

## Supporting Information

# Computational design of efficient and thermostable esterase for polylactic acid depolymerization

*Bin Xie<sup>a,e#</sup>, Jun Zhang<sup>c#</sup>, Huashan Sun<sup>a,e</sup>, Rongrong Bai<sup>a,e</sup>, Diannan Lu<sup>c</sup>, Yushan Zhu<sup>b</sup>,  
Weiliang Dong<sup>a,d,e\*</sup>, Zhou Jie<sup>a,d,e\*</sup>, Min Jiang<sup>a,d,e</sup>*

<sup>a</sup>State Key Laboratory of Materials-Oriented Chemical Engineering, College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University, Nanjing, 211800, P.R. China

<sup>b</sup>College of Life Science and Technology, National Energy R&D Center for Biorefinery, Beijing University of Chemical Technology, Beijing 100029, China

<sup>c</sup>Department of Chemical Engineering, Tsinghua University, Beijing 100084, China

<sup>d</sup>Jiangsu National Synergetic Innovation Center for Advanced Materials (SICAM), Nanjing Tech University, Nanjing 211800, P.R. China

<sup>e</sup>Key Laboratory for Waste Plastics Biocatalytic Degradation and Recycling, Nanjing Tech University, Nanjing 211816, Jiangsu, China

\* Corresponding author: Jie Zhou, [jayzhou@njtech.edu.cn](mailto:jayzhou@njtech.edu.cn);

Weiliang Dong, [dwl@njtech.edu.cn](mailto:dwl@njtech.edu.cn)

## Table of Contents

Supporting Tables .....	2
Supporting Figures .....	14
References .....	19

## Supporting Tables

**Table S1. Primers for designing single-point mutations.**

Mutation Sites	Primer name	Sequence
S128Y	F-S128Y	CGACGGTtatATCGATCCGAAGATCCTGTTCA
	R-S128Y	GATCGATataACCGTCGAGACACATCCGCGCC
S153L	F-S153L	AGCTTCTTcctgCCTCTCGCCAAGATGCTGG
	R-S153L	AGAGGcagGAAGAAGCTCGCCCGCG
P232A	F-P232A	AATCTGAAGgcgCTGCTGGTGCTGGTGGCG
	R-P232A	AGCAGcgcCTTCAGATTGGGTAACGCTTTCA
I245V	F-I245V	GCTgtgCCGCCATCGGTCGCGGTT
	R-I245V	ACCGATGGCGGcacAGCGCGGTCCCCCTCGGC
R276D	F-R276D	AGgatCCTGCGCTGATTGCCGCC
	R-R276D	AATCAGCGCAGGatcCTCCTCATGGGCCAGATGG
V202W	F-V202W	TGtggTCCAGCCCAAATCATGTGCGCCGAGCG
	R-V202W	ATTTGGGCTGGAccaCAGCTTGCCGTAGAGTTTGATCC
G139F	F-G139F	CCTCAATccaGCGTTTCTGCCCTATGGCGGCC
	R-G139F	GAAACGctggATTGAGGCTGAACAGGATCTTCG
F151L	F-F151L	AGCctaTTCTCGCCTCTCGCCAAGATGCTGGT
	R-F151L	AGAGGCGAGAAtagGCTCGCCGCGGGCCGCC
A267G	F-A267G	ATTCCGggtCTCGGCCATCTGGCCCATGAGGA
	R-A267G	TGGCCGAGaccCGGAATCCGCTCGATCACCGC

**Table S2-1. Potentially stabilizing mutations predicted by PROSS' calculations.**

Number	Wild Type	design_1	design_2	design_3	design_4	design_5	design_6	design_7	design_8	design_9	Variant
282	A									E	A282E
255	E									A	E255A
285	E									R	E285R
161	M									L	M161L
276	R									D	R276D
128	S									M	S128M
248	S									D	S248D
102	V									A	V102A
233	L								P	P	L233P
261	V								R	R	V261R
22	I							V	V	V	I22V
225	K							R	R	R	K226R
150	S							R	R	R	S150R
203	S							Q	Q	Q	S203Q
251	V							E	E	R	V251R
254	R						A	A	A	A	R254A
94	S						E	E	E	E	S94E
134	L	I	I		I					I	L134I
221	E					D	D	D	D	D	E221D

11	L					R	R	R	R	R	L11R
176	R					P	P	P	P	P	R176P
167	S					H	H	H	H	H	S167H
241	G		Q	Q	Q	Q			Q	Q	G241Q
265	I				L	L	L	L	L	L	I265L
196	K				A	A	A	A	A	A	K196A
159	L				M	M	M	M	M	M	L159M
206	N				G	G	H	H	H	H	N206H
154	P		A	A	A	A	A	A	A	A	P154A
36	S		D	D	D	D	D	D	D	D	S36D
193	A	R	R	R	R	R	R	R	R	R	A193R
238	A	V	V	V	V	V	V	V	V	V	A238V
267	A	G	G	G	G	G	G	G	G	G	A267G
38	A	D	D	D	D	D	D	D	D	D	A38D
151	F	L	L	L	L	L	L	L	L	L	F151L
139	G	P	P	P	P	P	P	P	P	P	G139P
222	P	T	T	T	T	T	T	T	T	T	P222T
232	P	A	A	V	A	V	V	V	V	V	P232V
160	V	A	A	A	A	A	A	A	A	A	V160A

**Table S2-2. Potentially stabilizing mutations predicted by Fireprot<sup>2</sup> calculations.**

Combined mutant: -35.20 kcal/mol (16 mutations)							
Chain	Mutation	Conserved	Correlated	BTC by majority	BTC by ratio	FoldX(kcal/mol)	Rosetta(kcal/mol)
A	I45L	Y	N	Y	N	0.12	-
A	S63A	N	N	Y	Y	0.22	-
A	Y66F	Y	N	Y	N	-0.61	-
A	Q104G	N	N	N	Y	-0.86	-
A	S128Y	N	N	N	N	-2.4	-2.75
A	Y144F	Y	N	Y	Y	0.4	-
A	F151L	N	N	Y	Y	-0.09	-
A	S153L	N	N	N	N	-1.49	-2.52
A	A173Y	N	N	N	N	-2.15	-3.66
A	G199Y	N	N	N	N	-1.55	-5.27
A	K200R	N	N	Y	Y	0.03	-
A	V202W	N	N	N	N	-1.57	-4.34
A	A209W	N	N	N	N	-1.64	-2.33
A	I245V	Y	N	Y	N	0.44	-
A	A267G	N	N	Y	Y	-2.3	-2.96
A	A281Y	N	N	N	N	-1.3	-4.73
Energy mutant: -33.75 kcal/mol (8 mutations)							
Chain	Mutation	Conserved	Correlated	FoldX(kcal/mol)	Rosetta(kcal/mol)		

A	S128Y	N	N	-2.4	-2.75
A	S153L	N	N	-1.49	-2.52
A	A173Y	N	N	-2.15	-3.66
A	G199Y	N	N	-1.55	-5.27
A	V202W	N	N	-1.57	-4.34
A	A209W	N	N	-1.64	-2.33
A	A267G	N	N	-2.3	-2.96
A	A281Y	N	N	-1.3	-4.73

Evolution mutant: -2.96 kcal/mol (9 mutations)

Chain	Mutation	BTC by majority	BTC by ratio	FoldX(kcal/mol)
A	I45L	Y	N	0.12
A	S63A	Y	Y	0.22
A	Y66F	Y	N	-0.61
A	Q104G	N	Y	-0.86
A	Y144F	Y	Y	0.4
A	F151L	Y	Y	-0.09
A	K200R	Y	Y	0.03
A	I245V	Y	N	0.44
A	A267G	Y	Y	-2.3

**Table S2-3. Potentially stabilizing mutations predicted by GRAPE<sup>3</sup> calculations, based respectively on FoldX, Rosetta, ABACUS method.**

Chain	WT	MUT	Num	Energy	SD	Method	WTNumMUT
A	A	D	18	-2.428	0.064	FoldX	A18D
A	R	L	83	-2.128	0.111	FoldX	R83L
A	A	M	38	-2.066	0.075	FoldX	A38M
A	L	E	11	-1.941	0.015	FoldX	L11E
A	P	M	14	-1.905	0.008	FoldX	P14M
A	T	L	81	-1.885	0.12	FoldX	T81L
A	S	L	63	-1.869	0.044	FoldX	S63L
A	V	R	31	-1.776	0.146	FoldX	V31R
A	Q	W	80	-1.635	0.023	FoldX	Q80W
A	P	V	60	-1.597	0.228	FoldX	P60V
A	D	V	95	-1.511	0.042	FoldX	D95V
A	L	G	86	-1.51	0.029	FoldX	L86G
A	G	E	77	-1.505	0.036	FoldX	G77E
A	G	T	26	-1.494	0.092	FoldX	G26T
A	Y	P	66	-1.452	0.052	FoldX	Y66P
A	L	Q	61	-1.412	0.082	FoldX	L61Q
A	P	T	37	-1.358	0.384	FoldX	P37T
A	A	I	97	-1.347	0.008	FoldX	A97I
A	P	N	40	-1.346	0.197	FoldX	P40N
A	L	Q	62	-1.332	0.065	FoldX	L62Q
A	H	W	30	-1.321	0.032	FoldX	H30W
A	W	H	29	-1.275	0.589	FoldX	W29H
A	D	L	9	-1.263	0.002	FoldX	D9L
A	G	I	57	-1.257	0.345	FoldX	G57I
A	P	S	82	-1.245	0.011	FoldX	P82S
A	L	Y	4	-1.227	0.006	FoldX	L4Y
A	E	V	17	-1.199	0.009	FoldX	E17V
A	V	S	69	-1.188	0.106	FoldX	V69S
A	F	S	78	-1.179	0.046	FoldX	F78S
A	S	W	94	-1.155	0.012	FoldX	S94W
A	A	E	70	-1.139	0.156	FoldX	A70E
A	L	T	96	-1.129	0.079	FoldX	L96T
A	H	F	76	-1.081	0.052	FoldX	H76F
A	G	F	25	-1.074	0.063	FoldX	G25F
A	H	S	15	-1.065	0.116	FoldX	H15S
A	S	Q	36	-1.022	0.206	FoldX	S36Q
A	H	V	65	-1.022	0.017	FoldX	H65V



Chain	WT	MUT	Num	Energy	SD	Method	WTNumMUT
A	R	P	83	-30.476	0.014	Rosetta	R83P
A	P	W	154	-13.757	0.005	Rosetta	P154W
A	P	Q	222	-11.069	0	Rosetta	P222Q
A	P	I	166	-10.38	0	Rosetta	P166I
A	S	L	153	-9.901	0.027	Rosetta	S153L
A	P	A	232	-8.371	0	Rosetta	P232A
A	S	Y	128	-8.215	0	Rosetta	S128Y
A	E	L	221	-7.501	0	Rosetta	E221L
A	V	W	202	-7.363	0.114	Rosetta	V202W
A	G	F	145	-7.053	0.006	Rosetta	G145F
A	A	F	238	-6.877	0	Rosetta	A238F
A	M	F	84	-6.544	0	Rosetta	M84F
A	D	Y	95	-6.532	0	Rosetta	D95Y
A	K	L	231	-6.319	0	Rosetta	K231L
A	M	W	34	-6.288	0	Rosetta	M34W
A	A	W	173	-6.269	0	Rosetta	A173W
A	G	F	199	-6.24	1.256	Rosetta	G199F
A	G	P	139	-6.224	0.03	Rosetta	G139P
A	G	F	241	-6.125	0.01	Rosetta	G241F
A	A	W	156	-5.953	0	Rosetta	A156W
A	G	Y	57	-5.808	0.183	Rosetta	G57Y
A	G	D	91	-5.808	0	Rosetta	G91D
A	H	F	53	-5.712	0.002	Rosetta	H53F
A	R	F	264	-5.61	0	Rosetta	R264F
A	A	W	149	-5.517	0.694	Rosetta	A149W
A	G	R	184	-5.37	0	Rosetta	G184R
A	T	P	48	-5.344	0.004	Rosetta	T48P
A	E	W	285	-5.297	0	Rosetta	E285W
A	G	F	49	-5.248	0.001	Rosetta	G49F
A	R	W	122	-5.23	0	Rosetta	R122W
A	A	Y	106	-5.22	0	Rosetta	A106Y
A	K	W	196	-5.173	0.007	Rosetta	K196W

Chain	WT	MUT	Num	Energy	SD	Method	WTNumMUT
A	Q	D	108	-16.567	0	ABACUS	Q108D
A	T	S	79	-9.675	0	ABACUS	T79S
A	G	L	139	-7.408	0	ABACUS	G139L
A	A	L	288	-6.784	0	ABACUS	A288L
A	E	I	263	-6.769	0	ABACUS	E263I
A	K	D	132	-6.402	0	ABACUS	K132D
A	S	D	136	-6.367	0	ABACUS	S136D
A	P	A	166	-5.95	0	ABACUS	P166A
A	A	N	51	-5.614	0	ABACUS	A51N
A	A	G	239	-5.279	0	ABACUS	A239G
A	A	G	117	-5.117	0	ABACUS	A117G
A	A	T	41	-5.11	0	ABACUS	A41T
A	F	L	141	-4.774	0	ABACUS	F141L
A	H	Q	273	-4.634	0	ABACUS	H273Q
A	L	A	120	-4.586	0	ABACUS	L120A
A	D	L	126	-4.543	0	ABACUS	D126L
A	G	L	194	-4.449	0	ABACUS	G194L
A	S	L	128	-4.315	0	ABACUS	S128L
A	Y	L	198	-4.27	0	ABACUS	Y198L
A	F	L	135	-4.269	0	ABACUS	F135L
A	N	S	138	-4.196	0	ABACUS	N138S
A	R	D	276	-3.757	0	ABACUS	R276D
A	F	A	152	-3.717	0	ABACUS	F152A
A	H	A	175	-3.58	0	ABACUS	H175A
A	G	F	10	-3.544	0	ABACUS	G10F
A	I	V	245	-3.537	0	ABACUS	I245V
A	H	V	207	-3.434	0	ABACUS	H207V
A	Q	Y	32	-3.288	0	ABACUS	Q32Y
A	W	I	6	-3.256	0	ABACUS	W6I
A	S	W	52	-3.224	0	ABACUS	S52W
A	R	T	28	-3.223	0	ABACUS	R28T
A	F	E	164	-3.151	0	ABACUS	F164E
A	Y	P	144	-3.12	0	ABACUS	Y144P
A	W	L	29	-3.113	0	ABACUS	W29L
A	L	I	236	-3.041	0	ABACUS	L236I

**Table S3. Structural basis and effect on thermostability of the single mutations for experimental tests predicted by three stability prediction tools.**

Mutations	Origin <sup>a</sup>	Location	Predicted improvement
S128Y	FireProt/GRAPE	loop in Hyd <sup>b</sup>	New H-bond interactions with the hydroxy Oxygen S94
G139P	PROSS/GRAPE	loop connecting Lid and Hyd	Increase of the rigidity of the loop
F151L	PROSS/FireProt	$\alpha$ -helix in Lid	Reduction of unfavorable hydrophobic buildup on exposed surfaces
S153L	FireProt/GRAPE	$\alpha$ -helix in Lid	Hydrophobic interactions
V202W	FireProt/GRAPE	$\alpha$ -helix in Lid	Entropic stabilization + Optimized surface charge Hydrophobic interactions and filling the deeply buried cavity
P232A	PROSS/GRAPE	loop in Hyd	Enhanced local hydrophobic packing
I245V	FireProt/GRAPE	loop in Hyd	Based on evolutionary information
A267G	PROSS/FireProt	loop in Hyd	Reduction of exclusion with P266
R276D	PROSS/GRAPE	$\alpha$ -helix in Hyd	New H-bond with backbone nitrogens of A278 and L279

<sup>a</sup> Single mutations predicted by at least two tools simultaneously were selected for experimental tests.

<sup>b</sup> The hydrolase domain is abbreviated as Hyd and the lid domain as Lid.

**Table S4. Overall scoring ranking of three mutations predicted by three predictive tools for disulfide bonds. The tools used to predict disulfide bond mutations were: Yosshi<sup>4</sup>, MAESTROweb<sup>5</sup>, and SSbondPre<sup>6</sup>.**

Predicted mutations	Yosshi's Rank	MAESTROweb's Rank	SSbondPre's Rank
M34-S63	22/59	2/106	13/51
A97-S128	35/59	1/106	9/51
L257-A260	17/59	41/106	50/51
A51-G77	1/59	25/106	

**Table S5. Folding free energy changes of experimental variants computed by Cartesian\_ddg<sup>7</sup>.**

Mutation	$\Delta\Delta G_{\text{fold}}$ (kcal/mol) <sup>a</sup>
S128Y	-8.46
G139P	-4.79
F151L	-1.53
S153L	-9.76
V202W	-7.47
P232A	-9.14
I245V	<b>1.63</b> <sup>b</sup>
A267G	-3.32
R276D	-0.56
F151L/A267G	-4.86
I245V/R276D	<b>1.05</b> <sup>b</sup>
I245V/S128Y	-6.66
S153L/R276D	-10.34
S153L/S128Y	-18.22
S153L/I245V	-7.68
S153L/I245V/R276D	-8.59
S153L/I245V/R276D/S128Y	-17.05
S153L/I245V/S128Y/V202W	-23.70
S153L/I245V/R276D/S128Y/V202W	-24.51

<sup>a</sup>  $\Delta\Delta G_{\text{fold}}$  is the difference of the folding energy of variant minus that of wild type. The value is the average of three rounds of Cartesian\_ddg calculations.

<sup>b</sup> The only two variants with positive  $\Delta\Delta G_{\text{fold}}$ , because I245V is a conserved mutation provided by evolutionary information.

**Table S6. Kinetic parameters of R5, WT derived from Michaelis-Menten experiments with 1-naphthylpropionate as substrate.**

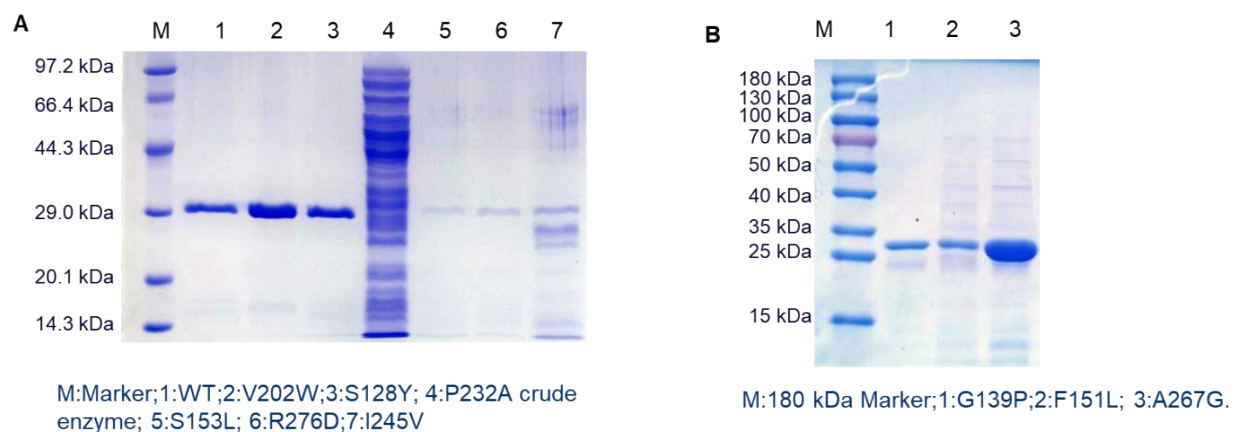
Protein	$V_{\max}$ ( $\mu\text{mol}/\text{min}/\text{mg}$ )	$K_m$ (mM)	$K_{\text{cat}}$ ( $\text{S}^{-1}$ )	$K_{\text{cat}}/K_m(\text{S}^{-1}\text{mM}^{-1})$
WT	2.22±0.11	0.74±0.1	1.17±0.06	1.59
R5	4.28±0.20	0.71±0.09	2.26±0.11	3.19

**Table S7. Kinetic parameters of R5, WT derived from the inverse Michaelis-Menten experiments<sup>8</sup> with PDLLA substrate concentration fixed at 10 g/L.**

Protein	$^{\text{Inv}}V_{\max}/S$ ( $\mu\text{mol g}^{-1} \text{s}^{-1}$ )	$^{\text{Inv}}K_m$ ( $\text{mg}_{\text{enzyme}} \text{g}_{\text{PLA}}^{-1}$ )
WT	0.1913	1.6380
R5	0.9229	1.1710

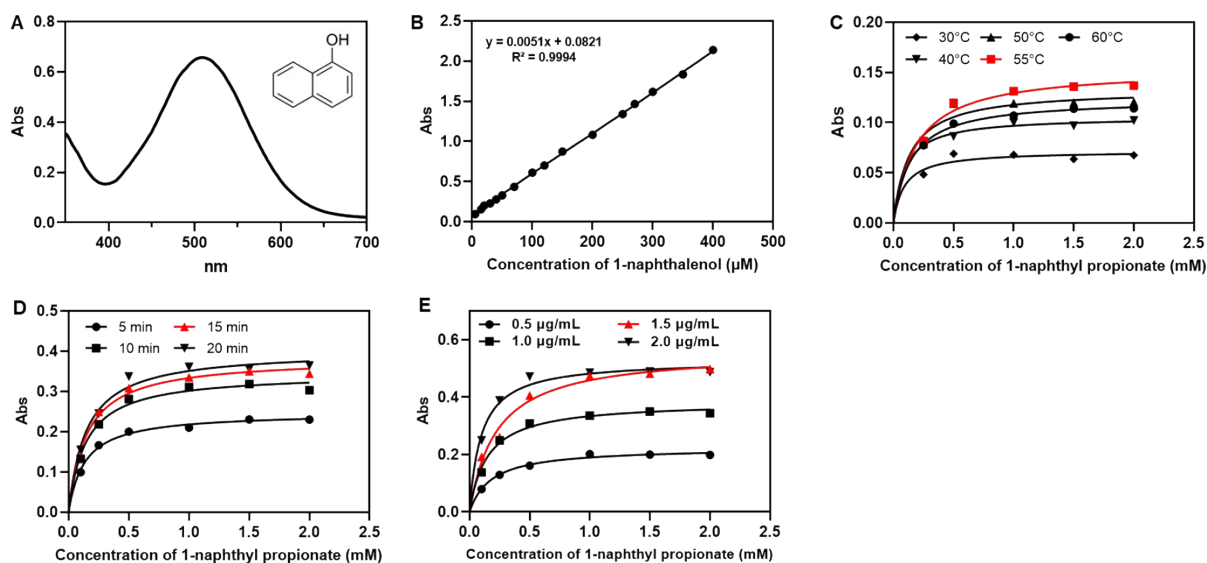
## Supporting Figures

**Figure S1.** SDS-PAGE<sup>9</sup> validation of the wild-type protein after single-point mutation. (A) The protein purification bands of V202W, S128Y, P232A crude enzyme extracts, along with S153L, R276D, and I245V. Notably, no corresponding protein band was found for the P232A mutation. (B) The protein purification bands for G139P, F151L, and A267G.



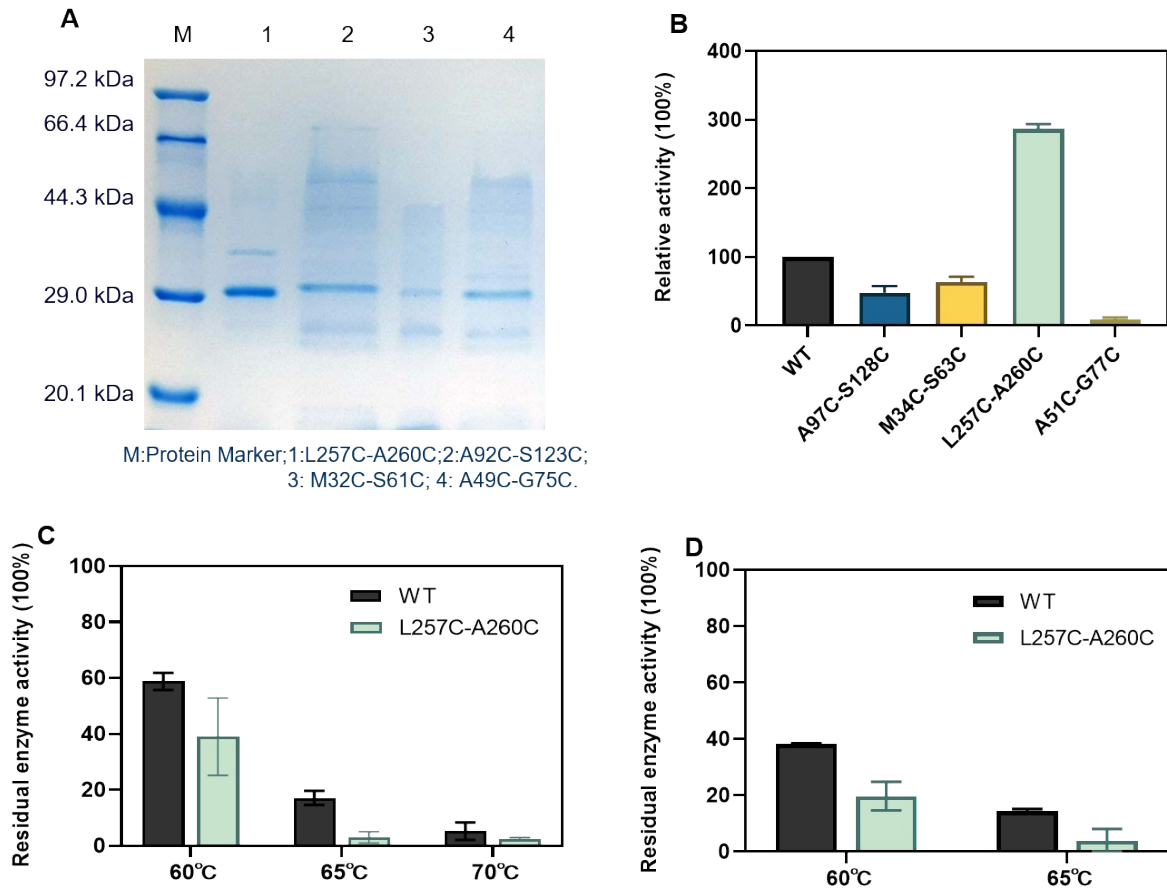
**Figure S2.** Optimization of the enzymatic activity assay system for 1-naphthylpropionate<sup>10</sup>. (A) Visible light wavelength scan of the product 1-naphthylpropionate formed after the reaction. Under alkaline conditions, the released product 1-

naphthol undergoes a catalytic reaction with ferricyanide in the presence of 4-aminoantipyrine, yielding a purple-red product with maximum absorption at a wavelength of 510 nm. (B) Standard curve established with different concentrations of 1-naphthol. (C) Absorbance intensity at 510 nm after 10 min reaction with 1-naphthylpropionate at different temperatures. (D) Absorbance intensity at 510 nm at different time points during the reaction with 1-naphthylpropionate. (E) Absorbance intensity at 510 nm after the reaction of different enzyme concentrations with 1-naphthylpropionate.



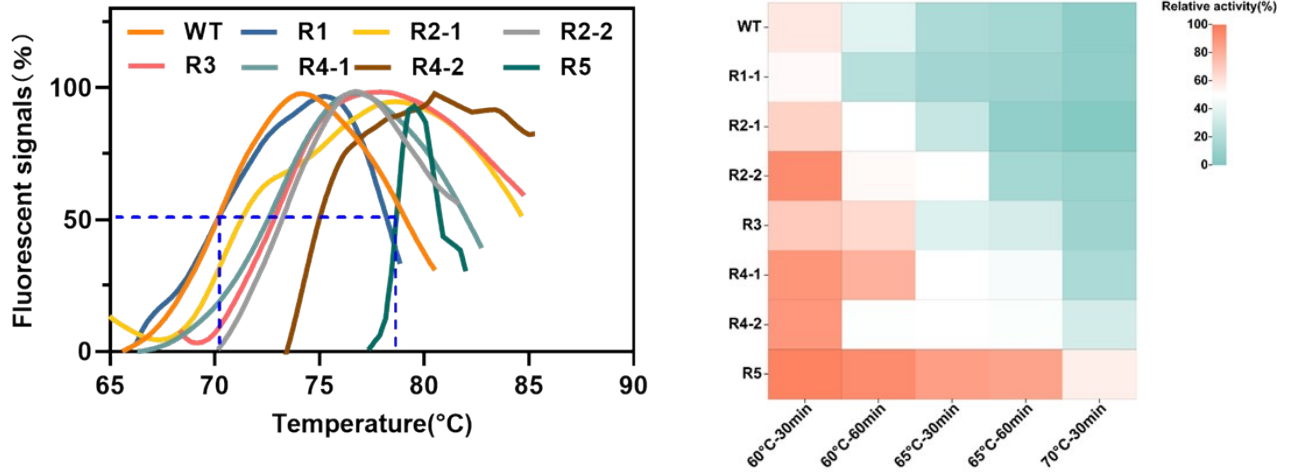
**Figure S3.** (A) Validation of the purified protein after disulfide bond combinatorial mutation with SDS-PAGE. (B) Evaluation of the enzyme activity of the mutants using 1-naphthyl propionate as a soluble substrate at 55°C. (C) Evaluation of the thermal stability of mutants after heat treatment at different temperatures for 30 min. (D) Evaluation of the thermal

stability of mutants after heat treatment at different temperatures for 60 min. Residual enzyme activity was assessed using 1-naphthyl propionate, with enzyme hydrolysis monitored by measuring absorbance at a wavelength of 510 nm.

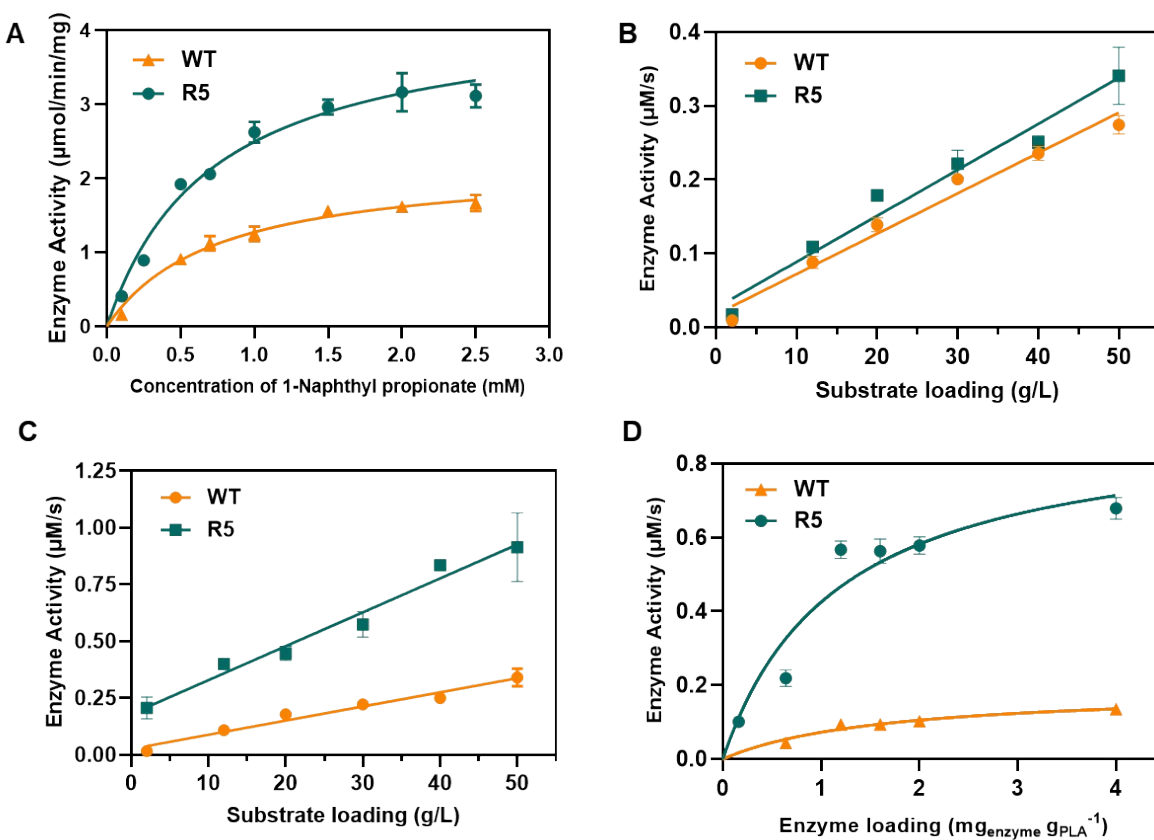




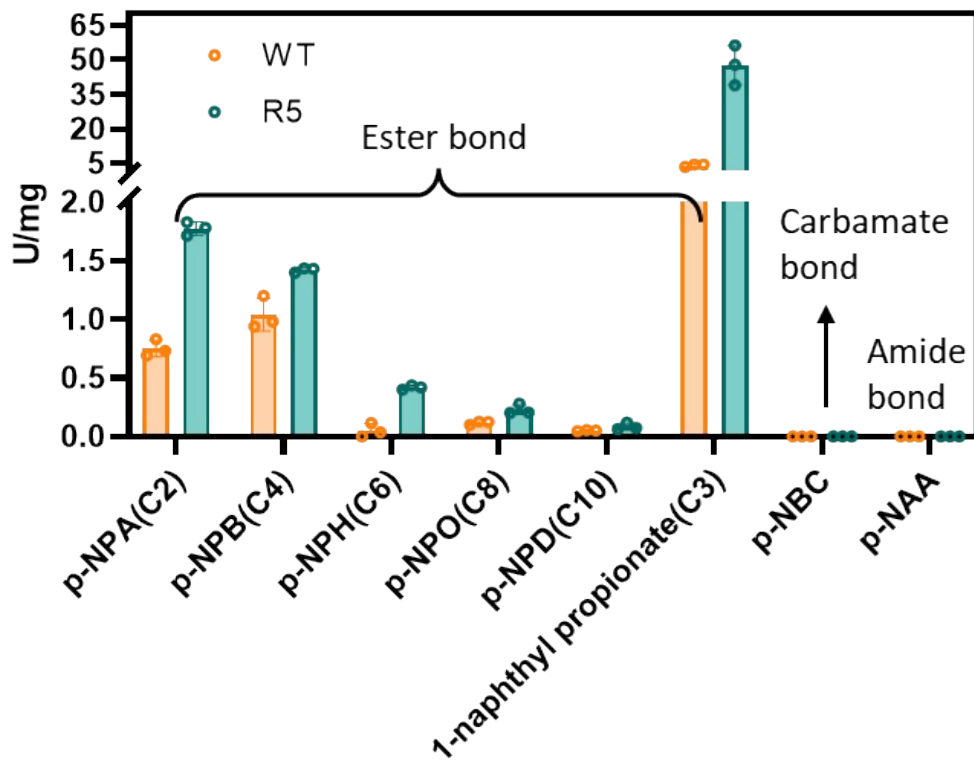
**Figure S4.** Comparison of the thermal stability properties between the mutants and the wild type (WT). (A) Fluorescence curves of all mutant proteins in DSF <sup>11</sup>(differential scanning fluorimetry) to calculate the temperature at which the fluorescence signal rises by half ( $T_m$ ). (B) Comparison of the thermal stability between variant R5 and WT after heat treatment at different temperatures for different minutes. 1-naphthylpropionate was used as the detection substrate.



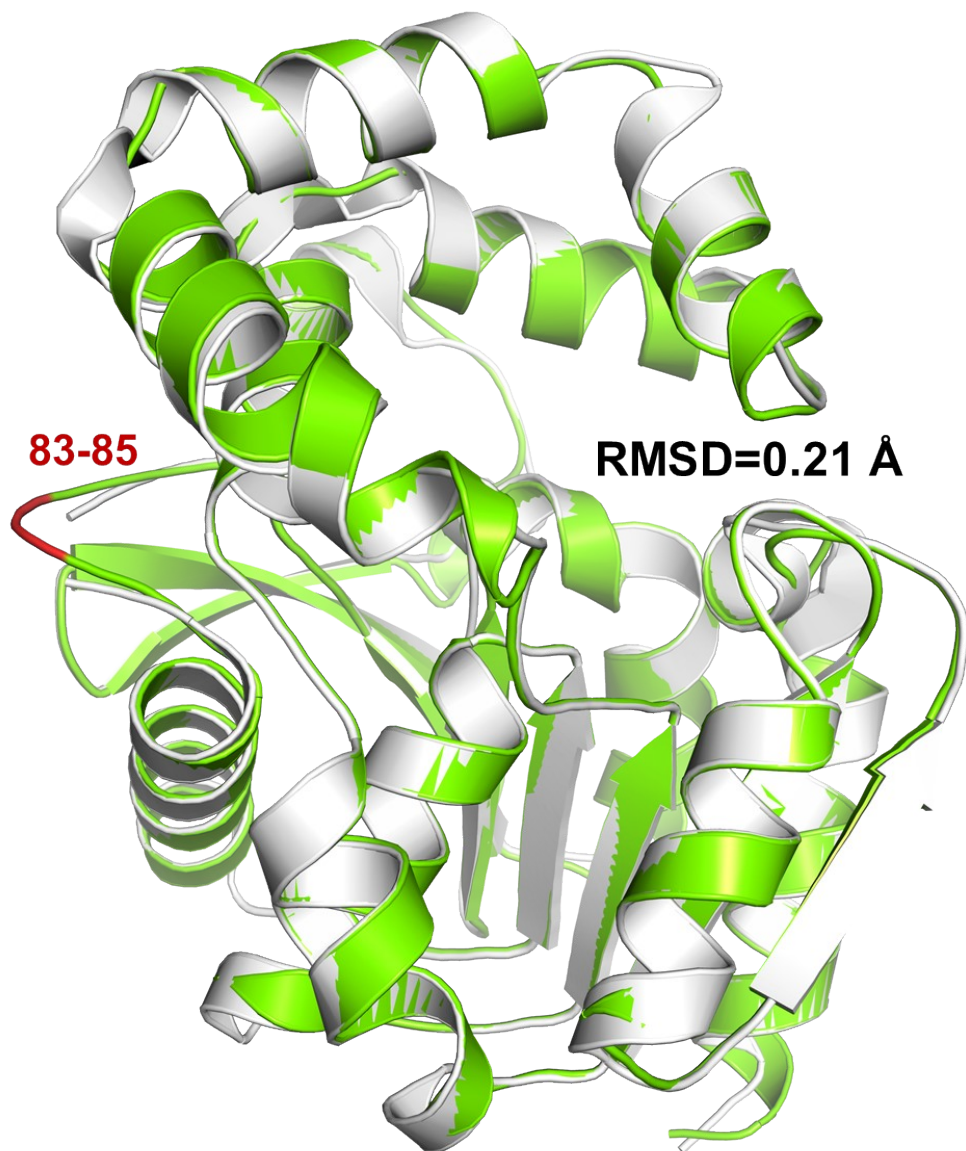
**Figure S5.** Characterization of WT and variant R5 kinetic parameters. (A) Conventional and Inverse (B-D) Michaelis-Menten plots of WT and R5, tested at 55°C with different substrate concentrations or different enzyme concentrations. Detection of conventional kinetic parameters of WT and R5 using 1-naphthylpropionate as a soluble substrate. (B) Fixed enzyme concentration of 12  $\mu\text{M}$ . (C) Fixed enzyme concentration of 37  $\mu\text{M}$ . (D) Inverse Michaelis-Menten<sup>8</sup> experiments assay with PLA substrate concentration fixed at 10 g/L. The concentration of lactic acid monomers is detected using HPLC<sup>5</sup>. The experiment is conducted in triplicate. Data are expressed as the mean  $\pm$  s.d. (standard deviation).



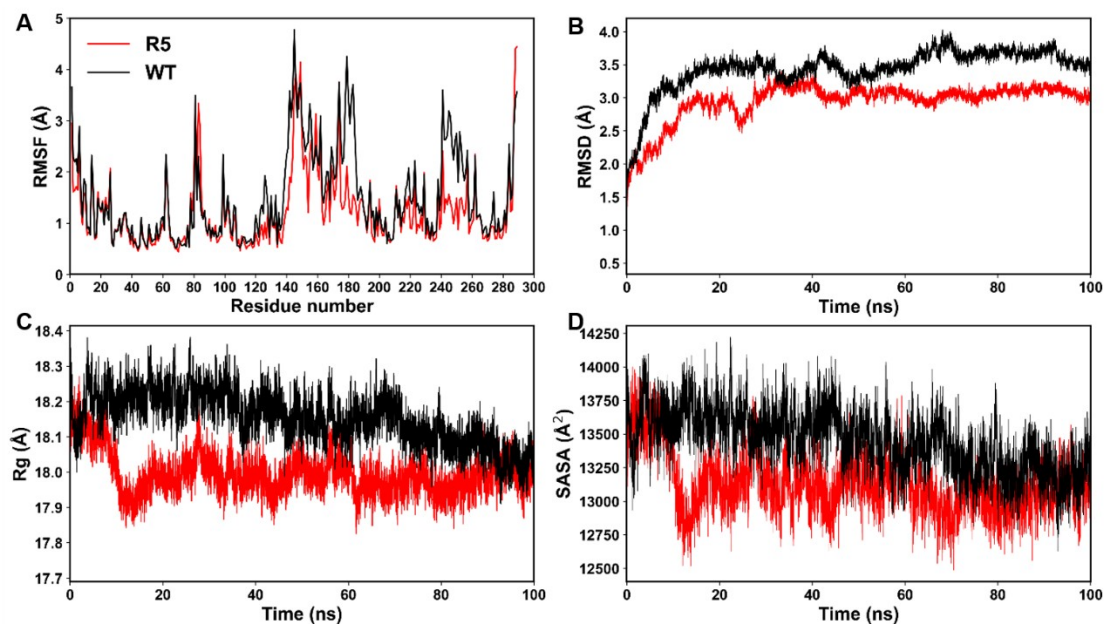
**Figure S6.** Substrate profiles for WT and variant R5 against model substrates. Model substrates include those commonly used for analyzing ester bond hydrolysis capability, such as p-nitrophenyl acetate (p-NPA, C2), p-nitrophenyl butyrate (p-NPB, C4), p-nitrophenyl hexanoate (p-NPH, C6), p-nitrophenyl octanoate (p-NPO, C8), p-nitrophenyl decanoate (p-NPD, C10), and 1-naphthyl propionate (C3). In addition, two other substrates were tested: tert-butyl (4-nitrophenyl) carbamate (p-NBC) containing a carbamate bond, and 4'-nitroacetanilide (p-NAA) containing an amide bond. Enzyme activity is defined as the amount of enzyme required to produce 1  $\mu\text{mol}$  of para-nitrophenol (pNP) or 1-naphthol per minute at 55°C, serving as one unit of enzymatic activity.



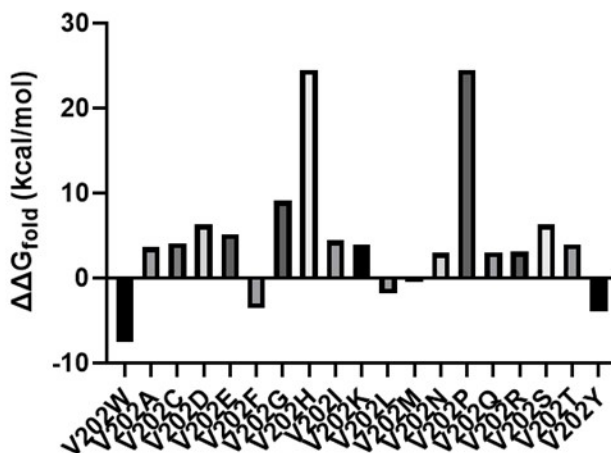
**Figure S7.** Structural alignment of the predicted structure of RPA1511 by AlphaFold2<sup>12</sup> with the crystal structure (PDB ID:4PSU<sup>10</sup>). The protein domains of predicted structure are shown as green cartoon and crystal structure as white cartoon. The missing 83-85 loop fragment is colored red. The Root Mean Square Deviation (RMSD) value between the two structures is only 0.21 Å.



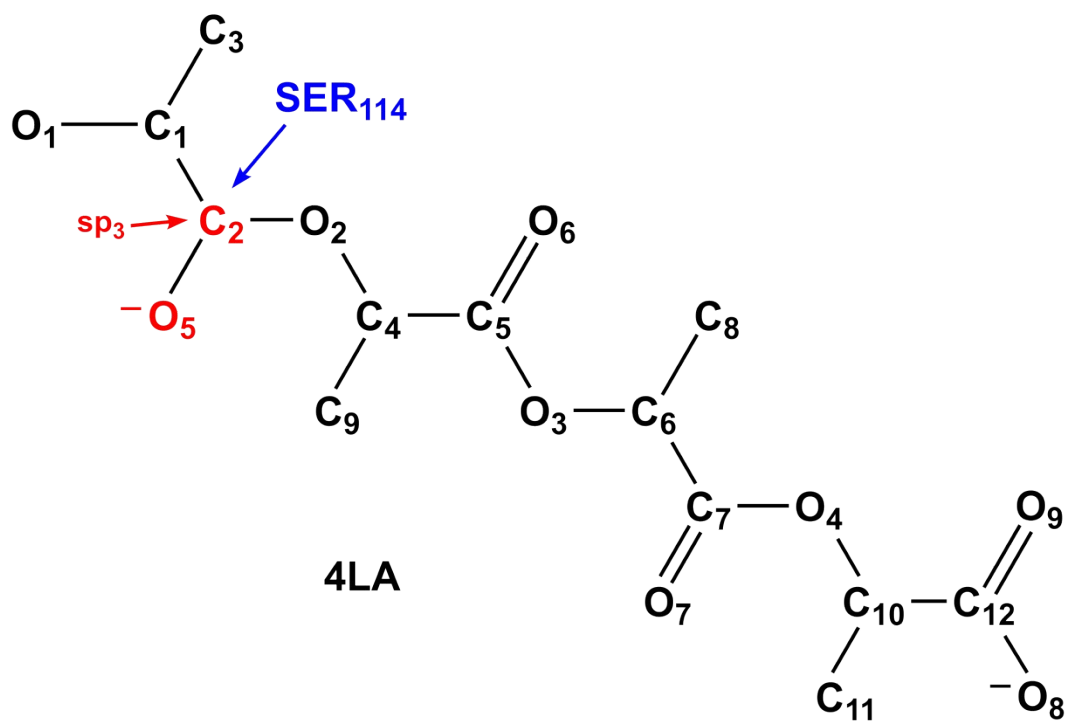
**Figure S8.** Comparison of the thermal stability of WT and the variant R5 based on MD simulations. (A) Mean root-mean-square fluctuation (RMSF) values for each residue over 100 ns. (B) The root-mean-square deviation (RMSD) values of the whole protein over 100 ns; (C) The radius of gyration (Rg) over 100 ns. (D) The solvent-accessible surface area (SASA) over 100 ns. Values correspond to the average of two independent 100 ns MD simulations.



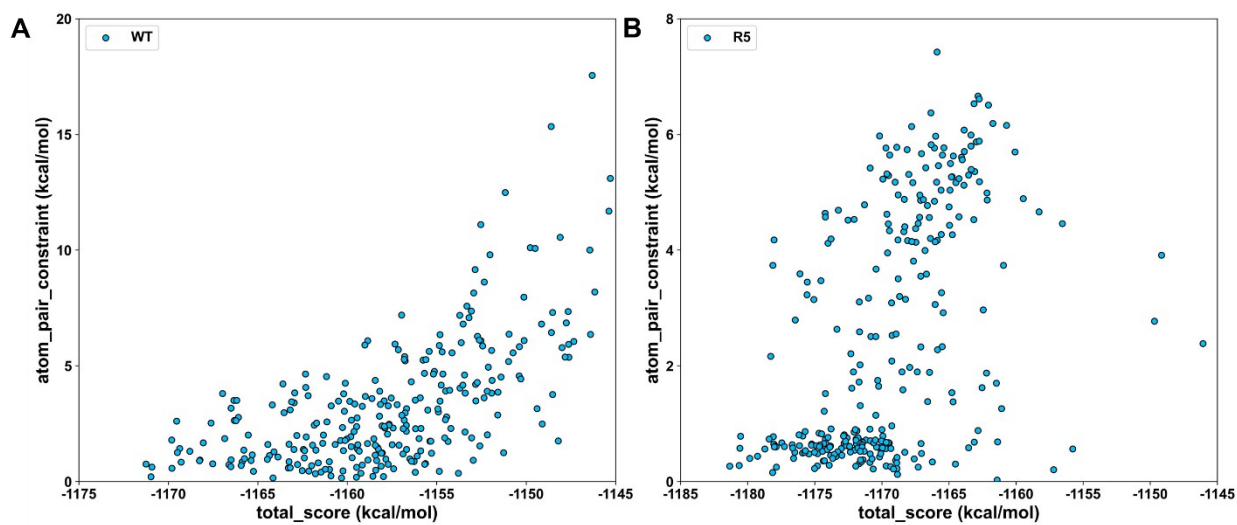
**Figure S9.** A single point mutation scan of V202 using the Rosetta cartesian\_ddG method to calculate the  $\Delta\Delta G_{\text{fold}}$  (kcal/mol).



**Figure S10.** Structural formula of the model substrate 4LA marked with heavy atom names. C<sub>2</sub> atom is the tetrahedral center carbon attacked nucleophilically by S114 and set to sp<sub>3</sub> hybridization. The negative charge on the O<sub>5</sub> atom is neutralized by the oxyanion hole residues.



**Figure S11.** Scatter plots of the total score and atom distance constraint score of WT (A) and R5 (B) docking results computed by Rosetta.<sup>13</sup> The “total\_score” represents the assessment of the total energy of docking models, while the “atom\_pair\_constraint” represents the penalty for the complex structure to be constrained.



## References

- 1 J. J. Weinstein, A. Goldenzweig, S. Hoch and S. J. Fleishman, *Bioinformatics*, 2021, **37**, 123–125.
- 2 M. Musil, J. Stourac, J. Bendl, J. Brezovsky, Z. Prokop, J. Zendulka, T. Martinek, D. Bednar and J. Damborsky, *Nucleic Acids Res*, 2017, **45**, W393–W399.
- 3 Y. Cui, Y. Chen, X. Liu, S. Dong, Y. Tian, Y. Qiao, R. Mitra, J. Han, C. Li, X. Han, W. Liu, Q. Chen, W. Wei, X. Wang, W. Du, S. Tang, H. Xiang, H. Liu, Y. Liang, K. N. Houk and B. Wu, *ACS Catal.*, 2021, **11**, 1340–1350.
- 4 D. Suplatov, D. Timonina, Y. Sharapova and V. Švedas, *Nucleic Acids Res*, 2019, **47**, W308–W314.
- 5 L. Josef, H.-F. Julia, L. Daniel and L. Peter, *Bioinformatics*, 2016, **32**, 1414–1416.
- 6 X. Gao, X. Dong, X. Li, Z. Liu and H. Liu, *Sci Rep*, 2020, **10**, 10330.
- 7 H. Park, P. Bradley, P. Greisen, Y. Liu, V. K. Mulligan, D. E. Kim, D. Baker and F. DiMaio, *J Chem Theory Comput*, 2016, **12**, 6201–6212.
- 8 J. Kari, M. Andersen, K. Borch and P. Westh, *ACS Catal.*, 2017, **7**, 4904–4914.
- 9 H. Schägger and G. von Jagow, *Analytical Biochemistry*, 1987, **166**, 368–379.
- 10 M. Hajighasemi, B. P. Nocek, A. Tchigvintsev, G. Brown, R. Flick, X. Xu, H. Cui, T. Hai, A. Joachimiak, P. N. Golyshin, A. Savchenko, E. A. Edwards and A. F. Yakunin, *Biomacromolecules*, 2016, **17**, 2027–2039.
- 11 F. H. Niesen, H. Berglund and M. Vedadi, *Nat Protoc*, 2007, **2**, 2212–2221.
- 12 M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov and M. Steinegger, *Nat Methods*, 2022, **19**, 679–682.
- 13 G. Lemmon and J. Meiler, *Methods Mol Biol*, 2012, **819**, 143–155.