Single Molecule Force Spectroscopic Investigation of the Siloxane Bond in Polydimethylsiloxane and Atomic Force Microscopy Imaging of Nucleosomal Arrays

Dissertation zur Erlangung des Doktorgrades Dr. rer. nat.
der Fakultät für Naturwissenschaften der Universität Ulm

vorgelegt von
Michael Roland Budde

aus
München

2015

Amtierender Dekan:
Prof. Dr. Joachim Ankerhold

Betreuer und Erstgutachter:
Prof. Dr. Jens Michaelis

Zweitgutachter:
Prof. Dr. Christoph Bräuchle
(Ludwig-Maximilians-Universität München)

Tag der Promotion:
20. Mai 2015

Abstract

Part I: Polydimethylsiloxane (PDMS) is a material that is widely used in research, medicine and industry. The goal of this thesis was to support the development and improvement of this material by researching a method that is able to quantitatively characterize the siloxane bond, which is the backbone of PDMS polymers, and it's rupture behavior. Single molecule force spectroscopic constant force measurements of covalently attached PDMS polymers were performed. Two different rupture populations were observed in the experiment. And it was found that the rupture behavior was independent of the polymer length, which means that only ruptures at or near the attachment points were measured. The obtained rupture data was then analyzed with Bayesian parameter estimation using the nested sampling algorithm. Three different bond rupture models were used to describe the rupture behavior of the siloxane bond: the Bell model, the CUSP model and the linear-cubic model. It was possible to estimate bond potential parameters with good quality for the Bell model. For the CUSP model and the linear-cubic model it was only possible to determine one of three parameters with sufficient quality. The reason for the insufficient parameter estimation quality was found to be a too small range of forces that was accessed during the constant force experiments. Constant speed data from a previous work was also analyzed with the same methods. None of the three bond rupture models described the constant speed data well. With the help of simulations it could be shown that the rupture populations from the constant force measurements and the rupture population from the constant speed measurements are different from each other and that there exist at least three different rupture populations in the same sample.

Part II: The 30 nm fiber is the presumed second level of condensation in chromatin after the beads-on-a-string conformation. Nucleosomal arrays are a good basis to systematically research the 30 nm fiber as they are a defined and controllable system. In this thesis nucleosomal arrays were imaged using the atomic force microscope. The nucleosomal arrays were imaged using contact mode in air and tapping mode in liquid. The preparation of samples, concentrations of nucleosomal arrays, salts and buffers were primarily optimized for good imaging results. The influence of ions on the conformation of the nucleosomal arrays was also examined.

Contents

Contents **1**

I Siloxane Bond Single Molecule Force Spectroscopy **5**

1 Introduction **7**

2 Theory and experimental fundamentals **11**

2.1 Atomic Force Microscope 11

2.2 Single-Molecule Force Spectroscopy 13

2.2.1 Constant Force Measurements 13

2.2.2 Constant Speed Measurements 17

2.2.3 Cantilever Calibration 18

2.3 Polydimethylsiloxane for single molecule force spectroscopy . 21

2.3.1 Properties 22

2.3.2 Sample Preparation 23

2.4 Bayesian Data Analysis 25

2.5 Nested Sampling 30

2.5.1 Sorted Likelihood Function 32

2.5.2 Algorithm 33

2.5.3 Creating a Representation of the Posterior 37

2.6 Bond Rupture Models 39

2.6.1 BELL model 40

2.6.2 CUSP model 40

2.6.3 Linear-Cubic Model 45

2.6.4 Potentials used in this work 49

2.6.5 Behavior of the CUSP potential with high barrier approximation and the linear-cubic potential without high barrier approximation . . . 50
2.6.6 Behavior of various covalent and non-covalent bond rupture potentials . . . 51
2.7 Parameter Estimation . . . 52
2.7.1 Parameter Estimation for Constant Force Bond Rupture Experiments . . . 52
2.7.2 Extension to Two Rupture Populations . . . 54
2.7.3 Parameter Estimation for Constant Speed Bond Rupture Experiments . . . 55
2.7.4 Parameter Estimation Result Visualization . . . 55
2.8 Data Simulation . . . 56
2.8.1 Constant Force Data . . . 58
2.8.2 Constant Speed Data . . . 59
2.9 Experimental Methods and Materials for PDMS Single Molecule Force Spectroscopy . . . 59
2.9.1 Atomic Force Microscope . . . 59
2.9.2 Commercial Software . . . 60
2.9.3 Cantilevers . . . 60
2.9.4 Experimental Procedures and Materials . . . 60

3 Force Spectroscopy Results 65
3.1 Constant Force Measurements . . . 65
3.1.1 Rupture Rate Length Independence . . . 65
3.1.2 Multiple Rupture Populations . . . 66
3.1.3 Bell Model Results . . . 70
3.1.4 CUSP Model Results . . . 72
3.1.5 Linear-Cubic Model Results . . . 73
3.1.6 Model Validity . . . 76
3.2 Factors influencing the parameter estimation quality . . . 79
3.2.1 Simulation and Parameter Estimation . . . 81
3.2.2 Influence of the Force Range . . . 83
3.2.3 Influence of the Amount of Data . . . 86
3.2.4 Influence of the Spring Constant Determination Error 89
3.2.5 Concurrent Parameter Estimation . . . 95

3.2.6 Summary of the Influencing Factors 103
3.2.7 Optimizing Constant Force Measurements 106
3.3 Constant Speed Measurements 108
3.3.1 BELL Model Results 108
3.3.2 CUSP Model Results 109
3.3.3 Linear-Cubic Model Results 110
3.3.4 Model Validity 111
3.4 Discussion and Comparison of the Parameter Estimation Results 113
3.4.1 Comparison of the Parameter Estimation Results with the Experimental Data 114
3.4.2 Multiple rupture processes in the same sample 117
3.4.3 Comparison of the Parameter Estimation Results with Theoretical Findings 122
3.4.4 Comparison of the Parameter Estimation Results with Silylester Force Spectroscopy Experiments 123

4 Summary and Outlook **127**

II Imaging of Nucleosomal Arrays **133**

5 Introduction **135**

6 Theory and Methods **139**
6.1 AFM Imaging 139
6.1.1 Imaging Modes 140
6.1.2 Imaging Media 141
6.2 DNA, Nucleosomes and Chromatin 141
6.2.1 DNA and nucleosomal array Properties 143
6.2.2 Attachment of the DNA and the nucleosomal arrays . 145
6.3 Methods and Materials for nucleosomal array Imaging 146
6.3.1 Atomic force microscope 146
6.3.2 Commercial software 146
6.3.3 Cantilevers 146
6.3.4 Experimental Procedures and Materials 146
6.3.5 Sample preparation 150

7 DNA and nucleosomal array Imaging Results **153**
7.1 Contact Mode Test Measurements of DNA and nucleosomal arrays in Air . . . 153
7.1.1 DNA . . . 154
7.1.2 Nucleosomal Arrays . . . 158
7.2 Tapping Mode Measurements of DNA and nucleosomal arrays in Liquid . . . 160
7.2.1 DNA . . . 160
7.2.2 Nucleosomal Arrays . . . 162
7.3 Conclusions . . . 169

8 Summary and Outlook **171**

III Appendix **175**

9 Software for Single Molecule Force Spectroscopy **177**
9.1 Experimental . . . 177
9.2 Simulation . . . 178
9.3 logL Functions . . . 182

List of Figures **187**

Bibliography **191**

Acknowledgement **203**

Declaration of Cooperations **205**

Curriculum Vitae **207**

Part I

Siloxane Bond Single Molecule Force Spectroscopy

Chapter 1

Introduction

The uses for silicone elastomers are manifold,[1] including molds and stamps for fabrication at the micro- and nanoscale,[2] microfluidics[3] or flexible electronics.[4] But they have some key disadvantages, which prevent an even broader adoption. Their low mechanical strength[5] and susceptibility to hydrolysis[6] are two of those disadvantages and closely connected to the strength and the rupture processes of the Si-O bond.

Single-molecule force spectroscopy has been used in the study of various polymers and also chemical bonds. The force-extension behavior of polymers such as DNA,[7] proteins[8] or polysaccharides[9] or the interaction of polymers with themselves[10] and with surfaces[11] has been used to identify the mechanical properties of these polymers.[12,13] The behavior of polymers in single-molecule force spectroscopy experiments can be explained by statistical physics models including the freely jointed chain (FJC),[14] the wormlike chain (WLC)[15] or the freely rotating chain[16] as well as the extendible expansions of the FJC and WLC for higher forces, which include an additional enthalpic term.[17] The behavior of polymers at high forces can also be simulated with ab initio molecular dynamics methods.[18,19] The properties of bonds have been examined in rupture experiments such as the unfolding of protein domains,[9] overstretching of DNA[20] and separation of receptor-ligand connections.[21] Using an atomic force microscope even covalent bonds in a polymer covalently attached to the cantilever and the substrate could be ruptured and the rupture forces for certain loading rates could be determined.[22] Single molecule bond rupture experiments subject the bonds to either varying loading rates, which influence the observed rupture forces, or

varying holding forces, which influence the observed rupture rates.[23,24]

The rupture behavior under these varying conditions can be described with phenomenological or microscopic models. Variation of the loading rate[25] or analysis of the probability density function of mean rupture forces[26–28] can yield information about the bond, such as the shape of its energy landscape or the critical force at which the energy barrier of the bond vanishes. Evans et al.[29] developed a first phenomenological approach to model the rupture behavior of bonds by assuming, that the rupture rate increases exponentially with the applied force and that properties of the bond are determined by two parameters: the rupture rate at zero force k_0 and the distance between the minimum and transition state, $x^{\ddagger}$.[24,30] This model is a good approximation only for low forces, since in experimental data considerable deviations from this exponential relation are observed at high forces. Microscopic models which reproduce this more complex behavior have been developed using the theory of diffusive barrier crossing under force by Kramers[22,23,31–33] combined with various energy potentials.[27,28,34,35]

Previous theoretical works[19,36,37] based on ab-initio density functional theory (DFT) have examined the molecular and mechanical properties of siloxane bonds in polydimethylsiloxane (PDMS) polymers in vacuum and in hexamethyldisiloxan (HMDS) solvent. They also found that water molecules can lower rupture forces via mechanically induced chemical reactions. The mechanical properties of PDMS polymers in HMDS and desorption behavior of PDMS polymers from surfaces were also the focus of previous practical experiments.[38–40] The extendible freely jointed chain model appeared to best describe the PDMS polymer chain and was used to analyze their extension with an atomic force microscope. The polymers were attached on both ends via Si-O bonds so that the whole system formed a chain of only Si-O bonds, which were assumed to be identical. Constant speed experiments were performed and least-squares fits were used to estimate the parameters of the Si-O bonds in the chain.[38,41]

In this work data from constant force measurements of covalently attached PDMS polymers is analyzed. Three different models, both phenomenological and microscopic, are used to interpret the data. Simulations are used to help compare the outcome for all the different models and measurement modes and ultimately to better understand the rupture process in the sample, which is more complex than anticipated. The properties of the

two microscopic models and especially the implications of those properties for the successful bond parameter estimation are also explored. Consideration is given to experimental aspects such as necessary amounts of data and holding force ranges, the erroneous calibration of the spring constant of the AFM cantilever, which is used to rupture the bond, and the analysis of two rupture populations at once. Furthermore the data from previous constant speed experiments[38,41] is reevaluated with the new analysis methods and the outcome is compared with the results from the new constant force measurements.

This work consists of two parts. This is part I, which is about the characterisation of the siloxane bond in poyldimethylsiloxane using single-molecule force spectroscopy. The following part is about imaging of nucleosomal arrays. After the introduction in this chapter (chapter 1) fundamental information about single-molecule force spectroscopy, analysis of force spectroscopy data and polydimethylsiloxan is given in chapter 2. Chapter 3 contains the analysis of polydimethylsiloxane force spectroscopy measurement data, simulations which were performed to better understand the factors that influence parameter estimation quality and finally a comparison of current results with results derived from previous experiments. Chapter 4 is the final chapter of part I and summarizes the previous chapters. It also gives a short outlook concerning possible future directions for this line of research.

Chapter 2 is subdivided into several sections. Section 2.1 describes the basic principle of the atomic force microscope. Then section 2.2 covers single-molecule force spectroscopy, the two main single-molecule force spectroscopy measurment modes as well as cantilever calibration, which is necessary for any force measurements in atomic force microscopes. The next section is section 2.3, which is about the polydimethylsiloxane samples. Section 2.4 is about Baeysian data analysis. Nested sampling is a Baeysian parameter estimation method used for data analysis in this work and is explained in section 2.5. Section 2.6 contains a description of three different bond rupture models. How Baeysian data analysis is applied to constant force and constant speed data can be found in section 2.7. Besides experimental data, simulated data is also used in this work. The simulation process for both constant force and constant speed data as well as input bond potentials are described in section 2.8. In the final section 2.9 experimental

methods and materials are described in more detail.

In chapter 3 the following sections can be found: In section 3.1 the parameter estimation results of the constant force data analysis are discussed as well as the observations that the rupture behavior of polydimethylsiloxane is not length dependent and that there are two different rupture populations. This section is followed by section 3.2 which is about various factors influencing the quality of the parameter estimation for constant force data. First the behavior of the parameter estimation under ideal conditions is examined (3.2.1). Then the influence of the range of holding forces (3.2.2), the amount of observed bond ruptures events (3.2.3), erroneous cantilever calibrations (3.2.4) as well as the concurrent estimation of two different bond potentials (3.2.5) are discussed. After that all factors are compared to each other (3.2.6) and from these comparisons conclusions are drawn for the optimization of constant force experiments (3.2.7). In section 3.3 data from previous constant speed experiments are analyzed with Bayesian data analysis. Finally in section 3.4 the results from the constant force and constant speed experiments are compared to each other and to results from other groups.

Chapter 2

Theory and experimental fundamentals

2.1 Atomic Force Microscope

Atomic force microscopy (AFM) is a form of scanning probe microscopy[42] developed by Binnig et al. in 1986.[43] In scanning probe microscopy a probe is systematically moved across a sample and the interactions between the probe and sample are recorded. Since the probe is only interacting with a small part of the sample it has to be moved across the sample surface in a scanning motion. This rasterization process by it's nature is drastically slower than imaging methods which capture the entire sample at once. On the other hand the scanning nature of this process allows for direct quantitative measurements of the spatial dimensions of the sample. Usually the probe is a very fine tip and depending on the microscopy method it can consist of various materials such as silicon, carbon nanotubes or metal coated glass fibers to allow for a multitude of interactions between the sample and the surface. The possible interactions between sample and tip are manifold: magnetic, electrostatic and contact interactions are only a small subset of these interactions. Tunneling of electrons for example is used in scanning tunneling microscopy (STM)[44,45] and in near-field scanning optical microscopy (NSOM)[46] the tip facilitates the detection of absorption, flourescence and polarization of light in the near-field with a high spatial resolution not limited by the wavelength of the light. The majority of interactions in scanning probe microscopy are of an extremely short range.

Combined with fine tips in the range of a few atoms or even single atoms it is thus possible to reach atomic resolutions with scanning probe microscopy. Finally the interactions not only allow the imaging of samples, but also the manipulation of a sample.

In atomic force microscopy the probe is located at the end of a cantilever. Forces interacting with the tip result in a mechanical deformation of the cantilever. The high spatial resolution of the atomic force microscope is possible because a light beam deflected from the backside of the cantilever magnifies this deformation reliably (Figure 2.1).[47] In addition to a sensitive detection of the cantilever deformation, precise movements of the probe in the picometer range are necessary to achieve atomic resolution. Usually piezoelectric elements can provide the necessary movement precision at reasonable speeds. While the main interaction in atomic force imaging is the contact between the sample and the cantilever tip, other forces such as van-der-Waals interactions or capillary forces between the cantilever tip and a surface water film have to be taken into account. Advantages of atomic force microscopy are that the measurements can be performed in various media such as liquids, gases and vacuum. Even large temperature ranges are accessible to atomic force microscopy. For the highest resolution measurements vaccuum and low temperatures are needed though. Another advantage of atomic force measurements is that biological samples from biomolecular complexes to cells and tissues can be imaged in a liquid environment. And with the exception of a sometimes necessary fixation on the substrate the samples mostly need no special prior preparation. This makes the imaging of active biological samples and their dynamic behavior possible. Depending on the properties of the tip the atomic force microscope can manipulate samples via mechanical deformation, localized heating, electric current and magnetic forces.

The cantilever, to which the probe is attached, is a very sensitive spring, which opens up another category of measurements if the cantilever is used to exert precisely measurable forces on attached molecules: single-molecule force spectroscopy. Knowledge of the deformation of the cantilever and the spring constant of the cantilever can be used to calculate the forces acting upon it. The following sections will focus on the properties and requirements of single-molecule force spectroscopy and how force spectroscopy can be used to determine the strength of bonds polymers.

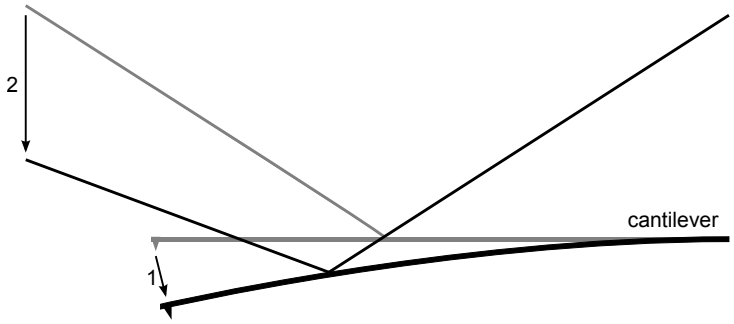


Figure 2.1: Schematic of the central atomic force microscope principle. A force is deforming the cantilever and a lightbeam is deflected from the back of the cantilever. The deformation of the cantilever (1) is amplified by the reflected light beam, which results in a comparatively large change in the deflection (2).

2.2 Single-Molecule Force Spectroscopy

Force spectroscopy is based on the observation of forces acting on a sample and the reaction of the sample to the forces. Single-molecule force spectroscopy specializes on the interaction of forces with a single molecule at a time and allows to examine very small systems such as single atomic bonds. Single-molecule force spectroscopy has been used to examine biological samples such as DNA[7] or proteins,[8] polysaccharides,[9] polymers interacting with themselves[10] or polymers interacting with surfaces.[11] Two basic methods of single-molecule force spectroscopy are constant force and constant speed measurements. Constant force measurements involve pulling a bond or polymer at a constant force and then observing the time until rupture occurs. In constant speed measurements the sample is stretched at a constant speed and then the force at the moment when rupture occurs is recorded. While these methods are conceptually simple, several complications have to be taken into account such as: attachment of the sample, the strength of all involved bonds, possible necessity of linkers or the precision of the measured forces.

2.2.1 Constant Force Measurements

Constant force measurements involve the stretching of a bond to a preset holding force and keeping that force constant until bond rupture occurs.

Consequently the data recorded in constant force experiments are holding forces and times until bond rupture. Higher holding forces lead to shorter times to rupture or in other words higher rupture rates. Such a force dependent rupture behavior can be described with bond rupture models. The practical execution of constant force experiments, however, makes other considerations necessary: The experiment is divided into an extension and a constant force holding phase. This requires a feedback mechanism in the instrument, which is more complicated than setting a cantilever movement speed in a constant speed experiment. The parameters for this feedback mechanism have to be chosen carefully to allow for a stretching phase, which is short in comparison to the expected rupture times. At the same time the feedback has to take into account the contour length distribution of the examined system, which can vary considerably, and has to be gentle enough to prevent premature rupture in the stretching phase before the holding force setpoint is reached. Furthermore the feedback mechanism has to be both sensitive and responsive enough to keep the holding force constant.

The rupture times recorded at a single holding force are distributed exponentially and thus only give information in the form of a rupture rate. But a single rupture rate is not sufficient information to determine the parameters of bond rupture models with more than one parameter. Thus it is necessary to measure at more than one holding force. On the other hand the acquisition of data is relatively straightforward, since both holding force and time to rupture can be read off directly from the acquired force over time information. Another advantage of constant force measurements is the long time for which the polymers can be held. This can potentially be used to investigate the influence of rare environmental interactions such as reaction with water molecules.

In practice the force spectroscopy measurements and the preceding calibration of the cantilever were performed in a closed chamber with an inlet and outlet for solvent exchange. The high volatility of the HMDS solvent as well as the sensitivity of the Si-O bond to water molecules make a closed chamber setup necessary. In an open setup, which uses HMDS drops, measurements are restricted to a time range of tens of minutes. On the other hand measurements of 24 hours and longer are possible in the closed chamber, as long as solvent is refilled every 3 to 4 hours. At the beginning of each single pull the cantilever is pressed on the surface for 500 ms with a

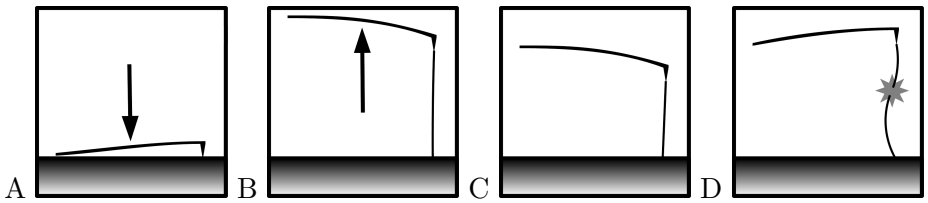


Figure 2.2: Single-molecule force spectroscopy measurement using constant force. (A) The cantilever tip is pressed on the surface to allow a polymer end to react and form a covalent attachment to the cantilever tip. (B) The attached polymer is quickly stretched until a set holding force is reached. (C) The polymer is held at the target force. (D) Bond rupture occurs. Not depicted: pull to 1.5 nN for holding forces below 1.5 nN.

nominal force of 5 nN. The constant force measurements (Figure 2.2) were performed using a feedback mechanism, which stretches polymers to a target force in a time shorter than 100 ms. The target forces were set in a range between ca. 1-2 nN. For forces below 1.5 nN the polymers were first stretched to a pre-pull force of 1.5 nN and then immediately relaxed to the target force. This step was necessary to reduce the number of non-covalently attached polymers, making such events negligible.[38,41] For each measurement the force as a function of time and the cantilever position as a function of time are recorded. The nominal force is kept constant using a feedback mechanism. After bond rupture has occurred and the cantilever has reached a distance of 2000 nm to the surface, which is twice the maximum length of the stretched polymers, the measurement data is saved and a new measurement is initiated by pressing the cantilever on the sample once again (Figures 2.3 and 2.4). Only events where the target force is reached and held for at least 100 ms were evaluated. If no polymer is attached, the target force or pre-pull force is never reached and the measurement is aborted after the distance between cantilever and surface exceeds 2000 nm. Events where the target force is reached, but held for a time below 100 ms, are recorded and then excluded during post processing of the data. Every 100 surface contacts the cantilever is automatically moved to a different random location on the sample surface to ensure a constant high availability of reactive polymer ends. The probability for observing covalent attachment of a polymer with a holding time longer than 100 ms was below ca. 1%, thus effectively minimizing the probability for observing multimolecular events.

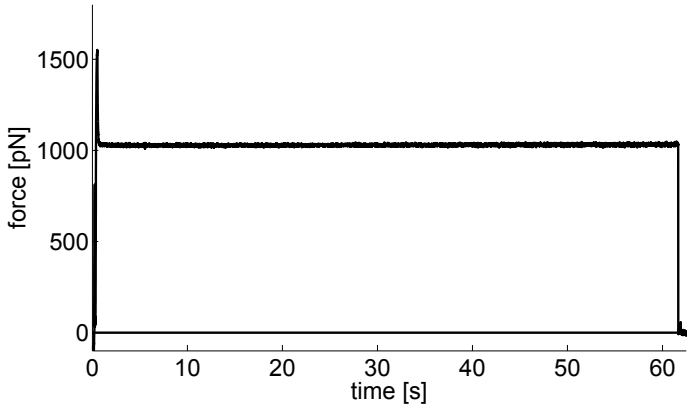


Figure 2.3: Raw force over time data of a polymer being held at a constant force until rupture occurs after roughly 60 seconds. The data bandwidth is 2 kHz. Because the set force is below 1.5 nN, the polymer is stretched to a force of 1.5 nN at the beginning.

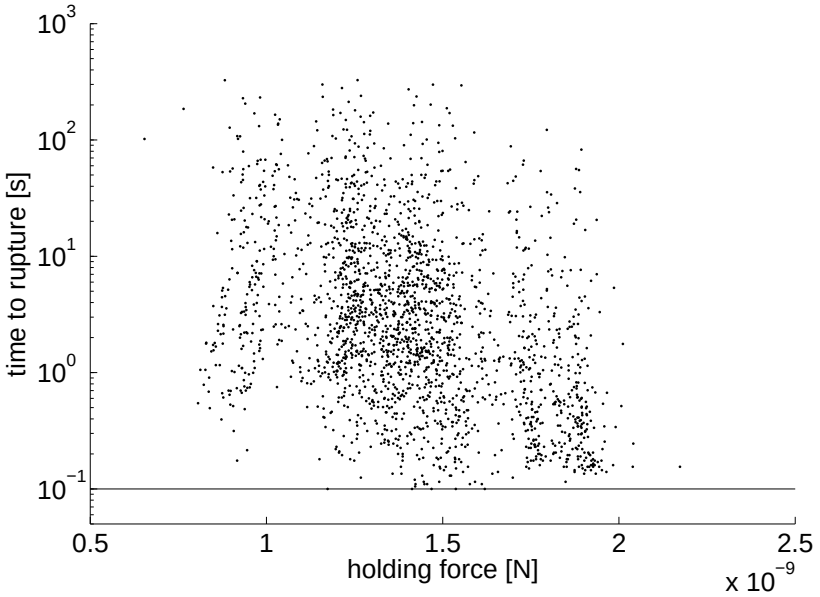


Figure 2.4: Constant force measurement data showing rupture times and corresponding holding forces. Each single molecule event is displayed by a dot. Holding forces below 100 ms duration have to be cut off, because those times become similar in length to the stretching process (horizontal line).

2.2.2 Constant Speed Measurements

The second measurement mode of single-molecule force spectroscopy are constant speed measurements. They involve pulling the polymers at a constant speed. Even though the method is called constant speed, the measure of importance is the increase of force over time acting on the pulled bond or bonds, which is called the loading rate. The measured data include force over time and force over distance data. From this data the rupture force can be determined. And with the usage of polymer models the contourlength of each polymer can be estimated. When the polymer length is given, the number of stretched Si-O bonds and finally the loading rate at bond rupture can be calculated. The variation of the pulling speed can be useful, but oftentimes when investigating polymers, the length distribution already provides a sufficiently large spread of loading rates.[38,41]

In comparison to constant force measurements pulling at a predefined speed is easier to accomplish experimentally than keeping a holding force constant via a feedback mechanism. The drawback is that the loading rate cannot be measured directly and has to be determined using polymer models and thus the analysis of the data is more complicated than the analysis of constant force data.

Furthermore it has to be taken into account that analysis of constant speed data assumes a constant loading rate at the moment of rupture. A constant loading rate corresponds to a linear increase of the force over time as well as distance, since the cantilever is moving away from the surface at a constant speed. Two processes contribute to the stretching of polymers: an enthropic process, which involves the reduction of conformational entropy of the polymer as it is stretched from a relaxed, randomly coiled form into a linear form, and an enthalpic process, where bond angles and lengths are deformed. At the beginning of the stretching process it is mainly enthropic, while the enthalpic contribution increases with the force. For the enthalpic process the polymer and the pulling cantilever can be treated as Hookian springs and thus this process contributes linearly to the observed force-distance curve. The entropic process on the other hand makes the usage of polymer models for necessary. The contribution of this process to the force-distance curve is not linear. Previous constant speed experiments on PDMS polymers have shown that even at high forces the entropic contribution leads to considerable deviations from a linear increase of the force over distance



Figure 2.5: Illustration of a typical force-distance curve (black line) measured when a polymer is stretched in a constant speed experiment. The light gray area below the curve corresponds to the enthropic contribution which is dominant at the beginning of the stretching process. The dark grey area is the contribution of the enthalpic process, which is linear in nature. Even at high forces there is a enthropic contribution.

(Figure 2.5). Only beyond a force of 3.8 nN the force-distance curve can be seen as approximately linear.[38,41]

2.2.3 Cantilever Calibration

Both the constant force (subsection 2.2.1) and the constant speed measurement mode (subsection 2.2.2) rely on precise force measurements. To achieve such precise force measurements with the atomic force microscope both the deformation of the cantilever and its spring constant have to be known accurately. The deformation of the cantilever can be measured with relative ease, but the determination of the spring constant is currently one of the more challenging parts of the force measurement process. The spring constant of a cantilever is strongly dependent on its dimensions with the thickness being especially influential. Even modern AFM probe manufacturing processes have difficulties in creating cantilevers of consistent dimensions. According to manufacturer specifications (e.g. Bruker), the spring constants of can-

tilevers from batch to batch can vary considerably, a factor of 4 between the maximum and minimum spring constant can be commonly found and even a factor of 8 is possible. With this much variation a reliable calibration of the cantilever spring constant is indispensable. The following will give an overview of cantilever calibration methods and then explains the thermal tune calibration method which was used in this work to calibrate all cantilevers.

Several methods to calibrate the AFM cantilever spring constant have been developed. The cantilever spring constant k can be predicted by measuring the cantilever dimensions including the thickness precisely and applying theoretical models.[48–50] Determining the thickness of the cantilever[51] as well as additional material layers such as reflective coatings, which have to be included in the calculations,[52] complicate the determination of k for this method. It is also possible to calculate k for rectangular cantilevers using measurements of just the length and width of the cantilever, the resonance frequency, the Q-factor and the viscosity of the surrounding medium.[49] It is not possible to use this method for non-rectangular cantilevers though. The spring constant of the cantilever can also be obtained by observing its deformation under a known force. Known forces can be applied by pressing the cantilever against other cantilevers with known spring constants[53] or by attaching small known masses to the cantilever.[54] Both methods rely heavily on precise positioning of the cantilevers to each other or of the weights on the cantilever. Also the mass of the attached weigths has to be known precisely. Weights are also used in a second method, which observes the change of the resonance frequency with repeated addition of weights.[55] For this method it is also necessary to precisely determine the positions and masses of the attached weights.

The method used in this work is called thermal tune method and is characterized by ease of use, speed and low amount of possible user error. For this method the cantilever is approximated by a harmonic oscillator, which is vibrating under the influence of thermal fluctuations in the environment under equilibrium conditions.[56] The hamiltonian H for a harmonic oscillator is the following

$$H = \frac{p^2}{2m} + \frac{1}{2} m \omega_0^2 z^2 \tag{2.1}$$

with m being the mass of the oscillator, ω_0 being the resonant angular frequency, p being the impulse operator and z being the position operator, which in this case corresponds to the displacement of the oscillator. According to the equipartition theorem the kinetic energy of each of the two quadratic terms is $k_BT/2$. The displacement term of the Hamiltonian is thus[56]

$$\left\langle \frac{1}{2} m\omega_0^2 z^2 \right\rangle = \frac{k_BT}{2} \tag{2.2}$$

With the relation of $\omega_0^2 = k/m$ the spring constant k can then be calculated from the mean-scquare displacement of the cantilever by[56]

$$k = \frac{k_BT}{\langle z^2 \rangle} \tag{2.3}$$

Under the assumption that other sources of noise, such as electronic noise, are negligible, the peak P in the power spectrum created by the thermal oscillations of the cantilever corresponds to the mean-square displacement of the cantilever $\langle z^2 \rangle$ and thus:[56]

$$k = \frac{k_BT}{P} \tag{2.4}$$

In practice the power spectrum of changes in the deflection, which are caused by the oscillations of the cantilever, has to be first recorded at a distance large enough to prevent interactions with the sample or substrate surface.

Since the deformation of the cantilever is measured as changes in the deflection of the light beam, the deflection of the light beam has to be set into relation with the deformation of the cantilever. This is achieved by pressing the cantilever into a hard substrate for a set distance prior to the recording of the power spectrum. From this the inverse optical lever sensitivity can be found via a fit to the measured change in deflection over distance. The oscillation modes of the cantilever have the shape of a Lorentzian under the conditions of a low viscosity environment such as air, high Q-factor and low amplitudes.[57] By substracting a term for instrumentation noise and then fitting a Lorentzian to the peak in the power spectrum, P, and applying equation 2.4 the spring constant k of the cantilever can be calculated.

The accuracy of this method overall has been evaluated in the past and

it has been estimated that the accuracy of this method is in the range of 10-20%.[49,58] An error of this magnitude for the spring constant might seem considerable, but is still a marked improvement over simply using the specifications of the manufacturers. Nonetheless the error is large enough to warrant close consideration of its influence. In this work the influene of this error has been examined and the results can be found in section 3.2.4.

The thermal tune method described above was used as implemented in the MFP3D AFM software version 0408. Even though the integration of the method in the software as well as it's general properties make it especially easy and fast to use with low probability for user error, it has to be noted that there are more sophisticated approaches to the thermal tune method, which correct several complications not addressed by the used method and which will reduce the aforementioned error of 10-20%:

- There are differences between the fundamental oscillatory mode of a cantilever and the oscillatory mode of an ideal spring.
- The measured displacement of the cantilever is also dependent not only on the displacement of the cantilever, but also on its angular deformation, both of which occur simultaneously. This can be corrected by application of simple beam theory, if the cantilever has a rectangular shape. For V-shaped cantilevers deviations of 25% have to be expected.[59]
- Furthermore there is an assumption that there is no overlap between the fundamental vibrational mode of the cantilever and its higher order modes. In air this assumption holds, but not necessarily for environments with higher viscosity. In such environments the shapes of the peaks of the thermal cantilever oscillations also begin to deviate from the Lorentzian which also leads to an increase of the error. Corrections for cantilever calibration in environments with high viscosity have also been discussed.[60]

2.3 Polydimethylsiloxane for single molecule force spectroscopy

As mentioned in the introduction (see chapter 1) silicone elastomers are a widely and commonly used group of materials.[1] They are used fabrication

at the micro- and nanoscale,[2] microfluidics[3] or flexible electronics[4] amongst many other applications. Their two key disadvantages of a low mechanical strength[5] and susceptibility to hydrolysis[6] are likely to be closely connected to the strength and the rupture processes of the Si-O bond, which is the basic building block of silicones. Thus, to understand the properties the properties and behavior of silicones on a fundamental level, it is necessary to closely examine the Si-O bond or siloxane bond. Single-molecule force spectroscopy can be used to characterize atomic bonds like the siloxane bond, but this method obviously requires a single molecule and not a solid state bulk sample consisting of a multitude of polymers. On the other hand the siloxane bond in the solid state is of interest. To approximate such a such solid state environment, where only a single molecule can be measured, a chain of siloxane bonds, which are also attached with siloxane bonds, in a chemically closely related solvent was used. The system was also chosen, because in future experiments the addition of water to the solvent would allow to examine the aforementioned susceptibility of silicones to hydrolysis. Other processes influencing the stability of silicones could also be examined by adding a variety of chemicals to the solvent. To achieve such a system polydimethylsiloxane (PDMS) polymers were used to provide the chain of siloxane bonds and hexamethyldisiloxane (HMDS) was used as the chemically related solvent.

2.3.1 Properties

Polydimethylsiloxane polymers consist of a backbone of Si-O bonds, with each silicone also having two methyl groups (Figure 2.6).[38,41] Each polymer has two different reactive ends to allow covalent attachment of the polymer to both the cantilever tip as well as to the substrate. One end is a -$SiCl_3$ group while the other end is a -$SiH(CH_3)_2$ group (Figure 2.6). These polymers were synthesized in two steps: The first step was an equilibration and condensation reaction of linear PDMS with -Si-H end groups and linear PDMS oligomers with silanol end groups catalyzed by phosphorous nitrile chloride. The product of this reaction is a mixture of PDMS polymers with mostly one Si-H group on one end and one silanol group on the other end.[61] In a second step the silanol groups react with $SiCl_4$ to form the aforementioned $SiCl_3$ end groups. The presence of these two endgroups is assessed by nuclear magnetic resonance measurements. The used PDMS polymers are

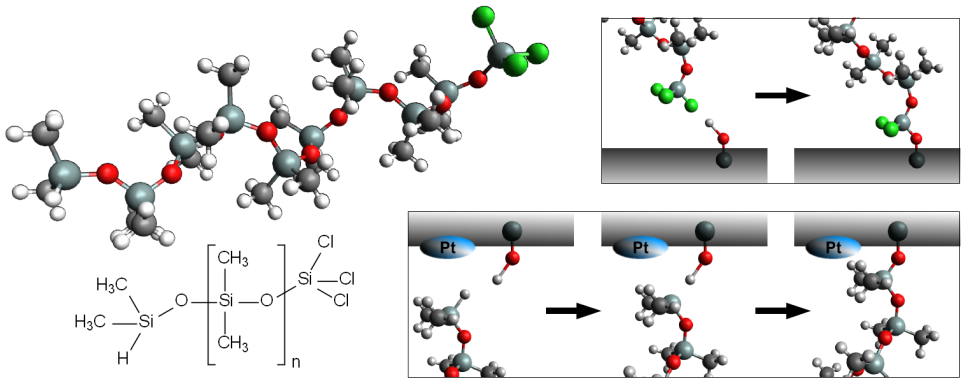


Figure 2.6: Two depictions of the polydimethylsiloxane polymer structure with the reactive endgroups -$CH(CH_3)_2$ on the left and -CCl_3 on the right (black: carbon, white: hydrogen, red: oxygen, grey: silicon, green: chlorine). Upper inset: proposed reaction of the polymer with the substrate, which occurs during the sample preparation step. Lower inset: proposed reaction with the cantilever tip, which is catalyzed by platinum clusters and only occurs near the contact area of the cantilever tip.

of a variable length between 50 and 1000 nm. Even the shortest length of 50 nm should be sufficient to minimize interactions between cantilever and substrate surface.[38,41] At these lengths PDMS is a viscuous liquid which is readily soluble in solvents such as hexamethyldisiloxane.

2.3.2 Sample Preparation

Covalent attachment is necessary for being able to perform force spectroscopy measurements on covalent bonds. Non-covalent attachment is much weaker than the covalent siloxane bond and thus would rupture before enough force is applied to rupture the siloxane bond.

Covalent attachment to the surface (insets in Figure 2.6) is provided by a spontaneous reaction of the -$SiCl_3$ group with the silanol groups of the substrate. Therefore oxidized silicon wafer were used (gift from Wacker Chemie AG).[38,41] The reaction yields a covalent Si-O bond between the substrate and the polymer as well as HCl as a byproduct. The Si-H bond on the other end of the polymer can be cleaved in a reaction catalyzed by platinum resulting in a free bond at the Si atom and a free proton. To create platinum clusters on a cantilever tip, it is sputtered with platinum. The conditions for the sputtering process were chosen to provide isolated

platinum clusters which leave the silicon nitride surface of the cantilever tip accessible.[38,41] The free bond at the Si atom, created by the cleaving off of the hydrogen atom, will now react with oxygen atoms, which can be regularly found on the surface of the silicon nitride cantilever tip as a result of oxidation, and form a second covalent bond. Because the covalently attached polymers are chains of Si-O-Si bonds which are attached on both ends via Si-O-Si bonds, only siloxane bonds should be involved in the force spectroscopy measurements. This should make it possible to only measure the rupture of Si-O bonds. However theoretical works suggest that the rupture forces of Si-O bonds in an attached polymer can vary considerably depending on the location, with bonds at the attachment points rupturing at lower forces than bonds in the polymer chain.[62] Another complication is that the environment of the bonds, present reactants or the strength of applied forces can change the rupture mechanisms.[63]

An important task is also the determination and optimization of the yield of the double covalent attachment as compared to the non-covalent sorption of the polymer to either the cantilever or the wafer surface. In previous single molecule measurements using the same attachment method the effectiveness of the covalent attachment has been shown as a shift from low rupture forces to high rupture forces when platinum clusters are added to the cantilever.[38,41] Nonetheless a portion of attached polymers that are measured are still non-covalently attached and these ruptures have to be prevented or separated from the ruptures of covalent bonds.

During the sample preparation the polymers are first attached to the wafer surface on one end, while the other end remains active. The cleaving off of the hydrogen atom at the remaining active polymer ends is triggered by the approach of the cantilever, because this reaction can only occur in the presence of platinum, which is located solely on the cantilever tip. The number of polymers, which can react with the cantilever tip when it is in contact with the sample surface, is limited to polymers in close proximity to platinum, i.e. polymers close to the contact point between wafer and tip. A large number of polymers can be measured in this way, because the contact area is very small and can be moved to another location on the sample, if not enough reactive polymer ends remain at a previously probed location. This highly localized nature of the Si-H cleavage reaction also makes measurements over a long time possible, since the Si-H bond is

relatively stable.

To provide a chemical environment for the force spectroscopy experiments which is similar to solid state silicone, HMDS is used as a solvent for the duration of the measurements as was previously mentioned.[38,41]

2.4 Bayesian Data Analysis

This section is based on introductions to this topic by Jaynes,[64] Sivia[65] and Muschielok.[66]

To gain information about the siloxane bond from force spectroscopy measurements it is necessary to compare the measured data to models for bond rupture using a rigorous statistical approach. Bayesian data analysis provides parameter estimations for various bond rupture models, which includes additional information such as the precision of the parameter estimation and correlations of the parameters among others. Bayesian data analysis or probabilistic data analysis is a form of probability calculus first proposed by Bayes in 1763.[67] Two axioms of probability calculus are the basis of Bayesian data analysis. The first axiom, which is called the sum rule, states that the sum of the probability $p(X|I)$ that proposition X is true and the probability $p(\overline{X}/I)$ that X is not true, when the background information I is given, has to be equal to 1:

$$p(X|I) + p(\overline{X}|I) = 1 \tag{2.5}$$

For a continuous variable X this can be formulated as:

$$\int dX p(X|I) = 1 \tag{2.6}$$

with $p(X|I)$ being the probability density function for X. The result of this is the normalization of the probability density $p(X|I)$.

The product rule is the second axiom and states that the probability $p(X, Y|I)$ that both the proposition X and the proposition Y are true when I is given, can be calculated by multiplying the probability $p(X|Y, I)$, that X is true when Y and I are given, with the probability $p(Y|I)$, that Y is true when I is given:

$$p(X, Y|I) = p(X|Y, I)p(Y|I) \tag{2.7}$$

X and Y can be easily switched to arrive at the following:

$$p(Y, X|I) = p(Y|X, I)p(X|I) \tag{2.8}$$

Since $p(X, Y|I)$ and $p(Y, X|I)$ are equal, Bayes' theorem can now be found by equating the right hand sides of the previous two formulas and rearranging them:

$$p(X|Y, I) = \frac{p(Y|X, I)p(X|I)}{p(Y|I)} \tag{2.9}$$

If X is called hypothesis θ and Y is called data d, the utility of Bayes' theorem becomes apparent:

$$p(\theta|d, I) = \frac{p(d|\theta, I)p(\theta|I)}{p(d|I)} \tag{2.10}$$

Bayes' theorem makes possible the calculation of the probability that a hypothesis is true for the given data. Bayes' theorem consists of four terms with the following names and implications:

- $p(\theta|I)$ is the prior probability and represents the state of knowledge prior to the data analysis.

- $p(d|\theta, I)$ is called the likelihood function and considers the experimental data when given a certain hypothesis and background information I.

- $p(d|I)$ is the evidence and is independent of the model and the model parameters. Once the data has been measured it can be used as a control.

- Finally $p(\theta|d, I)$ is called the posterior, which stands for the knowledge of the probability that the hypothesis is true for the given data.

From equations 2.5 and 2.7 it is also possible to derive the marginalization rule:

$$p(X|I) = \int_{-\infty}^{+\infty} dY p(X, Y|I) \tag{2.11}$$

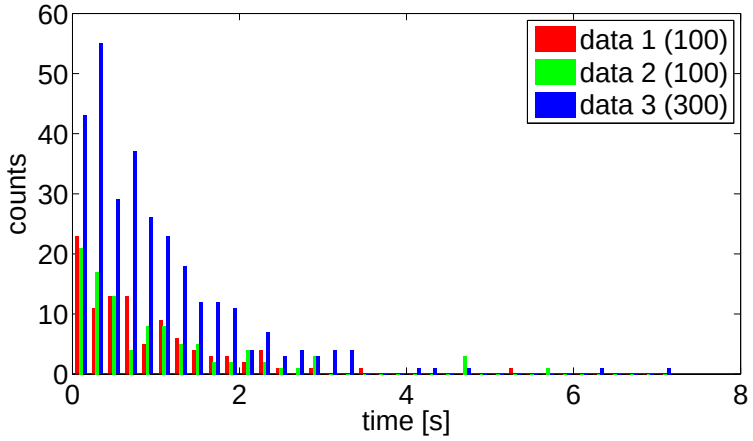


Figure 2.7: Three sets of simulated exponential decay data. All three sets were simulated with a rate of $1s^{-1}$. For the first two sets (1 = red, 2 = green) 100 events were simulated and for the third set (3 = blue) 300 events were simulated.

Marginalization is useful, when nuisance parameters such as background signal are present. By integration across all possible values of the nuisance parameter the marginal posterior can be calculated.

In the following Bayesian data analysis is explained in more detail using the example of analyzing data from a Poisson process to find the rate parameter. This example was chosen, as the rupture behavior of a bond under constant force is such a process.

Three independent data sets of a poisson process with a rate parameter k of 1 were simulated. Data sets 1 and 2 consist of 100 events, while data set 3 consists of 300 events. A histogram of all three decays is shown in figure 2.7.

For the prior $p(k|I)$ usually a uniform flat shape is chosen, as no information about the value of the parameter is available. This flat prior fulfills the condition $p(k|I) = constant$. Information such as the fact that no negative rate parameters are possible can be introduced as constraints to the prior. Prior experiments can also be used to shape the prior. If even the scale of the parameter is unknown, a prior shape has to be chosen, which is uniform across different scales of the prior. This prior is called the Jeffreys' prior and fulfills the condition $p(k|I) = constant/k$. In practice this can be achieved by using a prior which is flat in the logarithm of the parameter:

$$p(ln(k)|I) = constant \tag{2.12}$$

This is equivalent to Jeffreys' prior since:

$$p(ln(k)|I)d(ln(k)) = constant \cdot \frac{1}{k}dk = p(k|I)dk \tag{2.13}$$

The likelihood function for data created by a Poisson process is given by:

$$L = \prod_{i=1}^{N} k \exp^{-kt_i} \tag{2.14}$$

with N being the amount of measured events, t_i being the time when an event is measured and k being the rate constant. Some examples for the likelihood function at different decay rates k are shown in figure 2.8 It has to be noted that the likelihood function has to be normalized. This normalization can be recognized in figure 2.8 as the integral of the likelihood function over the parameter at all three rate parameters correspond to the value of 1.

The evidence term normalizes the product of likelihood and prior. The evidence is a constant since both the measured data and the background information are constant in the analysis. Normalization is necessary because the likelihood function is only normalized as a function of the data and not as a function for the parameters. The evidence can be calculated by integrating the product of the likelihood and the prior over all parameters:

$$p(d|I) = \int d\theta \, p(d|\theta, I)p(\theta|I) \tag{2.15}$$

Using the prior, the likelihood function and the evidence, all three data sets can be analyzed with Bayes' theorem to arrive at the corresponding posteriors. Since only one parameter is involved, evaluating the posterior at fixed intervalls of the parameter is feasible to calculate the posterior. The resulting posterior densities for the three data sets shown in figure 2.7 can be found in figure 2.9. While the posterior densities of all three measurements are clustered around the rate of 1, the maxima of the posterior densities are not centered on this value. The increased amount of data for the third simulation results in a smaller posterior density with a maximum closer to

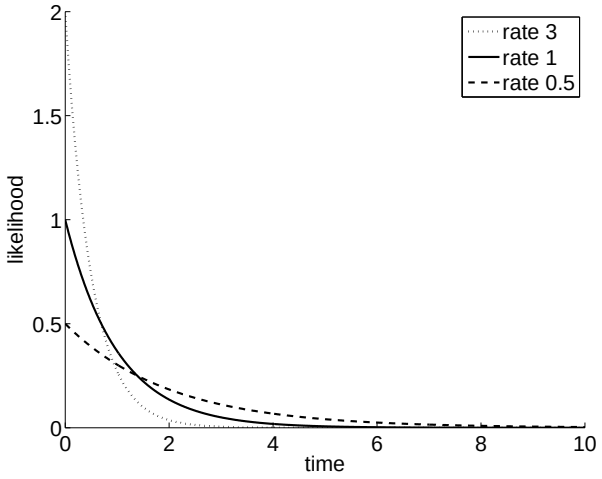


Figure 2.8: Likelihood functions for Poisson processes with three different rate parameters. The likelihood functions are normalized and as a result the areas below the functions are equal. For the highest rate at 3 s^{-1} (dashed line) a large amount of the area is located at low times when compared to the other two likelihood functions. For the lowest rate parameter at 0.5 s^{-1} (dotted line) the area below the likelihood function is distributed over a wider range of times when compared to the other two functions. As expected the likelihood function for the rate parameter 1 s^{-1} (solid line) is inbetween the two others.

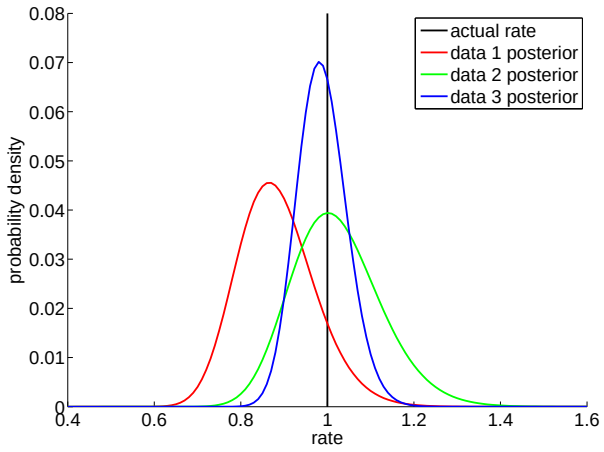


Figure 2.9: Posteriors calculated from three sets of simulated exponential decay data (Figure 2.7) using Bayesian data analysis. The actual rate 1 s^{-1} used for all three simulations is marked with a solid black line. For both small data sets (red and green) of 100 events the posterior contains a reasonable amount of probability density at the rate of 1 s^{-1}. For the large data set (blue) of 300 simulated events, the probability density distribution is narrower than for the small data sets. This is as expected, since a larger amount of data should lead to a more precise knowledge of the analyzed parameter. It is also expected that the maxima of posterior probability densities are closer to the actual rate the higher the amount of data is.

the true rate. This shows that a larger amount of data leads to better and more precise knowledge of the true value of the parameter.

2.5 Nested Sampling

This section is based on introductions to the nested sampling algorithm by Skilling[68] and Muschielok.[66]

The nested sampling algorithm by Skilling[68] is a Monte Carlo based Bayesian parameter estimation method (see section 2.4) and was used to calculate the posterior probability densities for the bond rupture model parameters. The applied implementation of the nested sampling algorithm was also used for FRET based nanopositioning experiments[69–72] and FCS analysis.[73] In brief, the nested sampling algorithm computes the posterior and evidence by compressing a set of objects, which are distributed across

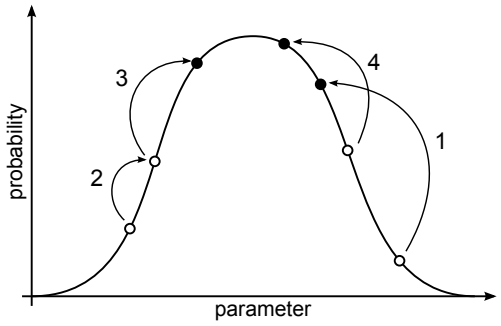


Figure 2.10: The nested sampling process compresses a set of objects (full circles) in the parameter space towards regions of high likelihood. This is demonstrated in four steps. Step 1: The object with the lowest likelihood value is removed and it's position is recorded as a sample (empty circle). Then a new independent object is created at a position with a higher likelihood then the new likelihood threshold. Step 2: The next lowest object is removed, a new threshold is set and a new independent object is generated. Step 3: The object created in step 2 is now the object with the lowest likelihood and immediately removed and replaced with a new object at a higher likelihood value. Step 4: Another object is now at the lowest likelihood value. It is removed, it's position in the parameter space is recorded as a sample and then it is replaced by a newly generated independent object.

the parameter space, towards regions of high likelihood values by means of an ever increasing likelihood threshold and removal of objects below this threshold and addition of objects above it (Figure 2.10). The recorded object positions in the parameter space sample the posterior and through this process the evidence can be estimated and a set of equally weighted samples can be found, which represent the posterior probability density. Nested sampling is especially useful for problems with many parameter dimensions or a posterior containing small regions with a very high probability density and large regions with a small probability density.

To be able to compute the evidence and the posterior probability density, it is necessary to use a sorted likelihood function. The concept of the sorted likelihood function will be explained in the following. Next the process of nested sampling is described. And finally a process to select equally weighted samples which represent the posterior probability density is shown.

2.5.1 Sorted Likelihood Function

To understand the sorted likelihood function one has to first consider the prior $p(\theta|I) \equiv \pi(\theta)$ to be subdivided into small elements with equal amounts of prior probability. Volume elements δV_k of the parameter space contain prior probability $\delta\pi_k$ according to:

$$\delta\pi_k = \int\limits_{\theta\in\delta V_k} d\theta\, \pi(\theta) \tag{2.16}$$

If the prior is flat, δV_k of equal size also contain equal amounts of prior probability. If this is not the case the volume elements have to be scaled according to the prior likelihood distribution, so that each element contains the same prior probability. Under this condition all $\delta\pi_k$ are equally sized and thus $\delta\pi_k \equiv \delta\pi$. Using this and the approximate likelihood L_k for each of the prior probability elements, an approximation of the evidence Z can be calculated:

$$Z \approx \sum_k L_k \delta\pi_k = \sum_k L_k \delta\pi \tag{2.17}$$

Next one has to consider that it is possible to arrange these prior probability elements according to their likelihood L_k in a descending order $L_k \geq L_{k+1}$. This also remains true when going to infinitely small prior probability elements $\delta\pi \to d\pi$. For this arrangement of infinitesimal prior probability elements the sorted prior mass ξ can be defined as the amount of prior probability in the parameter space where the likelihood is larger than a given value λ:

$$\xi(\lambda) = \int\limits_{L(\theta)>\lambda} d\theta\, \pi(\theta) \tag{2.18}$$

By inversion of equation 2.18 the sorted likelihood function (Figure 2.11) can be defined as:

$$L(\xi) \equiv \lambda \tag{2.19}$$

where the likelihood is a function of the sorted prior mass ξ. Since the elements $d\xi$ are equal to the previously defined $d\pi$ the area below $L(\xi)$ is equal to the evidence Z (Figure 2.11) according to:

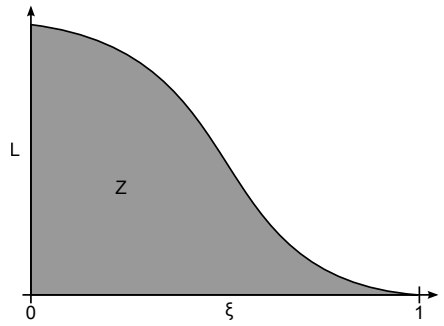


Figure 2.11: Sorted likelihood function with $L(\xi)$. The area below the function (grey) corresponds to the evidence Z. For illustrative purposes this diagram is not to scale. Usually most of the evidence Z is located at very low values of ξ.

$$Z = \int d\xi \, L(\xi) \tag{2.20}$$

2.5.2 Algorithm

The nested sampling method compresses a set of objects in the parameter space into regions of higher likelihood over multiple iterations (Figure 2.10). With each iteration of this process the object with the lowest likelihood is replaced by a new object with a position at a higher likelihood than the old object. The old discarded object is recorded as a sample. New independent objects, which satisfy this ever increasing minimum likelihood threshold, are generated by means of Markov chain Monte Carlo. From a sorted likelihood perspective the objects are at first distributed uniformly across the full range of ξ and then compressed towards lower values of ξ with each iteration. The threshold is first set at the highest value of ξ of all objects and moves towards lower values of ξ with each iteration (Figure 2.12).

The associated amounts of posterior probability mass for each sample j are defined as it's weight w_j and can also be calculated when ξ is known for a sample:

$$w_j = L_j \cdot (\xi_{j-1} - \xi_j) \tag{2.21}$$

The sum of the weights of all samples would then be the area below L(ξ)

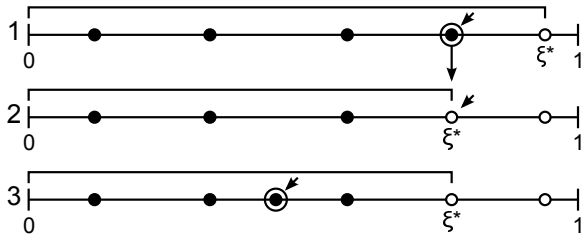


Figure 2.12: Nested sampling process viewed as compression of objects over ξ. The sorted likelihood function $L(\xi)$ is not shown, only the changes in objects, samples and thresholds in ξ for three examplary steps in the nested sampling process. Step 1: All four objects (filled circles) are located in an interval between 0 and ξ^*. The object with the lowest likelihood is identified by finding the object with the highest value of ξ. Step 2: The object is removed and determines the new threshold ξ^*. It's position is also recorded as a sample (empty circle). Step 3: A new independent object is added within the new threshold.

and thus the evidence is approximately:

$$Z \approx \sum_j w_j \tag{2.22}$$

Unfortunately it is not possible to directly determine ξ for a sample. But it is possible to calculate estimate values of ξ based on the average positions of the objects in ξ under an assumption of a uniform distribution in an interval of ξ and that this interval is shrinking by removing the objects with the highest values of ξ with each iteration. For this estimation it is advantageous to use the shrinking ratio $t_j = \xi_j/\xi_{j-1}$ of the prior masses in two consecutive iterations. If one considers the amount of objects uniformly distributed in an interval of $[0, \xi_{j-1})$ then the object with the highest value of ξ will be located closer to the threshold ξ the higher the amount of objects is. The distribution of shrinking ratios for M objects according to Skilling[68] is thus the following:

$$p(t_j) = M t_j^{M-1} \tag{2.23}$$

From this the mean and standard deviation for the logarithm of t_j (Figure 2.13) can be found:[68]

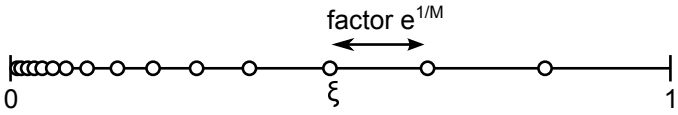


Figure 2.13: Nested sampling generates samples from objects which are compressed into an ever shrinking interval of 0 to ξ^*. This results in an increasing sample density at low values of ξ. On average the shrinking factor for each iteration of the nested sampling algorithm is $e^{1/M}$.

$$\langle \ln t_j \rangle = -1/M \tag{2.24}$$

$$\Delta \ln t_j = 1/M \tag{2.25}$$

Now the positions of the samples or in other words the objects which are discarded with each iteration can be estimated as the product of successive shrinking ratios (Figure 2.13):

$$\xi_j \approx \prod_{k=1}^{j} t_k \tag{2.26}$$

Taking into account that the shrinking ratio of each iteration is independently distributed according to equation 2.23 the mean and standard deviation of $\ln \xi_j$ for iteration j are:

$$\langle \ln \xi_j \rangle = -j/M \tag{2.27}$$

$$\Delta \ln \xi_j = \sqrt{j}/M \tag{2.28}$$

Inserting these estimates into equations 2.21 and 2.22 leads to the following estimations of the weight of a sample and the evidence:

$$w_j = L_j \cdot [\exp(\langle \ln \xi_{j-1} \rangle) - \exp(\langle \ln \xi_j \rangle)] = L_j \cdot \left[\exp\left(-\frac{j-1}{M} \right) - \exp\left(-\frac{j}{M} \right) \right] \tag{2.29}$$

$$Z \approx \sum_j w_j = \sum_j L_j \cdot \left[\exp\left(-\frac{j-1}{M}\right) - \exp\left(-\frac{j}{M}\right) \right] \tag{2.30}$$

Because these are estimations based on average sample positions, the mean and standard deviation of evidence have to be taken into account:

$$\langle \ln Z \rangle \approx \left(\sum_j w_j \right) \tag{2.31}$$

$$\Delta \ln Z \approx \sqrt{\frac{H}{M}} \tag{2.32}$$

with H being the information or negative entropy of the posterior relative to the prior:

$$H \approx \sum_j \frac{w_j}{Z} \ln\left(\frac{L_j}{Z}\right) \tag{2.33}$$

Hence the amount of uncertainty decreases with an increase of the number of objects M. The formulas are based on the logarithm of Z because the logarithm of the evidence distribution is symmetrical which is more convenient for giving a symmetric standard deviation.

To terminate the nested sampling process one can use the properties of $w_j = L_j \cdot (\xi_{j-1} - \xi_j)$ over the iterations. At the beginning the weights w_j are low because most of the objects are in regions with low L_j. Then the weights increase as the objects approach regions of high likelihood. At the same time $\xi_{j-1} - \xi_j$ shrinks with each iteration. Finally as the shrinkage of $\xi_{j-1} - \xi_j$ begins to dominate over the increases in L_j the maximum weight is reached and weights begin to decrease again.

The largest weights and thus contributions to Z are expected to occur around $\xi \approx \exp(-H)$, which is approximately at $j = MH$. Thus the calculation is best terminated, when the iteration number j significantly exceeds MH.[68] For this work a termination at $j = 2MH$ was sufficient. But it has to be noted, that the exact point of termination has to be adjusted on a case by case basis as the shape of the weight over iteration function can vary strongly. One example could be a small region of very high likelihood inside a large region of lower likelihood, which would lead to a small maximum

in weights followed by a dip in weights before an even larger maximum of weights is reached. This could lead to a too early termination of the algorithm if the chosen termination distance to $j = MH$ is not large enough.

2.5.3 Creating a Representation of the Posterior

The samples produced by the nested sampling algorithm can be used to create a representation of the posterior (see section 2.4). And the posterior contains the parameter estimations for the repective bond rupture model parameters.

This is possible, if those samples are of the same weight. In that case the density of the samples is proportional to the posterior probability density. The samples created with nested sampling do not possess equal weights and cannot be directly used to represent the posterior. But with staircase sampling[65] it is possible to create a set of equally weighted samples drawn from the original samples (Figure 2.14).

The eponymous staircase S_k is constructed by accumulating the normalized weight of the samples, which are sorted in ascending order of their iteration:

$$S_k = \sum_{j=1}^{k} \frac{w_j}{w_{max}} \tag{2.34}$$

Because of the normalization by the maximum weight w_{max} of the samples the maximum step size equals 1. Now a random number r between 0 and 1 is selected. Samples of equal weight are then chosen by selecting the sample belonging to the first step which rises above the value of $r + n$ with $n = 0, 1, 2, ...$ this is repeated until $r + n$ exceeds the maximum height of the steps. Because the maximum height between two steps equals one no sample can be drawn more than once (Figure 2.14).

The number of equally weighted samples that can be extracted from a set of samples is dependent on the maximum step size and the number of objects, as the number of objects also increases the number of samples that are created during the nested sampling process. With a higher number of equally weighted samples it is possible to represent the posterior more accurately, but on the other hand the higher number of objects also increases the computational cost. Thus it is necessary to determine the minimal num-

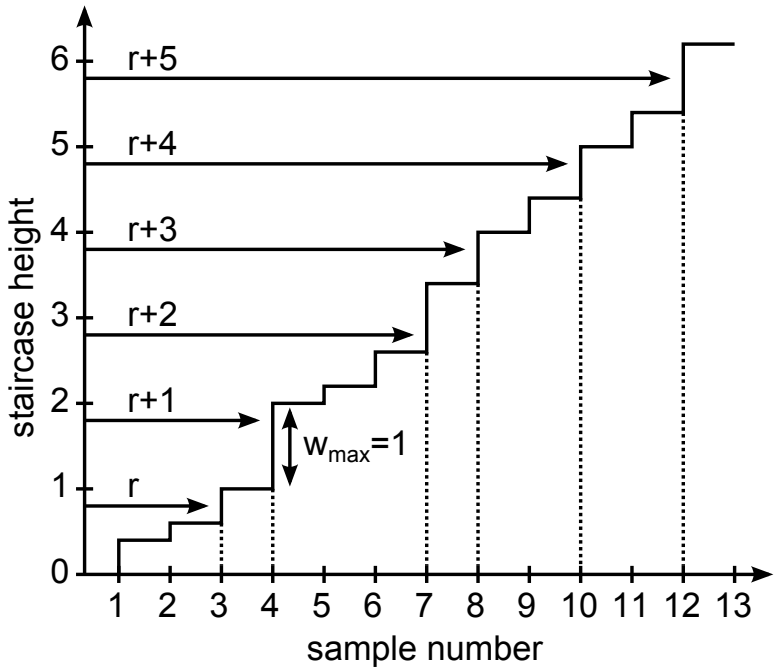


Figure 2.14: Staircase sampling for 13 samples. The sample from the fourth iteration (4) is the sample with the maximum weight and thus the height of this step in the staircase is 1 and all other steps are are of lower height than 1. r is a random number between 0 and 1. Equally weighted samples are determined by intersections at $r, r+1, r+2, ...$ with the staircase. Because the maximum step height is 1 and the height distance between each intersection is 1, no sample can be drawn more than once. In this example the samples drawn for the set of equally weighted samples are 3, 4, 7, 8, 10 and 12.

ber of objects that is sufficient to create a representation of the posterior. Any higher number of objects leads to an unneccessary expense of computational ressources. The determination of a sufficient number of objects is unfortunately subject to a process of trial and error, similarly to determining when the nested sampling process can be terminated. The number depends mainly on the parameter space and the posterior shape. As a rule of thumb the number of objects has to be high for large parameter spaces (number of parameters and the available range of values for each parameter) and for complex posterior shapes (e.g. very small and thin regions of high likelihood). If it is not possible to use previous similar calculations as guidance, one should start with a low number of objects and increase it generously with each try. In this work a number betweeen 200 and 1000 objects were used for the nested sampling, with 300 and 500 being the most common numbers of objects.

2.6 Bond Rupture Models

In order to gain insight into the strength of a bond probed by single molecule force spectroscopy, a bond rupture model has to be developed, which allows to determine characteristic parameters in conjunction with the nested sampling algorithm (see section 2.5). In this work a basic phenomenological and two different microscopic models for the rupture of bonds were examined and compared. The microscopic description of force dependent bond rupture is identical for all models:[22,23,31–33] diffusion through an energy potential along the reaction coordinate from an initial bound state in an energy well across an activation barrier. This energy landscape, which contains an energy well and an activation barrier, is deformed by application of force, which results in a tilt of the landscape. This tilt becomes more drastic with increasing forces and ultimately leads to the disappearance of the activation barrier, when the critical force F_{crit} is applied. The resulting behavior not only predicts rising rupture rates with increasing forces, but also describes a reduction of the distance between the minimum and the transition state under increasing forces, which is known as the Hammond effect.[74,75] In contrast to this model, the phenomenological Bell model does not take this change of distance into consideration.[30,76]

2.6.1 BELL model

A model using a phenomenological approach first proposed by Bell was developed by Evans et al.[29] under the assumption of an exponential increase of the rupture rate for increasing forces. The bond parameters for this model are the distance, $x^{\ddagger}$, between the energy minimum and the transition state and the rupture rate at zero force, k_0.[24,30] Smaller values for $x^{\ddagger}$ and smaller values for k_0 lead to a more stable the bond. The rupture rate as a function of the applied force, F, can then be computed by:[28]

$$k(F) = k_0 e^{Fx^{\ddagger}/k_B T} \tag{2.35}$$

where k_B is the Boltzmann constant and T the temperature, at which the experiment is performed.

This phenomenological model is well suited only for low forces, since it does not account for the force dependent change of the distance between the minimum and the transition state.[30,76] This changing distance leads to a deviation from the exponential increase of the rupture rate towards lower rupture rates than predicted.[74,75] Forces which increase k to values which are only two times higher than k_0 for example, typically already exhibit deviations of k in the range of 10 % when compared to more complex models.

Based on the rupture rate function of the Bell model, the likelihood function $p(F|v)$ for the loading rate dependent distribution of rupture forces for the Bell model can be calculated:[28]

$$p(F|v) = \frac{k(F)}{Kv} \times \exp\left(\frac{k_0}{x^{\ddagger}Kv}\right) \times \exp\left[-\frac{k(F)}{x^{\ddagger}Kv}\right] \tag{2.36}$$

with v being the speed at which the bond is pulled and K being the spring constant of the bond and thus the loading rate being Kv (Figure 2.15).

2.6.2 CUSP model

To more closely match the observed rupture behavior of bonds at higher forces, microscopic models have been developed using the theory of diffusive barrier crossing by Kramers,[27,31] various energy potential shapes[27,34] and the deformation of those potentials under force. In these microscopic models, the bond is diffusing through an energy landscape with an energy minimum and a maximum and bond rupture is defined as the moment when

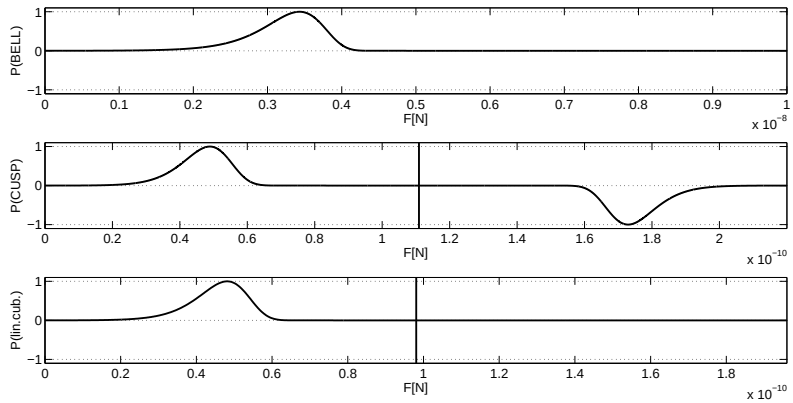


Figure 2.15: Constant speed rupture force likelihood function comparison. Likelihood functions for rupture forces in constant speed experiments for the Bell model, the CUSP model and the linear-cubic model (from top to bottom). For the Bell model the following parameters were used: $x^{\ddagger} = 10.5 \cdot 10^{-12}$ m and $k_0 = 0.01\ s^{-1}$. For the CUSP model the following parameters were used: $\Delta G^{\ddagger} = 6.1 \cdot 10^{-20}$ J, $x^{\ddagger} = 1.1 \cdot 10^{-9}$ m, $A = 2.7 \cdot 10^{4}$ s^{-1}. And for the linear-cubic model the following parameters were used: $\Delta G^{\ddagger} = 5.5 \cdot 10^{-20}$ J, $x^{\ddagger} = 1.0 \cdot 10^{-9}$ m, $D = 6.3 \cdot 10^{-16}$ $\mathrm{m}^2\mathrm{s}^{-1}$. For the CUSP and linear-cubic model the critical force was marked by a vertical line. The concept of a critical force does not apply to the Bell model.

the diffusion process has passed the maximum and a return to the minimum is not possible anymore. This condition that no back reaction should be possible is necessary, since at the point of bond rupture both ends of the bonds are rapidly moved apart by several processes: the force, that was stretching the bond, now pulls the ends apart instead, the intact parts of the polymer and linkers return to a much less linear relaxed conformation and the bonds, which were deformed during the stretching process, return to an energetically more favorable state. Taken together these processes make a back reaction very unlikely. Here, forces acting on the bond are represented by tilting the energy potential. High forces lead to a strong tilt of the energy landscape. This tilt not only lowers the height of the energy barrier, but also changes the distance between the minimum and the barrier and it also changes the slopes in the potential.[22,23,27] One of the potentials used is the CUSP potential, which is a harmonic potential with a cusp at the position of the energy maximum, which prevents back reactions, after the maximum has been passed, and is also a good approximation of the Morse potential for bonds (Figure 2.16):[27,28,35,77]

$$U_F(x) = \begin{cases} \Delta G^{\ddagger}(x/x^{\ddagger})^2 - Fx, & \text{if } x < x^{\ddagger} \\ -\infty & \text{else} \end{cases} \tag{2.37}$$

The height of the energy barrier here is $\Delta G^{\ddagger}$ and a higher $\Delta G^{\ddagger}$ results in a more stable bond and lower rupture rates. x is the position in the potential and corresponds to the extension of the bond. At a position x of zero the minimum of the potential is located. The left hand part of the term creates the basic shape of the potential, while the term $-Fx$ introduces the force dependent tilt of the potential. Kramers rate theory[31] is based on the Langevin equation, which describes Brownian motion in a potential and applies several simplifications: The crossed energy barrier is assumed to be high, which is valid in the case of single-molecule force spectroscopy, as long as the forces are small, and the system is assumed to be in the equilibrium state, which is true for constant speed measurements as the stretching process is much slower than the diffusion through the potential. And for constant force experiments this condition is also true, since the system is held at a certain force for long times.

To solve the Langevin Equation the fluctuation dissipation theorem or Einstein-Smoluchowski, the Focker-Planck equation related to the Langevin

equation and a Taylor-development of the potential at it's inflection point are used. Using this, an analytic solution for the rupture rate can be formulated:[27]

$$k(F) = A \times \exp\left(\frac{-\Delta G^{\ddagger}}{k_B T}\right) \times \left(1 - \frac{1}{2}\frac{F x^{\ddagger}}{\Delta G^{\ddagger}}\right) \times \exp\left\{\frac{\Delta G^{\ddagger}}{k_B T}\left[1 - \left(1 - \frac{1}{2}\frac{F x^{\ddagger}}{\Delta G^{\ddagger}}\right)^2\right]\right\} \tag{2.38}$$

A is the attempt frequency, i.e. more stable bonds have smaller values of A. The force at which the energy barrier completely vanishes is the critical force F_{crit}. At the point where $F = F_{crit}$, the slope of the linear force dependent tilting term, $-Fx$, becomes $-F_{crit}$ and cancels out the slope F_{crit} at the energy barrier, so that the slope at this position becomes horizontal. Because of the hyperbolic shape of the potential the minimum is found at the position with horizontal slope, this means the minimum now is at the position of the cusp and the maximum has vanished. Thus F_{crit} is the slope of the bond potential at the position $x^{\ddagger}$ is:[35]

$$F_{crit} = \lim_{x \to x^{\ddagger}} \frac{d}{dx} \Delta G^{\ddagger}(x/x^{\ddagger})^2 = 2\frac{\Delta G^{\ddagger}}{x^{\ddagger}} \tag{2.39}$$

The advantage of the CUSP model is, that in this model force now influences the distance between the minimum and maximum of the energy potential and as a consequence the rupture rate behaves more similar to experimental results than the Bell model.[35] The CUSP model can be typically used at holding forces five to ten times higher than the Bell model. The analytic nature of the model also provides exact results while being fast to calculate at the same time. There are drawbacks to this analytical model caused by the high barrier approximation and the cut off of the parabolic bond potential by a cusp, which leads to a discontinuity at this position. It also leads to a non-zero slope at the inflection point, while the energy barrier should be formed by a continuous change of the slope from positive to negative values.

These drawbacks manifest themselves at higher forces, where the bond begins to stabilize and rupture rates actually begin to decrease (Figure 2.17). This stabilization of the bond with increasing forces happens well below the

critical force and should be taken into account when the need arises to judge the validity of a fit, that uses this model. At high holding forces the validity of the CUSP model should be checked with care. 70 % of the critical force are used as an upper limit for the potentials in this work.

Following from the rupture rate function of the CUSP model, the likelihood function of the CUSP model for the loading rate dependent rupture is:[28]

$$p(F|v) = \frac{k(F)}{Kv} \times \exp\left[A \times \exp\left(\frac{-\Delta G^{\ddagger}}{k_B T}\right) \times \frac{1}{x^{\ddagger} K v}\right] \times \exp\left\{-\left[\frac{k(F)}{x^{\ddagger} K v}\right]\left[1 - \frac{1}{2}\frac{F x^{\ddagger}}{\Delta G^{\ddagger}}\right]^{-1}\right\} \tag{2.40}$$

Naturally also the likelihood function needed for the constant speed experiments is affected by the previously mentioned limitations of the CUSP model (Figure 2.15): The most obvious effect of the model limitations introduced by the CUSP potential shape and the high barrier approximation are the negative likelihood densities at forces higher than the critical force. But as long as the loading rates are low, the expected rupture forces are far away from those sections of negative likelihood density, which makes it safe to ignore them. Another way in which the limitations show up, is the area of the likelihood function. This area should be 1 by definition. But as soon as a considerable amount of likelihood density is located near the critical force, the area between zero force and the critical force shrinks below 1. The magnitude of this artifact is far from small and the area of the function continues to shrink with rising loading rates (Figure 2.18). The amount of negative likelihood values near the critical force also increases with higher loading rates. So for high loading rates, which result in rupture forces close to the critical force, the CUSP model with the high barrier approximation is not valid.

The high barrier approximation can be avoided, by applying the full mean-first-passage times formula (MFPT)[28] to the CUSP potential $U_F(x)$:

$$k(F) = D \times \left[\int_{x_{min}}^{x_{max}} \exp\left(\frac{U_F(x)}{k_B T}\right) \int_0^x \left(\frac{-U_F(y)}{k_B T}\right) \mathrm{d}y \mathrm{d}x\right]^{-1} \tag{2.41}$$

The parameter D is the diffusion coefficient and similar to A: A higher diffusion coefficient represents a less stable bond. The integration border x_{min} has to be chosen far enough left from the minimum so that the likelihood for the diffusive process to reach this region is low. The integration border x_{max} on the other hand has to be far enough to the right of the energy barrier, that a back reaction is unlikely. A consequence of these modifications is that an analytical expression for the holding force dependent rupture force cannot easily be found. Instead we had to use numerical methods.

As can be seen in figure 2.18 the shape of the CUSP potential itself proves problematic at higher forces: the rupture rate increases quickly to infinity when the critical force F_{crit} is approached even though a slowdown in the increase of the rupture rates with the force is what has to be expected according to experimental observations. Thus a model that describes bonds under forces close to the critical force has to abandon both the high-barrier approximation and the CUSP energy potential.

2.6.3 Linear-Cubic Model

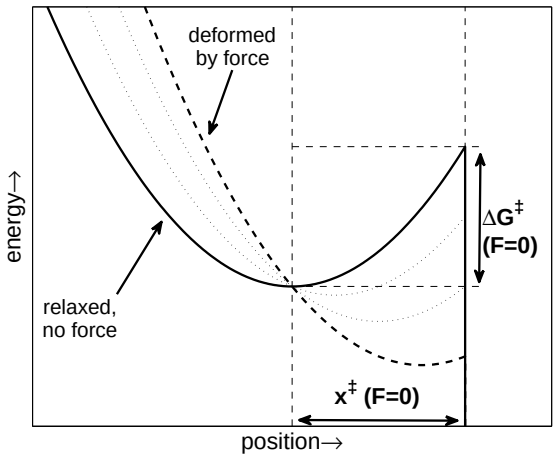
The limitations of the CUSP model mentioned in the previous subsection can be circumvented by both replacing the CUSP potential with a linear-cubic potential[27,28,77] and avoiding the high barrier approximation by using the full mean-first-passage time formula.[28,32,78–80] As shown above, avoiding the high barrier approximation alone is not sufficient (Figure 2.17). The force dependent linear-cubic potential is described by the following equation:[35]

$$U_F(x) = \frac{3}{2}\Delta G^{\ddagger}\frac{x}{x^{\ddagger}} - 2\Delta G^{\ddagger}\left(\frac{x}{x^{\ddagger}}\right)^3 - Fx \tag{2.42}$$

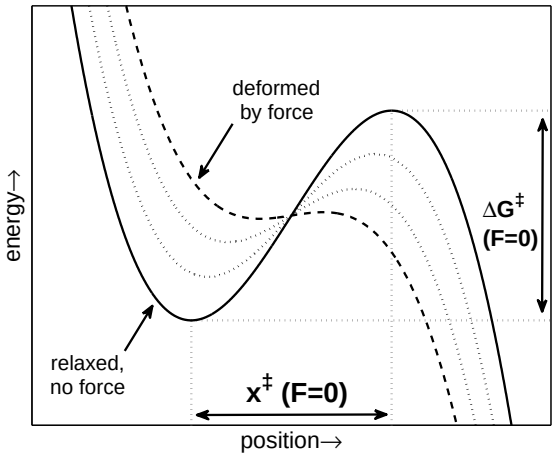
The first two terms (linear and cubic) together create the basic shape of the potential (Figure 2.16), while the third term (linear) provides the force dependent tilt of the potential. It's critical force is reached once the first term is fully compensated by the third term of the potential and both the minimum and maximum have vanished at the position of $x = 0$:

$$F_{crit} = \frac{3}{2}\frac{\Delta G^{\ddagger}}{x^{\ddagger}} \tag{2.43}$$

Analog to the CUSP potential the high barrier approximation can then be avoided by using the full mean-first-passage times formula for the force-



A



B

Figure 2.16: (A) Illustration of the CUSP energy landscape and its deformation resulting from application of force. The solid line represents the relaxed state of the CUSP potential with zero force acting on the bond. The distance between the minimum and the transition state is $x^{\ddagger}$ while the height of the energy barrier is $\Delta G^{\ddagger}$. The dashed line shows the CUSP potential being deformed by a force. The deformed potential has both a lower barrier height and a lower distance between the minimum and the energy barrier. The dotted lines are intermediate deformation steps at lower forces. (B) Illustration of the linear-cubic energy potential and it's deformation under force. The solid line represents the potential at zero force, the dashed line shows the potential being deformed by a holding force and the dotted lines show the influence of intermediate forces. It should be noted that similar to the CUSP model both the height of the energy barrier and the distance between the barrier and the well are reduced by rising forces.

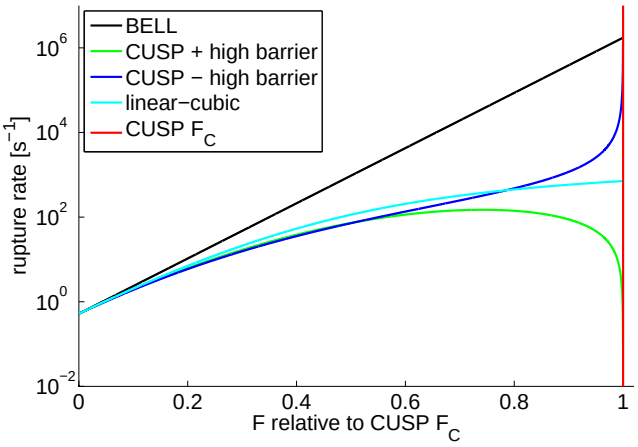


Figure 2.17: Force dependent rupture rate function for different bond rupture models. The Bell potential model (black), the high barrier approximation CUSP potential model (green), the CUSP potential model without the high barrier approximation (blue) and finally the linear-cubic potential model (cyan). The force is relative to the critical force of both CUSP potentials (red). The critical force of the linear-cubic model is at 0.75 of the CUSP critical force (not depicted). The Bell model predicts an exponential increase with increasing holding forces. The other three models are less simple and exhibit a slower increase of the rupture rates than the Bell model at low forces. But both CUSP models have a limited range of validity, which is well below their critical force and exceeding this range has severe consequences: with the high barrier approximation the rupture rates stop increasing and decrease instead and without the high barrier approximation the increase of rupture rates accelerates again instead of further slowing down. The linear-cubic model on the other hand continues to slow down the increase of the rupture rates even past its critical force as it approaches the diffusion rate.

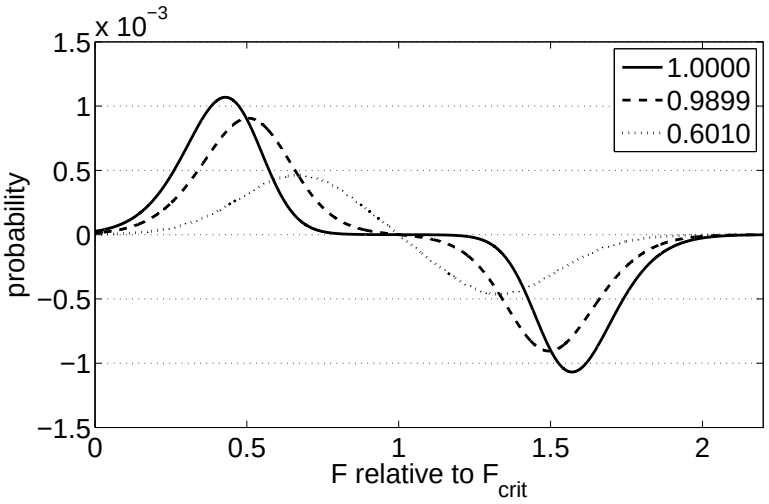


Figure 2.18: Probability density functions (PDF) of rupture forces for the CUSP model using the analytical expression for the high barrier approximation. The PDF shown for varying given rates demonstrates the limitations of the model for constant speed data analysis: Firstly, beyond the critical force the probability density function yields negative values for the likelihood function, which have to be ignored for the purpose of data analysis, i.e. the likelihood function is simply defined only for forces lower than F_{crit}. Secondly, at higher loading rates, with considerable probability densities close to the critical force, the likelihood function is no longer normalized to 1, but shrinks to lower values. The lowest loading rate (solid line) was chosen so that the bulk of the PDF is clearly separated from the critical force and the area between zero force and F_{crit} is practically 1. At a two times higher loading rate (dashed line), the area is at 0.9899. At a 10 times higher loading rate (dotted line) the area has already shrunk to 0.6010.

dependent rupture rate:[28,32,78–80]

$$k(F) = D \times \left[\int_{x_{min}}^{x_{max}} \exp\left(\frac{U_F(x)}{k_B T}\right) \int_0^x \left(\frac{-U_F(y)}{k_B T}\right) \mathrm{d}y \mathrm{d}x \right]^{-1} \tag{2.44}$$

In this case $U_F(x)$ corresponds to the linear cubic potential from equation 2.42. There is also no readily found analytical solution for this equation, which once again makes a numerical solution necessary. The integration borders have to be chosen so that the potential region left from the integral is unlikely to be reached and the right border should extend so far that a back reaction is unlikely. For the linear-cubic potentials and force ranges examined in this work a lower limit x_{min} of $1.5 \cdot x^{\ddagger}$ left from the potential minimum and an upper limit x_{max} of $1.5 \cdot x^{\ddagger}$ right from the potential maximum with a total width of 4 $x^{\ddagger}$ proved as sufficient when combined with a sample point density of 200 points distributed equidistantly across the range of the aforementioned 4 $x^{\ddagger}$.

This model has no obvious limitations up to the critical force, but the numerical approach has the drawback of a higher computational cost in comparison to the CUSP model and the results of this model are only approximate and rely on a sufficient sampling point density (Figure 2.17).

With $k(F)$ for the linear-cubic model and the assumption of a constant loading rate Kv during the pulling process, the loading rate dependent likelihood function of the linear-cubic model for the rupture forces (Figure 2.15) can be written as:[35]

$$p(F|v) = \frac{k(F)}{Kv} \times \exp\left[-\int_0^F \frac{k(F')}{Kv} \mathrm{d}F' \right] \tag{2.45}$$

The same advantages and drawbacks such as no limitation to low forces or the higher computational cost apply to the constant speed model.

2.6.4 Potentials used in this work

For the purpose of simulating various constant force and constant speed measurements the potential from the unfolding process of the fourth filamin domain (ddFLN4, $\Delta G^{\ddagger} = 6.1 \cdot 10^{-20}$ J, $x^{\ddagger} = 1.1 \cdot 10^{-9}$ m, $A = 2.7 \cdot 10^{4}$ s^{-1}), was chosen.[78,81] It is a non-covalent potential, which was used for first simulation tests and simulated data based on these parameters proved to behave

similar to experimental data. This should be true for many potentials as long as force ranges are compared relative to the respective critical forces of different potentials (i.e. 10 to 70% $F_{crit\ potential\ A}$ compared to 10 to 70% $F_{crit\ potential\ B}$, see subsection 2.6.5). Nevertheless using a more similar covalent potential would have been preferable. To maximize the comparability of both bond rupture models, the potential parameters for the linear-cubic model ($\Delta G^{\ddagger} = 5.5 \cdot 10^{-20}$ J, $x^{\ddagger} = 1.0 \cdot 10^{-9}$ m, $D = 6.3 \cdot 10^{-16}$ $m^2 s^{-1}$) have been determined by simulating ddFLN4 constant force data using the CUSP bond rupture model with the high barrier approximation in a force range of 10 to 70% of the respective critical force and then performing a parameter estimation using the linear-cubic bond rupture model without high barrier approximation. This way similar behavior of the potentials is ensured as long as the forces are well below the critical force. Additionally for the comparison of covalent and non-covalent bond potentials potentials found through the analysis of the unfolding of I27 subunits of the protein titin ($\Delta G^{\ddagger} = 8.1 \cdot 10^{-20}$ J, $x^{\ddagger} = 0.4 \cdot 10^{-9}$ m, $A = 4.8 \cdot 10^{4}$ s^{-1}),[27,82] of the rupture processes of the silyl ester bond of carboxymethyl-amylose polymers covalently attached to a silicium oxide substrate (CMA, Si-O bond, more frequent (72%) and less stable: $\Delta G^{\ddagger} = 6.31 \cdot 10^{-20}$ J, $x^{\ddagger} = 12.3 \cdot 10^{-12}$ m, $A = 1.9 \cdot 10^{6}$ s^{-1}, less frequent (28%) and more stable: $\Delta G^{\ddagger} = 6.14 \cdot 10^{-20}$ J, $x^{\ddagger} = 12.0 \cdot 10^{-12}$ m, $A = 5.5 \cdot 10^{4}$ s^{-1})[83–85] and of the rupture of the siloxane bond of covalently attached polydimethylsiloxane polymers (PDMS, Si-O bond, $\Delta G^{\ddagger} = 3.0 \cdot 10^{-20}$ J, $x^{\ddagger} = 1.2 \cdot 10^{-11}$ m, $A = 2.9 \cdot 10^{3}$ s^{-1})[38,41] were used.

2.6.5 Behavior of the CUSP potential with high barrier approximation and the linear-cubic potential without high barrier approximation

To better understand the CUSP model and the linear-cubic model and the bond rupture behavior they are predicting, the influence of the three potential parameters on the force dependent rupture rates were examined. $\Delta G^{\ddagger}$, $x^{\ddagger}$ and A of the CUSP bond rupture model were tested by halving and doubling the values of the parameters of the ddFLN4 potential(Figure 2.19). The most straightforward parameter is the attempt frequency A: the variation of A is shifting the rupture rates proportionally across the full force range. The energy barrier height $\Delta G^{\ddagger}$ on the other hand is more com-

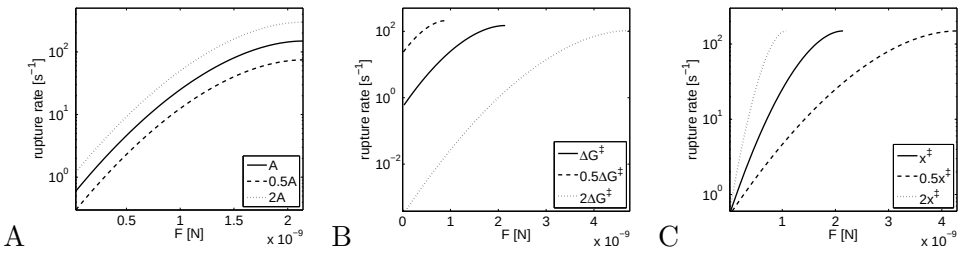


Figure 2.19: Parameter variation of the CUSP bond rupture model. Plot of the force dependence of the rupture rate of the ddFLN4 CUSP potential with high barrier approximation in comparison to variations of the same potential demonstrating the effects of the variation of each of the three potential parameters attempt frequency, barrier height and barrier position. The black line always represents the unmodified potential, the dashed line corresponds to a halved value of the parameter and the dotted line to a doubled value of the parameter. The force range is set between zero force and 70 % of the F_{crit} of the respective potential. The varied parameters are the attempt frequency A (A), the barrier height $\Delta G^{\ddagger}$ (B) and the distance between the minimum and the energy barrier $x^{\ddagger}$ (C).

plicated. A higher energy barrier leads to a more stable bond with high rupture rates being reached at several times higher forces. But the low force range at double $\Delta G^{\ddagger}$ also shows rupture rates which are several orders of magnitude lower compared to the unchanged $\Delta G^{\ddagger}$. This means the range of rupture rates is far wider for high $\Delta G^{\ddagger}$. While high values of $\Delta G^{\ddagger}$ lead to more stable bonds, for $x^{\ddagger}$ shorter distances stabilize the bond. In contrast to the barrier height, the variation of the barrier distance doesn't change the range of rupture rates. The linear-cubic model exhibits practically the same behavior as the CUSP model with D replacing A. Only $x^{\ddagger}$ exhibits an additional shift to higher rupture rates for lower values of $x^{\ddagger}$ (Figure 2.20).

2.6.6 Behavior of various covalent and non-covalent bond rupture potentials

In the following the behavior of potentials found through analysis of experimental data is explored, to give a baseline of expected bond rupture behavior and variance of this behavior for this work. The rupture rate functions for covalent bond and non-covalent bond potentials are dissimilar from each other (Figure 2.21): The two non-covalent bond potentials investigated

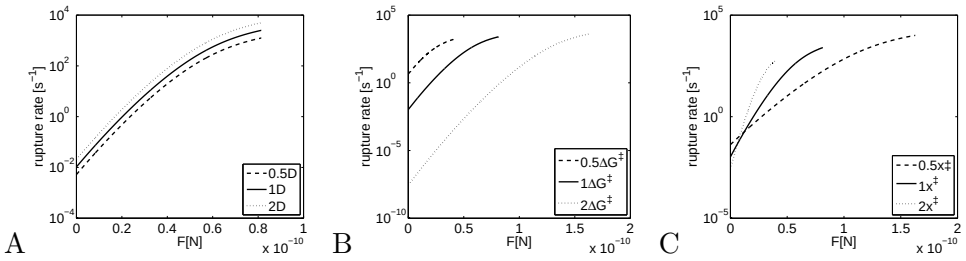


Figure 2.20: Parameter variation of the linear-cubic bond rupture model. Plot of the force dependence of the rupture rate of the ddFLN4 linear-cubic potential in comparison to variations of the same potential demonstrating the effects of the variation of each of the three potential parameters diffusion coefficient D (A), barrier height $\Delta G^{\ddagger}$ (B), and barrier position $x^{\ddagger}$ (C).The black line always represents the unmodified potential, the dashed line corresponds to a halved value of the parameter and the dotted line to a doubled value of the parameter. The force range is set between zero force and 100 % of the F_{crit} of the respective potentials. One remarkable difference to the behavior of the CUSP model with high barrier approximation is the crossover which occurs when $x^{\ddagger}$ is varied.

are those from titin[27,82] and from ddFLN4[78,81] and the three covalent bond potentials are from the two observed rupture processes observed in bond rupture experiments of the silyl ester bond in covalently attached CMA[83–85] and from the siloxane bond from covalently attached PDMS. How the parameters from the rupture of covalently attached PDMS are found is described in section 3.1. The covalent bond potentials have in common that their barrier distance $x^{\ddagger}$ is drastically shorter than for the non-covalent potentials. Overall the model predicts the rupture rates rising more slowly with rising forces and thus a higher stability of the covalent bond potentials as is expected. The range of rupture rates for PDMS is especially small when compared to the other potentials.

2.7 Parameter Estimation

2.7.1 Parameter Estimation for Constant Force Bond Rupture Experiments

With the nested sampling algorithm described in section 2.5, which is a Bayesian data analysis method (see section 2.4), and the bond rupture mod-

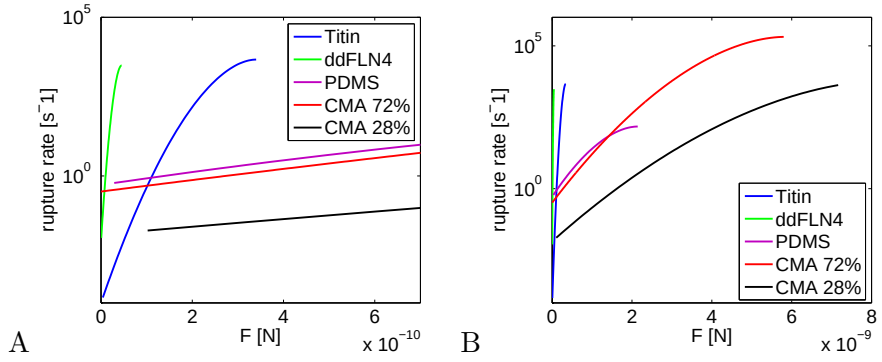


Figure 2.21: Comparison of three covalent and two non-covalent bond potentials over a range of holding forces from 0 to 6 nN. (A) shows a zoomed in section from 0 to 0.7 nN and (B) shows the full force range from 0 to 8 nN. Titin (blue line) and ddFLN4 (green line) are both examples for potentials of non-covalent bonds. The three covalent bond potentials are the two potentials from the two observed rupture processes of the silyl ester bond in covalently attached CMA (more frequent and less stable: red line; less frequent and more stable: black line) and the siloxane bond in covalently attached PDMS (magenta line). $x^{\ddagger}$ is short in comparison to the non-covalent potentials.

els described in section 2.6 (BELL, CUSP potential with high barrier approximation and linear-cubic potential without high barrier approximation) it is possible to estimate the parameters for those models using experimental force spectroscopy data. The bond rupture models predict force dependent rupture rates via their respective expressions for $k(F)$. For a single bond, that is held and has a rupture rate $k(F)$, the distribution of measured holding times will take the form of an exponential distribution. Thus one arrives at the following log likelihood function for an exponential decay with a rupture rate of $k(F)$:

$$\log L = \sum_{i=1}^{n} \log \left[k(F_n) \mathrm{e}^{-t_n k(F_n)} \right] \tag{2.46}$$

with t_n being the time to rupture of a polymer n held at the force F_n and $k(F_n)$ being the rupture rate at the force F_n for the respective rupture model. This function is used for all models. To reduce the computational cost for the numerical calculations of the linear-cubic model, rupture rates at twenty different equidistant forces between the minimum and maximum

holding force of the current force range are pre-calculated and used to interpolate values of k. For the linear-cubic model the integration borders we used were $3/2 \cdot x^{\ddagger}$ left from the minimum for x_{min} and $3/2 \cdot x^{\ddagger}$ right from the maximum for x_{max}. 200 equidistant sampling points were used for the numerical integration.

The Jeffreys prior[86] is flat over all scales (see section 2.4) and was thus used for the parameters $\Delta G^{\ddagger}$, A and D, where there was scarce knowledge of the scale of the parameter or where a low precision of the analysis results was expected. A flat prior was used for the parameter $x^{\ddagger}$, since the scale was known and the expected range of values for the position of the potential was considered small.

2.7.2 Extension to Two Rupture Populations

It became apparent that the constant force rupture data could not be well described by a single population (see subsection 3.1.2) and as a result the likelihood function had to be adapted to allow for the simultaneous estimation of the parameters for two populations with different energy landscapes:

$$\log L = \sum_{i=1}^{n} \log \left[k_1(F_n) \mathrm{e}^{-t_n k_1(F_n)} \times a + k_2(F_n) \mathrm{e}^{-t_n k_2(F_n)} \times (1-a) \right] \quad (2.47)$$

The different populations 1 and 2 have the potential parameters $\Delta G_1^{\ddagger}$, $x_1^{\ddagger}$, A_1 or D_1, $\Delta G_2^{\ddagger}$, $x_2^{\ddagger}$ and A_2 or D_2, respectively. The parameter a is between 0 and 0.5 and denotes the relative size of the two populations, where without loss of generality population 1 is always smaller or the same size as population 2.

The priors used for the two population parameter analysis are analog to the one population analysis, with the Jeffreys prior being used for most parameters except the parameters $x_1^{\ddagger}$, $x_2^{\ddagger}$ and a. A flat prior for a was selected, since not only the scale was known, but also roughly a value of 0.25 was expected.

As mentioned previously just analyzing the rupture rate at one holding force will not yield enough information to determine the three respective parameters of the two bond potentials. Instead measuring across a range of forces is necessary for the parameter estimation. To determine the required

size of the force range is one subject of this work and will be discussed below (see subsection 3.2.2).

2.7.3 Parameter Estimation for Constant Speed Bond Rupture Experiments

Bayesian parameter estimation (see section 2.4) and the nested sampling algorithm (see section 2.5) were also used to analyze the constant speed data from previously published constant speed experiments.[38,41] The likelihood functions for the Bell, CUSP and linear-cubic model were already shown in the bond rupture model section (see section 2.6). Normalization of those likelihood functions is necessary, since the experimental constant speed data is cut off below 0.5 nN to exclude rupture data from non-covalently attached polymers. To achieve this normalization and to also reduce the computational cost of the analysis, the experimental data was used to calculate a 2D histogram of loading rate and rupture force (Figure 2.22). For each bin the likelihood value is calculated. The likelihood values for each bin are normalized according to the number of counts in each row of bins with the same loading rate. Afterwards the $\log L$ value for each bin is calculated and multiplied with the respective number of bond ruptures in that bin. Initial tests indicated that the analysis results for the CUSP model have to be expected to possess a critical force close to the observed rupture forces or even lower than some of the highest measured rupture forces. Since the likelihood values beyond F_{crit} are negative for the CUSP model, all rupture forces higher than the current F_{crit} are ignored. For the sake of normalization bins with a negative likelihood value are considered to have a likelihood value of zero. The number of equidistant sampling points for the force integral of the linear-cubic model was set to 200.

2.7.4 Parameter Estimation Result Visualization

Projected credible volumes: The parameter probability density of the posterior (see section 2.4) is visualized by extracting a set of equally-weighted posterior samples from the total number of samples acquired by the nested sampling process. To find this set of samples of equal weight we selected appropriate ones by subjecting all samples to staircase sampling (see subsection 2.5.3).[65] The equally weighted samples are then used to create projected

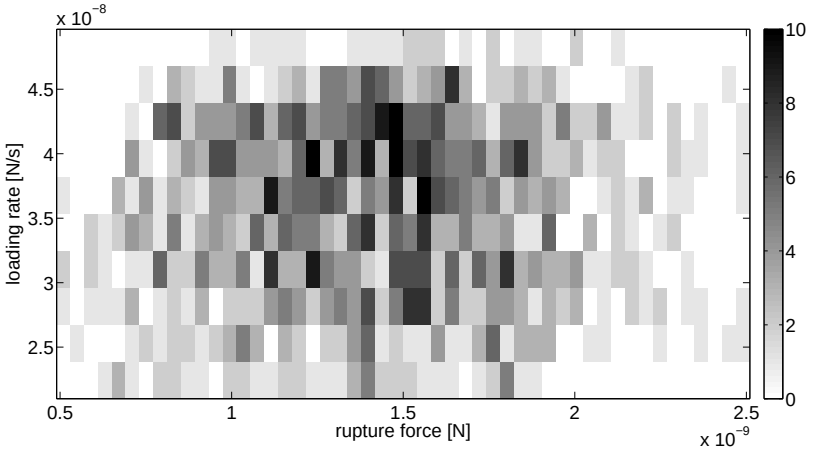


Figure 2.22: 2D histogram of constant speed bond rupture data. A relatively high number of columns for the rupture forces was chosen, to achieve a compromise between computational efficiency and accuracy.

credible volumes of the posterior probability density with contour lines at 0.5σ, 1σ and/or 2σ. This means that the contour lines in the projected credible volume plot are created by projecting the credible volumes which enclose 38.20 % (0.5σ), 68.27 % (1σ) or 95.45 % (2σ) of all equally weighted samples onto a twodimensional plane. Generally the parameter $\Delta G^{\ddagger}$ is on the x-axis of the plane and either $x^{\ddagger}$ or A/D are on the y-axis of the plane.

Marginal posterior probability distributions: To give more insight in the shape of the posterior, onedimensional histograms of the equally weighted samples can be shown in addition to the projected credible volume plots. The 1D histograms are created by only using one of the parameters of the equally weighted sample positions and thus represent the posterior with all but one parameter being marginalized. The 1D histograms are shown next to the axis of the projected credible volume plots that corresponds to the parameter that has not been marginalized in the histogram.

2.8 Data Simulation

Simulated data is useful in several ways. Simulated data can be used to test if parameter estimation algorithms are able to recover the input parameters of the simulation (see subsection 3.4.1) and also allows the comparison of different algorithms with regards to their performance. Furthermore mea-

surements using varying experimental parameters such as the amount of data, size of the range of pulling forces or loading rates, covalent and non-covalent bond potentials or even multiple rupture behaviors and more can be simulated and used as an input for the parameter estimation, which then allow to determine optimal measurement parameters for the experiment or if the accurate estimation of parameters is even possible (see section 3.2). An example for this would be the concurrent estimation of two bond potentials resulting in two different sets bond rupture potential parameters. Finally it is possible to examine the influence of properties of the measurement process itself on the parameter estimation result. An example for this is the determination of the cantilever spring constant and how various errors in the determination of it change the results of the data analysis. Examples for other properties of the measurement systems, which can be examined using simulations, are the temperature and it's stability or the noise and precision of the sensor which captures the deflection of the light beam reflected from the cantilever.

To simulate data a bond rupture model as well as bond potential parameters have to be given. The rupture model describes the bond rupture behavior according to the given parameters. It determines pulling forces or loading rates, calculates actual rupture times from resulting rupture rates or rupture forces from rupture force probability distributions and finally includes measurement process properties like the error of the cantilever spring constant if necessary. The models for the rupture bahavior of a bond are shared with the parameter estimation algorithms so the validity of the model and a correct implementation of those models have to be assumed. But it is possible to find an error, if it exists only in one implementation of either the simulation or analysis or if there are two different errors in each implementation. So if the input parameters of the simulation can be found by the parameter estimation it is ensured that both implementations of the rupture behavior are likely to be correct. For the simulation of pulling forces or loading rates one has to consider the experimental conditions and how close to match them. If data is distributed roughly evenly across all forces or loading rates a different distribution of rupture forces or loading rates has to be simulated than for the case when the amount of data decreases with an increase of the pulling forces or loading rates. These considerations are closely connected to the experimental conditions. For the simulation of data

from constant force and constant speed experiments the amount of data or the size and upper and lower bonds for the pulling forces or loading rates are of major importance since those parameter can help answer questions like: How much data at which holding forces has to be measured for a good estimation of the bond potential parameters?

2.8.1 Constant Force Data

To simulate constant force bond rupture data (see also section 9.2) a potential with the parameters $\Delta G^{\ddagger}$, $x^{\ddagger}$ and A or D, a lower and upper holding force limit and the total amount of simulated ruptures are chosen. The holding force dependent function for the rupture rate $k(F)$ is determined by the desired bond model, i.e. the Bell model, the CUSP model or the linear-cubic model. For the CUSP model the upper holding force limit choice has to be set with care to avoid the limitations of the model, which manifest well below the critical force. The simulation itself is repeating a loop for each needed rupture event, which contains a holding force F and a time to rupture t, until the predetermined amount of data has been created: a random holding force F is chosen from a flat distribution between the upper and lower force limits. Using $k(F)$ the rupture rate k corresponding to this force is then calculated. Finally the time to rupture t is randomly chosen from an exponential distribution with the rupture rate k. For two different rupture populations two different simulations are performed one after the other and the data sets are then concatenated. The order of rupture forces has no influence on the analysis result, since the $\log L$ values for all bond ruptures are summed up.

To examine the influence of errors in the determination of the spring constant (see subsection 2.2.3), simulations with built in spring constant determination errors are also needed. They are created by altering the actual holding force used to calculate k by an amount appropriate for the current spring constant error. Both the chosen and the actual holding force are recorded. The spring constant error can be altered for every bond rupture event according to a normal distribution or stays the same for the duration of the simulation. The order in which first the force for the erroneous spring constant is selected and then the actual force is calculated from this result is of importance, because it influences the range of forces at which bond ruptures are simulated (see subsection 3.2.4). In the used simulation order

the range of actually measured forces is larger than the desired range of forces, which would also happen in the experiment.

2.8.2 Constant Speed Data

Constant speed data simulation is also straightforward (see also section 9.2). At first a random loading rate is selected from a preset range of loading rates. Then a likelihood distribution of rupture forces is calculated for one of the corresponding bond rupture models. Then a rupture force is randomly selected according to this likelihood distribution. This is repeated until a sufficient number of bond rupture events has been simulated.

2.9 Experimental Methods and Materials for PDMS Single Molecule Force Spectroscopy

In the following the materials and methods for the sample preparation and for the measurements are shown. At first the AFM and the measurement setup can be found as well as the commercial software and the cantilevers that were used. Finally sample preparation and measurement procedures are shown including typical measurement parameters.

2.9.1 Atomic Force Microscope

The instrument used for the PDMS force spectroscopy measurements is the MFP3D AFM by Asylum Research. The xy-stage travel range is 90 μm travel in closed loop with a sensor noise below 0.5 nm averaged deviation in a 0.1 Hz to 1 kHz bandwidth. The z-stage in the AFM head has a travel range of 15 μm with a sensor noise below 0.25 nm averaged deviation in a 0.1 Hz to 1 kHz bandwidth. Z-height detection can be detected with a noise below 0.06 nm in a 0.1 Hz to 1 kHz bandwidth. The optical lever has sensor noise below 0.02 nm averaged deviation in a 0.1 Hz to 1 kHz bandwidth and has a class 1M superluminescent diode as a light source. The MFP3D is a bottom view model and is mounted on an Eclipse TE300 inverted microscope by Nikon. Isolation from acoustic noise is provided by a wooden box lined with acoustic foam. Vibrations are reduced by Newport optical tables on laminar flow isolators.

2.9.2 Commercial Software

Measurements and data aquisition on the AFM were performed using the MFP3D version 070111+217 software by Asylum Research. The MFP3D software is based on the Igor Pro version 5.05A programming software tool by WaveMetrics, Inc. Bayesian data analysis of the force spectroscopy data was performed under Matlab version 2010a by MathWorks in conjunction with Visual Studio 2008 and Visual Studio 2010 Express by Microsoft. For force spectroscopy data simulation and data visualization Matlab version 2011a by MathWorks was used.

2.9.3 Cantilevers

For the PDMS force spectroscopy measurements MLCT-AUHW triangular cantilevers by Veeco were used. The cantilevers had a spring constant of 0.03 N/m and a resonance frequency of 10 to 20 kHz in air.

2.9.4 Experimental Procedures and Materials

To catalyze the covalent attachment of the PDMS polymers to the cantilever tip, platinum has to be present. To achieve this platinum was sputtered onto the cantilever. Sputtering conditions were chosen so that the platinum forms clusters on part of the surface of the cantilever tip and not a closed layer. The formation of these clusters has been verified in previous works via scanning electron microscopy.[38,41]

The sample for the PDMS force spectroscopy measurements are oxidized silicon wafers with a layer PDMS polymers wich are covalently bound to the substrate on one end (Figure 2.23). The first preparation steps are performed under a nitrogen atmosphere in a plastic glove box. At first 30 μl of PDMS are dissolved in 4 ml HMDS in a small glass container. After allowing the PDMS to dissolve for 3 to 5 minutes, the solution is dropped onto the silicon wafer piece. The amount of solution should be sufficient to completely cover the surface of the wafer piece. For a wafer piece of ca. 1 cm^2 area 30 μl of solution are needed. After the solution has completely evaporated, the PDMS polymers should have covalently attached to the oxidized surface of the wafer piece and the substrate is rinsed with several milliliters of HMDS. Afterwards the substrate is placed in a large glass vessel with enough HMDS to cover it at all times and then placed on an

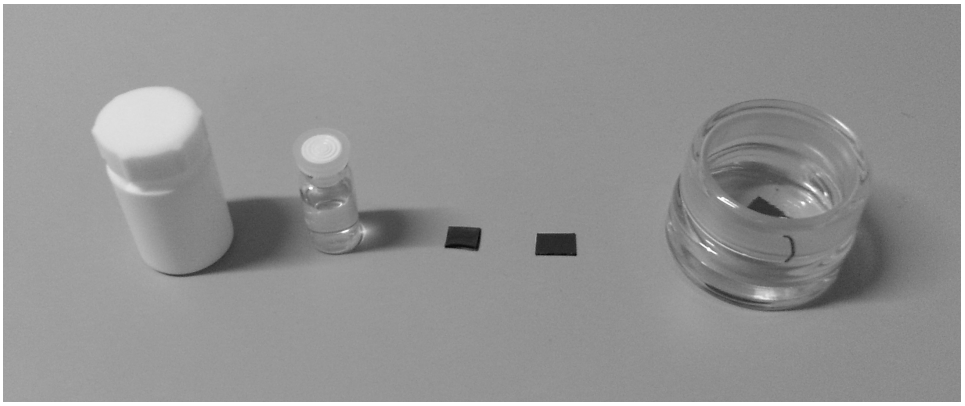


Figure 2.23: PDMS sample preparation. Preparation steps from left to right: A teflon vessel for PDMS storage, a small glass container used to dissolve PDMS in HMDS, an oxidized silicon wafer piece with PDMS solution, wafer piece after solution has evaporated, wafer piece in glass vessel for placement on orbital shaker over night. (For this picture water has been used instead of a solution of PDMS in HMDS.)

orbital shaker over night. This step removes non-covalently bound polymers. The next day the wafer piece is ready to be used for force spectroscopy measurements. PDMS, HMDS and the oxidized silicon wafers were provided by Wacker Chemie AG.

The PDMS constant force experiments were performed in HMDS to approximate a chemical environment similar to the solid state phase of silicone. HMDS is highly volatile and thus requires a a closed flow chamber (Figure 2.24). To close the flow chamber a silicone membrane has to be added to the AFM head (Figure 2.24). A new sample is taken out of its glass container and then rinsed once again with several milliliters of HMDS. It is then placed in the flow chamber and fixed in place by two neodymium magnets. Then the AFM head is lowered onto the flow chamber. The alignment of flow chamber, wafer piece and AFM head has to be carefully adjusted to allow for a closed seal of the chamber and to prevent the crashing of the AFM head. Finally the chamber is filled with HMDS and then the system is ready for measurements. Because the chamber is not fully closed, HMDS has to be refilled every 3 to 5 hours.

Constant force measurement preparation:

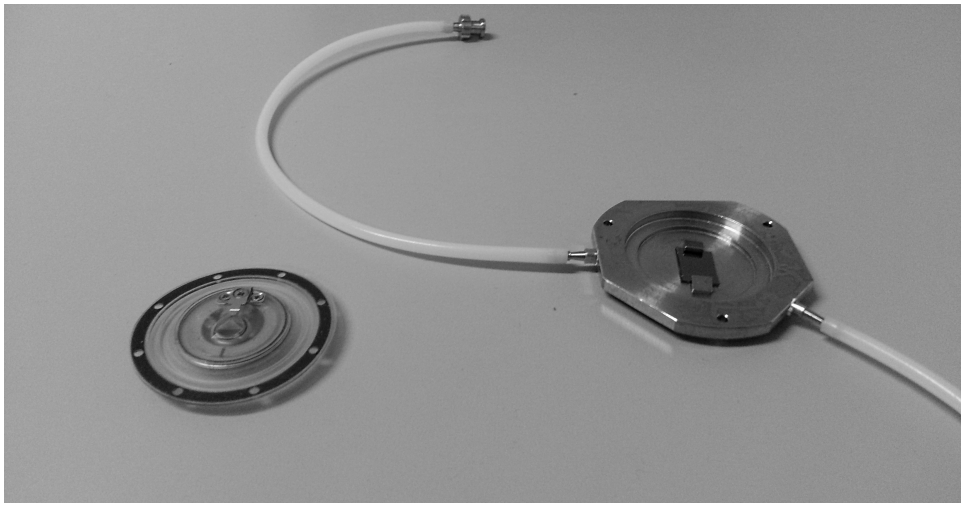


Figure 2.24: Left: AFM head with silicone seal, which reduces evaporation of HMDS solvent and prevents the solvent from entering the AFM interior. Right: Stainless steel flow chamber with silicone wafer piece held in place by two neodymium magnets.

1. Center light beam on cantilever using the microscope.
2. Place flow chamber with sample on xy-stage and fill with HMDS.
3. Recenter the light beam on the cantilever by maximizing the sum signal.
4. Measure a thermal of the thermal movement of the cantilever.
5. Measure the optical lever sensitivity by measuring a force distance curve of the cantilever being pressed into the subtrate surface.
6. Calibrate the cantilever spring constant by fitting the main resonance peak of the cantilever in the thermal.
7. Now the constant force measurement software can be loaded and the measurement parameters can be set.
8. The system is ready for constant force measurements.

Constant force measurement parameters:

The parameter Force on Surface determines with which force the cantilever is pressed onto the surface and the parameter Contact Time determines how long the duration of the contact is. Force Length and Seconds for Feedback set a time and a distance limit, which abort the measurement if they are exceeded to stop measurements with no attached polymer or measurements that would take too long. Gain is a parameter for the feedback loop. First Force, First Distance and First Gain are parameters for the phase of the measurement where the cantilever is pulled away from the surface until it detaches from the surface and the distance given by First Distance is reached. It should be noted that the the force set by First Force will not be reached in that phase. Prepull determines the force to which an attached polymer is pulled first, before it is held at the target force set by the Setpoint parameter. Data is recorded at the Sampling Frequency. For forces where long holding times are expected the Div parameter is used to reduce the amount of data. Otherwise the amount of data would be too large for the memory of the measurement computer. For a Div parameter of 2 data is reduced by half and for a Div parameter of 10 data is reduced by 10.

Typical parameters:

- Force on Surface: 5000 pN
- Contact Time: 5 s
- Force Length: 2000 nm
- Seconds for Feedback: 2000 s
- Gain: 15
- First Force: 2000 pN
- First Distance: 100 nm
- First Gain: 5
- Prepull: 1500 pN
- Setpoint: 800-2000 pN
- Sampling Frequency: 2000 Hz

- Div: 2 or 10

Chapter 3

Force Spectroscopy Results

In this thesis data from constant force single-molecule force spectroscopy measurements of PDMS was recorded and analyzed (see section 3.1). Then the parameters which influence the quality of the constant force data analysis are examined (see section 3.2). In section 3.3 constant speed data from previous works[38,41] is analyzed. The results of the constant force data analysis are finally compared to previous constant speed data[38,41] as well as to results of simulations[19,36,62] and also to the results of single-molecule force spectroscopy experiments on related Si-O bonds (see section 3.4).[84,85]

3.1 Constant Force Measurements

The results from the analysis of the data from the constant force measurements are subdivided into two major parts. At first two basic findings about the rupture behavior of the PDMS polymers are presented: the length independence of the rupture rates and the existence of two different rupture populations. After this the analysis results of the data using the Bell, CUSP and linear-cubic bond rupture model are shown in conjuction with the assessment of the validity of each respective analysis result.

3.1.1 Rupture Rate Length Independence

Samples of long polymers oftentimes show a large polydispersity. For the PDMS polymers (see section 2.3) in the sample used in the experiments the observed lengths fall mostly between 50 and 1000 nm (Figure 3.1). The longest contour length measured was 1350 nm and the mean contour length

was 199 nm. The lower limit of 50 nm was given by the minimum distance the cantilever had to move away from the surface to ascertain the actual detachment of the cantilever tip from the surface. Since the experiment was designed such that all bonds including the attachment points are siloxane bonds, all of these bonds were expected to rupture with the same probability thus leading to a length dependency of the rupture rates since larger polymers contain more Si-O bonds. However, over more than an order of magnitude in length no detectable change of the rupture rate is found (Figure 3.2, polymers held at forces between 1.3 to 1.5 pN were selected as this range provides a compromise between a large amount of data and a small force range). This length independence can only be explained, if the number of bonds which can be ruptured is constant for each polymer. This indicates that polymer rupture is caused by the rupture of a unique bond and independent of the polymer length. Most likely the rupture of one of the two terminal attachment points, as theoretical works have found that the Si-O bonds at the attachment point of a PDMS multimer to a β-cristobalite (100) surface are weaker than the Si-O bonds in the polymer chain.[62] In the following analysis of bond rupture behavior it will always assume that only a single bond is rupturable.

3.1.2 Multiple Rupture Populations

So far[38,41] analysis has assumed that only Si-O bonds are stretched and that thus the energetic landscape is similar for all bonds. Such a situation would lead to a mono-exponential distribution of rupture times for a certain holding force. In contrast when least-squares fits of cumulative probability distribution functions of mono-exponential and bi-exponential distributions to the cumulative probability distribution of experimental rupture times are performed, then the bi-exponential distribution clearly provides a better explanation for the data than the mono-exponential distribution (Figure 3.3; l2-norm of the residual 0.29 versus 1.48). For the the cumulative probability distribution of experimental rupture times, rupture events at forces between 1.3 nN to 1.5 nN were chosen as this representend the best compromise between a small force range and a sufficient amount of data.

The cumulative probability distribution of a mono-exponential distribution which is used for the least squares fit is:

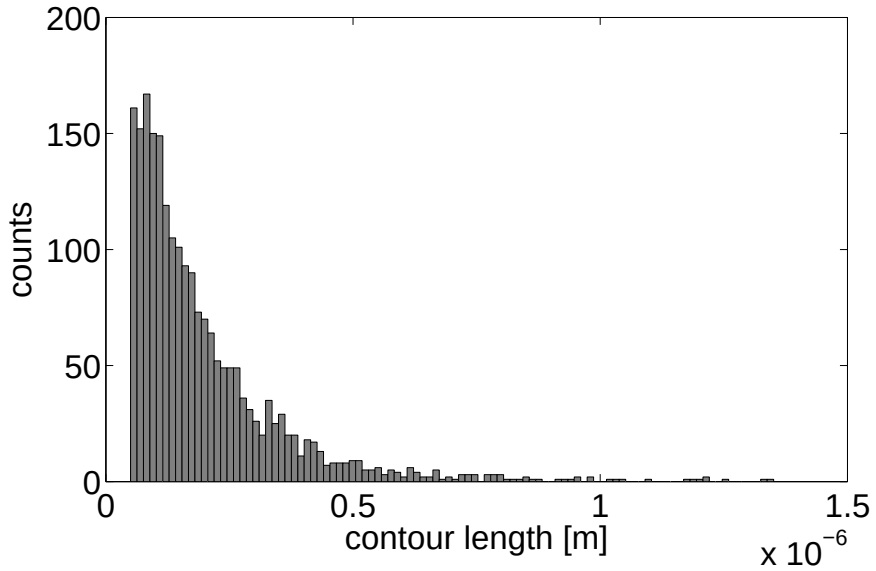


Figure 3.1: Histogram of contour lengths from polymers measured in constant force experiments. With increasing contour lengths the number of polymers sharply declines. No polymers with contour lengths below 50 nm are measured, because the cantilever has to move away from the surface by this distance to ensure detachment of the cantilever tip from the surface.

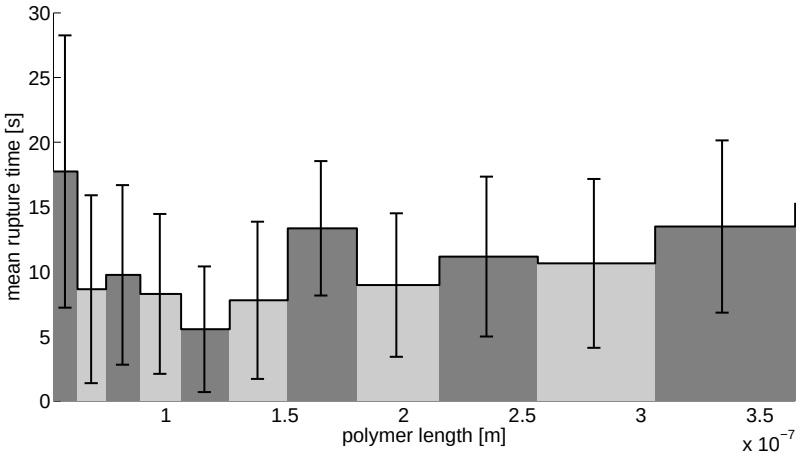


Figure 3.2: Mean rupture times versus polymer lengths of the experimental constant force data. Rupture times were collected between holding forces of 1.3 to 1.5 pN. The range of polymer lengths spans more than one order of magnitude, which also means that the number of bonds in the respective molecules varies by a factor larger than ten. The rupture times appear to be independent of the total number of bonds in a polymer.

$$P(t) = 1 - \mathrm{e}^{-kt} \tag{3.1}$$

and the cumulative probability distribution of a bi-exponential distribution for the least squares fit is:

$$P(t) = a(1 - \mathrm{e}^{-k_1 t}) + (1 - a)(1 - \mathrm{e}^{-k_2 t}) \tag{3.2}$$

with k, k_1 and k_2 being the respective rupture rates and t being the rupture time and a as well as $(a - 1)$ being the respective amplitudes.

This indicates that in the sample there are at least two different rupture populations with different parameters describing the potential energy landscape. The fit assuming a bi-exponential distribution of rupture events also indicates that the smaller population, corresponding to $33 \pm 0.47\%$ of the total bond ruptures, has a roughly ten times lower rupture rate than the bigger population. The existence of two rupture populations also makes their concurrent estimation in the data analysis process necessary.

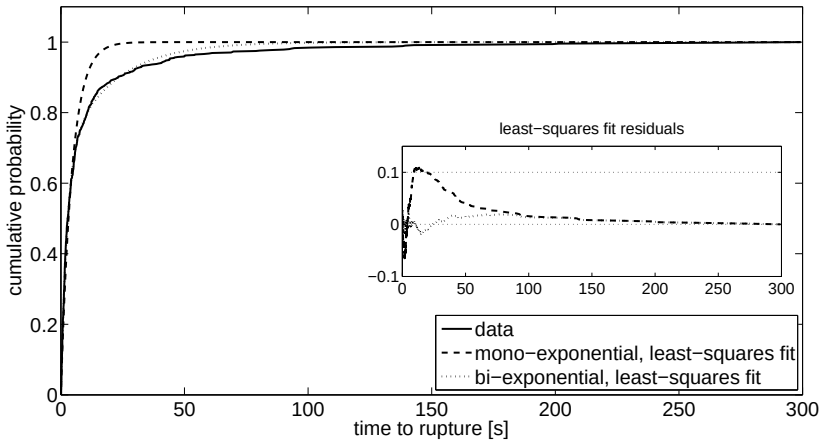


Figure 3.3: Cumulative probability distribution of rupture times between the holding forces of 1.3 to 1.5 nN (solid line) together with fits assuming a mono-exponential (dashed line) or a bi-exponential distribution (dotted line) of rupture times. The mean rupture time from the mono-exponential fit is 0.2237 ± 0.0020 s. The mean rupture times for the bi-exponential fit are 0.4119 ± 0.0033 s for the large population (67 ± 0.47%) and 0.0514 ± 0.0010 s for the small population (33 ± 0.47%).

3.1.3 Bell Model Results

The Bell model (see subsection 2.6.1) is the first of the three bond rupture models, that were used to analyze the data from constant force experiments (see subsection 2.2.1). It is phenomenological and does not take into account that forces pulling on a bond reduce the distance between minimum and maximum, which leads to a deviation from a purely exponential increase of rupture rates with the holding force. This deviation is also observed in nature, which means that the BELL model is only valid for low forces. A priori it is not known, whether the data measured from an unknown energy landscape is measured in a force range where the Bell model is still valid. Thus a comparison of the Bell model results with the results of other more complex bond rupture models is necessary: agreement between the Bell and the other models would suggest a range of low forces where the Bell model is still valid and a disagreement the opposite.

Analysis of the constant force bond rupture data (see subsection 2.2.1) using the nested sampling algorithm and the Bell model yields:

$x^{\ddagger} = 10.3 \pm 0.6$ pm (1σ) and $k_0 = 10.6 \cdot 10^{-3} \pm 2.4 \cdot 10^{-3}$ s^{-1} (1σ) for the big population and $x^{\ddagger} = 6.6 \pm 0.79$ pm (1σ) and $k_0 = 2.8 \cdot 10^{-3} \pm 0.78 \cdot 10^{-3}$ s^{-1} (1σ) for the small population (See figure 3.4).*

For both the small and the large population the shape of the posterior probability densities is small and well defined. For both populations there is a negative correlation between $x^{\ddagger}$ and k_0. The correlation is of the same degree for both populations (Figure 3.4). The small population accounts for 24 ± 1.5 % (1σ) of all rupture events. For this population both the rupture rate at zero force k_0 and the distance between minimum and energy barrier $x^{\ddagger}$ are smaller than the corresponding values for the large population, confirming our notion that the potential of the small population describes a more stable bond (see subsection 3.1.2). The proportion of 24.0 % is also comparable to the 33 % obtained from the fit of the cumulative distribution of rupture times assuming a bi-exponential decay of rupture times.

*Values and deviations given according to histograms of marginalized equally weighted samples.

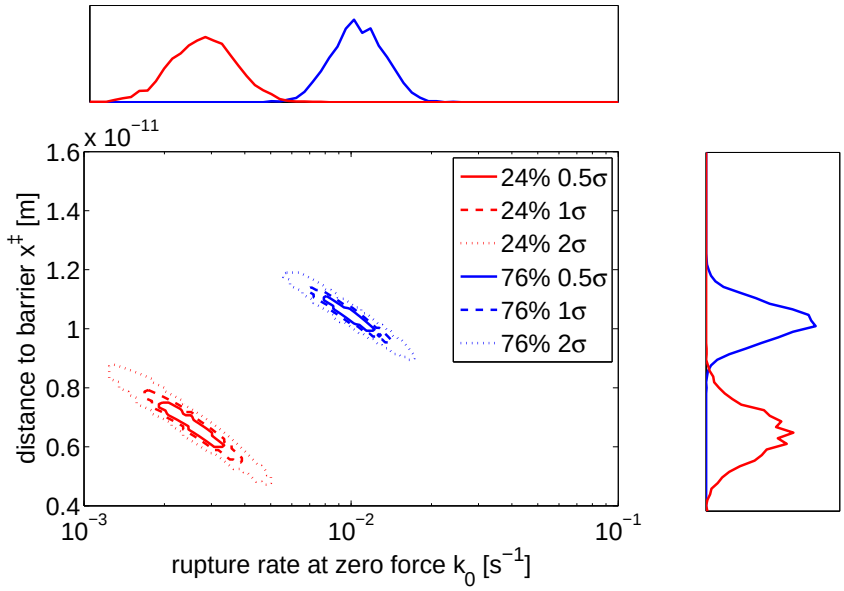


Figure 3.4: Concurrent parameter estimation of constant force data using the nested sampling algorithm of two bond potentials and the Bell model. Results shown as credible areas of the posterior probability density contoured at 0.5σ (solid line), 1σ (dashed line) and 2σ (dotted line). Marginal posterior probability distributions are shown above and to the right.

3.1.4 CUSP Model Results

In contrast to the Bell model in the previous subsection, the more complex CUSP model includes the force dependent change of the distance between the energetic minimum and the energy barrier in the bond potential and is valid for a larger range of forces (see subsection 2.6.2).

The CUSP model was fitted to the constant force data set described in subsection 2.6.1 using the nested sampling algorithm. The parameters of the CUSP model are barrier height, $\Delta G^{\ddagger}$, distance between energetic minimum and the barrier, $x^{\ddagger}$, and the attempt frequency, A. In the analysis result (see Figure 3.5) only $x^{\ddagger}$ is determined with a good precision of ca. 12 $\pm$ 0.9 pm (1σ) for the big population. For the small population in the part of the projected credible volume where $x^{\ddagger}$ is constant, the value of $x^{\ddagger}$ is ca. 7.7 $\pm$ 0.8 pm (1σ). For the $\Delta G^{\ddagger}$ and A the projected credible volumes of the posterior probability densities are correlated for both populations, i.e. an increase in $\Delta G^{\ddagger}$ is compensated by an increase of A and the other way around (Figure 3.5 C and D). Both $\Delta G^{\ddagger}$ and A are only limited at low energies or low attempt frequencies respectively, while the respective upper limit is simply given by the prior boundaries as indicated in figure 3.5. Especially A is poorly defined as the probable values for this parameter are spread out over many orders of magnitude. The projected credible volumes are spread out over a far wider range of the posterior space than the Bell model results (Figure 3.4).

As the prior size for the parameter A was chosen to be very large it is possible to give more precise values for $\Delta G^{\ddagger}$ by evaluating the posterior in a restricted range of values for A. Experimentally determined potentials (see subsection 2.6.4 and Eaton et al.[87]) are found to posses values for the parameter A between 10^3 s^{-1} to 10^7 s^{-1}. If one now considers the projected credible volumes of both rupture populations only in a range of 10^3 to 10^7 s^{-1} for A, then an upper and lower limit based on the maximum dimensions of the 1σ projected credible volume can be given for $\Delta G^{\ddagger}$ (Figure 3.6). For the large population the resulting range for $\Delta G^{\ddagger}$ is $5.3 \cdot 10^{-20}$ to $8.6 \cdot 10^{-20}$ J. For the small population the range for $\Delta G^{\ddagger}$ is $5.1 \cdot 10^{-20}$ to $9.0 \cdot 10^{-20}$ J. The relative location of the projected credible volume of the small population at lower attempt frequencies and higher energies still points to a more stable population and the proportion of 24 $\pm$ 1.5 % (1σ) also agrees with the results from 3.1.2 and 3.1.3.

If one compares the results for the large rupture population from this parameter estimation with the two potentials from the rupture processes of the silyl ester bond[83–85] (see subsection 2.6.4), it becomes apparent that the $x^{\ddagger}$ for both rupture processes (less stable: $12.3 \cdot 10^{-12}$ m; more stable: $12.0 \cdot 10^{-12}$ m) is close to the estimation result of 12 $\pm$ 0.9 pm and that the $\Delta G^{\ddagger}$ for both rupture processes (less stable: $6.31 \cdot 10^{-20}$ J; more stable: $6.14 \cdot 10^{-20}$ J) is within the estimated range of values for $\Delta G^{\ddagger}$ of $5.3 \cdot 10^{-20}$ to $8.6 \cdot 10^{-20}$ J. The estimated value range for A at the position of $\Delta G^{\ddagger} = 6.31 \cdot 10^{-20}$ J is about an order of magnitude smaller than $1.9 \cdot 10^{6}$ s^{-1} for the potential of the less stable rupture process, but for the more stable rupture process the position at $5.5 \cdot 10^{4}$ s^{-1} falls within the area of the 2σ projected credible volume. Thus there is some agreement between the large rupture population found in the constant force experiments and the small rupture population found in the silyl ester bond rupture experiments. The smaller, more stable population on the other hand is far less similar, as only the range of $\Delta G^{\ddagger}$ values matches, while $x^{\ddagger}$ is ca. 60% of the value of the silyl ester bond potentials and the value of A is about two orders of magnitude smaller when compared to the less stable rupture population and A for the more stable rupture population is outside of the 2σ projected credible volume.

3.1.5 Linear-Cubic Model Results

While the CUSP model in the previous subsection is valid for a larger range of forces than the Bell model (see subsection 3.1.3), the range of forces where it is valid is still limited because of the shape of the model and the high-barrier approximation it employs (see subsection 2.6.2). The linear-cubic model (see subsection 2.6.3) does not rely on the high barrier approximation and uses a different potential shape and is thus valid for an even larger range of forces.

The projected credible volume plots from the parameter estimation of the experimental constant force data using the linear-cubic model (Figure 3.7) are similar to the results using the CUSP model (Figure 3.5). $\Delta G^{\ddagger}$ and D show a strong correlation (Figure 3.7 B-D). With this the determination of the height of the energy barrier $\Delta G^{\ddagger}$ and the diffusion coefficient for both the large and the small population only show a well defined lower limit, while the upper limit is determined by the prior size. The diffusion coefficient D is distributed over several orders of magnitude similar to the case

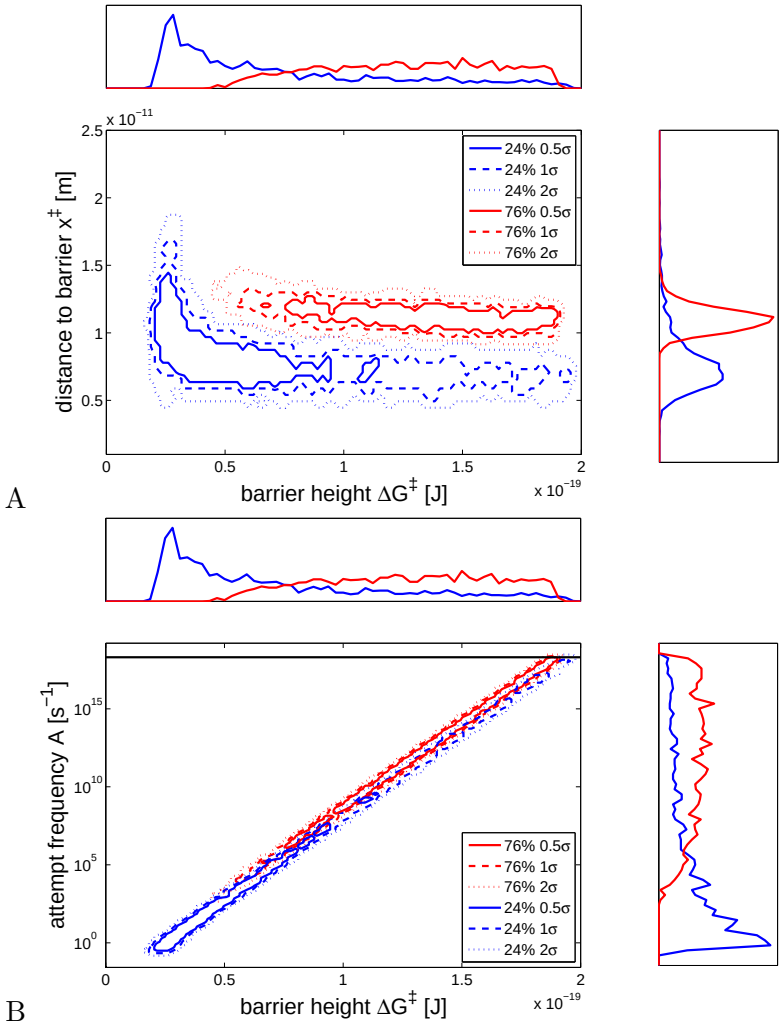


Figure 3.5: Constant force CUSP parameter estimation results. Projected credible volume contour plots of the posterior probability density contoured at 0.5σ (solid line), 1σ (dashed line) and 2σ (dotted line). Marginal posterior probability distributions are shown above and to the right. For the two rupture populations two different bond potentials have been estimated concurrently. (A) Results for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$. (B) Results for the parameters $\Delta G^{\ddagger}$ and A. The horizontal black line marks the upper boundary of the prior for A. Contourlines above the boundary are an artifact of the contouring process.

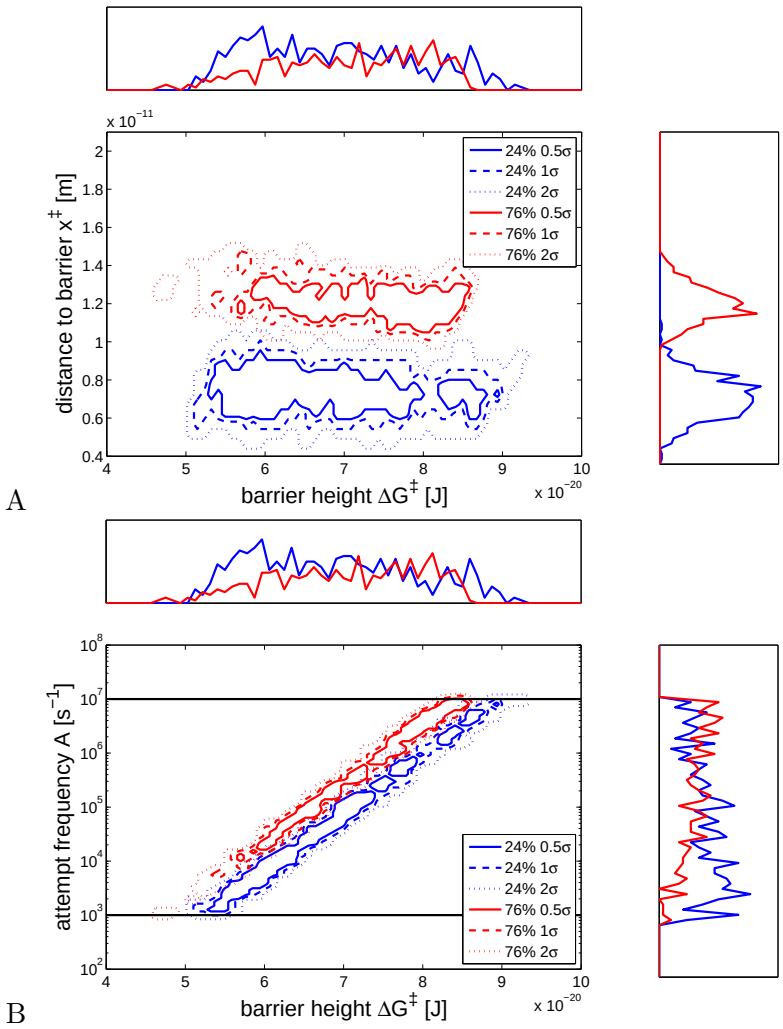


Figure 3.6: Constant force CUSP parameter estimation results with smaller prior boundaries for A. Projected credible volume contour plots of the posterior probability density contoured at 0.5σ (solid line), 1σ (dashed line) and 2σ (dotted line). Marginal posterior probability distributions are shown above and to the right. For the two rupture populations two different bond potentials have been estimated concurrently. The range for the parameter A was restricted to a range of 10^3 s^{-1} to 10^7 s^{-1}. This restriction keeps the values of the parameter A similar to results from experiments on other bond potentials. (A) Results for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$. (B) Results for the parameters $\Delta G^{\ddagger}$ and A. The horizontal black lines mark the boundaries of the prior for A. Contourlines above and below the boundaries are an artifact of the contouring process.

of the attempt frequency A for the CUSP potential. Once again only $x^\ddagger$ is determined with a sufficient precision of ca. 12 ± 0.8 pm (1σ) for the big population. For the small population two branches are visible. The first branch at an energy $\Delta G^\ddagger$ of ca. $2.5 \cdot 10^{-20}$ J, is missing in the result of the CUSP model. In this branch the diffusion coefficient is comparatively stable and an increase of $x^\ddagger$ is compensated by an increase in $\Delta G^\ddagger$. The value of $x^\ddagger$ in the second branch of the posterior density with constant $x^\ddagger$ is ca. 7.5 ± 0.7 pm (1σ). The shapes and relative locations of the posterior densities for both populations are also similar to the CUSP parameter estimation results and the small population is still more stable. The proportion of the smaller and more stable population is 24 ± 1.5 % (1σ) and in agreement with the previous results.

Furthermore if one uses onedimensional diffusion ($\tau = \langle x^2 \rangle / 2D$), $x^\ddagger = 12$ pm and the minimum value for $D \sim 1 \cdot 10^{-19}$ $\mathrm{m^1 s^{-1}}$ in the 1σ credible volume of the large rupture population, an attempt frequency of $A \sim 2 \cdot 10^3$ $\mathrm{s^{-1}}$ can be estimated. (A is calculated from the inverse of the mean liftime.) The minimum value for A in the 1σ credible volume of the parameter estimation using the CUSP model is $\sim 4 \cdot 10^3$ $\mathrm{s^{-1}}$. These values are separated by a factor of two, but with regards to the high uncertainity of the results for D and A in the respective models, their minimum values in the respective 1σ credible volumes are relatively similar.

3.1.6 Model Validity

The three bond rupture models (see section 2.6) used to analyze the constant force measurement data (subsections 3.1.3, 3.1.4 and 3.1.5) are valid for different ranges of forces. The following will now determine which of the three models are valid. This will also reveal if the forces measured in the experiment are mainly low, in this case all three models are valid, or span a large range of forces, in this case only the linear-cubic model is valid.

The critical force in both the CUSP model with high barrier approximation and the linear-cubic model are defined by the parameters $\Delta G^\ddagger$ and $x^\ddagger$ with $F_{crit} = 2\Delta G^\ddagger / x^\ddagger$ for the CUSP model and $F_{crit} = 3\Delta G^\ddagger / 2x^\ddagger$ for the linear-cubic model. Thus the lowest respective critical forces in the credible volumes of the posterior probability density can be found at positions of low $\Delta G^\ddagger$ and high of $x^\ddagger$. Because only $\Delta G^\ddagger$ and $x^\ddagger$ influence F_{crit} and only the minimal possible value of F_{crit} is of interest, just the borders of

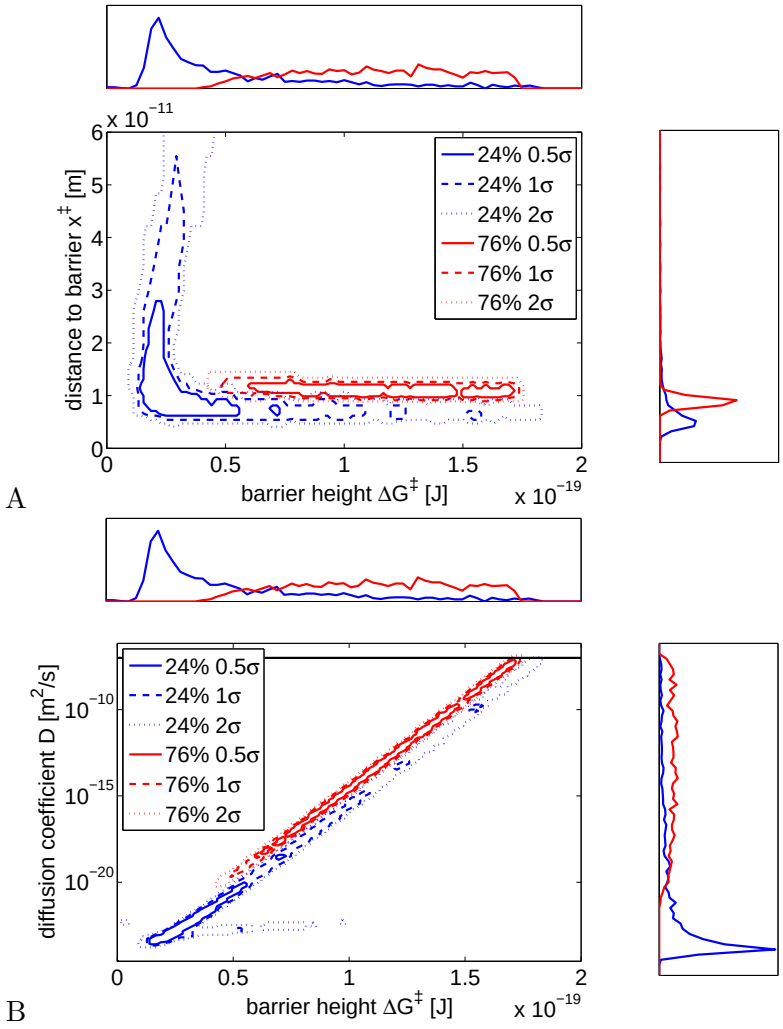


Figure 3.7: Constant force linear-cubic model parameter estimation results. Projected credible volumes of the posterior probability density contoured at 0.5σ, 1σ and 2σ. (A) Contour plots for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$. The projected credible volume for the smaller population of the linear-model is cut off for higher values of $x^{\ddagger}$. (B) Contour plots for the parameters $\Delta G^{\ddagger}$ and D. The horizontal black line marks the upper boundary of the prior for A. Contourlines above the boundary are an artifact of the contouring process.

the projected credible volume area in $\Delta G^{\ddagger}$ and $x^{\ddagger}$ can be searched for the lowest value of F_{crit}. For the CUSP model the position in the 1σ credible volume of the large population yielding the lowest possible F_{crit} results in a F_{crit} of ca. 8.8 nN at a $\Delta G^{\ddagger}$ of $5.3 \cdot 10^{-20}$ J and a $x^{\ddagger}$ of 12 pm. This F_{crit} is comparable to the F_{crit} of the potential for the slower, less frequent (28%) rupture process of the two silyl ester bond potentials[83–85] (see subsection 2.6.4), which is 10.2 nN. For the linear-cubic model the position in the 1σ credible volume of the large population yielding the lowest critical force found results in a F_{crit} of ca. 6.0 nN at a $\Delta G^{\ddagger}$ of $4.8 \cdot 10^{-20}$ J and a $x^{\ddagger}$ of 12 pm. The highest holding forces in the experimental data are at 2.2 nN, which is far below the critical forces of both bond rupture models. So both bond rupture models can be considered to be within their valid range. This also becomes apparent when the results for $\Delta G^{\ddagger}$ and $x^{\ddagger}$ of both models are compared with each other (Figure 3.8) as both models show a good overlap for both populations.

The lowest possible critical forces of 8.8 nN for the CUSP model and 6.0 nN for the linear-cubic model are also similar to theoretical findings which predict maximum rupture forces of 4.4 to 6.6 nN.[19,36,62]

Because experimental data was measured in a range of forces much smaller than the critical force of both the CUSP and the linear-cubic model, the Bell model should be also valid for the measured data. Using $k_0 = A \cdot \exp\left(-\Delta G^{\ddagger}/RT\right)$ the equally weighted samples from the parameter estimation using the CUSP model with high barrier approximation can be converted to the parameters of the BELL model and then used to calculate credible areas of the posterior probability density for the parameters k_0 and $x^{\ddagger}$ (Figure 3.9). Considering the existing overlap between the credible areas of the CUSP and BELL model results, it becomes clear, that the BELL model is indeed also valid for the measured constant force data.

Following from this the results of the BELL model, the CUSP model with high barrier approximation and the linear-cubic model can all be considered valid. Especially for $x^{\ddagger}$ of the large and small rupture populations all three models agree with each other. For $\Delta G^{\ddagger}$ of both the large and the small population there is agreement between the CUSP and the linear-cubic model, but for both rupture populations there is only a well defined lower limit, while the upper limit is given by using probable upper limits for A or D respectively. For A or D except for the correlation with $\Delta G^{\ddagger}$ the precision

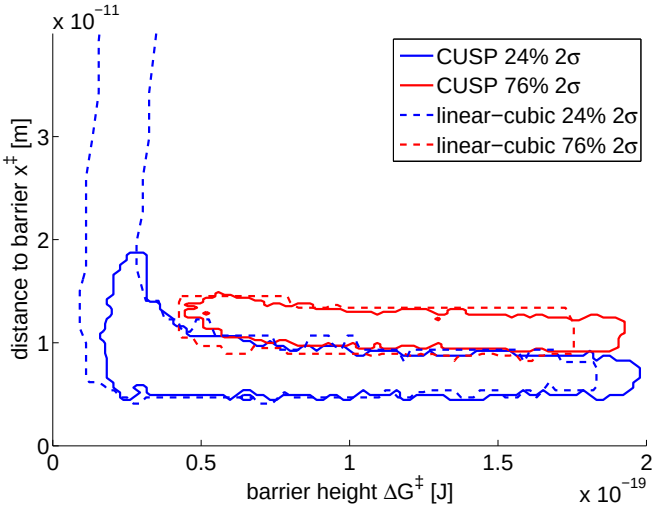


Figure 3.8: Comparison of the analysis results for the constant force data using either the CUSP or the linear-cubic model. Shown are the 2σ projected credible volumes of the posterior probability densities for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$ of the two rupture populations for both models. Both models show a good overlap for the large as well as the small population.

of the parameter estimation is very low. In conclusion the results from the Bell model are the most descriptive for the given experimental data with $x^{\ddagger} = 10.3 \pm 0.6$ pm (1σ) and $k_0 = 10.6 \cdot 10^{-3} \pm 2.4 \cdot 10^{-3}$ s^{-1} (1σ) for the big population and $x^{\ddagger} = 6.6 \pm 0.79$ pm (1σ) and $k_0 = 2.8 \cdot 10^{-3} \pm 0.78 \cdot 10^{-3}$ s^{-1} (1σ) for the small population.

3.2 Factors influencing the parameter estimation quality

For both microscopic bond rupture models (CUSP model in subsection 3.1.4 and linear-cubic model in subsection 3.1.5) the analysis has shown remarkable difficulties in the estimation of several of the bond potential parameters. This is obviously not a desired outcome and it is therefore important to determine the cause of these high uncertainites. Several possibilities come to mind such as the range of the holding forces, the amount of measured data, errors in the determination of the spring constant or the fact that not one, but two rupture populations have to be analyzed at the same time. The

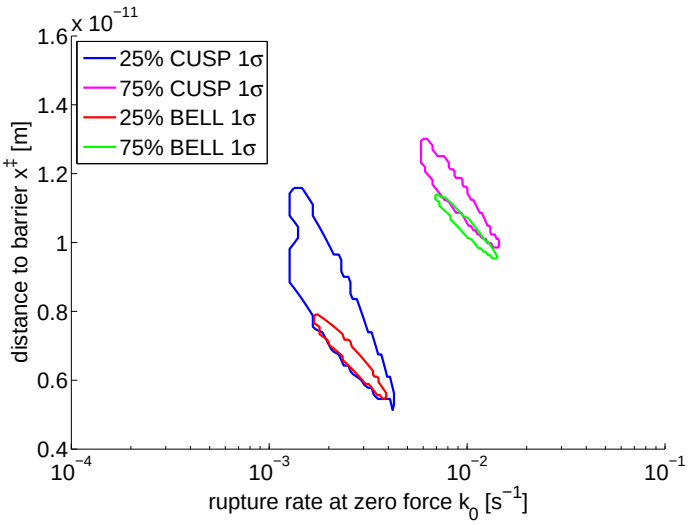


Figure 3.9: Comparison of the analysis results for the constant force data using either the CUSP or the BELL model. Equally weighted samples of the CUSP model have been converted to the parameters of the BELL model using $k_0 = A \cdot \exp\left(-\Delta G^{\ddagger}/RT\right)$. Shown are the 1σ projected credible volumes of the posterior probability densities for the parameters k_0 and $x^{\ddagger}$ of the two rupture populations for both models. Both models show overlap for the large as well as the small population.

influence of those factors will be tested using simulated data. But before a closer look is taken into the influence of these various factors, it has to be investigated first how precise the chosen simulation algorithm can determine the parameters of the bond, and what variations are to be expected from repetitive simulations, to make the separation of the random variation of the simulation and actual influences on the parameter estimation possible.

3.2.1 Simulation and Parameter Estimation

Before investigating how the number of measurements and other experimental parameters or the bond potential itself influence our ability to estimate the bond potential parameters, the behavior of the method to simulate data and how it influences the analysis result had to be examined first.

For that purpose a simulation was performed multiple times with the same bond potential and identical simulation parameters. The bond potential parameters were taken from the experimental analysis of the unfolding process of the fourth filamin domain (ddFLN4) (see subsection 2.6.4) and the force range was chosen to be 0 to 60 % of F_{crit} for the CUSP model with high barrier approximation and 10 to 70 % of F_{crit} for the linear-cubic model respectively. The range for the CUSP model was chosen to be from 0 to 60 % of F_{crit} instead of 10 to 70% to avoid the invalid force range close to the critical force as much as possible.[†] The ranges for both models possess a width of 60% within the force range from zero force to F_{crit}, which should be large enough for both models to allow for a more precise determination of the bond potential parameters than in the experimental data. The amount of data used was 2000 bond rupture events per simulated data set distributed evenly across the given force range. Overall 7 simulations with 2000 rupture events were simulated for both the CUSP and the linear-cubic model.

For the CUSP model (Figure 3.10) the input parameters are located within the 1σ credible volume 7 out of 7 times and within the 0.5σ credible volume 5 out of 7 times. For the linear-cubic model (Figure 3.11) the input parameters are located within the 2σ credible volume of the posterior probability density 7 out of 7 times, within the 1σ credible volume 6 out of 7 times and within the 0.5σ credible volume 5 out of 7 times. Statistically

[†]A repeat of the linear-cubic model simulations in the 0 to 60 % of F_{crit} force range would have been preferable, but was not possible because of time constraints.

the input parameter position is expected to be within the 0.5σ credible volume with a probability of ca. 38.2% and the parameter estimation seems to perform better than expected for both models. For the 1σ credible volume the statistical probability of finding the input paramter within the volume is 68.3% and here both models seem to also perform better than expected. Finally for the 2σ credible volume the probability is 95.4% and thus the fact that the input parameters are located within that volume for both models is expected. But the number of 7 simulations is relatively low and thus overall one can only say that the results of the parameter estimation are likely to behave as expected.

To further investigate the parameter estimation results, the centers of mass of the equally weighted samples from the analysis of the seven simulated data sets for each bond rupture model are used to calculate their mean absolute deviation from the input parameters as well as the deviation of the mean position of the centers of mass from the input parameters. The mean absolute deviation can give insight in the deviations of the parameter estimation results that have to be expected and the deviation of the mean of the centers of mass can be used to find and estimate systematic shifts. Ideally no shifts should exist.

The results for the CUSP model are:

CUSP model	$\Delta G^\ddagger$	$x^\ddagger$	A
mean absolute deviation	0.46%	0.75%	16.4%
deviation of the mean position	+0.40%	−0.36%	+10.5%

For the barrier height, $\Delta G^\ddagger$, and the distance to the barrier, $x^\ddagger$, the mean absolute deviation is below 1%. For the attempt frequency, A, the value is much higher at 16.4%. The absolute value of the deviation of the mean position of the centers of mass for all three parameters is smaller than each respective mean absolute deviation. For $x^\ddagger$ the difference is larger than a factor of 2, for A the difference is 56% and for $\Delta G^\ddagger$ the difference is 15%.

The results for the linear-cubic model are:

linear-cubic model	$\Delta G^\ddagger$	$x^\ddagger$	D
mean absolute deviation	1.54%	2.00%	14.3%
deviation of the mean position	-0.84%	+0.95%	-1.40%

The mean absolute deviations of the barrier height, $\Delta G^\ddagger$, and of the distance to the barrier, $x^\ddagger$, are about three times larger than the respective

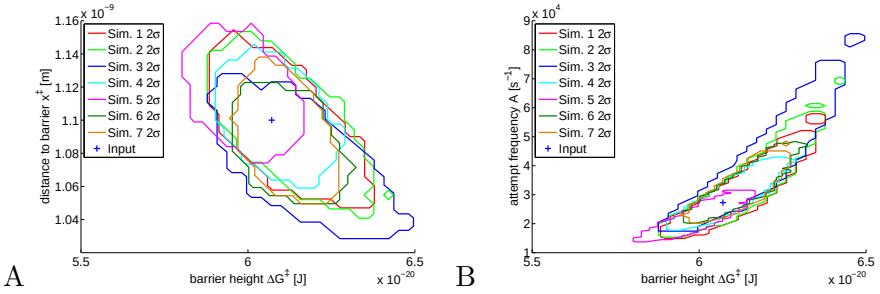


Figure 3.10: CUSP model: Analysis of repeated simulations. Contour plots of projected credible volumes of the posterior densities from the analysis results of seven repeated simulations (red, green, blue, cyan, magenta, dark green, orange) for the CUSP bond rupture model using the same ddFLN4 input potential parameters (blue cross). The contour lines represent the 2σ projected credible volume. The force range was chosen from 0 to 60 % of the critical force. For all seven simulations 2000 bond ruptures were simulated. (A) Results for $\Delta G^{\ddagger}$ and $x^{\ddagger}$ and (B) results for $\Delta G^{\ddagger}$ and A.

values for the CUSP model, but are still relatively small. Similar to A in the CUSP model, the mean absolute deviation of the diffusion coefficient, D, is much larger at 14.3%. The absolute values of the deviation of the mean position of the centers of mass for $\Delta G^{\ddagger}$ and $x^{\ddagger}$ are smaller than the mean absolute deviations by a factor of 2. For D the absolute value of the deviation of the mean position is even ten times smaller than the mean absolute deviation.

As noted further above, the number of 7 parameter estimations for each model is small, and so only rough estimations can be given. A systematic shift cannot be excluded with absolute certainity for any of the parameters of the two bond rupture models, but if it exists it is likely smaller than the random deviation of the results. The mean absolute deviations can be used to determine, if an observed difference between parameter estimation results is caused by actual differences in the analyzed data, such as data from two different rupture processes.

3.2.2 Influence of the Force Range

As measuring at only one holding force only yields a mean rupture rate (see subsection 2.2.1) and thus doesn't allow to uniquely determine all 3 parameters of the bond potential, a range of different forces has to be applied to

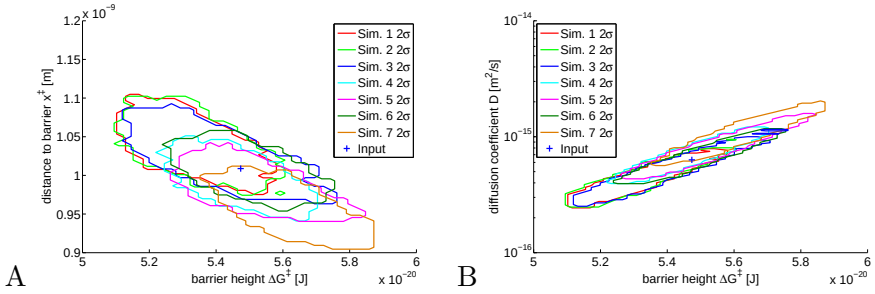


Figure 3.11: Linear-cubic model: Analysis of repeated simulations. Projected credible volumes of the posterior densities from the analysis of seven repeated simulations (red, green, blue, cyan, magenta, dark green, orange) using the ddFLN4 potential parameters (blue cross) for the linear-cubic bond rupture model contoured at 2σ. The force range was chosen from 10 to 70 % of the critical force. For all seven simulations 2000 bond ruptures were simulated. (A) Results for $\Delta G^{\ddagger}$ and $x^{\ddagger}$ and (B) results for $\Delta G^{\ddagger}$ and D.

the bond of interest. Measuring a given force range at higher or lower forces might also influence the quality of the data analysis. To be able to judge the importance of the range of holding forces, four different measurements were simulated using both the CUSP and the linear-cubic model with the ddFLN4 potential as input: for each simulation 2000 rupture events were simulated and the force ranges were chosen as 10-30 %, 30-50 %, 50-70 % and 10-70 % of the F_{crit} of each respective model.

In the following sizes of a 2σ credible volume are defined as the dimensions of a boundary box that is drawn around the respective 2σ credible volume. E.g. if a size of 1 is given for a parameter y, then 1 is the difference in y between the two points in the 2σ credible volume with the smallest and largest values for y.

For the CUSP bond rupture model a simulation of 2000 rupture events in a range of holding forces between 10 and 70% of F_{crit} leads to a parameter estimation with high accuracy: For $\Delta G^{\ddagger}$ the 2σ credible volume has a size of only $2.9 \cdot 10^{-21}$ J, for $x^{\ddagger}$ the size of the credible volume is $1.0 \cdot 10^{-10}$ m and for A it is $2.5 \cdot 10^{4}$ s^{-1} (Figure 3.12 A and B). For 2000 rupture events the three smaller force ranges 10-30 %, 30-50 % and 50-70 % of F_{crit} the accuracy of the parameter estimation decreases drastically. For $\Delta G^{\ddagger}$ the 2σ credible volume increases by a factor of ca. 40 for all three force ranges. One

has to take into account that the upper limit for $\Delta G^{\ddagger}$ for all force ranges is given by the upper boundary of the prior with regards to the parameter A. For the parameter $x^{\ddagger}$ there are differences in the precision for the different force ranges. The size of the credible volume for $x^{\ddagger}$ is the smallest for the force range 10-30 % with $5.1 \cdot 10^{-10}$ m, which is a size roughly five times larger than for the large force range. For the highest force range (50-70 %) the size increases to $1.1 \cdot 10^{-9}$ m, which is ten times larger than the size of the projected credible volume for the large force range. So for force ranges of the same size an increase in precision by a factor of two could be observed if the measurement is performed at low forces. For A the size of the credible volume for the 10-70 % is 13 to 15 orders of magnitude smaller than the size of the credible volumes for the three smaller force ranges (10-30 %, 30-50 % and 50-70 %). The sizes are only limited at low values of A ($3.8 \cdot 10^{2}$ s^{-1} for 10-30 %, $2.4 \cdot 10^{3}$ s^{-1} for 30-50 % and $1.3 \cdot 10^{4}$ s^{-1} for 50-70 %), while the upper limit is given by the upper prior boundary for A. While the shapes of the credible volumes for all three smaller force ranges vary, they intersect where the real parameter position of the potential is and the intersection of their volumes encloses the credible volume from the parameter estimation of the large force range. Which is expected, since the parameters of the bond potentials should be found by the estimations of all four froce ranges, albeit with a lower precision for the three smaller force ranges. The variation of the shapes is an indication that information about the three parameters can be found at varying precisions for the different force ranges. As the three bond parameters influence the force dependent bond rupture rate in specific different ways, it should be possible to use the shape of a credible volume to judge if the measured holding forces are low or high relative to the critical force.

For a simulation of 2000 rupture events using the linear-cubic bond rupture model very similar behavior to the CUSP model was observed, when the data was analyzed. For the large force range between 10 and 70% of F_{crit} the size of the 2σ credible volume is small for all three parameters: For $\Delta G^{\ddagger}$ the size is $6.1 \cdot 10^{-21}$ J, for $x^{\ddagger}$ it is $1.3 \cdot 10^{-10}$ m and for D it is $7.0 \cdot 10^{-16}$ $m^1 s^{-1}$ (Figure 3.12 C and D). For the smaller force ranges the precision of the parameter estimation decreases drastically. For $\Delta G^{\ddagger}$ the size of the credible volumes increases by a factor of ca. 20 for all force ranges. An upper limit for $\Delta G^{\ddagger}$ is present though, because of the prior

voundaries analogous to the results of the CUSP model analysis. For $x^\ddagger$ the lowest of the three force ranges (10-30 %) once again results in the most precise determination of the parameter with a size of the credible volume of $3.1 \cdot 10^{-10}$ m which is between two and three times larger than the size of the credible volume for the large force range (10-70 %). The highest force range (50-70 %) is the least precise range with regards to the parameter $x^\ddagger$ with a size of the 2σ credible volume of $1.5 \cdot 10^{-9}$ m which is about five times larger than the result of the 10-70 % force range and more than ten times larger than the result for the 10-70 % force range. For the parameter D the size of the 2σ credible volume for the large force range (10-70 % of F_{crit}) is 11 to 13 orders of magnitudes smaller than the size of the credible volumes for the three small force ranges (10-30 %, 30-50 %, 50-70 % of F_{crit}). Lower limits for D ($6.4 \cdot 10^{-18}$ $\mathrm{m^1 s^{-1}}$ for 10-30 %, $4.4 \cdot 10^{-17}$ $\mathrm{m^1 s^{-1}}$ for 30-50 % and $1.8 \cdot 10^{-16}$ $\mathrm{m^1 s^{-1}}$ for 50-70 %) can be found and upper limits are given by the prior boundaries for D. The shapes of the credible volumes from the three small ranges vary similar to the CUSP model results and intersect, where the true parameter position is located as expected.

Overall both bond rupture models have very similar results. Only for the parameter $\Delta G^\ddagger$ an increased precision for the CUSP model over the linear-cubic model could be argued, but to determine this with certainty a larger amount of simulations is necessary.

The range of forces is an important factor for the quality of the parameter estimation. With identical amounts of rupture events the precision of the parameter estimation can be improved drastically with an increase of the force range by a factor of 3: for the parameter $\Delta G^\ddagger$ the size of the credible volume is improved by a factor of 20 to 40, for the parameter $x^\ddagger$ it is improved by a factor 2 to 10 and for the parameters A and D the size can be reduced from many orders of a magnitude to within an order of magnitude.

3.2.3 Influence of the Amount of Data

For the quality of the parameter estimation, i.e. the accuracy and precision, the amount of data is naturally also of importance. For this purpose constant force measurements with varying amounts of data were simulated: The CUSP and linear-cubic bond rupture model and the experimentally determined parameters of the ddFLN4 potential (see subsection 2.6.4) were used over a force range of 10 to 70 % with respect to the critical force. 1000,

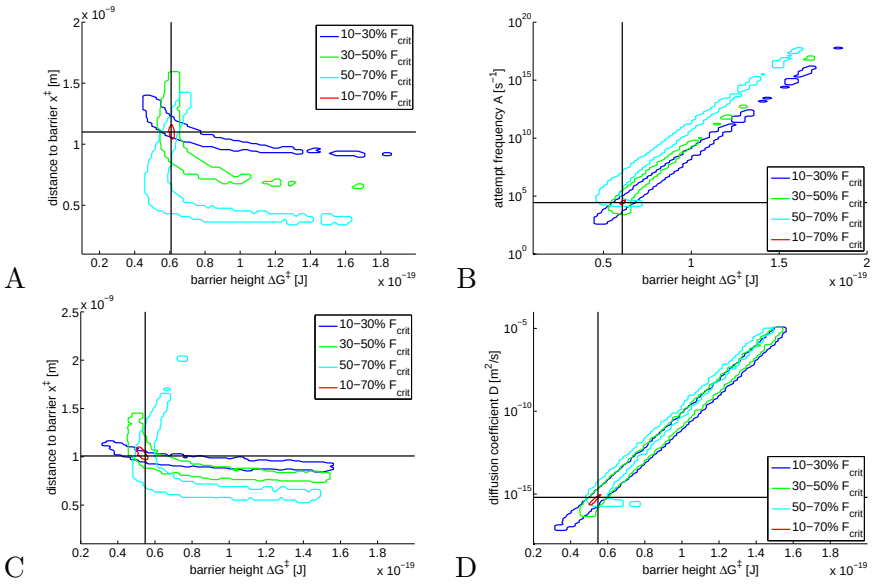


Figure 3.12: Parameter estimation results of analyzing simulated data distributed across four different force ranges. Projected credible volumes of posterior densities contoured at 2σ. The experimentally determined ddFLN4 potential parameters for the CUSP and the linear-cubic models were used for four simulations respectively and 2000 rupture events were simulated for each simulation. The force ranges are: 10-30 % F_{crit} (blue), 30-50 % F_{crit} (green), 50-70 % F_{crit} (cyan) and 10-70 % F_{crit} (red) with Fcrit being 110 pN for the CUSP model and 82 pN for the linear-cubic-model. (A) Analysis results for the CUSP model showing the projection for $\Delta G^{\ddagger}$ and $x^{\ddagger}$. (B) Analysis results for the CUSP model showing the projection for $\Delta G^{\ddagger}$ and A. (C) Analysis results for the linear-cubic model showing the projection for $\Delta G^{\ddagger}$ and $x^{\ddagger}$. (D) Analysis results for the linear-cubic model showing the projection for $\Delta G^{\ddagger}$ and D.

2000 and 4000 bond ruptures are compared. As in subsection 3.2.2 the sizes of a 2σ credible volume are defined as the dimensions of a boundary box that is drawn around the respective 2σ credible volume.

When using the CUSP bond rupture model for different numbers of rupture events (Figure 3.13), the results show that the parameter estimation can be improved with the amount of data. The size of the 2σ credible volume for the parameter $\Delta G^{\ddagger}$ shrinks from $4.0 \cdot 10^{-21}$ J for 1000 rupture events over $3.1 \cdot 10^{-21}$ J for 2000 rupture events to $2.5 \cdot 10^{-21}$ J for 4000 rupture events. For $\Delta G^{\ddagger}$ the credible volume size shrinks from 1000 to 2000 rupture events by 22% and from 2000 to 4000 rupture events by 18%. For the parameter $x^{\ddagger}$ the sizes of the 2σ credible volume are the following: $1.7 \cdot 10^{-10}$ m for 1000 rupture events, $1.2 \cdot 10^{-10}$ m for 2000 rupture events and $0.8 \cdot 10^{-10}$ m for 4000 rupture events. Thus the size of the credible volume for the parameter $x^{\ddagger}$ shrinks by 30% from 1000 to 2000 rupture events and also by 30% from 2000 to 4000 rupture events. For the parameter A the resulting sizes were $3.6 \cdot 10^{4}$ s^{-1} for 1000 rupture events, $2.6 \cdot 10^{4}$ s^{-1} for 2000 rupture events and $1.8 \cdot 10^{4}$ s^{-1} for 4000 rupture events. The size of the credible volume for A shrinks by 27% from 1000 to 2000 rupture events and by 29% from 2000 to 4000 rupture events. With each doubling of the amount of data the parameter estimation appears to improve by a factor between 18 and 30% for all three parameters.

Using the linear-cubic model for different numbers of rupture events (Figure 3.14) leads to the following results: For the parameter $\Delta G^{\ddagger}$ the size of the 2σ credible volume changes from $17.8 \cdot 10^{-21}$ J for 1000 rupture events over $6.3 \cdot 10^{-21}$ J for 2000 rupture events to $5.7 \cdot 10^{-21}$ J for 4000 rupture events. This is a reduction in size of 65% from 1000 to 2000 rupture events and 9% from 2000 to 4000 rupture events. The size of the credible volume for the parameter $x^{\ddagger}$ changes in the following manner: $2.2 \cdot 10^{-10}$ m for 1000 rupture events, $1.4 \cdot 10^{-10}$ m for 2000 rupture events and $1.0 \cdot 10^{-10}$ m for 4000 rupture events. This corresponds to a shrinking of the credible volume size by 36% from 1000 to 2000 rupture events and 25% from 2000 to 4000 rupture events. For the parameter D a big improvement of the credible volume size can be observed from 1000 to 2000 rupture events, but the results from the simulations with 2000 and 4000 rupture events are practically of the same size: $1.5 \cdot 10^{-14}$ $m^{1}s^{-1}$ for 1000 rupture events, $7.8 \cdot 10^{-16}$ $m^{1}s^{-1}$ for 2000 rupture events and $8.4 \cdot 10^{-16}$ $m^{1}s^{-1}$ for 4000 rupture events. From 1000 to

2000 rupture events the credible volume shrinks by 94.9% and from 1000 to 4000 rupture events the factor is 94.5%. 4000 rupture events are even slightly worse than 2000 rupture events for the determination of D. This difference is likely negligible though. The results from the 2000 and 4000 rupture event simulations thus show no improvement of the parameter estimation for $\Delta G^{\ddagger}$ and D. The visible improvement of the parameter $x^{\ddagger}$ is similar to the behavior in the CUSP model as it roughly improves by the same factor with each doubling of the amount of data (64% from 1000 to 2000 rupture events and 75% from 2000 to 4000 rupture events). The described results change less consistently than the results for the CUSP model. One reason is that the parameter estimation for the simulation with 1000 rupture events is of exceptional low precision. Similar outliers have been observed ocassionaly and a repeated simulation and analysis of the 1000 rupture events should lead to more consistent results.

While an increase of the amount of data improves the parameter estimation, compared to an increase of the range of holding forces, the impact is much more limited. Following from this it appears that a large amount of information can already be extracted with relatively few rupture events, if they are distributed across a large force range. In short a large force range at a low precision likely yields more information than a small force range with high precision. The improvement in the quality of the parameter estimation for both the CUSP and the linear-cubic model are similar. This similar behavior suggests, that in the exmined force ranges the behavior of both models is similar and likely describes the behavior of bond ruptures similarly well.

3.2.4 Influence of the Spring Constant Determination Error

In order to properly characterize the energy landscape of the bond of interest, the force measurements (see section 2.2) have to be precise. And for this to be the case the determination of the spring constant of the pulling system has to be precise as well. For atomic force microscopes the error in the calibration of the spring constant of the cantilever is typically in the range of 10 to 20 %, if a common calibration technique is used (see subsection 2.2.3).[52,56,88]

To examine the influence of an erroneous spring constant, bond rupture simulations which recorded both the actual and an erroneous force

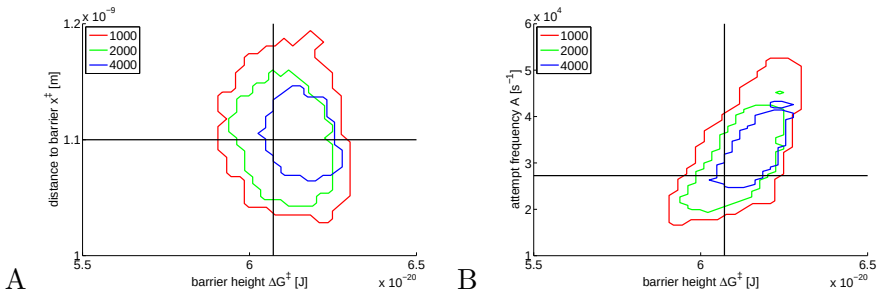


Figure 3.13: CUSP model: Influence of the amount of data. Projected credible volumes of the posterior density of the results of three simulations using the ddFLN4 CUSP bond rupture potential (intersection of the black lines) and three different amounts of data with contour lines at 2σ. The amount of simulated ruptures ranges from 1000 (red) over 2000 (green) to 4000 (blue) and the force range was chosen from 10 to 70 % of the critical force of 110 pN. (A) Results for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$. (B) Results for the parameters $\Delta G^{\ddagger}$ and D.

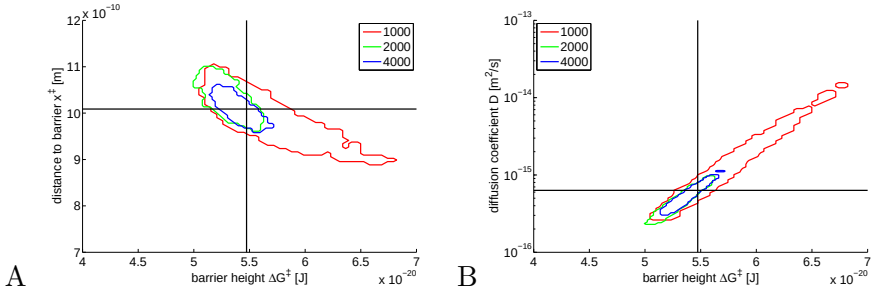


Figure 3.14: Linear-cubic model: Influence of the amount of data. Analysis results of three different sized data sets depicted as projected credible volumes contoured at 2σ. The three simulations used the linear-cubic model and the respective ddFLN4 potential parameters as input (intersection of the two black lines) and a force range between 10-70 % of the critical force Fcrit 82 pN. The amount of bond ruptures for each simulation were: 1000 (red), 2000 (green) and 4000 (blue). (A) Plot for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$. (B) Plot for the parameters $\Delta G^{\ddagger}$ and D.

for each bond rupture were performed. The simulations were done using the parameters from the experimentally determined ddFLN4 potential (see subsection 2.6.4) for both the CUSP and the linear-cubic model (see subsection 2.6.4). The force range was chosen to be between 0 to 60 % of the critical force for the CUSP model and between 10 and 70 % of the critical force for the linear-cubic model and the amount of simulated rupture events was 2000 per measurement set. The force range for the CUSP model was shifted to lower forces, to avoid simulating forces in the invalid range of the CUSP model, which is likely to occur for an upper force limit of 70 % of the critical force in combination with a spring constant error of 20 %. Data for four different cases was simulated:

- a variable spring constant error for each bond rupture, which is normal distributed with a width of 10 %
- a variable spring constant error for each bond rupture, which is normal distributed with a width of 20 %
- a static error of a 20 % too low spring constant
- a static error of a 20 % too high spring constant
- a static error of a 20 % too high spring constant

In normal use the number of bond ruptures that are measured, before the cantilever has to be calibrated or exchanged for a new one varies. The different cases were used to allow the estimation of the influence of the calibration error for various types of normal use. For a large amount of calibrations the influence will be closer to the normal distributed error and for very few calibrations the influence will be closer to either a constant too high or a constant too low spring constant. The two different widths of the distributions at 10 % and 20 % can be used to assess how improvements of the cantilever calibration will improve the parameter estimation overall.

For the case of a constantly changing spring constant error one would naively expect the errors to cancel out and result in a more spread out posterior density. But this is not the case. With varying errors of the spring constants a larger than determined range of forces is measured, while the rupture times are still recorded for the original force range. A force range is assumed which is smaller than the actually measured force range. All

simulations show shifts of the center of mass of the posterior probability density to erroneous spring constants and also changes of the uncertainty of the parameter estimation. To compare those changes in the uncertainty of the parameter estimation the change of the standard deviation is given. E.g. if the standard deviation of the parameter estimation with accurate spring constants is 10% and the standard deviation for the estimation with erroneous spring constants is 15%, then this would result in a change of the standard deviation of +50%.

For the three simulations using the CUSP model with high-barrier approximation and a normal distributed error of the spring constant of 10% (Figure 3.15) the shifts of the posterior density mass centers and changes in uncertainty were:

Sim. #	$\Delta G^{\ddagger}$ shift	$\Delta G^{\ddagger}$ uncert.	$x^{\ddagger}$ shift	$x^{\ddagger}$ uncert.	A shift	A uncert.
1	−3.63%	-3.73%	−0.47%	+5.69%	−43.7%	−9.34%
2	+0.33%	+27.8%	−4.48%	+6.87%	+16.0%	+29.8%
3	−4.46%	+1.79%	−1.90%	+6.68%	−48.9%	+1.09%

For the three simulations using the CUSP model with high barrier approximation and a normal distributed error of the spring constant of 20% (Figure 3.16) the shifts of the posterior density mass centers and changes in uncertainty were:

Sim. #	$\Delta G^{\ddagger}$ shift	$\Delta G^{\ddagger}$ uncert.	$x^{\ddagger}$ shift	$x^{\ddagger}$ uncert.	A shift	A uncert.
1	−20.6%	−9.20%	−4.91%	+30.9%	−95.4%	−25.5%
2	−14.1%	+31.2%	−6.98%	+43.1%	−87.1%	+21.9%
3	−8.69%	+147%	−17.2%	+70.0%	−62.2%	+187%

For the three simulations using the linear-cubic model and a normal distributed error of the spring constant of 10% (Figure 3.17) the shifts of the posterior density mass centers and changes in uncertainty were:

Sim. #	$\Delta G^{\ddagger}$ shift	$\Delta G^{\ddagger}$ uncert.	$x^{\ddagger}$ shift	$x^{\ddagger}$ uncert.	D shift	D uncert.
1	−2.68%	+35.9%	−2.49%	+26.3%	−27.1%	+41.8%
2	−2.44%	+39.6%	−2.81%	+20.6%	−24.6%	+41.2%
3	+1.87%	+72.1%	−3.71%	+23.7%	+31.6%	+191%

For the three simulations using the linear-cubic model and a normal distributed error of the spring constant of 20% (Figure 3.18) the shifts of the posterior density mass centers and changes in uncertainty were:

Sim. #	$\Delta G^{\ddagger}$ shift	$\Delta G^{\ddagger}$ uncert.	$x^{\ddagger}$ shift	$x^{\ddagger}$ uncert.	D shift	D uncert.
1	−17.8%	+73.1%	−10.8%	+78.0%	−89.2%	+61.3%
2	−18.2%	+83.2%	−15.3%	+73.3%	−89.4%	+70.1%
3	+5.97%	+430%	−22.3%	+86.8%	+356%	+858%

Because of the low number of simulations for each case and due to the fact that the results of the parameter estimations exhibit considerable variation, information had to be derived with caution. Because of that the usage of mean values was also avoided.

For the uncertainty of the parameter estimation, as in the change of the standard deviation from accurate to erroneous spring constants, two trends are clear for both the CUSP and the linear-cubic bond rupture model: the variation of the spring constant error will lead to an increase in uncertainty. This behavior is not surpirsing as the uncertainty should increase if erroneous but random information is superimposed into the data.

For the shifts of the posterior density mass centers there were two trends which became clearly apparent for both the CUSP and the linear-cubic model: parameters are likely to be underestimated, if variations of spring constants are assumed. The reason for a general underestimation of all three parameters ($\Delta G^{\ddagger}$, $x^{\ddagger}$ and A for the CUSP model and $\Delta G^{\ddagger}$, $x^{\ddagger}$ and D for the linear-cubic model) was not readily apparent, especially since the underestimation of $x^{\ddagger}$, A and D leads to the bonds appearing more stable, while the underestimation of $\Delta G^{\ddagger}$ leads to the bond potential appearing less stable. To better understand this one has to take into account how the variation of the potential parameters also influences the shape of the force dependent function for the rupture rate. (See also 2.6.5.) A simulation using experimentally determined parameters of the ddFLN4 potential, the high barrier approximation CUSP model and a normal distributed error of the cantilever spring constant of 10% was performed to simulate rupture events, which were randomly distributed across a force range of 10 to 60% of the critical force. Then the mean rupture rates for accurate and erroneous spring constants were calculated from the simulated rupture events (Figure 3.19).

As one can see the effect of a normal distributed spring constant error is not only an apparently more stable potential with lower rupture rates, but also a change in the shape of the function occurs. While a simple shift to lower rupture rates could be achieved by an underestimated A or D respectively, a change of curvature makes the change of two or three parameters necessary.

While the previous discussion has shown that a normal distributed error of the cantilever spring constant lets the bond appear to be stronger than it actually is, the effect of a constant error in the cantilever spring constant determination can let the analyzed bond potential appear either weaker or stronger than is actually the case as will be shown now. If the erroneous spring constant is higher than the actual spring constant the bond appears to be stronger, because holding forces are assumed to be higher than they actually are, and vice versa. When analyzing the distribution of rupture events using the CUSP model both $\Delta G^{\ddagger}$ and A stay unchanged and only $x^{\ddagger}$ is noticeably shifted to lower values by 19% (Figure 3.20). The opposite is the case for a spring constant error where the bond appears to be weaker and $x^{\ddagger}$ is shifted to higher values by 18%. If the simulated rupture events are analyzed using the linear-cubic model the bond appears stronger through both a shift to lower values of $x^{\ddagger}$ (21%) and D (42%), while there is practically no change to $\Delta G^{\ddagger}$ (Figure 3.21). Just as with the CUSP model the opposite is the case for a spring constant error where the bond appears to be weaker with a shift to higher values of 21% for $x^{\ddagger}$ and 47% for D and again no change for $\Delta G^{\ddagger}$. In comparison to normal distributed multiple spring constant errors, both the size and the shape of the 1σ posterior density volumes do not change, which is expected because of the common, equally strong shift for all rupture events. For both the CUSP and the linear-cubic model the overall uncertainty for the parameter estimation is higher, when the cantilever spring constant has been determined 20 % too high, since the actually measured range of holding forces is smaller than assumed.

The effects of multiple normal distributed errors of the spring constant and of a constant spring constant error are different and the kind of expected under- and overestimations depend on the number of times the spring constant had to be estimated for the measured data. For measurements with only one cantilever (or very few cantilevers) the estimation of the parameter $x^{\ddagger}$ can be conservatively expected to roughly fall between ± 20 % of the real value for both models. For D in the linear-cubic model the deviation can be

in the range of around ±45 %. And no change is expected for the parameter $\Delta G^{\ddagger}$. For measurements with many different cantilevers the parameters can be expected to be underestimated. For both potentials shifts to lower values have to be expected. $\Delta G^{\ddagger}$ can be shifted up to 20%, $x^{\ddagger}$ up to 22% and A or D up to 95%. The magnitudes of the given shifts might seem dramatic at first glance, but they have to be seen as estimates for the worst case. In a usual experimental situation where many cantilevers are used and the spring constant error can be kept close to 10% the shift that has to be expected is reduced to a few percentage points for $\Delta G^{\ddagger}$ and $x^{\ddagger}$ and about 45% for A and 30% for D. Furthermore one should note that more advanced calibration methods than the one used in this work (see subsection 2.2.3) will allow a more precise determination of the cantilever spring constant.

3.2.5 Concurrent Parameter Estimation

Lastly I wanted to test how the concurrent parameter estimation of two different potentials as well as the relative probability increases the uncertainty of the analysis. This is necessary because two rupture populations were observed in the constant force experiment(see subsection 3.1.2). In order to test this, data from constant force experiments with rupture events originating from two different potentials were simulated and analyzed both by the CUSP model (Figure 3.22) and by the linear-cubic model (Figure 3.23) were simulated. Both a concurrent and a separate analysis of the respective data sets were performed. The chosen parameters by the two bond potentials were: $\Delta G_1^{\ddagger} = 5.51 \cdot 10^{-20}$ J, $x_1^{\ddagger} = 11.9$ pm, $A_1 = 8000$ s^{-1}, $\Delta G_2^{\ddagger} = 3.25 \cdot 10^{-20}$ J, $x_2^{\ddagger} = 7.83$ pm, $A_2 = 7.4$ s^{-1} for the CUSP potential and $\Delta G_1^{\ddagger} = 5.51 \cdot 10^{-20}$ J, $x_1^{\ddagger} = 11.9$ pm, $D_1 = 7.51 \cdot 10^{-20}$ m^{2}s^{-1}, $\Delta G_2^{\ddagger} = 3.25 \cdot 10^{-20}$ J, $x_2^{\ddagger} = 7.83$ pm, $D_2 = 6.07 \cdot 10^{-23}$ m^{2}s^{-1} for the linear-cubic potential (marked by the intersections of the black lines in the Figures 3.22 and 3.23). The potential parameters were chosen so that F_{crit} of the large and F_{crit} of the small population are fairly close to each other so that the force range covered was similar for both bonds. Since the amount of data and the proportion of both populations in the experimental data are known, 1500 rupture events for population 1 and 500 bond rupture events for population 2 were simulated for both models. The force range for the simulated data was set to lie within 10 % and 70 % of F_{crit} of the large population and rupture events were evenly distributed over the entire force range. As expected the results

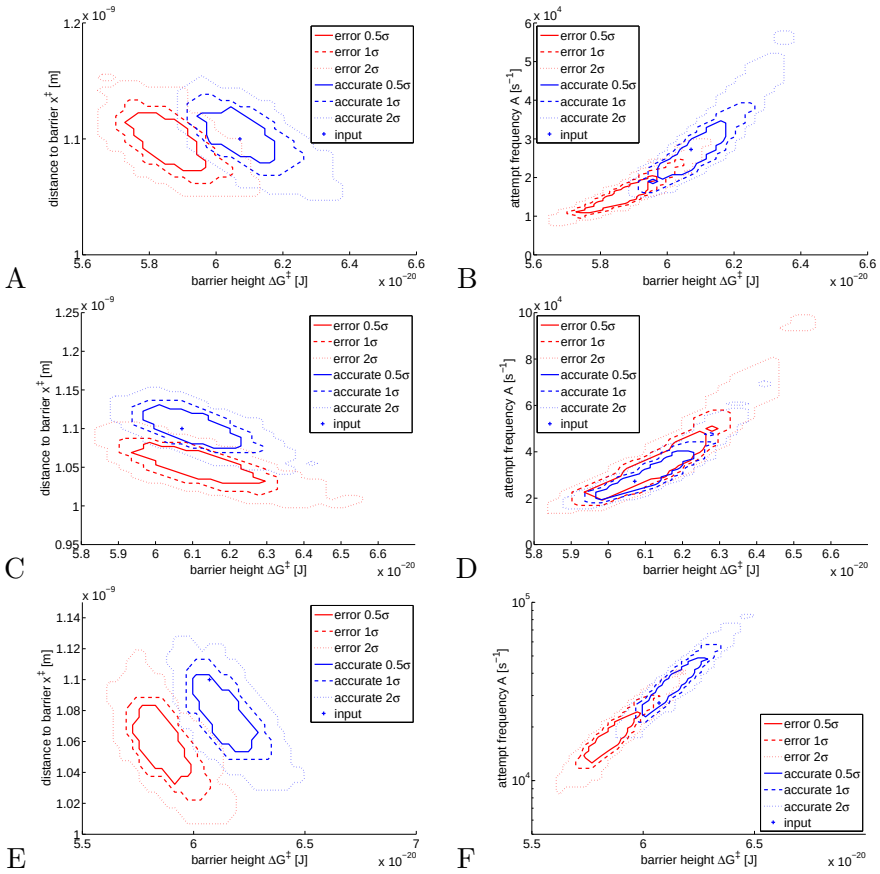


Figure 3.15: CUSP model: distributed spring constant error at 10 %. Results of three simulations (Simulation 1: A and B, simulation 2: C and D, simulation3: E and F) using the ddFLN4 CUSP bond rupture potential (intersection of the black lines) shown as projected credible volumes of the posterior probability distribution contoured at 0.5σ, 1σ and 2σ. Results are shown for analysis using accurate values for the cantilever spring constant (green) and for spring constants with a 10 % normal distributed error (red).

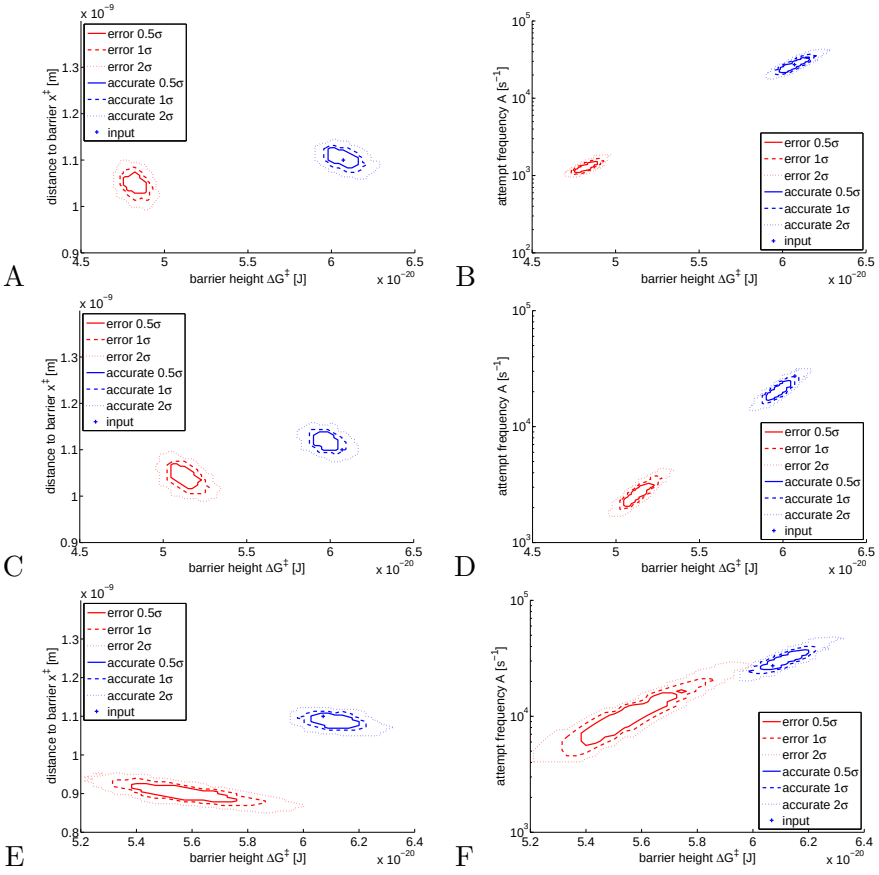


Figure 3.16: CUSP model: distributed spring constant error at 20 %. Results of three simulations (Simulation 1: A and B, simulation 2: C and D, simulation3: E and F) using the ddFLN4 CUSP bond rupture potential (intersection of the black lines) shown as projected credible volumes of the posterior probability density contoured at 0.5σ, 1σ and 2σ. Results are shown for analysis using accurate values for the cantilever spring constant (green) and for spring constants with a 20 % normal distributed error (red).

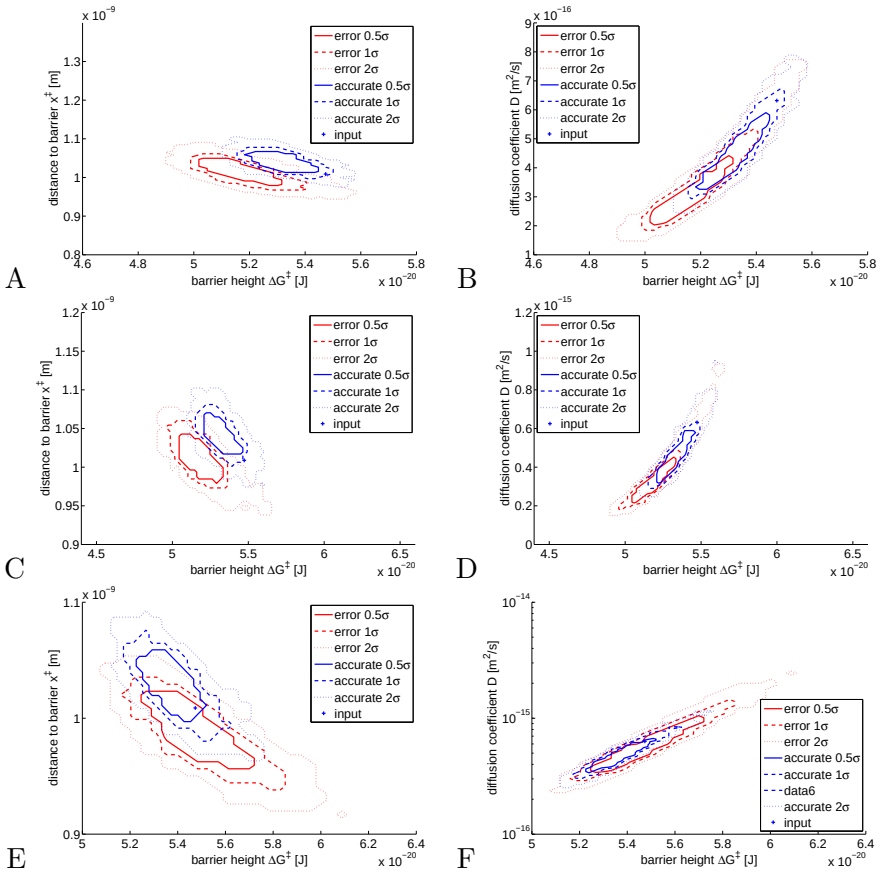


Figure 3.17: Linear-cubic model: distributed spring constant error at 10 %. Projected credible volumes of the posterior probability densities of the results of three simulations (Simulation 1: A and B, simulation 2: C and D, simulation3: E and F) using the ddFLN4 linear-cubic bond rupture potential (intersection of the black lines) contoured at 0.5σ, 1σ and 2σ. Results are shown for analysis using accurate values for the cantilever spring constant (green) and for spring constants with a 10 % normal distributed error (red).

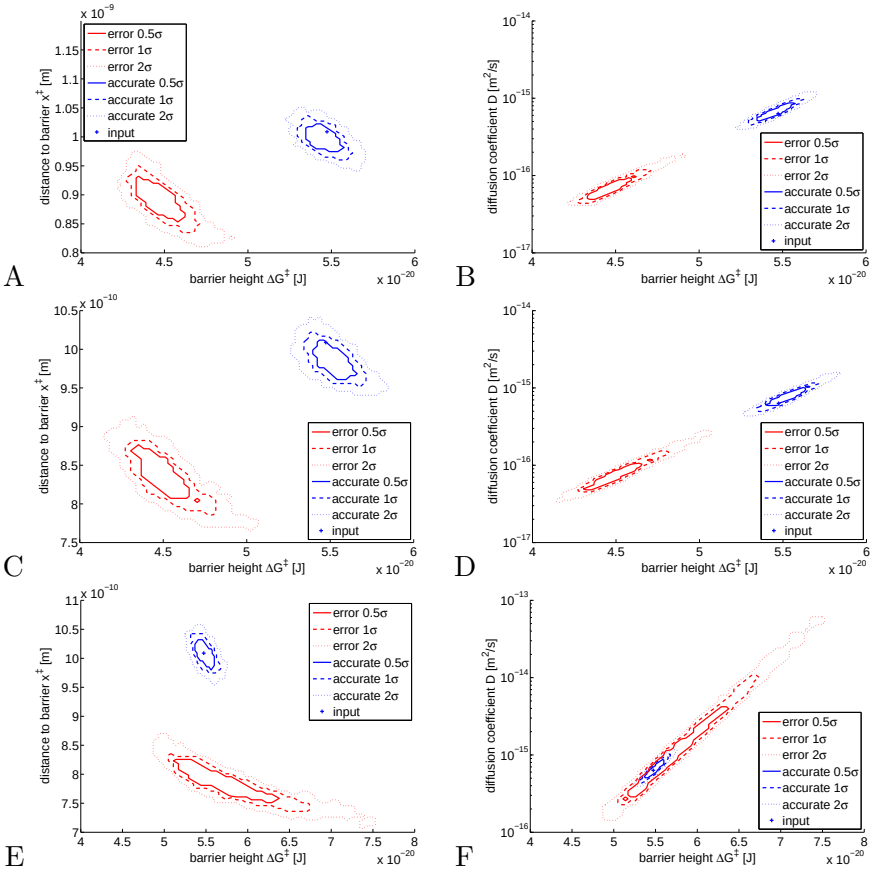


Figure 3.18: Linear-cubic model: distributed spring constant error at 20 %. Projected credible volumes of the posterior probability densities of three simulations (Simulation 1: A and B, simulation 2: C and D, simulation3: E and F) using the ddFLN4 linear-cubic bond rupture potential (intersection of the black lines) contoured at 0.5σ, 1σ and 2σ. Results are shown for analysis using accurate values for the cantilever spring constant (green) and for spring constants with a 20 % normal distributed error (red).

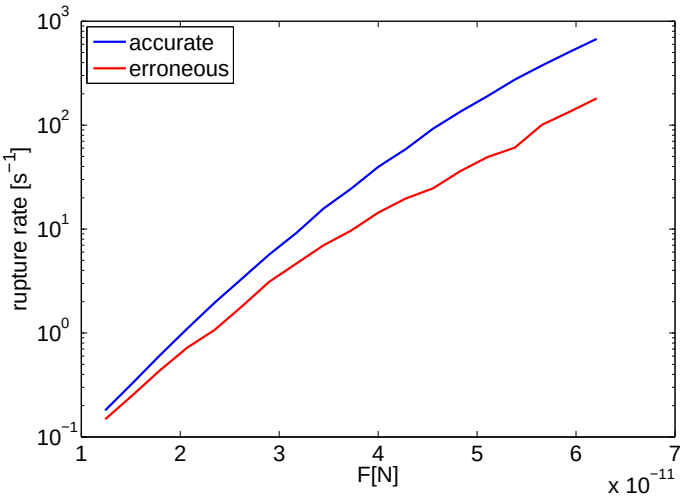


Figure 3.19: Effect of a normal distributed spring constant error on the observed force dependent rupture rates (red) compared to the actually occuring rupture rates (green). The observed bond potential appears more stable with lower rupture rates. The observed change of the function shape is not achievable through a shift of the rupture rates alone. Based on simulations of bond ruptures with holding forces randomly distributed across the shown force range using the CUSP bond rupture model. Mean times to rupture for a certain force interval have been used to determine the rupture rates.

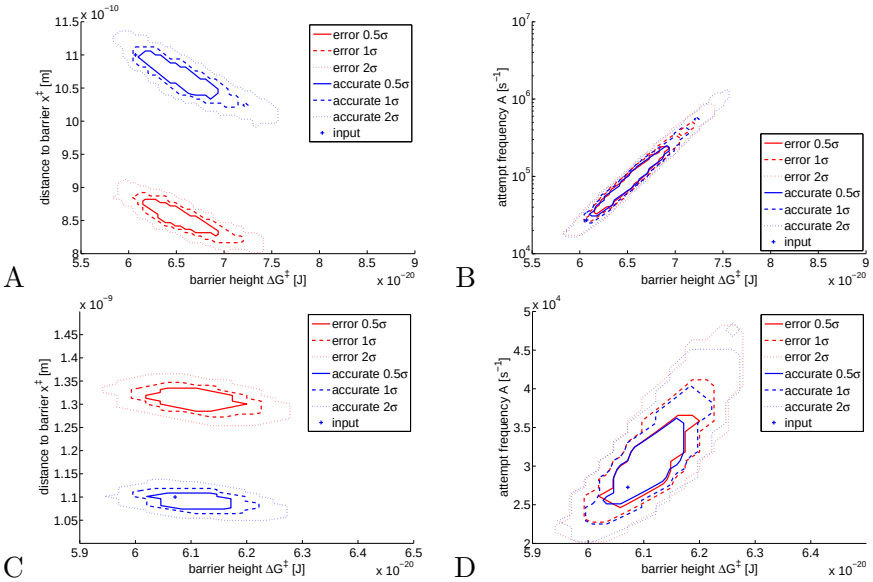


Figure 3.20: CUSP model: constant spring constant error. Analysis results of two simulations using the ddFLN4 CUSP bond rupture potential shown as projected credible volumes contoured at 0.5σ, 1σ and 2σ. The simulated data was analyzed using either an accurate value for the spring constant (green) or an erroneous value for the spring constant (red). In one case the erroneous spring constant value is 20 % too high (A and B) and in the other the value is 20 % too low (C and D).

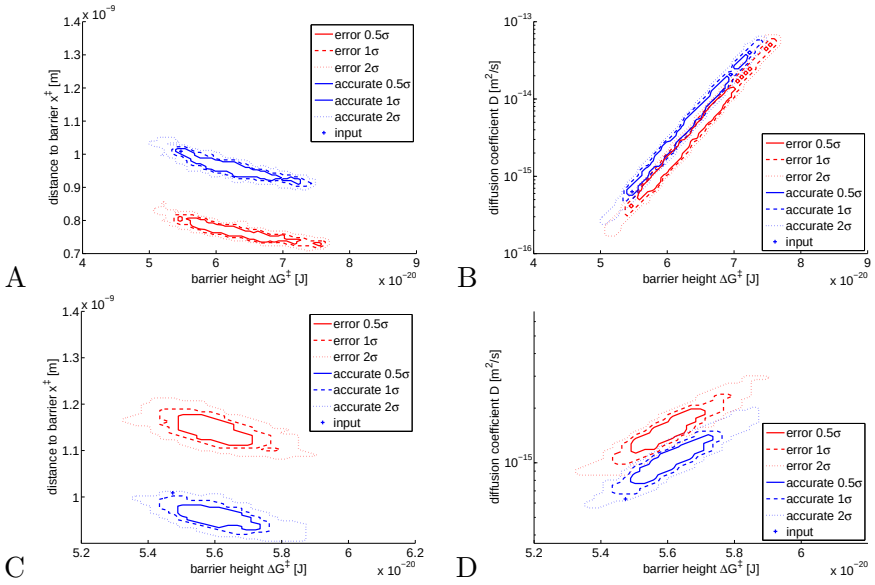


Figure 3.21: Linear-cubic model: constant spring constant error. Projected credible volumes of the posterior probability contoured at 0.5σ, 1σ and 2σ showing the analysis results of two simulations using the ddFLN4 linear-cubic bond rupture potential. The simulated data was analyzed using either an accurate value for the spring constant (green) or an erroneous value for the spring constant (red). In one case the erroneous spring constant value is 20 % too high (A and B) and in the other the value is 20 % too low (C and D).

of the concurrent analysis leads to generally larger uncertainty and thus increased sizes of the 1σ credible volume when compared to the separate analysis (The sizes of a 1σ credible volume are defined as the dimensions of a boundary box that is drawn around the respective 1σ credible volume. See subsection 3.2.2):

CUSP model	$\Delta G^{\ddagger}$	$x^{\ddagger}$	A
big pop.	+0.4%	+16%	+6.9%
small pop.	-5.1%	+11%	+16%
linear-cubic model	$\Delta G^{\ddagger}$	$x^{\ddagger}$	D
big pop.	+25%	+10%	+36%
small pop.	+4.2%	+9.1%	+9.3%

In all cases the larger uncertainties were still much smaller than those obtained in the parameter estimation of the experimental data. So under ideal conditions the concurrent parameter estimation of two bond populations and their proportion is possible and the quality of the parameter estimation should not suffer too much.

3.2.6 Summary of the Influencing Factors

Multiple reasons for the difficulty of obtaining good parameter estimations for the CUSP and the linear-cubic model are possible and have been examined in the previous subsections. They include an insufficient force range, an insufficient amount of data, the influence of an erroneous spring constant and the difficulty of estimating two rupture populations at once. Simulations indicate, that for an amount of 2000 bond ruptures, which is the number of ruptures that were acquired in the constant force measurements, a holding force range of 60 % of F_{crit} is necessary to achieve a good parameter estimation (see subsections 3.2.2 and 3.2.3).

When comparing to the measured data it becomes clear that this condition hadn't been fulfilled (see subsection 3.1.6). The lowest critical force in the 1σ posterior volume for the CUSP model can be estimated to be 8.8 nN and for the linear-cubic model to be 6.0 nN. In constant force experiments a force range between 0.8 and 2.2 nN amounting to $\sim$9% and 25% of F_{crit} for the CUSP potential and $\sim$13% and 37% of F_{crit} for the linear-cubic potential respectively was used. Thus the range of holding forces was much smaller

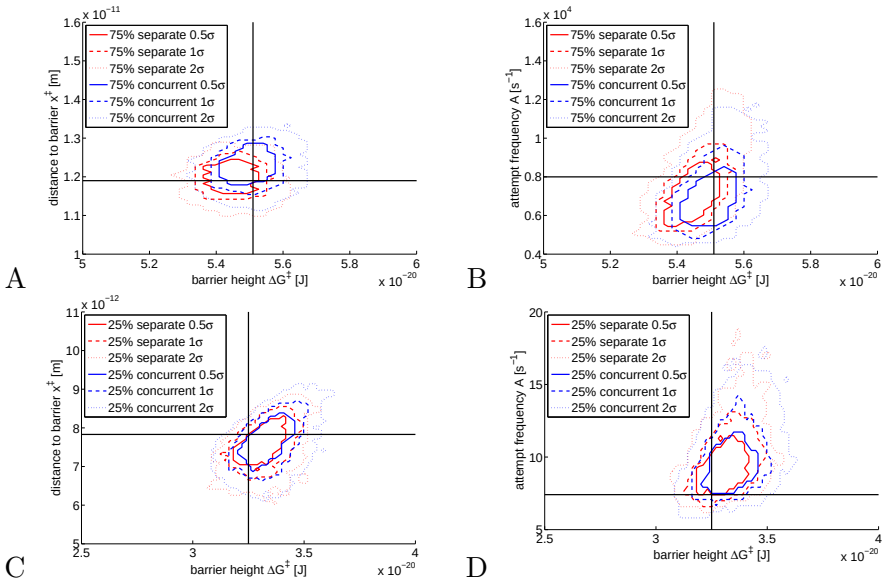


Figure 3.22: CUSP model: concurrent parameter estimation test. Comparison of the separate and concurrent analysis of simulated data with a large population (A+B) and a small population (C+D). The data was simulated for the CUSP potential with parameters identified by constant force experiments. Shown are projected credible volume of the psoterior probability density contoured at 0.5σ, 1σ and 2σ. A+C) Results for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$. B+D) Contour plots of the projected credible volumes for the parameters $\Delta G^{\ddagger}$ and A. In all cases the credible volumes for both kinds of analysis contain the input parameters of the simulated bond rupture potentials.

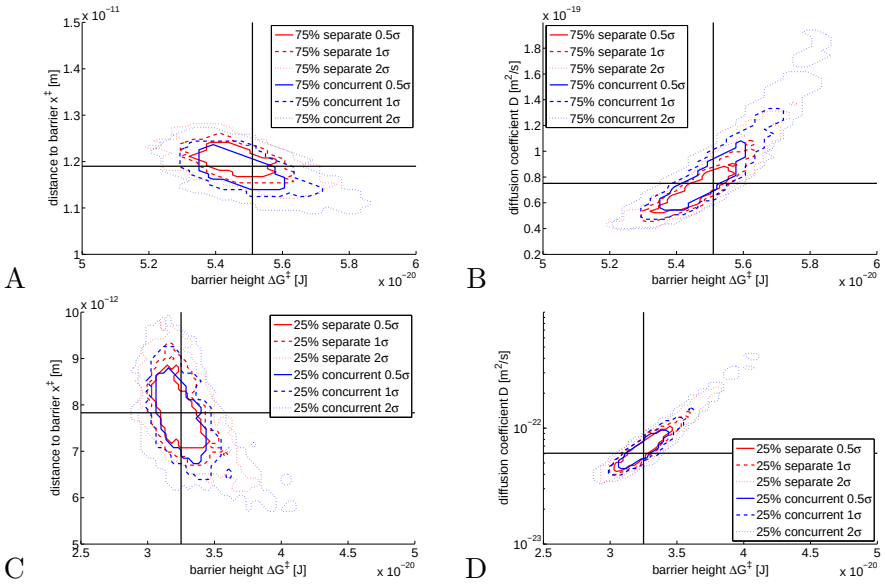


Figure 3.23: Linear-cubic model: concurrent parameter estimation test. Comparison of the separate and concurrent analysis of simulated data with a large population (A+B) and a small population (C+D). The data was simulated for the linear-cubic potential with parameters identified by constant force experiments. Shown are projected credible volumes of the posterior probability densities contoured at 0.5σ, 1σ and 2σ. A+C) Results for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$. B+D) Results for the parameters $\Delta G^{\ddagger}$ and D. In all cases the credible volumes for both kinds of analysis contain the input parameters of the simulated bond rupture potentials.

than necessary for an accurate parameter estimation given the amount of measured data.

Another factor for the unsatisfactory parameter estimation is an erroneous determination of the cantilever spring constant. While simulations could show that the effects of an erroneously determined spring constant are not small, they do not expand the prior density about many orders of magnitude for neither the parameter A nor the parameter D, so this factor is less likely to be responsible for the large uncertainties which were encountered in the analysis of the experimental data. Also simulations could show that for the amount of measured data a concurrent determination of two rupture populations is possible as well, in fact the presence of two populations only had a minor effect on the accuracy of the analysis.

Thus the insufficient range of holding forces remains as the most dominant factor for the observed large uncertainty of the experimental parameter estimation. Expansion to lower forces is possible, but will not yield a large expansion of the force range since the experimental force range is already very low even for the lowest possible critical forces. Measurement of bond ruptures at higher forces should be possible, however from 1 nN to 2 nN the rupture rates readily increased by an order of magnitude. For the experimental constant force data a further rate increase by about an order of magnitude is expected for each nanonewton by which the force range is extended. In order to observe these faster rupture rates the feedback mechanism has to be adapted. How much of this can be improved before technical limitations are reached is not clear. At least an increase to 3 nN should be possible, however. Increasing the amount of data is straightforward, but it is not clear how much data is actually necessary to achieve a good parameter estimation. Considering that the influence of the force range on the quality of the analysis is bigger than the amount of data, future measurements will have to primarily focus on the expansion of the accessible force range by implementing a faster feedback on modern AFMs which are equipped with faster electronics.

3.2.7 Optimizing Constant Force Measurements

To achieve a good determination of the parameters of a bond both the amount of measured data and the force range over which this data is recorded are of importance. Since the amount of data has a smaller effect on the

improvement of the analysis results than the force range, it is difficult to compensate for a small force range with a higher amount of data. Thus the force range for an experiment should be chosen as wide as possible. Difficulties in expanding this range can be encountered both on the lower and the higher force limit. At low forces the amount of bond ruptures which can be measured over time can be severely reduced. And at high forces the capability of the measurement instruments are the main obstacle, since they have to provide a short pull phase compared to the mean rupture rate while still being precise enough to not prematurely rupture the bond and keep the target force constant. While at high forces and accordingly high rupture rates the technical limitations of the instrument can be hard to overcome, at low forces increasing the measurement times and parallelizing those measurements are relatively straightforward options to compensate for low rupture rates. The measurement precision of the rupture times might also be of concern for high rupture rates. The rupture rates for various covalent and non-covalent bond potentials (see subsection 2.6.4) across the necessary range of holding forces, all span four to five orders of magnitudes which already might prove difficult to be covered by one instrument: For the non-covalent experimentally determined parameters of the ddFLN4 potential[78,81] the rupture rate at 10% of F_{crit} is $1.3 \cdot 10^{-1}$ s^{-1} and at 70% of F_{crit} the rupture rate is $2.1 \cdot 10^{3}$ s^{-1}. For the non-covalent experimentally determined parameters of the titin potantial[27,82] the respective rupture rates are $4.0 \cdot 10^{-3}$ s^{-1} and $2.4 \cdot 10^{3}$ s^{-1}. And for the experimentally determined parameters of the silyl ester bond potential[83–85] the rates are 5.6 s^{-1} and $1.4 \cdot 10^{5}$ s^{-1}. Since higher values of $\Delta G^{\ddagger}$ also lead to wider ranges of rupture rates, bond potentials with higher $\Delta G^{\ddagger}$ can be considered as more difficult to be examined than other potentials Especially for potentials with high values of $x^{\ddagger}$, the precision of the measured forces is also of importance, since for potentials with higher $x^{\ddagger}$ a certain change of the holding force has a large effect on the force range. The third parameter A or D respectively can hinder the bond parameter estimation if they are either too high or low. Finally the influence of the spring constant error should be kept in mind and how those errors are distributed over multiple separate measurements.

Since most of the time the parameters of the bond are the subject of examination and thus changing them is not possible, the highest priority remains accessing a force range as wide as possible, which might make even

the usage of different instruments necessary. Several thousand bond ruptures should be sufficient, if the accessible force range is large enough. Otherwise the amount of necessary measurements can be significantly higher.

As mentioned above, measuring bond rupture rates across four to five orders of magnitude might prove difficult experimentally. An alternative and more managable approach is the variation of the temperature as bond rupture behavior is also temperature dependent. Measurements at different temperatures between 295 and 320 K and over a small range of forces between 0.8 and 1.2 pN have been used for example, to determine the bond potential parameters for two rupture processes of the silyl ester bond.[85]

3.3 Constant Speed Measurements

In a previous work constant speed bond rupture data of PDMS polymers was analyzed by fitting the Bell model, the CUSP model with high barrier approximation and the the linear-cubic model with high barrier approximation to a histogram of experimentally determined rupture forces, while using a mean value for the loading rate.[38,41] It was found that the CUSP model fited the data well, while the Bell model and the linear-cubic model both failed to explain the data in the high force regime. Also for the CUSP model the rupture rate at zero force k_0 was found to be several orders of magnitude higher than expected.[38,41] Using the nested sampling algorithm for Bayesian parameter estimation (see section 2.4) allows to analyze the data without having to rely on a mean value for the loading rate which will be discussed in the following section.

3.3.1 BELL Model Results

The experimental data on single molecule bond rupture of PDMS molecules using constant speed published earlier[38,41] were analyzed using the Bell model (see subsection 2.6.1) and the nested sampling algorithm (see section 2.5). In previous works[38,41] the histogram fit results for the Bell model were $k_0 = 2.3\ \mathrm{s}^{-1}$ and $x^\ddagger = 10$ pm. The result for the parameter estimation using the nested sampling algorithm were $k_0 = 2.7 \pm 0.30\ \mathrm{s}^{-1}$ (1σ) and $x^\ddagger = 8.7 \pm 0.30$ pm (1σ) (See figure 3.24).[‡] The distance between minimum and

[‡]Values and deviations given according to histograms of marginalized equally weighted samples.

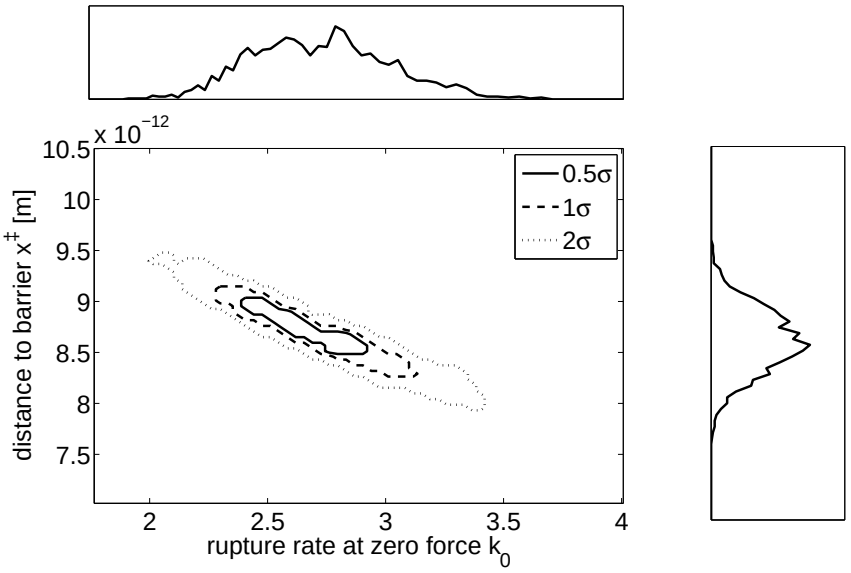


Figure 3.24: Concurrent parameter estimation of constant speed data using the nested sampling algorithm and the Bell model. Results shown as credible areas of the posterior probability density contoured at 0.5σ (solid line), 1σ (dashed line) and 2σ (dotted line). Marginal posterior probability distributions are shown above and to the right.

transition state for both methods are close to each other while the rupture rate for the nested sampling parameter analysis is slower by about 17 %. Even though the result of the new analysis has a slower rupture rate and is likely to be more trustworthy, because it uses the correct loading rate for each rupture event, the rupture rate is still much higher than expected. Further information about the validity of the result can be found below in 3.3.4.

3.3.2 CUSP Model Results

When analyzing the data with the CUSP model (see subsection 2.6.2), the parameters from the fit to the histogram[38,41] as well as the nested sampler (see section 2.5) were even more similar: The histogram fit in the earlier published works[38,41] yielded $\Delta G^{\ddagger} = 3.04 \cdot 10^{-20}$ J, $x^{\ddagger} = 21$ pm and $A = 0.86 \cdot 10^{3}$ s^{-1} and the Bayesian parameter estimation results were $\Delta G^{\ddagger} = 3.0{\cdot}10^{-20} \pm 9.0{\cdot}10^{-22}$ J (1σ), $x^{\ddagger} = 19 \pm 0.7$ pm (1σ) and $A = 1.0{\cdot}10^{3} \pm 96\,\mathrm{s}^{-1}$

(1σ) (See figure 3.25).[§] While the similarity of the results shows that the usage of a mean loading rate in the previous analysis was sufficient, the problem of too high rupture rates remains, since k_0 calculated from the nested sampling result is 0.60 s^-1, which is not slower but even faster than the histogram fit result of 0.52 s^{-1}. It should be noted that the Bell model result for $x^\ddagger$ is higher than the CUSP fit result and the opposite is the case for the rupture rates at zero force. Overall this leads to bond potentials of similar stability in the observed range of loading rates.

3.3.3 Linear-Cubic Model Results

The nested sampling results (see section 2.5) for the linear-cubic model (see subsection 2.6.3) are $3.10 \cdot 10^{-20} \pm 2.85 \cdot 10^{-21}$ J (1σ) for $\Delta G^\ddagger$, 34 ± 6.8 pm (1σ) for $x^\ddagger$ (See figure 3.25).[¶] These parameter estimation results are different to the results from the histogram fit[38,41] which are $2.6 \cdot 10^{-20}$ J for $\Delta G^\ddagger$ and 16 pm for $x^\ddagger$. But this difference is expected, as the histogram fit uses the linear-cubic model with high barrier approximation and the Baeysian parameter estimation uses the linear-cubic model without high barrier approximation. $\Delta G^\ddagger$ is less well defined than in the CUSP model results from the Baeysian parameter estimation, but the results are similar as $\Delta G^\ddagger$ for the linear-cubic model is about 3% higher than the CUSP model result. $x^\ddagger$ on the other hand is larger than the CUSP result by about a third of its value. The diffusion coefficient D has a value of $2.5 \cdot 10^{-20} \pm 2.7 \cdot 10^{-21}$ $m^2 s^{-1}$ (1σ) (See figure 3.25). Since this is not a common parameter between the high-barrier approximation CUSP and the linear-cubic model a direct comparison is not possible. But once again (see subsection 3.1.5) onedimensional diffusion ($\tau = \langle x^2 \rangle / 2D$) can be used to estimate an attempt frequency A from $x^\ddagger$ and D. The result of this estimation is ~ 40 s^{-1}, which is 25 times smaller than the result for the CUSP model. And even though the estimated parameter ranges based on the dimensions of the projected credible volumes for $\Delta G^\ddagger$ and $x^\ddagger$ in both models overlap, the projected credible volumes themselves (Figure 3.25 A) show no overlap. This is in stark contrast to the application of the high-barrier approximation CUSP and the linear-

[§] Values and deviations given according to histograms of marginalized equally weighted samples.

[¶] Values and deviations given according to histograms of marginalized equally weighted samples.

cubic model to the analysis of the constant force data, where the overlap between the results is very large. The following will explain in more detail why there is no overlap.

3.3.4 Model Validity

Just as with the results from the constant force experiments (see subsection 3.1.6), the validity of the three bond rupture models has to be assessed. When the results of the CUSP and the linear-cubic bond rupture models are compared (Figure 3.25), it appears that the projected credible volumes lead to value ranges of the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$ which are similar, but the projected credible volumes themselves show no overlap. The expectation is that the models should roughly agree similar to the parameter estimation results from the constant force data. An explanation for this not being the case is that one of the two models is not in its valid range for the measured data. The critical force for the high barrier approximation CUSP model is 3.1 nN and for the linear-cubic model F_{crit} is 1.4 nN. Even at 3.1 nN bond ruptures can be found in the experimental data and the peak of the population is roughly located at 1.5 nN (Figure 3.26).

It now has to be decided, which of the results of the models is trustworthy. The CUSP model has limitations for high forces. Those limitations also appear well below the critical force. The effects of those limitations manifest themselves as negative likelihood values at high experimental forces and the areas of the likelihood functions for the experimental loading rates being ca. 3 % smaller than 1 (Figure 3.26). Following from this the results from the CUSP model cannot be used as considerable amounts of data are not in the valid range of the model. The BELL model cannot be used either, because the forces where the BELL model is valid are even lower than for the CUSP model.

For the linear-cubic model there are no prinicipal limitations for the full force range from 0 to 100% of the critical force, but the critical force for this model itself is so low, that most of the data and even the peak of the measured rupture force data is in a region where the model predicts the barrier to have vanished. For a bond rupture model this is clearly a region of unphysical behavior. Thus also the linear-cubic model cannot be considered valid for the experimental constant force data.

All three models have to be considered not valid for the experimental

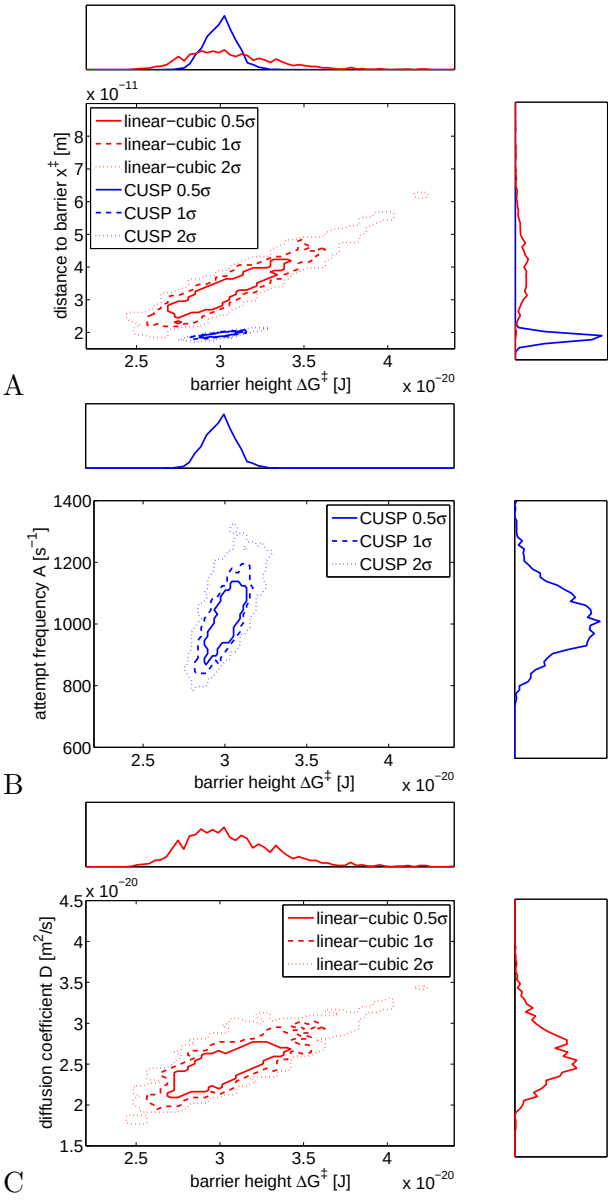


Figure 3.25: Constant speed parameter estimation results for the CUSP model and the linear-cubic model. 2σ projected credible volume plots contoured at 0.5σ, 1σ and 2σ from the analysis of constant speed data using the CUSP (green) and the linear-cubic (red) model. A) Comparison of the results from the CUSP and linear-cubic models for the parameters $\Delta G^{\ddagger}$ and $x^{\ddagger}$. B) CUSP model analysis result for the parameters $\Delta G^{\ddagger}$ and A. C) linear-cubic model analysis result for the parameters $\Delta G^{\ddagger}$ and D. The projected credible volume for the linear-cubic model is bigger than for the CUSP model, but nonetheless the projected credible volumes are clearly separated from each other with no overlap.

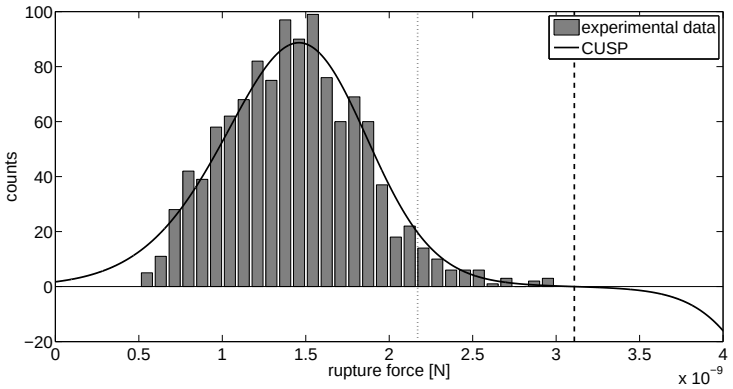


Figure 3.26: Probability density function fit of the CUSP model at a mean loadingrate $Kv = 35.3$ nN/s of the experimental constant force data. The vertical dashed line marks F_{crit} of the CUSP model fit. The vertical dotted line marks 70% of F_{crit}. Considerable amounts of the probability density are close to the critical force and the area of the function is smaller than 1. The experimental data shown in this figure has been cut off above 3.0 pN to avoid negative likelihood values during the parameter estimation. For the full data refer to figure 3.28 in 3.4.

constant speed data. But out of the three models, the CUSP model has the largest amount of data within it's valid range. Because of this the results from the CUSP model will be used for the purpose of comparison in the following sections.

3.4 Discussion and Comparison of the Parameter Estimation Results

Previously experimental constant force data was analyzed with three different models (see subsections 3.1.3, 3.1.4 and 3.1.5) and experimental constant speed data from previous works[38,41] was analyzed with these three models as well (see subsections 3.3.1, 3.3.2 and 3.3.3). In the following it will be evaluated how well the results from the constant force and the constant speed experiment agree with each other, with experimental results from other groups and with theoretical works.

In the first subsection the various potentials from the analysis of the constant force and constant speed data are used to simulate constant force

and constant speed data. The simulated data will then be compared to the experimental data to determine how well the experimentally determined potentials agree with the experimental data from both experiments. Furthermore experimentally determined silyl ester Si-O bond potentials[83–85] are used to also simulate constant force and constant speed data to serve as an external reference.

In the next subsection a closer look will be taken into the reasons why at least three different rupture processes were found in the constant force and constant speed experiments.

In the third and fourth subsection the results are compared to theoretical findings by Frank et al.[19,36,62,63] as well as experimental results by Clausen-Schaumann et al.[83–85]

3.4.1 Comparison of the Parameter Estimation Results with the Experimental Data

Several different bond potentials were found after analyzing the constant force data measured in this work (see subsections 3.1.3, 3.1.4 and 3.1.5) and the constant speed data (see subsections 3.3.1, 3.3.2 and 3.3.3) from previous experiments.[38,41] To see how well those potentials fit to the experimental data, constant speed and constant force data for the different bond potentials were simulated (Figure 3.27 and 3.28).

For the simulation of constant force and constant speed data the following parameters and models were used:

CF, Bell	k_0	$x^\ddagger$	
big pop.	$10.6 \cdot 10^{-3}$ s^{-1}	$10.3 \cdot 10^{-12}$ m	
small pop.	$2.8 \cdot 10^{-3}$ s^{-1}	$6.6 \cdot 10^{-12}$ m	
CF, CUSP	$\Delta G^\ddagger$	$x^\ddagger$	A
big pop.	$6.7 \cdot 10^{-20}$ J	$12 \cdot 10^{-12}$ m	$1.4 \cdot 10^{5}$ s^{-1}
small pop.	$6.2 \cdot 10^{-20}$ J	$7.7 \cdot 10^{-12}$ m	$1.1 \cdot 10^{4}$ s^{-1}
CF, linear-cubic	$\Delta G^\ddagger$	$x^\ddagger$	D
big pop.	$6.5 \cdot 10^{-20}$ J	$12 \cdot 10^{-12}$ m	$7.5 \cdot 10^{-19}$ m^{1}s^{-1}
small pop.	$3.2 \cdot 10^{-20}$ J	$7.5 \cdot 10^{-12}$ m	$6.1 \cdot 10^{-23}$ m^{1}s^{-1}
CS, CUSP	$\Delta G^\ddagger$	$x^\ddagger$	A
single pop.	$3.0 \cdot 10^{-20}$ J	$19 \cdot 10^{-12}$ m	$1.0 \cdot 10^{3}$ s^{-1}

CF refers to parameters determined through analysis of experimental constant force data (see subsections 3.1.3, 3.1.4 and 3.1.5). CS refers to parameters determined through analysis of experimental constant speed data (see subsection 3.3.2). The different potentials refer to the model that was used to analyze the experimental data as well as the model that was used to simulate constant speed and constant force data.

The Bell model parameters for the big and small population of the constant force experiment and the CUSP model parameters for the single population of the constant speed experiment are the mean values of the respective equally weighted samples (see subsection 3.1.3 for the Bell model parameters and subsection 3.3.2 for the CUSP model).

For the constant force experiment the CUSP model parameters and the linear-cubic model parameters could not be determined via the mean value of the equally weighted samples. For the CUSP model (see subsection 3.1.4) the equally weighted samples are distributed across several orders of magnitude of the parameter A for both the small and the large population. The same is the case with the linear-cubic model (see subsection 3.1.5) and the parameter D. This leads to the mean of the equally weighted samples being located far outside of any posterior probability density. The following alternative process was used for the the small and the large populations of both the CUSP model and the linear-cubic model: equally weighted samples located in the respective 0.5σ credible volumes were selected to provide the respective bond potential parameters.

Clausen-Schaumann et al.[83–85] (see subsection 2.6.4) examined the rupture processes of the silylester bond (Si-O bond) at the attachment points of carboxymethylated amylose polymers with amino linkers using constant speed and constant force spectroscopy measurements. The constant speed experiments were performed using an AFM with loading rates over several orders of magnitude between 1 nNs^{-1} and 1000 nNs^{-1} and at three different temperatures: 282 K, 193 K and 320 K. The constant force experiments were performed at a holding force of 0.8, 1.0 and 1.2 nN and at seven different temperatures between 295 and 320 K. Two rupture processes were found: One more frequent (72%) and less stable and the other less frequent (28%) and more stable. The two different processes are possibly linked to different chemical moieties close to the silylester bond. Here, I will use the experimentally determined parameters describing the two bond potentials

for the respective rupture processes to simulate constant force and constant speed data, which will serve as a reference for the Si-O bond behavior.

In the experimental data from the constant force measurements two rupture populations were found and consequently two bond potentials were fitted for the BELL model, the CUSP model and the linear-cubic model respectively. The total amount of simulated data was the same as the total amount of experimental data and the respective probabilities for the two potentials were also identical. Furthermore no data below a rupture time of 100 ms was simulated, as this was the cutoff point for rupture events in the experimental data.

For the experimentally determined silyl ester Si-O bond potentials reported by Clausen-Schaumann et al.[85] the CUSP model was used for the simulation, as the involved forces are well below the critical forces of the potentials, and the amount of simulated data was scaled to fit the proportion of the large rupture population in the experimental data. For the potential found through analysis of the constant speed data using the CUSP model the amount of simulated data was scaled in the same way. For the two silyl ester Si-O bond potentials and for the potential from the constant speed data analysis using the CUSP model constant force data was simulated with no restrictions regarding rupture times, as larger parts or even the bulk of the simulated rupture times were shorter than the cutoff in the experimental data. As those three potentials represent one rupture population and the experimental constant force data consists out of two rupture populations, the amount of data for those potentials was scaled to fit the amount of data of the large rupture population. This will allow for an easier comparison with this potential, as it is closer in stability to those three potentials than the more stable potential of the small constant force rupture population.

The simulated data based on the analysis of the constant force data is in good agreement with the experimental constant force data for all three bond rupture models, while the fit to the constant speed data results in simulated data with rupture times, which is two orders of magnitudes shorter than the experimental data (Figure 3.27). The data from the less stable silylester rupture process results in rupture times inbetween the experimental data and the simulated data from the linear-cubic constant speed fit. The data from the more stable silylester rupture process is similar to the portion of the experimental data belonging to the large rupture population.

For the simulation of constant speed data, the CUSP bond rupture potential from the parameter estimation of the constant speed data was used to simulate data. For the Si-O bond potentials reported by Clausen-Schaumann et al.[83–85] the CUSP bond rupture model was used to simulate constant speed data, because the resulting rupture forces were all well below the critical force. From the two potentials per bond rupture model (BELL, CUSP and linear-cubic) found by analyzing constant force data, only the potential representing the large rupture population was used to simulate data corresponding to the total amount of data. The respective more stable bond potentials of the small rupture population were not used for simulations, as they were expected to result in data less similar to the experimental data than the large rupture population. In all simulations no rupture data below a force of 0.5 pN was simulated as this was the cutoff point in the experimental data. The cutoff point was introduced to remove rupture events of non-covalently attached polymers.

The experimental constant speed data obtained previously[38,41] is matched well by the fit of the CUSP model to the constant speed data as expected, but the simulated forces from the three fits (BELL, CUSP, linear-cubic) to the constant force data are multiple times higher than the experimental data (Figure 3.28). The simulated rupture forces from the less stable silylester rupture process is inbetween the forces of the experimental data and the simulated data from the constant force data fits, while the simulated data from the more stable rupture process is similar to those simulated constant force data sets.

3.4.2 Multiple rupture processes in the same sample

The potentials found in the constant force measurements (see subsections 3.1.3, 3.1.4 and 3.1.5) lead to similar simulation results both for constant force and constant speed experiments for all three used bond rupture models (BELL, CUSP and linear-cubic) and can be seen as representing the same two rupture populations. The difference between these potentials and the CUSP model potential from the constant speed analysis (see subsection 3.3.2) is distinct and those potentials clearly represent different populations. The respective potentials of the smaller and more stable population from the analysis of the constant force data is even more different to the constant speed analysis result. Taken together at least three different bond rupture

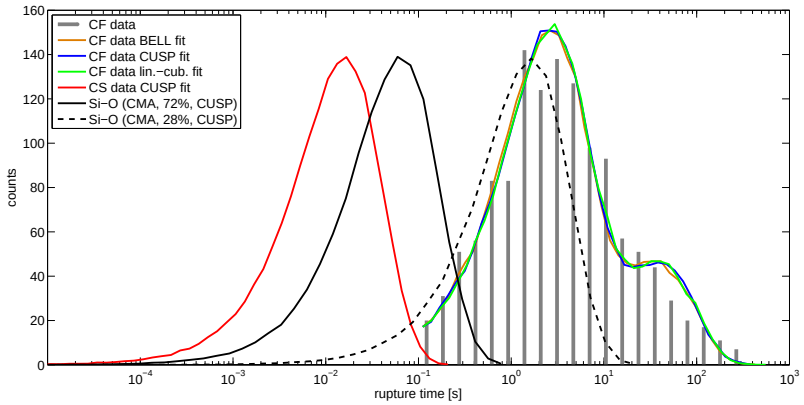


Figure 3.27: Experimental data from constant force measurements between 1.2 and 1.6 nN (grey) and simulated data from various model fits. The simulations using the BELL (orange, small and large population), the CUSP (blue, small and large population) and the linear-cubic (green, small and large population) model fits to the constant force data agree best with the experimental constant force data, while the simulation from the fit to the constant speed data (CUSP: red) is less stable than the experimental data. The simulated data from the more frequent (72%) and less stable rupture process of the two observed silyl ester bond potentials[83,84] has a stability inbetween the experimental data and the simulated data from the fit to the constant speed data (black line), while the simulated data for the less frequent (28%) and more stable rupture process is similar to the experimental data (black dashed line).

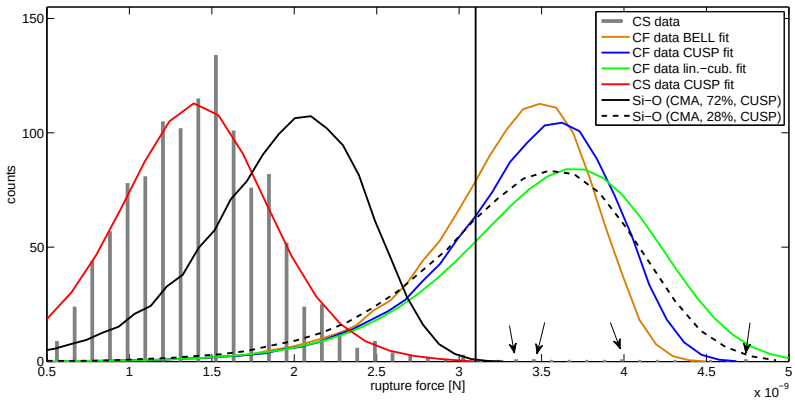


Figure 3.28: Experimental data from constant speed measurements obtained previously[38,41] (grey bars) and simulated data from various model fits. The simulated data from the CUSP (red) model fit to constant speed data agrees well with the experimental constant speed data. The BELL (orange, only large population), the CUSP (blue, only large population) and linear-cubic (green, only large population) model fits to the constant force data result in simulated data at much higher forces than seen in the experimental data. The simulated data from the more frequent (72%) and less stable rupture process of the two observed silyl ester bond potentials[83,84] is more stable than the experimental data (black line) and the simulated data for the less frequent (28%) and more stable rupture process is even more stable and also similar to the simulated data from the two fits to the constant force data (black dashed line). The vertical line marks the beginning of the high force region and experimental data in this region is highlighted with arrows.

potentials have been measured: one in the constant speed experiment[38,41] and two in the constant force experiment.

One explanation could be that the samples are different, but this is probably not the case, since extreme care was taken during polymer synthesis and sample preparation. Upon closer inspection, one finds that the constant speed data has some data in the range where bond ruptures for the constant force populations are expected (Figure 3.28), indicating that at least one of the two populations from the constant force experiment could have been observed, if more data had been acquired. Estimating from the total amount of data and the amount of data in the high force region above 3.1 pN (Figure 3.28), the proportion of the two populations should be roughly 500 to 1 though. Thus one could ask why the much more common, weaker population probed by the constant speed experiments could not be observed in the constant force experiments. The answer to that is that the prepull process, which was introduced to remove rupture events from non-covalent bonds, is mostly responsible for the exclusion of the constant speed potential rupture events from the data of the constant force experiment. Experimental force over time data of a typical prepull event (0 to 1500 to 1000 pN) can be used to estimate the survival rate of the less stable population versus the more stable one. (For simplicity's sake only the large population from the constant force experiment is used in this estimation.) At first the force over time data is converted into a sequence of N segments of equal time length Δt and of constant force (Figure 3.29, left hand side).

Then using the potential parameters from the constant speed rupture population and the CUSP bond rupture model, rupture rates k_m can be calculated for each segment (Figure 3.29, right hand side). The used parameters are: $3.0 \cdot 10^{-20}$ J for $\Delta G^{\ddagger}$, $19 \cdot 10^{-12}$ m for $x^{\ddagger}$ and $1.0 \cdot 10^{3}$ s^{-1} for A. In the following the abbreviation CFCUSP potential will be used.

The bond potential from the large rupture population of the constant force experiment and the CUSP bond rupture model can then be used to calculate another sequence of rupture rates. The following parameters were used for this purpose: $6.7 \cdot 10^{-20}$ J for $\Delta G^{\ddagger}$, $12 \cdot 10^{-12}$ m for $x^{\ddagger}$ and $1.4 \cdot 10^{5}$ s^{-1} for A. In the following the abbreviation CSCUSP potential will be used.

The CUSP model was used because this model is valid for the experimental constant force data and has the least amount of data in an invalid range for the experimental constant speed data. The probability $P_{survival}$ that a

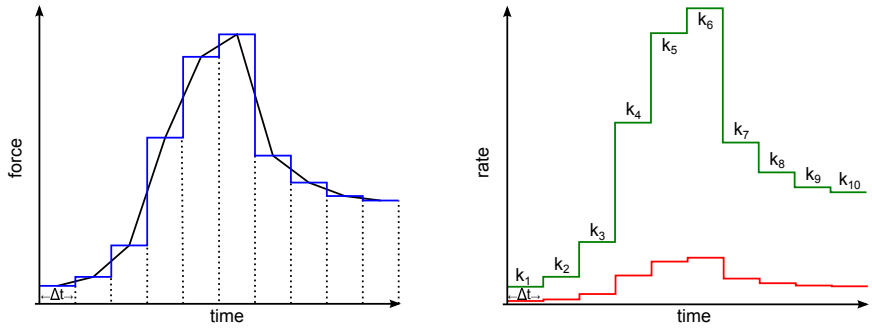


Figure 3.29: Scheme for the calculation of the chance of a bond to survive the prepull mechanism. Experimental data of a prepull event (black) is turned into segments of equal time Δt and constant force (blue). Using a bond potential and a bond rupture model, rupture rates (k_1, k_2, etc.) can be calculated from the force for each segment (green). A more stable potential leads to lower rupture rates (red).

bond of a certain potential survives this sequence can then be calculated using:

$$P_{survival} = \prod_{m=1}^{N} e^{-k_m \Delta t} = \exp\left(\sum_{i=m}^{N} -k_m \Delta t\right) \tag{3.3}$$

$P_{survival}$ for the CFCUSP potential is 0.93, while $P_{survival}$ for the CSCUSP potential is $3.2 \cdot 10^{-5}$. If prior to the pull the ratio of the less stable CSCUSP potential to the more stable CFCUSP potential was 500 to 1, as has been estimated previously, the prepull will turn this into $500 \cdot 3.2 \cdot 10^{-5}$ to $1 \cdot 0.93$. This means 1 surviving bond described by the CSCUSP potential per 58 surviving (large population) bonds described by the CFCUSP potential or 1.6% of all observed ruptures. For higher target forces the prepull process is to be expected to remove a larger proportion of the less stable bonds described by the CSCUSP potential. And for forces beyond 1500 pN the proportion of the bonds described by the CSCUSP potential can be expected to even further shrink. For target forces between 1000 and 800 pN the proportion of the bonds described by the CSCUSP potential will grow, but the amount of data in this force range is below 9 % of the total amount of data.

3.4.3 Comparison of the Parameter Estimation Results with Theoretical Findings

Frank et al. have made predictions[19,36,62] about the siloxane bond rupture behavior based on Car-Parinello molecular dynamics with density functional theory being used to describe the electrons. Constant speed single-molecule experiments at a pulling speed of 55 m/s were simulated with PDMS multimers of varying lengths (dimer to decamer) which were either isolated or attached to β-cristobalite (100) structures. The length of the multimers in the theoretical works is two to three orders of magnitude shorter than in the constant speed experiment, while the pulling speed of 55 m/s is seven orders of magnitude higher than in the constant speed experiment.

Frank et al. found that the maximum observed rupture forces are reduced with an increasing length of the multimer from 6.6 nN for the dimer, 5.2 nN for the trimer, 4.6 nN for the hexamer to 4.4 nN for the decamer. An explanation for this effect could be the redistribution of the added mechanical energy across several bonds. The diminishing decrease of the maximum rupture force implies, that the effect is limited over the number of bonds.[19] Furthermore it was found, that water can weaken the siloxane bond albeit not on the short timescale of the molecular dynamics calculations.[36] For the attached multimers two different maximum rupture forces were observed: a force of 4.4 pN for a bond rupture within the multimer, which is similar to the rupture in the isolated multimer and means that attachment likely doesn't affect the rupture behavior within the multimer, and a force of 3.5 pN at the attachment point of the siloxane molecule.[62]

In the data from the constant speed and constant force experiments no forces higher than the maximum rupture force for the decamer rupture (4.4 pN) and the attachment point rupture (3.5 pN) are found, with the exception of 2 rupture events out of 1154 in the constant speed data. And these high force rupture events are likely caused by the more stable rupture processes found in the constant force experiment. The distribution of the rupture forces in the constant speed data with a maximum at about 1.5 nN could agree with both a rupture at the attachment point as well as in the siloxane multimer, as the experimental pulling speeds are orders of magnitude smaller and the lengths of the polymers are orders of magnitude longer than in the theoretical works, which are both expected to reduce the

rupture forces. For the constant force data there is no disagreement with both reported maximum rupture forces, but this cannot be seen as an agreement either, as the upper limit for the force was chosen as an experimental parameter.

As previously mentioned in subsection 3.1.6, the parameter estimation results of the analysis of the constant force data with the CUSP model and the linear-cubic model can be seen as similar to these theoretical findings. Thus the rupture processes found in the constant force experiment could be related to the covalent bond ruptures from the CPMD calculations.[19,36,62] The CUSP model parameter estimation result from the analysis of constant speed data has a critical force which is slightly lower than the lowest reported theoretical maximum rupture force of the siloxane bond at 3.5 pN. Even though these results seem similar, this critical force is already reached and exceeded with experimental pulling speeds, while the pulling speeds for the theoretical values are seven orders of magnitude higher. Thus the constant speed experiment result does not agree well with the theoretical results.[19,36,62]

3.4.4 Comparison of the Parameter Estimation Results with Silylester Force Spectroscopy Experiments

Using constant speed and constant force single-molecule force spectroscopy measurements at varying temperatures Clausen-Schaumann et al. found two different rupture processes for the hydrolysis of the silylester bond in CMA, which was covalently attached to a silicium oxide substrate.[84,85] The two silylester hydrolysis processes were not distinguishable in the constant speed experiment. Only in the constant force experiments the rupture rates were distinct enough to discriminate between the two.

The less stable bond rupture process of the constant force data predicted by parameter estimations for all three bond rupture models has parameters similar to the more stable silylester hydrolysis process (Figure 3.27). This could suggest that water is also involved in the bond ruptures observed in the constant force experiment as described by Frank et al.[36] As the measurement environment was designed to exclude water as much as possible, this could explain why those rupture processes are much less frequent than the bond rupture process observed in the constant speed experiment. On the other hand, if the bond ruptures processes found in the constant force experiment

are water induced bond ruptures, then the rupture process observed in the constant speed experiment has to be a less stable one. What kind of process this might be remains unclear.

The two silylester rupture processes are caused by different chemical moieties[85] and demonstrate how strong the influence of the immediate environment of a bond can be. One might think, that the two rupture populations found in the constant force data are created by an analog cause, but it has to be taken into account that the rupture potentials for the silyl ester hydrolysis only differ in the attempt frequency, while the barrier distance and barrier height are practically the same. This is not the case for the two rupture populations found in the constant force experiments. All three applied bond rupture models agree, that the distances $x^{\ddagger}$ to the energy barrier for the two rupture processes are distinctly different.

Thus only differences in the attachment itself remain as an explanation for the complex rupture behavior. But those differences are not unlikely to exist, as there are two different attachment points in the sample system, which are created by different reactions (see subsection 2.3.2). An example for differences in the attachement could be caused by the highly reactive Cl atoms at the substrate attachment point. Of the two remaining Cl atoms either none, one or both could have reacted with water molecules or neighboring silanol groups (see subsection 2.3.2).

Parameter estimations using the experimental constant speed data all find very low critical forces that are reached with experimental measuring speeds. For the linear-cubic model the critical force is even 0.1 pN lower than the peak of the experimental rupture forces. The rupture potentials found in the constant force experiment as well as the rupture potentials for the silyl ester hydrolysis in CMA on the other hand are far from reaching their respective critical forces at the loading rates from the constant speed experiment. The maximum of the rupture forces predicted by the CUSP model with high barrier approximation for example is at 35% of it's critical force. For the linear-cubic model the number is 59%. As the silylester hydrolysis rupture potentials, the constant force rupture potentials as well as the constant speed rupture potentials are all expected to be related to the rupture of an Si-O bond, it seems unusual, that the critical force for the constant speed rupture potential is so much lower and also reached at such low loading rates, while this is not the case for the other four potentials.

This is expected as the three bond rupture models do not describe the experimental constant speed data well, which manifests as too low critical forces and following from that rupture forces being observed in invalid force ranges (see subsection 3.3.4).

It could be possible that analog to Clausen-Schaumann et al.[85] actually two rupture populations have been observed in the constant speed experiment, which are too similar in their rupture behavior to separate. It might also be possible that one cannot distinguish between the two rupture populations in constant force experiments. There were attempts to analyze the experimental constant speed data with a 2 rupture population approach, but the analysis was not successful for any of the three bond rupture models.

Another explanation could be that a more complex polymer model has to be used to determine the contour lengths of the PDMS polymers. In this work and previous works[38,41] the extensible freely jointed chain was used to fit the bahvior of the PDMS polymers. Works by other groups[16,18] have shown that polymers show a crossing over from extendible freely jointed chain behavior to freely jointed chain behavior at high forces. Taking into account this crossover in polymer behavior could lead to different contour lengths, following from this different loading rates and thus different loading rate dependent rupture force distributions. One or more of the three bond rupture models could possibly sufficiently describe these changed rupture force distributions.

Chapter 4

Summary and Outlook

The goal of this work was the investigation of the mechanical properties of the Si-O bond in the silicone elastomere PDMS. A quantification of the strength of this bond using several models could form the basis for future research which aims to improve the mechanical strength and stability of silicone materials. To achieve this goal constant force measurements were performed and analyzed using Bayesian parameter estimation and three different bond rupture models (see section 3.1). Previous constant speed experiments were reevaluated with Bayesian parameter estimation and the results compared to the results from the constant force experiments (see section 3.3).

Even though the measured PDMS polymer lengths vary by over an order of magnitude, no length dependence of the rupture rates could be observed (see subsection 3.1.1). This indicates that not all siloxane bonds in our sample are susceptible to rupture, but only the bonds at or close to one or both attachment points of the polymer. Moreover the data can best be fit to two bond rupture populations (see subsection 3.1.2). In combination with the knowledge that only the bonds at the attachment points rupture this suggests, that there are at least two different rupture mechanisms that were accessed in the constant force experiment.

The fits of the Bell, the CUSP and the linear cubic-model to the constant force data (see section 3.1) provide a good parameter estimation for the distance $x^{\ddagger}$ which is about 11 pm for the big population and about 7 pm for the small population. For the Bell model the values of the second parameter k_0 could be determined as $10.6 \cdot 10^{-3}$ s^{-1} for the big population and $2.8 \cdot$

10^{-3} s^{-1} for the small population. The two other parameters of the CUSP model ($\Delta G^{\ddagger}$ and A) and the linear-cubic model ($\Delta G^{\ddagger}$ and D) are spread out over a wide range of values and it is only possible to find lower limits for them. If a range of known values for A from other experiments is used to restrict the posterior to values of A between 10^3 to 10^7 s^{-1}, a range of values for $\Delta G^{\ddagger}$ of $5.3 \cdot 10^{-20}$ to $8.6 \cdot 10^{-20}$ J for the large population and $5.1 \cdot 10^{-20}$ to $9.0 \cdot 10^{-20}$ J for the small population can be given. Both the CUSP and the linear-cubic model are valid for the constant force data, since the highest forces are still very low when compared to the lowest possible critical forces. The Bell model is also valid at such low holding forces. This is reflected by the agreement of all three models with respect to their common parameter $x^{\ddagger}$. The results of the parameter estimation with the CUSP model can also be converted to the parameters of the BELL model and the resulting credible areas overlap with the credible areas of the BELL parameter estimation, which is further confirmation for the validity of the BELL model and the agreement of all three models with each other.

But these very low holding forces are also the reason for the difficulties of the parameter estimations (see section 3.2) for the CUSP and the linear-cubic model, as the low holding forces and small holding force range were determined to be the dominating influence in the quality of the parameter estimation. Simulations of constant force bond rupture experiments for varying amounts data, multiple rupture processes and errors in the cantilever spring constant determination all did not show as large of an influence on the parameter estimation quality as simulations with varying force ranges did. To allow the determination of the two potentials with three parameters, future measurements have to focus on expanding the force range to much higher forces. A higher amount of data will also prove useful. Since rupture rates have to be expected to vary considerably with the holding force, variable measurement temperatures should also be utilized.

The analysis of the constant speed data (see section 3.3) with the CUSP model with high barrier approximation and the nested sampling algorithm yielded the following parameter estimations: a barrier height, $\Delta G^{\ddagger}$, of $3.0 \cdot 10^{-20} \pm 9.0 \cdot 10^{-22}$ J (1σ), a distance to the energy barrier, $x^{\ddagger}$, of 19 ± 0.7 pm (1σ) and an attempt frequency, A, of $1.0 \cdot 10^3 \pm 96$ s^{-1} (1σ). While the fit of the CUSP potential matches the result of a previous analysis,[38] the CUSP model is not valid for the experimental constant speed data, because

the measured forces are close to F_{crit} of the found bond potential and the CUSP model is only valid for low forces, which are distant to F_{crit}. Since the forces are too high for the CUSP model, the Bell model is not valid as well. Using the linear-cubic model without high barrier approximation and the nested sampling algorithm yields a bond potential with a barrier height, $\Delta G^{\ddagger}$, of $3.1 \cdot 10^{-20}$ J $\pm$ 9.2 % (1σ), a distance to the barrier, $x^{\ddagger}$, of 34 pm $\pm$ 20 % (1σ) and a diffusion coefficient, D, of $2.5 \cdot 10^{-20}$ m^2s^{-1} (1σ). But the critical force for this set of parameters is even lower than the result for the analysis using the CUSP potential and a majority of the observed rupture events occur at rupture forces where according to the parameters and model no energy barrier exists. Thus also the results for the linear-cubic model have to be seen as invalid.

The estimated parameters from the constant force measurements would lead to two potentials that predict much higher bond rupture forces for the constant speed measurements than what was actually observed (see section 3.4). The separation is clear as the constant force potentials would result in rupture forces that are several times higher than the experimental data. Thus three different bond rupture populations were found in the examined samples, which are likely to correspond to three different bond rupture processes.

The rupture processes of the Si-O silyl ester bond analyzed by Clausen-Schaumann et al.[83,84] can also be used as a reference for comparisons (see subsection 3.4.4). The more frequent rupture process yields a potential that has a stability which lies between the potential found in the constant speed experiments and the various potentials of the large rupture population from the constant force experiments. The latter potential is also similar to the potential of the less frequent silylester hydrolysis process. The results by Clausen-Schaumann et al. also illustrate how strongly the rupture behavior of the Si-O bond can vary.

An explanation for the existence of different rupture processes could be chemical differences close to and/or at the attachment points of the polymer. Especially if one considers that both the -$SiCl_3$ and the -$Si(CH_3)_2$ radical used for covalent attachment are highly reactive and thus likely to be less selective. Possible alternative polymer attachments could be reactions with multiple silanol groups on the surface or reaction with Si or N on the surface of the cantilever tip.

As mentioned above the parameter estimation using the CUSP model as well as other models to analyze the constant speed data indicate, that in the constant speed experiment the critical force is reached and exceeded with experimental pulling speeds. This is not the case for the two rupture potentials found in the constant force data as well as for the two silylester hydrolisis potentials (see subsection 3.4.4). This deviation from other Si-O bond rupture processes could indicate different bond rupture processes or that the affected bond is strongly destabilized by the environment. A more likely explanation for this is that there wer actually two different rupture processes with overlapping rupture force distributions measured in the constant speed experiments. Two rupture populations could appear as a broader single rupture force distribution. If the data is then analyzed under the assumption that only one rupture population exists, it would lead to false results. So it is possible that not one, but two rupture processes were observed during the constant speed experiment.

Another explanation for the problems with the analysis of the constant speed data could be that a more complex mechanic polymer model is necessary to calculate the polymer contour lengths (see subsection 3.4.4). This could lead to a change in loading rates and thus possibly to a shape of the rupture force distributions that can then be better described by one or more of the three bond rupture models.

For future experiments a good parameter estimation of the two rupture processes found in the constant force experiments is a necessary next step. With regards to the constant speed experiment it has to be examined, if one or actually two rupture processes have been measureed. Both of these goals should be achievable with further constant force experiments, if the pulling process can be sped up sufficiently. Future constant force experiments should try to expand the accessible holding force range as much as possible and make use of the temperature dependence of the bond rupture processes at the same time. The prepull procedure employed in the constant force measurements in this work could also be used to exclude less stable populations from the measurement. Lastly usage of linkers could be used to attach the PDMS polymer in a single specific way. Reducing the chain of siloxane bonds to a single siloxane bond attached to cantilever and surface via linkers could also be used, but this would be a system further removed from the solid state silicone than the currently used one, which would be

undesirable as the examination of the mechanical stability of solid silicone materials is the ultimate goal of this work.

Part II

Imaging of Nucleosomal Arrays

Chapter 5

Introduction

The length of the DNA in comparison to the size of the nucleus in eukaryotes is astonishing: while human DNA has a length of several meters, it has to fit inside a nucleus with a diameter of several micrometers. In light of this fact it becomes clear that simply fitting the DNA inside of the nucleus alone makes a systematic compaction of the DNA necessary. This compaction process is facilitated by proteins, which interact with the DNA and form a large complex with it which is called chromatin.[89–91] The compaction of DNA in eukaryotic cells encompasses several organizational levels of increasing density. The inital stage of DNA compaction is wrapping of the DNA around protein complexes called histone octamers to form nucleosomes, which are connected by linking DNA strands.[91–97] The nucleosomes connected by the linker DNA resemble beads connected by a string which is why this stage is also called beads-on-a-string. The second step is presumably the formation of a fiber with an approximate diameter of 30 nm.[91,98] This is followed by several further steps of folding and attachment to protein scaffolds. And ultimately during mitosis the highest level of DNA compaction is reached in the form of the characteristic chromosome structure. On the other hand DNA has to remain accessible for the purposes of transcription, replication, repair etc. To enable these processes access to the bare DNA has to be provided by unpacking the chromatin on all structural levels down to the relocation or removal or histone octamers. Thus not only a dense compaction of the DNA but also the regulation of the compaction is necessary. In comparison to the beads-on-a-string level far less is known about the structure and the regulating processes of the 30 nm fiber. In fact to date no 30 nm fiber has

been observed in vivo. Several models for the structure of the fiber have been proposed in the past several decades and it has not yet become apparent which of those models is the correct one.[99–102] Regarding the regulation of the formation and alteration of the 30 nm fiber structure it is of interest which molecules interact with the fiber and under which conditions these interactions occur.

The atomic force microscope can help to give insights into both the structure of the 30 nm fiber and the interactions of the fiber with different molecules. The AFM is able to resolve single DNA strands in both air and liquid environments and thus could be also applied to study the larger beads-on-a-string and 30 nm fiber structures. When measuring in liquid it should be possible to observe the fiber under conditions where the structure of the fiber as well as its interactions are similar to what can be expected in vivo. The exact dimensions of the fiber depending on the length of the linker DNA and number of nucleosomes can help in the assessment of the various proposed models. Because atomic force microscopy is a probe based method, measurement speed has always been a limiting factor, but with current systems it has become possible to achieve time resolutions close to 24 frames per seconds which is commonly used in the recording of real time films. Lastly, the combination of atomic force measurements with optical methods such as flourescence microscopy, FRET and others is an option for future more advanced experiments. To provide the first stepping stone in the investigation of the structure and interactions of the 30 nm fiber using the atomic force microscope, the primary focus in this work is to establish sample preparation and measurement protocols, which allow the imaging of both the beads-on-a-string structure and the 30 nm fiber at first in air and then in liquid.

This is the second part of this work and consists of four chapters: after the introduction in chapter 5, chapter 6 introduces the theoretic and practical fundamentals necessary to understand the nucleosomal array and DNA imaging results in chapter 7. Chapter 8 summarizes the second part of this work and gives some perspectives on future directions and possibilities.

Chapter 6 has three sections: section 6.1 introduces AFM imaging. Section 6.2 is about DNA and nucleosomal arrays in general and about the samples which were measured specifically. In section 6.3 sample preparation and imaging protocols can be found.

The imaging results in chapter 7 are divided into three sections. Section 7.1 is concerned with test measurements, which were performed in air. Section 7.2 contains the results of in liquid measurements and the extent to which the goals of quality assurance and structure control could be achieved. Section 7.3 contains the conclusions drawn from the previous two sections of the chapter.

Chapter 6

Theory and Methods

In this chapter the necessary theory and methods to image DNA and nucleosomal arrays with the AFM will be introduced. In section 6.1 AFM imaging is described including the two used imaging modes and the two used imaging media. Then in section 6.2 basic information about DNA and chromatine is given as well as detailed information about the used samples and the attachment of the samples. Finally in section 6.3 experimental procedures and materials concerning sample preparation and imaging can be found.

6.1 AFM Imaging

The basic function of an AFM has been decribed previously in this work (see section 2.1) and thus the following will focus on AFM imaging.

Imaging with the AFM is a probe based microscopy method. Interactions between the cantilever tip of the AFM and the sample can be used to collect a multitude of information about the sample. The forces interacting between the tip and the sample are contact forces, van der Waals forces, capillary forces, electrostatic forces and magnetic forces among others. The threedimensional topography of the sample is just one example for information that can be gained via AFM imaging. Others are mechanical properties like hardness or even electromagnetic properties such as conductivity or magnetism. As a probe based method information can be collected only from a single point of the sample at a time which makes scanning necessary. The necessity of scanning is the reason why the speed of AFM imaging is an important limiting factor.

AFM imaging can be done in many different media such as liquids, gases or vacuum. For biological samples measurements in liquid are naturally preferable to examine them in a state as similar to in vivo as possible.

Another important property of AFM imaging is that the cantilever tip can change the sample via the aformentioned interactions. This can be beneficial if a manipulation of the sample such as cutting DNA strands or moving particles is desired, but it is also easily possible to deform the sample unintentionally. Thus the strength of the forces the sample is subjected to during a measurement have to be considered carefully and minimized as much as possible. This is especially true for biological samples which are soft in general.

6.1.1 Imaging Modes

Two different atomic force microscope imaging modes were used in this work: contact mode and tapping mode. Both methods can give information about the topography of the sample. How the two methods work and which properties they have, will be described in this subsection.

In contact mode (see section 7.1 for results) the cantilever is pressed onto the sample with a set force. During the scanning process a feedback mechanism is employed, which changes the z-position of the cantilever to keep the distance to the sample and thus the force constant. From the movement of the cantilever during the scanning process the absolute topography can be determined directly. Advantages of contact mode imaging are the simple implementation and analysis as well as ease of use. Preparation is minimal and calibration is not necessary. Disadvantages are high vertical and lateral forces, which can destroy the sample.[103] For soft biological samples this is especially problematic. If the sample is hard there is also the danger of the deformation of the tip. To minimize the occuring forces a cantilever with a soft spring constant has to be used. As soft cantilevers do not track the surface well, the imaging speed is further reduced.

In tapping mode (see section 7.2 for results) the cantilever is oscillated at a frequency close to the resonance frequency of the cantilever.[104] Close to the surface the interactions between cantilever tip and sample dampen the amplitude of the oscillation.[105,106] By keeping the amplitude of the oscillation constant with a feedback mechanism that moves the cantilever in the z-direction, the distance to the surface can be kept constant. Thus

the topography of the sample can be directly obtained from the movement of the cantilever. As the contact between tip and sample is only intermittent and lateral forces are minimized, tapping mode measurements are more gentle than contact mode measurements and thus it is easier to measure samples without destroying them. The forces on the sample are lower for softer cantilevers as well as smaller oscillation amplitudes which can be used to minimize those forces. Furthermore not only the amplitude of the oscillation can be used to obtain infromation from the sample. The phase of the oscillation can give insight in the mechanic properties of the sample and thus allows distinction between different materials of the sample.[106] One example for different materials in a sample can be the mica substrate and attached DNA which have different mechanical properties. A disadvantage of the method is that the process of finding the correct oscillation frequency and drive amplitudes in liquid can take some time and effort, which increases the time before a measurement can be started considerably.

6.1.2 Imaging Media

The medium in which atomic force imaging is performed has a big influence on both the imaging process itself and the structure and properties of the sample. DNA and nucleosomal arrays were imaged in air (see section 7.1 for results) and in water (see section 7.2 for results) in this work. Measurements in air have the advantage of being quick and easy to perform. The disadvantages are that the samples dry out and that the topology is obscured by an ever present film of water which interferes with the measurement via capillary forces. In liquid measurements allow the imaging of biological samples in a more natural environment and without the interference of a water film, but are more complex and elaborate. Thus in air measurements are used for test measurements to establish sample preparation protocols, while in liquid measurements are used to actually examine the samples.

6.2 DNA, Nucleosomes and Chromatin

Chromatin is a complex of DNA and proteins in the nucleus of eucaryotic cells. It has multiple functions: it compacts the DNA to allow it to fit inside the nucelus, protects the DNA from damage and also controls gene expression. Increased condensation of the chromatin during mitosis and meiosis

stabilizes the DNA to allow for the separation of the chromosomes.[89–91] Multiple levels of compaction are facilitated by the chromatin. On the lowest level DNA forms a complex with histone octamers, which are called nucleosomes.[91–97] The next level of condensation is presumably the 30 nm fiber, which received it's name due to it's diameter of roughly 30 nm.[91,98] The 30 nm fiber has not been found in vivo yet. Several higher levels beyond the 30 nm fiber exist and allow condensation up to a degree, where the characteristic X-shape of the chromosomes becomes visible during mitosis. But this work is only concerned with the aformentioned two lowest levels of condensation of the chromatin.

On the first level of condensation of the chromatin the DNA forms a complex with histone octamers. DNA has a diameter of roughly 2 to 2.6 nm and a length of 0.33 nm per base pair.[107,108] Atomic force microscopy studies[109–111] found that the persistence length of DNA is 50 nm at length scales above 30 nm. Histone octamers consist of the four different core proteins H2A, H2B, H3 and H4. A histone octamer is assembled from a tetramer with two copies of H3 and two copies of H4 and two H2A/H2B dimers.[91] A nucleosome is formed through interactions between the histone octamer and the minor groove of the DNA. In a nucleosome the 147 base pairs of the DNA are wound approximately 1.67 times around the histone octamer. Nucleosomes have the shape of a disc with a diameter of 10 nm and a height of 6 nm.[91–97] The nucleosomes are connected by stretches of unbound DNA, which are called linker. The length of the linker varies and can be changed by relocation of the nucleosomes. But it is possible to determine the position of nucleosome by using the so-called 601 sequence, which has a high affinity for the binding of histone octamers.[112,113] Because of the resemblance to a string with beads, this compaction level of chromatin is called beads-on-a-string.[91]

The 30 nm fiber is the second level of condensation of the chromatin. It has been found, that ion interactions are important for the stability of the 30 nm fiber.[114,115] Several structural models for the 30 nm fiber have been proposed: a solenoid,[99–101] a two-start helix with the linkers crossing the inside of the helix[99–101] and also a double helix of twisted tetranucleosomal units.[102] It yet remains unclear which of these structures are actually present in chromatin under biological conditions in vivo.

6.2.1 DNA and nucleosomal array Properties

Two different samples were imaged in this work: a DNA sample for test measurements (see subsections 7.1.1 and 7.2.1 for results) and the nucleosomal arrays (see subsections 7.1.2 and 7.2.2 for results).

The DNA sample for test measurements contained extracted genomic yeast DNA and was provided by Susanne Ude. The DNA was fragmented and thus DNA strands of varying lengths were present in the sample.

The nucleosomal arrays consist of DNA and histone octamers. A nucleosomal array should ideally contain a double stranded DNA and 25 histone octamers. Both a lower amount of histone octamers through incomplete assembly or disintegration after assembly as well as incomplete histone octamers were observed though and one goal of this work was to act as quality control for the nucleosomal arrays to help find optimal assembly and measurement conditions for complete nucleosomal arrays.

The double stranded DNA was of a total length of 4971 base pairs and contained 25 copies of the '601' sequence[112,113] with a length of 147 base pairs at regular intervals, spacers of 50 base pairs inbetween every two copies of the '601' sequence and ends with a total length of 96 base pairs (Figure 6.1). The DNA was developed in the group of Daniela Rhodes (MRC Laboratory of Molecular Biology, Cambridge) and was provided by the group of Peter Becker (Ludwig-Maximilians-Universität, Munich) with the consent of Daniela Rhodes.

The histones for the histone octamers were human wildtype and were expressed in E. coli (Figure 6.1). The plasmids for the histones were provided by the Group of Gernot Längst (University of Regensburg, Regensburg).

Production of the DNA and histone proteins as well as assembly of the histone octamers and the nucleosomal arrays were performed by Katarzyna Krzemien.

For the dimensions of the nucleosomal array in both the beads-on-a-string conformation as well as the proposed condensed conformations some estimations for the dimensions can be made, if one assumes a diameter of 2 nm for the DNA,[107,108] 0.33 nm length per base pair of the DNA[108] and a diameter of 11 nm for the nucleosomes.[91–97] For the beads on a string conformation and the assumption that each '601' sequence of 147 base pairs is completely wrapped around its corresponding histone octamer, the length of the array should be made up of the 24 spacers and the ends of the DNA

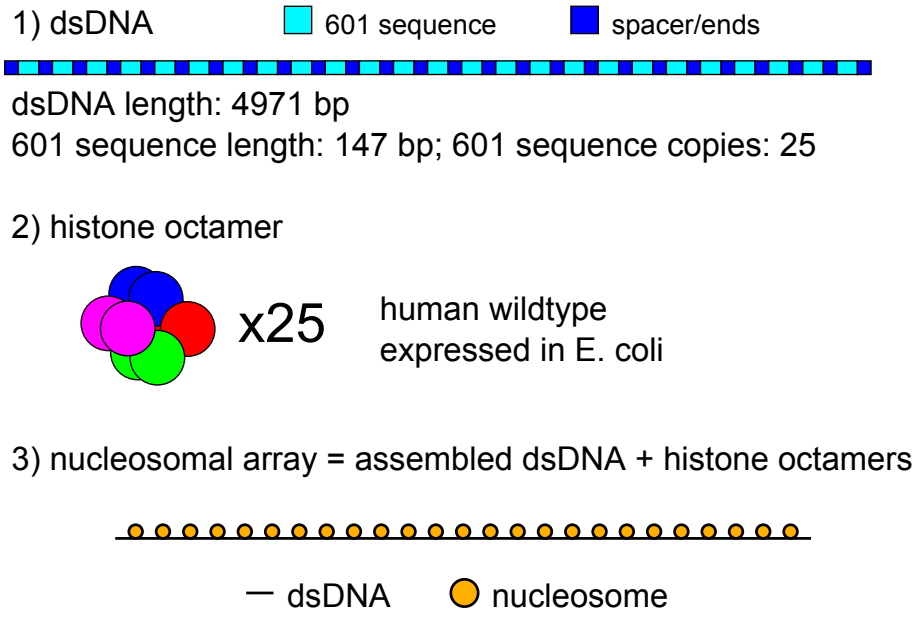


Figure 6.1: Nucleosomal array sample system with double stranded DNA and histone octamers. The DNA (1) contains 25 copies of the 601 sequence which is used to determine the position of the nucleosomes once the DNA (1) and histone octamers (2) are assembeled into the final nucleosomal array structure (3). The nucleosomes in the array should be regularly spaced with minimal sliding.

double strand. The resulting length of these 1296 base pairs correspond to a length of about 428 nm. Roughly this length should be also expected to be observed in the experiment. The bare DNA itself should have a length of 1640 nm. For the proposed condensed structures the diameter of the nucleosomes can be used as a point of reference. The width of the nucleosomal array should naturally be in the range of 30 nm since it's structure should correspond to the 30 nm fibre. Models for 22 nucleosomes[100] predict heights of either 33 nm for the two start model and 47 nm for the solenoid model of the 30 nm fiber and thus the condensed nucleosomal array with 25 nucleosomes should be expected to fall roughly in the same range, if slightly larger.

6.2.2 Attachment of the DNA and the nucleosomal arrays

To allow imaging with the AFM the DNA and the nucleosomal arrays have to be attached to a smooth surface. Electric charges are used to attach both the DNA and the nucleosomal arrays. Negative charges are provided by the phosphate groups in the DNA backbone which are present in both the DNA sample and the nucleosomal arrays. Thus the substrate surface has to be positively charged. Mica is a mineral with easily separatable layers. If mica is freshly cleaved, one can obtain an atomically smooth surface. Two methods based on this principle were used in this work and are described in this subsection. Both methods were used for in air (see section 7.1) and in liquid measurements (see section 7.2).

The surface of freshly cleaved mica is negatively charged and cannot be used directly for the attachment of the two samples. Two different methods can be used to achieve a smooth positively charged surface from cleaved mica: bivalent ions coating[116] and polylysine coating.[109,117] Bivalent ions such as Mg^{2+}, Ni^{2+} or Zn^{2+} are attracted by the negatively charged surface and form a layer which shields the negative charges and creates an overall positively charges substrate surface. Polylysine is a polymer with positive charges that also forms a layer of positive charges for the attachment of the sample.

The surface formed by bivalent ions is smoother than the surface formed by polylysine. Attachment using polylysine on the other hand is much stronger than attachment with bivalent ions. The strength of the attachment also varies between the various bivalent ions. As the goal of this work was the imaging of the nucleosomal arrays both in the beads-on-a-string conformation and the condensed conformation and it was not clear which attachment strength was necessary to image those conformation, several different bivalent ions as well as polysine were used to find the optimum attachment conditions for the two conformations.

Finally it has to be noted that the histone octamers are not directly attached to the surface. They are attached to the surface via their interaction with the DNA in the nucleosome. To enable the formation of nucleosomes with the negatively charged DNA, histones are positively charged. The positive charges of the substrate surface could thus possibly destabilize the nucleosomes.

6.3 Methods and Materials for nucleosomal array Imaging

In the following the measurement setup (see subsection 6.3.1), commercial software (see subsection 6.3.2) and used cantilevers (see subsection 6.3.3) are listed. Then general experimental materials and procedures are described (see subsection 6.3.4). Finally a detailed description of sample preparation procedures as well as preparation conditions is given (see subsection 6.3.5).

6.3.1 Atomic force microscope

The instrument used for nucleosomal array imaging is the MFP3D AFM by Asylum Research which allows for contact mode, AC Mode and force mode measurements among others as described previously in 2.9.1. A Newport RS4000 table was used to reduce external vibrations.

6.3.2 Commercial software

Nucleosome and DNA imaging on the AFM was performed using the MFP3D version 070111+217 software by Asylum Research as previously described in subsection 2.9.2.

6.3.3 Cantilevers

For nucleosomal array imaging two different cantilevers were used: For contact mode in air MLCT-AUHW rectangular cantilevers by Veeco were used with a spring constant of 0.02 N/m and a resonance frequency of 10 to 20 kHz in air. For tapping mode in liquid BL-AC40TS cantilevers by Olympus were used, which had a spring constant 0.09 N/m and a resonance frequency of 75 to 145 kHz in air and 17 to 45 kHz in water.

To determine the cantilever amplitude of the BL-AC40TS cantilever, the InvOLS of a single fresh cantilever was measured, as the cantilevers used for imaging were not available anymore. Since cantilevers vary the given values for cantilever amplitudes have to be treated as rough estimations.

6.3.4 Experimental Procedures and Materials

Muscovite mica is a sheet silicate with easily separable layers. Freshly cleaved mica provides a very smooth surface, which is ideal for AFM imag-

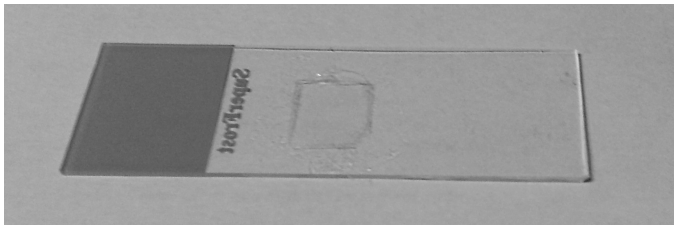


Figure 6.2: Substrate for imaging DNA and nucleosomal arrays. It is a piece of mica fixed on a glass cover slide with epoxy resin.

ing. The surface can also be functionalized in different ways. An easy to handle mica substrate can be created by cutting out a rectangular piece of mica from a mica sheet and then attaching it to a glass cover slide with epoxy (Figure 6.2). The amount of epoxy can be relatively small and should only cover the bottom of the mica piece. Cleaving off a few layers of mica can be done easily by attaching tape to the mica and then peeling it off carefully. Mica substrates for the measurements in this work are V1 mica discs from Scientific Services and V1 mica sheets from Ted Pella, Inc with a thickness of about 0.15 mm.

Microcrystals which form after some time in the buffer solutions disturb the imaging of the DNA and nucleosomal arrays. Two procedures were used to remove those crystals. Both should be done immediately before the sample preparation or one day prior to the measuremen at the earliest. The first method to remove the crystals is ultracentrifugation of the buffer solutions in Eppendorf reaction tubes in a Eppendorf 5417R Ultracentrifuge at 14000 rpm for 30 to 45 min. Afterwards the upper 2/3 of the buffer in the reaction tube can be transferred to a fresh reaction tube and then used. The second method is filtration of the buffers with Acrodisc 25 mm Syringe Filters by PALL with a 0.1 μm membrane. This method is fast and useful to treat larger volumes of buffer, but doesn't remove microcrystals as effectively as the first method.

Measuring DNA or nucleosomal array samples is straightforward (Figure 6.3): the glass slide with the mica substrate is placed on the xy-stage and secured with magents. Then the AFM head is lowered on the sample and the cantilever approach process can be started. For liquid samples the formation of a meniscus between sample and AFM head should be ensured. If this is not the case more measurement buffer has to be added in

between AFM head and sample. If a sample is measured for several hours measurement buffer should be added to the sample about every hour.

In air contact mode imaging preparation:

1. Center light beam on cantilever using the microscope.
2. Zero the deflection using the wheels on the AFM head.
3. Choose the setpoint at a low negative value.
4. Approach the surface with the cantilever. Close to the surfcae a snap in might occur, but this is not detrimental to the measurement as the cantilever tip is usually not damaged.
5. As soon as the surface is reached the setpoint should be adjusted so that manual z-movement via the wheel on the AFM head is tracked.
6. The system is now ready for imaging the sample.

In liquid tapping mode imaging preparation:

1. Lower the head on the sample and wet the cantilever. Ensure that a liquid meniscus has formed between sample and AFM head.
2. Center light beam on cantilever using the microscope.
3. Perform a tune over the range of expected resonance frequencies.
4. Identify a peak which is both not too small and clearly seperated from other peaks.
5. Set the driving frequency slightly lower than the maximum of this peak. The amplitude at this frequency should be 5 to 10 % smaller than the maximum of the peak.
6. Choose a set point at about 80 % of the free amplitude at the desired frequency.
7. Slowly approach the surface. Adjust drive amplitude and amplitude set point regularly during the approach as the oscillation of the cantilever changes during the approach. The drive amplitude should be

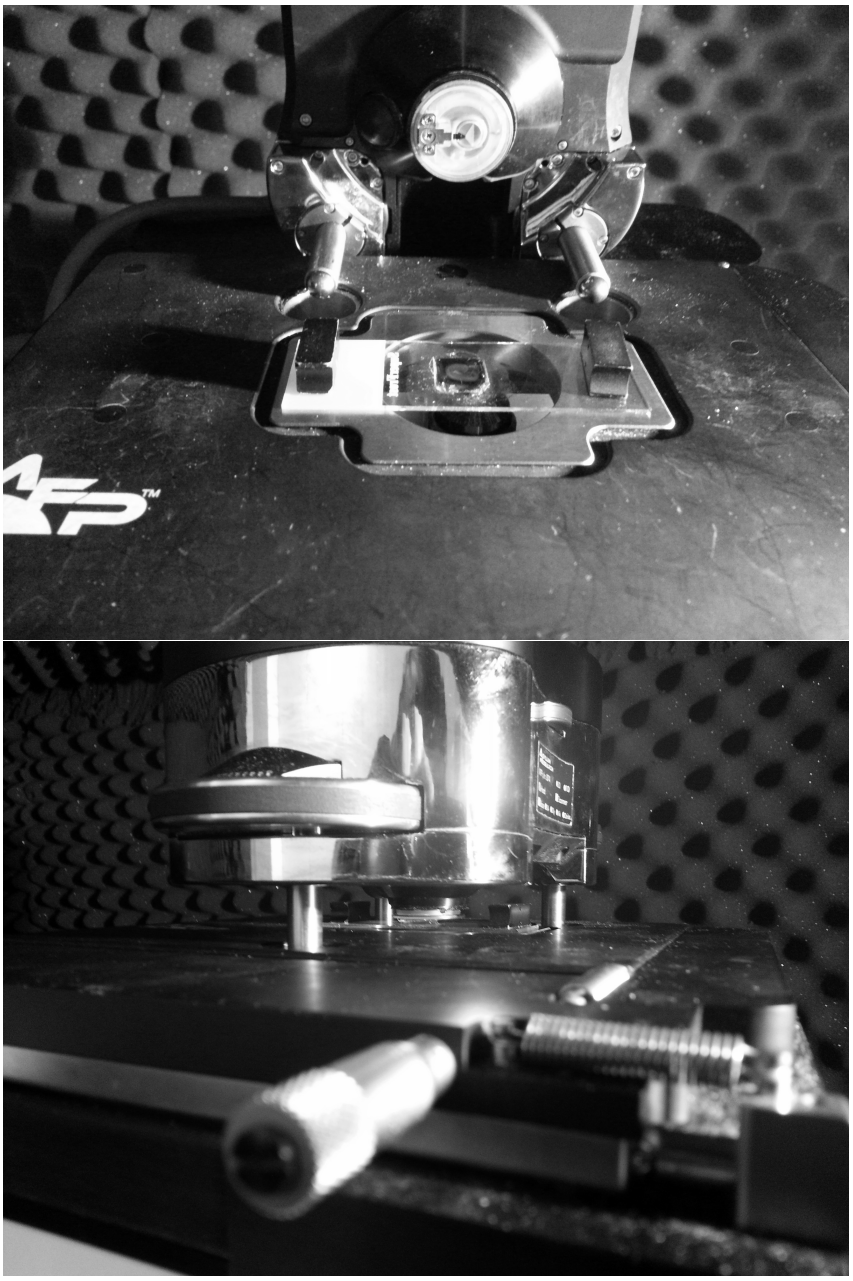


Figure 6.3: Nucleosome measurement setup. Above: Inserting the sample: The AFM head has been tilted away and the substrate with the sample has been placed on xy-stage. Below: AFM and sample are in place. A liquid meniscus is formed between substrate and AFM head. The sample can now be measured.

adjusted so that the amplitude of the oscillation remains constant. If not possible, the amplitude setpoint has to be adjusted to once again achieve 80% of the free amplitude. This is important to ensure a soft approach to the surface and prevent damagin the cantilever tip.

8. After surface contact the cantilever has to be retuned. A different peak might be more ideal near the surface.

9. The system is now ready for measurements.

Typical imaging parameters:

- Scan size: 500 nm to 5 μm
- Scan rate: 0.5 to 1 Hz
- Scan angle: 90 °
- Scan Points and Scan Lines: 512 or 1024
- Integral Gain: 10 - 30
- Drive amplitude: 0.5 to 2.5 V (can vary considerably)
- Set Point: 300 to 700 mV (can vary considerably depending on the free amplitude)

6.3.5 Sample preparation

Preparation for in air imaging

Sample molecules for contact mode measurements in air (see section 7.1 for results) can be attached with bivalent ions or polylysine: For bivalent ion attachment mica is cleaved and then deposition buffer containing the sample and bivalent ions is pipetted onto the substrate surface. A surface area of ˜1 cm^2 area can be covered with 40 μm of deposition buffer. After incubation for 5 min the buffer is rinsed off with Millipore Milli-Q filtered water. After the water has evaporated the sample can be measured.

For the attachment of DNA with bivalent ions the concentration of DNA in all attachment buffers was 1 ng/μl. 5 mM HEPES was used in all attachment buffers to stabilize the solution at a pH of 8.0. The first attachment

buffer used a bivalent ion concentration of 5 mM $MgCl_2$. In a second experiment the concentration was reduced to 1 mM $MgCl_2$. For a third experiment the bivalent ion was changed to 5 mM $ZnCl_2$. For the attachment of nucleosomal arrays with bivalent ions a concentration of nucleosomal arrays of 0.5 ng/μl was chosen for all attachment buffers. The concentration of the bivalent ions for the attachment was in all cases 1 mM $MgCl_2$. In the first experiment the solution was buffered at a pH 8.0 with 5 mM HEPES. In the second experiment the solution was buffered at pH 7.6 with 10 mM Tris HCl and additionally the attachment buffer contained 10 mM KCl and 1 mM EDTA.

For attachment via polylysine a 0.01 % polylysine solution has to be incubated on freshly cleaved mica for 30 s to form a closed polylysine layer and then be rinsed off with water twice. Here, also 40 μm of polylysine solution are sufficient for an area of around 1 cm^2. Afterwards the deposition buffer with the sample can be pipetted onto the mica surface. In analogy to to the bivalent ion method the buffer is rinsed off with water after an incubation period of 5 minutes. The sample can be measured as soon as the water has evaporated.

The attachment buffer for the attachment of both DNA and nucleosomal arrays to polylysine coated mica contained 10 mM Tris HCl and 1 mM EDTA at a pH of 7.6. For the DNA attachment the concentration of DNA in the buffer was 1 ng/μl of DNA and for the nucleosomal arrays the nucleosomal array concentration was 0.25 ng/μl.

Preparation for in liquid imaging

For tapping mode measurements in liquids (see section 7.2 for results) the two above mentioned attachment methods can be also used: Bivalent ion attachment is very simple. After cleaving the mica on the glass cover slide the deposition buffer with the sample and the bivalent atoms can be dropped on the substrate. After incubating the samples for 15 to 30 min they can be immaged directly without any further steps.

The sample buffer for bivalent ion attachment of DNA contained 1 ng/μl DNA, 1 mM $MgCl_2$ and 5 mM HEPES at pH 7.6. For the attachment of nucleosomal arrays three different sample buffers with different bivalent ions were used with a concentration of 1 mM $MgCl_2$, $ZnCl_2$ or $NiCl_2$. In all three sample buffers the concentration of nucleosomal arrays was 0.125 ng/μl, the

concentration of HEPES was 5 mM and the pH was 7.6.

For polylysine attachment a 0.01 % polylysine solution is incubated on the freshly cleaved mica for 30 s and then rinsed off twice with millipore water. Afterwards sample buffer is pipetted onto the substrate and incubated for 15 to 30 min. After incubation the sample is ready for measurements. All necessary liquid volumes for these to methods are the same: 40 μm for each 1 cm^2 of substrate surface area.

For the attachment of DNA to polylysine coated mica the sample buffer contained the following: 1 ng/μl DNA in 10 mM Tris HCl, 1 mM EDTA at pH 7.6. For the attachment of nucleosomes the sample buffer contained 0.125 ng/μl nucleosomal arrays and 10 mM Tris HCl at pH 7.6. To achieve more condensed forms of the nucleosomes varying amounts of bivalent and monovalent ions were added in three different experiments. In all three sample buffers the concentration of nucleosomal arrays was 0.125 ng/μl and the concentration of Tris HCl was 10 mM and the solution was at a pH of 7.6. The first deposition buffer also contained 1 mM $MgCl_2$ and 150 mM NaCl. The second buffer contained 2 mM $MgCl_2$ and 300 mM KCl and the third buffer 2 mM $NiCl_2$ and 250 mM KCl.

Chapter 7

DNA and nucleosomal array Imaging Results

To image nucleosomal arrays in different conformations, the imaging methods, the imaging environments as well as the attachment of the arrays were varied. In section 7.1 the results of contact mode measurements in air are described. In section 7.2 tapping mode in liquid measurement results are shown. Since attachment in all cases is facilitated through the attraction between negative charges of the DNA backbone and the positive charges on the substrate, DNA strands were used for preliminary tests to determine good starting points for the optimization of the nucleosomal array imaging conditions. The two different attachment methods used to attach the DNA and nucleosomal arrays are using either bivalent ions, which form positively charged islands on the negatively charged surface of freshly cleaved mica, or polylsine polymers, which form a positively layer on freshly cleaved mica. Additional parameters that were varied for optimal measurement conditions are: ion concentrations and the pH of the buffer.

7.1 Contact Mode Test Measurements of DNA and nucleosomal arrays in Air

Contact mode imaging (see subsection 6.1.1) is a common and simple to use method to image samples. In the following measurement results for bare DNA are shown, followed by results for the nucleosomal arrays, which were imaged using the best imaging conditions found through the bare DNA

measurements.

7.1.1 DNA

Attachment with Bivalent Ions

Using 1 ng/μl of DNA in 1 mM $MgCl_2$ and 5 mM HEPES at pH 8.0, single DNA strands can be measured as long thin structures with a length of several micrometers, which lies well within expectations for the measured DNA sample. The conformation of the DNA strands on the substrate surface is relaxed with an overall low curvature. This could be caused either by a combing effect of the evaporating water during the preparation of the sample or the diffusion of parts of the strand across the surface to reach an energetically more favorable state. (Figure 7.1)

Measurements with 1 ng/μl of DNA in 5 mM $MgCl_2$ and 5 mM HEPES at pH 8.0 and measurements with 1 ng/μl of DNA in 5 mM $ZnCl_2$ and 5 mM HEPES at pH 8.0 were not successful as only aggregates could be imaged.

Attachment with Polylysine

Polylysine provides a much stronger attachment of DNA to the substrate, but also introduces additional steps in the preparation process and results in a less smooth surface, which leads to lower quality AFM imaging results. Using 1 ng/μl of DNA in 10 mM Tris HCl and 1 mM EDTA at pH 7.6 on polylysine, attachment method, many DNA strands can be measured. The visible strands are numerous and their conformation on the substrate surface is strongly bent, when compared to the DNA attached via magnesium ions from above. The curvature of the structures is similar to the results of other works, which also imaged DNA on mica and reported a persistence length of 50 nm for DNA in solution.[109–111] The much less relaxed conformation of the DNA strands in comparison to attachment via bivalent ions can be explained by the much stronger attachment using polylysine, which prevents relaxation of the structure after surface contact.[109] A less strong attachment should be possible thrugh a reduction of the polylysine concentration.[117] (Figure 7.2 and figure 7.3)

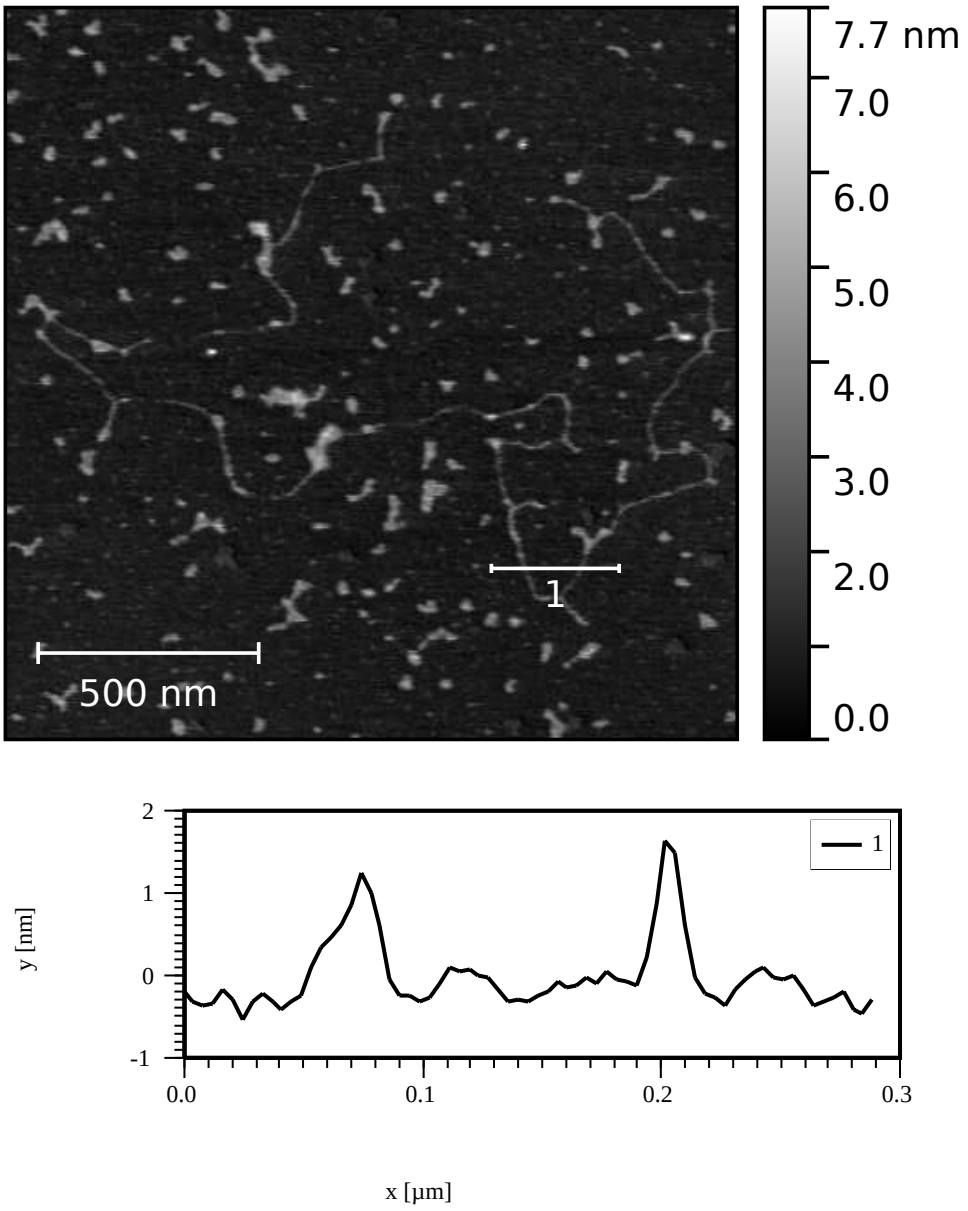


Figure 7.1: A single double stranded DNA. Contact mode in air image of DNA attached with a lower amount of magnesium ions and cross section. The cross section of two DNA strands shows that they are far wider than they are high. Their height is below 2 nm. Measurement parameters: cantilever: MLCT-AUHW(rectangular); scan lines: 512; scan points: 512; scan rate: 1.0 Hz; scan size: 2.1 μm; scan speed: 5.3 μm/s.

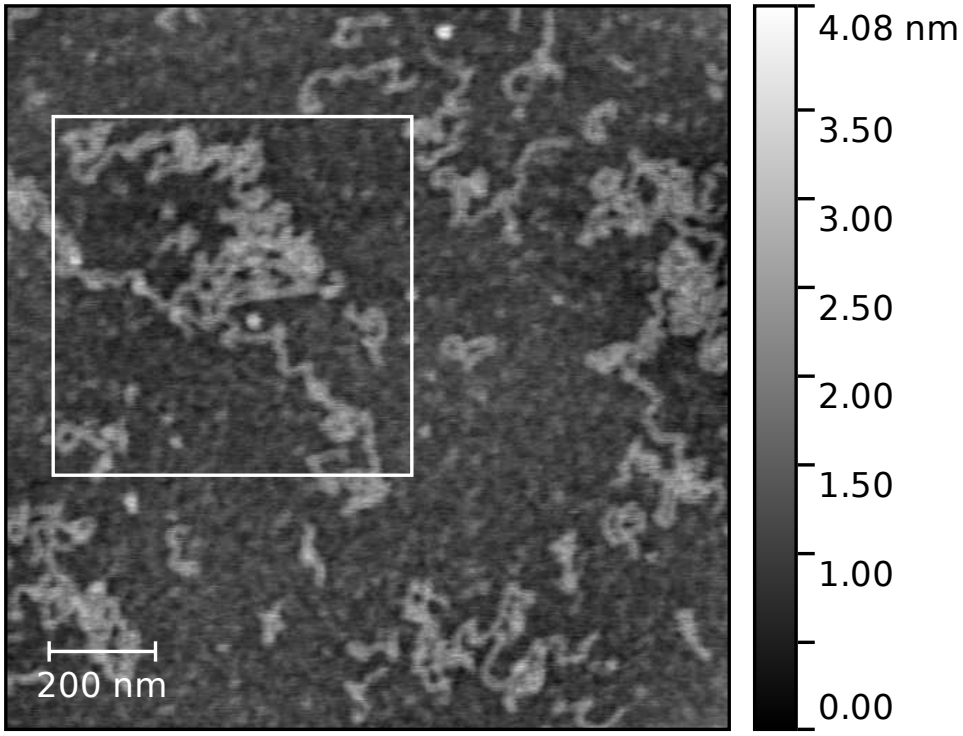


Figure 7.2: Finely bent strands of DNA. Contact mode in air image of DNA attached with polylysine. The rectangle marks a zoomed in region, which can be found in figure 7.3. Measurement parameters: cantilever: MLCT-AUHW(rectangular); scan lines: 512; scan points: 512; scan rate: 1.0 Hz; scan size: 1.3 μm; scan speed: 3.4 μm/s.

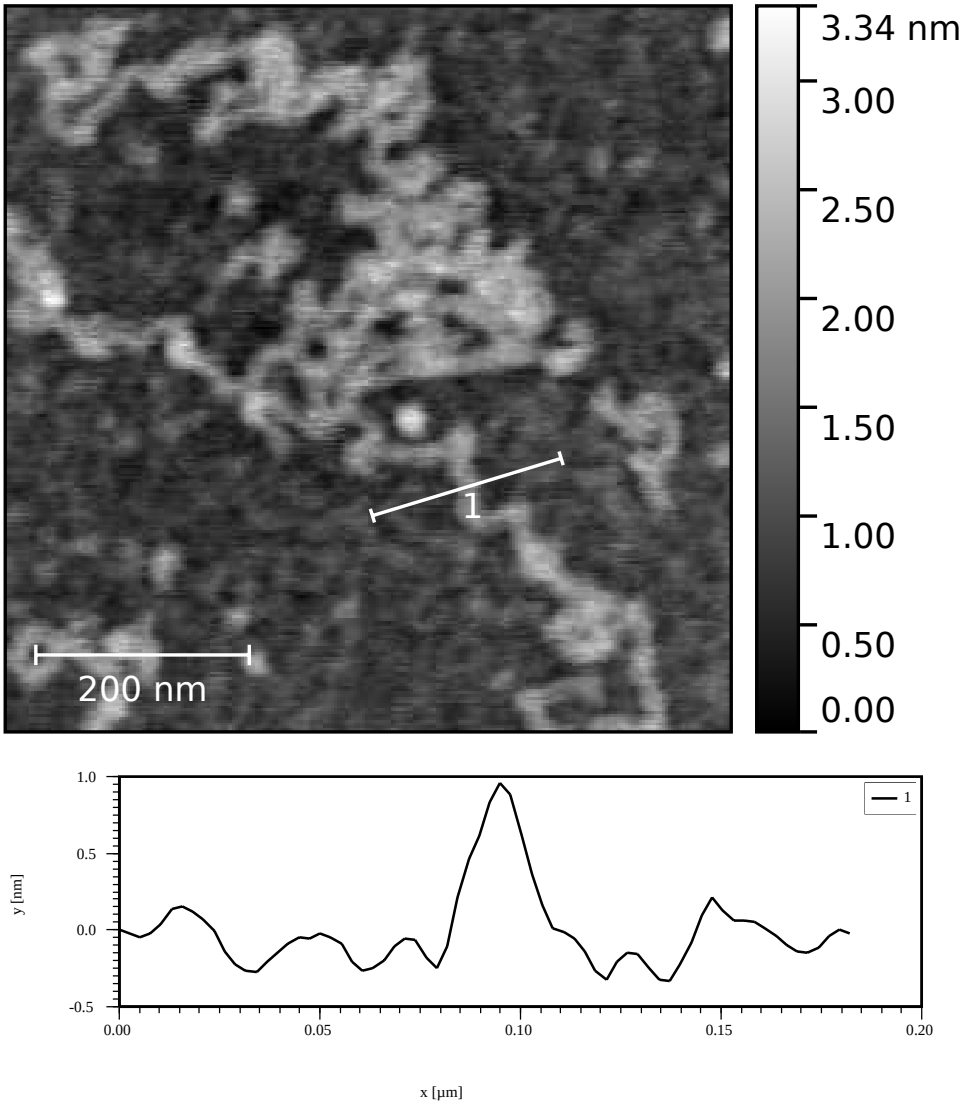


Figure 7.3: Zoomed in region of figure 7.2 with cross section. The height is below 2 nm.

Summary

The negative charges in the backbone of the bare DNA should also be available for attachment of the nucleosomal arrays and thus it appears that both attachment with covalent ions and polylysine should be useful for attaching the nucleosomal arrays. The difference in strength between those methods leads to remarkably different observable conformations of the DNA and should be taken into account for all measurements. Both the relaxation after attachment allowed by the bivalent ions and the fixation of structures close or identical to the conformation immediately after attachment could prove useful for imaging different forms of the nucleosomal array.

7.1.2 Nucleosomal Arrays

Attachment with Bivalent Ions

For 0.5 ng/μl nucleosomal arrays in 1 mM $MgCl_2$ and 5 mM HEPES at pH 8.0 and for 0.5 ng/μl nucleosomal arrays in 1 mM $MgCl_2$, 10 mM KCl, 10 mM Tris HCl and 1 mM EDTA at pH 7.6 only aggregates could be imaged.

Attachment with Polylysine

Because polylysine is used in this protocol, the attachment of the nucleosomal arrays should be much stronger than for methods using bivalent ions. In the sample with 0.25 ng/μl nucleosomal arrays in Tris/EDTA buffer at pH 7.6 on polylysine coated mica curled structures can be observed. These structures are shorter than the free DNA of the nucleosomal arrays would be and longer than the condensed form of the nucleosomal arrays. Lengths between 400 and 600 nm are observed, which are in agreement with the length of 428 nm expected for the beads-on-a-string conformation of the nucleosomal array. The bare DNA would be more than three times larger at 1640 nm and the condensed fiber would be expected to be an order of magnitude shorter. If the two structures in the middle are seen as one object, it's length of 900 nm lies inbetween the length for the beads-on-a-string conformation and the bare DNA and could represent an incomplete nucleosomal array. The height of the structures is roughly 1 nm and thus smaller than expected. (Figure 7.4) This is similar to the samples prepared using

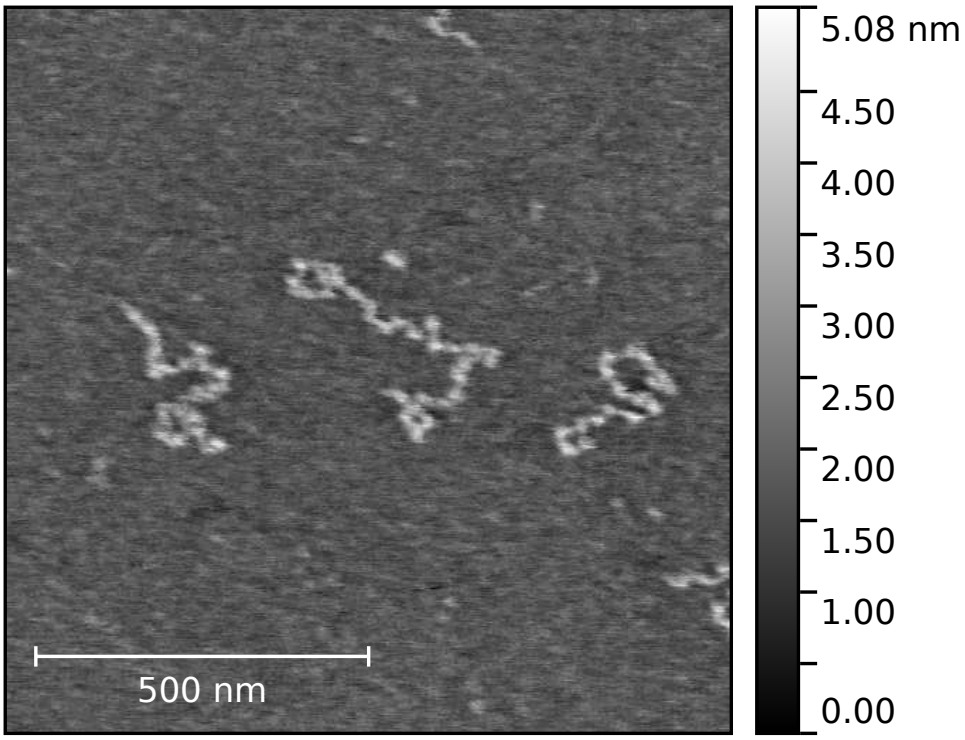


Figure 7.4: Multiple structures of a length between roughly 400 and 600 nm. Contact mode in air image of nucleosomal arrays attached with polylysine and cross section. From left to right the estimated lengths are 600 nm, 400 nm, 500 nm and 500 nm. The middle two structures could also be seen as one object with a length of 900 nm. Measurement parameters: cantilever: MLCT-AUHW(rectangular); scan lines: 512; scan points: 512; scan rate: 1.0 Hz; scan size: 1.1 μm; scan speed: 2.7 μm/s.

magensium ions and Tris/EDTA buffer (Figure 7.1) and likely also caused by deformation of the sample by the cantilever tip and/or by being dried out. The nucleosomal arrays are found in a beads-on-a-string conformation instead of a condensed, because in the buffer no ions are provided, that would promote the condensation of the molecule complexes. It was not possible to determine the number of nucleosomes due to the quality of the images.

Summary

It was possible to image the beads-on-a-string conformation of the nucleosomal arrays, but not the condensed conformation. It was found that a pH of

7.6 is necessary to prevent a dissociation of the DNA-histone complexes in solution. While the length and width of the structures were close to the expected values, the heights of the structures measured for both conformation are significantly lower than expected. Possible reasons for the low height are dehydration because of the measurement in air and the deformation of the sample by the cantilever.

7.2 Tapping Mode Measurements of DNA and nucleosomal arrays in Liquid

In tapping mode measurements (see subsection 6.1.1) the cantilever tip only touches the surface intermittently and minimizes lateral forces, which leads to less mechanical stress of the sample and less deformations. Imaging in liquid obviously matches the natural environment of the chromatin more closely than imaging in air, which should also lead to structures closer to expectations. A drawback is that microcrystals, protein aggregates and other contaminations aren't removed in a rinse off step, that was necessary for the contact mode in air measurements. Thus all solutions used in the measurement have to be purified using either filters or centrifugation.

7.2.1 DNA

Attachment with Magnesium Ions

In the sample with 1 ng/μl DNA in 1 mM $MgCl_2$ and 5 mM HEPES at pH 7.6 the DNA can be seen as long and thin structures. The lengths of the structures are in the expected range of lengths for the sample DNA. Like in the in air measurments with magnesium ion attachment the strands are in a similar conformation to DNA in air measurements with magnesium attachment. As combing effects through water evaporation can be ruled out, because the sample is never dried during preparation or measurement, the structure is likely a consequence of the diffusion of the strands across the substrate after attachment. This is only possible because the attachment the magnesium ions provide is weak. (Figure 7.5)

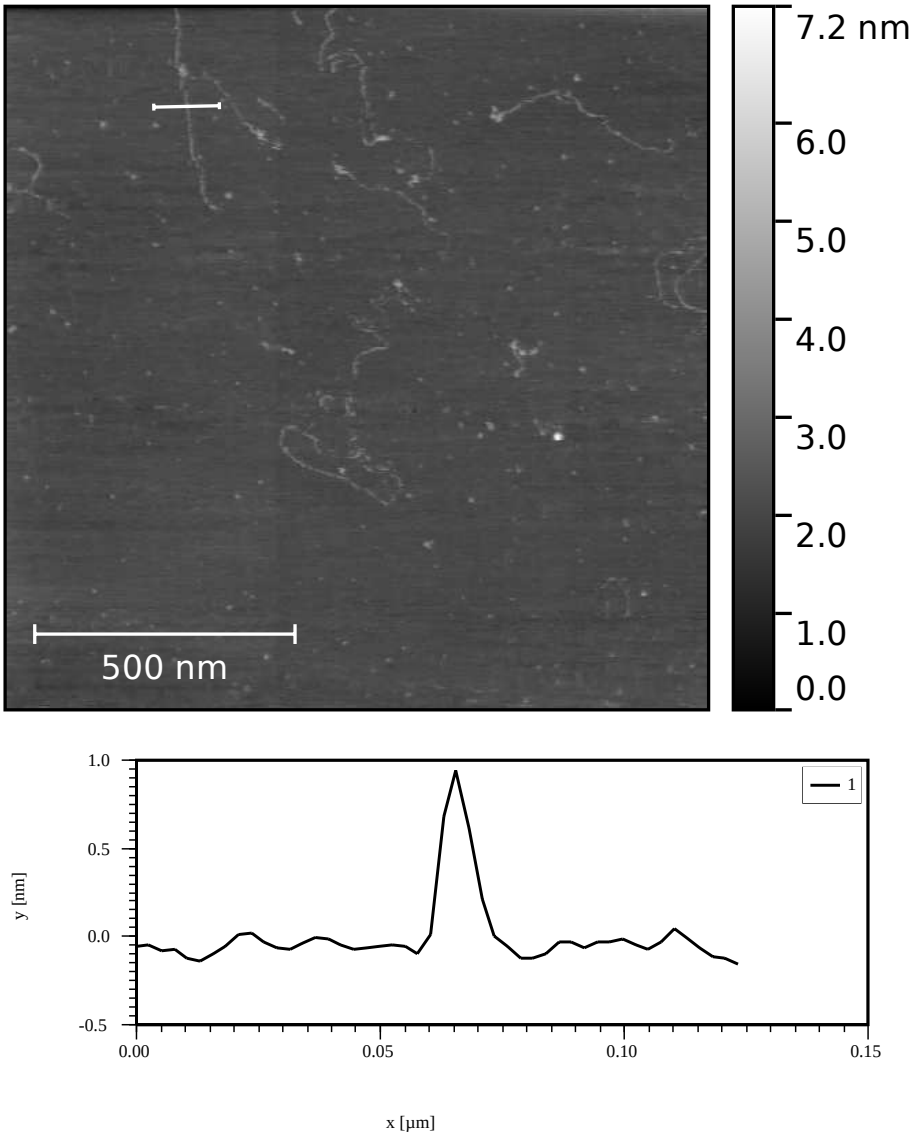


Figure 7.5: Multiple single DNA strands. Tapping mode in liquid image of DNA attached with magnesium ions with cross section. The height and width of the DNA strands are similar to the result of the analog contact mode measurement in air (see figure 7.1). Measurement parameters: cantilever: BL-AC40TS; drive amplitude: 2.0 V ; drive frequency: 20.1 kHz; amplitude setpoint: 400 mV; estimated cantilever amplitude (see subsection 6.3.3): 46.4 nm; scan lines: 512; scan points: 512; scan rate: 0.5 Hz; scan size: 1.3 μm; scan speed: 1.7 μm/s.

Attachment with Polylysine

Attaching DNA strands using polylysine in a sample with 1 ng/μl DNA in 10 mM Tris HCl, 1 mM EDTA at pH 7.6 on polylysine coated mica and measuring them with tapping mode imaging in liquid, leads to similar results to contact mode measurements in air with the same attachment method. The DNA strands are clearly visible and finely curled. Analog to the in air measurements the strong attachment also prevents diffusin of the DNA strands across the substrate surface. (Figure 7.6)

Summary

The results for measuring DNA using tapping mode in liquid are analog to the results of the contact mode measurements in air and the compositions of the buffers can be adopted without major changes.

7.2.2 Nucleosomal Arrays

Attachment with Bivalent Ions

For 0.125 ng/μl nucleosomal arrays in 1 mM $MgCl_2$/$ZnCl_2$/$NiCl_2$ and 5 mM HEPES at pH 7.6 no attempt at imaging the nucleosomal arrays attached with bivalent ions was successful. Several different ions were used: magnesium, zinc and nickel. The attachment of the molecule complexes using bivalent ions appears to be too weak.

Attachment with Polylysine

Using the stronger attachment provided by polylysine with 0.125 ng/μl nucleosomal arrays in 10 mM Tris HCl at pH 7.6 on polylysine coated mica, the nucleosomal arrays can be imaged. The beads-on-a-string conformation is clearly visible. Both the nucleosomes and the linking DNA can be seen. The number of nucleosomes is lower than expectated. Furthermore loops of DNA strands can be seen at positions where histone octamers are missing. These loops can be explained as nucleosomes attaching to the surface and then losing the histone octamer. This means that the number of nucleosomes in the arrays in solution has to be higher. The histone octamers could be removed either through interaction with the surface or with the cantilever tip. Multiple measurements have shown a further reduction of

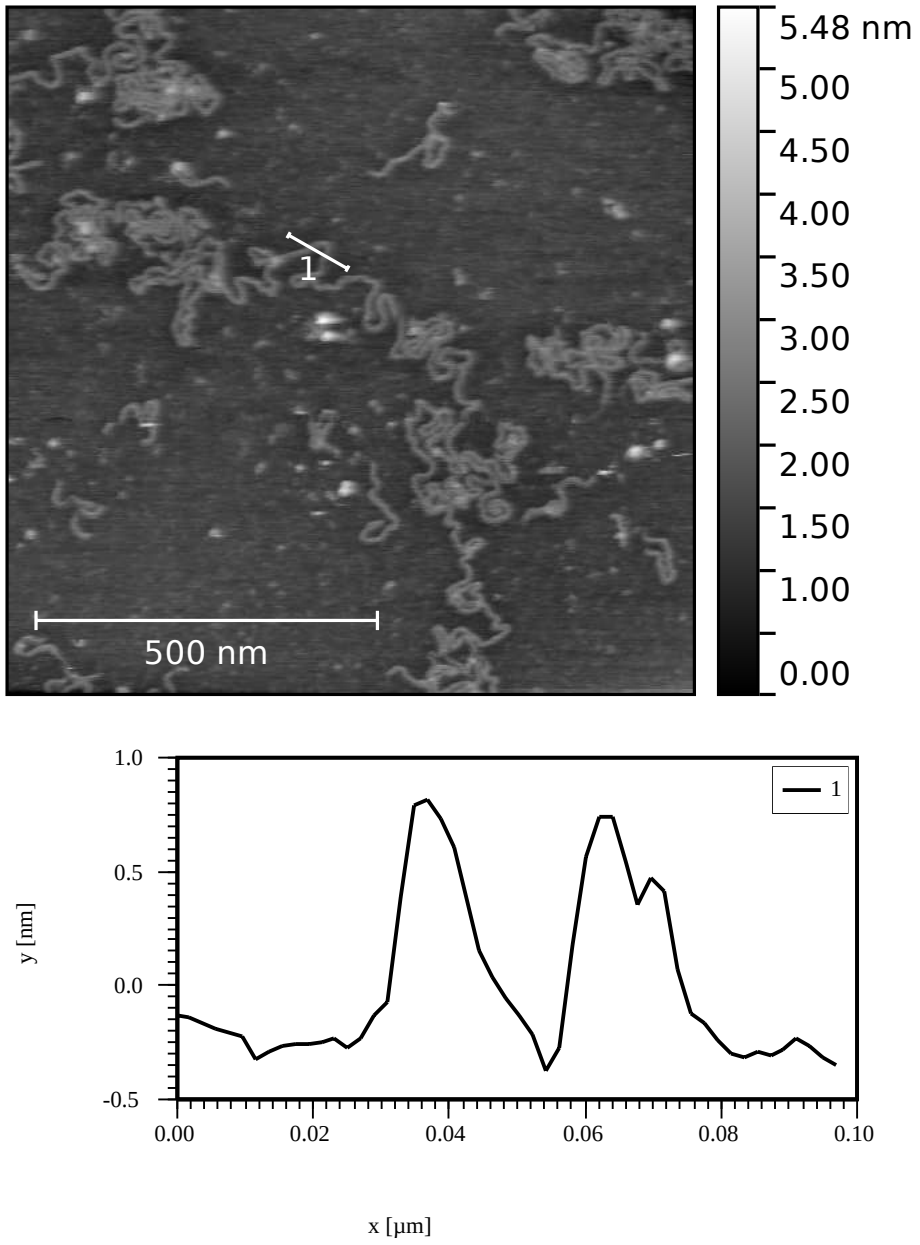


Figure 7.6: Multiple finely bent DNA strands. Tapping mode in liquid image of DNA attached with polylysine with cross section. This result is very similar to the corresponding contact mode in air measurement (see figure 7.2). The DNA double strands are strongly bent. Measurement parameters: cantilever: BL-AC40TS; drive amplitude: 1.5 V ; drive frequency: 25.5 kHz; amplitude setpoint: 400 mV; estimated cantilever amplitude (see 6.3.3): 46.4 nm; scan lines: 512; scan points: 512; scan rate: 1.0 Hz; scan size: 1.0 μm; scan speed: 2.5 μm/s.

the number of histone octamers, but it is not clear if the nucleosomes are lost over time or with increasing contacts with the cantilever tip. As in previous measurements with polylsine attachment the overall conformation of the molecule complexes has a high curvature and is likely close to the conformation directly after attachment. The curvature of the DNA between the nucleosomes is also similar to measurements of DNA attached to polylysine coated mica by Rivetti et al.[109] In comparison to in air contact mode measurements not only the resolution is better, but also the measured height of the nucleosomal arrays at 3 nm is twice as high. (Figure 7.7)

Attachment with Polylysine and High Sodium Ion Concentrations

To measure the condensed 30 nm fiber conformation of the nucleosomal arrays, magnesium and sodium ions have been added to the buffer for this sample: 0.125 ng/μl nucleosomal arrays in 1 mM $MgCl_2$, 150 mM NaCl and 10 mM Tris HCl at pH 7.6 on polylysine coated mica. The concentrations chosen should be sufficient to condense the molecule complexes when in solution.[115] Using this method, still only beads-on-a-string conformations of the complexes can be found. The conformations are more dense than in the measurements without condensation promoting ions in the buffer, but still not dense enough for the condensed form. The nucleosomes are still separated from each other. Considering the strong attachment polylysine provides, the imaged structures are likely to be projections of the structure of the nucleosomal arrays in solution onto the substrate surface. Thus it appears that some degree of compaction has occured, but as the width is larger than 30 nm, condensation into a 30 nm fiber was not achieved. It has also to be noted that the number of nucleosomes in the imaged structures is lower than 25. Empty stretches of DNA with lengths longer than the expected linker length of 16 nm also confirm this. Interactions of the histone octamers with the cantilever tip or substrate surface are possibly the cause for the missing nucleosomes. (Figure 7.8)

Attachment with Polylysine and High Potassium Ion Concentrations

For 0.125 ng/μl nucleosomal arrays in 2 mM $MgCl_2$, 300 mM KCl and 10 mM Tris HCl at pH 7.6 on polylysine coated mica no structures could be

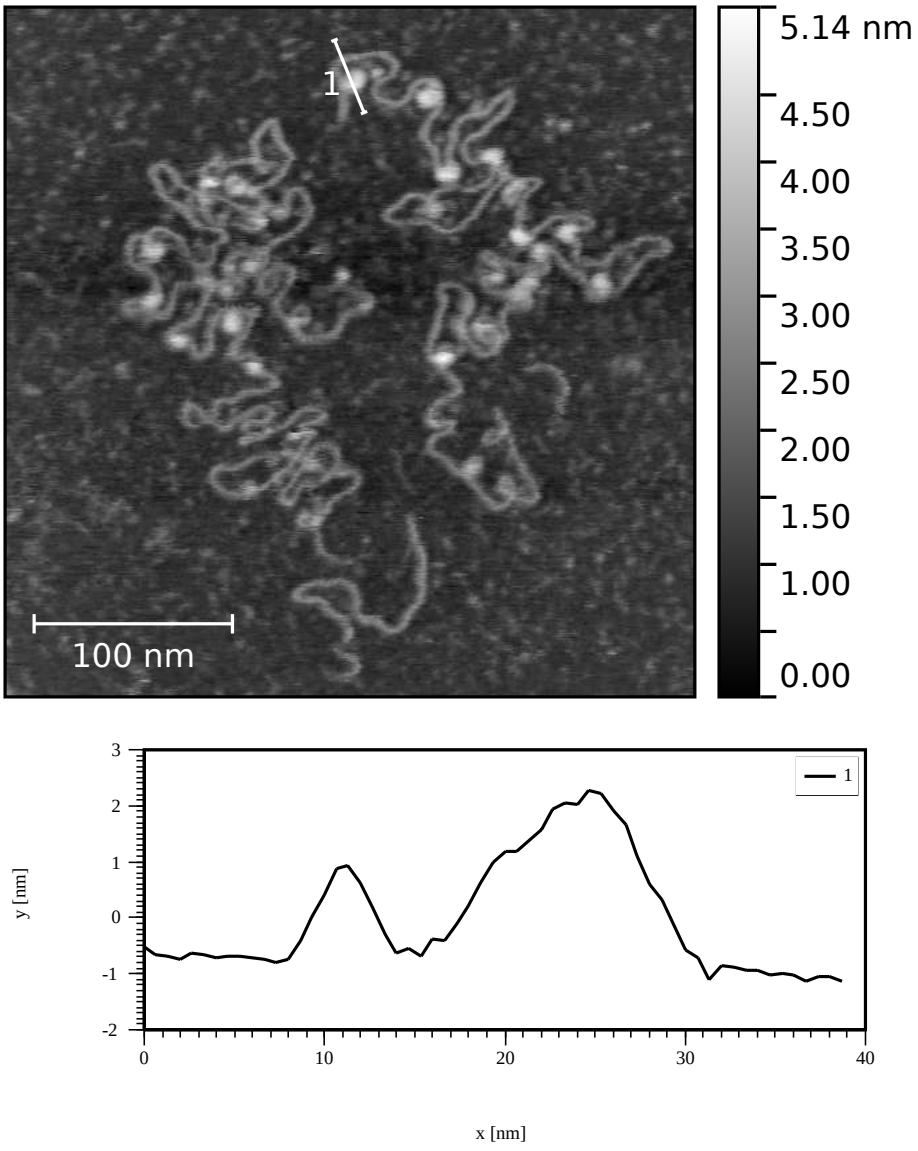


Figure 7.7: nucleosomal arrays in the beads-on-a-string conformation. Tapping mode in liquid image of nucleosomal arrays attached with polylysine with cross section. The number of nucleosomes is below 25 and loops of DNA can be seen where histone octamers are missing. In the cross section the width and height difference between the DNA double strand and the nucleosome is clearly visible. Measurement parameters: cantilever: BL-AC40TS; drive amplitude: 2.1 V ; drive frequency: 20.9 kHz; amplitude setpoint: 400 mV; estimated cantilever amplitude (see 6.3.3): 46.4 nm; scan lines: 512; scan points: 512; scan rate: 1.0 Hz; scan size: 3.4 μm; scan speed: 8.6 μm/s.

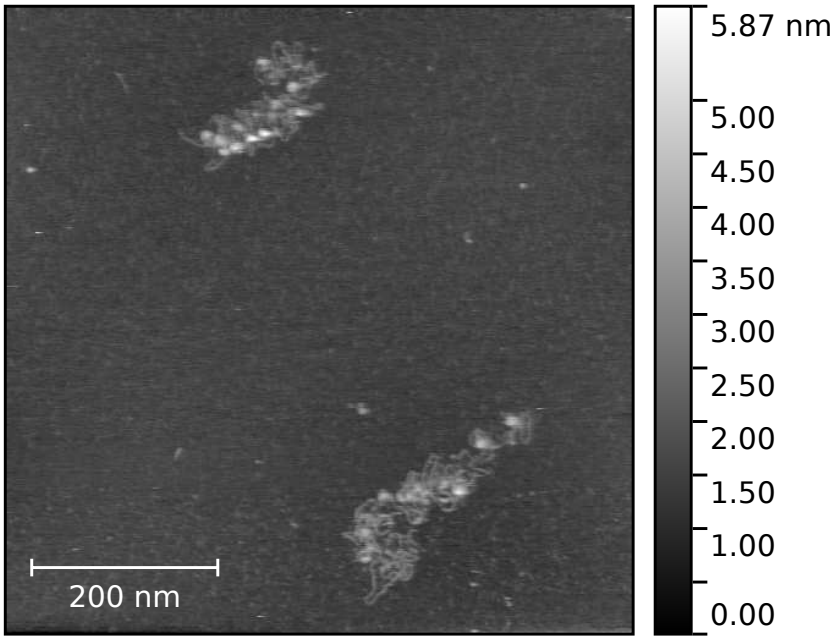


Figure 7.8: nucleosomal arrays in the beads-on-a-string form. Tapping mode in liquid image of nucleosomal arrays attached with polylysine and in a high concentration of sodium ions. Stronger condensation of the nucleosomal arrays than in figure 7.7. The numerous stretches of DNA that are longer than the linker length of 16 nm and the lower than expected number of visible nucleosomes suggest that a substantial amount of nucleosomes is missing in the imaged structures. Measurement parameters: cantilever: BL-AC40TS; drive amplitude: 1.5 V ; drive frequency: 29.2 kHz; amplitude setpoint: 310 mV; estimated cantilever amplitude (see 6.3.3): 36.0 nm; scan lines: 512; scan points: 512; scan rate: 0.5 Hz; scan size: 670 nm; scan speed: 840 nm/s.

successfully imaged.

Attachment with Nickel and High Potassium Ion Concentrations

Nickel is another bivalent ion that can be used to attach the nucleosomal arrays. The concentration of potassium was slightly lowered in comparison to the previous protocol: 0.125 ng/μl nucleosomal arrays in 2 mM $NiCl_2$, 250 mM KCl and 10 mM Tris HCl at pH 7.6 on polylysine coated mica. Blurred long structures with a length of 100 nm or more can be seen, which are about three times longer than the predicted lengths for the condensed nucleosomal arrays. With 15 nm the width of the structures is about half the expected value for the 30 nm fiber(see section 6.2.1). Thus the observed dimensions of the structures are more likely to correspond to nucleosomal arrays in the beads on a string conformation though. Furthermore the structures are interrupted as if the imaged nucleosomal arrays slide on the surface along with the movement of the cantilever. The blurry nature of the structures is thus likely caused by the nucleosomal arrays being attached only loosely to the surface. An explanation for the mobility of the nucleosomal arrays is the weak attachment via nickel ions. The smaller structures in this image all show the same orientation. This is likely caused by a contaminated or broken tip. (Figure 7.9)

Summary

Bivalent ions alone do not provide a strong enough attachment in liquid, so that DNA-histone complexes slide across the surface. While it was easily possible to image the beads-on-a-string conformation of the nucleosomal arrays using polylysine, it was not possible to image the 30 nm fiber. When a combination of polylysine and bivalent ions was used, some degree of condensation could be observed, but 30 nm fibers were not found. Higher concentrations of the ions, that were hoped to produce more stable condensed nucleosomal arrays, prevented the attachment of the nucleosomal arrays. The stability of the nucleosomes was low as no nucleosomal arrays with the full number of nucleosomes could be imaged. It is not clear wether the substrate or the cantilever tip are responsible for missing histone octamers. Structures that possibly represent nucleosomal arrays sliding across the surface could be observed when nickel ions were used for the attachment.

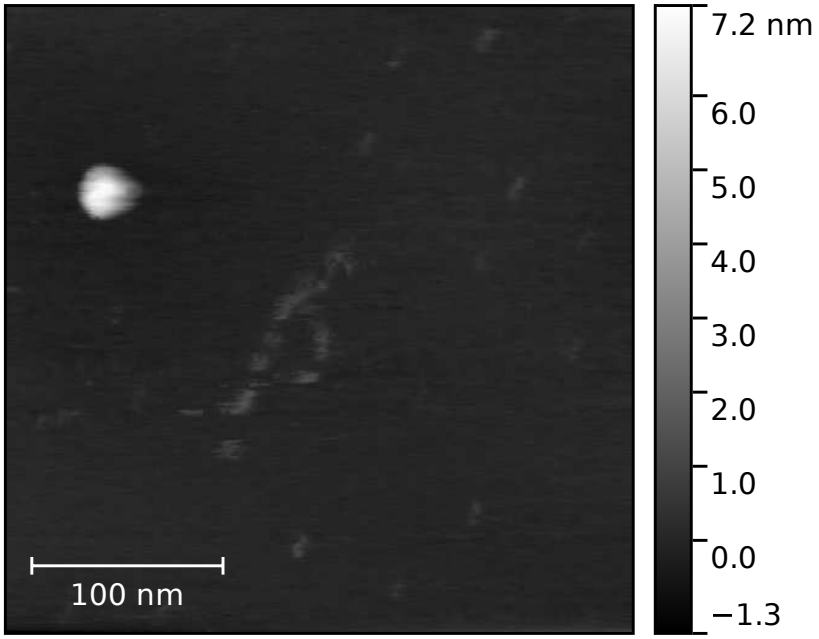


Figure 7.9: Long structures. Tapping mode in liquid image of nucleosomal arrays attached with nickel ions and in a high concentration of potassium ions. Neither length nor width match the 30 nm fiber. Measurement parameters: cantilever: BL-AC40TS; drive amplitude: 2.0 V ; drive frequency: 25.5 kHz; amplitude setpoint: 240 mV; estimated cantilever amplitude (see 6.3.3): 27.8 nm; scan lines: 512; scan points: 512; scan rate: 1.0 Hz; scan size: 326 nm; scan speed: 817 nm/s.

7.3 Conclusions

Using the presented protocols DNA strands and nucleosomal arrays in the beads-on-a-string conformation could be measured when attached with either bivalent ions or polylsine and using either contact mode in air imaging or tapping mode in liquid imaging methods. 30 nm fibers could not be imaged under any imaging conditions and imagind modes. Attachment strength emerged as the most important difference between the two attachment methods as expected. While the attachment through bivalent ions is relatively weak, the attachment through polylysine layers is so strong that attached molecules and molecules are fixed in their conformation upon contact with the substrate.

Even though multiple imaging methods, environments and various buffer compositions were used, a summary of parameters, conditions and concentrations for good measurement results can be given:

- **Sample concentration:**

  1 ng/μl for DNA and 0.125 ng/μl for nucleosomal arrays

- **Incubation times:**

  5 min for in air measurements and 15 min for in liquid measurements

- **Bivalent ion attachment buffer for DNA:**

  1-2 mM $MgCl_2$, 10 mM Tris HCl, 1 mM EDTA at pH 7.6

  This buffer can be used for DNA in air and in liquid measurements.

- **Polylysine attachment buffer:**

  10 mM Tris HCl, 1 mM EDTA at pH 7.6

  This buffer can be used for DNA in air and in liquid measurements and for beads-on-a-string nucleosomal array in air and in liquid measurements.

The two used attachment methods demonstrated, that the strength of the attachment they provide has a large influence on the structures that are imaged: attachment through bivalent ions is weak, which makes the measurement of large molecule complexes in liquid difficult, because they get displaced by interactions with the moving cantilever. On the other hand

the weaker attachment also means, that less stable conformations such as the condensed form of the nucleosomal arrays might retain their form when attached with bivalent ions. The strength of the polylysine attachment can be seen as an advantage, because it allows measurements of even very large molecule complexes in liquid and also prevents post-attachment rearrangements of conformations and thus conserves the conformation at the moment of the attachment.

Even though DNA and nucleosomal arrays could be measured and were found to possess the expected sizes in the twodimensional plane, the heights were in all cases much lower than expectations. Both environmental conditions, such as measurements in air, and mechanical deformation through the cantilever are likely the cause for the reduced height.

Chapter 8

Summary and Outlook

Chromatin (see section 6.2) has two important functions: the compaction of the DNA so that it fits inside the nucleus and also the control of the compaction so that the DNA remains accessible for processes like replication, transcription or repair. The most basic level of compaction are the nucleosomes and the highest level are the metaphase chromosomes. The second presumed compaction step after the formation of the nucleosomes involves the condensation of the nucleosomes into the so called 30 nm fiber. Several structures of the 30 nm fiber have been proposed with varying properties such as different diameters or independence or dependence of the diameter on the DNA linker length. Because the atomic force microscope is capable to resolve even doublestranded DNA well it can be expected to be a useful tool for the investigation of the basic compaction levels of the chromation.

The necessarily first step in the investigation of the beads-on-a-string structure and the 30 nm fiber (see section 6.2) with the AFM is being able to image those structures with the AFM. To image these structures, the molecule complexes have to be attached to the substrate surface (see subsection 6.2.2). One attachment principle utilizes the negative charge of the DNA backbone by creating a positively charged substrate surface, which can be achieved either with divalent cations or positively charged polymers. Because of the different attachment properties such as attachment strength of the ions and the polymers it is of course necessary to determine for both how useful they are for the imaging both the beads-on-a-string conformation and the 30 nm fiber conformation of the nucleosomal arrays. Since both attachment methods rely on the negative charges of the DNA backbone it

is possible to use doublestranded DNA as a more simple testing system. Two imaging methods (see subsection 6.1.1) and environments (see subsection 6.1.2) were chosen: contact mode in air measurements as the easiest and most straightforward way to image the DNA and chromatin structures and tapping mode in water measurements to provide a more gentle method in a more natural environment. Thus DNA, beads-on-a-string and the 30 nm fiber were to be imaged using either divalent cations or positively charged polymers for attachement and either contact mode in air or tapping mode in water as measurement methods.

For contact mode in air measurements (see section 7.1) it was possible to establish sample preparation and measurement protocols for DNA and nucleosomes in the beads-on-a-string conformation. For DNA both attachment methods worked similarly well, while only the positively charged polymers allowed imaging of the beads-on-a-string. Using the tapping mode in liquid measurements (see section 7.2) doublestranded DNA could be imaged with both attachment methods. Beads-on-a-string could also be measured using positively charged polymers for the attachment. But the 30 nm fiber could not be imaged. The images captured using attachment via divalent ions showed no easily identifiable structures. The observed structures were most likely caused by weakly attached nucleosomal arrays in a beads-on-a-string conformation being pushed across the surface. One remaining problem is the instability of the nucleosomes, that is caused by interactions with the cantilever tip and/or interactions with the positively charged substrate surface.

Nucleosomal arrays in the beads-on-a-string conformation could be easily imaged, but imaging of the 30 nm fiber has yet to be achieved. The measured structure of the 30 nm fiber can only be expected to be close to in vivo when using in liquid tapping mode imaging. To achieve this several possibilities exist: the structure of the fiber and nucleosomes could be stabilized by crosslinking or the fiber could be attached with other methods such as tethering. Also an AFM, which can provide a measurement that is as gentle as possible with softer cantilevers, smaller cantilever amplitudes and new measurement modes like peak force tapping, should improve the results. A further step following from the imaging of the 30 nm fiber in liquid is the investigation of dynamics and interactions of the 30 nm fiber. A possible combination with optical methods could allow even more detailed insights

in the structure, the dynamic behavior and interactions of both the 30 nm fiber and the beads-on-a-string structure.

Part III

Appendix

Chapter 9

Software for Single Molecule Force Spectroscopy

In the following software used to measure and process single molecule force spectroscopy data is described (section 9.1). Then software used to simulate force spectroscopy data is shown (section 9.2). In the last section the functions to determine logL can be found (section 9.3). For every software an accompanying function schematic is included.

9.1 Experimental

The measurement software for the AFM constant force experiments (Figure 9.1) has to fulfill several tasks to perform the measurements as described in subsection 2.2.1: regularly repositioning the cantilever to optimize the amount of accessible polymer ends with still reactive groups, contact with the surface area to allow the attachment of polymers, exclusion of noncovalently bound polymers, detection of covalently attached bonds and finally detection of the rupture of those covalently bound polymers. Moving the cantilever to a new location after 100 surface contacts has proven to be sufficient for the purpose of providing contact areas with reactive polymer ends. This position is chosen randomly since repeatedly choosing the same location is unlikely in the scope of this experiment. In earlier experiments it has been shown,[38] that almost all noncovalently bound polymers detach below a force of 1.5 pN and this is why the polymers are stretched to 1.5 pN first, before they are then held at their target force. This step is not neces-

sary for measurements with a target force higher than 1.5 pN. Bond rupture is detected as soon as the force on the cantilever is below a threshold value of 20% of the target force. During the whole process the elapsed time and the distance of the cantilever are monitored, if the threshold value for the distance is exceeded the algorithm concludes that no polymer is attached to the cantilever and the measurement is aborted. If the threshold value for the time is exceeded, the measurement is aborted to prevent a measurement taking too long.

While constant force data is relatively straighforward to analyze (Figure 9.2), some processing is still necessary to arrive at the data set of corresponding holding forces and times to rupture as seen in figure 2.4 in subsection 2.2.1. In the used AFM setup a the measured force is dependent on the distance of the cantilever fomr the substrate surface (Figure 9.3). Since the cantilever is mostly at the same position during the measurment this problem is mostly a problem for the analysis of the data. It can be corrected by fitting a linear function to the force-distance plot of the data (Figure 9.4).

9.2 Simulation

The simulation of data (see section 2.8) was not only necessary to test the data analysis software, but also to test the behavior of the parameter estimation for different bond rupture models and for different sets of data (see section 3.2). Another purpose of simulations is to allow comparisons between analysis results and experimental data as well as comparisons between the analysis results of constant force and constant speed data (see section 3.4).

To simulate constant force measurement data (Figure 9.5), for every data point a force is selected from a flat distribution of forces with set limits. This random force is then used to calculate a rupture rate, which then determines a exponential distribution of rupture times. The distribution is used to select a random rupture time. For the linear-cubic model no analytical formula is available and it has to be evaluated numerically. Careful consideration of the integration limits and sampling point density are necessary to balance accuracy and computational cost.

The simulation of constant speed data (Figure 9.6) selects random loading rates from a flat loading rate distribution and then generates a random

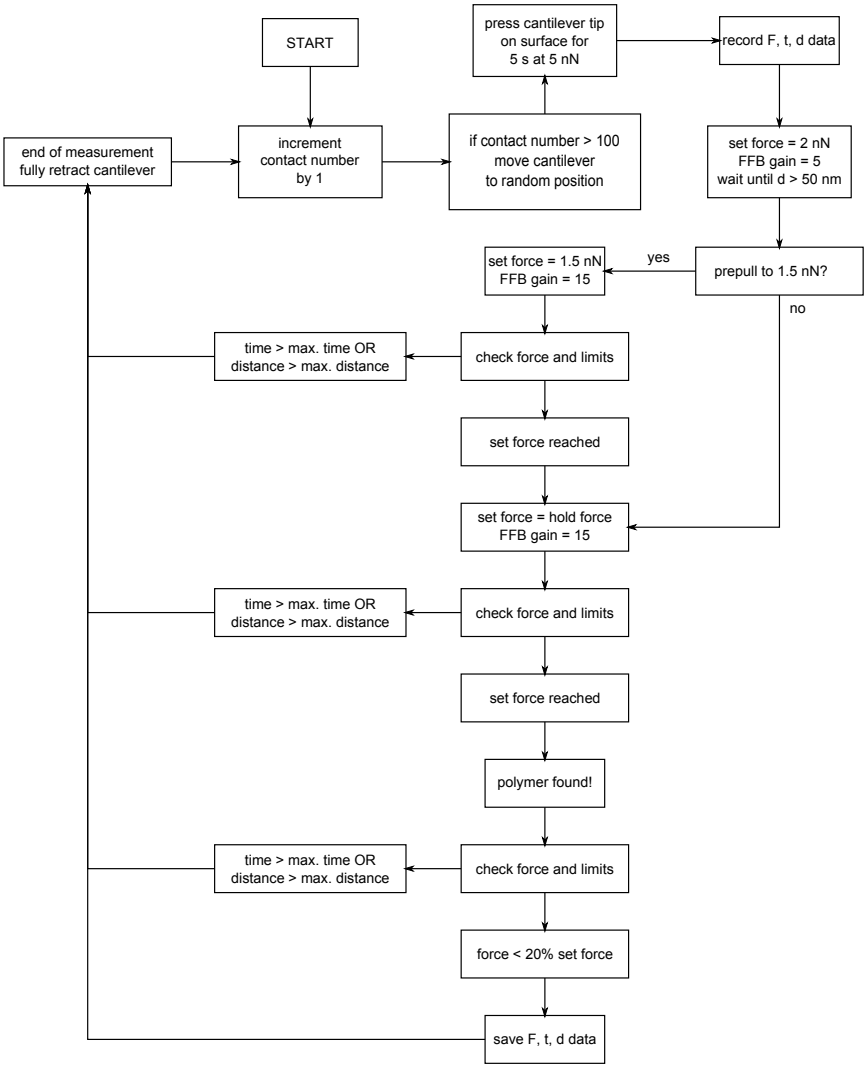


Figure 9.1: Constant Force Measurement algorithm scheme. This scheme includes the movement of the cantilever to new areas after 100 surface contacts, prepull for low holding forces to exclude non-covalently bound polymers and detection of covalently attached polymers.

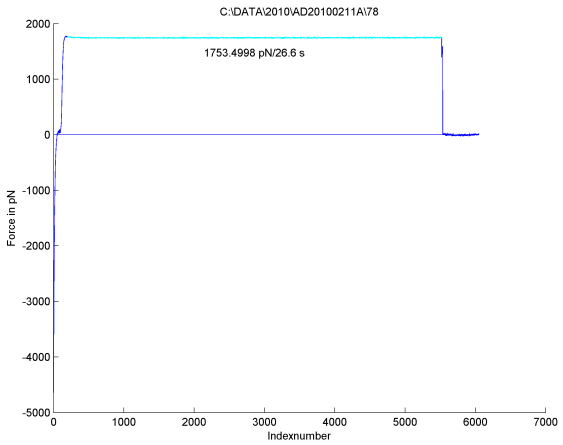


Figure 9.2: Constant force measurement data. The data is shown after correction, detection of polymer holding force and holding time. The detected polymer holding range of the F/t curve marked in cyan.

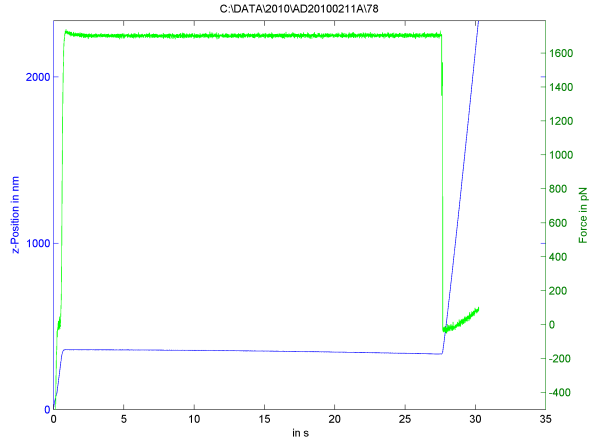


Figure 9.3: Raw data from a constant force measurement. Raw data with z-position over time in blue and force over time in green. A correction of the z-position by the feed back mechanism keeping the holding force constant can be seen.

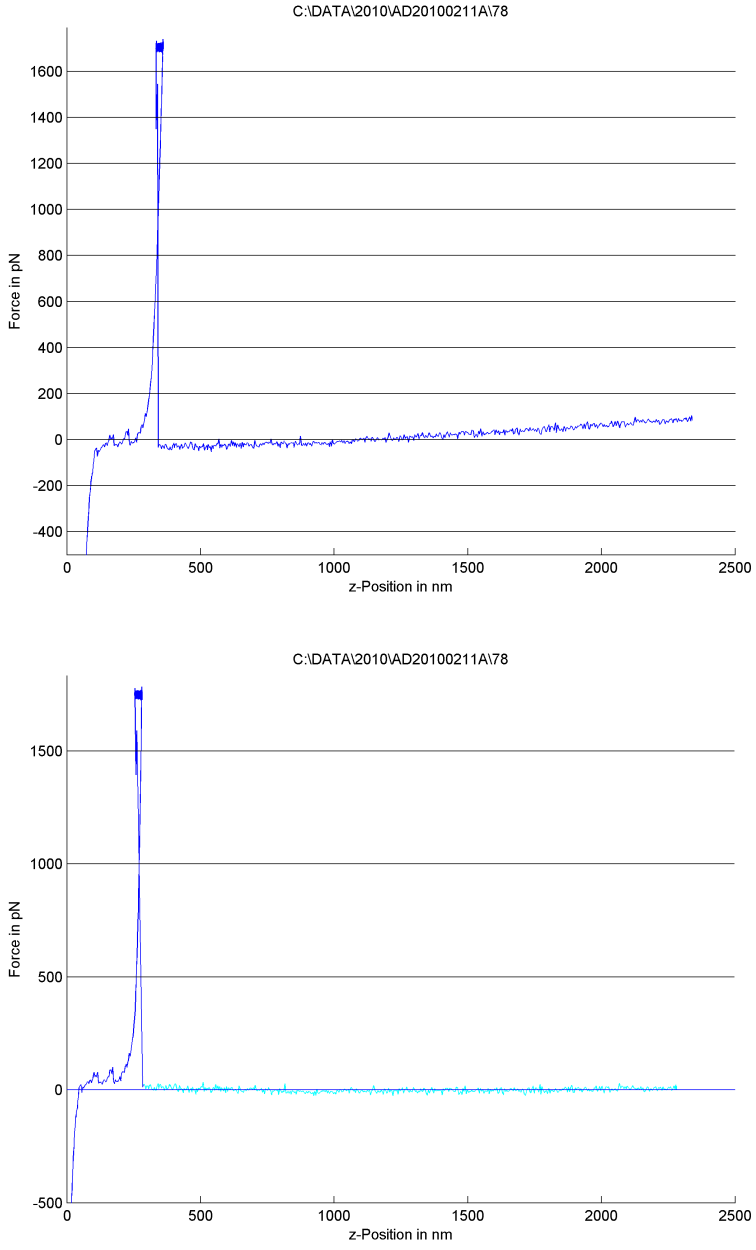


Figure 9.4: Constant force measurement data correction. A distance dependent change of the observed force is corrected by a line fit to the range of the force versus distance graph after the polymer holding section. The upper diagram is uncorrected data and the lower diagram is corrected data with the range used for the line fit marked in cyan.

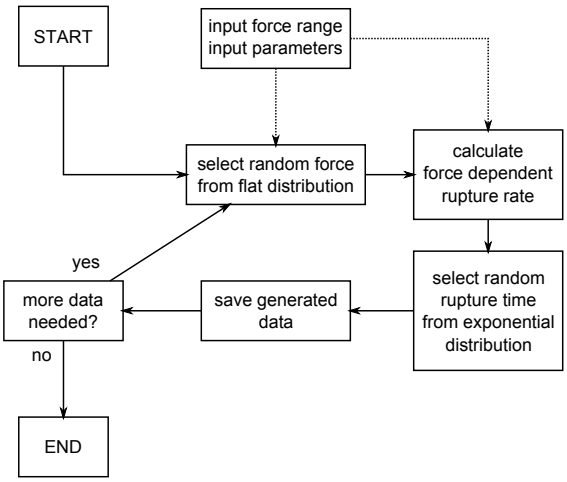


Figure 9.5: Constant force data simulation scheme. While there are analytical formulas to calculate the rupture rates for the BELL and the CUSP model, numerical integration is necessary to do the same for the linear-cubic model. Despite this difference, the general structure of the simulation is identical.

rupture force according to the used rupture model and bond potential parameters. As the shape of the distribution of rupture forces is more complex than a simple exponential distribution, a numerical approach was used. The numerical approach was faster to implement and could be used for all three bond rupture models as no analytical function of the rupture force distribution for the linear-cubic model was available.

9.3 logL Functions

Bayesian data analysis with the nested sampling algorithm uses likelihood functions to determine the likelihood of the fit of a set of parameters to the data (see section 2.7). In the case of the constant force analysis for the BELL and CUSP model (Figure 9.7) in subsections 3.1.3 and 3.1.4, the model is used to calculate the rupture rate for the holding force of every data point. Afterwards the likelihood value for that data point is calculated and added to the total logL. This process is repeated for every data point. No analytical solution is available for the linear-cubic model (see subsection 3.1.5) and the higher computational effort following from that has to be countered by time saving measures (Figure 9.8). At the cost

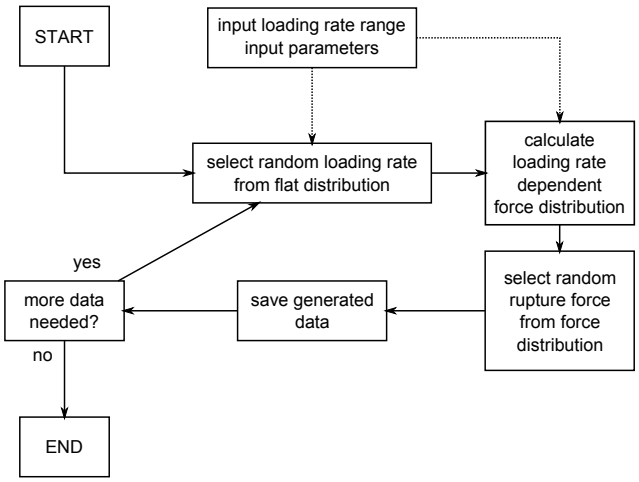


Figure 9.6: Constant speed data simulation scheme. Analytical formulas are used for the BELL and the CUSP model, while numerical integration has to be used for the linear-cubic model. With the exception of the calculation method of the rupture force distribution function, the simulation is identical for all three models.

of lower accuracy an array of rupture rates over a range of forces found in the experimental data is calculated once at the beginning. For every data point the rupture rate is calculated through interpolation which is much faster than the full calculation using the linear-cubic model. An amount of 20 sampling points has been found to be sufficient for the experimental data.

For the analysis of constant speed experiments (Figure 9.9) in section 3.3, the data has to be converted into a 2D histogram with rows for the loading rate and colums for the rupture forces (see figure 2.22 in subsection 2.7.3). This allows not only a faster analysis of the data but also a normalization of the likelihood results. The normalization is needed because the distribution of rupture forces is normalized only for the full force range and because there are necessary cutoffs for low and high forces (see figure 2.18 in subsection 2.6.2). This variable area in the observed rupture force range leads to a distortion of the likelihood results and can only be prevented by normalization of this area. For every row a correction factor is calculated, which is the area below the likelihood function in the experimental force range. Then for each bin the likelihood value is calculated according to its loading rate

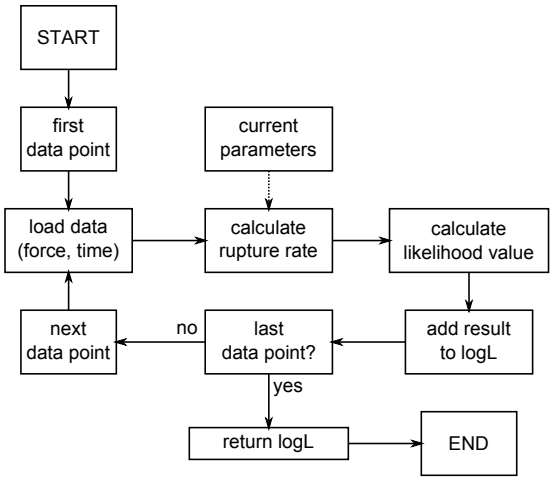


Figure 9.7: Constant force logL function of the BELL and CUSP model. The model is used to calculate the rupture rate for each data point.

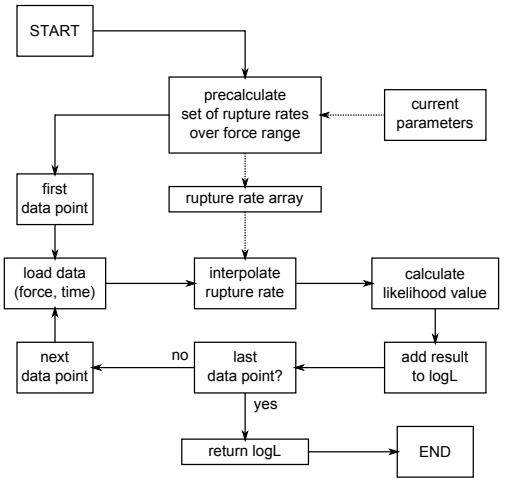


Figure 9.8: Constant force logL function of the linear-cubic model. In this case the model is used to calculate an array of rupture force once at the beginning of the function. The rupture rates for every datapoint are then interpolated to reduce the computational cost.

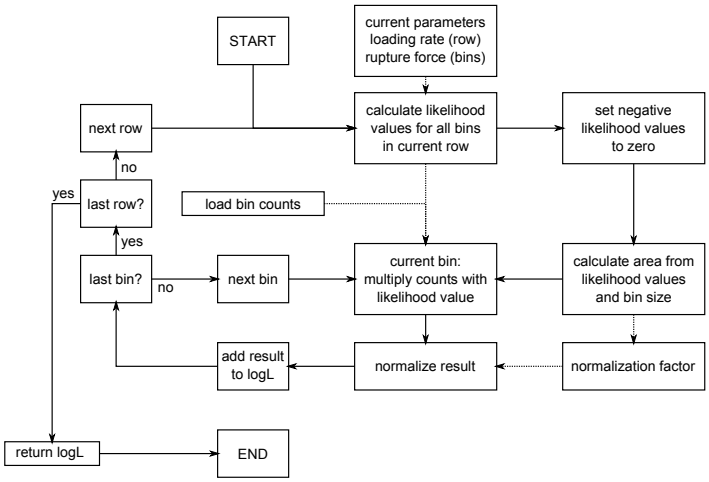


Figure 9.9: Constant speed logL function of the BELL and CUSP model. For this function the experimental data has to be converted into a 2D histogramm first. For each row of the histogram a correction factor is calculated to allow the normalization of the likelihood values of each bin.

and rupture force and multiplied by its counts. For the linear-cubic model (Figure 9.10) some precalculations are possible to reduce the higher computational time in comparison to the BELL and CUSP models. To find the rupture force distribution function knowledge about the function of rupture rates over force is needed. This function is independent from the loading rate and has to be calculated only once. Near and beyond the limits of the valid range of the CUSP model negative likelihood values can occur. The likelihood function treats those values as zero both for the calculation of the correction factor and the likelihood factor.

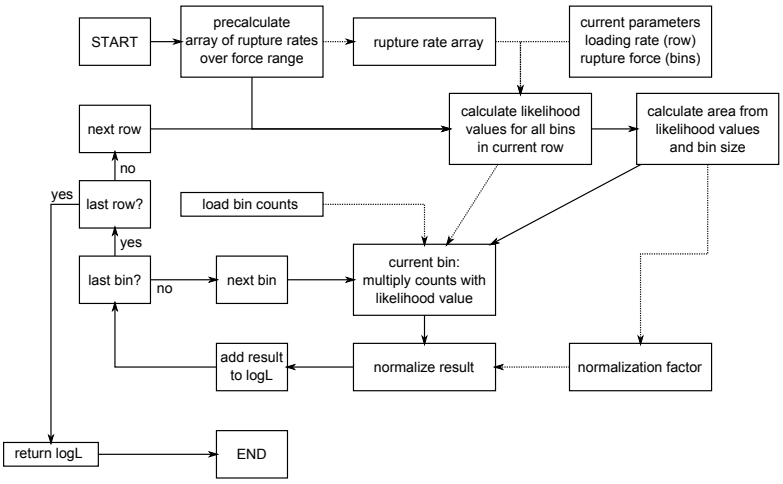


Figure 9.10: Constant speed logL function of the linear-cubic model. Here also the experimental data has to be in the form of a 2D histogram. Numeric integration makes calculation of force dependent rupture rates necessary. Since those rates are independent of the loading rate, the rates have to be calculated only once and not seperately for every row of the 2D histogram. This reduces the computational cost considerably.

List of Figures

2.1 AFM scheme 13
2.2 Constant force measurement scheme 15
2.3 Constant force data: force over time curve 16
2.4 Constant force raw data 16
2.5 Polymer stretching: Enthropic and enthalpic contributions . . 18
2.6 PDMS experimental system 23
2.7 Simulated exponential decay data. 27
2.8 Likelihood functions for exponential decays at different rates. 29
2.9 Posteriors for three sets of simulated exponential decay data. 30
2.10 Nested sampling process. 31
2.11 Sorted likelihood function. 33
2.12 Nested sampling over ξ. 34
2.13 Nested sampling: Sample distribution along ξ. 35
2.14 Staircase sampling. 38
2.15 Constant speed rupture force likelihood functions 41
2.16 CUSP and linear-cubic energy landscapes 46
2.17 Model comparison 47
2.18 Constant speed CUSP model limitations 48
2.19 High barrier approximation CUSP model parameter variation 51
2.20 Linear-cubic model parameter variation 52
2.21 Rupture behavior of various covalent and non-covalent bonds 53
2.22 2D histogram of constant speed data 56
2.23 PDMS sample preparation 61
2.24 Modified AFM head and PDMS flow chamber with sample . 62

3.1 Histogram of measured contour lengths 67
3.2 Rupture rate polymer length independence 68

3.3 Multiple rupture populations 69
3.4 Constant force BELL parameter estimation 71
3.5 Constant force CUSP parameter estimation 1 74
3.6 Constant force CUSP parameter estimation 2 75
3.7 Constant force linear-cubic parameter estimation 77
3.8 CUSP and linear-cubic model result comparison 79
3.9 CUSP and BELL model result comparison 80
3.10 CUSP model: Analysis of repeated simulations 83
3.11 Linear-cubic model: Analysis of repeated simulations 84
3.12 Influence of the range of forces 87
3.13 CUSP model: Influence of the amount of data 90
3.14 Linear-cubic model: Influence of the amount of data 90
3.15 CUSP model: distributed spring constant error 1 96
3.16 CUSP model: distributed spring constant error 2 97
3.17 Linear-cubic model: distributed spring constant error 1 . . . 98
3.18 Linear-cubic model: distributed spring constant error 2 . . . 99
3.19 Effect of a normal distributed spring constant error 100
3.20 CUSP model: constant spring constant error 101
3.21 Linear-cubic model: constant spring constant error 102
3.22 CUSP model: concurrent parameter estimation 104
3.23 Linear-cubic model: concurrent parameter estimation 105
3.24 Constant speed BELL parameter estimation 109
3.25 Constant speed parameter estimation 112
3.26 CUSP model limitations for constant speed data 113
3.27 Constant force: experimental and simulated data 118
3.28 Constant speed: experimental and simulated data 119
3.29 Prepull survival chance 121

6.1 Nucleosomal array sample system 144
6.2 Nucleosomal array substrate 147
6.3 Nucleosome measurement setup 149

7.1 Contact mode, DNA, 1 mM magnesium 155
7.2 Contact mode, DNA, polylysine 156
7.3 Contact mode, DNA, polylysine, zoomed in 157
7.4 Contact mode, nucleosomal arrays, polylysine 159
7.5 Tapping mode, DNA, 1 mM magnesium 161

7.6 Tapping mode, DNA, polylysine 163
7.7 Tapping mode, nucleosomal arrays, polylysine 165
7.8 Tapping mode, nucleosomal arrays, polylysine, high sodium . 166
7.9 Tapping mode, nucleosomal arrays, nickel, high potassium . . 168

9.1 Constant Force Measurement algorithm scheme 179
9.2 Constant force measurement data 180
9.3 Constant force measurement raw data 180
9.4 Constant force measurement data correction 181
9.5 Constant force data simulation scheme 182
9.6 Constant speed data simulation scheme 183
9.7 Constant force logL function: BELL and CUSP model 184
9.8 Constant force logL function: linear-cubic model 184
9.9 Constant speed logL function: BELL and CUSP model . . . 185
9.10 Constant speed logL function: linear-cubic model 186

Bibliography

[1] Tomanek, A. *Silicones and Industry* (Hanser, Munich, Germany, 1991).

[2] Rogers, J. A. & Nuzzo, R. G. Recent progress in soft lithography. *materialstoday* **8**, 50 (2005).

[3] McDonald, J. & Whitesides, G. M. Poly(dimethylsiloxane) as a material for fabricating microfluidic devices. *Acc. Chem. Res.* **35**, 491499 (2002).

[4] Kim, D. H., Ghaffari, R., Lu, N. & Rogers, J. A. Flexible and stretchable electronics for biointegrated devices. *Annu. Rev. Biomed. Eng.* **14**, 113 (2012).

[5] Flaningham, O. *Chem. Anal.* **112**, 135 (1991).

[6] Marceniec, B. a. G. J., Urbaniak, W. & Kornetka, T. W. *Comprehensive Handbook on Hydrosilation Chemistry* (Pergamon Press, Oxford, UK, 1992).

[7] Bustamante, C., Bryant, Z. & Smith, S. B. Ten years of tension: single-molecule DNA mechanics. *Nature* **421**, 423 (2003).

[8] Fisher, T. E., Oberhauser, A. F., Carrion-Vazquez, M., P. E., Marszalek & Fernandez, J. M. The study of protein mechanics with the atomic force microscope. *Trends Biochem. Sci.* **24**, 379 (1999).

[9] Rief, M., Oesterhelt, F., Heymann, B. & Gaub, H. E. Single molecule force spectroscopy on polysaccharides by atomic force microscopy. *Science* **275**, 1295 (1997).

[10] Scherer, A., Zhou, C. Q., Michaelis, J., Braeuchle, C. & Zumbusch, A. Intermolecular interactions of polymer molecules determined by single-molecule force spectroscopy. *Macromolecules* **38**, 9821 (2005).

[11] Hugel, T. & Seitz, M. The study of molecular interactions by afm force spectroscopy. *Macromol. Rapid Commun.* **22**, 989 (2001).

[12] Janshoff, A., Neitzert, M., Oberdorfer, Y. & Fuchs, H. Force spectroscopy of molecular systems—single molecule spectroscopy of polymers and biomolecules. *Angew. Chem., Int. Ed.* **39**, 3212 (2000).

[13] Ludwig, M. *et al.* Afm, a tool for single-molecule experiments. *Appl. Phys. A* **68**, 173 (1999).

[14] Smith, S. B., Finzi, L. & Bustamante, C. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science* **258**, 1122 (1992).

[15] Marko, J. F. & Siggia, E. D. Stretching DNA. *Macromolecules* **28**, 8759 (1995).

[16] Livadaru, L., Netz, R. R. & Kreuzer, H. J. Stretching response of discrete semiflexible polymers. *Macromolecules* **36**, 3732 (2003).

[17] Odijk, T. Stiff chains and filaments under tension. *Macromolecules* **28**, 7016 (1995).

[18] Hugel, T., Rief, M., Seitz, M., Gaub, H. E. & Netz, R. R. Highly stretched single polymers: Atomic-force-microscope experiments versus ab-initio theory. *Phys. Rev. Lett.* **94**, 048301 (2005).

[19] Lupton, E. M. *et al.* Stretching siloxanes: An ab initio molecular dynamics study. *Chem. Phys. Lett.* **414**, 132 (2005).

[20] Smith, S. B., Cui, Y. & Bustamante, C. Overstretching b-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271**, 795 (1996).

[21] Moy, V. T., Florin & E. L., H. E., Gaub. Intermolecular forces and energies between ligands and receptors. *Science* **266**, 257 (1994).

[22] Grandbois, M., Beyer, M., Rief, M., Clausen-Schaumann, H. & Gaub, H. E. How strong is a covalent bond? *Science* **283**, 1727 (1999).

[23] Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541 (1997).

[24] Bell, G. I. Models of the specifc adhesion of cells to cells. *Science* **200**, 618 (1978).

[25] Evans, E. Energy landscapes of biomolecular adhesion and receptor anchoring at interfaces explored with dynamic force spectroscopy. *Faraday Discuss.* **111**, 1 (1998).

[26] Friedsam, C., Wehle, A. K., Kuhner, F. & Gaub, H. E. Dynamic single-molecule force spectroscopy: bond rupture analysis with variable spacer length. *J. Phys.: Condens. Matter* **15**, 1709 (2003).

[27] Dudko, O. K., Hummer, G. & Szabo, A. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.* **96**, 108101 (2006).

[28] Dudko, O. K., Mathé, J., Szabo, A., Meller, A. & Hummer, G. Extracting kinetics from single-molecule force spectroscopy: Nanopore unzipping of DNA hairpins. *Biophys. J.* **92**, 4188 (2007).

[29] Evans, E., Berk, D. & Leung, A. Detachment of agglutininbonded red blood cells. i. forces to rupture molecular-point attachments. *Biophys. J.* **59**, 838 (1991).

[30] Eyring, H., Gershinowitz, H. & Sun, C. E. The absolute rate of homogeneous atomic reactions. *J. Chem. Phys.* **3**, 786 (1935).

[31] Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Phys.* **7**, 284 (1940).

[32] Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y. & Schulten, K. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **72**, 1568 (1997).

[33] Shapiro, B. E. & Qian, H. A quantitative analysis of single protein-ligand complex separation with the atomic force microscope. *Biophys. Chem.* **67**, 211 (1997).

[34] Hummer, G. & Szabo, A. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys. J.* **85**, 5 (2003).

[35] Dudko, O. K., Filippov, A. E., Klafter, J. & Urbakh, M. Beyond the conventional description of dynamic force spectroscopy of adhesion bonds. *PNAS* **100**, 11378 (2003).

[36] Lupton, E. M., Achenbach, F., Weis, J., Braeuchle, C. & Frank, I. J. Modified chemistry of siloxanes under tensile stress: Interaction with environment. *Phys. Chem. B* **110**, 14557 (2006).

[37] Pill, M. F., Schmidt, S. W., Beyer, M. K., Clausen-Schaumann, H. & Kersch, A. A dft model of mechanically activated silyl ester hydrolysis. *J. Chem. Phys.* **140**, 044321 (2014).

[38] Schwaderer, P. *et al.* Single-molecule measurement of the strength of a siloxane bond. *Langmuir* **24**, 1343 (2008).

[39] Sun, G. X. & Butt, H. J. Adhesion between solid surfaces in polymer melts: Bridging of single chains. *Macromolecules* **37**, 6086 (2004).

[40] Senden, T. J., di Meglio, J. M. & Auroy, P. Anomalous adhesion in adsorbed polymer layers. *Eur. Phys. J. B* **3**, 211 (1998).

[41] Schwaderer, P. U. *Investigation of the Strength of the PDMS Backbone under Tensile Stress and the Self-Assembly of Mesoporous Nanostructures* (Verlag Dr. Hut, München, Germany, 2009).

[42] Friedbacher, G. & Fuchs, H. Classification of scanning probe microscopies. *Pure and Appl. Chem.* **71**, 1337 (1999).

[43] Binnig, G., Quate, C. F. & Gerber, C. Atomic force microscope. *Phys. Rev. Lett.* **56**, 930 (1986).

[44] Binnig, G., Rohrer, H., Gerber, C. & Weibel, E. Surface studies by scanning tunneling microscopy. *Phys. Rev. Lett.* **49**, 57 (1982).

[45] Binnig, G., Rohrer, H., Gerber, C. & Weibel, E. Tunneling through a controllable vacuum gap. *Appl. Phys. Lett.* **40**, 178 (1982).

[46] Dürig, U., Pohl, D. W. & Rohner, F. Near-field optical-scanning microscopy. *J. Appl. Phys* **59**, 3318 (1986).

[47] Meyer, G. & Amer, N. M. Novel optical approach to atomic force microscopy. *Appl. Phys. Lett.* **53**, 1045 (1988).

[48] Clifford, C. A. & Seah, M. P. The determination of atomic force microscope cantilever spring constants via dimensional methods for nanomechanical analysis. *Nanotechnology* **16**, 1666 (2005).

[49] Sader, J. E., Larson, I., Mulvaney, P. & White, L. R. Method for the calibration of atomic force microscope cantilevers. *Review of Scientific Instruments* **66**, 3789 (1995).

[50] Sader, J. E., Pacifico, J., Green, C. P. & Mulvaney, P. General scaling law for stiffness measurement of small bodies with applications to the atomic force microscope. *J. Appl. Phys.* **97**, 124903 (2005).

[51] Gibson, C. T., Watson, G. S. & Myhra, S. Determination of the spring constants of probes for force microscopy/spectroscopy. *Nanotechnology* **7**, 259 (1996).

[52] Sader, J. E., Chon, J. W. M. & Mulvaney, P. Calibration of rectangular atomic force microscope cantilevers. *Rev. Sci. Instrum.* **70**, 3967 (1999).

[53] Cumpson, P. J., Clifford, C. A. & Hedley, J. Quantitative analytical atomic force microscopy: a cantilever reference device for easy and accurate afm spring-constant calibration. *Measurement Science and Technology* **15**, 1337 (2004).

[54] Senden, T. & Ducker, W. Experimental determination of spring constants in atomic force microscopy. *Langmuir* **10**, 1003 (1994).

[55] Cleveland, J. P., Manne, S., Bocek, D. & Hansma, P. K. A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. *Review of Scientific Instruments* **64**, 403 (1993).

[56] Hutter, J. L. & Bechhoefer, J. Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.* **64**, 1868 (1993).

[57] Heer, C. V. *Statistical Mechanic Kinetic Theory, and Stochastic Processes* (Academic, New York, USA, 1972).

[58] Gibson, C. T., Watson, G. S. & Myhra, S. Scanning force microscopy - calibrative procedures for best practice. *SCANNING* **19**, 564 (1997).

[59] Butt, H. J. & Jaschke, M. Calculation of thermal noise in atomic force microscopy. *Nanotechnology* **6**, 1 (1995).

[60] Pirzer, T. & Hugel, T. Atomic force microscopy spring constant determination in viscous liquids. *Review of Scientific Instruments* **80**, 035110 (2009).

[61] Hager, R., Schneider, O. & Schuster, J. Patent US 5,380,902 (1995).

[62] Lupton, E. M., Achenbach, F., Weis, J., Braeuchle, C. & Frank, I. J. Molecular origins of adhesive failure: Siloxane elastomers pulled from a silica surface. *Phys. Rev. B* **76**, 125420 (2007).

[63] Frank, I. J. & Hofbauer, F. Single-molecule mechanics: Breaking bonds at a stretch. *Nat. Chem.* **1**, 180 (2009).

[64] Jaynes, E. T. *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, UK, 2003).

[65] D., S. & J., S. *Data Analysis* (Oxford University Press, Oxford, USA, 2006).

[66] Muschielok, A. M. *Development and Application of a Quantitative Analysis Method for Fluorescence Resonance Energy Transfer Localization Experiments* (Verlag Dr. Hut, München, Germany, 2011).

[67] T., B. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London* **53**, 330 (1763).

[68] Skilling, J. Nested sampling for general bayesian computation. *Bayesian Anal.* **1**, 833 (2006).

[69] Muschielok, A. *et al.* A nano-positioning system for macromolecular structural analysis. *Nature Methods* **5**, 965 (2008).

[70] Andrecka, J. *et al.* Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase ii elongation complex. *Nucleic Acids Research* **37**, 5803 (2009).

[71] Muschielok, A. & Michaelis, J. Application of the nano-positioning system to the analysis of fluorescence resonance energy transfer networks. *J. Phys. Chem. B* **115**, 11927 (2011).

[72] Treutlein, B. *et al.* Dynamic architecture of a minimal RNA polymerase ii open promoter complex. *Molecular Cell* **46**, 136 (2012).

[73] Kuegel, W., Muschielok, A. & Michaelis, J. Bayesian-inference-based flourescence correlation spectroscopy and single-moledule burst analysis reveal the influence of dye selection on DNA hairpin dynamics. *ChemPhysChem* **13**, 1013 (2012).

[74] Hyeon, C. B. & Thirumalai, D. Forced-unfolding and forcequench refolding of RNA hairpins. *Biophys. J.* **90**, 3410 (2006).

[75] Koculi, E., Thirumalai, D. & Woodson, S. A. Counterion charge density determines the position and plasticity of RNA folding transition states. *J. Mol. Biol.* **359**, 446 (2006).

[76] Dembo, M., Torney, D. C., Saxman, K. & Hammer, D. The reaction-limited kinetics of membrane-to-surface adhesion and detachment. *Proc. R. Soc. Lond. B Biol. Sci.* **234**, 55 (1988).

[77] Dudko, O. K., Hummer, G. & Szabo, A. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *PNAS* **105**, 15755 (2008).

[78] Schlierf, M. & Rief, M. Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape. *Biophys. J.* **90**, 33 (2006).

[79] Szabo, A., Schulten, K. & Schulten, Z. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* **72**, 4350 (1980).

[80] Shillcock, J. & Seifert, U. Escape from a metastable well under a time-ramped force. *Phys. Rev. E* **57**, 7301 (1998).

[81] Schlierf, M. & Rief, M. Temperature softening of a protein in single-molecule experiments. *J. Mol. Biol.* **354**, 497 (2005).

[82] Oberhauser, A. F., Hansma, P. K., Carrion-Vazquez, M. & Fernandez, J. M. Stepwise unfolding of titin under force-clamp atomic force microscopy. *PNAS* **98**, 468 (2001).

[83] Schmidt, S. W., Beyer, M. K. & Clausen-Schaumann, H. Dynamic strength of the silicon carbon bond observed over three decades of force-loading rates. *JACS* **130**, 3664 (2008).

[84] Schmidt, S. W., Kersch, A., Beyer, M. K. & H., C.-S. Mechanically activated rupture of single covalent bonds: evidence of force induced bond hydrolysis. *PCCP* **13**, 5994 (2011).

[85] Schmidt, S. W., Filippov, P., Kersch, A., Beyer, M. K. & Clausen-Schaumann, H. Single-molecule force-clamp experiments reveal kinetics of mechanically activated silyl ester hydrolysis. *ACS Nano* **6**, 1314 (2012).

[86] Jeffreys, H. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. A Mathem. And Pys. Sci.* **186**, 453 (1946).

[87] Chung, H. S. & Eaton, W. A. Single-molecule fluorescence probes dynamics of barrier crossing. *Nature* **502**, 685 (2013).

[88] Jing, G. Y., Ma, J. & Yu, D. P. Calibration of the spring constant of afm cantilever. *J. Electron. Microsc.* **56**, 21 (2007).

[89] Ris, H. & Kubai, D. F. Chromosome structure. *Annu. Rev. Genet.* **4**, 263 (1970).

[90] Kornberg, R. D. Structure of chromatin. *Annu. Rev. Biochem.* **46**, 931 (1977).

[91] Cooper, G. M. *The Cell, 2nd edition, A Molecular Approach* (Boston University, Boston, USA, 2000).

[92] Davies, H. G. & Small, J. V. Structural units in chromatin and their orientation on membranes. *Nature* **217**, 1122 (1968).

[93] Olins, A. L. & Olins, D. E. Spheroidal chromatin units (nu bodies). *Science* **183**, 330 (1974).

[94] Finch, J. T. *et al.* Structure of nucleosome core particles of chromatin. *Nature* **269**, 29 (1977).

[95] Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. Structure of the nucleosome core particle at 7 Å resolution. *Nature* **311**, 532 (1984).

[96] Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251 (1997).

[97] Richmond, T. J. & Davey, C. The structure of dna in the nucleosome core. *Nature* **423**, 145 (2003).

[98] Robinson, P. J. & Rhodes, D. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr. Opin. Struct. Biol.* **16**, 336 (2006).

[99] Perisic, O., Collepardo-Guevara, R. & Schlick, T. Modeling studies of chromatin fiber structure as a function of DNA linker length. *J. Mol. Biol.* **403**, 777 (2010).

[100] Robinson, P. J. J., Fairall, L., Huynh, V. A. T. & Rhodes, D. EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure. *PNAS* **103**, 6506 (2006).

[101] Finch, J. T. & Klug, A. Solenoidal model for superstructure in chromatin. *PNAS* **73**, 1897 (1976).

[102] Song, F. *et al.* Cryo-em study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* **344**, 376 (2014).

[103] Radmache, M., Fritz, M. & Hansma, P. K. Imaging soft samples with the atomic-force microscope - gelatin in water and propanol. *Biophys. J.* **69**, 264 (1995).

[104] Bhushan, B., Schirmeisen, A., Anczykowski, B. & Fuchs, H. *Nanotribology and Nanomechanics - An introduction* (Springer, Hamburg, Germany, 2008).

[105] Anczykowski, B., Gotsmann, B., Fuchs, H., Cleveland, J. P. & B., E. V. How to measure energy dissipation in dynamic mode atomic force microscopy. *Appl. Surf. Sci.* **140**, 376 (1999).

[106] Garcia, R. *et al.* Identification of nanoscale dissipation processes by dynamic atomic force microscopy. *Phys. Rev. Lett.* **97**, 016103 (2006).

[107] Watson, J. D. & Crick, F. H. A structure for deoxyribose nucleic acid. *Nature* **171**, 737 (1953).

[108] Mandelkern, M., Elias, J. G., Eden, D. & Crothers, D. M. The dimensions of dna in solution. *J. Mol. Biol.* **152**, 153 (1981).

[109] Rivetti, C., Guthold, M. & Bustamante, C. Scanning force microscopy of dna deposited onto mica: Equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J. Mol. Biol.* **264**, 919 (1996).

[110] Marek, J. *et al.* Interactive measurement and characterization of dna molecules by analysis of afm images. *Cytometry, Part A* **63A**, 87 (2005).

[111] Wiggins, P. A. *et al.* High flexibility of dna on short length scales probed by atomic force microscopy. *Nat. Nanotechnol.* **1**, 137 (2006).

[112] Widom, J. A relationship between the helical twist of dna and the ordered positioning of nucleosomes in all eukaryotic cells. *PNAS* **89**, 1095 (1992).

[113] Lowary, P. T. & Widom, J. New dna sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol* **276**, 19 (1998).

[114] Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097 (2002).

[115] Poirier, M. G., Oh, E., Tims, H. S. & Widom, J. Dynamics and function of compact nucleosome arrays. *Nat. Struct. Mol. Biol.* **16**, 938 (2009).

[116] Cheng, H., Zhang, K., Libera, J. A., de la Cruz, M. O. & Bedzyk, M. J. Polynucleotide adsorption to negatively charged surfaces in divalent salt solutions. *Biophys. J.* **90**, 1164 (2006).

[117] Bussiek, M., Muecke, N. & Langowski, J. Polylysine-coated mica can be used to observe systematic changes in the supercoiled dna conformation by scanning force microscopy in solution. *Nucleic Acids Research* **31**, e137 (2003).

Acknowledgement

During the work on my thesis I have received help and support from many people and I want to thank every single one of them, because without them I wouldn't have been able to do this:

I want to thank Jens Michaelis for being a great supervisor. I never felt left alone and always supported and motivated. I learned very much through countless discussions, suggestions and hints. And I am also thankful that I learned to present my research more clearly and precisely. Furthermore I want to thank Jens for the Biophysics Institute he has built up in Ulm. It is a great place to work and learn. And also to have fun.

I thank Christoph Bräuchle for my time in the Physical-Chemistry I Institute in Munich. In his institute he has created an atmosphere of cooperation and friendliness despite it's large size. He was inspiring to everyone and not only in the sciences. I'm looking back with fondness on the years I spent working there. I am thankful to Christof Gebhardt for many insights and fruitful discussions.

This work would not have been possible without my cooperation partners and I'm grateful for their help: Frank Achenbach, Enno Funk and Johan Weis from Wacker Chemie AG for their help with the polydimethylsiloxane project. Daniela Rhodes from MRC Laboratory of Molecular Biology, Peter Becker from Ludwig Maximilian University of Munich and Gernot Längst from University of Regensburg for help with the nucleosomal array project. The samples for both projects are highly complex and very difficult to create and I fully appreciate how much work has been put into them.

I have received much help from Peter Schwaderer, Adam Muschielok, Wolfgang Kügel, Korbinian Paul-Yuan, Florian Drechsler, Christian Osseforth, Jochen Reichelt and Matthias Reisser. I want to thank you all for all the discussions we had and all the questions you answered. I want to thank

Katarzyna Krzemien for the help with the nucleosomal fibers. It was great to work together and I learned much.

I'm thanking furthermore all the people who keep the institutes running in both Munich and Ulm: Moritz Ehrl, Bernhard Gruchmann, Nadine Jakobi, Heidi Kohn, Carlheinz Röcker, Reinhilde Schiffner-Grossmann, Silke Steger, Uwe Theilen and Nadine Weidinger.

I want to also thank everyone else from the institutes in Munich and Ulm. You were the best colleagues I could imagine. Also the friendliest and funnest!

Finally I want to thank my whole family and all of my friends. You supported me in so many different ways, that I could fill many more pages with it. With you I was able to always look forward and keep on going with a smile on my face. Thank you!

Declaration of Cooperations

I have received help through cooperations with the following partners:

Frank Achenbach, Enno Funk and Johann Weis
Wacker Chemie AG, Munich
Sharing of knowledge, synthesis of polydimethylsiloxane samples, provision of force spectroscopy substrates and solvents.

Daniela Rhodes
MRC Laboratory of Molecular Biology, Cambridge
Development of the doublestranded DNA in the nucleosomal arrays.

Peter Becker
Ludwig Maximilian University of Munich, Munich
Provision of the doublestranded DNA in the nucleosomal arrays with consent of Daniela Rhodes.

Gernot Längst
University of Regensburg, Regensburg
Provision of histone plasmids for usage in the nucleosomal arrays.

Curriculum Vitae

Name:	Michael Roland Budde
Geburtsdatum:	20. 04. 1984
Geburtsort:	München

Beruf

2012-2014	Wissenschaftlicher Mitarbeiter an der Universität Ulm
2009-2012	Wissenschaftlicher Mitarbeiter an der Ludwig-Maximilians-Universität München

Studium

2012-2015	Fortsetzung der Promotion an der Universität Ulm
2009-2012	Beginn der Promotion an der Ludwig-Maximilians-Universität München
2005-2008	Master Chemie an der Ludwig-Maximilians-Universität München
2003-2005	Bachelor Chemie und Biochemie an der Ludwig-Maximilians-Universität München

Schulbildung

1994-2003	Thomas-Mann-Gymnasium München
1990-1994	Grundschule an der Berner Strasse München