Molecular Dynamics Study of the Mechanical Properties of Foldamers and the Metal Specificity of an Isatin Hydrolase

Dissertation

zur Erlangung des Grades
"Doktor der Naturwissenschaften"
im Promotionsfach Chemie

am Fachbereich Chemie, Pharmazie und Geowissenschaften
der Johannes Gutenberg-Universität
in Mainz

von

Lalita Shaki Uribe Ordoñez
geboren in Bucaramanga, Kolumbien

Mainz, 2016

Dekan:
1. Berichterstatter:
2. Berichterstatter:

Datum der mündlichen Prüfung: November 24th 2016

Die vorliegende Arbeit wurde in der Zeit von Oktober 2013 bis August 2016 am Institut für Physikalische Chemie der Johannes Gutenberg-Universität Mainz angefertig.

Molecular Dynamics Study of the Mechanical Properties of Foldamers and the Metal Specificity of an Isatin Hydrolase

Lalita Shaki Uribe Ordoñez

Johannes Gutenberg University of Mainz

Department of Chemistry
Institut for Physical Chemistry
Theoretical Chemistry Group

Graduate School of Excellence
Materials Science in Mainz

Mainz, 2016

Lalita Shaki Uribe Ordoñez
Molecular Dynamics Study of the Mechanical Properties
of Foldamers and the Metal Specificity of an Isatin Hydrolase
Johannes Gutenberg University of Mainz
Theoretical Chemistry Group
Institut for Physical Chemistry
Department of Chemistry
Duesbergweg 10-14
55128 Mainz

Contents

List of Abbreviations

AFM   Atomic force microscopy

AHDC   Average H-bond distance vs extension curves

ANC   Average fraction of native contacts vs extension curves

CN   Coordination number

CV   Collective variable

DFT   Density functional theory

DS   Dynamic strength

FE   Force vs extension

FF   Force field

FPMD   Force-probe molecular dynamics

H-bond   Hydrogen bond

HF   Hartree Fock

IH   Isatin hydrolase

MD   Molecular dynamics

NMR   Nuclear magnetic resonance

PMF   Potential of mean force

QM   Quantum mechanics

QM/MM Quantum-/molecular mechanics

RMSD   Root-mean-square deviation

SMFS   Single-molecule force spectroscopy

US   Umbrella sampling

WHAM Weighted histogram analysis method

Introduction

Folded polymers are widely found in nature, because they take part in all kind of vital processes such as catalysis, information storage, cellular signaling, molecular transport, and molecular recognition [1, 2]. Most biopolymers exhibit specific chemical functions thanks to their ability to produce well-defined folded structures. A very limited set of building blocks, 22 amino acids in proteins and four nucleic bases in DNA, is used by nature to control the conformation of its molecular machinery. This limited set of building blocks does not only respond to its suitability, but also complies with evolutionary constraints. In contrast, chemists count with a virtually unlimited set of building blocks. Thus, they are able to synthesize molecules that escape the evolutionary pressure and that can potentially mimic the properties of biopolymers and/or exhibit new features.

In general, non-natural folding oligomers are called foldamers. Several types of foldamers' backbones, *e.g.*, β-peptides, γ-peptides, aromatic foldamers, aminoxy acid oligomers, aza-peptides, aliphatic oligoureas, and foldamers based on heterocyclic skeletons have been successfully synthesized and proven to produce a large spectrum of folded motifs [3–7], *e.g.*, helices, linear strands, turns, and sheets. The large variety of available foldamers' backbones shows that folding may be governed by many different internal and external parameters. However, most foldamers have a certain degree of backbone rigidity in common that decreases the entropic cost of adopting an organized conformation [8]. Some internal factors governing foldamers' folding are: the shape of the molecule, its ability to establish attractive or repulsive intra-molecular non-covalent interactions and to form intra-molecular H-bonds. On the other hand, external factors may include solvent effects (such as hydrophobic effects), aggregation, host-guest complexation, and contacts with interfaces.

The large set of parameters which may rule the folding process makes the prediction of the folded structure that a given foldamer could adopt difficult. Even though, several studies have been carried out aiming at designing foldamers with predictable folded conformations (*e.g.*, see references [1, 9–11]). This effort reflects that the possibility of predicting the folded structure of a given backbone is a key element in foldamers' design.

The folding process of biopolymers and foldamers is still not fully understood. Moreover, how their folding pathway is shaped by their backbone architecture is also an open question [12, 13]. In protein research, nuclear magnetic resonance (NMR) and X-ray crystallography are very important tools to determine their structure and to study their folded conformations experimentally. In addition, recently, experimental setups aiming at studying single molecules have gained increasing importance [14].

Beside the dynamics of the unfolding of biomolecules chemically or thermally induced, those under the influence of external mechanical forces are relevant for the comprehension of several biological processes [15, 16]. In the last decades, single-molecule force

spectroscopy (SMFS) techniques, such as atomic force microscopy (AFM), have become standard tools to reveal detailed information of the mechanical unfolding of biopolymers such as DNA, RNA, proteins, and molecular assemblies [17–21]. The results obtained from SMFS experiments allow to determine the kinetic rates for all relevant transitions and to unravel the unfolding and folding pathways of a system. Moreover, the mechanical unfolding process of biopolymers has also been studied, with atomic resolution, using steered or force-probe molecular dynamics (FPMD) simulations [22–24]. In FPMD simulations, the application of an external force allows to mechanically unfold molecules and to study their unfolding pathway in detail, with a computational cost lower than that of molecular dynamics (MD) simulations in the absent of an external force.

There are different types of FPMD simulations similar to the types of methods used in single-molecule AFM [25]. In the so-called force-ramp mode, the external force is increased linearly in time and the unfolding or folding is observed at a certain molecular extension. The analysis of the distributions of unfolding forces, obtained in this mode, allows the determination of the unperturbed kinetic rates of the unfolding and folding processes. A more direct method for the determination of the rates is possible in the force-clamp mode. Here, a constant external force is applied and the transitions among the folded and unfolded conformations are monitored directly. In this mode, the kinetic rates are deduced from the waiting time distributions.

Analogously to protein research, NMR and X-ray crystallography are routinely used for determining the folded conformation of foldamers. Besides the experimental studies in which the chemical synthesis and the physical characterization are carried out, a number of theoretical investigations have been undertaken in the past, *e.g.*, the folding-unfolding transition of a class of foldamers, called β-peptides, and their thermodynamics have been studied by MD simulations [8, 26–29]. However, FPMD studies have been scarcely conducted in the area of foldamer research.

The goal of this thesis is to study the relation between the folding process of foldamers and their backbone architecture, through the investigation of their mechanical unfolding pathway, using molecular dynamics (MD) and FPMD simulations. Chapter 2 provides the theoretical foundation of this thesis. There, the basics of MD, FPMD, and quantum mechanics/molecular mechanics (QM/MM) simulations are explained together with the used data analysis methods. Chapter 3 comprises a detailed study of the mechanical unfolding pathway of a model β-peptide. It includes a discussion on how to choose the right force field to study non-natural oligomers and a demonstration of the information that can be extracted from FPMD simulations. In Chapter 4, the unfolding pathway of α- and β-peptides is compared at non-equilibrium and quasi-equilibrium conditions. The chain length dependence of the mechanical unfolding pathway of β-peptides is discussed in Chapter 5. Finally, in Chapter 6, the main parameters governing the folding process of foldamers are revealed through the comparison of the mechanical unfolding pathway of five oligomers with different backbone architectures.

Chapter 7 covers the work done as part of a collaboration with the group of Prof. Cascella of the Centre for Theoretical and Computational Chemistry at the University of Oslo. This work consists in the study of the high metal specificity of a class of metalloenzymes. Although this project does not concern the mechanical properties of foldamers, it is based on the same theoretical foundations. Two types of MD simulations were carried out for

this project, hybrid QM/MM simulations and QM/MM coupled with metadynamics simulations.

Metalloenzymes are enzymes with metal ion/s as cofactors. Metal ions act in different ways in enzymatic catalytic reactions, *e.g.*, as structural regulators, electron donors or acceptors or Lewis acids [30]. Because of these many roles, the activity of metal ions in metalloenzymes is difficult to understand and predict. In particular, the metal specificity of these enzymes is still poorly understood, because of their diverse behavior. The catalytic activity of some metalloenzymes does not depend on the metal ion acting as cofactor, since they show similar activity with a broad range of ions, even with ions not biologically available [31]. In contrast, some other metalloenzymes are highly specific for a metal ion and show no activity when bound to an ion different from their natural one [32]. In addition, an intermediate behavior has also been observed, some enzymes show similar activity with a group of ions, but are inhibited by others [33]. Comprehending the causes for the diverse metal specificity exhibited by metalloenzymes is crucial, because it allows to understand the role of metal ions in enzymatic catalysis and facilitates the design of artificial metalloenzymes.

In this work, the high specificity for manganese of isatin hydrolases (IH) from *Labrenzia aggregata* [34] is investigated using QM/MM simulations. An analysis of the binding site conformation adopted by IH with different metal ions reveals the importance of the binding site conformation for the catalytic activity. In addition, the mechanism of the hydrolysis of isatin in IH is revealed using QM/MM metadynamics simulations. The found hydrolysis mechanism differs from those of similar metallohydrolases [35] and explains the high metal specificity of IH.

Theoretical basis

<div style="text-align: right; font-size: 2em;">2</div>

The basis of molecular dynamics (MD) simulations is introduced in this chapter. The methodologies and techniques for standard MD simulations are described in sections 2.1-2.4. MD is a well established technique. Thus, here only the relevant information on MD is given, while a more comprehensive overview can be found in standard text books (for example see Reference [36]). In Section 2.5, a special type of MD simulations, used in this work, called force-probe molecular dynamics (FPMD) is described. In Section 2.6, two stochastic kinetic models for the conversion of rupture force distributions into kinetic rates are introduced. The concept of potential of mean force (PMF) and a method for its calculation are explained in the Section 2.7. The definition of the fraction of native contacts is introduced in Section 2.8. Finally, in the last two sections, 2.9 and 2.10, short descriptions of the quantum mechanics/molecular mechanics (QM/MM) and metadynamics methods are given.

2.1 Molecular dynamics

Molecular dynamics is a technique for computer simulations of ensembles of atoms and molecules, here referred to as particles in general. The aim of MD simulations is to understand the properties of assemblies of particles by studying their structure and interactions. In standard MD simulations, the interaction between particles and their movement is studied for a given period of time, through the numerical integration of Newton's equations of motion (see Section 2.3). The interaction, *i.e.*, the potential, between particles is defined by a set of mathematical functions and parameters called force field (FF) (see Section 2.2). A MD simulation can be carried out in different statistical ensembles, in which the number of particles, volume, energy, pressure, and/or temperature are kept constant (see Section 2.4).

A standard MD simulation consists of the following steps:

1. The target system is placed in a simulation box. This box must have the right shape and size to include all the particles involved in the simulation.

2. The solvent and ions (if required) are added to the simulation box, when the simulation is not carried out in vacuum.

3. A FF is chosen and the FF parameters for every particle in the system are defined. This set of parameters is called topology.

4. The system is equilibrated to obtain the desired thermodynamic ensemble.

5. The system's equations of motion are numerically integrated for a defined period of time and with a given time step.

MD simulations, contrary to most experiments, allow to obtain an atomistic description of a phenomenon. As a consequence, MD simulations are a powerful tool to understand different types of chemical and physical phenomena at a molecular level, as well as to explain experimental results and/or to validate theoretical models. Nowadays, the main limiting factor in a MD simulation is the computational cost. Unfortunately, for some systems the computational cost is still too high, because of their size or their slow dynamics. However, different methods for dealing with the size (such as coarse-grained [37, 38] or adaptive resolution MD [39]) or the slow dynamics (such as force probe MD (see section 2.5) or metadynamics (see section 2.10)) have been and are being developed.

2.2 Force fields

To introduce the concept of a force field, it is useful to analyze the requirements for the integration of the equations of motion of a system first. The main requirement is the definition of a Hamiltonian H comprising all energy contributions from the system's degrees of freedom. Such a Hamiltonian can be divided into a kinetic and a potential energy part,

$$H(\mathbf{p}, \mathbf{r}; m, s) = T(\mathbf{p}; m) + V(\mathbf{r}; s). \tag{2.1}$$

The kinetic energy T is a function of the momenta $\mathbf{p}$ and the masses m of all atoms in the system, and is normally independent of the atomic coordinates r,

$$T(\mathbf{p}; m) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} = \sum_{i=1}^{N} \frac{1}{2} m_i \mathbf{v}_i^2, \tag{2.2}$$

where N is the total number of atoms, $\mathbf{v}_i$ is the velocity of atom i, and $\mathbf{p}_i \equiv m_i \mathbf{v}_i$.

The potential energy V describes the interactions between particles as a function of the atomic coordinates $\mathbf{r}$ and the FF parameters s, and is usually expressed as a sum over bonded and non-bonded interactions,

$$V(\mathbf{r}; s) = V_{\text{bon}}(\mathbf{r}; s) + V_{\text{nbon}}(\mathbf{r}; s). \tag{2.3}$$

A FF consist of the definition of the bonded and non-bonded terms together with the parameters these terms depend on. In other words, a FF is a set of mathematical functions and parameters that describe the potential energy of an ensemble of particles.

2.2.1 GROMOS53A6 force field

The GROMOS53A6 FF [40] (GROMOS53A6FF) is a biomolecular FF whose parameterization is based on reproducing the free enthalpies of hydration and apolar solvation of a set of small polar molecules and amino acids. Since this is the main FF used in this work, its functional form is explained in detail.
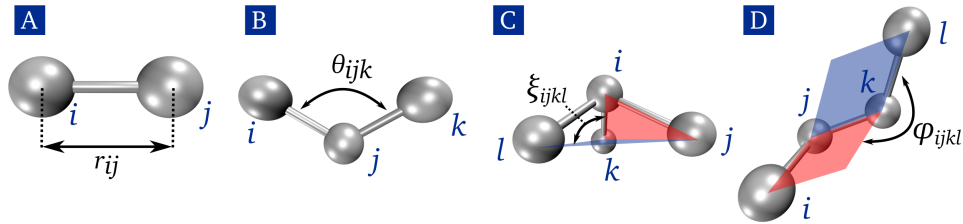
Bonded interactions

In the GROMOS53A6FF the bonded interactions V_{bon} are defined as:

$$V_{\text{bon}}(\mathbf{r};s) = V_{\text{bond}}(\mathbf{r};s) + V_{\text{angle}}(\mathbf{r};s) + V_{\text{har}}(\mathbf{r};s) + V_{\text{trig}}(\mathbf{r};s), \qquad (2.4)$$

where $V_{\text{bond}}(\mathbf{r};s)$, $V_{\text{angle}}(\mathbf{r};s)$, $V_{\text{har}}(\mathbf{r};s)$, and $V_{\text{trig}}(\mathbf{r};s)$ are the bond, angle, improper dihedral angle, and trigonometric (torsional) dihedral angle terms, as shown in Figure 2.1.

Fig. 2.1.: Schematic representation of the: **A.** bond interaction, **B.** angle interaction, **C.** improper dihedral interaction, and **D.** trigonometric dihedral interaction. The red planes are defined by atoms i, j, and k. The blue planes are defined by atoms j, k, and l.

The covalent bond interaction is defined by the anharmonic potential:

$$V_{\text{bond}}(\mathbf{r};s) = V_{\text{bond}}(\mathbf{r};K_b,b_0) = \sum_{n=1}^{N_b} \frac{1}{4} K_{b_n} [b_n^2 - b_{0_n}^2]^2. \qquad (2.5)$$

Here, $b_n \equiv r_{ij}$ is the bond length between atoms i and j (see Figure 2.1.A). The equilibrium bond length is b_0. The anharmonic bond constant K_b relates to the harmonic force constant by $K_{b_n}^{\text{harm}} = 2K_{b_n} b_{0_n}^2$. The sum goes over the total number of bonds N_b.

The covalent bond-angle interactions are calculated as the sum over N_θ bond angles,

$$V_{\text{angle}}(\mathbf{r};s) = V_{\text{angle}}(\mathbf{r};K_\theta,\theta_0) = \sum_{n=1}^{N_\theta} \frac{1}{2} K_{\theta_n} [\cos\theta_n - \cos\theta_{0_n}]^2. \qquad (2.6)$$

Here, K_θ and θ_0 are the angular force constant and the equilibrium angle, respectively. The nth angle θ_n is defined for the atoms i, j and k (see Figure 2.1.B) as

$$\cos\theta_{ijk} = \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}}{r_{ij}r_{kj}} \qquad (2.7)$$

The improper dihedral-angle interactions $V_{\text{har}}(\mathbf{r};s)$ help to keep a set of four atoms in a given configuration, for example, to maintain a tetrahedral configuration of a sp^3 carbon atom or to keep four atoms of a ring in a plane. The $V_{\text{har}}(\mathbf{r})$ is defined as

$$V_{\text{har}}(\mathbf{r};s) = V_{\text{har}}(\mathbf{r};K_\xi,\xi_0) = \sum_{n=1}^{N_\xi} \frac{1}{2} K_{\xi_n} [\xi_n - \xi_{0_n}]^2, \qquad (2.8)$$

with ξ_n being the dihedral angle defined by atoms i, j, k, and l (see Figure 2.1.C), calculated as

$$\xi_n = \text{sign}(\xi_n)\arccos\left(\frac{\mathbf{r}_{mj}\cdot\mathbf{r}_{qk}}{r_{mj}r_{qk}}\right), \tag{2.9}$$

where $\mathbf{r}_{mj} \equiv \mathbf{r}_{ij}\times\mathbf{r}_{kj}$, $\mathbf{r}_{qk}\equiv\mathbf{r}_{kj}\times\mathbf{r}_{kl}$ with the indices m and q defined through the cross products. The sign of the angle ξ_n is given by $\text{sign}(\xi_n) = \text{sign}(\mathbf{r}_{ij}\cdot\mathbf{r}_{qk})$.

Finally, the torsional dihedral-angle interactions are described by the trigonometric function,

$$V_{\text{trig}}(\mathbf{r};s) = V_{\text{trig}}(\mathbf{r};K_\varphi,\delta,m) = \sum_{n=1}^{N_\varphi}K_{\varphi_n}[1+\cos(\delta_n)\cos(m_n\varphi_n)], \tag{2.10}$$

where δ_n is the phase shift, m_n is the multiplicity of the torsional dihedral angle, and φ_n is the dihedral angle defined for atoms i, j, k, and l (see Figure 2.1.D), and is calculated using Equation 2.9.

Non-bonded interactions

In the GROMOS53A6FF, the non-bonded interactions V_{nbon} are calculated over pairs of non-bonded atoms. Although in principle all possible pairs of atoms should be included, in practice only the pairs within a cut-off distance R_{rf} and that are not first or second covalently bound neighbors are included. The non-bonded interactions are divided into Van der Waals and electrostatic interactions,

$$V_{\text{nbon}}(\mathbf{r};s) = V_{\text{electrostatic}}(\mathbf{r};s) + V_{\text{vdW}}(\mathbf{r};s). \tag{2.11}$$

The van der Waals (V_{vdW}) interactions are calculated as the sum over all non-bonded atoms pairs ij using a Lennard-Jones 12/6 potential,

$$V_{\text{LJ}}(\mathbf{r};s) = V_{\text{LJ}}(\mathbf{r};C12,C6) = \sum_{\text{pairs } i,j}\left(\frac{C12_{ij}}{r_{ij}^{12}} - \frac{C6_{ij}}{r_{ij}^{6}}\right). \tag{2.12}$$

The values adopted by the parameters $C12_{ij}$ and $C6_{ij}$ depend on the type of atoms involved and are obtained from the $C12_{ii}$, $C6_{ii}$ and $C12_{jj}$, $C6_{jj}$ parameters using the geometric combination rules,

$$C12_{ij} = \sqrt{C12_{ii}\cdot C12_{jj}} \qquad C6_{ij} = \sqrt{C6_{ii}\cdot C6_{jj}}. \tag{2.13}$$

The electrostatic interaction $V_{\text{electrostatic}}$ consists of three contributions. The first one is a sum over the Coulomb potential between interacting pairs of atoms,

$$V^{\text{C}}(\mathbf{r};s) = V^{\text{C}}(\mathbf{r};q) = \sum_{\text{pairs } i,j}\frac{q_iq_j}{4\pi\epsilon_0\epsilon_1}\frac{1}{r_{ij}} \tag{2.14}$$

where ϵ_0 is the dielectric permittivity of vacuum, ϵ_1 the relative permittivity of the medium, q_i the charge of atom i, and r_{ij} the distance between atom i and atom j.

The second contribution is a reaction-field contribution that represents the interaction of the atom i with the field induced due to the presence of the atom j outside of the cut-off

distance R_{rf} by a continuous dielectric medium. The last term is called distance-independent reaction-field and is a constant contribution to the energy for every pair of atoms. As it is independent of r_{ij}, it does not contribute to the force field, but it guarantees that the electrostatic energy is zero for atoms that are at the cutoff distance, helping to reduce possible noise coming from the cutoff.

2.3 Integration methods

In a MD simulation, the trajectories are obtained through the integration of Newton's equations of motion, for which there are several methods. In this work the Leap Frog algorithm [41] that can be derived from the Verlet scheme [42] was used. In the Verlet scheme, the coordinate $r(t)$ of a particle is written in terms of its Taylor expansion around time t,

$$r(t+\Delta t) = r(t) + v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2 + \frac{\Delta t^3}{3!}\dddot{r} + \mathcal{O}(\Delta t^4) \tag{2.15}$$

or,

$$r(t-\Delta t) = r(t) - v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2 - \frac{\Delta t^3}{3!}\dddot{r} + \mathcal{O}(\Delta t^4), \tag{2.16}$$

where f is the force and v is the velocity. The sum of these two equations yields

$$r(t+\Delta t) + r(t-\Delta t) = 2r(t) + \frac{f(t)}{m}\Delta t^2 + \mathcal{O}(\Delta t^4) \tag{2.17}$$

or,

$$r(t+\Delta t) \approx 2r(t) - r(t-\Delta t) + \frac{f(t)}{m}\Delta t^2. \tag{2.18}$$

Using this expression, the position at time $t+\Delta t$ can be calculated, if the force at time t and the positions at times t and $t-\Delta t$ are known. In most cases, the force f at t is calculated using a force field (see Section 2.2) and $r(t-\Delta t)$ are known from the previous time step.

The Verlet algorithm does not use the velocity to calculate the new position. However, it is possible to derive the velocity from the trajectory using,

$$v(t) = \frac{r(t+\Delta t) - r(t-\Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^2). \tag{2.19}$$

This expression for the velocity is only accurate to the first order. The Verlet's based algorithms have several advantages in comparison to other integration methods: the angular momentum is conserved, the trajectories are time-reversible, and the phase space area is preserved.

In the Leap Frog algorithm [41] the velocities are defined at half-integer time steps,

$$v(t-\Delta t/2) \equiv \frac{r(t) - r(t-\Delta t)}{\Delta t} \tag{2.20}$$

and

$$v(t+\Delta t/2) \equiv \frac{r(t+\Delta t) - r(t)}{\Delta t}. \tag{2.21}$$

From the latter equation the new positions can be calculated from the old positions and velocities

$$r(t + \Delta t) = r(t) + \Delta t \, v(t + \Delta t/2). \qquad (2.22)$$

From Equation 2.18 of the Verlet algorithm the following expression for the new velocities is obtained

$$v(t + \Delta t/2) = v(t - \Delta t/2) + \Delta t \frac{f(t)}{m}. \qquad (2.23)$$

The Leap Frog algorithm produces identical trajectories as the Verlet algorithm. However, in the Leap Frog scheme the velocities and positions are not defined at the same time. Therefore, the kinetic and potential energy are not defined at the same time neither and the total energy cannot be computed directly in the Leap Frog scheme. The main reason for using the Leap Frog algorithm and not the Verlet algorithm is that in the former the velocities are accurate up to third order, while in the second, as already mentioned, the velocities are only accurate up to first order.

2.4 Statistical ensembles

MD simulations can be carried out using different statistical ensembles. The used ensemble is chosen based on the system and the properties that are being investigated, *e.g.*, biological systems are usually studied using the NPT ensemble, where the number of particles N, pressure P, and temperature T are kept constant, because these are the conditions at which most biological processes happen in nature.

If an integration algorithm as the Verlet scheme (see Section 2.3) is used, it is possible to assume that the calculation is run in the microcanonical NVE ensemble, where the number of particle N, volume V, and energy E are kept constant. This assumption is valid, if the integration algorithm guarantees that the time averages show the same behavior as the averages over the space of all the system's states (phase space).

To perform MD simulations at constant N, V, and temperature T (called canonical or NVT ensemble) or at constant N, T, and pressure P (called isothermal–isobaric or NPT ensemble), it is necessary to use methods to control the temperature and the pressure. A method used to control the temperature is called a thermostat and a method used to control the pressure is called a barostat. The calculations presented in this work were mainly carried out in the NPT ensemble, using a velocity re-scaling method with a stochastic term to control the temperature [43] and the Parrinello-Rahman method [44, 45] to control the pressure. Both methods will be explained in the following subsections.

2.4.1 Velocity re-scaling thermostat

For the control of the temperature in MD simulations the velocity re-scaling thermostat proposed by Bussi, Donadio, and Parrinello [43] was used. In the simplest formulation of a velocity re-scaling thermostat, the velocities of all particles are rescaled by the same factor to enforce the system to match the average kinetic energy $\overline{K}$ at the target temperature

$(\overline{K} = N_f/2\beta$, where $\beta = 1/k_b T$, k_b is the Boltzmann constant, and N_f is the number of degrees of freedom). The scaling factor is defined as,

$$\alpha = \sqrt{\frac{\overline{K}}{K}}. \tag{2.24}$$

Because all particle velocities are rescaled with the same factor, the constraints on the bond lengths and the motion of the center of mass are not affected. The main disadvantage of this method is that the sampled ensemble is not explicitly known. For this reason, Bussi, Donadio, and Parrinello [43] proposed a modification of the velocity re-scaling method that ensures a canonical ensemble. In their algorithm, the system is evolved for a single time-step using the equations of motions. In a second step, the kinetic energy is calculated and consecutively evolved for another time step using an auxiliary continuous stochastic dynamics that does not modify the canonical distribution.

### 2.4.2 Parrinello-Rahman barostat

In MD simulations, the system is defined within a simulation box that contains all particles. The size and shape of this box depends on the system. Most pressure coupling methods modify the volume of the simulation box to enforce the system to conform to the target pressure. In this work the Parrinello-Rahman method [44, 45] was used. In this method, the box vectors $\mathbf{b}$ are re-scaled to fit a target pressure. As a consequence, the equations of motion of the particles are modified. The new equations of motion are

$$\frac{d^2 \mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} - \mathbf{M} \frac{d\mathbf{r}_i}{dt},$$
$$\mathbf{M} = \mathbf{b}^{-1} \left[\mathbf{b} \frac{d\mathbf{b'}}{dt} + \frac{d\mathbf{b}}{dt} \mathbf{b'} \right] \mathbf{b'}^{-1}. \tag{2.25}$$

2.5 Force-probe molecular dynamics

Steered or force-probe molecular dynamics (FPMD) is a type of MD, which aims at simulating single-molecule force spectroscopy (SMFS) experiments. There are several techniques that can be used to perform SMFS, *e.g.*, atomic force microscopy (AFM), optical tweezers, magnetic tweezers, acoustic force spectroscopy. An example of an experimental setup for AFM is shown in Figure 2.2. In AFM, a single molecule is attached to a cantilever at one side and to a pizoelectric (PZT) scanner at the opposite side (see Figure 2.2.A). In this setup, the cantilever acts as a piconewton (pN) force sensor, allowing to measure the force applied to the molecule. In the force-ramp mode the distance between the scanner and the cantilever is increased with a constant velocity. As showing in Figure 2.2.B, at the beginning of the pulling experiment the forces increases and the cantilever bends until a rupture event occurs. After the rupture event, the force drops and the cantilever retracts giving rise to a rip/maximum in the force. By performing the experiment several hundreds of times, it is possible to obtain a histogram of rupture forces.

In FPMD, a SMFS experiment is simulated. One of the main advantages of FPMD, in comparison to experiment, is that FPMD provides atomic resolution of the pulling process. This
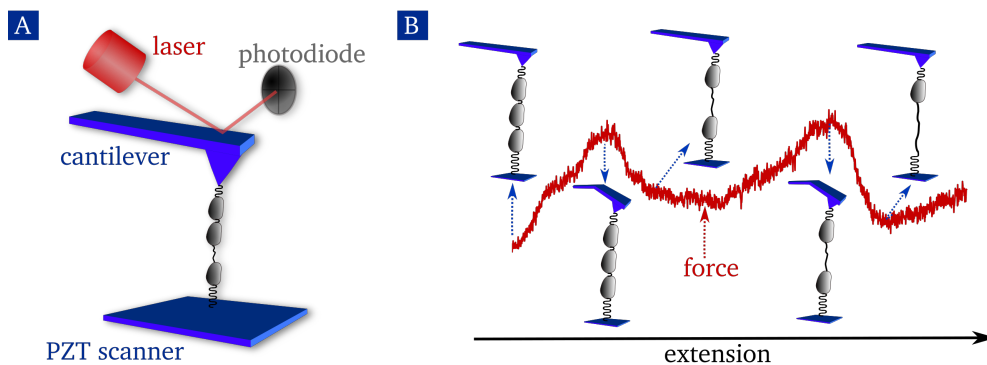
Fig. 2.2.: A. Schematic representation of an experimental setup for a single molecule AFM experiment. **B.** Example of a force curve obtained from AFM. Here extension refers to the distance between the cantilever and the PZT scanner. The blue arrows indicate the state of the system at the pointed force.

level of resolution allows to investigate the underlying atomistic dynamics and interactions that generate the measured forces. As in the experiment, in FPMD one side of the molecule is fixed (here called reference group) and the opposite side (here called pull group) is attached to a harmonic potential. The pull and the reference groups can be an atom or a group of atoms. The harmonic pulling potential (so called umbrella potential) is defined as

$$V_{\mathrm{umb}}(\mathbf{R}, t) = \frac{1}{2} K \left[(\mathbf{R} - \mathbf{R}^0) \cdot \hat{\mathbf{n}} - \mathbf{V} t \right]^2 , \tag{2.26}$$

where $\hat{\mathbf{n}}$ denotes the pulling direction, K is the harmonic force constant, and $\mathbf{V}$ is the pulling velocity. $\mathbf{R}^0$ is the initial position of the pull group's center of mass. The algorithm for FPMD is the same as for regular MD simulations, the only difference is that in FPMD, the umbrella potential is added to the potential calculated from the force field.

The optimal pulling velocity and harmonic constant to study the unfolding process of an oligomer depend on its size and the stability of its folded structure. In general, when a small velocity (*e.g.*, $V = 1 \times 10^{-3}$ nm/ns) and small harmonic constant (*e.g.*, $K = 100$ pN/nm) are used, the time-scale of the unfolding process is longer, but the resolution of the results increases. Therefore, a compromise between simulation time and resolution has to be found for each system. In this work, test simulations were performed with different parameters to find the optimal pulling velocity and harmonic constant for each system.

2.6 Stochastic kinetic models

By performing several FPMD simulations for a system it is possible to obtain rupture force distributions and from these distributions the kinetic rates of the associated rupture events. One way of transforming force distributions into kinetic rates consist in using a stochastic model of a diffusive barrier crossing in the presence of an external force. The simplest phenomenological model (Bell's model [46]) assumes that the bond rupture's rate follows an Arrhenius law and that the activation energy is decreased by the application of the external force. Under these assumptions Bell's formula writes

$$k(F) = k_0 \mathrm{e}^{F x^{\ddagger}} \tag{2.27}$$

where k_0 is the kinetic rate in the absence of external forces, F is the force and $x^{\ddagger}$ is the distance from the minimum to the energy barrier. Throughout this section the energy is expressed in $k_B T$ units, with k_B as the Boltzmann constant and T the absolute temperature.

More elaborated stochastic models apply Kramers theory [47] of diffusive barrier crossing to a simple model free-energy surface to obtain not only k_0 and $x^{\ddagger}$ but also the free energy of activation in the absence of external forces, $\Delta G^{\ddagger}$. In these models [48, 49], the soft spring approximation[1] is used and the escape rate k is found to be

$$k(F) = k(0)(\delta_F)^{1/\nu-1}e^{\beta \Delta G^{\ddagger}[1-(\delta_F)^{1/\nu}]},$$
$$\text{with } \delta_F = 1 + \frac{Kx^{\ddagger}}{2F_c} - \left(1 + \frac{(1-\nu)Kx^{\ddagger}}{2\nu F_c}\right)\frac{F}{F_c}, \tag{2.28}$$

where the critical force F_c at which the barrier vanishes is given by $F_c = \Delta G^{\ddagger}/(\nu x^{\ddagger})$. β is the inverse temperature ($\beta = 1/k_b T$) and ν is a parameter that defines the shape of the energy profile. Here, $\nu = 1/2$ for a harmonic well with a cusp-like barrier ($U_0(x) = \Delta G^{\ddagger}(x/x^{\ddagger})^2$ for $x < x^{\ddagger}$ and $-\infty$ for $x \geq x^{\ddagger}$) or $\nu = 2/3$ for a linear-cubic function ($U_0(x) = (3/2)\Delta G^{\ddagger}x/x^{\ddagger} - 2\Delta G^{\ddagger}(x/x^{\ddagger})^3$). In the limit $\Delta G^{\ddagger} \to \infty$, with $\nu = 1$, and in the soft spring limit $\frac{1}{2}K(x^{\ddagger})^2 \ll k_b T$, the previous model yields the phenomenological Bell model.

Equation 2.28 can be used to obtain $\Delta G^{\ddagger}$, $x^{\ddagger}$, and $k(0)$ of a rupture event by fitting $k(F)$ vs F curves. In the following sections, it is explained how these curves are obtained from FPMD simulations.

2.6.1 Lifetime method

From the models described in the previous section, the following relation for the lifetime τ in terms of the distribution of rupture forces $p(F)$ and the loading rate $\mu = K \times V$ is obtained [48]

$$\tau(F) = \int_F^{\infty} p(f)df/[\mu(F)p(F)]. \tag{2.29}$$

Setting in this equation $F = \langle F \rangle$ and approximating $p(F)$ by a normalized Gaussian distribution, an approximate relation between the lifetime of the folded state and the moments of the rupture force distributions is derived

$$\tau(\langle F \rangle) \simeq \sqrt{\left(\frac{\pi}{2}\right)(\langle F^2 \rangle - \langle F \rangle^2)}/\mu(\langle F \rangle). \tag{2.30}$$

Using this equation, the lifetimes $\tau(\langle F \rangle)$ are calculated from the rupture force distributions and the kinetic rates are obtained as the inverse of τ, $k(\langle F \rangle) = 1/\tau(\langle F \rangle)$. By calculating several rupture force distributions with different pulling parameters, *i.e.*, pulling velocity and harmonic constant, a $k(F)$ vs F curve is obtained.

---

[1]In the soft spring approximation the free-energy surface $U(x) = U_0(x) + K(x - Vt)^2/2$ along the pulling coordinate x is approximated as $U(x) \approx U_0(x) - F(t)x$. This is valid when $K << 2\Delta G^{\ddagger}/x^{\ddagger 2}$.

2.6.2 Zhang and Dudko method for dependent barriers

In the lifetime method, it is assumed that sequential transitions are independent. However, this assumption is not always valid. Therefore, in this work a second stochastic method, recently developed by Zhang and Dudko (ZD)[50] to avoid the assumption of independent barriers, was also used.

The ZD method is based on the fact that the probability $P_{ij}(t)dt$ that a molecule in state i undergoes a transition to state j between times t and $t + \delta t$ is proportional to the rate constant $k_{ij}(t)$ of this transition, the number of molecules $N_i(t)$ in state i, and the time interval dt

$$P_{ij}(t)dt = k_{ij}(t)N_i(t)dt, \qquad (2.31)$$

where $N_i(t)$ can be determined as the difference between the initial number N_i^0 of molecules in state i and the number of molecules that have left this state until time t. Considering that during the time interval $t + dt$ the applied force changes from F to $F + dF$, the number of the molecules that transit from i to j within the force interval $[F, F + dF]$ is

$$P_{ij}(F)|dF| = k_{ij}(F(t))N_i(t)dt. \qquad (2.32)$$

From this equation follows the relation between the $k_{ij}(F)$ and $P_{ij}(F)$

$$k_{ij}(F) = \mu_i \frac{P_{ij}(F)}{N_i(F)}, \qquad (2.33)$$

where $\mu_i = |\dot{F}(F)|_i$ is the loading rate in state i. A discretized version of this equation can be used to convert the rupture force histograms into kinetic rates

$$\text{rate}_{i \to j}(F) = \frac{|\text{loading rate in state } i \text{ at } F|}{\text{trajectories in state } i \text{ at } F} \times \frac{\text{counts in bin } F}{\text{bin width}}. \qquad (2.34)$$

2.7 Potential of mean force

The potential of mean force (PMF) is a potential energy profile, function of one or more defined reaction coordinates, obtained by integrating the mean force from an ensemble of configurations of a system. For example, the PMF can be used to study how the energy of a system changes as a function of the distance between two residues or how the energy changes when a system transits between two conformations. When the system can be assumed to be in equilibrium, the PMF approximates the free energy profile of a system as a function of the reaction coordinate. If the reaction coordinate is chosen properly, the PMF gives information about the shape of the free energy profile of a process, *i.e.*, the number and position of minima, and the number of energy barriers and their heights.

In this work, the PMFs of some of the studied oligomers were calculated to obtain information of their unfolding process under equilibrium conditions. There are different ways of calculating the PMF of a system such as thermodynamic integration and umbrella sampling (US). In this work, the weighted histogram analysis method (WHAM) combined with the umbrella sampling method [51] was used. Therefore, this method is described in the following sections.

2.7.1 The weighted histogram analysis method

The purpose of the weighted histogram analysis method (WHAM) [51] is to obtain an unbiased probability distribution $\rho_0(\eta)$ (where η is a reaction coordinate) from a series of simulations performed using a perturbing potential $W_i(\eta(\mathbf{R}))$. In these simulations, the potential energy U of the studied system is

$$U = U_0(\mathbf{R}) + W_i(\eta(\mathbf{R})), \tag{2.35}$$

where U_0 is the unperturbed potential of the system and $\mathbf{R}$ is the set of atomic coordinates. From such simulations, a set of biased probability distributions $\rho_i^{(b)}(\eta)$ is obtained. Each $\rho_i^{(b)}(\eta)$ is a normalized histogram of the values of η sampled during the simulation i. From $\rho_i^{(b)}(\eta)$ the unbiased probability distribution, $\rho_i^{(u)}(\eta)$, is defined as

$$\rho_i^{(u)}(\eta) = e^{\beta[W_i(\eta)-f_i]}\rho_i^{(b)}(\eta). \tag{2.36}$$

where f_i is the free energy related to adding $W_i(\eta(\mathbf{R}))$ to $U_0(\mathbf{R})$. f_i is defined as

$$e^{-\beta f_i} = \int d\eta \rho_0(\eta)e^{-\beta W_i(\eta)}. \tag{2.37}$$

The total probability distribution $\rho_0(\eta)$ can be obtained as a linear combination of the unbiased probabilities

$$\rho_0(\eta) = C\sum_{i=0}^{N} p_i(\eta)\rho_i^{(u)}(\eta), \tag{2.38}$$

where C is a normalization constant. The weights $\rho_i(\eta)$ are required to be normalized and are chosen so that the statistical error on the total probability distribution is minimized

$$\frac{\partial^2[\rho_0(\eta)]}{\partial \rho_i} = 0. \tag{2.39}$$

With this requirement, $\rho_0(\eta)$ takes the form

$$\rho_0(\eta) = C\sum_{i=1}^{N} \frac{n_i}{\sum_{j=1}^{N} n_j e^{\beta[W_j(\eta)-f_i]}} \rho_i^{(b)}(\eta). \tag{2.40}$$

Using the previous Equation and Equation 2.37, f_i can be calculated as

$$e^{-\beta f_k} = C\int d\eta \sum_{i=1}^{N} \frac{n_i e^{\beta W_k(\eta)}}{\sum_{j=1}^{N} n_j e^{\beta[W_j(\eta)-f_i]}} \rho_i^{(b)}(\eta). \tag{2.41}$$

As f_i appears on the left and on the right of this equation, it is possible to obtain f_i in an iterative self-consistent way.

2.7.2 The weighted histogram analysis method combined with umbrella sampling

As stated above, the weighted histogram analysis method (WHAM) combined with umbrella sampling (US) [51] was used to calculate the PMFs. In US, an umbrella potential is used to

restrain the system at a given value η_j of the reaction coordinate η. The reaction coordinate must be chosen such that it fully describes the transformation of the system from an initial to a final state of interest. In US, the perturbing potential is

$$W_{ij}(\mathbf{R}) = \lambda_i W_0(\mathbf{R}) + K_i(\eta(\mathbf{R}) - \eta_j)^2, \tag{2.42}$$

where

$$W_0(\mathbf{R}) = U_1(\mathbf{R}) - U_0(\mathbf{R}). \tag{2.43}$$

U_0 and U_1 are the potential energy of the initial and the final state. λ is a coupling parameter that describes the conversion of the system from the initial to the final state. Therefore, the first term in Equation 2.42 is the potential associated to the transformation of the system. The second term is the umbrella sampling potential.

By running several trajectories using an umbrella potential with different values of η, a series of histograms of η are obtained (see Figure 2.3.A) and transformed into the PMF using WHAM (see Figure 2.3.B).
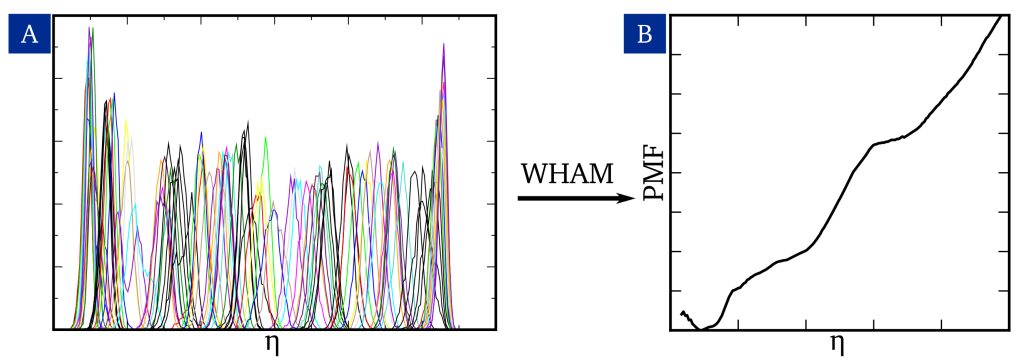
Fig. 2.3.: A. Histograms obtained from a series of umbrella sampling trajectories performed with different values of the reaction coordinate η. **B.** PMF obtained from the histograms in A using WHAM.

2.8 Fraction of native contacts

Protein folding theory assumes that only native contacts play a significant role in determining the folding mechanisms. This assumption is supported by the principle of minimal frustration, which states that evolution has designed the folding energy landscapes of proteins such that the correlation between energy and the nativeness of the structure is maximal [52]. In this context, the fraction of native contacts has been proposed as an order parameter for determining protein folding mechanisms [52].

In this work, the fraction of native contacts is used for the conformational analysis of MD and FPMD trajectories. For α-peptides, a native contact is a pair of heavy atoms that are separated in the native structure by a distance smaller than a threshold (in this work 0.35 nm) and that belong to amino acids more than four residues apart. For the non α-peptides oligomers, the same distance criterion is applied here. The required number of separating

residues was chosen to be equal to the number of residues per turn in each helix, two for β-peptoids, three for δ-peptides and β-peptides, and one for α/γ-peptides.

The fraction of native contacts is a measure of the number of native contacts present in any configuration and is defined as [52]

$$Q(X) = \frac{1}{N} \sum_{(i,j)} \frac{1}{1 + e^{[\gamma(r_{ij}(X) - \vartheta r_{ij}^0)]}}, \tag{2.44}$$

and takes values from 0 to 1, where 1 indicates a perfect native secondary structure, 0 a complete unfolded state, and intermediate values indicate some preservation of the native structure. The sum runs over the N pairs of contacts ij in the native structure, r_{ij} is the distance between atoms i and j in the configuration X, r_{ij}^0 is the same distance but in the native structure, γ is a smoothing parameter, here set to $5\,\mathring{A}^{-1}$, and ϑ is a factor that accounts for the fluctuations when a contact is formed and is set to 1.8 here.

2.9 Quantum-/molecular mechanics

The concept of quantum-/molecular mechanics (QM/MM) was introduced by Warshel and Levitt in 1976 [53]. They presented the first semi-empirical-QM/MM model and showed its application to the lysozyme enzymatic reaction. QM/MM is nowadays an established tool for modeling biomolecular processes, but it is also often used for investigating inorganic, organometallic, and solid state systems [54].

QM/MM is a hybrid molecular simulation method. This method combines the accuracy of quantum-mechanical techniques with the velocity of molecular mechanics (*i.e.*, MD) simulations. QM/MM simulations are used to study systems that due to their large size cannot be investigated with QM methods and that due to their complex electronic structure cannot be correctly described with classical mechanics, *e.g.*, a reaction happening in a solvent, the solvation of an ion, or the active site of an enzyme. In a routine QM/MM simulation, the system is separated into two regions, a QM region and a MM region (*e.g.* see Figure 2.4), and through the whole simulation this division is kept. The atoms in the QM region are treated using a QM method such as Hartree Fock (HF) or density functional theory (DFT), while the atoms in the MM region are described using classical mechanics as explained at the beginning of this chapter (see sections 2.1, 2.2). The whole system is evolved in time using classical equations of motion as described in sections 2.3 and 2.4.

To set up a QM/MM simulation three main issues have to be considered: i) how to choose the QM region, ii) how to describe the interaction between the QM and the MM regions, and iii) which QM and MM methods to use. There is no a unique answer to these questions since the answer depends on the size of the system and the properties to be studied. The choice of the QM region should follow a couple of rules: i) the region must include all atoms that require an explicit description of their electronic structure. In some cases it is advisable to perform calculations with differently sized regions. ii) If necessary, only single covalent bonds should be cut by the QM-MM division. iii) The region should be as small as possible to lower the computational costs of the simulations.
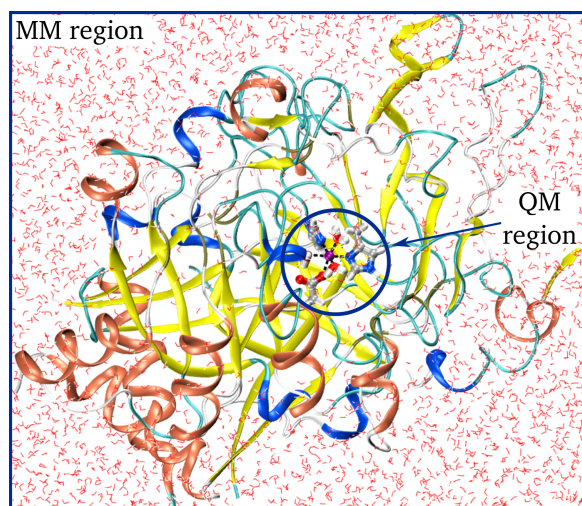
Fig. 2.4.: Example of the definition of the QM/MM region for the study of an enzymatic reaction. The red color lines represent water molecules, the enzyme is shown in cartoon, and the atoms in the QM region are represented by balls and sticks.

In most cases, it is not sufficient for calculating the total energy ($E_{QM/MM}$) to sum the QM energy (E_{QM}) and the MM energy (E_{MM}), but it is necessary to include an interaction term (E_{QM-MM}). There are several methods to calculated $E_{QM/MM}$ (for an overview see reference [55]). In this work, the so-called additive method is used and the $E_{QM/MM}$ is obtained as

$$E_{QM/MM} = E_{QM} + E_{MM} + E_{QM-MM}, \qquad (2.45)$$

E_{QM} is calculated using a QM method, *e.g.*, DFT, and E_{MM} is calculated using a FF (as described in Section 2.2). E_{QM-MM} is calculated as the sum of three terms

$$E_{QM-MM} = E^b_{QM-MM} + E^{vdW}_{QM-MM} + E^{el}_{QM-MM}. \qquad (2.46)$$

Here, E^b_{QM-MM}, E^{vdW}_{QM-MM}, and E^{el}_{QM-MM} are the bonded, the van der Waals, and the electrostatic interactions, respectively.

There are different methods for the calculation of E^{el}_{QM-MM}. The most commonly used one is the electrostatic or electronic embedding scheme [55]. In this scheme, the MM point charges are incorporated as one-electron terms in the QM Hamiltonian, by adding the following term

$$E^{el}_{QM-MM} = \sum_i^N \sum_j^L \frac{q_j}{r_i - R_j} + \sum_\alpha^M \sum_j^L \frac{q_j Q_\alpha}{R_\alpha - R_j}. \qquad (2.47)$$

Here, r_i is the position of electron i, q_j are the MM point charges located at R_j, and Q_α are the nuclear charges of the QM atoms at position R_α. The indices i, j, and α sum over the N electrons, L point charges, and M QM nuclei, respectively. This equation is given in atomic units.

The calculation of E^{vdW}_{QM-MM} is only performed at the MM level. Therefore, suitable FF parameters are also required for the atoms in the QM region. For the calculation of E^b_{QM-MM} there are also various possible methods [55]. In this work, the link atoms scheme was used. In this scheme, the free valency generated in the QM regime due to the cutting of covalent

bonds is filled by the addition of a monovalent atom, usually hydrogen, to electronically saturate the QM region. The added atoms are not considered in the MM calculation and the cut covalent bonds are described only at the MM level.

The choice of the QM method follows the same rules as for a regular QM calculation, *i.e.*, the size of the system, the computational time, and the required accuracy are considered. Regarding the MM part of the calculation, the main choice is which FF to be used.

2.10 Metadynamics

The concept of metadynamics was first introduced by Laio and Parrinello in 2002 [56], motivated by the limited time scales accessible in MD simulations, which is in the order of hundreds of nanoseconds for classical MD and in the range of picoseconds for QM/MM. However, most chemical phenomena such as reactions and structural transitions happen on time scales that are several orders of magnitude larger and are currently rare events in affordable simulation times. Metadynamics is an algorithm that allows to reconstruct the free energy landscape of a system and to decrease the time required to sample rare events at the classical or at the quantum level. In short, the idea is to fill the wells of the free energy surface (FES) of the process of interest with Gaussian hills.

The method assumes that the system can be described by a few collective variables. In the simulation, a positive Gaussian potential is added to the energy landscape of the system. By adding several Gaussian hills, the system is hindered from returning to its initial configuration and pushed to explore other regions of the energy landscape (see Figure 2.5.A). At some point of the simulation, the modified energy landscape becomes flat and the collective variables fluctuate rapidly (see Figure 2.5.B). The fast fluctuation indicates that the simulation has converged. At this point, the real energy landscape can be recovered as the complement of the sum of all the added Gaussian hills (see figures 2.5.C and D).
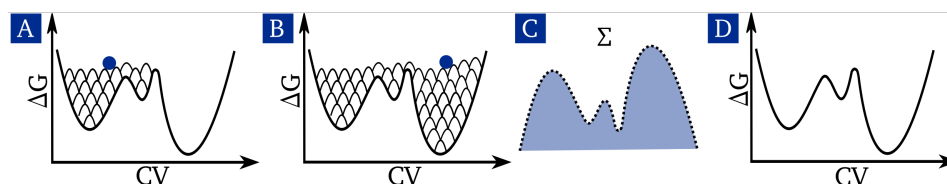
Fig. 2.5.: A. FES after some Gaussian hills has been added. B. FES after the metadynamics had converged. The blue circle indicates the position of the system. C. Sum of all added Gaussian hills. D. FES calculated as the complement of the sum of the Gaussian hills.

The simulation time and accuracy can be tuned by changing the time interval between the addition of two Gaussian hills and by modifying their height and width. For example, by increasing the size of the Gaussian a fast rough estimate of the energy landscape can be obtained.

The main advantage of metadynamics is that it does not require an initial estimate of the energy landscape. However, the right choice of the collective variables can be difficult, since

it highly depends on the studied system. If the collective variables are wrongly defined, the most probable outcome is that the process of interest will not be observed.

Mechanical unfolding pathway of a model β-peptide

<div style="text-align: right">3</div>

The β-peptides have been the most studied foldamer class, since Seebach and Gellman *et al.* reported for the first time in 1996 their ability to adopt helical conformations that are more stable than those of α-peptides [57]. β-peptides are derived from natural α-peptide sequences via backbone homologation. A single methylene insertion into each amino acid of an α-peptide leads to a β-peptide [1].

The extra carbon in β-peptides' backbone adds new variables which affect their secondary structure formation, *e.g.*, backbone substitution patterns and stereochemistry. Therefore, the universe of possible conformations that β-peptides may adopt is enlarged by the extra carbon [1]. As a consequence, whereas only two types of hydrogen bonded (H-bonded) helices are known for α-peptides (the α- and the 3_{10}-helix) [58], more than five types of H-bonded helices have been found for β-peptides [57, 59, 60]. For β-peptides, the helix type is mainly determined by the choice of the β-amino acid monomer. Cyclic ring constraints within the monomer of four, five, or six atoms promote the 10-helix, the 12-helix, and the 14-helix[1] (also called 3_{14}-helix[2]), respectively. Acyclic mono substituted residues tend to fold into 14-helices or 10/12-helices, depending on the substitution pattern [2]. From these helices, the 3_{14}-helix has been the most studied because of its high stability and its similarity with the α-helix [28, 29, 61].

Besides their structural versatility, there are two main aspects that make this kind of foldamer appealing: i) They are promising candidates for peptidomimetic, because their ability to form secondary structures and their resistance to proteases are comparable to those of natural peptides [26]. ii) Their folding time scales permit extensive theoretical studies of the folding process making them ideal test cases for the investigation of peptide folding [26].

Already very short β-peptides (6 to 7 monomers) have been found to adopt a stable 3_{14}-helix conformation in methanol (MeOH) [57]. Previous MD studies on β-peptides adopting a 3_{14}-helix have shown that the low flexibility of their backbone is the reason for the high stability of this helix. Keller *et al.* [26] showed that only one out of four dihedral angles in the β-peptide backbone can rotate freely, the one associated with the C_α-CO bond. Therefore, the high stiffness of the β-peptide's backbone stabilizes the helix conformation entropically.

In this chapter the study of the mechanical unfolding pathway of a model β-peptide (in the following referred to as heptamer) is presented. In section 3.1, the target system is

---

[1]10-, 12- and 14-helix mean that the intramolecular H-bonds found in the helices form cycles of 10, 12 and 14 atoms, respectively.

[2]In this case 3 indicates the number of residues per turn and 14 the number of atoms in the cycle formed by the H-bonds.

described. In section 3.2, the results of MD simulations using different FFs are presented and the criteria for choosing the right force field are discussed. The results of FPMD simulations of the heptamer are presented and discussed in Sections 3.3, 3.4, and 3.6. Finally, a summary and final discussion are given in Section 3.8.

3.1 The heptamer

The heptamer is a β-peptide with seven monomers(see Figure 3.1 for the sequence) that adopts a 3_{14}-helix conformation stabilized by five intramolecular H-bonds (shown in Figure 3.2). This β-peptide was chosen as the first model system for a series of reasons: i) The X-ray structure is available. ii) Previous MD simulations have reproduced its experimental structure [29, 57], determined a fast folding process (few nanoseconds), and revealed the reversibility of the unfolding process over a large temperature range [28, 29, 61]. iii) The heptamer's short chain-length allows to carry out a large amount of MD and FPMD simulations at an affordable computational cost.
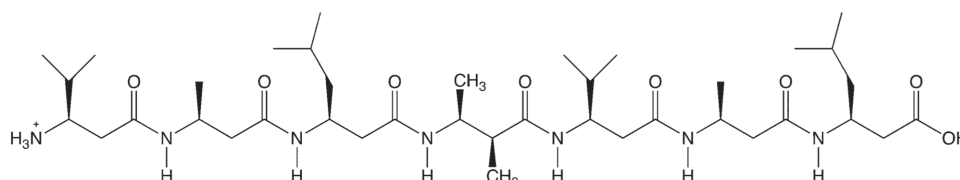
Fig. 3.1.: The heptamer's sequence.
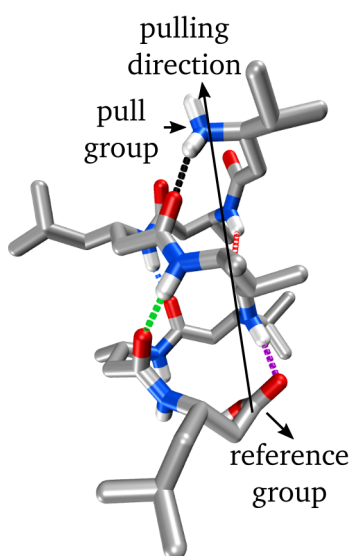
Fig. 3.2.: Schematic representation of the heptamer's 3_{14}-helix. The dashed lines represent H-bonds.

3.2 Choosing the right force field

As mentioned in Chapter 2, choosing the right FF to describe the system of interest is one of the main steps of preparing a MD simulation. The choice of the FF is critical and determines the results obtained, because there is no FF capable to describe every system under every condition. For this reason, it is important to clearly define which properties of the system are to be studied. This work is focused on the study of the mechanical unfolding pathway of foldamers. Therefore, the most relevant properties of the target systems are: the structure and stability of the folded and unfolded conformations. Thus, a force field capable of reproducing the stability of the helix conformation (observed experimentally [57]) and of obtaining the folded conformation from the unfolded conformation is desired.

Five different FFs (AMBER03 [62], CHARM27 [63], OPLS-AA [64], GAFF [65], and GRO-MOS53A6FF [40]) were chosen and tested according to the two criteria mentioned in the previous paragraph. For testing, two types of MD simulations were performed, both at temperatures ranging from 200 K to 400 K, in MeOH, and with a duration of 50 ns. In the first type, the simulations were started from the heptamer's helix conformation. In the second type, the simulations were started from a fully stretched unfolded conformation (all backbone dihedral angles were set to 180°). To display the results from the simulations the root-mean-square deviation (RMSD) was used. The RMSD is the average distance between atoms of two superimposed conformations. In this work, the RMSD was calculated between the backbone atoms of all non-terminal residues and for the superposition of the helix conformation and the conformation adopted by the oligomer at each time step. Thus, a small RMSD corresponds to a conformation close to the helix and a large RMSD to an unfolded conformation. For the following analysis, all conformations having a RMSD < 0.1 nm are considered to be folded, otherwise unfolded.

The RMSD vs time curves of the test simulations at $T = 300$ K are shown in Figure 3.3. The first set of trajectories (Figure 3.3, right) shows that only GAFF and GROMOS53A6FF are able to reproduce the stability of the helix conformation, *i.e.*, only in these simulations there is an equilibrium between the folded and unfolded conformations. In the second set of trajectories (Figure 3.3, left), only the GROMOS53A6FF is able to produce the helix from a fully unfolded conformation, *i.e.*, only for this FF the RMSD drops bellow 0.1 nm, during the simulation time.

At all the other studied temperatures a similar behavior was observed. Only the GRO-MOS53A6FF fulfilled the two criteria. For this FF, the increase of the temperature from 200 to 400 K was reflected in an increase of the RMSD fluctuations. Additionally, at 200 K no thermal folding/unfolding occurred, while at 400 K frequent folding and unfolding transitions were observed. The results with GROMOS53A6FF are in agreement with previous simulations of the heptamer in the same range of temperatures [28, 29]. In conclusion, this FF is the most suitable one for the simulations carried out in this work.
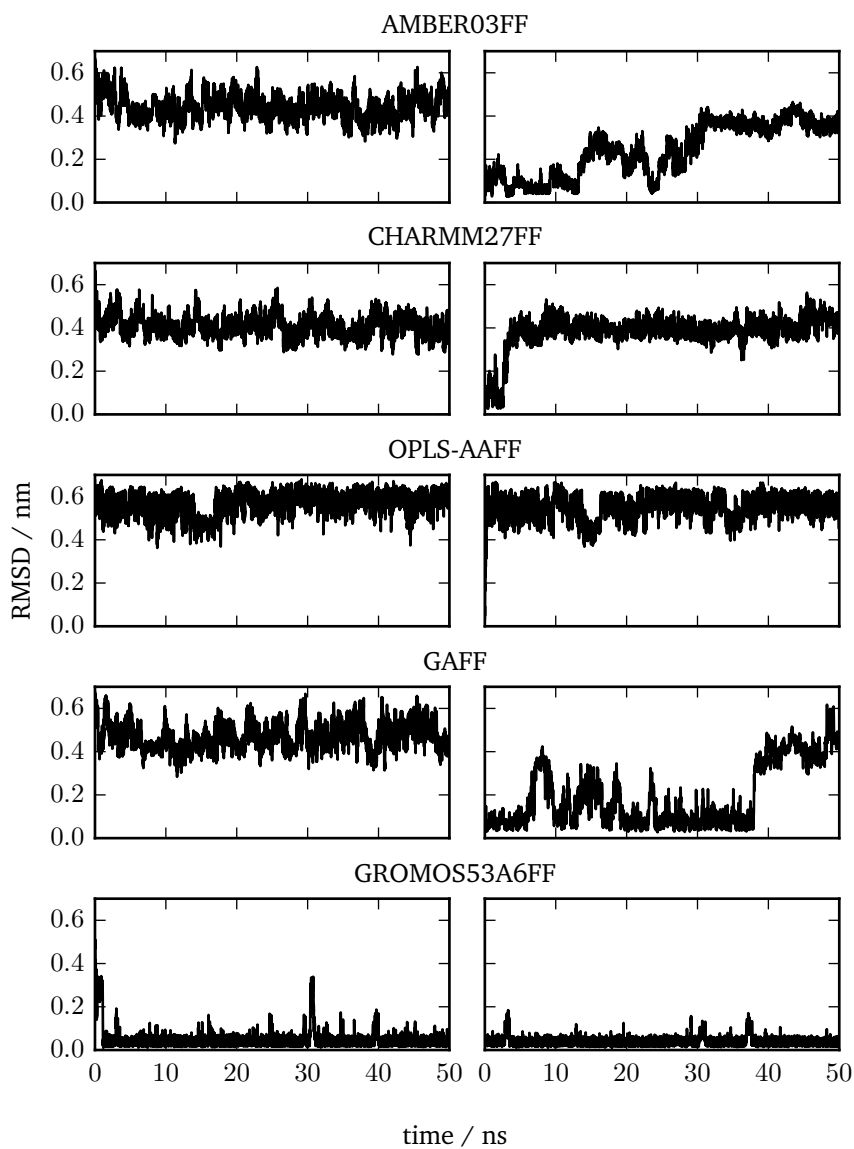
Fig. 3.3.: RMSD vs time curves for the heptamer using five different FFs at $T = 300\,\text{K}$. Left: simulations starting from a completely stretched conformation. Right: simulations starting from the helix conformation shown in Figure 3.2.

3.3 Force-probe molecular dynamics of the heptamer

FPMD simulations using the GROMOS53A6FF were performed to investigate the mechanical unfolding pathway of the heptamer. As explained in Section 2.5, in a FPMD simulation a harmonic potential (with spring constant K) attached to a pull group is moved with a constant velocity V while a reference group is kept fixed. For the heptamer, the pull group and the reference group are the N atom of the N-terminus and C atom of the C-terminus, respectively. Thus, the pulling direction corresponds to that of a vector pointing from the C atom to the N atom (see Figure 3.2). The distance between the pull and the reference group is named end-to-end distance. The same pull and reference groups were used in all FPMD simulations performed in this work.

All FPMD simulations of the heptamer were carried out at $T = 200\,\text{K}$, because at this temperature no thermal unfolding was observed in the equilibrium simulations (see Section 3.2). From the FPMD simulations a force vs extension (FE) curve is obtained. Such a curve is a plot of the force measured at the spring attached to the pull group as a function of the extension $x = Vt + x_i$, where t is the simulation time and x_i the initial extension. For the heptamer and all other oligomers studied in this work, x_i is the equilibrium distance between the N and C atoms of the N- and C-terminus. The maxima or rips in a FE curve correspond to rupture events. A rupture event occurs when a conformational change takes place, for example, the breaking of H-bonds or a large rotation of backbone bonds. As explained in Section 2.5, the force increases before a rupture event due to the stretching of the spring and the pulled molecule. When a rupture events occurs, *e.g.*, when there is a conformational change, the spring relaxes and the force drops, giving rise to the rips in the FE curve.
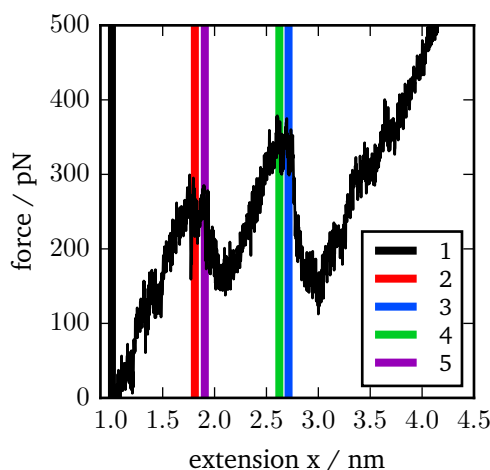
Fig. 3.4.: Typical FE curve using $K = 500\,\text{pN/nm}$, $V = 1\,\text{nm/ns}$ at $T = 200\,\text{K}$. The vertical lines indicate the extensions at which the opening of the respective H-bond is observed in the pulling trajectory. A threshold of 0.3 nm for the O-H distance was used to decide when a H-bond was opened. The color code for the H-bonds is the same as used in Figure 3.2.

A sample FE curve is shown in Figure 3.4, for $K = 500\,\text{pN/nm}$ and $V = 1\,\text{nm/ns}$, *i.e.*, loading rate $\mu = 5 \times 10^{11}\,\text{pN/s}$. Here, the colored vertical lines mark the extension at which a H-bond is opened, a O-H distance of 0.3 nm was used as threshold. Some general remarks about the sample FE curve are: i) The almost linear increase of the force before the first rupture event suggests that the molecular free energy landscape in the pulling direction is approximately harmonic. ii) All observed rupture events are related to the breaking of H-bonds, *i.e.*, all colored lines correspond to rips in the FE curve. iii) Several H-bonds open at the same time, suggesting that the helix unfolding is a cooperative process. iv) The opening of all H-bonds happens at similar forces. v) The opening of the N-terminus H-bond (1 black) is almost not observable, because it happens really early in the trajectory. vi) The outermost H-bonds (2 red and 5 purple) are broken first and the most inner H-bonds (4 green and 3 blue) are opened last.

The previously described behavior corresponds to a single trajectory with a single set of initial conditions. Since the opening of the H-bonds is a stochastic process, in order to obtain statistically relevant results, it is necessary to calculate a large number of trajectories with varying initial conditions. Thus, between 100 and 1000 simulations were carried out for each of the parameters sets shown in Table 3.1. Various loading rates were used to study its impact on the unfolding pathway and on the rupture forces.

For all loading rates, the same H-bond rupture order was observed: the N-terminus H-bond breaks immediately after starting pulling, then the outermost H-bonds break and finally the innermost H-bonds break. By analyzing the number of H-bonds formed between the heptamer and MeOH molecules during a pulling trajectory, the intramolecular N-terminus H-bond was found to have a low stability because of its ability to form intermolecular H-bonds with MeOH molecules. The H-bond rupture order was further studied at higher temperatures. 100 simulations were carried out for $V = 1\,\text{nm/ns}$ and $K = 500\,\text{pN/nm}$ at

Tab. 3.1.: H-bond rupture order at different loading rates ($\mu = K \times V$) in MeOH and in $CHCl_3$, determined as the most probable pathway from 100 or 1000 trajectories. Note that not all loading rates where investigated in $CHCl_3$.

		Breaking order	
K (pN/nm)	V (nm/ns)	MeOH	$CHCl_3$
1000	0.01	12543	54312
1000	0.05	12543	54312
1000	0.15	12543	54312
1000	0.50	12543	-
1000	1.00	12543	54312
1000	2.50	12543	54312
1000	10.0	12543	15432
1000	25.0	12543	15432
1000	50.0	12543	15432
500	0.10	12543	-
500	1.00	12543	54312
500	10.0	12543	-
200	0.10	12543	-
200	1.00	12543	-

$T = 298\,\text{K}$ and $T = 360\,\text{K}$, in MeOH. The same preferred pathway was found in these simulations.

In contrast, when the same simulations were performed in chloroform (CHCl_3) a different H-bond rupture order was observed. In CHCl_3, the unfolding also tends to start at the termini and propagates to the center, but the order depends on the loading rate. Table 3.1 summarizes the H-bond rupture order for each studied loading rate in MeOH and CHCl_3. Besides the polarity, the main difference between MeOH and CHCl_3 is that the former has the ability to form intermolecular H-bonds. Therefore, it is plausible that the differences in the H-bond rupture order are caused by the presence or lack of solvent-heptamer intermolecular H-bonds. This hypothesis is supported by the fact that in CHCl_3, the N-terminus H-bond is no longer always the first one to open.

For each loading rate, rupture force distributions for the opening of each H-bond were obtained. The rupture forces were determined as local maxima in the FE curves, in the vicinity of the extension at which the rupture of the corresponding H-bond takes place. In Figure 3.5, the rupture force distributions for each H-bond (except 1) from 1000 trajectories with $\mu = 5 \times 10^{11}\,\text{pN/s}$ are shown. The distribution for H-bond 1 (located at the N-terminus) is not shown, because it opens early in the pulling trajectory at a negligible breaking force. The rupture force distributions for the remaining four H-bonds behave very similar, showing that the four H-bonds have a similar strength.

As explained in Section 2.6, the kinetic rate of the associated rupture event can be determined from the force distribution. The kinetic rates found for each of the four H-bonds, using the method described in Section 2.6.1, are similar and yield an average of $\tau\,(360\,\text{pN})$ $\approx 0.2\,\text{ns}$. This analysis does not provide significant information about the unfolding pathway of the heptamer, besides the fact that all H-bonds break at similar forces and with similar rates. However, if the cooperative character of the unfolding process is taken into account, more relevant information about the unfolding pathway is obtained. This analysis is presented in the following section.

3.4 A two-barriers model

Statistical information about the unfolding pathway is obtained by calculating the average behavior of a large number of trajectories. In Figure 3.6, averages over 100 or 1000 FE curves (so called dynamic strength, DS) for various loading rates are shown. Each DS exhibits only two peaks. Hence, the detailed structure observed in single FE curves is averaged out and only two rupture events remain. The resolution of both peaks is higher for the simulations done with a stiffer spring, because in those cases the system is forced to follow the pulling potential more closely. The fact that the number of H-bonds is higher than the number of peaks implies that several H-bonds open at the same time and this reflects the cooperative character of the unfolding process.

Average H-bond distance vs extension curves (AHDC) were also calculated and are shown in Figure 3.7. By determining the extension at which each H-bond opens, *i.e.*, the extension at which each H-bond distance takes a value larger than 0.3 nm in the AHDCs, it becomes clear that the first rupture event corresponds to the opening of the outermost H-bonds (2 and 5)
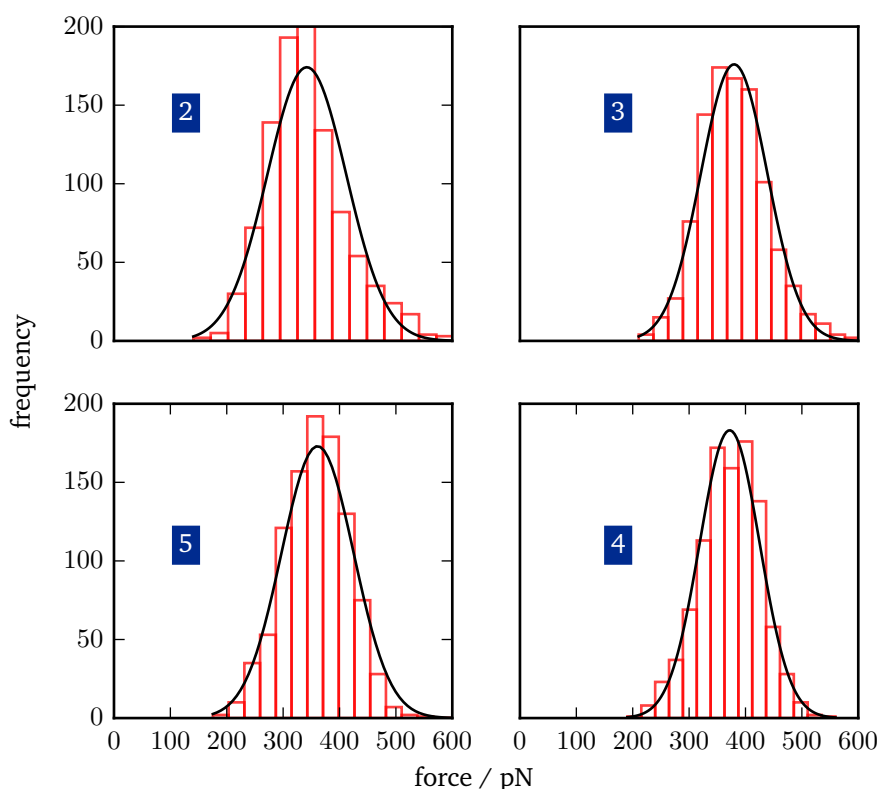
Fig. 3.5.: Rupture force distributions for the breaking of the individual H-bonds as indicated. The left panel shows the distributions for the outer H-bonds 2 and 5, and the right panels the ones for the inner H-bonds 3 and 4. Solid lines represent fits to a Gaussian with mean $\langle F \rangle$ and standard deviation σ: (2) $\langle F \rangle = 328.8$ pN, $\sigma = 58.9$ pN; (5) $\langle F \rangle = 360.7$ pN, $\sigma = 58.8$ pN; (3) $\langle F \rangle = 374.0$ pN , $\sigma = 56.8$ pN; (4) $\langle F \rangle = 374.4$ pN, $\sigma = 54.2$ pN.

and the second rupture event corresponds to the opening of the inner H-bonds (3 and 4). The breaking of H-bond 1 again happens very early and is practically not observable.

The two-barriers model is further supported by the average fraction of native contacts vs extension curves (ANC), shown in Figure 3.8. In these curves, a two-step decay is observed. Since the fraction of native contacts is a measure of the amount of native structure present in a conformation, the two-step decay in the ANCs implies that during its unfolding the heptamer adopts mainly three different conformations: the helix, an intermediate, and the unfolded conformation.

The DS, AHDC, and ANC indicate that the unfolding pathway of the heptamer is characterized by two energy barriers and a stable intermediate state. The first energy barrier is related to the opening of the outermost H-bonds and the second barrier to the opening of the central H-bonds. The analysis of the conformations adopted during a single pulling trajectory, shows that the intermediate conformation corresponds to a half-unfolded helix in which only the outermost H-bonds are opened.
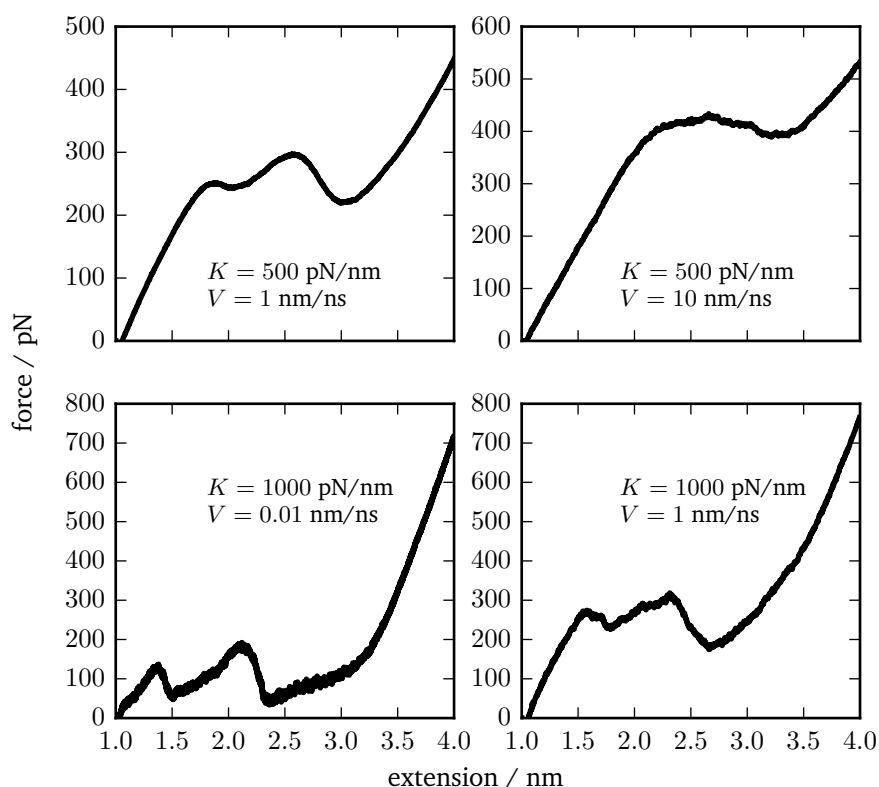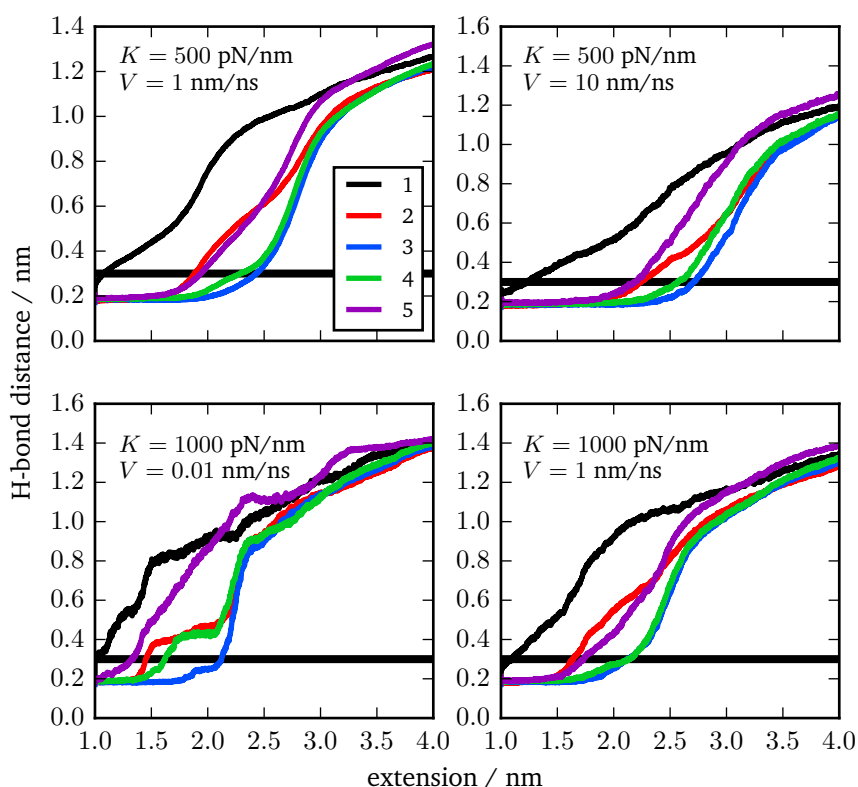
Fig. 3.6.: DS, *i.e.*, average over several FE curve at $T = 200\,\text{K}$ for various loading rates as indicated in the respective panel. For $K = 500\,\text{pN/nm}$ and $V = 1\,\text{nm/ns}$, an average over 1000 simulations was calculated; in all other cases, 100 simulations were used.

3.5 Potential of mean force of the heptamer

So far, the existence of a stable intermediate state in the unfolding pathway of the heptamer has been shown under non-equilibrium conditions. However, the validity of this finding under equilibrium conditions is not necessarily given. This validation was performed through the calculation of the PMF (see Section 2.7) of the unfolding pathway of the heptamer with the end-to-end distance as reaction coordinate. For this calculation, US+WHAM (see Section 2.7) was used.

The PMF can be interpreted as a free energy as a function of the reaction coordinate, when all other degrees of freedom (apart from the reaction coordinate) are not restrained and the system is given enough time to relax to equilibrium. This was accomplished by studying the effect of the length of the US simulations on the shape of the PMF. A length of 10 ns per US simulation was found to be ideal, since a further increase of the simulation time did not modify the calculated PMF significantly. However, the shape of the PMF was found to depend on the initial conditions, *i.e.*, the initial velocities, of the US simulations. Therefore, an average over 100 PMF was calculated to obtain statistical significance.

In Figure 3.9, the average PMF (black) and $U(x)$ (dashed lines) are shown. $U(x)$ is the overall potential in the soft spring limit obtained when an external force is applied to the
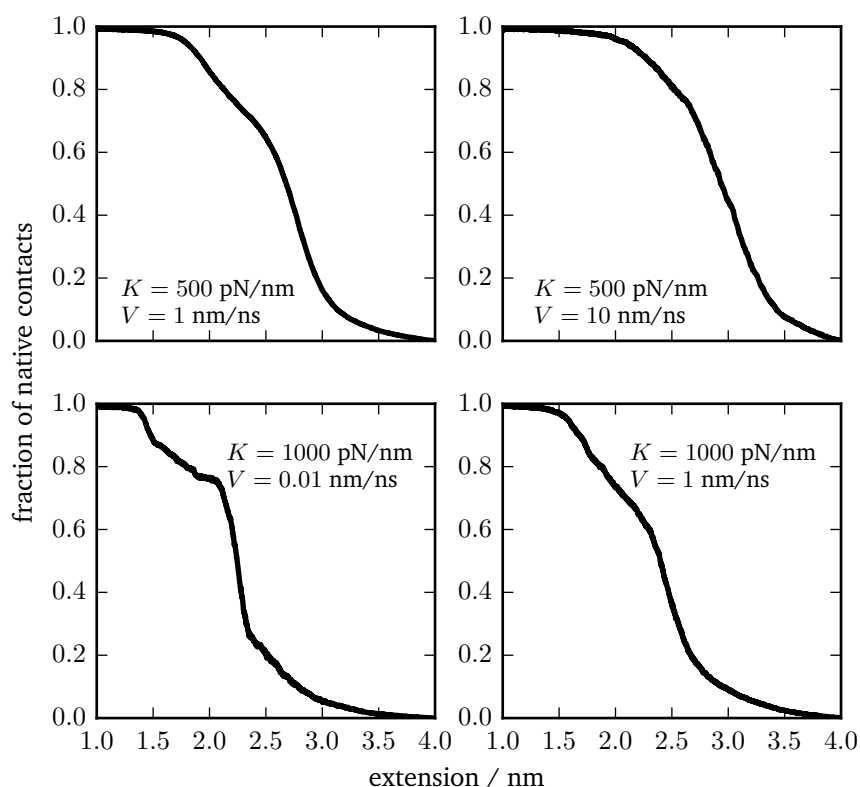
Fig. 3.7.: AHDC at $T = 200\,\mathrm{K}$ for various loading rates as indicated in the respective panel. The distance is calculated between the H atom and the acceptor atom. The black horizontal lines marks the threshold used to decide if a H-bond was opened. For $K = 500\,\mathrm{pN/nm}$ and $V = 1\,\mathrm{nm/ns}$, an average over 1000 simulations was calculated; in all other cases, 100 simulations were used. The color code for the H-bonds is the same used in Figure 3.2.

system, $U(x) = U_{\mathrm{PMF}}(x) - F \cdot x$. The PMF shows two shoulders and a minimum. The structures shown in Figure 3.9 represent the conformations adopted by the heptamer at the indicated distances. The structure corresponding to the intermediate state is a partially unfolded helix in which the outer H-bonds are opened and the inner H-bonds remain closed, equivalent to the conformation found in the FPMD simulations. The $U(x)$ curves show how the intermediate and the unfolded states are stabilized due to the effect of the pulling force. Furthermore, in the $U(x)$ curves the shoulders become real maxima. The form of the PMF and the $U(x)$ curves, together with the observed structure for the intermediate state, are consistent with the results of the FPMD simulations. This consistency proves that the unfolding pathway observed under non-equilibrium conditions is the same as the one observed under quasi-equilibrium conditions. In conclusion, the heptamer unfolds in two steps, *i.e.*, during its unfolding the heptamer overcomes two energy barriers.

3.6 Kinetics

By adopting the two energy barriers model, new rupture force distributions are calculated from the FE curves, an example is shown in Figure 3.10.A. Here, the rupture forces were

Fig. 3.8.: ANC at $T = 200\,\mathrm{K}$ for various loading rates as indicated in the respective panel. For $K = 500\,\mathrm{pN/nm}$ and $V = 1\,\mathrm{nm/ns}$, an average over 1000 simulations was calculated; in all other cases, 100 simulations were used.

determined as the closest maxima to the extension at which the maxima were found in the DSs. As expected, the mean rupture forces and the variance for both barriers are very similar to the corresponding values obtained for the rupture of each H-bond (Figure 3.5). Analogous to the case of the individual H-bond rupture force distributions, the distributions obtained within the two-barrier model can be used to calculate the kinetic parameters of the associated unfolding events.

There are several methods to calculate kinetic rates from rupture force distributions (see Section 2.6). A possible approach consists of fitting the mean forces $\langle F \rangle$ obtained with different loading rates to a suitable model. For small loading rates, stochastic models of diffusive barrier crossing are applicable. Within these models [48, 49], the force is proportional to the logarithm of the loading rate ($\mu = K \cdot V$), $\langle F \rangle \sim \ln(K \cdot V)^{\nu}$, with $\nu = 1/2$ or $1/3$. In contrast, for large loading rates, the bond rupture is assumed to be dominated by Stokes' friction [49]. The force is then proportional to the pulling velocity $\langle F \rangle \sim \gamma \cdot V$, where γ is the friction coefficient.

In Figure 3.10.B, the mean rupture force $\langle F \rangle$, for both rupture events, is plotted as a function of the loading rate on a logarithmic scale. The shape of the curves seems to indicate that, within the studied loading rate range, there is not a crossover from logarithmic to linear dependency. However, when the three points with largest loading rates are fitted to Stokes' model, a friction coefficient $\gamma \simeq 4.7\,\mathrm{pN \cdot s/m}$ and a diffusion coeffi-

cient $D = (k_B T)/\gamma \simeq 6\times10^{-10}$m^2/s are obtained (where k_B is the Boltzmann constant and $T = 200$ K). The fitted value for D is comparable to the experimental value for pure MeOH, $D(201$ K$) = 2.4\times10^{-10}$m^2/s [66], which indicates that the three larger loading rates are out of the diffusive barrier crossing regime. Therefore, these data points will be excluded from the further analysis, to guarantee that there is not frictional effects in the determined kinetic rates.

Two different methods were used to calculate kinetic rates from the rupture force distributions, the results are shown in Figure 3.11. The lines in this figure are fits to Equation 2.28. One of the methods (blue circles) (here called lifetime method and described in Section 2.6.1) uses Equation 2.30 to determine the lifetimes τ. The kinetics rates are calculated as the inverse, $k = 1/\tau$. In this method, the histograms for each rupture event are assumed to be produced by an independent single-barrier process and not from a multiple barrier process. As a consequence, in the lifetime method the consecutive energy barriers are assumed independent of each other. In contrast, in the other applied method (here called ZD and described in Section 2.6.2) the barriers are not assumed to be independent. For the heptamer's unfolding pathway, the assumption of independent barriers is expected not to be accurate, because the opening of the inner H-bonds (second barrier) only occurs after the opening of the outer H-bonds (first barrier).
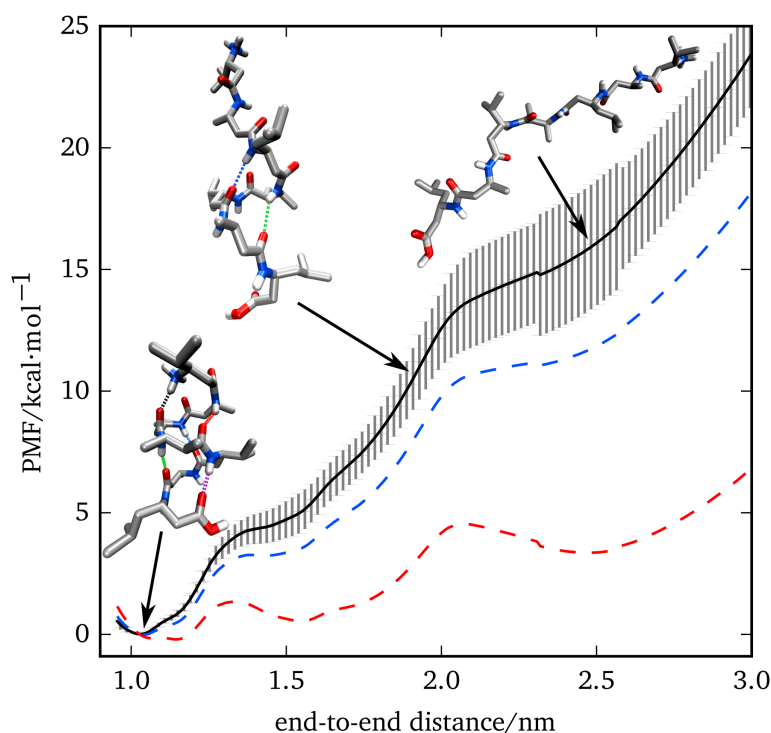
Fig. 3.9.: Average over 100 PMFs calculated at $T = 200$ K. Error bars are indicated in light grey. For the umbrella sampling, a force constant of $K = 500$ pN/nm was used. The dashed curves correspond to the effective potential resulting from the combination of the PMF and the pulling potential, $U(x) = U_{\text{PMF}}(x) - F \cdot x$ for $F = 2.9$ kcal$\cdot$ mol$^{-1}\cdot$ nm^{-1} (blue dashed line) and $F = 7.2$ kcal$\cdot$ mol$^{-1}\cdot$ nm^{-1} (red dashed line). The structures shown correspond to an end-to-end distance of 1.03, 1.95, and 2.62 nm.
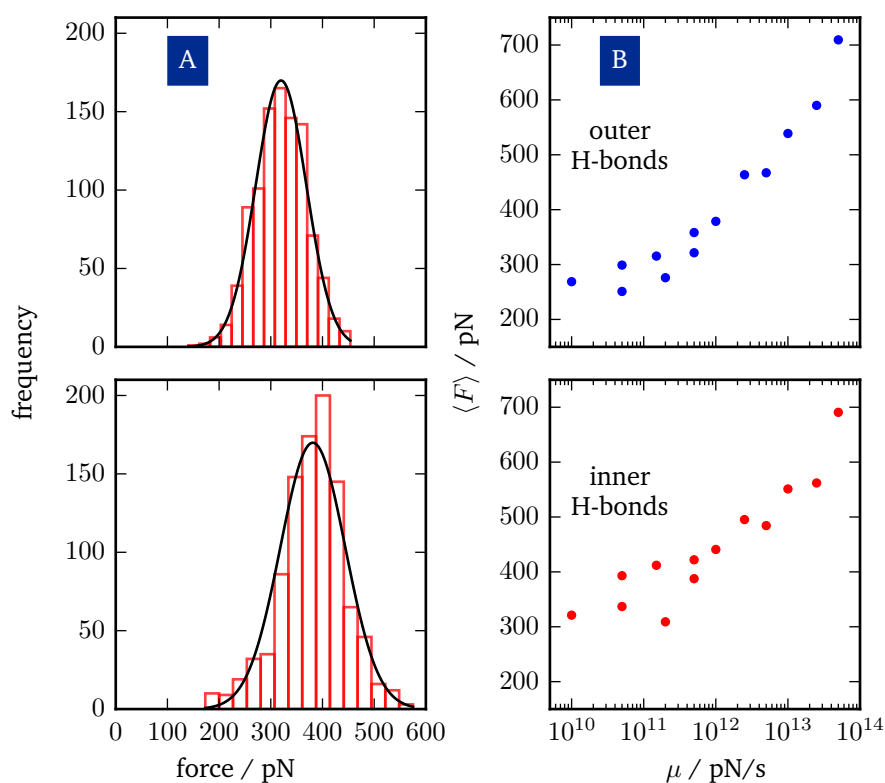
Fig. 3.10.: A. Rupture force distributions for the two unfolding steps, with $K = 500\,\text{pN/nm}$, $V = 1\,\text{nm/ns}$, and $T = 200\,\text{K}$. Upper panels: rupture of the outer H-bonds; lower panels: rupture of the inner H-bonds. Solid lines are fits to Gaussians: outer H-bonds: $\langle F \rangle = 321.7$ pN, $\sigma = 49.2$ pN; inner H-bonds: $\langle F \rangle = 387.9$ pN, $\sigma = 51.9$ pN. **B.** Mean rupture force vs loading rate for the two barriers. Note: for some loading rates, different combinations of K and V were used.

The red triangles and the green squares in Figure 3.11 are obtained using the ZD method for rupture force distributions calculated at the same loading rate ($\mu = 5 \times 10^{11}$ pN/s), but with different values of V and K. This loading rate was chosen because it is sufficiently small to avoid possible frictional effects, but large enough to allow several hundred simulations to be performed within an accessible computational time.

Both methods applied require different input information. The lifetime method requires the calculation of rupture force distributions for several loading rates. In order to obtain significant statistical values, for each rupture force distribution, at least 100 simulations need to be performed. Hence, if ten different loading rates are investigated, at least 1000 simulations are needed. Moreover, the simulations at small loading rates are computationally more expensive than the simulations at large loading rates, because at small loading rates the unfolding process takes places at larger time-scales. In contrast, the ZD method requires the calculation of a single rupture force distribution, *i.e.*, simulations with only one loading rate. However, the ZD method is more sensible to the quality of the force distribution. Thus, it is necessary to carry out around 1000 simulations for a chosen loading rate. If an intermediate loading rate is used, *e.g.*, $\mu = 5 \times 10^{11}$ pN/s, the ZD method is computationally less expensive than the lifetime method.
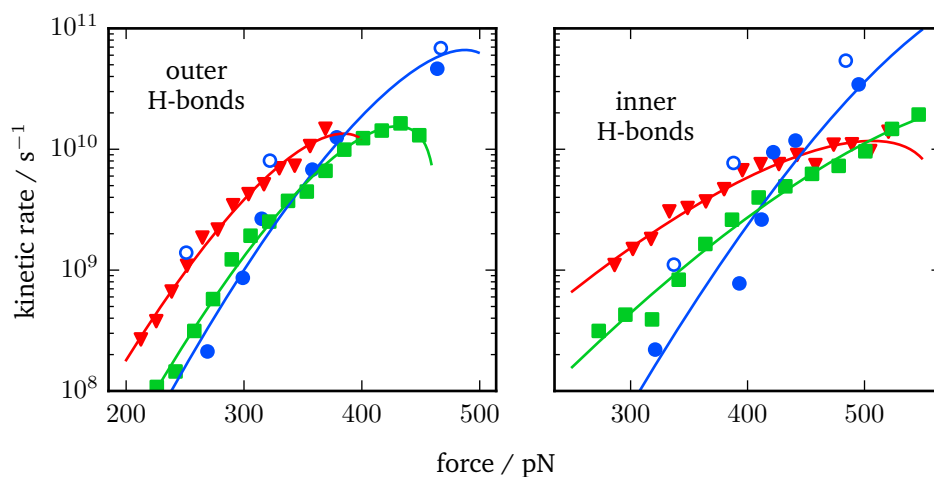
Fig. 3.11.: Kinetic rates as a function of the force for the breaking of the outer H-bonds (left) and the breaking of the inner H-bonds (right). Blue circles depict the rates obtained as the inverse lifetimes according to Equation 2.30. Full symbols: $K = 1000\,\mathrm{pN/nm}$, open symbols: $K = 500\,\mathrm{pN/nm}$. Red triangles and green squares symbols are obtained using the method of Zhang and Dudko [50] for $K = 500\,\mathrm{pN/nm}$, $V = 1\,\mathrm{nm/ns}$ (red) and $K = 1000\,\mathrm{pN/nm}$, $V = 0.5\,\mathrm{nm/ns}$ (green). The lines are fits to Equation 2.30 using $\nu = 2/3$. The fitted parameters are given in Table 3.2

The results of fitting the three data sets in Figure 3.11 to Equation 2.28 are compiled in Table 3.2. The fitted values obtained using the ZD method with two different pulling velocities are very similar, showing that Equation 2.28 applies under the simulated conditions, *i.e.*, Equation 2.28 is applicable under the investigated velocities and spring constants. Moreover, the agreement between the fitted values confirms that the soft spring approximation, made in the used stochastic model, is valid for the studied system and conditions.

While for the first barrier, both methods show similar fitted parameters, the results diverge for the second barrier. This facts shows that the kinetic rate of the first barrier is independent of the second barrier, while the kinetic rate of the second barrier depends on the

Tab. 3.2.: Kinetic parameters resulting from fitting the three data sets in Figure 3.11 to Equation 2.28. $k(0)$ is given in s^{-1}, $\Delta G^{\ddagger}$ in $\mathrm{kcal \cdot mol^{-1}}$, and $x^{\ddagger}$ in nm. $k(a)$: fitting parameters for the data set are obtained using the ZD method [50] for $K = 1000\,\mathrm{pN/nm}$ and $V = 0.5\,\mathrm{nm/ns}$; $k(b)$: as $k(a)$ for $K = 500\,\mathrm{pN/nm}$ and $V = 1\,\mathrm{nm/ns}$; $1/\tau$: fitting parameters to the data set are obtained using the lifetime method [48].

|		First barrier			Second barrier		
		$k(a)$	$k(b)$	$1/\tau$	$k(a)$	$k(b)$	$1/\tau$
$\nu = 2/3$	$k(0)$	1.7×10^6	9.6×10^5	1.1×10^6	3.4×10^6	9.1×10^6	8.9×10^4
	$\Delta G^{\ddagger}$	4.2	4.4	5	4.1	3.4	6.5
	$x^{\ddagger}$	0.11	0.12	0.12	0.07	0.07	0.11
$\nu = 1/2$	$k(0)$	7.4×10^5	4.8×10^5	3.1×10^5	2.7×10^6	3.0×10^6	8.3×10^4
	$\Delta G^{\ddagger}$	4.8	5	5.7	4.7	4.1	8.1
	$x^{\ddagger}$	0.13	0.14	0.14	0.07	0.09	0.11

kinetic rate of the first barrier. Therefore, the assumption of independent barriers made in the lifetime method only affects the determination of the second barrier's kinetic parameters.

If $\Delta G^{\ddagger}$ and $x^{\ddagger}$ are compared to the values obtained from the PMF (Figure 3.9), the $\Delta G^{\ddagger}$ values for both barriers are consistent. In contrast, the values for $x^{\ddagger}$ are only consistent for the first barrier. In the case of the second barrier, the values of $x^{\ddagger}$ obtained through the fitting procedure (Table 3.2) are too small in comparison to the value found in the PMF ($\sim 0.2\,\text{nm}$). A possible explanation for this discrepancy is the presumption of a specific form of the free energy profile, made in the used kinetic models, which might not be valid for the second barrier, since in the PMF the intermediate state minimum appears to be very shallow. Hence, a quadratic cusp or a linear-cubic function may not describe correctly the energy profile around the second barrier.

The discrepancy between the two methods shows the need to consider the possible dependency of consecutive energy barriers. In addition, the wrong description of the energy profile around the second barrier demonstrated that more complex models are required.

3.7 Reversibility of the mechanical unfolding

The reversibility of the heptamer's mechanical unfolding was evaluated both in MeOH and in CHCl$_3$ for the pulling velocities given in Table 3.1. In MeOH, the unfolding was found to be irreversible even at the slowest investigated pulling velocity (0.01 nm/ns). An analysis of the interactions with the solvent and the heptamer showed that the refolding in MeOH does not take place because the heptamer forms H-bonds with the MeOH molecules. These H-bonds cannot be easily opened under the conditions of reverse pulling, because there is not sufficient time for the system to relax. Unfortunately, lower pulling velocities, at which
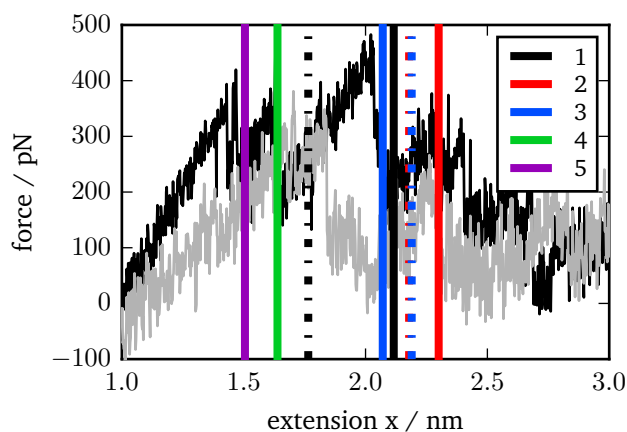
Fig. 3.12.: Typical FE curve using $K = 1000\,\text{pN/nm}$, $V = 0.1\,\text{nm/ns}$ at $T = 200\,\text{K}$ of an unfolding-refolding trajectory of the heptamer. The black curve corresponds to the unfolding and the gray curve to the refolding. The full/dashed lines indicate the extension at which the opening/forming of a H-bond is observed in the trajectory. A threshold of 0.3 nm for the O-H distance was used to decide when a H-bond was opened. The color code for the H-bonds is the same used in Figure 3.2.

the refolding might occur, cannot be investigated because the associated computational cost is too high.

In contrast to MeOH, in $CHCl_3$ a partial refolding was observed for slow pulling velocities ($> 0.1\,nm/ns$). A sample FE curve of an unfolding-refolding trajectory is shown in Figure 3.12. In the sample FE curve, the rupture and rebinding of the H-bonds is reflected as rips in the force. However, in the refolding curve there are less rips because only three out of five H-bonds are formed. Moreover, the forces at extension x_0 are not equal for the pulling and the refolding curves because the final conformation of the refolding trajectory is a partially folded helix.

3.8 Conclusions

In this chapter, a study of the mechanical unfolding pathway of a model β-peptide (the heptamer) in MeOH using FPMD is presented. In the first section, the choice of the right force field for describing the heptamer is illustrated. Results of temperature dependent MD simulations show that from the five studied force fields only the GROMOS53A6FF is capable of reproducing the heptamer's helix conformation and stability. The failure of FFs standardly used for the study of biomolecules (*e.g.*, Amber and CHARM) is not unexpected, because the parametrization of these FFs is known to be biased in favor of proteins' folded motifs [67]. In contrast, the failure of more general FFs, (*e.g.*, GAFF and OPLS) is surprising, given that these FFs have been developed for the study of organic molecules and peptides. The comparison between the different FFs shows how the transferability of most FFs is far from being ideal.

The FPMD trajectories are analyzed considering the individual opening of each H-bond. However, it is proven that the relevant information about the unfolding pathway is only obtained when the cooperative character of the unfolding process is taken into account. This is achieved by studying the statistical behavior of the unfolding pathway through the calculation of the DSs and ANCs. For most of the investigated loading rates, except the largest ones, the shape of the DSs and the ANCs indicates an unfolding pathway consisting of two energy barriers and an intermediate state. Such a complex unfolding pathway is unexpected for a small oligomer like the heptamer, especially considering that all H-bonds open at similar forces.

A conformational analysis of the pulling trajectories leads to the conclusion that the first and second barriers correspond to the opening of the outermost H-bonds and the inner H-bonds, respectively. In addition, the opening of the N-terminus H-bond is found to happen early in the pulling trajectory at negligible forces. The N-terminus H-bond low stability is explained by the ability of the N-terminus to form multiple inter-molecular H-bonds with the solvent.

The two barrier model is further supported by the calculation of the PMF. The heptamer's PMF displays two energy barriers separated by a stable intermediate state. A conformational analysis of the US trajectories, used for the calculation of the PMF, indicates that the stable intermediate state correspond to a half-unfolded helix in which the outermost H-bonds are opened. The agreement between the PMF and the FPMD simulations shows

that the unfolding pathway observed in the FPMD (under non-equilibrum conditions) is analogous to the one observed in equilibrium.

Using the two barrier model, rupture force distributions are calculated for a large range of loading rates ($\mu = 10 \times 10^{10}$ pN/nm -10×10^{14} pN/nm). The dependence of the mean rupture force on the loading rate is found logarithmic for most of the studied loading rates, as expected for pulling conditions at which the diffusive barrier crossing model applies [49]. The comparison of the kinetic rates calculated from the rupture force using two different approaches shows that the second energy barrier is not independent of the first one. This dependency indicates that the measured rupture force for the second barrier is not characteristic of this transition alone, but it is influenced by the first transition. The calculated kinetic rates are fitted to two different types of model free-energy surfaces. The fitted parameters for the energy landscape are found to be in agreement to the corresponding values determined from the PMF for the first barrier, but not for the second barrier. The less meaningful parameters for the second barrier are suspected to be a consequence of a complex form of the energy landscape around this transition.

The dependency of the second barrier on the first barrier is congruent with the statistical analysis of the H-bond rupture order, in which the central H-bonds are found to open only after the termini H-bonds have been opened. Therefore, the results obtained from the applied kinetic models support the conclusion that the heptamer exhibits a well defined non-trivial unfolding pathway. Moreover, the comparison of the H-bond rupture order in different solvents shows that, even though the order changes with the proticity of the solvent, the preference for an unfolding from the termini to the center remains. Therefore, the way in which the heptamer unfolds must correspond to intrinsic features of its backbone architecture. Such determining structural factors are investigated in Chapter 6.

Finally, the reversibility of the mechanical unfolding of the heptamer is investigated both in MeOH and CHCl$_3$. The unfolding is observed to be irreversible in MeOH and partially reversible in CHCl$_3$. The non-reversible unfolding process in MeOH is a consequence of the formation of inter-molecular H-bonds between the heptamer and MeOH molecules. The different behavior in MeOH and CHCl$_3$ evidences the importance of taking into account solvent effects when studying the unfolding pathway of short oligomers.

Comparison of the mechanical unfolding pathways of α- and β-peptides

As depicted in the previous chapter, β-peptides are strong candidates for peptide-mimetics applications. Therefore, after the detailed study of the mechanical unfolding of a model β-peptide (the heptamer), one of the main question that arises is: How similar are the folding processes of α- and β-peptides helices?

In this chapter, a study of the differences and similarities between the mechanical unfolding pathways of α- and β-peptides is presented. The model systems are described in Section 4.1. In Section 4.2, the results of FPMD simulations of the oligomers are discussed. The calculated PMFs are shown and analyzed in Section 4.3. The effect of temperature in the unfolding pathways is discussed in Section 4.4. Finally, a summary is presented in Section 4.5.

4.1 Description of the model systems

For the comparative study of α- and β-peptides, decalanine (α-Ala$_{10}$) was chosen as the α-peptide model, because it has been extensively studied theoretically and experimentally [68–72]. In addition, its chain length is short enough to render atomistic simulations affordable. As a β-peptide model, β-HAla$_8$ was chosen due to its similarity with α-Ala$_{10}$. The helix conformations adopted by the two oligomers are shown in Figure 4.1 and their characteristic features are compiled in Table 4.1. Both systems were studied in MeOH and in water.

Tab. 4.1.: Structural information of α-Ala$_{10}$ and β-HAlA$_8$. *E.g.*, a C=O $(i) \rightarrow$H-N$(i+4)$ is a H-bond between the C=O from residue i and the H-N from residue $(i+4)$

System	Conformation (handedness)	Number residues	# backbone atoms		# dihedral angles		H-bonds[1]
			Total	Per residue	Total	Per residue	
α-Ala$_{10}$	α-helix (right)	10	30	3	30	3	6 C=O $(i) \rightarrow$ H-N$(i+4)$
β-HAla$_8$	3_{14}-helix (left)	8	32	4	32	4	6 C=O$(i) \rightarrow$ H-N$(i-2)$

[1] Number and types of H-bonds.
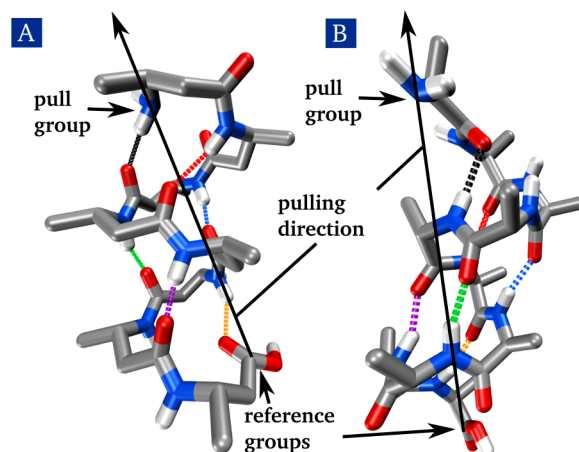
Fig. 4.1.: A. 3_{14}-helix conformation of β-HAla$_8$. **B.** α-helix conformation of α-Ala$_{10}$. Dotted colored lines represent the H-bonds that stabilize the helix conformations. Black arrows indicate the pulling direction.

4.2 FPMD simulations

FPMD simulations for α-Ala$_{10}$ and β-HAla$_8$ were performed at $T = 240$ K with $K = 1000$ pN/nm and $V = 0.1$ nm/ns. The temperature was chosen such that no thermal unfolding was observed. As in the case of the heptamer, the reference and the pulling groups are the N atom of the end terminus and the C atom of the C terminus.

A typical FE curve for each oligomer is shown in Figure 4.2. In the β-HAla$_8$ FE curve, the force reaches higher values and the extension interval at which the H-bonds opening occurs is larger. These facts indicate that the 3_{14}-helix adopted by β-HAla$_8$ has an overall higher mechanical resistance than the α-helix adopted by α-Ala$_{10}$. For β-HAla$_8$, the rupture events correspond to the opening of one or more H-bonds. In contrast, for α-Ala$_{10}$ the opening of the H-bonds is not reflected in the force, since there are no distinguishable rips/maxima. However, this is the behavior of a single FE curve.

Analogously to the methodology followed in the study of the heptamer (see Section 3.3), 100 FPMD simulations were performed for both oligomers, with $K = 1000$ pN/nm and $V = 0.1$ nm/ns at $T = 240$ K in order to get the statistical behavior of the unfolding pathway. The order in which the H-bonds open was studied by plotting the AHDCs, shown in Figure 4.3. Only the AHDCs in MeOH are shown because an equivalent behavior is observed in water. In the AHDCs for β-HAla$_8$, there are three opening events: i) the N-terminus H-bond opens right after the pulling starts, ii) the H-bonds 2, 5, and 6 open, and iii) the H-bonds 3 and 4 open. In contrast, in α-Ala$_{10}$'s AHDC there are no groups of H-bonds that open simultaneously, but there are approximately five opening events which start at the C-terminus and propagate to the N-terminus.

The DSs and ANCs for both oligomers, in MeOH and water, are shown in Figure 4.4 and Figure 4.5, respectively. In β-HAla$_8$'s DSs there are two rupture events. The comparison of the DS and the AHDC in MeOH indicates that the first rupture event corresponds to the opening of H-bonds 2, 5, and 6, and the second rupture event corresponds to the opening of H-bonds 3 and 4. The opening of the N-terminus H-bond is not observed in the DS,

because this H-bond breaks at very short extensions and low forces. The low stability of the N-terminus H-bond was also observed for the heptamer and was found to be a consequence of the ability of the N-terminus to form H-bonds with the solvent. In accordance to β-HAla$_8$'s DSs, the ANCs shows a two-step decay. The almost constant value of $\sim$0.5 in the ANCs, at extensions of 1.6 nm to 2.5 nm, indicates the existence of a stable intermediate conformation in β-HAla$_8$'s unfolding pathway.

In contrast to β-HAla$_8$'s DSs, in α-Ala$_{10}$'s DSs there are not distinguishable maxima. However, there is a plateau in the DS within the same extension interval ($\sim$1.4 nm to 4.0 nm) for which the opening of the H-bonds is observed in the AHDC. This plateau corresponds to a series of conformational changes that take place in a quasi-continuous way and at a constant force. There, the spring (represented by the pulling harmonic potential) is not stretched, because the pulled oligomer unfolds without measurable resistance. In agreement, α-Ala$_{10}$'s ANCs shows a one-step decay.

4.3 Potentials of mean force

To gain insight into the features of the energy landscapes and unfolding pathways under equilibrium conditions, the PMFs for both oligomers, in MeOH and water, were calculated using US+WHAM (see Section 2.7). Previous efforts had been made to determine the PMF of α-Ala$_{10}$ in vacuo and in water using different methodologies. At $T = 300$ K in vacuo, the PMF using the end-to-end distance as a reaction coordinate, shows a single minimum at an end-to-end distance that corresponds to the α-helix conformation [71–74]. Recently, Hazel, Chipot, and Gumbart (HCG) [69] calculated the PMF in water at $T = 300$ K. Their calculations indicate that when using the end-to-end distance as reaction coordinate, the conformation with the lower energy is not the α-helix but an unfolded conformation.
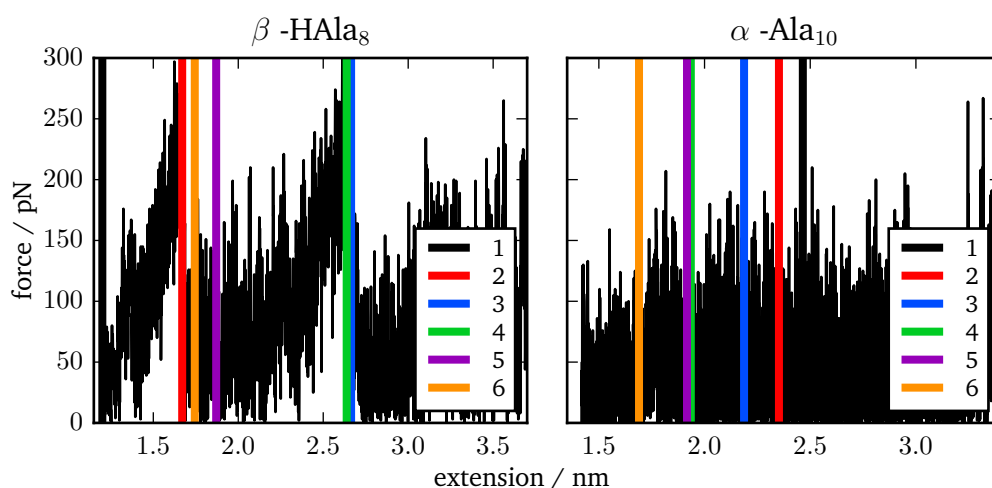
Fig. 4.2.: Typical FE curve using $K = 1000$ pN/nm and $V = 0.1$ nm/ns at $T = 240$ K (with a total simulation time of 35 ns). Left β-HAla$_8$ and right α-Ala$_{10}$. The vertical lines indicate the extensions at which the opening of the respective H-bond is observed in the trajectory. A threshold of 0.3 nm for the O-H distance was used to decide when a H-bond was opened. Each color corresponds to one of the six H-bonds shown in Figure 4.1
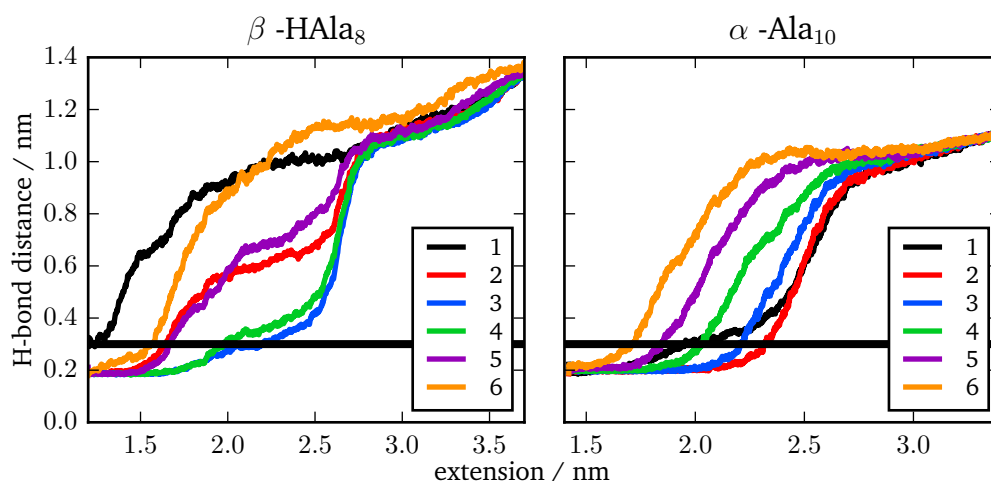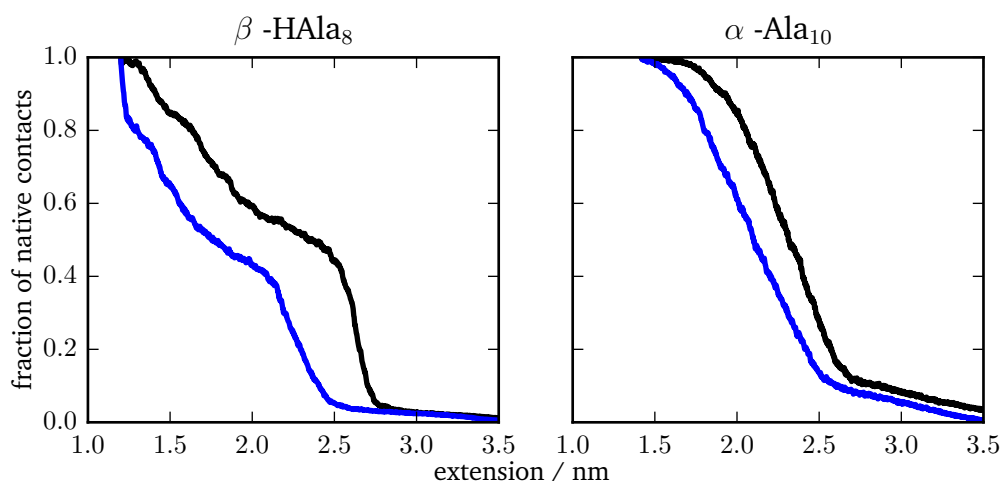
Fig. 4.3.: AHDC using $K = 1000\,\text{pN/nm}$ and $V = 0.1\,\text{nm/ns}$ at $T = 240\,\text{K}$, in MeOH. Left β-HAla$_8$ and right α-Ala$_{10}$. The plotted H-bond distance is the O-H distance. Each color corresponds to one of the six H-bonds shown in Figure 4.1. The black horizontal lines mark the threshold used to define a H-bond was opened
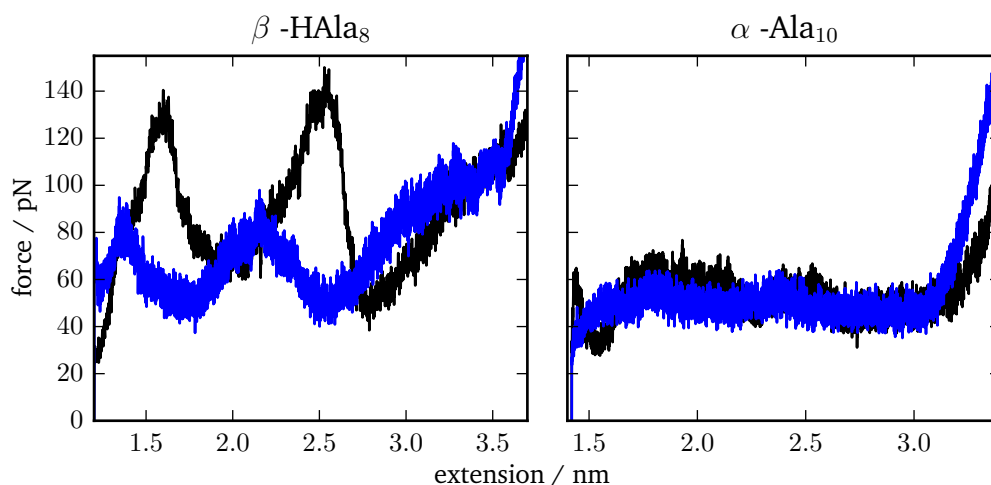
Fig. 4.4.: DS using $K = 1000\,\text{pN/nm}$ and $V = 0.1\,\text{nm/ns}$ at $T = 240\,\text{K}$. Left β-HAla$_8$ and right α-Ala$_{10}$. Black in MeOH and blue in water.

Furthermore, HCG showed that using the helix content as a second reaction coordinate results in an energy minimum corresponding to the α-helix, which is slightly more stable ($\sim 0.4\,\text{kcal} \cdot \text{mol}^{-1}$) than the local minimum corresponding to the unfolded conformation.

In this work, only the end-to-end distance was used as reaction coordinate to allow for a direct comparison with the results of the FPMD simulations. In Figure 4.6, the PMFs at $T = 300\,\text{K}$ (left) and $T = 240\,\text{K}$ (right), for α-Ala$_{10}$ and β-HAla$_8$ are shown. All the PMFs obtained for α-Ala$_{10}$ at $T = 300\,\text{K}$ have a similar shape and show a single minimum at an end-to-end distance that corresponds to an unfolded conformation. The similarities between the three PMFs indicate that there is no significant effect on the PMF by the change of the polarity of the solvent or by the presence of capping groups. The calculated PMFs are consistent with the results of HCG. Moreover, the fact that these results, obtained using the

Fig. 4.5.: ANC using $K = 1000\,\text{pN/nm}$ and $V = 0.1\,\text{nm/ns}$ at $T = 240\,\text{K}$. Left β-HAla$_8$ and right α-Ala$_{10}$. Black in MeOH and blue in water.

GROMOS53A6FF, are consistent with the ones from HCG obtained using the CHARM36 FF, further validates the FF used in this work.
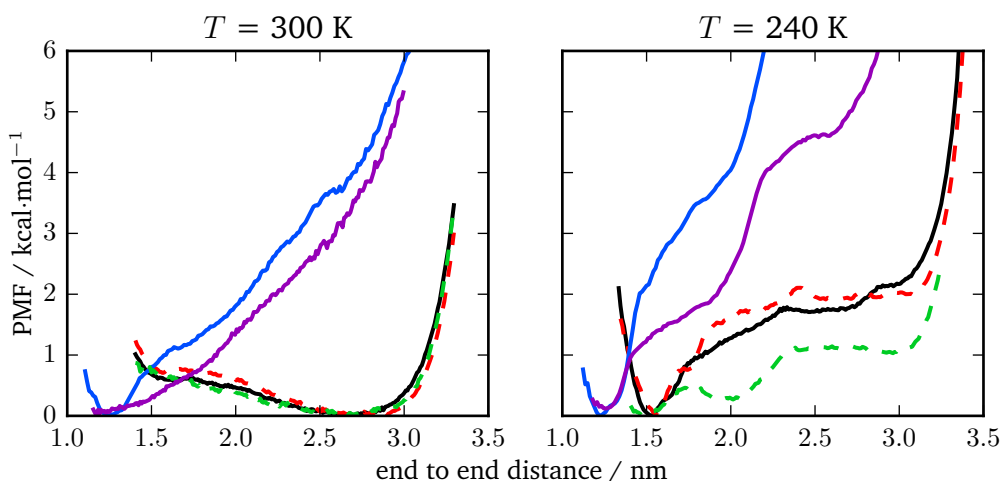
Fig. 4.6.: PMFs at $T = 300\,\text{K}$ (left) and $T = 240\,\text{K}$ (right). In the umbrella sampling a force constant of $K = 1000\,\text{pN/nm}$ was used. Full lines: Blue: β-HAla$_8$ in MeOH. Purple: β-HAla$_8$ in water. Black: α-Ala$_{10}$ in MeOH. Dashed lines: α-Ala$_{10}$ with an acetylated N-terminus and amidated C-terminus. Red: in MeOH. Green: in water.

The PMFs of α-Ala$_{10}$ at $T = 240\,\text{K}$ are different compared to the ones at $T = 300\,\text{K}$. At 240 K, the PMFs show a minimum at an extension that corresponds to the α-helix. The PMF in water (green line in Figure 4.6) additionally shows two energy barriers at $\sim 1.7\,\text{nm}$ and 2.4 nm and an intermediate state at $\sim 2.0\,\text{nm}$. By studying the conformations adopted during the US simulations, it was determined that the intermediate found in the α-Ala$_{10}$'s PMF in water corresponds to a half unfolded helix, in which some of the H-bonds are open. The α-helix is only more stable than the unfolded conformation by $\sim 1\,\text{kcal}\cdot\text{mol}^{-1}$ in water

and by $\sim 2\,\text{kcal} \cdot \text{mol}^{-1}$ in MeOH. The low stability of the α-helix formed by α-Ala$_n$ oligomers in water has also been observed experimentally. NMR studies [68] have indicated that short polyalanine peptides (α-Ala$_n$, $n = 3-9$) do not exist in water primarily as an α-helix but as polyproline II helix-like structures.

β-HAla$_8$'s PMF at $T = 300\,\text{K}$ in MeOH exhibit a global minimum at an extension that corresponds to the 3_{14}-helix and two not clearly defined shoulders at $\sim 1.5\,\text{nm}$ and $2.5\,\text{nm}$. At $T = 300\,\text{K}$, the PMF for β-HAla$_8$ in water shows a minimum at the extension corresponding to the 3_{14}-helix and no shoulders. However, at $T = 240\,\text{K}$, both in water and in MeOH, the PMF of β-HAla$_8$ has, beside the 3_{14}-helix minimum, two shoulders (*i.e.*, two energy barriers) at $\sim 1.5\,\text{nm}$ and $2.5\,\text{nm}$ and an intermediate state at $\sim 2.0\,\text{nm}$. In water, the energy barriers are lower, as expected given the experimental observation that β-peptide's helices are more stable in MeOH than in water [75]. The comparison of the PMFs for β-HAla$_8$ and α-Ala$_{10}$ shows that the 3_{14}-helix adopted by β-HAla$_8$ has a higher thermal stability than the α-helix adopted by α-Ala$_{10}$.

The calculated PMFs are consistent with the results of the FPMD simulations. As predicted from the DS, AHDC, and the ANC, at $T = 240\,\text{K}$ β-HAla$_8$ unfolds in a two-step fashion, passing through a stable intermediate conformation. In contrast, under the same conditions α-Ala$_{10}$ unfolds in a single step. The low stability of the α-helix formed by α-Ala$_{10}$, even at $T = 240\,\text{K}$, explains the absence of rupture events in α-Ala$_{10}$'s DS. Finally, the consistency between the results obtained from the calculation of the PMFs and those of the FPMD simulations shows the agreement between the observations made under non-equilibrium conditions with those made under quasi-equilibrium conditions.

4.4 Temperature effects in the unfolding pathways

The significantly different shape of α-Ala$_{10}$ PMFs at $T = 300\,\text{K}$ and $T = 240\,\text{K}$ indicates that the unfolding pathway of this oligomer depends on its thermal stability. In particular, the α-Ala$_{10}$ unfolding pathway is expected to be more complex at $T < 240\,\text{K}$, because the stability of the α-Ala$_{10}$'s helix is increased at lower temperatures. Thus, 100 FPMD simulations at $T = 200\,\text{K}$ were performed for α-Ala$_{10}$ and β-HAla$_8$ to study the effect of the temperature on their unfolding pathways. In this case, α-Ala$_{10}$ was studied only in water, whereas β-HAla$_8$ was studied only in MeOH. This choice of solvent was based on the higher stability of the α-helices and the 3_{14}-helices in the respective solvents [76].

Sample FE curves for both oligomers are shown in Figure 4.7. The FE curve for β-HAla$_8$ at $T = 200\,\text{K}$ is equivalent to that at $T = 240\,\text{K}$, being the only difference that at the lower temperature the rupture events are better defined because of the lower thermal fluctuations in the force. In contrast, the FE curve for α-Ala$_{10}$ is significantly different at $T = 200\,\text{K}$ in comparison to the one at $T = 240\,\text{K}$. At $T = 200\,\text{K}$, the opening of the H-bonds coincides with rupture events in the FE curves, whereas at $T = 240\,\text{K}$ not only the opening of the H-bonds does not coincide with rupture events, but there are no distinguishable rupture events.

The DS and AHDC for both oligomers, at $T = 200\,\text{K}$, are presented in Figure 4.8 and Figure 4.9, respectively. In β-HAla$_8$'s DS (Figure 4.8), there are two maxima and the first maximum has a shoulder. By comparing the extension at which the rupture events appear in the DS

and the extension at which the H-bonds open in the AHDC (Figure 4.9), the unfolding pathway of β-HAla$_8$ is found to consist of three steps: i) breaking of the N-terminus H-bond, ii) breaking of the outermost H-bonds (2, 5, and 6) and iii) breaking of the innermost H-bonds (3 and 4). Almost the same pathway was observed for β-HAla$_8$ at $T = 240$ K (see Section 4.2), the only difference is that at 240 K the breaking of the N-terminus H-bond was not observed, because of its low stability.

In an analogous way, the unfolding pathway of α-Ala$_{10}$ is found to consist of two steps: i) breaking of the outermost H-bonds (1,5,6) and ii) breaking of the innermost H-bonds (2, 3, and 4). This pathway is considerably different to the one observed at $T = 240$ K (see Section 4.2). At 240 K, the helix formed by α-Ala$_{10}$ was found not to be stable enough to observe any rupture events in the FE curves and the DS.
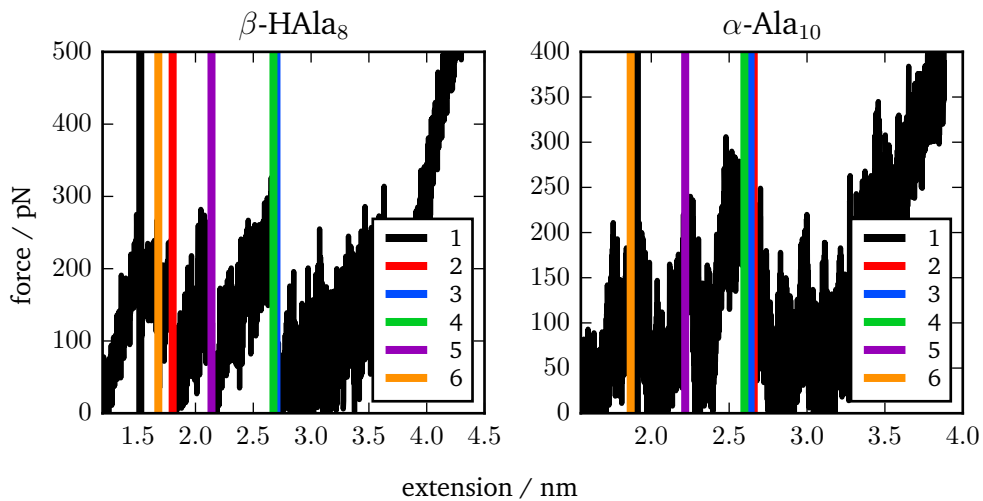
Fig. 4.7.: Typical FE curve using $K = 1000$ pN/nm and $V = 0.1$ nm/ns at $T = 200$ K, in MeOH for β-HAla$_8$ and in water for α-Ala$_{10}$. Vertical lines mark the opening the H-bonds. The color code corresponds to the one shown in Figure 4.1. A and B, respectively.
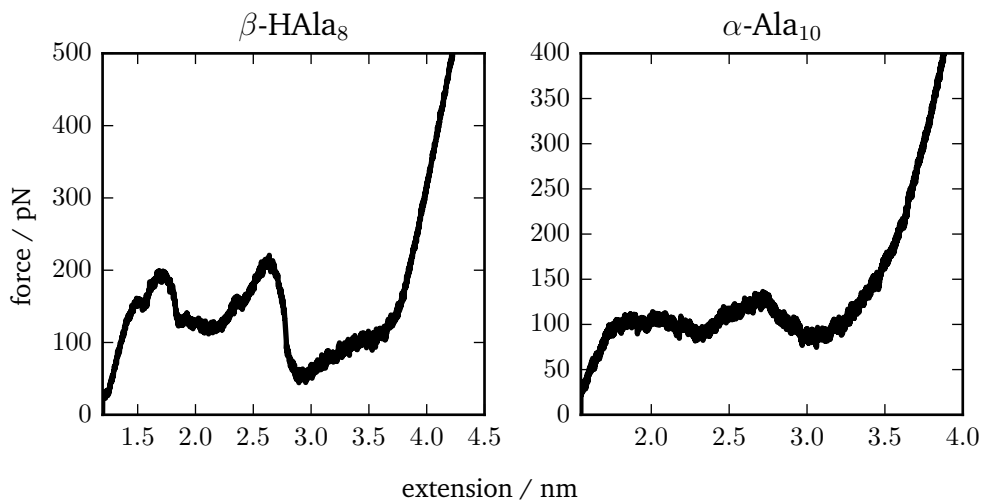

Fig. 4.8.: DS from 100 FPMD simulations, with $K = 1000$ pN/nm and $V = 0.1$ nm/ns at $T = 200$ K, in MeOH for β-HAla$_8$ and in water for α-Ala$_{10}$.

β-HAla$_8$'s and α-Ala$_{10}$'s DS have a similar shape, but the maxima in α-Ala$_{10}$'s DS are broader and not as well defined in comparison to the maxima in β-HAla$_8$'s DS. This difference is a consequence of the lower stiffness of α-Ala$_{10}$'s backbone. The effect of the backbone stiffness on the unfolding pathways is also evidenced in the ANCs at $T = 240\,\text{K}$ and $T = 200\,\text{K}$ (Figures 4.3 and 4.10). The ANCs of β-HAla$_8$ show a two-step decay, which implies a non-gradual increase in the distances between native contacts and a low flexibility of the backbone, because a stiff backbone reduces the fraction of conformational space that can be accessed by a system at a given temperature. In contrast, the ANCs of α-Ala$_{10}$ are smooth and show a constant decay, which implies a gradual increase in the distances between native contacts and a high flexibility of the backbone.
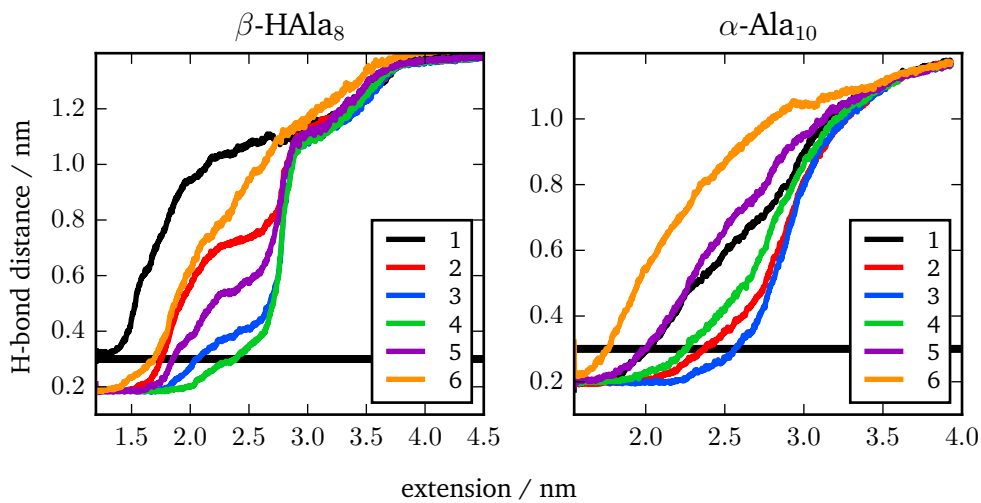
Fig. 4.9.: AHDC from 100 FPMD simulations, using $K = 1000\,\text{pN/nm}$ and $V = 0.1\,\text{nm/ns}$ at $T = 200\,\text{K}$, in MeOH for β-HAla$_8$ and in water for α-Ala$_{10}$. The H-bond distance correspond to the O-H distance. Each color corresponds to one of the H-bonds shown in Figure 4.1. The black horizontal lines mark the threshold (0.3 nm) used to decide that a H-bond is opened.
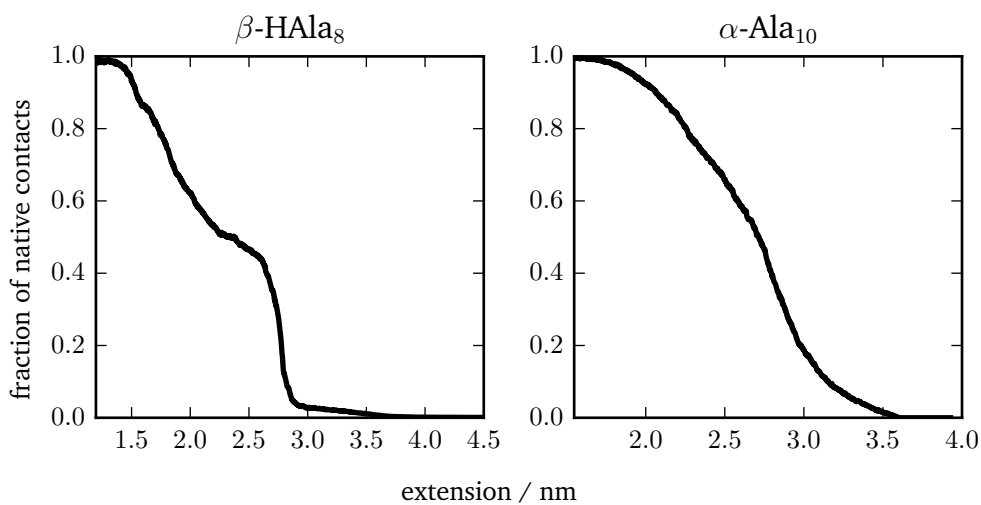
Fig. 4.10.: ANC using $K = 1000\,\text{pN/nm}$ and $V = 0.1\,\text{nm/ns}$ and $T = 200\,\text{K}$, in MeOH for β-HAla$_8$ and in water for α-Ala$_{10}$.

The fact that only the unfolding pathway of α-Ala$_{10}$ significantly differs between $T = 200\,\text{K}$ and $T = 240\,\text{K}$ indicates that the mechanical unfolding pathway of oligomers with low thermal stability changes more drastically when increasing the temperature. This hypothesis is further supported by the previous study of the heptamer, in which the unfolding pathway was found not to change at temperatures ranging between $200\,\text{K}$ and $360\,\text{K}$ (see Section 3.3).

4.5 Conclusions

In this chapter, a comparative study of the mechanical unfolding pathway of α- and β-peptides, in methanol and in water, using FPMD is presented. α-Ala$_{10}$'s α-helix and β-HAla$_8$'s 3_{14}-helix are used as model systems. The analysis of FE curves at $T = 240\,\text{K}$ and $T = 200\,\text{K}$ indicates that, despite that the two helices have a similar size and shape and the same number of H-bonds, the 3_{14}-helix has an overall higher thermal and mechanical stability. The lower thermal stability of the α-helix is further evidenced in the PMFs calculated at $T = 300\,\text{K}$ and $240\,\text{K}$. At $T = 300\,\text{K}$, in water and in MeOH, the PMFs only predict the 3_{14}-helix to be stable, as under these conditions a minimum corresponding to the α-helix is not observed. At $T = 240\,\text{K}$, both helices are found to be stable. However, the α-helix is only $\sim 1.0\,\text{kcal} \cdot \text{mol}^{-1}$ to $2.0\,\text{kcal} \cdot \text{mol}^{-1}$ lower in energy than the unfolded conformation.

The calculation of average properties, such as the DS and ANC, shows that the unfolding pathways of the two helices at $T = 240\,\text{K}$ are significantly different. The β-HAla$_8$'s 3_{14}-helix is found to unfold in a two-step fashion: first the outer H-bonds open and then the inner ones. In contrast, α-Ala$_{10}$'s α-helix is shown to unfold in a single step: the unfolding starts at the C-terminus and propagates fast to the N-terminus. However, at $T = 200\,\text{K}$, the unfolding pathway of both oligomers is shown to be similar, it starts at the termini and propagates to the center. The factors determining the way in which the two helices unfold are further discussed in Chapter 6.

Chain length dependence of the unfolding pathway of β-peptides

<div style="text-align: right">5</div>

In the previous two chapters, the unfolding pathways of small β- and α-peptides were investigated using FPMD. One of the most interesting results of these studies is the observation that the unfolding of small β-peptides starts at the helix termini and propagates to the helix center. However, the fact that two small β-peptides (the heptamer and the β-HAla$_8$) unfold in this manner does not imply that any β-peptide will show the same behavior. For this reason, a study of the chain length dependence of the mechanical unfolding pathway of β-peptides was carried out and the results are presented in this chapter. A short description of the systems can be found in Section 5.1. In Section 5.2, the results of the FMPD simulations and the calculated PMFs are presented and discussed. Finally, the conclusions of this chapter are given in Section 5.4.

5.1 Description of the model systems
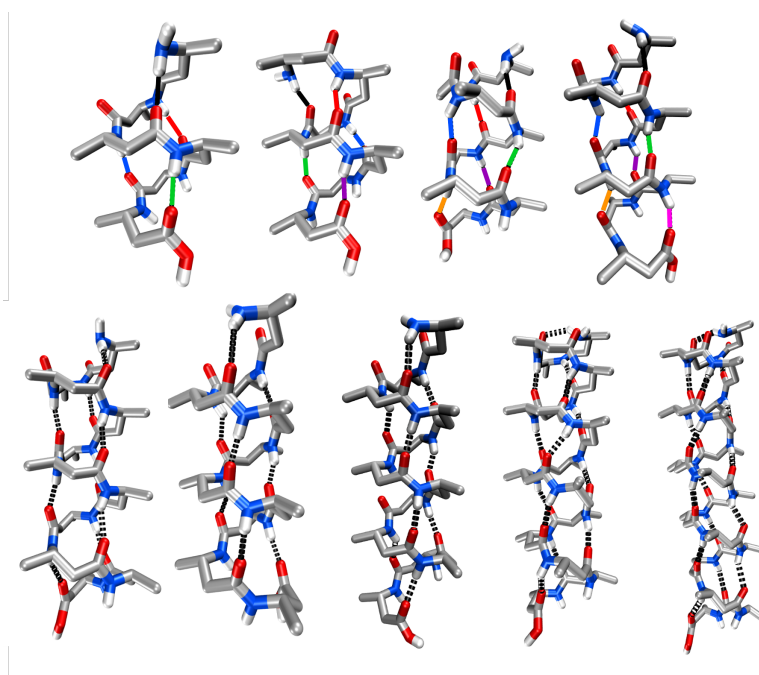
Fig. 5.1.: 3_{14}-helix conformation of β-HAla$_n$. Up: from left to right $n = 6,7,8,9$. Bottom: from left to right $n = 10,11,12,15,18$.

A series of β-HAla$_n$ oligomers with $n = 6$-12,15 and 18 were studied. These oligomers are constituted by β-HAla, which is the smallest β-amino-acid that promotes the 3_{14}-helix formation. The helix conformation adopted by each oligomer is depicted in Figure 5.1.

Since each turn of a 3_{14}-helix is constituted by three monomers, the number of turns in each oligomer is equal to $n/3$. Thus, only β-HAla$_n$ with $n = 6,9,12,15$, and 18 form helices with an integer number of turns. Additionally, as the 3_{14}-helix is stabilized by H-bonds of the type C=O(i) $\rightarrow$ H-N($i-2$), each β-HAla$_n$'s helix has n-2 H-bonds.

5.2 FPMD simulations

FPMD simulations of the β-HAla$_n$ oligomers were performed in MeOH, at $T = 240$ K, with $K = 1000$ pN/nm and $V = 0.1$ nm/ns. As in the case of the previous systems, 100 trajectories were calculated. To determine the order in which the H-bonds open for each β-HAla$_n$, the AHDCs were calculated. In Figure 5.2, the AHDCs for β-HAla$_n$ with $n = 6$-9 are depicted. The AHDCS show that the unfolding of all the oligomers starts at the termini and propagates to the center of the helix. For the remaining β-HAla$_n$ oligomers the same order was observed.

From the 100 simulations of each oligomer the DSs were obtained, shown in Figure 5.3. In this figure, an increasing number of maxima in the DS is observed with increasing chain length. Moreover, if the number of maxima in each DS is compared to the number of
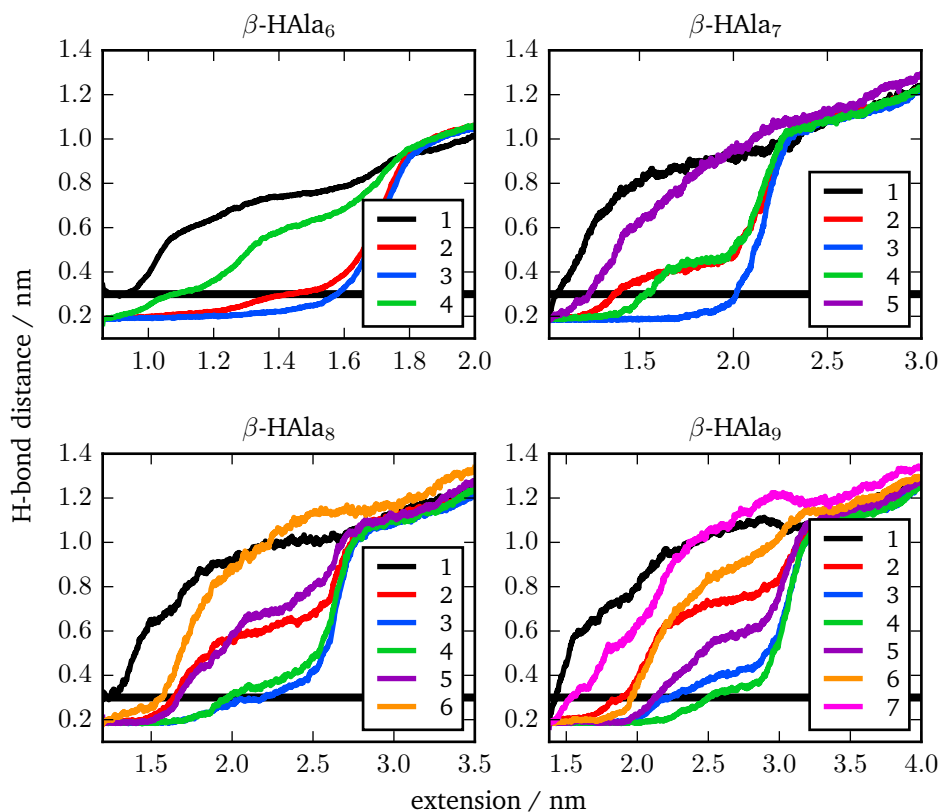
Fig. 5.2.: Average H-bond distance vs extension curves using $K = 1000$ pN/nm and $V = 0.1$ nm/ns at $T = 240$ K. The H-bond distance is the O-H distance. Each color corresponds to one H-bond shown in Figure 5.1. A threshold of 0.3 nm for the O-H distance was used to decide when a H-bond was opened (black horizontal lines).
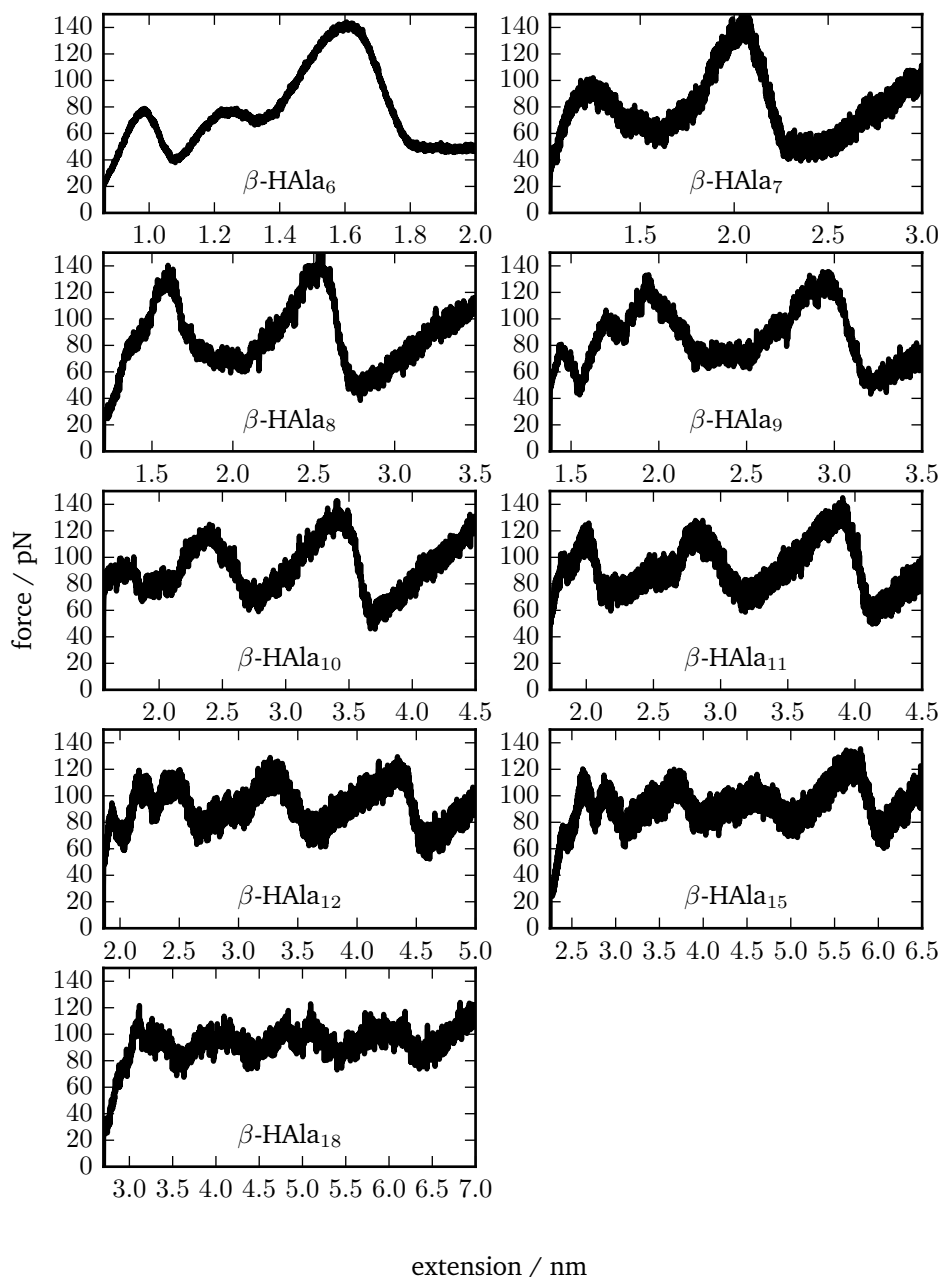
Fig. 5.3.: DS, *i.e.*, average over 100 FE curve for β-HAla$_n$ in MeOH, with $K = 1000\,\text{pN/nm}$ and $V = 0.1\,\text{nm/ns}$, at $T = 240\,\text{K}$.

complete turns in each helix, a clear trend is found. For oligomers with integer number of turns, there are $m + 1$ maxima with m the number of turns. In contrast, for oligomers with non-integer number of turns there are m maxima.

The DS of the oligomers with integer number of turns (*e.g.*, β-HAla$_6$) has one more maximum than the DS of the oligomers with non-integer number of turns (*e.g.*, β-HAla$_7$ and

β-HAla$_8$). By comparing the AHDCs of β-HAla$_6$ and β-HAla$_9$ with their DSs, it is found that the first maximum in the DSs corresponds to the breaking of the N-terminus' H-bond. The same is observed for β-HAla$_n$ with $n = 12$, 15, and 18. In contrast, the DSs for β-HAla$_n$ with $n = 7$, 8, 10, and 11 do not show a maximum corresponding to the breaking of the N-terminus' H-bond. Therefore, the extra maximum observed in the oligomers with integer number of turns is related to the opening of the N-terminus' H-bond. The strong correlation between the number of turns and the occurrence of the maximum related to the N-terminus indicates that the N-terminus H-bond is more stable in helices with integer number of turns. Not only the order in which the H-bonds open is the same for all oligomers, but the highest forces reached in the DSs are also similar, *i.e.*, ~ 120 pN to 140 pN. This indicates that the strength of all H-bonds, except the N-terminus ones, is independent of the chain length.

The dependence of the N-terminus H-bond stability on the number of turns of the helix may be related to the helix's dipole moment. Allison *et al.* [76] studied the dipole moment of the 3_{14}-helix adopted by β-peptides using MD simulations. Their calculations showed that the direction of the peptide bond dipole moment vector changes depending on the position of the bond in the helix. While the dipole moment of the central peptide bonds are parallel to the helix axis, the dipole moment vectors begin to point outward from the helix axis when moving away from the central residues. As a consequence, the sum of the peptide bond dipole moments is zero for a helix with integer number of turns, because the orientations of the dipole moment vectors cancel. In contrast, for a helix with non-integer number of turns the orientations of the dipole moment vectors do not cancel and the dipole moment is different from zero.

5.3 PMFs

The PMFs for all β-HAla$_n$ oligomers were calculated using US+WHAM at $T = 240$ K in MeOH, using the end-to-end distance as reaction coordinate. In Figure 5.4, the PMFs for β-HAla$_n$ with $n = 6$-9 are shown. These PMFs are consistent with the results of the FPMD simulations. For the three smallest oligomers there is an energy barrier in the PMF for each maximum in the DS. The PMF of β-HAla$_9$ shows almost a continuous line with some apparent shoulders. The almost vanishing energy barriers for β-HAla$_9$ are a hint that the longer the chain are, the smaller the barriers become. The PMF for β-HAla$_n$ with $n = 10$-12, 15, and 18 are not shown, because these PMF are mainly convex curves with small fluctuations which do not allow for a clear identification of any energy barrier.

For 6, 7, and 8 monomers, the extensions at which the maxima are found in the DS are consistent with the end-to-end distance at which the maxima are found in the PMFs. For example, for β-HAla$_6$, the three maxima in the DS are at ~ 1.0, 1.2, and 1.6 nm; in agreement the maxima in the PMF are at ~ 0.9, 1.2, and 1.5 nm. This result is remarkable because it shows that during the pulling simulation the oligomers follow the umbrella potential closely. Otherwise, the maxima in the DSs would be at larger extensions than the maxima in the PMF. This fact is also reflected in the average end-to-end distance vs extension curves shown in Figure 5.5. These curves are almost straight lines with slope 1. Hence, during the pulling simulations the end-to-end distance and the extension increase almost in the same way.
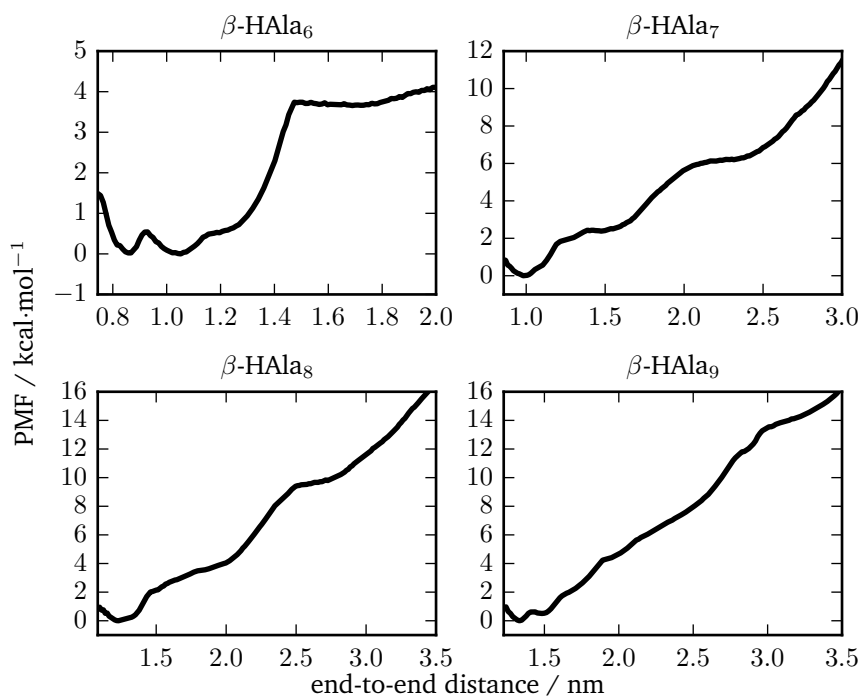
Fig. 5.4.: PMF at $T = 240\,\text{K}$ for β-HAla$_n$ with $n = $ 6-9 in MeOH. In the umbrella sampling, a force constant of $K = 1000\,\text{pN/nm}$ was used.
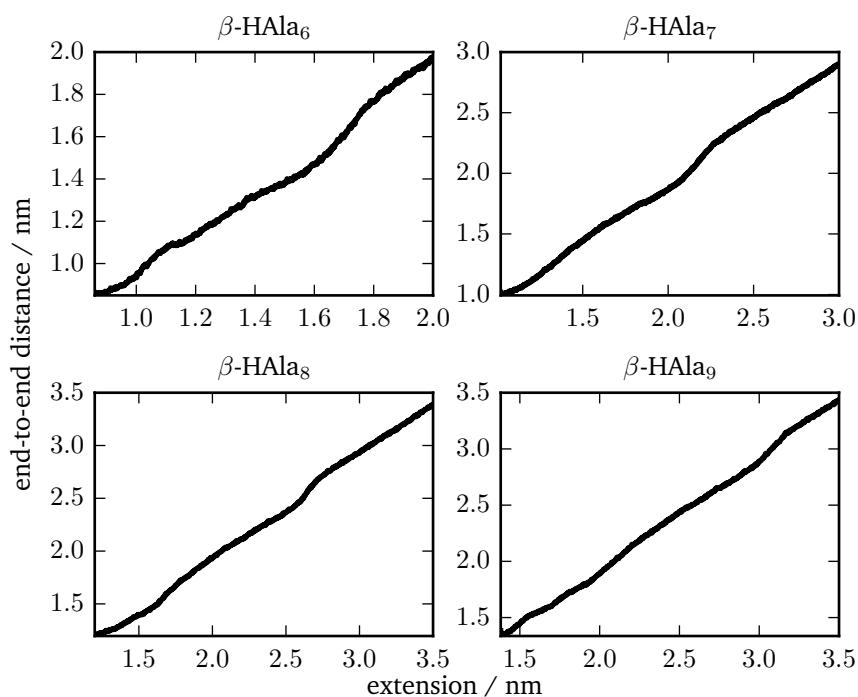
Fig. 5.5.: Average end-to-end distance vs extension curve for β-HAla$_n$, $n = $ 6-12,15,18 in MeOH, with $K = 1000\,\text{pN/nm}$ and $V = 0.1\,\text{nm/ns}$, at $T = 240\,\text{K}$.

5.4 Conclusions

In this chapter, the chain-length dependence of the unfolding pathway of β-peptides is studied, by performing FPMD simulations and calculating the PMFs of β-HAla$_n$ with $n = 6$-12,15,18. It is found that in the unfolding pathway of oligomers with integer number of helix's turns (β-HAla$_n$ with $n = 6,9,12,15,18$) there are $m + 1$ energy barriers, with m being the number of turns. m of the $m + 1$ energy barriers correspond to the unfolding of each turn of the helix. The additional energy barrier corresponds to the breaking of the N-terminus H-bond. In contrast, it is found that for oligomers with non-integer number of helix's turns (β-HAla$_n$ with $n = 7,8,10,11$) there are only m energy barriers, each of them corresponding to the unfolding of one helix turn. For these oligomers, the N-terminus H-bond opening is not observed. The fact that the number of energy barriers not only depends on the chain length but also on the completeness of the helix's turns indicates that the N-terminus H-bond is more stable in helices with integer number of turns.

Finally, the results of the FPMD simulations are found to be consistent with the PMFs of β-HAla$_n$ with $n = 6$-9, showing again that the FPMD simulations provide information about the energy landscape and unfolding pathways of short peptides, which is not only valid in non-equilibrium conditions but also in equilibrium.

Determining factors for the unfolding pathway of peptides, peptoids, and peptidic foldamers

<div align="right">

6

</div>

In the previous chapters, the unfolding pathways of α- and β-peptides have been extensively investigated. It has been shown that already short peptides exhibit a complex unfolding pathway that is shaped by the intrinsic stability of the folded conformations and their interaction with the solvent. However, a more comprehensive understanding of the impact of the backbone architecture on the shape of the unfolding pathway is missing.

In this chapter, general rules for the prediction of the unfolding pathway of oligomers are deduced by revealing the determining factors of this process for a set of oligomers using FPMD simulations. In Section 6.1, the main features of the backbone architecture of the studied oligomers are described. The results of the FPMD simulations are shown and analyzed in Sections 6.2-6.5. In Section 6.6, the mentioned rules are introduced. Finally, Section 6.7 summarizes the findings.

6.1 Description of the model systems

Five different classes of oligomers were studied: α-peptides, β-peptides, α/γ-peptides, δ-aromatic-peptides, and β-peptoids. From each class, a model oligomer was chosen based on two criteria: i) its chain length was short enough to allow extensive FPMD simulations and long enough to adopt a folded conformation. ii) Experimental observations have shown that the oligomer adopts a stable helix conformation (see references [57, 77–79]). Relevant structural information about the five studied oligomers is presented in Table 6.1 and a schematic representation of each helix conformation is shown in Figure 6.1.

The β-peptoid's helix was chosen because among the five helices, it is the only one that is not stabilized by intra-molecular H-bonds. However, the H-bonds in the remaining four helices are not of the same type and are not equally stable: δ-Chin$_8$'s H-bonds are of the type N $\rightarrow$ H-N and β-HAla$_8$'s, α-Ala$_{10}$'s, and the α/γ-peptide's H-bonds are of the type C=O $\rightarrow$ H-N. The N $\rightarrow$ H-N H-bonds are less stable, because of the acceptor atom's lower electronegativity. Although β-HAla$_8$, α-Ala$_{10}$, and the α/γ-peptide have the same type of H-bonds, the size of the ring formed by their H-bonds is different, *i.e.*, the ring is constituted by a different number of atoms. β-HAla$_8$'s H-bonds form a ring of 14 atoms, α-Ala$_{10}$'s H-bonds form a ring of 13 atoms, and the α/γ-peptide's H-bonds form rings of 10 or 12 atoms. In general, H-bonds forming smaller rings have a lower stability [78], because of their higher geometrical constraints. Therefore, the H-bonds in β-HAla$_8$ and α-Ala$_{10}$ are expected to be
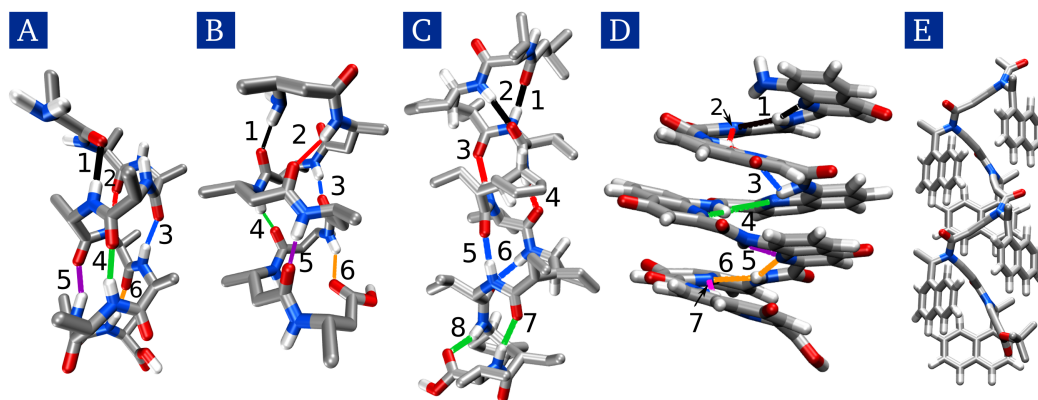
Fig. 6.1.: Schematic representation of: A) the α-helix of α-Ala$_{10}$. B) The 14-helix of β-HAla$_8$. C) The 12/10-helix of α/γ-peptide [78]; The H-bonds 1, 3, 5, and 7 are of the type C=O(i) →H-N($i+3$) and the H-bonds 2, 4, 6, and 8 are of the type C=O(i) →H-N($i-1$). D) The δ-Chin$_8$'s helix [79]; the substituents of the aromatic rings are not shown for simplicity. E) The β-peptoid's helix [77]. The numbered lines represent the H-bonds that stabilize the helix conformations.
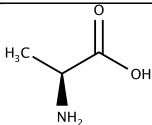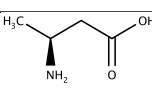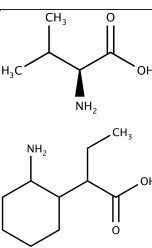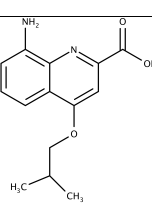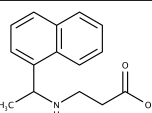
the most stable ones, followed by the α/γ-peptide's H-bonds, and the least stable H-bonds are those occurring in δ-Chin$_8$'s helix.

The backbone rigidity of the five helices is significantly different, because their backbone dihedral angles are not equally stiff. The constraints exerted by the amino acids' substituents are the main factor that affects the stiffness of the backbone dihedral angles. From the five oligomers, δ-Chin$_8$ is the one having the most constraining substituents, *i.e.*, aromatic rings, followed by the α/γ-peptide, which is half constituted by cyclic-γ-amino acids. The third one is the β-peptoid because of the large aromatic substituents. Finally, β-HAla$_8$ and α-Ala$_{10}$ do not have specially constraining groups. However, previous MD studies of Keller *et al.* [26] showed that, although β-peptides' backbones have four different types of dihedral angles, only one of them is a real degree of freedom, *i.e.*, only one of them can rotate freely. In contrast, it is common knowledge that in α-peptides only the dihedral angle associated to the peptide bond is highly constrained. Thus, in α-peptides two out of three dihedral angles are real degrees of freedom. Therefore, the β-peptides' backbone is stiffer than the α-peptides' backbone.

Another relevant difference between the oligomers are the sizes of their constituent monomers. α-Ala$_{10}$ has three backbone atoms per residue, β-HAla$_8$ and the β-peptoid have four, the α/γ-peptide has five, and δ-Chin$_8$ has six. In Section 6.5, the impact of the monomer size on the unfolding pathway of an oligomer is shown.

Finally, the handedness of the helices is another important feature to be compared. The handedness of the α-Ala$_{10}$'s, β-HAla$_8$'s, and α/γ-peptide's helices is determined by the chirality of their monomers, α-Ala$_{10}$'s helix is right handed and β-HAla$_8$'s and α/γ-peptide's helices are left handed. In contrast, the monomers of the other two oligomers, δ-Chin$_8$ and the β-peptoid, do not have backbone chiral centers, Hence, their right and left handed helices are energetically equal. In this work, δ-Chin$_8$'s left handed helix and the β-peptoid's right handed helix were studied. The choice of the handedness of these helices was arbi-

Tab. 6.1.: Structural information of the five studied oligomers.

Oligomer	Monomer/s	#[1]	H-bonds[2]	Nickname	Solvent
α-peptide		10	6 C=O (i) →H-N$(i+4)$	α-Ala$_{10}$	H_2O
β-peptide		8	6 C=O(i) →H-N$(i-2)$	β-HAla$_8$	CH_3OH
α/γ-peptide		9	4 C=O(i) →H-N$(i+3)$ 4 C=O(i) →H-N$(i-1)$	α/γ-peptide	$CHCl_3$
δ-aromatic-peptide		8	7 N(i) →H-N$(i+1)$	δ-Chin$_8$	$CHCl_3$
β-peptoid		6	-	β-peptoid	CH_3CN

[1] Number of residues. [2] Number and types of H-bonds; *e.g.*, a C=O (i) →H-N$(i+4)$ denotes a H-bond between the C=O from residue i and the H-N from residue $(i+4)$.

trary. However, tests made with δ-Chin$_8$'s right handed helix showed, as expected, that its behavior is identical to that of the left handed conformer.

From the five studied helices, four different unfolding pathways were identified. Each of them is presented separately in the following sections to emphasize their defining backbone features.

6.2 The β-peptoid - simultaneous unfolding of all turns

As mentioned in the previous section, the β-peptoid's helix is not stabilized by intra-molecular H-bonds. In fact, this oligomer was specially designed to adopt a helix conformation only stabilized by the high stiffness of the C_α-C-N-C_β-dihedral-angles [77], which is a consequence of the large and rigid substituents at the backbone N atoms. Through this chapter,

it will become evident how the lack of H-bonds makes the unfolding pathway of a helix fundamentally different from that of helices with H-bonds.

100 FPMD simulations were performed for the β-peptoid with $K = 1000\,\text{pN/nm}$ and $V = 0.01\,\text{nm/ns}$ at $T = 200\,\text{K}$. This temperature was used for all oligomers, because at these conditions non-thermal unfolding was observed. A typical FE curve (black) and the DS (red) for the β-peptoid are shown in Figure 6.2 (left). In both curves, at small extensions there is an interval ($\sim 2.0\,\text{nm}$ to $2.4\,\text{nm}$) in which the force is nearly constant. In this interval, several conformational changes take place without generating a rip in the force, indicating that the energy barrier associated with these events is practically negligible. At $\sim 4.2\,\text{nm}$ and at a force of $\sim 2000\,\text{pN}$ a second rupture event occurs. The large force suggests a high energy barrier associated with this event.
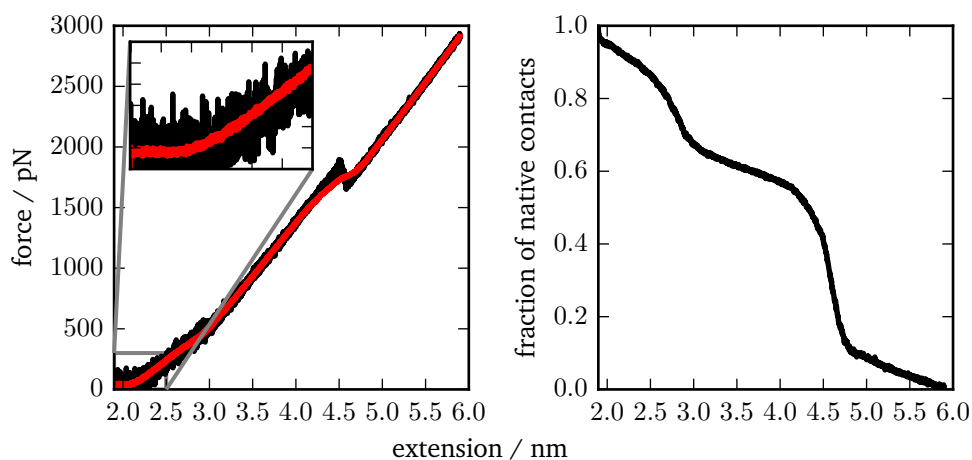
Fig. 6.2.: Results of the β-peptoid FPMD simulations using $K = 1000\,\text{pN/nm}$ and $V = 0.01\,\text{nm/ns}$ at $T = 200\,\text{K}$. Left: typical FE curve (black) and DS (red). Inset: zoom of the small force region. Right: ANC.

The two rupture events found in the DS are in agreement with the two step decay in the ANC, shown in Figure 6.2 (right). An ana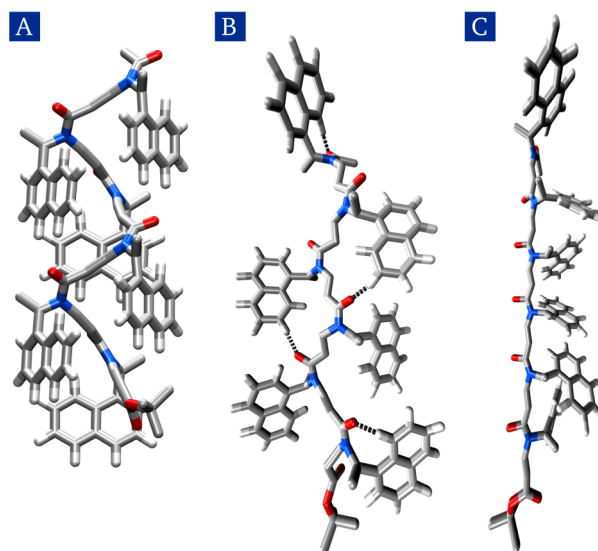lysis of the conformations adopted by the β-peptoid during a pulling trajectory shows that the first rupture event corresponds to the unfolding of the β-peptoid's helix (Figure 6.3.A) into the intermediate conformation (Figure 6.3.B). During this rupture event all turns of the helix unfold simultaneously through the rotation of most backbone bonds, except the N-CO bonds. The second rupture event corresponds to the transition of the β-peptoid from the intermediate (Figure 6.3.B) to the completely unfolded conformation (Figure 6.3.C), through a rotation of the N-CO bonds.

The high force ($\sim 2000\,\text{pN}$) at which the second rupture event occurs and the large constant interval between the two decays in the ANC indicate a highly stable intermediate. In addition, the large change in the fraction of native contacts (from 0.6 to almost 0) that takes place during the second rupture event suggests an abrupt transition from the intermediate to the unfolded state. The high stability of the intermediate state is mainly determined by the stiffness of the C_α-C-N-C_β-dihedral-angles. Additionally, the intra-molecular H-bonds formed between the carbonyl oxygens and the aromatic hydrogens further stabilize this conformation.

Fig. 6.3.: Schematic representation of the conformations adopted by the β-peptoid during its unfolding. A) Helix, B) intermediate, and C) unfolded conformation. The black dashed lines represent H-bonds.

The unfolding pathway of the β-peptoid can be illustrated by comparing its helix to a spring. When a spring is pulled, all turns of the spring are simultaneously stretched. Analogously, the turns of the helix unfold simultaneously through the rotation of most of the backbone bonds. However, the N-CO bonds do not rotate in the first unfolding event, because this rotation is hindered by the large aromatic substituents at the N atoms. Hence, the second unfolding event (rotation of the N-CO bonds) only occurs at high forces.
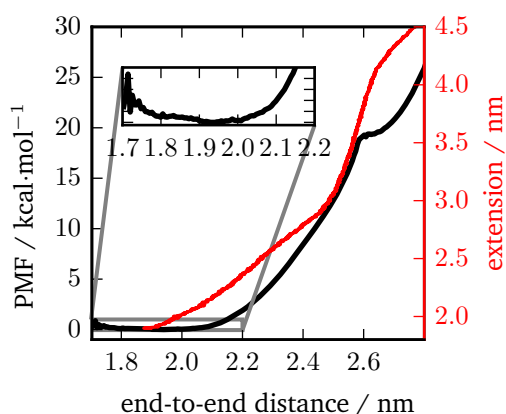
Fig. 6.4.: Black: PMF calculated with US+WHAM. The end-to-end distance is the distance between the reference and the pull atoms. Inset: small end-to-end distance region. Red: average end-to-end distance vs extension curve from the FPMD simulations.

The energy landscape of the unfolding pathway of the β-peptoid's helix was further investigated through the calculation of the PMF using US+WHAM (see Section 2.7). In the PMF,

shown in Figure 6.4, there is a very shallow minimum at $\sim 1.9\,\text{nm}$ (see inset Figure 6.4) and a shoulder, *i.e.*, an energy barrier, at $2.6\,\text{nm}$.

The comparison between the PMF and the results of the FPMD simulations is not straightforward. While the PMF is obtained under quasi-equilibrium conditions as a function of the end-to-end distance, the DS and ANC are obtained under non-equilibrium conditions as a function of the extension. However, the extensions at which the rupture events are observed in the DS and ANC can be translated to the end-to-end distance by plotting an average extension vs end-to-end distance curve (red line in Figure 6.4). Using this curve, the extensions at which the first rupture event occurs ($\sim 2.0\,\text{nm}$ to $2.4\,\text{nm}$) are found to correspond to the end-to-end distance at which the PMF is almost flat (see inset Figure 6.4). The lack of an energy barrier in the PMF associated to the first rupture event, indicates that under equilibrium conditions the β-peptoid can fluctuate freely between the helix and the intermediate conformations.

Analogously, the extension at which the second rupture event occurs ($4.2\,\text{nm}$) is found to correspond to the shoulder in the PMF. Hence, as expected given the shape of the DS, the transition from the intermediate to the unfolded conformation has a high energy barrier ($\sim 20\,\text{kcal} \cdot \text{mol}^{-1}$).

To further understand the stability of the β-peptoid's helix, temperature dependent MD simulations at $T = 200\,\text{K}$ and $298\,\text{K}$ were performed. The fraction of native contacts vs time curves from these trajectories are shown in Figure 6.5. At both temperatures, the β-peptoid fluctuates between several conformations but it never reaches a fully unfolded state, *i.e.*, the fraction of native contacts never goes under 0.6. This behavior shows that even at $T = 298\,\text{K}$ the β-peptoid cannot overcome the high energy barrier for the transition from the intermediate to the fully unfolded conformation. These results are in agreement with previous MD simulations of the β-peptoid with the OPLS force field [77].
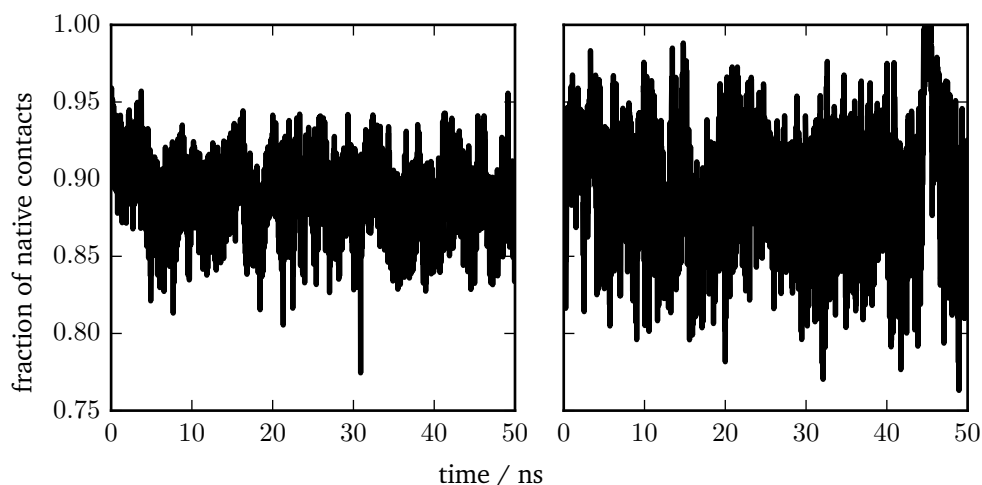
Fig. 6.5.: Fraction of native contacts vs time curves from MD simulations of the β-peptoid. Left: $T = 200\,\text{K}$ and right: $T = 298\,\text{K}$.

The results of FPMD, MD, and PMF are consistent, all indicate that the β-peptoid unfolds in a two-step fashion. Moreover, the energy barrier for the transition between the helix and

intermediate state is shown to be negligible. In contrast, the energy barrier of the transition between the intermediate and the unfolded state is found to be $\sim 20\,\text{kcal}\cdot\text{mol}^{-1}$. Therefore, only the second energy barrier contributes to the stability of the folded conformation. Thus, the β-peptoid's helix is not stable by itself, but the helix's stability is mainly determined by the stability of the intermediate state.

6.3 β-HAla$_8$ and α-Ala$_{10}$ - unfolding from the termini to the center

β-HAla$_8$'s and α-Ala$_{10}$'s unfolding pathways were studied at $T = 240\,\text{K}$ and $T = 200\,\text{K}$ in Chapter 4. In this chapter, the unfolding pathway at $T = 200\,\text{K}$ for both oligomers is further discussed. At $T = 200\,\text{K}$, β-HAla$_8$ and α-Ala$_{10}$ unfoldings start at the termini and propagates to the center. In order to understand why this is the preferred pathway for both oligomers, all other alternatives must be excluded. In the following, all possible pathways will be analyzed, focusing on β-HAla$_8$, because the same arguments apply to α-Ala$_{10}$.

A first possible pathway is the one found for the β-peptoid, which is the simultaneous unfolding of all turns. This pathway can be ruled out by revisiting the analogy of a spring being stretched. In the case of β-HAla$_8$'s helix, adjacent turns of the spring would be connected by H-bonds. Therefore, if all turns unfold simultaneously, all H-bonds will open also simultaneously, which requires a large amount of energy. Thus, under the simulated conditions this pathway is unlikely.

A second possible unfolding pathway starts with the opening of the H-bonds from one terminus propagating to the opposite terminus. Given that the same force is applied at both termini of the helix, this pathway would require that one of the terminus opens more easily than the other. Given that all H-bonds in β-HAla$_8$'s helix are equally strong, the opening of the H-bonds at one of terminus is not more favorable. Therefore, this case is also not probable.

Only two alternatives are left, a completely random unfolding and an unfolding that starts at the center of the helix and propagates to the termini. Both cases imply that a central H-bond can open before the outer ones. However, the study of the β-HAla$_8$ helix conformation shows that this opening is improbable. The opening of the central H-bond 3 (blue in Figure
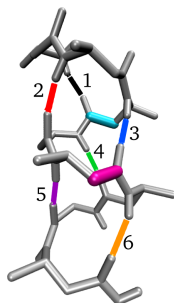

Fig. 6.6.: Schematic representation of β-HAla$_8$'s helix. The numerated colored lines represent H-bonds. The cyan and pink lines are C$_\alpha$-CO bonds.

6.6), requires the rotation of one of the two C_α-CO bonds (marked as cyan and pink in Figure 6.6). If the C_α-CO bond marked as pink rotates, both H-bond 4 (green) and 6 (orange) will open. On the other hand, the rotation of the C_α-CO bond marked as cyan would at least induce the opening of H-bond 2 (red). Thus, the opening of a central H-bond requires at least the opening of a second H-bond. Hence, more energy is required for the breaking of central H-bonds, in comparison to outer H-bonds. This argument is supported by the statistics of the H-bond breaking order. In 100 FPMD simulations, the opening of an inner H-bond before the opening of the outer H-bonds was not observed.

In conclusion, for β-HAla$_8$ and α-Ala$_{10}$, the most plausible unfolding pathway is the one observed in the FPMD. The termini H-bonds open first, followed by the opening of the adjacent H-bonds, and at the end the innermost H-bonds open.

6.4 The α/γ-peptide - unfolding from one terminus to the opposite one

In contrast to the other oligomers, the analysis of 100 FPMD simulations of the α/γ-peptide with $\mu = 100$ pN/ns ($K = 1000$ pN/nm and $V = 0.1$ nm/ns) was not conclusive. For this reason, 500 simulations with $\mu = 1 \times 10^3$ pN/ns ($K = 1000$ pN/nm and $V = 1$ nm/ns) were performed, where a higher velocity was used to lower the computational cost. The statistics from the 500 simulations showed that there are two equally probable unfolding pathways for the α/γ-peptide. The DSs and AHDCs for the two pathways are shown in Figure 6.7.
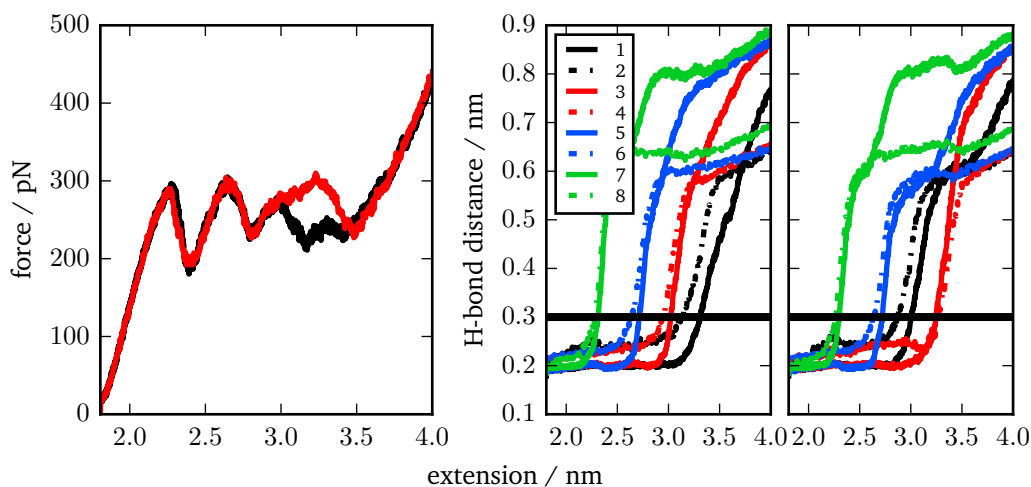
Fig. 6.7.: Results of the α/γ-peptide FPMD simulations, using $K = 1000$ pN/nm and $V = 1$ nm/ns at $T = 200$ K. Left: the DS for the two equally probable unfolding pathways. Center and right: AHDC, for the DS in black and the DS in red, respectively. The broken and continuous lines correspond to H-bonds of the type C=O(i) →H-N($i-1$) and C=O(i) →H-N($i+3$), respectively. The black horizontal lines mark the threshold (0.3 nm) used to decide when a H-bond is opened.

By comparing the DSs and the AHDCs, the first two steps of the two pathways are found equal: i) opening of H-bonds 8 and 7, ii) opening of H-bonds 6 and 5. The last two steps in one of pathway are: iii) opening of H-bonds 4 and 3, and iv) opening of H-bonds 2 and 1.

In the other pathway the order of the last steps is inverted and their maxima overlap giving rise to a single maximum (see red curve in Figure 6.7, left).

It is remarkable that in both pathways the α/γ-peptide unfolds one turn at a time through the almost simultaneous opening of two H-bonds, a C=O(i) →H-N($i-1$) (in short ($i-1$))) H-bond and a C=O(i) →H-N($i+3$) (in short ($i+3$)) H-bond. The opening of the pairs of H-bonds is accompanied by the rotation of the only flexible backbone bond of the γ-amino acids that constitutes the α/γ-peptide: the C_α-C_β bond. A similar unfolding pathway for another type of peptidic foldamer has been observed by Balamurugan *et al.* [80]. These authors studied the unfolding pathway of trans-$\alpha/\beta^{2,3}$-peptides using NMR spectroscopy and crystallography, finding that these peptides unfold one turn at a time, starting at the C-terminus and ending at the N-terminus.

The AHDCs curves (Figure 6.7) show that the α/γ-peptide unfolds following the path of the weaker H-bond, *i.e.*, at every rupture event the ($i-1$) H-bond is the first to open. The unfolding pathway of the α/γ-peptide can be explained by taking into account the relative strength of the H-bonds. Each of the amino acid in the α/γ-peptide forms one ($i+3$) H-bond (stronger) and one ($i-1$) H-bond (weaker). The pairs of H-bonds formed by an amino acids are shown with the same color in Figure 6.1. In the C-terminus, the outermost H-bond is the weaker one (H-bond 8), whereas in the N-terminus the outermost H-bond is the stronger one (H-bond 1). Thus, the unfolding starts at the terminus with the weaker H-bond: the C-terminus. The opening of the H-bond 8 is accompanied by a large rotation of the C_α-C_β bond in the 8th residue (see Figure 6.8). This rotation destabilizes the adjacent ($i+3$) H-bond (H-bond 7) and induces its opening. The following rupture events follow the same formula, a ($i-1$) H-bond opens and induces the opening of its adjacent ($i+3$) H-bond.
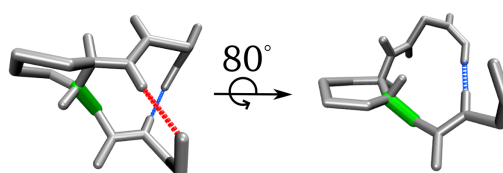
Fig. 6.8.: Schematic representation of the rotation of the C_α-C_β bond (green) within the 8th residue of the α/γ-peptide. The red and blue dashed lines represent the ($i-1$) H-bond and the ($i+3$) H-bond, respectively.

The H-bond pattern of the α/γ-peptide's helix has a second remarkable consequence, the turns of the helix can be seen as independent units, since the unfolding of a turn does not disrupt the adjacent turns. This implies that a central turn may unfold before the outer turns. However, this event has a low probability because when an inner C_α-C_β bond rotates all the outer part of the helix must rotate with it.

6.5 δ-Chin$_8$ - pulling velocity dependence

By comparing δ-Chin$_8$'s DS (Figure 6.9 left) and ANC (Figure 6.9 right) in an analogous way as for the other studied oligomers, the unfolding pathway of δ-Chin$_8$ is found to consist of four steps: i) opening of the N-terminus H-bond, ii) opening of H-bonds 2 and 7, iii) opening of H-bonds 3 and 4, and iv) opening of the H-bonds 5 and 6.
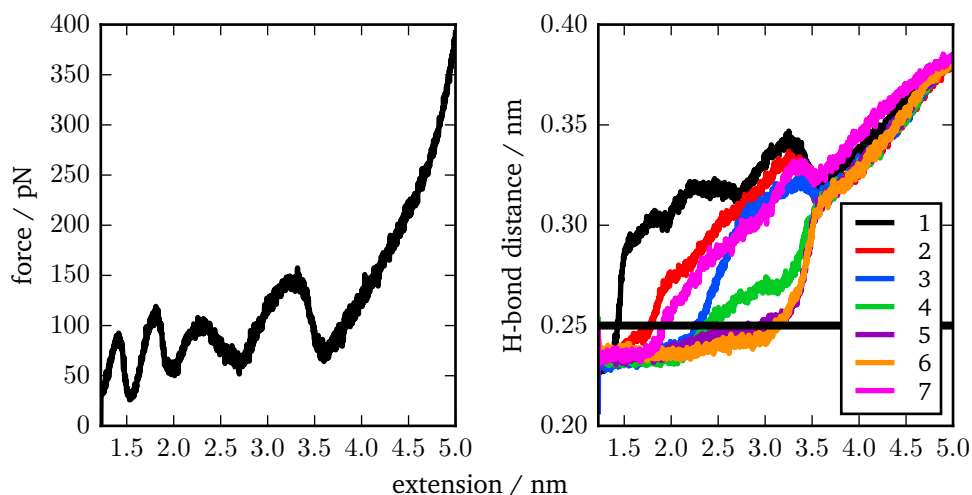
Fig. 6.9.: Left: DS and right: AHDC for δ-Chin$_8$, using $K = 1000\,\text{pN/nm}$ and $V = 0.1\,\text{nm/ns}$ at $T = 200\,\text{K}$. The black horizontal line marks the threshold (0.25 nm) used to decide when a H-bond is opened.

The anticipated high stiffness of δ-Chin$_8$'s backbone is reflected in the FPMD simulations, during its unfolding only one of the six different types of backbone bond rotates: the C$_\alpha$-CO bond. In agreement, Abramyan *et al.* [81] found in a recent metadynamics study, that during the handedness inversion of a series of δ-Chin$_n$ (n = 3,4,5,6) the only bonds that rotate are also the C$_\alpha$-CO bonds.

Despite the fact that δ-Chin$_8$'s helix is stabilized by only one type of H-bonds, its unfolding pathway is different than the one of β-HAla$_8$ and α-Ala$_{10}$. In the case of δ-Chin$_8$, the extremely stiff backbone, a large N to C asymmetry and a particular H-bond pattern, are the main factors that determine its unfolding pathway. In contrast to all the other studied H-bonded helices, the H-bonds in δ-Chin$_8$'s helix are not parallel to the helix axis, but they are parallel to the helix turns (see Figure 6.1). This particular pattern allows to see each H-bond as an independent unit, because the opening of a H-bond does not disturb the other H-bonds. Hence, the probability of a central H-bond opening before the outer ones is not zero.

The N to C asymmetry refers to the fact that the sequence of backbone atoms in δ-Chin$_8$ is inverted when seen from the C-terminus in comparison to when seen from the N-terminus. This asymmetry is larger for δ-Chin$_8$ than for β-HAla$_8$ and α-Ala$_{10}$, because each δ-Chin$_8$'s monomer has six backbone atoms, whereas each β-HAla$_8$'s and α-Ala$_{10}$'s monomer has four and three backbone atoms, respectively. A consequence of the N to C asymmetry is that the rotation of the C$_\alpha$-CO bond in δ-Chin$_8$'s N-terminus and C-terminus is not equal. The N-terminus's C$_\alpha$-CO bond rotation requires only the rotation of the aromatic ring and
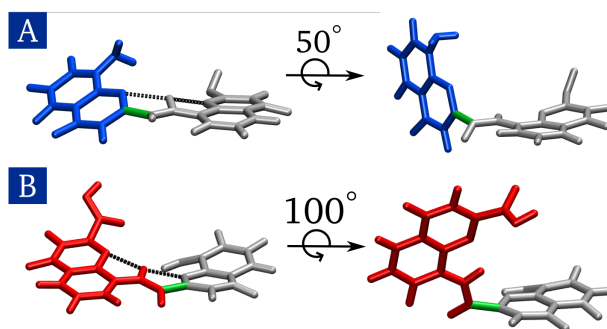
Fig. 6.10.: A) Schematic representation of the rotation of the C_α-CO (green) between δ-Chin$_8$'s N-terminus residue and 2nd residue. The group in blue is the one that rotates. B) Schematic representation of the rotation of the C_α-CO (green) between δ-Chin$_8$'s 6th residue and C-terminus residue. The group in red is the one that rotates. The black dashed lines represent H-bonds.

its substituents (shown in blue in Figure 6.10A). In contrast, the C-terminus's C_α-CO bond rotation requires the rotation of the whole C-terminus residue and the CO group of the 7th residue (shown in red in Figure 6.10B). The N-terminus rotation is energetically favored against the C-terminus rotation, because the rotation of a larger group requires more energy. Hence, the N-terminus H-bond opens first.

The effect of the asymmetry changes after the N-terminus opens, because the rotation of the C_α-CO bonds in the 2nd residue and in the C-terminus residue requires the movement of equivalent fragments of δ-Chin$_8$'s chain. This explains why the second unfolding event corresponds to the almost simultaneous opening of the H-bonds 2 and 7. However, if the size of the rotating fragments would be the only effect of the asymmetry, the opening of H-bonds 3 and 6 will be the third unfolding event and the opening of H-bonds 4 and 5 the fourth unfolding event. In contrast, the opening of H-bonds 3 and 4 is observed first and the opening of H-bonds 5 and 6 occurs at last. Therefore, there should be a second effect due to the N to C asymmetry.

The behavior of the rotation of the C_α-CO bonds was studied to reveal the nature of this second effect. It was found that when a H-bonds opens (except H-bond 7) the rotation around the C_α-CO bond shifts the C_β-C_α-CO-N-dihedral angle from 150° to 100° (*i.e.*, a 50° rotation), as shown in Figure 6.10.A. In contrast, when H-bond 7 opens, the rotation around the C_α-CO bond shifts the C_β-C_α-CO-N-dihedral angle from 150° to 50° (*i.e.*, a 100° rotation), as shown in Figure 6.10.B. The larger rotation accompanying the opening of H-bond 7 generates a conformation in which the opening of H-bond 6 is hindered. Hence, after the H-bond 7 opens the unfolding can only continue from the N-terminus side.

The effect of the N to C asymmetry was investigated in more detail by carrying out 100 FPMD simulations for δ-Chin$_8$ with $V = 10$ nm/ns or $V = 1$ nm/ns. Two interesting results where found in the statistical analysis of these simulations: i) the N-terminus is the first to open at all studied velocities, showing that the first effect of the N to C asymmetry does not depend on the pulling velocity. ii) For $V = 10$ nm/ns, the unfolding pathway is similar to the one of β-HAla$_8$ and α-Ala$_{10}$, from the termini to the center, showing that the second effect of the asymmetry is no longer observed at high pulling velocities.

The absence of the second effect due to the asymmetry at high pulling velocities is a consequence of the change in the behavior of the rotation of the C_α-CO bonds. At $V = 10$ nm/ns, the opening of the H-bonds is faster than the rotation of the C_α-CO bonds. This behavior does not allow the system to reach the conformation in which the C_β-C_α-CO-N-dihedral angle in the C-terminus takes a value of 50° (Figure 6.10B). Thus, at this velocity, the opening of H-bond 6 is not hindered by the opening of H-bond 7.

For the other oligomers, the effect of the pulling velocity on their unfolding pathways was also evaluated, with the same pulling velocity. For the other four studied oligomers no significant effect of the velocity was observed. Therefore, the different behavior of δ-Chin$_8$ is a consequence of the extra factor that affects its unfolding, the N to C asymmetry.

6.6 Rules for the prediction of the unfolding pathway of oligomers adopting a helix conformation

A series of rules for the prediction of the unfolding pathway of an oligomer can be derived from all results presented in the first part of this work. These rules depend on basic features of the oligomers backbone architectures, *e.g.*, the presence and stability of H-bonds, the H-bond pattern, the backbone stiffness and the N to C asymmetry. Hence, the following rules should be applicable to any oligomer that folds into a helix:

1. **All turns unfold simultaneously in helices without intra-molecular H-bonds (as the β-peptoid)**. The shape of the energy landscape of such a helices may be tuned by modifying the stiffness of the backbone's dihedral angles.

2. **Helices with one type of H-bonds (as α-Ala$_{10}$ and β-HAla$_8$) unfold starting at the termini and propagating to the center, independently of the backbone flexibility.** However, the overall stability of the helix and intermediate states does depend on the backbone flexibility.

3. **Helices with more than one type of H-bonds (as the α/γ-peptide) unfold following the path of the weaker H-bond.** Such a pathway might be tuned by changing the rigidity of the backbone.

4. **The unfolding pathway of helices with a stiff backbone and a large N to C asymmetry (as δ-Chin$_8$) might depend on the pulling velocity.** The effect of the N to C asymmetry is difficult to predict, because it depends on the system and the pulling velocity.

6.7 Conclusions

In this chapter, the unfolding pathway of four foldamers and one α-peptide are investigated using FPMD. The helix conformations of all studied oligomers are found to unfold in a cooperative manner, *i.e.*, every rupture event implies the opening of various H-bonds and/or the rotation of various backbone bonds. However, the backbone architecture of

each oligomer influences its unfolding pathway in a specific manner. For helices with intra-molecular H-bonds, the energy landscape is mainly shaped by an interplay between the H-bond stability and the backbone rigidity.

Four different unfolding pathways were identified from the five studied helices. The first pathway, follow by the β-peptoid, consists in the simultaneous unfolding of all the turns of the helix. In the second pathway, exhibited by β-HAla$_8$ and α-Ala$_{10}$, the helix unfolds starting from the termini and propagates to the center. In the third pathway, exhibited by the α/γ-peptide, the unfolding goes from one terminus to the other terminus. Finally, the fourth pathway, exhibit by δ-Chin$_8$, depends on the pulling velocity. At high velocities the unfolding is analogues to the one for β-HAla$_8$ and α-Ala$_{10}$ and at low velocities the unfolding is like the one for the α/γ-peptide.

Supported by the observations made in this work about the unfolding pathway of several oligomers, a set of rules for the prediction of the unfolding pathway of oligomers that adopt helix conformations are deduced and introduced in Section 6.6. These rules can be used for the design of new backbone architectures with predictable stability and mechanical properties.

QM/MM study of the metal specificity of isatin hydrolases from *Labrenzia aggregata*

<div style="text-align: right; color: red; font-size: 2em;">7</div>

Isatin is an indole derivative found in both prokaryotic and eukaryotic organisms [34] that is characterized by its high conjugation. In humans, isatin acts as an inhibitor of monoamine oxidase B and is the major component of the endogenous non-peptide inhibitor mix tribulin [82]. Isatin can be detected in blood, urine, and tissue and due to its hydrophobic nature it can cross the blood-brain barrier [83]. Although the neurological role of isatin is unclear, high levels of isatin have been reported in the blood of persons with Parkinson disease [83]. Isatin and its derivatives have been shown to possess a large range of biological and pharmaceutical applications (for some recent reviews see References [84, 85]). Among them, one of the most extensively studied is the antimicrobial activity [86–89].

Isatin hydrolases (IHs) constitute a class of metalloenzymes that catalyzes the hydrolysis of isatin (1H-indole-2,3-dione) into isatinate (2-(2-aminophenyl)-2-oxoacetate) (see Figure 7.1). The study of IHs has been encouraged by isatin's large spread in nature. The high resolution X-ray crystal structures of the two putative orthologs of IH (IH-a and IH-b, shown in Figure 7.2) in *Labrenzia aggregata* were recently obtained [34]. These enzymes consist of two monomers connected through β-hairpin domains. Each monomer has a metal center in which the catalytic reaction takes place. The binding site of IH-b with Mn^{2+} is shown in Figure 7.2. In IH-b, Mn^{2+} adopts an octahedral conformation with three waters, Asp79, His77, and His73, as ligands. In contrast, in IH-a the water opposite to His73 is replaced by Gln228.
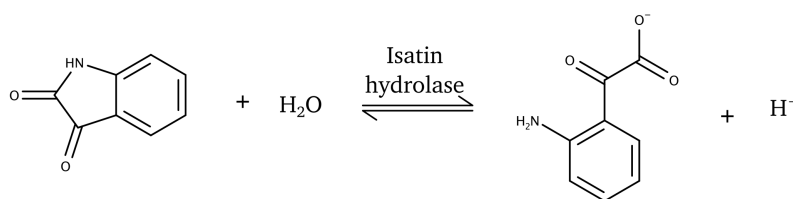
Fig. 7.1.: Isatin hydrolase catalyzes the hydrolysis of isatin to isatinate.

Recent experimental studies [32] showed that the catalytic activity of IHs highly depends on the metal bound to the binding site. The relative reaction rate of IH-b with a series of metals is summarized in Figure 7.3. This figure shows that the rate is the largest with Mn^{2+} and significantly lower with any other metal ion. A similar behavior for IH-a has been observed for most metals, except for the Cu^{2+}-loaded IH-a which shows almost no activity.
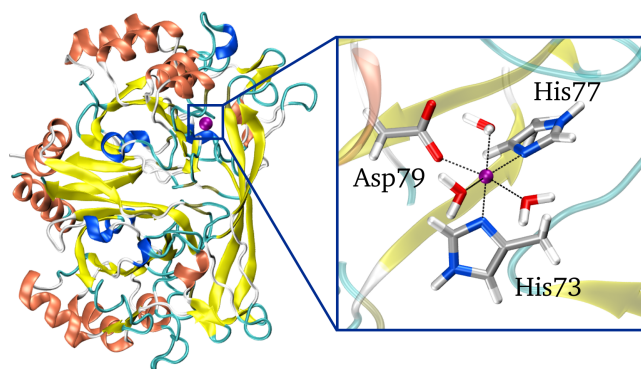
Fig. 7.2.: IH-b X-ray structure [34]. Inset: binding site with Mn^{2+}.
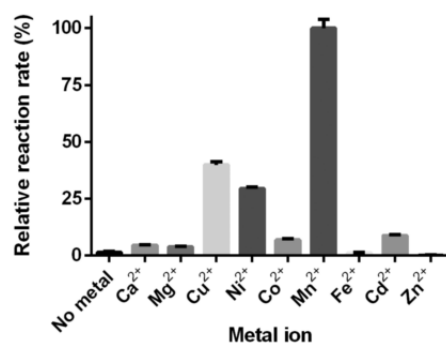
Fig. 7.3.: Metal dependency of IH-b relative reaction rate. Taken from Sommer *et al.* [32].

IH-a's and IH-b's high metal specificity is an unusual behavior for metalloenzymes in general. In particular, the complete inhibition of the catalytic power yield by the Zn^{2+}-loaded isoforms of both IH-a and IH-b is unexpected, because Zn^{2+} is one of the most common bio-active hydrolitic ions. A statistical study of all PDB entries by 2008 [30] showed that in more than 50% of the characterized metallohydrolases the natural cation is Zn^{2+}. In addition, this study showed that for most metalloenzymes that have Mn^{2+} as natural ion, the cation can be replaced by Zn^{2+}, Mg^{2+}, or Ca^{2+} without modifying the enzymes catalytic activity. Therefore, identifying the causes for the high metal specificity of IHs and the features that give rise to these enzymes' special behavior is of great interest.

To get a deep understanding of the catalytic process of IHs a methodology that provides an atomistic resolution of the binding site is required. Because this kind of resolution is not easily accomplished using experimental techniques, a computational investigation is the best approach. Here, QM calculations and QM/MM simulations are performed to reveal the causes of the high metal specificity of IHs and the mechanism of isatin's hydrolysis in this class of enzymes.

This study is divided in three parts. In the first part (Section 7.1), the results of QM calculations for a model system of IH-b binding site with Mn^{2+} are presented. In the second part (Section 7.2), the binding site conformation adopted with different metal ions is investigated. In the last part (Sections 7.3 and 7.4), the mechanism of the catalyzed hydrolysis

of isatin is studied using QM/MM metadynamics simulations. A summary of the findings is presented in Section 7.5.

7.1 Finding the lowest energy spin state of Mn^{2+}

The electronic configuration of Mn^{2+} is $[Ar]3d^5$. Hence, Mn^{2+} has three possible spin states: doublet, quartet, and sextet. The lowest energy spin state for tetrahedral and octahedral organometallic compounds of Mn^{2+} are usually the quartet and sextet state, respectively. As mentioned before, the experimentally observed conformation for the binding site of IH-b is octahedral. Thus, it is expected that the sextet is the lowest energy spin state when Mn^{+2} is bound to IH-b. However, given that the experimental conformation is not perfectly octahedral, it is also plausible that the quartet is the lowest state. Therefore, QM calculations were performed to confirm which is the lowest energy spin state of the Mn^{2+} when bound to the enzyme and to the substrate (isatin).

For this purpose, single-point energy QM calculations of a model system (shown in Figure 7.4) were carried out. These calculations were performed at the same level of theory as the QM/MM simulations (BLYP/DFT), at Hartree-Fock (HF) level, and at the LPNO-CCSD (local pair natural orbital coupled cluster with single and double excitations [90]) level of theory. The LPNO-CCSD calculations were performed to validate the results obtained with DFT, because DFT is known to often incorrectly describe the spin state of transition metals in organometallic complexes. The results of these calculations are summarized in Table 7.1.
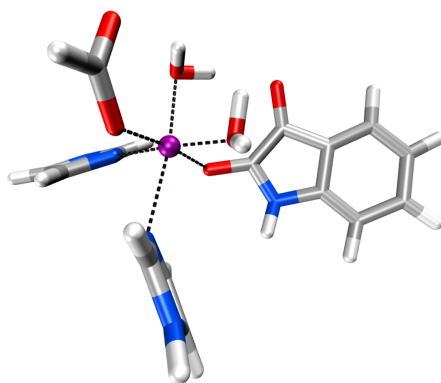
Fig. 7.4.: Model system for the QM calculations consisting of Mn^{2+} coordinated by isatin, two waters, a deprotonated formic acid molecule, and two imidazol molecules.

At all levels of theory and for both the protonated and deprotonated model system, the spin state with lowest energy is the sextet. However, the energy differences are smaller for DFT, indicating that DFT overstabilizes the quartet state with respect to the sextet state. DFT's underestimation of ΔE is probably caused by the large spin contamination of the quartet state, that is reflected in the large deviation of $\langle S^2 \rangle$ ($\langle S^2 \rangle_{protonated} = 4.44$ and $\langle S^2 \rangle_{deprotonated} = 4.63$) from the ideal value of 3.75. In contrast, the deviation of $\langle S^2 \rangle$ for the sextet ($\langle S^2 \rangle_{protonated} = 8.78$ and $\langle S^2 \rangle_{deprotonated} = 8.79$) is small in comparison with the

Tab. 7.1.: Energy difference (kcal·mol^{-1}) between quartet and sextet state for the model system as shown in Figure 7.4 (protonated) and with one of the water molecules replaced by a OH$^-$ group (deprotonated), calculated as $\Delta E = E_{\text{quartet}} - E_{\text{sextet}}$. All calculations were performed with the cc-pVDZ [91, 92] basis set.

		Level of theory	
System	HF	LPNO-CCSD	BLYP/DFT
Protonated	54.9	41.4	5.3
Deprotonated	1.3	10.9	3.0

ideal value of 8.75. Thus, DFT provides a good description of the electronic structure of the lowest energy spin state (*i.e.*, sextet state).

The results of the QM calculations confirm that the sextet is the lowest energy spin state of Mn^{2+} when bound to the enzyme and validate the level of theory chosen for the QM/MM simulations (BLYP/DFT).

7.2 QM/MM simulations for a series of metals

In order to understand the high metal specificity of IH, the conformation adopted by the binding site of IH-b with four different metal ions: Mn^{2+}, Cu^{2+}, Mg^{2+}, and Zn^{2+} was investigated using QM/MM simulations. These ions were chosen because they cover a large range of reaction rates (as shown in Figure 7.3) and have different electronic configuration and spin state. The QM region, for the QM/MM simulations, was chosen such that the metal ion and its ligands were included, as shown in Figure 7.5. The amino-acid ligands were cut by the QM-MM boundary at the C$_\alpha$-C$_\beta$ bonds and the link atoms scheme (see Section 2.9) was used to fill in the free valency.

For each metal, the conformation of the metal binding site was studied by placing the metal ion in the X-ray structure of the IH-b ortholog. QM/MM simulations were run until the binding site conformation was equilibrated, for a total simulation time of 5 - 20 ps. The final structures of the binding site for all metal ions are shown in Figure 7.5.
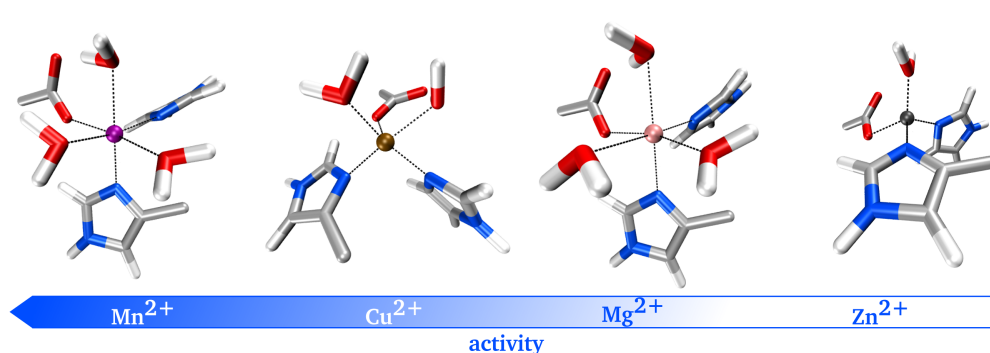
Fig. 7.5.: Final QM/MM binding site conformation with different metal ions. All ligands are neutral except the aspartic acid that has a charge of -1.

The final QM/MM conformation for Mn^{2+} and Mg^{2+} is the same as the one observed experimentally; both metal ions adopt an octahedral conformation bound to an aspartic acid, two histidines, and three waters. In contrast, the Cu^{2+} loses one of the waters observed in the experimental conformation and adopts a distorted trigonal bipyramidal conformation. Finally, the Zn^{2+} loses two waters and adopts a tetrahedral conformation.

The binding site conformations for IH-a can be safely assumed to be the same as found for IH-b, because the binding sites of both IH orthologs are very similar. The main modification in the binding site conformations when going from IH-b to IH-a is the replacement of one of the water ligands by a Gln. Therefore, in IH-a Mn^{2+} and Mg^{2+} are expected to have two water ligands, Cu^{2+} one water, and Zn^{2+} no water ligand.

The impact of the binding site conformation on the reaction rate is explained by considering the mechanism of isatin's hydrolysis. In water, the hydrolysis pathway has been found to depend on the pH [93], as shown in Figure 7.6. All three pathways start with the nucleophilic attack of a OH^- group to isatin's carbonyl, leading to the formation of an anionic tetrahedral intermediate. Afterwards, the pathways diverge according to the pH.
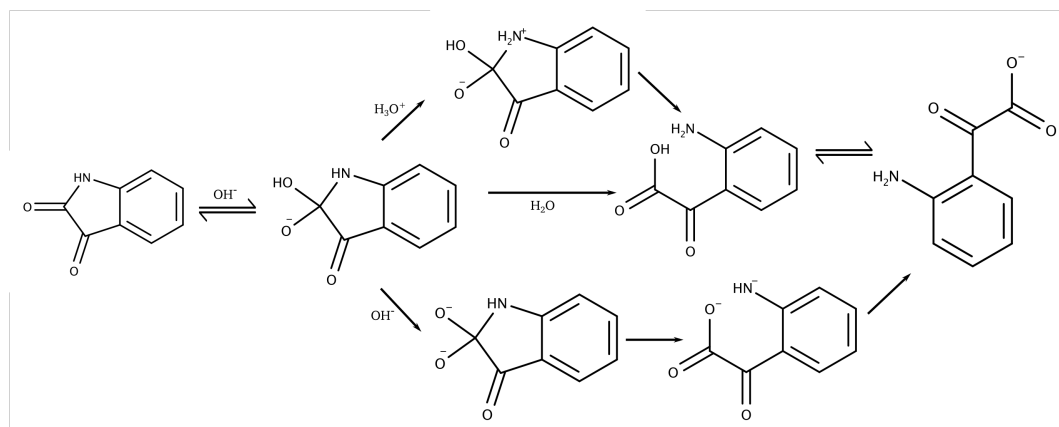
Fig. 7.6.: pH dependent hydrolysis mechanism of isatin in water. H_3O^+ indicates acid pH conditions, H_2O a neutral pH, and OH^- a basic pH.

In the case of the enzymatic hydrolysis, the metal center is expected to facilitate the deprotonation of one of the water ligands to create the OH^- group needed for the nucleophilic attack. Therefore, for the catalytic reaction to take place, at least the isatin and one water molecule (here called catalytic water) need to be bound to the metal ion at the same time. The tetrahedral conformation adopted by Zn^{2+} only allows for one and zero non-protein ligands, in IH-b and IH-a, respectively. Thus, a water molecule and isatin cannot be bound at the same time to Zn^{2+}, neither in IH-a nor IH-b. As a consequence, the Zn^{2+}-loaded IHs exhibit no activity.

The trigonal bipyramidal conformation adopted by Cu^{2+} accepts two and one non-protein ligands, in IH-b and IH-a, respectively. Therefore, a water molecule and isatin cannot be bound at the same time to Cu^{2+} in IH-a, but they can be bound to Cu^{2+} in IH-b. As a consequence, the Cu^{2+}-loaded IH-a shows no activity and Cu^{2+}-loaded IH-b exhibits some activity. However, the reaction rate of IH-b with Cu^{2+} is less than half of the one with Mn^{2+},

because the distortion in the binding site conformation (from an octahedral to a trigonal bipyramidal conformation) modifies the orientation of isatin and the catalytic water with respect to the position of the non-ligand residues that participate in the catalytic reaction.

Surprisingly, Mg^{2+} adopts the same conformation as Mn^{2+}. Hence, the low activity of IHs with Mg^{2+} cannot be explained based on the binding site conformation. A plausible explanation for Mg^{2+} low activity is that this ion fails to facilitate one or more steps of the reaction. To prove this hypothesis, the reaction mechanism and the free energy surface of the hydrolysis with Mg^{2+} and Mn^{2+} are investigated in the following section.

7.3 Activity - Mg^{2+} and Mn^{2+}

The activity of the IH-b with Mn^{2+} and Mg^{2+} was further investigated using QM/MM metadynamics simulations (see Section 2.10). Here, one of the three water ligands was substituted by isatin (see Figure 7.7) in the conformation suggested by the X-ray structure of a product analog bound to IH-b [34]. In comparison to the QM/MM calculations presented in the previous section, the QM region (shown in Figure 7.7) was extended to include other protein residues (Asp79, His212, and His83) that were assumed to participate in the catalytic reaction. The position of His212 relative to the catalytic water indicates that His212 may play the role of a base in the extraction of the proton from the catalytic water. Furthermore, the proximity of Asp198 to His212 suggests that the H-bond between these two residues could stabilize the protonated product (His212-H+). Finally, the position of His83 indicates that these residue might facilitate the deprotonation of isatinate, after the nucleophilic attack.
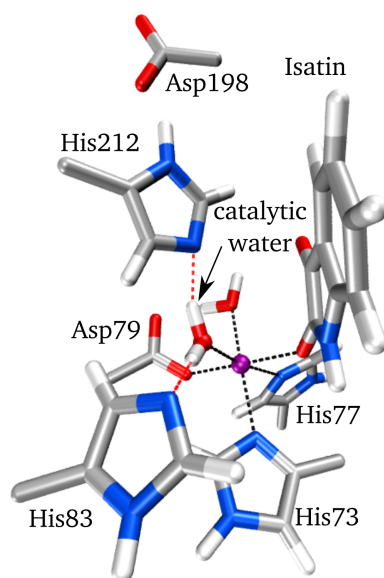
Fig. 7.7.: QM region for the metadynamics simulations. All ligands are neutral except the aspartic acid that has a charge of -1.

For the metadynamics simulations two collective variables (CV) were used, consisting of a combination of coordination numbers (CN) of covalent bonds being formed or broken during isatin's hydrolysis. A CN is calculated as:

$$CN_{ij} = \frac{1 - (d_{ij}/d_0)^p}{1 - (d_{ij}/d_0)^{p+q}}, \tag{7.1}$$

where d_{ij} is the distance between the atoms i and j, d_0 is a threshold distance that determines when a bond is broken, and p and q are constants that determine the steepness of the decay of CN_{ij} with respect to d_{ij}. The CN ranges from 0 (not bonded) to 1 (bonded).

The two CVs, shown in Figure 7.8.A, were defined as:

$$CV1 = CN_{N_{His212}\ldots H} - CN_{O\ldots H}$$
$$CV2 = CN_{O\ldots N_{isatine}}. \tag{7.2}$$

The metadynamics parameters used for CV1 are $p = 12$, $q = 14$, $d_0 = 1.10$ Å and for CV2 are $p = 12$, $q = 14$, $d_0 = 1.60$ Å. The values for p and q correspond to those typically used in metadynamics simulations [94].

The first variable (CV1) quantifies the degree of proton transfer between the catalytic water and His212. It takes negative values when the proton is bound to the water, zero when the H-bond between the water and His212 is formed, and positive values when the proton has been transferred to His212. The second variable (CV2) measures the degree of formation of the covalent bond between the nucleophile residue (OH^-) and the substrate (isatin). It takes values close to zero when the O-N bond has not yet been formed and values close to 1 when the O-N bond is formed.

The free energy surfaces (FES) obtained for Mg^{2+} and Mn^{2+}, from the metadynamics simulations, are shown in Figure 7.8.C. There, R, IS, and TI correspond to the reactant, an intermediate state, and the tetrahedral intermediate state, as shown in Figure 7.8.B. For Mg^{2+}, the transition from R to IS is endothermic and has an energy barrier of $13 \, \text{kcal} \cdot \text{mol}^{-1}$, whereas for Mn^{2+} the transition from R to IS is exothermic with an energy barrier of $7 \, \text{kcal} \cdot \text{mol}^{-1}$

The deprotonation of the catalytic water is observed in the FES of the two metals. However, for Mg^{2+} the deprotonation is endothermic and has a high energy barrier. In addition, the IS conformation is a saddle point and not a minimum in the Mg^{2+} FES, *i.e.*, the IS conformation is not stable in the Mg^{2+}-loaded IH. Furthermore, the transition from IS to TI, *i.e.*, the nucleophilic attack, is not observed in the Mg^{2+} FES. These results show that the low activity of the isatin hydrolase with Mg^{2+} is a consequence of the inability of this cation to activate the catalytic water necessary for the nucleophilic attack to take place.

The different behavior of Mn^{2+} and Mg^{2+} can be understood by considering the ions' polarizability and hydrolysis constants. Mg^{2+} is a hard ion with a pk_a of 11.4 [95]. In contrast, Mn^{2+} is a medium hard ion with a pk_a of 10.6 [95]. The lower pk_a of Mn^{2+} implies that the hydrolysis of a water molecule bound to this ion is favored against the one of a water molecule bound to Mg^{2+}. This observation is further supported by the ΔH_{298} value calculated by Trachtman *et al.* [96] for the proton transfer reaction $M[OH_2]^{n+} + H_2O \rightarrow M[OH]^{(n-1)+} + H_3O^+$, at the CCSD(T) level of theory with the 6-311++G** basis set.
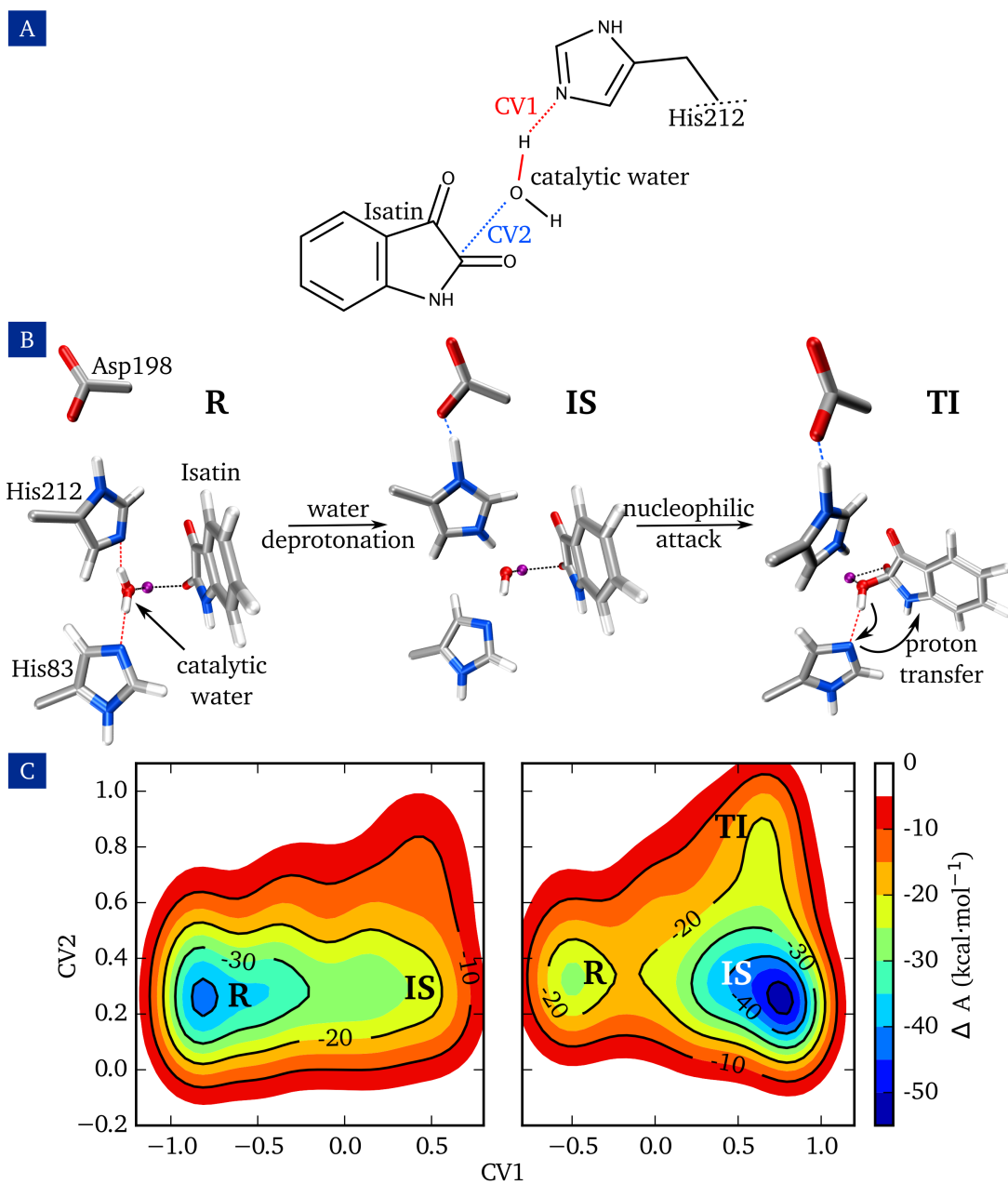
Fig. 7.8.: A. Collective variables as defined in Equation 7.2. **B.** Schematic representation of the conformations sampled during the metadynamics simulations with Mn^{2+}. **C.** Left: FES for Mg^{2+} after 2.8 ps and 112 deployed Gaussian hills. Right: FES for Mn^{2+} after 2.1 ps ps and 83 deployed Gaussian hills. R: reactant. IS: intermediate state. TI: tetrahedral intermediate.

For Mg^{2+} and Mn^{2+} the ΔH_{298} values are $-70.9\,\text{kcal}\cdot\text{mol}^{-1}$ and $-83.4\,\text{kcal}\cdot\text{mol}^{-1}$, respectively. The more negative value for Mn^{2+} shows that from an energetical point of view the deprotonation is more favorable for Mn^{2+}.

The fast activation of the catalytic water with Mn^{2+} was further confirmed by performing a QM/MM simulation without metadynamics, but with the same QM region (as shown Figure 7.7). In this simulation the deprotonation of the catalytic water was observed after only 5.6 ps, showing that the first step of the catalytic reaction with Mn^{2+} occurs spontaneously within an accessible simulation time. Therefore, the real Michaelis complex (*i.e.*, the enzyme-substrate complex) with Mn^{2+} corresponds to the state in which the water has been already deprotonated.

The nucleophilic attack is observed in the Mn^{2+} FES. However, the TI is not a minimum in the FES and the reaction does not proceed until the production of isatinate. This result indicates that there is at least one further step of the reaction mechanism which has not yet been considered in the metadynamics simulations. Thus, the reaction mechanism with Mn^{2+} is further investigated in the following section.

7.4 Reaction mechanism with Mn^{2+}

The missing step in the reaction mechanism with Mn^{2+} is revealed by taking into account the experimental observation that the activity of IH-b is the largest at pH 7 [32]. In the hydrolysis mechanism in water at pH 7 (see Figure 7.6), the step following the nucleophilic attack involves the protonation of isatin's N atom. Therefore, it is plausible that this is the step that follows the nucleophilic attack in the catalyzed hydrolysis.

A possible way to protonate isatin consist in a proton transfer from the OH^- group to isatin, facilitated by His83 (see Figure 7.8.B). A second possibility arises by considering the conformation in which the N-H bond in isatin points in the direction of His212 and not in the direction of His83 (see Figure 7.9.A). In this case, a hydrogen transfer could happen in a single step from His212 to isatin.

The two possibilities for the protonation of isatin were explored using metadynamics. In the first case (Figure 7.8.B), protonation was not observed. In the second case (Figure 7.9.A), a H-bond between His212 and isatin was observed, but the proton transfer did not take place. Therefore, the right binding mode of isatin is the one in which the N-H group of isatin points towards His212. Furthermore, the failure to observe the proton transfer suggests that this can only happen when the hybridization of the N atom of isatin has changed from sp^2 to sp^3 This change takes places when the C-N bond in isatin breaks.

Taking the previous observations into account, a second set of QM/MM metadynamics simulations was carried out, starting from the Michaelis complex. Three CVs were used as shown in Figure 7.10.A and defined as:

$$\begin{aligned}
CV1 &= CN_{C_{\text{isatin}}\cdots O} - CN_{C_{\text{isatin}}\cdots N_{\text{isatin}}} \\
CV2 &= CN_{N_{\text{His83}}\cdots H} - CN_{O\cdots H} \\
CV3 &= CN_{N_{\text{isatin}}\cdots H} - CN_{N_{\text{His212}}\cdots H}.
\end{aligned} \tag{7.3}$$
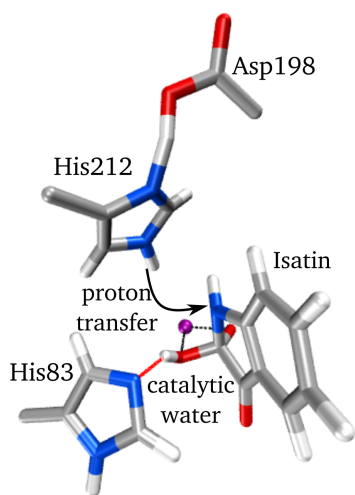
Fig. 7.9.: Alternative binding mode of the isatin to the binding site with Mn^{2+}. The colored dashed lines represent H-bonds.

For CV1, $p = 12$, $q = 14$, $d_0 = 1.6$ Å and for CV2 and CV3, $p = 12$, $q = 14$, $d_0 = 1.10$ 1.1 Å. CV1 measures both the degree of formation of the covalent bond between the substrate (isatin) and the nucleophile residue (OH^-) and the degree of elongation of isatin's C-N bond. CV1 takes negatives values when isatin's C-N bond is not yet broken and the C-O bond has not yet been formed, zero when the C-O bond is formed but the C-N bond is not yet broken, and positive values when the C-O bond is formed and the C-N is broken. CV2 quantifies the degree of proton transfer from the OH^- to His83. CV2 takes negative values when the proton is still bound to the OH^-, zero when the H-bond is formed, and positive values when the proton has been transfered to the His83. Finally, CV3 measures the degree of proton transfer from His212 to isatin. CV3 takes negative values when the proton is bound to His212, zero when the H-bond is formed, and positive values when the proton has been transfered to isatin.

The metadynamics simulation was first run for 1.05 ps, in which 41 Gaussian hills were deployed. During this simulation time, the tetrahedral intermediate (TI) shown in Figure 7.10.C was observed. However, the reaction did not proceed and the system returned to the reactant state (R, Michaelis complex). This behavior indicated that TI is only metastable and can easily recombine into R. Hence, a infinite potential wall was placed at CV1 for values less than zero to prevent the intermediate to return to R. The simulation was continued for 0.8 ps, in which 32 Gaussian hills were deployed. In this second run, the reaction was successfully sampled. The obtained FES and the observed reaction mechanism are depicted in Figure 7.10.B and C. In the FES, CV3 has been integrated out to decrease the dimensionality of the FES, this can be done without loosing information because the behavior of CV2 and CV3 was found to be very similar.

The first step of the reaction has an energy barrier of $4.0 \, kcal \cdot mol^{-1}$ and consists in the nucleophilic attack of the OH^- of isatin's carbonyl. The formed TI is, as expected, a saddle point in the FES, *i.e.*, TI is only metastable. The low stability of TI explains why only by placing a potential wall in CV1 it was possible to sample the reaction. In the second step, isatin's C-N bond opens and the transition state (TS) is reached. The energy difference bet-
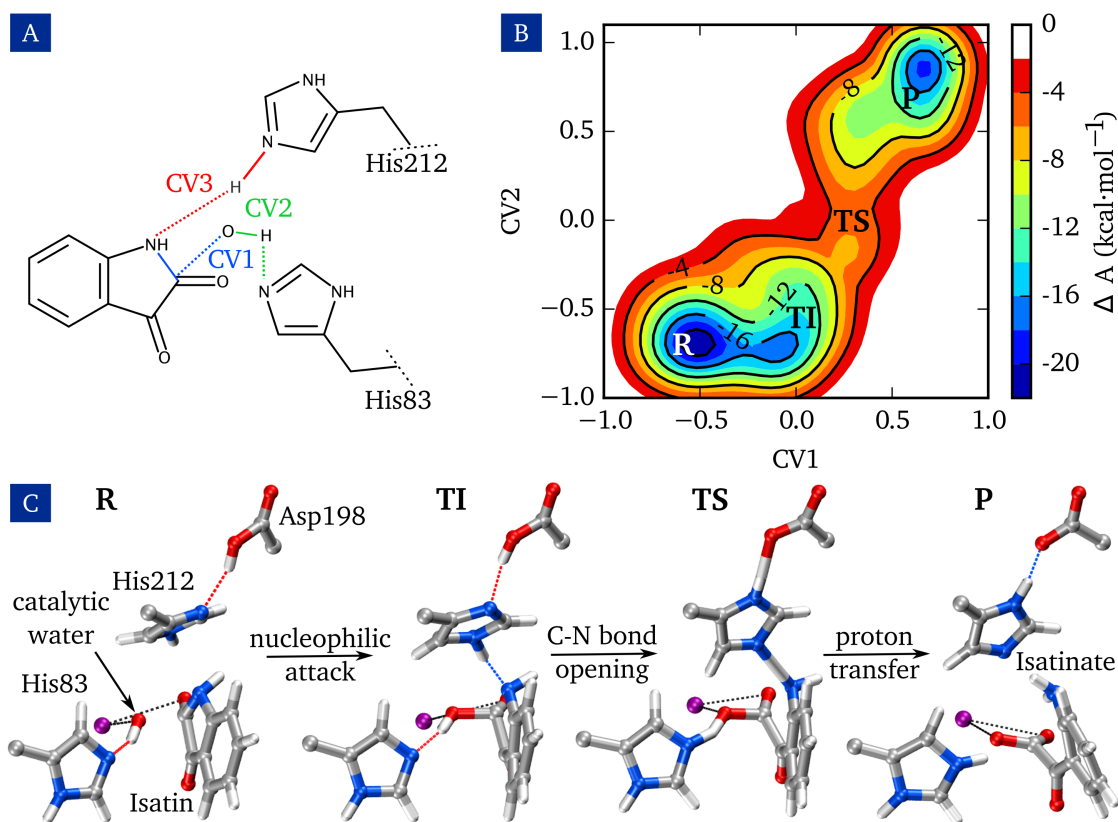
Fig. 7.10.: **A.** Collectives variables as defined in Equation 7.3. **B.** FES with Mn^{2+} after 2.7 ps and 109 deployed Gaussian hills; CV3 has been integrated out to decrease the dimensionality of the FES. **C.** Schematic representation of the hydrolysis mechanism catalyzed by Mn^{2+}; the colored dashed lines represent H-bonds. R: reactant. TI: tetrahedral intermediate. TS: transition state. P: product.

ween TI and TS is $12\,kcal\cdot mol^{-1}$. The TS is stabilized by the presence of two H-bonds, the first between isatin's N atom and the N-H of His212, and the second between the OH^- and the N-H of His83. The final step consists in the synchronous proton exchange between the two previously mentioned H-bonds. The protons transfer occurs without an energy barrier, because it only takes place once the TS has been reached. This prediction is supported by the experimental observation of the absence of a kinetic isotope effect on the reaction rate [99].

Further insights into the hydrolysis mechanism are obtained from the time evolution of the three CVs, shown in Figure 7.11. Between 0.1 ps and 0.5 ps CV1 takes values around zero, *i.e.*, the systems is in the TI state. However, the reaction does not proceed and the system returns to R. Only after placing the wall in CV1 at 1.05 ps, the transition from TI to TS is observed, *i.e.*, CV1 takes positive values. Almost immediately the H-bonds N_{His83}-H and N_{isatin}-H are formed (*i.e.*, CV2 and CV3 take values close to zero). The further elongation of isatin's C-N bond generates a fast increase in CV1 and induces the protons transfer (*i.e.*, CV2 and CV3 become positive). Hence, the time evolution of the CVs shows that the transition from TI to P follows a dissociative pathway in which the proton exchange only occurs after isatin's C-N bond is opened.
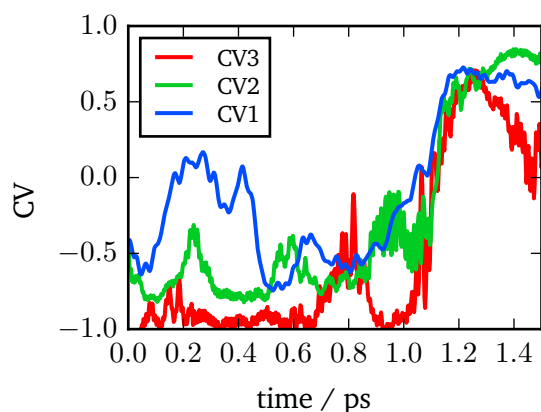
Fig. 7.11.: Time evolutions of the CVs in the second metadynamics simulation.

The poor Lewis base character of isatin's N-atom triggers the preference for such a dissociative pathway. In general, the pk_a value of aryl-amines is a few pH units larger than for alkyl-amines, because of the conjugation between the lone pair of the N-atom and the aromatic ring [97]. As a consequence, the formation of the H-bond and successive proton transfer from His212 to isatin can only occur after the electron density at the N-atom is increased due to the elongation of isatin's C-N bond. Moreover, the stabilization of the product requires the formation of the chelating carboxylate group through the deprotonation of the OH^- by His83.

The isomerization of isatinate

Isatinate has different conformers and the one obtained after the catalytic hydrolysis is not the most stable one. Therefore, a third set of metadynamics simulations was performed to study isatinate's isomerization. Two CVs, consisting on the two dihedral angles shown in Figure 7.12.A, were used. The FES for the isomerization process and the most stable conformers are shown in Figure 7.12.B and C.

After the hydrolysis reaction, isatinate adopts the initial conformation (I) shown in Figure 7.12.C (left). The FES shows that during its isomerization isatinate crosses several intermediate states and energy barriers, until it adopts the final conformation (F), shown in Figure 7.12.C (right). The conformation F is stabilized by an intra-molecular H-bond formed between the amino and one of the carbonyl groups. The isomerization of isatinate occurs through the simultaneous rotation of the two bonds described by CV1 and CV2, with a bicycle motion that resemble the one described for the isomerization of retinal in rhodopsin [98].

The final conformer found in the simulation is in good agreement with the binding geometry of the product observed in the X-ray conformation; a superposition of the final metadynamics conformation and the X-ray conformation is shown in Figure 7.13. This agreement proves that isatinate adopts the conformation F before it leaves the binding pocket.
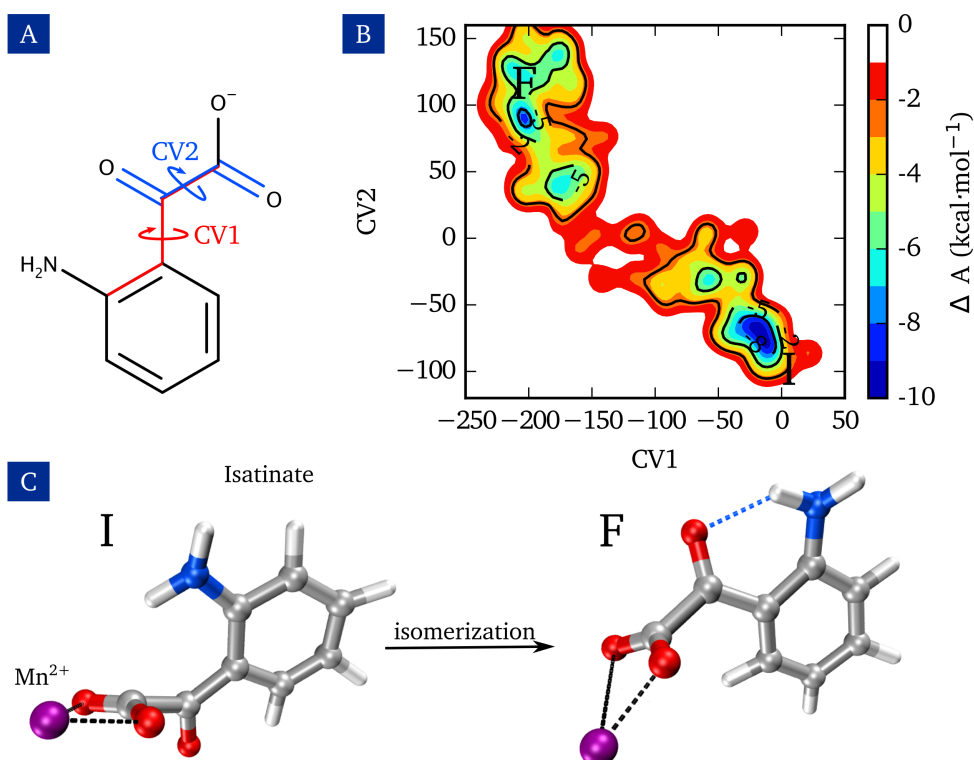
Fig. 7.12.: **A.** Collective variables used in the third set of metadynamics simulations. **B.** FES of the isomerization. **C.** Isatinate's most stable conformers. The colored dashed line represent a H-bond.
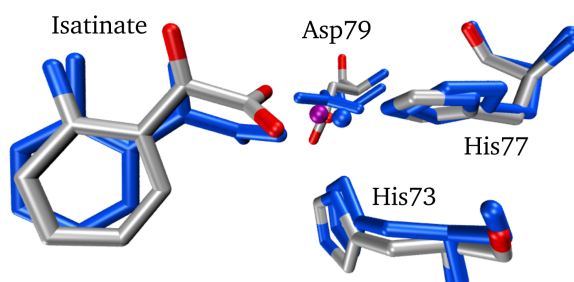
Fig. 7.13.: Superposition of the final metadynamics conformation (licorice colors) and the X-ray conformation (licorice blue). H atoms are omitted for simplicity.

7.5 Conclusions

In this chapter, the unusual high specificity of IH for Mn^{+2} has been investigated using QM calculations and QM/MM simulations. Furthermore, the mechanism of isatin's hydrolysis in IH is revealed using QM/MM metadynamics simulations.

QM calculations of a model system for the binding site with Mn^{+2} show that the sextet is the lowest energy spin state of this ion when bound to IH. Using this information, the conformation adopted by the binding site with four different metal ions (Mn^{2+}, Mg^{2+}, Zn^{2+}, and Cu^{2+}) was obtained from QM/MM simulations. For Zn^{+2} and Cu^{+2}, the higher or lower catalytic activity is fully explained by taking into account the binding site conformation.

However, the low activity of Mg^{+2} is found to be unrelated to its conformation, because it adopts an octahedral conformation equivalent to the one adopted by the ion with the highest catalytic activity, *i.e.*, Mn^{2+}.

The low catalytic activity of Mg^{+2} is further investigated by calculating the free energy surface of the catalytic hydrolysis and determining the reaction mechanism using QM/MM metadynamics simulations. The low activity of the Mg^{2+}-loaded IH is explained as a consequence of the failure of Mg^{+2} to activate the catalytic water, because Mg^{2+} is a hard acid with a large pk_a.

QM/MM metadynamics simulations were also performed for Mn^{+2}, allowing to deduce the following mechanism for the hydrolysis of isatin:

1. **Activation of the catalytic water.** This is an exothermic reaction which occurs spontaneously in a regular QM/MM simulation. Hence, the Michaelis complex of IH corresponds to a state in which the water has already been deprotonated.

2. **Nucleophilic attack.** The intermediate form after this step is only metastable.

3. **Isatin's C-N bond opening.** This is the rate-limiting step of the reaction.

4. **Isatin protonation and OH^- deprotonation.** This only occurs when the transition state has been reached.

5. **Isatinate isomerization.**

The observed mechanism for IHs differs from those for similar metallohydrolases in which the protonation of the amino group occurs before the C-N bond cleavage [35]. In IHs, the protonation can only take place after the bond cleavage, because of the poor Lewis base character of isatin's N-atom. Moreover, the high aromaticity of isatin increases the energy barrier for the opening of the C-N bond, since the breaking of this bond disrupts the conjugation between the N-atom and the aromatic ring.

The unusual specificity for manganese of IHs is more than a purely conformational effect, because it is intrinsically correlated to the particular mechanism of isatin's hydrolysis in IHs. Therefore, IHs constitutes a novel class of enzymes with a high metal selectivity, determined by both conformational and energetical effects.

Summary and outlook $\qquad$ 8

Computer simulations have become standard tools for the investigation of chemical systems and processes. In the recent years, the increase of computer power and the development of new methodologies has allowed to apply computer simulations to larger and more complex systems as well as to processes that occur on larger time scales. In this work, state of the art computer simulations are applied to two current topics in chemistry: oligomer folding and enzymatic catalysis. It is illustrated how computer-based studies can be used to explain experimental results, to get a deeper understanding of a chemical process, and to make predictions. Additionally, this work shows that there is a large spectrum of computational methods that allow to study systems with very different properties.

8.1 Mechanical properties of foldamers

The design of foldamers with predictable properties and folded conformations requires a deep understanding of the folding process and the identification of the internal and external factors influencing this process. Molecular dynamics (MD) studies of foldamers provide detailed atomistic information of the unfolding process. In particular, force-probe MD (FPMD) simulations enable the study of the mechanical unfolding pathway of oligomers and the elucidation of the main features of the energy landscape related to this process.

Through this work, the mechanical unfolding pathway of a series of foldamers and of an α-peptide adopting helix conformations have been investigated using MD and FPMD simulations. In a first instance, MD simulations performed with different force fields show how critical the choice of the right force field is to study a given system. Furthermore, these simulations depict the poor transferability of commonly used force fields such as AMBER and OPLS.

The statistical analysis of hundreds of FPMD simulations demonstrates that all the studied oligomers unfold in a cooperative manner following a nontrivial unfolding pathway. Although external factors such as temperature, solvent proticity, and pulling velocity are shown to influence the preference for an unfolding pathway, the preferred unfolding pathway of an oligomer is found to be mainly determined by intrinsic features of the oligomer's backbone architecture such as the presence of intra-molecular H-bonds and the backbone rigidity. A comparison of the behavior between the different systems allows to reveal the main structural determining factors of the unfolding energy landscape. The larger differences between unfolding pathways are found between helices with and without intramolecular H-bonds. Within the H-bonded helices, the interplay between H-bond stability and backbone stiffness is revealed. All these findings led to the proposal of a series of rules for the prediction of the unfolding pathway of oligomers adopting helix conformations.

In this dissertation, FPMD simulations are shown to be a powerful tool for the investigation of the unfolding process of oligomers. FPMD is shown to provide not only structural information of the unfolding pathway, but also information about the energy landscape and the kinetics. The investigation of the mechanical unfolding of small oligomers exhibiting nontrivial pathways carried out in this work marks a starting point for the understanding of the unfolding/folding process of foldamers. Moreover, this thesis sets a standard for future studies of polymer folding.

8.2 Metal specificity of an isatin hydrolase

The application of quantum-/molecular mechanics simulations (QM/MM) for the study of a complex biological process is illustrated in the second part of this thesis. The target system is a class of enzymes called isatin hydrolases (IHs) which has been experimentally observed to exhibit high specificity for Mn^{2+}. QM/MM simulations of the binding site of IH with different metals ions demonstrated that IH's low catalytic activity with some metals is a consequence of the binding site conformation. However, for Mg^{2+} the low activity is determined not to be influenced by the binding site conformation, but to be related to Mg^{2+} poor Lewis-acid character.

The mechanism of the hydrolysis of isatin in IH with Mn^{2+} has been determined using metadynamics QM/MM simulations. The predicted mechanism is significantly different from those of similar metallohydrolases and is found to be also the cause of IH's high metal specificity. IH's particular mechanism and unsual metal specificity indicates that IH constitute a novel class of enzymes. The knowledge acquired about IH constitutes a step towards the understanding of the role of metal ions in enzymatic catalysis.

8.3 Future perspectives

This work illustrates the application of computational simulations for the study of systems and processes that were not computationally accessible until a couple of years ago. The fast growth not only of the computational power, but especially the progress made in available methodologies and algorithms is bringing computational simulations to be recognized as equally valuable as experimental techniques. With the accessible systems and processes becoming more complex, the role of computational simulations in current highlight research topics, *e.g.*, protein folding, will continue to gain relevancy.

Bibliography

[1] Guichard, G.; Huc, I. *Chem. Commun.* **2011**, *47*, 5933.

[2] Kritzer, J. A.; Tirado-Rives, J.; Hart, S. A.; Lear, J. D.; Jorgensen, W. L.; Schepartz, A. *J. Am. Chem. Soc.* **2005**, *127*, 167.

[3] Fremaux, J.; Mauran, L.; Pulka-Ziach, K.; Kauffmann, B.; Odaert, B.; Guichard, G. *Angew. Chem. Int. Ed.* **2015**, *54*, 9816.

[4] Li, Y.; Dutta, T.; Gerasimchuk, N.; Wu, S.; Shetye, K.; Jin, L.; Wang, R.; Peng, Z. *ACS Appl. Mater. Interfaces* **2015**, *7*, 9372.

[5] Renfrew, P. D.; Craven, T. W.; Butterfoss, G. L.; Kirshenbaum, K.; Bonneau, R. *J. Am. Chem. Soc.* **2014**, *136*, 8772.

[6] Shin, Y.-H.; Mortenson, D. E.; Satyshur, K. A.; Forest, K. T.; Gellman, S. H. *J. Am. Chem. Soc.* **2013**, *135*, 8149.

[7] Sebaoun, L.; Maurizot, V.; Granier, T.; Kauffmann, B.; Huc, I. *J. Am. Chem. Soc.* **2014**, *136*, 2168.

[8] Hecht, S.; Huc, I., *Foldamers: Structure, Properties, and Applications*; Wiley-VCH: 2007, pp 1–33.

[9] Ducasse, L.; Castet, F.; Fritsch, A.; Huc, I.; Buffeteau, T. *J. Phys. Chem. A* **2007**, *111*, 5092.

[10] Dawson, S. J.; Mészáros, Á.; Pethő, L.; Colombo, C.; Csékei, M.; Kotschy, A.; Huc, I. *Eur. J. Org. Chem.* **2014**, *2014*, 4265.

[11] Qi, T.; Maurizot, V.; Noguchi, H.; Charoenraks, T.; Kauffmann, B.; Takafuji, M.; Ihara, H.; Huc, I. *Chem. Comm.* **2012**, *48*, 6337.

[12] Onuchic, J. N.; Wolynes, P. G. *Curr. Opin. Struct. Biol.* **2004**, *14*, 70.

[13] Thirumalai, D.; O'Brien, E. P.; Morrison, G.; Hyeon, C. *Annu. Rev. Biophys.* **2010**, *39*, 159.

[14] Borgia, A.; Williams, P. M.; Clarke, J. *Annu. Rev. Biochem.* **2008**, *77*, 101.

[15] Galera-Prat, A.; Gómez-Sicilia, A.; Oberhauser, A. F.; Cieplak, M.; Carrión-Vázquez, M. *Curr. Opin. Struct. Biol.* **2010**, *20*, 63.

[16] Woodside, M. T.; Block, S. M. *Annu. Rev. Biophys.* **2014**, *43*, 19.

[17] Ritort, F. *J. Phys. Condens. Matter* **2006**, *18*, R531.

[18] Evans, E. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 105.

[19] Noy, A.; Friddle, R. W. *Methods* **2013**, *60*, 142.

[20] Kumar, S.; Li, M. S. *Phys. Rep.* **2010**, *486*, 1.

[21] Žoldák, G.; Rief, M. *Curr. Opin. Struct. Biol.* **2013**, *23*, 48.

[22] Sotomayor, M.; Schulten, K. *Science* **2007**, *316*, 1144.

[23] Freddolino, P. L.; Harrison, C. B.; Liu, Y.; Schulten, K. *Nat. Phys.* **2010**, *6*, 751.

[24] Zhou, J. F.; Wu, Q. J.; Ye, Y. Z.; Tong, J. G. *Genetica* **2003**, *119*, 93.

[25] Merkel, R. *Phys. Rep.* **2001**, *346*, 343.

[26] Keller, B.; Gattin, Z.; van Gunsteren, W. F. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1677.

[27] Daura, X.; van Gunsteren, W. F.; Rigo, D.; Jaun, B.; Seebach, D. *Chem. Eur. J.* **1997**, *3*, 1410.

[28] Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *J. Mol. Biol.* **1998**, *280*, 925.

[29] Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem. Int. Ed.* **1999**, *38*, 236.

[30] Andreini, C.; Bertini, I.; Callavaro, G.; Holliday, G. L.; Thornton, J. M. *J. Biol. Inorg. CHem.* **2008**, *13*, 1205.

[31] Steitz, T. A.; Steitz, J. A. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6498.

[32] Sommer, T.; Bjerregaard-Andersen, K.; Simensen, S.-M.; Jensen, J. K.; Jochimsen, B.; Riss, P. J.; Etzerodt, M.; Morth, J. P. *ACS Chem. Neurosci.* **2015**, *6*, 1353.

[33] Yang, W.; Lee, J. Y.; Nowotny, M. *Mol. Cell* **2006**, *22*, 5.

[34] Bjerregaard-Andersen, K.; Sommer, T.; Jensen, J. K.; Jochimsen, B.; Etzerodt, M.; Morth, J. P. *J. Biol. Chem.* **2014**, *289*, 21351.

[35] Page, M. I.; Badarau, A. *Bioinorg. Chem. Appl.* **2008**, *2008*, 1.

[36] Frenkel, D.; Smit, B., *Understanding Molecular Simulation*; Academic Press: 2002.

[37] Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. *J. Phys. Condens. Matter* **2004**, *16*, R481.

[38] Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144.

[39] Praprotnik, M.; Site, L. D.; Kremer, K. *Annu. Rev. Phys. Chem.* **2008**, *59*, 545.

[40] Oostenbrink, C.; Soares, T.; van der Vegt, N. F.; van Gunsteren, W. F. *Eur. Biophys. J.* **2005**, *34*, 273.

[41] Hockney, R.; Eastwood., J., *Computer Simulations Using Particles.* New York, McGraw-Hill: 1981.

[42] Verlet, L. *Phys. Rev.* **1967**, *159*, 98.

[43] Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.

[44] Nosé, S.; Klein, M. *Mol. Phys.* **1983**, *50*, 1055.

[45] Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182.

[46] Bell, G. I. *Science* **1978**, *200*, 618.

[47] Kramers, H. *Physica* **1940**, *7*, 284.

[48] Dudko, O. K.; Hummer, G.; Szabo, A. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 15755.

[49] Dudko, O.; Hummer, G.; Szabo, A. *Phys. Rev. Lett.* **2006**, *96*, 108101.

[50] Zhang, Y.; Dudko, O. K. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16432.

[51] Souaille, M.; Roux, B. *Comput. Phys. Commun.* **2001**, *135*, 40.

[52] Best, R. B.; Hummer, G.; Eaton, W. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17874.

[53] Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.

[54] Thiel, W., *QM / MM Methodology : Fundamentals , Scope , and Limitations*; Grotendorst, J., Attig, N., Blügel, S., Marx, D., Eds.; Institute for Advanced Simulation, Forschungszentrum Jülich: 2009; Vol. 42, pp 201–214.

[55] Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.

[56] Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562.

[57] Seebach, D.; Ciceri, P. E.; Overhand, M.; Jaun, B.; Rigo, D. *HeIv. Chim. Acta* **1996**, *79*, 2043.

[58] Rathore, N.; Gellman, S. H.; de Pablo, J. J. *Biophys. J.* **2006**, *91*, 3425.

[59] Appella, D. H.; Christianson, L.; Klein, D. A.; Powell, D. R.; Huang, X.; Barchi, J. J.; Gellman, S. H. *Nature.* **1997**, *387*, 381.

[60] Appella, D. H.; Christianson, L.; Karle, I. L.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 6206.

[61] De Groot, B. L.; Daura, X.; Mark, E.; Grubmüller, H. *J. Mol. Biol.* **2001**, *309*, 299.

[62] Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Wei Zhang, R. Y.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999.

[63] Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400.

[64] Jorgensen, W. L.; Maxwell, D. S.; Tirado-rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

[65] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.

[66] Karger, N.; Vardag, T.; Lüdemann, H. *J. Chem. Phys.* **1990**, *93*, 3437.

[67] Beauchamp, K. A.; Lin, Y.-S.; Das, R.; Pande, V. S. *J. Chem. Theory Comput.* **2012**, *8*, 1409.

[68] Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H. *J. Am. Chem. Soc.* **2007**, *129*, 1179.

[69] Hazel, A.; Chipot, C.; Gumbart, J. C. *J. Chem. Theory Comput.* **2014**, *10*, 2836.

[70] Lin, Z.; Riniker, S.; Van Gunsteren, W. F. *J. Chem. Theory Comput.* **2013**, *9*, 1328.

[71] Park, S.; Schulten, K. *J. Chem. Phys.* **2004**, *120*, 5946.

[72] Ozer, G.; Quirk, S.; Hernandez, R. *J. Chem. Phys.* **2012**, *136*, 215104.

[73] Rohs, R.; Etchebest, C.; Lavery, R. *Biophys. J.* **1999**, *76*, 2760.

[74] Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 3559.

[75] Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219.

[76]Allison, J. R.; Müller, M.; van Gunsteren, W. F. *Protein Sci.* **2010**, *19*, 2186.

[77]Laursen, J. S.; Engel-Andreasen, J.; Olsen, C. A. *Acc. Chem. Res.* **2015**, *48*, 2696.

[78]Fisher, B. F.; Guo, L.; Dolinar, B. S.; Guzei, I. A.; Gellman, S. H. *J. Am. Chem. Soc.* **2015**, *137*, 6484.

[79]Jiang, H.; Léger, J.-M.; Huc, I. *J. Am. Chem. Soc.* **2003**, *125*, 3448.

[80]Balamurugan, D.; Muraleedharan, K. M. *Chem. Eur. J.* **2015**, *21*, 9332.

[81]Abramyan, A.; Liu, Z.; Pophristic, V. *Chem. Comm.* **2016**, *52*, 669.

[82]Glover, V.; Halket, J. M.; Watkins, P. J.; Clow, A.; Goodwin, B. L.; Sandier, M. *J. Neurochem.* **1988**, *51*, 656.

[83]Hamaue, N.; Yamazaki, N.; Terado, M.; Minami, M.; Ohno, K.; Ide, H.; Ogata, A.; Honma, S.; Tashiro, K. *Res. Commun. Mol. Pathol. Pharmacol.* **2000**, *108*, 63.

[84]Pakravan, P.; Kashanian, S.; Khodaei, M. M.; Harding, F. J. *Pharmecol. Rep.* **2013**, *65*, 313.

[85]Vine, K. L.; Matesic, L.; Locke, J. M.; Ranson, M.; Skropeta, D. *Anticancer Agents Med. Chem.* **2009**, *9*, 397.

[86]Pandeya, S. N.; Sriram, D.; Nath, G.; Declercq, E. *Eur. J. Pharm. Sci.* **1999**, *9*, 25.

[87]Patel, A.; Bari, S.; Talele, G.; Patel, J.; Sarangapani, M. *Iran. J. Pharm. Res.* **2006**, *4*, 249.

[88]Krishnan, S.; Saravanan, M.; Ramesh, A. *Eur. J. Med. Chem. 36* **2001**, *36*, 615.

[89]Pandeya, S. N.; Sriram, D.; Nath, G.; Clercq, E. D. *Il Farmaco* **1999**, *54*, 624.

[90]Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 034106.

[91]Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

[92]Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358.

[93]Casey, L. A.; Galtb, R.; Page, M. I. *J. Chem. Soc. Perkin Trans.* **1993**, *2*, 23.

[94]Biarnés, X.; Ardèvol, A.; Iglesias-Fernández, J.; Planas, A.; Rovira, C. *J. Am. Chem. Soc.* **2011**, *133*, 20301.

[95]Wulfsberg, G., *Principles of Descriptive Inorganic Chemistry*; University Science Books: 1991.

[96]Trachtman, M.; Markham, G. D.; Glusker, J. P.; George, P.; Bock, C. W. *Inorg. Chem.* **2001**, *4*, 5752.

[99]Morth, J. P. *Unpublished*.

[97]Perrin, D. D.; Dempsey, B.; Serjeant, E. P., *pK$_a$ Prediction for Organic Acids and Bases*; Springer Netherlands: 1981, pp 12–20.

[98]Schapiro, I.; Weingart, O.; Buss, V. *J. Am. Chem. Soc.* **2009**, *131*, 16.

[100]George D. Purvis, R. J. B. *J. Chem. Phys.* **1982**, *76*, 1910.

[101]VandeVondele, J.; Hutter, J. *J. Chem. Phys.* **2007**, *127*, 114105.

[102]Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2010**, *6*, 2348.

[103]Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703.

[104]Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.

[105]VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103.

[106]Neese, F. *WIREs. Comput. Mol. Sci.* **2012**, *2*, 73.

[107]Toukan, K.; Rahman, A. *Phys. Rev. B* **1985**, *31*, 2643.

[108]Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656.

[109]Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.

[110]Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

[111]Allen, M.; Tildesley, D., *Computer Simulations of Liquids*; Oxford, Oxford Science Publications: 1987.

[112]Hess, B.; Bekker, H.; Berendsen, H.; Fraajie, J. *J. Comput. Chem.* **1997**, *18*, 1463.

[113]Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950.

[114]Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. *J. Chem. Theory Comput.* **2011**, *7*, 4026.

Appendices

Computational details A

A.1 Molecular dynamic simulations

All MD and FPMD simulations were performed using the GROMACS 4.6.5 program package [109] and the GROMOS53A6FF [40, 108], except for some simulations that were carried out using other FFs as indicated in the text. When present, the water molecules were described using the SPC model [107]. The long-range Coulomb interactions were treated using the particle-mesh Ewald summation method [110], a cut-off of 1.4 nm was used for the short-range Coulomb interactions, and a dispersion correction was applied for the van der Waals interactions [111]. In all simulations, periodic boundary conditions were applied and all bonds were constrained to their equilibrium distances using the LINCS-algorithm [112]. The neighbor list was updated every 10 fs and a time-step of 2 fs was used.

All the simulations were carried out in the NPT ensemble and prepared in the following way:

1. An energy minimization was performed for all molecules.

2. Each oligomer was placed in a properly sized simulation box that was later filled with solvent molecules.

3. The energy of the system was minimized.

4. The system was equilibrated for 500 ps to the target temperature, using a velocity-rescaling thermostat [43] with a time constant of 0.1 ps.

5. The system was coupled to a barostat (Parrinello-Rahman barostat [45] with a time constant of 2 ps and a compressibility of $4.5 \times 10^{-5}\,\mathrm{bar}^{-1}$).

A.2 QM and QM/MM calculations

Single-point QM calculations were carried out with the Orca 2.0.3 program package [106], at the BLYP/DFT or CCSD [100] level using the cc-pVDZ basis set [91, 92] for all atoms. The QM/MM simulations were performed with the CP2K 2.5.1 package [105]. For the MM region, the Amberff99SB [113] force field was used. The QM region was studied at the BLYP/DFT level with the DZVP-MOLOPT-SR-GTH basis set [101], the GTH-BLYP core potential [103], and the cFIT3 auxiliary basis set [102]. The preparation of the simulations was performed in the following way:

1. The system was placed in a simulation box and solvated with water.

2. The system was equilibrated at MD level to $T = 300\,$K, keeping fixed the coordinates of the atoms of the active site.

3. The system was annealed from $T = 300\,$K to $T < 10\,$K, at QM/MM level.

4. The system was allowed to relax without any constraint within the NVE ensemble.

5. Step 3 and 4 were repeated until the temperature of the QM and the MM region reached the same stable value.

6. The temperature was slowly increased until $T = 300\,$K.

7. The QM/MM calculation was carried out at $T = 300\,$K within the NVT ensemble, using the Nosé–Hoover thermostat [104] and with a time-step of 0.25 fs.

A.2.1 QM/MM metadynamics simulations

All the metadynamics simulations were prepared and carried out as described above. The Gaussian hills, with a height of $2\,$kcal$\cdot$mol^{-1}, were spanned every 100 step (*i.e.*, every 250 fs) and a scaling factor of 0.15 was used for all the collective variables.

Force field parameters

<div style="text-align: right">**B**</div>

B.1 Force field parameters for the heptamer

The following section contains the force field atom types and charges used to describe the four different β-aminoacids present in the heptamer. Figure B.1 shows the labeling of the heavy atoms of the amino acids used for assigning the atom type and the charges. The GROMACS 4.6.5 nomenclature is employed.
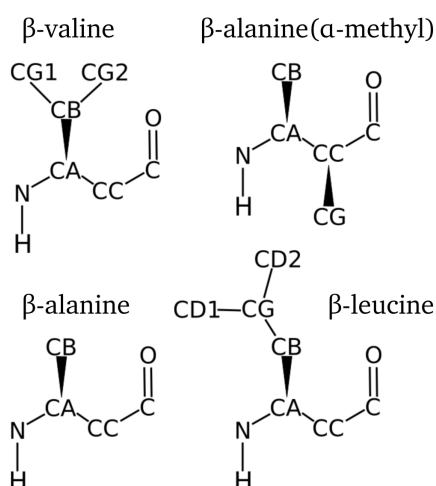
Fig. B.1.: Labeling of the heavy atoms of the amino acids used for assigning the atom type and the charges for the different force fields used. Hydrogens attached to carbons are omitted for clarity. They are labeled according to the carbon they are attached to.

B.1.1 Gromos force field

The parameters for the GROMOS53A6FF for the heptamer were determined using the Automated Topology Builder (ATB) and Repository version 1.2 [114].

B.1.2 Other force fields

Tab. B.1.: Atom types and charges of methanol.

	OPLS-AAFF [64]		CHARMM27FF [63]		AMBER03FF [62]		GAFF [65]	
Label	Atom-type	Charge	Atom-type	Charge	Atom-type	Charge	Atom-type	Charge
C	157	0.15	CT3	-0.04	CT	0.12	c3	0.12
HA	156	0.04	HA	0.09	H1	0.04	h1	0.04
HB	156	0.04	HA	0.09	H1	0.04	h1	0.04
HC	156	0.04	HA	0.09	H1	0.04	h1	0.04
O	154	-0.68	OH1	-0.65	OH	-0.65	oh	-0.65
H	155	0.42	H	0.42	HO	0.42	ho	0.42

Tab. B.2.: Atom types and charges of β-alanine.

	OPLS-AAFF		CHARMM27FF		AMBER03FF		GAFF	
Label	Atom-type	Charge	Atom-type	Charge	Atom-type	Charge	Atom-type	Charge
N	238	-0.5	NH1	-0.47	N	-0.49	n	-0.49
H	241	0.3	H	0.31	H	0.26	hn	0.26
CA	224	0.14	CT1	0.07	CT	0.34	c3	0.34
HA	140	0.06	HB	0.09	H1	0.05	h1	0.05
CB	135	-0.18	CT3	-0.27	CT	-0.20	c3	-0.20
HB1	140	0.06	HA	0.09	HC	0.05	hc	0.05
HB2	140	0.06	HA	0.09	HC	0.05	hc	0.05
HB3	140	0.06	HA	0.09	HC	0.05	hc	0.05
CC	136	-0.12	CT2	-0.18	CT	-0.35	c3	-0.35
HC1	140	0.06	HA	0.09	H1	0.09	hc	0.09
HC2	140	0.06	HA	0.09	H1	0.09	hc	0.09
C	235	0.5	C	0.51	C	0.59	c	0.59
O	236	-0.5	O	-0.51	O	-0.53	o	-0.53

Tab. B.3.: Atom types (A.-types) and charges of β-alanine(α-methyl).

Label	OPLS-AAFF A.-type	Charge	CHARMM27FF A.-type	Charge	AMBER03FF A.-type	Charge	GAFF A.-type	Charge
N	238	-0.5	NH1	-0.47	N	-0.44	n	-0.44
H	241	0.3	H	0.31	H	0.24	hn	0.24
CA	224	0.14	CT1	0.07	CT	0.19	c3	0.19
HA	140	0.06	HB	0.09	H1	0.10	h1	0.10
CB	135	-0.18	CT3	-0.278	CT	-0.26	c3	-0.26
HB1	140	0.06	HA	0.096	HC	0.07	hc	0.07
HB2	140	0.06	HA	0.096	HC	0.07	hc	0.07
HB3	140	0.06	HA	0.096	HC	0.07	hc	0.07
CC	136	-0.06	CT1	-0.096	CT	-0.09	c3	-0.09
HC	140	0.06	HA	0.09	H1	0.03	hc	0.03
CG	135	-0.18	CT3	-0.278	CT	-0.15	c3	-0.15
HG1	140	0.06	HA	0.096	HC	0.05	hc	0.05
HG2	140	0.06	HA	0.096	HC	0.05	hc	0.05
HG3	140	0.06	HA	0.096	HC	0.05	hc	0.05
C	235	0.5	C	0.51	C	0.52	c	0.52
O	236	-0.5	O	-0.51	O	-0.52	o	-0.52

Tab. B.4.: Atom types (A.-types) and charges of β-leucine.

Label	OPLS-AAFF A.-type	Charge	CHARMM27FF A.-type	Charge	AMBER03FF A.-type	Charge	GAFF A.-type	Charge
N	238	-0.5	NH1	-0.47	N	-0.40	n	-0.40
H	241	0.3	H	0.31	H	0.25	hn	0.25
CA	224	0.14	CT1	0.07	CT	0.08	c3	0.08
HA	140	0.06	HB	0.09	H1	0.10	h1	0.10
CB	136	-0.12	CT2	-0.18	CT	-0.12	c3	-0.12
HB1	140	0.06	HA	0.09	HC	0.04	hc	0.04
HB2	140	0.06	HA	0.09	HC	0.04	hc	0.04
CG	137	-0.06	CT1	-0.09	CT	0.23	c3	0.23
HG	140	0.06	HA	0.09	HC	-0.03	hc	-0.03
CD1	135	-0.18	CT3	-0.27	CT	-0.18	c3	-0.18
HD11	140	0.06	HA	0.09	HC	0.03	hc	0.03
HD12	140	0.06	HA	0.09	HC	0.03	hc	0.03
HD13	140	0.06	HA	0.09	HC	0.03	hc	0.03
CD2	135	-0.18	CT3	-0.27	CT	-0.12	c3	-0.12
HD21	140	0.06	HA	0.09	HC	0.02	hc	0.02
HD22	140	0.06	HA	0.09	HC	0.02	hc	0.02
HD23	140	0.06	HA	0.09	HC	0.02	hc	0.02
CC	136	-0.12	CT2	-0.18	CT	-0.16	c3	-0.16
HC1	140	0.06	HA	0.09	H1	0.04	hc	0.04
HC2	140	0.06	HA	0.09	H1	0.04	hc	0.04
C	235	0.5	C	0.51	C	0.56	c	0.56
O	236	-0.5	O	-0.51	O	-0.52	o	-0.52

Tab. B.5.: Atom types and charges of β-valine.

Label	OPLS-AAFF Atom-type	Charge	CHARMM27FF Atom-type	Charge	AMBER03FF Atom-type	Charge	GAFF Atom-type	Charge
N	238	-0.5	NH1	-0.47	N	-0.41	n	-0.41
H	241	0.3	H	0.31	H	0.23	hn	0.23
CA	224	0.14	CT1	0.07	CT	0.16	c3	0.16
HA	140	0.06	HB	0.09	H1	0.05	h1	0.05
CB	137	-0.06	CT1	-0.09	CT	0.15	c3	0.15
HB	140	0.06	HA	0.09	HC	0.01	hc	0.01
CG1	135	-0.18	CT3	-0.27	CT	-0.19	c3	-0.19
HG11	140	0.06	HA	0.09	HC	0.04	hc	0.04
HG12	140	0.06	HA	0.09	HC	0.04	hc	0.04
HG13	140	0.06	HA	0.09	HC	0.04	hc	0.04
CG2	135	-0.18	CT3	-0.27	CT	-0.17	c3	-0.17
HG21	140	0.06	HA	0.09	HC	0.04	hc	0.04
HG22	140	0.06	HA	0.09	HC	0.04	hc	0.04
HG23	140	0.06	HA	0.09	HC	0.04	hc	0.04
CC	136	-0.12	CT2	-0.18	CT	-0.25	c3	-0.25
HC1	140	0.06	HA	0.09	H1	0.08	hc	0.08
HC2	140	0.06	HA	0.09	H1	0.08	hc	0.08
C	235	0.5	C	0.51	C	0.56	c	0.56
O	236	-0.5	O	-0.51	O	-0.55	o	-0.55

B.2 Force field parameters for all monomers

The GROMOS53A6 [40] force field parameters for the monomers of all studied oligomers (except the heptamer) are compiled here in the GROMACS 4.6.5 format.

B.2.1 β-peptoid monomer

[PEP]
[atoms]
N N -0.31000 0
CD CH1 0.31000 0
CE CH3 0.00000 1
CF C 0.00000 2
CG1 C -0.14000 3
HG1 H 0.14000 3
CG2 C 0.00000 4
CI1 C -0.14000 5
HI1 H 0.14000 5
CI2 C 0.00000 6
CJ C -0.14000 7
HJ H 0.14000 7
CK1 C -0.14000 8
HK1 H 0.14000 8
CK2 C -0.14000 9
HK2 H 0.14000 9
CL1 C -0.14000 10
HL1 H 0.14000 10
CL2 C -0.14000 11
HL2 H 0.14000 11
CA CH2 0.00000 12
CC CH2 0.00000 13
C C 0.450 14
O O -0.450 14
[bonds]
N CA gb_21
N CD gb_21
CD CE gb_27
CD CF gb_27
CF CG1 gb_16
CF CG2 gb_16
CG1 CI1 gb_16
CI1 CJ gb_16
CG2 CI2 gb_16
CI2 CJ gb_16
CG2 CK2 gb_16
CI2 CK1 gb_16

CK1 CL1 gb_16
CK2 CL2 gb_16
CL1 CL2 gb_16
CG1 HG1 gb_3
CI1 HI1 gb_3
CJ HJ gb_3
CK1 HK1 gb_3
CK2 HK2 gb_3
CL1 HL1 gb_3
CL2 HL2 gb_3
CA CC gb_27
CC C gb_27
C O gb_5
C +N gb_11
[angles]
; ai aj ak gromos type
-C N CD ga_31
-C N CA ga_31
CD N CA ga_21
N CA CC ga_13
CA CC C ga_13
CC C O ga_30
CC C +N ga_19
O C +N ga_33
N CD CE ga_13
N CD CF ga_13
CE CD CF ga_15
CD CF CG1 ga_27
CD CF CG2 ga_27
CG2 CF CG1 ga_27
CF CG1 CI1 ga_27
CG1 CI1 CJ ga_27
CI1 CJ CI2 ga_27
CJ CI2 CG2 ga_27
CI2 CG2 CF ga_27
CF CG2 CK2 ga_27
CK1 CI2 CJ ga_27
CK1 CL1 CL2 ga_27
CL1 CL2 CK2 ga_27
CL2 CK2 CG2 ga_27
CK2 CG2 CI2 ga_27
CG2 CI2 CK1 ga_27

CI2 CK1 CL1 ga_27
HG1 CG1 CF ga_25
HG1 CG1 CI1 ga_25
HI1 CI1 CG1 ga_25
HI1 CI1 CJ ga_25
HJ CJ CI1 ga_25
HJ CJ CI2 ga_25
HK1 CK1 CI2 ga_25
HK1 CK1 CL1 ga_25
HK2 CK2 CG2 ga_25
HK2 CK2 CL2 ga_25
HL1 CL1 CK1 ga_25
HL1 CL1 CL2 ga_25
HL2 CL2 CK2 ga_25
HL2 CL2 CL1 ga_25

[impropers]
; ai aj ak al gromos type
N -C CA CD gi_1
C CC +N O gi_1
CF CG1 CG2 CD gi_1
CF CG2 CI2 CJ gi_1
CF CG1 CI1 CJ gi_1
CG1 CF CG2 CI2 gi_1
CG1 CF CI1 HG1 gi_1
CG1 CI1 CJ CI2 gi_1
CG2 CF CG1 CI1 gi_1
CG2 CF CI2 CK2 gi_1
CG2 CI2 CJ CI1 gi_1
CI1 CG1 CJ HI1 gi_1
CJ CI1 CI2 HJ gi_1
CK1 CG2 CJ CI2 gi_1
CK2 CL2 CL1 CK1 gi_1
CK2 CG2 CI2 CK1 gi_1
CG2 CK2 CL2 CL1 gi_1
CG2 CI2 CK1 CL1 gi_1
CL2 CL1 CK1 CI2 gi_1
CL2 CK2 CG2 CI2 gi_1
CK2 CL2 CG2 HK2 gi_1
CL2 CL1 CK2 HL2 gi_1
CL1 CK1 CL2 HL2 gi_1
CK1 CL1 CI2 HK1 gi_1

[dihedrals]
; ai aj ak al gromos type
-CC -C N CA gd_29
-C N CA CC gd_39
C CC CA N gd_34
CA CC C +N gd_10
-C N CD CF gd_39
N CD CF CG1 gd_10
-CC -C N CD gd_14
-C N CD CE gd_29
[exclusions]
; ai aj funct ; GROMOS 1-4
exclusions
CG1 CI2
CG1 CK2
CI1 CK1
CI1 CG2
CJ HK1
CJ CL1
CJ HG1
CI2 HI1
CI2 HK2
CI2 CL2
CI2 HL1
CG2 HJ
CG2 HG1
CG2 CL1
CG2 HL2
CG2 HK1
CK1 CK2
CK1 HJ
CK1 HL2
CK2 HL1
CL1 HK2
CL2 HK1
HG1 HI1
HI1 HJ
HK1 HL1
HL1 HL2
HL2 HK2

B.2.2 α-Ala

[ALA]
[atoms]
N N -0.31000 0

H H 0.31000 0
CA CH1 0.00000 1
CB CH3 0.00000 1
C C 0.450 2
O O -0.450 2
[bonds]
N H gb_2
N CA gb_21
CA CB gb_27
CA C gb_27
C O gb_5
C +N gb_10
[angles]
; ai aj ak gromos type
-C N H ga_32
-C N CA ga_31
H N CA ga_18
N CA CB ga_13
N CA C ga_13
CB CA C ga_13
CA C O ga_30
CA C +N ga_19
O C +N ga_33
[impropers]
; ai aj ak al gromos type
N -C CA H gi_1
CA N C CB gi_2
C CA +N O gi_1
[dihedrals]
; ai aj ak al gromos type
-CA -C N CA gd_14
-C N CA C gd_39
N CA C +N gd_40

B.2.3 β-HAla

[BAL]
[atoms]
N N -0.31000 0
H H 0.31000 0
CA CH1 0.00000 1
CB CH3 0.00000 1
CC CH2 0.00000 2
C C 0.450 3
O O -0.450 3
[bonds]
N H gb_2

N CA gb_21
CA CB gb_27
CA CC gb_27
CC C gb_27
C O gb_5
C +N gb_11
[angles]
; ai aj ak gromos type
-C N H ga_32
-C N CA ga_31
H N CA ga_25
N CA CB ga_13
N CA CC ga_13
CB CA CC ga_13
CA CC C ga_13
CC C O ga_30
CC C +N ga_19
O C +N ga_33
[impropers]
; ai aj ak al gromos type
N -C CA H gi_1
CA N CC CB gi_2
C CC +N O gi_1
[dihedrals]
; ai aj ak al gromos type
-CC -C N CA gd_14
-C N CA CC gd_39
C CC CA N gd_34
CA CC C +N gd_40

B.2.4 δ-Chin monomer

[CHI]
[atoms]
N NL -0.31 1
H H 0.31 1
CA C 0.00 2
CB1 C 0.05 3
CB2 C -0.14 4
HB2 HC 0.14 4
CC1 C 0 5
CC2 C -0.14 6
HC2 HC 0.14 6
CD C -0.14 7
HD HC 0.14 7
N2 NR -0.1 8

CE C 0.05 8
CF C -0.14 8
HF HC 0.14 8
CG C -0.14 8
HG HC 0.14 8
C C 0.45 10
O O -0.45 10
[bonds]
N H gb_2
N CA gb_10
CA CB1 gb_16
CA CB2 gb_16
CB1 CC1 gb_16
CB1 N2 gb_17
CB2 HB2 gb_3
CB2 CC2 gb_16
CC1 CD gb_16
CC1 CG gb_16
CC2 HC2 gb_3
CC2 CD gb_16
CD HD gb_3
N2 CE gb_17
CE CF gb_16
CE C gb_27
CF HF gb_3
CF CG gb_16
CG HG gb_3
C O gb_5
C +N gb_11
[angles]
-C N H ga_32
-C N CA ga_31
H N CA ga_32
N CA CB1 ga_27
N CA CB2 ga_27
CB1 CA CB2 ga_27
CA CB1 CC1 ga_27
CA CB1 N2 ga_27
CC1 CB1 N2 ga_27
CA CB2 HB2 ga_25
CA CB2 CC2 ga_27
HB2 CB2 CC2 ga_25
CB1 CC1 CD ga_27
CB1 CC1 CG ga_27
CD CC1 CG ga_27
CB2 CC2 HC2 ga_25
CB2 CC2 CD ga_27

HC2 CC2 CD ga_25
CC1 CD CC2 ga_27
CC1 CD HD ga_25
CC2 CD HD ga_25
CB1 N2 CE ga_25
N2 CE CF ga_27
N2 CE C ga_27
CF CE C ga_27
CE CF HF ga_25
CE CF CG ga_27
HF CF CG ga_25
CC1 CG CF ga_27
CC1 CG HG ga_25
CF CG HG ga_25
CE C O ga_30
CE C +N ga_19
O C +N ga_33
[impropers]
-C N CA CB2 gi_1
N -C CA H gi_1
CA N CB1 CB2 gi_1
CA CB1 CC1 CD gi_1
CA CB2 CC2 CD gi_1
CB1 CA CB2 CC2 gi_1
CB1 CC1 CD CC2 gi_1
CB1 CC1 CA N2 gi_1
CB1 CC1 CG CF gi_1
CB1 N2 CE CF gi_1
CB2 CA CB1 CC1 gi_1
CB2 CC2 CD CC1 gi_1
CB2 CA HB2 CC2 gi_1
CC1 HD CC2 CD gi_1
CC1 CB1 N2 CE gi_1
CC1 CG CF CE gi_1
CC1 CB1 CD CG gi_1
CC2 HC2 CB2 CD gi_1
N2 C CF CE gi_1
N2 CB1 CC1 CG gi_1
N2 CE CF CG gi_1
CE CG HF CF gi_1
CG CC1 CF HG gi_1
C CE +N O gi_1
[dihedrals]
-C N CA CB1 gd_39
N2 CE C +N gd_40
CE C +N +CA gd_14

B.2.5 α/γ-peptide monomers

[VAL]
[atoms]
N N -0.31000 0
H H 0.31000 0
CA CH1 0.00000 1
CB CH1 0.00000 1
CG1 CH3 0.00000 1
CG2 CH3 0.00000 1
C C 0.450 2
O O -0.450 2
[bonds]
N H gb_2
N CA gb_21
CA CB gb_27
CA C gb_27
CB CG1 gb_27
CB CG2 gb_27
C O gb_5
C +N gb_10
[angles]
; ai aj ak gromos type
-C N H ga_32
-C N CA ga_31
H N CA ga_18
N CA CB ga_13
N CA C ga_13
CB CA C ga_13
CA CB CG1 ga_15
CA CB CG2 ga_15
CG1 CB CG2 ga_15
CA C O ga_30
CA C +N ga_19
O C +N ga_33
[impropers]
; ai aj ak al gromos type
N -C CA H gi_1
CA N C CB gi_2
CA CG1 CG2 CB gi_2
C CA +N O gi_1
[dihedrals]
; ai aj ak al gromos type
-CA -C N CA gd_14
-C N CA C gd_39
N CA CB CG1 gd_34

N CA C +N gd_40

[GRSS]
[atoms]
N N -0.31000 0
H H 0.31000 0
CA C 0.00000 1
HA H 0.00000 1
CD1 CH2 0.00000 2
CD2 CH2 0.00000 2
CE1 CH2 0.00000 2
CE2 CH2 0.00000 2
CB C 0.00000 3
HB H 0.00000 3
CC CH1 0.00000 4
CF CH2 0.00000 5
CG CH3 0.00000 6
C C 0.450 7
O O -0.450 7
[bonds]
N H gb_2
N CA gb_11
CA HA gb_3
CB HB gb_3
CA CD1 gb_15
CD1 CE1 gb_26
CE1 CE2 gb_26
CE2 CD2 gb_26

CD2 CB gb_15
CA CB gb_16
CB CC gb_27
CC CF gb_27
CF CG gb_27
CC C gb_27
C O gb_5
C +N gb_11
[angles]
; ai aj ak gromos type
-C N H ga_32
-C N CA ga_31
H N CA ga_25
HA CA CB ga_25
HA CA N ga_25
HA CA CD1 ga_25
HB CB CA ga_25
HB CB CC ga_25
HB CB CD2 ga_25
N CA CB ga_19
N CA CD1 ga_31
CA CB CC ga_15
CA CD1 CE1 ga_8
CD1 CE1 CE2 ga_8
CE1 CE2 CD2 ga_8
CE2 CD2 CB ga_8
CD2 CB CA ga_8
CB CA CD1 ga_8

CD2 CB CC ga_15
CB CC CF ga_15
CC CF CG ga_15
CF CC C ga_15
CB CC C ga_15
CC C O ga_30
CC C +N ga_19
O C +N ga_33
[impropers]
; ai aj ak al gromos type
N -C CA H gi_1
CA N CD1 CB gi_2
CB CA CD2 CC gi_2
C CC +N O gi_1
[dihedrals]
; ai aj ak al gromos type
-CA -C N CA gd_14
-C N CA CB gd_39
C CC CB CA gd_34
CB CC C +N gd_40
CB CD2 CE2 CE1 gd_40
CD2 CE2 CE1 CD1 gd_40
CE2 CE1 CD1 CA gd_40
CE1 CD1 CA CB gd_40
CD2 CB CA CD1 gd_40
CA CB CD2 CE2 gd_40
CB CC CF CG gd_40
CC C +N +CA gd_14

B.3 Force field parameters for solvent molecules

The following tables contain the GROMOS 53A6 force field parameters of the solvent molecules used this work. The GROMACS 4.6.5 nomenclature is employed.

B.3.1 MeOH

[moleculetype]
; name nrexcl
MeOH 2

[atoms]
; nr type resnr residu atom cgnr charge mass
1 CMet 1 MeOH Me1 1 0.266000 15.035
2 OMet 1 MeOH O2 1 -0.674000 15.999
3 H 1 MeOH H3 1 0.408000 1.008

[constraints]
; ai aj funct c0
1 2 1 0.15300
2 3 1 0.10000
1 3 1 0.20770

[exclusions]
; ai aj ak
1 2 3
2 3 1
3 2 1

B.3.2 CHCl₃

[moleculetype]
; Name nrexcl
CHL 3

[atoms]
; nr type resnr residue atom cgnr charge mass
1 CLChl 1 CHL CL 1 -0.087 35.453
2 CChl 1 CHL C 1 0.179 12.011
3 HChl 1 CHL H 1 0.082 1.008
4 CLChl 1 CHL CL 1 -0.087 35.453
5 CLChl 1 CHL CL 1 -0.087 35.453

[bonds]
; ai aj funct
; 3 2 2 gb_39
2 1 2 gb_40
2 4 2 gb_40
2 5 2 gb_40
3 1 2 gb_47
3 4 2 gb_47
3 5 2 gb_47
1 4 2 gb_48
1 5 2 gb_48
4 5 2 gb_48

[angles]
; ai aj ak funct
3 2 1 2 ga_43
3 2 4 2 ga_43
3 2 5 2 ga_43

1 2 4 2 ga_44
1 2 5 2 ga_44
4 2 5 2 ga_44

B.3.3 CH₃CN

[moleculetype]
; name nrexcl
MeCN 2

[atoms]
; nr type resnr residu atom cgnr charge mass
1 D1 1 MeCN D1AN 1 0.000 9.49031
2 D2 1 MeCN D2AN 1 0.000 31.56239
3 MeAN 1 MeCN C3AN 1 0.206 0.00000
4 CAN 1 MeCN C4AN 1 0.247 0.00000
5 NAN 1 MeCN N5AN 1 -0.453 0.00000

[constraints]
; ai aj funct b0
1 2 1 0.26300

[dummies2]
; ai aj ak funct a
3 1 2 1 0.2652197
4 1 2 1 0.8203527
5 1 2 1 1.2652197

[exclusions]
3 4 5
4 5 3
5 4 3

Publications

C

1. *Mechanical unfolding pathway of a model β-peptide foldamer*
   L. Uribe, S. Jaschonek, J. Gauss, and G. Diezemann. J. Chem. Phys. **142**, 204901 (2015).

2. *Comparative study of the mechanical unfolding pathways of α- and β-peptides*
   L. Uribe, J. Gauss, and G. Diezemann. J. Phys. Chem. B **119**, 8313 (2015).

3. *Determining factors for the mechanical unfolding pathway of peptides, peptoids, and peptidic foldamers*
   L. Uribe, J. Gauss, and G. Diezemann. J. Phys. Chem. B, **120**, 10433 (2016).