain in the considered

chemical species, by quantum mechanics based strain analyses as recently developed [203, 204].

Stress-induced Long-range Ordering in Spider Silk The mechanical behavior of materials is ultimately defined by their underlying molecular or macroscopic structure. The enormous progress made in the last decades in understanding and designing highly complex composite materials was accompanied by a shift from mostly tuning chemical composition to additionally considering structural features. In composite materials, the combination of a stiff and soft phase allows to design materials that can be tuned at different levels of structural composition inspired by natural counterparts such as silk or nacre.

We here addressed the question of how the application of stress induces changes in structural features of a two-phase fiber, for the example of a spider silk fiber. Our bottom-up approach of creating a model of the spider silk fiber incorporates parameters from atomistic molecular dynamics simulations of the highly ordered crystalline and disordered amorphous phase and integrates them into a finite-element model of a single fibril. The arrangement of crystals within the fiber was adopted from available experimental data. We subsequently performed computational tensile tests, subjecting the fiber models to strain rates similar to experimental values and investigated the response on the rearrangement of the crystals.

Despite our model ultimately originating from a spider silk protein sequence, the several levels of coarse graining in the bottom-up approach lead towards a simplified and generalized two-phase model allowing a conceptual study of the structural reconfiguration of the two phases triggered by the influence of stress.

By calculating the scatter intensity of the crystallites in a similar fashion to the experimentally recorded small-angle neutron scattering spectra of the fibers' cyrstalline phase, we found a remarkable agreement of this strain response. Together with the ability to quantify local crystallinity values from our model we could identify the presence of a strain-induced ordering of the crystal positions along the fiber axis. This ordering should be interpreted as an expansion of regions of the softer amorphous phase, rather than crystals being attracted towards each other. Furthermore, mechanical properties such as ultimate tensile strength and rupture strain and structural parameters such as the changes in average crystal tilt angles were found to be in agreement with previous experimental findings, corrobating the predictive power of our simple model.

Incorporating the gained knowledge into future designs of composite materials composed of two or more phases having different stiffnesses requires further systematic studies and parameter ranges such as volume ratios, shape factors and internal factors such as friction in-between the different phases. Also a sound theoretical description beyond the simple toy model presented in this thesis would enhance our understanding of stress-induced long-range ordering.

Taken together, the computational studies allow short glimpses into the wealth of strain promoted processes in molecular systems. We can expect a steady increase in

examples where strain very decisively influences structure, reactivity and function of molecules and materials, and believe that the methods employed within this thesis are a valuable addition to the experimental and simulation techniques required to tackle the challenge of comprehensively understand and make use of molecular strain. Expanding our mechanistic understanding of strain promoted processes will help to explore strain in chemistry, material science and beyond.

Acknowledgments

FOREMOST, I would like to express my sincere thanks to my two thesis advisors Prof. Dr. Dieter HEERMANN and Prof. Dr. Frauke GRÄTER from the University of Heidelberg. Without their support and guidance during my study and research this thesis would have not been possible.

Frauke has been my day-to-day supervisor and provided guidance throughout the years this research was undertaken. I am truly grateful for her sincere dedication, motivation and pragmatism. She never ceased to impress me with her swift and profound understanding of the most diverse scientific problems that arose during this work. I cannot imagine having a better supervisor than her.

I am very grateful to all the members of the (former) Klaus-Tschira-Lab within the CAS-MPG Partner Institute for Computational Biology in Shanghai, who helped to get me started in the world of computational biophysics and to understand biophysical systems, simulation and analysis techniques. Especially Dr. Scott EDWARDS and YANG Hui who were of invaluable help and never got tired of answering my many questions.

From HITS I am grateful for the amazing working environment, established by all the people working there, scientific or not. But first and foremost I have to thank the whole Molecular Biomechanics group. I am lucky to have found such a passionate and diverse group of people. Everyone was always happy to discuss problems and provide help in their very diverse fields of expertise.

For experimental data and fruitful discussions on the quantum chemistry part, I would like to acknowledge Dr. Edward LEMKE and his group member Dr. Ivana NIKIĆ, as well as Dr. Carsten SCHULTZ and his group member Dr. Jan-Erick HOFFMANN from EMBL Heidelberg. I thank Prof. Dr. Andreas DREUW from the University of Heidelberg for valuable discussions. For experimental data and many discussions on the silk fiber part, I would like to express my gratitude to Dr. Imke GREVING and Prof. Dr. Martin MÜLLER from the Helmholtz-Zentrum Geesthacht/Hamburg.

On a more personal note, I would like to thank Dr. Yang Hui, Roland Darjes and Martin Goutard for reminding me of the life outside of the PhD and among many others, made my stay in Shanghai and China unforgettable. I thank Sebastian Seifert for welcoming me in Mannheim from the very first day. I would especially like to thank Mercadante for the many breaks and discussions, cursing sessions and hopeless arguments and lastly for getting chemistry and pasta right. My eternal thanks goes to my family, my parents Marion and Andreas and my siblings Konstanze and Robert who are and were always there for support.

I am deeply indebted to my other half, Melanie, who supported and encouraged me at all times throughout the period of this Ph.D. research. You suffered a long time but accompanied me throughout this adventure and I look very much forward to our next one!

And finally, I would like to thank the Klaus-Tschira-Foundation, the Heidelberg Institute for Theoretical Studies and Toyota Motor Europe who provided full funding for this work.

Heidelberg, January 2017 Johannes Andreas Wagner

List of Figures

List of Tables

Abbreviations

BSSE	basis set superposition error
BCN	bicyclonyne
DFT	density functional theory
FE	finite element
FEA	finite element analysis
FEM	finite element method
FMO	frontier molecular orbital
GGA	generalized gradient approximation
GH	global hybrid
GTO	gaussian type orbital
HF	Hartree-Fock
KS	Kohn-Sham
LCAO	linear combination of atomic orbitals
LDA	local density approximation
LSDA	local spin density approximation
MD	molecular dynamics
QM	quantum mechanics
RSH	range-separated hybrid
SANS	small-angle neutron-scattering
SAXS	small-angle X-ray scattering
SCF	self consistent field
SCO	strained cyclooctyne
STO	slater-type orbital
TCO	trans-cyclooctene
TET	tetrazine
TS	transition state

List of Symbols

$\top$	Transposition
det	Determinant
∇	Nabla as column vector
Δ	Laplace operator
$\delta(.)$	Dirac delta function
$\|.\|$	Vector norm
k_B	Boltzmann constant
h	Planck constant
$\hbar$	reduced Planck constant
R_∞	Rydberg constant
Ha, H	Hartree
m_p	mass of proton
u	unified atomic mass unit
e	elementary charge
μ_N	nuclear magneton
E	Energy
t	Time
T	Temperature
σ	Stress
ε	Strain
η	Dynamical viscosity
N	Number of particles
$\hat{H}$	Hamiltonian
$\mathbf{k}$	Wavevector of incoming neutron
$\mathbf{k}'$	Wavevector of scattered neutron
$\mathbf{q}$	Momentum transfer vector of neutron
λ	Wavelength of neutron
I	Intensity
T	Transmission

Bibliography

[1] K. Gunasekaran, B. Ma, and R. Nussinov. Is allostery an intrinsic property of all dynamic proteins? *Proteins: Struct., Funct., Bioinf.*, 57(3):433–443, 2004. (Cited on page 1.)

[2] W. Stacklies, C. Seifert, and F. Gräter. Implementation of force distribution analysis for molecular dynamics simulations. *BMC Bioinf.*, 12(1):101, 2011. (Cited on page 1.)

[3] S. Xiao, S. A. Edwards, and F. Gräter. Nano-scale mechanics of nacre: Forces at protein-crystal linkages and flaws. *Biophys. J.*, 102(3):733a–734a, 2011. (Cited on page 1.)

[4] E. R. Cruz-Chu, S. Xiao, S. Patil, K. Gkagkas, and F. Gräter. Organic filling mitigates flaw-sensitivity of nanoscale aragonite. *ACS Biomater. Sci. Eng.*, 2016. (Cited on page 1.)

[5] D. E. Ingber. Tensegrity: the architectural basis of cellular mechanotransduction. *Annu. Rev. Physiol.*, 59(1):575–599, 1997. (Cited on page 1.)

[6] S. A. Edwards, J. A. Wagner, and F. Gräter. Dynamic prestress in a globular protein. *PLoS Comput. Biol.*, 8(5):e1002509, 2012. (Cited on pages 1, 8 and 23.)

[7] A. D. McNaught and A. Wilkinson. *Compendium of chemical terminology*, volume 1669. Blackwell Science Oxford, 1997. (Cited on pages 2 and 105.)

[8] LLC. Schrödinger. The pymol molecular graphics system, version 1.8. November 2015. (Cited on pages 2 and 105.)

[9] E. V. Anslyn and D. A. Dougherty. *Modern physical organic chemistry*. University Science Books, 2006. (Cited on pages 2 and 105.)

[10] O. Miyashita, J. N. Onuchic, and P. G. Wolynes. Nonlinear elasticity, protein-quakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 100(22):12570–12575, 2003. (Cited on page 2.)

[11] A. del Sol, C.-J. Tsai, B. Ma, and R. Nussinov. The origin of allosteric functional modulation: multiple pre-existing pathways. *Structure*, 17(8):1042–1050, 2009. (Cited on page 2.)

[12] H. Cao and A. G. Evans. Electric-field-induced fatigue crack growth in piezoelectrics. *J. Am. Ceram. Soc.*, 77(7):1783–1786, 1994. (Cited on page 3.)

[13] B. I. Costescu and F. Gräter. Graphene mechanics: II. Atomic stress distribution during indentation until rupture. *Phys. Chem. Chem. Phys.*, 16(24): 12582–12590, 2014. (Cited on page 3.)

[14] W. G. Johnston and J. J. Gilman. Dislocation velocities, dislocation densities, and plastic flow in lithium fluoride crystals. *J. Appl. Phys.*, 30(2):129–144, 1959. (Cited on page 3.)

[15] D. M. Marsh. Plastic flow in glass. In *Proc. R. Soc. A*, pages 420–435. The Royal Society, 1964. (Cited on page 3.)

[16] V.V. Martynov and V.V. Kokorin. The crystal structure of thermally-and stress-induced martensites in ni2mnga single crystals. *J. Phys. III*, 2(5):739–749, 1992. (Cited on page 3.)

[17] G. Liu, X. Zhang, and D. Wang. Stress induced reversible crystal transition in polymers. *Polym. Int.*, 64(8):951–956, 2015. (Cited on page 3.)

[18] L. Morbitzer and H. Hespe. Correlations between chemical structure, stress-induced crystallization, and deformation behavior of polyurethane elastomers. *J. Appl. Polym. Sci.*, 16(10):2697–2708, 1972. (Cited on page 3.)

[19] S. Toki, T. Fujimaki, and M. Okuyama. Strain-induced crystallization of natural rubber as detected real-time by wide-angle x-ray diffraction technique. *Polymer*, 41(14):5423–5429, June 2000. (Cited on pages 3 and 79.)

[20] B. Huneau. Strain-induced crystallization of natural rubber: a review of x-ray diffraction investigations. *Rubber Chem. Technol.*, 84(3):425–452, 2011. (Cited on page 3.)

[21] P. Asztalos, C. Parthier, R. Golbik, M. Kleinschmidt, G. Hübner, M. S. Weiss, R. Friedemann, G. Wille, and K Tittmann. Strain and near attack conformers in enzymic thiamin catalysis: X-ray crystallographic snapshots of bacterial transketolase in covalent complex with donor ketoses xylulose 5-phosphate and fructose 6-phosphate, and in noncovalent complex with acceptor aldose ribose 5-phosphate. *Biochemistry*, 46(43):12037–12052, 2007. (Cited on page 3.)

[22] D. A. Davis, A. Hamilton, J. Yang, L. D. Cremar, D. Van Gough, S. L. Potisek, M. T. Ong, P. V. Braun, T. J. Martínez, S. R. White, et al. Force-induced activation of covalent bonds in mechanoresponsive polymeric materials. *Nature*, 459(7243):68–72, 2009. (Cited on page 3.)

[23] C. R. Hickenboth, J. S. Moore, S. R. White, N. R. Sottos, J. Baudry, and S. R. Wilson. Biasing reaction pathways with mechanical force. *Nature*, 446 (7134):423–427, 2007. (Cited on page 3.)

[24] R. M. Alexander and H. C. Bennet-Clark. Storage of elastic strain energy in muscle and other tissues. *Nature*, 265(5590):114–117, 1977. (Cited on page 4.)

[25] K. B. Wiberg. The concept of strain in organic chemistry. *Angew. Chem. Int. Ed. (English)*, 25(4):312–322, 1986. (Cited on page 4.)

[26] N. J. Agard, J. A. Prescher, and C. R. Bertozzi. A strain-promoted [3+ 2] azide- alkyne cycloaddition for covalent modification of biomolecules in living systems. *âJ. Am. Chem. Soc.*, 126(46):15046–15047, 2004. (Cited on page 4.)

[27] M. B Smith and J. March. *March's advanced organic chemistry: reactions, mechanisms, and structure*. John Wiley & Sons, 2007. (Cited on page 4.)

[28] F. Jensen. *Introduction to computational chemistry*. John Wiley & Sons Ltd, 1999. (Cited on pages 9, 10 and 11.)

[29] B. Mennucci. Polarizable continuum model. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2(3):386–404, 2012. (Cited on pages 9 and 17.)

[30] M. Born and V. Fock. Beweis des adiabatensatzes. *Zeitschrift für Physik*, 51 (3-4):165–180, 1928. (Cited on page 10.)

[31] D. R. Hartree. The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods. In *Mathematical Proc. Cambridge Phil. Soc.*, volume 24, pages 89–110. Cambridge Univ Press, 1928. (Cited on page 11.)

[32] V. Fock. Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. *Zeitschrift für Physik*, 61(1-2):126–148, 1930. (Cited on page 11.)

[33] P. M. W. Gill. Density functional theory (dft), hartree-fock (hf), and the self-consistent field. *J. Chem. Phys*, 100:5066–5075, 1994. (Cited on page 13.)

[34] P. Echenique and J. L. Alonso. A mathematical and computational review of hartree–fock scf methods in quantum chemistry. *Mol. Phys.*, 105(23-24): 3057–3098, 2007. (Cited on page 13.)

[35] J. Foresman and E. Frish. Exploring chemistry. *Gaussian Inc., Pittsburg, USA*, 1996. (Cited on page 13.)

[36] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phy. Rev. Lett.*, 91(14):146401, 2003. (Cited on page 14.)

[37] Y. Zhao and D. G. Truhlar. Applications and validations of the minnesota density functionals. *Chem. Phys. Lett.*, 502(1):1–13, 2011. (Cited on page 14.)

[38] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *J. Chem. Phys.*, 132(15):154104, 2010. (Cited on pages 14 and 15.)

[39] M. Cossi, V. Barone, R. Cammi, and J. Tomasi. Ab initio study of solvated molecules: a new implementation of the polarizable continuum model. *Chem. Phys. Lett.*, 255(4-6):327–335, 1996. (Cited on page 17.)

[40] B. Mennucci, E. Cances, and J. Tomasi. Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: theoretical bases, computational implementation, and numerical applications. *J. Phys. Chem. B*, 101(49):10506–10517, 1997. (Cited on page 18.)

[41] K. Fukui, T. Yonezawa, and H. Shingu. A molecular orbital theory of reactivity in aromatic hydrocarbons. *J. Chem. Phys.*, 20(4):722–725, 1952. (Cited on page 18.)

[42] K. Fukui. The role of frontier orbitals in chemical reactions (nobel lecture). *Angew. Chem. Int. Ed. (English)*, 21(11):801–809, 1982. (Cited on page 18.)

[43] C. Rostgaard, K. W. Jacobsen, and K. S. Thygesen. Fully self-consistent gw calculations for molecules. *Phys. Rev. B*, 81(8):085103, 2010. (Cited on page 19.)

[44] J. D. Hunter. Matplotlib: A 2d graphics environment. *Comput. In Sci. & Eng.*, 9(3):90–95, 2007. doi: 10.1109/MCSE.2007.55. (Cited on pages 20, 85 and 106.)

[45] F. M. Bickelhaupt. Understanding reactivity with kohn–sham molecular orbital theory: E2–sn2 mechanistic spectrum and other concepts. *J. Comput. Chem.*, 20(1):114–128, 1999. (Cited on page 21.)

[46] 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, May 2009. (Cited on pages 23, 81, 83, 89 and 91.)

[47] S. P. Patil, B. Markert, and F. Gräter. Rate-dependent behavior of the amorphous phase of spider dragline silk. *Biophys. J.*, 106(11):2511–2518, June 2014. (Cited on pages 23, 83, 89 and 92.)

[48] S. P. Patil, S. Xiao, K. Gkagkas, B. Markert, and F. Gräter. Viscous friction between crystalline and amorphous phase of dragline silk. *PLoS ONE*, 9(8): e104832, August 2014. (Cited on pages 23 and 89.)

[49] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*, volume 1. Academic press, 2001. (Cited on page 23.)

[50] L. Monticelli and E. Salonen. *Biomolecular simulations: methods and protocols*. Humana Press New York, 2013. (Cited on page 23.)

[51] M. J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GROMACS development team. *GROMACS User Manual version 2016.2*. www.gromacs.org, 2016. (Cited on page 23.)

[52] R. W. Hockney, S. P. Goel, and J. W. Eastwood. Quiet high-resolution computer models of a plasma. *J. Comput. Phys.*, 14(2):148–158, 1974. (Cited on page 25.)

[53] C. M. Baker. Polarizable force fields for molecular dynamics simulations of biomolecules. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 5(2):241–254, 2015. (Cited on page 25.)

[54] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard. Reaxff: a reactive force field for hydrocarbons. *J. Phys. Chem. A*, 105(41):9396–9409, 2001. (Cited on page 25.)

[55] W. F. Van Gunsteren and H. J. C. Berendsen. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.*, 34(5):1311–1327, 1977. (Cited on page 26.)

[56] B. Hess, H. Bekker, H. J . C. Berendsen, J. G. E. M. Fraaije, et al. Lincs: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, 18(12): 1463–1472, 1997. (Cited on page 26.)

[57] H. J. C. Berendsen, J. P. M. van Postma, W. F. van Gunsteren, A. R. H. J. Dinola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81(8):3684–3690, 1984. (Cited on page 26.)

[58] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126(1):014101, 2007. (Cited on page 26.)

[59] T. J. Lane, Di. Shukla, K. A Beauchamp, and V. S. Pande. To milliseconds and beyond: challenges in the simulation of protein folding. *Curr. Opi. in Struc. Biol.*, 23(1):58–65, 2013. (Cited on page 26.)

[60] M. Feig, R. Harada, T. Mori, I. Yu, K. Takahashi, and Y. Sugita. Complete atomistic model of a bacterial cytoplasm for integrating physics, biochemistry, and systems biology. *J. Mol. Graphics and Model.*, 58:1–9, 2015. (Cited on page 26.)

[61] G. Groenhof. Introduction to qm/mm simulations. *Biomolecular simulations: methods and protocols*, pages 43–66, 2013. (Cited on page 26.)

[62] S. J. Marrink, J. H. Risselada, S. Yefimov, P. D. Tieleman, and A. H. De Vries. The martini force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B*, 111(27):7812–7824, 2007. (Cited on page 26.)

[63] R. J. Wagner, L. Ma, F. Tavazza, and L. E. Levine. Dislocation nucleation during nanoindentation of aluminum. *J. Appl. Phys.*, 104(11):114311, 2008. (Cited on page 26.)

[64] P. Haupt. *Continuum mechanics and theory of materials*. Springer Science & Business Media, 2013. (Cited on page 27.)

[65] Livermore Software Technology Corporation. *LS-DYNA Theory Manual*. Livermore Software Technology Corporation, 2016. (Cited on pages 27, 31, 35 and 82.)

[66] Sandeep Parasharam Patil. Multiscale modeling of spider dragline silk. Dr. IAM-01 (2015), Aachen, 2015. URL https://publications.rwth-aachen.de/record/483749. Zugl.: Aachen, Techn. Hochsch., Diss., 2015. (Cited on page 27.)

[67] Livermore Software Technology Corporation. *LS-DYNA KEYWORD USER'S MANUAL, Volume II Material Models*. Livermore Software Technology Corporation, 2015. (Cited on pages 30 and 31.)

[68] L. R. Herrmann and F. E. Peterson. A numerical procedure for viscoelastic stress analysis. In *7th meeting of ICRPG mech. behavior work. group, Orlando, FL, CPIA Publ.*, number 177, pages 60–69, 1968. (Cited on page 30.)

[69] K. Hackl and M. Goodarzi. An introduction to linear continuum mechanics. *Ruhr-University Bochum*, 2010. (Cited on page 31.)

[70] O. C. Zienkiewicz and R. L. Taylor. *The finite element method: solid mechanics*, volume 2. Butterworth-heinemann, 2000. (Cited on page 31.)

[71] D. V. Hutton. *Fundamental of Finite Element Analysis, Elizabeth A*. Jonse Ltd., 2004. (Cited on page 31.)

[72] D. L. Logan. *A first course in the finite element method*. Cengage Learning, 2011. (Cited on page 31.)

[73] P. J. Frey and P.-L. George. *Mesh generation: application to finite elements*. ISTE, 2007. (Cited on page 33.)

[74] C. Geuzaine and J.-F. Remacle. Gmsh: A 3-d finite element mesh generator with built-in pre- and post-processing facilities. *Int. J. Numer. Meth. Eng.*, 79(11):1309–1331, September 2009. (Cited on pages 33, 82 and 106.)

[75] S. Rebay. Efficient unstructured mesh generation by means of delaunay triangulation and bowyer-watson algorithm. *J. Comput. Phys.*, 106(1):125–138, 1993. (Cited on page 33.)

[76] J. Schöberl. Netgen an advancing front 2d/3d-mesh generator based on abstract rules. *Comput. Vis. Sci.*, 1(1):41–52, 1997. (Cited on pages 33 and 82.)

[77] J.-F. Remacle, J. Lambrechts, B. Seny, E. Marchandise, A. Johnen, and C. Geuzainet. Blossom-quad: A non-uniform quadrilateral mesh generator using a minimum-cost perfect-matching algorithm. *Int. J. Numer. Methods Eng.*, 89(9):1102–1119, 2012. (Cited on page 33.)

[78] R. Schneiders. Algorithms for quadrilateral and hexahedral mesh generation. *Proc. of the VKI Lect. Ser. on Comput. Fluid Dynamic, VKI-LS*, 4, 2000. (Cited on page 33.)

[79] K. Nakamura and Particle Data Group. The review of particle physics. *J. Phys. G: Nucl. Part. Phys.*, 37(and 2011 partial update for the 2012 edition): 075021, 2010. URL `http://pdg.lbl.gov`. (Cited on pages 37 and 113.)

[80] S. W. Lovesey. *Theory of Neutron Scattering from Condensed Matter*, volume 1 and 2. Oxford University Press, 2004. (Cited on page 37.)

[81] M. Hennig. *Dynamics of Globular Proteins in Crowded Electrolyte Solutions - Studied by Neutron Scattering*. PhD thesis, Universität Tübingen, 2011. (Cited on pages 37 and 113.)

[82] V. F. Sears. Neutron scattering lengths and cross sections. *Neutron news*, 3 (3):26–37, 1992. (Cited on page 37.)

[83] E. F. Pettersen, T. D Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. Ucsf chimeraâa visualization system for exploratory research and analysis. *J. Comput. Chem.*, 25(13):1605–1612, 2004. (Cited on pages 42 and 106.)

[84] H. C. Kolb, M. G. Finn, and B. K. Sharpless. Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed. (English)*, 40(11): 2004–2021, 2001. (Cited on pages 41, 43 and 44.)

[85] J. E. Moses and A. D. Moorhouse. The growing applications of click chemistry. *Chem. Soc. Rev.*, 36(8):1249–1262, 2007. (Cited on page 43.)

[86] H. C. Kolb and B. K. Sharpless. The growing impact of click chemistry on drug discovery. *Drug Disc. today*, 8(24):1128–1137, 2003. (Cited on page 43.)

[87] R. Schirrmacher, C. Wangler, and E. Schirrmacher. Recent developments and trends in 18f-radiochemistry: syntheses and applications. *Mini-Rev. Org. Chem.*, 4(4):317–329, 2007. (Cited on page 43.)

[88] K. Nwe and M. W. Brechbiel. Growing applications of âclick chemistryâ for bioconjugation in contemporary biomedical research. *Cancer Biother. Radiopharm.*, 24(3):289–302, 2009. (Cited on page 43.)

[89] L. Liang and D. Astruc. The copper (i)-catalyzed alkyne-azide cycloaddition (cuaac)âclickâ reaction and its applications. an overview. *Coordination Chemistry Reviews*, 255(23):2933–2945, 2011. (Cited on pages 43 and 44.)

[90] J.-P. Meyer, P. Adumeau, J. S. Lewis, and B. M. Zeglis. Click chemistry and radiochemistry: The first 10 years. *Bioconjugate Chem.*, 2016. (Cited on page 43.)

[91] J. A. Prescher and C. R. Bertozzi. Chemistry in living systems. *Nat. Chem. Biol.*, 1(1):13–21, 2005. (Cited on page 43.)

[92] H. Nandivada, X. Jiang, and J. Lahann. Click chemistry: Versatility and control in the hands of materials scientists. *Adv. Mater.*, 19(17):2197–2208, September 2007. (Cited on page 43.)

[93] W. H. Binder and R. Sachsenhofer. 'click' chemistry in polymer and material science: An update. *Macromol. Rapid Commun.*, 29(12–13):952–981, July 2008. (Cited on page 43.)

[94] M. A. Azagarsamy and K. S. Anseth. Bioorthogonal click chemistry: an indispensable tool to create multifaceted cell culture scaffolds. *ACS Macro Lett.*, 2(1):5–9, 2012. (Cited on page 43.)

[95] W. Xi, T. F. Scott, C. J. Kloxin, and C. N. Bowman. Click chemistry in materials science. *Adv. Funct. Mater.*, 24(18):2572–2590, 2014. (Cited on page 43.)

[96] R. Huisgen. Introduction, survey, mechanism. In *1, 3-Dipolar Cycloaddition Chemistry*, pages 1–176. Wiley, 1984. (Cited on page 43.)

[97] A. D. McNaught and A. Wilkinson. *IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book").* Blackwell Scientific Publications, Oxford, 1997. (Cited on page 43.)

[98] M. Nic, J. Jirat, B. Kosata, and A. Jenkins. *XML online corrected version: http://goldbook.iupac.org.* 2006-. ISBN ISBN 0-9678550-9-8. doi: doi:10.1351/goldbook. (Cited on page 43.)

[99] I. Fleming. *Pericyclic reactions.* Oxford University Press, USA, 2015. (Cited on page 43.)

[100] Neil E Schore. Transition metal-mediated cycloaddition reactions of alkynes in organic synthesis. *Chemical Reviews*, 88(7):1081–1119, 1988. (Cited on page 43.)

[101] G. C. Tron, T. Pirali, R. A. Billington, P. L. Canonico, G. Sorba, and A. A. Genazzani. Click chemistry reactions in medicinal chemistry: Applications of the 1,3-dipolar cycloaddition between azides and alkynes. *Med. Res. Rev.*, 28 (2):278–308, 2008. (Cited on page 43.)

[102] V. V. Rostovtsev, L. G. Green, V. V. Fokin, and B. K. Sharpless. A stepwise huisgen cycloaddition process: copper (i)-catalyzed regioselective âligationâ of azides and terminal alkynes. *Angew. Chem.*, 114(14):2708–2711, 2002. (Cited on page 43.)

[103] F. Himo, T. Lovell, R. Hilgraf, V. V. Rostovtsev, L. Noodleman, B. K. Sharpless, and V. V. Fokin. Copper (i)-catalyzed synthesis of azoles. dft study predicts unprecedented reactivity and intermediates. *âJ. Am. Chem. Soc.*, 127(1):210–216, 2005. (Cited on page 43.)

[104] J. Sauer and H. Wiest. Diels-alder-additionen mit âinversemâ ï¸ elektronenbedarf. *Angew. Chem.*, 74(10):353–353, 1962. (Cited on page 44.)

[105] G. J. Bodwell, Z. Pi, and I. R. Pottie. Electron deficient dienes. 2. one step synthesis of a coumarin-fused electron deficient diene and its inverse electron demand diels-alder reactions with enamines. *Synlett.*, 1999(04):477–479, 1999. (Cited on page 44.)

[106] N. J. Agard, J. A. Prescher, and C. R. Bertozzi. A strain-promoted [3 + 2] azideâalkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.*, 126(46):15046–15047, November 2004. (Cited on pages 44 and 58.)

[107] T. Misteli and D. L. Spector. Applications of the green fluorescent protein in cell biology and biotechnology. *Nature Biotech.*, 15(10):961–964, 1997. (Cited on page 44.)

[108] P. J. Hudson and C. Souriau. Engineered antibodies. *Nature Med.*, 9(1): 129–134, 2003. (Cited on page 44.)

[109] N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo, and C. R. Bertozzi. A comparative study of bioorthogonal reactions with azides. *ACS Chem. Biol.*, 1(10):644–648, November 2006. (Cited on pages 44 and 58.)

[110] M. D. Best. Click chemistry and bioorthogonal reactions: Unprecedented selectivity in the labeling of biological molecules. *Biochem.*, 48(28):6571–6584, July 2009. (Cited on page 44.)

[111] L. I. Willems, W. A. van der Linden, N. Li, K.-Y. Li, N. Liu, S. Hoogendoorn, G. A. van der Marel, B. I. Florea, and H. S. Overkleeft. Bioorthogonal chemistry: Applications in activity-based protein profiling. *Acc. Chem. Res.*, 44 (9):718–729, September 2011. (Cited on page 44.)

[112] C. P. Ramil and Q. Lin. Bioorthogonal chemistry: strategies and recent developments. *Chem. Commun.*, 49(94):11007, 2013. (Cited on page 44.)

[113] D. S. Liu, A. Tangpeerachaikul, R. Selvaraj, M. T. Taylor, J. M. Fox, and A. Y. Ting. Diels–alder cycloaddition for fluorophore targeting to specific proteins inside living cells. *âJ. Am. Chem. Soc.*, 134(2):792–795, 2012. (Cited on page 49.)

[114] N. K. Devaraj, R. Weissleder, and S. A. Hilderbrand. Tetrazine-based cycloadditions: Application to pretargeted live cell imaging. *Bioconjugate Chem.*, 19 (12):2297–2299, December 2008. (Cited on page 49.)

[115] R. K. Iha, K. L. Wooley, A. M. Nystrom, D. J. Burke, M. J. Kade, and C. J. Hawker. Applications of orthogonal "click" chemistries in the synthesis of functional soft materials. *Chem. Rev.*, 109(11):5620–5686, November 2009. (Cited on page 49.)

[116] M. R. Karver, R. Weissleder, and S. A. Hilderbrand. Bioorthogonal reaction pairs enable simultaneous, selective, multi-target imaging. *Angew. Chem. Int. Ed. (English)*, 51(4):920–922, January 2012. (Cited on page 49.)

[117] I. Nikic, T. Plass, O. Schraidt, J. Szymanski, J. A. G. Briggs, C. Schultz, and E. A. Lemke. Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy. *Angew. Chem. Int. Ed. (English)*, 53(8):2245–2249, January 2014. (Cited on pages 49, 51, 56, 58 and 99.)

[118] J. Sauer, D. K. Heldmann, J. Hetzenegger, J. Krauthan, H. Sichert, and J. Schuster. 1,2,4,5-tetrazine: Synthesis and reactivity in [4+2]cycloadditions. *Eur. J. of Org. Chem.*, 1998(12):2885–2896, December 1998. (Cited on page 49.)

[119] M. L. Blackman, M. Royzen, and J. M. Fox. Tetrazine ligation: Fast bioconjugation based on inverse-electron-demand dielsâalder reactivity. *J. Am. Chem. Soc.*, 130(41):13518–13519, October 2008. (Cited on page 49.)

[120] A. Borrmann, S. Milles, T. Plass, J. Dommerholt, J. M. M. Verkade, M. Wiessler, C. Schultz, J. C. M. van Hest, F. L. van Delft, and E. A. Lemke. Genetic encoding of a bicyclo[6.1.0]nonyne-charged amino acid enables fast cellular protein imaging by metal-free ligation. *ChemBioChem*, 13(14):2094–2099, September 2012. (Cited on pages 49, 50, 58, 62, 107 and 113.)

[121] T. Plass, S. Milles, C. Koehler, J. Szymanski, R. Mueller, M. Wiessler, C. Schultz, and E. A. Lemke. Amino acids for diels-alder reactions in living cells. *Angew. Chem. Int. Ed. (English)*, 51(17):4166–4170, April 2012. (Cited on pages 49 and 58.)

[122] R. Selvaraj and J. M. Fox. Trans-cyclooctene–a stable, voracious dienophile for bioorthogonal labeling. *Curr. Opin. Chem. Biol.*, 17(5):753–760, October 2013. (Cited on page 49.)

[123] M. R. Karver, R. Weissleder, and S. A. Hilderbrand. Synthesis and evaluation of a series of 1,2,4,5-tetrazines for bioorthogonal conjugation. *Bioconjugate Chem.*, 22(11):2263–2270, November 2011. (Cited on page 49.)

[124] F. Schoenebeck, D. H. Ess, G. O. Jones, and K. N. Houk. Reactivity and regioselectivity in 1,3-dipolar cycloadditions of azides to strained alkynes and alkenes: A computational study. *J. Am. Chem. Soc.*, 131(23):8121–8133, June 2009. (Cited on page 49.)

[125] Y. Liang, J. L. Mackey, S. A. Lopez, F. Liu, and K. N. Houk. Control and design of mutual orthogonality in bioorthogonal cycloadditions. *J. Am. Chem. Soc.*, 134(43):17904–17907, October 2012. (Cited on pages 49, 53 and 56.)

[126] W. Chen, D. Wang, C. Dai, D. Hamelberg, and B. Wang. Clicking 1,2,4,5-tetrazine and cyclooctynes with tunable reaction rates. *Chem. Comm.*, 48 (12):1736–1738, 2012. (Cited on pages 49 and 55.)

[127] F. Liu, R. S. Paton, S. Kim, Y. Liang, and K. N. Houk. Diels-alder reactivities of strained and unstrained cycloalkenes with normal and inverse-electron-demand dienes: Activation barriers and distortion/interaction analysis. *J. Am. Chem. Soc.*, 135(41):15642–15649, October 2013. (Cited on page 49.)

[128] J. Dommerholt, O. van Rooijen, A. Borrmann, C. F. Guerra, M. F. Bickelhaupt, and F. L. van Delft. Highly accelerated inverse electron-demandcycloaddition of electron-deficient azides withaliphatic cyclooctynes. *Nat. Commun.*, 5:1–7, October 2014. (Cited on page 49.)

[129] D. H. Ess and K. N. Houk. Distortion/interaction energy control of 1,3-dipolar cycloaddition reactivity. *J. Am. Chem. Soc.*, 129(35):10646–10647, September 2007. (Cited on page 49.)

[130] A. Talbot, D. Devarajan, S. J. Gustafson, Israel Fernandez, F. M. Bickelhaupt, and D. H. Ess. Activation-strain analysis reveals unexpected origin of fast reactivity in heteroaromatic azadiene inverse-electron-demand diels-alder cycloadditions. *J. Org. Chem.*, 80(1):548–558, January 2015. (Cited on page 49.)

[131] J. Li, S. Jia, and P. R. Chen. Diels-alder reaction–triggered bioorthogonal protein decaging in living cells. *Nat. Chem. Biol.*, pages 1–5, November 2014. (Cited on page 51.)

[132] J.-E. Hoffmann, T. Plass, I. Nikic, I. Valle Aramburu, C. Koehler, H. Gillandt, E. A. Lemke, and C. Schultz. Highly stabile trans-cyclooctene amino acids for reactions in living cells. *Chem. Eur. J.*, 2015. submitted. (Cited on pages 51, 56, 57, 60, 62, 100 and 113.)

[133] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9): 1157–1174, April 2004. (Cited on pages 51 and 53.)

[134] Y. Zhao and D. G. Truhlar. The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four m06-class functionals and 12 other functionals. *Theor. Chem. Acc.*, 120(1-3):215–241, 2008. (Cited on pages 51 and 53.)

[135] Y. Lan, L. Zou, Y. Cao, and K. N. Houk. Computational methods to calculate accurate activation and reaction energies of 1,3-dipolar cycloadditions of 24 1,3-dipoles. *J. Phys. Chem. A*, 115(47):13906–13920, December 2011. (Cited on pages 51 and 53.)

[136] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, and K. A. Lyssenko. Density functional theory is straying from the path toward the exact functional. *Science*, 355(6320):49–52, 2017. ISSN 0036-8075. doi: 10.1126/science.aah5975. URL http://science.sciencemag.org/content/355/6320/49. (Cited on page 51.)

[137] E. G. Hohenstein, S. T. Chill, and D. C. Sherrill. Assessment of the performance of the m05-2x and m06-2x exchange-correlation functionals for noncovalent interactions in biomolecules. *J. Chem. Theory Comput.*, 4(12):1996–2000, December 2008. (Cited on page 53.)

[138] L. Goerigk and S. Grimme. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.*, 13(14):6670, 2011. (Cited on page 53.)

[139] Y. Zhao and D. G. Truhlar. Exploring the limit of accuracy of the global hybrid meta density functional for main-group thermochemistry, kinetics, and noncovalent interactions. *J. Chem. Theory Comput.*, 4(11):1849–1868, November 2008. (Cited on page 53.)

[140] M. Walker, A. J. A. Harvey, A. Sen, and C. E. H. Dessent. Performance of m06, m06-2x, and m06-hf density functionals for conformationally flexible anionic clusters: M06 functionals perform better than b3lyp for a model system with dispersion and ionic hydrogen-bonding interactions. *J. Phys. Chem. A*, 117 (47):12590–12600, 2013. (Cited on page 53.)

[141] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople. Self-consistent molecular orbital methods. xx. a basis set for correlated wave functions. *J. Chem. Phys.*, 72(1):650, 1980. (Cited on page 53.)

[142] M. J. *et al.* Frisch. Gaussian 09 Revision D.01, 2009. Gaussian Inc. Wallingford CT 2009. (Cited on page 53.)

[143] J. Wang, W. Wang, P. A. Kollman, and D. A. Case. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.*, 25(2):247–260, October 2006. (Cited on page 53.)

[144] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.*, 4(1):17, 2012. (Cited on page 53.)

[145] N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler, and L. Kronik. Dispersion interactions with density-functional theory: Benchmarking semiempirical and interatomic pairwise corrected density functionals. *J. Chem. Theory Comput.*, 7(12):3944–3951, December 2011. (Cited on page 53.)

[146] R. Sedlak, T. Janowski, M. Pitonak, J. Rezac, P. Pulay, and P. Hobza. Accuracy of quantum chemical methods for large noncovalent complexes. *J. Chem. Theory Comput.*, 9(8):3364–3374, August 2013. (Cited on page 53.)

[147] N. M. O'boyle, A. L. Tenderholt, and K. M. Langner. cclib: A library for package-independent computational chemistry algorithms. *J. Comput. Chem.*, 29(5):839–845, 2008. (Cited on page 53.)

[148] J. Dommerholt, S. Schmidt, R. Temming, L. J. A. Hendriks, F. P. J. T. Rutjes, J. C. M. van Hest, D. J. Lefeber, P. Friedl, and F. L. van Delft. Readily accessible bicyclononynes for bioorthogonal labeling and three-dimensional imaging of living cells. *Angew. Chem. Int. Ed. (English)*, 49(49):9422–9425, September 2010. (Cited on page 58.)

[149] J. I. García, J. A. Mayoral, and L. Salvatella. The source of the endo rule in the diels- alder reaction: Are secondary orbital interactions really necessary? *Eur. J. Org. Chem.*, 2005(1):85–90, 2005. (Cited on page 58.)

[150] D. T. Grubb and L. W. Jelinski. Fiber morphology of spider silk: the effects of tensile deformation. *Macromolecules*, 30(10):2860–2867, 1997. (Cited on pages 79, 80, 81, 82, 83 and 89.)

[151] D. T. Grubb and G. Ji. Molecular chain orientation in supercontracted and re-extended spider silk. *Int. J. Biol. Macromol.*, 24(2-3):203–210, 1999. (Cited on page 79.)

[152] M. J. Hill and A. Keller. Direct evidence for distinctive, stress-induced nucleus crystals in the crystallization of oriented polymer melts. *J. Macromol. Sci., Part B: Phys.*, 3(1):153–169, August 2006. (Cited on page 79.)

[153] C. Riekel, C. Bränden, C. Craig, C. Ferrero, F. Heidelbach, and M. Müller. Aspects of x-ray diffraction on single spider fibers. *Int. J. Biol. Macromol.*, 24(2):179–186, 1999. (Cited on pages 79 and 83.)

[154] A. Glišović, T. Vehoff, R. J. Davies, and T. Salditt. Strain dependent structural changes of spider dragline silk. *Macromolecules*, 41(2):390–398, January 2008. (Cited on pages 79 and 80.)

[155] F. A. McClintock. A criterion for ductile fracture by the growth of holes. *J. Appl. Mech.*, 35(2):363–371, June 1968. (Cited on page 79.)

[156] P. F. Thomason. A theory for ductile fracture by internal necking of cavities. *J. Inst. Metals*, 96(12):360–365, 1968. (Cited on page 79.)

[157] J. R. Rice and D. M. Tracey. On the ductile enlargement of voids in triaxial stress fields. *J. Mech. Phys. Solids*, 17(3):201–217, 1969. (Cited on page 79.)

[158] H. M. Simpson, A. Sosin, G. R. Edwards, and S. L. Seiffert. New mechanism for internal friction. *Phys. Rev. Lett.*, 26(15):897–900, April 1971. (Cited on page 79.)

[159] J. R. Townsend, M. Schildcrout, and C. Reft. Mechanical studies of irradiation-induced defects in cu and w. *Phys. Rev. B*, 14(2):500–516, July 1976. (Cited on page 79.)

[160] A. S. Nowick, B. S. Berry, and J. L. Katz. Anelastic relaxation in crystalline solids. *J. Appl. Mech.*, 42(3):750–751, 1975. (Cited on page 79.)

[161] F. Zemlin, R. Schuster, E. Beckmann, J. L. Carrascosa, J. M. Valpuesta, and G. Ertl. Stress-induced recrystallization of a protein crystal by electron irradiation. *Nature*, 399(6731):51–54, May 1999. (Cited on pages 80 and 98.)

[162] M. Xu and R. V. Lewis. Structure of a protein superfiber: spider dragline silk. *Proc. Natl. Acad. Sci. U.S.A.*, 87(18):7120–7124, September 1990. (Cited on pages 80 and 89.)

[163] 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(23):3295–3303, 1999. (Cited on pages 80, 85, 89, 94 and 113.)

[164] J. Gatesy, C. Hayashi, D. Motriuk, J. Woods, and R. Lewis. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science*, 291 (5513):2603–2605, 2001. (Cited on pages 80 and 89.)

[165] A. D. Parkhe, S. K. Seeley, K. Gardner, L. Thompson, and R. V. Lewis. Structural studies of spider silk proteins in the fiber. *J. Mol. Recogn.*, 10(1): 1–6, January 1997. (Cited on page 80.)

[166] D. V. Mahoney, D. L. Vezie, R. K. Eby, W. W. Adams, and D. Kaplan. Aspects of the morphology of dragline silk of nephila clavipes. *ACS Symp. Ser.*, 544: 196–210, 1994. (Cited on pages 80 and 91.)

[167] Z. Yang, D. T. Grubb, and L. W. Jelinski. Small-angle x-ray scattering of spider dragline silk. *Macromolecules*, 30(26):8254–8261, December 1997. (Cited on pages 80, 83 and 91.)

[168] C. Riekel and F. Vollrath. Spider silk fibre extrusion: combined wide-and small-angle x-ray microdiffraction experiments. *Int. J. Biol. Macromol.*, 29 (3):203–210, 2001. (Cited on pages 80 and 91.)

[169] R. Valluzzi and H.-J. Jin. X-ray evidence for a super-secondary structure in silk fibers. *Biomacromolecules*, 5(3):696–703, May 2004. (Cited on pages 80 and 91.)

[170] D. Sapede, T. Seydel, V. T. Forsyth, M. M. Koza, R. Schweins, F. Vollrath, and C. Riekel. Nanofibrillar structure and molecular mobility in spider dragline silk. *Macromolecules*, 38(20):8447–8453, October 2005. (Cited on pages 80, 81 and 91.)

[171] N. Du, X. Yang Liu, J. Narayanan, L. Li, M. L. M. Lim, and D. Li. Design of superior spider silk: From nanostructure to mechanical properties. *Biophys. J.*, 91(12):4528–4535, December 2006. (Cited on pages 80 and 91.)

[172] M. Cetinkaya, S. Xiao, and F. Gräter. Effects of crystalline subunit size on silk fiber mechanics. *Soft Matter*, 7(18):8142–8148, 2011. (Cited on pages 81 and 83.)

[173] M. Cetinkaya, S. Xiao, B. Markert, W. Stacklies, and F. Gräter. Silk fiber mechanics from multiscale force distribution analysis. *Biophys. J.*, 100(5): 1298–1305, March 2011. (Cited on pages 81, 82, 83 and 89.)

[174] C. Fu, Z. Shao, and F. Vollrath. Animal silks: their structures, properties and artificial production. *Chem. Commun.*, (43):6515–6529, 2009. doi: 10.1039/B911049F. URL `http://dx.doi.org/10.1039/B911049F`. (Cited on page 81.)

[175] R. Madurga, T. A. Blackledge, B. Perea, G. R. Plaza, C. Riekel, M. Burghammer, M. Elices, G. Guinea, and J. Pérez-Rigueiro. Persistence and variation in microstructural design during the evolution of spider silk. *Sci. Rep.*, 5:14820, 2015. (Cited on page 81.)

[176] G. R. Plaza, J. Pérez-Rigueiro, C. Riekel, B. G. Perea, F. Agulló-Rueda, M. Burghammer, G. V. Guinea, and M. Elices. Relationship between microstructure and mechanical properties in spider silk fibers: identification of two regimes in the microstructural changes. *Soft Matter*, 8(22):6015–12, 2012. (Cited on pages 81 and 89.)

[177] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez. Packmol: a package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.*, 30(13):2157–2164, 2009. (Cited on page 82.)

[178] S. E. Benzley, E. Perry, K. Merkley, B. Clark, and G. Sjaardama. A comparison of all hexagonal and all tetrahedral finite element meshes for elastic and elasto-plastic analysis. In *Proc., 4th Int. Meshing Roundtable*, volume 17, pages 179–191. Sandia National Laboratories Albuquerque, NM, 1995. (Cited on page 82.)

[179] S. Panagiotis. Development of material structure design method combining evolutionary algorithms and finite elements. application for optimisation of bio-inspired nano-composite fibres. Master's thesis, National Technical University of Athens, Greece, 2014. (Cited on page 82.)

[180] S. P. Patil, F. Gräter, and B. Markert. Refining a bottom-up computational approach for spider silk fibre mechanics. *Proc. 3rd GAMM Seminar Contin. Biomech.*, II–21:75–87, November 2012. (Cited on pages 83 and 89.)

[181] A. H. Simmons, C. A. Michal, and L. W. Jelinski. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science*, 271(5245):84, 1996. (Cited on page 83.)

[182] I. Krasnov, T. Seydel, M. Greving, I.and Blankenburg, F. Vollrath, and M. Müller. Strain-dependent fractional molecular diffusion in humid spider silk fibres. *J. R. Soc. Interface*, 13(122):20160506, 2016. (Cited on pages 83, 88, 91, 92 and 98.)

[183] R. V. Lewis. Spider silk: the unraveling of a mystery. *Acc. Chem. Res.*, 25 (9):392–398, 1992. (Cited on page 85.)

[184] D. H. Hijirida, K. G. Do, C. Michal, S. Wong, D. Zax, and L. W. Jelinski. 13c nmr of nephila clavipes major ampullate silk gland. *Biophys. J.*, 71(6):3442, 1996. (Cited on page 85.)

[185] C. Dicko, D. Knight, J. M. Kenney, and F. Vollrath. Secondary structures and conformational changes in flagelliform, cylindrical, major, and minor ampullate silk proteins. temperature and concentration effects. *Biomacromolecules*, 5(6):2105–2115, 2004. (Cited on page 85.)

[186] T. Lefevre, J. Leclerc, J.-F. Rioux-Dubé, T. Buffeteau, M.-C. Paquin, M.-E. Rousseau, I. Cloutier, M. Auger, S. M. Gagné, S. Boudreault, et al. Conformation of spider silk proteins in situ in the intact major ampullate gland and in solution. *Biomacromolecules*, 8(8):2342–2344, 2007. (Cited on page 85.)

[187] B. An, J. E. Jenkins, S. Sampath, G. P. Holland, M. Hinman, J. L. Yarger, and R. Lewis. Reproducing natural spider silksâ copolymer behavior in synthetic silk mimics. *Biomacromolecules*, 13(12):3938–3948, 2012. (Cited on page 85.)

[188] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A fresh approach to numerical computing. *CoRR*, abs/1411.1607, November 2014. URL `http://arxiv.org/abs/1411.1607`. (Cited on page 85.)

[189] S. Carmichael and C. Viney. Molecular order in spider major ampullate silk (dragline): Effects of spinning rate and post-spin drawing. *J. Appl. Polym. Sci.*, 72(7):895–903, 1999. (Cited on page 85.)

[190] D. R. Drodge, B. Mortimer, C. Holland, and C. R. Siviour. Ballistic impact to access the high-rate behaviour of individual silk fibres. *J. Mech. Phys. Solids*, 60(10):1710–1721, 2012. (Cited on page 85.)

[191] J. Guan, D. Porter, and F. Vollrath. Silks cope with stress by tuning their mechanical properties under load. *Polymer*, 53(13):2717–2726, 2012. (Cited on page 85.)

[192] S. Ulrich, A. Glišović, T. Salditt, and A. Zippelius. Diffraction from the β-sheet crystallites in spider silk. *Eur. Phys. J. E Soft Matter*, 27(3):229–242, 2008. (Cited on page 86.)

[193] I. Krasnov, T. Seydel, I. Greving, M. Blankenburg, F. Vollrath, and M. Müller. Strain-dependent fractional molecular diffusion in humid spider silk fibres. *J. R. Soc. Interface*, 13(122):20160506, 2016. (Cited on page 87.)

[194] R. W. Work and P. D. Emerson. An apparatus and technique for the forcible silking of spiders. *J. Arachnol.*, 10(1):1–10, Spring 1982. (Cited on page 87.)

[195] T. Seydel, W. Knoll, I. Greving, C. Dicko, M. M. Koza, I. Krasnov, and M. Müller. Increased molecular mobility in humid silk fibers under tensile stress. *Phy. Rev. E Stat. Nonlin. Soft Matter Phys.*, 83(1 Pt 2):016104, January 2011. (Cited on pages 88 and 91.)

[196] C. Dewhurst. Grasp, 2015. ILL, www.ill.eu/instruments-support/instruments-groups/groups/lss/grasp. (Cited on page 88.)

[197] T. Seydel, K. Kölln, I. Krasnov, I. Diddens, N. Hauptmann, G. Helms, M. Ogurreck, S.-G. Kang, M. M. Koza, and M. Müller. Silkworm silk under tensile strain investigated by synchrotron x-ray diffraction and neutron spectroscopy. *Macromolecules*, 40(4):1035–1042, 2007. (Cited on page 88.)

[198] B. M. Ginzburg and S. Tuichiev. Microdeformational behavior of oriented semicrystalline polymers. *J. Macromol. Sci., Part B: Phys.*, 31(3):291–317, 1992. doi: 10.1080/00222349208215518. (Cited on page 92.)

[199] M. Denny. The physical properties of spiderâs silk and their role in the design of orb-webs. *J. exp. Biol*, 65(2):483–506, 1976. (Cited on pages 94 and 113.)

[200] T. Köhler and F. Vollrath. Thread biomechanics in the two orb-weaving spiders araneus diadematus (araneae, araneidae) and uloborus walckenaerius (araneae, uloboridae). *J. Exp. Zool. A. Ecol. Genet. Physiol.*, 271(1):1–17, 1995. (Cited on pages 94 and 113.)

[201] T. Vehoff, A. Glišović, H. Schollmeyer, A. Zippelius, and T. Salditt. Mechanical properties of spider dragline silk: humidity, hysteresis, and relaxation. *Biophys. J.*, 93(12):4425–4432, 2007. (Cited on pages 94 and 113.)

[202] T. H. Ware, J. S. Biggins, A. F. Shick, M. Warner, and T. J. White. Localized soft elasticity in liquid crystal elastomers. *Nat. Commun.*, 7, 2016. (Cited on page 98.)

[203] T. Stauch and A. Dreuw. Quantum chemical strain analysis for mechanochemical processes. *Acc. Chem. Res.*, 2017. (Cited on page 101.)

[204] W. Li, S. A. Edwards, L. Lu, T. Kubar, S. P. Patil, H. Grubmüller, G. Groenhof, and F. Gräter. Force distribution analysis of mechanochemically reactive dimethylcyclobutene. *ChemPhysChem*, 14(12):2687–2697, 2013. (Cited on page 101.)