Supporting Information

Precise redesign for improving enzyme robustness based on coevolutionary analysis and multidimensional virtual screening

Jie Luo[1], Chenshuo Song[1], Wenjing Cui[1], Qiong Wang[1], Zhemin Zhou*[1], Laichuang Han*[1]

[1] Key Laboratory of Industrial Biotechnology (Ministry of Education), School of Biotechnology, Jiangnan University, Wuxi, Jiangsu, China.

*corresponding authors：

Zhemin Zhou,

Key Laboratory of Industrial Biotechnology (Ministry of Education), School of Biotechnology, Jiangnan University.

Wuxi, Jiangsu, 214122, China

E-mail: zhmzhou@jiangnan.edu.cn

Laichuang Han,

Key Laboratory of Industrial Biotechnology (Ministry of Education), School of Biotechnology, Jiangnan University.

Wuxi, Jiangsu, 214122, China

E-mail: hanlaichuang@jiangnan.edu.cn

Contents

Supplementary Tables ... 3

Table S1. Plasmids and Bacterial strains used in this study. ... 3

Table S2. Primers used in this study. ... 4

Table S3. Coevolutionary residue pairs selected from EVcouplings software. ... 5

Table S4. Screening mutants for molecular dynamics simulation. ... 6

Table S5. Characterization of multidimensional dynamic parameters of mutants. ... 8

Table S6. Mutant grouping analysis. ... 9

Table S7. The list of mutations presented in mutants M4, M5 and M6. ... 10

Table S8. Protonation states of amino acids at different pH condition. ... 10

Supplementary Figures ... 11

Figure S1. Interaction details of selected coevolutionary residue pairs. ... 11

Figure S2. Evaluation of the folding free energy of mutants. ... 12

Figure S3. Correlation analysis of four parameters of mutants. ... 13

Figure S4. Original image of SDS-PAGE of purified WT and mutants. ... 14

Figure S5. Melting temperature derivative data of WT and mutants. ... 14

Figure S6. Kinetics characteristic analysis for WT and mutants. ... 15

Figure S7. The variation in RMSD between the WT and M6 at different temperatures and pH conditions. ... 15

Figure S8. Folding free energy of different enzymes. ... 16

Figure S9. The fluctuation of mutant sequences at 300K simulation. ... 16

Reference ... 17

Supplementary Tables

Table S1. Plasmids and Bacterial strains used in this study.

Plasmid and strains	Relevant characteristics	source
Plasmid		
pBP43-WapA-NK	*E.coli-B.subtilis* shuttle vector, Amp^R in *Escherichia coli*, Kan^R in *Bacillus subtilis*, RepB, P43,signal peptide-WapA	Lab stock
Strains		
E. coli JM109	*recA*1, *supE*44 *endA*1 *hsdR*17 ($^{r-}$k,m^{+} k) *gyrA*96 *relA*1 thi (*lac-proAB*) F'[*traD*36 *proAB*$^{+}$ *lacI*^{q} *lacZ* ΔM15]	Lab stock
B. subtilis comK	Derive from *B.subtilis 168*, *lacA::erm R-PxylA-comK*	Lab stock[1]

Table S2. Primers used in this study.

Primer	Primer Sequence (5'-3')
P-T22F-F	CACTCTCAAGGCTACTTTGGCTCTAACGTAAAAGTAGCTGTTATCG
P-T22L-F	CACTCTCAAGGCTACCTTGGCTCTAACGTAAAAGTAGCTGTTATCG
P-T22Y-F	CTTCACTCTCAAGGCTACTATGGCTCTAACGTAAAAGTAGCTGTTATC
P-T22-R	GTAGCCTTGAGAGTGAAGAGCC
P-S87-R	TGGCGCTACGCCCAGAAC
P-S87F-F	GGGCGTAGCGCCAAGCGCATCATTATATGCAGTAAAAGTGCTTG
P-S87L-F	GGGCGTAGCGCCACTCGCATCATTATATGCAGTAAAAGTGCTTG
P-S87W-F	GGGCGTAGCGCCATGGGCATCATTATATTGTGTAAAAGTGCTTGATTC
P-S87Y-F	GGGCGTAGCGCCATACGCATCATTATATGCAGTAAAAGTGCTTG
P-K27F-F	CACAGGCTCTAACGTATTTGTAGCTGTTATCGACAGCGGAATTG
P-K27L-F	CACAGGCTCTAACGTACTGGTAGCTGTTATCGACAGCGGAATTG
P-K27N-F	CACAGGCTCTAACGTAAATGTAGCTGTTATCGACAGCGGAATTG
P-K27H-F	CACAGGCTCTAACGTACATGTAGCTGTTATCGACAGCGGAATTG
P-K27R-F	CACAGGCTCTAACGTAAGAGTAGCTGTTATCGACAGCGG
P-K27Y-F	CACAGGCTCTAACGTATATGTAGCTGTTATCGACAGCGGAATTG
P-K27-R	TACGTTAGAGCCTGTGTAGCCTTG
P-S89F-F	GCGCCAAGCGCATTTTTATATGCAGTAAAAGTGCTTGATTCAACAGG
P-S89L-F	GCGCCAAGCGCACTGTTATATGCAGTAAAAGTGCTTGATTCAACAGG
P-S89W-F	GCGCCAAGCGCATGGTTATATGCAGTAAAAGTGCTTGATTCAACAGG
P-S89Y-F	GCGCCAAGCGCATATTTATATGCAGTAAAAGTGCTTGATTCAACAGG
P-S89-R	TGCGCTTGGCGCTACG
P-H67A-F	GTTCTCACGGTACGGCGGTCGCCGGTACGATTGCC
P-H67A-R	CGTACCGTGAGAACTGCCG
P-S207A-F	GGCGTGTCCATCCAAGCGACACTTCCTGGAGGCACTTAC
P-S207V-F	GGCGTGTCCATCCAAGTTACACTTCCTGGAGGCACTTAC
P-S207-R	TTGGATGGACACGCCAGGAG
P-N269W-F	CTTTCTACTATGGAAAAGGGTTAATCTGGGTACAAGCAGCTGCACAATAAG
P-T255F-R	CCCTTTTCCATAGTAGAAAGAGTTTCCAAGATAAAATGCAGTGCTTTCTAAACGATC

	A
P-N269D-F	CTTTCTACTATGGAAAAGGGTTAATCGACGTACAAGCAGCTGCACAATAAG
P-N269I-F	CTTTCTACTATGGAAAAGGGTTAATCATTGTACAAGCAGCTGCACAATAAG
P-T255Y-R	CCCTTTTCCATAGTAGAAAGAGTTTCCAAGATAATATGCAGTGCTTTCTAAACGATC A
P-K141V-F	ATCAACTACTGTTTTCAGCGCTGTAG
P-K141V-R	ATCAACTACTGTTTTCAGCGCTGTAG
P-D140A-F	CTGAAAACAGTAGTTGCGAAAGCGGTTTCCAGCGGTATC
P-D140A-R	AACTACTGTTTTCAGCGCTGTAG
P-S173L-F	CCCTGCAAAATATCCTCTGACTATTGCAGTAGGTGCGGTAAAC
P-S173L-R	AGGATATTTTGCAGGGTAGCCG

Table S3. Coevolutionary residue pairs selected from EVcouplings software.

Residue pairs	Score	Distance between Cα (Å)
L96-G102	42.85	4.8
L209-G215	33.92	6.6
T22-S87	33.44	5.3
A230-V270	31.66	5.6
V93-W113	31.54	8.1
A230-I268	29.39	8.3
A179-F189	28.95	8.1
K27-Y91	28.88	7.4
E112-K141	27.51	8.4
V28-I72	27.10	8.3
G178-M199	27.02	4.5
K27-S89	26.92	5.0
A169-A176	26.73	4.1
D120-L235	26.67	8.4
D140-S173	26.52	6.4
A187-V203	26.45	6.8
H67-M222	26.16	8.0
H67-S207	25.72	5.4
V180-M199	25.09	6.6
T255-N269	23.78	6.8

Table S4. Screening mutants for molecular dynamics simulation.

Coevolutionary site	Mutants	ddG (REU)	Molecular dynamic simulation
L96-G102	G102F	-1.8	√
	G102W	-1.6	
T22S87	T22FS87F	-2.1	√
	T22FS87L	-2.0	√
	T22FS87W	-2.0	√
	T22FS87Y	-2.1	√
	T22HS87F	-1.7	
	T22HS87L	-1.7	
	T22HS87Y	-1.7	
	T22LS87F	-1.9	
	T22LS87L	-1.6	√
	T22LS87W	-1.9	√
	T22LS87Y	-1.9	
	T22QS87L	-1.6	
	T22YS87F	-2.2	√
	T22YS87L	-2.1	√
	T22YS87M	-1.6	
	T22YS87W	-2.2	√
	T22YS87Y	-2.2	√
E112K141	E112WK141V	-1.6	√
K27-S89	K27FS89F	-2.2	√
	K27FS89L	-2.6	√
	K27FS89M	-1.9	
	K27FS89W	-2.1	√
	K27FS89Y	-2.2	√
	K27HS89F	-1.9	
	K27HS89L	-2.1	√
	K27HS89W	-1.7	
	K27HS89Y	-1.9	
	S89F	-1.7	
	S89L	-2.0	√
	S89W	-1.7	
	S89Y	-1.8	
	K27LS89E	-1.7	
	K27LS89F	-2.5	√
	K27LS89H	-1.7	
	K27LS89I	-1.6	
	K27LS89L	-2.1	√
	K27LS89N	-1.7	

	K27LS89Q	-1.9	
	K27LS89V	-1.8	
	K27LS89W	-2.5	√
	K27LS89Y	-2.6	√
	K27MS89F	-1.7	
	K27MS89L	-1.7	
	K27MS89W	-1.6	
	K27MS89Y	-1.7	
	K27NS89F	-1.9	
	K27NS89L	-2.1	√
	K27NS89N	-1.6	
	K27NS89W	-1.8	
	K27NS89Y	-2.0	√
	K27RS89F	-1.7	
	K27RS89L	-2.0	√
	K27RS89W	-1.6	√
	K27RS89Y	-1.7	
	K27WS89F	-1.6	
	K27WS89L	-2.0	√
	K27WS89W	-1.7	
	K27WS89Y	-1.7	
	K27YS89F	-2.2	√
	K27YS89L	-2.6	√
	K27YS89M	-1.8	
	K27YS89W	-2.0	√
	K27YS89Y	-2.2	√
D140-S173	D140AS173L	-1.6	√
H67-M222	H67AM222F	-1.6	√
H67-S207	H67AS207A	-2.1	√
	H67AS207T	-2.7	√
	H67AS207V	-3.2	√
	H67FS207D	-1.8	
	H67HS207N	-1.6	
T255N269	T255FN269W	-2.2	√
	T255FN269Y	-1.6	
	T255IN269D	-1.7	√
	T255IN269L	-1.6	√
	T255WN269W	-1.7	
	T255YN269I	-1.7	√
	T255YN269L	-1.7	
	T255YN269W	-1.9	

Note: REU means the Rosetta Energy Unit.

Table S5. Characterization of multidimensional dynamic parameters of mutants.

Entry	RMSD(Å)	ΔRMSD(Å)	Rg(Å)	ΔRg(Å)	H-bond	ΔH-bond
G102F	1.99±0.36	0.36	17.16±0.09	0.04	46.72±6.03	-2.03
T22FS87F	1.67±0.23	0.04	17.13±0.09	0.01	49.16±6.05	0.41
T22FS87L	1.44±0.19*	-0.19	17.01±0.07*	-0.11	50.26±6.16*	1.51
T22FS87W	1.46±0.19*	-0.17	17.06±0.08*	-0.06	51.23±6.11*	2.48
T22FS87Y	1.57±0.19	-0.06	17.10±0.08	-0.02	48.55±6.61	-0.20
T22LS87L	1.63±0.26	0.00	17.12±0.10	0.00	49.56±6.23*	0.81
T22LS87W	1.44±0.18*	-0.19	17.05±0.08*	-0.07	49.13±6.02	0.38
T22YS87F	1.50±0.19	-0.13	17.08±0.08	-0.04	49.83±5.97*	1.08
T22YS87L	1.41±0.13*	-0.22	17.05±0.07*	-0.07	49.80±5.96*	1.05
T22YS87W	1.46±0.16*	-0.17	17.06±0.07*	-0.06	48.12±5.99	-0.63
T22YS87Y	1.43±0.19*	-0.20	17.05±0.08	-0.07	49.76±6.22*	1.01
E112WK141V	1.53±0.25	-0.10	17.06±0.08*	-0.06	49.45±6.04	0.70
K27FS89F	1.66±0.25	0.03	17.13±0.09	0.01	48.29±6.17	-0.46
K27FS89L	1.63±0.22	0.00	17.11±0.09	-0.01	49.44±6.17	0.69
K27FS89W	1.65±0.20	0.02	17.10±0.08	-0.02	48.98±6.00	0.23
K27FS89Y	1.42±0.13*	-0.21	17.07±0.07*	-0.05	48.75±5.97	0.00
K27HS89L	1.58±0.35	-0.05	17.10±0.13	-0.02	49.11±6.13	0.36
S89L	1.40±0.11*	-0.23	17.04±0.07*	-0.08	48.74±6.16	-0.01
K27LS89F	1.52±0.22	-0.11	17.08±0.09	-0.04	48.86±5.99	0.11
K27LS89L	1.45±0.16*	-0.18	17.05±0.07*	-0.07	48.56±6.18	-0.19
K27LS89W	1.39±0.13*	-0.24	17.03±0.07*	-0.09	48.80±6.06	0.05
K27LS89Y	1.37±0.13*	-0.26	17.02±0.06*	-0.10	49.02±6.08	0.27
K27NS89L	1.58±0.31	-0.05	17.04±0.08*	-0.08	48.02±6.08	0.27
K27NS89Y	1.41±0.17*	-0.22	17.04±0.08*	-0.08	49.77±6.04*	1.02
K27RS89L	1.63±0.18	0.00	17.11±0.08	-0.01	50.09±6.15*	1.34
K27RS89W	1.81±0.36	0.18	17.09±0.08	-0.03	49.46±6.03	0.71
K27WS89L	1.66±0.35	0.03	17.11±0.12	-0.01	48.14±6.06	-0.61
K27YS89F	1.50±0.15	-0.13	17.08±0.08	-0.04	48.34±6.25	-0.41
K27YS89L	1.36±0.12*	-0.27	17.04±0.07*	-0.08	49.96±6.20*	1.21
K27YS89W	1.31±0.17*	-0.32	17.03±0.07*	-0.09	49.08±6.07	0.33
K27YS89Y	1.41±0.13*	-0.22	17.04±0.06*	-0.08	49.86±6.13*	1.11
D140AS173L	1.59±0.19	-0.04	17.10±0.08	-0.02	48.93±6.03	0.18
H67AS207A	1.70±0.16	0.07	17.07±0.08*	-0.05	49.54±6.20*	0.79
H67AS207T	1.60±0.25	-0.03	17.07±0.09*	-0.05	50.04±6.20*	1.29
H67AS207V	1.49±0.18	-0.14	17.05±0.08*	-0.07	49.79±6.06*	1.04
H67AM222F	1.58±0.19	-0.05	17.11±0.08	-0.01	49.49±5.97	0.74
T255FN269W	1.65±0.21	0.02	17.09±0.07	-0.03	48.30±6.00	-0.45
T255IN269D	1.59±0.20	-0.04	17.12±0.08	0.00	50.01±6.13*	1.26
T255IN269L	1.45±0.23*	-0.18	17.05±0.09*	-0.07	49.08±6.11	0.33

T255YN269I	1.62±0.22	-0.01	17.09±0.07	-0.03	49.42±6.08	0.66

The parameter of MD simulations employed the Kruskal-Wallis test for analysis at 5% level.

* Denotes significant differences in values corresponding to the mutant compared to the WT (P<0.05).

Table S6. Mutant grouping analysis.

Overlapping mutants in three groups	Others	Filter out
T22FS87L	T22LS87W	G102F
T22FS87W	T22YS87W	T22FS87Y
T22YS87L	T22YS87F	K27FS89F
T22YS87Y	K27LS89L	K27FS89W
K27NS89Y	T22FS87F	K27WS89L
K27YS89L	T22LS87L	K27HS89L
K27YS89Y	K27LS89Y	D140AS173L
H67AS207V	K27YS89W	T255FN269W
	K27NS89L	
	K27LS89W	
	K27LS89F	
	K27FS89L	
	K27RS89L	
	K27RS89W	
	K27YS89F	
	K27FS89Y	
	S89L	
	H67AM222F	
	H67AS207T	
	H67AS207A	
	E112WK141V	
	T255IN269L	
	T255IN269D	
	T255YN269I	

Overlapping mutants in three groups means the intersection of three sets in a Venn diagram. Others means the mutants outside of three sets but at least one set is satisfied. Filter out means the mutants do not belong to any set.

Table S7. The list of mutations presented in mutants M4, M5 and M6.

Entry	Mutations
M4[2]	P57CS62AA92CN181DY217KN218SN259P
M5	T22FS87LT255IN269D
M6	T22FP57CS62AS87LA92CN181DY217KN218ST255IN269DN259P

Table S8. Protonation states of amino acids at different pH condition.

pH	Protonation states	Residues	
		WT	M6
pH3	ASPP	32, 41, 60, 120, 248	32, 41, 60, 120, 181, 248, 269
	GLUP	54, 112, 156, 195, 251	112, 156, 251
	HSP	17, 39, 64, 67, 226, 238	17, 39, 64, 67, 226, 238
	HSD	-	-
	HSE	-	-
pH7	ASPP	-	-
	GLUP	-	-
	HSP	64	64
	HSD	67, 226	226
	HSE	17, 39, 238	17, 39, 67, 238

Note: The Protonation states of each residue were determined by the PDB2PQR (https://server.poissonboltzmann.org/pdb2pqr)[3]. ASPP and GLUP means the protonated Asp and Glu, respectively. HSP means the protonated His; HSD means the neutral His, proton on ND1; HSE means the neutral His, proton on NE2.

Supplementary Figures

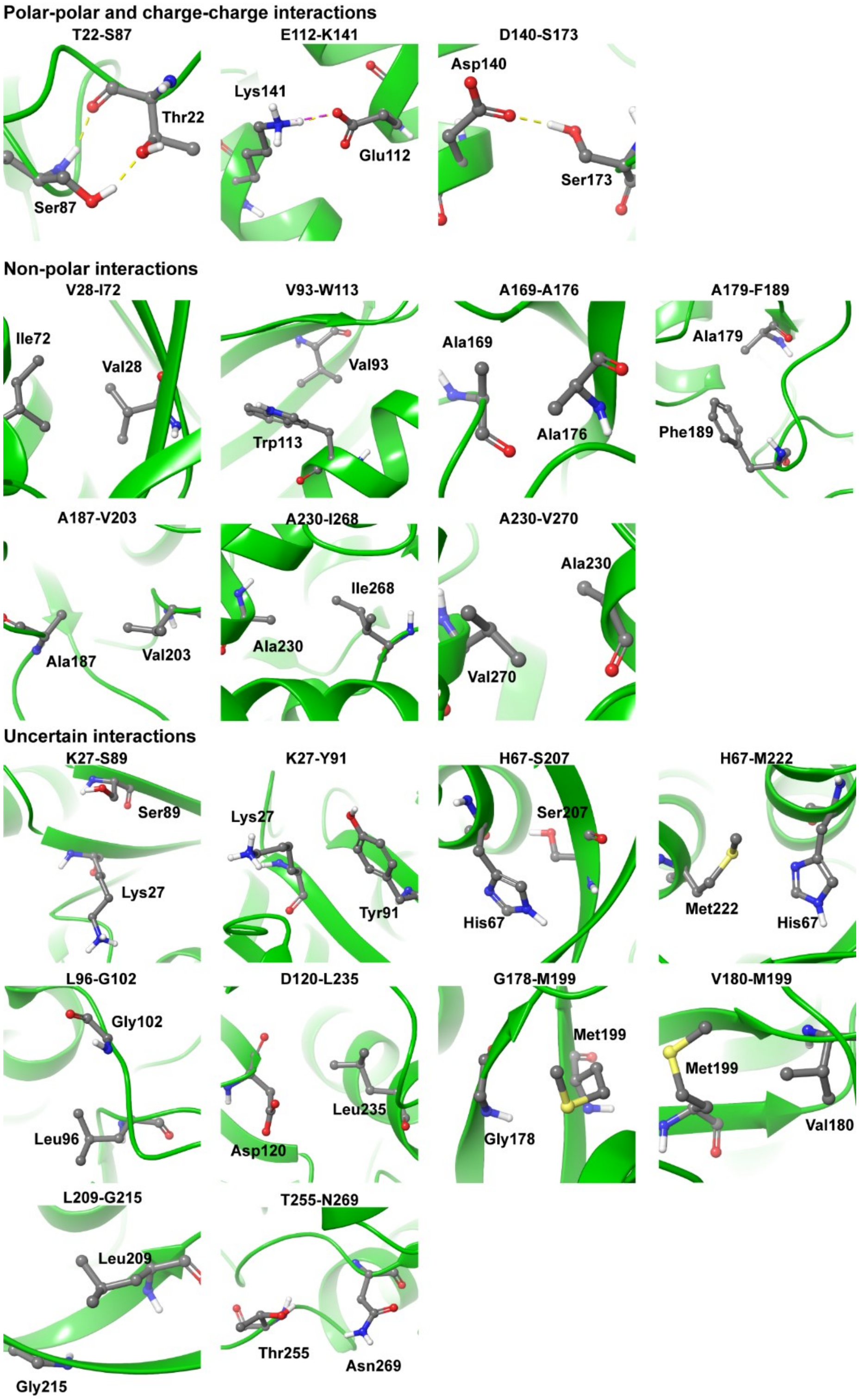

Figure S1. Interaction details of selected coevolutionary residue pairs.

Hydrogen bonding interactions are highlighted by dashed lines in yellow, and salt-bridge interactions are in magenta.

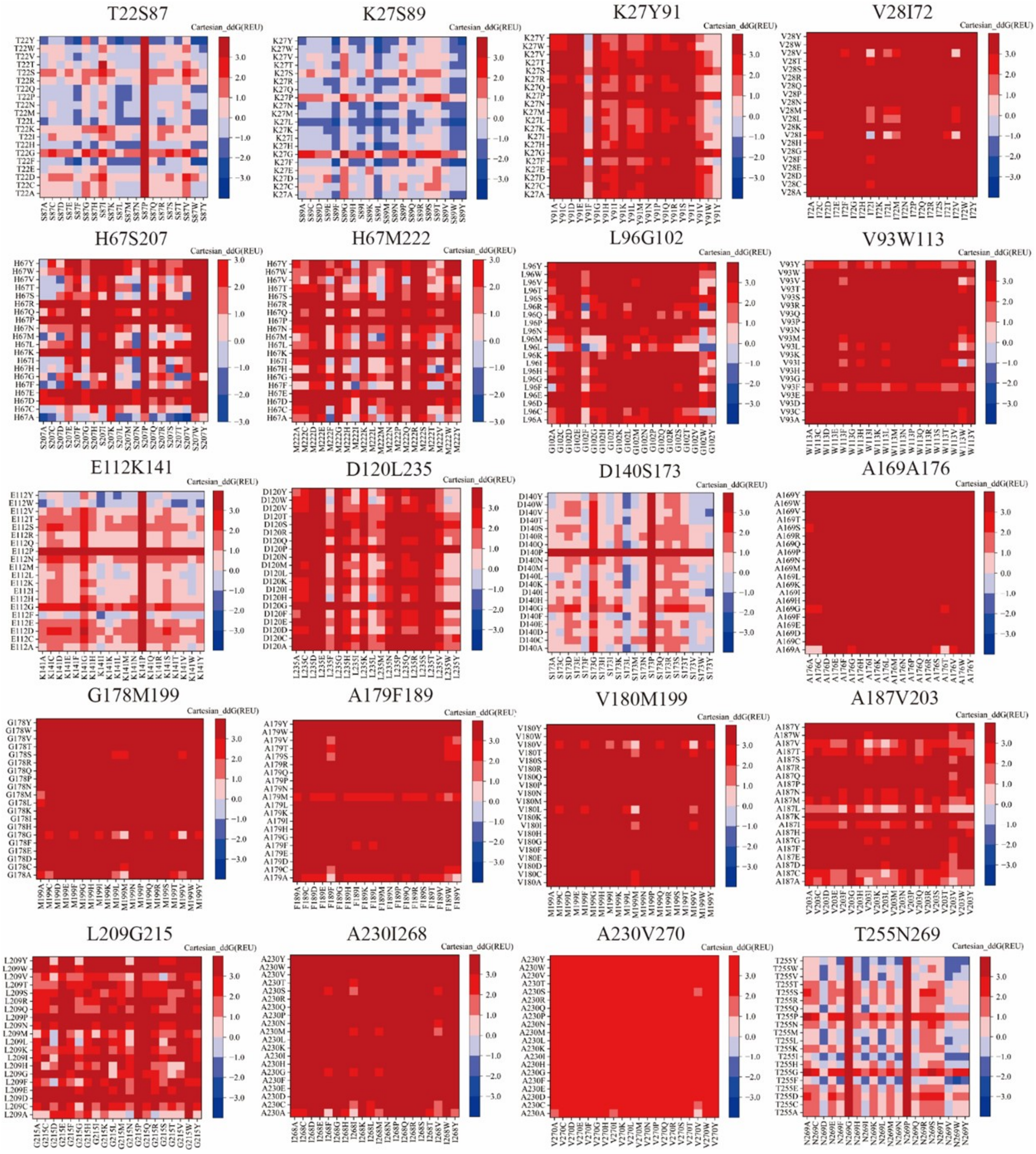

Figure S2. Evaluation of the folding free energy of mutants.

Each ΔΔG of the total 7980 mutations in the virtual mutant library were calculated. REU means the Rosetta Energy Unit, Range from -3.2 to 3.0.

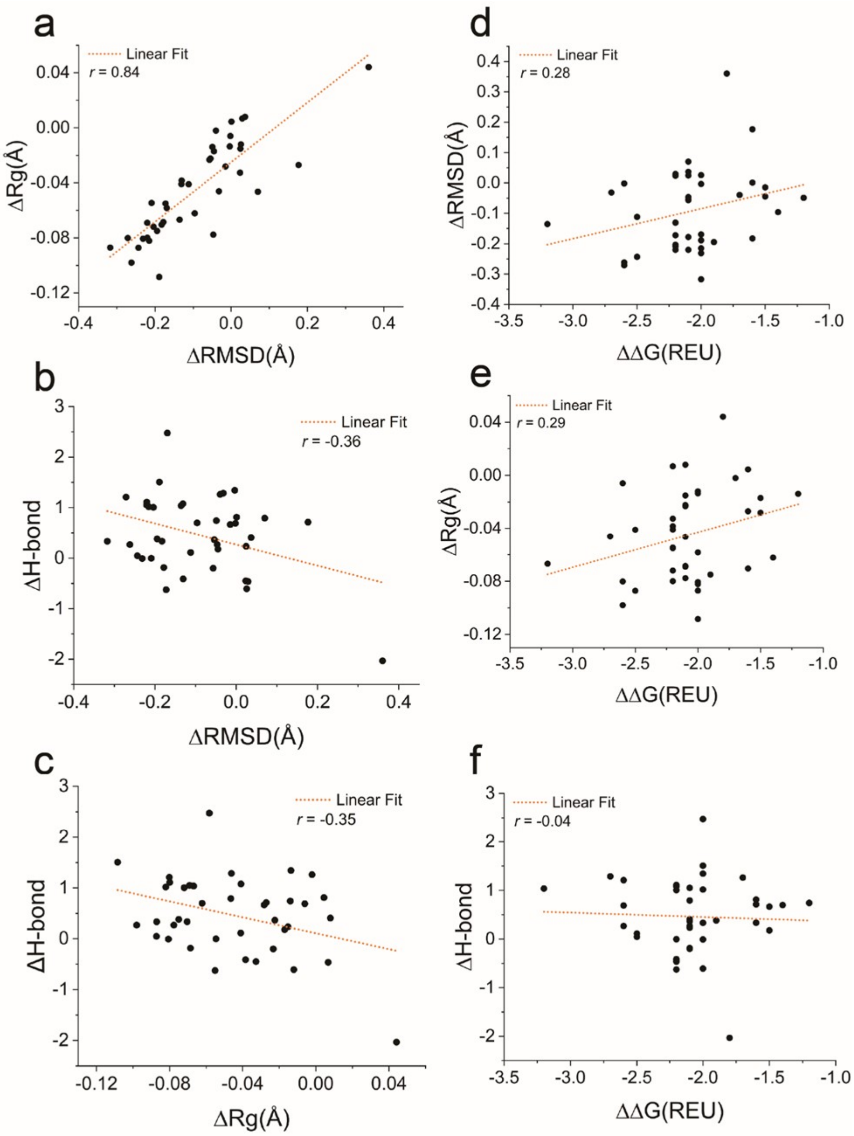

Figure S3. Correlation analysis of four parameters of mutants.

Pearson's chi squared test based on mutant parameter values.

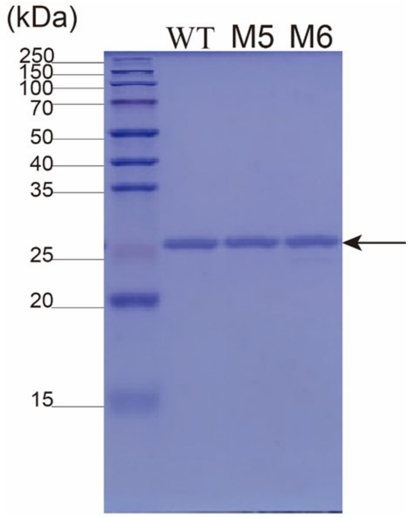

Figure S4. Original image of SDS-PAGE of purified WT and mutants.

The black arrow indicates the position of Nattokinase on SDS-PAGE gel.

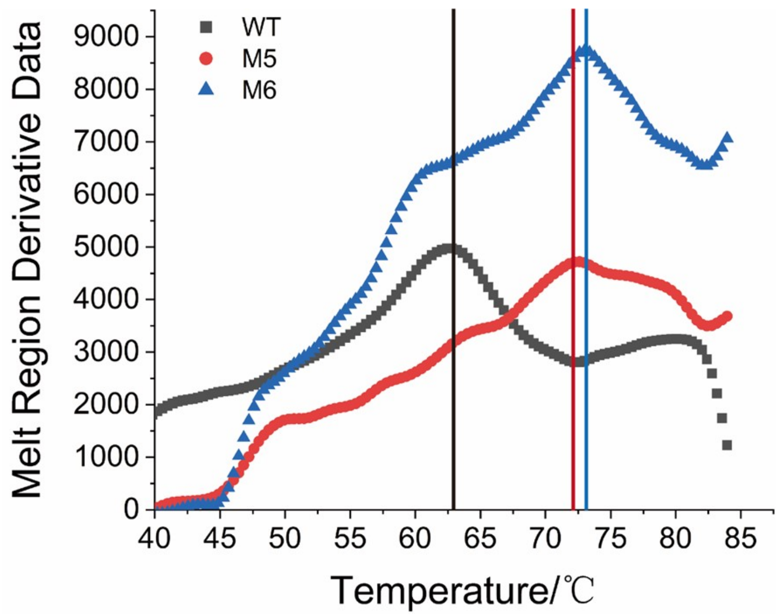

Figure S5. Melting temperature derivative data of WT and mutants.

The fluorescence was tracked using a built-in detector. The StepOnePlus™ real-time PCR system calculated the derived melting temperature data. T_m is represented by the three different vertical lines indicating the temperatures at which absolute fluorescence change rates decrease sharply.

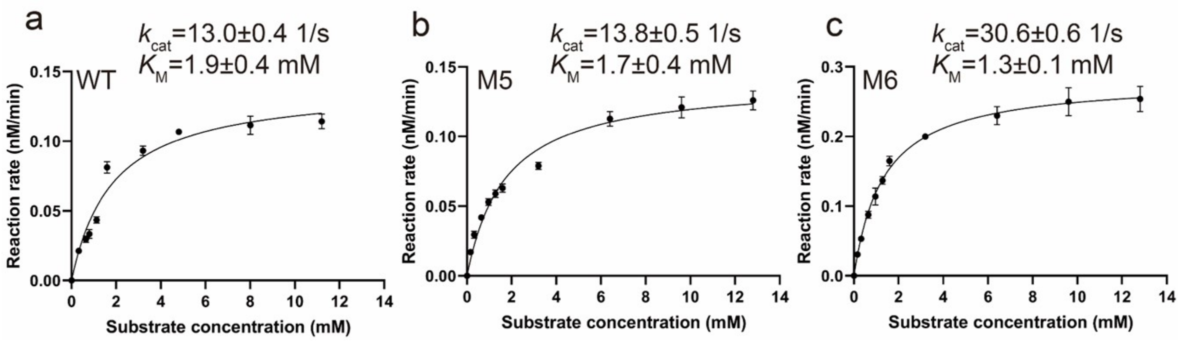

Figure S6. Kinetics characteristic analysis for WT and mutants.

The reaction rate was represented in nM.min^{-1}per nM enzyme.

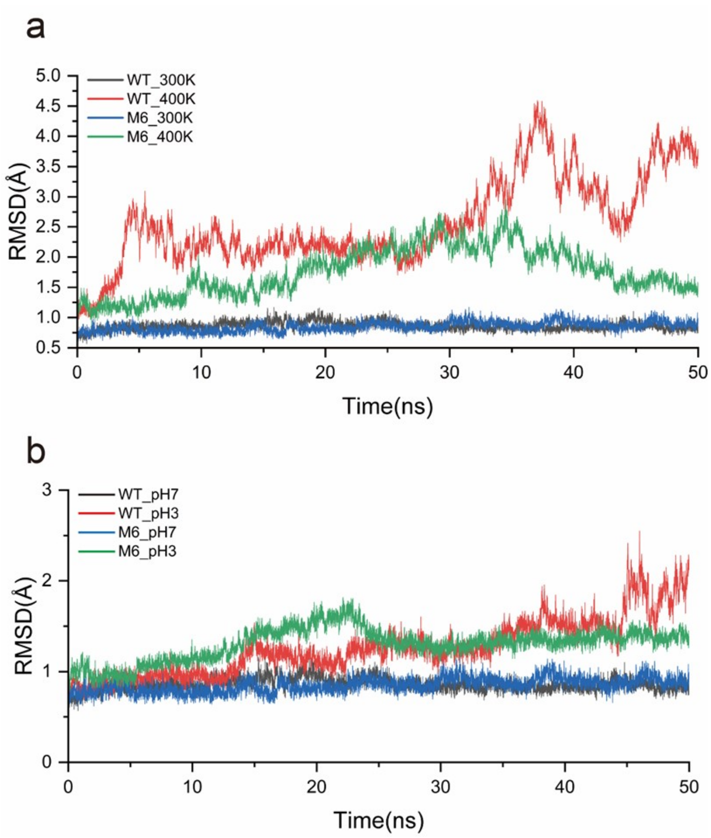

Figure S7. The variation in RMSD between the WT and M6 at different temperatures and pH conditions.

Comparison of the RMSD values between the WT and mutant M6 under various simulated temperatures (a) and pH (b) condition.

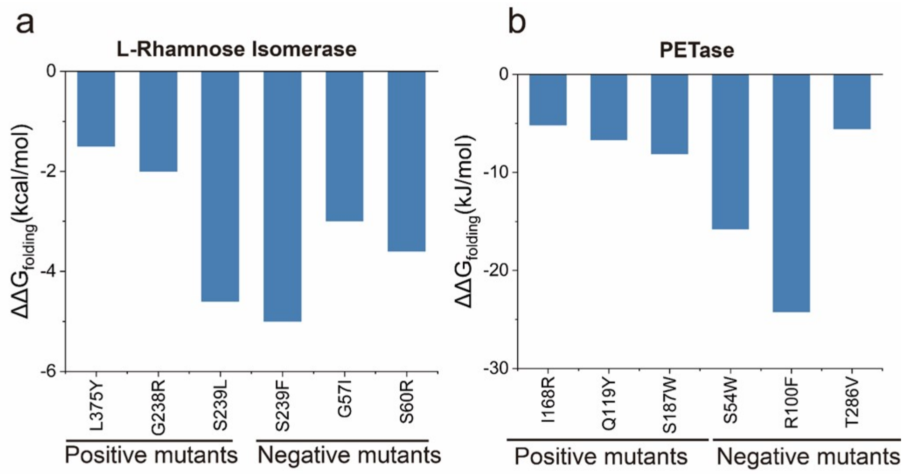

Figure S8. Folding free energy of different enzymes.

Folding free energy of L-Rhamnose Isomerase (a) [4] and PETase(b) [5].We selected two enzymes from existing reports to validate the multidimensional parameter screening system.

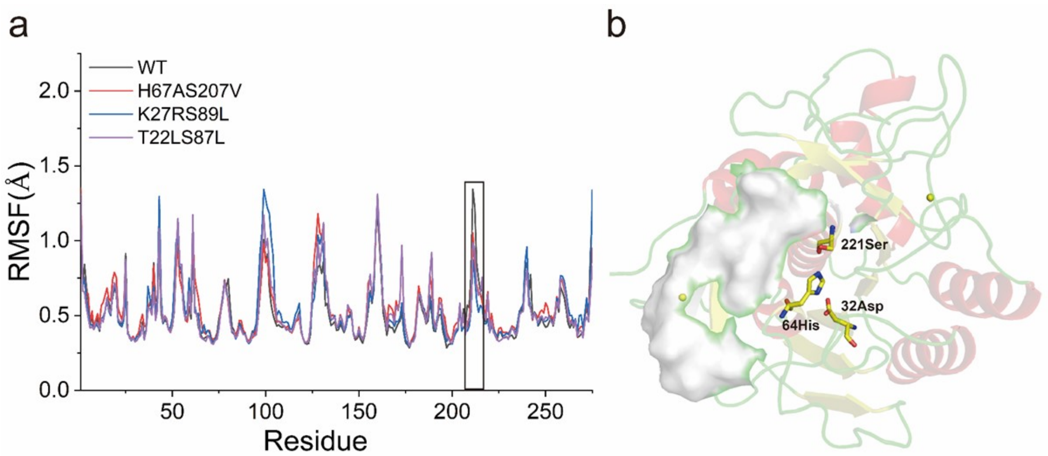

Figure S9. The fluctuation of mutant sequences at 300K simulation.

(a) Comparison of RMSF values of WT and mutants under 300K condition. Black box expression the region from 210 to 220 exhibits increased rigidity in mutants compared to the WT. (b) The three-dimensional structure of NK. The gray surface represents region from 210 to 220. 32Asp, 64His, 221Ser represent catalytic region of NK.

Reference

1. L. Han, W. J. Cui, F. Y. Suo, S. N. Miao, W. L. Hao, Q. Q. Chen, J. L. Guo, Z. M. Liu, L. Zhou and Z. M. Zhou, Development of a novel strategy for robust synthetic bacterial promoters based on a stepwise evolution targeting the spacer region of the core promoter in Bacillus subtilis, *Microb Cell Fact*, 2019, **18**.
2. J. Luo, C. S. Song, W. J. Cui, L. C. Han and Z. M. Zhou, Counteraction of stability-activity trade-off of Nattokinase through flexible region shifting, *Food Chem*, 2023, **423**.
3. E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. H. Chen, K. Liles, M. J. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E. Nielsen, T. Head-Gordon, W. H. Geng, R. Krasny, G. W. Wei, M. J. Holst, J. A. McCammon and N. A. Baker, Improvements to the APBS biomolecular solvation software suite, *Protein Sci*, 2018, **27**, 112-128.
4. M. J. Wei, X. Gao, W. Zhang, C. Li, F. P. Lu, L. J. Guan, W. D. Liu, J. W. Wang, F. H. Wang and H. M. Qin, Enhanced Thermostability of L-Rhamnose Isomerase for D-Allose Synthesis by Computation-Based Rational Redesign of Flexible Regions, *J Agr Food Chem*, 2023, **71**, 15713-15722.
5. Y. L. Cui, Y. C. Chen, X. Y. Liu, S. J. Dong, Y. E. Tian, Y. X. Qiao, R. Mitra, J. Han, C. L. Li, X. Han, W. D. Liu, Q. Chen, W. Q. Wei, X. Wang, W. B. Du, S. Y. Tang, H. Xiang, H. Y. Liu, Y. Liang, K. N. Houk and B. Wu, Computational Redesign of a PETase for Plastic Biodegradation under Ambient Condition by the GRAPE Strategy, *Acs Catal*, 2021, **11**, 1340-1350.