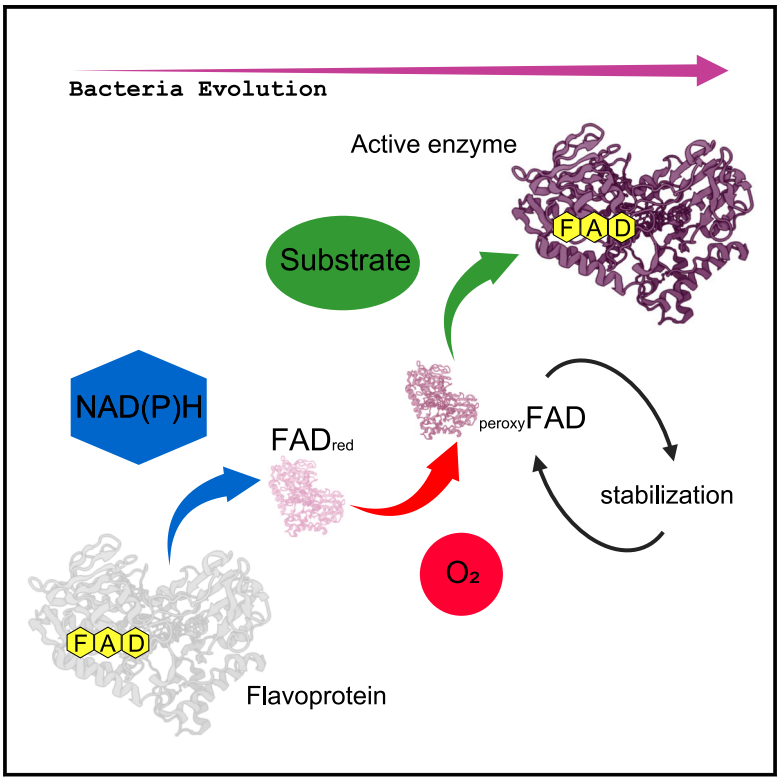
Cell Reports

Report

Evolution of the catalytic mechanism at the dawn of the Baeyer-Villiger monooxygenases

Graphical abstract



Highlights

- BVMOs' mechanism evolved through a series of slightly overlapped, discrete steps
- Catalytic elements of the mechanism were recruited and later function optimized
- The hydride transfer process precedes establishment of the nicotinamide specificity
- AncBVMOs evolved oxygen reactivity contingent on the hydride donor specificity

Authors

Guang Yang, Ognjen Pećanac, Hein J. Wijma, Henriëtte J. Rozeboom, Gonzalo de Gonzalo, Marco W. Fraaije, Maria Laura Mascotti

Correspondence

m.l.mascotti@rug.nl

In brief

Yang et al. report how the complex catalytic mechanism of bacterial BVMOs evolved. AncBVMOs increased their activity over time contingent on the hydride donor usage and improvement of oxygen reactivity. The network of residues in the active site, the $NADP^+$ cofactor, and global conformational changes govern the functionality.

Yang et al., 2024, Cell Reports *43*, 114130
May 28, 2024
https://doi.org/10.1016/j.celrep.2024.114130

CellPress

Cell Reports

CellPress
OPEN ACCESS

Report

Evolution of the catalytic mechanism at the dawn of the Baeyer-Villiger monooxygenases

Guang Yang,[1] Ognjen Pećanac,[1] Hein J. Wijma,[1] Henriëtte J. Rozeboom,[1] Gonzalo de Gonzalo,[2] Marco W. Fraaije,[1] and Maria Laura Mascotti[1,3,4,5,*]

[1]Molecular Enzymology Group, University of Groningen, 9747 AG Groningen, the Netherlands
[2]Departamento de Química Orgánica, Universidad de Sevilla, and Centro de Innovación en Química Avanzada (ORFEO−CINQA), 41012 Sevilla, Spain
[3]IMIBIO-SL CONICET, Facultad de Química Bioquímica y Farmacia, Universidad Nacional de San Luis, San Luis, Argentina
[4]Present address: IHEM CONICET, Universidad Nacional de Cuyo, M5502JMA Mendoza, Argentina
[5]Lead contact
*Correspondence: m.l.mascotti@rug.nl
https://doi.org/10.1016/j.celrep.2024.114130

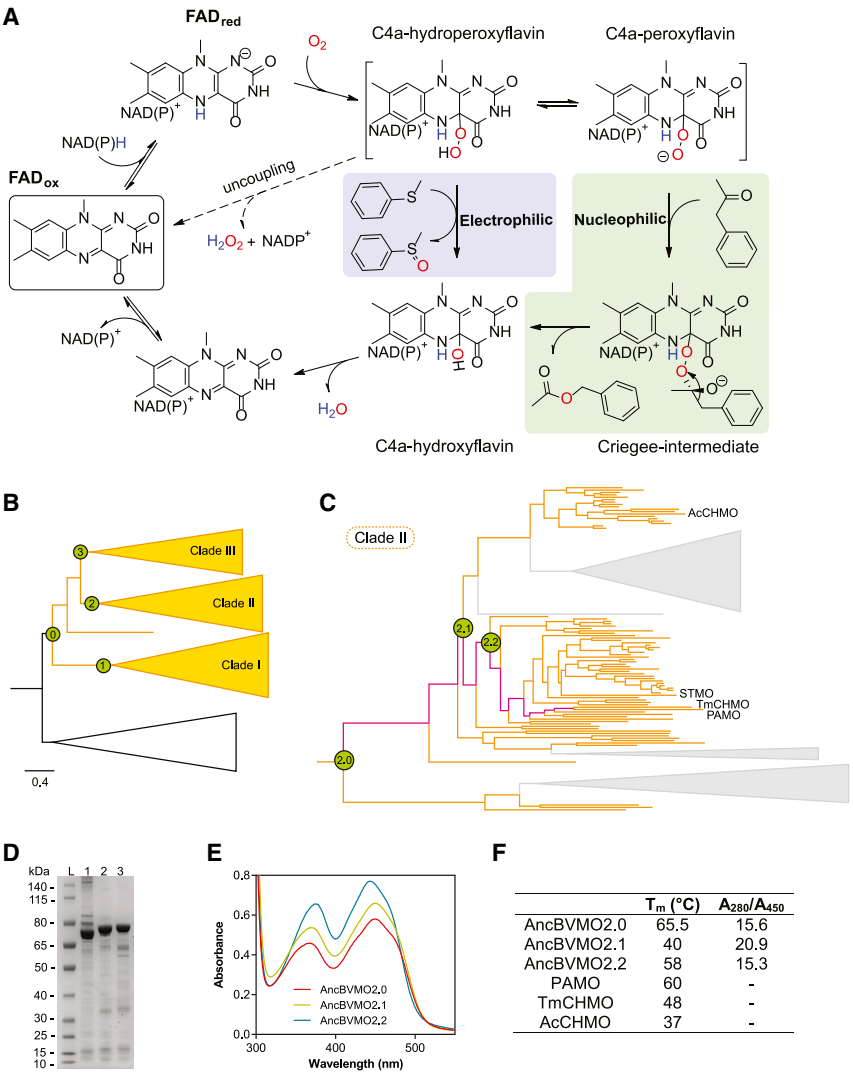
SUMMARY

Enzymes are crucial for the emergence and sustenance of life on earth. How they became catalytically active during their evolution is still an open question. Two opposite explanations are plausible: acquiring a mechanism in a series of discrete steps or all at once in a single evolutionary event. Here, we use molecular phylogeny, ancestral sequence reconstruction, and biochemical characterization to follow the evolution of a specialized group of flavoprotein monooxygenases, the bacterial Baeyer-Villiger monooxygenases (BVMOs). These enzymes catalyze an intricate chemical reaction relying on three different elements: a reduced nicotinamide cofactor, dioxygen, and a substrate. Characterization of ancestral BVMOs shows that the catalytic mechanism evolved in a series of steps starting from a FAD-binding protein and further acquiring reactivity and specificity toward each of the elements participating in the reaction. Together, the results of our work portray how an intrinsically complex catalytic mechanism emerged during evolution.

INTRODUCTION

Flavoprotein monooxygenases are key enzymes for cell survival.[1] They fulfill a plethora of roles, ranging from xenobiotic detoxification to hormone biosynthesis.[1–4] Several different classes of flavoprotein monooxygenases can be recognized based on sequence features and their mechanistic properties.[1] The class B flavoprotein monooxygenases (also known as two dinucleotide binding domain monooxygenases, tDNBD MOXs) are widespread in nature and have been studied in detail.[1,5] They display a catalytic mechanism consisting of a series of events involving several actors: a tightly bound flavin adenine dinucleotide (FAD) cofactor, a dissociable nicotinamide coenzyme as hydride donor, O_2 as oxidant, the substrate/product pair, and a constellation of residues in the active site facilitating the oxygenating process.[6] The first step in the catalytic cycle involves the reduction of FAD, and this then reacts with molecular oxygen to form the key reactive enzyme intermediate, a (hydro)peroxyflavin (Figure 1A). At this point, the enzyme is primed to react with a substrate that can reach the active site.[7] Dependent on the nature of the substrate, two fundamentally different chemistries are possible. Heteroatom-containing molecules are oxidized via an electrophilic mechanism. This is the canonical activity performed by two kinds of enzymes in the family, the flavin-containing monooxygenases (FMOs)[3,8] and the N-hydroxyl monooxygenases (NMOs).[4,9] However, if a ketone reaches the active site, there is a specialized group called Baeyer-Villiger monooxygenases (BVMOs), which can catalyze its oxidation via a nucleophilic mechanism. A substrate-flavin adduct—the Criegee intermediate—is formed that allows the migration of a carbon for the insertion of an oxygen atom next to the carbonyl moiety (BV oxidation).[2,10] Regardless of the type of reaction catalyzed, the last step is the release of the oxidized hydride donor molecule ($NADP^+$), which often defines the overall rate of catalysis.[11,12] How such a complex catalytic mechanism arose during evolution is an open question.

For some enzymes with simpler mechanisms, like the cyclohexadienyl dehydratases[13] or the chalcone isomerases,[14] it was found that they evolved from non-catalytic, organic molecule-binding proteins via reshaping and further refinement of the active site. For the above-mentioned flavoprotein monooxygenases, it is harder to envision a sequence of gradual steps allowing activity to emerge, as each catalytic element is essential for the mechanism. However, the fact that other enzymes belonging to the tDNBD superfamily do not use oxygen as an electron acceptor,[5] such as the alkylhydroperoxide reductases, disulfide reductases, and NADH dehydrogenases,[15–17] suggests that the monooxygenases may somehow have gained the ability to use O_2 at some point in their history. One further level of complexity is posed by the two possible chemistries upon the formation of the key reactive enzyme intermediate. In principle, the difference relies on the protonation state of the enzyme



	T_m (°C)	A_{280}/A_{450}
AncBVMO2.0	65.5	15.6
AncBVMO2.1	40	20.9
AncBVMO2.2	58	15.3
PAMO	60	-
TmCHMO	48	-
AcCHMO	37	-

Figure 1. The BVMOs diverged in three clades: clade II AncBVMOs can be purified recombinantly

(A) The catalytic cycle of BVMOs.

(B) Condensed phylogeny of BVMOs. BVMOs (yellow clades), FMOs (outgroup, in black), and reconstructed ancestors (numbered green circles). The scale bar indicates substitutions per site. Also see Table S1 and Figure S1.

(C) Closer view of clade II. The gray condensed clades include only fungal sequences. The trajectory followed is shown in pink. Also see Figure S2.

(D) SDS-PAGE analysis of purified AncBVMOs. Lanes: (L) protein ladder, (1) AncBVMO2.0, (2) AncBVMO2.1, (3) AncBVMO2.2.

(E) Absorbance spectra of AncBVMOs recorded with 50–80 μM enzyme samples.

(F) Melting temperatures and A_{280}/A_{450} ratios of AncBVMOs. T_m values were determined in triplicate and represented as mean.

intermediate, being the hydroperoxyflavin required for the electrophilic mechanism (FMOs and NMOs) or the peroxyflavin for the nucleophilic one (BVMOs).[10,12,18,19] However, the sequence determinants for that functional diversity seem intricate and likely to involve hard-to-predict long-range interactions.[20] Interestingly, the ability to perform BV oxidations has emerged at least three times in the class B flavoprotein monooxygenases. The FMO5 clade of tetrapods (one of the metazoan paralogs involved in detoxification and defense metabolism)[20] and the YUCCAs from plants (involved in auxin biosynthesis)[21,22] carry out this activity as a main functionality. The third case is the BVMO group. FMO5s and YUCCAs are restricted to specific taxonomic groups, while BVMOs have been more successful in spreading, with representatives in numerous bacteria, fungi, and archaea.

Here, we investigated the evolution of the catalytic mechanism in a clade of bacterial BVMOs. We reconstructed several ancestors from the lineage leading to the modern and well-characterized BVMOs: PAMO (phenylacetone monooxygenase from *Thermobifida fusca*),[23–26] *Tm*CHMO (a cyclohexanone monooxygenase from *Thermocrispum municipale*),[27] and STMO (a steroid monooxygenase from *Rhodococcus rhodochrous*).[28] We found that the catalytic mechanism was established over a process of stepwise recruiting of catalytic elements and subsequent functional optimization.

RESULTS

The BVMOs diverged in three clades: Clade II AncBVMOs can be produced recombinantly

We performed a robust phylogenetic analysis by constructing a representative dataset including all characterized sequences to the date and homologs from bacterial classes, eukaryotic phyla, and archaeal groups (~500 sequences) (Table S1). To validate the identity of the hits as true BVMOs, the presence of the canonical fingerprints[29] was analyzed. The robustness of the phylogeny was assessed by using 500–1,000 bootstrap replicates on each run, and the topology of the tree was stable upon independent inferences as well as to the addition/removal of some sequences.

The phylogeny of BVMOs (Figures 1B and S1) shows a complex topology, mainly due to duplication and loss events as well as horizontal gene transfer events. The taxonomic distribution, though broad, is largely biased toward soil microorganisms like fungal and bacterial species, and this resonates with the role of class B flavoprotein monooxygenases in detoxification and defense metabolism.[1–4] Three major clades, evenly populated, are identified, named here I–III. Clade I (transfer bootstrap expectation [TBE] = 1) is largely composed by proteobacterial and actinobacterial sequences. The antitubercular prodrug—ethionamide—activator (EtaA) from *Mycobacterium tuberculosis*[30,31] and the *n*-alkane degrading almA from *Acinetobacter* sp.[32] are

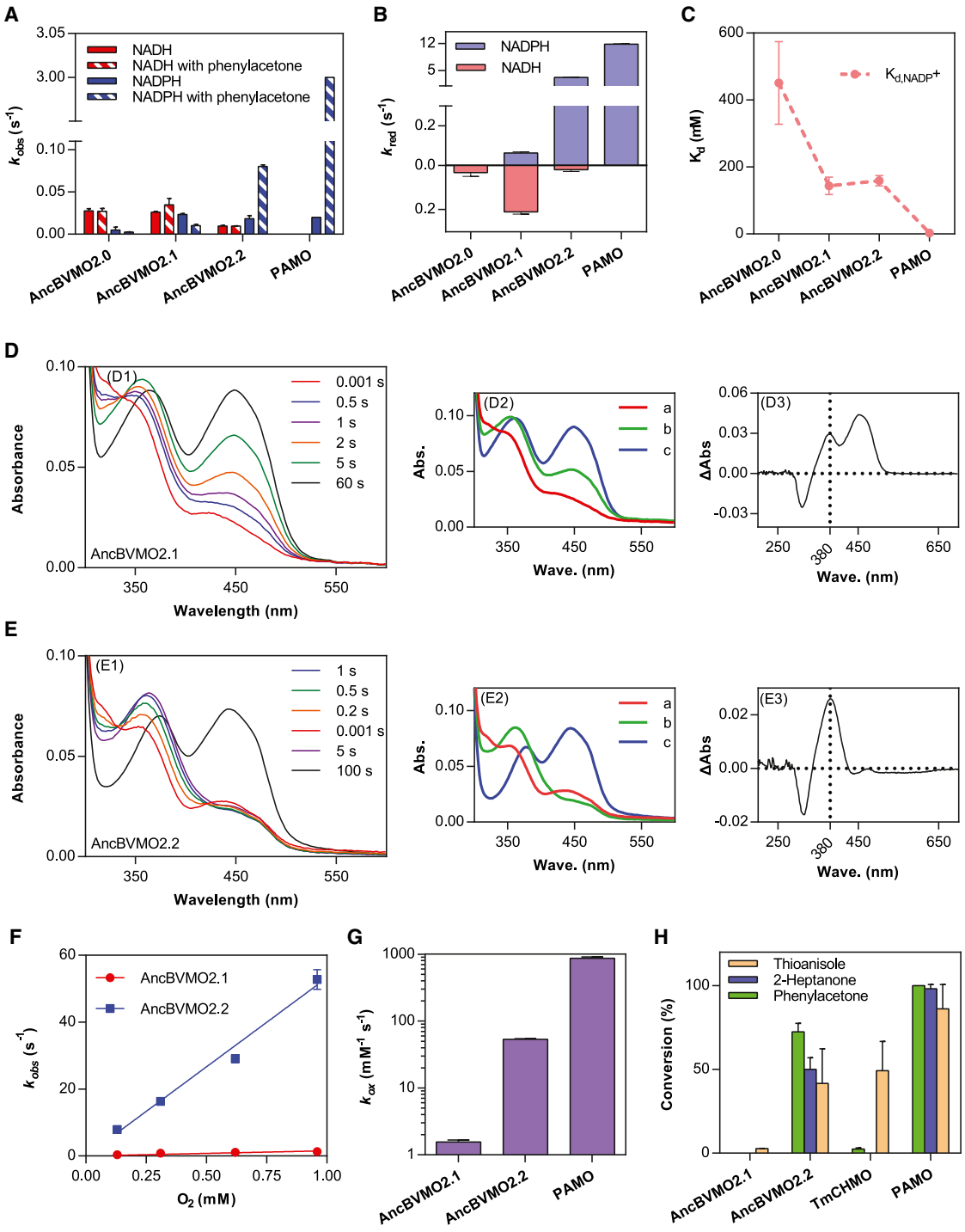


Figure 2. The hydride transfer process was established before the specificity for the coenzyme, and the reactivity against oxygen is contingent on this

(A) The uncoupling rates were determined in duplicate and are presented as mean ± SD. Values for PAMO are from Brondani et al.[43]

(B) Reduction rate constants (k_{red}) with NAD(P)H. The data were collected in triplicate and are presented as mean ± SD.

(C) The evolutionary trend of dissociation constants (K_d) of $NADP^+$.

(D and E) C4a-(hydro)peroxyflavin intermediate formation of AncBVMO2.1 (D) and AncBVMO2.2 (E). (D1) and (E1) show recorded spectra of reduced AncBVMOs after mixing with O_2 containing buffer. (D2) and (E2) show deconvoluted spectra fitted into a two-step process (a $\xrightarrow{k1}$ b $\xrightarrow{k2}$ c). (D3) and (E3) show absorbance change between species a and b, which indicates the formation of the intermediate.

(legend continued on next page)

found here. On average, clade I sequences are shorter (~490 amino acids) than the BVMOs from clades II and III. The latter two share a common ancestor (TBE = 0.99). Clade II (TBE = 0.95) includes most of the experimentally characterized BVMOs to date (Table S1; Figure S2), such as cyclohexanone monooxygenase (*Ac*CHMO) from *Acinetobacter* sp.,[33] PAMO,[11] STMO,[34] and the camphor monooxygenase (OTEMO) from *Pseudomonas putida*,[35] plus the fungal enzymes $BVMO_{Af1}$ from *Aspergillus fumigatus*[36] and PockeMO from *Thermothelomyces thermophila*.[37] Clade II sequences show the typical BVMO length of ~540 amino acids. Clade III BVMOs are found in a broader range of organisms, including other bacterial classes (like *Deinococcus* and *Chloroflexi*) and yeasts (*Ustilago maydis* and *Yarrowia lipolytica*[38]), as well as stramenophiles and arthropods. The sequence length among these enzymes varies greatly (500–600 aa).

Aiming to unveil the path for the emergence of the BVMO phenotype, we performed a reconstruction campaign following the major divergence points in the phylogeny (Figure 1B). Acknowledging the complex evolutionary history and the low posterior probabilities (PPs) of the reconstruction at the deep nodes (Figure S3), we approached the targeted nodes at three different stages to gain insight into their viability for experimental characterization. First, the common ancestor of the whole BVMO clade diverging from FMO-like sequences, AncBVMO0, was reconstructed (Figure 1B). This showed minimal expression in different constructs, and as the isolated protein did not contain any flavin cofactor, it was discarded for further studies. Next, the ancestors from each of the three diverging clades were targeted: AncBVMO1, AncBVMO2, and AncBVMO3. AncBVMO1 and AncBVMO3 showed good expression levels but were largely insoluble. To isolate them, the buffer had to be supplemented with a detergent, suggesting they might be membrane associated or require interaction with other biomolecules (something that has been speculated for some extant members of these clades, like EtaA).[30,31] In addition, these proteins easily aggregated just after the purification, which prevented further studies. In contrast, the ancestor of clade II, AncBVMO2 (from here on called AncBVMO2.0), showed good soluble expression and FAD retention after purification under standard conditions. Hence, we decided to closely follow the evolutionary trajectory of this clade and targeted the intermediate ancestors AncBVMO2.1 and AncBVMO2.2 (Figure 1C; Figure S3). They showed good expression levels (30–80 mg of purified protein per liter) (Figure 1D) and stability in solution over time. The binding of FAD was confirmed by the intense yellow color of the enzyme solutions and their ultraviolet-visible (UV-Vis) absorption spectra (Figure 1E). Also, they showed good thermal stabilities compared with extant BVMOs (Figure 1F). By size-exclusion chromatography, AncBVMO2.1 and AncBVMO2.2 were estimated to have a molecular mass of 60 kDa, suggesting that they are monomers in solution, like most extant BVMOs.[25,39,40] However, AncBVMO2.0 showed multiple oligomeric states with estimated molecular masses of 68, 135, 280, and 370 kDa, corresponding to mono-, di-, tetra-, and hexameric species.

To assess the robustness of the reconstruction, we reconstructed the alternative versions of all three studied ancestors: AltAncBVMO2.0, AltAncBVMO2.1, and AltAncBVMO2.2 (Figure S3A). These sequences include the second-best state (PP > 0.2) at all ambiguous sites (maximum *a posteriori* [*map*] state PP < 0.8).[41] These three AltAncBVMOs could also be expressed and purified and showed the same physicochemical features compared with their corresponding *map* ancestors (Figures S3B and S3C), confirming the robustness of the reconstruction.

AncBVMOs increase their oxyfunctionalization activity over time

To assess the phenotype of AncBVMOs, conversions were performed using mixtures of ketones or sulfides and NADPH as the electron donor (the preferred coenzyme of extant BVMOs) (Table S2). AncBVMO2.0 showed no detectable conversion of any of the 24 tested substrates, while AncBVMO2.1 was only active on sulfides reaching low conversions (10%–20%). In contrast, AncBVMO2.2 showed activity on several aliphatic and aromatic ketones (45%–100%) as well as some sulfides. To rule out that the lack of BV oxidation activity of AncBVMO2.0 and AncBVMO2.1 was due to a different cofactor specificity, the experiment was carried out with NADH as the hydride donor. While both ancestors were inactive, AncBVMO2.2 was able to perform some conversions, albeit with lower efficiency (8%–30% conversion) (Table S2). These results suggest that the trajectory of clade II started from a FAD-binding protein, virtually inactive. The heteroatom oxygenating activity emerged next, and the BV oxidation activity was developed in the further stage, while the sulfoxidation activity was enhanced. The AltAncBVMOs showed a similar activity profile, confirming a common phenotype with the AncBVMOs (Table S2).

The hydride transfer process precedes the establishment of the nicotinamide cofactor specificity

We explored the evolution of the nicotinamide specificity across the trajectory. As BVMOs can consume hydride donors without performing catalysis (*i.e.*, uncoupling activity), we initially analyzed the uncoupling rates using either NADPH or NADH (Figure 2A). All three AncBVMOs were found to display uncoupling activity with both coenzymes at typical rates for BVMOs (0.01–0.03 s^{-1}),[11,42] except for AncBVMO2.0, showing very slow uncoupling with NADPH (0.002–0.008 s^{-1}) (Table 1). Interestingly, there is a clear trend in the cofactor preference. AncBVMO2.0 was more effective in consuming NADH, AncBVMO2.1 was nearly indifferent, while AncBVMO2.2 showed the highest uncoupling rate with NADPH. To assess the relevance of these differences, measurements were performed in the presence of a canonical substrate, phenylacetone. AncBVMO2.2 showed a

(F) k_1 value vs. O_2 concentration (0.13, 0.31, 0.62, 0.96 mM after mixing). Measurements were done in duplicate or triplicate. The data are presented as mean ± SD.
(G) The oxygenation rate constants (k_{ox}) obtained from (F). The value for PAMO was obtained from Brondani et al.[43]
(H) Conversions. The reactions were performed in duplicate. The data are presented as mean ± SD.

Table 1. Kinetic parameters of AncBVMOs

	AncBVMO2.0	AncBVMO2.1	AncBVMO2.2	PAMO[a]
Pre-steady-state kinetics				
k_{red}, NADH (s^{-1})	0.032 ± 0.017[b]	0.21 ± 0.01	0.019 ± 0.008[b]	N/D
K_d, NADH (μM)	–	<15	<50	N/D
k_{red}, NADPH (s^{-1})	N/A	0.063 ± 0.005	3.2 ± 0.1	11.8 ± 0.1
K_d, NADPH (μM)	N/A	<3.5	<15	0.7 ± 0.3
K_d, $NADP^+$ (μM)	451 ± 124	144 ± 26	159 ± 16	3.3 ± 0.3
k_1 (a → b) (s^{-1})	–	0.28 ± 0.14	7.9 ± 0.3	113[c]
k_2 (b → c) (s^{-1})	–	0.27 ± 0.01	0.026 ± 0.003	N/D
k_{ox} (mM^{-1} s^{-1})	–	1.55 ± 0.12	53.3 ± 1.4	870 ± 35
Steady-state kinetics				
k_{unc}, NADH (s^{-1})	0.028 ± 0.002	0.026 ± 0.001	0.010 ± 0.001	N/D
k_{unc}, NADPH (s^{-1})	0.005 ± 0.002	0.023 ± 0.001	0.019 ± 0.003	0.014
k_{cat}, NADH (s^{-1})	0.027 ± 0.002[d]	0.026 ± 0.001[d]	0.010 ± 0.001[d]	0.39[e]
k_{cat}, NADPH (s^{-1})	0.003 ± 0.001[d]	0.025 ± 0.002[d]	0.080 ± 0.001[d]	–
k_{cat}, PA (s^{-1})	–	–	0.064 ± 0.002	3.1 ± 0.2
K_M, NADPH (μM)	–	–	2.4 ± 0.5	0.7 ± 0.1
K_M, PA (μM)	–	–	1.7 ± 0.3	80 ± 0.6

N/A, not active; N/D, not determined; PA, phenylacetone.
[a]Reported in Torres Pazmiño et al.[11]
[b]The values are the k_{obs} measured with 250 μM NADH.
[c]The value was calculated based on 0.13 mM oxygen and the k_{ox} of PAMO.
[d]The values were the k_{obs} measured with 100 μM NAD(P)H and 2 mM phenylacetone.
[e]Reported in Dudek et al.[44]

clearly enhanced NADPH consumption rate, while, as expected, AncBVMO2.0 and AncBVMO2.1 showed values in the same range as the uncoupling rates.

Pre-steady-state kinetic measurements (Table 1)[11] confirmed that AncBVMO2.2 is quite efficient in using NADPH as a hydride donor (k_{red} of 3.2 ± 0.1 s^{-1} and K_d < 15 μM); with NADH, the reduction rates were too low to measure reliably (<0.05 s^{-1}). AncBVMO2.1 seems to accept hydrides from both cofactors with high affinity (K_d < 15 μM), albeit at much slower rates. The observed rates for AncBVMO2.0 with NADH were too slow to be reliably fitted (<0.05 s^{-1}), while with NADPH the enzyme was not reduced. Therefore, the main explanation for AncBVMO2.0 inactivity is the virtually negligible reduction rate.

A plausible evolutionary path for nicotinamide cofactor usage would involve first the cofactor binding and hydride transfer regardless of the nature of the donor (AncBVMO2.0 → AncBVMO2.1) and then the reduction efficiency improved with a clear preference for NADPH (AncBVMO2.1 → AncBVMO2.2) (Figure 2B). However, there seems to be a conflict between the conversion results and the pre-steady-state kinetics: why can AncBVMO2.1 be reduced by NAD(P)H but still not perform oxygenations well? A hint might be given by the role of NADP(H) in inducing a conformational change of the $NADP^+$-flavoenzyme complex, which promotes the formation of the C4a-(hydro)peroxyflavin intermediate and stabilization thereof.[11,12,45] In fact, $NADP^+$ remains bound to the enzyme during the whole catalytic cycle and is only released in the last, rate-limiting, step.[12] As the actual nicotinamide cofactor release rate cannot be measured in a direct manner, the dissociation constants (K_d) for $NADP^+$ were determined (Figure 2C). The affinity for $NADP^+$ gradually increased from AncBVMO2.0 to extant BVMOs. However, for NAD^+, no binding was observed for all three AncBVMOs. This indicates that, although AncBVMO2.0 can be reduced by NADH, the NAD^+ does not remain bound during the catalytic circle.

The inactivity of AncBVMO2.0, and the poor activity of AncBVMO2.1, seemed puzzling. Therefore, we designed a set of experiments to find alternative hydride donors or acceptors inspired by remote homologs from the tDNBD superfamily.[5] As hydride donors, amines and glycerol were tested. Yet, none of these was reactive with AncBVMO2.0 or AncBVMO2.1. As electron acceptors, transition metals, muconic acid, and cysteamine were employed, as well as aromatic carbonyl compounds, in either aerobic or anaerobic condition, while providing NADH or NADPH. However, no positive hits could be detected (Table S3).

Finally, the redox potentials were investigated. We found that the E° values for the three ancestral flavoenzymes are similar: −169, −180, and −208 mV for AncBVMO2.0, AncBVMO2.1, and AncBVMO2.2, respectively. These are higher than the redox potential for FAD (−219 mV), meaning that the observed changes in reactivity are not related to the capacity for accepting electrons from any of the tested hydride donors.

The oxygen reactivity is contingent on the hydride donor specificity

To understand the underlying reason for the activity change between AncBVMO2.1 and AncBVMO2.2, we explored the reactivity of the reduced flavoenzymes with molecular oxygen.

The enzymes were first anaerobically reduced by NADPH (AncBVMO2.2) or NADH (AncBVMO2.1) and then mixed with buffer containing O_2. The observed spectral changes, reporting on the process in which the reduced flavin cofactor reacts with dioxygen, could be fitted into a two-step process (a → b → c) (Figures 2D and 2E) for both ancestors. The first step represents formation of the C4a-(hydro)peroxyflavin with a specific absorbance peak at 380 nm[12,20] (Figures 2D2 and 2E2: a → b). The observed rates (k_1) indicated a 30 times faster intermediate formation for AncBVMO2.2 vs. AncBVMO2.1 (Table 1). The second step involves the decay of the intermediate to the oxidized FAD, resulting in an increase in the absorbance at 450 nm (Figures 2D2 and 2E2: b → c). The observed rate (k_2) of AncBVMO2.1 was 10 times faster than that of AncBVMO2.2, indicating that the intermediate formed by AncBVMO2.1 is not very stable (Figures 2D3 and 2E3). This explains the lack of BV oxidation activity of this AncBVMO and its poor efficiency in sulfoxidations. Further, the rates at which the reduced protein-bound flavins react with dioxygen (k_{ox}) were determined. By plotting the k_1 values over different O_2 concentrations (Figure 2F), the rate constants were obtained. The k_{ox} values show a clear tendency of increasing oxygen reactivity from the AncBVMOs toward extant ones (Table 1; Figure 2G), indicating that they improved their ability to use molecular oxygen as substrate over time.

Finally, a detailed kinetic analysis (Table 1) and conversions were performed using AncBVMO2.2 and the two well-studied descendants, PAMO and *Tm*CHMO (Figures 2H; Table S2). AncBVMO2.2's preferred substrate is phenylacetone, yet it is 50 times slower than PAMO[11] (Table 1). Also, it accepts a wide range of sulfides with moderate/good enantioselectivities (Table S4). A striking observation is that activity on cyclic ketones is absent, which is in contrast with the substrate profile of *Tm*CHMO. According to AlphaFold models, a particular loop present at the entrance of the substrate tunnel in PAMO (441–444 residues), which is missing in *Tm*CHMO, is also present in the AncBVMOs (Figure S4A). Other CHMOs found in clade II that also diverged from AncBVMO2.1, like *Ac*CHMO from *Acinetobacter* sp., lack this bulge as well. It has been demonstrated that engineering this loop affects the substrate scope of PAMO, converting it into a BVMO with a substrate profile like that of CHMOs.[46–48]

BVMO functionality is determined by an entangled network of residues and conformational changes

Eight residues are essential for BVMO activity[27,48]: two from the active site (D66 and R337, PAMO numbering),[11,49] four from the NADPH binding pocket (S196,[50] R217,[44] H(Q)220,[44] and W501[40,51]), and two involved in tuning the nicotinamide oxidation, associated with the uncoupling (T218[50] and K336[44]). Overall, AncBVMOs showed very strong conservation of these residues, displaying high probabilities on their reconstruction (Figure 3A). Consequently, the observed differences among them cannot be explained by substitutions directly involved in the catalytic mechanism. Therefore, based on the AlphaFold models (Figures 3B and 3C), we explored the historical sequence space in the vicinity of the active site and close to the key catalytic elements by introducing single mutations in AncBVMO2.0 and AncBVMO2.1 based on the AncBVMO2.2 sequence and performed conversions and kinetic measurements with purified enzymes (Figure 3C). However, none of the tested mutants displayed BV activity (Figure S4B), suggesting that the emergence of the BV oxidation itself must be due to a more complex process requiring remote substitutions, synergistic effects, or even conformational changes.

The overall structural changes along the AncBVMOs' evolutionary trajectory were explored by molecular dynamics (MD). AlphaFold models were docked with FAD and both possible hydride donors, NADPH and NADH, and MD simulations (0.6 μs) were run. No significant differences were observed for the distance between the C4 hydrogen of the NADPH/NADH to the N5 of the FAD cofactor, suggesting that more distant conformational changes may facilitate substrate binding and BV oxidation. The simulations showed that enzyme mobility gradually decreases from the eldest AncBVMO toward the extant BVMOs (Figures 3D and S5). AncBVMO2.0 showed the most fluctuating structure, indicating a poor interaction with both ligands, and the regions with the highest conformational flexibility were those forming the FAD, NAD, and substrate binding pockets (Figures 3D and S5B). AncBVMO2.1 seems to gain in rigidity upon binding of NADH, while AncBVMO2.2 resembles the root-mean-square fluctuation (RMSF) profile of an extant BVMO upon binding with NADPH. These results, though not conclusive, suggest that the dynamics of the three-dimensional conformation of the enzyme contribute to the emergence of the BV oxidation activity. This is likely because the active-site cavity acquires and maintains the appropriate geometry to accommodate formation of the Criegee intermediate when in complex with $NADP^+$.

DISCUSSION

Here, we have explored how the complex catalytic mechanism of BVMOs emerged and matured over time. Based on our investigation, the evolutionary trajectory of the clade II BVMOs involves, at early stages, a FAD-binding ancestor (AncBVMO2.0) that could be reduced by NADH inefficiently, suggesting a marginal NADH oxidase activity. The next clear step in the trajectory was the recruitment of a nucleotide hydride donor, transforming the marginally active ancestor into a poor NADH oxidase, as AncBVMO2.1 could be reduced at low rates by NADH and NADPH. Thus, in principle, the chemical nature of the donor would be less important than the actual hydride transfer process. We evidenced in the transition from AncBVMO2.0 to AncBVMO2.1, the disappearance of the tendency to form oligomers with the acquisition of increasing conformational stability. The last step toward a fully functional BVMO (AncBVMO2.2) was the recruitment of O_2 as an electron acceptor. AncBVMO2.1 has no well-established preference for NADPH as a hydride donor; its reactivity with O_2 is very poor, and the reactive intermediate species formed cannot survive long enough in the active site. In part, this might be due to the low affinity for $NADP^+$, which affects the formation of the catalytic network at the active site.[11,12,45,52] In contrast, in AncBVMO2.2, the O_2 usage improves concomitant with the NADPH preference established (or vice versa). Interestingly, although AncBVMO2.2 seems to show an overall low catalytic efficiency, the functional

A

Role	PAMO		TmCHMO		AncBVMO2.2		AncBVMO2.1		AncBVMO2.0		Ref.
	Site	State	Site	State	Site	State	Site	State	Site	State	
Overall activity	66	**D**	59	**D**	60	**D** (1)	60	**D** (1)	60	**D** (1)	11,47
	337	**R**	329	**R**	328	**R** (1)	328	**R** (1)	328	**R** (1)	11,47
NAD(P)H binding	196	**S**	188	**S**	189	**S** (1)	189	**S** (0.99)	189	**A** (0.84)	48
	217	**R**	209	**R**	210	**R** (1)	210	**R** (1)	210	**R** (1)	49
	220	**H**	212	**Q**	213	**N** (1)	213	**N** (1)	213	**S** (0.66)[a]	49
	501	**W**	492	**W**	492	**W** (1)	492	**W** (1)	492	**W** (1)	40,50
NAD(P)H oxidation	218	**T**	210	**T**	211	**T** (1)	211	**T** (1)	211	**T** (1)	48
	336	**K**	328	**K**	327	**K** (1)	327	**K** (1)	327	**K** (1)	49

For AncBVMOs, the PP of the reconstruction of each site is given in parentheses next to the state.
[a]Alt state N (0.21)

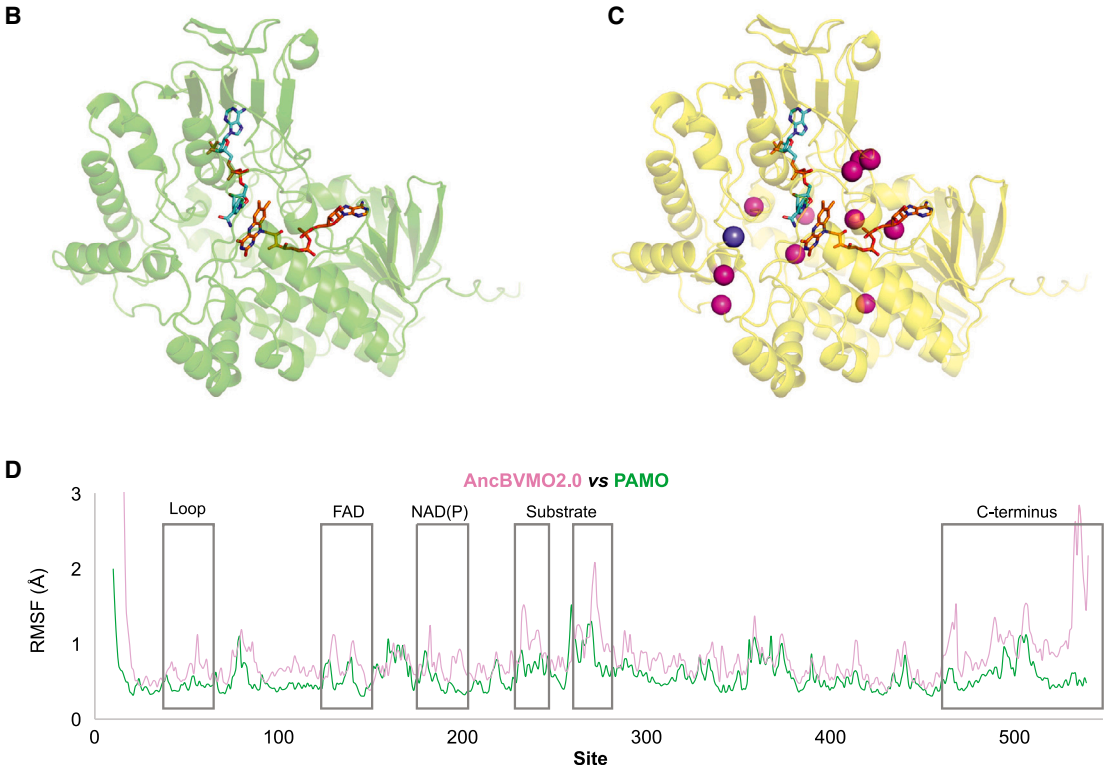


Figure 3. BVMO functionality is determined by an entangled network of residues and conformational changes

(A) State of key residues for BVMO functionality in AncBVMOs.

(B and C) AlphaFold models of AncBVMO2.2 (B) and AncBVMO2.1 (C). FAD (orange sticks) and NAD(P)H (cyan sticks) were docked. The purple and magenta spheres in (C) represent the locations of the mutations in AncBVMO2.1 and AncBVMO2.0, respectively.

(D) RMSF profiles of PAMO with NADPH (green) vs. AncBVMO2.0 with NADH (pink). The most flexible regions are marked with gray squares.

phenotype is already the same as that of extant BVMOs. This suggests that further fine-tuning of the established gears allowed catalysis to proceed in more efficient ways (mostly accelerating the overall turnover).

A conundrum rises with the identified evolutionary trajectory and the phylogeny of the BVMOs. Did BV functionality evolve more than once (at least three times) in the BVMO history? Although BV functionality can emerge by parallel mechanisms—as demonstrated by the FMOs displaying BV activity—it seems non-parsimonious for the BVMO clade. A simple explanation could be that the observed activity is an artifact due to uncertainty in the reconstruction. However, we reconstructed the AltAncs, and these showed a consistent phenotype with the *map* Ancs, ruling out this possibility. Therefore, it is possible that ancestral BV activity is displayed on unknown substrates, or in different microenvironments, that we are unable to detect. However, this directly points to the second issue; would it be possible that O_2 reactivity was acquired/improved independently by modern lineages of monooxygenases? Considering that the remote homologs from the tDNBD superfamily[5] show other acceptors than molecular oxygen, it is plausible that AncBVMOs showed low O_2 affinity—likely inherent to the reactivity of the reduced flavin—and this improved over time as a consequence of the higher O_2 availability. Unfortunately,

AncBVMO1 and AncBVMO3 could not be characterized to capture the overall mechanistic dynamics of the family.

In summary, we have identified a trajectory for a catalytic mechanism to emerge consisting in the individual recruitment of the catalytic elements followed by steps of joint optimization of their specific roles in an entangled fashion, simultaneously occurring with—or as a consequence of—the acquisition of conformational stability in the structural fold, resulting in fully functional enzymes prone to posterior fine-tuning.

Limitations of the study

The phenotype measurements of AncBVMOs were performed *in vitro*, under laboratory conditions, which do not resemble the cellular environment that the ancestors likely experienced, constituting a source of bias/uncertainty. Moreover, the expression was done in *E. coli*, under aerobic conditions, and thus the folding and overall conformational dynamics of the ancestors might be affected. Indeed, we could not obtain AncBVMO1 and AncBVMO3.

Unfortunately, crystallization trials failed for the three AncBVMOs, and this prevented us from understanding if the observed phenotypes are due to structural changes. Although MD simulations were performed, those results should be considered only as indicative, as the starting structures were AlphaFold models.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Molecular phylogeny and ancestral sequence reconstruction
  - Cloning, expression and purification
  - Substrates screening and conversions
  - Size exclusion chromatography
  - Steady-state kinetics
  - $NAD(P)^+$ titration
  - Pre-steady-state kinetics
  - Redox potential measurement
  - Site-directed mutagenesis
  - Structural modeling
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Quantification
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114130.

ACKNOWLEDGMENTS

M.L.M. would like to acknowledge the Thornton lab's members at U Chicago (2018 cohort) for the scientific discussions that inspired this work. G.Y. was financially supported by a CSC scholarship. This work was funded by the European Union's Horizon 2020 Research and Innovation program under grant agreement 847675 COFUND project oLife.

AUTHOR CONTRIBUTIONS

All listed authors performed experiments and/or analyzed data. M.L.M. and M.W.F. proposed the project. M.L.M. performed molecular phylogenetics and ASR. G.Y. designed the experiments under the guidance of M.L.M. O.P. made the AncBVMO2.1 mutant and measured its activities, reduction/oxidation rates, and alternative activities under the supervision of G.Y. H.J.W. generated the AlphaFold models and performed docking and molecular dynamic simulations. H.J.R. performed size-exclusion chromatography. G.d.G. performed the conversions of sulfides. G.Y. performed all other experiments. G.Y., M.L.M., and M.W.F. analyzed the data. G.Y. and M.L.M. wrote and edited the manuscript and prepared the figures. M.W.F. and M.L.M. revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 3, 2023
Revised: February 15, 2024
Accepted: April 4, 2024
Published: April 18, 2024

REFERENCES

1. van Berkel, W.J.H., Kamerbeek, N.M., and Fraaije, M.W. (2006). Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. J. Biotechnol. *124*, 670–689. https://doi.org/10.1016/j.jbiotec.2006.03.044.
2. Fürst, M.J.L.J., Gran-Scheuch, A., Aalbers, F.S., and Fraaije, M.W. (2019). Baeyer-Villiger monooxygenases: Tunable oxidative biocatalysts. ACS Catal. *9*, 11207–11241. https://doi.org/10.1021/acscatal.9b03396.
3. Cashman, J.R., and Zhang, J. (2006). Human flavin-containing monooxygenases. Annu. Rev. Pharmacol. Toxicol. *46*, 65–100. https://doi.org/10.1146/annurev.pharmtox.46.120604.141043.
4. Olucha, J., and Lamb, A.L. (2011). Mechanistic and structural studies of the N-hydroxylating flavoprotein monooxygenases. Bioorg. Chem. *39*, 171–177. https://doi.org/10.1016/j.bioorg.2011.07.006.
5. Ojha, S., Meng, E.C., and Babbitt, P.C. (2007). Evolution of function in the "two dinucleotide binding domains" flavoproteins. PLoS Comput. Biol. *3*, e121–e1280. https://doi.org/10.1371/journal.pcbi.0030121.
6. Ballou, D.P., and Entsch, B. (2013). The reaction mechanisms of groups A and B flavoprotein monooxygenases. In Handbook of Flavoproteins: Complex Flavoproteins, Dehydrogenases and Physical Methods (Walter de Gruyter GmbH), pp. 1–28. https://doi.org/10.1515/9783110298345.1.
7. Cashman, J.R. (2005). Some distinctions between flavin-containing and cytochrome P450 monooxygenases. Biochem. Biophys. Res. Commun. *338*, 599–604. https://doi.org/10.1016/j.bbrc.2005.08.009.
8. Ziegler, D.M. (1988). Flavin-containing monooxygenases: Catalytic mechanism and substrate specificities. Drug Metab. Rev. *19*, 1–32. https://doi.org/10.3109/03602538809049617.
9. Badieyan, S., Bach, R.D., and Sobrado, P. (2015). Mechanism of N -hydroxylation catalyzed by flavin-dependent monooxygenases. J. Org. Chem. *80*, 2139–2147. https://doi.org/10.1021/jo502651v.
10. Walsh, C.T., and Chen, Y.-C.J. (1988). Enzymic Baeyer–Villiger oxidations by flavin-dependent monooxygenases. Angew. Chem., Int. Ed. Engl. *27*, 333–343. https://doi.org/10.1002/anie.198803331.
11. Torres Pazmiño, D.E., Baas, B.J., Janssen, D.B., and Fraaije, M.W. (2008). Kinetic mechanism of phenylacetone monooxygenase from *Thermobifida fusca*. Biochemistry *47*, 4082–4093. https://doi.org/10.1021/bi702296k.
12. Sheng, D., Ballou, D.P., and Massey, V. (2001). Mechanistic studies of cyclohexanone monooxygenase: Chemical properties of intermediates involved in catalysis. Biochemistry *40*, 11156–11167. https://doi.org/10.1021/bi011153h.

13. Clifton, B.E., Kaczmarski, J.A., Carr, P.D., Gerth, M.L., Tokuriki, N., and Jackson, C.J. (2018). Evolution of cyclohexadienyl dehydratase from an ancestral solute-binding protein article. Nat. Chem. Biol. *14*, 542–547. https://doi.org/10.1038/s41589-018-0043-2.

14. Kaltenbach, M., Burke, J.R., Dindo, M., Pabis, A., Munsberg, F.S., Rabin, A., Kamerlin, S.C.L., Noel, J.P., and Tawfik, D.S. (2018). Evolution of chalcone isomerase from a noncatalytic ancestor article. Nat. Chem. Biol. *14*, 548–555. https://doi.org/10.1038/s41589-018-0042-3.

15. Bieger, B., and Essen, L.-O. (2001). Crystal structure of the catalytic core component of the alkylhydroperoxide reductase AhpF from *Escherichia coli*. J. Mol. Biol. *307*, 1–8. https://doi.org/10.1006/jmbi.2001.4441.

16. Karplus, P.A., and Schulz, G.E. (1987). Refined structure of glutathione reductase at 1.54 Å resolution. J. Mol. Biol. *195*, 701–729. https://doi.org/10.1016/0022-2836(87)90191-4.

17. Sousa, F.M., Sena, F.V., Batista, A.P., Athayde, D., Brito, J.A., Archer, M., Oliveira, A.S.F., Soares, C.M., Catarino, T., and Pereira, M.M. (2017). The key role of glutamate 172 in the mechanism of type II NADH:quinone oxidoreductase of *Staphylococcus aureus*. Biochim. Biophys. Acta Bioenerg. *1858*, 823–832. https://doi.org/10.1016/j.bbabio.2017.08.002.

18. Tinikul, R., Lawan, N., Akeratchatapan, N., Pimviriyakul, P., Chinantuya, W., Suadee, C., Sucharitakul, J., Chenprakhon, P., Ballou, D.P., Entsch, B., and Chaiyen, P. (2021). Protonation status and control mechanism of flavin–oxygen intermediates in the reaction of bacterial luciferase. FEBS J. *288*, 3246–3260. https://doi.org/10.1111/febs.15653.

19. Huijbers, M.M.E., Montersino, S., Westphal, A.H., Tischler, D., and Van Berkel, W.J.H. (2014). Flavin dependent monooxygenases. Arch. Biochem. Biophys. *544*, 2–17. https://doi.org/10.1016/j.abb.2013.12.005.

20. Bailleul, G., Yang, G., Nicoll, C.R., Mattevi, A., Fraaije, M.W., and Mascotti, M.L. (2023). Evolution of enzyme functionality in the flavin-containing monooxygenases. Nat. Commun. *14*, 1042. https://doi.org/10.1038/s41467-023-36756-x.

21. Mashiguchi, K., Tanaka, K., Sakai, T., Sugawara, S., Kawaide, H., Natsume, M., Hanada, A., Yaeno, T., Shirasu, K., Yao, H., et al. (2011). The main auxin biosynthesis pathway in Arabidopsis. Proc. Natl. Acad. Sci. USA *108*, 18512–18517. https://doi.org/10.1073/pnas.1108434108.

22. Dai, X., Mashiguchi, K., Chen, Q., Kasahara, H., Kamiya, Y., Ojha, S., DuBois, J., Ballou, D., and Zhao, Y. (2013). The biochemical mechanism of auxin biosynthesis by an arabidopsis YUCCA flavin-containing monooxygenase. J. Biol. Chem. *288*, 1448–1457. https://doi.org/10.1074/jbc.M112.424077.

23. Gonzalo, G.d., Pazmiño, D.E.T., Ottolina, G., Fraaije, M.W., and Carrea, G. (2005). Oxidations catalyzed by phenylacetone monooxygenase from *Thermobifida fusca*. Tetrahedron Asymmetry *16*, 3077–3083. https://doi.org/10.1016/j.tetasy.2005.08.004.

24. Dudek, H.M., de Gonzalo, G., Pazmiño, D.E.T., Stepniak, P., Wyrwicz, L.S., Rychlewski, L., and Fraaije, M.W. (2011). Mapping the substrate binding site of phenylacetone monooxygenase from *Thermobifida fusca* by mutational analysis. Appl. Environ. Microbiol. *77*, 5730–5738. https://doi.org/10.1128/AEM.00687-11.

25. Fraaije, M.W., Wu, J., Heuts, D.P.H.M., Van Hellemond, E.W., Spelberg, J.H.L., and Janssen, D.B. (2005). Discovery of a thermostable Baeyer-Villiger monooxygenase by genome mining. Appl. Microbiol. Biotechnol. *66*, 393–400. https://doi.org/10.1007/s00253-004-1749-5.

26. Pazmiño, D.T., Snajdrova, R., Rial, D., Mihovilovic, M., and Fraaije, M. (2007). Altering the substrate specificity and enantioselectivity of phenylacetone monooxygenase by structure-inspired enzyme redesign. Adv. Synth. Catal. *349*, 1361–1368. https://doi.org/10.1002/adsc.200700045.

27. Romero, E., Castellanos, J.R.G., Mattevi, A., and Fraaije, M.W. (2016). Characterization and crystal structure of a robust cyclohexanone monooxygenase. Angew. Chem. *128*, 16084–16087. https://doi.org/10.1002/ange.201608951.

28. Morii, S., Sawamoto, S., Yamauchi, Y., Miyamoto, M., Iwami, M., and Itagaki, E. (1999). Steroid monooxygenase of *Rhodococcus rhodochrous*: Sequencing of the genomic DNA, and hyperexpression, purification, and characterization of the recombinant enzyme. J. Biochem. *126*, 624–631. https://doi.org/10.1093/oxfordjournals.jbchem.a022494.

29. Fraaije, M.W., Kamerbeek, N.M., van Berkel, W.J.H., and Janssen, D.B. (2002). Identification of a Baeyer-Villiger monooxygenase sequence motif. FEBS Lett. *518*, 43–47. https://doi.org/10.1016/S0014-5793(02)02623-6.

30. Fraaije, M.W., Kamerbeek, N.M., Heidekamp, A.J., Fortin, R., and Janssen, D.B. (2004). The prodrug activator EtaA from *Mycobacterium tuberculosis* is a Baeyer-Villiger monooxygenase. J. Biol. Chem. *279*, 3354–3360. https://doi.org/10.1074/jbc.M307770200.

31. Vannelli, T.A., Dykman, A., and Ortiz De Montellano, P.R. (2002). The antituberculosis drug ethionamide is activated by a flavoprotein monooxygenase. J. Biol. Chem. *277*, 12824–12829. https://doi.org/10.1074/jbc.M110751200.

32. Throne-Holst, M., Wentzel, A., Ellingsen, T.E., Kotlar, H.K., and Zotchev, S.B. (2007). Identification of novel genes involved in long-chain n-alkane degradation by *Acinetobacter* sp. strain DSM 17874. Appl. Environ. Microbiol. *73*, 3327–3332. https://doi.org/10.1128/AEM.00064-07.

33. Chen, Y.C., Peoples, O.P., and Walsh, C.T. (1988). *Acinetobacter* cyclohexanone monooxygenase: Gene cloning and sequence determination. J. Bacteriol. *170*, 781–789. https://doi.org/10.1128/jb.170.2.781-789.1988.

34. Franceschini, S., Van Beek, H.L., Pennetta, A., Martinoli, C., Fraaije, M.W., and Mattevi, A. (2012). Exploring the structural basis of substrate preferences in Baeyer-Villiger monooxygenases: Insight from steroid monooxygenase. J. Biol. Chem. *287*, 22626–22634. https://doi.org/10.1074/jbc.M112.372177.

35. Leisch, H., Shi, R., Grosse, S., Morley, K., Bergeron, H., Cygler, M., Iwaki, H., Hasegawa, Y., and Lau, P.C.K. (2012). Cloning, Baeyer-Villiger biooxidations, and structures of the camphor pathway 2-oxo-Δ3-4,5,5-trimethylcyclopentenylacetyl-coenzyme A monooxygenase of *Pseudomonas putida* ATCC 17453. Appl. Environ. Microbiol. *78*, 2200–2212. https://doi.org/10.1128/AEM.07694-11.

36. Mascotti, M.L., Kurina-Sanz, M., Juri Ayub, M., and Fraaije, M.W. (2014). Insights in the kinetic mechanism of the eukaryotic Baeyer-Villiger monooxygenase $BVMO_{Af1}$ from *Aspergillus fumigatus* Af293. Biochimie *107*, 270–276. https://doi.org/10.1016/j.biochi.2014.09.005.

37. Fürst, M.J.L.J., Savino, S., Dudek, H.M., Gómez Castellanos, J.R., Gutiérrez de Souza, C., Rovida, S., Fraaije, M.W., and Mattevi, A. (2017). Polycyclic ketone monooxygenase from the thermophilic fungus *Thermothelomyces thermophila*: A structurally distinct biocatalyst for bulky substrates. J. Am. Chem. Soc. *139*, 627–630. https://doi.org/10.1021/jacs.6b12246.

38. Bordewick, S., Beier, A., Balke, K., and Bornscheuer, U.T. (2018). Baeyer-Villiger monooxygenases from *Yarrowia lipolytica* catalyze preferentially sulfoxidations. Enzym. Microb. Technol. *109*, 31–42. https://doi.org/10.1016/j.enzmictec.2017.09.008.

39. Messiha, H.L., Ahmed, S.T., Karuppiah, V., Suardíaz, R., Ascue Avalos, G.A., Fey, N., Yeates, S., Toogood, H.S., Mulholland, A.J., and Scrutton, N.S. (2018). Biocatalytic routes to lactone monomers for polymer production. Biochemistry *57*, 1997–2008. https://doi.org/10.1021/acs.biochem.8b00169.

40. Mirza, I.A., Yachnin, B.J., Wang, S., Grosse, S., Bergeron, H., Imura, A., Iwaki, H., Hasegawa, Y., Lau, P.C.K., and Berghuis, A.M. (2009). Crystal structures of cyclohexanone monooxygenase reveal complex domain movements and a sliding cofactor. J. Am. Chem. Soc. *131*, 8848–8854. https://doi.org/10.1021/ja9010578.

41. Mascotti, M.L. (2022). Resurrecting enzymes by ancestral sequence reconstruction. In Enzyme Engineering: Methods and Protocols, pp. 111–136. https://doi.org/10.1007/978-1-0716-1826-4_7.

42. Li, G., Fürst, M.J.L.J., Mansouri, H.R., Ressmann, A.K., Ilie, A., Rudroff, F., Mihovilovic, M.D., Fraaije, M.W., and Reetz, M.T. (2017). Manipulating the stereoselectivity of the thermostable Baeyer-Villiger monooxygenase

TmCHMO by directed evolution. Org. Biomol. Chem. *15*, 9824–9829. https://doi.org/10.1039/c7ob02692g.

43. Brondani, P.B., Dudek, H.M., Martinoli, C., Mattevi, A., and Fraaije, M.W. (2014). Finding the switch: Turning a Baeyer-Villiger monooxygenase into a NADPH Oxidase. J. Am. Chem. Soc. *136*, 16966–16969. https://doi.org/10.1021/ja508265b.
44. Dudek, H.M., Torres Pazmiño, D.E., Rodríguez, C., De Gonzalo, G., Gotor, V., and Fraaije, M.W. (2010). Investigating the coenzyme specificity of phenylacetone monooxygenase from *Thermobifida fusca*. Appl. Microbiol. Biotechnol. *88*, 1135–1143. https://doi.org/10.1007/s00253-010-2769-y.
45. Ryerson, C.C., Ballou, D.P., and Walsh, C. (1982). Mechanistic studies on cyclohexanone oxygenase. Biochemistry *21*, 2644–2655. https://doi.org/10.1021/bi00540a011.
46. Yang, G., Cang, R., Shen, L.Q., Xue, F., Huang, H., and Zhang, Z.G. (2019). Expanding the substrate scope of phenylacetone monooxygenase from *Thermobifida fusca* towards cyclohexanone by protein engineering. Catal. Commun. *119*, 159–163. https://doi.org/10.1016/j.catcom.2018.10.022.
47. Reetz, M.T., and Wu, S. (2008). Greatly reduced amino acid alphabets in directed evolution: Making the right choice for saturation mutagenesis at homologous enzyme positions. Chem. Commun., 5499–5501. https://doi.org/10.1039/b813388c.
48. Balke, K., Beier, A., and Bornscheuer, U.T. (2018). Hot spots for the protein engineering of Baeyer-Villiger monooxygenases. Biotechnol. Adv. *36*, 247–263. https://doi.org/10.1016/j.biotechadv.2017.11.007.
49. Malito, E., Alfieri, A., Fraaije, M.W., and Mattevi, A. (2004). Crystal structure of a Baeyer-Villiger monooxygenase. Proc. Natl. Acad. Sci. USA *101*, 13157–13162. https://doi.org/10.1073/pnas.040453810.
50. Beier, A., Bordewick, S., Genz, M., Schmidt, S., van den Bergh, T., Peters, C., Joosten, H.J., and Bornscheuer, U.T. (2016). Switch in cofactor specificity of a Baeyer–Villiger monooxygenase. Chembiochem *17*, 2312–2315. https://doi.org/10.1002/cbic.201600484.
51. Fürst, M.J., Fiorentini, F., and Fraaije, M.W. (2019). Beyond active site residues: overall structural dynamics control catalysis in flavin-containing and heme-containing monooxygenases. Curr. Opin. Struct. Biol. *59*, 29–37. https://doi.org/10.1016/j.sbi.2019.01.019.
52. Polyak, I., Reetz, M.T., and Thiel, W. (2012). Quantum mechanical/molecular mechanical study on the mechanism of the enzymatic Baeyer-Villiger reaction. J. Am. Chem. Soc. *134*, 2732–2741. https://doi.org/10.1021/ja2103839.
53. Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010). CD-HIT Suite: A web server for clustering and comparing biological sequences. Bioinformatics *26*, 680–682. https://doi.org/10.1093/bioinformatics/btq003.
54. Katoh, K., Rozewicki, J., and Yamada, K.D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Briefings Bioinf. *20*, 1160–1166. https://doi.org/10.1093/bib/bbx108.
55. Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
56. Lemoine, F., Domelevo Entfellner, J.B., Wilkinson, E., Correia, D., Dávila Felipe, M., De Oliveira, T., and Gascuel, O. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. Nature *556*, 452–456. https://doi.org/10.1038/s41586-018-0043-0.
57. Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*, 1586–1591. https://doi.org/10.1093/molbev/msm088.
58. Kumar, S., Stecher, G., Suleski, M., and Hedges, S.B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. Mol. Biol. Evol. *34*, 1812–1819. https://doi.org/10.1093/MOLBEV/MSX116.
59. Kumar, S., Suleski, M., Craig, J.M., Kasprowicz, A.E., Sanderford, M., Li, M., Stecher, G., and Hedges, S.B. (2022). TimeTree 5: An expanded resource for species divergence times. Mol. Biol. Evol. *39*, msac174. https://doi.org/10.1093/molbev/msac174.
60. Riebel, A., De Gonzalo, G., and Fraaije, M.W. (2013). Expanding the biocatalytic toolbox of flavoprotein monooxygenases from *Rhodococcus jostii* RHA1. J. Mol. Catal. B Enzym. *88*, 20–25. https://doi.org/10.1016/j.molcatb.2012.11.009.
61. Eick, G.N., Bridgham, J.T., Anderson, D.P., Harms, M.J., and Thornton, J.W. (2017). Robustness of reconstructed ancestral protein functions to statistical uncertainty. Mol. Biol. Evol. *34*, 247–261. https://doi.org/10.1093/molbev/msw223.
62. Ueoka, R., Meoded, R.A., Gran-Scheuch, A., Bhushan, A., Fraaije, M.W., and Piel, J. (2020). Genome mining of oxidation modules in trans-acyltransferase polyketide synthases reveals a culturable source for lobatamides. Angew. Chem., Int. Ed. Engl. *59*, 7761–7765. https://doi.org/10.1002/anie.201916005.
63. Gran-Scheuch, A., Trajkovic, M., Parra, L., and Fraaije, M.W. (2018). Mining the genome of *Streptomyces leeuwenhoekii*: Two new type I Baeyer–Villiger monooxygenases from Atacamam Desert. Front. Microbiol. *9*, 1609. https://doi.org/10.3389/fmicb.2018.01609.
64. Maklashina, E., and Cecchini, G. (2020). Determination of flavin potential in proteins by xanthine/xanthine oxidase method. Bio. Protoc. *10*, e3571. https://doi.org/10.21769/bioprotoc.3571.
65. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.
66. Orru, R., Dudek, H.M., Martinoli, C., Torres Pazmiño, D.E., Royant, A., Weik, M., Fraaije, M.W., and Mattevi, A. (2011). Snapshots of enzymatic Baeyer-Villiger catalysis: Oxygen activation and intermediate stabilization. J. Biol. Chem. *286*, 29284–29291. https://doi.org/10.1074/jbc.M111.255075.
67. Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. Proteins: Structure. Proteins *77*, 114–122. https://doi.org/10.1002/prot.22570.
68. Krieger, E., and Vriend, G. (2015). New ways to boost molecular dynamics simulations. J. Comput. Chem. *36*, 996–1007. https://doi.org/10.1002/jcc.23899.
69. Krieger, E., Darden, T., Nabuurs, S.B., Finkelstein, A., and Vriend, G. (2004). Making optimal use of empirical energy functions: Force-field parameterization in crystal space. Proteins: Structure. Proteins *57*, 678–683. https://doi.org/10.1002/prot.20251.
70. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., and Pedersen, L.G. (1995). A smooth particle mesh Ewald method. J. Chem. Phys. *103*, 8577–8593. https://doi.org/10.1063/1.470117.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
AncBVMO2.0	This paper	N/A
AncBVMO2.1	This paper	N/A
AncBVMO2.2	This paper	N/A
Alt_AncBVMO2.0	This paper	N/A
Alt_AncBVMO2.1	This paper	N/A
Alt_AncBVMO2.2	This paper	N/A
AncBVMO0	This paper	N/A
AncBVMO1	This paper	N/A
AncBVMO2	This paper	N/A
AncBVMO3	This paper	N/A
Xanthine oxidase	Sigma-Aldrich	X1875-50UN
Phosphite dehydrogenase (PTDH)	Fraaije Lab	N/A
Oligonucleotides		
Primer: A2.1_ XX _F, see Table S5	This paper	N/A
Primer: A2.1_ XX _R, see Table S5	This paper	N/A
Primer: A2.0_ R218A _F, see Table S5	This paper	N/A
Primer: A2.0_ R218A _R, see Table S5	This paper	N/A
Recombinant DNA		
pBAD plasmid with BsaI restriction sites and a N-terminal 6×His-SUMO tag	Fraaije Lab	N/A
Recombinant plasmids (of AncBVMOs)	This paper	N/A
Deposited data		
AncBVMO2.0_bacteria	This paper	OR264814
AncBVMO2.1_bacteria	This paper	OR264815
AncBVMO2.2_bacteria	This paper	OR264816
Software and algorithms		
CD-HIT	Li et al.[53]	http://weizhongli-lab.org/cd-hit/
MAFFT v7	Katoh et al.[54]	https://mafft.cbrc.jp/alignment/server/index.html
RAxML v8.2.10	Stamatakis et al.[55]	https://github.com/stamatak/standard-RAxML
BOOSTER (online)	Lemoine et al.[56]	https://booster.pasteur.fr/
FigTree v1.4.4	Github	https://github.com/rambaut/figtree/
pamlX v4.9a (codeml module)	Ziheng Yang[57]	https://github.com/abacus-gene/paml
TimeTree	Kumar et al.[58,59]	https://timetree.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Maria Laura Mascotti (m.l.mascotti@rug.nl).

Materials availability

All material newly generated in this study are available freely upon request to the lead contact.

Data and code availability

- Ancestral sequences experimentally characterized in this work have been deposited in GenBank under the following accession codes: AncBVMO2.0_bacteria (OR264814), AncBVMO2.1_bacteria (OR264815), AncBVMO2.2_bacteria (OR264816). The accession codes are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Molecular phylogeny and ancestral sequence reconstruction

A dataset of BVMOs was constructed by BLASTp homology searches using as query the sequences of experimentally characterized BVMOs (Table S1) in GenBank non-redundant protein sequences (nr) and Uniprot KB databases. Searches were performed restricting the taxonomy to all classes of Bacteria and the domains Archaea and Eukarya. Then, the collected sequences were gathered and vetted for the presence of the two conserved Rossmann folds (GxGxxG/A) and the BVMO fingerprints ((A/G)GxWxxxx(F/Y)P(G/M)xxxD and FxGxxxHxxxW(P/D))[29,60]. Redundancy was removed (cut-off < 95% identity) employing CD-HIT[53] and partial sequences were also discarded. The final dataset contained 507 BVMO sequences (Table S1). As outgroup, validated FMO sequences were added. Multiple sequence alignments (MSAs) were constructed in MAFFT v7[54] and manually trimmed for single sequence extensions/insertions and gaps. To construct the phylogeny a conservative trimming of the indels was performed by removing them only when present in less than 10% of the total sequences in the dataset. The maximum likelihood phylogeny was inferred in RAxML v8.2.10[55] (HPC-PTHREADS module) employing the fast-climbing algorithm (-f a), 500 bootstraps replicates, under gamma distribution and automated model selection. The tree was subjected to transfer bootstrap expectation (TBE) in BOOSTER (online)[56] to obtain the phylogeny with TBE values as support. FigTree v1.4.4 was employing to visualize and edit the tree.

Ancestral sequence reconstruction was performed, using the untrimmed MSA (547 sequences, 676 sites) and the topology of the obtained phylogeny, as marginal reconstruction in pamlX v4.9a (codeml module)[57]. Sequences were analyzed using the LG substitution matrix, an empirical equilibrium amino acid frequencies (model = 2) and 16 gamma categories (α value re-estimated). The posterior probability distribution at the selected nodes (AncBVMO0, AncBVMO1, AncBVMO2(2.0), AncBVMO3, AncBVMO2.1 and AncBVMO2.2) was recorded. The length of the targeted ancestors was determined by Fitch parsimony[41]. Sites reconstructed with PP < 0.8 and with a second-best state displaying PP > 0.2, were considered ambiguously reconstructed. To address the robustness of the reconstruction[61], the sequences of the alternative ancestors AltAncBVMO2.0, AltAncBVMO2.1 and AltAncBVMO2.2 were also reconstructed.

Cloning, expression and purification

The synthetic genes ordered as lyophilized powder were resuspended to a final concentration of 10 ng/μL in sterile Milli-Q water. Golden Gate cloning method was employed to clone the genes. A modified pBAD plasmid with BsaI restriction sites and a N-terminal 6×His-SUMO tag was used as the recipient vector. The cloning mixture contained: 75 ng of pBAD vector, around 40 ng of genes (to reach a molar ratio of 2:1 between gene inserts and the vector), 15 U BsaI-HF V2, 15 U T4 DNA ligase, 1× T4 ligase buffer and nuclease-free Milli-Q water to a final volume of 20 μL. A T100™ Thermal Cycler (BIO-RAD) was employed and the cloning mixtures underwent the following cycling steps: i) 30 cycles of holding at 37 °C for 5 min and then 16 °C for 10 min; ii) incubation at 55 °C for 10 min; iii) incubation at 65 °C for 20 min. Then, 3.5 μL of the assembly mix were used to transform under standard procedure 100 μL *E. coli* NEB 10-beta $CaCl_2$-competent cells. The recovered cells were plated on LB-agar plates containing 50 μg/mL ampicillin and incubated at 37 °C overnight. Several single colonies were picked to grow in LB medium. The plasmids were extracted and verified by sequencing (Eurofins Genomics). The strains containing the correct recombinant plasmid were stored in 15 % glycerol at -70 °C.

For expression, pre-cultures were prepared in LB medium with 50 μg/mL ampicillin overnight. 400 mL of Terrific Broth medium containing 50 μg/mL ampicillin were inoculated with 4 mL of the preculture and incubated at 37 °C for around 4 h. L-arabinose was added to a final concentration of 0.02 % (*w/v*) to induce the cells when the OD_{600} reached 0.6 – 0.8. Cultures were incubated at 24 °C in a shaker (135 rpm) for 20 h. After that, cells were harvested by centrifugation at 5000 g for 10 min. The pellets were resuspended in 50 mM TRIS-HCl buffer, pH 8.0, containing 150 mM KCl and 10 % glycerol (Buffer B) and supplemented with 0.5 mM dithiothreitol, 0.5 mM phenyl methyl sulfonyl fluoride and 100 μM FAD. Cells were disrupted by sonication (70 % amplitude, 2 sec on/ 3 sec off, for 15 min). Cell-free extracts (CFE) were obtained by centrifugation at 14000 g, 4 °C, for 40 min and filtered by 0.45 μm filters. The CFEs were loaded into the column containing Ni Sepharose® 6Fast Flow (Cytiva) equilibrated with Buffer B. The non-bound proteins were washed out with 10 column volume (CV) of Buffer B while the weakly bound proteins were removed with 10 CV of Buffer B containing 25 mM imidazole. The target enzymes were eluted with Buffer B containing 500 mM imidazole. The eluted fractions with yellow color were pooled and loaded into a HiPrep 26/10 desalting column (Cytiva) to remove the imidazole. The purity of the enzymes was analyzed by SDS-PAGE (Figure 1D). The purified enzymes were frozen in small aliquots with liquid nitrogen and stored at -70 °C. To determine the concentration of holoenzymes, the enzymes were first incubated at 95 °C for 20 min to denature them and then centrifuged at 12000 g for 10 min. The free FAD amount in the supernatant was calculated by UV-Vis spectroscopy in a Jasco V-660 spectrophotometer based on the extinction coefficient of FAD (ε_{FAD} = 11.3 mM^{-1} cm^{-1} at 450 nm). This is the same than the concentration of holoenzymes.

Substrates screening and conversions

To determine the functionalities of the AncBVMOs, 24 compounds including ketones, sulfides and steroids were screened (Table S2). These compounds were divided into four groups (6 compounds in each) based on their classifications for easy isolation on GC analysis. The compounds in the same group were mixed with DMSO as solvent to prepare the substrate mixes (mix 1 - 4)[62,63]. The final concentration of each substrate was 0.5 mM (except for benzyl phenyl sulfide which was 0.25 mM due to its poor solubility). The reaction mixtures (400 μl final volume) included 2.0 μM AncBVMOs, 100 μM NAD(P)H, 5.0 μM PTDH, 20 mM sodium phosphite, 5 % DMSO and buffer B in 2 mL sealed vials. The reactions were set in a shaker (135 rpm) at 24 °C for 20 h in duplicate. PTDH and sodium phosphite were used to regenerate the cofactor NAD(P)H.

To determine the activities, phenylacetone, 2-heptanone and thioanisole were selected to conduct conversions with AncBVMOs, AncBVMO mutants and extant BVMOs (PAMO and *Tm*CHMO, both produced in-house). Reaction mixtures contained 2.0 mM of each single substrate, 1.0 μM enzyme, 200 μM NADPH, 2.5 μM PTDH while other conditions were the same as the conversions above. For anaerobic conversions, the dissolved oxygen in the buffer was removed by flushing with nitrogen for 20 min. All conversion measurements were done in duplicate. The conversions were analyzed by a GC-MS in a QP2010 Ultra apparatus (SHIMADZU) equipped with an Agilent HP-5MS column (30 m x 0.25 mm x 0.25 μm). The samples were first extracted with the same volume of ethyl acetate (EtOAc) containing 0.02 % (*v/v*) mesitylene as internal standard and dried with $MgSO_4$. The GC-MS was set with the injector and detector temperature at 260 °C, a split ratio of 5.0, and an injection volume of 1 μL. For substrate screening using substrate mixtures, the column temperature was held at 50 °C for 4 min, increased by 10 °C/min to 250 °C and held for 5 min. For the single substrate conversions, the temperature was held at 50 °C for 4 min, increased by 10 °C/min to 150 °C, then kept increasing by 20 °C/min to 250 °C and held for 2 min. The identity of substrates and corresponding products was assigned by comparison with the mass spectra libraries NIST11s.lib and NIST11.lib. Conversion percentages were calculated based on the consumption of the substrates normalized with the internal standard.

Given that AncBVMO2.2 could convert sulfides, a series of sulfides was tested for enantioselectivity (Table S4). All conversion reactions were carried out in duplicate under the following conditions: 1.0 μM of enzyme, 200 μM NADPH, 4.0 μM PTDH, 10 mM sodium phosphite and 5 mM substrates in 50 mM TRIS-HCl, pH 8.0, containing 1 % DMSO as cosolvent. The reactions were stirred at 30 °C for 24 – 50 h. Once finished, the reactions were extracted with EtOAc (2 x 0.6 mL), dried onto Na_2SO_4 and the samples were directly analysed by GC-MS and HPLC to determine the level of conversion, as well as the enantiomeric excesses, of the sulfoxides obtained based on the calibration curves of products. GC-MS analyses were performed with a GC Hewlett Packard 7890 Series II equipped with a Hewlett Packard 5973 chromatograph MS (Agilent Technologies) using a HP-5MS cross-linked methyl siloxane column (30 m x 0.25 mm x 0.25 μm, 1.0 bar N_2). Substrates and products were quantified by using calibration curves. The following temperature program was employed: 50°C (5 min), 10 °C/min to 200°C (7 min). HPLC analyses to determine the enantiomeric excesses were carried out with a Thermo-Fischer UltiMate 3000 chromatograph equipped with a Thermo UltiMate detector. To determine the enantiomeric excesses of chiral sulfoxides Chiralcel OD (25 cm x 0.46 cm, Daicel), Chiralcel OB (25 cm x 0.46 cm, Daicel), Chiralpak IA (25 cm x 0.46 cm, Daicel) or Chiralpak IC (25 cm x 0.46 cm, Daicel) columns were employed, using *n*-hexane/IPA mixtures as eluent. The degree of conversion in the sulfoxidation of pyrmetazole and pantoprazole sulfide were determined by HPLC employing a calibration curve.

Size exclusion chromatography

AncBVMO2.0, AncBVMO2.1 and AncBVMO2.2 were analyzed by gel filtration chromatography at 280 K using a Superdex200 10/300 GL column (Cytiva), equilibrated with 20 mM HEPES buffer, pH 7.3, containing 150 mM NaCl. The absorbance was monitored at 254, 280 and 445 nm.

Steady-state kinetics

Steady-state kinetic parameters of AncBVMO2.2 were obtained by monitoring the depletion of nicotinamide cofactor, NADPH, at 340 nm with a Jasco V-660 spectrophotometer. The observed rate constants were the initial rates calculated based on the extinction coefficient of NADPH ($\epsilon_{NADPH,340}$= 6.22 mM^{-1} cm^{-1}). The kinetics assays were done by mixing 1.0 μM enzyme with varied concentrations of phenylacetone (0 – 2.0 mM) and NADPH (0 - 200 μM) in Buffer B at 25 °C in triplicate. For kinetics towards NADPH, the phenylacetone concentration was fixed to 2.0 mM, while for kinetics towards phenylacetone, the NADPH concentration was fixed to 100 μM. The kinetic parameters were calculated with the fitting data using Michaelis-Menten equation in GraphPad Prism v6.07. The uncoupling rates with 100 μM NAD(P)H were determined in absence of substrates in triplicate.

$NAD(P)^+$ titration

$NAD(P)^+$ affinity was determined by titration experiments. The AncBVMOs stock solutions were concentrated to ~ 60 μM and a high concentration $NAD(P)^+$ stock (10 mM) was prepared. Small amounts (0.2 μl – 2 μl) of the $NAD(P)^+$ were added into AncBVMOs solution to reach a certain $NAD(P)^+$ concentration gradually from 0 to 1000 μM. The spectral changes caused by the binding of enzyme and $NAD(P)^+$ were recorded by a Jasco V-660 spectrophotometer every time after adding the $NAD(P)^+$ stock. The absorbance changes were plotted over the wavelengths to find the point where the absorbance changed the most. Then, the $NAD(P)^+$ affinity could be

visualized by plotting the absorbance changes at the selected wavelength over the NAD(P)$^+$ concentrations and fitted using the Michaelis-Menten equation (eq1) to calculate the K_d values in GraphPad Prism v6.07.

$$k_{obs} = \frac{k_{cat}[S]}{K_m+[S]} \quad \text{(Equation 1)}$$

Pre-steady-state kinetics

Pre-steady-state kinetics were measured in a SX20 stopped-flow spectrometer equipped with either the photodiode array (PDA) detector or the photomultiplier tube (PMT) module (Applied Photophysics, Surrey, UK). To create anaerobic conditions, all components and buffer were flushed with nitrogen for 15 min in sealed vials and supplemented with 5 mM glucose and 0.3 μM glucose oxidase (*Aspergillus niger* type VII, Sigma-Aldrich). To confirm the cofactor usage of the AncBVMOs, reductive half-reactions were performed with PMT monitoring the FAD reduction at 450 nm. The absorbance change at 450 nm was recorded after mixing final 3.5 μM AncBVMOs and various NAD(P)H concentrations (final 0 – 500 μM) anaerobically at 25 °C. The observed rates (k_{obs}) were obtained by fitting traces to exponential functions. The k_{red} and K_d values were calculated by fitting k_{obs} against NAD(P)H concentrations with Michaelis-Menten equation (eq1). All measurements were conducted in triplicate.

To observe the formation of C4a-(hydro)peroxyflavin intermediate, the reduced AncBVMOs were mixed with dioxygen and monitored with the PDA detector. Anaerobically, 10 μM AncBVMO were reduced by adding 1.2 equivalent of NAD(P)H to the enzyme solution. The yellow color would fade away if the AncBVMOs are fully reduced. The spectral changes were recorded in time immediately after mixing with buffer containing various concentrations of dioxygen. Four different dioxygen concentrations (0.13, 0.31, 0.62, 0.96 mM after mixing) were used by mixing the anaerobic enzyme solution with (1) air-saturated buffer; (2) a mix of equal volumes of argon-saturated buffer and O_2-saturated buffer; (3) O_2-saturated buffer; (4) O_2-saturated buffer on ice. The observed rates (k_{obs}) were obtained by fitting traces to exponential functions. The k_{ox} values were the slopes of the curves obtained by plotting k_{obs} against O_2 concentrations. The data were analyzed by Pro-Data Viewer v4.2.12, Pro-Kineticist v1.0.13 (Applied Photophysics, Surrey, UK) and GraphPad Prism v6.07 (La Jolla, CA, USA). All measurements were conducted in duplicate or triplicate.

Redox potential measurement

To determine the redox potentials of AncBVMOs, the xanthine/xanthine oxidase method was employed[64]. The enzyme concentration used was 10 μM. 20 μM anthraquinone-2,6-disulphonate (AQ26DS, E°= -184 mV) was found to be the best dye for AncBVMO2.0 and AncBVMO2.1 while 20 μM anthraquinone-2-sulfonate (AQ2S, E°= -225 mV) was suitable for AncBVMO2.2. The measurements were done with a Jasco V-660 spectrophotometer and spectra were collected every 90 s until benzyl viologen started to get reduced. Because the spectra of the dyes partially overlap with the enzymes, the AncBVMOs and the two dyes were reduced separately to find the isosbestic points (AncBVMO2.0: 334 nm; AncBVMO2.1 and AncBVMO2.2: 340 nm; AQ2S: 353 nm; AQ26DS: 354 nm). The absorbance changes at these wavelengths were used for the calculations.

Site-directed mutagenesis

Site-directed mutagenesis was performed by using the Quick-Change method. Primers are shown in Table S5. The PCR mixture contained 1× Pfu Ultra II Hotstart Master Mix, 25 ng template plasmid, 0.2 μM forward and reverse primers. It was subjected to the following procedure in a T100™ Thermal Cycler: activation at 95 °C for 3 min, a cycle (denaturation of template DNA at 95 °C for 30 s, annealing at 60 °C for 30 s, extension at 72 °C for 3.5 min) repeated 30 times and the final extension at 72 °C for 5 min. Next the mixture was incubated with DpnI at 37 °C for 2 h to digest the templates. The template-free PCR mixture was then transformed into *E. coli* NEB 10-beta $CaCl_2$-competent cells. Several colonies were picked and verified by Sanger sequencing and further stored in 15 % glycerol at -70 °C.

Structural modeling

AlphaFoldv2.1.0 was used to predict the three-dimensional structures of the ancestral proteins[65]. AlphaFold generated structures for the five most likely alignments to homologous sequences and used these, together with known protein database structures up to the release date of the 10th of December 2021, to create 3D models. For each ancestral variant, the highest scoring model was selected. NAD(P)H and flavin were inserted by superposition with an X-ray structure of *Thermobifida fusca* BVMO (2YLX) with bound NADP[66]. All modeled complexes were monomers. The generated NADPH- and NADH-complexed structures were subjected to energy optimization of the NADPH/NADH cofactor (the protein structures were frozen at this stage) using the YASARA forcefield[67].

Molecular dynamics (MD) simulations were carried out under Yasara[68] using the Yamber3 force field[69]. The MD simulations of the extant BVMOs with NADH/NADPH bound were performed using structure PDB ID: 2YRL, with all crystallographic waters removed (the AlphaFold models also lack crystallographic water). A cubic cell with periodic boundary extended 7.5 Å beyond any protein atom. The cell was neutralized with up to 0.5 % of NaCl ions. Long range electrostatics, beyond 7.62 Å, were calculated with the particle mesh Ewald method[70]. Prior to the MD simulations, the entire simulation cell was subjected to an energy minimization. The MD simulation was started at 5 K, which was gradually increased to 298 K in the first 30 ps of each trajectory. Subsequently, another 19.97 ns was used to allow the structure to equilibrate while the subsequent 10 ns was used to calculate the root mean square fluctuation (RMSF) for each residue. A total of 20 independent trajectories, started with different initial velocities, were run for each modeled

enzyme variant and snapshots were collected every 25 ps. Thus, in total 0.6 μs of MD simulations was run per enzyme complex of which the last 8000 snapshots (= 20 x 10,000/25) were used to determine the RMSF.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification

All conversions were conducted in duplicate with the conversion percentages calculated based on the depletion of substrates. The area of the substrate peak of GC chromatograph was normalized by using an internal standard. All the conversion percentages are displayed as mean values ± standard deviation (SD).

Statistical analysis

Uncoupling rates measurements were conducted in duplicate, while all (pre-) steady-state kinetic measurements were done in triplicate. All observed rates are displayed as mean values ± standard deviation (SD).

Cell Reports, Volume *43*

Supplemental information

Evolution of the catalytic mechanism at the dawn of the Baeyer-Villiger monooxygenases

Guang Yang, Ognjen Pećanac, Hein J. Wijma, Henriëtte J. Rozeboom, Gonzalo de Gonzalo, Marco W. Fraaije, and Maria Laura Mascotti

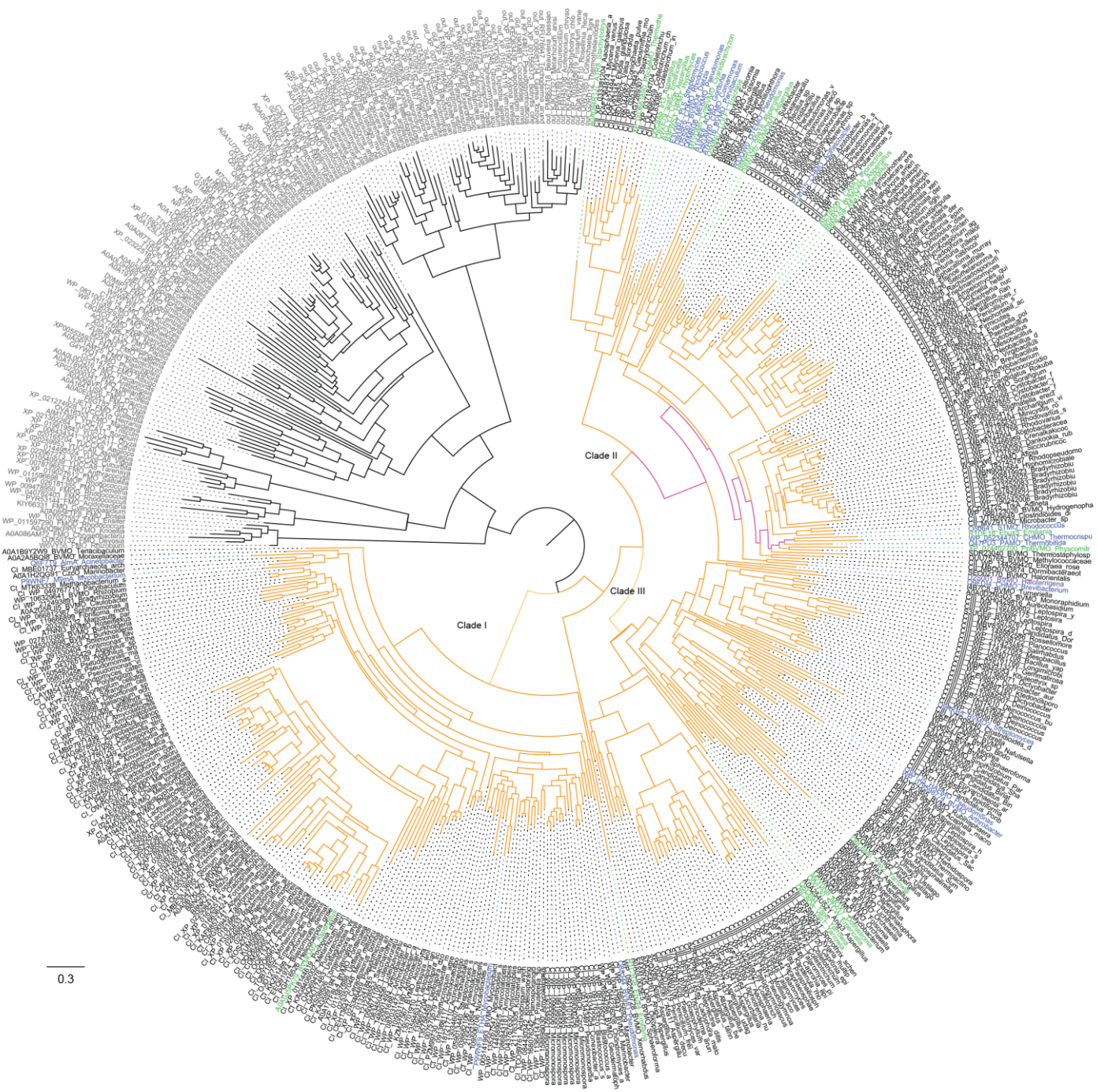


Figure S1. Phylogeny of BVMO family, related to Fig 1.

Evolutionary history of the BVMO family. The tree was constructed in RaxML (500 rapid bootstraps, LG model with empirical frequencies). BVMO sequences are shown in yellow branches and FMOs (outgroup) in black branches. The tree is shown as midpoint rooted. The sequences which have been experimentally characterized, or for which there is experimental evidence of their activity, are shown in blue (from Bacteria), green (from Eukarya) and violet (from Archaea). Clade designations are shown at the deep divergence points. The evolutionary trajectory explored here is shown in pink. The scale bar indicates substitutions per site.

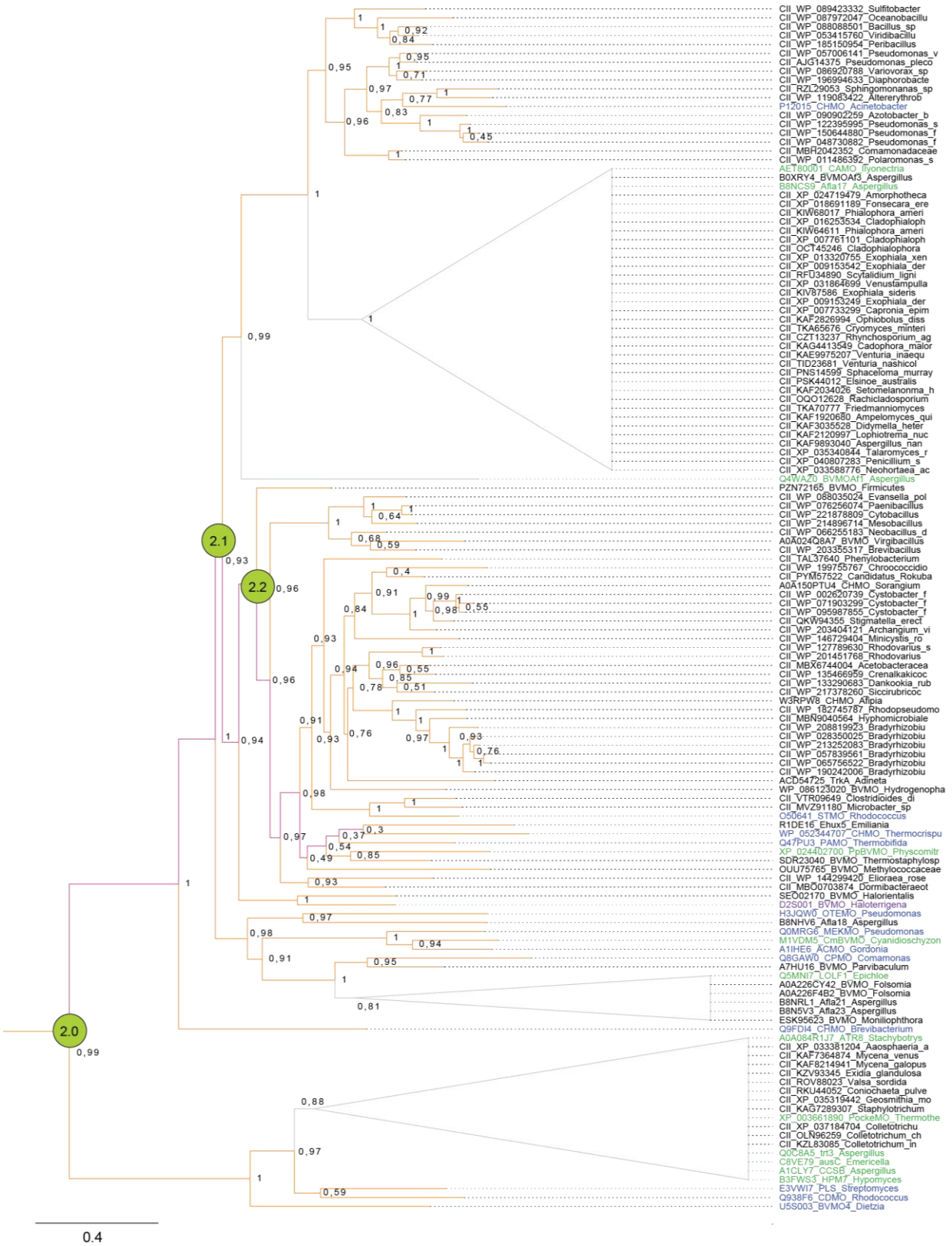


Figure S2. Phylogeny of BVMO family, related to Fig 1.

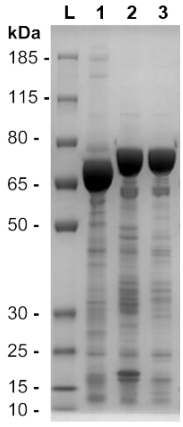
Closer view to clade II with the resurrected nodes shown in green circles. The support values (TBE, calculated with BOOSTER) are shown at the nodes.

A

Reconstruction	Overall PP	Length	Ambiguous sites[#]	Substitutions*
AncBVMO0	0.63	525	146 (28%)	-
AncBVMO1	0.73	492	112 (23%)	252 (to AncBVMO0)
AncBVMO2 (2.0)	0.73	535	139 (26%)	279 (to AncBVMO0)
AncBVMO3	0.80	495	75 (15%)	173 (to AncBVMO0)
AltAncBVMO2.0	-	535	-	-
AncBVMO2.1	0.90	535	62 (12%)	169 (to AncBVMO2.0)
AltAncBVMO2.1	-	535	-	-
AncBVMO2.2	0.93	535	48 (9%)	63 (to AncBVMO2.1)
AltAncBVMO2.2	-	535	-	-

[#]*map* state PP < 0.8 & Alt state PP > 0.2
*Number of substitutions to the previous targeted ancestor in the phylogeny

B



C

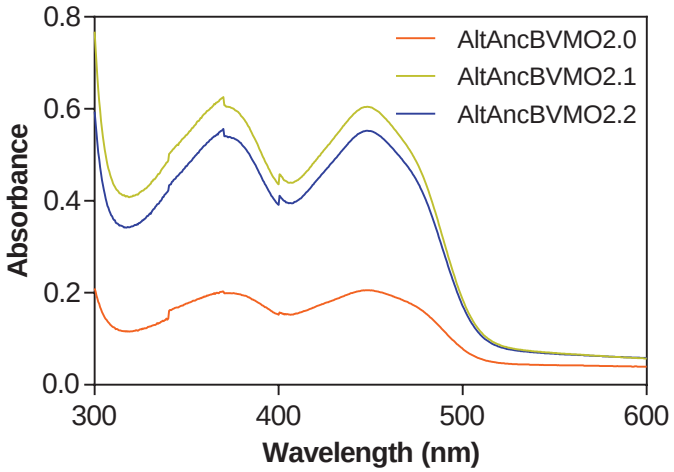


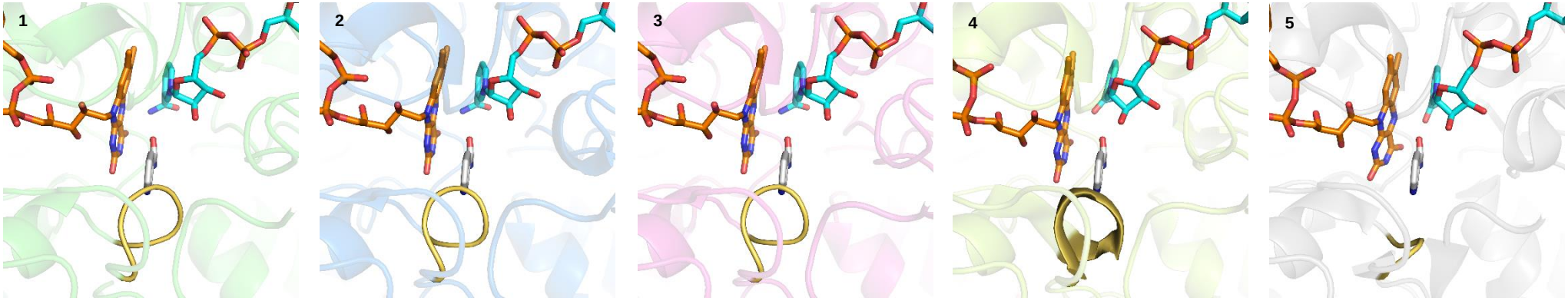
Figure S3. Ancestral sequence reconstruction, related to Fig 1.

(A). The overall posterior probabilities ($\overline{PP}$) for the reconstruction of each targeted node are given along with data about the accuracy of the reconstruction.

(B). SDS-PAGE analysis of purified AltAncBVMOs. Lanes: L) protein ladder; 1) AltAncBVMO2.0; 2) AltAncBVMO2.1; 3) AltAncBVMO2.2.

(C). Absorbance spectra of AltAncBVMOs. The spectra were recorded in a Jasco V-660 spectrophotometer with enzyme sample in 50 mM Tris-HCl buffer, pH 8.0, containing 150 mM KCl and 10 % glycerol at 25 °C.

A



B

Site	AncBVMO2.0			AncBVMO2.1			AncBVMO2.2			PAMO		Structural feature	Experimentally assayed		
	State	PP	AltPP	State	PP	AltPP	State	PP	AltPP	Site	State		Enzyme	Mutant	Activity
144	T	0.994		T	0.81		V	1		150	S	FAD binding site	A2.1	T144V	No
146	A	0.89		L	0.32	M(0.28)	C	1		152	Q	FAD binding site	A2.1	L146C	No
150	P	1.00		P	1		A	0.76	P(0.27)	156	P	Linker region	A2.1	P150A	No
169	A	0.90		A	0.81		G	0.99		175	G	Close to FAD pocket	A2.1	A169G	No
170	R	0.94		R	0.91		H	0.96		176	N	Close to FAD pocket	A2.1	R170H	No
218	R	0.97		A	0.99		A	1		225	A	Linker region	A2.0	R218A	No
253	E	0.20		H	0.5	Q(0.12)	P	0.99		260	Q	Substrate tunnel	A2.1	H253P	No
279	L	0.15		R	0.42	H(0.24)	S	0.64	G(0.21)	287	D	Close to substrate binding site	A2.1	R279S	No
326	C	0.99		A	0.98		T	0.99		335	T	Linker region, close to R337	A2.1	A326T	No
434	Y	0.50	F(0.30)	F	0.95		L	1		443	L	Substrate binding site	A2.1	F434L	No
444	Q	1.00		Q	0.94		H	1		453	H	Close to substrate binding site	A2.1	Q444H	No
505	L	0.51		L	0.28	I(0.26)	F	1		514	F	Side chain pointing to substrate binding site	A2.1	L505F	No

Figure S4. Structural differences between AncBVMOs and extant derived BVMOs, related to Figs 3A, B & C.

(A). The presence/absence of the loop at the entrance of the substrate tunnel of AncBVMOs and extant BVMOs. The catalytic cores of AncBVMO2.0 (**1**, green), AncBVMO2.1 (**2**, blue), AncBVMO2.2 (**3**, purple), PAMO (**4**, lemon, PDB= 4D04) and TmCHMO (**5**, gray, PDB= 5M10) are displayed. The loop influencing the substrate profile is highlighted in yellow. FAD and NADPH are displayed as sticks in orange and cyan respectively. A molecule of the substrate nicotinamide bound in the substrate pocket of TmCHMO is displayed in white sticks and showed superimposed into all other structures presented. Images were created in PyMOL Graphics System v2.4.0.

(B). Mutational analysis of the substitutions to find the sequence determinants of functionality. The structural features were analyzed based on AlphaFold predicted structures docked with FAD and NAD(P).

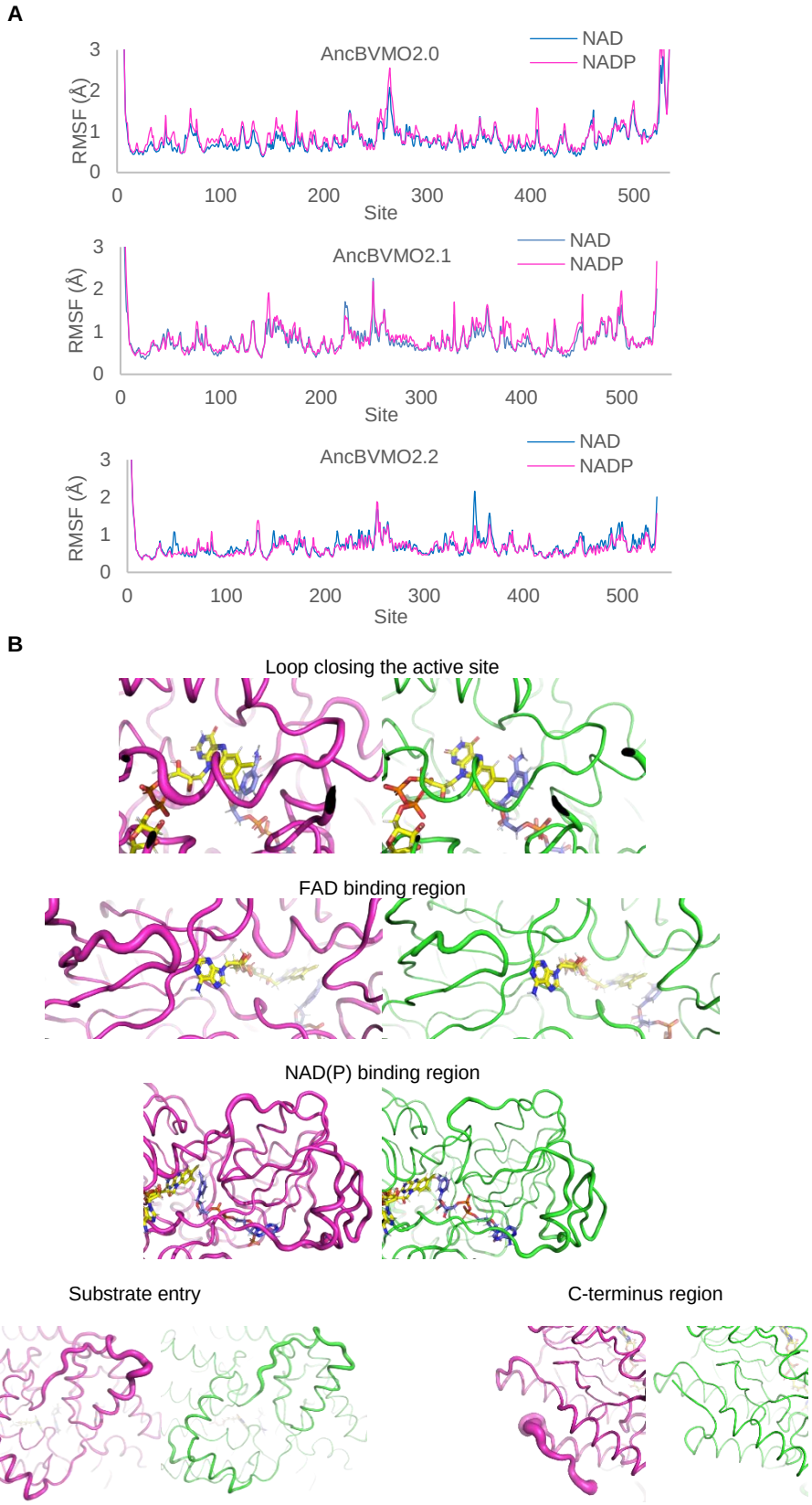


Fig S5. MD simulation analyses of AncBVMOs with FAD and NADH/NADPH docked, related to Fig 3D.

(A). The graphs show the RMSF values for each of the AncBVMOs when bound to NADH (blue) or NADPH (magenta).

(B). Flexible regions depicted in the AncBVMO2.0 model (hotpink) and PAMO (green). FAD (yellow) and NAD(P)H (blue) are shown as sticks. The thickness of the shown backbones linearly corresponds to the RMSF values on an absolute scale.

Table S2. Screening and conversions of AncBVMOs with NAD(P)H, related to Fig 2 & STAR Methods. The screening reactions contained 2 μM AncBVMOs, 0.5 mM substrate (except for benzyl phenyl sulfide which was 0.25 mM because of the poor solubility), 100 μM NAD(P)H, 5 μM PTDH, 20 mM sodium phosphite, 5 % DMSO in 50 mM Tris-HCl buffer, pH 8.0, containing 150 mM KCl and 10 % glycerol. A negative control was set without any enzyme. For single substrate reactions, 2 mM single substrate, 1 μM enzyme, 200 μM NADPH, 2.5 μM PTDH were used while other conditions are the same. The 400 μL reaction mixture were prepared in sealed 2 mL vials. Reactions were carried out at 24 °C for 20 h in duplicate. The conversions were analyzed by GC-MS and the reported percentages are presented as mean ± SD based on the depletion of substrates.

Conversions using NADPH					
Mix	**#**	**Substrate**	**A2.0**	**A2.1**	**A2.2**
1	1	Bicyclo[3.2.0]hept-2-en-6-one	0	1.7 ± 0.2	53.4 ± 1.5
	2	Cyclohexanone	0	0	0
	3	2-Heptanone	0	0	>99.9
	4	Phenylacetone	0	0	>99.9
	5	Thioanisole	0	22.3 ± 2.7	99.9 ± 0.1
	6	Benzyl phenyl sulfide	0	10 ± 2.7	98.9 ± 0.1
2	7	Cyclopentanone	0	0	0
	8	Cycloundecanone	0	0	0
	9	4-Methylcyclohexanone	0	0	0
	10	2-Phenylcyclohexanone	0	0	0
	11	4-Hydroxyacetophenone	0	0	0
	12	Stanolone	0	0	0
3	13	Cyclooctanone	0	0	0
	14	Cyclododecanone	0	0	0
	15	4-Phenylcycloheanone	0	0	0
	16	Cyclohexyl methyl ketone	0	0	44.8 ± 5.1
	17	3-Ocatanone	0	0	77.7 ± 1.9
	18	Androstenedione	0	0	0
4	19	2-Dodecanone	0	0	99.3 ± 0.7
	20	1,6-Dihydrocarvone	0	0	0
	21	4-Hydroxycyclohexanone	0	0	0
	22	(±)-Wieland-Miescher ketone	0	0	0
	23	Cyclopentadecanone	0	0	0
	24	Androsta-1,4-diene-3,17-dione	0	0	0
Conversions using NADH					
1	1	Bicyclo[3.2.0]hept-2-en-6-one	0	0	1.0 ± 0.1
	2	Cyclohexanone	0	0	0
	3	2-Heptanone	0	0	7.8 ± 0.1
	4	Phenylacetone	0	0	33.0 ± 1.4
	5	Thioanisole	0	0	18.2 ± 0.3
	6	Benzyl phenyl sulfide	0	0	1.7 ± 0.1

Comparison of conversions between AncBVMOs and AltAncBVMOs			
Enzymes	**Phenylacetone**	**2-Heptanone**	**Thioanisole**
PAMO	> 99.9	98.1 ± 1.9	86.2 ± 10.3
TmCHMO	2.5 ± 0.4	0	49.3 ± 12.3
AncBVMO2.0	0	0	0
AncBVMO2.1	0	0	2.7 ± 0.1
AncBVMO2.2	72.5 ± 3.6	50.0 ± 5.0	41.6 ± 16.7
AltAncBVMO2.0	0	0	0
AltAncBVMO2.1	0	0	0
AltAncBVMO2.2	40.5 ± 2.9	25.8 ± 0.4	25.7 ± 0.3

Table S3. Other donors/acceptors tested for AncBVMOs, related to Fig 2. Alternative donors were tested by monitoring FAD reduction after mixing 5.0 µM AncBVMOs and 2.0 mM trimethylamine or diethylamine or 10 % glyrerol, in triplicate. Phenylacetone, benzaldehyde and acetophenone were assessed by conversions of 2.0 mM substrates with/without oxygen in duplicate. The anaerobic condition was created by flushing the reaction mixture with nitrogen for 10 min before the reaction was initated by addition of enzymes. Metal ions were tested as electron acceptors by mixing them with reduced AncBVMOs in anaerobic conditions in triplicate. Muconic acid and cysteamine were tested in triplicate by monitoring the consumption of NAD(P)H after mixing 1.0 µM enzyme with 200 µM NADPH and 2.0 mM substrate.

	Compounds	**Enzymes**	**Donor**	**Presence of O_2**
Alternative Donors	Trimethylamine	AncBVMO2.0, AncBVMO2.1	-	Yes
	Diethylamine			
	Glycerol			
Alternative Acceptors	Phenylacetone	AncBVMO2.1, AncBVMO2.2	NADH for AncBVMO2.1 NADPH for AncBVMO2.2	No
	Benzaldehyde			
	Acetophenone			
	Mn^{2+} ($MnCl_2$)	AncBVMO2.1		
	Zn^{2+} ($ZnCl_2$)			
	Cu^{2+} ($CuSO_4 \cdot 5H_2O$)			
	Ni^{2+} ($NiCl_2$)			
	Fe^{3+} ($FeCl_3$)			
Other substrates	Muconic acid		NAD(P)H	Yes
	Cysteamine			

Table S4. Sulfoxidations catalyzed by AncBVMO2.2, related Fig 2H. All conversion reactions were carried out using the following conditions: 1.0 µM AncBVMO2.2, 0.20 mM NADPH and 5.0 mM substrates in 50 mM Tris-HCl, pH 8.0 containing 1 % DMSO as cosolvent. NADPH was regenerated by supplementing the reaction mixtures with 4.0 µM PTDH and 10 mM sodium phosphite. The reactions were stirred at 30 °C for the times indicated in the table. Samples were analyzed by GC-MS to determine conversions based on the calibration curves of products and chiral HPLC to determine the enantiomeric excesses and the conversions of pyrmetazole and pantoprazole sulfide based on the calibration curves of the corresponding sulfoxides. The data are represented by mean values of duplicate experiments.

#	Substrate	t (h)	c (%)	Sulfone (%)	*ee* (%)	Config
1	PhSMe	24	93	<3	71	*S*
2	PhSCH2Cl		<3	<3	n.d.	*n.d.*
3	PhSEt		97	<3	80	*S*
4	PhSCH=CH_2		23	<3	78	*S*
5	PhSPr		72	<3	65	*S*
6	PsSCyPr		34	<3	21	*S*
7	PhSEtOH		21	<3	57	*R*
8	PhSEtCl		10	<3	47	*R*
9	NafSMe		28	<3	56	*S*
10	BnSMe		>97	7	90	*S*
11	BnSEt		63	<3	60	*R*
12	BnSiPr		33	<3	29	*R*
13	PhSBn		16	<3	35	*S*
14	pO2NPhSMe		22	<3	40	*S*
15	pNCPhSMe		30	<3	61	*S*
16	oClPhSMe		34	<3	47	*S*
17	mClPhSMe		87	<3	85	*S*
18	pClPhSMe		89	<3	77	*S*
19	pBrPhSMe		73	<3	63	*S*
20	mBrPhSMe		77	<3	66	*S*
21	pMePhSMe		92	<3	70	*S*
22	pMeOPhSMe		97	<3	78	*S*
23	pAllylOPhSMe		52	<3	65	*S*
24	pHOPhSMe		21	<3	49	*S*
25	pAcHNPhSMe		11	<3	39	*S*
26	pH2NPhSMe		<3	<3	n.d.	*n.d.*
27	Pyrmetazole	50	4	<3	6	*S*
28[a]	Pyrmetazole	48	5	<3	9	*S*
29	Pantoprazole sulfide		<3	<3	n.d.	*n.d.*
30	CySMe	24	97	<3	39	*S*
31	PhSOMe		<3	-	n.d.	*n.d.*
32	pMePhSOMe		5	-	<3	*n.d.*
33	BnSOME		29	-	7	*S*

n.d.: Not determined

[a]Reaction temperature was 40 °C

Table S5. Quick Change primers designed for site-directed mutagenesis, related to STAR Methods.

Primers	Sequence (5' to 3')
A2.1_A169G_F	TGTATCACACTGGGCGTTGGCCACA
A2.1_A169G_G	TGTGGCCAACGCCCAGTGTGATACA
A2.1_A326T_F	CACCCTTTTGGGACGAAGCGACCGC
A2.1_A326T_R	GCGGTCGCTTCGTCCCAAAAGGGTG
A2.1_F434L_F	AGTCCTTCTGTGCTTAGTAATATGC
A2.1_F434L_R	GCATATTACTAAGCACAGAAGGACT
A2.1_H253P_F	TTCCTTTTGATCCCTCAACTAAATC
A2.1_H253P_R	GATTTAGTTGAGGGATCAAAAGGAA
A2.1_L146C_F	ACGGCAACCGGTTGCTTATCGGCACCA
A2.1_L146C_R	TGGTGCCGATAAGCAACCGGTTGCCGT
A2.1_L505F_F	AAGCCCCGTACTTTTATGCCCTATG
A2.1_L505F_R	CATAGGGCATAAAAGTACGGGGCTT
A2.1_P150A_F	CTGTTATCGGCAGCAAACATACCCG
A2.1_P150A_R	CGGGTATGTTTGCTGCCGATAACAG
A2.1_Q444H_F	TATTGAGCAGCATGTAGAATGGATA
A2.1_Q444H_R	TATCCATTCTACATGCTGCTCAATA
A2.1_R170H_F	ATCACACTGCGCATTGGCCACATGG
A2.1_R170H_R	CCATGTGGCCAATGCGCAGTGTGAT
A2.1_R279S_F	AAAGGCGGATTTTCATTTATGGGCAC
A2.1_R279S_R	GTGCCCATAAATGAAAATCCGCCTTT
A2.1_T144V_F	CTTATTACGGCAGTCGGTCTGTTATC
A2.1_T144V_R	GATAACAGACCGACTGCCGTAATAAG
A2.1_T255E_F	TTTGATCACTCAGAAAAATCAGCAATG
A2.1_T255E_R	CATTGCTGATTTTTCTGAGTGATCAAA
A2.0_R218A_F	CGTGGTGCCTGCGAACAATAGACCGACAGATC
A2.0_R218A_R	GTCTATTGTTCGCAGGCACCACGTAGCTCGGGG