Prediction of Noncompetitive Inhibitor Binding Mode Reveals

Promising Site for Allosteric Modulation of Falcipain-2

Jorge Enrique Hernández González,[†] Lilian Hernández Alvarez,[†] Pedro Geraldo Pascutti,[‡] Vitor B. P. Leite^{†,§}

[†]Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista Júlio de Mesquita Filho, Rua Cristóvão Colombo, 2265, Jardim Nazareth, São José do Rio Preto, São Paulo, CEP 15054-000, Brazil.

[‡]Laboratório de Dinâmica e Modelagem Molecular, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Ave. Carlos Chagas Filho, 373, CCS-Bloco D sala 30, Cidade Universitária Ilha de Fundão, Rio de Janeiro, CEP 21941-902, Brazil.

[§]Center for Theoretical Biological Physics, Rice University, Houston, Texas, 77005, United States.



Figure S1. Allosteric inhibitors of FP-2 reported in literature. Three scaffolds of allosteric inhibitors of mature FP-2 have been described so far: *i*) suramin, *ii*) heme and *iii*) chalcone derivatives. Inhibitors of scaffolds *i*) and *ii*) trigger a substrate-excess inhibition of the enzyme, which requires the binding of a second substrate molecule that acts as modulator. On the other hand, the chalcone derivative (Cpd66) exerts a classical non-competitive inhibition on FP-2.



Figure S2. FP-2 tryptic peptides displaying differences in their hydrolysis rates in the presence of either a competitive or a non-competitive inhibitor. The tryptic peptide shown in blue is protected from trypsin degradation in the presence of an orthosteric ligand, whereas the one depicted in orange becomes protected when the non-competitive inhibitor Cpd66 binds the enzyme.¹ Trypsin cleavage sites are highlighted by depicting the residues forming the scissible peptide bond as sticks. In addition, the residues of the catalytic diad, C42 and H174, are represented as sticks. The primary sequences of the tryptic peptides are also included in the figure.

Text S1. Umbrella sampling protocol

After careful orientation using gmx editconf of GROMACS v5.1.4, the FP-2:Cpd66 and FP-2:peptide:Cpd66 complexes were embedded in 7x7x10 nm boxes. Of note, identical orientation of both complexes inside the box was guaranteed by structural alignment prior to solvation with TIP3P water molecules. A neutralizing amount of NaCl was added to each, so that the final concentration was 0.1 M. The systems were then subjected to 50000 steps of steepest descents energy minimization, followed by 100 ps of NVT and NPT equilibrations in order to reach a final temperature and pressure of 298 K and 1 bar, respectively. The heating process during the NVT equilibration was conducted using a linear temperature gradient, from 50 to 298 K, the target temperature being achieved at 90 ps. The Berendsen thermostat and barostat were used for temperature and pressure control, respectively, during the equilibrations.² Pulling simulations were

carried out during 25 ns at a pulling rate of 0.192 nm/ns on the z axis. The reference atoms for pulling and US simulations were the carbonyl carbon atom of residue I146 and the alkenyl carbon atom of Cpd66 bonded to the furanyl ring (Fig. S1). Harmonic restraints on the orthogonal axes of the distance vector between the reference atoms were set during all pulling and US simulations. The orthogonal restraint constant was 8000 kJ·mol⁻¹·Å⁻² during pulling simulations, and was lowered to 4000 kJ·mol⁻¹·Å⁻² during US simulations. The equilibrium values for the xy distance components were calculated by averaging the results from five 25 ns replicate MD simulations of the complexes carried out with no restraints, after discarding the first 5 ns of each. In addition, the protein's six external degrees of freedom were restrained to avoid rotation and translation with respect to the center of mass without affecting the conformational equilibrium in both, pulling and US simulations.³ This was achieved by attaching the protein to three dummy atoms, which were kept fixed in the simulation cell by means of position restraints of 21000 kJ·mol⁻¹·Å⁻². The six restraints involving the three dummy atoms (D1, D2 and D3) and three protein atoms (carbonyl carbons of I57, S147 and A88) corresponded to D1-I57 distance, D2-D1-I57 and D1-I57-S147 angles, and D3-D2-D1-I57, D2-D1-I57-S147 and D1-I57-S147-A88 dihedrals. A harmonic constant of 21000 kJ·mol⁻¹·Å⁻² was employed for the D1-I57 distance restraint, and 418.4 kJ·mol⁻ ¹rad⁻², for the angular and dihedral restraints. Frames were collected from the pulling simulations every 0.1 nm intervals along the z axis until reaching a distance of 3 nm between the ligand and the protein reference atoms. Beyond that distance, frames were collected every 0.2 nm. In the 0-1.20 nm interval, a restraint constant of 6000 kJ·mol⁻¹·Å⁻² was set for every umbrella window, whereas a value of 1000 kJ·mol⁻¹·Å⁻² was employed for the remaining ones. The stiffer constant allowed us to keep the ligand reference atom close to the equilibrium position while still partially buried in the binding pocket. The starting frame of each window was equilibrated in the NPT

ensemble during 200 ps at temperature 298 K and pressure 1 bar. The production runs were conducted in the NPT ensemble at 298 K using the velocity rescaling and Parrinello-Rahman algorithms for temperature and pressure control, respectively.^{4,5} Up to 1.20 nm, all windows were simulated during 25 ns during the production runs, the remaining ones being simulated over 15 ns. Data collection started after discarding the first 5 ns. The gmx wham program of GROMACS v5.1.4 was then used to calculate the PMF along the reaction coordinate.⁶ Standard errors of the mean were estimated by conducting 1000 bootstrapping simulations taking into account the autocorrelation time for each window.

SI

In all simulations, a distance cut-off of 1.0 nm was set for electrostatic and van der Waals nonbonded interactions. Particle Mesh Ewald (PME) was used to handle long-range electrostatics. Periodic boundary conditions were employed in all cases. The equation of motion was solved using the leap frog algorithm with a Δt of 2 fs,⁷ while constraining the bond lengths with the Linear Constraint Solver (LINCS).⁸



Figure S3. Occlusion by residue K34 of site 3 internal cavity in FP-2 crystal structure 2OUL. The internal cavity walls are shown in gray. Residue K34, which acts as a lid, and the acidic residues stabilizing its side-chain conformation are depicted as sticks.



Figure S4. Formation of a transient pocket in the site 3 region of FP-3. A) Surface representation of a central structure of FP-3 having an open pocket conformation. The surface region corresponding to the transient pocket entrance is depicted in orange. A protein slice after 90° rotation shows the inner space of the pocket. The pocket internal volumetric density is represented in light blue. B) Detailed structural representation of the open pocket conformation. Residues lying at the pocket entrance are depicted as cyan sticks surrounded by a transparent orange surface. Residues forming the pocket internal wall are shown as green sticks. The internal surface of the pocket is colored in gray. Hydrogen bonds occurring at the pocket entrance are indicated with yellow dashed lines. C) Graphs showing the E54(CD)-K36(NZ) distance and site 3 internal volume time profiles during the concatenated MD simulations of FP-3. D) Volume versus E54(CD)-K36(NZ) distance graph. Dots depicted in orange correspond to frames bearing an open pocket. For FP-3, the conditions established to define an open cavity were 10 Å >E54(CD)-K36(NZ) distance ≥ 80 Å³, or E54(CD)-K36(NZ) distance ≥ 10 Å and a pocket inner volume ≥ 70 Å³.



Figure S5. Best docking pose of Cpd66 into site 3. A) Two views of the best docking pose showing the electrostatic surface of FP-2 calculated with the APBS plugin of pymol 2.1.0. The lower image corresponds to a protein slice after performing a 90° rotation with respect to the upper representation. The surface is colored according to the gradient of electrostatic potential expressed in k_bT/e units, where k_b , *T* and *e* stand for the Boltzmann's constant, the temperature (298.15 K) and the electron charge, respectively. **B**) Detailed views of the complex interface. Residues interacting with Cpd66 are shown as sticks. Yellow dashed lines indicate hydrogen bonds occurring at the interface. The inhibitor is depicted as yellow sticks in all cases.



Figure S6. Time evolution of the FP-2:Cpd66 complex determined by docking during four independent MD simulations. A), B), C) and D) represent independent MD simulations started from the same structure (t=0) but using different randomly-generated velocities. The electrostatic surface of the protein was colored according to the gradient shown in Fig. S5. Two views of last frame obtained from each simulation are depicted and labeled accordingly. On top of the first view, the ΔG_{eff} mean value corresponding to the last 50 ns of each trajectory is provided. Graphs displaying the time profiles of RMSD values calculated for the Cpd66 heavy atoms after fitting each trajectory to the corresponding starting frames and with respect to the protein's backbone atoms, are shown in all cases. Simulations were halted at different times, when sufficient data were obtained to definitely propose the most stable binding mode.



Figure S7. Comparison of the docking and MD-generated conformations of Cpd66 in complex with **FP-2.** The FP-2:Cpd66 complex used as starting structure for the MD simulations is represented as green cartoon (protein) and yellow sticks (ligand). The complex structure obtained from the MD simulation leading to the most stable binding mode is shown as cyan cartoon (protein) and salmon sticks (ligand). Note that the hydrogen bond (yellow dashed line) mediated by T31 and E52 breaks when Cpd66 adopts the final conformation.



Figure S8. Contact between residue K203 and the six-membered aryl ring of Cpd66 in the FP-2:Cpd66 complex. Residues K203 and N204, whose peptide bond is susceptible to trypsin cleavage, and Cpd66 are shown in surface representation. The distribution of distances between the heavy atoms of K203 and those of the six-membered aryl ring of Cpd66 is shown on the right. The distances between centers-of-mass of the previously-mentioned atom groups were calculated during the 1.2 µs concatenated MD simulations of the FP-2:Cpd66 complex.



Figure S9. Exchange of water molecules mediating the hydrogen bond between Cpd66 and G29. A) Frames collected from an MD simulation of the FP-2:Cpd66 complex showing the water exchange during an interval of 620 ps are sequentially represented. Note that t=0 does not indicate the first frame of the trajectory, but the one taken as a reference. The water molecules are colored differently to allow their identification throughout the frames. B) Same as A) but showing the protein surface (gray) created by residues R25, G29, T31, and F215. A hole on this surface is transiently created, thus allowing the water exchange.



Figure S10. RMSD time profiles for the Cpd66 and peptide during the replicate MD simulations of the studied complexes. RMSD values calculated during the MD simulations of **A**) the FP-2:Cpd66 complex, **B**) the FP-2:peptide complex and **C**) the FP-2:peptide:Cpd66 complex. Time profiles of RMSD values were calculated for the Cpd66 and the peptide heavy atoms after fitting each trajectory to the corresponding starting frames and with respect to the protein's backbone atoms.

Text S2. Compound 66 binding induces subtle perturbations in the FP-2 conformational space The allosteric inhibition exerted by Cpd66 on FP-2 must lead to some changes in the protein, ultimately responsible for slowing down the enzymatic activity. We started the characterization of such changes by means of coarse measurements, such as backbone RMSD and RMSF calculations throughout the MD simulations (Fig. S11). Interestingly, the Cpd66 binding causes a shift in the backbone RMSD histograms calculated with respect to the FP-2 crystal structure (PDB: 2OUL) toward the right, when compared to the systems lacking the inhibitor (Figs. S11A and S11B). Therefore, when Cpd66 is bound, the enzyme occurs in conformations that are, on average, more divergent with respect to the crystal structure than in the absence of this ligand. However, this effect is largely caused by the reorganization of the loops surrounding Cpd66 and do not reflect necessarily distal changes. In addition, the calculation of RMSD histogram widths, i.e., the standard deviations, reveals that the global diversity of FP-2 conformations slightly narrows in the presence of Cpd66 (Figs. S11A and S11B). In other words, the protein conformations sampled during the MD simulations of the FP-2:Cpd66 and FP-2:peptide:Cpd66 complexes tend to be less diverse in terms of backbone RMSD values if compared to those generated for free FP-2 and the FP-2:peptide complex.



Figure S11. Comparison of RMSD distributions and per-residue RMSF values for the studied systems. A) and **B)** backbone RMSD distributions for the indicated systems with respect to the FP-2 crystal structure 2OUL. The standard deviations of the RMSD distributions for the FP-2, FP-2:Cpd66, FP-2:peptide and FP-2:peptide:Cpd66 systems are 0.214, 0.189, 0.227 and 0.187 Å, respectively; and the mean values, 1.40, 1.74, 1.40 and 1.69 Å, respectively. **C)** and **D)** Per-residue RMSF values calculate for the protein's backbone with respect to the average position of each residue during the concatenated 1.2 μs MD simulations conducted for each system. The arrows indicate regions that appreciably changed their flexibility in the compared systems.

The analysis of per-residue backbone RMSF values allowed us to identify protein regions that changed their local flexibility in the presence of the inhibitor (see arrows, Figs S11C and S11D). The binding of Cpd66 to FP-2 promotes the most appreciable RMSF variations between the apo and holo systems in the following residues: 33-39, 58-63, 106-109, 111-114, 117-120, 166-168 and 206-210 (Fig. S11C), all of them belonging to loops. Remarkably, the flexibility change in the

latter loop, which contains residues W206 and W210, might also explain, to some extent, the variation in Trp intrinsic fluorescence emission between free FP-2 and the FP-2:Cpd66 complex,¹ as microenvironment modifications are expected to occur due to an increase in loop motion upon ligand binding. Residue W24, besides interacting directly with Cpd66, contacts some residues belonging to the segments 33-39 and 58-63, whose RMSF profiles varied because of the ligand binding (Fig. S11C). This reinforces our conclusion that changes in the microenvironment of W24 are likely to be critical to explain the shift of Trp emission maximum observed in the fluorescence experiments.¹

For the FP-2:peptide and FP-2:peptide:Cpd66 complexes, the main RMSF variations are found in the positions: 10-21, 27-36, 62-70, 100-107, 154-157, 205-211 and 233-235, all lying in loop regions as well (Fig. S11D). Some of the previously-mentioned loop segments either contain residues directly interacting with the inhibitor in both FP-2:Cpd66 and FP-2:peptide:Cpd66 complexes, e.g., those involving positions from 27 to 39, or lie close to the interface, e.g., those comprising the 58-70 and 205-211 positions. However, the remaining loops are located farther from the allosteric pocket, thus indicating that the perturbations triggered by the inhibitor can propagate toward distal regions of the protein. Moreover, the binding of Cpd66 tends to increase the fluctuations of various loops (Figs. S11C and S11D), especially those in vicinity of the compound. Of note, the previous result is not in contradiction with the previous RMSD analysis, which showed a decreased conformational diversity in the presence of the inhibitor. The RMSF measures the average fluctuation of each atom (or set of atoms) during the MD simulation with respect to a reference position, whereas the RMSD quantifies the overall divergence of each conformation of the ensemble with respect to the reference structure. Therefore, flexible but small regions do not necessarily lead to large RMSD variations. Conversely, less flexible but larger

regions will have higher weights in the calculations and, consequently, their contribution to the final RMSD value is likely to be more significant.

In addition, we performed a PCA to compare the motions of the studied systems along the main PCs.⁹ As shown in Fig. S12A, the eigenvalues associated with the PCs of free FP-2 rapidly decrease up to the fourth PC (Fig. S12A). This indicates that the first three PCs are associated with the largest motions of the protein and account for nearly 42% of the cumulative fluctuation (Fig. S12A). Of note, PC1, PC2 and PC3 are associated with the collective motions of various enzyme loops that also display large RMSF values (Figs. S11C, S12B and S12C). The 2D projections of the FP-2 and FP-2:Cpd66 complex trajectories onto the PC1, PC2 and PC2, PC3 subsets of the free enzyme show a significant overlap; thereby indicating that protein motions along the selected PCs are rather similar in both systems. However, FP-2 samples, in general, motions of larger amplitude in the absence of the inhibitor, especially along PC3 (Fig. S12D). Similar conclusions can be drawn from the 2D projections of the FP-2:peptide and FP-2:peptide:Cpd66 trajectories onto the same PC subsets employed before (Fig. S12D). Therefore, the Cpd66 binding tends to slightly narrow the concerted movements of the protein along the first three PCs regardless of the presence of the peptide in the active site. Nonetheless, the peptide seems to counter this effect to some extent, as greater overlaps of the 2D projections are observed for the FP-2:peptide and FP-2:peptide:Cpd66 complexes (Fig. S12D).



Figure S12. PCA results for the studied systems. A) Eigenvalues of the first twenty eigenvectors corresponding to the 1.2 μ s concatenated MD simulations of free FP-2. The cumulative fluctuations of the first five eigenvectors are shown beside their corresponding points in the graph. B) Per-residue components of PC1, PC2 and PC3 of free FP-2. Peaks displaying the largest vector components are labeled using Roman numbers. C) Structural representation of the collective motions in free FP-2 along PC1, PC2 and PC3. Regions corresponding to the peaks shown in B) are labeled accordingly. D) 2D projections of the protein motions along the PC1, PC2 and PC2, PC3 subsets of free FP-2 PCs for the indicated systems. The PCA was conducted for the Ca atoms of the protein.



Figure S13. Generalized correlations for the studied systems. *GC* matirices for **A**) free FP-2 (upper triangle) and the FP-2:Cpd66 complex (lower triangle), and **B**) the FP-2:peptide (upper triangle) and the FP-2:peptide:Cpd66 complexes (lower triangle). RMSD of *GC* matrices for **C**) free FP-2 (upper triangle) and the FP-2:Cpd66 complex (lower triangle), and **D**) for the FP-2:peptide (upper triangle) and the FP-2:peptide:Cpd66 complexes (lower triangle), and **D**) for the FP-2:peptide (upper triangle) and the FP-2:peptide:Cpd66 complexes (lower triangle). *GC*_{*ij*} RMSD values were calculated from the four replicate MD simulation of each system.



Figure S14. Distribution of raw and filtered generalized correlations for the four studied systems. A) Free FP-2, B) FP-2:Cpd66 complex, C) FP-2:peptide complex and D) FP-2:peptide:Cpd66 complex. The bars in blue correspond to distributions of raw $GC \ge 0.5$. The bars colored in light purple correspond to distributions of $GC \ge 0.5$ in at least one replicate MD simulation of each system and filtered using residueresidue contact maps, which zero all correlations of distant residues. The overlap between both distributions is highlighted in dark purple in each graph. Two residues are considered to be in contact if their heavy atoms are closer than 5.0 Å during, at least. 75% of the simulation time. Note that for $GC \ge 0.6$, both distributions largely match, thus indicating that long-range correlations have been greatly ruled out when using this higher GC cut-off.

Community	ED 2	ED 2.Cmd((ED 2montido	ED 2montido. Cud66
No.	FP-2	FP-2:Cp000	FP-2:pepude	FF-2:pepude:Cpuoo
1	1:15, 126:144,	1:15.	1:16.	1:16.
	236:241.			
2	16:33, 179:202,	16:33, 145,	17:35, 156:164,	17:23, 179:202,
	208:226.	178:204, 215:226.	179:203, 205:226.	217:226.
3	34:35, 145:150,	34, 63:74, 96:124.	36:63 , 238,	24:33.
	174:178 , 203:207 ^a ,		240:241.	
	231:232, 234:235.			
4	36:41 , 76:96.	35, 153:167,	64:74, 97:126.	34, 63:73, 95, 97:123.
		205:214.		
5	42:63 .	36:41 , 75:95.	75:96.	35, 154:164, 203:216 .
6	64:75, 97:125.	42:62.	127:130, 145:155,	36:47 , 74:94, 96.
			165:178, 204,	
			227:237, 239.	
7	151:173, 227:230,	125:127, 238:241.	131:144.	48:62, 124:126, 238,
	233.			240:241.
8	-	128:130, 146:152,	-	127:144, 237, 239.
		168:177, 227:237.		
9	-	131:144.	-	145:153, 165:178 ,
				227:236.

Table S1. Residue composition of each community of the studied systems using a 5.0 Å distance cutoff for residue-residue contact definition

^aResidue segments containing key residues for catalysis, Q36, C42, H174 and N204, are highlighted in bold.



Figure S15. Dependency of modularity values on the number of communities for the analyzed systems. A) Free FP-2, **B**) FP-2:Cpd66 complex, **C**) FP-2:peptide complex and **D**) FP-2:peptide:Cpd66 complex. The optimal number of communities for each system is the smallest one yielding a nearly maximal modularity (a 0.01 tolerance from the modularity absolute maximum was chosen in each case). The calculated numbers of communities are indicated between parentheses in each graph.



Figure S16. Community analysis for the studied systems using a distance cut-off of 4.5 Å to define residue-residue contacts. 2D diagrams showing the organization of communities for each system are depicted on the left. The communities are represented by circles whose diameters are proportional to the number of residues contained within each. The lines account for intercommunity correlations, which are proportional to the line width. For convenience, communities are colored and numbered differently. The labels of residues Q36, C42, H174 and N204, are shown beside their corresponding communities. On the right, the community organization is represented as a diagram mimicking the 3D-structure of FP-2, with nodes colored according to their community membership. The dots and gray lines stand for the network nodes (C α atoms) and their pairwise *GCs*, respectively.

Community	ED 2	ED 2.Cnd66	ED Junantida	ED Imontidor Condee
No.	FP-2	FF-2:Cp000	rr-2:pepude	гг-2:pepuae:0раоо
1	1:16, 126:127,	1:16.	1:16.	1:16.
	238:241.			
2	17:35, 145,	17:25, 157:164,	17:28, 156:164,	17:23, 178:202,
	174:226 ^a .	179:204, 206:228.	179:203, 208:226.	217:226.
3	36:41 , 76:96.	26:33, 35:36 , 205.	29:35, 205:207.	24:33.
4	42:63 .	34, 52, 63:74,	36:39 , 41:63 , 127,	34, 64:74, 97:123.
		97:124.	238:241.	
5	64:75, 97:125.	37:40, 75:96	40, 75:96	35, 154:164, 203:216
6	128:144.	41:51 .	64:74, 97:126.	36:47 , 75:96.
7	146:149, 232:237.	53:62.	128:130, 146:155,	48:63, 124:126, 238,
			165:178, 204,	240:241.
			227:237.	
8	150:173, 227:231.	125:129, 237:241.	131:145.	127:130, 235:237,
				239.
9	-	130, 145:156,	-	131:144.
		165:178, 229:236.		
10	-	131:144.	-	145:153, 165:177 ,
				227:234.

Table S2. Residue composition of each community of the studied systems using a 4.5 Å distance cutoff for residue-residue contact definition

^aResidue segments containing key residues for catalysis, Q36, C42, H174 and N204, are highlighted in bold.

Text S3: Betweenness centrality analysis

Our results show that the centrality profiles of the analyzed systems are perturbed by the presence of the inhibitor (Fig. S17). For free FP-2 and the FP-2:Cpd66 complex, the residues undergoing the largest variations are: M2, K37, S41, S47, I48, I93, P105, S108, Y123, I125, K126, I148, A175 and F236, which decreased their centralities upon the inhibitor binding; and Q1, G29, V33, K34, G40, F45, S46, S50, E67, E69, L140, F142, K203 and W206, which possess higher values in the FP-2:Cpd66 complex (Fig. S17A). Some of the previous residues, e.g., G29, V33, K34, F45, S46, E67 and K203, lie at the allosteric pocket (Fig. 2B) and, interestingly, all belong to the group of residues whose centralities augmented appreciably when the inhibitor is bound. On the other hand, most of the residues displaying large centrality variations are not in direct contact with Cpd66. In

fact, the perturbations caused by the inhibitor are not limited to the allosteric pocket vicinity, but are spread throughout the protein structure (compare the distribution of per-residue centralities in free FP-2 and the FP-2:Cpd66 complex, Fig. S17B). Remarkably, some of the residues undergoing large centrality variations are located within the active site, e.g., G40, S41, A175 and W206, thus providing further clues about the impact of Cpd66 on this region of the protein.

The comparison of per-residue centralities for the FP-2:peptide and FP-2:peptide:Cpd66 complexes leads to conclusions that are roughly similar to those obtained in the absence of the peptide. However, it is worth saying that the centrality differences between the peptide-bound complexes become more apparent (compare both graphs in Fig. S17A). The following residues decreased significantly their centralities upon Cpd66 binding to the FP-2:peptide complex: C39, G40, S41, C42, S50, C80, G97, D109, S108, Y123, G124, I148, A175, L178, F236, I237 and P238. On the contrary, other residues, comprising W24, G29, V33, K34, K37, F45, S46, E67, E69, G83 and N204, notably increased their centralities in the presence of the inhibitor (Fig. S17A). Many of these residues display a similar behavior in FP-2 and the FP-2:Cpd66 complex.



Figure S17. Betweenness centralities for the FP-2 residues in four different states. A) Graphs show the per-residue centrality values for the indicated systems. Residues displaying large centrality variations between the compared systems are labeled. B) Distribution of per-residue centralities on the FP-2 structure in the different systems (see labels). The C α atoms of FP-2 are depicted as spheres and colored according to their normalized centralities (see color gradient bars). Cpd66 (not shown) binds along the FP-2 central α -helix.



Figure S18. Titration curves that were reliably estimated in only one system. In every graph, the fraction of deprotonation (f_d) was plotted versus the pH. Calculated pK_a 's and Hill's cooperativity indices (n) are indicated. 95% confidence intervals for the estimated values were calculated by error bootstrapping. Residues marked with asterisks were not accurately titrated.



Figure S19. Titration curves displaying the largest shifts upon Cpd66 binding to FP-2. In every graph, the fraction of deprotonation (f_d) was plotted versus the pH. Calculated pK_a 's and cooperativity indices (n) are indicated. 95% confidence intervals for the estimated values were calculated by error bootstrapping.



Figure S20. Residues displaying significant pK_a shifts upon the binding of Cpd66 to FP-2. The residues undergoing such variations are depicted as cyan sticks. The inhibitor is shown in yellow and the protein backbone structure, as white cartoon.



Figure S21. RMSD values for the peptide at different pHs. The RMSDs were calculated with respect to the peptide starting conformation in the FP-2:peptide complex at pH=5.5 after fitting all frames to FP-2 initial structure. As the titration was performed by changing the pH in discrete steps of 0.1, the graph was built by assigning evenly distributed pH values between *i* and *i*+0.1 to the frames sequentially collected at pH *i*. At pH<4.8 (see dashed line) the peptide loses affinity for the enzyme's active site.

References

- Bertoldo, J. B.; Chiaradia-Delatorre, L. D.; Mascarello, A.; Leal, P. C.; Cordeiro, M. N.; Nunes, R. J.; Sarduy, E. S.; Rosenthal, P. J.; Terenzi, H. Synthetic compounds from an in house library as inhibitors of falcipain-2 from *Plasmodium falciparum*. J. Enzyme Inhib. Med. Chem. 2015, 30, 299-307.
- (2) Berendsen, H. J.; Postma, J. v.; van Gunsteren, W. F.; DiNola, A.; Haak, J. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.
- (3) Heinzelmann, G.; Henriksen, N. M.; Gilson, M. K. Attach-Pull-Release calculations of ligand binding and conformational changes on the first BRD4 bromodomain. J. Chem. Theory Comput. 2017, 13, 3260-3275.
- (4) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- (5) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182-7190.
- (6) 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. J. Chem. Theory Comput. 2010, 6, 3713-3720.
- (7) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 2015, *1*, 19-25.
- (8) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463-1472.
- (9) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential dynamics of proteins *Proteins* 1993, 17, 412-425.