Diploma Thesis

# The Unfolding Pathway of Ubiquitin in various chemical environments

Florian Dommert

Ludwigs-Maximilians University Munich

Chair for Applied Physics

Prof. Dr. Hermann E. Gaub

July 2007

# Contents

# 1 Introduction

Research in every aspect was affected drastically by the invention of the computer, because of the arising possibilities for fast and precise processing of experimental data. In physics, increasing performance of computers yielded a further possibility to connect theory and experiments. Today many interesting physical quantities not available from experiments are accessible in a computer simulation, modelling the corresponding physical system.

Nowadays the computational power enables us to model and simulate the dynamics of proteins and even small parts of large biological systems, like a cell membrane, on the nanosecond timescale. The existing experimental possibilities to probe the molecular dynamics of a protein on the atomic level suffer from appropriate resolution in space and time. Computer simulations aim at modelling these experiments for a further elucidation of the processes studied in the experiments.

An important aspect investigated in biophysics is the analysis of the protein folding mechanism. The time-scale for the folding of a protein reaches from nanoseconds to seconds or longer, spanning nine orders of magnitude. Here, the limitations of MD simulations arise from the necessary simulation time for a folding process. Only partially folded structures can be derived from an MD simulation [1], so we need another way to improve the understanding of the folding events. Despite the inability to reproduce the folding pathway of the protein, unfolding may be observed to acquire information about different folding states of the protein. Comparison with according experiments

reveals, whether the unfolding simulations describe unfolding in an appropriate way. Hence, it is possible to characterize a current state of the protein. Different kinds of molecular dynamics (MD) simulations [2, 3] provide insight into the unfolding mechanism, examined in experiments. Several possibilities, like for example chemical and thermal denaturation, exist to unfold a protein. In each case the unfolding process is driven via an increase of the Gibb's free energy $G$ in the given examples due to the additon of denaturants, like urea or guanidine hydrochloride, to a water solution or an increased temperature of the system.

Our MD simulations aim at modelling Atomic Force Microscopy (AFM), which provides another experimental possibility to unfold a protein and obtain information about the profile of the energy landscape $G$ of the protein during unfolding. Two kinds of experimental AFM setups are established. Both stress the protein applying a mechanical force, yielding a shift of the Gibb's free energy landscape $G$ towards higher energies for the folded state and hence increase the probability for an unfolding of the protein in a short time. Both methods apply the mechanical force by picking up one end of the protein with a cantilever that acts like a harmonic spring, whereas the other end is bound to a substrate and the applied force is measured through the deflection of the cantilever. The two methods differ in the time course of the applied forces. In the first method, the cantilever is moved with constant velocity in a certain direction, yielding a time-dependent force that is measured in the course of the experiment providing information about the time-dependent $G$ during unfolding. The other method applies a constant force at one end of the protein by a continuous adjustment of the cantilever, which induces a time-independent shift of $G$. This method primarily aims at an analysis of $G$ by the results of the time course of the end to end distance of the protein. However, both methods lack direct information about the molecular mechanisms, corresponding to the changes in force or in the end to end distance. Here, MD simulations provide a possibility to complement the information gained from AFM studies, and

allow a description of the processes on the atomic level.

In our studies we modelled the introduced AFM methods. However expermiments and our simulations differ regarding the application of the pulling potential, that acts at both termini of the protein in contrast to experimental setups, which perform pulling at one end of the protein. Furthermore in case of a moving pulling potential, the velocity of the potential is about nine orders of magnitude faster than in experiments, to obtain unfolding in a timespan feasible for computation. However in simulations and experiments, unfolding of a protein is favoured by the identical physical mechanism; due to the pulling, the Gibb's free energy of the protein shifted upwards resulting in an increase of the probability for a conformational change towards the unfolded state.

Recent experimental studies on the protein ubiquitin examined its unfolding kinetics [4] and monitored the unfolding pathway via the end to end distance of the protein [5]. For a further elucidation of the folding process computational simulations of ubiquitin [1] were performed that describe primary folding events on the atomic level. Refolding of several partially unfolded structures, derived from MD simulations applying a constant force of 500 pN, was reached by lowering the force to 23 or 40 pN. After a fast hydrophobic collapse of the elongated structures two different states were observed, a kinetically trapped state that did not reveal attempts to form native contacts of ubiquitin and an intermediate state that also was described in experiments. This intermediate state exhibited strong fluctuations of the end to end distance of the $C_\alpha$ atoms at the termini of the protein while funneling along a complex pathway towards the minimum of the Gibb's free energy $G$. It turned out that this intermediate state features properties of a molten-globule like state, characterized by continuous changes of the secondary structure conformations and a lack of a stable folding nucleus. The molten-globule character of the intermediate and the strong fluctuations of the end to end distance are also seen in refolding experiments on ubiquitin, denaturated with chemical or thermodynamical methods, and elucidate the high dynamics of the inter-

mediate state.

Motivated by these studies, we performed MD simulations on ubiquitin, to gain information about the influence of different methods favouring unfolding of ubiquitin by changing the Gibb's free energy landscape $G$. To this end, we simulated both AFM methods introduced above in different solvent environments of the protein. The solvents differed in the concentration of the widely used denaturant urea. Two reasons give rise to the choice of urea. First, it is known that urea is a strong denaturant for ubiquitin. An 8 M urea concentration degenerates ubiquitin to a random coil at a pH value of 2 [6]. Second, the mechanism yielding the degeneration of a protein by urea is not completely understood [7, 8], despite its common usage.

In our simulations we have applied pulling at the protein to completely destroy its secondary structure in different urea concentrations to compare the resulting time courses until a complete loss of the secondary structure elements. Our studies aimed at providing additional information about the influence of urea on the protein during an enforced unfolding process.

Many proteins do not exhibit a complete denaturation in the presence of urea leading to residual secondary structure elements. However, complete denaturation in urea together with its small size renders ubiquitin a good candidate for our MD simulations.

The globular protein ubiquitin marks proteins for degradation and hence belongs to a class of proteins necessary for the degeneration cycle of proteins.

In its native environment ubiquitin is composed of repeated single domains, each consisting of 76 amino acids (fig. 1.1). Its tertiary structure is dominated by a core built up of a mixed $\beta$-sheet, consisting of five $\beta$-strands, and an $\alpha$-helix (fig. 1.2). Many residues are of hydrophobic nature and arrange in a hydrophobic core. Such a hydrophobic core is characteristic for globular proteins and a stabilizing factor for the tertiary structure. The cyan dotted spheres in fig. 1.3 provide an idea of their position. Ubiquitin is very stable under physiological conditions; denaturation does

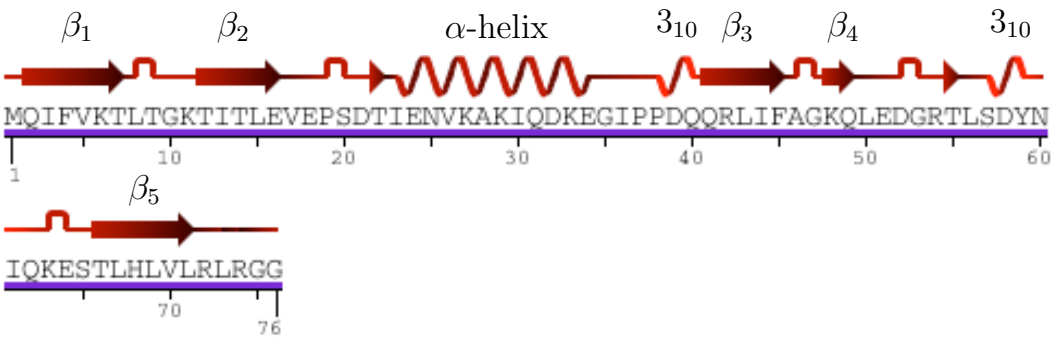



***Figure 1.1:*** Sequence and secondary structure of ubiquitin. The tertiary structure of the protein mainly arises from the arrangement of the strands $\beta_1$ to $\beta_5$, the $\alpha$-helix, and the two $3_{10}$-helices. Figure adapted from [9].

not start until 80 °C and is almost pH-independent [10, 11]. Further structural stability is provided by a dense hydrogen bonding pattern, due to a high amount of secondary structure elements. Additional to the hydrogen bonds connecting the residues of the $\alpha$-helix and the $\beta$-strands, the backbone nitrogen atoms of the residues 23 and 24 form hydrogen bonds with the carbonyl oxygen atoms of the residues 52 and 54. Furthermore hydrogen bonds between the backbone nitrogen atoms of the residues 56 and 57 and the carbonyl oxygen atoms of the residues 18 and 21 strengthen the tertiary structure.

We aim at analyzing the unfolding pathway of ubiquitin in different urea concentrations with MD simulations, which allow direct observation of the unfolding process on the atomic level. Application of different methods to stress the protein by modelled mechanical forces, enables us to compare the unfolding events with the time course of the forces and their influence on the unfolding process. This is expected to gain insight into the Gibb's free energy landscape. Furthermore with the addition of the denaturant urea we want to derive information about its influence on the unfolding process and its interaction mechanism with the solvent and the protein. To this end we will analyse the course of the secondary structure, the energetics of the hydrogen bonds, providing

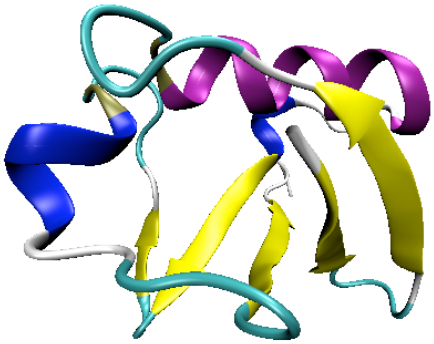


***Figure 1.2:*** Tertiary structure of ubiquitin. $\beta$-strands are yellow-colored, the $\alpha$-helix is purple-colored, $\alpha_{3_{10}}$ helices are shown in blue and loops in cyan.

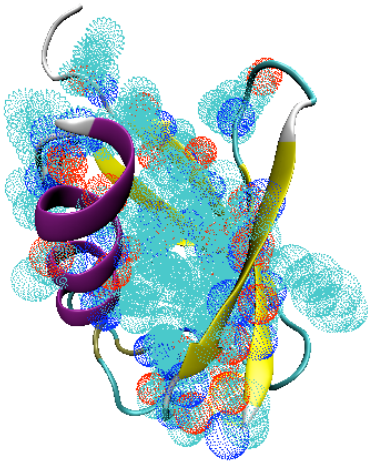


***Figure 1.3:*** The hydrophobic core of ubiquitin. The cyan dotted spheres indicate the hydrophobic surface of the amino acid side chains. Most of them accumulate in the core surrounded by the $\beta$-strands (yellow) and the $\alpha$-helix (purple), stabilizing the tertiary structure of the protein.

information about a change of the enthalpic part of $G$, and the surface of the residues accessible to the solvent, that helps to get an idea about a changing entropy of the system.

# 2 Methods

## 2.1 Molecular Dynamics Simulations

Molecular Dynamics (MD) studies aim at elucidating processes at the atomic level. They enable us to obtain information about many physical quantities that are difficult or impossible to access experimentally. Particularly examinations of biological systems like a solvated protein, can be refined with the help of computer simulations. However, some problems arise: the creation of an accurate model, the sampling of the configurational space of the investigated protein and the the limited computational resources that restrict the size of the simulation system. This chapter describes the basics of MD simulations and our way to adress the problems mentioned above.

One main concern in MD simulations is the reduction of the computational cost. To this end several possibilities exist, which affect the accuracy of the model in different ways. Major difference arises in the model for the interaction of the protein with its environment. Here two different possibilities are available; either the interaction between protein and solvent is included in a model of the protein (implicit solvent) or the interaction between protein and solvent in a simulation is explicitly calculated (explicit solvent). The first method considerablely reduces the computational effort, but the results strongly depend on the protein model. As the interaction of urea with the protein and water is not clear at all, we used an explicit solvent model. Furthermore this model allows a more precise description, but the computational cost is much higher,

due to the explicit treatment of the solvent molecules. Accordingly, the simulation times with explicit solvent are restricted to the nanosecond timescale.

### 2.1.1 Principles

To gain an atomic description of a protein motion, the Schrödinger equation for the system has to be solved:

$$\left(\hat{T}_e + \hat{T}_K + \hat{V}_e + \hat{V}_K + \hat{V}_{eK}\right)\Psi = i \cdot \hbar \frac{\partial}{\partial t}\Psi, \quad \text{with } \Psi = \Psi\left(\vec{x}_1, ..., \vec{x}_{N_e}, \vec{X}_1, ..., \vec{X}_{N_K}\right). \tag{2.1}$$

The vectors $\vec{x_i} \in \mathbb{R}^3$, $(i \in \{1...N_e\})$, and $\vec{X_j} \in \mathbb{R}^3$, $(j \in \{1...N_K\})$ denote the positions of the electrons with mass $m_e$ and nuclei with mass $M_K$, respectively, the operator for the kinetic energies $\hat{T}$ for the electrons is,

$$\hat{T}_e = \frac{-\hbar^2}{2} \sum_{i=1}^{N_e} \frac{1}{m_e} \frac{\partial^2}{\partial \vec{x_i}^2}, \tag{2.2}$$

and for the nuclei

$$\hat{T}_K = \frac{-\hbar^2}{2} \sum_{j=1}^{N_K} \frac{1}{M_K} \frac{\partial^2}{\partial \vec{X}_j^2}. \tag{2.3}$$

The operators $\hat{V}_e = \hat{V}_e\left(x_1, ..., x_{N_e}\right)$ and $\hat{V}_K = \hat{V}_K(X_1, ..., X_{N_K})$ describe the potential energy $V$ in terms of the coordinates of the electrons and nuclei, respectively, arising from their interaction among each other. To include the energy, that results from the interaction between eletrons and nuclei the operator $\hat{V}_{eK} = \hat{V}_{eK}\left(x_1, ..., x_{N_e}, X_1, ..., X_{N_K}\right)$ is present.

For all but the most simple cases the Schrödinger equation cannot be solved analytically, therefore since its introduction in 1927 theoretical physics searches for appropiate approximations [12]. Until M. Born and R. Oppenheimer provided a more rigorous explanation [13], a separability of the wavefunction $\Psi$ in $\chi(X_1, ...X_{N_K})$ and $\varphi(\vec{x}_1, ..., \vec{x}_{N_e})$ was justified with the thermodynamic equipartition of the velocities of the electrons

and nuclei. Due to the strongly differing masses of electrons and nuclei, their velocities scale with $\sqrt{\frac{m_e}{m_K}}$. For this reason the reaction of the electrons, following a movement of the nuclei, is assumed to be instantaneous. Hence, a description of the molecule in its center of mass system with a wavefunction separated in a function of the nucleus coordiantes and a function of the electron coordinates, parametrized by the nucleus coordinates is justified. Finally an orthogonal transformation can be found, yielding the complete separability of the the coordinates of the electrons and nuclei in the wavefunction. It should be mentioned that the calculations of M. Born and R. Oppenheimer do not only explain the separability of the wavefunctions, they firstly predicted the order of the energy spectrum from a diatomic molecule by its vibrational and rotational modes. For our aims the separability of the wavefunction allows us to treat the motion of the considered atoms with the coordinates of their nuclei.

Another approximation is to describe the potential $V$ of our examined system in mathematically simple terms to obtain an easily computable and parametrized expression, the so called force field. The potential energy V and the resulting forces $\vec{F_k}$ mainly depend on the Coulomb, Van der Waals, and intramolecular interactions, that are defined by the bond configuration of the molecule. Due to the long range interaction character of the Coulomb forces, their calculation turns out to be the most time consuming part of the computations.

Potentials are taken into account with the help of different energy terms, depending on the model for the protein-solvent system. Accordingly a variety of possibilities for the implementation of a force-fields exists. In our simulations we used the OPLS-AA force field[14], defined as follows:

$$V_{nonbonded} = \sum_{i<j} \left[ q_i q_j \frac{e^2}{r_{ij}} + 4\epsilon_{ij} \left( \frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}} \right) \right] f_{ij} \tag{2.4}$$

$$f_{ij} = \begin{cases} 0 & \text{pair } ij \text{ connected by a valence bond or a valence bond angle} \\ 0.5 & \text{1,4 interaction (separated by exactly 3 bonds)} \\ 1.0 & \text{otherwise} \end{cases}$$

$$V_{bonded} = \sum_{bonds} K_r \left(r - r_{eq}\right)^2, \tag{2.5}$$

$$V_{angle} = \sum_{angles} K_\Theta \left(\Theta - \Theta_{eq}\right)^2, \tag{2.6}$$

$$V_{torsion} = \frac{1}{2}\sum_i V_{1,i}\left[1 + \cos(\phi_i)\right] + V_{2,i}\left[1 - \cos(2\phi_i)\right] + V_{3,i}\left[1 + \cos(3\phi_i)\right]. \tag{2.7}$$

The nonbonded terms include the electrostatic and the Van der Waals interactions, where $f_{ij}$ defines their strength and rules out the directly bonded pairs. Harmonic potentials approximate the energies, arising from the deviations in the equilibrium distances $r_{eq}$ and equilibrium angles $\Theta_{eq}$ of the bond conformations. The dihedral configuration energy (2.7) depends on the potentials $V_{1,i}, V_{2,i}, V_{3,i}$ of the bonded atoms $i$. In contrast to force-fields, using only a single term for the diheral configuration energy, the three terms allow a more accurate description of the energy.

The third important approximation in MD simulatons is the classical treatment of the molecular motion. Newton's second law is applied to the masses of the nuclei and its numerical integration yields the trajectories of the nuclei in an MD simulation:

$$M_K \cdot \frac{d^2\vec{X}_K}{\mathrm{dt}^2} = -\frac{\partial V(\vec{X}_1, ..., \vec{X}_N)}{\partial \vec{X}_K} \tag{2.8}$$

$$= \vec{F}_K(\vec{X}_1, ..., \vec{X}_N). \tag{2.9}$$

In 1967 L. Verlet[15] published a simple algorithm to solve this set of coupled differential equations by discretizing the time t and expanding up to second order in t:

$$\vec{X}_K(t + \Delta t) = 2\vec{X}_K(t) - \vec{X}_K(t - \Delta t) + \frac{\vec{F}_K\left(\Delta t\right)^2}{m_K} \tag{2.10}$$

The resulting data of a simulation, integrated over n time steps $\Delta$t, describe the trajectories of the molecules,

$$\vec{X}_K(t_i), \qquad t_i = i \cdot \Delta t, \quad i = 1, ..., n, \tag{2.11}$$

providing a possibility to calculate many thermodynamic and mechanic quantities, not accessible to measurements in experiments. Today most MD program packages use a modified version of the Verlet algorithm, but the discretization in time and the expansion up to second order is still the main idea to allow fast and precise computation.

### 2.1.2 Force Probe Molecular Dynamics simulations

To model the effect of the mechanical stress harmonic potentials, modelling a spring with spring constant $k$, are attached to the terminus atoms of the protein (fig. 2.1). They are moved in a specified direction $\hat{n}$ with velocity $v$, hence the potential energy $V_\alpha$ on the respective atoms is modified by including the term:

$$V_{\alpha,\text{eff}}(t) = \frac{1}{2}k \cdot \left(\vec{X}_K(t) - \hat{n} \cdot v \cdot t\right)^2. \tag{2.12}$$


***Figure 2.1:*** Setup of an FPMD simulation: Two springs are pulling the termini of the protein each in opposite directions. This way friction forces arising from a movement of the protein through the solvent are reduced.

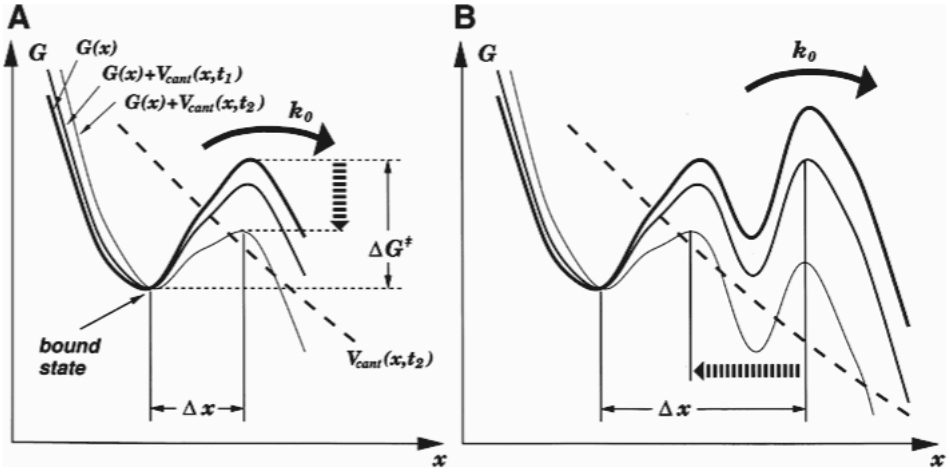



***Figure 2.2:*** A one-dimensional Gibb's free Energy landscape $G$ [16]. Both plots depict the deformation of the Gibb's free Energy $G$ during a pulling simulation. Plot A describes a simple two state unfolding mechanism. The energy landscape in plot B represent a more complicated unfolding process implying the possibility for an intermediate state.

Figure 2.2 shows a simplified one dimensional energy landscape with the reaction coordinate x, a parameter describing the state of the protein. Due to the pulling, the height of the energy barrier between the native and the unfolded state becomes lower and the accessible configuration space of the protein is growing, increasing the probability of an unfolding event.

This kind of simulations corresponds to Atomic Force Microscopy (AFM) experiments but is performed with pulling velocities $v$ nine to ten orders of magnitude larger, due to restrictions on the simulation time. However, the faster pulling velocities cause friction between solvent and the protein that is not negligible in the FPMD simulations unlike the AFM experiments. To this end the modification in the application of the forces was introduced, because experience showed that pulling at both termini in opposite directions reduces friction, due to the absence of a pulling of the protein

through the solvent. Anyway, the results of the simulations and the experiments are not comparable directly and an appropriate description of the effect of friction on the calculated forces is inevitable. A model for a two energy state system [17], describing the protein unfolding attempt frequency $k_0$ in terms of the reciprocal thermal energy $\beta = \frac{1}{k_B T}$ and Kramer's prefactor[18] $\omega_0$, provides an approximation of the attempt frequency for unfolding:

$$k_0 = \omega_0 \cdot \exp(-\beta \Delta G^{\neq}). \tag{2.13}$$

As an unfolding attempt decreases the probability to find the system in the folded state, with the assumption of a linear decrease in the barrier height $\Delta G^{\neq}(t) = \Delta G^{\neq} - kvt\Delta x$, yields a differential equation describing the probability $P(t)$ to find the system in the native state. In the model $\Delta x$ corresponds to the distance between the minimum of $G$ and the following maximum.

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = -P(t) \cdot k_0(t) \tag{2.14}$$

$$= -P(t) \cdot \omega_0 \cdot \exp(-\beta(\Delta G^{\neq} - kvt\Delta x)). \tag{2.15}$$

This approximation holds as long as $\Delta G^{\neq} - k \cdot v \cdot \Delta x \cdot t \gg k_B T$ and back reactions are negligible. Integrating eq. 2.15 yields, regarding the boundary condition $P(t = 0) = 1$, to find the system in the folded state:

$$P(t) = \exp\left[\frac{\omega_0}{\beta kv\Delta x} e^{-\beta \Delta G^{\neq}} \left(1 - e^{\beta kvt_D \Delta x}\right)\right]. \tag{2.16}$$

Substituting the rupture force $F_D = kvt_D$ in eq. 2.16 and differentiating the result in respect to $F_D$ reveals the most probable pulling force $F_{max}$ at the time of an unfolding event, according to $\frac{\mathrm{d}P(F_D)}{\mathrm{d}F_D} = 0$:

$$F_{\max} = \frac{1}{\beta \Delta x} \cdot \ln(\frac{\beta kv\Delta x}{k_0}). \tag{2.17}$$

For an explicit treatment of the friction forces a linear term with a friction constant $\gamma$ is added heuristically:

$$F_{\max} = \gamma \cdot v + \frac{1}{\beta \Delta x} \cdot \ln\left(\frac{\beta kv\Delta x}{k_0}\right). \tag{2.18}$$

Equation (2.18) provides the possibility to extrapolate the computed forces to the regime, where friction is negligible — the situation given in the AFM experiments — hence rendering a direct comparison possible.

### 2.1.3 Force Clamp Molecular Dynamics Simulations

In the previous section was described how to deal with the friction forces and the time-dependent shift of the Gibb's free energy $G$, occuring in FPMD simulations. Both strongly affect the unfolding mechanism. To minimize their influence on the unfolding mechanism of the protein, the idea of the Force Clamp experiments was realized in simulations. Force Clamp experiments are carried out with AFM by pulling the protein with a cantilever applying a constant force ($k_0 = const.$), yielding a constant overall decrease of $\Delta G^{\neq}$. To realize this situation in the Force Clamp Molecular Dynamics (FCMD) simulations, the position of the harmonic potential is calculated and adjusted every integration step to keep the distance between the potential and the considered atoms constant. The advantage arises in the improved comparability of experiment and simulations. Since the constant shift of the Gibb's free energy that favours unfolding is realized in nearly identical ways in simulation and experiment, no special analysis of the friction forces $F_\gamma$ in the FCMD simulations is required. Additionally in this kind of simulations unfolding is not dominated by the mechanical deformation of the tertiary structure, due to the increasing distance of the pulling potentials. Instead, the denaturation mechanism is based on the lowering of $\Delta G^{\neq}$, which perturbs the native energy landscape to a lesser extent. Apart from the differences in the method to apply a force the simulation setup corresponds to the FPMD simulations.

For an FCMD simulation information about the rupture force $F_{rup}$ is required. FCMD simulations with a force much lower than $F_{rup}$ yield unfolding attempt frequencies $k_0$, which are too low to obtain an unfolding event in a computationally feasible time-span. An application of a force too strong can yield unfolding pathways

of the protein, that are artificial, because of the strongly deformed Gibb's free energy landscape.

## 2.2 The Aquaeous Urea Solution

FPMD and FCMD simulations provide us a tool to denature the protein in a time-span feasible for computational studies. Additionally, chemical denaturants like urea (fig. 2.3) reveal a possibility to influence the protein unfolding behavior. Urea is one of the proteolysis end products in mammals, some plants and many fungi [19]. The French chemist Rouelle firstly extracted urea with hot alcohol from evaporated urine in 1773 [20].

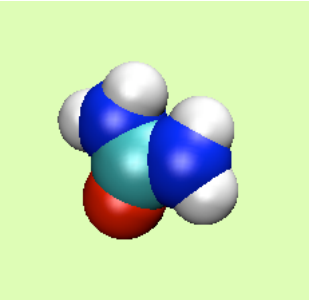


***Figure 2.3:*** The urea molecule $(H_2N)_2\,CO$

Urea is widely used as a denaturant. However, despite its common application it is not clear how urea influences unfolding. On the one hand, there could be direct interaction within urea and the protein [7]. This model suggest, that interaction between the urea molecules and the less polar atoms of the protein backbone provide the denaturating effect of urea. On the other hand unfolding can be affected by altering of the solvent environment, due to interaction of urea with water [8]. Several experiments and MD studies led to two models for the hydrogen bonding dynamics in an aqueous urea solution, according to which the urea concentration strongly effects the action mechanism. One model suggests that the urea molecules fit into the hydrogen bond

network of water. No distortion of the orientational dynamics of the water dynamics occurs and the entropy of the system stays stable. This model is suitable to describe low concentrations of urea. At high urea concentration however, the urea molecules start to interact strongly with each other resulting in a urea aggregation, predicted by the other model. This narrows the water configurational space decisively with the effect of lower solvent entropy due to urea aggregation.

In the native state, the tertiary structure of ubiquitin generates a hydrophobic core, because most of the hydrophobic residues point inwards. Unfolding increases the solvent accessible surface area of these residues and narrows the configurational space of the water molecules, which lowers the entropy of the solvent. If the stability of ubiquitin is strongly dependent on the solvent entropy, high urea concentrations should destabilize the protein, due to the lower absolute entropy of the whole system. Otherwise if the stability of the protein depends mostly on the enthalpic part of the Gibb's free Energy a change in urea concentration should not effect the unfolding process much. In case that the distortion of the hydrogen bonding pattern of the water molecules in high urea concentration is of major relevance, the question arises whether the disturbed electrostatic environment or a change in the viscosity of the solution influences the unfolding.

## 2.3 Simulation Details

The simulations are performed in a box, filled with the protein and a solvent. To minimize artifacts from the limited system size, periodic boundary conditions (PBC) are used in the calculation of the forces. The PBC feature a wrapping of the simulation box with copies of itself, allowing an appropriate treatment of the forces by the minimal image convention (MIC). This convention constrains the number of atoms for the determination of the potential energy V to the bulk within a box, centered at the

position of the atom considered for the integration.

As we use PBC in our simulations, the time consuming calculation of the electrostatic interaction can be modified to reach faster computation time with only a slight loss in accuracy. To this aim the calculation of the electrostatic potential is split up. Within a certain cut-off radius $r_{cutoff}$ the potenial is calculated explicitly. Beyond this radius the charges are assigned to a grid via an adequate distribution that is subjected to a Fourier transformation. The two parts of the electrostatic potential converge fast, because the cut-off radius does restrict explicit calculation to a small space and the presence of the PBC allows an easily computable expression for the potential in the reciprocal space, dependent on the used distribution function. This method is an enhanced Ewald summation[21] and called Particle Mesh Ewald (PME) algorithm[22]. In contrast to an explicit calculation the number of computational operations is reduced from $\mathcal{O}(N^2)$ to $\mathcal{O}\left(N \log N\right)$.

All simulations represent an NpT ensemble, so at first the methods to keep the pressure and temperature constant should be explained. To obtain constant temperature and pressure during the MD simulation all the atoms in the simulation box are coupled to an external heat bath via a Berendsen thermostat and barostat [23].

A constant temperature requires a constant mean kinetic energy. To this end in every simulation step the atom velocities are linearly scaled to yield a constant temperature:

$$v \rightarrow \lambda \cdot v, \tag{2.19}$$

$$\lambda = \sqrt{1 + \frac{\Delta t}{2\tau_T}\left(\frac{T_0}{T} - 1\right)}. \tag{2.20}$$

Similarly, the pressure scaling corresponds to a variation of the box size and coordinates. Introducing a pressure coupling constant $\tau_P$, the transformation for a cubic box

with lenght $l$ and the coordinates $X_K$ can be written as follows:

$$\vec{X}_K \rightarrow \mu\vec{X}_K, \tag{2.21}$$

$$l \rightarrow \mu l, \tag{2.22}$$

$$\mu = \sqrt[3]{1 - \frac{\Delta t}{\tau_P}(P - P_0)}\,. \tag{2.23}$$

The coupling constants have to be set appropiate to the system size, to take the energy flux through the system into account. For our setup we choose $\tau_P = 1\,\text{ps}$ and $\tau_T = 0.1\,\text{ps}$.

## 2.4 System Setup

Ubiquitin renders a good simulation system, due to its small size of 76 residues. For this reason no big simulation box is required and the computational resources allow simulations on the nanosecond timescale. Additionally, a structure model, obtained from a x-ray diffraction pattern, exists[11], containing structure information about the individual residues with a resolution of 1.8 Å. These conditions alleviate the set up of a realistic simulation system in several steps.

First the crystal structure of the protein has to be solvated in a simulation box. To this end we chose a water model, suitable for the applied force field. Our choice, the TIP4P model water consists of two hydrogens, one oxygen, and an additional charged dummy atom to approximate the water dipole moment more accurately. In this case our simulation setup differs from the methods of [1], but this change is required to achieve comparability of the interaction in the solvent with the results of [7].

The simulation box was cubic and of 5.9 nm length filled with about 5000 TIP4P water molecules. Finally to get a physiological environment, we added sodium and chloride ions to obtain a 150 mM NaCl solution.

The x-ray-scattering method does not provide information about the protonation

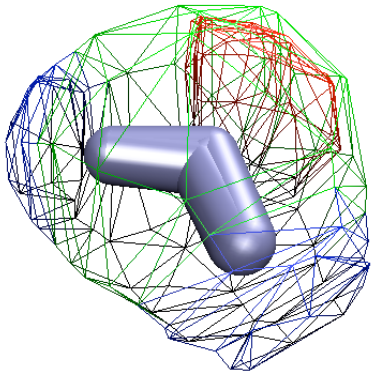


***Figure 2.4:*** TIP4P water model. The coloured grid illustrates the electrostatic influence of the different atoms. In addition to the two hydrogens (blue) and the oxygen (red), an added dummy atom (green) provides an improved approximation of the water dipole moment.

states of the native state of the protein. To overcome this problem we used the program packages WHAT IF [24] and Delphi. They enable us to calculate the electrostatic potentials at the residue sites and the $pK_a$ in the solution, resulting in information about the protonation states of the protein. We found the histidines being in the protonated state.

Now the protein is ready for further steps preparing the producing simulations. To obtain a solved protein from the crystal structure, surrounded by water, sodium, and chloride ions, first we minimized the potential energy of the crystal structure via a Monte-Carlo method. This deepest descent energy minimization scans the dihedral configurational space of the protein to reach a minimum in the potential energy (fig. 2.5 ). As yet the water molecules are positioned uniformly with random orientations, but are not necessarily thermally equilibrated. To reach the thermal equilibrium within the solvent molecules, the next preparation step aims at a uniform distribution of the solvent molecules only affected by the presence of the protein. To

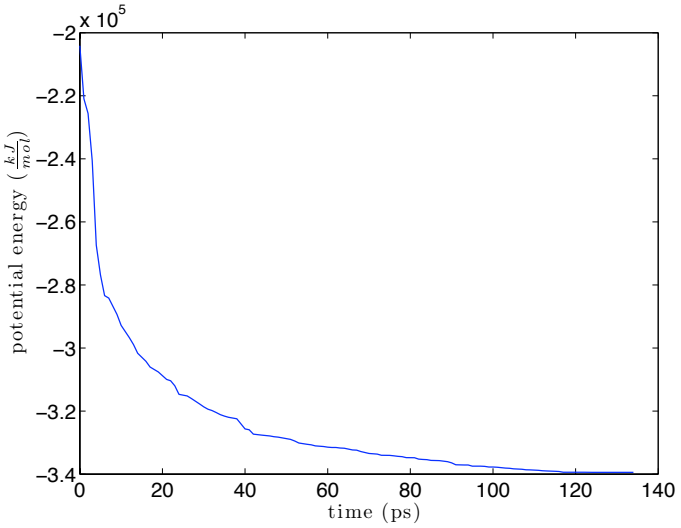



***Figure 2.5:*** During a deepest descent energy minimization the potential energy of the whole system is converging to a minimum.

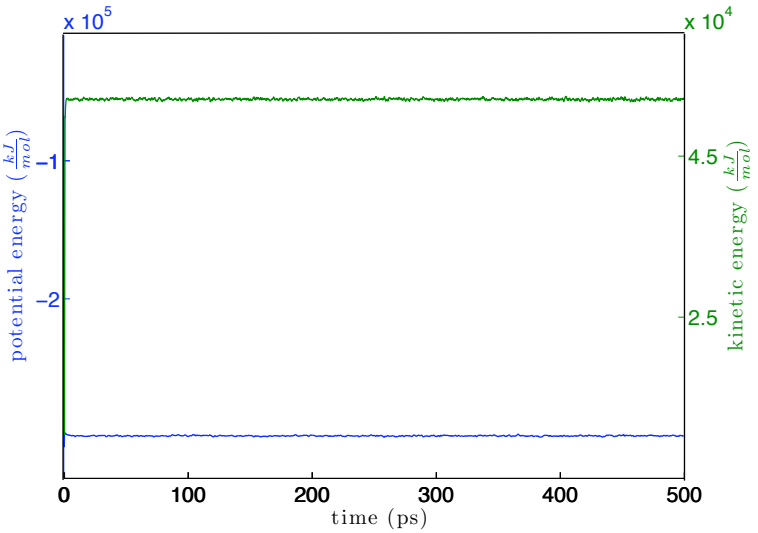



***Figure 2.6:*** Potential (blue) and kinetic (green) energy of the system in the simulation during the equilibration of the water. The increase in kinetic energy the potential energy is decreasing, corresponding to the equilibration of the solvent.

avoid denaturation of the protein, its atom positions are held fixed, and a simulation lasting 500 ps (fig. 2.6) allows the water molecules to reach thermal equilibrium.

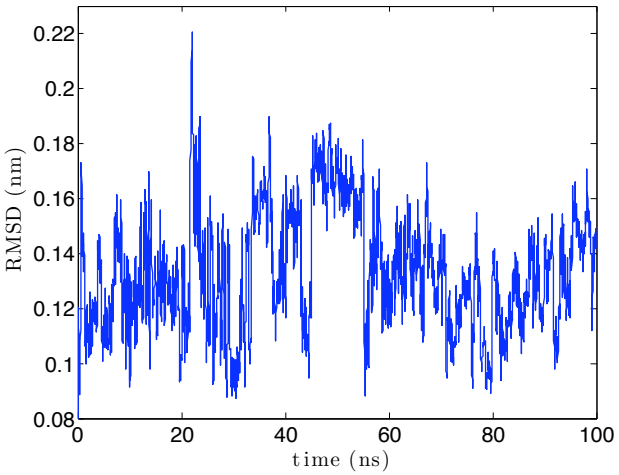



***Figure 2.7:*** The plot represents the root mean square deviation (RMSD) of the $C_\alpha$ backbone configuration during a system equilibration in respect of the initial structure. Large fluctuations occur, corresponding to conformational changes of the protein structure.

The next step serves to equilibrate the whole system and from now the atom movement is only restricted due to constrained bond lengths, which are held fixed using the LINCS algorithm[25]. During a 100 ns MD simulation the protein and the solvent merge to an equilibrated state, while the protein structure is changing to fit perfectly into its environment. The root mean square deviation (RMSD) calculated in respect to the protein $C_\alpha-$backbone reveals the difference between crystal and current structure (fig. 2.7). To find a structure similar to the crystal structure, we chose initial structures with a low RMSD from the last 40 ns of the equilibration during the MD simulation. The obtained structures were centered in an enlarged $5.9 \times 5.9 \times 18.0\,\mathrm{nm}^3$ simulation box, to enable an elongation of the protein due to unfolding, and solvent was added. Again the solvent is equilibrated by energy minimization and a simulation

with restrained protein atom positions.

Finally the system is in an equilibrated state and prepared for the actual investigations. As mentioned we intend to unfold ubiquitin by applying mechanical force in several solvent environments. To this end after the selection of the three starting structures the described equilibration process for the solvent has been performed using a 3 M, 7 M, and 9 M urea solution in addition to the pure water box.

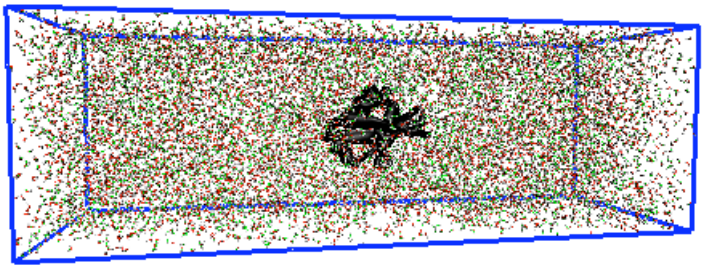


***Figure 2.8:*** Simulation box containing the protein (black) and about 18000 solvent molecules, depending on the urea concentration. For further FPMD and FCMD simulations, twelve different structures with variable solvent concentration and protein conformation were used.

Each of the twelve structures was subjected to an FCMD simulation applying 500 pN force to the $C_\alpha$ terminus atoms for 18 ns. This simulation setup was expected to be adequate to observe an unfolding process [1] in respect to the simulation time and the applied force. The pulling was carried out in two opposite directions parallel to the $z$-axis, where most space is available. In contrast to a setup applying only one pulling potential to a bound protein, our setup reduces the occurring friction forces. In the different FPMD simulation the direction of the force was kept the same, but the spring modelling the potential was moving with different velocities of 1, 2, 5 and 10 $\frac{\mathrm{m}}{\mathrm{s}}$. It is expected that under these conditions the protein changes from the native folded to

the unfolded state. Further, ten FPMD simulations with the same structure in pure water were performed to analyze the variance of the rupture force and alleviate error estimation for the force extrapolation.

For our simulations we used a modified code of the MD program package GROMACS 3.3.1 [26]. Because of mechanisms concerning the Fourier transformation in the PME algorithm GROMACS was compiled with the same FFTW2 library [27] like F. Gräter [1] to maintain comparability, despite longer computation times.

MD simulations produce a huge amount of data and FPMD additionally maintain information about the spring position. However, GROMACS does not provide a tool for the analysis of the spring positions. To this end and to provide an efficient way for handling the data a tool in the C-language was written using the Message Passing Interface (MPI) [28]. To visualize the trajectories of the molecules VMD [29] was used.

## 2.5 Analysis Methods

### 2.5.1 Secondary structure determination

Maps of the change of the secondary structure in time provide information about the dynamics of the hydrogen bonds, not accurately accessible from direct observation. For our analysis we used the established DSSP algorithm [30] , which localizes the motifs by calculating the electrostatic energies between the according elements of the protein backbone. To this end a partial charge $q_1 = 0.42\,\mathrm{e}$ and $-0.42\,\mathrm{e}$ is assigned to the carbon and oxygen, respectively, and a partial charge $q_2 = 0.20\,\mathrm{e}$ and $-0.20\,\mathrm{e}$ is assigned to the hydrogen and nitrogen, respectively, followed by a straightforward calculation of the energy $E$:

$$E = f \cdot q_1 q_2 \cdot \left\{ \frac{1}{r_{\mathrm{ON}}} + \frac{1}{r_{\mathrm{CH}}} - \frac{1}{r_{\mathrm{OH}}} - \frac{1}{r_{\mathrm{CN}}} \right\}. \tag{2.24}$$

The units of the radii are in Å and $f = 1390 \frac{\mathring{A}}{C^2}$ is a dimensional factor yielding the energy unit of $\frac{\text{kJ}}{\text{mol}}$. For $E < -2.1 \frac{\text{kJ}}{\text{mol}}$, a hydrogen bond between the corresponding atom groups is identified and finally, the secondary structure of the protein is derived from the configuration of these hydrogen bonds.

### 2.5.2 Hydrogen Bond Energies

To calculate hydrogen bond energies between arbitrary atoms of the simulation system, we used the established method in [31], based on the distance $d$ in Å between the proton donor and proton acceptor:

$$E_{HB} = -\frac{1}{2}\left(50 \cdot 10^3 \frac{\text{kJ}}{\text{mol}}\right) \cdot e^{-36d}. \tag{2.25}$$

### 2.5.3 Contact Maps

Contact maps were obtained by averaging the distance between the residues over a time window of 1 ns. Complementing the secondary structure maps, the contact maps provide information about the orientation of the different secondary structure motifs.

### 2.5.4 Solvent Accessible Surface Area

Calculation of the solvent accessible surface area of the protein bases on the Double Cubic Lattice Method [32]. Firstly, the algorithm divides the volume around the protein into cubic boxes, which only contain one protein atom, treated as a sphere. By projection of the overlapping spheres onto the surface of the cubes, the surface accessible to the solvent is calculated. An implementation of this algorithm is present in the used MD programm package GROMACS, which we use for our analysis.

# 3 Results and Discussion

## 3.1 The Force Profile of Ubiquitin in the FPMD simulations

The resulting force profile (fig. 3.1) of an FPMD simulation and a comparsion with the time course of the simulation allows to identify rupture events and relate them to involved residues. Figure 3.1 shows a typical force profile and illustrates the main features of the performed simulations.

For the reaction coordinate, the end to end distance of the pulled atoms was chosen to provide an appropiate description for the state of the protein, allowing the comparison of FPMD simulations with different pulling velocities.

As can be seen in figure 3.2 A-C, the rupture forces in 0 M, 3 M, 7 M, and 9 M urea-water solution only slightly differ. Raising the pulling velocity $v_{\text{pull}}$ from $1\,\frac{\text{m}}{\text{s}}$ to $5\,\frac{m}{s}$ increases the maximum force by about 250 pN. However, increasing the velocity further to $10\,\frac{\text{m}}{\text{s}}$ does not reveal drastic changes. In contrast, in the course of an unfolding process the decrease of the force after the maximum is less pronounced for higher $v_{\text{pull}}$. Figure 3.2 D depicts the scattering of the maximum forces, occuring in FPMD simulations with a pulling speed of $10\,\frac{\text{m}}{\text{s}}$. It reveals the strong influence of the starting structure on the time course of the simulations. For the 0 M environment, one structure could not be analyzed due to an error in the simulation setup pulling in wrong directions, hence only two results are available for a discussion.

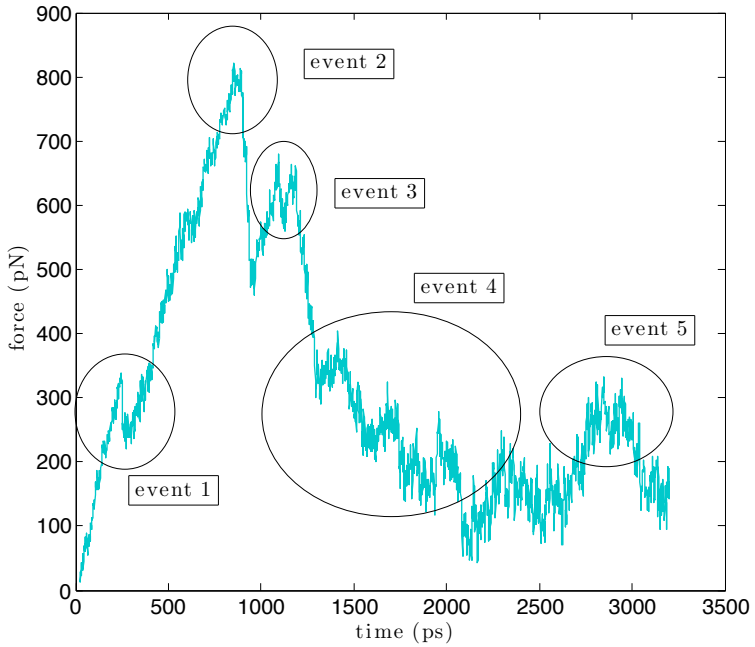



***Figure 3.1:*** Force profile resulting from an FPMD simulation in pure water with a pulling velocity of $1\,\frac{m}{s}$. The encircled areas mark the different characteristic rupture events of the unfolding process, described in fig. 3.6.

The rupture force peak seems to be slightly increased in case of a low 3 M urea concentration. Though, a further increase of the concentration weakens this effect or the effect even vanishes. However, urea is known as a denaturant, implying a decrease of the rupture force with increasing urea concentration, an effect not observable in the force profiles of our simulations.

To estimate the error or the variance in our model ten FPMD simulations in water solution were performed, differing only in the distribution of the starting velocities of the particles corresponding to Poisson-Boltzmann. A pulling speed of $10\,\frac{\mathrm{m}}{\mathrm{s}}$ was applied to minimize the required computation time. Figure (3.3) presents the results of the ten runs. The trajectories elucidate the occurring trend of the force. The rupture events

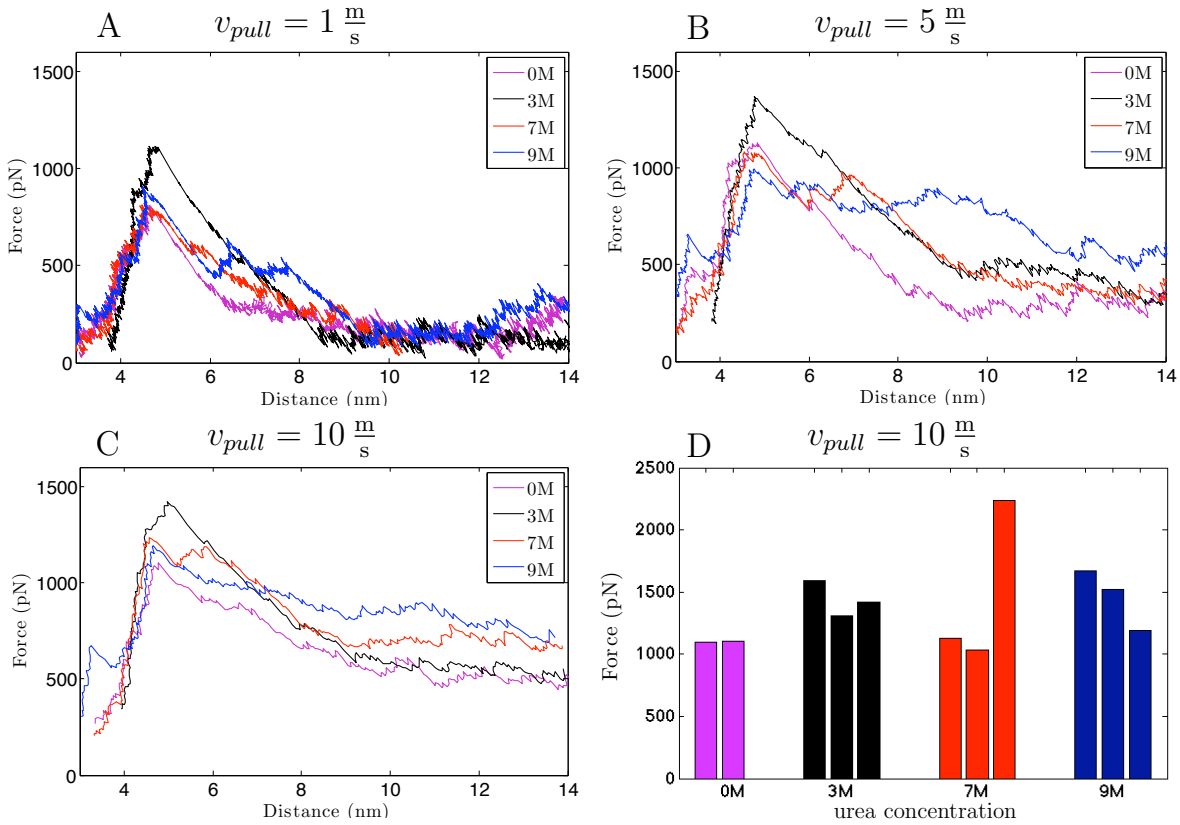



***Figure 3.2:*** Force vs. distance plots. Plots A to C illustrate the dependence of the forces due to a change of $v_{\mathrm{pull}}$ and the urea concentration. Plot D depicts the scattering of the rupture forces for different structures obtained with $v_{\mathrm{pull}} = 10\,\frac{\mathrm{m}}{\mathrm{s}}$.

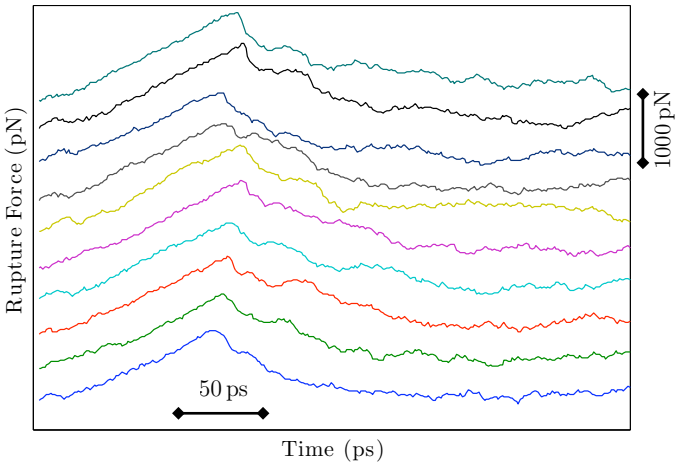



***Figure 3.3:*** Force profiles of one starting structure with different distributions of the molecular velocities at the beginning of the simulation. The system produces a homogenous unfolding pattern, due to the identical starting structure.

occured at slightly differing points in time and at different rupture forces. The maximal forces in the simulations ranged from 1160 to 1350 pN and the end to end distances of the pulled atoms at the corresponding moments in time, from 4.70 to 4.85 nm.

The course of the simulation obviously does not depend critically on the starting velocities (fig. 3.3) of the solvent molecules. However, the strongly differing rupture forces in simulations with different starting structures reveal a strong dependence of the rupture forces on the starting structure (fig. 3.4). As expected, the selection of the starting structure from the equilibration (fig. 2.7) strongly influences the behaviour of the protein in respect to rupture forces and the corresponding rupture distance in the FPMD simulations.

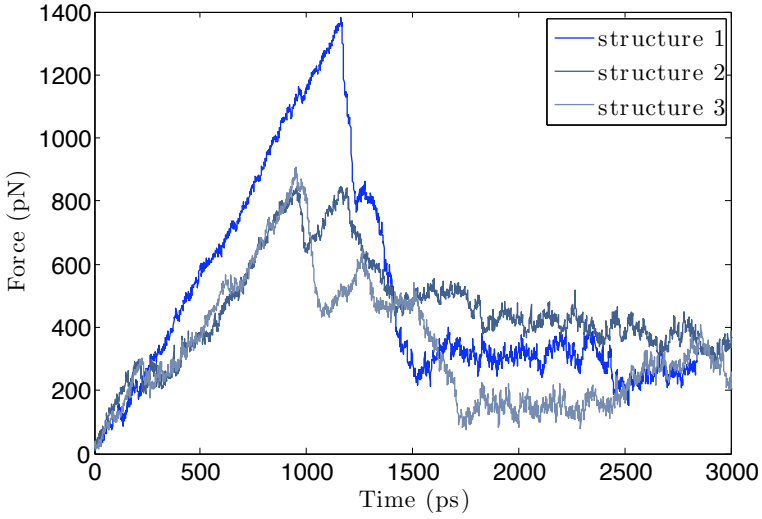



***Figure 3.4:*** Force vs time plot gained from FPMD simulations ($v_{pull} = 1\,\frac{\mathrm{m}}{\mathrm{s}}$) in a 9 M urea solution.

## 3.2 Force fit from the FPMD simulations

To extrapolate the occuring friction forces to a regime, comparable with AFM experiments, we choose pulling velocities $v_{pull} = 1,\ 2,\ 5$ and $10\,\frac{\mathrm{m}}{\mathrm{s}}$. A Non Linear Least Squares fit of the maximum forces, averaged over the three starting structures, to eq. 2.18, corresponding to the first dissociation event in the course of unfolding, and the pulling velocities should provide the friction $\gamma$ and $k_0$ as well as $\Delta x$ in the different solvent environments.

All FPMD simulations, performed with equal velocity and urea concentration, reveal a strong variance in the observed rupture force. Unfortunately, the number of the simulations did not suffice to gain a reliable average of the rupture forces and fit them using the model resulting in eq. 2.18. Furthermore, fig. 3.5 reveals, that all simulations have been performed in the high friction regime, rendering an appropiate extrapolation to the low friction regime impossible. As can easily be seen from the plot, already small changes in the data lead to strongly differing values for the friction constant $\gamma$, $k_0$, and

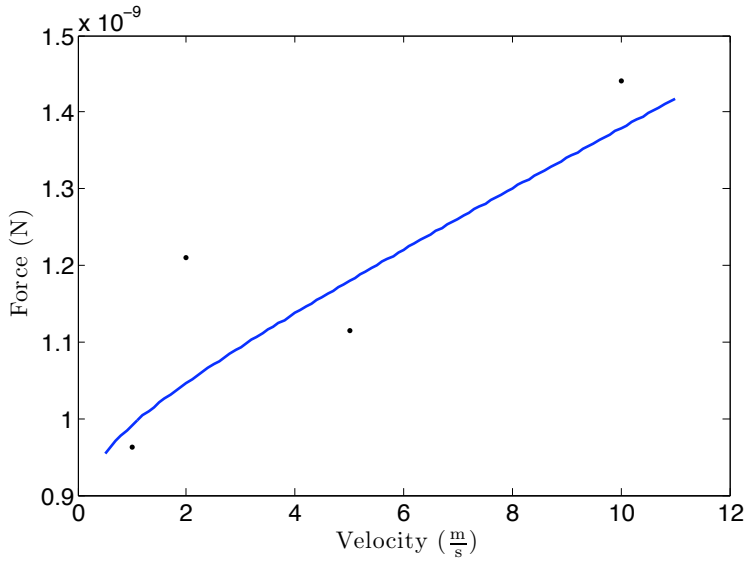



***Figure 3.5:*** Force fit to eq. 2.18. The strong variance of the rupture forces renders a reliable fit impossible.

$\Delta x$. To provide an example, omitting a single data point that does not follow the trend of an increasing force with increasing velocity, can yield negative values for the friction constant $\gamma$, which is not consistent with physical reality.

The application of the high pulling velocities entails very strong friction forces. Due to the fast deformation of the Gibb's free energy landscape, kinetically trapped transition states might arise, because the protein is not in thermal equilibrium with its solvent environment, causing the extreme high rupture forces. Because of the inability to extrapolate the rupture forces to the low friction regime comparison with experiments is not possible. However, this result reveal, that in our further analysis of the FPMD simulations we have to keep the influence of the very strong friction at the back of our mind.

## 3.3 The Unfolding Pathway of Ubiquitin

### 3.3.1 Results of the FPMD simulations

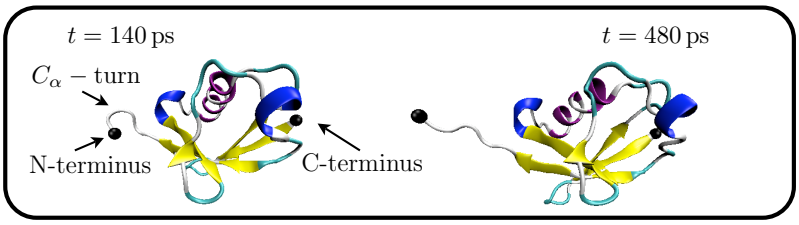
In addition to the rupture forces, also available from experiments, snapshots of the unfolding process can be obtained and compared to the events from the force profile. Figure 3.6 shows snapshots from a trajectory of an FPMD simulation pulling the terminal atoms with a velocity of $1\,\frac{m}{s}$ in a water solution. They reveal the stepwise conformational changes of ubiquitin during the simulation.

The first unfolding event occurs within a time of about 340 ps (fig. 3.6, event 1). The $C_\alpha$-turn breaks up and the amino acid chain starting at residue 71 is almost completely stretched. This turn is contained in the crystal structure of the Protein Database but does not occur in every starting structure of the FPMD simulation.

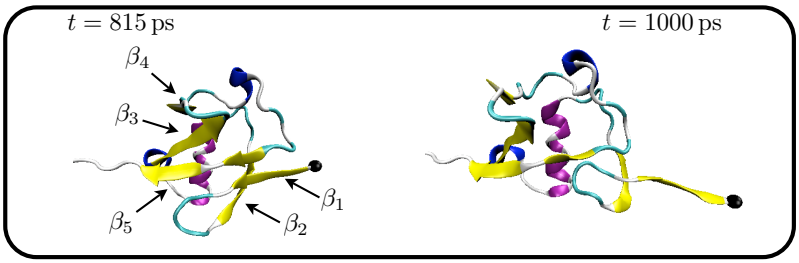
In the course of the FPMD simulations the arrangement of the $\beta$-strands changes drastically (fig. 3.6, event 2). Two snapshots taken at 815 and 1000 ps simulation time highlight the rupture of the protein structure, determined by the arrangement of $\beta$-strands 1 to 5, a $\alpha$-helix and two $3_{10}$-helices. This second event leads to a loss of contact between the loop connecting the $\beta_1$ and $\beta_2$ strand and the $\alpha$ helix. Furthermore the contact between the $\beta_1$ and $\beta_5$ strand is lost. This structural change introduces the separation into two small clusters formed by strand $\beta_1$ and $\beta_2$, and the remaining secondary structure elements of ubiquitin. The $\beta_1$- and $\beta_5$-strand align completely along the pulling direction, whereas the remainder of the protein maintains its three dimensional structure.

Further pulling on the terminal atoms separates the two stuctural clusters completely (fig. 3.6, event 3). The structural change of the $\beta_1/\beta_2$ cluster concentrates on the alignment along the pulling direction, whereas the conformational dynamic of the other cluster is dominated by a rearrangement of $\beta_3/\beta_4/\beta_5$.

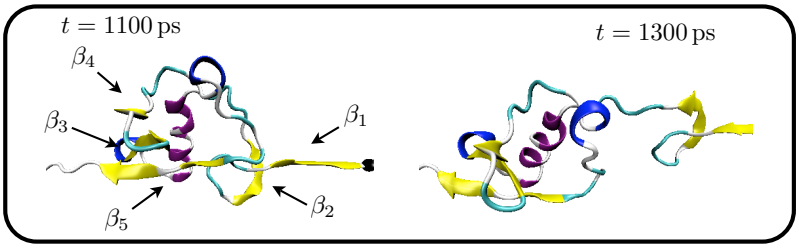
After the partitioning of the two clusters the larger cluster containing the $\alpha$-helix



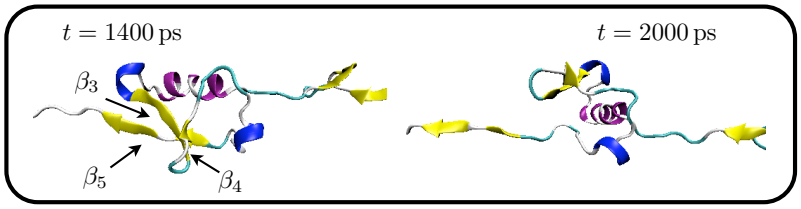

Event 1. The small turn next to the $C_\alpha$ atom of the N-terminus looses its bond conformation due to the pulling



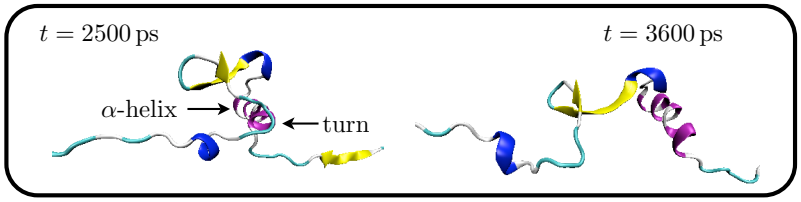

Event 2. Primary partitioning of the protein structure. The core-shaping conformation of the $\beta_1$, $\beta_2$ and $\beta_5$-strand is altered drastically, especially the orientation of $\beta_1$ relative to $\beta_2$, and $\beta_1$ relative to $\beta_5$.




Event 3. Separation into a $\beta_1/\beta_2$ cluster and a cluster of the remaining structure elements is completed.




Event 4. The parallel alignment of $\beta_3$ relative to $\beta_5$ vanishes.




Event 5. The remaining tertiary structure gets lost by a sudden disrupture of the common axis of the $\alpha$-helix and the turn.

***Figure 3.6:*** Snaptshots depicting the unfolding in a FPMD simulation with pulling velocity $v_{pull} = 1\,\frac{\mathrm{m}}{\mathrm{s}}$.

rearranges. The contact between $\beta_3$ and $\beta_5$ opens up and instead $\beta_3$ and $\beta_4$ align antiparallel (fig. 3.6, event 4). From now on the unfolding proceeds in a rotation around the axis formed by the $\alpha$-helix and the bend between $\beta_4$ and the blue $3_{10}$-helix in the lower part of the figure.

At the end of the FPMD simulations the tertiary structure becomes an uncorrelated chain of secondary structure elements (fig. 3.6, event 5). The last rupture events destroy the alignment of the $\beta$-strands and the collective axis along the helix and the turn gets lost. Further pulling yields a complete stretching of the residue chain without any secondary structure elements.

This unfolding sequence was found in all performed FPMD simulations. Apart from the first step, which can only arise if the starting contains the bend at the N terminus, the protein unfolding follows this pathway independent of the urea concentration and pulling velocity. The points in time of the events differ in the simulations, but the order of the events remains unchanged and no dependence on the urea concentration is detectable.

### 3.3.2 The Unfolding Process in the Force Clamp MD simulations

In contrast to the FPMD simulations a characterization of the unfolding process by different rupture forces is not possible, due to the constant pulling force. However in this case the end to end distance of the $C_\alpha$ termini allows a reliable characterization the state of the protein. In our FCMD simulations we compare in unfolding in pure water with unfolding in solutions with different urea concentrations (3 M, 7 M, 9 M). Unfolding was only obtained in pure water and 7 M urea solution (fig. 3.7).

The unfolding process in an FCMD simulation proceeds differently compared to the FPMD simulations. The first event is visible as the step in the end to end distance at $t = 3$ ns and $t = 5$ ns in the 7 M and 0 M urea solution, respectively. For the two solvents this intermediate state lasts for different times $t_{\mathrm{IM}}$; 1.7 ns in 0 M and 0.3 ns in 7M urea

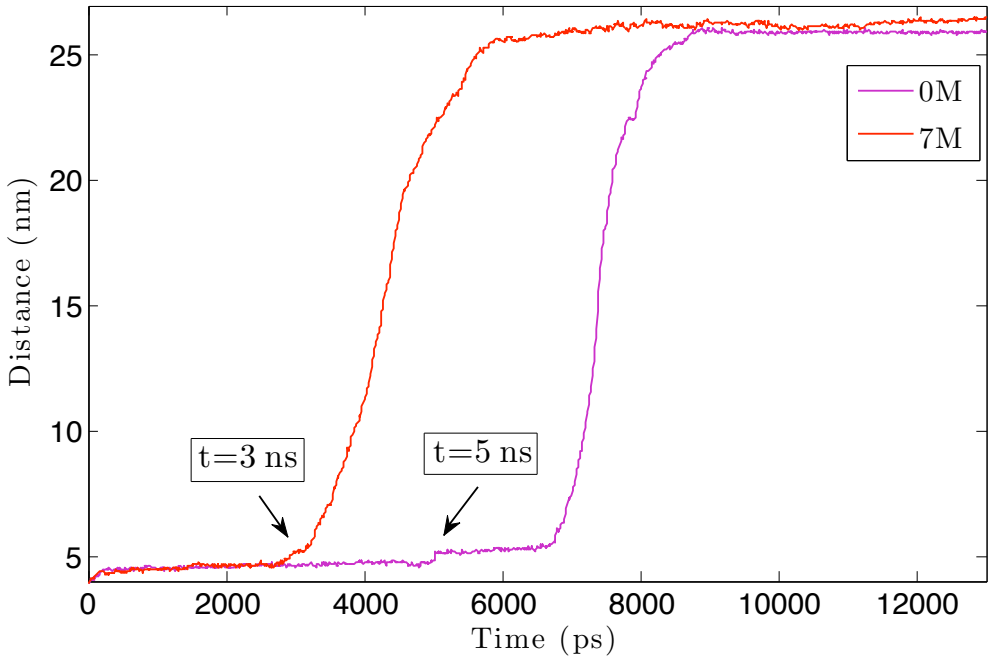
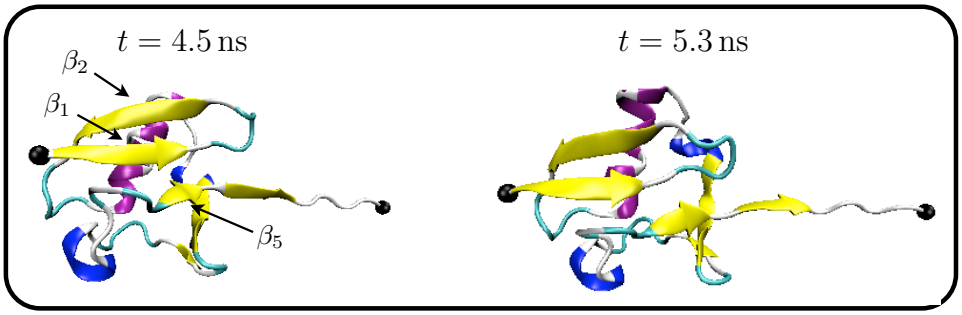


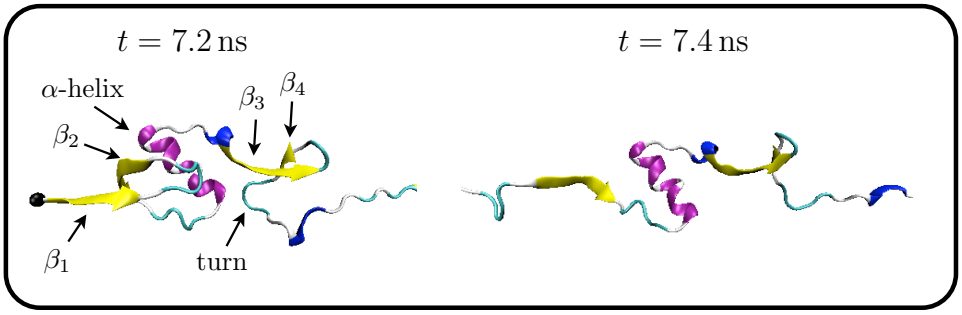

***Figure 3.7:*** Change of the $C_\alpha$ termini end to end distance in time. The protein elongation up to 26 nm proceeds in two steps. In pure water (purple) the distance increased after 5 ns from 4.5 to 5.2 nm, and 1.7 ns later the remaining unfolding events occur. In the 7 M solution unfolding starts earlier at $t = 3$ ns and the first unfolding phase timespan shortens to 0.3 ns.

concentration. It results from the displacement of the $\beta_1$ and $\beta_5$-strand (fig. 3.8, top). Subsequently the arrangement of $\beta_5$ , $\beta_3$, and $\beta_1$-strands gets lost sequentially in the corresponding order (fig. 3.8). A small cluster, consisting of the $\alpha$-helix and the strands $\beta_1$ and $\beta_2$, resists the force for the longest timespan and maintains the remaining three dimensional structure of the protein. Finally, also the parallel alignment of the $\alpha$-helix and $\beta_2$-strand gets lost.

Comparison of the unfolding process in the two solvent environments discloses two differences. The first distinct elongation occurs at different points in time and the




During the first slow unfolding event the $\beta_1$ and $\beta_5$-sheets are pulled apart. This introduces the decay of the parallel alignment of $\beta_1/\beta_2/\beta_5$.




At $t = 7.2\,\mathrm{ns}$ the main unfolding process of the protein begins. After the disorientation of the $\alpha$-helix and the turn, firstly the alignment of $\beta_1/\beta_2$ vanishes.

***Figure 3.8:*** Snaptshots describing the unfolding in a FCMD simulation in a 0 M urea solution.

duration of the decomposition of $\beta_1$ and $\beta_5$ is different. However only two unfolding events are available for analysis, rendering a meaningful evaluation impossible.

### 3.3.3 Comparison of the unfolding pathway in the FPMD and FCMD simulations

Unfolding of ubiquitin was simulated using two different methods: pulling at the termini of the protein by a moving harmonic potential (FPMD) and alleviating the unfolding process by applying a constant force on the termini of the protein (FCMD). It turned out, that two different unfolding pathways occur, depending on the method to apply mechanical stress to the protein. The difference in the unfolding process mainly regards the dissociation of the $\beta$-sheet. In the FPMD simulations a loss of tertiary structure is introduced by a rupture of the $\beta$-sheet between the strands $\beta_1$ and $\beta_5$, resulting in a separation of the tertiary structure into an antiparallel $\beta$-sheet, consisting of the strands $\beta_1$ and $\beta_2$, and a protein part consisting of the remaining secondary structure elements. In contrast, our FCMD simulations reveal a disconnection of the $\beta_5$-strand from the $\beta$-sheet, yielding a separation of $\beta_5$ from the remaining tertiary structure. However, the urea concentration does not affect the unfolding pathway in both cases.

Another difference between the FPMD and FCMD simulations arises in the time course of ubiquitin unfolding. In the course of an FPMD simulation a rupture of the bonds stabilizing the tertiary structure occurs continuously due to the linear increase of the distance of the harmonic potential, pulling at the termini of the protein. An approximately linear increase of the end to end distance of the termini of the protein can be observed. However, the course of the end to end distance of the termini in an FCMD simulation (fig. 3.7) reveals a distinct increase of about 0.7 nm, introducing the unfolding of ubiquitin. The steplike increase arises from a translational displacement of the strands $\beta_3$ and $\beta_5$, provided by a rupture of the hydrogen bonds between ARG42 and VAL70 and a reformation of a hydrogen bond between ARG42 und LEU69 (fig. 3.9).

Despite different unfolding pathways arising in the FPMD and FCMD simulations, both have in common that final loss of tertiary structure occurs, when the connection

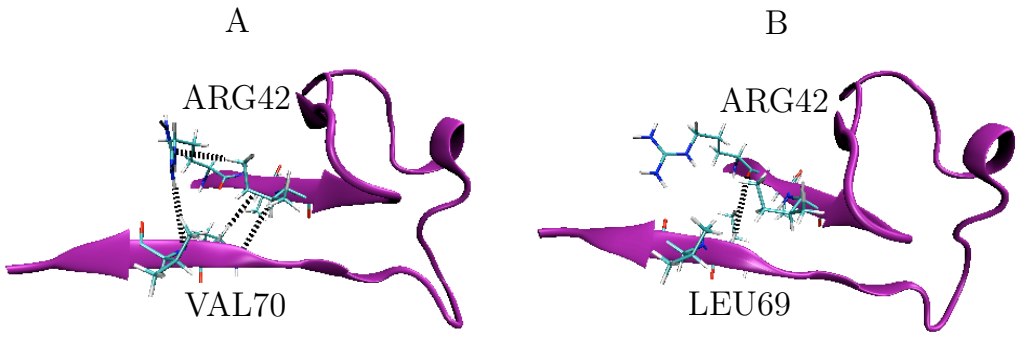



***Figure 3.9:*** Segment of ubiquitin, depicting the residues 41-71. In the FCMD simulations, unfolding is introduced by the translational displacement of the strands $\beta_3$ and $\beta_5$. The bonds (dotted lines) between ARG42 und VAL70 (A) breakt and a new bond between ARG42 and LEU69 (B) is formed.

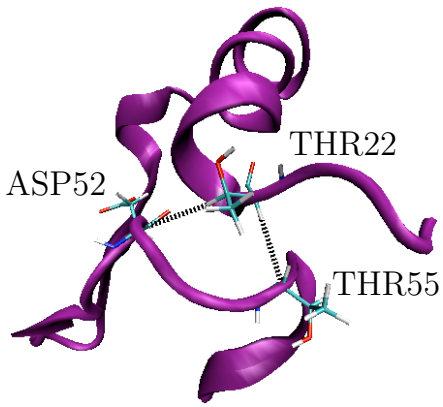



***Figure 3.10:*** Segment of ubiquitin, depicting the residues 20-60. The residues THR22, ASP52, and ASP55 provide the high stability of the arrangement between the $\alpha$-helix and the turn from residue 52 to 54, due to two hydrogen bonds (dotted line).

between the $\alpha$-helix and a turn, involving the residues 52 to 55, gets lost (fig. 3.6, 3.8). Figure 3.10 provides an explanation for the strong bonding between these two secondary structure elements of the protein. Two hydrogen bonds between THR22 and ASP52, and between THR22 and THR55, maintain the remaining tertiary structure until their final destruction.

## 3.4 Unfolding in terms of contact and secondary structure maps

The previous sections described the analysis of ubiquitin unfolding with the help of forces and structural snapshots. It turned out, that the sequential loss of tertiary structure occurs differently in the FPMD and FCMD simulations. To obtain more detailed information on the change of the secondary structure, we derived contact maps and secondary structure maps from the trajectories.

### 3.4.1 Results of the FPMD simulations

Figure 3.11 A depicts the time course of the secondary structure pattern of the protein in an FPMD simulation ($0\,\mathrm{M}, v_{\mathrm{pull}} = 1\,\frac{\mathrm{m}}{\mathrm{s}}$). At $t = 0\,\mathrm{s}$ many residues are involved in secondary structure motifs, but in the course of the simulation the decay of the $\beta$-strands (red) is cleary observable. Furthermore the figure 3.11 A alleviates the identification of the secondary motifs in the corresponding contact maps, which are symmetric to the diagonal (blue) from the lower left to the upper right depicting the self contact of the residues. Furthermore a homogenous broadening of the diagonal reflects helical secondary structure motifs. The domains, orthogonal and parallel to the diagonal depict the antiparallel and parallel alignment of the $\beta$-strands, respectively. Further dashed encircled areas indicate the interaction between the $\beta$-strands and the $\alpha$-helix. The

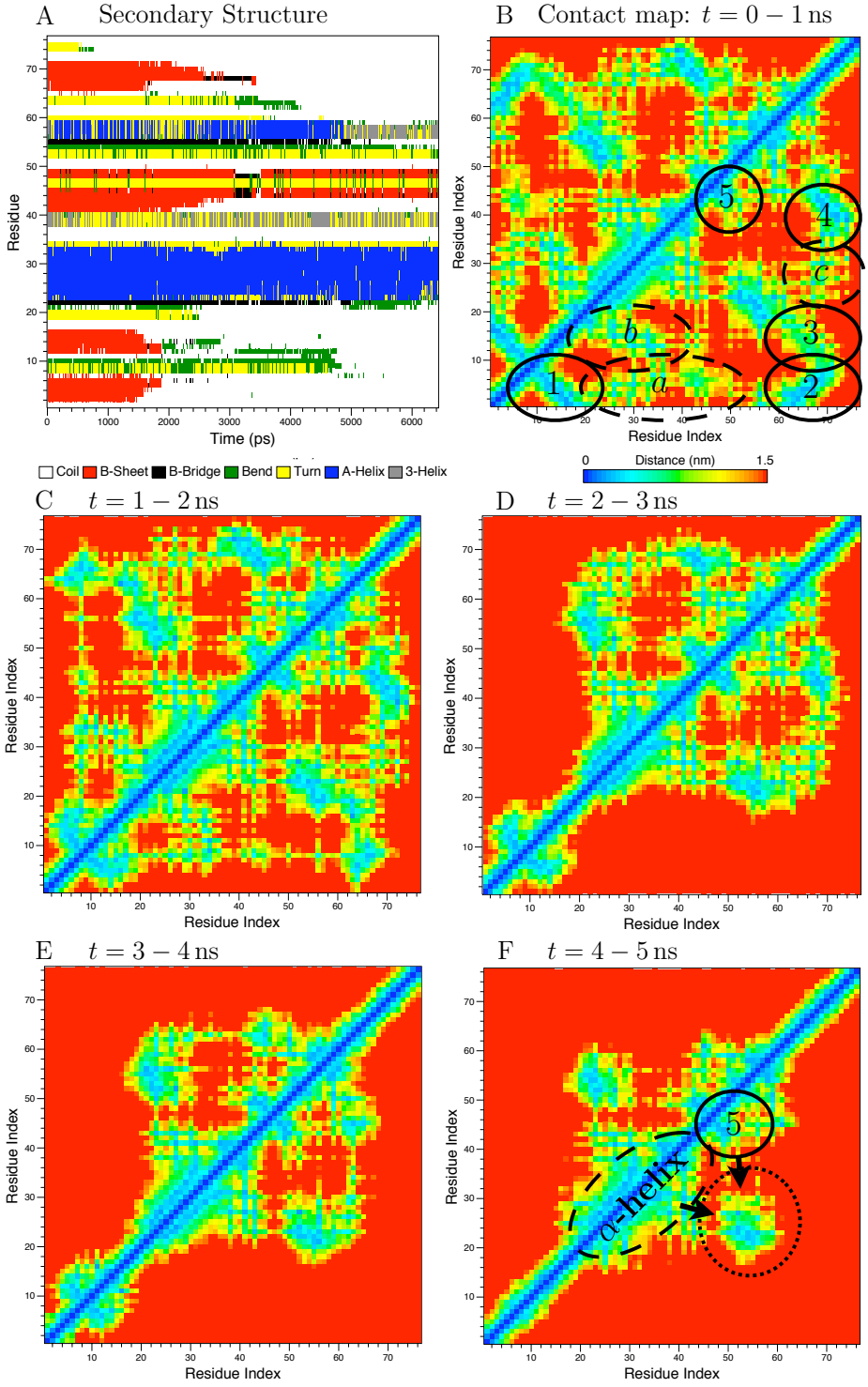



***Figure 3.11:*** Dynamic of the secondary structure (A) and contact maps (B-F) of an FPMD simulation $(0\,\mathrm{M}, v_{\mathrm{pull}} = 1\frac{\mathrm{m}}{\mathrm{s}})$

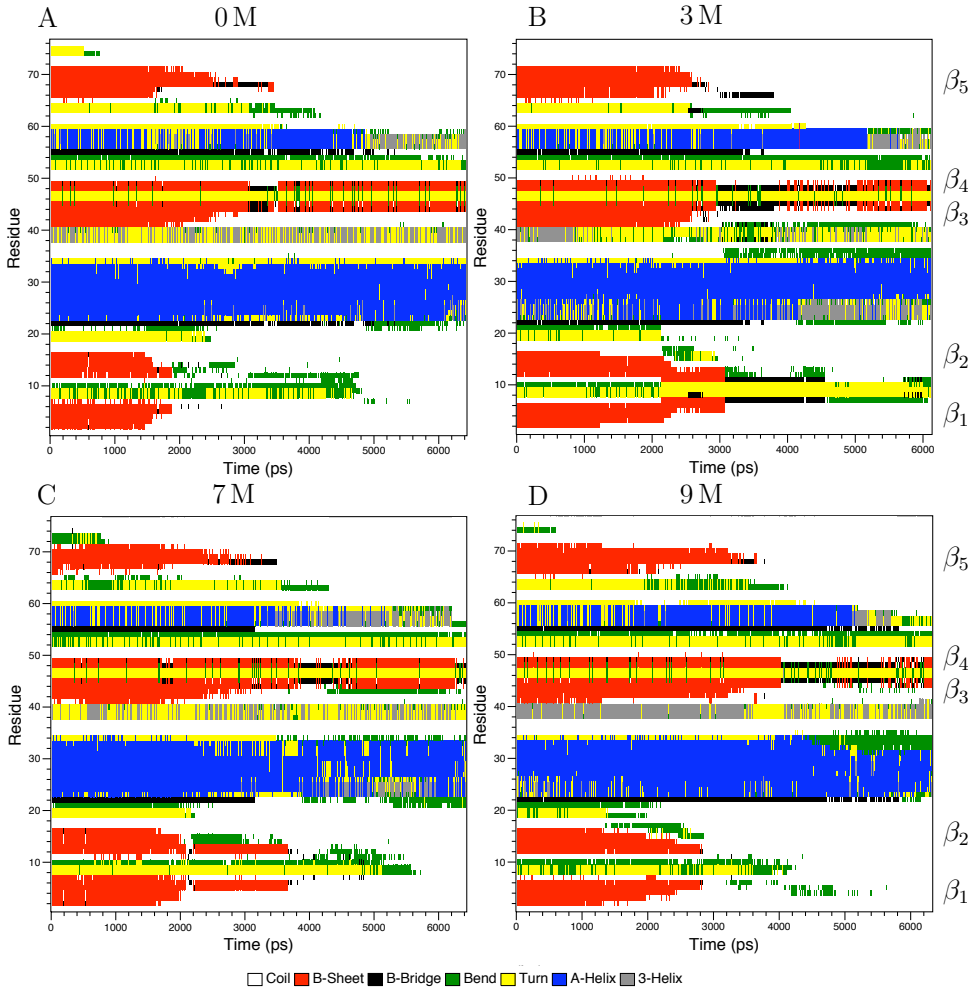



***Figure 3.12:*** Comparison of the evolution of the secondary structure in several urea solutions in FPMD simulations ($v_{\text{pull}} = 1\,\frac{\text{m}}{\text{s}}$).

contact map (fig. 3.11 B) for $t = 0$-$1$ ns displays the large antiparallel $\beta$-sheet, with the solid encircled areas revealing the arrangement of the different $\beta$-strands to each other. With orientation on the secondary structure map, area 1 reflects the arrangement of $\beta_1/\beta_2$. Area 2, parallel to the diagonal, depicts the parallel alignment of the strands $\beta_1/\beta_5$, followed by area 3 visualizing the dense packing and antiparallel arrangement of $\beta_2/\beta_5$. Finally the areas 4 and 5 reflect the antiparallel alignment of $\beta_3/\beta_4$ and $\beta_3/\beta_5$, respectively. The areas a to c reveal the $\alpha$-helix as the secondary structure motif providing further stabilization of the tertiary structure by its tight arrangement to $\beta_1$ (a), $\beta_2$ (b), and $\beta_5$ (c), that enables the protein to form further intramolecular hydrogen bonds.

After a fast partioning of the protein in the course of the simulations (fig. 3.11 B - D), the decay of the remaining tertiary structure (fig.3.11 D- F) proceeds slower. Particularly the conformation of the $\alpha$-helix, the strands $\beta_3$ and $\beta_4$ (fig. 3.11 F, encircled area), and the turn at the residues 51 - 53 remains stable, revealing distances between the involved residues in the range of the length of a hydrogen bond ($\approx 0.15 - 0.30$ nm).

In section 3.3.1 the observation of the unfolding process via the snapshots and the force profile of the FPMD simulations does not reveal any differences due to a changing urea concentration of the solvent environment. However, a comparison of the different secondary structure maps allows a more detailled analysis. Figure 3.12 depicts the course of the secondary structure motifs under different urea concentrations, applying a pulling speed of $1\,\frac{\mathrm{m}}{\mathrm{s}}$, and they reveal an influence of the urea on the unfolding process. The difference mainly arises in the stability of the $\beta$-strands. Particularly, the motif $\beta_1/\beta_2$ resists the forces for a longer time, if urea is present. However increasing the urea concentration weakens this effect.

### 3.4.2 Results of the FCMD simulations

In comparison to the FPMD simulations the FCMD reveal an altered unfolding pathway. In addtion to fig. 3.11 a contact map of a structure, in an FCMD simulation (fig. 3.13) containing a 7 M urea solution, helps to describe the unfolding processes. The presented plots are derived before (fig. 3.13 A) and shortly after the beginning (fig. 3.13 B) of unfolding.

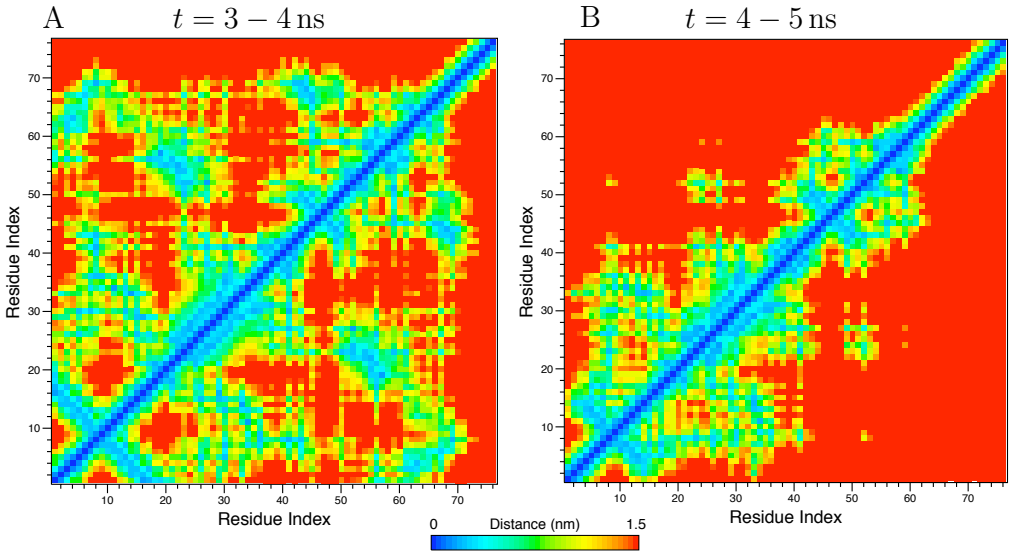



***Figure 3.13:*** Contact maps of the FCMD simulation with a 7 M solvent environment at the beginning of the unfolding process.

An effect, highlighted by the contact maps, concerns the partitioning of the protein into two clusters. In the FPMD simulations the tertiary structure of ubiquitin splitted up in a sequence of the $\alpha$-helix and the motif $\beta_3/\beta_4/\beta_5$, and the strands $\beta_1$ and $\beta_2$. However, in the FCMD simulations the pulling separates the $\beta$-sheet into the $\beta_5$-strand and a cluster consisting of the remaining secondary structure elements (1 to 64).

Despite the different unfolding pathways in the FPMD and FCMD simulations, the secondary structure maps of the structures, that unfold in the FCMD simula-

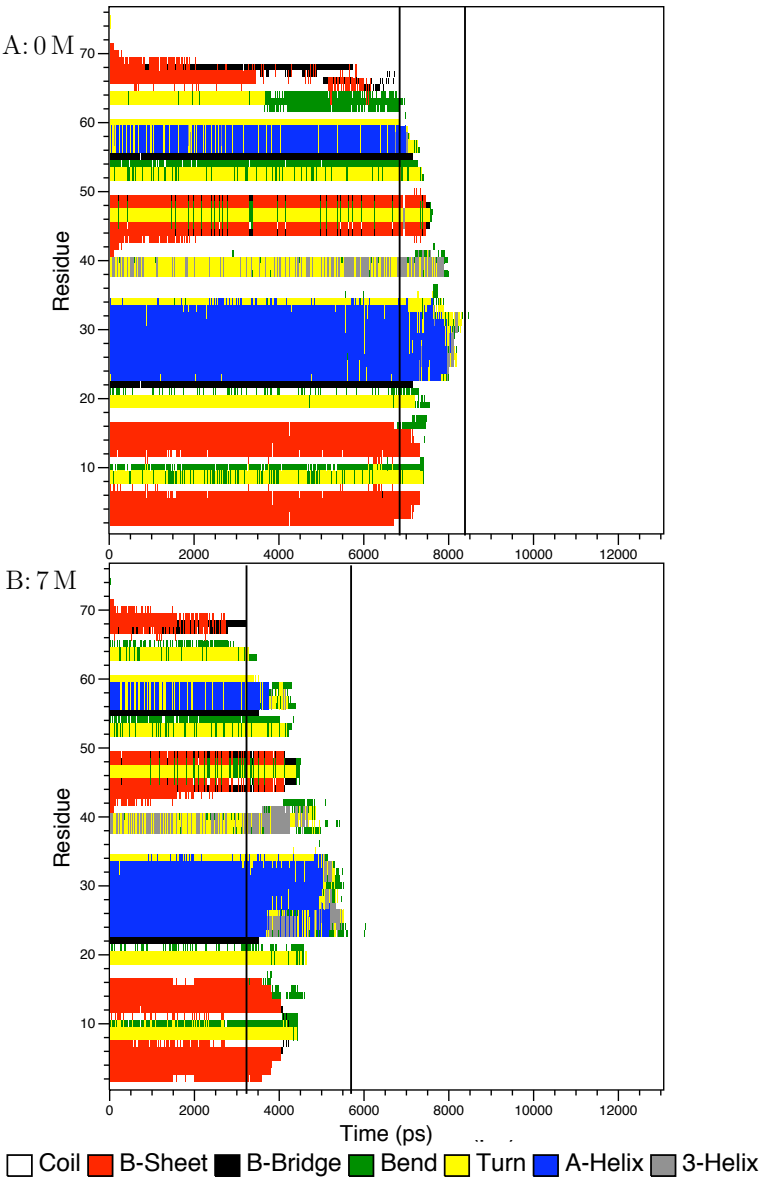



***Figure 3.14:*** Secondary structure maps derived from the FCMD simulations yielding an unfolding process. The left bar indicates the onset of unfolding, whereas the right bar depicts the complete loss of the secondary structure. The distance between the bars is smaller in the 0 M urea solution revealing a faster unfolding process compared to the 7 M environment (B).

tions (fig. 3.14), display an analogue influence of urea in respect to the FPMD simulations. The distance between the two black bars, reflects the duration for the decay of the cluster after the partioning of the tertiary structure. Comparing the distances in the 0 M (fig. 3.14, A) and 7 M (fig. 3.14 B) environment also reveals hindering of the unfolding process, due to a presence of urea. However, the proceeding in 7 M environment starts earlier than in pure water.

### 3.4.3 Discussion

The FCMD and FPMD simulations reveal a slower decay of secondary structure, if the solvent contains urea. This finding is somewhat unexpected. As mentioned in section 3.2 kinetic factors dominate interaction of the protein and the solvent. Hence, the destabilizing mechanism of the denaturant urea [8, 7] might not affect the interaction between the protein and the solvent. However, FPMD simulations in high urea concentrations reveal a faster loss of the $\beta$-motifs than structures simulated with low urea concentrations, which might hint, that the effect of solvent-water interaction is outbalanced by the denaturating effect of urea. With a possible onset of aggregation the viscosity of the liquid does not alter, but some urea molecules are able to interact directly with the protein. This effect might allow rupture forces and unfolding velocities, which compare to the values for a 0 M water solution and obeys the denaturant characteristics of urea despite an increase in viscosity of the solvent due to an increasing urea concentrations [33]. However this effect could also arise simply from the selection of the system, as also a large scatter in the forces is observed. Thus more averaging is required to obtain reliable information.

## 3.5 Hydrogen Bond Energies

As ubiquitin has a very large amount of secondary structure motifs, which stabilize the tertiary structure by hydrogen bonds between them, we compared the hydrogen bond energies for the different solvent environments in the simulations. To this end the time course of the mean energy per hydrogen bond between the different components of the modelled system was averaged over the three initial structures, simulated in the different solutions. This way the influence, arising from a change in the urea concentration, may be evaluated. A comparison of low and high pulling velocities does not reveal significant differences in the absolute hydrogen bond energy values of the several solvent environments (data not shown). Hence, we restrict the description of the results to a detailled dissection of the FPMD simulations, performed with $v_{\mathrm{pull}} = 1\,\frac{\mathrm{m}}{\mathrm{s}}$ and the FCMD simulations, where averaging was only applied to folded structures and not to the unfolded structures (one in 0 M and one in 7 M).

### 3.5.1 Analysis of the interaction between solvent molecules

Examinations of the hydrogen bond energies between the water molecules allow elucidating a possible disturbance of the hydrogen bonding pattern of water due to the presence of urea, which may destabilize the structure of the protein. Figure 3.15 depicts the mean values of the hydrogen bond energy in the FPMD and FCMD simulations. Obviously, the strength of a hydrogen bond between water molecules increases with increasing urea concentration, highlighted by the clearly separated curves (fig. 3.15 A). In the FCMD simulations (fig. 3.15 B), nearly equal values as in FPMD simulations arise for the energies and also the magnitude of the fluctuations is at the same level.

Gaining information about the interaction of the urea molecules with each other may help in understanding the influence of urea on unfolding. Hence, we turn to the energies of the hydrogen bonding pattern between urea molecules. In comparison to

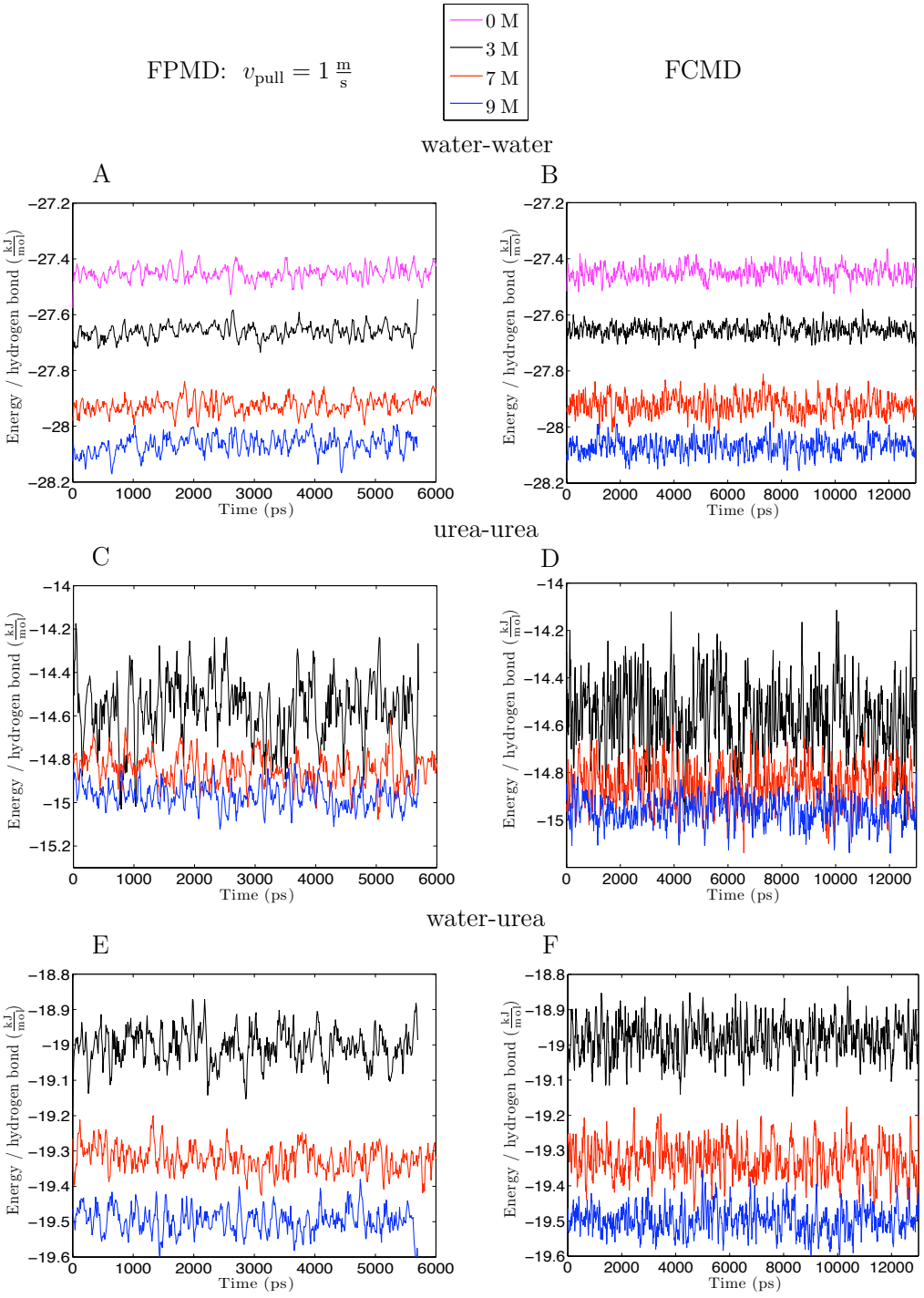



***Figure 3.15:*** Hydrogen bond energies between the solvent components. The results were averaged over the starting structure (n=2 for FCMD, 0 M and 7 M; otherwise n=3).

the hydrogen bond energies between the water molecules, the bond energy between urea molecules is about -14.5 $\frac{\text{kJ}}{\text{mol}}$ in the FPMD and FCMD simulations (fig. 3.15 C and D, respectively), which is only half of that between water molecules. However, the difference in energy between the 3 M, 7 M, and 9 M solvent environment decreases, while the order of energies in respect to the urea concentration remains. Further distinction arises in the magnitude of the fluctuation of the hydrogen bond energies, compared to the fluctuations of the energies between the water molecules. Increasing the urea concentration leads to a decrease of the magnitude of the fluctuations of the curves, an effect that is detectable in the FPMD simulations as well as in the FCMD simulations.

It seems, that the urea concentration is affecting the hydrogen bond energies between the solvent molecules. Hence, further information about the energy between water and urea molecules is indispensable. The course of the energies in the different solvent environments does not differ greatly between the FPMD simulations (fig. 3.15 E) and the FCMD simulations (fig. 3.15 F). The values of the hydrogen bond energies between the water and urea molecules range in the middle of the energy values of bonds between water molecules or between urea molecules. Regarding the order of the values of the hydrogen bond energies, in respect to the molar concentration of urea, the pattern is analogous to the energies between congeneric molecules.

The presented results of the hydrogen bond energies between the different molecules in the solvent suggest, that urea strengthens the hydrogen bonds. The strongest interaction occurs between the water molecules. It increases with increasing urea concentration. This strong interaction is expected to decrease the entropy of the solvent due to a restricted mobility of all molecules in it, yielding a decrease of its available configurational space.

### 3.5.2 Analysis of the interaction within the protein and between the protein and the solvent molecules

It is interesting to observe the time evolution of the hydrogen bond energies in FPMD and FCMD simulations in order to relate it to protein unfolding (fig. 3.16). In FPMD simulations energies vary by an amount of about $4\,\frac{\mathrm{kJ}}{\mathrm{mol}}$ during unfolding (fig. 3.16 A), independent of the urea concentration and no tendency for a stabilization of hydrogen bonds in different urea concentrations is observable. In contrast, as expected energies do not change in case of the FCMD simulations (fig. 3.16 B), which do not unfold during the simulation time.

Due to the rupture of a great amount of hydrogen bonds during the unfolding process, the possibilities for the protein to form bonds with solvent molecules increase. Because they are the main solvent component, water molecules have the highest probability to form new hydrogen bonds with the protein. Hence, a discussion of the hydrogen bond energies between the protein and water, is expected to help understanding the denaturation process.

Figures 3.16 C and D depict the time course of the hydrogen bond energies between the protein and water, gained from the FPMD and FCMD simulations, respectively. Both plots reveal a difference in energy of about $1\,\frac{\mathrm{kJ}}{\mathrm{mol}}$ between the simulations in pure water and the simulations, containing urea. In the FCMD simulations the energy values remain constant throughout the simulation apart from small fluctuations. In contrast, the FPMD simulations show a slightly increasing course of the energy, particularly in the first half of the simulations. This increase of energy stops when the collective axis of the $\alpha$-helix and the turn gets lost, depicted in fig. 3.6 (event 5, $t \approx 2500\,\mathrm{ps}$). From this time on the protein is almost completely unfolded and all the hydrophobic residues are able to connect to the hydrogen bond network of the solvent. In the process the order of the energy values in respect to the urea concentration converges to the order of the energy values between the solvent components. Due to the relatively weak bonds

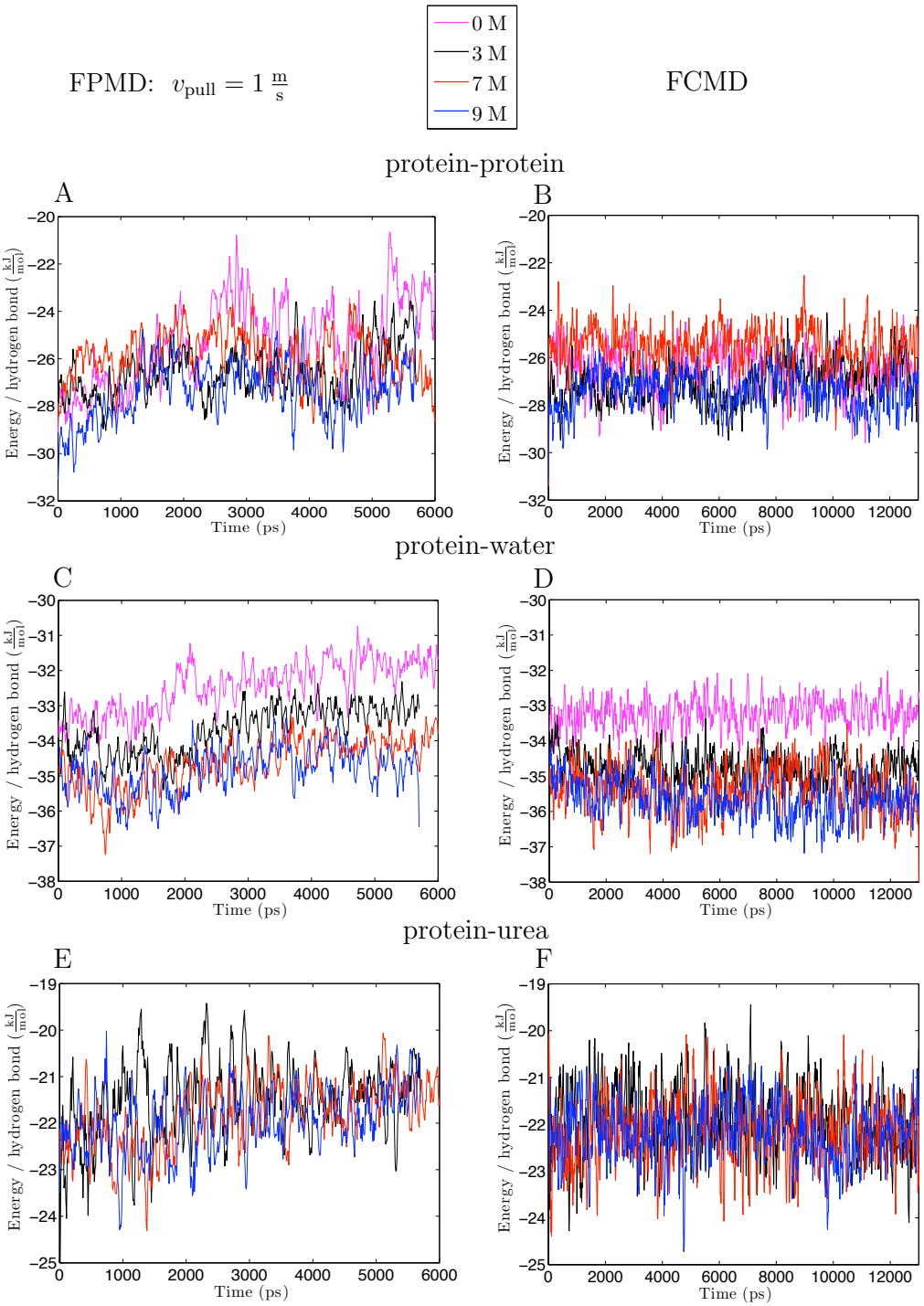



***Figure 3.16:*** Hydrogen bond energies within the protein and between protein and the solvent. Left column FPMD simulations, right column FCMD simulations.

of hydrophobic residues to water molecules, the mean energy per hydrogen bond is increased in respect to the initial structure. This effect provides the explanation for the increasing energy during the unfolding process, which is however not observed in fig. 3.16 D, because the plot represents those structures, that remain in the folded state throughout the FCMD simulation.

To complete the observation of the energy of the hydrogen bonds between the different components of our simulation system, we also examine the interaction between urea and the protein. Like all the other types of interaction discussed, the values of the hydrogen bond energy between urea and the protein are in the same order of magnitude in the FPMD (fig. 3.16 E) and FCMD (fig. 3.16 F) simulations and no differences due to a changing urea concentration arise. However, a tendency of increased energy per hydrogen bond in the course of unfolding is observable, because an increasing number of less polar residues is exposed to the solvent, forming weaker hydrogen bonds. Hence the mean energy bond increases.

The characterization of the interaction within the protein and between the protein and the solvent by the hydrogen bond energies, reveals the influence of urea concentration. The results of the hydrogen bond energies between the protein and urea molecules do not reveal significant differences due to a changing urea concentration. In contrast, for both FPMD and FCMD simulations, interactions between the protein and water show a visible effect of urea on the stability of the hydrogen bonds. Hydrogen bond energies between protein and water are lower in the presence of urea. However, in contrast to fig. 3.15 increasing urea concentrations did not lead to a clear gradient of decreasing bond energies.

Two possibilities exist for an explanation of the influence of urea on unfolding of a protein. Firstly, a direct interaction between the protein and urea might be a driving factor for the denaturation of the protein. However, a denaturating effect of urea based on direct interaction, as observed in [7], was not observed in our work. Secondly, urea

might affect unfolding via an indirect mechanism due to a distortion of the hydrogen bonding pattern of water as the dominant factor for the denaturative character of urea [8]. This distortion might lead to the increasing strenght of the hydrogen bonds between protein and water because the water molcules may favour a bonding to the protein instead to a bonding to urea. However, this interpretation is only speculative and does not explain why increasing urea concentration does not further stabilize the hydrogen bond significantly. As the values of the hydrogen bond energies between the protein and water are nearly equal in the FPMD and FCMD simulations (fig. 3.16 C and D, respectively), friction effects can be excluded, because they are absent in the FCMD simulations.

## 3.6 Solvent accessible surface area

Observations on the energies of the hydrogen bonds particularly yield information on the electrostatic effect of urea on ubiquitin. Therefore the course of the solvent accessible surface area (SASA) of the residues of the protein was analyzed, to obtain information about the influence of urea on the entropy of the system. For FPMD simulations (fig. 3.17), we averaged the results the simulations with different structures, while in FCMD simulations no averaging was carried out because only two structures unfolded. We distinguish between the SASA of the hydrophobic and hydrophilic residues to outline the influence of hydrophobic effects.

All FPMD simulations, independent of the pulling velocity, revealed the same characteristics of the SASA in the course of the simulation. Hence we restrict the discussion and analysis to the simulation with a pulling velocity of $1\,\frac{\mathrm{m}}{\mathrm{s}}$. No differences in the SASA of the hydrophobic residues (fig. 3.17 A), due to a changing urea concentration arise. In the first part of the unfolding procedure between $t = 0\,\mathrm{ps}$ and $t \approx 1800\,\mathrm{ps}$ no increase is observable. At $t \approx 1800\,\mathrm{ps}$ the decay of the core-shaping tertiary struc-

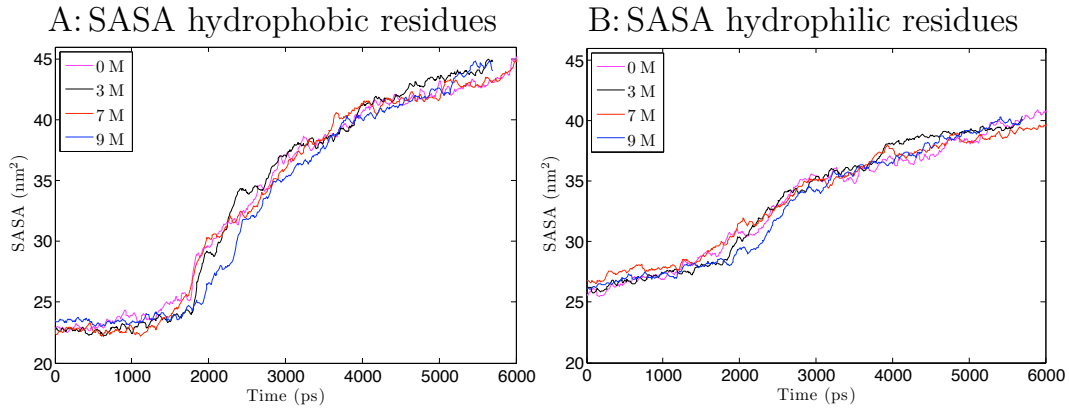



***Figure 3.17:*** Solvent accessible surface area (SASA) in an FPMD simulation ($v_{pull} = 1\,\frac{\mathrm{m}}{\mathrm{s}}$). Plot A and Plot B depict the different different characteristics of the hydrophobic and hydrophilic residues, respectively.

ture, consisting of the strands $\beta_3$, $\beta_4$, $\beta_5$, and the $\alpha$-helix, begins and the SASA of the hydrophobic residues is rapidly increasing. In the further course in time the strong increase of the SASA slows down towards the end of the simulations. In case of the hydrophilic residues (fig. 3.17 B), the increase of the SASA does not occur as fast as for the hydrophobic residues. The main difference to the hydrophobic residues arises in the scaling, due to the high fraction of hydrophobic residues in the amino acid chain of the protein ubiquitin in comparison to the hydrophilic residues.

In the folded state, the hydrophobic residues of the protein are not exposed to the solvent, minimizing the free energy of the system. During the enforced unfolding process the hydrophobic core of the protein breaks up leading to a fast increase of the SASA of the hydrophobic residues (fig. 3.6 event 4), due to the exposure of the residues in the hydrophobic core to the solvent. Finally, when nearly all tertiary structure is lost (fig. 3.6 event 5), the increase in SASA becomes slower.

Examination of the course of the SASA in the FCMD simulations reveals the difference of the SASA regarding the folded and unfolded state of the protein. The SASA

of the hydrophobic residues increases from 25 nm$^2$ to about 55 nm$^2$ in the 0 M as well as in the 7 M urea solution (fig. 3.18 A). However, after the complete elongation of the protein an obvious difference arises in the further course of the SASA in the two different solvent environments. In case of the 7 M environment, the SASA stays constant after complete elongation, whereas in the 0 M urea concentration the SASA decreases about 5 nm$^2$. The increase of the SASA of the hydrophilic residues (fig. 3.18 B) displays an equal character, but the increase in the SASA is only 20 nm$^2$. As expected and similar to the FPMD simulations, the course does not differ much from the SASA of the hydrophobic residues.

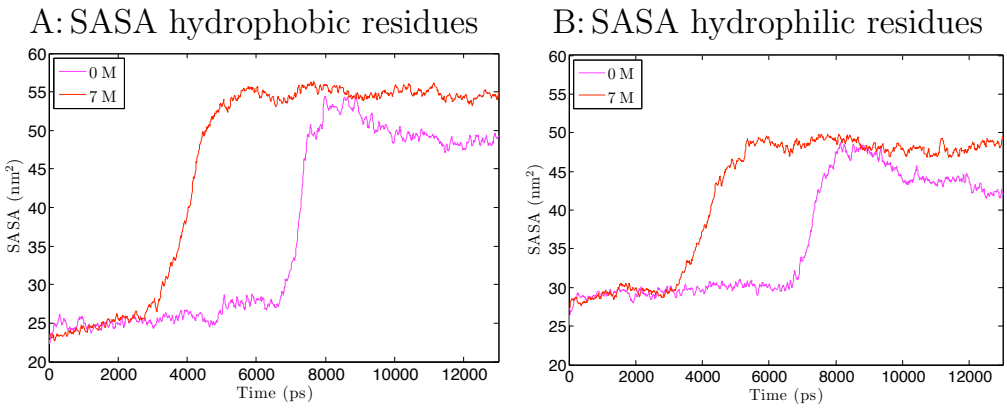



***Figure 3.18:*** Solvent accessible surface area (SASA) in the FCMD simulation, which yielded unfolding of the protein.

The sparse results of the FCMD simulations only allow a preliminary analysis of the influence of urea, but the characteristic decrease of the SASA in pure water after the gained maximum indicates the influence of urea on the interaction between protein and solvent. An explanation for the decrease is a partial refolding of the protein in water due to the high fraction of hydrophobic residues in the amino acid chain of the protein. Driven by the hydrophobic effect, the elongated protein attempts to minimize its SASA in respect to the water molecules, resulting in the observed decrease of the

SASA. The resulting slight difference in the end to end distance of the terminus $C_\alpha$ atoms of the protein for the unfolded state in a 0 M and 7 M solution, as depicted in fig. 3.7 effective from $t \approx 8000$ ps, corroborates this explanation. In case of a 7 M urea solution it seems that the hydrophobic effect is not as pronounced allowing a larger elongation of the protein, resulting in a larger SASA.

# 4 Summary and Conclusion

Motivated by recent simulation [1] and experimental [4, 5] studies on the protein ubiquitin applying AFM methods we performed several Molecular Dynamics simulations to explore the unfolding pathway of ubiquitin in different chemical environments. To this end we simulated two different AFM methods enforcing the unfolding of ubiquitin. First, we unfolded ubiquitin by pulling at the termini of the protein with potentials moving with constant velocity (FPMD) in opposite directions. Second, the termini were stressed with constant forces (FCMD) of 500 pN in opposite directions.

The FPMD simulations enable us to obtain an idea of the time course of the rupture forces. However, in FPMD simulations the energy landscape of $G$ is time dependent. With proceeding simulation time the energy landscape $G$ is strongly disturbed, in contrast to FCMD simulations, in which the shift of the perturbed $G$ remains constant throughout a simulation. Furthermore, results from FPMD simulations are strongly influenced by friction forces, which is not the case for FCMD simulations.

The influence of urea on the unfolding process is of interest, because urea is widely used as denaturant, but the interaction mechanism is not clear at all [7, 8]. To elucidate the influence of urea, four different concentrations, 0 M, 3 M, 7 M and 9 M solutions, in the setup of the simulations were compared.

The analysis of the results of the FPMD simulations with different initial structures revealed a very strong scattering of the rupture forces for the intial structures of the simulations. All starting structures differ only little in the RMSD of the $C_\alpha$-backbone,

but large differences in respect to the rupture forces in the FPMD simulations are observable (fig. 3.4). This finding corroborates the small variance of the time course of the forces if the same initial structures are simulated with different starting velocities of the solvent molecules (fig. 3.3). However, although the points in time of the several unfolding events vary as well as the corresponding rupture forces, all FPMD simulations even when performed with different structures, equal in respect to the order of rupture events, independent of the urea concentration.

The influence of urea on the velocity of the loss of the secondary structure motifs is somewhat unexpected; no direct relation between rupture forces and increasing urea concentration is observable. However, the presence of some urea molecules (3 M) seems to stabilize the protein, though it is actually know for its denaturant characteristics.

Yet in both cases the expected influence of urea might simply be hidden by the large scatter of the forces.

The extremely high forces in combination with only a slight increase with the changing of the pulling velocities from $1\,\frac{\mathrm{m}}{\mathrm{s}}$ to $10\,\frac{\mathrm{m}}{\mathrm{s}}$ reflects that the simulated system is strongly influenced by friction effects of the solvent during the unfolding process [33]. To analyse whether the unexpected influence of urea is an artefact, due to high friction occuring in FPMD simulations, we try to extrapolate the rupture force to the low friction regime. However, a reliable fit of the forces to the model resulting in eq. 2.18 and hence extrapolation to the low friction regime, governing AFM experiments is impossible due to the strong variance of the different simulations. Assuming a strong influence of hydrodynamic effects a speculative explanation for the absence of the denaturative effects of urea is possible. An increasing urea concentration leads to a non linear increase of the viscosity of the solvent [33], which might outbalance the denaturative effects of urea in a pulling simulation. However, at high urea concentrations, more urea molecules aggregate at the surface of the protein and may affect it in a denaturative way, such that the rupture forces decrease despite an increasing viscosity

of the solvent. But also in this case the effect could simply have a statistical character.

Furthermore, the strongly differing force profiles of the different starting structures corroborate that quantitative conclusions about the course of Gibb's free energy during unfolding are infeasible. A possible explanation might arise from a strong deformation of the Gibb's free energy landscape, due to the fast pulling. The possibility to overcome the energetic barriers for an unfolding event is strongly increased resulting in the differing points in time and forces for the unfolding events (fig. 3.4).

A different unfolding pathway of ubiquitin found in the FCMD simulations, compared to the FPMD simulations, reflects the influence of the two methods to apply a force on the termini of the protein on the Gibb's free energy landscape. Here, only two simulations yield an unfolding event and conclusions are only speculative. However, the observed unfolding events might reflect the denaturative character of urea in respect to the secondary structure of the protein. The unfolding of the protein in a solution containing urea starts much earlier than the observed unfolding event in pure water. Furthermore, the FCMD simulation in urea leading to unfolding of ubiquitin shows a larger elongation of the protein in comparison to the unfolding in pure water. These two effects might support our interpretation for the unexpected behaviour of the protein in the different urea solutions in the FPMD simulations, because in the FCMD simulations friction and hydrodynamic effects do not play a decisive role until the main unfolding event starts. In 7 M urea concentration the elongation proceeds with a velocity of several $\frac{\mathrm{m}}{\mathrm{s}}$. Hence the longer unfolding time of the structure in a 7 M urea solution in comparison to pure water might point to a possible influence of the viscosity of the solvent. However this effect could also depend on the mechanical stress and once again the statistics of the simulated process are too bad for a reliable conclusion.

The results of the energetic environment of the protein in terms of hydrogen bond energies and the solvent accessible surface area of the protein do not allow a clear conclusion about the interaction mechanism between urea, the protein, and water molecules.

As proposed by one model of the urea-water interaction [8] low urea concentrations strongly disturb the hydrogen bonding pattern of water, yielding a decrease in the solvent entropy, an effect that indirectly affects the stability of the protein. However, current MD studies on urea [7] suggest, that urea directly affects the protein. Our results show a stabilization of the bonds between protein and water in the presence of urea. With increasing urea concentration more urea molecules may replace the weak hydrogen bonds between water and the less polar residues. Hence, the average energy of the hydrogen bonds decreases. However, this effect decreases with increasing urea concentration, because most less polar residues are saturated already at low urea concentrations. This interpretation points again to a direct interaction mechanism of urea and the protein.

In summary, our studies show the high dynamics of the protein ubiquitin and a complex interaction of the protein with urea. The denaturative character of urea is not observed in the simulations, particularly low urea concentrations seem to have a stabilizing effect on the protein (high rupture force), but this effect might be explained with the help of the friction between the protein and the solvent due to fast pulling. Hence, an improved insight into the interaction mechanism between the solvent and the protein could be gained by much slower FPMD simulations, which minimize the hydrodynamic effects and elucidate the influence of urea on the Gibb's free energy landscape of the protein ubiquitin. Additional FCMD simulations, which last more than 20 ns or application of a higher force on the protein, could provide more unfolding events providing a possibility to refine the information on the energy landscape. Definitely many simulations with different starting structures have to be performed to get a deeper insight into the statistics of the unfolding process in the various chemical environments.

# Acknowledgements

Very many thanks to Prof. Dr. Hermann Gaub and Prof. Dr. Helmut Grubmüller for providing me the possibility to join Mr. Grubmüllers group at the department of Theoretical and Computational Biophysics of the Max-Planck Institute for Biophysical Chemistry in Göttingen and their continuous support in every aspect of this work. Thanks to all group members for their patience, when the cluster nodes were filled up with my simulation jobs. Special thanks to Frauke Gräter and Wolfram Stacklies for providing parts of the source code for the FCMD simulations. Furthermore special thanks to Ira Tremmel and Martin Stumpe for very generative discussions on this work. Additionally many thanks to my family, which supported me in all unscientific aspects during my whole studies. Finally I would like to thank my physics teacher Mr. Kühnel, for his encouraged school, that sparked my interest, resulting in the physics studies.

# Bibliography

[1] Frauke Gräter and Helmut Grubmüller. Fluctuations of primary ubiquitin folding intermediates in a force clamp. *Journal of Structural Biology*, 157:557–569, 2007.

[2] W. F. van Gunsteren and H. J. C. Berendsen. Computer simulation of molecular dynamics: Methodology, application, and perspectives in chemistry. *Angewandte Chemie*, 29:992–1023, 1990.

[3] W. F. van Gunsteren et. al. Biomolecular modeling: Goals, problems, perspectives. *Angewandte Chemie*, 45:4064–4092, 2006.

[4] Michael Schlierf, Hongbin Li, and Julio M. Fernandez. The unfolding kinetics of ubiquitin captured with single molecule force clamp techniques. *Proceedings of the National Academy of Sciences of the United States of America*, 101(19):7299–7304, May 2004.

[5] Julio M. Fernandez and Hongbin Li. Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*, 303:1674–1678, March 2004.

[6] Julia Wirmer, Wolfgang Peti, and Harald Schwalbe. Motional properties of unfolded ubiquitin: a model for a random coil protein. *Journal of Biomolecular NMR*, 35:175–186, 2006.

[7] Martin C. Stumpe and Helmut Grubmüller. Aqueous urea solutions: Structure,

energetics and urea aggregation. *Journal of Chemical Physics B*, 111:6220–6228, 2007.

[8] A. Idrissi. Molecular structure and dynamics of liquids: aqueous urea solutions. *Spectrochimica Acta Part A*, 61:1–17, 2005.

[9] Sequence details of ubiquitin 1UBQ. http://www.pdb.org.

[10] Chia-Lin Chyan, Fan-Chi Lin, Haibo Peng, Jian-Min Yuan, Chung-Hung Chang, Sheng-Hsien Lin, and Guoliang Yang. Reversible mechanical unfolding of single ubiquitin molecules. *Biophysical Journal*, 87:3995–4006, 2004.

[11] Senadhi Vijay-Kumar, Charles E. Bugg, and William J. Cook. Structure of ubiquitin refined at 1.8 Å resolution. *Journal of Molecular Biology*, 194:531–544, 1987.

[12] George A. Hagedorn and Alain Joye. Mathematical analysis of the Born-Oppenheimer approximations. http://www.math.vt.edu/people/hagedorn/, 1991.

[13] M. Born and M. Oppenheimer. Zur Quantentheorie der Molekeln. *Annalen der Physik*, 84(4):457–484, 1927.

[14] George A. Kaminski, Richard A. Friesner, Julia Tirado Rives, and William L. Jorgensen. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *Journal of Physical Chemistry*, B(105):6474–6487, February 2001.

[15] L. Verlet. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159(1):98–103, July 1967.

[16] Helmut Grubmüller. Force probe molecular dynamics simulations. *Methods in Molecular Biology: Protein - Ligand Interactions; Methods and Applications*, 305:493–515, 2005.

[17] G. I. Bell. Models for specific adhesion of cells to cells. *Science*, 200:618–627, 1978.

[18] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, VII(4):284–304, 1940.

[19] Wolfgang Walter. *Lehrbuch der Organischen Chemie*. S. Hirzel Verlag, Stuttgart, 21 edition, 1988.

[20] William Robert Fearon. The biochemisty of urea. *Physiological Reviews*, VI(3):399–439, July 1926.

[21] P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 64:253–287, 1921.

[22] R. W. Hockney and J. W. Eastwood. *Computer simulation using particles*. McGraw-Hill, New York, 1981.

[23] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J.R. Haak. Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics*, 81:3684–3690, October 1984.

[24] G. Vriend. What if: A molecular modeling and drug design program. *Journal of Molecular Graphics*, 8:52–56, 1990.

[25] Berk Hess, Henk Bekker, Herman J.C. Berendsen, and Johannes G. E. M. Fraaije. Lincs: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472, 1997.

[26] David van der Spoel, Erik Lindahl, Berk Hess, Aldert R. van Buuren, Emile Apol, Pieter J. Meulenhoff, D. Peter Tieleman, Alfons L.T.M Sijbers, K. Anton Feenstra, Rudi van Drunen, and Herman J.C. Berendsen. Gromacs user manual 3.3. http://www.gromacs.org, 2005.

[27] Matteo Frigo. A fast fourier transform compiler. In *Proceedings of the 1999 ACM SIGPLAN*, pages 1–12. MIT Laboratory for Computer Science, May 1999.

[28] MPI-2: Extensions to the Message-Passing interface. Message Passing Interface Forum, July 1997.

[29] W. Humphrey, A. Dalke, and K. Schulten. VMD - Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14(1):33–38, 1996.

[30] Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22:2577–2637, 1983.

[31] E. Espinosa, E. Molins, and C. Lecomte. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chemical Physics Letters*, 285:170–173, 1998.

[32] Frank Eisenhaber, Philip Lijnzaad, Patrick Argos, Chris Sander, and Michael Scharf. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *Journal of Computational Chemistry*, 16(3):273–284, 1995.

[33] Kazuo Kawahara and Charles Tanford. Viscosity and density of aqueous solutions of urea and guanidine hydrochloride. *Journal of Biological Chemistry*, 241(13):3328–3232, 1966.

Erklärung zur Diplomarbeit

Hiermit versichere ich, die vorliegende Diplomarbeit selbständig angefertigt und keine anderen als die angegeben Quellen und Hilfsmittel benutzt zu haben.

München, den 9.7.2007

(Florian Dommert)