

# **Molecular Dynamics simulations of enforced functional protein motions**

Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultäten

der Georg-August-Universität zu Göttingen

vorgelegt von

**Frauke Gräter, geb. Meyer**

aus Holzminden

Göttingen, 2005

---

D7  
Referent: Prof. Dr. Jürgen Troe  
Korreferent: PD Dr. Helmut Grubmüller  
Tag der mündlichen Prüfung:

# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>                                       | <b>7</b>  |
| <b>2</b> | <b>Theory and Methods</b>                                 | <b>17</b> |
| 2.1      | Molecular Dynamics Simulations . . . . .                  | 17        |
| 2.2      | Monte Carlo Simulations . . . . .                         | 24        |
| 2.3      | Replica-Exchange Molecular Dynamics Simulations . . . . . | 25        |
| 2.4      | Simulating force-induced transitions . . . . .            | 26        |
| 2.4.1    | Force Probe Molecular Dynamics Simulations . . . . .      | 26        |
| 2.4.2    | Theory of force probe experiments . . . . .               | 28        |
| 2.4.3    | Essential Dynamics Simulations . . . . .                  | 30        |
| 2.5      | Free energy calculations . . . . .                        | 31        |
| 2.5.1    | Free energies from non-equilibrium trajectories . . . . . | 32        |
| 2.5.2    | Free energies from equilibrium trajectories . . . . .     | 34        |
| 2.5.3    | Free energy perturbation . . . . .                        | 35        |
| <b>3</b> | <b>Enforced Titin Kinase Activation</b>                   | <b>37</b> |
| 3.1      | Titin and titin kinase . . . . .                          | 37        |
| 3.2      | Simulation details . . . . .                              | 42        |
| 3.2.1    | MD and Force Probe MD simulations . . . . .               | 42        |
| 3.2.2    | Force Probe Monte Carlo Simulations . . . . .             | 44        |
| 3.3      | Results . . . . .   | 45        |
| 3.3.1    | Equilibration and free dynamics . . . . .                 | 45        |

## CONTENTS

---

|          |  |           |
|----------|--|-----------|
| 3.3.2    | Unfolding mechanism . . . . .                                  | 46        |
| 3.3.3    | Open-closure motion . . . . .                                  | 49        |
| 3.3.4    | $\beta$ -sheet ruptures . . . . .                              | 51        |
| 3.3.5    | Release of auto-inhibition . . . . .                           | 57        |
| 3.3.6    | Comparison to twitchin kinase unfolding . . . . .              | 59        |
| 3.4      | Discussion and Conclusion . . . . .                            | 61        |
| <b>4</b> | <b>Titin kinase mutants</b>                                    | <b>67</b> |
| 4.1      | Why mutants? . . . . .   | 67        |
| 4.2      | Choice of mutants and simulation details . . . . .             | 68        |
| 4.3      | Results . . . . .  | 69        |
| 4.3.1    | $\alpha$ R2-mutants . . . . .                                  | 69        |
| 4.3.2    | $\beta$ R1-mutant . . . . .                                    | 72        |
| 4.4      | Discussion and Conclusion . . . . .                            | 76        |
| <b>5</b> | <b>Bombykol release from BmPBP</b>                             | <b>79</b> |
| 5.1      | Olfaction and the bombykol-BmPBP complex . . . . .             | 79        |
| 5.2      | Simulation details . . . . .                                   | 83        |
| 5.2.1    | MD and replica-exchange MD simulations . . . . .               | 83        |
| 5.2.2    | Essential dynamics and force probe MD simulations . . . . .    | 85        |
| 5.2.3    | Free energy calculations . . . . .                             | 86        |
| 5.3      | Results and Discussion . . . . .                               | 87        |
| 5.3.1    | Equilibrium dynamics of the bound complex . . . . .            | 87        |
| 5.3.2    | Ligand release pathways from essential dynamics . . . . .      | 92        |
| 5.3.3    | Ligand release pathways from force probe simulations . . . . . | 95        |
| 5.3.4    | Bombykal versus bombykol binding . . . . .                     | 102       |
| 5.4      | Conclusions and Outlook . . . . .                              | 105       |



---

|          |   |            |
|----------|---|------------|
| <b>6</b> | <b>Photoswitch mechanism</b>                    | <b>109</b> |
| 6.1      | The fluorescent protein asFP595 . . . . .       | 109        |
| 6.2      | Simulation details . . . . .                    | 112        |
| 6.2.1    | Chromophore force field . . . . .               | 112        |
| 6.2.2    | MD simulations . . . . .                        | 113        |
| 6.3      | Results . . . . .                               | 116        |
| 6.3.1    | Enforced chromophore isomerization . . . . .    | 116        |
| 6.3.2    | Spontaneous chromophore isomerization . . . . . | 117        |
| 6.4      | Discussion and Conclusion . . . . .             | 118        |
| <b>7</b> | <b>Summary and Conclusion</b>                   | <b>121</b> |
|          | <b>Acknowledgment</b>                           | <b>127</b> |
|          | <b>Appendix A</b>                               | <b>129</b> |

## CONTENTS

---

# 1

## Introduction

---

Proteins are the working horses of life, enabling living organisms to move, metabolize and sense, in short, to function. To fulfill such diverse tasks, proteins take up a specific fold, *i. e.*, a region in a complex high-dimensional free energy landscape, which they dynamically explore [1].

The process on which the function of a protein is based, such as a large-scale domain motion, an enzymatic reaction, association, or dissociation, is a particular move on this free energy landscape between different protein states. The energy barriers, valleys and plateaus along the pathway govern the protein dynamics, resulting in the molecular mechanism and kinetics as function demands. This coupling between dynamics and function is an established theoretical concept [2]. Its quantitative understanding, however, is far from complete, and requires the comprehensive characterization of the underlying molecular processes and energetics.

Static three-dimensional structures of folded proteins are experimentally readily accessible in atomic detail, mainly by means of X-ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy. Monitoring proteins as their structure evolves in time, the basis of protein function, became possible only recently, *via* advanced experimental techniques, among others NMR relaxation measurements [3–5], and fluorescence spectroscopy [6, 7]. Single-molecule force probe experiments, such as atomic-force microscopy (AFM), give indirect insight into protein motion by reporting the mechanical force required to induce protein motions [8–10].

Experimental techniques, however, to date are not able to directly provide an atomistic picture of protein dynamics. Here, simulations come into play

to complement experiments with a microscopic interpretation. Molecular dynamics (MD) simulations are widely used to monitor protein motion on femtosecond to microsecond time scales, and are described in Section 2.1. To investigate specific protein transitions, including those that occur on even longer time scales, a number of MD simulation techniques has been developed that force the system to proceed along a defined pathway within a reasonable simulation time. The method of choice depends on the questions to be answered regarding the functional motion under investigation. By means of a set of advanced MD simulation methods, this thesis addresses various types of transitions: first, force-induced protein unfolding, second, protein-ligand unbinding, both ubiquitous processes investigated for particular systems, a muscle protein and an odorant protein, and third, an isomerization reaction in a chromoprotein. The next two paragraphs shortly outline AFM experiments and the MD based methods applied for the investigation of these enforced processes.

### **Enforcing protein motions experimentally**

The AFM, originally developed as an imaging technique, was more recently also used to manipulate single protein molecules and to measure the involved mechanical force [10–12]. AFM experiments have been performed primarily to study the properties of proteins with mechanical functions, namely filaments and molecular motors. Above all, since force enables to explore the energy landscape of functional motions in proteins in general, the AFM also proved useful to gain insight into unfolding and folding pathways, ligand recognition or the molecular basis of elasticity. Representative examples of AFM single-molecule experiments are the unfolding of the protein filament titin [11, 13–16] and the protein ubiquitin [17–19], the stretching of DNA and sugar [20–23], and the dissociation of adhesion complexes [24–26].

In a typical AFM single molecule experiment, as shown in Fig. 1.1, the molecule under investigation is attached *via* polymeric linkers on one side to a surface and on the other to the cantilever, a flexible arm holding a sharp tip. The surface is moved away from the tip with constant velocity, thus inducing the unfolding or unbinding of the substrate clamped in between. The force as a function of the cantilever position is obtained from the bending of the cantilever. The resulting maximum and shape of the force profile allows to draw conclusions on the mechanical stability of the system, the unfolding or unbinding pathway and, eventually, on involved intermediates. As single molecule experiments, AFM experiments yield forces of individual rupture



events, and after a series of runs statistical information such as distributions, rather than merely mean values as obtained from bulk experiments. Comparison of individual measurements to molecular simulations thus provides insightful information on a single molecule level, which instead is lost when exclusively considering ensemble averages.

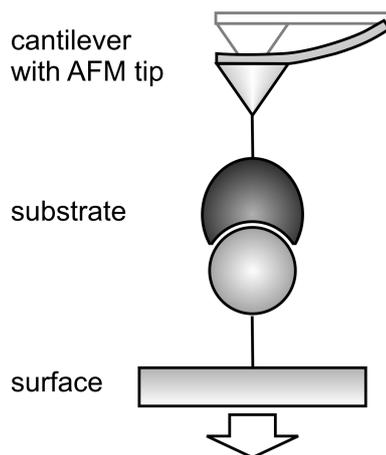


Figure 1.1: Schematic drawing of a typical AFM setup, here considering the example of the dissociation of a protein-ligand complex. The surface is moved downwards with constant velocity, exerting a pulling force onto the attached complex. The bending angle of the cantilever reports on this force as a function of surface position.

A comprehensive picture of the enforced protein motions is optimally obtained by the combination of different experimental approaches. Thermodynamic and kinetic quantities, such as free energies and activation free energies of folding and binding, are a valuable source of data molecular simulations can in principle be compared to, providing the simulations allow the estimation of these equilibrium properties. Finally, site-directed mutagenesis of individual residues is an established experimental tool to seek out the residues important for domain motions or ligand binding of proteins.

Titin, *e. g.*, was intensively studied by single-molecule force experiments like AFM and others, which were complemented by results from mutants and thermodynamic and kinetic data (see Chapter 3 for details) [27–29]. Again, a microscopic interpretation of the macroscopic observables is left to complementary simulations.

## Enforcing protein motions by simulations

Available simulation methods to enforce the functional motion of a protein mainly differ in the definition of the transition coordinate and the way the system is forced to proceed along this coordinate. The coordinate can simply be a vector in Cartesian space, a distance or angle, a collective coordinate of motion, etc. Provided the initial and final states of the motion are at hand, some approaches make use of both, by defining a transition coordinate that connects the two [30–32]. To enforce the transition, one or more atoms are subjected to a constraining force or a harmonic potential, as in Force Probe Molecular Dynamics (FPMD) simulations [33], Umbrella sampling [32], or Essential Dynamics (ED) [34, 35]. Other methods adaptively modify the potential energy landscape of the system by increasing the energy of the current local minimum to enable the system to escape from it [36, 37]. In the Free Energy Perturbation (FEP) method, the force field of the system itself is perturbed, *i. e.*, interpolated between two structures or topologies [38, 39]. Principles and applicability of the approaches employed in this thesis, FPMD, ED and FEP, are shortly sketched in the following; methodological aspects are explained in detail in Chapter 2.

By closely mimicking single molecule AFM experiments, FPMD simulations are a valuable tool to simulate enforced motions in proteins. For force-induced protein unfolding [40–42], the protein termini are connected to 'virtual' springs. Moving the springs away from each other results in a pulling force that induces unfolding. Similarly, the rupture of a protein-ligand complex can be simulated by subjecting the ligand to a pulling force that enforces the dissociation of the ligand from the protein [43–46]. FPMD simulations thus directly model the setup of AFM experiments, with the springs modeling the flexible cantilever. The mechanism of the enforced process can be related to the involved forces and energetics to identify the mechanically stable units of the system. Additionally, structural factors determining the mechanical stability and, in case of protein-ligand complexes, the binding specificity can be deduced. Conclusions can be further refined by comparison to point mutants [13, 16, 44], and finally validated (or falsified) by comparison to single-molecule force probe experiments.

In ED simulations, the system is enforced to undergo functional motions by using the knowledge of the protein or protein-ligand complex dynamics at equilibrium. The system is driven along one or more major modes of motion observed in free dynamics. Again, the unfolding or unbinding mechanism, involved forces, and energetics can be obtained. ED sampling has as yet



proved useful for an extended sampling of protein conformational space [35, 47, 48], but in this thesis was applied to ligand dissociation from a protein.

FEP simulations are the third technique employed here to enforce functional motions, also in the context of protein-ligand binding. They allow to calculate free energy changes of transformations between chemically different states, and were originally used to calculate hydration and relative binding free energies [49, 50]. Here, FEP was not only applied in this context, but also to assess the mechanism of the transition between conformationally different states [51]. This requires a unique definition of the initial and final states and the transition coordinate.

**Aim of this thesis**

This thesis aims at contributing to the understanding of protein function by elucidating the molecular mechanism of three types of transitions in different proteins, force-induced unfolding of the muscle protein titin kinase (Chapters 3 and 4), ligand release from a pheromone-binding protein (PBP, Chapter 5), and photo-induced isomerization of a chromoprotein (Chapter 6), by means of MD simulations. Each of these studies addresses particular biological questions concerning the physiological role of the protein under investigation. The biological background and questions addressed are further detailed in the introductory sections of the corresponding chapters.

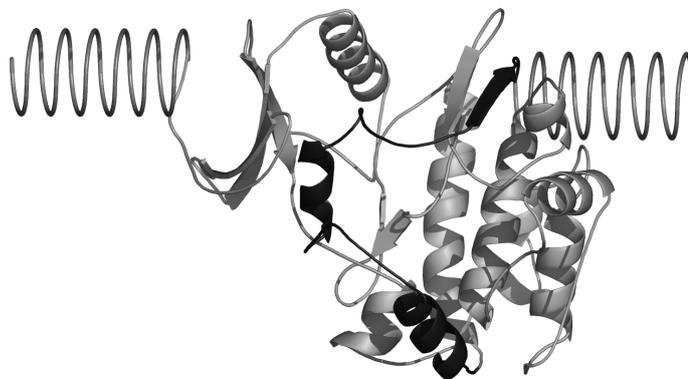


Figure 1.2: The muscle protein titin kinase. The enzyme is auto-inhibited by blocking its active site with a C-terminal auto-regulatory tail (dark). FPMD simulations of the wild-type and mutants were employed to characterize the relieve of auto-inhibition.

Folding and unfolding are the most fundamental functional motions of proteins, those that bestow and seize protein function. Titin, a filament of the muscle cell, partially unfolds upon muscle cell stretching and refolds upon relaxation, thereby acting as a molecular spring under mechanical load. A kinase domain, titin's only catalytic domain, is located near the C-terminus of the filament, and thus is in a prominent position to sense mechanical force in the muscle cell.

Further insight into the putative role of titin kinase as a force sensor was gained from its crystal structure, which is shown in Fig. 1.2 [52]. Activation of titin kinase requires substantial motions of the protein to relieve the active site blockage by its auto-regulatory tail. Is titin kinase activated when unfolding under mechanical force and thus senses mechanical load in the muscle? Our FPMD simulations aimed at characterizing the enforced



motion — the unfolding and eventually activation mechanism — of titin kinase, to challenge the hypothesis of titin kinase as a force sensor.

The molecular process of unfolding is governed by the underlying protein structural properties responsible for force resistance. To complement and validate the mechanism of wild type titin kinase unfolding obtained from FPMD simulations, we additionally selected point mutants on the basis of the previous FPMD simulations of the wild-type, and subjected them to FPMD simulations. Crucial residues of the structural motifs that were found to be mainly responsible for the stability of titin kinase were selected for mutation. Comparison to the wild-type simulations aimed at refining the picture of force-induced titin kinase activation.

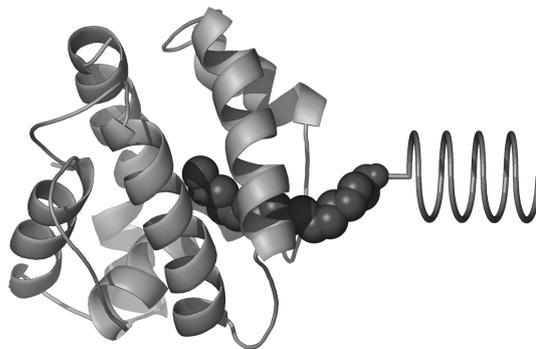


Figure 1.3: The olfactory BmPBP and the pheromone bombykol (spheres) during force-induced dissociation as observed during FPMD and ED. The simulations aim at identifying physiologically relevant dissociation pathways.

Apart from functional motions of a protein by itself, protein dynamics also govern the interaction with potential binding partners. Olfaction recently increasingly gained attention as one of the most elaborate examples for protein function involving ligand binding. A ligand involved in olfaction is the pheromone bombykol, a hormone for the sexual communication of the silk moth *Bombyx mori*. After absorption, the hydrophobic bombykol is encapsulated by the pheromone-binding protein (BmPBP) and carried to the pheromone receptor to elicit neuronal response. The remarkable mode of function of BmPBP, both as a carrier and a selective filter, remains elusive. Olfaction requires BmPBP to uptake and release bombykol notably fast in conjunction with a fair selectivity. How is this achieved by BmPBP, which does not exhibit any obvious ligand entry/exit route out of the central cavity? Here, we aimed at characterizing the binding properties of BmPBP by identifying ligand release pathways on the basis of forces and energetics, by

means of FPMD and ED simulations. We also investigate the degree and determining factors of ligand selectivity of BmPBP using conventional MD and FEP simulations.



Figure 1.4: The fluorescent protein asFP595 of the sea anemone *Anemonia sulcata*. The photo-switching mechanism was elucidated by inducing the underlying isomerization reaction of the chromophore (spheres) in FPMD simulations, and by additional simulations of the spontaneous isomerization.

A confined, yet fateful transformation, as is characteristic for light-induced processes, occurs in the protein asFP595, a fluorescent protein of the GFP family. This protein undergoes a reversible isomerization of its chromophore from the trans to the cis conformation upon light absorption, and thereby becomes fluorescent. What is the exact molecular mechanism of this reversible switching motion, and what is the role of the protein matrix? We aimed here at exploring the different conceivable pathways of the isomerization by FPMD and free MD simulations to estimate the discriminative impact of the protein environment on them.

The studies described above were carried out in close collaboration with experimental groups headed by M. Gautel, King's College London, K. E. Kaissling, MPI for Ornithology, Seewiesen, M. Wahl, and S. Hell at the MPI for Biophysical Chemistry, Göttingen. Simulations were compared to results from single-molecule experiments, spectroscopic, thermodynamic or kinetic measurements, and mutant experiments, obtained at the laboratories of the collaborators and others.

The systems under investigation are not only of biological interest, but rep-



resent test cases for the applied methods, such as FPMD, ED, and FEP simulations. Comparing the simulation results obtained with the different techniques with each other, with other test cases, and, most importantly, with experiments, contributes to the validation and improvement of the approaches and spot their strength and potential pitfalls.

These investigations finally also aimed at adding to the picture of protein dynamics in general. What brings about the mechanical stability of proteins? What determines molecular recognition? Or what is the role of a protein environment in a specific reaction? By looking at particular proteins, this thesis shall give further insight into some of the more basic principles of functional protein motions.

1. INTRODUCTION

---

## 2

# Theory and Methods

---

This Chapter outlines the general theoretical framework of this thesis and the basic methods applied. Details on the particular methods and the simulation setup used for the different biological systems are given in the respective Chapters 3 to 6.

## 2.1 Molecular Dynamics Simulations

The general concern of this thesis is to describe the molecular dynamics of proteins, *i. e.*, the structure and properties of proteins at atomic detail as they evolve in time. Proteins typically consist of several tens to several hundred amino acid residues, amounting to thousands of atoms, and feature dynamics from bond vibrations at the femtosecond timescale to large-scale conformational changes occurring within microseconds or even more slowly. The only established simulation method that can cope with the high complexity of protein structures and yet is able to assess relevant time scales is Molecular Dynamics (MD) simulations. This section shortly outlines the principles and approximations on which MD simulations rest. It specifically highlights those aspects of MD simulations which were of particular concern for this thesis without the attempt to be comprehensive. Special MD-based simulation techniques applied, FPMD, ED, and FEP, are described in Sections 2.4 and 2.5. For a comprehensive description of MD, we refer to recent reviews [53] and relevant text books [54, 55].

### Approximation I: Born-Oppenheimer

The dynamics of any system is described by the time-dependent Schrödinger equation,

$$H\psi = i\hbar\frac{\delta\psi}{\delta t}, \quad (2.1)$$

where  $H$  denotes the Hamiltonian, the sum of potential and kinetic energy,  $\psi$  the wave function, and  $\hbar = h/2\pi$  with  $h$  Planck's constant. The wave function comprises all particles of the system, *i. e.*, it is a function of the coordinates and momenta of both, nuclei and electrons. Due to the much lower mass and consequently much higher velocity of electrons compared to nuclei, electrons can often be assumed to instantaneously follow the nuclear motion. Therefore, in the Born-Oppenheimer approximation [56], the total wave function  $\psi_{\text{tot}}$  is separated into the nuclear  $\psi_{\text{n}}$  and electronic wave function  $\psi_{\text{e}}$ ,

$$\psi_{\text{tot}}(R, r) = \psi_{\text{n}}(R)\psi_{\text{e}}(R; r), \quad (2.2)$$

where  $R$  and  $r$  are the coordinates and momenta of the nuclei and electrons, respectively. Thus, the electronic wave function  $\psi_{\text{e}}$  only parametrically depends on the position, not on the dynamics, of the nuclei. As a result of this approximation, Eq. 2.2 separates into two equations, a time-dependent Schrödinger equation for the motion of the nuclei, and a time-independent Schrödinger equation for the electronic dynamics.

### Approximation II: Force fields

The Born-Oppenheimer approximation allows to treat the electronic wave function of a system as a function of only the nuclear coordinates. Within this framework, the calculation of electronic energies requires the solvation of the time-independent Schrödinger equation for the electrons. This, however, is prohibitive due to the large number of electrons in proteins.

As a second approximation, therefore, a classical force field for the calculation of the potential energy of the system as a function of nuclear coordinates is derived from fitting to the quantum-mechanical ground state energy. The force field is based on a simple 'ball-and-spring' model for atoms connected *via* bonds [57]. For details and recent advancements we refer to the review of Jørgensen et al. [58].

Fig. 2.1 illustrates the components of the force field comprising bonded interactions, namely energy terms for bonds,  $V_{\text{b}}$ , angles,  $V_{\text{a}}$ , proper and improper dihedrals,  $V_{\text{dih}}$ , and additional non-bonded interactions, specifically

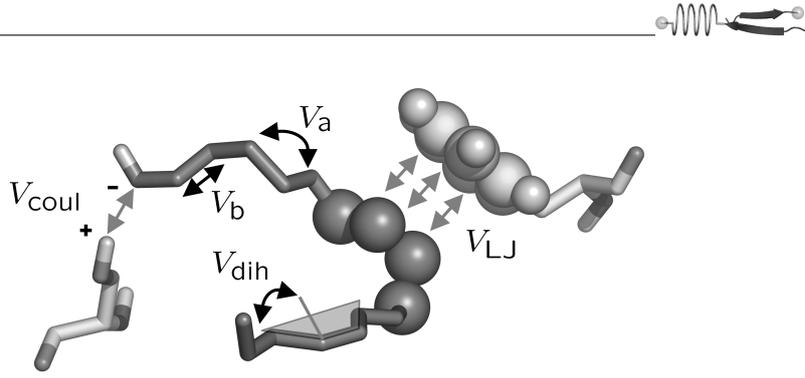


Figure 2.1: Scheme to illustrate the components of a typical mechanical force field as used for the simulations in this thesis considering as example the pheromone bombykol and selected protein side chains. Black arrows: bonded interactions:  $V_b$ , bond-stretching potential,  $V_a$ , angle-bending potential,  $V_{dih}$ , dihedral (out-of-plane) potential. Gray arrows: non-bonded interactions:  $V_{Coul}$ , Coulomb potential,  $V_{LJ}$ , Lennard-Jones potential.

non-polar Lennard-Jones interactions,  $V_{LJ}$ , and electrostatic interactions between partial charges of the atoms,  $V_{Coul}$ . It is given by

$$\begin{aligned}
 V(R) &= V_b + V_a + V_{dih} + V_{Coul} + V_{LJ} & (2.3) \\
 &= \sum_{\text{bonds}} \frac{k_i}{2} (l_i - l_{i,0})^2 \\
 &+ \sum_{\text{angles}} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 \\
 &+ \sum_{\text{dihedrals}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \\
 &+ \sum_{\text{atoms } i} \sum_{\text{atoms } j > i} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi r_{ij}}.
 \end{aligned}$$

The force field parameters for bonded interactions comprise the equilibrium bond length  $l_{i,0}$  and angle  $\theta_{i,0}$ , the respective force constants  $k_i$ , and the multiplicity  $n$ , the torsional barrier height  $V_n$  and the phase  $\gamma$  of the dihedrals. Non-bonded interactions are parametrized in terms of partial charges  $q_i$  for Coulombic interactions, and the parameters  $\epsilon_{ij}$  and  $\sigma_{ij}$ , defining the depth and width of the Lennard-Jones potential, respectively [59]. Parameters are derived by fitting to experimental thermodynamic or structural quantities, and/or to higher (quantum) level calculations [58].

A number of force fields have been developed differing in the strategy of parametrization, and the exact formulation of some of the energy terms in Eq. 2.3 [60–63]. Here, titin kinase and the pheromone-binding protein were

modeled using the GROMOS96 force field [62]. It has been developed and refined for proteins by comparison to thermodynamic properties, mainly solvation free energies, of amino acids, and belongs to the class of united-atoms force fields, in which aliphatic non-polar hydrogens are unified with the carbon they are bound to. The resulting CH, CH<sub>2</sub>, and CH<sub>3</sub> groups are treated as a single 'atom' with appropriately adjusted mass, charge and Lennard-Jones parameters. This united-atom approach proved useful for the simulation of protein dynamics over more than a decade and is under constant development, and also was applied for the simulations of titin kinase and BmPBP. The OPLS force field [63] was developed for proteins on the basis of quantum mechanical calculations of small chemical entities, and treats all atoms explicitly. It here was used for the simulation of the photoisomerization in asFP595, which required a consistent and accurate all-atom force field for both protein and chromophore.

The force fields described above have been developed for proteins on the basis of their building blocks, amino acids. Non-protein parts of the system, such as cofactors, ligands, or non-native amino-acids, require additional parameters. Depending on the desired accuracy, they can be either derived *ab initio* from experimental data and quantum mechanical calculations, or, if possible, simply adopted from parameters of amino acids or other molecules available in the force field. Bombykol, the physiological ligand of the pheromone-binding protein, has been parametrized to simulate possible pathways of ligand release (Chapter 5). While its polar head is similar to the serine side chain, the two conjugated double bonds of the apolar tail are unique. Available serine parameters for the former, and crystallographic data and quantum mechanical results for the latter were applied. The covalently bound chromophore of asFP595 (Chapter 6) is formed from three amino acids, methionine, tyrosine, and glycine. It shows similarity to these residues and histidine, and thus was parametrized in accordance to the OPLS force field of these residues. Charges of the chromophore, instead, were derived quantum-mechanically.

The molecular mechanical force fields applied in this thesis suffice to describe the potential energy of proteins for the simulation of protein dynamics. Their application to quantum processes involving substantial changes of the wave function, such as chemical reactions and photo-excitation, however, are not forthright feasible. Nevertheless, the force fields in principle permit to implicitly incorporate quantum effects to some extent. Here, excited state properties were successfully modeled within the classical OPLS force field, to study the photo-induced isomerization of the fluoroprotein asFP595, as



described in Chapter 6. This example demonstrates the usefulness, and, at the same time, the limit of today's molecular mechanical force fields.

### Approximation III: Classical dynamics for the nuclei

The size of proteins prohibits the solution of the time-dependent Schrödinger equation for the motion of the nuclei. As the third approximation, the nuclear dynamics are therefore described classically by Newton's second law,

$$-\frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2}, \text{ or} \quad (2.4)$$

$$F_i = m_i a_i, \quad (2.5)$$

where  $V$  is the potential energy of the system, as given in Eq. 2.3, as a function of all atomic coordinates,  $r_i$  and  $m_i$  are the coordinates and the mass of atom  $i$ . The force  $F_i$  acting on this atom at position  $r_i$  determine its acceleration  $a_i$ , *i. e.*, which in turn yields the change of the velocity and position of atom  $i$  within a given time step  $dt$ . For the efficient numerical integration of the Newton's equations of motion, the Verlet algorithm [64] is applied here.

The integration time step has to be chosen such that it is small in comparison to the fastest motions of the systems. Bond vibrations involving hydrogen occur within several femtoseconds, and restrict the time step to  $\sim 1$  fs. A number of algorithms to constrain bond lengths have been developed that allow larger time steps. MD simulations presented in this thesis apply the LINCS algorithm and a time step of 2 fs [65].

### Simulation details

The natural environment of proteins is, apart from membrane proteins and interactions to other macromolecules, water. For simulating a model system as close as possible to the *in vivo* system, water molecules and sodium chloride in physiological concentration are added to the system for solvating the protein. Explicit solvation of the protein poses the problem, how to treat the system boundary such that, without the introduction of large artefacts, the number of solvent molecules, and thereby the computational expense for the solvation, are kept small. Two major approaches are the use of a boundary potential that maintains the volume and shape of the simulation system, or the use of periodicity. Periodic boundary conditions are achieved by infinitely replicating the simulation box, and evaluating the forces by

taking interactions of a particle within a single copy, with its nearest image and additional long-range interactions (see below) into account. The periodicity removes errors from artificial boundaries, but might entail internal correlations if the box size is so small that the mirror particles 'see' each other. For the MD simulations of this thesis, cuboid boxes with a minimal distance of the protein to the borders of 1.2 nm were used.

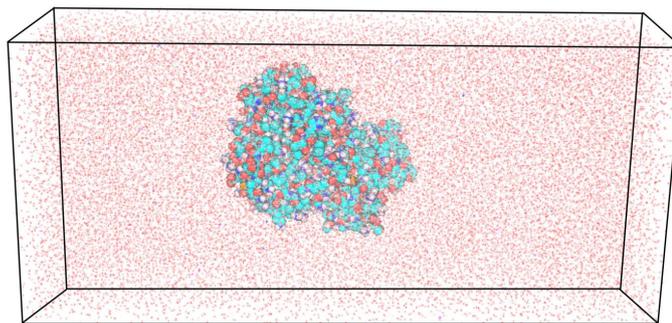


Figure 2.2: Simulation system of the FPMD simulations. Protein atoms are shown as spheres, water molecules as lines.

Simulating transitions involving large-scale conformational changes, such as protein unfolding or ligand unbinding, requires a simulation box large enough to accommodate each of the conformations occurring during the transition. The simulation system for examining force-induced unfolding of titin kinase has therefore been chosen such that the box dimension in the direction the force acts onto the protein is sufficiently large, here 186 Å (Fig. 2.2). For this purpose, the simulation system obtained from equilibrating the protein in water was enlarged by adding a box of water and ions on both side of the titin kinase, and re-equilibrating the solvent only. The same procedure was applied to the bombykol-PBP system by enlarging the simulation box on the side the ligand is pulled out of the protein cavity. Additionally, in the case of titin kinase, parts of the unfolded peptide segments were deleted before crossing the box borders to continue the unfolding simulation without the need to enlarge the box again. We refer to the respective sections for details (3.2.1 and 5.2.2).

The computationally most expensive part of the force calculation at each integration step is the evaluation of the non-bonded interactions, the last sum of Eq. 2.3, as it requires  $N^2$  summations for the  $N$  atoms of the system. A simple and routinely applied way to enhance the computational efficiency is to restrict the calculation of non-bonded interactions to atoms within a



defined cut-off, typically in the range of 1.0 to 1.4 nm [61]. In contrast to the fast decaying Lennard-Jones potential (with  $r^{-6}$ ), the use of a cut-off for the Coulomb potential that slowly decays with  $r^{-1}$  has been shown to cause severe errors [66, 67].

We therefore applied the grid-based Ewald summation [68]. Originally formulated for the calculation of electrostatic potentials in periodic crystals, this method splits the electrostatic interaction into two contributions, of which the interactions within the cut-off are calculated directly, *via*  $\sum_i \sum_{j>i} \frac{q_i q_j}{4\pi r_{ij}}$ , as given in Eq. 2.3, and the long-range interactions outside of the cut-off are calculated in the reciprocal space. Here, the related Particle Mesh Ewald (PME) method as implemented in Gromacs was employed, which, with a scaling of  $N \log N$ , is again more efficient than the Ewald summation by using fast Fourier transformations for the calculation of the reciprocal sum [69, 70]. The FEP method, described in detail in Section 2.5.3 below, involves the smooth conversion of partial charges in case they differ in the initial and final state. The evaluation of the electrostatic contribution to the total free energy change on the basis of PME is not yet implemented in Gromacs. This currently prohibits the application of PME in this case. For the free energy calculations of bombykol and bombykal binding to BmPBP and of the trans-cis isomerization of asFP595, a cut-off was employed, of which the associated inaccuracies remain to be evaluated by comparison to PME.

Due to the approximations in evaluating non-bonded interactions, and the numerical integration, the total energy of a simulation system is not *per se* conserved. To adjust the simulation temperature  $T$  accordingly, the system was coupled to a heat bath of the reference temperature  $T_0$  with the coupling time constant  $\tau_T$ . Here, using the algorithm of Berendsen et al. [71], the atomic velocities were rescaled with

$$v' = v \sqrt{1 + \frac{\Delta t}{\tau_T} \left( \frac{T_0}{T} - 1 \right)}, \quad (2.6)$$

where  $\Delta t$  is the integration time step. The coupling strength is set by the coupling constant  $\tau_T$ , here 0.1 ps.

Temperature coupling assumes the system to be thermally equilibrated, and removes temperature differences. At thermal equilibrium, the different parts constituting the simulation system, protein, water, and ions, are at the same temperature, which in most of the simulations presented here is separately adjusted for these parts. In contrast, a system in which part of the particles actually *are* at a different temperatures is the chromophore of asFP595

after absorption of a photon. Electronic excitation is followed by a relaxation from the Franck-Condon region on the excited state surface. A significant part of the photonic energy thereby is converted into kinetic energy of the chromophore. To simulate the dynamics of the 'hot' chromophore, as described in detail in Section 6.2.2, correspondingly higher velocities were assigned to these atoms. The chromophore was decoupled from the heat bath; and thermal equilibrium sets in more slowly by collisions with the particles surrounding the chromophore. The possibility of coupling parts of the system to separate heat baths of different temperature was also useful for replica-exchange MD simulations of the bombykol-BmPBP system (see Section 2.3).

## 2.2 Monte Carlo Simulations

MD methods simulate the system as it evolves in time by calculating the motion of the atoms from the forces acting on them (Eq. 2.5). The system then moves on the potential energy surface and overcomes barriers not larger than a few  $k_B T$ . Alternatively, the phase space of the system can be explored by a random walk on the potential energy surface. In Monte Carlo (MC) methods, a new point in phase space is generated by subjecting one or more atoms of the system to a random move. Coordinates are modified by adding a random Gaussian distributed step, with a given average step size and an equal probability of forward and backward steps. The new conformation is accepted if the potential energy  $E$  decreases in this MC step, *i. e.*, if  $E(i+1) > E(i)$ . Otherwise, the step is accepted if the Boltzmann factor,  $e^{-\beta(E(i+1)-E(i))}$ , is larger than a random number between 0 and 1, otherwise rejected. The Metropolis criterion [72], which generates a Boltzmann-distributed conformational ensemble, in short reads

$$P = \min(1, e^{-\beta(E(i+1)-E(i))}), \quad (2.7)$$

where  $P$  is the acceptance probability of an attempted step, and  $\beta$  denotes the inverse temperature,  $\beta = \frac{1}{k_B T}$ , with  $k_B$  the Boltzmann constant.

MC methods are less computationally expensive than MD methods, given the same potential energy function to model the system, since forces do not have to be evaluated. The high dimensionality of the phase space of proteins, however, renders the application of MC methods to the simulation of protein motions difficult, since the generation of new conformations with a sufficiently high acceptance rate and average step size is non-trivial.



Nevertheless, MC simulations proved useful for systems with less degrees of freedom.

In this thesis, force probe MC (FPMC) simulations were carried out to validate the conclusions derived from FPMD simulations of titin kinase (Chapter 3). FPMD simulations suggested the different mechanism of  $\beta$ -sheet ruptures as the main cause for their different mechanical resistance. To test if this is actually the case, the all-atom molecular mechanical model has been simplified further to yield a two-dimensional potential energy surface of the  $\beta$ -sheets, which was explored using the MC algorithm and subjecting the  $\beta$ -sheets to an additional pulling force. Details are given in Sections 3.2.2 and 3.3.4.

## 2.3 Replica-Exchange Molecular Dynamics Simulations

Within the short nanosecond time scale accessible to conventional MD simulations, the system often cannot overcome larger energy barriers to regions of the configurational space that are sampled at physiological conditions. In this case, the obtained conformational ensemble might not be quite complete. Several MD simulation methods to enhance conformational sampling have been developed and extensively reviewed [73–75], *inter alia* replica-exchange MD (REMD) simulations, which make use of increased temperatures and were originally developed for folding simulations of small peptides [76–79]. More recent applications are, *e. g.*, the simulation of peptide aggregation [80], of membrane insertion of a peptide [81], or the calculation of NMR order parameters [82].

In REMD simulations, a number of copies (replicas) of the system is simulated simultaneously at different temperatures with conventional MD. Pair-wise exchange of replicas is attempted repeatedly after a number of MD steps. This allows the system that is eventually trapped within an energy minimum at low temperature, to escape this minimum at higher temperatures by surmounting the barrier towards another minimum. The exchange probability is calculated using the Metropolis criterion as in MC simulations (see Eq. 2.7).

Successful exchanges of replicas, the prerequisite for efficient sampling, require a sufficient overlap of their potential energy distributions. The overlap is large enough for satisfying acceptance ratios (10 – 50 %) in case of systems

with a small number of degrees of freedom and small temperature differences between the replicas. Conventional REMD simulations are therefore restricted to systems comprising not more than a few thousand atoms, using implicit solvent and/or small peptides, typically in the context of peptide folding [77, 78]. In this thesis, however, the large size of the system under investigation, the average sized BmPBP in explicit solvent, required an additional reduction in the number of degrees of freedom for a sufficient energy overlap. To this end, the solvent was simulated at the same, physiological, temperature in all replicas, while the protein and ligand, the part of the system of interest, were simulated at different temperatures. By coupling these two parts of the system to separate heat baths (Eq. 2.6), the temperature difference was maintained.

## 2.4 Simulating force-induced transitions

By means of the three approximations described above, and due to the ever increasing computational power at hand, protein dynamics can nowadays be routinely assessed on nanosecond time scales. Some of the processes we are interested in here, such as ligand release or protein unfolding, however occur more slowly. The strategy applied here to also describe events that are slow from the viewpoint of molecular simulations, is to accelerate the transition by means of additional forces that act into the designated direction.

A number of simulation methods have been developed for this purpose that differ in the choice of the reaction coordinate along which the process is enforced to proceed, and the way the force is applied. In this thesis, two of them, FPMD (and FPMC) and ED simulations, were used and are described below.

### 2.4.1 Force Probe Molecular Dynamics Simulations

In FPMD simulations, parts of the system, one or more pull groups, are enforced to move along a pre-defined pulling direction. This allows to simulate otherwise slow processes within the accessible time scale of MD simulations, provided the reaction coordinate can be approximated by a linear pulling direction in Cartesian coordinates. Additionally, FPMD is particularly appropriate for the simulation of proteins of which the dynamics under force as such are physiologically and/or experimentally relevant. Proteins are subjected to force (i) *in vivo* — as molecular force sensors or molecular



machines with other mechanical functions, *e. g.*, ubiquitin and immunoglobulin — and (ii) in single-molecule experiments, using AFM (see Chapter 1), optical/magnetic tweezers, or others.

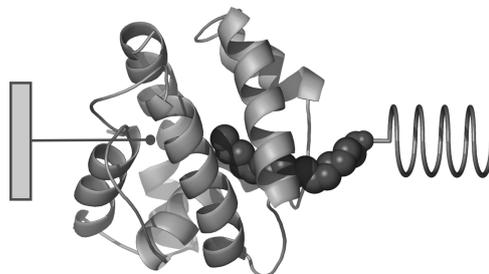


Figure 2.3: Example for FPMD simulations. Bombykol dissociating from its receptor, BmPBP, when subjected to a harmonic spring potential. The COM of the protein BmPBP is kept fixed.

Fig. 2.3 shows an example of an FPMD simulation system.

One or more pull groups  $i$ , comprising one or more atoms of the system, are subjected to a harmonic spring potential,

$$V_{\text{spring},i}(t) = k_0[z_i(t) - z_{\text{spring},i}(t)]^2, \quad (2.8)$$

where  $k_0$  is the force constant of the spring,  $z_i(t)$  the position of the center-of-mass (COM) of the pulled atoms, and  $z_{\text{spring},i}$  the position of the spring.

Mechanical stress is applied by moving the spring(s) with constant velocity  $v$  along the pulling direction,  $z_{\text{spring},i}(t) = z_i(0) \pm \Delta z(t)$ , where  $\Delta z(t) = vt$  is the spring dislocation. Experiments often refer to the loading rate,  $\frac{dF}{dt}$ , which for weak springs is defined as  $k_0v$ . Forces

$$F_i = k_0 [z_i(t) - z_{\text{spring},i}(t)] \quad (2.9)$$

at the pull groups are recorded every time step. Eqs. 2.8 and 2.9 identically apply to FPMD simulations. In this case, the spring is moved with a constant rate per attempted step, including those rejected according to the Metropolis criterion.

For the force-induced unfolding of titin kinase (Chapter 3), the  $C_\alpha$ -atoms of the terminal residues,  $i = C, N$ , were subjected to harmonic potentials, which are moved in opposite directions. This 'two-side pulling' approach differs from the usual FPMD simulations for protein unfolding, where only one terminus is pulled, and the other one is held fixed. This has the advantage that friction arising from dragging the protein through water is significantly

reduced. Bombykol unbinding (Chapter 5) was induced by subjecting the COM of bombykol to the moving harmonic potential, while constraining the COM of the protein to the initial position to prevent it from being dragged through the water together with the ligand.

Force profiles are obtained using Eq. 2.9. Fluctuations at the resonance frequency of the spring lead to artifacts in the profiles, and were suppressed by smoothening the forces with a Gaussian function. Rupture or unfolding forces are the maxima of the force profiles, the error of which is estimated from the uncertainty in rupture forces after smoothening with varying Gaussian distribution widths [42, 43].

Parts of the peptide chains have been omitted in FPMD and FPMC simulations, in the first case for keeping the simulation system reasonably sized, and in the second case for a further simplification of the  $\beta$ -sheet model. To account for their elasticity, the worm-like-chain (wlc) model was applied. The wlc model was originally developed to describe the elasticity of polymer chains, and is routinely used to fit force profiles from protein unfolding events [11]. It relates the force  $F_{\text{wlc}}$  to the extension of the chain,  $x$ , with

$$F_{\text{wlc}} = \frac{1}{\beta p} \left[ \frac{1}{4(1-x/L)} - \frac{1}{4} + \frac{x}{L} \right], \quad (2.10)$$

where  $L$  denotes the contour length of the peptide chain,  $\beta$  the inverse temperature, and  $T$  the temperature,  $p = 0.4 \text{ nm}$  the typical persistence length of peptides. Deviations from the wlc model can be interpreted as particular mechanical resistance apart from entropic effects along the unfolding path.

## 2.4.2 Theory of force probe experiments

As stated previously, MD simulations of larger systems are in general restricted to nanosecond time scales. AFM experiments, as they are under way for titin kinase, instead enforce the process to occur several orders of magnitude slower. Simulations and experiments yield unfolding forces that can be directly compared after extrapolating the forces obtained from FPMD simulations to experimental time scales. For the systems investigated in this thesis, experimental forces are, however, not (yet) available for comparison.

A quantitative relation between the loading rate,  $k_0 v$  and the rupture force that allows this extrapolation has been previously deduced from rate theory [44, 83]. It is based on the assumption that the mechanical load does not alter the unfolding pathway, *i. e.*, the overall shape of the energy profile,

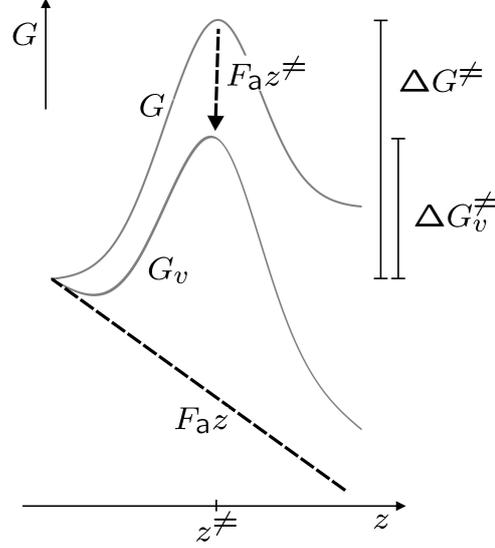


Figure 2.4: The activating force  $F_a$  acting along the reaction coordinate  $z$  accelerates the transition by lowering the activation free energy  $G^\ddagger$  to  $G_v^\ddagger$ .

but only the barrier heights. Kramers' approximation [84],

$$k = \omega_0 e^{-\beta \Delta G^\ddagger} \quad (2.11)$$

relates the rate of a transition  $k$  to the activation free energy,  $\Delta G^\ddagger$ , the maximal free energy along a given reaction coordinate. Kramers' pre-factor  $\omega_0$  is the attempt frequency of the system due to thermal fluctuations to rescue the initial valley and to overcome the activation barrier, and is related to the curvature of the energy profile at the initial minimum. We now consider a force-activated process, where the reaction coordinate  $z$  equals the pulling direction, and the free energy profile  $G(z)$  of unfolding is modified by the additional activating force,  $F_a$  according to  $G_v(z) = G(z) - F_a z$ . As shown in Fig. 2.4, the activation free energy is reduced to

$$\Delta G_v^\ddagger = \Delta G^\ddagger - z^\ddagger F_a, \quad (2.12)$$

where  $z^\ddagger$  is the position of the activation barrier on the reaction coordinate. Under the assumption that the total external force,  $F_{\text{ext}}$ , comprises the activating force  $F_a$  and an additional friction force, one obtains

$$F_{\text{ext}}(v) = \gamma v + \frac{1}{\beta z^\ddagger} \ln \frac{v}{k \Delta z^\ddagger}, \quad (2.13)$$

where  $\Delta z^\ddagger$  is the scatter width of  $z^\ddagger$ , and given by  $v/k_v$ , and  $\gamma$  is the friction coefficient.

Enforcing the reaction on small time scales, *i. e.*, with high pulling velocities  $v$ , yields forces dominated by friction. In this thesis, we aimed at assessing lower pulling velocities to obtain forces that linearly depend on  $\ln v$  (compare Eq. 2.13), and therefore can be compared to experimentally observed forces of the same and related proteins.

### 2.4.3 Essential Dynamics Simulations

In FPMD simulations as described in Section 2.4.1, the reaction coordinate along which the system is driven is defined externally. Only in some cases, the pulling direction is known *a priori*. In the case of titin kinase (Chapter 3) the pulling direction simply equals the vector that connects both protein termini. In contrast, for the identification of a ligand exit pathway out of the protein (Chapter 5), the appropriate pulling direction is unknown, a problem that can be only partially solved by sampling and by comparing different directions.

For the identification of ligand release pathways in BmPBP, alternatively, the reaction coordinate was deduced from the system under investigation itself. In ED simulations, the system is driven along a reaction coordinate that represents one or several of the main modes of collective motion observed during free dynamics of the system. The main modes, or principal components, are calculated by diagonalizing the covariance matrix. This Principal Component Analysis (PCA) [85] gives the eigenvectors, of which those with the highest eigenvalues correspond to the collective motion of the system with largest amplitude. PCA is in general a useful analysis tool to detect major conformational changes in proteins apart from high-frequency fluctuations (see, *e. g.*, Section 3.3.3). It is however restricted to linear collective coordinates. Recently, an approach to assess general non-linear collective coordinates was developed by Schröder et al. (submitted), which was applied here to detect the non-linear correlated bombykol-BmPBP dynamics (Section 5.3.1).

Fig. 2.5 shows the basic principle of ED sampling. One or more eigenvectors are defined as the subspace in which sampling shall be enhanced. Each of the eigenvectors of this subspace is sampled along a given direction, and with a given step size. This is either achieved by enforcing a step-wise fixed increase, or, alternatively, by accepting only those MD steps that show no change or any increase along the subset of collective coordinates. The latter restricts the system to a smaller extent and therefore was applied here (see Section 5.2.2). The perpendicular degrees of freedom are unperturbed and

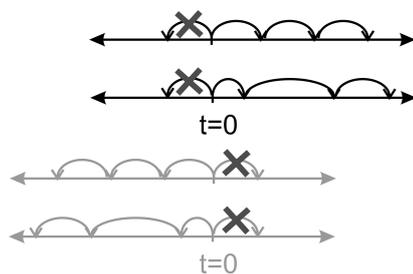


Figure 2.5: Principle of ED. Linear sampling along a chosen eigenvector in forward (black) or backward (gray) direction, with fixed (top) or variable (bottom) increment.

can equilibrate during the ED simulation. The parameter that influences the simulation time, in which the motion of a certain amplitude occurs, is given by the step size, which, in the latter case of an acceptance criterion, represents an optional upper limit for the progressive motion.

## 2.5 Free energy calculations

Forces, as obtained from FPMD simulations of protein unfolding or protein-ligand dissociation, measure the mechanical stability of the system. They thereby give insight into the properties of the system under mechanical load, *i. e.*, under non-equilibrium conditions. The free energy of a system, in contrast, is an equilibrium quantity, a state function independent of the pathway leading to this state. The free energy of state X is defined as  $G^X = -\frac{1}{\beta} \ln Z^X$ , where  $Z^X$  is the partition function,  $Z^X = \int e^{-\beta H^X(r)} dr$ , with phase space  $r$ , the Hamiltonian  $H^X(r)$ , and the inverse temperature  $\beta = \frac{1}{k_B T}$ . The calculation of the free energy from its partition function, however, is prohibitive, due to the complexity of the Hamiltonian in the high-dimensional phase space of a protein. Instead, free energy differences between a state A and a state B,  $\Delta G = G^B - G^A$ , can be more readily estimated from driving the system from A to B. Three calculation methods, the Jarzynski relation, potentials of mean force (PMF), and FEP, were applied in this thesis to the bombykol-BmPBP system (Chapter 5), and are outlined in the following Sections 2.5.1 to 2.5.3.

### 2.5.1 Free energies from non-equilibrium trajectories

Driving the system from state A to state B in finite time, *i. e.*, under non-equilibrium conditions, involves friction. Part of the force externally applied to induce the transition does not add to the free energy difference but irreversibly goes into dissipated heat. Thus, the spent mechanical work equals to or is larger than the difference in free energy,  $\Delta G \leq W$ , where the equality holds for infinitely slow transitions at equilibrium.

The Jarzynski equality,

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta G}, \quad (2.14)$$

instead holds also for the non-equilibrium case [86]. This remarkable relation in principle allows to calculate free energy differences from non-equilibrium experiments and simulations [87,88]. It was applied here to calculate the free energy for the dissociation of bombykol from BmPBP from the forces obtained from FPMD simulations. Free energy calculations using the Jarzynski equality, however, are hampered by the slow convergence of the exponential average, especially when applied to processes far from equilibrium. To discuss the errors involved in the calculations, its original derivation [86] is briefly shown here.

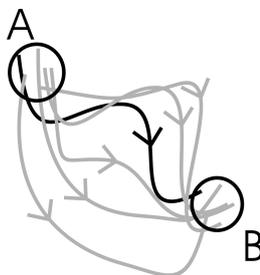


Figure 2.6: Ensemble of trajectories for the transition from A to B for the free energy calculation using the Jarzynski equation.

As shown in Fig. 2.6, we consider an ensemble of paths leading from state A to state B within a switching time  $t_s$ . The switching parameter  $\lambda$  parametrizes the reaction coordinate from the initial ( $t = 0, \lambda = 0$ ) to the final state ( $t = t_s, \lambda = 1$ ). The work applied for a single trajectory of this ensemble to the point  $z$  in phase space is the force, *i. e.*, the derivative of the Hamiltonian with respect to  $\lambda$ , integrated over the whole path,

$$w(z, t_s) = \int_0^1 \frac{\delta H_\lambda(z, t)}{\delta \lambda} d\lambda. \quad (2.15)$$



The Boltzmann average of the work over the ensemble of trajectories is given by

$$\langle e^{-\beta W} \rangle = \int f(z, t_s) e^{-\beta w(z, t_s)} dz, \quad (2.16)$$

where  $f(z, t_s)$  denotes the phase space density. According to the Liouville theorem of Hamiltonian systems, the phase space density is conserved in time along a trajectory,

$$f(z, t) = f(z_0, 0) = \frac{e^{-\beta H_0(z_0)}}{Z_0}. \quad (2.17)$$

With this relation and the work along one trajectory

$$w(z, t_s) = \int_0^1 \frac{\delta H_\lambda}{\delta \lambda} d\lambda \quad (2.18)$$

$$= H_1(z, t_s) - H_0(z_0), \quad (2.19)$$

one obtains

$$\langle e^{-\beta W} \rangle = \int Z_0^{-1} e^{-\beta H_1(z)} dz \quad (2.20)$$

$$= Z_1 Z_0^{-1} \quad (2.21)$$

$$= e^{-\beta \Delta G}, \quad (2.22)$$

the above Jarzynski relation.

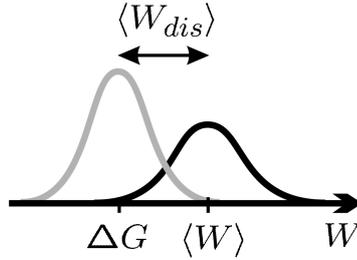


Figure 2.7: Work distributions to illustrate the convergence problem of Jarzynski free energy estimates. The distribution of the exponential work (gray) is shifted with respect to the work distribution (black). The respective averages, the free energy,  $\Delta G$ , and the average work,  $\langle W \rangle$ , respectively, differ in the average dissipated work,  $\langle W_{dis} \rangle$ .

The Jarzynski equation holds for infinite sampling of transitions from A to B. As illustrated in Fig. 2.7, the exponential average is particularly sensitive towards rare trajectories in the region of low values of the work distribution. The higher the irreversible part of the work, the more is the work distribution

shifted with respect to the distribution of the exponential work, and the slower does the sampling converge. The shift equals the dissipated work,  $W_{dis}$ .

Convergence was here tried to be improved by calculating the free energy difference not directly from the exponential average, but from the cumulant expansion [86, 89, 90],

$$\log\langle e^x \rangle = \langle x \rangle + \frac{1}{2}\langle (x - \langle x \rangle)^2 \rangle + \dots, \quad (2.23)$$

up to the second order, which gives

$$\begin{aligned} \Delta G &= \langle W \rangle - \frac{1}{2}\beta\langle (W - \langle W \rangle)^2 \rangle \\ &= \langle W \rangle - \frac{1}{2}\beta\sigma_W^2, \end{aligned} \quad (2.24)$$

where  $\sigma_W$  is the variance of the work distribution. By using the second order cumulant expansion, the free energy estimate is not dominated by the tails, but by the width of the distribution, which can be more easily estimated from a rough sampling. As can be seen from Eq. 2.24, the average dissipated work,  $\langle W_{dis} \rangle = \langle W \rangle - \Delta G$ , is directly related to the variance of the work distribution,  $\langle W_{dis} \rangle = \frac{1}{2}\beta\sigma_W^2$  (see Fig. 2.7).

The Jarzynski relation allows to calculate free energies from simulations far from equilibrium. This saves computational power, that in turn, however, has to be invested into sufficient sampling of transition paths. Previous studies showed that free energy calculations on the basis of the Jarzynski equation are comparable to conventional equilibrium methods in terms of accuracy at similar computational expense [88, 90]. In this thesis, free energy profiles calculated with Eq. 2.14 from FPMD simulations for bombykol release from BmPBP were compared to the PMF obtained from ED simulations (see next section for details), and found to achieve comparable accuracy (Chapter 5).

## 2.5.2 Free energies from equilibrium trajectories

ED simulations yielded non-equilibrium forces necessary to drive the system forward or backward along a principal mode. Equilibrium forces were estimated from fixing the essential coordinate at equidistant values  $z_n$ , while allowing the remaining  $3N - 1$  degrees of freedom, with  $N$  the number of atoms, to equilibrate. The mean force of each of the  $n$  points along the principal mode then was approximated by the force average of this window. An initial equilibration phase of each window was excluded from the calculation



of the average force. The PMF is given by the sum over the average forces of the  $n$  windows,

$$\Delta G_{\text{PMF}} = \sum_n \langle F \rangle \Delta z, \quad (2.25)$$

where  $dz$  is the distance between sampling points along the principal mode  $z$ , and  $\langle F \rangle$  the mean force. It here was calculated for the bombykol release from BmPBP, with the first principal mode serving as the reaction coordinate  $z$ .

Convergence can be slow due to non-equilibrium effects and large force fluctuations. It was assessed by comparing PMFs obtained from windows of different lengths and parts of the trajectories. Free energy profiles and convergence for the release of bombykol from BmPBP are discussed in Section 5.3.2

### 2.5.3 Free energy perturbation

The calculation of the free energy difference between two different states A and B from their partition function, with  $\Delta G = RT \ln \langle e^{\beta H^A(r) - H^B(r)} \rangle$  requires infinite sampling of the phase space and is therefore computationally prohibitive, as stated before. Instead, an initial system A (parametrized with  $\lambda = 0$ ) was slowly transformed, or 'mutated' into a system B ( $\lambda = 1$ ) [38,39], and the free energy difference then was calculated according to

$$\Delta G_{\text{FEP}} = \int_0^1 \left\langle \frac{\delta H(r)}{\delta \lambda} \right\rangle_\lambda d\lambda. \quad (2.26)$$

This formulation assumes the switching to be sufficiently slow to ensure equilibrium conditions (compare the non-equilibrium case of Eq. 2.15).

The perturbation might simply comprise the change in one or a few conformational degrees of freedom the two states differ in. A rotation around selected dihedral angles of the asFP595 chromophore by means of FEP yielded its isomerization and the involved forces (Chapter 6). Concerning protein-ligand complexes, the point of interest here was the comparison of the binding affinity of two different ligands to a protein, as for the binding of bombykol and bombykal to BmPBP (Chapter 5). Fig. 2.8 shows the thermodynamic cycle that allows to calculate the difference of the binding free energies of two different ligands A and B,

$$\Delta \Delta G_b = \Delta G_b^B - \Delta G_b^A \quad (2.27)$$

from the free energy changes upon 'mutating' the ligands in their unbound states,  $\Delta G_{\text{aq}}^{A2B}$ , and bound to the protein,  $\Delta G_{\text{complex}}^{A2B}$ , with

$$\Delta \Delta G_b = \Delta G_{\text{aq}}^{A2B} - \Delta G_{\text{complex}}^{A2B}. \quad (2.28)$$

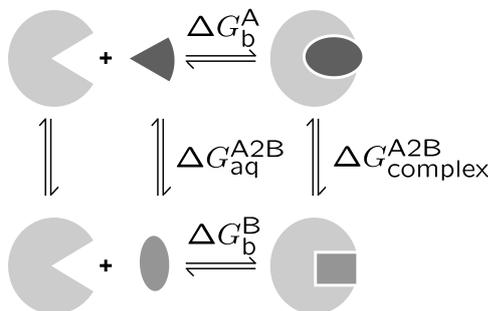


Figure 2.8: Thermodynamic cycle to calculate the difference in binding free energies of two ligands from FEP simulations of the unbound and bound ligands into each other, respectively, using Eq. 2.28.

The 'mutation' is introduced by interpolating between the two ligand topologies, here between bombykol and bombykal. Within the time course of one FEP simulation, the interpolation parameter  $\lambda$  is gradually changed from 0 to 1 with an identical increase of  $\lambda$  at each MD step (conventionally referred to as slow-growth thermodynamic integration). The bombykol (ol) and bombykal (al) bonded potential energies are parametrized by  $\lambda$  yielding a total bonded potential  $V_b$ ,

$$V_b = (1 - \lambda)V_{b,ol} + \lambda V_{b,al}. \quad (2.29)$$

The hydrogen atom of the bombykol hydroxyl group is mutated to a dummy atom, *i. e.*, its charges and Lennard-Jones parameters are changed to zero. The appearance or disappearance of this atom close to  $\lambda = 0$  and  $\lambda = 1$  gave rise to singularities of the non-bonded potentials when interpolated linearly. To remove this problem, non-bonded interactions were calculated with soft-core potentials [91–93], here given by

$$V_{nb} = (1 - \lambda)V_{nb,ol}(r_{ol}) + \lambda V_{nb,al}(r_{al}) \quad (2.30)$$

with the soft-core distances

$$r_{al} = (\alpha\sigma_{al}^6(1 - \lambda)^2 + r^6)^{1/6} \quad (2.31)$$

and

$$r_{ol} = (\alpha\sigma_{ol}^6(1 - \lambda)^2 + r^6)^{1/6}. \quad (2.32)$$

Here,  $r$  is the inter-atomic distance used for calculating the conventional non-bonded potentials. The soft core parameter  $\alpha$  controls the height of the potential at zero distance and was set to 1.51 [93]. The interaction radius  $\sigma$  is given by the Lennard-Jones parameters, with  $(C_{12}/C_6)^{1/6}$ , for  $C_{12}, C_6 \neq 0$ , otherwise  $\sigma = 0.3$ .



## 3

# Enforced Titin Kinase Activation

---

### 3.1 Titin and titin kinase

”According to Greek Mythology, the Universe was formed and controlled by Titans - godlike giants who created order from chaos. In the same way, the most gigantic of all proteins, titin, brings order to sarcomeres (the contractile units of the striated muscle) and organizes the machinery for condensation of chromosomes in dividing cells.” (A. R. Means [94])

The basic contractile unit of muscle is the sarcomere, which contains two major types of protein filament, thin and thick (Fig. 3.1a). The thin and thick filaments are crosslinked at the Z-disc and M-line and mainly consist of the motor proteins myosin and actin, which undertake the task of generating an active force for muscle contraction.

Titin, the third type of filament of the sarcomere, is a giant protein of approx. 3 MDa, and the longest covalently linked protein known. In contrast to the motor protein filaments, and in spite of — or due to — its large size, it was first described as late as 1977 (as connectin) [95]. Within the past decade, since the amino acid sequence of titin is known [96], much insight

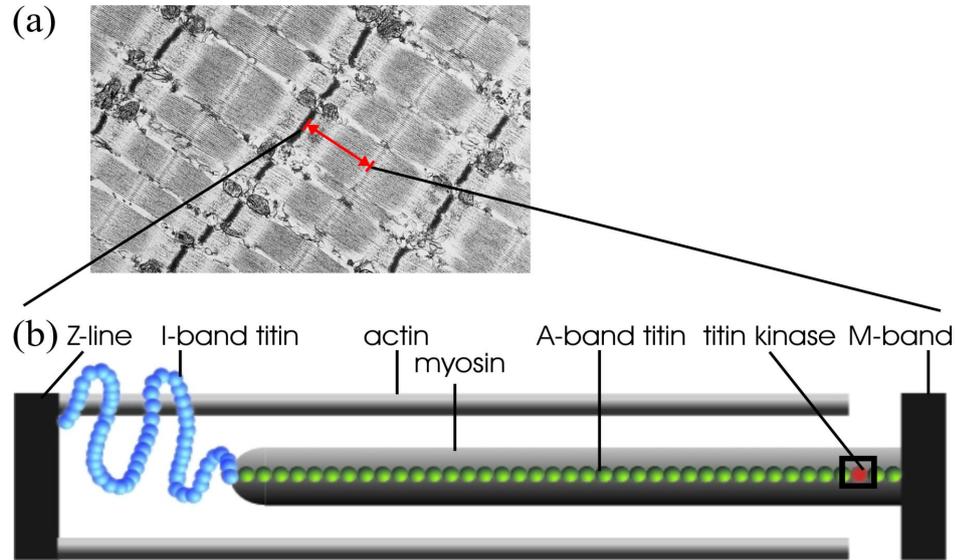


Figure 3.1: Muscle and the muscle filament titin. (a) Electron-micrograph of striated muscle, kindly provided by Mathias Gautel. (b) Half of the sarcomere and the position of titin kinase (red) are sketched. Tandem titin domains are shown as spheres.

into its structure and mechanical properties has been gained. It spans half the sarcomere from the Z-line to the M-band, thereby connecting the other two major muscle filaments, actin and myosin, with each other, as shown in Fig. 3.1b [97,98]. Titin is composed of a series of covalently bound domains (blue and green spheres in Fig. 3.1), of which the number and sequence vary for different muscle types. Human cardiac titin consists of approximately 300 repeating immunoglobulin (Ig) and fibronectin III (Fn3) domains. It shows additional insertions, namely the PEVK segments, an elastic proline-, glutamate-, valine-, and lysine-rich region, and the other unique segments N2A and N2B [99].

Upon muscle cell stretch the titin filament generates passive tension by straightening and partial unfolding of its domains that restores the sarcomere length and realigns the muscle filaments after stress release [100,101]. It thereby keeps the filamentous components organized and plays a major role in muscle elasticity [102].

Single-molecule studies of the mechanical stability and mechanical unfolding of titin have revealed the molecular basis of titin elasticity. I-band Ig and Fn3, and PEVK domains have been extensively studied by optical tweezers and atomic force microscopy (AFM) [8, 9, 11, 103–105]. For a description of



single molecule AFM experiments, see Chapter 1. A sequence of unfolding events of individual domains along the titin filament at forces in the 20 to several 100 pN range were resolved by these techniques. MD simulations [40, 41, 106, 107] identified hydrogen bond ruptures between terminal  $\beta$ -strands as the initial events of Ig unfolding, leading to an unfolding intermediate. Further pulling yields unfolding of the domain in an all-or-none fashion [13, 108]. The high rupture forces required for Ig and Fn3 unfolding found in these experiments indicate that unfolding, if any, occurs only under extreme physiological tension to prevent irreversible damage [102, 109].

Near titin's carboxy terminus, located at the M-band, titin contains its only catalytic domain, the titin kinase (red ball in Fig. 3.1). The crystal structure of the titin kinase domain has been solved [52] and is shown in Fig. 3.2. It exhibits a kinase domain fold that, with a  $\beta$ -sheet rich smaller lobe and an  $\alpha$ -helix rich bigger lobe, is typical for giant kinases and was first found for CAPK [110]. As a member of the serine/threonine kinases, the active site is located between the small and the big lobe. It comprises the catalytic residue, Asp127, that transfers the phosphate moiety from ATP to a serine or threonine of the protein substrate, as well as the well-conserved ATP binding residues. In contrast to most other kinases, titin kinase is autoinhibited by its C-terminal regulatory tail (shown in red), in a way similar to twitchin kinase [111]. The  $\alpha$ R2 helix binds to the kinase ATP binding site and specifically interacts with the residues involved in catalysis. The C-terminal part of the regulatory tail,  $\beta$ R1, forms a  $\beta$ -sheet with  $\beta$ C10 and  $\beta$ C11 (yellow in Fig. 3.2). Similarly, the N-terminal part of titin kinase forms a  $\beta$ -sheet comprising the strands  $\beta$ C1,  $\beta$ C2, and  $\beta$ C3.

Titin kinase activation therefore requires release of autoinhibition, in particular the removal of the auto-inhibitory tail from the active site. It has been shown to occur in differentiating myocytes upon phosphorylation of Tyr170 and  $\text{Ca}^{2+}$ /Calmodulin binding to  $\alpha$ R1, the N-terminal  $\alpha$ -helix of the regulatory tail [52].

The role of titin kinase is largely unknown, and has been only partly uncovered by experiments showing that it influences myofibrillogenesis in differentiating muscle cells by phosphorylation of telethonin [52, 113]. Recent studies also indicate a connection between titin kinase and the signalling pathway of muscle-specific ring finger proteins (MURF) [114–116]. One of the three-members MURF protein family, MURF-1, interacts with the titin domains adjacent to the N-terminus of titin kinase, and downstream, in form of heterodimers with MURF-3, with transcriptional regulators. Deletion of a large segment of M-line titin, including the binding sites for muscle-specific ring

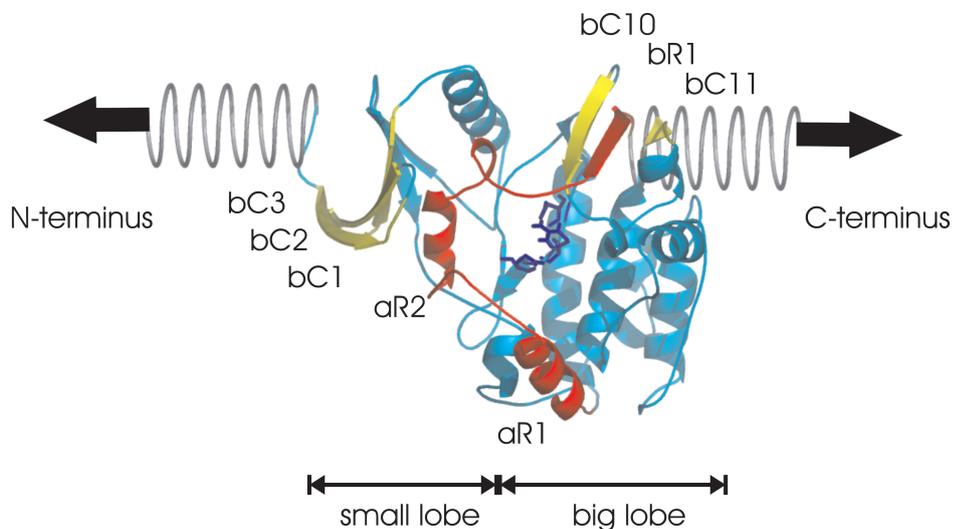


Figure 3.2: Titin kinase [1TKI [52]] in its autoinhibited conformation. The catalytic site is shown in blue, the regulatory tail in red, and the terminal  $\beta$ -sheets in yellow. The harmonic pulling potentials used for the force probe simulations described in the Methods section are symbolized as springs. All proteins were plotted with Pymol [112].

finger proteins (MURF), myomesin and DRAL as well as titin kinase indeed led to low expression and poor organization of myofibrils [117, 118]. These findings point at a role of titin kinase in muscle gene expression signaling in the adult muscle.

Considering the exceptional location of the kinase, which is integrated into the molecular titin 'spring', it has been speculated (M. Gautel, priv. communication) that titin might play the role of a force sensor that passes the information of stress level in the titin filament on *via* substrate phosphorylation. In this scenario, as depicted in Fig. 3.3, structural rearrangements within the titin kinase domain driven by the stretching forces along the titin molecule result in a release of auto-inhibition. The force-induced active conformation allows ATP binding and phosphorylation of the substrate. Very recent findings indeed connect titin kinase to signalling pathways controlling muscle protein expression [119]. Titin kinase was found to bind nbr1, a zinc-finger protein that is involved in a signalling cascade down to the the serum response transcription factor in the nucleus. Transcription of muscle protein was enhanced upon mechanical load on titin. This mechanism corroborates the hypothesis of titin kinase as a force sensor.

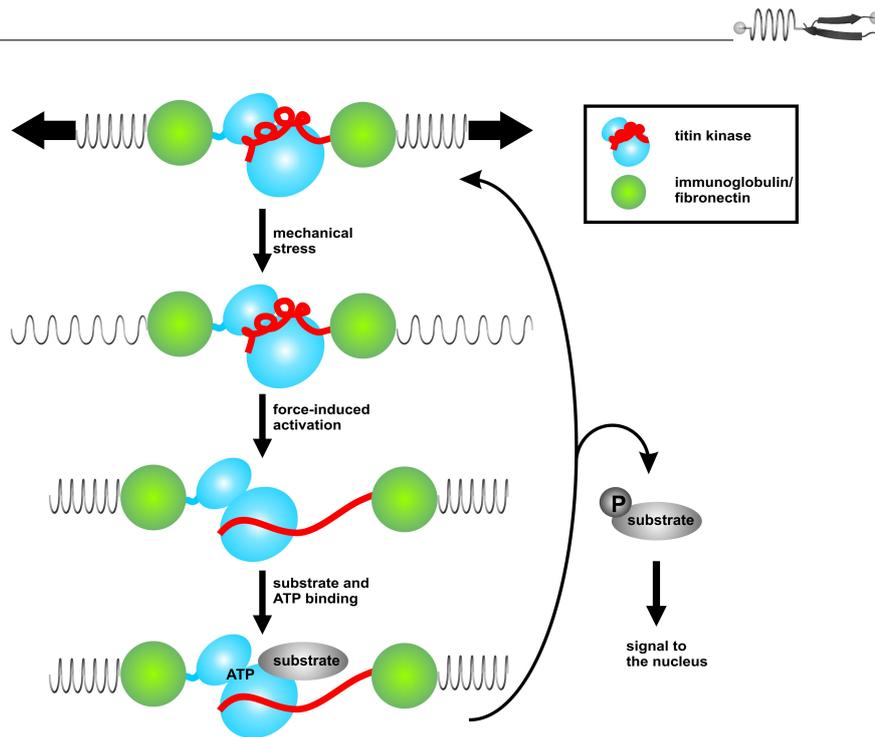


Figure 3.3: Titin kinase as a force sensor. The hypothesis of a force-induced release of the regulatory tail from the active site, that then can bind ATP and substrate for phosphoryl-transfer, is depicted schematically.

Whether this is actually the case and how kinase activation is triggered by release of autoinhibition is the subject of this Chapter. To this aim, FPMD simulations (see Section 2.4.1) of titin kinase partial unfolding were carried out by subjecting both titin kinase termini to a harmonic pulling potential which mimics the tension in a stretched titin filament. The hypothesis of titin kinase as a force sensor would only be corroborated in case the active site of titin kinase is found to not at once fall apart, as observed for Ig and Fn domains, but instead stays intact in the long run during partial unfolding. The simulations aim at (a) predicting the force required for mechanical unfolding, (b) examining the force-induced structural changes in the regulatory tail and kinase domain with regard to kinase activity, and (c) gaining a closer insight into the dependence of force resistance on structural properties.

The titin kinase construct used for crystallization [52] lacks 24 N-terminal residues of the kinase domain. This sequence is resolved in the twitchin kinase crystal structure (pdb-entry 1KOB [111]) and shows high sequence similarity in these two proteins (Fig. 3.4). Twitchin is yet another filament of the muscle, similar to but shorter than titin, and located in the A-band





the steepest descent algorithm was followed by a 400 ps MD simulation with harmonical constraints on the protein heavy atoms with a force constant of  $k = 1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  to equilibrate water and ions. A subsequent MD simulation of 7.5 ns length was performed to equilibrate the whole system, during which the protein backbone RMSD was monitored.

For FPMD simulations a simulation system large enough to accommodate the unfolded protein is required. After rotating the equilibrated system such that the termini were aligned along the box  $z$ -axis, water and ions were added to give a cuboid box of  $78 \times 84 \times 186 \text{ \AA}^3$  in size, containing  $\sim 120,000$  atoms (see Fig. 2.2). A subsequent 200 ps MD simulation with harmonic constraints on the protein heavy atoms during the first 100 ps was carried out. The final structure of this run served as starting structure for the FPMD simulations.

A number of FPMD simulations with different pulling velocities  $v$  were performed,  $0.004 \text{ \AA/ps}$ ,  $0.008 \text{ \AA/ps}$ ,  $0.02 \text{ \AA/ps}$  (twice),  $0.05 \text{ \AA/ps}$  (twice),  $0.1 \text{ \AA/ps}$ , and  $0.5 \text{ \AA/ps}$ , with simulation times ranging from 150 ps to 22 ns, totaling 42 ns of FPMD simulation time.

Simulations were stopped after the distance of one of the termini to the box borders dropped below  $7 \text{ \AA}$ . In order to further extend the FPMD simulations of  $0.004 \text{ \AA/ps}$  and  $0.008 \text{ \AA/ps}$  pulling velocity, pulling was continued with truncated protein structures, and the elasticity of the omitted peptides was described by a wlc model (Section 2.4.1). Here, residues 1–19 and 308–321 of the final snapshots with  $\Delta z_i(t) = 65 \text{ \AA}$  obtained from the  $0.008 \text{ \AA/ps}$  and  $0.004 \text{ \AA/ps}$  simulations were deleted. Terminal ammonium and carboxylate groups of residues 20 and 307, respectively, were built, and SPC water and two sodium ions to substitute the net negative charge of the removed peptide chains were added. The resulting system was energy minimized, and solvent and ions equilibrated with harmonic restraints on the protein heavy atoms as described above.  $C_\alpha$ -atoms of residues 20 and 307 were then subjected to a harmonic spring potential with modified spring constant  $k'_0$  to account for the elasticity of the deleted peptide chains. Accordingly,  $k'_0$  was calculated from the force  $F_{\text{wlc}}$  present at the peptide fragments alone using the wlc model (Eq. 2.10) with

$$k'_0 = \frac{F_{\text{wlc}}}{\Delta d}, \quad (3.1)$$

where  $\Delta d$  denotes the end-to-end distance of the peptide chain.

The contour length  $L$  was chosen as the number of amino acids times length of one amino acid. The resulting spring constant of  $\sim 100 \text{ kJ/mol/nm}^2$  for

the deleted peptide chains is smaller than  $k_0$ , and thus was chosen as the spring constant  $k'_0$  for further pulling.

The simulation setup for twitchin kinase was chosen similar to the setup of titin kinase. The structure was solvated in SPC water with physiological ion strength in a cubic box of  $101 \times 80 \times 83 \text{ \AA}^3$  in size. After energy minimization and a 400 ps MD simulation with harmonic constraints on the protein heavy atoms, a 5.5 ns equilibration run of the whole system was performed. The  $74 \times 85 \times 140 \text{ \AA}^3$  box for the subsequent FPMD simulation contained the equilibrated twitchin kinase structure, with its termini aligned along the z-axis, water and ions ( $\sim 89,000$  atoms). Added water and ions were equilibrated in a 200 ps MD simulation as described above. A single FPMD simulation of 6 ns in length was performed, as also described above, with two pull groups, the  $C_\alpha$ -atoms of residues 1 and 353, a pulling velocity of  $0.008 \text{ \AA/ps}$ , and a spring constant of  $500 \text{ kJ/mol/nm}^2$ .

Protein domains and the hinge axis involved in the hinge motion as obtained from PCA were determined using the program DynDom [125].

Hydrogen bond potentials were calculated from hydrogen-acceptor distances  $\Delta s$  according to Espinosa et al. [126] with

$$E_{\text{pot}} = -\frac{1}{2}(50 \times 10^{-3} \text{ kJ/mol})e^{-36\Delta s}. \quad (3.2)$$

To obtain a measure for the width of the hydrogen bond potential decay,  $a_0$ , hydrogen bond energy traces were fitted with the function  $\tilde{E}_{\text{pot}}(\Delta z) = (\exp(\Delta z - a_2)/a_0 + a_1)^{-1}$ , where  $\Delta z$  is the distance between the current and initial location of the spring.

### 3.2.2 Force Probe Monte Carlo Simulations

To test if the conclusion from the FPMD simulations that the different strength of the terminal  $\beta$ -sheets originates from their different rupture mechanism, two-dimensional models of the traversed energy landscapes were studied by FPMC simulations (for MC and FPMC simulations, see Sections 2.2 and 2.4.1). The rupture of the two  $\beta$ -sheets with six inter-strand hydrogen bonds each (between  $\beta C1$  and  $\beta C2$ , and between  $\beta R1$  and  $\beta C10$ , respectively, see Fig. 3.2) was modeled using a properly chosen two-dimensional effective potential  $G(x_1, x_2)$  (see 3.3.4 and Fig. 3.11 there). The two reaction coordinates for the concurrent and sequential rupture,  $x_1$  and  $x_2 - x_1$ , respectively, were defined as the distances between the terminal amino acids of the  $\beta$ -strands pulled apart. The total length of the two



connected  $\beta$ -sheets then is given by  $x_2$ . The origin was chosen such that the initial  $\beta$ -sheet lengths were zero.

For the FPMD simulations, a harmonic spring potential as a function of the spring position and  $x_2$  was added in a similar manner as in the FPMD simulations (Eq. 2.8), using the same spring constant of 500 kJ/mol/nm<sup>2</sup>. The short peptide sequences between the terminal sheets and the spring in the FPMD simulations are absent in the MC simulations. To take their elasticity into account, their force constants were calculated according to Eq. 2.10 as described previously, but were found to be an order of magnitude larger than the force constant of the spring, and thus to be negligible. Each MC simulation was started with a particle at the minimum of the energy surface  $G(x_1, x_2)$  (reaction coordinate  $x_1 = 0$  and  $x_2 - x_1 = 0$ ). Random Gaussian distributed steps backward and forward along the two reaction coordinates  $x_1$  and  $x_2 - x_1$  with an average step size of 0.0015 nm were accepted or rejected according to the Metropolis criterion [72] at 300 K. Each step was accompanied by a movement of the spring to a larger  $x_2$  with a constant step size. A set of FPMD simulations with spring velocities varying between  $10^{-5}$  and  $10^{-2}$  nm/step was performed, and the spring position at which the  $\beta$ -sheets rupture was monitored and compared to the respective FPMD results.

## 3.3 Results

### 3.3.1 Equilibration and free dynamics

The structure of titin kinase during equilibration and the subsequent 7.5 ns long free MD simulation was compared to the crystal structure. The root mean-square deviation (RMSD) from the crystal structure leveled off at 2.0 Å for the protein backbone atoms. The root mean-square fluctuations per residue showed a correlation coefficient = 0.7 with the crystallographic B-factors, indicating that the overall protein structure and dynamics as observed by X-ray crystallography is also captured in the simulation (data not shown).

The only significant deviation from the X-ray structure was seen for the catalytic site residues. According to the crystal structure, Asp127, the catalytic base, and Arg129, both conserved residues of serine/threonine kinases, form a double salt bridge. At the nanosecond time scale of the free dynamics simulation, this salt bridge is reversibly broken and mostly substituted by a

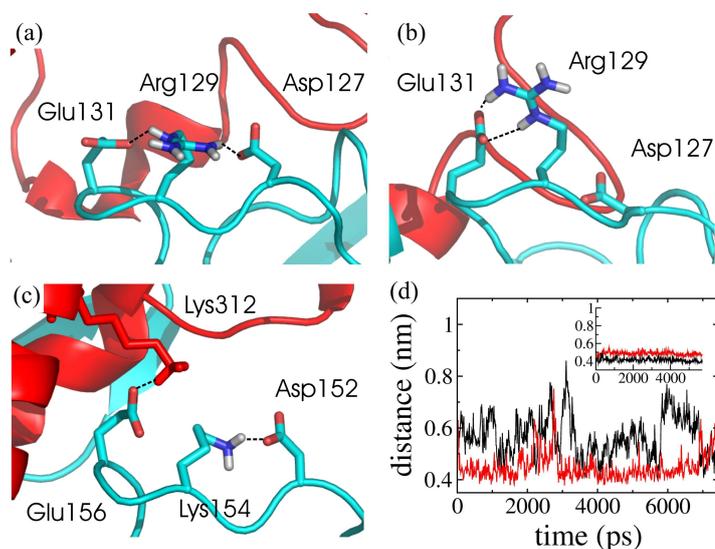


Figure 3.5: (a)-(c) Snapshots of titin kinase at 3.7 ns and 6.0 ns, and of twitchin kinase at 2.7 ns. (d) Comparison of the side chain fluctuations of the titin kinase (main panel) and twitchin kinase catalytic sites (inset). Center-of-mass distances between the side chains of Asp127 and Arg129 (black) and of Arg129 and Glu131 (red) of titin kinase are shown. The inset shows the respective distances between Asp152 and Lys154 (black) and Lys154 and Glu156 (red) of twitchin kinase.

Glu131-Arg129 salt bridge (Fig. 3.5). This catalytic site 'wobbling' was not seen for the free MD simulation of twitchin kinase, where the homologue Asp152-Lys154 salt bridge stayed intact.

### 3.3.2 Unfolding mechanism

Force probe unfolding simulations have been carried out with pulling speeds from  $0.004 \text{ \AA/ps}$  to  $0.5 \text{ \AA/ps}$ ; the slowest pulling simulation covered 22 ns. This time scale is four to six orders of magnitude faster than the time scale *in vivo* and of AFM experiments. In order to analyze how the unfolding forces and mechanism vary with the time scale of the unfolding, FPMD simulations were carried out with varying pulling velocities covering a range of nearly two orders of magnitude. Unfolding forces observed for different velocities are shown in Fig. 3.6a. As can be seen, these forces are nearly equal for the N and C-termini, respectively, which shows that in all cases the protein position between the springs is close to equilibrium. Furthermore, the expectation that both side pulling reduces friction was justified.

Eq. 2.13 describes a linear dependency of the rupture force on  $v$ , in Fig. 3.7



seen for  $v > 0.05 \text{ \AA/ps}$ , the friction-dominated regime, and a logarithmic dependency for  $v \leq 0.05 \text{ \AA/ps}$ , the thermally activated regime. Typically FPMD does not allow for a straightforward extrapolation to experimental time scales. Therefore, Heymann et al. have suggested to use the spontaneous unfolding rate  $k_0$  as an additional experimental input [127]. Unfortunately, this rate has not yet been measured for titin kinase, so that a prediction of rupture forces at the experimental time scale is not possible at present.

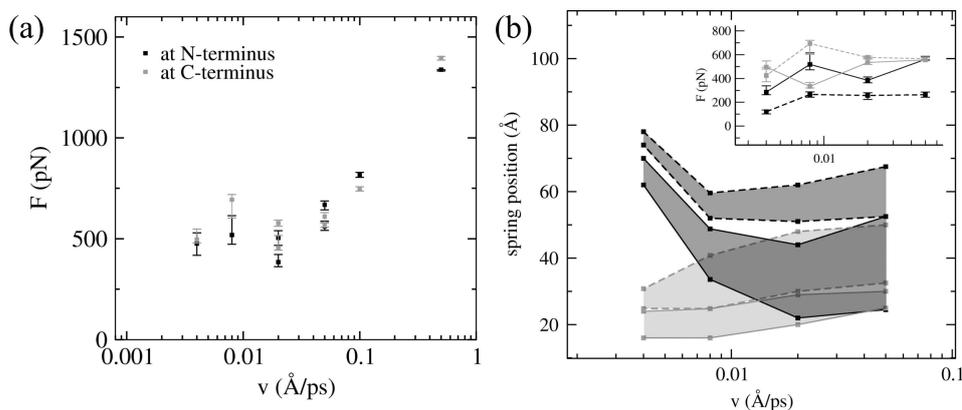


Figure 3.6: (a) Maximal rupture forces as a function of pulling velocity  $v$  for the N-terminus (black) and C-terminus (red). Error bars indicate the uncertainty in the maximal rupture force due to the choice of a Gaussian distribution width for smoothing (see Methods section). (b) Spring positions (dislocation with respect to the starting position) at which the first (solid) and final (dashed) rupture of two N-terminal (black) and two C-terminal (red)  $\beta$ -sheets occur, as a function of the pulling velocity for the four slowest pulling simulations. The inset shows the maximal rupture force observed during the period of the four  $\beta$ -sheet ruptures.

The rupture events observed in the FPMD simulations of the logarithmic regimes were compared with each other and found to be similar with respect to the main features, as described below. If not otherwise mentioned, the results of the  $0.004 \text{ \AA/ps}$  simulation are therefore presented.

Unfolding dynamics are mainly governed by the rupture of the terminal anti-parallel  $\beta$ -sheets, shown in yellow in Fig. 3.2, in particular by the rupture of hydrogen bonds between the strands  $\beta C1$ ,  $\beta C2$  and  $\beta C3$  at the N-terminus (I) and those between  $\beta C10$ ,  $\beta C11$  and  $\beta R1$  at the C-terminus (II) (Fig. 3.7). The largest force is required for the C-terminal  $\beta C10$ - $\beta R1$  rupture, which is preceded by the  $\beta C11$ - $\beta R1$  rupture at lower force. At the N-terminus rupture of the  $\beta C1$ - $\beta C2$   $\beta$ -sheet is followed by the  $\beta C2$ - $\beta C3$  rupture. Another main rupture event is the loss of secondary structure of the inhibitory tail

### 3. ENFORCED TITIN KINASE ACTIVATION

sequence  $\alpha$ R2 which coincides with its detachment from the ATP binding groove (III).

As the only major difference of the unfolding pathways between slower and faster pulling simulations, we observed the sequence of events I-III (Fig. 3.7). The FPMD simulations with  $v = 0.05 \text{ \AA}/\text{ps}$  and  $v = 0.004 \text{ \AA}/\text{ps}$  shall serve as the most pronounced examples. At a velocity of  $0.05 \text{ \AA}/\text{ps}$ , N- and C-terminal  $\beta$ -sheet ruptures occur simultaneously, after the spring positions have moved by approx.  $35 \text{ \AA}$  each. At the end of the simulation,  $\alpha$ R2 has lost its secondary structure only partially and did not substantially dislocate from the ATP binding site between small and big lobe. Instead,  $\alpha$ R2-small lobe interactions are partially disrupted due to the dislocation of  $\beta$ C1 and  $\beta$ C2 at the time of N-terminal  $\beta$ -sheet rupture.

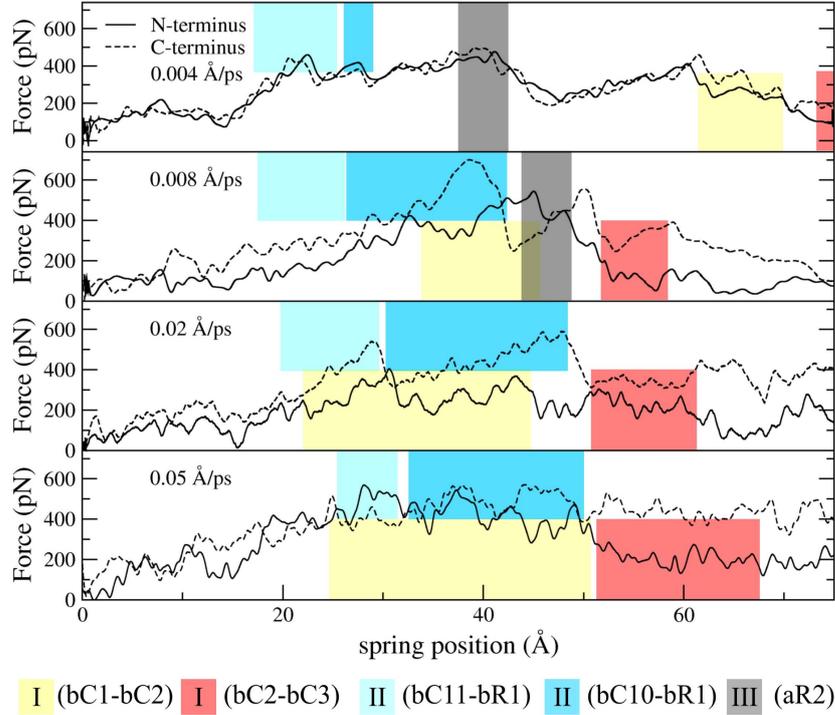


Figure 3.7: Force profiles and main rupture events for FPMD simulations of the low velocity regime. N-terminal ruptures (I) are shown in yellow and red, C-terminal ruptures (II) in light and dark cyan and the  $\alpha$ R2 rupture (III) in gray.

In contrast, during the FPMD simulation with  $v = 0.004 \text{ \AA}/\text{ps}$ , ruptures of the C-terminal  $\beta$ -sheet at a spring position of  $15 - 30 \text{ \AA}$  clearly precedes that of the N-terminal  $\beta$ -sheet ( $62 - 78 \text{ \AA}$ ). As a consequence of the higher force resistance of the  $\beta$ C1- $\beta$ C2  $\beta$ -sheet relative to the  $\beta$ C10- $\beta$ R1  $\beta$ -sheet,  $\alpha$ R2



leaves the active site prior to ruptures at the N-terminus. The force peak at a spring location of 42 Å can be attributed to the simultaneous rupture of interactions between  $\alpha$ R2 and the small lobe, of which the  $\beta$ -sheet still is taking up its original conformation and hydrogen bond pattern. With regard to the similarity of the N and C-terminal  $\beta$ -sheets — both are antiparallel and contain six intrastrand hydrogen bonds — the different mechanical strength is remarkable and will be inspected more closely below.

In general, as shown in Fig. 3.6, a decrease in pulling velocity leads to a stabilization of the N-terminal  $\beta$ -sheet with respect to the C-terminal  $\beta$ -sheet. However, a similar tendency is not observed for the forces required for these ruptures (Fig. 3.6, inset). Here, the relative strength of  $\beta$ -sheets are not reflected by the rupture force but rupture time.

Except for the terminal  $\beta$ -strands and the inhibitory tail, the protein secondary and tertiary structure remains remarkably stable. After a dislocation of 92 Å of each spring, the titin kinase keeps its native secondary structure fold for 50 % of its residues, or for 64 % of the residues not yet pulled out of the protein core.

### 3.3.3 Open-closure motion

For four FPMD simulations at low pulling velocity, principal components of the titin kinase dynamics have been calculated to characterize the main collective motions induced by the applied mechanical stress. The FPMD runs were used to calculate the covariance matrix of the protein backbone atoms, and after projection onto the largest eigenvectors, similar global domain motions were found for each of the four trajectories. In all cases, the eigenvector contributing mainly to the total protein motion (with a contribution of 76-86% to the total fluctuations) describes the movement of the terminal protein sequences apart from each other, and, in this sense, is the trivial one. The more relevant second largest eigenvector, with a contribution of 7-8%, describes an open-closure motion involving the active site (Fig. 3.8). During the 0.004 Å/ps simulation, e.g., the projection of the trajectory onto the second eigenvector (Fig. 3.8a) decreases until a spring position of 60 Å is reached. Structural inspection (Fig. 3.8b) shows that this drop describes an opening motion of the active site. The subsequent rupture of the N-terminal  $\beta$ -sheet then yields the main drop in force, leading to a closure motion of the small and big lobe. This relaxation of the strained protein conformation is reflected by a subsequent increase of the projection of the second eigenvector, and thereby restores the initial binding site geometry.

### 3. ENFORCED TITIN KINASE ACTIVATION

The open-closure motion closely resembles the one observed also in the conformational transition from inactive to active states in other protein kinases like insulin receptor tyrosine kinase [128].

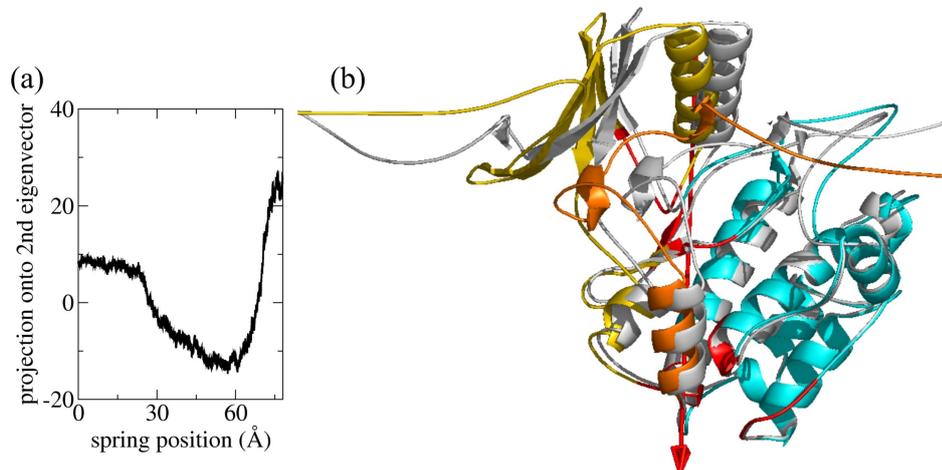


Figure 3.8: Open-closure motion of the small and big lobe during pulling. (a) Projection onto the second eigenvector as a function of simulation time. (b) Overlay of the extreme projections of the titin kinase motion on the second eigenvector (gray: closed conformation, colored: open conformation). The hinge axis between the two dynamic domains as determined with DynDom [125] (cyan: big lobe; yellow and orange: small lobe and regulatory tail) is depicted as red arrow; hinge residues are shown in red.

According to the domain analysis of the second eigenvector with DynDom [125], the whole regulatory tail including the  $\alpha R1$  and  $\alpha R2$  segments moves collectively together with the small lobe (Fig. 3.8b). Consequently,  $\alpha 2$  maintains its tight interactions to small lobe residues until rupture of those at a spring position of 44 Å, *i.e.*, briefly before the closure motion sets in. In contrast to  $\alpha R2$ , which at a spring position of 90 Å has moved to an RMSD of 40 Å,  $\alpha R1$  shows a deviation from its initial conformation by less than 5 Å. Thus, in spite of the concerted motion with  $\alpha R2$ , the total fluctuations of  $\alpha R1$  are significantly smaller. The two main hinge points for the open-closure motion are located at (i) the big lobe regulatory tail interconnection N-terminal of  $\alpha R1$  (residues 272-273) and neighboring sequences, and (ii) at residues 60-66 (small lobe) and residues 141-145 (big lobe), which is the only region at which the big and small lobe stay tightly aligned during force-induced unfolding (Fig. 3.8b).

Comparison of the PCA (see Section 2.4.3) of all FPMD simulations shows that both, the eigenvectors, which characterize the nature of the main col-



lective motions, as well as the associated eigenvalues, which measure the amplitudes of these motions, are largely unaffected by the choice of the pulling velocity. The most pronounced change is seen for the first eigenvector, which contributes between 86% ( $v = 0.05 \text{ \AA/ps}$ ) and 76% ( $v = 0.004 \text{ \AA/ps}$ ) to the overall dynamics. This finding is in line with the observation that also the rupture forces drop only slightly below  $0.005 \text{ \AA/ps}$  (Fig. 3.6a) as expected for a low unfolding rate. Taken together, our results indicate that the FPMD simulations presented here, even though restricted to nanoseconds, capture at least a relevant fraction of the conformational dynamics that are expected at the physiological and experimental time scales.

### 3.3.4 $\beta$ -sheet ruptures

#### Molecular dynamics simulation results

To elucidate the molecular origin of the pronounced and unexpected asymmetry in the strength of the two terminal  $\beta$ -sheets, we compared their hydrogen bond rupture pattern. Figure 3.9 shows the hydrogen potential energy as a function of time for the ruptures at the N- and C-terminus.

As can be seen, hydrogen bonds between the strands  $\beta C1$  and  $\beta C2$  rupture concurrently within a very short time interval, in response to a shearing force acting parallel to the  $\beta$ -strands in opposite directions (*cf.* the sketch and snapshots in Fig. 3.10). After rupture, the strands slide along each other until they snap in again by forming a new set of hydrogen bonds (see Fig. 3.9). In contrast, the  $\beta$ -strands at the C-terminus, oriented approximately vertically with respect to the pulling direction, rupture in a step-wise manner, resulting in a zipper-like opening of the  $\beta$ -sheet. Here, no transient non-native hydrogen bonds are formed (Fig. 3.10). As already observed for the overall conformational motions, also this pronounced contrast of concurrent versus sequential rupture of the  $\beta$ -sheets is seen in all FPMD simulations, independent of the pulling velocity.

Assuming that the two  $\beta$ -sheet ruptures occur independently from each other, the order of ruptures can be used to characterize the relative mechanical stability of the respective  $\beta$ -sheets — the weaker  $\beta$ -sheet should rupture first. Indeed, the  $\beta$ -sheet that is oriented parallel to the pulling direction ruptures at higher loads, and thus is more force resistant, than the  $\beta$ -sheet oriented vertical to the pulling direction (Fig. 3.6). This is not unexpected, since the concurrent rupture of N hydrogen bonds requires a higher force than a single bond rupture within a sequential rupture, as it

### 3. ENFORCED TITIN KINASE ACTIVATION

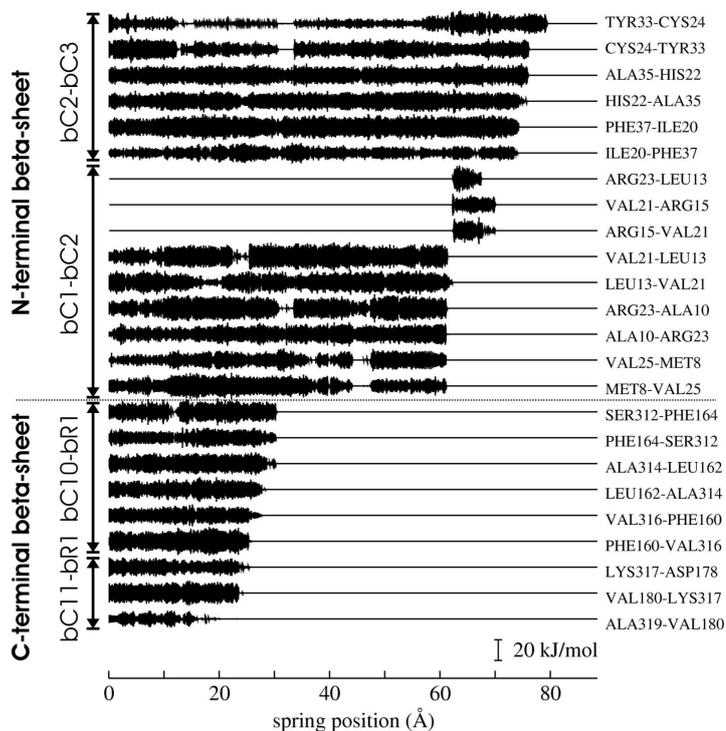


Figure 3.9: Hydrogen bond interactions at the terminal  $\beta$ -sheets. Shown are hydrogen bond strengths for each hydrogen bond between the  $\beta$ -strands as indicated. As can be seen,  $\beta C1$ - $\beta C2$  ruptures concurrently, the other sheets sequentially. For concurrent rupture, subsequent transient hydrogen bonds are formed.

has been shown previously [40, 108, 129, 130]. Interestingly, the difference in stability vanishes for increasing pulling velocity (Fig. 3.6).

The question arises, if the different rupture behavior at the termini is indeed sufficiently and quantitatively explained by the different orientation of the  $\beta$ -sheets, or if other factors also contribute. How would one expect them to scale towards experimental pulling velocities and does this explanation also hold for the physiological time scales? To answer these questions and to understand the underlying physics, we have performed FPMC simulations. Here, the system is not simulated in atomic detail, but instead only the different energy landscapes underlying the concurrent and sequential rupture mechanisms are included.

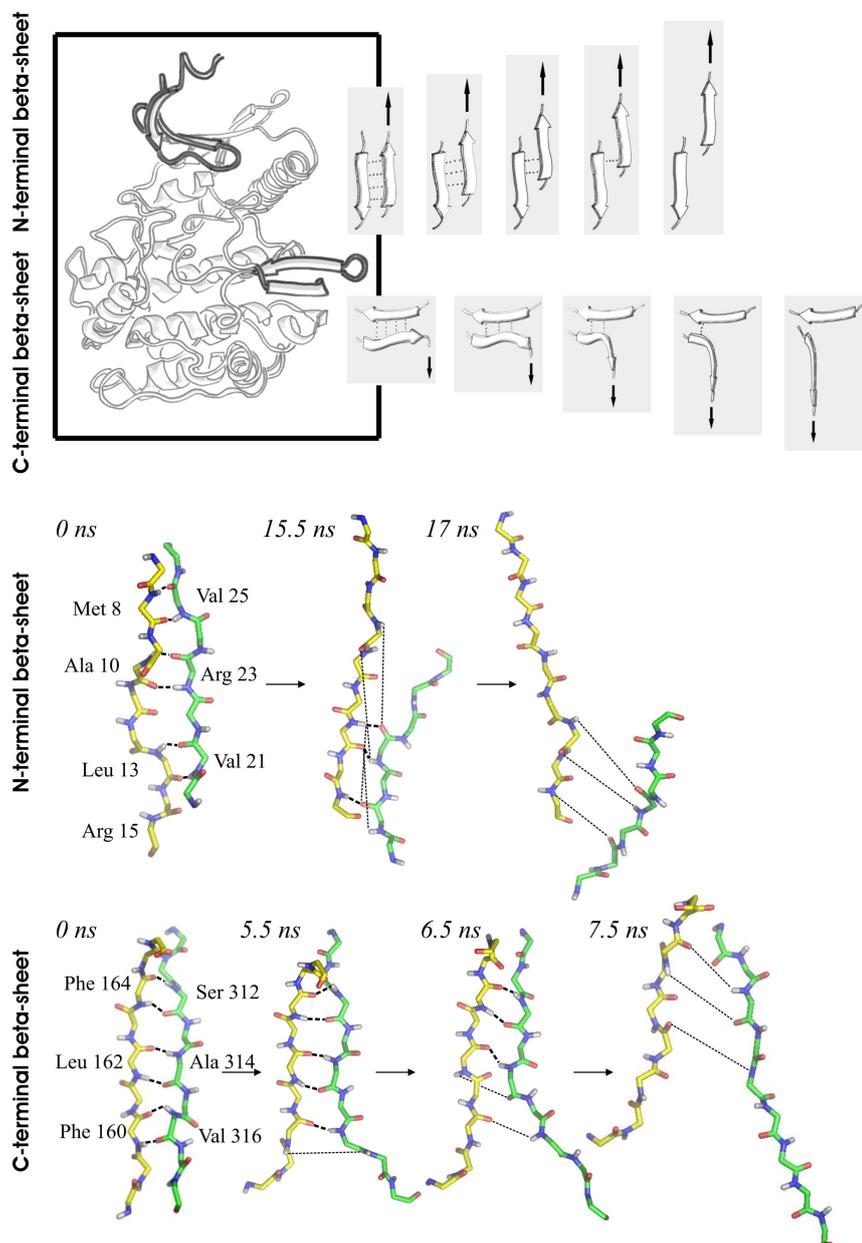


Figure 3.10: Differences of hydrogen bond rupture patterns for the two terminal  $\beta$ -sheets. Upper panel: Sketch of the sequence of ruptures at the N-terminus (top) and the C-terminus (bottom). Lower panel: snapshots of the N-terminal  $\beta$ C1 and  $\beta$ C2 (yellow and green, respectively, top); snapshots of the C-terminal  $\beta$ C10 and  $\beta$ R1 (yellow and green, respectively, bottom). Hydrogen bonds are drawn as dashed (intact) and dotted (ruptures briefly before) dashed lines.

### Monte Carlo simulation results

As a simple model we describe sequential and concurrent rupture by the one-dimensional energy profiles shown in Fig. 3.11b. The average potential energy difference between an intact and broken single hydrogen bond was set to  $-23$  kJ/mol, the average value obtained for the relevant bonds from the MD simulations according to Eq. 3.2. For sequential rupture (Fig. 3.11a top), a step-wise increasing hydrogen bond potential was chosen (Fig. 3.11b). A single rupture event involves the rupture of one hydrogen bond, and, therefore, leads to an increase in potential energy along the reaction coordinate  $x_2 - x_1$  by  $-23$  kJ/mol with a low activation barrier. Each rupture event of the concurrent rupture (Fig. 3.11a bottom), in contrast, involves a barrier with a height corresponding to the number  $n$  of hydrogen bonds presently formed, as all hydrogen bonds rupture at each sliding event along  $x_1$ . The subsequent drop accounts for the formation of  $n - 1$  transient hydrogen bonds. For both types of rupture, the change along the reaction coordinate was set to 1 nm, the average increase of this distance observed in the MD simulations during a complete rupture. Comprisin these two processes, the system moves within a 2-dimensional energy landscape (Fig. 3.11c).

In the FPMD simulations,  $\beta$ -sheet rupture is followed by a fast unfolding of the released peptide chain. To account for its elasticity in the MC simulations, a wlc element was attached to each  $\beta$ -sheet. This element was modeled by allowing steps along each reaction coordinate after the respective hydrogen bond potential has been traversed ( $x_1 > 1$  nm and  $x_2 - x_1 > 1$  nm respectively) on a wlc energy surface (Eq. 2.10) with a contour length of 2.5 nm.

Figure 3.11d shows the spring positions at which the first and last hydrogen ruptures of the  $\beta$ -sheets occur for varying spring velocities, which is to be compared to the MD results shown in Fig. 3.6b. As expected, for low pulling velocities, the parallel occurs after the sequential rupture, with the two ruptures well separated from each other, in correspondence to the FPMD results. The trajectory of an FPMD run in this regime leads along the border of the energy profile of Fig. 3.11c (first along  $x_2 - x_1$  with  $x_1 = 0$ , then along  $x_1$  with  $x_2 - x_1 = 0$ ).

The increase of the spring positions with pulling velocity that was seen in the FPMD simulations is also observed here. In particular, for large pulling velocities the rupture processes take longer, in qualitative agreement with the MD simulations (Fig. 3.6b). Below  $v = 50 \times 10^{-5}$  nm/step, the two different positions at which the two  $\beta$ -sheets rupture do not change considerably,

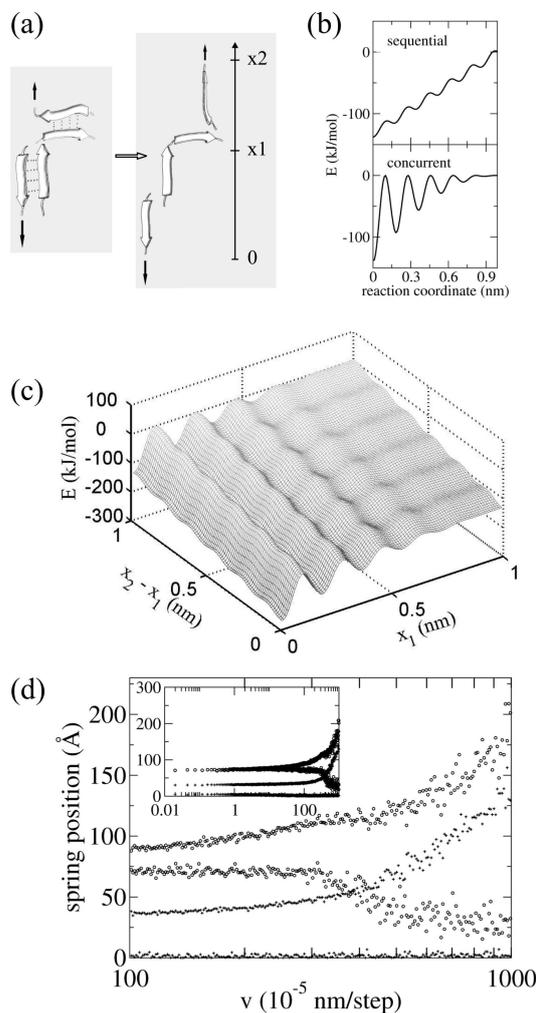


Figure 3.11: FPMC simulations with model energy landscapes. (a) Sketch of the model system: two attached  $\beta$ -sheets rupturing concurrently (top) and sequentially (bottom). (b) Assumed one-dimensional energy profiles for the sequential and concurrent rupture of a  $\beta$ -sheet with six hydrogen bonds. (c) Combined hydrogen bond potential energy surface  $G(x_1, x_2 - x_1)$ . Here,  $x_1$  is the reaction coordinate of the sequential,  $x_2 - x_1$  that of the concurrent rupture, respectively. combining the profiles of (b). (d) Spring positions at which the first (lower curve) and final rupture (upper curve) of the  $\beta$ -sheets occur as a function of the pulling velocity for the concurrent (circle) and the sequential (plus) rupture. The inset shows data for a larger velocity range.

which suggests that this sequence of events, as seen in the MD simulations, is likely to be observed also at physiological and experimental time scales. For large velocities, the spring positions of the final C-terminal rupture and

the starting N-terminal rupture approach each other and finally overlap. The simultaneous rupture of both  $\beta$ -sheets corresponds to a trajectory that diagonally traverses the energy surface of Fig. 3.11c. The cross-over takes place when the velocity, at which the spring is dislocated each step, exceeds the average step size, *i.e.*, when Stokes friction sets in, which in the MC simulations is implicitly described via the diffusion coefficient that is set by the ratio between the squared average positional step size and the time step. Above this point, unfolding can not keep up with the fast movement of the spring. In agreement with the MD simulation results (3.6), the cross-over occurs within one order of magnitude of pulling velocities.

### Individual hydrogen bond ruptures

A striking difference between the dynamics of concurrent and sequential rupture was found in the individual rupture events as quantified by the steepness of the hydrogen bond potential energy increase with time. In the FPMD simulations, the dislocation of the spring during which a rupture event takes place was two orders of magnitudes larger for the sequential as compared to the concurrent rupture. In other words, hydrogen bonds rupture within a significantly longer time interval in a sequential than in a concurrent fashion.

Is this differential behavior as found for the simple two-dimensional model also in agreement with the MD simulations? We studied the rupture dynamics of the  $\beta$ -sheet hydrogen bonds during the FPMD simulations by examining the potential energy of individual hydrogen bonds. The interval during which a hydrogen bond potential drops from the value of 23 kJ/mol in average to 0 kJ/mol, was assessed by fitting a properly chosen function to the hydrogen bond energy trace (see Methods section and Fig. 3.12). As can be seen, hydrogen bonds that rupture concurrently indeed show a smaller width  $a_0$  of the potential decrease than those that rupture sequentially.

A more sudden rupture of a sheet under longitudinal force indeed is expected, since a weakening of the hydrogen bonds between the strands leads to the complete loss of interactions and a large sliding movement until the strands snap in again. A pair of hydrogen bond partners breaking apart during a sequential rupture, instead, stay in proximity until the adjacent hydrogen bond breaks in the next step. This results in a comparably slow hydrogen bond potential decay. The width  $a_0$  decreases with decreasing pulling velocity for the same reason that causes the decrease of maximal rupture force (Fig. 3.6); with a lower loading rate a rupture becomes likely

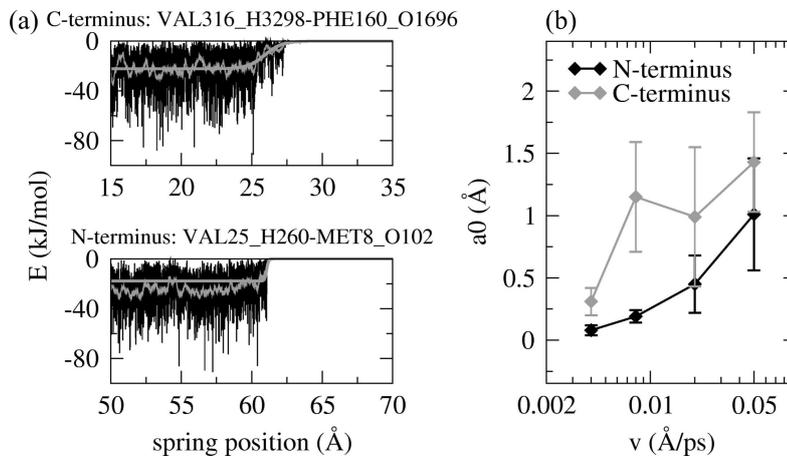


Figure 3.12: Difference between fast and slow hydrogen bond ruptures. (a) Hydrogen bond potentials (black), running average over 50 ps and fit to the running average with a logistic function (gray, see Section 3.2). Shown are two representative hydrogen bond ruptures at the N and C-terminus, respectively. (b) Width  $a_0$  of the fitting function as a measure for the abruptness of the rupture process. Values are averaged over the hydrogen bonds of the sheet, error bars give the variance among them.

already at lower loads due to thermal activation. In summary, the difference in the abruptness of the hydrogen bond potential decay between concurrent and sequential rupture as suggested by the MC simulations was also found in the corresponding FPMD simulations, thus providing further support that our simple two-dimensional model captures the underlying physics of the rupture process.

### 3.3.5 Release of auto-inhibition

We now focus at the active site to investigate whether the unfolding indeed leads to a relieve of auto-inhibition while leaving the active site intact. If not noted otherwise, we subsequently report the results for the simulations with a pulling velocity of  $0.004 \text{ \AA}/\text{ps}$ . Similar results have been obtained for the other simulations in the low velocity regime.

The titin kinase regulatory tail (residues 275-321), particularly the  $\alpha\text{R2}$  helix (residues 296-301), tightly binds to the active site via multiple hydrogen bonds and van der Waals contacts. Fig. 3.13 shows the course of the van der Waals interaction energies between residues of the  $\alpha\text{R2}$  and the small lobe. As can be seen, the van der Waals contacts rupture simultaneously at

### 3. ENFORCED TITIN KINASE ACTIVATION

a spring position of 44 Å. These ruptures are the main trigger for  $\alpha$ R2 release from the active site groove. Together with the rupture of the intra-helical hydrogen bonds of  $\alpha$ R2, this event gives rise to a force peak. Interactions of  $\alpha$ R2 to the active site also include those to side chains of residues known to be involved in ATP binding (Lys36, Glu51, Asp87, Glu147, data not shown). In all simulations the inhibition of residues Glu51 and Glu147 is removed prior to the rupture of van der Waals contacts of to the small lobe; the Glu51-Ala304 hydrogen bond ruptures at a spring position of 36 Å, the Glu147-Gly303 hydrogen bond ruptures at 30 Å, respectively. An early rupture of the initial Lys36-Gly302 hydrogen bond at 10 Å is followed by the formation of a Lys36-Ile299 hydrogen bond, which ruptures at 45 Å, *i.e.*, immediately after the main rupture of van der Waals contacts to the small lobe. Asp87 forms a transient hydrogen bond to regulatory tail Ser295, which finally breaks at a spring position of 42 Å.

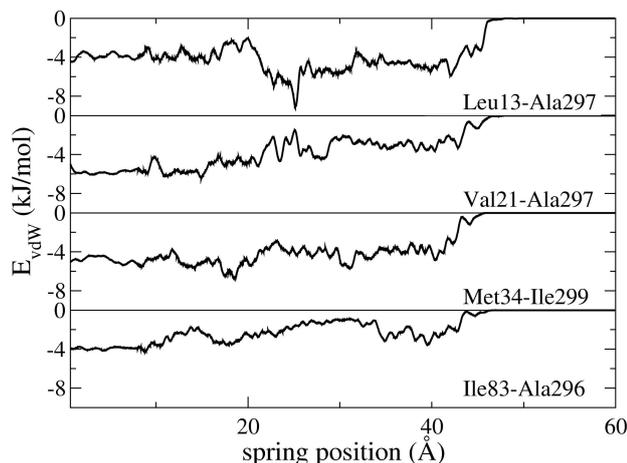


Figure 3.13: One-step rupture of the inhibitory tail from the small lobe. The van der Waals interaction energy for residue pairs of the small lobe (1-83) and  $\alpha$ R2 (296-301) smoothed by Gaussian filtering with a band width of 100 ps is shown.

In summary, together with the detachment of the hydrophobic contacts between the regulatory tail and the small lobe, autoinhibition of the ATP binding residues is removed. In spite of the variability of side chain conformations, all low velocity FPMD simulations share the relevant specific residue-residue interactions, although the detailed sequence of ruptures may differ, due to variations of the spring position at rupture by  $\pm 20$  Å.

Is the structure of the active site after removal of the auto-inhibitory tail sufficiently intact for the proposed force-triggered kinase activity? To answer this question, the geometry of the active site was characterized by monitoring



the distances between  $C_\alpha$  atoms of the catalytic and ATP binding residues during the FPMD simulation (Fig. 3.14a). As can be seen, all distances stay close to the distances seen in the titin kinase crystal structure [52] and in the crystal structures of its homologues twitchin kinase [111] and extracellular regulated kinase (ERK-2) [131]. The only exception is the distance between Lys36 and Asp127, which increases by about 2 Å. Further deviations from the crystal structure are transient, *e.g.*, for Lys36-Asp127 at 75 Å, and disappear during the subsequent conformational relaxation that takes place after the stress is reduced. We note that these transient deviations are smaller for slow pulling velocities, which suggests that they may be irrelevant or even absent at physiological time scales. These results suggest that the active site integrity, though temporarily perturbed, is maintained to a sufficient extent to allow kinase activity. Since Tyr170 remains with its 4' hydroxyl group oriented towards the catalytic base Asp127, the semi-opened catalytic cleft raises the possibility of an intramolecular autophosphorylation upon ATP binding.

The Gly-rich loop (residues 14-19), located in the small lobe between the  $\beta$ -strands  $\beta C1$  and  $\beta C2$ , plays a role in accommodating ATP in the active site groove [132, 133]. This loop remains stable upon mechanical relieve of autoinhibition, but does not remain close to the active site, when the small lobe unravels. As shown in Fig. 3.14b, it is rather pulled apart as soon as the  $\beta C1$ -  $\beta C2$  sheet ruptures at a spring position of about 60 Å. As a consequence of the N-terminal  $\beta$ -sheet destabilization with increasing pulling velocity (Fig. 3.6), the Gly-rich loop is removed from its active site location prior to the release of auto-inhibition for high pulling velocities, while for small velocities such as 0.004 Å/ps, release occurs from an intact N-terminal  $\beta$ -sheet including the Gly-rich loop.

### 3.3.6 Comparison to twitchin kinase unfolding

The titin kinase domain contains an N-terminal coil that was not included in the FPMD simulations due to the lack of structural information on this sequence. To test if the absence of this coil entails artifacts in the observed unfolding mechanism, a 6 ns FPMD simulation of twitchin kinase including the N-terminal coil was carried out. During the FPMD simulation, twitchin kinase shows little conformational changes in the protein core, and instead separates from the N-terminal random coil (residues 1-24), giving rise to a force maximum of  $430 \pm 21$  pN at 2.6 ns (Fig. 3.15). This force maximum can be attributed to ruptures of mainly four polar interactions, namely

### 3. ENFORCED TITIN KINASE ACTIVATION

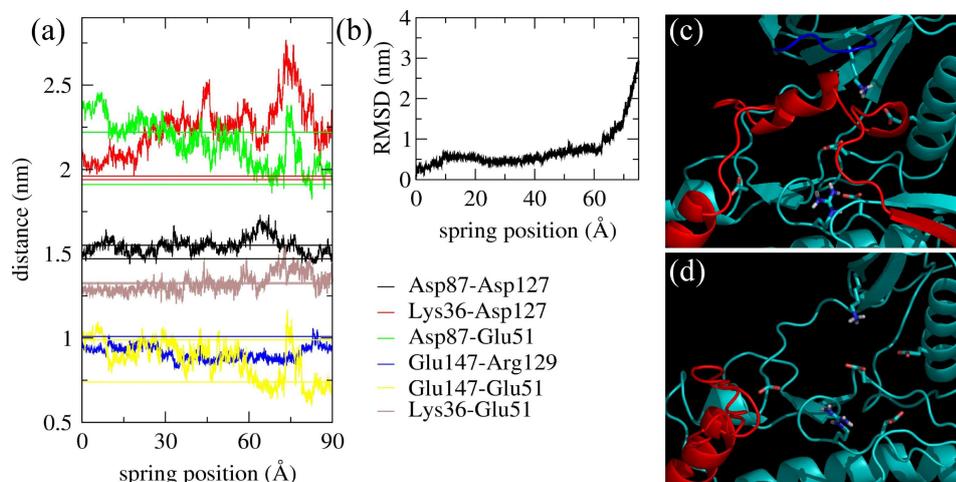


Figure 3.14: Integrity of the active site. (a)  $C_{\alpha}$ -distances of ATP binding residues. Straight lines are the equivalent distances in twitchin kinase [111] and extracellular regulated kinase (ERK-2) [131] (b) Displacement of the Gly-rich loop shown as root-mean square deviation of residues 14-19 from the starting structure. (c) and (d) Snapshots of the active site before (0 ns) and after (22 ns) the release of auto-inhibition. ATP binding and catalytic residues are shown as sticks, the regulatory tail in red, the Gly-rich loop in blue.

Asp5-Ser109, Asp5-Lys163, Gln18-Leu91, Gln18-His92. All of the involved residues are strictly conserved in titin kinase, except Gln18, which in the titin kinase is a functionally equivalent lysine (Fig. 3.4). Therefore, a similar force pattern can be expected for the respective unfolding of a titin kinase domain including the additional sequence.

To evaluate the error in the protein conformational changes due to the lack of the N-terminal coil in the titin kinase structure, the dislocation of the terminal  $\beta$ -strands during pulling was compared to the one observed during twitchin kinase pulling (Fig. 3.15). In both kinases, dislocation of the C-terminal  $\beta$ R1 strand occurs sooner and to a larger extent than dislocation of the N-terminus. Thus the conformational changes of the  $\beta$ -sheets are not significantly affected by the presence of the additional N-terminal sequence.

The initial pulling direction, defined as the vector connecting the twitchin kinase termini, differs from that chosen for titin kinase unfolding by  $100^{\circ}$ . With regard to the anisotropy of the mechanical stability of  $\beta$ -sheets the question arises whether the fact that the N-terminal coil is missing in the titin kinase structure affects the FPMD results discussed here. Comparison with the twitchin kinase simulation allows to address this uncertainty. As can be seen in Fig. 3.16, detachment of the N-terminal coil of twitchin ki-

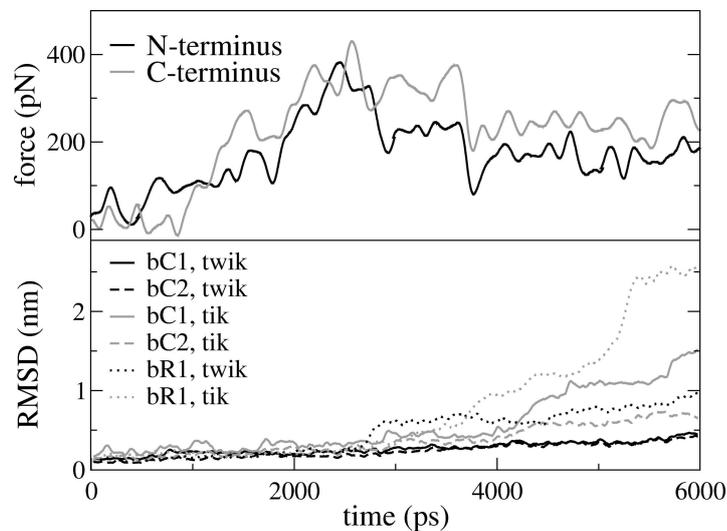


Figure 3.15: Upper panel: Force profile as obtained from the twitchin kinase pulling simulation with  $0.008 \text{ \AA}/\text{ps}$  velocity after Gaussian filtering. Lower panel: Comparison of the terminal  $\beta$ -sheet dislocation between twitchin kinase and titin kinase during the first 6 ns of the  $0.008 \text{ \AA}/\text{ps}$  pulling simulation.

nase induces an overall rotation of the whole protein, such that the pulling direction becomes similar to that of titin kinase. Therefore, we consider artifacts from the initial difference in the pulling directions unlikely.

### 3.4 Discussion and Conclusion

We have carried out force probe simulations of titin kinase in order to test the hypothesis that this titin domain is a force sensor that can be activated by mechanical stress. Our results suggest that mechanical strain on this domain leads to a sequence of conformational changes that result in an open conformation of the catalytic cleft leaving the catalytic core intact. Titin kinase is the only kinase among  $\text{Ca}^{2+}$ /Calmodulin activated kinases, of which the catalytic site Lys is substituted by Arg [52]. The heterogeneous hydrogen bond network of the titin kinase catalytic site as observed during the free MD simulation can be partially ascribed to the larger size of Arg comparing to Lys, and its higher potency to form multiple salt bridges. Furthermore, in contrast to titin kinase, Glu156 of twitchin kinase forms a salt bridge to Lys312, a residue of the inhibitory tail, and thereby is directed away from Lys154. These findings may point at a role of the conserved Glu in influencing the catalytic base activity of Asp127 in titin kinase, in addition to

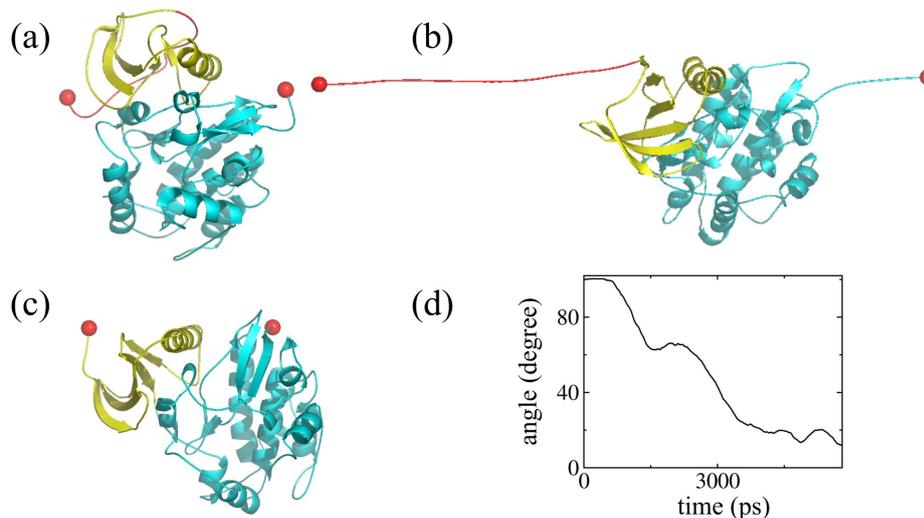


Figure 3.16: Comparison of the pulling geometry between twitchin kinase and titin kinase. (a) and (b) Snapshots of the twitchin kinase pulling simulation at 0 ns and 6 ns. (c) Titin kinase starting structure. In (a)-(c) the big lobe is shown in cyan, the small lobe in yellow, and the pulled  $C_{\alpha}$ -atoms as spheres in red. (d) Angle between the force vector during the twitchin kinase pulling simulation and the initial force vector for titin kinase pulling simulation.

its participation in ATP binding, as generally assumed for serine/threonine kinases [110].

Performing FPMD simulations of protein unfolding requires a box length longer than the length of the unfolded protein. The use of a system size able to accommodate the unfolded protein throughout the FPMD simulation, however, is due to the large amount of solvent molecules computationally impractical. Elongating solvent spheres or repeated box enlargement during the simulation have been applied to tackle this problem [108, 134]. Here, cutting terminal protein segments and accounting for their elasticity by adjusting the spring constant proved to be an easy and efficient method to reduce the box length in pulling direction.

Information on the mechanical function of a protein can be obtained from its response to force. A number of proteins, those that *in vivo* are subjected to force (titin Ig/Fn3 domains) as well as those that are not, have been investigated by pulling experiments (AFM and optical tweezers) and MD. In AFM experiments,  $\beta$ -sandwich proteins such as titin Ig and Fn3 have been found to unfold at significantly higher forces (130-300 pN [11, 135, 136] for I-band Ig, 180 pN in average for central A-band Fn3 [135]) than the non-



mechanical protein barnase (50-100 pN [134]) and the cytoskeleton component spectrin [25-35 pN [137]]. By comparing those findings with our FPMD results, the force resistance of titin kinase can be classified. The maximal rupture force of 670 pN as found for titin kinase unfolding at 0.05 Å/ps (corresponding to 0.1 Å/ps if one terminus is fixed and the other one pulled) lies below the unfolding forces of Ig27 domains (1350 pN [108]) and Ig1 domains (1000 pN [138]) found for pulling one terminus with 0.1 Å/ps (Fig. 3.6). Fn3 domains have been shown to unfold at forces of 1000 – 1300 pN at pulling velocities of 0.01 Å/ps [139], and therefore are another example of titin domains that are more force resistant than titin kinase (500 pN at 0.004 Å/ps). Barnase, instead, was found to unfold at lower forces (270 pN at 0.01 Å/ps [134]) than titin kinase (610 pN at 0.008 Å/ps). Thus, titin kinase takes up a medium position, ranking below the mechanically most stable  $\beta$ -sandwich proteins Ig and Fn3 and above those containing  $\alpha$  helical secondary structure. Thus, for the physiological and experimental time scales, a force in the range between 50 and 150 pN is expected from our simulations. This force is exactly what is to be expected for a force sensor reporting large mechanical stress that may unfold Ig domains of titin.

Our FPMD results show that the largest unfolding forces are required to rupture inter-strand hydrogen bonds. Therefore, the terminal  $\beta$ -sheets can be considered as the mechanical elements mainly responsible for the remarkable force resistance. With a typical kinase fold, titin kinase is at first sight not expected to have specifically evolved for the sake of force resistance. From this point of view, the force resistance and the significant catalytic site stability are unexpected.

For slow pulling velocities ( $\leq 0.05$  Å/ps), the N-terminal  $\beta$ C1- $\beta$ C2 sheet is significantly more stable than the C-terminal  $\beta$ R1- $\beta$ C10 sheet. As discussed below, the order of their rupture, *i.e.*, their relative force resistance, is the main feature of the force-induced activation mechanism. The precedence of the C-terminal sequential over the N-terminal concurrent rupture might be attributed to the orientation of the  $\beta$ -sheets. To get quantitative insight into the physics of this mechanism and the resulting differential stabilities, we have considered a two-dimensional model of the rupture of both terminal  $\beta$ -sheets, for which we have carried out FPMC simulations. The FPMC simulations yield a picture in quantitative agreement with the  $\beta$ -sheet rupture events observed during the FPMD simulations: At pulling velocities well below the friction dominated regime, the rupture of the  $\beta$ -sheet oriented vertically to the pulling direction is clearly favored over the one oriented in parallel. At high velocities, however, the force resistances approach each

other, resulting in a rupture of the two  $\beta$ -sheets at similar spring positions. The good agreement of the FPMC results with those observed in the FPMD simulations of the full titin kinase indicates that our simplistic two-dimensional model, although it neglects other protein domains and solvation, captures the main physics underlying the differential force resistance of differently oriented  $\beta$ -sheets.

In particular, the differential force resistance of the terminal  $\beta$ -sheets is attributed to their different orientation to force and consequential unfolding mechanism. Whereas the N-terminal sheet can rotate such that its strands are positioned in parallel to the pulling direction,  $\beta$ C1 of the C-terminal  $\beta$ -sheet lies fixed upon the kinase domain and prescribes a vertical orientation towards the pulling direction. The parallel (in contrast to the vertical) orientation stabilizes the N-terminal over the C-terminal  $\beta$ -sheet. It is the depleting effect of the large spring potential onto the potential energy surface that diminishes the differential resistance of the two  $\beta$ -sheets at high pulling velocities.

We conclude that the catalytic site of the titin kinase is stabilized by an N-terminal  $\beta$ -sheet that is aligned in parallel to force, whereas the C-terminal  $\beta$ -sheet, being vertically aligned, functions as a gate that readily opens to allow the dissociation of  $\alpha$ R2 from the active site. Hence, by means of pulling geometry, the active site is protected against disintegration on the N-terminal side, and the release of auto-inhibition is facilitated on the C-terminal side. The FPMC simulations suggest that at experimental and physiological time scales, *i.e.*, for pulling velocities four to five orders of magnitude below the cross-over regime, the difference in  $\beta$ -sheet stabilities is still present. Thus, the force-induced activation mechanism as described here should hold also at physiologically relevant slow time scales.

Studies on ubiquitin [140] and a lipoyl domain of pyruvate dehydrogenase, E2lip3 [130], have recently uncovered similar examples showing that beyond the number and nature of intra-strand hydrogen bonds, their relative orientation towards the pulling direction defines the stability of a  $\beta$ -sheet. The authors conclude that the same protein can exhibit different levels of force resistance depending on the pulling direction. Here, a different orientation of  $\beta$ -sheets of otherwise similar topology leads to a difference in mechanical domain stabilities within the same protein. The differential mechanical stability is required for the mechanical function. This finding demonstrates the importance of the dependency of the force resistance on the pathway when elucidating the unfolding mechanism.



In addition to the auto-inhibition of the active site by the regulatory tail in the crystal structure, the catalytic base Asp127 is blocked by Tyr170, located at the P+1 loop, via a hydrogen bond to its carboxylate group. It is assumed that phosphorylation of Tyr170 removes this blockage and thus is required for titin kinase activation [52]. Whereas in FPMD simulations of high pulling velocity transient rupture of this interaction is observed, it is fully maintained during the 0.004 Å/ps simulation. Hence, the FPMD results do not suggest a force-induced release of the Asp127 blockage.

Release of auto-inhibition by  $\alpha$ R2 dislocation leads to a structure of a principally intact kinase active site. Even at the final extension of 220 Å, the catalytic site integrity is striking, except that the Gly-rich loop, positioned above the ATP binding pocket and at the N-terminal  $\beta$ -sheet, is largely dislocated. However, a continuation of the trend towards N-terminal  $\beta$ -sheet stabilization with further lowering of the pulling velocity can be assumed. We speculate that at low velocities (which are computationally not yet within reach), the auto-regulatory tail vacates the active site prior to the displacement of the Gly-rich loop together with the N-terminal  $\beta$ -strand. Binding of  $\text{Mg}^{2+}$ /ATP to the exposed substrate binding cleft is likely to provide further stabilizing interactions for the ATP-binding lobe. Even after breakage of both terminal  $\beta$ -sheets and removal of the autoinhibitory tail, the secondary and tertiary structure of the kinase core remain remarkably intact. This is exactly what must be expected in light of the idea that the titin kinase is a force sensor activated by mechanical stress.

Remarkably, other domains behave quite differently. The Ig domains, *e.g.*, spontaneously disintegrate after rupture of  $\beta$ -sheets that are quite similar to those of the titin kinase [40, 108]. Thus, the stability of the kinase core and, in particular, the active site, is quite an unusual feature, which adds considerable weight to our results in favor of the force sensor hypothesis. The conformation of the regulatory domain under mechanical activation, with  $\alpha$ R2 and  $\beta$ R1 unfolded and an open ATP-binding site, is exactly the conformation found recently to bind tightly to the titin kinase associated signalsome [119, 141]. This complex is involved in mechanically modulated regulation of protein turnover and gene transcription in muscle. We therefore propose that this postulated semi-open conformation, as depicted in Fig. 3.3, can indeed be induced mechanically and thus can provide direct feedback of mechanical stress to the kinase signaling pathway. Thus, the results presented here strongly support the assumption of a force-induced kinase activation [42].

The half-open, presumably active titin kinase conformation as suggested by

### 3. ENFORCED TITIN KINASE ACTIVATION

---

simulations remains to be tested experimentally. Single molecule experiments of titin kinase are currently under way to measure rupture forces and to eventually detect intermediates [142], which then can be compared to those found in our simulations.

## 4

# Tuning titin kinase mechanical resistance by point mutations

---

### 4.1 Why mutants?

Mechanical stability of a protein is brought about by its overall structure. However, as we have seen for titin kinase, resistance towards force can often be ascribed to specific modules of the protein fold. Depending on the direction and magnitude of mechanical load, unfolding involves a sequence of ruptures, often involving considerable energy barrier(s). Unfolding intermediates withstand high mechanical load prior to the passage of these major barriers. This is reflected by maxima of the force profile from single-molecule experiments or FPMD simulations. In addition, FPMD simulations also directly yield the rupture event giving rise to the maxima, at atomic detail. They thereby allow to identify those structural motifs of the protein which mainly affect the unfolding resistance upon mechanical load.

The identification of mechanically stable units by force probe simulations can be validated by additional simulations of force-induced unfolding of protein mutants. The change in rupture force upon introduction of a point-mutation to abate crucial protein interactions gives further insight into the energetic barriers and intermediates along the pulling direction. This has been demonstrated previously for antigen-antibody complexes, immunoglobulin and fibronectin domains [13, 16, 28, 44, 143].

The force-induced unfolding of titin kinase as observed in FPMD simulations was discussed in Chapter 3. The mechanical stability of titin kinase was

found to be comparable to other domains of titin, namely immunoglobulin (Ig) and fibronectin (Fn) domains. By single-molecule experiments, mutational studies and simulations, the remarkable force resistance of Ig and Fn to fulfill their mechanical function was traced back to a certain topology of their  $\beta$ -sheet fold [13,144]. How does titin kinase, which exhibits the typical  $\alpha$ -helix rich fold of a kinase, achieve the observed mechanical stability? How is enzymatic function, *i. e.*, the active site, combined with mechanical function? Our simulations sought out the N- and C-terminal  $\beta$ -sheets and the central  $\alpha$ -helix of the auto-inhibitory tail as the major modules giving rise to energy barriers on the unfolding pathway. Here, on the basis of the picture of enforced titin kinase unfolding, mutants are suggested and subjected to MD and FPMD simulations. The observed free and force-induced unfolding dynamics of the mutants were compared to the wild-type (wt) simulations to validate and refine the suggested unfolding mechanism.

## 4.2 Choice of mutants and simulation details

Our studies on the wt titin kinase revealed two structural motifs mainly responsible for the remarkable mechanical stability of the enzyme: First, the initial major rupture of the C-terminal  $\beta$ -sheet, involving  $\beta$ C10, adjacent  $\beta$ C11, and  $\beta$ R1 (see Fig. 3.2). Second, the rupture of the  $\alpha$ R2-helix of the regulatory tail, from the active site, to which it is tightly bound *via* hydrophobic interactions, requires significant mechanical load (Fig. 3.7 and Fig. 3.13). To reduce the hydrophobic interaction between the active site and the tightly bound  $\alpha$ R2 helix of the regulatory tail, Ile299 and Ala297 were mutated to glycine, yielding the two mutants I299G or A297G, respectively. A third mutant with altered topology of the C-terminal  $\beta$ -sheet was obtained by mutating Val316 to a proline residue (V316P), which is known to destabilize  $\beta$ -sheets. The location of the mutated residues, Ala297, Ile299, and Val316, in the titin kinase fold is shown in Fig. 4.1.

Mutations of Ile299, Ala297 to glycine, and Val316 to proline were carried out starting from a wt structure taken from the equilibration as described in Section 3.2.1. The I299G and A297G mutations were obtained by deleting the spare side chain atoms. For the V316P mutant, proline was modeled into  $\beta$ -strand  $\beta$ R1 with Sybyl [145]. The box size, and water and ions coordinates of the systems were directly adopted from the wt snapshot. After minimization, an MD simulation of 2 ns length with harmonic restraints on the protein heavy atoms was performed. After an 8 ns equilibration, the

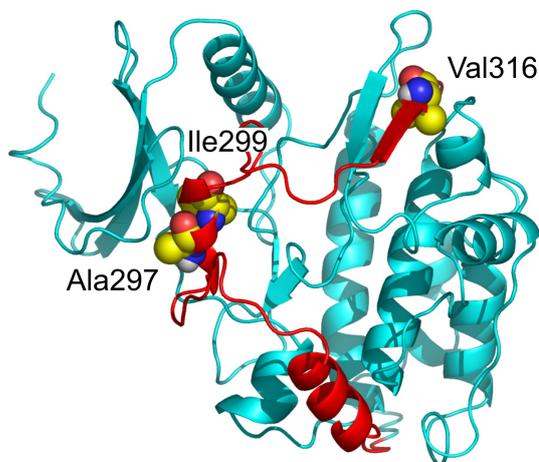


Figure 4.1: Location of titin kinase residues subjected to point mutations (shown as spheres) at the  $\alpha$ R2-helix and  $\beta$ -strand  $\beta$ R1 of the auto-regulatory tail (red).

mutant was subjected to FPMD simulations (see Fig. 2.4.1) with a pulling velocity of  $0.004 \text{ \AA/ps}$ , the lowest velocity used for the wt. Simulation parameters for the MD and FPMD simulations are identical to those described in Section 3.2.1.

## 4.3 Results

### 4.3.1 $\alpha$ R2-mutants

Both residues Ala297 and Ile299 are located within the  $\alpha$ R2-helix. Fig. 4.2 shows the structural changes in the protein upon point mutation of one of these residues to glycine. As expected for  $\alpha$ -helices, introduction of glycine destabilizes the secondary structure to a large extent. As a consequence, within a simulation time of 7 ns, the RMSD of this helix with respect to the wt increases to as much as 0.45 nm (Fig 4.2a) in both cases. The helical content of the regulatory tail (Fig 4.2b) simultaneously decreases in the region of the mutated residues.  $\alpha$ R2 unwinds partially or, in the case of I299G, completely. The free dynamics of the titin kinase with both Ala297 and Ile299 mutated to glycine were also monitored during MD simulations. For this double-mutant, a similar but more pronounced tendency of  $\alpha$ R2 disintegration was observed (data not shown).

The wt features an  $\alpha$ R2 conformation which forms numerous hydrophobic contacts with the active site, thereby substituting ATP and inhibiting ATP

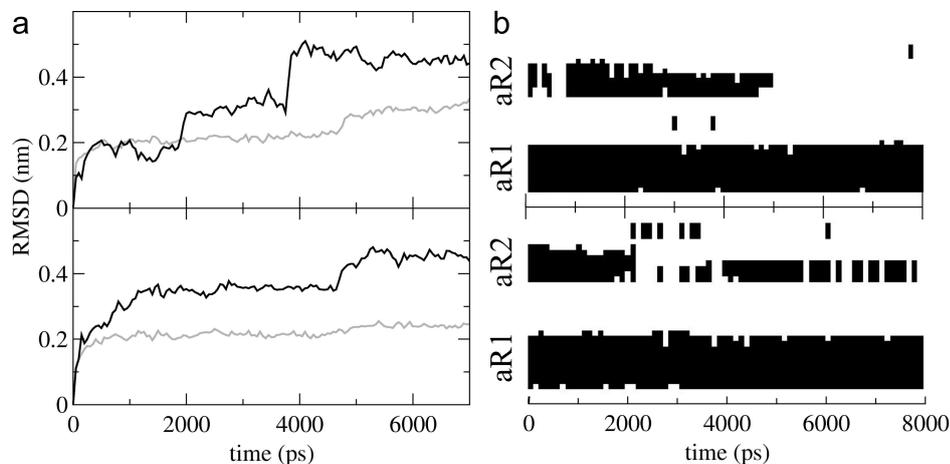


Figure 4.2: Equilibration of titin kinase after the point mutations I299G (upper panel) and A297G (lower panel). (a) RMSD of the protein (gray) and  $\alpha$ R2 (black) with respect to the equilibrated titin kinase wt structure. (b) Secondary structure (helical: black, non-helical white) of the regulatory tail (residues 274-305). Both mutations induce dislocation and complete (I299G) or partial (A297G) disruption of the  $\alpha$ R2 helix.

binding. The hydrophobic surface of the  $\alpha$ R2 helix is complementary to the active cleft, thereby ensuring high affinity binding

Fig. 4.3a compares the Lennard-Jones interaction energy between  $\alpha$ R2 and the protein for the wt and the A297G and I299G mutants. First, upon mutation, the hydrophobic interaction of the mutated residue itself to residues at the cleft is lost (time 0). Second, the tight binding of  $\alpha$ R2 to the cleft is further weakened as reflected by a further decrease in favorable Lennard-Jones energy during equilibration of the mutants. As shown in Figs. 4.3b-d, this is due to the unwinding of the helix and the subsequent loss of shape complementary to the active site cleft upon mutation.

The free dynamics of the titin kinase A297G and I299G mutants allow indirect conclusions on their force-resistance. In both cases, the intra-helical hydrogen bonds of  $\alpha$ R2 as well as the hydrophobic interactions to the protein core were substantially reduced. The sequence forming  $\alpha$ R2 in the wt might be only partially folded to an  $\alpha$ -helix bound to the active site, or even might not fold at all. Therefore, force-induced unfolding can be expected to give rise to a lower force maximum, which for the wt was traced back to the  $\alpha$ R2 rupture, as shown in Fig. 3.7.

Our result that the  $\alpha$ R2 mutants are less force-resistant than the wt is

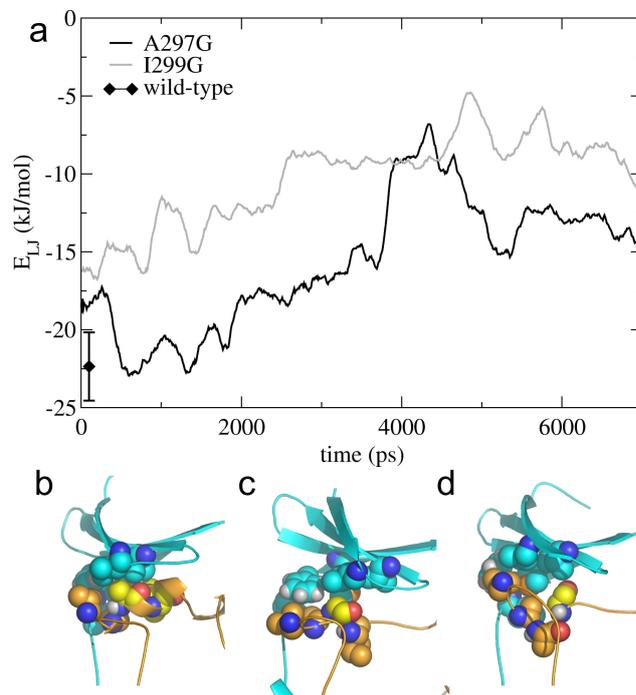


Figure 4.3: Weakening of hydrophobic contacts between the regulatory and catalytic domain upon point mutation. (a) Total Lennard-Jones energy between  $\alpha$ R2 and the catalytic lobe. The respective energy of the wt is given as reference. (b)-(d) Snapshots of titin kinase wt (1TKI), and mutants A297G and I299G (each after 7 ns of equilibration), respectively. Residues at the hydrophobic  $\alpha$ R2-catalytic lobe interface are depicted as spheres. Cyan: catalytic domain, orange: regulatory domain, yellow: A297 and I299 (b), G297 (c), G299 (d).

based on the assumption that the mutants only alter the local structure of the regulatory tail in the sequence adjacent to the mutated residue. The instant unwinding within the short nanosecond time scale and the involved sharp step-wise increase of the structural deviation from the wt (Fig. 4.2a), however, indicate that both mutations induce a significant destabilization of the wt fold. The simulations can thus not be assumed to be converged to equilibrium mutant conformations, which are also experimentally unknown. An equilibrated mutant structure as starting conformation was not accessible for subsequent FPMD simulations, which therefore were not performed. No proper assumptions on the unfolding pathway and forces can be made for the unlikely but not excludable case that the mutations also affect the integrity of the small lobe or the C-terminal  $\beta$ -sheet (compare Fig. 3.2).

### 4.3.2 $\beta$ R1-mutant

The first major rupture event during enforced wt unfolding is the sequential rupture of the  $\beta$ -sheet at the C-terminus (Fig. 3.7). The V316P mutant reduces the number of hydrogen bonds of this  $\beta$ -sheet, and is therefore expected to facilitate its rupture and consequently the release of the regulatory tail, which is covalently bound to it, from the active site. Is this indeed the case, and does the V316P mutation alter the force profile and underlying unfolding mechanism also in other respects?

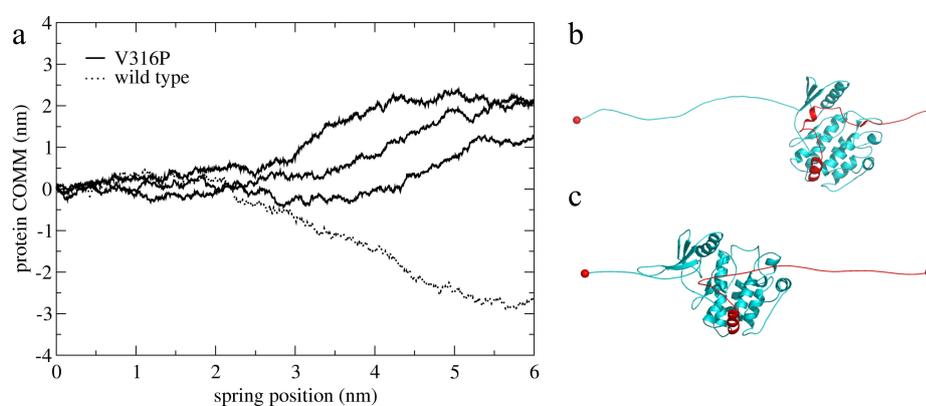


Figure 4.4: Motion of the protein core during FPMD simulations. (a) Center of mass motion (COMM) during force-induced unfolding of the V316P mutant (3 simulations) and the wt. Pulling velocity for all simulations is  $4 \cdot 10^{-4}$  nm/ps. (b) and (c) Snapshots of the V316P mutant and wt titin kinase at a spring position of 6 nm. Cyan: catalytic domain, red: regulatory domain, spheres: pulled atoms.

During enforced unfolding of titin kinase, each of the rupture events releases a protein segment from the core, and the protein adjusts its position so as to balance the force the two termini are subjected to. The COM motion gives insight into the release of peptide chains upon rupture at the protein ends, and thus is a measure for the differential stability of the N- and C-terminus. It is shown in Fig. 4.4 for the wt and V316P mutant. As discussed in Section 3.3.2, rupture of the C-terminal  $\beta$ -sheet precedes unfolding of the N-terminal  $\beta$ -sheet. This results in a motion of the protein core in the same direction as the N-terminal spring. In contrast, the V316P mutant follows the C-terminal spring, reflecting that, while mainly the N-terminal  $\beta$ -sheet unfolds, the C-terminus stays intact. This means that the V316P mutation does not, as expected, further destabilize the C-terminal sheet, but instead, astonishingly, just the opposite is observed: the proline residue within the C-terminal  $\beta$ -sheet apparently increases the force-resistance of the module.



This and the following results were reproduced in three independent FPMD simulations starting from different snapshots of the V316P equilibration. If not otherwise mentioned, representative results from a single trajectory are shown.

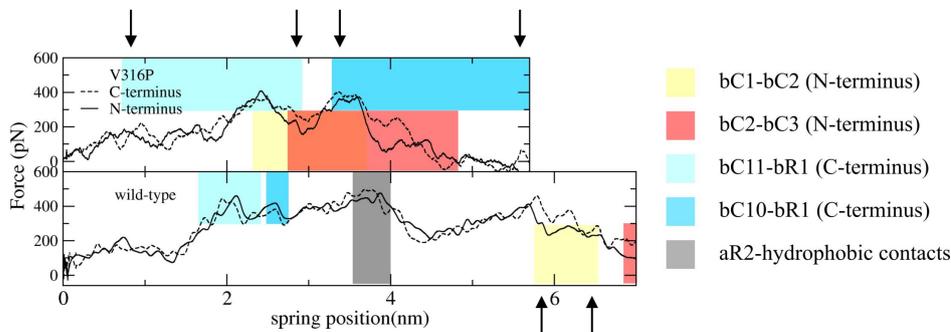


Figure 4.5: Force-extension profiles during FPMD of the V316P mutant and the wt. Areas indicate the main hydrogen bond and hydrophobic rupture events. Arrows indicate sliding-snapping events during concurrent rupture (see text and Fig. 4.7 for details).

The sequence of rupture events and the corresponding force profile of V316P and the wt are shown in Fig. 4.5. Overall, V316P unfolding gives rise to lower rupture forces ( $F_{\max}(N) = 387 \pm 16$  pN,  $F_{\max}(C) = 388 \pm 6$  pN, average and standard deviation over three simulations) than the wt ( $F_{\max}(N) = 476$  pN,  $F_{\max}(C) = 496$  pN). Furthermore, forces decay already at lower extensions in the case of V316P. Thus, as expected the V316P mutation impairs mechanical stability.

As can be seen, in the V316P mutant, the  $\beta C11-\beta R1$  sheet starts to rupture first, similar to the wt. However, the further rupture of the  $\beta C11-\beta R1$  and  $\beta C10-\beta R1$  sheets is remarkably slow, continues over 5 ns of the FPMD simulation, and is not completed within the simulation time. It is *accompanied* and not, as in the wt, *followed* by the rupture of  $\beta C1-\beta C2$  and  $\beta C2-\beta C3$  at the N-terminus.

This is also confirmed by the changes in the secondary structure during unfolding, as shown in Fig. 4.6. The mutation of V316 to Pro within  $\beta R1$ , as expected, perturbs the  $\beta C10-\beta R1$  sheet locally. Opening of the sheets at the N-terminus occurs, while the part downstream of P316 stays intact up to the simulated extension of 6 nm. As a consequence, also  $\alpha R2$  does not detach from the catalytic domain, as indicated by the perpetuation of helical structure.

The opposite was observed for the wt, with the C-terminal zipper-like open-

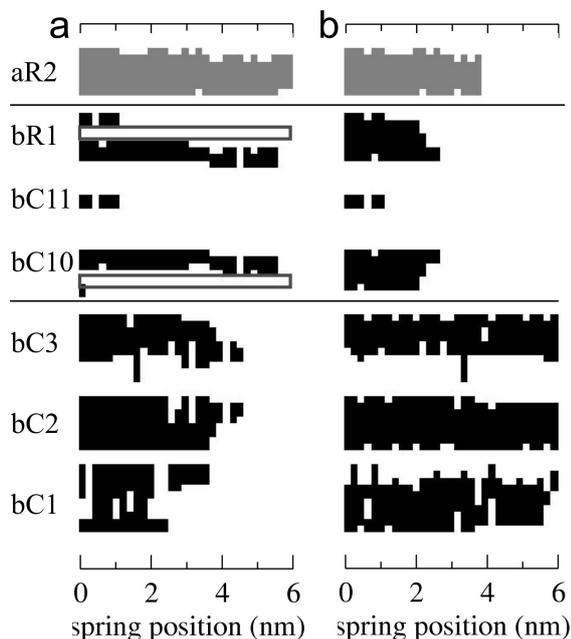


Figure 4.6: Secondary structure changes during FPMD simulations of the V316P mutant (a) and the wt (b). Black:  $\beta$ -sheet, gray:  $\alpha$ -helix, gray box: interruption of the  $\beta$ -sheet structure due to the V316P mutation.

ing of  $\beta$ C10- $\beta$ R1 followed by  $\alpha$ R2 rupture and detachment, while the N-terminal sheets stay intact (Fig. 3.9). In summary, surprisingly, the V316P mutation stabilizes the C-terminal with respect to the N-terminal  $\beta$ -sheet. Consequently, unraveling the N-terminal small lobe precedes detachment of the auto-regulatory tail from active site cleft between the small and big lobes (compare Fig. 3.2). For the V316P mutant, enforced unfolding does not yield an active conformation of titin kinase like for the wt (Fig. 3.14d), but instead results in disintegration of the small lobe and active site prior to the relief of auto-inhibition. In other words, our simulations predict that V316P is not activated under force.

What is the molecular basis of the unexpected order of mechanical  $\beta$ -sheet stability? Fig. 4.7 shows the sequence of hydrogen bond ruptures in the terminal sheets as found for the V316P mutant. The two parts of the  $\beta$ C10- $\beta$ R1 sheet interrupted by Pro316 break independently from each other, each of them concurrently. This is preceded by a sequential opening of the N-terminal  $\beta$ -sheet. For the wt, the mechanism of rupture was found to occur exactly just vice versa (compare Fig. 3.9). Previous FPMC simulations (see Section 3.3.4) have attributed the differential mechanical resistance to a differential rupture mechanism. In this light, the swap of relative stabilities of

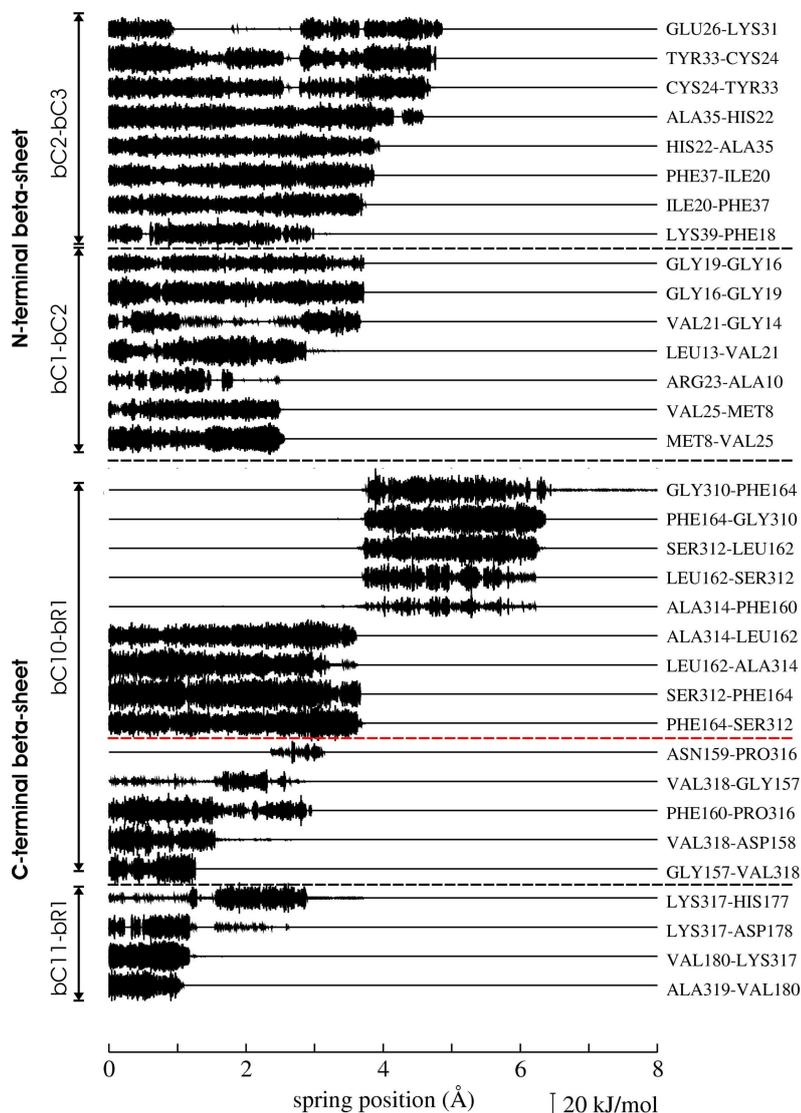


Figure 4.7: Potential energies of inter-strand hydrogen bonds during FPMD simulations. N-terminal sheets rupture sequentially, C-terminal sheets concurrently. The  $\beta$ C10- $\beta$ R1 sheet is split into two by the insertion of Pro316 (red dashed line).

the C- and N-terminal  $\beta$ -sheets by the V316P mutation can be traced back to the swap of rupture mechanisms. Here, the C-terminal  $\beta$ -sheet is stabilized over the N-terminal one, because its sliding-snapping mechanism requires a higher mechanical load than zipper-like opening at the N-terminus.

For the wt, the orientation of the  $\beta$ -sheet towards the pulling direction was found to determine the rupture mechanism. The question arises, if the reason for the concurrent rupture of the C-terminal  $\beta$ -sheet is a different

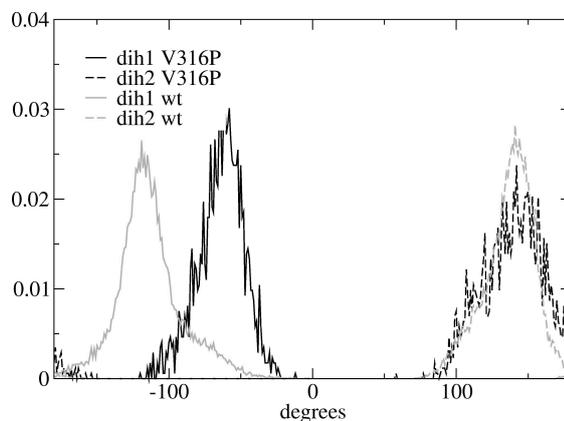


Figure 4.8: Backbone dihedral distribution for the titin kinase wt and V316P mutant in comparison. Dihedral 1: K315(C)-P316(N-CA-C), dihedral 2: P316(N-CA-C)-K317(N). Dihedral 1 differs, leading to a misalignment of the  $\beta$ -sheet parts upstream and downstream of P316.

alignment of the sheet with respect to the pulling direction due to the V316P mutation. Fig. 4.8 compares backbone dihedrals comprising residue 316 in wt and mutant titin kinase. The dihedral to upstream Lys315 is rotated by 90 degree upon mutation. The backbone therefore adopts a conformation that significantly differs from the wt. A possible explanation for the observed concurrent rupture of this  $\beta$ -sheet is its different orientation towards the pulling direction. It can be assumed that the N-terminal  $\beta$ -sheet started to rupture before it was properly aligned parallel to the pulling direction. It thus underwent sequential rupture and was overall less mechanically stable. In summary, the V316P mutation only locally altered the titin kinase fold but resulted in an unexpected unfolding pathway under force, that featured an opposite force resistance, rupture sequence and rupture pattern of the terminal  $\beta$ -sheets.

## 4.4 Discussion and Conclusion

The simulations of enforced unfolding of wt titin kinase identified the terminal  $\beta$ -sheets and the interaction of the regulatory tail to the active site as the structural modules responsible for the mechanical stability and force-sensor function of the kinase. The differential stability of the terminal sheets was found to be essential for the force-induced activation and thus physiologically important. This study aimed at tuning — more specifically reducing — the force-resistance of titin kinase by point mutations at the mechani-



cally crucial modules to validate and modify the unfolding mechanism and its determining factors as concluded from the wt simulations.

We found the  $\alpha$ R2 mutations A297G and I299G to cause, as expected,  $\alpha$ -helix unwinding and partial detachment of the active site. This suggests these mutants — if they are stable at all — to be less force-resistant than the wt titin kinase, which remains to be tested experimentally. Hence, observing in single-molecule force probe experiments lower rupture forces than the wt for both mutants would validate the unfolding mechanism of titin kinase as obtained from FPMD simulations. Assuming that this validation is in fact possible, a directed tuning, or more specifically lowering, of the mechanical stability by eliminating stabilizing interactions for these two mutants can be achieved.

The V316P mutant was expected to unfold similarly to the wt, but with a C-terminal  $\beta$ -sheet rupture at even lower mechanical load, *i. e.*, occurring earlier. Surprisingly, the opposite effect was found, due to the reversal of the rupture mechanisms of the terminal  $\beta$ -sheets. This has three striking consequences. First, this example confirms our previous result that the force resistance of  $\beta$ -sheets is primarily determined by the mechanism of hydrogen bond rupture, more than by the number of inter-strand hydrogen bonds. The V316P mutation reduces the number of hydrogen bonds in the  $\beta$ -sheet by one, but nevertheless stabilizes it, as it allows for the concurrent rupture of the henceforth two independent  $\beta$ -sheets.

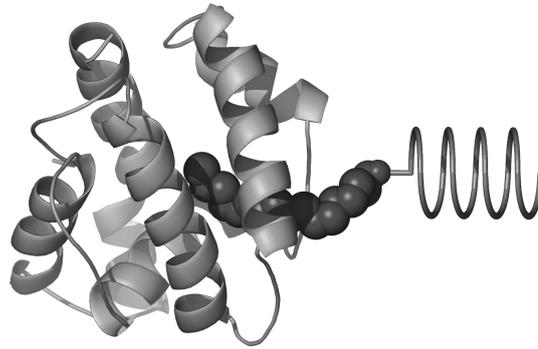
Second, this study underlines the impact of the pulling direction, *i. e.*, the reaction coordinate, along which unfolding is enforced, onto the resulting rupture mechanism. This has been previously shown for wt titin kinase (Chapter 3), ubiquitin [140], and E2lip3 [130]. In these examples, however, different  $\beta$ -sheets of the respective protein subjected to force are compared. The mutant example here is particularly striking, since it is the same  $\beta$ -sheet in the same protein fold that differs from the wt by a single hydrogen bond and backbone dihedral that gives rise to a different rupture mechanism.

Third, the unexpected behavior of the V316P mutant demonstrates that the presence/absence of crucial interactions and the observed forces do not directly correlate, but instead are related to each other in a more complex way. Since protein secondary structure motifs mechanically depend on each other, their mechanical resistance is not simply additive. The stability of a protein under force can therefore in general not be directly tuned by switching the stabilizing factors on or off. This poses a major challenge of systematically engineering a protein with respect to its mechanical properties.

#### 4. *TITIN KINASE MUTANTS*

---

The titin kinase mutants studied here gave further insight into the stabilizing role of the mechanically crucial modules along the unfolding pathway. Our results remain to be finally compared to experiment. AFM experiments of wt and mutant titin kinase are currently under way in the group of H. Gaub, LMU Munich.



## 5

# Ligand release pathways in the pheromone binding protein of *Bombyx mori*

---

### 5.1 Olfaction and the bombykol-BmPBP complex

Olfactory systems of animals share general, yet largely unknown, mechanisms to achieve their required sensitivity and specificity. L. Buck and R. Axel were awarded the Nobel Prize in Medicine in 2004 for their pioneering work on odorant receptor (OR) proteins. A certain OR is exclusively expressed in special olfactory cells that thereby only respond to specific odorants, resulting in an overall 'odorant pattern' in response to an odorant blend [146]. Another common principle of olfaction in insects is the assistance by odorant-binding proteins (OBPs), which are an integral part of the odorant signalling pathway, and are restrictively expressed in olfactory tissues [147, 148]. OBPs are nanocapsules that solubilize small hydrophobic molecules of the odorant blend by binding them into their central cavity [149–152]. The odorant molecule is carried by the OBP to the neuronal membrane and presented to the OR for receptor activation (Fig. 5.1a). For

the function of an OBP as a carrier and for putative additional roles in odorant discrimination, receptor activation and odorant deactivation, its ligand binding properties and uptake/release mechanism need to be individually tuned [153]. How this is achieved is yet to be elucidated and one aim of this thesis.

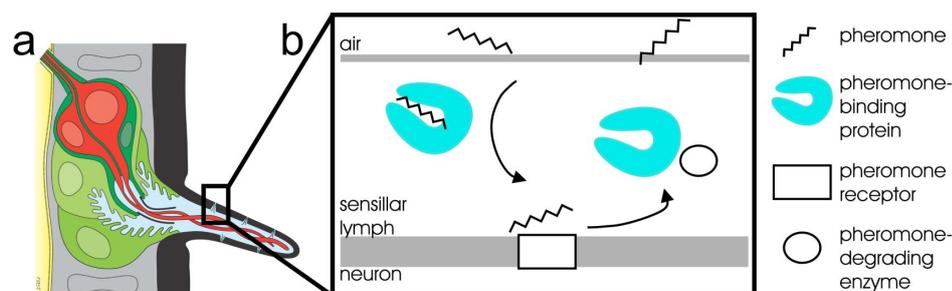


Figure 5.1: Olfactory system of *Bombyx mori*. (a) Schematic drawing of an olfactory hair of a moth antenna. Blue: sensillar lymph, red: neuron, black: cuticle with pores for pheromone entrance. Kindly provided by R. A. Steinbrecht W. S. Leal. (b) Pheromone reception involves the entrance of bombykol into the sensillar lymph, transport of bombykol by the pheromone binding protein to the pheromone receptor, stimulation of the receptor by bombykol and bombykol degradation.

The olfactory communication system of insects is an ideal model system for studying olfaction at molecular detail. The spectrum is able to discriminate subtle differences in the chemistry of the small organic pheromone molecules, a special class of odorant molecules for sexual attraction [154, 155]. In the well-studied case of the silk moth, *Bombyx mori*, the pheromone bombykol, produced and released by the female insect, is detected over large distances by the olfactory antennae of the male. As sketched in Fig. 5.1b, the pheromone binding protein from *Bombyx mori* (BmPBP) acts as the transporter for bombykol through the sensillar lymph of the antennae to the pheromone receptor [147, 156, 157].

Fig. 5.2 shows the crystal structure of bombykol-complexed BmPBP, the first three-dimensional structure of an insect OPB solved [158]. The six  $\alpha$ -helices of PBP, stabilized by three interhelical disulphide bonds, form a rigid capsule with a mainly apolar core and a highly charged polar surface. Bombykol is completely enclosed within the hydrophobic cavity. Its hydroxyl group forms a hydrogen bond to one of the few internal polar residues, Ser56 (yellow in Fig. 5.2b and c). Bombykol binding to BmPBP is extremely strong with a dissociation constant  $K_D$  of 105 nM [159] at physiological pH (previous measurements gave 60 nM and 640 nM, respectively [160, 161]). The NMR structure of uncomplexed BmPBP shows a highly similar fold

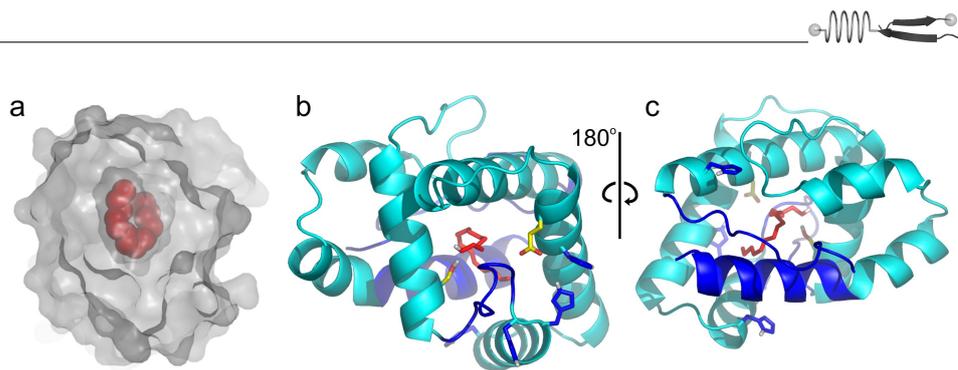


Figure 5.2: (a) Bombykol (red spheres) is completely surrounded by its receptor BmPBP (gray surface). (b) and (c) The bombykol-PBP complex (1DQE [158]) viewed from the front and back, respectively. Bombykol is shown in red, hydrogen bond partners Ser56 and Asp98 are shown in yellow, and His residues as blue sticks. Regions presumably involved in ligand passage, the front lid (residues 60-68) and the terminal fractions (residues 1-14 and 126-137), respectively, are colored blue.

with an empty yet preformed cavity to accommodate bombykol [162].

As for many other OBPs in general, a crucial yet unsolved question is the mechanism of ligand entrance and exit in BmPBP. Obviously, ligand uptake and release must be fast, *i. e.*, it must occur within a ms time scale or less to guarantee an equally fast nerve signal decay for sufficiently frequent readjustments of the flight direction. Yet, as can be seen in Figure 5.2a, the protein conformation in the X-ray structure builds up a tight envelope around bombykol without an obvious entrance/exit gate. To resolve this puzzle, we here address the question at which site of the structure BmPBP provides a gate for efficient bombykol uptake and release. Apparently, this requires a sufficiently low energetic barrier.

BmPBP undergoes a conformational transition from the structure at neutral pH to an acidic fold at pH 4.5 [163, 164] with low binding affinity [165]. Concerning the correlation of the pH-dependent conformational transition and the ligand uptake and release, two possible exit gates have been envisaged, ligand passage along the front lid formed by residues 60-68 ('front' pathway, Fig. 5.2b), or along the N-terminal and C-terminal chains located at the back of BmPBP ('back' pathway, Fig. 5.2c) [158, 164]. A previous MD study [166] pointed, though only indirectly, towards the front lid as the favoured entrance and exit route of the ligand.

This thesis aims at a characterization of and the discrimination between the two putative ligand passage ways by means of MD simulations of the bombykol release from BmPBP. To this end, the pathways for dissociation towards the front and back, as obtained from essential dynamics

(ED, Section 2.4.3) [34, 35] and force probe MD simulations (FPMD, Section 2.4.1) [43, 167], were compared in terms of their mechanisms and forces.

Besides their function as a passive detergent and carrier, an additional role of OBPs, and BmPBP in particular, in odorant discrimination has been suggested [161, 168, 169]. Such a pre-filtering function of OBPs would require a sufficient binding selectivity for the relevant odorant molecules. Due to the extremely low solubility of ligands, biochemical measurements of the binding specificity of OBPs are problematic [161]. Biochemical and mass-spectroscopic ligand binding studies supported the hypothesis that OBPs contribute, at least to some extent, to the specificity of odorant reception [170–172]. A previous quantum mechanical study of bombykol-BmPBP binding suggested a high selectivity, for which mainly hydrophobic interactions were found to be responsible [173]. However, this study was restricted to the static crystal structure, hence omitting entropic contributions to the binding free energy and, in particular, the hydrophobic effect. In this study, we therefore also address the question of ligand specificity in the case of BmPBP by characterizing the dynamics and energetics of bombykol-BmPBP interactions, including explicitly the solvent.

A second route to the study of binding specificity is opened by the observation that not only bombykol, but also its oxidized form, the aldehyd bombykal, is a component of the pheromone gland. These two highly similar components differ in their physiological response, and thus represent an interesting model system to study selectivity for olfaction (Fig. 5.1a). Each olfactory hair consists of two olfactory cells, each of which detects one of the two pheromones. These cells share the sensillar lymph including the PBP, but express two different ORs, which specifically respond to either bombykol or bombykal [157, 174]. To what extent BmPBP assists the ORs in the discrimination between bombykol and its oxidized form, the aldehyd bombykal, is largely unknown [169]. To assess the specificity of pheromone binding to BmPBP for these two pheromones, the dynamics and protein-ligand interactions of BmPBP complexed by bombykol and bombykal, respectively, are here compared. In addition, as a direct measure for binding specificity, their difference in binding free energy is calculated.



## 5.2 Simulation details

### 5.2.1 MD and replica-exchange MD simulations

Protonation states of titratable groups were determined by calculating  $pK_a$  values for 1DQE [158] with Whatif [123] and its interface to DelPhi [124]. None of the five histidine residues of 1DQE was found to be cationic at pH 7. The GROMOS96 force field [62] for the protein and the SPC water model [122] were used. Molecular mechanical force field parameters for bombykol (listed in Tables 7.1 and 7.2 in the Appendix) and bombykal (listed in Tables 7.3 and 7.4 in the Appendix) were determined as follows. Non-bonded and bonded parameters for the saturated part of the aliphatic chain and for the hydroxyl group were adopted from chemically similar fragments, *i. e.*, from serine and threonine, as parametrized in the GROMOS96 force field. For the  $\pi$ -conjugated part of the hydrocarbon chain, Lennard-Jones parameters of the CH-groups were taken from the GROMOS96 force field. Merz-Kollman charges [175] were calculated at the B3LYP/6-31G(d) level. They were found to deviate only slightly from zero (-0.17-0.16) for a CH group of the hydrocarbon chain and showed high conformational dependency. Therefore, zero charges were applied except for the polar part of bombykol.

Two parameter sets, p1 and p2, have been set up that differ in the parametrization of the conjugated double bonds. Bond lengths and angles of p1 are those found for bombykol in 1DQE [158]. Force constants of p1 were adopted from quantum mechanical calculations of model systems for the retinal chromophore [176]. More specifically, the corresponding bond and angle force constants were adopted from [177], the dihedral constant for torsion around a double bond from dodecapentene [178], for torsion around the central bond of the  $\pi$ -system from butadiene [179]. For set p2, all bonded parameters were taken from the GROMOS96 force field for retinol [62]. Parameter set p1 was also used for bombykal except for the polar head group. The angle including the carbonyl group was set to 125 degrees, as taken from crystallographic studies and Hartree-Fock calculations of acetaldehyde [180,181]. Other parameters, such as the bond length and bond and angle force constants, were adopted from the GROMOS96 force field for carbonyl groups [62].

Four simulation systems were setup, as detailed in Table 5.1. All four systems were subjected to the following simulation protocol. The respective protein was solvated in SPC water [122] molecules in a cubic box of size  $7.1 \times 6.7 \times 7.0 \text{ nm}^3$ . Twelve chloride and 20 sodium ions were added to yield a

zero net charge and a physiological ion strength. All four systems comprised  $\sim 32,000$  atoms each. After 1000 steps of steepest descent energy minimization, the solvent and ions were equilibrated during a 0.5 ns MD simulation with the protein heavy atoms subjected to harmonical constraints with a force constant of  $k = 1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ . Subsequently, all four systems were equilibrated as shown in Table 5.1.

All simulations were carried out using the MD software package GROMACS 3.2.1 [121] except for the FPMD simulations, for which GROMACS 3.1.4 was used due to technical reasons. Simulation time lengths comprising production and equilibration of the system are given in Table 5.1. All simulations were run in the NPT ensemble. The temperature was kept constant at  $T = 300 \text{ K}$  by coupling to a Berendsen thermostat with a coupling time of  $\tau_T = 0.1 \text{ ps}$  [71]. The pressure was kept constant at 1 bar by coupling to a Berendsen barostat with  $\tau_p = 1.0 \text{ ps}$  and a compressibility of  $4.5 \cdot 10^{-5} \text{ bar}^{-1}$  [71]. All bonds were constrained using the LINCS algorithm [65]. An integration step of 2 ps was used. Non-bonded interactions were calculated using a cutoff of 10 Å. Except from free energy perturbation simulations (see below), long-range electrostatic interactions were calculated by Particle-Mesh Ewald summation [69] with a grid spacing of 0.12 nm and cubic interpolation.

| No. | starting structure                                  | time  | remarks                          |
|-----|---|-------|----------------------------------|
| MD1 | 1DQE  | 45 ns | bombykol set p0, for REMD and ED |
| MD2 | 1DQE  | 45 ns | bombykol set p1                  |
| MD3 | snapshot from 1, bombykol as hydrogen bond donor    | 60 ns | bombykal                         |
| MD4 | snapshot from 1, bombykol as hydrogen bond acceptor | 60 ns | bombykal                         |

Table 5.1: Details of MD simulation systems for equilibration

Replica-exchange MD (REMD) simulations (see Section 2.3) with six replica, each 18 ns in length, have been performed, starting from simulation system MD1. Simulated temperatures were 200, 280, 360, 440, 540 and 600 K, respectively. BmPBP and bombykol were coupled to a heat bath at the respective temperature, whereas solvent and ions were coupled independently to a 300 K heat bath, with a coupling time of  $\tau_T = 0.1 \text{ ps}$  [71]. In this way,



the number of degrees of freedom of the system adding to the difference of potential energies of two replicas is reduced. This narrows the energy distributions, *i. e.*, increases their overlap, and thus results in a reasonable acceptance ratio of approx. 8% at the desired temperature spacing, in spite of the large system size. Exchanges of configurations were attempted every 2 ps.

## 5.2.2 Essential dynamics and force probe MD simulations

Along the first eigenvector (EV) obtained from the principal-component analysis [85,182,183] of the REMD trajectory, essential dynamics (ED) simulations [34,35] were performed (see Section 2.4.3). At each simulation step, the system was forced to increase or keep the distance to its starting position along the first EV, while the other degrees of freedom were unperturbed (Fig. 2.5, bottom). The maximal step size along the EV was restricted to  $10^{-5}$  nm/step to ensure that the system proceeds slowly enough to allow equilibration of the remaining degrees of freedom.

Two ED simulations were carried out. For the first, unbinding towards the front lid was enforced by increasing the position along the first EV; for the second, unbinding towards the back was enforced by decreasing it. Sampling of forces during the ED simulations was improved by carrying out additional MD simulations at equidistant points along the first EV. Snapshots of the two ED trajectories, in total 34, were taken and simulated for 6 ns each. During each of these simulations, the position on the first EV was kept fixed and the other degrees of freedom were equilibrated. Constraint forces, acting on the first EV, were recorded every 100 MD steps, allowing the reconstruction of a potential of mean force (PMF), as detailed in Section 2.5.2. Convergence of forces and free energies was monitored.

As starting structures for FPMD simulations, four snapshots were taken from MD1. For each of the snapshots, the system size was increased by adding water and ions at the side to which bombykol was pulled out of the cavity such that the box can accommodate the dissociated protein and ligand, yielding box dimensions of approximately  $7.1 \times 6.7 \times 8.1$  nm<sup>3</sup>, containing  $\sim 38,000$  atoms. Additional water molecules and ions were equilibrated during a 100 ps MD simulation with positional constraints on the protein and ligand as described above. In subsequent FPMD simulations, the center of mass of bombykol was subjected to a harmonic spring potential, which was moved along the pulling direction with constant velocity, as described in Section 2.4.1. The force constant of the spring was set to  $500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ ,

the pulling velocity to 1 m/s. Simulations with ten different pulling directions were performed, starting from the four snapshots. As indicated in Fig. 5.9, one pulling vector was chosen such that it points directly towards the front and back pathways. The others were defined to form an angle of 55 degree with respect to this vector. The center of mass of the protein was fixed to prevent the protein from being dragged by the ligand motion. Forces were recorded at each MD step and smoothed with a Gaussian filter of 1.6 Å width [42].

### 5.2.3 Free energy calculations

The difference between the binding free energies of bombykal (al) and bombykol (ol),  $\Delta\Delta G_b = \Delta G_{b,al} - \Delta G_{b,ol}$ , was calculated from the free energies  $\Delta G_{\text{FEP}}(\text{prot})$  and  $\Delta G_{\text{FEP}}(\text{aq})$ . The latter correspond to the free energy difference between the two ligands in aqueous solution and in the protein complex, respectively, and were calculated using free energy perturbation (FEP) (for details, see Section 2.5.3).

The underlying thermodynamic cycle is shown in Fig. 5.14a (compare Fig. 2.8).  $\Delta G_{\text{FEP}}(\text{prot})$  was obtained from simulating the (unphysical) pathway from bombykol to the bombykal when bound to the protein by 'mutating' the ligand from bombykol to bombykal.  $\Delta G_{\text{FEP}}(\text{aq})$  was obtained in the same way, but for the ligand in solution. The 'mutation' was done by interpolating between the two ligand topologies [38,39] as follows. Within the time course of one FEP simulation, the scaling factor  $\lambda$  was gradually changed from 0 to 1 with an identical increase of  $\lambda$  at each MD step. The bombykol and bombykal bonded potential energies were linearly scaled by  $\lambda$  yielding a total bonded potential  $V_b$ ,

$$V_b = (1 - \lambda)V_{b,ol} + \lambda V_{b,al}. \quad (5.1)$$

The hydrogen atom of the bombykol hydroxyl group was mutated to a dummy atom, *i. e.*, its charges and Lennard-Jones parameters were changed to zero. The appearance or disappearance of this atom close to  $\lambda = 0$  and  $\lambda = 1$  gave rise to singularities of the non-bonded potentials when interpolated linearly. To remove this problem, non-bonded interactions were calculated with soft-core potentials [91,92]. The soft core parameter  $\alpha$  controls the height of the potential at zero distance and was set to 1.51 [93]. The interaction radius  $\sigma$  is given by the Lennard-Jones parameters, with  $(C_{12}/C_6)^{1/6}$ , for  $C_{12}, C_6 \neq 0$ , otherwise  $\sigma = 0.3$ .



The starting structure for the bombykal-BmPBP equilibration MD3 (Table 5.1) also served as starting structure for the FEP calculations. For the FEP calculation of the ligand in solution, bombykol was taken from the crystal structure 1DQE [158], solved in SPC water in a  $4.0 \times 3.4 \times 3.8 \text{ nm}^3$  box and equilibrated for 5 ns after a 200 ps MD simulation with harmonic constraints on the bombykol atoms. FEP simulations of the ligand in solution and in the complex were carried out for 1 ns, 2 ns and 5 ns each, and each FEP simulation was performed in both directions, to check for the convergence of the free energy calculations. The 1 ns and 5 ns simulations of the complex were done three times with different starting velocities to estimate the error due to insufficient sampling of the highly heterogeneous protein-ligand interactions. The simulation protocol described above for the equilibrium MD simulations was applied, except for the non-bonded interactions, which here were calculated using a cut-off of 1.5 nm for the Coulombic potential, 1.2 nm for the Lennard-Jones potential. Due to technical restrictions, PME was not applied for the FEP simulations.

## 5.3 Results and Discussion

### 5.3.1 Equilibrium dynamics of the bound complex

The BmPBP and bombykol dynamics and interactions in the bound state were characterized by putative MD simulations to suggest conceivable pathways for ligand entrance and exit. Fig. 5.3a quantifies the fluctuations during the free dynamics of the protein backbone and bombykol compared to experimental temperature factors from X-ray scattering [158]. As expected, the helical segments, especially when linked *via* disulphide bonds, show only little fluctuations. Regions of particularly high flexibility (shaded in gray) are the terminal residues 1-24 and 125-137, even though the N-terminal fragment is  $\alpha$ -helical and contains one of the disulphide bonds. In addition, the histidine-rich loop (residues 60-68) and the loop proximal to it (residues 99-106) are remarkably mobile. This agrees well with the experimental temperature factors (cyan line), which were also found to be high in these regions. The correlation coefficient between the fluctuations from simulation and the temperature factors is 0.5. Overall, the equilibrium protein dynamics fit into the picture of a rigid scaffold as suggested by the available experimental BmPBP structures. The observed flexibility also suggests the front loop and the terminal fragments at the back as sufficiently flexible lids for ligand entrance and exit.

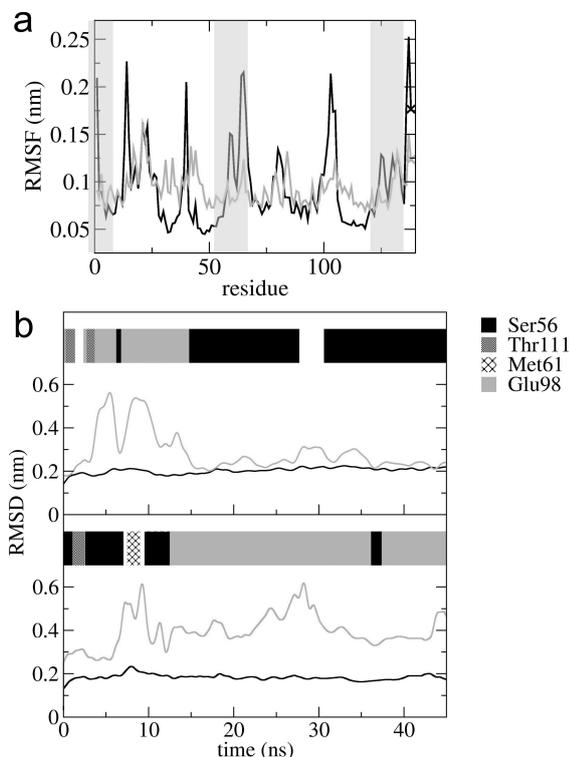


Figure 5.3: Equilibration of the bombykol-PBP complex. (a) RMS fluctuation of PBP (black curve) and bombykol (black circle) during the final 40 ns of equilibration in comparison to the temperature factors of 1DQE, after conversion to fluctuations according to [184] (gray). Gray areas indicate regions along which ligand binding and release might occur (terminal fractions and lid). (b) RMSD of the protein backbone atoms (black curve) and bombykol (gray curve) from the crystal structure during two molecular dynamics simulations. Major bombykol-PBP hydrogen bonds are indicated. Results from two independent MD simulations, using parameters p1 (top) and p2 (bottom) are shown.

As indicated in Fig. 5.3a, bombykol with a root-mean-square fluctuation (RMSF) of 0.17 nm showed strikingly large fluctuations, though apparently encapsulated in the cavity. This finding, too, is in agreement with the X-ray experiment, which gives significantly larger crystallographic temperature factors for bombykol compared to the average temperature factor. The nature of this mobility is detailed in Fig. 5.3b. During the two independent MD simulations with different bombykol parameters for the  $\pi$ -system, p1 and p2, the native bombykol-Ser56 hydrogen bond was transiently and reversibly broken and substituted by bonds to other side chains in the cavity, such as Glu98, Met61, and Thr111. This involved dislocation of the hydrophobic part of bombykol as well, as reflected by the observed high



root-mean-square deviation (RMSD) from the crystal structure (cyan lines in Fig. 5.3b). In one extreme case, bombykol left the binding cavity to some extent and stuck out along the front lid towards the solvent, mostly forming hydrogen bonds to Glu98 and/or water. This outward movement was directed along the putative front lid-pathway and involved partial dislocation of the lid.

To further characterize this collective motion, the curvilinear principal coordinate of the protein that optimally correlates to the ligand motion was calculated as suggested recently by Schröder et al. (submitted). Indeed, significant correlations were found, most notably for the lid dynamics. Fig 5.4a shows representative snapshots of bombykol and the lid after projection onto the non-linear mode of largest amplitude. The mode represented an opening motion of the protein by a lid flip in conjunction with the ligand dynamics, to give way to the half-dissociated ligand.

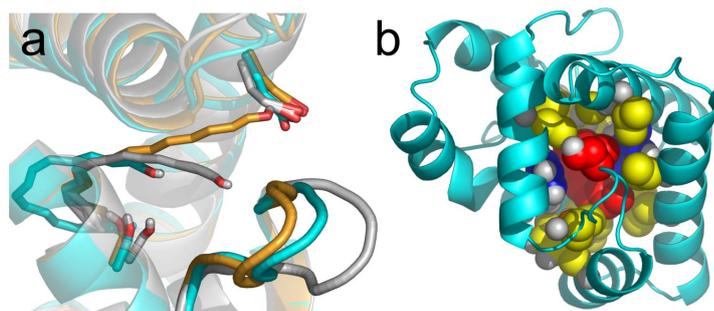


Figure 5.4: Equilibration of the bombykol-PBP complex. (a) Correlation of protein and ligand dynamics. Representative structures along the non-linear collective mode that best correlates to the ligand motion, are shown. It exhibits a lid opening in conjunction with partial ligand dissociation towards the front. (b) Bombykol in the hydrophobic cavity of PBP. Residues with significant contributions to the hydrophobic interactions of the protein to bombykol are shown as spheres. Bombykol: red, BmPBP residues with Lennard-Jones interaction energies to bombykol > 5 kJ/mol: yellow, > 20 kJ/mol: (Phe12 and Phe118): blue.

The high mobility of bombykol is also reflected for the  $C_{14}$ -chain of bombykol by the numerous hydrophobic contacts formed to BmPBP during the MD, as depicted in Fig. 5.4b. Many residues, shown as yellow spheres, of the binding pocket that are highly conserved across general OBPs and PBPs [158] were found to significantly contribute to the total Lennard-Jones (LJ) interaction between protein and ligand. The interaction energies between the ligand and the individual residues typically showed high fluctuations. In the half-dissociated conformation of bombykol with a hydrogen bond to Glu98,

*e. g.*, Val114, located near the front loop, compensated the loss of interaction at the end of bombykol (to residues Leu8 and Leu90), thereby stabilizing the half-dissociated state. With such a high variability and heterogeneity in the ligand polar and non-polar interaction network, bombykol can favorably bind to BmPBP in significantly different conformations implying a considerable entropic component to the binding free energy. This suggests that also changes in the polar head group, the position and stereochemistry of the double bonds, or the chain length do not have *per se* to result in a complete loss of binding affinity. Instead, BmPBP features a moderate ligand specificity that does not suffice to provide the remarkably high olfactory selectivity on its own, but adds to it as a first filter.

The conventional MD simulation described so far also served as a test of the bombykol force field parameters. The experimentally determined structure was reproduced during large parts of the MD simulations, with reversible hydrogen bond rupture and formation. The observation of a highly flexible ligand in the rigid pocket of BmPBP was robust with regard to the ligand force field used (compare Fig. 5.3b). For the subsequent simulations, set p1 was used.

The ligand motion observed in the above MD simulations suggested the front lid as a putative exit pathway as was also found in the previous MD study [166]. However, one has to note that within the short nanosecond time scale of our conventional MD simulations, the complex might not overcome larger energy barriers to regions of the configurational space that are sampled at physiological conditions. In particular, a possible role of the terminal fragments as a possible second lid could not be excluded so far. To enhance sampling, REMD simulations with increased temperatures were therefore performed (Fig. 5.5). With temperature steps of 80 K, exchange attempts of adjacent replicas were accepted with ratios of around 8 %, due to a sufficient overlap of the potential energy distributions (Fig. 5.5a). Thus, simulating all replicas with the same solvent temperature of 300 K resulted in a reasonable acceptance ratio at the desired temperature spacing in spite of the large system size. Fig. 5.5b shows the exchange of a selected replicum within different temperatures (upper panel) and the consequent changes in RMSD of bombykol for two example temperatures (lower panel). The frequent exchanges enable bombykol to effectively sample the conformational space when bound to BmPBP.

As for the conventional 300 K MD simulations described above, partial unbinding towards the front lid was observed (Fig. 5.6a). Additionally, some REMD trajectories also showed a turn of bombykol in the pocket and move-

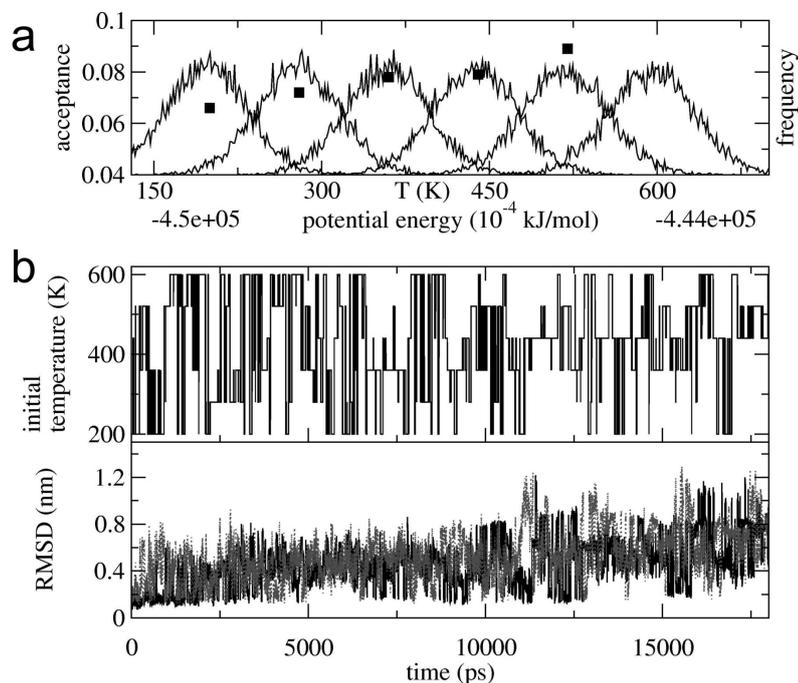


Figure 5.5: Replica-exchange molecular dynamics. (a) Potential energy distribution of the six replicas (black lines) and acceptance ratio for the exchange of the replica at temperature  $T$  with the replica at temperature  $T + 80$  K (squares). (b) Efficiency of REMD for enhanced sampling. Upper panel: exchange of the replica with initial temperature  $T = 200$  K between the six temperatures during REMD. Lower panel: RMSD of bombykol from the crystal structure of example replica during REMD with the initial temperatures as indicated.

ments towards the back passage way near the protein termini, though to a smaller extent (Fig. 5.6b). The partial unbinding towards front and back was found to be represented by the same collective mode of fluctuation (first principal mode obtained from PCA) in opposite directions (Fig. 5.6c). In both directions, conformations of extreme amplitude along the first eigenvector (EV) exhibited reduced protein-ligand interactions and increased solvent exposure of the ligand. This correlated motion of protein and ligand comprises both BmPBP and bombykol dynamics, to which the protein with 87 % of the total fluctuations substantially contributes.

Total unbinding involving complete loss of protein-ligand interactions and the release of the whole ligand into water was not observed, as it would presumably require still more sampling and, hence, much longer simulation times.

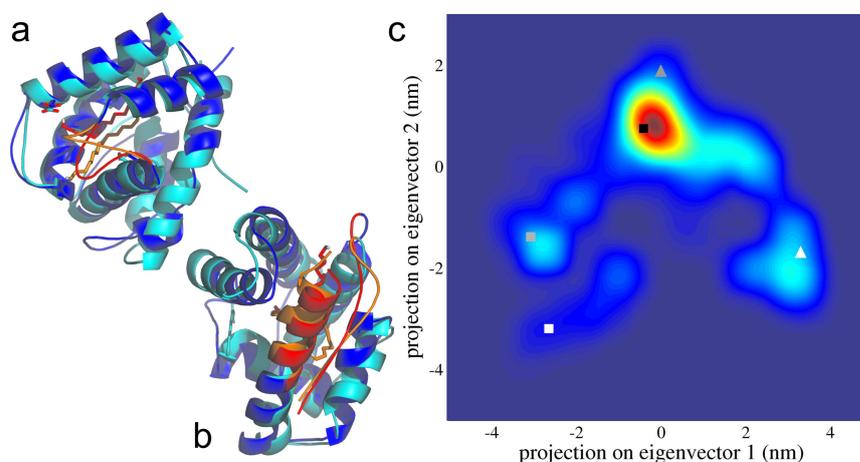


Figure 5.6: Principal component analysis of the REMD. (a) and (b) Representative snapshots of the REMD simulation showing ligand release towards the front and back, respectively. Bombykol and the parts of the protein serving as lids during release are shown in orange and red, respectively. (a) Projection of the REMD trajectory onto the first two EVs. The first EV represents partial ligand release towards the front lid (high projection) and towards the termini at the back (low projection). Structures shown in (a) and (b) are shown as gray triangles (front) and squares (back), respectively, the crystal structure as black square.

### 5.3.2 Ligand release pathways from essential dynamics

To further alleviate the sampling limitations, ED simulations were performed to study full ligand release. The REMD simulations revealed the first EV as a suitable reaction coordinate for the ligand motion towards the back and front exit. This EV therefore was chosen as the coordinate along which ED sampling in positive (front exit) and negative (back exit) direction was carried out. By constraining the system to move along the first EV only in the defined direction, the system was forced to sample a much larger region of conformational space as compared to the free dynamics simulations. This indeed lead to ligand release towards the front and back, as can be seen from snapshots of example trajectories (Fig. 5.7a). ED sampling along the second EV (Fig. 5.6c), another mode of protein-ligand motion possibly involved in ligand binding and release, did not result in ligand release. As a measure for the barriers to ligand exit, the forces along the first EV were compared for the front and back exit as shown in Fig. 5.7b. Unfortunately, as can be seen in the Figure, the large fluctuations render it difficult to infer any significant preference for one of the release pathways.

To resolve this problem, further MD simulations were carried out for fixed

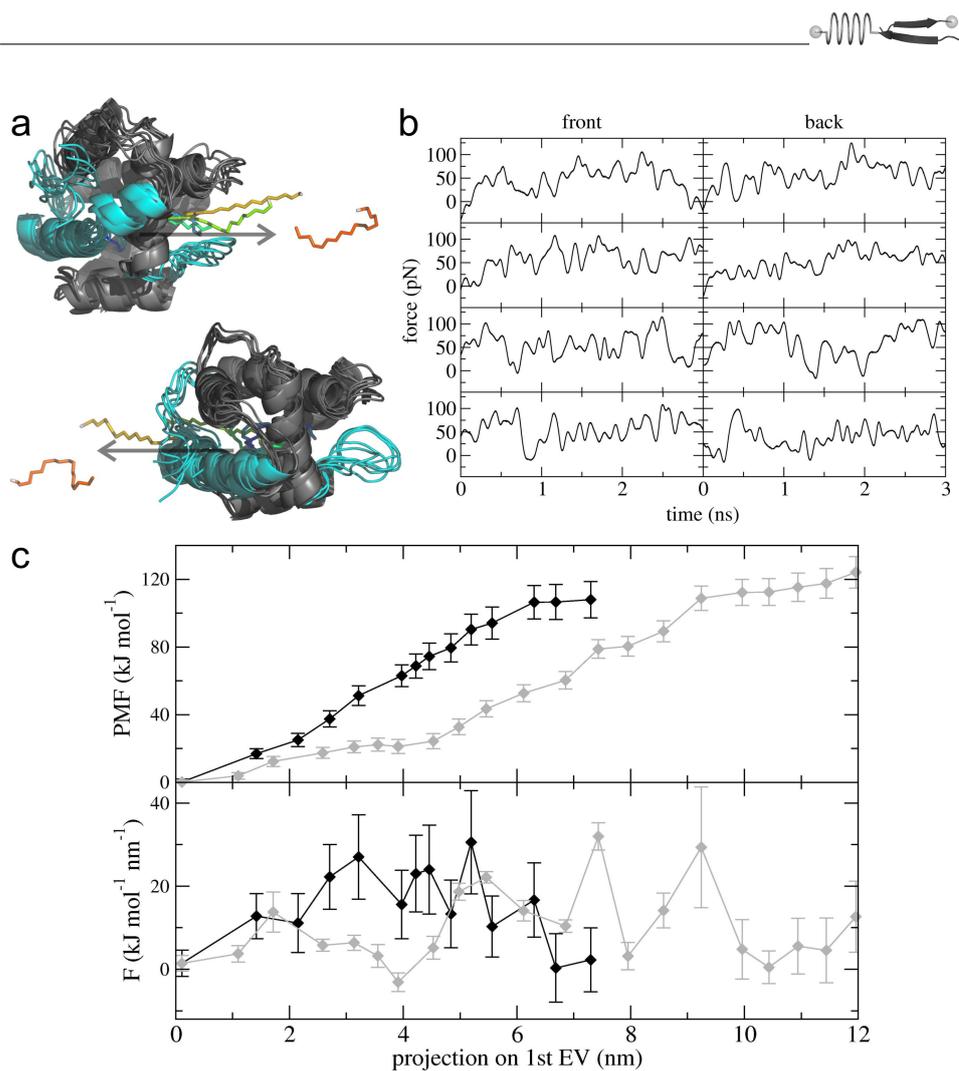


Figure 5.7: Ligand release probed by ED simulations. (a) Snapshots during dissociation to the front (top) and the back (bottom) from one representative trajectory. Colored: bombykol, cyan: front loop and terminal helix and coil. (b) Forces along the first EV to the front and back during four ED runs each. (c) Extended sampling of forces at discrete positions along the first EV. PMF (upper panel) and forces (lower panel) along the dissociation towards the front (black) and back (gray), averaged over the last 2 ns of the sampling. Error bars indicate the standard deviation of averages over 0.2 ns windows.

reaction coordinates along the two exit pathways. On the basis of the obtained forces and potentials of mean force (PMF), shown in Fig. 5.7c, the exit of bombykol along the front and back lid is equally likely. Both pathways exhibit a rupture force of  $\sim 30 \pm 12 \text{ kJ mol}^{-1} \text{ nm}^{-1}$  (lower panel). The derived activation free energies of  $108$  and  $124 \pm 12 \text{ kJ mol}^{-1}$ , respectively, for front and back release were also found to agree within the expected stan-

dard deviation (upper panel). Thus, none of the two sites is preferred over the other, and hence can both serve as passage ways for bombykol release.

We note that the PMF overestimates the experimental dissociation free energy ( $42 \text{ kJ mol}^{-1}$ ) by more than a factor of two, likely due to insufficient sampling and non-equilibrium effects. In particular, the ensemble of the unbound ligand and protein is typically larger and consequently sampled to a smaller extent than the bound state. Here, the accurate calculation of the dissociation free energy was presumably specifically complicated by the high degree of ligand conformational flexibility. Furthermore, the protein might be enforced to undergo unfavorable conformational changes to permit ligand release that are irreversible on the time scale of the simulations, and therefore misleadingly add to the free energy difference. A non-optimal reaction coordinate might be another reason for the present overestimation of the PMF. However, we here aimed at the comparison of the two putative pathways. Since the forces and free energies can be assumed to involve errors of the same magnitude for both pathways (see below), the difference can be expected to be significantly more accurate than the absolute values.

The convergence of the PMF during the accumulation of forces were examined as shown in Fig. 5.8. As before, sampling is particularly problematic due to the fluctuation of the forces, which, on a ps time scale, were about an order of magnitude larger than the forces itself. Reliable estimates therefore required averaging over windows of several nanoseconds in length, of which the convergence was then examined. To this end, for each of the point along the first eigenvector, forces were calculated from averaging over 0.2 ns windows at different positions along the ED trajectories, and the resulting PMFs were compared (Fig. 5.8a). The PMF initially decreased indicating equilibration of the system, and after several ns somewhat levelled off, though associated with considerable fluctuation at this window length. In addition, the dependence of the PMF on the window length was tested (Fig. 5.8b). Within a range of window lengths of several ns, as a compromise between high fluctuations of the forces using shorter times, and inclusion of unequilibrated ED frames when using longer times, the PMF was reasonably robust.

As expected, since the essential dynamics rest upon the principal modes identified in the free dynamics of the bombykol-BmPBP complex, the ligand was forced to proceed along those pathways to dissociate, at which the bound state dynamics already hinted at. More specifically, ligand release towards the front and back in all ED simulations involves conformations similar to the one shown in blue in Fig. 5.6a and b.

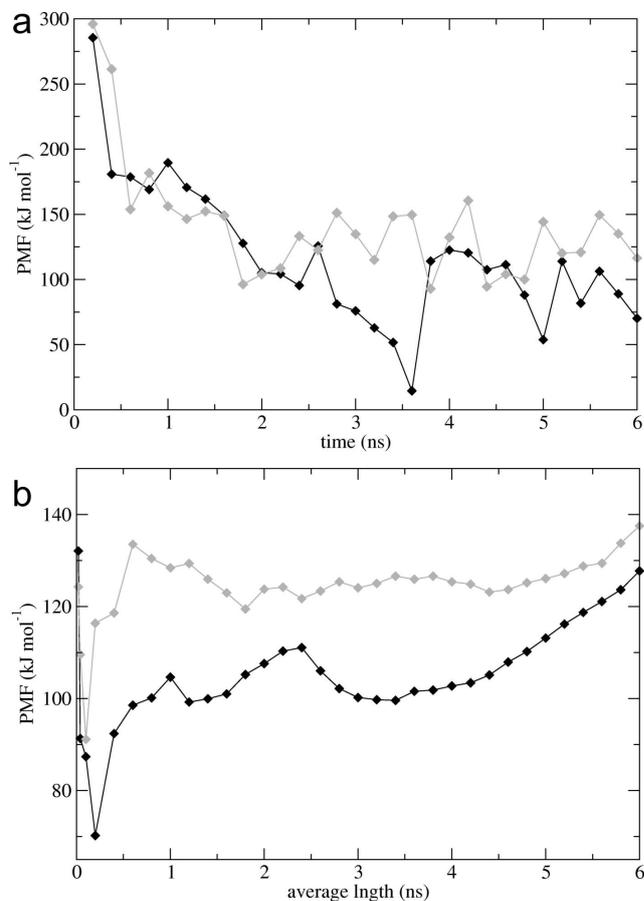


Figure 5.8: Convergence of the PMF derived from ED simulations towards the front (black curve) and back (gray curve) with extended sampling. (a) PMF as a function of the time at which the forces for PMF calculation were collected (in 0.2 ns windows). (b) PMF as a function of the length of the window to collect forces. The last part of the trajectories is chosen in each case.

### 5.3.3 Ligand release pathways from force probe simulations

To probe a larger variety of ligand release pathways and to reduce the bias introduced above by sampling along the first eigenvector, as a third approach, bombykol was pulled out of the cavity of BmPBP by means of a pulling force in a predefined direction. Here, in contrast to the ED simulations above, the protein was completely free to adapt to the exit motion of the ligand. In a first series of FPMD simulations, four different bombykol-BmPBP conformations observed during the conventional MD simulations served as starting structures. The pulling directions were chosen such that the center of mass of bombykol was pulled directly towards the approximate

center of the putative lid at the front and of the N and C-terminal helix and tail at the back, respectively. These directions are shown as solid arrows in Fig. 5.9.

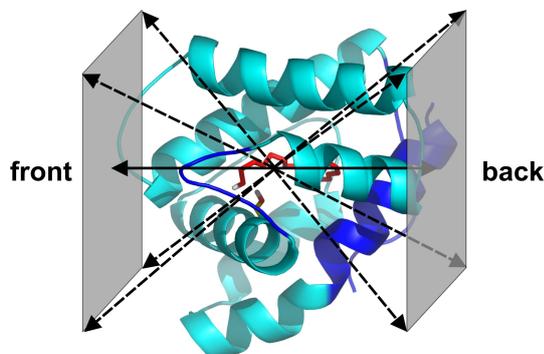


Figure 5.9: Pulling directions used in FPMD simulations to enforce ligand release.

Fig. 5.10a compares the force profiles resulting from FPMD simulations of the four starting structures. The respective intermediate conformations at the force maximum are shown in Fig. 5.10b. For pulling towards the front, irrespective of the initial complex conformation, unbinding occurred via a transient Glu98-BOM hydrogen bond. Formation of this bombykol-BmPBP hydrogen bond was accompanied by the rupture of the intramolecular Glu98-Leu68 hydrogen bond (Fig. 5.10b and c), which clamps the front loop in complexed and vacant BmPBP [158,162]. The front loop thereby transiently gained flexibility, allowed ligand passage and finally is fixed again upon re-formation of the Glu98-Leu68 hydrogen bond. The semi-dissociated state with the polar head of bombykol pointing partially out of the cavity has been also observed in the conventional MD and REMD simulations (Figs. 5.4a and 5.6). This state remained stable up to a force of  $\sim 300$  pN. It is the final detachment of the hydrophobic ligand from the cavity, in particular from Trp110, which is directly located at the lid, and its solvation with water that gave rise to the sudden force drop.

For bombykol release along the front lid, the mechanism and the involved forces are consistent with the findings obtained from sampling along the first EV using ED (see above). Combining the above results, we suggest that for front exit — if it occurs — the pathway is well defined. In contrast, unbinding trajectories along the back pathway were found to be more diverse, also pointing towards a larger entropic contribution to the unbinding free energy barrier [127]. The different starting structures yielded a variety of exit pathways near the termini, even though bombykol was pulled towards

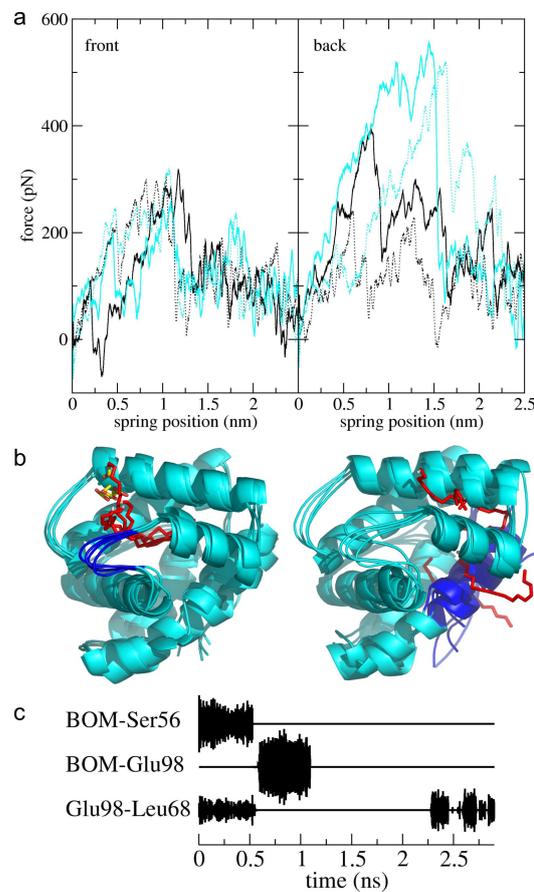


Figure 5.10: Ligand release probed by FPMD simulations. (a) Force profiles from pulling towards the front lid (left panel) and the termini at the back (right panel) along the main pulling direction (solid arrows in (a)), obtained from four different trajectories each. (b) Snapshots at which forces are maximal during force-induced unbinding along the main pulling direction. Pulling towards the front yields similar rupture forces and unbinding trajectories that have a transient hydrogen bond to Glu98 in common (left), pulling towards the back yields diverging forces and pathways (right). Red: bombykol, yellow: Glu98, blue: regions involved in ligand passage, the front lid and the terminal fractions. (c) Hydrogen bond ruptures involving the front loop (Glu98 and Leu68) during ligand release to the front.

the same direction. Since force peaks for backwards unbinding scattered within  $\sim 100$  pN, they did not allow for discriminating between unbinding towards the front and back on the basis of these few trajectories.

In order to elucidate the dependance of the trajectories and forces on the chosen pulling direction, for both putative exits, a second series of FPMD simulations using the aforementioned snapshots and, in addition, five differ-

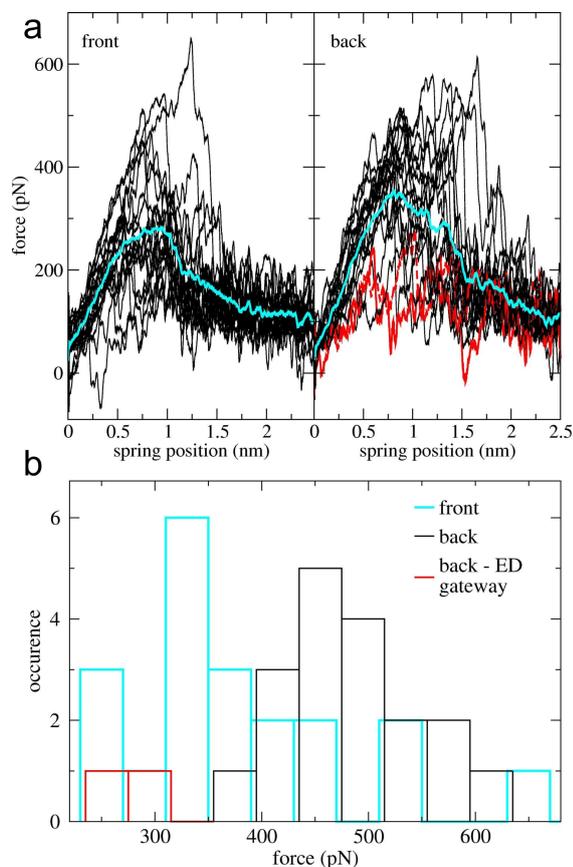


Figure 5.11: Ligand release probed by FPMD simulations along multiple directions. (a) Force profiles from pulling towards the front (left panel) and back (right panel), each along five different directions (solid and dashed arrows in Fig. 5.9). The average over all twenty profiles is plotted in gray. The two profiles with minimal rupture forces for back release are colored red. (b) Histogram of rupture forces observed in FPMD simulations.

ent pulling directions (dashed and solid arrows in Fig. 5.9) were performed. Fig. 5.11a compares the force profiles for the obtained ensemble of trajectories. For the release to the front and back, rupture forces scattered significantly (black lines in Fig. 5.11a), and yield a slightly lower average for release towards the front exit (cyan lines).

As before, front release exhibited a uniform mechanism in all cases. Independent from the exact pulling direction, it occurred along the front lid as described above. Rupture forces far beyond 300 pN were observed in those cases which required a relative rotation of protein and ligand to superimpose the ligand and the lid along the pulling vector. An unfeasible definition of



the pulling direction gives rise to additional friction, an artifact leading to an overestimation of the forces.

For back release, each of the five directions yielded a range of rupture forces and different release mechanisms. However, for two of them strikingly lower forces were sufficient (Fig. 5.11a, right panel). Interestingly, the underlying exit passageway of these two — and only these two — candidates was in accordance with the pathway identified in the ED simulations. This pathway involved a turn of the bombykol molecule within the cavity, resulting in the formation of a transient bombykol-Ser9 hydrogen bond. The flexible C-terminal coil gave way for final ligand release. Trp37 here played a crucial role as it constituted the endpoint of the hydrophobic bombykol-protein interaction during dissociation. The significantly lower rupture force and the consistency with the findings from ED sampling suggest that, among the ligand release trajectories towards the back, this pathway is favored by bombykol.

Figure 5.11b shows a histogram with the rupture forces for bombykol dissociation to the front and back exit binned into 40 pN intervals. We found that on average, front pulling ( $\sim 320$  pN) required slightly lower forces than back pulling ( $\sim 450$  pN). However, as discussed above, only part of the front and back pathway ensembles can be taken into account for this comparison, namely the two trajectories of back release with lowest rupture forces (Fig. 5.11b, red) and those trajectories of front release with rupture forces of  $\sim 300$  pN and below. Thus, a similar range of rupture forces was covered by these representatives.

The obtained work associated with ligand release to the front and back is compared in Fig. 5.12. The two back release trajectories with lowest rupture forces (right panel, red) arising from ligand dissociation in accordance with the ED simulation yielded work profiles similar in range to the work profiles obtained for front release (left panel). We again conclude that none of the two exits, front or back, can be excluded. We note that the calculated work inherently comprises non-equilibrium effects and hence is, as expected, in average larger than the previously estimated PMF (Fig. 5.7c) and the free energies derived from the work (see below).

A very recent and quite unexpected result by Jarzynski permits to calculate the PMF even from an ensemble of non-equilibrium work profiles, provided sufficient sampling is available (see Section 2.5.1 for details) [86,87]. To allow the application of this result, we performed a third set of FPMD simulations, now exclusively applying the previously identified prevalent pulling direction

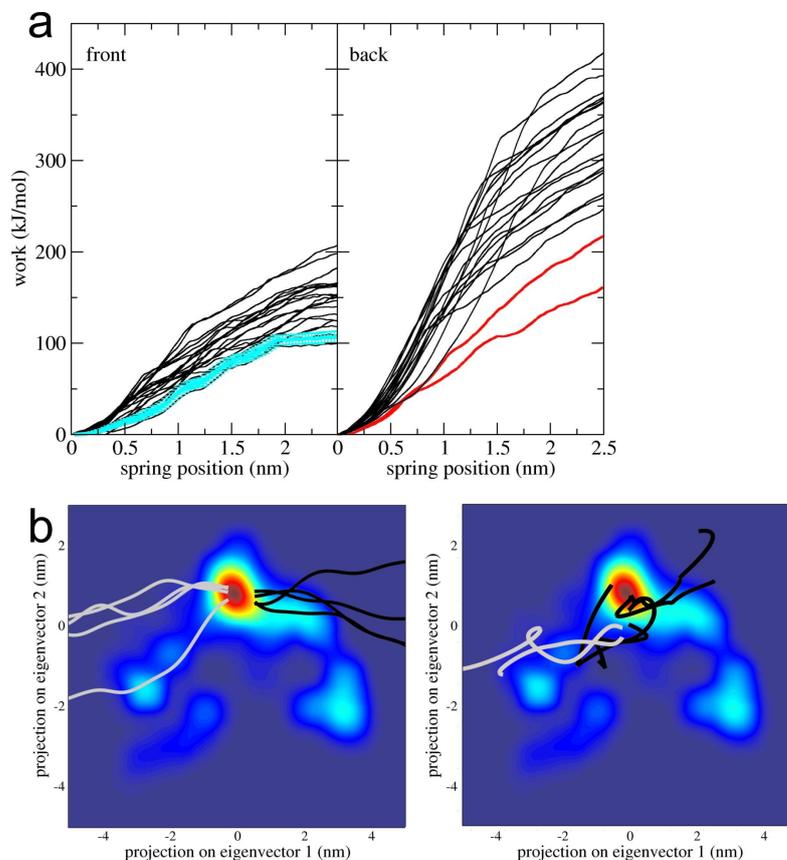


Figure 5.12: (a) Work and free energies associated with ligand release towards the front and back. Black: work, cyan: free energy profile calculated with the Jarzynski equation, red: preferred bombykol release pathways to the back (as in Fig. 5.11). The work was obtained from release towards the front along a single preferred direction, towards the back along different pulling directions and pathways. (b) Comparison of protein/ligand motion during bombykol release obtained from ED (left panel) and FPMD (right panel). Trajectories were projected onto the EVs from REMD simulations (Fig. 5.6). Gray: back release, black: front release.

(solid arrows in Fig. 5.9). Since back release was not controllable by the pulling direction, and thus a prevalent direction could not be identified, the PMF could not be assessed for the back exit. As can be seen from the PMF derived for front release (cyan in Fig. 5.12a), a free energy profile similar to the PMF from the ED simulation is obtained (Fig. 5.7c), again with an overestimation of the dissociation free energy. The results for the PMF using cumulant expansion [88,185] (data not shown) is highly sensitive towards the number of included trajectories, suggesting that the free energy estimate indeed is not yet sufficiently converged. A successful application



of the Jarzynski equality would require larger ensembles of ligand release trajectories or smaller pulling velocities to reduce non-equilibrium effects or both.

In summary, for front exit, the ligand release simulations suggested a unique mechanism along the putative front lid, for which indications were found also in the free dynamics of bombykol-BmPBP bound state. A half-dissociated state with a transient hydrogen bond from bombykol to BmPBP and a detached front loop represents an intermediate. Among the numerous pathways obtained for back release, a gate framed by the two termini could be identified as the most likely exit. Trp110 and Trp37, respectively, were identified as crucial residues stabilizing bombykol at the front and back exit just before its final release into the solvent. They thus can be assumed to lower the energy barrier for ligand binding and unbinding. Together with the residues forming transient hydrogen bonds to bombykol during release to the front and back, namely Glu98 and Ser9, respectively, these residues are promising candidates for point mutations with the objective of testing our predictions and further exploring the two pathways.

Even though the ED and FPMD simulations are conceptually different, they share the course of the forces, PMFs and the associated ligand/protein dynamics of the two identified exit gateways. This is also reflected by the projection of dissociation trajectories onto the first two EVs from REMD, as shown in Fig. 5.12b. The motion along the first EV, as enforced in ED simulations (left panel), is reproduced in the pulling simulations for the back exit and to a certain extent for the front exit (right panel). In view of the significant contribution of protein fluctuations to the first EV, which were unbiased during the FPMD simulations, this is an astonishing agreement. The two different simulation methods to probe unbinding do not only have the main ligand direction in common, which obviously is similar in ED and FPMD, but also the underlying protein main motion, and, thus, predict similar exit pathways. In the FPMD simulations, the front lid is detached from the neighboring helix only reversibly to temporarily give way to bombykol dissociation. Therefore, trajectories for front unbinding show only a transient motion to higher values of the first EV. The overall consistent picture provides confidence for the identified opening lids, namely the front His-rich loop and the terminal helix and coil with the respective intermediate states. Our results suggest that in fact both exit pathways may be physiologically relevant.

### 5.3.4 Bombykal versus bombykol binding

The question to what extent PBPs contribute to the selectivity in pheromone reception was addressed here by comparing the binding of bombykol and bombykal to BmPBP. These two components of the pheromone gland only differ in their polar head (Fig. 5.13a). The carbonyl group of bombykal, in contrast to bombykol, can interact only as hydrogen bond acceptor with BmPBP. This, *e.g.*, excludes anionic side chains such as Glu98 from interacting with bombykal. Do the polar interactions of bombykal, and consequently also its binding mode and position in the pocket, for this reason differ from bombykol? Figure 5.13b shows the RMSD and hydrogen bond partners of bombykal during free molecular dynamics starting from two different equilibrated BmPBP-bombykol structures with a Ser56-bombykol hydrogen bond.

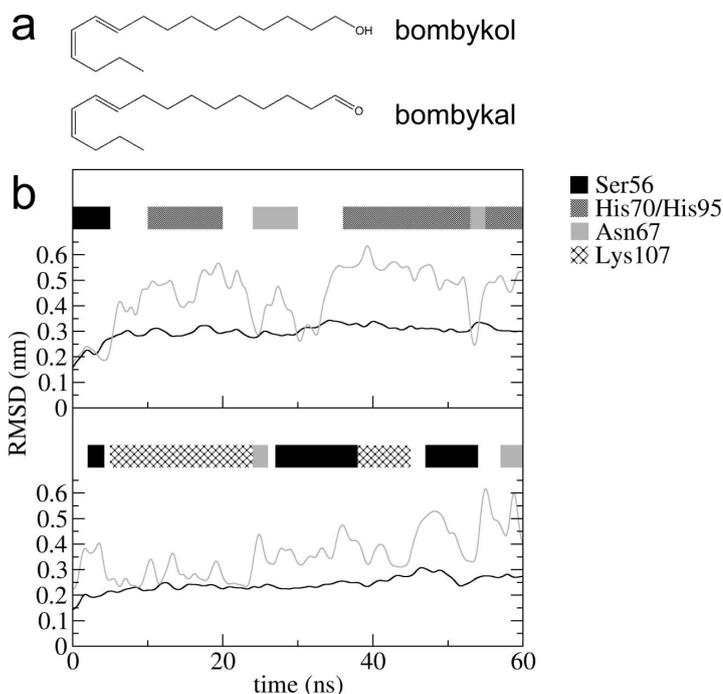


Figure 5.13: Bombykal-BmPBP free dynamics. (a) Schematic drawing of bombykol and bombykal, both components of the *Bombyx mori* pheromone blend. (b) RMSD of the protein backbone atoms (black curve) and bombykal (gray curve) from the bombykol-BmPBP crystal structure during two molecular dynamics simulations starting from different ligand conformations. Bombykal-PBP hydrogen bonds are indicated above.

Both simulations show a stable protein structure (black) and for the ligand (gray) significant deviations from the Ser56-bound initial structure, which



are similar in magnitude and fluctuations to those observed for BmPBP-bombykol (Fig. 5.3). Bombykal reversibly forms and breaks hydrogen bonds to Ser56, as bombykol, and also to His70, His95, Asn67 and Lys107. Thus, bombykal is able to form polar interactions with BmPBP during the bulk of the free dynamics, albeit under formation of hydrogen bonds to other side chains due to the different chemical preference of the carbonyl group. As a consequence of the similar ligand position within the cavity, hydrophobic protein interactions were found to be similar for bombykol and bombykal.

These results indicate similar binding modes of bombykol and bombykal in the cavity of BmPBP. Since bombykal only functions as a hydrogen bond acceptor, bombykal-BmPBP interactions can be expected to be to some extent less favorable. The same, however, also holds true for the ligand-water interactions in the unbound state, thus probably compensating for the observed difference in the bound state.

To quantitatively assess the difference in binding affinity, FEP calculations were performed (Fig. 5.14). In these simulations, bombykol in the bound and unbound state was slowly converted to bombykal, as described in Section 2.5.3. According to the thermodynamic cycle shown in Fig. 5.14a, the associated free energy changes  $\Delta G_{\text{FEP}}(\text{prot})$  and  $\Delta G_{\text{FEP}}(\text{aq})$  yield the difference in binding free energies,  $\Delta\Delta G_{\text{b}}$ , which are given in Fig. 5.14b. Simulations of bombykol perturbed to bombykal (and vice versa) in the unbound state gave a free energy change  $\Delta G_{\text{FEP}}(\text{aq})$  of 9.5 kJ/mol for a simulation time of 20 ns. With a standard deviation of 0.2 kJ/mol (0.4 kJ/mol for 2 ns simulations), the FEP simulations were found to converge well for the ligand in water. As expected, for bombykal a less favorable solvation free energy than its hydroxyl counterpart was found, corresponding to its limited ability to form hydrogen bonds with water.

The change in free energy during FEP calculations of bombykol to bombykal within the binding pocket of BmPBP,  $\Delta G_{\text{FEP}}(\text{prot})$ , was  $3.9 \pm 1.5$  kJ/mol (5 ns simulation;  $3.8 \pm 2.7$  for 1 ns simulation). Sampling of the BmPBP-bombykol and BmPBP-bombykal conformations during the FEP simulations did not converge, as reflected by the deviation of free energy changes for different starting structures. With regard to the time scale of several nanosecond for hydrogen bond rupture and formation during the free dynamics of bombykol and bombykal (Fig. 5.3 and 5.13), longer simulation times would be necessary to sufficiently sample the conformational space of the bound ligand for a more robust free energy estimate. Here, however, an accuracy of 1.5 kJ/mol for  $\Delta G_{\text{FEP}}(\text{prot})$  suffices for the comparison of the binding free energies of bombykol and bombykal.

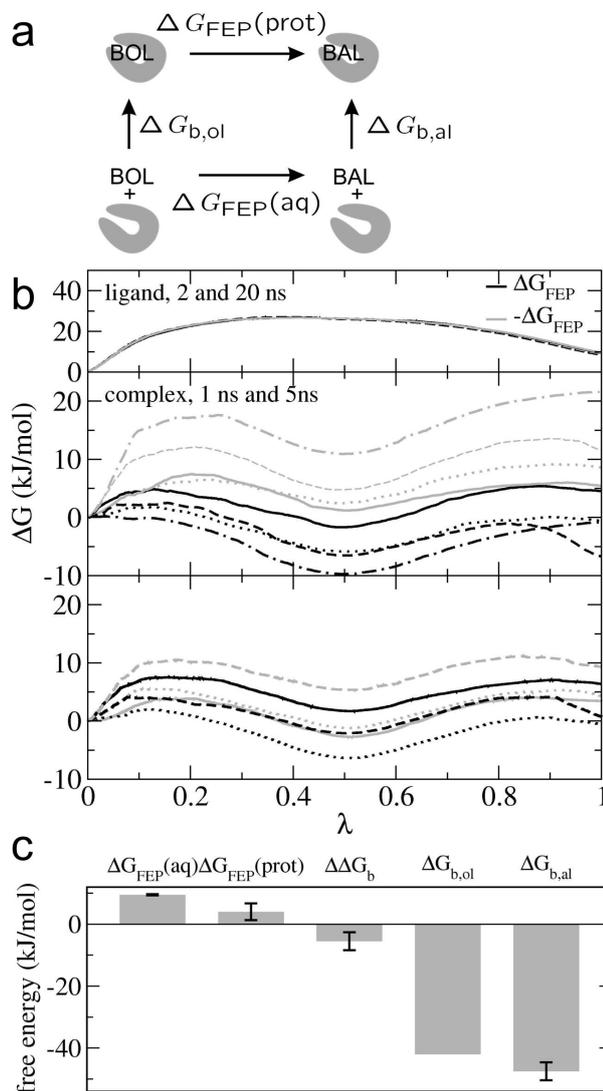


Figure 5.14: Calculation of the relative binding free energy of bombykol and bombykal to BmPBP. (a) Thermodynamic cycle (see Methods). The relative binding free energy is calculated as  $\Delta \Delta G_{\text{b}} = \Delta G_{\text{b,al}} - \Delta G_{\text{b,ol}} = \Delta G_{\text{FEP}}(\text{prot}) - \Delta G_{\text{FEP}}(\text{aq})$ . (b) Free energy change upon 'mutating' bombykol to bombykal (black) and back (gray) as obtained from FEP calculations. Results for the simulations of the ligand in water (upper panel), and the ligand bound to the protein (central and lower panel, for 1 ns and 5 ns simulation length, respectively) are shown. Identical line patterns refer to the same cycle of forward and backward simulation. (c) Calculated free energy changes as obtained from FEP simulations (compare (b)) and the resulting binding free energy of bombykal with respect to the experimental binding free energy of bombykol.



The positive free energy change during the FEP simulations of the bound state suggested a weaker interaction of bombykal to BmPBP than bombykol. This is consistent with the weaker interactions to the binding pocket observed already for the free bombykal-BmPBP dynamics. Since this effect was even stronger for the ligand in the unbound state, the effective binding free energy of bombykal to BmPBP is lower (Fig. 5.14c). In other words, the effect that bombykal is less well solvated in water compensates for the less favorable ligand-protein interactions comparing to bombykol, ultimately resulting in a more favorable bombykal-BmPBP binding.

According to our FEP calculations, the difference in binding free energies for bombykal and bombykol is  $-5.6 \pm 1.7$  kJ/mol. With a binding free energy of -40 kJ/mol for bombykol [159], this gives -45.6 kJ/mol for bombykal binding (Fig 5.14c). It should be noted that, apart from the statistical error discussed above, uncertainties in the free energies also may arise from inaccuracies in the ligand force field parameters. Furthermore, the bombykal-BmPBP starting structure does not rest on experimental data, but was derived from simulated equilibrium conformations of bombykol-BmPBP. For these reasons, our result of a slightly higher affinity of BmPBP for bombykal than for bombykol should be regarded with caution, as the difference was found to be small (only few  $k_B T$ ).

Summing up, the results suggest that bombykal is favorably bound by BmPBP, with an affinity similar to or even higher than bombykol. Thus, the subtle chemical difference of the female pheromone blend compounds bombykol and bombykal is most likely not sensed by BmPBP. That each pheromone component excites a separate cell type must, therefore, be attributed to the specificity of the respective pheromone receptor molecules. Validation by experiments is yet to be accomplished.

## 5.4 Conclusions and Outlook

The function of the olfactory protein BmPBP as a carrier for its physiological binding partner, the pheromone bombykol, is closely linked to the mechanism and underlying energetics of bombykol uptake and release. Our simulations identified two ligand release gateways, the front His-rich lid and a gateway framed by the terminal helix and coil. Both, forces and free energies involved in ligand release indicate that the two pathways are equally likely. Within an uncertainty in the relative free energy barriers of a few  $k_B T$ , we thus suggest that both pathways are indeed physiologically rele-

vant. This conclusion rests on two different simulation techniques to enforce unbinding, ED and FPMD, which differ fundamentally in the reaction coordinate along which unbinding is enforced, here an intrinsic mode of motion or an externally defined direction. The quantitative agreement of the mechanism and energetics obtained with the different approaches underscores our conclusion of two relevant ligand release pathways.

Our simulations were challenged by the high conformational flexibility of the bound and unbound ligand. No preference of conformational states or pathways could be inferred from single trajectories and forces, which highly depended on the starting configuration and velocities. Due to the unspecific character of binding, extensive sampling was required in all cases. Besides its high biological relevance, the bombykol-BmPBP system thus also represents a demanding, yet intriguing test system for the calculation of binding free energies.

What might be the physiological purpose of two opposite pathways? The requirements for PBPs are in fact manifold. Their function as a carrier requires pheromone binding to be (i) tight to protect against degrading proteins and to achieve a high pheromone sensitivity, (ii) selective for serving as a pre-filter with certain specificity in addition to the receptor, and (iii) sufficiently fast (on a millisecond time scale) such that the kinetics of olfactory reception allows for a fast adjustment of the flight direction of the moth. Encapsulation of the whole ligand by an apparently impermeable protein envelope supports tight and specific binding, but sterically impedes ligand entrance and exit by severely restricting the access to the cavity comparing to solvent exposed binding pockets. This affects particularly ligand uptake, which requires bombykol diffusion towards the entrance lid. BmPBP exhibits a more or less uniformly negative electrostatic potential (data not shown) and a net charge of -8 at pH 7. Diffusion of bombykol, which is virtually apolar, towards its receptor BmPBP hence is undoubtedly not directed by electrostatic funneling towards the entrance, as found for acetylcholine esterase [186], and eventually represents the rate-limiting step.

We therefore speculate that the physiological significance of two lids is to ensure a sufficiently fast rate of ligand uptake. By providing two pathways for bombykol entrance, the binding free energy barrier is lowered, and the area of access for bombykol to the BmPBP cavity by random diffusion is enlarged. Electrostatic funneling, another way to ensure fast ligand binding and unbinding, enables acetylcholine esterase to rapidly take up and release *different* molecules, namely the substrate and products, *via* two opposite one-way gates [186]. Here, in the case of a carrier protein, fast binding and



unbinding of the *same* molecule, instead, is guaranteed by two nondirectional pathways.

Furthermore, two alternative pathways warrant robustness of the binding properties of BmPBP against point mutations that eliminate one of the routes. The previous finding that mutating Trp37, a residue located at the back gateway, leaves the binding energetics and kinetics unchanged, was unexpected [159], since this residue, conserved among PBPs, is thought to be physiologically important. Now, in the light of the two-pathway model, this robust behavior of BmPBP towards mutation at only one of the pathways is explained.

In conclusion, the identified two discrete pathways to a central cavity provide the moth with the advantage of facilitating ligand entrance and exit without the loss of binding affinity and selectivity. This hypothesis can in principle be validated experimentally by point mutations. Potential candidates, as discussed above, are mutations of Trp37 and Trp110 to, *e. g.*, alanine, which are expected to eliminate the stabilizing influence of the Trp-bombykol hydrophobic interactions onto the half-dissociated intermediate state. Alternatively, a complete steric blockage of the gates might be achieved by fixing the opening lids of the exit gates by disulphide bonds. The temporary hydrogen bond partners of bombykol during dissociation, Glu98 at the front and Ser9 at the back, together with an appropriate proximal bonding partner are also suitable for this purpose. The fourfold cysteine mutation should result in a substantial hindrance of bombykol uptake.

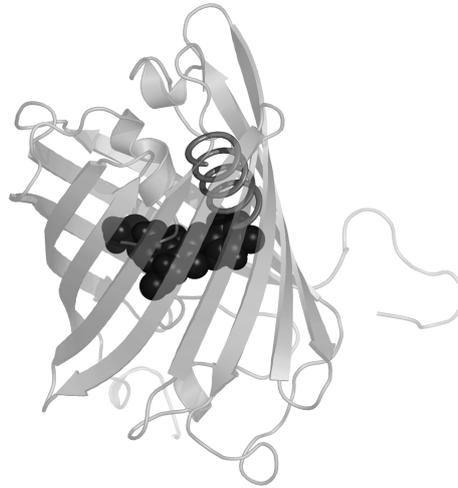
If our model of two relevant pathways is correct, one would expect a substantial change of the binding affinity and kinetics only if *both* pathways are eliminated by point mutations. A mutation at only *one* of the exits, instead, should not significantly alter the binding kinetics and energetics. Together, this suggests a harsh test for our two-pathway model. The proposed mutants remain to be tested in simulations and experiment.

We here aimed at elucidating the pathway for bombykol release at neutral pH, at which bombykol is reversibly taken up and released on a millisecond time scale [159]. A related open question is how BmPBP undergoes the conformational transition towards its acidic low-affinity structure and at what point this transition induces ligand release. This will be addressed in further investigations.

Pheromones for sexual attraction differ within insect species. Excitation by only those pheromone components constituting the pheromone blend of the own species requires highly selective reception. Is this achieved only by

the pheromone receptor or also by peri-receptor events? Our simulations depict BmPBP as a versatile carrier that can accommodate a certain subset of ligands. A distinct but not extreme binding specificity of BmPBP agrees with the scenario of PBPs acting as a pre-filter [187]. According to this scenario, BmPBP filters out a subset of components by protecting them from degradation and carrying them to the receptor. It is the combined BmPBP *and* receptor specificity which, in two sequential steps, yields the ultra-high selectivity of the olfactory receptor cell.

The discrimination of molecules in the two-step filtering by BmPBP and the receptor was examined here for bombykol and bombykal. According to our simulation results, BmPBP does not differentiate significantly between the hydroxyl and carbonyl group, but delivers both bombykol and bombykal to the receptor. The pheromone receptor as the second filter yields the selective power required for the excitation by only a single specific pheromone component. Concerning the physiological function of two layers of filters, one can speculate that their interplay not only significantly increases specificity, but also facilitates its fine-tuning and control. *E. g.*, to encode the required specificity into the different olfactory proteins across insect species during evolution, two filters can be individually subjected to modifications to tailor their affinities. Calculating relative free energies proved useful to assess the filtering function of the first filter, BmPBP, for bombykol and bombykal. Extending the calculations to other pheromone components of *Bombyx mori* will yield a more complete picture of the pheromone discriminative role of BmPBP.



## 6

# Photoswitch mechanism of the fluorescent protein asFP595

---

### 6.1 The fluorescent protein asFP595

Nature has equipped certain organisms of the deep sea with fluorescent proteins, of which the green fluorescent protein (GFP) of the jellyfish is the most widely known [188, 189]. GFP and its homologues feature a  $\beta$ -barrel fold, the center of which is occupied by the chromophore, the light absorbing and emitting moiety (Fig. 6.1a). GFP-like proteins are widely used in cell biology as photolabels to track protein expression, protein trafficking, and protein-protein interactions *in vivo*. Their fluorescence spectrum and kinetics are defined by the spectroscopic properties of the chromophore and the proximal protein scaffold and can thus be selectively tuned by protein engineering.

More recently, chromophoric proteins have been discovered that can be

## 6. PHOTOSWITCH MECHANISM

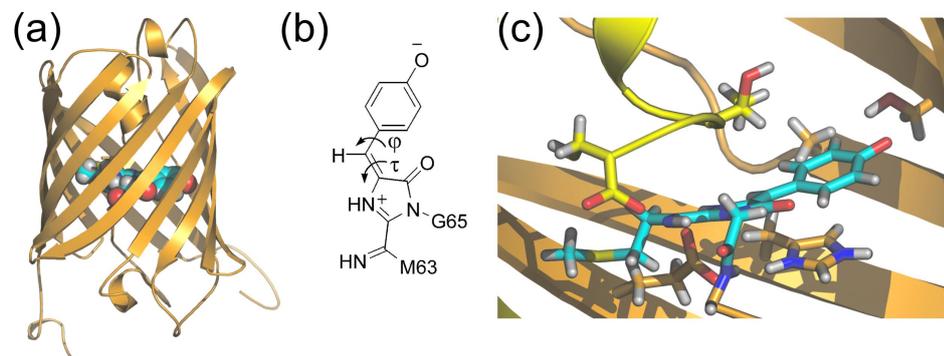


Figure 6.1: The asFP595 protein structure. (a) asFP595 (pdb entry 2A50) [190]. The chromophore is shown as spheres. (b) MYG chromophore and relevant dihedrals. (c) Close-up view of asFP595. The chromophore and important interacting protein residues are shown as sticks. Hydrogen bonds are formed from MYG to Ser, Arg, Glu, Cys. Colors: cyan - MYG, asFP595 chain A - yellow, chain B - orange.

switched. Upon light absorption, their fluorescent state can be reversibly turned on and off [191–194]. A prominent example is asFP595, the fluorescent protein of the sea anemone *Anemonia sulcata* [191]. It is switched on into the fluorescent (kindling) state by absorbing light with an excitation maximum at 575 nm. The kindling state fluoresces with a maximum at 595 nm, and undergoes relaxation to the off-state with a life time in the order of seconds. It can also be switched off instantly with blue light. Reversible photoswitching from an off- (dark) to an on- (fluorescent) state is of great interest with regard to the application of switchable chromoproteins as tags for *in vivo* fluorescent microscopy with improved sub-wavelength resolutions in the nanometer scale [195].

The asFP595 structure has recently been solved by X-ray crystallography [190, 196], and is shown in Fig. 6.1. As can be seen, asFP595 exhibits the typical GFP-like  $\beta$ -barrel fold. The covalently bound fluorophore MYG originates from the Met64-Tyr65-Gly66 sequence which is post-translationally modified to a 2-iminomethyl-5-(4-hydroxybenzylidene)imidazolinone (Fig. 6.1b) system by cyclization and oxidation [197, 198]. The chromophore formation also involves backbone rupture between the residues upstream of MYG.

MYG is similar to the GFP chromophore except that the conjugated  $\pi$ -system is extended by an additional imino group. It should be noted that the replacement of the imino group by a carbonyl group subsequent to the cyclization could not be excluded on the basis of the crystallographic data, and indeed was suggested by more recent studies on the MYG chromophore in solution [198]. However, the impact of the chemical nature of this moiety



on the isomerization mechanism can be assumed to be small. As shown in Fig. 6.1c, it is firmly positioned in the central cavity of the  $\beta$ -barrel by forming an extensive hydrogen bond network and distinct hydrophobic interactions with nearby protein side chains.

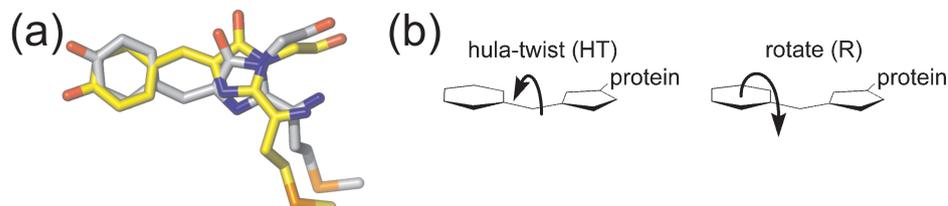


Figure 6.2: The photoswitch involves trans-cis-isomerization of the MYG chromophore (a) MYG chromophores from the A143S crystal structures before (cyan, off-state) and after irradiation (yellow, on-state), superpositioned after least-square fitting the chromophore atoms. (b) The two possible isomerization mechanisms, rotation R and hula-twist HT, described in the text.

The A143S mutant of asFP595 has a longer lifetime of the stabilized kindling state, which allowed to solve both the non-fluorescent and fluorescent structure by switching the mutant on and off within the same protein crystal [190]. As shown in Fig. 6.2a, switching is associated with a change in the chromophore structure. The off-state corresponds to a trans-configuration of MYG, the on-state to a cis-configuration. As revealed by the mutant and wt structure, the A143S mutation stabilizes the on-state by means of an additional hydrogen bond to the hydroxyl group of MYG.

According to the experimental structural and mutational data described above, the photoswitch to a fluorescent state upon light absorption involves a trans-cis isomerization. The detailed molecular mechanism that connects these two end-states, however, remains unknown. Two principally different mechanisms can be envisaged that lead to the cis conformer, namely rotation around both dihedrals of the ring-bridging CH-group,  $\tau$  and  $\phi$  (hula-twist, HT, Fig. 6.2b, left), or rotation around  $\tau$  only (R, Fig. 6.2b, right).

For the rotate mechanism, the MYG tyrosine ring can either rotate towards the initially coplanar His197 ( $R^{\text{top}}$ ) or towards the other side ( $R^{\text{bot}}$ ). Likewise, during HT isomerization, the bridging methylene group can move along a top ( $HT^{\text{top}}$ ) or bottom ( $HT^{\text{bot}}$ ) pathway.  $R^{\text{top}}$  and  $R^{\text{bot}}$  require a space-demanding rotation of the tyrosyl ring involving the transient loss of hydrophobic interactions, whereas  $HT^{\text{top}}$  and  $HT^{\text{bot}}$  leave the chromophore virtually in place and only involves movement of the central CH-bridge.

In the studies on titin kinase and BmPBP, FPMD simulations proved suit-

able to investigate the functional process of interest in these proteins, unfolding and unbinding, respectively. Also here, FPMD simulations render themselves useful to identify the preferred isomerization mechanism of the asFP595 photoswitch. The chromophore is subjected to a pulling force such that the isomerization along one of the above described pathways is enforced; and the forces opposing the reaction are compared. Thus, using FPMD simulations, we exclusively focus on the steric influence of the protein scaffold around the chromophore onto the enforced photoswitch process.

A second set of MD simulations studies the spontaneous photo-induced isomerization. The photo-excited state is accounted for by means of a suitably chosen force field with modified atomic velocities, charges and dihedral potentials. Both simulation techniques identify the bottom HT mechanism with the ring-bridging CH-group rotating towards His197 below the chromophore as the favored pathway.

## 6.2 Simulation details

All asFP595 simulations described here were carried out in close collaboration with Lars V. Schäfer. He also kindly provided the results from the quantum mechanical calculations discussed below. My focus was on the development of an appropriate chromophore force field, and the simulation of the isomerization reaction using FPMD and MD.

### 6.2.1 Chromophore force field

Partial charges of a model trans and cis chromophore in the ground and excited state were calculated quantum mechanically by Lars V. Schäfer using the CHELPG scheme of the Gaussian03 program [199]. The chromophore structures were geometry-optimized at the B3LYP/G-31+G\* level. The partial charges for the electronic ground state were calculated using the DFT wave function, whereas the excited state partial charges are based on a configuration interaction singles (CIS) calculation. Heavy-atom bond lengths and angles were taken from the X-ray structure, those involving hydrogen atoms from the QM-optimized structures. For the force constants and Lennard-Jones parameters, OPLS force field parameters of residues with similar chemical nature were adopted (His, Tyr, Met, styrene). Except for partial charges, the same parameters were used for the cis-state.



## 6.2.2 MD simulations

All simulations were carried out with the Gromacs 3.2.1 simulation suite [121]. The simulations were started from the 1.3 Å crystal structure of asFP595 (PDB code 2A50) [190]. Protonation states of the standard amino acids were calculated using the program WHATIF [123] and its interface to Delphi [124]. Because of its close proximity to MYG, the protonation state of His197 is crucial, and was predicted to be cationic. To discriminate between the possible chromophore protonation states, simulations of the neutral state (Imidazolinone-N, Tyr-OH), the anionic state (Imidazolinone-N, Tyr-O<sup>-</sup>), and the zwitterionic state (Imidazolinone-NH<sup>+</sup>, Tyr-O<sup>-</sup>) were performed, and the protein-chromophore interactions were monitored. All hydrogen bonds from the chromophore to the protein stayed intact only in case of the zwitterionic protonation state, pinpointing the zwitterion, henceforth used in all simulations, to be the correct chromophore protonation state in the asFP595 'off'-state.

The OPLS force field [63] was applied. A monomer of the protein was solvated in a 8.9 x 7.9 x 7.9 nm<sup>3</sup> box of TIP4P water molecules. 25 sodium and 28 chloride ions were added to the simulation system to compensate for the overall positive charge of the protein and to mimic physiological conditions. This yielded a total system size of 71734 atoms. Otherwise, simulation parameters were identical to those described for the simulation of titin kinase in Section 3.2.1. Initially, the system was energy-minimized (steepest descent, 1000 steps), before equilibrating the solvent for 200 ps with positional restraints on protein heavy atoms. Then, the whole system was equilibrated (0.1 ns at 300 K).

In subsequent simulations, trans-cis isomerization of the chromophore was induced by means of FPMD simulations (see Section 2.4.1) [43]. For this purpose, the FEP algorithm of the Gromacs simulation package was used. Simulation details were as described in Section 3.2.1, apart from the calculation of non-bonded interactions. Here, no PME, but a cut-off of 1.5 nm and 1.2 nm for electrostatic and Lennard-Jones interactions, respectively, was used. The R mechanism along the top or bottom path was enforced by shifting the minimum of the  $\tau$  dihedral potential by 180 degrees backward or forward, respectively. The minima of both dihedral potentials  $\tau$  and  $\phi$  were shifted simultaneously in equal measure to enforce the HT mechanism. To induce the isomerization at the short timescale of photo-induced processes (here 20 ps were chosen), a Ryckaert-Bellemans (RB) constant of 140kJ/mol for  $\tau$  (R mechanism) was found necessary.

To yield forces along the isomerization pathway that can be compared with each other, we applied an RB constant of 100kJ/mol for  $\tau$  and  $\phi$  to enforce HT, as can be inferred from geometric considerations. Dihedral potentials with the same central atoms as the perturbed ones were set to zero. Excited state partial charges were smoothly converted into their cis ground state values during the simulations. This prohibited the usage of PME, see above. The force was recorded at each integration step. The force values shown in Fig. 6.4 were calculated with respect to a semicircular coordinate with a radius of 0.5 nm, described by the idealized rotation of the center of the chromophore phenyl moiety upon isomerization. For the corresponding simulations in water, the chromophore model obtained from the quantum mechanical calculations was used, with the full Met63 residue included.

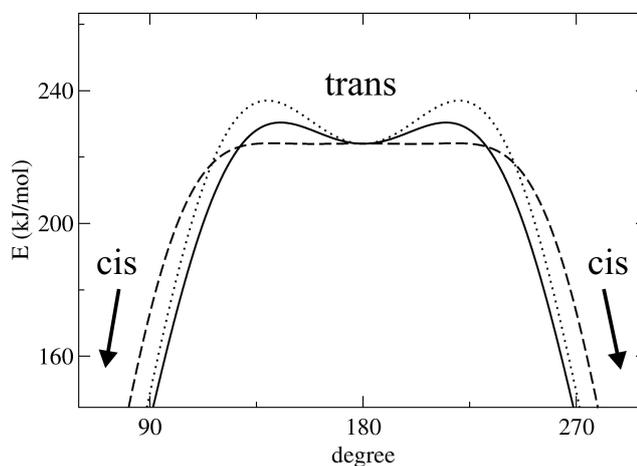


Figure 6.3: Dihedral potentials used for the free MD simulations. Solid line,  $\tau$ ; dashed line,  $\phi$ . The potential energy barriers of both dihedrals were varied between 0 (dashed) and 7 kcal/mol (dotted line) in test simulations.

To monitor spontaneous isomerization starting from a trans conformation in the excited state, free MD simulations have been carried out. The applied simulation protocol was similar to the one used for equilibration (see above), apart from the force field description of the chromophore. A semi-quantitative description of the transition from an excited-state potential through a conical intersection back to the ground state, as previously applied to rhodopsin [200], was developed and is shown in Fig. 6.3. In this model the transition of one of the dihedrals from the off-state configuration to a 90 degree twisted state involves a barrier of 8.4kJ/mol ( $\tau$ ) and 0 kJ/mol ( $\phi$ ), as revealed by quantum mechanical calculations of the similar GFP chromophore [201]. Further rotation to a planar dihedral conformation



was assumed to proceed downhill towards the ground state *via* a putative conical intersection or thermal deactivation. The model further comprised excited state CIS charges (see above), and an additional kinetic energy of the chromophore (21 kJ/mol [201]) to account for the relaxation from the Franck-Condon region after excitation.

## 6.3 Results

### 6.3.1 Enforced chromophore isomerization

To examine the — obstructive or supportive — influence of the protein matrix onto the chromophore during isomerization, FPMD simulations along the four conceivable pathways were performed, the results of which are shown in Fig. 6.4.

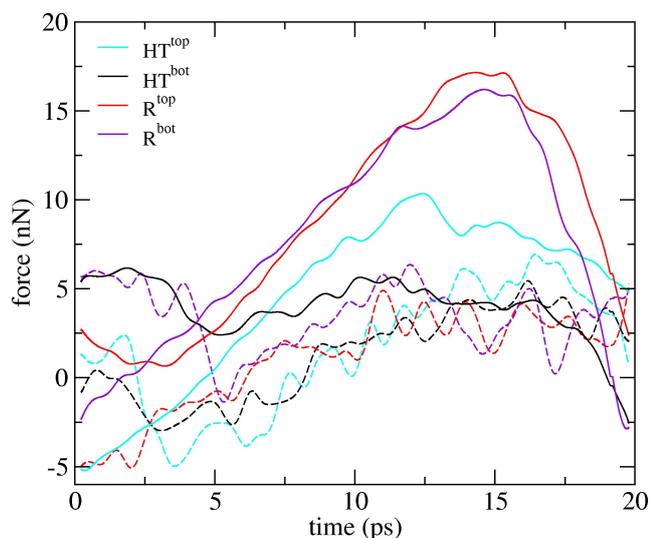


Figure 6.4: Forces due to the protein matrix opposing chromophore isomerization were calculated by non-equilibrium FPMD simulations and averaged over 10 trajectories for the four possible pathways (solid lines). Control simulations of the chromophore in water show similar forces for the R and HT mechanisms (dashed curves).

As expected, both rotation mechanisms,  $R^{\text{top}}$  and  $R^{\text{bot}}$ , are significantly more opposed by forces stemming from the protein matrix than the HT mechanism. The rotation of the tyrosyl ring is strongly opposed by the protein matrix, that solidly clasps the chromophore *via* numerous polar and non-polar interactions (compare Fig. 6.1c).

Residue-wise interaction energies of MYG with the protein matrix were calculated to identify residues giving rise to major energy barriers. In case of the isomerization along R, Thr60 and His197 showed a substantial increase in the interaction energy to MYG during isomerization, since their interaction to the chromophore is transiently disrupted during the rotation mechanism.



As shown in Fig. 6.4, the HT mechanism, in contrast, involved remarkably smaller forces as it impaired MYG-protein interactions only marginally. HT, as expected, thus is suggested to be less sterically hindered, as it comprises space-saving rotation of only the methine bridge *via* simultaneous rotation of both dihedrals  $\tau$  and  $\phi$ . Since HT does not require the rotation of the MYG tyrosyl moiety, the favorable  $\pi$ -stacking interaction with His197 is maintained throughout the HT isomerization. Concerning the direction of the rotation, the forces obtained from FPMD simulations suggest HT<sup>bot</sup> to be favored over HT<sup>top</sup>.

Here, since the simulations aimed at elucidating the role of the protein during trans-cis isomerization, intrinsic chromophore properties are purposefully not taken into account. Control simulations in water indeed show, as expected, no preference of any of the mechanisms on the basis of forces. In conclusion, our results suggest the space-saving HT, more specifically HT<sup>bot</sup>, as the switching mechanism favored by the protein matrix.

### 6.3.2 Spontaneous chromophore isomerization

To identify the preferred pathway with respect to the intrinsic chromophore properties, additional MD simulations of the spontaneous isomerization reaction were performed. To this end, the excited state chromophore was described using the semi-quantitative model specified in Section 6.2.2

With the additional kinetic energy of the chromophore stemming from relaxation from the Franck-Condon region, the chromophore explored the excited state surface in all degrees of freedom including the  $\tau$  and  $\phi$  dihedral angles. As shown in Fig. 6.5 (upper panel), within a time scale of 100 to 500 fs, the chromophore in the protein spontaneously isomerized by eventually overcoming both of the barriers. Each of the simulations showed a concerted rotation around  $\tau$  and  $\phi$  corresponding to the HT<sup>bot</sup> mechanism.

The semi-quantitative nature of the approach entails an uncertainty in the rotational energy barriers of the excited state. To account for this uncertainty, simulations were repeated with the two barrier heights for rotation around  $\tau$  and  $\phi$  varying between 0 and 7 kcal/mol (Fig. 6.3). Irrespective of the barrier heights simulations yielded again exclusively HT<sup>bot</sup> as the strictly preferred pathway. The isomerization time scales, however, varied, but were not the point of interest here.

The protonation of the ground and excited state chromophore are both not accessible experimentally. The results presented here were obtained for the

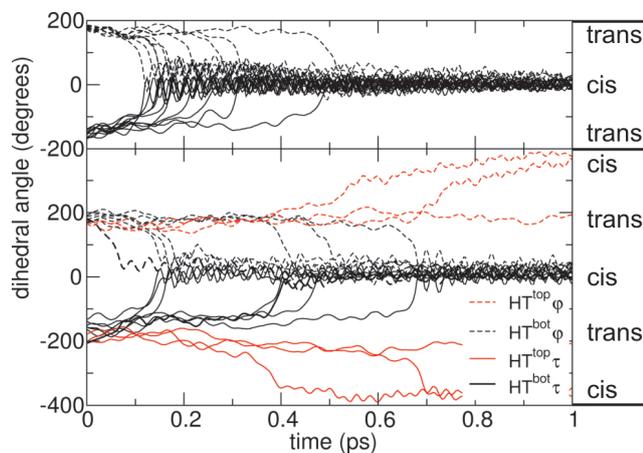


Figure 6.5: Spontaneous trans-cis-isomerization during free excited state MD simulations for the chromophore within the protein matrix (upper panel) and in water (lower panel), monitored via the dihedral angles  $\tau$  (dashed curves) and  $\phi$  (solid curves). The protein favors the  $\text{HT}^{\text{bot}}$  mechanism, with both dihedral angles rotating simultaneously within a narrow time frame. Simulation of protein-free isomerization of the chromophore in water also follows HT, and both directions (top and bottom) are observed.

zwitterionic protonation state. The corresponding anionic and neutral states could be excluded, as they showed rapid disintegration of the hydrogen bond network during MD simulations (see 6.2.2 for details). Another possible protonation state with an anionic chromophore and a protonated Glu215, thus having the same hydrogen bond network as the zwitterionic state, can be envisaged. Spontaneous isomerization of this state yielded highly similar spontaneous  $\text{HT}^{\text{bot}}$  trajectories.

The HT isomerization mechanism was also observed for the chromophore in water. However, without the restricting protein matrix, the direction of the rotation (top or bottom) as well as time scales varied, and the torsion around  $\tau$  and  $\phi$  was less correlated (Fig. 6.5, lower panel). In accordance with the above described FPMD simulations, the spontaneous isomerizations strongly suggest the  $\text{HT}^{\text{bot}}$  mechanism as the pre-dominant isomerization pathway in asFP595.

## 6.4 Discussion and Conclusion

Our FPMD simulations, on the basis of the involved forces, suggest the protein matrix surrounding the MYG chromophore of asFP595 to favor photo-



induced isomerization of MYG along the hula-twist ( $\text{HT}^{\text{bot}}$ ) mechanism. Enforcing the chromophore to isomerize *via* the R mechanism revealed a substantial hindrance of the tyrosyl rotation by the protein matrix. Simulations of the spontaneous isomerization of the excited state taking also the chromophore properties into account, confirmed this finding.

On the basis of our simulation results, one can speculate that the protein functions as a scaffold guiding the chromophore along the  $\text{HT}^{\text{bot}}$  isomerization pathway. It thereby restricts the ensemble of transition paths and dictates the kinetics of photoswitching. Indeed, simulations of the spontaneous isomerization of the MYG chromophore in water yielded a more heterogeneous picture in terms of the mechanism and time scale.

The simulations can be considered as a first step towards a rational tuning of the spectroscopic properties, such as the quantum yield and fluorescence life time, of asFP595 by protein engineering. A strategic modification of the switching properties will ultimately allow for the application in cell imaging and protein tracking with nanoscale resolution [195].

The molecular mechanical model of the excited state here was successfully applied to study the mechanism of spontaneous isomerization. However, the semi-quantitative description can not be assumed to accurately describe the process in detail, including the chromophore geometry and energetics during switching and its time scale. Hybrid quantum/molecular mechanical simulations are required to gain further insight into the photoisomerization, but are beyond the scope of this thesis. Open questions yet to be addressed are the presence of a putative conical intersection and, in this regard, the fate of the putative chromophore intermediate with both dihedrals twisted. Another remaining question is the reason for the fluorescent property of the *cis*-conformation, *i. e.*, the factors determining the kinetics of back-isomerization to the off-state. The hypothesis of a proton transfer in the on-state [192] is yet to be confirmed.

6. PHOTOSWITCH MECHANISM

---

# 7

## Summary and Conclusion

---

A major goal of protein science is to explore the coupling of protein motion and function. The underlying molecular basis in terms of protein dynamics, however, is mostly not directly accessible by experiments. This thesis aimed at characterizing the functional motions of three individual proteins, titin kinase, BmPBP and asFP595, by means of molecular dynamics simulations. Our findings were tested by and complemented experimental results, with the final goal to contribute to the understanding of the proteins' function — as a putative force sensor, a ligand filter and carrier, and a molecular switch, respectively.

**Muscle protein titin kinase.** Titin kinase, the enzymatic domain of the muscle protein titin, has been proposed to function as a force sensor by undergoing force-induced unfolding presumably associated with its activation. In the first part of this thesis, the force-induced unfolding and activation was studied by FPMD simulations. We found evidence for a sequential, mechanically induced opening of the catalytic site, which itself is left nearly unperturbed, as required for titin kinase activity. In the light of FPMD simulations of other titin domains revealing an abrupt disintegration of the domains under force, this distinctive behavior of titin kinase was unexpected. These results strongly support the hypothesis of titin kinase as a force sensor, fulfilling its function *via* an unfolding motion thoroughly designed to convert mechanical load into a biochemical signal.

The FPMD simulations suggest the rupture of two terminal  $\beta$ -sheets as the primary unfolding steps. The mechanically stable N-terminal  $\beta$ -sheet protects the active site from disintegration, while the lower mechanical strength

of the C-terminal  $\beta$ -sheet facilitates the release of the auto-inhibitory tail from the active site. The differential force resistance of the two  $\beta$ -sheets thus is crucial for the titin kinase activation, and was attributed to their contrastive topology. A simple FPMC model provided strong evidence that the topology of the two  $\beta$ -sheets is indeed the dominant factor for their differential mechanical strength. Our findings confirmed the general concept that the mechanical stability of proteins — particularly those with mechanical function — is achieved and controlled by their specific structural and topological properties.

Additional simulations of the enforced unfolding of three titin kinase mutants focussed on the structural factors determining mechanical strength. Mutants were chosen on the basis of the unfolding mechanism suggested by the titin kinase wild-type simulations. The mutant simulations demonstrated that the identified stabilizing structural units do not merely contribute to the force resistance of the protein in an additive manner. They instead disclosed a more complex and highly non-local dependency of the mechanical properties on the structure and unfolding motion as a whole.

FPMD simulations of protein unfolding are associated with three common problems. First, since time scales accessible to MD simulations are restricted to tens or hundreds of nanoseconds, unfolding is enforced several orders of magnitude faster than in experiments. Here, a range of pulling velocities was applied to seek out those mechanistic patterns of the activation mechanism that can be assumed to be also relevant under experimental conditions. FPMC simulations of a simplified  $\beta$ -sheet model allowed to further extend the accessible time scales towards those of experiments. Second, enforcing protein motions at nanosecond time scales causes artificially high friction, which here was reduced by two-side pulling of the titin kinase. Third, due to the substantial change of the protein extension during unfolding, the setup of a computationally feasible simulation system is problematic. Here, the simulation box was kept reasonably sized by deletion of external peptide fragments during pulling. Taken together, FPMD simulations, here in conjunction with FPMC simulations, again proved particularly adequate to assess enforced unfolding, and thereby designated themselves as a powerful tool to study other proteins with putative force sensing function, such as myosin light chain kinase. Forces still remain to be compared to AFM experiments.

Our simulations suggest titin kinase to be less force resistant than the immunoglobulin and fibronectin domains of titin. Titin kinase thus can be assumed to report on mechanical load that is lower than the load causing



irreversible unfolding of the titin molecule, as expected for a force sensor. Further insight into the sequence of domain unfolding events and involved forces can be gained from a simplified model of the whole titin molecule based on data from AFM experiments and FPMD simulations. This is the objective of future studies.

**Bombyx mori pheromone binding protein.** A second objective of this thesis was to reveal the structural and dynamical properties qualifying the olfactory protein BmPBP to function as a molecular filter and carrier for its physiological binding partner, the pheromone bombykol. By means of FPMD and ED simulations, two opposite ligand release pathways, along a flexible front lid and along the termini at the back, were identified to be equally likely. Free energy calculations supported this finding and ascribed the major energetic barriers to the underlying protein dynamics involving intramolecular hydrogen bond ruptures and the final loss of protein-ligand non-polar interactions.

We concluded that the two distinct pathways are both physiologically relevant, enabling BmPBP to take up and release the ligand sufficiently fast. The two individual pathways additionally provide an increased robustness of BmPBP towards point mutations, resulting in an increased capability to adjust the energy landscape of binding under evolutionary pressure.

A marginal difference in the binding affinity between the similar ligands bombykol and bombykal along with manifold, albeit restricted, binding modes of these ligands in the BmPBP pocket were obtained from FEP calculations and MD simulations of the complexes. This hints at a moderate, yet existent ligand selectivity of BmPBP. The versatile binding cavity along with two ligand binding pathways, allowing for the fast, tight and selective binding, render BmPBP particularly suitable as an efficient filter and carrier for bombykol and similar compounds.

The available experimental structures of BmPBP reveal a transition of the protein at low pH to a remarkably different conformation associated with a loss of bombykol affinity. How ligand release is linked to the pH-induced transition will be addressed in further investigations.

**Fluorescent protein asFP595.** The fluorescence of the chromoprotein asFP595 is reversibly switched on and off upon light-induced isomerization. As the third contribution, this thesis aimed at clarifying the underlying mechanism of the trans-cis isomerization reaction of asFP595. In particular, we wanted to gain insight into the influence of the protein matrix

surrounding and solidly embedding the chromophore onto the isomerization mechanism.

In FPMD simulations, side chains in close proximity to the chromophore were found to sterically favor the hula-twist mechanism, a mere space-saving rotation of a ring bridging group, over the rotation of a whole tyrosyl moiety during isomerization. Simulations of the spontaneous isomerization confirmed this finding, and yielded a mechanism in which the direction and duration is confined and controlled by the protein. The protein environment thus has the function to guide the reaction in terms of mechanism and kinetics.

The investigation of the three enforced processes, unfolding, ligand unbinding and a photoreaction, by means of MD simulations yielded a microscopic picture of the underlying mechanisms. By exploring the relevant regions of the energy landscape, the simulations shed light on the molecular properties of the protein that guide the process along the pathway as function requires. Thus, in conjunction with results from experiments, our simulations significantly contributed to the understanding of the respective protein function, and the structural requirements to fulfill it.

This in turn is prerequisite for the rationally tailored modification of the protein's properties by genetic engineering, active compounds, and other means, in the field of medicine, pharmacy, or biotechnology. More precisely, potential fields of application are the therapeutic intervention in titin kinase associated myopathies [119, 141], the targeted improvement of the photo-switching properties of asFP595 for fluorescence imaging applications [190], or the rational design of carriers, detergents or insecticides on the basis of PBPs [202].

The applications of this thesis also hint at the current limits of MD simulations. First, the direct environment of the protein — apart from water — in the cell, such as titin and myosin domains adjacent to titin kinase, or the charged neuronal membrane close to BmPBP, was not incorporated into the simulation setup due to the lack of computational power and structural data. Their influence on the observed molecular mechanism and energetics remains to be investigated by experiments and simulation. Second, free energy calculations were considerably hampered by the high flexibility of the system under investigation. Along with the access to longer simulation times, an advancement in sampling techniques, such as the usage of non-linear transition coordinates, will reduce non-equilibrium effects and improve the accuracy of free energy estimates. Third, the potential inaccu-



racies and limits in deriving appropriate force fields manifested themselves in this thesis for the pheromone components bombykol and bombykal and the ground and excited states of the asFP595 chromophore. The extended range and longer time-scales of applications in the future will further reveal the shortcomings and thereby assist the ongoing development of force fields [58], *e. g.*, by incorporating polarizability [203], and of hybrid quantum mechanical/molecular mechanical models [204, 205].

Thus, main future challenges, among others, are to increase the accuracies of force fields, to tackle the protein sampling problem for reliable free energy calculations, and to simulate larger protein or protein-membrane assemblies. Meeting these challenges requires more experimental protein structures and further advancements in the experimental assessment of protein dynamics, for an increasingly quantitative comparison of simulations with experiments for validation. Single molecules will certainly play a pronounced role in this endeavor.

Providing their comprehensive application and associated development, the potential of molecular dynamics simulations, as demonstrated in this thesis, as a microscopic complement to experiments then will ultimately result in a more fundamental molecular understanding of protein motion and function and in the ability to predict and rationally modify the underlying processes.

7. SUMMARY AND CONCLUSION

---

## Acknowledgment

The work for this thesis was encouraged and supported by a number of people to whom I would like to express my gratitude at this point.

First of all, I would like to thank Dr. Helmut Grubmüller, with the reasons for my acknowledgment being manifold. He accepted me in his group at the MPI for Biophysical Chemistry, and suggested me to study force-induced protein dynamics, by means of the exciting examples of titin kinase, BmPBP and asFP595. It was his steady support of my plans to carry out part of the work at the Shanghai Institute of Materia Medica (SIMM) that finally enabled me to do so. He at all times had an open ear for any questions regarding the work, in Göttingen and over the distance from Shanghai. Discussions and critical comments on manuscripts and on this thesis were always instructive, stimulating, and motivating. His support, at the same time, came along with the freedom to follow own ideas.

I also thank Prof. Dr. Jürgen Troe for his willingness to take over the position of the primary referee for this thesis and for his helpful suggestions.

Part of this work was carried out in the group of Hualiang Jiang at the SIMM, Chinese Academy of Sciences, Shanghai, P. R. China. I thank Jianglaoshi to accept me in his group as a guest researcher. I gratefully acknowledge the warm welcome and great support by his group, especially by Jianhua Shen, Hong Liu, Kunqian Yu, Jinzhi Tan, and Chunying Niu. In particular, I warmly thank Yechun Xu for helping me through the work and life in China, the great cooking, our trips, and all the fun.

I very much thank the whole Theoretical Molecular Biophysics department, its former and current members, for all the discussions, the help with Gromacs, the replies to emergency calls from Shanghai, fruit salads, and wetting canoeing. I especially would like to name Bert de Groot for many ideas regarding the BmPBP project and others, Lars V. Schäfer for our diverting team work on over and over astonishing asFP595, Ulrich Zachariae for carefully reading this thesis, Evi Heinemann, Rainer Böckmann, Matthias Müller, Gunnar Schröder, Oliver Slawik, Oliver Lange, and Martin Stumpe.

I also thank the experimentalists who stimulated the simulations of this

### *Acknowledgment*

---

thesis. The titin kinase project was initiated by M. Gautel, King's College London. AFM experiments of titin kinase and mutants are under way in the group of H. Gaub, LMU Munich. K. E. Kaissling, MPI for Ornithology, Seewiesen, addressed us with the olfaction project on BmPBP. S. Hell and M. Wahl of the MPI for Biophysical Chemistry suggested us to study asFP595. I thank all of them for the numerous discussions, for assisting the physiological interpretation of the results, and for carefully reading the manuscripts.

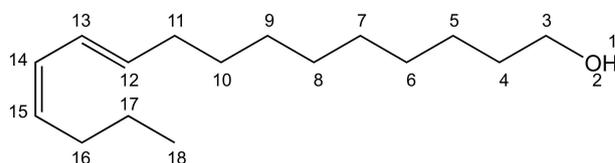
I would like to express my thanks to the Boehringer Ingelheim Foundation (B.I.F.) to support my thesis in form of a PhD scholarship, not only financially but in many respects. The B.I.F. made my stay in China, my participation in international conferences and in the Hirscheegg seminars possible. I am most grateful to Hermann Fröhlich, Monika Beutelspacher and Claudia Walther for their help and motivation, and the enjoyable times in Hirscheegg.

Support by the Max-Planck-Society and the 863 Hi-Tech Program of China by providing computing facilities used for this thesis is gratefully acknowledged.

This thesis would not be on hand in this form without those persons, my family and friends, who supportively accompanied me before and during this work. I am indebted to my parents for encouraging and backing me in all years, and for granting the unrestricted freedom to follow my interests, at whatever places. Thanks to my sisters Imke and Ute, for all good advice and enjoyable times.

Finally, I sincerely thank my husband for his boundless support, his never-ending patience, and the marvelous time we together had and have in Tübingen, Heidelberg, Shanghai, Göttingen and Stuttgart.

## Appendix A Bombykol and bombykal force field



Non-bonded parameters

| atom | atom type | charge | $C_6$        | $C_{12}$      |
|------|-----------|--------|--------------|---------------|
| 1    | H         | 0.398  | 0.0          | 0.0           |
| 2    | OA        | -0.548 | 0.0022619536 | 1.505529e-06  |
| 3    | CH2       | 0.150  | 0.0071048041 | 2.5775929e-05 |
| 4    | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 5    | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 6    | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 7    | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 8    | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 9    | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 10   | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 11   | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 12   | CH1       | 0.000  | 0.0037797904 | 1.1377129e-05 |
| 13   | CH1       | 0.000  | 0.0037797904 | 1.1377129e-05 |
| 14   | CH1       | 0.000  | 0.0037797904 | 1.1377129e-05 |
| 15   | CH1       | 0.000  | 0.0037797904 | 1.1377129e-05 |
| 16   | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 17   | CH2       | 0.000  | 0.0071048041 | 2.5775929e-05 |
| 18   | CH3       | 0.000  | 0.0099161764 | 3.3570436e-05 |

Table 7.1: Force field parameters for bombykol (I)

Appendix A

---

Bonded parameters

| Bonds      |    |                       |                 |  |        |   |
|------------|----|-----------------------|-----------------|--|--------|---|
| atom types |    | $b_{0,p1}$ (nm)       | $b_{0,p2}$ (nm) | $f_c$ (kJ mol <sup>-1</sup> nm <sup>-2</sup> ) |        |   |
| 1          | 2  | 0.100                 | 0.100           | 1.57e+07                                       |        |   |
| 2          | 3  | 0.143                 | 0.143           | 8.18e+06                                       |        |   |
| 3          | 4  | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 4          | 5  | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 5          | 6  | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 6          | 7  | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 7          | 8  | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 8          | 9  | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 9          | 10 | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 10         | 11 | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 11         | 12 | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 12         | 13 | 0.134                 | 0.136           | 1.02e+07                                       |        |   |
| 13         | 14 | 0.145                 | 0.148           | 0.76e+07                                       |        |   |
| 14         | 15 | 0.134                 | 0.136           | 1.02e+07                                       |        |   |
| 15         | 16 | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 16         | 17 | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| 17         | 18 | 0.153                 | 0.153           | 7.15e+06                                       |        |   |
| Angles     |    |                       |                 |  |        |   |
| atom types |    | $\theta_{0,p1}$ (deg) |                 | $\theta_{0,p2}$ (deg)                          |        | $f_c$ (kJ mol <sup>-1</sup> rad <sup>-2</sup> ) |
| 1          | 2  | 3                     | 109.5           | 109.5  | 450.00 |   |
| 2          | 3  | 4                     | 109.5           | 109.5  | 520.00 |   |
| 3          | 4  | 5                     | 109.5           | 109.5  | 520.00 |   |
| 4          | 5  | 6                     | 109.5           | 109.5  | 520.00 |   |
| 5          | 6  | 7                     | 109.5           | 109.5  | 520.00 |   |
| 6          | 7  | 8                     | 109.5           | 109.5  | 520.00 |   |
| 7          | 8  | 9                     | 109.5           | 109.5  | 520.00 |   |
| 8          | 9  | 10                    | 109.5           | 109.5  | 520.00 |   |
| 9          | 10 | 11                    | 109.5           | 109.5  | 520.00 |   |
| 10         | 11 | 12                    | 109.5           | 109.5  | 520.00 |   |
| 11         | 12 | 13                    | 126.0           | 120.0  | 560.00 |   |
| 12         | 13 | 14                    | 126.0           | 120.0  | 560.00 |   |
| 13         | 14 | 15                    | 126.0           | 120.0  | 560.00 |   |
| 14         | 15 | 16                    | 126.0           | 120.0  | 560.00 |   |
| 15         | 16 | 17                    | 109.5           | 109.5  | 520.00 |   |
| 16         | 17 | 18                    | 109.5           | 109.5  | 520.00 |   |

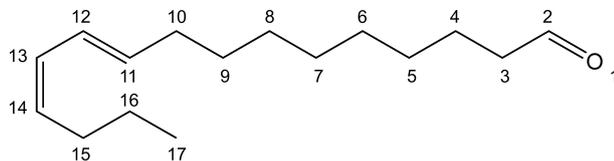


---

| Proper dihedrals |    |    |    |                  |  |  |              |
|------------------|----|----|----|------------------|--|--|--------------|
| atom types       |    |    |    | $\theta_0$ (deg) | $f_{c,p1}$ (kJ mol <sup>-1</sup> rad <sup>-2</sup> ) | $f_{c,p2}$ (kJ mol <sup>-1</sup> rad <sup>-2</sup> ) | multiplicity |
| 1                | 2  | 3  | 4  | 0                | 1.26   | 1.26   | 3            |
| 2                | 3  | 4  | 5  | 0                | 5.92   | 5.92   | 3            |
| 3                | 4  | 5  | 6  | 0                | 5.92   | 5.92   | 3            |
| 4                | 5  | 6  | 7  | 0                | 5.92   | 5.92   | 3            |
| 5                | 6  | 7  | 8  | 0                | 5.92   | 5.92   | 3            |
| 6                | 7  | 8  | 9  | 0                | 5.92   | 5.92   | 3            |
| 7                | 8  | 9  | 10 | 0                | 5.92   | 5.92   | 3            |
| 8                | 9  | 10 | 11 | 0                | 5.92   | 5.92   | 3            |
| 9                | 10 | 11 | 12 | 0                | 5.92   | 5.92   | 3            |
| 10               | 11 | 12 | 13 | 0                | 5.92   | 5.92   | 3            |
| 11               | 12 | 13 | 14 | 180              | 18.288   | 33.50  | 2            |
| 12               | 13 | 14 | 15 | 180              | 9.330  | 16.70  | 2            |
| 13               | 14 | 15 | 16 | 180              | 18.288   | 33.50  | 2            |
| 14               | 15 | 16 | 17 | 0                | 5.92   | 5.92   | 3            |
| 15               | 16 | 17 | 18 | 0                | 5.92   | 5.92   | 3            |

---

Table 7.2: Force field parameters for bombykol (II)



Non-bonded parameters

| atom | atom type | charge | $C_6$        | $C_{12}$      |
|------|-----------|--------|--------------|---------------|
| 1    | O         | -0.380 | 0.0022619536 | 1.5055290e-06 |
| 2    | CH1       | 0.380  | 0.0037797904 | 1.1377129e-05 |
| 3    | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 4    | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 5    | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 6    | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 7    | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 8    | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 9    | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 10   | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 11   | CH1       | 0      | 0.0037797904 | 1.1377129e-05 |
| 12   | CH1       | 0      | 0.0037797904 | 1.1377129e-05 |
| 13   | CH1       | 0      | 0.0037797904 | 1.1377129e-05 |
| 14   | CH1       | 0      | 0.0037797904 | 1.1377129e-05 |
| 15   | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 16   | CH2       | 0      | 0.0071048041 | 2.5775929e-05 |
| 17   | CH3       | 0      | 0.0099161764 | 3.3570436e-05 |

Table 7.3: Force field parameters for bombykal (I)



---

Bonded parameters

Bonds

| atom types | $b_0$ (nm) | $f_c$ (kJ mol <sup>-1</sup> nm <sup>-2</sup> ) |           |
|------------|------------|--|-----------|
| 1          | 2          | 0.123  | 1.66e+07  |
| 2          | 3          | 0.153  | 7.15e+06  |
| 3          | 4          | 0.153  | 7.15e+06  |
| 4          | 5          | 0.153  | 7.15e+06  |
| 5          | 6          | 0.153  | 7.15e+06  |
| 6          | 7          | 0.153  | 7.15e+06  |
| 7          | 8          | 0.153  | 7.15e+06  |
| 8          | 9          | 0.153  | 7.15e+06  |
| 9          | 10         | 0.153  | 7.15e+06  |
| 10         | 11         | 0.153  | 7.15e+06  |
| 11         | 12         | 0.134  | 1.02e+07  |
| 12         | 13         | 0.145  | 0.764e+07 |
| 13         | 14         | 0.134  | 1.02e+07  |
| 14         | 15         | 0.153  | 7.15e+06  |
| 15         | 16         | 0.153  | 7.15e+06  |
| 16         | 17         | 0.153  | 7.15e+06  |

Angles

| atom types | $\theta_0$ (deg) | $f_c$ (kJ mol <sup>-1</sup> rad <sup>-2</sup> ) |
|------------|------------------|---|
| 1 2 3      | 125.0            | 685.00  |
| 2 3 4      | 109.5            | 520.00  |
| 3 4 5      | 109.5            | 520.00  |
| 4 5 6      | 109.5            | 520.00  |
| 5 6 7      | 109.5            | 520.00  |
| 6 7 8      | 109.5            | 520.00  |
| 7 8 9      | 109.5            | 520.00  |
| 8 9 10     | 109.5            | 520.00  |
| 9 10 11    | 109.5            | 520.00  |
| 10 11 12   | 126.0            | 560.00  |
| 11 12 13   | 126.0            | 560.00  |
| 12 13 14   | 126.0            | 560.00  |
| 13 14 15   | 126.0            | 560.00  |
| 14 15 16   | 109.5            | 520.00  |
| 15 16 17   | 109.5            | 520.00  |

---

| Proper dihedrals |    |    |    |                  |   |              |
|------------------|----|----|----|------------------|---|--------------|
| atom types       |    |    |    | $\theta_0$ (deg) | $f_c$ (kJ mol <sup>-1</sup> rad <sup>-2</sup> ) | multiplicity |
| 1                | 2  | 3  | 4  | 0                | 1.00  | 6            |
| 2                | 3  | 4  | 5  | 0                | 5.92  | 3            |
| 3                | 4  | 5  | 6  | 0                | 5.92  | 3            |
| 4                | 5  | 6  | 7  | 0                | 5.92  | 3            |
| 5                | 6  | 7  | 8  | 0                | 5.92  | 3            |
| 6                | 7  | 8  | 9  | 0                | 5.92  | 3            |
| 7                | 8  | 9  | 10 | 0                | 5.92  | 3            |
| 8                | 9  | 10 | 11 | 0                | 5.92  | 3            |
| 9                | 10 | 11 | 12 | 0                | 5.92  | 3            |
| 10               | 11 | 12 | 13 | 180              | 18.288  | 2            |
| 11               | 12 | 13 | 14 | 180              | 9.330   | 2            |
| 12               | 13 | 14 | 15 | 180              | 18.288  | 2            |
| 13               | 14 | 15 | 16 | 0                | 5.92  | 3            |
| 14               | 15 | 16 | 17 | 0                | 5.92  | 3            |

Table 7.4: Force field parameters for bombykal (II)

# Bibliography

- [1] H. Frauenfelder, S. G. Sligar und P. G. Wolynes. The energy landscapes and motions of proteins. *Science*, 254:1598–1603, 1991.
- [2] M. Karplus und J. Kuriyan. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. USA*, 102:6679–6685, 2005.
- [3] E. Z. Eisenmesser, D. A. Bosco, M. Akke und D. Kern. Enzyme dynamics during catalysis. *Science*, 295:1520–1523, 2002.
- [4] J. G. Kempf und J. P. Loria. Protein dynamics from solution NMR - Theory and applications. *Cell Biochem. Biophys.*, 37:187–211, 2003.
- [5] R. Bruschweiler. New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins. *Curr. Opin. Struct. Biol.*, 13:175–183, 2003.
- [6] S. Weiss. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nature Struct. Biol.*, 7:724–729, 2000.
- [7] B. Schuler. Single-molecule fluorescence spectroscopy of protein folding. *Chem. Phys. Chem.*, 6:1206–1220, 2005.
- [8] L. Tskhovrebova, J. Trinick, J. Sleep und R. M. Simmons. Elasticity and unfolding of single molecules of the giant muscle protein titin. *Nature*, 387:308–312, 1997.
- [9] M. S. Z. Kellermayer, S. B. Smith, C. Bustamante und H. L. Granzier. Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science*, 276:1112–1116, 1997.
- [10] V. T. Moy, E. L. Florin und H. E. Gaub. Intermolecular forces and energies between ligands and receptors. *Science*, 266:257–259, 1994.
- [11] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez und H. E. Gaub. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science*, 276:1109–1112, 1997.
- [12] T. E. Fisher, P. E. Marszalek und J. M. Fernandez. Stretching single molecules into novel conformations using the atomic force microscope. *Nature Struct. Biol.*, 7:719–724, 2000.

## BIBLIOGRAPHY

---

- [13] P. E. Marszalek, H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten und J. M. Fernandez. Mechanical unfolding intermediates in titin modules. *Nature*, 402:100–103, 1999.
- [14] A. F. Oberhauser, C. Badilla-Fernandez, M. Carrion-Vazquez und J. M. Fernandez. The mechanical hierarchies of fibronectin observed with single-molecule AFM. *J. Molec. Biol.*, 319:433–447, 2002.
- [15] H. B. Li, A. F. Oberhauser, S. D. Redick, M. Carrion-Vazquez, H. P. Erickson und J. M. Fernandez. Multiple conformations of PEVK proteins detected by single-molecule techniques. *Proc. Natl. Acad. Sci. USA*, 98:10682–10686, 2001.
- [16] S. P. Ng, R. W. S. Rounsevell, A. Steward, C. D. Geierhaas, P. M. Williams, E. Paci und J. Clarke. Mechanical unfolding of TNfn3: The unfolding pathway of a fnIII domain probed by protein engineering, AFM and MD simulation. *J. Molec. Biol.*, 350:776–789, 2005.
- [17] M. Carrion-Vazquez, H. B. Li, H. Lu, P. E. Marszalek, A. F. Oberhauser und J. M. Fernandez. The mechanical stability of ubiquitin is linkage dependent. *Nature Struct. Biol.*, 10:738–743, 2003.
- [18] J. M. Fernandez und H. B. Li. Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*, 303:1674–1678, 2004.
- [19] M. Schlierf, H. B. Li und J. M. Fernandez. The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proc. Natl. Acad. Sci. USA*, 101:7299–7304, 2004.
- [20] S. Smith, Y. Cui und C. Bustamante. Overstretching b-DNA: the elastic response of individual double stranded and single-stranded DNA molecules. *Science*, 271:795–799, 1996.
- [21] M. Rief, F. Oesterhelt, B. Heymann und H. E. Gaub. Single molecule force spectroscopy reveals conformational change in polysaccharides. *Science*, 275:1295–1297, 1997.
- [22] B. D. Sattin, A. E. Pelling und M. C. Goh. DNA base pair resolution by single molecule force spectroscopy. *Nucleic Acids Res.*, 32:4876–4883, 2004.
- [23] Z. Y. Lu, W. Nowak, G. R. Lee, P. E. Marszalek und W. T. Yang. Elastic properties of single amylose chains in water: A quantum mechanical and AFM study. *J. Am. Chem. Soc.*, 126:9033–9041, 2004.
- [24] E. L. Florin, V. T. Moy und H. E. Gaub. Adhesion forces between individual ligand-receptor pairs. *Science*, 264:415–417, 1994.
- [25] U. Dammer, O. Popescu, P. Wagner, D. Anselmetti und H. J. Güntherodt. Binding strength between cell-adhesion proteoglycans measured by atomic-force microscopy. *Science*, 267:1173–1175, 1995.



- [26] F. Kuhner, L. T. Costa, P. M. Bisch, S. Thalhammer, W. M. Heckl und H. E. Gaub. Lexa-DNA bond strength by single molecule force spectroscopy. *Biophys. J.*, 87:2683–2690, 2004.
- [27] S. B. Fowler und J. Clarke. Mapping the folding pathway of an immunoglobulin domain: Structural detail from phi value analysis and movement of the transition state. *Structure*, 9:355–366, 2001.
- [28] P. M. Williams, S. B. Fowler, R. B. Best, J. L. Toca-Herrera, K. A. Scott, A. Steward und J. Clarke. Hidden complexity in the mechanical properties of titin. *Nature*, 422:446–449, 2003.
- [29] C. F. Wright, A. Steward und J. Clarke. Thermodynamic characterisation of two transition states along parallel protein folding pathways. *J. Molec. Biol.*, 338:445–451, 2004.
- [30] J. Schlitter, M. Engels und P. Kruger. Targeted molecular-dynamics - a new approach for searching pathways of conformational transitions. *J. Mol. Graph.*, 12:84–89, 1994.
- [31] G. N. Patey und J. P. Valleau. Monte-Carlo method for obtaining interionic potential of mean force in ionic solution. *J. Chem. Phys.*, 63:2334–2339, 1975.
- [32] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman und J. M. Rosenberg. The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. *J. Comp. Chem.*, 13:1011–1021, 1992.
- [33] H. Grubmüller. Force probe molecular dynamics simulations. *Methods Mol. Biol.*, 305:493–515, 2005.
- [34] A. Amadei, A. B. M. Linssen, B. L. de Groot, D. M. F. van Aalten und H. J. C. Berendsen. An efficient method for sampling the essential sub-space of proteins. *J. Biom. Str. Dyn.*, 13:615–626, 1996.
- [35] B. L. de Groot, A. Amadei, D. M. F. van Aalten und H. J. C. Berendsen. Towards an exhaustive sampling of the configurational spaces of the two forms of the peptide hormone guanylin. *J. Biom. Str. Dyn.*, 13:615–626, 1996.
- [36] H. Grubmüller. Predicting slow structural transitions in macromolecular systems - conformational flooding. *Phys. Rev. E*, 52:2893–2906, 1995.
- [37] A. Laio und M. Parrinello. Escaping free-energy minima. *Proc. Natl. Acad. Sci. USA*, 99:12562–12566, 2002.
- [38] P. A. Bash, U. C. Singh, F. K. Brown, R. Langridge und P. A. Kollman. Calculation of the relative change in binding free-energy of a protein-inhibitor complex. *Science*, 235:574–576, 1987.
- [39] T. Simonson, G. Archontis und M. Karplus. Free energy simulations come of age: Protein-ligand recognition. *Acc. Chem. Res.*, 35:430–437, 2002.
- [40] H. Lu, B. Isralewitz, A. Krammer, V. Vogel und K. Schulten. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.*, 75:662–671, 1998.

## BIBLIOGRAPHY

---

- [41] B. Isralewitz, M. Gao und K. Schulten. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.*, 11:24–230, 2001.
- [42] F. Gräter, J. H. Shen, H. L. Jiang, M. Gautel und H. Grubmüller. Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations. *Biophys. J.*, 88:790–804, 2005.
- [43] H. Grubmüller, B. Heymann und P. Tavan. Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*, 271:997–999, 1996.
- [44] B. Heymann und H. Grubmüller. Molecular dynamics force probe simulations of antibody/antigen unbinding: Entropic control and non-additivity of unbinding forces. *Biophys. J.*, 81:1295–1313, 2001.
- [45] M. Gao, D. Craig, V. Vogel und K. Schulten. Simulated unbinding RGD peptide ligand from integrin  $\alpha(v)\beta(3)$ . *Biophys. J.*, 84:396A–396A, 2003.
- [46] L. L. Shen, J. H. Shen, X. M. Luo, F. Cheng, Y. C. Xu, K. X. Chen, E. Arnold, J. P. Ding und H. L. Jiang. Steered molecular dynamics simulation on the binding of NNRTI to HIV-1 RT. *Biophys. J.*, 84:3547–3563, 2003.
- [47] B. L. de Groot, A. Amadei, R. M. Scheek, N. A. J. vanNuland und H. J. C. Berendsen. An extended sampling of the configurational space of HPr from E-coli. *Proteins*, 26:314–322, 1996.
- [48] I. Daidone, A. Amadei, D. Roccatano und A. Di Nola. Molecular dynamics simulation of protein folding by essential dynamics sampling: Folding landscape of horse heart cytochrome c. *Biophys. J.*, 85:2865–2871, 2003.
- [49] W. L. Jorgensen und C. Ravimohan. Monte-carlo simulation of differences in free-energies of hydration. *J. Chem. Phys.*, 83:3050–3054, 1985.
- [50] B. L. Tembe und J. A. McCammon. Ligand receptor interactions. *Comp. & Chem.*, 8:281–283, 1984.
- [51] W. L. Jorgensen und J. K. Buckner. Use of statistical perturbation-theory for computing solvent effects on molecular-conformation - butane in water. *J. Phys. Chem.*, 91:6083–6085, 1987.
- [52] O. Mayans, P. F. M. van der Ven, M. Wilm, A. Mues, P. Young, D. O. Fürst, M. Wilmanns und M. Gautel. Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature*, 395:863 – 397, 1998.
- [53] T. Hansson, C. Oostenbrink und W. F. van Gunsteren. Molecular dynamics simulations. *Curr. Opin. Struct. Biol.*, 12:190–196, 2002.
- [54] A. R. Leach. *Molecular Modelling: Principles and Applications*. Pearson Education Limited, Essex, England, 2001.
- [55] F. Jensen. *Computational Chemistry*. John Wiley and Sons Ltd, Sussex, England, 2001.



- [56] P. W. Atkins. *Molecular Quantum Mechanics*. Oxford University Press, Oxford, 1983.
- [57] U. Dinur und A. T. Hagler. *New approaches to empirical force fields*. Wiley-VCH, New York, USA, 1991.
- [58] W. L. Jorgensen und J. Tirado-Rives. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. USA*, 102:6665–6670, 2005.
- [59] J. E. Lennard-Jones. On the forces between atoms and ions. *Proc. R. Soc. London Series A - Math. and Phys.*, 109:584–597, 1925.
- [60] W. L. Jorgensen und J. Tirado-Rives. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. USA*, 102:6665–6670, 2005.
- [61] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan und M. Karplus. Charmm - a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.*, 4:187–217, 1983.
- [62] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott und I. G. Tironi. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland, 1996.
- [63] W. L. Jorgensen und J. Tirado-Rives. The OPLS potential functions for proteins - energy minimizations for crystals of cyclic-peptides and crambin. *J. Am. Chem. Soc.*, 110:1657–1666, 1988.
- [64] L. Verlet. Computer experiments on classical fluids .i. thermodynamical properties of lennard-jones molecules. *Phys. Rev.*, 159:98–103, 1967.
- [65] B. Hess, H. Bekker, H. J. C. Berendsen und J. G. E. M. Fraaije. LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.*, 18:1463–1472, 1997.
- [66] M. Saito. Molecular dynamics simulations of proteins in solution - artifacts by the cutoff approximation. *J. Chem. Phys.*, 101:4055–4061, 1994.
- [67] M. Saito. Molecular dynamics simulations of proteins in solution with all degrees of freedom and long-range Coulomb interactions. *J. Phys. Chem.*, 99:17043–17048, 1995.
- [68] P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Physik (Leipzig)*, IV:253–287, 1921.
- [69] T. Darden, D. York und L. Pedersen. Particle Mesh Ewald — An  $N \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.*, 98:10089–10092, 1993.
- [70] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee und L. G. Pederson. The smooth particle mesh Ewald method. *J. Chem. Phys.*, 103:8577, 1995.

## BIBLIOGRAPHY

---

- [71] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola und J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81(8):3684–3690, 1984.
- [72] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller und E. Teller. Equations of state applications by fast computing machines. *J. Chem. Phys.*, 21:1087–1091, 1957.
- [73] A. Mitsutake, Y. Sugita und Y. Okamoto. Generalized-ensemble algorithms for molecular simulations of biopolymers. *Biopolymers*, 60:96–123, 2001.
- [74] K. Tai. Conformational sampling for the impatient. *Biophys. Chem.*, 107:213–220, 2004.
- [75] C. A. Hixson, J. Chen, Z. N. Huang und R. A. Wheeler. New perspectives on multiple-copy, mean-field molecular dynamics methods. *J. Mol. Graph. Mod.*, 22:349–357, 2004.
- [76] Y. Sugita und Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, 314:141–151, 1999.
- [77] A. E. García und K. Y. Sanbonmatsu. Exploring the energy landscape of a beta hairpin in explicit solvent. *Proteins*, 42:345–354, 2001.
- [78] K. Y. Sanbonmatsu und A. E. García. Structure of met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. *Proteins*, 46:225–234, 2002.
- [79] A. E. Garcia und J. N. Onuchic. Folding a protein in a computer: An atomic description of the folding/unfolding of protein a. *Proc. Natl. Acad. Sci. USA*, 100:13898–13903, 2003.
- [80] J. Gsponer, U. Haberthur und A. Caffisch. The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion sup35. *Proc. Natl. Acad. Sci. USA*, 100:5154–5159, 2003.
- [81] W. Im und C. L. Brooks. Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA*, 102:6771–6776, 2005.
- [82] R. B. Best, J. Clarke und M. Karplus. What contributions to protein side-chain dynamics are probed by NMR experiments? A molecular dynamics simulation analysis. *J. Molec. Biol.*, 349:185–203, 2005.
- [83] B. Heymann. *Molekulardynamiksimulationen kraftmikroskopischer Einzelmolekülexperimente. Dissertation.* D. Reidel Publishing Company, Georg-August-Universität Göttingen, 1999.
- [84] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica (Utrecht)*, VII:284–304, 1940.



- [85] A. Amadei, A. B. M. Linssen und H. J. C. Berendsen. Essential dynamics of proteins. *Proteins*, 17:412–425, 1993.
- [86] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, 1997.
- [87] G. Hummer und A. Szabo. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl. Acad. Sci. USA*, 98:3658–3661, 2001.
- [88] S. Park, F. Khalili-Araghi, E. Tajkhorshid und K. Schulten. Free energy calculation from steered molecular dynamics simulations using Jarzynski’s equality. *J. Chem. Phys.*, 119:3559–3566, 2003.
- [89] R. W. Zwanzig. High-temperature equation of state by a perturbation method. 1. Nonpolar gases. *J. Chem. Phys.*, 22:1420–1426, 1954.
- [90] J. Gore, F. Ritort und C. Bustamante. Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements. *Proc. Natl. Acad. Sci. USA*, 100:12564–12569, 2003.
- [91] M. Nilges, G. M. Clore und Gronenborn A. M. Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. *FEBS Lett.*, 239:129136, 1988.
- [92] T. C. Beutler, A. E. Mark, R. C. Vanschaik, P. R. Gerber und W. F. van Gunsteren. Avoiding singularities and numerical instabilities in free-energy calculations based on molecular simulations. *Chem. Phys. Lett.*, 222:529–539, 1994.
- [93] H. Schäfer, W. F. van Gunsteren und A. E. Mark. Estimating relative free energies from a single ensemble: Hydration free energies. *J. Comp. Chem.*, 20:1604–1617, 1999.
- [94] A. R. Means. Muscle proteins - the clash in titin. *Nature*, 395:846–847, 1998.
- [95] K. Maruyama, S. Matsubara, R. Natori, Y. Nonomura, S. Kimura, K. Ohashi, F. Murakami, S. Handa und G. Eguchi. Connectin, an elastic protein of muscle. *J. Biochemistry*, 82:317–350, 1977.
- [96] S. Labeit und B. Kolmerer. Titins - giant proteins in charge of muscle ultrastructure and elasticity. *Science*, 270:293–296, 1995.
- [97] D. O. Fürst, M. Osborn, R. Nave und K. Weber. The organization of titin filaments in the half-sarcomere revealed by monoclonal-antibodies in immunoelectron microscopy - a map of 10 nonrepetitive epitopes starting at the Z-line extends close to the M-line. *J. Cell Biol.*, 106:1563–1572, 1988.
- [98] L. Tskhovrebova und J. Trinick. Titin: properties and family relationships. *Nat. Rev. Mol. Cell. Biol.*, 4:679–689, 2003.

## BIBLIOGRAPHY

---

- [99] S. Labeit und B. Kolmerer. Titins, giant proteins in charge of muscle ultra-structure and elasticity. *Science*, 270:293–296, 1995.
- [100] J. Trinick und L. Tskhovrebova. Titin: a molecular control freak. *Trends in Cell Biol.*, 9:377–380, 1999.
- [101] M. S. Z. Kellermayer und L. Grama. Stretching and visualizing titin molecules: combining structure, dynamics and mechanics. *J. Muscle Res. Cell Motility*, 23:499–511, 2002.
- [102] W. A. Linke und J. M. Fernandez. Cardiac titin: molecular basis of elasticity and cellular contribution to elastic and viscous stiffness components in myocardium. *J. Muscle Res. Cell Motility*, 23:483–497, 2002.
- [103] H. Li, W. A. Linke, A. F. Oberhauser, M. Carrion-Vazquez, J. G. Kerkvliet, H. Lu, P. E. Marszalek und J. M. Fernandez. Reverse engineering of the giant muscle protein titin. *Nature*, 418:998–1002, 2002.
- [104] S. B. Fowler, R. B. Best, J. L. T. Herrera, T. J. Rutherford, A. Steward, E. Paci, M. Karplus und J. Clarke. Mechanical unfolding of a titin ig domain: Structure of unfolding intermediate revealed by combining AFM, molecular dynamics simulations, NMR and protein engineering. *J. Molec. Biol.*, 322:841–849, 2002.
- [105] W. A. Linke, M. Kulke, H. B. Li, S. Fujita-Becker, C. Neagoe, D. J. Manstein, M. Gautel und J. M. Fernandez. PEVK domain of titin: An entropic spring with actin-binding properties. *J. Struct. Biol.*, 137:194–205, 2002.
- [106] E. Paci und M. Karplus. Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. *J. Molec. Biol.*, 288:441–459, 1999.
- [107] M. Rief und H. Grubmüller. Force spectroscopy of single biomolecules. *Chem. Phys. Chem.*, 3:255–261, 2002.
- [108] H. Lu und K. Schulten. The key event in force-induced unfolding of titin’s immunoglobulin domains. *Biophys. J.*, 79:51–65, 2000.
- [109] P. M. Williams, S. B. Fowler or R. B. Best, J. L. Toca-Herrera, K. A. Scott, A. Steward und J. Clarke. Hidden complexity in the mechanical properties of titin. *Nature*, 422:446–469, 2003.
- [110] D. Bossemeyer, R. A. Engh, V. Kinzel, H. Postingl und Hueber R. Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with Mn<sup>+2</sup> adenylyl imidodiphosphate and inhibitor peptide PKI(5-24). *EMBO J.*, 12:849–859, 1993.
- [111] B. Kobe, J. Heiherhorst, S. C. Feil, M. W. Parker und G. M. Benian. Giant protein kinases: domain interactions and structural basis of autoregulation. *EMBO J.*, 15:6810–6821, 1996.



- [112] W. L. DeLano. *PyMOL Manual*. DeLano Scientific, San Carlos, CA, USA. Internet: <http://www.delanoscientific.com>, 2001.
- [113] G. Nicholas, M. Thomas, B. Langley, W. Somers, K. Patel, C. F. Kemp, M. Sharma und R. Kambadur. Titin-cap associates with, and regulates secretion of, myostatin. *J. Cell. Physiology*, 193:120–131, 2002.
- [114] T. Centner, J. Yano, E. Kimura, A. S. McElhinny, K. Pelin, C. C. Witt, M. L. Bang, K. Tromitas, H. Granzier, C. C. Gregorio, H. Sorimachi und S. Labeit. Identification of muscle specific ring finger proteins as potential regulators of the titin kinase domain. *J. Molec. Biol.*, 306:717–726, 2001.
- [115] A. S. McElhinny, K. Kakinuma, H. Sorimachi, S. Labeit und C. C. Gregorio. Muscle-specific RING finger-1 interacts with titin to regulate sarcomeric m-line and thick filament structure and may have nuclear functions via its interaction with glucocorticoid modulatory element binding protein-1. *J. Cell Biol.*, 157:125–136, 2002.
- [116] V. Pizon, A. Iakovenko, P. F. M. van der Ven, R. Kelly, C. Fatu, D. O. Fürst, E. Karsenti und M. Gautel. Transient association of titin and myosin with microtubules in nascent myofibrils directed by the MURF2 RING-finger protein. *J. Cell Science*, 115:4469–4482, 2002.
- [117] G. Miller, H. Musa, M. Gautel und M. Peckham. A targeted deletion of the c-terminal end of titin, including the titin kinase domain, impairs myofibrillogenesis. *J. of Cell Science*, 116:4811–4819, 2003.
- [118] M. Gottardt, R. E. Hammer, N. Hübner, J. Monti, C. C. Witt, M. McNabb, J. A. Richardson, H. Granzier, S. Labeit und J. Herz. Conditional expression of mutant m-line titins results in cardiomyopathy with altered sarcomere structure. *J. Biol. Chem.*, 278:6059–6065, 2003.
- [119] S. Lange, F. Q. Xiang, A. Yakovenko, A. Vihola, P. Hackman, E. Rostkova, J. Kristensen, B. Brandmeier, G. Franzen, B. Hedberg, L. G. Gunnarsson, S. M. Hughes, S. Marchand, T. Sejersen, I. Richard, L. Edstrom, E. Ehler, B. Udd und M. Gautel. The kinase domain of titin controls muscle gene expression and protein turnover. *Science*, 308:1599–1603, 2005.
- [120] G. M. Benian, X. X. Tang und T. L. Tinley. Twitchin and related giant IG superfamily members of c-elegans and other invertebrates. *Adv. Biophys.*, 33:183–198, 1996.
- [121] E. Lindahl, B. Hess und D. van der Spoel. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.*, 7:306–317, 2001.
- [122] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren und J. Hermans. *Interaction model for water in relation to protein hydration*. D. Reidel Publishing Company, Dordrecht, The Netherlands, 1981.
- [123] G. Vriend. WHAT IF: a molecular modelling and drug design program. *J. Mol. Graph.*, 8:52–56, 1990.

## BIBLIOGRAPHY

---

- [124] A. Nicholls, K. A. Sharp und B. Honig. *DelPhi V3.0*. Columbia University, New York, 1990.
- [125] S. Hayward und H. J. C. Berendsen. Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme. *Proteins*, 30:144–154, 1998.
- [126] E. Espinosa, E. Molins und C. Lecomte. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem. Phys. Lett.*, 285:170–173, 1998.
- [127] B. Heymann und H. Grubmüller. AN02/DNP unbinding forces studied by molecular dynamics AFM simulations. *Chem. Phys. Lett.*, 303:1–9, 1999.
- [128] S. R. Hubbard. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J.*, 16:5572–5581, 1997.
- [129] R. Rohs, C. Etchebest und R. Lavery. Unraveling proteins: a molecular mechanics study. *Biophys. J.*, 76:2760–2768, 1999.
- [130] D. J. Brockwell, E. Paci, R. C. Zinober, G. S. Beddard, P. D. Olmsted, D. A. Smith, R. N. Perham und S. E. Radford. Pulling geometry defines the mechanical resistance of a  $\beta$ -sheet protein. *Nature Struct. Biol.*, 10:731–737, 2003.
- [131] F. Zhang, A. Strand, D. Robbins, M. H. Cobb und E. J. Goldsmith. Atomic structure of the MAP Kinase ERK2 at 2.3 Å resolution. *Nature*, 367:704–711, 1994.
- [132] D. R. Knighton, J. Zheng, L. F. Ten Eyck, V. A. Ashford, N.-H. Xuong, S. S. Taylor und J. M. Sowadski. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*, 253:407–414, 1991.
- [133] D. Bossemeyer. The glycine-rich sequence of protein kinases: a multifunctional element. *TIBS*, 19:201–205, 1994.
- [134] R. B. Best, B. Li, A. Steward, V. Daggett und J. Clarke. Can non-mechanical proteins withstand force? Stretching barnase by atomic force microscopy and molecular dynamics simulation. *Biophys. J.*, 81:2344–2356, 2001.
- [135] M. Rief, M. Gautel, A. Schemmel und H. E. Gaub. The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy. *Biophys. J.*, 75:2008–3014, 1998.
- [136] H. Li und J. M. Fernandez. Mechanical design of the first proximal IG domain of human cardiac titin revealed by single molecule force spectroscopy. *J. Molec. Biol.*, 334:75–86, 2003.
- [137] M. Rief, J. Pascual, M. Saraste und Hermann E. Gaub. Single molecule force spectroscopy of spectrin repeats: Low unfolding forces in helix bundles. *J. Molec. Biol.*, 286:553–561, 1999.



- [138] M. Gao, H. Lu und K. Schulten. Unfolding of titin domains studied by molecular dynamics simulations. *J. of Muscle Res. and Cell Motility*, 23:513–521, 2002.
- [139] M. Gao, D. Craig, O. Lequin, I. D. Campbell, V. Vogel und K. Schulten. Structure and functional significance of mechanically unfolded fibronectin type III<sub>1</sub> intermediates. *Proc. Natl. Acad. Sci. USA*, 100:14784–14789, 2003.
- [140] M. Carrion-Vazquez, H. Li, H. Lu, P. E. Marszalek, A. F. Oberhauser und J. M. Fernandez. The mechanical stability of ubiquitin is linkage dependent. *Nature Struct. Biol.*, 10:738–743, 2003.
- [141] L. Tshkhovrebova und J. Trinick. Muscle disease: a giant feels the strain. *Nature Medicine*, 11:478–479, 2005.
- [142] F. Oesterhelt, D. Oesterhelt, M. Pfeiffer, A. Engel, H. E. Gaub und D. J. Muller. Unfolding pathways of individual bacteriorhodopsins. *Science*, 288:143–146, 2000.
- [143] H. B. Li, M. Carrion-Vazquez, A. F. Oberhauser, P. E. Marszalek und J. M. Fernandez. Point mutations alter the mechanical stability of immunoglobulin modules. *Nature Struct. Biol.*, 7:1117–1120, 2000.
- [144] C. D. Geierhaas, E. Paci, M. Vendruscolo und J. Clarke. Comparison of the transition states for folding of two ig-like proteins from different superfamilies. *J. Molec. Biol.*, 343:1111–1123, 2004.
- [145] Tripos Associates, St. Louis, MO. *Sybyl Molecular Modeling System, version 7.0*.
- [146] L. Buck und R. Axel. A novel multigene family may encode odorant receptors - a molecular-basis for odor recognition. *Cell*, 65:175–187, 1991.
- [147] R. G. Vogt und L. M. Riddiford. Pheromone binding inactivation by moth antennae. *Nature*, 293:161–163, 1981.
- [148] E. Bignetti, A. Cavaggioni, P. Pelosi, K. C. Persaud, R. T. Sorbi und R. Tirindelli. Purification and characterization of an odorant-binding protein from cow nasal tissue. *Eur. J. Biochemistry*, 149:227–231, 1985.
- [149] R. A. Steinbrecht, M. Laue und G. Ziegelberger. Immunolocalization of insect odorant-binding proteins - a comparative-study. *Chemical Senses*, 20:109–110, 1995.
- [150] S. W. Kruse, R. Zhao, D. P. Smith und D. N. M. Jones. Structure of a specific alcohol-binding site defined by the odorant binding protein lush from drosophila melanogaster. *Nature Struct. Biol.*, 10:694–700, 2003.
- [151] A. Lartigue, A. Gruez, S. Spinelli, S. Riviere, R. Brossut, M. Tegoni und C. Cambillau. The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism. *J. Biol. Chem.*, 278:30213–30218, 2003.

## BIBLIOGRAPHY

---

- [152] A. Lartigue, A. Gruez, L. Briand, F. Blon, V. Bezirard, M. Walsh, J. C. Pernellet, M. Tegoni und C. Cambillau. Sulfur single-wavelength anomalous diffraction crystal structure of a pheromone-binding protein from the honeybee *Apis mellifera* L. *J. Biol. Chem.*, 279:4459–4464, 2004.
- [153] R. A. Steinbrecht. Odorant-binding proteins: Expression and function. *Olfaction And Taste Xii*, 855:323–332, 1998.
- [154] O. Karlson und M. Lüscher. Pheromones: A new term for a class of biologically active substance. *Nature*, 183:55–56, 1959.
- [155] K. Mori. Chirality and insect pheromones. *Chirality*, 10:578–586, 1998.
- [156] K. E. Kaissling. Olfactory perireceptor and receptor events in moths: A kinetic model. *Chemical Senses*, 26:125–150, 2001.
- [157] T. Sakurai, T. Nakagawa, H. Mitsuno, H. Mori, Y. Endo, S. Tanoue, Y. Yasukochi, K. Touhara und T. Nishioka. Identification and functional characterization of a sex pheromone receptor in the silkworm *Bombyx mori*. *Proc. Natl. Acad. Sci. USA*, 101:16653–16658, 2004.
- [158] B. H. Sandler, L. Nikonova, W. S. Leal und J. Clardy. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chemistry & Biology*, 7:143–151, 2000.
- [159] W. S. Leal, A. M. Chen, Y. Ishida, V. P. Chiang, M. L. Erickson, T. I. Morgan und J. M. Tsuruda. Kinetics and molecular properties of pheromone binding and release. *Proc. Natl. Acad. Sci. USA*, 102:5386–5391, 2005.
- [160] K.-E. Kaissling, U. Klein, J.J. de Kramer, T.A. Keil, S. Kanaujia und J. Hemberger. *Insect Olfactory Cells: Electrophysiological and Biochemical Studies*. Proc. of the Intern. Symp. in Memory of D. Nachmansohn, Berlin, 1985.
- [161] G. H. Du und G. D. Prestwich. Protein-structure encodes the ligand-binding specificity in pheromone binding-proteins. *Biochemistry*, 34:8726–8732, 1995.
- [162] D. Lee, F. F. Damberger, G. H. Peng, R. Horst, P. Guntert, L. Nikonova, W. S. Leal und K. Wuthrich. NMR structure of the unliganded *Bombyx mori* pheromone-binding protein at physiological pH. *FEBS Letters*, 531:314–318, 2002.
- [163] H. Wojtasek und W. S. Leal. Conformational change in the pheromone-binding protein from *Bombyx mori* induced by pH and by interaction with membranes. *J. Biol. Chem.*, 274:30950–30956, 1999.
- [164] R. Horst, F. Damberger, P. Luginbuhl, P. Guntert, G. Peng, L. Nikonova, W. S. Leal und K. Wuthrich. NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *Proc. Natl. Acad. Sci. USA*, 98:14374–14379, 2001.
- [165] G. D. Prestwich, G. H. Du und S. Laforest. How is pheromone specificity encoded in proteins. *Chemical Senses*, 20:461–469, 1995.



- [166] T. Nemoto, M. Uebayashi und Y. Komeiji. Flexibility of a loop in a pheromone binding protein from *bombyx mori*: a molecular dynamics simulation. *Chem-Bio Informatics J.*, 2:32–37, 2002.
- [167] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono und K. Schulten. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.*, 72:1568–1581, 1997.
- [168] B. Pophof. Moth pheromone binding proteins contribute to the excitation of olfactory receptor cells. *Naturwissenschaften*, 89:515–518, 2002.
- [169] B. Pophof. Pheromone-binding proteins contribute to the activation of olfactory receptor neurons in the silkworms *antheraea polyphemus* and *bombyx mori*. *Chemical Senses*, 29:117–125, 2004.
- [170] N. J. Oldham, J. Krieger, H. Breer, A. Fishedick, M. Hoskovec und A. Svatos. Analysis of the silkworm moth pheromone binding protein - pheromone complex by electrospray-ionization mass spectrometry. *Angew. Chem. Int. Ed. Engl.*, 39:4341–4343, 2000.
- [171] N. J. Oldham, J. Krieger, H. Breer und A. Svatos. Detection and removal of an artefact fatty acid from the binding site of recombinant *bombyx mori* pheromone-binding protein. *Chemical Senses*, 26:529–531, 2001.
- [172] R. Maida, G. Ziegelberger und K. E. Kaissling. Ligand binding to six recombinant pheromone-binding proteins of *antheraea polyphemus* and *antheraea pernyi*. *J. Comparative Physiology B-Biochem. Systemic And Environmental Physiology*, 173:565–573, 2003.
- [173] V. Klusak, Z. Havlas, L. Rulisek, J. Vondrasek und A. Svatos. Sexual attraction in the silkworm moth: Nature of binding of bombykol in pheromone binding protein - an ab initio study. *Chemistry & Biology*, 10:331–340, 2003.
- [174] T. Nakagawa, T. Sakurai, T. Nishioka und K. Touhara. Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science*, 307:1638–1642, 2005.
- [175] U. C. Singh und P. A. Kollman. A combined abinitio quantum-mechanical and molecular mechanical method for carrying out simulations on complex molecular-systems - applications to the  $\text{ch}_3\text{cl} + \text{cl}^-$  exchange-reaction and gas-phase protonation of polyethers. *J. Comp. Chem.*, 7:718–730, 1986.
- [176] J. Baudry, S. Crouzy, B. Roux und J. C. Smith. Quantum chemical and free energy simulation analysis of retinal conformational energetics. *J. Chem. Information Comp. Sciences*, 37:1018–1024, 1997.
- [177] M. Nina, B. Roux und J. C. Smith. Functional interactions in bacteriorhodopsin - a theoretical-analysis of retinal hydrogen-bonding with water. *Biophys. J.*, 68:25–39, 1995.
- [178] M. Said, D. Maynau, J. P. Malrieu und M. A. García Bach. A nonempirical heisenberg hamiltonian for the study of conjugated hydrocarbons. ground-state conformational studies. *J. Am. Chem. Soc.*, 106:571–579, 1984.

## BIBLIOGRAPHY

---

- [179] M. Nina, J. C. Smith und B. Roux. Ab-initio quantum-chemical analysis of schiff-base water interactions in bacteriorhodopsin. *Theochem - J. Mol. Struct.*, 105:231–245, 1993.
- [180] R. M. Ibberson, O. Yamamuro und T. Matsuo. Crystal structures and phase behaviour of acetaldehyde-d(4): a study by high-resolution neutron powder diffraction and calorimetry. *J. Mol. Structure*, 520:265–272, 2000.
- [181] G. L. Gutsev und L. Adamowicz. Electronic and geometrical structure of dipole-bound anions formed by polar-molecules. *J. Phys. Chem.*, 99:13412–13421, 1995.
- [182] A. E. García. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.*, 68:2696–2699, 1992.
- [183] H. Grubmüller, N. Ehrenhofer und P. Tavan. Conformational dynamics of proteins: Beyond the nanosecond time scale. In M. Peyard, editor, *Proceedings of the Workshop ‘Nonlinear Excitations in Biomolecules’, May 30–June 4, 1994, Les Houches (France)*, Seiten 231–240. Centre de Physique des Houches (France), Springer-Verlag, 1995.
- [184] S. T. Wlodek, T. W. Clark, L. R. Scott und J. A. McCammon. Molecular dynamics of acetylcholinesterase dimer complexed with tacrine. *J. Am. Chem. Soc.*, 119:9513–9522, 1997.
- [185] G. Hummer. Fast-growth thermodynamic integration: Error and efficiency analysis. *J. Chem. Phys.*, 114:7330–7337, 2001.
- [186] D. R. Ripoll, C. H. Faerman, P. H. Axelsen, I. Silman und J. L. Sussman. An electrostatic mechanism for substrate guidance down the aromatic gorge of acetylcholinesterase. *Proc. Natl. Acad. Sci. U. S. A.*, 90(11):5128–5132, 1993.
- [187] W. S. Leal. Pheromone reception. *Chemistry Of Pheromones And Other Semiochemicals Ii*, 240:1–36, 2005.
- [188] D. Davenport und J. A. C. Nicol. Luminescence in hydromedusae. *Proc. R. Soc. London Series B - Biol. Sci.*, 144:399–411, 1955.
- [189] R. Y. Tsien. The green fluorescent protein. *Ann. Rev. Biochemistry*, 67:509–544, 1998.
- [190] M. Andresen, M. Wahl, A. Stiel, F. Gräter, L. V. Schäfer, G. Weber, C. Eggeling, H. Grubmüller, S. W. Hell und S. Jakobs. Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proc. Natl. Acad. Sci. USA*, 102:13070–13074, 2005.
- [191] K. A. Lukyanov, A. F. Fradkov, N. G. Gurskaya, M. V. Matz, Y. A. Labas, A. P. Savitsky, M. L. Markelov, A. G. Zaraisky, X. N. Zhao, Y. Fang, W. Y. Tan und S. A. Lukyanov. Natural animal coloration can be determined by a nonfluorescent green fluorescent protein homolog. *J. Biol. Chem.*, 275:25879–25882, 2000.



- [192] D. M. Chudakov, A. V. Feofanov, N. N. Mudriku, S. Lukyanov und K. A. Lukyanov. Chromophore environment provides clue to "kindling fluorescent protein" riddle. *J. Biol. Chem.*, 278:7215–7219, 2003.
- [193] R. Ando, H. Mizuno und A. Miyawaki. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science*, 306:1370–1373, 2004.
- [194] S. Habuchi, R. Ando, P. Dedecker, W. Verheijen, Mizuno H., A. Mizawaki und Hofkens J. Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa. *Proc. Natl. Acad. Sci. USA*, 102:9511–9516, 2005.
- [195] S. W. Hell. Toward fluorescence nanoscopy. *Nature Biotechnology*, 21:1347–1355, 2003.
- [196] M. L. Quillin, D. A. Anstrom, X. K. Shu, S. O'Leary, K. Kallio, D. A. Chudakov und S. J. Remington. Kindling fluorescent protein from *Anemonia sulcata*: Dark-state structure at 1.38 Angstrom resolution. *Biochemistry*, 44:5774–5787, 2005.
- [197] V. E. Zagranichny, N. V. Rudenko, A. Y. Gorokhovatsky, M. V. Zakharov, T. A. Balashova und A. S. Arseniev. Traditional GFP-type cyclization and unexpected fragmentation site in a purple chromoprotein from *Anemonia sulcata*, asFP595. *Biochemistry*, 43:13598–13603, 2004.
- [198] I. V. Yampolsky, S. J. Remington, V. I. Martynov, V. K. Potapov, S. Lukyanov und K. A. Lukyanov. Synthesis and properties of the chromophore of the asFP595 chromoprotein from *Anemonia sulcata*. *Biochemistry*, 44:5788–5793, 2005.
- [199] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez und J. A. Pople. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004.
- [200] J. Saam, E. Tajkhorshid, S. Hayashi und K. Schulten. Molecular dynamics investigation of primary photoinduced events in the activation of rhodopsin. *Biophys. J.*, 83:3097–3112, 2002.

## BIBLIOGRAPHY

---

- [201] M. E. Martin, F. Negri und M. Olivucci. Origin, nature, and fate of the fluorescent state of the green fluorescent protein chromophore at the CASPT2//CASSCF resolution. *J. Am. Chem. Soc.*, 126:5452–5464, 2004.
- [202] F. A. De Wolf und G. M. Brett. Ligand-binding proteins: Their potential for applications in systems for controlled delivery and uptake of ligands. *Pharm. Rev.*, 52:207–236, 2000.
- [203] T. A. Halgren und W. Damm. Polarizable force fields. *Curr. Opin. Struct. Biol.*, 11:236–242, 2001.
- [204] A. Warshel. Computer simulations of enzyme catalysis: Methods, progress, and insights. *Ann. Rev. Biophys. Biomol. Struct.*, 32:425–443, 2003.
- [205] L. Ridder und A. J. Mulholland. Modeling biotransformation reactions by combined quantum mechanical/molecular mechanical approaches: From structure to activity. *Curr. Topics Med. Chem.*, 3:1241–1256, 2003.