

# **New approaches for the analysis of force spectroscopy data and uncovering of free energy landscapes**

Inaugural-Dissertation

zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

**Adeline Bieker**  
aus München

Düsseldorf, November 2014

aus dem Institut für Physikalische Biologie  
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der  
Mathematisch-Naturwissenschaftlichen Fakultät der  
Heinrich-Heine-Universität Düsseldorf

Referent: Dr. Philipp Oesterhelt

Korreferent: Prof. Dr. Dieter Willbold

Tag der mündlichen Prüfung:

*„Es ist nicht genug, zu wissen, man muss es auch anwenden;  
es ist nicht genug, zu wollen, man muss es auch tun.“*

Johann Wolfgang von Goethe  
(aus: Wilhelm Meisters Wanderjahre)

# ERKLÄRUNG

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Düsseldorf, den 20.11.2014

Adeline Bieker

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Physical principles</b>	<b>3</b>
2.1	Atomic force microscope . . . . .	3
2.1.1	The general setup . . . . .	4
2.1.2	Tip-sample interactions . . . . .	6
2.2	Force Spectroscopy . . . . .	8
2.2.1	Recording a force-distance curve . . . . .	9
2.3	The cantilever . . . . .	11
2.3.1	Determination of the spring constant . . . . .	11
2.4	Elasticity of polymers . . . . .	13
2.4.1	Freely-jointed chain model . . . . .	14
2.4.2	Worm like chain model . . . . .	16
<b>3</b>	<b>Biological Principles</b>	<b>19</b>
3.1	Proteins . . . . .	19
3.1.1	The transcription activator PhoB . . . . .	19
3.1.2	Membrane protein Bacteriorhodopsin . . . . .	22
3.1.2.1	Unfolding of BR in a force-distance curve . . . . .	23
<b>4</b>	<b>Materials and Methods</b>	<b>27</b>
4.1	Monte Carlo simulations of force spectroscopy data . . . . .	27
4.1.1	Simulating the setup of cantilever and PEG . . . . .	27
4.2	$\chi^2$ -fits and parameter uncertainties . . . . .	29
4.3	Force spectroscopy measurements on PhoB-DNA complexes . . . . .	30
4.4	Concentration-dependent measurements on BR . . . . .	31
<b>5</b>	<b>Revealing the energy landscape from force spectroscopy data</b>	<b>33</b>
5.1	Model-free calculation of force-dependent dissociation rate . . . . .	37

## Contents

5.2	Determination of the energy landscape parameters . . . . .	38
5.2.1	Testing the analysis strategy with simulated data . . . . .	42
5.2.1.1	Considering one retraction velocity . . . . .	42
5.2.1.2	Considering three retraction velocities . . . . .	46
5.2.2	Applying the model to PhoB-DNA force spectroscopy data . . . . .	51
<b>6</b>	<b>Dependency of the stability of BR on the concentration of compatible solutes</b>	<b>57</b>
6.1	Analysis of each unfolding step . . . . .	58
6.1.1	Uncertainty of the forces derived from simulation . . . . .	60
6.1.2	Uncertainty of the persistence length . . . . .	64
6.1.2.1	Influence of the WLC fit on the obtained persistence length . . . . .	64
6.2	Results and discussion . . . . .	66
	Publication . . . . .	69
<b>7</b>	<b>Conclusion and Outlook</b>	<b>79</b>
7.1	Analysing single-molecule force spectroscopy data . . . . .	79
7.2	Differentiation of the single unfolding steps to determine the concentration-dependent influence of compatible solutes . . . . .	80
	<b>Summary</b>	<b>83</b>
	<b>Zusammenfassung</b>	<b>84</b>
	<b>Appendix</b>	<b>87</b>
	Appendix A . . . . .	87
	<b>Danksagung</b>	<b>90</b>
	<b>Bibliography</b>	<b>93</b>

# 1 Introduction

The cell is the structural unit of life. It is highly complex and organized. For understanding the metabolic processes and related diseases of an organism, investigating the protein-protein and DNA-protein interactions on their structural level is essential. The possibility to have an impact on metabolic processes opens a broad field of drug development, aimed at specific targets to prevent unwanted interactions causing side effect.

One of the first developed biopharmaceutical drugs was human insulin for patients suffering from diabetes mellitus whose metabolism is not able anymore to produce its own insulin or just produces it with an reduced effectiveness in the islets of Langerhans in the pancreas. Challenging are diseases with an overproduction of misfolded proteins. One example is the amyloid beta peptide involved in Alzheimer's disease. This peptide is self-aggregating, leading to oligomeres which are believed to be toxic and leading to the damage of neurons [BKS12]. To stop or delay this process is only possible by understanding how the amyloid beta peptide interacts with itself.

The invention of the atomic force microscope [BQG86] was an important milestone for investigating living biological samples in aqueous solutions. This microscope allows also force spectroscopy measurements for observing the force-dependent dissociation of two single molecules like a receptor-ligand pair, or the stretching of a single, chain-like macromolecule. The first receptor-ligand pairs investigated by this technique were biotin and avidin by Florin et al. [FMG94] and biotin and streptavidin by Lee et al. [LKC94]. They determined the dissociation force of the complex which can not be determined in ensemble measurements.

Evans and Ritchie proposed that this forced dissociation process can be viewed as a thermally activated escape over a potential barrier of the energy landscape describing the molecular interaction [ER97]. This thermally activated escape is ubiquitous in biological, physical and chemical processes and its understanding therefore of importance [HTB90]. The aim of this work is to develop a model to reveal the free energy landscape of the molecular interaction from force spectroscopy data and obtain a pos-

## 1 Introduction

sible curvature of the free energy landscape. Furthermore I discuss the reliability of this reconstruction from simulated single-molecule force spectroscopy experiments, dependent on the accessible force range, by considering a single or multiple retraction velocities. To test my analysis strategy I analyse the data of the intermolecular interaction of the point mutant R176A of the protein PhoB with DNA from the bacteria *Escherichia coli*.

Force spectroscopy measurements allow also to determine the elasticity of the stretched molecule. This is used to identify the influence of compatible solutes on each unfolding step, recorded in a force-distance curve, of the membrane protein Bacteriorhodopsin. To determine the reliability of the changes in the dissociation force, I calculate statistical errors for the measured force histograms including the thermal influence.

I start with an introduction to the physical principles, focusing especially on the atomic force microscope and force spectroscopy measurements. Furthermore I explain the determination of the spring constant of the cantilever and the modelling of polymers which is essential for the analysis of force-distance curves. The data used and analysed to test my analysis strategy is from measurements on the transcription factor PhoB. This protein and the investigated membrane protein Bacteriorhodopsin are presented in chapter 3 describing the biological principles. The simulation of force spectroscopy data is explained in the materials and methods in chapter 4. The results of this work are given in the following two chapters, before concluding in chapter 7 with an outlook on future perspectives.

## 2 Physical principles

In this chapter I introduce the physical bases for single-molecule force spectroscopy: the atomic force microscope and the measurable forces, the performing of force spectroscopy measurements, the determination of the spring constant of the cantilever. I conclude with polymer models which are necessary for the analysis of force spectroscopy data.

### 2.1 Atomic force microscope

The invention of an absolute new approach to image and measure surfaces by Binnig and Rohrer [BR86] in 1981 lead to a revolution in investigating materials: the information is not gained by an electromagnetic wave as before in light and electron microscopy, but by an interaction between a sensor with a needle-like tip and the surface. The first interaction used was a tunneling current between the sensor and a surface which is why this method was called Scanning Tunneling Microscope. The investigated surfaces had to be made from or covered with conducting material and the measurements had to be performed in air or vacuum. With this method it was possible to reach an atomic resolution for the first time which earned the Nobel prize in 1986.

Soon another interaction was used, namely the force which acts between the spring like sensor and the surface. This allows to perform measurements in aqueous solutions which broadens the range of investigated materials to living biological samples in buffer reaching a resolution underneath the diffraction limit of optical systems. This method is called atomic force microscopy, invented by Binnig, Quate and Gerber [BQG86].

In this chapter I will introduce the physical background of atomic force microscopy, its general setup, and the measurable forces.

### 2.1.1 The general setup

The underlying principle of the atomic force microscope (AFM) is the detection of the mechanical force acting between a sample and a sensor. This microfabricated sensor is a so-called cantilever which consists of a rectangular or triangular beam with a pyramidal or cone-shaped tip at the end. By approaching this cantilever towards a sample it deflects according to Hooke's law when the tip gets in proximity of the surface. While scanning with the tip over the sample, the cantilever deflection changes due to the topography of the sample. A topographic image of the sample is therefore obtained by plotting the deflection of the cantilever versus its position on the sample [BJ95]. To that aim the atomic force microscope has to consist out of two main components:

1. the piezoelectric scanner which allows movements of the cantilever with sub-nanometer accuracy in all three dimensions. This scanner is made from a piezoelectric material which expands or contracts proportionally to an applied voltage. For each movement on one of the three axes the scanner has one independent piezoelectric crystal.
2. a detection system to measure the deflection of the cantilever. Typically an optical method is used which consists out of a laser and an array of four photodiodes. The laser beam is focused on the backside of the free end of the cantilever. The backside is made of silicium or coated with gold and therefore the laser is reflected and guided to the position-sensitive photodiode [Duf02].

Such a general setup of an AFM is shown in figure 2.1.

In the following the detection of the movements by the photodiode is explained in more detail: The photodiode has four detection fields. Each field translates the impinging light into an electric current via the photoeffect. For the detection of the movements of the cantilever the following voltages are observed, see the labeling in fig 2.1:

## 2 Physical principles

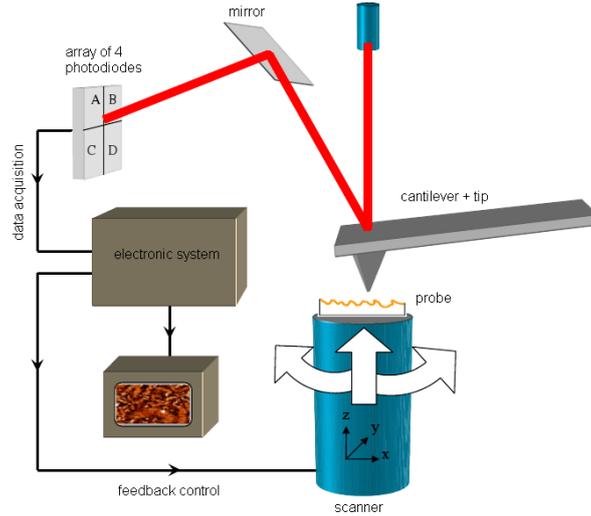


Figure 2.1: Setup of an AFM [Har09]. The labeling of the four photodiode arrays is shown.

- The sum of the four voltage fields is

$$U_{sum} = U_A + U_B + U_C + U_D. \quad (2.1)$$

This parameter has to be maximised at the beginning of each measurement by adjusting the position of the laser spot on the cantilever.

- The vertical deflection is often called the *deflection signal*.

$$U_{deflection} = U_A + U_B - (U_C + U_D). \quad (2.2)$$

For calibrating the photodiode, the fields get positioned such that the laser spot is in the middle between the both upper and lower fields. So at the beginning the signal is  $U_{Deflection} = 0$ .

- A horizontal deflection of the cantilever gets detected by the *lateral signal*

$$U_{lateral} = U_A + U_C - (U_B + U_D). \quad (2.3)$$

For calibration the laser spot has to be positioned to the middle between the both left and both right fields. Then the lateral signal is  $U_{Lateral} = 0$ .

When a force is acting on the cantilever, it deflects which leads to a shift of the position of the laser spot on the photodiode fields, thereby inducing  $U_{deflection}$  and  $U_{lateral}$ . Furthermore the *deflection signal* is used to control the movement of the cantilever in

## 2 Physical principles

vertical direction, see figure 2.1. When the forces acting on the cantilever exceed a defined value, the controller retracts the cantilever from the surface until the deflection reaches an acceptable value again. This feedback loop saves the cantilever from damage.

Because in this work the imaging modes of the AFM are not used I will not explain them here but refer for more information to the literature like for example [EW10]. In the next section I will explain which forces the AFM can detect.

### 2.1.2 Tip-sample interactions

By approaching the cantilever towards a surface, several forces of different strength and interaction range act on the cantilever. By these attractive or repulsive forces the cantilever gets deflected out of his idle position. In figure 2.2 the interaction range of the different forces is shown. In the following the cause of these forces is explained in order of their interaction range.

#### Fluid film damping

This boundary only occurs for measurements in Tapping Mode [Eli95, EW10]. A damping layer of air is developed when the oscillating cantilever comes to within 10 microns of the sample surface. While the cantilever moves downwards during the oscillation, the air is squeezed and while moving upwards a partial vacuum occurs. This phenomenon disappears when the cantilever passes this boundary.

#### Electrostatic forces

Between two electric charges  $Q_1$  and  $Q_2$  which have a distance  $r$  from each other acts the Coulomb force

$$|\vec{F}_C| = \frac{1}{4\pi\epsilon_0\epsilon} \frac{Q_1 Q_2}{r^2}. \quad (2.4)$$

$\epsilon_0$  is the dielectric constant of vacuum and  $\epsilon$  the one of the surrounding media. Electric charges can accumulate on the cantilever and the sample surface and interact with each other. This electrostatic interaction can be influenced by the choice of the cantilever material, the material of the surface or by the surrounding media. They get reduced by conducting material. In liquids they can get controlled by the concentration of ions.

## 2 Physical principles

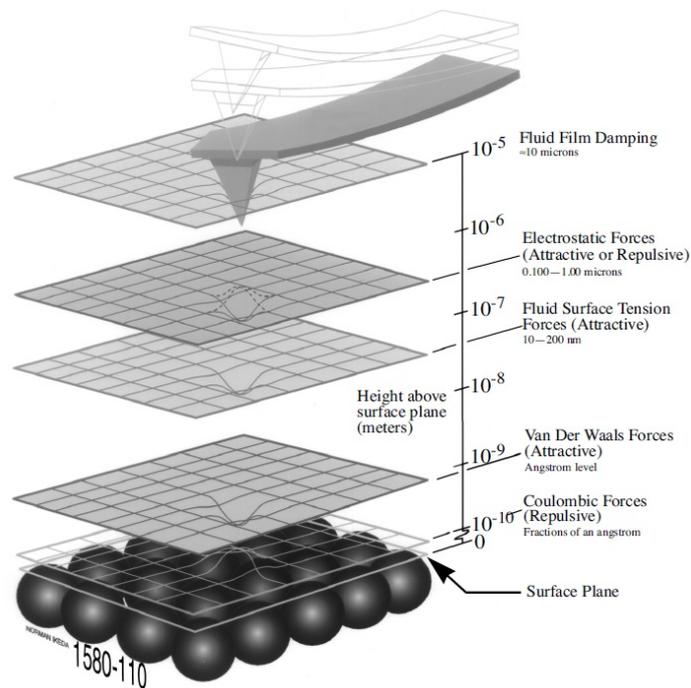


Figure 2.2: Interaction range of the different interaction forces [Ell95]: The electrostatic forces occur with this interaction length in air. In liquids the interaction length is shorter. The fluid film tension does not occur in liquids.

### Fluid surface tension

On air every surface is covered by a thin water layer. A capillary tube is formed when the cantilever is 10-200 nm above the surface. This leads to an additional attractive force. This force is in the range of 10 nN to 100 nN [JNOF00]. This effect does not occur here because the measurements used in this work are made in liquids.

### Van der Waals forces

The following forces are combined to the Van der Waals force:

- the interaction between two dipoles (dipole dipole force),
- the interaction between a dipole and an induced dipole in a polarisable molecule (dipole induced-dipole force), and
- the dispersion force which is the interaction between two induced dipoles (induced-dipole induced-dipole force).

## 2 Physical principles

These interactions have in common that the interaction energy is inversely proportional to the 6th power of the distance  $r$  between the atoms or molecules:

$$V(r) = -\frac{C}{r^6}, \quad (2.5)$$

with a constant  $C$ . For the dissociation of molecules which are bound to each other by Van der Waals force an energy between 1-2 kJ/mol is needed [Toe04]. Thereby the Van der Waals force is a relatively weak, non-covalent interaction with a low interaction range of Angström level [Ell95].

An example for the effect of Van der Waals forces is graphite. The layers of C-atoms are hold together by this force. That these forces are very weak can be recognised in that while writing with a pencil the layers remove easily.

### Pauli repulsion (Coulombic forces)

The Pauli principle says that electrons which are at the same place have to differ at least in one quantum number. When the tip of the cantilever touches the surface the electron orbitals overlap. While overlapping of fully loaded electron orbitals the electrons are forced to jump to free orbitals of higher energy. The increase of energy of the system leads to a strong repulsive force. The interaction range is less than 0.1 nm [Bar06]. Pressing the cantilever on the surface beyond this level leads to a damage of the cantilever.

## 2.2 Force Spectroscopy

With an atomic force microscope it is not only possible to image surfaces but also to determine forces which act between the cantilever and the surface. For that purpose the cantilever is regarded as a spring . With Hooke's law,

$$\vec{F} = -k_{cant} \Delta\vec{z}, \quad (2.6)$$

it is possible to determine the force  $F$ . The spring constant of the cantilever is  $k_{cant}$  and  $\Delta z$  is the displacement of the cantilever out of its idle position. An atomic force microscope allows to detect forces in the pico-Newton range.

The force spectroscopy has a broad range of application, for example one can investigate the unfolding of proteins (see chapter 6) or the interaction between a receptor-ligand-pair (see section 5.2.2). For the determination of specific dissociation events of

## 2 Physical principles

the molecules under investigation, they have to be bound covalently to the surface and to the cantilever.

For the investigation of receptor-ligand interactions, like the protein PhoB and DNA used in this work, the protein is immobilised to the surface, while the DNA gets bound to the cantilever. To observe the specific dissociation events and to discriminate them from other events like adhesion, polyethyleneglycol (PEG) linkers are used between the DNA and the cantilever. The length of the PEG-linker used in this work is around 27 nm. The DNA gets a higher steric flexibility from the PEG-linker which means that the binding possibility of the DNA to the protein gets increased due to the larger spatial accessibility. The stretchability of the PEG leads to a specific dissociation event in the force-distance-curve which allows the discrimination of this specific event from unspecific interactions.

### 2.2.1 Recording a force-distance curve

In the following the principle of force-distance curves is explained using the example of DNA and protein as a receptor-ligand interaction. The functionalised cantilever is approached to the surface. At the same time the deflection of the cantilever is detected in dependency of the piezoposition which corresponds to the distance between tip and surface, see figure 2.3.

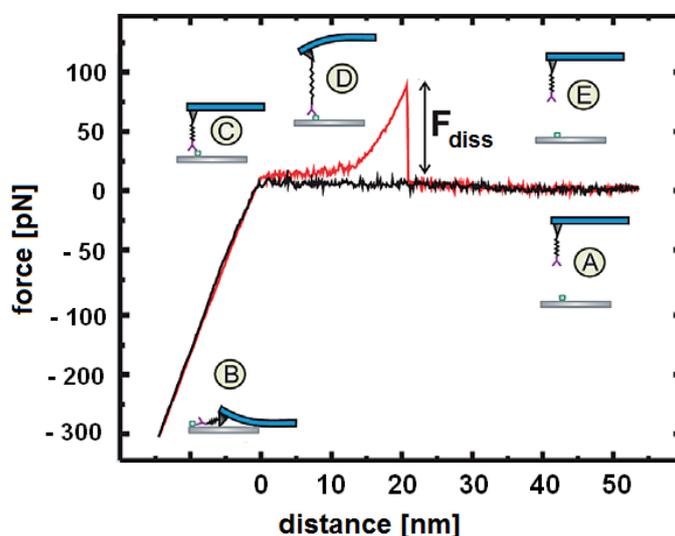


Figure 2.3: Force-distance curve, adapted from [EGH06]. Explanations are in the text.

## 2 Physical principles

- A. Before the contact of tip and surface the cantilever is not deflected because no force acts on it.
- B. When the tip touches the surface, the cantilever starts to bend upwards until the maximally allowed force is reached. During the contact the molecules can bind to each other. Then the cantilever is removed from the surface.
- C. At first the cantilever goes back in its idle position. If a bond was formed during the contact, the PEG-linker is stretched while retracting the cantilever and the cantilever starts to bend downwards.
- D. The increasing force on the complex leads to a dissociation of DNA and protein. The cantilever jumps back to its idle position.
- E. With increasing displacement to the surface the cantilever remains in its idle position.

The bending of the cantilever towards the surface is displayed as positive values. Between steps C and D unspecific interactions, so-called adhesion events, can occur. In this case the tip adheres to the surface and the cantilever bends upwards. Once the retraction force is stronger than the adhesion force the cantilever jumps back to its idle position. If no interaction occurs the cantilever goes back to its idle position immediately after leaving the surface. In figure 2.4 a force-distance-curve with an adhesion event and a following dissociation event of a DNA-protein-complex is shown.

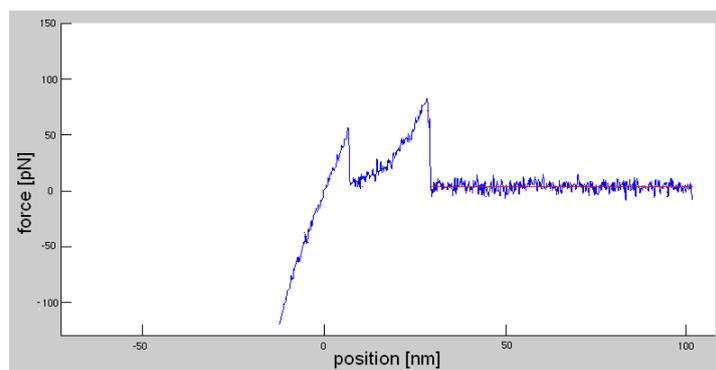


Figure 2.4: *Force-distance curve with an adhesion event (from my Master thesis) [Bie11].*

The force can be determined from the deflection signal using the spring constant, see the next section. How a force curve of the unfolding of a protein looks like is

described in subsection 3.1.2 after introducing the three-dimensional structure of the protein Bacteriorhodopsin which is essential to understand the force curve.

### 2.3 The cantilever

The investigation of the interaction of single molecules needs the application and detection of forces on a molecular complex. Usually the needed element for this investigation has a microscopic dimension [BMW00].

For the AFM this central element is the cantilever. It gets produced industrially and is available with a great variety of properties like size, geometry, material and thereby spring constant and resonance frequency. These properties have to be adjusted to the underlying system of research and the forces to be measured. If an unapproximated cantilever is used it is possible that the interaction of interest can not be measured.

To give an example the bond strength of a noncovalent bond like between protein and DNA is in the pico-Newton range. So the spring constant of the cantilever has to be in the same range of pN/nm. If the spring constant is higher, the interaction is not measurable because the cantilever is too stiff to show a bending. A detailed description of the motion of the cantilever can be found in [Bie11].

#### 2.3.1 Determination of the spring constant

The cantilever is used as a spring to detect forces via Hooke's law as mentioned before in section 2.2. The displacement  $\Delta z$  is measured so the spring constant has to be known in order to estimate the force.

In the production of cantilevers the lateral dimensions are well defined but the vertical thickness is subject to fluctuations. Since the spring constant depends on the third power of the thickness, this leads to significant uncertainty in the spring constant [SLMW95, CS05]. Therefore the spring constant of each cantilever has to be determined experimentally. For this determination several methods were developed depending on the dimension of the spring constant. I introduce here the Thermal Noise Method invented by Hutter and Bechhoefer [HB93] which is mostly used for cantilevers with a pN/nm range in AFM force spectroscopy measurements. More information about other methods can be found for example in [HB93, LM02].

## 2 Physical principles

The Thermal Noise Method relies on the equipartition theorem which says that the thermal energy per degree of freedom is  $k_B T/2$ . Assuming the cantilever to be an ideal spring with spring constant  $k_{cant}$ , the cantilever can be described as a harmonic oscillator with the Hamilton function

$$\hat{H} = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 q^2. \quad (2.7)$$

$q$  is the generalised space coordinate and  $p$  is the momentum. With the equipartition theorem follows

$$\frac{1}{2}m\omega_0^2 \langle q^2 \rangle = \frac{1}{2}k_B T. \quad (2.8)$$

With

$$\omega_0 = \sqrt{\frac{k_{cant}}{m}} \quad (2.9)$$

this results in

$$k_{cant} = \frac{k_B T}{\langle q^2 \rangle}. \quad (2.10)$$

Therefore the measurement of the thermal fluctuations of the free end of the cantilever allows the determination of the spring constant. The thermal fluctuations are the displacements of the cantilever out of his resting position due to the brownian motion of the surrounding medium which is also called thermal noise.

The mean square displacement  $\langle q^2 \rangle$  corresponds to the area under the resonance peak in the thermal noise spectra. For this spectra the squared amplitude is plotted in dependency on the frequency less the white noise background. The peak in this spectrum has a Lorentzian profile (a mathematical derivation can also be found in [Bie11]). By a Lorentzian fit to this peak the area under the peak which is used as a measure of the energy in the resonance and so  $\langle q^2 \rangle$  can be determined.

The photodiode detects voltage changes. Therefore at first the measured energy spectrum has the unit  $V^2/Hz$  what has to be converted into  $nm^2/Hz$ . For that purpose a force-distance curve has to be taken. In the region of the force curve where the tip is in contact to the surface the force curve shows a straight line with a slope in the unit  $V/nm$ . This has to be done for each cantilever because the slope differs for each cantilever due to vertical thickness fluctuations. With the reciprocal of this parameter follows

$$\text{Deflection}[nm] = \text{InvOLS} \left[ \frac{nm}{V} \right] \cdot \text{Deflection}[V]. \quad (2.11)$$

*InvOLS* the so-called inverse optical cantilever sensitivity. So the spring constant can be determined by

$$k_{cant} = \frac{k_B T}{\langle U^2 \rangle \text{InvOLS}^2 \chi^2}, \quad (2.12)$$

## 2 Physical principles

with the mean square cantilever displacement  $\langle U^2 \rangle$  in Volt.  $\chi^2$  is a correction factor which contains different contributions, for example from the fact that the photodiode detects in the strict sense the decline of the end of the hanger of the spring and not the displacement of the cantilever [BJ95]. Furthermore the shape of the cantilever is considered in the correction factor. Normally the first resonance frequency is used for the calibration of the cantilever because it has the best signal to noise ratio. For V-shaped cantilever (triangular beam) which are used here the correction factor is 0,97 [SDH01] for the first resonance frequency.

The standard deviation of the Thermal Noise Method is around 10-15% just like the other methods. The advantage is that the cantilever is not damaged during this calibration.

When the spring constant has been determined the experimental curves can be converted to the force signal and analysed further. The next step in the analysis is the evaluation of the force curves because of the need to distinguish between different events. For example adhesion events should be discarded. The difference of these events is their stretching behaviour. In the following section I therefore introduce the mathematical description of polymer stretching.

### 2.4 Elasticity of polymers

In the analysis and simulation of force spectroscopic data it is important to be able to describe mathematically the stretching behaviour of polymers. In the experimental setup polymers are used because they allow to distinguish specific dissociation events from adhesion events due to their specific stretching behaviour. Also the number of simultaneously bound polymers can be determined. In the simulation the polymer models are necessary for a realistic reproduction of the experimental setup.

In the following I introduce two basic polymer models and their extensions which are often used in AFM force spectroscopy for various biological and synthetic polymers ([JNOF00, LRO<sup>+</sup>99, ROHG97, ORG99, RGO<sup>+</sup>97]). Furthermore I give an overview of their applications in different force spectroscopy measurements.

### 2.4.1 Freely-jointed chain model

The freely-jointed chain (FJC) is the simplest model of a polymer. The polymer segments have a fixed length and are linearly connected. All bond and torsion angles are equiprobable. In this model there are no interactions between the segments apart from their connectivity between neighbours and the segments can penetrate right through each other. Therefore this model of a polymer can be described by a simple random walk.

An FJC polymer is an entropic spring: there are fewer ways to arrange the chain the more it is extended, leading to a decrease in entropy and thus an increase in free energy. In other words, the restoring force increases as the length increases.

In the discrete FJC, the segment length is called Kuhn length  $l_K$  and describes the stiffness of the chain. The larger the Kuhn length, the stiffer is the chain.

The dependency of the extension  $x$  on an external force  $F$  acting on the polymer is based on the extended Langevin function ( $L(x) = \coth(x) - 1/x$ ):

$$x(F) = L_c \cdot L\left(\frac{Fl_K}{k_B T}\right), \quad (2.13)$$

introduced by Smith et al. [SCB96]. In Eq. (2.13)  $L_c$  is the contour length of the completely stretched filament.

Force spectroscopy measurements on dextran by Rief et al. [ROHG97] necessitated an expansion of this model by an enthalpic elasticity due to a detected stretchability of the monomers. They introduced the modified freely-jointed chain by adding a further extension of a segment elasticity  $k_s$  to Eq. (2.13):

$$x(F) = L_c \left[ \coth\left(\frac{Fl_k}{k_B T}\right) - \frac{k_B T}{Fl_k} \right] + \frac{nF}{k_s}. \quad (2.14)$$

$n$  is the number of monomers in the filament. Here both  $n$  and the segment elasticity  $k_s$  are related to monomers and not to the Kuhn segments.

#### PEG

The polymer which is mostly used in receptor-ligand force spectroscopy measurements is polyethylene glycol (PEG), as already mentioned in section 2.2.1. The first measurements on pure PEG were done by Oosterhelt et al. [ORG99]. To describe the polymer extension behaviour the FJC model needs a further expansion. Oosterhelt et al. discover a conformational transition of the monomers which can be explained by a shift from a shorter helical state towards an elongated planar state. The new contour

## 2 Physical principles

length describing the PEG is then given by:

$$L_C = N_{planar} \cdot L_{planar} + N_{helical} \cdot L_{helical}. \quad (2.15)$$

Furthermore is the number of monomers constant

$$N_S = N_{planar} + N_{helical} \quad (2.16)$$

and the ratio of the populations  $N_{helical}/N_{planar}$  is Boltzmann distributed

$$\frac{N_{helical}}{N_{planar}} = e^{\frac{\Delta G}{k_B T}}. \quad (2.17)$$

The applied force alters the difference in free energy according to:

$$\Delta G(F) = (G_{planar} - G_{helical}) - F \cdot (L_{planar} - L_{helical}). \quad (2.18)$$

Together with Eq. (2.14) this results in the following extension-force relation:

$$x(F) = N_s \cdot \left( \frac{L_{planar}}{e^{-\Delta G/k_B T} + 1} + \frac{L_{helical}}{e^{\Delta G/k_B T} + 1} \right) \cdot \left[ \coth \left( \frac{Fl_k}{k_B T} \right) - \frac{k_B T}{Fl_k} \right] + \frac{nF}{k_s}. \quad (2.19)$$

The Kuhn length  $l_k$  for a PEG linker is 0.7 nm and its segment elasticity  $k_s$  is 150 N/m per monomer.  $L_{planar}$  is 3.58 Å,  $L_{helical}$  is 2.7 Å and  $\Delta G$  is 3  $k_B T$  [ORG99]. Remarkably this model describes with this extension-force relation quite well the behaviour of the real PEG linker.

In experiments often not just one single PEG linker binds between the cantilever and the surface, it is necessary to take this into account. With the assumption that the force is equally distributed to all  $N_{\#}$  simultaneously bound PEG linkers the force  $F$  in Eq. (2.19) is replaced by  $F/N_{\#}$  [FNY12].

$$x(F) = N_s \cdot \left( \frac{L_{planar}}{e^{-\Delta G/k_B T} + 1} + \frac{L_{helical}}{e^{\Delta G/k_B T} + 1} \right) \cdot \left[ \coth \left( \frac{Fl_k}{N_{\#} \cdot k_B T} \right) - \frac{N_{\#} \cdot k_B T}{Fl_k} \right] + \frac{nF}{N_{\#} \cdot k_s}. \quad (2.20)$$

With this equation the experimental force-distance curves will be fitted, using the number of bound PEGs  $N_{\#}$  and the number of monomers  $n$  as fitting parameters. All other parameters are well known and determined by experiments for PEG linker [ORG99]. In figure 2.5 the influence of the number of bound PEGs on the shape of a force-distance curve is shown; this illustrates that the impact on the shape is high enough to distinguish between the different numbers of bound PEG.

## 2 Physical principles

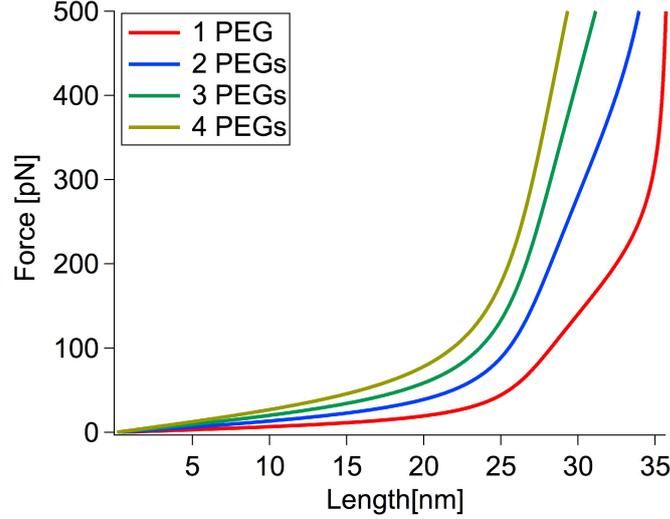


Figure 2.5: Force-extension-curves for different numbers ( $N_{\#}=1, 2, 3, 4$ ) of simultaneously bound and dissociated PEGs.

This extended FJC model is applicable to analyse experimental force spectroscopy data from receptor-ligand interactions when PEG linkers are used and dominate the stretching in the force-distance curve. In the next section another model is introduced which is better suited to describe the stretching in the force-distance curves of biopolymers like DNA or the unfolding of proteins.

### 2.4.2 Worm like chain model

Biopolymers consist of monomers connected by chemical bonds, which have fixed bond angles to each other and are thereby limited in their three-dimensional movements. Therefore the freely-rotating chain model assumes that the polymer consists of  $N$  segments with the length  $l$  which have a fixed angle  $\Theta$  to each other but are able to rotate freely around their longitudinal axis. The polymer has a total length  $L$  and a so-called persistence length  $l_p$  which is defined as the sum of projections of the bond vectors of all conformation to the first vector:

$$l_p = l \sum_{j=2}^N \cos \Theta_j. \quad (2.21)$$

The persistence length is a parameter describing the stiffness of the polymer which can be visualised by how long the orientation of the chain persists through its length.

## 2 Physical principles

Now taking the limit  $N \rightarrow \infty$ ,  $l \rightarrow 0$  and therefore  $\Theta \rightarrow 0$ , but keeping a constant total length  $L$  and constant persistence length  $l_p$  leads to the worm-like chain (WLC) model. Therefore the characteristic parameters describing the polymer are the total length  $L$  and the persistence length  $l_p$  (these parameters correspond in the FJC model to the number of segments  $N$  and the Kuhn length  $l_K$ ). The WLC is a continuous freely rotating chain and is also called Kratky-Porod model, detailed information to this model can be gained for example from [Yam97].

The force-extension relation for the WLC is [MS95]

$$F(x) = \frac{k_B T}{l_p} \left[ \frac{1}{4 \left(1 - \frac{x}{L}\right)^2} + \frac{x}{L} - \frac{1}{4} \right]. \quad (2.22)$$

This model is used for describing the stretching of molecules like double-stranded DNA, RNA and proteins. It is applied in chapter 6 to fit the unfolding events of the membrane protein Bacteriorhodopsin to identify changes in the persistence length. In this project we investigate the dependency of each unfolding step of Bacteriorhodopsin on the concentration of compatible solutes.



## 3 Biological Principles

In this section I give an overview about the used biological systems in this work.

### 3.1 Proteins

Considering the dry weight of normal cells, the quantity of protein is more than 50 % [BS99]. Therefore proteins are one of the most prevalent biological molecules. They are modular built from single components, the amino acids, whose sequence determines essentially the structure and therefore characteristics of the protein. Their tasks vary from biocatalytics over regulatory transmission of signals to structure and stability of cells.

Below I introduce the transcription activator PhoB and afterwards the transmembrane protein Bacteriorhodopsin, both relevant for the present work.

#### 3.1.1 The transcription activator PhoB

Almost any regulatory process in an organism involves specific recognition and binding of one molecule to another, for example a protein to DNA, which regulates the complete cell cycle and gene expression. The gene expression is the process by which the genetic information is translated to the corresponding gene product, for example a protein. Due to changes of the environment, for example a varying supply of nutrients, it must be possible for a cell to adjust the gene expression. This process is called gene regulation. On each step of the gene expression regulatory factors can influence and control this process. The most important step is for most of the genes the start of the transcription [Wat11]. This is the point to decide whether a gene is expressed

### 3 Biological Principles

or not. Regarding energy efficiency this is the most favourable point for regulation because a waste of energy and nutrients is avoided.

This decision is not made on the gene itself but on the promotor which is close to the gene. A key protein can bind to the promotor and change the conformation of the DNA to enable or suppress the transcription. These key proteins are called transcription factors. A blocking transcription factor is called transcription repressor and an activating one transcription activator.

I used in the project for our new model and analysis procedure the data of the receptor-ligand interaction of the transcription activator PhoB with DNA to test the model. In the following I give a short overview about this protein, for more information about the system and the data see [Bie11, RWK<sup>+</sup>13].

For the bacteria *Escherichia coli* (*E. coli*) more than 30 genes coding for proteins which are responsible for the admission and utilisation of phosphorus are regulated by the two-component system PhoR-PhoB [HW10]. Phosphorus is an essential element in an organism. It is needed for the synthesis of DNA, carbohydrates and membrane lipids, as well as for the energy metabolism in form of adenosine triphosphate (ATP). The concentration of phosphorus in the environment of *E. coli* is detected by the sensor kinase PhoR which is anchored in the cell membrane and controls the activity of PhoB.

When the phosphorus concentration outside the bacteria is less than  $4 \cdot 10^{-6}$  mol/L [Wan93, BSMS05], PhoR is present in activated form. In this conformation PhoR transfers a  $\gamma$ -phosphate residue of one ATP to PhoB. This phosphorylation changes the conformation of PhoB in the way that two PhoBs can bind to the binding site of the DNA. This induces a conformational change of the binding site of the DNA and enables the start of the transcription. When the external concentration of phosphorus increases PhoR dephosphorylates PhoB. Thereby the binding affinity of PhoB to DNA is lowered and therefore the transcription of associated genes decreases.

The transcription activator PhoB is built of 229 amino acids which are divided into two structural domains, see figure 3.1. One is the so-called regulatory phosphorylation domain in the N-terminus of amino acids 1-124. PhoR can interact with this domain and induce the conformational changes so that PhoB is present in activated or inactivated form. The second domain is the DNA-binding domain (DBD) in C-terminus of amino acids 131-229. The secondary structure of the DBD consists of 7

### 3 Biological Principles

$\beta$ -sheets and 3  $\alpha$ -helices in the order of

$$\beta^1 \beta^2 \beta^3 \beta^4 \alpha^1 \beta^5 \alpha^2 \alpha^3 \beta^6 \beta^7$$

[Wan93].

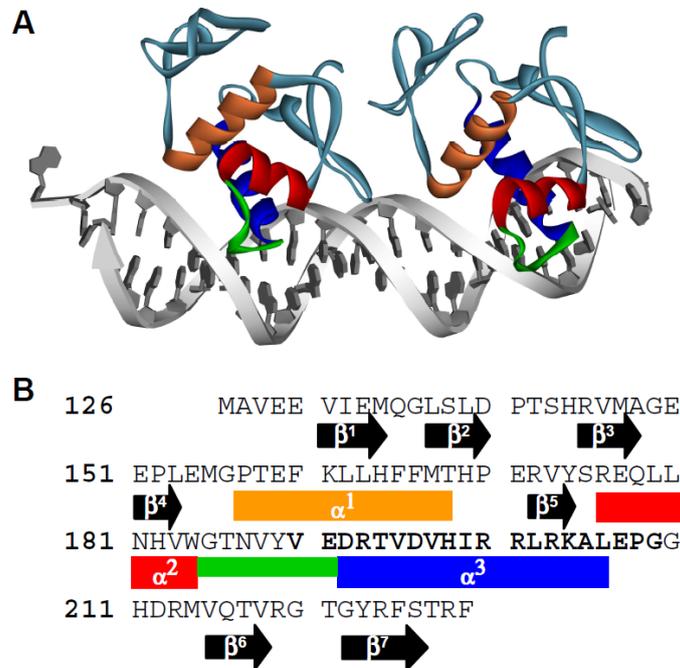


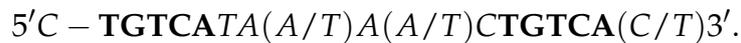
Figure 3.1: *A*: Crystall structure of the binding of two PhoB-DBDs to the *pho* box, see text. The recognition helix  $\alpha^3$  is blue,  $\alpha^1$  orange and  $\alpha^2$  red, while the loop is green. *B*: The amino acid sequence of the DBD. Arrows depict  $\beta$ -sheets and boxes  $\alpha$ -helices [Wol09].

The interaction of PhoB with the DNA occurs via a *winged helix-turn-helix* motif. This motif consists of two  $\alpha$ -helices which are connected via a turn of a short sequence of amino acids. The binding to the DNA occurs via a recognition helix which interacts with the DNA-bases. This bonding is stabilised by a second helix.

For PhoB the recognition helix is the  $\alpha^3$ -helix and  $\alpha^2$  the stabilising one. The turn is replaced by a transactivation loop between  $\alpha^3$  and  $\alpha^2$  and is responsible for the interaction with the RNA-polymerase.  $\beta^6$  and  $\beta^7$  form a structure which binds to the minor groove of the DNA.

### 3 Biological Principles

The binding section of the DNA is called *pho* box and contains two binding sites with the specific binding sequence *TGTC*A. Two PhoBs bind as a dimer to the *pho* box. The DNA-sequence of the *pho* box is [MAK<sup>+</sup>96]



The binding of PhoB to the *TGTC*A-motif in the major groove and the A/T-rich minor groove is essential for the overall recognition process [RWK<sup>+</sup>13].

#### 3.1.2 Membrane protein Bacteriorhodopsin

Each cell is a structural well-defined unit enclosed by a cell membrane. This membrane separates the cytosolic components from the extracellular environment and forms a physical barrier. The membrane consists of a lipid bilayer containing specialised proteins which enable the cell for example to exchange substances, to communicate with other cells and to detect changes in the environmental conditions. These so-called membrane proteins are essential for the controlling of metabolic processes including energy production, photosynthesis and osmotic balance.

Approximately one third of all proteins are membrane proteins. Most pharmaceutical drugs use membrane proteins as target but little is known about the three-dimensional structure, the folding process and the interaction with other molecules like drugs of these proteins. This lack of knowledge is caused by the difficulty to isolate and purify these hydrophobic proteins. They are not water soluble and they denature irreversible outside the membrane.

*Bacteriorhodopsin* (BR) of the photosynthetic bacteria, *Halobacterium salinarium*, is one of the most studied transmembrane proteins. This results from the advantage that this protein forms trimers which are packed into a two-dimensional hexagonal crystal in the cell membrane, called purple membrane. BR is a light-driven proton pump that captures light and converts the electromagnetic energy into an electrochemical proton gradient which is used in the ATP metabolism. This process needs a retinal which upon light activation is isomerised and leads to a conformational rearrangement of the protein. This results in a pumping of one proton into the cell.

The retinal is embedded in seven closely packed  $\alpha$ -helices named from A to G, see figure 3.2 B. Bacteriorhodopsin has 248 amino acids whose sequence is displayed in

### 3 Biological Principles

figure 3.2 A. On the cytoplasmic side of the membrane are the loops connecting the  $\alpha$ -helices A-B, C-D, E-F as well as the C-terminus. This is the side which is needed for the force spectroscopy experiments.

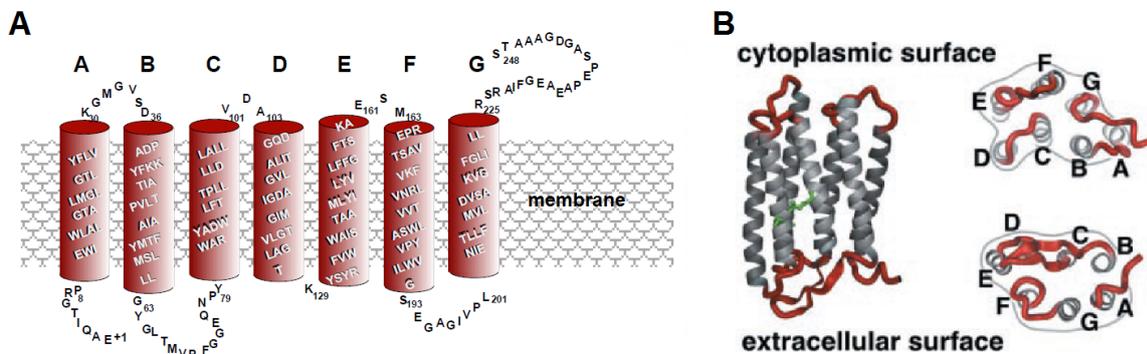


Figure 3.2: Model of the three-dimensional structure of BR. (A) BR is a 248 amino acid membrane protein that consists of seven transmembrane  $\alpha$ -helices, which are connected by loops. (B) Three-dimensional model and top and bottom view show spatial arrangement of the helices. Helices F and G are neighboring helices A and B and thus can stabilize them [OOP<sup>+</sup>00].

In the next section I give an overview of a BR force curve, for more details see inter alia [OOP<sup>+</sup>00, Oes00].

#### 3.1.2.1 Unfolding of BR in a force-distance curve

As explained in the following, a force curve of BR looks quite different from the already explained PEG-based receptor-ligand force curve, see section 2.2.1, with one dissociation event.

The tip is approached towards the surface. During the contact the protein can bind unspecifically to the tip with the parts of the protein on the cytoplasmic side of the membrane. This induces strict criteria for the selection of the force curves because it is possible unknown which loop or if the C-terminus binds to the tip. It is also possible that more than one BR binds unspecifically to the tip. The maximum length

### 3 Biological Principles

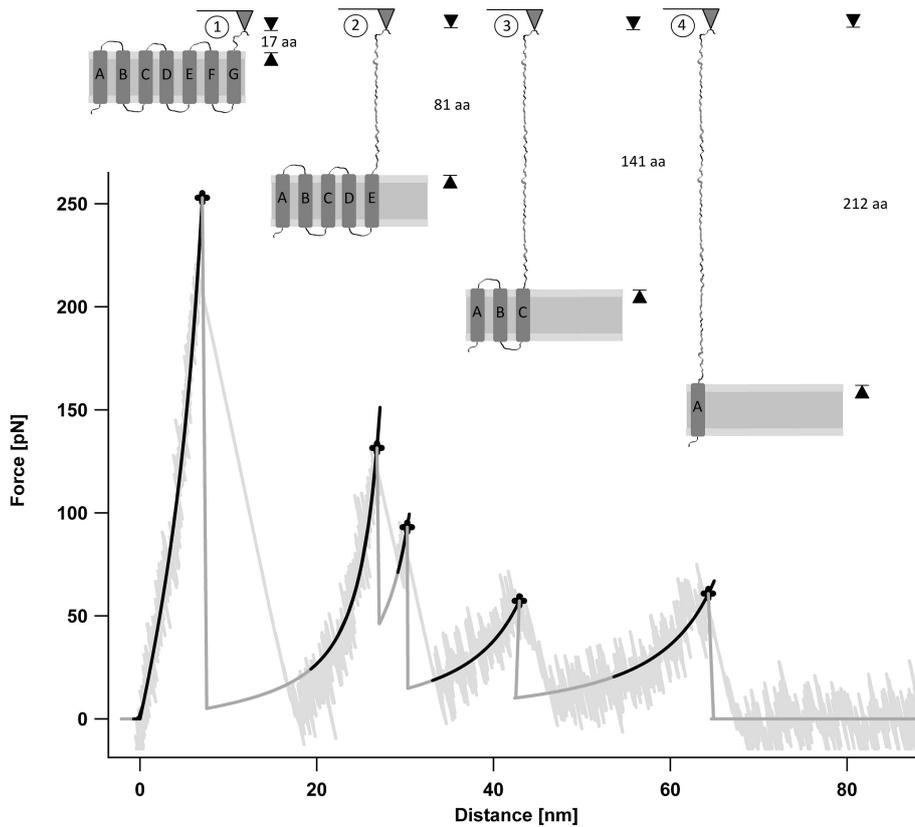


Figure 3.3: Model to explain the peaks in the force spectra as the sequential extraction and unfolding of a single BR. If a force is applied on the C-terminus, helices F and G will be pulled out of the membrane and unfold. Upon further retraction, the unfolded chain will be stretched and a force will be applied on helices D and E until they are extracted from the membrane. Thus, peak 2 reflects unfolding of helices D and E and peak 3 reflects unfolding of helices B and C. Peak 4 shows extraction of the last remaining helix A [OOP<sup>+</sup>00].

of the amino acid sequence is only measurable when the C-terminus binds to the tip and the N-terminus is still fixed in the membrane. The complete interaction length is therefore a strict criterion to select the force curves and has to be in the range of the maximum length. BR has with its 248 amino acids a maximum length of 90 nm, given by a length of around 0.36 nm per amino acid [LNC98].

When BR is unfolded from C- to N-terminus a typical pattern of four peaks occurs in the force curve:

### 3 Biological Principles

1. A force acts on helix G until helices G and F are extracted out of the membrane. The cantilever goes back to its idle position.
2. During further retraction of the cantilever, the force on the protein increases again and stretches it. This second unfolding step has a length corresponding to 88 amino acids which is exactly the length of helix G, loop G-F, helix F and loop F-E. Therefore the force is acting on helix E and leads to the unfolding of helices E and D. The tip snaps back to its idle position.
3. The already unfolded and now stretched amino acid sequence is 148 amino acids long corresponding to the length from helix G to the loop D-C. The force acts on helix C and leads to the unfolding of helices C and B.
4. The last stretching step has a length of 219 amino acids which corresponds to the length of BR just without the remaining helix A. When helix A is extracted out of the membrane the cantilever goes back to its idle position because there is no bond between tip and sample surface anymore.

The bacteria *Halobacterium salinarium* live in brine pools with high temperature, high exposure to sunlight and saturated saline conditions. The bacteria are therefore in the need of a process to maintain the osmotic balance without interfering with essential cellular processes. This is achieved inter alia by the stabilisation of the membrane proteins by compatible solutes, which are small, zwitterionic, organic molecules. This process is mostly not understood which makes BR interesting for our study to investigate the influence of compatible solutes on each unfolding step.



## 4 Materials and Methods

This chapter specifies the Monte Carlo method to simulate force spectroscopy data used in this work. Furthermore I introduce the statistical methods used during the tests of the developed analysis approach with simulated data.

### 4.1 Monte Carlo simulations of force spectroscopy data

The simulations are all made on an iMAC, Mac OS X Lion, Version 10.7.5. The software used for the analysis and simulation of force spectroscopy data is Igor Pro, Version 6.32A, from WaveMetrics Inc., Lake Oswego, Oregon 97035, USA. The main used mathematical equations are introduced in the following section and in 5.

To test our new analysis approach I simulate and analyse force spectroscopy data. Also in the concentration-dependency of the membrane protein Bacteriorhodopsin on osmolytes I simulated data to determine the influence of the thermal fluctuations on the uncertainty of the force. In the next section I explain the use of Monte Carlo simulations to generate force curves with a PEG linker. The used parameters are given in each section when force spectroscopy data is simulated.

#### 4.1.1 Simulating the setup of cantilever and PEG

A measured force curve plots the detected force versus the position of the piezo element  $l_{pp}$  which is the sum of the polymer length  $x(F)$  and the cantilever deflection  $x_{defl}(F)$ . The deflection of the cantilever is

$$x_{defl}(F) = F/k_{cant} \quad (4.1)$$

## 4 Materials and Methods

with  $k_{cant}$  denoting the spring constant of the cantilever. The relation between the polymer length and the force for PEG is given in Eq. (2.19).

The noise of the cantilever depends on the temperature and whether a polymer is bound between tip and surface. With no interaction between cantilever and surface the noise is just the thermal noise given by the equipartition theorem. Then the amplitude  $A$  of the cantilever deflection is

$$A = \sqrt{\frac{k_B T}{k_{cant}}}. \quad (4.2)$$

The binding of the cantilever to the surface via a polymer limits its movements related to thermal noise. The more the cantilever stretches the polymer, the less is the noise. The stiffness of the polymer is force-dependent and described by the spring constant of the polymer derived from the slope of the polymer length  $x(F)$

$$k_{Pol} = \frac{dx(F)}{dF}. \quad (4.3)$$

Then the spring constant of the combined system is

$$k_{sys} = k_{Pol} + k_{cant}, \quad (4.4)$$

which is used to add a gaussian distributed thermal noise to the cantilever deflection. Following the equipartition theorem the amplitude  $A$  of the cantilever deflection is

$$A = \sqrt{\frac{k_B T}{k_{sys}}}. \quad (4.5)$$

The determination of the dissociation point is based on the lifetime of the bond

$$\tau(F) = 1/k_{diss}(F), \quad (4.6)$$

with the Arrhenius equation describing the dissociation rate

$$k_{diss}(F) = \nu_0 \cdot e^{-\frac{\Delta G(F)}{k_B T}}. \quad (4.7)$$

$\nu_0$  is the so-called attempt frequency and represents the frequency of collisions between reactant molecules. The energy  $\Delta G(F)$  describes the strength of the bond and is force dependent. This is described in detail in chapter 5.

The dissociation probability is calculated as

$$p_{diss}(F) = 1 - e^{-k_{diss}(F) \cdot \Delta t} \quad (4.8)$$

## 4 Materials and Methods

with the time per data point  $\Delta t$  (which is the reciprocal data rate) [HA92]. At each time step the probability of the bond dissociation is compared to random, equally distributed numbers  $P_{random}$  between 0 and 1. The first time step with  $p_{diss}(F) > P_{random}$  is then the data point in the simulated force curve where the dissociation occurs.

In figure 4.1 I give an example of a simulated force-distance curve with parameters derived by analysing an experimental PEG-based force-distance curve of my Master thesis (describing the interaction of PhoB with DNA). The stretching parts of the curves overlay perfectly. The dissociation point and therefore the dissociation force depends on statistics.

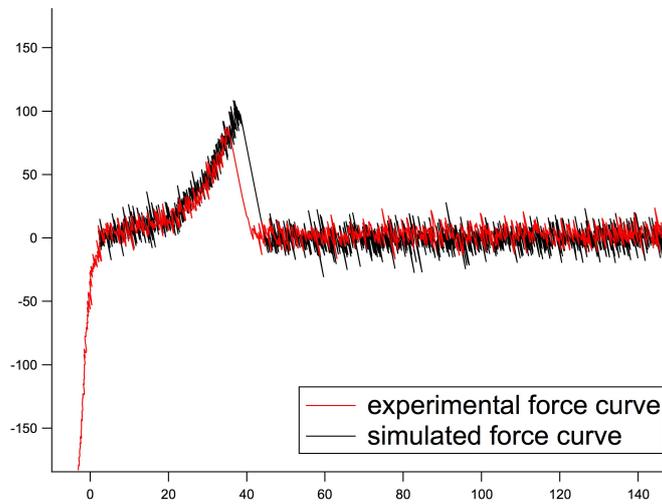


Figure 4.1: *Overlay of an experimental and a simulated force-distance curve of a PEG with a contour length of 39 nm, which corresponds to 111 monomers, a spring constant of the cantilever of 16 pN/nm and an InVOLS value of 40 nm/V. The retraction velocity is 500 nm/s.*

### 4.2 $\chi^2$ -fits and parameter uncertainties

Due to problems of Igor Pro to find the best fit and calculate the uncertainties of correlated parameters in the fit routine, I perform the fits in my analysis approach (see section 5.2.1) with Mathematica 7 (Wolfram Research).

## 4 Materials and Methods

With  $N$  simulated data points  $(x_i, y_i, \sigma_{y_i})$ ,  $n$  parameters  $p$  and the fit function  $y(x, p)$  the  $\chi^2$  is defined as

$$\chi^2(p) = \sum_i^N \left( \frac{y_i - y(x_i, p)}{\sigma_{y_i}} \right)^2 \quad (4.9)$$

for independent data points and gaussian distributed  $\sigma_{y_i}$ .

The minimum of this function,

$$\chi_{min}^2 = \min(\chi^2(p), p) = \chi^2(p_0), \quad (4.10)$$

gives information about the fit quality. Typically one expects the minimum of the order of the degrees of freedom, defined as  $N - n$ .

To visualise the correlation between the parameters I make 2-dimensional contour plots with different confidence levels. I define

$$\Delta\chi^2(p) = \chi^2(p) - \chi_{min}^2 \quad (4.11)$$

and determine for the parameters  $a \in p$

$$\Delta\chi^2(a) = \min[\Delta\chi^2(p), p \neq a]. \quad (4.12)$$

The condition

$$\Delta\chi^2(a) \leq x \quad (4.13)$$

defines a region in the (a)-parameter space, corresponding to a certain confidence level. For 2 parameters  $x = 2.28$  corresponds to the 68%-confidence interval and  $x = 5.99$  to the 95%-confidence interval [BZ06]. To determine the uncertainties of the parameter estimates I proceed analogously, minimizing  $\Delta\chi^2(p)$  for two parameters and choosing  $x=1$  and  $x=4$  for the 68%- and 95%-confidence interval, respectively.

I neglect the uncertainties of the forces and do not consider other statistical errors from the measurements or the FJC fit to the force-distance curves.

### 4.3 Force spectroscopy measurements on PhoB-DNA complexes

The intermolecular force spectroscopy measurements on PhoB and DNA were done during my Master thesis and the data were kindly provided by Prof. Dr. Anselmetti,

Institute for Experimental Biophysics and Applied Nanoscience, University of Bielefeld. The protocols for the surface and cantilever preparation as well as the used chemicals can be found in [Bie11] and [RWK<sup>+</sup>13].

### **4.4 Concentration-dependent measurements on BR**

These measurements were performed by Arpita Roychoudhury during her PhD and also the sorting of the data and fitting with the WLC model were done by her. Detailed information can be found in her dissertation [Roy13].



## 5 Revealing the energy landscape from force spectroscopy data

Dynamic force spectroscopy investigates the dissociation of non-covalent bonds under an external force, for example the interaction of ligand and receptor, see section 3.1.1. These measurements enable the direct observation of the force needed to dissociate the two bound molecules. However, it turns out that the mean dissociation force is not a characteristic parameter of the investigated receptor-ligand interaction, because it depends also on the retraction velocity which is explained in the following. Therefore there is the need to find a model to describe the intermolecular interactions with parameters which unambiguously characterize the interaction. Furthermore the determination of these parameters from force spectroscopy data should be possible.

The first interaction model of two free molecules in a solution was designed by Kramers in 1940 who introduced the so-called energy landscape [Kra40]: If two molecules are bound they are situated in a local minimum and for dissociation they have to cross an energy barrier by thermal fluctuations. Bell, Evans and Ritchie adapt this model to force-induced measurements and define three parameters of the energy landscape: the energy barrier height  $\Delta G_0$ , the distance between the minimum and the maximum of the energy barrier  $\Delta x_0$  (potential width) and the dissociation rate at zero force  $k_0$ .

With the assumptions that the potential has linear decrease between the minimum and the maximum, see figure 5.1, and that the applied force increases linearly with time when pulling with a constant speed, the pulling force adds an energy term. This term is the product of the applied force and the position  $x$  of the molecular complex in the energy landscape. This leads to a reduced energy barrier of the binding energy landscape by  $F \cdot \Delta x$  which is the distance between the minimum and the maximum. From the assumption of a linear dependency between minimum and maximum follows that  $\Delta x$  is independent of the force, so  $\Delta x = \Delta x_0$  (see the black line in figure 5.1).

## 5 Revealing the energy landscape from force spectroscopy data

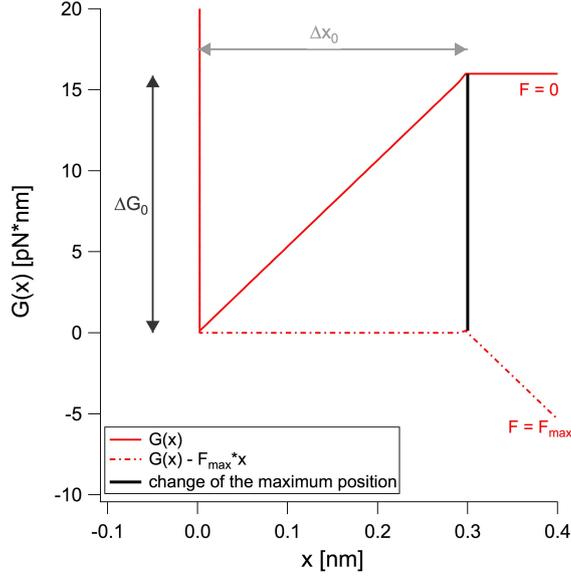


Figure 5.1: Shape of the energy landscape used by Evans et al. [ER97]: Between the minimum and the maximum is a linear increase. At zero force the height of the barrier is  $\Delta G_0$  and the potential width is  $\Delta x_0$ . When a force is applied on the molecular complex the energy barrier is lowered until it is 0 (dashed line). This figure also illustrates that  $\Delta x$  is independent of  $F$  and remains at  $\Delta x_0$  (black line).

Thus

$$k_{diss}(F) = \nu_0 e^{\frac{-\Delta G(F)}{k_B T}} = \nu_0 e^{\frac{-\Delta G_0 - F \cdot \Delta x_0}{k_B T}}, \quad (5.1)$$

with the Boltzmann constant  $k_B$  and temperature  $T$ . The attempt frequency  $\nu_0$  is actually temperature dependent because it is related to molecular collisions. The unit of  $k_{diss}$  depends on the order of the reaction. In AFM-force spectroscopy the survival probability  $p(t)$  of a bond under an external force can be described as a first-order Markov process [VK92]:

$$\dot{p}(t) = -k_{diss}(t) \cdot p(t). \quad (5.2)$$

Therefore the unit of  $k_{diss}$  and also of the attempt frequency  $\nu_0$  is reciprocal seconds ( $s^{-1}$ ).

This model does not include to determine the dissociation rate directly from the force spectroscopy data. They use the variation of the retraction velocity  $v$  in the measurements. Furthermore it is possible to obtain the dissociation force and the slope directly prior to the dissociation, called  $k_{eff}$  (which is the change in force with time), for each force curve. This  $k_{eff}$  is used to determine the so-called loading rate, defined

## 5 Revealing the energy landscape from force spectroscopy data

as [Eva01, Eva98]

$$r = k_{eff} \cdot v. \quad (5.3)$$

For each retraction velocity a distribution of loading rates is obtained and the most probable loading rate  $r^*$  (determined by a gaussian fit) is used in the subsequent analysis.

Under the assumption that the force loading rate is constant for each retraction velocity, Evans showed [ER97] that the most probable rupture force  $F^*$  depends on the loading rate  $r^*$  according to

$$F^* = \frac{k_B T}{\Delta x_0} \ln \left( \frac{\Delta x_0 \cdot r^*}{k_0 \cdot k_B T} \right). \quad (5.4)$$

$k_0$  and  $\Delta x_0$  can therefore be determined by measuring  $F^*$  and  $r^*$  for different retraction velocities and fitting the obtained data with Eq. (5.4).

The assumption of a linear decrease between minimum and maximum of the energy landscape leads to a combination of the parameters  $\Delta G_0$  and  $\nu_0$ , which is

$$k_0 = \nu_0 e^{-\frac{\Delta G_0}{k_B T}}. \quad (5.5)$$

In summary,  $k_0$  and  $\Delta x_0$  are obtained from the data with this model which is called Bell-Evans model in the following. This model describing force spectroscopy measurements does not include to determine  $k_{diss}(F)$  from the data.

To show that the reduction of the loading rate for one retraction velocity to a single data point is not a good description of this dependency I generated figure 5.2. For this figure I simulated for each of the 6 different retraction velocities (50 nm/s, 100 nm/s, 500 nm/s, 1000 nm/s, 3000 nm/s and 5000 nm/s) 1000 force curves and used from the simulation the values for the rupture force to display them versus the loading rates.

The parameters in the simulation are the following: I used the WLC stretching behaviour for the polymer with a total length of  $L = 50$  nm, a persistence length of  $l_c = 0.6$  nm and a temperature of  $T = 300$  K. For the energy landscape I used the Bell-Evans (linear) potential with a potential barrier  $\Delta G_0 = 16k_B T$ ,  $\Delta x_0 = 0.3$  nm and  $\nu_0 = 10^6$  s<sup>-1</sup>.

The difficulties associated with the loading rate can be avoided experimentally by using another AFM measurement technique, called force-clamp technique [SKW<sup>+</sup>08]. There, the force applied to the molecular system is held constant while measuring the

## 5 Revealing the energy landscape from force spectroscopy data

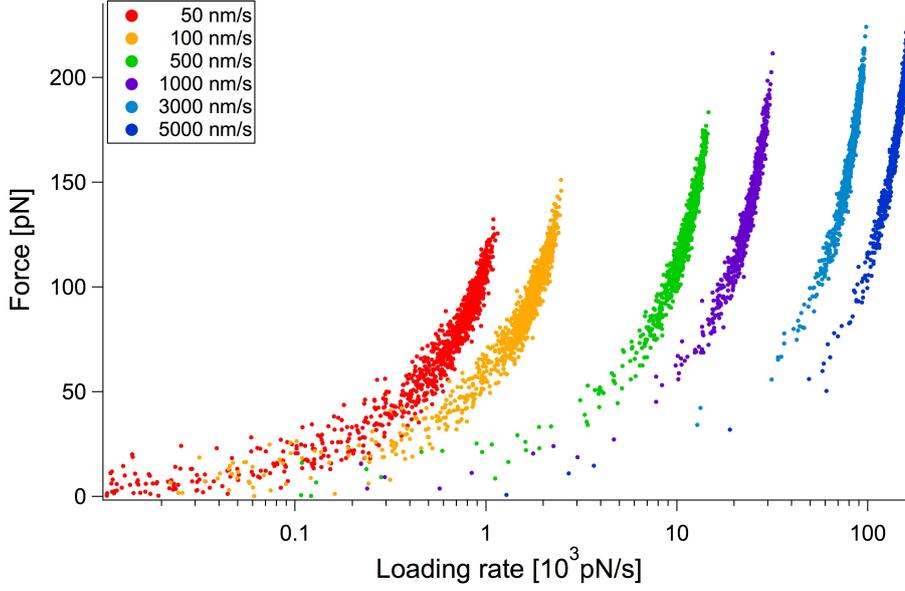


Figure 5.2: Illustration of the dissociation force dependend on the loading rate for each force curve. In the Bell-Evans model a most probable dissociation force and a most probable loading rate is obtained for each distribution corresponding to one retraction velocity.

bond survival time. The probability  $P$  to find the dissociation after the time  $t_i$  follows an exponential decay,

$$P(t_i|k_{diss}) = 1 - e^{-t_i \cdot k_{diss}} \quad (5.6)$$

with  $i = 1, \dots, N$  for the  $i$ th measurement of the survival time of the bond.

The probability to find the survival times  $\{t_1, \dots, t_N\}$  in  $N$  measured dissociation events is given by the product

$$P(\{t_1, \dots, t_N\}|k_{diss}) = \prod_i P(t_i|k_{diss}) = P(\bar{t}|k_{diss})^N \quad (5.7)$$

with the average survival time

$$\bar{t} = \sum_i \frac{t_i}{N}. \quad (5.8)$$

The maximal probability for  $P(t_{av}|k_{diss})^N$  is at

$$k_{diss} = \frac{1}{\bar{t}} = \sum_i \frac{N}{t_i}. \quad (5.9)$$

In the next section I explain how the force-clamp analysis procedure can be adapted to force spectroscopy data to extract the force-dependent offrate. For a detailed description see Oberbarnscheidt et al. [OJO09] and Serpe et al. [SKW<sup>+</sup>08], who introduced simultaneously this approach.

## 5.1 Model-free calculation of force-dependent dissociation rate

When measuring force-distance curves, each data point corresponds to a force  $F_{ext}$  applied over a time interval  $\Delta t$  which can be regarded as an individual force-clamp experiment. With a fixed sampling rate  $\nu$ , the time intervals are evenly spaced with  $\Delta t = 1/\nu$ . The dissociation probability of the bond while the external force  $F_{ext}$  is acting on it can be approximated as  $k_{diss}(F_{ext}) \cdot \Delta t$ . Since the data points are independent from each other, all data points taken at the same force can be merged. The dissociation rate  $k_{diss}(F_{ext})$  is the number of dissociation events per time:

$$k_{diss}(F_{ext}) = \frac{N(F_{ext})}{M(F_{ext}) \cdot \Delta t}. \quad (5.10)$$

$N(F_{ext})$  is the number of ruptures measured at the force  $F_{ext}$  which is represented by the dissociation force histogram.  $M(F_{ext})$  is the number of data points with no dissociation taken at this force  $F_{ext}$ . So  $M(F_{ext}) \cdot \Delta t$  is the total time of the bond surviving the external force  $F_{ext}$ .

This method extracts the force-dependent dissociation rate directly from the data and all data points of all force curves are taken into account. Since every force curve is influenced by thermal fluctuations, it is important to determine the truly acting forces on the bond: Assuming the bending of the cantilever towards the surface is displayed as positive values, see figure 2.3, then a bending towards the surface shows up as a positive pulling force but in fact the length of the polymer is shortened and therefore the force on the bond is reduced. So the truly acting force is given by the average force at the actual tip surface distance. This average force can be determined by the vertical projection of each data point in the force curve onto a fit which describes the force derived from the average cantilever deflection.

With this technique it is possible to calculate the force-dependent dissociation rate directly from the force curves. The aim of this work is to develop a model to determine

the energy landscape from this force-dependent dissociation rate. This is discussed in the next section.

## 5.2 Determination of the energy landscape parameters

The force-dependent dissociation rate can be obtained from force spectroscopy data. Its connection to the energy landscape is given by the Arrhenius equation,

$$k_{diss}(F_{ext}) = \nu_0 e^{-\frac{\Delta G(F_{ext})}{k_B T}} \quad (5.11)$$

(see also Eq. (5.1)) with  $\Delta G(F_{ext})$  describing the dependency of the energy barrier height on the external force  $F_{ext}$ . To explain the way to derive an equation describing  $\Delta G(F_{ext})$  I start with defining the energy landscape. I generalize this simple linear dependency between minimum and maximum of the potential by considering higher orders  $n$ :

$$G(x) = \begin{cases} \Delta G_0 \left[ \left( \frac{x}{\Delta x_0} \right)^n - 1 \right], & \text{for } x \leq \Delta x_0 \\ 0, & \text{for } x \geq \Delta x_0. \end{cases} \quad (5.12)$$

For  $\Delta G_0 = 16k_B T$  and  $\Delta x_0 = 0.3 \text{ nm}$   $G(x)$  is shown in figure 5.3 for  $n = 1, \dots, 5$ .

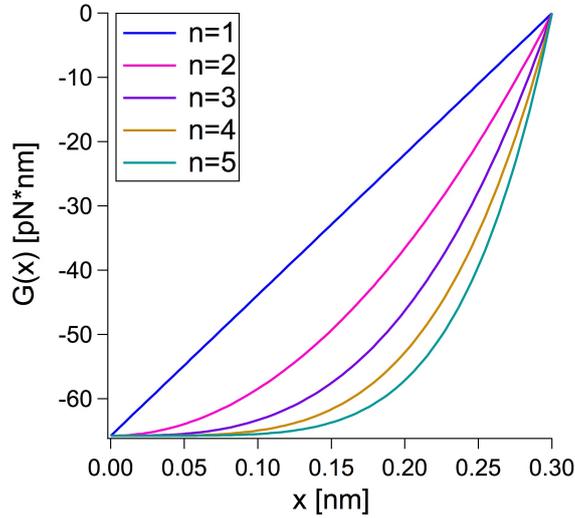


Figure 5.3: The energy landscape  $G(x)$  for  $n = 1, \dots, 5$  following Eq. (5.12) with  $\Delta G_0 = 16k_B T$  and  $\Delta x_0 = 0.3 \text{ nm}$ .

## 5 Revealing the energy landscape from force spectroscopy data

By pulling on the molecular complex during force spectroscopy measurements a force  $F_{ext}$  acts on the bond. This adds energy to the system:

$$G(x, F_{ext}) = G(x) - F_{ext} \cdot x. \quad (5.13)$$

Due to the fact that we do not get direct information about the  $x$ -values of  $G(x)$  out of our force-distance curves, the dependency of  $G$  has to be expressed just by  $F_{ext}$ . So we need a relation between  $x$  and  $F_{ext}$ :

The force  $F$  between two molecules is

$$F = -\frac{d}{dx}G(x, F) = -G'(x) + F_{ext}. \quad (5.14)$$

With the assumption that the structural changes within the molecular complex take place much faster than the changes in the pulling force, which means that the complex is in equilibrium all the time, the force  $F$  between the two molecules vanishes. From this follows that

$$F_{ext} = G'(x) = n \frac{\Delta G_0}{\Delta x_0} \left( \frac{x}{\Delta x_0} \right)^{n-1}. \quad (5.15)$$

By rearranging Eq. (5.15) an equation describing the dependency of the distance between minimum and maximum  $x$  on the force  $F$  is obtained

$$x(F_{ext}) = \left( \frac{\Delta x_0 F_{ext}}{n \Delta G_0} \right)^{\frac{1}{n-1}} \Delta x_0. \quad (5.16)$$

By inserting (5.16) and (5.12) in Eq. (5.13) we gain an expression for  $G$  depending on  $F_{ext}$ :

$$G(F_{ext}) = \Delta G_0 \left[ \left( \frac{\Delta x_0 F_{ext}}{n \Delta G_0} \right)^{\frac{n}{n-1}} - 1 \right] - F_{ext} \Delta x_0 \left( \frac{\Delta x_0 F_{ext}}{n \Delta G_0} \right)^{\frac{1}{n-1}}. \quad (5.17)$$

This equation describes the position of the minimum of  $G$  for varying  $F_{ext}$ . The dissociation rate  $k_{diss}(F_{ext})$  is the potential barrier height  $\Delta G(F_{ext}) = G_{max}(F_{ext}) - G_{min}(F_{ext})$ .

With

$$G_{max}(F_{ext}) = G(\Delta x_0, F_{ext}) = -\Delta x_0 F_{ext}, \quad (5.18)$$

the potential barrier height is

$$\Delta G(F_{ext}) = \Delta G_0 - \Delta x_0 F_{ext} + \left[ \frac{\Delta x_0^{\frac{n}{n-1}}}{\Delta G_0^{\frac{1}{n-1}}} \left( \frac{1}{n^{\frac{1}{n-1}}} - \frac{1}{n^{\frac{n}{n-1}}} \right) \right] F_{ext}^{\frac{n}{n-1}}. \quad (5.19)$$

The maximal difference to the linear case is the quadratic energy landscape with  $n = 2$ . This offers in principle the possibility to extract additional information from the

## 5 Revealing the energy landscape from force spectroscopy data

force dependence, as long as the terms can be separated experimentally. This is easier when the exponents are more different. The power of  $F$  in the third term for  $n = 3$  is  $3/2$  and  $4/3$  for  $n = 4$ . So the difference in this term decreases with increasing  $n$ . As less the contribution of the third term as more the parameters  $\Delta G_0$  and  $\nu_0$  are only of significance combined to the parameter

$$k_0 = \nu_0 \cdot e^{-\frac{\Delta G_0}{k_B T}}. \quad (5.20)$$

Interestingly  $\Delta G(F_{ext})$  converges to

$$\lim_{n \rightarrow \infty} \Delta G(F_{ext}) = \Delta G_0. \quad (5.21)$$

This means that for high  $n$  the solution converges to a force-independent potential well.

For the quadratic energy landscape with  $n = 2$  the equation for  $\Delta G(F_{ext})$  is

$$\Delta G(F_{ext}) = \Delta G_0 - \Delta x_0 F_{ext} + \frac{\Delta x_0^2}{4\Delta G_0} F_{ext}^2. \quad (5.22)$$

Inserting Eq. (5.19) in the Arrhenius-equation (5.11) I obtain an equation for the force-dependent dissociation rate. For the orders  $n = 1, \dots, 5$   $k_{diss}(F_{ext})$  is shown in figure 5.4 with  $\Delta G_0 = 16k_B T$  and  $\Delta x_0 = 0.3$  nm.

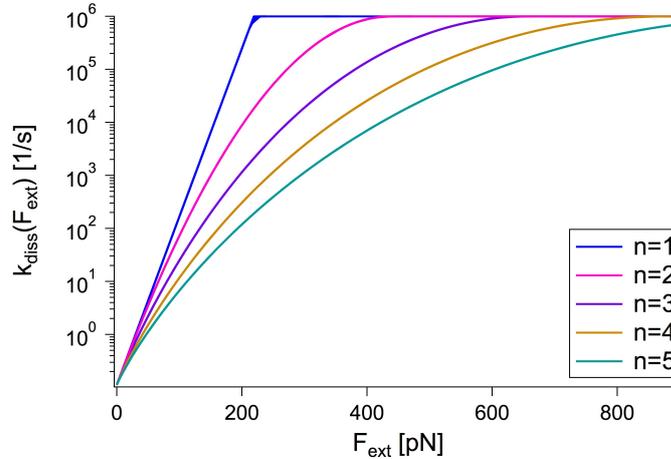


Figure 5.4:  $k_{diss}(F_{ext})$  with  $n = 1, \dots, 5$ ,  $\Delta G_0 = 16k_B T$  and  $\Delta x_0 = 0.3$  nm, following the Arrhenius equation with Eq. (5.19) for the energy barrier dependence on the external force.

A difference to the Bell-Evans analysis procedure with  $\Delta x(F_{ext}) = \Delta x_0 = \text{constant}$ , is the dependency of the potential width on force. In figure 5.5 this is illustrated for

## 5 Revealing the energy landscape from force spectroscopy data

a harmonic energy landscape ( $n=2$ ). Be aware that the  $G(x)$  in my model is defined differently to Eq. (5.12), see figure 5.3. I choose here this different definition of the energy landscape for illustration purpose, only.

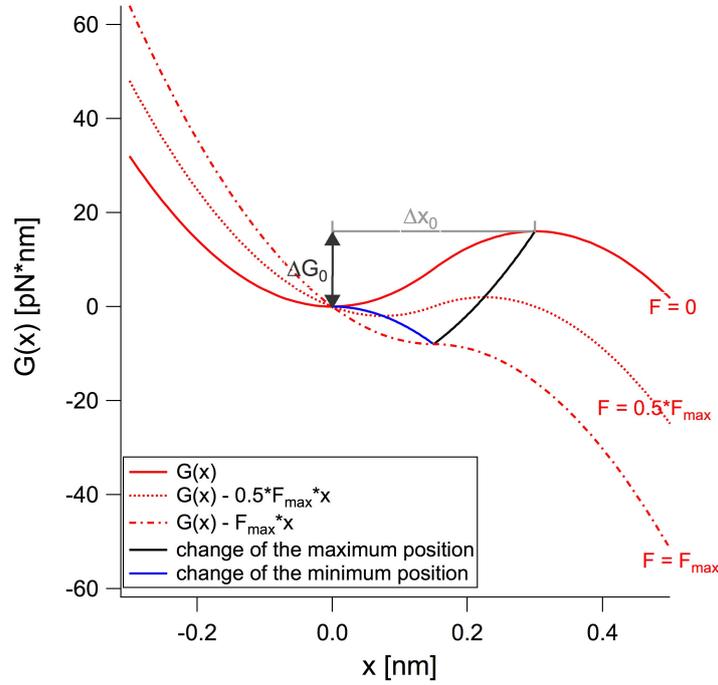


Figure 5.5: Illustration of the dependency of  $\Delta x$  on the applied force  $F_{ext}$ . The black curve shows the change in the minimum and maximum of the potential until they get reduced to one inflexion point.

Ideally I could determine  $G(x)$  directly from the force spectroscopic data without a fit. This is not possible because I can not get direct information about  $\Delta x(F)$ . So I have to make an assumption on the potential shape that defines the dependency of  $\Delta x(F_{ext})$  and therefore  $\Delta G(F_{ext})$ . Then we can fit the dissociation rate data and determine the parameters  $\nu_0$ ,  $\Delta G_0$  and  $\Delta x_0$ . Furthermore  $\chi_{min}^2$  gives information about the goodness of the fit.

In the next section I test our model with simulated data to estimate the goodness of the parameter determination and the fit quality.

## 5.2.1 Testing the analysis strategy with simulated data

To test this analysis approach I simulate data with known parameters and a given energy landscape. Then I analyse the simulated data with our approach and fit my model to the obtained dissociation rate. From this fit it is possible to determine the quality of this approach by  $\chi_{min}^2$  and by comparison of the parameters obtained for different fit orders  $n$  of  $\Delta G(F_{ext})$  with the ones used in the simulation. In the following I denote  $n$  as  $n_{fit}$  when I use Eq. (5.19) for the fitting, and  $n_{sim}$  when I use it for the simulation.

For the simulation I use the following parameters for the PEG based polymer: The polymer has a contour length of 37 nm. All the other parameters for the PEG are from Oesterhelt et al. [ORG99]. For the spring constant of the cantilever I use a value of 16 pN/nm and an *InvOLS* value of 45 nm/V (see Eq. (2.11)). The energy landscape parameters are  $\Delta G_0 = 16k_B T$ ,  $\Delta x_0 = 0.3$  nm (based on the values of titin [RGO<sup>+</sup>97]) and  $\nu_0 = 10^6$  1/s (based on the measurements of Scheuring et al. on titin [RGC<sup>+</sup>13]).

### 5.2.1.1 Considering one retraction velocity

First I present the results when only one retraction velocity is measured. I use here the retraction velocity of 500 nm/s. I simulate 2000 force curves with always only one bound and dissociated PEG. For the underlying energy landscape I set  $n_{sim} = 2$ . Then I fit them with the PEG model (see section 2.4.1) and calculate the dissociation rate as explained in the introduction of this chapter. The standard deviation is calculated as

$$\sigma_{k_{diss}(F_{ext})} = \frac{\sqrt{N(F_{ext})}}{M(F_{ext}) \cdot \Delta t}. \quad (5.23)$$

#### **Test with a quadratic potential: $n_{sim} = 2$**

In figure 5.6 the force-dependent dissociation rate obtained from the force curves and the theoretical distribution of the dissociation rate is shown for the quadratic case with  $n_{sim} = 2$ .

The calculated dissociation rate reflects the theoretical distribution (black line). In the force spectroscopy measurement with a constant retraction velocity just a section of

## 5 Revealing the energy landscape from force spectroscopy data

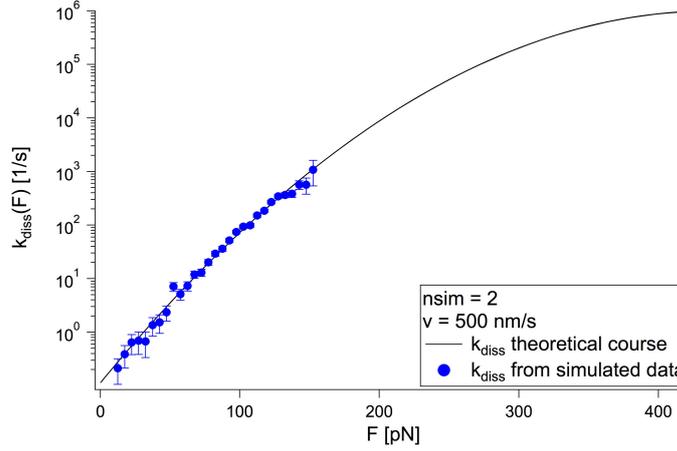


Figure 5.6: Force-dependent dissociation rate for a retraction velocity of 500 nm/s from  $N=2000$  force curves, gaining 29 data points. The black line shows the theoretical distribution of the force-dependent dissociation rate, following Eq. (5.11) and the corresponding force-dependent potential barrier height  $\Delta G(F_{ext})$  (5.19) with  $n_{sim} = 2$ .

the force-dependent dissociation rate can be accessed because the accessible forces depend on the dissociation probability, see Eq. (4.8). This probability is dependent on the time per data point which is controlled by the retraction velocity in the experiment. The slower the cantilever is retracted the earlier the bond can dissociate by thermal influence. Therefore lower dissociation forces are obtained. However, with low retraction velocities dissociation events with low dissociation forces get lost when the bond dissociates within the thermal noise.

I fit the force-dependent dissociation rate with Eq. (5.11) and the corresponding force-dependent potential barrier height  $\Delta G(F_{ext})$  (5.19) for  $n_{fit} = 1, \dots, 5$ . From these fits I get  $\chi_{min}^2$ , describing the goodness of the fit, and the parameters  $\Delta G_0$ ,  $\Delta x_0$  and  $\nu_0$ .

In figure 5.7  $\chi_{min}^2$  in dependency of the fit order  $n_{fit}$  is shown for the simulated data with  $n_{sim} = 2$ . The degrees of freedom are  $29 - 3 = 26$ . This is approximately the expected  $\chi_{min}^2$ -value. The  $\chi_{min}^2 = 71.2$  for  $n_{fit} = 1$  compared to the other  $\chi_{min}^2$  of  $n_{fit} = 2, \dots, 5$  shows that it is possible to distinguish between the linear potential and the higher order potentials.  $n_{fit} = 2$  has the expected minimum in  $\chi_{min}^2 = 23.7$  but the differentiation to  $n_{fit} = 3, 4, 5$  with  $\chi_{min}^2 \approx 25$  each is difficult..

It is important to have in mind why the determination of the parameter is possible: The logarithm of the dissociation rate (Eq. (5.11) with Eq. (5.19)) can be expressed as

## 5 Revealing the energy landscape from force spectroscopy data

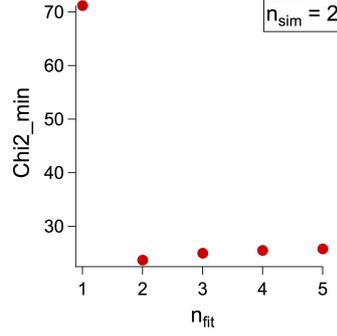


Figure 5.7:  $\chi^2_{min}$  for different fit orders  $n_{fit}$  for  $n_{sim} = 2$ .

a polynomial in  $F_{ext}$ :

$$\ln k_{diss}(F_{ext}) = a + b \cdot F_{ext} + c \cdot F_{ext}^{\frac{n}{n-1}} \quad \text{for } n \geq 2, \quad (5.24)$$

with

$$a = \ln v_0 - \frac{\Delta G_0}{k_B T}, \quad (5.25)$$

$$b = \frac{\Delta x_0}{k_B T} \quad \text{and} \quad (5.26)$$

$$c = -\frac{1}{k_B T} \left[ \left( \frac{\Delta x_0}{\Delta G_0} \right)^{\frac{1}{n-1}} \left( \frac{1}{n^{n-1}} - \frac{1}{n^{\frac{n}{n-1}}} \right) \right]. \quad (5.27)$$

The fit routine effectively optimizes the parameters  $a, b$  and  $c$ , from which the values for the parameters of  $\Delta x_0$ ,  $\Delta G_0$  and  $v_0$  are determined, according to the solution of the system of equations (5.25) - (5.27).

For the Bell-Evans case ( $n = 1$ ) the fit equation is

$$\ln k_{diss}(F_{ext}) = a + b \cdot F_{ext} \quad (5.28)$$

with

$$a = \ln v_0 - \frac{\Delta G_0}{k_B T} \quad \text{and} \quad (5.29)$$

$$b = \frac{\Delta x_0}{k_B T}. \quad (5.30)$$

Due to this reduction of parameters it is not possible to separate  $\Delta G_0$  and  $v_0$  for the Bell-Evans case with  $n = 1$ . Only the introduction of the dependency of  $\ln k_{diss}(F_{ext})$

## 5 Revealing the energy landscape from force spectroscopy data

on higher orders of  $F_{ext}$  leads to the third equation and therefore allows for the separation of these parameters, always keeping in mind that for high  $n$  the correlation of the parameters increases again.

When considering only one retraction velocity the range of accessible forces is limited. This limitation and the influence of the thermal noise on the FJC fit to the force curves leads to the indistinguishability of high order potentials ( $n_{fit} > 2$ ). Therefore I decide not to analyse the parameters for this case.

### Test with a cubic potential: $n_{sim} = 3$

I test also the case for an underlying energy landscape with  $n_{sim} = 3$  in the simulation. All the other parameters stay the same. The force-dependent dissociation rate for this case is shown in figure 5.8.

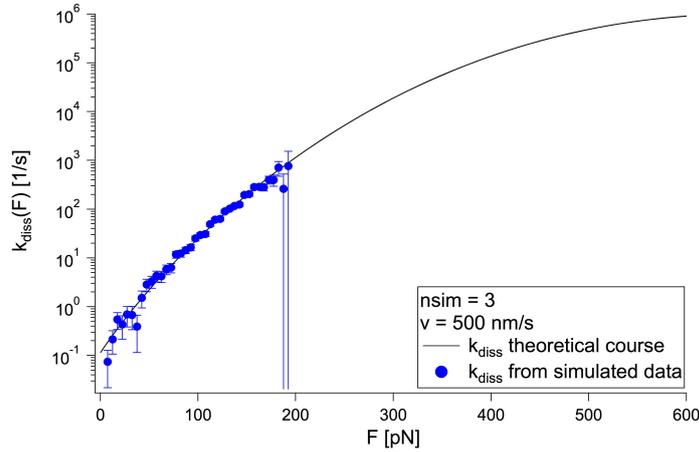


Figure 5.8: Force-dependent dissociation rate for a retraction velocity of 500 nm/s from  $N=2000$  force curves with  $n_{sim} = 3$ . The black line shows the theoretical distribution of the force-dependent dissociation rate, following Eq. (5.11) and the corresponding force-dependent potential barrier height  $\Delta G(F_{ext})$  (5.19) with  $n_{sim} = 3$ .

Again, just a limited range of the force-dependent dissociation rate is obtained from the data, well described by the theoretical distribution. The results for  $\chi_{min}^2$  using different orders  $n_{fit}$  are shown in figure 5.9. The number of degrees of freedom is here  $38 - 3 = 35$ .  $n_{fit} = 4, 5$  with  $\chi_{min}^2 = 33.2$  are again not differentiable from  $n_{fit} = n_{sim} = 3$  with  $\chi_{min}^2 = 33.3$ , but an exclusion of  $n_{fit} = 1$  is also possible here.

## 5 Revealing the energy landscape from force spectroscopy data

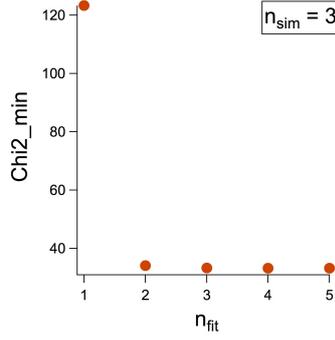


Figure 5.9:  $\chi^2_{min}$  for different fit orders  $n_{fit}$  for  $n_{sim} = 3$ .

Therefore there is a need for broadening the force-range in the measurements to obtain better insight into the potential. This can be done by varying the retraction velocity which influences the dissociation probability as explained before and leads to different force-ranges in the measurements. In the next section I test the fit routine when three different retraction velocities are considered.

### 5.2.1.2 Considering three retraction velocities

In order to test if a differentiation of the higher orders of the potential becomes possible with a higher amount of data points in the dissociation rate  $k_{diss}$  I combine the data of three retraction velocities which broadens the force range. A higher retraction velocity allows to determine higher forces and vice versa, as explained in the previous section. For that purpose I simulate for each of the retraction velocities 20 nm/s, 500 nm/s and 10000 nm/s 2000 force curves. Again I analyse the force curves and determine the dissociation rates for each retraction velocity.

To combine the dissociation rate for all retraction velocities, I determine the dissociation rate as the weighted mean for a single bin,

$$k_{diss,all} = \frac{\sum_i \frac{k_{diss,i}}{\sigma_{k_{diss,i}}^2}}{\sum_i \frac{1}{\sigma_{k_{diss,i}}^2}} \quad \text{with} \quad \sigma_{k_{diss,all}} = \frac{1}{\sqrt{\sum_i \frac{1}{\sigma_{k_{diss,i}}^2}}}. \quad (5.31)$$

### Test with a quadratic potential: $n_{sim} = 2$

In figure 5.10 this combined force-dependent dissociation rate is shown for an un-

## 5 Revealing the energy landscape from force spectroscopy data

derlying potential with  $n_{sim} = 2$ . By considering three retraction velocities the force range is broadened to 20 – 230 pN.

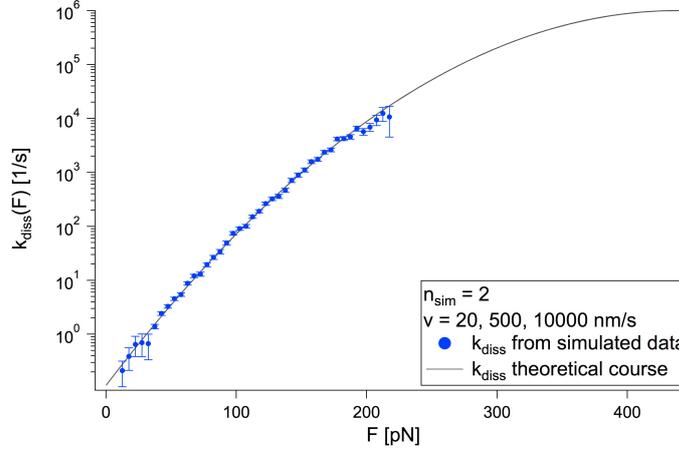


Figure 5.10: Force-dependent dissociation rate for the retraction velocities of 20, 500 and 10000 nm/s, using  $N=2000$  force curves for each retraction velocity with  $n_{sim} = 2$ . The black line shows again the theoretical distribution of the force-dependent dissociation rate, following Eq. (5.11) and the corresponding force-dependent potential barrier height  $\Delta G(F_{ext})$  (5.19) with  $n_{sim} = 2$ . The dissociation rate has 42 data points.

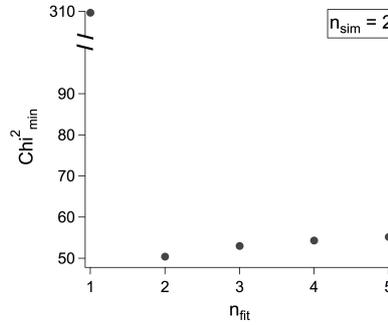


Figure 5.11:  $\chi^2_{min}$  of the different fit orders  $n_{fit}$  for  $n_{sim} = 2$ .

After fitting I obtain  $\chi^2_{min}$  and the parameters  $\Delta x_0$ ,  $\Delta G_0$  and  $v_0$ .  $\chi^2_{min}$  is illustrated for the different fit orders in figure 5.11. For  $n_{fit} = 1$ , which corresponds to the linear decrease between potential minimum and maximum, comparable to the Bell-Evans approach, the fit does not describe the data well as demonstrated by  $\chi^2_{min} = 309.7$  for 40 degrees of freedom. The fit with  $n_{fit} = 2$  yields the best fit with  $\chi^2_{min} = 50.4$ . For

## 5 Revealing the energy landscape from force spectroscopy data

larger  $n_{fit}$  the fit worsens.

For the visualisation of the uncertainties and correlation of the parameters I make contour plots, as explained in 4.2, which are shown in figure 5.12.

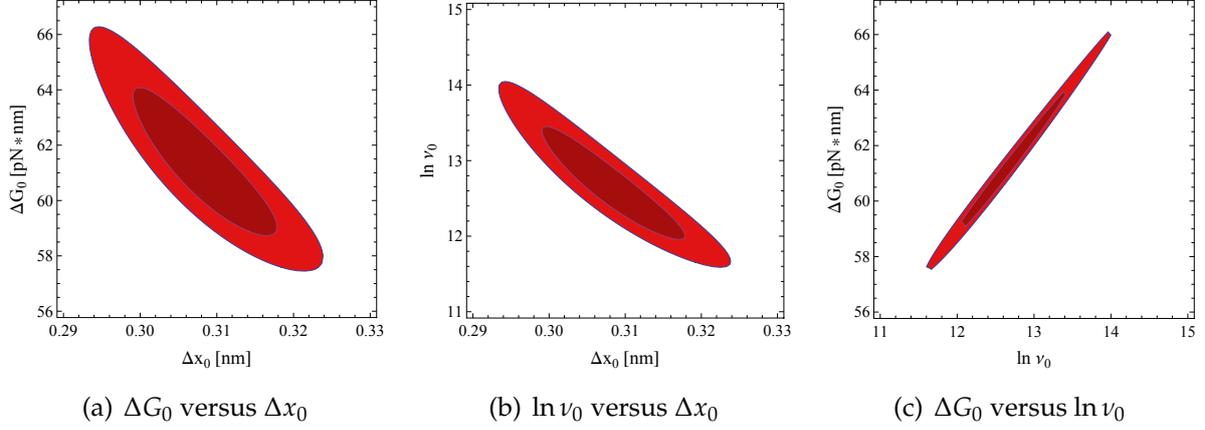


Figure 5.12: Contour plots for the three parameters  $\Delta G_0$ ,  $\Delta x_0$  and  $\ln \nu_0$  with  $n_{sim} = 2$  and  $n_{fit} = 2$ . Red illustrates the 68%-confidence interval and orange the 95%-confidence interval. All parameters show a correlation. The simulation values are  $\Delta G_0 = 65.7$  pN·nm,  $\Delta x_0 = 0.3$  nm and  $\ln \nu_0 = 13.8$  and lie all in the 95%-confidence interval.

The parameters show a strong correlation. The uncertainties obtained here are much bigger than the uncertainties with the assumption of uncorrelated parameters in Igor Pro. I determine the individual uncertainties as described. I summarize the estimated parameters and their standard deviations in figure 5.13.

The parameters  $\Delta G_0$ ,  $\Delta x_0$  and  $\ln \nu_0$  show that the fit with the quadratic potential reflects best the simulated parameters in comparison to the other fit orders, see figure 5.13. The fit yields  $\Delta x_{0,fit} / \Delta x_{0,sim} = 1.02$ ,  $\Delta G_{0,fit} / \Delta G_{0,sim} = 0.93$  and  $\ln \nu_{0,fit} / \ln \nu_{0,sim} = 0.92$ . The extracted potential width  $\Delta x_0$  is linearly dependent on the fit order. This means that an incorrect choice for  $n_{fit}$  will lead to a bias in this parameter. The same is true for  $\Delta G_0$  and  $\ln \nu_0$ , albeit with a different dependence. All parameters show a strong correlation, which is to be expected from Eq. (5.24). I do not show the parameters for  $n_{fit} = 1$  because this distribution does not describe the data.

## 5 Revealing the energy landscape from force spectroscopy data

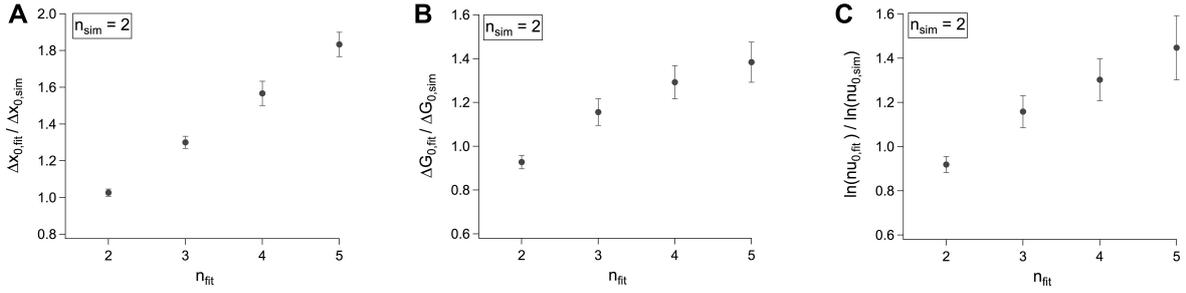


Figure 5.13: Comparison of the determined parameters in the fit with the values of the simulation with a given quadratic potential. A) The potential barrier height  $\Delta G_0$ , B) the potential width  $\Delta x_0$  and C) the attempt frequency  $\ln \nu_0$ , all normalised to the simulation values.

### Test with a cubic potential: $n_{sim} = 3$

Again I test also the case with  $n_{sim} = 3$ . The corresponding force-dependent dissociation rate is shown in figure 5.14. By considering the three retraction velocities (20 nm/s, 500 nm/s and 10000 nm/s) the force-range is broadened to 20 – 300 pN.

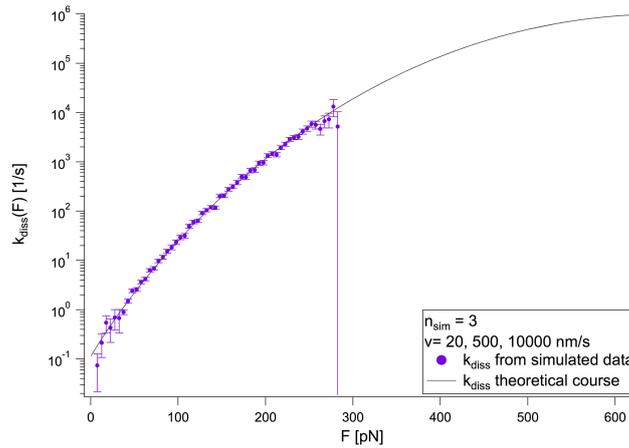


Figure 5.14: Force-dependent dissociation rate for the retraction velocities of 20, 500 and 10000 nm/s, using  $N=2000$  force curves for each retraction velocity with  $n_{sim} = 3$ . The black line shows the theoretical distribution of the force-dependent dissociation rate, following Eq. (5.11) and the corresponding force-dependent potential barrier height  $\Delta G(F_{ext})$  (5.19) with  $n_{sim} = 3$ . The dissociation rate has 48 data points.

After fitting I obtained the values for  $\chi_{min}^2$ , shown in figure 5.15. Here the linear

## 5 Revealing the energy landscape from force spectroscopy data

fit describes the data even worse than for the  $n_{sim} = 2$ . Also the fit based on the quadratic potential  $n_{fit} = 2$  does not reproduce the data well, yielding  $\chi_{min}^2 = 64$ . The orders  $n_{fit} = 4, 5$  yield a slightly better  $\chi_{min}^2$  than  $n_{fit} = 3$  which is due to the increasing indifference in the order of  $F_{ext}$  in the fit equation (5.24). Therefore the orders  $n_{fit} = 1$  and 2 can be rejected, but a differentiation between  $n_{fit} = 3, 4$  and 5 is not possible.

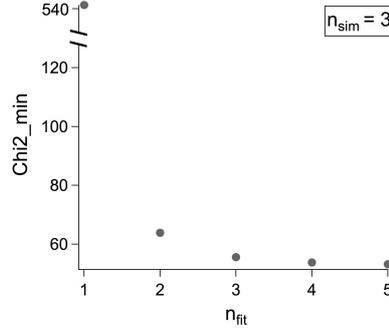


Figure 5.15:  $\chi_{min}^2$  of the different fit orders  $n_{fit}$  for  $n_{sim} = 3$ .

Again I make contour plots, see figure 5.16, to visualise the uncertainties and correlations of the parameters.

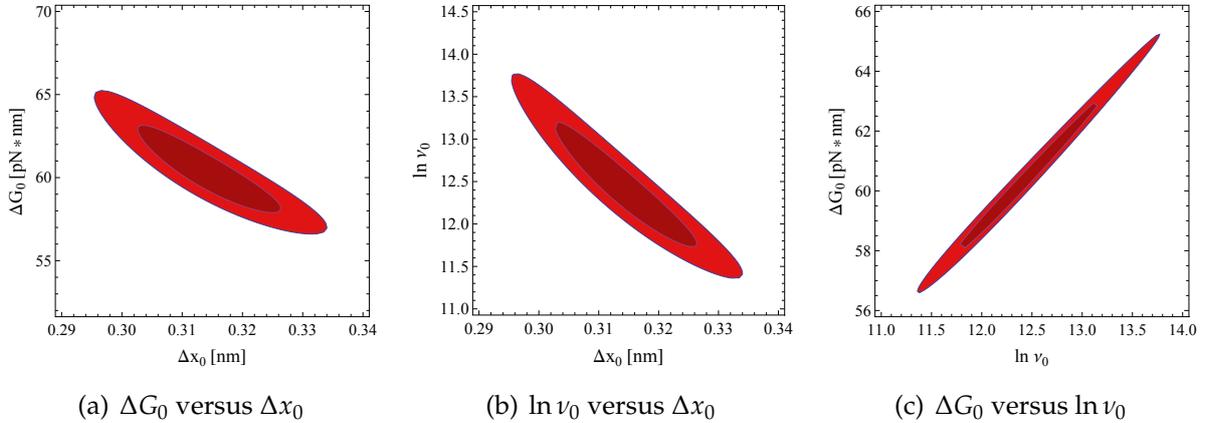


Figure 5.16: Contour plots for the three parameters  $\Delta G_0$ ,  $\Delta x_0$  and  $\ln \nu_0$  with  $n_{sim} = 3$  and  $n_{fit} = 3$ . Red illustrates the 68%-confidence interval and orange the 95%-confidence interval. All parameters show a correlation. The simulation values are  $\Delta G_0 = 65.7$  pN · nm,  $\Delta x_0 = 0.3$  nm and  $\ln \nu_0 = 13.8$  which are hardly in the 95%-confidence interval.

## 5 Revealing the energy landscape from force spectroscopy data

I determine the uncertainties from the 1-dimensional plots and display the parameters with their standard deviations in figure 5.17. The parameters show a strong correlation for  $n_{sim} = 3$ , too. Also all parameter show a dependency on the fit order what leads to a bias with an incorrect choice for  $n_{fit}$ .

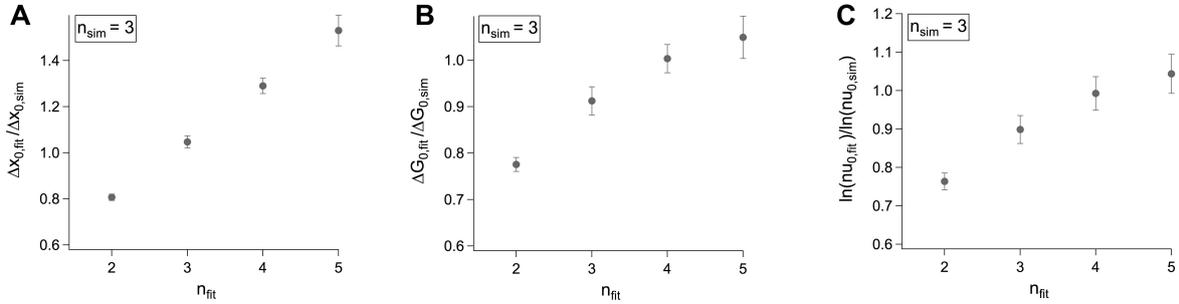


Figure 5.17: Comparison of the parameters determined in the fit with the values of the simulation for a cubic potential. A) The potential barrier height  $\Delta G_0$ , B) the potential width  $\Delta x_0$  and C) the attempt frequency  $\nu_0$ , all normalised to the simulation values.

The dependency of the parameters on  $n_{fit}$  shows that care must be taken when assuming a potential for this kind of analysis. On the other hand, the knowledge of a valid range of one or more parameters leads to an increase of knowledge for the potential. The valid range could be obtained for example from different methods like NMR spectroscopy.

### 5.2.2 Applying the model to PhoB-DNA force spectroscopy data

In this section I use the data of the point mutant R176A of the protein PhoB, measured during my Master thesis, to test the analysis approach of Oberbarnscheidt et al. [OJO09], determining the force-dependent dissociation rate. Furthermore I test my analysis procedure for the parameter determination of the energy landscape by fitting these data and comparing with the parameters obtained from the analysis with the Bell-Evans model, published in [RWK<sup>+</sup>13].

The force spectroscopy data of R176A is interesting to analyse because the binding behaviour of the alanine mutant provides valuable insight into the type of interac-

## 5 Revealing the energy landscape from force spectroscopy data

tions involved in the recognition process. Residue arginine 176 interacts with the minor groove by hydrogen bonds [YOI<sup>+</sup>08]. Arginine 176 and tyrosine 223 also interact with each other to stabilize the orientation and providing an optimal distance between helix  $\alpha_2$  and the C-terminal hairpin [BSGRC02, YOI<sup>+</sup>08]. The CD spectroscopy measurements on the mutant R176A do not show any binding of the mutant to the DNA anymore [RWK<sup>+</sup>13]. In the force spectroscopy measurement Ritzefeld et al. see a slight binding affinity with a dissociation rate of  $1.8 \text{ s}^{-1}$  which gives a lifetime of 0.5 s of the R176A-DNA interaction. The wildtype has a lifetime of 400 s. Due to this remarkable reduction of the lifetime of the PhoB-DNA complex the residue arginine 176 is important for the binding of the protein PhoB to the DNA.

I analyse the force spectroscopy data by fitting the extended FJC model with multiple bonds, described by Eq. (2.20), to the dissociation events, which is different from the fit used in [Bie11]. Only the curves corresponding to a single dissociation event are used in the following analysis. From these I calculate the force-dependent dissociation rate.

As shown in section 5.2.1.2 it is advisable to use more than one retraction velocity to obtain a broader force range for a better determination of the energy landscape parameters. I analyse here the retraction velocities 50 nm/s, 500 nm/s and 4000 nm/s, including the lowest and highest retraction velocities measured. The retraction velocities 500 nm/s and 4000 nm/s are measured with the same cantilever with a spring constant of 16 pN/nm. The measurement with a retraction velocity of 50 nm/s is performed with a different cantilever with a spring constant of 14.6 pN/nm. The force-dependent dissociation rates obtained for the three retraction velocities are shown in figure 5.18.

The force-dependent dissociation rates show a different distribution for the retraction velocity of 50 nm/s compared to the retraction velocities of 500 nm/s and 4000 nm/s. To verify if the DNA is bound to the PEG, I analyse the contour lengths of the FJC fits to the dissociation events in the force curves. The PEG-linker has a contour length of 27 nm and the DNA a length of 10 nm. Therefore the expected contour length is around 37 nm. I build histograms for the contour lengths for the different retraction velocities, see figure 5.19, and fit them with a gaussian distribution to determine the most probable contour length. The most probable contour lengths for the retraction velocities 500 nm/s and 4000 nm/s are in agreement with the expected value of 37 nm. Due to the increased contour length of around 10 nm, the displacement of the

## 5 Revealing the energy landscape from force spectroscopy data

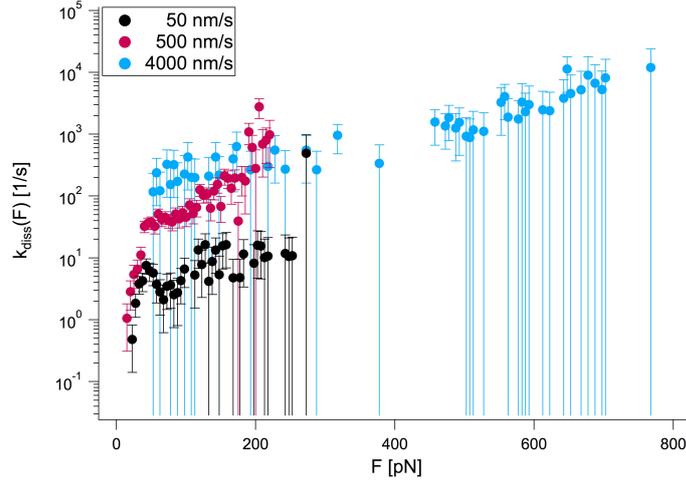


Figure 5.18: Force-dependent dissociation rate of the analysed retraction velocities  $v = 50 \text{ nm/s}$ ,  $500 \text{ nm/s}$  and  $4000 \text{ nm/s}$ .

distribution of the force-dependent dissociation rate and no contribution to a broadening of the force-range, I decide not to include the data of the retraction velocity of  $50 \text{ nm/s}$  in the following analysis.

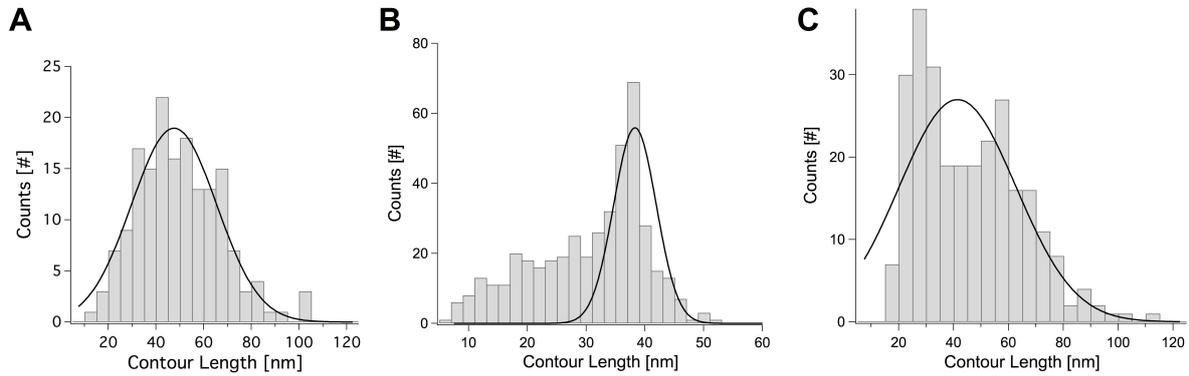


Figure 5.19: Histograms for the determination of the most probable contour length of A)  $50 \text{ nm/s}$ , B)  $500 \text{ nm/s}$  and C)  $4000 \text{ nm/s}$ . The gaussian fit is displayed in black. The expected contour length is  $37 \text{ nm}$  and for  $50 \text{ nm/s}$  I obtain a most probable contour length of  $(45 \pm 2) \text{ nm}$ , for  $500 \text{ nm/s}$  a contour length of  $(36.0 \pm 0.5) \text{ nm}$  and for  $4000 \text{ nm/s}$  a contour length of  $(39 \pm 2) \text{ nm}$ .

The data for the retraction velocities of  $500 \text{ nm/s}$  and  $4000 \text{ nm/s}$  are combined as described in Eq. (5.31), and shown in figure 5.20. These force-dependent dissociation

## 5 Revealing the energy landscape from force spectroscopy data

rates show a linear dependency on the force. Therefore I fit the dissociation rate, using Eq. (5.11), Eq. (5.19) and  $n_{fit} = 1$  which corresponds to a linear decrease between minimum and maximum in the energy landscape as in the Bell-Evans analysis approach. With  $n_{fit} = 1$  it is only possible to obtain the parameters for the potential width  $\Delta x_0$  and the dissociation rate at zero force  $k_0$ , see Eq. (5.5). From the fit I obtain the following parameters:

$$k_0 = (27 \pm 1) \text{s}^{-1}$$

and

$$\Delta x_0 = (0.31 \pm 0.1) \text{\AA}.$$

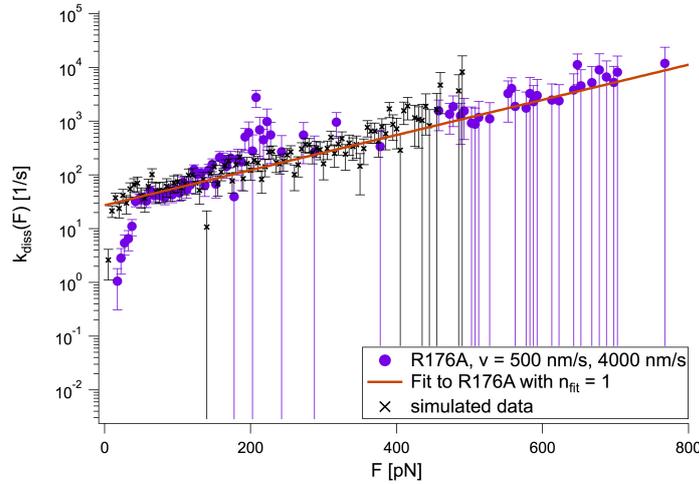


Figure 5.20:

The dissociation rate  $k_0$  is one order of magnitude larger than  $k_0 = (1.8 \pm 1.2) \text{ nm}$  [RWK<sup>+</sup>13], obtained by the Bell-Evans analysis procedure (fitting the most probable dissociation force in dependency of the most probable loading rate). The value for the potential width is significantly shorter than the values typically observed for interactions between protein and DNA [KCB<sup>+</sup>04, VNA05] and the obtained value for  $\Delta x_0$  of around 1 nm by the Bell-Evans analysis procedure [Bie11]. Interestingly, this value is in the range of the barrier width per base pair of 0.7 Å obtained by Strunz et al. in the DNA unbinding experiments [SOSG99]. Also the interaction between carboxyl groups (COOH) has a comparable potential width of 0.6 Å for fast loading rates higher than  $10^{-8} \text{ N/s}$ . To explain this short potential width I see two possibilities. First, it could be characteristic for the reduced binding affinity of this mutant to the DNA. In the CD spectroscopy measurements even no binding is found. Second, the short potential width could be determined because no specific interaction is detected

## 5 Revealing the energy landscape from force spectroscopy data

in the measurements. It is known that PEG-linker can bind unspecifically between surface and tip of the cantilever, used by Oesterhelt et al. in their PEG stretching measurements [ORG99]. But they do not report the strength of the nonspecific binding of the PEG to the cantilever. In the measurements used here, the DNA is bound to the PEG-linker and the most probable contour length is comparable to the one expected with successful DNA-binding to the PEG, see figure 5.19 B) and C). Therefore another possibility is that we see here the nonspecific interaction between DNA and gold [RCSG99] and the DNA binding to the mutant R176A is completely disrupted.

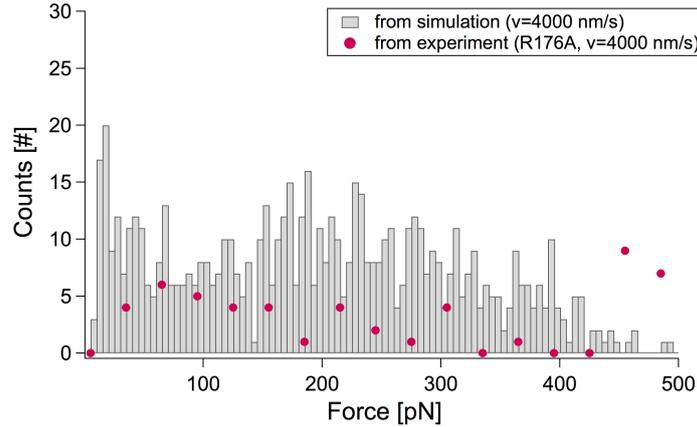


Figure 5.21: Histogram of the dissociation forces for the experimental and simulated data, with a retraction velocity of 4000 nm/s. The binning of the histograms is different due to the different number of counts per histogram. Remarkably, both distributions are not gaussian.

To test the obtained values for  $\Delta x_0$  and  $k_0$ , I simulate force curves with these values, a PEG-linker with a contour length of 37 nm and a retraction velocity of 4000 nm/s. The obtained data points for the analysis of the simulated data are shown in figure 5.20. The simulated data show the same distribution as the experimental data and also a similar broadness and uncertainties. Remarkably, the dissociation force histogram for the simulated data is very broad and does not show a gaussian distribution. The same holds for the experimental data, see figure 5.21, which also supports the thesis that we do not see a specific interaction here.

From the test of my analysis approach with simulated data, see section 5.2.1.2, we know that the wrong order in the fit leads to a bias in the parameters. A too small fit order leads to a too small potential width  $\Delta x_0$ . Therefore I fit also the data with  $n_{fit} = 2$ , obtaining a  $\Delta x_0 = 0.34 \text{ \AA}$ , which does not influence the value sufficiently

## *5 Revealing the energy landscape from force spectroscopy data*

that an order of magnitude could be achieved. The uncertainties here are directly from the fit.

The observed force dependency of the dissociation rate is linear and no curvature can be observed in the accessible force range, see figure 5.20. The force-dependent dissociation rate in the measurements of Berkemeier et al. is linear over the experimentally accessible force range [BBX<sup>+</sup>11], too. This is the first and so far only publication which uses the analysis approach to determine the dissociation rate directly from force spectroscopy data. A verification if the force-dependent dissociation rate shows a curvature and therefore the underlying energy landscape has a non-linear decrease between its minimum and maximum can only be achieved by analysing more force spectroscopy data with this method.

## 6 Dependency of the stability of BR on the concentration of compatible solutes

Protecting effects of compatible solutes on soluble proteins have been studied to a fairly good extent [OSF<sup>+</sup>03, LS06]. Due to the technical challenges of isolation, purification and folding of a membrane protein into its natural conformation and to find measuring conditions which do not denature the membrane protein, studies on membrane proteins are rare so far. In the paper [RBHO13], we report the results of single molecule unfolding experiments on the mechanical stability of the membrane protein BR in dependency on the concentration of the compatible solutes ectoine, betaine and taurine. The mechanical effects are represented by the change in dissociation force and persistence length of each unfolding step of BR. The dissociation force is a measure for the protein stability against unfolding and the persistence length describes the tendency of the protein to form a compact coil. When the persistence length is reduced, the protein is more coiled up.

Arpita Roychoudhury performed the measurements during her PhD on BR without osmolytes (0 M) and on BR with 1 mM, 10 mM, 100 mM and 1000 mM of each of the osmolytes betaine, ectoine and taurine, see [Roy13].

First I explain how I analyse the data to distinguish between the different unfolding steps to determine the influence of the compatible solutes on the dissociation force and the persistence length of each unfolding step. Furthermore I analyse the influence of the number of received unfolding events per unfolding step on the uncertainty of the force with Monte Carlo simulations, which is explained in section 6.1.1. Then follows the explanation how the uncertainty of the persistence length of the single peak analysis is determined. Concluding I present and discuss the results of the concentration-dependent measurements of compatible solutes on BR.

## 6.1 Analysis of each unfolding step

The force curves have been selected by virtue of their typical unfolding pattern with four main unfolding events and a total length of around 70 – 90 nm (see section 3.1.2). We fit these peaks with the worm-like chain model (see section 2.4.2). From these fits we obtain the persistence length, considering only events with a  $\chi^2 \leq 5$ . From these data we generate scatter plots of the persistence length versus the fit length, see figure 6.1 B, and of the dissociation force versus fit length, see figure 6.1 A, where the latter is determined directly from the force curve.

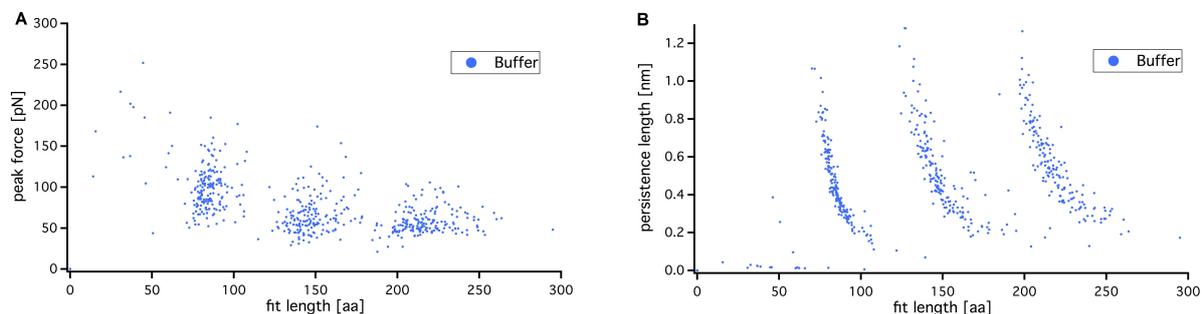


Figure 6.1: Scatter Plots for BR with no osmolytes (0M). A) shows the dissociation force vs the fit length. B) shows the persistence length vs the fit length. In both graphs the first unfolding step is missing (which should be at a fit length of around 17 amino acids). Therefore just 3 distributions are visible.

Because of the short dissociation length of around 10 nm of the first unfolding step of helices G and F, see figure 3.3, the quality of the WLC fits is low ( $\chi^2 > 5$ ). Therefore the corresponding unfolding step, expected at a fit length of about 17 amino acids, is missing in figure 6.1 A and B.

My task is to analyse every distribution and determine the mean dissociation force and the mean persistence length with the corresponding uncertainties for each unfolding step. To group the events according to their corresponding unfolding step I differentiate the distributions by their fit length and create histograms out of the individual sections for the dissociation forces and persistence lengths for each osmolyte concentration. The events with a fit length between 80 and 130 amino acids are assigned to the second unfolding step. To the third unfolding step I count all the events

## 6 Dependency of the stability of BR on the concentration of compatible solutes

with a fit length between 130 and 200 amino acids and to the fourth one the events with a fit length of 200 to 250 amino acids. The number of dissociation events for each unfolding step and concentration of each osmolyte can be found in the appendix A.

From these distributions I create histograms corresponding to individual unfolding steps, and determine the mean values by gaussian fits. This procedure is illustrated in figure 6.2 for the data of 1 M betaine.

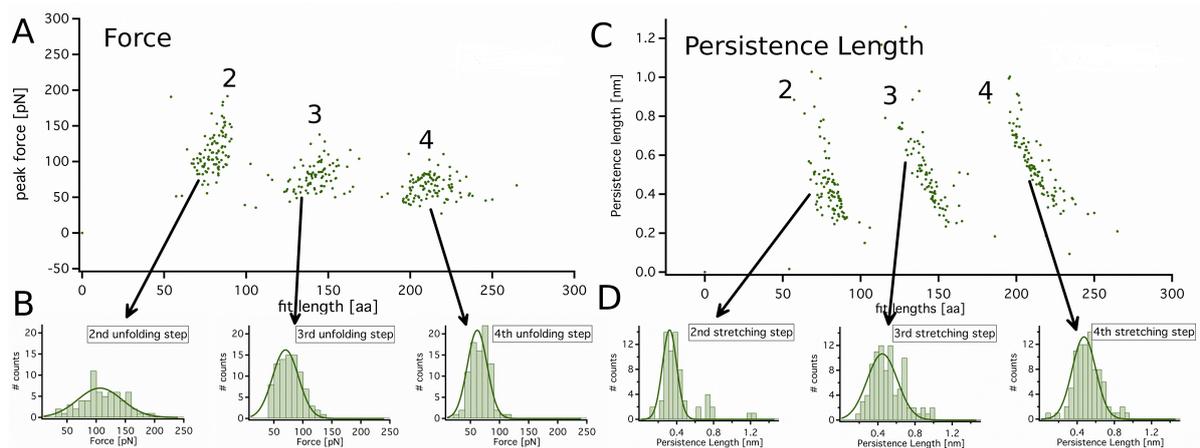


Figure 6.2: Explanation of our analysis procedure for the different unfolding steps on the basis of 1 M betaine. A) Scatter Plot of the peak force versus the fit length. We get three distributions for the pairwise unfolding of the helices E-D (2nd unfolding step), C-B (3rd unfolding step) and for the last unfolding of the helice A (4th unfolding step). The histogramming for each unfolding step leads to the histograms in B). A gaussian fit reveals the most probable dissociation force. C) Scatter Plot of the persistence length versus the fit length. We get three distributions for the stretching of the unfolded parts: the second distribution is the stretching of helices G-F, the third one for G-F and E-D, and the fourth one for the stretching of G-F, E-D and C-B. D) Histograms of the persistence length for 1M betaine for the 2nd, 3rd and 4th stretching step.

### 6.1.1 Uncertainty of the forces derived from simulation

I analyse the influence of the number of received unfolding events per unfolding step on the uncertainty of the force. This is due to the reduced number of unfolding events in the experiments with osmolytes because the osmolytes reduce the binding probability of the tip of the cantilever to the protein significantly.

Due to the fact that the extraction of dissociation forces of an unfolding step is not only dependent on statistics but also on thermal fluctuations, we get a distribution of unfolding forces dependent on the stretching behaviour of the polymer and the thermal influence. The higher the temperature the broader is the histogram. The stretching behaviour leads to an asymmetric distribution: The probability that the unfolding step occurs at small forces is higher than at high forces. In figure 6.3 a histogram is shown to illustrate this asymmetric distribution which I derive from 1000 simulated dissociation forces.

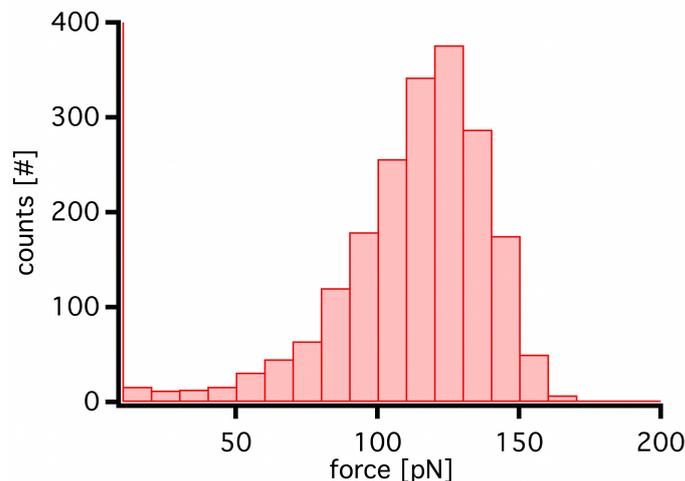


Figure 6.3: *Histogram of 1000 simulated dissociation forces: Smaller dissociation forces are more likely than high dissociation forces.*

To include the thermal influences on the dissociation force and investigate the dependence on the number of measured events, I simulate force spectroscopy data. I use the WLC model which fits best for protein unfolding with a persistence length of 0.6 nm and a total length of  $L = 50$  nm, which corresponds to the second stretching step (helices G and F) of BR with 81 amino acids (see figure 3.3). For the spring constant of the cantilever I use a value of 25 pN/nm and an *InvOLS* value of 45 nm/V

## 6 Dependency of the stability of BR on the concentration of compatible solutes

(see Eq. (2.11)). Furthermore I set the retraction velocity to 400 nm/s. These values are in accordance with the experimental situation. For the energy landscape I assume a harmonic potential with  $n_{sim} = 2$  (see Eq. (5.19)) with  $\Delta G_0 = 16k_B T$ ,  $\Delta x_0 = 0.3$  nm and  $\nu_0 = 10^6$  s<sup>-1</sup>.

Each force curve is fitted with the WLC model and the dissociation forces are obtained. Ten sets of dissociation forces are produced, each with the following numbers of dissociation events: 25, 50, 100 or 200. For the concentration of 1 mM betaine and the third unfolding step we just have 12 unfolding events with  $\chi^2 < 5$  in the WLC fit (see appendix A). For this case I produce 20 sets of simulated dissociation forces. I histogram each set and each histogram is fitted with a gaussian distribution.

In figure 6.4, the 10 histograms containing 200 counts are shown in A and for 100 counts in B. The influence of noise can be seen but still the distribution is similar for all 10 histograms. I apply to each histogram a Gaussian fit

$$y = A \cdot e^{-\left(\frac{x-x_0}{FWHM}\right)^2}. \quad (6.1)$$

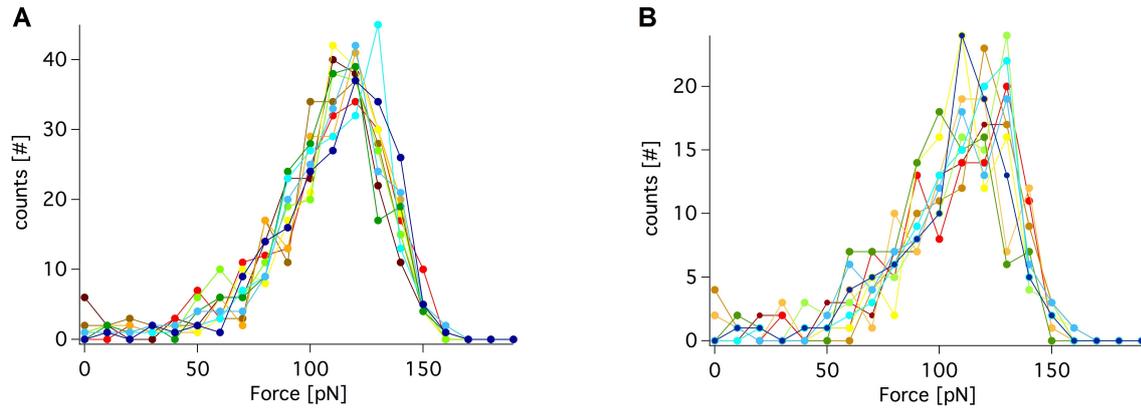


Figure 6.4: A) Overlay of 10 histograms of simulated forces containing 200 counts. B) Overlay of 10 histograms of simulated forces containing 100 counts.

In figure 6.5, I display the obtained mean forces ( $x_0$  in Eq. (6.1)) and their errors derived from these fits.

For the calculation of the mean force and its error, I determine the arithmetic mean

## 6 Dependency of the stability of BR on the concentration of compatible solutes

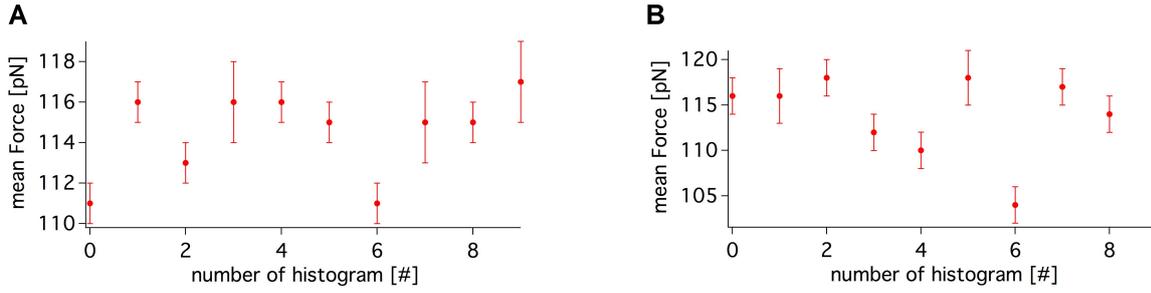


Figure 6.5: A) Mean force of the histograms shown in figure 6.4 A with 200 counts, derived from a Gaussian fit, and its standard deviation. B) Mean force of the histograms shown in figure 6.4 B with 100 counts, derived from a Gaussian fit, and its standard deviation.

and its standard deviation:

$$\bar{F} = \frac{1}{N} \sum_{i=0}^N F_i \quad \text{with} \quad \sigma_F = \sqrt{\frac{1}{N-1} \sum_{i=0}^N (F_i - \bar{F})^2}. \quad (6.2)$$

From this I derive for the histogram containing 200 counts a mean force of 115 pN and a standard deviation of 2 pN. For the histogram with 100 counts I get a mean force of 114 pN and a standard deviation of 4 pN.

Because the procedure is the same for all counts, I show finally just the result for the histogram with the lowest number of counts, namely 12 counts. Based on the low number I decide to use here 20 sets, see figure 6.7.

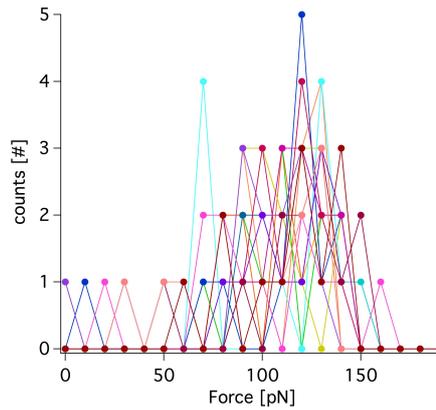


Figure 6.6: Overlay of 20 histograms of simulated forces containing 12 counts, each.

Although the histograms do not contain many data points, their distribution is not

## 6 Dependency of the stability of BR on the concentration of compatible solutes

flat. The obtained mean forces by the gaussian fit, see figure 6.7, show that the first histogram cannot be fitted in a proper way, but all the other histograms are still fine. With the same calculation procedure as above I get for the histogram containing 12 counts a mean force of 113 pN and a standard deviation of 9 pN.

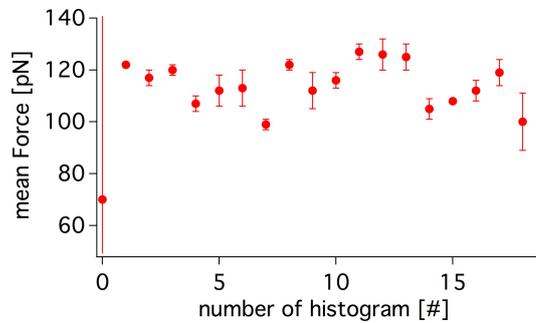


Figure 6.7: Mean forces of the histograms shown in figure 6.3, derived from a Gaussian fit, and their standard deviations.

To summarize these results I show in figure 6.8 the dependency of the standard deviation of the force on the number of counts in the histogram.

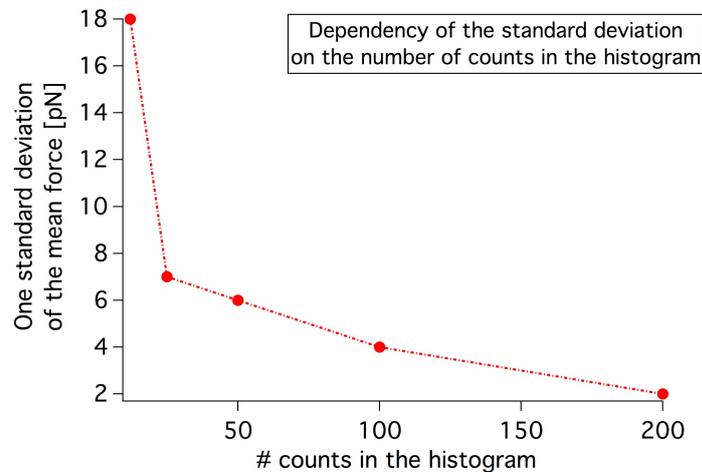


Figure 6.8: Graph of the standard deviations dependent on the number of counts in a histogram.

From this graph I derive tabular 6.1 which we use to get the uncertainty of the forces of our experimental results, shown in section 6.2, as the determined standard deviation from the simulations.

counts	$\sigma_F$ [pN]
12 - 15	9
16 - 20	8
21 - 30	7
31 - 66	6
67 - 88	5
89 - 113	4
114 - 165	3
166 - 200	2
> 200	2

Table 6.1: Standard deviation of the force in dependency on the number of data counts.

### 6.1.2 Uncertainty of the persistence length

The persistence length values for each unfolding step are derived from WLC fits to the experimentally obtained force curves. I perform a Gaussian fit to each of the persistence length histogram from the single peak analysis to determine the standard deviation as  $\sigma_{l_p} = FWHM / \ln(4)$ .

#### 6.1.2.1 Influence of the WLC fit on the obtained persistence length

When we look at the scatter plots of the persistence length for the different concentrations of osmolytes, we notice a curvature in the distributions for the different unfolding steps: High persistence lengths are obtained for lower fit lengths and vice versa. To decide if this is based on the intramolecular interaction or influenced by our analysis procedure I simulate and analyse data to observe the persistence length parameter. From this data I create scatter plots as before for the experimental data. I use the same parameters for the simulation as for the determination of the standard deviation of the force in section 6.1.1. The noise is calculated as  $\sqrt{k_B T / k_{cant}}$ .

What I vary first is the sampling frequency to obtain the influence of the number of data points on the persistence length, see the medium blue and light blue data in figure 6.9. The increase in the number of data points spreads the distribution but cannot explain the broad distribution of the experimental data shown in yellow (for BR with-

## 6 Dependency of the stability of BR on the concentration of compatible solutes

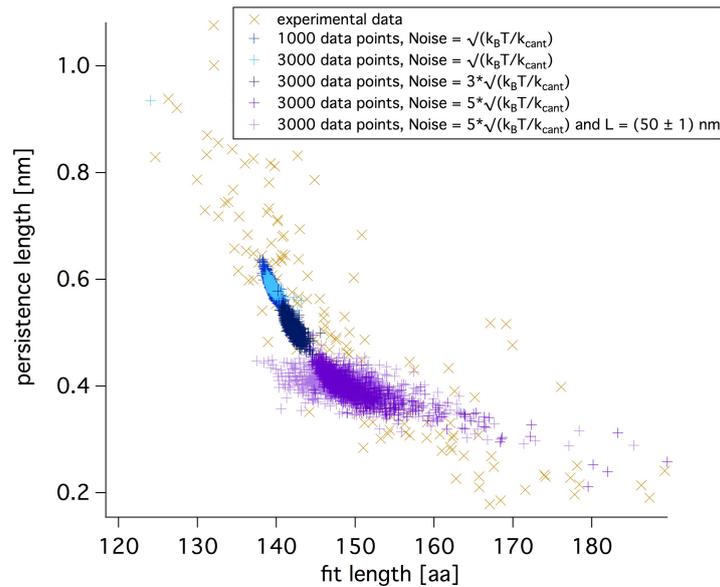


Figure 6.9: Comparison of the simulated data with experimental data of BR without osmolytes (yellow) by plotting the persistence length in dependency on the fit length. The influence of the parameter variation in the simulation is tested. The variation in the number of data points cannot reproduce the experimental distribution (medium and light blue distribution). The increase of the thermal noise shifts the distribution to lower persistence length values and also broadens the distribution. The variation in the total length of the simulated polymer by plus-minus 1 nm leads to a similar width of the distribution compared to the experimental data.

out osmolytes). The input value of the persistence length of 0.6 nm is reproduced.

Then I vary the thermal noise of the cantilever: I increase the value by three times. The dark blue distribution in figure 6.9 corresponds to this simulation. This change shifts the distribution to smaller persistence length values and broadens it. The increase of the noise by five times intensifies this effect (dark purple distribution). By variation of the total length of the WLC polymer by equally distributed values by plus-minus 1 nm I achieved a similarly broad variation in the fit length as in the experimental data.

These simulations show that the experimentally obtained distribution are likely to stem from the analysis procedure. This can be explained by an inaccurate WLC fit which is discovered by these simulations: In the dissociation point the forces are asymmetrically distributed to lower forces. Therefore the fit routine drags the fit

## 6 Dependency of the stability of BR on the concentration of compatible solutes

curve down which influences the value of the persistence length, too. Thus a high rupture force leads to a lower persistence length. This is already observed by Kühner and Gaub [KG06].

The fit routine is the same for all osmolyte concentrations and unfolding steps.

## 6.2 Results and discussion

I present here the results obtained by my analysis presented in the last section.

### Influence of concentration changes on the dissociation force

The concentration-dependent impact of compatible solutes on the dissociation force for each peak is presented in figure 6.10.

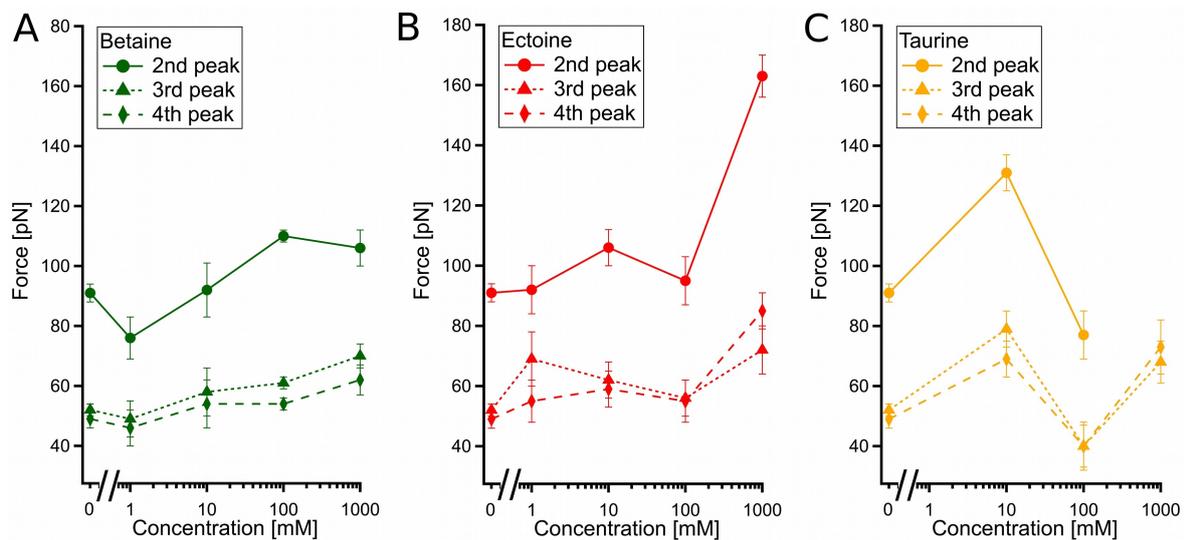


Figure 6.10: Dependency of the single peak dissociation force on osmolyte concentration for the second, third and fourth peak. The unfolding forces are the mean of the gaussian fits. The error bars are obtained from simulations, see section 6.1.1. A) For betaine we observe a slight increase in the unfolding force. B) Ectoine shows a non-linear dependency. C) For taurine the pattern is the same for all unfolding steps. For the 1M unfolding step for the second peak we did not observe enough events.

## 6 Dependency of the stability of BR on the concentration of compatible solutes

As explained in section 6.1, we do not have enough first unfolding steps with high quality WLC fits. Therefore we see in figure 6.10 the forces for the second, third and fourth unfolding peak (derived from the gaussian fits). The first data point for each peak is the buffer solution in absence of osmolytes (concentration of osmolytes: 0 M). For all osmolytes and concentrations the second peak has a higher unfolding force than the third and fourth ones confirming previous studies [OOP<sup>+</sup>00].

For betaine we can observe a slight increase in the unfolding force for all three peaks with increasing concentration. Furthermore the influence of the concentration on the unfolding forces is similar for all peaks.

For ectoine we see an increase as well, but a non-linear dependency of the unfolding force on the concentration. Also here the influence of the concentrations is approximately independent of the peak.

For taurine we also observe a non-linear dependency, although with higher fluctuations. A general increase cannot be noticed. Due to lower quality of the data, we could not analyse the second peak for 1 M for taurine. The force pattern is the same for all peaks.

All in all we see a general increase for betaine and ectoine in unfolding forces for all peaks. Additionally, for each of the three osmolytes the effect on the different peaks changes in a similar way with the concentration.

Kurz et al. discuss two effects how osmolytes influence the measured forces [Kur08]: On the one hand, additional hydrogen bonds can occur because of the binding of osmolytes to the protein. This leads to new intermediate peaks in the force curve. If these new bonds are close to already existing bonds, they will dissociate simultaneously, which leads to higher unfolding forces of the existing peaks. On the other hand, there can be an entropic effect: a change in the Gibbs free energy can make the transition from the folded to the unfolded state more unlikely. This also increases the forces of existing peaks. We cannot observe significantly additional intermediate unfolding peaks. This leads us to the assumption that there is no specific binding of the osmolytes to the protein and therefore the stability of BR is enhanced by the compatible solutes in general. Consequently the increase in force is most probably caused by a raise in the Gibbs free energy.

### **Influence of concentration changes on the persistence length**

The dependency of the persistence length on the increasing osmolyte concentration is shown in figure 6.11. For ectoine and taurine we observe a slight decrease in persistence length with increasing concentration. The pattern of the persistence length for the different peaks varies with the concentration.

## 6 Dependency of the stability of BR on the concentration of compatible solutes

For the concentration of betaine we detect no clear dependency of the persistence length. Also the pattern varies for each stretching step.

However, these trends are not statistically significant, therefore no clear conclusion can be drawn. The widths of the persistence length distributions in the histograms lead to the sizable errors.

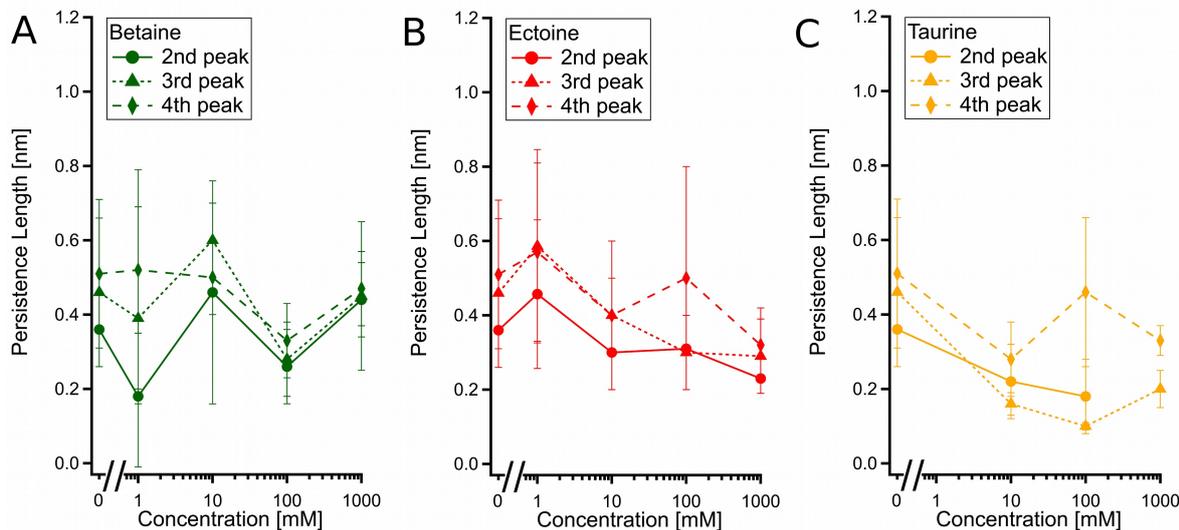


Figure 6.11: Dependency of single peak persistence length on osmolyte concentration for the second, third and fourth unfolding step for A) betaine, B) ectoine and C) taurine. The error bars are the standard deviation obtained from the gaussian fits.

The persistence length is a measure of the stiffness of the part of the protein which is already unfolded out of the membrane. The possible slight decrease for ectoine and taurine could be explained by a restructuring of water molecules via hydrogen bonds while the osmolytes are excluded from the vicinity of the protein [XT97a, XT97b, YJN07]. This forces the protein to acquire a more coiled conformation by burying its hydrophobic parts [Bur00, Sch02], leading to the decrease in persistence lengths. The persistence length is around 4 Å, see 6.11, which is a well-known measure for unfolded proteins [RGO<sup>+</sup>97], especially for BR [MKO<sup>+</sup>02]. Due to the known inaccuracy of WLC fit, see section 6.1.2.1, the errors are large because of the broad distributions. This could be improved in future analyses.

Arpita Roychoudhury\*, Adeline Bieker, Dieter Häussinger and Filipp Oesterhelt

## Membrane protein stability depends on the concentration of compatible solutes – a single molecule force spectroscopic study

**Abstract:** Compatible solutes are small, uncharged, zwitter ionic, osmotically active molecules produced and accumulated by microorganisms inside their cell to counteract different kinds of environmental stress. They enhance protein stability without interfering with the metabolic pathways even at molar concentrations. In this paper, we report the stabilizing effects of compatible solutes, ectoine, betaine and taurine on membrane protein bacteriorhodopsin at different concentrations. Using atomic force microscopy based single molecule force spectroscopy the impact of the osmolytes was quantified by measuring the forces required to pull the protein out of the membrane and the change in the persistence lengths of the unfolded polypeptide chain. Increase in unfolding forces were observed, indicating the strengthening of intramolecular interactions, which are vital for protein stability. The decrease in persistence lengths was recorded and showed increasing tendencies of the polypeptide strand to coil up. Interestingly, it was revealed that these molecules have different stabilizing effects on protein unfolding at different concentrations. The results show that the unfolding of single protein provides insight to the structure-dynamic relationship between the protein and compatible solute molecules at sub-nanometer scale. This also helps to understand the molecular mechanism involved in protein stabilization by organic osmolytes.

**Keywords:** atomic force microscopy; compatible solutes; membrane proteins; osmolytes; protein unfolding; single molecule force spectroscopy.

\*Corresponding author: **Arpita Roychoudhury**, Institute for Physical Biology, Heinrich-Heine-University, Universitätsstrasse. 1, Geb. 26.12, D-40225 Düsseldorf, Germany, e-mail: Arpita.Roychoudhury@uni-duesseldorf.de

**Adeline Bieker and Filipp Oesterhelt:** Institute for Physical Biology, Heinrich-Heine-University, Universitätsstrasse. 1, Geb. 26.12, D-40225 Düsseldorf, Germany

**Dieter Häussinger:** Department of Gastroenterology, Hepatology and Infectious Diseases, Heinrich-Heine-University, Moorenstrasse 5, D-40225 Düsseldorf, Germany

### Introduction

Compatible solutes are small, low molecular weight, naturally occurring compounds that are known to stabilize proteins against external stresses (Brown and Simpson, 1972; Clark, 1985). They are zwitter ionic with no net charge and belong to different categories, e.g., amino acids, their derivatives, methylamines, polyols, etc. (da Costa et al., 1998). They are accumulated inside the cell during adverse conditions such as extreme temperatures, pressures, chemical denaturation, UV radiation and the presence of high salt concentration or other solutes in their environment (Rockel et al., 2007; Warskulat et al., 2007a, 2008; Kurz, 2008).

They are known as compatible solutes because of their compatibility with the metabolism, as they do not interfere with the cellular composition and function even when they are accumulated in the cell at molar concentrations (Yancey et al., 1982; Burg, 1995). Several external stresses create a harsh condition for the cellular macromolecules, threatening the loss of their biological properties and functions (Pace, 1975; Privalov and Makhataadze, 1993; Xie and Freire, 1994). Hence, many compatible solutes are accumulated and used by the cells to counteract such denaturing effects by stabilizing the intracellular macromolecules, such as proteins.

A number of experimental studies involving ensemble measurements have shown that accumulation or addition of compatible solutes reduce and prevent structural destabilization and help restore the stability of the protein (Lee and Timasheff, 1981; Yancey et al., 1982; Foord and Leatherbarrow, 1998; Knapp et al., 1999; Lamosa et al., 2003). Organic osmolytes have also proven to restore healthy conditions and to counteract degenerative diseases, e.g., taurine deficiency diseases, which result in the loss of eyesight and reduced cardiac activity (Warskulat et al., 2007b; Delic et al., 2010).

Experimental works done by the groups of Galinski, Timasheff and Yancey show that compatible solutes like ectoine increase the melting temperature of the proteins, thereby rendering them thermo-dynamically more stable.

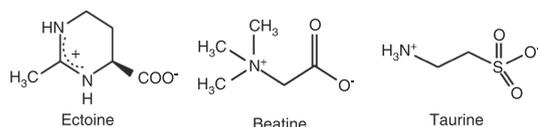
The stabilizing osmolytes exert a force that causes the protein to co-operatively fold into a native-like functional state from an unstructured conformation in aqueous environment (Murphy and Freire, 1992; Baskakov and Bolen, 1998; Burg, 2000; Schellman, 2002). Several studies have demonstrated that the stabilizing property of naturally occurring osmolytes correlate with the preferential exclusion of these osmolytes, from the vicinity of unfolded or denatured protein. This results in the formation of a denser and more structured water shell around the stretched protein (preferential hydration) (Xie and Timasheff, 1997a,b; Yu et al., 2007).

The protein backbone is osmophobic in nature (Cioni et al., 2005), so when the concentration or number of compatible solute molecules increase in the solvent-accessible area of the protein, the protein acquires a more compact conformation by hiding the backbone into its core in order to reduce exposure to the added solutes (Baskakov et al., 1998; Bolen, 2001; Yancey, 2005). This raises the enthalpy of the system and as a result the unfolded state of the protein will be more destabilized due to its greater solvent exposed area in turn resulting in a stabilization of the folded conformation.

In our studies, we used three different compatible solutes, namely ectoine, betaine and taurine (see Figure 1) and studied their effect on the mechanical stabilization of bacteriorhodopsin (BR) using unfolding experiments carried out by single molecule force spectroscopy (SMFS).

Ectoine is a heterocyclic amino acid accumulated in bacterial cells that maintains osmotic balance and serves as a protective agent for cells against hostile conditions (Galinski et al., 1985; Knapp et al., 1999). It is most commonly derived from the halophototrophic bacteria *Halomonas elongata*.

Betaine is known more for its role in osmo-regulation. Mammalian cells can accumulate betaine in the cytoplasm (Burg et al., 2007). It is mostly synthesized from choline in mammalian liver and kidney and known to be used as medicine, as a PCR rate enhancer, for tissue protection, and regulating eicosanoid formation in macrophages (Timasheff, 1992; Zhang et al., 1996).



**Figure 1** Chemical structure of compatible solutes used in the study. All of these three compatible solutes fall under the class of amino acids and their derivatives.

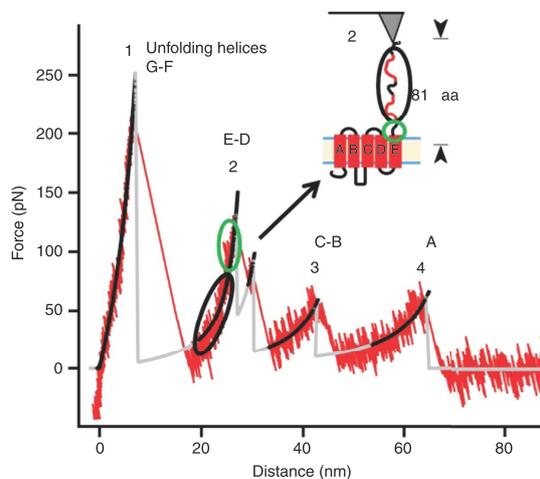
Taurine is an essential nutrient accumulated in mammalian liver and kidney for mammalian growth, particularly for neurological development. It helps in regulating the level of water and mineral salts in the blood. Sodium- and chloride-dependent taurine transporter TauT is known to carry out the osmoregulation during hypertonic condition (Häussinger, 1998).

Proteins are complex bio macromolecules that perform fundamental functions for the viability of the cell. Of all the proteins, transmembrane proteins constitute about 30% population and perform key roles as transporters, receptors, sensors, ion channels, catalysts and so on. But at the same time, the proteins belonging to this class are less stable and are difficult to handle. Hence, stabilization of membrane proteins is of great significance for carrying out various cellular activities.

Bacteriorhodopsin is a seven alpha helical transmembrane protein that belongs to the family of microbial rhodopsins and is found in the cellular membrane of the archaeon *Halobacterium salinarum* (Haupts et al., 1999). It contains a covalently bound retinal molecule, which absorbs light thereby contributing to the process of photosynthesis by the cell (Pedersen, 1995).

For SMFS an atomic force microscope (AFM) is used due to its ability to perform measurements on single proteins at natural conditions and to manipulate the matter at the nanoscale (Bustamante et al., 2000; Janovjak et al., 2006; Neuman and Nagy, 2008). A further benefit is that the protein BR is a well-established model system for AFM measurements (Oesterhelt et al., 2000). The unfolding of a membrane protein requires the pairwise extraction of its helices out of the membrane into the aqueous environment. A typical force curve for BR unfolding is shown in Figure 2. With this technique we can measure the unfolding curves of single proteins and determine for each unfolding step the dissociation force and force-related properties, e.g., persistence length. The force applied by the cantilever is used as external stress and the influence of different osmolytes at different concentrations on the unfolding curves is investigated.

Protecting effects of compatible solutes on soluble proteins have been studied to a fairly good extent, however studies on membrane proteins have not been concentrated on so far. Here, we report the findings of single molecule unfolding experiments of the membrane protein BR and the influence of the compatible solutes ectoine, betaine and taurine on its mechanical stabilization in presence of these osmolytes at various concentrations. The effects are represented by the change or difference in unfolding forces of individual BR molecule and their persistence lengths.



**Figure 2** Schematic representation of a single Bacteriorhodopsin molecule unfolding in presence of osmolytes. This scheme depicts the pairwise extraction of individual BR molecules out of the membrane. The green circle depicts the region where force acts on the protein to pull it out of the membrane. When a force is applied to the terminus, helices F–G will be pulled out of the membrane and get unfolded. Upon further retraction, the unfolded chain will be stretched and a force will be applied on helices D and E until they are extracted from the membrane and the same for C–B and A. The black ellipse denotes the stretched part with which persistence length is correlated. This value is obtained by fitting the curve with the worm like chain (WLC) model (described in detail in Materials and methods). The number of amino acids of the stretched chain can be counted using the analysis software IGOR. Each of the four peaks correspond to a defined chain length.

## Results

The organic osmolyte ectoine is known to stabilize the membrane protein BR (Roychoudhury et al., 2012). We extended the study to the osmolytes betaine and taurine. Furthermore we investigated the impact of the osmolyte concentration on the unfolding of BR.

For the analysis of the data we just included the force curves, which showed all four unfolding peaks. This proved that the protein was picked up at the terminus and unfolded completely.

The peaks were fitted with the worm-like chain (WLC) model. Only peaks with a good fit quality were considered (see the Materials and methods section for detailed information).

From the fit we derived the unfolding forces and persistence lengths for the four main peaks of each unfolding event. To determine whether osmolytes affect all unfolding steps similarly, or if the effect is different for each step,

we compared the unfolding forces and persistence lengths of the different peaks.

Persistence length and peak force can be attributed as a measure of the stability of the protein in the membrane. The more stable the protein is against external denaturing stress, the more force is needed to unfold it. The persistence length describes the thermodynamic stability. As external stress we used the force applied by the tip.

## Influence of concentration changes

For the distinction between the different unfolding peaks, we plotted two 2D scattered diagrams: the first as a function of (unfolding) peak forces vs. fit lengths and the second as persistence lengths vs. fit lengths (Figures 3A and C). We expect to see four distributions indicating the unfolding by the pairwise extraction of helices G–F, E–D, C–B and A of the protein. In Figures 3A and C, we can see just three distinct distributions for the force as well as the persistence length. The first distribution is missing because selection procedure only considered high quality fits.

In a next step we histogrammed the unfolding forces and persistence lengths according to their unfolding step (Figures 3B and D). The most probable unfolding force or persistence length is obtained as the position of the center of the peak by a Gaussian fit. This procedure was performed for all osmolytes and concentrations.

To determine how much the center of the Gaussian curve differs we simulated force curves.

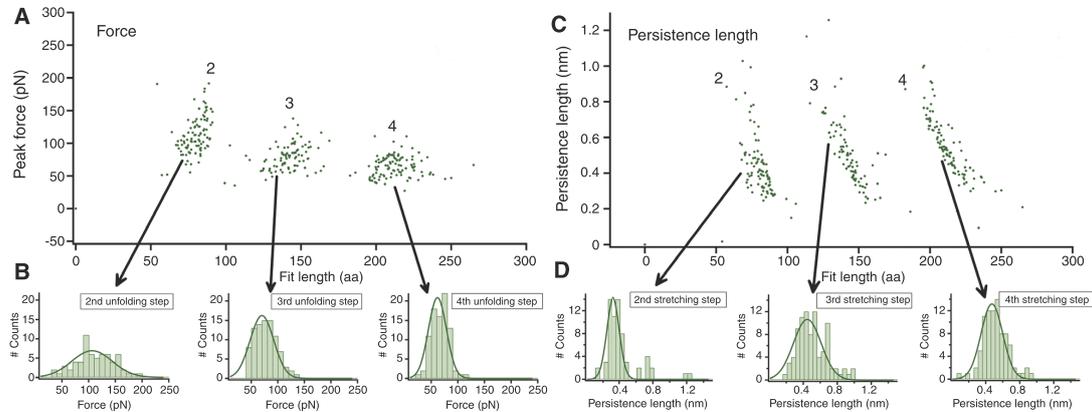
Then we analyzed the dependency of the center on the number of data counts in a histogram (see the Materials and methods section for detailed information). This error is used to estimate the significance of the forces. The error for the most probable persistence length is the standard deviation.

## Influence of concentration changes on unfolding forces

The concentration-dependent impact of compatible solutes on peak forces is presented in Figure 4. We see the forces for the second, third and fourth unfolding peak (derived from the Gaussian fits). As there is not enough data on the first unfolding step, these data are not shown.

The first data point is for the buffer solution in absence of any osmolyte ( $c = 0$  M). In general, the second peak has always a higher unfolding force than the third and fourth peaks, confirming previous studies (Oesterhelt

## 6 Dependency of the stability of BR on the concentration of compatible solutes

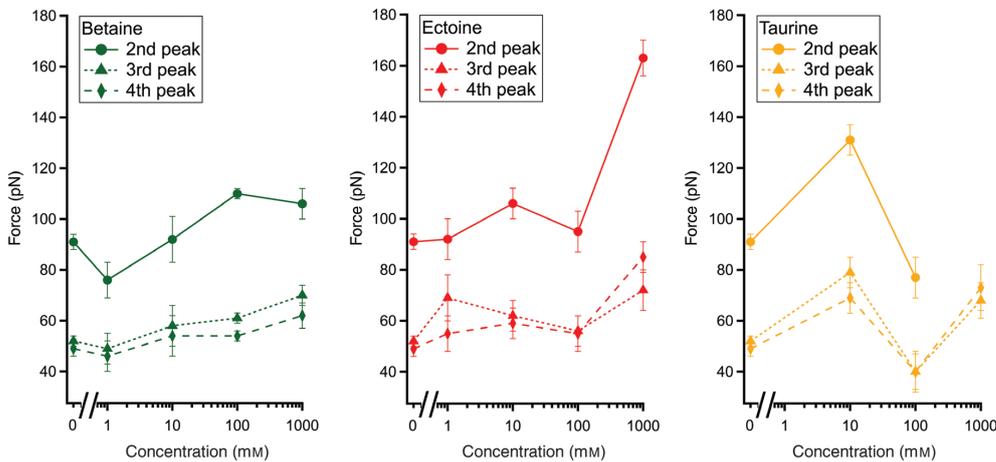


**Figure 3** Explanation of our analysis procedure for the different unfolding steps on the basis of 1 M betaine. (A) Scatter plot of the peak force vs. the fit length. We get three distributions for the pairwise unfolding of the helices E–D (second unfolding step), C–B (third unfolding step) and for the last unfolding of the helix A (fourth unfolding step). The histogramming for each unfolding step leads to the histograms shown in panel (B). A Gaussian fit reveals the most probable dissociation force. (C) Scatter plot of the persistence length vs. the fit length. We get three distributions for the stretching of the unfolded parts: the second distribution is the stretching of helices G–F, the third one for G–F and E–D, and the fourth one for the stretching of G–F, E–D and C–B. (D) Histograms of the persistence length for 1M betaine for the 2nd, 3rd and fourth stretching step.

et al., 2000). For betaine we can observe a slight increase in the unfolding force for all three peaks with increasing concentration. For ectoine we see an increase but a non-linear dependency of the unfolding force on the concentration. For taurine we also observe a non-linear dependency, although with higher fluctuations. A general

increase cannot be noticed. Because of lower quality data, we could not analyze the second peak at 1 M for taurine.

All in all, we see a general increase for betaine and ectoine in unfolding forces for all peaks. Additionally, for each of the three osmolytes the effect on the different peaks changes in a similar manner with the concentration.



**Figure 4** Dependency of single peak unfolding force on osmolyte concentration for the second, third and fourth peak. The unfolding forces are the mean of the Gaussian fits. The error bars are obtained from simulations, more explanations can be found in Materials and methods part. (A) For betaine we observe a slight increase in the unfolding force and the same pattern for all unfolding steps. (B) Ectoine shows within its peaks a similar course. The dependency is non-linear and shows a dip for 100 mM. (C) For taurine arises this dip, too. The pattern is also the same for all unfolding steps. For the 1 M unfolding step for the second peak we did not observe enough data. The amount of unfolding events per peak and concentration is listed in the Supplementary Information.

### Influence of concentration changes on persistence length

The dependency of the persistence length on the increasing osmolyte concentration is shown in Figure 5. For ectoine and taurine we observe a decrease in persistence length with increasing concentration. In contrast we detect almost no effect for betaine. The width of the persistence length distribution lead to the dimension of the errors (see the Materials and methods section for detailed information).

### Influence of concentration changes on the binding probability

The binding probability was reduced for all osmolytes except for 100 mM betaine. For ectoine and taurine we found a reduction of the binding probability to <25% for all concentrations compared to the buffer solution (0 M). Betaine showed also a reduction in most cases (for the number of data counts see SI).

## Discussion

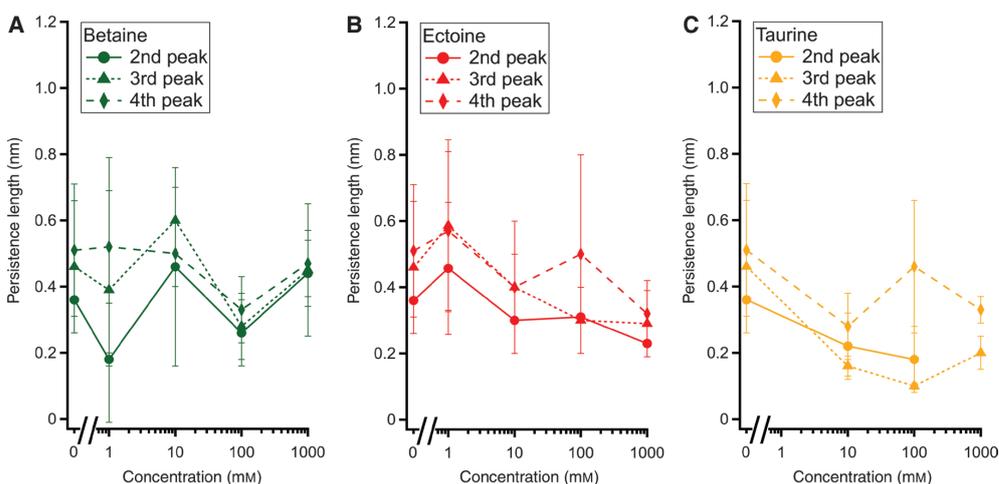
In this study, we investigated concentration-dependent effects of the compatible solutes ectoine, betaine and

taurine on the membrane protein bacteriorhodopsin with SMFS. Until now only ensemble measurements were performed (Nishiguchi and Somero, 1992; Oberdörfer et al., 2003; Rösger et al., 2007; Khan et al., 2010).

Bacteriorhodopsin represents one of the most extensively studied and best investigated membrane proteins, making it ideal for the study of various life processes (Oesterhelt, 1998; Haupts et al., 1999; Wagner et al., 2013). Hence it was suitable to be taken up for our study involving compatible solutes, which are known to stabilize protein structures and support folding (Somero, 1986; Timasheff, 1992; Burg and Ferraris, 2008; Hoffmann et al., 2009). However, very few studies have been carried out to investigate the effects of compatible solutes on membrane proteins (Flores Jiménez et al., 2010). The application of single molecule techniques for such stabilization studies is quite novel. Furthermore, little is known about the complex mechanism involved in this stabilization (Arakawa and Timasheff, 1985; Bolen, 2004; Street et al., 2006; Kurz, 2008).

The results of this force spectroscopic study of protein unfolding revealed stabilizing effects of the compatible solutes on BR: an increase in the unfolding forces, a decrease in persistence length and a reduction of the binding probability.

Due to already known non-specific surface interactions between the purple membrane and the AFM tip (Janovjak et al., 2003), there is a superposition of these adhesion peaks and the first unfolding peak. Therefore,



**Figure 5** Dependency of single peak persistence length on osmolyte concentration for the second, third and fourth unfolding step. The error bars are the standard deviation obtained from the Gaussian fits. (A) Betaine shows a non-linear dependency on concentration and the pattern for all unfolding steps varies. (B) Ectoine reveals a slight concentration-dependent decrease in the persistence length. (C) For taurine we observe also a slight concentration-dependent decrease and a similar pattern for all unfolding steps.

the fit quality for the first peaks decreased and consequently these peaks were deselected. Consequently we had not enough data for the first peak. Thus the first peak is not included in the 2D scatter plots and not further analyzed.

Because of our selection procedure (see the Materials and methods section for detailed information), the number of events for the second, third and fourth peaks within one concentration can differ (see Supplementary information).

### Increase in unfolding forces in presence of compatible solutes

In the concentration-dependent analysis of the forces we find that all peaks are affected in a similar manner for all osmolytes. In general the unfolding forces are increased. Two effects that influence measured forces are known. On the one hand, additional hydrogen bonds can occur because of the binding of osmolytes to the protein (Kurz, 2008). This leads to new intermediate peaks in the force curve. If these new bonds are close to already existing bonds, they will dissociate simultaneously, which leads to higher unfolding forces of the existing peaks. On the other hand, there can be an entropic effect: a change in the Gibbs free energy can make the transition from the folded to the unfolded state more unlikely (Kurz, 2008). This also increases the forces of existing peaks.

In our results we cannot observe additional intermediate unfolding peaks. This leads us to the assumption that there is no specific binding of the osmolytes to the protein. Consequently the increase in force is caused by a raise in the Gibbs free energy.

### Influence of compatible solutes on persistence length

The persistence length is a mechanical property related to the measure of the stiffness of the part of the polypeptide, which is already unfolded out of the membrane. Upon addition of the osmolytes the water molecules are restructured via hydrogen bonds while excluding the osmolytes of the vicinity of the protein. This forces the protein to acquire a more coiled conformation by burying its hydrophobic parts. This leads to the observed decrease in persistence lengths.

The persistence length is around 4 Å (see Figure 5), which is a well-known measure for unfolded proteins (Rief et al., 1997), especially for BR (Müller et al., 2002).

In the scatter plot for the persistence length (see Figure 3C) the distribution is bent. This results from the WLC fit: the WLC model is known (Kühner and Gaub, 2006; Kühner et al., 2007) to give different persistence lengths when it is fitted to different force ranges. Thus a high rupture force leads to a lower persistence length. Therefore the errors are overestimated because of the broad distributions.

### Influence of the osmolytes on the binding probability

We found a reduction in the binding probability of the tip to the terminus of the protein by a factor of three to ten independently of the osmolytes and their concentration. The tip of the cantilevers used is made of silicon nitride, which is a solid state surface. Thus the decreased binding probability indicates a general effect of osmolytes to reduce the interaction between proteins and solid state surfaces.

The increase in binding probability for 100 mM betaine needs to be further investigated.

With our concentration-dependent measurements we could show that BR is affected by the compatible solutes, however there is no specific binding of the solutes to the protein. Hence, we conclude that the stability of BR is enhanced by the compatible solutes in general.

### Protein stability is governed by water structure

Compatible solutes dissolve well in aqueous solutions even at very high concentrations. This is caused by their favorable interaction with water. Compatible solutes form hydrogen bonds with the water molecules so that a densely arranged water cluster is formed (Wiggins, 1990; Zou et al., 2002; Bennion and Daggett, 2003; Klimov et al., 2004). The observed reduction in the binding probability can be explained by this effect: because of the rearrangement of the water by the osmolytes, the water shell around the protein surface becomes more structured, therefore decreasing the unspecific tip-protein interaction.

During the unfolding of the protein a greater structuring of the hydration shell around the unfolded part occurs and results in the thermodynamic stabilization of the protein.

The compatible solutes prefer interacting with the surrounding water molecules, than with the polypeptide backbone, preserving the preferential hydration of the protein (Bolen and Baskakov, 2001; Roberts, 2005). The

exclusion of compatible solutes from the protein surface increases the concentration of solutes in the bulk solution. This results in an increase of surface tension of water, which drives the protein into a more compact structure.

In accordance with the preferential exclusion model (Liu and Bolen, 1995; Plaza del Pino and Sanchez-Ruiz, 1995; Timasheff, 2002a,b; Kurz, 2008), the results of our experiments show a stabilization of the membrane protein structure in presence of osmolytes. They also substantiate the fact that different osmolytes have different stabilizing effects on unfolding forces and persistence lengths. Osmolytes facilitate folding of BR, and additionally support it in withstanding the unfolding stress by creating a denser water shell around it. Our experimental results show a decrease in the persistence lengths of the stretched amino acid chain of BR, which indicates that under the influence of osmolytes there is a stronger tendency of the entropic coil to collapse and fold into a dense conformation by expelling the solute molecules from the exposed protein surface area.

### Stabilization of membrane proteins in addition to globular proteins

In general, osmolytes are widely known for their stabilizing properties on globular proteins (Lippert and Galinski, 1992; Oberdörfer et al., 2003, Lentzen and Schwarz, 2006) and in stabilizing the structure of lipid bilayers (Harishchandra et al., 2010). Oberdörfer et al. (2003) previously showed that there were no significant changes in the unbinding force of the FN-III domains under the influence of the compatible solute ectoine. The geometry of the stabilizing intramolecular hydrogen bonds might be one explanation that the FN-III domains did not show any increase in unfolding forces. On BR we observed an increase in unfolding forces in presence of ectoine as well as betaine and taurine (Figure 5). For the rise in the force, several hydrogen bonds need to be broken simultaneously. This leads to the high enthalpic contribution to the unfolding force and at the same time, as mentioned before, the absence of any additional unfolding peak indicate that no new and specific stabilizing interactions were established between the protein and the compatible solutes.

This suggests that stabilization occurs by entropic contributions, which may be weak compared to the strong enthalpic contribution to the system. Furthermore, compatible solutes also act as a stabilizing agent against denaturing stress (in this case, the unfolding force applied) for membrane proteins as well as for globular proteins.

Because of the increasing clinical relevance of compatible solutes in understanding and curing human diseases, we hope that our results and findings will help and lead to further investigations of osmolytes and their use in drug development with respect to the important role of membrane proteins as transporters and drug targets.

## Materials and methods

### Sample preparation

Native purple membrane patches (see Acknowledgements) from *Halobacterium salinarum* were dissolved in solution buffer. The sample of bacteriorhodopsin diluted in buffer, was then adsorbed onto freshly cleaved mica surfaces (Müller et al., 1997).

The effect of all the three compatible solutes on mechanical unfolding of membrane protein were studied under several concentrations (1 mM, 10 mM, 100 mM and 1 M) in a solution of 20 mM Tris +300 mM KCl buffer in nanopure water ( $R > 18 \text{ M } \Omega \text{ cm}$ ) at a physiological pH of 7.8. The sample containing the purple membranes was diluted such that the solution is not too dense and the adsorption time was chosen to be around 15 min to ensure high coverage, but no overlapping of membrane patches. After 15 min of incubation time, the sample was washed with the same buffer for three times to wash away the unadsorbed protein and the sample was then immersed in the standard buffer with or without the compatible solutes at above-mentioned concentrations for concentration-dependent unfolding experiments. We carried out the same number of pulling experiments for each condition.

Ectoine was provided by BiTop, Witten, Germany and betaine and taurine were purchased from Sigma Aldrich, Germany. All of these had analytical grade purity of 99.9%.

### Atomic force microscopy and single molecule force spectroscopy

The membrane patches were imaged using an AFM from MFP 3D, Asylum Research (Santa Barbara, CA) instrument and Olympus OMCL TR400 silicon nitride cantilevers. Imaging was performed in fluid using contact mode with the same cantilever like in the force spectroscopy measurements. An image of such a membrane patch is given in the Supplementary Information.

After imaging, mechanical unfolding experiments were carried out on single proteins by selecting a patch and thereafter directing the tip of the cantilever on top of the selected patch.

The attachment of the terminus to the tip was obtained via non-specific adsorption by pressing the cantilever onto the sample. Force curves were recorded at a retraction speed of 400 nm/s. The stiffness of the cantilever exhibited by the measure of the spring constant was calculated for each unloaded cantilever by positioning it away from the surface and taking the mean of thermal vibration signal and applying the equipartition theorem (Hutter and Bechhöfer, 1993). The spring constant calculated by this method was measured to be around 20 pN/nm.

### Data analysis

Force vs. extension curves were recorded in repeated pull and release cycle using self-written procedures implemented in the Asylum AFM software. Selected force vs. extension curves were corrected for virtual deflection of the cantilever, displacement of the cantilever, tip-surface conversion and other factors using the graph correction options in Igor. Afterwards the curves were analyzed by fitting each force peak with WLC model with a monomer length of 3.6 Å in Igor using self-written procedures. The WLC model gives the simple approximate force extension relationship of BR unfolding, given by:

$$F = KB T/p (1/4(1-x/L0)^{-2} - 1/4 + x/L0).$$

Here  $F$  is the applied force,  $x$  is the end to end distance,  $KB$  is the Boltzmann constant,  $p$  is the persistence length and  $L0$  is the contour length, i.e., the length of the completely stretched polymeric chain. The WLC model was used to fit the force and persistence length in the protein unfolding force curve; although our system is a complicated one and this model is a simple model based on assumptions, which do not fully comply in accordance with the requirements of our system. But the WLC has been often found to have given reliable results for studies involving force spectroscopic techniques for stretching polypeptide molecules.

The recorded force curves showed the typical characteristic pattern of four main peaks (Figure 2) (Oesterhelt et al., 2000). In these force curves, each main peak represents the extraction of an alpha helical pair of the protein out of the membrane, which is typical for various seven transmembrane helical proteins of the rhodopsin family (Sharma et al., 2006; Klare et al., 2008; Oberbarnscheidt et al., 2009). Additionally, force curves reveal further side peaks representing unfolding intermediates revealing different unfolding pathways that differ depending on the experimental condition (Müller et al., 2002; Janovjak et al., 2006; Kedrov et al., 2007; Cisneros et al., 2008).

From the number of amino acids obtained from each fit, counting back from the C-terminus, it can be calculated which part of the protein is left in the membrane as an intermediate.

Persistence length ( $P$ ) is derived from fitting the curvature of each peak and it is a measure of elasticity or bending rigidity of the protein. The unfolding force ( $F$ ) is the measure of mechanical force required to extract the protein out of the membrane. It is derived by fitting the highest point of each peak before it shows a rupture resulting in the drop of unfolding force. Localization of stabilizing

intramolecular interactions reveal that helices G and F, D and E, and B and C, respectively, mostly unfold pair wise, sometimes showing less stable intermediates where only one helix is extracted out of the membrane. The seventh helix, A, is then extracted from the membrane in the last step.

### Error calculation

Because of the thermodynamic influence on the dissociation the unfolding force does not have a distinct value but a broad distribution. Therefore we performed Monte Carlo simulations of the dissociation curves to determine the standard deviation of the force histogram without any experimental error. After analyzing the simulated curves, we made histograms with as many curves that were taken for plotting the histograms from the experimentally recorded data. This we performed ten times and fitted a Gaussian curve to them, which we used to calculate the mean force and its standard deviation. We used this standard deviation as the error of the force in the graph showing unfolding force for each intermediate peak dependent on concentration (Figures 4 and 5).

For the error of the persistence length of the single peak analysis, we used the standard deviation value of the Gaussian fit to the persistence length histogram.

**Acknowledgments:** We gratefully acknowledge support (and training) from the International NRW Research School BioStruct, granted by the Ministry of Innovation, Science and Research of the State North Rhine-Westphalia, the Heinrich-Heine-University of Düsseldorf, the Entrepreneur Foundation at the Heinrich-Heine-University of Düsseldorf and the German Research Foundation (DFG). Native purple membrane patches of bacteriorhodopsin were a gift from the group of Prof. Dr. Norbert Hampp, Department of Chemistry, University of Marburg, 35032 Marburg, Germany.

Received May 8, 2013; accepted September 4, 2013; previously published online September 7, 2013

## References

- Arakawa, T. and Timasheff, S.N. (1985). The stabilization of proteins by osmolytes. *Biophys. J.* 47, 411–414.
- Baskakov, I.V. and Bolen, D.W. (1998). Forcing thermodynamically unfolded proteins to fold. *J. Biol. Chem.* 273, 4831–4834.
- Baskakov, I.V., Wang, A., and Bolen, D.W. (1998). Trimethylamine-N-oxide counteracts urea effects on rabbit muscle lactate dehydrogenase function: a test of the counteraction hypothesis. *Biophys. J.* 74, 2666–2673.
- Bennion, B.J. and Daggett, V. (2003). The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Acad. Sci. USA* 100, 5142–5147.
- Bolen, D.W. (2001). Protein stabilization by naturally occurring osmolytes. *Methods Mol. Biol.* 168, 17–36.
- Bolen, D.W. and Baskakov, I.V. (2001). The osmophobic effect: natural selection of a thermodynamic force in protein folding. *J. Mol. Biol.* 310, 955–963.
- Bolen, D.W. (2004). Effects of naturally occurring osmolytes on protein stability and solubility: issues important in protein crystallization. *Methods* 34, 312–322.
- Brown, A.D. and Simpson J.R. (1972). Water relations of sugar-tolerant yeasts: the role of intracellular polyols. *J. Gen. Microbiol.* 72, 589–591.
- Burg, M.B. (1995). Molecular basis of osmotic regulation. *Am. J. Physiol.* 268, F983–F996.
- Burg, M.B. (2000). Macromolecular crowding as a cell volume sensor. *Cell. Physiol. Biochem.* 10, 251–256.

- Burg, M.B., Ferraris, J.D., and Dmitrieva, N.I. (2007). Cellular response to hyperosmotic stresses. *Physiol. Rev.* *87*, 1441–1474.
- Burg, M.B. and Ferraris, J.D. (2008). Intracellular organic osmolytes: function and regulation. *J. Biol. Chem.* *283*, 7309–7313.
- Bustamante, C., Macosko, J.C., and Wuite, G.J.L. (2000). Grabbing the cat by the tail: manipulating molecules one by one. *Nat. Rev. Mol. Cell Biol.* *1*, 130–136.
- Cioni, P., Bramanti, E., and Strambini, G.B. (2005). Effects of sucrose on the internal dynamics of azurin. *Biophys. J.* *88*, 4213–4222.
- Cisneros, D.A., Oberbarnscheidt, L., Pannier, A., Klare, J.P., Helenius, J., Engelhard, M., Oesterhelt, F., and Müller, D.J. (2008). Transducer binding establishes localized interactions to tune sensory rhodopsin II. *Structure* *16*, 1206–1213.
- Clark, M.E. (1985). The osmotic role of amino acids: Discovery and function. In: *Transport processes, ionic- and osmoregulation*. Gilles, R. and Gilles-Baillien, M. (eds.) (New York, USA: Springer), pp. 412–423.
- da Costa, M.S., Santos, H., and Galinski, E.A. (1998). An overview of the role and diversity of compatible solutes in Bacteria and Archaea. *Adv. Biochem. Eng. Biotechnol.* *61*, 117–153.
- Delic, D., Warskulat, U., Borsch, E., Al-Qahtani, S., Al-Quraishi, S., Häussinger, D., and Wunderlich, F. (2010). Loss of ability to self-heal malaria upon taurine transporter deletion. *Infect. Immun.* *78*, 1642–1649.
- Flores Jiménez, R.H., Do Cao, M.A., Kim, M., and Cafiso, D.S. (2010). Osmolytes modulate conformational exchange in solvent-exposed regions of membrane proteins. *Protein Sci.* *19*, 269–278.
- Food, R.L. and Leatherbarrow, R.J. (1998). Effect of osmolytes on the exchange rates of backbone amide protons in proteins. *Biochemistry* *37*, 2969–2978.
- Galinski, E.A., Pfeiffer, H.P., and Trüper, H.G. (1985). 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinocarboxylic acid: a novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. *Eur. J. Biochem.* *149*, 135–139.
- Harishchandra, R.K., Wulff, S., Lentzen, G., Neuhaus, T., and Galla, H.J. (2010). The effect of compatible solute ectoines on the structural organization of lipid monolayer and bilayer membranes. *Biophys. Chem.* *150*, 37–46.
- Haupts, U., Tittor, J., and Oesterhelt, D. (1999). Closing in on bacteriorhodopsin: progress in understanding the molecule. *Annu. Rev. Biophys. Biomol. Struct.* *28*, 367–399.
- Häussinger, D. (1998). Osmoregulation of liver cell function: signalling, osmolytes and cell heterogeneity. *Contrib. Nephrol.* *123*, 185–204.
- Hoffmann, E.K., Lambert, I.H., and Pedersen, S.F. (2009). Physiology of Cell Volume Regulation in Vertebrates. *Physiol. Rev.* *89*, 193–277.
- Hutter, J.L. and Bechhöfer, J. (1993). Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.* *64*, 1868–1873.
- Janovjak, H., Kessler, M., Oesterhelt, D., Gaub, H.E., and Müller, D.J. (2003). Unfolding pathways of native bacteriorhodopsin depend on temperature. *EMBO J.* *22*, 5220–5229.
- Janovjak, H., Kedrov, A., Cisneros, D.A., Sapra, K.T., Struckmeier, J., and Müller, D.J. (2006). Imaging and detecting molecular interactions of single transmembrane proteins. *Neurobiol. Aging* *27*, 546–561.
- Kedrov, A., Janovjak, H., Sapra, K.T., and Müller, D.J. (2007). Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. *Annu. Rev. Biophys. Biomol. Struct.* *36*, 233–260.
- Khan, S.H., Ahmad, N., Ahmad, F., and Kumar, R. (2010). Naturally Occurring Organic Osmolytes: From Cell Physiology to Disease Prevention. *IUBMB Life.* *62*, 891–895.
- Klare, J.P., Chizhov, I., and Engelhard, M. (2008). Microbial rhodopsins: scaffolds for ion pumps, channels, and sensors. *Results Probl. Cell Differ.* *45*, 73–122.
- Klimov, D.K., Straub, J.E., and Thirumalai, D. (2004). Aqueous urea solution destabilizes Abeta(16–22) oligomers. *Proc. Natl. Acad. Sci. USA* *101*, 14760–14765.
- Knapp, S., Ladenstein, R., and Galinski, E.A. (1999). Extrinsic protein stabilization by the naturally occurring osmolytes beta-hydroxyectoine and betaine. *Extremophiles* *3*, 191–198.
- Kurz, M. (2008). Compatible solute influence on nucleic acids: many questions but few answers. *Saline Systems* *4*, 6. DOI: 10.1186/1746-1448-4-6.
- Kühner, F. and Gaub, H.E. (2006). Modelling cantilever-based force spectroscopy with polymers. *Polymer* *47*, 2555–2563.
- Kühner, F., Morfill, J., Neher, R.A., Blank, K., and Gaub, H.E. (2007). Force-Induced DNA Slippage. *Biophys. J.* *92*, 2491–2497.
- Lamosa, P., Turner, D.L., Ventura, R., Maycock, C., and Santos, H. (2003). Protein stabilization by compatible solutes. Effect of diglycerol phosphate on the dynamics of *Desulfovibrio gigas* rubredoxin studied by NMR. *Eur. J. Biochem.* *270*, 4604–4614.
- Lee, J.C. and Timasheff, S.N. (1981). The stabilization of proteins by sucrose. *J. Biol. Chem.* *256*, 7193–7201.
- Lentzen, G. and Schwarz, T. (2006). Extremolytes: Natural compounds from extremophiles for versatile applications. *Appl. Microbiol. Biotechnol.* *72*, 623–634.
- Lippert, K. and Galinski, E.A. (1992). Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl. Microbiol. Biotechnol.* *37*, 61–65.
- Liu, Y. and Bolen, D.W. (1995). The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry* *34*:12884–12891.
- Murphy, K.P. and Freire, E. (1992). Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Protein Chem.* *43*, 313–361.
- Müller, D.J., Amrein, M., and Engel, A. (1997). Adsorption of biological molecules to a solid support for scanning probe microscopy. *J. Struct. Biol.* *119*, 172–188.
- Müller, D.J., Kessler, M., Oesterhelt, F., Möller, C., Oesterhelt, D., and Gaub, H.E. (2002). Stability of bacteriorhodopsin alpha-helices and loops analyzed by single-molecule force spectroscopy. *Biophys. J.* *83*, 3578–3588.
- Neuman, K.C. and Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Meth.* *5*, 491–505.
- Nishiguchi, M.K. and Somero, G.N. (1992). Temperature- and concentration-dependence of compatibility of the organic osmolyte beta-dimethylsulfoniopropionate. *Cryobiology* *29*, 118–124.
- Oberbarnscheidt, L., Janissen, R., Martell, S., Engelhard, M., and Oesterhelt, F. (2009). Single-molecule force spectroscopy measures structural changes induced by light activation and transducer binding in Sensory Rhodopsin II. *J. Mol. Biol.* *394*, 383–390.
- Oberdörfer, Y., Schrot, S., Fuchs, H., Galinski, E.A., and Janshoff, A. (2003). Impact of compatible solutes on the mechanical properties of fibronectin: a single molecule analysis. *PCCP* *5*, 1876–1881.

- Oesterhelt, D. (1998). The structure and mechanism of the family of retinal proteins from halophilic archaea. *Curr. Opin. Struct. Biol.* 8, 489–500.
- Oesterhelt, F., Oesterhelt, D., Pfeiffer, M., Engel, A., Gaub, H.E., and Müller, D.J. (2000). Unfolding pathways of individual Bacteriorhodopsins. *Science* 288, 143–146.
- Pace, C.N. (1975). The stability of globular proteins. *Crit. Rev. Biochem.* 3, 1–43.
- Pedersen, P.L. (1995). In: *Biochemistry*, 2nd edn. Voet, D. and Voet, J.G. (eds.) (New York, USA: Wiley).
- Plaza del Pino, I.M. and Sanchez-Ruiz, J.M. (1995). An osmolyte effect on the heat capacity change for protein folding. *Biochemistry* 34, 8621–8630.
- Privalov, P.L. and Makhatadze, G.I. (1993). Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration. *J. Mol. Biol.* 232, 660–679.
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J.M., and Gaub, H.E. (1997). Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* 276, 1109–1112.
- Roberts, M.F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Systems* 1, 5. DOI:10.1186/1746-1448-1-5.
- Rockel, N., Esser, C., Grether-Beck, S., Warskulat, U., Flögel, U., Schwarz, A., Schwarz, T., Yarosh, D., Häussinger, D., and Krutmann, J. (2007). The osmolyte taurine protects against ultraviolet B radiation-induced immunosuppression. *J. Immunol.* 179, 3604–3612.
- Rösgen, J., Pettitt, B.M., and Bolen, D.W. (2007). An analysis of the molecular origin of osmolyte-dependent protein stability. *Protein Sci.* 16, 733–743.
- Roychoudhury, A., Häussinger, D., and Oesterhelt, F. (2012). Effect of the compatible solute ectoine on the stability of the membrane proteins. *Protein Pept. Lett.* 19, 791–794.
- Schellman, J.A. (2002). Fifty years of solvent denaturation. *Biophys. Chem.* 96, 91–101.
- Sharma, A.K., Spudich, J.L., and Doolittle, W.F. (2006). Microbial rhodopsins: functional versatility and genetic mobility. *Trends Microbiol.* 14, 463–469.
- Somero, G.N. (1986). Protons, osmolytes, and fitness of internal milieu for protein function. *Am. J. Physiol.* 251, R197–R213.
- Street, T.O., Bolen, D.W., and Rose, G.D. (2006). A molecular mechanism for osmolyte-induced protein stability. *Proc. Natl. Acad. Sci. USA* 103, 13997–14002.
- Timasheff, S.N. (1992). A physicochemical basis for the selection of osmolytes by nature. In: *Water and life*, G.N. Somero, ed. (New York: Springer), pp. 70–84.
- Timasheff, S.N. (2002a). Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. Acad. Sci. USA* 99, 9721–9726.
- Timasheff, S.N. (2002b). Protein hydration, thermodynamic binding, and preferential hydration. *Biochemistry* 41, 13473–13482.
- Wagner, N.L., Greco, J.A., Ranaghan, M.J., and Birge, R.R. (2013). Directed evolution of bacteriorhodopsin for applications in bioelectronics. *J. R. Soc. Interface.* 10, 20130197.
- Warskulat, U., Brookmann, S., Reinen, A., and Häussinger, D. (2007a). Ultraviolet B radiation induces cell shrinkage and increases osmolyte transporter mRNA expression and osmolyte uptake in HaCaT keratinocytes. *Biol. Chem.* 388, 1345–1352.
- Warskulat, U., Heller-Stilb, B., Oermann, E., Zilles, K., Haas, H., Lang, F., and Häussinger, D. (2007b). Phenotype of the taurine transporter knockout mouse. *Methods Enzymol.* 428, 439–458.
- Warskulat, U., Brookmann, S., Felsner, I., Brenden, H., Grether-Beck, S., and Häussinger, D. (2008). Ultraviolet A induces transport of compatible organic osmolytes in human dermal fibroblasts. *Exp. Dermatol.* 17, 1031–1036.
- Wiggins, P.M. (1990). Role of water in some biological processes. *Microbiol. Rev.* 54, 432–449.
- Xie, D. and Freire, E. (1994). Molecular basis of cooperativity in protein folding. V. Thermodynamic and structural conditions for the stabilization of compact denatured states. *Proteins Struct. Func. Genet.* 19, 291–301.
- Xie, G.F. and Timasheff, S.N. (1997a). Mechanism of the stabilization of ribonuclease A by sorbitol: preferential hydration is greater for the denatured than for the native protein. *Protein Sci.* 6, 211–221.
- Xie, G.F. and Timasheff, S.N. (1997b). The thermodynamic mechanism of protein stabilization by trehalose. *Biophys. Chem.* 64, 25–43.
- Yancey, P.H., Clark, M.E., Hand, S.C., Bowlus, R.D., and Somero, G.N. (1982). Living with water stress: evolution of osmolyte systems. *Science* 217, 1214–1222.
- Yancey, P.H. (2005). Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J. Exp. Biol.* 208, 2819–2830.
- Yu, I., Jindo, Y., and Nagaoka, M. (2007). Microscopic understanding of preferential exclusion of compatible solute ectoine: direct interaction and hydration alteration. *J. Phys. Chem. B.* 111, 10231–10238.
- Zhang, F., Warskulat, U., Wettstein, M., and Häussinger, D. (1996). Identification of betaine as an osmolyte in rat liver macrophages (Kupffer cells). *Gastroenterology* 110, 1543–1552.
- Zou, Q., Bennion, B.J., Daggett, V., and Murphy, K.P. (2002). The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea. *J. Am. Chem. Soc.* 124, 1192–1202.

# 7 Conclusion and Outlook

Single-molecule force spectroscopy represents a powerful tool to investigate intra- and intermolecular interactions by the determination of kinetic and thermodynamic properties. In the first part of this work, I dealt with the challenge of theoretically modeling, analysing and interpreting single-molecule force spectroscopy data. The second part investigated the concentration-dependent influence of osmolytes on each single unfolding step of Bacteriorhodopsin in force spectroscopy data. In the following I discuss the parts separately.

## 7.1 Analysing single-molecule force spectroscopy data

The dissociation of chemical bonds induced by external forces is a thermally activated process and can be described with the help of Kramers' rate theory [ER97]. Due to the mathematical inverse problem in force spectroscopy measurements, the interpretation of experimental force spectroscopy data is still challenging. I developed a model to fit the force-dependent dissociation rate, allowing to consider a possible non-linear increase between the potential minimum and the potential barrier. If experimental data show a bending in the force-dependent dissociation-rate, this is evidence that the underlying energy landscape does not have a linear increase between the minimum and maximum. By the assumption of a simple decrease based on the order  $n$ , the second order shows the biggest difference to the linear case in the force-dependent dissociation rate, and allows the separation of the parameter  $\Delta G_0$ , describing the potential barrier height, and the parameter  $\nu_0$ , describing the attempt frequency. An increase in the order  $n$  leads again to a stronger correlation of these parameters and also an increasing indistinguishability of the underlying order of the potential. Furthermore I analysed the influence of the assumed shape of the energy landscape on the parameter determination. The parameter  $\Delta x_0$ , describing the potential width,

is linearly dependent on the fit order; similarly,  $\Delta G_0$  and  $\nu_0$  show a strong dependence. This leads to an under- or overestimation of these parameters if the wrong fit order is chosen.

The application of the analysis strategy combined with a better FJC fit revealed a remarkably short lifetime of the mutant-DNA-complex compared to the binding of the wildtype to DNA. This is presumably in accordance with an unspecific binding of the DNA to gold and the binding of the mutant to the DNA is completely disrupted.

High-speed force spectroscopy can broaden the force-range much more than conventional AFMs can, because retraction velocities up to 4000  $\mu\text{m/s}$  are measurable [RGC<sup>+</sup>13]. This will promote the investigation of inter- and intramolecular interactions. However, have to bear in mind the assumption that structural changes within the molecular complex are much faster than the changes in the pulling force. It is possible that this assumption is not realistic anymore for measurements with high retraction velocities. Therefore the description of high-speed force spectroscopy will require a modified modeling of the force-dependent molecular dissociation.

### **7.2 Differentiation of the single unfolding steps to determine the concentration-dependent influence of compatible solutes**

Compatible solutes are known to have a general stabilising effect on globular proteins [OSF<sup>+</sup>03, LS06], but the stabilising mechanism is still under investigation, see for example [BB98, Bol01, VLL09]. Because of severe technical difficulties, the use of single-molecule techniques is still in its infancy for membrane proteins. This is the first concentration-dependent study of the influence of osmolytes on a membrane protein.

By the analysis of each unfolding step the effect of osmolytes on the different structural parts gives an insight on the interaction of the osmolytes with the protein. The dissociation forces showed for each osmolyte-concentration and unfolding step the same pattern. We could not observe additional intermediate unfolding steps significantly. This let us to assume that there is no specific binding to the protein. Furthermore a slight increase in force with increasing concentration of osmolytes was

## 7 Conclusion and Outlook

observed which is a consequence of the stabilising effect. The persistence length showed a slight decrease with increasing concentration of osmolytes. This indicates an increased tendency of the protein to coil up outside the membrane. A more coiled structure supports the refolding of the protein.

For better insight the data amount has to be increased. Furthermore the WLC fit routine has to be improved. This would decrease the uncertainties in the persistence length. Also the developed analysis strategy could be used to get information about the influence of the concentration of osmolytes on the energy landscape.

This work shows that single-molecule force spectroscopy is an important method to achieve a better insight in molecular interactions, especially allowing measurements on challenging biological samples like membrane proteins. These results hold the promise of improved understanding of inter- and intramolecular interactions on the structural level yielding eventually advances in the development of specific drugs to influence the metabolism.



## Summary

The cell is the structural unit of life. For understanding the metabolic processes and related diseases, it is important to understand inter- and intramolecular interactions. This will lead to a development of drugs against specific targets like a protein to prevent unwanted interactions causing side effect. Single-molecule force spectroscopy is a common technique to investigate inter- and intramolecular interactions. Specifically, it can be used to observe the dissociation of the molecular bond in dependence on an external force. This forced dissociation process can be viewed as a thermally activated escape over a potential barrier of the energy landscape describing the molecular interaction, but the determination and interpretation of parameters describing the energy landscape obtained from force spectroscopy data is challenging. I develop a model to fit the force-dependent dissociation rate to determine the potential width  $\Delta x_0$ , the barrier height  $\Delta G_0$  and the attempt frequency  $\nu_0$ . To that aim I consider a generalized dependence of the energy landscape on the interaction coordinate, characterized by the order  $n$ . By testing the model with simulated data, I reveal a linear dependency of  $\Delta x_0$  on the fit order; similarly,  $\Delta G_0$  and  $\nu_0$  show a strong dependence. This leads to an under- or overestimation of these parameters if the wrong fit order is chosen. The application of this analysis strategy to experimental data combined with an improved freely jointed chain fit to experimental data reveals a remarkably short lifetime of the R176A-DNA-complex compared to the binding of the wildtype to DNA. Upon further analysis, this is interpreted as an unspecific binding of the DNA to gold, therefore the binding of the mutant to the DNA seems to be completely disrupted.

The atomic force microscope is one of the few methods allowing to investigate the unfolding of a single membrane protein. In the second part of this work we use this technique for the first time to analyse the concentration-dependent influence of compatible solutes on the unfolding of Bacteriorhodopsin. By the analysis of each unfolding step the effect of osmolytes on the different structural parts gives an insight on the interaction of the osmolytes with the protein. The dissociation forces show for each osmolyte-concentration and unfolding step the same pattern. We do not observe significantly additional intermediate unfolding steps, which we interpret as the absence of specific binding to the protein. Furthermore a slight increase in force with increasing concentration of osmolytes is observed indicating the stabilising effect of osmolytes. The persistence length shows a possible slight decrease with increasing concentration of osmolytes. This indicates an increased tendency of the protein to

coil up outside the membrane which supports the refolding of the protein. These results hold the promise of improved understanding of inter- and intramolecular interactions yielding eventually advances in the development of specific drugs.

## Zusammenfassung

Die Zelle ist das strukturelle Bauelement von Leben. Um die Stoffwechselprozesse und damit verbundene Krankheiten verstehen zu können, ist es wichtig, die inter- und intramolekularen Wechselwirkungen zu verstehen. Dies trägt zur Entwicklung von Medikamenten bei, die an spezifische Ziele wie zum Beispiel ein Protein angreifen, um ungewollte Wechselwirkungen zu vermeiden, die zu Nebenwirkungen führen. Die Einzelmolekülkraftspektroskopie ist eine weitverbreitete Methode, inter- und intramolekulare Wechselwirkungen zu untersuchen. Dabei wird die Dissoziation der molekularen Wechselwirkung abhängig von einer externen Kraft beobachtet. Diese induzierte Dissoziation kann als thermisch aktivierter Zerfall über eine Potentialbarriere der Energielandschaft, die die molekulare Wechselwirkung beschreibt, beschrieben werden, aber die Bestimmung und Interpretation der Parameter, die die Energielandschaft beschreiben, ist immer noch eine Herausforderung. Ich entwickle ein Modell, um durch einen Fit an die kraftabhängige Dissoziationsrate die Potentialbreite  $\Delta x_0$ , die Barrierenhöhe  $\Delta G_0$  und die sogenannte "attempt frequency"  $\nu_0$  zu bestimmen. Dafür berücksichtige ich eine verallgemeinerte Abhängigkeit der Energielandschaft von der Wechselwirkungskordinate, charakterisiert durch die Potenz  $n$ . Durch die Untersuchung des Modells mit simulierten Daten zeige ich eine lineare Abhängigkeit von  $\Delta x_0$  von der Fitpotenz auf. Gleichmaßen weisen  $\Delta G_0$  und  $\nu_0$  eine starke Abhängigkeit von der Fitpotenz auf. Dies führt zu einer Unter- oder Überschätzung der Parameter, wenn die falsche Fitpotenz gewählt wird. Die Anwendung der Analysestrategie auf experimentelle Daten in Verbindung mit einer besseren Freely Jointed Chain-Fitroutine deckt eine bemerkenswert verkürzte Lebensdauer des R176A-DNA-Komplexes im Vergleich zur Bindung des Wildtyps an DNA auf. Nach einer weiterführenden Analyse, ist diese Beobachtung als unspezifische Wechselwirkung der DNA mit Gold interpretiert, während die Bindung der Mutante an die DNA komplett unterbrochen zu sein scheint.

Das Rastersondenmikroskop ist eine der wenigen Methoden, die die Untersuchung der Entfaltung eines einzelnen Membranproteins erlaubt. Im zweiten Teil dieser Ar-

beit haben wir diese Methode das erste Mal verwendet, um den konzentrationsabhängigen Einfluss von kompatiblen Soluten auf die Entfaltung von Bacteriorhodopsin zu untersuchen. Die Analyse der einzelnen Entfaltungsschritte erlaubt einen Einblick in die Wechselwirkung der Osmolyte mit den verschiedenen Strukturelementen des Proteins. Die Dissoziationskraft zeigt für jeden Entfaltungsschritt und jede Osmolytkonzentration dasselbe Muster. Wir konnten keine signifikante Entstehung von Zwischenschritten in der Entfaltung beobachten. Das führt uns zu der Annahme, dass die Osmolyte nicht spezifisch an das Membranprotein binden. Außerdem konnten wir eine leichte Zunahme der Dissoziationskraft mit steigender Osmolytkonzentration beobachten, was auf den Stabilisierungseffekt von Osmolyten hinweist. Die Persistenzlänge zeigt eine mögliche Reduzierung mit steigender Osmolytkonzentration. Dies weist auf eine steigende Tendenz des Proteins hin, sich außerhalb der Membran mehr zu winden, was die Rückfaltung des Proteins in die Membran unterstützt. Diese Ergebnisse versprechen ein verbessertes Verständnis von inter- und intramolekularen Wechselwirkungen, das zu einem eventuellen Fortschritt in der Entwicklung von spezifischen Medikamenten beitragen wird.



# Appendix A

## Number of data counts for each osmolyte concentration and unfolding step

Osmolyte	Concentration [mM]	total amount of force curves	selected curves with full length rupture
Buffer	0	18000	180
Betaine	1	17000	44
	10	14000	20
	100	13000	278
	1000	17000	89
Ectoine	1	18000	24
	10	15000	59
	100	17000	31
	1000	19000	31
Taurine	1	16000	8
	10	17000	66
	100	17000	22
	1000	19000	22

Osmolyte	Peak	Concentration [mM]	# counts	Osmolyte	Peak	Concentration [mM]	# counts
Betaine	2nd	0	147	Ectoine	2nd	0	147
		1	22			1	22
		10	13			10	31
		100	224			100	16
		1000	65			1000	24
	3rd	0	180		3rd	0	180
		1	44			1	12
		10	20			10	55
		100	199			100	31
		1000	89			1000	20
	4th	0	149		4th	0	149
		1	40			1	24
		10	18			10	59
		100	278			100	26
		1000	87			1000	31

Osmolyte	Peak	Concentration [mM]	# counts
Taurine	2nd	0	147
		10	42
		100	17
		1000	8
	3rd	0	180
		10	40
		100	16
		1000	22
	4th	0	149
		10	66
		100	22
		1000	17

# Danksagung

*hinfallen aufstehen*

*Krone richten*

*weitergehen*

Fußmatte, Impressionen Katalog, Sommer 2014

Während meiner Zeit im Institut für Physikalische Biologie an der Heinrich Heine Universität habe ich vielfältige Unterstützung beim Aufstehen und Weitergehen und Durchhalten erhalten. Dafür möchte ich mich jetzt hier namentlich bedanken.

Für das interessante Thema, die Betreuung und Unterstützung, sowie die Bereitstellung des finanziellen Rahmens bedanke ich mich bei Dr. Philipp Oesterhelt.

Ein Dank geht auch an Prof. Dr. Dieter Willbold als Zweitgutachter, für seine Unterstützung, für die Möglichkeit, die Arbeit in seinem Institut der Physikalischen Biologie anfertigen zu dürfen, und sein Interesse an der Arbeit.

Für das Korrekturlesen der Arbeit danke ich Dr. Luitgard Nagel-Steger, Martin Wolff und Dr. Martin Jung, die mir die unterschiedlichen Blickwinkel der Biologie, Pharmazie und theoretischen Physik auf meine Arbeit nähergebracht haben.

Für die Bereitstellung der PhoB-Daten danke ich Prof. Dr. Dario Anselmetti und Dr. Volker Walhorn von der Universität Bielefeld.

Ich danke Dr. Arpita Roychoudhury für die Zusammenarbeit und Möglichkeit, an kraftspektroskopischen Daten von Membranproteinen arbeiten zu dürfen.

Ich danke Anna Bronder, Stephanie Schiefer und Mario Schneider, sowie Arpita Roychoudhury für die tolle Zusammenarbeit, eure Unterstützung im Labor und im Allgemeinen, eure Gesprächs- und Diskussionsbereitschaft, sowie eurer Bereitschaft zur

Teilnahme an weiterführenden, teambildenden Maßnahmen. Ihr habt maßgeblich zum Durchhalten beigetragen.

Ein besonderer Dank geht an meine langjährigen Bürokollegen Dr. Oliver Bannach, Dr. Alexander Brener, Shantha Elter, Dr. Lothar Gremer und Elke Reinartz für die nette Atmosphäre, die vielen und spontanen Gespräche, Aufmunterungen und Ratschläge, sowie die sich daraus entwickelnden Freundschaften, die Kochgruppe und die freizeitleichen Nebenprojekte.

Vielen Dank auch an Dr. Luitgard Nagel-Steger und apl. Prof. Dr. Ing. Gerhard Steger für die Unterstützung, ein offenes Ohr und Hilfsbereitschaft.

Ein herzlicher Dank geht auch an die TechnikerInnen Bernd Ester, Elke Reinartz und Ilka Ostermann, sowie Barbara Schulten und unsere Sekretärin Heidi Gruber, ohne deren Arbeit das reibungslose Funktionieren des Instituts nicht möglich wäre.

Ich danke Dr. Cordula Kruse für die wertvollen Ratschläge, die Unterstützung während der gesamten Zeit der Promotion und die Sorge während des Zusammenschreibens.

Dr. Christian Dumpitak danke ich für die Unterstützung, sowie seine Bereitschaft, einen Teamentwicklungsworkshop für uns durchzuführen.

Dank geht auch für die maßgebliche Unterstützung an meine Mentorin Prof. Dr. Janine Splettstößer, Chalmers University, sowie Dr. Bernt-Michael Hellberg, Universität Bielefeld.

Vielen Dank für die Unterstützung und ihr Verständnis an meine Eltern, Geschwister und meinen Neffen, die mich während der Promotion nicht allzu häufig zu Gesicht bekommen haben.

Besonderer Dank geht an meinen Freund Martin Jung, der mich ungefragt, andauernd, ohne sich zu beschweren und bis zum Ende unterstützt hat. Danke für deine Geduld, deine theoretischen Ratschläge und Hilfestellungen, für die technische Unterstützung und jederzeitiges Beantworten aller  $\LaTeX$ -Fragen.

Als letztes möchte ich allen danken, die durch Spaziergänge in und um den Botanischen Garten, sowie dem Sport zu meiner Entspannung beigetragen haben.



# Bibliography

- [Bar06] P. Barth. Hochauflösende Strukturierung von Siliziumoberflächen mittels Mikrokontaktdruck-Technik untersucht mit Rasterkraftmikroskopie. *Dissertation, Universität Ulm*, 2006.
- [BB98] I. Baskakov and D. W. Bolen. Forcing thermodynamically unfolded proteins to fold. *J. Biol. Chem.*, 273(9):4831–4834, 1998.
- [BBX<sup>+</sup>11] F. Berkemeier, M. Bertz, S. Xiao, N. Pinotsis, M. Wilmanns, F. Gräter, and M. Rief. Fast-folding  $\alpha$ -helices as reversible strain absorbers in the muscle protein myomesin. *Proc Natl Acad Sci USA*, 108(34):14139–14144, 2011.
- [Bie11] A. Bieker. Untersuchung der PhoB-Interaktion mit der kleinen Furche der DNA mittels Einzelmolekül-Kraftspektroskopie. *Universität Bielefeld, Masterarbeit*, 2011.
- [BJ95] H. J. Butt and M. Jaschke. Calculation of thermal noise in atomic force microscopy. *Nanotechnology*, 6(1-7), 1995.
- [BKS12] I. Benilova, E. Karran, and B. De Strooper. The toxic  $\alpha\beta$  oligomer and Alzheimer's disease: an emperor in need of clothes. *Nature Neuroscience*, 15(3):349–357, 2012.
- [BMW00] C. Bustamante, J. C. Macosko, and G. J. L. Wuite. Grabbing the cat on the tail: manipulating molecules one by one. *Nature*, 1(130-136), 2000.
- [Bol01] D. W. Bolen. Protein stabilization by naturally occurring osmolytes. *Methods Mol. Biol.*, 168:17–36, 2001.

- [BQG86] G. Binnig, C. F. Quate, and C. Gerber. Atomic Force Microscope. *Phys. Rev. Lett.*, 56:930–933, 1986.
- [BR86] G. Binnig and H. Rohrer. Scanning tunneling microscopy. *IBM Journal of Research and Development*, 4(30), 1986.
- [BS99] J. Bille and W. Schlegel. *Medizinische Physik*. Springer Verlag, 1999.
- [BSGRC02] A. G. Blanco, M. Sola, F. X. Gomis-Rueth, and M. Coll. Tandem dna recognition by phob, a two-component signal transduction transcriptional activator. *Structure (London)*, 10(5):701–713, 2002.
- [BSMS05] P. Bachhawat, G. V. T. Swapna, G. T. Montelione, and A. M. Stock. Mechanism of activation for transcription factor phob suggested by different modes of dimerization in the inactive and active states. *Structure (London)*, 13(19):1353–1363, 2005.
- [Bur00] M. B. Burg. Macromolecular Crowding as a Cell VolumeSensor. *Cell. Physiol. Biochem.*, 10(5-6):251–256, 2000.
- [BZ06] G. Bohm and G. Zech. *Einführung in die Statistik und Messwertanalyse für Physiker*. DESY, 2006.
- [CS05] C.A. Clifford and M.P. Seah. The determination of atomic force microscope cantilever spring constants via dimensional methods for nanomechanical analysis. *Nanotechnology*, 16(1666-1680), 2005.
- [Duf02] Y. F. Dufrene. Atomic force microscopy, a powerful tool in microbiology. *Journal of bacteriology*, 184(19):5205–5213, 2002.
- [EGH06] A. Ebner, H. J. Gruber, and P. Hinterdorfer. Molekulare Kraftmikroskopie und Kraftspektroskopie. *BIOspektrum*, 12(5), 2006.
- [Ell95] V. Ellings. *Digital Instruments Nanoscope Comand Reference Manual*. Santa Barbara CA, 1995.
- [ER97] E. Evans and K. Ritchie. Dynamic Strength of Molecular Adhesion Bonds. *Biophysical Journal*, 72(4), 1997.

- [Eva98] E. Evans. Energy landscapes of biomolecular adhesion and receptor anchoring at interfaces explored with dynamic force spectroscopy. *Faraday Discuss.*, (111):1–16, 1998.
- [Eva01] E. Evans. Probing the relation between force–lifetime–and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.*, 30:105–128, 2001.
- [EW10] P. Eaton and P. West. *Atomic Force Microscopy*. Oxford University Press, 2010.
- [FMG94] E.-L. Florin, V. T. Moy, and H. E. Gaub. Adhesion forces between individual ligand-receptor pairs. *Science*, 264(5157):415–417, 1994.
- [FNY12] R. W. Friddle, A. Noy, and J. J. De Yoreo. Interpreting the widespread nonlinear force spectra of intermolecular bonds. *PNAS*, 109(34):13573–13578, 2012.
- [HA92] D. A. Hammer and S. M. Apte. Simulation of cell rolling and adhesion on surfaces in shear flow: general results and analysis of selectin-mediated neutrophil adhesion. *Biophys. J.*, 63(1):35–57, 1992.
- [Har09] A. Harder. Kraftmikroskopische und kraftspektroskopische Untersuchungen an Proteoglykanen. *Diplomarbeit, Universitaet Bielefeld*, 2009.
- [HB93] J.L. Hutter and J. Bechhoefer. Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.*, 64(1868-1873), 1993.
- [HTB90] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory: fifty years after kramers. *Rev. Mod. Phys.*, 62(2):251–342, 1990.
- [HW10] Y. J. Hsieh and B. L. Wanner. Global regulation by the seven-component pi signaling system. *Y.-J. Hsieh and B. L. Wanner*, 13(2):198–203, 2010.
- [JNOF00] A. Janshoff, M. Neitzert, Y. Oberderfer, and H. Fuchs. Kraftspektroskopie an molekularen Systemen - Einzelmolekuelspektroskopie an Polymeren und Biomolekuelen. *Angewandte Chemie*, 112(3346 - 3374), 2000.

- [KCB<sup>+</sup>04] F. Kühner, L. T. Costa, P. M. Bisch, S. Thalhammer, W. M. Heckl, and H. E. Gaub. LexA-DNA Bond Strength by Single Molecule Force Spectroscopy. *Biophys. J.*, 87(4):2683–2690, 2004.
- [KG06] F. Kühner and H. E. Gaub. Modelling cantilever-based force spectroscopy with polymers. *Polymer*, 47(7):2555–2563, 2006.
- [Kra40] H. A. Kramers. Brownian Motion in a Field of Force and the Diffusion Model of Chemical Reactions. *Physica*, 7(4), 1940.
- [Kur08] M. Kurz. Compatible solute influence on nucleic acids: Many questions but few answers. *Saline Systems*, 6(4), 2008.
- [LKC94] G. U. Lee, D. A. Kidwell, and R. J. Colton. Sensing discrete streptavidin-biotin interactions with atomic force microscopy. *Langmuir*, 10(2):354–357, 1994.
- [LM02] R. Levy and M. Maaloum. Measuring the spring constant of atomic force microscope cantilevers: thermal fluctuations and other methods. *Nanotechnology*, 13(33-37), 2002.
- [LNC98] A. L. Lehninger, D. L. Nelson, and M. M. Cox. *Prinzipien der Biochemie*. Spektrum der Wissenschaft, 1998.
- [LRO<sup>+</sup>99] H. Li, M. Rief, F. Oesterhelt, H. E. Gaub, X. Zhang, and J. Shen. Single-molecule force spectroscopy on polysaccharides by AFM - nanomechanical fingerprint of  $\alpha$ -(1,4)-linked polysaccharides. *Chem. Phys. Lett.*, 305:197–201, 1999.
- [LS06] G. Lentzen and T. Schwarz. Extremolytes: Natural compounds from extremophiles for versatile applications. *Appl. Microbiol. Biotechnol.*, 72(4):623–634, 2006.
- [MAK<sup>+</sup>96] K. Makino, M. Amemura, T. Kawamoto, S. Kimura, H. Shinagawa, A. Nakata, and M. Suzuki. Dna binding of phob and its interaction with rna polymerase. *J Mol Biol*, 259(1):15–26, 1996.
- [MKO<sup>+</sup>02] D. J. Müller, M. Kessler, F. Oesterhelt, C. Möller, D. Oesterhelt, and H. E.

- Gaub. Stability of bacteriorhodopsin  $\alpha$ -helices and loops analyzed by single-molecule force spectroscopy. *Biophys. J.*, 83(6):3578–3588, 2002.
- [MS95] J. F. Marko and E. D. Siggia. Stretching DNA. *Macromolecules*, 28(26):8759–8770, 1995.
- [Oes00] F. Oesterhelt. Kraftinduzierte Sekundärstrukturänderungen in einzelnen Molekülen. *LMU München, Dissertation*, 2000.
- [OJO09] L. Oberbarnscheidt, R. Janissen, and F. Oesterhelt. Direct and model free calculation of force-dependent dissociation rates from force spectroscopic data. *Biophys. J.*, 97(9):L19–L21, 2009.
- [OOP<sup>+</sup>00] F. Oesterhelt, D. Oesterhelt, M. Pfeiffer, A. Engel, H. E. Gaub, and D. J. Müller. Unfolding Pathways of Individual Bacteriorhodopsins. *Science*, 288(5463):143–146, 2000.
- [ORG99] F. Oesterhelt, M. Rief, and H. E. Gaub. Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water. *New J. Phys.*, 1:6.1–6.11, 1999.
- [OSF<sup>+</sup>03] Y. Oberdörfer, S. Schrot, H. Fuchs, E. Galinski, and A. Janshoff. Impact of compatible solutes on the mechanical properties of fibronectin: a single molecule analysis. *PCCP*, 5(9):1876–1881, 2003.
- [RBHO13] A. Roychoudhury, A. Bieker, D. Häussinger, and F. Oesterhelt. Membrane protein stability depends on the concentration of compatible solutes - a single molecule force spectroscopic study. *Biol. Chem.*, 394(11):1465–1474, 2013.
- [RCSG99] M. Rief, H. Clausen-Schaumann, and H. E. Gaub. Sequence-dependent mechanics of single dna molecules. *Nature Structural Biology*, 6(4):346–349, 1999.
- [RGC<sup>+</sup>13] F. Rico, L. Gonzalez, I. Casuso, M. Puig-Vidal, and S. Scheuring. High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science*, 342(6159):741–743, 2013.

- [RGO<sup>+</sup>97] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. Reversible unfolding of individual titin immunoglobulin domains by afm. *Science*, 276(5315):1109–1112, 1997.
- [ROHG97] M. Rief, F. Oesterhelt, B. Heymann, and H. E. Gaub. Single Molecule Force Spectroscopy on Polysaccharides by Atomic Force Microscopy. *Science*, 275(5304):1295–1297, 1997.
- [Roy13] A. Roychoudhury. Compatible solute induced stabilisation of membrane proteins studied by single molecule force spectroscopy. *Dissertation, Heinrich-Heine University Düsseldorf*, 2013.
- [RWK<sup>+</sup>13] M. Ritzefeld, V. Walhorn, Ch. Kleineberg, A. Bieker, K. Kock, Ch. Herrmann, D. Anselmetti, and N. Sewald. Cooperative binding of PhoB(DBD) to its cognate DNA sequence—a combined application of single-molecule and ensemble methods. *Biochem.*, 52(46):8177–8186, 2013.
- [SCB96] S. B. Smith, Y. Cui, and C. Bustamante. Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science*, 271(5250):795–799, 1996.
- [Sch02] J. A. Schellmann. Fifty years of solvent denaturation. *Biophys. Chem.*, 96(2-3):91–101, 2002.
- [SDH01] R. W. Stark, T. Drobek, and W. M. Heckl. Thermomechanical noise of a free v-shaped cantilever for atomic-force microscopy. *Ultramicroscopy*, 86(207-215), 2001.
- [SKW<sup>+</sup>08] M. J. Serpe, F. R. Kersey, J. R. Whitehead, S. M. Wilson, R. L. Clark, and S. L. Craig. A simple and practical spreadsheet-based method to extract single-molecule dissociation kinetics from variable loading-rate force spectroscopy data. *J. Phys. Chem. C*, 112(49):19163–19167, 2008.
- [SLMW95] J.E. Sader, I. Larson, P. Mulvaney, and L.R. White. Method for the calibration of atomic force microscope cantilevers. *Rev. Sci. Instrum.*, 66(3789-3798), 1995.

- [SOSG99] T. Strunz, K. Orozlan, R. Schäfer, and H. J. Güntherodt. Dynamic force spectroscopy of single DNA molecules. *Proc. Natl. Acad. Sci. USA*, 96(20):11277–11282, 1999.
- [Toe04] A. Toepel. *Chemie und Physik der Milch: Naturstoff- Rohstoff- Lebensmittel*. Behr's Verlag, 2004.
- [VK92] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Amsterdam, 1992.
- [VLL09] P. Venaketsu, M. J. Lee, and H. M. Lin. Osmolyte counteracts urea-induced denaturation of alpha-chymotrypsin. *J. Phys. Chem. B*, 113(15):5327–5338, 2009.
- [VNA05] D. V. Vezenov, A. Noy, and P. Ashby. Chemical force microscopy: probing chemical origin of interfacial forces and adhesion. *Journal of Adhesion Science and Technology*, 19:189–240, 2005.
- [Wan93] B. L. Wanner. Gene regulation by phosphate in enteric bacteria. *J. Cell. Biochem.*, 51(1):47–54, 1993.
- [Wat11] J. D. Watson. *Molekularbiologie*. Pearson Studium, 2011.
- [Wol09] K. Wollschlaeger. Modulation der Bindung des Transkriptionsfaktors PhoB aus *Escherichia coli* an DNA. *Dissertation, Universitaet Bielefeld*, 2009.
- [XT97a] G. F. Xie and S. N. Timasheff. Mechanism of the stabilization of ribonuclease a by sorbitol: preferential hydration is greater for the denatured than for the native protein. *Protein Sci.*, 6(1):211–221, 1997.
- [XT97b] G. F. Xie and S. N. Timasheff. The thermodynamic mechanism of protein stabilization by trehalose. *Biophys. Chem.*, 64(1-3):25–43, 1997.
- [Yam97] H. Yamakawa. *Helical Wormlike Chains in Polymer Solutions*. Springer Verlag, 1997.
- [YJN07] I. Yu, Y. Jindo, and M. Nagaoka. Microscopic understanding of preferen-

tial exclusion of compatible solute ectoine: direct interaction and hydration alteration. *J. Phys. Chem. B*, 111(34):10231–10238, 2007.

- [YOI<sup>+</sup>08] T. Yamane, H. Okamura, M. Ikeguchi, Y. Nishimura, and A. Kidera. Water-mediated interactions between DNA and PhoB DNA-binding/transactivation domain: NMR-restrained molecular dynamics in explicit water environment. *Proteins: Structure, Function, and Bioinformatics*, 71(4):1970–1983, 2008.