Accessing Chemo- and Regioselective Benzylic and Aromatic Oxidations by Protein Engineering of an Unspecific Peroxygenase

Anja Knorrscheidt, Jordi Soler, Nicole Hünecke, Pascal Püllmann, Marc Garcia-Borràs, Martin Weissenborn

Unspecific peroxygenases (UPOs) enable oxyfunctionalisations of a broad substrate range with unparalleled activities. Tailoring these enzymes for chemo- and regioselective transformations represents a grand challenge due to the difficulties in their heterologous productions. Herein, we performed a protein engineering in S. cerevisiae with the novel MthUPO. Experimental approaches were combined with computational modelling resulting in the screening of more than 5,300 transformants. This protein engineering led to a significant reshaping of the active site as elucidated by molecular dynamics. The k_{cat}/K_m was improved by 16.5-fold. Variants were identified with high chemo- and regioselectivities in the oxyfunctionalisation of aromatic and benzylic carbons, respectively. The benzylic hydroxylation was demonstrated to perform with excellent enantioselectivities of 95 % ee. Additionally, the first reported effective exchange of the conserved catalytic Glu residue was observed.

File list (2)

Manuscript_Knorrscheidt et al.pdf (566.94 KiB)	view on ChemRxiv • download file
Supporting_Knorrscheidt et al.pdf (5.26 MiB)	view on ChemRxiv • download file

Accessing Chemo- and Regioselective Benzylic and Aromatic Oxidations by Protein Engineering of an Unspecific Peroxygenase

Anja Knorrscheidt,[a] Jordi Soler,[c] Nicole Hünecke,[a] Pascal Püllmann,[a] Marc Garcia-Borràs*[c] and Martin J. Weissenborn*[a, b]

[a]   Anja Knorrscheidt, Nicole Hünecke, Pascal Püllmann, Jun.-Prof. Dr. Martin J. Weissenborn
        Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry
        Weinberg 3, 06120 Halle, Germany
        E-Mail: martin.weissenborn@ipb-halle.de
[b]   Jun.-Prof. Dr. Martin J. Weissenborn
        Institute of Chemistry, Martin Luther University Halle-Wittenberg
        Kurt-Mothes-Str. 2, 06120 Halle, Germany
[c]   Jordi Soler, Dr. Marc Garcia-Borràs
        Institut de Química Computacional i Catàlisi and Departament de Química, Universitat de Girona
        Carrer Maria Aurèlia Capmany 69, Girona 17003, Catalonia, Spain
        E-Mail: marc.garcia@udg.edu

        Supporting information for this article is given via a link at the end of the document.

Abstract: Unspecific peroxygenases (UPOs) enable oxyfunctionalisations of a broad substrate range with unparalleled activities. Tailoring these enzymes for chemo- and regioselective transformations represents a grand challenge due to the difficulties in their heterologous productions. Herein, we performed a protein engineering in *S. cerevisiae* with the novel *Mth*UPO. Experimental approaches were combined with computational modelling resulting in the screening of more than 5,300 transformants. This protein engineering led to a significant reshaping of the active site as elucidated by molecular dynamics. The k_{cat}/K_m was improved by 16.5-fold. Variants were identified with high chemo- and regioselectivities in the oxyfunctionalisation of aromatic and benzylic carbons, respectively. The benzylic hydroxylation was demonstrated to perform with excellent enantioselectivities of 95 % *ee*. Additionally, the first reported effective exchange of the conserved catalytic Glu residue was observed.

Unspecific peroxygenases (UPOs) recently arose as dream biocatalysts for oxyfunctionalisation reactions.[1] They enable the direct transformation of oxygen from hydrogen peroxide to a broad substrate scope capable of reaching impressive total turnover numbers (TONs) beyond 200,000.[2] There is a vast diversity of UPO sequences available within the fungal kingdom with more than 4,300 unique annotated sequences.[3] However, the widespread application and subsequent protein engineering of UPOs are hampered by their challenging heterologous production in fast-growing host organisms.[4] To address these challenges, Alcalde and co-workers[5] as well as our group,[6] developed yeast-based heterologous production setups. These systems gave access to novel UPOs as well as the opportunity for protein engineering. Amongst the recently discovered enzymes, a UPO derived from the thermophilic fungus *Myceliophthora thermophila* (*Mth*UPO) achieved the thus far highest UPO shake flask production yields. In contrast to the predominantly studied *Aae*UPO derived from *Agrocybe aegerita*, the *Mth*UPO revealed the oxidation of naphthalene to 1,4-naphthoquinone as one of two main products.[6-7] These 1,4-naphthoquinone core structures are intensely studied moieties with application in the industry as an intermediate for the anthraquinone synthesis[8] and are widely distributed in the plant and animal kingdom.[9] In nature, naphthoquinones occur in the human body as the blood clotting precursor vitamin K3,[10] in plants as carcinogenesis plumbagin analogues[11] or in insects as chemical defence.[12]

We hypothesised that *Mth*UPO's ability to perform efficient aromatic as well as benzylic oxyfunctionalisations could provide a starting point to engineer chemoselective UPO variants for either specificity. Using the colourimetric assay 5-nitro-1,3-benzodioxole (NBD),[5] around 5,300 transformants were screened to obtain improved variants exhibiting complementary regio- and chemoselectivities as well as improved enantioselectivities. Computational modelling described the molecular basis behind the observed catalytic and selectivity improvements.

We commenced our studies by generating a homology model for the *Mth*UPO structure based on the solved crystal structure of the UPO from *Marasmius rotula* (*Mro*UPO, PDB: 5FUJ, 29% identity and 41 % similarity, see SI for details). The NBD substrate was docked into a representative structure of the most populated cluster, which aided in identifying substrate-binding residues (Fig. 1a). Nine positions (L56, F59, L60, L86, F154, T155, S159, A161, L206, see Fig. S9) were randomised by the Golden Mutagenesis technique.[13] The mutant library was transformed in *Saccharomyces cerevisiae* producing the corresponding variants. We screened the library using the colourimetric NBD assay[14] in combination with the recently established split-GFP analysis in yeast.[15] The combination of these two assays enabled the distinction between substrate conversion and secretion of the respective enzyme. Variations at position L56 substantially influenced the production of the
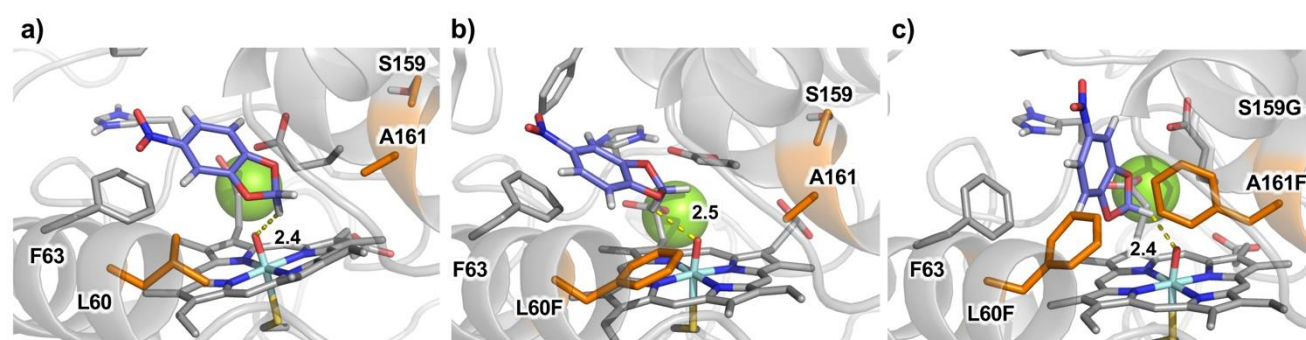
Figure 1. Active site arrangement of *Mth*UPO and evolution of NBD's catalytically relevant binding modes in **a**) wildtype and the variants **b**) L60F and **c**) L60F/S159G/A161F as observed from MD simulations. Mutated positions are highlighted in orange.

enzyme. Only 52 % of the transformants displayed a split-GFP signal (Tab. S3). Site saturation at position L60 yielded improved variants with superior TONs for NBD conversion. The variants L60M (1.2-fold relative to the wildtype), L60Q (1.3-fold) and L60F (2.7-fold) showed the most noticeable improvements (Tab. 1 and S4). Variant L60F also exhibited a 1.4-fold improved catalytic efficiency (k_{cat}/K_m, Tab. 2). The variant library of position F154, which is located at the entrance channel, turned out to be a pivotal position for the NBD conversion. Even though 81 % of the variants were secreted according to the split-GFP signal, only the rediscovered wildtype enzymes displayed activity (Fig. S1).

To gain insights into the highly conserved catalytic residue E158, we tested the variants E158A and E158D. E158A led to the expected loss of NBD activity. To retain the deprotonation of the peroxo-iron complex, we introduced an aspartate as a shortened amino acid.

Astonishingly, variant E158D proved to be active with a decelerated but prolonged product formation. Whereas the *Mth*UPO wildtype reached the maximum product formation after a few minutes, the variant E158D converted the substrate overnight yielding 1.5-fold improved TONs relative to the wildtype. This variant represents the first exchange of the catalytic glutamate residue while preserving the UPO activity. Computational analysis based on MD simulations revealed that the E158D mutation induces higher rigidity to the α-7-helix than the wildtype (Figs. S10 and S11), which might be responsible for its prolonged conversion.

The 23.6-fold decreased k_{cat} value of variant E158D (Tab. 2) indicates that the mutation has a severe impact on the rate-limiting step of the reaction. This impact is most likely related to the H_2O_2 activation via proton transfer steps that are mediated by this residue.[16] To assess the influence of the peroxide species with these variants, we employed cumene hydroperoxide and *tert*-butyl hydroperoxide as oxygen source. These experiments revealed diminished activities using the alternative peroxides with the wildtype and E158D and no activity with E158A (Fig. S3).

The catalytic glutamic acid residue in UPOs is well known to coordinate to histidine or arginine stabilising the negative charge formed at the acid-base catalyst.[1, 17] MD simulations proved that H88 can establish persistent H-bond interactions with both E158 and E158D (Fig. S11). We therefore further investigated interactions of the positions E158 and H88 by generating the variants H88A and H88A/E158D. Variant H88A led to no product formation while variant H88A/E158D was able to convert NBD with a 4.5-fold decreased TON relative to the wildtype. MD simulations indicated that in the absence of H88 (H88A) the α-7-helix becomes even more flexible and disordered, whereby the side chain of position E158 is barely preorganised to activate H_2O_2. On the other hand, the shorter side chain of E158D in H88A/E158D is still well-positioned to generate compound 1 (Fig. S12).

To continue with the mutagenesis approach, we grouped two amino acid residues and saturated them simultaneously with a reduced codon degeneracy (NDT), thereby obtaining double mutants with improved TONs (Tab. S4). The improved double mutated variants were L60F/F154I and L60F/F154V (1.2-fold improvement compared to the wildtype, Tab. S4).

Table 1. The catalytic activity of MthUPO variants for the hydroxylation of NBD.[a]

Catalyst	Conversion [%]	TOF [min⁻¹]	TON
*Mth*UPO WT	29	72	4340
*Mth*UPO L60F	77	194	11610
*Mth*UPO E158D[b]	44	6	6560
*Mth*UPO H88A/E158D[c]	32	16	970
*Mth*UPO L60F/S159G/A161F	84	210	12590
*Mth*UPO F59Q/L60M/S159G/F154A	77	192	11540
*Mth*UPO F59Q/L60F/S159G[d]	76	379	22760

TOF = turnover frequency, TON = turnover number, standard deviation < 3.2 %, [a] Reaction conditions: 20 nM *Mth*UPO variant, 300 μM NBD, 1 mM H_2O_2, 100 mM KPi buffer (pH 7), 5 % acetone (v/v), measurement conditions: absorbance was measured at 425 nm for one hour in triplicates, values were calculated with the corrected extinction coefficient of 10870 $M^{-1} \cdot cm^{-1}$ (see Supporting Information), [b] reaction time over night, [c] 100 nM *Mth*UPO variant, [d] 10 nM *Mth*UPO

Table 2. Biochemical characterisation of the *Mth*UPO wildtype and the evolved variants towards the substrate NBD.[a]

Catalyst	T_m [°C]	K_m [μM]	k_{cat} [s⁻¹]	K_{cat}/K_m [M⁻¹s⁻¹] x 10^4
*Mth*UPO WT	63.5	386	7.1	1.9
*Mth*UPO L60F	55.7	110	2.9	2.7
*Mth*UPO E158D	60.8	186	0.3	0.1
*Mth*UPO L60F/S159G/A161F	56.9	422	132.2	31.3
*Mth*UPO F59Q/L60M/S159G/F154A	54.7	290	59.4	20.5
*Mth*UPO F59Q/L60F/S159G	55.8	303	47.0	15.5

Standard deviation <18 %, [a] Values were calculated with the corrected extinction coefficient of 10870 M⁻¹*cm⁻¹, reaction conditions and Michaelis-Menten plots: see Supporting Information.

The residues from the single and double saturation libraries, which had a positive or neutral influence on the NBD conversion (Tab. S4) were selected for recombination. These residues were F59Q, L60F/Q/M, A57I, F154I/V, E158D, S159N/G and A161I/F. Inspired by Reetz' single-, double-, and triple-code saturation mutagenesis approaches,[18] we chose to recombine all residues and variations in all combinatorial possibilities including the respective wildtype amino acid residue. This combination led to 864 mutants and required the screening of more than 2,300 transformants.

The recombination library resulted in the discovery of triple and quadruple mutations with up to 16.5-fold improved catalytic efficiency (k_{cat}/K_m, Tab. 2). All of the most active variants harboured amino acid exchanges at the L60 position (L60F/M) and the mutation S159G. The kinetic measurements revealed 8.2-fold (F59Q/L60F/S159G), 10.8-fold (F59Q/L60M/S159G/F154A) and 16.5-fold (L60F/S159G/A161F) increased k_{cat}/K_m values relative to the wildtype (Tab. 2). Whereas the K_m value of NBD was decreased or similar to the wildtype, the values were

significantly decreased for H_2O_2 (Tab. S5). The k_{cat} for NBD, however, was substantially improved for the identified triple and quadruple mutants with an 18.6-fold increase for the variant L60F/S159G/A161F relative to the wildtype.

Docking calculations were used to obtain starting points for MD simulations in order to analyse the binding of NBD in the wildtype and the variants L60F and L60F/S159G/A161F (Figs. 1, S13 and S14). These simulations revealed a switch in the binding mode of NBD when moving from wildtype to L60F and L60F/S159G/A161F (Fig. S13). Due to the inclusion of bulkier residues in the inner active site (first L60F, then A161F), the NBD substrate was reoriented from its original more buried binding pose in the wildtype to partially occupying the entrance channel in L60F/S159G/A161F. This new binding pose allows the NBD to better approach the Fe=O active species in a near attack conformation, thus facilitating the oxidation reaction (Fig. S14).

To determine the influence of the engineered variants on the regio- and chemoselectivity, we investigated the oxyfunctionalisation of naphthalene and its derivatives (Scheme 1). The most active variants were selected, and an initial screening was performed by GC-MS, resulting in significant changes in the selectivity (Fig. S6-S8). Utilising a syringe pump setup achieved more than 10,000 TONs with variant F59Q/L60F/S159G for the conversion of naphthalene to 1,4-naphthoquinone **1a** (Scheme 2a). These TONs represent a 10-fold increase to the previously published *Mth*UPO wildtype activity.[6] By increasing H_2O_2 equivalents, 2-hydroxy-1,4-naphthoquinone **1b** was formed as a by-product, which is a natural dye known as Lawsone.[19]

The biotransformation of 2-methylnaphthalene with the wildtype led predominantly to 6-methyl-1,4-naphthoquinone (Fig. S7). Variant L60F/S159G/A161F was able to shift the major product formation to 2-methyl-1,4-naphthoquinone
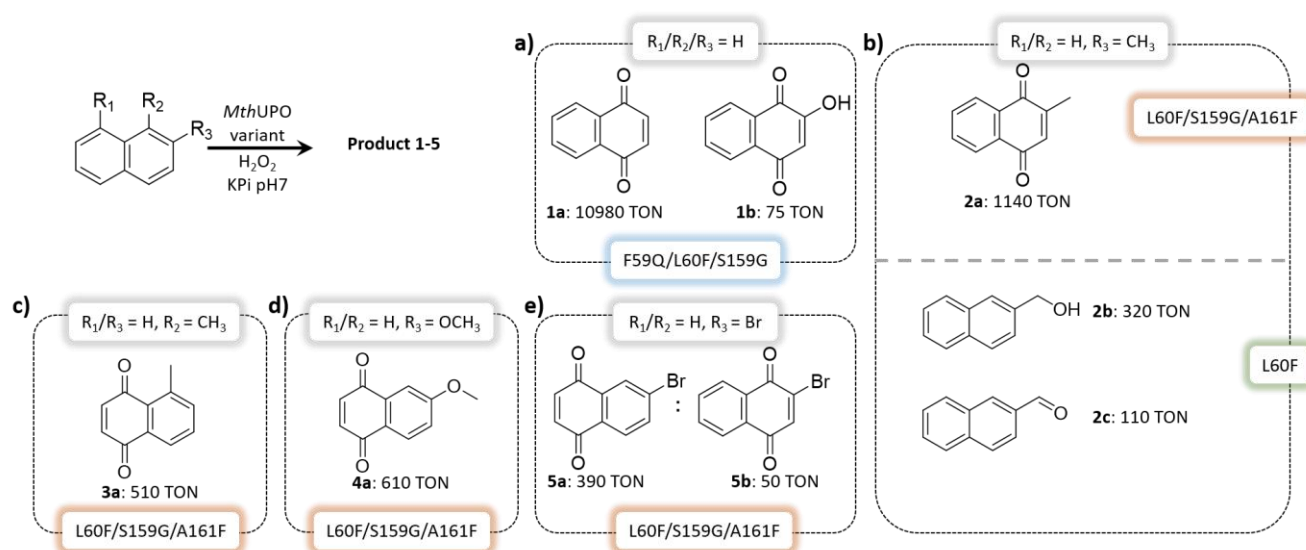
Scheme 1: Catalytic activity of *Mth*UPO variants for the hydroxylation of naphthalene and its derivatives (reaction conditions are shown in Tab. S6, standard deviation <5.2 %)

3

(**2a**), also known as vitamin K3, demonstrating the regioselectivity of this variant (Scheme 2b).

We were pleased to note that we could also identify a variant with the preference for the methyl hydroxylation of 2-methylnaphthalene. Variant L60F showed an altered chemoselectivity dramatically suppressing the aromatic hydroxylation and accessing the hydroxylation of the methyl group (**2b**) as the major product. Also the overoxidation to the aldehyde was observed (**2c**, Scheme 2b). MD simulations with the variant L60F and 2-methylnaphthalene demonstrated that only the methyl group was able to approach the active Fe=O species in a catalytically competent pose due to the presence of the bulky phenylalanine residue L60F (Fig. S15). When 2-methylnaphthalene was bound into the variant L60F/S159G/A161F (Fig. S16), however, a binding pose similar to the previously observed for NBD in this variant was observed (Figs. 1 and S13). This binding mode is promoted by hydrophobic interactions occurring in the newly engineered active site, which is dominated by the presence of several aromatic residues (F59, F60, F63, F154, and F161). Within this new binding pose, the substituted aromatic ring 1 of 2-methylnaphthalene is placed close enough to the Fe=O catalytic species to react while keeping the 2-methyl group away from it. This different binding mode of the L60F variant is responsible for the observed switch in chemoselectivity (Figs. S16 and S17).

When the methyl substitution was changed from 2-methyl to 1-methylnaphthalene a decrease in the activity of 50 % was observed yielding to the opposite unsubstituted ring 2 oxidation (Scheme 2c). More surprisingly, also substitutions in position 2 at the naphthalene led to the oxidation of the unsubstituted ring generating the products 6-methoxy- (**4a**, Scheme 2d) and 6-bromo-1,4-naphthoquinone (**5a**, Scheme. 2e) with diminished TONs relative to 2-methylnaphthalene. For 2-bromo-naphthalene also 2-bromo-1,4-naphthoquinone (**5b**) was detected as a by-product. MD simulations with variant L60F/S159G/A161F and 2-methoxy-naphthalene (Fig. S18) revealed a preferential binding in a catalytically competent pose equivalent to the binding mode observed for 2-methyl naphthalene (Fig. S17). However, the bulky methoxy group at the 2-position sterically clashes with the L60F residue, and the naphthalene core needs to rotate slightly when approaching the Fe=O active species to place the 2-methoxy group away from the L60F and the heme cofactor. This rotation brings the unsubstituted aromatic ring 2 closer to the Fe=O active species, as opposite to 2-methylnaphthalene (Fig. S17), inducing a switch in the regioselectivity (Fig. S18). A similar behaviour was also observed for 1-methylnaphthalene when bound in variant L60F/S159G/A161F in a catalytically competent pose (Fig. S19).
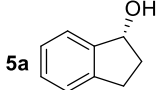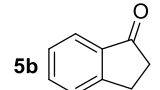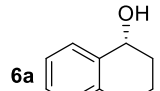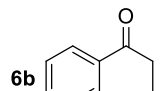
Intrigued by the chemoselective benzylic hydroxylation of L60F, we tested this variant with indane and tetralin (1,2,3,4-tetrahydronaphthalene). L60F was able to convert indane with more than 8,000 TONs and an improved enantioselectivity for the (*R*)-1-indanol (**5a**) from 85 % *ee*

(wildtype) to 95 % *ee* (Tab. 3 and S7). Interestingly, further examination of the most active variants L60F/S159G/A161F and F59Q/L60M/S159G/F154A revealed the excess in formation of the (*S*)-enantiomer (Tab. S7).

For the bioconversion of tetralin to the alcohol **6a**, an improved enantioselectivity for the (*R*)-enantiomer was detected for L60F (74 % *ee*) relative to the wildtype (45 % *ee*). Similar to indane, a variant could be identified, which forms predominantly the (*S*)-enantiomer (F59Q/L60M/S159G/F154A). The overoxidation to the ketone as a major product was achieved by the same variant with TONs of 440.

MD simulations with indane bound within the L60F variant (Fig. S20) characterised a preferential binding pose of the substrate that resembles the previously observed pose for NDB and 2-methylnaphthalene in L60F/S159G/A161F (Figs. S13 and S16). In this case, indane mainly interacts with the aromatic rings of F63 and L60F, establishing C-H···π interactions. These hydrophobic interactions keep the substrate in a preferred binding pose where only the pro-*R* 1-H-indane is close enough and well-aligned to the Fe=O active species to be efficiently hydroxylated (Fig. S20).

Table 3. The catalytic activity of *Mth*UPO variants towards benzylic hydroxylation yielding chiral products.[a]

R	Catalyst	Product	TON	% *ee*[d]
CH₂	*Mth*UPO L60F[a]	**5a**	8160	95 (*R*)
CH₂	*Mth*UPO L60F[a]	**5b**	450	-
(CH₂)₂	*Mth*UPO L60F[b]	**6a**	860	74 (*R*)
(CH₂)₂	*Mth*UPO F59Q/L60M/ S159G/F154A[c]	**6b**	440	-

TON = turnover number, standard deviation < 6.5 %, reaction conditions: [a] 100 nM *Mth*UPO variant, 1 mM indane, 1 mM H₂O₂, 100 mM KPi buffer (pH 7), 5 % acetone (v/v), 1 h at 25 °C in triplicates, [b] 100 nM *Mth*UPO variant, 1 mM 1,2,3,4-tetrahydronaphthalene, 1 mM H₂O₂, 100 mM KPi buffer (pH 7), 5 % acetone (v/v), 1 h at 25 °C in triplicates, [c]] 100 nM *Mth*UPO variant, 1 mM 1,2,3,4-tetrahydronaphthalene, 2 mM H₂O₂, 100 mM KPi buffer (pH 7), 5 % acetone (v/v), 2 h at 25 °C in triplicates, [d] determined by chiral GC.

4

In the present work, the novel *Mth*UPO was engineered using *S. cerevisiae* as expression host and NBD as an assay system. By initial saturation of relevant amino acid residues in single and double saturation mutageneses, key residues were identified. This mutagenesis also led to an increased understanding of the highly conserved catalytic glutamate, its stabilisation by histidine, and—most surprisingly—the generation of the active variant E158D.

A recombination library of the best performing variants revealed substantially improved variants with triple and quadruple mutations. Variant L60F demonstrated a high chemoselectivity for the oxidation at the benzylic position, whereas L60F/S159G/A161F selectively formed aromatic ring oxidations with 2-methylnaphthalene. L60F also revealed an improved enantioselectivity up to 95 % *ee* for the conversion of indane to its benzylic hydroxylated derivative. MD simulations described a reshaping of *Mth*UPO's active site during evolution due to the inclusion of bulkier residues, which induced novel binding modes for the studied substrates, and rationalised enzyme-substrate specific interactions providing the molecular basis for the experimentally observed oxidation patterns.

These results demonstrate the high versatility and evolvability of *Mth*UPO to achieve regio-, chemo- and enantioselective biocatalytic oxyfunctionalisations.

[1] M. Hobisch, D. Holtmann, P. Gomez de Santos, M. Alcalde, F. Hollmann, S. Kara, *Biotechnol. Adv.* **2020**, doi: 10.1016/j.biotechadv.2020.107615.

[2] Y. Ni, E. Fernandez-Fueyo, A. Gomez Baraibar, R. Ullrich, M. Hofrichter, H. Yanase, M. Alcalde, W. J. van Berkel, F. Hollmann, *Angew. Chem. Int. Ed.* **2016**, *55*, 798-801.

[3] M. Hofrichter, H. Kellner, R. Herzog, A. Karich, C. Liers, K. Scheibner, V. W. Kimani, R. Ullrich, in *Grand Challenges in Fungal Biotechnology* (Ed.: H. Nevalainen), Springer International Publishing, Chamb, **2020**, pp. 369-403.

[4] A. Conesa, F. van De Velde, F. van Rantwijk, R. A. Sheldon, C. A. van Den Hondel, P. J. Punt, *J. Biol. Chem.* **2001**, *276*, 17635-17640.

[5] P. Molina-Espeja, E. Garcia-Ruiz, D. Gonzalez-Perez, R. Ullrich, M. Hofrichter, M. Alcalde, *Appl. Environ. Microbiol.* **2014**, *80*, 3496-3507.

[6] P. Püllmann, A. Knorrscheidt, J. Münch, P. R. Palme, W. Hoehenwarter, S. Marillonnet, M. Alcalde, B. Westermann, M. J. Weissenborn, *bioRxiv* **2020**, doi: 10.1101/2020.1107.1122.216432.

[7] P. Molina-Espeja, M. Canellas, F. J. Plou, M. Hofrichter, F. Lucas, V. Guallar, M. Alcalde, *ChemBioChem* **2016**, *17*, 341-349.

[8] J. Grolig, R. Wagner, in *Ullmann's Encyclopedia of Industrial Chemistry, Vol. 23*, Wiley-VCH, Weinheim, **2000**, pp. 733-737.

[9] I. Hook, C. Mills, H. Sheridan, in *Studies in Natural Products Chemistry, Vol. 41* (Ed.: R. Atta ur), Elsevier, **2014**, pp. 119-160.

[10] S. Shukla, C.-P. Wu, K. Nandigama, S. V. Ambudkar, *Mol. Cancer Ther.* **2007**, *6*, 3279.

[11] J. J. Inbaraj, C. F. Chignell, *Chem. Res. Toxicol.* **2004**, *17*, 55-62.

[12] L. V. Vujisić, D. Ž. Antić, I. M. Vučković, T. L. Sekulić, V. T. Tomić, B. M. Mandić, V. V. Tešević, B. P. M. Ćurčić, V. E. Vajs, S. E. Makarov, *Chem. Biodivers.* **2014**, *11*, 483-490.

[13] P. Püllmann, C. Ulpinnis, S. Marillonnet, R. Gruetzner, S. Neumann, M. J. Weissenborn, *Sci. Rep.* **2019**, *9*, 10932.

[14] a) P. Molina-Espeja, E. Garcia-Ruiz, D. Gonzalez-Perez, R. Ullrich, M. Hofrichter, M. Alcalde, *Appl. Environ. Microbiol.* **2014**, *80*, 3496-3507; b) M. Poraj-Kobielska, M. Kinne, R. Ullrich, K. Scheibner, M. Hofrichter, *Anal. Biochem.* **2012**, *421*, 327-329.

[15] J. Santos-Aberturas, M. Dorr, G. S. Waldo, U. T. Bornscheuer, *Chem. Biol.* **2015**, *22*, 1406-1414.

[16] S. Hayakawa, H. Matsumura, N. Nakamura, M. Yohda, H. Ohno, *FEBS J.* **2014**, *281*, 1409-1416.

[17] a) K. Piontek, E. Strittmatter, R. Ullrich, G. Grobe, M. J. Pecyna, M. Kluge, K. Scheibner, M. Hofrichter, D. A. Plattner, *J. Biol. Chem.* **2013**, *288*, 34767-34776; b) X. Yi, A. Conesa, P. J. Punt, L. P. Hager, *J. Biol. Chem.* **2003**, *278*, 13855-13859.

[18] a) Z. Sun, R. Lonsdale, G. Li, M. T. Reetz, *ChemBioChem* **2016**, *17*, 1865-1872; b) Z. Sun, P. T. Salas, E. Siirola, R. Lonsdale, M. T. Reetz, *BIOB* **2016**, *3*, 44.

[19] J. R. Widhalm, D. Rhodes, *Hortic. Res.* **2016**, *3*, 16046.

6

Accessing Chemo- and Regioselective Benzylic and Aromatic Oxidations by Protein Engineering of an Unspecific Peroxygenase

Anja Knorrscheidt,[a] Jordi Soler, [c] Nicole Hünecke, [a] Pascal Püllmann, [a] Marc Garcia-Borràs*[c] and Martin J. Weissenborn*[a, b]

[a.] Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle (Saale), Germany, martin.weissenborn@ipb-halle.de.

[b.] Institute of Chemistry, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 2, 06120 Halle (Saale), Germany.

[c.] Institut de Química Computacional i Catàlisi and Departament de Química, Universitat de Girona, Carrer Maria Aurèlia Capmany 69, Girona 17003, Catalonia, Spain.

Table of contents

1. General information

Chemicals. Solvents were used as provided without further purification from Carl Roth (Karlsruhe, DE) as GC ultra-grade. The commercially available compounds were also used without further purification from the following suppliers: *N,O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA, Macherey-Nagel, Düren, DE), hydrogen peroxide solution (30 % (*w/w*) in H_2O, Sigma-Aldrich, St. Louis, US), cumene hydroperoxide (contains 20 % aromatic hydrocarbon, TCI, Tokyo, JP), *tert*-butyl hydroperoxide (70 % in water, TCI, Tokyo, JP), 1,2-(methylenedioxy)-4-nitrobenzene (NBD, 98 %, Sigma-Aldrich, St. Louis, US), naphthalene (99 %, Sigma-Aldrich, St. Louis, US), (*R*)-(-)-1-Indanol (99 %, Sigma-Aldrich, St. Louis, US), (*S*)-(+)-1,2,3,4-tetrahydro-1-naphthol (98 %, TCI, Tokyo, JP), 1,2,3,4-tetrahydronaphthalene (99 %, Fluka, Buchs, CH), 1-indanone (99 %, Sigma-Aldrich, St. Louis, US), indane (95 %, TCI, Tokyo, JP), 1-hydroxyindane (99 %, TCI, Tokyo, JP), *α*-tetralone (96 %, Fluka, Buchs, CH), 2-methylnaphthalene (97 %, Sigma-Aldrich, St. Louis, US), 2-naphthaldehyde (98 %, TCI, Tokyo, JP), 2-hydroxymethylnaphthalene (97 %, abcr, Karlsruhe, DE), 1-methylnaphthalene (96 %, TCI, Tokyo, JP), 2-methyl-1,4-naphthoquinone (98 %, Sigma-Aldrich, St. Louis, US), 1,4-naphthoquinone (97 %, Sigma-Aldrich, St. Louis, US), 2-methoxy-1,4-naphthoquinone (98 %, Sigma-Aldrich, St. Louis, US), 1,2,3,4-tetrahydro-1-naphthol (97 %, Sigma-Aldrich, St. Louis, US), 2-bromo-1,4-naphthoquinone (98 %, Sigma-Aldrich, St. Louis, US), 4-octanone (98 %, TCI, Tokyo, JP), 2-methoxynaphthalene (98 %, TCI, Tokyo, JP), 2-bromonaphthalene (97 %, Fluka, Buchs, CH). As a buffer system, 50 mM potassium phosphate (KPi) pH 7 was utilised as an aqueous phase for the bioconversions.

Gas chromatography-mass spectrometry (GC-MS). Measurements were performed on a Shimadzu GCMS-QP2020 NX instrument (Shimadzu, Kyoto, JP) with a Lipodex E column (25 m x 0.25 mm, Macherey-Nagel, Düren, DE) for the chiral and on an SH-Rxi-5Sil MS (30 m x 0.25 mm, Shimadzu, Kyoto, JP) for the achiral measurements, whereas helium was utilised as the carrier gas. The samples were injected splitless (1 µl) with a liner temperature of 280 °C. The interface temperature was set to 290 °C. Ionisation was obtained by electron impact with a voltage of 70 V, and the temperature of the ion source was 250 °C. The oven temperature profile for each compound is shown in table S1. The detector voltage of the secondary electron multiplier was adjusted in relation to the tuning results with perfluorotributylamine. The GC-MS parameters were controlled with GCMS Real Time Analysis, and for data evaluation, GCMS Postrun Analysis (GCMSsolution Version 4.45, Shimadzu, Kyoto, JP) was used. Calibration and quantification were implemented in SIM (selected ion monitoring) mode with the corresponding *m/z* traces as shown in table S1 in triplicates. As internal standard 4-octanone (1 mM, *m/z* 128) in EtOAc was utilised. The product formation of 5-Methyl-1,4-naphthoquinone (**3a**) was confirmed by the consistent literature fragmentation pattern[1]: MS (EI) *m/z* 172, 144, 118, 116, 115, 90. For product quantification of 5-methyl-1,4-naphthoquinone (**3a**), 6-methoxy-1,4-naphthoquinone (**4a**, MS (EI) *m/z*: 188, 160, 134, 106, 63) and 6-bromo-1,4-naphthoquinone (**5a**, MS (EI) *m/z*: 384, 382, 296, 294, 266, 264, 73) the corresponding structural isomers menadione (**2a**), 2-methoxy-1,4-naphthoquinone (**4b**) and 2-bromo-1,4-naphthoquinone (**5b**) were utilised. Compound **1b** was derivatised according to the standard procedure of Macherey-Nagel (Düren, DE) with BSTFA: therefore the water was removed by an Eppendorf concentrator 5301 (Hamburg, DE) under vacuum at 60 °C for 1.5 h. The resulting residue was dissolved in 100 µl pyridine with ultrasound for five minutes, followed by addition of 100 µl BSTFA. The derivatisation reaction was accomplished at 80 °C for 20 min and the solution was directly utilised for GC-MS analysis resulting in the literature known fragmentation pattern[2]: MS (EI) *m/z* 231, 203.

Table S1 Instrumental parameters for GC-MS measurements

Products	GC-MS	Column	Temperature program
1,4-Naphthoquinone (*m/z* 158)	Achiral	SH-Rxi-5Sil MS	40 °C 7 °C/min to 190 °C 100 °C/min to 300 °C hold 10 min
Silylated 2-Hydroxy-1,4-naphthoquinone (Lawsone, *m/z* 231)	Achiral	SH-Rxi-5Sil MS	70 °C 20 °C/min to 270 °C 100 °C/min to 300 °C hold 4 min
2-Methyl-1,4-naphthoquinone (Menadione, *m/z* 172)	Achiral	SH-Rxi-5Sil MS	40 °C 7 °C/min to 190 °C 100 °C/min to 300 °C hold 10 min
2-Naphthalenemethanol (*m/z* 158) 2-Naphthaldehyde (*m/z* 156)	Achiral	SH-Rxi-5Sil MS	40 °C 7 °C/min to 190 °C 100 °C/min to 300 °C hold 10 min
5-Methyl-1,4-naphthoquinone (*m/z* 172)	Achiral	SH-Rxi-5Sil MS	40 °C 7 °C/min to 190 °C 100 °C/min to 300 °C hold 2 min
6-Methoxy-1,4-naphthoquinone (*m/z* 188)	Achiral	SH-Rxi-5Sil MS	40 °C 7 °C/min to 190 °C 100 °C/min to 300 °C hold 10 min
Silylated 6-Bromo-1,4-naphthoquinone (*m/z* 384)	Achiral	SH-Rxi-5Sil MS	70 °C 20 °C/min to 270 °C 100 °C/min to 300 °C hold 4 min
1-Hydroxyindan (*m/z* 133) 1-Indanol (*m/z* 132)	Achiral (quantification)	SH-Rxi-5Sil MS	40 °C 7 °C/min to 190 °C 100 °C/min to 300 °C hold 2 min
	Chiral (*ee* determination)	Lipodex E	95 °C 0.5 °C/min to 110 °C 100 °C/min to 200 hold 2 min
1,2,3,4-Tetrahydro-1-naphthol (*m/z* 148) α-Tetralone (*m/z* 146)	Achiral (quantification)	SH-Rxi-5Sil MS	40 °C 7 °C/min to 190 °C 100 °C/min to 300 °C hold 2 min
	Chiral (*ee* determination)	Lipodex E	75 °C 1 °C/min to 120 °C 100 °C/min to 200 hold 2 min

2. Mutagenesis of the *Mth*UPO and screening procedure

Single site saturation mutagenesis. Mutagenesis was performed using the Golden Mutagenesis technique[3] combined with the "22c-trick"[4] for residue randomisation. The *Mth*UPO wildtype gene combined with a *Sce*–α Galactosidase signal peptide at the *N*-terminus and a TwinStrep-GFP11 purification/detection tag at the *C*-terminus was chosen as genetic template[3] targeting the following amino acid residues within the active site or the entrance channel for randomisation: L56, F59, L60, L86, F154, T155, S159, A161, L206. As backbone structure, the expression plasmid pAGT572_Nemo 2.0 was utilised, enabling bacterial antibiotic selection (ampicillin resistance) and yeast auxotrophy selection (URA3 marker). As a placeholder for the target gene sequence, a lacZ cassette (approx. 600 bp) is integrated, which enables β- galactosidase based blue/white selection of transformants based on the conversion of X-Gal. For the mutagenesis, the Golden Mutagenesis protocol was applied.[3] The resulting Golden Gate setup mixture was transformed in chemically competent *E.coli* DH10B cells. After heat shock transformation (90 seconds at 42 °C) and recovery (1h at 37 °C) the mixture (approx. 320 μL) was split into two fractions, 60 μL were plated on selective LB agar plates (50 μg × mL⁻¹ X-Gal; 100 μg × mL⁻¹ carbenicillin; 150 μM IPTG) and the remaining volume used to directly inoculate 4 mL TB

Medium (100 μg × mL^{-1} carbenicillin). The plasmid DNA mixture was isolated from the liquid culture using a NucleoSpin Plasmid Kit (Macherey-Nagel, Düren, DE). The efficiency of the Golden Gate reaction was evaluated based on the performed blue/white screening and the Golden Mutagenesis analysing feature "Quick Quality Control"[5] monitoring the codon distribution at the previously mentioned positions. The resulting plasmid DNA mixture was then transformed into chemically competent yeast cells (INVSc1 strain) by polyethylene glycol/lithium acetate transformation. For transformation, an amount of 100 ng of the plasmid preparation was added to 10 μL of salmon sperm DNA (10 mg/mL; Sigma Aldrich, Hamburg, DE) and mixed. This mixture was then added to a thawed aliquot of INVSc1 cells on ice. 600 μL of transformation buffer (40 % (v/v) poly ethylene glycol 4000; 100 mM lithium acetate; 1 mM EDTA; 10 mM Tris-HCl pH 7.4) were added and the cells incubated under rigid shaking (30 °C; 850 rpm) for 30 min. Afterwards, 70 μL of pure DMSO were added and the cells incubated for a further 15 min at 42 °C without shaking. Finally, the cells were precipitated by short centrifugation, the supernatant discarded, and the cell pellet resuspended in 350 μL sterile ddH$_2$O. Different volumes were plated on Synthetic Complement (SC) Drop Out plates supplemented with 2 % (w/v) glucose as carbon source and lacking Uracil as an auxotrophic selection marker. Plates were incubated for at least 48 hours at 30 °C till clearly background distinguishable white colonies appeared.[3]

Double site saturation mutagenesis. The single-site saturation mutagenesis led to the identification of relevant amino acid residues, which were combined in a double site saturation mutagenesis approach based on spatial proximity and their potential interactions. The positions L60/F154, I52/A57, F59/L150, S159/A161, F63/L86, were therefore simultaneously randomised using an NDT codon degeneracy utilising the previously mentioned Golden Mutagenesis protocol.

Generation of a recombination library of the best performing variants. The best performing variants from the single, as well as from the double site saturation mutagenesis were chosen for recombination. For this approach, the best performing amino acid residues were selected and randomised. The following positions were chosen: L60 (WT/F/Q/M), F59 (WT/Q), A57 (WT/I), F154 (WT/V/I), E158 (WT/D), S159 (WT/G/N), A161 (WT/I/F). The library was generated based on the previously mentioned Golden Mutagenesis protocol yielding 864 possible combinations.

Point mutagenesis. By introducing a single amino acid change using one point mutation, we wanted to investigate the catalytic acid residue E158 and the coordinating residue H88. Therefore the variants E158A, E158D, H88A, H88A/E158D were created.

Table S2 Overview of generated libraries.

Library #	Mutagenesis approach	Degeneracy	Number of screened variants	Statistical library coverage [%][a]
1	Single saturation (9 positions)	22c-trick	For each position: 88 In total: 792	98
2	Double saturation (5 grouped positions)	NDT	For each position: 440 In total: 2200	95
3	Recombination (864 possible combinations)	Best performing variants from library 1/2	In total: 2376	94

[a] The statistical library coverage was calculated using the equation for fractional library completeness[6]: L = - V ln(1-F); L – Number of screened variants, V – number of possible variants, F – library coverage

Microtiter plate cultivation of *S. cerevisiae.* For *Mth*UPO production in microtiter plate format specialised 96 half deep well plates were utilised. The model type CR1496c was purchased from EnzyScreen (Heemstede, NL) and plates were covered with fitting CR1396b Sandwich cover for cultivation. Plates and covers were flushed before every experiment thoroughly with 70 % ethanol and

air-dried under a sterile bench until usage. In each cavity, 220 µL of minimal expression medium were filled and inoculated with single, clearly separated yeast colonies using sterile toothpicks. The minimal selective expression medium (1x concentrated SC Drop stock solution lacking uracil, 2 % (w/v) galactose, 71 mM potassium phosphate buffer pH 6.0, 3.2 mM magnesium sulfate, 3.3 % (v/v) ethanol, 50 mg/L haemoglobin, 25 mg/L chloramphenicol) was freshly prepared out of sterile stock solutions immediately before each experiment, mixed and added to the cavities. After inoculation of the wells the plates were covered, mounted on CR1800 cover clamps (EnzyScreen) and incubated in a Minitron shaking incubator (Infors, Bottmingen, SUI) for 72 h (30 °C; 230 rpm). After cultivation, the cells were separated from the peroxygenase containing supernatant by centrifugation (3400 rpm; 45 min; 4 °C).[3] The cell pellet was resuspended in the remaining supernatant, and glycerol was added to achieve a final concentration of 25 % (v/v). The sealed microtiter plate was frozen by liquid nitrogen and stored at – 80 °C as mother plate for subsequent hit verification.

Activity Screening via NBD assay in microtiter plate format. The utilisation of 5-nitro-1,3-benzodioxole for a colourimetric screening approach has been described before yielding the chromophore 4-nitrocatechol.[7-8] The conditions were slightly modified: after centrifugation 20 µl of the supernatant from the microtiter plate cultivation was transferred to a polypropylene 96 well screening plate (Greiner Bio-One, Kremsmünster, AT) and 180 µl of the master mix were added (end concentrations: 100 mM KPi buffer pH 7, 300 µM NBD, 1 mM H_2O_2, 5 % (v/v) acetone). The absorption was detected for one hour at 425 nm (interval: 30 s) starting directly after the addition of the master mix at an absorbance reader (TECAN, Grödig, AT). The reaction endpoint was determined overnight. Improved *Mth*UPO variants were identified by comparing the absorption values of the NBD assay and the fluorescence values of the split-GFP assay with the parental variant.

Table S3 Overview of secreted variants within the single saturated mutagenesis libraries.

Saturated position	Secreted variants [%]
L56	52
F59	93
L60	75
L86	76
F154	81
T155	83
S159	83
A161	72
L206	69

Split-GFP assay. Protein normalisation was performed employing the principle of a split GFP normalisation assay as described by Santos-Aberturas *et al.*[9] with slight modifications. The complementation fragment sfGFP 1–10 was cloned into the Golden Mutagenesis plasmid pAGM22082_cRed[3] for T7 promoter controlled expression in *E.coli*. For measurement 20 µl of yeast expression supernatant was transferred to a previously BSA blocked 96 well Nunc MaxiSorp Fluorescence plate (ThermoFisherScientific, Waltham, US) and 180 µl of sfGFP 1–10 inclusion body preparation added. Immediate fluorescence values (GFP fluorophore: excitation wavelength: 485 nm; emission wavelength: 535 nm; top read mode) was measured using a 96 well plate fluorescence reader (TECAN, Grödig, AT). After storage over two nights (at 4 °C) final fluorescence values were measured. Protein quantities were then normalised based on the relative fluorescence increase (differential values) and in comparison to the empty plasmid backbone. Based on the split-GFP assay, the percentage of secreted variants could be calculated (table S3).
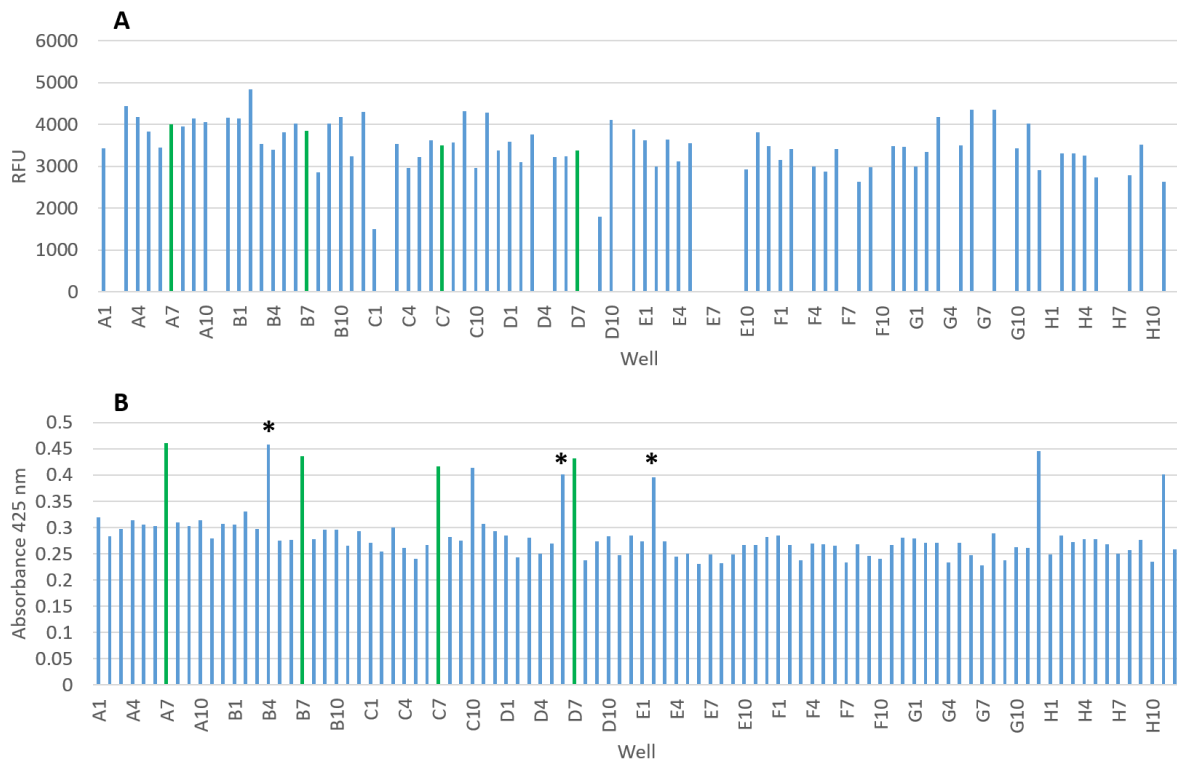
Figure S1 Screening results of saturated position F154, A) split-GFP results depicts secreted variants whereas B) illustrates the active variants within the NBD assay, green = WT as control, * = sequenced variants which confirmed the WT.

Automated data evaluation and verification. For the microtiter plate screening, including the NBD and split-GFP assay, an automated data evaluation by R Studio was utilised. Thereby, the best performing variants were identified based on their respective endpoint after one hour and overnight, their NBD slope and their NBD/GFP correlation when compared to the parental variant. For the data verification, the best performing variants were reproduced in a microtiter plate setup in triplicates. If the improved activity was confirmed, the protein was cultivated in shake flask scale and purified for further characterisation.

Shake flask cultivation and protein purification. For the preculture 50 mL of SC Drop out selection media (+ 2 % (w/v) Raffinose as non-repressible carbon source and 25 mg/L chloramphenicol) was inoculated with one single colony derived from a selection plate (SC Drop; -Uracil) and grown for 48 h (30 °C; 160 rpm; 80 % humidity). The main expression culture was inoculated with a starting optical density of 0.3. For large scale peroxygenase production rich non-selective expression medium (20 g/L peptone; 10 g/L yeast extract; 2 % (w/v) galactose; 71 mM potassium phosphate buffer pH 6.0; 3.2 mM magnesium sulfate; 3.3 % (v/v) ethanol; 25 mg/L chloramphenicol) was utilised. Cultivation was performed in 2.5 L Ultra yield flasks (Thomson Instrument, Oceanside, US) in a final culture volume of 500 mL per flask after sealing the flask with breathable Aeraseal tape (Sigma Aldrich, Hamburg, DE) allowing for gas exchange. The main cultures were incubated for further 72 h (25 °C; 110 rpm; 80 % humidity). After cultivation, the cells were separated from the peroxygenase containing supernatant by centrifugation (4300 rpm; 35 min; 4 °C). The supernatant was concentrated approx. 10-fold using ultrafiltration. Therefore a Sartocon Slice 200 membrane holder (Sartorius, Göttingen, DE) was equipped with a Sartocon Slice 200 ECO Hydrosart Membrane (10 kDa nominal cut-off; Sartorius) within a self-made flow setup. The flow system for ultrafiltration was operated by an EasyLoad

peristaltic pump (VWR International, Darmstadt, DE). The cleared supernatant (500 mL) was concentrated to a volume of 50 mL, and 900 mL of purification binding buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl) were added as a buffer exchange step. This sample was then concentrated to a final volume of 50 mL. Protein purification was implemented utilising the C-terminal attached double Strep II Tag (WSHPQFEK), coined TwinStrep® (Iba Lifesciences, Göttingen, DE). As a column, Gravity Strep-Tactin®XT Superflow® columns (5 mL; Iba Lifesciences) were chosen. In the first step, the column was equilibrated with 5 column volumes (CVs) binding buffer. The concentrated sample (50 mL) was filter sterilised (0.2 µm syringe filter) and applied to the column. After application, the column was washed with 6 CVs binding buffer. Elution was performed based on binding competition with biotin, therefore approx. 2 CV of elution buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl; 50 mM biotin) were applied to the column. The pooled elution fraction was then dialysed overnight (4 °C) against 5 L of storage buffer (100 mM potassium phosphate pH 7.0) using ZelluTrans dialysis tubing (6-8 kDa nominal cut-off; Carl Roth).[3] The enzyme concentration of the dialysed samples was determined by colourimetric BCA assay utilising a Pierce™ BCA Protein Assay Kit (ThermoFisherScientific, Waltham, US) following the instructions of the manufacturer. The samples were stored at − 20 °C until further utilisation.

3. Characterisation of *Mth*UPO variants

NBD assay with purified enzyme for TOF/TON determination. The purified *Mth*UPO variants were measured via NBD assay under the following conditions: 20 nM *Mth*UPO variant (exception: H88A/E158D with 500 nM and F59Q/L60F/S159G with 10 nM), 300 µM NBD, 1 mM H_2O_2, 100 mM KPi buffer (pH 7), 5 % acetone (v/v). The turnover numbers (TON), turnover frequency values and conversions were determined after 1h except for E158D where the overnight endpoint was determined. For determination of the catalytic performance, the TON, TOF and conversion were calculated based on the corrected extinction coefficient of 4-nitrocatechol. The actual extinction coefficient was calculated by a calibration curve of 4-nitrocatechol in 100 mM KPi buffer (pH 7) with 5 % acetone (v/v) (Fig. S2 A) leading to $\varepsilon_{425nm, \text{4-nitrocatechol}}$ = 11289 $M^{-1}*cm^{-1}$. This coefficient was corrected by the extinction coefficient of NBD $\varepsilon_{425nm, \text{NBD}}$ = 419 $M^{-1}*cm^{-1}$ (Fig. S2 B) yielding to the corrected coefficient of $\varepsilon_{425nm, \text{corr}}$ = 10870 $M^{-1}*cm^{-1}$, which can be utilised for the catalytic performance calculations:

$$TON = \frac{[product]}{[catalyst]} \left[\frac{mol}{mol}\right]$$

$$TOF = \frac{TON}{time} [min^{-1}]$$
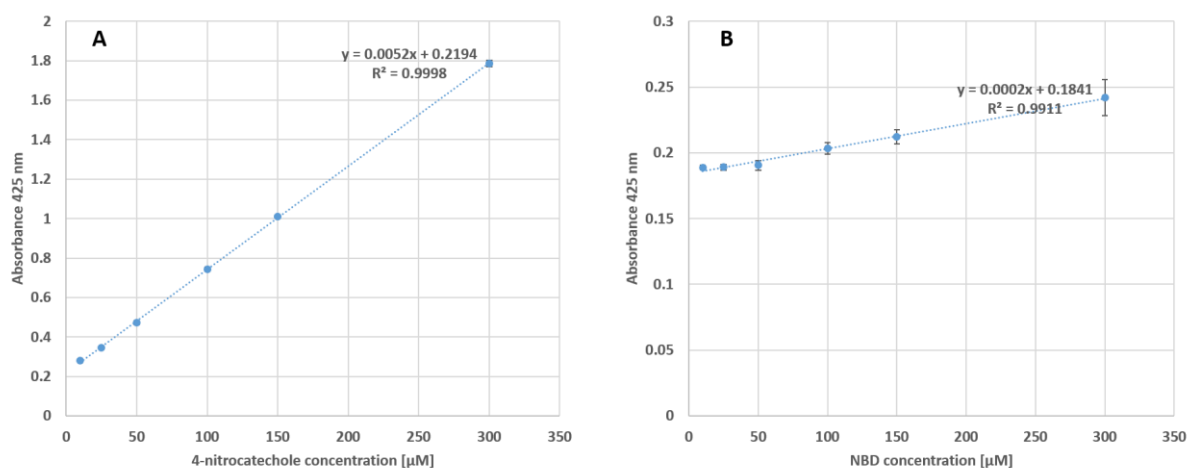
Figure S1. Calibration curves of A) 4-nitrocatechol and B) NBD under NBD-assay reaction conditions.

Table S4 Catalytic activity of purified *Mth*UPO variants for the hydroxylation of NBD.[a]

Catalyst	TOF [min⁻¹]	TON	Conversion [%]
*Mth*UPO WT	72	4340	29
*Mth*UPO L60M	83	5000	33
*Mth*UPO L60Q	94	5680	38
*Mth*UPO L60F	194	11610	77
*Mth*UPO F59Q	72	4360	29
*Mth*UPO E158D[b]	6	6560	44
*Mth*UPO L60F/F154I	90	5380	36
*Mth*UPO L60F/F154V	87	5200	35
*Mth*UPO I52/A57I	71	4290	29
*Mth*UPO S159N/A161F	75	4510	30
*Mth*UPO S159G/A161I	75	4520	30
*Mth*UPO L60F/S159G/A161I	198	11850	79
*Mth*UPO L60F/S159G/A161F	210	12590	84
*Mth*UPO F59Q/L60M/S159G/F154A	192	11540	77
*Mth*UPO F59Q/L60F/A161I	199	11950	80
*Mth*UPO F59Q/L60F/S159G	379	22760	76

TOF = turnover frequency, TON = turnover number, standard deviation < 3.2 %, [a] Reaction conditions: 20 nM *Mth*UPO variant, 300 µM NBD, 1 mM H_2O_2, 100 mM KPi buffer (pH 7), 5 % acetone (v/v), measurement conditions: absorbance was measured at 425 nm for one hour in triplicates, values were calculated with the corrected extinction coefficient of 10870 $M^{-1}*cm^{-1}$, [b] reaction time over night, [c] 10 nM *Mth*UPO variant.

NBD assay using altered peroxide sources. The purified *Mth*UPO wildtype, E158D, E158A were measured via NBD assay under the following conditions for one hour: 500 nM *Mth*UPO variant, 300 µM NBD, 1 mM cumene or *tert*-butyl hydroperoxide, 100 mM KPi buffer (pH 7), 5 % acetone (v/v). The product formation is illustrated in Figure S3.
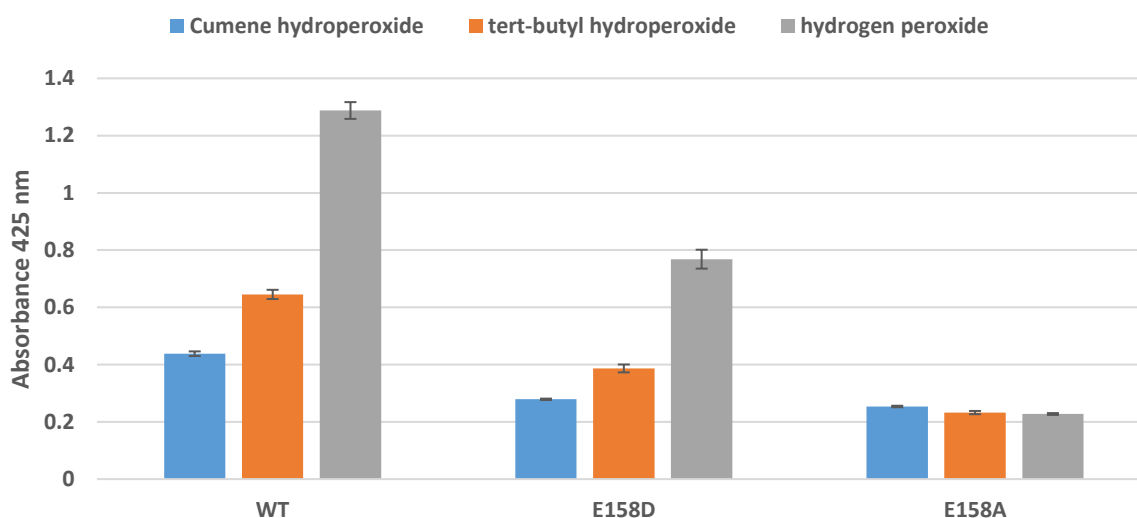
Figure S3 Comparison of the *Mth*UPO wildtype, E158D and E158A underutilisation of cumene, *tert*-butyl hydroperoxide and hydrogen peroxide within the NBD assay.

Determination of kinetic parameters. To determine the kinetic parameters K_m and k_{cat} the purified protein samples were utilised. The best performing variants from the NBD screening were thereby compared to each other: WT (20 nm), L60F (NBD variation: 5nM, H_2O_2 variation: 20 nM), E158D (NBD variation: 250 nM, H_2O_2 variation: 500 nM), L60F/S159G/A161F (20 nM), F59Q/L60M/S159G/F154A (20 nM) and F59Q/L60F/S159G (20 nM). For the K_m determination of the corresponding substrate, the second substrate was utilised in its saturation concentration (except for E158D). The velocity quantification was achieved in the linear range of the 4-nitrocatechol using the corrected $\varepsilon_{425nm, corr} = 10870$ $M^{-1}*cm^{-1}$ by applying automated pathlength correction in the microtiter plate. The nonlinear regression using the Michaelis-Menten model was performed with the aid of Sigmaplot (Version 14.0, Germany) yielding kinetic parameters v_{max}, K_m, R^2. The K_m and v_{max} value for E158D could not be determined with the nonlinear curve-fitting due to high required H_2O_2 concentration quenching the reaction. Therefore, a Lineweaver-Burk plot was utilised to apply the Michaelis-Menten model for linearisation. By using this model, it is not necessary to achieve the saturation concentration. The corresponding values and plots are illustrated in Figure S4 and S5.

Table S5 Kinetic parameters of the *Mth*UPO variants for H_2O_2 as a substrate.[a]

Catalyst	K_m [mM]	k_{cat} [s^{-1}]	k_{cat}/K_m [M^{-1}s^{-1}] x 10^4
*Mth*UPO WT	0.45 ± 0.10	5.2 ± 0.10	1.3
*Mth*UPO L60F	0.93 ± 0.150	1.1	0.1
*Mth*UPO E158D	19.7 ± 2.0	2.3	1.8
*Mth*UPO L60F/S159G/A161F	1.04 ± 0.10	93.4	9.0
*Mth*UPO F59Q/L60M/S159G/F154A	2.63 ± 0.55	70.6	2.7
*Mth*UPO F59Q/L60F/S159G	3.13 ± 0.43	24.7	0.8

Standard deviation < 20 %, TOF = turnover frequency, TON = turnover number, [a] Reaction conditions: 20 nM *Mth*UPO variant, 300 µM NBD, 1 mM H_2O_2, 100 mM KPi buffer (pH 7), 5 % acetone (v/v), measurement conditions: absorbance was measured at 425 nm for one hour in triplicates, values were calculated with the corrected extinction coefficient of 10870 M^{-1}*cm^{-1}, [b] reaction time over night, [c] 10 nM *Mth*UPO variant.
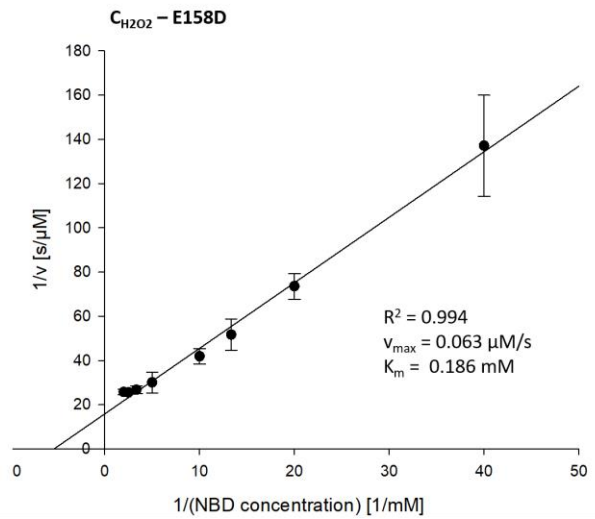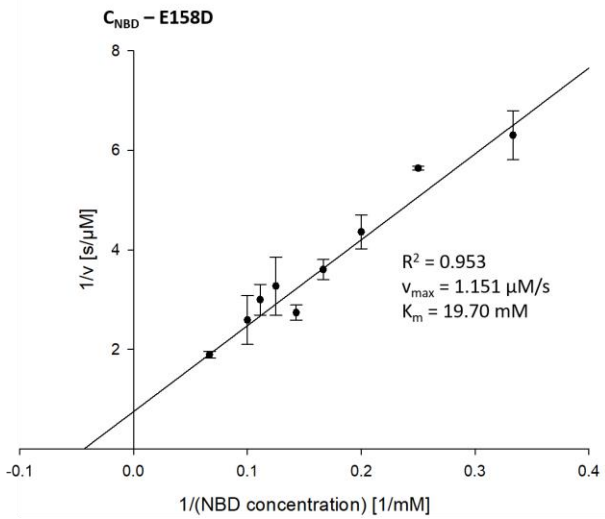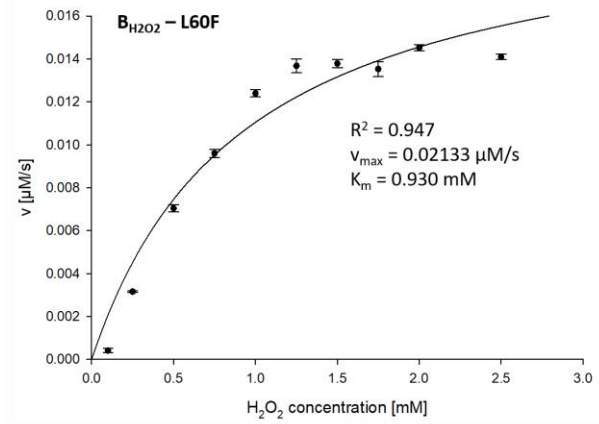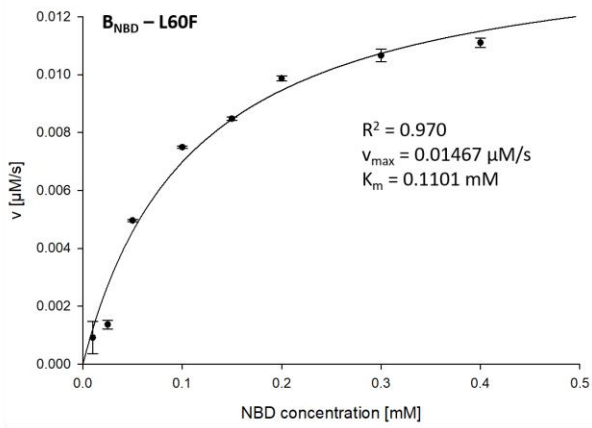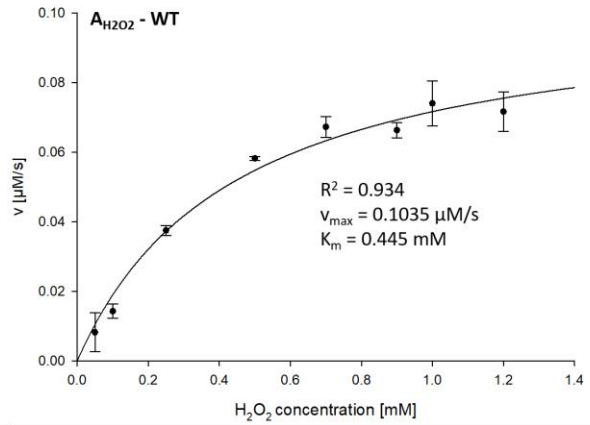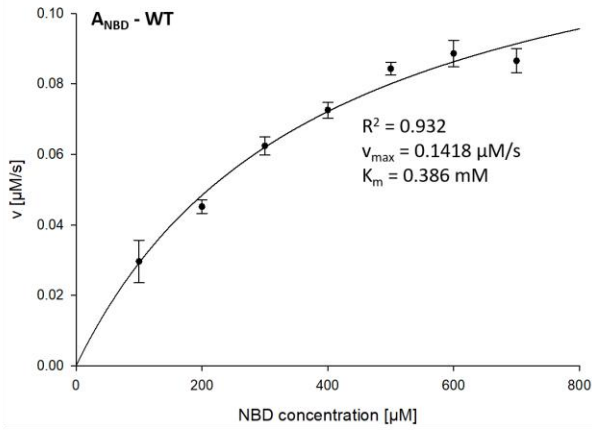
Figure S4. Michaelis-Menten plots of the best performing variants and their calculated kinetic parameters.

D_NBD – L60F/S159G/A161F

$R^2 = 0.985$
$v_{max} = 2.645 \ \mu M/s$
$K_m = 0.422 \ mM$

D_H2O2 – L60F/S159G/A161F

$R^2 = 0.977$
$v_{max} = 1.869 \ \mu M/s$
$K_m = 1.041 \ mM$

E_NBD – F59Q/L60M/S159G/F154A

$R^2 = 0.929$
$v_{max} = 1.188 \ \mu M/s$
$K_m = 0.296 \ mM$

E_H2O2 – F59Q/L60M/S159G/F154A

$R^2 = 0.940$
$v_{max} = 1.412 \ \mu M/s$
$K_m = 2.636 mM$

F_NBD – F59Q/L60F/S159G

$R^2 = 0.961$
$v_{max} = 0.940 \ \mu M/s$
$K_m = 0.303 \ mM$

F_H2O2 – F59Q/L60F/S159G

$R^2 = 0.954$
$v_{max} = 0.494 \ \mu M/s$
$K_m = 3.126 \ mM$
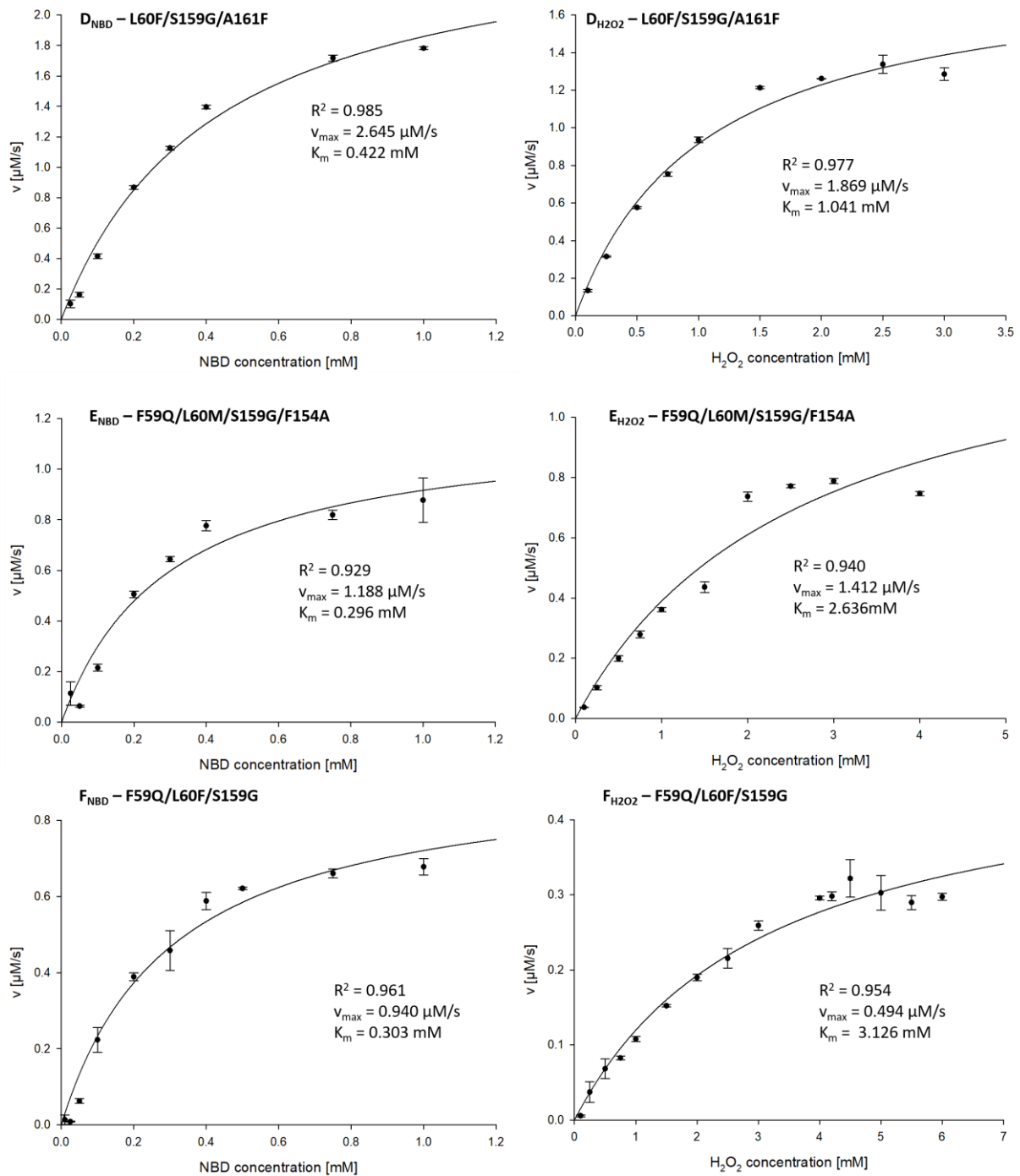
Figure S5. Michaelis-Menten plots of the best performing variants and their calculated kinetic parameters.

Thermostability measurements. Thermostability measurements of the purified enzymes were performed by Differential Scanning fluorimetry (DSF) on a Prometheus NT.48 nanoDSF instrument (NanoTemper Technologies GmbH, München, DE) in storage buffer (100 mM Tris-HCl pH 7.0). Approximately 10 μL of sample volume were loaded into a Prometheus NT.48 High Sensitivity Capillary (NanoTemper Technologies GmbH). Protein unfolding was subsequently monitored by following the ratio of intrinsic protein tyrosine and tryptophan fluorescence at 350 nm to 330 nm over time, increasing the temperature from 20 °C to 95 °C with a heating ramp of 1 °C per minute. The melting temperature corresponds to the maximum of the first derivative of the 350/330 nm ratio. All measurements were performed at least in triplicates.[10]

Initial oxyfunctionalisation comparison between *Mth*UPO and the best performing variants from the NBD assay. For the identification of the best performing variant towards the hydroxylation of naphthalene and its derivatives, an initial approach with direct addition of H_2O_2 (without a syringe pump setup) was chosen. The following conditions were adjusted: 500 nM *Mth*UPO variant, 1 mM H_2O_2, 1 mM of the corresponding substrate (see Fig. S6-S8), 100 mM KPi pH 7, 1 h at 25 °C. The reaction was quenched by the addition of 400 μl *n*-hexane with benzyl alcohol as the internal standard. Corresponding samples were analysed in the GC/MS scan mode, and the products were identified by library comparison. For quantification and product confirmation, the syringe pump setup was utilised.
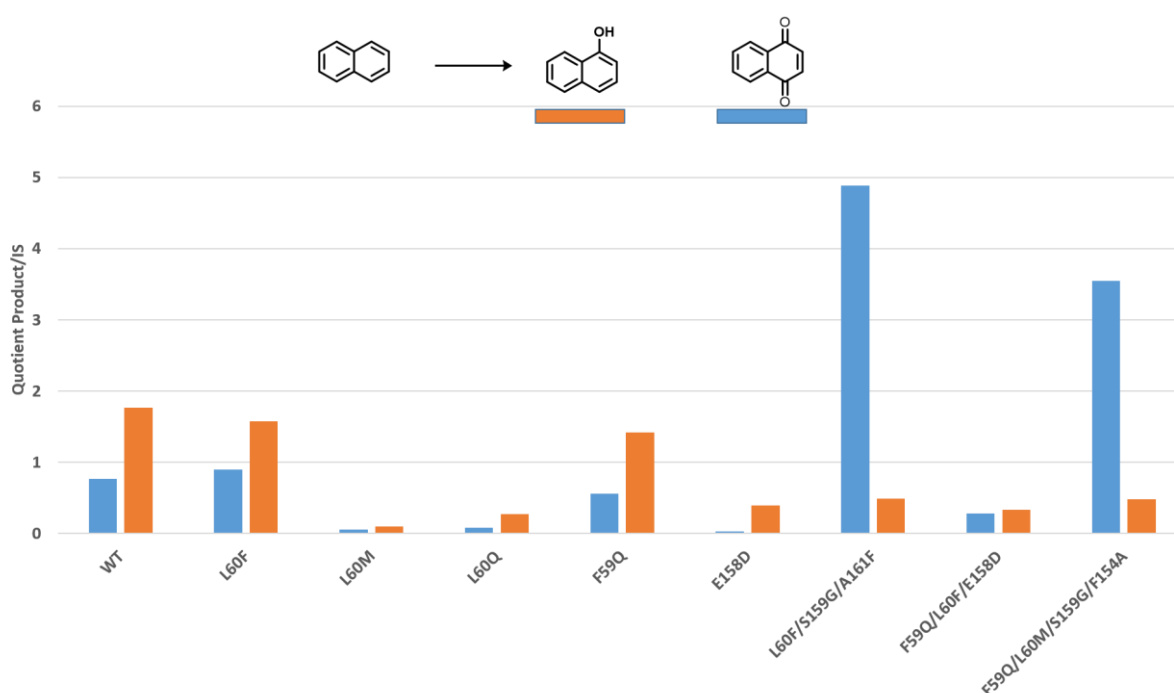
Figure S6 Product distribution of *Mth*UPO variants towards the bioconversion of naphthalene.
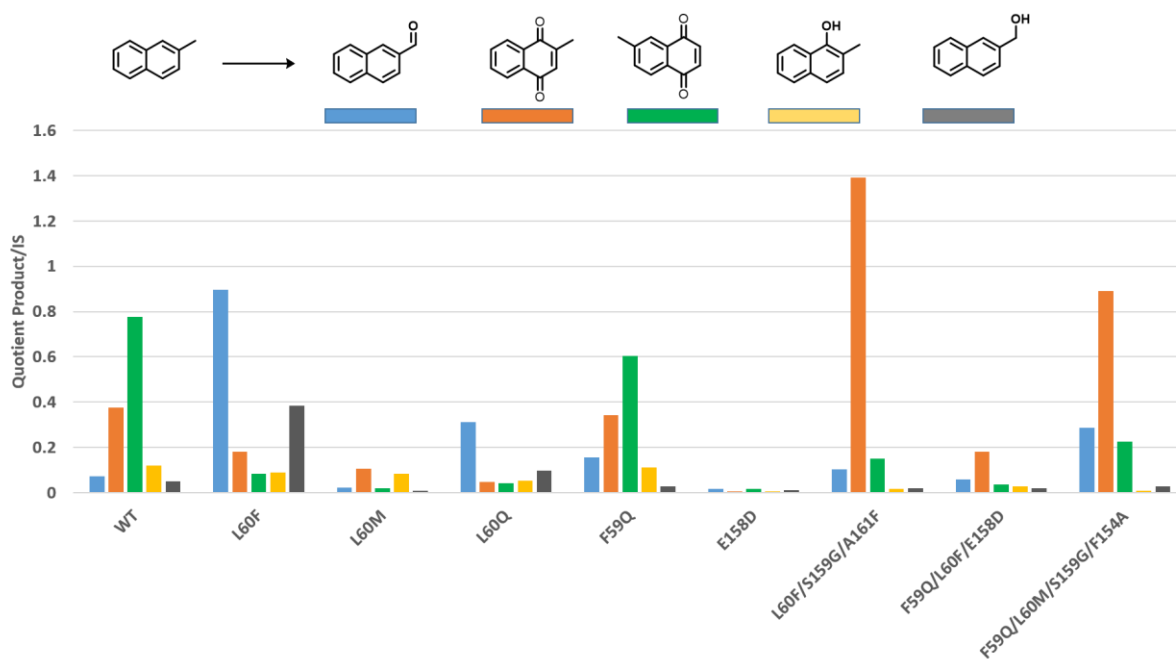
Figure S7 Product distribution of *Mth*UPO variants towards the bioconversion of 2-methylnaphthalene.
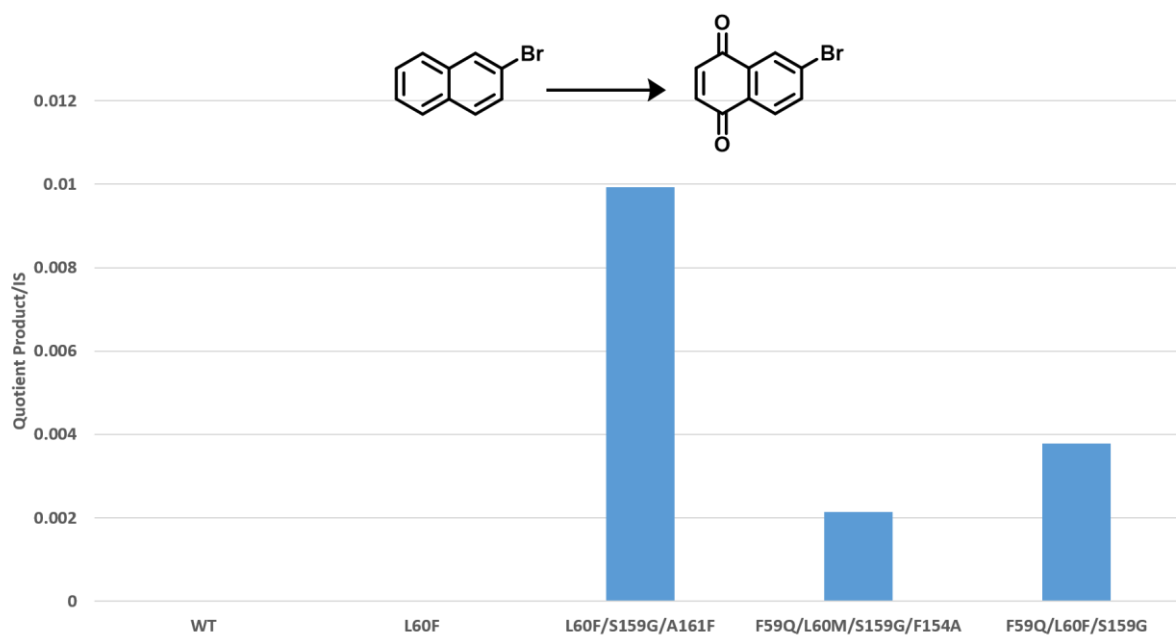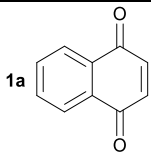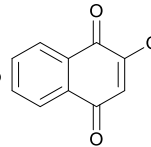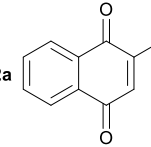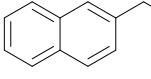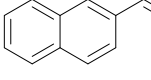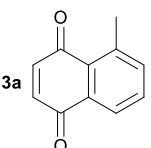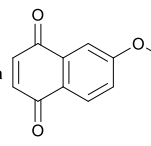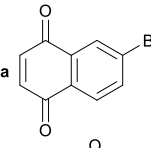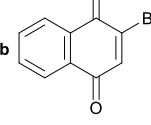
Figure S8 Product distribution of *Mth*UPO variants towards the bioconversion of 2-bromonaphthalene.

Bioconversion of naphthalene and naphthalene derivatives by a syringe pump system. With a syringe pump system, the product formation could be increased compared to the direct addition of H_2O_2 (initial approach). The conditions were adjusted to each reaction setup and are shown in table S6. For the H_2O_2 addition, a programmable syringe pump from Chemyx Inc. (Model: Fusion 101R, Stafford, US) was utilised. The reaction and addition were performed under continuous stirring at room temperature. The extraction was accomplished by the addition of 400 µl of EtOAc (containing 1 mM 4-octanone as internal standard) at the end of the reaction. After 30 s of vortexing, the organic layer was transferred for GC/MS analysis.
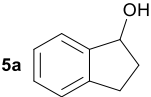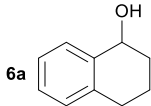
Table S6 Overview of reaction conditions for the hydroxylation of naphthalene and naphthalene derivatives.

Product	Catalyst	Reaction conditions
1a	MthUPO F59Q/L60F/S159G	1 mM naphthalene, 4 mM H_2O_2, 50 nM catalyst, 100 mM KPi pH 7, 5 % (v/v) acetone, V_{total} = 400 µl, syringe pumpe addition of a 16 mM H_2O_2 stock solution (100 µl stock solution, 100 µl/h), 30 min additional stirring
1b	*Mth*UPO F59Q/L60F/S159G	1 mM naphthalene, 5 mM H_2O_2, 500 nM catalyst, 100 mM KPi pH 7, 5 % (v/v) acetone, V_{total} = 400 µl, syringe pumpe addition of a 10 mM H_2O_2 stock solution (200 µl stock solution, 200 µl/h), additional stirring overnight
2a	*Mth*UPO L60F/S159G/A161F	1 mM 2-methylnaphthalene, 4 mM H_2O_2, 100 nM catalyst, 100 mM KPi pH 7, 5 % (v/v) acetone, V_{total} = 400 µl, syringe pumpe addition of a 8 mM H_2O_2 stock solution (200 µl stock solution, 100 µl/h), 1 additional stirring
2b 2c	*Mth*UPO L60F	1 mM 2-methylnaphthalene, 4 mM H_2O_2, 500 nM catalyst, 100 mM KPi pH 7, 5 % (v/v) acetone, V_{total} = 400 µl, syringe pumpe addition of a 8 mM H_2O_2 stock solution (200 µl stock solution, 100 µl/h), 30 min additional stirring
3a	*Mth*UPO L60F/S159G/A161F	1 mM 1-methylnaphthalene, 4 mM H_2O_2, 500 nM catalyst, 100 mM KPi pH 7, 5 % (v/v) acetone, V_{total} = 400 µl, syringe pumpe addition of a 8 mM H_2O_2 stock solution (200 µl stock solution, 200 µl/h)
4a	*Mth*UPO L60F/S159G/A161F	1 mM 2-methoxynaphthalene, 4 mM H_2O_2, 500 nM catalyst, 100 mM KPi pH 7, 5 % (v/v) acetone, V_{total} = 400 µl, syringe pumpe addition of a 8 mM H_2O_2 stock solution (200 µl stock solution, 200 µl/h), additional stirring overnight
5a 5b	*Mth*UPO L60F/S159G/A161F	1 mM 2-bromonaphthalene, 4 mM H_2O_2, 500 nM catalyst, 100 mM KPi pH 7, 5 % (v/v) acetone, V_{total} = 400 µl, syringe pumpe addition of a 8 mM H_2O_2 stock solution (200 µl stock solution, 200 µl/h), 1 h additional stirring

Standard deviation <5.2 %.

Bioconversion of indane and 1,2,3,4-tetrahydronaphthalene. For the bioconversion of these substrates, no syringe pump system was needed. The reaction was started by addition of H_2O_2. The hydroxylation and overoxidation to the ketone were performed under the following conditions: 100 nM L60F, 1 mM indane, 1 mM H_2O_2, 100 mM KPi pH 7, 5 % acetone (*v/v*), 1 h at 25 °C in triplicates. For the hydroxylation of 1,2,3,4-tetrahydronaphthalene, the conditions were adapted from the indane bioconversion. The best performing overoxidation of 1,2,3,4-tetrahydronaphthalene to the *α*-tetralone was accomplished by *Mth*UPO F59Q/L60M/S159G/F154A with a final concentration of 2 mM H_2O_2. The extraction was achieved by the addition of 400 µl of EtOAc (containing 1 mM 4-octanone as internal standard) at the end of the reaction. After 30 s of vortexing, the organic layer was transferred for GC/MS analysis. Quantification of the products were carried out on an SH-Rxi-5Sil MS column under the previously mentioned GC/MS conditions. The *ee* determination was carried out on a Lipodex E column.

Table S7 Enantioselectivity of the evolved variants towards the hydroxylation of indane and 1,2,3,4-tetrahydronaphthalene.

Product	Catalyst	Configuration	*ee* [%]
	*Mth*UPO WT	*R*	84
	*Mth*UPO L60F	*R*	95
5a	*Mth*UPO L60F/S159G/A161F	*S*	14
	*Mth*UPO F59Q/L60M/S159G/F154A	*S*	14
	*Mth*UPO F59Q/L60F/S159G	*R*	91
	*Mth*UPO WT	*R*	45
	*Mth*UPO L60F	*R*	74
6a	*Mth*UPO L60F/S159G/A161F	*R*	6
	*Mth*UPO F59Q/L60M/S159G/F154A	*S*	19
	*Mth*UPO F59Q/L60F/S159G	*R*	50

4. Computational simulations and calculations

Homology model and Molecular Dynamics (MD) simulations. Homology model for *Mth*UPO structure (283 AA) has been constructed based on the solved crystal structure of the UPO from *Marasmius rotula* (*Mro*UPO, PDB: 5FUJ, 29% identity and 41% similarity) using the homology server SWISS-MODEL.[11] The resulting homology model has been further refined with extensive MD simulations based on 5 independent replicas of 1000 ns each, accumulating a total of 5 μs of simulation time. MD simulations in explicit water were performed using the AMBER18 package.[12] Parameters for the different substrates (5-Nitro-1,3-benzodioxole (NBD), 2-methylnaphthalene, 1-methylnaphthalene, 2-methoxynaphthalene, and indane) were generated within the antechamber[13] module in AMBER18 package using the general AMBER force field (gaff),[14] with partial charges set to fit the electrostatic potential generated at the B3LYP/6-31G(d) level by the RESP model.[15] The charges were calculated according to the Merz–Singh–Kollman scheme[16-17] using the Gaussian 09 package. Parameters for the heme compound I (Cpd I) and the axial Cys were taken from reference.[18] The protein was solvated in a pre-equilibrated cubic box with a 12-Å buffer of TIP3P[19] water molecules using the AMBER18 leap module, resulting in the addition of ~17,500 solvent molecules. The systems were neutralised by addition of explicit counterions (Na^+ and Cl^-). All subsequent calculations were done using the AMBER force field 14 Stony Brook (ff14SB).[20] A two-stage geometry optimisation approach was performed. The first stage minimises the positions of solvent molecules and ions imposing positional restraints on solute by a harmonic potential with a force constant of 500 kcal mol^{-1} Å^{-2}, and the second stage is an unrestrained minimisation of all the atoms in the simulation cell. The systems were gently heated using six 50 ps steps, incrementing the temperature by 50 K for each step (0–300 K) under constant-volume and periodic-boundary conditions. Water molecules were treated with the SHAKE algorithm such that the angle between the hydrogen atoms was kept fixed. Long-range electrostatic effects were modelled using the particle-mesh-Ewald method.[21] An 8 Å cutoff was applied to Lennard–Jones and electrostatic interactions. Harmonic restraints of 30 kcal·mol^{-1} were applied to the solute, and the Langevin equilibration scheme was used to control and equalise the temperature. The time step was kept at 1 fs during the heating stages, allowing potential inhomogeneities to self-adjust. Each system was then equilibrated for 2 ns with a 2 fs time step at a constant pressure of 1 atm and temperature of 300 K without restraints. Once the systems were equilibrated in the NPT ensemble, production trajectories were then run under the NVT ensemble and periodic-boundary conditions. In particular, a total of 5000 ns for wildtype *Mth*UPO were accumulated from 5 independent replicas (1000 ns each); 4000 ns were accumulated for E158D variant from independent 5 replicas (3x 1000 ns + 2x 500 ns); and 1500 ns were accumulated for H88A and H88A/E158D variants from 3 independent replicas (500 ns each). Finally, a total of 750 ns from 3 independent replicas were accumulated for L60F and L60F/S159G/A161F variants (250 each). An extensive conformational sampling based on longer and more MD replicas for wildtype *Mth*UPO has been carried out to refine the initial homology model. The more visited conformation of wildtype *Mth*UPO (based on backbone clustering analysis) has been used as a starting point to computationally generate the other variants. Trajectories were processed and analyzed using the cpptraj[22] module from Ambertools utilities.

Docking and protocol used for substrate bound MD simulations. Docking calculations were performed using AutoDock Vina.[23] The most populated clusters (based on backbone clustering analysis) obtained from MD simulations carried out in the apo state were used, and docking predictions were then utilised as starting points for substrate-bound MD simulations. To avoid substrate diffusion outside the enzyme active site and to sample catalytically competent binding poses, a 100 kcal·mol−1·Å−2 restraint is applied when the distances between the center of mass of the substrate and the O atom of the Fe=O in Cpd I are greater than 6Å (for indane) or 6.7Å (for naphthalene derivatives). No restraints were used in NBD substrate-bound simulations. Same protocol for MD

simulations described above has been employed, accumulating a total of 300 ns of production trajectories from 3 independent replicas for all substrate-bound studied.

Quantum Mechanics (QM) calculations. Density Functional Theory (DFT) calculations were carried out using Gaussian09.[24] Geometry optimisations and frequency calculations were performed using (U)B3LYP[25-27] functional with 6-31G(d) basis set on all atoms. The stationary points were verified as minima by a vibrational frequency analysis. Enthalpies and entropies were calculated for 1 atm and 298.15 K. Single point energy calculations were performed using the functional (U)B3LYP with the Def2TZVP basis set on all atoms, and within the CPCM polarisable conductor model (dichloromethane, $\varepsilon = 8.9$)[28-29] to have an estimation of the dielectric permittivity in the enzyme active site.[30] Bond Dissociation Energies (BDEs) were calculated as the standard enthalpic change of the following process at 298 K (Indane → Indane$^{\bullet}$ + H$^{\bullet}$), which provides an estimation of the strength of the C-H bond under study. Different electronic states (singlet close-shell for indane and doublet for both radical species) have been considered.
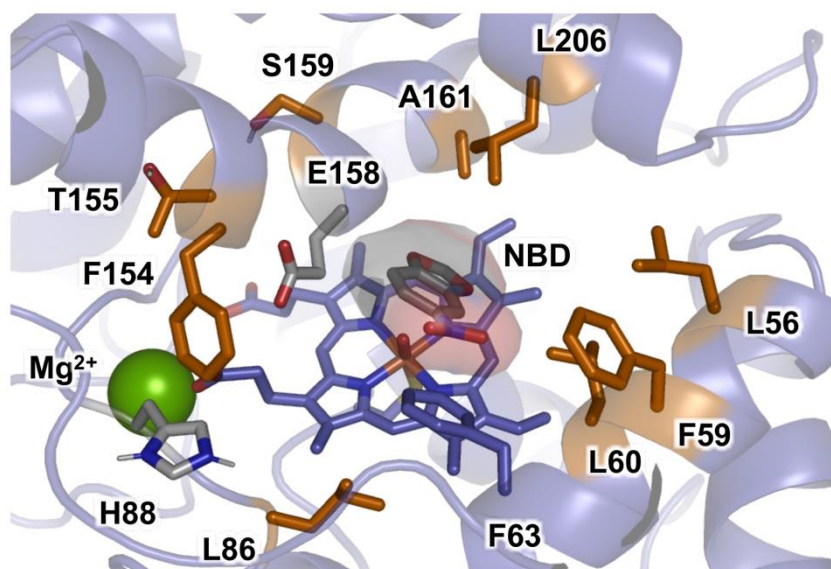
Figure S9 *Mth*UPO active site arrangement with NBD substrate bound obtained from MD simulations based on the generated homology model. Important active site and catalytic residues (in grey) are shown in stick format. The nine positions initially randomised in round 1 are highlighted in orange (L56, F59, L60, L86, F154, T155, S159, A161, L206).
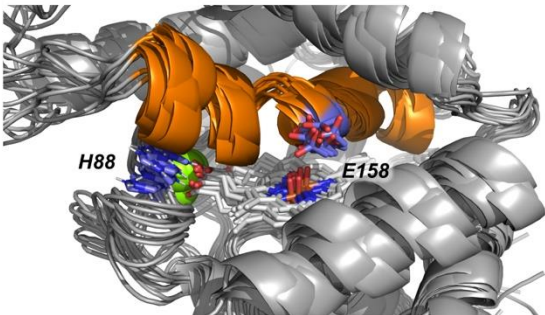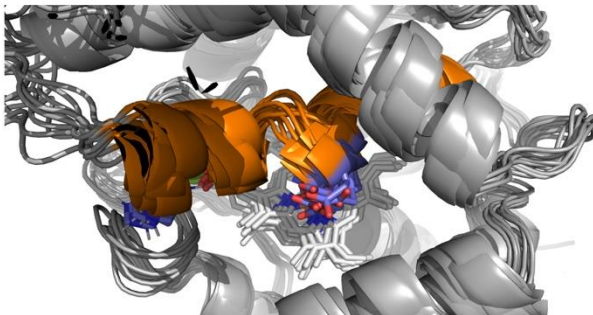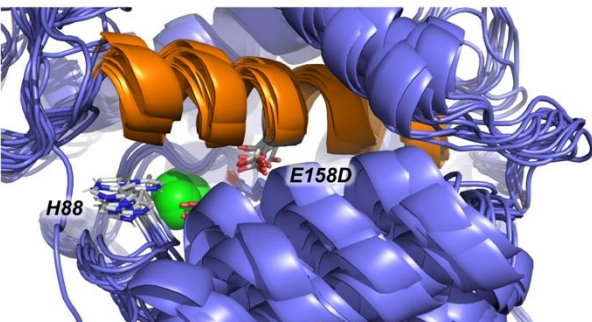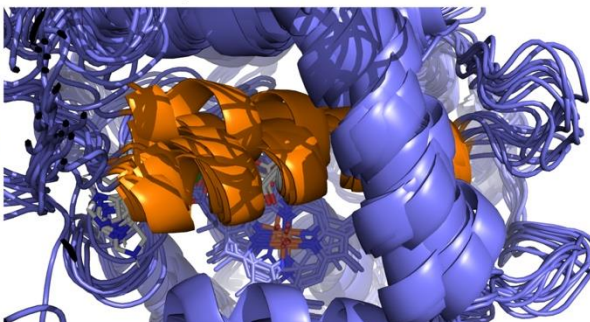
Figure S10 Overlay of 10 representative snapshots from MD simulation trajectories of wildtype and E158D variant. E158-H88 and E158D-H88 catalytic residues are highlighted in purple and gray, respectively, and α-helix-7 is highlighted in orange. MD simulations indicate that E158D side chain and α-helix-7 in E158D variant are more rigid than E158 α-helix-7 in wildtype enzyme that exhibit more flexibility (see also Fig. S11).

A)

Replica 1

Replica 2

Replica 3

Replica 4

Replica 5
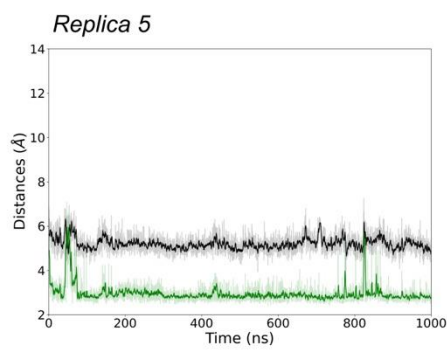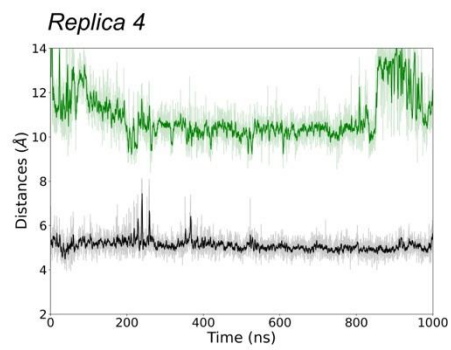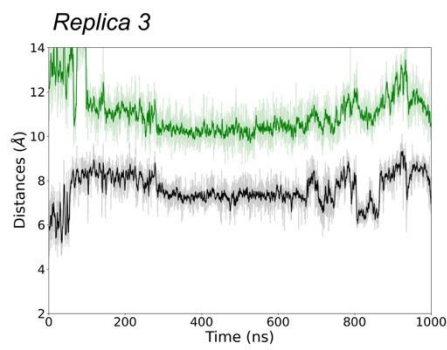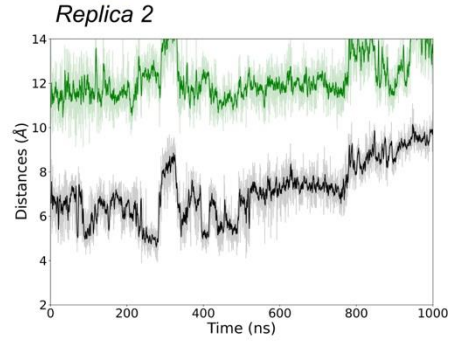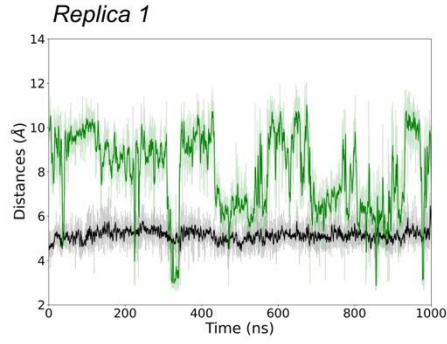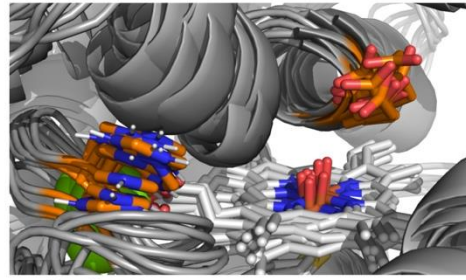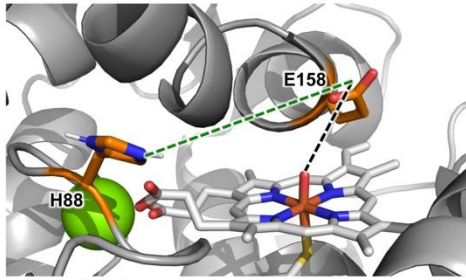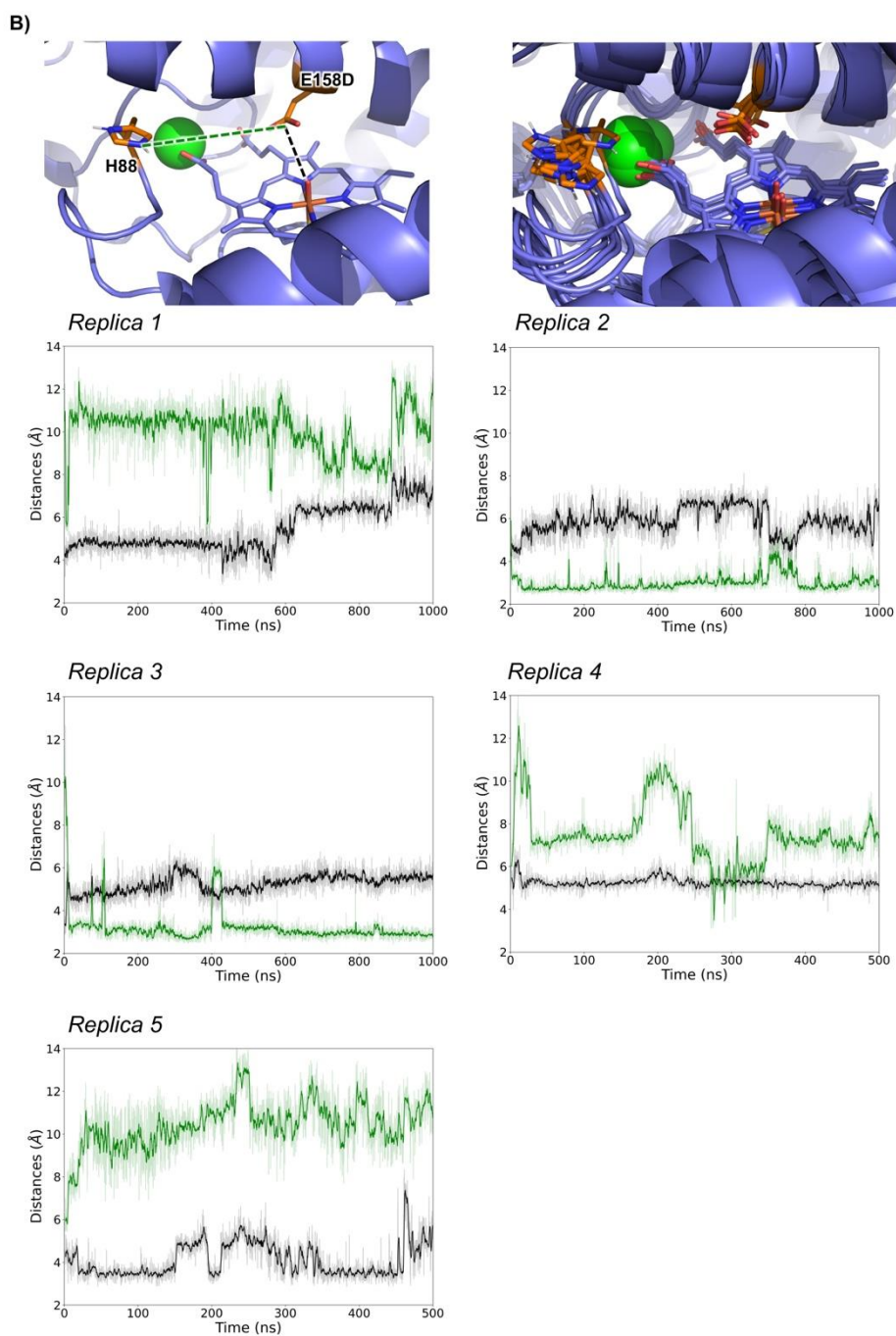
Figure S11 Key catalytic distances explored by **A**) wildtype and **B**) E158D *Mth*UPO variant analyzed from MD simulations. Black lines describe the evolution of the distance between O(Fe=O) and the carboxylic acid from E158/E158D sidechains along time; green lines describe the evolution of the distance between H88 ε-nitrogen and the carboxylic acid from E158/E158D sidechains along time. An overlay of representative snapshots from MD simulations are provided, where catalytic residues H88 and E158(D) are highlighted in orange. Catalytic distances (short distances between E158(D) and H88 and Fe=O) are explored in both systems.

A)
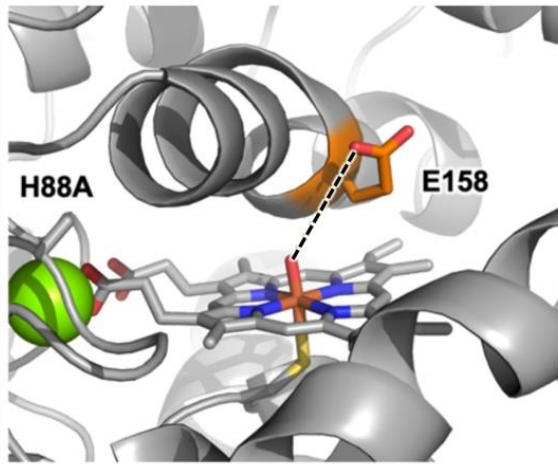
Replica 1
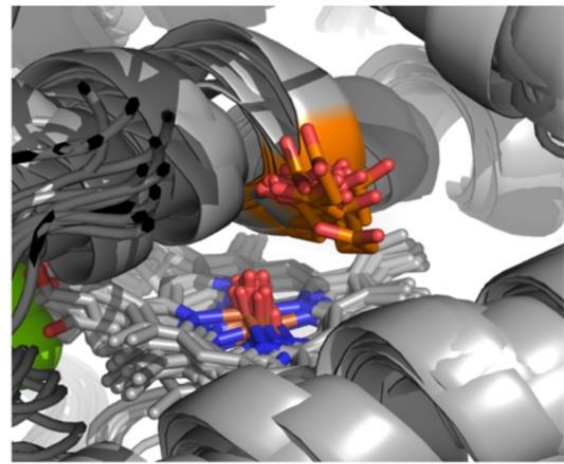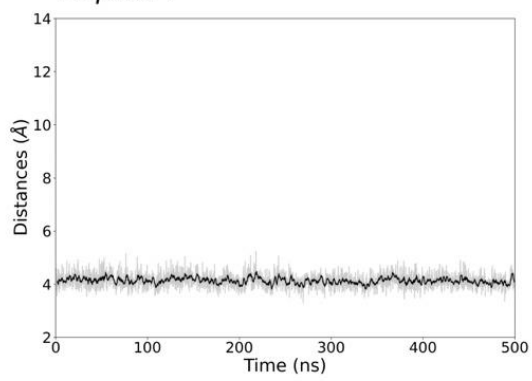
Replica 2
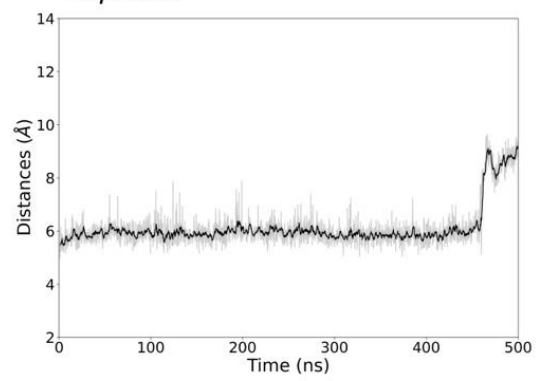
Replica 3
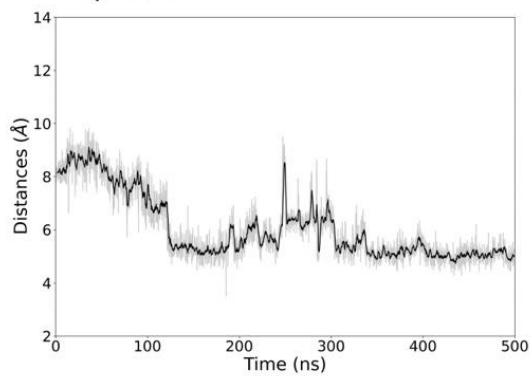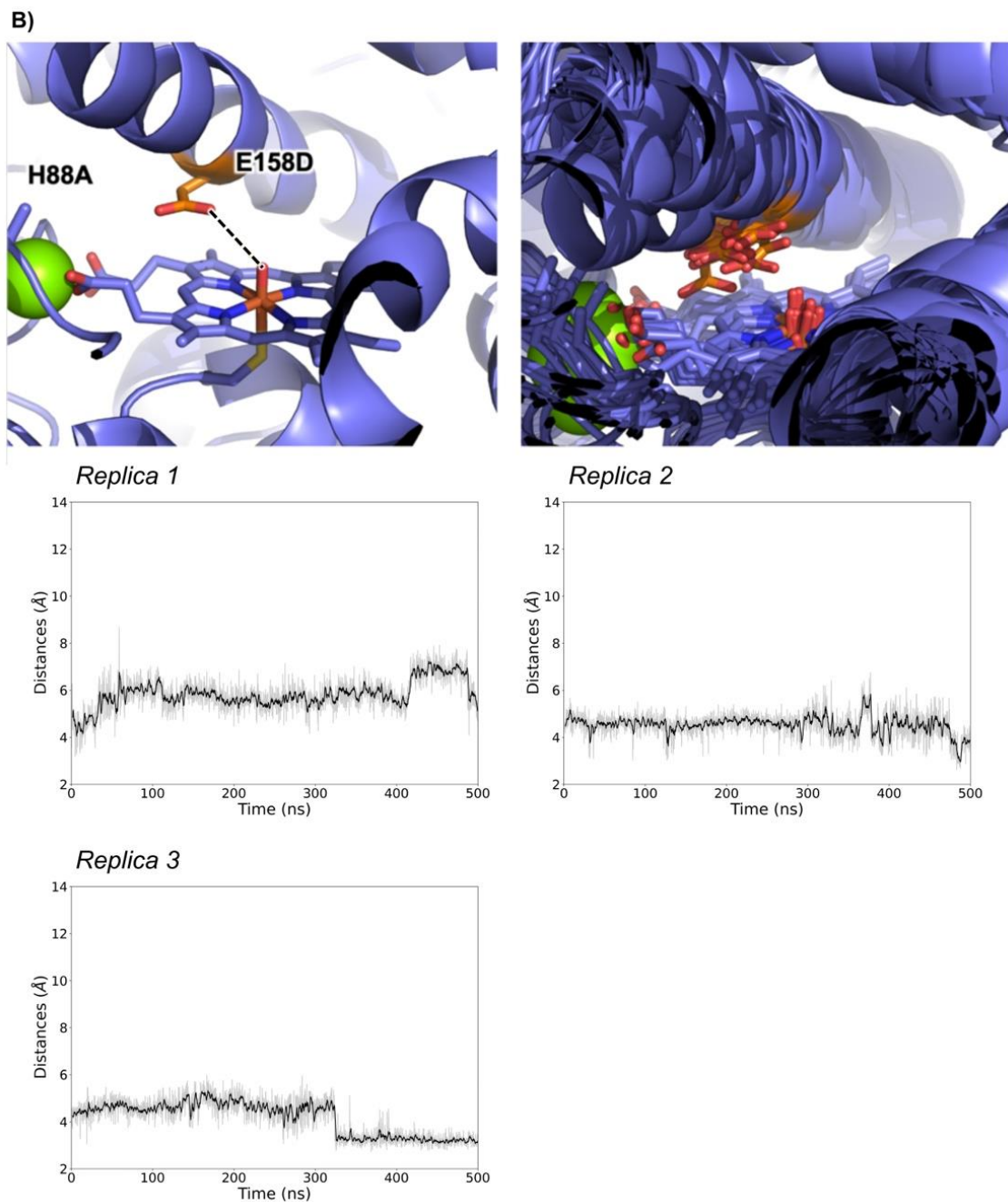
Figure S12 Relevant catalytic distances explored by **A**) H88 and **B**) H88/E158D *Mth*UPO variants analyzed from MD simulations. Black lines describe the evolution of the distance between O(Fe=O) and the carboxylic acid from E158/E158D sidechains along time. An overlay of representative snapshots from MD simulations are provided, where catalytic residues E158(D) are highlighted in orange. Catalytic distances (short distances between E158(D) and Fe=O) are explored in both systems, being the less flexible E158D sidechain better preorganised for catalysis.
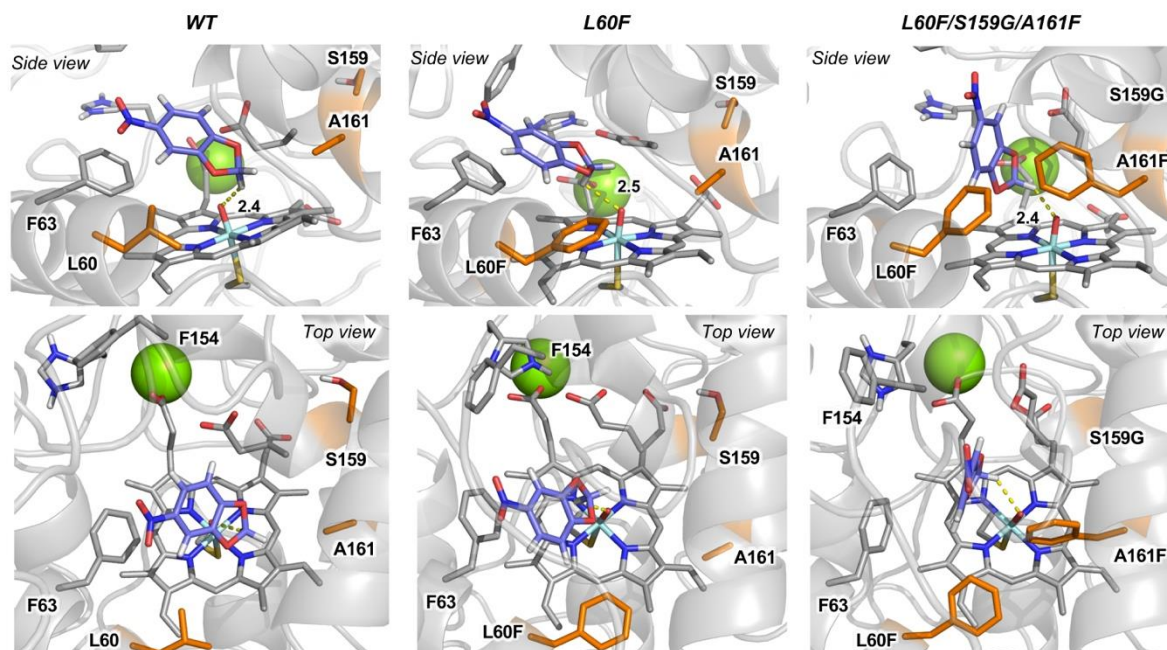
Figure S13. Evolution of NBD catalytically relevant binding modes in wildtype, L60F and L60F/S159G/A161F variants as observed from MD simulations (see **Figure S14**). Mutated positions are highlighted in orange. NBD explores substantially different near attack conformations (NACs) in each variant due to the new introduced mutations. L60F displaces NBD from a more buried binding pose in wildtype, to a new binding mode that increases the aromatic interactions with residues F63 and newly introduced L60F. Finally, A161F and S159G mutations led to a significantly reduced active site that forces NBD to explore a new binding mode, perpendicular to the heme, that facilitates its interaction with the catalytic Fe=O species (see also **Figure S14**).

A)
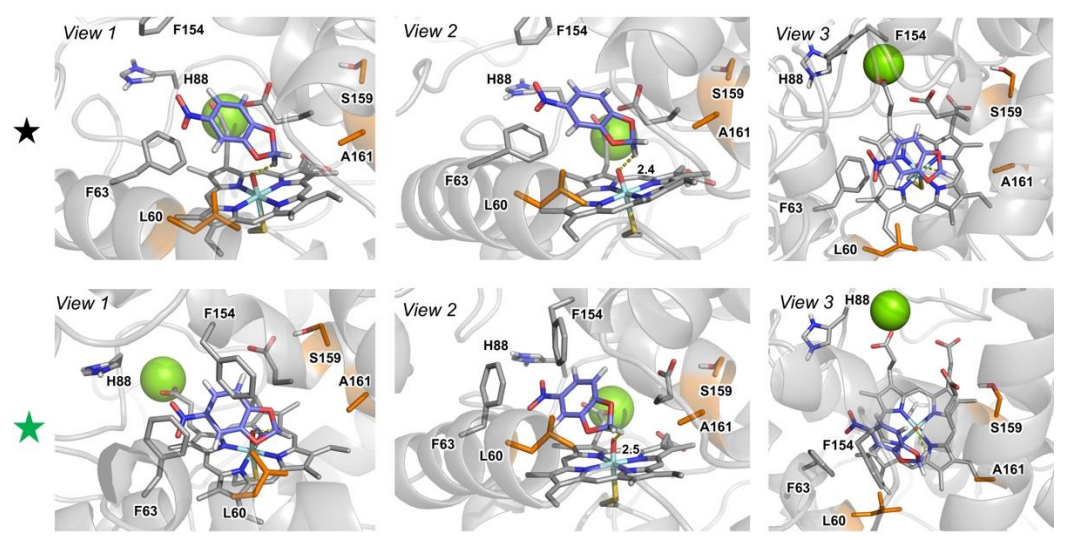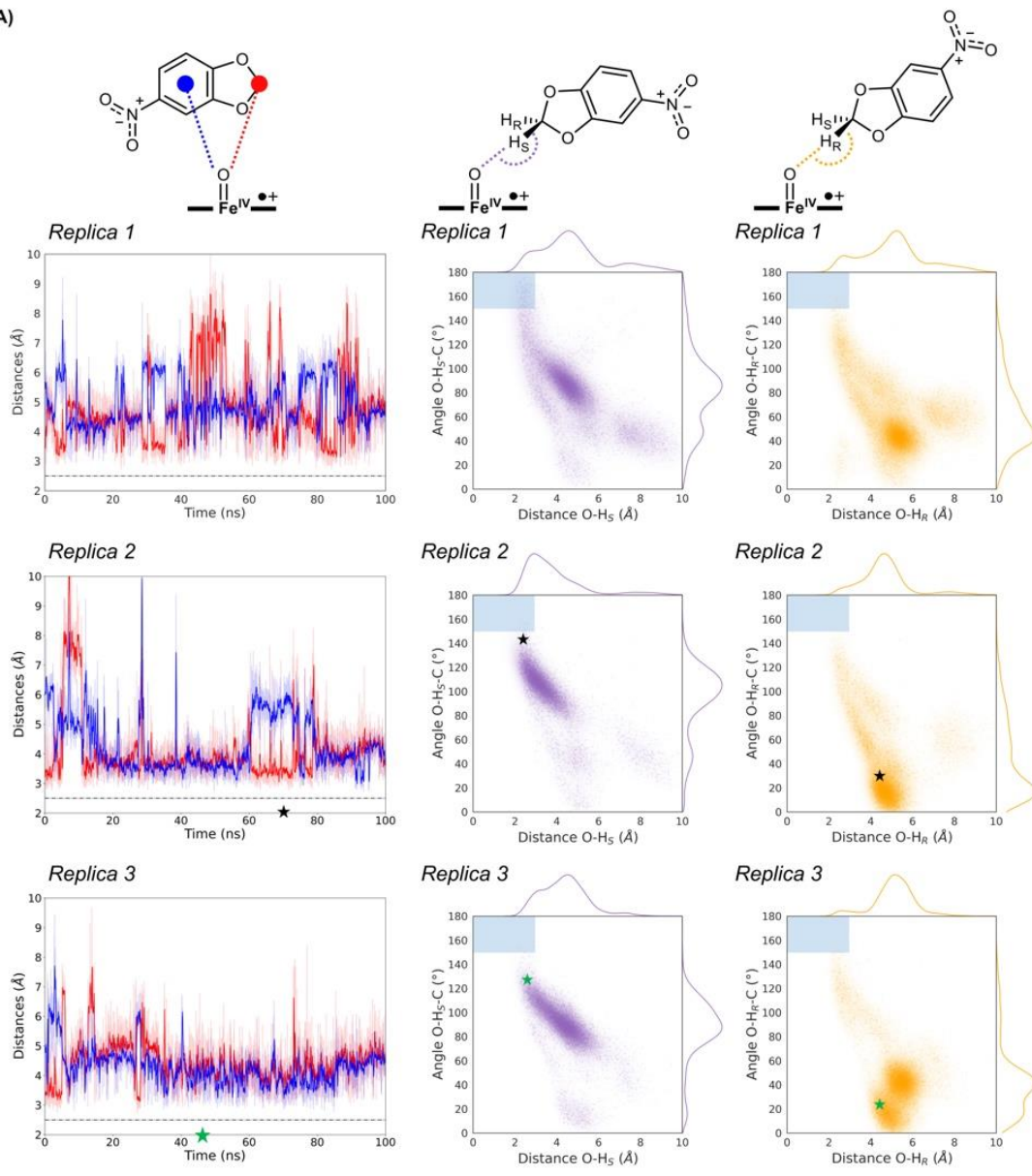
B)
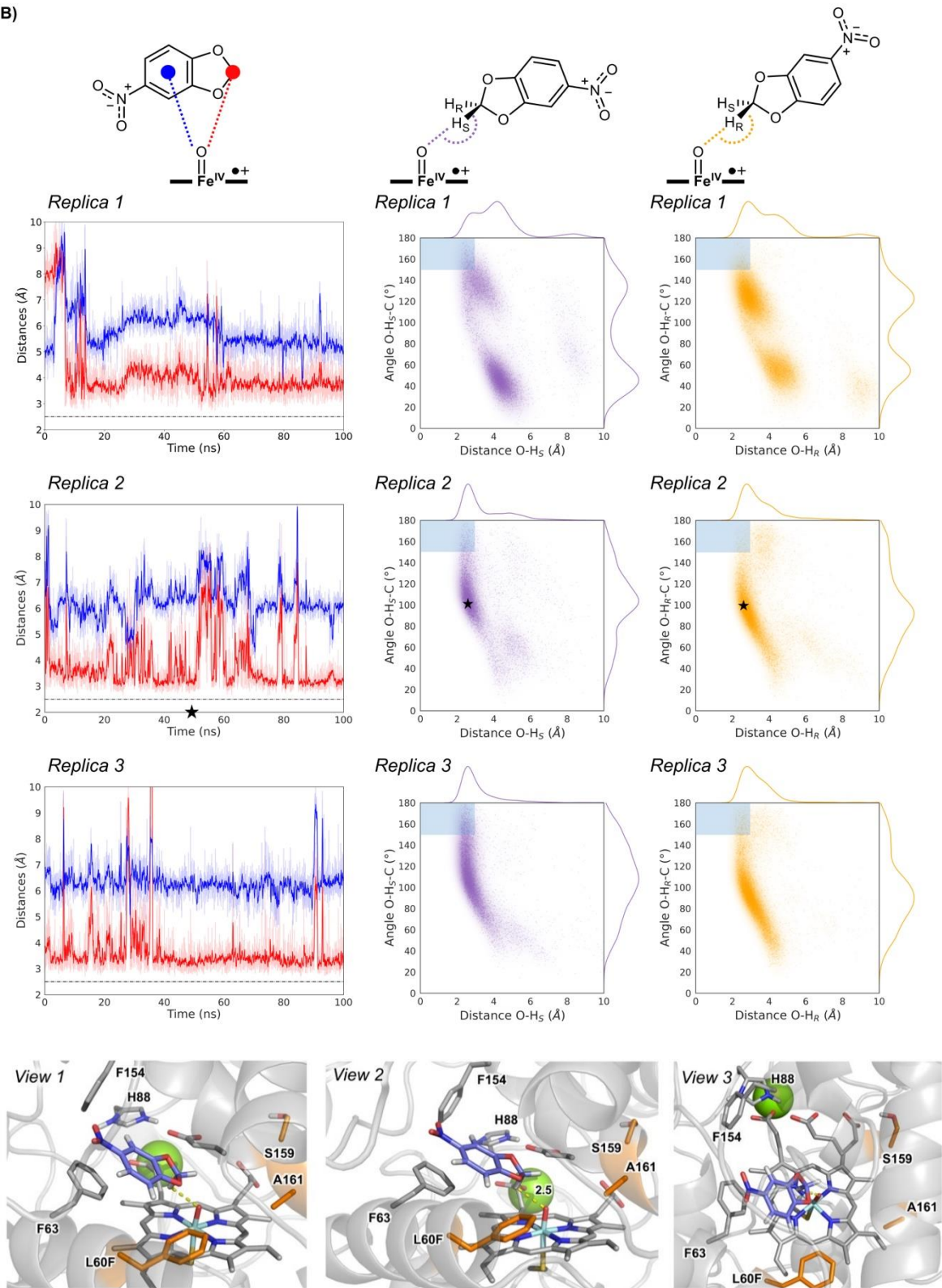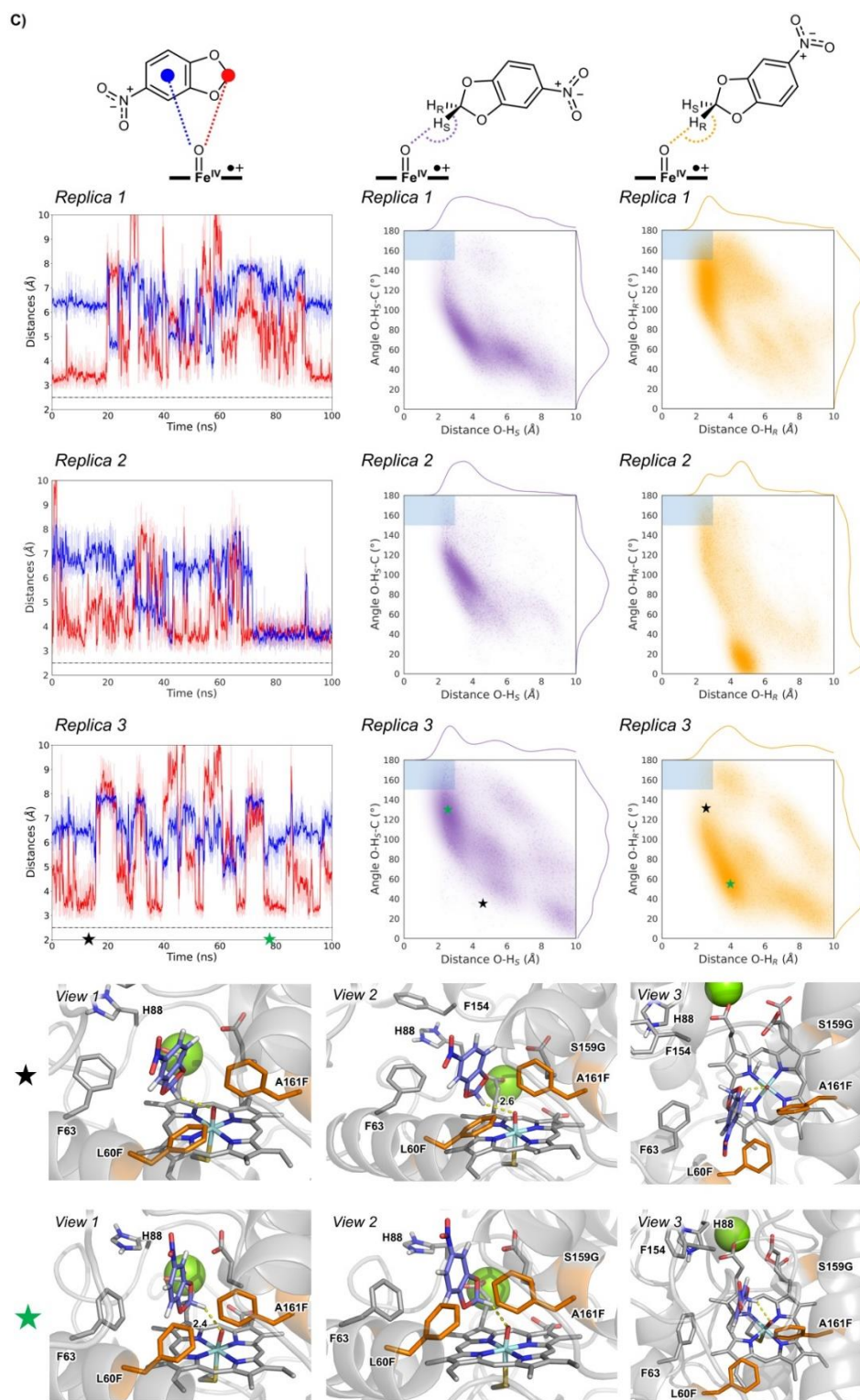
Figure S14. Analysis of NBD binding modes through three independent MD replicas in: **A**) wildytpe; **B**) L60F variant; and **C**) L60F/S159G/A161F variant. Key distances relevant for hydroxylation and aromatic oxidation are monitored along MD simulations, as described in the schemes. Fe=O – H(CH) distance and angle of attack (O-H-C) are used as geometric parameters to characterise near attack conformations for effective C-H hydroxylation in heat maps. Representative snapshots from MD trajectories (highlighted with a "star" symbol) that describe reactive near attack conformations explored during MDs are shown. Distances and angles are given in angstroms (Å) and degrees (º), respectively. Angle vs. distance heat maps show that NBD explores more catalytically competent poses in L60F/S159G/A161F and L60F variants than in the WT, in line with the higher activity experimentally observed. The differences in the NBD binding poses in the three variants are discussed in **Figure SX13**.
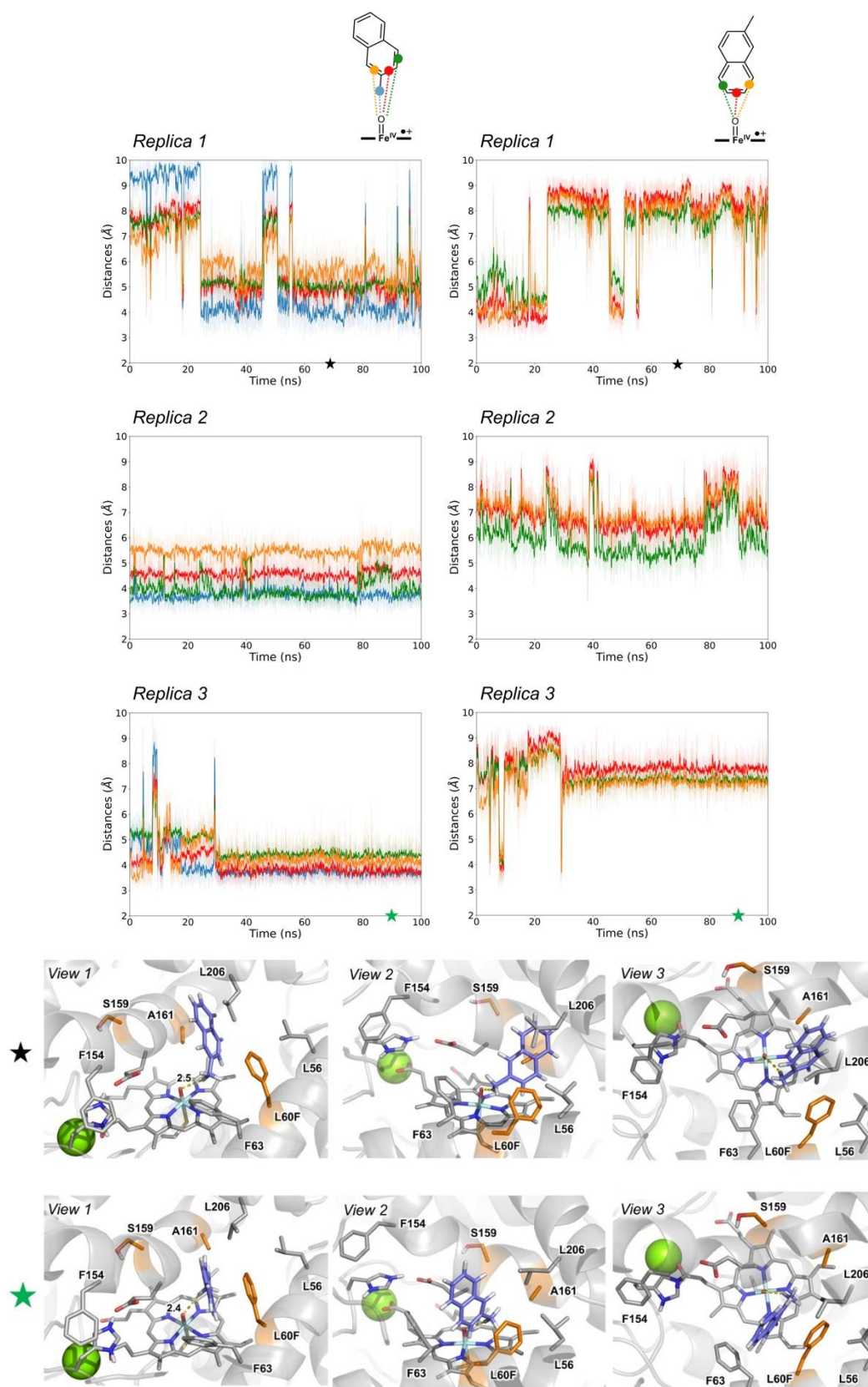
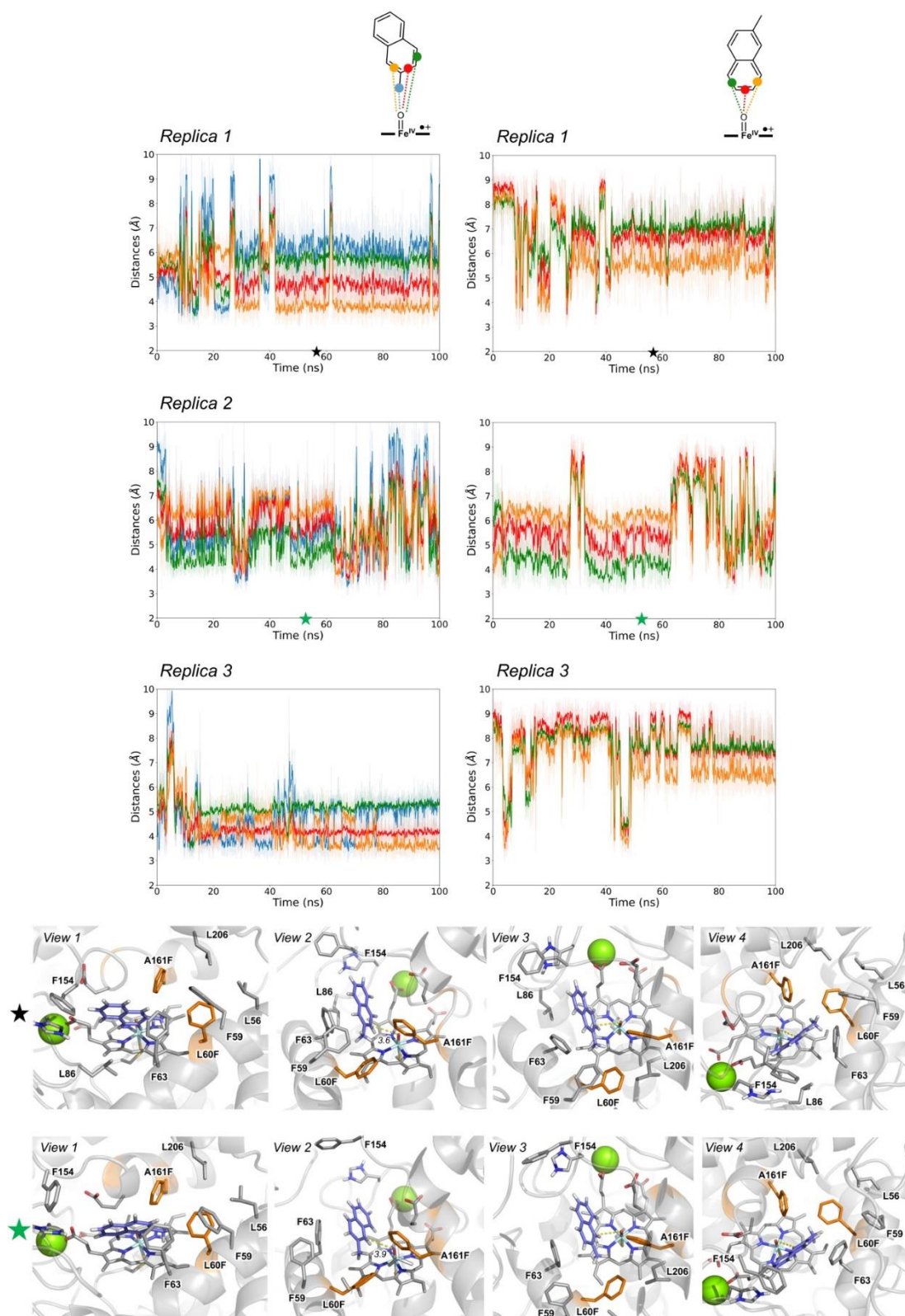**Figure S15** Analysis of 2-methylnaphthalene binding modes through three independent MD replicas in L60F variant. Key distances relevant for hydroxylation and aromatic oxidations are monitored along MD simulations, as described in the schemes. Representative snapshots from MD trajectories (highlighted with a "star" symbol) that describe reactive near attack conformations explored during MDs are shown. Distances and angles are given in angstroms (Å) and degrees (º), respectively. 2-methylnaphthalene bound in L60F variant predominantly explores catalytically relevant binding poses in which the 2-methyl group is placed in a near attack conformation respect to the Fe=O active species. Differences observed for naphthalene derivatives binding modes in different MthUPO variants are discussed in Figure S17.

Figure S16 Analysis of 2-methylnaphthalene binding modes through three independent MD replicas in L60F/S159G/A161F variant. Key distances relevant for hydroxylation and aromatic oxidations are monitored along MD simulations, as described in the schemes. Representative snapshots from MD trajectories (highlighted with a "star" symbol) that describe reactive near attack conformations explored during MDs are shown. Distances and angles are given in angstroms (Å) and degrees (º), respectively. 2-methylnaphthalene bound in L60F/S159G/A161F variant predominantly explores catalytically relevant binding poses in which the substituted aromatic ring is placed in a near attack conformation respect to the Fe=O active species. Differences observed for naphthalene derivatives binding modes in different MthUPO variants are discussed in Figure S17.
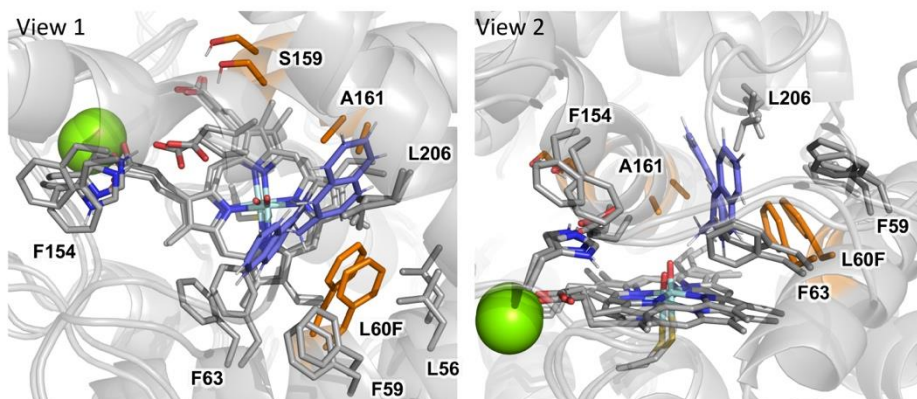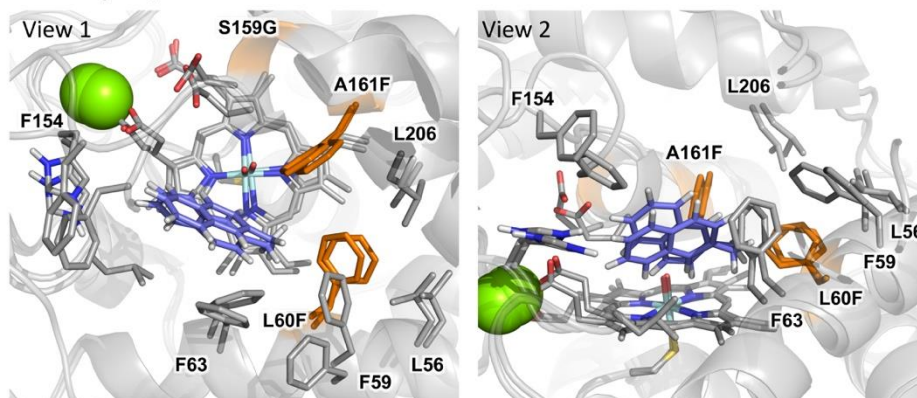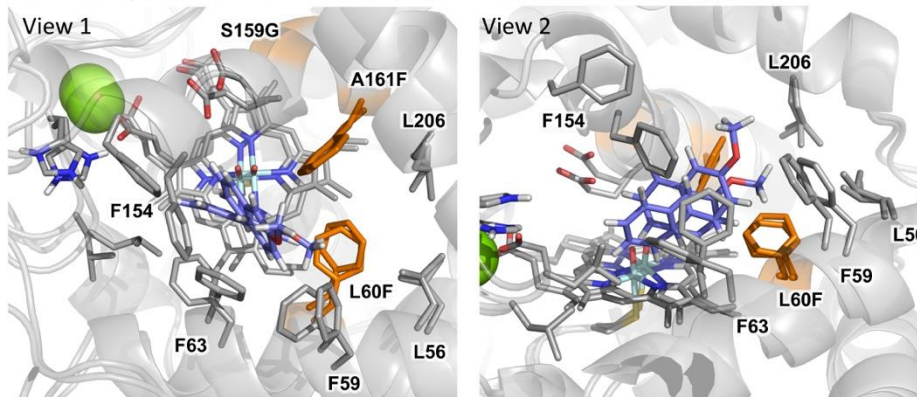
Figure S17 Differences in catalytically relevant binding modes of 2-substituted naphthalene derivatives (2-methlynaphthalene and 2-methoxynaphthalene) in L60F and L60F/S159G/A161F variants as observed from MD simulations (see **Figure S15 S16 and S18**). Mutated positions are highlighted in orange. 2-methlynaphthalene explores substantially different near attack conformations (NACs) in each variant due to the new introduced mutations. In L60F, 2-methlynaphthalene explores catalytically competent poses in which the 2-methyl group is suitable to directly interacts with Fe=O active species. On the other hand, in L60F/S159G/A161F variant, the substrate is displaced form the former binding position to a new one that resembles the binding mode observed by NBD in this triple mutant (Figure S13). This new binding mode, induced by the presence of bulky A161F mutation, allows the direct interaction between the substituted aromatic ring of 2-methlynaphthalene and the catalytic Fe=O species (see Figure S15). When 2-methoxynaphthalene is bound in L60F/S159G/A161F variant, it occupies the same binding position as 2-methyl derivative. However, because the more bulkier 2-methoxy group, the substrate slightly rotates and preferentially explores catalytically relevant conformations in which the 2-methoxy group is placed far from the heme group and the unsubstituted aromatic ring is placed closer to the Fe=O. Because of this reorientation, the regioselectivity of the oxidation reaction changes from preferential functionalisation at the substituted aromatic ring in 2-methylnaphtalene to the oxidation at the unsubstituted one in 2-methoxynaphthalene when L60F/S159G/A161F variant is used.
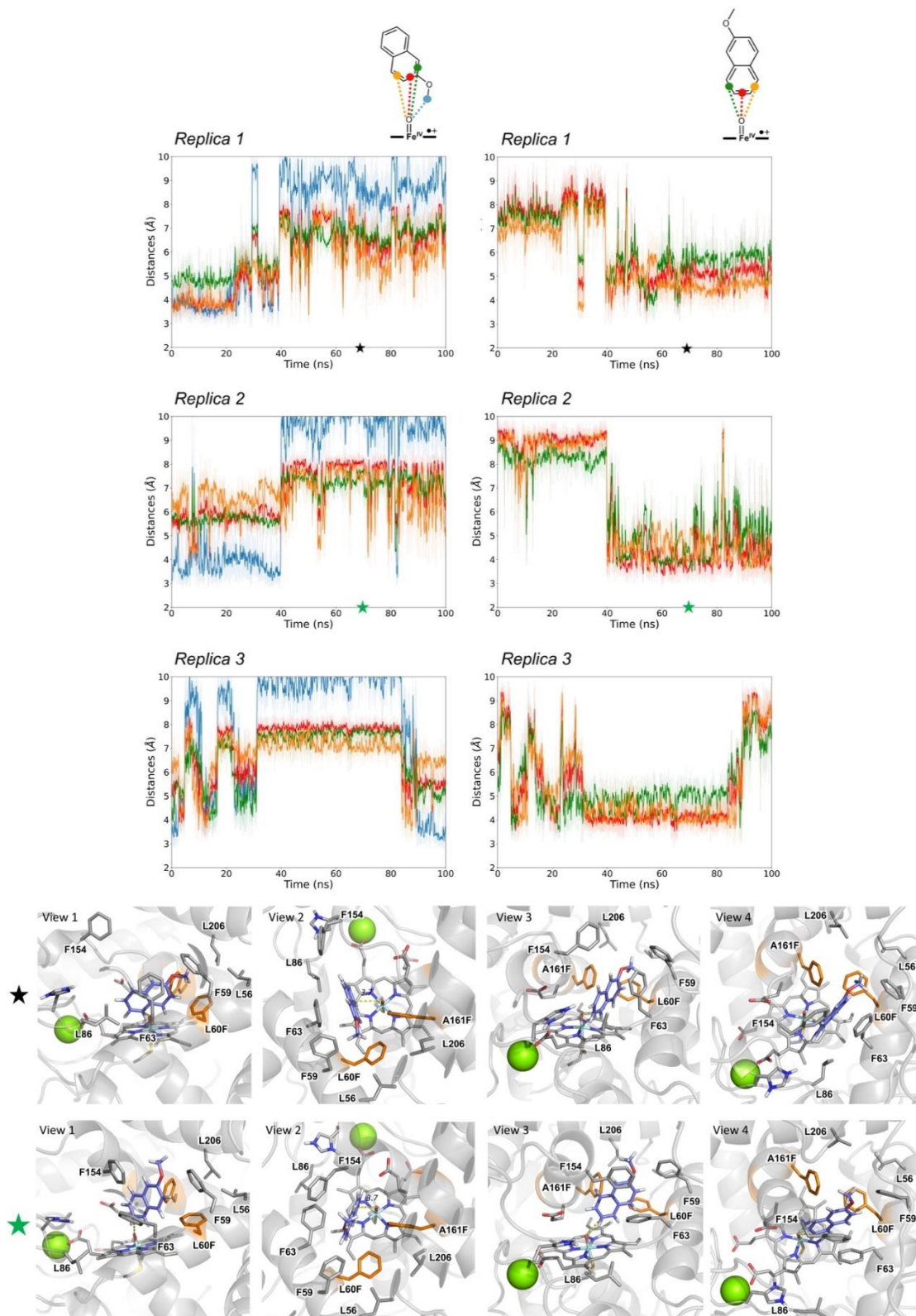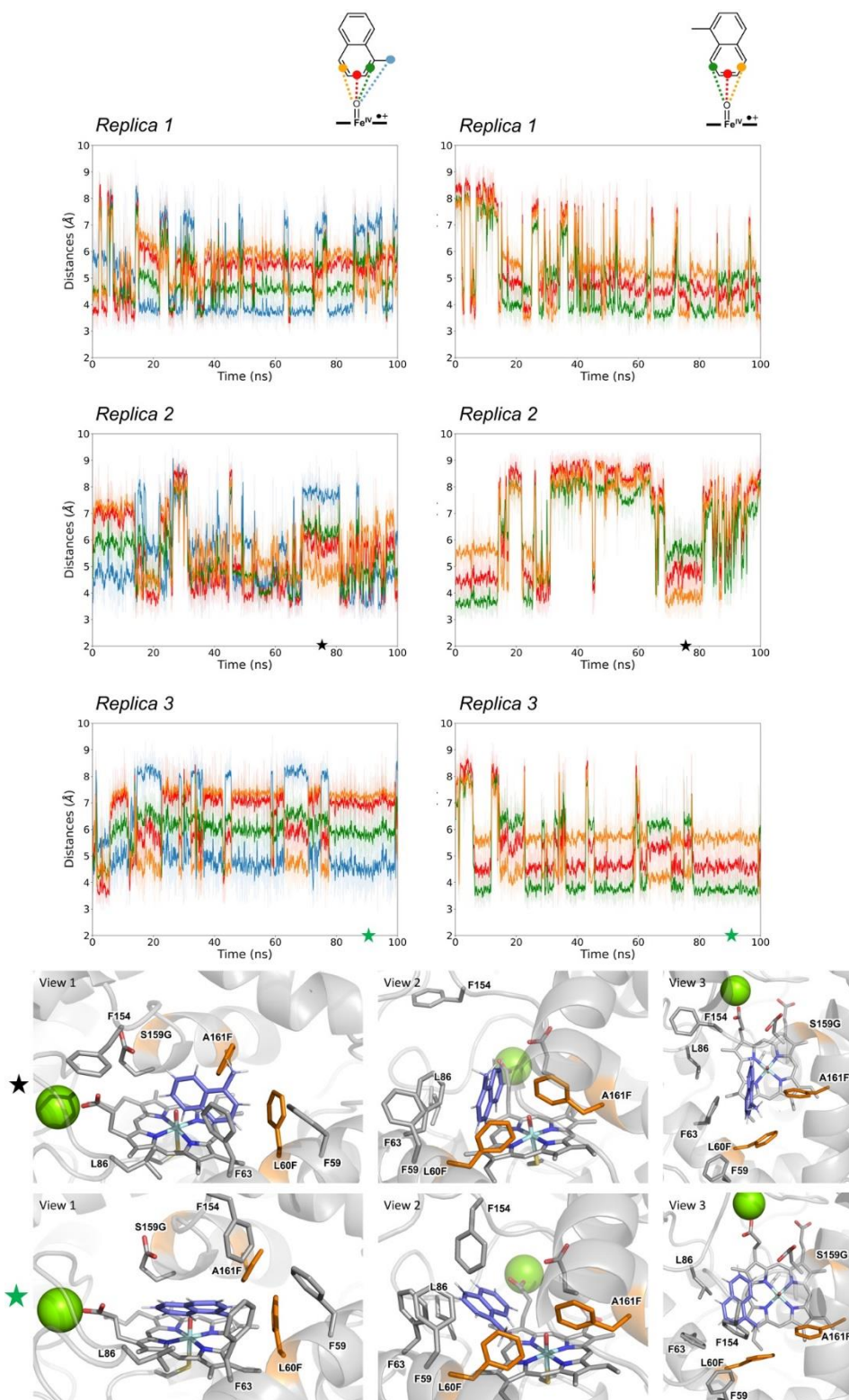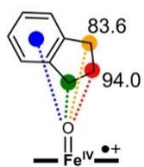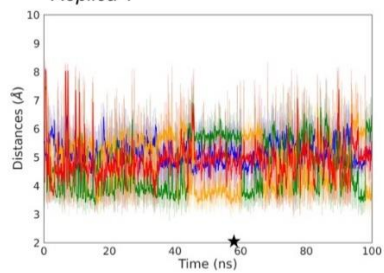
Figure S18. Analysis of 2-methoxynaphthalene binding modes through three independent MD replicas in L60F/S159G/A161F variant. Key distances relevant for hydroxylation and aromatic oxidations are monitored along MD simulations, as described in the schemes. Representative snapshots from MD trajectories (highlighted with a "star" symbol) that describe reactive near attack conformations explored during MDs are shown. Distances and angles are given in angstroms (Å) and degrees (º), respectively.2-methoxynaphthalene bound in L60F/S159G/A161F variant predominantly explores catalytically relevant binding poses in which the unsubstituted aromatic ring is placed in a near attack conformation respect to the Fe=O active species. Differences observed for naphthalene derivatives binding modes in different MthUPO variants are discussed in **Figure S17**.

Figure S19. Analysis of 1-methylnaphthalene binding modes through three independent MD replicas in L60F/S159G/A161F variant. Key distances relevant for hydroxylation and aromatic oxidations are monitored along MD simulations, as described in the schemes. Representative snapshots from MD trajectories (highlighted with a "star" symbol) that describe reactive near attack conformations explored during MDs are shown. Distances and angles are given in angstroms (Å) and degrees (º), respectively. 1-methylnaphthalene bound in L60F/S159G/A161F variant predominantly explores catalytically relevant binding poses in which the unsubstituted aromatic ring is placed in a near attack conformation respect to the Fe=O active species.
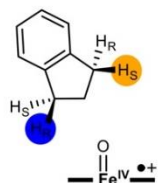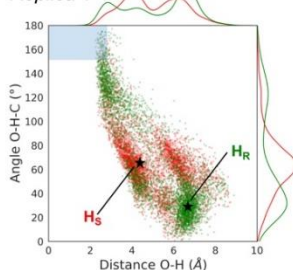
A)

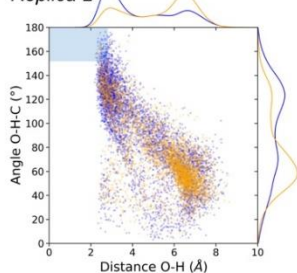*Replica 1*
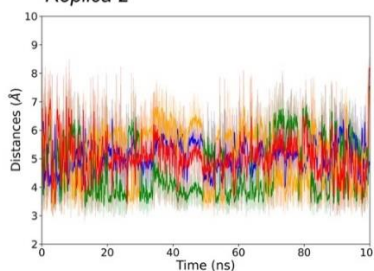
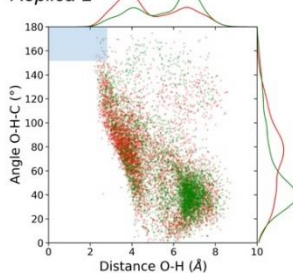*Replica 2*   *Replica 3*
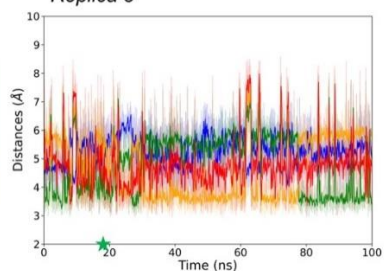
B)
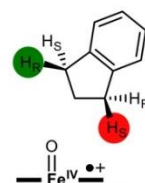
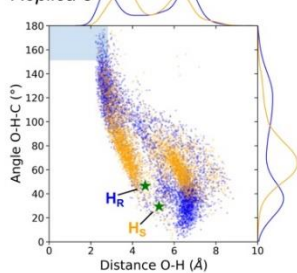*Replica 1*   *Replica 1*
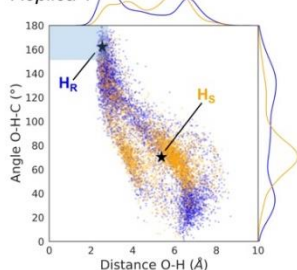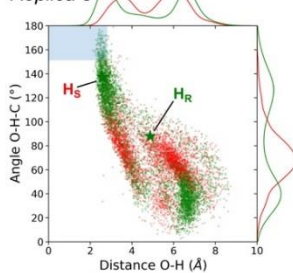
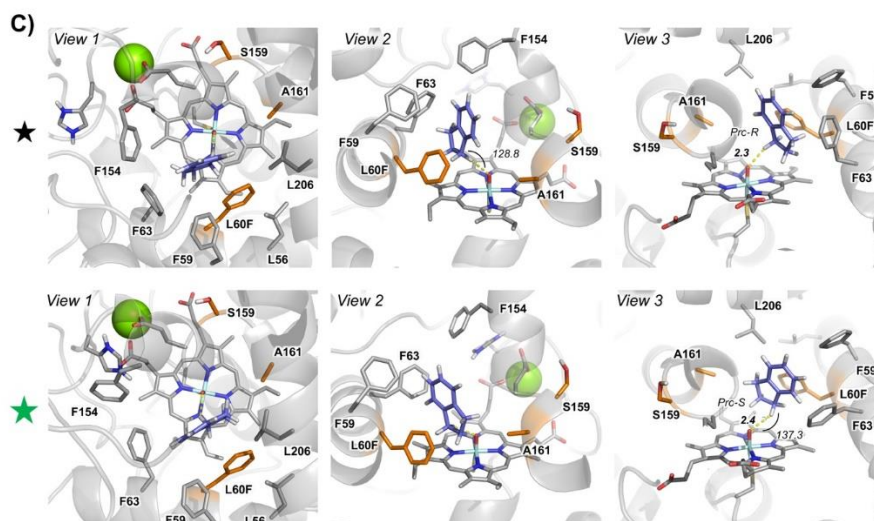*Replica 2*   *Replica 2*

Replica 3   *Replica 3*

Figure S20. A) Analysis of indane binding modes through three independent MD replicas in L60F variant. Key distances relevant for hydroxylations and aromatic oxidations are monitored along MD simulations, as described in the schemes. Bond Dissociation Energies (BDEs) for C–H bonds at C1 and C2 positions of indane are reported ((U)B3LYP/Def2TZVP/PCM(CH$_2$Cl$_2$)// (U)B3LYP/6-31G(d)/PCM(CH$_2$Cl$_2$), in kcal·mol^{-1}). **B)** Stereoselectivity analysis of indane oxidation at C1 position from MD simulations. Catalytic distances between O(Fe=O) – H and attack angles (O(Fe=O) – H – C) in the two equivalent *pro*-R and *pro*-S C–H bonds have been monitored, as described in the schemes. **C)** Representative snapshots from MD trajectories (highlighted with an "star" symbol) that describe reactive near attack conformations explored during MDs. Distances and angles are given in angstroms (Å) and degrees (º), respectively. Indane bound in L60F variant predominantly explores catalytically relevant binding poses in which the C1 positions are placed in a near attack conformation respect to the Fe=O active species (**A** and **B**). Calculated C1–H BDEs are significantly lower than those estimated for C2–H, indicating a higher intrinsic reactivity at C1 position. Deeper analysis of the C1 hydroxylation stereopreferences from MD simulations (**B**) demonstrated that *pro*-R C1–H bonds preferentially explore much better near attack conformations (i.e. shorter distances and more optimal attack angles) than *pro*-S C1–H, in line with the high stereoselectivity observed for L60F variant.

References

[1]   E. Rommel, J. Wirz, *Helv. Chim. Acta.* **1977**, *60*, 38-42.
[2]   M. N. Bakola-Christianopoulou, V. P. Papageorgiou, K. K. Apazidou, *Phosphorus, Sulfur, Silicon Rel. Elem.* **1994**, *88*, 53-65.
[3]   P. Püllmann, C. Ulpinnis, S. Marillonnet, et al., *Sci. Rep.* **2019**, *9*, 10932.
[4]   S. Kille, C. G. Acevedo-Rocha, L. P. Parra, et al., *ACS Synth. Biol.* **2013**, *2*, 83-92.
[5]   C. G. Acevedo-Rocha, M. T. Reetz, Y. Nov, *Sci. Rep.* **2015**, *5*, 10654.
[6]   W. M. Patrick, A. E. Firth, J. M. Blackburn, *Protein Eng.* **2003**, *16*, 451-457.
[7]   P. Molina-Espeja, E. Garcia-Ruiz, D. Gonzalez-Perez, et al., *Appl. Environ. Microbiol.* **2014**, *80*, 3496-3507.
[8]   M. Poraj-Kobielska, M. Kinne, R. Ullrich, et al., *Anal. Biochem.* **2012**, *421*, 327-329.
[9]   J. Santos-Aberturas, M. Dorr, G. S. Waldo, et al., *Chem. Biol.* **2015**, *22*, 1406-1414.
[10]   P. Püllmann, A. Knorrscheidt, J. Münch, et al., *bioRxiv* **2020**, 2020.2007.2022.216432.
[11]   A. Waterhouse, M. Bertoni, S. Bienert, et al., *Nucleic Acids Res.* **2018**, *46*, W296-W303.
[12]   I. Y. B.-S. D.A. Case, S.R. Brozell, D.S. Cerutti, et al., *Amber 2018*, University of California, San Francisco, **2018**.
[13]   P. Li, K. M. Merz, *J. Chem. Inf. Model.* **2016**, *56*, 599-604.
[14]   J. Wang, R. M. Wolf, J. W. Caldwell, et al., *J. Comput. Chem.* **2004**, *25*, 1157-1174.
[15]   C. I. Bayly, P. Cieplak, W. Cornell, et al., *J. Phys. Chem. A* **1993**, *97*, 10269-10280.
[16]   B. H. Besler, K. M. Merz Jr, P. A. Kollman, *J.Comp. Chem.* **1990**, *11*, 431-439.
[17]   U. C. Singh, P. A. Kollman, *J. Comput. Chem.* **1984**, *5*, 129-145.
[18]   K. Shahrokh, A. Orendt, G. S. Yost, et al., *J. Chem. Theory Comput.* **2012**, *33*, 119-133.
[19]   W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, et al., *J. Chem. Phys.* **1983**, *79*, 926-935.
[20]   J. A. Maier, C. Martinez, K. Kasavajhala, et al., *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.
[21]   T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089-10092.
[22]   D. R. Roe, T. E. Cheatham, *J. Chem. Theory Comput.* **2013**, *9*, 3084-3095.
[23]   O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455-461.
[24]   M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al., Wallingford, CT, **2016**.
[25]   A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098-3100.
[26]   A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648-5652.
[27]   C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B Condens Matter* **1988**, *37*, 785-789.
[28]   V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995-2001.
[29]   M. Cossi, N. Rega, G. Scalmani, et al., *J. Comput. Chem.* **2003**, *24*, 669-681.
[30]   C. N. Schutz, A. Warshel, *Proteins* **2001**, *44*, 400-417.