Supplementary Information

Structure-guided manipulation of the regioselectivity of the cyclosporine A hydroxylase CYP-sb21 from *Sebekia benihana*

Fengwei Li[a,b]*, Li Ma[a], Xingwang Zhang[a], Jingfei Chen[b], Feifei Qi[b], Yinyue Huang[a], Zepeng Qu[b], Lishan Yao[b], Zhang Wei[a,b], Eung-Soo Kim[c]*, Shengying Li[a,b,d]*

[a]State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong 266237, China

[b]Shandong Provincial Key Laboratory of Synthetic Biology, CAS Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, Shandong 266101, China

[c]Department of Biological Engineering, Inha University, Incheon 22212, South Korea

[d]Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, Shandong 266237, China

*Address correspondence to: lifengwei@sdu.edu.cn (F.L.), eungsoo@inha.ac.kr (E.-S.K.), lishengying@qibebt.ac.cn (S.L.)

Table S1. Key residues around the substrate binding pocket

Autodock analysis	MD analysis	kJ/mol
83 Ile	83 Ile	
90 Tyr	90 Tyr	
91 Leu	91 Leu	-9.74
96 Leu	96 Leu	
184 Pro	184 Pro	
187 Met	187 Met	-5.68
188 Aal	188 Aal	-0.95
191 Val	191 Val	-5.46
238 Leu	238 Leu	
239 Thr	239 Thr	
242 Leu	242 Leu	-12.37
243 Ala	243 Ala	
246 Glu	246 Glu	-6.92
247 Thr	247 Thr	
290 Val	290 Val	-6.13
292 Gly	292 Gly	-0.71
293 Thr	293 Thr	
295 Val	295 Val	-2.58
294 Arg	294 ARG	-8.96
316 Leu	316 Leu	-6.58
396 Met	396 Met	-9.75
397 Pro	397 Pro	-19.63
398 Ala	398 Ala	
399 Ser	399 Ser	
	179 Met	-12.73
	34 Pro	-6.44
	33 Phe	
	97 Asp	
	40 Ile	
	319 Leu	
	291 Gln	
	289 Pro	
	31 Ala	
	240 Leu	
Total 24AA	**Total 34AA**	

The key amino acids identified only by MD analysis are colored in red. The energies of amino acids with the greatest energy change (top 10) and non-conservative in Figure 2 are listed. Amino acids that are non-conservative as shown in Figure 2 are marked by green shadows.

Table S2. The number of hydrogen bonds between CsA and different cyclophilins

Hydrogen bonds	Proteins	PDB ID
6	BmCYP-1	1C5F
5	murine_CyPC	2RMC
5	CypD	2Z6W
6	PfCyP19	1QNG
5	SmBz	4IPZ

Table S3. Oligonucleotide primers used in this study

Mutant sites	Primer (5' - 3')
E246W	F:GGTCACTGGACCACTGCTCACCTGATCGGTAACGGTACCGCTGC
	R:AGTGGTCCAGTGACCTGCCAGAACCAGGGTCAGTACCATCGTAAT
E246Y	F:GGTCACTACACCACTGCTCACCTGATCGGTAACGGTACCGCTGC
	R:AGTGGTGTAGTGACCTGCCAGAACCAGGGTCAGTACCATCGTAAT
E246L	F:GGTCACCTGACCACTGCTCACCTGATCGGTAACGGTACCGCTGC
	R:AGTGGTCAGGTGACCTGCCAGAACCAGGGTCAGTACCATCGTAAT
E246A	F:GGTCACGCAACCACTGCTCACCTGATCGGTAACGGTACCGCTGC
	R:AGTGGTTGCGTGACCTGCCAGAACCAGGGTCAGTACCATCGTAAT
G292A	F: GTGCAGGCAACTCGTGTTCGTTATGCAGCAGAGGATGTCGAA
	R: ACGAGTTGCCTGCACCGGACCACACCAACGCATCAGCTCGT
R294K	F: GGTACTAAAGTTCGTTATGCAGCAGAGGATGTCGAACTGG
	R: CGAACTTTAGTACCCTGCACCGGACCACACCAACGCATCAG
R294L	F: GGTACTCTGGTTCGTTATGCAGCAGAGGATGTCGAACTGG
	R: CGAACCAGAGTACCCTGCACCGGACCACACCAACGCATCAG
M179Q	F: GCTAAAAACGTATCCCTGTCTCCGGGTGCAATGGCCGAGC
	R: GGATACGTTTTTAGCACCCCATTCGCGCCACAGCGGACGG
M179R	F: GCTAAACGTGTATCCCTGTCTCCGGGTGCAATGGCCGAGC
	R: GGATACACGTTTAGCACCCCATTCGCGCCACAGCGGACGG
M187Q	F: GGTGCAAACGCCGAGCCGGTTATCAGCATGGTTGATTACA
	R: CTCGGCGTTTGCACCCGGAGACAGGGATACCATTTTAGCA
M187R	F: GGTGCACGTGCCGAGCCGGTTATCAGCATGGTTGATTACA
	R: CTCGGCACGTGCACCCGGAGACAGGGATACCATTTTAGCA
G292F	F: GTGCAGTTCACTCGTGTTCGTTATGCAGCAGAGGAT
	R: ACGAGTGAACTGCACCGGACCACACCAACGCATCAGCT
G292V	F: GTGCAGGTTACTCGTGTTCGTTATGCAGCAGAGGAT
	R: ACGAGTAACCTGCACCGGACCACACCAACGCATCAG
M316I	F: GCAGTGATTGCAGTACTGGTATCTGCGAACTATGAT
	R: TACTGCAATCACTGCTTCGCCACGCTTAACGGTCAT
M316Q	F: GCAGTGCAAGCAGTACTGGTATCTGCGAACTATGAT
	R: TACTGCTTGCACTGCTTCGCCACGCTTAACGGTCAT
V191L	F: GAGCCGCTGATCAGCATGGTTGATTACATCCACGAT
	R: GCTGATCAGCGGCTCGGCCATTGCACCCGGAGACAG
V191F	F: GAGCCGTTCATCAGCATGGTTGATTACATCCACGAT
	R: GCTGATGAACGGCTCGGCCATTGCACCCGGAGACAG
V295L	F: ACTCGTCTGCGTTATGCAGCAGAGGATGTCGAACTG
	R: ATAACGCAGACGAGTACCCTGCACCGGACCACACCA
V295F	F: ACTCGTTTCCGTTATGCAGCAGAGGATGTCGAACTG
	R: ATAACGGAAACGAGTACCCTGCACCGGACCACACCA
M396R	F: CAGCTGCGTCCGGCCTCTTGGCGCCTGGCATCCCT
	R: AGGCCGGACGCAGCTGACGTTCCAGTTCTTCCGGCGCAA
M396L	F: CAGCTGCTGCCGGCCTCTTGGCGCCTGGCATCCCT
	R: AGGCCGGCAGCAGCTGACGTTCCAGTTCTTCCGGCGCAA

The mutant sites are shown in red and underlined.

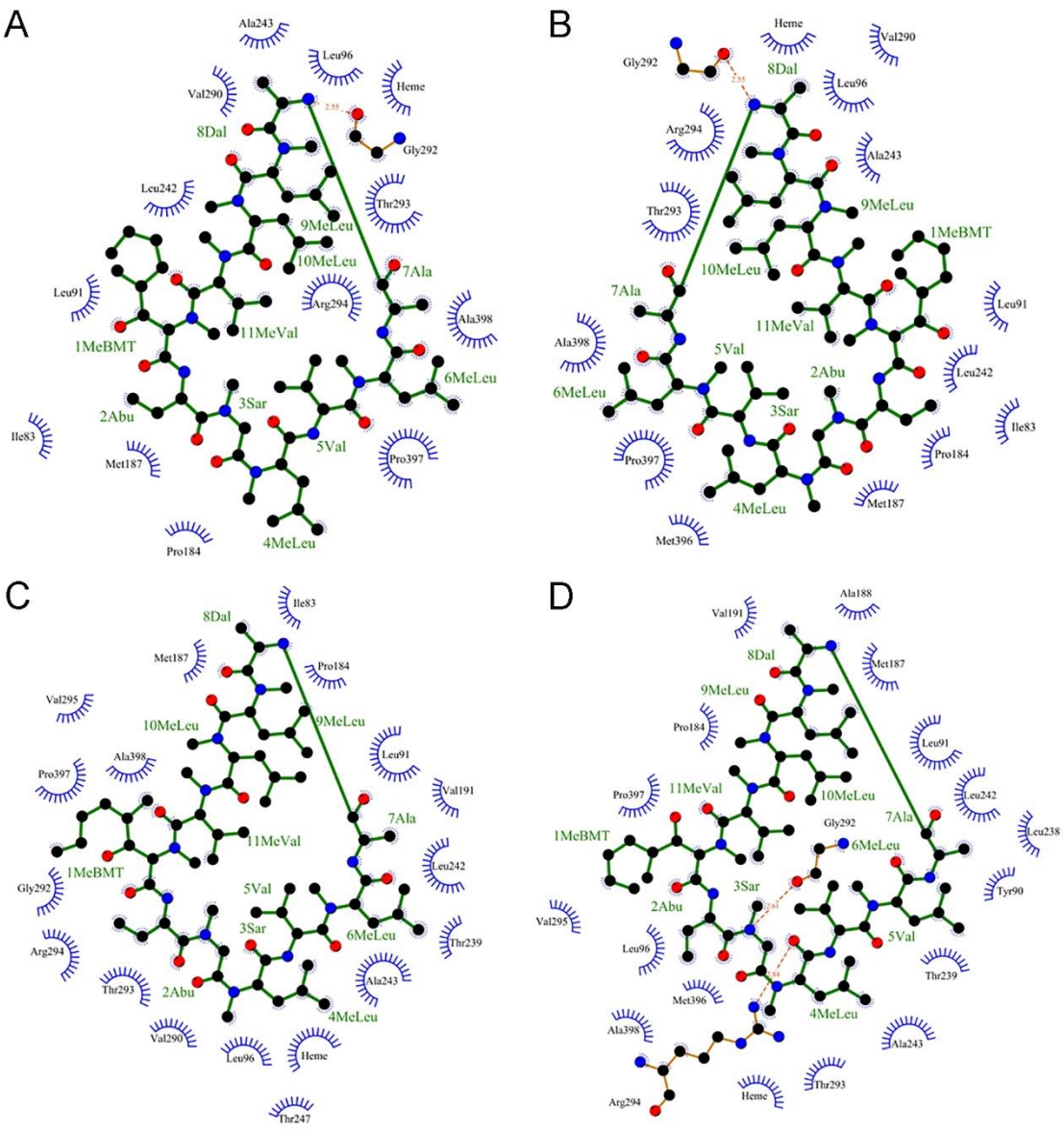

Figure S1. A. Key amino acids in pose 2 of CsA. **B.** Key amino acids in pose 6 of CsA. **C.** Key amino acids in pose 8 of CsA. **D.** Key amino acids in pose 9 of CsA. Hydrogen bonds are labelled by red dashed lines.

Figure S2. Key amino acids revealed in the top 10 poses of CsA (the distances to substrate are within 5 Å). Red boxes represent the key amino acids seen in all poses; green boxes represent the key amino acids seen in a part of poses; and yellow boxes represent the key amino acids only seen in pose 8 or pose 9.

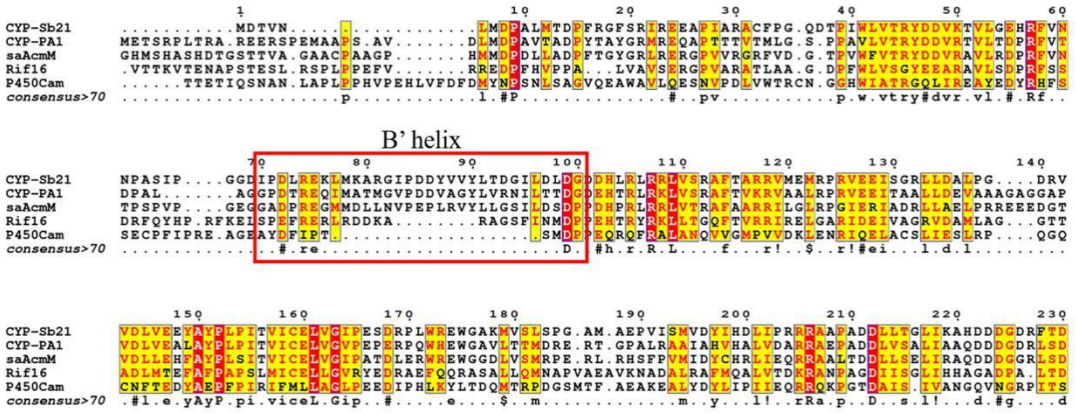

Figure S3. Multiple protein sequence alignment of CYP-sb21, CYP-pa1, saAcmM, Rif16, and P450cam with substrates in different sizes and shapes. The substrates of CYP-sb21, CYP-pa1, saAcmM, Rif16, and P450cam are cyclosporine A (MW 1202.6), cyclosporine A (MW 1202.6), actinomycin D (MW 1255.4), rifamycin L (MW 755.8), and camphor (MW 152.2), respectively. The B' helix region is boxed in a red rectangle. Sequence analysis was performed using Expresso through the T-COFFEE online service, and the figure was prepared using ESPript 3.0.

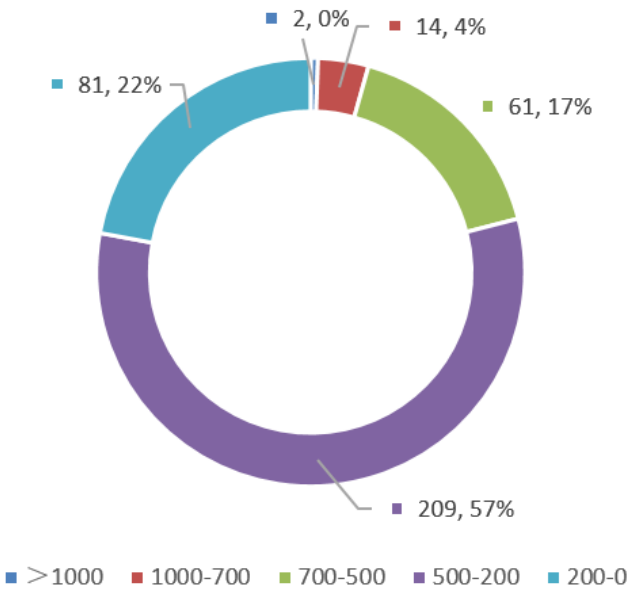

Figure S4. Molecular weight distribution of all P450 substrates/ligands in PDB databank (http://www.rcsb.org/).

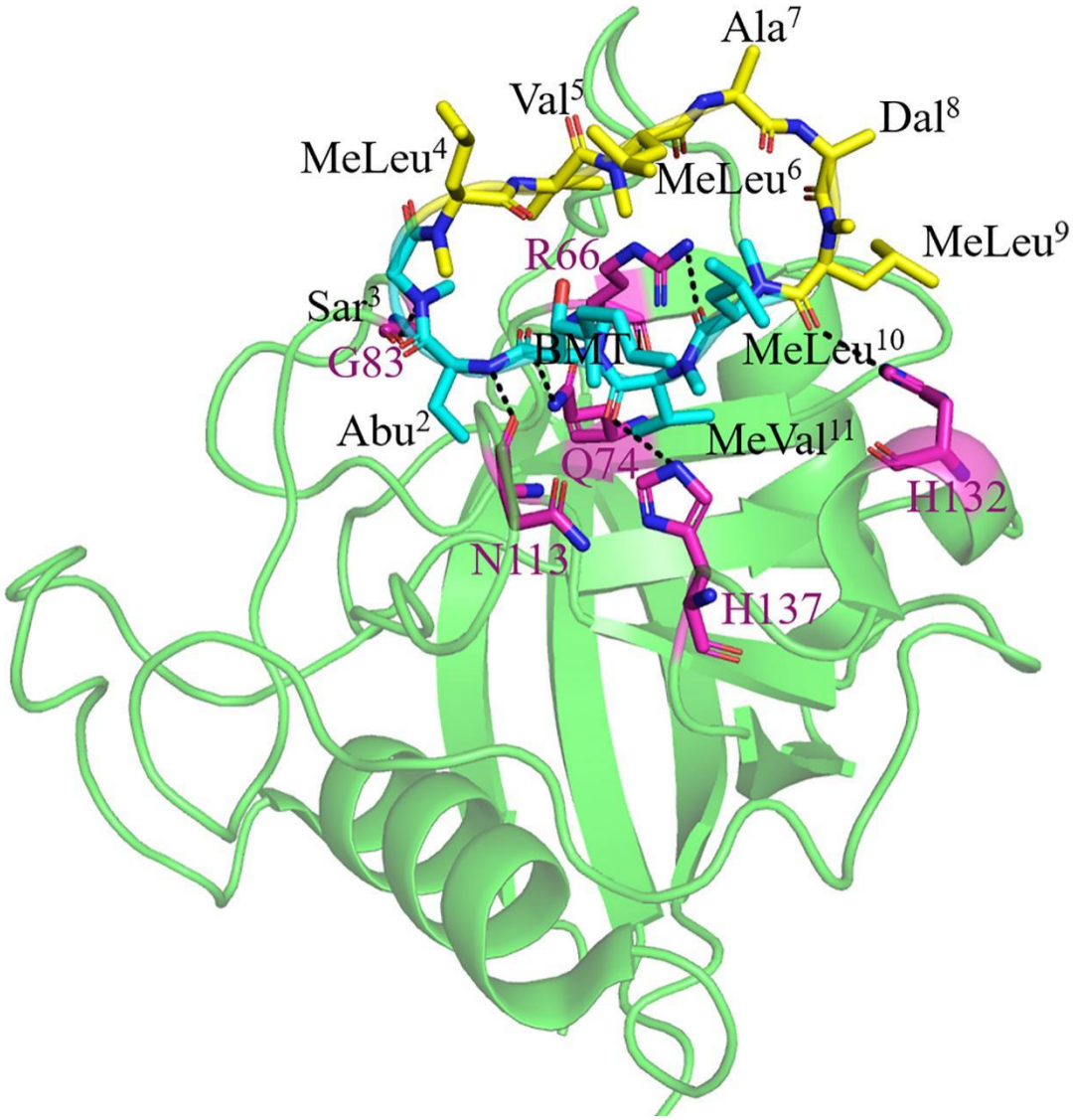

Figure S5. The 3D structure of CsA in complex with the cyclophilin BmCYP-1 (PDB ID: 1C5F). CsA is shown as sticks. The amino acids bind in and out of binding pocket are colored in cyan and yellow, respectively. The amino acids contact CsA through hydrogen bonds (labelled by black dashed lines) are shown as sticks in magenta.