Supplementary Material

for

How mutational epistasis impairs predictability in protein evolution and design

Charlotte M. Miton[1] and Nobuhiko Tokuriki[1,*].

Table of Contents

A. Supplementary figures

Figure S1: Nine evolutionary trajectories toward new function.
Figure S2: Round-per-round epistatic constraints in evolutionary trajectories.
Figure S3: Localization of epistatic mutations on the protein crystal structures.
Figure S4: Evolution of mutational distribution, round-per-round.
Figure S5: Further examples of direct and indirect interactions causing epistasis.

B. Supplementary tables

Table S1: Summary of the mutational effects and distances for each evolutionary trajectory.
Table S2: Summary of round-per-round improvements for each evolutionary trajectory.

A. Supplementary figures

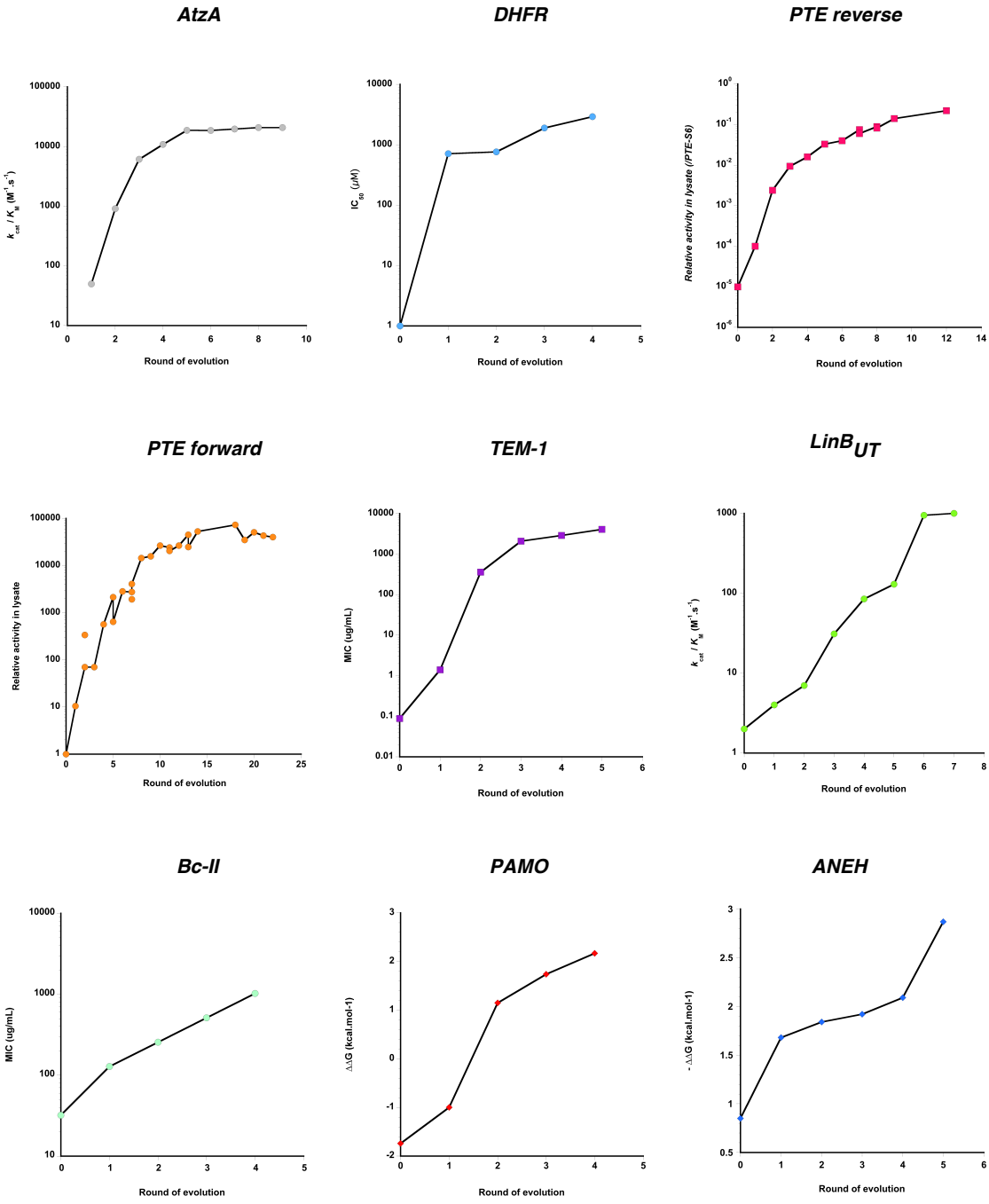

Figure S1. Nine evolutionary trajectories toward new function. Five trajectories are constrained by diminishing returns, *i.e.* the first steps yield large improvements that vanished toward the end of the evolution: AtzA, DHFR, PTE-rev, PTE-for and TEM-1. By contrast, $LinB_{UT}$, Bc-II, PAMO and ANEH display more linear trajectories, suggesting that some evolutionary potential may still exist.

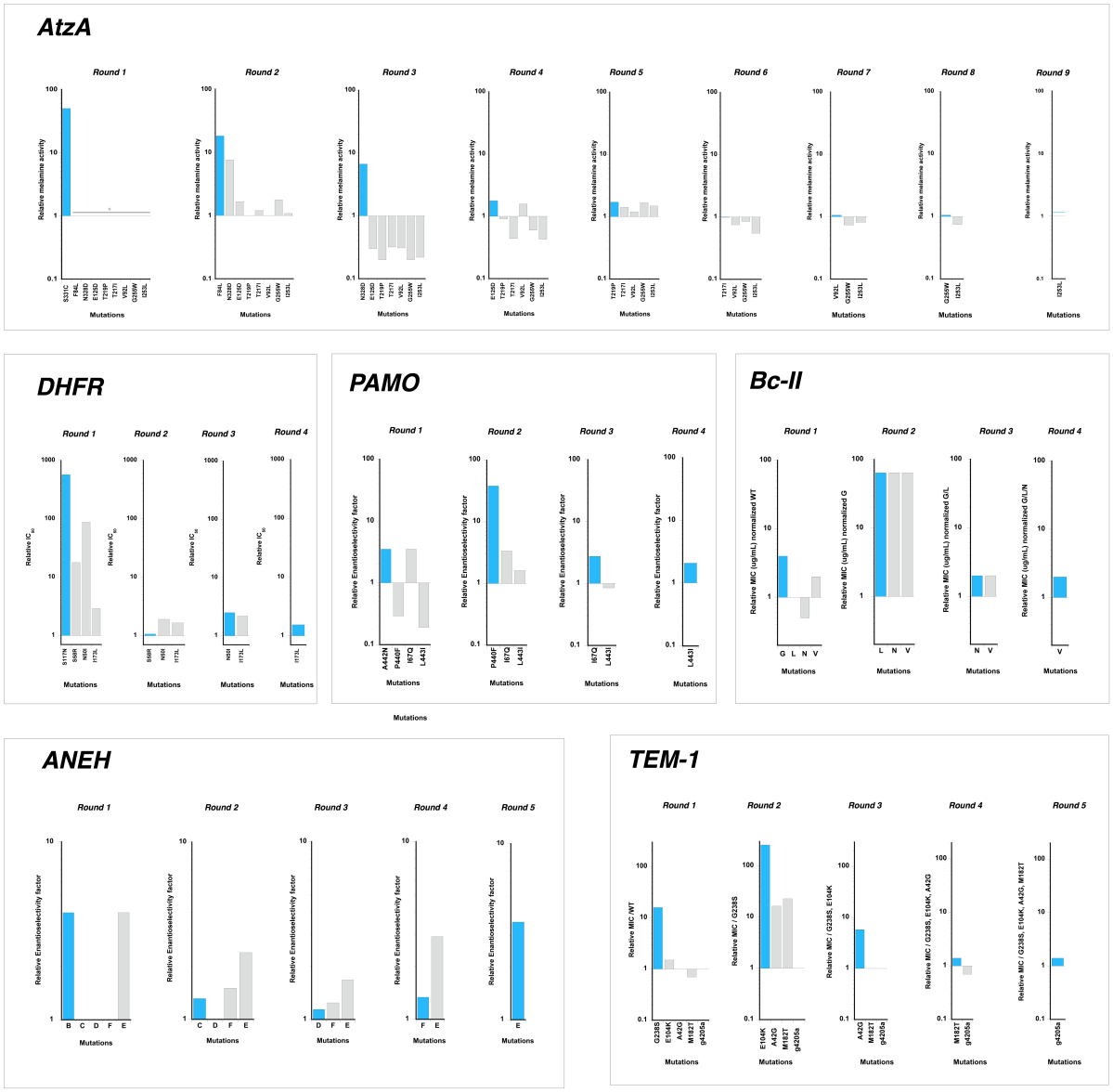

Figure S2. Round-per-round epistatic constraints in evolutionary trajectories. Fold change in activity for all possible mutational combinations in the evolutionary trajectories of AtzA, DHFR, PAMO, Bc-II, ANEH and TEM-1. Mutations fixed at the next round are shown as a blue bar. All possible combinations, not selected but still available at a given round, are shown as grey bars. Non-detectable activity/binding levels are shown as stars.

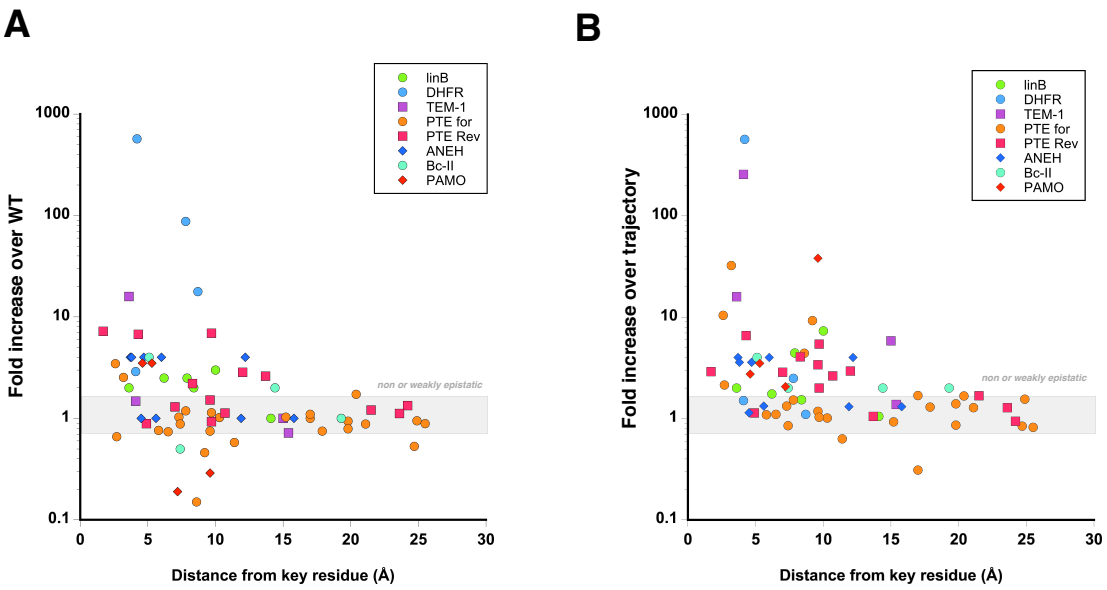

Figure S3. Localization of epistatic mutations on the protein crystal structures. Mutational effect *versus* distance between each mutation and a key active site residue on (A) the wild-type background, and (B) as it occurs in the trajectory. Distances are measured in angstroms (Å). The grey rectangle defines mutations for which the epistasis effects falls between 0.7 and 1.5-fold change, *i.e.* non-functional or non-epistatic mutations.

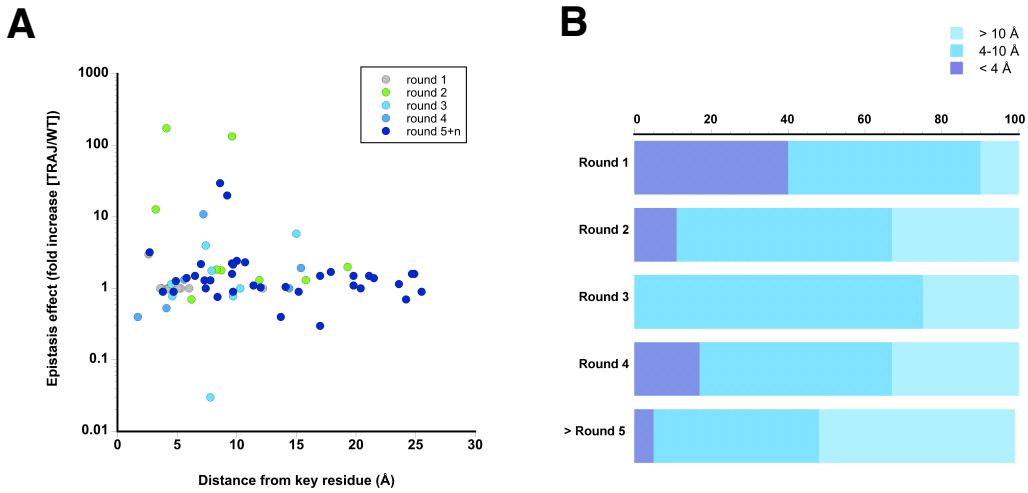

Figure S4. Evolution of mutational distribution, round-per-round. (A) Epistasis effect of each mutation *versus* distance from a key catalytic residue colored by round of appearance in the evolutionary trajectory (from round 1, to 5 and more, grey to dark blue). (B) Percentage of mutations localized within less than 4 Å (dark blue), between 4-10 Å (marine blue) and over 10 Å (light blue) at each round of evolution for 59 mutations across nine trajectories.

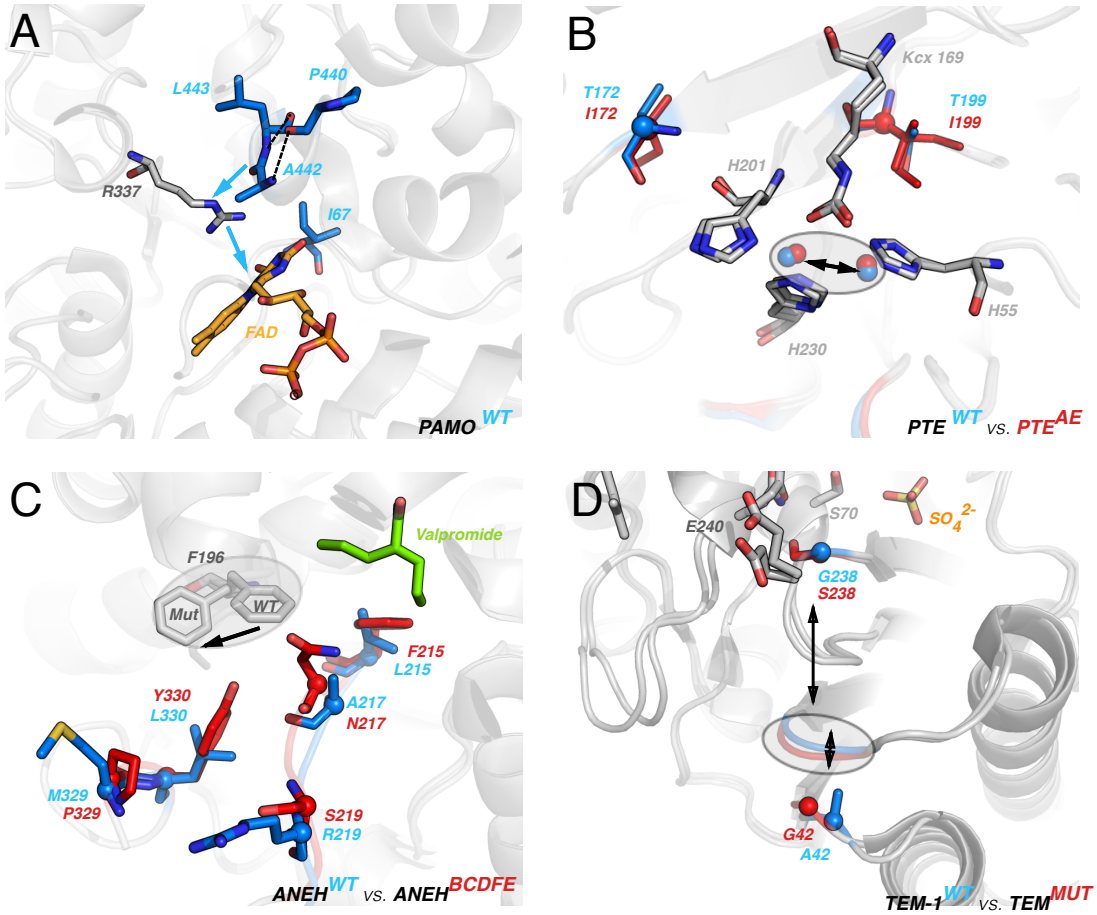

Figure S5. Further examples of direct and indirect interactions causing epistasis. (A) In PAMO^{WT}, the proximity of mutations P440F, A442N and L443I suggest that they interact (blue arrows) with each other (through hydrogen bonding interactions, dotted black lines). Yet these mutations may not directly interact with the substrate or the FAD cofactor (orange sticks), rather they appear to contribute to a functional change *via* repositioning of R337 [1]. (B) Similar to Bc-II evolution, PTE^{AE} (evolved mutant) exhibits a shift in distance between two metal ions that seems to result from an epistatic change in dynamics, acted upon by T199I and T172I. (C) Indirect interactions may cause epistasis between clusters B and C in ANEH, *via* a shift of active site residues such as F196 (grey sticks) between ANEH^{WT} and ANEH^{BCDFE}. (D) Long-range effects in TEM-1 between mutations A42G and G238S may propagate through a change in loop 266-271 conformation.

B. Supplementary tables

Table S1. Summary of the mutational effects and distances for each evolutionary trajectory.

AtzA

			WT	TRAJECTORY
Round of appearance	Mutation	k_{cat}/K_M ($M^{-1}.s^{-1}$) [a]	Fold increase over WT [a,b]	Fold increase over previous round [a,b]
$AtzA^{WT}$		*n.d.*	1.00	1.00
1	**S331C**	50	> 50	> 50
2	**F84L**	910	*n.d.*	18.2
3	**N328D**	6100	*n.d.*	6.7
4	**E125D**	10898	*n.d.*	1.8
5	**T219P**	18701	*n.d.*	1.7
6	**T217I**	18604	*n.d.*	1.0
7	**V92L**	19700	*n.d.*	1.1
8	**G255W**	20760	*n.d.*	1.1
$TriA^{WT}$	**I253L**	20810	*n.d.*	1.0
Additive model (null)			*> 50*	
Obs. fold change final mut (TriA)/WT (AtzA)		> 20810		

[a] The hydrolysis of melamine is not detectable in $AtzA^{WT}$ nor in any variants besides $AtzA^{S331C}$.

[b] The fitness is the Michaelis parameter k_{cat}/K_M ($M^{-1}.s^{-1}$).

***P. vivax* DHFR**

			WT	TRAJECTORY	Epistasis effect	
Round of appearance	Mutation	IC_{50} (µM) [a]	Fold increase over WT [a,b]	Fold increase over previous round [a]	Ratio Traj/WT	Distance between mut AA and key residue [b]
$DHFR^{WT}$		1.2	1.00	1.00	1.00	-
1	**S117N**	712	570.5	570.5	1.00	4.2
2	**S58R**	761	17.8	1.1	0.06	8.7
3	**N50I**	1899	87.4	2.5	0.03	7.8
4	**I173L**	2926	2.9	1.5	0.53	4.1
Additive model (null)			2572294			
Obs. fold change final mut/WT		2343				

[a] IC_{50} measured in *S. cerevisiae*.

[b] Distance between mutated residues and the pyrimethamine ligand on the WT structure (pdb: 2bl9).

***S. japonicum UT26* Lin_{BUT}**

			WT	TRAJECTORY	Epistasis effect	
Round of appearance	Mutation	k_{cat}/K_M ($M^{-1}.s^{-1}$) [a]	Fold increase over WT [a,b]	Fold increase over previous round [a]	Ratio Traj/WT	Distance between mut AA and key residue [b]
Lin_{BUT}^{WT}		2	1.00	1.00	1.00	-
1	**A247H**	4	2	2.00	1.00	3.6
2	**I134V**	7	2.5	1.75	0.70	6.2
3	**I138L**	31	2.5	4.43	1.77	7.9
4	**A112V**	85	*n.d.*	2.74	-	10.5
5	**M253I**	130	2	1.53	0.76	8.4
6	**A135T**	950	3	7.31	2.44	10
Lin_{BMI}^{WT}	**A81T**	1000	1	1.05	1.05	14.1
Additive model (null)			75			
Obs. fold change final mut/WT		500				

[a] The fitness is the Michaelis parameter k_{cat}/K_M ($M^{-1}.s^{-1}$) of the second step of the reaction PCHL->TDCL.

[b] Distance between mutated residues and the 1,2-dichloropropane ligand on the WT structure (pdb: 2bfn).

A

***B. diminuta* PTE S6 variant** *(forward trajectory-2NH)*

Round of appearance	Variant	Mutation	Lysate activity relative to PTE^{WT} [a]	WT: Fold increase over PTE^{WT} [b]	TRAJECTORY: Fold increase over parent enzyme [c]	Epistasis effect: Ratio Traj/WT	Distance between mut AA and key residue [e]	Calculation for Trajectory [d]
0	PTE^{WT}		1.00	1.00	1.00	1.00	-	-
1	R1	**H254R**	10.42	3.47	10.42	3.0	2.6	*R1 / PTE^{WT}*
2	R2a	**F306L**	70	-	-	-	-	
2	R2b	**D233E**	337	2.54	32.31	12.7	3.2	*R2b / R1*
3	R3	**I274S**	70	1.02	1.01	1.0	10.3	*R3 / R2a*
4	R4	**-**	564	-	-	-	-	-
5	R5a	**T172I**	2137	0.15	4.41	29.6	8.6	*R6 / R5b*
5	R5b	**S269T**	642	1.03	1.33	1.3	7.3	*R6 / R5a*
6	R6	**-**	2835	-	-	-	-	-
7	R7a	**M138I**	2752	1.00	0.31	0.3	17	*R8 / R8+I138m*
7	R7b	-	4065	-	-	-	-	-
7	R7c	**T199I**	1924	0.46	9.25	20.0	9.2	*R8 / R8+I199t*
8	R8	**-**	14462	-	-	-	-	-
9	R9	**L272M**	15707	0.76	1.09	1.4	5.8	*R9 / R8*
10	R10	**A80V**	26569	1.10	1.69	1.5	17	*R10 / R9*
11	R11a	**S111R**	24245	0.75	1.30	1.7	17.9	*R12 / R11b*
11	R11b	**A204G**	20516	0.74	1.10	1.5	6.5	*R12 / R11a*
12	R12	**-**	26594	-	-	-	-	-
13	R13a	**L271F**	44983	0.66	2.14	3.2	2.7	*R14 / R13b*
13	R13b	**L130V**	24812	0.75	1.18	1.6	9.6	*R14 / R13a*
14	R14	**-**	53031	-	-	-	-	-
18	R18	**A49V**	72912	0.88	1.28	1.5	21.1	*R18 / R18+V49a*
		K77E		0.94	1.40	1.5	19.8	*R18 / R18+E77k*
		L140M		1.73	1.67	1.0	20.4	*R18 / R18+M140l*
		I313F		1.19	1.52	1.3	7.8	*R18 / R18+F313i*
19	R19		35019	-	-	-	-	
		S137T		0.79	0.86	1.1	19.8	*R20 / R20+T137s*
		Q180H		0.58	0.63	1.1	11.4	*R20* / R20+H180q*
20	R20		50819	-	-	-	-	
		T45A		0.95	1.55	1.6	24.9	*R20 / R20+A45t*
		E144V		0.53	0.84	1.6	24.7	*R20 / R20+V144e*
		M314T		1.14	1.03	0.9	9.7	*R20 / R20+T314m*
		I341T		0.89	0.82	0.9	25.5	*R20 / R20+T341i*
21	R21	**S102T**	43382	0.88	0.85	1.0	7.4	*R21 / R20*
22	PTE^{AE}	**V176M**	40322	1.03	0.93	0.9	15.2	*PTE^{AE} / R21*
Additive model (null)				0.0521				
Obs. fold change final mut/WT			40322					

[a] The fitness is the catalytic activity in crude lysate relative to PTE^{WT}.

[b] Value obtained experimentally by contructing the single mutation on PTE^{WT} background.

[c] Value calculated from experimental data in Table A and B, with the equations provided in the last column, see note [d].

[d] Equations used to calculate the fold increase over the parent enzyme in the previous colum, see note [c].

[e] Distance between mutated residues on PTE^{WT} (PTE-S6, pdb: 4pcp) and the overlaid 2-naphtyl hexanoate (2NH) analogue ligand from (pdb: 4E3T).

B

Variant	2NH: Lysate activity relative to PTE^{WT} [a]
R18 +V49A	56822
R8 +I138M	47156
R8 +T199I	1563
R14 +I313F	25024
R18 +E77K	52038
R18 +M140L	43569
R18 +F313I	48012
R20 +A45T	32699
R20 +T137S	59273
R20 +V144E	60205
R20 +T314M	49256
R20 +T341I	62199
*R20**	29888
R20+H180Q	47132

R20 + H180Q was measured in a separate experiment and normalized by R20* re-assayed in the same conditions

A ***B. diminuta*** **PTE AE variant** *(Reverse trajectory-paraoxon)*

				WT	TRAJECTORY	Epistasis effect		
Round of appearance	Variant	Mutation	Lysate activity relative to PTE^{AE} [a]	Fold increase over PTE^{AE} [b]	Fold increase over parent enzyme [c]	Ratio Traj/WT	Distance between mut AA and key residue [e]	Calculation for Trajectory [d]
0	*PTE*^{AE}		1.00	1.00	1.00	1.00	-	-
1	revR1	**S308C**	6.62	6.74	6.60	0.98	4.3	PTE^{AE}+*s*308C / PTE^{AE}*
1	revR1	**-**	10	-	-	-	-	-
2	revR2	**I172T**	312	2.21	4.06	1.84	8.3	revR3 / revR3+*t*172I
2	revR2	**V130M**		6.92	5.42	0.78	9.7	*rev*R3 / revR3+M130V
3	revR3	**F271L**	905	7.23	2.90	0.40	1.7	revR3 / revR2
4	revR4	**T314M**	1357	1.52	3.40	2.24	9.6	revR4 / revR3+*p*135S
4	revR4	**P135S**		2.61	1.05	0.40	13.7	revR4 / revR3+*t314M*
5	revR5	**A203E**	4572	1.30	2.86	2.19	7	revR6+K293*m* / revR4
6	revR6	**M293K**	3660	1.34	0.94	0.70	24.2	revR6 / revR6+K293*m*
7	revR7a	**G174D**	9823	1.13	2.64	2.33	10.7	revR9 / revR9+D174*g*
7	revR7b	**H180Q**	7170	2.85	2.94	1.03	12	revR9 / revR9+*q*180H
8	revR8a	**-**	13780	-	-	-	-	-
8	revR8b	**S258N**	9361	0.93	2.00	2.15	9.7	revR9 / revR9+N258*s*
9	revR9	**Y156H**	18414	1.12	1.28	1.15	23.6	revR9 / revR9+H156y
12	*neo*PTE	**I306M**	19968	0.89	1.14	1.27	4.9	*neo*PTE / *neo*PTE+M306I
12	*neo*PTE	**V49A**		1.21	1.68	1.39	21.5	*neoPTE* / neo*PTE+*a*49V
Additive model (null)				18653				
Obs. fold change final mut/WT			19968					

[a] The fitness is the catalytic activity in crude lysate relative to PTE^{AE}.

[b] Value obtained experimentally by contructing the single mutation on PTE^{AE} background, extracted from reference [38].

[c] Value calculated from experimental data in Table A,B and C with the equations provided in the last column, see note [d].

[d] Equations used to calculate the fold increase over the parent enzyme in the previous colum, see note [c].

[e] Distance between mutated residues on PTE^{AE} (PTE-R22, pdb: 4pcn) and the overlaid paraoxon analogue ligand taken from pdb: 2r1n.

B

	Paraoxon
Variant	Lysate activity relative to PTE^{AE} [a]
PTE^{AE}	1
revR3+*t*172I	223
revR3+M130V	167
revR3+*p*135S	399
revR3+T314*m*	1291
revR6+K293*m*	3875
revR9+D174*g*	6984
revR9+*q*180H	6262
revR9+N258*s*	9218
*neo*PTE+M306I	17560
revR9+H156*y*	14393

C

	Paraoxon
Variant	Lysate activity relative to PTE^{AE}* or *neo*PTE* [a]
PTE^{AE}*	1
PTE^{AE}+*s*308C	6.6
*neo*PTE*	1
*neo*PTE+*a*49V	0.59

PTE^{AE}* and *neo*PTE* were remeasured a posteriori in the same conditions for a49V and s308C normalization only

T. fusca PAMO

Round of appearance	Mutation	E-Value ($E_{R/S}$)	WT	TRAJECTORY	Epistasis effect	
			Fold increase over Wt [a]	Fold increase over previous round [a]	Ratio Traj/WT [b]	Distance between mut AA and key residue [c]
PAMO^{WT}		0.05	1.00	1.00	1.00	-
1	**A442N**	0.18	3.51	3.51	1.00	5.3
2	**P440F**	6.98	0.29	38.24	134.13	9.6
3	**I67Q**	19.21	3.51	2.75	0.78	4.6
4 (*ZGZ-2*)	**L443I**	39.82	0.19	2.07	10.90	7.2
Additive model (null)			0.67			
Obs. fold change final mut/WT		765				

[a] The original fitness, $\Delta\Delta G_{R/S}$ (kcal/mol) was converted to the enantioselectivity factor $E_{R/S}$ with Equation (1) and (2), see Material and Methods.

[b] The ratio is the enantioselectivity factors provided by each mutation as it occurs in the trajectory (E_{TRAJ}) over the effect on the WT background (E_{WT}), E_{TRAJ}/E_{WT}, as described in reference [41].

[c] Distance between mutated residue and R337 in the WT structure (pdb: 2ylr).

A. niger ANEH

Round of appearance	Mutation	E-Value ($E_{S/R}$)	WT	TRAJECTORY	Epistasis effect	
			Fold increase over WT [a,b]	Fold increase over previous round [a]	Ratio Traj/WT [b]	Distance between mut AA and key residue [c]
ANEH^{WT}		4.00	1	1	1	-
B	**L215F**	16	4.00	4.00	1.00	3.7
	A217N					6
	R219S					12.2
C	**M329P**	21	1.00	1.31	1.31	15.8
	L330Y					11.9
D	**C350V**	24	1.00	1.14	1.14	4.5
F	**L249Y**	32	1.00	1.33	1.33	5.6
E	**T317W**	115	4.00	3.59	0.90	3.8
	T318V					4.7
Additive model (null)			16			
Obs. fold change final mut/WT		29				

[a] The original fitness, $\Delta\Delta G_{R/S}$ (kcal/mol) was converted to the enantioselectivity factor $E_{R/S}$ with Equation (1) and (2), see Material and Methods.

[b] The ratio is the enantioselectivity factors provided by each mutation as it occurs in the trajectory (E_{TRAJ}) over the effect on the WT background (E_{WT}), E_{TRAJ}/E_{WT}, as described in reference [39].

[c] Distance between mutated residues and the valpromide inhibitor in the WT structure (pdb: 3g0i).

B. cereus Bc-II

Round of appearance	Mutation	MIC (μg.ml^{-1})	WT	TRAJECTORY	Epistasis effect	
			Fold increase over WT [a,b]	Fold increase over previous round [a]	Ratio Traj/WT	Distance between mut AA and key residue [b]
Bc-II^{WT}		32	1	1	1	-
1	**G262S**	128	4	4.0	1.0	5.1
2	**L250S**	256	1	2.0	2.0	19.3
3	**N70S**	512	0.5	2.0	4.0	7.4
4 (*GLNV*)	**V112A**	1024	2	2.0	1.0	14.4
Additive model (null)			4			
Obs. fold change final mut/WT		32				

[a] The fitness is the MIC (μg.ml^{-1}).

[b] Distance between mutated residues and Zn^{2+} (II) on the WT structure (pdb: 1bc2) .

E. coli TEM-1

Round of appearance	Mutation	MIC (μg.ml^{-1})	WT	TRAJECTORY	Epistasis effect	
			Fold increase over WT [a,b]	Fold increase over previous round [a]	Ratio Traj/WT	Distance between mut AA and key residue [b]
TEM-1^{WT}		0.088	1	1	1	-
1	**G238S**	1.4	15.91	15.91	1.00	3.6
2	**E104K**	360	1.48	257.14	174.07	4.1
3	**A42G**	2100	1.00	5.83	5.83	15
4	**M182T**	2900	0.72	1.38	1.93	15.4
5	**g4205a**	4100	1.00	1.41	1.41	-
Additive model (null)			17			
Obs. fold change final mut/WT		46591				

[a] The fitness is the MIC (μg/mL).

[b] Distance between mutated residues on the WT structure (pdb: 1btl) and the overlaid N-(benzyloxycarbonyl)amino]methyl]phosphate, a transition state analogue from (pdb: 1axb).

Table S2. Summary of round-per-round improvements for nine evolutionary trajectories.

AtzA

		Fold increase normalized on n-1 mutations								
Round of evolution	**Mutation fixed**	***Round 1***	***Round 2***	***Round 3***	***Round 4***	***Round 5***	***Round 6***	***Round 7***	***Round 8***	***Round 9***
AtzA^{WT}		*n.d.*								
1	**S331C**	50.00	1.00							
2	**F84L**	*n.d.*	18.20	1.00						
3	**N328D**	*n.d.*	7.60	6.70	1.00					
4	**E125D**	*n.d.*	1.66	0.30	1.79	1.00				
5	**T219P**	*n.d.*	1.00	0.20	0.91	1.72	1.00			
6	**T217I**	*n.d.*	1.20	0.32	0.44	1.42	0.99	1.00		
7	**V92L**	*n.d.*	1.00	0.31	1.59	1.20	0.75	1.06	1.00	
8	**G255W**	*n.d.*	1.77	0.20	0.60	1.68	0.85	0.74	1.05	1.00
TriA^{WT}	**I253L**	*n.d.*	1.08	0.22	0.43	1.49	0.55	0.81	0.74	1.00

***P. vivax* DHFR**

		Fold increase normalized on n-1 mutations			
Round of evolution	**Mutation fixed**	***Round 1***	***Round 2***	***Round 3***	***Round 4***
DHFR^{WT}		1.00			
1	**S117N**	570.50	1.00		
2	**S58R**	17.79	1.07	1.00	
3	**N50I**	87.41	1.91	2.50	1.00
4	**I173L**	2.90	1.66	2.18	1.54

***T. fusca* PAMO**

		Fold increase normalized on n-1 mutations			
Round of evolution	**Mutation fixed**	***Round 1***	***Round 2***	***Round 3***	***Round 4***
PAMO^{WT}		1.00			
1	**A442N**	3.51	1.00		
2	**P440F**	0.29	38.24	1.00	
3	**I67Q**	3.51	3.37	2.75	1.00
4 (*ZGZ-2*)	**L443I**	0.19	1.63	0.85	2.07

***A. niger* ANEH**

		Fold increase normalized on n-1 mutations				
Round of evolution	**Mutation fixed**	***Round 1***	***Round 2***	***Round 3***	***Round 4***	***Round 5***
ANEH^{WT}		1.00				
1	**B**	4.00	1.00			
2	**C**	1.00	1.31	1.00		
3	**D**	1.00	1.00	1.14	1.00	
4	**F**	1.00	1.50	1.24	1.33	1.00
5 (*BCDFE*)	**E**	4.00	2.38	1.67	2.92	3.59

***B. cereus* Bc-II**

		Fold increase normalized on n-1 mutations			
Round of evolution	**Mutation fixed**	***Round 1***	***Round 2***	***Round 3***	***Round 4***
Bc-II^{WT}		1.00			
1	**G**	4.00	1.00		
2	**L**	1.00	64.00	1.00	
3	**N**	0.50	64.00	2.00	1.00
4 (*GLNV*)	**V**	2.00	64.00	2.00	2.00

***E. coli* TEM-1**

		Fold increase normalized on n-1 mutations				
Round of evolution	**Mutation fixed**	***Round 1***	***Round 2***	***Round 3***	***Round 4***	***Round 5***
TEM-1^{WT}		1.00				
1	**G238S**	15.91	1.00			
2	**E104K**	1.48	257.14	1.00		
3	**A42G**	1.00	16.43	5.83	1.00	
4	**M182T**	0.72	22.86	1.00	1.38	1.00
5	**g4205a**	1.00	1.00	1.00	0.71	1.41

REFERENCES

1. Zhang ZG, Lonsdale R, Sanchis J, Reetz MT. Extreme synergistic mutational effects in the directed evolution of a baeyer-villiger monooxygenase as catalyst for asymmetric sulfoxidation. J Am Chem Soc 2014;136(49):17262-72.