Supporting Information

Engineering of a Baeyer-Villiger Monooxygenase reveals key residues for asymmetric oxidation of omeprazole sulfide

Shiyu Wei, Guochao Xu*, Lu zhang, Jieyu Zhou, Ye Ni *

Key laboratory of industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, Jiangsu, China

*Corresponding author

Prof. Ye Ni, E-mail: yni@jiangnan.edu.cn; Phone: (+86) 0510-85918206; ORCID: 0000-0003-4887-7517

Dr. Guochao Xu, E-mail: guochaoxu@jiangnan.edu.cn; ORCID: 0000-0001-9784-5648

Table of Contents

Materials and methods

Chemicals and materials. All other commercial chemicals were purchased from Sigma-Aldrich (Stockholm, Sweden), Aladdin (Beijing, China) or Macklin (Shanghai, China). All PCR reagents and enzymes were purchased from TaKaRa Biotechnology Co. (Dalian, China). Primers were synthesized by Talen-bio Biological Technology Co. (Shanghai, China).

Homology modeling and structure analysis. Homology modeling of *Ra*BVMO was performed using AlphaFold2 software [1] (https://alphafold.ebi.ac.uk/). The structure of *Ra*BVMO was evaluated by SAVES v6.0 [2] (https://saves.mbi.ucla.edu). CAVER 3.0 [3] (http://www.caver.cz) was used to identify the tunnels existing in *Ra*BVMO. The *Ra*BVMO mutants were generated using Chimera (version 1.16). Molecular docking analysis was performed by Discovery studio 4.5 and the docking results were selected according to their binding affinities and molecule conformations.

Library construction. Variants were generated using whole-plasmid mutagenesis protocol with the plasmid of the WT and corresponding mutants as a template. Saturation mutagenesis was then performed on selected sites using the NNK codon. Sites exhibiting over 50% activity or 25% *ee* were subjected to further iterative combinatorial mutagenesis (ISM). The primers used in the library creation process are listed in the Supporting Information (**Table S1**). The PCR procedures were as follows: 95 °C for 2 min, followed by 25 cycles of 94 °C for 30 s, 50 °C for 30 s, and 68 °C for 3 min and 40 s. After Dpn I digestion at 37 °C for 1h, the PCR products were transformed into competent *E. coli* BL21 (DE3) cells and plated onto LB agar plates supplemented with 50 μg/mL kanamycin. The cells were spread on the LB-agar plates supplemented with 50 $\mu g \cdot mL^{-1}$ kanamycin for overnight incubation at 37 °C.

Library screening. Colonies were randomly picked and introduced into 96-well plates that containing 300 μL of LB medium (10 $g \cdot L^{-1}$ NaCl, 10 $g \cdot L^{-1}$ tryptone, 5 $g \cdot L^{-1}$ yeast extract) with 50 $\mu g \cdot mL^{-1}$ kanamycin. After overnight growth at 37 °C, with shaking at 180 rpm, 100 μL overnight cultures were inoculated into 4 mL of fresh LB medium containing 50 $\mu g \cdot mL^{-1}$ kanamycin in a new 24 deep-well plate. For protein expression, IPTG was added to a final concentration of 0.2 mM after cultivation at 25°C for 16 h. Then the cultures were centrifuged at 4000 rpm and at 4 °C for 15 min to collect cells. Subsequently discarding the supernatants, the cells were lysed with 500 μL lysis buffer (pH 9.0 Tris-HCl buffer, 750 $mg \cdot mL^{-1}$ lysozyme) at 4 °C for 4 h and then centrifuged at 4 °C and 4000 rpm for 15 min. 400 μL of each supernatant were transferred into a new 24 deep well plate to prepare the reaction mixture. In each well of the plate, 500 μL of reaction mixture was added, containing 1 mM NADPH, and 1 mM substrate with 5% (v/v) methyl alcohol. The reaction plates were incubated at 30 °C with shaking at 180 rpm for 4 h. After 500 μL of acetonitrile was added into each well, the reaction plates were centrifuged at 4 °C and 4000 rpm for 30 min. Then 200 μL of the supernatants in each well were subjected to HPLC analysis. The mutants with improved selectivity or activity were picked out, and the results were confirmed by repeating the reaction in Eppendorf tubes. The mutations in the gene of the mutants were confirmed by sequencing.

Enzyme expression and purification. The single colonies of *Ra*BVMO WT and mutants were inoculated into 4 mL of LB media supplemented with 50 $\mu g \cdot mL^{-1}$ kanamycin. After overnight growth at 37 °C, 1 mL of culture were transferred into 100 mL of LB media containing 50 $\mu g \cdot mL^{-1}$ kanamycin as an inducer at 25 °C with shaking at 180 rpm for 16 h for enzyme expression. Cell pellets were harvested by centrifugation and washed once with 100 mM, pH 9.0 Tris-HCl buffer. The pellets were suspended in 10 mL of 100 mM Tris-HCl buffer (pH 9.0) and were disrupted by sonication. The supernatant containing the crude enzyme was obtained by centrifugation at 4 °C and 8000 rpm for 30 min. After going through the filter, the supernatant was subjected to affinity chromatography (HisTrap™ FF column, 5 mL) for purification. The impurity was removed by flushing the column with 50 mM imizadol, and the enzyme was eluted by 25 mM Tris-HCl buffer containing 250 mM imizadol and 0.5 M NaCl. The elution fractions were subjected to a desalination column for removal of salt and imizadol. The purified enzymes were dissolved in 100 mM, pH 9.0 Tris-HCl buffer and stored at -80 °C for use. The purity of the enzyme was confirmed by SDS-PAGE (**Fig S6** in Supporting Information). The concentration of the purified enzyme was determined by a Bradford protein assay method, following the provided protocol.

Specific activity and enantioselectivity assays. The specific activity and stereoselectivity of *Ra*BVMO variants toward cyclohexanone (**1a**), thioanisole (**1b**), methyl(naphthalen-2-yl)sulfane (**1c**), benzyl(phenyl)sulfane (**1d**) and omeprazole sulfide (**1e**) were measured using purified enzymes. In a 500 μL reaction system, 2 mM substrate **1a** or **1b** (1 mM for **1c** or **1e**), 2 mM NADPH for **1a** or **1b** (1 mM NADPH for **1c** or **1e**), and diluted enzyme were mixed in Tris-HCl (100 mM, pH 9.0). After the reaction was performed at 30 °C for 30 min, the reaction mixture was extracted with an equal volume of ethyl acetate. The

specific activity and enantioselectivity were determined by GC or HPLC, as described elsewhere [4, 5]. All the measurements were carried out in triplicate.

Determination of kinetic parameters. The kinetic constants (*K*m and *k*cat) of MT2 were assayed by varying the concentration of omeprazole sulfide (**1e**) from 0.01 to 1 mM with 1 mM NADPH present in each reaction. The reaction time was 30 min at 30 ℃ and the conversion rate relative to substrate was maintained within 20%. The kinetic parameters were calculated according to non-linear curve fitting with Michaelis–Menten equation. All samples were determined by HPLC and all experiments were carried out in triplicate.

Asymmetric oxidation of omeprazole sulfide for synthesis of *S*-omeprazole. A 10 mL of reaction mixture was composed of 0.35-, 0.6- or 1-g/L omeprazole sulfide (**1e**) as substrate (with 5% (v/v) MeOH), glucose (1.5 equiv), purified enzyme M2T (4 g/L), glucose dehydrogenase (GDH, 10 g/L), $NADP^+$ (0.2 mM), and Tris-HCl buffer (100 mM, pH 9.0) and was shaken at 180 rpm and 30 °C. Samples were intermittently removed and extracted to analyze the conversion rate by HPLC.

Molecular simulations. Molecular docking analysis was performed by Discovery studio 4.5. The model structure of *Ra*BVMO containing FAD and NADP was used. The receptor model was built using the Protein Preparation Wizard [6]. Protonation states were determined using PROPKA [7, 8]. Prior to docking, each protein structure was energy minimized using the OPLS2005 force field [9], with all heavy atoms restrained to an RMSD of 0.3 Å.

Table S1 PCR primers used for mutants construction.

Mutants	Primers[a]
C67NNK	GCGCGTACCGAT**NNK**GAGGGC
	GTAGTAGCCCTC**MNN**ATCGGT
E68NNK	CGTACCGATTGC**NNK**GGCTAC
	GTAGTAGTAGCC**MNN**GCAATC
F442NNK	CCACAGAGCCCA**NNK**ACCAAT
	TGGGATATTGGT**MNN**TGGGCT
T443NNK	CAGAGCCCATTC**NNK**AATATC
	CGGTGGGATATT**MNN**GAATGG
N444NNK	AGCCCATTCACC**NNK**ATCCCA
	GCACGGTGGGAT**MNN**GGTGAA
R337NNK	TTCGGTGCGAAA**NNK**CCACCA
	ACCGCTTGGTGG**MNN**TTTCGC
P338NNK	GGTGCGAAACGC**NNK**CCAAGC
	GGTACCGCTTCC**MNN**CGCTTT

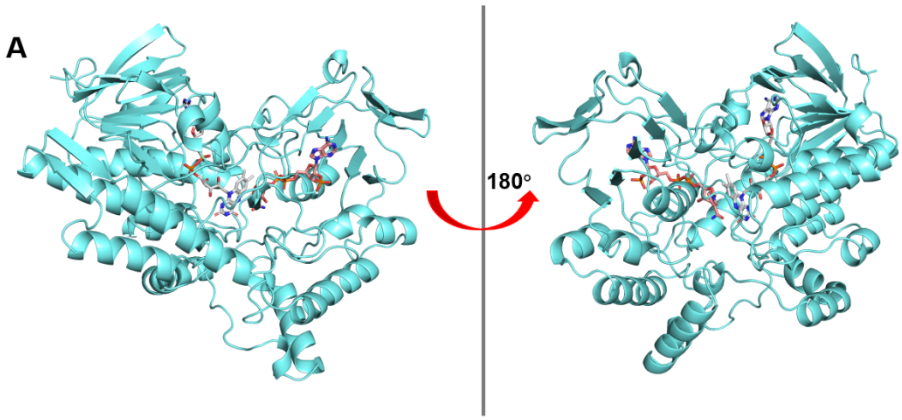
[a]The mutated sites are underlined.

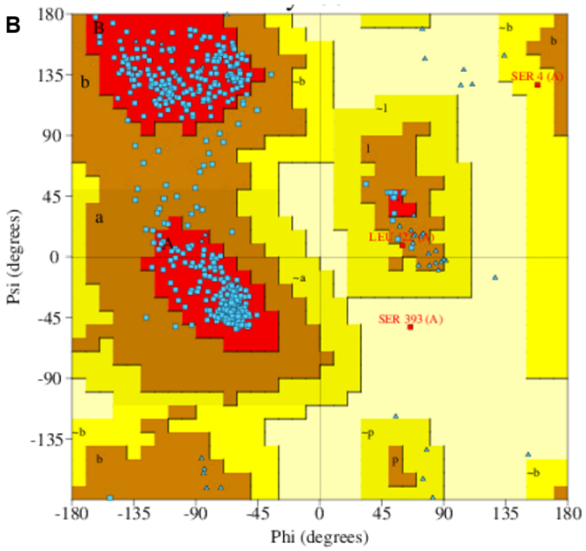
Table S2 HPLC conditions and retention times for substrates 1b-1e used in the study.

Substrate	Flow (mL min^{-1})	T (ºC)	Column	Eluent[a]	Product(s) t_R (min)
1b	0.8	30	Chiralcel OD-H	*n*-hexane-IPA 90:10	14.6 (*R*); 18.8 (*S*)
1c	0.8	30	Chiralcel OD-H	*n*-hexane-IPA 95:5	39.9 (*R*), 43.2 (*S*)
1d	0.8	30	Chiralcel OD-H	*n*-hexane-IPA 95:5	26.7 (*R*); 32.1 (*S*)
1e	0.5	30	Chiralcel AD-3	*n*-hexane-IPA 50:50	11.9 (*R*); 14.3 (*S*)

[a] Experiments were performed with isocratic eluent.

Supporting Figures





C **Ramachandran Plot statistics**

	No. of residues	%-tage
Most favoured regions [A,B,L]	421	90.0%
Additional allowed regions [a,b,l,p]	44	9.4%
Generously allowed regions [~a,~b,~l,~p]	2	0.4%
Disallowed regions [XX]	1	0.2%
Non-glycine and non-proline residues	468	100.0%
End-residues (excl. Gly and Pro)	3	
Glycine residues	51	
Proline residues	31	
Total number of residues	553	

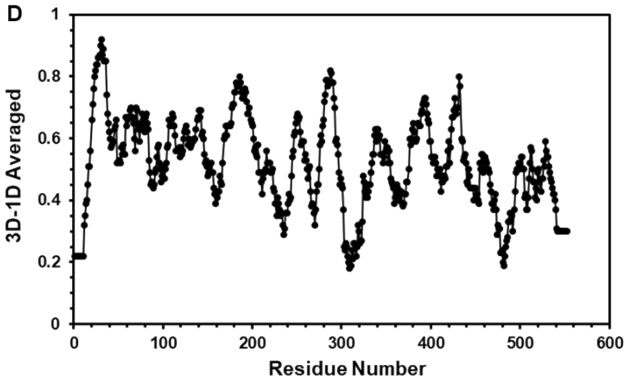
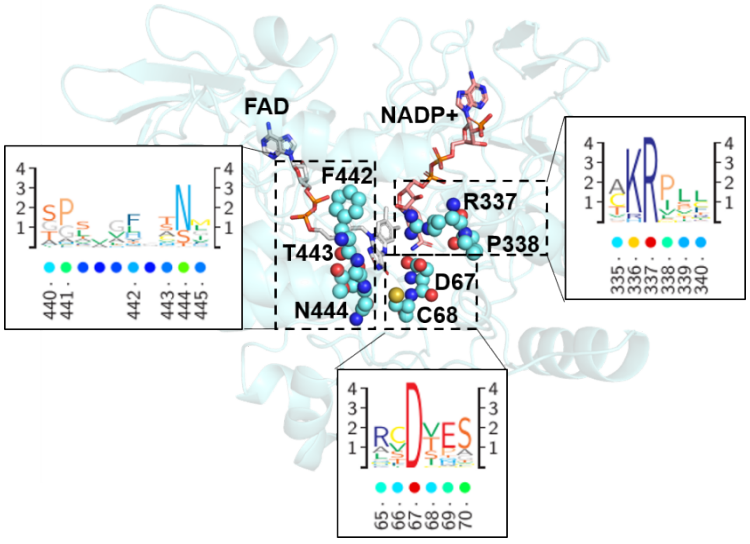


Fig. S1 Procheck-Ramachandran plots and 3D-profile window plots of the Verify 3D server for the strucure model of *Ra*BVMO by homology modelling and the corresponding Ramachandran plot statistics. **A**: Three-dimensional structure of *Ra*BVMO; **B** and **C**: Ramachandran plots for structural model of *Ra*BVMO; **D**: 3D-profile window plots for the homology model of *Ra*BVMO.

The Ramachandran plot distribution indicated that structure of *Ra*BVMO had over 90% of residues in the most favored orientation. In the structure of *Ra*BVMO the only residue (Ser^{393}) in disallowed region was later shown to be away from the catalytic cavity. Therefore, we could conclude that the structural model agreed well with the expected conformational space for residues. In addition, the 3D-1D averaged profile analysis of the Verify 3D server showed a reliable folding in most regions of the structure of *Ra*BVMO (99.28% of the residues had an averaged 3D-1D score >= 0.2). In summary, the overall assessment indicated that model of *Ra*BVMO had a satisfactory accuracy for further structure-based analysis.



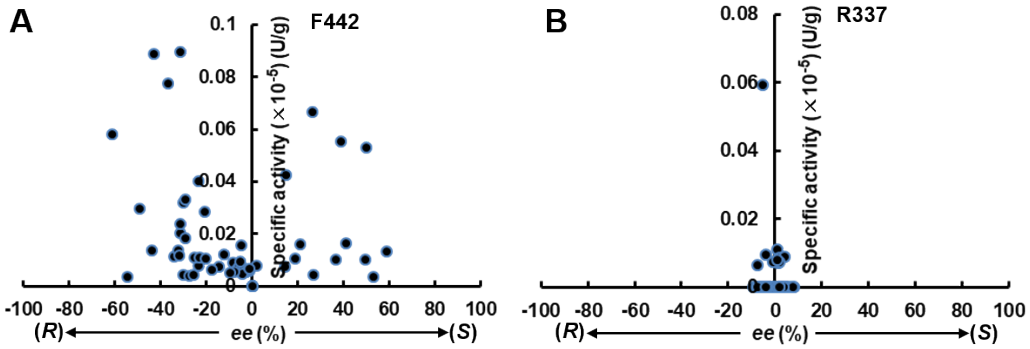
154

155 **Fig. S2** Conservation analysis of seven key sites (67, 68, 337, 338, 442, 443 and 444) of *Ra*BVMO. The residue frequencies
156 were constructed with WebLogo.

157

158

159



160

161 **Fig. S3** Screening results of saturation mutagenesis libraries of sites 442 and 337.

162

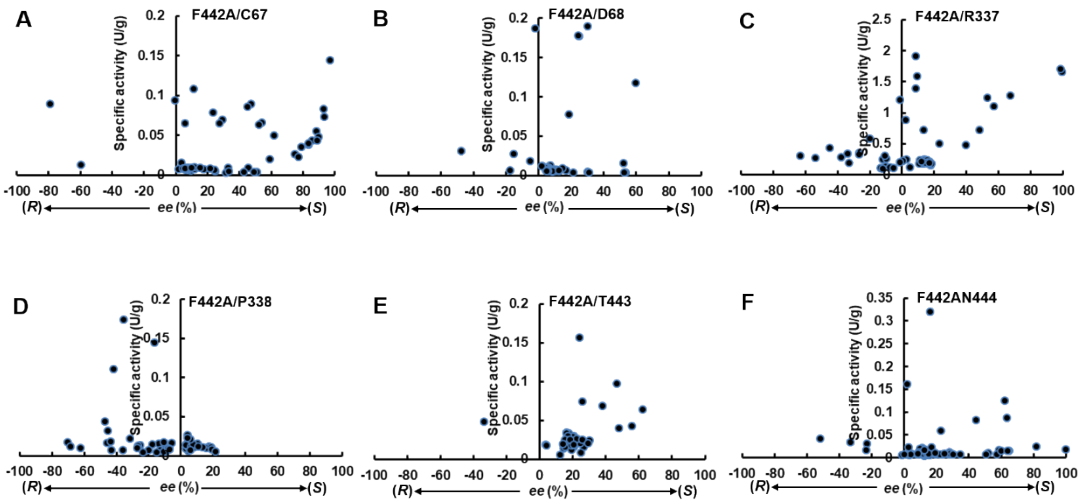


Fig. S4 Screening results of iterative saturation mutagenesis libraries of sites 67, 68, 337, 338, 443 and 444 using MT1 (F442A) as template.

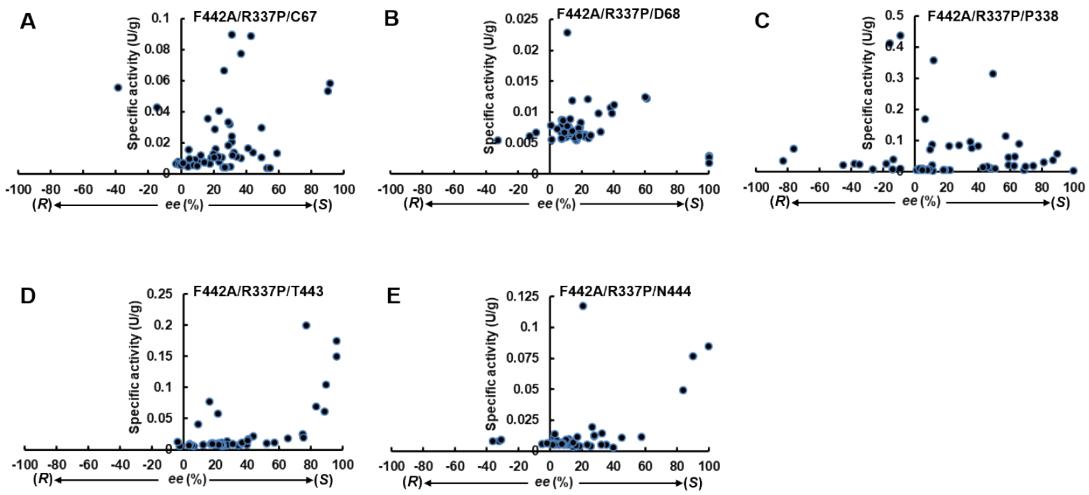


Fig. S5 Screening results of iterative saturation mutagenesis libraries of sites 67, 68, 338, 443 and 444 using MT2 (F442A/R337P) as template.

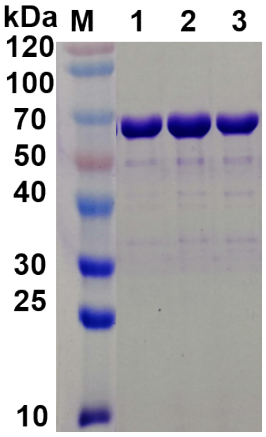


Fig. S6 SDS-PAGE (10%) analysis of the purified wild-type of *Ra*BVMO and varients. M, molecular weight marker; lane 1, wild-type *Ra*BVMO; lane 2, mutant F442A; lane 3, mutant F442A/R337P.

The wild-type of *Ra*BVMO and variants were overexpressed in *E. coli* BL21(DE3) using pET28a as expression vector. The recombinant wild-type of *Ra*BVMO and varients with an N-terminal His-tag was purified to electrophoretic homogeneity by nickel affinity chromatography. The purified enzyme migrated as a single band with a molecular weight of approximately 61 kDa on SDS-PAGE (**Fig. S6**), consistent with the theoretical molecular mass of the His-tagged protein.

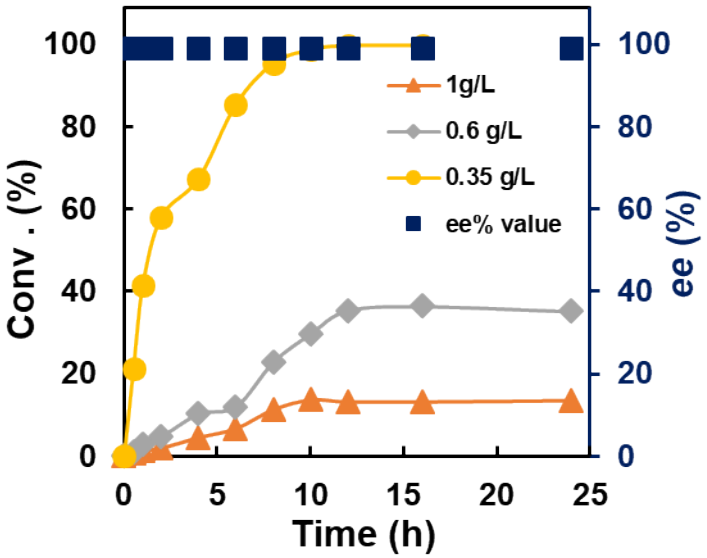


Fig. S7 Time course of sulfoxidation of omeprazole sulfide catalyzed by mutant MT2. A 10 mL of reaction mixture was composed of 0.35-, 0.6- or 1-g/L omeprazole sulfide (**1e**) as substrate (with 5% (v/v) MeOH), glucose (1.5 equiv), purified enzyme M2T (4 g/L), glucose dehydrogenase (GDH, 10 g/L), $NADP^+$ (0.2 mM), and Tris-HCl buffer (100 mM, pH 9.0) and was shaken at 180 rpm and 30 °C for 8 h. Conversion and enantiomeric excess were determined by HPLC.

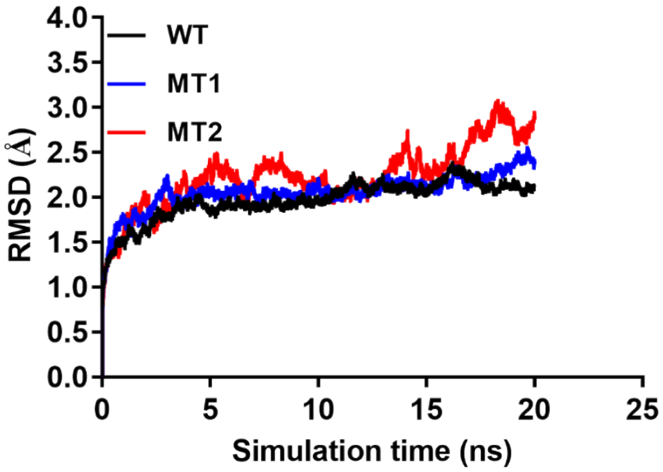
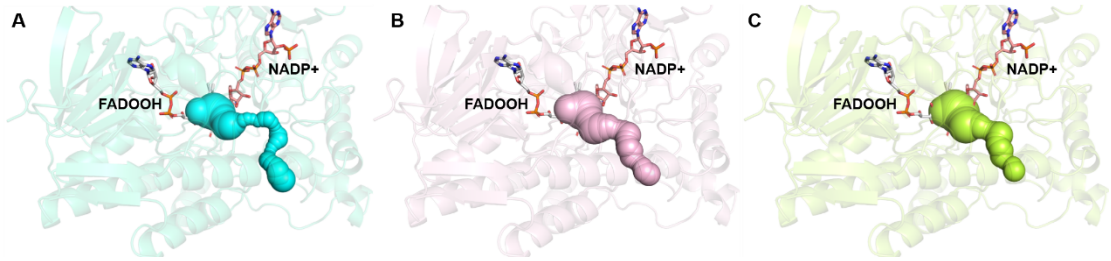


Fig. S8 Time profile of the root mean square deviations (RMSD) for the protein backbone and substrate during 20 ns MD of *Ra*BVMO WT and and its mutants MT1 (F442A) and MT2 (F442A/R337P).



196

197 **Fig. S9** CAVER analysis of substrate tunnel of *Ra*BVMO WT and its mutants based homology models. A: *Ra*BVMO WT; B:
198 MT1 (F442A); C: MT2 (F442A/R337P).

199

200

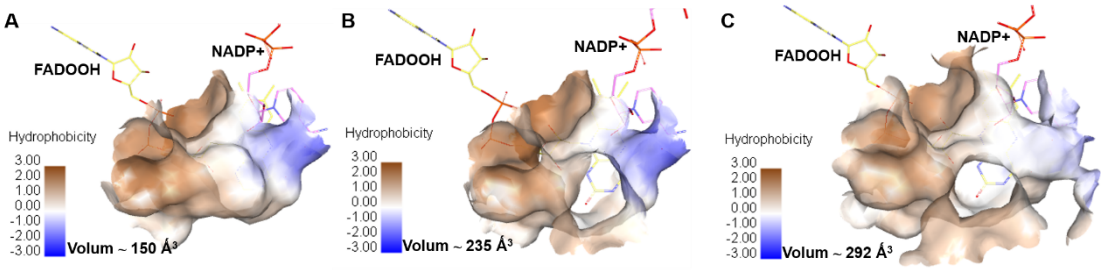


Fig. S10 Comparison of binding pocket of *Ra*BVMO WT and its mutants. A: *Ra*BVMO WT; B: MT1 (F442A); C: MT2 (F442A/R337P).

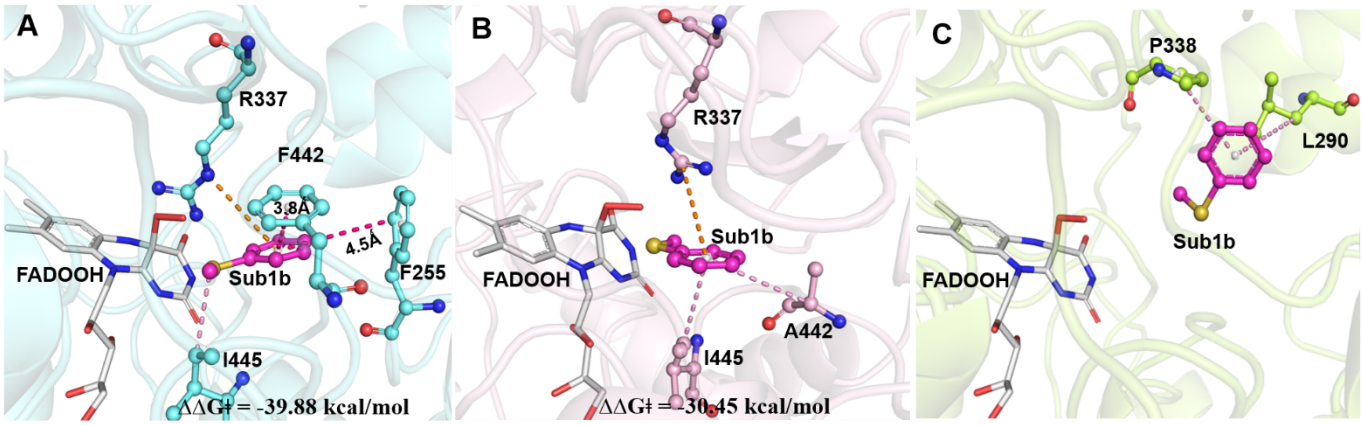
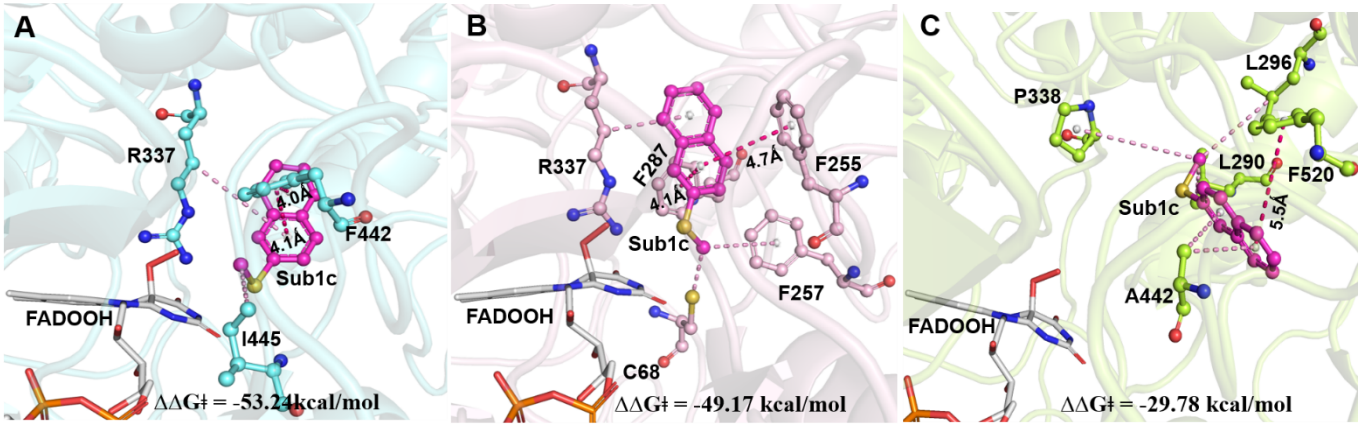


Fig. S11 Molecular docking of the substrate 1b into the active sites of *Ra*BVMO WT and its mutants. A: wild-type *Ra*BVMO; B: MT1 (F442A); C: MT2 (F442A/R337P). Hotpink dashed line: Pi-Pi T-shaped or Pi-Pi stacked; pink dashed line: alkyl or Pi-alkyl; orange dashed line: Pi-Cation.



212

213 **Fig. S12** Molecular docking of the substrate 1c into the active sites of *Ra*BVMO WT and its mutants. A: wild-type *Ra*BVMO;
214 B: MT1 (F442A); C: MT2 (F442A/R337P). Hotpink dashed line, Pi-Pi T-shaped or Pi-Pi stacked; pink dashed line, alkyl or Pi-
215 alkyl.

216

217

218

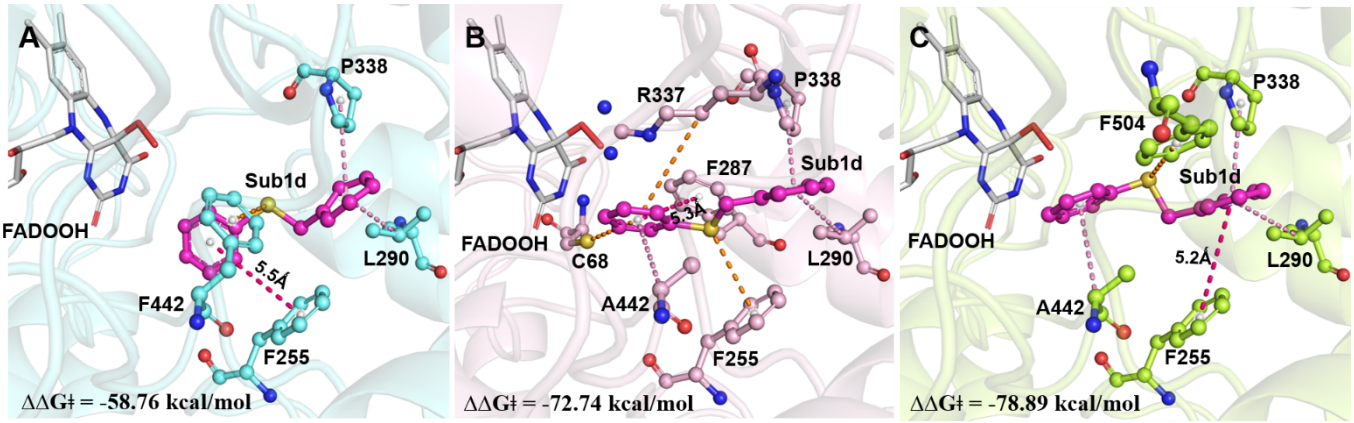


Fig. S13 Molecular docking of the substrate 1d into the active sites of *Ra*BVMO WT and its mutants. A: wild-type *Ra*BVMO; B: MT1 (F442A); C: MT2 (F442A/R337P). Hotpink dashed line, Pi-Pi T-shaped or Pi-Pi stacked; pink dashed line, alkyl or Pi-alkyl; orange dashed line, Pi-Sulfur or Pi-Cation.

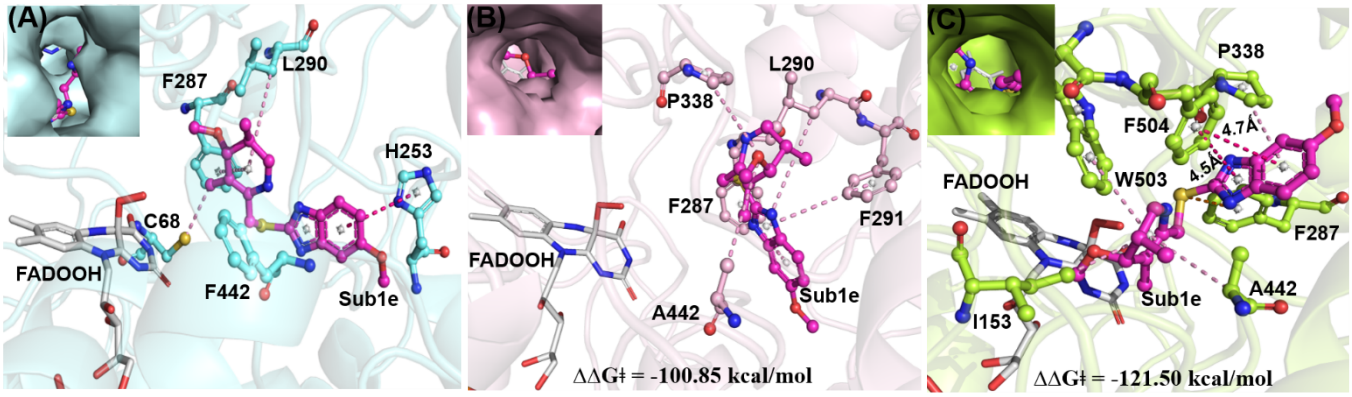
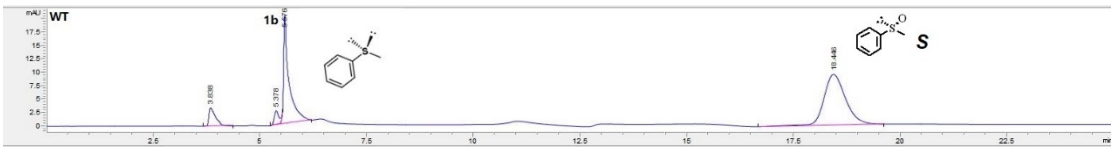
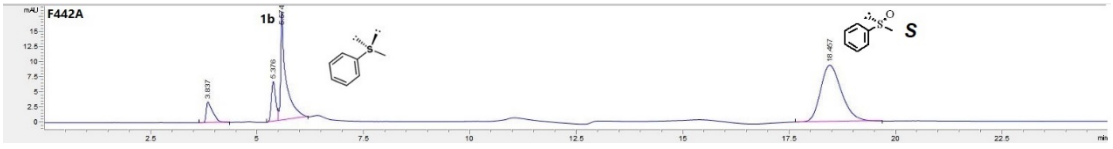


Fig. S14 Docking of substrate **1e** into the active sites of *Ra*BVMO WT and mutants. A: WT *Ra*BVMO; B: MT1 (F442A); C: MT2 (F442A/R337P). Hot pink dashed lines, Pi-Pi T-shaped or Pi-Pi stacked; pink dashed line, alkyl or Pi-alkyl; green dashed lines, hydrogen bond; orange dashed lines, Pi-Sulfu.

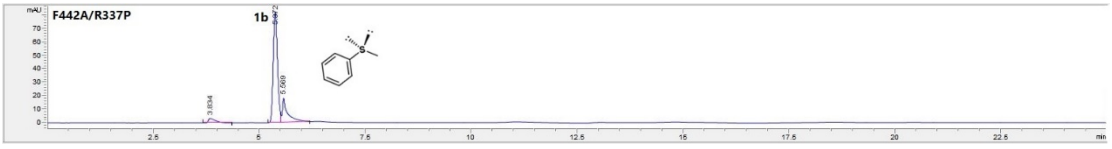
228



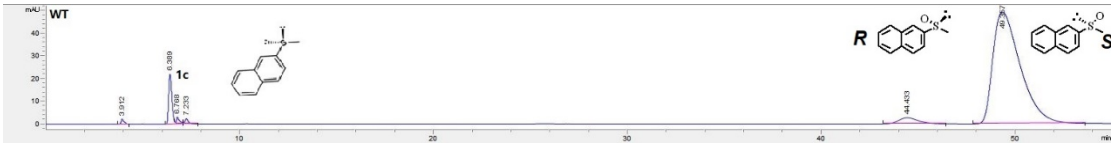
229



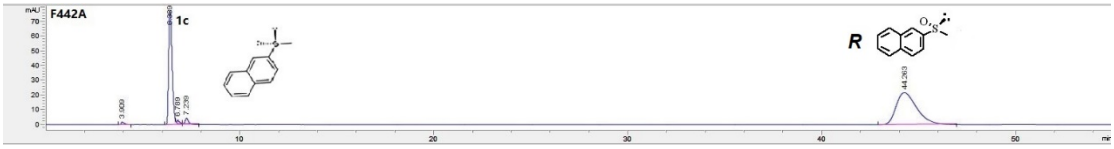
230



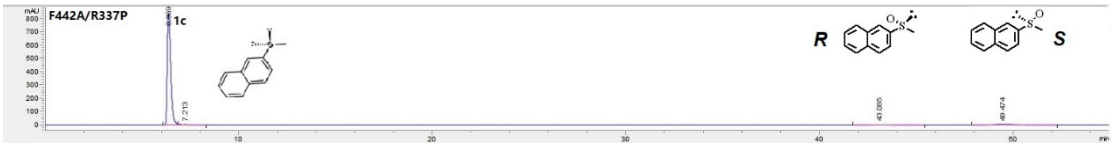
231



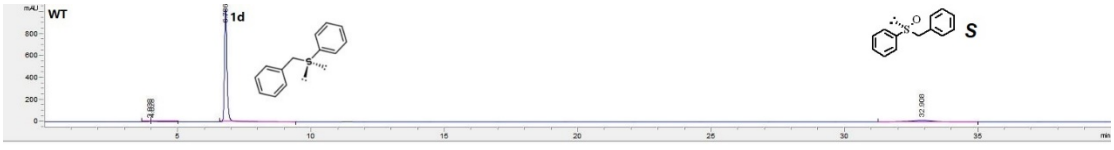
232



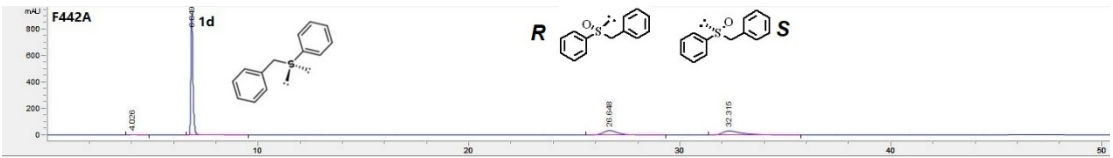
233



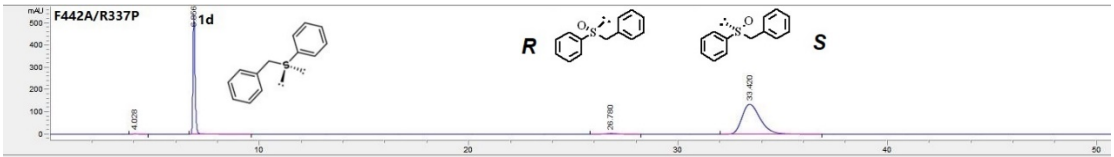
234

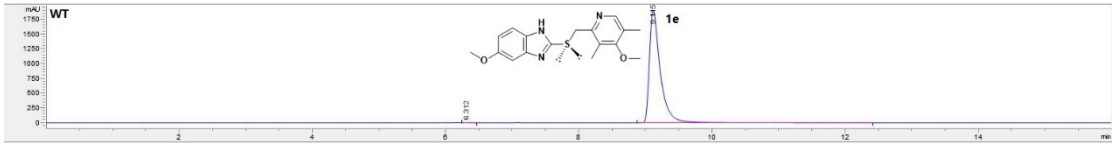


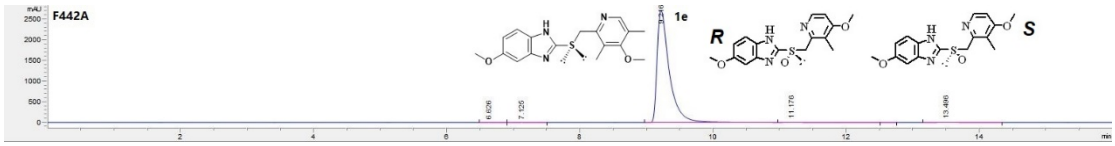
235



236







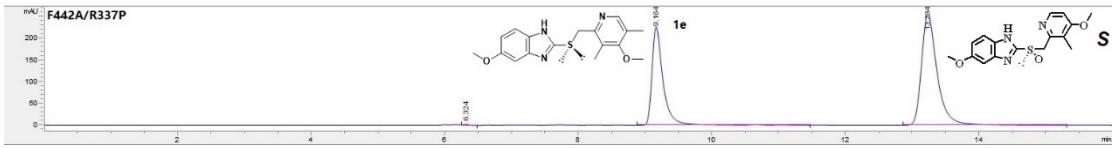
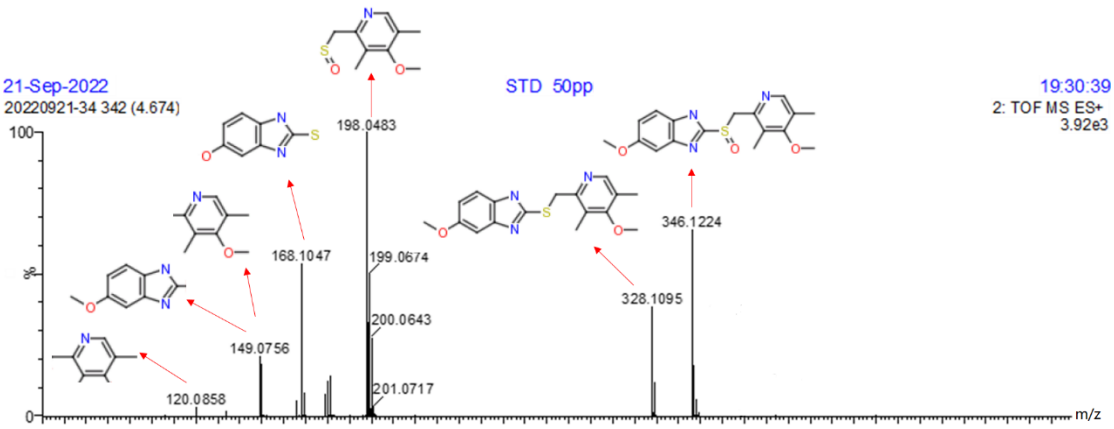


Fig. S15 Chiral HPLC spectrum of *Ra*BVMO WT and its mutants catalyzed reactions with substrates 1b-1e.



243

244 **Fig. S16** LC-MS spectrum for the reaction mixture of oxidative conversion of **1e** by MT2.

245

246

Reference

1. J. Jumper, R. Evans, A. Pritzel, T. Green and D. Hassabis, *Nature*, 2021, 1-11.
2. D. E. Kim, C. D. ylan and B. David, *Nucleic Acids Research*, 2004, **32**, 526-531.
3. E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor and J. Damborsky, *PLoS Comput Biol*, 2012, **8**, e1002708.
4. Y. Zhang, Y.-Q. Wu, N. Xu, Q. Zhao, H.-L. Yu and J.-H. Xu, *ACS Sustainable Chemistry & Engineering*, 2019, **7**, 7218-7226.
5. H. M. Dudek, M. J. Fink, A. V. Shivange, A. Dennig, M. D. Mihovilovic, U. Schwaneberg and M. W. Fraaije, *Appl Microbiol Biotechnol*, 2014, **98**, 4009-4020.
6. G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, *J Comput Aided Mol Des*, 2013, **27**, 221-234.
7. C. R. S. Mats H. M. Olsson, Michal Rostkowski, and and J. H. Jensen*, *J. Chem. Theory Comput*, **2011**, *7*, 525–537.
8. C. R. Sondergaard, M. H. Olsson, M. Rostkowski and J. H. Jensen, *J Chem Theory Comput*, 2011, **7**, 2284-2295.
9. J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio and R. M. Levy, *J Comput Chem*, 2005, **26**, 1752-1780.