1 Introduction

2 Motivation

As described in **??** the process of FAK clustering and its effect upon activation, especially on an atomistic scale, are still not understood and part of current research. In this thesis the results of MD simulations with the Martini force field (see **??**) are presented, which address the clustering process of FAK molecules. In this context Martini is a necessary simplification due to the large number of particles in systems containing sever FAK molecules.

However previous work in the group [**sara**] obtained difficulties in the use of Martini for simulations of FAK on a PI(4,5)P_2 containing membrane. In some simulations equivalent to setup 3 (see section 4.4) except for an external force the protein rapidly changed its inclination to the membrane. In the following part this shall be characterised by the angle β between the z-axis and the vector connecting F1 and F2, $\vec{d}_F$.

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

sara simulated five different copies for $10\,\mu$s each. The resulting distributions of the angle can be seen in **??**. Exemplary the red line shows a mean value of 90 deg, which is what meant with a falling of FAK in the following part. The angle changed in less than XXX and stayed constant for the remaining simulation time.
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 [1]. The first one is, that FAK falls to both extensive sites, which means that the key residues for an interaction of the extensive site of the kinase with the membrane proposed by Feng and Mertz [1] are located on top of the FAK and not on the side of the membrane. Indeed contact analysis showed, that virtually all residues on the surface (in both, FERM domain and kinase) were interacting with the membrane. Another one is, that this behaviour was not observed in equivalent all atom simulations in C36 ($1.5\,\mu$s in total). Here only two maxima were observed around 8 deg and around 20 deg, the largest observed angle was 40 deg.

In the course of this project several efforts were made to understand the cause of this falling, e. g. to review the binding of the basic patch in Martini, which is presented in **??**. However the reason could not have proven beyond doubt. In order to still perform reasonable simulations of multiple FAK an external force was applied to each FAK molecule. This is called stabilizing force in the following parts.
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. An illustration of the force can be found in **??**. For the determination of z_0 only the green and the blue distribution from **??** were considered, because the large angles observed in the other distributions have not been observed in C36 simulations. The mean value of $\vec{d}_{F,z}$ for these two distributions is $2.228\,\text{nm}$, which was therefore set as z_0.
@Csaba: sorry, but how did we get this force constant of 1000 out of the distributions?

In the following two chapters the used methods, i.e. MD simulation, are explained and the used setups introduced. Afterwards the obtained results are presented. For this purpose FAK in solution and FAK bound to PI(4,5)P_2 are analysed and compared to known information from experiments or other simulations. Also the impacts of the stabilizing force onto the simulations are commented. At last the focus is set on interactions between multiple FAK molecules.

3 Methods

3.1 MD

3.2 Free energy calculations

3.3 ConAn :)

4 Setup

4.1 Protein structure

All simulations have been done with starting configurations adapted from previous work in the group (C36 forcefield: Zhou et al. [4], Martini forcefield: **sara**). These configurations contain only a FERM-kinase fragment without the FAT domain and its linker (residues 35 to 686, PDB 2J0J [2]).
As explained in **??** the secondary structure of proteins have to be stabilized in Martini using elastic networks. This was set up by **sara**. It acts only between residues of the same domain with a force constant of 830 $kJmol^{-1}nm^{-2}$. Therefore the interface between FERM domain and kinase is not affected and the linker is still flexible.

4.2 Setup 1 - FAK in solution

This setup contains a single FAK molecule in Martini structure, which was placed in a waterbox with ions to eliminate net charges. After a short equilibration the system was simulated for $20\,\mu s$ at a temperature of $300K$. The standard parameters of the Martini forcefield were used as simulation 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), which contains the basic patch, was used. The lobe was placed above a single $PI(4,5)P_2$ embedded into a phosphatidylethanolamine (POPE) membrane (see **??**). After a short equilibration the F2 lobe was pulled slowly away from the membrane using a distance pull between the COM of the F2 lobe and the COM

of PI(4,5)P_2. This simulation was used to retrieve starting conformations for the umbrella windows.
The starting configuration was set up in C36 and transferred to a Martini (with PW) structure with provided transformation tools [**PW**, 3]. In the Martini structures the mentioned elastic network was applied.
The number of umbrella windows was chosen accordingly to the sampling (between 90 and 120 windows). Each window was shortly equilibrated and then simulated for 6 ns in C36 and 10 ns in Martini (with PW). From the trajectories of the umbrella windows the free energy profile was calculated using GROMACS WHAM implementation [**gromacsWHAMpaper**]. For each forcefield five different profiles were obtained to estimate the statistical error.
The presented results are based on a total simulation time of $6.33\,\mu$s for Martini, $5.64\,\mu$s for Martini with PW and $3.88\,\mu$s for C36. The temperature in the simulation was 300K.

4.4 Setup 3 - FAK on a PI(4,5)P_2 membrane

Setup 3 was adopted from **sara** and contains a single FAK molecule in Martini structure, which was placed on a phosphatidylcholine (POPC) and PI(4,5)P_2 (15%) membrane. Ions were added to eliminate net charges. In contrast to **sara** the stabilizing force explained in chapter 2 was applied to the protein.
Five different copies were simulated for $10\,\mu$s each with a temperature of 300K.

4.5 Setup 4 - FAK cluster

In order to set up a cluster of multiple FAK molecules 25 different frames were chosen from setup 3 equidistant in time. The frames were arranged on a 5x5 grid. The stabilizing force was applied on each FAK molecule independently. After a short equilibration the system was simulated for $9\,\mu$s. Five different copies (regarding to the arrangement of the frames) were set up, resulting in a total simulation time of $45\,\mu$s. The temperature was set to 300K.

5 Results

5.1 FAK in solution

In this section the conformation of FAK in absence of PI(4,5)P_2 and other FAK molecules is analyzed. For this purpose the simulation data of setup 1 were used (see section 4.2).

First the COM distances of F1 to the N-lobe ($d_{\text{F1-N}}$) and F2 to the C-lobe ($d_{\text{F2-C}}$) are considered. In **??** a hexagonal binning plot of both values can be found, which indicates, that there are two different states: one in which $d_{\text{F2-C}}$ is larger (spot 1) and $d_{\text{F1-N}}$ smaller and one the other way around (spot 2). The systems starts in spot 1 and goes to spot 2 after $7\,\mu$s (this transition goes along with several frequent transitions), where it stays until the end of the simulation.
jing performed equivalent simulations in C36 and obtained

$$d_{\text{F1-N}} = (3.30 \pm 0.40\,\text{std})\text{nm} \tag{5.1}$$

$$d_{\text{F2-C}} = (3.15 \pm 0.15\,\text{std})\text{nm} \tag{5.2}$$

These values can not be classified in one of the two spots as both distances are lower than those obtained in Martini.

A contact map of the interface between the FERM domain and the kinase for frames of spot 2 can be found in **??**. Two contact areas can be identified. The first one (area 1) is located between F1 and the N-lobe/activation loop. It shows i.e. contacts between Y^{576} and Y^{577} and residues of the FERM domain. The minimal distance in this area, between residue H^{41} and Y^{576}, is 0.45 nm with an RMSF value of 0.03 nm. This reflects the burying of the activity regulating residues in closed state.
The second contact area (area 2) 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}. 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. These two residues have been reported as important actors in the interface by showing, that a mutation disturbs the interface and enhances the activation of FAK.
The linker show contacts with both domains. Interestingly the minimal distance in the marked areas (area 3 and area 4) occur between the autophosphorylation site Y^{397} and H^{58} (0.45 nm, RMSF 0.03 nm) in F1 or Y^{576} (0.50 nm, RMSF 0.10 nm) in the kinase. This is consistent with the concept, that autophosphorylation is prevented in closed conformation by a binding of the linker to the FERM domain.

In contrast to **??**, the contact map for frames of spot 1 show less contacts between F2 and the C-lobe, i.e. around the mentioned residues Y^{180} to M^{183}. A few additional contacts appear between F1 and the N-lobe, but these are only minor spots.

5.2 FAK binding to PI(4,5)P_2

5.2.1 Free energy profile of basic patch

In the context of the investigation of FAK falling the free energy profile of the basic patch in F2 was investigated. Because this might be of interest for other applications as well, a short report on the obtained results is given at this point. The used setup is setup 2 described in section 4.3.
The average profile of the five copies together with the standard deviation for the forcefields C36, Martini and Martini with PW can be found in **??**. Here the range between $z = 6$ nm and $z = 7$ nm was defined as zero point.
C36 and Martini show both a similar energy depth of $A \approx 17$kJ/mol and the same slope between $z = 3$ nm and $z = 4$ nm. Certainly Martini results in larger values for the free energy, which could originate from an underestimation of electrostatic forces due to the unpolar water beads [**introOrPaper**]. For values above $z = 4$ nm the free energy profile flattens out in C36 while a kink can be observed in Martini. This is due to the different treatment of long range electrostatics: Martini uses a cut off radius, C36 uses PME.
Also Martini with PW uses PME for long range electrostatic interactions and indeed it fits much better to the C36 profile for larger distances. However the bind-

ing strength is crucially underestimated in Martini with PW.
It is remarkable, in which extent the Martini force field redraws the results from all-atom simulations. A contribution to this might be, that the parameters for Martini were obtained a.o. from free energy calculations and measurements **??**.

In the used starting configurations the proteins are already bound to PI(4,5)P_2. Therefore a correct binding strength and the shape near the minimum is of larger interest than a correct sampling for further distances. In addition Martini with PW showed a much higher computational cost. That is why in the following simulations only Martini has been considered.

5.2.2 Impacts of the applied force

In setup 3 and setup 4 each protein was stabilized with an external force acting on the FERM domain (see chapter 2). These forces have of course impacts on the FAK molecules, which are analysed below.

Distance between F1 and F2

First the COM distance of F1 and F2, namely d_F, is considered. In **??** a hexbinning plot of d_F against the force can be found. The plot shows, that some combinations (below the orange line) are never visited. For a given force, d_F seems to have a lower boundary. This boundary is defined by $d_{F,\text{min}} = z_0 - \Delta z$, where z_0 is the reference distance of the force and Δz the elongation (proportional to the force), and is reached, if the vector between F1 and F2, $\vec{d}_F$, is parallel to the z-axis.
The plot indicates, that d_F and the force are hardly correlated, which is tested with a linear regression. To not bias the regression by $d_{F,\text{min}}$ only points with a force lower than $0\,\text{kJ/mol/nm}$ are taken into account. With a p-Value of 23% for the H0 hypothesis slope $m = 0$, the distance d_F and the applied force can indeed be seen as uncorrelated.

Inclination angle of FAK

The applied force is directly linked to the angle between the distance vector of F1 and F2 and the z-axis, namely β. In order to examine how much the forces bias β, the distribution of β for setup 3 (with external force) are compared to equivalent

simulations except the stabilizing force (simulations mentioned in chapter 2).
The distribution of is very similar to those obtained for the unbiased systems, albeit the mean value is determined by the choice of z_0. The variance for the biased system is about $18\,\text{deg}$ and similar to those of the unbiased systems. This indicates, that the fluctuations around the mean value are not restricted crucially by the applied force.

The force itself has a mean value of $1.47\,\text{kJ/mol/nm}$ and a standard deviation of $13.13\,\text{kJ/mol/nm}$. It is skewed to positive values, large deviations with a negative sign are restricted due to $z_0 + \Delta z \leq d_F$. Neither additional maxima nor minima are observable in the distribution.

Intramolecular distances

At last the influence of the force on conformational changes inside FAK shall be investigated.
First all distances between interacting residues are tested on correlation with the applied force. For this 10 different proteins without neighbours were picked out of the trajectories from setup 4, each for $1\,\mu\text{s}$. In **??** the calculated Pearson correlation coefficients for significant correlations are shown (only for residues contributing to the interface). Out of these possible residue pairs 270 show a correlation larger than 0.3 or less than -0.3. The slope of these pairs is of magnitude ± 20force unit/nm. So a change of the force of one standard deviation would lead to a change of roughly $0.6\,\text{nm}$, which declines the sensitivity to conformational changes a lot.
The sign of the correlation seems to be systematic. Interactions corresponding to the N-lobe (lower part of **??**) is dominated by negative values, while the interactions corresponding to the C-lobe show both. However also in this region, i.e. the spot around residues Y^{180} and R^{634}, spots of the same correlation sign can be found.
Because the COM distance between F1 and the N-lobe as well as the COM distance between F2 and the C-lobe are used as indicators for conformational changes as well, their dependence upon the force are also important. As already shown in section 5.1 these distances cluster in two spots. To test correlations a linear regression of both distances was performed separately for the two spots. All four regressions revealed a significant correlation. However the maximum slope was

0.0015nm/force unit. A change of 13 kJ/mol/nm, which corresponds to one standard deviation in force, would therefore lead to a change of 0.02 nm. This correlation is therefore negligibly.

It is important to see, that in this section only the observed quantities were tested for correlation with the force. It is however possible and probable, that due to the restriction configurations or states are completely missing while others become overpopulated. This effect can not be estimated, but should be kept in mind.

5.2.3 Conformational changes

In this section the conformation of FAK bound to PI(4,5)P_2 (FAK-PIP) is compared to the observations from section 5.1 (FAK-SOL). For this purpose the simulation data of setup 4 was used with the condition, that other FAK molecules are more than 2 nm away (0 neighbours). The contact map is based on the same dataset, which was used in section 5.2.2.

Analogously to section 5.1 the distribution of the COM distances is presented in **??** as an hexagonal binning plot. Again, different spots can be obtained. The spot with most encounters (spot 1) as well as the second spot (spot 2) are located at small values for both, $d_{\text{F1-N}}$ and $d_{\text{F2-C}}$. These spots show also smaller distances than FAK-SOL. In addition to this two spots appear, one at larger $d_{\text{F2-C}}$ (spot 3) and one at larger $d_{\text{F1-N}}$ (spot 4). These are however less populated and not as concentrated as spot 1 and 2. While spot 4 could be identified with spot 2 of FAK-SOL, spot 3 show completely new values for both COM distances.

In **??** the difference of the contact maps of FAK-SOL and FAK-PIP can be found, again only for the interface. The distances between F2 and the C-lobe tend to get smaller, even if contacts around R^{665} show another trend. Also the contact between F1 and the N-lobe/activation loop show smaller distances. The RMSF values are increasing for all residue pairs.
Remarkable changes occur in the linker region. The residues around the autophosphorylation site Y^{397} increase their distances to residues M^{384} to S^{390} by up to 0.9 nm. Also the RMSF values are increased in the linker by up to 0.40 nm near Y^{397}.

With these observations the enhanced autophosphorylation of FAK bound to PI(4,5)P_2 can be comprehended. However the impacts on the interface between the FERM domain and the kinase are quit small. One reason for this could be the applied elastic network. Since electrostatics are treated with a cutoff radius, long range conformational changes have to be transferred along the residues. Therefore the choice of the correct force constant is of large importance to obtain allosteric effects in Martini.

5.3 Multiple FAK interactions

In this section the interactions occuring between multiple FAK molecules are analysed, for which the data from setup 4 is used. At this point the reader shall be reminded, 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 make an important difference for clustering processes.
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 had 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 an agglomeration (see **??**). Nevertheless in none FAK molecule the FERM domain dissociated from the kinase, which means no activation took place.

For further analysis interactions between two proteins are classified into 9 different interaction types (see **??**). Two proteins are interacting, if their minimal distance is smaller than 1.5 nm and are called neighbours in this case. In contrast a protein belongs to a cluster, if it has at least one neighbour inside the cluster. A single protein without neighbours has clustersize 1.
First of all the mean number of neighbours is examined. One can see in **??** a fast rising in the number of neighbours in the beginning and a flattening after 6 ns. The average of the five copies is at the end of the simulation 1.86. This indicates a tendency to chains of FAK.
In **??** the average number of encounters of the different interaction types is plotted against the time. It shows, that FERM-kinase interactions (type 3) occur the most, while the others occur equally often. Only type 6 and type 7 (interactions, in which all four domains are involved) occur much less than the others.
From these observations one could draw the conclusion, that the preferred ar-

rangement of FAK molecules is a chain, in which the FERM domain interacts with the kinase of the next molecule.