

Abstract

Focal adhesion kinase (FAK) is a tyrosine kinase associated to focal adhesions with a downstream effect on, for example, cell migration and surviving. Its activation is linked to integrin signalling whereby phosphatidylinositol-4,5-bisphosphate ($\text{PI}(4,5)\text{P}_2$) acts as a moderator. $\text{PI}(4,5)\text{P}_2$ has allosteric effects on FAK and is required for activation. Moreover, it induces clustering of FAK molecules, but the impact of clustering on the conformation of FAK is still not understood. Since an overexpression of FAK is associated with invasive tumours, insights in the clustering process could give rise to new cancer treatments. In this project, we performed MD simulations with the coarse-grained force field MARTINI. In previous studies, an unnaturally falling of FAK in MARTINI was observed. Although we are able to exclude the main binding site of FAK to $\text{PI}(4,5)\text{P}_2$ as the problem, the cause could not be identified. Therefore, we introduced a stabilizing force acting on the FAK molecules. Our observations support the conclusion that FAK clusters arrange to chain-like structures. However, we did not observe a significant change of activation-associated quantities for protein instances inside the clusters.

Focal adhesion kinase (FAK) ist eine Tyrosinkinase, die vermehrt in Fokalen Adhäsionen zu finden ist und regulierende Effekte auf z.B. Zellmigration und Zellüberlebensfähigkeit hat. Die Aktivierung von FAK wird u.a. durch Signalgebungen von Integrin-Molekülen hervorgerufen, wobei phosphatidylinositol-4,5-bisphosphate ($\text{PI}(4,5)\text{P}_2$) als Mediator auftritt. $\text{PI}(4,5)\text{P}_2$ bewirkt allosterische Effekte in FAK und ist notwendig für die Aktivierung. Zudem induziert $\text{PI}(4,5)\text{P}_2$ die Bildung von Clustern. Allerdings ist der Einfluss der Clusterbildung auf die Konformation von FAK noch nicht vollständig verstanden. Da in invasiven Tumoren häufig eine überdurchschnittlich hohe Konzentration von FAK zu finden ist, könnte ein tieferes Verständnis der Clusterbildung ein wichtiger Schritt auf dem Weg zu neuartigen Krebstherapien sein. Die vorliegende Arbeit stützt sich auf MD Simulationen mit dem Kraftfeld MARTINI. In vorangegangenen Untersuchen wurde ein unnatürliches Umfallen von FAK auf die Seite beobachtet. Obwohl wir die Hauptbindungsstelle von FAK zu $\text{PI}(4,5)\text{P}_2$ als Problem ausschließen können, konnte der tatsächliche Grund des Fallens nicht identifiziert werden. Daher führten wir eine Stabilisierungskraft ein, die auf die FAK-Moleküle wirkt. Unsere Beobachtungen lassen den Schluss zu, dass sich FAK-Cluster in kettenartigen Strukturen bilden. Allerdings konnte keine signifikante Änderung von Observablen, die mit Aktivierung von FAK assoziiert werden, für Proteine aus solchen Strukturen festgestellt werden.

1 Introduction

Focal adhesions (FA) are macromolecular protein complexes which act as a connection hub between the cell, especially the cytoskeleton, and the extracellular matrix (ECM). They enable the cell to exert tension forces, but can also transduce mechanical stimuli from the ECM to the inside of the cell and integrate this information. One important protein associated to FAs is the focal adhesion kinase (FAK). FAK occurs in several signalling pathways and is a key player in integrating extracellular stimuli. It is of large interest not least because often an overexpression of FAK can be found in cancer cells, and understanding the activation processes and dynamics of FAK could give rise to new cancer treatments [41].

1.1 Structure

FAK consists of four domains: (i) a FERM (4.1 protein, ezrin, radixin and moesin) domain at the N-terminus, (ii) a tyrosine kinase, (iii) a proline-rich region and (iv) a focal adhesion targeting (FAT) domain at the C-terminus (see Figure 1.1).

FERM is a common protein domain which links proteins to membranes by binding to various phospholipids [10] and consists of three subdomains: F1, F2 and F3. In the F2 subdomain, a basic patch ($^{216}\text{KAKTLRK}^{222}$) can be found, which is a prominent binding site for phosphatidylinositol-4,5-bisphosphate (PI(4,5)P₂).

The kinase consists of a C-lobe, an activation loop and an N-lobe. Catalytic activity of the kinase is mainly regulated by the phosphorylation of Y⁵⁷⁶ and Y⁵⁷⁷, which are located in the activation loop [8]. The kinase also provides binding sites for PI(4,5)P₂. One is located next to the basic patch of the FERM domain, but others (namely R⁵⁰⁸, R⁵¹⁴, K⁵¹⁵, K⁶²¹ and K⁶²⁷) can be found on the side of the kinase [14, 18].

The FERM domain and the kinase are connected by a linker region. In contrast to other kinase domains, the main autophosphorylation site of FAK, Y³⁹⁷, can be found in this region and not in the kinase itself [17].

The FAT domain is linked to the kinase by a flexible proline-rich region. FAT links to FAs by interacting with talin and paxillin which are integrin-associated proteins [2].

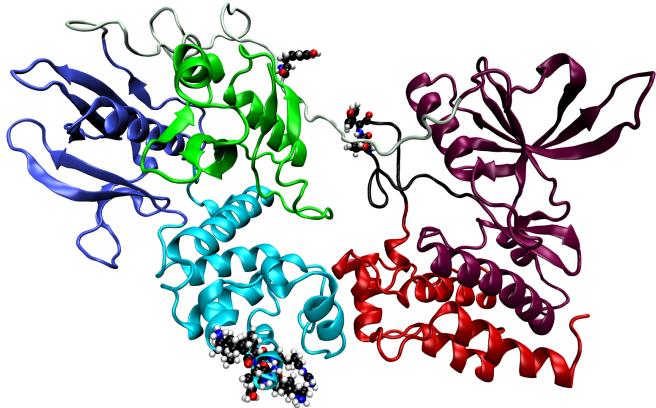


Figure 1.1: **Structure of FAK**

The FERM domain consisting of F1 (green), F2 (iceblue) and F3 (blue) connected via the linker to the kinase consisting of N-lobe (violet), activation loop (black) and C-lobe (red). The basic patch (F2), Y³⁹⁷ (linker) and Y⁵⁶⁷, Y⁵⁷⁷ (activation loop) are shown in atomistic representation.

1.2 PI(4,5)P₂ binding and activation

It is known that integrin signalling due to cell adhesion triggers activation of FAK and that PI(4,5)P₂ is an important mediator in this signalling pathway. FAK interacts in several ways with PI(4,5)P₂, which is discussed in the following.

In the autoinhibited conformation, the FERM domain shields the active site of the kinase. An activation of FAK therefore requires the dissociation of the FERM and the kinase [28]. PI(4,5)P₂ is a phospholipid which is locally generated in FAs due to integrin signalling [27]. Its charge depends strongly on the pH, but at physiological conditions a net charge of -4 is the preferred state. However, in presence of basic amino acids (Arg, Lys), the deprotonated state gets promoted, resulting in a net charge of -5 [32]. The electrostatic binding of PI(4,5)P₂ to the basic patch in the F2 subdomain results in allosteric changes, which also influence the interface between the F1 subdomain and the N-lobe. Moreover, the linker region gets less strongly bound, which promotes autophosphorylation of Y³⁹⁷. However, the PI(4,5)P₂ binding alone has no effect on the catalytic activity, which suggests that the FERM domain is still associated to the kinase [17, 42]. For activation, an additional stimulus, either biochemical or mechanical, is needed.

If Y³⁹⁷ is phosphorylated, it becomes a suitable binding site for SH2 or SH3, which are subdomains of proteins of Src family tyrosine kinases. Due to the conformational changes induced by PI(4,5)P₂, this kinase has access to Y⁵⁷⁶ and Y⁵⁷⁷. As said, the phosphorylation of these residues makes FAK fully active, resulting in dissociation of the FERM domain and the kinase [17].

Mechanical forces can lead to dissociation of the FERM domain from the kinase as well. Forces acting on the FAT domain are transduced to the interface of the FERM domain and the kinase because the linker is connected to the kinase, while the FERM domain is anchored onto a PI(4,5)P₂ containing membrane. These forces can lead to activation of FAK. In that way, it is acting as a mechanical sensor [43].

The binding sites on the side of the kinase were identified by computer simulations [14] and have been confirmed recently in experiments [18]. The findings from Hall and Schaller [18] show that these residues are required for catalytic activity of the kinase, and that they bind to phospholipids *in vivo*. However, since the catalytic activity is not regulated by PI(4,5)P₂, this binding was hypothesized to act as a stabilisation of the active state only [18].

1.3 Dimerization, clustering and autophosphorylation

Autophosphorylation of Y³⁹⁷ is an important event in FAK activation. It has been shown that this happens in intact cells *in trans*, for which a self-association of FAK is required [38].

The FERM domain induces a dimerization of FAK by interacting with the FERM domain of a neighbouring FAK molecule. The interaction emerges around W²⁶⁶ in the connected FERM domains and is stabilised by an interaction of the FAT domain with the basic patch of the other FERM domain, respectively. The presence of W²⁶⁶ is also required for autophosphorylation of FAK.

PI(4,5)P₂ is not needed for the dimerization of the FERM domains. However, an enriched FAK concentration is needed to observe FAK dimers in cells, which is the case at FAs. It is still unclear how the dimer is stabilised at membranes, where the basic patch is also required for ligand binding [7].

It has been shown that *in trans* autophosphorylation of FAK is promoted by dimerization [24].

Albeit PI(4,5)P₂ is not required for dimerization, it induces clustering of several FAK molecules on the membrane *in vitro* [17]. Because dimers support autophosphorylation of FAK, it is not surprising that the same effect is observed in clusters. However, these clusters can trigger additional biochemical stimuli [24] and may play an important role in the scaffolding function of FAK for FAs [17].

2 Motivation

As described in section 1.3, intermolecular interactions of FAK molecules can effect their activation. However, the process is still not understood, especially on an atomistic scale. In this thesis, we used MD simulations with the MARTINI force field to investigate these interactions. MARTINI lacks chemical details, but it is a necessary simplification since systems with several FAK molecules involve a large number of particles. Unfortunately, previous work in the group revealed problems in the use of MARTINI regarding simulations of FAK, which are summarized in the following [3].

Becker et al. [3] obtained in simulations of a single FAK molecule on a PI(4,5)P₂-containing membrane rapid changes in the inclination of the protein with respect to the membrane. In the following, this inclination is characterized by the angle β between the z-axis and the vector connecting F1 and F2, \vec{d}_F . β is given as

$$\cos(\beta) = \frac{\vec{d}_{F,z}}{d_F}, \quad d_F = |\vec{d}_F| \quad (2.1)$$

The distributions of β for different simulations (10 μ s each) are presented in Figure 2.1a. Considering the red line, a mean value of 90° can be observed. The FAK molecule changed from an upright position to this state in less than 50 ns and stayed constant for the remaining simulation time. We refer to this behaviour as a fall of FAK in the following.

There are several reasons why this is rather an artefact of the MARTINI force field than a possible binding pose of FAK to the membrane as suggested by Feng and Mertz [14]. First, FAK fell to both sides, which means that the interaction sites for PI(4,5)P₂ proposed by Feng and Mertz [14] were also located on top of the protein instead of at the protein-membrane interface. Indeed, contact analysis between the protein and PI(4,5)P₂ lipids showed that virtually all residues on the surface (in both, FERM domain and kinase) were interacting with the membrane. A second reason is that this behaviour was not observed in equivalent all-atom simulations in CHARMM36 (1.5 μ s in total). Here, only two maxima were observed around 8° and around 20°, the largest observed angle in all frames was 40°.

In the course of this project, we carried out several simulations to understand the cause of this falling. However, we were not able to prove the reason beyond doubt. In order to still perform reasonable simulations of multiple FAK molecules, we decided to apply

a stabilising force on each FAK molecule.

The force is acting onto F1 and F2 parallel to the z-axis and is proportional to the deviation of their z-distance Δz from a reference distance z_0 . This is illustrated in [Figure 2.1b](#). For the determination of z_0 we took only the green and the blue distribution from [Figure 2.1a](#) into account, because the large angles observed in the other distributions have not been observed in the CHARMM36 simulations. The mean value of $\vec{d}_{F,z}$ for these two distributions is 2.228 nm, which we therefore set as z_0 . From the standard deviation of $\vec{d}_{F,z}$, $\Delta z = 0.11$ nm, we estimated the force constant equivalent k_i of the fluctuation around the mean angle:

$$k_B T = \frac{1}{2} k_i \Delta z^2 \quad (2.2)$$

$$k_i = \frac{2k_B T}{\Delta z^2} = 412 \frac{\text{kJ}}{\text{mol nm}^2}. \quad (2.3)$$

In order to not sharpen the distribution, we chose $k = 100 \frac{\text{kJ}}{\text{mol nm}^2}$ for the stabilising force.

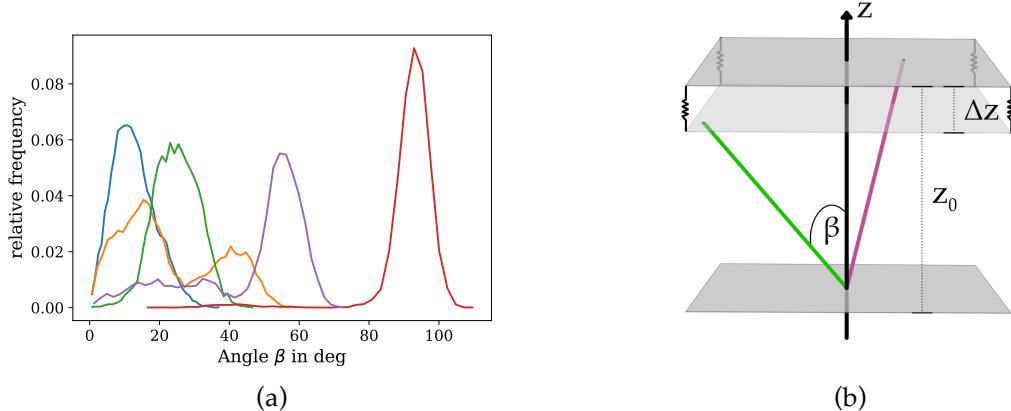


Figure 2.1: Inclination angle of FAK

(a): Distributions of β obtained by Becker et al. [3]. The red curve shows the distribution for a fallen FAK. The blue and green curve are consistent with equivalent all-atom simulations in CHARMM36. (b): Illustration of the stabilising force. The violet line represents a orientation of \vec{d}_F which wouldn't lead to a force ($\Delta z = 0$). The green line represents a orientation which leads to an uplift ($\Delta z > 0$, the springs are tensioned). The force is proportional to Δz .

3 Methods

Before we can present the used setups and the obtained results, the most important methods used in the course of the project are explained. These involve molecular dynamics simulations (MD) as well as umbrella sampling. Furthermore, the concepts used for the classification of multiple FAK interactions are introduced.

3.1 Molecular dynamics simulations

Because the measurable time and length scale of biological experiments are usually larger than those of the system, the dynamics in the system can only be measured indirectly. To get a more precise insight into the specific processes, molecular dynamics simulations (MD) are a suitable tool.

In the following section, the main concepts of MD and characteristics of the used models are outlined with a special focus on the MARTINI force field. An exhaustive exploration of these methods, however, would be outside the scope of this thesis. Because GROMACS (GROningen MACHine for Chemical Simulations) [1, 4] was used as MD engine in this thesis, the following explanations refer to GROMACS conventions and features.

3.1.1 The physics of MD

In MD, atoms are modelled as solid beads, which comprise the nucleus as well as the electron shell, and follow classical equations of motion. Due to this implicit treatment of electrons, quantum mechanical (QM) effects, such as excitation or electron transfer processes, are not accessible in MD. However, the beads are parametrized by effective parameters motivated by QM or experiments [33, p. 127-128].

Newton's equation of motion can be turned into two first order differential equations:

$$\frac{d\vec{r}_i}{dt} = \vec{v}_i \quad (3.1)$$

$$m_i \frac{d\vec{v}_i}{dt} = \vec{F}_i, \quad (3.2)$$

which can be integrated numerically with e.g. Leapfrog- or Verlet integration scheme. Both are time reversible and symplectic, which ensures a small long-term error in energy

conservation [16, p. 72-74]. The force acting on particle i is given by the gradient of the potential V at its position. In MD, bonded and non-bonded interactions contribute to this potential, but also external forces can be applied:

$$\vec{F}_i = -\frac{\partial V}{\partial \vec{r}_i} = -\frac{\partial}{\partial \vec{r}_i} (V_{\text{bonded}} + V_{\text{non-bonded}} + V_{\text{external}}). \quad (3.3)$$

Bonded interactions

Bonded interactions act intramolecularly and describe chemical bonds. They can occur between two, three or four particles and refer to bond stretching, bending and torsion respectively. An illustration can be found in [Figure 3.1a](#).

Deviations of the bond length r from an equilibrium distance r_0 result in potential energy, which is usually described by a harmonic oscillator with the force constant k_{dist} :

$$V_{\text{dist, bond } i} = \frac{k_{\text{dist}}}{2} (r - r_0)^2. \quad (3.4)$$

For larger deviations, a Morse potential, assuming an exponential decay of the potential energy, is more precise, but has a much higher computational cost.

Bending of a chemical bond refers to deviations of the angle θ between three bonded partners from an equilibrium angle θ_0 . The resulting potential is usually described by a harmonic oscillator as well:

$$V_{\text{angle}} = \frac{k_{\text{angle}}}{2} (\theta - \theta_0)^2. \quad (3.5)$$

The dihedral angle describes the angle between two planes, each going through three beads, and having two beads in common. Therefore, it can be understood as a torsion angle, but can also be used to preserve plane rings and the chirality of four particle groups. The resulting potential energy is usually approximated with a periodic approach:

$$V_{\text{dihedral, periodic}} = \frac{k_{\text{dihedral}}}{2} (1 + \cos(n\phi - \phi_0)), \quad (3.6)$$

where k_{dihedral} describes the energy barrier for turning the dihedral angle, n the number of minima in the energy function (multiplicity) and ϕ_0 a phase factor [1, p. 71-83].

Non-bonded interactions

Non-bonded interactions are present between all atoms in the system and act pairwise. In MD, Pauli repulsion, van der Waals (vdW) forces and electrostatic forces are taken into

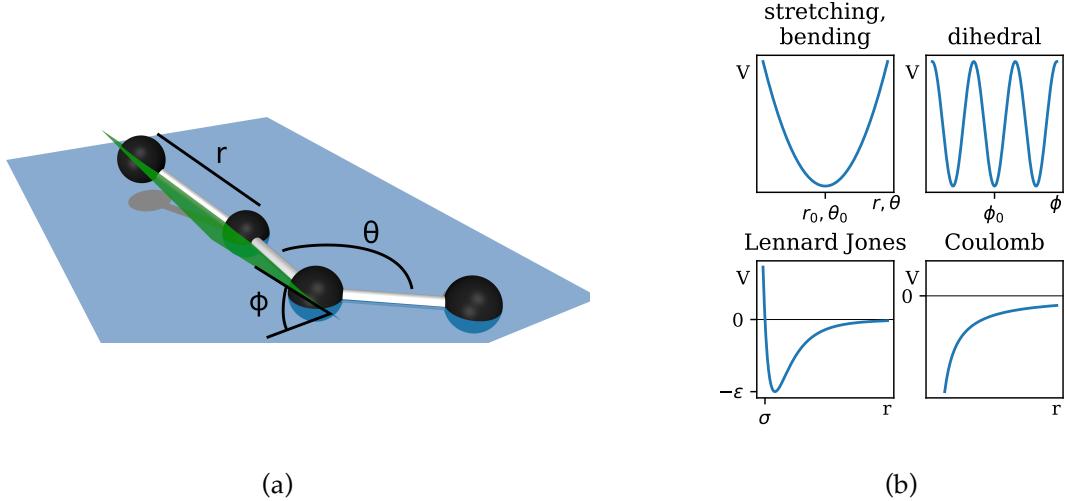


Figure 3.1: **Interactions in MD**

(a): Illustration of bonded interactions: bond stretching (r), bending (θ) and dihedral angle Φ (angle between blue and green plane). (b): schematic course of contributing potentials

account. The Lennard-Jones potential combines Pauli repulsion (r^{-12} term) and the vdW force (r^{-6} term):

$$V_{\text{Lennard-Jones}} = \sum_{\text{non-bonded pairs } i,j} 4\epsilon \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right). \quad (3.7)$$

ϵ is related to the potential depth and σ to the potential range (see Figure 3.1b).

The Coulomb potential is given by

$$V_{\text{Coulomb}} = \frac{q_1 q_2}{4\pi\epsilon_0\epsilon_r r}, \quad (3.8)$$

where q_1, q_2 are the (partial) charges of the interacting particles, r their distance and ϵ_r the relative dielectric constant [1, p. 65-71].

In general, non-bonded interactions act between all atoms in the system implying a very large computational cost.

The easiest solution for this problem is to use a cut-off radius r_c . Particles beyond this radius are not taken into account. This can be implemented very efficiently with a Verlet neighbour lists. For each particle, a neighbour list is created, which contains all particles inside a second radius r_v with $r_v > r_c$. This reduces the number of distance calculations a lot. The lists are updated if the maximal displacement in the system is larger than $r_v - r_c$. The application of a cut-off radius is suitable especially for Lennard-Jones potential, as it decays very rapidly and r_c can be chosen very small [33, p. 144].

Because the electrostatic potential is proportional to $1/r$, the use of a cut-off radius would

lead to large jumps in the potential. That is why long range interactions have to be considered, which can be effectively done with Particle Mesh Ewald (PME) summation [11]. Particle mesh methods in general split up the electrostatic potential into a short-range and a long-range part via a switching function. The short-range part can be calculated with a small cut-off radius in real space. The long-range part, however, is calculated by solving the Poisson equation of the actual charge distribution, for which a discrete grid (mesh) is used. In contrast to other particle mesh methods, this grid is transformed to Fourier space in PME with FFT techniques. Here, the solution of the Poisson equation is a sum over the grid points. Afterwards, the potential can be transformed back into real space. The use of PME requires of course periodic boundary conditions, which are described in section 3.1.3 [33, p. 246-251].

External forces

With GROMACS, it is possible to perform pulling forces onto groups of atoms in the system. In this thesis, pulling was used to bias distances between groups. For this, GROMACS provides an option to apply an umbrella potential to two groups, which yields a force proportional to the deviation of the distance between the groups from a reference distance. The force can be applied on one or two spatial dimensions only or along a pre-defined vector and the reference distance can change in time [1, p. 154-159].

3.1.2 Force fields

The parameters for the potentials described above are provided by force fields. There is a wide range of force fields, which are optimized for different application fields. In this thesis, two different ones, CHARMM36 and MARTINI, are employed.

Force fields map atoms (particles in physical system) onto beads (particles used in the simulation). This mapping can take the chemical environment of the particle into account, but can also neglect details by e.g. mapping several atoms onto one bead (coarse graining).

Force fields can be distinguished in non-polarizable and polarizable force fields. The latter offer the possibility to model electronic polarization, e.g. with additional shell beads. Orientational and geometric polarisation are also accessible with non-polarizable force fields since they only rely on partial charges [33, p. 215-217]. Both, MARTINI and CHARMM36, are non-polarizable force fields.

Force fields not only define the atom types and their properties, but also all parameters for the calculation of the potential, namely force constants, equilibrium distances, etc. Therefore, force fields define the physics of the system.

All-atom and the CHARMM36 force field

CHARMM36 [6, 25] (C36) is part of the CHARMM (*Chemistry at HARvard Macromolecular Mechanics*) MD engine and was published in 2010. In C36, all atoms are considered (all-atom force field). Parameters were mainly optimized to structural experimental data, such as nuclear magnetic resonance (NMR) or X-ray data, but also QM and semi-empirical QM calculations were used (e.g. for dihedral angles of the sidechains of proteins [6] and partial charges of lipids [25]).

All simulations for optimizing have been done with a time step of 2 fs. Therefore, very detailed dynamics are still included in the simulations.

For all simulations, the TIP3P water model [23] was used, which was also used in the parametrisation of C36. Here, each water molecule is modelled with three partially charged beads.

Coarse graining and the MARTINI force field

The MARTINI force field [31, 22, 29] is one of the most famous coarse-graining force fields. Most beads are constituted of four heavy atoms (see Figure 3.2), which implies a loss of chemical information, but also an enormous reduction of computational cost. With this approach much larger time- and spatial scales are accessible for MD simulations.

The parametrisation of MARTINI is mainly based on reproducing free energies, e.g. partitioning free energies and dimerization free energies of amino acid side chain analogues. For lipids, also thermodynamic properties, such as area per lipid, have been considered. The parameters were optimized with a time step of 20 fs up to 30 fs [22, 29].

There are several side effects of coarse graining. Due to the mapping of four atoms to one bead, different atomistic structures can end up in the same MARTINI structures, which can be problematic e.g. for lipid tails. Also the secondary structure of proteins is affected as it becomes less stable in coarse-grained models. Therefore, it has to be constrained with

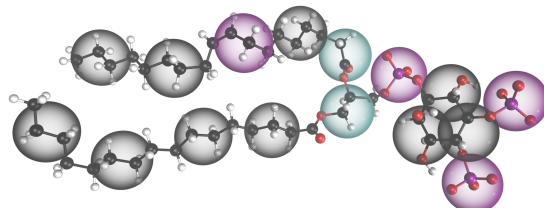


Figure 3.2: **MARTINI bead mapping**

MARTINI structure of PI(4,5)P₂ with underlying atomistic structure. Four heavy atoms are mapped to one MARTINI bead. The colour indicates the bead type.

elastic networks (additional bonds between backbone beads) and cannot change during the simulation [30].

The decrease in degrees of freedom in the system also smooths the energy surface in coarse-grained models. Because smaller local irregularities in the energy surface, which would slow down the evolution of the system, are smoothed out, coarse graining speeds up the dynamics of the system. The speed-up factor is not constant, but can be relatively well approximated by factor of 4 (obtained in most diffusion simulations) [30, 31].

Coarse graining also has an effect on the entropy and the temperature dependency of the system. In NpT ensembles, the Gibbs free energy G is given by

$$G = H - TS, \quad (3.9)$$

where S is the entropy and H the enthalpy. Due to a lower number of degrees of freedom in coarse grained systems, the configurational entropy is reduced. Because MARTINI is tuned to free energy calculations, this implies also a reduction of the enthalpy. The temperature dependency is therefore systematically biased [30].

Consistently with the four to one mapping, solvent beads in MARTINI represent four water molecules. The MARTINI water has a freezing temperature of ca. 300 K and the freezing process is irreversible. Therefore, antifreeze beads were introduced, which have a larger σ_{LJ} regarding interactions with normal solvent beads. By inserting these antifreeze beads, the lattice conformation is disturbed and the freezing temperature reduced [31].

In contrast to water, the solvent beads are not (orientationally) polarizable. For this reason, electrostatic screening has to be treated implicitly, which is done by a relative dielectric constant $\epsilon_r = 15$. However, this approximation fails near polar beads, resulting in an underestimation of the interaction strength between polar beads [30]. To forgo implicit screening, a polarizable water model was introduced in MARTINI, which is denoted as PW in the following. In PW, one water bead consists of three beads, two of them carrying charges. This model enables orientational polarizability and decreases the freezing point as well. This, however, comes with a higher computational cost [40].

As reviewed by Marrink and Tieleman [30], MARTINI has been already used to study protein-membrane interactions as well as protein oligomerization on membranes. Also changes in tertiary structure have been addressed using MARTINI.

3.1.3 External constraints

Periodic boundary conditions

Because the simulation of an open system is not possible, boundary conditions have to be considered. Closed boundaries often lead to surface interaction artefacts and are therefore

in most cases not suitable for MD simulation. That is why usually periodic boundary conditions (PBC) are used.

To use PBC, the shape of the simulation box has to have a space filling geometry (e.g. rectangular or rhombic dodecahedron). With periodic boundary conditions, images of the simulation box are repeated in every direction. If a particle leaves the simulation box, its periodic image is coming in from the opposite side. In that way, the number of particles is kept constant while surface interactions are avoided.

A particle interacts only with the nearest image of another particle, which means that particles near boundaries can interact with periodic images of other particles instead of the real particle. Moreover, molecules can have long-range interactions with their own periodic images, which leads to artefacts and has to be considered when choosing the size of the simulation box [33, p. 141-142].

Thermostats

Integration schemes in MD are designed to conserve the energy of a system. However, in real biological systems, not the energy but rather the temperature is kept constant. This can be achieved with thermostats. For the simulations in this thesis, the Berendsen thermostat [5] (BT), the Parrinello-Bussi thermostat [12] (PBT) and the Nosè-Hoover thermostat [34, 20] (NHT) were used. BT couples the system to a heat bath by rescaling the velocities of the particles, which results in an exponential decay of temperature deviations. The coupling strength is given by the time constant of this decay. The rescaling involves a transfer of kinetic energy from internal degrees of freedom to translational and rotational kinetic energy of the system's COM [19]. In order to prevent this transfer, PBT extends BT by a stochastic term. Both BT and PBT are useful for equilibration runs or non-equilibrium MD simulations, as they are stable upon large deviations [5, 1, p. 31]. On the other hand, NHT extends the Hamiltonian of the system by a friction term representing a heat bath. The friction parameter follows a differential equation depending on the temperature deviation. This ansatz samples the phase space more accurately, making NHT suitable for production runs. However, NHT gets unstable for large deviations [1, p. 32-33].

Often, groups are coupled to independent thermostats. This is helpful because the heat exchange between for instance proteins and solvents is often not correct. Therefore, proteins would cool down and the solvent would heat up [1, p. 34].

Barostats

In biological systems, often not the volume but the pressure is constant, which can be achieved in MD with a barostat. For equilibration runs, the Berendsen barostat [5] was

used, which couples the system, analogously to the BT, to an external pressure by rescaling the positions of the particles [1, p. 36]. For production runs, the Parinello-Rahman barostat [36, 35] was used, because it samples the phase space more accurately. This barostat allows also a rotation of the position vectors. The appropriate matrix follows a differential equation depending on the current deviation of the pressure from the external pressure. However, large deviations lead to oscillations in the box [1, p. 36].

GROMACS provides the possibility to couple the z-direction independently from the x- and y-direction, which is called semiisotropic pressure coupling. This feature is useful for membrane and pulling simulations, because the dynamics differ a lot between these axes [1, p. 36].

3.2 Free energy calculations

In order to understand state transitions in physical systems, the free energy is a key quantity. It is directly linked to the probability distribution for different states and other quantities can be derived easily.

The free energy depends on the partition function and therefore on the considered thermodynamic ensemble. In this section, only the Helmholtz free energy A , which refers to a canonical ensemble, is examined, but the concepts can be applied to other ensembles as well.

Partition function

In the microcanonical ensemble, the probability that a system enters a microstate with energy $E' = E(\mathbf{q}, \mathbf{p})$ (\mathbf{q}, \mathbf{p} are the positions and momenta of the particles respectively), is equal for all $|E' - E| < dE$ and 0 otherwise. Therefore, the partition function Ω is given as

$$\Omega(N, V, E) = C_0 \int \delta(\mathcal{H}(\mathbf{q}, \mathbf{p}) - E) d\mathbf{q}d\mathbf{p}, \quad (3.10)$$

where $\mathcal{H}(\mathbf{q}, \mathbf{p}) = U(\mathbf{q}) + K(\mathbf{p})$ is the Hamiltonian and C_0 a proportionality constant, in which the smallest phase space volume and the indistinguishability of particles have to be taken into account [9, p. 16].

In the canonical ensemble, however, the temperature T is kept constant instead of the energy. Therefore, the partition function has to include all possible energies weighted

with their probability given by the Boltzmann factor:

$$Q(N, V, T) = \int \exp(-\beta E) \Omega(N, V, E) dE \quad (3.11)$$

$$= C_0 \int \exp(-\beta \mathcal{H}(\mathbf{q}, \mathbf{p})) d\mathbf{q} d\mathbf{p}, \quad (3.12)$$

where $\beta = \frac{1}{k_B T}$. The configurational integral Z is defined as

$$Z(N, V, T) = \int \exp(-\beta U(\mathbf{q})) d\mathbf{q}. \quad (3.13)$$

It is important to see that \mathcal{H} only depends \mathbf{p}^2 , so the integral over the momenta can always be solved analytically by turning it into a Gaussian integral. This implies that for two related systems in which the particle masses are the same, the integral over \mathbf{p} does not change, and therefore

$$\frac{Q_2}{Q_1} = \frac{Z_2}{Z_1} \quad (3.14)$$

holds [9, p. 17].

The Helmholtz free energy A is defined as

$$A = -\beta \ln(Q). \quad (3.15)$$

Usually, the exact value of the free energy is unknown, but the main interest lies in free energy differences between two states of a system. The difference ΔA can be written as

$$\Delta A = A_2 - A_1 = -\beta \ln(Q_2) + \beta \ln(Q_1) = -\beta \ln\left(\frac{Q_2}{Q_1}\right) = -\beta \ln\left(\frac{Z_2}{Z_1}\right), \quad (3.16)$$

and the normalisation constant C_0 in Equation 3.10 therefore cancels out.

Free energy in MD and Umbrella Sampling

The free energy of a system as a function of a parameter ξ is given as

$$A(\xi) = -\beta \ln(\rho(\xi)). \quad (3.17)$$

$\rho(\xi)$ is the probability distribution of ξ and can be easily obtained in MD simulation by counting the number of the appropriate visited states. ξ is called reaction coordinate and could be e.g. the distance between two molecules. However, in reality $A(\xi)$ can have big barriers, and because the potential energy U is sharply distributed around its mean in NVT simulations, $\rho(\xi)$ can hardly be sampled at a realistic computational cost.

One way to overcome this sampling problem is umbrella sampling [37]. In this approach, the path ξ is split up into distinct windows $[\xi_0, \xi_n]$. To each window i , a biasing potential

$\hat{U}_i(\xi)$ can be applied to ensure a sufficient sampling of ξ around ξ_i . This changes the potential energy to

$$U_{B,i}(\mathbf{q}) = U(\mathbf{q}) + \hat{U}_i(\xi(\mathbf{q})) \quad (3.18)$$

After a simulation with the biased potential $U_{B,i}$ the unbiased measured probability distribution $\tilde{\rho}_i(\xi)$ has to be reconstructed from the observed biased one, $\tilde{\rho}_{B,i}(\xi)$. At this point, only the result is given; the full derivation can be found in Frenkel and Smit [16, p. 179-181].

$$\tilde{\rho}_i(\xi) = \exp\left(-\beta\left(\Delta A_i - \hat{U}_i(\xi)\right)\right) \tilde{\rho}_{B,i}(\xi) \quad (3.19)$$

Of course ΔA_i is not known, but assuming that $\rho(\xi)$ is a continuous function, the results from the single windows can be combined and afterwards normalized.

With this method, however, only two probability distributions can be considered in the overlapping region. Another problem is that the sampling in the tails is usually poor and statistical errors, which propagate through all overlapping regions, can become very large [9, pp. 236-238]. Therefore, umbrella sampling is usually combined with the weighted histogram analysis method (WHAM) [15, 26]. This algorithm is able to combine several probability distributions in one overlapping region and is designed to keep the statistical errors small. The main idea is to combine distributions linearly with an additional weighting factor $\omega_i(\xi)$ to the total probability density $\tilde{\rho}$. The weighting factors ω_i are chosen iteratively in a way that the overall statistical error is minimized.

3.3 Contact Analysis

In the following part the quantities and tools used for contact analysis are briefly introduced.

Intramolecular distances have been analysed with CONAN. This analysis tool measures inter-residue distances and performs statistical analysis on them. CONAN is still under development and not published yet.

The contact area of the FERM-kinase interface can be determined with the solvent accessible surface area (SASA) [13] of the domains involved:

$$CA = \frac{1}{2} (\text{SASA}_{\text{FERM}} + \text{SASA}_{\text{kinase}} - \text{SASA}_{\text{FERM-kinase}}). \quad (3.20)$$

For the calculation of SASA values GROMACS sasa tool was used. The v.d.W radii were adapted to MARTINI beads.

For an appropriate description of intermolecular interactions between FAK molecules,

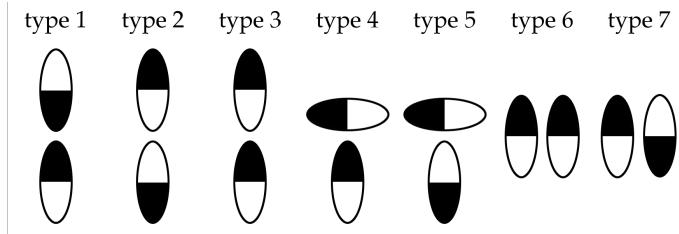


Figure 3.3: **Neighbouring types**

The black part refers to the FERM domain, the white to the kinase.

we defined the following terms:

Interaction Proteins or part of proteins interact if their minimal distance is smaller than a cut-off distance (here 1.5 nm).

Neighbour Two proteins are neighbours if they are interacting with each other. One protein can have several neighbours. For a more detailed characterisation we defined the following neighbouring types (see Figure 3.3):

type 1: only the FERM domain interacts with only the FERM domain of the neighbouring protein

type 2: only the kinase interacts with only the kinase of the neighbouring protein

type 3: only the FERM domain interacts with only the kinase of the neighbouring protein

type 4: the FERM domain is interacting with both, the FERM and kinase of the neighbouring protein

type 5: the kinase is interacting with both, the FERM and kinase of the neighbouring protein

type 6: the FERM domain is interacting with the FERM domain of the neighbouring protein and the kinase is interacting with the kinase of the neighbouring protein

type 7: the FERM domain is interacting with the kinase of the neighbouring protein and the kinase is interacting with the FERM domain of the neighbouring protein

It is important to see that the probability of the different types is initially biased. Assuming that both proteins can have four different orientations (FERM on top, on right, on bottom or on left) and that the proteins do not influence the orientation of their neighbours 16 permutations of the two orientations are possible. Only one permutation would result in a type 1 neighbour (FERM domains pointing towards each other), but, for example, two permutations would result in a type 3 neighbour (both FERM domains point to the right or both point to the left). **Cluster** Neighbouring proteins form a cluster. A protein belongs to a cluster if it has at least one neighbour inside the cluster. One protein can only belong to one cluster. The clustersize is the number of proteins belonging to the cluster.

4 Setup

4.1 Protein structure

All simulations were carried out with starting configurations adapted from previous work in the group (C36 forcefield: Zhou et al. [42], MARTINI forcefield: Becker et al. [3]). These configurations contain only a FERM-kinase fragment without the FAT domain and its linker (only residues 35 to 686, PDB 2J0J [28]).

As explained in section 3.1.2, the secondary structure of proteins has to be stabilized in MARTINI using elastic networks. This was set up by Becker et al. [3] to act only between backbone beads of the same domain which are within a cut-off radius of 1 nm. Therefore, the interface between FERM domain and kinase is not affected and the linker is still flexible. The force constant is $830 \text{ kJmol}^{-1}\text{nm}^{-2}$.

4.2 Setup 1 - FAK in solution

Setup 1 refers to a MARTINI simulation of a single FAK molecule in waterbox. NaCl ions were added to neutralize the charge (see Figure 4.1).

After a short equilibration the system was simulated for $20 \mu\text{s}$ at a temperature of 300 K. We used the default parameters of the MARTINI forcefield as input parameters.

4.3 Setup 2 - Free energy of basic patch

For this setup only a part of F2 (residues 107 to 219, referred as F2 lobe in the following) was used. The lobe contains the basic patch and has a net charge of -5. Therefore no additional ions were needed.

We placed the lobe as a CHARMM36 structure above a single PI(4,5)P₂ embedded into a phosphatidylethanolamine (POPE) membrane (see Figure 4.2). After a short equilibration, the F2 lobe was pulled slowly away from the membrane using a distance pull between the COM of the lobe and the COM of PI(4,5)P₂. From this simulation, we retrieved starting conformations for the umbrella window. The number of umbrella windows was chosen accordingly to the sampling (between 90 and 120 windows). Each window was shortly

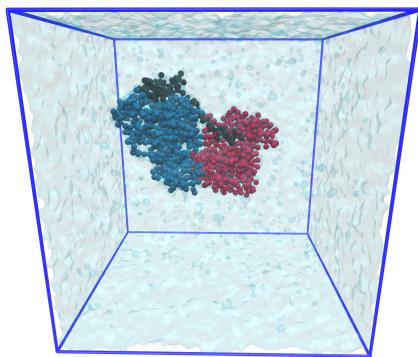


Figure 4.1: **Setup 1 - FAK in solution**

The MARTINI structure of FAK (FERM blue, kinase red and linker black, 1486 beads) was put in a waterbox (ca. 50000 solvent beads) with ions (23 sodium beads and 20 chlorine beads). The box dimensions are 18.8 nm x 18.8 nm x 18.8 nm.

equilibrated and afterwards simulated for 6 ns. From the trajectories of the umbrella windows, the free energy profile was calculated using GROMACS WHAM implementation [21]. The pulling and umbrella sampling was carried out five times to estimate the statistical error.

The starting configuration was transferred to a MARTINI structure (with both standard water model and PW) with the provided transformation tools [39]. Afterwards the elastic network was applied. Analogously to the simulation in CHARMM36, we retrieved starting configurations for the umbrella windows and simulated them for five independent copies. The simulation time for one umbrella window was increased to 10 ns.

The presented results are based on a total simulation time of 3.88 μ s for C36, 6.33 μ s for MARTINI, 5.64 μ s for MARTINI with PW. The temperature in the simulations was kept at 300 K. For all three force fields the default parameters were used.

4.4 Setup 3 - FAK on a PI(4,5)P₂ membrane

Setup 3 is a MARTINI simulation adopted from Becker et al. [3]. It contains a single FAK molecule which was placed on a phosphatidylcholine (POPC) and PI(4,5)P₂ membrane (PI(4,5)P₂ concentration 15%). NaCl ions were added to neutralize the system (see Figure 4.3). In contrast to Becker et al. [3], we applied the stabilizing force explained in chapter 2 to the protein.

Five independent copies were simulated for 10 μ s each. The temperature was kept at 300K and the default parameters of MARTINI were used.

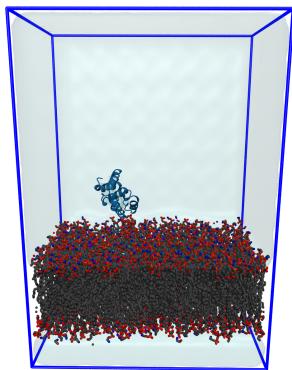


Figure 4.2: **Setup 2 - Free energy of basic patch**

The CHARMM36 structure of the lobe (1936 atoms) above one PI(4,5)P₂ (143 atoms) embedded in a POPE membrane (509 lipids, ca. 63600 atoms). The box was filled up with ca. 63000 water molecules. The corresponding MARTINI structure contains 269 beads for the lobe, ca. 6100 beads for the membrane and ca. 16200 solvent beads/PW molecules. The box dimensions are 14.0 nm x 9.0 nm x 20.0 nm.

4.5 Setup 4 - FAK cluster

From each copy in setup 3, we cut out five frames. All these 25 frames were arranged on a 5x5 grid ([Figure 4.4](#)). Each of the 25 proteins was stabilized with the external force independently. After a short equilibration the system was simulated for 9 μ s. We set up 25 different copies (regarding the arrangement of the frames) which resulted in a total simulation time of 45 μ s. The temperature was kept at 300 K and the default parameters of MARTINI were used.

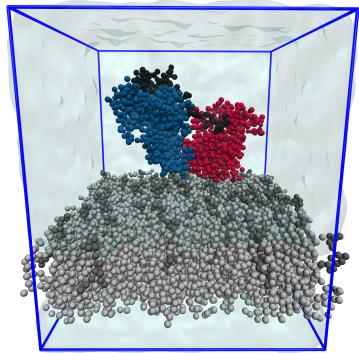


Figure 4.3: **Setup 3 - FAK on a PI(4,5)P₂ membrane**

The setup involves one FAK molecule (FERM blue, kinase red and linker black, 1486 beads) on top of a PI(4,5)P₂ containing membrane (667 POPE lipids, 54 PI(4,5)P₂ lipids, ca. 22000 membrane beads in total). The surrounding solution consists of ca. 21400 solvent beads and 473 ions. The box dimensions are 15.2 nm x 15.2 nm x 17.0 nm.

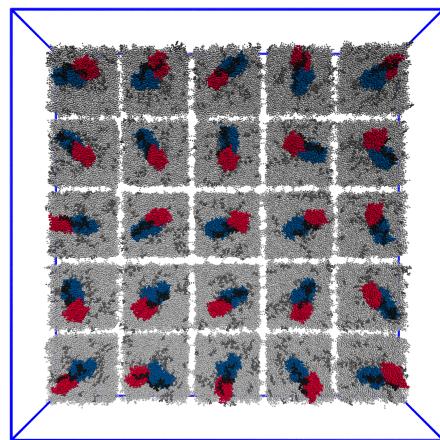


Figure 4.4: **Setup 4 - FAK cluster**

The setup is a combination of 25 frames from section 4.4. It consists of ca. 807000 beads in a box of size 80.0 nm x 80.0 nm x 17.0 nm. The spaces between the frames were needed during setup to avoid an overlapping of the beads, but disappears during the equilibration.

5 Results

In the first part of the project, we investigated the cause of the falling of FAK. Therefore, we want to begin the presentation of our results in [section 5.1](#) with the main finding of these studies, namely that the falling is not due to an underestimation of the binding energy in MARTINI. Afterwards, we analyse the observations of FAK in solution to develop a reference state for later comparisons ([section 5.2](#)). Before we consider FAK molecules on membranes, the applied stabilising force and its impact on used observables is examined in [section 5.3](#). The investigation of FAK molecules on the membrane is split up into two parts. In [section 5.4](#), the obtained conformational changes of FAK molecules bound to a PI(4,5)P₂-containing membrane are presented. In the last section of this chapter, [section 5.5](#), the focus is then on interactions between multiple FAK molecules.

5.1 Free energy profile of basic patch

The basic patch is the main binding site of FAK to PI(4,5)P₂ and the membrane. The fact that the falling of FAK was only observed in MARTINI suggest that this binding is underestimated in the coarse-grained force field. Therefore, we calculated the free energy profile for MARTINI and compared it to equivalent simulations in CHARMM36. In the following, a short report on the results is given.

The profiles are obtained from setup 2. The reaction coordinate is the z component of the COM distance between PI(4,5)P₂ and the protein fragment. For each forcefield, CHARMM36, MARTINI and MARTINI with PW, the average profile out of the five copies together with the standard deviation can be found in [Figure 5.1](#). The free energy profiles were shifted to be zero on average in the range $6 \text{ nm} \leq z \leq 7 \text{ nm}$.

Both force fields, CHARMM36 and MARTINI, show a similar energy depth of $\approx 17 \text{ kCal/mol}$ and a similar slope between $3 \text{ nm} \leq z \leq 4 \text{ nm}$. The slight underestimation of the MARTINI-profile can be attributed to the proposed underestimation of electrostatic forces due to the non-polar water beads (see [section 3.1.2](#)). However, the distance between profiles is smaller than the statistical error and not significant. The difference in the range $4 \text{ nm} \leq z \leq 6 \text{ nm}$ originates from the different treatment of long-range electrostatics: whereas MARTINI uses a cut-off radius, CHARMM36 uses PME.

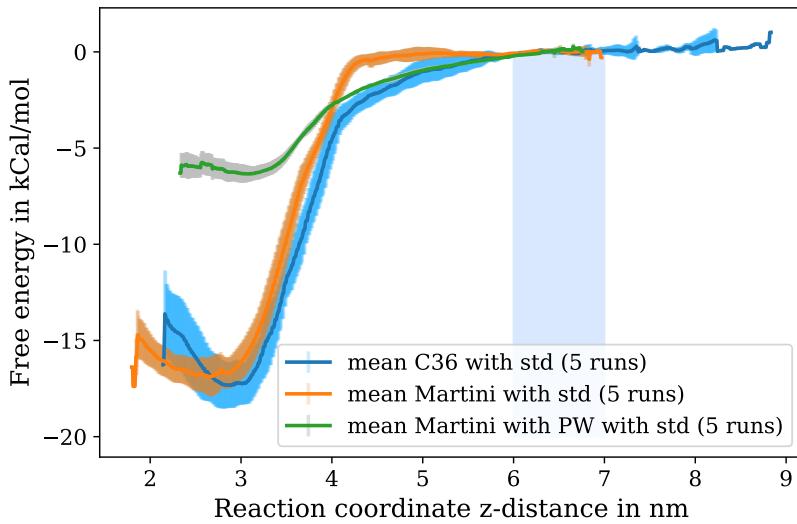


Figure 5.1: Free energy profile of basic patch

For each forcefield the average and standard deviation of the five copies is presented.
All profiles are shifted to be zero on average in the blue region.

Also MARTINI with PW uses PME for long-range electrostatic interactions, and indeed it fits much better to the CHARMM36 profile for larger distances. However, the binding strength is crucially underestimated in Martini with PW.

In the used starting configurations, the proteins are already bound to PI(4,5)P₂. Therefore, a correct binding strength and the shape near the minimum is of larger interest than a correct sampling of farther distances. In addition, MARTINI with PW required a much higher computational effort. Hence, we consider only the standard water model in the following simulations.

5.2 FAK in solution

In this section, we analyse the conformation of solute FAK (FAK-SOL) in order to retrieve a reference state for later comparisons. Since the secondary structure of the two domains is fixed due to the elastic network, the focus is on the FERM-kinase interface.

The first approach to describe the FERM-kinase interface is to consider the COM distances of F1 to the N-lobe (d_{F1-N}) and F2 to the C-lobe (d_{F2-C}).

The two-dimensional histogram of these distances reveals two different states (see Figure 5.2a). Spot 1 refers to conformations which are partially opened at the F2 - C-lobe interface, but close at the F1 - N-lobe interface. In contrast to this, the conformations of spot 2 refers to states in which F2 and the C-lobe get closer while the distance between F1 and the N-lobe is increased. The corresponding 3D structures are shown in Figure 5.3. They imply that spot 1 refers to a configuration in which the kinase is slightly tilted against the FERM domain while it is in line for configurations associated with spot 2.

We observed several transitions between the spots during the simulation, which indicates a sufficiently long simulation time. In total, 47.4% of the obtained distances were located in spot 1 and 52.6% in spot 2.

A second important quantity associated to the FERM-kinase interface is the contact area (CA) of the interface. Yet, the observed states cannot be identified in the distribution of the CA values (see Figure 5.2b). Spot 2 shows only a slightly larger mean value of 27.6 nm^2 as compared to 27.1 nm^2 in spot 1.

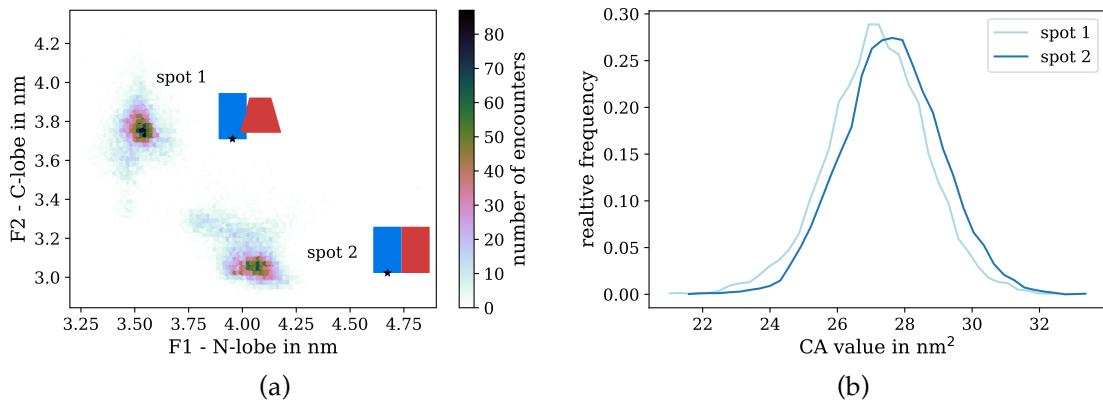


Figure 5.2: Domain distances and contact area of FAK-SOL

(a): two-dimensional hexagonal binning plot of d_{F1-N} and d_{F2-C} , which implies two states. The tilting of the kinase (red) against the FERM domain (blue) is sketched beside (basic patch is marked with star). (b): Distribution of the CA value for the two obtained spots. The mean value of spot 1 is shifted by ca. 2% to smaller values.

Finally, the contact map showing inter-residue distances of the FERM-kinase interface is considered. Both contact maps, for frames of spot 1 and frames of spot 2, show similar features, except for area 2 which is only visible in frames of spot 2. Here, the contacts disappear in spot 1 because the kinase is tilted against the FERM domain. Hence, we only refer to the contact map of spot 2 in the following.

The average contact map for frames in spot 2 is shown in [Figure 5.4](#). Two regions can be identified. The first one (area 1) is located between F1 and the N-lobe/activation loop. It especially shows contacts between Y^{576}/Y^{577} and residues of the FERM domain. The minimal distance in this area, occurring between residue H^{41} and Y^{576} , is 0.45 nm with a RMSF value of 0.03 nm. This area reflects the burying of the activity-regulating residues in the closed state (see [Figure 5.3c](#)).

The second region (area 2, obtained in spot 2 only) is located between F2 and the C-lobe. The spots occur around the residues Y^{180} and D^{200} of F2 as well as F^{596} and R^{665} of the C-lobe. The minimal distance in this area occurs between Y^{180} and F^{596} with 0.45 nm and an RMSF value of 0.02 nm. Mutation experiments showed these two residues, Y^{180} and F^{596} , to have an important effect upon the interface [28], which fits to these observations. The linker shows contacts with both domains. The minimal distances in area 3 and area 4 occur between the autophosphorylation site Y^{397} and H^{58} of F1 (0.45 nm, RMSF 0.03 nm) as well as Y^{397} and Y^{576} of the kinase (0.50 nm, RMSF 0.10 nm). Furthermore, several interactions occur between the residues in the linker itself, namely, between S^{379} to V^{389} and T^{394} to I^{400} (area 5). Also in this area, the minimal distance occurs between Y^{397} and T^{386} (0.52 nm, RMSF 0.04 nm). The density in this area implies that the linker forms a ball, which is slightly plunged into the FERM-kinase interface (regarding area 3 and area 4). This can be seen also in the 3D structure in [Figure 5.3b](#). Our observations support the thesis that autophosphorylation is prevented in the closed conformation by a binding of the linker to the FERM domain [42].

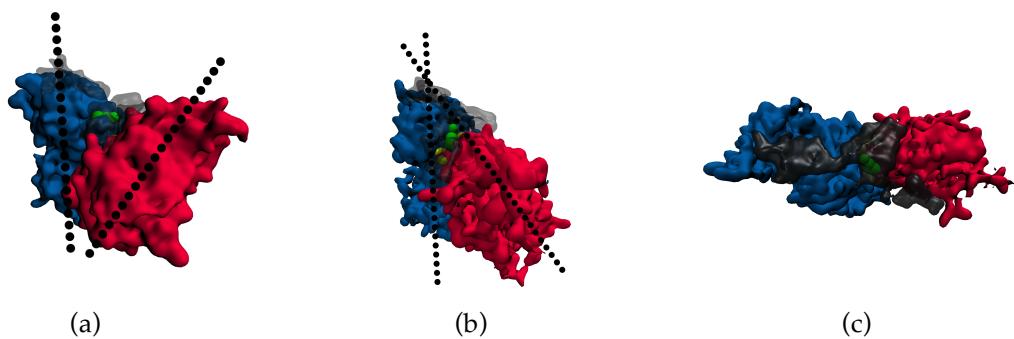


Figure 5.3: 3D structure of the two states in FAK-SOL

The 3D structures show the FERM domain (blue), the kinase (red) and the linker (black, transparent). The autophosphorylation site Y^{397} (green) is plunged into the interface and hidden by the linker. Y^{576} and Y^{577} (yellow) are shielded by the FERM domain. (a, spot 1): The kinase is tilted against FERM. (b, spot 2): Both domains are in line. (c, spot 2): Y^{397} , Y^{576} and Y^{577} are shielded by the linker.

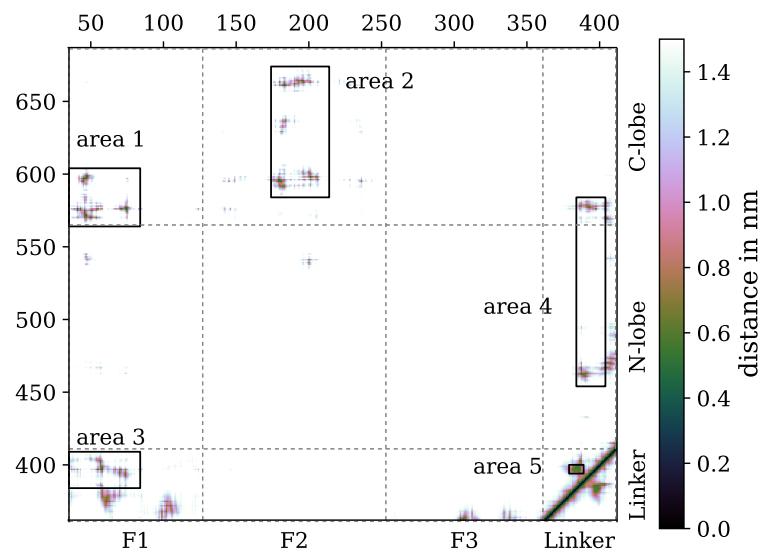


Figure 5.4: **Contact map of FAK-SOL**

Average contact map of the FERM-kinase interface and the linker region for frames of spot 2. The contact map for frames of spot 1 is similar, but lacks the contacts in area 2. In the contact map the burying of the active site (area 1) as well as the hiding of the autophosphorylation site Y³⁹⁷ (area 3, 4 and 5) can be seen.

5.3 Impact of the stabilising force

In setup 3 and setup 4, we stabilised each protein with an external force acting on the FERM domain (see [chapter 2](#)). To ensure that this restraining does not cause major artefacts, the force-dependency of the used observables is examined below.

The force has a mean value of 2.44 pN and a standard deviation of 21.80 pN. Interestingly, the positive mean value shows that the proteins are rather pushed downwards than pulled up. This indicates that the stabilising force does not have to hold the protein upright the whole time.

Linear regressions show that all of the quantities d_F (distance between F1 and F2), d_{F1-N} , d_{F2-C} and CA have a negligible correlation to the applied force. Here, negligible means that either the regression result was not significant or that the obtained slope was so small, that a change of one standard deviation in force would not change the quantity noticeably.

We also tested the inter-residue distances for correlation with the applied force. To this end, 10 different proteins without neighbours were monitored, each for 1 μ s. For each residue pair in this dataset, we performed a linear regression. [Figure 5.5](#) shows the calculated Pearson correlation coefficient (only significant correlations with Pearson $|r| > 0.3$). The mean value of the slope for the positively correlated pair distances is 33.7 pN/nm and -32.7 pN/nm for the negatively correlated pair distances. Thus, the force can influence residue pairs contributing to the interface. However, we do not expect major changes in the contact regions, since the majority of residue pairs show only weak correlations.

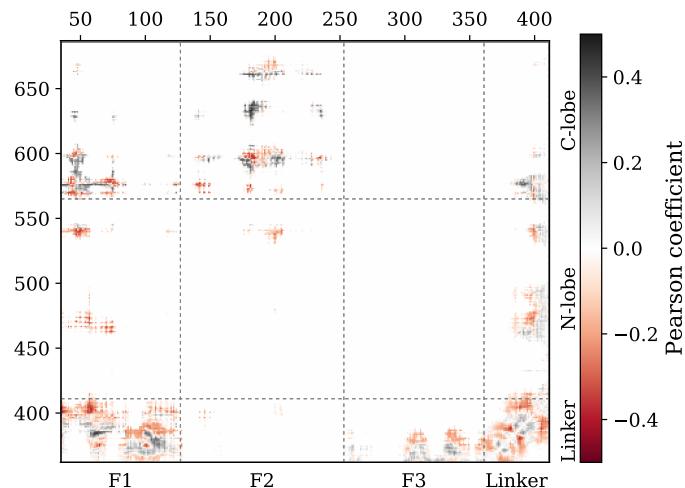


Figure 5.5: Correlations in contact map

The contact map shows correlations between inter-residue distances and the stabilising force. The mean slope is 33.2 pN/nm, the maximal Pearson $r < 0.5$.

In summary, we do not expect large perturbations of our observables due to the force. However, this approach has limitations. First, there could be binding poses in multiple FAK interactions requiring a large inclination of the FAK molecule. These states would be suppressed by the force. Another limitation is that the force does not only prevent tilts around the long axis of FAK (falling to sideways), but also around the short axis, which happens e.g. in FERM-FERM dimers. Lastly the reference to the z-axis is problematic. A reference to the membrane might be better, because it would involve membrane curvature as well.

5.4 Conformational changes on a membrane

In this section, we compare the conformation of FAK bound to a PI(4,5)P₂-containing membrane (FAK-MEM) to the observations of FAK-SOL in order to get an insight into conformational changes due to PI(4,5)P₂. To this end, FAK structures of setup 4 were used with the condition that the FAK molecules have no neighbours.

The distribution of d_{F2-C} shown in [Figure 5.6a](#) lacks the peak observed for FAK-SOL at 3.8 nm which has been associated with spot 1 (see [Figure 5.2](#)). Moreover, the distribution of d_{F1-N} shows a distinct peak at values associated to spot 2 observed in FAK-SOL (4.0 nm). This implies that the tilting of the kinase domain is suppressed in presence of the membrane, which is reasonable since both, FERM domain and kinase, have a docking site for PI(4,5)P₂. However, there is a significant increase in the contact area CA of ca. 10% to 30.8 nm² (see [Figure 5.6b](#)), which suggest a closure of the two domains.

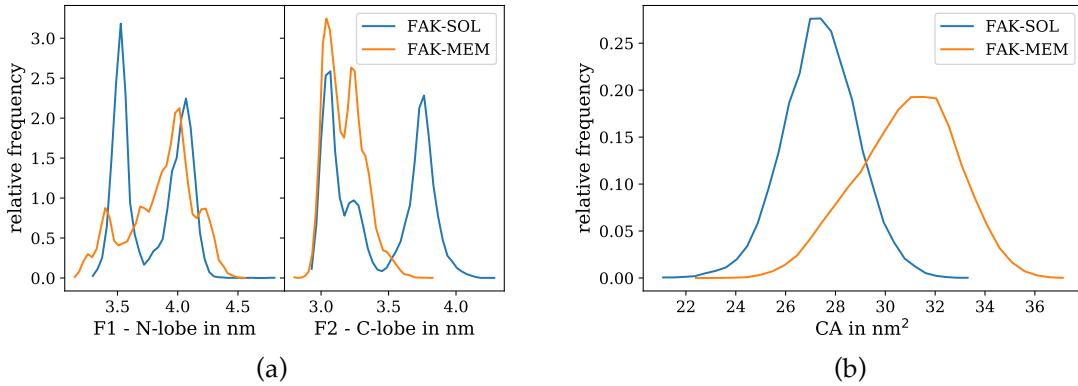


Figure 5.6: Domain distances and contact area of FAK-MEM

(a): distribution of d_{F1-N} and d_{F2-C} in comparison to FAK-SOL. The preferred state corresponds to spot 2 observed in FAK-SOL. (b): CA in comparison to FAK-SOL. The mean value is increased by ca. 10%.

The difference contact map in [Figure 5.7](#) shows the difference of inter-residue distances in comparison to FAK-SOL. At the interface of F2 and the C-lobe (area 1), the residue pairs get closer. In area 2, both increases and decreases in distances can be found. Regarding the 3D structure ([Figure 5.8](#)), one can see that although the upper part of the interface (contacts between P⁴⁹ to S⁵⁴ and the kinase, yellow) gets closer, the region below (contacts between E⁴⁴ to E⁴⁸ and the kinase, magenta) gets farther away. This induces a spreading of the interface at the activation loop including the activity-regulating residues Y⁵⁷⁶ and Y⁵⁷⁷.

Also in the linker region, conformational changes can be observed. Area 3 indicates that the Y³⁹⁷-containing ball structure observed in FAK-SOL disappears. As shown in [Figure 5.8](#), the linker is pulled back to the FERM domain, which results in an exposed posi-

tion of the autophosphorylation site Y³⁹⁷.

As suggested by findings of Goñi et al. [17] and Zhou et al. [42], the FERM domain does not dissociate from the kinase because of binding to PI(4,5)P₂. However, this binding leads to conformational changes resulting in a partial opening of the FERM-kinase interface at the activation loop and the exposure of the autophosphorylation site Y³⁹⁷.

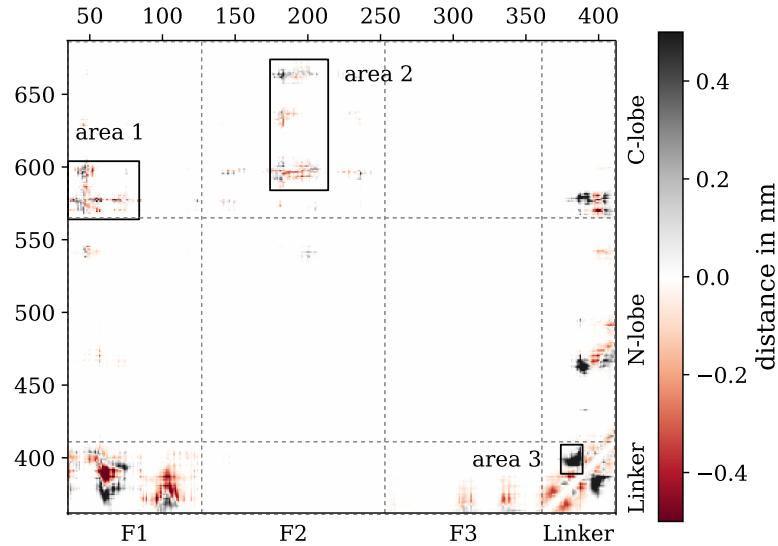


Figure 5.7: Contact map of FAK-MEM

The contact map shows the difference FAK-MEM - FAK-SOL of the inter-residue distances for the FERM-kinase interface. Conformational changes (area 1 and 2) as well as the exposure of Y³⁹⁷ (area 3) can be seen. The underlying dataset for FAK-MEM is the same as for the correlation contact map in [Figure 5.5](#).

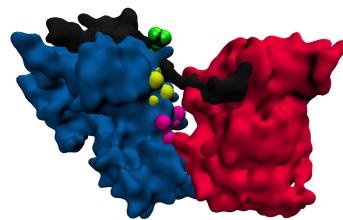


Figure 5.8: 3D structure of FAK-MEM

The 3D structure shows the FERM domain (blue), the kinase (red) and the linker (black, transparent). The membrane is not shown. The autophosphorylation site Y³⁹⁷ (green) is exposed. Residues around S⁴⁷ (magenta) get farther away from the kinase, while residues around W⁵² (yellow) get closer.

5.5 Multiple FAK interactions

The focus in this section is on interactions obtained from setup 4 occurring between multiple FAK molecules. At this point, we want to remind the reader that the used protein structure lacks the FAT domain, which is in full length FAK connected to the kinase via a linker region. This might have a significant effect on clustering processes.

5.5.1 Dimerization of FAK

First, we want to investigate interactions between two FAK molecules. To this end, we filtered the dataset for interactions in which both partners have exactly one neighbour (dimers). It turns out that FERM-kinase dimers (type 3 neighbours) occur the most, followed by FERM-FERM dimers (type 1 neighbours). Probably, this does not reflect an energetic preference but is rather attributed to the initial biased preference of type 1 interactions (see [section 3.3](#)). Regarding our observations, the number of encounters of FERM-FERM dimers plus the number of encounters of kinase-kinase dimers is, indeed, very similar to the number of encounters of FERM-kinase dimers. Actually, it is difficult to retrieve the energetic preference of the dimerization-type in MARTINI. Since two proteins can only interact within the cut-off radius of long range interactions, the proteins forming a dimer are already very close before starting to interact and might not adjust their orientations.

Because FERM-FERM dimers have been experimentally investigated by Brami-Cherrier et al. [7], a short comparison between the obtained results and the experimental findings follows. For this purpose, we chose the FERM-FERM dimer with the longest lasting time as a representative example.

The contact map of the interaction surface is shown in [Figure 5.9b](#). From this, the importance of residue W²⁶⁶ can be confirmed (orange region). Beside this, also the residues Y²⁸² to K²⁸⁶ (green region) and T²⁹¹ to A²⁹⁴ (red region) appear as important actors at the interaction surface since they interact with W²⁶⁶ and among each other. However, the RMSF value of these distances is larger than for the pairs involving W²⁶⁶.

At this point, it has to be remarked that other FERM-FERM dimers in the simulation also show contacts between F1 and F3. Yet, since these contacts are not visible in other samples (e.g. the chosen one), they are not presented here.

Comparing the FERM-kinase interfaces of the involved proteins to FAK-MEM, we noticed a decrease of the inter residue distances by ca. 0.2 nm. This can be attributed to the fact that FERM-FERM dimerisation requires an uplift of F3, because the interaction interface points towards the membrane in FAK-MEM. Since the FERM domain is anchored on the membrane via the basic patch, this would push the kinase partly into the membrane (see

Figure 5.9a). Therefore, a force is acting on the kinase, which is a possible reason for the observed closure.

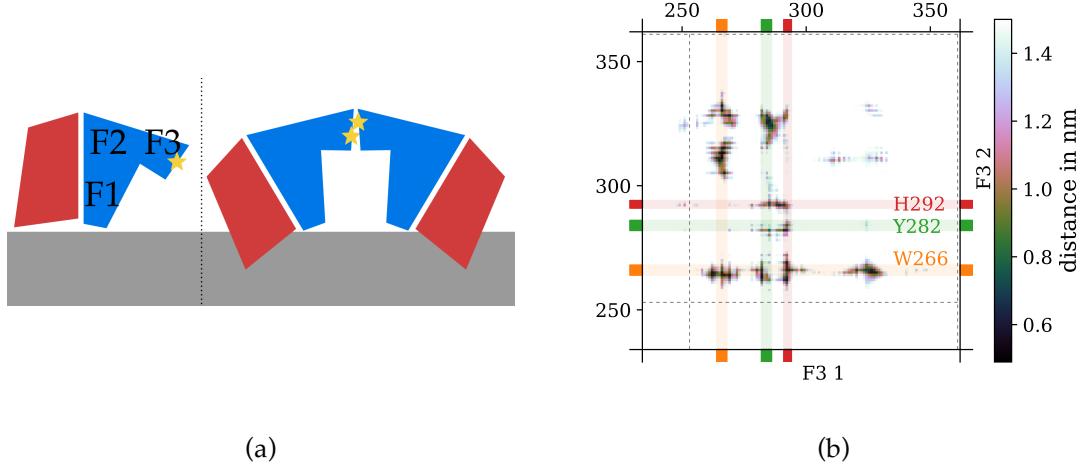


Figure 5.9: **FERM-FERM dimers**

(a): Sketch of FAK-MEM (left) the interaction site points to the membrane. Therefore, an uplift of F3 is required for dimerization (right), which pushes the kinase into the membrane. W²⁶⁶ is marked with a star. (b): Contact map of the interaction site between the two FERM domains (only F3 is shown). W²⁶⁶ (orange) forms a lot of contacts with the neighbouring FAK molecule. In addition, the regions around Y²⁸² (green) and H²⁹² (red) are important contributors.

5.5.2 FAK clusters

The characterisation of the emerged FAK clusters is very difficult as they differ a lot in size and shape. The largest cluster observed in setup 4 has a size of 21 proteins, while there are other proteins, which did not join any cluster at all. Present shapes of the clusters include long chains as well as ring-like conformations or just agglomerations (see **Figure 5.10**).

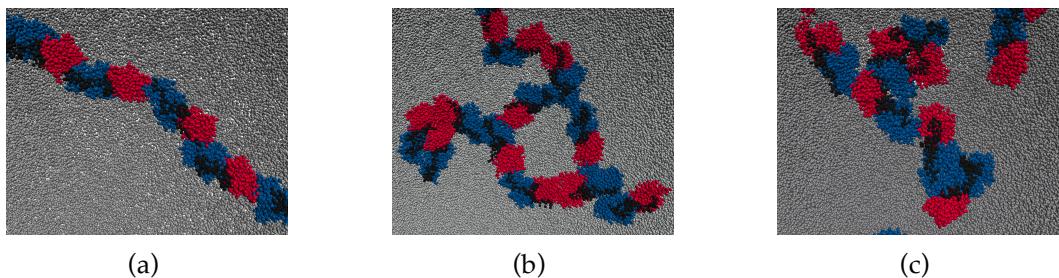


Figure 5.10: **Obtained cluster shapes**

FAK molecules (FERM blue, kinase red, linker black) in different clusters on a membrane (grey). The shapes differ a lot and are hard to classify. Examples are chains (a), ring-like structures (b) and agglomerations (c)

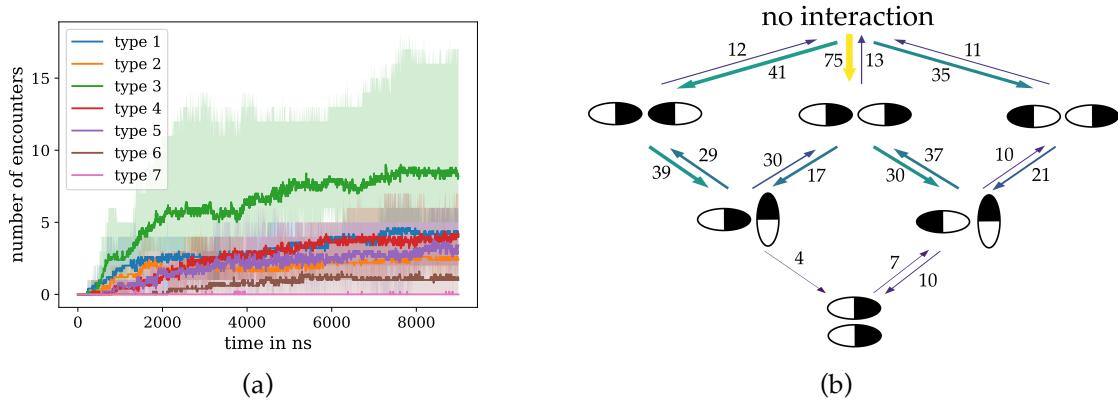


Figure 5.11: Evolution of neighbouring types

(a): Mean number of encounters over the five copies for the different neighbouring types. The filled area shows the range from minimum to maximum. (b): Transition diagram of the neighbouring types. Width and colour of the arrows correspond to the total number of observed transitions, which is also given beside the arrows. For simplicity, only transitions which occurred more than 5% of the maximum (at least 4 times) are shown.

In order to classify the occurring interactions, we analysed the time evolution of the neighbouring types as well as the transitions among them (see Figure 5.11). These plots indicate that interactions involving two domains in total (type 1, 2 and 3) occur much more often than interactions involving four domains (type 6 and 7). Especially type 7 neighbours seem to be unstable, since they always lose contact immediately after the formation (see Figure 5.11a, pink line). Therefore, a tendency for interactions along the long axis of FAK can be deduced.

Furthermore, the mean number of neighbours is considered. This quantity rises fast in the beginning and flattens after $6\ \mu s$. The average over the five copies is 1.86 neighbours at the end of the simulations.

From these observations, the conclusion can be drawn that the preferred arrangement of FAK molecules are chains along their long axis. The FAK molecules inside the chain can interact either as type 3 neighbours (FK chains) or as type 1 and type 2 neighbours (FFKK chains). Combinations of these two are also possible.

5.5.3 Activation due to clustering

Lastly, the impact of clustering on FAK activation is addressed. Since activation means the dissociation of the FERM domain from the kinase, we analysed our data with respect to the CA values, d_{F1-N} and d_{F2-C} .

Unfortunately, in no FAK molecule did a full dissociation take place at any time. In addition, the CA values seem to be independent of the number of neighbours, the interaction type and the clustersize. Because only a small number of FAK molecules was put into

the system and the clustering process was not finished at the end of the simulations, this does not mean that activation due to clustering is impossible. In the hope of seeing trends regarding to activation, a more detailed analysis for the different proteins was performed.

Motivated from subsection 5.5.2, we focus on FAK molecules inside chains. A FAK molecule can be seen as a chain member if it has exactly two neighbours and if these neighbours are not neighbours of one another. For FK chains, only type 3 interactions were allowed, for FFKK chains both, type 1 and type 2. We compare the results for FK chains and FFKK chains to FAK-MEM (see Figure 5.12 and Figure 5.13, respectively).

Figure 5.12 indicates that an arrangement of FAK molecules to FK chains does not influence the mean value of CA. The distribution gets slightly sharper compared to FAK-MEM, which could imply a stabilisation of the interface. Not even the distributions of d_{F1-N} and d_{F2-C} differ noticeably from those obtained in FAK-MEM.

Neither does an arrangement into FFKK chains affects the value of CA (Figure 5.13) significantly. However, the distribution of d_{F1-N} indicates a trend to larger values in comparison to FAK-MEM, which are linked to a more open state of FAK.

The filtering leads to a smaller sampling of the distributions of d_{F1-N} , d_{F2-C} and CA. For example, for FFKK chains ca. 40% of the underlying dataset (21.7 μ s of simulation time in total) comes from one FAK instance only. Therefore, the statistical significance of the obtained difference doubtful.

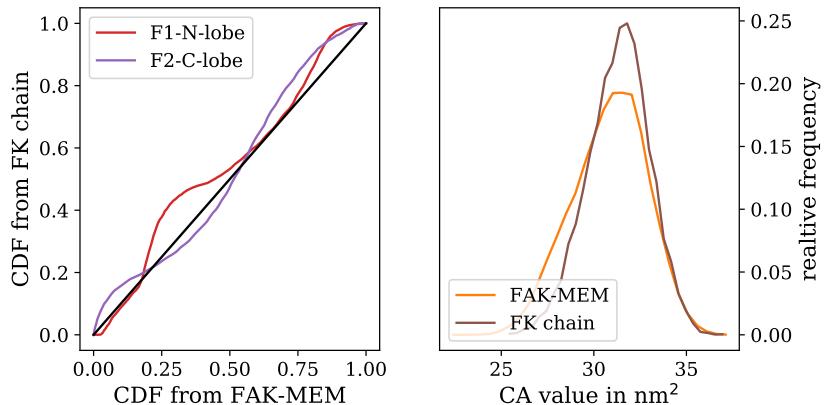


Figure 5.12: Domain distances and contact area of FK chains

(left): The Q-Q plot of d_{F1-N} and d_{F2-C} distributions in comparison to FAK-MEM imply a strong similarity. (right): The CA value doesn't change noticeably in comparison to FAK-MEM.

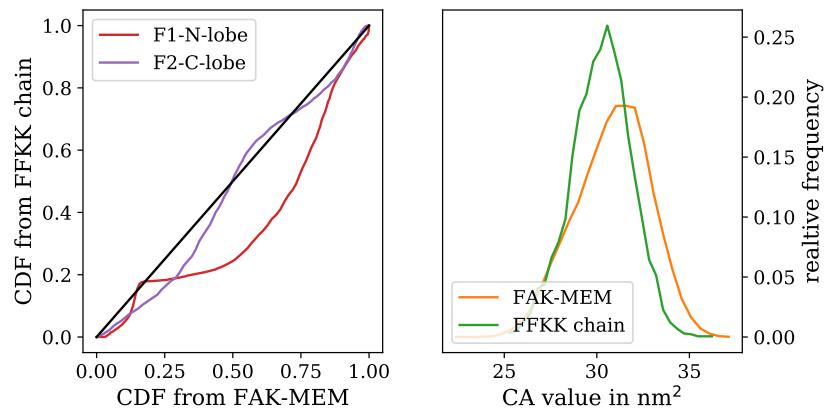


Figure 5.13: Domain distances and contact area of FFKK chains

(left): The Q-Q plot of d_{F1-N} and d_{F2-C} distributions in comparison to FAK-MEM imply a tendency to larger distances of F1 and the N-lobe for FFKK chain members. (right): The CA value decreases slightly by 2%..

6 Conclusion

The aim of the present work was to gain insight into conformational changes of FAK due to PI(4,5)P₂ binding and interactions between multiple FAK molecules. To this end, we used coarse-grained MD simulations with the MARTINI force field.

Previous work in the group revealed unnatural behaviour of FAK bound to a PI(4,5)P₂-containing membrane in MARTINI simulations, namely, a falling sideways on the membrane. We were able to prove that this is not induced by an underestimation of PI(4,5)P₂ binding at the basic patch. Indeed, MARTINI was able to reproduce the free energy profile of this binding obtained from a reference simulation in CHARMM36 remarkably well. However, the cause of the falling is still an outstanding question and should be addressed in further studies.

In this project, we introduced a workaround, namely, a stabilising force acting on the FERM domain of the FAK molecules. In section 5.3, we showed that this approach does not influence the observables used in the remaining part, but we are aware of its limitations (discussed in section 5.3) and still investigate possible alternatives, for example, flat-bottom pulling [1, p. 156-158].

From the configurations obtained for FAK in solution, we identified important residues contributing to the FERM-kinase interface. The observations fit well with experimental studies of Lietha et al. [28]. Also the burying of the active site of the kinase as well as the hiding of the autophosphorylation site were observed in the simulations.

We compared these results to conformations obtained in FAK molecules bound to a PI(4,5)P₂-containing membrane and identified configurational changes. They involve not only the promotion of the autophosphorylation site, but also a partial opening of the FERM-kinase interface at the activation loop of the kinase. These changes are consistent with previous studies on allosteric effects of PI(4,5)P₂ binding to FAK by Goñi et al. [17] and Zhou et al. [42].

In section 5.5, we investigated the interactions between multiple FAK molecules on a membrane. Our observations do not support a conclusion on the energetic preference of the different dimer types. Yet, we were able to confirm the importance of W²⁶⁶ in FERM-FERM dimerization proposed by Brami-Cherrier et al. [7]. Regarding more than two FAK

molecules, we observed a tendency to aggregate in chain-like structures. These structures were investigated for activation-promoting features, like increased domain distances or a smaller contact area, but no significant differences have been observed.

Several aspects of FAK clustering are still elusive and have to be clarified in further studies. First, the local aggregation and arrangement of FAK molecules is of special interest and could be investigated with further simulations of dimers, trimers and tetramers. With a deep understanding of these interactions, simulations of larger oligomers with a higher level of coarse graining could become accessible. Moreover, a more detailed analysis of the role of the membrane, especially the PI(4,5)P₂ concentration and membrane curvature, might contribute to the understanding of FAK clustering processes.

References

- [1] M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GROMACS development team. *GROMACS User Manual version 2016.2.* 2016.2. www.gromacs.org, 2017.
- [2] Stefan T Arold, Maria K Hoellerer, and Martin E.M Noble. “The Structural Basis of Localization and Signaling by the Focal Adhesion Targeting Domain”. In: *Structure* 10.3 (2002), pp. 319–327.
- [3] Sara Becker, Csaba Daday, and Frauke Gräter. “Clustering of FAK on a PIP2 enriched membrane”. Feb. 2017.
- [4] H.J.C Berendsen, D van der Spoel, and R van Drunen. “GROMACS: A message-passing parallel molecular dynamics implementation”. In: *Computer Physics Communications* 91.1 (1995), pp. 43–56.
- [5] HJC Berendsen, JPM POSTMA, WF VANGUNSTEREN, A DINOLA, and [No Value] HAAK. “Molecular-Dynamics with coupling to an external bath”. In: *Journal of Chemical Physics* 81.8 (1984), pp. 3684–3690.
- [6] Robert B. Best, Xiao Zhu, Jihyun Shim, Pedro E M Lopes, Jeetain Mittal, Michael Feig, and Alexander D. MacKerell. “Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ 1 and χ 2 Dihedral Angles”. In: *Journal of Chemical Theory and Computation* 8.9 (2012), pp. 3257–3273.
- [7] Karen Brami-Cherrier et al. “FAK dimerization controls its kinase-dependent functions at focal adhesions”. In: *The EMBO Journal* 33.4 (2014), pp. 356–370.
- [8] M B Calalb, T R Polte, and S K Hanks. “Tyrosine phosphorylation of focal adhesion kinase at sites in the catalytic domain regulates kinase activity: a role for Src family kinases”. In: *Molecular and Cellular Biology* 15.2 (1995), pp. 954–963.
- [9] Christophe Chipot and Andrew Pohorille. *Free Energy Calculations. Theory and Applications in Chemistry and Biology.* Springer Series in CHEMICAL PHYSICS ; 86. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
- [10] Athar H Chishti et al. “The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane”. In: *Trends in Biochemical Sciences* 23.8 (1998), pp. 281–282.

- [11] Tom Darden, Darrin York, and Lee Pedersen. "Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems". In: *The Journal of Chemical Physics* 98.12 (1993), pp. 10089–10092.
- [12] Davide Donadio, Michele Parrinello, and Giovanni Bussi. "Canonical sampling through velocity rescaling". In: *Journal of Chemical Physics* 126.1 (2007).
- [13] Frank Eisenhaber, Philip Lijnzaad, Patrick Argos, Chris Sander, and Michael Scharf. "The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies". In: *Journal of Computational Chemistry* 16.3 (1995), pp. 273–284.
- [14] Jun Feng and Blake Mertz. "Novel Phosphotidylinositol 4,5-Bisphosphate Binding Sites on Focal Adhesion Kinase". In: *PloS one* 10.7 (2015).
- [15] Alan M Ferrenberg and Robert H Swendsen. "Optimized Monte Carlo data analysis". In: *Physical Review Letters* 63.12 (1989), pp. 1195–1198.
- [16] Daan Frenkel and Berend Smit. *Understanding molecular simulation. from algorithms to applications*. 2. ed. Computational science series. San Diego [u.a.]: Academic Press, 2002.
- [17] Guillermina M Goñi, Carolina Epifano, Jasminka Boskovic, Marta Camacho-Artacho, Jing Zhou, Agnieszka Bronowska, M Teresa Martín, Michael J Eck, Leonor Kremer, Frauke Gräter, et al. "Phosphatidylinositol 4, 5-bisphosphate triggers activation of focal adhesion kinase by inducing clustering and conformational changes". In: *Proceedings of the National Academy of Sciences* 111.31 (2014), pp. 3177–3186.
- [18] Jessica E Hall and Michael D Schaller. "Phospholipid binding to the FAK catalytic domain impacts function". In: *PloS one* 12.2 (2017).
- [19] Stephen C Harvey, Robert K.-Z Tan, and Thomas E Cheatham. "The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition". In: *Journal of Computational Chemistry* 19.7 (1998), pp. 726–740.
- [20] William G Hoover. "Canonical dynamics: Equilibrium phase-space distributions". In: *Physical Review A* 31.3 (1985), pp. 1695–1697.
- [21] Jochen S Hub, Bert L de Groot, and David van der Spoel. "g_wham-A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates". In: *Journal of Chemical Theory and Computation* 6.12 (2010), p. 3713.
- [22] Djurre H de Jong, Gurpreet Singh, W. F. w Bennett, Clement Arnarez, Tsjerk A Wassenaar, Lars V Schafer, Xavier Periole, D. Peter Tieleman, Siewert J Marrink, and Lars V Schäfer. "Improved Parameters for the Martini Coarse-Grained Protein Force Field". In: *Journal of Chemical Theory and Computation* 9.1 (2013), pp. 687–697.
- [23] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein. "Comparison of simple potential functions for simulating liquid water". In: *The Journal of Chemical Physics* 79.2 (1983), p. 926.

- [24] Ben-Zion Katz, Shingo Miyamoto, Hidemi Teramoto, Muriel Zohar, Dmitry Krylov, Charles Vinson, J.Silvio Gutkind, and Kenneth M Yamada. "Direct transmembrane clustering and cytoplasmic dimerization of focal adhesion kinase initiates its tyrosine phosphorylation". In: *BBA - Molecular Cell Research* 1592.2 (2002), pp. 141–152.
- [25] Jeffery B Klauda, Richard M Venable, J Alfredo Freites, Joseph W O'Connor, Douglas J Tobias, Carlos Mondragon-Ramirez, Igor Vorobyov, Alexander D MacKerell Jr, and Richard W Pastor. "Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types". In: *The journal of physical chemistry. B* 114.23 (2010), p. 7830.
- [26] Shankar Kumar, John M Rosenberg, Djamal Bouzida, Robert H Swendsen, and Peter A Kollman. "The weighted histogram analysis method for free-energy calculations on biomolecules." In: *Journal of Computational Chemistry* 13.8 (1992), pp. 1011–1021.
- [27] Kyle R Legate, Seiichiro Takahashi, Navid Bonakdar, Ben Fabry, David Boettiger, Roy Zent, and Reinhart Fässler. "Integrin adhesion and force coupling are independently regulated by localized PtdIns(4,5)2synthesis. PIPK γ accelerates adhesion formation". In: *The EMBO Journal* 30.22 (2011), pp. 4539–4553.
- [28] Daniel Lietha, Xinming Cai, Derek F.J Ceccarelli, Yiqun Li, Michael D Schaller, and Michael J Eck. "Structural Basis for the Autoinhibition of Focal Adhesion Kinase". In: *Cell* 129.6 (2007), pp. 1177–1187.
- [29] Cesar A Lopez, Zofie Sovova, Floris J van Eerden, Alex H de Vries, and Siewert J Marrink. "Martini Force Field Parameters for Glycolipids". In: *Journal of Chemical Theory and Computation* 9.3 (2013), pp. 1694–1708.
- [30] Siewert J Marrink and D. Peter Tieleman. "Perspective on the Martini model". In: *Chemical Society Reviews* 42.16 (2013), pp. 6801–6822.
- [31] Siewert J Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H de Vries. "The MARTINI force field: Coarse grained model for biomolecular simulations". In: *Journal of Physical Chemistry B* 111.27 (2007), pp. 7812–7824.
- [32] Stuart McLaughlin, Jiyao Wang, Alok Gambhir, and Diana Murray. "PIP2 and Proteins: Interactions, Organization, and Information Flow". In: *Annual Review of Biophysics and Biomolecular Structure* 31.1 (2002), pp. 151–175.
- [33] Luca Monticelli and Emppu Salonen, eds. *Biomolecular Simulations - Methods and Protocols*. Vol. 924. Methods in Molecular Biology. Hamana Press, 2013.
- [34] Shūichi Nosé. "A molecular dynamics method for simulations in the canonical ensemble". In: *Molecular Physics* 52.2 (1984), pp. 255–268.
- [35] Shuichi Nosé and M.L Klein. "Constant pressure molecular dynamics for molecular systems". In: *Molecular Physics* 50.5 (1983), pp. 1055–1076.

- [36] M Parrinello and A Rahman. "Polymorphic transitions in single crystals: A new molecular dynamics method". In: *Journal of Applied Physics* 52.12 (1981), pp. 7182–7190.
- [37] G.M Torrie and J.P Valleau. "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling". In: *Journal of Computational Physics* 23.2 (1977), pp. 187–199.
- [38] Madeleine Toutant, Alicia Costa, Jeanne-Marie Studler, Gress Kadaré, Michèle Carnaud, and Jean-Antoine Girault. "Alternative Splicing Controls the Mechanisms of FAK Autophosphorylation". In: *Molecular and Cellular Biology* 22.22 (2002), pp. 7731–7743.
- [39] Tsjerk A Wassenaar, Kristyna Pluhackova, Rainer A Böckmann, Siewert J Marrink, and D Peter Tieleman. "Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models". In: *Journal of Chemical Theory and Computation* 10.2 (2014), p. 676.
- [40] Semen O Yesylevskyy, Lars V Schafer, Durba Sengupta, Siewert J Marrink, and Lars V Schäfer. "Polarizable Water Model for the Coarse-Grained MARTINI Force Field". In: *PLOS Computational Biology* 6.6 (2010).
- [41] Jihe Zhao and Jun-Lin Guan. "Signal transduction by focal adhesion kinase in cancer". In: *Cancer and Metastasis Reviews* 28.1 (2009), pp. 35–49.
- [42] Jing Zhou, Agnieszka Bronowska, Johanne Le Coq, Daniel Lietha, and Frauke Gräter. "Allosteric Regulation of Focal Adhesion Kinase by PIP2 and ATP". In: *Biophysical Journal* 108.3 (2015), pp. 698–705.
- [43] Jing Zhou, Camilo Aponte-Santamaría, Sebastian Sturm, Jakob Tómas Bullerjahn, Agnieszka Bronowska, and Frauke Gräter. "Mechanism of Focal Adhesion Kinase Mechanosensing". In: *PLoS computational biology* 11.11 (2015).