Supporting Information

Computational Study of an Acyltransferase Domain from a Modular Polyketide Synthase

Huining Ji,[a] Ting Shi,[a] Lei Liu,[a] Fa Zhang,[a] Wentao Tao,[a] Qing Min[c], Zixin Deng,[a] Linquan Bai,[a] Yilei Zhao*[a] and Jianting Zheng*[a,b]

[a] State Key Laboratory of Microbial Metabolism, and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

[b] Joint International Research Laboratory of Metabolic & Developmental Sciences, Shanghai Jiao Tong University, Shanghai, China

[c] Pharmacy School, Hubei University of Science and Technology, Hubei, Xianning 437100, China

*Correspondence: Yilei Zhao, E-mail: yileizhao@sjtu.edu.cn, Jianting Zheng, E-mail: jtzheng@sjtu.edu.cn

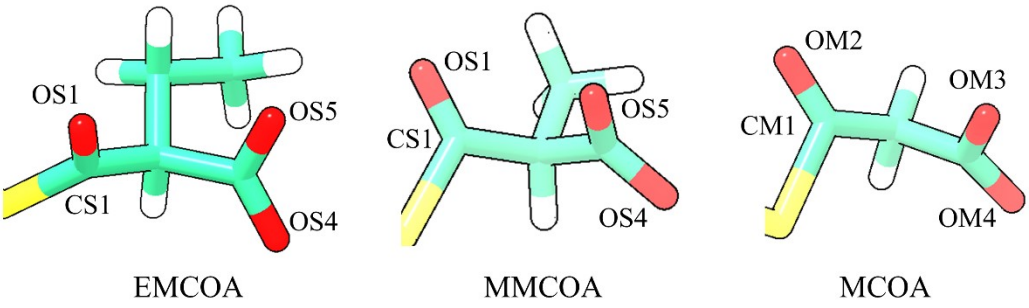

Figure S1. Names of the key atoms of the substrate acyl groups.

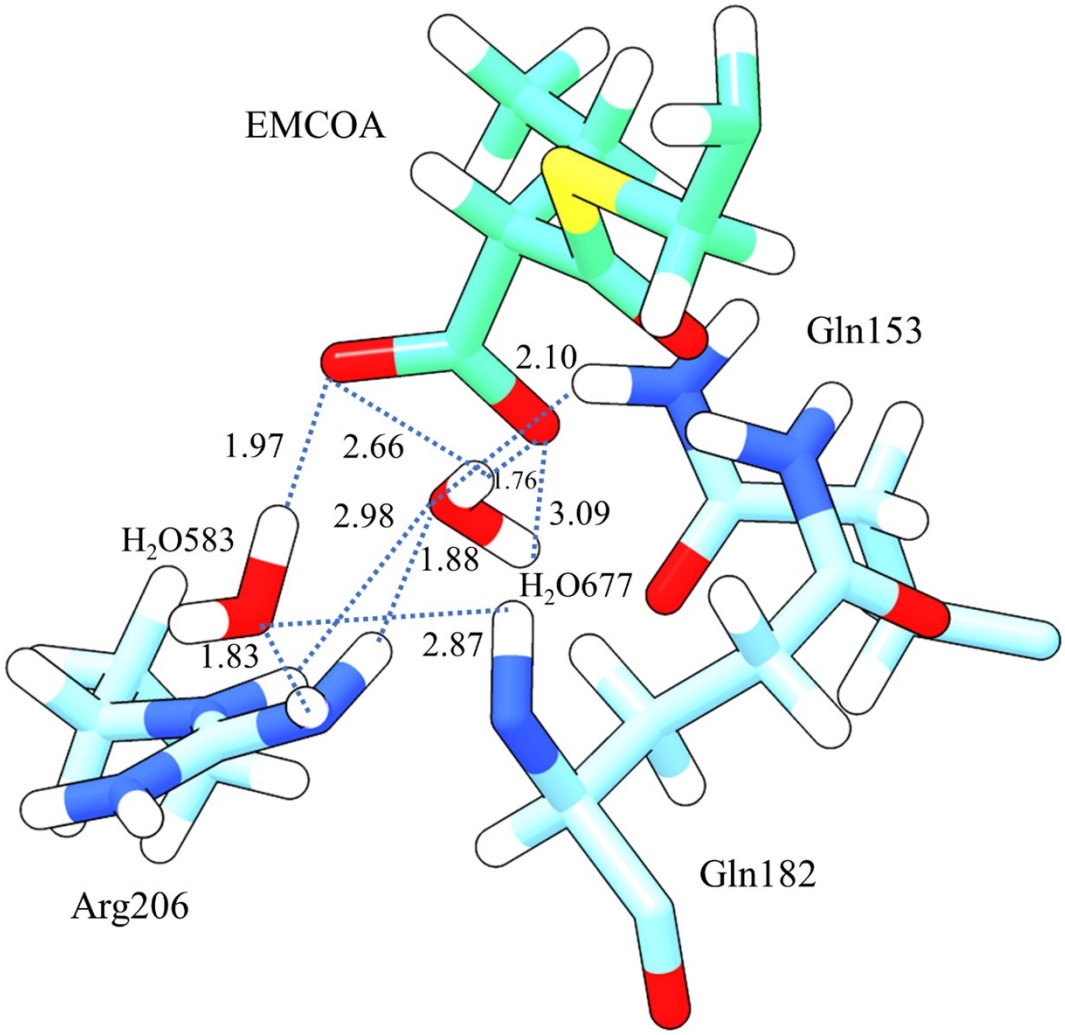

Figure S2. The hydrogen-bond network of H_2O583 and H_2O677 which bound to Gln153, Gln182 and Arg206 stabilized the substrate EMCOA in the active site of the represent structure. The distances are given in Å.

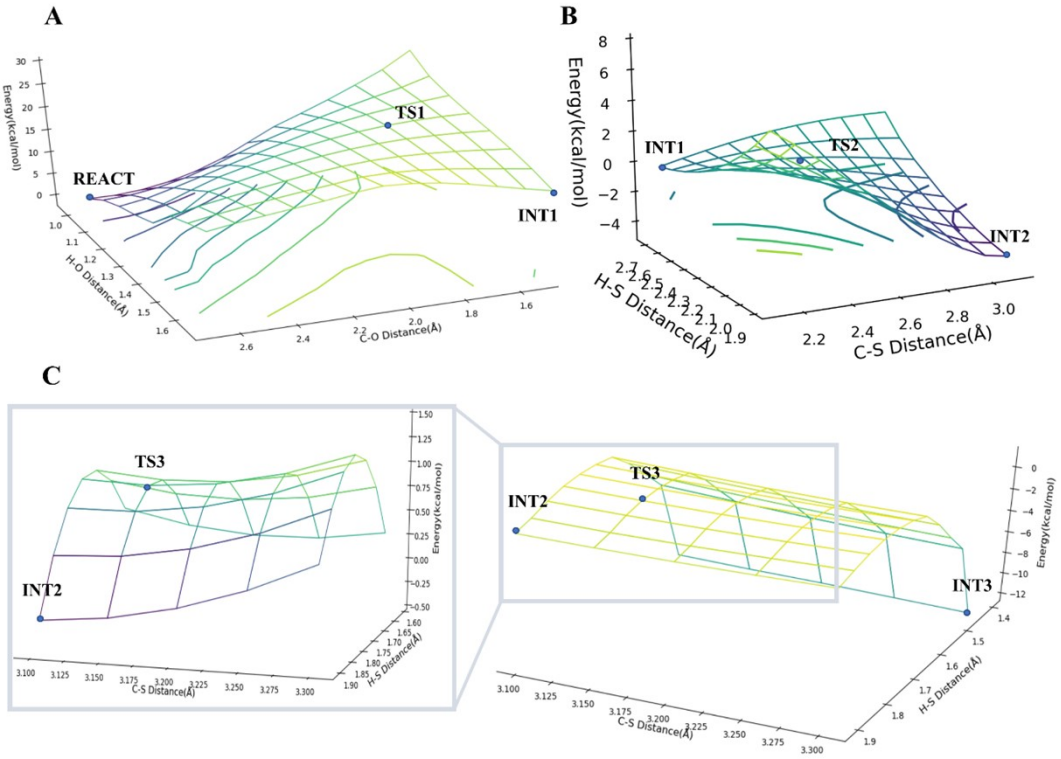

Figure S3. Two-dimensional potential energy surfaces (PESs) in the SalAT14-EMCOA system on the first half of transacylation: (A) the nucleophilic attack on the thioester carbonyl group of the substrate EMCOA; (B) the breakdown of the tetrahedral intermediate; (C) the formation of the intermediate3, the detail of the TS3 is inserted into the left panel. The energy is in kcal/mol, and the distances are all in Å. Besides, in the absence of the enzyme, the TS1 activation barrier for EMCOA is 40.45 kcal/mol, for MMCOA is 42.81 kcal/mol, and for MCOA is 41.78 kcal/mol. The high energy barriers were mainly due to the use of PM7.

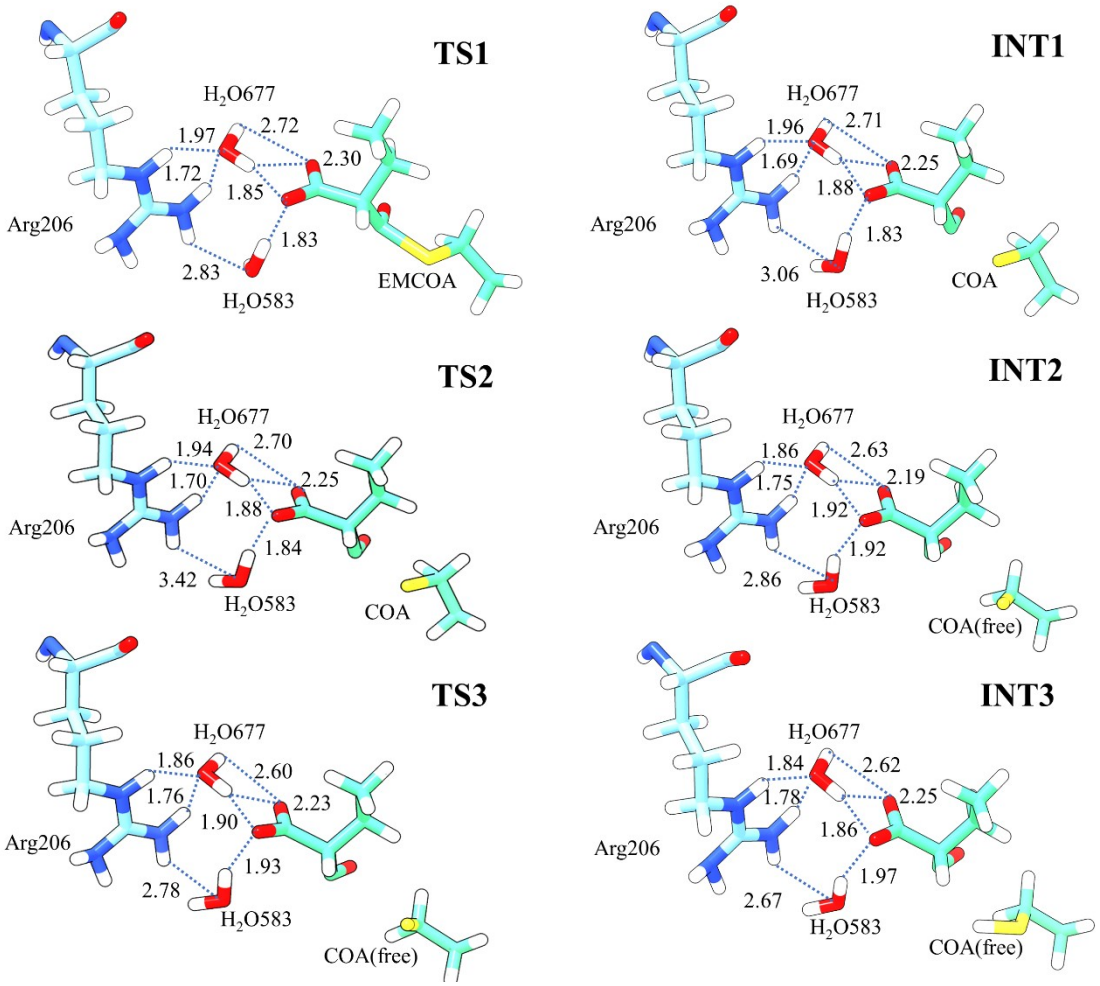

Figure S4. The indirect effects of Arg206 on the EMCOA through H_2O583 and H_2O677 during TS1, INT1, TS2, INT2, TS3 and INT3. The distances are given in Å.

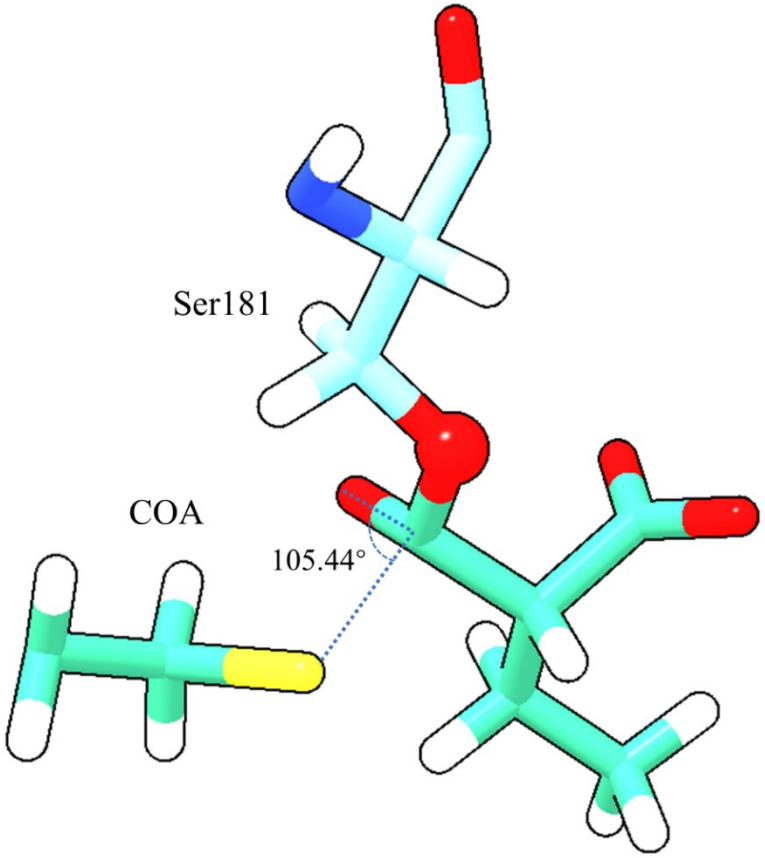

Figure S5. The angle of S-C-O in TS2.

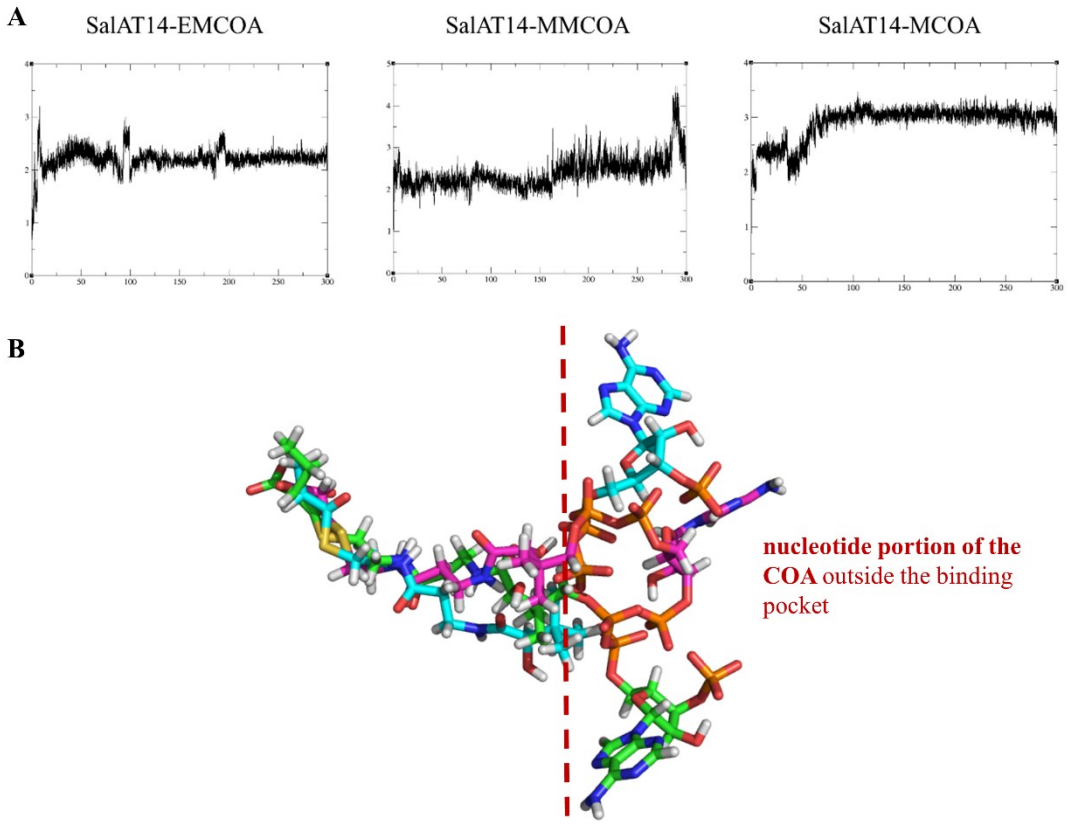

Figure S6. (A) The RMSD values of 300 ns MD simulations of the binding pockets (residues within 4 Å distance of the substrate and the substrate parts inside the binding pockets) from SalAT14-EMCOA, SalAT14-MMCOA and SalAT14-MCOA systems; (B) The superposition of three substrates from the representative structures of three systems. The adenosine nucleotide of the COA is exposed to the surrounding water environment and outside the binding pocket.

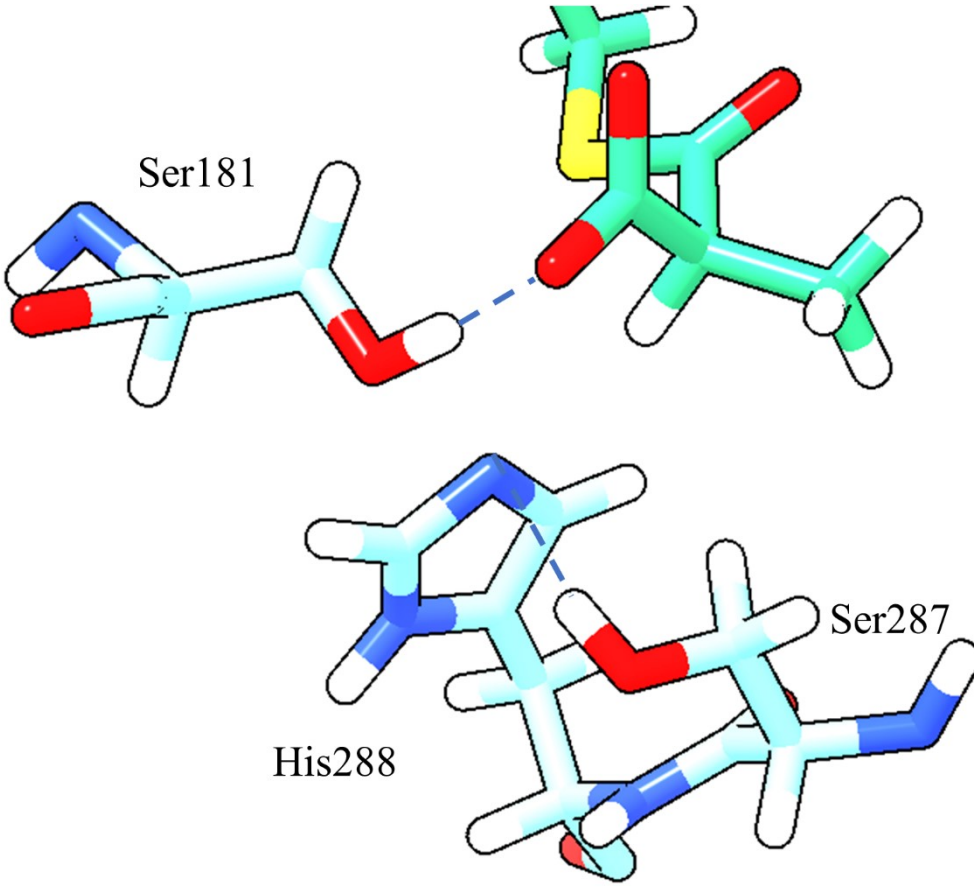

Figure S7. The phenomenon that the Hγ atom of Ser181 interacts with the carboxyl group of MMCOA and the Nε atom of His288 forms a strong hydrogen bond with Ser287. The truncated substrate MMCOA is cyan and the residues are blue.

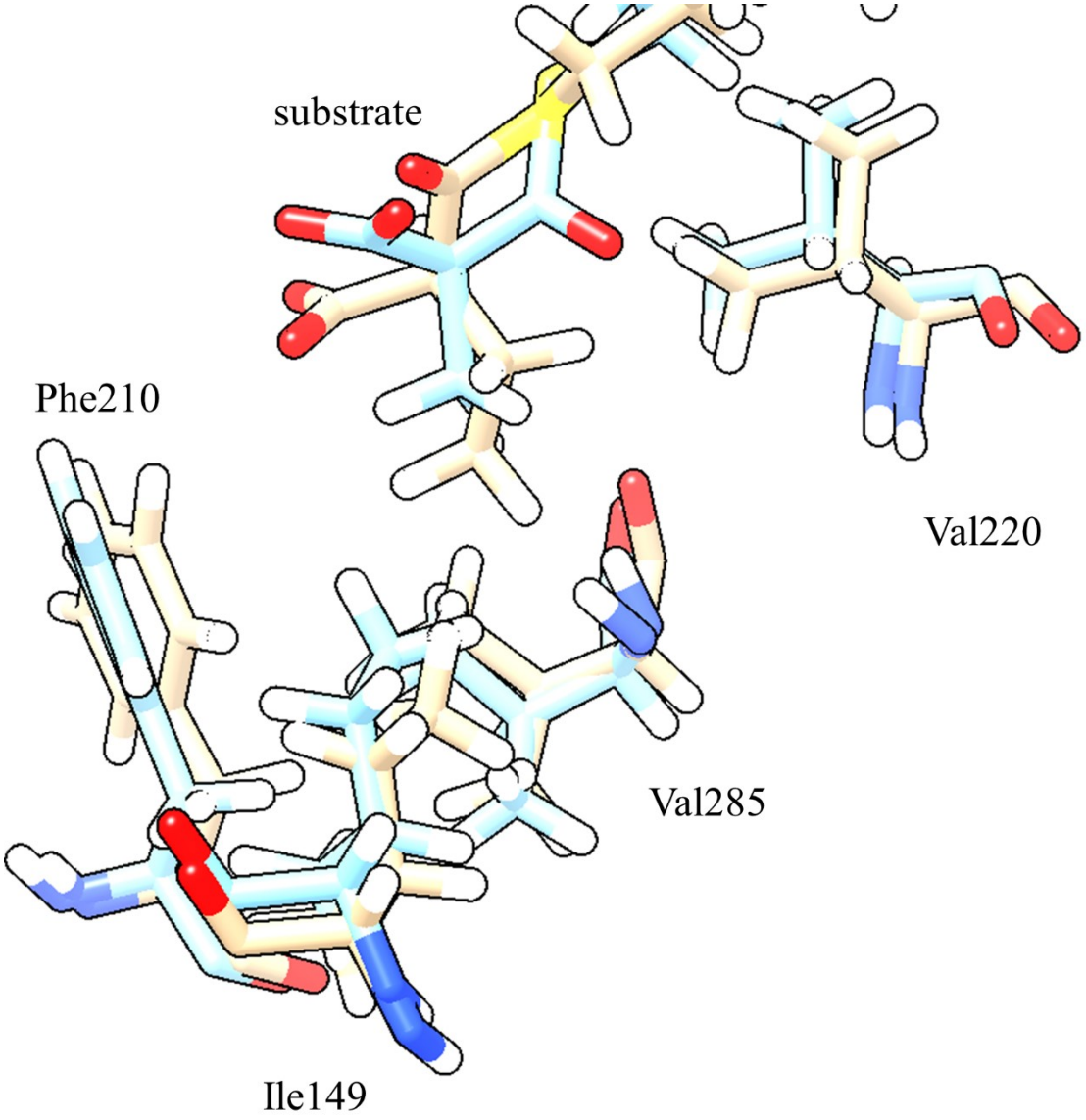

Figure S8. The overlap of the SalAT14-EMCOA (light khaki) and SalAT14-MMCOA (blue) in the hydrophobic pocket of SalAT14.

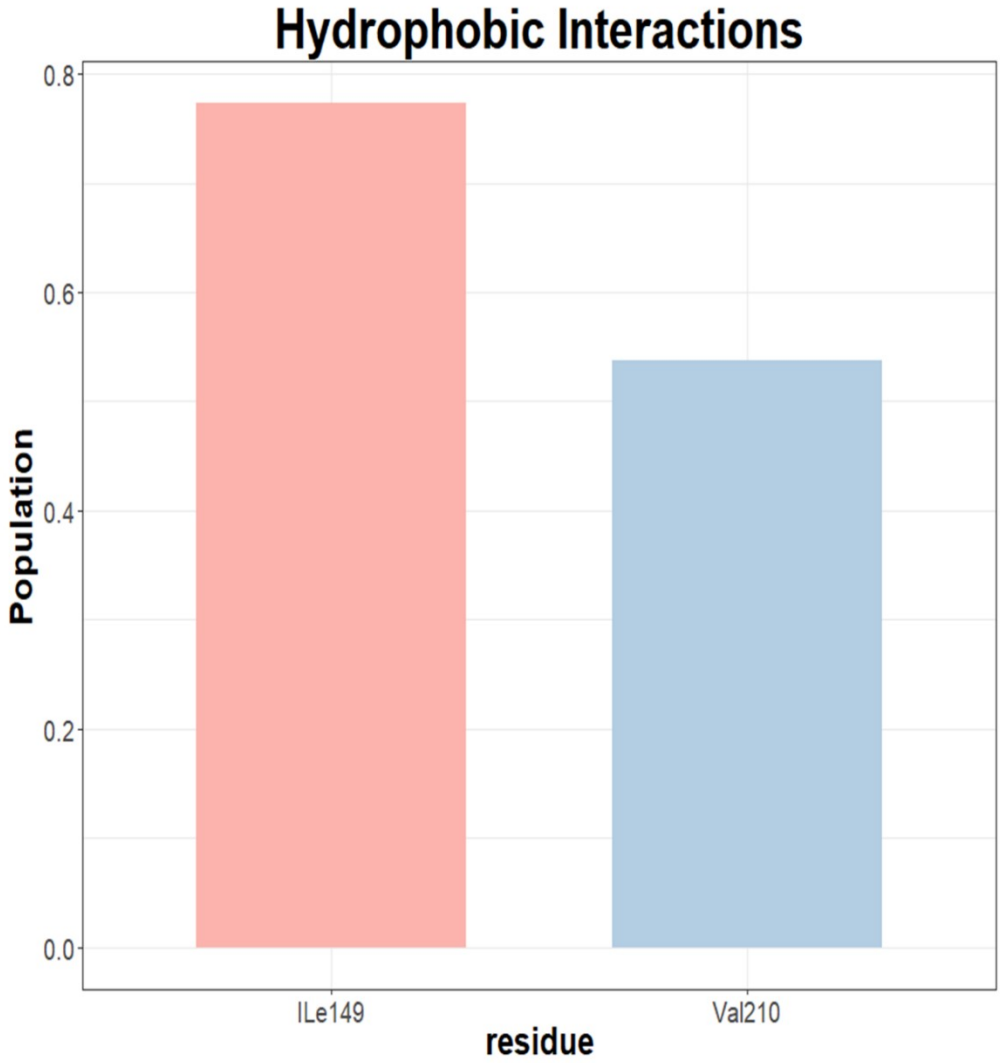

Figure S9. Hydrophobic Interactions between MMCOA and SalAT14(V285Y/F210V/V220M).

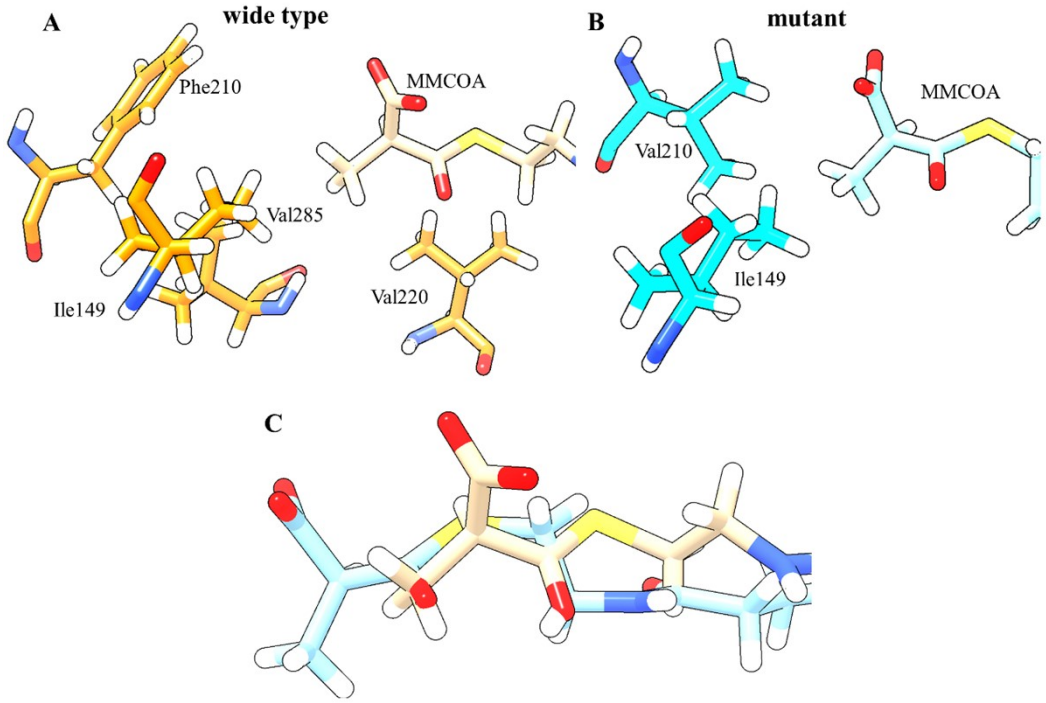

Figure S10. SalAT14-MMCOA (A) and SalAT14(V285Y/F210V/V220M)-MMCOA (B) in the hydrophobic pockets; (C) the overlap of the MMCOA in wide type SalAT14 (light khaki) and SalAT14(V285Y/F210V/V220M) (blue).

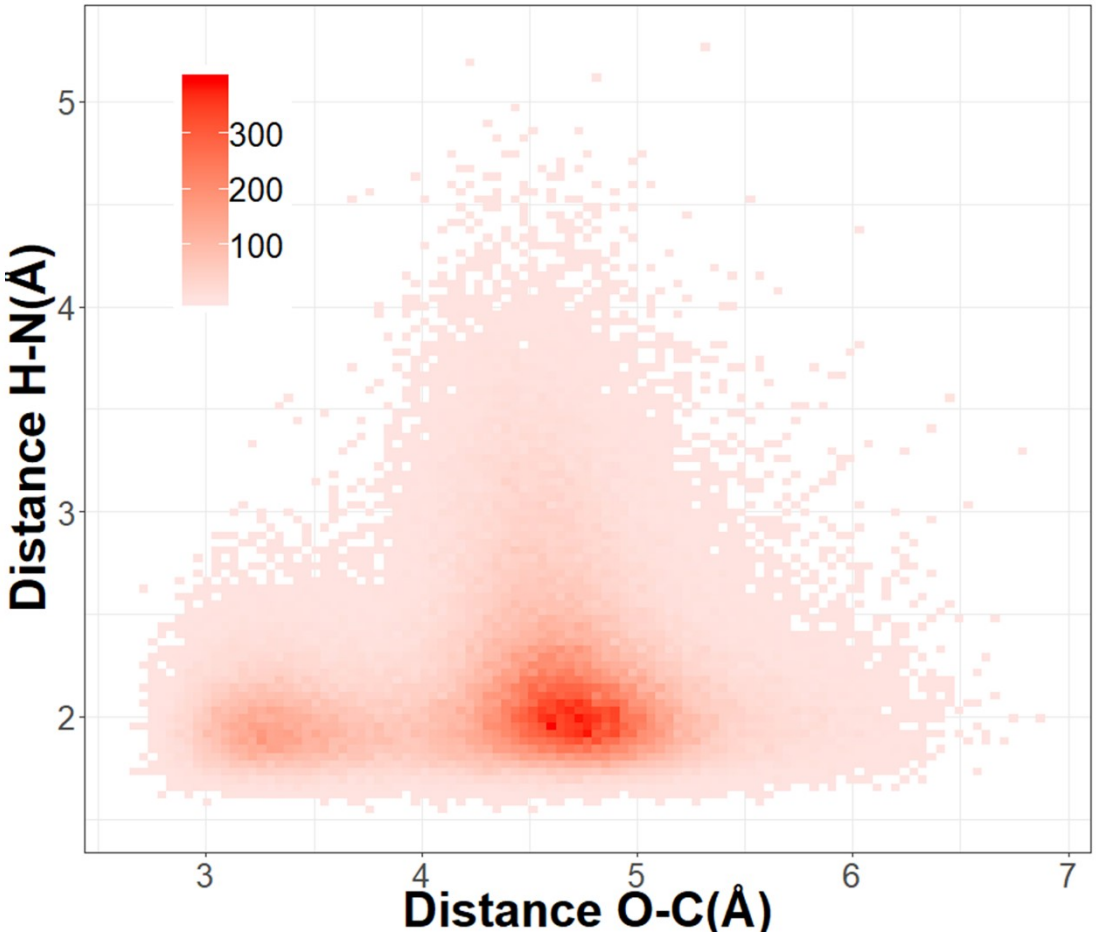

Figure S11. Conformer populations with d(N_{ε}−HG) and d(CS1−OG) distances obtained from 300ns MD simulation in SalAT14(V285Y/F210V/V220M)-MMCOA.

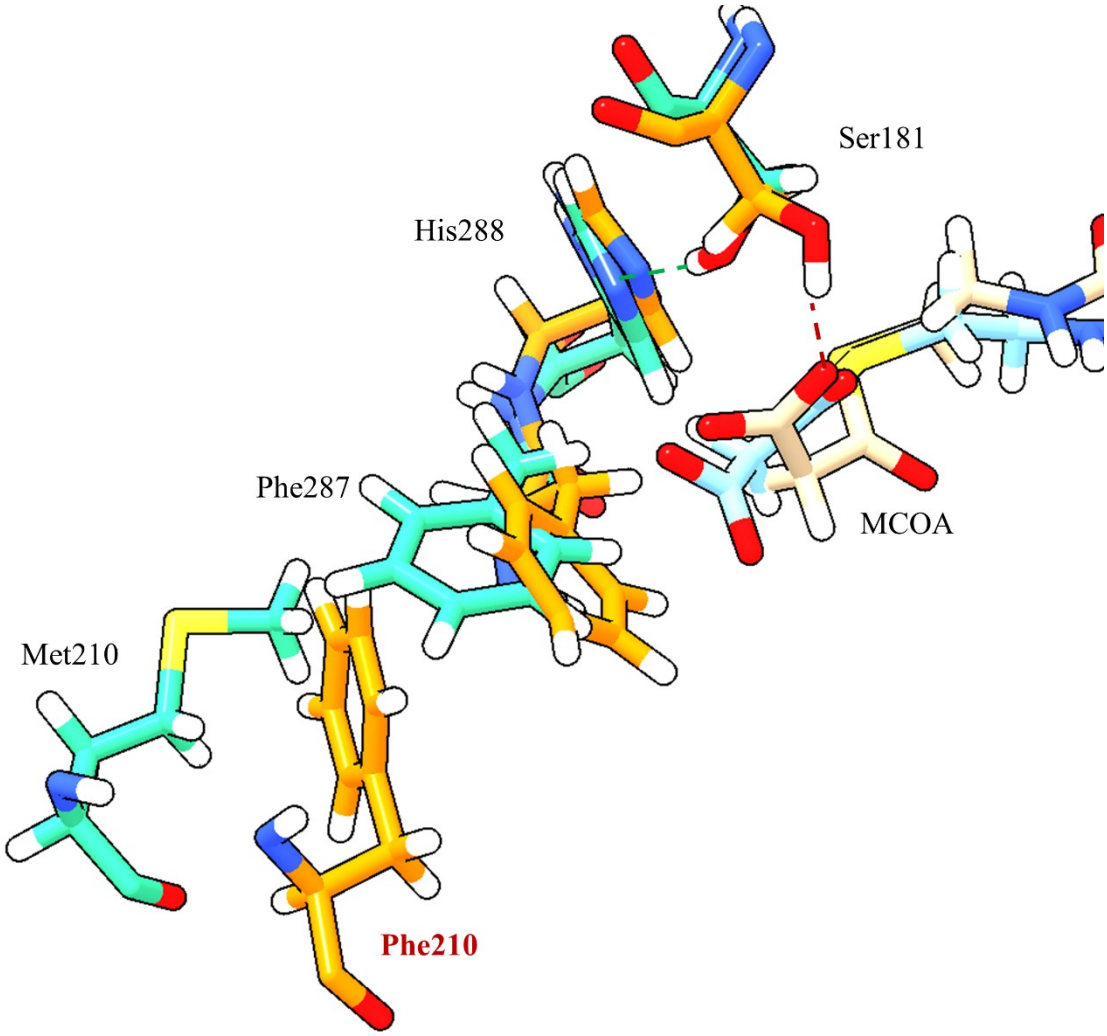

Figure S12. The overlap of SalAT14(V285H/S287F/Q182I)-MCOA (light khaki for MCOA and orange for residues) and SalAT14(V285H/S287F/Q182I/F210M)-MCOA (blue for MCOA and cyan for residues). In SalAT14(V285H/S287F/Q182I)-MCOA, the Hγ atom of Ser181 interacts with the carboxyl group of MCOA. In SalAT14(V285H/S287F/Q182I/F210M)-MCOA, the Hγ atom of Ser181 interacts with the residue His288.

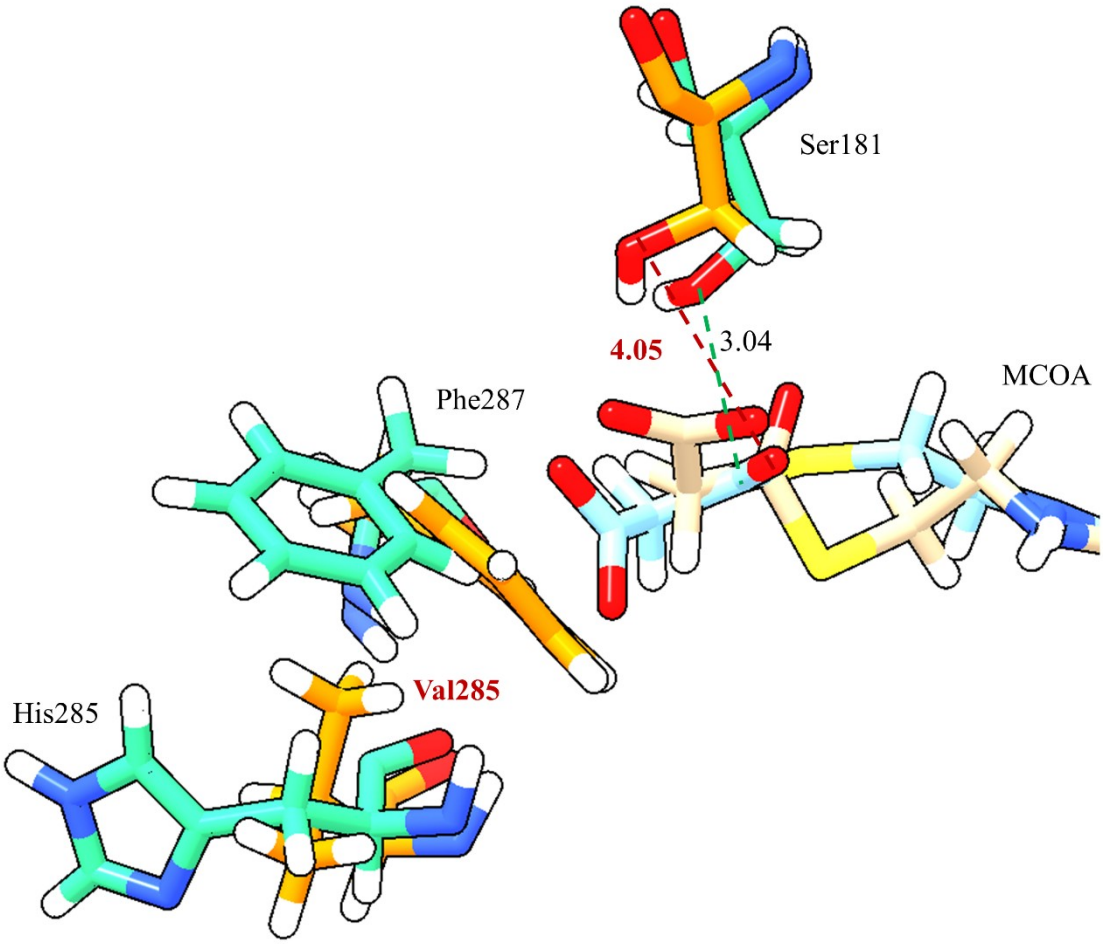

Figure S13. The overlap of representative structures of SalAT14(S287F/Q182I/F210M)-MCOA (light khaki for MCOA and orange for residues) and SalAT14(V285H/S287F/Q182I/F210M)-MCOA (blue for MCOA and cyan for residues). In SalAT14(S287F/Q182I/F210M)-MCOA, the distance between the Oγ atom of Ser181 and the carbonyl carbon of the MCOA is much longer than that in SalAT14(V285H/S287F/Q182I/F210M)-MCOA. The distances are given in Å.

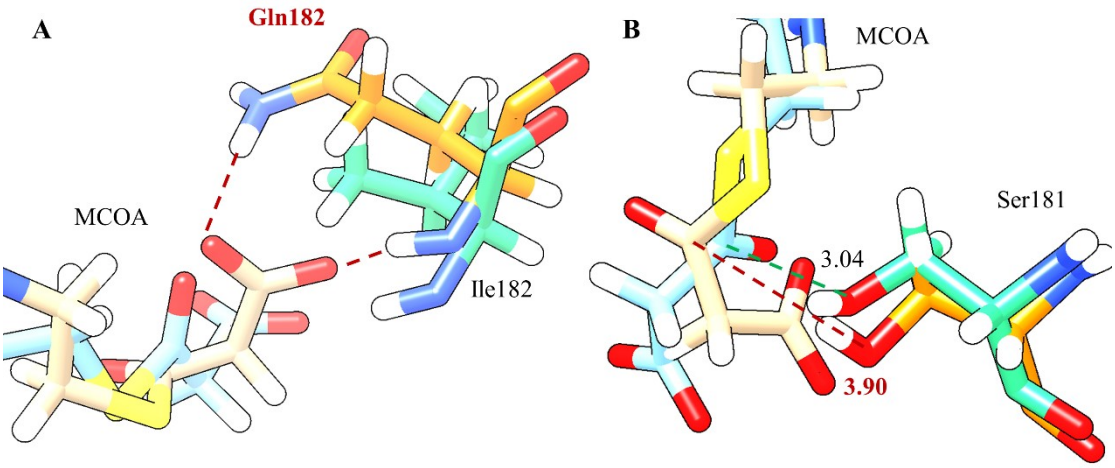

Figure S14. The overlap of representative structures of SalAT14(V285H/S287F/F210M)-MCOA (light khaki for MCOA and orange for residues) and SalAT14(V285H/S287F/Q182I/F210M)-MCOA (blue for MCOA and cyan for residues): (A) in SalAT14(V285H/S287F/F210M)-MCOA, the Gln182 interacts with the carboxyl group of MCOA; (B) in SalAT14(V285H/S287F/F210M)-MCOA, the distance between the Oγ atom of Ser181 and the carbonyl carbon of the MCOA is much longer than that in SalAT14(V285H/S287F/Q182I/F210M)-MCOA. The distances are given in Å.

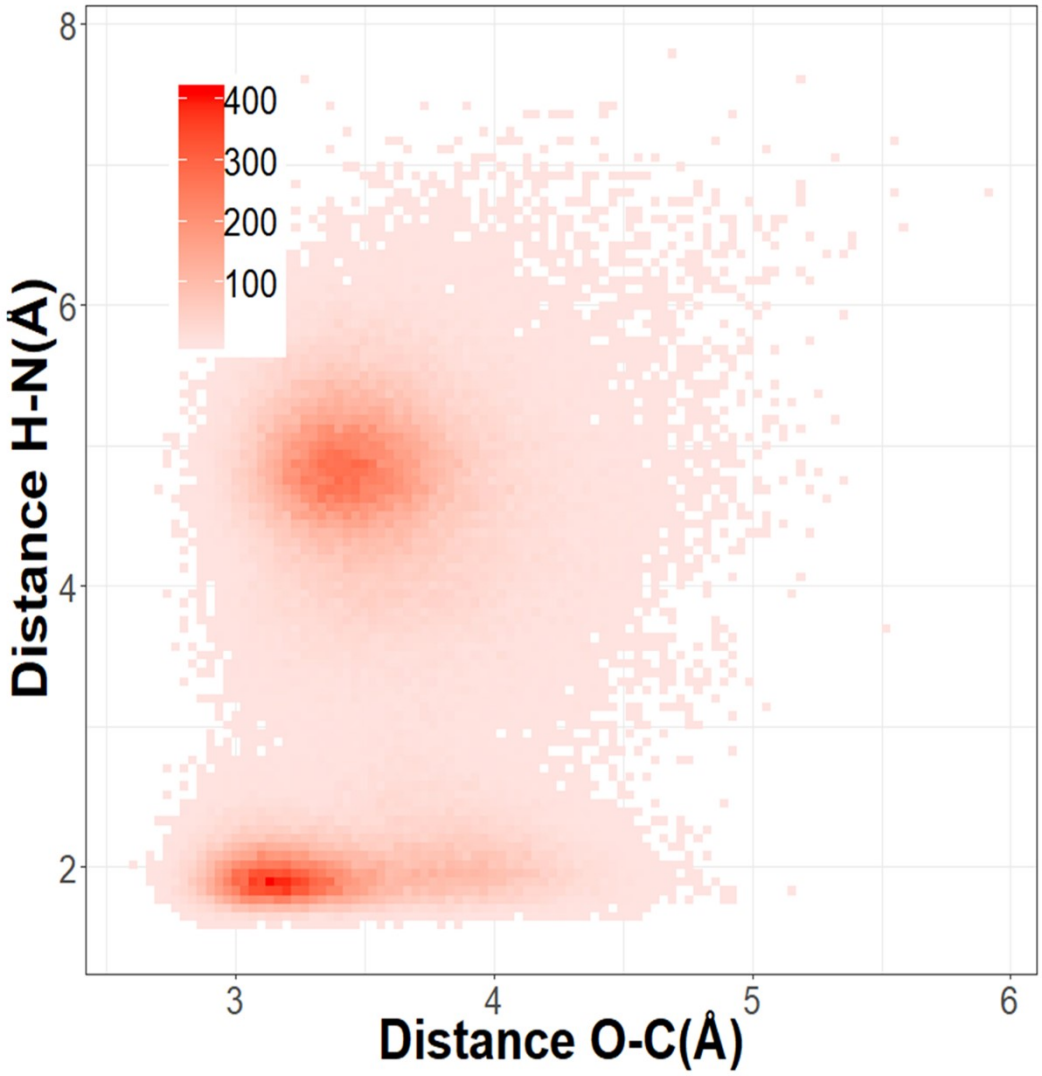

Figure S15. Conformer populations with d(N_{ε}−HG) and d(CM1−OG) distances obtained from 300ns MD simulation in SalAT14(V285H/S287F/Q182I/F210M)-MCOA.

Table S1. Statistics associated with the hydrogen bonds between the substrates and the protein residues for the 300ns MD simulations of the SalAT14-EMCOA system, the SalAT14-MMCOA system and the SalAT14-MCOA system.

Acceptor	Donor	Percentage	Acceptor	Donor	Percentage	Acceptor	Donor	Percentage
OS1 of EMCOA	Gln96H	83.11%	OS1 of MMCOA	Gln96H	41.48%	OM2 of MCOA	Gln96H	19.75%
OS1 of EMCOA	Gln182HE22	57.81%	OS1 of MMCOA	Gln182HE22	16.75%	OM2 of MCOA	Gln182HE22	12.74%
OS4 of EMCOA	Gln182HE22	83.07%	OS4 of MMCOA	Gln96H	47.15%	OM3 of MCOA	Gln182HE22	19.70%
OS5 of EMCOA	Ser287HG	65.68%	OS4 of MMCOA	Gln182HE22	80.00%	OM4 of MCOA	Ser287HG	82.57%
OS4 of EMCOA	Gln153HE21	55.92%	OS5 of MMCOA	Ser287HG	58.39%	OM3 of MCOA	Gln153HE21	76.83%
OS5 of EMCOA	H_2O 583H2	88.81%	OS4 of MMCOA	Gln153HE21	24.24%	OM4 of MCOA	H_2O583H2	26.47%
OS5 of EMCOA	H_2O 677H1	21.27%	OS5 of MMCOA	H_2O 583H2	46.84%	OM4 of MCOA	H_2O677H1	58.70%
H_2O 677O	Arg206HE	93.58%	OS5 of MMCOA	H_2O 677H1	24.02%	H_2O677O	Arg206HE	20.85%
H_2O 677O	Arg206HH21	97.67%	OS5 of MMCOA	Ser181HG	79.91%	H_2O677O	Arg206HH21	42.41%
H_2O 583O	Gln182H	42.35%	His288NE2	Ser287HG	80.74%	H_2O583O	Gln182H	14.12%
H_2O 583O	H_2O 677H1	67.04%	H_2O 677O	Arg206HH21	32.85%	H_2O583O	H_2O677H1	27.80%
Gln153OE1	H_2O 677H2	80.53%	H_2O 583O	Gln182H	20.51%	Gln153OE1	H_2O677H2	13.50%
Ser181OG	H_2O 583H2	65.68%	H_2O 583O	H_2O 677H1	37.65%	Ser181OG	H_2O583H2	19.34%
H_2O 583O	Arg206HH21	91.64%	Ser181OG	H_2O 583H2	58.39%	H_2O583O	Arg206HH21	83.47%
			H_2O 583O	Arg206HH21	49.86%	OM2 of MCOA	Gln96H	19.75%

Table S2. Statistics associated with the hydrogen bonds between the MMCOA and the protein residues for the 300ns MD simulation of the SalAT14(V285Y/F210V/V220M)-MMCOA system

Acceptor	**Donor**	**Percentage**
OS1 of MMCOA	Gln96H	11.74%
OS1 of MMCOA	Gln182HE21	93.90%
OS4 of MMCOA	Gln182HE22	99.96%
OS5 of MMCOA	Ser287HG	99.99%
OS4 of MMCOA	Gln153HE21	99.99%
OS4 of MMCOA	Arg206HE	98.51%
OS5 of MMCOA	Arg206HE	98.54%
OS5 of MMCOA	Arg206HH21	99.37%
OS4 of MMCOA	Arg206HH21	99.99%

Table S3. Statistics associated with the hydrogen bonds between the MCOA and the protein residues for the 300ns MD simulation of the SalAT14(V285H/S287F/Q182I/F210M)-MCOA system

Acceptor	**Donor**	**Percentage**
OM2 of MCOA	Gln96H	89.71%
OM3 of MCOA	Gln96HE22	39.02%
OM4 of MCOA	Gln153HE21	38.22%
OM3 of MCOA	Gln153HE22	26.34%
OM4 of MCOA	Arg206HE	88.75%
OM4 of MCOA	Arg206HH21	99.05%

Table S4. Reaction energy barriers at different levels with M06-2x method for SalAT14-EMCOA. The energy barriers are given in kcal/mol.

	6-311+g(d)	**6-311+g(d,p)**	**6-311+g(2df,2p)**
R	0.0	0.0	0.0
TS1	19.5	19.8	22.2
INT1	17.5	19.2	21. 3
TS2	17.7	19.5	21.5
INT2	13.6	14.6	14.8
TS3	14.8	15.2	15.4
INT3	1.2	1.5	0.8

Table S5. The clustering statistics of the 50ns MD simulation which was selected to display the representative structure of SalAT14-EMCOA.

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	20005	0.800	0.981	0.131	16111	1.182
1	3621	0.145	1.003	0.158	2045	1.302
2	662	0.026	0.876	0.203	3920	1.304
3	380	0.015	0.703	0.094	9368	1.426
4	332	0.013	0.622	0.031	14515	1.350