Supporting Information

Rational Design of a (S)-Selective-Transaminase for Asymmetric Synthesis of (1S)-1-(1,1'-biphenyl-2-yl)ethanamine

Daniel F.A.R. Dourado [1,2‡], Stefan Pohle [2‡], Alexandra T.P. Carvalho [1,2‡], Dharmendra S. Dheeman [2], Jill M. Caswell [2], Timofey Skvortsov[1,2], Iain Miskelly[2], Rodney T. Brown [2], Derek J. Quinn [2], Christopher C.R. Allen[3], Leonid Kulakov [3], Meilan Huang [1,], & Thomas S. Moody [2]*

[1] School of Chemistry and Chemical Engineering, Queen's University Belfast, David Keir Building, Stranmillis Road, Belfast BT9 5AG, Northern Ireland, UK

[2] Almac Sciences, Department of Biocatalysis and Isotope Chemistry, 20 Seagoe Industrial Estate, Craigavon BT63 5QD, Northern Ireland, UK

[3] School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland, UK

‡ These authors contributed equally to this work

* Correspondence and requests for materials should be addressed to M.H.

(email: m.huang@qub.ac.uk**)**

Table S1 Rationally designed *V. fluvialis* TAm mutants. In the first *in vitro* screening round 72 mutants were tested while 41 mutants constituted the second round. The study is focused on residues: G55, L56, W57, H83, B: R88, V153, K163, A228, E257, V258, I259, V422, R415. Additionally, we also mutated residues that interact with PMP: Y150, S118, S284, and L287 (interacts with S284), and residues that interact with L56 and with B: R88: N58, F19 and F85.

First screening round	Second screening round
F19K/W57G	W57F/I259A/V422A
F19R/W57G	W57F/I259M/V422A
G55A/W57G	W57F/R415A
G55S/W57G	W57F/R88H
L56A/W57G	W57F/R88H/V153S/K163F/I259A/R415A/V422A
L56F/W57G	W57F/R88H/V153S/K163F/I259M/R415A/V422A
L56I/W57G	W57G/ K163F/I259M/V422A
L56V/W57G	W57G/I259A/R415A
W57A	W57G/I259A/R415A/V422A
W57A/N58V	W57G/I259A/R415G/V422A
W57A/R415L	W57G/I259A/R415P/V422A
W57F	W57G/I259M/ R415G/V422A
W57G/A228S	W57G/I259M/ R415P/V422A
W57G/A228S/V422A	W57G/I259M/R415A
W57G/A228T	W57G/I259M/R415A/V422A
W57G/A228V	W57G/K163F/I259A/V422A
W57G/E257A	W57G/K163F/R415A
W57G/E257A/V258A	W57G/R415A/V422A
W57G/E257G	W57G/R88H /I259A/R415A/V422A
W57G/E257G/V258A	W57G/R88H/I259A
W57G/E257N	W57G/R88H/I259A/V422A
W57G/F85A	W57G/R88H/I259M
W57G/F85G	W57G/R88H/I259M/R415A/V422A
W57G/F85V	W57G/R88H/I259M/V422A
W57G/G55A/S284G	W57G/R88H/K163F
W57G/H83I	W57G/R88H/R415A
W57G/H83L	W57G/R88H/R415G
W57G/H83M	W57G/R88H/R415P

W57G/H83N	W57G/R88H/V153S
W57G/H83P	W57G/R88H/V153S/I259A/R415A/V422A
W57G/I259A	W57G/R88H/V153S/I259M/R415A/V422A
W57G/I259A/V422A	W57G/R88H/V153S/K163F/I259A/R415A/V422A
W57G/I259A/V422Y	W57G/R88H/V153S/K163F/I259A/R415G/V422A
W57G/I259M	W57G/R88H/V153S/K163F/I259A/R415P/V422A
W57G/I259M/V422A	W57G/R88H/V153S/K163F/I259M/R415A/V422A
W57G/I259Q	W57G/R88H/V153S/K163F/I259M/R415G/V422A
W57G/I259V	W57G/R88H/V153S/K163F/I259M/R415P/V422A
W57G/K163F	W57G/R88H/V422A
W57G/K163L	W57G/V153S/I259A/V422A
W57G/L287I	W57G/V153S/I259M/V422A
W57G/L287V	W57G/V153S/R415A
W57G/L56A/I259M	
W57G/L56A/I259V	
W57G/N58A	
W57G/N58I	
W57G/N58P	
W57G/N58T	
W57G/N58V	
W57G/R415A	
W57G/R415F	
W57G/R415L	
W57G/R415P	
W57G/R415V	
W57G/R415W	
W57G/R88H	
W57G/R88K	
W57G/R88Q	
W57G/S118A	
W57G/S284A	
W57G/S284G	
W57G/V153A	
W57G/V153K	
W57G/V153R	

W57G/V153S	
W57G/V258A	
W57G/V258I	
W57G/V258T	
W57G/V422A	
W57G/V422I	
W57G/V422L	
W57G/V422Y	
W57G/Y150F	

Table S2 *V. fluvialis* TAm WT and engineered variants reaction rates and kinetic parameters towards acetophenone.

Enzyme	**Wild type**	**W57G/ R415A**	**W57G/ K163F/ R415A**	**W57G/ I259M/ R415A**	**W57F/R88 H/V153S/ K163F/I25 9M/R415A /V422A**
K_M **[mM]**	14.7	4.53	12.7	5.80	6.70
K_{cat} **[s^{-1}]**	$1.83x10^{-2}$	$2.28x10^{-2}$	$5.80x10^{-2}$	$1.73x10^{-2}$	$1.16\ x10^{-1}$
K_{cat}/K_M **[$M^{-1}\cdot s^{-1}$]**	1.24	5.03	4.55	2.98	16.6
Fold increase in K_{cat}/K_M [a]	1	4	4	2	13
Reaction rate [mM/h][b]	0.442	0.706	1.24	0.469	2.93
Reaction rate fold increase	1	2	3	1	7

[a] Fold change relative to WT values.

[b] Reaction rates are initial reaction rates. Reaction conditions: 20 mM substrate, 1M IPA, 0.01 mM enzyme, 0.1 M K·phosphate buffer (pH 8.0), 10% DMSO (v/v), 40 °C. Reactions were performed in duplicates.

Table S3 Residual activities of the WT enzyme and best mutants. Residual activity is defined as the [amount of product formed in a reaction with preincubated enzyme]/[amount of product formed in control reaction]*100%.

Variant	**40°C - 18 h**	**50°C - 18 h**
W57G/R415A	91.1%	28.2%
W57G/K163F/R415A	109%	34.1%
W57G/I259M/R415A	38.8%	16.4%
W57F/R88H/V153S/K163F/ I259M/R415A/V422A	97.9%	89.6%

Table S4 Amount of amine product formed (mM) in the reactions with the preincubated enzymes.

Enzyme	40°C - 0 h	40°C - 18 h	50°C - 0 h	50°C - 18 h
Wild type	0.000	0.000	0.000	0.000
	0.000	0.000	0.000	0.000
	0.000	0.000	0.000	0.000
W57G/R415A	0.0425	0.0306	0.0364	0.0093
	0.0299	0.0320	0.0322	0.0096
	0.0315	0.0321	0.0334	0.0099
W57G/K163F/R415A	0.108	0.114	0.149	0.0506
	0.110	0.116	0.150	0.0488
	0.102	0.120	0.149	0.0535
W57G/I259M/R415A	0.0220	0.0074	0.0880	0.0150
	0.0182	0.0082	0.0838	0.0132
	0.0197	0.0076	0.0840	0.0137
W57F/R88H/V153S/K163F/ I259M/R415A/V422A	0.744	0.705	1.14	0.990
	0.730	0.717	1.12	1.01
	0.736	0.742	1.12	1.03

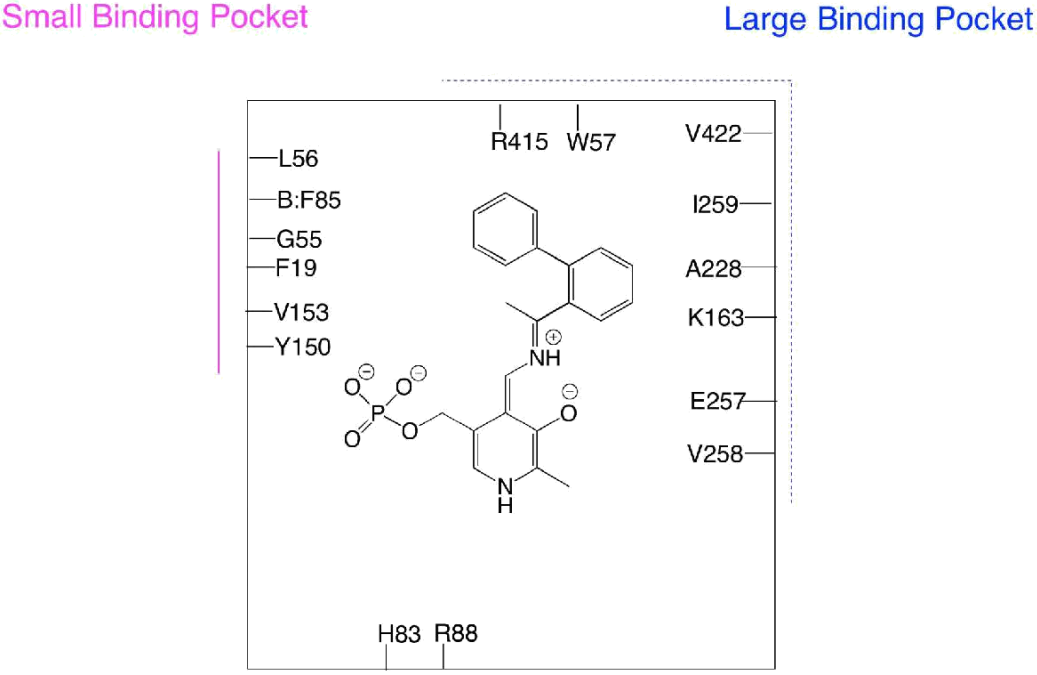
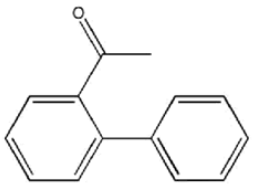


Figure S1. The large and small pockets of *V. fluvialis* amine transaminase.

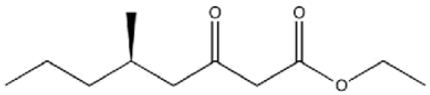
a



2-acetylbiphenyl

Chemical Formula: $C_{14}H_{12}O$
Exact Mass: 196.09
Molecular Weight: 196.25

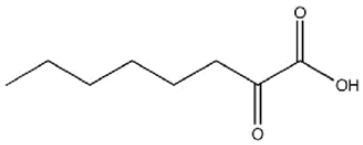
b



(R)-ethyl 5-methyl 3-oxooctanoate

Chemical Formula: $C_{11}H_{20}O_3$
Exact Mass: 200.14
Molecular Weight: 200.28

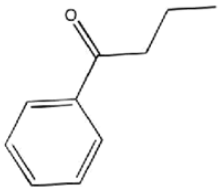
c



2-oxo-octanoate

Chemical Formula: $C_8H_{14}O_3$
Exact Mass: 158.09
Molecular Weight: 158.20

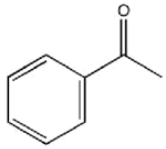
d



butyrophenone

Chemical Formula: $C_{10}H_{12}O$
Exact Mass: 148.09
Molecular Weight: 148.21

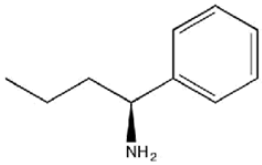
e



acetophenone

Chemical Formula: C_8H_8O
Exact Mass: 120.06
Molecular Weight: 120.15

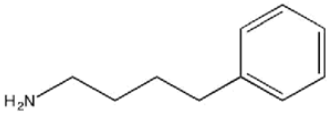
f



(S)-1-phenylbutylamine

Chemical Formula: $C_{10}H_{15}N$
Exact Mass: 149.12
Molecular Weight: 149.24

g



4-phenylbutylamine

Chemical Formula: $C_{10}H_{15}N$
Exact Mass: 149.12
Molecular Weight: 149.24

Figure S2. Substrates studied in the engineering of (S)-selective-TAms. **a)** Substrate studied in the present research (ketone **1**); **b**)-**f**) Substrates studied in previous literature. **b**) Substrate studied in Ref. 6e; **c**) Substrate studied in Ref. 6f; **d-e**) Substrates studied in Ref. 6g; **f**) Substrate studied in Refs. 10&11; **g)** Substrate studied in Ref. 8.

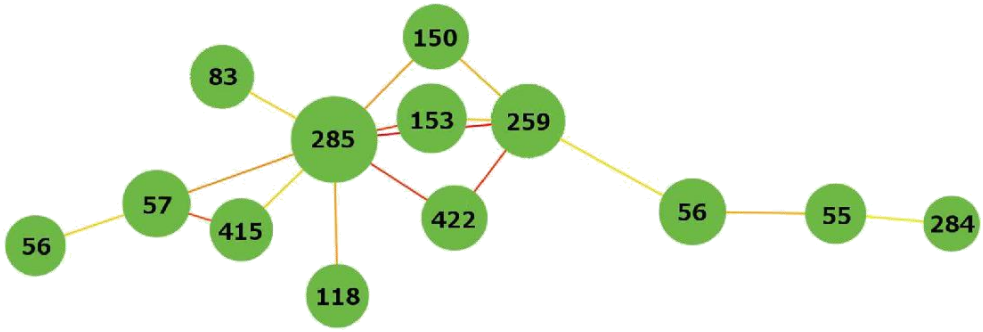
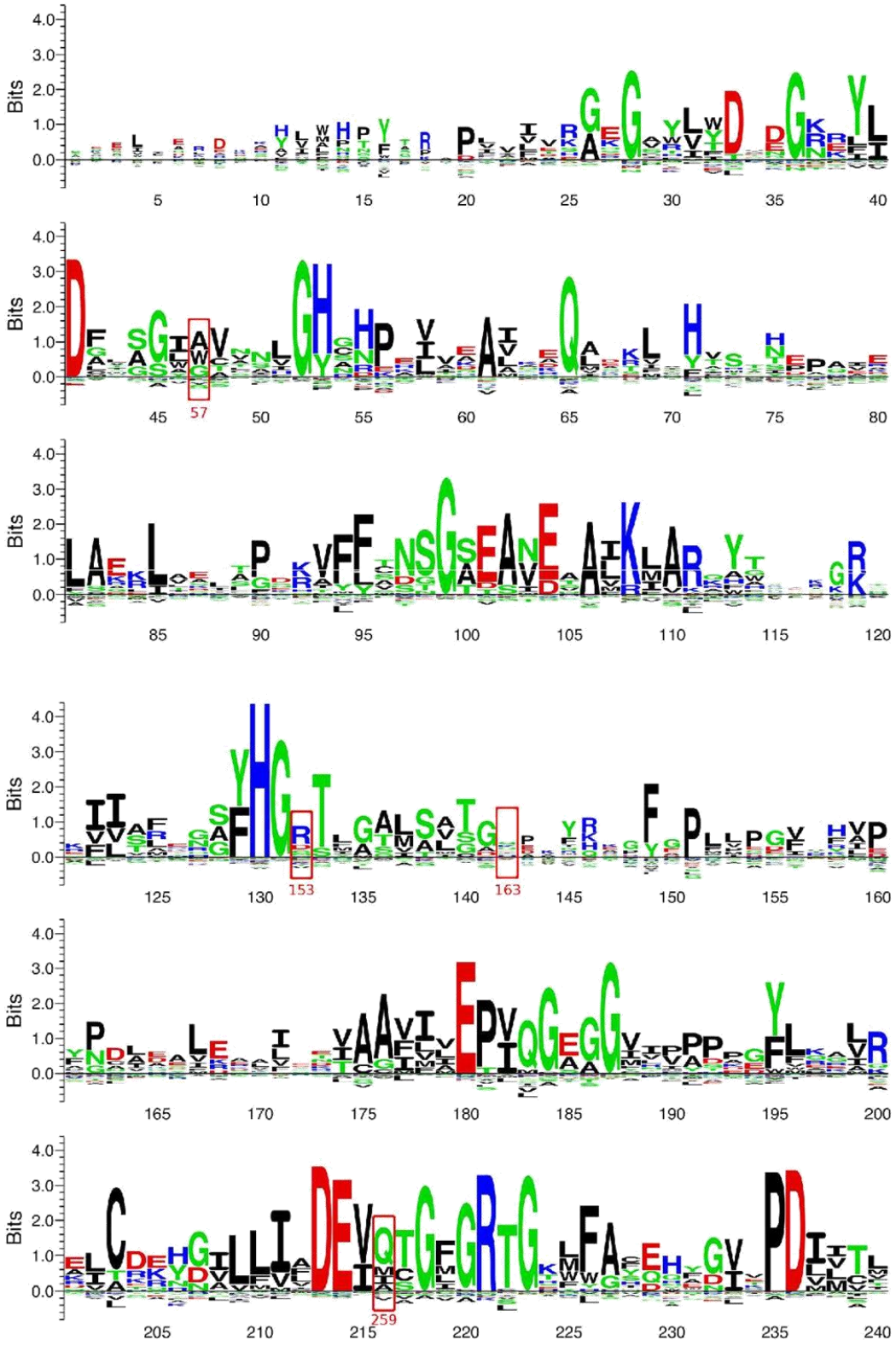


Figure S3. Co-evolution network of pyridoxal-phosphate-dependent aminotransferase super family (30,000 sequences) with a coefficient higher than 0.8.


Bits
4.0
3.0
2.0
1.0
0.0
5 10 15 20 25 30 35 40
45 57 50 55 60 65 70 75 80
85 90 95 100 105 110 115 120
125 130 153 135 140 163 145 150 155 160
165 170 175 180 185 190 195 200
205 210 215 259 220 225 230 235 240

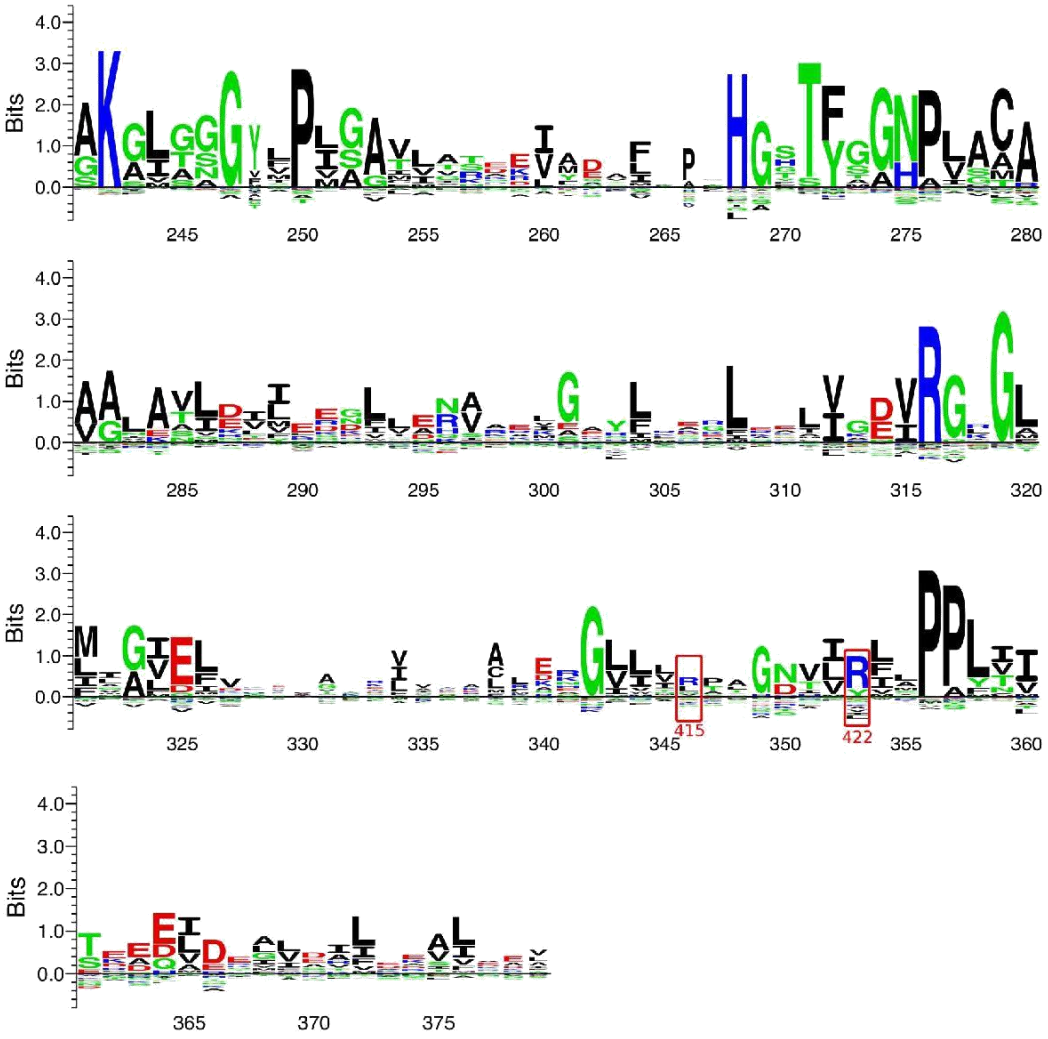


Figure S4. Amino acid distribution of the Pyridoxal-phosphate-dependent aminotransferase super family. Total number of sequences: 10623; the sequences coverage is at least 60 % of the *V. fluvialis* Tam sequence; sequence identity >25% and < 95%; Residue positions mutated in the best variants are highlighted with redboxes.

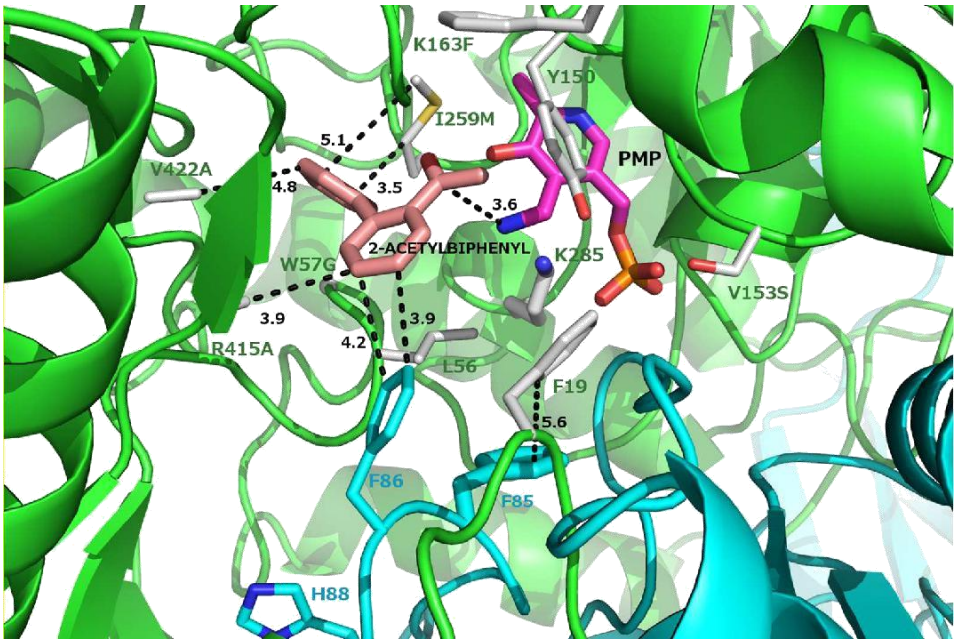


Figure S5. MD reference structure of ketone **1** docked into *V. fluvialis* TAm W57G/R88H/V153S/K163F/I259M/V422A/R415A in presence of PMP. The active centre residues are represented by sticks with the carbons of Chain A (green ribbon) coloured in grey and the carbons of Chain B (cyan ribbon) coloured in cyan. Relevant distances are shown (Å).

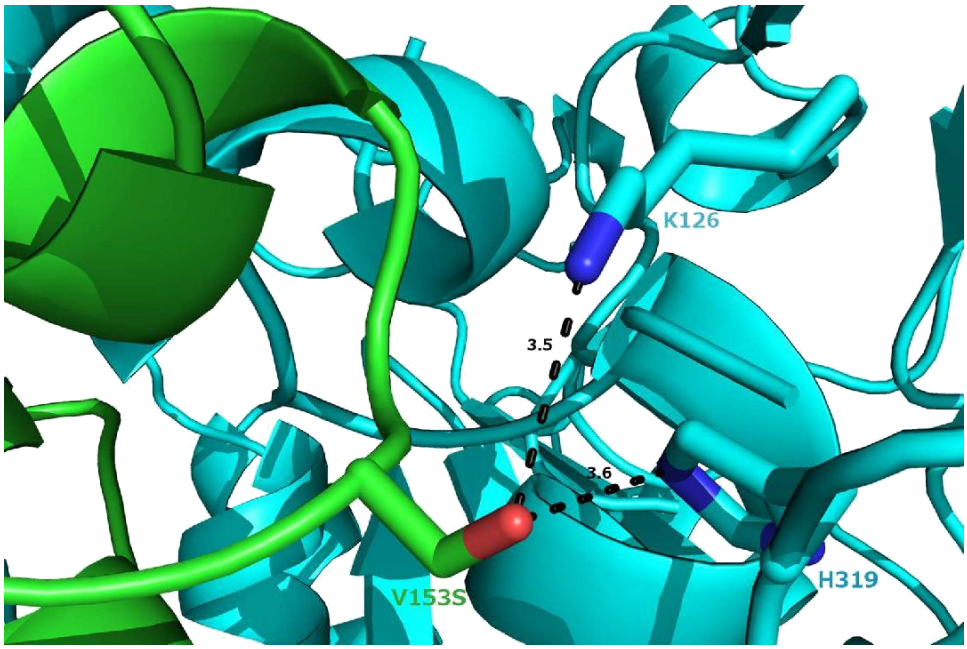


Figure S6. The predicted interactions along the dimeric interface of the *V. fluvialis* TAm V153S mutant from ZEMu calculation. The mutant structure was built based on the crystal structure of the WT *V. fluvialis* TAm (pdb: 4E3Q). Chain A is shown by green ribbon and Chain B is shown by cyan ribbon. Relevant distances are shown (Å).

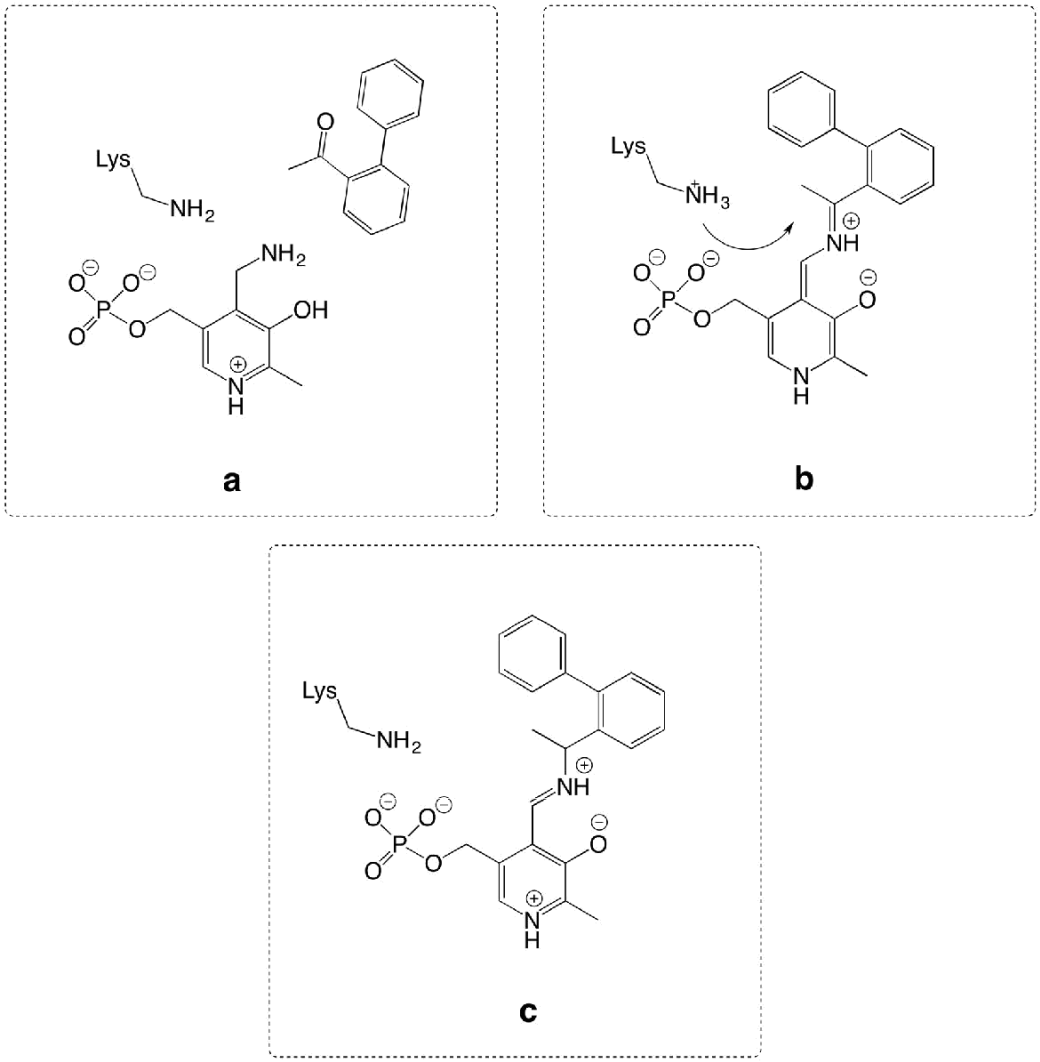


Figure S7. Intermediates in the transamination reaction pathway modelled in this study: (**a**) E: PMP; (**b**) Planar quinonoid; (**c**) External aldimine.

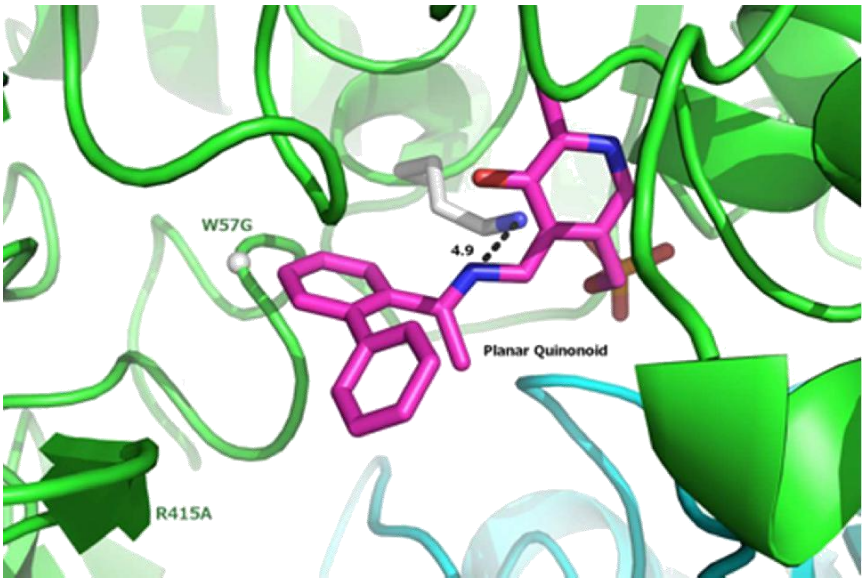


Figure S8. *V. fluvialis* TAm W57G/R415A docked into the planar quinonoid. The active centre residues are represented by sticks with the carbons of Chain A (green ribbon) coloured in grey and the carbons of Chain B (cyan ribbon) coloured in cyan. Relevant distances are shown (Å).

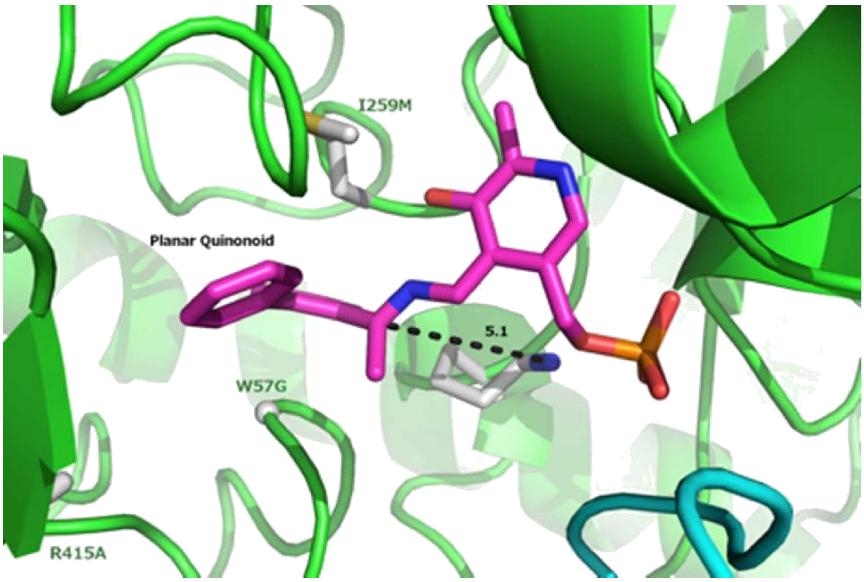


Figure S9. *V. fluvialis* TAm W57G/R415A/I259M docked into the planar quinonoid. The active centre residues are represented by sticks with the carbons Chain A (green ribbon) coloured in grey and the carbons of Chain B (cyan ribbon) coloured in cyan. Relevant distances are shown (Å).

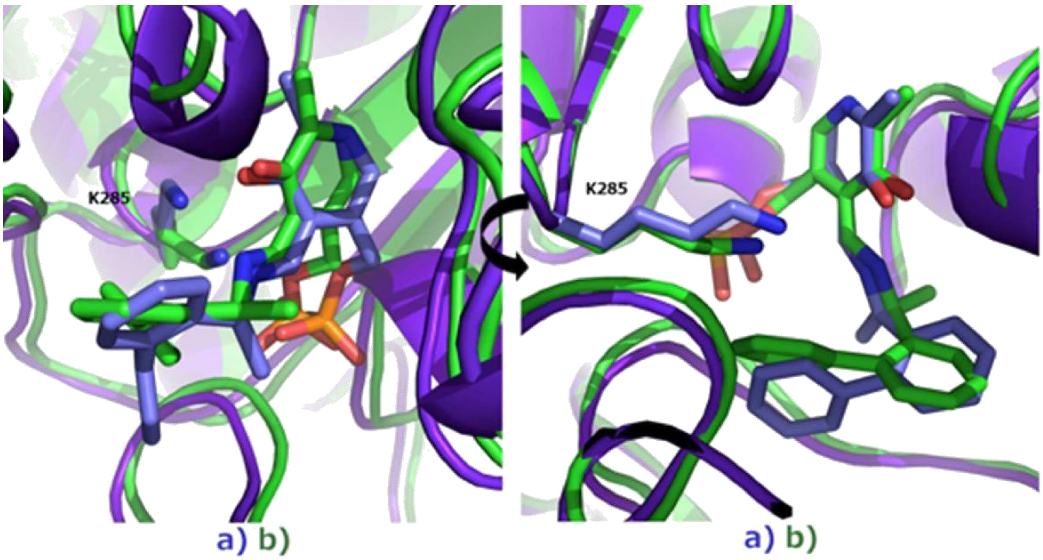


Figure S10. Superimposition of the MD reference structures of *V. fluvialis* TAm W57F/R88H/V153S/K163F/I259M/R415A/V422A with **(a)** planar quinonoid and **(b)** pro – (S) external aldimine (main chain rms = 1.2 Å). Two different orientations are shown.

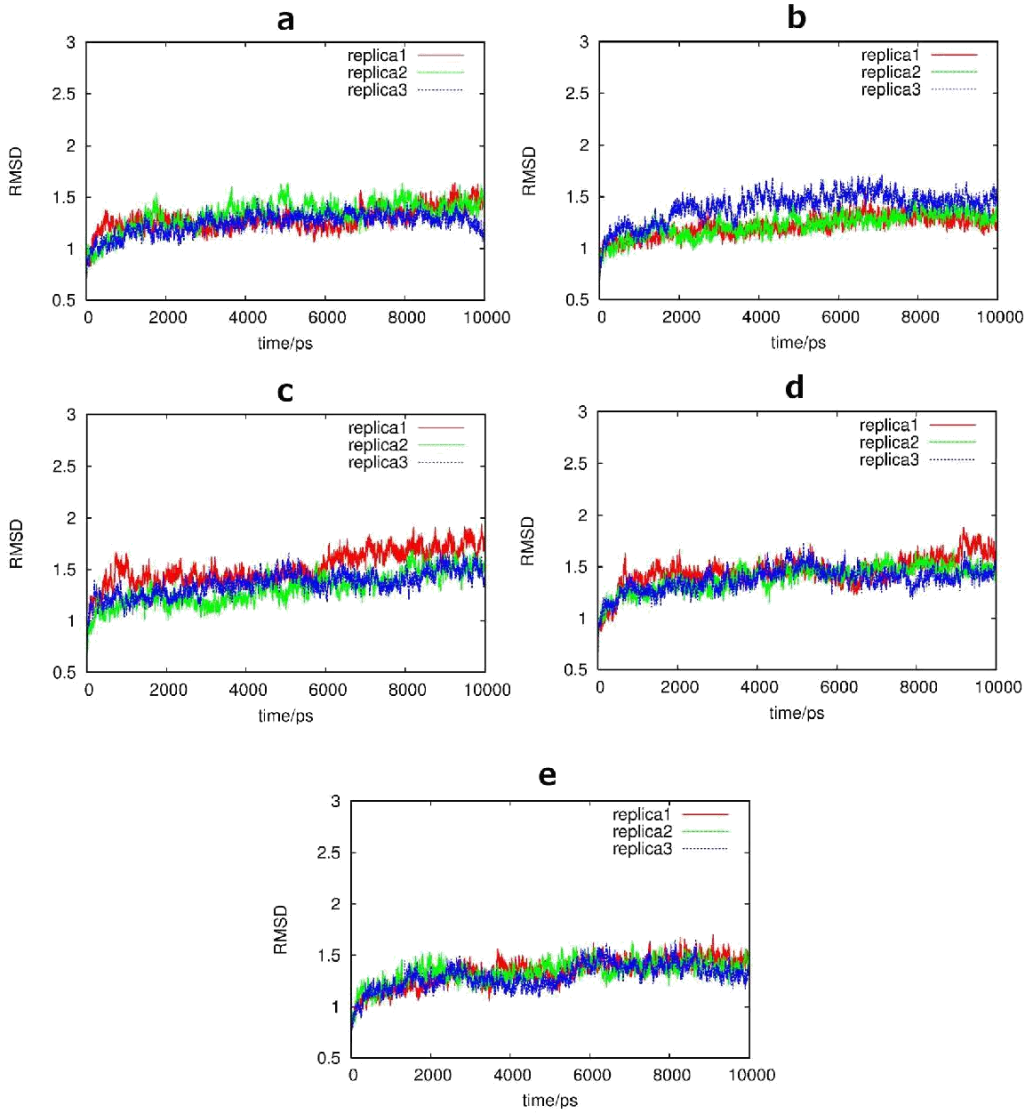


Figure S11. RMSD of the alpha-carbon atoms for (**a**) *V. fluvialis* TAm WT; (**b**) *V. fluvialis* TAm W57G/R415A; (**c**) *V. fluvialis* TAm W57G/ K163F/R415A; (**d**) *V. fluvialis* TAm W57G/ I269M/R415A; (**e**) *V. fluvialis* TAm W57F/ R88H/V153S/K163F/I259M/R415A/V422A. All the complexes contain ketone **1** and the PMP cofactor. No significant difference was found in the respective MD replicas.

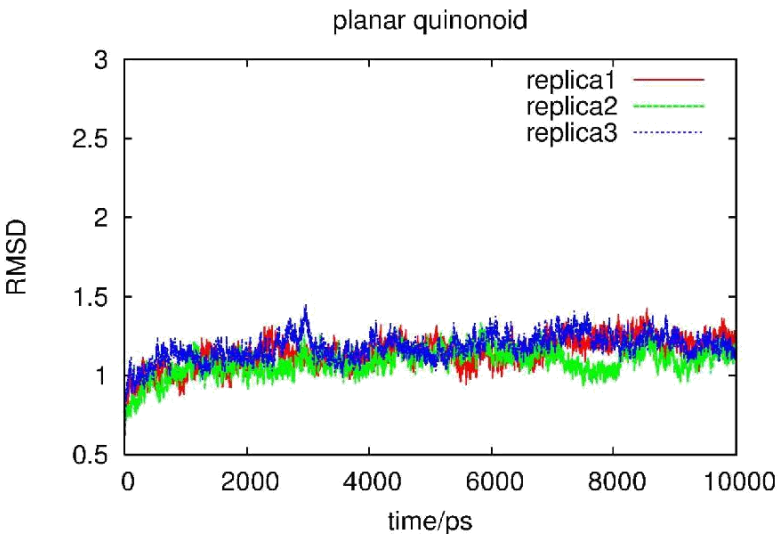


Figure S12. RMSD of the alpha-carbon atoms for the *V. fluvialis* TAm W57F/ R88H/V153S/K163F/I259M/R415A/V422A bound with the planar quinonoid. No significant difference was found in the respective MD replicas.

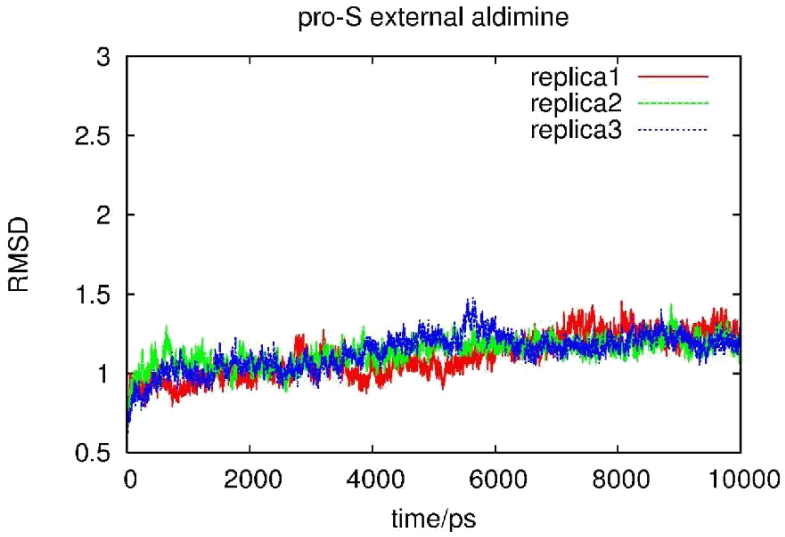


Figure S13. RMSD of the alpha-carbon atoms for the *V. fluvialis* TAm W57F/ R88H/V153S/K163F/I259M/R415A/V422A bound with the pro (S) external aldimine. No significant difference was found in the respective MD replicas.

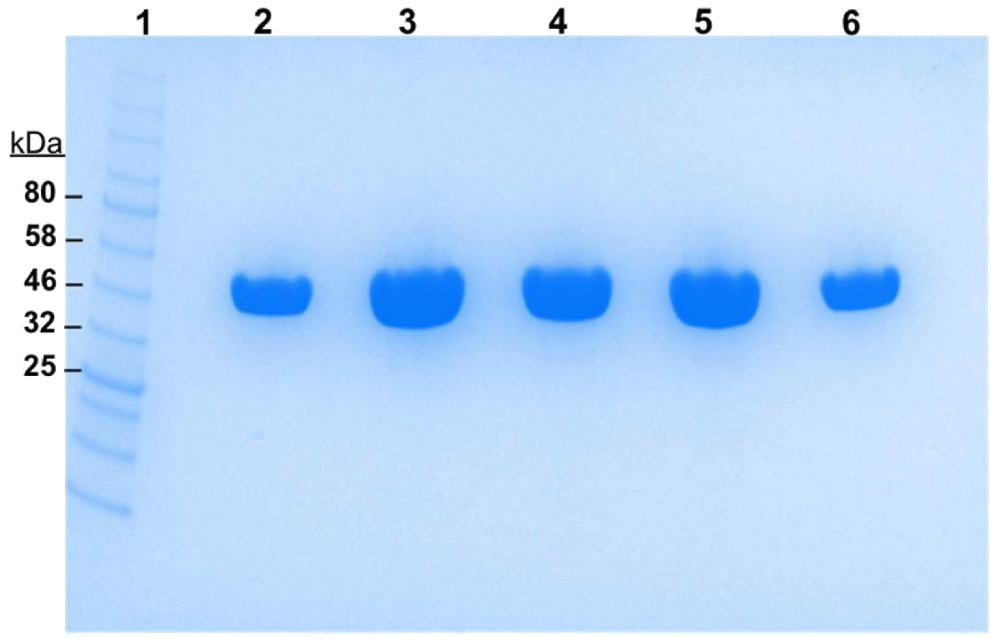


Fig S14. SDS-PAGE analysis of purified WT *V. fluvialis* TAm and its four variants. The enzymes were purified using immobilised metal affinity chromatography (IMAC) as described in material and methods. (1) Molecular weight markers; (2) WT; (3) W57G/R415A; (4) W57G/K163F/R415A; (5) W57G/I259M/R415A; (6) W57F/R88H/V153S/K163F/I259M/R415A/V422A.

$$\Delta G = (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) + \left(V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta S_{conf}\right)$$

Equation S1. Autodock 4.2 free energy equation. ΔS_{conf} is the entropy lost upon binding; V corresponds to the pairwise energy terms; L refers to the ligand and P the protein Ref. 24.

$$V = W_{vdw} \sum_{i.j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond} \sum_{i.j} E_{(t)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec} \sum_{i.j} \frac{q_i q_j}{e(r_{ij}) r_{ij}} + W_{sol} \sum_{i.j} \left(S_i V_j + S_j V_i\right) e^{\left(-r_{ij}^2/2\sigma^2\right)}$$

Equation S2. Autodock 4.2 pairwise energy terms. *Ws* are weighting constants optimized to calibrate a data set of 188 diverse protein–ligand complexes of known structure and binding energy. The first term is the Lennard-Jones potential with A and B based on the amber force field; the second term is a directional H-bond term; the third term describes electrostatics with the Coulomb potential; the fourth corresponds to the desolvation potential (Ref. 24).