Modulating chemoselectivity in a Fe(II)/ α-ketoglutarate dependent dioxygenase for the oxidative modification of a non-proteinogenic amino acid

Fabian Meyer,[+ a] *Raphael Frey,*[+ a c] *Mathieu Ligibel,*[b] *Emine Sager,*[b] *Kirsten Schroer,*[b] *Radka Snajdrova,*[b] *and Rebecca Buller**[a]

[a] Competence Center for Biocatalysis, Institute of Chemistry and Biotechnology, Zurich University of Applied Sciences, Einsiedlerstrasse 31, 8820 Wädenswil, Switzerland

[b] Novartis Institutes for BioMedical Research, Global Discovery Chemistry, 4056 Basel, Switzerland

[c] present address: Bioconjugates R&D, Lonza Ltd, Rottenstrasse 6, 3930 Visp, Switzerland

+ These authors contributed equally to the work.

* rebecca.buller@zhaw.ch

Abstract:

Modification of aliphatic C-H bonds in a regio- and stereoselective manner can pose a formidable challenge in organic chemistry. In this context, the use of non-heme iron and α-ketoglutarate-dependent dioxygenases (αKGDs) represents an interesting tool for C-H activation as this enzyme family can catalyze a broad set of synthetically valuable reactions including hydroxylations, oxidations and desaturations. The consensus reaction mechanism of αKGDs proceeds via the formation of a Fe(IV)-oxo complex capable of hydrogen atom transfer (HAT) from a sp^3-hybridized substrate carbon center. The resulting substrate radical and Fe(III)-OH cofactor is considered to be the branch point toward the possible reaction outcomes which are determined by the enzyme's active site architecture. To date, the modulation of the reaction fate in Fe(II)/ α-ketoglutarate-dependent dioxygenases via enzyme engineering has been mainly elusive. In this study, we therefore set out to engineer the L-proline cis-4-hydroxylase SmP4H from *Sinorhizobium meliloti* for selective oxidative modifications of the non-proteinogenic amino acid L-homophenylalanine (L-hPhe) to produce pharmacological relevant small molecule intermediates. Using structure-guided directed evolution, we improved the total turnover number, the k_{cat} as well as the k_{cat}/K_m of the hydroxylation reaction yielding the desired γ-hydroxylation product by approximately 10-fold, >100-fold and > 300-fold, respectively. Notably, the exchange of only one amino acid in the active site (W40Y) allowed us to re-program the natural hydroxylase to predominantly act as a desaturase, presumably through tyrosine's capability to serve as a catalytic entity in the reaction mechanism. An investigation of the substrate scope revealed additional acceptance of the non-canonical amino acids L-homotyrosine and *(S)*-α-amino-3,4-dichlorobenzenebutanoic acid by SmP4H variants.

Keywords: non-heme Fe(II)/ α-ketoglutarate-dependent dioxygenases, C-H functionalization, enzyme engineering, biocatalysis, non-proteinogenic amino acid, chemoselectivity

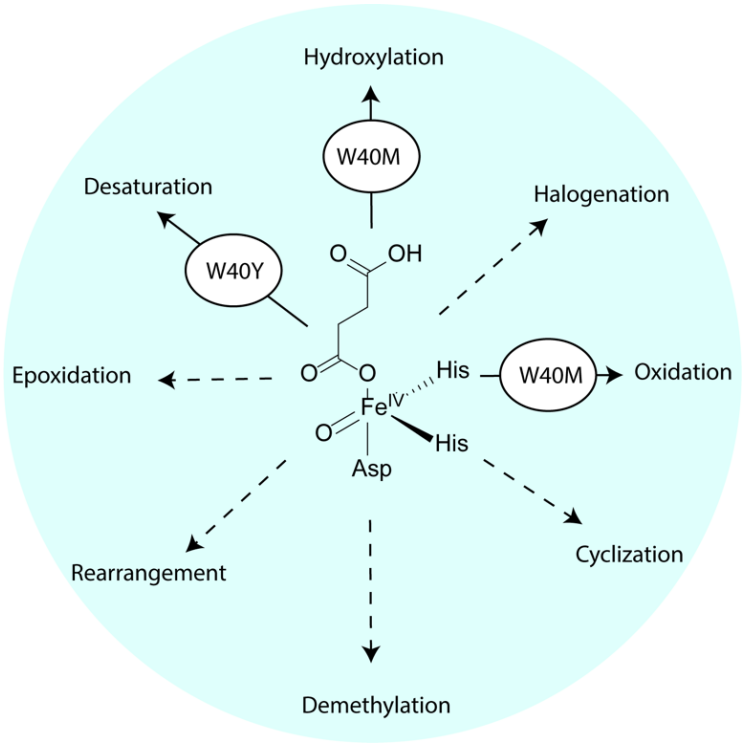
Introduction:

The selective modification of unactivated sp3-hybridized carbons poses a considerable challenge for traditional synthetic chemistry often requiring the use of toxic and/or difficult-to-handle chemicals. Particularly, the development of chemoselective catalysts that can oxidize aliphatic C-H bonds in the presence of more electron-rich functional groups such as olefins, aromatic rings, and nitrogen atoms remains highly desirable [1]. Non-heme iron(II)/α-ketoglutarate (Fe(II)/α-KG)-dependent dioxygenases have recently emerged as promising biocatalysts for mild and selective C-H functionalizations [2–4]. The versatile Fe(II)/α-KG-dependent dioxygenases only require α-ketoglutarate and O_2 as co-substrates to catalyse a plethora of different reactions including hydroxylations, epoxidations, halogenations, oxidations, desaturations, endoperoxidations, ring-forming and re-arrangement reactions [2,3,5,6]. The enzymes have been applied in chemo-enzymatic natural product synthesis, for example for the production of podophyllotoxin, an important lead compound in drug discovery [7,8], or for the commercial derivatization of amino acids and derivatives such as L-pipecolic acid, L-proline or L-isoleucine [9,10]. The catalytic mechanism employed by Fe(II)/α-KG-dependent dioxygenases involves initial formation of a substrate radical by abstraction of an H-atom through a high-valent Fe(IV)-oxo species, similar to P450 enzymes (Figure S1) [11]. Subsequent events in the enzyme active site determine the fate of the substrate radical and, hence, the different reaction outcomes (Scheme 1). In case of hydroxylases, for example, the iron-coordinated hydroxyl (Fe(III)-OH) reacts with the carbon substrate radical leading to the formation of a new C-O bond [11]. In reaction outcomes other than hydroxylation, rebound of the hydroxyl group is often suppressed by the enzyme, thereby enabling a different fate of the substrate radical, such as rebound of a halogen [12,13] or olefin formation in case of desaturases [14]. Further processing of the hydroxylated intermediate by the Fe(II)/α-KG-dependent

dioxygenases is another way to generate diversity [15]. Often, these functional differences are governed by subtle structural dissimilarities in the enzyme active sites, affecting the positioning of the substrate in relation to the catalytic residues and co-factors, and are not readily predictable by inspection of the enzyme structure alone [16,17]. Directed evolution approaches could, therefore, be a promising strategy to not only improve the dominant Fe(II)/α-KG-dependent dioxygenases activity for non-native substrates but access new reactivities altogether. Intrigued by the idea to engineer variants of the same parental enzyme capable to carry out multiple types of sp^3-carbon functionalizations, we set out to build a Fe(II)/α-KG-dependent dioxygenases toolbox for the derivatization of non-native substrates with the aim to create valuable building blocks for the chemical and pharmaceutical industry.

Taking a first step in this direction, we describe the set-up and screening of a library consisting of wildtype Fe(II)/α-KG-dependent dioxygenases and the subsequent use of a semi-rational protein engineering approach to adapt a proline hydroxylase from *Sinorhizobium meliloti* (SmP4H) [18] for efficient and selective oxidative modifications of the non-proteinogenic amino acid L-homophenylalanine (L-hPhe, 1). γ-Hydroxylated α-amino acids are interesting synthons in the manufacture of valuable intermediates and building blocks. While access to β-hydroxylated α-amino acids can be obtained via e.g. threonine aldolases [19], we describe here the use Fe(II)/α-KG-dependent dioxygenases as a novel method to access γ-hydroxylated α-amino acids and therefore further expand the enzymatic toolbox for the synthesis of highly functionalized non-proteinogenic amino acids. The targeted product, the γ-hydroxylated L-hPhe (2), is a valuable intermediate in the synthesis of potential kynureninase or kynurenine monooxygenase inhibitors [20–22]. Kynureninases and kynurenine monooxygenase are a group of enzymes involved in the kynurenine pathway which is the major pathway for tryptophan catabolism in mammals. Several intermediates of this

pathway are potentially toxic and have been implicated in the progression of neurodegenerative disorders, psychiatric disorder and cancer, triggering the search for potent inhibitors of the responsible enzymes [22,23].



Scheme 1 Catalytic mechanism of Fe(II)/ α -KG-dependent enzymes involves formation of a high-valent Fe(IV)-oxo species capable of formation of a substrate radical by H-atom abstraction. Subsequent catalytic steps define the different reaction outcomes in dependence of the active site environment of the corresponding enzyme. In case of SmP4H, the position W40 represents a functional junction and its selective substitution allows fine-tuning of the chemoselectivity of the enzyme.

While the utility of Fe(II)/α-KG-dependent amino acid hydroxylases has been demonstrated for the biocatalytic production of other non-canonical amino acids, the substrate scope in these examples often remained limited to the native amino acid substrate or structurally close

congeners [9,24,25]. Starting from a small promiscuous hydroxylation activity of SmP4H, a natural proline hydroxylase, towards L-hPhe (1), we identified enzyme variants with significantly improved hydroxylation efficiency, enabling the biocatalytic production of γ-hydroxylated L-hPhe (2), a versatile precursor molecule that can be further transformed chemically and enzymatically. Applying acidic conditions, one-pot self-lactonization of the obtained product was possible to give the lactone product, (3S,5R)-3-amino-5-phenyldihydrofuran-2(3H)-one (6). Furthermore, close inspection of the biocatalytic reaction revealed an enzymatic cascade leading to two additional products, the corresponding γ-ketone (3) ((S)-2-amino-4-oxo-4-phenylbutanoic acid) and β-hydroxy-γ-ketone (4) ((2S)-2-amino-3-hydroxy-4-oxo-4-phenylbutanoic acid). Importantly, the introduction of a new catalytic entity, a tyrosine (W40Y), allowed us to tune the chemistry taking place in the active site turning SmP4H into a predominant desaturase, giving almost exclusively 3,4-desaturated L-hPhe (5) (Figure 1).

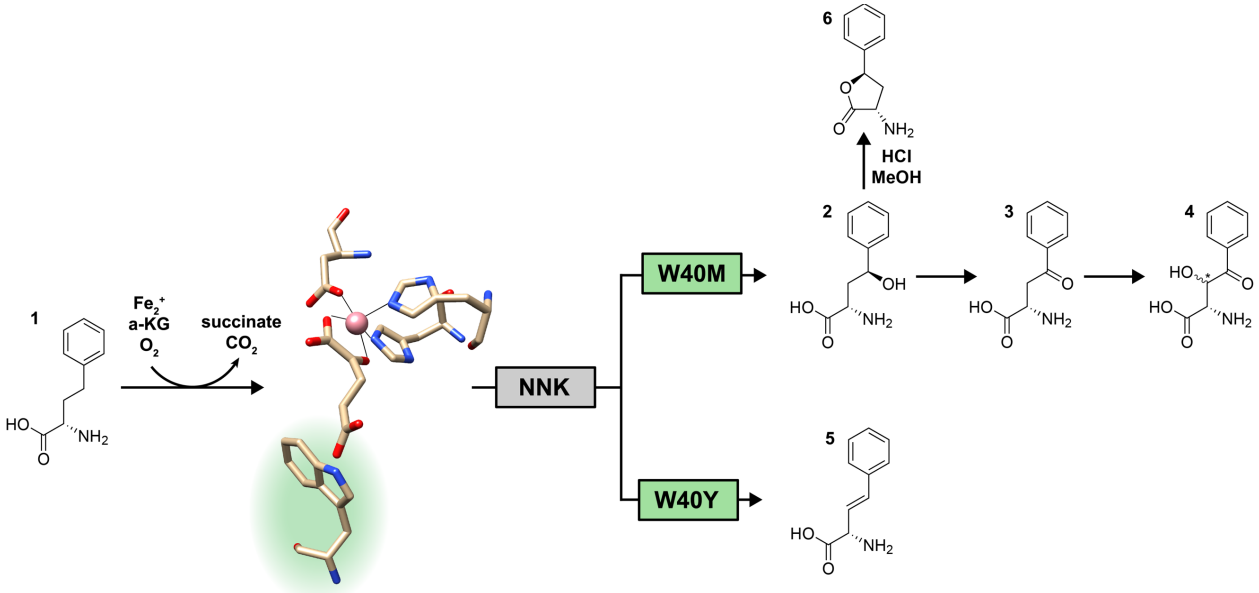


Figure 1 A minimal set of mutations leads to altered chemoselectivity in SmP4H resulting in different modifications of L-homo phenylalanine.

Results and Discussion:

We embarked on our enzyme engineering campaign by assembling a panel of 36 functionally characterized Fe(II)/α-KG-dependent dioxygenases (Figure 2a and Table S1). To identify a suitable starting scaffold for evolution, we screened our wildtype enzyme library, overexpressed in *E.coli* BL21(DE3), for hydroxylase activity towards L-hPhe in 96-well plate crude lysate assay using liquid-chromatography-coupled to mass spectrometry (LC-MS) to detect product formation (Figure 2b). While most enzymes did not show any activity towards the target molecule, the biocatalytic reactions carried out with proline hydroxylase SmP4H led to the formation of a small amount of a product with $\Delta m/z = +16$ to the substrate, indicating modification of L-hPhe with a hydroxyl group (Figure 2c). Enzymatic origin of this activity was confirmed by running reactions with empty vector and in the absence of the co-factor α-KG, in which only traces of hydroxylated product were detected (Figure 2d). Isolation of the putative hydroxylation product from upscaled biocatalytic reactions allowed confirmation of the structure using nuclear magnetic resonance (NMR) analysis and revealed hydroxylation to regio- and stereoselectively occur at the aliphatic sp^3-γ-carbon (Figure 1). Enzymatic functionalization at this position complements existing chemical methods to access unactivated aliphatic C-H bonds in the presence of electron-rich functional groups, like aromatic rings [1,26]. For example, the chemical derivatization method of L-hPhe using a glycine-imine cobalt(III) complex with acetophenone followed by reduction developed by Laval et al. led to a racemic mixture of γ-hydroxylated L-hPhe [27].

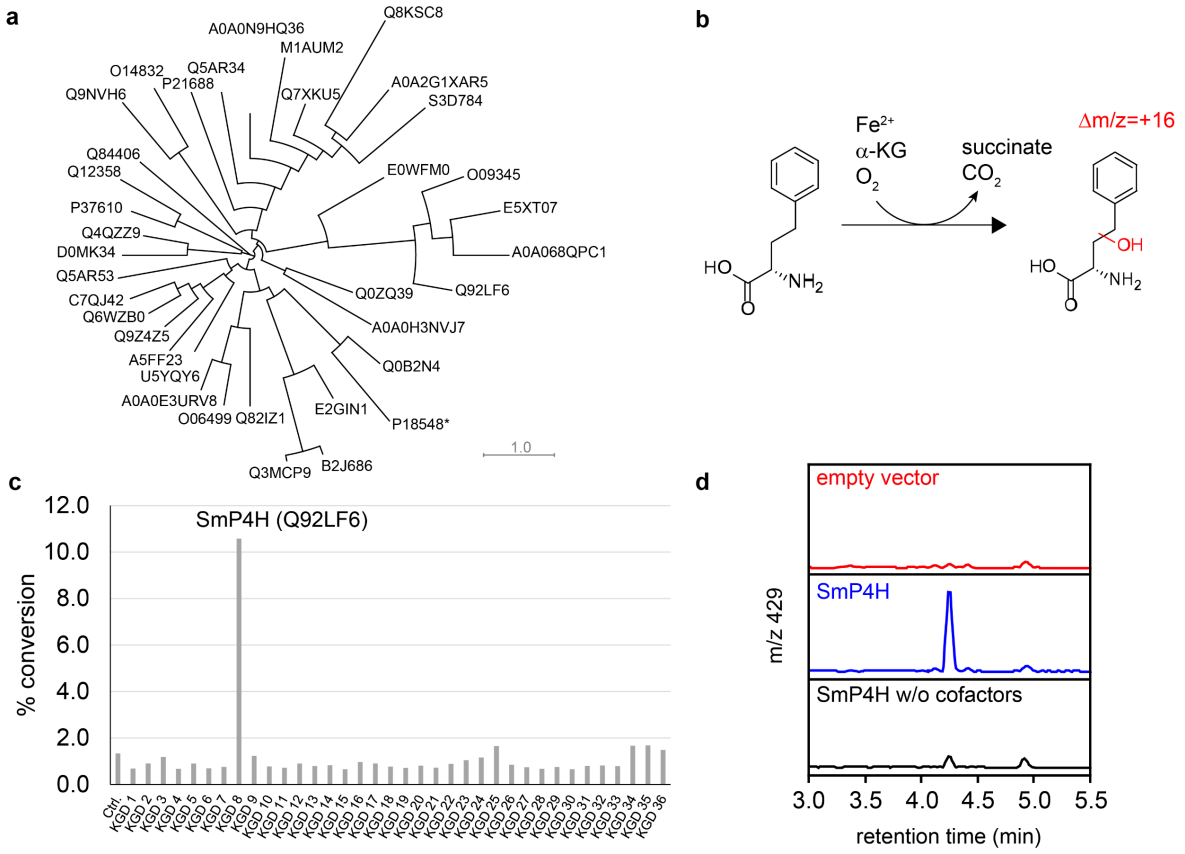


Figure 2 Screening approach to identify a suitable Fe(II)/α-KG-dependent dioxygenases with hydroxylase activity towards L-homophenylalanine. a) Fe(II)/α-KG-dependent dioxygenase library consisting of 36 enzymes were selected to represent a diverse phylogenetic set of functional enzymes. b) Target reaction for oxidative modification of L-homophenylalanine. c) Fe(II)/α-KG-dependent dioxygenase library screen led to the identification of L-homophenylalanine hydroxylation activity by the enzyme SmP4H from *Sinorhizobium meliloti* (Table S1). d) Selective ion monitoring (SIM) of dansyl-derivatized γ-hydroxylated L-hPhe; red, empty vector control; blue, expressed SmP4H in reaction mix; black, expressed SmP4H in reaction mix but without cofactors.

While SmP4H showed high regioselectivity, hydroxylation activity was low (approximately 10% conversion under assay conditions). We, therefore, opted for a structure-guided semi-rational

protein engineering approach to improve hydroxylation efficiency and generate an enzyme for the preparative biocatalytic production of hydroxylated L-hPhe. In the absence of a crystal structure for SmP4H, a homology model based on the structure of a homologous proline hydroxylase (MlP4H) (66% sequence identity, Figure S2) [25] was built using SwissModel [28] and the target molecule L-hPhe was docked in the active site (Figure 3). In one of the docking solutions, the molecule is located such that its carboxyl and amino group were directed towards Arg118 and Asp108 respectively. A similar positioning of the native substrate L-proline is observed in the crystal structure of MlP4H and enables electrostatic and H-bonding interactions to lock the amino acid substrate near the catalytically active Fe(II) species [25]. Using this docking model, we identified 11 sites in the vicinity of the target molecule. For each of the 11 sites, we individually created site-saturated mutagenesis libraries using NNK degenerate codons encoding all 20 amino acids.

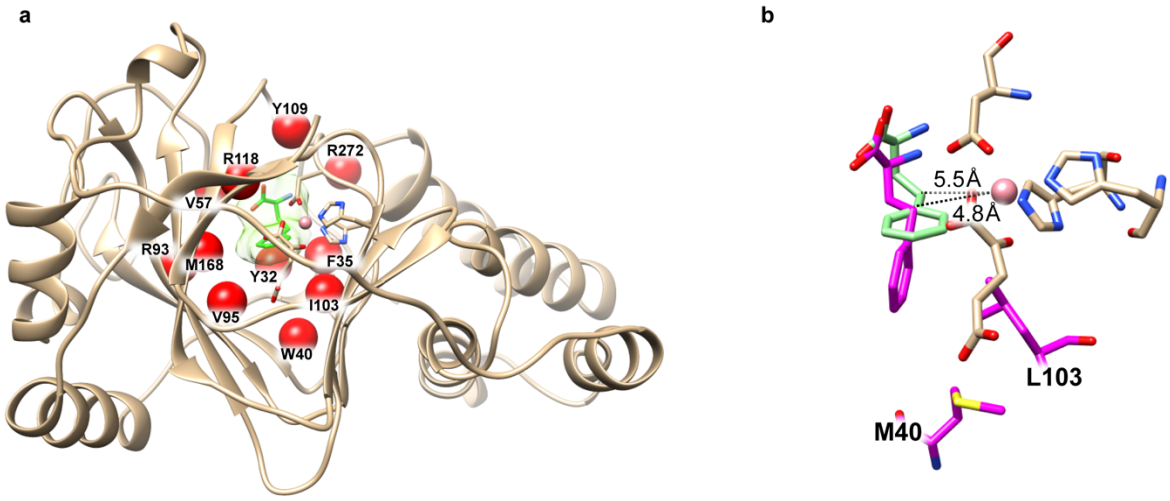


Figure 3 a) A substrate-docked homology model of SmP4H wildtype. The red spheres indicate the residues targeted for site direct mutagenesis using NNK libraries. b) Active site of improved variant W40M I103L showing the original docking solution for L-hPhe in the parent enzyme (green) and the new docking solution found in the double mutant (magenta). The acquired

mutations W40M and I103L enable a new position of the substrate presumably shortening the distance of the β-carbon atom to the Fe(II) center from 5.5 to 4.8 Å.

Screening of the 1100 variants was performed using the LC-MS based crude cell lysate assay. Mutations at two positions, W40 and I103, were identified that gave up to 4.5-fold improved conversion (Figure S3). Particularly, smaller amino acid residues at these positions seemed to improve hydroxylation, putatively by enabling better accommodation of the phenyl moiety of L-hPhe (Table S4). The best variant in this first round of engineering, W40M, was subsequently characterized as a purified enzyme and showed an approximately 8-fold higher apparent k_{cat} compared to the wild type (Table 1). In order to find a potentially more beneficial combination of mutations, we created a 2-site randomized library using degenerate codons DBK[29] and DYA at positions W40 and I103, respectively (Figure S4). The best variant emerging from the second round of engineering, W40M I103L, showed an additional ca. 2-fold increased conversion to the hydroxylated product compared to its parent W40M in the crude lysate screen which was verified by kinetic analysis with purified enzymes. Overall, the double variant (W40M I103L) exhibited an > 100-fold improved k_{cat} and a > 300-fold improved k_{cat}/K_m compared to the wild type enzyme (Table 1 and Figure S5). Additionally, the turnover number (TTN) of the double variant (W40M I103L) for the hydroxylation of L-hPhe was increased by 10-fold to a value of approximately 39 during the course of the evolution. For comparison, the turnover number of the wildtype SmP4H for its native substrate proline was found to be 257 ± 10.

Table 1: Total turnover number and apparent K_m, k_{cat} and k_{cat}/K_m of engineered SmP4H variants for the hydroxylation of L-hPhe.

Variant	K_m [mM]	k_{cat} [min^{-1}]	rel. k_{cat}	k_{cat}/K_m [$M^{-1}s^{-1}$]	rel. k_{cat}/K_m	TTN	rel. TTN
WT	1.10 ± 0.24	0.015 ± 0.001	1	0.2 ± 0.5	1	4.1 ± 0.6	1.0
W40M	0.24 ± 0.04	0.124 ± 0.003	8	8.6 ± 0.2	38	10.6 ± 1.1	2.6
W40M I103L	0.40 ± 0.08	1.680 ± 0.068	112	70.0 ± 14.0	308	38.7 ± 3.5	9.4

Mean values (n = 3 biological replicates) are shown. Reaction conditions: 2 mM L-hPhe, 10 mM ascorbic acid, 1 μM enzyme, 2h at 20°C.

Besides its utility as a potential precursor for the synthesis of kynureninase monooxygenase inhibitors, γ-hydroxylated L-hPhe can also undergo self-lactonization under acidic conditions to give lactone (6), a structural analogue of pharmacological relevant compounds [30]. Using 20 μM W40M I103L and 0.5 mM L-hPhe, lactone (6) could be generated from γ-hydroxylated L-hPhe in a one-pot reaction by replacing water with organic solvent under acidic conditions. Isolation of lactone (6) from upscaled chemoenzymatic reactions allowed confirmation of the structure using nuclear magnetic resonance (NMR). Under these conditions a yield of 94% lactonization was achieved further underscoring the chemoenzymatic utility of C-H functionalization by the W40M I103L variant.

Conversion of L-hPhe with the evolved SmP4H variants, however, did not exclusively yield γ-hydroxylated L-hPhe. Both in the crude lysate and *in vitro* experiments, we observed two side-products, γ-ketone (3) and β-hydroxy-γ-ketone (4) (Figure 1). The identity of γ-ketone (3) was

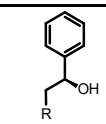
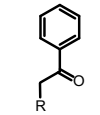
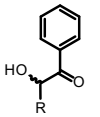
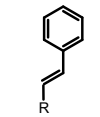
confirmed by comparison to a commercially available standard by LC-MS, whereas the structure of the β-hydroxy-γ-ketone (4) was solved by NMR analysis. Our experiments suggest that these products result from further oxidation of the γ-hydroxylated product (2) (Figure S6 and Figure S7). Using the commercially available standard for ketone (3) we could show that this molecule is a well-accepted substrate for the engineered SmP4H variants and is hydroxylated surprisingly efficiently by W40M I103L to give β-hydroxy-γ-ketone (4) (Figure S8). All enzymatically functionalized compounds can easily be separated chromatographically and serve as valuable synthons in medicinal chemistry.

Based on these findings, we revaluated the oxidative modification of L-hPhe and the conversion into the different products using purified enzymes (Table 2). Comparison of the different SmP4H variants shows that we were able to improve production of γ-hydroxylated L-hPhe in a significant manner from 1.6% to 35.3% (W40M I103L). Improved conversion was achieved as well in case of the subsequent oxidation to product (3) and formation of β-hydroxy-γ-ketone (4). While in case of W40M accumulation of γ-ketone (3) was observed, the W40M I103L variant showed increased conversion towards β-hydroxy-γ-ketone (4). Michaelis-Menten analysis confirmed that the additional mutation of I103L leads to an increased catalytic efficiency towards the β-hydroxylation of the γ-ketone (3) (Figure S8). This multifunctional character of SmP4H is not uncommon for this class of enzymes. Several examples of Fe(II)/α-KG-dependent dioxygenases exist that catalyse a hydroxylation and sequential oxidation to the ketone, the aldehyde or the carboxylic acid [15,31,32] or possess multiple activities including hydroxylation, desaturation and epoxidation such as PtlD [33] involved in the neopentalenoketolactone biosynthesis [34].

Intrigued by the catalytic promiscuity of the Fe(II)/α-KG-dependent dioxygenase SmP4H, we searched the enzyme library screening data for alternative products, hoping to identify SmP4H

variants capable of chemistries other than L-hPhe hydroxylation. Interestingly, we found a variant, W40Y, which gave a product (>80% selectivity) corresponding to a loss of 2 m/z units (Figure S9). Using NMR, we could solve the structure of the obtained product and identified the compound as 3,4-desaturated L-hPhe (5) (Figure 1). Kinetic analysis of variant W40Y allowed us to determine k_{cat} (0.83 ± 0.02 min^{-1}) and K_m (0.12 ± 0.02 mM) values, which are in a similar range as seen for the hydroxylase variants (Figure S13). This report therefore is, to our knowledge, the first example of remodelling a Fe(II)/α-KG-dependent hydroxylase into a desaturase by enzyme engineering. Notably, our results additionally highlight that the here-described engineered SmP4H variant represents a rare case of a desaturase that acts on C-C unit without an activating α-heteroatom [14], a synthetically very desirable reaction. To date, only PrhA [35], BcmF [36] and KabC [37] have been described to desaturate entirely unactivated C-C units in their natural substrates, whereas native VioC [14] and CarC [38] have been reported to exhibit non-α-heteroatom assisted desaturation activity in addition to hydroxylation activity when acting on non-native compounds.

Table 2: Product distribution of oxidatively modified L-homophenylalanine generated by wildtype and evolved variants based on extracted ion chromatograms.

	Product	WT	W40M	W40M I103L	W40Y	W40F
2		1.6%	17.8%	35.3%	5.9%	21.8%
3		4.3%	17.0%	4.1%	5.1%	4.8%
4		0.3%	7.3%	29.8%	6.8%	26.9%
5		0.1%	1.9%	0.7%	50.7%	0.8%

Reaction conditions: 1mM L-hPhe, 10 mM ascorbic acid, 20 μM enzyme, 60 min at 20°C. Mean values (n = 3 biological replicates) are shown. R= NH_2-CH-COOH

To better understand the role of the introduced tyrosine residue in the context of SmP4H catalysed L-hPhe desaturation, we carried out in-depth docking analyses. These studies revealed that the tyrosine residue may cause the aromatic ring of the substrate to shift slightly leading to a conformation that exposes both β- and γ-H atoms to the Fe(II) center (Figure S10). Similar substrate placements - i.e. proximity of multiple substrate carbons to the active iron center - have also been found in Fe(II)/α-KG-dependent dioxygenases which have been described to carry out multiple reaction chemistries such as VioC, capable of hydroxylation and desaturation [14,39,40]. As a consequence, the newly introduced tyrosine residue might lead to a substrate placement relative to the active Fe-center in which two sequential H-atom transfers (HAT) are favoured over rebound of a hydroxyl group allowing di-radical dimerization to give the desaturated product (Figure 4a).

In this scenario, however, the intermediate Fe(III)-OH, expected to be a modest oxidant, would be capable to catalyse the second HAT step at the carbon adjacent to the radical center. In addition, the SmP4H variant W40Y would be strikingly efficient in favouring the second HAT step over hydroxyl rebound (Table 2).

More likely, the introduced tyrosine not only influences the interaction of substrate and iron center sterically, but actively contributes to the desaturation mechanism: The product distribution of SmP4H variant W40F, in which position 40 was substituted with a phenylalanine instead of a tyrosine, supports this hypothesis as it showed almost no desaturated product (Table 2). Even though a structural effect of the tyrosine substitution cannot be excluded, the predominant hydroxylation activity of variant W40F make it plausible that the additional hydroxyl-group of the tyrosine residue is mechanistically relevant. Consequently, the desaturation reaction could take place via an intermediate hydroxylation of L-hPhe followed by a deprotonation event in which tyrosine acts as the potential base (Figure 4b). However, using β- or γ-hydroxylated L-hPhe as substrate in biocatalytic reactions with variant W40Y did not yield any desaturated product suggesting that a cation-triggered desaturation pathway might be a more likely scenario (Figure 4c). Here, the conversion of the intermittent L-hPhe radical to 3,4-desaturated L-hPhe would involve electron transfer from the radical to the Fe(III)-OH species followed by deprotonation via tyrosine, acting as a base. Precedent for the required unusual low pK_a of the phenolic hydroxyl to perform general base catalysis has been suggested, for example, in the case of nitrile hydratases[41,42], alanine racemases[43,44] and the 2-((R)-2-hydroxypropylthio)ethanesulfonate dehydrogenase[45]. In this reaction pathway, the phenyl-moiety of the substrate would stabilize the formed carbocation, mechanistically resembling the α-heteroatom effect described in natural desaturases.

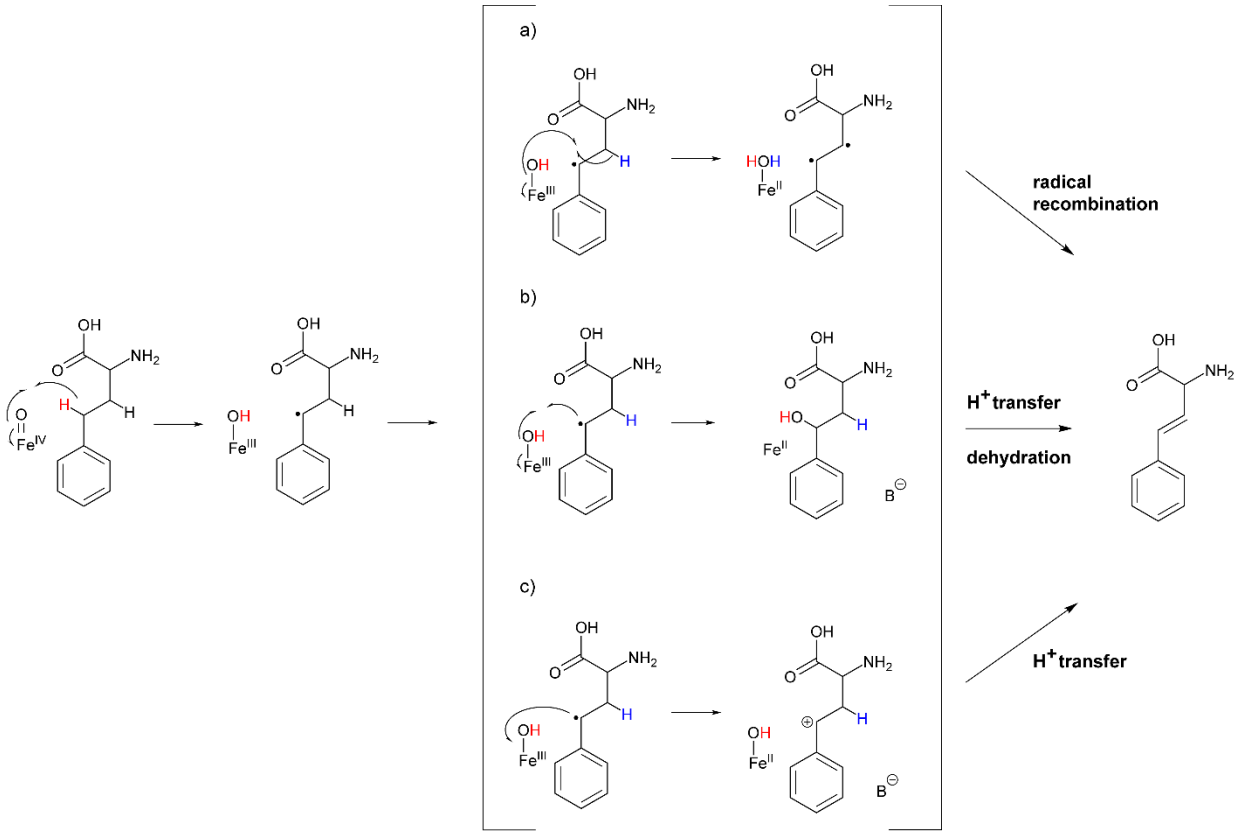


Figure 4 Proposed pathways for SmPH4 W40Y catalyzed desaturation. After CX- H abstraction, the reaction can proceed through a) a second hydrogen abstraction (2-HAT) and a subsequent di-radical recombination, b) hydroxylation and dehydration, or c) carbocation formation and subsequent H^+ removal by a base to generate the C=C bond.

In our quest to synthesize even more potent kynurenine monooxygenase inhibitors such as (S)-2-amino-4-(3,4-dichlorophenyl)-4-oxobutanoic acid (IC_{50} of 0.2 μM compared to compound 3 with an IC_{50} of 7.6 μM)[22,46], we probed the substrate promiscuity of our evolved Fe(II)/α-KG-dependent dioxygenase SmP4H with the non-canonical amino acids L-homotyrosine and *(S)*-α-amino-3,4-dichlorobenzenebutanoic acid. The substitutions present on the aryl rings of the molecules were additionally expected to modulate the properties of any arising radical at the γ -position, with the corresponding L-homotyrosine radicals being more nucleophilic and the *(S)*-α-amino-3,4-

dichlorobenzenebutanoic acid radicals being more electrophilic than the respective L-homophenyl radical. Strikingly, we were able to detect formation of products with $\Delta m/z = +16$ indicating the installation of a hydroxyl group into both additional non-native substrates. Whereas in SmP4H W40M I103L catalyzed reactions of L-homotyrosine the biocatalytic cascade halted at the first hydroxylation step (Compound 7, Figure 5), the respective reactions with *(S)*-α-amino-3,4-dichlorobenzenebutanoic acid predominantly led to the formation of a product with $\Delta m/z = +30$ to the substrate (Figure 5). In addition, we found that wildtype SmP4H only very poorly accepts L-homotyrosine as a substrate (conversion 0.5 %) but readily oxidizes *(S)*-α-amino-3,4-dichlorobenzenebutanoic acid leading to the prevalent formation of products with $\Delta m/z = +16$ (Compound 7, Figure 5) and $\Delta m/z = +14$ (Compound 8, Figure 5), similarly reflecting the intrinsic reactivity of the substrate (Table S5, Table S6 and Figure S14).

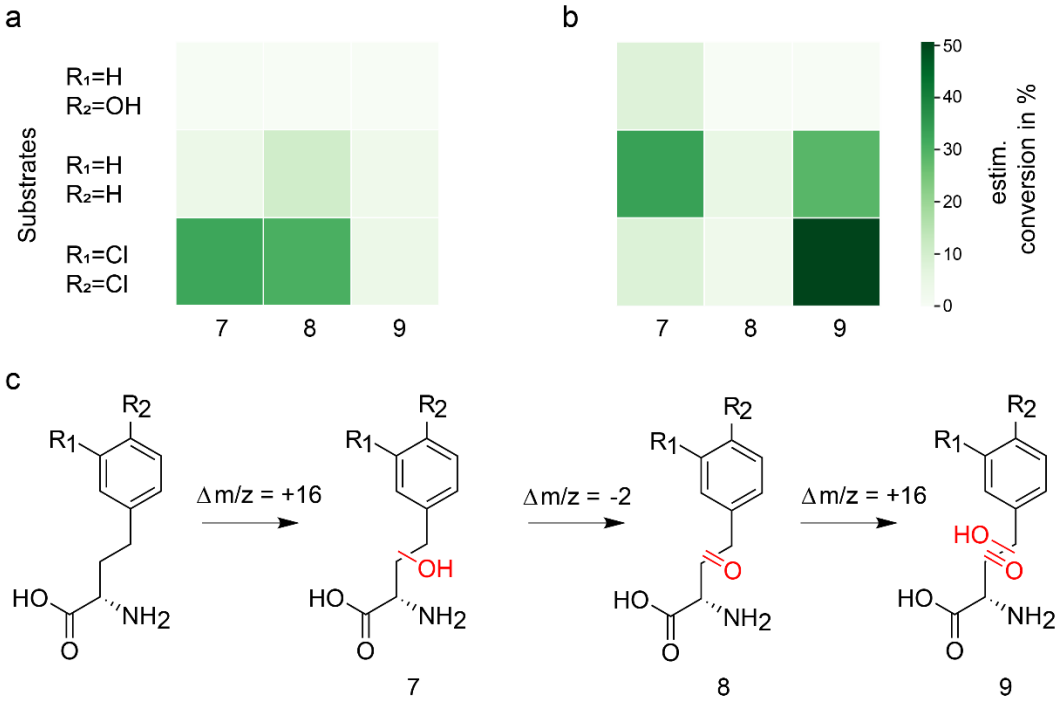


Figure 5 Substrate screen with wildtype SmP4H and engineered variant W40M I103L investigating conversion of L-homotyrosine, L-homophenylalanine and (*S*)-*α*-amino-3,4-dichlorobenzenebutanoic acid. a) Conversions with wildtype SmP4H. b) Conversions with engineered variant SmP4H W40M I103L. c) Proposed reaction cascade based on Δm/z values in analogy with confirmed structures from L-homophenylalanine (R_1=H, R_2=H). Conversions are estimated through comparison with the negative control omitting the enzyme and the calculations assume similar ionization efficiency of products originating from the same substrate.

Conclusion:

The plethora of reactions accessible via Fe(II)/α-KG-dependent dioxygenases make them highly interesting catalysts for the production of difficult-to-synthesize products. By setting-up and screening a wildtype library of functionally characterized Fe(II)/α-KG-dependent dioxygenases, we identified an enzyme capable of performing oxidative modifications on the non-proteinogenic amino acid L-homophenylalanine at the β- and γ- position as well as accepting L-homotyrosine and *(S)*-α-amino-3,4-dichlorobenzenebutanoic acid as substrates leading to valuable kynurenine monooxygenase inhibitors and their building blocks. Functionalization of unactivated C(sp^3)–H bonds in the presence of more electron-rich functional groups, like aromatic rings, remains a challenge in organic chemistry – however, as we show here, this can be overcome by means of a suitable biocatalyst [1,26].

In addition, we showcase a rare example of engineering a Fe(II)/α-KG dioxygenase toward higher activity [47]: Not only did we successfully identify enzyme variants with a >100-fold improved apparent k_{cat}, a > 300-fold improved k_{cat}/K_m and a 10-fold increased total turnover number for the hydroxylation of L-hPhe, but we also revealed the importance of position W40 in SmP4H which serves as a functional junction for the reaction outcome. By substituting position W40 with specific amino acids, we were able to improve L-hPhe hydroxylation activity while almost completely abolishing acceptance of the native substrate L-proline (W40M) (Figure S11) or we could trigger the formation of an alternative, desaturated product (W40Y). Notably, the functional switch between a hydroxylase and a desaturase was achieved by exchanging a single amino acid supporting the possibility to deliberately access the chemical versatility of alpha-ketoglutarate dependent dioxygenases through redesign of the active site architecture. This finding is underscored by further recent studies on Fe(II)/α-KG-dependent dioxygenases in which it was

shown that only a small number of mutations was necessary to change the reaction outcome or to increase substrate promiscuity [24,48–51]. Identifying such hotspots is an important step to provide starting points for further engineering studies and acts as a basis to obtain deeper insights into the structure-function relationship of this enzyme class.

At present, the mechanism of olefin installing desaturation catalyzed by Fe(II)/α-KG-dependent dioxygenases remains poorly understood [14,52]. As discussed above, the engineered SmP4H variant W40Y may follow one of several hypothetical reaction mechanisms similar to what has been described for the VioC catalysed desaturation of L-homoarginine [14]. Notably, in several recent mechanistic analyses of Fe(II)/α-KG-dependent dioxygenases strong evidence for a cation mechanism was presented, such as for example in an enzyme-catalysed 1,2-phosphono-migration reaction [53] or the desaturation of cyclopeptin by AsqJ [54]. In addition, recent insights into the pyrrolidine and olefin formation mechanism of Fe(II)/α-KG-dependent dioxygenases KabC in the biosynthesis of kainic acid further allude to the possibility that a cation-triggered desaturation may be a general reaction principle used by these enzymes[37].

To further elucidate the underlying mechanistic motifs allowing for olefin insertion between unactivated carbons, we aligned the structures of literature-known desaturases (PrhA (5YBN), VioC (6DAX) and CarC (4OJ8)) with the model of our engineered SmP4H variant. Strikingly, in the case of CarC, a similar relative placement of tryptophan 40 in SmP4H and the mechanistically important tyrosine 165 in CarC was observed (Figure S12). CarC catalyzes two subsequent reactions, a stereoinversion reaction turning (3*S*, 5*S*)-carbapenam to (3*S*, 5*R*)-carbapenam and a subsequent desaturation reaction to (5*R*)-carbapenem[39] (Figure S12). In CarC, tyrosine 165 has been shown to be involved in the stereoinversion reaction which is not catalytic but proceeds stoichiometrically. Since multiple turnovers of desaturation are possible in case of the SmP4H

variant, however, the tyrosyl residue in our SmP4H variant probably does not act as a hydrogen atom donor to quench the substrate radical as proposed in the CarC stereoinversion [39].

At this point, we can therefore only speculate if position W40 might be of similar relevance in other alpha-ketoglutarate-dependent dioxygenases. It is noteworthy, though, that in our wildtype dioxygenase library position W40 is conserved in 5 out of 36 enzymes, indicating that the elucidation of this position's exact mechanistic or structural function in future enzyme engineering studies might be, in fact, a worthwhile pursuit.

ASSOCIATED CONTENT

Corresponding Author

*Rebecca Buller, Competence Center for Biocatalysis, Institute of Chemistry and Biotechnology, Zurich University of Applied Sciences, Einsiedlerstrasse 31, 8820 Wädenswil, Switzerland
Email: rebecca.buller@zhaw.ch

Notes

The authors declare no competing financial interest.

Supporting Information

This information is available free of charge on the ACS Publications website.

Experimental section, additional data, and figures including NMR elucidation of product structures, uniprot numbers of enzyme wildtype library, primer sequences, sequencing results of evolution rounds, sequence alignments, product distribution tables, biocatalysis results with L-proline, kinetic characterization of SmP4H variants, raw MS-spectra of biotransformations, docking studies with engineered variants.

ACKNOWLEDGMENT

The authors would like to thank Etienne Richard, Michael Brand and Flavio Gall for help with isolation of compounds prior to NMR analysis. This work was supported by Innosuisse—Swiss Innovation Agency (Grant No. 28385.1 PFLS-LS).

References

(1) White, M. C. Adding Aliphatic C-H Bond Oxidations to Synthesis. *Science* **2012**, *335*, 807–809.

(2) Peters, C.; Buller, R. M. Industrial Application of 2-Oxoglutarate-Dependent Oxygenases. *Catalysts* **2019**, *9*.

(3) Herr, C. Q.; Hausinger, R. P. Amazing Diversity in Biochemical Roles of Fe(II)/2-Oxoglutarate Oxygenases. *Trends in Biochemical Sciences*. 2018, pp 517–532.

(4) Zwick, C. R.; Renata, H. Harnessing the Biocatalytic Potential of Iron- and α-Ketoglutarate-Dependent Dioxygenases in Natural Product Total Synthesis. *Nat. Prod. Rep.* **2020**, *37*, 1065–1079.

(5) Matsuda, Y.; Bai, T.; Phippen, C. B. W.; Nødvig, C. S.; Kjærbølling, I.; Vesth, T. C.; Andersen, M. R.; Mortensen, U. H.; Gotfredsen, C. H.; Abe, I.; Larsen, T. O. Novofumigatonin Biosynthesis Involves a Non-Heme Iron-Dependent Endoperoxide Isomerase for Orthoester Formation. *Nat. Commun.* **2018**, *9*, 2587.

(6) Voss, M.; Honda Malca, S.; Buller, R. M. U. Exploring the Biocatalytic Potential of Fe/α-Ketoglutarate Dependent Halogenases. *Chem. Eur. J.* **2020**, *26*, 7336.

(7) Li, J.; Zhang, X.; Renata, H. Asymmetric Chemoenzymatic Synthesis of (−)-Podophyllotoxin and Related Aryltetralin Lignans. *Angew. Chemie Int. Ed.* **2019**, *58*, 11657–11660.

(8) Lazzarotto, M.; Hammerer, L.; Hetmann, M.; Borg, A.; Schmermund, L.; Steiner, L.;

Hartmann, P.; Belaj, F.; Kroutil, W.; Gruber, K.; Fuchs, M. Chemoenzymatic Total Synthesis of Deoxy-, Epi -, and Podophyllotoxin and a Biocatalytic Kinetic Resolution of Dibenzylbutyrolactones. *Angew. Chemie Int. Ed.* **2019**, *58*, 8226–8230.

(9) Chen H, Bong YK, Cabirol FL, Prafulchandra AG, Li T, Moore JC, Quintanar-Audelo M, Hong Y, Collier SJ, S. D. Biocatalysts and Methods for Hydroxylation of Chemical Compounds. US2017/0121744 A1. Codexis Inc., 2015.

(10) Smirnov, S. V.; Samsonova, N. N.; Kotliarova, V. A.; Rushkevich, N. Y.; Fedorina, E. A.; Sokolov, P. M.; Kolokolova, A. V.; Ogawa, J.; Hibi, M.; Shimizu, S.; Imabayashi, Y.; Suzuki, S.; Sugiyama, M. Method for Producing Hydroxylated Amino Acids. WO2011021717A2, 2011.

(11) Martinez, S.; Hausinger, R. P. Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-Dependent Oxygenases. *J. Biol. Chem.* **2015**, *290*, 20702–20711.

(12) Galonić, D. P.; Barr, E. W.; Walsh, C. T.; Bollinger, J. M.; Krebs, C. Two Interconverting Fe(IV) Intermediates in Aliphatic Chlorination by the Halogenase CytC3. *Nat. Chem. Biol.* **2007**, *3*, 113–116.

(13) Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. Structural Basis for Halogenation by Iron-and 2-Oxo-Glutarate-Dependent Enzyme WelO5. *Nat. Chem. Biol.* **2016**, *12*, 636–640.

(14) Dunham, N. P.; Chang, W. C.; Mitchell, A. J.; Martinie, R. J.; Zhang, B.; Bergman, J. A.; Rajakovich, L. J.; Wang, B.; Silakov, A.; Krebs, C.; Boal, A. K.; Bollinger, J. M. Two Distinct

Mechanisms for C-C Desaturation by Iron(II)- and 2-(Oxo)Glutarate-Dependent Oxygenases: Importance of α-Heteroatom Assistance. *J. Am. Chem. Soc.* **2018**, *140*, 7116–7126.

(15) Islam, M. S.; Leissing, T. M.; Chowdhury, R.; Hopkinson, R. J.; Schofield, C. J. 2-Oxoglutarate-Dependent Oxygenases. *Annu. Rev. Biochem.* **2018**, *87*, 585–620.

(16) Gao, S. S.; Naowarojna, N.; Cheng, R.; Liu, X.; Liu, P. Recent Examples of α-Ketoglutarate-Dependent Mononuclear Non-Haem Iron Enzymes in Natural Product Biosyntheses. *Nat. Prod. Rep.* **2018**, *35*, 792–837.

(17) Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chemie Int. Ed.* **2018**, *57*, 4143–4148.

(18) Hara, R.; Kino, K. Characterization of Novel 2-Oxoglutarate Dependent Dioxygenases Converting l-Proline to Cis-4-Hydroxy-l-Proline. *Biochem. Biophys. Res. Commun.* **2009**, *379*, 882–886.

(19) Ligibel, M.; Moore, C.; Bruccoleri, R.; Snajdrova, R. Identification and Application of Threonine Aldolase for Synthesis of Valuable α-Amino, β-Hydroxy-Building Blocks. *Biochim. Biophys. Acta, Proteins Proteomics* **2020**, *1868*, 140323.

(20) Robert S., P.; Lakshmi V., C.; Phillips, S. R.; Cyr, V. L. Inhibitors of Kynureninase. WO1995011878A1, 1996.

(21) Moroni, F.; Carpenedo, R.; Chiarugi, A. Kynurenine Hydroxylase and Kynureninase Inhibitors as Tools to Study the Role of Kynurenine Metabolites in the Central Nervous

System. *Adv. Exp. Med. Biol.* **1996**, *398*, 203–210.

(22) Phillips, R. S.; Iradukunda, E. C.; Hughes, T.; Bowen, J. P. Modulation of Enzyme Activity in the Kynurenine Pathway by Kynurenine Monooxygenase Inhibition. *Front. Mol. Biosci.* **2019**, *6*, 3.

(23) Jacobs, K. R.; Castellano-Gonzalez, G.; Guillemin, G. J.; Lovejoy, D. B. Major Developments in the Design of Inhibitors along the Kynurenine Pathway. *Curr. Med. Chem.* **2017**, *24*, 2471–2495.

(24) Wetzl, D.; Bolsinger, J.; Nestl, B. M.; Hauer, B. α-Hydroxylation of Carboxylic Acids Catalyzed by Taurine Dioxygenase. *ChemCatChem* **2016**, *8*, 1361–1366.

(25) Koketsu, K.; Shomura, Y.; Moriwaki, K.; Hayashi, M.; Mitsuhashi, S.; Hara, R.; Kino, K.; Higuchi, Y. Refined Regio- and Stereoselective Hydroxylation of l -Pipecolic Acid by Protein Engineering of l -Proline Cis -4-Hydroxylase Based on the X-Ray Crystal Structure. *ACS Synth. Biol.* **2015**, *4*, 383–392.

(26) Hartwig, J. F.; Larsen, M. A. Undirected, Homogeneous C−H Bond Functionalization: Challenges and Opportunities. *ACS Cent. Sci.* **2016**, *2*, 281–292.

(27) Laval, G.; Clegg, W.; Crane, C. G.; Hammershøi, A.; Sargeson, A. M.; Golding, B. T. Assembly of Polyamines via Amino Acids from Three Components Using Cobalt(III) Template Methodology. *Chem. Commun.* **2002**, 1874–1875.

(28) Bienert, S.; Waterhouse, A.; De Beer, T. A. P.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository-New Features and Functionality. *Nucleic Acids Res.* **2017**,

45, D313–D319.

(29) Reetz, M. T.; Carballeira, J. D. Iterative Saturation Mutagenesis (ISM) for Rapid Directed Evolution of Functional Enzymes. *Nat. Protoc.* **2007**, *2*, 891–903.

(30) Zabska, R.; Jakóbiec, T.; Dobek, R.; Wilimowski, M.; Barczyńska, J.; Kedzierska-Goździk, L.; Wojewódzki, W.; Rutkowska, M.; Duś, E.; Szelag, A. Synthesis and Pharmacological Properties of Derivatives of Alpha-Amino-Beta-(p-Chlorobenzoyl)-Propionic Acid and Alpha-Amino-Gamma-(p-Chlorophenyl)-Tetrahydrofuran-2-One. *Pol. J. Pharmacol. Pharm.* **1991**, *43*, 271–280.

(31) Simmons, J. M.; Müller, T. A.; Hausinger, R. P. FeII/α-Ketoglutarate Hydroxylases Involved in Nucleobase, Nucleoside, Nucleotide, and Chromatin Metabolism. *Dalton Trans.* **2008**, 5132–5142.

(32) Li, W.; Zhang, T.; Ding, J. Molecular Basis for the Substrate Specificity and Catalytic Mechanism of Thymine-7-Hydroxylase in Fungi. *Nucleic Acids Res.* **2015**, *43*, 10026–10038.

(33) Deng, Q.; Liu, Y.; Chen, L.; Xu, M.; Naowarojna, N.; Lee, N.; Chen, L.; Zhu, D.; Hong, X.; Deng, Z.; Liu, P.; Zhao, C. Biochemical Characterization of a Multifunctional Mononuclear Nonheme Iron Enzyme (PtlD) in Neopentalenoketolactone Biosynthesis. *Org. Lett.* **2019**, *21*, 7592–7596.

(34) You, Z.; Omura, S.; Ikeda, H.; Cane, D. E. Pentalenolactone Biosynthesis. Molecular Cloning and Assignment of Biochemical Function to PtIH, a Non-Heme Iron Dioxygenase of Streptomyces Avermitilis. *J. Am. Chem. Soc.* **2006**, *128*, 6566–6567.

(35) Matsuda, Y.; Iwabuchi, T.; Fujimoto, T.; Awakawa, T.; Nakashima, Y.; Mori, T.; Zhang, H.; Hayashi, F.; Abe, I. Discovery of Key Dioxygenases That Diverged the Paraherquonin and Acetoxydehydroaustin Pathways in Penicillium Brasilianum. *J. Am. Chem. Soc.* **2016**, *138*, 12671–12677.

(36) Meng, S.; Han, W.; Zhao, J.; Jian, X.; Pan, H.; Tang, G. A Six-Oxidase Cascade for Tandem C–H Bond Activation Revealed by Reconstitution of Bicyclomycin Biosynthesis. *Angew. Chemie Int. Ed.* **2018**, *57*, 719–723.

(37) Chen, T. Y.; Xue, S.; Tsai, W. C.; Chien, T. C.; Guo, Y.; Chang, W. C. Deciphering Pyrrolidine and Olefin Formation Mechanism in Kainic Acid Biosynthesis. *ACS Catal.* **2021**, *11*, 278–282.

(38) Zhou, J.; Kelly, W. L.; Bachmann, B. O.; Gunsior, M.; Townsend, C. A.; Solomon, E. I. Spectroscopic Studies of Substrate Interactions with Clavaminate Synthase 2, a Multifunctional R-KG-Dependent Non-Heme Iron Enzyme: Correlation with Mechanisms and Reactivities. *J. Am. Chem. Soc.* **2001**, *123*, 7388–7398.

(39) Chang, W. C.; Guo, Y.; Wang, C.; Butch, S. E.; Rosenzweig, A. C.; Boal, A. K.; Krebs, C.; Bollinger, J. M. Mechanism of the C5 Stereoinversion Reaction in the Biosynthesis of Carbapenem Antibiotics. *Science* **2014**, *343*, 1140–1144.

(40) Su, H.; Sheng, X.; Zhu, W.; Ma, G.; Liu, Y. Mechanistic Insights into the Decoupled Desaturation and Epoxidation Catalyzed by Dioxygenase AsqJ Involved in the Biosynthesis of Quinolone Alkaloids. *ACS Catal.* **2017**, *7*, 5534–5543.

(41) Mitra, S.; Holz, R. C. Unraveling the Catalytic Mechanism of Nitrile Hydratases. *J. Biol. Chem.* **2007**, *282*, 2020.

(42) Hopmann, K. H.; Himo, F. Theoretical Investigation of the Second-Shell Mechanism of Nitrile Hydratase. *Eur. J. Inorg. Chem.* **2008**, 1406–1412.

(43) Ondrechen, M. J.; Briggs, J. M.; Mccammon, J. A. A Model for Enzyme-Substrate Interaction in Alanine Racemase. *J. Am. Chem. Soc.* **2001**, *123*, 2830–2834.

(44) Watanabe, A.; Yoshimura, T.; Mikami, B.; Esaki, N. Tyrosine 265 of Alanine Racemase Serves as a Base Abstracting A-Hydrogen from L-Alanine: The Counterpart Residue to Lysine 39 Specific to D-Alanine. *J. Biochem* **1999**, *126*, 786.

(45) Clark, D. D.; Ensign, S. A. Characterization of the 2-[(R)-2-Hydroxypropylthio]Ethanesulfonate Dehydrogenase from Xanthobacter Strain Py2: Product Inhibition, PH Dependence of Kinetic Parameters, Site-Directed Mutagenesis, Rapid Equilibrium Inhibition, and Chemical Modification. *Biochemistry* **2002**, *41*, 2727–2740.

(46) Speciale, C.; Wu, H. Q.; Cini, M.; Marconi, M.; Varasi, M.; Schwarcz, R. (R,S)-3,4-Dichlorobenzoylalanine (FCE 28833A) Causes a Large and Persistent Increase in Brain Kynurenic Acid Levels in Rats. *Eur. J. Pharmacol.* **1996**, *315*, 263–267.

(47) Frey, R.; Hayashi, T.; Buller, R. M. Directed Evolution of Carbon–Hydrogen Bond Activating Enzymes. *Curr. Opin. Biotechnol.* **2019**, *60*, 29–38.

(48) Nakashima, Y.; Mori, T.; Nakamura, H.; Awakawa, T.; Hoshino, S.; Senda, M.; Senda, T.;

Abe, I. Structure Function and Engineering of Multifunctional Non-Heme Iron Dependent Oxygenases in Fungal Meroterpenoid Biosynthesis. *Nat. Commun.* **2018**, *9*, 104.

(49) Hayashi, T.; Ligibel, M.; Sager, E.; Voss, M.; Hunziker, J.; Schroer, K.; Snajdrova, R.; Buller, R. Evolved Aliphatic Halogenases Enable Regiocomplementary C-H Functionalization of an Added-Value Chemical. *Angew. Chemie Int. Ed.* **2019**, *58*, 18535.

(50) Mitchell, A. J.; Dunham, N. P.; Bergman, J. A.; Wang, B.; Zhu, Q.; Chang, W. C.; Liu, X.; Boal, A. K. Structure-Guided Reprogramming of a Hydroxylase to Halogenate Its Small Molecule Substrate. *Biochemistry* **2017**, *56*, 441–444.

(51) Yan, W.; Song, H.; Song, F.; Guo, Y.; Wu, C. H.; Sae Her, A.; Pu, Y.; Wang, S.; Naowarojna, N.; Weitz, A.; Hendrich, M. P.; Costello, C. E.; Zhang, L.; Liu, P.; Jessie Zhang, Y. Endoperoxide Formation by an α-Ketoglutarate-Dependent Mononuclear Non-Haem Iron Enzyme. *Nature* **2015**, *527*, 539–543.

(52) Boal, A. K.; Bollinger, J. M.; Chang, W. C. Assembly of the Unusual Oxacycles in the Orthosomycin Antibiotics. *Proc. Natl. Acad. Sci.* **2015**, *112*, 11989–11990.

(53) Chang, W. C.; Dey, M.; Liu, P.; Mansoorabadi, S. O.; Moon, S. J.; Zhao, Z. K.; Drennan, C. L.; Liu, H. W. Mechanistic Studies of an Unprecedented Enzyme-Catalysed 1,2-Phosphono-Migration Reaction. *Nature* **2013**, *496*, 114–118.

(54) Liao, H.-J.; Li, J.; Huang, J.-L.; Davidson, M.; Kurnikov, I.; Lin, T.-S.; Lee, J. L.; Kurnikova, M.; Guo, Y.; Chan, N.-L.; Chang, W. Insights into the Desaturation of Cyclopeptin and Its C3 Epimer Catalyzed by a Non-Heme Iron Enzyme: Structural Characterization and

Mechanism Elucidation. *Angew. Chemie Int. Ed.* **2018**, *57*, 1831–1835.