TECHNISCHE UNIVERSITÄT MÜNCHEN

Department Chemie
Lehrstuhl für Theoretische Chemie

In *Silico* Prediction of Dissolution Rates of Pharmaceutical Ingredients: Micro-Kinetic Model Based on Spiral Dissolution

DISSERTATION

Berna Doğan

TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Theoretische Chemie

In *Silico* Prediction of Dissolution Rates of Pharmaceutical Ingredients: Micro-Kinetic Model Based on Spiral Dissolution

Berna Doğan

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften

genehmigten Dissertation

Vorsitzender: Univ.-Prof. Dr. Ville R. I. Kaila

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Karsten Reuter
2. Assistant Prof. Dr. Carlo Camilloni

Die Dissertation wurde am 26.02.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 17.06.2016 angenommen.

Abstract

The quest to discover substances that have healing powers have always been an integral part of human life. The understanding of human body functions and response to the healing compounds open a new era for this discovery process. Not only the discovery but the development and design of what is nowadays known as drug compounds have started to be modeled in a systematic way. The most encountered drawback in the drug development process is that the drug candidates fail to reach to the targeted part in the body or even remain inactive mostly due to the poor bio-availability, hence being inadequate candidates. Due to an *in-vivo-in-vitro* correlation between dissolution rates and bio-availability of drug formulations, dissolution testing is considered as a potential tool for drug development and design process. However, the dissolution testing itself is time consuming and costly, requiring the drug formulations to be synthesized beforehand. On the other hand, an *in silico* method that is able to predict the dissolution rates accurately would greatly accelerate the drug development process and reduce the costs to take a new formulation to the market.

To this end, in this work, a computational protocol is proposed to predict the dissolution rates of molecular crystals that found usages in pharmaceutical industry. Considering that organic crystals would dissolve along screw dislocations at low undersaturations, the classical spiral model of Burton-Cabrera-Frank (BCF) along with modifications as well as extensions is exploited. Other than readily available crystallographic parameters, only step velocities and critical lengths of step edges are required for the model. The step velocities are estimated considering the kink free energies of formation and detachment rate constants from the kink sites along the step edges formed by the screw dislocations on the crystal faces. As the estimation of the critical lengths of step edges also depends on kink free energies, only two parameters are necessary in addition to crystallographic constants. They are evaluated by accurate molecular dynamic (MD) simulation techniques. Then, the spiral dissolution model can be built to predict the dissolution rates of crystal faces. Depending on the interactions between molecules, different spiral shapes may be observed for the crystal faces.

In this work, as test cases for the spiral dissolution model, two model organic crystals are considered: aspirin, an active pharmaceutical ingredient (API) and alpha-lactose monohydrate, a widely used excipient. The predicted dissolution rates for faces of aspirin and alpha-lactose monohydrate show good agreement with measured dissolution rates, around one order of magnitude differences are observed. By employing several approximations to the protocol

developed, it is even possible to predict dissolution rates without extensive computational simulations. This renders the approach suitable for fast computational screenings of compounds used in drug formulations, and hence could notably reduce the time and cost of the drug development and design processes. The extensions of the protocol to other compounds would be more or less straightforward; however, the ability of the protocol to handle drug formulations that actually contain API and excipient together still remains to be tested.

Zusammenfassung

Die Aufgabe, Substanzen mit heilenden Kräften zu entdecken war schon immer ein integraler Teil des menschlichen Lebens. Das wachsende Verständnis der Funktionen und Heilungsmechanismen des menschlichen Körpers eröffnet eine neue Ära für eben solche Entdeckungen. Nicht nur die Entdeckung, sondern auch die Entwicklung und das Design von einzelnen Medikamentenbestandteilen erfolgen heutzutage zunehmend systematisch. Die in der Arzneimittelherstellung am häufigsten auftretende problematik besteht darin, dass die ausgewählten Kandidaten nicht den gezielten Körperteil erreichen oder aufgrund der geringen Bio-Verfügbarkeit inaktiv bleiben. Folglich stellen sie sich als inadäquat heraus. Aufgrund von *in-vivo-in-vitro* Korrelationen zwischen der Auflösungsrate und der Bio-Verfügbarkeit von Arzneimittelzusammensetzungen können Tests zur Ermittlung der Auflösung ein geeignetes Mittel für die weitere Entwicklung von Arzneistoffen sein. Das Testen der Auflösung ist jedoch zeitaufwendig und kostspielig, weil die Arzneimittel zuvor von Hand synthetisiert werden müssen, ehe die Tests durchgeführt werden können. Im Gegensatz dazu würde eine *in silico* Methode, welche die Auflösungsrate akkurat vorhersagen kann, die Dauer des Prozesses sehr verkürzen und zudem viele Kosten sparen.

In dieser Arbeit wird ein rechner-gestütztes Protokoll zur Vorhersage von Auflösungsraten molekularer Kristalle, welche in der Pharmaindustrie verwendet werden, vorgeschlagen. In der Hinsicht, dass sich organische Kristalle bei niedriger Untersättigung entlang von Schraubenversetzungen auflösen, findet hier das klassische Spiralen-Modell nach Burton-Cabrera-Frank (BCF) mit einigen Änderungen und Erweiterungen Verwendung. Anders als bei leicht verfügbaren kristallographischen Parametern sind in diesem Modell nur Stufen-Geschwindigkeiten und kritische Längen von Stufenkanten erforderlich. Die Stufen-Geschwindigkeiten werden beruhend auf der freien Energie der Kinken Erzeugung und Ratenkonstanten für die Ablösung von den Kinkenplätze bestimmt. Da die kritische Länge der Stufenkanten auch von der freien Kinken-Energie abhängt, werden zusätzlich zu den kristallographischen Konstanten somit nur zwei Größen benötigt. Sie werden unter Einsatz von akkuraten molekular-dynamischen (MD) Simulationstechniken berechnet. Im Anschluss kann das Spiralen-Auflösungs-Modell zur Vorhersage der Auflösungsrate von kristallinen Oberflächen eingesetzt werden. Je nach Interaktion zwischen den Molekülen kann man verschiedene Spiral-Formen der Kristalloberflächen beobachten.

In dieser Arbeit werden als Testbeispiele für das Spiralen-Auflösungs-Modell zwei organische Kristalle betrachtet: Aspirin, ein aktiver pharmazeutischer Stoff (API) und Alpha-Laktose Monohydrat, ein weit verbreiteter Hilfsstoff. Die vorhergesagten Auflösungsraten für die Oberflächen von Aspirin und Alpha-Laktose-Monohydrat zeigen gute Übereinstimmung mit den gemessenen Auflösungsraten im Rahmen einer Größenordnung. Mit Hilfe einiger Näherungen zu dem entwickelten Protokoll können leicht ungenauere Auflösungsraten sogar ohne umfangreiche rechnerische Simulationen vorhergesagt werden. Somit können Werte für die Bestandteile von Arzneimitteln schnell ermittelt und folglich die Kosten und der Zeitaufwand signifikant reduziert werden. Das Protokoll kann für eine große Klasse pharmazeutischer Bestandteile eingesetzt werden. Für Stoffe, die API und Hilfsstoffe gemeinsam beinhalten, benötigt das Protokoll noch Tests und gegebenenfalls Erweiterungen.

Contents

List of Figures

List of Tables

Abbreviations

ACPYPE	**A**nte**C**hamber **PY**thon **P**arser Interfac**E**
Alpha-LM	**Alpha-L**actose **M**onohydrate
AMBER	**A**ssociated **M**odel **B**uilding with **E**nergy **R**efinement
API	**A**ctive **P**harmaceutical **I**ngredients
ASA	**A**cetyl **S**alicylic **A**cid (aspirin as commercially known)
B3LYP	**B**ecke-**3** parameter-**L**ee-**Y**ang-**P**arr hybrid semi-empirical functional
BAR	**B**ennett **A**cceptance **R**atio
BCF	**B**urton-**C**abrera-**F**rank spiral model
BT	**B**ulk **T**ransport
CHARMM	**C**hemistry at **HAR**vard **M**acromolecular **M**echanics
COM	**C**enter **O**f **M**ass
CV	**C**ollective **V**ariable
DFT	**D**ensity **F**unctional **T**heory
ESP	**E**lectro**S**tatic **P**otential
FHI-aims	**F**ritz-**H**aber-**I**nstitut ***ab initio*** **m**olecular **s**imulations package
GAFF	**G**eneralized **A**MBER **F**orce **F**ield
GGA	**G**eneral **G**radient **A**pproach
GROMACS	**GRO**ningen **MA**chine for **C**hemical **S**imulations package
LDA	**L**ocal **D**ensity **A**pproximation
kMC	**k**inetic **M**onte-**C**arlo method
MD	**M**olecular **D**ynamics
MP2	**M**øller–**P**lesset perturbation theory **2**nd order
MZ	**M**eta-stable **Z**one boundary
OPLS-AA	**O**ptimized **P**otentials for **L**iquid **S**imulations-**A**ll **A**tomic
PBC	**P**eriodic **B**ond **C**hains
PBE	**P**erdew-**B**urke-**E**rnzerhof semi-local functional
PES	**P**otential **E**nergy **S**urface
PME	**P**article **M**esh-**E**wald method
PRD	**P**arallel **R**eplica **D**ynamics
QM	**Q**uantum **M**echanics
QM/MM	hybrid **Q**uantum **M**echanical and **M**olecular **M**echanics methods

R.E.D server	The **R**ESP and **E**SP charge **D**river server
RESP	**R**estraint **E**lectro**S**tatic **P**otential
SPC	**S**imple **P**oint **C**harge
TI	**T**hermodynamics **I**ntegration method
vdW	**v**an **d**er **W**aals interactions

1 Introduction

The birth of pharmaceutical industry as known today has been led by the discovery of the effect of so called 'active principles' inside herbs and/or potions that were used as traditional remedies before the twentieth century [1]. Therefore, in the beginning the focus was on the extraction of these active principles or ingredients, the pure chemicals responsible for the medicinal properties. However, with the rapid development of synthetic organic chemistry, minor changes to the natural products by mostly altering the functional groups were made in order to improve their effectiveness and/or reduce the undesired side-effects. Later, the analogues of the natural products that are described as drug prototypes or candidates are designed in which only the essential part of the molecule is retained. Most of these studies, however, are carried out on a trial and error basis until scientists started to have better understanding of the body functions and response at both cellular and molecular levels [1-2]. This understanding has affected the production and development of new drug candidates in a way that they are intended to be more suitable for targets in the body by identifying and interacting with them. By this way, the previous trial and error approach to discover new candidates was improved, though the method could still not be considered as infallible.

However, the discovery of lead or active pharmaceutical ingredients (API) was not enough itself to ensure that the drug candidate would be able to identify, interact, and block etc. the target part in the body. Many of the potential drug candidates have failed as a result of difficulties in penetrating barriers or not arriving or even remaining inactive at the site of action, mostly due to poor pharmacokinetics and bioavailability [3-4]. As one of the essential tools in pharmacokinetics studies, bio-availability relates the fraction of API reaching the blood supply in a given period of time to the amount originally administered [2]. The bio-availability of an API is determined by a variety of factors, including physical properties such as hydrophobicity, pK_a, solubility, and dissolution rate of the crystalline form of the API [5-6]. The excipients chosen to be included in drug formulations and traditionally considered as inert compounds have also been shown to affect the bioavailability [7]. Among these factors, the dissolution rate has been demonstrated to be directly correlated to bio-availability [8-9]. This *in vitro-in vivo* correlation between dissolution rates and bio-availability of the drug formulations led dissolution testing to become a part of the standard procedure for quality control tests [10-14]. In addition to being a useful control tool, dissolution testing can also be a predictive tool for the efficiency of new drug formulations [15] and provide useful information during different stages of drug development [16]. However, being

an empirical method, it has the drawback that for every dissolution test, the sample product must have already been physically made. This significantly adds to the time and cost of taking a new drug to market. A theoretical model, on the other hand, can help to explain experimental observations and to predict the likely outcome of a new formulation at the design stage. By the use of a reliable such *in silico* method, it would be possible to reduce the number of physical tests thst have to be conducted as well as the total cost and time of drug development [17]. Additionally, *in silico* models may help to understand the underlying processes during dissolution and to be able to control them since pharmaceutical industry not only demands formulations with high efficiency but also wants to manipulate the release of APIs in order to maintain constant drug concentration [9].

With the advances in computational power, there have been some modelling studies of the dissolution process on a molecular level [18-23], in addition to previous conceptual models [24-27]. However, most of the studies on a molecular level focus on small molecules like urea [22] or ionic compounds like NaCl [20] or $BaSO_4$ [23]. Very few of the studies with emphasis on solid/liquid interfaces are on molecular crystals that have found applications in pharmaceutical industry such as aspirin [19] or acetaminophen [18]. Though they have provided insightful information of dissolution process using common molecular dynamics (MD) techniques especially about ‘corner and edge effects’, there is up to date no established micro-kinetic model with predictive quality for dissolution rates that could replace experimental studies.

In contrast, for the reverse process of dissolution, i.e. crystallization, mechanistic models in particular by Doherty *et al*. have recently been developed to calculate the microscopic growth rates of individual facets of organic crystals [28-34]. Using the widely recognized spiral growth model of Burton-Cabrera-Frank (BCF) [35] they related the growth rates of the relevant facets to predict the morphology of the crystals in good agreement with experiments. The model Burton, Cabrera and Frank developed more than six decades ago is considered the principal growth model at low supersaturations. Based on BCF, crystals of organic molecules grow via spirals emanating from pervasive screw dislocations [34]. These defects are the primary growth parts on surfaces and specifically kink sites on step edges are the incorporation sites for molecules. Therefore, growth kinetics depends on step kinetics on edges which was justified by different studies [33, 36-37]. Though the BCF mechanistic considerations were originally for growth, the dissolution process is nothing but the ‘reverse’ process of growth and can equally captured by considering the retraction of steps. In addition parameters can nowadays also be determined by more accurate MD-based methods and therewith lend material-specific predictive quality to the BCF model.

Hence, in this work, the dissolution mechanism at low undersaturations will be considered as the reverse of the spiral growth mechanism. The aim of this thesis is the development of a predictive-quality computational approach for the absolute dissolution rates of crystals by such a micro-kinetic model description of the dissolution process. The method is applied to two prototypical compounds; aspirin as an active pharmaceutical ingredient and alpha-lactose monohydrate, an excipient used in drug formulations. The ambitious aim is to be able to predict the dissolution rates for novel drug candidates *in silico*, without synthesizing them and therewith to reduce the tests necessary for the drug development process. This thesis is structured as follows:

Part I of this work will be an overview of the growth and dissolution models and Chapter 3 is specifically concerned with the BCF spiral model [35]. The theory modifications and advancements of this classical model for the case of crystal dissolution will also be emphasized in this chapter.

In Part II, the theoretical background information is outlined. Starting from the description of classical equation of motions, the basic concepts of molecular dynamic simulations are explained. The temperature and pressure control during the simulations via thermostats and barometers are discussed. Chapter 5 will be concerned with the differentiation of *ab initio* models and semi-empirical pre-defined force fields. In Chapter 6, a brief overview over thermodynamics and statistical mechanics in the context of molecular simulations is given, focusing in particular on free energy calculations as well as accelerated molecular dynamics techniques as employed in our computational approach to dissolution.

Part III will focus on the development of force fields and the description of the micro-kinetic model based on spiral dissolution. In the first chapter of this part, Chapter 7, the methods for parameter evaluation, namely the calculation of kink free energies and detachment rate constants from kink sites, will be explained. In Chapter 8, after the development and testing of force fields is found to be adequate, the model is applied for the prediction of dissolution rates for aspirin, an active pharmaceutical ingredient. In a second demonstration of the model, in Chapter 9, the dissolution of alpha-lactose monohydrate, a widely used excipient, is considered. For both cases, the spiral forms that can be observed on the considered facets are discussed in detail after the evaluation of the dissolution rates for such spiral forms. In addition, possible approximations to the method without significant sacrifices in accuracy are considered so that the model would be applicable for fast screening approaches. Chapter 10 will summarize the findings and provide an outlook towards future work.

PART I
Crystal Growth and Dissolution

Models

2 An Overview of Mechanisms

Crystal growth and dissolution involves the transfer of matter between solid and liquid phases due to the imbalance of potentials at the interface [38]. Although both phenomena proceed through a number of processes (i.e. adsorption, solvation/desolvation etc.), the major rate-limiting processes involved are either the volume diffusion (also referred to as bulk transport) or interfacial processes [38-39]. Volume diffusion or bulk transport is the transfer of solute molecules to and from the interface. The interfacial processes consist mainly of surface diffusion and the exchange processes of surface molecules. During crystal growth from solution molecules transported from the solution are incorporated into the crystal surface, whereas during dissolution the reverse process of dispersion of crystal surface molecules are observed. The overall growth and dissolution rate is limited by the slower of the two process types. Hence, there are two main mechanisms for crystal growth and dissolution: Rough (normal) growth/dissolution that is transport limited, or lateral (layered) growth/dissolution that is surface reaction limited [28, 40].

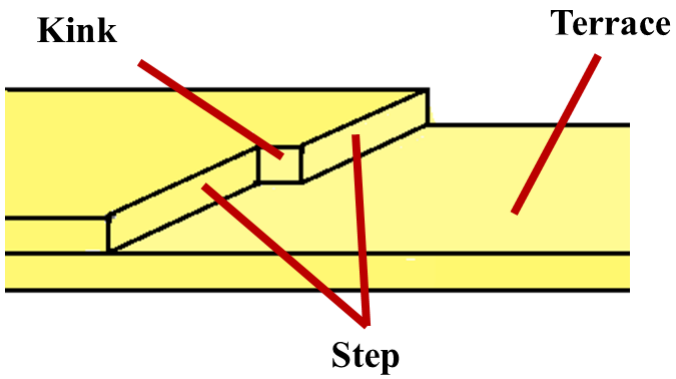

Figure 2.1: Illustration of relevant defect structures: Steps and kink sites at a crystal surface.

Exactly which process and mechanism become the dominating one depend on both the external factors, specifically super/undersaturation and temperature at the surface, and the internal factors, mainly the crystal faces considered [28]. The crystal faces are classified depending on their morphological importance and solid-state conditions as proposed by Hartman and Perdok [41]. Their proposed periodic bond chain (PBC) theory involves the description of solid-state conditions by considering the noncovalent bond interactions. Using PBC theory, faces can be classified based on the number of coplanar PBCs they contain within a slice of thickness d_{hkl}. If a face contains two or more PBC vectors, then it is called an F face. Such faces are stable and macroscopically flat. When a face contains only one PBC vector, then this face is an S face. On S faces, surface

defects, mainly steps and kink sites, are present as illustrated in Figure 2.1.These defect sites are the primary growth/dissolution sites. Crystal faces that do not contain any PBC vector are called K faces, as they are rough and completely *k*inked. While S and K faces always grow/dissolve via rough growth/dissolution mechanisms, the mechanism for F faces may change. At lower temperature and super/undersaturation i.e. below the roughening transition [42], the growth/dissolution mechanism is layered growth/dissolution. Beyond the roughening transition the mechanism instead is rough and the crystal face would have a rough surface macroscopically. In Figure 2.2 the dependence of the mechanism on the supersaturation for crystal growth is illustrated; an equivalent graph but depending on undersaturation would be valid for crystal dissolution [28-29].

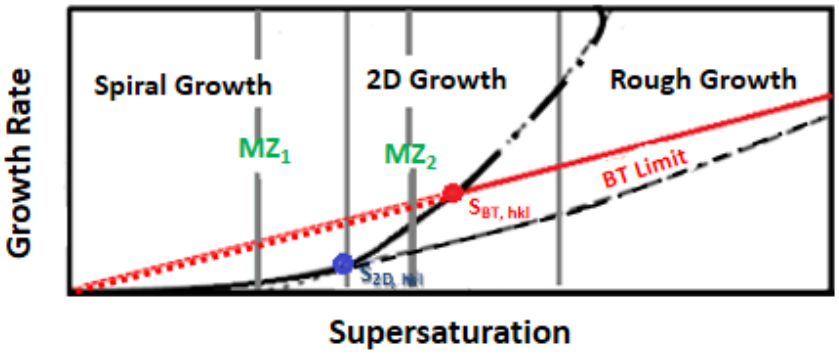

Figure 2.2: Growth mechanisms for a flat (F) face, as a function of supersaturation. The solid line is the growth rate. The red line indicates the bulk transport (BT) limit; the black curves represent the mechanism-dependent surface integration limits; the points $S_{2D,hkl}$ and $S_{BT,hkl}$ indicate the onset of supersaturation for growth by 2D nucleation and bulk transport limited growth, respectively; the vertical green lines marked MZ_1 and MZ_2 are two possible positions for the metastable zone boundary. $S_{2D,hkl}$ does not necessarily occur at lower supersaturation than $S_{BT,hkl}$; does. Depending on system conditions, $S_{BT,hkl}$ can occur at lower supersaturation than $S_{2D,hkl}$ does. Figure inspired by ref. [29].

During the layered growth/dissolution mechanism, the growth/dissolution proceeds via the movement of steps containing kink sites. The step motion on the faces is along the surface lateral direction until they reach a boundary and layers are formed by continuous movements [40]. This mechanism is slower than rough growth/dissolution since the density of kink sites is much lower on F faces. Nevertheless, kinks are still present and/or formed by thermal fluctuations at temperatures other than 0 K. Layered growth/dissolution occurs either via 2D-nucleation or via spiral formation, depending on the method of edge generation. The individual molecules (or rather growth/dissolution units) adsorbed onto the surfaces may diffuse towards each other, agglomerate

and islands (for growth) or pits (for dissolution) may be formed. Once these 2D-structures have grown larger than their critical size, it becomes thermodynamically favorable for other molecules to attach to these islands or pits. Based on classical nucleation theory, the free energy barrier to form a post-critical 2D-structure from a perfect surface can be calculated as

$$\Delta G^{\text{island/pit}} = \Delta n_{\text{cryst}}\,\Delta\mu + \sum_{i=1}^{N_{\text{steps}}} l_i \Delta G_i^{\text{step}} \qquad 2.1$$

Here, $\Delta\mu$ is the chemical potential difference between a molecule in the crystal and in solution at the relative saturation value $S = c/c_{\text{sat}}$. It can be related to the saturation value as $\Delta\mu \simeq k_B T \ln(\sigma)$ where k_B is the Boltzmann constant, T is the temperature, c and c_{sat} are the actual and saturation concentrations. Δn_{cryst} denotes the change in the number of crystal molecules upon creating the 2D defect; N_{steps} is the number of step edges taken into account and $\Delta G_i^{\text{step}}$the free energy cost for creating the surrounding edges of length l_i [43]. The barrier also depends on the shape of the islands/pits adding the shape factor into the equation. Depending on the relative saturation σ the barrier may decrease and easily be overcome. However, for low super/undersaturations, the barrier will be very large and the formation of 2D structures will not be possible. Yet, it is known that crystal growth and dissolution are still observed at these conditions [35]. In these cases, screw dislocations, a common topological defect of crystal lattices [40] act as continuum sources of edges that are exposed to solution at each dislocation. When the lengths of the edges reach the critical values, the edges will start moving and expose new edges. By repeating this movement and exploitation of edges, spiral-like structures evolve on the faces. At higher super/undersaturation, the free energy barrier for 2D nucleation will decrease and at a critical super/undersaturation value $S_{2\text{D},hkl}$ (see Figure 2.2) the mechanism dominating the dissolution/growth becomes the 2D nucleation mechanism [29]. In the next chapter, the spiral mechanism proposed first by Burton, Cabrera and Frank in 1951 will be explained in detail along with the mechanistic growth models developed by the group of Doherty [28, 32, 44-45]. Though a kinetic asymmetry of growth and dissolution rates at equal distances from the equilibrium saturation ratio has been indicated in different studies [33, 46-47], the differences especially in the regime of low super/undersaturation are small, around one order of magnitude. Also, these differences in rates do not mean that the mechanism cannot be microscopically reversed to each other [48]. Therefore, in this work, dissolution is considered as the time-reversed process of crystal growth and the here described concepts of growth mechanisms are transferred in an analogous way to the dissolution process [43].

3 Spiral Growth and Dissolution

In 1931, Volmer and Schultze [49] observed the crystal growth for iodine at low supersaturation $S \approx 1.1$ that is below the minimum value predicted for growth to occur. As this observation could not be explained by a 2D nucleation mechanism, an alternative layered growth mechanism was proposed by Frank [50]. He proposed that when dislocations of 'screw' form are present on the face, the face would always expose terraces on which growth could continue. The need for fresh new nuclei or islands would never arise under these conditions. If these types of dislocations emerge on a crystal face, then that face could grow perpendicularly 'up a spiral staircase'. Later, together with Burton and Cabrera, he further developed this spiral mechanism [35], which is nowadays widely recognized as BCF model after their developers. Since its proposition more than six decades ago, several modifications and extensions have been presented, including the works done by Chernov [51], Nielsen [52] and Synder and Doherty [32]. Though the emphasis for these modifications was on the spiral growth mechanism, Christoffersen *et al*. [48, 53] and Zhang and Nancollas [33] have also considered the dissolution process along with the growth process but for ionic compounds mostly. On the other hand, different crystals and specifically organic crystals have been considered for the development of growth models. More mechanistic and simpler methods especially for computer simulations have hitherto been presented for the growth process. Hence, in this thesis, mostly these methods are considered by assuming the dissolution process to be the time-reverse of the growth process and transferring the concepts analogously to the dissolution case. When there are some changes that needed to be taken into account when considering the reverse process, they will be mentioned explicitly in the following explanations.

The growth/dissolution rate through the spiral mechanism can be expressed as

$$R = \frac{h}{\tau}, \tag{3.1}$$

where τ is the rotation time of the crystal and h is the step height. The expression for the rotation time of a convex N-sided polynomial spiral is

$$\tau = \sum_{i=1}^{N} \frac{l_{c,i+1}\ \sin(\alpha_{i,i+1})}{v_i}, \tag{3.2}$$

with $l_{c,i+1}$ the critical length of edge $i+1$, $\alpha_{i,i+1}$ the angle between edges i and $i+1$, and v_i the step velocity of edge i. The numbering scheme is such that the first edge to move is labelled as 1 and the last is labelled as N. Therefore, each edge i exposes edge $i+1$. The rotation direction is defined by the advancement of the outward normal of edge $i+1$ relative to that of edge i and

hence spirals rotate in either clockwise or anticlockwise direction. The step velocity profile as first proposed by Kaischew and Budevski [54] and later by Voronkov [55] is used

$$\begin{aligned} v &= 0 \qquad l \leq l_c \,, \\ v &= v_\infty \qquad l \leq l_c \,, \end{aligned} \tag{3.3}$$

which assumes that a spiral edge starts moving at a constant velocity only after reaching a critical length (Figure 3.1).

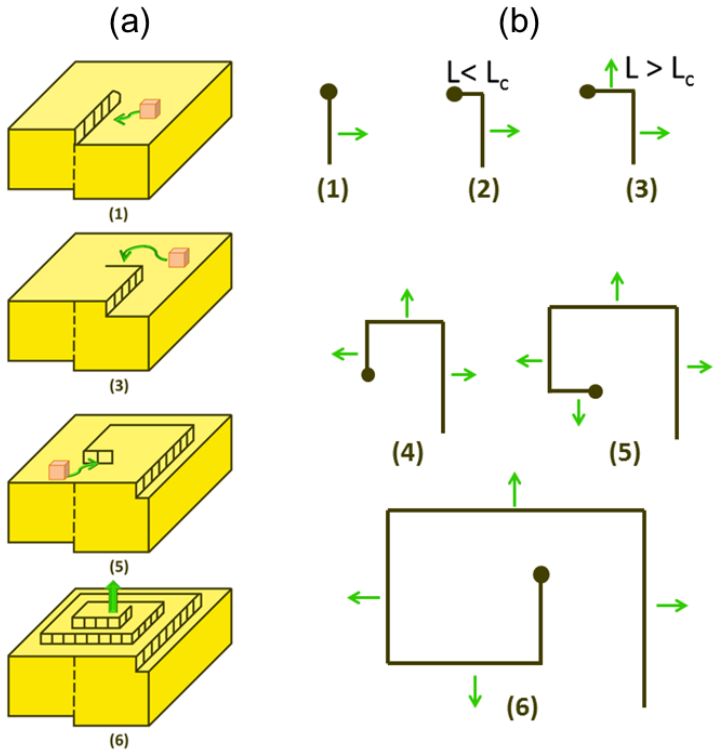

Figure 3.1: Schematics of growth of a square spiral (a) perspective view and (b) top view. The conditions for step movement are explained in the text.

Predicting the critical length of a spiral edge is required to determine edge movement, as the edges are assumed to have a velocity of zero at below critical lengths (eq. 3.3). There are a variety of methods for the determination of critical lengths including kinetic methods [56] as well as various thermodynamic methods [32, 51, 57-59]. For this thesis, the thermodynamic approach of Lovette and Doherty [59] is applied. In this approach, the decrease or increase in the free energy of the system with the addition/retraction of a crystal row is considered. If the addition/retraction results in a decrease in free energy, only then the edge is assumed to grow. The critical length is the length at which the free energy of the system is zero under the addition/retraction of a row of molecules. Based on assumptions, the expression for the critical length is [59]

$$l_{c,i} = \frac{2a_{e,i}\Delta G_{\text{kink}}}{RT\ln(S)}, \qquad 3.4$$

with $a_{e,i}$ the distance between molecules along edge i (see Figure 3.2), ΔG_{kink} is again the work or energy change required to create one kink site along an edge by rearrangement of the structure [32, 35], and S is the saturation ratio. As the saturation ratio is less than one under dissolution conditions, eq. 3.4 would give negative values for dissolution. To get positive dissolution rates, a negative sign thus has to be added either while evaluating the critical lengths or while calculating the overall dissolution rate, R [60].

The calculation of step velocities depends on the crystal structure. Most of the studies mentioned before consider simple cubic lattice crystals, i.e. so-called *Kossel* crystals for model development. The work of Kuvadia and Doherty [61] is therefore especially significant since they not only consider *non-Kossel* crystals but also classify them based on the arrangements of the molecules. As shown by their study, the behavior of crystals differs depending on the lattice crystal structure and molecular arrangement. For this reason it is first worthwhile to clarify these concepts and highlight the differences in crystal growth between them.

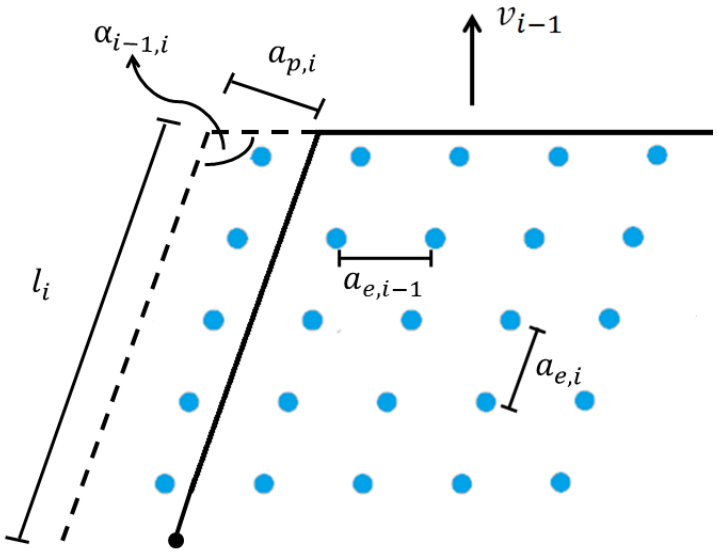

Figure 3.2: Schematic top view of a crystal face (with blue points denoting individual molecular units in the crystal), illustrating the crystallographic variables used for critical length and step velocity calculations. The dashed lines show the advance of edge $i-1$.

A *Kossel* crystal is defined as a simple cubic crystal lattice structure with equal bond strengths in all six directions [61]. Except for one of the polymorphs of the metal polonium [62], no real crystal structure fulfills the *Kossel* crystal requirements. All such *Non-Kossel* crystals are thus

further classified as centrosymmetric or non-centrosymmetric depending on the arrangements of the molecules constituting the lattice [44]. Centrosymmetric molecules/crystal systems [63] contain an inversion center such that for every point (x, y, z) in the molecule/crystal structure there is an indistinguishable point (-x, -y, -z). Non-centrosymmetric molecules/crystal systems on the other hand do not contain an inversion center. Here, it is necessary to note the distinction between the symmetry of molecules (or growth/dissolution units) and the symmetry of crystal systems. A centrosymmetric crystal system means that the space group of the crystal system contains an inversion center(s). Centrosymmetric molecules on the other hand have a center of symmetry inside their own structure irrespective of whether the crystal structure belongs to a centrosymmetric or non-centrosymmetric space group. The growth/dissolution units may not always be single molecules but sometimes they can be dimers, tetramers etc. These kinds of structures may also be centrosymmetric though the molecule itself is not. The symmetry of the growth/dissolution units dictates whether the crystal structure behaves in *Kossel* or *non-Kossel* like manner. Depending on this manner, the growth/dissolution units experience different driving forces, i.e. isotropic or non-isotropic, and the step velocity expression changes.

3.1 Step Velocities for Centrosymmetric Growth/Dissolution Units

When crystal structures contain centrosymmetric units, the bonds are equal in strength *along a single direction* on all facets [44-45]. At any kink site, exactly half of the bonds available is exposed to solution, whereas the remaining half is still interacting with the solid [32]. The energy for these kink sites is correspondingly half of the lattice energy. Therefore, these kink sites are defined classically as *half crystal position* [61]. As such, the same bonds are exposed at any kink site on the face and hence all kink sites are equivalent. They experience an isotropic driving force irrespective of the edge direction considered. In fact, different faces for crystal structures containing centrosymmetric units also expose half of the bonds at any kink sites. There is only one type of kink site for the whole crystal lattice. Hence, the driving force for incorporation/disincorporation into/out of kink sites is the same for every edge. The spiral growth/dissolution is isotropic, implying it will have an equal effect on all orientations. It can be said that these crystals behave in a *Kossel-like* fashion. Therefore, the approaches used for *Kossel* crystals would be valid.

The step velocity of the ith edge for centrosymmetric systems is given by [32, 44]

$$v_i = a_{p,i} v_0 \left(\frac{a_e}{\delta}\right)_i \exp\left(\frac{\Delta U}{k_B T}\right) V_m c_{sat} (S-1), \qquad 3.5$$

where $a_{p,i}$ is the distance an edge i propagates with the addition of a single row of growth units (see Figure 3.2) and v_0 is the frequency of attachment and detachment attempts. The average spacing between kinks is δ and the molecular spacing along the step is a_e; thus, the quantity $\frac{a_e}{\delta}$ is the probability of finding a kink at any location along the step edge; in other words, a measure of the density of kink sites (ρ_i) on the particular edge. c and c_{sat} are actual and equilibrium concentrations of the solute in the solution near the edge that is growing (with $S = c/c_{\mathrm{sat}}$ the saturation ratio as before) and V_{m} is the molar volume of the solute. Finally, ΔU is the energy barrier for the incorporation of molecules into a kink site with k_{B} as Boltzmann constant and T as temperature. For the derivation of this step velocity expression, it is assumed that classical kinks along step edges are the primary sites of incorporation/disincorporation into/from the crystal and the detachments from non-kink sites as well as the attachment to non-kink sites are neglected.

The general step velocity equation, eq. 3.5, can be rewritten by grouping the terms for kink density and kink rates [44-45]

$$v_i = a_{p,i}\rho_i u_i \,, \tag{3.6}$$

where ρ_i and u_i are kink density and the net kink rate for edge i, respectively. The density of kinks along an edge can be calculated as

$$\rho_i = [1 + 0.5\exp(\Delta G_{\mathrm{kink}}/RT)]^{-1}, \tag{3.7}$$

where ΔG_{kink} is the work or energy change required to create one kink site along an edge by rearrangement of the structure [32, 35], R and T are gas constant and temperature respectively. It should also be noted that kink rearrangements are assumed to occur faster than the incorporation/disincorporation into/out of kink sites. Therefore, kinks are assumed to be in their most probable distribution and are Boltzmann distributed.

The net kink rate or the net flux is the difference between two opposing fluxes; the incorporation of solute molecules into kink sites and the disincorporation of solute molecules from the kink sites [29]

$$u_i = j_i^+ - j_i^- \,, \tag{3.8}$$

where j_i^+ and j_i^- are the fluxes into and out of kink sites for edge i, respectively. These fluxes are given by the following equations [64]

$$j_i^+ = V_{\mathrm{m}}\, c\, k_i^+ = x\, k_i^+ \,, \tag{3.9}$$

and

$$j_i^- = (1 - V_{\mathrm{m}}\, c)k_i^- = (1 - x)k_i^-, \tag{3.10}$$

with k_i^+ and k_i^- as the rate constants for the attachment and detachment into and out of the kth kink site along a step edge i, and V_{m} is the molar volume of the solution. $V_{\mathrm{m}}c$ is then the mole fraction

of solute molecules in the solution and can be denoted by x. Here the presence of external species such as impurities or additives are ignored for simplicity. At equilibrium, the opposing fluxes balance each other and edges are stagnant with $v = 0$. Hence, the following equation derives

$$\left(1 - x_{\mathrm{eq}}\right)k_i^- = x_{\mathrm{eq}}k_i^+, \tag{3.11}$$

with x_{eq} as the equilibrium mole fraction. Using this balance, the following relationship between equilibrium mole fraction and detachment and attachment rates can be obtained

$$x_{\mathrm{eq}} = \frac{k_i^-}{(k_i^+ + k_i^-)}, \tag{3.12}$$

The mole fraction of solute molecules at any time in solution can also be expressed depending on the saturation concentration and equilibrium mole fraction,

$$x = Sx_{\mathrm{eq}}\ , \tag{3.13}$$

where $x_{\mathrm{eq}} = V_m c_{\mathrm{sat}}$. When eq. 3.12 and eq. 3.13 are combined and put into eq. 3.9 and 3.10, the fluxes into/out of kink sites can be written for different saturation constants. When these new flux expressions are put into the net kink rate expression, eq. 3.8, and after some rearrangements and cancelations, a simple expression for the kink rate of centrosymmetric growth units is obtained. This net kink rate for the spiral growth mechanism

$$u_i = (S - 1)k_i^-, \tag{3.14}$$

just depends on the detachment rates and saturation ratio.

For the spiral dissolution mechanism, eq. 3.14 will again give negative values since the saturation constant will be lower than one. Therefore, the net kink rate for dissolution is considered as the subtraction of fluxes into kink sites from fluxes out of kink sites in eq. 3.8 and the simple expression for the edge velocity for spiral dissolution is

$$u_i = (1 - S)k_i^-. \tag{3.15}$$

3.2 Step Velocities for Non-Centrosymmetric Growth/Dissolution Units

The crystal systems that contain non-centrosymmetric growth units have facets with complex bonding structures [44]. The bond strength *along a single direction* may not necessarily be same, i.e. the bond chains may be comprised of bonds of unequal strength. Moreover, each half of the slide broken by these bond chains is not essentially identical. It is common for adjacent molecules to have different sets of bonds and behave as different growth/dissolution units. Hence, kink sites are not *half-crystal positions* and more than one type of kink sites can be expected along the same edges on faces. The different faces would reveal different kink types and as a whole, the crystal

lattice would contain more than one kink type. As a consequence of this, the driving force for growth/dissolution is anisotropic, and the incorporation or disincorporation of solute molecules into or out of kink sites will depend on the type of kink sites.

For crystals with non-centrosymmetric growth/dissolution units, the step velocity expression, (eq. 3.6) needs to be re-assessed to account for this anisotropic driving force. The kink density cannot be calculated using the classical Boltzmann formula, eq. 3.7. Instead, the different microstates (ways of rearrangement from a straight edge) that correspond to each kink site type must be counted. By using the Boltzmann distribution the probabilities for each kink type can be evaluated and the sum of all their probabilities provides the total number density of all kink sites for an edge [44]. The net kink rate also needs to account for the transformations of kink sites. Depending on the termination of bond chains, the kink site type changes, hence detachment/attachment of kink site molecules will result in different types of kink site molecules. Therefore, the probabilities of each kink type and the corresponding states along the edges have to be evaluated first. Then, the fluxes into and out of the specific positions are considered. For position k (that is the bond chain terminated by kink type k) the two opposing fluxes are incorporation of solute molecules into the kth kink position and disincorporation of molecules from the $k+1$th kink site to lead to a chain being terminated by the kth kink type. The net kink rate from an edge i is then

$$u_i = (P_k\, j_k^+ - \, P_{k+1}\, j_{k+1}^-)_i, \qquad 3.16$$

where j_k^+ and j_k^- are the fluxes into and out of the kth kink site, and P_k and P_{k+1} are the probability of kth site and $k+1$ occurring on edge i, respectively. As the attachment rate is assumed to be independent of the exact position of attachment [65], eq. 3.16 is rewritten as

$$u_i = (P_k\, j^+ - \, P_{k+1}\, j_{k+1}^-)_i\,. \qquad 3.17$$

It should be noted that the kink rate, u_i, is the same regardless of which kink sites are chosen in eq. 3.16; that is, the kink rate is for an entire edge with all its kink sites [44].

The probabilities of each state can be calculated by setting up a steady-state balance. The net rate of transition into the kth kink site is exactly equal to the net rate of transition out of the kth kink site at steady state. Then a 1D master equation for the steady-state can be written as [45]

$$P_k(j^+ + \, j_k^-) = P_{k+1}\, j_{k+1}^- + \, P_{k-1}\, j^+\,. \qquad 3.18$$

Solving this equation by using the condition that the sum of all probabilities has to be equal to one and putting the resulting probabilities into eq. 3.16, the net kink rate of edge i is

$$u_i = \frac{(j^+)^n - \prod_{k=1}^{n} j_{k,i}^-}{\sum_{r=1}^{n} (j^+)^{n-r} \, (j_i^-)^{r-1}} \,. \tag{3.19}$$

Here n is the number of kink types along edge i and

$$(j_i^-)^{r-1} = \sum_{k=1}^{n} (j_k^- \, j_{k+1}^- \dots \, j_{k+r-2}^-)_i \,. \tag{3.20}$$

The fluxes into and out of kink sites can still be evaluated using the previous eqs. 3.9 and 3.10, however they should account for specific detachment rates for the kth kink type

$$j^+ = x \, k^+, \tag{3.21}$$

and

$$j_{k,i}^- = (1 - x) k_{k,i}^- \,. \tag{3.22}$$

Using eq. 3.13 for x, the flux equations can finally be rewritten as

$$j^+ = S x_{\mathrm{eq}} \, k^+, \tag{3.23}$$

and

$$j_{k,i}^- = (1 - S x_{\mathrm{eq}}) k_{k,i}^- \,. \tag{3.24}$$

Further simplifications in the case of non-centrosymmetric units are not possible. However, to complete the calculation of the net kink rate the equilibrium mole fraction x_{eq} also needs to be calculated. It cannot be replaced by the macroscopic solubility as it varies from surface to surface. Kim *et al*. [45] have provided a way to calculate x_{eq} considering the relationship between the free energy required to remove a partially solvated growth unit from the kink site, ΔW_k, and x_{eq} in the equilibrium state. The edge specific mole fraction of solute molecules in solution can be calculated in a self-consistent way as

$$-\frac{1}{n} \sum_{k=1}^{n} \left(\frac{\Delta W_k}{k_{\mathrm{B}} T} \right)_i = \ln \left(\frac{x_{\mathrm{eq}}}{1 - x_{\mathrm{eq}}} \right)_i \,. \tag{3.25}$$

It is noteworthy to mention again that here the mole fractions of external species are ignored, and only solute molecules that can incorporate into kink sites are considered.

For the spiral dissolution mechanism, again the net kink rate for dissolution is considered as the subtraction of fluxes into kink sites from fluxes out of kink sites in eq. 3.19 and the expression for the edge velocity for spiral dissolution is

$$u_i = \frac{\prod_{k=1}^{n} j_{k,i}^- - (j^+)^n}{\sum_{r=1}^{n} (j^+)^{n-r} \, (j_i^-)^{r-1}} \,. \tag{3.26}$$

where $(j_i^-)^{r-1}$ is again calculated by eq. 3.20.

PART II
Theoretical Background

4 Basics of Molecular Dynamics Simulations

One of the principal computer simulation methods is molecular dynamics (MD). By allowing molecules to interact for a fixed period of time, they provide a view of the dynamical evaluation of molecular systems. It is not only a promising method for complementing and understanding experiments, but may provide new routes for experiments such as the prediction of structural features for proteins [66]. Since the first introduction of MD simulations by Alder and Wainwright in the late 1950's [67-68], it has been used for a large variety of systems to investigate different aspects such as structure, thermodynamics etc. As such the number of simulation techniques has expanded enormously; specialized MD methods are developed for application to particular problems.

In this part, first the basic concepts of MD simulations are introduced and then common interaction models, i.e. potentials used in MD simulations, are summarized. A brief overview of thermodynamics and statistical mechanics in the context of MD simulations is given. Particular focus is on free energy calculations and sampling methods.

4.1 Equations of Motions

For a classical system of N particles with coordinates $\boldsymbol{R} = (r_1, r_2, \dots r_N)$ and potential energy $V(\boldsymbol{R})$, the atomic momenta can be introduced as $\boldsymbol{P} = (p_1, p_2, \dots p_N)$. The kinetic energy may then be written as $T(\boldsymbol{R}) = \sum_{i=1}^{N} |p_i|^2/2m_i$. The total energy or Hamiltonian would be written as the sum of kinetic and potential terms $\mathcal{H} = T + V$. The Hamiltonian is obtained from the Lagrange function $\mathcal{L} = T - V$ via Legendre transformation by relating it to the original formulation of Lagrangian mechanics. The classical equations of motions are derived as

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial r_i} = f_i, \quad \dot{r}_i = \frac{\partial \mathcal{H}}{\partial p_i} = \frac{p_i}{m_i}. \tag{4.1}$$

where $f_i = -\nabla_i V$ denotes the forces acting on particle i. This way, Newton's second law

$$m_i \ddot{r}_i = f_i, \tag{4.2}$$

is retrieved. Eq. 4.2 can in principle be integrated, assuming the knowledge of $V(\boldsymbol{R})$ and initial conditions $(\boldsymbol{R}(t=0), \boldsymbol{R}(t=0))$. Combined with Newton's first law, the classical motion of the system at all times can be derived. There are many existing methods to perform a step-by-step numerical integration of eqs. 4.1 and 4.2. Here, only the ones considered in this work are explained briefly.

4.1.1 The Leap Frog Algorithm

This simple algorithm [69] uses positions $\boldsymbol{R}$ at time t and velocities $\boldsymbol{V}(=\dot{\boldsymbol{R}})$ at time $t-\frac{1}{2}\delta t$. It then updates positions and velocities using the forces $F(t)$ determined by the positions at time t. Its mathematical formulations are written as

$$\boldsymbol{V}\left(t+\frac{1}{2}\delta t\right)=\boldsymbol{V}\left(t-\frac{1}{2}\delta t\right)+\frac{\delta t}{m}F(t)\,, \tag{4.3}$$

and

$$\boldsymbol{R}(t+\delta t)=\boldsymbol{R}(t)+\delta t\,\boldsymbol{V}\left(t+\frac{1}{2}\delta t\right), \tag{4.4}$$

The leap-frog algorithm produces trajectories that are identical to a Verlet algorithm [70]. Eq. 4.4 can also be rewritten as

$$\boldsymbol{R}(t+\delta t)=2\boldsymbol{R}(t)-\boldsymbol{R}(t-\delta t)+\delta t^{2}\frac{F(t)}{m}+O(\delta t^{4}), \tag{4.5}$$

with a convergence of $O(\delta t^{4})$.

4.1.2 The Stochastic Leap Frog Algorithm

Stochastic or Langevin dynamics extends Newton's equation by an approximate account of the effect of the environment on a molecular system by adding a friction and a noise term

$$m_i\,\ddot{r}_i=\;f_i-\;m_i\xi_i v_i+n_i(t)\,. \tag{4.6}$$

Here, ξ_i is the friction constant and $n_i(t)$ is the random force. The exact derivation of positions and velocities are complex and depend on the random numbers generated. As such the reader is referred to work of Van Gunsteren and Berendsen [71] for further details. It is noted here that their algorithm produces third-order accuracy for any value of $\delta t\xi$.

4.2 Thermostats and Barostats

Experiments are usually conducted under defined external conditions. In some of them the temperature is controlled by keeping the system in equilibrium with a heat bath and/or the pressure of the system is controlled. In simulations, it is also possible to control these parameters and mimic the experiments. To help a simulation sample the correct ensemble (i.e. canonical, *NVT*, or isothermal-isobaric, *NPT*) thermostats and barostats are designed. Several models have been developed to control the system temperature and pressure. Here, only the thermostats and barostats used in this work are discussed briefly.

4.2.1 Berendsen Thermostat

The Berendsen algorithm [72]mimics weak coupling to an external heat bath with given temperature T_0. In this approach, the temperature of the system obeys the following equation

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}, \tag{4.7}$$

where τ is the temperature time constant and the temperature deviation decays exponentially with time constant as expected. The fluctuations of the Berendsen thermostat suppress the fluctuations of the kinetic energy. Therefore, a proper canonical ensemble is not generated and rigorously the sampling will be incorrect. For very large systems, the ensemble averages will not be affected significantly though, except for the distribution of the kinetic energy itself. However, fluctuation properties, such as the heat capacity, will be affected.

This scheme scales all velocities after integration at every n_{TC} (i.e. temperature coupling step) by a time dependent factor of

$$\lambda = \left[1 + \frac{n_{\mathrm{TC}}\,\delta t}{\tau_T}\left(\frac{T_0}{T(t+\frac{1}{2}\delta t)} - 1\right)\right]^{1/2}, \tag{4.8}$$

where τ_T is close but not exactly equal to τ. The relationship between them is defined by

$$\tau = 2C_{\mathrm{v}}\tau_T / N_{\mathrm{df}} k_{\mathrm{B}}\,, \tag{4.9}$$

with C_{v} the total heat capacity of the system, k_{B} the Boltzmann constant and N_{df} the total number of degrees of freedom. Though compared to other temperature coupling schemes the Berendsen thermostat has the disadvantage of not giving the correct ensemble, it has the advantage in terms of having a user adaptable coupling strength. It is mostly used to reach the target temperature in relaxation/equilibration simulations.

4.2.2 Velocity Rescaling Thermostat

This thermostat is essentially the same as the Berendsen thermostat except an additional stochastic term that ensures a correct kinetic energy distributiton [73]

$$dT(R) = (T_0(R) - T(R))\frac{dt}{\tau_T} + 2\sqrt{\frac{T(R)\,T_0(R)}{N_{\mathrm{f}}}}\,\frac{\mathrm{d}W}{\sqrt{\tau_T}}, \tag{4.10}$$

where T(R) is the kinetic energy, N_{f} the number of degrees of freedom and $\mathrm{d}W$ a Wiener process, i.e. a continuous-time stochastic process. No additional parameter, except for a random seed is

necessary. This thermostat is advantageous compared to the Berendsen thermostat since it produces a correct canonical ensemble. In addition, it still has the advantage of the Berendsen thermostat: first-order decay of temperature deviations and no oscillations.

4.2.3 Nosé-Hoover Thermostat

To enable exact ensemble simulations, a temperature coupling approach introduced by Nosé [74] and Hoover [75] can be employed. In this approach, the particles' equation of motions are replaced by:

$$\ddot{\boldsymbol{R}}(t) = \frac{F(t)}{m} - \xi(t)\dot{\boldsymbol{R}}(t)\,, \tag{4.11}$$

and

$$\frac{d\xi(t)}{dt} = \frac{T - T_0}{Q}. \tag{4.12}$$

In eq. 4.11 and 4.12, ξ(t) is a scaling factor, $\ddot{\boldsymbol{R}}(t) = \dot{\boldsymbol{V}}(t)$ is the acceleration, and Q is called a 'mass parameter' defined as $Q = T_0\,\tau_T^2$. The integration of eq. 4.12 is embedded into the velocity verlet scheme in which the velocities are scaled twice during one time step. The major shortcoming of this thermostat is that the actual time it takes for relaxation is large as it produces an oscillatory relaxation.

There are several other approaches for temperature control [76-77] that can be suitable for different systems under certain conditions. The thermostats summarized in this part are the ones used in this work.

4.2.4 Berendsen Barostat

The Berendsen algorithm [72]rescales the coordinates and box vectors at every n_{PC} (i.e. pressure coupling step), with a scaling matrix μ. The pressure of the system follows the equation

$$\frac{dP}{dt} = \frac{P_0 - T}{\tau_p}\,, \tag{4.13}$$

with P_0 the reference pressure. The scaling matrix elements are given by

$$\mu_{ij} = \delta_{ij} - \frac{n_{\mathrm{PC}}\,\delta t}{3\tau_p}\,\beta_{ij}\{P_0 ij - P_{ij}(t)\}. \tag{4.14}$$

β_{ij} is the isothermal compressibility of the system. When scaling completely anisotropically, the system has to be rotated to obey the box vector conditions. The actual scaling vector μ' is

$$\mu' = \begin{pmatrix} \mu_{xx} & \mu_{xy} + \mu_{yx} & \mu_{xz} + \mu_{zx} \\ 0 & \mu_{yy} & \mu_{yz} + \mu_{zy} \\ 0 & 0 & \mu_{zz} \end{pmatrix}. \quad 4.15$$

The velocities are neither scaled nor rotated. Though the relaxation time is short for the Berendsen barostat and it yields the correct average pressure, the exact *NPT* ensemble is not given.

4.2.5 Parrinello-Rahman Barostat

When the fluctuations in pressure or volume are important e.g. to calculate thermodynamic properties, especially for small systems, then the weak coupling scheme of Berendsen coupling will not be enough. In the Parrinello-Rahman approach [78], the box vectors as represented by the matrix b obey the matrix equation of motion

$$\frac{db^2}{dt^2} = V\, W^{-1} b'^{-1} (P - P_{\text{ref}}). \quad 4.16$$

The volume of the box is denoted as V, and W is a matrix parameter that determines the strength of the coupling. The matrices P and P_{ref} are the current and reference pressures, respectively.

The equations of motions for the particle are also changed. The parameter matrix is calculated by

$$(W^{-1})_{ij} = \frac{4\pi^2 \beta_{ij}}{3\tau_p^2\, L}, \quad 4.17$$

with L the largest box matrix element. Just as for the Nosé-Hoover thermostat, the Parrinello-Rahman time constant is not equivalent to the relaxation time used in the Berendsen pressure coupling algorithm. If the pressure is far from the target pressure, very large oscillations in box vectors could be observed. In that case, it is better to first start with a weak coupling thermostat to reach the target pressure and then switch to the Parrinello-Rahman algorithm.

As pressure coupling could be completely anisotropic, it is also possible to include only selected components of the pressure tensor in the coupling. Especially for solid-liquid interfaces pressure control should best be applied only along the direction of the surface normal, while the lateral dimensions of the cell should be kept constant.

4.3 Periodic Boundary Conditions

The number of particles that can be simulated in MD simulations is limited, especially by the complexity of the treatment of interactions. Hence, extended systems of atoms are often simulated

within a space-filling box, which is surrounded by translated copies of itself. This way, unwanted boundaries of a finite system are replaced by periodic boundary conditions. The interactions are now mostly short-ranged. The calculation of these interactions obeys the minimum image convention; i.e. each particle only interacts with the closest particle or its periodic image of the remaining particle system. In order to calculate long-ranged and electrostatic interactions, different strategies have been developed. It should be noted that, the application of the periodic boundary conditions concept does not automatically prevent finite-size effects, as the periodicity of the repeated images might still influence the system's behavior, in particular for small cells. Thus, it is always advised to test for the convergence of certain properties with respect to the system size so that significant artefacts would be excluded.

5 Interaction Models/Potentials

The definition of a potential function through which the particles interact with each other is the central requirement for a MD simulation. All properties of the simulated system such as forces, energies or thermodynamic properties are determined by the interaction model applied. Hence, it is extremely important to choose an appropriate potential/model. There are a variety of potentials existing in the literature at many levels of accuracy. The choice of the potential will depend on the properties and systems to be investigated, as well as the numerical efficiency desired.

Interaction potentials are commonly distinguished between *ab initio* methods that are fully quantum mechanically described i.e. based on quantum mechanics, and empirical, i.e. pre-defined force fields that are based on a classical potential definition. There are also semi-empirical potentials in which the parameters of the potential are derived from *ab initio* calculations. Another promising interaction method, especially for large systems, is a hybrid quantum mechanical and molecular mechanical approach (hybrid QM/MM) [79]. As the description of all methods is beyond the scope of this work, only *ab initio* methods and empirical potentials will be summarized with special focus on the models used in this work.

5.1 *Ab Initio* Methods

Newton's equation of motion is not exactly followed on atomic and subatomic levels; the system rather obeys the laws of quantum mechanics. The analogue of Newton's law is Schrödinger's equation for quantum systems such as atoms and molecules whether free, bound, or localized. The fundamental difference between the two equations is the representation of the particles. In Schrödinger's equation the particles, specifically the electrons, are represented in terms of wave functions, not by the positions and momenta of point-like particles anymore

$$\hat{H}\psi_i = E_i\psi_i \qquad 5.1$$

where ψ_i and E_i are the wavefunction and energy eigenvalue of state i. $\hat{H} = \hat{T} + \hat{V}$ is the Hamiltonian operator, analogous to its classical definition.

Unfortunately, there is no analytical solution for systems bigger than the hydrogen atom for eq. 5.1. Hence, a standard quantum mechanical approach is to partition the problem into simple mathematically feasible subsystems. Due to the heavy mass of the atomic cores, electrons and (heavier) nuclei move on different time scales. Therefore, their motions can often be separated in an adiabatic approach. In the well-known Born-Oppenheimer approximation [80], the resulting

electronic Hamilton operator includes the positions of nuclei merely as parameters, by treating the electrostatic interactions between nuclei and electrons as an external field $v(r)$

$$\hat{H}_{\mathrm{BO}}^{\mathrm{elec}} = -\frac{1}{2}\sum_{i}\nabla_i^2 + \frac{1}{2}\sum_{i,j\ i\neq j}\frac{1}{|r_i - r_j|} + v(r) = \hat{T} + \hat{V}_{\mathrm{ee}} + \hat{V}_{\mathrm{ext}}\,. \qquad 5.2$$

$\hat{V}_{\mathrm{ee}}$ is the electron-electron interaction energy and $\hat{V}_{\mathrm{ext}}$ is the nuclei-electron energy. The corresponding wavefunction $\psi(\{r_i\})$ depends only on the coordinates of the electrons. A usually more efficient way, especially for *ab initio* molecular dynamics is the Car-Parrinello method [81]. It is also based on the Born-Oppenheimer approximation, but in addition it includes the electronic degrees of freedom via (fictitious) dynamical variables into the classical dynamics.

5.1.1 Density Functional Theory

The wave function is the key variable in many-body theory for quantum many-particle systems. However, the calculation of it is generally complicated and computationally demanding. It is also a quantity only defined up to an arbitrary phase factor and not a physical observable. Therefore, a more accessible quantity can be considered and this quantity is the electron density, $\rho(r)$. In density functional theory (DFT) all energy terms are correspondingly rewritten based on the density

$$E[\rho] = T[\rho] + V_{\mathrm{ee}}[\rho] + V_{\mathrm{ext}}[\rho] = T[\rho] + V_{\mathrm{ee}}[\rho] + \int drv(r)\rho(r). \qquad 5.3$$

The ground-state energy of the system is determined by the minimization of the energy functional with respect to $p_0(r)$, the ground state density function. Though in principle, eq. 5.3 is exact, the electron-electron interaction potential $V_{\mathrm{ee}}\,[\rho]$ is still unknown. Kohn and Sham [82] proposed to replace the exact system by an auxiliary system of uncorrelated electrons that generate the same density as the system of interacting ones. By this approach, the calculation of the kinetic energy operator $\hat{T}$ is simplified and the electron-electron interaction is reduced to a Coulombic Hartree potential V_H. However, in order to account for the difference between the original system and the auxiliary system of independent electrons, another energetic contribution, the so-called *exchange-correlation potential* is introduced, $V_{\mathrm{xc}}[\rho]$. There are different approximations to this term; the most common ones are the local density approximation (LDA) also proposed by Kohn and Sham [82], the generalized gradient approximation (GGA) such as used in the PBE functional [83], and hybrid functionals, such as B3LYP [84]. The advantages and drawbacks for each approximation

depend on the particular case and material. In this work, mostly the PBE functional is used for the quantum calculations and the B3LYP functional is also considered for one of the cases.

Further practical approximations are common in state-of-the-art DFT codes to increase thecomputational efficiency. One fundamental approximation is the expansion of the wave functions in an appropriate basis so that mathematical operations are performed easier. The two most popular types of basis sets are (simplified) atomic orbitals and plane waves. While atomic orbital basis sets are well suited for isolated atoms or molecules, for crystalline solids use of plane wave basis sets may be advantageous due to their periodicity. The precision of this expansion can be controlled by the number of basis functions included. Another approximation that is particularly common in DFT with plane wave basis sets is the so-called *frozen core approximation* via the introduction of pseudopotentials [85]. There are some extensions to DFT to increase the accuracy of the calculations depending on specific systems. One such extension considered in this work is the description of van-der-Waals interactions for which different approximations are already proposed but only one of them is considered in this work (chapter 8.1).

Though many other solutions for the Schrödinger equation exist [86-91], only one of the common ones, DFT, is introduced here. The reader is kindly referred to the works cited for a discussion of the other methods.

5.2 Empirical Potentials

The system sizes that can be simulated using *ab initio* methods with feasible computational costs are rather small. Even without considering the dynamics, i.e. only energy calculations of systems with more than one thousand atoms would generally be challenging to compute. With dynamic simulations using the Car-Parrinello approach sizes are even restricted to a few hundred atoms to achieve reasonable simulation times of 10-20 ps. Hence, the usage of *ab initio* molecular dynamics is limited for systems of the type of interest to this work. Fortunately, for certain systems as the ones of interest here quantum effects may be neglected with less severe consequences. The well-defined potential energy landscape (PES) for many molecules and solids can be produced by pre-defined interaction potentials as long as no chemical reaction is involved. Such interaction models are commonly referred as force fields and can be defined either on a semi-empirical basis by derivation of the potential energy from quantum mechanical equations or by mapping the potential energy to appropriate functions fully empirically. The parameters for these functions should be adjusted by fitting to suitable target values such as DFT energies, experimental observables or

atomic and molecular structures. These reference values are chosen according to properties one is ultimately interested in investigating. The parameterized force field should finally be tested carefully by comparing to experiments if available and/or to *ab initio* calculations to check if target quantities can be approximately reproduced.

The basic functional form of a force field potential consists of bonded terms associated with chemical bonds, bond angles, and bond dihedral interactions of atoms that are linked by covalent bonds. In addition it contains non-bonded or noncovalent terms that describe the long-range electrostatic and van der Waals forces. The total potential energy is thus defined as

$$V_{\text{total}} = V_{\text{bond}} + V_{\text{angle}} + V_{\text{dihedrals}} + V_{\text{electrostatic}} + V_{\text{van der Waals}} . \quad 5.4$$

A very large number of force fields suitable for different kinds of systems, investigations etc. exist in the literature. There are polarizable force fields [92] to treat polarization and charge delocalization, coarse-grained force fields [93-94] to increase simulation times for large systems, and reactive force fields that allow to simulate some specific chemical reactions such as hydrocarbon reactions [95] or water proton transfer [96]. In this work, only classical force fields specially developed for biological systems, i.e. biological force field methods, will be considered.

5.2.1 Biological Force Fields

The use of force fields for the simulation of biological system is a well-established tool, specifically for structure-activity relationships [97]. Hence, a variety of force fields for these systems are developed and the most common ones are AMBER [98-99], CHARM [100] and OPLS-AA [101]. In the majority of this work a generalized form of AMBER force field is used. CHARMM and OPLS-AA force fields are considered for comparisons. Here, the functional form of the AMBER force field is introduced. The functional forms for other force fields are the same or very similar. They typically only use different parametrization values.

The functional form of the AMBER force field is given by

$$V = \sum_{bonds} k_{\mathrm{b}}(r_{ij} - r_0)^2 + \sum_{angles} k_{\theta}(\theta_{ij} - \theta_0)^2 + \sum_{dihedrals} \frac{k_{\Phi}}{2} \left[1 + \cos\left(n\phi_{ijkl} - \gamma\right)\right] + \sum_{pairs} 1/\varepsilon_{ij} \left[\left(\frac{A_{ij}}{r_{ij}^{12}}\right) - \left(\frac{B_{ij}}{r_{ij}^{6}}\right)\right] + \sum_{pairs} \frac{q_i q_j}{r_{ij}}. \quad 5.5$$

Bond stretching is modeled by a harmonic potential with spring constant k_{b} and equilibrium bond length r_0. Angle bending terms account for the stiffness of two connected bonds by including an

angle force constant k_θ and equilibrium angle θ_0. Dihedral angles describe the torsion of four atoms around the central bond by the dihedral force constant k_ϕ, the periodicity n and the phase angle γ. A, B and q are coefficients for the non-bonded potentials. The electrostatic interactions are described by Columb-interactions between partial charges on the particles (last term in eq. 5.5); the van der Waals (vdW) and all remaining interactions are described by Lennard-Jones potentials (fourth term in eq. 5.5). The partial charges are usually determined from *ab initio* calculations by the restraint electrostatic potential (RESP) method [102], a restraint ESP approach. Using this approach, the charges are determined by restrained fitting of a set of point charges located at the positions of the nuclei to reproduce the potential outside the van der Waals radii of the atoms.

The treatment of water molecules is necessary when they are present as solvents or even as crystal waters (such as crystal water in the lattice of alpha-lactose monohydrate). To define a reasonable model for liquid water, an intermolecular potential function for water dimers should be accurately described. Common water models are the simple point charge (SPC) model, the TIP3P or the TIP4P models [103]. The AMBER as well as CHARMM force fields mostly employ TIP3P with constrained bonds and angles. TIP4P model is advised to be used with the OPLS-AA force field.

6 Free Energy Calculations and Accelerated Molecular Dynamics Methods

6.1 Free Energy Calculations

Quantities commonly measured in experiments are macroscopic observables such as entropy, internal energy etc., i.e. thermodynamic functions. As simulations are performed for a limited collection of possible microstates, it would not be proper to predict these observables directly. Hence, to bridge the gap in between, statistical mechanics are exploited, depending on the ensemble chosen.

In statistical mechanics, all possible microstates spanning the phase space $\Gamma\{\boldsymbol{R};\boldsymbol{P}\}$ in a canonical ensemble can be described as

$$Q = \int\int \exp\big(-\beta\, \mathcal{H}(\boldsymbol{R},\boldsymbol{P})\big)\, d\boldsymbol{R}\, d\boldsymbol{P}\,, \tag{6.1}$$

where $\mathcal{H} = \mathcal{T} + \mathcal{V}$ is again the Hamiltonian of the system and $\beta = 1/k_{\mathrm{B}}T$ is the inverse of thermal energy with k_{B} as Boltzmann constant and T the temperature. Q is called *partition function* or *the sum over states.* As shown in eq. 6.1, it is obtained by integrating over the phase space for systems with continuous energies. The probability of finding a microstate i with corresponding Hamiltonian $\mathcal{H}(\boldsymbol{R}_i,\boldsymbol{P}_i)$ is given by:

$$P(\boldsymbol{R}_i,\boldsymbol{P}_i) = \frac{1}{Q}\exp(\mathcal{H}(-\beta\boldsymbol{R}_i,\boldsymbol{P}_i))\,. \tag{6.2}$$

The associated thermodynamic potential for the canonical ensemble is the Helmholtz free energy $G = U - TS$ with internal energy U and the entropy of the system S. (The Gibbs free energy function on the other hand contains the enthalpy instead of the internal energy, as it is associated with an *NPT* ensemble). Relating thermodynamics to statistical mechanics, the free energy can be expressed in terms of the partition function

$$G = k_B T \ln Q\,. \tag{6.3}$$

The direct evaluation of eq. 6.3 using molecular dynamics is challenging as high-energy states would be difficult to reach, hence a correct partition function cannot be evaluated. Fortunately, absolute free energies of the system are rarely necessary; rather free energy differences between different subsystems are desired to be evaluated. This difference can be obtained by defining the partition functions for the corresponding subsystems. For instance, subsystems A and B with partition functions Q_A and Q_B, respectively, would have a free energy difference

$$\Delta G = k_B T \ln \frac{Q_A}{Q_B}. \quad 6.4$$

As the integrals over phase space are not straightforward to calculate, instead a continuous reaction coordinate λ can be considered and the free energy is calculated as

$$G(\lambda) = -k_B T \ln P(\lambda) + G_0 , \quad 6.5$$

with G_0 an arbitrary constant, and $P(\lambda) = Q(\lambda)/Q$ the probability distribution. The partition function $Q(\lambda)$ is described as

$$Q(\lambda) = \int \exp\big(-\beta\, \mathcal{H}(\boldsymbol{R}, \boldsymbol{P})\big)\, \delta\big(\lambda - \lambda(R)\big) d\boldsymbol{R}\, d\boldsymbol{P}. \quad 6.6$$

While considering free energy differences instead of absolute free energies simplifies the numerical treatment of the simulation, still a suitable method should be devised for practical applications of free energy calculations.

The evaluation of eq. 6.5, in principle, is straightforward and by running a long enough equilibrium simulation, the probability distribution $P(\lambda)$ can be obtained by counting the occurrences associated with each value of λ. However, in practice, this would not be the approach chosen especially when free energy landscapes are not flat but structured i.e. rather high barriers between states are present. Then , for most or all of the simulation time (depending on the times achievable by simulations), the reaction coordinate would be trapped in a local free energy minimum and the global free energy profile would not be sampled efficiently. This is actually one of the main challenging issues of present-day MD simulations. There are several strategies to overcome this problem for free energy calculations such as thermodynamic perturbation [104], thermodynamic integration (TI) [105], the Bennett Acceptance Ratio (BAR) [106], or biased calculations [107]. In this work the BAR method will be discussed briefly as it is the main method used in this work.

6.1.1 The Bennett Acceptance Ratio

As one of the earliest free energy methods, the BAR method extracts data from multiple states to estimate the free energy differences. In this approach, *configuration information from both states* is required for estimating free energy differences rather than the configuration of just a single state as in the case of the TI method.

In BAR and also other alchemical free energy calculations such as TI, the free energy difference calculations are broken down into multiple smaller sub-steps. This approach improves the similarity, i.e. the phase space overlap of the two end states involved, and yields a better

convergence of the free energy results. To do so, intermediate states between the two end states i and j have to be simulated. The intermediate states are created by mixing the energy parameters of the states i and j. A so-called coupling parameter λ that ranges between 0 and 1 is used as mixture ratio. The free energy difference of interest is then calculated by summing free energy differences between the intermediate states

$$\Delta G_{i,j} = \sum_{l=1}^{n_{\lambda-1}} \Delta G_{l,l+1} \,, \tag{6.7}$$

with n_λ being the number of intermediate states. These free energy differences between intermediates are calculated as

$$\Delta G_{l,l+1} = -k_B T \ln \frac{n_{\lambda_{l+1}}}{n_{\lambda_l}} + C \,, \tag{6.8}$$

where C is a constant that can be found based on an iterative process [106].

During free energy calculations where particles grow out of nothing or particles disappear, using the simple linear interpolation of the Lennard-Jones and Coulomb potentials are observed to lead to poor convergence. To get around the numerical instabilities so-called soft core potentials [108] for intermediate λ values are used.

6.2 Accelerated Molecular Dynamics

As already mentioned in chapters 5.2 and the previous chapter 6.1, MD simulations especially for large systems require massive amounts of computational power for simulation times of even 10-20 ps. The processes one desires to investigate such as protein folding, detachment from surfaces etc. often take much larger times, around microseconds to seconds. The common feature of these slower processes (compared to reachable MD simulation time scales) is that they are infrequent or rare events. Many different methodologies are developed for these rare events systems and they are all categorized under accelerated MD in this work. It should be noted that not all of them are designed to overcome sampling issue of MD unlike metadynamics, but to boost up the system so that barriers can be crossed. By this way, infrequent events can be observed and higher simulation time scales can be achieved. In this work, metadynamics [109], parallel replica dynamics (PRD) [110] and hyperdynamics [111]methods are discussed briefly.

6.2.1 Metadynamics

As an enhanced sampling method, in metadynamics the sampling of a system is improved by the introduction of an additional bias potential (or force). This potential acts on a selected number of degrees of freedom, often referred to as collective variables (CVs). To prevent the system from revisiting previous locations and so push it out of a free energy basin, the bias potential is constructed in a way that it repels the reaction coordinate from already visited locations, i.e. it is history-dependent. The exact explanation and the derivation can be found in the works of Laio and Parrinello in ref. [112] and [113]. In practice, during metadynamics Gaussian shaped hills centered at the CV are added to the potential of the system. The definition of the bias potential is given as

$$V_{\text{bias}}(\vec{s},t) = \sum_{k\tau<t} W(k\tau) \exp\left(-\sum_{i=1}^{d} \frac{(s_i - s_i^{(0)}(k\tau))^2}{2\sigma_i^2}\right). \qquad 6.9$$

W is the height of the Gaussian hills, τ is the time step between each k steps, s^0 specifies the position of the centroid of the Gaussian with s being the positions during simulations, σ is the width of the Gaussian hills. However, this original approach mostly suffers from convergence issues. Therefore, well-tempered metadynamics [114] based on deposition of adaptive bias potentials in terms of hill heights along the CV is preferred to extract exact free energy profiles.

6.2.2 Parallel Replica Dynamics

Introduced by Voter [110], the concept behind PRD simulations is trivial. Using the power of parallel computers, many replicas of the system can be followed to parallelize time evolution. Based on the number of processors available, the system is replicated on each of these processors. Then, these replicas should be uncorrelated, which is usually achieved by assigning different initial velocities. The trajectories obtained by simulations (after the dephasing) are checked for the desired transition event and as soon as such a transition observed, the simulation should be stopped. The trajectory times over all processors are added and simulation clock is advanced by this sum. Then, if necessary, one can continue the trajectory in which the transition is observed and replicate it to start new parallel simulations. This approach gives exact dynamics. However, the boost is limited by the number of processes. The dephasing times should be small enough to gain good efficiencies by this method.

6.2.3 Hyperdynamics

The concept of hyperdynamics [111] is in some respect similar to metadynamics. However, instead of filling the basin with Gaussian hills, the simulations are started with a suitable pre-defined bias potential. By this way, the chance of the system leaving the basin is increased. An illustration of the hyperdynamics method is depicted in Figure 6.1 and compared to a classical MD simulation without any acceleration method. The design of the bias potential is the most critical point in this approach. It should be zero at all dividing surfaces so as not to bias rates along different pathways. Considering that in principle one does not have any *a priori* information about the system such as neighboring states or dividing surfaces, it is rather difficult to design a bias potential that is suitable for the system as well as computationally efficient. A few different forms have already been proposed and tested, see ref. [115]. In this work, the proposed bias potential design would be different from the methods described in the literature (see chapter 7.2). It will actually involve the metadynamics technique.

The advantage of the hyperdynamics technique compared to PRD is that the boost factor is not dependent on the number of processors involved. This actually decreases the computational cost and allows the acceleration of the system without performing too many simulations. Especially for systems with really high barrier energies for which the chances of observing a transition event may be very low. Then again, the PRD method gives the exact dynamics of the system, hence in a sense it is more accurate than hyperdynamics or any other acceleration method. Hence, if possible, the hyperdynamics simulations and most importantly the bias potential design method should be checked against PRD simulations for suitable model systems, maybe with lower barriers.

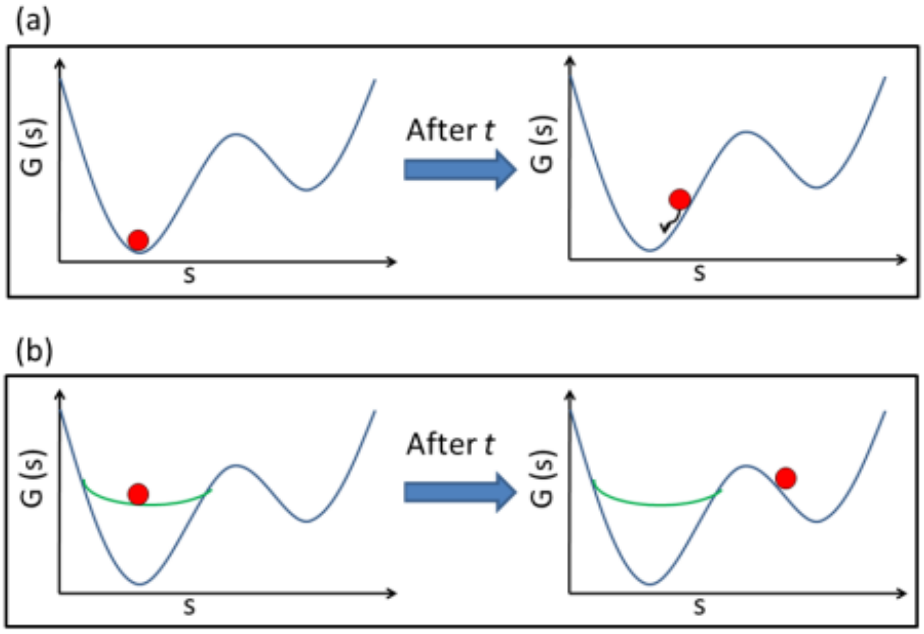

Figure 6.1: (a) Illustration of the free energy diagram for a classical molecular dynamics simulation along a simple one dimensional reaction coordinate *s*. (b) Illustration of the free energy diagram for a hyperdynamics simulation performed along the same reaction coordinate as (a). The green line depicts the added bias potential and the red circle denotes the system state at a time step *t*.

PART III
Development and Application of the Spiral Dissolution Model

The spiral growth model originally proposed by Burton, Crabrera and Frank has been extended and modified during the last six decades. In particular, the BCF model was advanced into a molecularly-resolved mechanistic crystal growth model by *Doherty et al.* for the prediction of crystal morphologies. Yet, they did not exploit their model for an absolute growth/dissolution rate prediction of real crystal structures, as they used crude methods to approximate the parameters entering the model (see chapter 3). There is no conceptual barrier to use more accurate methods to evaluate the parameters though and then employ the model for absolute growth/dissolution rate predictions. As the importance of dissolution testing during drug development is already well-established in the introduction part of this work, the focus of this work is in the dissolution process. The differences in the expressions to evaluate the parameters in the context of the dissolution process instead of growth were also already explained in chapter 3. Hence, this work aims to develop a computational protocol that would be able to predict dissolution rates for pharmaceutical ingredients by exploiting the spiral model and extensions of the model. In this part, the procedures to calculate the model parameters are explained. After that, the spiral dissolution model is applied to two prototype compounds, aspirin and alpha-lactose monohydrate (alpha-LM). This thesis will strongly build on the works in refs. [43, 116-117]. Part of this work, the fast screening approaches for predicting the dissolution rates for both compounds in chapter 8 and 9, is submitted for publication. Another publication with the detailed alpha-LM results is in preparation.

7 Micro-Kinetic Spiral Dissolution Model

The spiral growth developed by Synder and Doherty [32] and further extended by Kuvadia and Doherty [44] was explained in chapter 3 in detail together with an emphasis on the changes necessary to treat the dissolution process. Apart from crystallographic parameters, critical lengths and step velocities are required to build the model. Particularly, the two parameters ΔG_{kink}, the free energies necessary to create kink sites along edges, and k^{-}, the detachment rate constants from kink sites, need to be evaluated for centrosymmetric units. In this work, the model built will always consider the step velocities using eq. 3.15 even though the molecules considered themselves may be non-centrosymmetric. The workaround is to consider centrosymmetric dissolution units instead. As already mentioned the parameters have hitherto only been estimated using crude methods like bond breaking energies. However, the aim of this work is to predict the absolute dissolution rates with sufficient accuracy using molecular dynamics techniques. In the next parts, the evaluation of kink free energies of formations and detachment rate constants from kink sites will be explained.

7.1 Kink Free Energy Calculations

To evaluate kink free energies, previous methods used to calculate binding free energies mostly for ligand/receptor complexes are modified [43, 118-120]. In these methods, the overall process of free energy difference between initial and final states is decomposed into several stages and then, each state is computed separately. This so-called ‘thermodynamic cycle’ involves the slow decoupling of the non-bonded interactions between the ligand and its surrounding. To preserve the reversibility of the thermodynamics cycle, additional sets of potentials are employed to restrain the orientation and position of the ligand. The cost of these restraint potentials also has to be taken into account. Instead of a ligand molecule, kink molecules along step edges are considered here.

In Figure 7.1, the application of this method is illustrated with a top view in which the upper terrace part of the surface is represented as a box. The straight edge, the initial state (on top left), and the edge with kink sites, the final state (on top right), are represented as rectangular dark blue boxes. Since the initial state is a straight edge and the final state is an edge with kink sites, the free energy contribution for the initial state is denoted with a ‘step’ subscript and for the final state one, it is denoted with a ‘kink’ subscript. Decoupling and restraining energy contributions are denoted by ‘dec’ and ‘rest’ superscripts, respectively. The thermodynamic circle starts by identifying the

kink molecules (represented as light blue boxes in Figure 7.1) and restraining them in their positions. Here, the positions of the heavy atoms as well as any hydrogen atoms of hydroxyl groups moieties are restraint using absolute position restraints via harmonic potentials. Then, the non-bonded interactions between selected molecules and their surrounding are decoupled by using a parameter, λ. The decoupling free energy change, $\Delta G_{\mathrm{step}}^{\mathrm{dec}}$ is obtained. As the position restraints are fully employed during this process, the additional energy necessary to restrain the kink molecules in the perfect crystal positions also has to be evaluated. It is denoted as $\Delta G_{\mathrm{step}}^{\mathrm{rest}}$ in Figure 7.1. After the decoupling process, the kink molecules can be considered in vacuum since there is no interaction with their surroundings. Therefore, it is possible to transfer or rotate them without additional cost of energy. In the next step, these molecules are thus shifted and rearranged into the final kink positions. Then, the interactions between the kink molecules in their new arrangement and surrounding are switched on, associated with a free energy change of -$\Delta G_{\mathrm{kink}}^{\mathrm{dec}}$. As last step, the position restraints on the kink molecules are removed at a cost of -$\Delta G_{\mathrm{kink}}^{\mathrm{rest}}$. The overall kink free energy of formation along an edge, ΔG_{kink} is therefore calculated as:

$$\Delta G_{\mathrm{step}}^{\mathrm{rest}} + \Delta G_{\mathrm{step}}^{\mathrm{dec}} - \Delta G_{\mathrm{kink}}^{\mathrm{rest}} - \Delta G_{\mathrm{kink}}^{\mathrm{dec}}/4\,. \qquad 7.1$$

The sum of contributions is divided by four as four kink sites are created in Figure 7.1, but only the cost of creating one kink site along an edge is required.

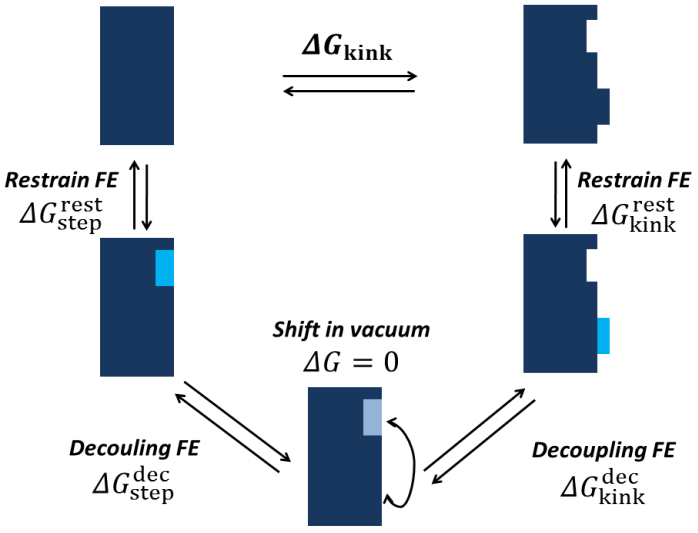

Figure 7.1: Thermodynamic cycle to calculate the kink formation free energy. ‘dec’ and ‘rest’ superscripts refer to decoupling and restraint energy, respectively. See text for explanation of free energy contributions.

In the thermodynamic cycle, there are two different free energy contributions, decoupling and restraint free energy differences. The calculations of these two energy contributions are done by different free energy calculation methods (chapter 6.1). For the decoupling free energy calculations, the BAR method [106] explained in chapter 6.1.1 is employed. To remind again, in this method, a coupling/switching parameter λ is employed so that the Hamiltonian of the initial system slowly changes to the Hamiltonian of the final system. As suggested for complex systems [119, 121], the switching parameter is separated into an electrostatic and Lennard-Jones part, $\lambda = \lambda_{\mathrm{el}}, \lambda_{\mathrm{LJ}}$. First, the electrostatic contributions are decoupled by applying a linear coupling parameter λ_{el} to the electrostatic contribution of the potential energy

$$U^{\mathrm{el}}(r_{ij}\lambda_{\mathrm{el}}) = (1-\lambda_{\mathrm{el}})\frac{q_i q_j}{r_{ij}}. \qquad 7.2$$

q_i (q_j) denotes the partial charge of atom i (j), the former being part of the decoupling set, while the latter belongs to the remaining part of the system.

Subsequently, the Lennard-Jones interactions are decoupled through the coupling parameter λ_{LJ} with the electrostatic interactions remaining in the $\lambda_{\mathrm{el}} = 1$ state. During this process, soft-core potentials are applied to ensure that particles are not sitting on top of each other [108] (chapter 6.1)

$$U^{\mathrm{LJ}}(r_{ij}\lambda_{\mathrm{LJ}}) = (1-\lambda_{\mathrm{LJ}})U_{ij}^{LJ}(\alpha\sigma_{ij}^{6}\lambda_{\mathrm{LJ}} + r)\,. \qquad 7.3$$

Here, σ_{ij} denotes the standard Lennard-Jones parameter, α is a soft-core parameter, U_{ij}^{LJ} represents the original Lennard-Jones potential between particles *j* and *i*.

The treatment of λ points may be asymmetric, i.e. different number of intermediate states may be used for Lennard-Jones and electrostatic decoupling, $n_{\lambda_{\mathrm{LJ}}} \neq n_{\lambda_{\mathrm{el}}}$. In fact in most cases, higher numbers of λ_{LJ} points are used for Lennard-Jones decoupling as the accuracy at large λ_{LJ} is affected by the onset of penetration of solvent into the decoupled region. This procedure is applied for the calculation of both decoupling free energy contributions, i.e. $\Delta G_{\mathrm{step}}^{\mathrm{dec}}$ and $\Delta G_{\mathrm{kink}}^{\mathrm{dec}}$.

The free energy contributions due to position restraints are obtained by the thermodynamic integration method [104] i.e. by integrating the ensemble-averaged derivatives of potential energy over the spring constant k. Then, the restraining free energy is calculated according to

$$\Delta G^{\mathrm{rest}} = \int_0^{k_{\max}} \frac{1}{2}\,dk\ \langle(\vec{R} - \vec{R}_0)^2\rangle_k\,, \qquad 7.4$$

where $\vec{R}$ and $\vec{R}_0$ are the position vectors of the restraint atoms during the simulations and of the perfect crystal, i.e. reference positions, respectively. It is also possible to estimate the difference between positions using the root mean-square displacement, ξ, to the reference positions and in

practice these deviations between perfect kink site positions and positions encountered during simulations are used.

7.2 Detachment Rate Constant Calculations

In the studied systems spontaneous detachment processes for single trajectories from kink sites are not observed for several nanoseconds (see below). It is thus obvious that detachment is a rare event and accelerated MD methods (chapter 6.2) are better suited for the evaluation of detachment rates. The method chosen in this work is hyperdynamics [111], however it is also combined with metadynamics methods [109].

The design of a bias potential is a prerequisite for the hyperdynamics method. It must be appropriate for the process considered and determines the numerical efficiency. Previous bias potentials were governed by atomic bonds or dihedral torsions. Applying such bias to non-bonded potential energies particularly in an explicit solvent environment is prevented by the large fluctuations affecting these energies. Hence, the metadynamics technique is used to design a custom-tailored bias potential. The proposed algorithm is complementing the on-the-fly combination of both techniques recently suggested by Tiwary and Parrinello to simulate back-and-forth transitions between bound states [122]. The approach used for detachment rate constant calculations is a two-step step algorithm:

1. Using the metadynamics technique a suitable bias potential acting on previously selected collective variables (CVs) is adaptively constructed. By the aid of this bias potential, the molecule is pushed from its initial state *A* to the new state *B*. The time-dependent evolution of the bias potential is truncated before the first escape event out of state A, to fulfill the requirement of zero bias at the barrier region [111].
2. Using the resulting static bias potential, the hyperdynamics simulation step then generates an ensemble of biased trajectories starting from different conditions in state *A*.

The detachment rate then can be determined using the following equation:

$$k^{-} = \kappa \frac{k_{\mathrm{bias}}^{\mathrm{TST}}}{\langle e^{-\beta V_{\mathrm{bias}}} \rangle_{\mathrm{A,bias}}} = \kappa \frac{k_{\mathrm{bias}}^{\mathrm{TST}}}{f_{\mathrm{acc}}} \qquad 7.5$$

where κ is the transmission coefficient, $k_{\mathrm{bias}}^{\mathrm{TST}}$ is the biased rate constant, V_{bias} is the bias potential and $\beta = 1/k_{\mathrm{B}}T$ with k_{B} being the Boltzmann constant and T the temperature. $k_{\mathrm{bias}}^{\mathrm{TST}}$ is calculated from an ensemble of biased simulation trajectories, where a bias potential V_{bias} suitably accelerates

the escape from the free energy basin *A* [117]. f_{acc} is the so-called acceleration or boost factor. In contrast to the on-the-fly combination of both steps, as suggested in ref. [122], the proposed approach does not rely on the same CVs accelerating both the forward and reverse transition in order to observe a large number of transitions. Instead the focus is only on the forward reaction, the detachment from the kink sites. For the present bound-to-free transition and the associated largely changing entropic contributions, this aspect of the approach becomes particularly beneficial [117]. Moreover, the criterion of zero-bias potential at the barrier location is controlled better by the separation of the two steps. To account for barrier re-crossing effects, the trajectory segments that follow the first barrier crossing can be exploited to explicitly calculate a transmission coefficient κ. The proposed method was tested against parallel replica dynamics which give the exact dynamics (see chapter 6.2.2) and excellent agreement was observed [117].

In the following chapters, the application of the computational protocol as well as the spiral dissolution model is explained, starting with the development and test of force fields for both prototypical compounds. Then, the computed kink free energy values and detachment rate constants will be discussed. The last part for each compound will be the building of the spiral model and a comparison for building the model using different detachment rate constants will be given. Finally, the theoretically calculated dissolution rates will be compared to available experimental values.

8 Calculation of the Dissolution Rate of Aspirin

Aspirin, also known as acetylsalicylic acid (ASA), is an active pharmaceutical compound often used to treat pain, fever, inflammation or to prevent blood clotting. Hence, it is a pharmaceutically relevant compound to use as a model system. It is also an experimentally well-known substance, though it exhibits higher degrees of freedom compared to previously studied systems. Nevertheless, it is an excellent prototype compound to consider. After the development and testing of the force fields for aspirin, the chosen force field is used for calculating the parameters to build the model.

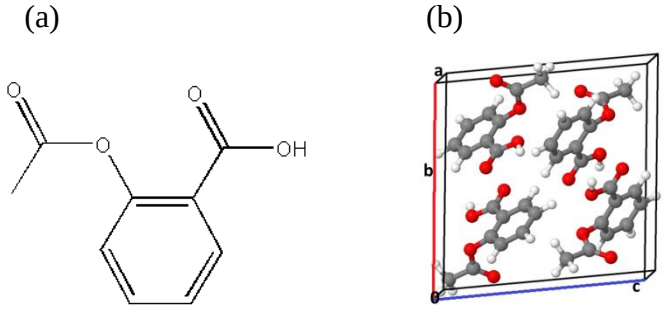

Figure 8.1: (a) The molecular structure of aspirin. (b) The unit cell structure of aspirin taken from ref. [123].

8.1 Development and Testing of the Force Fields

Different force fields, namely generalized AMBER (GAFF), CHARMM, and OPLS-AA force fields were evaluated for their suitability in simulating the dissolution dynamics [116]. This part of the work was performed mostly by Maximilian Greiner, Ekaterina Elts, and Julian Schneider and was published in ref. [116]. In the remaining part of this chapter the force field built by them is used. Hence, to validate this choice of force field, the development and testing of the force field results are displayed and discussed briefly in this part, too.

The generalized AMBER (GAFF) [99] parameters for the force field are obtained by ACPYPE [124], a program based on ANTECHAMBER [125] and partial atomic charges are calculated using the R.E.D. server [126]. The CHARMM force field parameters are generated using the SWISSPARAM server [127]. For the generation of the OPLS-AA [128] (all atomic version) force field, a commercial program Maestro from the SCHRÖDINGER software package [18] is used.The TIP3P [129] water model is used for both solvent water molecules and hydrate water molecules in the unit cell. The unit cell representations of the model substance aspirin in its I form

[123] is taken from literature. The classical force field MD simulations are performed using the GROMACS package [130-132]. First, the ability of the force field to correctly reproduce the lattice parameters is checked. For this reason, an optimization of the bulk crystal structures and the corresponding lattice parameters is performed. The starting point is the experimental crystal structure, and then the system is slowly cooled from 100 to 0 K in a 50 ps simulation while applying a Berendsen barostat at zero pressure. Specifically, the barostat is coupled to the lengths of the cell vectors as well as to the β-angle, whereas the remaining two angles were fixed to preserve the monoclinic crystal system. After the optimization of the cell vectors, all atomic positions were again relaxed using the conjugate gradient method with a maximum force threshold of 0.01eV/Å. Apart from fixing the α- and β-angles during the MD simulations, no additional symmetry constraints were imposed on the structure for the optimization procedure. Table 8.1 displays the results for the different force fields. For aspirin, the calculated lattice parameters and interaction energies were in good agreement with the experimental reference for the GAFF force field; the angle β is also better represented by GAFF as compared to the experimental structure.

Table 8.1: Comparison of experimental and calculated lattice parameters for aspirin.

	Lattice Constants (Å)			Cell angle (degrees)
	A	B	c	β
Exp.[a]	11.40	6.60	11.50	95.6
GAFF	11.47	6.42	11.24	96.3
CHARM	11.59	6.40	11.22	91.9
OPLS-AA	13.47	5.52	11.18	86.3

[a]Ref. [123]

To further test the force fields, the intermolecular interaction energies between molecules are calculated with the force fields and the resulting energies are compared to quantum mechanical calculations (QM). The interaction energies are calculated by displacing one of the two molecules rigidly in the direction connecting their centers-of-mass in the starting structure and calculating the energy relative to the energy of the two isolated molecules [116]. The QM calculations are performed with the all-electron full potential code FHI-aims [133-134]. Electronic exchange and correlation was treated at the GGA level using the PBE [83] and the description of van der Waals (vdW) interactions are corrected with the Tkatchenko-Scheffler dispersion correction approach [135]. All sampling calculations are done with the "tier2" basis set with the numerical integrations

at "tight" settings. From the crystal configuration of aspirin, representative dimer structures are chosen. Additional QM energetic calculations employed wave-function based second-order Møller–Plesset perturbation theory (MP2) [136]. The interaction energies versus the relative distance to the starting structure are displayed in Figure 8.2b, with the aspirin dimer interactions chosen for the displacement shown in Figure 8.2a. As can be seen from the figure, CHARMM shows the best quantitative agreement with both DFT and MP2 calculations. Whereas GAFF also achieves a rather good agreement, OPLS-AA underestimates the binding energy. Considering that GAFF produces best lattice parameters compared to the experimental values and gives a good quantitative agreement also for the interaction energies as compared to QM methods, the choice of the force field for this work was GAFF for aspirin.

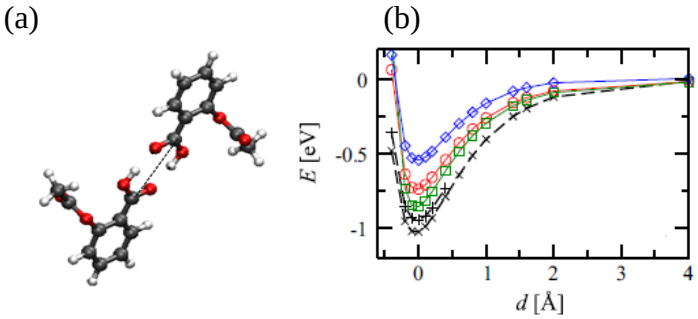

Figure 8.2: (a) The dimer structure of aspirin with dashed lines representing the displacement direction (b) Interaction energies versus displacement *d* with respect to the dimer geometry for aspirin. Reference values calculated from DFT+vdW (black dashed lines, X), MP2 (black dashed lines, +) and resulting energies from classical force fields GAFF (red circles), CHARMM (green squares), and OPLS (blue diamonds). Taken with permission from [116].

8.2 Spiral Dissolution Model for the Aspirin(001) Faces

The most common and the here chosen crystal structure for aspirin belongs to the $P2_1/c$ space group [137]. The unit cell contains four non-centrosymmetric molecules, two pairs of dimers. The face considered in this work is the (001) face, since it is one of the dominating facets of aspirin crystals grown from solution [137]. Though along the surface normal the unit cell is composed of two molecular layers, each layer exposes the same terminations. Considering the strong hydrogen bond interactions between dimers within a single layer, growth or dissolution is expected to proceed in single layers. Therefore, the step height is considered as the height of a single molecule and this also reduces the complexity of our model. The step edge directions are chosen by considering the strong intermolecular interactions, i.e. by considering the periodic bond chain (PBC) theory of Hartman and Perdok [138]. Based on PBC theory, there are three strong bond chains on this face: [100], [010] and [110] that each expose straight, close-packed interactions.

The step edge directions chosen then would be [100], [010] and $[1\bar{1}0]$ and their symmetry related counterparts, i.e. opposite step edge directions. As the kink sites would be the primary dissolution sites and the spiral dissolution model depends on the free energy of kink formation and detachment rate constants of kink site molecules, the kink sites along these edges need to be created.

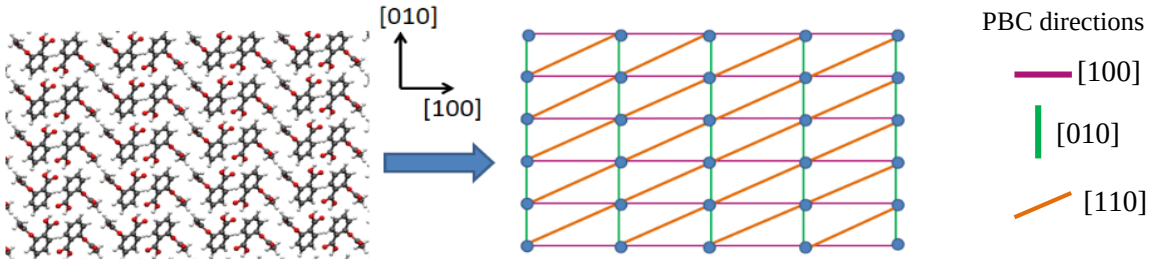

Figure 8.3: Top view and bonding structure of the (001) face of aspirin. On the right side, the blue filled circles represent the center of mass of the aspirin dimers and colored lines represent the bond chains along the shown directions. Aspirin dimers are considered for drawing PBCs.

Aspirin itself is a non-centrosymmetric molecule meaning that there could be more than one kink type along edges. Considering all these kink types would lead to a complete model, but would not be computationally efficient. Instead a simplified and more efficient model can be achieved by considering aspirin dimers as dissolution units. These dimers may be hydrogen bonded as seen in Figure 8.2a or have weaker electrostatic or Lennard-Jones interactions. The criterion is that after the detachment of the first molecule, the counter-part will detach immediately at least on MD simulation time scales. In other words, the kink structure should be meta-stable on the scale of free energy calculations (chapter 8.2.1) To further reduce the model, the following criteria are considered: Detachment/attachment of the kink site molecule always results in the same type of kink site molecule and the edge terminations before and after re-arrangement must be the same [43]. For instance, symmetry-related opposite edges (like [100] and $[\bar{1}00]$) expose different terminations. In order to ensure a precise and well-defined calculation of defect formation free energies, kink sites are therefore always created along the same edge.

It is noted that aspirin even though weak is still an acid, and therefore engages in proton transfer reactions, i.e. the deprotonation of the carboxylic acid moiety [139]. However, it is beyond the capabilities of the common classical force fields to simulate these reactions and consequently such effects are disregarded in the simulations in this work. A large fraction of dissolved aspirin molecules would become de-protonated in aqueous solution at neutral pH due to the aspirin acidity. Strictly speaking, the simulation model thus corresponds to an acidic environment. Different

solution pH values may alter the interfacial behavior to some extent. Yet, experimental results indicate that the differences in solubility and dissolution rate between neutral and acidic solution are small [140]. Thus, it is assumed that the proton transfer reaction takes place primarily in bulk solution when the carboxylic acid groups are well accessible to water molecules and that the detachment process of initially neutral crystal molecules itself does not involve deprotonation to a significant amount. Moreover, for the same reason of force field limitations, any possible dissolution mechanism proceeding via hydrolysis is not taken into account [43].

8.2.1 Kink Free Energies of Formation

This part of the work was performed by Chen Zheng as a part of his master thesis [141] and was published in [43] by Schneider, Zheng and Reuter. As the spiral dissolution model builds on the kink densities and kink free energies of formations, the details of the simulations and results for aspirin are also summarized briefly in this work. Please refer to the publication for further details.

Simulation Details and Model Systems

The GAFF previously parameterized for aspirin is applied for all calculations again with the GROMACS package. The thermodynamic cycle explained in chapter 7.1 is employed for the free energy calculations. A further notice that should be mentioned is that the simulation cells considered are one half of the one represented in Figure 7.1. In other words, from Fig. 7.1 it can be seen that half of the simulation cell would be idle when the systems with positive and negative kink sites together are considered. Hence, the entire system has been chosen split into two separate simulation cells, each containing half of the step edge. This way the costs of the simulations are reduced considerably. The top view of a possible kinked step edge direction [100] is shown as an example for such system separation in Figure 8.4. For all simulations a slab model is employed with periodic boundary conditions at all directions. The surface models are again built by the GAFF force previously parameterized (chapter 8.1). The crystal slab thickness amounts to four unit cells, i.e. eight molecular layers. The interface systems are prepared by filling the vacuum gap between opposite slab surfaces with pre-equilibrated water molecules. The thickness of the water layer is chosen sufficiently large in order to retain a significant region of bulk-like water molecules in between the two slab surfaces. The latter is assessed by monitoring the water density profile.

All the systems considered are first simulated at constant temperature, 300 K, and volume for 100-200 ps with position restraints employed on all heavy atoms of the aspirin molecules. During

this time the water molecules and surface atoms are allowed to equilibrate. Then, the pressure of the system along the surface normal is adjusted to 1 atm by 3 ns long simulations without any position restraints. The box vectors along the lateral directions are fixed, but along the surface normal direction, i.e. the water layer direction, the final height is determined after the pressure adjustment. An additional 2-3 ns simulation at constant volume and temperature is run before starting the thermodynamic cycle to determine the average positions of kink molecules. During equilibration, the temperature is controlled by Nose-Hoover [142] thermostat by coupling the kinetic energy at a relaxation time of 0.5 ps and the barostat employed is Parrinello-Rahman [78]. The production runs are 3 ns long and they are performed at constant temperature and volume at the box height determined beforehand. The atoms of the bottom layer of the aspirin slabs are fixed at their positions to mimic a bulk crystal substrate for the last equilibration under constant volume and pressure and for the production simulations. During the decoupling process, the new thermostat suggested by Bussi *et al*. [73] that rescales the velocities for stochastic processes, is employed. Lennard-Jones potentials and the short-range Columbic potentials are truncated at a cutoff distance of 1 nm, and long-range interactions are calculated via the particle-mesh-Ewald (pme) [143-144] method. The long-range interactions are updated at every 10 steps.

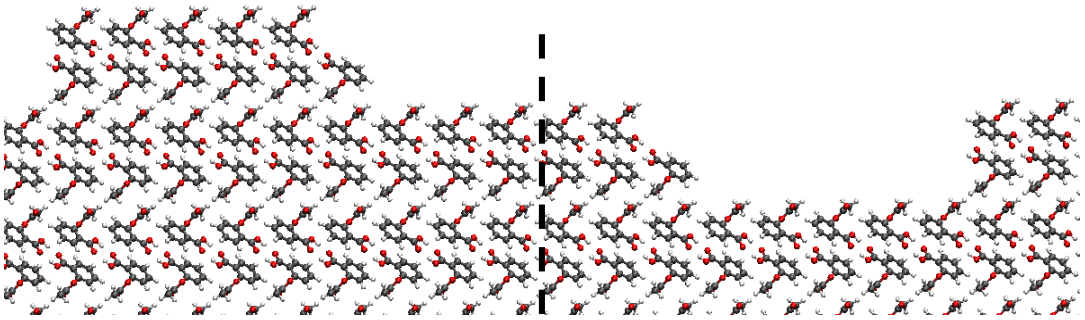

Figure 8.4: Top view of step edge direction [100] for an aspirin(001) face with the dashed lines indicating the system separation.

As mentioned in chapter 7.1, electrostatic and Lennard-Jones interactions are decoupled separately as is suggested for complex systems [119], and the number of λ points used for each decoupling is asymmetric. For the decoupling of electrostatic interactions, intermediate λ_{el} points between 0 to 1 states are placed at a spacing of 0.1 with an additional point at $\lambda_{\mathrm{el}} = 0.05$ to improve sampling of the steep initial free energy gradient. Since the decoupling of Lennard-Jones potentials is trickier and affected by solvent penetration most of the time, more λ_{LJ} points are employed. Specifically, between 0.6 to 1 states, four to six more intermediate values are placed. The soft-core

parameter α (see eq. 7.3) used during the Lennard-Jones interactions decoupling is set at the GROMACS default value of 0.3 [43].

To calculate the restraining free energy contributions, thermodynamic integration is carried out numerically by using the trapezoidal rule, based on seven k-values for eq. 7.4: 0, 50, 100, 250, 500, 750, and 1000 kJ/mol/nm^2. To sample the steep gradient at small k values, a finer spacing is employed in this region. A final value of $k_{\max}$ = 1000.0 kJ/mol/nm^2 is found sufficient to provide a balance between preserving the molecular arrangement and a reasonably low restraint free energy, associated with accordingly low error values.

Results and Discussion

All contributions of kink free energies are displayed in Table 8.2. for six step edge directions. The error estimates in the decoupling free energy calculation are done by splitting the data into blocks taking into account time correlations, and determining the free energy differences over those blocks and assuming the blocks are independent. The final error estimate is determined from the average variance over five blocks [145]. The total error of restraint free energy contributions is also calculated by block averaging via error propagation of the statistical uncertainties on individual generalized force values (see eq. 7.4). It is seen that symmetry related edges, i.e. opposite step edge directions have similar kink free energies, whereas higher differences in kink free energies are observed for different edge directions such as $[100]$ and $[010]$. Based on the kink densities, the edge directions $[100]$ and $[\bar{1}00]$ are quite smooth, straight edges, with a low number of kinks. Along $[1\bar{1}0]$ and $[\bar{1}10]$ edges on the other hand the number of kinks is very high, hence these are kinked edges. The remaining edge directions $[010]$ and $[0\bar{1}0]$ fall in between these two kinked and straight structures.

Table 8.2: Free energy components during the calculation of kink free energies for kinks within step edges of several orientations, as well as final kink free energies per kink site

step direction	$\Delta G_{\text{step}}^{\text{rest}}$ (kJ/mol)	$\Delta G_{\text{step}}^{\text{decou}}$ (kJ/mol)	$\Delta G_{\text{kink}}^{\text{rest}}$ (kJ/mol)	$\Delta G_{\text{kink}}^{\text{decou}}$ (kJ/mol)	ΔG_{kink} (kJ/mol)
$[100]$	93.6 ± 0.5	1308 ± 1.4	102.6 ± 0.6	1244.5 ± 1.4	13.7 ± 0.6
$[\bar{1}00]$	106.6 ± 0.6	1267.1 ± 1.3	115.3 ± 0.7	1198.3 ± 1.6	14.9 ± 0.6
$[010]$	55.6 ± 0.4	574.8 ± 1.0	56.2 ± 0.4	557.3 ± 1.2	4.2 ± 0.5
$[0\bar{1}0]$	56.1 ± 0.4	572.3 ± 2.7	61.8 ± 0.5	553.9 ± 2.2	3.2 ± 0.9
$[1\bar{1}0]$	61.6 ± 0.4	812.1 ± 2.4	67.4 ± 0.5	800.5 ± 0.8	1.5 ± 0.7
$[\bar{1}01]$	75.0 ± 0.5	798.9 ± 3.0	85.8 ± 0.6	784.6 ± 1.8	0.9 ± 0.9

8.2.2 Rate Constants of Detachment Processes from Kink Sites

The two-step algorithm explained in chapter 7.2 and already successfully applied to a kink site molecule along the step edge [100] aspirin (001) face [117] is applied for the calculation of the detachment rate constants from kink sites along other step edges on this face. The criteria for creating kink structures are already explained (beginning part of chapter 8.2). Hence, again only meta-stable kink structures are considered, i.e. the same kink structures as considered in the kink free energy calculations. This way, the number of kink types along the edges is reduced to opposite pairs of kink site structures with equal probabilities of forming. Since aspirin is a low-symmetry molecule, opposite kink site molecules along the same edges expose different terminations, though the kink site molecule before and after detachment/attachment will still be same type of kink molecule and the same bonds will be broken/formed for both kink site molecules. These kink sites may nevertheless have different detachment rate constants, especially in a solvent environment. Therefore, these opposite kink site molecules are labelled as type A or type B, for simplicity. They may not be considered as different kink types based on the definition of Kuvadia *et al*. [44], as the detachment/attachment would result in the same kink site structure and the same types of bonds are broken/formed for both kink types. In other words, the kink sites do not transform from one type to another during incorporation/disincorporation processes. Hence, it is not necessary to write a master equation for the net flux as done by Kuvadia *et al*. [44] for the cases of there being more than one kink type along edges, cf. eq. 3.18 in chapter 3.2. In the cases considered here, detachment/attachment of kink type A will also result in kink type A until the last two remaining molecules; kink type A on one side and kink type B on the other site. Therefore, it is still feasible to calculate the net flux by considering the equation for centrosymmetric cases with only one kink type per edge (detailed explanation is given in chapter 8.2.3).

Simulation Details and Model Systems

For the detachment rate simulations, the model systems with positive kinked step edges are prepared. The preparation method is the same one as for the free energy calculations by building the models from GAFF. The simulation cells used, however, are slightly smaller than the ones used for the free energy calculations, i.e. the slab thickness amounts only to four layers. The water density profile is again checked for bulk-like water properties. The effect of using smaller cells is tested to avoid finite size effects for the bias potential construction, i.e. the metadynamics simulations, by comparing the bias potential value before the first escape events of both small and

large systems. Since the results show almost equal bias potential values, the smaller cells are used for the ensuing hyperdynamics simulations to increase the simulation times. Before starting the metadynamics simulations, the system is again equilibrated in an *NVT* ensemble with all atoms of aspirin molecules restraint for around 100-200 ps. The equilibrium box height is determined as it was done for the free energy calculations. For both metadynamics and hyperdynamics simulations, the box height is then fixed and the atoms of the bottom layer molecules are restrained again to mimic the bulk crystal.

The simulations have been implemented in the PLUMED plugin interfaced with GROMACS [146]. The center of mass (COM) vector for the kink site molecule considered is used as CV for metadynamics. Instead of accumulating Gaussian hills on selected CVs continuously, a push and equilibrate method is employed to increase the reproducibility of the bias potential construction [117]. This approach comprises an alternating sequence of adding new hills, to push the system away from the current position in CV space, and equilibrating the system inside the new augmented energy landscape without depositing additional hills. As the temporary and final escape processes predominantly take place during the equilibration phase, this procedure facilitates the preparation of a more homogeneous and compact bias potential. By this way, the system is prevented to escape due to the accumulation of hills on only some part of the free energy landscape without exploring the other part of the landscape. The deposition frequency of hills during push sequences is always 1 ps^{-1}. The heights and widths of hills deposited are slightly changed depending on the kink system considered so that the transition region as well as truncation point could be determined better. Generally, widths of hills are between 0.03-0.04 nm; heights of hills are 0.2-0.5 kJ/mol, mostly starting with smaller hill heights and switching to higher height. In some cases, the equilibration phase time is also increased for further equilibration of accumulated hills. The hyperdynamics simulations are initiated for different conformations that are obtained from snapshots from long equilibration runs. They are all assigned new velocities to further de-correlate the snapshots and run for 15 ns to obtain long biased escape times. During metadynamics and hyperdynamics simulations, the temperature is again controlled by Nose-Hoover thermostat with 0.5 ps relaxation time [142]. Here, also Lennard-Jones potentials and the short-range Columbic potentials are truncated at a cutoff distance of 1.0 nm; long range interactions are calculated via particle-mesh-Ewald (pme) [143-144] method. The time step is of 0.001 fs and the long range interactions are again updated at every 10 steps.

Results and Discussion

Top views of the aspirin step edge structures with the employed dissolution units, i.e. aspirin dimers, marked with solid lines are displayed for each edge direction together with the trajectories of the COM-components observed during metadynamics and the reweighted free energy profile, as well as the bias potential values encountered during the hyperdynamics simulations. The detachment of the first kink site molecule (labeled by bold atomic spheres) is always considered as the slower process, which is followed by the fast detachment of the second molecule. The kink site molecules marked by blue solid lines, kink site type of B, generally have slower detachment rate constants than the kink site molecules marked by green solid lines, kink type A. This can be rationalized based on the accessibility of solvent water molecules to hydrogen bonds, as shown in Figure 8.25. Since the water access to hydrogen bonds for the dissolution units considered for edges $[\bar{1}10]$ and $[1\bar{1}0]$ is the same for opposite kink sites, only the detachment of one kink type for these two edges is considered. For the $[1\bar{1}0]$ step edge, parallel replica dynamics simulations (PRD) [110] are performed along with hyperdynamics simulations to check the validity of the bias potential considered, as the bias potentials encountered during hyperdynamics simulation have a different shape than the ones for other edges. All detachment rate results evaluated can be found in Table 8.13. In addition, detachment rate constants for the second molecules in the dimeric dissolution units for the slow detaching kink type B are also calculated to check that it is indeed detaching faster than the first molecule in the dimeric units.

The trajectories obtained from the hyperdynamics simulations are analyzed to establish the free energy profiles first. To map the dissolution process onto a one-dimensional reaction coordinate, which accounts for the anisotropic COM-fluctuations, a dimensionless order parameter d is defined as

$$d^2 = \left(\vec{r} - \vec{r}_0\right)^T . \underline{C}^{-1} . (\vec{r} - \vec{r}_0) , \qquad 8.1$$

where $\underline{C}$ is the covariance matrix of the COM-fluctuations from the average position and is calculated from the equilibrium MD simulations, $\vec{r}$ and $\vec{r}_0$ are the positions of the kink site molecule encountered during the hyperdynamics simulations and the average positions the kink site molecule, respectively. The number of hyperdynamics simulations employed is also provided in the figure captions.

The evaluation of rate constants is based on Transition State Theory (TST) [147] which is essentially the following formalism for the calculation of transition rate constant from state *A* to *B*:

$$k_{AB}^{\mathrm{TST}} = \langle \dot{\lambda} \ominus (\dot{\lambda}) \rangle_{\lambda^*} \frac{\exp[-\beta G(\lambda^*)]}{\int_{\lambda_{\min}}^{\lambda^*} \exp[-\beta G(\lambda)]} \,. \tag{8.2}$$

$G(\lambda)$ denotes the free energy profile along a suitably chosen reaction coordinate λ, and λ^* parameterizes the transition state dividing surface, separating initial state *A* from final state *B*, while $\ominus$ represents the step function. To account also for dynamical re-crossings of escape trajectories after passing the barrier, this rate constant is multiplied by a time-dependent transition constant, κ. The reactive flux from state *A* to *B* is

$$\kappa(t) = \frac{\langle \dot{\lambda}(\mathbf{R}_0) \ominus [\lambda(\mathbf{R}_t) - \lambda^*] \rangle_{\lambda^*}}{\langle \dot{\lambda}(\mathbf{R}_0) \ominus [\dot{\lambda}(\mathbf{R}_0)] \rangle_{\lambda^*}}. \tag{8.3}$$

The subscript 0 refers to the point of barrier crossing and the subscript λ^* denotes the transition state ensemble. In practice, the biased rate constants $k_{\mathrm{bias}}^{\mathrm{TST}}$ in eq. 7.5 is calculated by the biased mean-first passage times (MFPT) over a suitable chosen boundary, i.e. dimensionless parameter d in this case. It is also possible to reweight the biased MFTP and directly calculate k^{TST} (see Supporting Information of ref. [117] for more details) and this is actually done in practice here.

The rate constants from parallel replica dynamics simulations are obtained by dividing the number of escape events by the simulation time.

$$k^{\mathrm{PRD}} = \frac{n_{\mathrm{escape}}}{t_{\mathrm{total}}} = \frac{n_{\mathrm{escape}}}{\left(n_{\mathrm{replica}} t_{\mathrm{replica}}\right)}, \tag{8.4}$$

where n_{replica} is the number of replicas used and t_{replica} is the simulation time of each replica. Multiplication of them gives the total time accumulated in the parallel replica simulations.

The errors in the detachment rate constants can be roughly estimated from the deviations of the mean value, which can be taken by averaging the detachment rates constants for different choices of transition barrier. Three such locations are considered, $d = 15, 17, 19$.

1. Step edge [100]

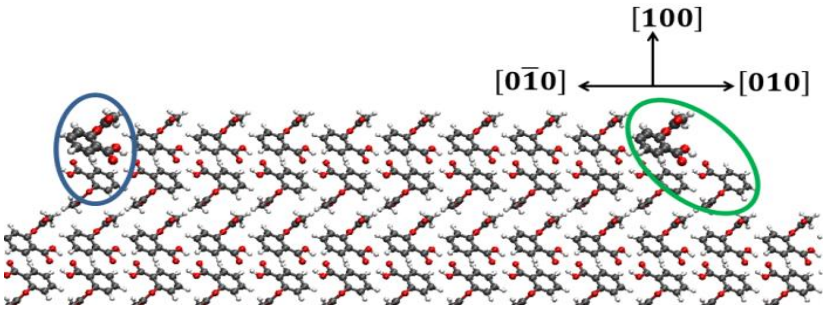

Figure 8.5: Top view of the [100] edge direction of the aspirin (001) face. For clarity only the terrace top layer above the step edge is displayed and (green and blue) solid lines highlight the considered dimeric dissolution unit at the two kink sites shown. C, O,H atoms are shown as gray, red and white spheres, respectively. The detachments of molecules shown with bold atomic spheres are considered as the rate determining.

a) Kink Type A

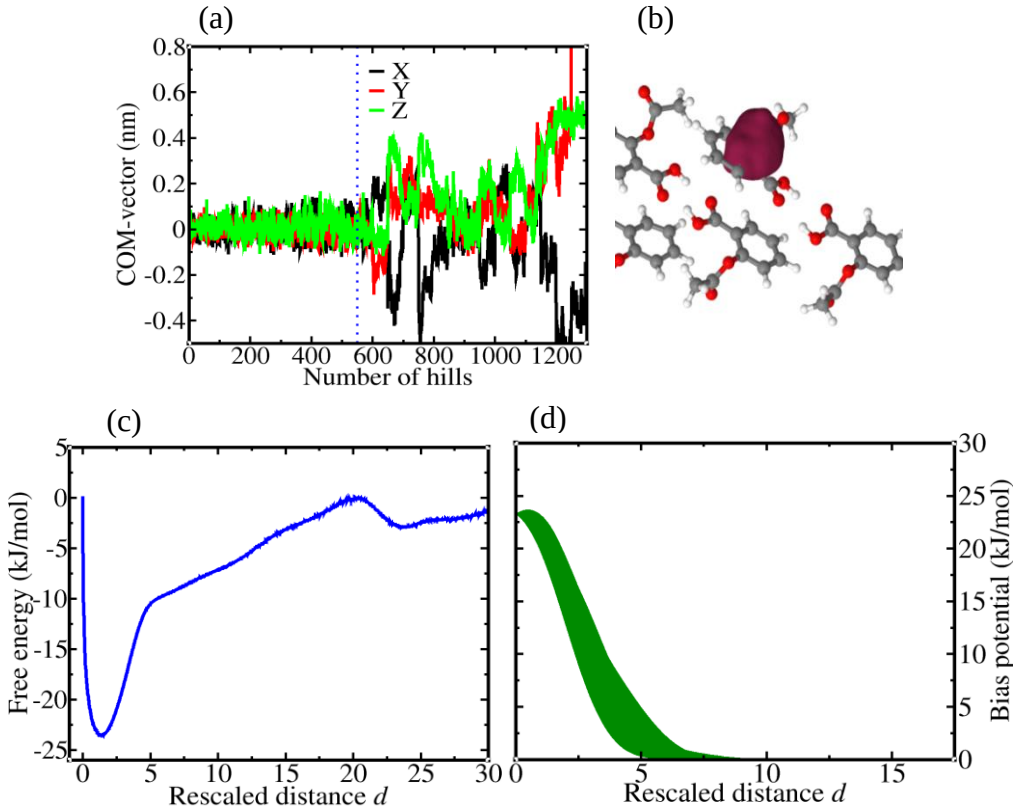

Figure 8.6: (a) Trajectory of COM components for the kink type A molecule along the [100] edge of the aspirin (001) face. The truncation point at 550 hills is marked by dotted lines. (b) The outer boundary of the isosurface of the fixed bias potential. (c) Free energy profile obtained from hyperdynamics simulations by reweighting. (d) The encountered bias potential values versus the dimensionless order parameter, *d*, during the hyperdynamics simulations. The rate constant obtained by 64 hyperdynamics simulations is same as previously reported in ref. [117]. See Table 8.13.

Table 8.3: TST rate constants, transmission coefficients, and reactive-flux rate constants, boost factor, number of detachment events, and barrier-to-well height for different choices of the transition barrier for the kink type A molecule along the [100] edge of the aspirin (001) face.

Barrier location	k_{TST} (s^{-1})	$\kappa(t)$	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	4.62×10^{7}	0.02	7.79×10^{5}	1.28	85.24	2241	20.46
$d = 17$	2.12×10^{7}	0.04	9.31×10^{5}	1.07	83.61	1004	21.53
$d = 19$	1.15×10^{7}	0.06	7.64×10^{5}	1.31	82.69	541	23.08

b) Kink Type B

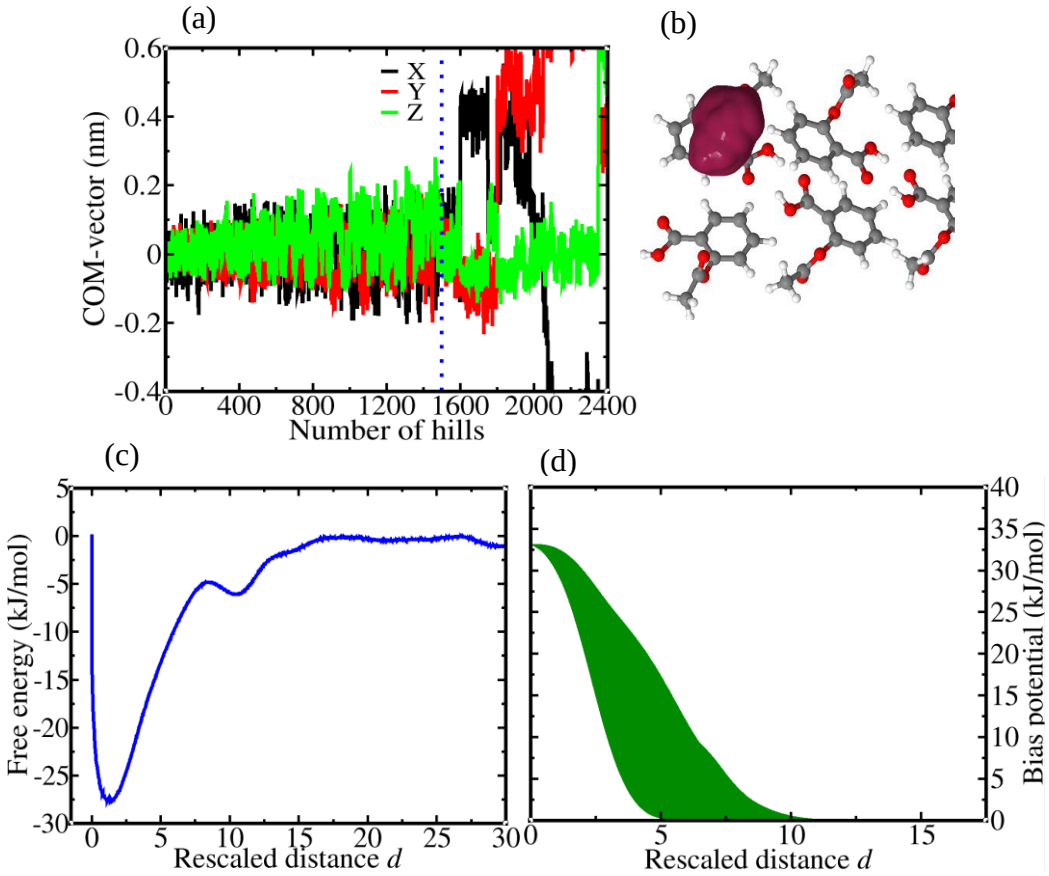

Figure 8.7: Same as Figure 8.6 for the kink site molecule considered. The truncation point at 1500 hills is marked by dotted lines and 64 hyperdynamics simulations are performed.

Table 8.4: Same as Table 8.3 for the considered kink site molecule.

Barrier location	k_{TST} (s^{-1})	$\kappa(t)$	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	2.23×10^{6}	0.04	8.82×10^{4}	11.34	1472.56	1460	26.40
$d = 17$	1.04×10^{6}	0.08	8.72×10^{4}	11.46	1430.83	630	27.46
$d = 19$	5.47×10^{5}	0.17	9.56×10^{4}	10.46	1403.36	316	27.54

c) Second molecule in dissolution unit

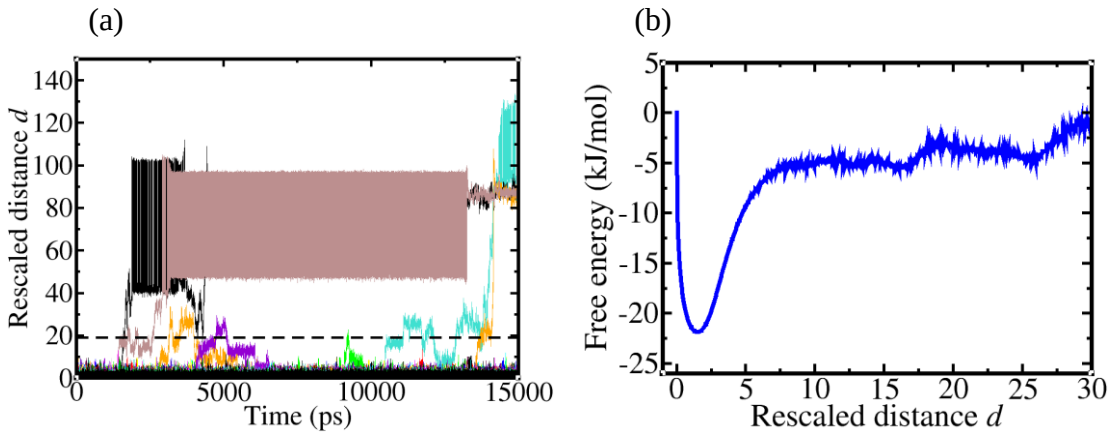

Figure 8.8: (a) The trajectories obtained from parallel replica dynamics (PRD) for the second molecule in the dissolution dimer unit for the [100] step edge. There detachment events are observed from 64 simulations. (b) Free energy profile obtained by reweighting of PRD simulations.

2. Step edge [$\bar{1}00$]

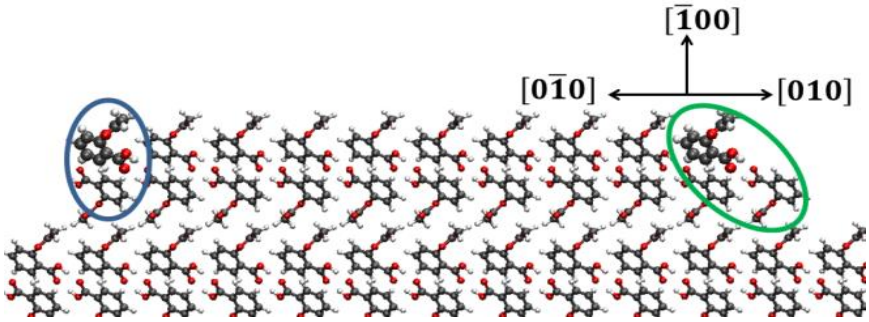

Figure 8.9: Same as Figure 8.5 for the [$\bar{1}$00] edge direction of the aspirin (001) face with dissolution units considered marked by solid lines.

a) Kink Type A

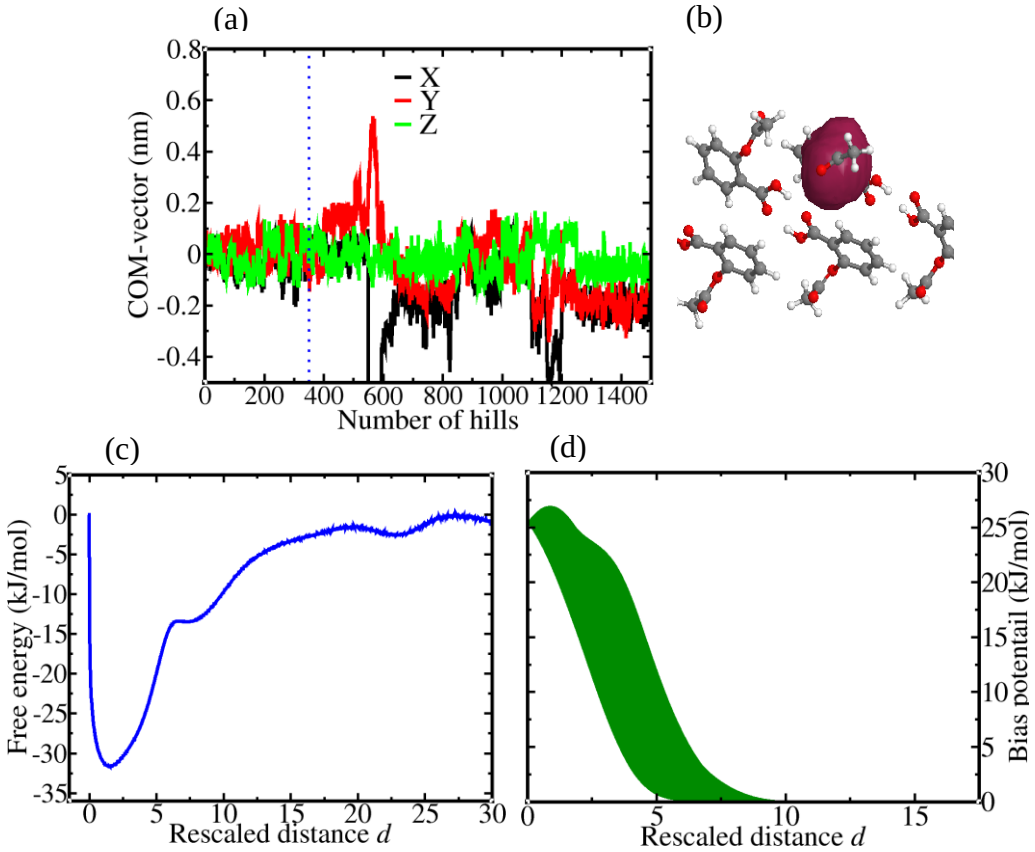

Figure 8.10: Same Figure 8.6 for the kink type molecule considered. The truncation point is at 350 hills and 128 hyperdynamics simulations are performed.

Table 8.5: Same as Table 8.3 for the considered kink site molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	4.69×10^6	0.01	4.11×10^4	24.36	411.26	2228	28.34
$d = 17$	2.13×10^6	0.01	3.36×10^4	29.72	409.75	915	29.24
$d = 19$	1.18×10^6	0.03	3.67×10^4	27.23	408.76	407	30.06

b) Kink Type B

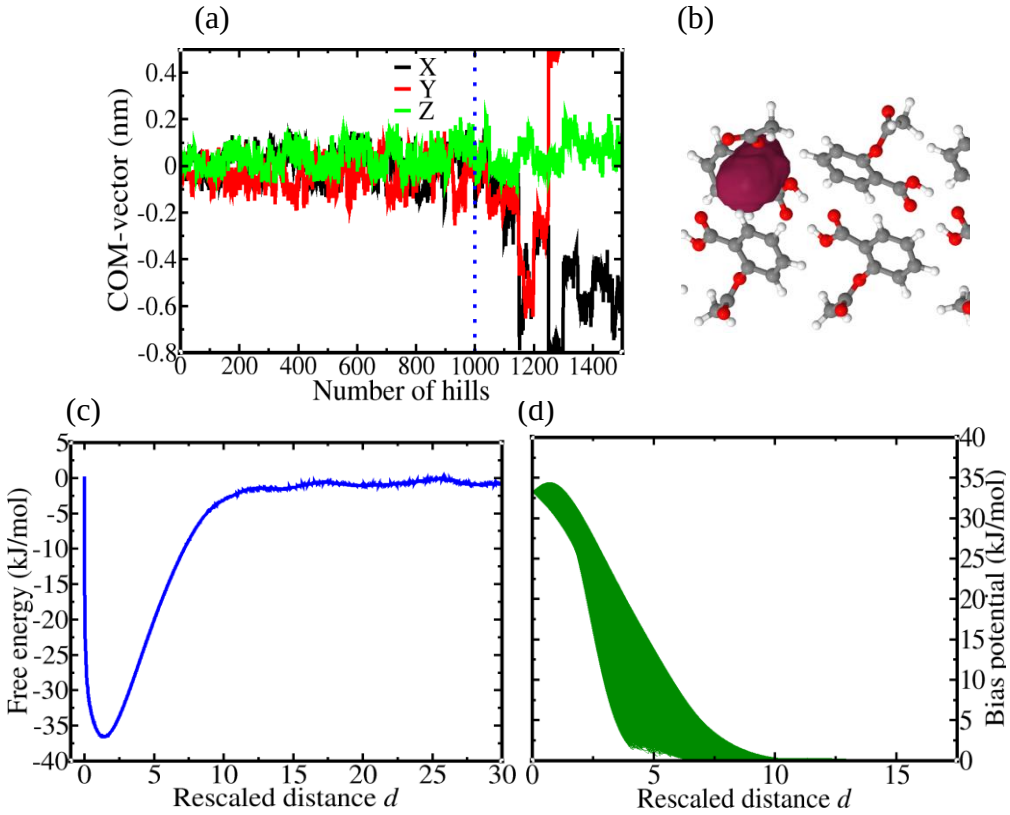

Figure 8.11: Same as Figure 8.6 for the kink type molecule considered. The truncation point is at 1000 hills and 64 hyperdynamics simulations are performed.

Table 8.6: Same as Table 8.3 for the considered kink site molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	5.14×10^5	0.01	7.72×10^3	129.46	5927.69	1020	35.27
$d = 17$	2.13×10^5	0.04	9.67×10^3	103.43	5907.17	260	35.91
$d = 19$	1.16×10^5	0.12	1.36×10^3	73.67	5889.99	112	35.66

c) Second molecule in dissolution unit

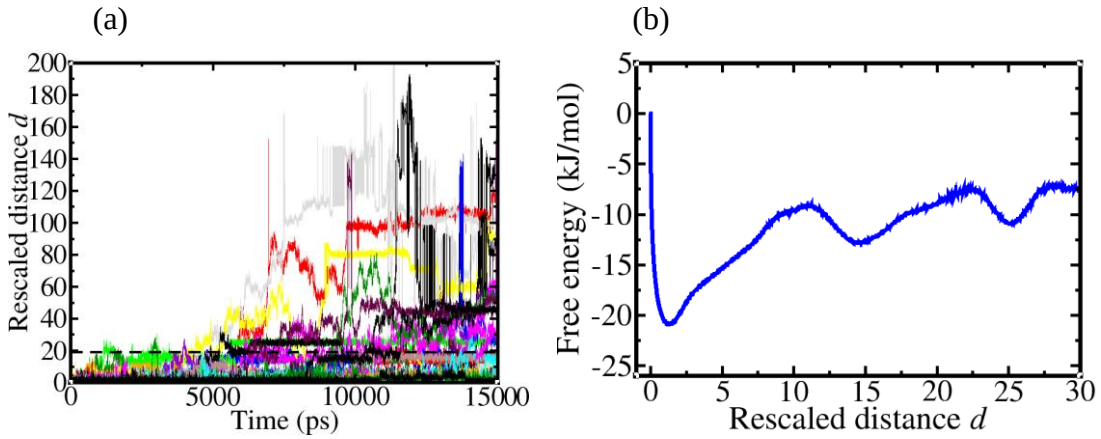

Figure 8.12: (a) The trajectories obtained from parallel replica dynamics (PRD) for the second molecule in the dissolution dimer unit for the $[\bar{1}00]$ step edge, nine detachment events are counted from 64 simulations. (b) Free energy profile obtained by reweighting of PRD simulations.

3. Step edge [010]

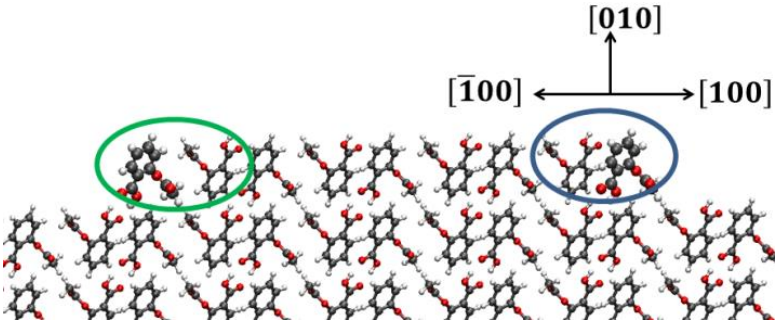

Figure 8.13: Same as Figure 8.5. Top view of the [010] edge direction of the aspirin (001) face with dissolution units considered marked by solid lines.

a) Kink Type A

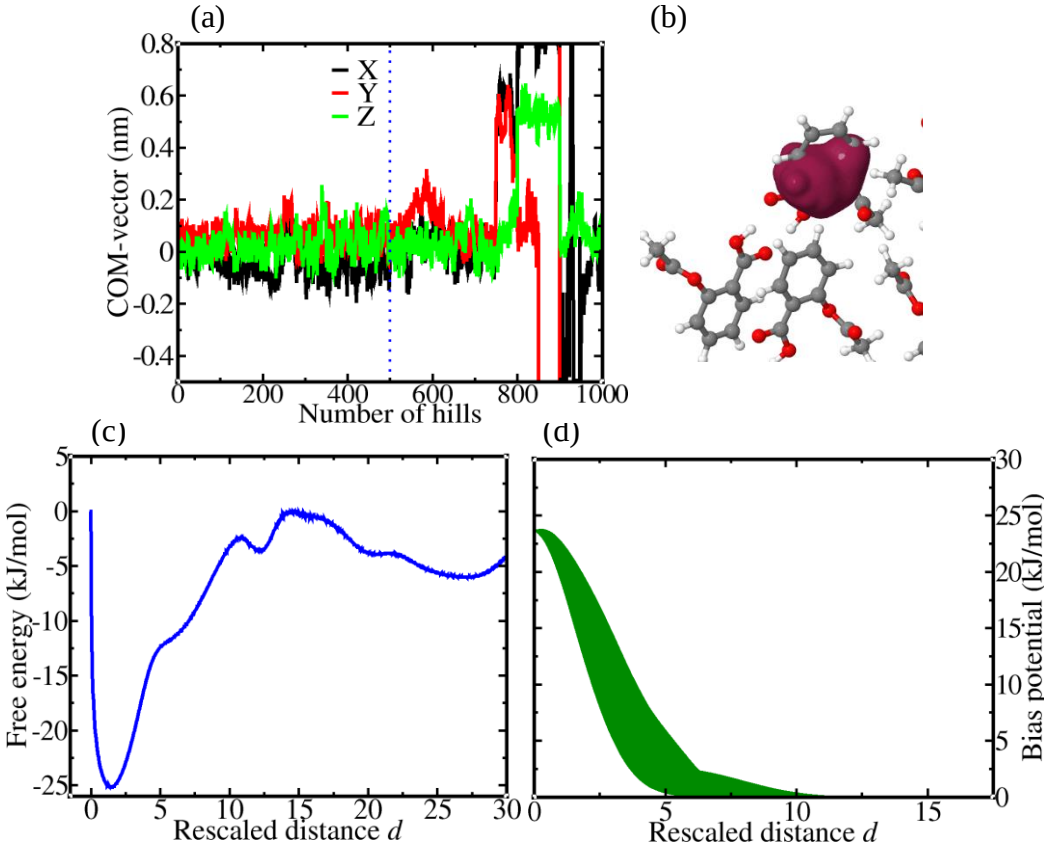

Figure 8.14: Same as Figure 8.6 for the kink type molecule considered. The truncation point is at 500 hills and 128 hyperdynamics simulations are performed.

Table 8.7: Same as Table 8.3 for the considered kink site molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	1.97×10^{7}	0.05	9.44×10^{5}	1.06	79.77	1397	24.96
$d = 17$	1.13×10^{7}	0.08	9.62×10^{5}	1.04	79.43	690	24.31
$d = 19$	7.85×10^{6}	0.15	1.15×10^{6}	0.87	79.10	424	22.12

b) Kink Type B

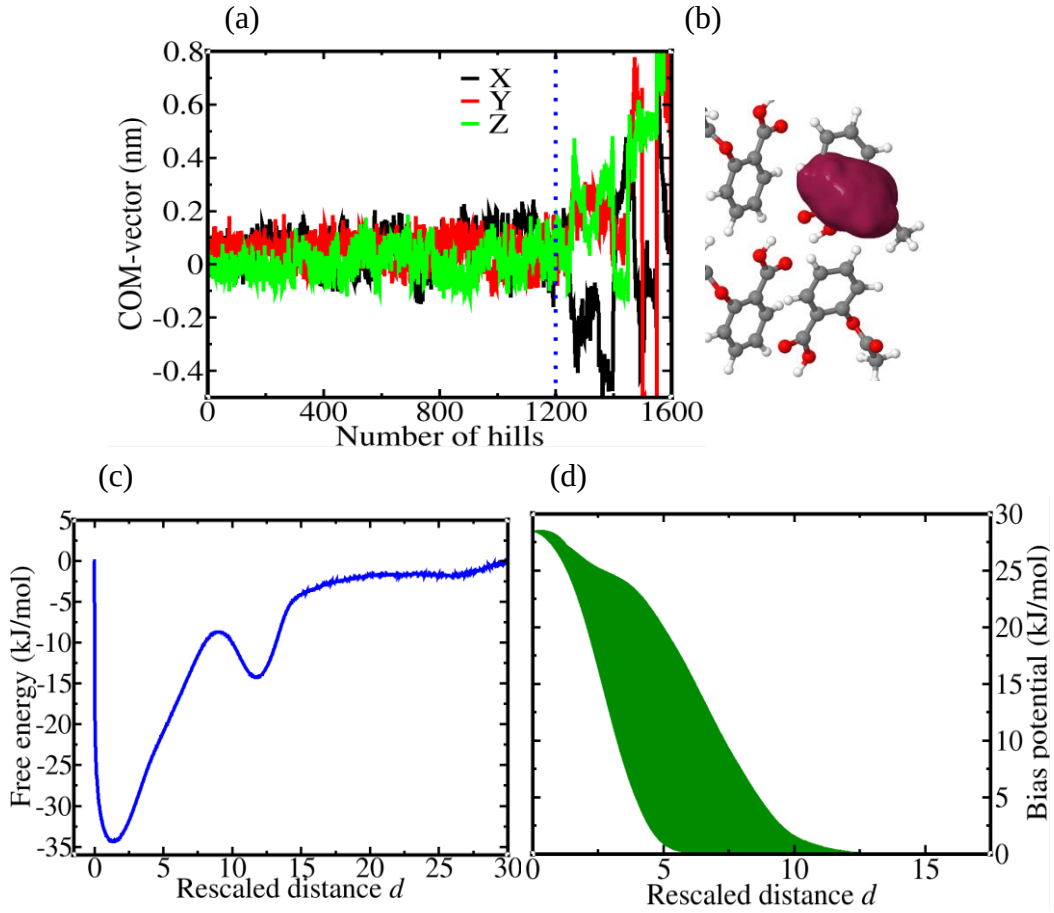

Figure 8.15: Same as Figure 8.6 for the kink type molecule considered. The truncation point is at 1200 hills and 128 hyperdynamics simulations are performed.

Table 8.8: Same as Table 8.3 for the considered kink site molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	1.12×10^{6}	0.02	2.50×10^{4}	39.96	1134.56	1232	30.23
$d = 17$	5.44×10^{5}	0.09	4.80×10^{4}	20.85	1129.27	312	31.65
$d = 19$	2.97×10^{5}	0.21	6.22×10^{4}	16.07	1126.00	131	32.32

c) Second molecule in dissolution unit

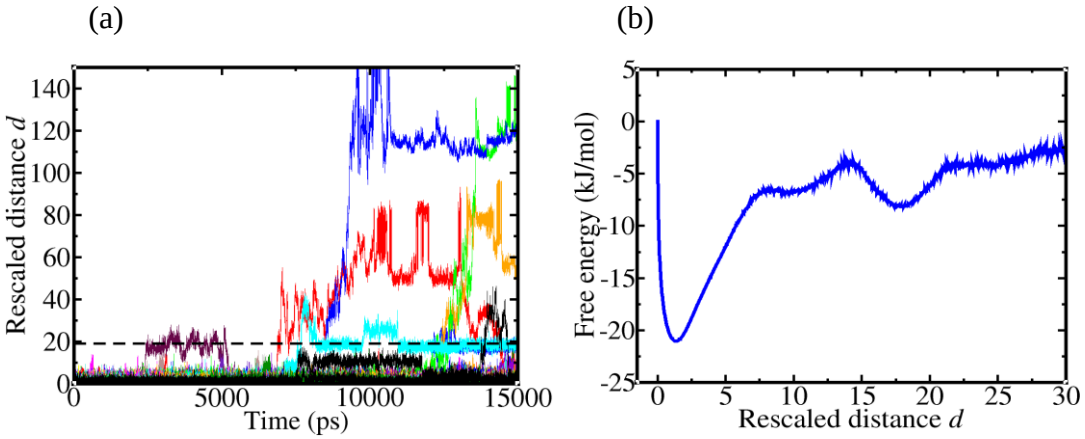

Figure 8.16: (a) The trajectories obtained from parallel replica dynamics (PRD) for the second molecule in the dissolution dimer unit for the [010] step edge. Four detachment events are counted from 64 simulations. (b) Free energy profile obtained by reweighting of PRD simulations.

4. Step edge $[0\bar{1}0]$

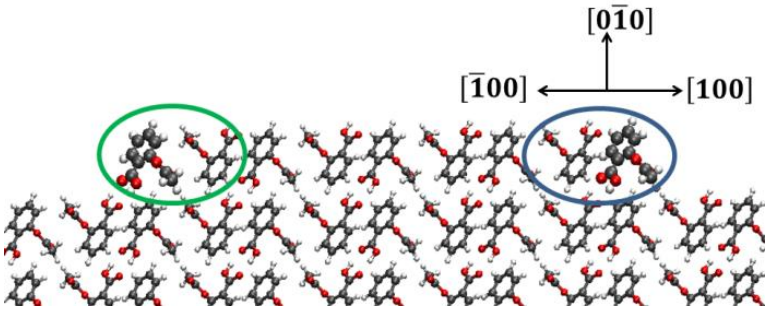

Figure 8.17: Same as Figure 8.5. Top view of the $[0\bar{1}0]$ edge direction of the aspirin (001) face with dissolution units considered marked by solid lines.

a) Kink Type A

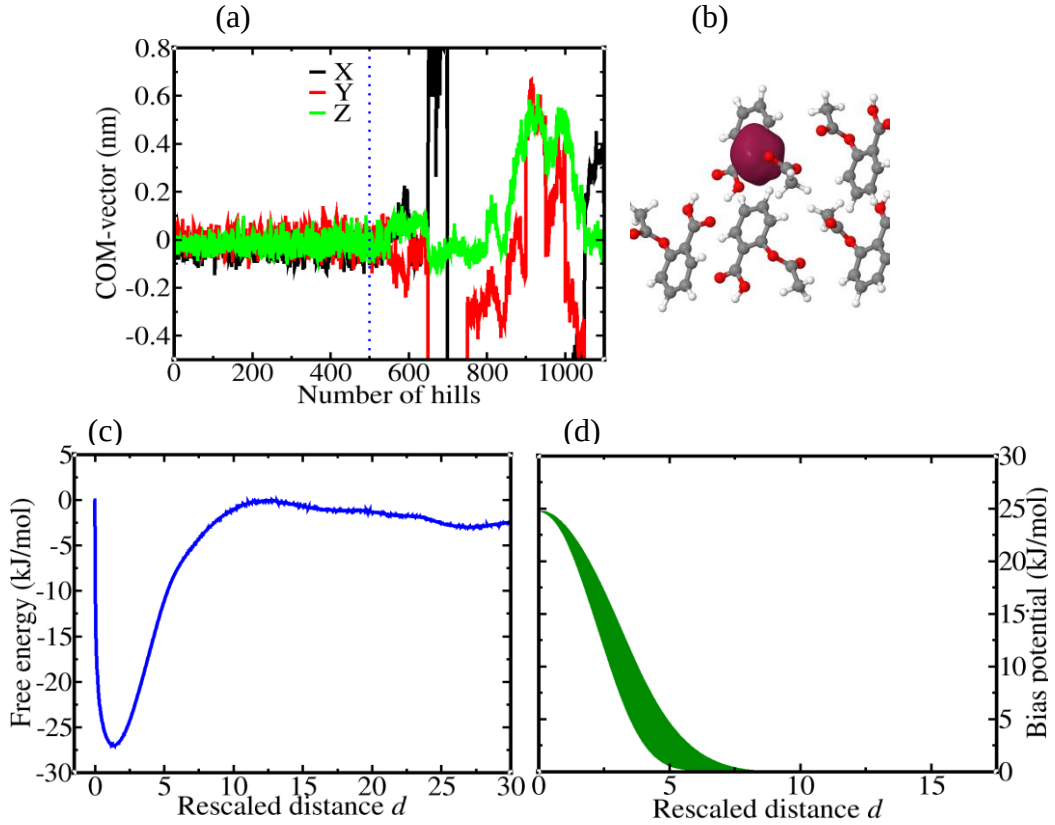

Figure 8.18: Same as Figure 8.6 for the kink type molecule considered. The truncation point is at 500 hills and 128 hyperdynamics simulations are performed.

Table 8.9: Same as Table 8.3 for the considered kink site molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	2.03 x 10^6	0.13	2.56 x 10^5	3.91	375.83	305	26.28
$d = 17$	1.50 x 10^6	0.18	2.77 x 10^5	3.60	372.93	207	25.93
$d = 19$	1.13 x 10^6	0.28	3.23 x 10^5	3.10	370.62	145	25.92

b) Kink Type B

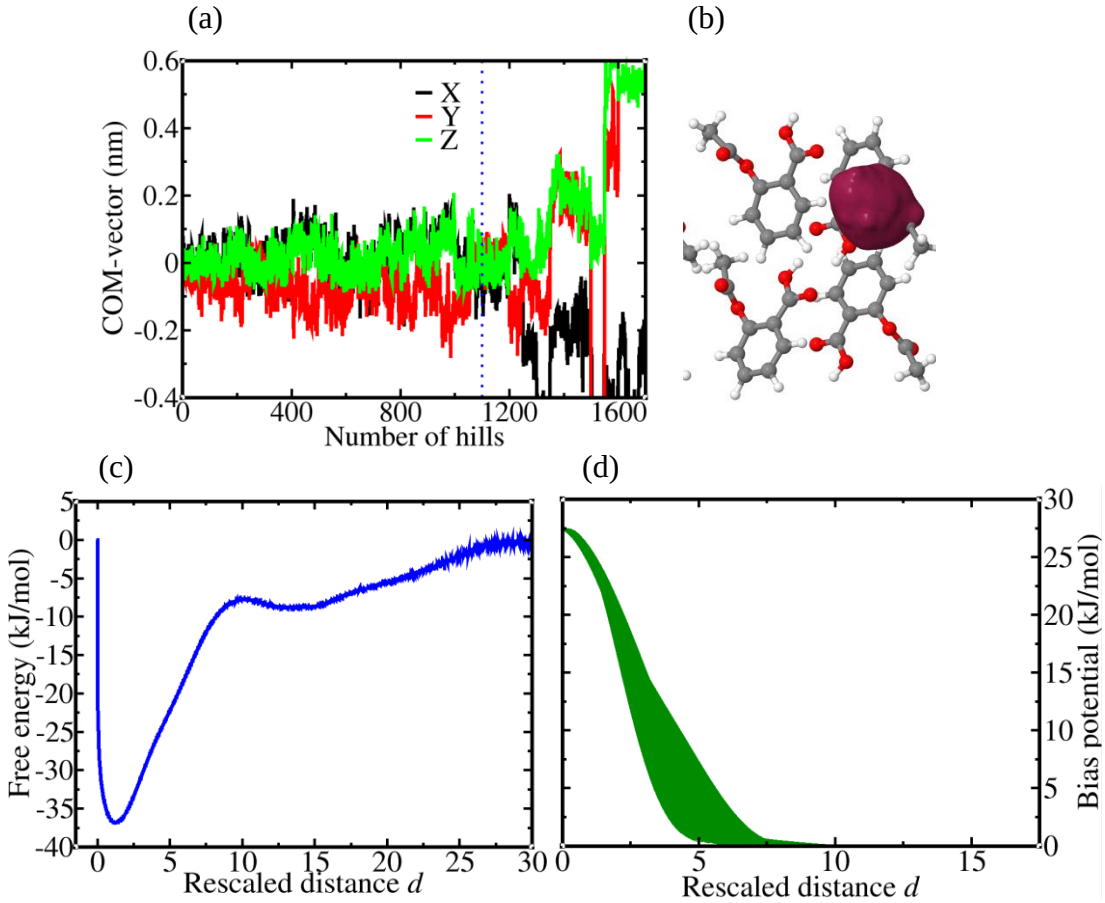

Figure 8.19: Same as Figure 8.6 for the kink type molecule considered. The truncation point is at 1100 hills and 128 hyperdynamics simulations are performed.

Table 8.10: Same as Table 8.3 for the considered kink site molecule

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	1.92×10^{7}	0.002	3.01×10^{4}	33.21	653.13	4156	28.24
$d = 17$	1.31×10^{7}	0.006	7.84×10^{4}	12.75	648.07	1875	29.53
$d = 19$	4.32×10^{6}	0.01	6.75×10^{4}	14.82	645.32	610	30.76

c) Second molecule in the dissolution unit

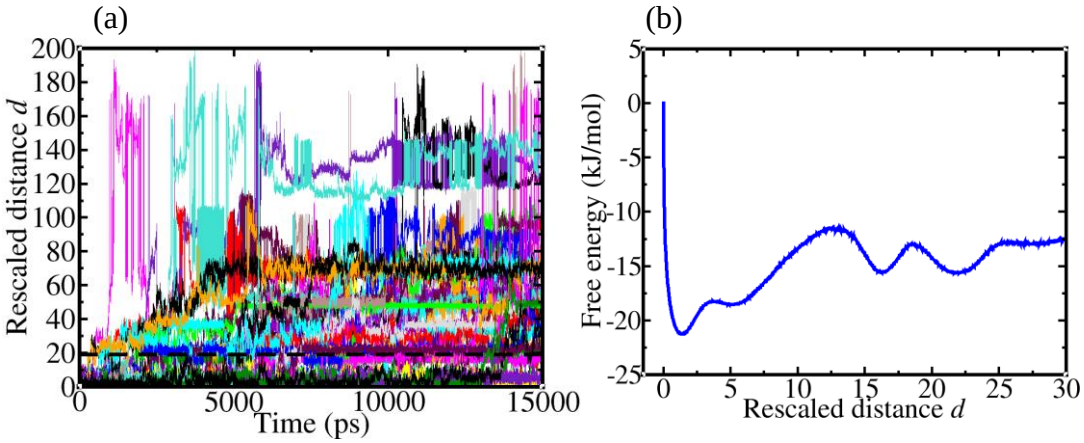

Figure 8.20: (a) The trajectories obtained from parallel replica dynamics (PRD) for the second molecule in the dissolution dimer unit for the $[0\bar{1}0]$ step edge. Twenty nine detachment events are counted from 64 simulations. (b) Free energy profile obtained by reweighting of PRD simulations.

5. Step edge $[1\bar{1}0]$

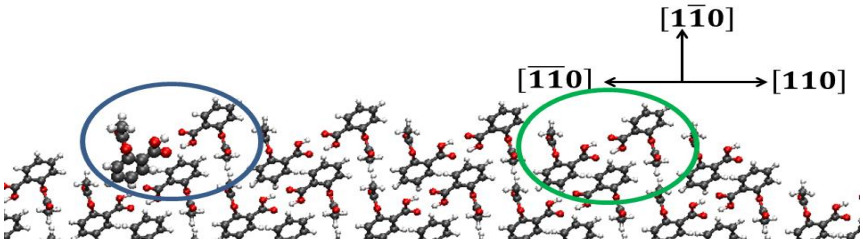

Figure 8.21: Same as Figure 8.5. Top view of the $[1\bar{1}0]$ edge direction of the aspirin (001) face with dissolution units considered marked by solid lines.

a) Kink Type A

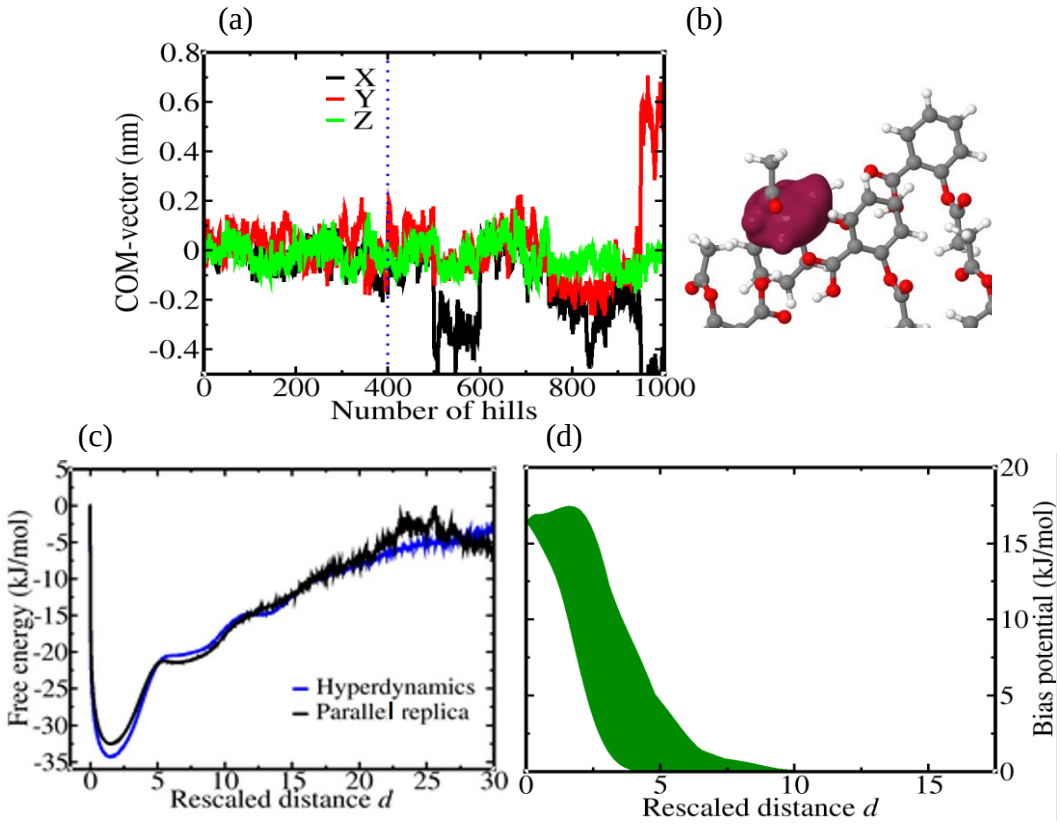

Figure 8.22: Same as Figure 8.6 for the kink type molecule considered. The truncation point is at 400 hills and 128 hyperdynamics simulations are performed. For parallel replica dynamics simulations, the same 128 initial configurations are run and one detachment event is observed.

Table 8.11: Same as Table 8.3 for the considered kink site molecule. The results obtained from PRD are also displayed.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	5.09×10^7	0.002	1.21×10^5	8.28	38.77	1165	21.88
$d = 17$	1.54×10^7	0.01	1.45×10^5	6.89	40.48	639	24.33
$d = 19$	7.98×10^7	0.03	2.43×10^5	4.11	40.41	241	25.56
PRD ($d = 17$)	-	-	5.21×10^5	1.92	-	1	22.79

6. Step edge $[\bar{1}10]$

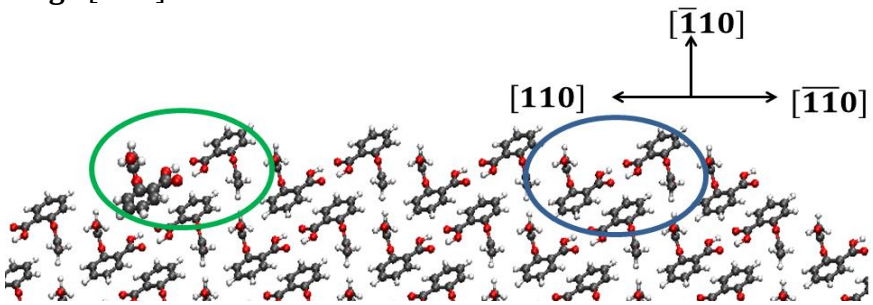

Figure 8.23: Same as Figure 8.5. Top view of the $[\bar{1}10]$ edge direction of the aspirin (001) face with dissolution units considered marked by solid lines.

a) Kink Type A

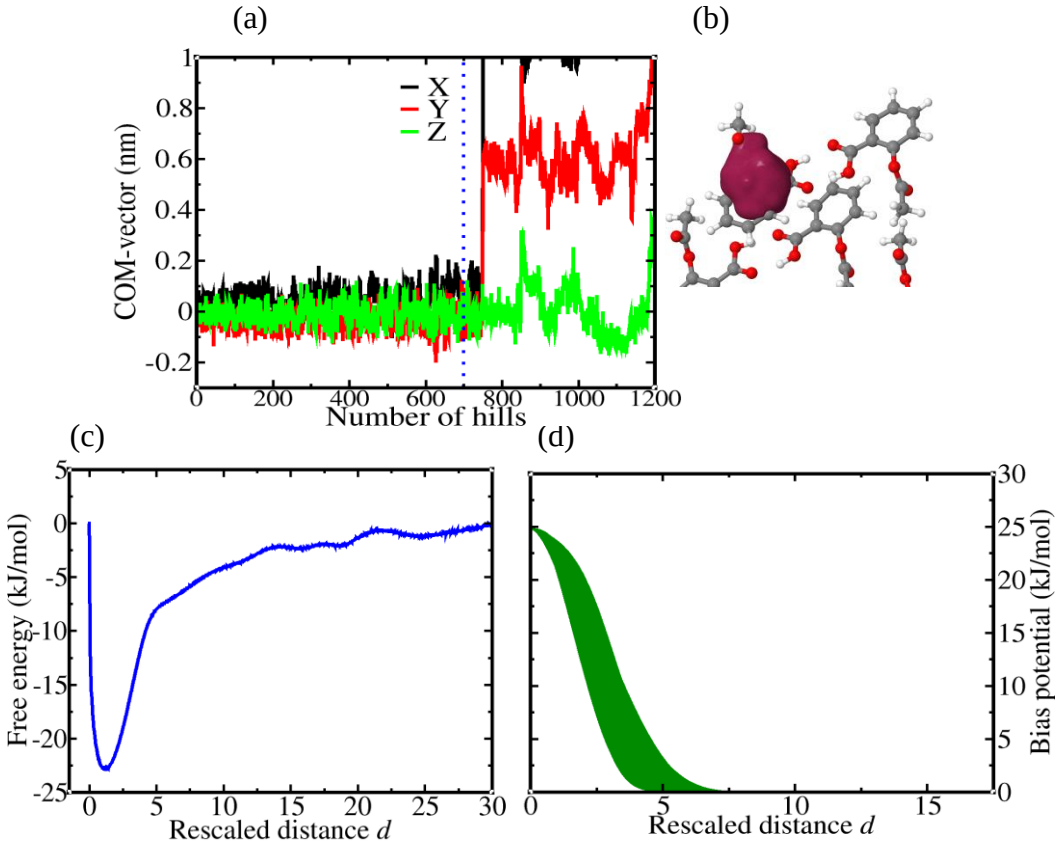

Figure 8.24: Same as Figure 8.6 for the kink type molecule considered. The truncation point is at 700 hills and 64 hyperdynamics simulations are performed.

Table 8.12: Same as Table 8.3 for the considered kink site molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
$d = 15$	1.51×10^{7}	0.06	9.96×10^{5}	1.00	131.38	812	20.59
$d = 17$	7.29×10^{7}	0.11	8.28×10^{5}	1.21	126.43	383	20.84
$d = 19$	4.40×10^{6}	0.21	9.13×10^{5}	1.10	122.35	218	20.90

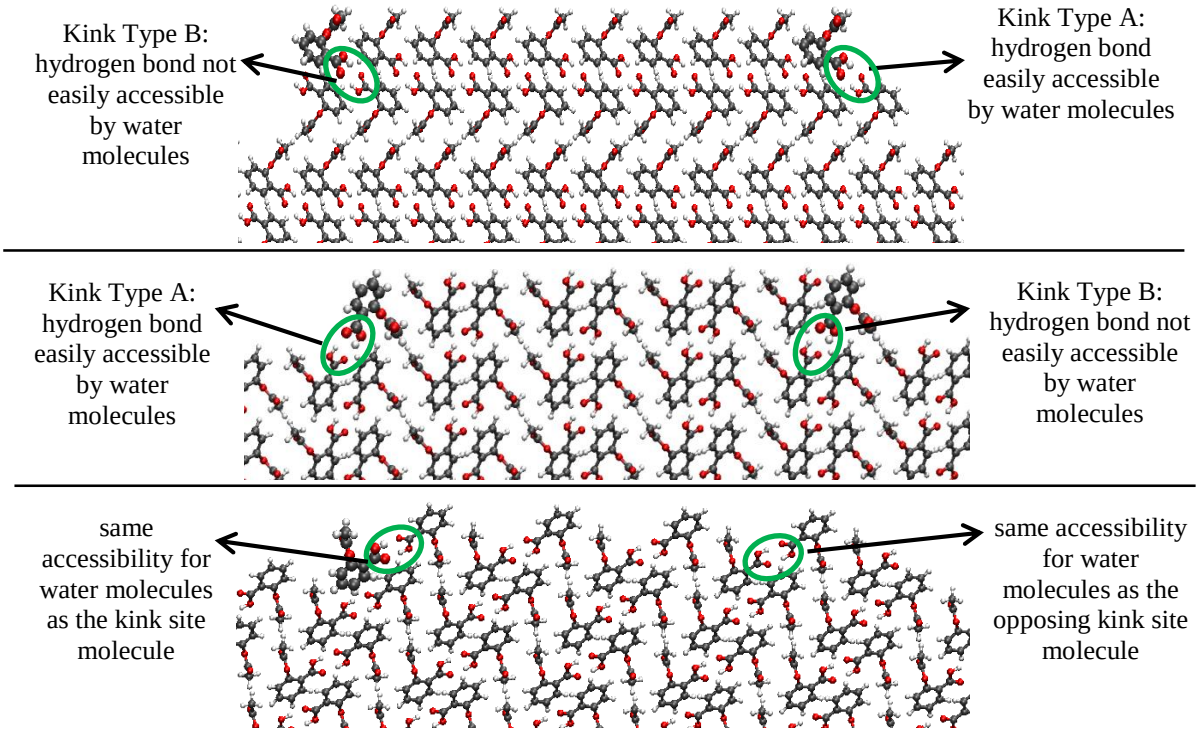

Figure 8.25: The accessibility of water molecules to intermolecular bonds between kink site molecules and their neighbors are shown for different kink types along aspirin step edges [100], [010] and [$\bar{1}$10], displayed from upper to lower parts, respectively. The same explanation is valid for the opposite edges [$\bar{1}$00], [0$\bar{1}$0] and [1$\bar{1}$0].

Table 8.13. Kink free energies and detachment rate constants with errors displayed in parentheses for different kink types along all PBC edges of the aspirin (001) face.

Edge direction	ΔG_{kink} (kJ/mole)	Detachment rate constants (s^{-1})		
		Kink Type A	Kink Type B	Second molecule in dimeric units of kink type B
[100]	13.7±0.6	7.66×10^5 ($\pm 7.57 \times 10^4$)	8.72×10^4 ($\pm 3.76 \times 10^3$)	3.13×10^6
[$\bar{1}$00]	14.9±0.6	3.36×10^4 ($\pm 3.04 \times 10^3$)	9.67×10^3 ($\pm 2.43 \times 10^3$)	9.38×10^6
[010]	4.2±0.5	9.62×10^5 ($\pm 9.46 \times 10^4$)	4.80×10^4 ($\pm 1.53 \times 10^4$)	4.17×10^6
[0$\bar{1}$0]	3.2±0.9	2.77×10^5 ($\pm 2.26 \times 10^4$)	7.84×10^4 ($\pm 2.07 \times 10^4$)	3.02×10^7
[1$\bar{1}$0]	1.5±0.7	5.21×10^5 ($\pm 5.29 \times 10^4$)		-
[$\bar{1}$10]	0.9±0.9	8.28×10^5 ($\pm 6.87 \times 10^4$)		-

8.2.3 Predicted Dissolution Rates

All necessary parameters to build the spiral dissolution model are either calculated, i.e. kink free energies and detachment rate constants, or can be easily extracted from the crystal structures, i.e. the angles between edges, step height etc. (see chapter 7 for model parameters). The last piece

needed to build the model is determining the shape of the spiral that is expected to be formed on the surface. In other words, it is necessary to determine the number of edges involved in the spiral formation, N in eq. 3.2,. This number is not necessarily the same as the total number of edges on a given face. Considering the periodic bond chains (PBCs), different spirals may be formed; six-fold or four-fold spirals depending on critical lengths and step velocities of all edges are considered. Since an edge will not start moving until it reaches its critical length, it could be expected that edges with lower critical lengths would start moving first and be the active edges. The edges with higher critical lengths may not even move nor be included in the calculation of rotation time, τ in this case (see eq. 3.2). However, since critical lengths and also step velocities depend on the undersaturation, the situation may differ at higher undersaturations. As the undersaturation increases, the critical lengths each edge needs to reach become smaller whereas the step velocities get higher. Then, the edges with lower critical lengths may have velocities that are too high and other edges could have critical lengths that are reachable, consequently edges with too high velocities may even disappear. Due to this fact, they again would not be included while calculating τ; hence four-fold spirals may be formed. The rotation direction of the spirals is another point to consider while forming the model. The dissolution rate resulting from the clockwise and anticlockwise spiral rotation directions can be different since the rotation direction will dictate which edge velocity is paired with the corresponding next edge's critical length as can be seen from eq. 3.4 [32]. This difference may even be more pronounced when the opposing edges do not have the same step velocities. Thus, the characteristic spiral rotation times, τ, for each direction and for all spirals, six-fold or four-fold, are considered and their dissolution rates are evaluated for both rotation directions. The smaller value of rotation time corresponds to larger values of dissolution rates (eq. 3.1) and as such the spiral shapes with lower τ would be more active. Hence, these active spirals determine the dissolution rates.

As explained in chapter 7, the net flux is calculated by eq. 3.15, the equation proposed for centrosymmetric molecules or dissolution units. Still, the effects of non-centrosymmetry of the molecule should be taken into account. But it is obvious that at the end only one detachment rate constant per edge enters into eq. 3.15. In principle, for reactions with more than one step such as in this case, rate-determining steps are mostly considered as slower reaction steps. However, in this case, there can only be a limited number of molecules between opposing kink sites. When the kink site molecule with higher detachment rate constant leaves its position, a slower detaching molecule would still be in its original kink position. More than one molecule can thus detach from faster detaching kink sites until one detachment event takes place on the slower detaching kink

site. This depends on the ratio between fast and slow detachment processes, called m here, and the average number of molecules between kink sites $\bar{n}$, that is the reciprocal of the kink density $\bar{n} = 1/\rho$, and indirectly it thus depends on the kink free energies (see eq. 3.7). In the case that the ratio m is very high and/or $\bar{n}$ is very small, the dissolution of the edge may even entirely proceed by the faster process with the faster detachment process dissolving also the last remaining kink site molecule that originally has a lower detachment constant. The faster of the two processes should then be considered for the flux calculations. When this is not the case, then the net flux has to be calculated explicitly. How this is done can be explained by the following example. Considering an edge with 'low enough' densities hence high $\bar{n}$, and with a ratio m of ten, the explicit calculation is performed as follows. While ten molecules detach with the faster detachment rate constant, one molecule detaches from the opposing kink site with the slower rate constant. Thus, in total eleven molecules are detached and as ten of them detached with the faster rate, the contribution of the fast process would be $10/11$ for the total rate. On the other hand, the slower process contribution would be $1/11$. By multiplying the contributions with the detachment rates respectively and summing the resulting values, it is possible to get a final detachment rate that considers both processes. Here, the 'low enough' definition would also depend on the ratio m of the two processes. If the ratio of the processes is ten but $\bar{n}$ is for instance seven, then the faster process would again dominate. In fact, for this case, half of the kink structure would be inactive and the calculation of the net flux can only consider the rate constant of the faster process. However, while calculating the overall step velocity, the result should be divided by two to account for the inactivity of half of the kink sites, or this division can be done while getting the reassessed rate constants as in Table 8.14. For aspirin (001) and specifically for the edges $[100]$ and $[\bar{1}00]$, the overall detachment rates are calculated explicitly, since the kink free energies are high enough that the average number of molecules between kink sites is higher than the ratio of the faster to the slower process. For edges $[010]$ and $[0\bar{1}0]$, this is not the case. The detachment rate is dominated by the faster process. Hence, the half of the faster detachment rate constant is taken. For the remaining two edges this is not necessary as only the detachment rate constants for one kink type along edges are evaluated due to them having similar solvent accessibility as depicted in Figure 8.25. All the parameters are displayed in Table 8.14.

The dissolution rates for different shapes of spirals are calculated by taking all these different kink types along the edges into account. They are displayed in Figure 8.26b. Only the results for active spirals in terms of rotation direction (see chapter 7) are displayed and the shapes for these spirals are shown in part a of Figure 8.26. The step height, h is taken as 6.0 Å, half of the lattice

parameter along the (001) direction. These spirals are named as follows to make the discussion easier: Spiral 1 depicted in green is a four-fold spiral along [010], $[\bar{1}10]$, $[0\bar{1}0]$ and $[1\bar{1}0]$ edges and rotating anti-clockwise. Spiral 2 depicted in red is a four-fold spiral along [010], $[\bar{1}00]$, $[0\bar{1}0]$ and [100] edges and rotating anti-clockwise. Spiral 3 depicted in blue is a four-fold spiral along $[\bar{1}10]$, [100], $[1\bar{1}0]$ and $[\bar{1}00]$ edges and rotating clockwise, and finally spiral 4 depicted in black is a six-fold clockwise rotating spiral along all available edges on the face. There are two available experimental data sets for the dissolution rate of the aspirin (001) face; one carried out in pure water by Kim *et al*. [148] and another one in 0.05 M hydrochloric acid by Danesh *et al.* [149]. However, for neither of the experiments, the undersaturation values are specified. Therefore, they are plotted as dashed horizontal lines in Figure 8.26. As the spiral model is assumed to be valid for low undersaturation values, the dissolution rates are calculated for undersaturation percentages up to 15%. The high error values are acknowledged, however it should be noted that the error calculation for the detachment rates is crude, based only on the average value of the detachment constants at different d values.

Table 8.14: Kink free energies, ΔG_{kink}; kink densities ρ; the average number of molecules between kink sites, $\bar{n}$; the ratio of faster to slower detachment rate constants m ; and the reassessed detachment rate constants with error values in parentheses for all PBC edges.

Edge direction	ΔG_{kink} (kJ/mol)	Kink density, ρ	$\bar{n}$	Ratio, m	Reassessed detachment rate constants (s^{-1})
[100]	13.7±0.6	0.0082	122.5	8.0	6.90×10^{5} ($\pm 7.58 \times 10^{4}$)
$[\bar{1}00]$	14.9±0.6	0.0051	197.6	3.5	2.83×10^{4} ($\pm 3.14 \times 10^{3}$)
[010]	4.2±0.5	0.2707	3.7	20.0	4.81×10^{5} ($\pm 4.73 \times 10^{4}$)
$[0\bar{1}0]$	3.2±0.9	0.3567	2.8	3.5	1.39×10^{5} ($\pm 1.04 \times 10^{4}$)
$[1\bar{1}0]$	1.5±0.7	0.5229	1.9	-	5.21×10^{5} ($\pm 5.29 \times 10^{4}$)
$[\bar{1}10]$	0.9±0.9	0.5823	1.7	-	8.28×10^{5} ($\pm 6.87 \times 10^{4}$)

From Figure 8.26, it is seen that spiral 1, a four-fold spiral, displays the fastest dissolution rate. This spiral is formed by the edges that have lower critical lengths and higher step velocities, mostly due to higher kink free densities. As explained before, spiral 1 would be the most active spiral. The exact comparison of predicted and experimental dissolution rates is not trivial due to the unknown degree of undersaturation. Moreover, it is noted that the solution concentration entering the spiral model refers to the local value in proximity of the interface instead of the bulk solution concentration. Measured values may typically include contributions from diffusive transport especially at higher undersaturation that are thus ignored in the model proposed here. In light of

these uncertainties, the calculated dissolution rates compare quite favorably, crossing both of the experimental data lines (spiral 1) or at least one of the lines (all spirals) at some undersaturation degrees.

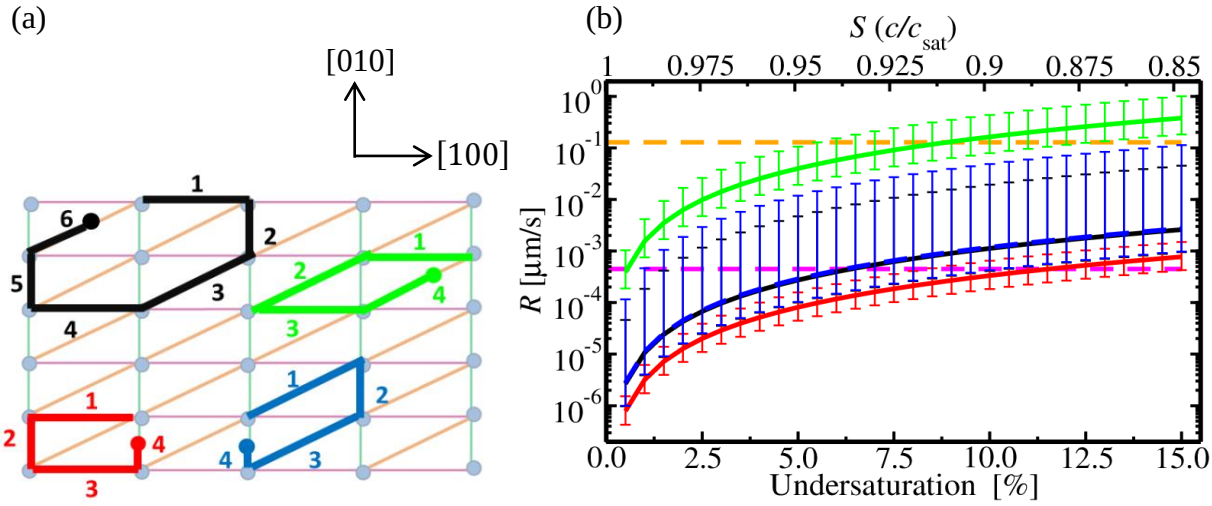

Figure 8.26: (a) Periodic bond chains (PBCs) for the aspirin(001) face together with the spiral shapes considered on this face; green spiral (spiral 1), red spiral (spiral 2), blue spiral (spiral 3) and black spiral (spiral 4). See text for the edges corresponding to the different spiral shapes. (b) Calculated dissolution rates for the different spiral shapes considered. The orange horizontal dashed line corresponds to the experimental value carried out in water by Kim *et al.* [148], while the pink horizontal dashed line is for the experimental value conducted in 0.05 M HCl by Danesh *et al.* [149]. The predicted dissolution rates are displayed by the colors that correspond to the colors of the spiral shapes shown in part (a).

Table 8.15: Parameters used to build spiral model 1 for the aspirin (001) face. Distance retraced due to the loss of a single dissolution unit $a_{p,i}$, distance between dissolution units along edges $a_{e,i}$, step edge velocity v_i, critical lengths l_i, angles between edges i and $i+1$, $\alpha_{i,i+1}$, rotation time τ, and the dissolution rate, R at $c/c_{\text{sat}} = 0.995$.

Edge	$a_{p,i}$ (Å)	$a_{e,i}$ (Å)	v_i (Å/s)*	l_i (Å)	$\alpha_{i,i+1}$ (°C)	τ (s)	R (μm/s)
Edge1=[010]	6.4	11.4	4179.97	7726.65	150.6	1.76	3.24×10^{-4}
Edge2=[$\bar{1}$10]	6.5	13.1	15784.24	1899.07	29.4		
Edge3=[0$\bar{1}$0]	6.4	11.4	1588.45	5886.97	150.6		
Edge4=[1$\bar{1}$0]	6.5	13.1	8919.43	3165.11	29.4		

* detachment rate constants in Table 8.14 are used.

Table 8.16: Same Table 8.15 for spiral 2.

Edge	$a_{p,i}$ (Å)	$a_{e,i}$ (Å)	v_i (Å/s)*	l_i (Å)	$\alpha_{i,i+1}$ (°C)	τ (s)	R (μm/s)
Edge1=[010]	6.4	11.4	4179.97	7726.65	90.0	756.00	7.94×10^{-7}
Edge2=[$\bar{1}$00]	11.4	6.4	8.18	15396.31	90.0		
Edge3=[0$\bar{1}$0]	6.4	11.4	1588.45	5886.97	90.0		
Edge4=[100]	11.4	6.4	321.70	14156.34	90.0		

* detachment rate constants in Table 8.14 are used.

Table 8.17: Same Table 8.15 for spiral 3.

Edge	$a_{p,i}$ (Å)	$a_{e,i}$ (Å)	v_i (Å/s)*	l_i (Å)	$\alpha_{i,i+1}$(°C)	τ (s)	R (μm/s)
Edge1=[$\bar{1}$10]	6.5	13.1	15784.24	1899.07	119.3	213.18	2.81×10^{-6}
Edge2=[100]	11.4	6.4	321.70	14156.34	60.7		
Edge3=[1$\bar{1}$0]	6.5	13.1	8919.43	3165.11	119.3		
Edge4=[$\bar{1}$00]	11.4	6.4	8.18	15396.31	60.7		

* detachment rate constants in Table 8.14 are used.

Table 8.18: Same Table 8.15 for spiral 4.

Edge	$a_{p,i}$ (Å)	$a_{e,i}$ (Å)	v_i (Å/s)	l_i (Å)	$\alpha_{i,i+1}$ (°C)	τ (s)	R (μm/s)
Edge1=[010]	6.4	11.4	4179.97	7726.65	90.0	224.54	2.67×10^{-7}
Edge2=[100]	11.4	6.4	321.70	14156.34	119.3		
Edge3=[1$\bar{1}$0]	6.5	13.1	8919.43	3165.11	150.7		
Edge4=[0$\bar{1}$0]	6.4	11.1	1588.45	5886.97	90.00		
Edge5=[$\bar{1}$00]	11.4	6.4	8.18	15396.31	119.3		
Edge6=[$\bar{1}$10]	6.5	13.1	15784.24	1899.07	150.7		

* detachment rate constants in Table 8.14 are used.

The whole number of simulations performed for predicting these dissolution rates is very high as can already be deduced by the number of hyperdynamics simulations (written under the figure captions for each kink site detachment rate calculation) and the number of λ points for the kink free energy calculation for each decoupling contribution determination. With nowadays computer power these numbers of simulations are tractable, especially when using supercomputers, though they are still computationally demanding, expensive and in a sense time-consuming. Hence, the

assessment of further approximations is appropriate, especially for fast screening approaches. Therefore, the edge specific kink free energies and detachment constants are re-analyzed and approximations to the method without significant losses in accuracy are considered.

It is shown that the edges with higher step velocities and lower critical lengths form the spiral shape that has the fastest dissolution rate, at least at low undersaturation. For a fast screening approach it would be enough to determine the dissolution rate for this spiral shape as it is the dominant spiral. As mostly the higher step velocities and specifically the kink free energies are the reason for the high dissolution rates, a first approximation that can be considered is based on the differences between kink free energies. It is also stressed that symmetry related edges have similar kink free energies, whereas differences in kink free energies are higher for symmetrically non-related edge directions, such as [100] and [010]. Hence, identifying the edges forming this most active spiral shape *before* starting to the simulations for free energy calculations would help to decrease the number of free energy calculations which are expensive to perform especially for even more complex systems (see alpha-lactose monohydrate case, chapter 9). A potential approach would be to use the crude methods such as the solid-state interaction energies applied by Doherty *et al.* (see ref. [28, 44, 150] and within) to predict the crystal morphology. Their method successfully predicts the relative step velocities of different edges and hence relative growth rates of different facets which actually depend on kink free energies (see eq. 3.5 and 3.6). It is essentially a rigid crystal approach meaning that dynamics are not considered, just the interaction energies between molecules along different PBCs directions are calculated. Hence, it is extremely inexpensive and just a single work station would be enough for these calculations. Using this approach, the trend in the kink free energies is reproduced correctly (only opposite edge site effects cannot be reproduced but this would not have a big consequence as symmetry related edges have similar kink free energies). The kink free energies calculated by this crude method are at least one order of magnitude higher than the kink free energies calculated by the method used in this work. The specific values are 52.9 kJ/mol for the [100] edge, 35.8 kJ/mol for the [010] edge and 30.2 kJ/mol for the $[1\bar{1}0]$ edge. As can be seen, the more pronounced difference between [100] and the other edges are well reproduced. Hence, using this simpler approach the edges forming the fastest dissolving spiral shape to be focused on in further calculations can be identified.

A second approximation arises from the reassessed dissolution rates. They are dominated by the faster of the detachment processes from opposite kink sites along the same edges. Hence, it is possible to ignore the kink type differences along the edges and only evaluate the detachment rate for the faster processes. This approximation would at first be trickier than the first one. One needs

to consider and check the structures of the kink types so that it would be possible to identify the faster of the detachment processes for the kink types along same edges. It has already been shown in this work that the faster processes for the kink types are the ones for which the solvent molecules can access the bonds easier (see Figure 8.25). Though it may not be so obvious for some crystal structures and would require detailed inspection, it will reduce the number of simulations, specifically the hyperdynamics simulations. Using this approximation, the dissolution rates are re-calculated for all spiral shapes. The dissolution rates calculated by this approximation are displayed in Figure 8.27b now omitting the error bars for clarity. It is nevertheless clear that with this approximation, the rates obtained are within the error range as compared to the dissolution rates obtained with the reassessed detachment rate constants, displayed in Figure 8.27a also without error bars for comparison.

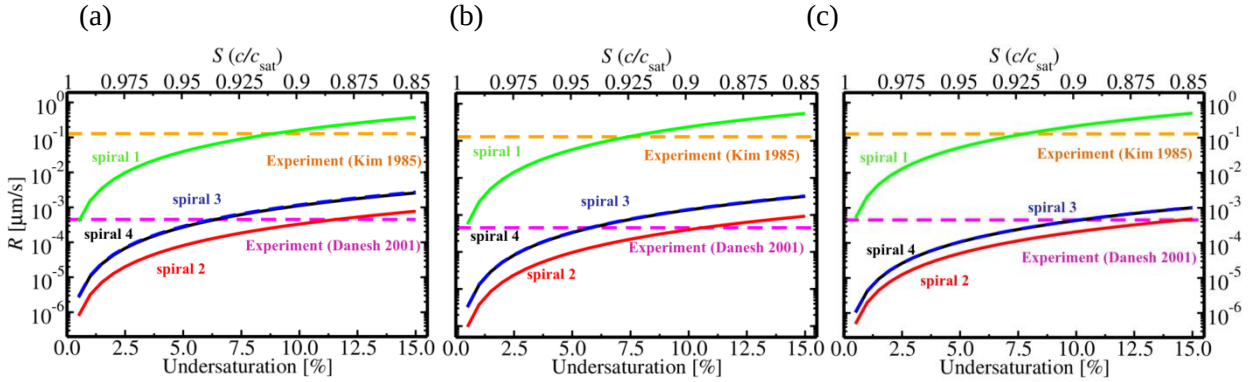

Figure 8.27: (a) The dissolution rates obtained by considering the different kink type effect along an edge for all spiral shapes considered for the aspirin (001) face. This plot is the same as Figure 8.26b without error bars. (b) The dissolution rates evaluated for a fast screening approach using only the second approximation, see text. (c) The dissolution rates calculated using both second and third approximations, see text.

The final approximation that has been considered in this work is to evaluate the dissolution rates without considering opposing edges, i.e. only considering the kink free energy and detachment rate constant of the [010] edge and using these values also for the opposing edge direction $[0\bar{1}0]$. If only the step edge directions with lower step velocities are used for this approximation, the rates calculated would still be in the error range; less than one order of magnitude change in the rates are observed. In most cases, the edges with lower step velocities correspond to higher kink free energies. For [010] and $[0\bar{1}0]$ edge directions, it is the other way around but kink free energies are still within the error ranges of each other. Hence, it is also possible to just consider the edge directions with higher $\Delta G_{\mathrm{kink},i}$ values and re-evaluate the dissolution rate constants. In addition, as in the case of the second approximation for kink types along edges, by checking the structural differences between opposing step edges, the edges with

higher kink free energies and/or lower step velocities can be identified before starting the simulations. For aspirin, this identification is again based on solvent accessibility of the edges. Edges that have the acetyl moieties pointing downward to the crystal surface at kink structures, i.e. $[\bar{1}00]$, $[0\bar{1}0]$ and $[1\bar{1}0]$, would be less accessible by solvent molecules as they expose more hydrophobic terminations and hence have lower step velocities. The re-evaluated dissolution rates are based on exclusively relying on the molecular data for these edges. The exception of $[010]$ and $[0\bar{1}0]$ edges, i.e. $[0\bar{1}0]$ edge having lower kink free energy and also lower step velocity is also considered. Using the molecular data for $[010]$ would still give dissolution rates at the same order of magnitude with the one calculated by explicitly considering slower step edges. However, in Figure 8.27c, only the re-evaluation of dissolution rates relying on the molecular data for the $[0\bar{1}0]$ edge is displayed, since this edge would be detected as the slower edge before starting the simulations. The comparison of dissolution rates calculated with and without these approximations is compiled in Figure 8.27.

Through these approximations, the computational cost can be decreased significantly. Hence, the corresponding process time per tested system would become insignificant on present-day computing platforms, while still accurate values compared to measured dissolution rates are obtained.

8.3 Summary of Chapter 8

In this chapter, using the micro-kinetic spiral model, dissolution rates for the aspirin (001) face are calculated for different spiral shapes. Specifically the two material parameters, kink free energies of formation and detachment rate constants from kink sites, are evaluated using accurate MD-based methods to be able to predict absolute dissolution rates. Though for the considered meta-stable kink structures, there is just one kink type meaning that the detachment/attachment of solute molecules always results in the same type of kink molecule as before, opposing kink sites along an edge reveal different terminations. Hence, they detach at different rates and this difference is traced back to solvent accessibility. The step velocities are calculated using the step velocity equation for the centrosymmetric dissolution unit case, i.e. eq. 3.15. The differences in opposing kink sites along an edge are still considered in the calculation of the step velocities by considering the ratio of faster to slower detachment rate constants and the average number of molecules between kink sites. All edges along periodic bond chains are considered for possible spiral shapes. This results in three four-fold and one six-fold spiral shapes, for which dissolution rates are

discussed. Due to the asymmetric nature of the opposing edges, spirals rotating in opposite directions (clockwise or anti-clockwise) do not have the same dissolution rates. The spiral shape and rotation with the highest dissolution rate will dominate the dissolution process. Since the undersaturation degrees at which two available experimental studies of aspirin (001) dissolution rates were carried are unknown, comparison to experiment is rather difficult. Still, the predicted dissolution rates compare quite favorably, around the same order of magnitude with the experimental value(s). Edges with lower kink free energies/higher kink densities are found to form spirals that are faster than other spirals. For a fast screening approach, such spirals can be identified on the basis of lower-level parameters derived e.g. via solid-state bond energies. Further approximations can lead to a significant additional reduction of the computational cost, such as a neglect of the differences of opposing kink sites and opposing step edges. The dissolution rates re-evaluated within these approximations are still within the same order of magnitude as before. All in all, the method established in this work appears thus robust and promising to predict the dissolution rates of an active pharmaceutical compound. In the next chapter this assessment is further validated by application to a second prototype compound, and in particular a compound used as excipient.

9 Calculation of the Dissolution Rate of Alpha-Lactose Monohydrate

In addition to the active ingredients, drug formulations also contain excipients that are defined as any component other than active principles but still intentionally included in the formulations. Though traditionally considered as inert substances included in formulations to provide the desired weight, volume, consistency or flavor [151], excipients have been shown to actually influence the quality and even bio-availability of the final formulation [7]. Hence, it is valuable to also consider the dissolution of excipients. For this purpose and also to validate the applicability of the computational approach developed, a widely used excipient, alpha-lactose monohydrate (alpha-LM), is chosen as a prototype excipient compound in this work. The fact that alpha-LM is also commonly used in food and confectionery industry to increase the storage life of products and that there is a single crystal experimental study of its dissolution [152] make it particularly appealing as a model compound to consider. The first part of this chapter will again focus on the development and testing of the force fields. The validated force field is thereafter employed to calculate the parameters needed to build the spiral dissolution model.

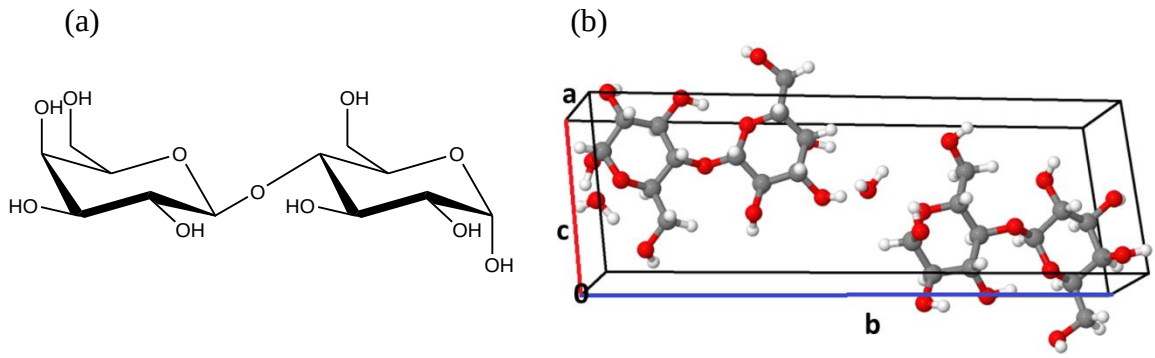

Figure 9.1: (a) Molecular structure of alpha-lactose. (b) The unit cell of alpha-lactose monohydrate taken from ref. [153].

9.1 Development and Testing of Force Fields

Alpha-LM is the crystalline solid form of the milk sugar, lactose, obtained when the crystallization from solution is performed at temperatures below 93.5°C. In its crystalline structure, each lactose molecule is associated with one molecule of water so that water molecules are also an integral part of the lattice structure. The water and lactose molecules crystallize into a monoclinic structure of

the $P2_1/c$ space group with each unit cell composed of two asymmetric units with two lactose and two water molecules [153]. The experimental X-ray unit cell structure of alpha-LM taken from literature is displayed in Figure 9.1 [153]. The force fields chosen for testing are the generalized AMBER force field (GAFF) [99] and the CHARMM force field. Contrary to the aspirin case (see chapter 8.1) the OPLS-AA force field is not considered for testing as it had not shown favorable results for aspirin and also for other crystals [116]. The parameters for GAFF are again obtained via the ACPYPE [124], and ANTECHAMBER [125] programs and the partial atomic charges are calculated using the R.E.D. server [126]. The SWISSPARAM server is used for the generation of the CHARMM force field parameters [127]. The TIP3P [129] water model is used for both solvent water molecules and the hydrate water molecules in the unit cell. First, the capabilities of the two force fields to correctly reproduce lattice parameters are tested. An annealing procedure similar to the one applied to theaspirin lattice structure is again applied for alpha-LM. The experimental structure is optimized by slowly cooling the system from 100 to 10 K in a 70 ps MD simulation while applying a Berendsen barostat of zero pressure. Again the barostat was coupled to the lengths of the cell vectors as well as to the β-angle to preserve the monoclinic crystal system. The remaining two angles were fixed [116]. All atomic positions are relaxed using the conjugate gradient method with a maximum force threshold of 0.01eV/Å. The unit cell parameters can be found in Table 9.1 for both experimental and optimized cell structures. For both force fields, the agreement with experimental structure is quite good with only 1-2% deviations in the parameters.

Table 9.1: Comparison of experimental and calculated lattice parameters for alpha-LM

	Lattice Constants (Å)			Cell angle (degrees)
	A	b	C	β
Exp.[a]	4.78	21.54	7.76	105.91
GAFF	4.75	21.78	7.64	104.24
CHARMM	4.78	21.83	7.88	104.79

[a] Ref. [153]

As done for aspirin, a further testing of the force fields is again performed against QM calculated intermolecular interaction energies. However, due to the crystal water in the unit cell structure, it would not be sufficient to only consider the interactions between lactose molecules. Instead, also the interactions between lactose and water molecules are evaluated. The interaction energies are again calculated by displacing one of the two molecules rigidly in the direction connecting their centers of mass in the starting structure and calculating the energy relative to the

energy of the two isolated molecules [116] (see chapter 8.1 also). As before the all-electron DFT code FHI-aims [133-134] is used to perform the QM calculations. Electronic exchange and correlation was treated at the GGA functional level using the PBE functional [83] and the vdW dispersion correction approach due to Tkatchenko and Scheffler [135]. For lactose-water interaction energies the B3LYP hybrid functional [154], is also employed for further testing. The "tier2" basis set with "tight" settings is employed. The interaction energies versus the relative distance to the starting structure are displayed in Figure 9.2 for the lactose-lactose interaction and in Figure 9.3 for the lactose-water interaction. It is seen that both GAFF and CHARMM give good agreement compared to DFT+vdW interaction energies. The shape of the lactose-water plot is almost exactly the same for DFT+vdW and GAFF, whereas the lactose-lactose interactions are also fairly well reproduced, especially at high displacement values. The agreement of the CHARMM data with the DFT+vdW energies is also reasonably good. However, CHARMM underestimates the binding energies slightly. Similar to the situation for aspirin, GAFF is correspondingly chosen for the alpha-LM dissolution calculations.

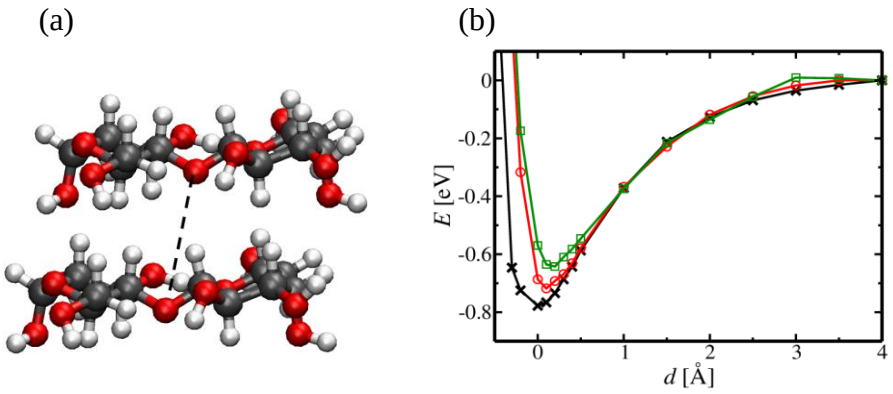

Figure 9.2: (a) Lactose-lactose interaction considered for the displacement *d*, with the center of displacement represented by a dashed line between the molecules. (b) Calculated interaction energies versus displacement *d*. Reference values calculated from DFT+vdW (black dashed lines, X) and resulting energies from the classical force fields GAFF (red circles), CHARMM (green squares).

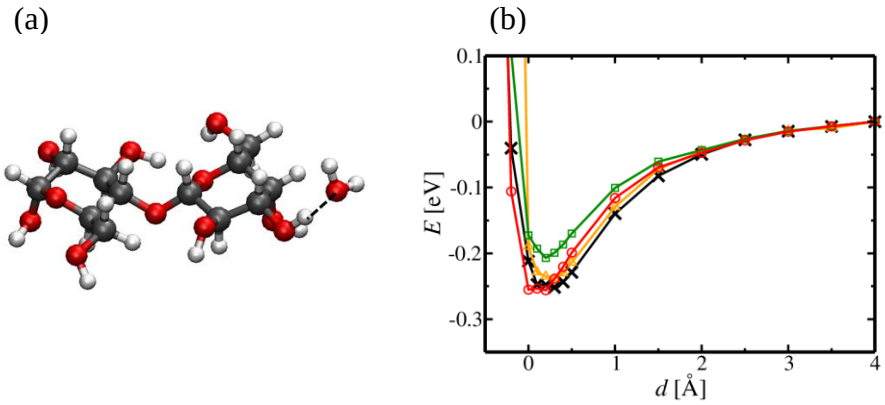

Figure 9.3: (a) Lactose-water interaction considered for the displacement d, with the center of displacement represented by a dashed line between the molecules. (b) Calculated interaction energies versus displacement d. Reference values calculated from DFT(PBE)+vdW (black dashed lines, X), DFT(B3LYP)+vdW (orange triangles) and resulting energies from the classical force fields GAFF (red circles), CHARMM (green squares).

9.2 Spiral Dissolution Model for the Alpha-Lactose Monohydrate (010) Face

In this work, the (010) face of alpha-LM is chosen since it is one of the dominant facets grown in aqueous solutions and also the slowest dissolving face [152, 155]. Lactose molecules in the unit cell are elongated along the [010] direction, so that there is only one growth unit on this (010) face. Moreover, the two lactose molecules along the surface normal have the same surface terminations, i.e., the step height can be considered to be one layer height. Again, based on the PBCs (see chapter 8.2), the strong intermolecular interactions are identified. There are three different straight chain directions: [100], [001] and [101]. As alpha-LM is also a low-symmetry molecule, opposite edge directions expose different bond terminations. Therefore in total, again six step edge directions have to be considered for the spiral model.

Though alpha-LM is also a non-centrosymmetric molecule (see chapter 3) and the unit cell contains two asymmetric units, on the (010) face only one of these units are exposed. In other words, after the detachment of one kink site molecule the second molecule will be the same even without considering dimeric units as in the case of aspirin (see chapter 8.2). There are obviously also crystal water molecules to consider as they are integral parts of the crystal structure. However, their attachment/detachment will generally be very fast. The interaction energy for the lactose-water complex is relatively low compared to the lactose-lactose complex (see Figure 9.2 and 9.3). Hence, the detachment of lactose molecules from any kink site will be much slower than the release of the crystal water molecules from their corresponding lactose molecules. Additionally, in a water environment the mole fraction of water molecules is obviously much higher, so that the

incorporation of water molecules into their respective kink sites should also be an undeniably faster process compared to the incorporation of lactose molecules. For these reasons, even though lactose-water complexes are considered as dissolution units to build the model, in practice only kink free energies and detachment rate calculations for lactose molecules are required. In the corresponding MD calculations, the effect of opposite kink sites along the same step edges are again taken into account by explicitly considering the detachment of both kink site molecules.

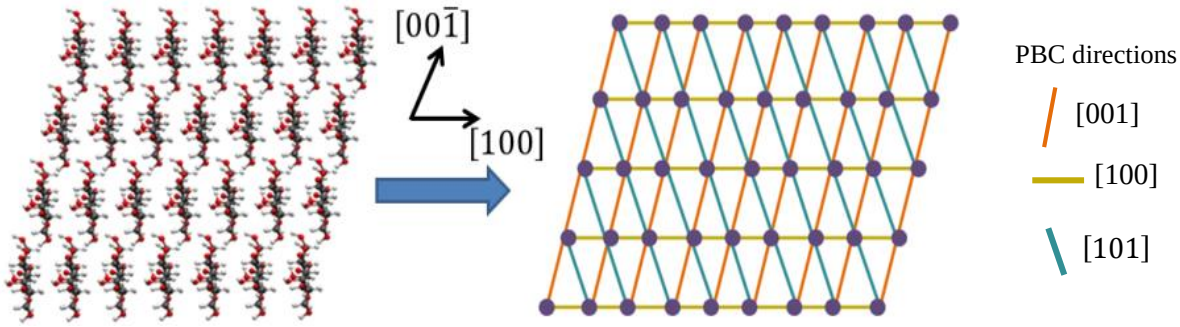

Figure 9.4: Top view and bonding structure of the (010) face of alpha-LM. On the right side, the PBCs along the shown directions are displayed.

9.2.1 Kink Free Energies of Formation

Kink sites are created along the six step edges to calculate the kink free energies for each edge. Again the thermodynamic cycle explained in chapter 7.1 is employed. First, the simulation details are summarized. Then, the results will be presented along with a brief discussion of all kink free energy contributions.

Simulation Details and Model Systems

The generalized AMBER force field (GAFF) with the TIP3P water model MD simulations are performed within the GROMACS package. As before, the simulation cells are split into two separate parts to decrease the computational cost. This system separation is depicted in Figure 9.5 for the kinked step edge direction [00$\bar{1}$] as example. The surface models are built with a crystal slab thickness amounting to two unit cells, i.e. four molecular layers. The vacuum gap between opposite slab surfaces are filled with pre-equilibrated water molecules to prepare the interface systems. The water layer thickness is assessed by monitoring the water density profile so that the chosen thickness is sufficiently large to retain bulk-like water properties.

The other simulation details are similar to the ones applied for aspirin (chapter 8.2.1). All the systems considered are first run at constant temperature, 300 K, and volume for 100 ps with position restraints employed on all heavy atoms of the lactose molecules to allow water molecules as well as surface atoms to equilibrate. Then, the pressure of the system along the surface normal is adjusted to 1 atm by 3 ns long simulations without any position restraints. The box vectors along the lateral directions are fixed, while along the surface normal direction, i.e. water layer direction, the final height is determined after the pressure adjustment. An additional 2 ns simulation at constant volume and temperature is run before starting the thermodynamic cycle to determine the average positions of kink molecules. During equilibration, the temperature is controlled through a Nose-Hoover [142] thermostat by coupling the kinetic energy at a relaxation time of 0.5 ps and the barostat employed is Parrinello-Rahman [78]. The production runs are 3 ns long without pressure coupling but at constant temperature and the volume determined before. The atoms at the bottom ring of the bottom molecular layer, i.e. the glycose moiety atoms, are fixed at their positions to mimic a bulk crystal substrate during the last equilibration under constant volume and pressure and production simulations. Again the new thermostat suggested by Bussi *et al*. [73] that rescales the velocities for stochastic processes is employed during the decoupling processes. Lennard-Jones potentials and the short-range Columbic potentials are truncated at a cutoff distance of 1.2 nm instead of 1.0 nm as in the aspirin case; long-range interactions are calculated via the particle-mesh-Ewald (pme) [143-144] method. The time step is 0.001 fs and the long-range interactions are updated at every 10 steps.

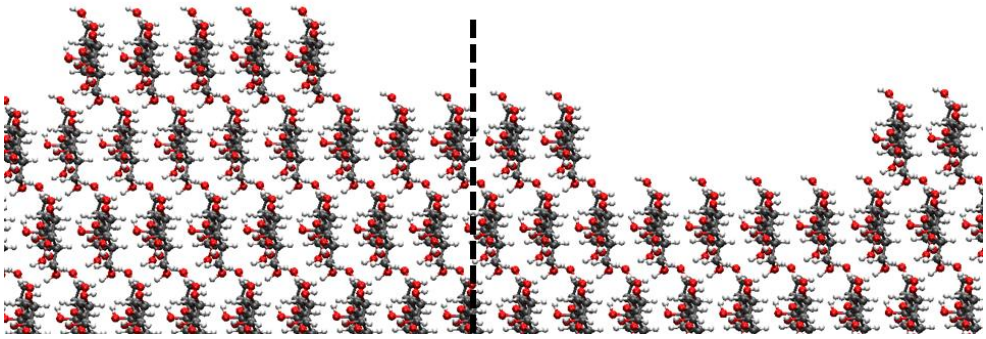

Figure 9.5: Top view of the $[00\bar{1}]$ step edge direction for the alpha-LM (010) face with dashed lines indicating the system separation. For clarity only the molecules at the upper terrace are shown.

The electrostatic and Lennard-Jones interactions are again decoupled separately as suggested for complex systems [119] and the number of λ points used for each decoupling are asymmetric (chapter 7.1 and 8.2.1). For the decoupling of electrostatic interactions, intermediate λ_{el} points

between the 0 and 1 states are placed at a spacing of 0.1 with an additional point at $\lambda_{\mathrm{el}} = 0.05$ to improve the sampling of the steep initial free energy gradient. Additional three to four λ_{el} points are added between states 0.6 to 1 due to the more complex nature of lactose molecules as compared to aspirin. Since the decoupling of Lennard-Jones potentials is trickier and affected by solvent penetration most of the time, even more λ_{LJ} points are employed. Specifically between 0.6 to 1 states, at least six more intermediate values are placed to increase the overlap between states. The soft-core parameter α (see eq. 7.3) used during the Lennard-Jones interaction decoupling is set at the GROMACS default value of 0.3 [43].

The restraining free energy contributions are again evaluated by thermodynamic integration. Specifically, integration is carried out numerically by using the trapezoidal rule, based on six k-values for eq. 7.4: 0, 50, 100, 250, 500, and 1000 kJ/mole/nm^2. The final value of the spring constant is again $k_{\max}$ = 1000.0 kJ/mole/nm^2, since this yields a balance between preserving the molecular arrangement and a reasonably low restraint free energy with associated low error values.

An exception treatment is employed for the kink free energy simulations at the $[001]$ step edge due to the enhanced flexibility of the kink site molecules along this edge. Hence, achieving a sampling with average root-mean square ξ deviations to calculate the restraining potential using eq. 7.4 was challenging. Mostly the atoms belonging to the galactose moiety in the upper part of the molecule exhibit more flexibility. Therefore, for the calculation of this free energy contribution, the atoms in kink site molecules are considered separately depending on them belonging to 'upper' and 'lower' parts of the molecule as shown in Figure 9.6. Though this separation is only shown for a single kink site molecule in the figure, it should be noted that the same atoms on the other molecules in the kink structure are also subject to this separation. The green colored atoms in the figure are considered to belong to the upper part of the molecule and their flexibilities are higher than the ones for the lower part atoms. Therefore, first the potential energy of restraining the lower part atoms, $\Delta G_{\mathrm{kink}}^{\mathrm{bottom}}$, is evaluated by integrating at different spring constants as already explained in chapter 7.1. In order to obtain the remaining part of the free energy contribution to restrain the whole molecules, an additional simulation is performed with restrain potentials fully switched on the lower part of the molecules and to enhance the sampling, the well-tempered metadynamics technique is applied [122]. This technique is an improvement over metadynamics such that the obtained estimate of the free energy value converges to the exact result in the long time limit [114] (see chapter 7.1). Here, as collective variables, the upper part (green colored) atoms of all kink site molecules are chosen and the targeted mean square deviation (MSD), ξ^2, option is used for the PLUMED [146] plugin within the GROMACS package. The

perfect kink site molecules are considered as reference positions. The Gaussian hills of width 0.002 nm and height 0.8 kJ/mole are deposited with a frequency of 0.5 ps^{-1}. The bias factor for the well-tempered simulations is 5 with 300 K as the simulation temperature. For the targeted-MSD type collective variable, the alignments to the reference structure are prevented for the center of mass and the rotation of atoms. A converged free energy profile, $G(\xi^2)$ is obtained and can be calculated by numerically integrating over the ξ^2 probability distribution [119]

$$\exp\left(-\beta\Delta G_{\text{kink}}^{\text{upper}}\right) = \frac{\int_0^{\xi^2_{\max}} d\xi^2 \exp\left(-0.5\beta k^{\text{upper}} N_{\text{atoms}}^{\text{upper}} \xi^2\right)\exp(-\beta G(\xi^2))}{\int_0^{\xi^2_{\max}} d\xi^2 \exp(-\beta G(\xi^2))}, \qquad 9.1$$

with $\beta = 1/k_B T$. The spring constant k^{upper} is taken as 1000.0 kJ/mole/nm^2, and $N_{\text{atoms}}^{\text{upper}}$ is 65 in total. This calculated free energy contribution is added to the value obtained by restraining the lower part of the kink site molecules. This way, the total restraining energy contribution is obtained for this special case in a more accurate way than it can be calculated without separating the atoms into parts. The separated kink free energy contributions $\Delta G_{\text{kink}}^{\text{bottom}}$ and $\Delta G_{\text{kink}}^{\text{upper}}$ are also included in Table 9.2 as a footnote.

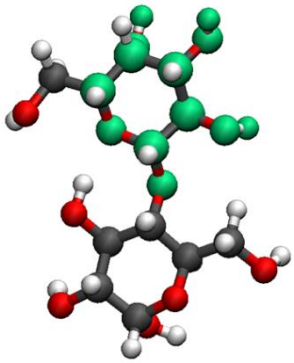

Figure 9.6: Separation of atoms in lactose molecules in the kink structures at the [001] step edge. Atoms displayed by green colors are considered as upper part atoms, all remaining atoms belong to the lower part.

Results and Discussion

Table 9.2 displays each free energy contribution along with the statistical errors that are calculated via error propagation for the six step directions. It can be deduced from the results that individual decoupling free energy contributions are really high, whereas the restraining contributions are comparably small. This shows that most of the cost for the kink site creations arises due to the decoupling of the interactions. However, the major part of these high free energy contributions cancels out in the free energy differences. The statistical errors are again calculated by block

averaging for both decoupling and restraining free energy contributions (see chapter 8.2.1). Based on the results for the kink free energies, it can be said that kink free energies for opposing step directions, like for [100] and [$\bar{1}$00], are again only slightly different. In contrast, the differences in kink free energies are more pronounced for edges that are not symmetrically related. The most stable edge terminations are the symmetry-related [001] and [00$\bar{1}$]. All the other edge directions have smaller kink free energy values which are close to each other. They lead to high kink densities which shows that these edges may not be stable as straight edges; they are probably kinked edges especially when compared to [001] and [00$\bar{1}$] edges.

Table 9.2: Free energy components during the calculation of kink free energies for kinks within step edges of several orientations, as well as the final kink free energies per kink site for the alpha-LM (010) face.

Step direction	$\Delta G_{\text{step}}^{\text{rest}}$ (kJ/mol)	$\Delta G_{\text{step}}^{\text{dec}}$ (kJ/mol)	$\Delta G_{\text{kink}}^{\text{rest}}$ (kJ/mol)	$\Delta G_{\text{kink}}^{\text{dec}}$ (kJ/mol)	ΔG_{kink} (kJ/mol)
[100]	136.35 ± 1.68	1264.03 ± 2.18	125.14 ± 1.44	1265.01 ± 0.77	2.56 ± 0.80
[$\bar{1}$00]	161.27 ± 2.24	1219.46 ± 2.44	207.35 ± 2.60	1162.76 ± 3.89	2.66 ± 1.43
[001]	119.64 ± 1.06	1286.96 ± 2.89	181.25 ± 0.87*	1186.12 ± 1.63	9.80 ± 0.90
[00$\bar{1}$]	155.30 ± 1.44	1361.19 ± 1.59	165.76 ± 2.80	1320.00 ± 2.79	7.68 ± 1.12
[$\bar{1}$01]	258.35 ± 2.21	1152.83 ± 1.48	200.36 ± 2.13	1200.51 ± 1.22	2.58 ± 0.90
[10$\bar{1}$]	167.65 ± 1.31	1240.45 ± 1.86	170.49 ± 1.69	1221.85 ± 2.08	3.94 ± 0.88

*$\Delta G_{\text{kink}}^{\text{bottom}}$ = 105.24 kJ/mol and $\Delta G_{\text{kink}}^{\text{upper}}$ = 76.01 kJ/mol. See text for explanation.

9.2.2 Rate Constants of Detachment Processes from Kink Sites

Again, the two-step algorithm explained in chapter 7.2 is applied for the calculation of the detachment rate constants from kink sites for all relevant step edges. For these calculations, the prepared simulation cells contain only positive kinked step edge structures, i.e. the left part of the simulation box shown in Figure 9.5. Though there is only one dissolution unit, the lactose-water complexes per edge for this face cause opposite kink site molecules along the same edges to expose different terminations as explained before. The kink site molecule type before and after detachment/attachment is still the same and the same bonds are broken/formed in both kink site molecule cases. Still, these kink sites may have different detachment rate constants due to the environment they are in. Hence, again they are labelled as kink type A or type B, for simplicity. The detachment rate constants for both kink types are evaluated by employing the two-step algorithm. The net flux is calculated using the equation for centrosymmetric cases.

Simulation Details and Model Systems

The preparation of the simulation cells is the same as the preparation of the aspirin simulation cells for the detachment rate calculations, chapter 8.2.2. The slab thickness amounts to four layers and all surface models are built by the parameterized GAFF/TIP3P combination. The water density profile is again checked so that bulk-like water properties are observed between opposite surfaces. The simulation cells used are slightly smaller than the ones used for free energy calculations. The slab thickness is the same but in the lateral directions, the cell dimensions are smaller with fewer rows between step edges. The effect of using smaller cells is explicitly tested to avoid finite size effects in the bias potential construction, i.e. metadynamics simulations are performed comparing the bias potential value before the first escape events for both small and large systems. Since the results show almost equal bias potential values, the smaller cells are used for the hyperdynamics simulations to increase the simulation times. Before starting the metadynamics simulations, the system is again equilibrated in the *NVT* ensemble with all atoms of lactose molecule restrained for around 100-200 ps. The equilibrium box height is determined as it is done for the free energy calculations. For both metadynamics and hyperdynamics simulations, the box height is fixed and the bottom cyclic part atoms of the bottom layer molecules, i.e. the glucose moieties of the bottom layer molecules, are fixed in their position, again to mimic a bulk crystal.

All the simulations have been implemented in the PLUMED plugin interfaced with GROMACS [146]. Instead of using the center-of-mass (COM) vector of the entire kink site molecule, only the COM of the glucose moiety of the kink site molecules is considered as collective variable (CV). This is again due to the more flexible nature of the galactose moiety atoms belonging to the upper part of the molecule. The push and equilibrate approach employed previously for aspirin is again used for the bias potential construction [117] (and see chapter 8.2.2). The deposition of hills for 30 ps is followed by an equilibration phase of 100 ps. The deposition frequency of hills during push sequences is always 1 ps^{-1}. The width of Gaussian hills deposited is 0.03 nm and the height of the hills is 0.3 kJ/mole. In contrast to the aspirin case, using the same width and height of hills for all considered kink molecules proved sufficient to determine the truncation point. The hyperdynamics simulations are initiated for different conformations obtained from snapshots from long equilibration runs at every 50 ps. They are all assigned new velocities to further de-correlate the snapshots and run for 15 ns to obtain long biased escape times. During the metadynamics and hyperdynamics simulations, the temperature is again controlled through a Nose-Hoover thermostat with 0.5 ps relaxation time [142]. Here, Lennard-Jones potentials and the short-range Columbic potentials are truncated at a cutoff distance of 1.2 nm; long-range

interactions are calculated via the particle-mesh-Ewald (pme) [143-144] method. The time step is 0.001 fs and the long-range interactions are again updated at every 10 steps.

Results and Discussion

The following figures display top views of the alpha-LM step edge structures with the dissolution units, i.e. the lactose-water complex, marked with solid lines for each edge direction. The trajectories of COM-components observed during the metadynamics simulations and the reweighted free energy profile, as well as the bias potential values encountered during the hyperdynamics simulations can be found in these figures for each edge direction. Again it is observed that detachment of kink type B (marked by blue solid lines) is slower than the detachment of kink type A molecules (marked by green solid lines). This difference especially for $[100]$, $[\bar{1}00]$, $[\bar{1}01]$ and $[\bar{1}01]$ edges, can be explained by solvent access to intermolecular bonds between kink site molecules and their neighbors, see Figure 9.25. For edges $[001]$ and $[00\bar{1}]$, the exposure to water is the same on both kink sites along the edges, hence the rate constants are almost the same for these edges. All detachment rate results are finally compiled in Table 9.15.

The trajectories obtained from the hyperdynamics simulations are first analyzed to establish the free energy profiles. The dimensionless order parameter d is defined as in eq. 8.1. In the alpha-LM case, always 64 independent hyperdynamics simulations are performed. The evaluation of rate constants is based on Transition State Theory (TST) [147] and calculated by eq. 8.2. The time-dependent transition constant κ is evaluated by eq. 8.3 to also account for dynamical re-crossings of escape trajectories after passing the barrier. The errors in the detachment rate constants can be roughly estimated from the deviations of the mean value which can be taken by averaging the detachment rate constants for different choices of transition barrier. Three such locations are considered; d=15, 17, 19.

1. Step edge [100]

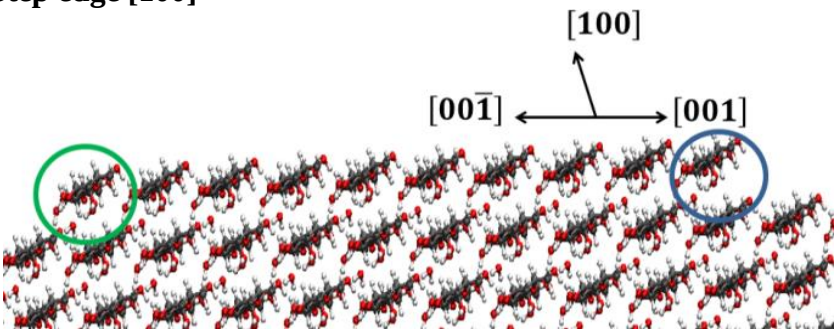

Figure 9.7: Top view of the [100] step edge of the alpha-LM (010) face. For clarity only the terrace top layer above the step edge is displayed and (green and blue) solid lines highlight the considered dimeric dissolution unit at the two kink sites shown. C, O, H atoms are shown as gray, red and white spheres, respectively. The detachments of molecules shown with bold spheres are considered as rate determining.

a) Kink Type A

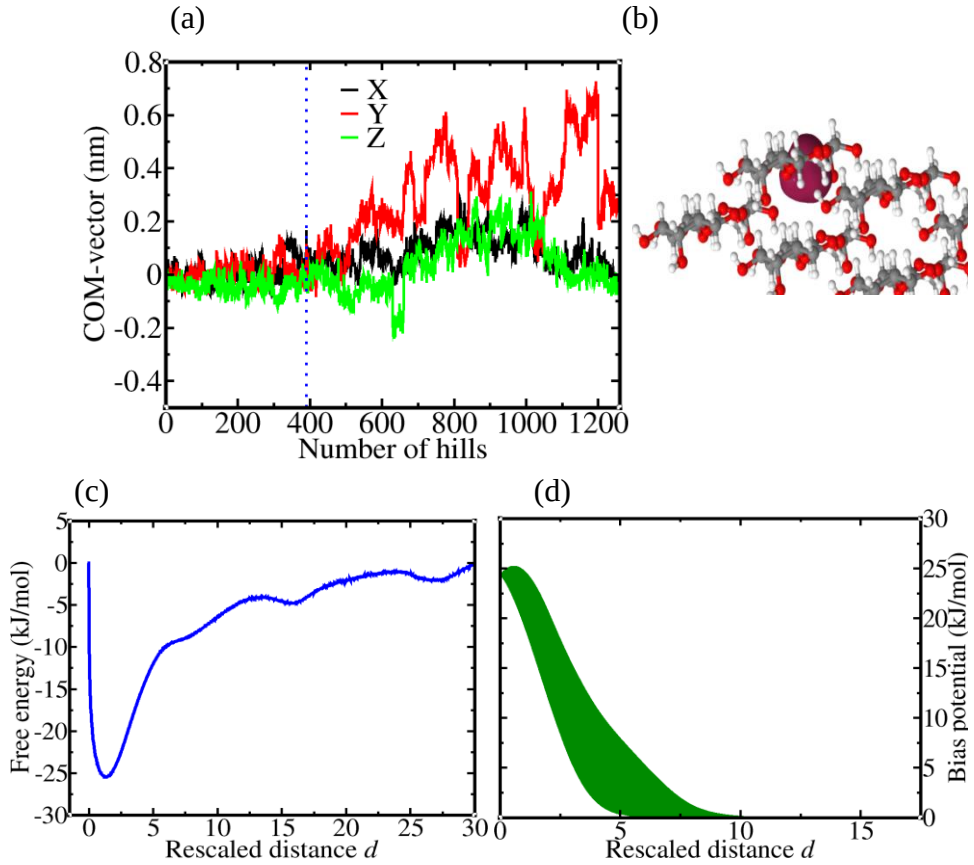

Figure 9.8: (a) Trajectory of COM components for the kink type A molecule along edge [100] of the alpha-LM (010) face. The truncation point at 390 hills is marked by dotted lines. (b), (c) and (d) are the isosurface of fixed bias potential, reweighted free energy profile of the hyperdynamics simulations, and the bias potential encountered during the hyperdynamics simulations, respectively.

Table 9.3: TST rate constants, transmission coefficients, and reactive-flux rate constants, boost factor, number of detachment events, and barrier-to-well height for different choices of the transition barrier for the kink type A molecule along edge [100] of the alpha-LM (010) face.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	2.09 x 10^7	0.02	4.47 x 10^5	2.24	129.66	1172	20.93
d = 17	1.09 x 10^7	0.05	5.55 x 10^5	1.80	125.41	428	21.51
d = 19	4.98 x 10^6	0.14	7.04 x 10^5	1.42	123.19	159	23.02

b) Kink Type B

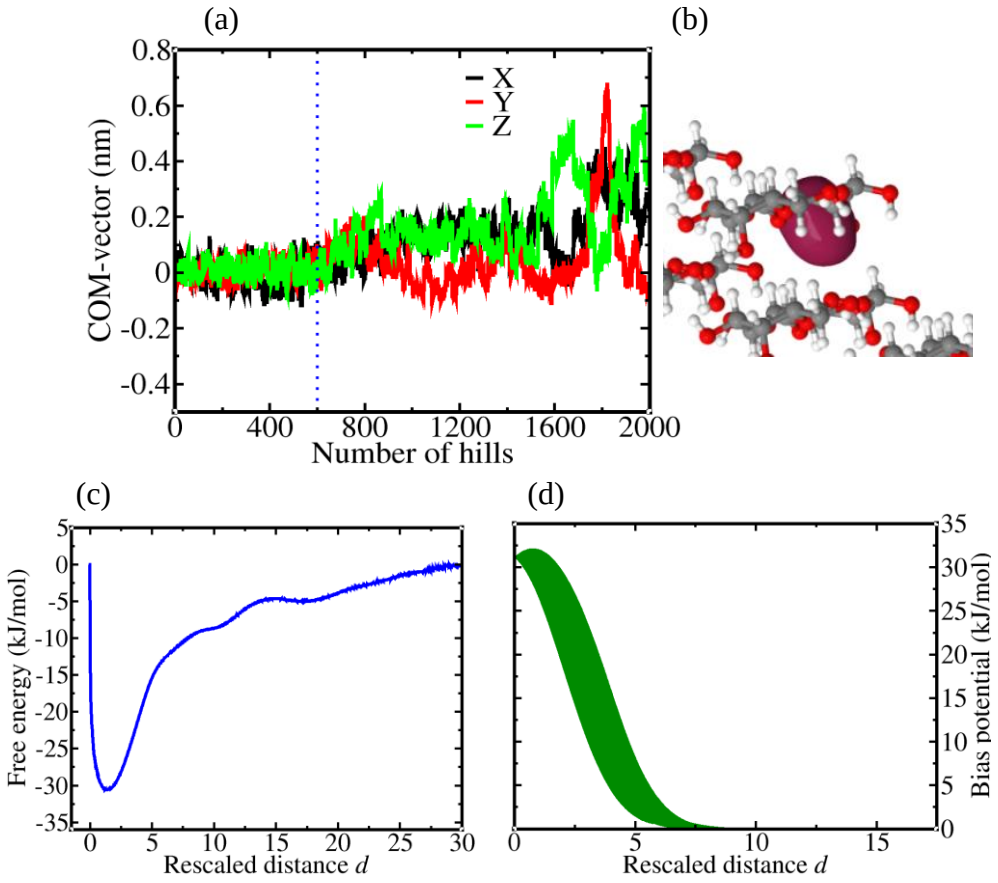

Figure 9.9: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 600 hills and marked by dotted lines.

Table 9.4: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	3.27 x 10^6	0.02	6.70 x 10^4	14.92	619.71	940	25.98
d = 17	1.44 x 10^6	0.05	6.51 x 10^4	15.35	606.59	290	25.62
d = 19	6.85 x 10^6	0.08	5.35 x 10^4	18.67	594.76	126	26.11

2. Step edge [$\bar{1}$00]

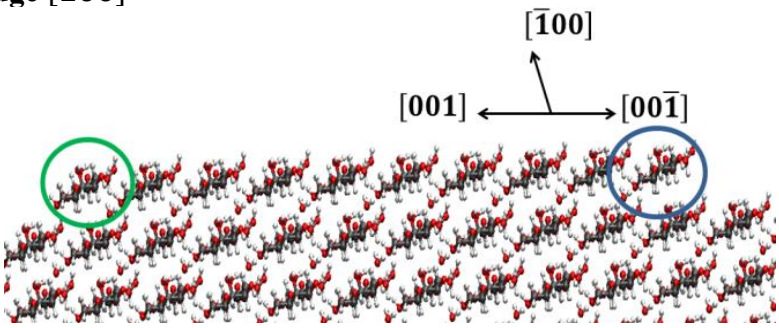

Figure 9.10: Same as Figure 9.7. Top view of the [$\bar{1}$00] edge of the alpha-LM (010) face with the considered dissolution units marked by solid lines.

a) Kink Type A

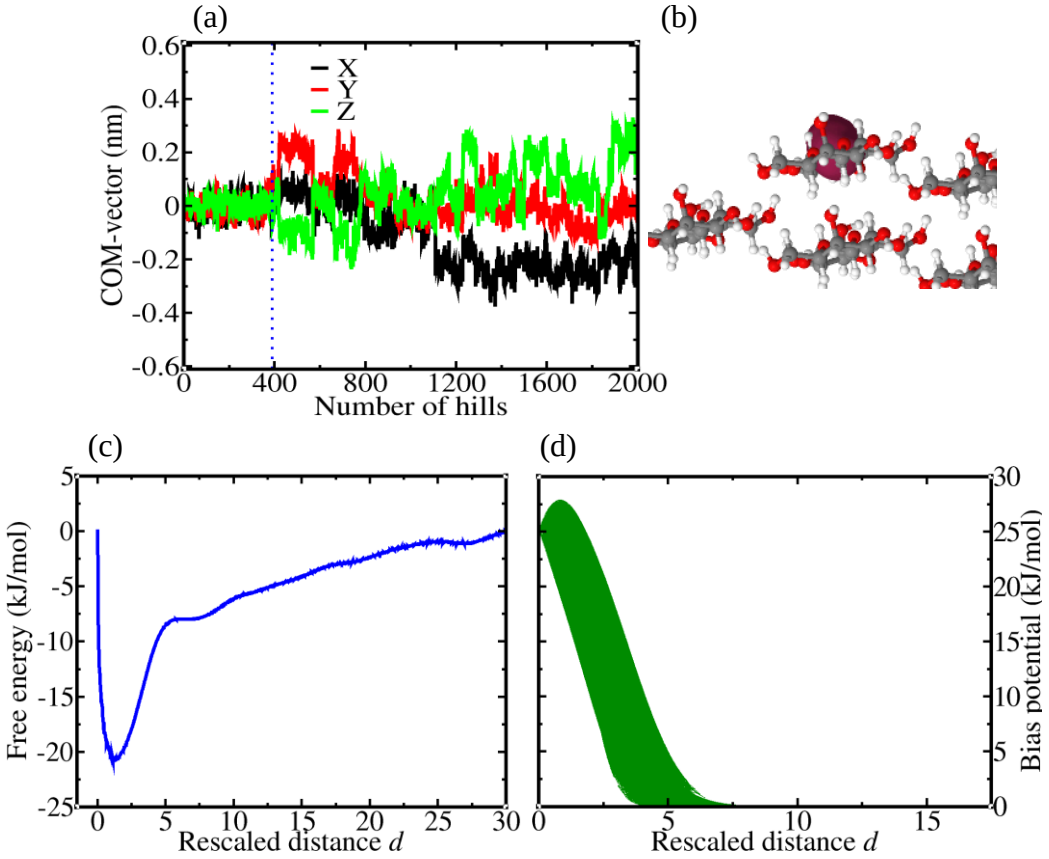

Figure 9.11: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 390 hills and marked by dotted lines.

Table 9.5: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	4.57 x 10^7	0.05	2.41 x 10^6	0.41	32.65	597	16.60
d = 17	2.05 x 10^7	0.12	2.48 x 10^6	0.40	31.44	234	17.56
d = 19	1.37 x 10^7	0.21	2.86 x 10^6	0.35	30.61	130	17.98

b) Kink Type B

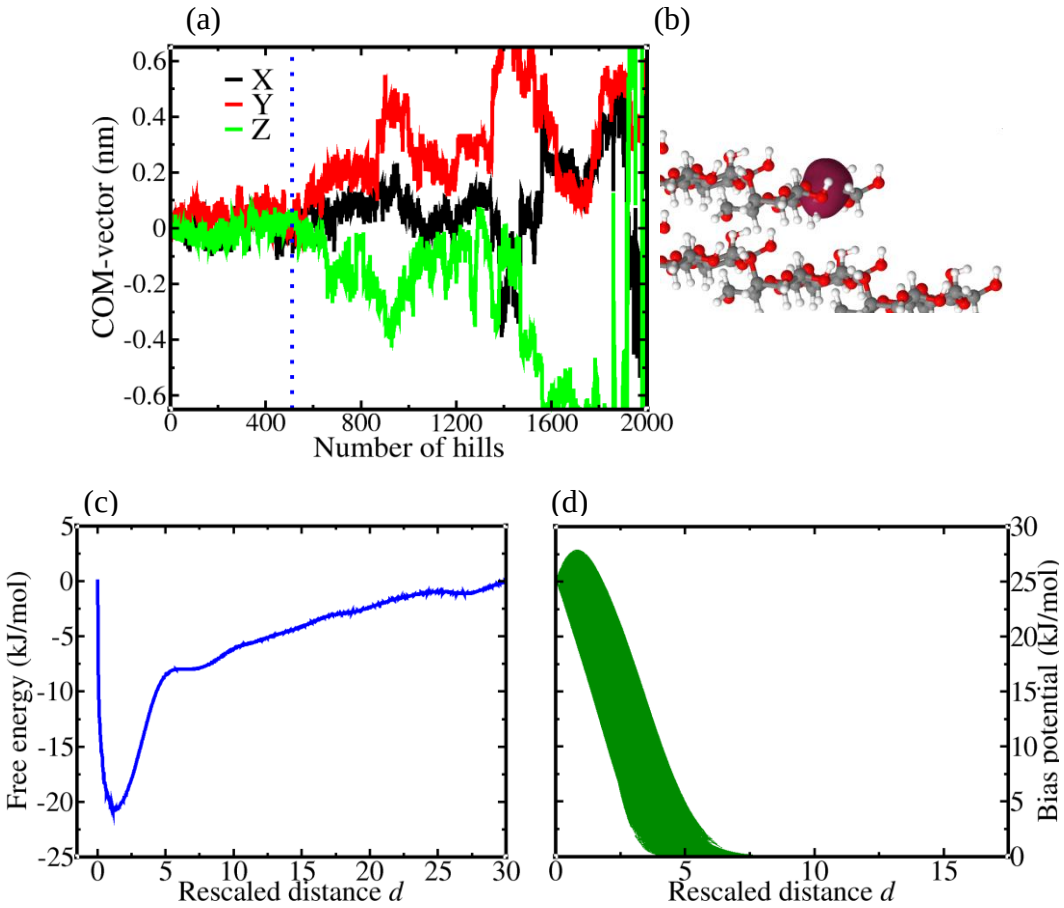

Figure 9.12: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 510 hills and marked by dotted lines.

Table 9.6: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	2.10 x 10^7	0.02	3.45 x 10^5	2.90	183.59	2207	21.71
d = 17	5.66 x 10^6	0.05	2.93 x 10^5	3.41	179.26	575	22.96
d = 19	2.46 x 10^6	0.12	3.04 x 10^5	3.29	175.67	207	22.50

3. Step edge [001]

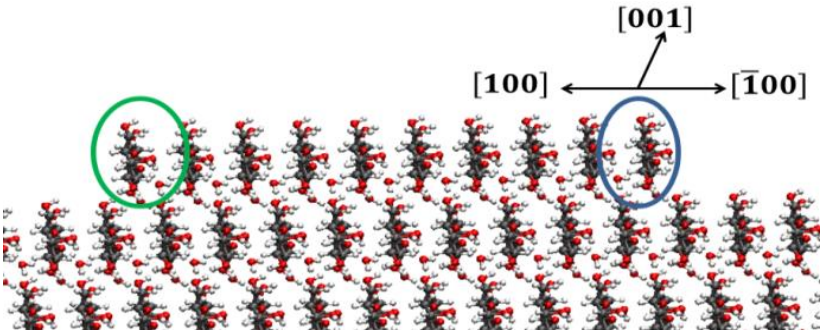

Figure 9.13: Same as Figure 9.7. Top view of the [001] edge of the alpha-LM (010) face with the considered dissolution units marked by solid lines.

a) Kink Type A

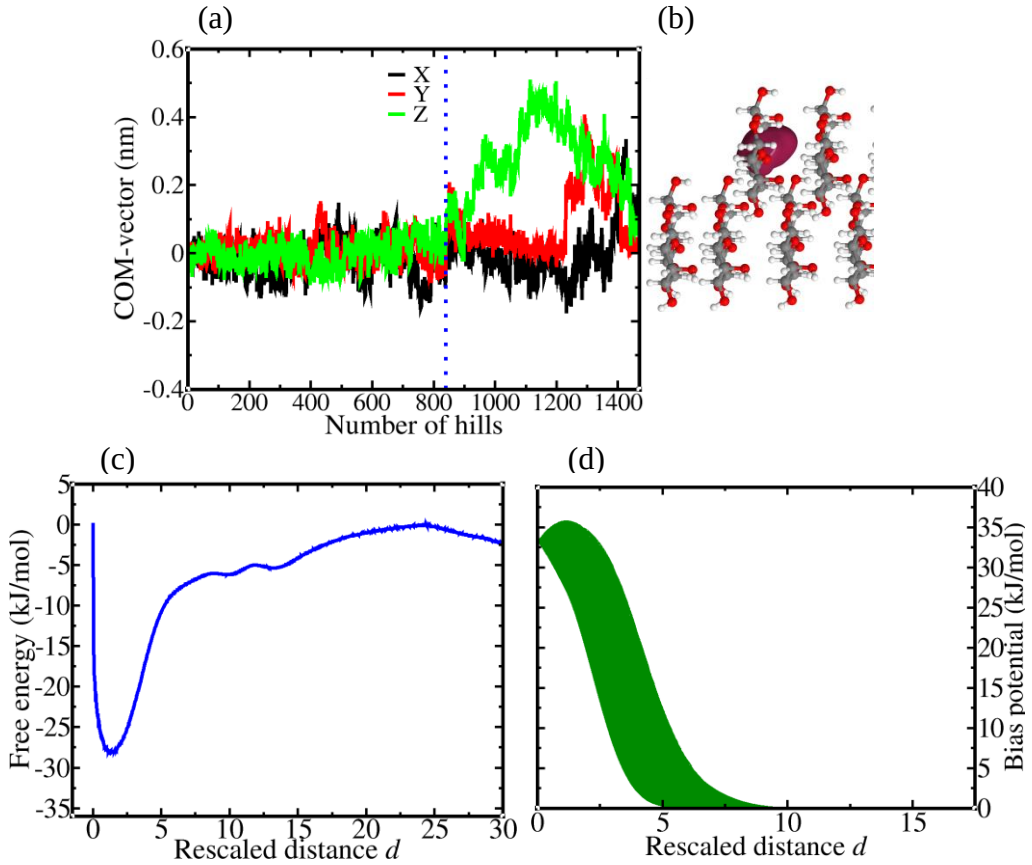

Figure 9.14: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 840 hills and marked by dotted lines.

Table 9.7: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	6.23 x 10^6	0.02	1.03 x 10^5	9.70	770.64	1963	23.95
d = 17	2.10 x 10^6	0.05	1.22 x 10^5	8.20	738.68	523	25.62
d = 19	8.77 x 10^6	0.12	1.03 x 10^5	9.75	721.97	191	26.76

b) Kink Type B

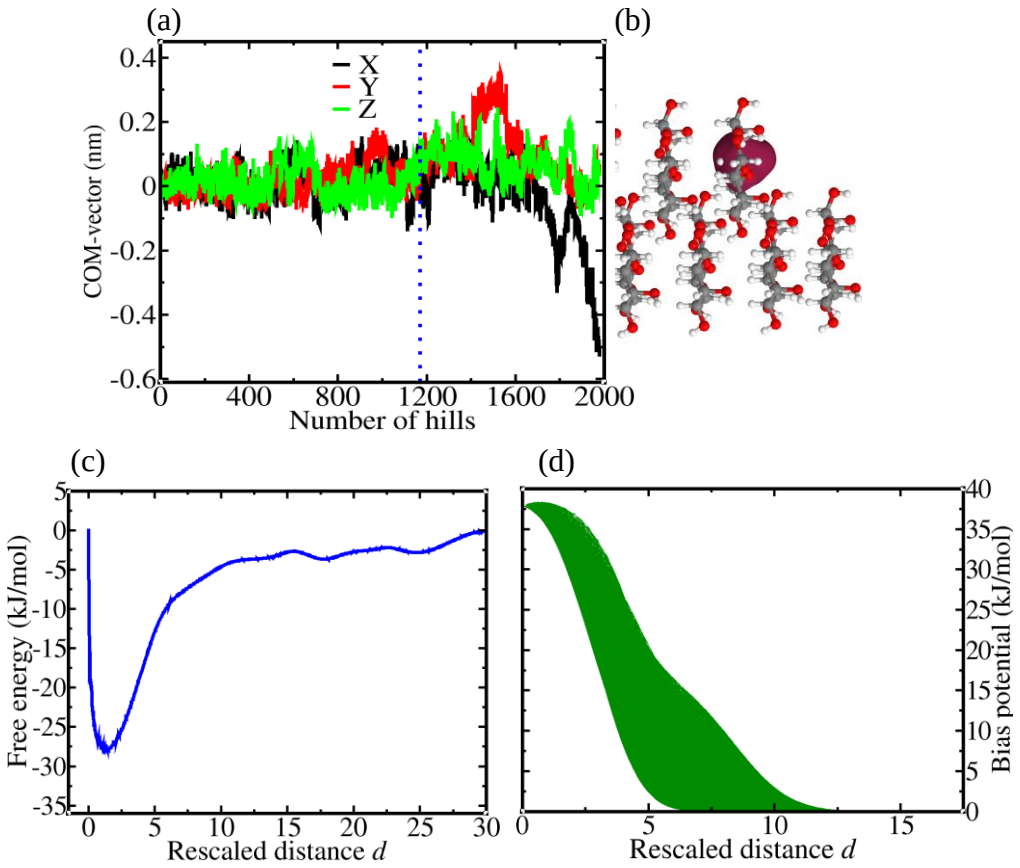

Figure 9.15: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 1170 hills and marked by dotted lines.

Table 9.8: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	2.61 x 10^7	0.003	7.98 x 10^4	12.54	1465.08	11758	25.17
d = 17	8.54 x 10^6	0.01	8.60 x 10^4	11.63	1368.30	3110	24.49
d = 19	2.98 x 10^6	0.03	8.81 x 10^4	11.34	1269.08	950	24.72

4. Step edge $[00\bar{1}]$

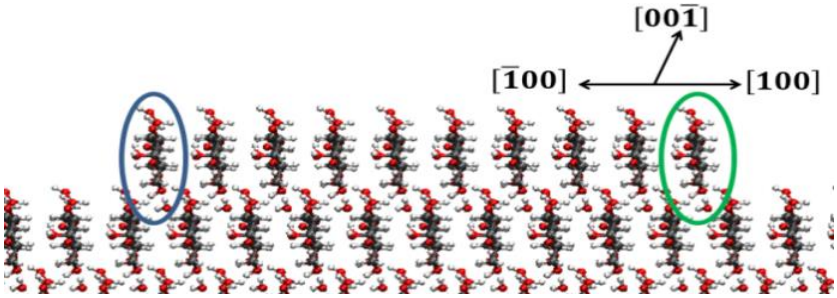

Figure 9.16: Same as Figure 9.7. Top view of the $[00\bar{1}]$ edge of the alpha-LM (010) face with the considered dissolution units marked by solid lines.

a) Kink Type A

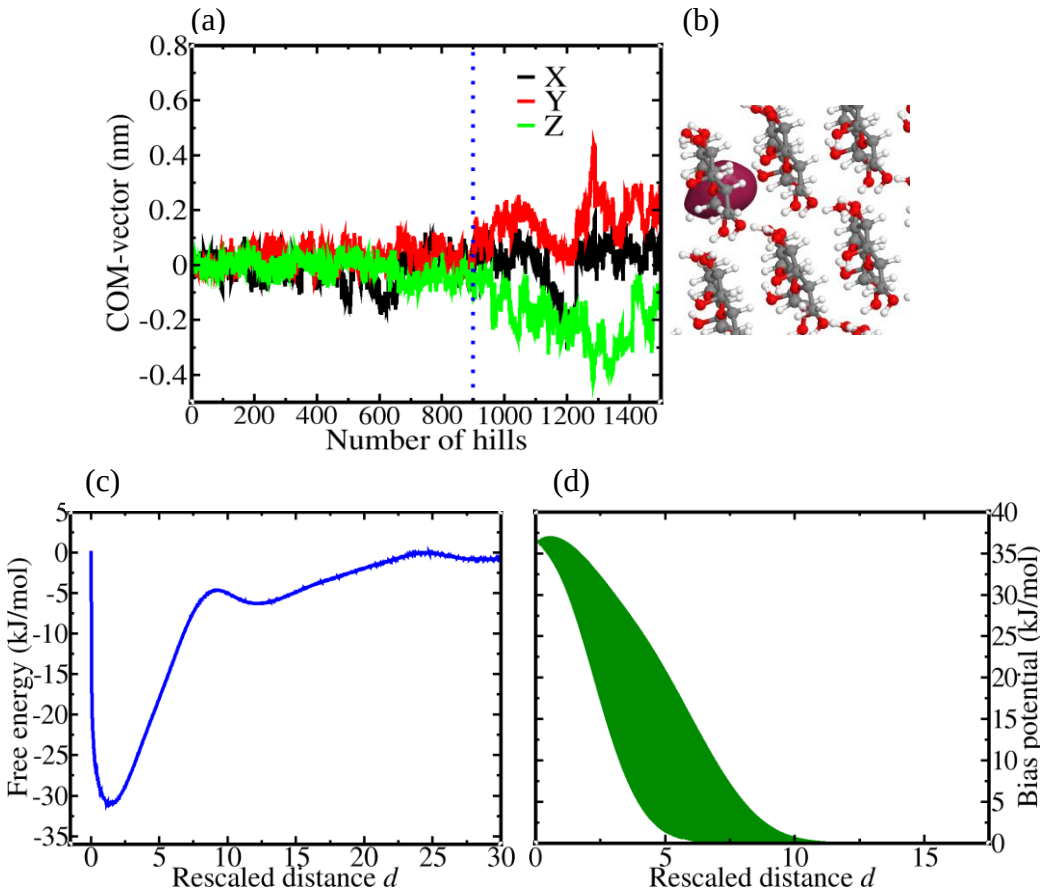

Figure 9.17: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 900 hills and marked by dotted lines.

Table 9.9: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	2.07 x 10^7	0.001	2.30 x 10^4	43.46	2902.40	30789	26.19
d = 17	5.26 x 10^6	0.004	2.01 x 10^4	49.76	2741.45	7918	27.89
d = 19	1.38 x 10^6	0.02	2.49 x 10^4	40.16	2667.47	2052	28.97

b) Kink Type B

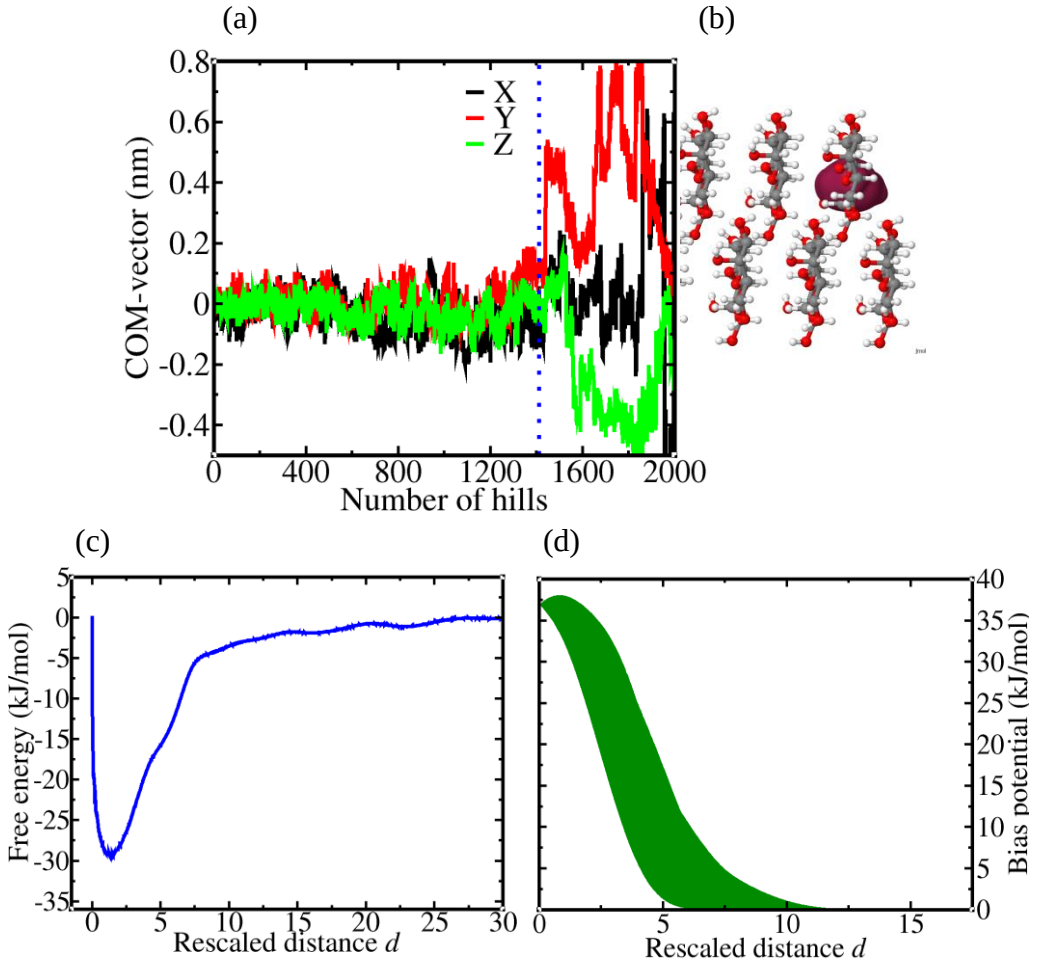

Figure 9.18: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 1410 hills and marked by dotted lines.

Table 9.10: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	$\kappa(t)$	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	3.74×10^{6}	0.01	3.70×10^{4}	27.05	3135.52	4325	27.61
d = 17	8.31×10^{5}	0.05	4.04×10^{4}	24.75	3007.30	835	27.64
d = 19	2.92×10^{5}	0.14	4.12×10^{4}	24.26	2916.70	282	28.29

5. Step edge $[10\bar{1}]$

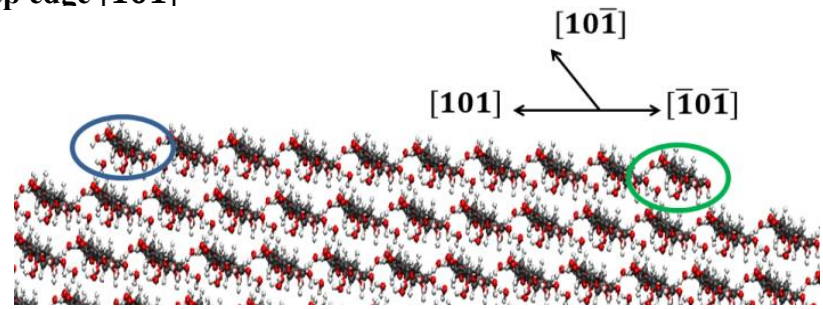

Figure 9.19: Same as Figure 9.7. Top view of the $[10\bar{1}]$ edge of the alpha-LM (010) face with the considered dissolution units marked by solid lines.

a) Kink Type A

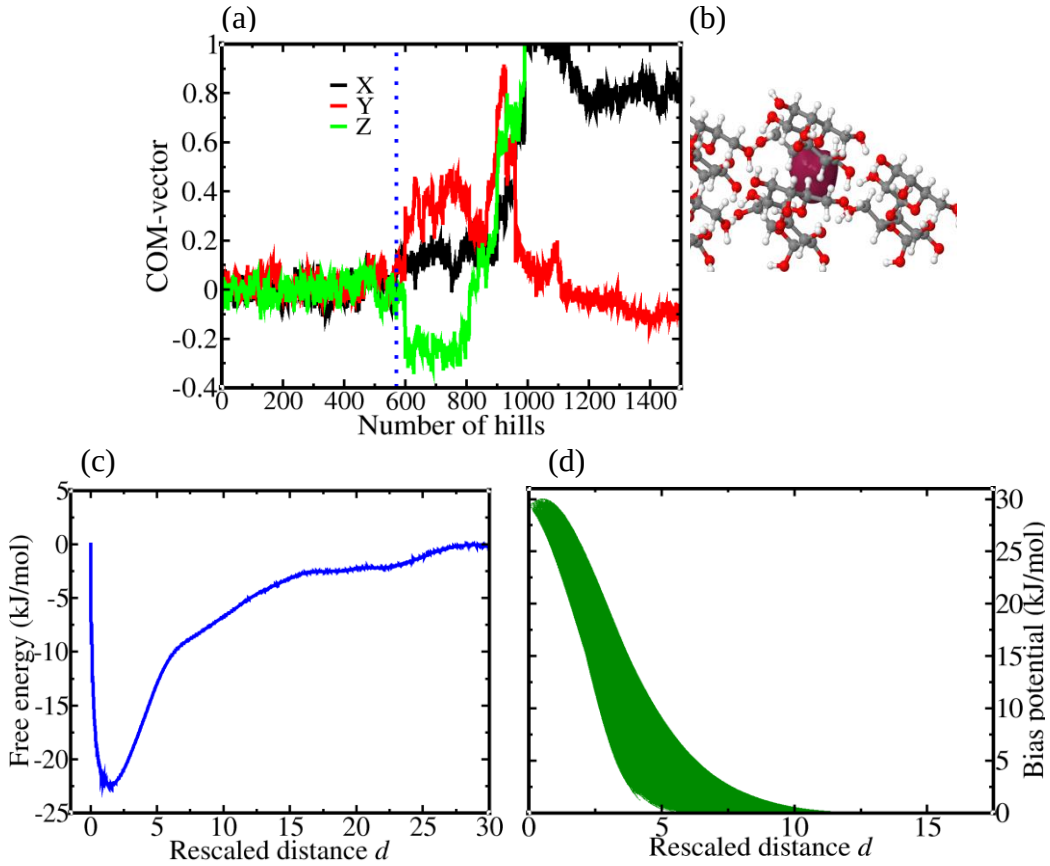

Figure 9.20: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 570 hills and marked by dotted lines.

Table 9.11: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	5.69 x 10^7	0.01	6.86 x 10^5	1.46	71.83	2224	19.37
d = 17	1.92 x 10^7	0.05	9.23 x 10^5	1.08	70.57	642	19.98
d = 19	9.24 x 10^6	0.12	1.12 x 10^6	0.89	69.41	241	20.05

b) Kink Type B

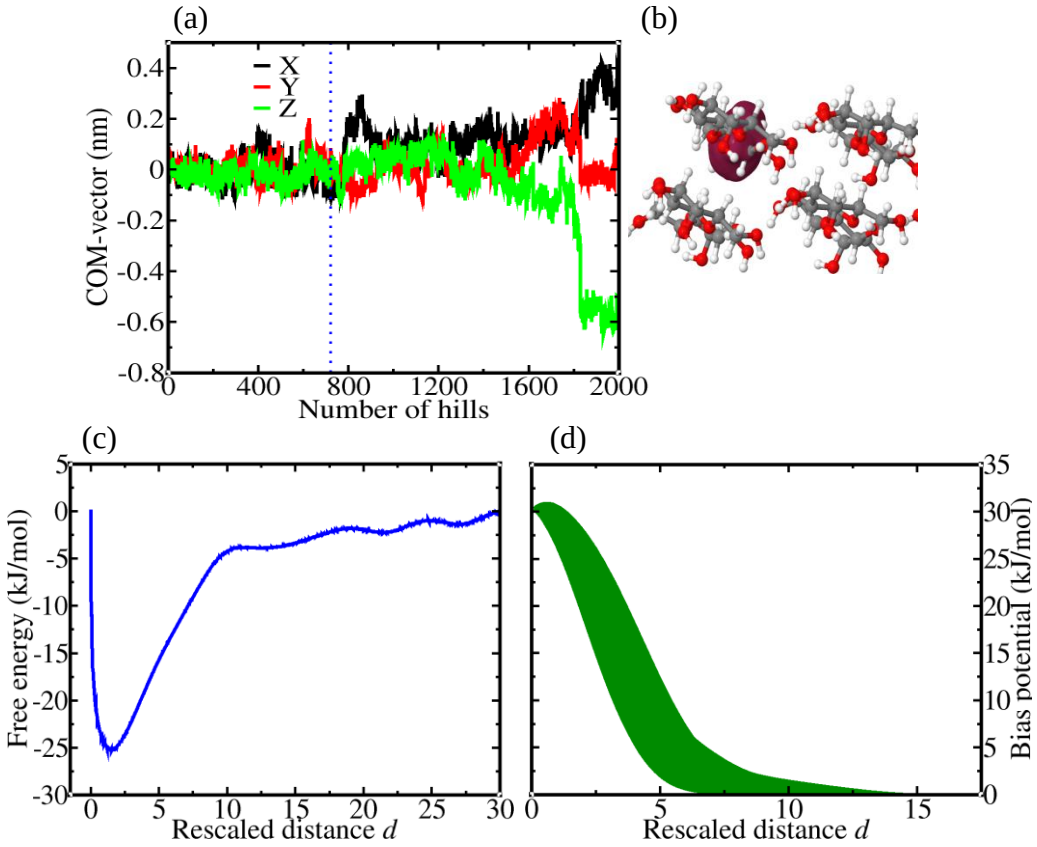

Figure 9.21: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 720 hills and marked by dotted lines.

Table 9.12: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	1.89 x 10^8	0.002	2.98 x 10^5	3.36	276.23	20373	21.96
d = 17	1.03 x 10^8	0.003	3.12 x 10^5	3.21	269.51	8693	22.88
d = 19	2.72 x 10^7	0.01	3.90 x 10^5	2.57	265.47	2015	23.53

6. Step edge $[\bar{1}01]$

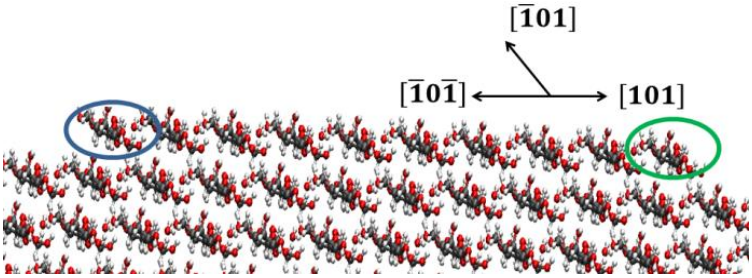

Figure 9.22: Same as Figure 9.7. Top view of the $[10\bar{1}]$ edge of the alpha-LM (010) face with the considered dissolution units marked by solid lines.

a) Kink Type A

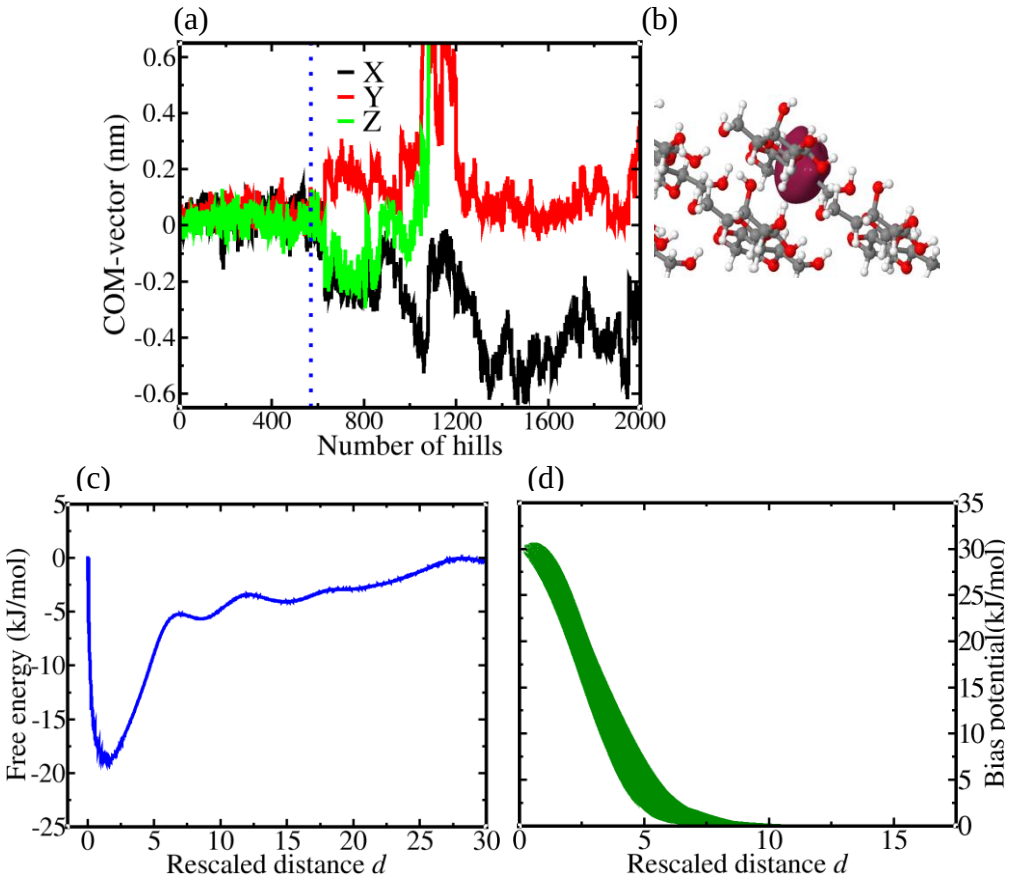

Figure 9.23: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 570 hills and marked by dotted lines.

Table 9.13: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	2.09 x 10^8	0.01	1.66 x 10^6	0.60	54.85	4234	15.00
d = 17	9.51 x 10^7	0.02	1.83 x 10^6	0.55	49.20	1499	15.64
d = 19	4.47 x 10^7	0.04	1.92 x 10^6	0.52	46.01	678	16.15

b) Kink Type B

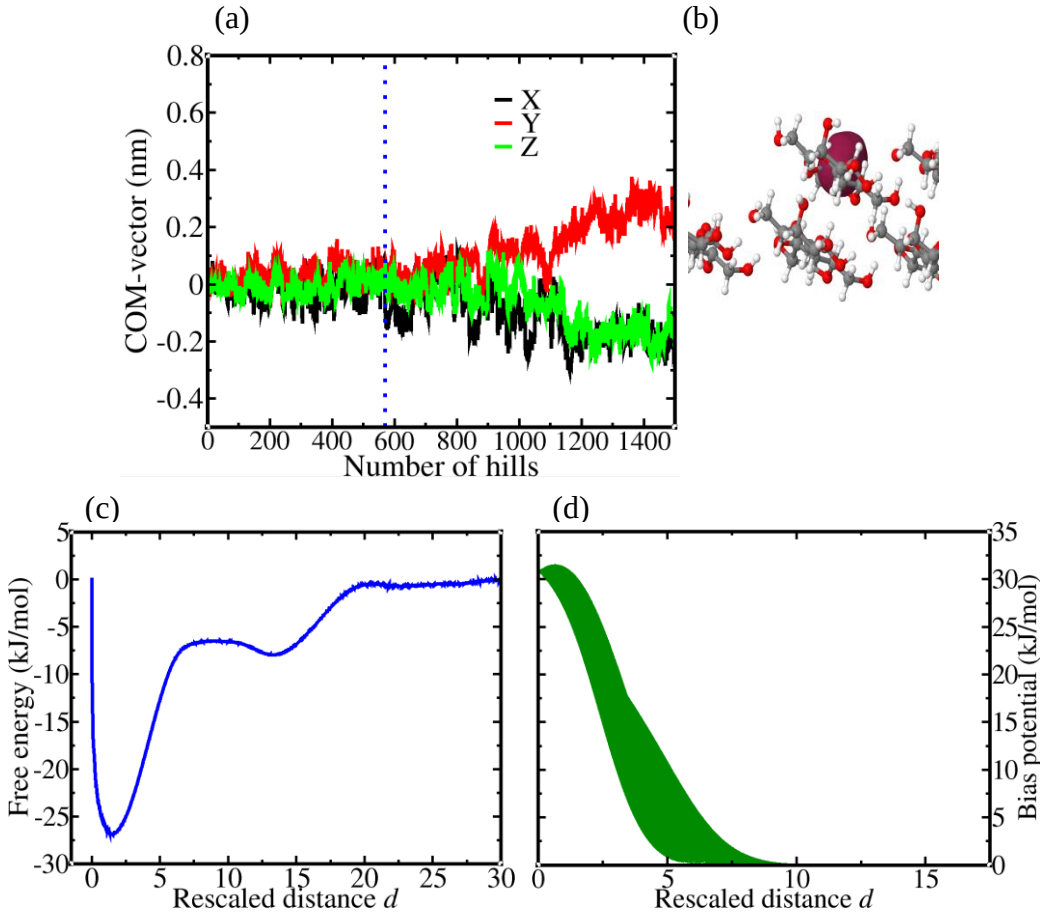

Figure 9.24: Same as Figure 9.8 for the kink type molecule considered. The truncation point is at 570 hills and marked by dotted lines.

Table 9.14: Same as Table 9.3 for the considered kink type molecule.

Barrier location	k_{TST} (s^{-1})	κ(t)	k (s^{-1})	k^{-1} (μs)	Boost factor (f_{acc})	n_{detach}	Barrier to well height (kJ/mole)
d = 15	2.94×10^{7}	0.004	1.18×10^{5}	8.50	427.37	5967	20.08
d = 17	9.61×10^{6}	0.01	9.41×10^{4}	10.63	393.13	1975	23.14
d = 19	2.71×10^{6}	0.04	1.08×10^{5}	9.26	384.56	526	25.85

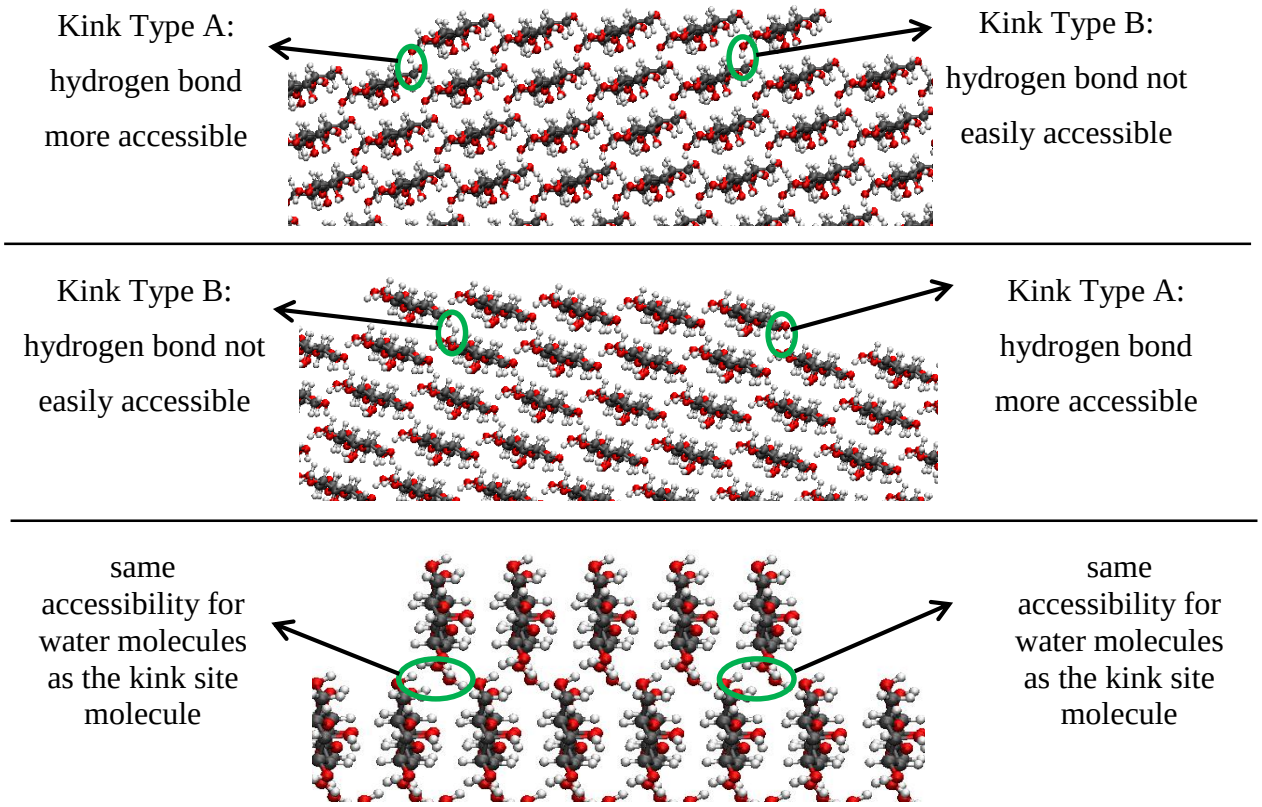

Figure 9.25: The accessibility of water molecules to intermolecular bonds between kink site molecules and their neighbors is shown for different kink types along the [100], [10$\bar{1}$] and [001] step edges of alpha-LM, displayed from upper to lower parts, respectively. The same explanation is valid for opposite edges [$\bar{1}$00], [$\bar{1}$01] and [00$\bar{1}$].

Table 9.15: Kink free energies and detachment rate constants of different kink types along all PBC edges for alpha-LM(010).

Edge direction	ΔG_{kink} (kJ/mole)	Detachment rate (s^{-1})	
		Kink Type A	Kink Type B
[100]	2.56±0.80	5.55×10^5 ($\pm 1.05 \times 10^5$)	6.51×10^4 ($\pm 5.95 \times 10^3$)
[$\bar{1}$00]	2.66±1.43	2.48×10^6 ($\pm 2.00 \times 10^5$)	2.93×10^5 ($\pm 2.24 \times 10^4$)
[001]	9.81±0.90	1.22×10^5 ($\pm 9.02 \times 10^3$)	8.60×10^4 ($\pm 3.55 \times 10^3$)
[00$\bar{1}$]	7.68±1.12	4.04×10^4 ($\pm 1.84 \times 10^3$)	2.49×10^4 ($\pm 1.98 \times 10^3$)
[10$\bar{1}$]	3.94±0.88	9.23×10^5 ($\pm 1.76 \times 10^5$)	2.69×10^5 ($\pm 4.04 \times 10^4$)
[$\bar{1}$01]	2.58±0.90	1.83×10^6 ($\pm 1.04 \times 10^5$)	9.41×10^4 ($\pm 9.66 \times 10^3$)

9.2.3 Predicted Dissolution Rates

As was the case for aspirin, the spiral dissolution model for alpha-LM draws only on the parameters calculated in the preceding section or on parameters that can easily be extracted from the crystal structure. As in the case of aspirin, the last thing needed to calculate the dissolution rate

based on the spiral model is to determine the number of edges that forms the spiral shape(s) on the faces. As explained in chapter 8.2 not all the edges have to be necessarily involved in the spiral formation and in the concomitant dissolution rate calculation. Which edges would be included depends on the critical lengths and step velocities of the edges. Considering the PBCs of the alpha-LM (010) face, four spiral shapes are possible; one of them is a six-fold and three of them are four-fold spirals. Though the characteristic rotation time, τ, in eq. 3.2 is calculated for both clockwise and anti-clockwise spiral rotation directions for all spiral shapes, only the most active spirals are displayed in this work as they correspond to lower τ values and dominate the dissolution rate calculations.

The net fluxes for the step edges are evaluated using eq. 3.15. Again, the low symmetry of the lactose molecules is reflected in differences of the detachment rate constants of opposite kink site molecules along the same edges seen in Table 9.15. Hence, the effective detachment rates are re-assessed as explained for the aspirin case. For each edge, the detachment rate reassessed depends on the ratio between faster and slower detachment processes of opposite kink site molecules, m, and the average number of molecules between kink sites, $\bar{n}$ (the reciprocal of the kink density, $\bar{n} = 1/\rho$) which itself depends on the kink free energies. When $\bar{n}$ is higher than the ratio m for the edges, then the final detachment rate is evaluated explicitly. In this case, the contributions of faster and slower detachment rate constants are calculated, and then each rate constant is multiplied by their contributions. By adding these resulting values the re-assessed detachment rates are attained. When this is not the case, i.e. the ratio, m is so large that all the kink sites and molecules in between them detach by the faster process, then the detachment rate constant for this faster process is divided by two to obtain the final detachment rate constants for the edges. It is also possible to do the division while calculating the step velocities; the reason for the division by two is the inactivity of half of the kink sites. In case of alpha-LM, for edges $[001]$, $[00\bar{1}]$ and $[10\bar{1}]$, the final re-assessed detachment rate constants are calculated explicitly as explained. For all remaining edges, the retreat of the step edge would be by the faster processes; hence half of the faster processes are taken as final detachment rate constants for these edges. In Table 9.16 these values are compiled together with ratios m and the number of molecules between kink sites, $\bar{n}$.

The calculated dissolution rates for different shapes of active spirals are displayed in Figure 9.27b. The shapes of these active spirals are shown in Figure 9.26a. The step height is again taken as half of the lattice parameter along the surface direction, i.e. h =10.9 Å. The different spiral shapes are color-coded: Spiral 1 depicted in green is a four-fold spiral along $[\bar{1}00]$, $[\bar{1}01]$, $[100]$ and $[10\bar{1}]$ edges and rotating clockwise. Spiral 2 depicted in red is a four-fold spiral along

$[00\bar{1}]$, $[10\bar{1}]$, [001] and $[\bar{1}01]$ edges and rotating anti-clockwise. Spiral 3 depicted in blue is a four-fold spiral along $[00\bar{1}]$, [100], [001] and $[\bar{1}00]$ edges and rotating clockwise, and finally spiral 4 depicted in black is a six-fold anti-clockwise rotating spiral along all available edges on the face. For alpha-LM, there is a single crystal experiment carried out by Raghavan *et al*. that measures the dissolution rate of the (010) face in water [152]. The dissolution rates are again calculated for undersaturations up to 15% as the experimental values are for this undersaturation range and the spiral dissolution is expected to be the dominating mechanism at low undersaturations. It should be noted that the error calculation for the detachment rates is only based on the average value of the detachment constants at different d values, hence high errors are associated with the dissolution rates.

Table 9.16: Kink free energies, ΔG_{kink}, kink densities ρ, the average number of molecules between kink sites $\bar{n}$, the ratio of faster to slower detachment rate constants m, and the reassessed detachment rate constants with error values in parentheses of all PBC edges of alpha-LM.

Edge direction	ΔG_{kink} (kJ/mole)	Kink density, ρ	$\bar{n}$	Ratio, m	Reassessed detachment rate constants (s^{-1})
[100]	2.56±0.80	0.4178	2.4	8.5	2.77×10^5 ($\pm 5.27 \times 10^4$)
$[\bar{1}00]$	2.66±1.43	0.4081	2.4	8.5	1.23×10^6 ($\pm 9.99 \times 10^4$)
[001]	9.81±0.90	0.0377	11.9	1.6	1.07×10^5 ($\pm 6.09 \times 10^3$)
$[00\bar{1}]$	7.68±1.12	0.0842	26.5	2.0	3.37×10^4 ($\pm 1.42 \times 10^3$)
$[10\bar{1}]$	3.94±0.88	0.2918	3.4	2.8	7.21×10^5 ($\pm 1.47 \times 10^5$)
$[\bar{1}01]$	2.58±0.90	0.4158	2.4	19.5	9.16×10^5 ($\pm 5.22 \times 10^4$)

As can be seen from Figure 9.26b and similar to the aspirin case, the four-fold spiral 1 is the fastest dissolving spiral shape. As for aspirin it is formed by edges with low critical lengths and high step velocities. However, in case of alpha-LM, the differences between step velocities or critical lengths of edges are not as pronounced as in the case of aspirin. This is actually the reason why for almost all spiral shapes the predicted dissolution rates are in agreement within one order of magnitude to the experimental values, especially at higher undersaturations. However, the fact that experimental values do not follow an exponential trend may be an indication that different spiral shapes would be dominant at different undersaturation ranges. Notwithstanding, it would be ambitious to draw conclusions with absolute certainty as the experiments could be affected by mass transfer limitations in the fluid as discussed for aspirin. Nonetheless, by assuming fast diffusion behavior especially at low undersaturations, the preliminary comparison of experimental and simulation results is very promising for the spiral dissolution model considered in this work.

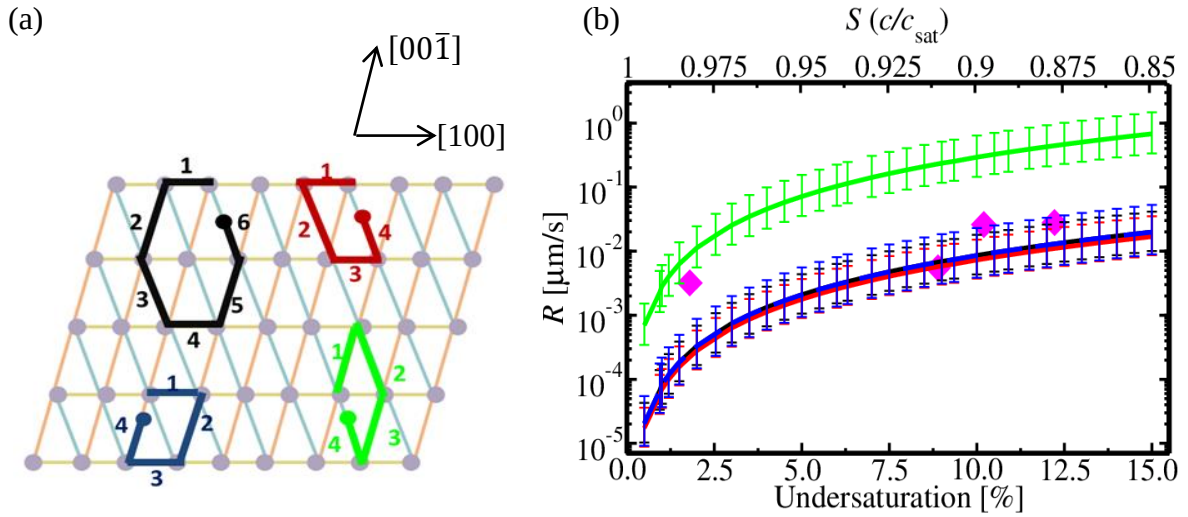

Figure 9.26: (a) The periodic bond chains (PBCs) at the alpha-LM (010) face together with the spiral shapes considered. See text for the edges forming the different spiral shapes and for the naming of spirals. (b) Calculated dissolution rates for the different spiral shapes considered. The pink colored diamonds are the experimental dissolution rates measured by Raghavan *et al.* [152] in single-crystal experiments. The predicted dissolution rates are displayed by colors that correspond to the colors of the shapes shown in part a: the green line is for spiral 1, the red is for spiral 2, blue is for spiral 3 and black is for spiral 4.

Table 9.17: Parameters used to build spiral model 1 for the alpha-LM (010) face. Distance retraced due to the loss of a single dissolution unit $a_{p,i}$, distance between dissolution units along edges $a_{e,i}$, step edge velocity v_i, critical lengths l_i, angles between edges i and $i+1$, $\alpha_{i,i+1}$, rotation time τ , and the dissolution rate, R at $c/c_{\text{sat}} = 0.995$.

Edge	$a_{p,i}$ (Å)	$a_{e,i}$ (Å)	v_i (Å/s)*	l_i (Å)	$\alpha_{i,i+1}$ (°C)	τ (s)	R (µm/s)
Edge1=[$\bar{1}$00]	4.6	7.6	11619.60	3265.72	144.6	1.56	6.99 x 10^{-4}
Edge2=[10$\bar{1}$]	4.4	7.9	4968.80	5041.16	35.4		
Edge3=[100]	4.6	7.6	2665.64	3143.55	144.6		
Edge4=[$\bar{1}$01]	4.4	7.9	8434.09	3297.22	35.4		

*detachment rate constants in Table 9.16 are used.

Table 9.18: Same Table 9.17 for spiral 2.

Edge	$a_{p,i}$ (Å)	$a_{e,i}$ (Å)	v_i (Å/s)*	l_i (Å)	$\alpha_{i,i+1}$ (°C)	τ (s)	R (µm/s)
Edge1=[00$\bar{1}$]	7.4	4.7	104.99	5872.30	111.2	62.65	1.74 x 10^{-5}
Edge2=[$\bar{1}$01]	4.4	7.9	8434.09	3297.22	68.8		
Edge3=[001]	7.4	4.7	149.54	7498.67	111.2		
Edge4=[10$\bar{1}$]	4.4	7.9	4968.80	5041.16	68.8		

* detachment rate constants in Table 9.16 are used.

Table 9.19: Same as Table 9.17 for spiral 3.

Edge	$a_{p,i}$ (Å)	$a_{e,i}$ (Å)	v_i (Å/s)*	l_i (Å)	$\alpha_{i,i+1}$ (°C)	τ (s)	R (μm/s)
Edge1=$[00\bar{1}]$	7.4	4.7	104.99	5872.30	104.2	53.28	2.04×10^{-5}
Edge2=[100]	4.6	7.6	2665.64	3143.55	75.8		
Edge3=[001]	7.4	4.7	149.54	7498.67	104.2		
Edge4=$[\bar{1}00]$	4.6	7.6	11619.60	3265.72	75.8		

* detachment rate constants in Table 9.16 are used.

Table 9.20: Same as Table 9.17 for spiral 4.

Edge	$a_{p,i}$ (Å)	$a_{e,i}$ (Å)	v_i (Å/s)*	l_i (Å)	$\alpha_{i,i+1}$ (°C)	τ (s)	R (μm/s)
Edge1=$[00\bar{1}]$	7.4	4.7	104.99	5872.30	104.2	53.89	2.02×10^{-5}
Edge2=$[\bar{1}00]$	4.6	7.6	11619.60	3265.72	144.6		
Edge3=$[\bar{1}01]$	4.4	7.9	8434.09	3297.22	111.2		
Edge4=[001]	7.4	4.7	149.54	7498.67	104.2		
Edge5=[100]	4.6	7.6	2665.64	3143.55	144.6		
Edge6=$[10\bar{1}]$	4.4	7.9	4968.80	5041.16	111.2		

* detachment rate constants in Table 9.16 are used.

Also in the case of alpha-LM, the total number of simulations performed to predict the dissolution rates is rather large. In fact it is even more expensive in comparison to aspirin due to the use of more λ points for the kink free energy calculation for each decoupling contribution which themselves are the most expensive simulations. For a fast screening approach, it would certainly be preferable to lower this computational burden without significant sacrifices in accuracy. Hence, the edge specific kink free energies and detachment constants are re-analyzed to consider the same kinds of approximations as was done for aspirin in chapter 8.2.3

For a fast screening approach, it would be enough to determine the dissolution rate for the fastest spiral shape as this will dominate the dissolution process. As this spiral is formed by the edges with the highest step velocities, hence by edges with lowest kink free energies, the first approximation would again be to determine which edges these would be *before* starting the free energy calculations. Using crude methods such as solid-state interaction energies was shown to be

effective for relative step velocity calculations by Doherty *et al*. for crystal morphology predictions [28, 44, 150] and also by the application to the aspirin case in chapter 8.2.3 of this work. With this low-cost method opposite edge site effects cannot be considered as bond breaking/forming on opposite step edges actually involves the same bonds. The absolute energies calculated by this approximate method are considerably different to the kink free energies calculated using the thermodynamic cycle. The specific values are as follows 33.81 kJ/mole for the [001] edge, 23.54 kJ/mole for the [100] edge and 19.35 kJ/mole for the $[10\bar{1}]$ edge. Nevertheless, the relative trend in the kink free energies is reproduced correctly, proving that with this less expensive approach, the edges that form the most active spiral can be identified beforehand.

Comparing the detachment rate constants in Table 9.15 for different kink sites and the reassessed detachment rate constants in Table 9.16 one can see that the effects of faster detachment processes in the overall calculation of the net flux are more pronounced than the effects of slower ones. Hence, re-evaluation of dissolution rates by only considering the faster detachment processes would still give comparable results to the exact dissolution rates calculated by using the reassessed detachment rate constants. As already mentioned in the aspirin case, this approximation would at first appear trickier than the first approximation. However, it was already shown that the solvent accessibility is the source of the differences in detachment rates of opposite kink site molecules (see Figure 9.25). Though it may be at the cost of human time, the identification of faster detachment processes would also allow to reduce the number of simulations, specifically the hyperdynamics simulations. With this approximation applied, the dissolution rates are re-evaluated for all spiral shapes. However, in practice one would only do this for spiral 1 for a fast screening approach as the edges forming this shape were already identified with the first approximation. In Figure 9.27b, the dissolution rates calculated applying this second approximation are displayed. Comparing to the original dissolution rates displayed in Figure 9.27a it is clear that the rates obtained are within the error range of the latter. Hence, also this second approximation would be reasonable to apply.

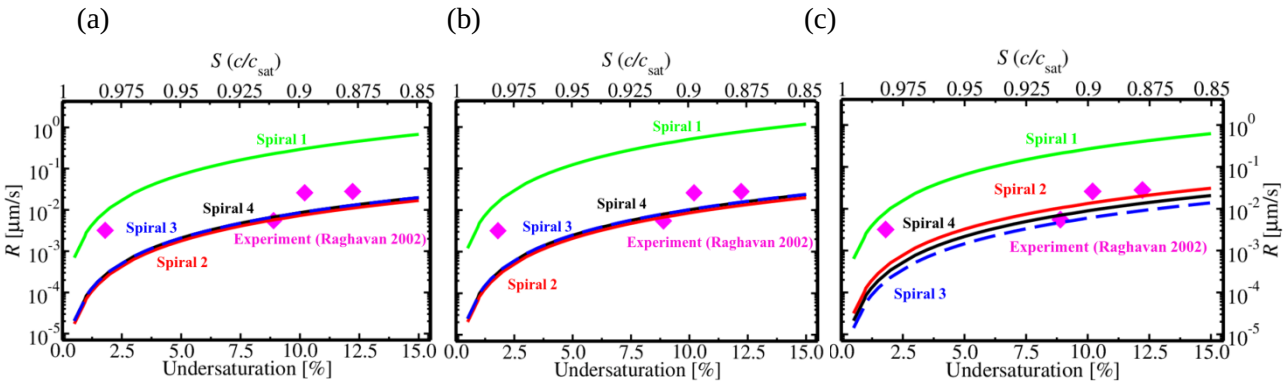

Figure 9.27: (a) Dissolution rates obtained by accounting for the different kink types along an edge for all spiral shapes considered for the alpha-LM (010) face. Same as Figure 9.26b without error bars. (b) Dissolution rates evaluated for a fast screening approach using the second approximation, see text. (c) Dissolution rates calculated using both second and third approximations, see text.

It is even possible to further reduce the number of required simulations by neglecting the opposite step edge effects. This means that for instance only the kink free energy and detachment rate constant of the [010] edge is calculated and these values are also used for the opposing $[0\bar{1}0]$ edge direction. In other words, asymmetries of the spiral shapes are ignored. In the aspirin case, a correlation was observed between step velocities and kink free energies such that edges with higher kink free energies have lower step velocities, at least in most cases. For alpha-LM, however, such a correlation is trickier to deduce. First, the difference between kink free energies of [100] and $[\bar{1}00]$ edge directions is only about 0.1 kJ/mole; a negligible difference. This renders it difficult to decide which edge to choose for these step directions based on only the kink free energies. Then, for the edges [001] and $[00\bar{1}]$, the correlation between step velocities and kink free energies is the other way around, i.e. the $[00\bar{1}]$ edge with the lower kink free energy has a higher step velocity as the detachment rate constant is higher for this edge. Nevertheless, even though the kink free energies of these edges are not within the error ranges of each other, the actual ratio of step velocities is only 1.4. Hence, it still does not make too much of a difference to nevertheless consider the edge directions with higher $\Delta G_{\mathrm{kink},i}$ values for this third approximation. In fact the dissolution rates calculated by applying this approximation are indeed within the error range of the ones calculated by using the correct reassessed rate constants. It is also possible to choose the edges with the lower step velocities from the opposite edge directions and arrive again at dissolution rates that are within the error ranges. The choice of the edges with lower velocities may actually be preferred due to the possibility of identifying them beforehand. It is possible to detect the edges that would have lower step velocities without actually performing simulations just

on the basis of solvent accessibility. Edges that have the hydroxyl groups pointing towards the solvents are more hydrophilic and hence more accessible by solvent water molecules. Correspondingly, it is easier for molecules on these edges to dissolve and so they have higher velocities. In the case of alpha-LM, the edges with higher velocities are [001], [$\bar{1}$00] and [$\bar{1}$01]. The opposite edges of these edges considered are the ones that provide less access to solvent molecules and have slower velocities. Therefore, the molecular data for [001], [100] and [10$\bar{1}$] edge directions are used exclusively to re-evaluate the dissolution rates. The comparisons of the approximate dissolution rates to the ones without any approximation can be found in Figure 9.27. By considering all these approximations, the computational demand is reduced so significantly that it will be trivial for nowadays computers to perform these simulations without taking long times.

9.3 Summary of Chapter 9

In this chapter, the computational protocol explained in chapter 7 and established for the aspirin (001) face in chapter 8.2 has been adapted for the alpha-LM (010) face to prove its applicability for different systems. Again, by using the MD based methods previously developed [43, 117], kink free energies and detachment rate constants for specific edges were accurately evaluated. The net fluxes for edges are estimated by eq. 3.15 that is the equation for the net flux of centrosymmetric dissolution units. However, the low symmetry and complex nature of lactose are still taken into account through the dependence on the ratio of faster to slower detachment rate constants and number of molecules between kink sites. As in the aspirin case, the periodic bond chains on this face reveal that different spiral shapes are possible: three four-fold and one six-fold spiral shapes. For all these shapes, dissolution rates are determined for both rotation directions, however only the ones with the higher dissolution rates, i.e. the more active spirals, were displayed and discussed. The predicted dissolution rates compare again quite favorably to available experimental data. They are within the same order of magnitude, though it is acknowledged that contributions from diffusive transport processes are disregarded in the model. By again considering some approximations based on the structural differences of kink site molecules and edge directions, the computational cost of the overall simulations are reduced significantly. This renders the method suitable as a fast computational screening approach for drug formulations.

10 General Conclusions and Outlook

In this thesis, a computational protocol to predict the dissolution rates of organic crystals at low undersaturations is proposed based on the classic spiral growth model of Burton, Cabrera and Frank [35], widely known as BCF model. As the time-reverse process of crystal growth, the BCF model can readily be applied to the dissolution process. Simple extensions to the original conceptual model had already allowed the prediction of crystal morphology as explained in chapter 3. The appealing elegance of such models is that they require only a minimal number of material-specific quantities to be determined. Specifically, kink free energies and detachment rate constants for edges are needed to estimate step velocities and critical lengths of the edges along with some readily available crystallographic data. In these existing mechanistic models, however, low-level methods such as solid-state interaction energies were employed to obtain rough estimates of these parameters estimations, since only relative growth rates of faces are necessary for morphology predictions. These methods do not allow an accurate determination of these parameters, which are needed to quantitatively predict absolute dissolution rates of individual faces. Hence, in this thesis MD-based methods are advanced and employed for the evaluation of these parameters as explained in detail in chapter 7 and refs. [43, 117]. By transferring all the concepts developed for growth in an analogous way to the dissolution process, a micro-kinetic spiral dissolution model is thus built to predict the dissolution rates.

The application of the model to two prototype compounds was described in the remaining chapters of part III. First, in chapter 8, the model is applied to aspirin, a widely used and also experimentally well-studied active pharmaceutical ingredient (API). Specifically the (001) face of aspirin is chosen, representing a dominant face of crystals grown from solutions [148]. The second application of the model is demonstrated in chapter 9 for alpha-lactose monohydrate (alpha-LM), a widely used excipient in pharmaceutical industry and also again an experimentally well-studied organic crystal. In this case, the chosen face is the (010) face, again representing a dominant face of crystals grown from solution [155] and a face for which dissolution rates at different undersaturations have been measured [152]. By considering the periodic bond chains (PBC) [41] on the selected faces, the step edge directions are determined and kink sites are created along these edges. In the aspirin case especially, the low symmetry of the molecule may offer several possibilities to create kink arrangements and thereby complicate the model. However, it was observed that not all these kink types are stable on MD simulation times. Hence, only meta-stable kink structures that will control the dissolution rates are considered and this way the model is

simplified. Still, even then opposing kink site molecules along the same step edges reveal different terminations for both crystal faces and lead to detachment at different rates. These differences can be traced back to differences in the solvent accessibility to intermolecular interactions between kink site molecules and their neighbors. In the full model these differences are explicitly considered while calculating the net flux for edges and step velocities. In contrast, there is no necessity to consider a master equation that accounts for transformations of kink sites from one to another as presented by Kuvadia *et al.* [44]. By focusing on meta-stable kink structures, always equivalent kink site molecules are attained after attachment/detachment processes from the same side of the edges. Therefore, the effects of these opposing kink sites along edges can effectively be accounted for through the ratio of faster to slower detachment rate constants and the average number of molecules between kink sites. This allows to explicitly calculate the net flux.

The overall dissolution rate depends on the specific form and shape of the dissolving spiral. In this work all shapes resulting from edges along the PBCs of the face are considered. This leads to six-fold and four-fold spirals for both aspirin (001) and alpha-LM (010). The spiral rotation direction also affects the dissolution rate, as it dictates which edge velocity is paired with the corresponding critical length of following edge. The spiral shape with the highest dissolution rate in terms of both step edges involved and rotation direction will be the most active spiral and will dominate the overall dissolution rate. This most active spiral shape is actually formed by edges with lowest critical lengths and highest step velocities both of which themselves depend on undersaturation ratios. The therewith predicted dissolution rates of the overall crystal face compare quite favorably, within about an order of magnitude, with existing experimental data. For alpha-LM, this comparison is more straightforward as experimental dissolution rates have been reported for specific undersaturations. In case of alpha-LM such specification has not been provided, allowing only for a qualitative comparison. One should also keep in mind that the measured rates are likely affected by contributions from diffusive processes in the solution, which are neglected in the model considered here. Despite all these limitations, the *in silico* protocol developed in this work appears promising to yield reliable micro-kinetic dissolution rates from molecular simulations.

There is a disadvantage or shortcoming of the proposed protocol, however, in terms of computational demand and time. This demand is very high when explicitly determining the necessary parameters for all considered edge directions. For a fast screening approach for either potential new APIs or excipients, it is desirable to predict rates without such high computational costs. Hence, several approximations to the model are proposed that impose only insignificant

losses in the accuracy. Starting with low-level methods as already proposed by Doherty *et al*. [28, 32, 44], the edges forming the most active spiral, i.e. the spiral form with the highest dissolution rate, can be identified before expensive MD calculations are started. Neglecting the differences between opposite kink site molecules and just focusing on the faster of the detachment processes from kink sites along any given step edge would mean even fewer simulations to determine detachment rate constants. As such, which kink site molecule will have a higher detachment rate constant can already be determined beforehand by structural differences attributed to solvent accessibility. A last approximation considered is to ignore the opposite step edge direction effects and just consider the molecular data of one edge direction for symmetrically related edges such as $[010]$ and $[0\bar{1}0]$. Edges with lower step velocities should thereby be chosen as their contribution to the overall dissolution rate will be more pronounced. This identification of the edges with lower step velocities can be done beforehand, i.e. without any simulation, again by considering the structural differences between edges. For instance, edge directions with more hydrophobic terminations will have lower step velocities. At the end, only the molecular data for a limited number of edges is needed to predict the dissolution rates. Though these approximations involve human input, they thus lower the demand of computer time and power considerably. This renders the approach suitable for fast computational screenings of compounds used in drug formulations, and hence can help in reducing the time and cost of drug development and design processes.

As an outlook, it is clear that extension of the protocol to other organic compounds will be more or less straightforward. It is acknowledged that it may not always be possible to reduce the number of kink sites to a pair of opposite kink sites for edges of some crystal structures. Then, to calculate the step velocities, the mole fractions of solute molecules near the edges are required as seen in eqs. 3.23 and 3.24. The equilibrium mole fraction of solute molecules actually depends on the dissolution free energy that is the energy required to remove a surface molecule from its kink site position (see eq. 3.25). This can also be calculated by the thermodynamic cycle depicted in chapter 7.1 of this work, albeit the scheme would be slightly different as explained in ref. [117]. The attachment rate constants can even be evaluated by using these dissolution free energies and the detailed balance equation for incorporation and disincorporation events. By this way, even for crystal structures and faces that have more than one kink types per edges and where kink sites transform from one type to another, it is possible to calculate the dissolution rates of faces. The much more consequential complication and limitation of the here developed model is the necessity to know the bulk crystal structure and the faces dominating the dissolution process. For yet unsynthesized drug candidates this information is not generally available. For a wide applicability

of the here developed protocol it would thus need to be combined with a global crystal structure determination approach – an equally active field of research. Last, not least, the present model only provides intrinsic dissolution rates, i.e. it neglects any limitations on the measured rates by fluid transport. Hence, a coupling of the here devised microkinetic model to continuum fluid dynamics will be another interesting step to capture such effects.

References

[1] W. Sneader, *Drug Discovery: A History*, John Wiley & Sons Ltd., Chichester, 2005, (pages 1,5).

[2] G. L. Patrick, *An introduction to medicinal chemistry*, 4 ed., The Oxford University Press, Oxford, 2009, (pages 185,192,225).

[3] H. van de Waterbeemd, D. A. Smith, K. Beaumont, D. K. Walker, *Property-based design: Optimization of drug absorption and pharmacokinetics*, J. Med. Chem. **44**, 1313-1333 (2001).

[4] H. van de Waterbeemd, E. Gifford, *ADMET in silico modelling: towards prediction paradise?*, Nature reviews. Drug discovery **2**, 192-204 (2003).

[5] N. N. Song, S. Y. Zhang , C. X. Liua, *Overview of Factors Affecting Oral Drug Absorption*, Asian Journal of Drug Metabolism and Pharmacokinetics **4**, (2004).

[6] H. van de Waterbeemd, E. Gifford, *ADMET in silico modelling: Towards prediction paradise?*, Nature Reviews Drug Discovery **2**, 192-204 (2003).

[7] G. Pifferi, P. Santoro, M. Pedrani, *Quality and functionality of excipients*, Farmaco **54**, 1-14 (1999).

[8] G. L. Amidon, H. Lennernas, V. P. Shah, J. R. Crison, *A Theoretical Basis for a Biopharmaceutic Drug Classification - the Correlation of in-Vitro Drug Product Dissolution and in-Vivo Bioavailability*, Pharm. Res. **12**, 413-420 (1995).

[9] A. Dokoumetzidis, P. Macheras, *A Century of Dissolution Research: From Noyes and Whitney to the Biopharmaceutics Classification System*, Int. J. Pharm. **321**, 1-11 (2006).

[10] F. A. Campagna, E. Nelsons, G. Cureton, R. A. Mirigian, *Inactive Prednisone Tablets USP XVI*, J. Pharm. Sci. **52**, 605-606 (1963).

[11] G. Levy, *Effect of Dosage Form Properties on Therapeutic Efficiancy of Tolbutamide Tablets*, Canadian Medical Association Journal **90**, 978-979 (1964).

[12] G. Levy, A. J. M., D. C. White, *Effect of Certain Tablet Formulation Factors on Dissolution Rate of the Active Ingredient II*, J. Pharm. Sci. **52**, 1047-1051 (1963).

[13] C. MacLeod, H. Rabin, J. Ruedy, M. Caron, D. Zarowny, R. O. Davies, *Comparative Bioavailability of Three Brands of Ampicillin*, Canadian Medical Association Journal **107**, 203-209 (1972).

[14] E. Nelson, *Solution Rate of Theophylline Salts and Effects from Oral Administration*, J. Am. Pharm. Assoc. **46**, 607-614 (1957).

[15] C. Tedeschi, V. Clement, M. Rouvet, B. Valles-Pamies, *Dissolution tests as a tool for predicting bioaccessibility of nutrients during digestion*, Food Hydrocolloids **23**, 1228-1235 (2009).

[16] T. Cheng, R. Lozano, M. Yun, T. Mirza, R. Löbenberg, B. Nickerson, V. Gray, W. Qingxi, *The Value of In Vitro Dissolution in Drug Development*, Pharm. Technol. **33**, 52-64 (2009).

[17] X. Jia, R. A. Williams, *From Microstructures of Tablets and Granules to Their Dissolution Behaviour*, Dissolution Technologies **13**, 11-19 (2006).

[18] Y. Gao, K. W. Olsen, *Molecular dynamics of drug crystal dissolution: simulation of acetaminophen form I in water*, Mol. Pharm. **10**, 905-917 (2013).

[19] M. Greiner, E. Elts, H. Briesen, *Insights into pharmaceutical nanocrystal dissolution: a molecular dynamics simulation study on aspirin*, Mol. Pharm. **11**, 3009-3016 (2014).

[20] G. Lanaro, G. N. Patey, *Molecular dynamics simulation of NaCl dissolution*, J. Phys. Chem. B **119**, 4275-4283 (2015).

[21] D. Toroz, R. B. Hammond, K. J. Roberts, S. Harris, T. Ridley, *Molecular dynamics simulations of organic crystal dissolution: The lifetime and stability of the polymorphic forms of para-amino benzoic acid in aqueous environment*, J. Cryst. Growth **401**, 38-43 (2014).

[22] S. Piana, J. D. Gale, *Understanding the barriers to crystal growth: Dynamical simulation of the dissolution and growth of urea from aqueous solution*, J. Am. Chem. Soc. **127**, 1975-1982 (2005).

[23] A. G. Stack, P. Raiteri, J. D. Gale, *Accurate Rates of the Complex Mechanisms for Growth and Dissolution of Minerals Using a Combination of Rare-Event Theories*, J. Am. Chem. Soc. **134**, 11-14 (2012).

[24] A. Barat, H. J. Ruskin, M. Crane, *Probabilistic methods for drug dissolution. Part 2. Modelling a soluble binary drug delivery system dissolving in vitro*, Simul. Model. Pract. Theory **14**, 857-873 (2006).

[25] P. Costa, J. Manuel, S. Lobo, *Modeling and comparison of dissolution profiles*, Eur. J. Pharm. Sci. **13**, 123-133 (2001).

[26] B. Narasimhan, *Mathematical models describing polymer dissolution: consequences for drug delivery*, Adv. Drug Del. Rev. **48**, 195-210 (2001).

[27] J. Siepmann, *Mathematical modeling of controlled drug delivery*, Adv. Drug Del. Rev. **48**, 137-138 (2001).

[28] M. A. Lovette, A. R. Browning, D. W. Griffin, J. P. Sizemore, R. C. Snyder, M. F. Doherty, *Crystal Shape Engineering*, Ind. Eng. Chem. Res. **47**, 9812-9833 (2008).

[29] M. A. Lovette, M. F. Doherty, *Predictive Modeling of Supersaturation-Dependent Crystal Shapes*, Cryst. Growth Des. **12**, 656-669 (2012).

[30] M. A. Lovette, M. Muratore, M. F. Doherty, *Crystal Shape Modification Through Cycles of Dissolution and Growth: Attainable Regions and Experimental Validation*, American Institute of Chemical Engineers Journal **58**, 1465-1474 (2012).

[31] P. Smereka, *Spiral crystal growth*, Physica D **138**, 282-301 (2000).

[32] R. C. Snyder, M. F. Doherty, *Predicting crystal growth by spiral motion*, Proc. R. Soc. London, Ser. A **465**, 1145-1171 (2009).

[33] J. Zhang, G. H. Nancollas, *Kink Density and Rate of Step Movement during Growth and Dissolution of an AB Crystal in a Nonstoichiometric Solution*, J. Colloid Interface Sci. **200**, 131-145 (1998).

[34] J. J. De Yoreo, P. G. Vekilov, *Principles of Crystal Nucleation and Growth*, Rev. Mineral. Geochem. **54**, 57-93 (2008).

[35] W. K. Burton, N. Cabrera, F. C. Frank, *The Growth of Crystals and the Equilibrium Structure of Their Surfaces*, Philos. Trans. R. Soc. London, Ser. A **243**, 299-358 (1951).

[36] A. A. Chernov, L. N. Rashkovich, P. G. Vekilov, *Steps in Solution Growth: Dynamics of Kinks, Bunching and Turbulence*, J. Cryst. Growth **275**, 1-18 (2005).

[37] H. M. Cuppen, H. Meekes, W. J. P. van Enckevort, E. Vlieg, *Kink incorporation and step propagation in a non-Kossel model*, Surf. Sci. **571**, 41-62 (2004).

[38] V. K. W. Cheng, Computer simulation studies on crystal growth and dissolution kinetics in *Research in Chemical Kinetics* (Eds.: R. G. Compton, G. Hancock), Elsevier, Amsterdam, **1994**, pp. 81-146.

[39] K. Tsukamoto, 24 - In Situ Observation of Crystal Growth and Flows by Optical Techniques in *Handbook of Crystal Growth (Second Edition)* (Ed.: T. Nishinaga), Elsevier, Boston, **2015**, pp. 1031-1060.

[40] T. Nishinaga, *Handbook of Crystal Growth: Fundamentals*, Elsevier Science, 2014, (pages 1032).

[41] P. Hartman, W. G. Perdok, *On the relations between structure and morphology of crystals. I*, Acta Crystallogr. **8**, 49-52 (1955).
[42] P. Nozieres, F. Gallet, *The roughening transition of crystal-surfaces .I. Static and dynamic renormalization theory, crystal shape and facet growth*, Journal De Physique **48**, 353-367 (1987).
[43] J. Schneider, C. Zheng, K. Reuter, *Thermodynamics of Surface Defects at the Aspirin/Water Interface*, J. Chem. Phys. **141**, 124702 (2014).
[44] Z. B. Kuvadia, M. F. Doherty, *Spiral Growth Model for Faceted Crystals of Non-Centrosymmetric Organic Molecules Grown from Solution*, Cryst. Growth Des. **11**, 2780-2802 (2011).
[45] S. H. Kim, P. Dandekar, M. A. Lovette, M. F. Doherty, *Kink Rate Model for the General Case of Organic Molecular Crystals*, Cryst. Growth Des. **14**, 2460-2467 (2014).
[46] V. K. W. Cheng, *Kinetic Asymmetry Between Crystal-Growth and Dissolution On A Flat Surface - Computer-Simulation Study at High-Temperatures* Journal of the Chemical Society-Faraday Transactions **87**, 2467-2472 (1991).
[47] V. K. W. Cheng, E. C. M. Tang, T. B. Tang, *The Kinetic Asymmetry Between Nucleation Growth and Dissolution - A Monte-Carlo Study*, J. Cryst. Growth **96**, 293-303 (1989).
[48] J. Christoffersen, M. R. Christoffersen, *Spiral Growth and Dissolution Models with Rate Constants Related to the Frequency of Partial Dehydration of Cations and to the Surface Tension*, J. Cryst. Growth **87**, 41-50 (1988).
[49] M. Volmer, W. Schultze, *Condensation of crystals*, Zeitschrift Fur Physikalische Chemie-Abteilung a-Chemische Thermodynamik Kinetik Elektrochemie Eigenschaftslehre **156**, 1-22 (1931).
[50] F. Frank, *The influence of dislocations on crystal growth*, Discuss. Faraday Soc. **5**, 48-54 (1949).
[51] A. A. Chernov, *Modern crystallography III. Crystal growth*, Springer-Verlag, Berlin, 1984, (pages 152).
[52] A. E. Nielsen, *Electrolyte crystal growth mechanisms*, J. Cryst. Growth **67**, 289-310 (1984).
[53] M. R. Christoffersen, J. Dohrup, J. Christoffersen, *Kinetics of growth and dissolution of calcium hydroxyapatite in suspensions with variable calcium to phosphate ratio*, J. Cryst. Growth **186**, 283-290 (1998).
[54] R. Kaischew, E. Budevski, *Surface processes in electrocrystallization*, Contemporary Physics **8**, 489-516 (1967).
[55] V. Voronkov, *Dislocation mechanism of growth with a low kink density*, Sov. Phys. Cryst **18**, 19-23 (1973).
[56] V. Voronkov, *The movement of an elementary step by means of the formation of one-dimensional nuclei*, Sov. Phys. Cryst **15**, 8-13 (1970).
[57] D. Winn, M. F. Doherty, *A new technique for predicting the shape of solution-grown organic crystals*, AlChE J. **44**, 2501-2514 (1998).
[58] T. N. Thomas, T. A. Land, T. Martin, W. H. Casey, J. J. DeYoreo, *AFM investigation of step kinetics and hillock morphology of the {100} face of KDP*, J. Cryst. Growth **260**, 566-579 (2004).
[59] M. A. Lovette, M. F. Doherty, *Reinterpreting edge energies calculated from crystal growth experiments*, J. Cryst. Growth **327**, 117-126 (2011).
[60] J. Zhang, G. H. Nancollas, *Kink densities along a crystal surface step at low temperatures and under nonequilibrium conditions*, J. Cryst. Growth **106**, 181-190 (1990).
[61] W. Kossel, *Zur Theorie des Kristallwachstums*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **206**, 135-143 (1927).

[62] I. V. Markov, *Crystal Growth for Beginners: Fundamentals of Nucleation, Crystal Growth and Epitaxy*, 2003, (pages 15).
[63] R. J. D. Tilley, *Crystals and Crystal Structures*, Wiley, 2006, (pages 75, 76).
[64] J. P. Sizemore, M. F. Doherty, *A stochastic model for the critical length of a spiral edge*, J. Cryst. Growth **312**, 785-792 (2010).
[65] G. H. Gilmer, P. Bennema, *Computer Simulation of Crystal Surface Structure and Growth Kinetics*, J. Cryst. Growth **13-14**, 148-153 (1972).
[66] M. N. Offman, M. Krol, I. Silman, J. L. Sussman, A. H. Futerman, *Molecular Basis of Reduced Glucosylceramidase Activity in the Most Common Gaucher Disease Mutant, N370S*, J. Biol. Chem. **285**, 42105-42114 (2010).
[67] B. J. Alder, T. E. Wainwright, *Phase Transition for a Hard Sphere System*, J. Chem. Phys. **27**, 1208-1209 (1957).
[68] B. J. Alder, T. E. Wainwright, *Studies in Molecular Dynamics .1. General Method*, J. Chem. Phys. **31**, 459-466 (1959).
[69] R. W. Hockney, S. P. Goel, J. W. Eastwood, *Quiet High-Resolution Computer Models of a Plasma*, J. Comput. Phys. **14**, 148-158 (1974).
[70] L. Verlet, *Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules*, Phys. Rev. **159**, 98-+ (1967).
[71] W. F. Van Gunsteren, H. J. C. Berendsen, *A leap-frog algorithm for stochastic dynamics*, Molecular Simulation **1**, 173-185 (1988).
[72] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *Molecular dynamics with coupling to an external bath*, The Journal of Chemical Physics **81**, 3684-3690 (1984).
[73] G. Bussi, D. Donadio, M. Parrinello, *Canonical sampling through velocity rescaling*, J. Chem. Phys. **126**, 014101 (2007).
[74] S. Nose, *A molecular dynamics method for simulations in the canonical ensemble (Reprinted from Molecular Physics, vol 52, pg 255, 1984)*, Mol. Phys. **100**, 191-198 (2002).
[75] W. G. Hoover, *Canonical dynamics: equilibrium phase-space distributions.*, Phys. Rev. A **31**, 1695-1697 (1985).
[76] H. C. Andersen, *Molecular dynamics simulations at constant pressure and/or temperature*, J. Chem. Phys. **72**, 2384-2393 (1980).
[77] W. F. Vangunsteren, H. J. C. Berendsen, *Algorithms for brownian dynamics*, Mol. Phys. **45**, 637-647 (1982).
[78] M. Parrinello, A. Rahman, *Polymorphic transitions in single crystals: A new molecular dynamics method*, J. Appl. Phys. **52**, 7182-7190 (1981).
[79] A. Warshel, M. Levitt, *Theoretical Studies of Enzymic Reactions - Dielectric, Electrostatic and Steric Stabilization of Carbonium-Ion in Reaction of Lysozyme*, J. Mol. Biol. **103**, 227-249 (1976).
[80] M. Born, R. Oppenheimer, *Zur Quantentheorie der Molekeln*, Annalen der Physik **389**, 457-484 (1927).
[81] D. J. E. Callaway, A. Rahman, *Unified Approach for Molecular Dynamics and Density-Functional Theory*, Phys. Rev. Lett. **49**, 613-616 (1982).
[82] W. Kohn, L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev. **140**, A1133-A1138 (1965).
[83] J. P. Perdew, K. Burke, M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **77**, 3865-3868 (1996).
[84] A. D. Becke, *A new mixing of Hartree–Fock and local density-functional theories*, The Journal of Chemical Physics **98**, 1372-1377 (1993).

[85] B. Meyer, *The pseudopotential plane wave approach*, NIC Series **31**, 71-83 (2006).
[86] E. Schrödinger, *Quantisierung als eigenwertproblem*, Annalen der physik **385**, 437-490 (1926).
[87] D. J. Griffiths, *Introduction to quantum mechanics*, Pearson Education India, 2005, (pages
[88] M. Caffarel, P. Claverie, *Development of a pure diffusion quantum Monte Carlo method using a full generalized Feynman–Kac formula. I. Formalism*, The Journal of Chemical Physics **88**, 1088-1099 (1988).
[89] L. Brillouin, *La mécanique ondulatoire de Schrödinger; une méthode générale de résolution par approximations successives*, CR Acad. Sci **183**, 24-26 (1926).
[90] H. A. Kramers, *Wellenmechanik und halbzahlige Quantisierung*, Z. Phys. **39**, 828-840 (1926).
[91] G. Wentzel, *Eine Verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik*, Z. Phys. **38**, 518-529 (1926).
[92] T. A. Halgren, W. Damm, *Polarizable force fields*, Curr. Opin. Struct. Biol. **11**, 236-242 (2001).
[93] A. Korkut, W. A. Hendrickson, *A force field for virtual atom molecular mechanics of proteins*, Proc. Natl. Acad. Sci. U.S.A. **106**, 15667-15672 (2009).
[94] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries, *The MARTINI force field: coarse grained model for biomolecular simulations*, The Journal of Physical Chemistry B **111**, 7812-7824 (2007).
[95] A. C. T. van Duin, S. Dasgupta, F. Lorant, W. A. Goddard, *ReaxFF: A Reactive Force Field for Hydrocarbons*, The Journal of Physical Chemistry A **105**, 9396-9409 (2001).
[96] D. W. M. Hofmann, L. Kuleshova, B. D'Aguanno, *A new reactive potential for the molecular dynamics simulation of liquid water*, Chem. Phys. Lett. **448**, 138-143 (2007).
[97] A. D. Mackerell, Jr., *Empirical force fields for biological macromolecules: overview and issues*, J. Comput. Chem. **25**, 1584-1604 (2004).
[98] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *A second generation force field for the simulation of proteins, nucleic acids, and organic molecules*, J. Am. Chem. Soc. **117**, 5179-5197 (1995).
[99] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *Development and testing of a general amber force field*, J. Comput. Chem. **25**, 1157-1174 (2004).
[100] A. D. MacKerell, D. Bashford, M. Bellott, R. Dunbrack, J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. a. Ha, *All-atom empirical potential for molecular modeling and dynamics studies of proteins*, The journal of physical chemistry B **102**, 3586-3616 (1998).
[101] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids*, J. Am. Chem. Soc. **118**, 11225-11236 (1996).
[102] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - The RESP Model*, J. Phys. Chem. **97**, 10269-10280 (1993).
[103] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *Comparison Of Simple Potential Functions For Simulating Liquid Water*, J. Chem. Phys. **79**, 926-935 (1983).
[104] R. W. Zwanzig, *High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases*, The Journal of Chemical Physics **22**, 1420-1426 (1954).
[105] J. G. Kirkwood, *Statistical mechanics of fluid mixtures*, The Journal of Chemical Physics **3**, 300-313 (1935).

[106] C. H. Bennet, *Efficient Estimation of Free Energy Differences from Monte Carlo Data*, J. Comput. Phys. **22**, 245-268 (1976).
[107] C. Bartels, *Analyzing biased Monte Carlo and molecular dynamics simulations*, Chem. Phys. Lett. **331**, 446-454 (2000).
[108] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, W. F. van Gunsteren, *Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations*, Chem. Phys. Lett. **222**, 529-539 (1994).
[109] A. Barducci, M. Bonomi, M. Parrinello, *Metadynamics*, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **1**, 826-843 (2011).
[110] A. F. Voter, *Parallel replica method for dynamics of infrequent events*, Physical Review B (Condensed Matter) **57**, R13985-13988 (1998).
[111] A. Voter, *Hyperdynamics: Accelerated Molecular Dynamics of Infrequent Events*, Phys. Rev. Lett. **78**, (1997).
[112] A. Laio, M. Parrinello, *Escaping free-energy minima*, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562-12566 (2002).
[113] A. Laio, F. L. Gervasio, *Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science*, Rep. Prog. Phys. **71**, (2008).
[114] A. Barducci, G. Bussi, M. Parrinello, *Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method*, Phys. Rev. Lett. **100**, (2008).
[115] A. F. Voter, F. Montalenti, T. C. Germann, *Extending the time scale in atomistic simulation of materials*, Annual Review of Materials Research **32**, 321-346 (2002).
[116] M. Greiner, E. Elts, J. Schneider, K. Reuter, H. Briesen, *Dissolution study of active pharmaceutical ingredients using molecular dynamics simulations with classical force fields*, J. Cryst. Growth **405**, 122-130 (2014).
[117] J. Schneider, K. Reuter, *Efficient Calculation of Microscopic Dissolution Rate Constants: The Aspirin-Water-Interface*, J. Phy. Chem. Lett. **5**, 3859 (2014).
[118] Y. Deng, B. Roux, *Computations of Standard Binding Free Energies with Molecular Dynamics Simulations*, J. Phys. Chem. B **113**, 2234-2246 (2009).
[119] J. Wang, Y. Deng, B. Roux, *Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials*, Biophys. J. **91**, 2798-2814 (2006).
[120] H. J. Woo, B. Roux, *Calculation of absolute protein-ligand binding free energy from computer simulations*, Proc. Natl. Acad. Sci. U.S.A. **102**, 6825-6830 (2005).
[121] S. Boresch, F. Tettinger, M. Leitgeb, M. Karplus, *Absolute Binding Free Energies: A Quantitative Approach for Their Calculation*, J. Phys. Chem. B **107**, 9535-9551 (2003).
[122] P. Tiwary, M. Parrinello, *From Metadynamics to Dynamics*, Phys. Rev. Lett. **111**, 230602 (2013).
[123] Y. Kim, K. Machida, T. Taga, K. Osaki, *Structure redetermination and packing analysis of aspirin crystal*, Chem. Pharm. Bull. (Tokyo) **33**, 2641 (1985).
[124] A. W. Sousa da Silva, W. F. Vranken, *ACPYPE - AnteChamber PYthon Parser interface*, BMC Res. Notes **5**, 367 (2012).
[125] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *Automatic atom type and bond type perception in molecular mechanical calculations*, J. Mol. Graphics Modell. **25**, 247-260 (2006).
[126] E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak, F. Y. Dupradeau, *R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments*, Nucleic Acids Res. **39**, W511-517 (2011).
[127] V. Zoete, M. A. Cuendet, A. Grosdidier, O. Michielin, *SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules*, J. Comput. Chem. **32**, 2359-2368 (2011).

[128] W. L. Jorgensen, D. S. Maxwell, J. TiradoRives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids*, J. Am. Chem. Soc. **118**, 11225-11236 (1996).
[129] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *Comparison of simple potential functions for simulating liquid water*, The Journal of Chemical Physics **79**, 926 (1983).
[130] H. J. C. Berendsen, D. van der Spoel, R. van Drunen, *GROMACS: A message-passing parallel molecular dynamics implementation*, Comput. Phys. Commun. **91**, 43-56 (1995).
[131] B. Hess, *GROMACS 4: Algorithms for Highly Efficient,*
Load-Balanced, and Scalable Molecular Simulation, J. Chem. Theory Comput. **4**, 435-447 (2008).
[132] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. Berendsen, *GROMACS: fast, flexible, and free*, J. Comput. Chem. **26**, 1701-1718 (2005).
[133] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler, *Ab initio molecular simulations with numeric atom-centered orbitals*, Comput. Phys. Commun. **180**, 2175-2196 (2009).
[134] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, M. Scheffler, *Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, RPA, MP2 andGWwith numeric atom-centered orbital basis functions*, New Journal of Physics **14**, 053020 (2012).
[135] A. Tkatchenko, M. Scheffler, *Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data*, Phys. Rev. Lett. **102**, 073005 (2009).
[136] C. Møller, M. S. Plesset, *Note on an Approximation Treatment for Many-Electron Systems*, Phys. Rev. **46**, 618-622 (1934).
[137] Y. Kim, K. Machida, T. Taga, K. Osaki, *Structure Redetermination and Packing Analysis of Aspirin Crystal*, CHEMICAL & PHARMACEUTICAL BULLETIN **33**, 2641-2647 (1985).
[138] P. Hartman, W. G. Perdok, *On the Relations Between Structure and Morphology of Crystals. II*, Acta Crystallogr. **8**, 521-524 (1955).
[139] L. Edwards, *The hydrolysis of aspirin. A determination of the thermodynamic dissociation constant and a study of the reaction kinetics by ultra-violet spectrophotometry*, Trans. Faraday Society **46**, 723-735 (1950).
[140] L. Edwards, *The dissolution and diffusion of aspirin in aqueous media*, Trans. Faraday Society **47**, 1191-1210 (1951).
[141] C. Zheng, *Thermodynamics of Defects at the Aspirin-Water Interface*, M.Sc. Thesis, Tehnical University of Munich (Munich), **2014**.
[142] S. Nosé, *A unified formulation of the constant temperature molecular dynamics methods*, The Journal of Chemical Physics **81**, 511 (1984).
[143] T. Darden, D. York, L. Pedersen, *Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems*, The Journal of Chemical Physics **98**, 10089-10092 (1993).
[144] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *A smooth particle mesh Ewald method*, The Journal of Chemical Physics **103**, 8577-8593 (1995).
[145] H. J. C. Berendsen, Gromacs User Manual Version 4.6.3, www.gromacs.org (2013)
[146] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, M. Parrinello, *PLUMED: A portable plugin for free-energy calculations with molecular dynamics*, Comput. Phys. Commun. **180**, 1961-1972 (2009).

[147] P. Hänggi, P. Talkner, M. Borkovec, *Reaction-rate theory: fifty years after Kramers*, Rev. Mod. Phys. **62**, 251 (1990).
[148] Y. Kim, M. Matsumoto, K. Machida, *Specific surface energies and dissolution behavior of aspirin crystal*, Chem. Pharm. Bull. **33**, 4125-4131 (1985).
[149] A. Danesh, S. D. Connell, M. C. Davies, C. J. Roberts, S. J. B. Tendler, P. M. Williams, M. J. Wilkins, *An In Situ Dissolution Study of Aspirin Crystal Planes (100) and (001) by Atomic Force Microscopy*, Pharm. Res. **18**, 299-302 (2001).
[150] M. A. Lovette, M. F. Doherty, *Needle-Shaped Crystals: Causality and Solvent Selection Guidance Based on Periodic Bond Chains*, Cryst. Growth Des. **13**, 3341-3352 (2013).
[151] R. C. Rowe, P. J. Sheskey, M. E. Quinn, *Handbook of Pharmaceutical Excipients*, 6th ed., The Pharmaceutical Press and American Pharmacists Association, London and Washington DC, 2009, (pages
[152] S. L. Raghavan, R. I. Ristic, D. B. Sheen, J. N. Sherwood, *Dissolution Kinetics of Single Crystals of alpha-Lactose Monohydrate*, J. Pharm. Sci. **91**, 2166-2174 (2002).
[153] J. H. Smith, S. E. Dann, M. R. J. Elsegood, S. H. Dale, C. G. Blatchford, *Alpha-Lactose Monohydrate: A Redetermination at 150 K*, Acta Crystallogr., Sect. E: Struct. Rep. Online **61**, o2499-o2501 (2005).
[154] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields*, The Journal of Physical Chemistry **98**, 11623-11627 (1994).
[155] S. L. Raghavan, R. I. Ristic, D. B. Sheen, J. N. Sherwood, *Morphology of Crystals of alpha-Lactose Hydrate Grown from Aqueous Solution*, J. Chem. Phys. **104**, 12256-12262 (2000).

Acknowledgements

I would like to first thank Prof. Dr. Karsten Reuter for accepting me to this fascinating albeit sometime frustrating project. I appreciate his support and suggestions related to science or not. Of course this work would not be completed without the ideas, support and even guidance of Dr. Julian Schneider. It was a delight to work and discuss ideas with him. In addition many thanks to Chen Zheng or as he would like to be called Frank for especially for his work regarding the aspirin part of this work.

Next, I would like to thank my family; my mother Birsel Doğan, my father Ahmet Doğan and my sisters Ebru Demirkaya and Beyzanur Doğan. My parents' support and advice sometimes idle chat were the only things to help me in frustrating times. I am grateful for their endless and unconditional love. Also I am very happy for their infinite patience with me while trying to decide what to do next and where to go. Though I know they didn't always agree with my decisions, they still supported me so thanks for being the best parents ever. Of course I am also grateful for my awesome sisters' support and love and thank them from the bottom of my hearth.

I would also like to specifically express my gratitude to our secretary Ruth Mösch for her determination to her work and her helps relating to any problems I have regardless of considering if it is her job description. Technical support from Dr. Christoph Scheurer, Dr. Max Hoffman and Christoph Schober is gratefully acknowledged.

All the members of our group in München are all remarkable and I am very happy to meet them all. Specifically I would like to thank Dr. Tongyu Wang for being a really nice friend to me, office mate and person in general. Of course, my special thanks also go to my other office mates Dr. Max Hoffman, Christoph Schober and Georg Michelitsch for some interesting discussions and maintaining a quiet and relatively tidy office space. The lunch times especially in the first half my doctorate would not be so much fun without Dr. Reini Maurer, Dr. Jelena Jelic, Vanessa Bukas, Dr. Tuğba Davran-Candan, Dr. Katherina Diller and many more people. In addition to our lunch group, I would also like to thank Dr. Dennis Palagin and Dr. Daniel Berger and again to Prof. Dr. Karsten Reuter for making me feel welcome during my first short stay in München for the job-interview.

Of course I cannot forget about my dear friends since high school Işıl Kaya, Funda Özdemir and Emel Gültekin and my college friends Esra Cabalar Balcı and Merve Coşar our shared memories. I am thankful to be able to call them my friends and just for their understandings. My flat mates and especially Reyhan Korkmaz were one of the greatest motivators for me and an

excellent cook and baker. I am happy to have shared a home with her and also to call them my friends also. I thank many of my Turkish friends to make me feel more at home in Germany and helping me with especially translations of documentation and other paper works.

Finally, I should mention the financial support by Deutsche Forschungsgemeinschaft is gratefully acknowledged, as is generous CPU time at the Leibniz Rechenzentrum der Bayerischen Akademie der Wissenschaften as well as at Partnership for Advanced Computing in Europe (PRACE) for awarding access to Curie at TGCC, France. The funding by TUM Gradutae School is also acknowledged and much appreciated.