VIP Biocatalysis Very Important Paper

How to cite: *Angew. Chem. Int. Ed.* **2023,** *62,* e202300657
International Edition: doi.org/10.1002/anie.202300657
German Edition: doi.org/10.1002/ange.202300657

Structural and Mechanistic Studies on Substrate and Stereoselectivity of the Indole Monooxygenase *Vp*IndA1: New Avenues for Biocatalytic Epoxidations and Sulfoxidations

Julia Kratky, Daniel Eggerichs, Thomas Heine, Sarah Hofmann, Philipp Sowa, Renato H. Weiße, Dirk Tischler, and Norbert Sträter**

Abstract: Flavoprotein monooxygenases are a versatile group of enzymes for biocatalytic transformations. Among these, group E monooxygenases (GEMs) catalyze enantioselective epoxidation and sulfoxidation reactions. Here, we describe the crystal structure of an indole monooxygenase from the bacterium *Variovorax paradoxus* EPS, a GEM designated as *Vp*IndA1. Complex structures with substrates reveal productive binding modes that, in conjunction with force-field calculations and rapid mixing kinetics, reveal the structural basis of substrate and stereoselectivity. Structure-based redesign of the substrate cavity yielded variants with new substrate selectivity (for sulfoxidation of benzyl phenyl sulfide) or with greatly enhanced stereoselectivity (from 35.1 % to 99.8 % *ee* for production of (1*S*,2*R*)-indene oxide). This first determination of the substrate binding mode of GEMs combined with structure-function relationships opens the door for structure-based design of these powerful biocatalysts.

Introduction

Group E monooxygenases (GEMs) produce optically active epoxides as precursors for pharmaceuticals, agrochemicals, pigments and other biotechnologically valuable compounds.[1] These enzymes open up novel synthetic pathways, e.g. for the production of indigo without toxic by-products compared to traditional chemical synthesis.[2] In addition, the degradation of environmentally harmful species like styrene leads to an increased scientific focus on this class of enzymes.[3] GEMs accept a broad substrate spectrum ranging from substituted styrenes and indole derivatives to aryl alkyl sulfides. Rational design and directed evolution was employed to improve the performance of these monooxygenases or alter their substrate specificity.[4–7] Most of these studies focused on *Pseudomonas* GEMs, although other enzymes of this class exhibit higher catalytic efficiency.[8]

Based on their substrate specificity, GEMs have been subdivided into styrene monooxygenases (SMO) and indole monooxygenases (IMO).[6] The SMO from *Pseudomonas putida* (*Pp*StyA)[9] and the IMO from *Rhodococcus opacus* 1CP have been used in whole-cell as well as cell-free biotransformations.[3] The high enantioselectivity of these proteins towards formation of the (*S*)-enantiomer of e.g. methyl phenyl sulfoxide (MPSO) let them move into the spotlight of investigation for the production of enantiomerically pure products in industry.[3,10] The Gram-negative betaproteobacterium *Variovorax paradoxus* EPS contains two IMO enzymes: a two-component fusion protein with a reductase (*Vp*IndA2B) and the monooxygenase *Vp*IndA1.[6] For both enzymes, indole degradation or detoxification is probably the physiological role.[4,11] Nevertheless, *Vp*IndA1 was found to act as a highly efficient aryl alkyl sulfoxidase ($>3\ U\ mg^{-1}$).[6] Thus, we were interested to uncover the mechanistic details of this outstanding biocatalyst.

GEMs catalyze the enantioselective conversion of styrene and chemical analogues to the corresponding epoxides under the use of $FADH_2$, most likely using a mechanism similar to other flavoprotein monooxygenases.[12] An NADH-utilizing flavin reductase (designated as StyB for SMO) commonly provides $FADH_2$ for this reaction. Notably, in artificial systems, two-component enzymes can also utilize $FADH_2$ from the solution.[13] Thus, cofactor reduction by chemical methods is possible. So far, (*i*) the reduction of FAD by chemical agents (e.g. sodium thiosulfate or ruthenium complexes),[14,15] (*ii*) the use of NAD(P)H-mimics[16] or (*iii*) the electrochemical reduction of the cofactor[17] were applied successfully. In the presence of dioxygen, peroxyflavin is formed, which reacts with the electrophilic or nucleophilic substrates to form hydroxyfla-

[*] Dr. J. Kratky, Dr. R. H. Weiße, Prof. Dr. N. Sträter
Institute of Bioanalytical Chemistry, Leipzig University
Deutscher Platz 5, 04103 Leipzig (Germany)
E-mail: strater@bbz.uni-leipzig.de

D. Eggerichs, P. Sowa, Prof. Dr. D. Tischler
Microbial Biotechnology, Ruhr-Universität Bochum
Universitätsstr. 150, 44780 Bochum (Germany)
E-mail: dirk.tischler@rub.de

Dr. T. Heine, S. Hofmann, Prof. Dr. D. Tischler
Environmental Microbiology, TU Bergakademie Freiberg
Leipziger Str. 29, 09599 Freiberg (Germany)

vin and the monooxygenated product. Release of water and reduction of the released FAD closes the catalytic cycle.

A crystal structure of a GEM enzyme could so far only be determined for *Pp*StyA, but binding of FAD or substrates could not be achieved in this study.[18] Also, for the structurally related FMO families GFM and GGM, no productive substrate binding modes of FAD together with a substrate were reported so far, despite the determination of several enzyme structures. Substrate binding information is essential for enabling sophisticated structure-based enzyme design for biocatalytic transformations. We therefore set out to determine the crystal structure of *Vp*IndA1 in complex with FAD and substrates and used computational molecular mechanics (MM) force field methods to rationalize the basis of substrate- and stereoselectivity. The predictions were verified by kinetic data obtained by Michaelis–Menten and rapid mixing kinetics to solidify our conclusions.

Results and Discussion

*Crystal structures of Vp*IndA1 in the unliganded state and with bound substrates

Crystals of *Vp*IndA1 diffracted up to 1.5 Å resolution and contained one protein chain in the asymmetric unit (Table S3). *Vp*IndA1 consists of two subdomains. One subdomain has the classical fold of a CATH 3.50.50.60 domain, which is similar to a Rossmann fold and responsible for binding the FAD cofactor.[19] The substrate binding subdomain is positioned on top of the isoalloxazine ring (Figure 1). *Pp*StyA is the closest homolog with known structure, but the two enzymes share only 28.2 % sequence identity and differ by an rms deviation of 1.78 Å (Figure S1A). Only few residues of the substrate-binding cavity are conserved (Figure S1B). Heine et al.[8] described a signature motif of six residues to distinguish SMO from

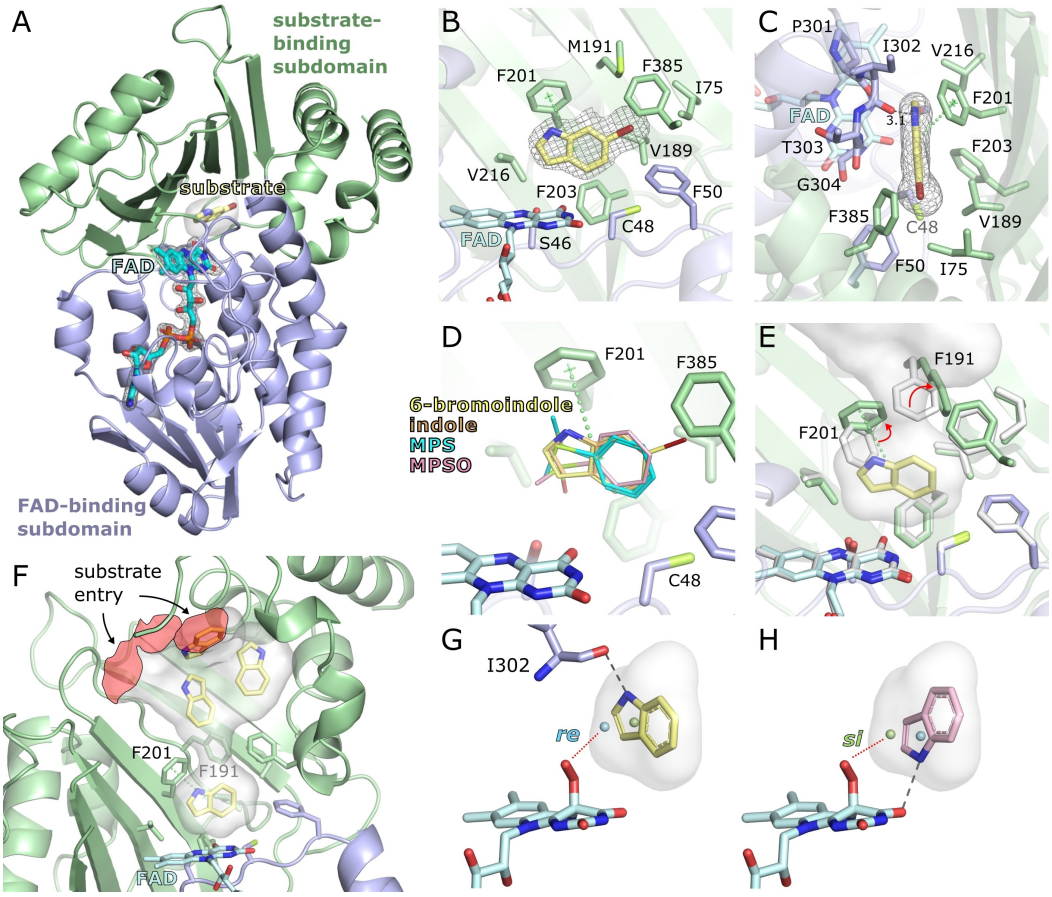


Figure 1. Substrate binding of *Vp*IndA1. A) The substrate-binding site is buried at the interface between the FAD-binding subdomain (blue) and the substrate-binding subdomain (green). The polder omit electron density of FAD is shown at a contour level of 5.0 σ. B) 6-bromoindole bound to the hydrophobic substrate-binding cavity of *Vp*IndA1^{F191M}. The polder omit electron density is contoured at 5.0 σ. C) Side-view of the binding mode of 6-bromoindole. D) Superposition of the binding modes of 6-bromoindole, indole, MPS (two alternate conformations), and MPSO. E) Superposition of unliganded *Vp*IndA1^{wt} (white) and *Vp*IndA1^{wt}×indole (colored). F201 and F191 change conformation to enable substrate binding to the hydrophobic cavity (white transparent). F) The substrate binding site is connected to a larger cavity or tunnel that provides access to the solvent. The channel is shown in transparent white for the crystal structure of *Vp*IndA1^{wt}×indole, in which three indole molecules have bound in this tunnel-like cavity. (G and H) Energy minimization of two alternative indole binding modes. The lowest energy binding mode (G, $E \equiv 0$ kcal mol^{-1}) of indole corresponds to the crystallographically observed pose. In the alternate binding mode (H, $E = 2.6$ kcal mol^{-1}), the indole ring is rotated such that the N1 can form a hydrogen bond to the FAD cofactor.

IMO enzymes. Interestingly, these residues are mostly not in direct contact with the substrate but form a second shell around the binding pocket (Figure S1CD).

Determination of crystal structures with a productive binding mode of substrate and FAD cofactor for the oxidation step appears to be difficult for many related flavoprotein monooxygenases, in particular for the families GEM, GFM and GGM, for which, to the best of our knowledge, no such structures are available despite successful structure analysis of several enzyme members. For *Vp*IndA1, determination of a substrate binding mode was also a challenge. Structures of the wildtype enzyme (*Vp*IndA1^{wt}) co-crystallized with indole, 6-bromoindole or MPS resulted in substrate binding around 13 Å away from the active site in a tunnel-like cavity of the protein (further results below). In the active site of a *Vp*IndA1^{wt}×indole complex obtained in the presence of 1 mM indole, weak density indicated low occupancy binding of the substrate and alternate conformations of L193 and F191 in the vicinity of F201. The latter residue appeared to sterically hinder substrate binding. We suspected that a conformational change of F201 and F191 might be necessary for substrate binding. F191 is conserved in IMOs among different species, whereas the SMO *Pp*StyA has a methionine at this position (Figure S1B). We therefore prepared an F191M variant of *Vp*IndA1 resembling *Pp*StyA, which is more active in the epoxidation and sulfoxidation of many substrates compared to *Vp*IndA1.[7]

Indeed, substrate binding to the active site of the F191M variant was observed for 6-bromoindole, MPS and the product methyl phenyl sulfoxide (MPSO) at a distance of around 4 Å to the reactive C4a of FAD (Figure 1B and C, Figure S2). The substrate cavity has a pronounced hydrophobic character formed at the FAD-distal side by side chains of the seven-stranded β-sheet of the substrate-binding subdomain and at the FAD-proximal side by backbone atoms of residues 301–304, FAD, F50 and F385. The only polar contact is formed between the NH-group of the indole ring and the carbonyl oxygen of I302 (Figure 1C). With this polar interaction and its flat elongated shape, the substrate cavity has a good complementary fit to indole, in agreement with the assumed physiological role of this enzyme as an indole monooxygenase.[4,11] A superposition of the observed substrate binding modes (Figure 1D) shows that the phenyl rings occupy similar binding modes that are characterized by a stacking interaction between the backbone amide of T303/G304 of the FAD-subdomain and the side chains of F201 and V189 of the substrate-binding subdomain (Figure 1C).

Finally, we also obtained a productive substrate binding mode for the wildtype enzyme by co-crystallizing *Vp*IndA1 with a concentration of 18 mM indole (Figure S2). Comparison of the apo and substrate-bound structures demonstrates that a concerted conformational switch of the side chains of F201 and F191 is necessary to enable substrate binding, as F201 in the apo structure would clash with the pyrrole ring of indole (Figure 1E). In the substrate-bound *Vp*IndA1^{wt} structure, the F191 side chain occupies a similar position as methionine in the F191M variant. Thus, the F191M variant facilitates binding of substrates since the methionine is not impeding the conformational change of F201. This is in accordance with lower K_M values found for enzyme variants with F191M substitution in Michaelis–Menten kinetics (further results below). As a result, substrate binding modes could be analyzed at lower substrate concentrations. However, for substrates such as indene or styrene, we could not observe significant occupancy of the active site even with the F191M variant.

A large hydrophobic funnel-shaped pocket or tunnel is present next to the small substrate binding cavity (Figure 1F). In many structures, co-crystallized ligands bound to this pocket. In crystal structures of the wildtype enzyme, binding was often observed only in this pocket, while the active site cavity was not or only partially occupied. The pocket is open to the solvent and has a narrow connection to the catalytic binding site. It is thus tempting to speculate that the tunnel might provide substrate or dioxygen access to the active site, when the FAD binding cavity is occupied by the cofactor. For monoamine oxidases (GGM family) and 3-hydroxybenzoate hydroxylase (GAM family) the presence of similar tunnel structures has been proposed to function in substrate access.[20] For *Pp*StyA highly cooperative binding of substrate and oxidized FAD was observed, leading to the suggestion that this cooperativity serves to maintaining the substrate in the vicinity of the active site until reduced FAD binds.[18] The substrate binding sites in the hydrophobic tunnel might help to store the substrates in the proximity of the catalytic substrate cavity, if the substrate has to leave the latter binding site for access of dioxygen.

The F191M mutation reduces turnover of different substrates by 50–80 % (Table 1) but has little influence on enantioselectivity (Table 2). Since the ground state conformations of F191 and F201 prevent substrate binding at close distance to the FAD peroxy intermediate, it can be assumed from the crystal structures that productive substrate binding requires a conformational rearrangement of these residues to a conformation of higher energy (Figure 1E), thus reducing substrate affinity. Possible catalytic functions of this affinity-decreasing ground state conformation are (*i*) to prevent binding and turnover of small low-affinity ligands (only the natural substrates exert enough binding energy to induce the conformational switch), (*ii*) that the dynamical switch increases the rate of exchange of products and substrates or (*iii*) that the tendency of F201 to change to the conformation where it is oriented into the substrate cavity helps to bring the bound substrate in closer proximity to the reactive hydroperoxy group of the activated FAD intermediate. To better understand the influence of the F191/F201 conformational change on *Vp*IndA1 catalysis, we studied enzyme kinetics by steady-state Michaelis–Menten and rapid-mixing kinetics.

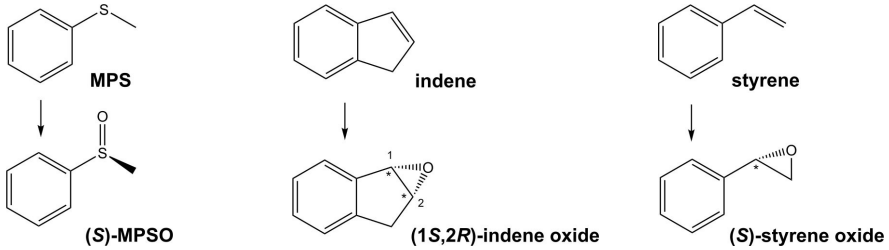
Michaelis–Menten kinetics show substrate inhibition and cooperative binding

In agreement with the difficulties we and others faced to determine co-crystal structures with substrates, the K_M-value

Table 1: Activity data of *Vp*IndA1 variants. The assay was composed of 20 mM Tris HCl pH 7.5, 10 mM BNAH, 10 vol.% EtOH, 650 U mL^{-1} bovine liver extract catalase, 50 μM FAD cofactor, 1.5 mM substrate and an appropriate amount of the respective *Vp*IndA1 variant. The assay was performed at 30 °C and product formation was determined photometrically for indole derivatives and via HPLC for the other substrates.

Variant	Activity [$\Delta E^*(\text{min mg})^{-1}$]		Specific activity [mU mg^{-1}]		
	Indole	6-bromoindole	MPS	Indene	BPS
*Vp*IndA1^{wt}	1.15 ± 0.03	0.20 ± 0.01	933 ± 13	733.04 ± 23.68	4 ± 2
F191M	0.49 ± 0.14	0.10 ± 0.03	227 ± 25	271.02 ± 11.05	2.7 ± 0.3
F191M/F201M	0.93 ± 0.05	0.28 ± 0.04	220 ± 24	–	4.3 ± 0.8
F191M/F201V	0.30 ± 0.07	0.10 ± 0.01	249 ± 17	99.99 ± 9.42	23.3 ± 1.1
F191M/F201A	0.41 ± 0.07	0.10 ± 0.05	107 ± 20	53.22 ± 14.03	36.8 ± 2.1
F191M/F201L/I302V	0.23 ± 0.01	0.41 ± 0.10	642 ± 32	1529.95 ± 5	2.8 ± 0.2
F191M/F201M/I302V	0.33 ± 0.01	0.59 ± 0.02	101 ± 27	59.14 ± 7.57	2.6 ± 0.2

Table 2: Stereoselectivity of *Vp*IndA1 variants. The assay was composed of 20 mM Tris HCl pH 7.5, 10 mM BNAH, 10 vol.% EtOH, 650 U mL^{-1} bovine liver extract catalase, 50 μM FAD cofactor, 2 mM of substrate and 3.5 μM *Vp*IndA1 variant. The assay was performed at 25 °C and product formation was determined via HPLC. The enantiomeric excess was measured by chiral GC-FID after 2 h of conversion.



ee [%]
conversion rate [mU mg^{-1}]

Variant	MPS	3-Br-MPS	indene	styrene	3-bromostyrene
wt	**99.06 ± 0.05**	**97.10 ± 0.24**	**35.09 ± 0.42**	**80.33 ± 0.03**	**69.19 ± 0.16**
	767.1 ± 43.7	569.2 ± 26.4	138.5 ± 24.7	305.8 ± 23.0	158.1 ± 8.2
F191M	**97.75 ± 0.03**	**98.36 ± 0.15**	**31.05 ± 0.70**	**78.43 ± 0.05**	**83.24 ± 0.40**
	141.2 ± 4.4	165.1 ± 2.7	48.1 ± 3.5	77.6 ± 8.3	48.4 ± 2.7
F191M/F201L/I302V	**99.02 ± 0.08**	**96.95 ± 0.28**	**99.80 ± 0.20**	**87.09 ± 0.06**	**78.24 ± 0.06**
	610.5 ± 10.0	542.7 ± 44.2	290.2 ± 4.0	256.6 ± 13.8	254.3 ± 27.8
F191M/F201A	**99.70 ± 0.12**	**96.90 ± 0.40**	**78.20 ± 12.4**	**87.20 ± 0.03**	**77.20 ± 0.05**
	512.0 ± 15.0	389.3 ± 9.2	23.7 ± 2.6	211.1 ± 5.0	35.9 ± 1.4

of *Vp*IndA1^{wt} for styrene is with 2.5 ± 0.5 mM quite high (Table S9). Michaelis–Menten kinetics of styrene turnover confirmed the expected increase in substrate affinity of the F191M variant by more than 6-fold lower K_M value, while the turnover rate dropped about 7-fold, which is consistent with the turnover numbers measured at high substrate concentrations (Table 1).

Substrate inhibition was observed for all tested variants and substrates except for benzyl phenyl sulfide (BPS) turnover by the F191M/F201A variant. Substrate inhibition is a frequent phenomenon in enzyme catalysis and it is often linked to additional substrate binding sites, that sterically block substrate binding or product release.[20] The presence of multiple substrate binding sites in the hydrophobic channel providing access to the catalytic site suggests a similar inhibition mechanism for *Vp*IndA1 (Figure 1F). One indole molecule binds close to F191 and would block the entrance to the substrate binding site. In this position it might prevent product release and thus cause inhibition. In contrast to the K_M-value, the substrate inhibition constant was not affected by the F191M mutation for styrene oxidation. An additional change introduced with the F191M mutation was a pronounced S-shaped activation profile. We used a kinetic model including substrate inhibition and cooperativity for the determination of the kinetic parameters.[21] In addition to cooperative binding in oligomeric enzymes, sigmoidal substrate binding curves have also been observed in monomeric enzymes as a result of slow conformational transitions even in enzymes with a single substrate binding site.[22] Due to the promiscuous binding of the substrates at multiple sites of the hydrophobic funnel-shaped pocket next to the catalytic site, we speculate that the substrate might also act as an activating effector at one or more of these sites. The cooperativity is highest for *Vp*IndA1^{F191M} and *Vp*IndA1$^{F191M/F201A}$ ($n = 3.91 \pm 0.42$ and 2.35 ± 0.68, respectively), whereas the wild-type enzyme shows low cooperativity for styrene turnover but significant substrate activation for indene ($n = 1.14 \pm 0.04$ and 1.77 ± 0.17, respectively). A similar behavior was observed for the conversion of indene by *Vp*IndA1$^{F191M/F201L/I302V}$ ($n = 1.60 \pm 0.30$).

Rapid mixing kinetics indicate that the F191/F201 conformational switch facilitates product release

The molecular mechanisms of the styrene monooxygenase from *Pseudomonas putida* S12 (*Pp*StyA) and the natural fusion enzyme of monooxygenase and reductase from *Rhodococcus opacus* 1CP (*Ro*StyA2B) were already studied by stopped-flow kinetics.[14,18,23] Further detailed mechanistic studies were conducted on related flavoprotein monooxygenases such as *p*-hydroxybenzoate hydroxylase (PHBH).[24] Therefore, we focused on the effect of the mutations on key steps of *Vp*IndA1 catalysis in the present work. For *Pp*StyA, a five-step mechanism was proposed, where four steps involve oxidized flavin cofactors (FAD_{OOH}, FAD_{OH} or FAD_{ox}) and are thus spectroscopically accessible in stopped-flow experiments. Assuming a similar overall mechanism for *Vp*IndA1 (Table S10), the substrate can bind to the enzyme (k_1) after binding of FAD_{red}, which is followed by the reaction with dioxygen to form peroxy FAD (k_2). The rapid decay of FAD_{OOH} (k_3) accompanies product formation and gives rise to the highly fluorescent FAD_{OH} species. This species is stabilized by the bound product and changes its spectroscopic properties upon product release (k_4) before FAD_{ox} is yielded after release of water (k_5). As $FADH_2$ possesses no significant absorption above 350 nm, k_1 (substrate binding) cannot be observed spectrophotometrically. We observed spectral changes which fit a mathematical model with three consecutive reaction steps where the wavelength dependent changes relate flavin species in the last three mechanistic steps (k_3 to k_5). As a comparably fast rate of 147 s^{-1} was reported for k_2 of *Pp*StyA,[23] it is likely that this step was not visible in our experimental setup. To draw conclusions from the effects of different mutations on the enzymatic mechanism, we compared the rates for oxygenation, product release and FAD oxidation for five substrates by the most promising variants (Table S5).

After oxygen transfer from FAD_{OOH} to the substrate, the following product release may be promoted by the switch to the ground state conformation in the wildtype enzyme where F201 displaces the product in the catalytic pocket and F191 closes the entrance to the active site. For the *Vp*IndA1^{F191M} variant, the product release rate k_4 was found to be two-times slower for all tested substrates compared to *Vp*IndA1^{wt} (Table S5), supporting such a function of F191/F201 in product release. Furthermore, a significantly reduced oxidation rate k_5 of the FAD cofactor was observed for this variant. Since water release is the rate limiting step of the catalytic mechanism, this explains the overall reduced activities of this variant (Tables 1, S9, S10). At the same time, the conversion rate k_3 was found to be twice as fast in the case of styrene. This is in accordance with the lower K_M value of styrene turnover for this variant as easier substrate access facilitates a faster oxygenation (Figure S17). In contrast, for indene, the closest tested analogue to the natural substrate indole, or for MPS, k_3 is reduced for the *Vp*IndA1^{F191M} variant. These differences are likely due to the fact, that the F191M mutation does not only affect the affinity of substrate binding but may also cause less productive binding modes.

We also investigated rates of FAD cofactor reactions in the absence of substrates for comparison. *Vp*IndA1^{F191M} and *Vp*IndA1$^{F191M/F201L/I302V}$ have an open pocket where dioxygen can enter more freely causing a twice as fast rate of peroxy FAD formation than in *Vp*IndA1^{wt} (Table S11). However, the degree of FAD_{OOH} stabilization (k_6) varies significantly between all enzyme variants. While for *Vp*IndA1$^{F191M/F201L/I302V}$ a rapid decay of the intermediate was observed, k_6 is rather slow for *Vp*IndA1^{F191M}. This finding correlates with the overall activity of the enzyme variants where a slower conversion of all substrates was observed for *Vp*IndA1^{F191M} (Tables 1, S9).

*Molecular mechanics (MM) force field calculations rationalize the excellent stereoselectivity of Vp*IndA1, which is governed by the confined hydrophobic substrate binding pocket

An important basis for enzyme design is to rationalize the substrate, regio- and stereoselectivity of the catalyzed reactions. These selectivities are determined by the presence of catalytically productive low energy binding modes. Binding of the to be oxidized atom or bond in near contact to the FAD peroxy intermediate controls turnover, regio- and stereoselectivity. We determined the regio- und stereoselectivity for different substrates of epoxidation and sulfoxidation reactions (Table 2) and analyzed the structural basis of this selectivity. To assess the estimated binding energy and binding geometry of the reactants, the C4a-hydroperoxy FAD cofactor and the substrates were modeled in different binding modes and minimized using the Amber10 and EHT force fields for protein and ligands, respectively.

For 6-bromoindole and indole, the lowest energy substrate binding mode is unambiguously determined by the crystal structures. These structures predict formation of (2*R*,3*R*) indole oxide, probably with a high diastereomeric excess (Figure 1G, H, S3). Alternative flipped binding modes have estimated energy differences of >2.6 kcal mol^{-1} to the crystallographically observed ground state structure (Figure S3). Unfortunately, the stereochemical outcome cannot be determined experimentally as indole oxide is unstable and converts to indoxyl and indigo. Overall, the MM modelling is in good accordance with the crystallographic results (Figure S3).

For methyl phenyl sulfide (MPS), two alternate binding modes are observed in the crystal structure of *Vp*IndA1^{F191M}×MPS, in which the methyl group is located approximately in the plane formed by the rest of the molecule (Figure 1D). The force field calculation in the presence of FAD_{OOH} shows that the conformation and binding mode that has a suitable geometric arrangement for formation of (*S*)-methyl phenyl sulfoxide (MPSO) is present at the lowest energy (Figure 2). In contrast, formation of the *R*-configuration of the product is hindered by slightly less favorable geometry but mostly by the much too high energy (predicted $\Delta E > 10$ kcal mol^{-1}) of this binding mode. To orient the pro-(*R*) lone pair of the sulfur in close proximity to the reactive FAD-hydroperoxy group, the methyl group

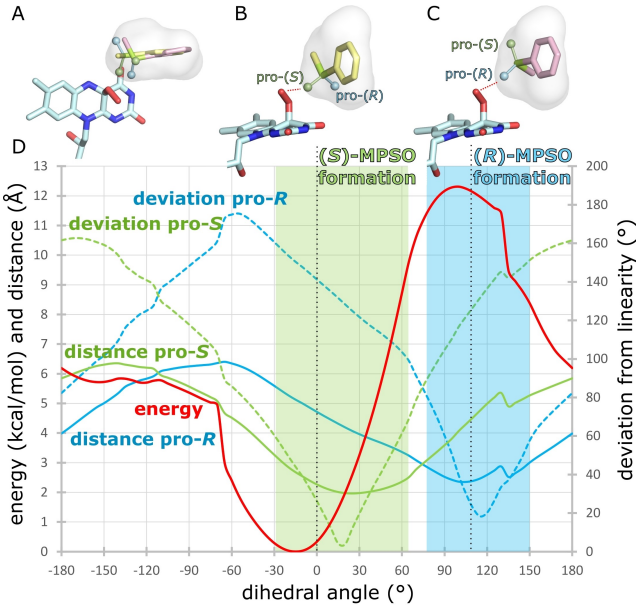


Figure 2. Torsion angle analysis of methyl phenyl sulfide binding. A) Superposition of the most productive binding modes resulting in the formation of (*S*)-MPSO (yellow) and (*R*)-MPSO (pink). B) Minimal distance of the pro-(*S*) lone pair electrons (green sphere) to the reactive FAD-hydroperoxy oxygen is obtained for torsion angles (C_{phenyl}–S bond) around 20°. C) Reasonable geometry for the positioning of the pro-(*R*) lone pair electrons (blue sphere) to yield the (*R*)-configured product is obtained at torsion angles around 110°. D) Plot of the distance of the lone pair electrons (LP, located at an arbitrary distance of 1.5 Å to the tetrahedrally configured sulfur atom) to the distal oxygen of the FAD-hydroperoxy group (as indicated by the red broken line in panels B and C). The deviation of the S–LP···O angle from linearity is shown as broken lines. The red line indicates the total energy relative to the lowest energy state.

has to rotate out of the plane formed by the phenyl ring, which causes unfavorable steric interactions. This finding is in excellent agreement with the experimentally observed formation of >99% *ee* for formation of (*S*)-MPSO (Table 2).

In contrast to MPS or indole, styrene can adopt two catalytically productive binding modes, that lead to different stereochemical outcomes, with a low energy difference (Figure S4). This agrees with the low *ee*-values of ≈30% observed for the wildtype enzyme or the F191M variant (Table 2).

Structure-based redesign of the substrate-binding cavity yields variants with greatly enhanced stereoselectivity or accepting novel substrates

To extend the catalytic scope of *Vp*IndA1, we prepared and analyzed several variants with mutations within the substrate binding cavity (Table 1 and 2). The mutations were selected on the basis of the crystallographically observed substrate binding modes and sequence alignments. F201 was chosen as it is in direct contact with the substrate near the atoms that are oxidized by the enzyme. I302 is located next to F201 and its carbonyl group is positioned to directly interact with substrates. Both amino acids are highly conserved among different GEMs (Figure S5). In addition, all variants contained the F191M mutation to facilitate substrate binding in the active site for structural analysis. The purified enzyme variants were characterized for their turnover of indole, 6-bromoindene, BPS, MPS and indene (Tables 1, 2, S9, and S10) and for the stereoselectivity (Table 2).

With the variants studied in this work, it was not possible to improve the transformation of MPS compared to the wildtype enzyme (Tables 1 and 2). In contrast, turnover of BPS was increased almost 10-fold by the F201A mutation (at 30 °C). The achieved activity is in the same range as the highest determined rates reported previously for other GEMs (IndA from *Gemmobacter nectariphilus* and IndA1 from *Streptomyces auratus* also reach 36 mU mg^{-1}) (Table 1).[8] While an interpretable co-crystal structure with the substrate BPS could not be obtained, co-crystallization of *Vp*IndA1$^{F191M/F201A}$ with the product benzyl phenyl sulfoxide (BPSO) as a racemic mixture resulted in a crystal structure with (*S*)-BPSO bound to the substrate cavity (Figure 3). Unfortunately, the *ee* value for this reaction could not be determined. The binding mode of the phenyl group resembles that in the *Vp*IndA1^{F191M}×MPS structure, while the benzyl substituent occupies a cavity that was generated by the removal of the F201 phenyl ring in the F201A variant.

We further analyzed the turnover of indole and 6-bromoindole to indigo and 6,6′-dibromoindigo (tyrian purple), respectively. *Vp*IndA1 converts indole well, while conversion of 6-bromoindole is less effective (Table 1). With the *Vp*IndA1$^{F191M/F201M/I302V}$ variant it was possible to improve the production rate of 6,6′-dibromoindigo almost 3-fold. With this activity, production of the high value dye tyrian purple would be about twice as efficient as the biotransformation with StyA from *Gordonia polyisoprenivorans*.[25]

A remarkable improvement is observed for the epoxidation of indene for the *Vp*IndA1$^{F191M/F201M/I302V}$ variant (Table 1). In addition to a significant activity increase (2-

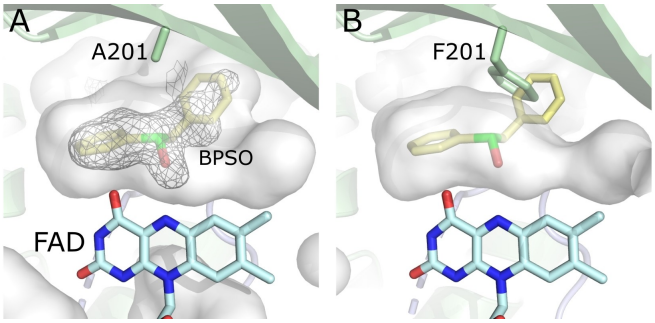


Figure 3. Binding mode of BPSO to the designed variant F191M/F201A for BPS sulfoxidation. A) Crystal structure of the *Vp*IndA1$^{F191M/F201A}$×-BPSO variant. The F201A mutation generates space for the benzyl group. BPSO was refined to an occupancy of 0.7 and the polder omit electron density map is shown at a contour level of 4.0 σ. Weak density of the benzyl ring indicates flexibility. B) The F201 residue of the wildtype enzyme would clash with the benzyl group of BPSO.

fold compared to the wildtype enzyme and 4- to 6-fold relative to the F191M variant), the *ee*-value for formation of (1*S*,2*R*)-indene oxide is dramatically enhanced from 35.1 % to 99.8 % (Table 2). The molecular origin of this improvement in the stereoselectivity is not readily explained by our simple geometric model of substrate orientation and estimation of relative binding energy alone (Figure S6). However, the analysis shows that due to a narrowing of the substrate pocket near the F201L mutation site, the indene molecule in the *si*-face orientation (Figures S6F and S7) rotates to achieve a low-energy binding pose. As a result, the nucleophilic double-bond of indene is positioned closer to the plane of the isoalloxazine ring of FAD and further away from the carbonyl group of P301. QM/MM calculations on the reaction mechanism of *p*-hydroxybenzoate hydroxylase (PHBH) indicated an important role of the carbonyl group of this proline (corresponding to P301 in *Vp*IndA1) in interacting with the reactive OH-group of hydroperoxy-FAD.[26] The closer proximity of the reacting double bonds of indene (Figure S6) and also styrene (Figure S3B) to the carbonyl oxygen of P301 in the binding modes leading to the formation of the experimentally observed product stereoisomers explains the observed product configurations better than the distance to the distal hydroperoxy group alone.

Michaelis–Menten kinetics show that the positive effect of the F191M mutation on the K_M value is observed to a lesser extent for *Vp*IndA1$^{F191M/F201L/I302V}$ catalyzing indene turnover (Table S9, Figure S18). On the other hand, V_{max} is increased 3.7-fold by the triple mutation. The stopped-flow kinetic analysis demonstrates that the molecular origin of this improvement lies not in the otherwise rate limiting product and water release as the reaction rates observed by rapid mixing kinetics are indistinguishable from the wildtype (Table S10). However, compared to the F191M variant, proper product release is restored in *Vp*IndA1$^{F191M/F201L/I302V}$. Additionally, increased k_3 rates were observed for styrene and indene (+47 % and +6 %, respectively), while the rate for MPS remained unchanged and decreased for 3Br-MPS (−60 %). These differences indicate a change in substrate fit without changes in the product release mechanism.

For the *Vp*IndA1$^{F191M/F201A}$ variant, effects of the reaction rates were observed according to the conversion rates determined (Table 2). Compared to the wildtype enzyme, the rate-limiting rates are reduced for MPS, 3Br-MPS and styrene as is the overall activity of the enzyme variant (Table S10). For indene, k_3 and k_4 were found to be twice as fast although the overall conversion rate is lower than for the wildtype. Here, the open pocket may cause unproductive reactions where the substrate binds as effector and increases the rate of hydrogen peroxide formation without substrate conversion. For BPS, only two catalytic rates were observed. As F201 is replaced by an alanine and the benzyl moiety of BPS occupies the space of F201 (Figure 3), it is likely that the product release mechanism is altered for this variant causing an overall change in the observable enzymatic mechanism.

The products of indene biotransformation by *Vp*IndA1 are 2-indanone and indene oxide (Figure S8). Formation of 2-indanone has also been observed for related GEM enzymes[2,9] and was attributed to the enzymatic activity of other enzymes in the host cell (for whole-cell biocatalysis or crude cell extracts)[2] or spontaneous decomposition.[10] However, several other studies identified only the epoxide as final product.[7,15,25,27] We determined that purified *Vp*IndA1 produces approximately 386 μM of 2-indanone as well as 151 μM indene oxide within 5 min of reaction time from an initial indene concentration of 2.5 mM. In the presence of *N*,*N*-diethyldithiocarbamate (DTC), which converts indene oxide to stable thioesters,[25] 2-indanone is not formed from indene oxide in the absence of the enzyme (Figure S9), whereas upon oxidation of indene with *Vp*IndA1, significant formation of 2-indanone is observed even in the presence of DTC (Figure S9). This result may be explained by *Vp*IndA1-catalyzed formation of 2-indanone. Alternatively, the enzyme prevents derivatization of the indene oxide by DTC while the product is bound in the active cavity or other enzyme pockets allowing for spontaneous conversion of indene oxide to 2-indanone.

All variants showing enhanced turnover rate for the substrates investigated in this work contain the F191M mutation. Although this mutation was introduced to facilitate substrate bound crystallization, rapid mixing kinetics highlight the importance of the interplay of residues in position 191 and 201. For the *Vp*IndA1$^{F191M/F201M/I302V}$ variant, it is likely that the F191M mutation is required for optimal performance as the triple mutation assembles partially the amino acid composition of a styrene monooxygenase. Nevertheless, it is still possible that some variants are even more active without the F191M mutation, as this mutation alone reduces activity compared to the wildtype enzyme (Tables 1, S9).

Conclusion

In this work, we could obtain first structural information on substrate binding of a group E flavoprotein monooxygenase, the indole monooxygenase *Vp*IndA1. The substrate cavity has a flat elongated shape that perfectly fits the bicyclic indole scaffold. A stacking interaction with the backbone amide of T303/G304 dominates the substrate binding mode together with the steric fit to the confined hydrophobic cavity. This binding mode is preserved in other substrates such as indene, styrene or aromatic thioethers. Based on the structural data, enzyme kinetics and MM modelling strategies, we could rationalize the substrate- and stereoselectivity of *Vp*IndA1 and uncover key steps of the molecular mechanism. For substrate binding and product release, a conformational rearrangement of residues in position 191 and 201 is required. In the wildtype enzyme, the coupled F191/F201 conformational switch might act as a gatekeeper to prevent binding and oxidation of smaller, less affine substrates or it might facilitate product release to increase the speed of catalytic turnover. The observed reduction of the k_{cat} value for the F191M mutation supports the latter function.

We used this structural and mechanistical information to achieve a substantial improvement of substrate acceptance

(for BPS turnover) or stereoselectivity (for highly diastereoselective production of (1*S*,2*R*)-indene oxide) with only a limited number of mutations. The fundamental understanding of *Vp*IndA1 catalysis presented in this work opens the door for further rational design of enzyme variants among the family of flavoprotein monooxygenases. Hereby, MM methods would result in even more efficient design protocols for engineering chemo-, regio- and stereoselectivity. Additionally, the evaluation of larger datasets and machine-learning algorithms may result in further information for guiding enzyme design. Our data thus further paves the way for rational design and investigation of new opportunities of GEMs in the field of green chemistry as well as white biotechnology to produce synthetically highly valuable epoxides and sulfoxides.

Acknowledgements

The authors thank M. Schwarze for assistance during the expression and purification of the mutants. We appreciate the funding of this project by the European Social Fund and the Saxonian Government (SMWK TG70 and GETGEO-WEB: 100101363). Furthermore, D. Eggerichs was supported by the Deutsche Bundesstiftung Umwelt (20019/625) and the Deutsche Forschungsgemeinschaft (RTG 2341 MiCon). We thank the MX Laboratory at the Helmholtz Zentrum Berlin (BESSY II) for beam time as well as for travel support. We also acknowledge the EMBL beamlines of the DESY synchrotron in Hamburg for beamtime. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The coordinates and reflection data of the crystal structures described in this manuscript are available in the Protein Data Bank. Further primary data are shown in the Supporting Information and are available from the authors upon request.

Keywords: Biocatalysis · Flavoenzyme · Monooxygenase · Protein Crystallography · Structure-Based Enzyme Design

[1] a) M. Breuer, K. Ditrich, T. Habicher, B. Hauer, M. Kesseler, R. Stürmer, T. Zelinski, *Angew. Chem. Int. Ed.* **2004**, *43*, 788; *Angew. Chem.* **2004**, *116*, 806; b) C. E. Paul, D. Eggerichs, A. H. Westphal, D. Tischler, W. J. van Berkel, *Biotechnol. Adv.* **2021**, *51*, 107712; c) J. Dong, E. Fernandez-Fueyo, F. Hollmann, C. E. Paul, M. Pesic, S. Schmidt, Y. Wang, S. Younes, W. Zhang, *Angew. Chem. Int. Ed.* **2018**, *57*, 9238; *Angew. Chem.* **2018**, *130*, 9380.

[2] K. E. O'Connor, A. D. Dobson, S. Hartmans, *Appl. Environ. Microbiol.* **1997**, *63*, 4287.

[3] D. Tischler, R. Kermer, J. A. D. Groening, S. R. Kaschabek, W. J. H. van Berkel, M. Schloemann, *J. Bacteriol.* **2010**, *192*, 5220.

[4] T. Heine, W. J. H. van Berkel, G. Gassner, K.-H. van Pee, D. Tischler, *Biology* **2018**, *7*, 42.

[5] a) M. L. Corrado, T. Knaus, F. G. Mutti, *ChemBioChem* **2018**, *19*, 679; b) H. Lin, D.-F. Tang, A. A. Q. Ahmed, Y. Liu, Z.-L. Wu, *J. Biotechnol.* **2012**, *161*, 235; c) A. A. Qaed, H. Lin, D.-F. Tang, Z.-L. Wu, *Biotechnol. Lett.* **2011**, *33*, 611; d) C. Tan, X. Zhang, Z. Zhu, M. Xu, T. Yang, T. Osire, S. Yang, Z. Rao, *Microb. Cell Fact.* **2019**, *18*, 12; e) J. Nikodinovic-Runic, L. Coulombel, D. Francuski, N. D. Sharma, D. R. Boyd, R. M. O. Ferrall, K. E. O'Connor, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 4849.

[6] D. Tischler, R. Schwabe, L. Siegel, K. Joffroy, S. R. Kaschabek, A. Scholtissek, T. Heine, *Molecules* **2018**, *23*, 809.

[7] L. J. Gursky, J. Nikodinovic-Runic, K. A. Feenstra, K. E. O'Connor, *Appl. Microbiol. Biotechnol.* **2010**, *85*, 995.

[8] T. Heine, A. Scholtissek, S. Hofmann, R. Koch, D. Tischler, *ChemCatChem* **2020**, *12*, 199.

[9] H. D. Dunn, T. Curtin, M. A. O'riordan, P. Coen, P. M. Kieran, D. M. Malone, K. E. O'Connor, *FEMS Microbiol. Lett.* **2005**, *249*, 267.

[10] H. Toda, R. Imae, N. Itoh, *Tetrahedron: Asymmetry* **2012**, *23*, 1542.

[11] Q. Ma, X. Zhang, Y. Qu, *Front. Microbiol.* **2018**, *9*, 2625.

[12] a) M. M. E. Huijbers, S. Montersino, A. H. Westphal, D. Tischler, W. J. H. van Berkel, *Arch. Biochem. Biophys.* **2014**, *544*, 2; b) P. Chenprakhon, T. Wongnate, P. Chaiyen, *Protein Sci.* **2019**, *28*, 8; c) E. Romero, J. R. Gómez Castellanos, G. Gadda, M. W. Fraaije, A. Mattevi, *Chem. Rev.* **2018**, *118*, 1742; d) R. Teufel, *Arch. Biochem. Biophys.* **2017**, *632*, 20.

[13] W. Zhang, F. Hollmann, *Chem. Commun.* **2018**, *54*, 7281.

[14] D. Tischler, M. Schlömann, W. J. H. van Berkel, G. T. Gassner, *FEBS Lett.* **2013**, *587*, 3848.

[15] F. Hollmann, P.-C. Lin, B. Witholt, A. Schmid, *J. Am. Chem. Soc.* **2003**, *125*, 8209.

[16] C. E. Paul, D. Tischler, A. Riedel, T. Heine, N. Itoh, F. Hollmann, *ACS Catal.* **2015**, *5*, 2961.

[17] F. Hollmann, K. Hofstetter, T. Habicher, B. Hauer, A. Schmid, *J. Am. Chem. Soc.* **2005**, *127*, 6540.

[18] U. E. Ukaegbu, A. Kantz, M. Beaton, G. T. Gassner, A. C. Rosenzweig, *Biochemistry* **2010**, *49*, 1678.

[19] M. L. Mascotti, M. Juri Ayub, N. Furnham, J. M. Thornton, R. A. Laskowski, *J. Mol. Biol.* **2016**, *428*, 3131.

[20] B. Wu, *Drug Metab. Rev.* **2011**, *43*, 440.

[21] V. J. LiCata, N. M. Allewell, *Biophys. Chem.* **1997**, *64*, 225.

[22] C. M. Porter, B. G. Miller, *Bioorg. Chem.* **2012**, *43*, 44.

[23] A. Kantz, G. T. Gassner, *Biochemistry* **2011**, *50*, 523.

[24] a) S. Montersino, D. Tischler, G. T. Gassner, W. J. H. van Berkel, *Adv. Synth. Catal.* **2011**, *353*, 2301; b) B. Entsch, L. J. Cole, D. P. Ballou, *Arch. Biochem. Biophys.* **2005**, *433*, 297.

[25] T. Heine, C. Großmann, S. Hofmann, D. Tischler, *Biol. Chem.* **2019**, *400*, 939.

[26] L. Ridder, J. N. Harvey, I. M. C. M. Rietjens, J. Vervoort, A. J. Mulholland, *J. Phys. Chem. B* **2003**, *107*, 2118.

[27] a) H. Lin, J. Qiao, Liu, Wu, Zhong-Liu, *J. Mol. Catal. B* **2010**, *67*, 236; b) A. Schmid, K. Hofstetter, H.-J. Feiten, F. Hollmann, B. Witholt, *Adv. Synth. Catal.* **2001**, *343*, 732.

Manuscript received: January 13, 2023
Accepted manuscript online: February 10, 2023
Version of record online: March 16, 2023