Biophysical Journal, Volume *122*

Supplemental information

FERM domains recruit ample PI(4,5)P_2s to form extensive protein-membrane attachments

Thomas Ehret, Tim Heißenberg, Svenja de Buhr, Camilo Aponte-Santamaría, Claudia Steinem, and Frauke Gräter

Supporting Information - FERM domains recruit ample PI(4,5)P_2s to form extensive protein-membrane attachments

Thomas Ehret,[†] Tim Heißenberg,[‡] Svenja de Buhr,[†] Camilo Aponte-Santamaría,[†] Claudia Steinem,[*,‡,¶] and Frauke Gräter[*,§,†,¶]

†*Heidelberg Institute for Theoretical Studies (HITS), Heidelberg*

‡*Institute of Organic and Biomolecular Chemistry, University of Göttingen, Göttingen*

¶*Max Planck School Matter to Life*

§*Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Heidelberg*

E-mail: csteine@gwdg.de; frauke.graeter@h-its.org

SI Methods

Preparation of the ezrin-FERM-membrane starting complex.

To assemble the ezrin-FERM-membrane complex, serving as the starting point of all ezrin-FERM simulations, we exploited the high sequential and structural similarity between the ezrin-FERM structure (PDB accession code 4RMA;[1] solved through x-ray diffraction at a resolution of 1.7 Å) and the radixin-FERM-inositol-(1,4,5)-triphosphate complex (PDB accession code 1GC6;[2] solved in x-ray diffraction at a resolution of 2.9 Å). The ezrin-FERM-membrane starting structure was produced in the following way:

1. A preliminary PIP_2-containing 'helper membrane' was generated using the web interface of the CHARMM-GUI Membrane Builder.[3,4] It solely served as a full-length PIP_2 source and orientation reference.

2. Structural alignment of the radixin-FERM-inositol-(1,4,5)-triphosphate complex and the helper membrane based on the coordinates of a PIP_2 approximately located at the center of the helper membrane and the coordinates of the inositol-(1,4,5)-triphosphate molecule within the radixin-FERM-inositol-(1,4,5)-triphosphate complex. This was accomplished using the alignment capabilities of UCSF Chimera.[5]

3. Removal of protein-membrane clashes introduced in step 2 by alternate application of small-angle rotations and translations in z-direction (orthogonal to the membrane) to the radixin-FERM-PIP2 complex (realized in VMD[6]).

4. Structural alignment of the ezrin-FERM domain with the radixin-FERM-PIP2 complex (realized in VMD).

5. Clashfree generation of a new membrane around the PIP2-molecule that was docked to the pre-oriented ezrin-FERM-PIP2 complex.

The sequence analysis tool EMBOSS-NEEDLE[7] computes a sequence similarity of $\sim 96\%$ and sequence identity of $\sim 86\%$ based on the amino acid sequences belonging to the selected FERM domain structures of ezrin and radixin. The minimal RMSDs for the inositol-(1,4,5)-triphosphate-PIP2 fit and the ezrin-radixin FERM fit were 1.35 Å and 0.66 Å, respectively. After solvation and addition of sodium and chloride ions at the target concentration of 150 mM the ezrin-FERM simulation system comprised 127557 atoms.

Preparation of the FAK-FERM-membrane starting complex.

The FAK-FERM-membrane complex was assembled in a similar way from the FAK-FERM structure (PDB accession code 6CB0;[8] solved through x-ray diffraction at a resolution of 1.97 Å) and a previously modeled FAK-dimer-PIP2 complex based on a cryo-EM structure of an FAK-dimer (resolution ~ 6 Å) that was subjected to a PIP_2 docking procedure.[9] The FAK-FERM-membrane complex used for all FAK-FERM simulations was built in three steps:

1. Alignment of the FAK-FERM domain with the FERM domain of one of the FAK-monomers comprising the above-mentioned FAK-dimer (realized in VMD).

2. Cutting out of the coordinates of the superimposed FAK-FERM domain (residues 31 to 375) plus the PIP_2-molecule docked at this side of the FAK-dimer.

3. Assembly of the FAK-FERM-membrane complex by uploading the structure obtained in step 2 to the CHARMM-GUI Membrane Builder (analogously to step 5 of the complex assembly for ezrin).

After solvation and addition of sodium and chloride ions at the target concentration of 150 mM the FAK-FERM simulation system comprised 171301 atoms.

Production runs

The interaction of the ezrin- and FAK-FERM domains with the PIP_2-containing lipid bilayers was simulated for 8.1 μs and 6.9 μs in 10 independent replicates, respectively. The precise duration of each replicate is provided in Table S1. During these production runs lipid and protein coordinates were written every 40 ps. To save disk storage, the coordinates of ions (and water), however, were written only once every 10 ns. While this amount of sampling was sufficient for most of analysis performed in this study, it was not sufficient for an appropriate sampling of differently sized PIP_2-clusters required for the intended evaluation of cluster-

wise charge (Figure 4C of main text). For this reason, 6 replicates per FERM domain were continued for 100 ns writing lipid, protein and ion coordinates once every 10 ps.

Table S1: Simulation data overview. The durations of the individual replicates in ns are compiled for both simulation systems. The first six replicates of simulation systems (marked with a "*") were continued for further 100 ns, with more frequent coordinate write-out of the lipid, protein and ion coordinates.

Replicate	1*	2*	3*	4*	5*	6*	7	8	9	10
Ezrin-FERM	900	800	800	800	800	800	800	800	800	800
FAK-FERM	756.12	649.56	791.68	695.52	697.68	686.6	687.44	643.2	648.44	646.6

Although the simulation systems were well adjusted to the target temperature $T = 303\,\mathrm{K}$ and pressure $p = 1\,\mathrm{bar}$ upon completion of the 6-step CHARMM equilibration protocol of $\sim 1.1\,\mathrm{ns}$ duration, the systems have not yet reached an equilibrium state in the sense that quantities characterizing the FERM-PIP_2 interaction have stabilized. Contact maps, encoding the interaction between the FERM domains and PIP_2-molecules in the simulation systems, make for a reasonable measure to quantitatively assess equilibration on the basis of stabilization of FERM-PIP_2 contact formation. We therefore used the contact analysis tool ConAn[10] to determine the number of contacts for each residue-lipid pair. From these contact maps, we computed the stoichiometry as well as the RMSD between the contact map at time t and the initial contact map ($t = 0$),

$$\mathrm{RMSD}_{\mathrm{cmap}} = \|\mathrm{M(t)} - \mathrm{M(t=0)}\|_{\mathrm{Frobenius}}\,, \tag{1}$$

to compute the development of the contact map RMSD over time for each replicate (Figure S1).

Based on these contact map RMSD time series we decided to conservatively omit the first 600 ns of the each ezrin-FERM replicate (Figure **??**A) and the first 450 ns of each FAK-FERM replicate (Figure **??**B) from data analysis. The equilibrated portion of the simulation data thus amounts to 2.1 μs and 2.4 μs for the ezrin- and FAK-FERM domains. To ensure that the remaining simulation time for each replicate is sufficient we determined the residence

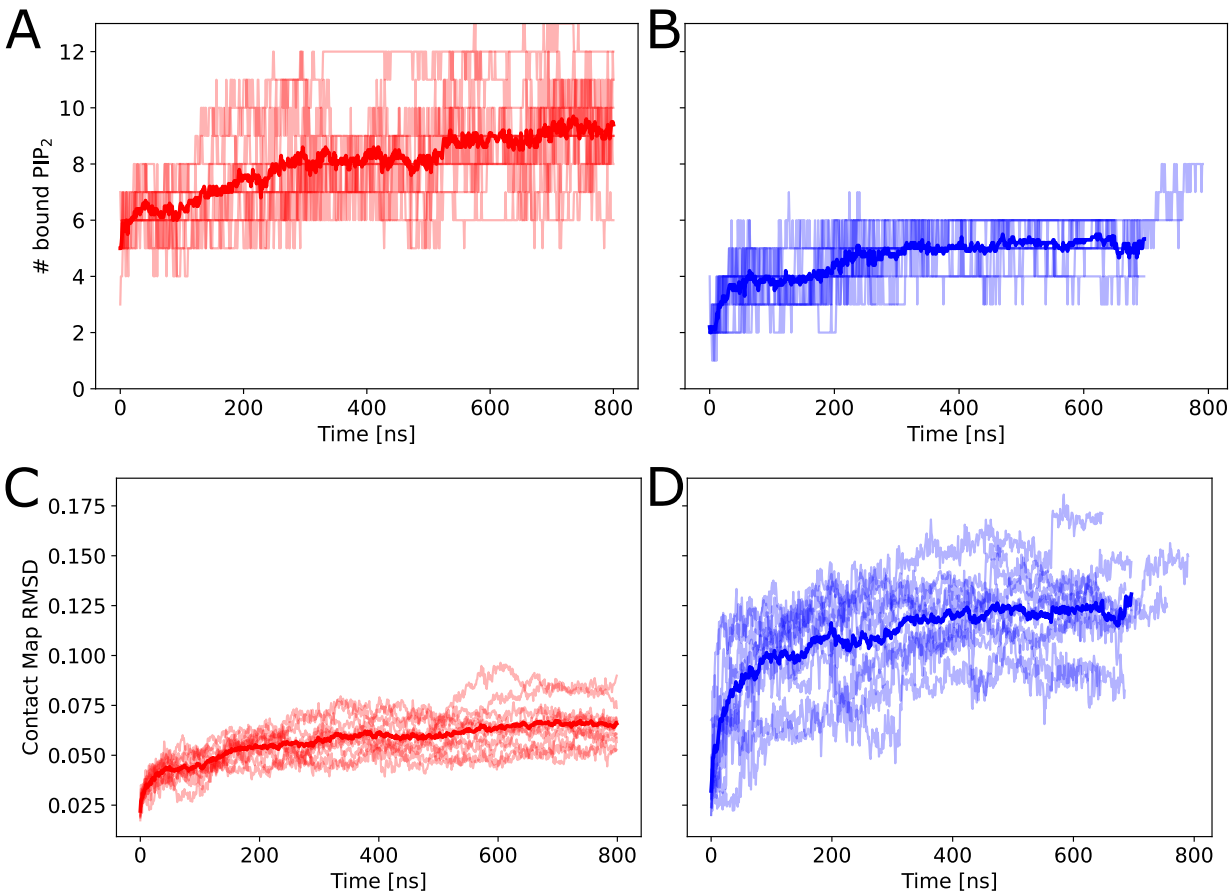

Figure S1: Progression of contacts over time. The upper panel (A,B) shows how the number of bound PIP_2s increases initially. The lower panel (C,D) displays the contact map RMSD with respect to the initial frame of each simulation. Left panels (A,C) represent data for the ezrin-FERM, right panel (B, D) for the FAK-FERM domain. The thin lines represent time series of the individual replicates. Thick lines mark average time series computed from 10 replicates per FERM domain. The time step of RMSD calculation is $\Delta t = 1\,\mathrm{ns}$.

times of each PIP_2 lipid at the FERM domains (Fig. S2). We considered the time the lipids were bound to at least one protein residue. Thus, we only capture unbinding events if a PIP_2 molecule leaves the FERM-membrane interface entirely, but miss unbinding events in the case that PIP_2 immediately binds to another protein residue. The latter likely occurs at the center of the interfaces due to the large amounts of basic residues located there (Fig. 4) and likely amounts for the few observed long residence times spanning large parts of the overall simulation time. However, the vast majority of PIP_2 moleculess stays bound to protein residues for less than 100 ns. Therefore, we conclude that our simulations were conducted for a long enough time despite the large portion discarded for equilibration purposes.

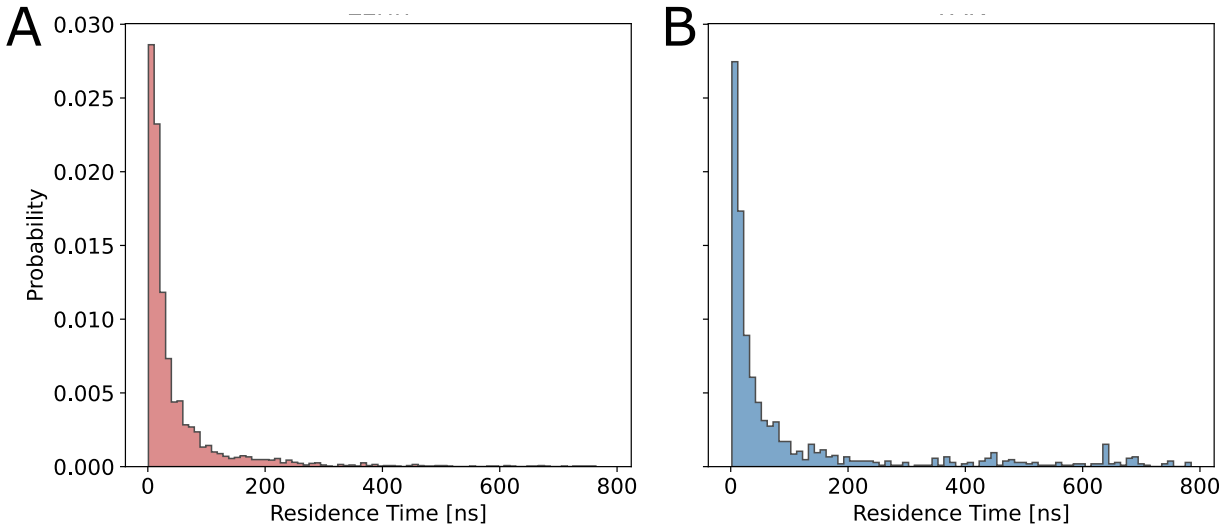

Figure S2: Distribution of PIP_2 residence times at (A) the ezrin-FERM and (B) the FAK-FERM domain.

FERM Domain Positional Variations on the Membrane Surface

To evaluate whether the two FERM domains stayed in their upright position or whether they used potential additional PIP_2 binding sites, we analyzed the angle formed between the protein and the membrane surface (Fig. S3). The angle time traces show fluctuations around 90°, revealing that both FERM domains remain upright albeit with extensive rocking

motions.

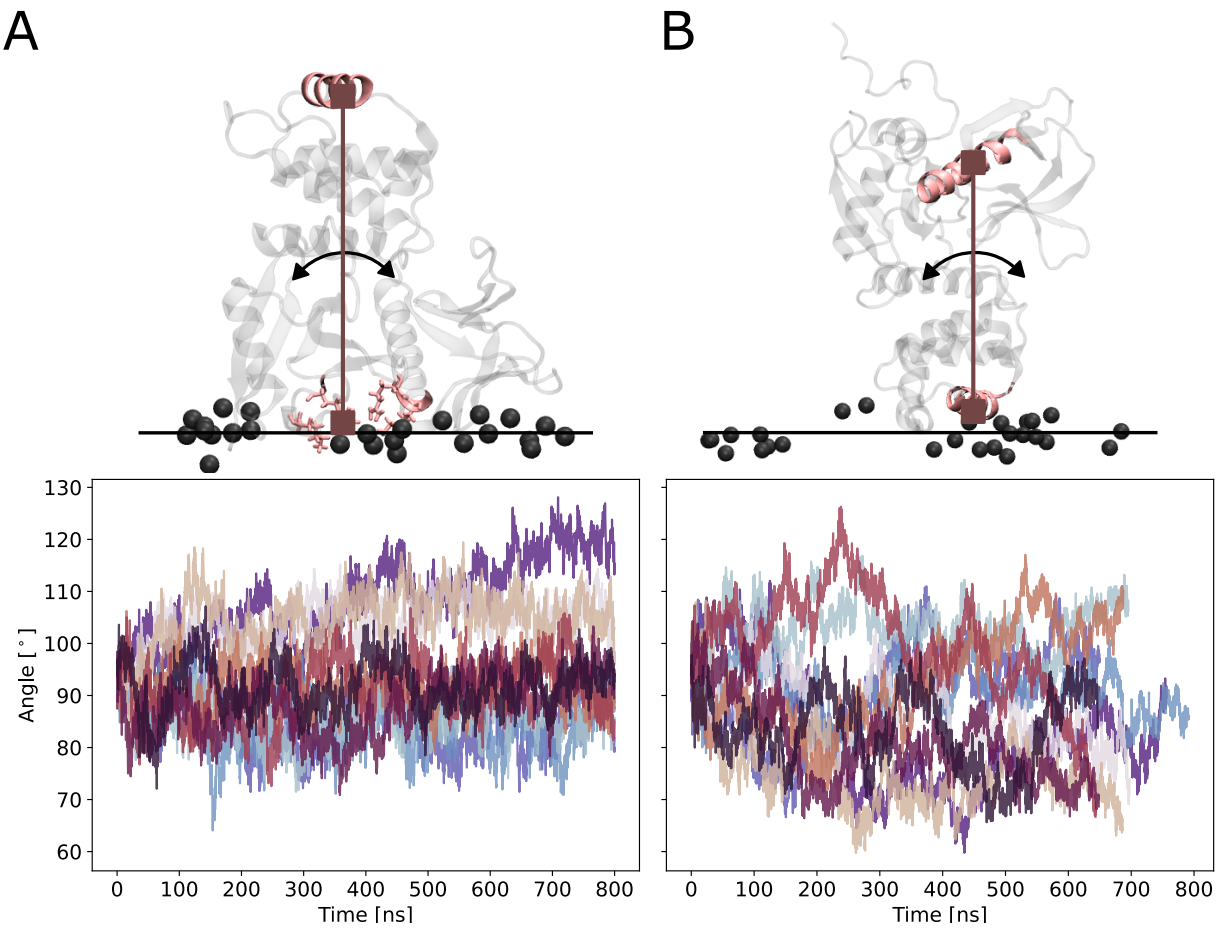

Figure S3: Angle between the membrane located in the xy-plane (black line) and the FERM domain (grey). The top panel illustrates an exemplary simulation frame for each protein. The residues used to define the protein positional vectors (dark pink line) are highlighted in light pink. Black spheres depict the phosphate groups of PIP_2's inositol ring. Each colored line in the bottom panel represents an individual simulation replicate. (A) For ezrin, the vector describing the position of the FERM domain is defined by the basic cleft (COM of residues 60, 63, 278 and 279) and a helix on top of the F2 lobe with respect to the membrane (COM of residues 167-178). (B) The positional vector for FAK is given by the basic patch helix (COM of residues 216-230) and the F3-lobe helix (COM of residues 334-352).

Evaluation of FERM-PIP_2 stoichiometry

Taking into account that the negative charge on a PIP_2 head group is carried by its 4', 5'- and 1'-phosphates (Figure S4), we took a "phosphate-centered" approach when counting the number of PIP_2-molecules bound to a FERM domain at a given time. That is, we considered a PIP_2-molecule to be bound to a FERM domain if the center of one of its phosphate groups, described by the respective position of its P4, P5 or P phosphorus atom, got within 6 Å of the protonated end of a lysine or arginine residue on the FERM domain. Considering the assignment of partial charges according to the CHARMM topology, we assumed the protonated ends to approximately sit at the NZ nitrogen atom for lysines, and at the center between the NH1 and NH2 nitrogen atoms of the guanidium group of arginines (Figure S4).

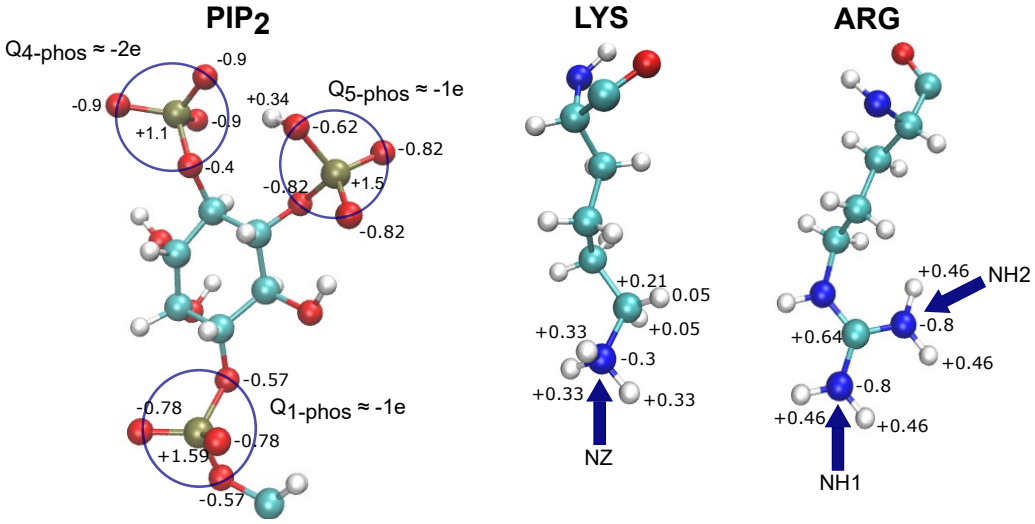

Figure S4: Partial charge assignment. The atoms of a PIP_2 head group, lysine and arginine are labeled according to their partial charges in units of the elementary charge e as defined by the CHARMM topology. Oxygen, phosphorus, carbon, hydrogen, and nitrogen are colored in red, tan, light blue, white and dark blue, respectively. The phosphate groups of the PIP_2-molecule are indicated by blue circles. The partial charges of the labeled atoms at the ends of the lysine and arginine roughly sum to $+1e$. Please note that the relative sizes of the displayed molecules are not true to scale.

The above-mentioned distance cutoff of 6 Å was inferred from the radial distribution functions (RDFs) of the different PIP_2-phosphates around the protonated ends of the lysine and arginine residues (Figure S5C). It was chosen such that secondary peaks reflecting weaker

interaction of 4'-phosphates with lysines/arginines would also be considered when counting the number PIP_2-molecules bound to the FERM domain.

The RDFs were calculated with the GROMACS function *gmx rdf*.[11] Given a selection set

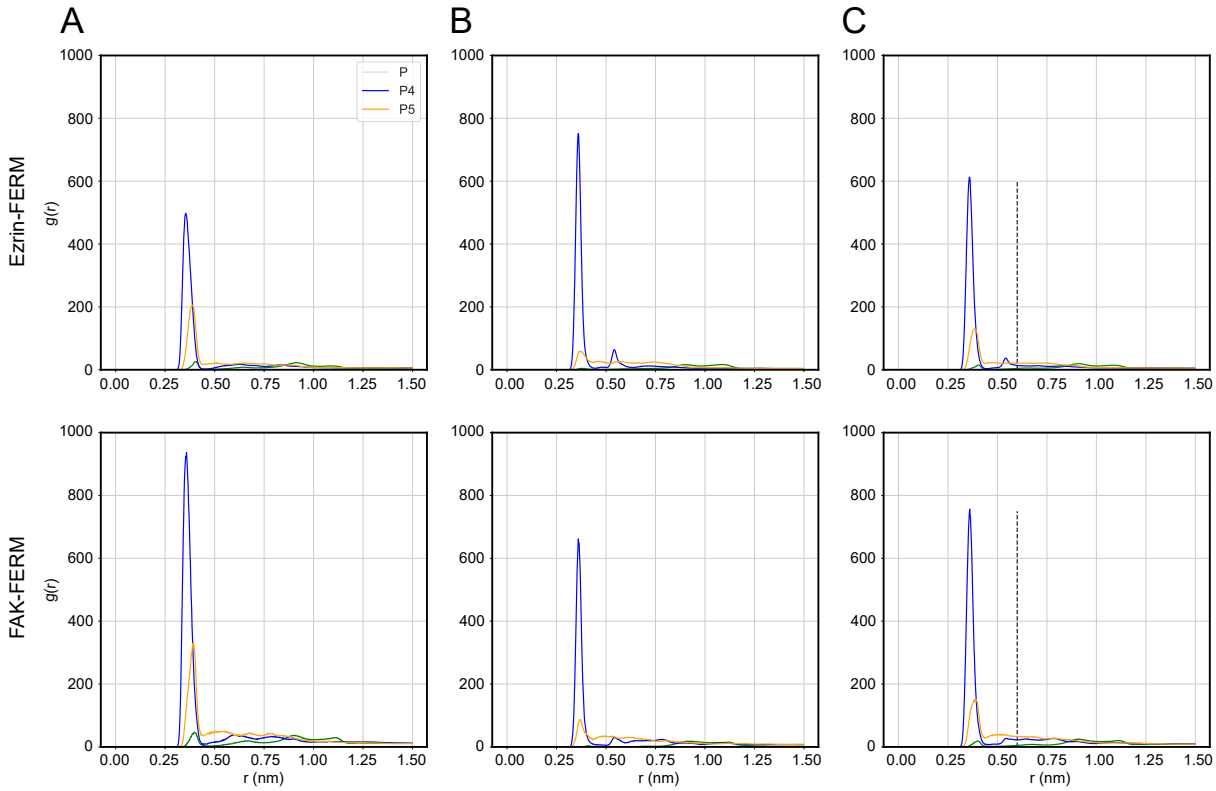

Figure S5: Radial distribution functions of PIP_2-phosphates around lysines and arginines. From left to right the 1-, 4- and 5-phosphate distributions around the lysine NZ-atoms (A), the arginine NH1-/NH2-atoms (B) and the whole set of NZ-/NH1-/NH2-atoms (C) are plotted. The upper and lower panels distinguish between the ezrin and FAK binding poses. The 4-phosphate, 5-phosphate and 1-phosphate RDFs are colored in blue, orange and green, respectively. The RDF curves shown represent averages across the 10 replicates simulated in NaCl per binding pose. For the RDF calculation a maximal sphere size of $r_{\max} = 1.5\,\text{nm}$ and shell thickness of $\Delta r = 0.002\,\text{nm}$ were specified. The lysine and arginine residues included into the RDF calculations were selected after a first rough analysis of contacts within the rather large interaction cutoff of 0.9 nm. A particular lysine/arginine was included when it was observed at least once within the required interaction range. The vertical black dashed line indicates the distance cutoff below which a phosphate was considered to stably interact with a protonated lysine/arginine end.

of atoms B and a reference set of atoms A, *gmx rdf* computes the RDFs of B-atoms around

A-atoms according to

$$g_{\mathrm{AB}}(r) = \frac{\langle n_B(r) \rangle}{\langle n_B \rangle_{\mathrm{local}}} = \frac{1}{\langle n_B \rangle_{\mathrm{local}}} \frac{1}{\mathrm{N_A}} \sum_{\mathrm{i \in A}}^{\mathrm{N_A}} \sum_{\mathrm{j \in B}}^{\mathrm{N_B}} \frac{\delta\left(r_{\mathrm{ij}} - r\right)}{4\pi r^2}, \tag{2}$$

where $\langle n_B \rangle$ denotes the average number density of B-atoms at distance r from A-atoms and $\langle n_B \rangle_{\mathrm{local}}$ represents the number density of B-atoms as averaged over all spheres of radius r_{max} around the A-atoms. In reality, GROMACS calculates an RDF histogram by counting the number of selection set atoms (B-atoms) found in spherical shells of default thickness $\Delta r = 0.002\,\mathrm{nm}$ around the reference set atoms (A-atoms) up to the above-mentioned maximal distance, which is defined by the user. This count is then divided by the number of reference atoms $\mathrm{N_A}$, the shell volume $4\pi r^2 \Delta r$ and the average particle density of the selection atoms $\langle n_B \rangle_{\mathrm{local}}$, rendering $g_{\mathrm{AB}}(r)$ a dimensionless function. When given an entire trajectory *gmx rdf* will also average across the frames comprising the trajectory.

For the computation of the RDFs of the $\mathrm{PIP_2}$-phosphate groups around the lysine and arginine residues of the FERM domain, we chose as representative selection sets

$$\mathrm{B} \in \left\{ \underbrace{\left\{P^{(1)}, \ldots, P^{(\mathrm{N_{PIP2}})}\right\}}_{\text{1-phosphate}}, \underbrace{\left\{P_4^{(1)}, \ldots, P_4^{(\mathrm{N_{PIP2}})}\right\}}_{\text{4-phosphate}}, \underbrace{\left\{P_5^{(1)}, \ldots, P_5^{(\mathrm{N_{PIP2}})}\right\}}_{\text{5-phosphate}} \right\}, \tag{3}$$

marking the centers of the 1-, 4- or 5-phosphate groups of all $\mathrm{N_{PIP2}} = 13$ $\mathrm{PIP_2}$-molecules in the system. To represent the lysine and arginine residues that served as the references for the RDF calculation, we used the atom reference sets

$$\mathrm{A} \in \left\{ \underbrace{\left\{NZ^{(1)}, \ldots, NZ^{(\mathrm{N_{LYS}})}\right\}}_{\text{LYS}}, \underbrace{\left\{NH1^{(1)}, NH2^{(1)}, \ldots, NH1^{(\mathrm{N_{ARG}})}, NH2^{(\mathrm{N_{ARG}})}\right\}}_{\text{ARG}}, \right.$$
$$\left. \underbrace{\left\{NZ^{(1)}, \ldots, NZ^{(\mathrm{N_{LYS}})}, NH1^{(1)}, NH2^{(1)}, \ldots, NH1^{(\mathrm{N_{ARG}})}, NH2^{(\mathrm{N_{ARG}})}\right\}}_{\text{LYS + ARG}} \right\}, \tag{4}$$

where the atoms NZ, NH1 and NH2 as shown in Figure S4, mark the approximate locations of the positive charges carried by the lysine and arginine residues. Note that by specifying the two nitrogen atoms NH1 and NH2 in the reference atom sets for arginine, we do not double the RDF-counts for arginines as compared to the RDF counts for lysine because the RDF as defined in equation (2) is normalized by the number of atoms in the reference set.

Counting in this way the number of PIP_2-molecules bound to the FERM domains for each frame of the equilibrated portion of the simulated trajectories allowed to compute normalized histograms of stoichiometry for both FERM domains (Figure 2A/B of the main text). Based on these discrete probability distributions, mean and standard deviation (STD) were calculated according to

$$\bar{N} = \sum_{\mathrm{n}=\mathrm{n}_{\min}}^{\mathrm{n}_{\max}} p_{\mathrm{n}} n \tag{5}$$

$$\sigma_{\bar{N}} = \left(\sum_{\mathrm{n}=\mathrm{n}_{\min}}^{\mathrm{n}_{\max}} p_{\mathrm{n}} (n - \bar{N})^2 \right)^{1/2}, \tag{6}$$

leading to the values

$$\bar{\mathrm{N}}_{\mathrm{pip2}} = \begin{cases} 7.4 \pm 1.3 & \text{for the ezrin-FERM binding pose} \\ 4.6 \pm 0.8 & \text{for the FAK-FERM binding pose} \end{cases}.$$

reported in the main text.

Contact Analysis

Analysis of residue-wise FERM-PIP_2 contacts and the contributions of FERM-PIP_2 contacts to the overall FERM-membrane contacts was performed with ConAn. Precisely, in what is called the "asymmetric mode", designed for the analysis of inter-molecular interactions, we

ran ConAn on the concatenation of the equilibrated portions of the MD trajectories.
For determination of residue-wise FERM-PIP_2 contacts we peformed three such runs, specifying as the two molecule groups the residue ID range capturing the 13 PIP_2-molecules in the system and

1. ... residues 2 - 86 corresponding to the F1-lobe of the ezrin-FERM domain
2. ... residues 204 - 297 corresponding to the F3-lobe of the ezrin-FERM domain
3. ... residues 180 - 227 corresponding to the F2-lobe of the FAK-FERM domain.

The parameters $d_{\text{inter}} = 0.4\,\text{nm}$ and $d_{\text{high}} = 0.6\,\text{nm}$ introduced in the methods section of the main text correspond to the ConAn input parameters "TRUNC_INTER" and "TRUNC_INTER_HIGH".
To estimate the contribution of FERM-PIP_2 contacts to the overall number of contacts formed between the FERM domains and the membrane, the above asymmetric runs were repeated, this time including also the POPC-molecules of the protein-proximal membrane leaflet into the contact analysis.

Charge compensation analysis

Our analysis of charge compensation in the FERM-membrane interaction interface and on the level of PIP_2-clusters includes charges contributed by lysine and arginine residues, sodium and chloride ions and PIP_2-molecules considered to be in interaction based on distance criteria as explained below. Taking into account the five theoretically possible types of intermolecular hydrogen bond formation between adjacent PIP_2-molecules, 4/5-phosphate-4/5-phosphate, 4/5-phosphate-hydroxyl, hydroxyl-hydroxyl, 4/5-phosphate-1-phosphate, and hydroxyl-1-phosphate,[12] our criterion for clustering PIP_2-molecules together was to require at least one oxygen-oxygen encounter of adjacent PIP_2-molecules within $r_{\text{O}-\text{O}} \leq 0.33\,\text{nm}$, which proved to capture hydrogen bonding well in previous MD study on PIP_2-clustering.[13]

(i) Computation of ion concentration profiles

To compute concentration profiles for the sodium and chloride ions, the simulation system was cut into n slices of thickness $\Delta z = 0.25\,\mathrm{nm}$ along the z-direction, i.e. orthogonal to membrane plane. The ion concentration profiles $c_{\mathrm{Na^+}}(z)$ and $c_{\mathrm{Cl^-}}(z)$ were then computed from the trajectory counts $\mathrm{N}_{\mathrm{ion}}^{(\mathrm{i})}(z)$, $\mathrm{i} = 1, \ldots, \mathrm{n}$ of both ion species in the different slices via

$$c_{\mathrm{ion}}(z) = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{\mathrm{N}_{\mathrm{ion}}^{(\mathrm{i})}(z)}{\mathrm{N_A} A \Delta z}, \tag{7}$$

where $\mathrm{N_A}$ and A denote Avogadros's number and the slice area respectively (Figure S6).

The concentration profiles represent averages across the above-mentioned 6 x 100 ns simulation bits with higher coordinate write-out frequency appended to the first six replicates for each binding pose. For reference, also the mass densities of both FERM domains reflecting the characteristic orientations of the two FERM-domains with respect to the plasma membrane are plotted. In both simulation systems the ion concentration far from the membrane settled at $c_0 = 200\,\mathrm{mM}$ and also the peak sodium concentrations, reached at the approximate average z-coordinate of the PIP_2 head groups, was similar.

(ii) Charge compensation in the FERM-membrane interaction interface

The minutiae of the analysis of the total charge in the FERM-PIP_2 interaction interface are described next. Since the system was prepared to be overall charge-neutral, enlarging a search volume around the PIP_2-molecules checked for charge neutralization must at some point yield a neutral net charge. We thus had to make a choice for the size of the search volume that adequately reflects the zone of FERM-PIP_2 interaction. We defined the search volume as follows:

1. Identify all positively charged protein residues and PIP_2-molecules that form phosphate-lysine/arginine contacts in the above sense.

2. Enlarge search spheres of radius $\mathrm{r_{sph}}$ around the protonated ends of the lysines/arginines

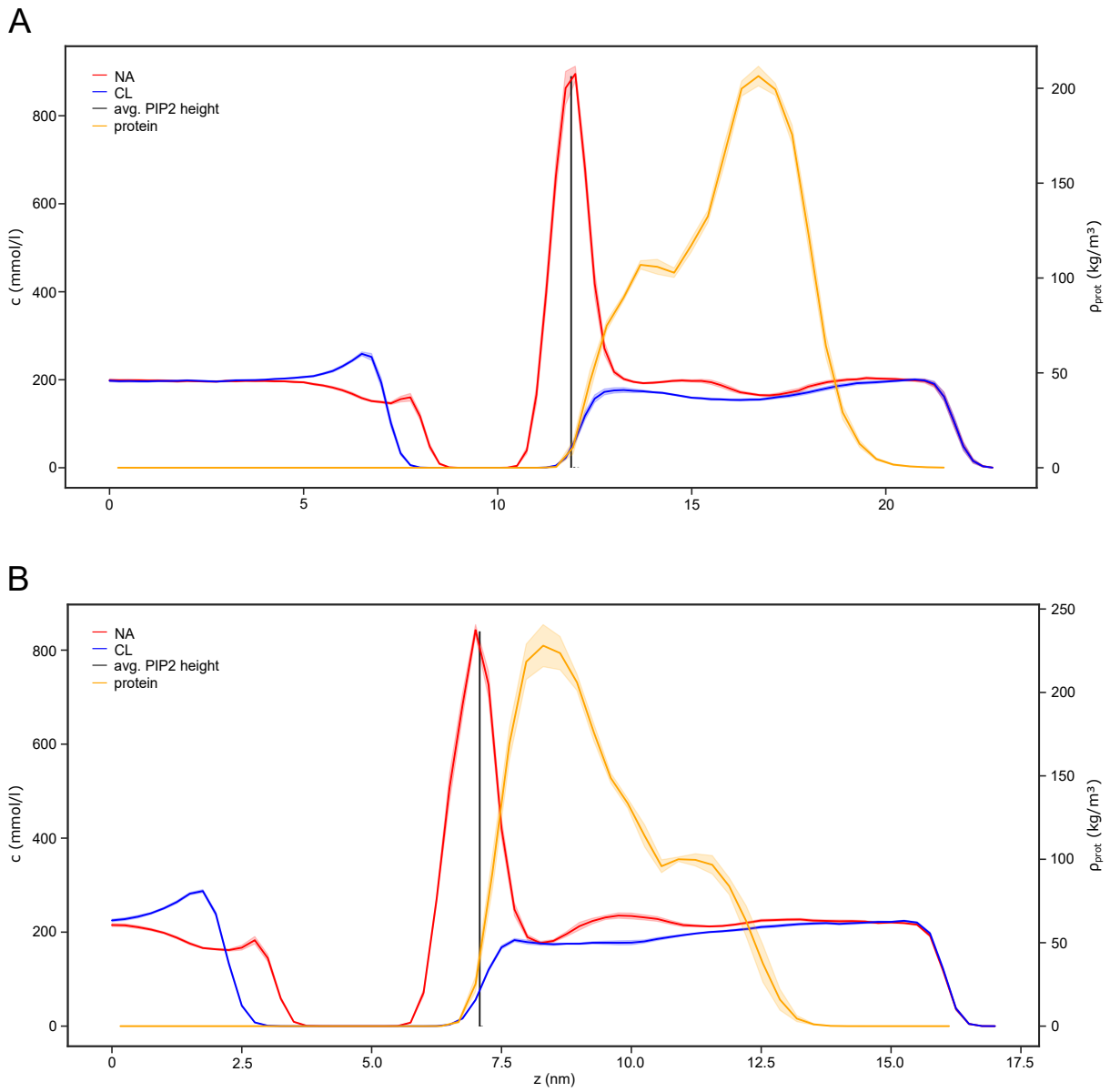

Figure S6: Ion concentration profiles of sodium and chloride along the z-direction computed from (A) the FAK-FERM simulations, and (B) the ezrin-FERM simulations. The profiles for sodium (red) and chloride (blue) represent averages across the above-mentioned 6 x 100 ns simulation bits with higher coordinate write-out frequency appended to the first six replicates for each binding pose. The shaded areas mark the standard deviation zones. For reference, the mass density profiles of both proteins are plotted in orange.

and search for chloride ions.

3. Enlarge search spheres of radius r_{sph} around the P-, P4- and P5-atoms at the centers of the PIP_2-phosphates and search for sodium ions.

4. Add the charges of protein residues and PIP_2-molecules in contact, the charges of chloride ions from 2. and the charges of the sodium ions from 3. to obtain Q_{tot}.

Evaluating time-averages of the total charge found according to the above procedure for the 10 repeat simulations available per binding pose, one obtains 10 independent Q_{tot}-values per binding pose, from which we calculated a global mean and standard deviation of the mean. This calculation was repeated for different search sphere radii between 0.5 nm and 1.0 nm (Figure S7).

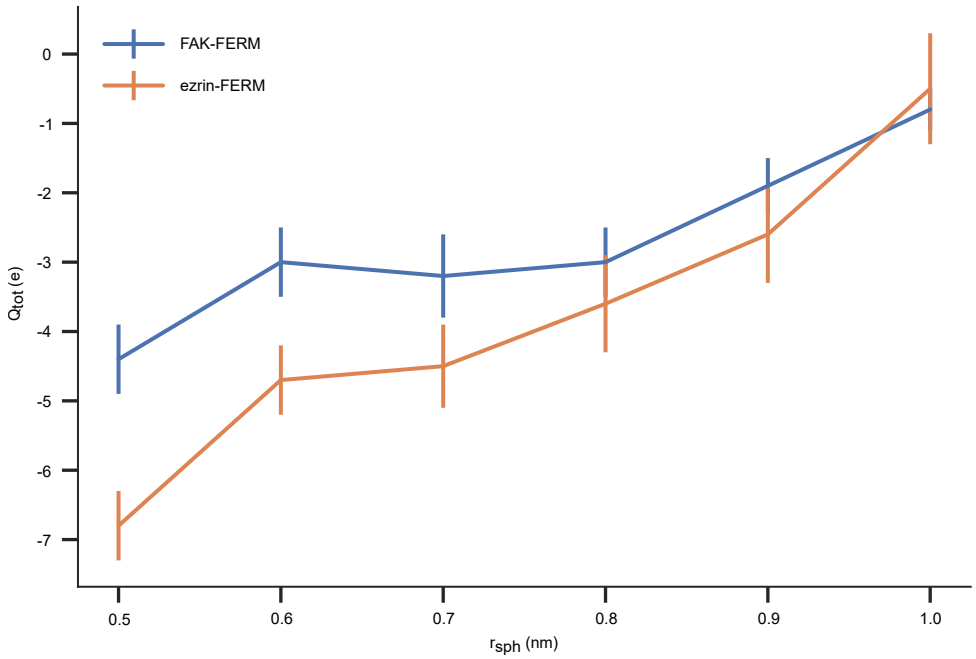

Figure S7: Total interface charge as a function of the search sphere radius.

Evaluating $Q_{tot}(r_{sph})$ at our previously defined cutoff $r_{cut} = 0.6$ nm for phosphate-lysine/arginine interaction, with the idea to include only stably interacting charges into the total charge calculation, yields total interface charges of $Q_{tot}^{ezr}(r_{cut}) \approx -5e$ and $Q_{tot}^{FAK}(r_{cut}) \approx -3e$ for the ezrin- and FAK-FERM domains as reported in the main text of the paper. As ex-

pected, $Q_{\mathrm{tot}}(\mathrm{r_{sph}})$ approaches zero as $\mathrm{r_{sph}}$ is increased. The observation of a slightly negative charge excess in the binding interface zone persists, however, in the intermediate range $\mathrm{r_{sph}} \in \{0.6, 0.7, 0.8, 0.9\}$ nm.

(iii) Charge compensation of PIP_2-molecules and clusters

To analyze charge compensation more locally on the level of individual PIP_2-molecules or clusters, we divided the 13 PIP_2-molecules present in both systems at each frame of the 6 x 100 ns simulation bits per binding pose with higher coordinate write-out frequency into clusters. To this end we deployed the hierarchical single-linkage clustering algorithm, using as a linking criterion an oxygen-oxygen encounter within a distance of $r_{\mathrm{O-O}} \leq 0.33\,\mathrm{nm}$, representing inter-molecular hydrogen bonding. To evaluate the total charge of a PIP_2-cluster we proceeded similarly to the protocol for evaluating charge compensation in the FERM-membrane interface. Precisely, total charge on the cluster level was computed by:

1. identifying the cluster size,
2. counting the number of sodium ions, chloride ions, or protonated ends of lysines/arginines within a distance of $\mathrm{r_{sph}}$ of the phosphate groups of the PIP_2-molecules constituting the cluster, and
3. adding the thus collected charges to the charge $Q_{\mathrm{clust}} = \mathrm{N_{pip2}} * (-4e)$ carried by the $\mathrm{N_{pip2}}$ PIP_2-molecules in the respective cluster.

As for the analysis of the total interface charge in the previous paragraph, we set $\mathrm{r_{sph}} = 0.6\,\mathrm{nm}$.

(iv) Sodium ions substantially aid compensation of PIP_2 charge.

As PIP_2-molecules were simulated with a net charge of $-4\,e$ per molecule, reflecting the

experimentally determined[12] PIP_2 net charge at pH 7.0, the rather high average binding counts we observed for the ezrin- and FAK-FERM domains correspond to a negative charge of $\sim -30\,e$ for the ezrin-FERM domain and $\sim -18\,e$ for the FAK-FERM domain. By contrast, the number of positively charged FERM residues in pronounced interaction ($\bar{n}_{res} \geq 0.5$ avg. contacts/frame) with PIP_2 amounts to only 15 for the ezrin-FERM domain and 7 for the FAK-FERM domain (Figure 4E/F). Consequently, the bound PIP_2-molecules clearly overcompensate the protein surface charges by approximately $-10\,e$ to $-15\,e$. Given this considerable excess of negative charge, the question arises if and how it is compensated to avoid a potential destabilization of FERM-membrane binding. Sodium concentration is maximal at the mean height of the PIP_2-headgroups (Figure S6), suggesting that sodium ions might enter the FERM-membrane binding interface. Chloride concentration, on the other hand, quickly decays toward the lipid bilayer and due to the comparatively large size of the chloride ions is strongly suppressed in the binding interface. To quantify the contribution of sodium ions in the compensation of PIP_2 charge, we decomposed the total charge in the binding interface into its different contributions (Figure S8B). The total interface charge was computed as the sum of protein, PIP_2, sodium and chloride charges.

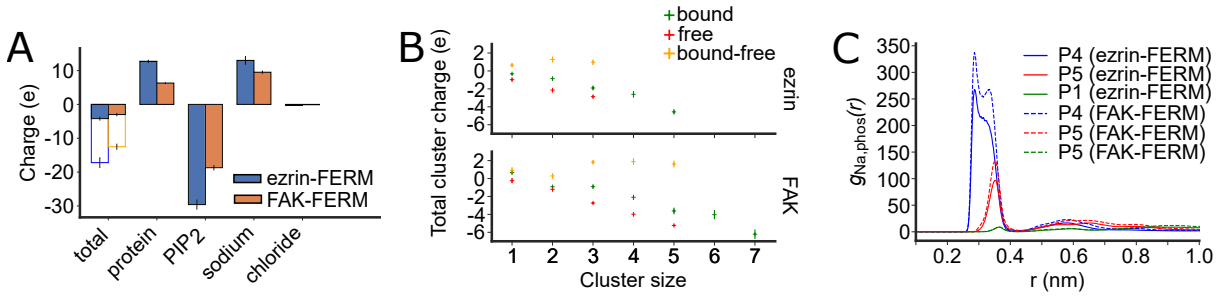

Figure S8: Sodium aids the compensation of PIP_2 charge in the binding interface and on the level of individual PIP_2-clusters. (A) Decomposition of total interface charge. Error bars denote standard errors of the mean computed from $n = 10$ replicates per FERM-domain. (B)Total charge of PIP_2-clusters as a function of cluster size. Error bars denote standard errors of the mean obtained in block averaging. (C) Radial distribution of sodium around phosphorus atoms at the center of the 1'-, 4'- and 5'-phosphate groups of PIP_2.

Scrutinizing the total interface charge in this manner, we find the large amount of negative charge carried by the PIP_2-molecules reduced to $-5\,e$ and $-2\,e$ in the ezrin-FERM

and FAK-FERM simulations due to positive FERM residues and sodium ions contributing positive charge to the interface in roughly equal parts. We hence conclude that sodium ions enter the FERM-membrane binding interface, where they almost completely compensate the excess of negative charge caused by the large stoichiometry of FERM-PIP_2 binding and thus avoid potential destabilization of FERM-membrane binding due to electrostatic repulsion between PIP_2-molecules.

(v) PIP_2 diffusion is electrostatically biased toward positive FERM domain surface. Having found sodium ions to considerably contribute to charge neutralization of the FERM-membrane binding interface and with in principle enough sodium ions in the system to compensate the negative charge of the laterally diffusing PIP_2-molecules, we wondered what drives PIP_2-molecules in such high numbers to the FERM domain surface in the first place. To answer this question we investigated charge compensation of PIP_2-molecules on a more local scale. Taking into account the formation of PIP_2-clusters due to inter-molecular hydrogen bonding,[13] we analyzed the total charge of PIP_2-clusters as a function of cluster size, distinguishing between protein-bound and freely diffusing PIP_2-clusters (Figure S8C). To start with, we note that the average total charge of freely diffusing PIP_2-clusters decreases linearly at a rate of roughly $-1\,e$ per PIP_2-molecule in a cluster. Recalling the net negative charge of PIP_2 to be $-4\,e$, this is only possible when sodium ions and PIP_2-molecules on average associate at a ratio of $\sim$3:1. The radial distribution functions (RDFs) of sodium around the 4'-, 5'- and 1'-phosphate groups of the PIP_2-molecules in the system (Figure S8D) corroborate this reasoning with a double peak in the sodium RDF around the 4'-phosphate and a single peak in the sodium RDF around the 5'-phosphate of PIP_2. Consequently, freely diffusing PIP_2-molecules and clusters, despite their partial neutralization by sodium ions, retain a negative net charge that increases linearly with growing cluster size. This biases lateral diffusion of PIP_2-molecules and clusters toward regions on the FERM domain sur-

face that still excess positive charge in the form of lysine and arginine residues. Calculating the difference between the total charges of protein-bound and freely diffusing PIP_2-clusters (Figure S8C), one finds that upon binding the ezrin- or FAK-FERM domains PIP_2-clusters gain on average $+1\,e$ to $+2\,e$ of charge, which in view of their previous excess of negative charge should be energetically favorable.

Calculation of FERM-Membrane Binding Free Energy

The Gibbs free energy of binding ΔG was calculated using Umbrella Sampling as described previously[14] (Fig. S9). We randomly selected three starting frames for each protein from the equilibrium MD simulations, in which the FERM domains where bound by the average number of PIP_2s. We pulled on the protein backbone and lower leaflet of the membrane in opposite directions with a constant velocity of 0.05 m/s and a spring constant of 500 kJ/mol/nm^2. To avoid shearing of the FERM-domain lobes and removal of individual lipids from the membrane we additionally restrained the distance of the four closest heavy atom pairs between the F1 und F3 lobes with a force constant of 50 kJ/mol/nm^2 and the position of the PIP_2 head groups to the height of the membrane plane with a force constant of 1000 kJ/mol/nm^2. Umbrella sampling windows were extracted from the pulling trajectories with a spacing of 0.1 nm during the first 1 nm distance increase between the center of mass of the protein and lower membrane leaflet and 0.2 nm afterwards to cover a distance of at least 4 nm. We simulated each window for 50 ns and discarded the first 10 ns prior to analysis with the weighted histogram analysis method.[15]

Calculation of protein area coverage and packing parameter

To calculate the area A_{one} covered by each FERM domain, we determined the density of the domains using the Gromaps package[16] and projected these densities onto the membrane

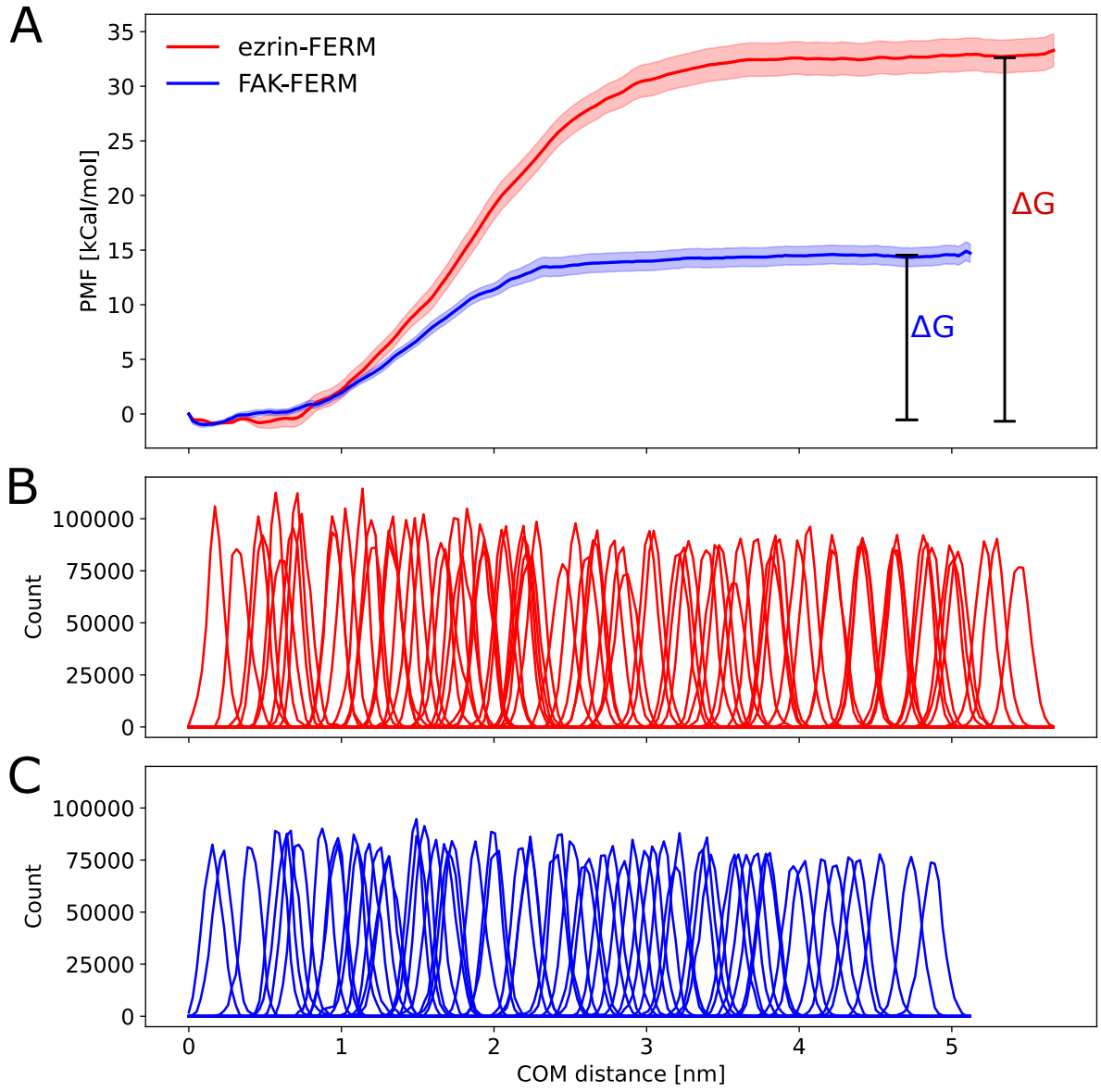

Figure S9: PMF-trace and histograms of Umbrella Sampling calculations. (A) PMF-traces determined by the weighted histogram analysis method. ΔG is calculated as the difference between the lower and upper plateau of the curves. Panels (B) and (C) show the corresponding histograms for ezrin and FAK, respectively.

plane. For estimation of the packing parameter α, we subjected an array of 64 domains to lateral pressure.

$$\alpha = \frac{64 A_{\mathrm{one}}}{A_{\mathrm{Array}}} \tag{8}$$

with A_{Array} being the total area of the protein array.

Model for binding of proteins to multiple surface-bound ligands

The protein has n binding sites and a footprint of A_{P}. The ligand has a footprint of $A_{\mathrm{L}} < A_{\mathrm{P}}$. The area covered by of one FAK-FERM was $A_{\mathrm{one}} = 18.6 \pm 0.9\,\mathrm{nm}^2$ and for ezrin-FERM $A_{\mathrm{one}} = 15.1 \pm 1.0\,\mathrm{nm}^2$ which was divided by the maximal surface coverage ($\alpha_{\mathrm{FAK-FERM}} = 0.73 \pm 0.04$ and $\alpha_{\mathrm{ezrin-FERM}} = 0.50 \pm 0.02$) to get the respective A_{P}. The values for the area covered by one FERM-domain and the maximal surface coverage are derived from the MD simulations.

Protein coverage

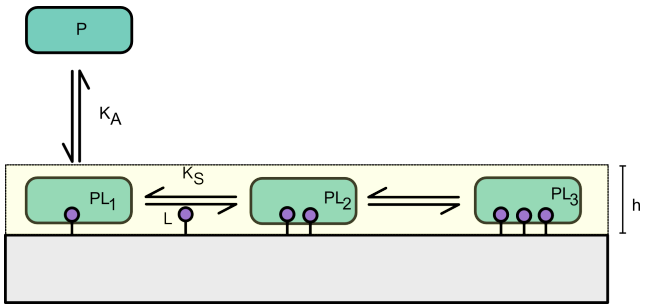

Figure S10: Adsorption of proteins P onto specific sites (ligands L).

Here, we sum only the *occupied* binding sites. The protein may cover 1, 2 or more ligands.

$$\Gamma_{\mathrm{P}} = \frac{N_{\mathrm{P}}}{A} = \Gamma_{\mathrm{PL}_1} + \Gamma_{\mathrm{PL}_2} + ... + \Gamma_{\mathrm{PL_n}} \tag{9}$$

where N_{P} is the number (moles) of protein on the surface of area A. Γ denotes the respective surface concentration.

Conservation of ligands

The number of ligands is constant. Initially, i.e., prior to binding the number was N_{L0}:

$$N_{\mathrm{L0}} = N_{\mathrm{PL}_1} + 2N_{\mathrm{PL}_2} + 3N_{\mathrm{PL}_3} + ... + nN_{\mathrm{PL_n}} + N_{\mathrm{L}} \tag{10}$$

N_{L} is the number of non-bound ligands on the surface. Written in surface concentrations we obtain:

$$\Gamma_{\mathrm{L0}} = \Gamma_{\mathrm{PL}} + 2\Gamma_{\mathrm{PL}_2} + 3\Gamma_{\mathrm{PL}_3} + ... + n\Gamma_{\mathrm{PL_n}} + \Gamma_{\mathrm{L}} \tag{11}$$

Multiple binding reactions

$$\mathrm{PL} + \mathrm{L} \rightleftharpoons \mathrm{PL}_2, \mathrm{PL}_2 + \mathrm{L} \rightleftharpoons \mathrm{PL}_3, \mathrm{PL_{n\text{-}1}} + \mathrm{L} \rightleftharpoons \mathrm{PL_n} \tag{12}$$

The first mass of action reaction with the surface binding constant $K_{\mathrm{S}}[\mathrm{m}^2/\mathrm{mol}]$ reads:

$$K_{\mathrm{S}} = \frac{\Gamma_{\mathrm{PL}_2}}{\Gamma_{\mathrm{PL}}\Gamma_{\mathrm{L}}} \tag{13}$$

The second reaction reads:

$$K_{\mathrm{S}} = \frac{\Gamma_{\mathrm{PL}_3}}{\Gamma_{\mathrm{PL}_2}\Gamma_{\mathrm{L}}} \tag{14}$$

and consequently we find for the last reaction after plugging in all the previous reactions $\Gamma_{\mathrm{PL(n-1)}}$:

$$\Gamma_{\mathrm{PLn}} = (K_{\mathrm{S}}\Gamma_{\mathrm{L}})^{n-1}\Gamma_{\mathrm{PL}_1} \tag{15}$$

Note that the association constant $K_{\mathrm{A}}[\mathrm{m}^3/\mathrm{mol}]$ reads:

$$K_{\mathrm{A}} = \frac{\Gamma_{\mathrm{PL}}}{c_{\mathrm{P}}\Gamma_{\mathrm{L}}} \tag{16}$$

It is possible to transform between two and three dimensions a protein encounters through the simple expression $K(2D) = K(3D)/h$, where h is the confinement length. The basic idea is that if two interacting species are confined to a region of length $h \approx 1 - 10\,\mathrm{nm}$ along an axis perpendicular to the plane of a membrane, then they are effectively confinded to a volume $V = Ah$, where A is the area per molecule. This simple procedure turns a 2D system into a 'quasi-3D system' because there is now a volume associated with each molecule even when it is constrained to a planar membrane. The result is not influenced by the height between $1 - 10\,\mathrm{nm}$. The protein concentration in solution can then be written as:

$$c_{\mathrm{P}} = \frac{N_{\mathrm{P}}}{Ah} = \frac{\Gamma_{\mathrm{P}}}{h} \tag{17}$$

inserting in equation (16) gives:

$$K_{\mathrm{A}} = \frac{\Gamma_{\mathrm{PL}} h}{\Gamma_{\mathrm{P}} \Gamma_{\mathrm{L}}} = h K_{\mathrm{S}} \tag{18}$$

That means if a protein binds to the first ligand it moves into the surface slab, where the second site encounters a higher apparent concentration of ligands. Mathematically this is equivalent to the binding of the first ligand increasing the association constant for the second site.

K_{A} was calculated with the equation $\Delta G = -RT \ln(K_{\mathrm{A}})$ where we inserted the Gibbs free energy change from the MD simulations at 21 °C ($\Delta G_{\mathrm{FAK-FERM}} = -12.4 \pm 2.2\,\mathrm{kcal/mol}$, $\Delta G_{\mathrm{ezrin-FERM}} = -30.5 \pm 4.4\,\mathrm{kcal/mol}$).

Protein concentration on the surface

Plugging equation (15) into the conservation equation (11) we arrive at:

$$\Gamma_{\mathrm{L0}} = \Gamma_{\mathrm{PL}} + \sum_{i=2}^{n} i (K_{\mathrm{S}} \Gamma_{\mathrm{L}})^{i-1} \Gamma_{\mathrm{PL}} + \Gamma_{\mathrm{L}} \tag{19}$$

or

$$\Gamma_{\mathrm{PL}} = \frac{\Gamma_{\mathrm{L0}} - \Gamma_{\mathrm{L}}}{\sum_{i=1}^{n} i (K_{\mathrm{S}} \Gamma_{\mathrm{L}})^{i-1}} \tag{20}$$

and eventually by summing up the expression for $\Gamma_{\mathrm{PL_2}}, \Gamma_{\mathrm{PL_3}} ... \Gamma_{\mathrm{PL_n}}$ plugged into equation (9):

$$\Gamma_{\mathrm{P}} = \frac{(\Gamma_{\mathrm{L0}} - \Gamma_{\mathrm{L}})(1 + K_{\mathrm{S}} \Gamma_{\mathrm{L}} + ... + (K_{\mathrm{S}} \Gamma_{\mathrm{L}})^{n-1})}{\sum_{i=1}^{n} i (K_{\mathrm{S}} \Gamma_{\mathrm{L}})^{i-1}} \tag{21}$$

or

$$\Gamma_{\mathrm{P}} = \frac{(\Gamma_{\mathrm{L0}} - \Gamma_{\mathrm{L}}) \sum_{i=1}^{n} (K_{\mathrm{S}}\Gamma_{\mathrm{L}})^{i-1}}{\sum_{i=1}^{n} i(K_{\mathrm{S}}\Gamma_{\mathrm{L}})^{i-1}} \tag{22}$$

Translating into bound fractions $\Theta_{\mathrm{P}} = \Gamma_{\mathrm{P}}/\Gamma_{\mathrm{P}}^{\mathrm{max}} = \Gamma_{\mathrm{P}}(N_{\mathrm{A}}A_{\mathrm{P}})$ since $\Gamma_{\mathrm{P}}^{\mathrm{max}} = 1/(N_{\mathrm{A}}A_{\mathrm{P}})$:

$$\Theta_{\mathrm{P}} = A_{\mathrm{P}}N_{\mathrm{A}}(\Gamma_{\mathrm{L0}} - \Gamma_{\mathrm{L}}) \frac{\sum_{i=1}^{n} (K_{\mathrm{S}}\Gamma_{\mathrm{L}})^{i-1}}{\sum_{i=1}^{n} i(K_{\mathrm{S}}\Gamma_{\mathrm{L}})^{i-1}} \tag{23}$$

Figure S11 shows the input data of the procedure for the FAK-FERM derived from experimental data from which the free PIP_2 concentration (ligand) was calculated with an assumed stoichiometry of $n = 2$ and $n = 3$. If $n = 2$ is increased to $n = 3$, the free PIP_2 concentration reaches values below zero indicating that all protein bindings sites are occupied.

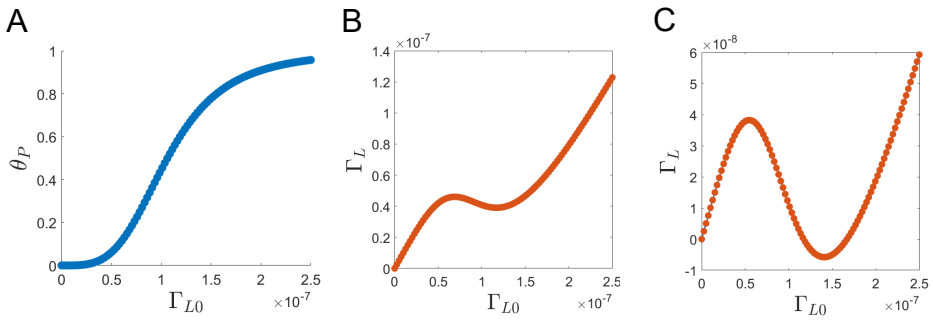

Figure S11: Simulation of free PIP_2 concentration. (A) Input data - FAK-FERM coverage as a function of surface concentration of PIP_2. (B) Free ligand concentration as a function of PIP_2 content with an assumed stoichiometry of $n = 2$. (C) Free ligand concentration as a function of PIP_2 content with an assumed stoichiometry of $n = 3$.

References

(1) Phang, J. M.; Harrop, S. J.; Duff, A. P.; Sokolova, A. V.; Crossett, B.; Walsh, J. C.; Beckham, S. A.; Nguyen, C. D.; Davies, R. B.; Glöckner, C., et al. Structural characterization suggests models for monomeric and dimeric forms of full-length ezrin. *Biochemical Journal* **2016**, *473*, 2763–2782.

(2) Hamada, K.; Shimizu, T.; Matsui, T.; Tsukita, S.; Tsukita, S.; Hakoshima, T. Structural basis of the membrane-targeting and unmasking mechanisms of the radixin FERM domain. *The EMBO journal* **2000**, *19*, 4449–4462.

(3) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry* **2008**, *29*, 1859–1865.

(4) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y., et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Journal of chemical theory and computation* **2016**, *12*, 405–413.

(5) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004**, *25*, 1605–1612.

(6) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **1996**, *14*, 33–38.

(7) Madeira, F.; Park, Y. M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R.; Potter, S. C.; Finn, R. D., et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research* **2019**, *47*, W636–W641.

(8) Marlowe, T.; Dementiev, A.; Figel, S.; Rivera, A.; Flavin, M.; Cance, W. High resolution crystal structure of the FAK FERM domain reveals new insights on the Druggability of tyrosine 397 and the Src SH3 binding site. *BMC molecular and cell biology* **2019**, *20*, 1–9.

(9) Goñi, G. M.; Epifano, C.; Boskovic, J.; Camacho-Artacho, M.; Zhou, J.; Bronowska, A.; Martín, M. T.; Eck, M. J.; Kremer, L.; Gräter, F., et al. Phosphatidylinositol 4, 5-bisphosphate triggers activation of focal adhesion kinase by inducing clustering and

conformational changes. *Proceedings of the National Academy of Sciences* **2014**, *111*, E3177–E3186.

(10) Mercadante, D.; Gräter, F.; Daday, C. CONAN: a tool to decode dynamical information from molecular interaction maps. *Biophysical journal* **2018**, *114*, 1267–1273.

(11) Abraham, M.; Van Der Spoel, D.; Lindahl, E.; Hess, B., et al. GROMACS user manual version 5.0. 4. *Sweden: Royal Institute of Technology and Uppsala University* **2014**,

(12) Kooijman, E. E.; King, K. E.; Gangoda, M.; Gericke, A. Ionization properties of phosphatidylinositol polyphosphates in mixed model membranes. *Biochemistry* **2009**, *48*, 9360–9371.

(13) Han, K.; Gericke, A.; Pastor, R. W. Characterization of specific ion effects on PI (4, 5) P2 clustering: molecular dynamics simulations and graph-theoretic analysis. *The Journal of Physical Chemistry B* **2020**, *124*, 1183–1196.

(14) Lemkul, J.; Bevan, D. Assessing the Stability of Alzheimer's Amyloid Protofibrils Using Molecular Dynamics. *Journal of Physical Chemistry B* **2010**, *114*, 1652–1660.

(15) Hub, J. S.; de Groot, B. L.; Van Der Spoel, D. g_{wham} - A free weighted histogram analysis implementation including robust error and autocorrelation estimates. *Journal of Chemical Theory and Computation* **2010**, *6*, 3713–3720.

(16) Briones, R.; Blau, C.; Kutzner, C.; de Groot, B. L.; Aponte-Santamaría, C. GROmaρs: A GROMACS-Based Toolset to Analyze Density Maps Derived from Molecular Dynamics Simulations. *Biophysical Journal* **2019**, *116*, 4–11.