Supporting Information

Salt-Bridged Mediated Conformational Dynamics in the

Figure-of-Eight Knotted Ketol Acid Reductoisomerase (KARI)

Sanjib Thakuria[a] and Sandip Paul[a,*]

[a]*Department of Chemistry, Indian Institute of Chemistry, Guwahati, Assam, India-781039*

E-mail: sandipp@iitg.ac.in

1 Equations

The initial step in Principal Component Analysis (PCA) is the generation of the covariance matrix, C, from the trajectories of our Cartesian coordinate space using the C_α atoms of the protein. Subsequently, a least square fit superimposition (rms fit) onto the reference structure is executed to remove the rotational and vibrational modes. Finally, we attain our PCs by diagonalizing this covariance matrix.[1,2]

$$C = R\lambda R^{\mathrm{T}} \tag{S1}$$

where λ and R contain the eigenvalues and the corresponding eigenvectors of matrix C repsectively.

The following equation is used for the calculation of Dynamic Cross-Correlation Map (DCCM):[3,4]

$$C_{\mathrm{ij}} = \frac{< \Delta r_{\mathrm{i}}.\Delta r_{\mathrm{j}} >}{\sqrt{< \Delta r_{\mathrm{i}}.\Delta r_{\mathrm{j}} >< \Delta r_{\mathrm{i}}.\Delta r_{\mathrm{j}} >}} (i,j = 1,2,3,...,3N) \tag{S2}$$

where Δr_{i} and Δr_{j} is the mean displacement of the ith and j^{th} residue respectively, while N denotes the total number of $C\alpha$ atoms. C_{ij} ranges from -1 to +1.

The harmonic potential V_{i} at window i is expressed as

$$V_{\mathrm{i}} = 1/2k(x - r_{\mathrm{i}})^2 \tag{S3}$$

where k is the force constant with r_{i} being the reference position. Weighted Histogram Analysis Method (WHAM)[5,6] was utilised for the calculation of the unbaised probability distribtion in x referred as P(x). The final PMF was obtained as

$$PMF = -K_{\mathrm{B}}TlnP(x) \tag{S4}$$
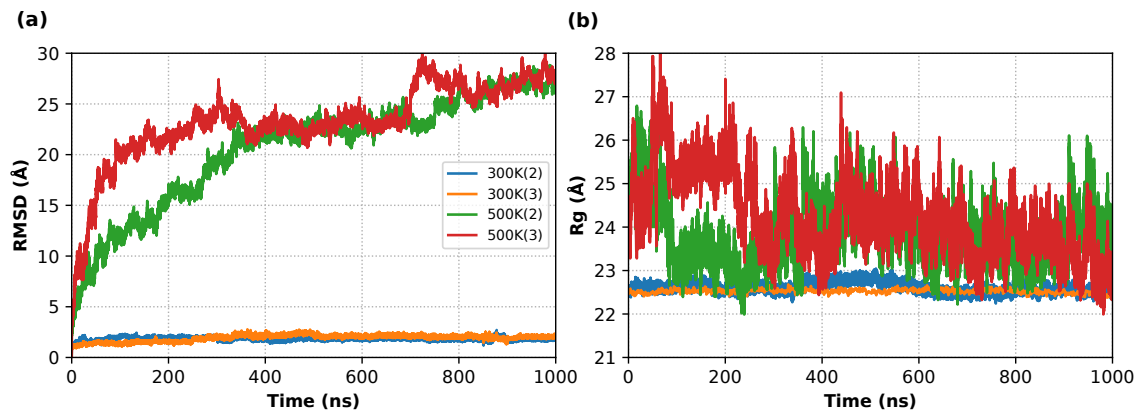
2 Denaturation of the protein

Figure S1: Time evolution of (a) Root Mean Square Deviation (RMSD) and (b) Radius of Gyration (R_g) of the backbone atoms of 3fr8 at temperatures of 300K and 500K for trajectories 2 and 3 respectively.
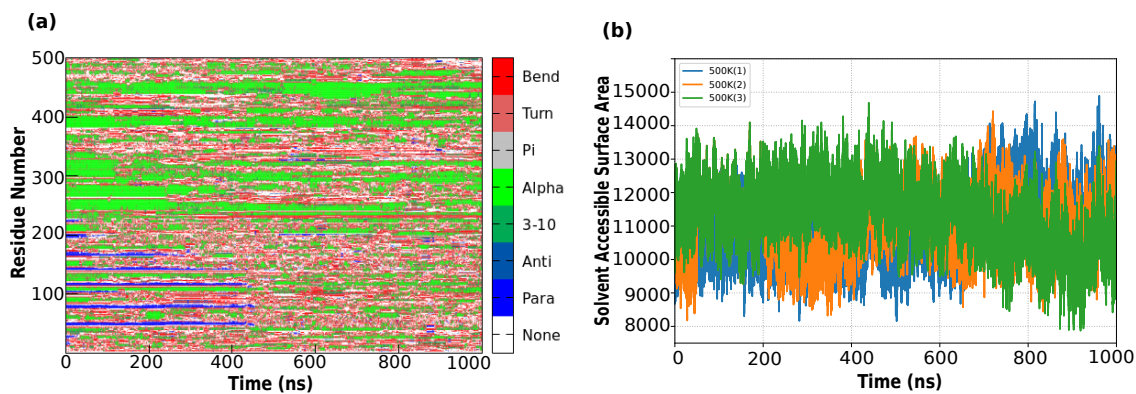

Figure S2: Time evolution of (a) secondary structure of 3fr8 using DSSP algorithm at 500K (b) Solvent Accessible Surface Area (SASA) of the knotted domain.
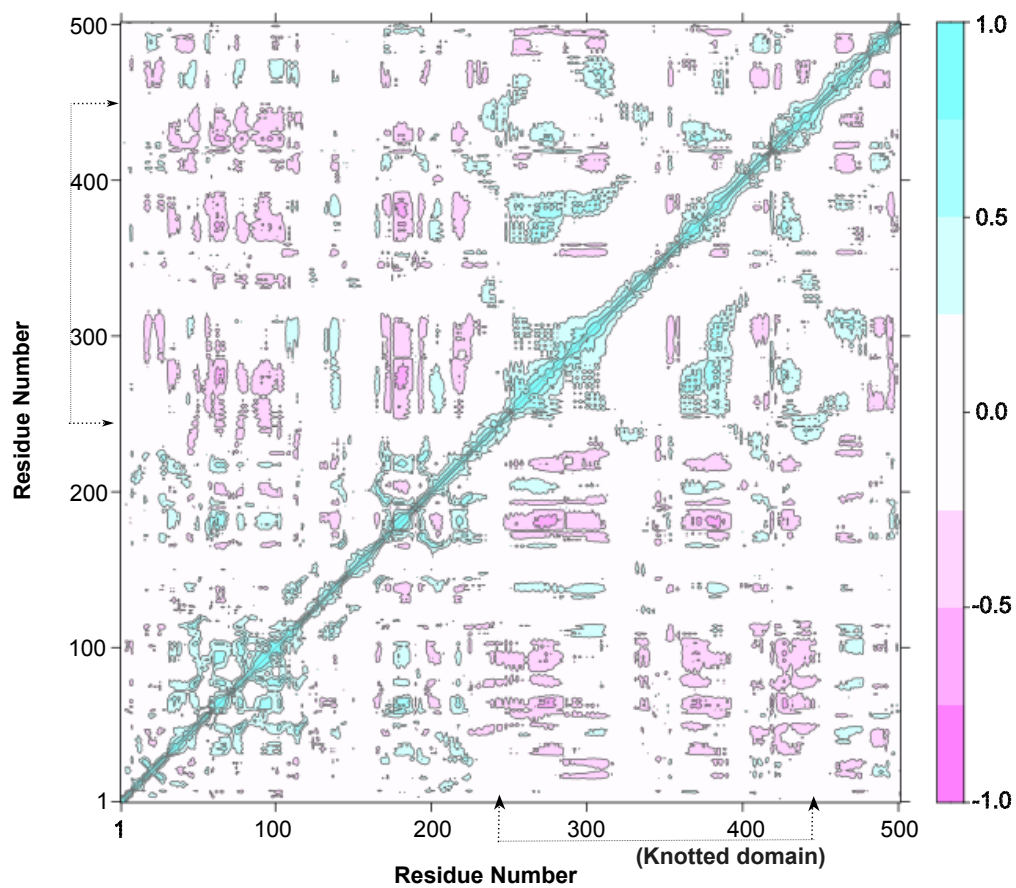
3   DCCM and Correlation Cutoffs

Figure S3: The DCCM map of 3fr8 indicating the knotted regions (residues 239 to 442).
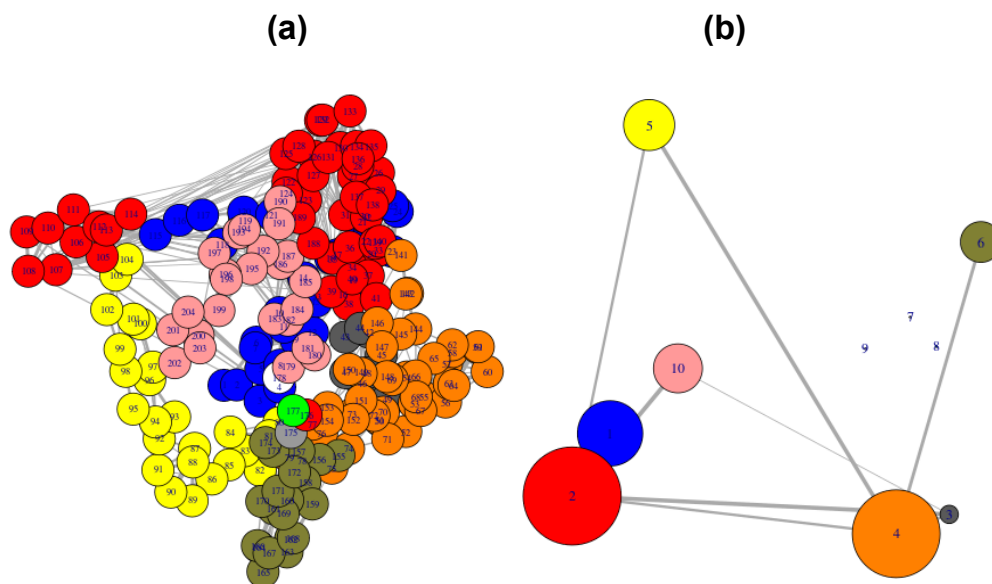
Figure S4: Full Residue Community Network (a) and the corresponding Minimalized Network of 3fr8$_{knot}$.

Table S1: Modularity scores and community nodes for the correlation cutoff for 3fr8$_{knot}$

cf	Community Nodes	Modularity
0.3	9	0.26
0.4	9	0.39
0.5	12	0.47
0.6	10	0.70
0.7	18	0.65

4 Clusters

Cluster 1 (30%)

—	**GLU15-ARG19**
—	**GLU28-ARG123**
—	**GLU29-LYS130**
—	**GLU37-ARG135**
—	**GLU61-LYS64**
—	**ASP88-ARG96**
—	**GLU129-ARG132**
—	**GLU160-ARG163**

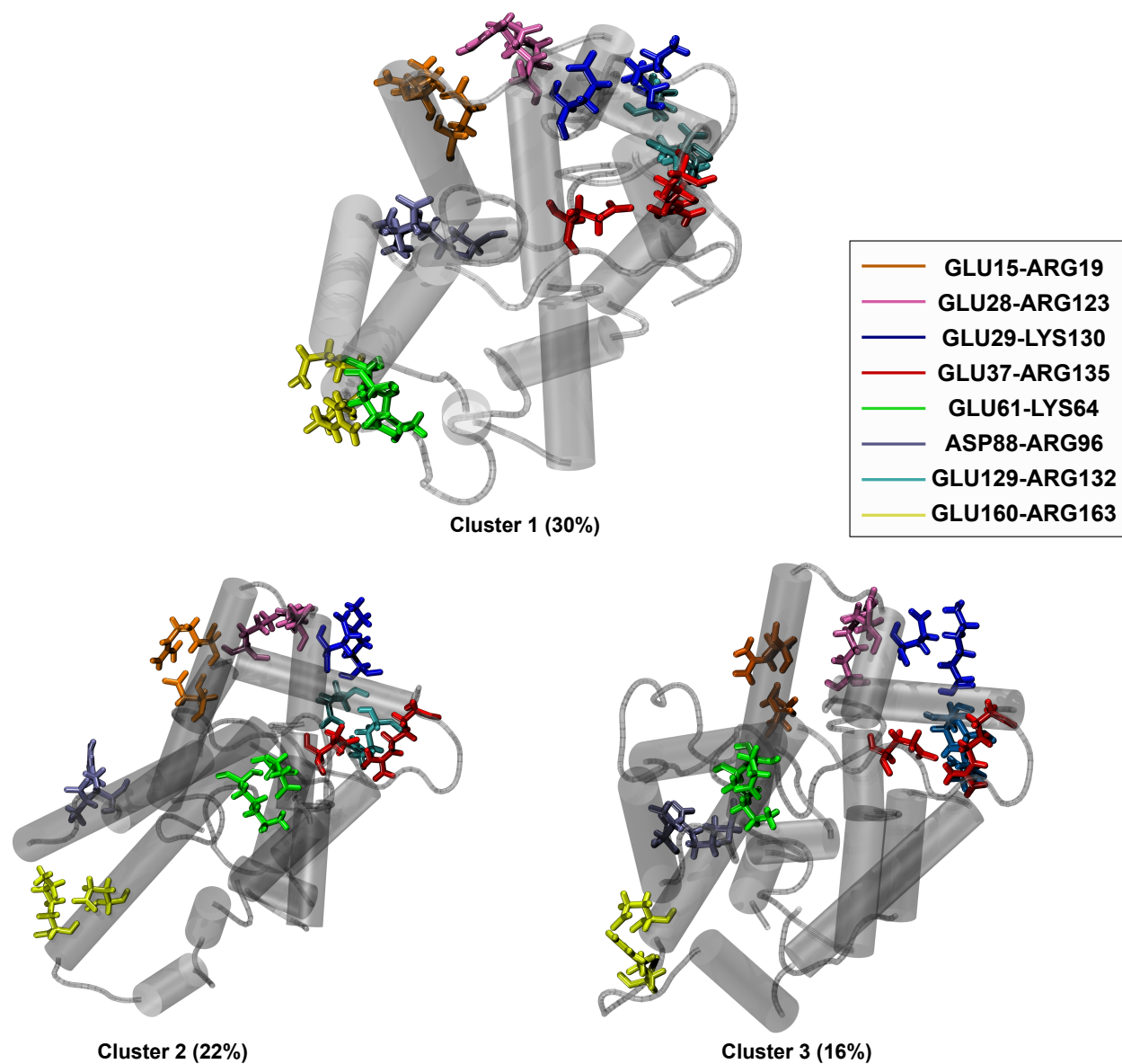**Cluster 2 (22%)**

Cluster 3 (16%)

Figure S5: Snapshots of the three most populated clusters highlighting the major salt bridges.

Table S2: Occupancy of the salt bridges with increasing temperatures

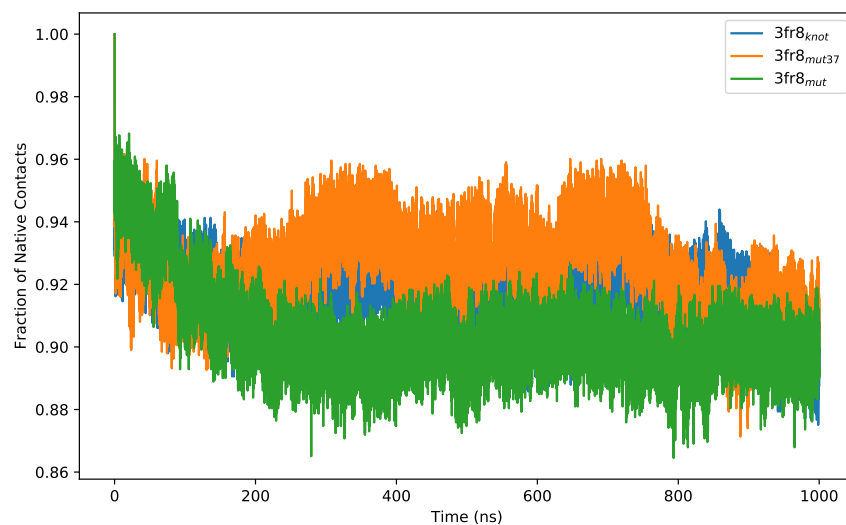| Salt Bridges	Occupancy		
	300K	400K	500K
GLU15:ARG19	0.75	0.56	0.17
GLU28:ARG123	0.70	0.31	
GLU29:LYS130	0.55	0.21	0.05
GLU37:ARG135	0.91	0.61	
GLU61:LYS64	0.37	0.39	0.29
ASP88:ARG96	0.25	0.06	
GLU129:ARG132	0.69	0.52	0.49
GLU160:ARG163	0.59	0.45	0.42

5 Mutated Systems

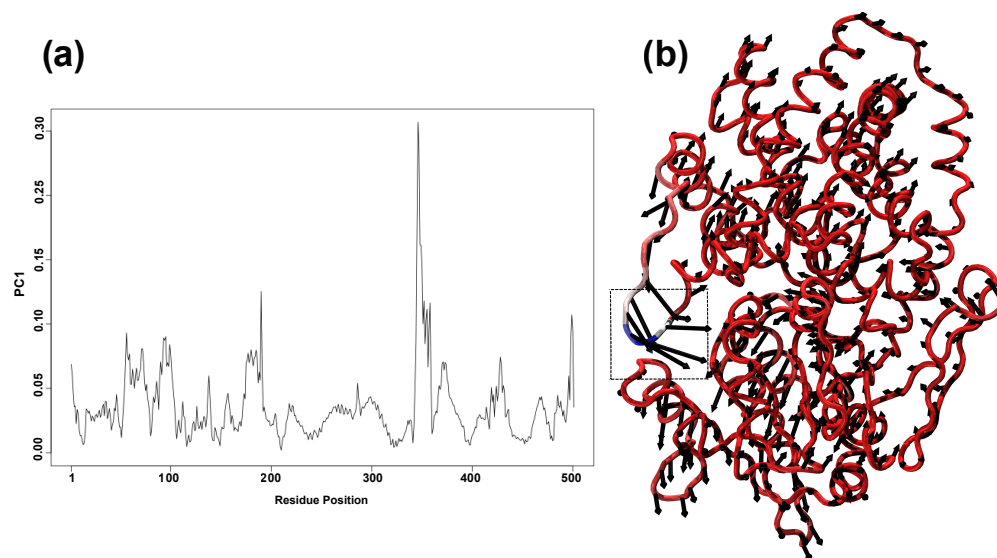Figure S6: Native Contact Ratio of the three protein at 300K.

Figure S7: Contribution of residues to the first Principal Component (PC1) (a) and Porcupine plots of PC1(b) for 3fr8 respectively. The rectangular box in (b) denotes the regions (residue 340 to 357) of high mobility.
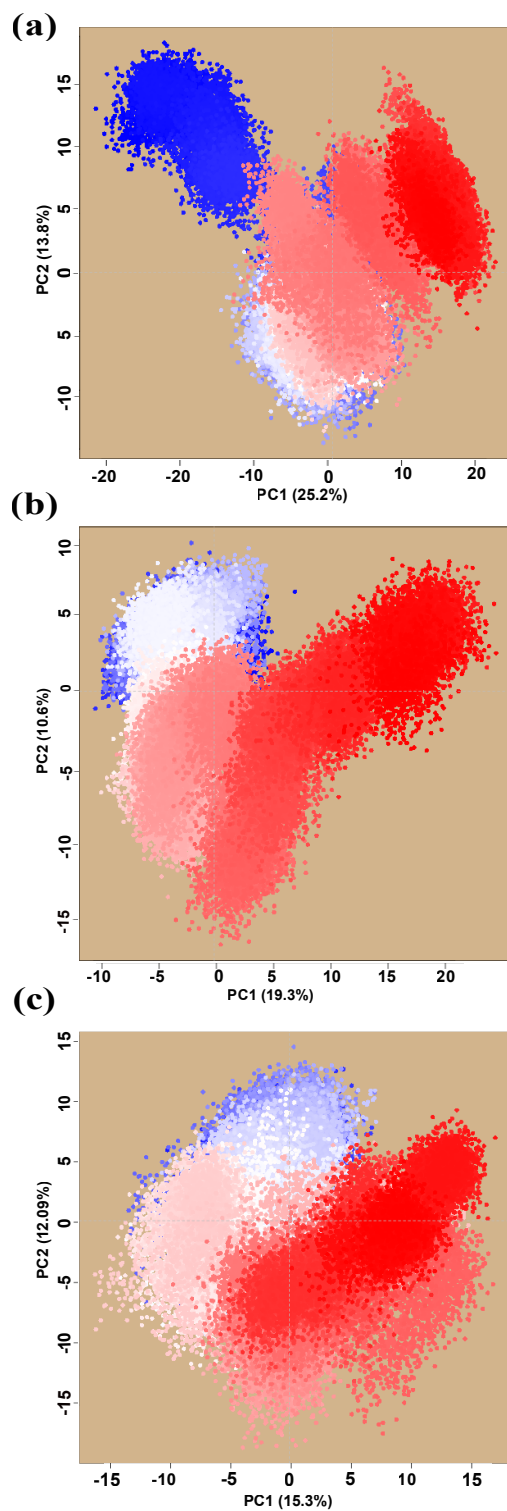
Figure S8: Principal Component Analysis (PCA) using first Principal Component (PC1) and second Principal Component (PC2) after removal of the β-turn residues 100-117 for (a) $3\text{fr}8_{knot}$, (b) $3\text{fr}8_{mut37}$, and (c) $3\text{fr}8_{mut}$ at 300K respectively.
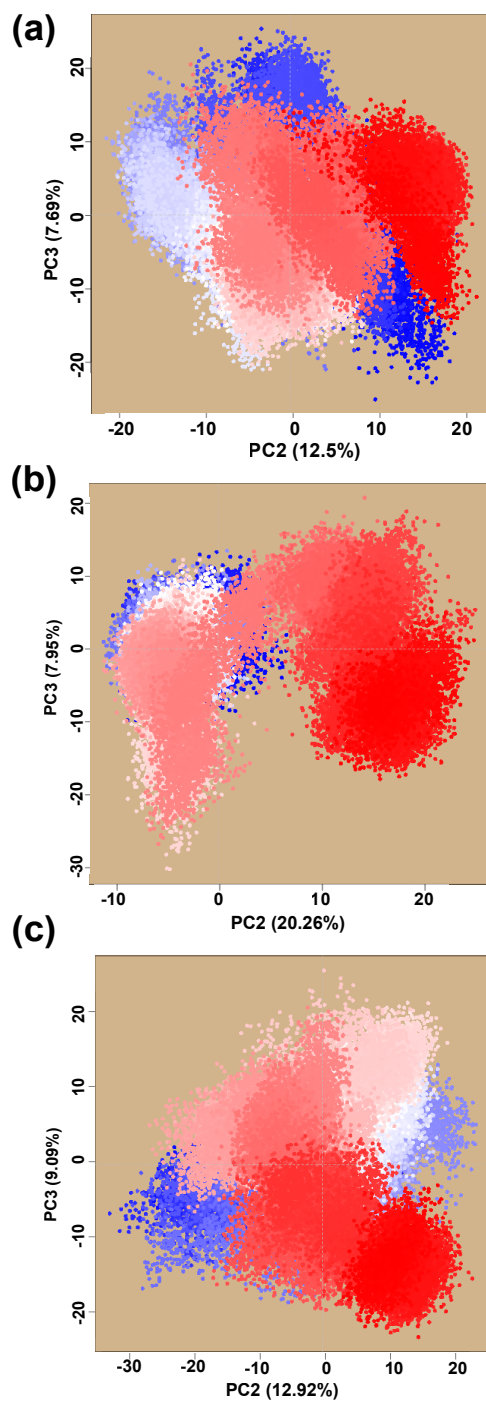
Figure S9: Principal Component Analysis (PCA) using second Principal Component (PC2) and third Principal Component (PC3) for (a) 3fr8$_{knot}$, (b) 3fr8$_{mut37}$, and (c) 3fr8$_{mut}$ at 300K respectively.
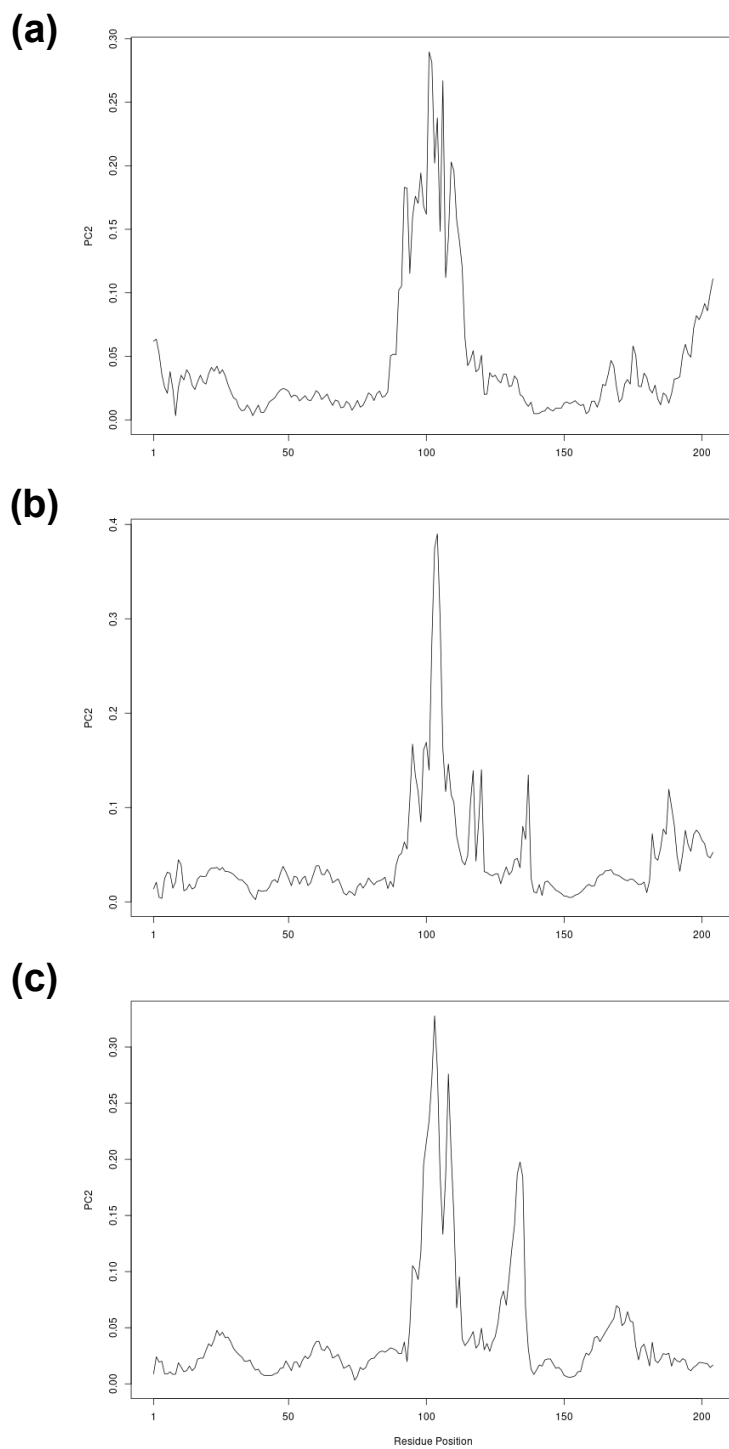
Figure S10: Contribution of residues to the second Principal Component (PC2) for (a) 3fr8$_{knot}$, (b) 3fr8$_{mut37}$, and (c) 3fr8$_{mut}$ at 300K respectively.

Table S3: Knotted regions of monomeric 3fr8 and the mutants at 500K for trajectory 2

Time (ns)	Core Length			Core Range		
	3fr8	3fr8-mut37	3fr8-mut	3fr8	3fr8-mut37	3fr8-mut
0	203	203	203	239-442	239-442	239-442
100	222	221	208	226-448	223-444	237-445
200	329	335	335	112-441	101-436	105-440
300	356	348	360	87-443	89-437	72-432
400	406	-	-	33-439	-	-
500	410	384	386	48-448	59-443	55-441
600	408	424	-	47-455	45-469	-
700	418	-	411	37-455	-	47-458
800	476	-	-	17-493	-	-
900	484	476	478	7-491	10-486	7-485
1000	469	-	484	23-492	-	8-492

Table S4: Knotted regions of monomeric 3fr8 and the mutants at 500K for trajectory 3

Time (ns)	Core Length			Core Range		
	3fr8	3fr8-mut37	3fr8-mut	3fr8	3fr8-mut37	3fr8-mut
0	203	203	203	239-442	239-442	239-442
100	232	228	226	213-445	232-460	234-460
200	348	327	329	89-437	232-459	128-457
300	384	360	354	59-443	73-433	87-441
400	402	402	393	48-450	52-454	56-449
500	418	411	415	37-455	48-459	35-450
600	408	415	-	47-455	37-452	-
700	389	-	402	50-439	-	50-452
800	-	418	420	-	38-456	41-461
900	409	470	478	36-445	15-485	15-493
1000	484	469	476	10-494	22-491	11-487

References

(1) Mesentean, S.; Fischer, S.; Smith, J. C. Analyzing large-scale structural change in proteins: Comparison of principal component projection and sammon mapping. *Proteins: Struct., Funct., Bioinf.* **2006**, *64*, 210–218.

(2) Skjaerven, L.; Martinez, A.; Reuter, N. Principal component and normal mode analysis of proteins; a quantitative comparison using the GroEL subunit. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 232–243.

(3) Arnold, G. E.; Ornstein, R. L. Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: cytochrome P450BM-3. *Biophys. J.* **1997**, *73*, 1147–1159.

(4) Junaid, M.; Shah, M.; Khan, A.; Li, C.-D.; Khan, M. T.; Kaushik, A. C.; Ali, A.; Mehmood, A.; Nangraj, A. S.; Choi, S., et al. Structural-dynamic insights into the H. pylori cytotoxin-associated gene A (CagA) and its abrogation to interact with the tumor suppressor protein ASPP2 using decoy peptides. *J. Biomol. Struct. Dyn.* **2018**,

(5) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(6) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. Multidimensional free-energy calculations using the weighted histogram analysis method. *J. Comput. Chem.* **1995**, *16*, 1339–1350.