Accessing Enantiopure Epoxides and Sulfoxides: Related Flavin-Dependent Monooxygenases Provide Reversed Enantioselectivity

Thomas Heine,[a] Anika Scholtissek,[a] Sarah Hofmann,[a] Rainhard Koch,[b] and Dirk Tischler*[a, c]

Enantioselective organic compounds are of major importance for the chemical and pharmaceutical industry. Flavin-dependent group E monooxygenases, composed of monooxygenase and reductase, are known to perform epoxidation of substituted alkenes as well as sulfoxidation in a regio- and enantioselective fashion. Group E is divided into styrene monooxygenases (SMO) and indole monooxygenases (IMO). Hitherto mainly SMOs have been characterized. In this study, we assayed 31 monooxygenases from both types, while 23 of which showed activity. They almost exclusively produced (S)-stereoe oxide at high enantio-meric excess with maximum activities of 0.73 μmol min⁻¹ mg⁻¹ (kcat = 0.54 s⁻¹). In case of sulfoxidation, we found that the enantioselectivity is contrary between both types. IMOs preferably produce the (S)-enantiomer while SMOs have a tendency to produce the (R)-enantiomer. Sequence analysis and molecular docking of substrates allowed identifying fingerprint motives: SMO N46-V48-H50-Y73-H76-S96 and IMO S46-Q48-M50-V/I73-I76-A96. These form an essential part of the active site while the loop (AS44-51) interacts with the co-substrate and other amino acids direct the substrate. The motives clearly distinguish group E monooxygenases and define the enantioselectivity and thus direct biotechnological applications. Two-hour biotransformations with several sulfides in conjunction with upscale experiments (10 and 100 mg scale) resulted in the identification of promising candidates for the realization of biocatalytic processes.

Introduction

Enzyme driven catalysis is of increasing interest for biochemical and industrial applications.[1,2,3] Biocatalysts can work under mild conditions, they utilize non-toxic cofactors and often produce less waste, or side products, as they are often chemo-, regio- and enantioselective towards a substrate or product. Some enzymes allow reactions that are not possible under regular conditions. This includes monooxygenases that introduce a single oxygen atom from O₂ into an organic substrate, which is formally spin-forbidden. They employ an (in)organic cofactor, like heme, copper or flavins to activate molecular oxygen for catalysis.[4] Especially flavin-dependent monooxygenases are attractive, as they are involved in various biological processes and oxygenation reactions, which are of relevance for industrial applications.[5] For instance, they catalyze hydroxylation, Baeyer-Villiger oxidation, sulfoxidation, epoxidation and (de)halogenation reactions.[6] Flavin-dependent monooxygenases were categorized into eight groups according to their fold and function.[7] Four of these groups comprise two-component monooxygenases, which require for an associated reductase that provides the reduced flavin co-substrate via consumption of reduced nicotinamide adenine dinucleotide. These two-component monooxygenases have attracted much attention in the last decade.[8,9]

Among these, group E monooxygenases (GEMs) perform epoxidation of substituted alkenes and sulfoxidation in a regio- and enantioselective fashion (Scheme 1). GEMs were initially designated as styrene monooxygenases (SMO) consisting of two subtypes (E1 and E2).[10] However, recent studies provided evidence for a distinct physiological role of both subgroups and resulted in a reclassification into SMOs (previously E1-type) and indole monooxygenases (IMO; previously E2-type).[9,11–14]

There is already a huge set of substrates that are known to be converted by these enzymes in a regio- and enantioselective fashion, which illustrates their biotechnological potential (Scheme 1).[15–18,19,20–31] Upscaling of biotransformations and liquid- two-phase systems[26,28,29,32–35,36,37,40] as well as enzyme cascade applications[20,41] and co-substrate regeneration experiments[18,23,26,32,42,43] were performed to fathom their applicability in an industrial process. As a consequence, they are
Currently considered as candidates for biotechnological processes. However, limited activity, stability, and conversion of bulky substrates have been issues within this group of flavoproteins. Furthermore, changing or even optimizing the enantioselectivity of these biocatalysts is of high interest. Random and rational design of GEMs showed that there is space for optimization of these enzymes with regard to activity, specificity, and enantioselectivity. Artificial fusion proteins were generated and several mutations were identified that changed and/or improved biocatalytic properties. These resulted in GEMs with, inter alia, up to 12-fold higher activity and improved stability, change in substrate preference towards bulkier substrates, reversed enantioselectivity and higher catalytic efficiency. However, up to now it was not possible to predict activity or even enantioselectivity of these enzymes. This is needed to provide access to functional biocatalysts for novel processes to be designed. Most noticeable is the so far outstanding selectivity during epoxidation leading to only (S)-styrene oxide. However, just recently an (R)-selective epoxidase was discovered showing sequence homology to GEMs.

Further, some of the introduced mutations that resulted in improved GEMs occur naturally in homologous proteins. As this indicates, that a high biotechnological potential might be yet unexplored, we cloned and expressed 31 putative SMOs and IMOs to span the whole range of GEMs. Selected candidates were characterized in detail and besides styrene, we applied 11 sulfides as model substrates for the biosynthesis of chiral sulfoxides (Scheme 2), which are of great value for industrial, pharmaceutical, and agrochemical applications (reviewed in Refs [45, 48]). Sulfide biotransformations were scaled up using the most promising SMOs together with the nicotinamide cofactor mimic 1-benzyl-1,4-dihydronicotinamide (BNAH) for regeneration of reduced FAD.

Results

Exploration of GEMs

Previous phylogenetic analysis of GEMs already showed the clustering into the two designated sub-types: SMOs and IMOs. Due to the large number of uncharacterized, putative GEMs, the question arose whether there might exist an unexploited source of biotechnologically interesting catalysts. By expanding the number of utilized sequences for the phylogenetic tree, a more detailed view on group E is now available (Figure 1). Therewith, it was possible to identify and select candidates from various branches for a screening. In this study, 23 of the selected GEMs showed activity in an indigo-blue assay from which 18 were expressed from PCR amplified...
or synthesized genes, purified and further characterized. A detailed list of the selected enzymes, including abbreviations and conducted experiments can be found in Table S2. The inactivity of the other candidates can have various reasons (discussed in the supporting information section).

Active GEMs require for regeneration of reduced FAD. Under cell-free conditions, we used either a reductase (PfStyB, RoStyBart) or 1-benzyl-1,4-dihydronicotinamide (BNAH) as stoichiometric reductant, respectively. The reaction conditions were applied as described previously. A condensed
excerpt of the screening results is presented in Table 1. The detailed compilation of the generated data is deposited in Dataset S1 in the supporting information section.

All assayed GEMs converted 1 into the product (S)-styrene oxide. The epoxidation activities (initial rates) of these are in a range between 0.015 to 0.734 μmol min⁻¹ mg⁻¹ (Table 1). However, there is no clear correlation between type and taxonomy of the bacterial source towards the activity, which is reached by the enzymes. All IndA2B-like variants show at least an order of magnitude lower activity for the epoxidation when compared to the fastest monoxygenases, while the highest activity towards styrene was measured for NfIndA1. IndA- and IndA1-like enzymes reach comparable activities as StyA-like variants, and fast epoxidases are usually more active sulfoxidases.

Exceptions are variants from V. paradoxus EPS that perform very well in sulfoxidation although the activity with styrene is rather low. Interestingly, up to 25-times higher activities are reached for sulfoxidation by these enzymes.

With this screening it was possible to identify active GEMs that were further analyzed in long-term biotransformations.

Two-Hour Biotransformation Assays

Two-hour biotransformations of styrene and aryl alkyl sulfides were done to gain information about the long-term performance and the substrate scope of GEMs (Figure 1, Table 1 and Dataset S1). Thereby, no overoxidation to sulfoxones was observed when applying sulfides as substrates.

The biotransformation rates over two hours of 1, 2 and 11 decreased (2 to 37-fold) compared to the initial rates of each enzyme, what might be caused by product inhibition or instability of the enzymes (Dataset S1).

Except for VpIndA2B, with only 45%, the enantiomeric excess (ee) for the production of styrene oxide is above 87%. As known from previous studies, GEMs are able to and almost exclusively produce enantiopure (S)-epoxides. The most remarkable observation in this study is that IMOs usually produce the (S)-enantiomeric sulfoxides while SMOs have a preference to produce the (R)-enantiomers. GsStyA and PaStyA represent an exception, as mainly the (S)-enantiomeric sulfoxides are produced but only with a low ee. In general, the ee of most StyA-like enzymes is not very promising for an application and barely raises up to 70%. However, GpStyA, which is from a phylogenetic point of view an outsider within the SMO-type, can produce (R)-sulfoxides with high purity.

IMOs reach in general higher enantiomeric excess for sulfoxidations compared to SMOs and almost solely produce the (S)-enantiomer. Exceptions in this pattern were only observed for RoIndA2B and IMOs from V. paradoxus EPS, which produce the (R)-enantiomers of 2* and 3*, respectively. These exceptions need further investigations and here especially structural investigations will need to be performed to understand this. R. opacus 1CP has an SMO as well as an IndA1/
IndA2B system. However, independent from the type, they show a rather low ee in sulfoxidation reactions. The IMOs perform poorly compared to the similar system in the β-proteobacterium *V. paradoxus*. Especially IndA2B systems of actinobacteria have low activity and stability when purified and most of them were therefore not considered for the two-hour biotransformation assays.

The extended reaction period together with the application of various substrate allowed identifying promising GEMs for biotechnological applications. Two candidates were subsequently utilized for a (semi)-preparative synthesis.

Preparative Synthesis of Sulfoxides

Preparative synthesis experiments were conducted for production of 5* with purified *GpStyA* and *DalIndA* as catalysts, respectively (Table 2). Both enzymes offer a comparatively high activity together with a promising stereoselectivity. 5 was selected as model-substrate as it contains a relative large cavity on the putative positioning of the substrates in the active site. Understanding this might allow for the generation of active, stable and stereoselective monooxygenases. For that purpose, it would be helpful to include sequence and structural information.

Molecular Docking of Ligands Into the Putative Substrate-Binding Site of an SMO

The finding of a reversed enantioselectivity of different GEMs raised the question which key residues or motives are relevant for substrate positioning in the active site. While the location of the isoalloxazine is principally in accordance with the prediction, the adenine part of FAD is probably not correctly docked. Comparison to PHBH and MBHB suggest a stretched binding mode and not a stacked binding as found in our model. This is probably due to the limitations of molecular docking on a static protein structure and might be circumvented by molecular dynamics simulations that allow the adenine part to find its correct positioning.

Moreover, the position of the FAD co-substrate raises questions. While the location of the isoalloxazine in the active site is closer situated to the FAD binding site and the residues Y73, H76 and S96 that have been shown to impact the catalytic activity of the enzyme. In our model, these residues are about 4 to 5 Å away from the substrate, which is in accordance with previous predictions. *PpStyA* S12 has a large cavity on the surface within the dimer interface that is supposed to bind one FAD molecule per monomer. Beyond the isoalloxazine ring of the flavin co-substrate, each monomer of the SMO possesses a tight channel that leads to a hydrophobic cavity representing the putative substrate-binding site (Figure 2). Three binding modes were observed for styrene and the 11 sulfides (Scheme S1). Except for 6, best docking results for para-substituted sulfides indicate binding with the para-substituent facing to the entry of the substrate channel (binding mode 2). This is contrary to binding mode 1 and might be due to steric hindrance or intermolecular forces pushing the *para*-substituent out of the pocket. Bulky substrates as 10 and 12 fit well into the active site with the sulfur pointing away from the FAD co-substrate, what might explain rather slow conversion rates.

Moreover, the position of the FAD co-substrate raises questions. While the location of the isoalloxazine is principally in accordance with the prediction, the adenine part of FAD is probably not correctly docked. Comparison to PHBH and MBHB suggest a stretched binding mode and not a stacked binding as found in our model. This is probably due to the limitations of molecular docking on a static protein structure and might be circumvented by molecular dynamics simulations that allow the adenine part to find its correct positioning.

However, the docking studies allowed getting an indication on the putative positioning of the substrates in the active site.

### Table 2. Cell-free biotransformation of S by *GpStyA* and *DalIndA* in 10 and 100* mg scale.

| Enzyme | Conversion [%] | ee [%] | Total product [mg] | Activity [U mg⁻¹] | Product-catalyst ratio [mg mg⁻¹] | Productivity [mg mg⁻¹ h⁻¹] | TTN [min⁻¹] | Isolated yield [%] |
|--------|----------------|-------|-------------------|------------------|-------------------------------|-----------------------------|--------------|-----------------|
| *GpStyA* | 100            | R 93 | 12.3              | 0.13             | 2.5                           | 1.5                         | 708          | 7.1             | 95              |
| *DalIndA* | 100            | S > 99| 12.7              | 0.11             | 1.8                           | 1.1                         | 521          | 5.2             | 51              |
| *DalIndA* | 100            | S > 99| 126.3             | 0.06             | 1.8                           | 0.6                         | 520          | 2.9             | 53              |

[a] Calculated from analytical data. [b] Enzyme activity = total product-catalyst amount · conversion time⁻¹ [µmol mg⁻¹ min⁻¹]. [c] Product-catalyst ratio = total product-catalyst amount [µmol mg⁻¹]. [d] Productivity = product-catalyst ratio · conversion time⁻¹ [µmol mg⁻¹ h⁻¹]. [e] TTN = total product-catalyst amount [µmol] / [f] TF = TTN · conversion time⁻¹ [min⁻¹].
Interestingly, Ab might explain that SMOs nonetheless produce (standing enantioselectivity. However, anomalies in the sulfoxidation enantioselectivity might also result in a different FAD-protein interaction. Further, GrStyA forms only (S)-sulfoxides although it contains a tyrosine at position 73. This is in accordance with biotransformations of 4 with the highly similar R10StyA.\[24] Also other SMOs like PaStyA produce mainly the (S)-enantiomer. Consequently, the reversed stereoselectivity cannot solely be attributed to residue 73 and the binding mechanism is more complex. Comparison of the substrate binding site between SMOs and IMOs revealed several differences in hydrophobicity and charge distribution that might also affect the orientation of the ligand (Table S1 and Dataset S1).\[71]

However, the distance of the substrates to the C4a-hydroperoxide is too long for interaction since loop 1 and 2 hinder access towards the cavity. Hence, we propose a structural rearrangement within this region. Different conformations and substantial movements upon flavin and substrate binding are indeed known for the structural homolog PHBH.\[71] A similar movement of loop 1 (Figure 2B) might allow access of FAD into the cavity (Figure 2A). If assuming a similar position of the isoaalloxazine ring of PHBH from P. aeruginosa in the PaStyA S12, the C4a-hydroperoxide then would be about 4 to 5 Å closer to the substrate and in a range that is reasonable for direct oxygen transfer. In accordance with that, mutations within loop 1 and loop 2 showed already an impact on the enzyme activity.\[22,46] It should be mentioned that the loops that interact with the isoaalloxazine are composed differently in SMOs and IMOs, which might also result in a different FAD-protein interaction.

**Comparison of the Substrate Cavity of SMOs and IMOs**

The results of the substrate docking, the multiple sequence alignment (Figure S2) and previous studies on mutation of GEMs were used to identify important amino acid residues relevant for catalysis. Interestingly, most of the highly conserved residues in GEMs are not located within the active site of the enzyme (Figure S3A). They are rather located within the FAD-binding domain and at the edge of the active site. Vice versa, if only considering residues that are highly conserved in the subtypes (SMOs and IMOs) then these are occupied by different amino acids (see alignment in Figure S2). Further, many of them are situated in or around the substrate-binding site (Figure S3B). It is likely that these amino acid changes affect the characteristics of the substrate binding cavity and biochemical properties of the enzymes (Table S1). This concerns in particular loop 1 as well as Y73, H76, and S96 which are involved in substrate binding.

Especially Y73 has an orientation that might cause non-covalent interaction with the sulfur atom of the substrates.\[22] Thus, we propose that a hydrogen bond is formed between the sulfur and the hydroxyl group of Y73 that assists in the positioning of the sulfides in SMOs so that the (R)-enantiomer is preferably produced. As styrene does not contain a sulfur atom, there is no similar interaction towards the vinyl side chain what might explain that SMOs nonetheless produce (S)-styrene oxide. Interestingly, AblIndA and VplndA2B do not offer this outstanding enantioselectivity.

However, anomalies in the sulfoxidation enantioselectivity were found that also require for further investigation (e.g. the stereoselectivity towards 3 in case of RoStyA, PStyA, VplndA1, and VplndA2B). Except for GpStyA and SfStyA, SMOs reach rather low ee values. Further, GrStyA forms only (S)-sulfoxides although it contains a tyrosine at position 73. This is in accordance with biotransformations of 4 with the highly similar R10StyA.\[24] Also other SMOs like PaStyA produce mainly the (S)-enantiomer. Consequently, the reversed stereoselectivity cannot solely be attributed to residue 73 and the binding mechanism is more complex. Comparison of the substrate binding site between SMOs and IMOs revealed several differences in hydrophobicity and charge distribution that might also affect the orientation of the ligand (Table S1 and Dataset S1).\[71]
Discussion

Accessing optically pure epoxides and sulfoxides via biocatalytic approaches is a current issue\(^{[32]}\) and several mono- and dioxygenases, as well as peroxxygenases and epoxide hydro- lases, have been employed to reach this goal.\(^{[40,46,74,75]}\) However, there are still hindrances concerning process stability, due to inactivation or degradation of the enzyme, limited substrate scope and/or insufficient product purity and quality. While biocatalysts are yet attractive for the generation of green processes, it is, therefore, necessary to look for alternative enzymes that can meet the requirements mentioned above.

Including this study, 48 GEMs have been investigated so far. However, previous studies focused mainly on SMOs and especially on those from *Pseudomonas* and *Rhodococcus* species. The phylogenetic analysis (Figure 1) reveals that therefore the insight into this enzyme group was narrowed, emphasizing its unexploited potential. This hypothesis is supported by the just recent identification of a GEM that is able to convert styrene into the \((R)\)-enantiomer of styrene oxide.\(^{[47]}\) An extended phylogenetic tree comprising the novel \((R)\)-selective epoxide and related enzymes show that GEMs can be found as a separate cluster within the branch of SMOs (Figure S1). The same can be found for the recently identified IndA from *Burkholderia* sp. IDO3 and its related monoxygenases.\(^{[37]}\) However, both novel GEMs cluster far away from other SMOs as well as IMOs.

Herein, we identified 23 enzymes that showed activity on indole, which is structurally similar to styrene. It is converted by all so far investigated GEMs and it can be reasoned that all representatives of group E will accept this substrate due to substrate promiscuity. This turns it into the most suitable screening substrate among GEMs as a rapid indigo formation can be determined due to epoxidase activity.

The highest activity for an SMO was determined for PtStyA with 2.1 mmol min\(^{-1}\) mg\(^{-1}\).\(^{[24]}\) However, this high activity rapidly decreases after the first 5 min due to product inhibition. Another StyA that originates from an uncultured bacterium was reported to have an activity of 0.8 mmol min\(^{-1}\) mg\(^{-1}\) in cell extract.\(^{[30]}\) This calculation based on the estimation of the StyA protein content in the cell extract from SDS-PAGE. Thus, it remains to prove this determined activity with purified enzyme. A comparison of all epoxidases with our data shows that the so far characterized GEMs from *Pseudomonas* and *Rhodococcus* have a rather week performance in epoxidation as well as sulfoxidation (Figure 1, Table 1 and Dataset S1).\(^{[17,27]}\)

The rates are one order of magnitude lower if compared to the most active enzymes of this study and the enantioselectivity is rather low in case of sulfoxidation (Table 1 and Dataset S1). For instance, the SMOs SfStyA and GpStyA exhibit higher activities than commonly used variants from *Pseudomonas* and *Rhodococcus* of up to 0.57 mmol min\(^{-1}\) mg\(^{-1}\). The most active IMOs reach similar or higher rates. In contrast, one-component IMOs are rather slow and often hard to express and purify. Further, they are often unstable as can be seen for the VpindA2B, which has remarkably high initial sulfoxidation activity but cannot maintain these over a long period.\(^{[38]}\) This instability might originate from the C-terminal fusion to the reductase component that could cause un- or misfolding. This could depend on the protein environment, e.g. in the cell or in a cell-free buffer system. However, artificially fused GEMs performed well in whole-cell systems, which might have a favorable effect on the stability.\(^{[32,43]}\) Most of the characterized group E members show higher rates when sulfides are applied as substrate and are therefore rather sulfoxidases than epox- idases, although these are supposed to be no natural substrates. A restricted epoxidation activity is reasonable from a physiological point of view, to prevent cell damage caused by excess of these reactive products. However, it remains to be clarified why sulfoxidation proceeds faster. One reason could be a different orientation of the substrate within the active site. A more general explanation might be a better accessibility of the electrons of the sulfide compared to those of the carbon-carbon double bond. This intrinsic chemical reactivity of the substrate was already discussed before.\(^{[31]}\) The latter hypothesis is supported by the preference for sulfoxidation over epoxidation in other enzymes.\(^{[26]}\)

GEMs reach turnover numbers of 0.02 to 0.54 s\(^{-1}\). Although this seems to be rather low, a compilation of rates of flavin dependent monooxygenases shows that GEMs are in a comparable range as those (Table S3). Except for some out- liners, a rather moderate activity seems to be a common feature of such monooxygenases. As the nature of substrates and reaction mechanisms are different, it is hard to generalize this finding. However, a reason for this might be the usage of non- natural substrates or xenobiotics and therefore the enzymes are not perfectly adapted to such compounds, yet. In addition, the low solubility of those substrates in water could be limiting. Due to the aromatic nature of many substrates, they are often hardly accessible and difficult to degrade, as they are not very reactive. Further, there might be physiological reasons. Some of them initiate secondary metabolite production and therefore might have a regulatory role. And especially true for GEMs, the epoxide products are known to be very toxic and can cause cell damage.\(^{[41]}\) Therefore, it is reasonable that the monoxygenase is rather slow compared to the following enzymes of a metabolic pathway. For example, the epoxide converting isomerase (SOI) is supposed to be the most active one of the upper styrene degradation pathway.\(^{[31]}\) However, from a bio- technological point of view further investigations are desirable to rationalize the rate limiting factors.

For GEMs, there is no clear correlation between activity and phylogenetic origin of the enzyme, as for instance GrStyA is highly similar to the *Rhodococcus*-like epoxidases but reaches higher activities. The same is true for IMOs and can be an indication that even a few mutations can have a huge impact on the activity and specificity of the enzyme. This observation is in accordance with several random and rational mutation studies on StyAs from *Pseudomonas*.\(^{[21,22,38]}\) For instance, Tan et al. just recently identified a residue that sits on top of the substrate cavity and affects activity by influencing FAD access to the active site.\(^{[46]}\) They were able to show by saturation mutagenesis that glycine at this position, which is conserved in over 96% of GEMs, preserves highest activity and stability.\(^{[46]}\)
In contrast, the preference to produce a certain sulfoxide enantiomer is highly related to the group E subtype, while all still produce the (S)-enantiomer of styrene oxide (Table 1). To rationalize this, we performed molecular docking studies to get an indication about the situation of the substrates in the putative active site of SMOs (Figure 2). Three binding modes were observed for the docked substrates (Scheme S1). Such effects of substituents of the substrates were described for 

\[
\text{PpStyA CA-3 before and were explained by steric hindrances.}^{[31]} \text{ Nikodinovic-Runic and co-workers also compared the sulfoxidation of the wild-type and a generated mutant SMOeng R3-11.}^{[26,30]} \text{ Herein, they showed an change in activity while the stereoselectivity was unaffected.}^{[31]} \text{ Interestingly, the mutations were located far away from the active site. In contrast, Lin et al. reported an effect for an Y73V mutant in 2012. This residue is located in the active site and they found a reversed stereoselectivity for 1-phenylcyclohexene to the (R,R)-enantiomer of the epoxide and ascribed that to the structural flexibility of this particular substrate.}^{[21]} \text{ However, our study suggests that this might be a general pattern, as SMOs also have some preference to build the (R)-enantiomer of sulfoxides. The mutation also occurs naturally, as all SMOs contain a Tyr whereas IMOs harbor a Val or ile at position 73.}

Based on sequence alignment, docking studies and the biocatalytic data of investigated GEMs, we can propose important amino acids within the active site of the enzyme. For that purpose, we included information of previous studies\cite{21,31} to draw a more precise picture of the substrate-binding and its effects on enantioselectivity. Considering this, it is now possible to determine the type of GEM and predict functionality, based on an amino acid motif: \[
\text{SMO N46-V48-H50-Y73-H76-S96 and IMO S46-Q48-M50-V173-I76-A96. The motif comprises loop 1 as well as 3 amino acids, which are important for substrate binding with respect to the hydroperoxyflavin attacking the substrate (Figure S3C). It has to be mentioned that this motif is not conserved in GEMs from phylogenetic branches containing the recently identified (R)-selective epoxidase from \textit{Streptomyces} \textit{sp. NRRL S-31} and the IMO from \textit{Burkholderia} \textit{sp. ID03} (Figure S1 and S2). This might be explained by the evolutionary progress and stabilization of the conserved residues in SMOs and IMOs. However, further explanations, e.g. another physiological role, are possible, what remains to be proven.}

The preparative synthesis of 5* showed that GEMs are promising candidates with respect to other oxygenases concerning conversion rate and stereoselectivity (Table S4). However, especially the enzyme production and stability might be addressed in future to force an application. Besides the long-term stability, one of the biggest issues of applying these proteins in industrial scale reactions is the need for an expensive cofactor. This can be abolished for GEMs by the utilization of cheaper nicotinamide mimics.\cite{18,77} These mimics are further suitable to generate simple screening systems for these enzymes. Paul and co-workers demonstrated with \textit{RoIndA1} that preparative sulfoxidation is feasible by using BNAH as reductant.\cite{18} Only one other cell-free (semi)-preparative systems was published so far, which used \textit{PtStyA} in a biphasic system to produce epoxides at gram scale.\cite{31} Due to varying experiment setups and substrates, it is hard to compare between the different experiments. Nonetheless, it can be stated that the production rates are in a similar range (This study: 0.06–0.13 U mg\(^{-1}\);\cite{26,31} 0.11–0.12 U mg\(^{-1}\);\cite{26} 0.48 U mg\(^{-1}\)).

Most (semi)-preparative scale reactions with GEMs were done by using whole-cell biocatalysts what can be beneficial to keep the costs low and bypass the need for a regeneration system.\cite{26,32,33,34,35} All of these focused on epoxidation reactions and so far no large scale sulfoxidation with whole cells has been performed. A whole-cell process can also have positive effects on the stability and the long-term performance of the enzyme. Henceforward the focus will become more on preparative synthesis of the considered reactions. Especially \textit{GpStyA} and \textit{DaIndA} are interesting candidates for further studies as they have a reasonable (long-term) activity and are highly enantioselective over a broad range of applied substrates.

### Conclusions

Our study is the biggest survey on GEMs and demonstrates that the activity and enantioselectivity of the so far characterized representatives are surpassed by several homologs. Although there is no correlation between type and activity, a suitable representative can be found for a specific substrate and a specific application. Especially \textit{GpStyA} (SMO) and \textit{DaIndA} (IMO) are interesting candidates for the synthesis of enantiopure (R)- and (S)-sulfoxides due to their remarkable stereoselectivity and activity. Moreover, the recent identification of an (R)-selective monooxygenase now also provides access to (R)- and (S)-epoxides via biocatalysts from group E. Thus, an expanded scope of GEM allows taking the next step towards an industrial application of these oxidoreductases.

### Experimental Section

#### Database Search and Phylogenetic Analysis

Putative group E monooxygenases were identified \textit{in silico} by using the BLAST algorithm\cite{69} on the non-redundant protein sequence database (NCBI). 150 representatives with a sequence identity above 20% on aa level were aligned using the ClustalW algorithm and after phylogenetic analysis the maximum likelihood tree was constructed applying the MEGAX software.\cite{70} One-component IMOs were truncated for the reductase part. From the resulting tree, 31 putative and known members were selected equally from all branches for further characterization (Table S2).

#### Molecular Docking of Ligands Into the Putative Substrate-Binding Site

Receptor and ligand files were prepared with MGLTools 1.5.6 and molecular docking was performed on the AutoDock Vina platform.\cite{69,70} The crystal structure of \textit{StyA} from \textit{P. putida} S12 (PDB ID: 3JHM)\cite{52} served as receptor. Chain B was removed and polar hydrogens were added to the monomer. As the crystal structure of the receptor does not contain any substrate or flavin co-substrate,
were synthesized as follows: purchased from Sigma (Steinheim, Germany), TCI (Eschborn, Germany) 4-methoxy phenyl methyl sulfoxide (4), phenyl methyl sulfoxide (3), 4-chlorophenyl methyl sulfoxide (7), 4-bromo phenyl methyl sulfoxide (8), ethyl phenyl sulfoxide (9), diphenyl sulfoxide (10), benzyl methyl sulfoxide (11), benzyl phenyl sulfoxide (12).

Products: styrene-7,8-oxide (1), phenyl methyl sulfoxide (2), 4-methyl phenyl methyl sulfoxide (4), 4-methoxyphenyl methyl sulfoxide (5), 4-fluoro phenyl methyl sulfoxide (6), 4-chlorophenyl methyl sulfoxide (7), 4-bromo phenyl methyl sulfoxide (8), ethyl phenyl sulfoxide (9), diphenyl sulfoxide (10), benzyl methyl sulfoxide (11), benzyl phenyl sulfoxide (12).

1, 1*, 2, 2*, 3, 3*, 4, 5, 6, 7, 8, 9, 10, 10*, 11, 12, and cofactors were purchased from Sigma (Steinheim, Germany), TCI (Eschborn, Germany) and Carl Roth (Karlsruhe, Germany). All other sulfoxides were synthesized as follows: A solution of sulfoxide (7 mmol in 100 ml methanol) was gently mixed with 20 ml water and 10 ml titanium(III) chloride (16% aqueous solution). Hydrogen peroxide (3.2 ml 30% aqueous in 15 ml methanol) was drop-wise added while the solution was constantly stirred at 20°C. After the substrate was completely converted, (about 25 min) the reaction was stopped by adding 50 ml water. Products obtained were extracted by chloroform (three times) and the combined extracts were subsequently dried over anhydrous magnesium sulfate. Thereafter, chloroform was removed under reduced pressure. Products were analyzed as described elsewhere. Restriction enzymes were received from MBI Fermentas (St. Leon-Rot, Germany) and New England Biolabs GmbH (Frankfurt am Main, Germany). Oligonucleotides and synthetic genes were synthesized by Eurofins MWG Operon (Ebersberg, Germany).

Gene Synthesis, Expression and Purification of Group E Monooxygenases

SMO and IMO gene constructs were obtained by PCR from genomic DNA or purchased from Eurofins MWG (Ebersberg) as described earlier. Gene expression, protein purification, storage and quantification of recombinant synthesized enzymes was done as described previously. Successful expression and monooxygenase activity was determined in LB media by means of indigo blue formation from indole originating from tryptophan metabolism.

Enzymatic Assays and Analytic

All measurements were done in triplicates with purified enzyme. The enzymatic assay and standard HPLC analytic was performed as described previously. The standard assay contained the following components in a final volume of 250 μl: Tris-HCl (20 mM, pH 7.5), DTT (1 mM), glycerol (5% v/v), catalase (650 U), BNAH (10 mM), FAD (50 μM), substrate (2 mM) and an appropriate amount of a GEM preparation. Initial reaction rates of selected monooxygenases were determined for styrene, PMS and BMS. Additionally, two-hour biotransformations were performed for styrene and all sulfoxides mentioned above using the same reaction conditions. Enantiomeric excess of obtained products were analyzed as described previously for epoxides and sulfoxides.

Preparative Synthesis

Preparative synthesis for the production of 5° was realized at 10 mg scale with the SMO GpStyA and the IMO DanlDA and 100 mg scale for DanlDA, respectively. All assays were done in triplicates. The reaction set up contained the same components and conditions as for the standard enzymatic assay (BNAH as e-donor, 0.5 mg ml⁻¹ GpStyA or 0.7 mg ml⁻¹ DanlDA). The substrate/BNAH mixture was supplied in portion of 1 ml 100/250 mM every 10 min to a final concentration of about 7 mM S. Samples were obtained over a period of 100 to 180 min to follow product formation via HPLC. After complete conversion, the product 5° was extracted from the remaining reaction solution with 3-times 10 ml dichloromethane in a separation funnel. The extract was concentrated in a rotary evaporator (240 rpm, 20°C, Laborata 4000-efficient, Heidolph) and re-dissolved in methanol. All samples were applied to standard HPLC and GC analysis.

Acknowledgements

The authors thank Caroline E. Paul for providing BNAH for co-substrate regeneration and Stefan R. Kaschabek for the synthesis of sulfoxides that are not commercially available. Further thanks go to Lucas Siegel and Phillip Rathsock for support during the analytical measurements. We appreciate the funding of this project by the European Social Fund and the Saxonian Government (GETGEDWEB: 100101363).

Conflict of Interest

The authors declare no conflict of interest.

Keywords: Styrene Monoxygenase · Asymmetric Epoxidation and Sulfoxidation · Biocatalysis · Substrate docking · Flavoprotein phylogeny

[1] J. Dong, E. Fernández-Fueyo, F. Hollmann, C. E. Paul, M. Pesic, S. Schmidt, Y. Wang, S. Younes, W. Zhang, Angew. Chem. Int. Ed. 2018, 57, 9238–9261.
[2] a) U. T. Bomscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, Nature 2012, 485, 185–194; b) A. Schmid, J. S. Dordick, B. Hauer, A. Kienzer, M. Wubbolts, B. Witholt, Nature 2001, 409, 258–268; c) M. Breuer, K. Dittrich, T. Habicher, B. Hauer, M. Kesseler, R. Stürmer, T. Zelinski, Angew. Chem. Int. Ed. 2004, 43, 788–824; d) G. Hughes, J. C.
A. A. Qaed, H. Lin, D.-F. Tang, Z.-L. Wu, H. Lin, D.-F. Tang, A. A. Q. Ahmed, Y. Liu, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, X. Zhang, J. Jing, L. Zhang, Z. Song, H. Zhou, M. Wu, Y. Qu, L. Liu, D. E. Torres Pazmiño, M. Winkler, A. Glieder, M. W. Fraaije, M. M. E. Huijbers, S. Montersino, D. Tischler, T. Heine, W. H. J. van Berkel, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.

A. R. Kaiser, H. Lin, D.-F. Tang, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.

A. R. Kaiser, H. Lin, D.-F. Tang, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.

A. R. Kaiser, H. Lin, D.-F. Tang, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.

A. R. Kaiser, H. Lin, D.-F. Tang, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.

A. R. Kaiser, H. Lin, D.-F. Tang, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.

A. R. Kaiser, H. Lin, D.-F. Tang, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.

A. R. Kaiser, H. Lin, D.-F. Tang, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.

A. R. Kaiser, H. Lin, D.-F. Tang, Z.-L. Wu, G.-H. Lin, H.-P. Chen, H.-Y. Shu, S.-W. Lee, J. Biotechnol. Bioeng. 2015, 123, 201–213.
[61] T. Heine, C. Großmann, S. Hofmann, D. Tischler, BIospektrum 2018, 24, 446–448.
[62] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, Mol. Biol. Evol. 2018, 35, 1547–1549.
[63] T. Heine, C. Großmann, S. Hofmann, D. Tischler, BIol. Chem. 2019, 400, 939–950.
[64] J. A. D. Gröning, S. R. Kaschabek, M. Schlömann, D. Tischler, Arch. Microbiol. 2014, 196, 829–845.
[65] a) R. Z. Harris, S. L. Newmyer, P. R. Ortiz De Montellano, J. BIol. Chem. 1993, 268, 1637–1645; b) E. O’Reilly, M. Corbett, S. Hussain, P. P. Kelly, D. Richardson, S. L. Flitsch, N. J. Turner, Catal. Sci. Technol. 2013, 3, 1490; c) J.-D. Zhang, A.-T. Li, Y. Yang, J.-H. Xu, Appl. Microbiol. Biotechnol. 2010, 85, 615–624; d) G. Carrea, B. Redigolo, S. Riva, S. Colonna, N. Gaggéro, E. Battistel, D. Bianchi, Tetrahedron: Asymmetry 1992, 3, 1063–1068.
[66] G. de Gonzalo, M. Fürst, M. Fraaije, Catalysts 2017, 7, 288.
[67] D. L. Gatti, B. Entsch, D. P. Ballou, M. L. Ludwig, 1996, p-Hydroxybenzoate hydroxylase complexed with 4-4-hydroxybenzoate at pH 5.0 (PDB ID: 1UV), The Protein Data Bank, doi: 10.2210/pdb1UV/pdb.
[68] T. Hiromoto, S. Fujiwara, K. Hosokawa, H. Yamaguchi, 2006, Crystal structure of 3-hydroxybenzoate hydroxylase from Comamonas testosteroni, in complex with the substrate (PDB ID: 2DKH), The Protein Data Bank, doi: 10.2210/pdb2DKH/pdb.
[69] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, J. Comput. Chem. 2009, 30, 2785–2791.
[70] O. Trott, A. J. Olson, J. Comput. Chem. 2010, 31, 455–461.
[71] B. Entsch, L. J. Cole, D. P. Ballou, Arch. Biochem. Biophys. 2005, 433, 297–311.
[72] H. S. Biswal, in Challenges and Advances in Computational Chemistry and Physics, Vol. 19 (Ed.: S. Scheiner), Springer, Cham, 2015.
[73] a) R. N. Patel, Green Biocatalysis, John Wiley & Sons, Hoboken, NJ, 2016; b) T. Matsui, Y. Dekishima, M. Ueda, Appl. Microbiol. Biotechnol. 2014, 98, 7699–7706.