Supplementary Information (SI)

Modulation of conformational landscape of PDZ3 domain by perturbation on distal non-canonical α3 helix: Decoding microscopic mechanism of allostery in PDZ3 domain

Subhajit Sarkar[1], Saikat Dhibar, Biman Jana[1*]

[1]School of Chemical Science, Indian Association for the Cultivation of Science, Jadavpur, Kolkata-700032

Corresponding Email: pcbj@iacs.res.in.

This pdf file includes:

Supporting Information:

Figures

1. Figure S1: The schematic picture of calculating the binding free energy of different wild and mutated PDZ3-CRIPT ligand systems using the thermodynamic integration method.
2. Figure S2: Two dimensional distance plot obtained from each 2μs simulation trajectory of (a) 1BFE E395A (b) 1BFE Y397E (c) 1BFE R399A (d) 1BFE E401A.
3. Figure S3: Principal Component Analysis based on Cα RMSF data obtained from each 2μs simulation trajectory of (a)1BFE E395A , (b) 1BFE Y397E (c) 1BFE R399A ,(d)1BFE E401A.
4. Figure S4: Representative structures of (a) Closed state (b) Open state (c) Extended open state (d) Wide open state.
5. Figure S5: Cα covariance matrix plot obtained from each 2μs trajectory of (a) 1BFE E395A (b) 1BFE Y397E (c) 1BFE R399A (d) 1BFE E401A.
6. Figure S6: The binding mode of the ligand in the different cavities highlighting the relevant interactions in the (a) Closed state (b) Crystal open state (c) Extended open state (d) Wide open state is shown.
7. Figure S7: Reweighted free energy landscape and Implied timescale plot for building the Markov State Model.

Table

8. Table S1: Comparison of Closed state crystal structure only PDZ3-CRIPT interaction vs Wide-open state only PDZ3-CRIPT interaction.
9. Table S2: Comparison of Closed state crystal structure only PDZ3-CRIPT interaction vs Extended-open state only PDZ3-CRIPT interaction.
10. Table S3: Comparison of Closed state crystal structure only PDZ3-CRIPT interaction vs Crystal-open state only PDZ3-CRIPT interaction.

Method Details

Mutational Frustration calculation method

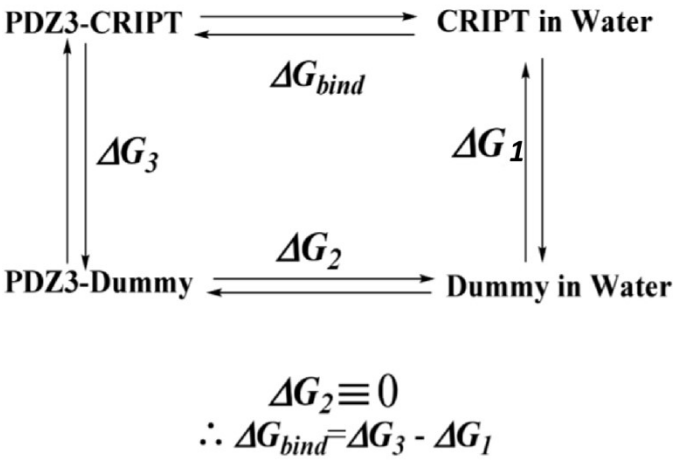


Figure S1: Thermodynamic cycle to calculate relative binding free energy, $\mathbf{\Delta G_{bind}}$ is the relative binding free energy of the protein-ligand (PDZ3-CRIPT) system.

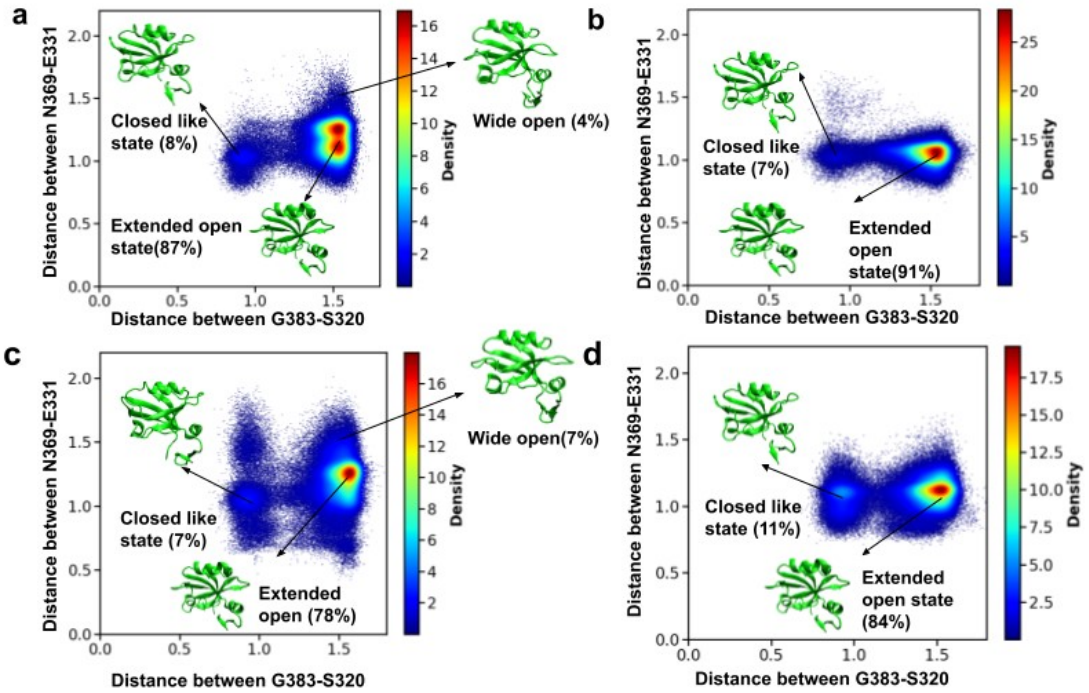


Figure S2: Two dimensional distance plot obtained from each 2µs simulation trajectory of (a)1BFE E395A , (b)1BFE Y397E (c) 1BFE R399A ,(d)1BFE E401A .The center of geometry distance (nm) between residue G383-S320 and residue N369-E331 is represented by X and Y axis respectively. Representative structures and population of each cluster is mentioned alongside.

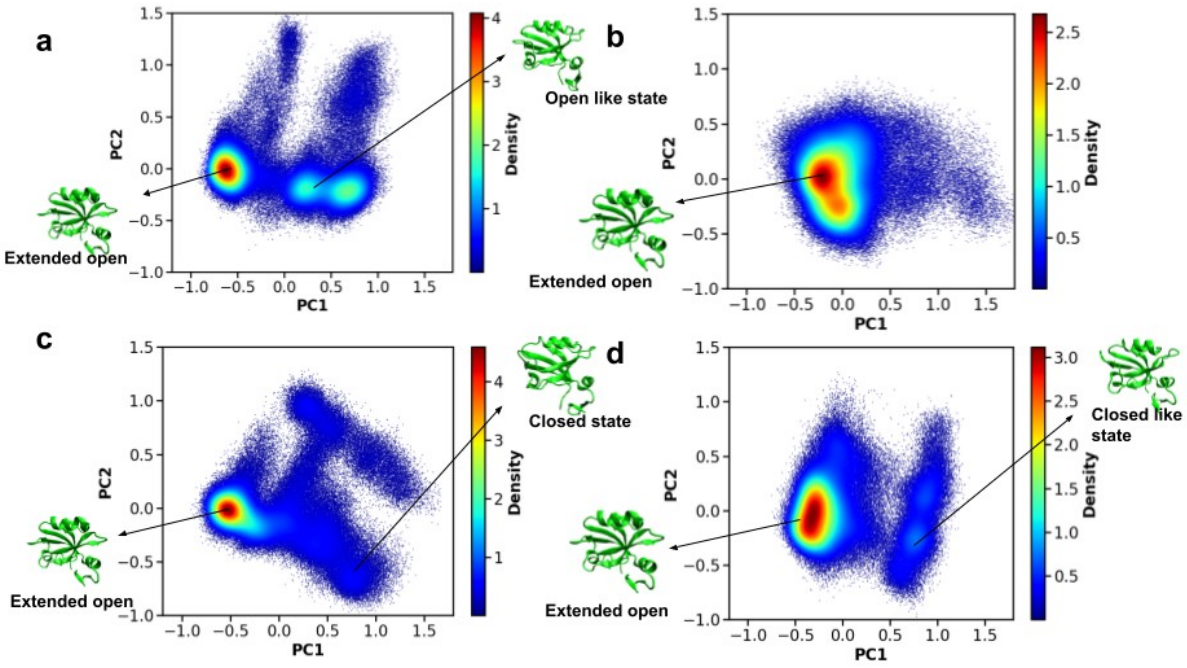


Figure S3: Principal Component Analysis based on Cα RMSF data obtained from each 2μs simulation trajectory of (a)1BFE E395A , (b)1BFE Y397E (c) 1BFE R399A ,(d)1BFE E401A .The representative structures of the major populated states are mentioned alongside.

Main differences between different metastable states based on the center of geometry distance G383-S320 and N369-E331

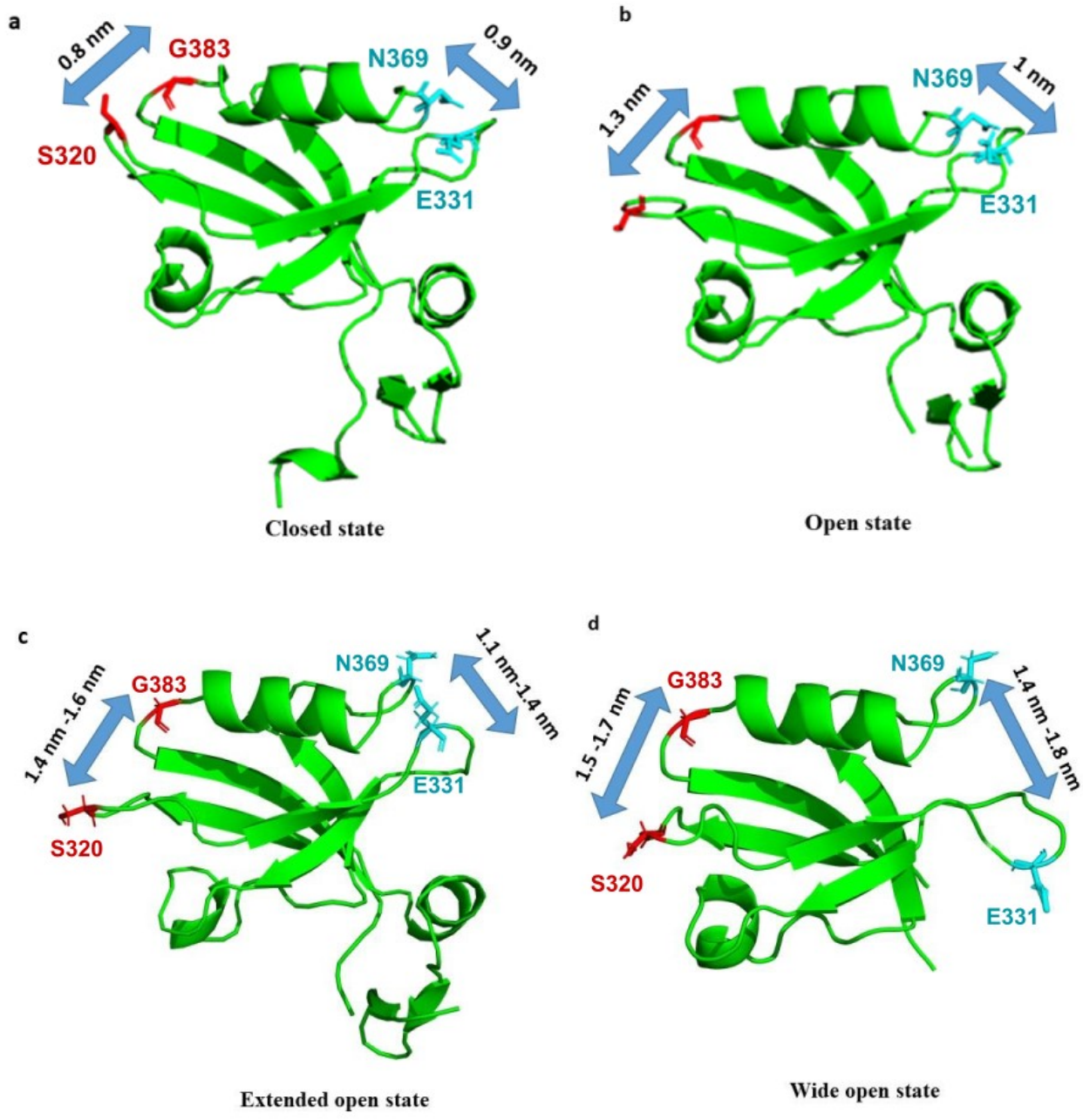


Figure S4**:** Representative structures of (a) Closed state (b) Open state (c) Extended open state (d) Wide open state. The center of geometry distance (nm) between residue G383-S320 and residue N369-E331 is mentioned alongside the representative figures. The

closed state (a) present in the wild PDZ3 is completely absent in the α3 truncated variation. The wide open state (d) present in the α3 truncated variation is absent in the wild PDZ3. The presence of multiple structural ensembles in the loop position based landscape of PDZ3 domain supports the conformational ensemble view of allostery.

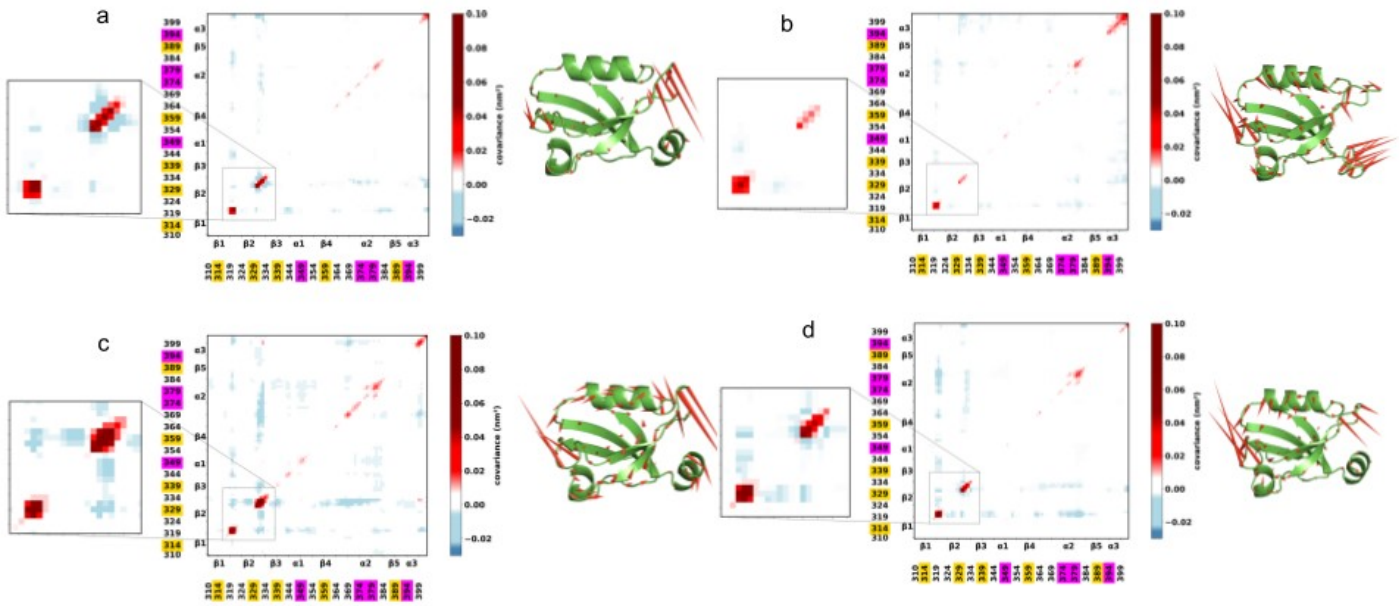
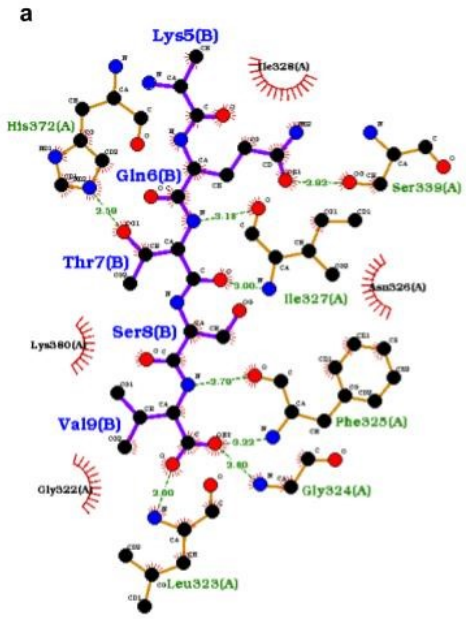
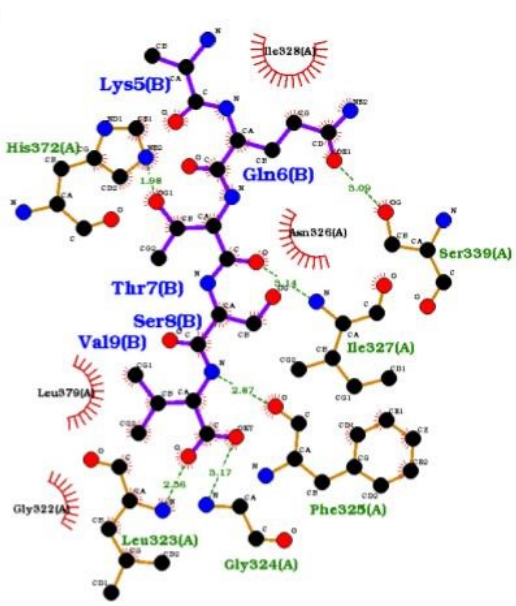


Figure S5: Correlated motion of β1-β2 and β2-β3 loops were lost due to mutation in the distant α3 helix. a,b,c.d represents the Cα covariance matrix plot obtained from each 2μs trajectory of 1BFE E395A ,1BFE Y397E,1BFE R399A,1BFE E401A mutated systems corroborated with porcupine projection plots respectively.



Closed state



Crystal open state

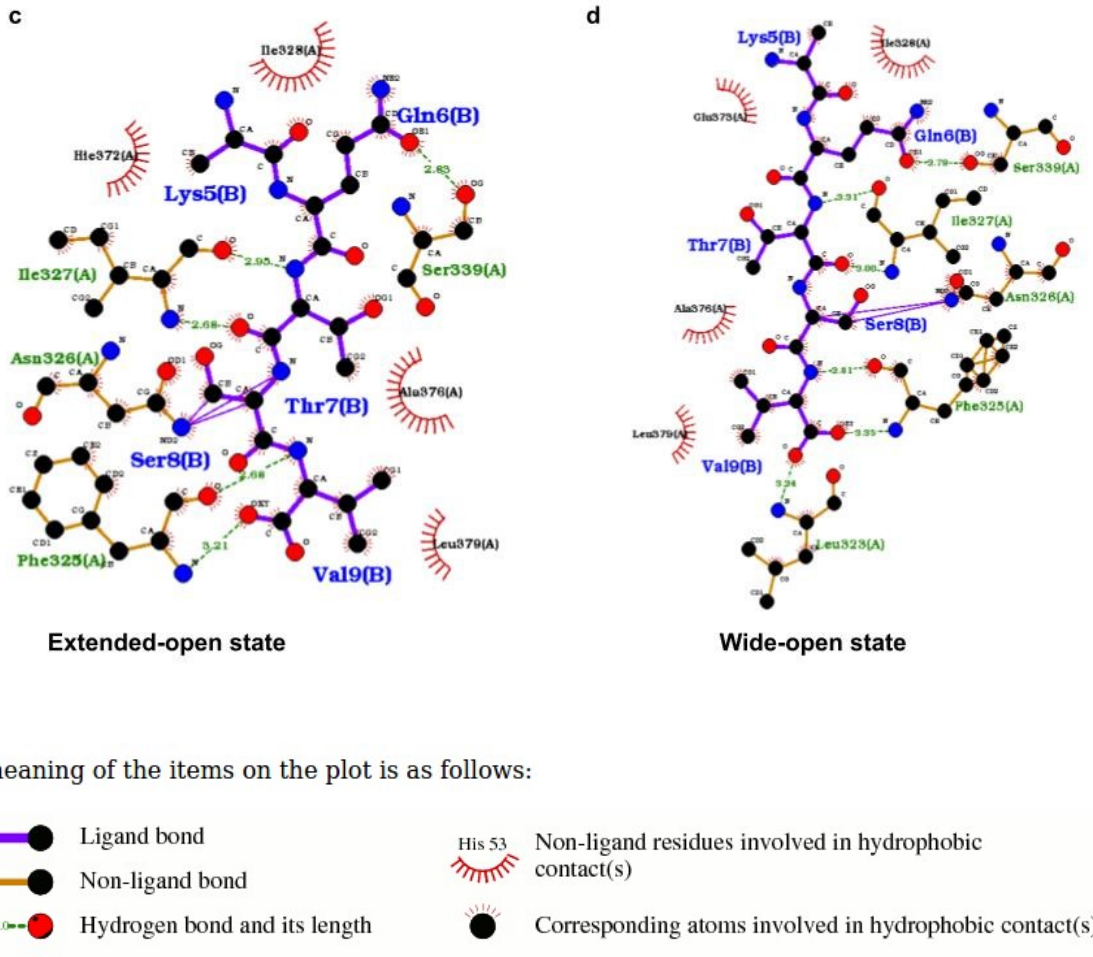


Figure S6: The binding mode of the ligand in the different cavities highlighting the relevant interactions in the (a) Closed state (b) Crystal open state (c) Extended open state (d) Wide open state is shown.

Table S1: Comparison of Closed state crystal structure only PDZ3-CRIPT interaction vs Wide-open state only PDZ3-CRIPT interaction

Closed state crystal structure only PDZ3-CRIPT interaction	**Wide-open state only PDZ3-CRIPT interaction**
Val9(B)–Gly324(A)(Hydrogen Bonding)	**Lys5 (B)- Glu373 (A) (Hydrophobic)**
Thr(B)–His372 (A) (Hydrogen Bonding)	**Val9 (B)-Ala376 (A) (Hydrophobic)**
Ser8 (B) – Lys380 (A) (Hydrophobic)	**Val9 (B)-Leu379 (A) (Hydrophobic)**
Val9 (B) – Gly322 (A) (Hydrophobic)	

Table S2: Comparison of Closed state crystal structure only PDZ3-CRIPT interaction vs Extended-open state only PDZ3-CRIPT interaction

Closed state crystal structure only PDZ3-CRIPT interaction	**Extended-open state only PDZ3-CRIPT interaction**
His372(A)-Thr7(B) (Hydrogen bonding)	**His372(A)-Thr7(B) (Hydrophobic)**
Lys380(A)-Ser8(B) (Hydrophobic)	**Leu379(A)-Val9(B) (Hydrophobic)**
Gly322(A)-Val9 (B) (Hydrophobic)	**Ala376(A)-Thr7(B) (Hydrophobic)**
Leu323(A)-Val9(B) (Hydrogen bonding)	
Gly324(A)-Val9(B) (Hydrogen bonding)	

Table S3: Comparison of Closed state crystal structure only PDZ3-CRIPT interaction vs Crystal-open state only PDZ3-CRIPT interaction

Closed state crystal structure only PDZ3-CRIPT interaction	Crystal-open state only PDZ3-CRIPT interaction
Lys380(A)-Ser8(B) (Hydrophobic)	Leu379(A)-Val9(B) (Hydrophobic)

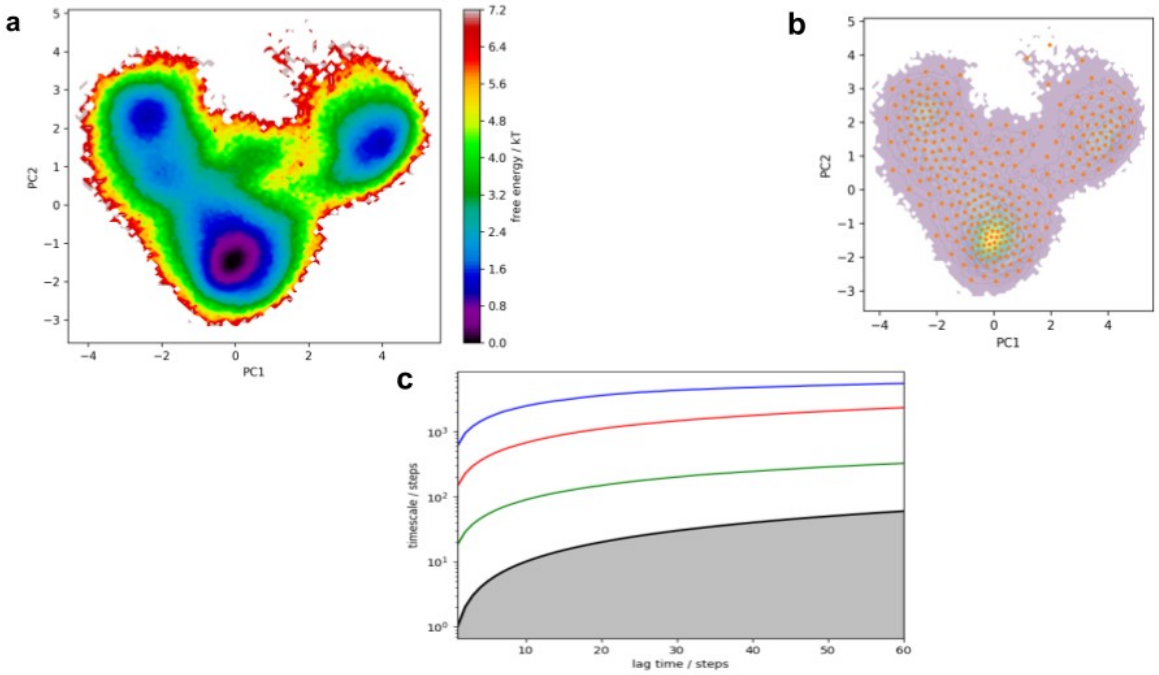


Figure S7: (a) Reweighted free energy landscape which was used to build the Markov State Model. (b) 300 Microstates generated from K-means clustering. (c) Implied timescale plots to choose the specific lag time. The implied time scale plot reveals timescale reaches convergence after 30 steps.

Mutational Frustration calculation method

Mutational frustration quantifies how favorable native amino acids are compared to other potential amino acids at the same position. To assess how frustrated the interactions are in a given structure, the protein sequence is systematically perturbed and the resulting total energy change is calculated. The amino acids in a specific contact pair are mutated with other amino acids to generate a set of decoys, for which the protein's total energy is recalculated. Based on the native amino acid frequency distribution of the particular protein under consideration, sequence space is randomly sampled, generating one thousand appropriately distributed decoys for each contact. A histogram of the energy of the decoys is compared to the distribution to the native energy, H^N. The "frustration index" (Fi,j) for the contact between the amino acids i,j is defined as a Z-score of the energy of the native pair compared to the N decoys.

$$F_{ij} = (\mathcal{H}_{ij}^{N} - < \mathcal{H}_{i',j'}^{U} >) / \sqrt{1/N \sum_{k=1}^{n} (\mathcal{H}_{i',j'}^{U} - < \mathcal{H}_{i',j'}^{U} >)^2}$$

Where $H^{U}_{i',j'}$ is the energy of the decoy and $H^{N}_{i,j}$ is the native state energy. The frustration index measures of how favorable a particular interaction is relative to the set of all possible interactions in that location, normalized using the variance of that distribution.

A contact is classified as 'minimally frustrated' when its native energy falls at the lower end of the decoy energy distribution, characterized by a frustration index (Z-score) of 0.78 or above. This suggests that the majority of other amino acid pairs at that location are less favorable. In contrast, a contact is considered 'highly frustrated' if its native energy is positioned at the upper end of the distribution, with a frustration index below -1. This indicates that most alternative amino acid pairs at that site are more conducive to folding by more than one standard deviation. Contacts with native energies that lie between these extremes are labeled as 'neutral'.