Electronic Supplementary Material (ESI) for RSC Advances.

Understand R882H mutation effect of DNA methyltransferase DNMT3A: a combination of molecular dynamics simulation and QM/MM calculation

Lanxuan Liu,[1] Ting Shi,[1] Kendall N. Houk,[2] Yi-Lei Zhao*[1, 2]

[1]*State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China*

[2]*Department of Chemistry and Biochemistry, University of California, Los Angeles, California, USA.*

*To whom correspondence should be addressed:

Prof. Yi-Lei Zhao

School of Life Sciences and Biotechnology

Shanghai Jiao Tong University

800 Dongchuan Road, Shanghai 200240, China

Tel/Fax: +86-21-34207190;

Email: yileizhao@sjtu.edu.cn

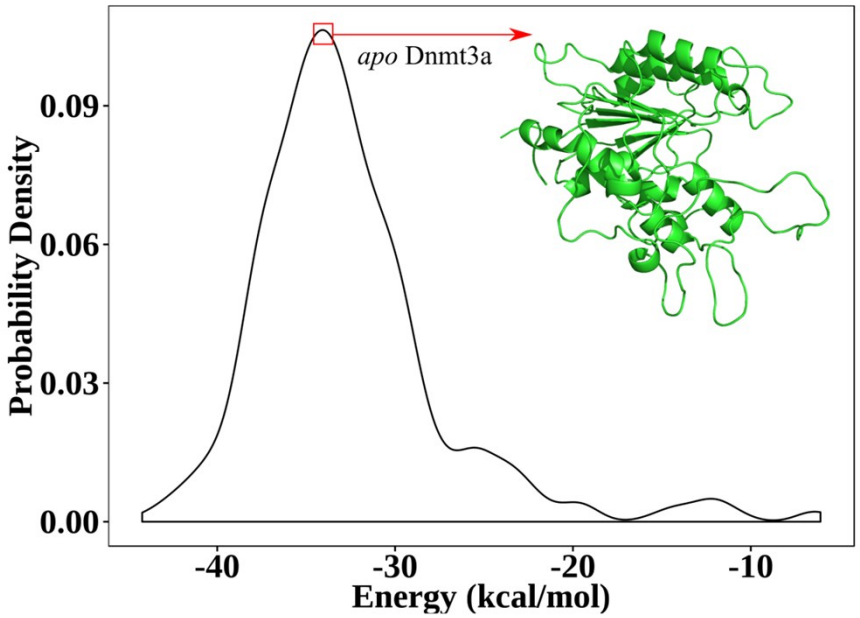

Fig. S1 *WT-apo* DNMT3A modeled by Rosetta.

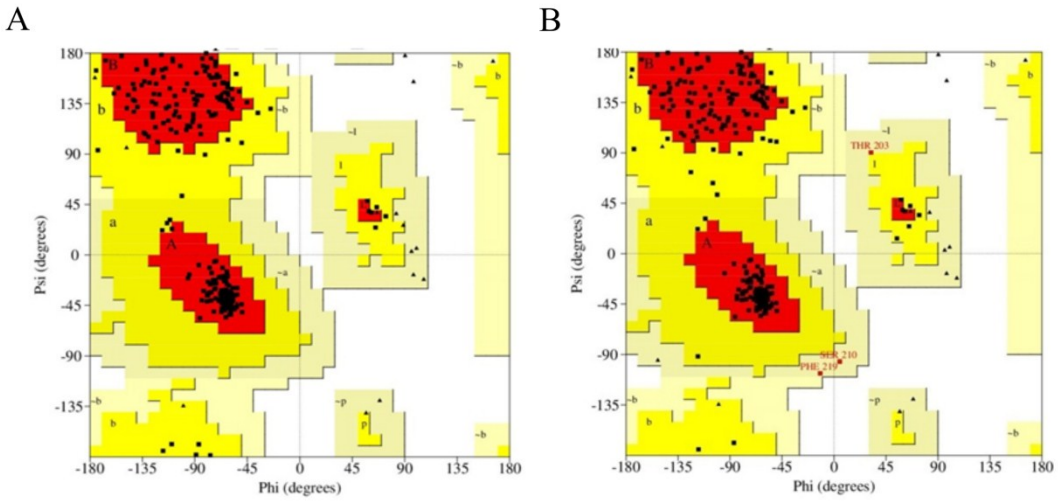

Fig. S2 Ramachandran plot of homolgy A) *apo*- and B) *holo*- DNMT3A model.

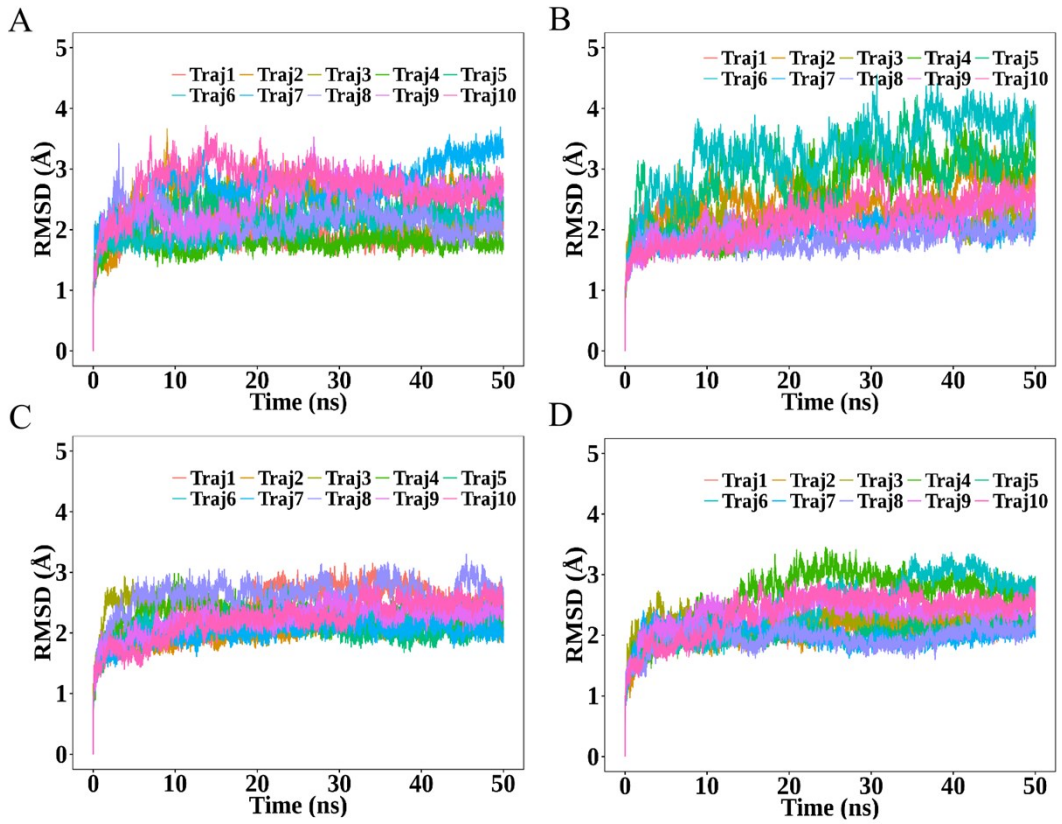

Fig. S3 RMSD values of A) *WT-apo*, B) *Mut-apo*, C) *WT-holo* and D) *Mut-holo* DNMT3A (homology model).

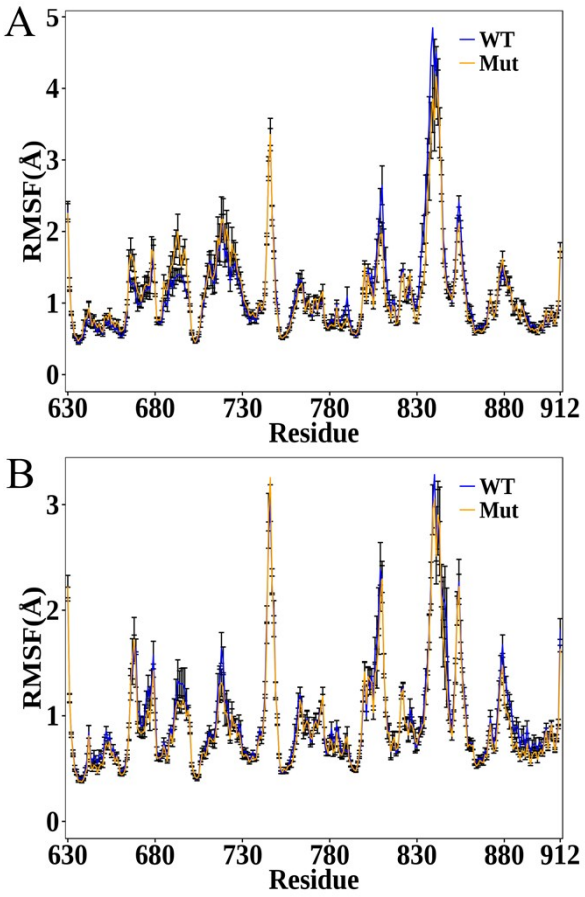

Fig. S4 Fluctuations between homology WT and R882H-mutated of A) *apo*-DNMT3A, B) *holo*-DNMT3A.

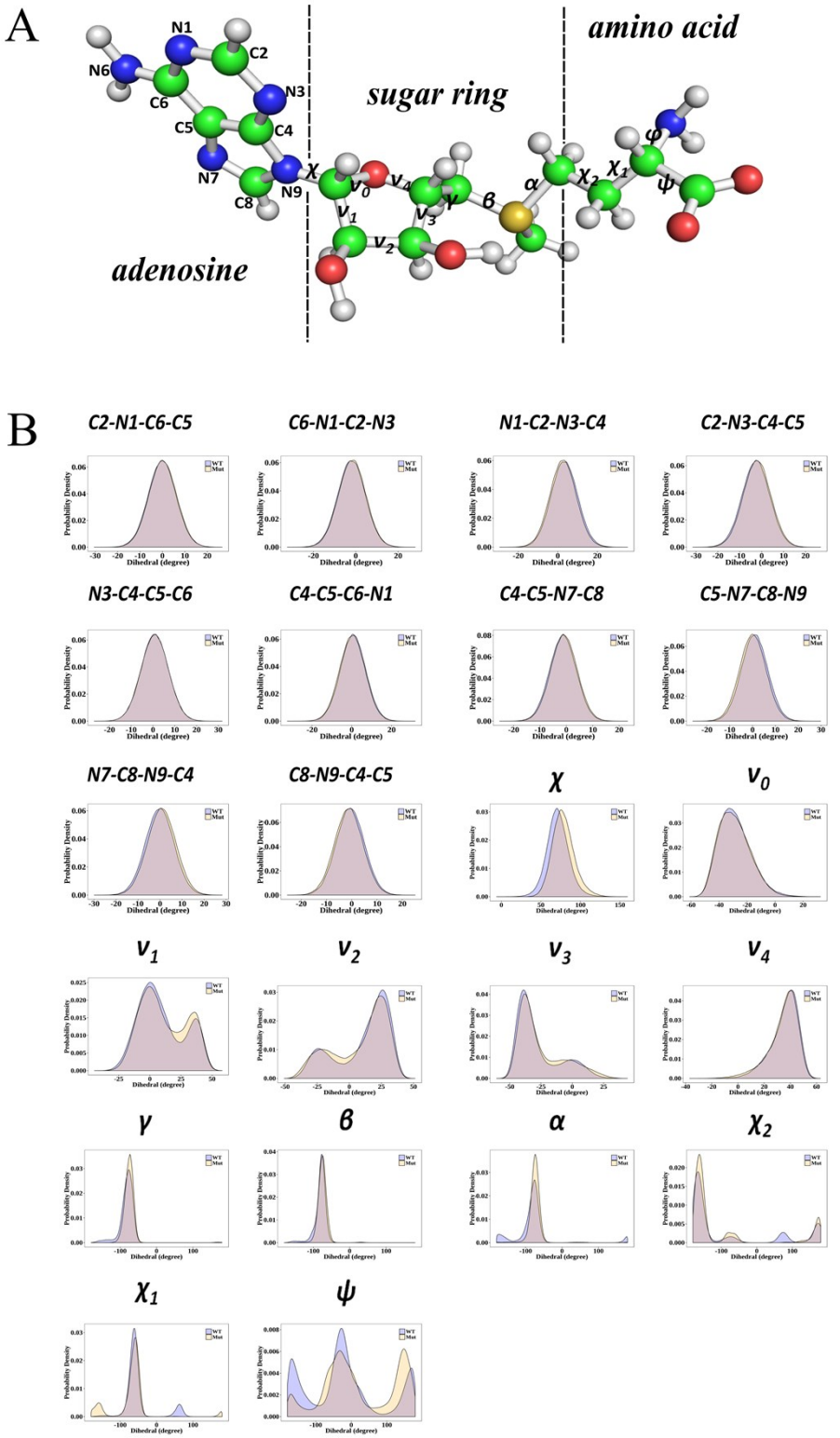

Fig. S5 Conformational analysis of cofactor SAM of homology *holo*-DNMT3As. A) structural partition and dihedral definition and B) population distributions of dihedral angles of the concatenated cMD trajectories.

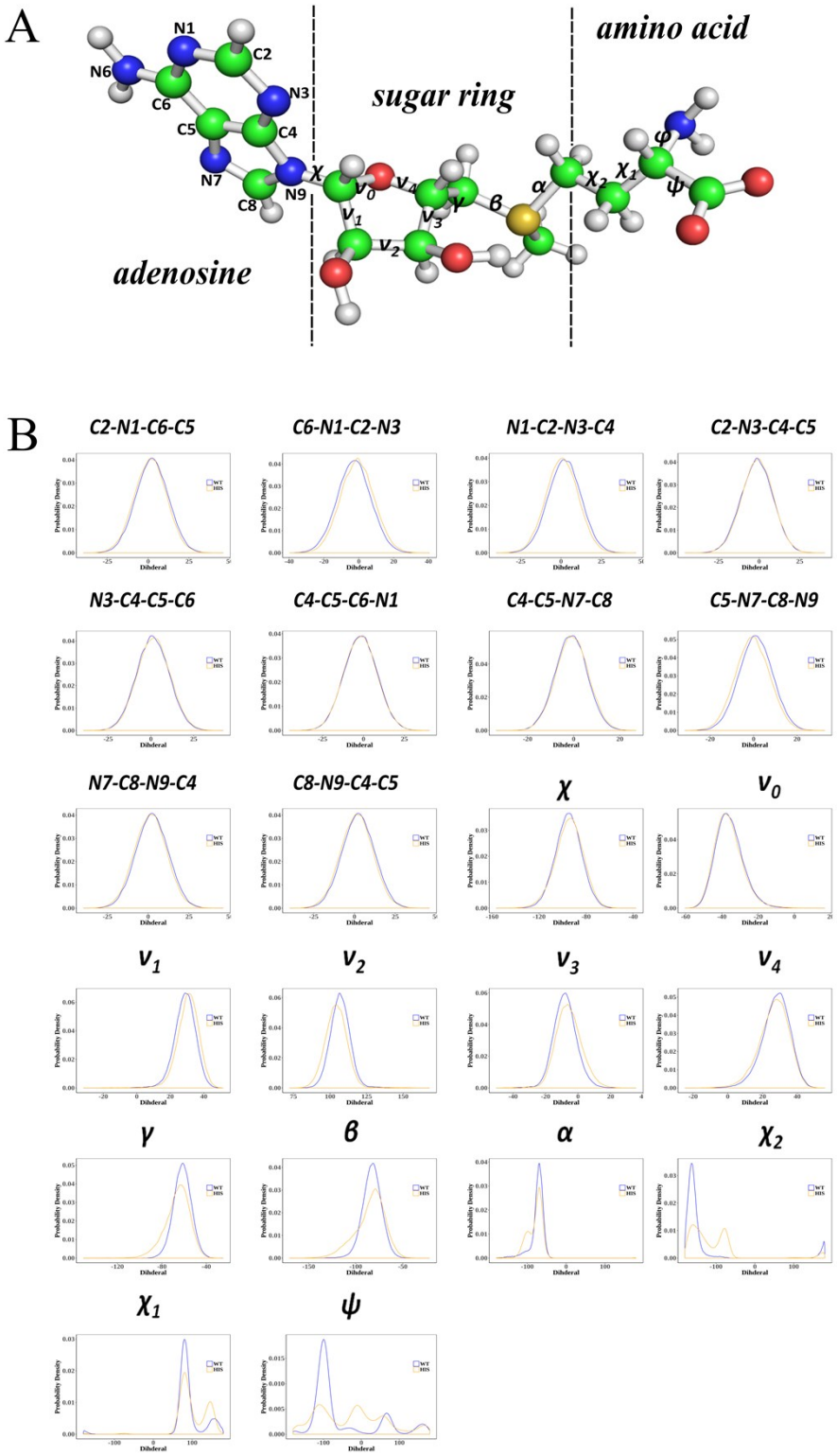

Fig. S6 Conformational analysis of cofactor SAM of crystal *holo*-DNMT3As. A) structural partition and dihedral definition and B) population distributions of dihedral angles of the accelerated aMD trajectories.

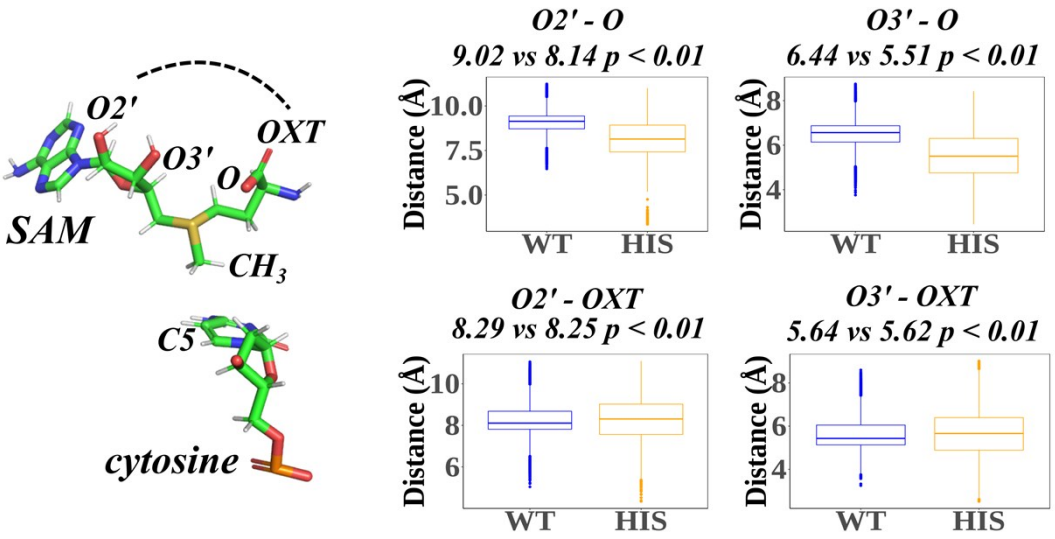

Fig. S7 Statistical comparison of the distance between oxygen atoms of the hydroxyl on the ribose ring and the oxygen atoms on the amino acid terminal portion of SAM in crystal DNMT3A *holo*-models.

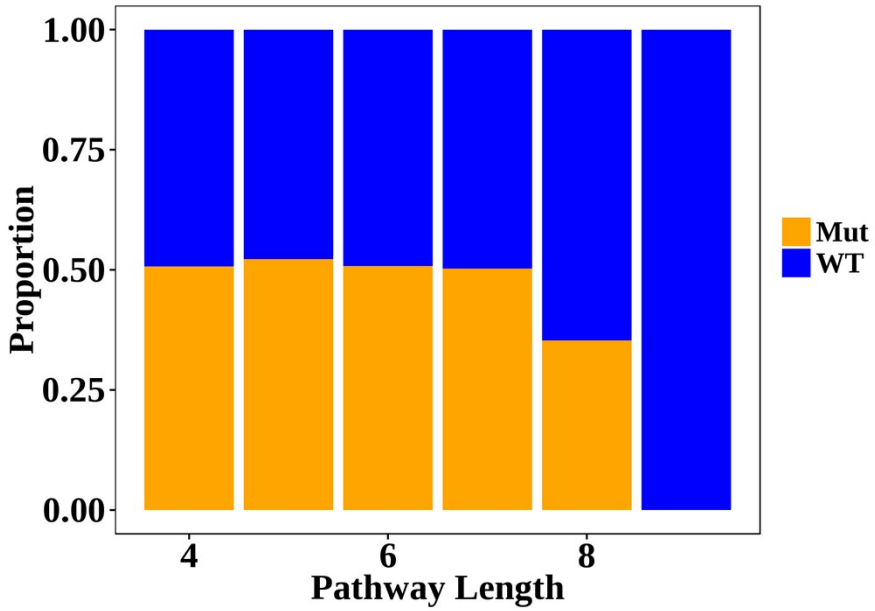

Fig. S8 Pathway length distribution from mutational site Arg882/His882 to Arg891 in homology *holo*-models.

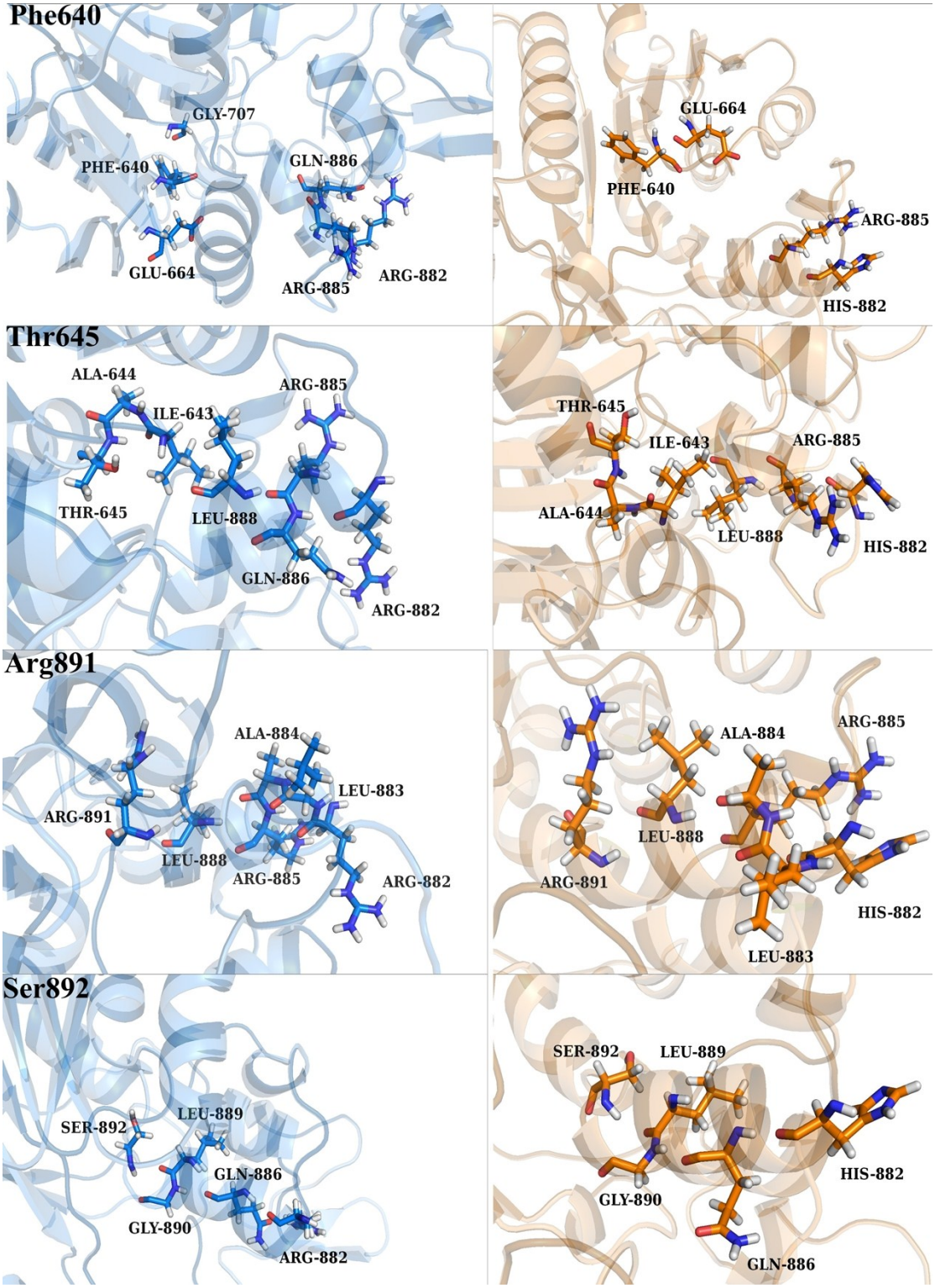

Fig. S9 Most probable pathway from the mutational site Arg/His882 to Phe640, Thr645, Arg891, and Ser892 for homology *apo*-DNMT3A.

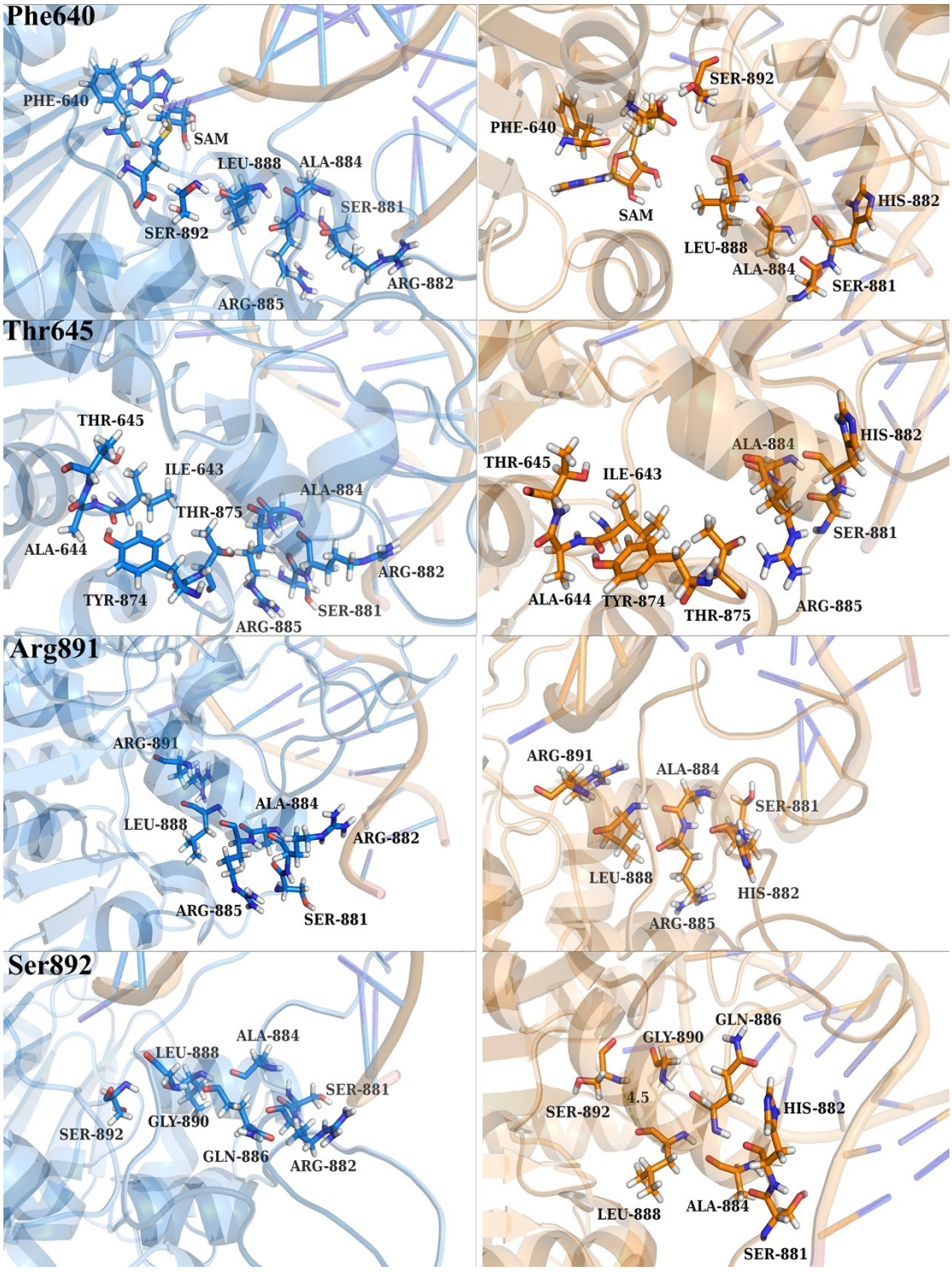

Fig. S10 Most probable pathway from the mutational site Arg/His882 to Phe640, Thr645, Arg891, and Ser892 for homology *holo*-DNMT3A.

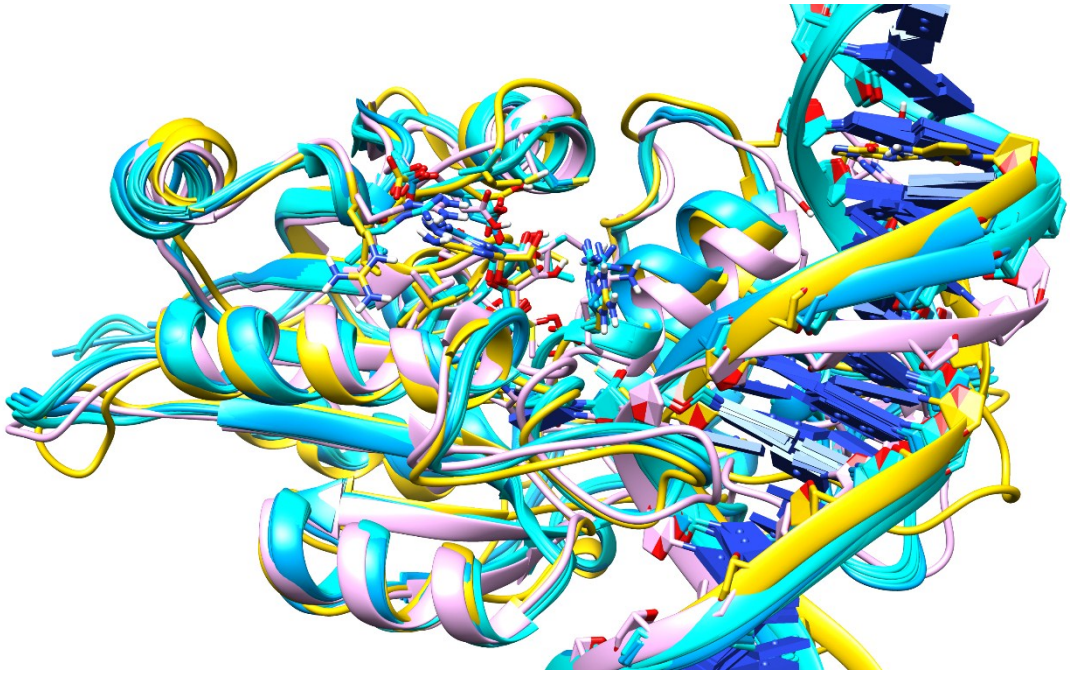

Fig. S11. The structural alignment of the crystal structures (PDB#: 5YX2, 6BRR, and 6F57), and the calculated least-RMSD snapshot to the average structure. Blue: the six holo complexes from the crystal structures, copper: the holo-model of WT, purple: the holo-model of R882H.

Table S1 Binding free energies between DNMT3A and cofactor SAM of multiple cMD trajectories (Unit = kcal/mol)

WT-*holo*													
Energy	Traj1	Traj2	Traj3	Traj4	Traj5	Traj6	Traj7	Traj8	Traj9	Traj10	Mean	SD	SE
Evdw	-44.60	-53.86	-50.55	-49.69	-42.01	-52.43	-49.34	-47.49	-50.12	-49.41	-48.95	3.50	1.11
Eele	-69.96	-66.84	-65.12	-73.76	-73.16	-60.63	-73.22	-55.72	-64.00	-57.03	-65.94	6.66	2.11
Egb	77.47	69.36	76.66	82.11	77.92	65.52	84.07	71.90	72.82	71.35	74.92	5.77	1.82
Esurf	-5.35	-6.29	-6.73	-6.24	-5.05	-6.13	-6.31	-5.70	-6.41	-6.20	-6.04	0.52	0.16
ΔH	-42.44	-57.64	-50.75	-47.58	-42.30	-53.66	-44.80	-37.00	-47.70	-41.29	-46.52	6.24	1.97
TΔS	-26.97	-30.08	-28.97	-29.93	-27.40	-28.37	-33.96	-26.67	-27.88	-29.46	-28.97	2.12	0.67
ΔG	-15.47	-27.56	-21.78	-17.65	-14.84	-25.29	-10.84	-10.36	-19.82	-11.82	-17.54	6.00	1.90
Mut-*holo*													
Energy	Traj1	Traj2	Traj3	Traj4	Traj5	Traj6	Traj7	Traj8	Traj9	Traj10	Mean	SD	SE
Evdw	-50.09	-48.06	-46.72	-50.01	-51.45	-52.66	-50.92	-53.65	-52.48	-54.17	-51.02	2.38	0.75
Eele	-70.41	-51.78	-62.60	-63.99	-66.96	-64.67	-62.36	-41.40	-75.57	-74.94	-63.47	10.36	3.28
Egb	85.49	67.43	80.93	77.02	73.71	74.00	66.02	61.31	87.27	86.77	76.00	9.17	2.90
Esurf	-6.27	-5.54	-5.27	-6.32	-6.34	-6.20	-5.99	-6.56	-6.27	-6.28	-6.10	0.40	0.13
ΔH	-41.28	-37.94	-33.67	-43.30	-51.03	-49.54	-53.24	-40.30	-47.05	-48.62	-44.60	6.30	1.99
TΔS	-27.36	-25.82	-25.67	-30.10	-32.31	-29.19	-25.57	-24.69	-28.56	-29.51	-27.88	2.46	0.78
ΔG	-13.91	-12.13	-8.00	-13.20	-18.72	-20.35	-27.67	-15.61	-18.49	-19.12	-16.72	5.42	1.71

Table S2 Amino acid contributions of DNMT3A-SAM binding free energy (Unit=kcal/mol)

Residue	WT-*holo*	Mut-*holo*
SAM	-16.7±3.2	-15.1±3.1
Phe640	-4.0±0.6	-4.1±0.5
Thr645	-1.3±1.2	-0.9±0.7
Glu664	-1.5±0.6	-1.5±0.3
Val665	-2.9±0.6	-2.7±0.5
Asp686	-3.8±0.7	-4.5±0.8
Arg891	-0.9±2.0	-0.8±0.8
Ser892	-2.1±1.4	-2.5±0.5
Trp893	-2.8±0.9	-2.9±1.4

Table S3 Hydrogen bonding occupancy between residues of DNMT3A and SAM in all cMD and aMD trajectories

Acceptor	Donor	WT-*holo*	R882H-*holo*
Phe640@O	SAM@N	11.4%±3.4%	10.3%±3.6%
SAM@OXT	Thr645@$O_{\gamma 1}$	34.8%±6.4%	27.2%±8.7%
Asp686@$O_{\delta 1}$	SAM@N6	39.2%±6.9%	12.1%±5.7%
Asp686@$O_{\delta 2}$	SAM@N6	15.3%±6.0%	41.1%±6.3%
SAM@O3'	Arg891@ N_{ε}	10.9%±7.4%	5.7%±3.0%
SAM@O3'	Arg891@$N_{\eta 2}$	7.7%±5.5%	2.5%±1.8%
SAM@O3'	Arg891@$N_{\eta 1}$	0.5%±0.4%	0.8%±0.5%
SAM@O2'	Arg891@N_{ε}	0.2%±0.2%	0.1%±0.1%
SAM@O2'	Arg891@$N_{\eta 1}$	0.5%±0.3%	0.0%±0.0%
SAM@O2'	Arg891@$N_{\eta 2}$	7.2%±6.6%	2.9%±2.7%
SAM@OXT	Arg891@N_{ε}	1.8%±1.8%	5.3%±3.2%
SAM@OXT	Arg891@$N_{\eta 1}$	1.7%±1.0%	7.0%±5.4%

SAM@OXT	Arg891@$N_{\eta 2}$	0.4%±0.4%	0.5%±0.4%
SAM@O	Arg891@N_{ε}	2.4%±2.1%	7.9%±5.1%
SAM@O	Arg891@$N_{\eta 1}$	0.9%±0.9%	1.3%±1.2%
SAM@O	Arg891@$N_{\eta 2}$	0.0%	1.1%±1.1%
SAM@O	Ser892@O_{γ}	48.7%±9.9%	61.2%±10.9%
SAM@OXT	Ser892@O_{γ}	46.5%±8.6%	53.6%±9.9%
SAM@N	Trp893@$N_{\varepsilon 1}$	11.5%±4.3%	5.9%±2.3%
SAM@O	Trp893@$N_{\varepsilon 1}$	7.9%±5.3%	0.3%±0.2%
SAM@OXT	Trp893@$N_{\varepsilon 1}$	7.7%±5.5%	0.4%±0.3%
SAM@N	Trp893@N	1.8%±1.6%	12.9%±8.2%
SAM@O	Trp893@N	24.3%±7.0%	20.2%±8.4%
SAM@OXT	Trp893@N	14.2%±5.3%	26.0%±10.3%

(SAM@OXT represents the carboxyl oxygen, and SAM@O represents the carbonyl oxygen of the amino acid terminal of SAM; R882-*holo* is the representation of combining both Mut-*holo* of the homology model and HIS-*holo* of the crystal model)

Table S4 Energy barriers calculated by the different combination of functional and basis set. (Unit = kcal/mol)

Basis set	M06-2X		ωB97x-D	
	WT-*holo*	Mut-*holo*	WT-*holo*	Mut-*holo*
6–31G(d)	16.0	19.1	18.8	22.5
6–311G(d)	17.4	19.8	19.8	23.0
6–311+G(d)	18.7	20.1	21.2	23.3

Table S5 Energies of different layers on PRS and TS compared to the reference (Unit = kcal/mol)

	Layers	PRS	TS
WT-*holo*	QM layer	-15.1	-4.5
	MM layer	2.7	7.0
Mut-*holo*	QM layer	-20.0	1.0
	MM layer	-2.3	-7.7

Table S6 Population comparison of near attack conformers (NAC) parameters

	homology		***crystal***	
Parameters	WT-*holo*	Mut-*holo*	WT-*holo*	HIS-*holo*
C4-C5-CH_3	33.1%	12.5%	56.5%	8.9%
C6-C5-CH_3	63.4%	80.0%	75.2%	75.5%
Both angles	23.8%	8.9%	44.5%	7.4%
O4'-C1'-N1-C2	60.4%	61.1%	75.8%	70.4%

(Both angles represent those conformers that satisfy criteria of both C4-C5-CH_3 and C6-C5-CH_3)

Table S7 Center of mass distance between sidechains along the residues consisting of quasi-"Newton's cradle" model on H11 helix (Unit = angstrom)

	DNMT3A	882-883	883-887	887-891
homology	WT-*holo*	8.0±1.2	5.8±1.8	7.8±2.1
	Mut-*holo*	7.2±0.9	5.2±1.2	8.2±1.7
crystal	WT-*holo*	5.5±0.5	6.0±0.8	6.4±0.8
	HIS-*holo*	5.6±0.4	5.5±0.8	7.2±0.7