Single Molecule Force Spectroscopy on Protein-Nucleicacid-Complexes

by
Alexander Fuhrmann

Diploma Thesis

Faculty of Physics

University of Bielefeld

October 2006

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstängig verfasst und dabei keine weiteren als die angegebenen Hilfsmittel verwendet habe.

Bielefeld, den 31.10.2006

(Alexander Fuhrmann)

1 Introduction

Most physiological processes in the cell are controlled by proteins. However, not all reactions are carried out at the same time, throughout the whole life of a cell or of an organism. Thus gene expression has to be regulated strictly. Several mechanisms of regulation can be observed: transcriptional-, post-transcriptional- and post-translational regulation. Transcriptional regulation is mostly achieved by promoter binding of proteins, whereas post-transcriptional regulation requires binding of proteins to the mRNA. Post-translational regulation needs the modification of already expressed proteins.

Although these processes take place on the single molecule level these interactions between biomolecules are usually studied on huge ensembles. One of these commonly used methods to analyse the interaction of nucleic acids and proteins is the electrophoretic mobility shift assay (EMSA). Therefore the two binding partners are incubated in an appropriate buffer system and the reaction products are subsequently separated on polyacrylamide or argarose gels. Under application of an external electrical field the separation depends on size, charge and shape. Due to its larger size, a complex of protein and RNA (resp. DNA) will run slower than the free, unbound components: a so called band-shift happens. By labelling the RNA (resp. DNA) (or protein) with a radioactive or fluorescent label the covered distance of the molecules can be visualized.
The specificity of binding can also be obtained by competition experiments. For this increasing amounts of free (unlabeled) homologous or heterologous binding partners (e.g. unlabeled RNA) are added to the binding reactions. If increasing amounts of the homologous, but not of the heterologous, competitors can displace the labelled RNA (resp. DNA) complex it can be considered as specific.
This method provides information about the equilibrium constant (the ratio of forward and reverse reactions in thermal equilibrium) of the RNA-protein interactions investigated although the knowledge about the molecular mechanisms of RNA-protein binding is pretty poor.

This is the point where another tool can demonstrate its vantages, the atomic force microscope (AFM). While all mentioned methods analyse the binding processes in an ensemble, this instrument is so sensitive that it can observe the (un-)binding characteristics of biomolecules at the single molecule level, a method called single molecule force spectroscopy (SMFS). One binding partner is attached via a long polymer linker to a sharp tip of a force sensor (cantilever) while the other binding partner is immobilized, also via a linker, on the sample surface. The sample is then moved up and down by a piezo element and if two respective binding partners find each other during the contact time of the cantilever on the sample (dwell time), the cantilever will be bent during the retraction. Usually the unbinding process (as well as the binding) is driven by thermal fluctuations driving the system across the activational barrier of the binding potential. This remains, in principal, untouched under an externally applied force but this force has a severe influence on the stability of the bond. It is thought that the activational barrier deforms under a pulling force and thus makes it easier to override the energy barrier. The common standard theory describing these processes suffers of several inconsistencies. These problems will not only be discussed in the theoretical part of this work, furthermore a promising theoretical approach accounting these inconsistencies will be described. Although the advantages of these theoretical improvements were known, there was still no applicable method to analyse the experimental data according to this theory.

One major part of this work is the development of new methods and tools for the analysis of the data gathered in the experimental part that will not only meet the demands of this new theory but also has many advantages for the whole data analysis. These will be demonstrated at a protein-DNA interaction belonging to the category of the transcriptional gene regulation. Beside the automation of the data analysis this novel software enables the qualitative as well as the quantitative separation of different binding modes of the investigated protein-RNA interactions, appertaining to the post-transcriptional regulation, that could not even be noticed before.

2 The investigated biological Systems

In the experimental part of this thesis two different biological systems responsible for gene regulation were investigated. The first one, quorum sensing in bacteria *Sinorhizobium melilotiti*, is part of the transcriptional control and is thus a protein-DNA interaction. The second biological system of interest is a negative feedback loop in the circadian clock of the plant *Arabidopsis thaliana* based on post-transcriptional regulation and thus a protein-RNA interaction.

2.1 Quorum Sensing in *Sinorhizobium meliloti*

The regulation of gene expression is a control mechanism allowing a cell to respond on chemical signals or environmental changes by adapting the expression of genes. The first (and most important) step in gene regulation occurs at the transcriptional level: Transcription can be increased by positive regulation (activation) or decreased by negative regulation (repression).
One particular form of gene regulation in bacteria is quorum sensing (QS), i.e. a population density–dependent transcription controlled by low molecular-weight compounds called autoinducers. QS is known to regulate many different physiological processes, including the production of secondary metabolites, conjugal plasmid transfer, swimming, swarming, biofilm maturation, and virulence in human, plant, and animal pathogens (1, 2). Many QS systems involve *N*-acyl homoserine lactones (AHLs) as signal molecules (3). These AHLs vary in length, degree of substitution and saturation of the acyl chain (Fig. 2.5).

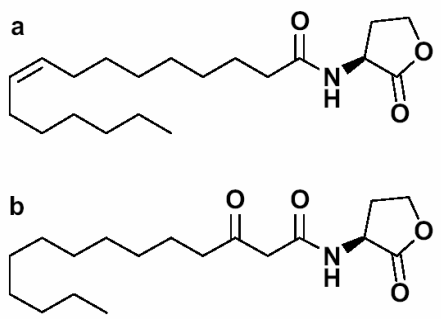

Figure 2.5: Acyl homoserine lactones (AHLs)
Of the synthesized AHLs, those with modifications in the acyl side chain are shown: (a) *N*-[(9Z)-hexadec-9-enoyl]-L-homoserine lactone ($C_{16:1}$-HL) and (b) *N*-(3-Oxotetradecanoyl)-L-homoserine lactone (oxo-C_{14}-HL).

Bacterial cell walls are permeable to AHLs, either by unassisted diffusion across the cell membrane (for shorter acyl chain length) or active transport (possibly for longer acyl chain length). With an increasing number of cells AHLs accumulate both, intracellularly and extracellularly. Once a threshold concentration is reached, they act as co-inducers, usually by activating LuxR-type transcriptional regulators. *Sinorhizobium meliloti* is a common Gram-negative soil and rhizosphere bacterium serving as a biological model system in the study of nitrogen fixation. It has the ability to induce the formation of nodules on the roots of

Medicago, Melilotus and Trigonella sp. where differentiated bacteria called bacteroids fix atmospheric nitrogen to ammonia in symbiotic association with certain genera of these leguminous plants.

In *S. meliloti* Rm1021, a QS system consisting of the AHL synthase SinI and the LuxR-type AHL receptors SinR and ExpR was identified (4). SinI is responsible for production of several long-chain AHLs (C_{12}-HL to C_{18}-HL) (5). The presence of a second QS system, the Mel system, controlling the synthesis of short-chain AHLs (C_6-HL to C_8-HL) was suggested (5). In addition to SinR, five other putative AHL receptors, including ExpR, were identified (6). As originally described for the model QS LuxI/LuxR system of *Photobacterium fischeri*, it is assumed that the LuxR-type regulators are activated by binding of specific AHLs (7). Once activated, the expression of target genes is regulated by binding upstream of the promoter regions of these genes (8). The first target genes identified for the *S. meliloti* Sin system were the *exp* genes mediating biosynthesis of the exopolysaccharide galactoglucan. The expression of the *exp* genes not only relies on a sufficient concentration of Sin system-specific AHLs but also requires the presence of the LuxR-type AHL receptor ExpR (6, 9). Data of transcriptomics and proteomics approaches suggested that the majority of target genes of the Sin system is controlled by ExpR (10, 11). The *S. meliloti* 1021 wild type strain carries an inactive *expR* gene due to disruption of its coding region by insertion element IS*Rm*2011-1 (6). However, the spontaneous dominant mutation *expR101* resulting from precise reading frame-restoring excision of the insertion element from the coding region unraveled the role of *expR* in regulation of galactoglucan biosynthesis (6). ExpR is highly homologous to the *Vibrio fischeri* LuxR. Activated LuxR type regulators usually bind to a consensus sequence known as the lux box, typically located upstream of the promoters of its target genes (8). However, the DNA binding site of ExpR has not yet been identified.

2.2 RNA-Protein Interactions

Binding of proteins to RNA molecules is a common process in cells. For the assembly of these so called ribonucleoprotein particles (RNP) the direct binding of an RNA-binding protein to RNA is required. This binding is achieved by specialized RNA-binding domains. The most common and best-studied example is the RNA recognition motif (RRM), which can be found in different classes of RNA-binding proteins. Extensively studied representatives of this class are for example the spliceosomal protein U1A, the heterogeneous nuclear ribonucleoprotein A1 or the *Drosophila* sex determination switch factor Sex-lethal. In a second step other proteins can be incorporated into the complex by means of protein-protein interaction (12). The complex formation of RNA and proteins serves different purposes of post-transcriptional gene regulation: stabilization, protection, package, transport, processing or degradation. RNA-protein interactions are required for a wide range of regulatory processes in the cell and thus essential for survival. One of these processes includes the regulation of the circadian rhythm in the *Arabidopsis thaliana* plant.

2.1.1 Introduction Circadian Clocks

Every living creature on earth is exposed to periodical changes of the environmental conditions mainly caused by the rotation of the earth about its own axis (day-night rhythms) and around the sun (seasons) (Fig. 2.1). Periodical changes with a 24 hours period length are called circadian rhythms (latin: circa – around, dies – day). In 1729 the French astronomer

Jean Jacques d'Ortous de Mairan discovered that the opening and closing of the leafs of the *Mimosa* plant could not only be explained by periodically changes of light intensity because the opening and closing of the leafs in a 24 hours rhythm was also observed in darkness suggesting an intrinsic rhythm operating this process (13).
About three centuries later it is now possible to research these intrinsic rhythms at the molecular level.

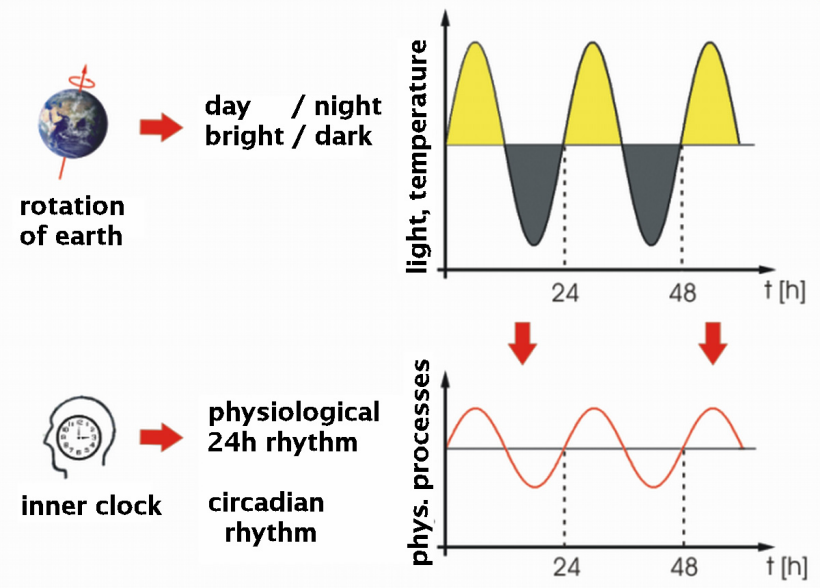

Figure 2.1: Illustration of circadian rhythm

The following section will give a short summary of how circadian clocks in plants work in general. Then the scientific-historic path leading to this general model will be described for the *Arabidopsis thaliana* model system, together with current scientific results.

2.1.2 Basic Model of Circadian Clocks in Plants

Underlying the periodical changes in plantmetabolism is a complex mechanism of gene regulation building up a molecular oscillator. In principle the oscillator consists of clock proteins whose abundances change within a day. These clock proteins feed back on each other by means of positive or negative transcriptional control. The timed transcription and translation of the clock genes leads to robust self sustained rhythms of protein abundance. These rhythms are adjusted to the environmental conditions like light and temperature via so called input pathways. In these pathways, e.g. photoreceptors mediate the information from outside the plant to the central oscillator and “entrain” the clock.
The clock genes of the central oscillator directly control output genes or genes of secondary oscillators, so called “slave” oscillators, by promoter binding. These slave oscillators themselves show a negative autoregulatory feedback loop. Secondary oscillators are needed for an expansion of the clock signal on the way to the output genes (14).

2.1.3 Discovering the Circadian Clock of the *Arabidopsis thaliana* Plant

The *Arabidopsis thaliana* plant is an important model systems in plant sciences because of its beneficial qualities for experiments (e.g. fast growth, easy handling). With about 125 mega base pairs and five chromosomes it has a relatively small genome for plants, so it is not astonishing that it was the first plant genome sequenced (15, 16).

While the presence of inner clocks in plants was known for quite a long time (as mentioned above) by observing the macroscopic changes, in 1985 the first molecular evidence for circadian rhythm was found in the form of time dependent gene activity (17).
The level of mRNA for the light harvesting chlorophyll (LHC) a/b binding protein (and also other mRNAs) has been analyzed under light-dark and constant light conditions. While the mRNA can hardly be detected during night time, the concentration rises about 2 hours before sunrise, having a maximum about noon. It was found that this rhythm also persists under constant light conditions.

The gene of a protein called luciferase, occurring in fireflies, implanted into the *Arabidopsis thaliana* genome, made the plants luminescent especially during morning time. In search of mutants disturbing the natural clock, seeds of these transgenic plants were exposed to chemicals causing genetic mutations. One mutant plant showed the desired character of having a shortened period which could be easily observed in a change of the bioluminescence and also leaf movement period. This mutant has been named "*timing of CAB expression*" (*toc1*-mutant) (18).
In another experiment, the *late elongated hypocotyls (LHY)* transcript, encoding a transcription factor, could be identified to be a player in the circadian clock. *Arabidopsis* plants growing up in a long-day environment (16 hours of light per day) start flowering earlier than those growing up in shortened day conditions (8 hours light). *lh*y mutants, however, ignore the difference of light time per day by starting flowering in both cases at the same time (in both cases plants showed the short day phenotype) (19).
Overexpression LHY results in a damping of the *LHY* transcript and leaf movement oscillation (19).
By searching for proteins binding to the LHC-promoter sequence, another circadian regulated transcription factor was found: Both the "circadian clock associated" (CCA1) mRNA and protein show a circadian oscillation and CCA1 overexpression causes a damping of the *CCA1*-mRNA oscillation (20).
From these experimental assessments the following pathway for the central oscillator of the *Arabidopsis thaliana* clock could be identified. The nuclear protein TOC1 activates the expression of *LHY* and *CCA1* while LHY and CCA1 in turn repress *TOC1* transcription, compare Fig. 2.3.

While the central clock became more and more understood, a circadian regulated glycine-rich RNA-binding protein (GRP) was discovered in *Sinapis alba*. Homology analysis in *Arabidopsis thaliana* revealed at least two corresponding genes: *AtGRP7* and *AtGRP8* (21, 22).
The protein *At*GRP7 shows a circadian oscillation with a maximum of transcript in the evening times (where LHY and CCA1 concentrations are low) (Fig. 2.2) (23).

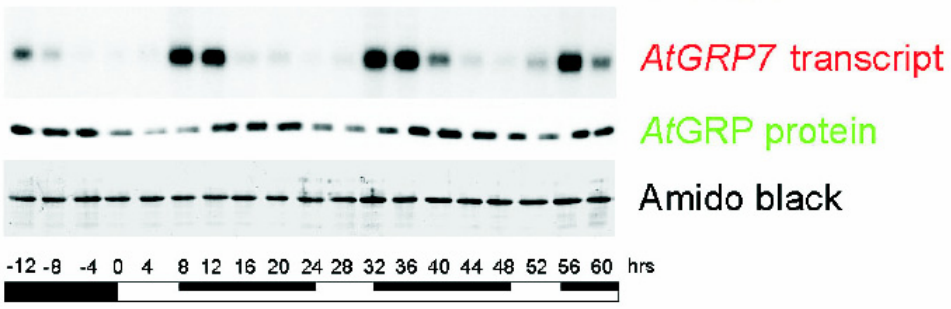

Figure 2.2: The oscillation of AtGRP7 transcript remains even under conditions of constant light (taken of (23))

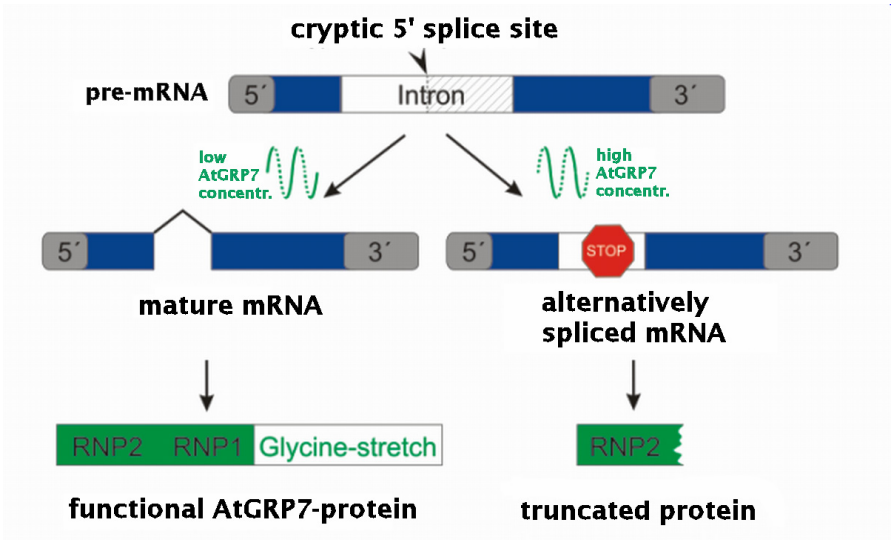

Figure 2.3: Alternative splicing of AtGRP7 pre-mRNA

The oscillation of the RNA and its protein were found to be interdependent. When the protein concentration increases, the transcript concentration decreases suggesting a negative feedback loop of the protein and its own transcript (Fig. 2.2 and 2.4). This supposition could be confirmed by transgenic plants having a constitutive high *At*GRP7 expression (23). In these plants the amount of endogenous transcript dampens to nearly undetectable levels, confirming the assumption that *At*GRP7 regulates the abundance of its own transcript. Additionally the transcript shows a different size, due to alternative splicing upon *At*GRP7 overexpression. This ~150bp longer transcript has a reduced half life responsible for its low steady state abundance and a pre-mature stop codon resulting in the production of no or unfunctional protein (Fig. 2.3).

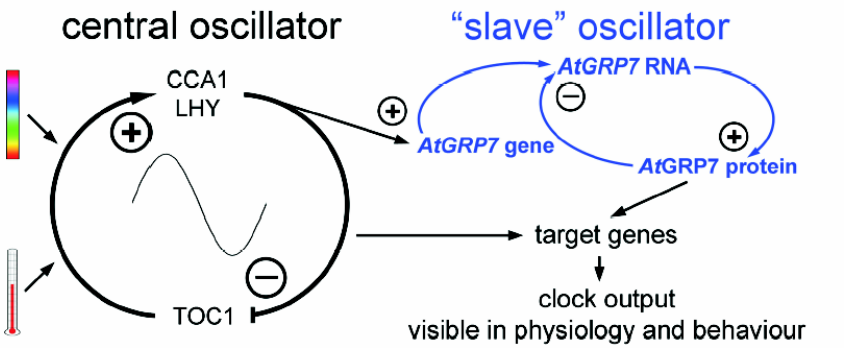

Figure 2.4: Schematic illustration of the central oscillator and the "slave" oscillator

Thus the mechanism causing the oscillation of AtGRP7 seems to rely on the ability of this protein to bind to its own transcript at high protein levels. This binding at the pre-mRNA leads to the attachment of further factors (e.g. spliceosomes) changing the mRNA maturation. In case of low protein concentrations, the pre mRNA is processed correctly and functional protein is made. Increasing *At*GRP7 levels, however, lead to the binding of *At*GRP7 to its own transcript, resulting in the production of the alternatively spliced mRNA form and thus to the loss of functional protein.

Binding sites for *At*GRP7 have been identified in intron and 3'UTR by means of radioactive EMSA with recombinant GST-fusion protein. Mutation of the binding sequence has been shown to abolish binding nearly completely.
Despite its autoregulatory function, AtGRP7 was also shown to regulate the expression of the related protein AtGRP8 via the same mechanism (24).

3 Force Spectroscopy

The single molecule technique used in this work is force spectroscopy (FS) based on the atomic force microscope (AFM). In this chapter first a brief introduction to the AFM will be given, followed by the theoretical basics needed for this thesis. In consideration of the fact that during this work the AFM was mainly used as a very sensitive force sensor measuring unbinding forces of single biomolecules, the main part is focused on the theoretical topics concerning these problems. A typical force distance curve will be discussed, followed by the method applied to acquire the force information from the deflection signal of the cantilever. While these annotations can be conceived as standard in up to date atomic force microscopy, the last section, dealing with the theory of single molecule force spectroscopy, introduces a new promising theoretical approach yielding much better results, especially in combination with the new methods of analysis as applied to the data in chapter 6.

3.1 AFM based Single Molecule Force Spectroscopy

3.1.1 Scanning Probe Microscopy

The scanning tunneling microscope (STM) was invented in 1981 by Binnig and Rohrer (25, 26). The STM makes use of the quantum mechanical tunnelling effect of electrons. An atomically sharp tip is moved over the surface of the sample while the tunnelling current is detected. The main reason why the STM is not suitable for biological probes (with some exceptions) is that most biological materials are insulators.
In 1986 Binnig, Quate and Gerber invented the atomic force microscope (AFM) which does not need an electric conductor as sample (Fig. 3.1) (27). The tip of the AFM is mounted on a cantilever which behaves like a flat spring according to Hooke´s law. The deflection of the cantilever is measured by a laser spotted on the top of the cantilever which reflects the laser beam onto an array of four photodiodes. In this way, both vertical and lateral deflection of the cantilever can be detected.
The AFM is mainly used to scan surfaces line by line. Two main imaging modes can be distinguished. The simplest mode is the contact mode where the tip is pressed on the surface with a constant force. This method can hardly be applied on fragile biomolecules like DNA because of a layer of condensed water that is always on a sample and pulls the tip on the surface (capillary forces). The tip would move these soft and not strongly attached biomolecules across the surface. Immersing both, tip and sample, in water reduces these capillary forces drastically.
Another way to avoid this problem is to use mainly attractive forces that the tip experiences near by the surface. At a certain distance the tip experiences an attractive force while upon further approach to the surface the attractive force decreases, until at a certain distance the force becomes repulsive and the tip is retracted. This particular character of the acting forces is used in the so called dynamic modes, where the cantilever is oscillated close to its resonance frequency by a piezo, and the distance between tip and surface is controlled by the frequency or amplitude change of the cantilever that occurs when approaching the surface.

The dynamic mode commonly used to scan biological samples is the tapping mode where the tip oscillates ("taps") for only a few micro seconds in the vicinity of the surface which even does not affect biomolecules. It not only yields the topography of the probe, but provides additional information about the phase shifts (of the cantilever oscillation) caused by different materials.

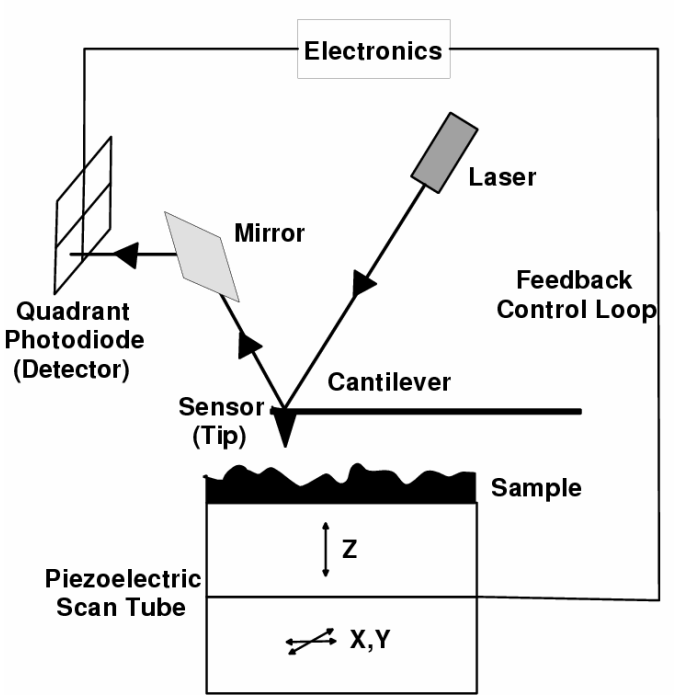

Figure 3.1: Schematic buildup of an atomic force microscope

3.1.2 Force Distance Measurements

In this work the main use of the AFM was to measure forces between receptor-ligand systems. This method, termed AFM force spectroscopy, will now be explained in detail.

A typical (adhesion) force distance curve can be divided into 7 steps (Fig. 3.2).

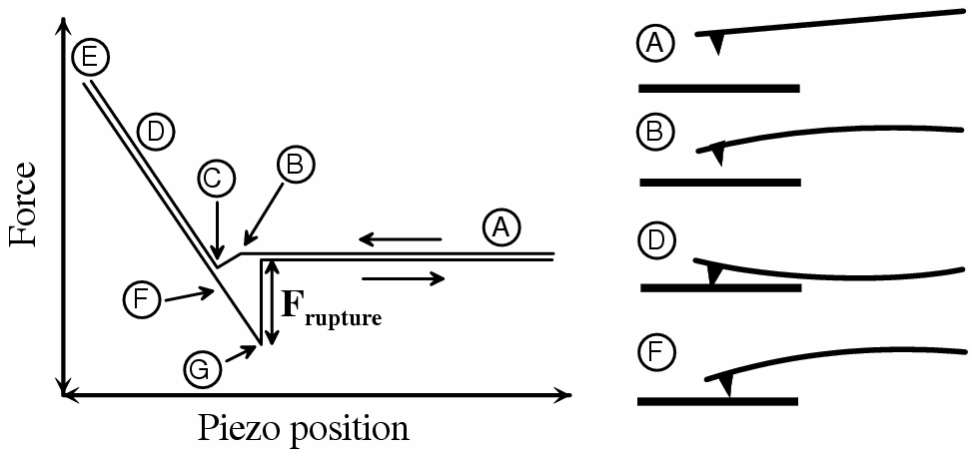

Figure 3.2: Force-distance curve (schematic representation)

In the beginning (A), the cantilever is far away from the surface, approaching the sample with a constant speed. Upon further approach to the sample the cantilever will be deflected towards the surface due to attractive forces (e.g. van der Waals forces, electrostatic forces resulting from unlike net charges) (B). The force at (B) is not always attractive, it can also be repulsive due to electrostatic forces resulting from like net charges or other effects. However, the systems analysed in this work generally exhibited weak attractive forces. (C) is the point where the tip "snaps" into contact with the surface. Pushing the sensor further into the sample results in the proportional relationship between deflection and z-piezo movement as predicted by Hooke´s Law (D). The slope of this straight line is called "sensor response" and will be necessary to convert the voltage signal from the photodiodes into force information (see below). The point of return where the piezo stops approaching and the movement is reversed can be set at a default force value (if the spring constant of the cantilever is given), ensuring in this way that the tip always stays in contact for a well defined time (E). In some experiments presented later in this work the resuming time of the tip at the surface (dwell time) was varied. The retraction in the first part resembles the approach, but at (F) the cantilever experiences an attractive force (adhesion) causing the tip to remain in contact with the sample until the elastic restoring force exceeds the attractive force (G). During further retraction the cantilever oscillates freely.
If this cycle is repeated very often at different retract velocities, it is called dynamic force spectroscopy (DFS). The features of force distance curves in a typical single molecule force spectroscopy (SMFS) experiment will be addressed later.

During the last two decades, DFS has developed into a highly sensitive tool for the investigation of the interaction of single biomolecules (28, 29), from complementary DNA strands (30) to ligand-receptor pairs (31-36) and cell adhesion molecules (37). Only most recently, protein-DNA interaction has come under survey by DFS (38-42). In particular, DFS data has proven to complement the information gained from conventional molecular biology experiments in a detailed study of DNA binding of the regulator ExpG activating transcription of the *exp* genes (41).

3.1.3 Calibration of the Sensors

As mentioned before, the voltage information from the detector needs to be converted into a force information. For this purpose the spring constant (the intrinsic stiffness) of the cantilever must be evaluated. There are several techniques to do so: The spring constant can be computed from the geometric and material properties (length, thickness, density, elastic modulus) of the cantilever (43); it can be determined by coupling the cantilever with an additional load (44) or another spring (45). An alternative (used in this work) is to derive the spring constant from the analysis of its thermal noise spectrum (46-48). The Hamiltonian for the system, assuming only oscillations with small amplitude, is given by

$$H = \frac{p^2}{2m_{eff}} + \frac{1}{2} m_{eff} \omega_0^2 q^2 \qquad (3.1)$$

where m_{eff} designates the effective mass, q the displacement, p the linear momentum, and ω_0 the resonance frequency of the cantilever. The relation between thermal energy $k_B T$ and the mean square displacement $\langle q^2 \rangle$ is established by the equipartition theorem:

$$\tfrac{1}{2} m_{eff} \omega_0^2 \left\langle q^2 \right\rangle = \tfrac{1}{2} k_B T \tag{3.2}$$

Using $\omega_0^2 = k / m_{eff}$, one obtains for the spring constant:

$$k = \frac{k_B T}{\left\langle q^2 \right\rangle} \tag{3.3}$$

The Langevin equation for the motion of an externally driven harmonic oscillator in the presence of friction is

$$\frac{d^2 q}{dt^2} + 2\lambda \frac{dq}{dt} + \nu_r^2 q = \frac{F}{m_{eff}} e^{i\nu t} \tag{3.4}$$

where ν is frequency, ν_r the resonance frequency, λ the damping constant, and F the external driving force. The solution for stationary oscillation in thermal equilibrium with the surroundings in the case of small damping ($2\lambda << \nu_r$) and for $\nu \approx \nu_r$ is approximately given by a Lorentzian profile:

$$q(t) = \frac{F}{2 m_{eff} \nu_r \sqrt{(\nu_r - \nu)^2 + \lambda^2}} \cos(\nu t + \delta) \tag{3.5}$$

Measurements of the time-dependent square displacement, i.e. the Fourier transform of $q^2(t)$, yields another Lorentzian for $\hat{q}^2(\nu)$, which can be obtained by a fit to the experimental data of the kind

$$\hat{q}^2(\nu) = q_0 + \frac{A}{(\nu_r - \nu)^2 + B} \tag{3.6}$$

where q_0, A and B are fitting constants. The mean square displacement of the cantilever can now be obtained from the integral

$$\left\langle q^2 \right\rangle = \int_0^\infty (\hat{q}^2(\nu) - q_0) d\nu \tag{3.7}$$

and the spring constant from (3.3).

3.2 Forces and kinetics in SMFS

The experimental data provided by intermolecular SMFS are dissociation forces and molecular elasticities of the whole molecular system, including linker and cantilever. The main question is how these experimental data could lead to quantitative information like off rates or dissociation lengths. Unfortunately the first and very simple assumption that the rupture forces are a direct measure of bond strength was disproven by the experimental data yielding a statistical distribution of rupture forces. It was also found that the distribution of the rupture forces varies with the retract velocity. Since these conclusions, the theoretical interpretation of the rupture forces is a nontrivial task.

The first breakthrough in SMFS was the paper by Evans and Ritchie (49) based on a model by Bell (50) giving rise to quantitative information about the analyzed systems. They formulated that the dissociation under an externally applied force corresponds to the thermally activated decay of a metastable state, which can be described within the framework of classical reaction rate theory.

Before showing how the externally applied force on a molecular bond can be explained by means of classical reaction rate theory, a brief summary of chemical ensemble reaction rate theory will be given.

3.2.1 Kinetics and Thermodynamics

Alternative überschrift (einfach von oben übernehmen): ensemble reaction rate theory

The interaction between a ligand L and its corresponding receptor R can be described by:

$$[L]+[R] \underset{k_{off}^0}{\overset{k_{on}^0}{\rightleftharpoons}} [L \cdot R] \tag{3.8}$$

The brackets [] stand for the respective concentrations of the free ligand L (e.g. the RNA), the free receptor R (e.g. the binding protein) and the complex between both $L \cdot R$. While the on-rate (measured in L M^{-1} s^{-1}) describes the kinetic rate of the forward reaction, the off-rate (measured in s^{-1}) describes the backward reaction. The mean life time of a bond is given by the inverse of the off-rate $\tau = (k_{off}^0)^{-1}$.
Under conditions of constant temperature T and constant pressure the free reaction enthalpy is defined by the difference of Gibbs free enthalpy of the end and initial state:

$$\Delta G = \Delta H - T\Delta S = G_2 - G_1 \,, \tag{3.9}$$

with ΔH the reaction enthalpy and ΔS the entropy change of the system.
The dissociation constant K_D is defined by the ratio of reactants to products:

$$\frac{[L][R]}{[L \cdot R]} = \frac{k_{off}^0}{k_{on}^0} = K_D \,. \tag{3.10}$$

By introducing the free standard enthalpy ΔG^0 (under conditions of constant temperature and pressure) the reaction enthalpy can be written as:

$$\Delta G = \Delta G^0 - RT \ln K_D \,, \tag{3.11}$$

where $R = N_A k_B = 8.314$ $JK^{-1}mol^{-1}$ is the molecular gas constant.
In thermal equilibrium the reaction enthalpy ΔG equals zero and thus, the well-known relationship of the standard free enthalpy for the dissociation process and the dissociation constant results:

$$\Delta G^0 = RT \ln K_D \tag{3.12}$$

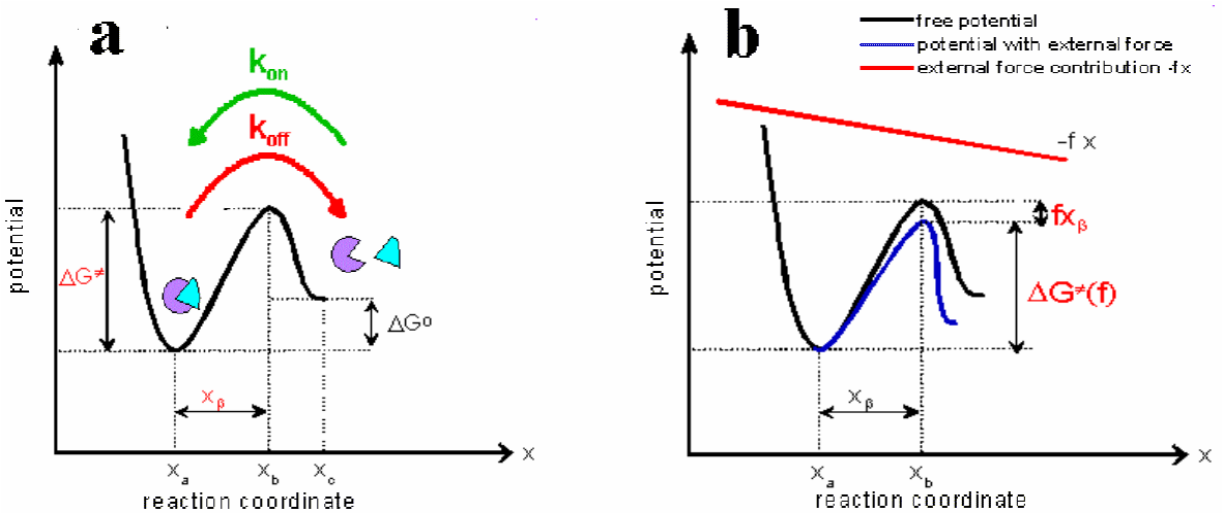

Figure 3.3: Schematic illustration of the energy barriers
(a) Only thermally driven dissociation of a complex from a metastable bound state via a potential barrier of height $\Delta G^{\neq}$ to the unbound state. (b) Externally applied force f lowers the height of the potential barrier.

The dissociation of the complex can be seen as overriding the potential barrier by thermal fluctuations, the nomenclature can be obtained from Fig. 3.3. The quantitative correlation of the height of the potential barrier $\Delta G^{\neq}$ and the velocity of dissociation, and hence the off-rate, was discovered by Arrhenius:

$$k_{off}^{0} = C\, e^{-\beta \Delta G^{\neq}} \quad . \tag{3.13}$$

where $\beta = 1/k_B T$, with k_B the Boltzmann constant and T the temperature.
Although the exponential relationship of (3.13) was discovered by Arrhenius in the late 19th century, it took more than 50 years until Kramers could derive the first expression for the proportional constant C from statistical mechanics. Friction force and the shape of the energy landscape are the main variables which determine the proportional constant (51).

3.2.2 Dissociation under externally applied forces – The standard theory

The dissociation of the complex driven by the application of a constant external force can also be described within the framework of classical reaction rate theory. Figure 3.3 (b) shows a model how the externally applied force lowers the potential barrier, where the reaction length (dissociation length) $x_\beta = x_B - x_A$, (Bell-model), is the difference between the maximum of the potential barrier and the minimum of the meta stable state along the reaction coordinate (49, 52). Although the pre-exponential factor varies under external force, it is kept constant in the model because the equation is dominated by the exponential term. In the Bell-model the reaction length (dissociation length) x_β is also kept constant (52). There are several subsequent works assuming a variation of x_β under external force or other assumptions but none of them is able to solve the problems that will be discussed in section 3.2.3 (53-59).
The applied force f (projected onto the direction of the reaction coordinate) influences the activational barrier therefore:

$$\Delta G^{\neq}(f) = \Delta G^{\neq} - f x_\beta \,. \tag{3.14}$$

Inserting (3.14) into (3.13) and neglecting the force dependence of the pre-exponential factor yields an expression for the off-rate as a function of externally applied force, known as Bell-rate:

$$k_{off}(f) = k_{off}^{0} e^{\beta f x_\beta}. \tag{3.15}$$

The acting force on the bond molecules is not constant but the changes are very slow in comparison to molecular relaxation processes so that the reaction kinetics can be approximated by:

$$\frac{dp(t)}{dt} = -k_{off}(f(t))p(t) \tag{3.16}$$

where p(t) denotes the survival probability of the bond.
Another common assumption developed by Evans and Ritchie is that the force f(t) depends solely on the total extension s=vt of all elastic components (molecules, linker, cantilever etc.):

$$F(s) = F(\mathrm{v}\cdot t) = f(t) \tag{3.17}$$

where F(s) is independent of the retract velocity.
In particular it is assumed that if the retract velocity v is kept constant, the force acting on the complex varies temporally according to:

$$f(t) = \kappa_{eff}\,\mathrm{v}\cdot t = r\cdot t, \tag{3.18}$$

where κ_{eff} denotes the effective spring constant for the system which is derived from the spring constant k of the cantilever and the elasticity of the polymer linker attached to the tip. r is called loading-rate.
With (3.16) the formal solution of the survival probability of the bond under an externally applied force f for any $k_{off}(f)$ and F(s) is given by:

$$p_{\mathrm{v}}(f) = \exp\left\{-\frac{1}{\mathrm{v}}\int_{f_{\min}}^{f} df' \frac{k_{off}(f')}{F'(F^{-1}(f'))}\right\} \tag{3.19}$$

with $p_v(f(t))=p(t)$ and $p(t=0)=p_v(f=f_{min})=1$. Here, f_{min} is the threshold value below which dissociation forces cannot be distinguished from thermal fluctuations. Additionally, it is assumed that F(s) is strictly monotonic increasing (for the inverse F^{-1} to be existent).
Using the two assumptions x_β=constant and $f(t) = r\cdot t$, it is possible to integrate the survival probability for the bonds (3.19):

$$\mathrm{p}_{\mathrm{v}}(f) = \exp\left\{-\frac{k_{off}}{\mathrm{v}\kappa_{eff}} \frac{e^{\frac{x_\beta f}{k_B T}} - e^{\frac{x_\beta f_{\min}}{k_B T}}}{\frac{x_\beta}{k_B T}}\right\} \tag{3.20}$$

Using the Bell-rate (3.15), the most probable rupture force $\hat{F}$ at a given loading rate r can be derived from the maximum of the distribution $-dp_v(f)/df$:

$$\hat{F} = \frac{k_B T}{x_\beta} \ln\left(\frac{x_\beta r}{k_{off}^{0} k_B T}\right) \tag{3.21}$$

This relation is the basis for standard dynamic fore spectroscopy according to Evans and Ritchie. By varying the pulling velocity over several magnitudes, force distributions for each loading rate can be obtained and then be analyzed by plotting the most probable rupture forces semilogarithmically against their respective loading rates. According to (3.21) x_β can be estimated from the slope of a linear fit to the data and finally, by extrapolating the regression line to zero force ($\hat{F}$ =0), a value for the off-rate can be gathered.

Up to this point it has always been assumed that there exists only one single, well defined energy barrier along the dissociation path, although the existence of additional intermediate energy barriers cannot be neglected.
The simplest case of such a intermediate barrier is shown in Fig. 3.4.

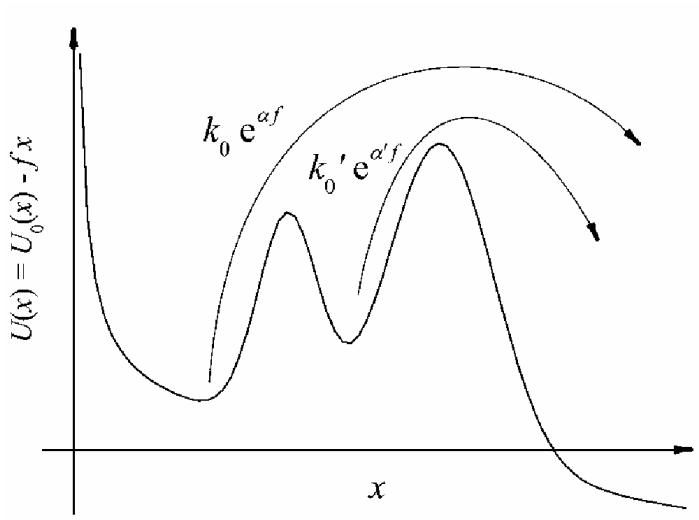

Figure 3.4: Illustration of an intermediate energy barrier

This intermediate energy barrier leads obviously to different off-rates and has been quite often discussed (57). In (60) it was exemplified that this extension cannot be the reason for the discovered problems of the standard theory (see below), although it is quite an interesting option for the explanation of some other phenomena.

3.2.3 Inconsistencies of the standard theory

Since the time-force interrelationship of the rupture curves (i.e. F(s)) is of importance for the survival probability of the bonds, it is necessary to fit these curves properly. According to the standard theory which assumes a linear force extension (3.17) the common way was to fit a straight line at the point of rupture a few nanometers into the loading. While this fit shows good results for many systems in the higher force area (starting at about 60pN), it is improper in the low force range.
In chapter 5.1.1 it will be shown that the force distance characteristic with F(s)=F(vt) are in good agreement with a second grade polynomial:

$$F(s) = a_0 + a_1 s + a_2 s^2 \tag{3.22}$$

leaving the independence of the pulling velocity untouched. Although this modification has some effects on the used derivation (especially the integral of (3.19) cannot be solved analytically anymore (at least no solution is known)) this is not the main reason for the problems that will be discussed now.

In 2003 a severe discrepancy between the experimental data and the standard theory was discovered (61).

Starting from the survival probability of the bonds $p_v(f)$ at pulling velocity v (3.19), a function g(f) can be defined as:

$$-v \ln p_v(f) = g(f) \tag{3.23}$$

Under the assumptions (3.16) and (3.17) this g(f) should be of the form:

$$g(f) = \int_{f_{min}}^{f} df' \frac{k(f')}{F'(F^{-1}(f'))} \tag{3.24}$$

and thus independent of the pulling velocity.
This correlation cannot only be calculated, but also estimated from the experimental data. With a given data set consisting of N_v rupture forces f_n (n=1, …,N_v; $f_n > f_{min}$ for all n) measured at given retract velocity v the true probability of bond survival $p_v(f)$ can be estimated to

$$\tilde{p}_v(f) = \frac{1}{N_v} \sum_{n=1}^{N_v} \Theta(f_n - f) \tag{3.25}$$

where $\Theta(x) = \int_{-\infty}^{x} \delta(y)\,dy$ is the Heaviside step function $\left[\Theta(x<0)=0, \Theta(x>0)=1\right]$. For every finite number of rupture events N_v formula (3.25) is the best estimate for $p_v(f)$ without making any further assumptions. It allows an estimate of the "true" integral in equation (3.24):

$$\tilde{g}_v(f) = -v \ln \tilde{p}_v(f)\,. \tag{3.26}$$

Following the assumptions of Evans and Ritchie (3.16) and (3.18) the expression in equation (3.23) must be, apart from statistic fluctuations, independent of the pulling velocity v. Plotting the functions $\tilde{g}_v(f)$ resulting from different retract velocities in a diagram, Fig. 3.4, should collapse to a single master curve. That this is definitely not the case was shown in (62). Moreover, not one single ligand-receptor system analyzed by force spectroscopy was found to show this predicted collapse. Even experiments using a micropipette rather than an AFM as the force transducer showed this severe divergence for different retract velocities (63).

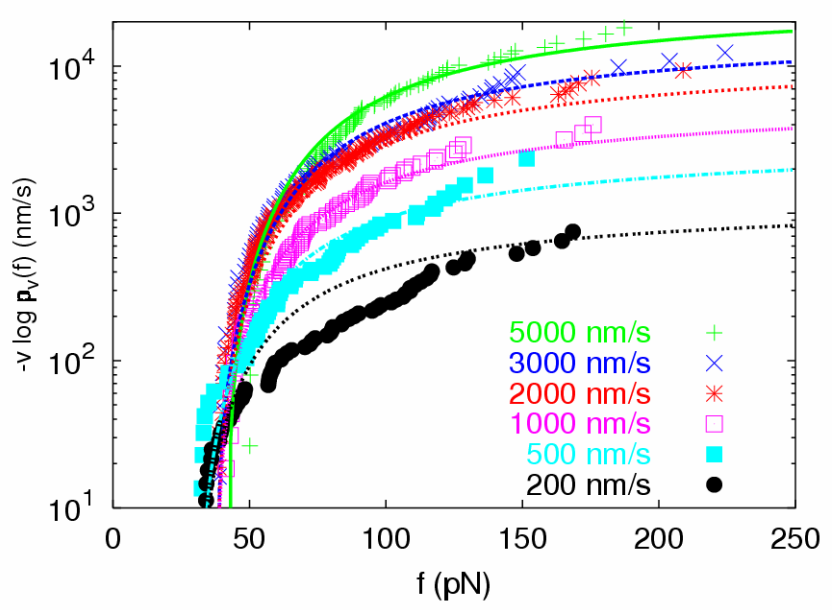

Figure 3.5: Plot of v log(p_v(f)) for different pulling velocities

The data is obtained from ExpR-DNA with effector C_{10}-HL (cf. Fig. 6.2). According to the standard theory the force distributions should collapse to a single master curve and thus be independent of the pulling velocity.

The plot also includes the estimated survival probabilities *(3.27)*, calculated with the theory of bond heterogeneity, for each pulling velocity (dotted lines). The estimated lines do not fall together in the low force regime because of different f_{min} for different retract velocities (cf. Fig. 6.2)[1].

3.2.4 Theory of heterogeneity of chemical bonds

In (60) a new promising approach is presented which postulates a heterogeneity of bonds to account for the several observed inconsistencies. The main assumption is that both, x_β and k_{off}^0 are statistically distributed.

There are several reasons justifying such an intrinsic random distribution of the rate of dissociation but not all of them need to be present in a real SMFS experiment:

1. geometrical variations like, e.g., different orientations of the complex relative to the direction of the applied pulling force;
2. Random variations and fluctuations of the local molecular environment by ions, water, and solvent molecules locally modulating ionic strength, pH, and electric fields, which may influence the dissociation process of the molecular complex;
3. Structural fluctuations due to thermal activation may lead to different conformations of a (macro-) molecule;
4. complex bio-molecules may have more than one binding domain;
5. unspecific events could be misinterpreted as “true” specific binding events.

The new methods for the analysis of the force distance curve presented later in this work may give some new answers to these problems. Nevertheless, these ideas can be quantified by an ad hoc ansatz as mentioned above assuming a statistically distributed α ($\alpha = x_\beta / k_B T$) and also

[1] *Of course, it was ascertained that even for same f_{min} the curves do not collapse to a single curve*

k_{off}^0 , while leaving the Bell approximation untouched. Thus the values for the parameters α and k_{off}^0, combined in $\vec{\lambda} = (k_{off}^0, \alpha)$, change at every repetition of the experiment according to a certain (conditional) probability density $\rho(\vec{\lambda};\vec{\mu})$, depending itself on some fit parameters $\vec{\mu}$.
Hence, only an averaged survival probability $\overline{p}_v$ can be measured:

$$\overline{p}_{\mathrm{v}}(f;\vec{\mu}) = \frac{\int d\vec{\lambda}\, \rho(\vec{\lambda};\vec{\mu})\, p_{\mathrm{v}}(f;\vec{\lambda})}{\int d\vec{\lambda}\, \rho(\vec{\lambda};\vec{\mu})\, p_{\mathrm{v}}(f_{\min};\vec{\lambda})} \tag{3.27}$$

where $p_v(f;\vec{\lambda})$ is in principal the survival probability from (3.18):

$$p_{\mathrm{v}}(f;\vec{\lambda}) = \exp\left\{-\frac{1}{\mathrm{v}}\int_{f_0}^{f} df' \frac{k_{off}(f')}{F'(F^{-1}(f'))}\right\} \tag{3.28}$$

with the little difference that f_{min} is replaced by the force f_0 at the beginning of the pulling. The actual value depends on the experimental actualities. However, a common value for theoretical calculations is zero, meaning that no force acts on the bond before pulling starts. The denominator in (3.27) accounts for normalization.
Since there is no linear dependence assumed, this integral cannot be solved analytically (better: no solution is known) anymore which means that the values of the numerical solution have to be inserted in (3.27).

In (60) it could be shown that randomization of k_{off}^0 does not improve the consistence of the theory with experimental data. A randomization of the linker elasticity as proposed by Friedsam et al.(55, 64) essentially corresponds to a randomization of the off-rate and does not explain this phenomenon. Since the randomization of k_{off}^0 has not as much implications as the randomization of α because of the exponential characteristics of the Bell-rate (3.14), k_{off}^0 will be taken as a fix parameter.
The probability distribution of α is assumed as a (truncated) Gaussian:

$$\rho(\alpha) = \rho(\alpha; a, \sigma) = C\, \exp\left\{-(\alpha - a)^2 / 2\sigma^2\right\} \Theta(\alpha)\,. \tag{3.29}$$

Although it was shown that the exact form of the distribution is not very important, it should be mentioned that this Gaussion fitted the experimental data best among the others tried.
Succeeding these assumptions the three fit parameters k_{off}^0, a and σ ($\vec{\mu} = (k_{off}^0, a, \sigma)$) determine the averaged survival probability (3.27).

The negative derivative of the survival probability (3.27) to force,

$$\wp(f;\vec{\mu}) = -d\overline{p}_{\mathrm{v}}(f;\vec{\mu})/df\,, \tag{3.30}$$

yields the probability density to observe a rupture force. For N independent experiments the probability to measure a set of rupture forces $\{f_i\}$ is the product of the single probabilities:

$$\wp(\{f_i\};\vec{\mu}) = \prod_{i=1}^{N} \wp(f_i;\vec{\mu})\,. \tag{3.31}$$

The final challenge is then to acquire the parameter $\vec{\mu}$. The probability (3.31) is assumed as a function of $\vec{\mu}$ and the 'true' parameters are estimated by the parameters maximizing this function. In the literature (65) this is known as a maximum likelihood estimate, accordingly the broadness of the distribution is used to calculate the statistical uncertainties of the parameters. Since the off rate is estimated in its logarithmical expression, $\ln(k_{off}^0)$, the statistical uncertainties, are left as $\Delta\ln(k_{off}^0)$ while the off rate can be easily derived from the natural logarithm.
For a detailed discussion of this approach please refer to (66).

4 Materials and Methods

4.1 Immobilization of the biomolecules

In single molecule force spectroscopy experiments it is necessary that the binding of the molecules to their respective surfaces (tip and sample) is much stronger than the expected rupture forces. The formation of covalent bonds between the attached molecules ensures that this requirement is achieved.

The main focus on the experimental side is laid onto the RNA-protein interactions, but additionally some control measurements of a DNA-protein interaction had to be performed. In consequence, the description of the DNA-ExpR preparation part has been abbreviated but can be found in full detail in (67).

4.1.1 Preparation for RNA-AtGRP7/8 experiments

Si_3N_4 cantilevers (Microlevers, Veeco Intruments) were first dipped for about 3 seconds in concentrated nitric acid for activation and then incubated with a solution of 2% aminopropyltriethoxysilane (Sigma-Aldrich, Seelze, Germany) in dry toluene (Fluka, Seelze, Germany) for 90 to 120 minutes at ambient temperature. The used RNA oligos (Table 4.1) with a thiol (-SH) group (biomers.net –The Biopolymer Factory, Ulm, Germany) were heated up to a temperature of 70° C for 3 minutes and then the PEG linker (N-hydroxysuccinimid-poly(ethylenglycol)-maleimid (NHS-PEG-MAL, MW 3.4 kD; Nektar, Huntsville, Alabama, USA)) was added in a 1:1 molar ratio. After washing the cantilevers with toluene and autoclaved water, the tips were incubated with 0.5 pmol/µl RNA and PEG linker solution for at least 90 minutes at 4° C. The functionalized cantilevers were then washed with the binding buffer [20 mM HEPES-KOH, pH7,5, 100 mM NaCl, 1 mM $MgCl_2$, 0,01% NP40] and could be used for force experiments.

used label	"official" label	sequence (5' -.....-3')
RNA7	GRP7_UTR-SH	AUUUUGUUCU GGUUCUGCUU UAGAUUU
mutant RNA7	GRP7_UTR_GMut-SH	AUUUUAUUCU AAUUCUACUU UAGAUUU
RNA8-938	GRP8_938-SH	CGUUUGGUUU ACUUUUUUGA UGAAACA
mutant RNA8-938	GRP8_938_GMut-SH	CGUUUAAUUU ACUUUUUUAA UAAAACA
RNA8-886	GRP8_886-SH	GUUUUUGGUU UAGAUUUGGU UUUGUGU
mutant RNA8-886	GRP8_886_GMut-SH	GUUUUUAAUU UAAAUUUAAU UUUAUGU
poly(A)	Poly_A	AAAAAAAAAA AAAAAAAAAA AAAAAAA
poly(U)	Poly_U	UUUUUUUUUU UUUUUUUUUU UUUUUUU

Table 4.1: The used RNA oligos

The AtGRPs were recombinatly expressed in *E. coli* with an N-terminal Gluathione S-Transferase (GST) tag for subsequent purification via affinity-chromotography. The purified fusion protein was tested in radioactive EMSA for activity. The comperative value for these activity test was the cleaned up AtGRP protein without GST-tag. No difference could be detected concerning the activity. The protein expression, purification and testing was done by Jan Schöning (department of molecular cell physiology).
For immobilization of the protein on the mica surface (Provac, Florida, USA), the cross-linker BS3 (bis(sulfosuccinimidyl)suberate) (Sigma-Aldrich, Seelze, Germany) was used. The fusion protein bears several target amino groups for the cross-linker, thus decreasing the probability of binding directly to amino groups in close vicinity to the RNA binding site. However, since the cross-linker is a homo-bifunctional molecule, there is the unwanted side effect that two or more proteins may bind to each other, causing long protein-linker chains. To avoid this problem the proteins were immediately applied on the sample after incubation with the linkers. A protein-to-linker molar ratio of 1:1 at a concentration of 0.5 to 1 pmol/µl showed good results while a ratio of 1:5 (five-fold excess of linker) tends to result in multiple rupture events.

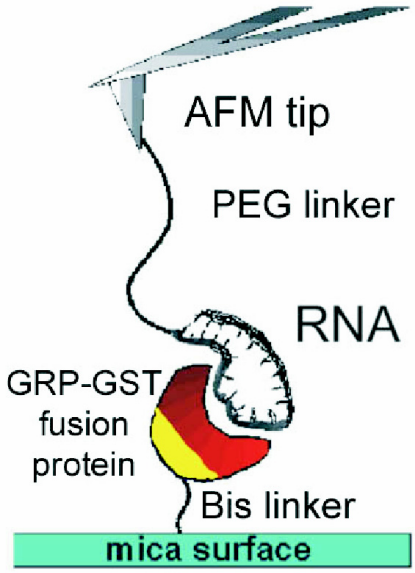

Figure 4.1: Illustration of the immobilization

4.1.2 The preparation for ExpR-DNA

The DNA fragments with a length of 216 bp were functionalized with a thiol (-SH) group by using a modified primer in the PCR. The silanized tips (see above) were incubated in phosphate buffer including the PEG linker (1mM NHS-PEG-MAL in 0.1M KH_2PO_4, pH 8.0) (Nektar) for about 2 hours. After washing the tips with phosphate buffer they were incubated over night at 4° C with a solution of 10ng/µl SH-DNA in this buffer solution.

The proteins ExpR (MW 28kDa) bearing 11 lysine residues were also provided by Matthew McIntosh. The proteins were attached to the mica surface via the BS3 crosslinker (Sigma-Aldrich) in a 1:5 ratio (five-fold excess of linker). Coupling with the BS3 linker needs a buffer which is free of primary amino groups; after the proteins are linked to the surface, however, it is possible to perform the experiments in the standard binding buffer (for ExpR:

100 mM KH_2PO_4, pH 7.5) without loss of their functional properties. The DNA fragments and proteins were provided by Matthew McIntosh (department of genetics).

4.2 Instruments

The first series of experiments was performed with a commercial AFM head (Multimode, Veeco Instruments) using a 16 bit AD/DA card (PCI 6052E, National Instruments) and a high-voltage amplifier (600H, NanoTechTools) controlled by a home-built software based on LabView (National Instruments). For the second series, an MFP-3D (Asylum Research) was used. The calibration of the cantilevers and all force spectroscopy experiments were done with the provided software based on Igor Pro 5 (Wavemetrics). The whole data analysis was done with matlab (MathWorks), although the Igor data files (Waves) needed to be converted in matlab-readable files. Due to the fact that the programming for data analysis (e.g. automation) was a major part of this work, all important improvements in the data analysis programs will be addressed in detail in the next chapter.

5 Data Analysis

The analysis of force spectroscopy (FS) experiments is still a difficult but also very important task. While the standard experiments are relatively easy to perform with a modern AFM (the sample only needs to be driven via the vertical axis) the analysis of this data is not only a time consuming but also tricky job that needs a lot of experience. At the beginning of this work the analysis of the force distance curves was done with a matlab program written mainly by Robert Ros with modifications by Rainer Eckel. Although the program could detect rupture events automatically, every single force curve has to be watched to confirm the found rupture manually. Beside the rupture force the program took the slope[2] of a fitted line about 4nm from the rupture event. The forces and slopes were used for the analysis according to the standard theory of Evans and Ritchie (3.21).
During this work a novel data analysis software was developed. This software yields many improvements compared to the old analysis program. The most important ones are centralised in 5.1. The principle procedure developed for the analysis of single molecule force experiments (SMFS) will be described and some details will be discussed in 5.2.1-3.

5.1 The main improvements at a glance

1. The software has two graphical user interfaces (GUIs) that enable even a novel experimenter to do a complete data analysis since all needed parameters can be set from the GUIs. Of course, all presented plots shown in this work (except the illustrations and Fig. 6.12) can be done with a few mouse clicks from these GUIs.
2. The software prepares the "raw" force distance curves with a very high accuracy while the range of tolerance can also be set. At this point artefacts like noise, oscillations of the signal or adhesion are detected. After this is done, a subroutine does a scan for rupture events. It can also be selected whether the program only searches for the last occurring rupture (as used in the later presented results) or all ruptures.
3. The detected rupture curves (resp. the loading) are fitted also completely automated. Beside the fitting parameters and the force, 9 additional parameters are saved for every rupture found. This allows a fast characterization of the ruptures after the scan of the complete data set has finished.
4. The high accuracy of the detected rupture events gives rise to a new kind of plot developed during this work. This plot even allows to identify different binding modes of the RNA-protein complex that are not noticeable in a standard histogram plot.
5. To meet the demands of the (modified) theory (cf. Chapter 3) it is also possible to find (resp. define) a master curve. The software then automatically compares every detected rupture curve with this master curve and decides which curves are within the tolerances.
6. Then, for the final step, the prepared data needs to be copied (manually) to a program written in the C programming language by S. Getfert that does the maximum likelihood estimate to gather the quantitative results.

[2] usually this is called "elasticity" but since the molecular elasticity is actually the reciprocal of this value it is termed "stiffness" in this work

5.2 The three parts of Data Analysis

The new data analysis provided in this work can be divided in 3 parts:

1. finding rupture events and classify them by avoiding noise and double rupture events
2. getting the "master curve" and compare the rupture curves with it
3. final data analysis based on the modified theory

5.1.1 Part 1

At first it should be mentioned that the data delivered by the two different AFMs (Multimode and MFP-3D) have little differences. The only difficulty with the MFP-3D data is that the force curves are only saved in the programming language (Igor Pro) provided by this AFM. The Igor "Waves" were converted into matlab files keeping all the important additional information (e.g. spring constant, dwell time, temperature).
After converting the deflection signal of the Multimode´s data into forces by sensor response and spring constant the data of both instruments were alike.

The aim in single molecule force spectroscopy is to have only one molecule binding at one molecule on the sample. If the cantilever pulls at more than one molecule the acting force on each molecule cannot be separated and thus, these values must not be used for the statistics. One common method to avoid these double or even multiple rupture events is to lower the concentration of the two binding partners. On the other hand, this also decreases the overall rupture probability and thus one of the difficult tasks of the experiment is to find a suiting medium. Since the data analysis has been automated it would be possible to lower the concentration much more without losing time during the data analysis. Of course, this is not the most elegant way and thus, also a routine that ensures that only single molecules events are detected, was developed[3]. There can be two different kinds of flawed detection of multiple ruptures. The first and relatively easy one to detect is indicated if the cantilever does not immediately jump back to the base line after the detected rupture happened. The other one is quite tricky to detect in an automated analysis. If the cantilever pulls at two molecules but one dissociates during the load, the forces of this rupture event are mostly low. A fitting routine (see below, $\hat{f}_{min}$) for the loading that automatically stops fitting if the loading is disturbed of other molecules could be developed.
Of course, these verifications do not exclude every artefact but all parameters are saved and enable a check as described in *Part 2* of the data analysis.

[3] If multiple molecule bonds disrupt simultaneously (resp. chronologically close together) there is no way to detect this as a multiple rupture event but since the dissociation process is still driven by thermal fluctuations this is very implausible.

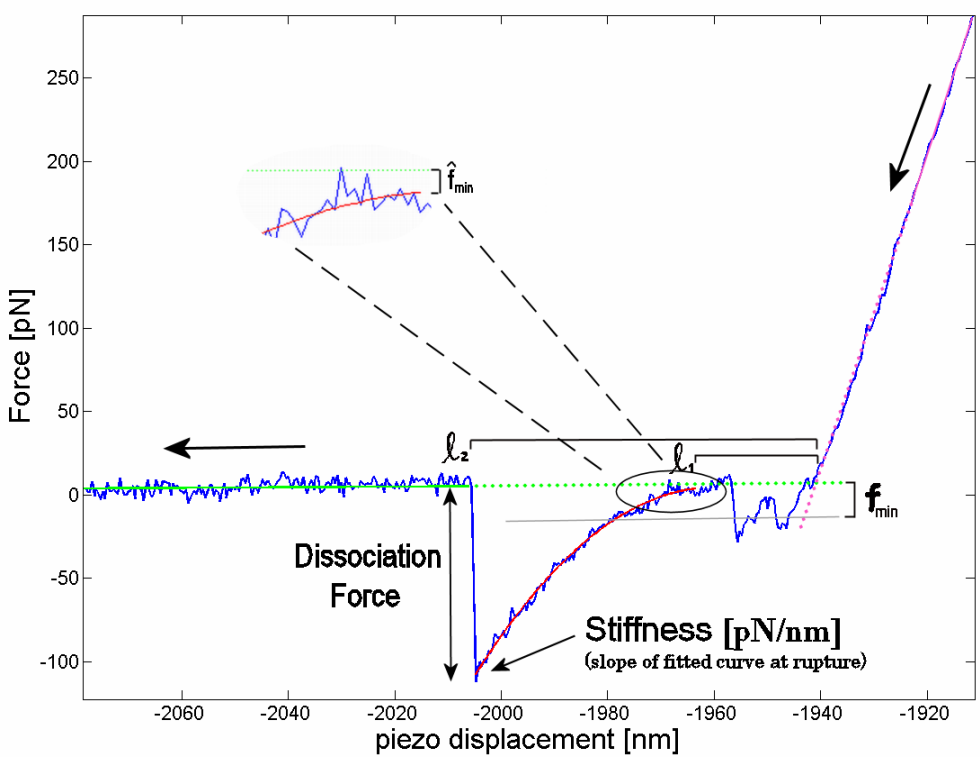

Figure 5.1: Typical force distance curve (only retractive part shown) with an illustration of the analysis method

The base line (green) is extrapolated until the point of intersection with the extrapolated line (magenta) of the contact area ("sensor response"). The (estimated) beginning of the load is the starting point of the fitted (red) curve at ℓ_1. The slope of the fitted rupture curve at extension ℓ_2 is called "Stiffness" and the "Dissociation Force" is obtained by the difference of the force value of the fitted curve (red) at extension ℓ_2 and the (green) base line. The pathway of the rupture curve from ℓ_1 until ℓ_2 is called loading (resp. the load). Values for f_{min} can be set manually. For further data analysis (*Part 2*) only curves beginning above the f_{min} line are accepted ($\hat{f}_{min} < f_{min}$).

For the further theoretical analysis of the gathered data the time dependent development of the rupture curves is of major concern giving the necessity to fit into the noisy rupture curves. In the former analysis a straight line was assumed. Since the time dependence is of great importance it is necessary to know at which time point (resp. extension) forces affect the bond. For this a (force-) offset f_{min}, dependent on noise and other artefacts, is introduced at which all accepted rupture curves definitely have to exist (cf. Fig 5.1, force difference of green dotted base line and grey "offset" line). The value for the force offset of a fitted rupture curve, $\hat{f}_{min}$, is the difference of the base line and the beginning of the load (cf. Fig. 5.1, force difference of green dotted base line and fitted curve (red) at ℓ_1.

Due to noise and other artefacts (increasing with the pulling velocity) the value for f_{min} increases also from 32 pN for the low velocities (< 500 nm/s) to 45 pN for the high ones (> 6 µm/s) simply to get more accepted rupture events for better statistics. With a dataset providing many ruptures it was ascertained that the choice of the offsets does not affect the further data analysis, of course, only as long as the aimed peaks do not vanish.

Although most systems in SMFS consist of many molecule types, like linkers, DNA and proteins, an adopted polymer model (like worm like chain (WLC) for DNA) should be the best way for fitting. The cantilever movement could be easily accounted for and added

afterwards because it has an intense influence on the time-force dependence and should not be neglected. On the other hand, an (automated) WLC fit would be quite difficult while providing only little improvement on the method presented, and since many thousands curves need to be analyzed it is pretty unrealistic to do it in reasonable time.
In this work the rupture curves are fitted by a simple second grade polynomial. These polynomials describe the pathways during the load very well (cf. Fig. 5.1) and have the very important advantage of allowing an easy and, most important, automated analysis of the force distance curves taken. The second postulate of the standard theory according to Evans and Ritchie (3.18) remains unaffected by a non-linear time-force dependence. A numeric solution of the integral in (3.20) is possible.

The polynomial fit also leads to data for rupture force and stiffness being closer to the true values. For rupture forces the value of the polynomial is taken which compensates noise that causes a lower rupture force (cf. Fig. 4.2, force difference of fitted rupture curve (red) from green dotted zero line at ℓ_2). Due to thermal oscillation the tip is always deflected by moving up and down, but also the force is measured by the deflection. So when the cantilever moves up during oscillation the measured force decreases while the real force acting on the molecules increases. This is the reason why most rupture events appear at "lower" forces[4].
Taking the slope of the polynomial is obviously a more authentic way than fitting a line into the last points of the rupture curve.
Beside some parameters indicating the quality of the rupture events (e.g. noise, double rupture events, adhesion) the coefficients of the fitted polynomials are saved.
Although the program could correct the movement of the cantilever during a rupture curve the data presented is not corrected in this way because the theoretical analysis requires all aspects, including cantilever movement, to describe the processes.
This first part of the data analysis has a scan rate of about 200 force distance curves per minute with a standard PC (Igor Waves need much more time due to the complex data conversion). This means that it enables the user to get the reaction length and off-rates according to the standard theory of Evans and Ritchie (3.21) in less than one hour, even with higher quality. Analysing each force curve "by hand" with the former analysis software usually takes more than one day of concentrated work for the same result.

5.1.2 Part 2

One requirement for further theoretical analysis is that all rupture curves describe the same pathway (resp. have the same time-force dependence) during the load. The task of this part of the data analysis is to make sure that there is really only one such pathway, called master curve, and to exclude all artefacts that always appear in experimentally gathered data.

[4] By mischance this effect does not appear in Fig. 5.1.

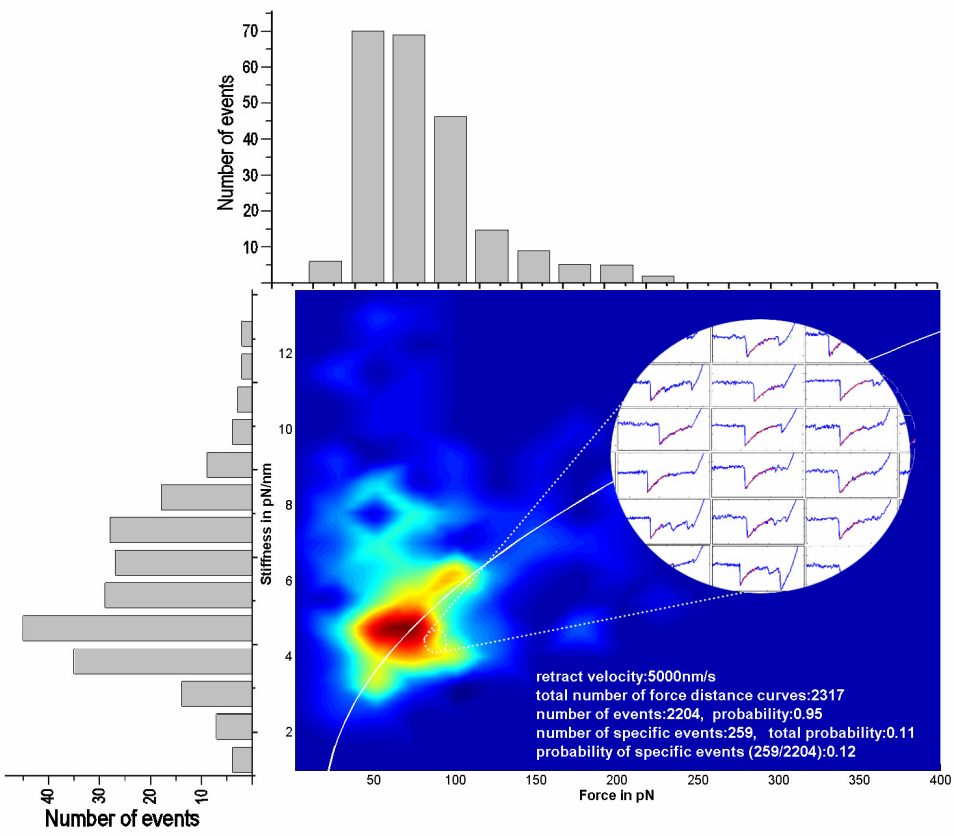

Figure 5.2: Plot of stiffness against rupture force as a 2d-histogram
The 2d-histogram also includes the respective values for the master curve (white solid line). The software allows also to look at the rupture curves of any area being of interest which is also illustrated here. Further explanations in text.

Taking the slope of the polynomial at point of rupture gives rise to a new kind of plot of stiffness against force. The values are binned in a 2d-histogram (Fig. 5.2) and only for better visualization a matlab program can do a fit between each element resulting in a height coded colour plot which can also be displayed as a 3D object (Fig. 5.3).
If only one process during force spectroscopy would happen the plots should only show one single peak for a certain pulling velocity. Plotting all pulling velocities in one graph should consequently lead to one universal (master) curve (depending mainly on the linkers and the cantilever). For some data (i.e. ExpR-DNA) these single peaks were found while some plots (i.e. RNA-AtGRP) suggest at least two peaks and some few plots can only be explained by introducing a second master curve. To check whether the accepted ruptures for this kind of plot are "contaminated" with artefacts like noise, flawed fits or multiple rupture events it is possible to select an area in the 2d-histogram and then, all the rupture curves of this spot can be plotted as illustrated in Fig. 5.2. If this plot shows artefacts, the parameters can be readjusted until only ruptures without artefacts are accepted. If noise or double rupture events are the main problem the analysis of *Part 1* should be done again to allow a comparison of the data being described in the following paragraph.

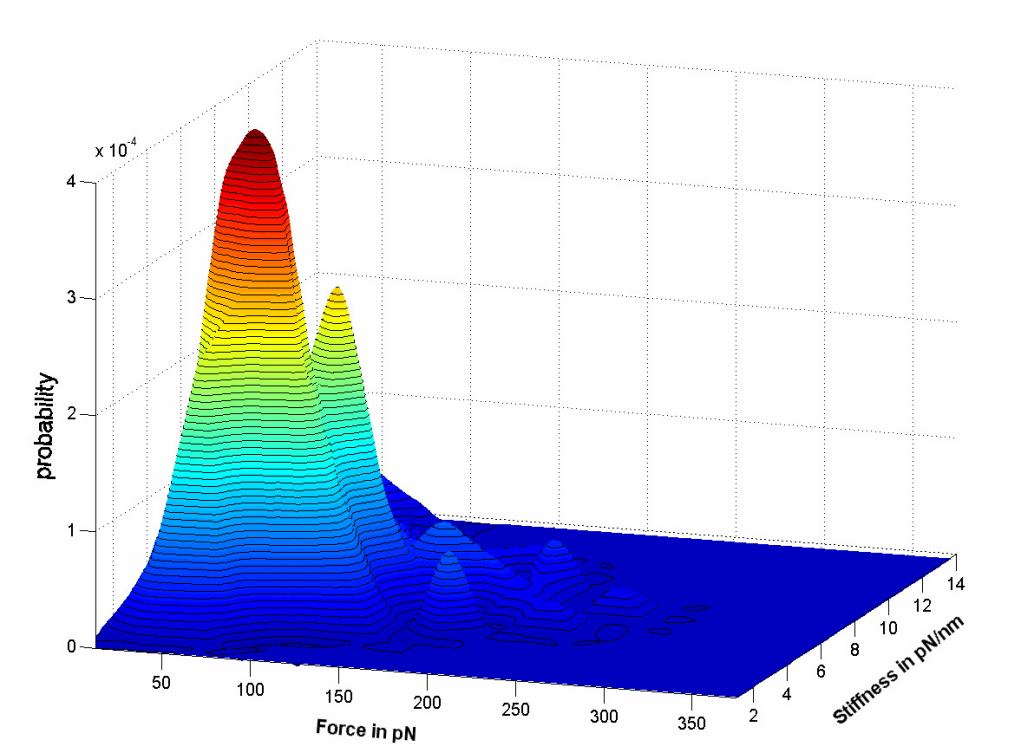

Figure 5.3: Same data as for Fig. 5.2 but plotted as a 3d object.

The program provides also some additional information. The number of analysed force distance curves at the displayed pulling velocity (number of total force distance curves), the number of detected rupture events (number of events) and the respective probability are plotted. The "number of specific events" gives information about the ruptures having a set of parameters being within the adjusted tolerances. Only these ruptures are used for the 2d-histogram plots. The ratio of detected ruptures in *Part 1* to the accepted ruptures is also plotted and gives rise to the character of the processes involved. By normalizing to this value the difference of binding of the wild type RNA in comparison with the mutant RNA will be shown in Fig. 6.3.

The 2d-histogram plots also give rise to a much higher "resolution" in the low force regime since here the main changes occur in the slope of the load. Especially for the protein-RNA interactions the reliability of rupture forces in this low force regime is not as good as the information of the "stiffness" because of some (mainly) attracting forces deflecting the cantilever[5]. These forces, depending on the surface charges, seem to attract, and thus bend, the cantilever toward the sample. This force seems to be relatively constant in the low force regime of one pull, but varies depending on the x-y-position (presumably due to the proteins on the surface[6]). This results in an additional broadening of the distribution of rupture forces. Because of the apparent constancy of attracting forces in this regime it does not seem to cause such an additional variation of the slope (resp. stiffness).

To get a master curve (or even two) all polynomials of one binning field can be plotted, then a force offset can be set at which all curves are automatically scrolled and another polynomial is fitted through the selected curves, being then the "master curve" (cf. Fig. 5.4, (b) red line). The curves of all other fields can be compared with this master curve. The decision which curves are accepted is controlled by 2 parameters. One sets the maximum relative deviation of the slope at rupture point from the slope of the master curve at this point. The other one sets

[5] Due to the short overall length of the linker-RNA-protein chain, many rupture events appear quite near to the surface.

[6] It could be observed that the rupture probability and the adhesion are interdependent.

the value for an interval laid around the master curve (cf. Fig. 5.4 (b) green lines). All force curves cutting this interval will not be accepted. For long force extension curves (=high rupture forces) this method works without problems but to analyze a common data set some improvements had to be done.
Sometimes the fit of the rupture curve in *Part 1* stops before reaching the force offset f_{min} due to noise. Losing those rupture would be a waste so all rupture curves not reaching the offset are refitted until reaching f_{min} ($\hat{f}_{min} = f_{min}$) while the other ones are kept untouched conserving their greater information. The reason for the offset is only to make sure the rupture curves existed already at this point. If a rupture curve existed even longer the fit will be much better showing better results especially in the low force area (cf. Fig. 4.2, the fitted curve (red) would be left untouched because $\hat{f}_{min}$ is lower than f_{min}).
The effects of taking only the rupture curves along the master curve are illustrated in Fig. 5.4. While the histogram of the "raw" data provided by *Part 1* of data analysis shows quite a wide distribution (a), the histogram of only the accepted curves is much narrower (c).

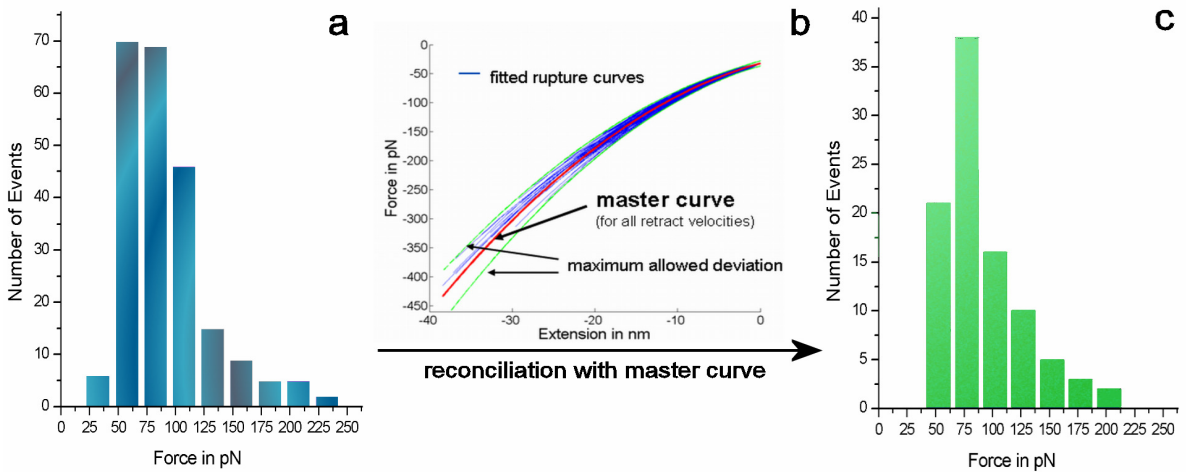

Figure 5.4: Reconciliation of the "raw" data (a) with the master curve (illustration) (b) yields a much clearer histogram (c)

Plotting every single pulling velocity side by side in a 2d-histogram shows that there is always a little, but definitely existing, deviation from the (universal) master curve. For slow velocities the peaks are always a little bit left shifted to the master curve while it is for higher velocities the other way round. As mentioned these deviations are relatively small and it cannot clearly be said whether the reason lies on the AFM side (the inertia of the cantilever should not be the reason since the "zero" line is fitted into the data) or on the physical properties of the single molecules (e.g. stretching of the linker).

5.1.3 Part 3

The gathered information, the parameters of the master curve, accepted rupture forces with the corresponding retract velocities and force offsets (f_{min}) are taken to estimate σ, a and k_{off}^{0} with a program written by Sebastian Getfert (department of condensed matter theory) based on the modified theory briefly explained in part 3.2.4 of this work. This program is written in the C programming language. The parameters are estimated by maximising the likelihood function which is done numerically employing a commercial minimization algorithm (e04ucf) from the NAG library.

6 Results and Discussion

The investigated biomolecular interactions in this thesis belong to two different levels of gene regulations. At first the force spectroscopy data taken of the specific binding of the transcriptional regulatory protein ExpR at the sinR/sinI-DNA sequence of *Sinorhizibium meliloti* is applied to the introduced methods of data analysis. The second analyzed biological system is part of the post-transcriptional control of the *Arabidopsis thaliana*. For this protein-RNA interaction the introduced 2d-histogram plots will demonstrate their advantages enabling the identification of different binding modes of this complex.
The principal setup for both systems was alike. The DNA (resp. RNA) was attached to the tip via a long PEG-linker while the proteins were immobilized on a mica surface via a short linker.

6.1 ExpR-DNA

6.1.1 Effector molecules activate the binding

For the first force spectroscopy measurement DNA-fragments were covalently bound to the tip and the proteins in the same way to the surface. The retract velocity was always kept constant (2 µm/s) as well as the approach velocity (5 µm/s). The resulting rupture force distributions are plotted in Fig. 6.1 (a), showing only some few interactions. After adding 10 µM $C_{16:1}$-HL the binding probability rises drastically and a familiar distribution of forces is established, (b).
This demonstrates that there are only some rupture events due to unspecific attraction, p.e. electrostatic interactions, but the probability of binding is highly increased in the presence of a proper effector. The effector molecules stimulate the binding of the protein ExpR to its DNA-target sequence.

6.1.2 Specificity of binding

To verify that the observed interactions belong to a specific binding, a competition experiment was performed (Fig. 6.1(c-d)). Adding one binding partner (free DNA fragment) in excess during the experiment while leaving the $C_{16:1}$-HL concentration constant blocks available binding sites of the proteins and thus the rupture probability decreases (c). Washing the sample carefully with binding buffer in order to remove all free molecules and adding $C_{16:1}$-HL at prior concentration restores the initial unbinding characteristics (d). Here, a very often appearing effect can be observed: after washing the probes with binding buffer in a competition experiment the overall rupture probability increases (although everything lies within the range of statistical fluctuations). This effect will be discussed in the RNA section where sufficient data was gathered which gives rise to a 2d-histogram plot providing more information.

Although this principal procedure demonstrates the specificity of binding in a sufficient way, an additional experiment with another DNA fragment attached to the cantilever was performed. This fragment is commonly used in control experiments to ensure that the observed bindings depend only on the sequence of the DNA. This EBNA (Epstein-Barr virus nuclear antigen) DNA has no ExpR-binding sequence and thus, there were not any specific interactions with the protein even in the presence of 10 µM $C_{16:1}$-HL. In 834 force distance curves taken, only 5 rupture events were detected (e).
These experiments, shown in Fig 6.1 (a-d), were all performed with the same cantilever, and the same protein sample as used for the EBNA-DNA experiment (e).

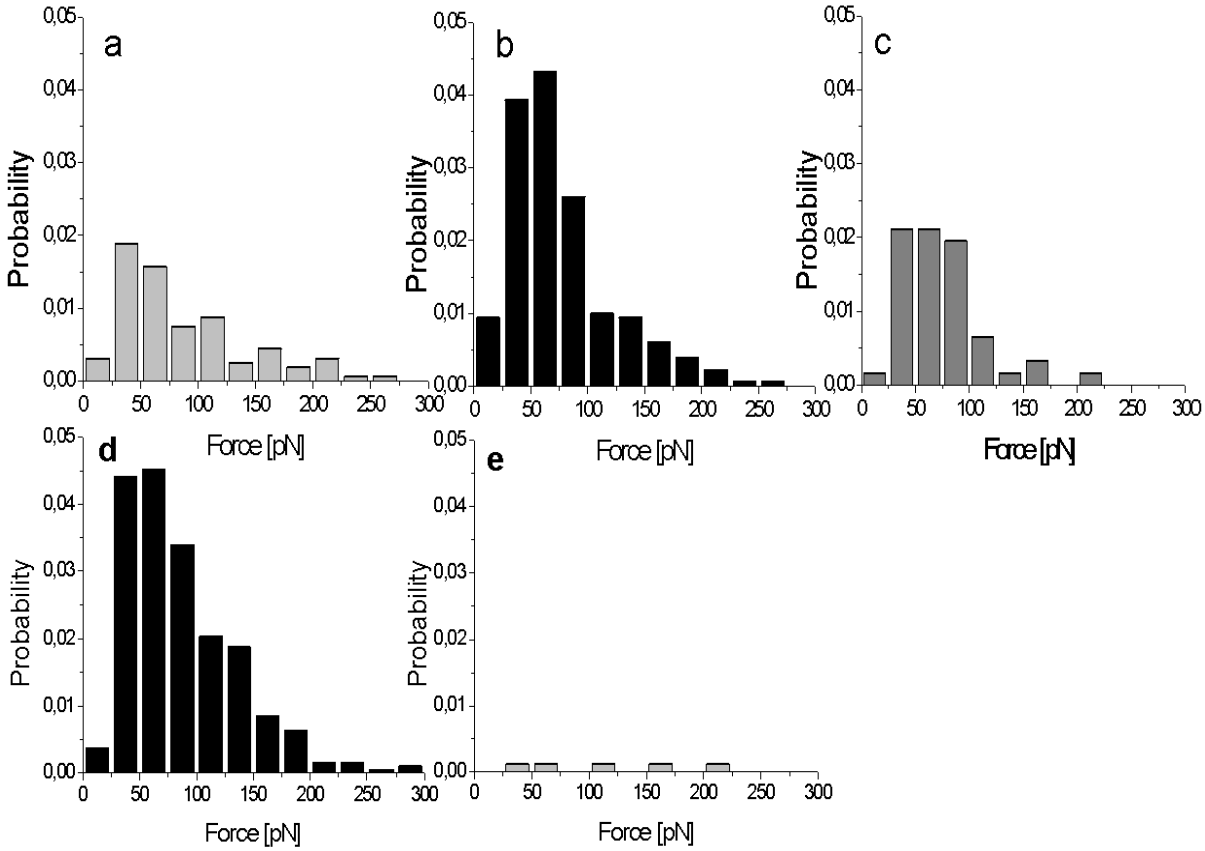

Figure 6.1: Force spectroscopy control experiments.
(**a**) Distribution of the dissociation forces of DNA-$(His)_6$ExpR complexes in the absence of any AHL. (**b**) After adding 10 µM $C_{16:1}$-HL. (**c**) Competition with free DNA binding fragment (10 ng/µl). (**d**) After washing with binding buffer solution. (**e**) Additional control experiment with EBNA-DNA immobilized on the AFM-tip and $C_{16:1}$-HL activated $(His)_6$ExpR.

6.1.2 DFS results for different methods of analysis

The novel analysis software yields much more precise and reliable data than its forerunner. Beside these improvements it was possible to automate the data analysis and the decision of which rupture curve is taken for the further analysis does not depend on the individual experiences of the user anymore but rather on parameters that can be set equal for a complete experimental series. As described in chapter 5 all parameters and the whole program settings are saved, enabling the user not only to repeat the data analysis but furthermore to exclude

artefacts and other failures[7]. Although the inconsistencies of the standard theory of Evans and Ritchie were shown in Chapter 3 in the first two sets of rows in Table 6.1 the results of this theoretical analysis are shown. The experimental data for both methods of analysis was taken by F. Bartels. The data analysis with the standard analysis (cf. Table 6.1)[8] was done by F. Bartels with the former analysis software "by hand" while the results plotted in the second set of rows ("automated analysis – standard theory") was performed with the novel software. Strictly speaking these two different methods can hardly be compared with each other since one of them depends only on apparent selection criteria while the other one provides in principle objective selection criteria. Due to too much noise it was not possible to analyse the data for the effectors C_{12}-HL and oxo-C_{14}-Hl with the new software.
The main problem for the further data analysis according to the theory of bond heterogeneity is that the number of ruptures[9] is quite at a critical level for sufficient statistical analysis. Only the control experiment of $C_{16:1}$-HL (indicated with *), contributed by the author, provided enough rupture curves to meet the demands of the new analysis method. Thus the statistical uncertainties are quite high for the other values. These data are plotted in the last set of rows ("automated analysis – theory of bond heterogeneity"). For reasons mentioned in chapter 3.2.4 the statistical uncertainties for the off-rates are in a logarithmical expression.

	standard analysis		**automated analysis**		**automated analysis**			
AHL	**standard theory**		**standard theory**		**theory of heterogeneity of bonds**			
	x_β [Å]	k^0_{off} [s^{-1}]	x_β [Å]	k^0_{off} [s^{-1}]	x_β [Å]	σ [Å]	k^0_{off} [s^{-1}]	$\Delta\ln(k^0_{off})$
C_8-HL	5.7 ± 0.3	0.47 ± 0.15	5,0 ± 0,8	1,76 ± 0,90	3.7 ± 0.6	1.5 ± 0.3	3.2×10^{-1}	0.95
C_{10}-HL	5.2 ± 0.3	1.43 ± 0.45	4,6 ± 0,6	1,05 ± 1.75	6.0 ± 1.0	2.3 ± 0.4	2.1×10^{-2}	1.56
C_{12}-HL	2.9 ± 0.2	5.40 ± 1.03	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
oxo-C_{14}-HL	3.5 ± 0.3	3.48 ± 0.62	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
$C_{16:1}$-HL	3.9 ± 0.6	2.19 ± 1.88	5.2 ± 1.0	0.16 ± 0.13	4.8 ± 1.5	2.2 ± 0.7	8.3×10^{-2}	2.62
$C_{16:1}$-HL*			4.5 ± 0.6	0.47 ± 0.22	3.4 ± 0.5	1.6 ± 0.3	4.2×10^{-1}	1.02
C_{18}-HL	4.7 ± 0.8	1.32 ± 1.27	2.5 ± 1.6	1.20 ± 8.55	1.2 ± 0.3	0.6 ± 0.2	1.3×10^{1}	0.47

Table 6.1: Comparison of the results of the different analysis methods.

The first thing one notices by observing the values for the different methods is that the values of the heterogeneity of bonds theory (especially the off-rates) vary more than the others when changing the effector. This effect was also observed with other data sets (not presented) and indicates the improvement in the evaluation of the off-rates. Although all estimated values for the different effectors vary quite a lot for the different analysis methods these variations lie

[7] Of course, only detected ruptures are saved. By decreasing the tolerances for the detection of ruptures one can increase the total number of found ruptures until the highest possible value.
[8] The respective values had to be recalculated because of a little mistake happened in the calculation of the off-rates.
[9] The data presented before was only the one gathered in *Part 1* of data analysis. Now, after setting different f_{min} and reconciliation with the master curve many rupture events are not accepted.

within the statistical uncertainties (except the data for C_{18}-HL) which are quite high due to the lack of recorded experimental data.

Nevertheless, the plots of Figure 6.2 (and Fig. 3.5) demonstrate the advantages of the new analysis method. For C_{10}-HL the histograms (only ruptures along master curve) are plotted in Fig. 6.2. The solid red lines show the densities of the rupture events predicted by the heterogeneous bond model for the estimated parameters while the dotted yellow ones are the densities predicted by the standard theory (x_β and k_{off} from Tabl. 6.1) divided by factor 3 for better visibility of the other curves. While f_{min} for the standard analysis equals zero, the values for the new analysis method increase with the pulling velocity, varying between 32 to 45 pN. Therefore the solid lines do not start at zero and thus, strictly speaking, the dotted lines do not correspond to the plotted histograms. Nevertheless, the improvement of the new analysis method seems to be outstanding (only the values for 200 nm/s suffer of too few ruptures). Please also compare with Fig. 3.5 where the survival probability is plotted against force.

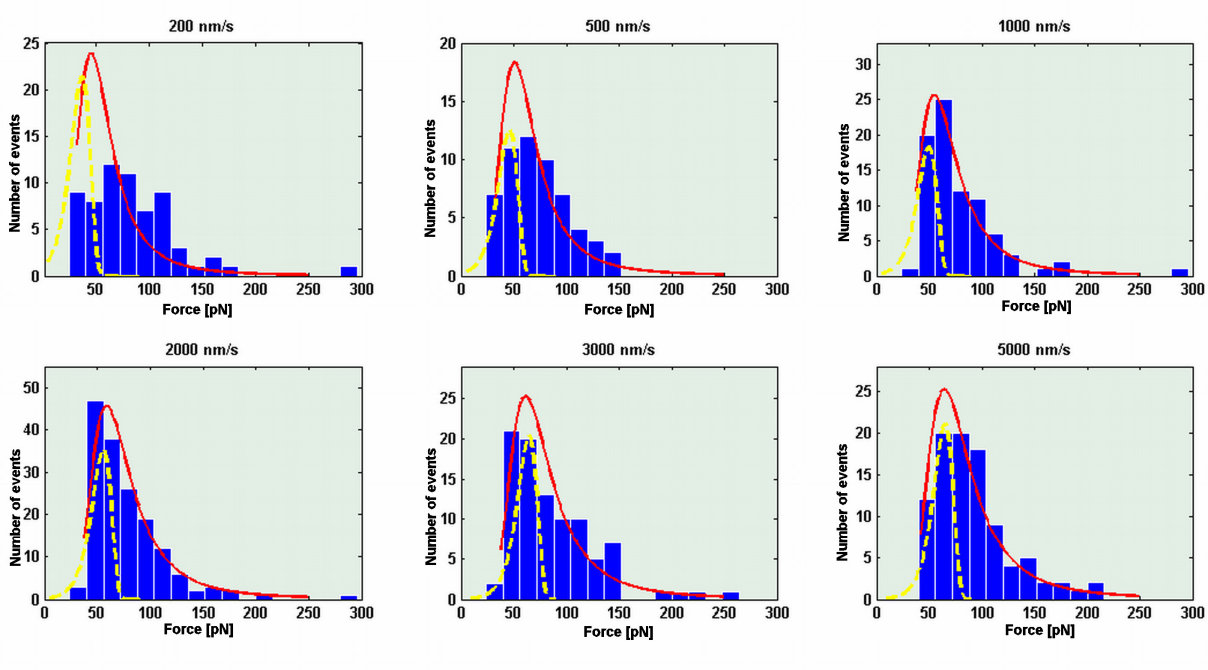

Figure 6.2: Force distributions for ExpR-DNA in presence of C_{10}-HL
The solid red lines are the estimated distributions for theory of bond heterogeneity while the dotted yellow lines are the calculated distributions (rescaled) of the standard theory. (cf. Fig. 3.4)

6.2 AtGRP-RNA

SMFS on DNA-protein interactions have been successfully performed and these interactions show a high specificity of binding like the ExpR-DNA complex. Protein-RNA interactions are very interesting on their own, this work also presents the first DFS experiment successfully yielding also quantitative information about these interactions. Compared with DNA, the RNA protein interactions are much more difficult to investigate for mainly two reasons.
The first one is that the RNA is much more sensitive to pollution, especially RNAse. The second reason lies within the nature of RNA-protein interactions. Although biologists can show that the binding complexes can be quite stable and also specific (the binding disappears completely by introducing point mutations into the RNA as well as the protein), the physical mechanisms of these bindings remain quite unknown. As a matter of course, this work will

not discover the chemical interactions that form the basis of this binding but will reveal a closer look at the different aspects of RNA-protein complex formation that in turn gives rise for further investigations. Although the term unspecific vs. specific is used quite often, the differences of these cannot be clearly specified.
Due to the charges of RNA, "electrostatic interactions" make the distinction to specific binding quite difficult (since the detailed processes are not yet known, the differentiation between specific and unspecific binding is not only difficult in the physical context but also difficult from a linguistic point of view).
One important factor for the protein-RNA binding is the secondary structure of the RNA. To ensure a stable confirmation and also to minimize the (unknown) electrostatic interactions, the used RNA oligos are quite short (27 bases) which means that the only possible confirmation should be a simple loop (or none, of course) as predicted to prevail in the wild type RNA. The used RNA is listed in Table 4.1.

6.2.1 Specificity of binding I – unspecific vs. specific binding

In all experiments the concentrations of RNA and protein were kept constant. Fig. 6.3 shows 2d-histogram plots for different systems.
In the first part (*Part 1*) of data analysis ruptures are identified and characterized automatically by the novel software, this is the so called number of events. In the second part (*Part 2*), the plotting of these 2d-histograms, only rupture curves having parameters in a well defined range are accepted, e.g. f_{min} or stiffness. Comparing this value with the number of specific events yields a parameter that is (nearly) independent of concentration fluctuations in the different experiments. Normalizing the height colour coding to this probability of specific events (same code for the experimental series) thus enables to compare different biological systems as done in Fig. 6.3. These plots can also be compared with the EMSA plots kindly provided by J. Schöning (Fig. 6.4). Please note that these EMSA experiments have been performed with AtGRP7. A set for AtGRP8 remains to be done, but since the binding mechanisms for both proteins are pretty similar, the experiments for AtGRP8 should yield comparable results.

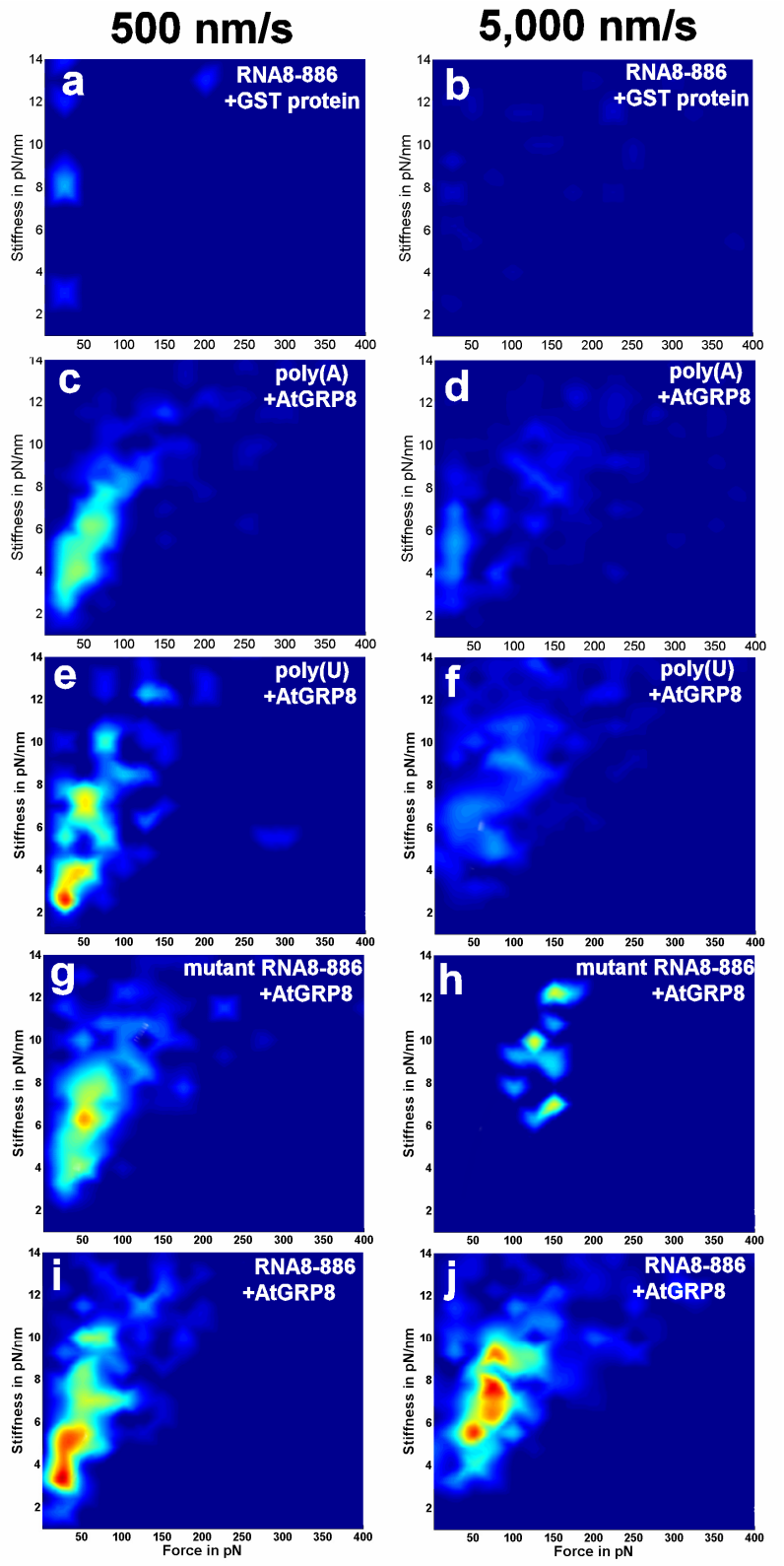

Figure 6.3: Normalized 2d-histograms of different binding partners at 500 nm/s and 5 µm/s

The AtGRPs are attached to a GST protein that was not only needed for purification but also enabled the binding of AtGRP via the GST protein and linkers to the surface[10]. With EMSA

[10] Although the GST protein does not occur in the labelling, all tested AtGRPs were attached to this GST protein (in strict sense it always should have been labelled "GST-AtGRP fusion protein")

experiments it was demonstrated that the GST protein without AtGRP does not interact with RNA (cf. Fig. 6.4).
The first pair of plots (on the left column the pulling velocity is always 500 nm/s, on the right 5,000 nm/s) shows the RNA8-886 ("wild type" RNA of AtGRP8) immobilized at the tip (the principal setup remains untouched during these experiments) and the GST protein immobilized on the sample surface. At both pulling speeds (nearly) no interaction can be observed. The fact that even the wild type RNA does not show any measurable interaction with this protein, while even poly(A) and poly(U) do with the RNA binding protein (see below), gives rise to the assumption that only some proteins, mainly RNA binding proteins, have net charges that interfere with RNA.
In the second row of Fig. 6.3 (c-d) the set for a poly(A) strand and AtGRP8 is plotted. While one can see some fairly broad distributed rupture events at a speed of 500 nm/s there are only 64 at 5,000 nm/s accepted for the plot although 738 events were found in the first part (*Part 1*) of the data analysis. Due to this strange behaviour this kind of binding is definitely related to the category of unspecific interactions.
This peculiarity may give some information about the shape of the energy barrier for this kind of binding but was not investigated further in this work.
In the next row of Fig. 6.3 the unbinding characteristics of poly(U) can be seen. The plot for 500 nm/s indicates 2 spots, but these events show a relatively broad distribution. At 5,000 nm/s the rupture events are also broadly distributed over the whole area between 20 and 200 pN. The poly(U) investigated under EMSA does show a slight binding but not rudimentary sharp peaks as for the wild type, Fig. 6.4 (b).
The mutant RNA8-886, Fig. 6.3 (g-h), shows one single peak at 500 nm/s but a strange behaviour at higher speeds. The mutant RNA will be under further investigation in the following paragraphs. The "wild type" RNA8-886 is the only one showing not only "specific interactions" at low pulling velocities but also at 5 µm/s (Fig. 6.3 (i-j), please note that all plots have been normalized).

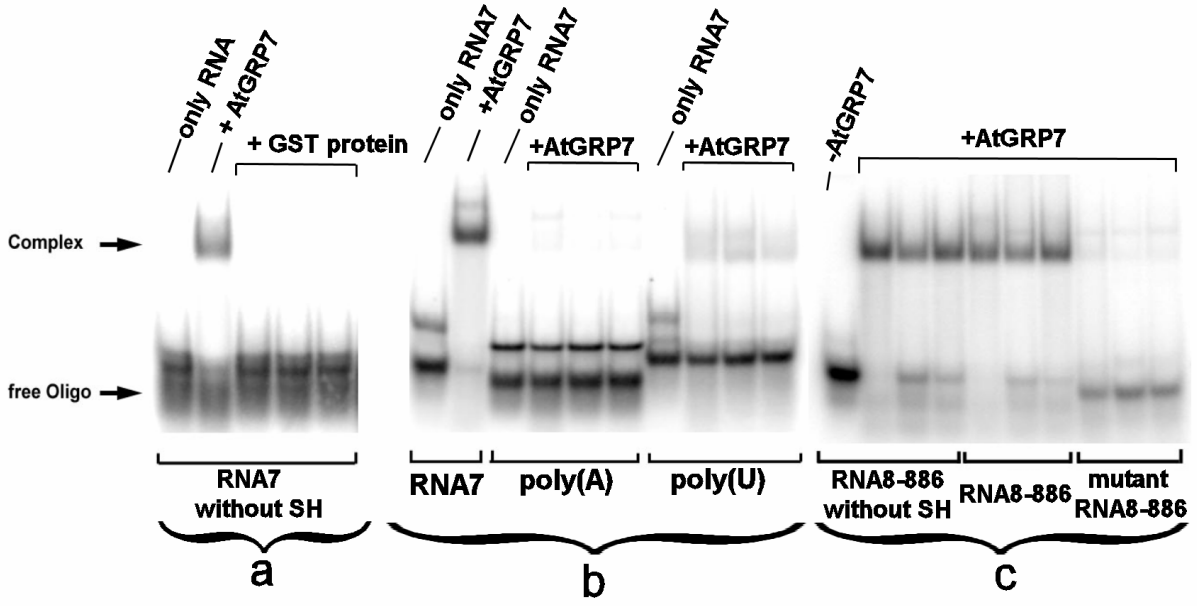

Figure 6.4: EMSA experiments (a-c) with different binding partners (provided by J. Schöning)

For retract velocity 1,776 nm/s a separation of ruptures by dwell time (the time the cantilever spent in contact to the surface) could be done. While taking all rupture events results in two peaks, plotting the data for dwell time lower than 93 ms shows a weaker second spot (Fig. 6.5). Plotting only the data with a dwell time higher than 93 ms gives rise to only one spot,

leaving only some interactions in the area of the weaker spot. This data was recorded with the same tip and sample. This dwell time dependent binding characteristic indicates a different binding affinity for these two binding modes.

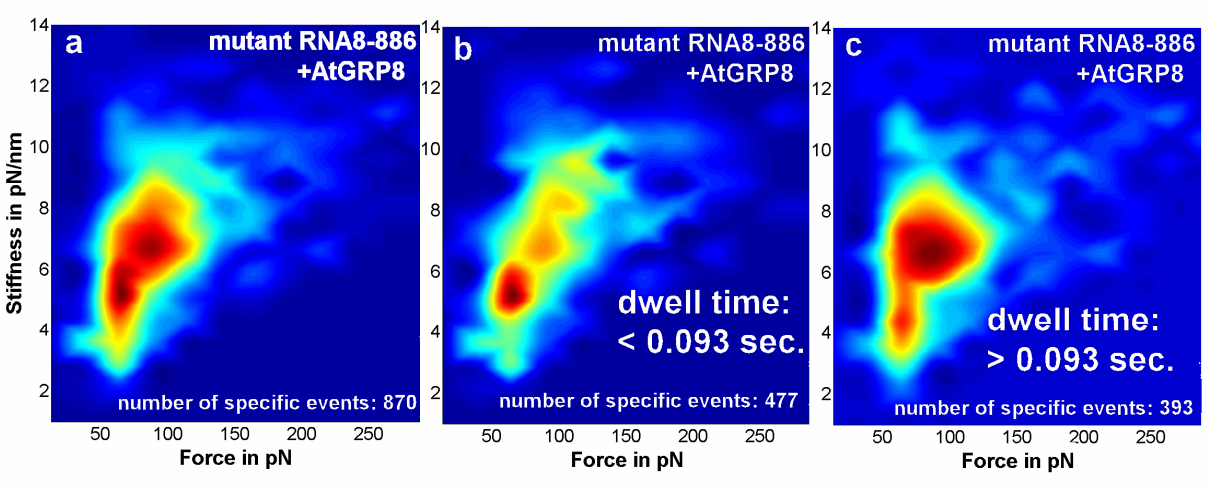

Figure 6.5: Effects of different dwell times on the binding of mutant RNA8-886 AtGRP8

The next plot (Fig. 6.6) shows the data of a complete DFS experiment analyzing the (un-)binding characteristics of mutant RNA8-938 and AtGRP8 with pulling velocities varying from 200 to 6,000 nm/s. All plots have been normalized to the specificity of binding. While at low pulling velocities of 200 and 500 nm/s at least one distinct peak can be obtained, the distribution of rupture events gets broader with increasing velocity ending in a (relatively) equal distribution at 6 µm/s. Maybe further, theoretical, investigations can provide some assumptions for the shape of this kind of energy barrier causing these effects.

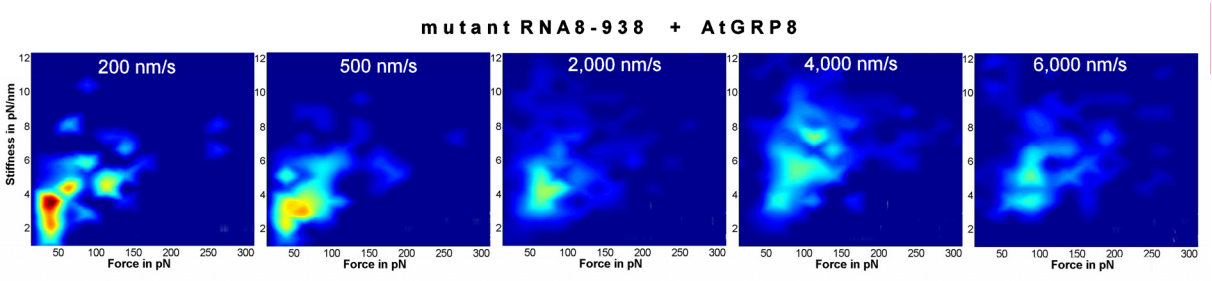

Figure 6.6: (Un-)binding characteristics of mutant RNA8-938 AtGRP8 at different pulling speeds

By introducing the 2d-histogram plot the broad distribution of the standard histograms (cf. Fig. 6.7, 6.11) could be addressed to unspecific events (indicated by the (relatively) equal distribution of the ruptures on the area of the 2d-histogram plots) and specific events (indicated by single or even multiple peaks). The effects of dwell time in correlation with pulling velocity indicate an unusual shape of the energy barrier of the mutant RNA and AtGRP binding.
Nevertheless, the classification and explanation of the protein-RNA interactions in unspecific and specific binding will be a nontrivial task for further experiments and theoretical analysis.

6.2.2 Specificity of binding II – competition experiments

For all different biological systems investigated several competition experiments were performed. It was discovered very early that a competition with free proteins does not make sense because of the binding affinity of the free proteins to free BIS-linker-ends. Consequently the rupture probability increases drastically when adding free protein but does not vanish after the washing with binding buffer. Thus only RNA oligos having same strand length and sequence, but without SH group, were used for the competition experiments.

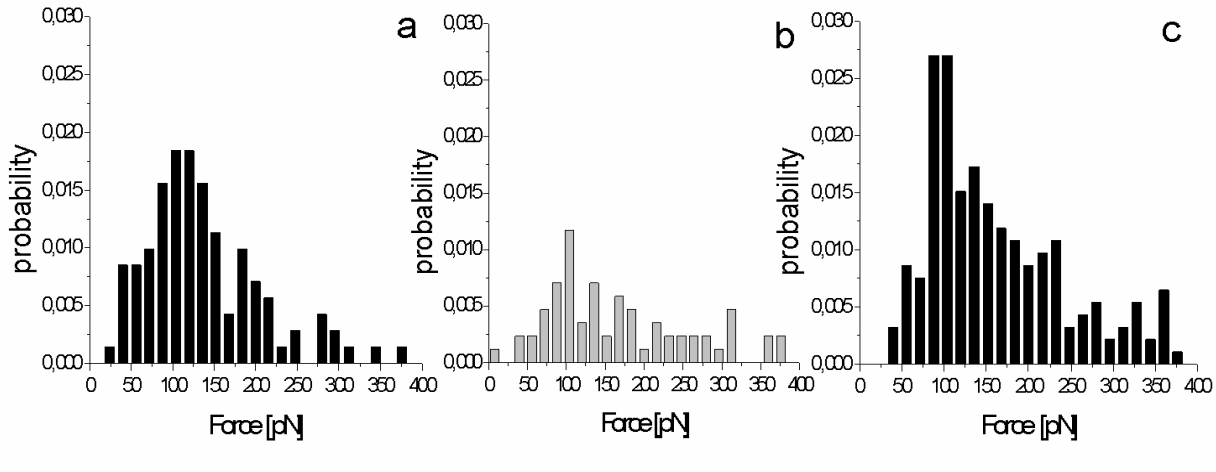

Figure 6.7: Competition presented in standard histogram plot
(a) Force distribution of RNA8-938 AtGRP7 at 2 µm/s pulling velocity. (b) After adding free RNA8-938 (without –SH) the rupture probability decreases. (c) Washing with binding buffer shows (even more) interaction as seen at the initial state.

Fig. 6.7 shows one of these competition experiments in a standard histogram plot. Adding free RNA reduces the rupture probability drastically while the binding could be restored after washing with binding buffer. Although the amount of added RNA was with 75 µM quite high not all binding could be eliminated. The advantages of a 2d-histogram of this data (RNA8-938 AtGRP7) can be easily recognized in Fig. 6.8. The plot shows for this specific complex a relatively unusual appearance which was not under further investigation since a RNA sequence having a higher binding affinity was found (see below). These measurements have been done before the advantages of the 2d-histogram plot had been discovered, being the reason why the number of force distance curves taken is relatively low (but high enough for common histogram plots) which increases the risk of flawed fits due to statistical fluctuations. While the total binding probability only halves, the distribution of rupture events gets broader and no peaks can be obtained. After washing with binding buffer the rupture probability increases over the initial point (A).

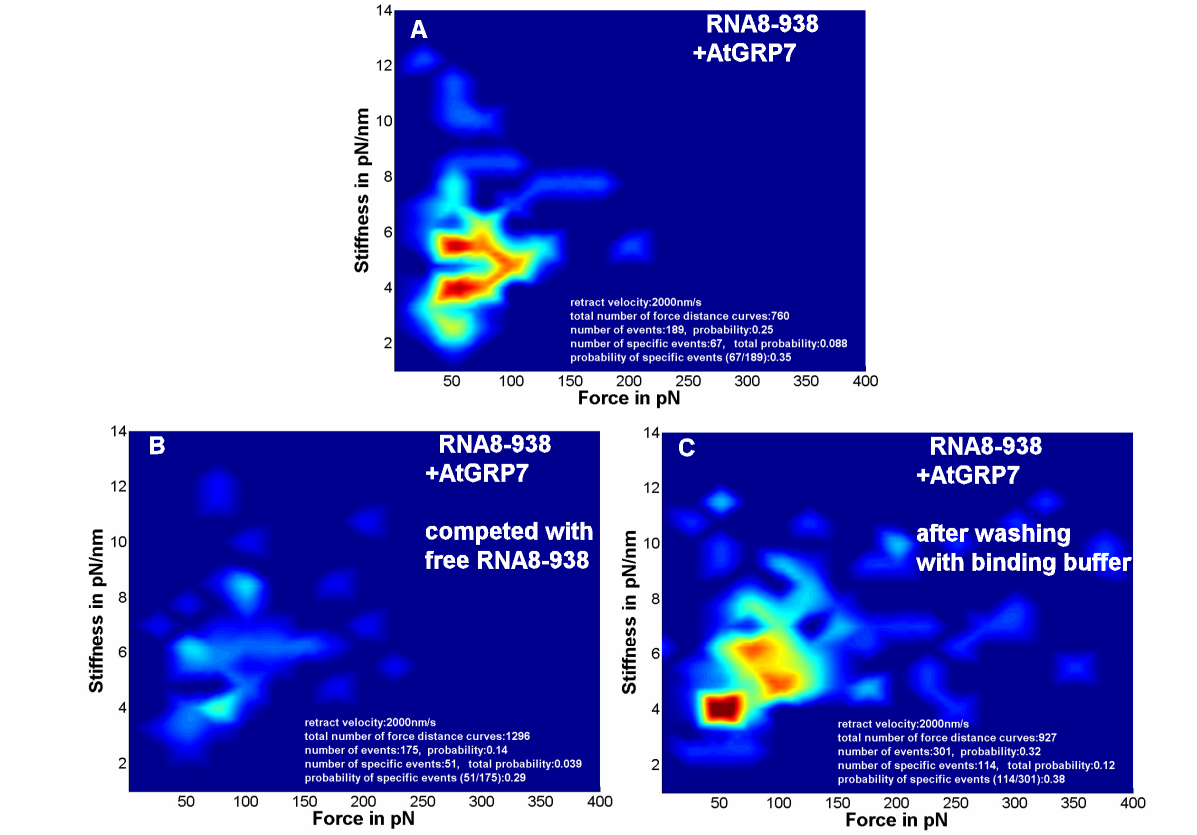

Figure 6.8: Competition experiment with "wild type" RNA8-938 and AtGRP7
(A) Initial state at 2,000 nm/s pulling velocity (B) After adding free RNA8-938 (without SH). (C) After washing with binding buffer.
Same data as for Fig. 6.7. The experiment is suffering from statistical fluctuations because of too few force distance curves taken. The increment of the total rupture probability (C) compared to (A) is remarkable but often observed in competition experiments.

The next competition experiment presented in Fig. 6.9 shows that mutant RNA8-938 binding to AtGRP8 can also be competed with free RNA8-938. This means that (at least) the mutant RNA has the same binding side at the protein as wild type RNA.

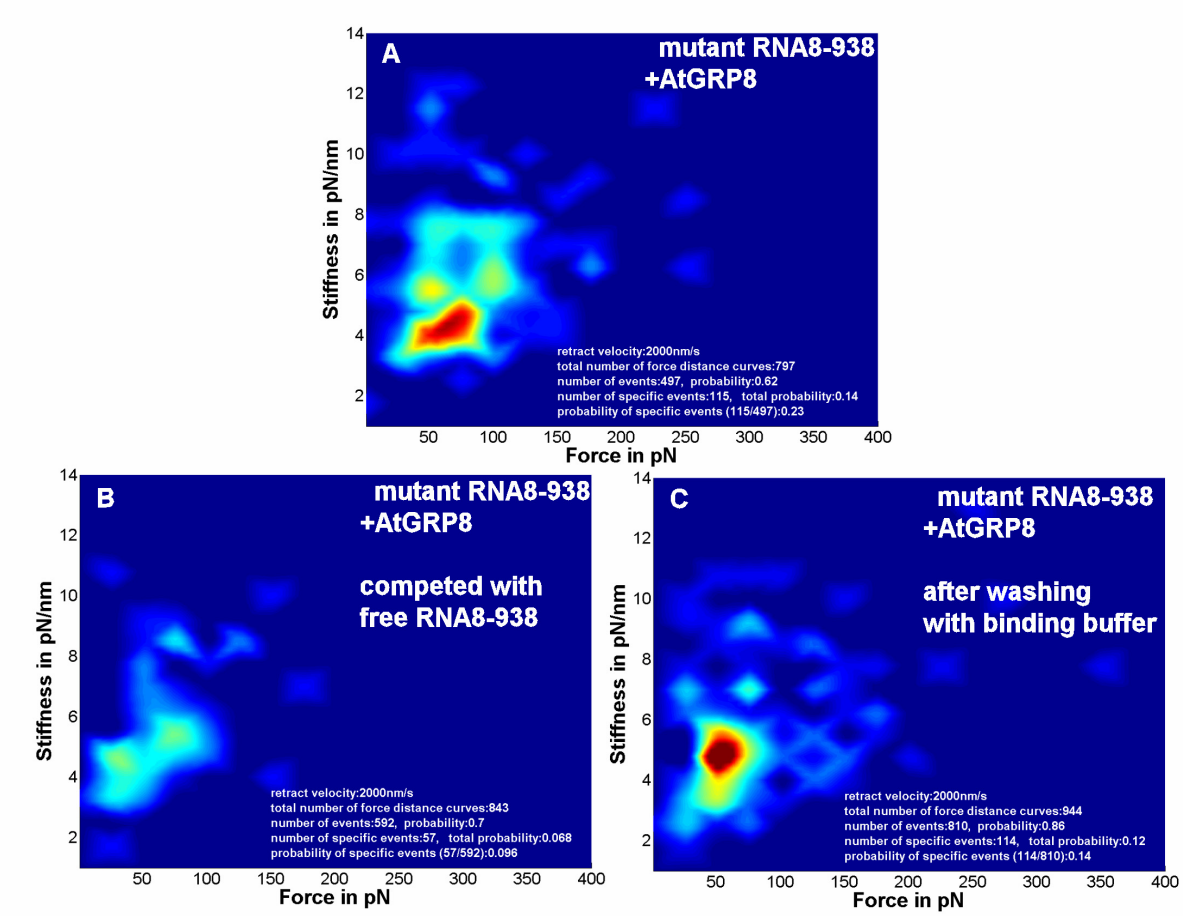

Figure 6.9: Competition experiment with mutant RNA8-938 and AtGRP8 with free RNA8-938 (without SH)

The last competition experiment presented (Fig. 6.10) is the most interesting one, not only because it is the first one providing enough data for sufficient statistics. While the amount of free RNA added was relatively high in the experiments presented before, here only ~1/10 of the usual amount of free RNA8-886 was added to the system of RNA8-886 and AtGRP8. At least the chosen pulling velocity of 500 nm/s maybe chosen too low since the peaks at 5,000 nm/s seem to be much further separeted. Nevertheless, after adding free RNA8-886 the peak indicated by a stiffness of 5.0 pN/nm disappears while the lower one at 3.6 pN/nm seems to be untouched (the slight up movement of this peak may be caused by the fitting routine in the absence of the second peak. The binding probability decreases from 0.17 to 0.13 but the probability of specific events increases from 0.53 to 0.61 which asks for further investigations. Although these two values for the washed system could be restored nearly exactly (while the rupture probability at the competition experiments increased always after washing) the restored peak now seems to be more dominant.

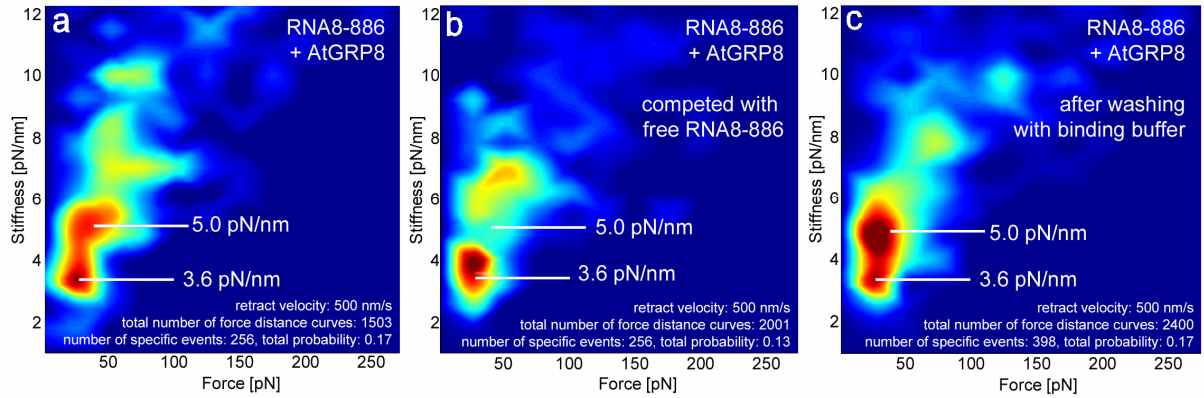

Figure 6.10: Competition experiment of RNA8-886 AtGRP8

Also for competition experiments the new kind of plot demonstrates its potential. While in standard histogram plots only an overall reduction of the rupture probability is observed, these plots provide information about these binding modes. With the discovery of these binding modes one upcoming question was whether these binding modes are caused by different binding sites or a second energy barrier. In the last presented competition experiment the amount of free RNA was quite low and blocked only one binding mode, leaving the other one untouched. This result suggests a second binding site of the RNA AtGRP complex. It can be assumed that the one that is blocked is the one that has a higher binding affinity and thus a lower off-rate. Another plausible explanation is that the probability of having only completely functional proteins is not very high and that degenerated proteins do not have such a high binding affinity. It also could be demonstrated that the mutant RNA shares (at least) one of these binding sites with its respective wild type. Thus it is possible that the peak at 3.6 pN/nm belongs to a weaker and less specific interaction that does not include a positioning of the RNA into the binding site of the AtGRP.

6.2.3 Quantitative Results

There were 3 different RNAs under investigation that interact strongly with AtGRP7 and AtGRP8. RNA7 and RNA8-938 were thought to be the section responsible for the binding at their respective proteins but than it could be shown by means of radioactive EMSA that RNA8-886 binds much better. The binding characteristics of this RNA8-886 to AtGRP8 will be discussed first and then, not only for the sake of completeness, the results for RNA7 and RNA8-938.

The RNA8-886 AtGRP8 complex is with as much as ~ 94,000 force distance curves the best investigated in this work. This amount of data allows a very detailed analysis of this binding complex by separating the data with respect to the dwell time into 4 classes. There have been several separations of different dwell times leading to similar results. Although other dwell-times than the presented showed clearer results concerning the vanishing of the multiple peaks (and leaving only one) the presented fragmentation was chosen in an impartial way having an equal distribution of the number of ruptures (for every division slightly more than 400 (only from 0 to 0.1 seconds dwell time the number is 309)) ensuring highest possible statistical reliability.
To focus on the main point and in order to enable one to compare the 1d-histograms shown with these 2d-histograms the ruptures plotted in Fig. 6.11 are only those who have been accepted to be close enough at the pathway of the master curve. The plots for dwell times less

than 0.1 seconds are not shown since it is (nearly) impossible to realize contact times of the cantilever of less than 100ms in the lower pulling velocity regime. The plots shown in Fig. 6.11 indicate at least 3 different binding modes. Concerning the results of the competition experiments at least two of them cannot be explained by a secondary potential barrier.

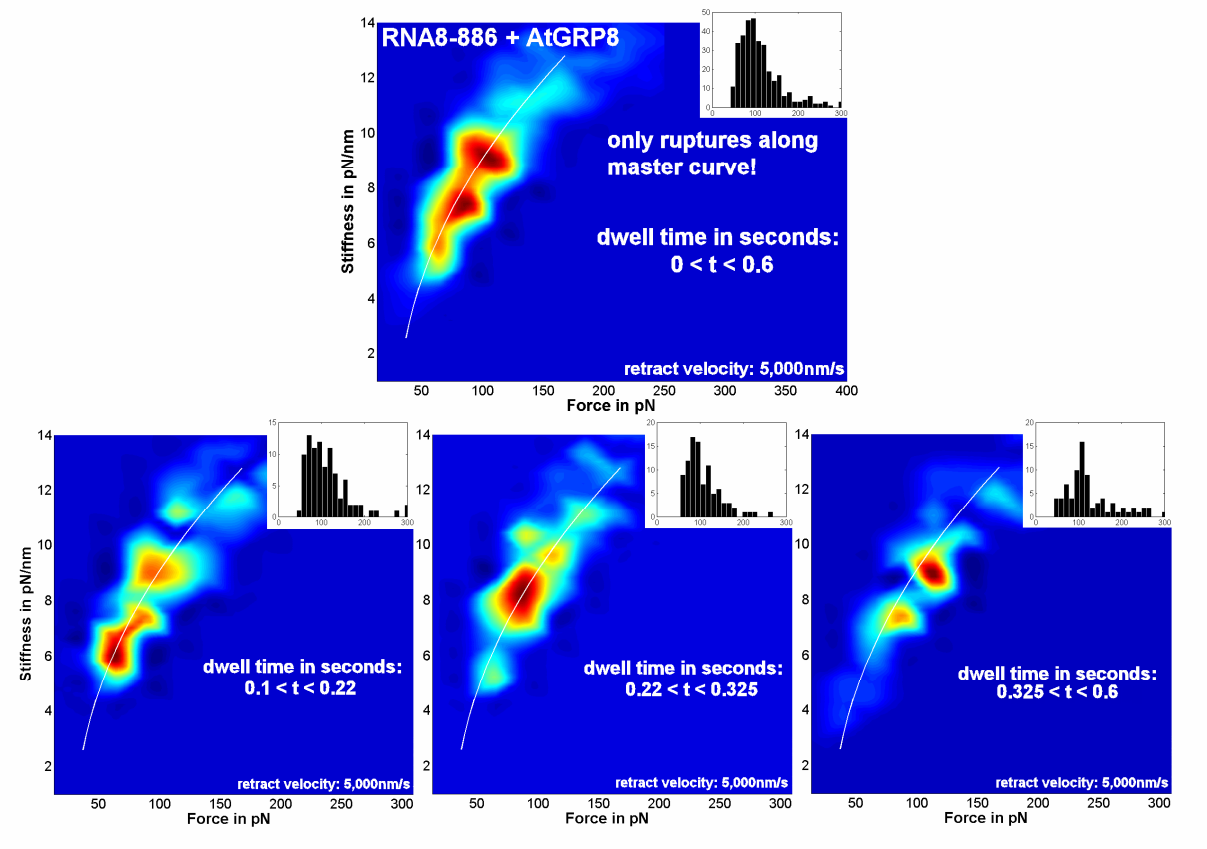

Figure 6.11: Only ruptures, separated after dwell-time, along the master curve are plotted.

While these plots provide a lot of qualitative information it is also possible to do a quantitative dwell-time dependent analysis separated by dwell time. Again, the data for dwell times less than 100 ms is not used. Table 6.2 shows the results for this data analysis.

Oligo	Protein	dwell-time [s]	x_β [Å]	σ [Å]	k^0_{off} [s^{-1}]	$\Delta\ln(k^0_{off})$	mean lifetime [s]
RNA7	AtGRP7	$0.14 \leq t < 2.61$	4.1 ± 0.7	1.8 ± 0.4	8.9×10^{-2}	1.3	11
RNA7	AtGRP8	$0.1 \leq t < 1.85$	4.6 ± 0.9	1.5 ± 0.4	1.5×10^{-1}	1.5	6.5
RNA8-938	AtGRP7	$0.11 \leq t < 4.81$	2.5 ± 0.8	0.8 ± 0.2	4.7×10^{-1}	0.7	2.1
RNA8-886	AtGRP8	$0.1 \leq t < 0.6$	5.7 ± 0.3	1.9 ± 0.2	2.8×10^{-2}	0.6	36
RNA8-886	AtGRP8	$0.1 \leq t < 0.22$	2.8 ± 0.3	1.0 ± 0.2	1.8	6.3	0.5
RNA8-886	AtGRP8	$0.22 \leq t < 0.325$	5.7 ± 0.9	2.1 ± 0.4	1.0×10^{-2}	1.7	92
RNA8-886	AtGRP8	$0.325 \leq t < 0.6$	9.3 ± 2.8	4.3 ± 1.9	3.8×10^{-5}	6.3	2.6×10^{4}

Table 6.2: Results of the analysed RNA – AtGRP combinations

As mentioned in the theory part it is a drawback of the methods of analysis that the most important parameter (for biological usage), the off-rate, is the one that can be calculated with least accuracy. Nevertheless the relative errors between the different results for the dwell

times seem to be low enough to have a clear significance in the differences of the off rates. Since the calculations have been performed for several dwell times showing similar results, statistical fluctuations of the data set provided cannot be neglected but should not be predominant.

The results presented for RNA7 and RNA8-938 are not separated according to dwell-time and the contact times have not been under nearer investigations but the range demonstrates that the rather low off-rates (which are still in the range of the ExpR-DNA complex) cannot be explained by low contact times.
The data presented for the RNA8-886 AtGRP8 complex is much more interesting. The differentiation to different dwell-times, which have been chosen impartially, seems also to allow a quantitative separation of the different binding modes (Tab. 6.2). While the estimated off-rates for the dwell times below 325ms are in the range of 1 to 10^{-2} s^{-1}, the estimated ones for higher dwell times are pretty low with 10^{-5} s^{-1}. This would correspond to a mean life time of some hours which is much more than the results of further studies of similar systems would expect. However, considering the statistical uncertainties, these values can vary even over some magnitudes
The last presented Figure (Fig. 6.12) demonstrates the advantages of the new analysis method very well. Both plots show an impressive consistence of the estimated theoretical values and the experimental data. The plots were done with the not dwell time separated data set of RNA8-886 AtGRP8 providing up to ~600 rupture events at 1,776 and 5,000 nm/s.

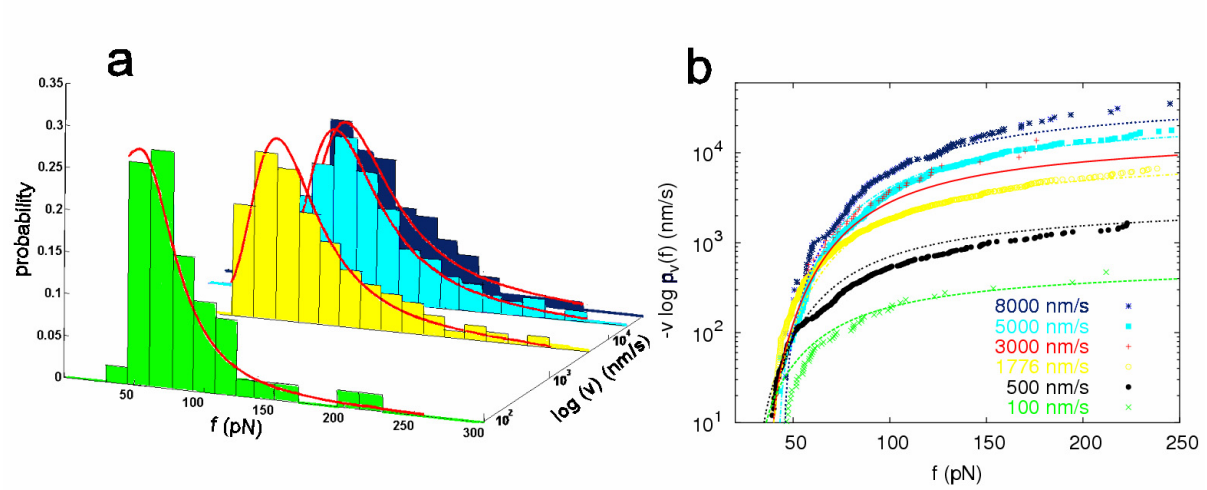

Figure 6.12: Two different kind of plots of the same data (RNA8-886 AtGRP8, not dwell time separated)
(a) Estimated distributions (red line) are the negative derivatives of the survival probability $p_v(f)$ (3.28) to force for each pulling velocity.
(b) The function $-v\log(p_v(f))$ plotted for different pulling velocities (concern that f_{min} (here) depends on retract velocity)

7 Conclusions and Outlook

In this work a new method of data analysis for SMFS experiments was presented and also applied on two different biological systems. The novel software gives rise to an automated analysis of force distance curves and thus will enable future users to analyse much more different SMFS data (at the same time). Apart from this acceleration, the program also yields much more reliable information about the processes involved, not only because all program settings are saved (and the data analysis can thus be re-done easily) but mainly because much more parameters (e.g. rupture length, the values of the fitted polynomials, information about the quality of the force distance curves) are extracted for each curve. Although the advances of assuming a bond heterogeneity were known, it was not possible to analyse the experimental data according to this theory in an adequate way. The novel analysis software enables now the implementation of the theory of bond heterogeneity in an applicable way, which was demonstrated at the ExpR-DNA complex. Certainly, it will be a new standard in future SMFS experiments.
This work also presents the first DFS experiment successfully performed on a RNA-protein complex not only qualitatively but also quantitatively. By comparing this DFS data with the EMSA experiments the stability of the RNA-AtGRP complex formations was demonstrated. The combination of all the advances of the software enables a 2d-histogram plot giving rise to a much higher "resolution" in the low force regime. While the standard force histograms sometimes seem to be only a little broader than usual, this new kind of plot even allows a differentiation of binding modes of the RNA-AtGRP complexes. It was possible to block only one binding mode during a competition experiment while leaving the other one unaltered, indicating different binding modes beside the main binding site, the RNA recognition motif (RRM), of the AtGRP. By varying the dwell time these effects were observed, too, but could also be analysed quantitatively, with estimated mean lifetimes of up to some hours (keeping in mind that these values can vary over some magnitudes). While up to this point only different mutations of RNA have been analysed, a mutant AtGRP may offer detailed information about the reasons for the binding modes and thus the mechanisms of protein-RNA binding.
Although these effects of the RNA-protein interactions are pretty interesting and ask for more experiments the processes occurring in FS experiments of ligand-receptor complexes remain still to be clarified. To uncover these problems some much simpler model systems should be investigated. For such a system it might be feasible to investigate the possible sources of bond heterogeneity step by step as listed in chapter 3.2.3 and address these results to more complex systems like RNA-protein interactions. Here especially the influence of geometrical distortions of the binding molecules on the distribution of the rupture forces could be investigated by reducing the other problems by using such a simple molecular system.
However, since the AFM that will be used for further experiments works on another programming language, all software should be converted to this enabling the user to do the data analysis during the experiment in real time. Then it would be possible to decide immediately whether the gathered data is sufficient for statistical analysis, or whether the amount of free binding partners during a competition is too high or too low, etc.. It should also be possible to implement routines that accelerate the experimental part. If no rupture events occur or the tip is placed over an area contaminated with artefacts (e.g. partial damages of the surface, protein clusters) the sample could be automatically re-positioned until significant rupture events occur and then stay there until a certain number of curves are taken. All these improvements will have impact for the investigation of molecular binding

mechanisms, but this software package will also make FS based screening applications possible. The technique to spot some thousand different DNA mutants on a chip is commonly used. With an AFM mounted on an inverted microscope these spots could be identified and analysed via force spectroscopy at a satisfying rate.

Bibliography

(1) Gonzalez JE, Marketon MM. Quorum sensing in nitrogen-fixing rhizobia. Microbiol Mol Biol Rev 2003;67(574):592.

(2) Swift Sl, Downie JA, Whithead N, Barnard AML, Salmond GPC, Williams P. Quorum sensing as a population density dependent determinant of bacterial physiology. Adv Microb Phys 2001;45:199-270.

(3) Fuqua C, Parsek MR, Greenberg EP. Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. Annu Rev Genet 2001;35:439-68.

(4) Marketon MM, Gonzalez JE. Identification of two quorum-sensing systems in *Sinorhizobium meliloti*. J Bacteriol 2002;184:3466-75.

(5) Marketon MM, Groquist MR, Berhard A, Gonzalez JE. Characterization of the *Sinorhizobium meliloti sinR*/*sinI* locus and the production of novel N-acyl homoserine lactones. J Bacteriol 2002;184:5686-95.

(6) Pellock BJ, Teplitski JM, Boinay RP, Bauer WD, Walker GC. A LuxR homolog controls production of symbiotically active extracellular polysaccharide II by *Sinorhizobium meliloti*. J Bacteriol 2002;184:5067-76.

(7) Hanzelka BL, Greenberg EP. Evidence that the N-terminal region of the *Vibrio fischeri* LuxR protein constitutes an autoinducer-binding domain. J Bacteriol 1995;177:815-7.

(8) Stevens AM, Doland KM, Greenberg EP. Synergistic binding of the *Vibrio fischeri* LuxR transcriptional activator domain and RNA polymerase to the *lux* promoter region. Natl Acad Sci USA 1994;91:12619-23.

(9) Marketon MM, Glenn SA, Eberhard A, Gonzalez JE. Quorum sensing controls exopolysaccharide production in *Sinorhizobium meliloti*. J Bacteriol 2003;185:325-31.

(10) Gao M, Chen H, Eberhard A, Gronquist MR, Robinson JB, Rolfe BG, et al. *sinI*- and *expR*-dependent quorum sensing in *Sinorhizobium meliloti*. J Bacteriol 2005;187:7931-44.

(11) Hoang X, Becker A, Gonzalez JE. The LuxR Homolog ExpR, in Combination with the Sin Quorum Sensing System, Plays a Central Role in *Sinorhizobium meliloti* Gene Expression. J Bacteriol 2004;186:5460-72.

(12) Marchand V, Mougihng A, Mereau A, Branlant C. Study of RNA-Protein Interactions and RNA Structure in Ribonuleoproteoin Particles. In: Hartmann RK, Bindereif A,

Schön A, Westhof E, editors. Handbook of RNA Biochemistry. Wiley-VCH; 2005. p. 172-204.

(13) De Marian JJdO. Observation botanique. Histoir de l'Academie Royale des Science 1729;35-9.

(14) Staiger D. Biologische Zeitmessung bei Pflanzen. Biologie in unserer Zeit 2000;76-81.

(15) Schoof H, Zaccaria P, Gundlach H, Lemcke K, Rudd S, Kolesov G, et al. MIPS Arabidopsis thaliana Database (MAtDB): an integrated biological knowledge resource based on the first complete plant genome. Nucleic Acids Res 2002 Jan 1;30(1):91-3.

(16) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 2000 Dec 14;408(6814):796-815.

(17) Kloppstech K. Diurnal and circadian rhythmicity in the expression of light-induced plant nuclear messenger RNAs. Planta 1985;165:502-6.

(18) Millar AJ, Carré IA, Strayer CA, Chua NH, Kay SA. Circadian clock mutants in Arabidopsis identified by luciferase imaging. Science 1995;267:1161-3.

(19) Schaffe R, Ramsay N, Samach A, Corden S, Putterill J, Carre IA, et al. The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. Cell 1998;93:1219-29.

(20) Wang Z-Y, Tobin EM. Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. Cell 1998 Jun 26;93:1207-17.

(21) van Nocker S, Viestra RD. Two cDNAs from Arabidopsis thaliana encode putative RNA binding proteins containing glycine-rich domains. Plant Mol Biol 1993;4:695-9.

(22) Carpenter CD, Kreps JA, Simon AE. Genes Encoding Glycine-Rich Arabidopsis thaliana Proteins with RNA-Binding Motifs Are Influenced by Cold Treatment and an Endogenous Circadian Rhythm. Plant physiology 1994;104:1015-25.

(23) Heintzen C, Nater M, Apel K, Staiger D. AtGRP7, a nuclear RNA-binding protein as a component of a circadian-regulated negative feedback loop in Arabidopsis thaliana. Proc Natl Acad Sci U S A 1997 Jun 2;94:8515-20.

(24) Staiger D, Zecca L, Wieczorek Kirk DA, Apel K, Eckstein L. The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. Plant J 2003;33:361-71.

(25) Binnig G, Rohrer H, Gerber C, Weibel E. Surface studies by scanning tunneling microscopy. Phys Rev Lett 1982;49:57-61.

(26) Binnig G, Rohrer H, Gerber C, Weibel E. Tunneling through a controllable vacuum gap. Appl Phys Lett 1982;40:178-80.

(27) Binnig G, Quate CF, Gerber C. Atomic Force Microscope. Phys Rev Lett 1986;56(9):930-3.

(28) Bustamante C, Macosko JC, Wuite GJ. Grabbing the cat by the tail: manipulating molecules one by one. Nat Rev Mol Cell Biol 2000;1:130-6.

(29) Zlatanova J, Lindsay SM, Leuba SH. Single molecule force spectroscopy in biology using the atomic force microscope. Progress in Biophysics and Molecular Biology 2000;74:37-61.

(30) Lee GU, Chrisey LA, Colton RJ. Direct measurement of the forces between complementary strands of DNA. Science 1994;266:771-3.

(31) Florin E-L, Moy VT, Gaub HE. Adhesion forces between individual ligand-receptor pairs. Science 1994;264:415-7.

(32) Hinterdorfer P, Baumgartner W, Gruber H, Schilcher K, Schindler H. Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. Proc Natl Acad Sci USA 1996;93:3477-81.

(33) Dammer U, Hegner M, Anselmetti D, Wagner P, Dreier M, Huber W, et al. Specific antigen/antibody interactions measured by force microscopy. Biophys J 1996;70:2437-41.

(34) Ros R, Schwesinger F, Anselmetti D, Kubon M, Schäfer R, Plückthun A, et al. Antigen binding forces of individually addressed single-chain Fv antibody molecules. Proc Natl Acad Sci USA 1998;95:7402-5.

(35) Evans E, Leung A, Heinrich V, Zhu C. Mechanical switching and coupling between two dissociation pathways in a P-selectin adhesion bond. Proc Natl Acad Sci USA 2004;101(31):11281-6.

(36) Eckel R, Ros R, Decker B, Mattay J, Anselmetti D. Supramolecular Chemistry at the Single Molecule Level. Angew Chem Int Ed 2005;(44):484-8.

(37) Dammer U, Popescu O, Wagner P, Anselmetti D, Güntherodt H-J, Misevic GN. Binding strength between cell adhesion proteoglycans measured by atomic force microscopy. Science 1995;267:1173-5.

(38) Bartels FW, Baumgarth B, Anselmetti D, Ros R, Becker A. Specific binding of the regulatory protein ExpG to promoter regions of the galactoglucan biosynthesis gene cluster of *Sinorhizobium meliloti*--a combined molecular biology and force spectroscopy investigation. J Struct Biol 2003;143(2):145-52.

(39) Koch SJ, Wang MD. Dynamic force spectroscopy of protein-DNA interactions by unzipping DNA. Phys Rev Lett 2003;91(2):208103-1-208103-4.

(40) Kühner F, Costa LT, Bisch PM, Thalhammer S, Heckl WM, Gaub HE. LexA-DNA bond strength by single molecule force spectroscopy. Biophys J 2004 Oct;87(4):2683-90.

(41) Baumgarth B, Bartels FW, Anselmetti D, Becker A, Ros R. Detailed studies of the binding mechanism of the Sinorhizobium meliloti transcriptional activator ExpG to DNA. Microbiology 2005;(151):259-68.

(42) Eckel R, Wilking SD, Becker A, Sewald N, Ros R, Anselmetti D. Single-molecule experiments in synthetic biology: An approach to the affinity ranking of DNA-binding peptides. Angew Chem Int Ed 2005;44:3921-4.

(43) Neumeister JM, Ducker WA. Lateral, normal, and longitudinal spring constants of atomic force microscopy cantilevers. Rev Sci Instr 1994;64:2527-31.

(44) Cleveland JP, Manne S, Bocek D, Hansma PK. A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. Rev Sci Instr 1993;64(2):403-5.

(45) Gibson CT, Watson GS, Myhra S. Determination of the spring constants of probes for force microscopy/spectroscopy. Nanotechnology 1996 Sep;7(3):259-62.

(46) Hutter JL, Bechhoefer J. Calibration of atomic-force microscope tips. Rev Sci Instr 1993;7:1868-73.

(47) Roters A, Johannsmann D. Distance-dependent noise measurements in scanning force microscopy. J Phys :Condens Matter 1996;8:7561-77.

(48) Walters DA, Cleveland JP, Thomson NH, Hansma PK, Wendman MA, Gurley G, et al. Short cantilevers for atomic force microscopy. Rev Sci Instr 1996 Oct;67(10):3583-90.

(49) Evans E, Ritchie K. Dynamic strength of molecular adhesion bonds. Biophys J 1997;72(4):1541-55.

(50) Bell GI. Models for the specific adhesion of cells to cells. Science 1978;200:618-27.

(51) Kramers HA. Brownian motion in a field of force and the diffusion model of chemical reactions. Physica VII 1940;4:284-304.

(52) Bell GI. Models for the specific adhesion of cells to cells. Science 1978;200(12 May):618-27.

(53) Braun O, Hanke A, Seifert U. Probing molecular free energy landscapes by periodic loading. arXiv 2004;cond-mat:0402496.

(54) Evans E. Probing the relation between force-lifetime-and chemistry in single molecular bonds. Annu Rev Biophys Biomol Struct 2001;30:105-28.

(55) Friedsam C, Wehle AK, Kühner F, Gaub HE. Dynamic single-molecule force spectroscopy: bond rupture analysis with variable spacer length. J Phys :Condens Matter 2003;15:S1709-S1723.

(56) Heymann B, Grubmüller H. Dynamic force spectroscopy of molecular adhesion bonds. Phys Rev Lett 2000;84(26):6126-9.

(57) Merkel R, Nassoy P, Leung A, Ritchie K, Evans E. Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. Nature 1999;397:50-3.

(58) Rief M, Fernandez JM, Gaub HE. Elastically coupled two-level systems as a model for biopolymer extensibility. Phys Rev Lett 1998;81(21):4764-7.

(59) Strunz T, Oroszlan K, Schumakovitch I, Güntherodt H-J, Hegner M. Model ernegy landscapes and the force-induced dissociation of ligand-receptor bonds. Biophys J 2000;79(3):1206-12.

(60) Raible M, Evstigneev M, Bartels FW, Eckel R, Nguyen-Duong M, Merkel R, et al. Theoretical analysis of single-molecule force spectroscopy experiments: Heterogeneity of chemical bonds. Biophysical Journal 2006 Jun;90(11):3851-64.

(61) Evstigneev M, Reimann P. Dynamic force spectroscopy: Optimized data analysis. Phys Rev E 2003;68:045103-1-045103-4.

(62) Raible M, Evstigneev M, Reimann P, Bartels FW, Ros R. Theoretical analysis of dynamic force spectroscopy experiments. J Biotechnol 2004;112:13-23.

(63) Nguyen-Duong M, Koch KW, Merkel R. Surface anchoring reduces the lifetime of single specific bonds. 61 ed. 2003. p. 845-51.

(64) Kuhner F, Costa LT, Bisch PM, Thalhammer S, Heckl WM, Gaub HE. LexA-DNA bond strength by single molecule force spectroscopy. Biophys J 2004 Oct;87(4):2683-90.

(65) Sivia D. Data Analysis - A Bayesian tutorial. Oxford: Clarendon Press; 1996.

(66) Getfert S, Reimann P. Bayesian inference Applied to single -molecule force spectroscopy. to be submitted 2006.

(67) Bartels FW. Kraftspektroskopische Bindungsstudien an einzelnen Protein-DNA-Komplexen. Dissertation, Universität Bielefeld; 2005.

Acknowledgements

Wie bereits in einigen Kapiteln kurz erwähnt entstand diese Arbeit in enger Kooperation mit anderen Arbeitsgruppen. Am Lehrstuhl für molekulare Zellphysiologie bedanke ich mich vor allem bei Jan Schöning für stetige Hilfe in allen biologischen (und auch anderen) Belangen, sowie Martina Lummer für unterhaltsame Schnitzeljagden nach gut versteckter RNA und bei der Lehrstuhlinhaberin Dorothee Staiger.
In der theoretischen Physik möchte ich mich vor allem bei Sebastian Getfert und Peter Reimann bedanken. Sebastian hat mich bei der Entwicklung des 2. Teils der Datenanalyse (*Part 2*) vor allem durch kleine Programmteile unterstützt. Das Programm für die Abschätzung der zu ermittelnden Parameter im letzten Analyseschritt ist dagegen ohne mein Zutun von ihm entwickelt worden.
Am Lehrstuhl für Genetik geht mein Dank an Matthew McIntosh der bereit war für die notwendigen Kontrollexperimente für die ExpR-DNA Bindung nochmals alle notwendigen biologischen Proben zu bereitzustellen, sowie natürlich an Anke Becker.

Der gesamten Arbeitsgruppe auf D1 möchte ich für die angenehme Arbeitsatmosphäre und die stetige Hilfsbereitschaft danken, ohne die das Experimentieren nur schwerlich möglich gewesen wäre. Daniel Wesner danke ich für Rat und Tat bei den nicht immer ganz so erfolgreichen Versuchen DNA abzubilden, Jan Regtmeier für großzüge Unterstützung bei den alltäglichen Problemen des Experimentierens und Volker Walhorn für die schnelle Lösung der immer wieder auftretenden technischen Probleme des AFM. Auch Gabi Kromi soll für die unbürokratische Abwicklung der bürokratischen Hürden gedankt sein. Besonderer Dank gilt Rainer Eckel, der mich mit tatkräftiger Unterstützung jeglicher Art die gesamte Diplomarbeit über betreut hat. Als Geheimtipp für programmiertechnische Fragen (insbesondere natürlich bei matlab) erwähne ich den Diplomandenraum auf M6 (besonders Falk Schubert), in dem sich immer ein hilfsbereiter Informatiker gefunden hat.
Dario Anselmetti möchte ich für die Möglichkeit danken an seinem Lehrstuhl diese Diplomarbeit anfertigen zu dürfen.
Zuletzt möchte ich mich bei Robert Ros für die nicht nur formale sondern vor allem auch engagierte Betreuung dieser Arbeit danken.

Appendix

List of Publications

1. Frank W. Bartels, Matthew McIntosh, Alexander Fuhrmann, Christoph Metzendorf, Patrik Plattner, Norbert Seewald, Dario Anselmetti, Robert Ros and Anke Becker. Effector-stimulated single molecule protein-DNA interactions of a quorum sensing system in Sinorhizobium meliloti. Biophysical Journal (submitted).

Contributions to Conferences

1. Alexander Fuhrmann, Jan Schoening, Rainer Eckel, Sebastian Getfert, Dario Anselmetti, Dorothee Staiger and Robert Ros. Single Molecule Protein-RNA Interactions. *International Conference on Nanoscience and Technology*, July 30 – August 4, 2006, Basel, Switzerland. (Poster)

2. Theoretical Part: Sebastian Getfert, Martin Raible, Mykhaylo Evstigneev and Peter Reimann.
Experimental Part: Alexander Fuhrmann, Rainer Eckel, D. Anselmetti, and Robert Ros. Heterogeneity of Chemical Bonds in Dynamic Force Spectroscopy Experiments. *Theory of single molecule experiments and simulations,* September 26 – 29, 2006, Lyon, France. (Poster)

3. Alexander Fuhrmann, Jan Schoening, Rainer Eckel, Sebastian Getfert, Dario Anselmetti, Dorothee Staiger and Robert Ros. Single Molecule Protein-RNA Interactions. *Scanning Probe Microscopies and Organic Materials XV*, September 27 – 29, Dresden, Germany. (Talk)

4. Jan Schoening, Alexander Fuhrmann, Corinna Streitner, Martina Lummer, Dario Anselmetti, Robert Ros and Dorothee Staiger. The RNA-Binding Protein AtGRP7 – Molecular Characterisation and RNA-Binding Properties. *4th Meeting of the GBM study section 'RNA-Biochemistry', RNA Biochemistry & Workshop microRNAs Kassel*, October 12 – 15, 2006, Kassel, Germany. (Poster)