Supplementary Information

A Distal Regulatory Strategy of Enzyme: From Local to Global Conformational Dynamics

Xue Peng,‡[a] *Chenlin Lu,*‡ [a] *Jian Pang,*[a] *Zheng liu,*[a] *and Diannan Lu*[a*]

State Key Lab of Chemical Engineering, Ministry of Science and Technology

Department of Chemical Engineering, Tsinghua University

Beijing, 100084, China

Corresponding author: ludiannan@tsinghua.edu.cn

Xue Peng and Chenlin Lu contributed equally to the article.

Contents

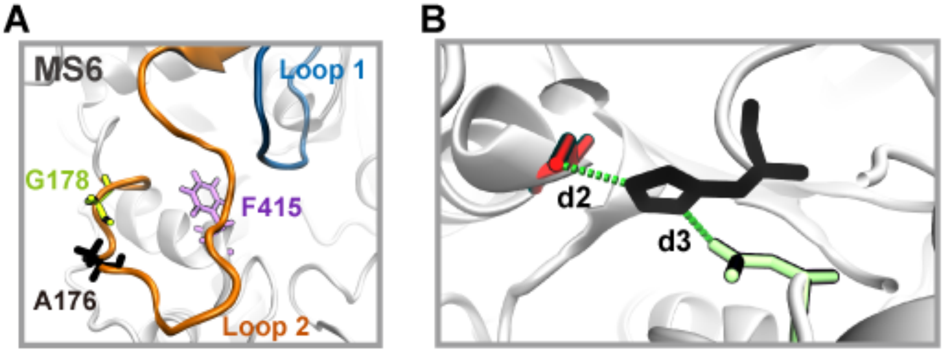

Figure S1. (A) The spatially location of A176, G178, and F415 on the protein. **(B)** Definition of Distance d2 and Distance d3. S225, D492 and H528 are shown in red, green, and black sticks, respectively. The protein is colored in gray.

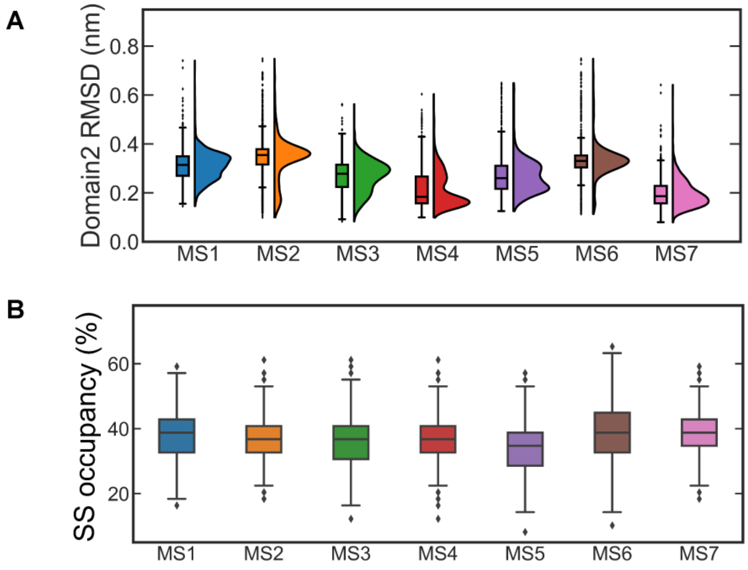

Figure S2. **(A)** The RMSD values of the Domain 2 in MS1–MS7. **(B)** The secondary structure occupancies of Domain 2 in MS1–MS7.

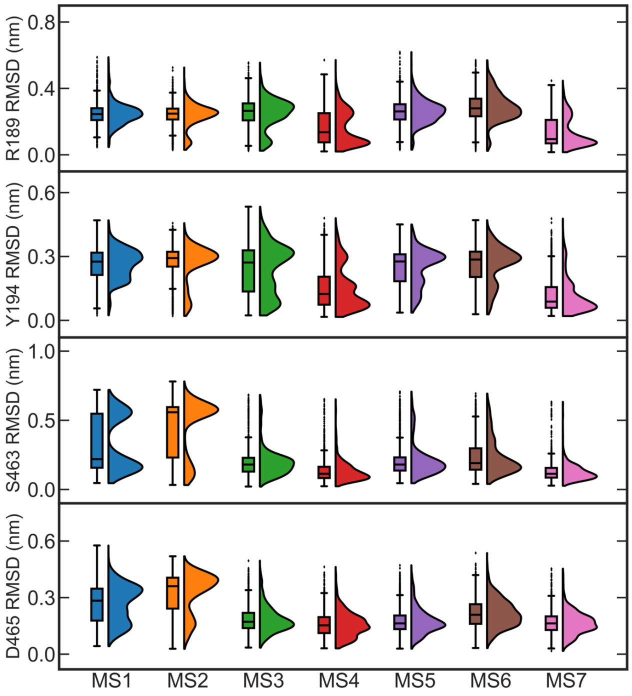

Figure S3. RMSD distribution of residues in Domain 2 with large variations in MS1–MS7.

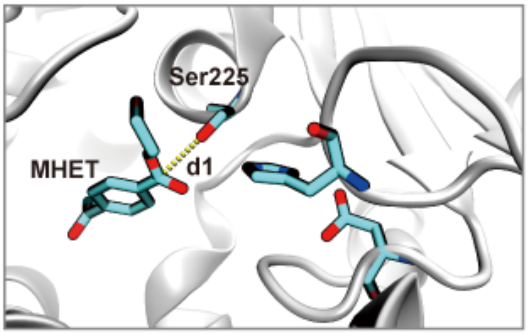

Figure S4. Definition of Distance d1.

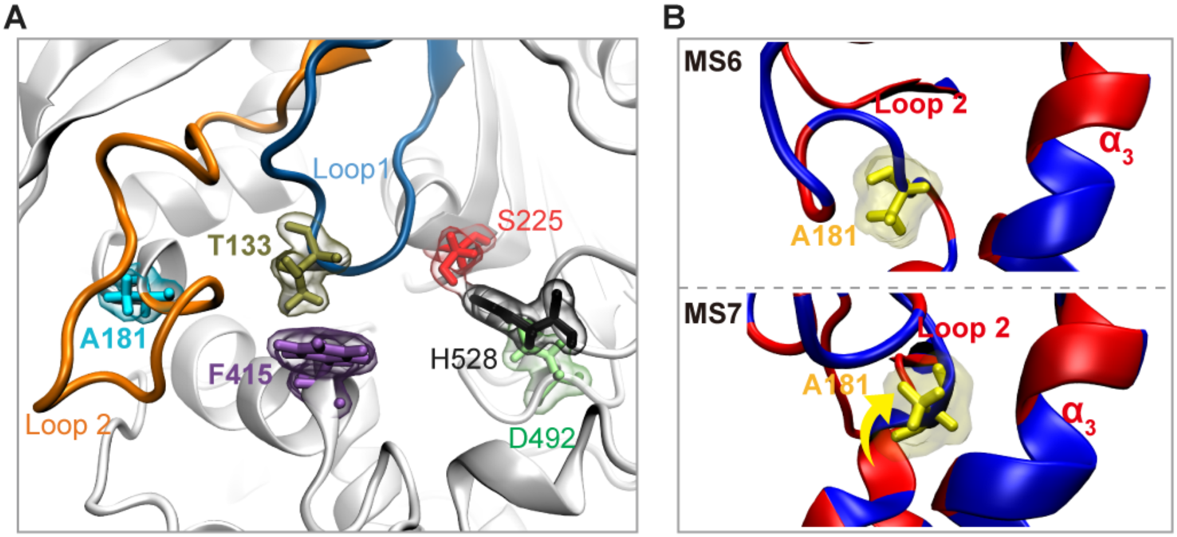

Figure S5. **(A)** The spatially location of key residues on the protein. **(B)** The rotation of A181 along with the transition from MS6 to MS7. The protein backbone is shown in the cartoon, with hydrophobic and hydrophilic regions colored in blue and red, respectively.

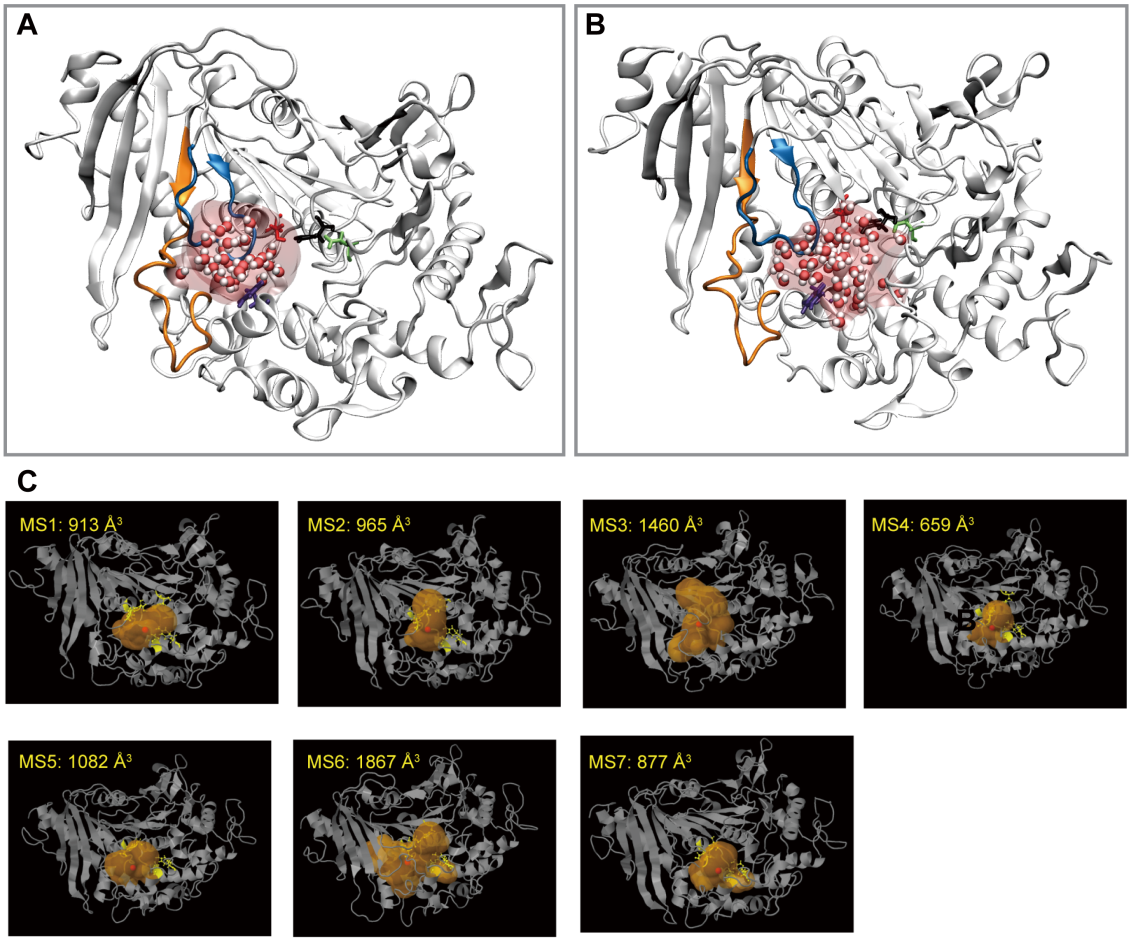

Figure S6. **(A)** Position of Cavity 1 (red bubble) in MHETase. **(B)** Position of Cavity 2 (red bubble) in MHETase. S225, D492, H528, and F415 are shown in red, green, black, and purple sticks, respectively. The Loop 1, Loop 2 and protein is colored in blue, orange, and gray. Water molecules are shown as a three-dotted sphere model. **(C)** The volume of the cavity that is composed of residue G132, S225, E226, R411, S416, H528 for the representative conformation in MS1 to MS7 were visualized and quantatively characterized using Caver Web server.[1]

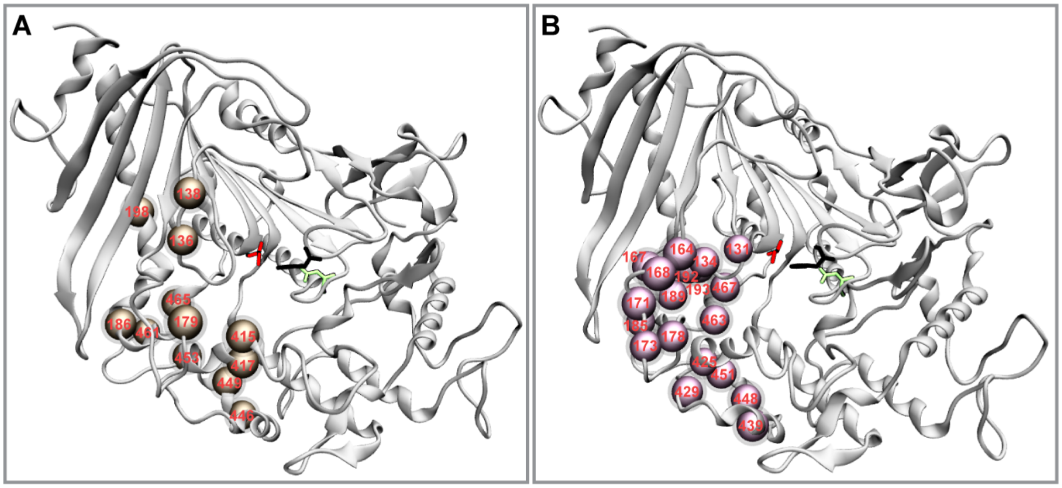

Figure S7. The key residues (colored spheres) of MS6 **(A)** and MS7 **(B)**, identified in H-bond occupancy.

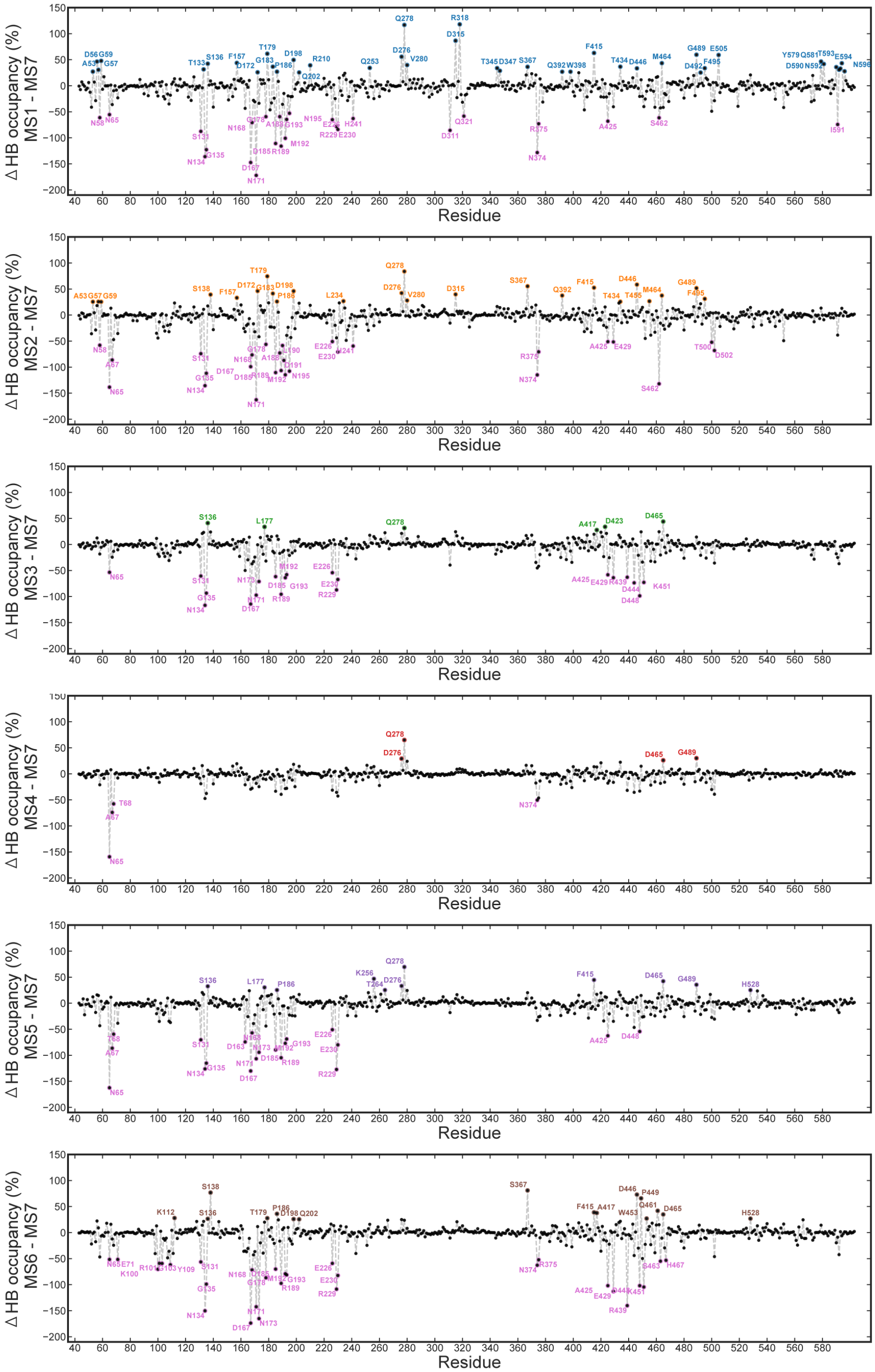

Figure S8. The differences in the H-bond occupancy between MS1-MS6 and MS7. The residues with large differences are labeled in the figure, according to the corresponding metastable state colors.

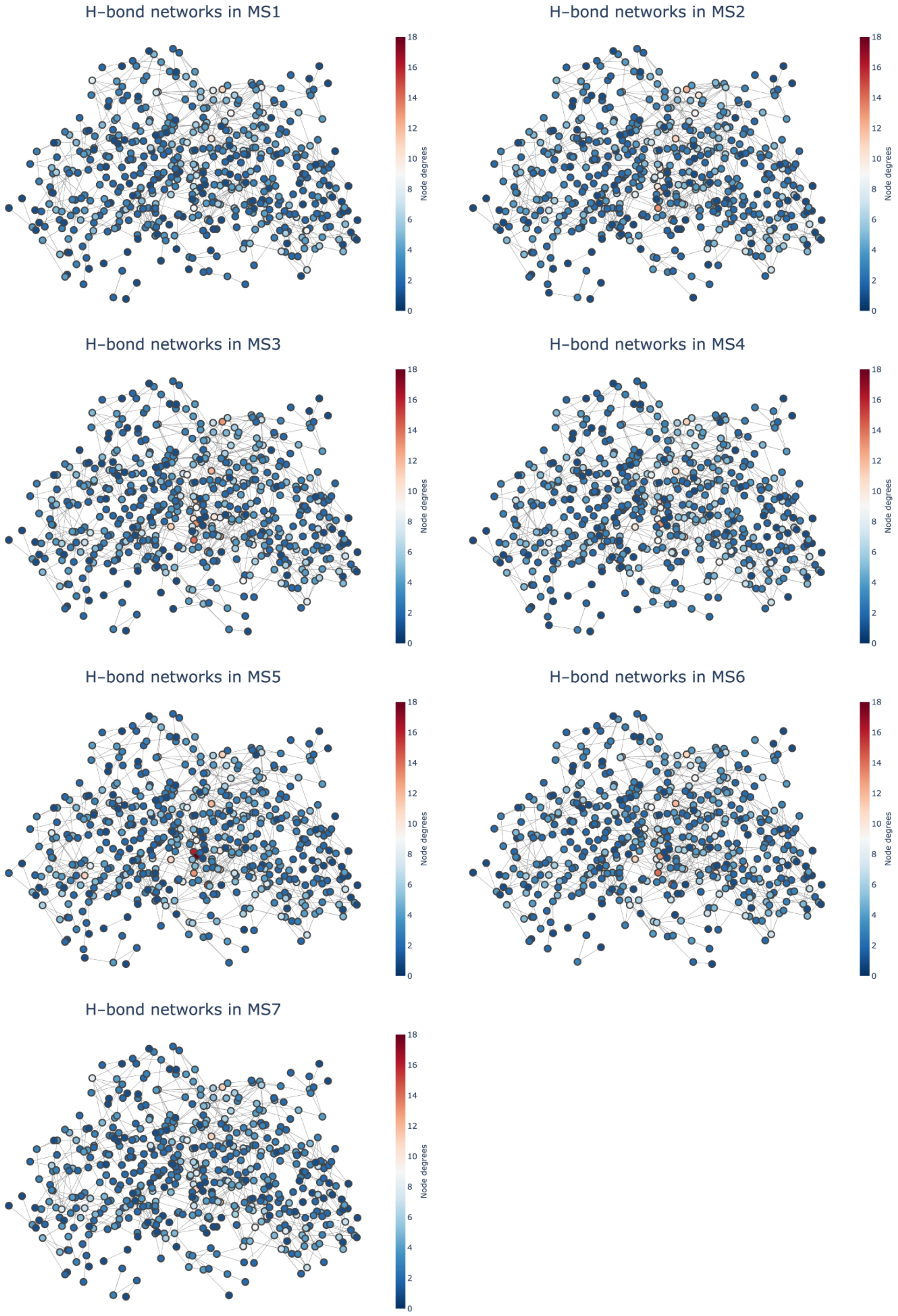

Figure S9. The H-bond networks in MS1-MS7. Residues are represented by circles and H-bonds are connected by straight lines. Interactive plot can be found in the supplementary HTML files, available at https://github.com/thuresearch/MHETase_SI.

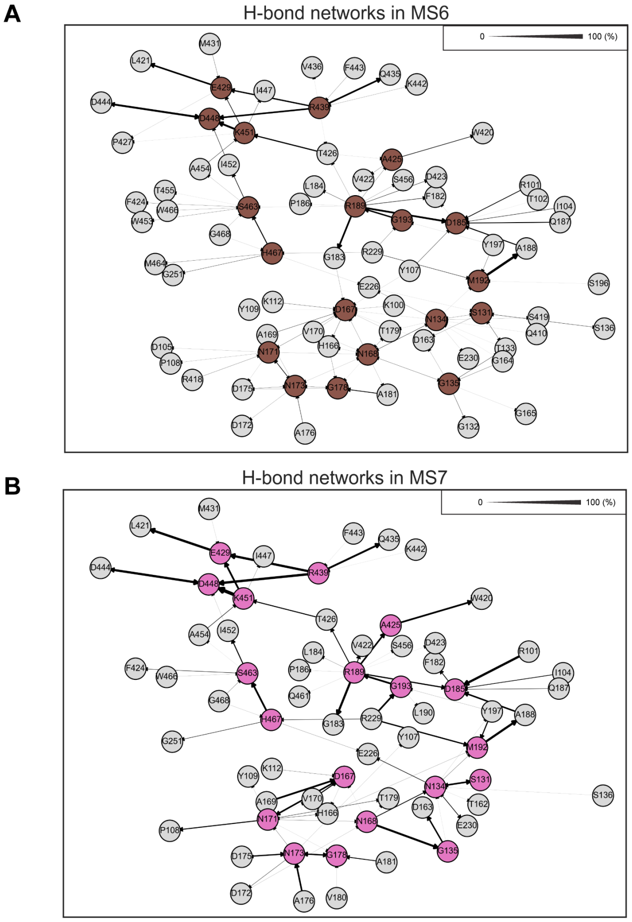

Figure S10. H-bond networks of key residues identified of MS7 in MS6 **(A)** and MS7 **(B)**. The thickness of the line represents the H-bond occupancy formed in all trajectories, with thicker lines indicating a higher probability of H-bond formation.

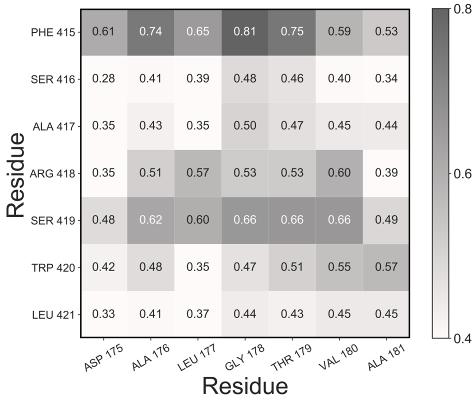

Figure S11. Dynamic cross-correlation plots of residues pairs within residue 415–421 and residue 175–181 in Free-MHETase. The darker the grayscale the stronger the correlation is.

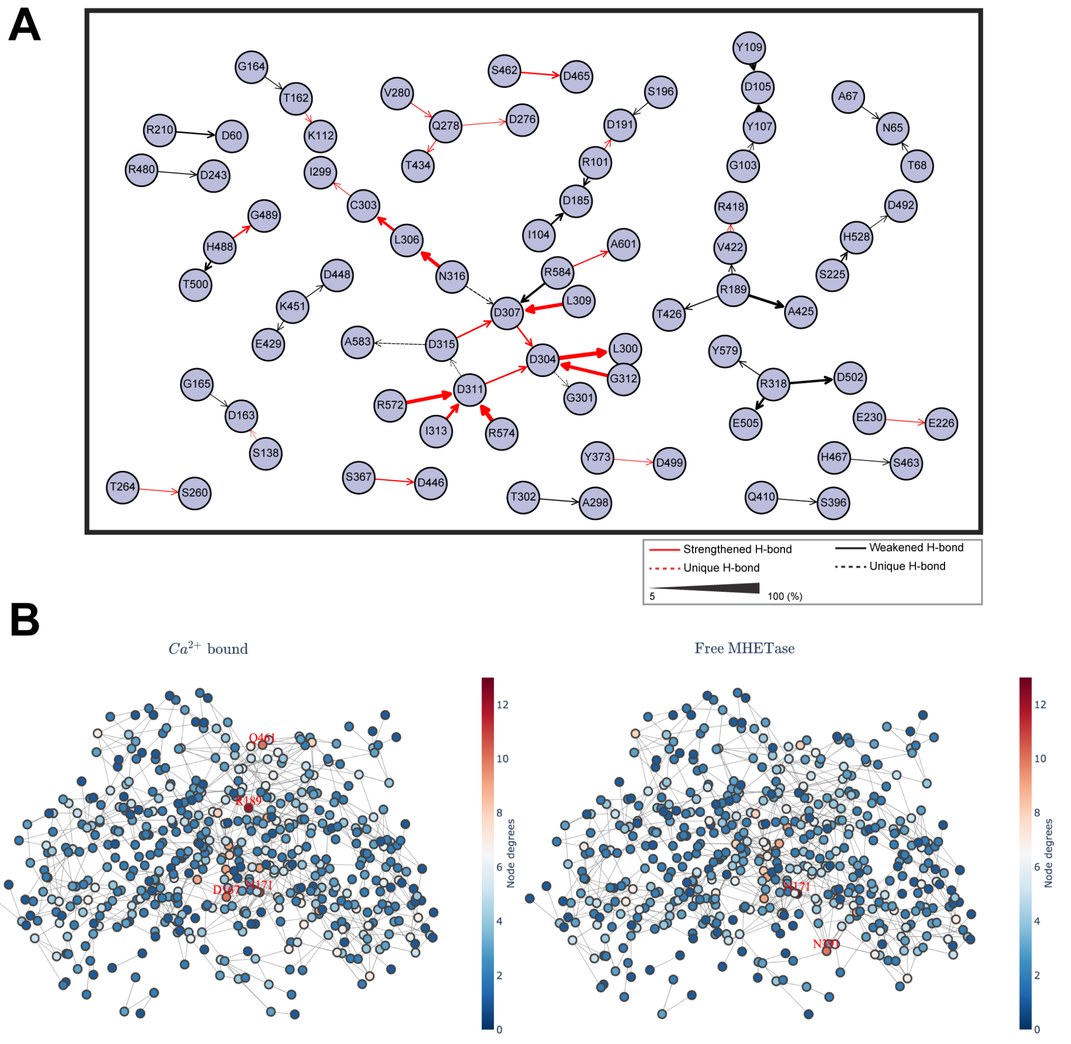

Figure S12. **(A)** Changes in H-bond networks after Ca^{2+} deportation. The thickness of the line is proportional to the amount of variation with the red line representing a strengthened H-bond and the black line representing a weakened H-bond. The dashed lines represent unique H-bonds that do not form in the other system. **(B)** The H-bond networks in Ca-MHETase and Free-MHETase. Residues are represented by circles and H-bonds are connected by straight lines. Interactive plot can be found in the supplementary HTML files, available at https://github.com/thuresearch/MHETase_SI.

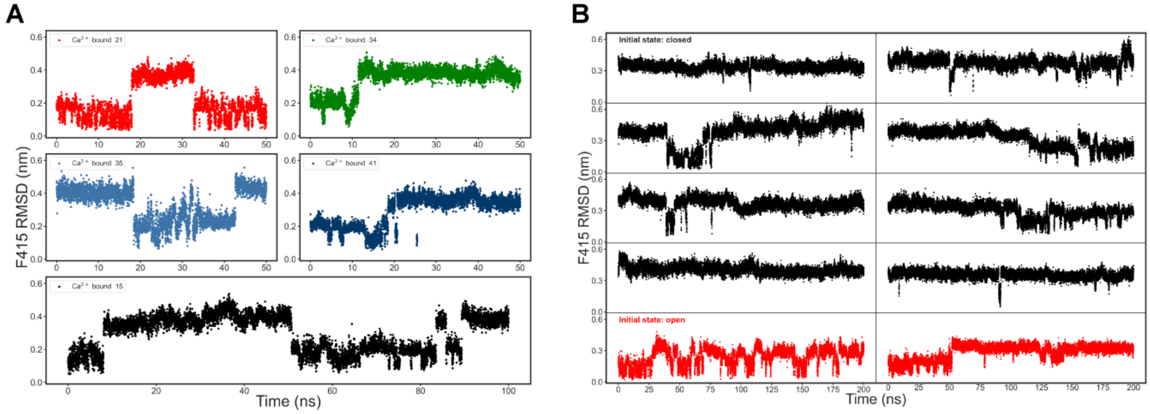

Figure S13. **(A)** Typical F415 backbone RMSD changes during the opening process in the first two rounds of MD simulations. **(B)** The F415 backbone RMSD changes in the third sampling round, the first eight trajectories begun with a closed state of F415 (black line) and the last two trajectories begun with an open state (red line).

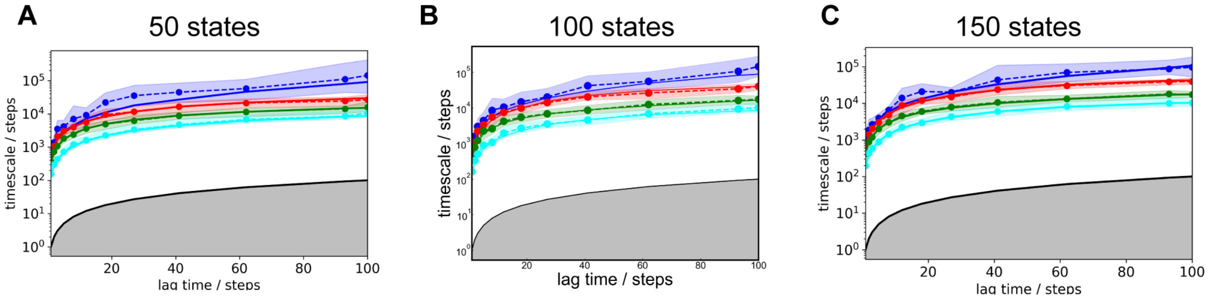

Figure S14. The implied timescale plots as a function of different number of microstates, 50 states **(A),** 100 states **(B),** and 150 states **(C)**, respectively.

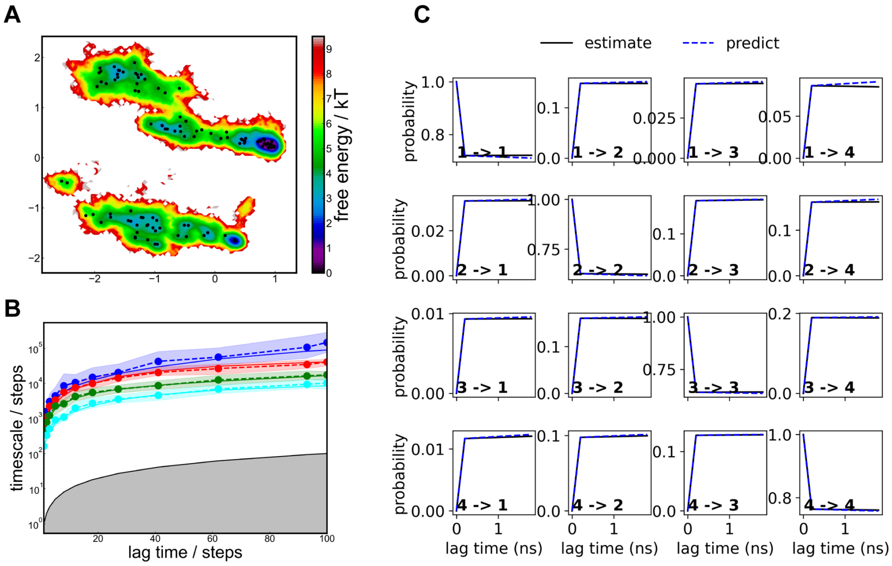

Figure S15. **(A)** Free energy projections of K-means clusters. **(B)** The implied timescale plots and the 1σ confidence interval. **(C)** The Chapman–Kolmogorow test of the Markov model.

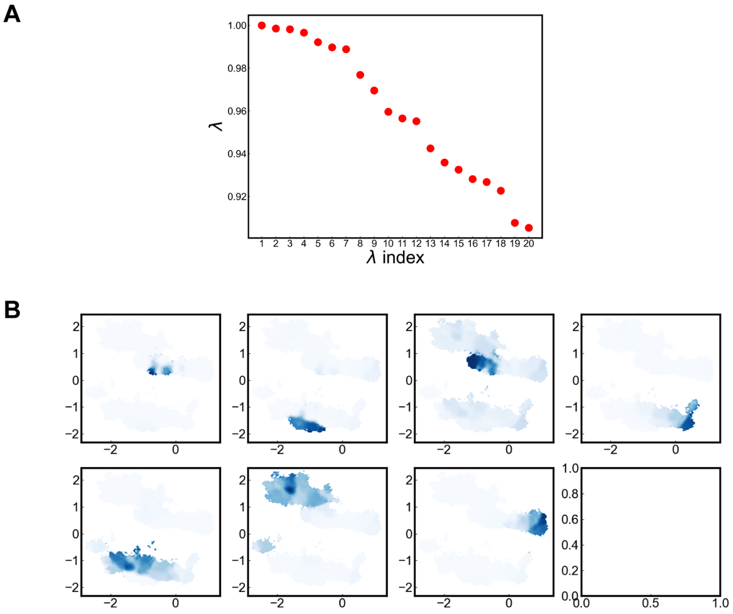

Figure S16. **(A)** Eigenvalues with a gap after the seventh point of the Markov state model transition matrix. **(B)** The conformational density plots of the seven metastable states with PCCA+ algorithm.

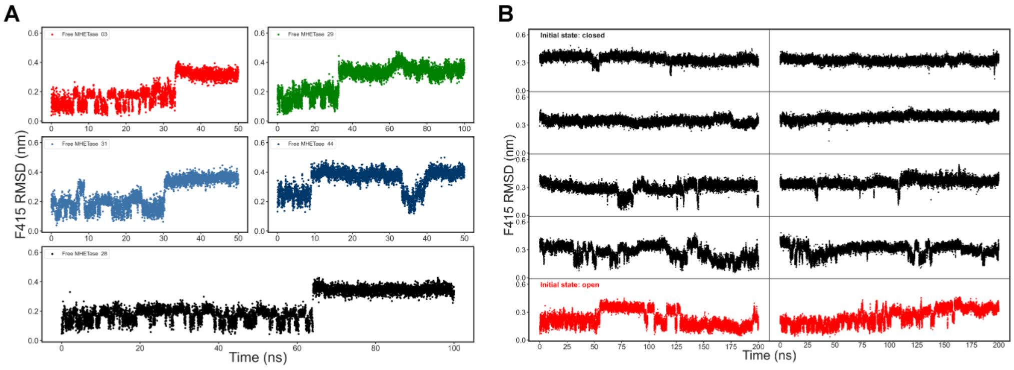

Figure S17. **(A)** Typical F415 backbone RMSD changes during the opening process in the first two rounds of MD simulations in Free-MHETase. **(B)** The F415 backbone RMSD changes in the third sampling round in Free-MHETase, the first eight trajectories begun with a closed state of F415 (black line), and the last two trajectories begun with an open state (red line).

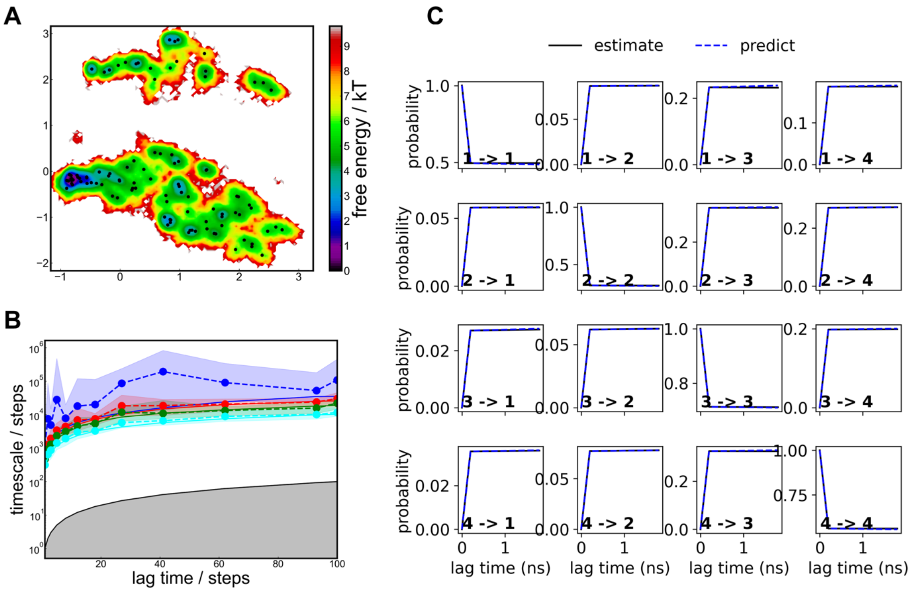

Figure S18. **(A)** Free energy projections of K-means clusters. **(B)** The implied timescale plots and the 1σ confidence interval. **(C)** The Chapman–Kolmogorow test of the Markov model. The model is Free-MHETase.

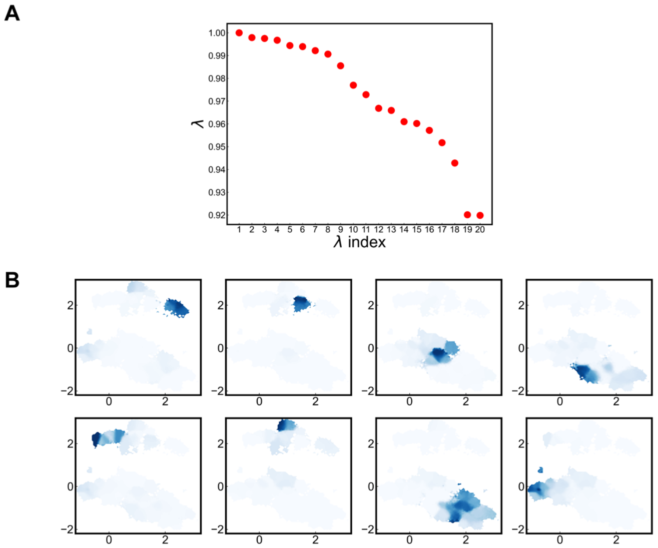

Figure S19. **(A)** Eigenvalues with a gap after the seventh point of the Markov state model transition matrix. **(B)** The conformational density plots of the seven metastable states with PCCA+ algorithm. The model is Free-MHETase.

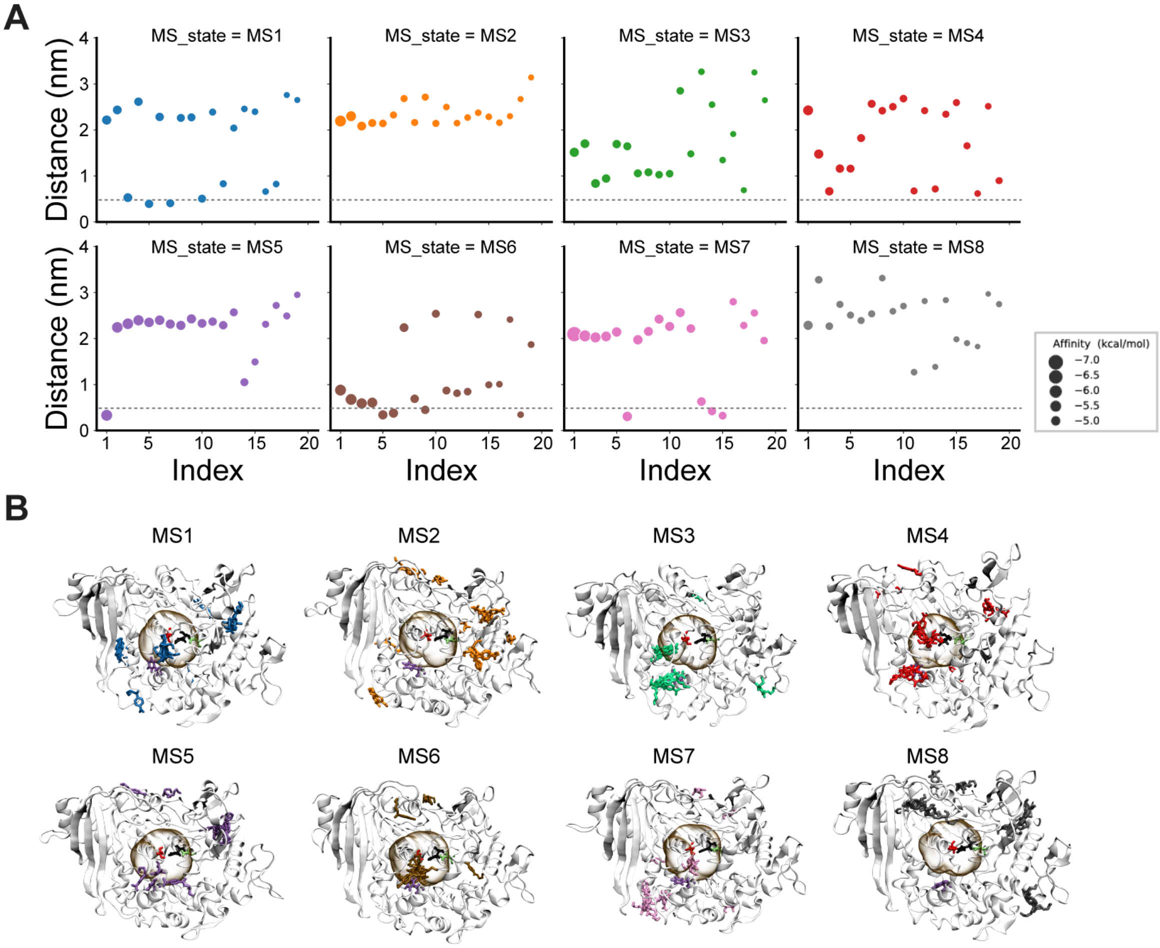

Figure S20. **(A)** Statistics of molecular docking results of MS1-MS8 for Free-MHETase. The x-axis represents the index of the binding pose (from 1st to 20th), where the y-value is the distance between the hydroxyl oxygen atom of the Ser225 and the carbonyl carbon atom of the substrate MHET (SI, Figure S4). A pose is marked as active if the distance is below the dashed line (cut off: 5 Å). The scatter size is proportional to the binding affinity free energy. **(B)** The top 20 docking poses of MS1–MS8 for Free-MHETase. The transparent bubble represents the space at a distance of 7 Å from the Ser225 hydroxyl oxygen atom.

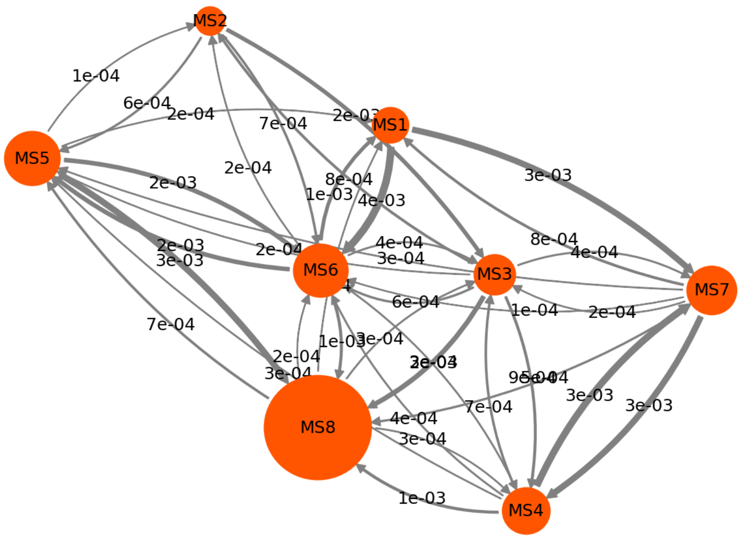

Figure S21. MSM model for Free-MHETase with eight metastable states. The circle size represents the equilibrium probability of the corresponding state: 5.3 ± 6.0% (MS1), 3.4 ± 4.9% (MS2), 6.4 ± 4.3% (MS3), 8.6 ± 1.0% (MS4), 11.9 ± 1.7% (MS5), 11.8 ± 1.7% (MS6), and 9.3 ± 3.8% (MS7), and 43.3 ± 4.1% (MS8).

Table S1 Simulation details

System	**Sampling cycle #**	**Sim #**	**Initial protein structure**	**Length (ns)**	**F415 state**
Ca-MHETase	1	1	6JTU	100	Open
		2	6JTU	100	Open
		3	6JTU	100	Open
		4	6JTU	100	Open
		5	6JTU	100	Open
		6	6JTU	100	Open
		7	6JTU	100	Open

		8	6JTU	100	Open
		9	6JTU	100	Open
		10	6JTU	100	Open
	2	11	Sim 1 6041 Frame	100	Open
		12	Sim 2 5101 Frame	100	Open
		13	Sim 3 6889 Frame	100	Open
		14	Sim 4 6254 Frame	100	Open
		15	Sim 5 5504 Frame	100	Open/Closed 1
		16	Sim 6 7127 Frame	100	Open
		17	Sim 7 6675 Frame	100	Open
		18	Sim 8 7788 Frame	100	Open
		19	Sim 9 6903 Frame	100	Open
		20	Sim 10 6264 Frame	100	Open
		21	Sim 11 5278 Frame	50	Open/Closed 1

		22	Sim 11 6360 Frame	50	Open
		23	Sim 11 8680 Frame	50	Open
		24	Sim 12 4822 Frame	50	Open
		25	Sim 12 4362 Frame	50	Open
		26	Sim 12 3756 Frame	50	Open
		27	Sim 13 5019 Frame	50	Open
		28	Sim 13 2310 Frame	50	Open
		29	Sim 13 6702 Frame	50	Open/Closed 1
		30	Sim 14 5367 Frame	50	Open
		31	Sim 14 7139 Frame	50	Open
		32	Sim 14 4422 Frame	50	Open
		33	Sim 15 2420 Frame	50	Open
		34	Sim 15 6842 Frame	50	Open/Closed 1

		35	Sim 15 9948 Frame	50	Closed 1/ Open
		36	Sim 16 4182 Frame	50	Open
		37	Sim 16 7022 Frame	50	Open
		38	Sim 16 6355 Frame	50	Open
		39	Sim 17 2589 Frame	50	Open
		40	Sim 17 5575 Frame	50	Open
		41	Sim 17 7484 Frame	50	Open/Closed 1
		42	Sim 18 4054 Frame	50	Open
		43	Sim 18 7973 Frame	50	Open
		44	Sim 18 9363 Frame	50	Open
		45	Sim 19 2338 Frame	50	Open
		46	Sim 19 6820 Frame	50	Open
		47	Sim 19 , 5895 Frame	50	Open

		48	Sim 20 5882 Frame	50	Open
		49	Sim 20 6090 Frame	50	Open
		50	Sim 20 9519 Frame	50	Open
	3	51	Sim 15 3500 Frame	200	Closed 1
		52	Sim 15 10000 Frame	200	Closed 1
		53	Sim 21 2600 Frame	200	Closed 1/ Open
		54	Sim 29 3000 Frame	200	Closed 1/ Open
		55	Sim 29 4000 Frame	200	Closed 1/ Open
		56	Sim 34 5000 Frame	200	Closed 1/ Open
		57	Sim 35 1000 Frame	200	Closed 1
		58	Sim 41 5000 Frame	200	Closed 1
		59	Sim 46 4000 Frame	200	Open/Closed 1
		60	Sim 47 4000 Frame	200	Open/Closed 1

Free-MHETase	1	1	6JTU	50	Open
		2	6JTU	50	Open
		3	6JTU	50	Open/Closed 1
		4	6JTU	50	Open
		5	6JTU	50	Open
		6	6JTU	50	Open
		7	6JTU	50	Open
		8	6JTU	50	Open
		9	6JTU	50	Open
		10	6JTU	50	Open
	2	11	Sim 1 3834 Frame	100	Open
		12	Sim 2 2431 Frame	100	Open
		13	Sim 3 4019 Frame	100	Closed 1/Open
		14	Sim 4 3010 Frame	100	Open
		15	Sim 5 2072 Frame	100	Open
		16	Sim 6 4540 Frame	100	Open
		17	Sim 7 3750 Frame	100	Open
		18	Sim 8 2062 Frame	100	Open

		19	Sim 9 3860 Frame	100	Open
		20	Sim 10 4749 Frame	100	Open
		21	Sim 11 4089 Frame	100	Open
		22	Sim 11 5991 Frame	100	Open
		23	Sim 11 8733 Frame	100	Open
		24	Sim 12 6084 Frame	100	Open
		25	Sim 12 6762 Frame	100	Open
		26	Sim 12 8336 Frame	100	Open
		27	Sim 13 4356 Frame	100	Open
		28	Sim 13 4721 Frame	100	Open/Closed 1
		29	Sim 13 4429 Frame	100	Open/Closed 1
		30	Sim 14 7230 Frame	100	Open
		31	Sim 14 5566 Frame	50	Open/Closed 1

		32	Sim 14 7928 Frame	50	Open
		33	Sim 15 7623 Frame	50	Open
		34	Sim 15 7318 Frame	50	Open
		35	Sim 15 6828 Frame	50	Open
		36	Sim 16 8042 Frame	50	Open/Closed 1
		37	Sim 16 4186 Frame	50	Open
		38	Sim 16 5867 Frame	50	Open
		39	Sim 17 3928 Frame	50	Open
		40	Sim 17 2081 Frame	50	Open
		41	Sim 17 4569 Frame	50	Open
		42	Sim 18 3810 Frame	50	Open
		43	Sim 18 8049 Frame	50	Open
		44	Sim 18 5578 Frame	50	Open/Closed 1

		45	Sim 19 9350 Frame	50	Open
		46	Sim 19 7315 Frame	50	Open
		47	Sim 19 , 6813 Frame	50	Open
		48	Sim 20 5492 Frame	50	Open
		49	Sim 20 5120 Frame	50	Open
		50	Sim 20 9519 Frame	50	Open
	3	51	Sim 28 8000 Frame	200	Closed 1
		52	Sim 28 9500 Frame	200	Closed 1
		53	Sim 29 8300 Frame	200	Closed 1
		54	Sim 29 9920 Frame	200	Closed 1
		55	Sim 36 2000 Frame	200	Closed 1/ Open
		56	Sim 36 2800 Frame	200	Closed 1
		57	Sim 44 2500 Frame	200	Closed 1/ Open

		58	Sim 44 2800 Frame	200	Closed 1/ Open
		59	Sim 45 2500 Frame	200	Open/Closed 1
		60	Sim 50 2500 Frame	200	Open/Closed 1

Table S2 RMSD of eight metastable states in Free-MHETase with the references of seven metastable states in Ca-MHETase.

RMSD		**Ca-MHETase**						
		MS1	**MS2**	**MS3**	**MS4**	**MS5**	**MS6**	**MS7**
Free-MHETase	**MS1**	0.303445	0.305995	0.312023	0.287989	0.31662	0.330853	0.286973
	MS2	0.317544	0.327186	0.306193	0.283021	0.33001	0.336131	0.292809
	MS3	0.307879	0.30185	0.29416	0.268462	0.311884	0.324962	0.278758
	MS4	0.33638	0.335532	0.311916	0.299728	0.32534	0.347704	0.301385
	MS5	0.333374	0.335743	0.321672	0.287308	0.337353	0.350142	0.292249
	MS6	0.330571	0.327854	0.306022	0.282105	0.341825	0.338864	0.290315
	MS7	0.29154	0.292791	0.306369	0.275362	0.304206	0.312149	0.27423
	MS8	0.293589	0.298865	0.30354	0.251464	0.316657	0.321113	0.263475

References

[1] Stourac, J., Vavra, O., Kokkonen, P., Filipovic, J., Pinto, G., Brezovsky, J., Damborsky, J., & Bednar, D. (2019). Caver Web 1.0: identification of tunnels and channels in proteins and analysis of ligand transport. Nucleic Acids Research, 47(W1), W414–W422.