Engineering of an epoxide hydrolase for efficient bioresolution of bulky pharmaco substrates

Xu-Dong Kong, Shuguang Yuan, Lin Li, She Chen, Jian-He Xu and Jiahai Zhou

Correspondence e-mail: jiahai@mail.sioc.ac.cn or jianhexu@ecust.edu.cn

Contents

Additional Materials and Methods

General

HPLC was performed on a Shimadzu LC-10AT. Optical rotations were measured in a 10 cm cell on a Rudolph Research Autopol I automatic polarimeter. The ^{1}H NMR spectra were recorded on an Agilent 400 MHz spectrometer, using the δ scale (ppm) for chemical shifts.

Expression and Purification of *Bm*EH

After 12 h induction at 16 °C with 0.5 mM isopropyl β-D-1-thiogalactopyranoside, the cells were harvested by centrifugation at 6,000 × *g* and resuspended in a sodium phosphate lysis buffer (50 mM, pH 7.5) containing 500 mM NaCl, 10 mM imidazole and 5 mM β-mercaptoethanol. The pellets were disrupted by high-pressure homogenizer and the cell debris was removed by centrifugation at 31,000 × *g* for 40 min. The soluble protein sample was loaded onto a nickel affinity column (GE Healthcare) and washed with 10-250 mM imidazole solution containing 500 mM NaCl and 50 mM sodium phosphate (pH 7.5). Proteins from the flow through were pooled and digested by thrombin for 5 h at 4 °C to remove the His tags. The obtained solution was concentrated and loaded onto a Superdex 75 Hiload 16/60 column (GE Healthcare) and eluted with 25 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM DTT. Fractions containing highly pure *Bm*EH were pooled and concentrated to 20 mg ml^{-1}.

Site-directed Mutagenesis

Mutations were introduced by PCR into the pET28-*Bm*EH template DNA (1) using the QuickChange (Stratagene) according to the manufacturer's instructions. The primers used in this study are listed in Table S4. The sequence of each mutant was confirmed by DNA sequencing. The mutant proteins were expressed and purified similarly as the wild type *Bm*EH. Among them, variants M145A, L168A, L132A, and L206A had comparable solubility as that of the wild-type *Bm*EH, while the others were expressed partially as inclusion bodies. Only the soluble fractions were chosen for purification on a nickel affinity column.

Activity Measurement with Pure Enzymes

The specific activities were measured by monitoring the conversion of substrate by HPLC. Pure enzymes were added to potassium phosphate buffer (100 mM, pH 7.0) containing 2 mM substrate with a total volume of 500 μl (10% of DMSO, 0.02% of Tween-80), and the reaction was performed at 30 °C with shaking (1000 rpm). The samples of 100 μl were withdrawn at different intervals and mixed with 400 μl of methanol to terminate the reaction. The resulting solution was analyzed by RP-HPLC (C18 column). One EH unit (U) was defined as the amount of enzyme required for the hydrolysis of 1 μmol of substrates per minute under the assay conditions.

Computational Modeling

Structures of the *Bm*EH-POA complex and the F128A-(*R*)-NPD complex were imported into the Maestro 9.2 program (2). Crystallographic solvent molecules in the binding pocket were kept. The ProPrep tool in Schrodinger software treats the p*K*a of each residue at pH 7.0 usually, which is valid for most cases. Hydrogen atoms were added to the structure according to the pH environment as reported in literature (2). The protein preparation utility in Maestro 9.2 was used to run a restrained minimization, which removed unfavorable steric contacts and improved the quality of the protein hydrogen-bonding network without large rearrangements of the protein heavy atoms.

All MD simulations were performed using the Schrodinger 2011 based Desmond package (3). The OPLS_2005 force field was applied for protein and ligand was recognized automatically by Desmond. Crystal waters at binding pocket were retained. The protein was placed in a trigonal box and solvated with TIP3P, water molecules solvated with 0.15 M NaCl with a distance of 10 Å between any protein atom and the box edge, resulting in roughly 28,000 atoms. Chloride counter-ions were added to compensate for the net positive charge of the protein. Unfavorable contacts in each system were relieved using 5000 steps of conjugate gradient energy minimization. The whole system was gradually heated from 10 K to 300 K over 100 ps at constant volume. 300 ps length equilibration step at constant pressure (NPT ensemble) was executed with position restraints on all heavy protein and ligand atoms to allow relaxation of the solvent molecules. Next, 80 ns for the *Bm*EH-POA complex and 80 ns for the F128A-(*R*)-NPD complex production MD on each system were executed, respectively. During all simulations, pressure and temperature were coupled separately for protein and

solvent atoms (including ions) to an external bath using the Berendsen coupling method at 300 K and 1 bar. Periodic boundary conditions were imposed in all three directions. All bond lengths to hydrogen atoms were constrained using M-SHAKE. The simulation time step was 2 fs. Van der Waals and short-range electrostatic interactions were cut off at 9 Å, while Long-range electrostatic interactions were calculated using the Particle Mesh Ewald. All analyses including movie-making were performed in VMD 1.91 (4).

CAVER Analysis of *Bm*EH Substrate Pathway

CAVER software package (5) facilitates the identification of routes leading from cavities in the interior of proteins to bulk solvent outside the protein matrix, via a reciprocal distance search algorithm. 100 structures of *Bm*EH extracted throughout the whole MD simulation trajectories were used for CAVER input files. The starting point for substrate access pathway search in this study was the midpoint of the active site cavity conducted in Pymol (http://www.pymol.org/) manually. Output from CAVER was finally visualized and rendered in Pymol.

References

1. Zhao J, *et al.* (2011) An unusual (*R*)-selective epoxide hydrolase with high activity for facile preparation of enantiopure glycidyl ethers. *Adv Synth Catal* 353(9): 1510-1518.
2. Banks JL, *et al.* (2005) Integrated modeling program, applied chemical theory (IMPACT). *J Comput Chem* 26(16): 1752-1780.

3. Shan Y, *et al.* (2009) A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proc Natl Acad Sci USA* 106(1): 139-144.
4. Humphrey W, Dalke A, & Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14(1): 33-38.
5. Eva Chovancov áAP, *et al.* (2012) CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput Biol.* 8(10):e1002708.

Table S1. Key residues for catalysis in *Bm*EH and other EH structures.

Enzyme (PDB ID)	Catalytic triad			Binding motif		Oxyanion hole	
*Bm*EH (4G00)	D97	H267	D239	Y144	Y203	W98	F30
*Hss*EH (1VJ5)	D333	H523	D495	Y381	Y465	W334	F265
*Mm*sEH (1CR6)	D333	H523	D495	Y381	Y465	W334	F265
*Mt*EHB (2ZJF)	D104	H333	D302	Y164	Y272	W105	F36
*St*EH (2CJP)	D105	H300	D265	Y154	Y235	W106	F32
*An*EH (3G0I)	D192	H374	D348	Y251	Y314	I193	W117
*Ar*EH (1EHY)	D107	H275	D245	Y152	Y215	F108	W30

Table S2. Activity of *Bm*EH variants for NGE and PGE.

Mutants	Specific activity for NGE (μmol/min/mg protein)	Specific activity for PGE (μmol/min/mg protein)
Wild type	0.53	57.4
F128A	22.4	1.34
L132A	7.08	14.1
M145A	13.5	0.90
M145F	0.10	12.8
L168A	0.27	40.9
L206A	0.29	1.69
L219A	0.71	18.8
F220A	0.10	7.47
F242A	0.045	0.42
F128A/M145A	0.31	0.04
T241R/L168E	0.024	11.4

Table S3. Data collection and refinement statistics.

Name	SeMet-*Bm*EH	Native-*Bm*EH	*Bm*EH-POA complex	M145A	F128A-(*R*)-NPD complex
Data collection					
Wavelength (Å)	0.9791	0.9791	1.5418	1.5418	1.5418
Space group	$P2_12_12_1$	$P2_12_12_1$	$P4_12_12$	$P2_12_12_1$	$P4_12_12$
Cell dimensions (Å)	a=63.6 b=75.7 c=121.4	a=63.2 b=76.0 c=120.8	a=108.5 b=108.5 c=116.8	a=63.8 b=77.7 c=120.0	a=109.1 b=109.1 c=117.5
Resolution (Å)[a]	1.85 (1.92-1.85)	1.75 (1.81-1.75)	1.95 (2.02-1.95)	1.70 (1.76-1.70)	2.90 (3.00-2.90)
No of measured reflections	544886	749929	669808	416525	121961
No of unique reflections [a]	50337 (4941)	59830(5452)	51362(5050)	65165 (6501)	16484 (1601)
Redundancy [a]	10.8 (11.0)	12.5 (7.4)	13 (12.4)	6.4 (6.4)	7.4 (7.5)
Completeness (%)[a]	99.9 (100)	99.0 (91.5)	99.7(99.6)	98.2 (99.1)	100.0 (100.0)
Average (I/σ) [a]	39.8 (5.5)	26.5 (2.5)	24.6 (4.1)	16 (4.3)	8.3 (2.6)
R_{merge} (%)[a, b]	0.093 (0.498)	0.090 (0.643)	0.094(0.640)	0.050 (0.439)	0.221 (0.790)
Refinement					
No of reflections	49252	57830	50352	64070	15898
R_{work}/R_{free} [c]	0.1771/0.2109	0.1861/0.2330	0.1695/0.2071	0.1575 /0.2091	0.1560/0.2316
No of non-H atoms					
protein	4784	4862	4788	4740	4737
waters	345	423	493	394	100
ligand	0	0	68	18	42
Average B factor [A^2]	29.33	30.93	29.62	27.92	33.02
RMS deviations					
Bond lengths (Å)	0.007	0.007	0.007	0.005	0.008
Bond angles (°)	1.039	0.995	1.035	0.939	1.117
Ramachandran plot favored (%)	96.9	97.6	97.0	97.0	95.4
Ramachandran plot allowed (%)	2.9	2.1	2.7	2.7	4.4
Ramachandran plot outliers (%)	0.2	0.3	0.3	0.3	0.2

[a] Numbers in parentheses are values for the highest-resolution shell.

[b] $R_{merge} = \sum_{hkl}\sum_i|I_i - \langle I\rangle|/\sum_{hkl}\sum_i|\langle I\rangle|$, where I_i is the intensity for the *i*th measurement of an equivalent reflection with indices h, k, and l.

[c] R_{free} was calculated with the 5% of reflections set aside randomly throughout the refinement.

Table S4. Primers used for site-directed mutagenesis.

Primer	5' to 3' Sequence
F128A forward	CCGCACCCTTATACAGCTATGCGTGAGC
F128A reverse	GCTCACGCATAGCTGTATAAGGGTGCGG
L132A forward	CATTTATGCGTGAGGCTAGAACAAATAAAAATC
L132A reverse	GATTTTTATTTGTTCTAGCCTCACGCATAAATG
M145A forward	GCAAGCGAGTATGCGAAATGGTTTC
M145A reverse	GAAACCATTTCGCATACTCGCTTGC
L168A forward	GGTTTGCGTAAGGCTGTGATCGACC
L168A reverse	GGTCGATCACAGCCTTACGCAAACC
L168E forward	GGTTTGCGTAAGGAAGTGATCGACC
L168E reverse	GGTCGATCACTTCCTTACGCAAACC
L206A forward	CCGCAATGCAAAGATTTTCACGGAAG
L206A reverse	CTTCCGTGAAAATCTTTGCATTGCGG
L219A forward	CCTTCGAAGAAAATCTGCGTTTCCGTTAGAAG
L219A reverse	CTTCTAACGGAAACGCAGATTTTCTTCGAAGG
F220A forward	CGAAGAAAATCTTTGGCTCCGTTAGAAGAAG
F220A reverse	CTTCTTCTAACGGAGCCAAAGATTTTCTTCG
T241R forward	GGAAATCAAGATCCTAGGTTTATGCCTG
T241R reverse	CAGGCATAAACCTAGGATCTTGATTTCC
F242A forward	GGGGAAATCAAGATCCTACTGCTATGCCTG
F242A reverse	CAGGCATAGCAGTAGGATCTTGATTTCCCC

Supporting Figure Legends

Figure S1. Structural comparison of *Bm*EH with human epoxide hydrolase and several other α/β hydrolases. Superposition of *Bm*EH (blue) with human soluble epoxide hydrolase (3ANT, green), a putative acyl transferase from *Salmonella enteric* (4NVR, orange), fluoroacetate dehalogenase RPA1163 (3R41, yellow), and aryl esterase from *Pseudomonas fluorescens* (3T4U, cyans). All structures have a conserved catalytic domain with α/β hydrolase fold, while more differences are observed in lid domain.

Figure S2. Overall perspective of POA binding ($2|F_o|-|F_c|$ density at 1σ; blue mesh) in the *Bm*EH-POA complex structure. Bound POAs in zone 1, zone 2 and zone 3 are shown in yellow, green and cyan, respectively.

Figure S3. Zone 2 is accessible from zone 1, and very close to but detached from zone 3. The regions of Zone 1, Zone 2 and Zone 3 were defined by yellow, red and green dash lines, respectively.

Figure S4. Putty cartoon of B-factor variation on the mean structure of *Bm*EH, presented from low to high (thin to thick). Zone 1 and zone 2 are regions with low B-factor while zone 3 is a high B-factor area. Residues 125-156 and 204-225 are colored in blue and yellow, respectively.

Figure S5. Molecular dynamics simulations of *Bm*EH-POA complex. *A*) Root-mean-square fluctuation of backbone atoms (N, Cα and C atoms) during the 80 ns molecular dynamics (MD) simulations. Residue 123-156 and 204-225 are framed with green dash and blue dash,

respectively. *B*) Comparison of the Root-mean-square deviation of backbone heavy atoms during 80 ns MD simulations within reference at the crystal structure. Black: the overall *Bm*EH; green: residues 123-156; blue: residues 204-225.

Figure S6. Reaction and catalytic mechanism catalyzed by *Bm*EH. *A*) The bioresolution of PGE or NGE catalyzed by *Bm*EH. *B*) The molecular mechanism of *Bm*EH catalysis. After the initial binding of the substrate in the active center, the carboxylate side chain of the catalytic nucleophile Asp97 in *Bm*EH attacks the unsubstituted carbon atom of the oxirane ring to form a covalent ester intermediate. The process is facilitated through simultaneous proton donation by Tyr144 or Tyr203. The second step is the hydrolysis of the ester intermediate, which is attacked by a water molecule activated through proton abstraction with the aid of the His267/Asp239 charge-relay system. Therefore, the inactivation mutant H267F can capture the transition state by terminating the reaction right before the hydrolysis of the ester intermediate.

Figure S7. Detection of epoxide hydrolysis reaction progress by stopped-flow spectrofluorimeter. *A*) Fluorescence spectrum of NGE and NPD excited at 310 nm. *B*-*C*) Kinetic trace of 25 μM wild-type (*B*), F128A (*C* red line) or M145A (*C* green line) catalyzed hydrolysis of (*R*)-NGE (125 μM). Each kinetic trace was the average data of five individual experiments. *D*) Kinetic trace of H267F (25 μM) catalyzed hydrolysis of (*R*)-NGE (25 μM). We proposed the decrease of fluorescence signal implied the formation of intermediate and the slow increase of signal was due to the hydrolysis of intermediate.

Figure S8. Kinetic mechanism proposed for enzymatic reactions.

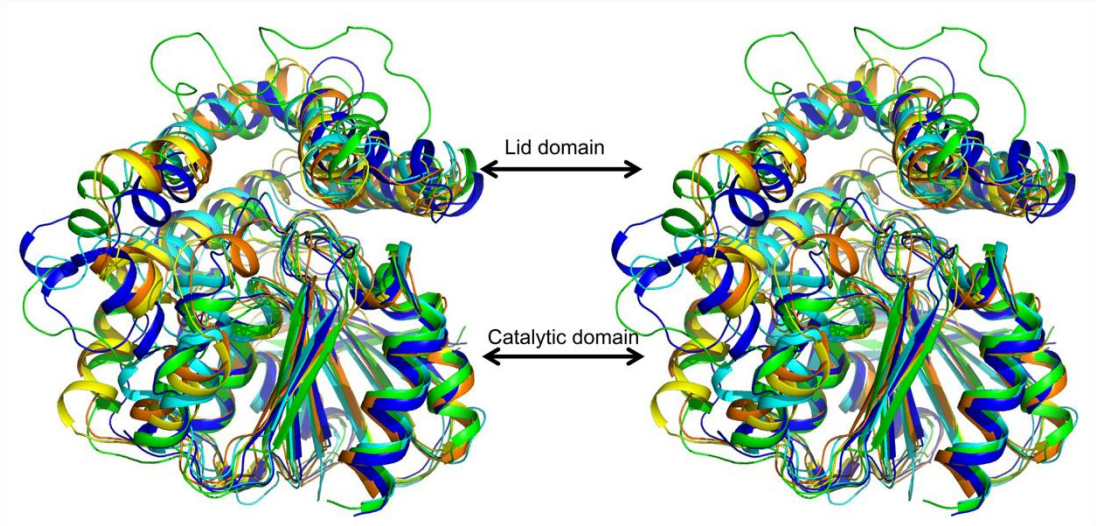

Figure S1.

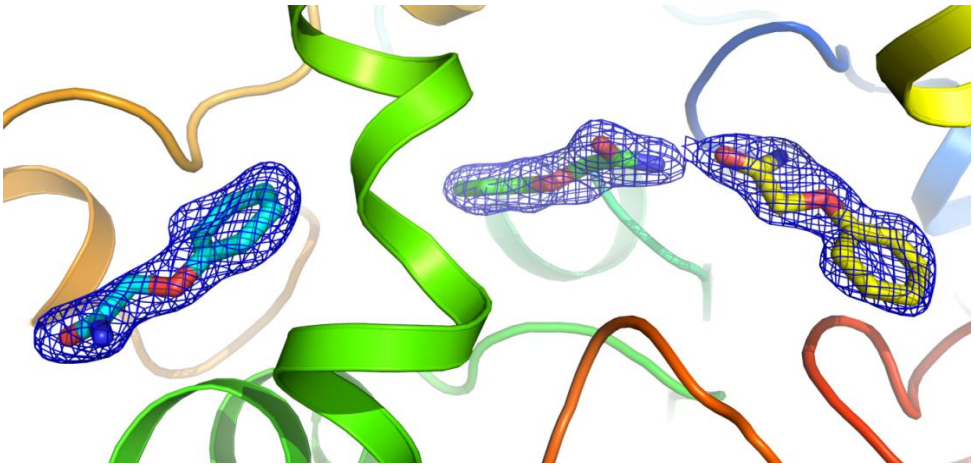

Figure S2.

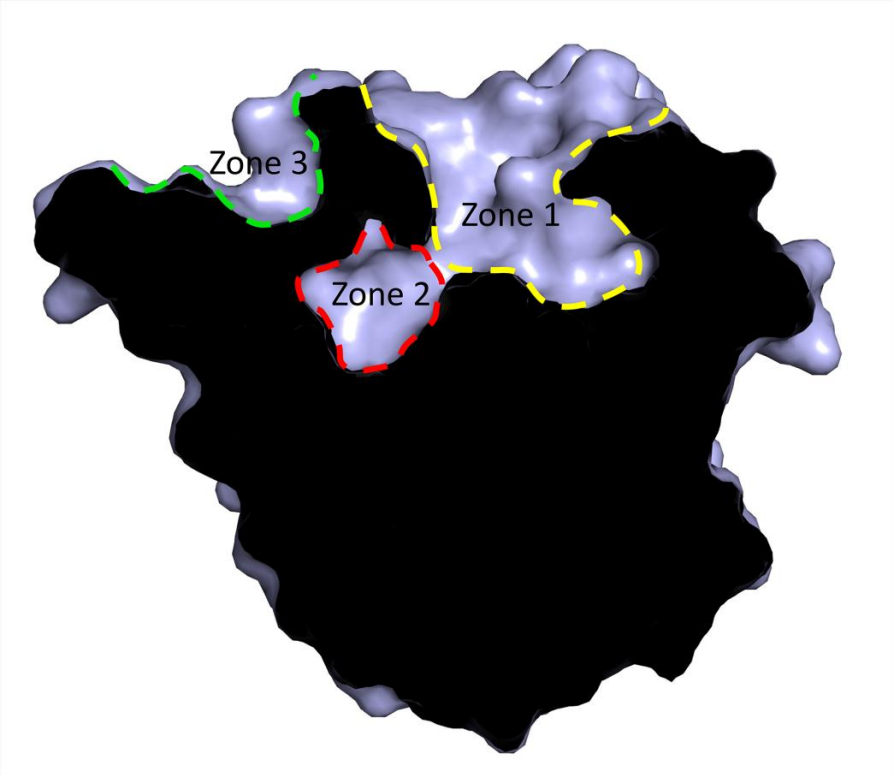

Figure S3.

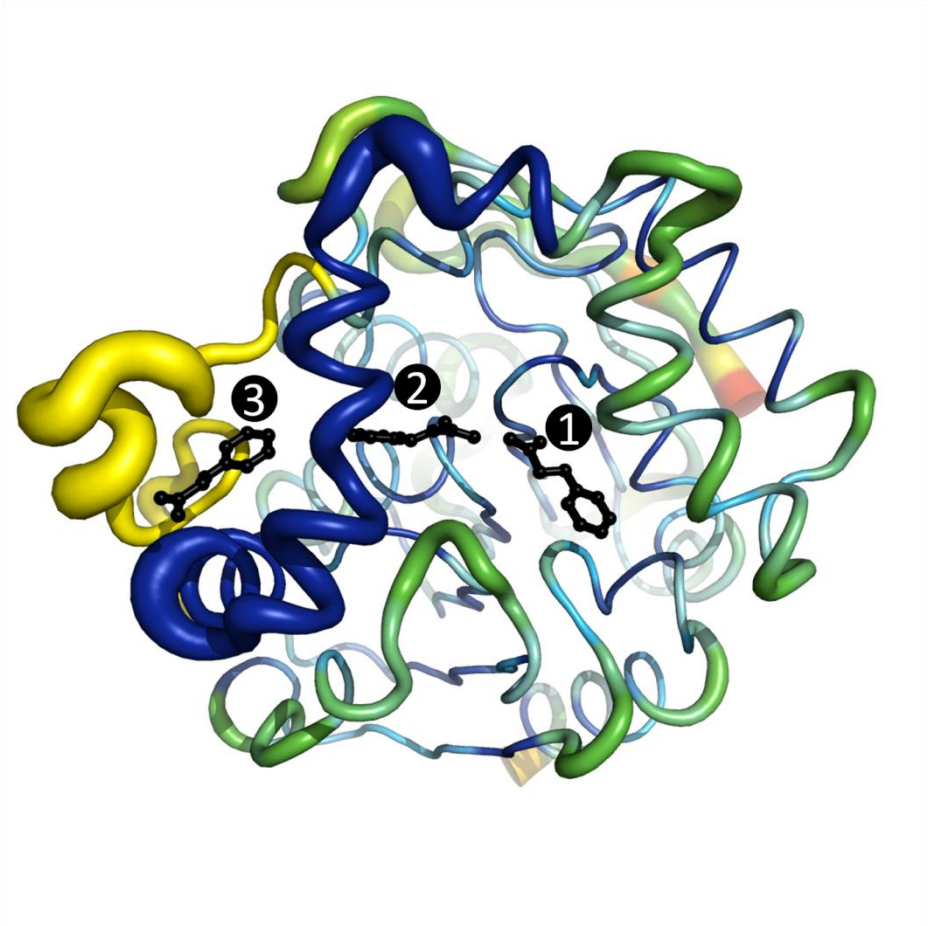

Figure S4.

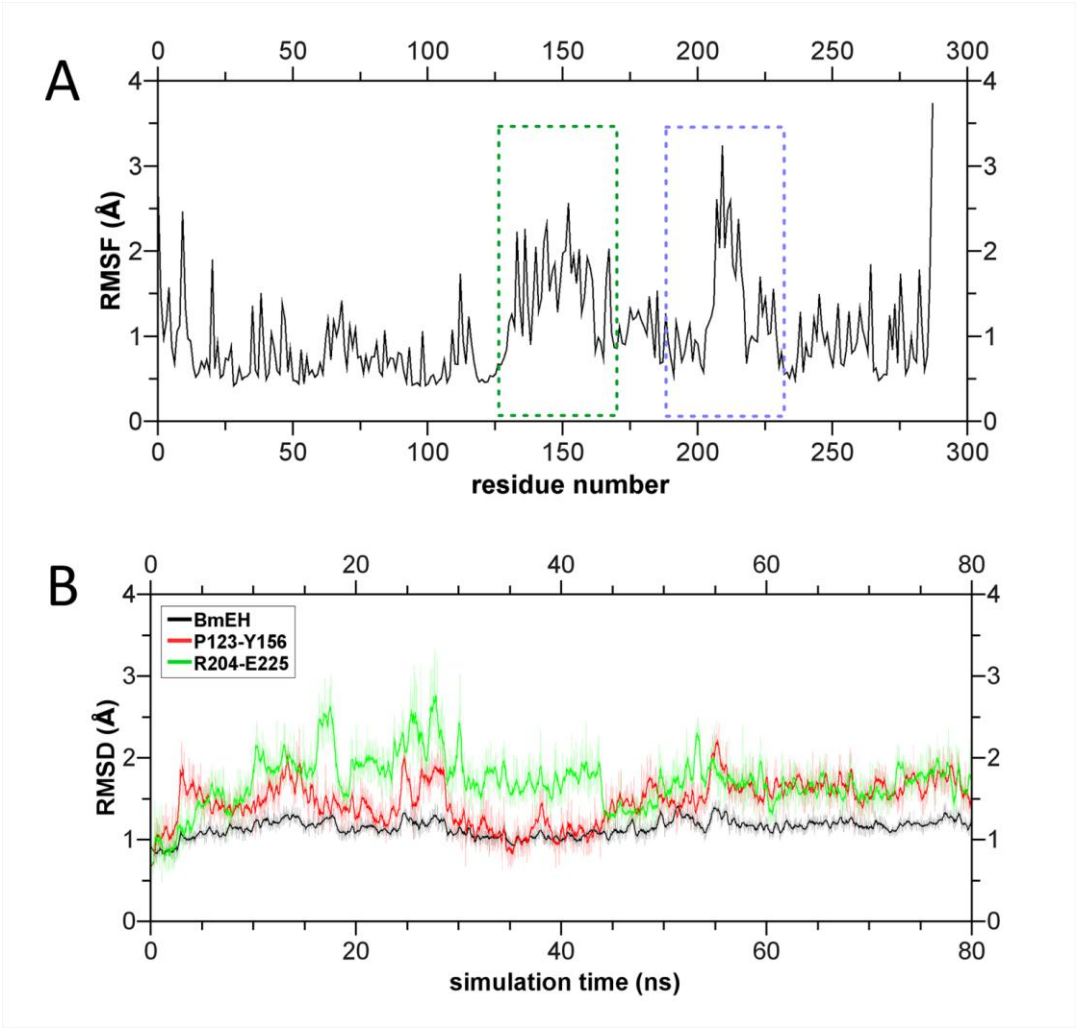
A
RMSF (Å)
residue number
B
RMSD (Å)
simulation time (ns)
BmEH
P123-Y156
R204-E225

Figure S5.

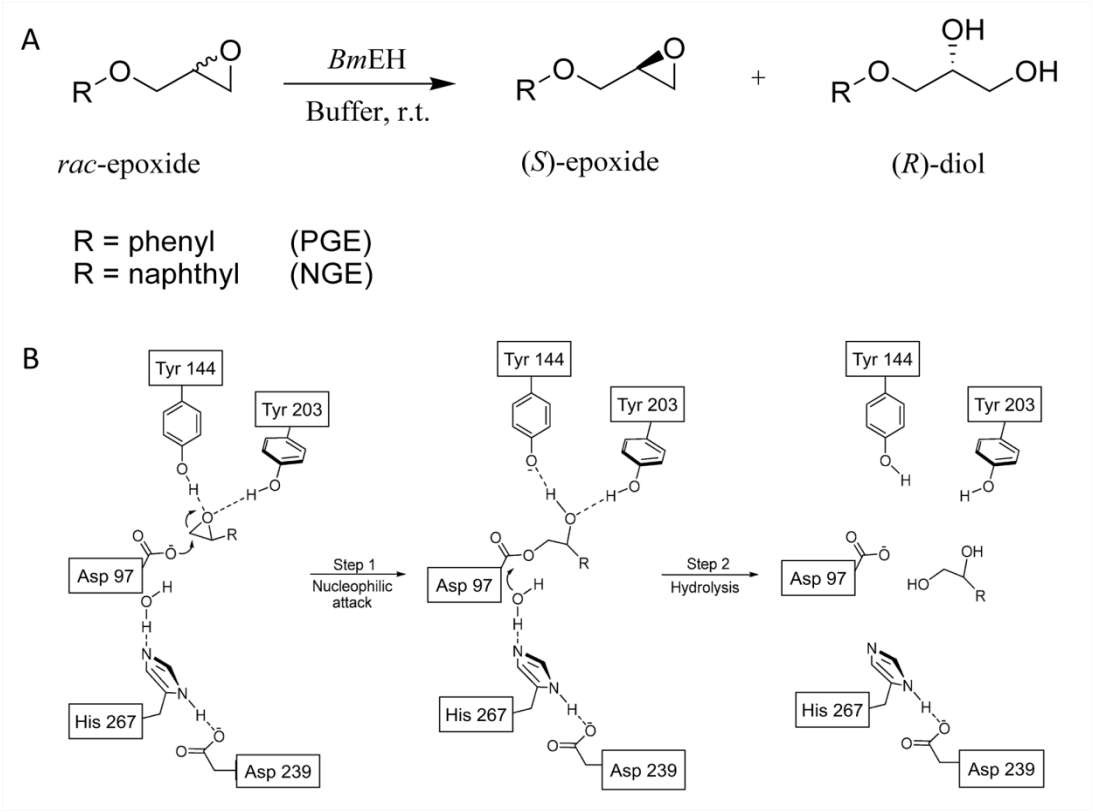
A
BmEH
Buffer, r.t.
rac-epoxide
(S)-epoxide
(R)-diol
R = phenyl (PGE)
R = naphthyl (NGE)
B
Tyr 144
Tyr 203
Asp 97
His 267
Asp 239
Step 1
Nucleophilic attack
Step 2
Hydrolysis

Figure S6.

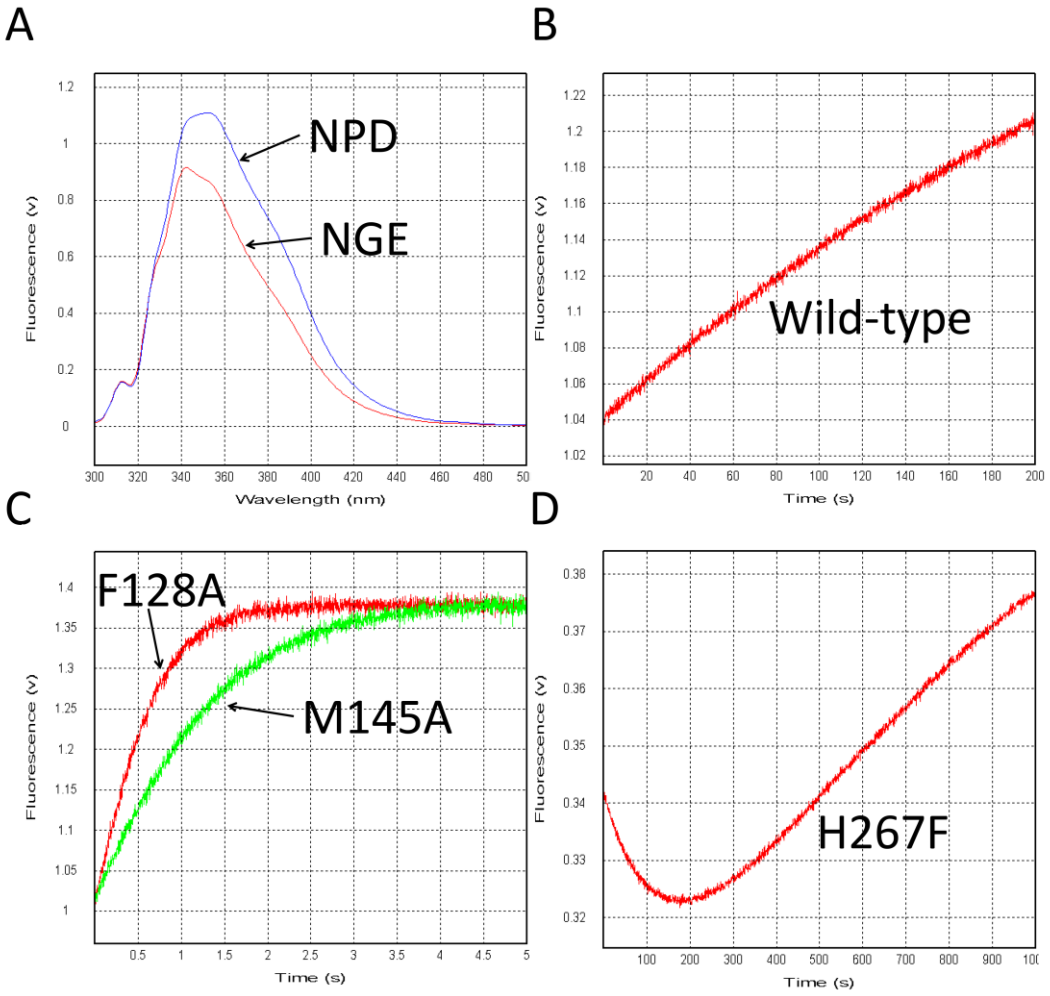
A
B
C
D
NPD
NGE
Wild-type
F128A
M145A
H267F
Fluorescence (v)
Wavelength (nm)
Time (s)

Figure S7

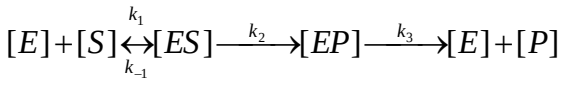

Figure S8.