Accessing the Biocatalytic Potential for C−H-Activation by Targeted Genome Mining and Screening

Lisa Marie Schmitz,[a] Jonas Schäper,[a] Katrin Rosenthal,[a] and Stephan Lütz*[a]

Cytochrome P450 monooxygenases (P450s) are ubiquitous hemeproteins that insert oxygen specifically into substrates leading to diverse chemical transformations. Utilizing their capabilities, microbial whole-cell biocatalysts are applied in pharmaceutical and fine chemical industry to produce biomolecules and drug metabolites. In order to synthesize novel bioactive compounds there is a great demand to identify P450s with new reaction and substrate scope. In this study, genome mining and an activity screening were successfully combined to discover so far underutilized biocatalysts. The screening revealed the expected broad range of reactions, such as hydroxylations, dealkylations, reductions and desaturations. For *Actinosynnema mirum* and ritonavir the biotransformation was transferred to a preparative scale resulting in a ritonavir conversion of 90 % after 48 h and 13 different metabolites analyzed by LC-MS² and NMR. These results clearly demonstrate the potential of the underlying approach to identify promising whole cell biocatalysts with good conversion and product scopes.

Introduction

Since their discovery five decades ago P450s have emerged as one of the most studied families of enzymes due to their ability to catalyze the selective oxyfunctionalization of various substrates. Starting from an initial formal oxene transfer, P450s are able to catalyze many different chemical transformations, such as dealkylations, epoxidations or deaminations.[1] Linked to their natural role in detoxification reactions and oxyfunctionalization of active pharmaceutical ingredients (API) in the phase I metabolism, they serve as useful biocatalysts for an effective enzymatic one-step conversion to high value hydroxylation products. The synthesis of drug metabolites during an early stage of drug development is useful to determine the metabolic fate of a drug. The products can be used for structure elucidation and pharmacological activity or toxicity studies which is a central aspect in the approval of APIs.[2] Furthermore, the modification and diversification of already existing drug candidates by P450s can lead to even more active compounds.[3] While some chemical oxidation reactions for the synthesis of metabolites are known, the biologically relevant metabolites are mostly obtained by a biotransformation[4] either by the human P450s or their microbial equivalents, which typically exhibit higher specific activities and stabilities often resulting in larger product quantities.[5]

In order to optimize the catalytic activity or broaden the substrate scope of P450 reactions, either new P450s can be identified or already known enzymes can be optimized by protein engineering. The most prominent enzyme used for engineering approaches is the CYP102 A1 (also referred to as P450 BM3) from *Bacillus megaterium*, since the wildtype enzyme exhibits a high expression level in *E. coli* and a high catalytic activity.[6] Mutants of CYP102 A1 were shown to metabolize a variety of drugs, such as naproxen, ibuprofen,[7] amodiaquine[8] or astemizole[9] producing the main human metabolites. Besides protein engineering to receive improved, tailor made biocatalysts, a likewise promising approach is exploiting the natural diversity of P450s to find catalysts for selective oxidations. Nowadays, several online databases exist that host the information of thousands of new, but underexplored P450 enzymes. Prominent databases are the P450 database maintained by David Nelson, which provides information on over 20,000 P450 genes[10] and CYPED with nearly 16,000 sequences.[11] Due to an increasing availability of microbial whole genome sequences and with that an increasing number of annotated P450s, it is not the matter of gene sequence availability any more but to correctly assign the gene-function relationship.[12] Some species, such as the actinomycetes are known to have a high number of P450 enzymes which corresponds to their natural rich secondary metabolism.[13] Different actinomycetes have already been described to successfully produce drug metabolites. In a whole cell biotransformation the conversion of cyclosporine was screened with 1,237 actinomycetes. The screening resulted in nine strains showing good bioconversion rates of cyclosporine into the two human hydroxylated (AM1 and AM9) and N-demethylated (AM4) derivatives.[14] In a recent study a novel P450 from the actinomycete *Streptomyces platensis* was identified showing

[a] L. M. Schmitz, J. Schäper, Dr. K. Rosenthal, Prof. S. Lütz
     Chair for Bioprocess Engineering
     Department of Biochemical and Chemical Engineering
     TU Dortmund University
     Emil-Figge-Straße 66
     Dortmund 44227 (Germany)
     E-mail: stephan.luetz@tu-dortmund.de

This manuscript is part of the Special Issue on the Sustainable and Affordable Chemistry to Meet Future Challenges in the Pharmaceutical Industry.

high activity for the conversion of several pharmaceutical compounds while producing equivalent metabolites to the human metabolizing P450s 3A4, 2C8, 2C19, and 2D6.[15]

Compared to bacteria, fungi, especially filamentous fungi, are known for even larger amounts of P450 which is related to their ability to inhabit ecological niches and their production of specialized metabolites.[16] Similar to actinomycetes, fungi have also been used for the production of drug metabolites. A potential new antimalarial compound cladosporin was incubated with 96 microbial strains, 48 fungi and 48 actinomycetes, to receive closely related but more active variants of cladosporin giving an excellent example of a late-stage functionalization with P450 catalysis.[17] The biotransformation generated a wide set of cladosporin analogues that could be used for structure-activity studies. All of these examples show that there is a huge potential within microorganisms to find efficient biocatalysts with new P450 reactivities.

In this study, we present a combined approach of a bioinformatic sequence analysis and medium-throughput activity screening of promising candidates in order to identify new P450s catalyzing novel reactions. Genomes were analyzed and P450 coding sequences were identified. Since the gene sequences alone do not reveal information about the functionality and activity of the P450s, 84 strains were tested in a medium-throughput screening towards the conversion of seven pharmaceutical compounds, namely amodiaquine, cyclosporine, tamoxifen, haloperidol, ritonavir, rapamycin and testosterone. Using this combined approach, we were able to generate a platform of novel whole-cell biocatalysts for hydroxylation reactions with some strains showing a broad reaction spectrum and wide substrate acceptance. For the actinomycete *Actinosynnema mirum*, and the substrate ritonavir, whole cell biotransformation conditions were optimized and transferred to a fermentative scale. A ritonavir conversion of approximately 90% was achieved and in total 13 ritonavir metabolites could be found after 48 h of biotransformation. Further characterization by tandem LC-MS (LC-MS^2) and NMR analyses revealed among others the three main human metabolites.

Results and Discussion

A Genome Mining Approach Identified Strains with a high Hydroxylation Potential

As all P450s share a common structural fold of the heme center and the surrounding binding pocket, a genome analysis for the presence of conserved patterns and key-motifs is a promising approach to identify novel P450 enzymes.[18]

To determine the number of P450 genes encoded per strain, matches per NCBI taxonomy-database entry were counted. Two different approaches were performed based on two different databases, namely protein data deposited on UniProtKB (PROSITE) and the CYPED database. The data-base analysis resulted in a list of organisms with their putative number sequences coding for P450s. Both approaches resulted in different total amounts of sequences, however mostly showing

the same tendency of sequence distributions between the species. Sequence analysis for bacterial organisms resulted in a total number of up to 320 P450 sequences per organism for the CYPED and 215 sequences per organism for the PROSITE database. High values above 200 resulted from the phenomenon that entries were counted per NCBI Taxonomy ID, which can refer to a sequenced strain, a whole species or a genus and therefore can cause an overestimation of sequences. In particular, the genome mining showed that actinomycetes from the families Streptomycetaceae and Mycobacteriaceae possess the highest numbers of P450, corresponding to the expectations. Actinobacteria especially from these families are known for their large number of P450s with 18 found for example in *Streptomyces coelicolor*, 33 in *Streptomyces avermitilis*[19] and 70 in *Mycobacterium rhodesiae* NBB3.[20] In this study, we found up to 80 total P450 sequences in one strain of the Streptomycetaceae, while up to 145 sequences were found in a strain of the Mycobacteriaceae with the PROSITE analysis. The CYPED approach resulted in up to 82 total sequences for strains belonging to the Streptomycetaceae (except for 141 in *Streptomyces griseus*) and up to 120 sequences for strains belonging to the Mycobacteriaceae.

In addition to the total count of sequences, the numbers of sequences belonging to homologous families and superfamilies were determined using the CYPED database. Analysis of the homologous families and superfamilies revealed the highest number of superfamilies for Mycobacteriaceae with 30 superfamilies followed by Streptomycetaceae with 18 superfamilies. In literature, analyses of mycobacterial CYPomes, mostly of the pathogenic bacterium *Mycobacterium tuberculosis*, revealed a high diversity within the found sequences and a localization widely distributed across the genome, making determination and assignment of the catalytic function difficult.[13] Streptomycetaceae are one of the greatest producers of diverse natural products requiring many genes encoding for P450s with tremendous diversity in secondary metabolite gene clusters.[21] Furthermore, numerous transformations of xenobiotic compounds for industrial and environmental applications are known for *Streptomyces* species.[22] Mycobacteriaceae and Streptomycetaceae are therefore promising candidates to identify P450s with novel substrate spectrum and reactivity.

Genome mining of fungi revealed even larger numbers of P450s per organism. The mining of fungal genomes for P450s revealed up to 120 sequences per strain with the CYPED approach and up to 232 sequences per strain for the PROSITE genome mining approach. The genus *Fusarium* and *Colletotrichum/Glomerella*, both groups containing plant pathogen species, showed the highest number of total sequences. High numbers of superfamilies (up to 36) were also present in these families. Many fungi are capable to degrade environmental pollutants, break down organic materials and produce secondary metabolites to adapt to various environmental conditions. This results in P450s with an enormous diversity to be found in fungal genomes.[23] Especially multiple phytopathogenic and plant degrading fungi have multiple P450s.[24] The highest number with 250 P450 encoding genes is described for the brown-rot fungus *Postia placenta*.[25]

The genome mining thus identified promising candidates for a subsequent activity screening on pharmaceutical substances. The number of candidates for this activity screening was further reduced based on different criteria. As first criterion, they had to have a high number of total P450 sequences. Besides the number of total sequences, a high number of superfamilies was supposed to guarantee high diversity in reactivity as second criterion. As criterion three, good cultivability and four, novelty were chosen; hence strains were picked which had not been described as oxidation catalysts before. 89 bacterial and fungal strains fulfilling these criteria were selected. If possible the exact strain was included; on some strains with limited availability a closely related strain was included. Chosen bacterial strains for the activity screening mainly belonged to the orders Streptosporangiales, Pseudonocardiales, Streptomycetales, Corynebacteriales with highest total sequences for the Streptomycetales and Corynebacteriales (Figure 1A). For fungi, strains with high numbers of P450 sequences did not belong to some main orders but were more widely distributed across different classes. The chosen fungi mainly belonged to the Ascomycota and the classes Dothideomycetes, Eurotiomycetes and Sordariomycetes showing highest total sequence numbers within the Sordariomycetes, especially for the order Glomerellales (Figure 1B).

## Screening Revealed Broad Reactivity on Model Substrates

To determine functionality and reactivity, selected strains from the genome mining approach were tested for the conversion of seven representative substrates. From the 89 selected strains 84 strains were cultivable with the applied conditions. The media NL148 s (fungi) or NL148sb (bacteria) were used for the screening, if not indicated otherwise. After 48 h and 72 h of biotransformation the product profile was analyzed via LC-MS and new products with their respective mass-to-charge ratio (m/z) were counted. Natural products formed by the strain were distinguished from the product profile by a negative control (culture without substrate). The number of formed products from all tested substrates was plotted against number of sequences determined by the genome mining approach (Figure 2).

Among the tested bacteria (Figure 2 A), strains belonging to the Pseudonocardiales, Streptosporangiales, and Streptomycetales showed the highest product spectrum on the tested substrates with up to 52 formed products. Between 8 and 29 sequences were found on average for these groups using the CYPED approach. In contrast, strains from the Corynebacteriales (45363, 45147, 44199, 43066, 44338, 45407, 43239, 44107, 44383, 44892, 45091, 44980, 43826 and 45343) generally showed the lowest metabolite numbers with 15 formed products for the soil bacterium *Gordonia rhizophera* (44383) as highest value within this group. In general, no correlation between sequence number and number of formed products could be observed. On the one hand, especially the class of Myctobacteria revealed high numbers of total sequences and superfamilies of P450s, however, resulted in the lowest product

formation with a maximum of nine formed total products. The strain *Mycobacterium hassiacum* (44199) for example was found to have 45 total sequences in the PROSITE and 53 total sequences and 24 superfamilies in the CYPED approach, however, in the screening only five total products were identified. On the other hand, for the strain *Streptomyces platensis* (40041) for which genome mining resulted in 8 total sequences for CYPED and 13 total sequences belonging to three found superfamilies for PROSITE, 31 products were found. Also regarding the substrate acceptance, Mycobacteria did not show a broad specificity, with conversion of only up to three of the seven substrates, regardless of the high number of total sequences and superfamilies. Additionally, mostly substrates with lower molecular mass, namely haloperidol and testosterone, were converted. Genomic analyses from literature showed that many mycobacterial P450s appear to be unique within mycobacteria catalyzing still unknown reactions.[26] The lacking conversion of the tested substrates could be explained by a narrow substrate acceptance, as mycobacterial P450s act rather specific.

In contrast to the Mycobacteria, *Actinomadura rifamycini* (43936), *Kutzneria albida* (43870) and *Actinosynnema mirum* (43827) belonging to the Streptosporangiales (43936) or Pseudonocardiales (43827, 43870) accepted all seven substrates. Three strains from these groups and three strains belonging to the Streptomycetales accepted six substrates. Even though these strains accepted all or nearly all substrates, especially ritonavir was broadly converted with 18 products for example for *K. albida* and 11 products for *A. mirum* (see supporting information). Among others, these were single hydroxylations (m/z = 737.3), double hydroxylations (m/z = 753.3) and dealkylations (m/z = 582.3) or combinations of single hydroxylation and dealkylation or demethylation (m/z = 596.3 and m/z = 723.3). Besides ritonavir, cyclosporine, one of the largest molecules, was broadly metabolized. For *A. rifamycini* six products were identified, ranging from single oxidation (m/z = 610.0), double oxidation (m/z = 618.0) and triple oxidation (m/z = 626.0) to several combinations of demethylation and oxidations (m/z = 603.0 or m/z = 611.0). Despite belonging to the same phylum Actinomycetales, Mycobacterium and Streptomyces species were shown to have few similarities in the P450 profile.[27] Many widespread P450 families such as the families CYP105 and CYP107 appear to be unique in Streptomycetaceae and other related actinomycetes. Characterization of these families revealed roles in oxidative tailoring of secondary metabolites and the metabolism of xenobiotics and thereby showing broad substrate specificities and catalytic reactivity.[13] Therefore, broad substrate acceptance of these enzymes catalyzing diverse reactions besides native catalytic activity is most likely. Well characterized P450s from these families are involved in the synthesis of the antibiotics filipin (CYP105D1)[28] and pikromycin (CYP107B)[29] but were also shown to catalyze hydroxylation of vitamin D (CYP105 A1).[30]

For fungi the number of total products compared to the bacterial strains was even lower, though genome mining resulted in higher numbers for P450 sequences (Figure 2 B). Comparable to the results for Mycobacteria, substrates with low
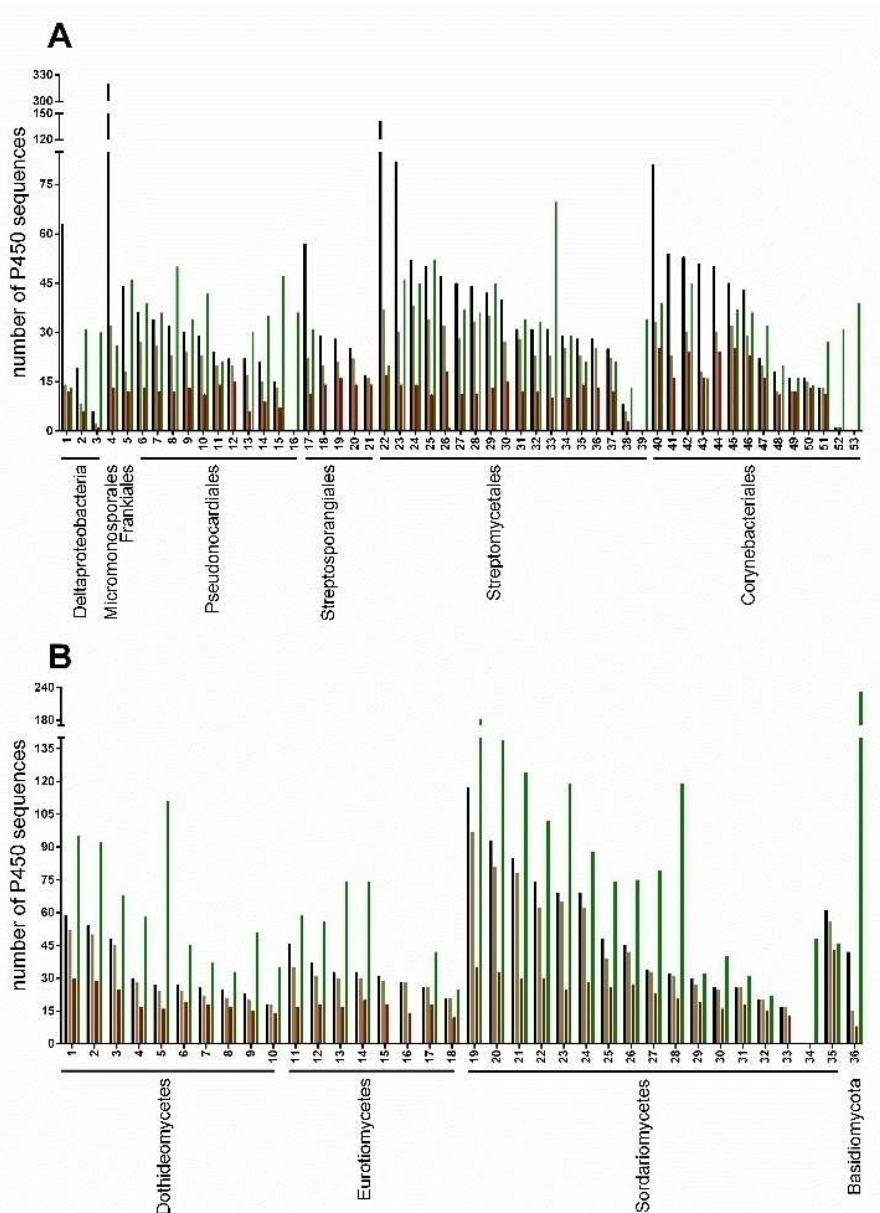
Figure 1. Bacterial and fungal strains selected for activity screening. Shown are total numbers of found P450 sequences for CYPED (black) and PROSITE (green) and the number of homologous families (brown) and superfamilies (red). A: bacterial strains **1** *Bradyrhizobium elkanii*, **2** *Cystobacter fuscus* DSM 2262, **3** *Sphingomonas macrogoltabidus*, **4** *Salinispora arenicola* CNS-205, **5** *Frankia* sp. Eul1c, **6** *Amycolatopsis vancoresmycina* DSM 44592, **7** *Amycolatopsis rifamycinica* NT 19, **8** *Kutzneria* sp. 744, **9** *Amycolatopsis japonica* DSM 44213, **10** *Amycolatopsis decaplanina* DSM 44594, **11** *Saccharomonospora marina* XMU15, **12** *Pseudonocardia acacia*, **13** *Actinosynnema mirum* DSM 43827, **14** *Lentzea albida* DSM-44437, **15** *Kutzneria albida* DSM 43870, **16** *Amycolatopsis regifaucium*, **17** *Streptosporangium roseum*, **18** *Actinomadura oligospora* A80190.1, **19** *Actinomadura rifamycini* DSM 43936, **20** *Microtetraspora glauca*, **21** *Actinomadura atramentaria*, **22** *Streptomyces griseus*, **23** *Streptomyces albulus*, **24** *Streptomyces rapamycinicus* NRRL 5491, **25** *Streptomyces albus*, **26** *Streptomyces sclerotialus*, **27** *Streptomyces viridochromogenes* Tue 494, **28** *Streptomyces violaceusniger*, **29** *Streptomyces iranensis* HM 35, **30** *Streptomyces xanthophaeus*, **31** *Streptomyces ipomoeae* 91–03, **32** *Streptomyces cattleya* DSM 46488, **33** *Streptomyces acidiscabies* 104, **34** *Streptomyces decoyicus*, **35** *Streptomyces katrae* NRRL ISP-5550, **36** *Streptomyces ochraceiscleroticus* NRRL ISP-5594, **37** *Streptomyces aurantiacus* ISP 5412, **38** *Streptomyces platensis*, **39** *Streptomyces cellostaticus*, **40** *Mycobacterium gilvum* Spyr1, **41** *Nocardia rhamnosiphila* NRRL B-24637, **42** *Mycobacterium hassiacum* DSM 44199, **43** *Rhodococcus erythropolis*, **44** *Mycobacterium tusciae* JS617, **45** *Mycobacterium aromaticivorans* JS19b1, **46** *Mycobacterium phlei* DSM 43239, **47** *Rhodococcus wratislaviensis* NBRC 100605, **48** *Gordonia rhizosphera* NBRC 16068, **49** *Rhodococcus triatomae* DSM-44892, **50** *Rhodococcus imtechensis* RKJ 300, **51** *Rhodococcus* sp. EsD8 (BCP1), **52** *Mycobacterium chlorophenolicum* DSM 43826, **53** *Mycobacterium llatzerense*. B: fungal strains **1** *Bipolaris zeicola* 26-R-13, **2** *Bipolaris victoriae* FI3, **3** *Pyrenophora tritici-repentis* Pt-1C-BFP, **4** *Pseudocercospora fijiensis* CIRAD86, **5** *Neofusicoccum parvum* UCRNP2, **6** *Aureobasidium subglaciale* EXF-2481, **7** *Aureobasidium pullulans* EXF-150, **8** *Aureobasidium melanogenum* CBS 110374, **9** *Coniosporium apollinis* CBS 100218, **10** *Baudoinia compniacensis* UAMH 10762, **11** *Capronia epimyces* CBS 606.96, **12** *Penicillium roqueforti* FM164, **13** *Phialophora europaea* CBS 101466, **14** *Penicillium stipitatum* ATCC 10500, **15** *Cladophialophora yegresii* CBS 114405, **16** *Capronia coronata* CBS 617.96, **17** *Penicillium oxalicum* 114–2, **18** *Uncinocarpus reesii* 1704 UAMH 1704, **19** *Colletotrichum gloeosporioides* Nara gc5, **20** *Colletotrichum fioriniae* PJ7, **21** *Colletotrichum orbiculare* 104-T, **22** *Colletotrichum sublineola*, **23** *Colletotrichum higginsianum* IMI 349063, **24** *Colletotrichum graminicola* M1.001, **25** *Trichoderma virens* Gv29-8, **26** *Metarhizium robertsii* ARSEF 23, **27** *Podospora anserina* S28, **28** *Eutypa lata* UCR-EL1, **29** *Sordaria macrospora* k-hell, **30** *Verticillium alfalfae* VaMs.102, **31** *Myceliophthora thermophila* ATCC 42464, **32** *Neurospora tetrasperma* FGSC 2508, **33** *Ophiostoma piceae* UAMH 11346, **34** *Acremonium chrysogenum* DSM 880, **35** *Beauveria bassiana* ARSEF 2860, **36** *Moniliophthora roreri* MCA 2997
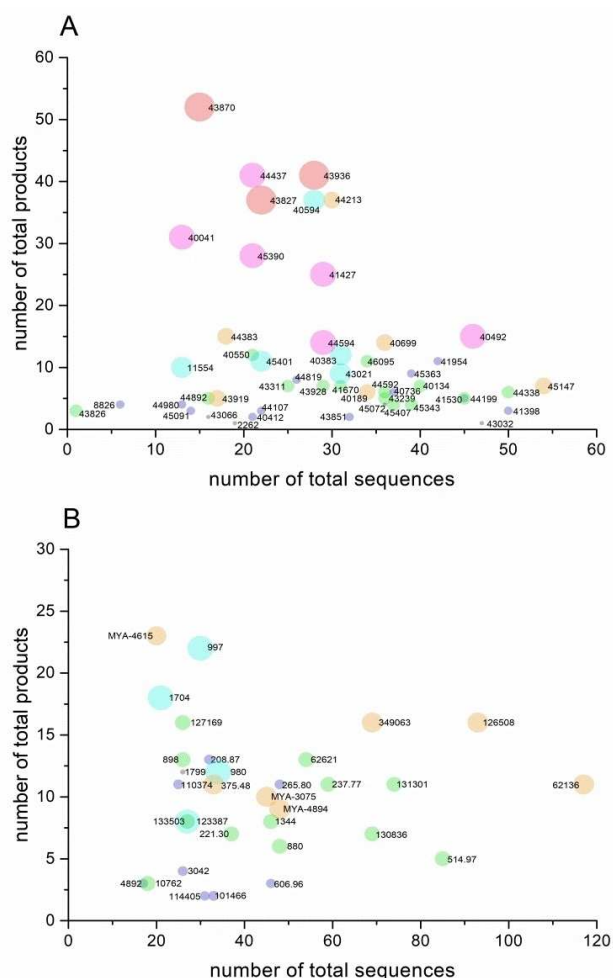
Figure 2. Conversion of the seven tested substrates. The total product formation of all substrates is plotted against the number of found sequences from both genome mining approaches. As both approaches did not cover the same strains, genomic data analysis results from both approaches (CYPED and PROSITE) were included. If both approaches resulted in a different number of sequences for the same strain, always the lowest number of found sequences was plotted on the x-axis. A: bacterial strains. B: fungal strains. Size of dots (big to small) and color indicates the number of accepted substrates. Red: seven, magenta: six, cyan: five, orange: four, green: three, blue: two, gray: one.

molecular mass, such as testosterone, were converted by nearly all fungal strains, whereas cyclosporine, the largest molecule, was not converted at all. The strain *Neurospora tetrasperma* (MYA-4615) for example formed 16 products for the substrate testosterone, with several hydroxylation (m/z = 305.1) and hydroxylation/desaturation products (m/z = 303.1). Possibly, biotransformation of testosterone is likely, as testosterone has a similar structure compared to ergosterol, a major component of eukaryotic cell membranes from yeasts and filamentous fungi. The P450 enzymes participating in the ergosterol synthesis, CYP51 and CYP61, might also act on testosterone as substrate. CYP51 and CYP61 are highly conserved in the entire kingdom of fungi.[31]

Despite a low number of formed products, fungi also showed a narrower substrate acceptance with five accepted

substrates as maximum for *Sordaria macrospora* (997), *Uncinocarpus reesii* (1704), and *Podospora pauciseta* (980). From these three strains, *S. macrospora* (997) and *U. reesii* (1704) formed the highest number with up to 22 total products. Only for *Neurospora tetrasperma* (MYA-4615) a higher number of 23 products was detected. Additionally, *S. macrospora* showed high intensities of the formed products, especially for the potential desethylamodiaquine (m/z = 328.0), tamoxifen demethylation (m/z = 358.1) and tamoxifen single oxidation (m/z = 388.1).

The screening revealed novel biocatalysts showing a broad reactivity and different specificities for the tested substrates. Bacteria showed higher product formation of the tested substrates compared to the fungi, however on average more P450 sequences were found in fungi by the genomic analysis. However, it became also apparent that number and P450 diversity are not sufficient to predict the potential for oxyfunctionalization. Especially, some bacterial strains from the Streptosporangiales, Pseudonocardiales and Streptomycetales and some filamentous fungi from the Eurotiomycetes and Sordariomycetes were potent biocatalysts to catalyze diverse chemical transformations with a potential for further optimization of the biotransformation parameters.

Improvement of the Ritonavir Biotransformation Conditions for the Actinomycete *Actinosynnema Mirum*

One strain from the initial screening was chosen to optimize the biotransformation parameters in order to develop a scale up to a 3 L fermenter. *A. mirum* was chosen since it showed high reactivity in the activity screening and had a broad substrate spectrum accepting all seven substrates regardless of size and properties. *A. mirum* is a gram-positive actinobacterium belonging to the order of Pseudonocardiales which forms aerial and substrate mycelia.[32] The genome is fully assembled. *A. mirum* was shown to be capable of synthesizing natural products, such as nocardicin antibiotics[33] and ansamitocin P-3.[34] For biotransformation, the substrate ritonavir was chosen, as a high metabolite conversion of this substrate was detected in the initial screening. In contrast to the described inhibitory effect on the human major drug-metabolizing CYP3A4 and CYP3A5 isoforms,[35] this effect of ritonavir on the responsible P450 seems not to be present.

Optimization of substrate conversion and product scope was performed in the BioLector, as it offers the possibility of parallel cultivation combined with online monitoring of the biomass formation. Studies have shown that the substrate conversion rate is significantly influenced by the composition of media.[14] Therefore, in a first experiment different media for the pre- and main-cultivation were tested. Ritonavir was added in the main cultivation after 72 h. Highest total biomass formation was achieved with terrific broth (TB) medium (Figure 3). Second highest biomass concentration was reached in the GYM (glucose, yeast, malt) medium followed by the NL148sb medium, M9 und CY medium. No growth could be detected in oatmeal medium, although, excellent growth on oatmeal agar is described in literature.[32] The growth rates (μ_{max}) were
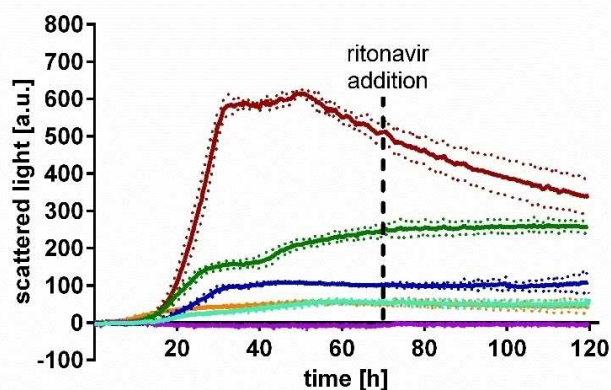
Figure 3. Growth curves of medium variation for *Actinosynnema mirum*. Cultivation was performed for 120 h in the BioLector with 800 µL filling volume of TB medium (red), GYM medium (green), NL148sb medium (blue), CY medium (orange), M9 medium (cyan), and oatmeal medium (magenta). Ritonavir with a concentration of 0.1 mg mL^{-1} was added after 72 h of cultivation in the stationary phase (dashed line). Values and standard deviation are taken from six replicates.

0.28 h^{-1} (TB medium), 0.27 h^{-1} (GYM medium), 0.36 h^{-1} (NL148sb medium), 0.27 h^{-1} (CY medium) und 0.30 h^{-1} (M9 medium). Calculation of growth rate for the oatmeal medium was not possible. In literature, *A. mirum* was shown to hydrolyze starch, casein, tyrosine and gelatin.[32] The media TB, GYM, NL148sb and CY contain casein-peptone, while starch was additionally present in NL148sb medium.

Contrary to the good growth of *A. mirum* in TB medium no ritonavir conversion after 48 h could be observed (Table 1). Besides cultivation in TB medium, cultivation in M9 medium resulted in no product formation. This however, might be linked to low biomass formation. Biotransformation in NL148sb resulted in the three metabolites M1, M2, and M5 (for chemical structures and nomenclature see Table 1 and Figure 7) and in CY medium cultivation resulted in one metabolite (M5). Interestingly, for cultivation in oatmeal medium the three metabolites M1, M2, and M5 were detected. These transformations might either result from the cells used for inoculation, or from slightly growing cells, which could not be detected due to the high turbidity of the medium. In microscopic analysis after 48 h cells could be observed in the oatmeal medium. Most metabolites (M1 to M7) were detected after the

cultivation in GYM medium. GYM medium, also named International Streptomyces Project (ISP) medium 2, is a standard medium for morphological studies and characterization of *Streptomyces* species. It is the only tested medium containing malt extract. Furthermore, it contains yeast extract. Yeast extract was also present in NL148sb medium in which cultivation did also result in a good product formation. Possibly, components from either yeast extract or malt extract are responsible for P450 gene induction.

After cultivation in GYM medium, eight different products were identified. The m/z ratio of 582.3 for M1 represents the des-isopropylthiazolyl product of ritonavir and the m/z ratio of 580.3 for M2 represents the decarbamoylated product, which were described earlier in conversion of ritonavir with human liver microsomes.[36] M3 with the m/z ratio 707.3 indicates an *N*-demethylation of ritonavir and M4 (m/z ratio of 723.3) an additional oxidation ($+16$) of M3. M5 with the m/z ratio of 737.3 ($+16$) indicates a single oxidation of ritonavir. Several single oxidations for ritonavir are described[37] with the isopropylthiazolyl oxidation product as the main human metabolite with around 30% of the total initial dose.[36b] Metabolite M6 with an m/z ratio of 753.3 ($+32$) is a double hydroxylation product of ritonavir. M7 with an m/z ratio of 596.3 indicates a single oxidation with the removal of the carbamoyl group produced either from metabolite M2 or M5.[36b] M8 with an m/z ratio of 735.3 is possibly corresponding to an oxidation of metabolite M5 to a ketone. This metabolite was described before in a biotransformation of ritonavir with a P450 of the strain *Streptomyces platensis*.[15] Furthermore, products with the m/z ratio of 606.3 and 622.3 could be identified which were described as degradation products of ritonavir before and were also present in the sample directly taken after substrate addition.[38] These products were excluded from the product profile.

As cultivation in GYM medium resulted in the best product formation, GYM medium was used to characterize the biotransformation in more detail. The conversion of ritonavir over 176 h was determined using the BioLector system (Figure 4). For eukaryotic P450 expression it is described that gene expression can be induced by the presence of the substrate.[39] To test an effect on conversion, ritonavir was added in the exponential phase after 20 h. Early addition of the substrate did not show a negative effect on growth. A biomass concentration of 3.3 g$_{CDW}$L^{-1} was reached after 40 h. The substrate ritonavir was fully converted after 48 h of cultivation. In further experi-

Table 1. Product formation for the cultivation on different media. Metabolites M1 to M8 are shown with their retention times for the different media.

| Medium	M1 582.3	M2 580.3	M3 707.3	M4 723.3	M5 737.3	M6 753.3	M7 596.3	M8 735.3
NL148sb	6.27	–	8.16	–	7.18 6.96	–	–	8.18
TB	–	–	–	–	–	–	–	
GYM	6.27	6.67	8.19	6.54	7.18 6.96	6.27	5.22	6.72 8.18
CY	–	–	–	–	7.24	–	–	
M9	–	–	–	–	–	–	–	
Oatmeal	6.07	6.68	–	–	7.27	–	–	
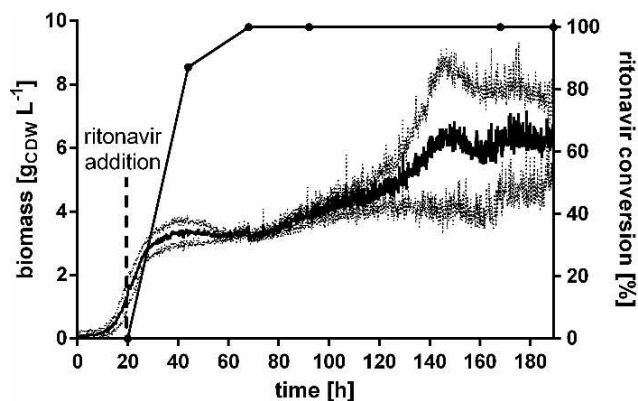
Figure 4. Growth and ritonavir conversion of cultivation on GYM medium over 196 h. 0.1 mg mL^{-1} ritonavir were added in the exponential phase after 20 h of cultivation. Cultivation was performed using the BioLector system in an 800 μL scale. Ritonavir conversion (connected dots), biomass (solid line). Values and standard deviation are taken from triplicates.

Figure 5. Ritonavir conversion in preparative scale in 1.7 L culture volume. *A. mirum* was cultivated in GYM medium and the substrate ritonavir in a final concentration of 0.1 mg mL^{-1} was added after 20 h of cultivation. Shown are growth (green), glucose consumption (red) and conversion of ritonavir (black). Data points were calculated from duplicates.

ments a repeated addition of ritonavir after 44 h resulted in a slower conversion of ritonavir (data not shown). Biotransformation with P450s is further dependent on sufficient cofactor regeneration. To maintain cell metabolism and to guarantee sufficient cofactor supply, glucose as energy source was added in the stationary phase. Blank *et al.* determined a clear correlation of glucose uptake rate and NADH regeneration rate and could also observe a linear dependency of glucose uptake rate and biocatalytic activity when using resting cells.[40] This shows that the biocatalytic performance of the whole-cell biocatalysts relies on the metabolic activity. However, in this study a further addition of glucose did not show further improvement of the conversion (data not shown) indicating that cofactor regeneration was not limiting. Minura *et al.* furthermore showed that the addition of 5 aminolevulinic acid (5-ALA), a precursor of the porphyrin biosynthesis, led to the up-regulation of the P450 CYP2A6 activity and prodrug activation.[41] Addition of 5-ALA in the exponential phase, however, did not show further improvement of the conversion (data not shown). As the ritonavir addition in the exponential phase did not show a negative effect on growth and a good substrate conversion was reached, a batch cultivation with GYM medium and substrate addition after 20 h was utilized for the scale-up.

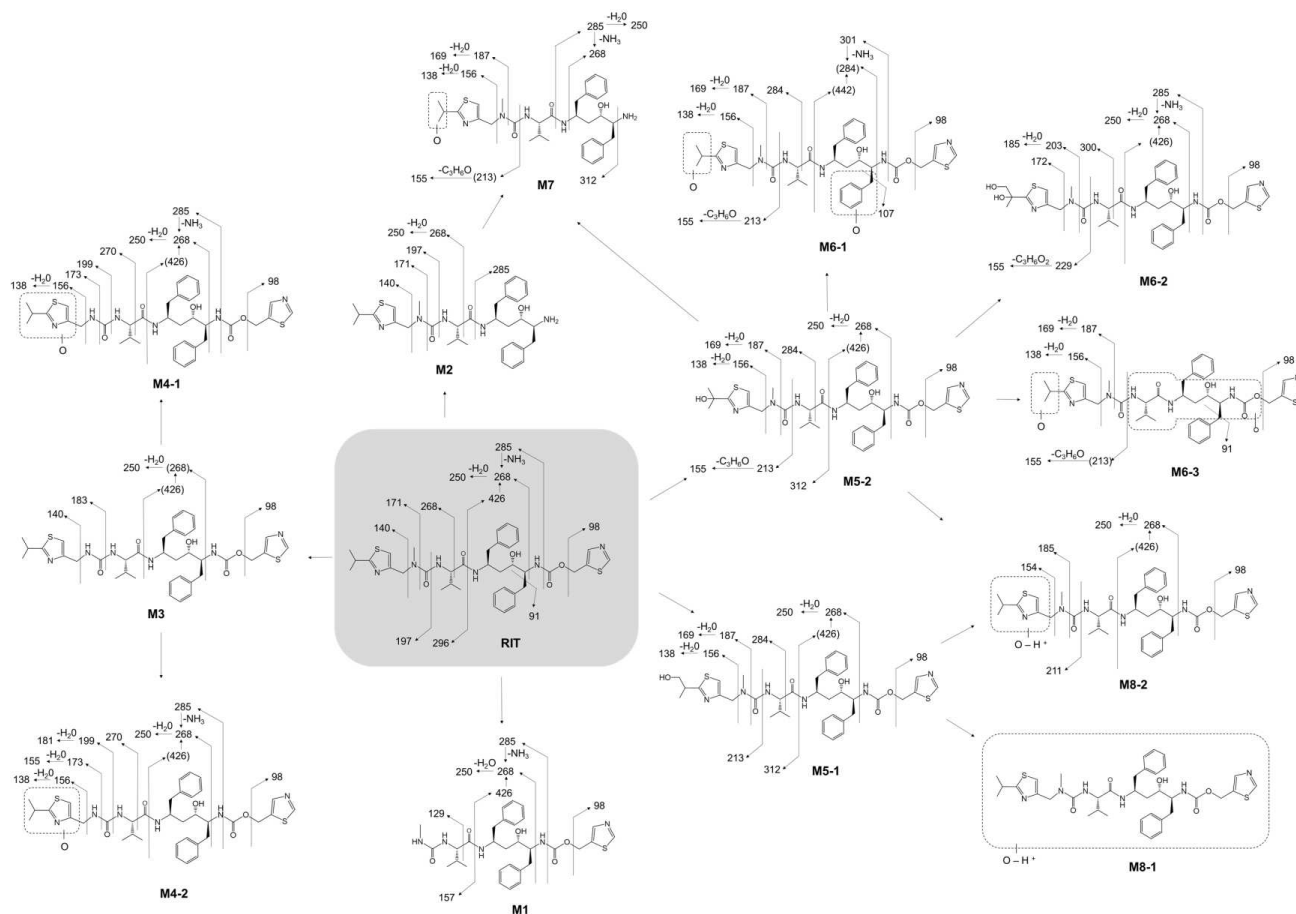## Ritonavir Biotransformation on Preparative Scale

In a next step the conditions from the BioLector cultivation were transferred to a 1.7 L preparative scale (Figure 5). Fermentation of *A. mirum* showed good scalability as a similar growth behavior could be observed. The final biomass concentration of 4.3 g$_{CDW}$ L^{-1} was 1 g$_{CDW}$ L^{-1} higher compared to the results in the BioLector which might be a result from better oxygen availability of the stirred system. Glucose was depleted after 44 h of cultivation. After 24 h already 46 % of the initial 170 mg of ritonavir were converted, 90 % after 48 h. After 68 h
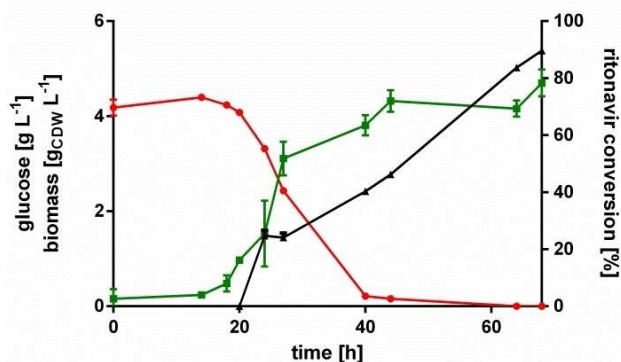
the fermentation broth was centrifuged and the biomass, as well as the supernatant were separately extracted with ethyl acetate with an extraction efficiency of 70 % calculated from the transition of the internal standard darunavir. After the extraction of both, soluble phase and biomass, it could be observed that the highest amount of unconverted substrate remained in the biomass phase (10.4 mg). With a ritonavir solubility of 0.30 mg L^{-1} in water,[42] precipitation of ritonavir and accumulation in the biomass is likely. The poor solubility makes high substrate loading challenging.

After 48 h the metabolites M1 to M8 were formed (Figure 6). For further LC-MS^2 and NMR analysis the soluble phase was fractionated into nine fractions (F1–F9) by preparative HPLC. For the single oxidation products, M5-1 and M5-2, LC-MS^2 analysis revealed two different oxidation products on the isopropylthiazolyl moiety (Figure 7). Metabolite M5-2 was the
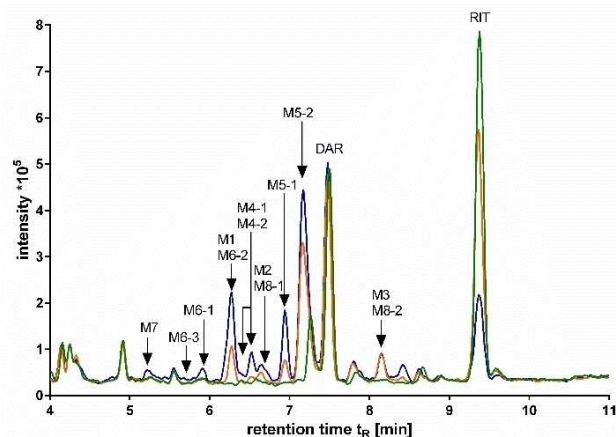
Figure 6. Product formation of biotransformation in fermentative scale. Shown is the TIC of the mass chromatogram. The metabolites M1 to M8 were formed. Green: 0 h, red: 24 h, blue: 48 h of conversion. Arrows indicate found metabolites.

Figure 7. Metabolites formed in the biotransformation on preparative scale. Localization of conversion was determined by LC-MS^2 or NMR analysis.

main metabolite formed and structure elucidation with NMR revealed the isopropylthiazolyl oxidation product at the isopropyl moiety. After fractionation 32 mg of M5-2 with a purity of 97.7 % could be isolated corresponding to a yield of 19 %. As pure isolation of the other metabolites was challenging the percentage of the metabolite in the different fractions was determined (Table 2). NMR analysis in combination with LC-MS^2 analysis of metabolite M5-1 indicated the hydroxylation at one of the methyl groups at the isopropyl moiety. Oxidation of metabolite M5-1 to a ketone (M8-2) was confirmed by LC-MS^2. LC-MS^2 of M-7 determined the single oxidation at the isopropyl moiety with the removal of the carbamoyl group indicating a further reaction product of M5-2. Characterization of M6-2 via LC-MS^2 showed both oxidations on the isopropylthiazolyl moiety. M6-2 is therefore a further product of M5-2. The structure was confirmed by ^{1}H NMR and ^{13}C NMR indicating one hydroxylation at the CH of the isopropyl moiety and the second hydroxy group at one of the methyl groups. For the two additional double hydroxylation products M6-1 and M6-3 an NMR analysis was not possible. However, analysis by LC-MS^2 identified the first hydroxylation on the isopropyl moiety. The second hydroxylation position was verified as one of the ethylbenzyl rings for metabolite M6-1. For M6-3 an exact localization was not possible. The des-isopropylthiazolyl prod-

uct (M1) and the decarbamoylated product (M2) were verified by LC-MS^2. Also the demethylation of M-3 was verified by LC-MS^2. LC-MS^2 of M4-1 and M4-2 verified the demethylation,

Table 2. Recovery of all metabolites with their m/z ratio and retention time (t_R) in the fractions after preparative HPLC. Determined was the mass of the whole fraction, total area of DAD detection (254 nm) and the percentage of the specific metabolite in the fraction. n.d. stands for not defined.

Fraction	Metabolite	m/z	t_R [min]	Mass of fraction [mg]	Total area	Area from total [%]
F1	M7	596	5.21	11.2	n.d.	n.d.
F2	M7	596	5.21	14.5	n.d.	n.d.
F3	M2	580	6.64	7.7	n.d.	n.d.
	M6-3	753	5.77		n.d.	n.d.
F4	M6-1	753	5.91	9.4	19.30	7.1
F5	M6-2	753	6.21	10.2	76.86	60.4
	M4-1	723	6.40		8.90	7.2
F6	M1	582	6.28	20.6	133.62	50.9
	M4-2	723	6.52		88.52	33.7
F7	M1	582	6.28	13.2	95.37	37.8
	M4-2	723	6.52		5.61	2.2
	M8-1	735	6.70		n.d.	n.d.
	M5-1	737	6.94		132.92	57.7
F8	M5-2	737	7.16	32.6	584.45	97.7
F9	M3	707	8.13	4.7	n.d.	n.d.
	M8-2	735	8.15		n.d.	n.d.

however, a differentiation of the hydroxylation positions was not possible. Both metabolites show hydroxylation on the isopropylthiazolyl moiety. This indicates that the same hydroxylation positions as found in the single hydroxylation products (M5-1 and M5-2) are present.

In general the scale-up in a stirred system showed a positive effect on biomass formation and ritonavir conversion. *A. mirum* is therefore a promising candidate for further biotechnological applications. As a full assembly of the genome is available, *A. mirum* is also a promising strain to undergo molecular analysis of the P450 enzymes.

Conclusions

Cytochrome P450 monooxygenases possess the to date unsurpassed ability to selectively insert oxygen into many complex substrates at ambient conditions. Thus, this efficient one-step conversion resulting in regio- and stereospecific products is a useful approach for the production of oxyfunctionalized molecules with desired pharmacological properties. In order to take full advantage of their catalytic potential, it is still essential to identify and investigate new P450 monooxygenases. Many genome sequencing projects revealed genes encoding putative P450s. However, it is still challenging to determine the specific functionality of a single P450 by structure comparison. Analyzing the CYPome of wild-type strains using a bioinformatics approach to identify promising whole-cell biocatalysts is a favorable alternative.

The data presented within this study demonstrates that strain selection by genome analysis combined with a medium-throughput screening enables the identification of new highly efficient biocatalysts for hydroxylation reactions. Using this approach we were able to generate a platform of novel whole-cell biocatalysts useful for oxfunctionalization of diverse compounds. This platform contains strains with high reactivity and broad acceptance for substrates of various sizes and with different properties. A further refinement of the search algorithms taking more parameters than number of sequences into account might be useful to come up with an even more tailored set. Furthermore, we were able to demonstrate for the conversion of ritonavir with the actinomycete *A. mirum* that the transfer of the biotransformation from the initial screening to a preparative scale already yields good product formation. A conversion of 90% and the formation of 13 different metabolites after 48 h was reached. Preparative HPLC could generate sufficient purity of the products for LC-MS^2 analysis or the structure elucidation by NMR. For the main metabolite M5-2 we were able to isolate 32 mg with a purity of 97.7% after first fractionation. The good scalability of the biotransformation shows the potential of the strain collection for further applications in biotechnology and synthetic biology. For that, optimization of the biotransformation, especially higher substrate loading, combined with reaction engineering approaches and efficient downstream processing has to be carried out.

Experimental Section

Genome Data Mining for the Identification of P450s

Genomes of sequenced organisms were mined using two different approaches based on different databases. In the first approach the conversed cysteine heme ligand signature specific for P450s (code "PDOC00081") that is deposited on the PROSITE was matched with protein sequences stored on the protein database UniProtKB. Per NCBI-Taxonomy ID the number of found P450 gene entries was counted resulting in a list of organisms ordered by their number of total P450 genes. Using the second approach based on the CYPED data base, gene sequences per NCBI-taxonomy ID were determined. Additionally, sequences could be categorized belonging to super- and homologous families. A sequence similarity cut-off of 55% classified the sequences in homologous families, whereas a cut-off of 40% placed the sequences into the same superfamily.

Growth and Screening Conditions

Growth of bacterial strains was performed in standard NL148 medium for precultivation and NL148sb medium for the main culture. NL148 contained glucose (22 g L^{-1}), casein peptone (8 g L^{-1}), meat extract (4 g L^{-1}), yeast extract (0.5 g L^{-1}), sodium chloride (1.5 g L^{-1}) and 3-(*N*-morpholino)propanesulfonic acid (MOPS) (21 g L^{-1}). NL148sb contained soluble starch (20 g L^{-1}), casein peptone (4 g L^{-1}), soya peptone (4 g L^{-1}), meat extract (4 g L^{-1}), yeast extract (0.5 g L^{-1}) and sodium chloride (1.5 g L^{-1}) and MOPS (21 g L^{-1}). The pH was adjusted to 6.5 for NL148 and 7.5 for NL148sb. Both media were supplemented with 1 mL L^{-1} trace element solution A consisting boric acid (0.1 g L^{-1}), iron sulphate (5 g L^{-1}), potassium iodide (0.05 g L^{-1}), cobalt chloride×6 H$_2$O (2 g L^{-1}), copper sulfate×5 H$_2$O (0.2 g L^{-1}), manganese chloride× 4H$_2$O (2 g L^{-1}), zinc sulfate×7 H$_2$O (4 g L^{-1}) and 1 mL sulfuric acid (97%). The standard medium used for the growth of fungi was NL148 s medium containing soluble starch (20 g L^{-1}), casein peptone (8 g L^{-1}), meat extract (4 g L^{-1}), yeast extract (0.5 g L^{-1}) and sodium chloride (1.5 g L^{-1}) and MOPS (21 g L^{-1}). The pH was adjusted to 6.5. NL148 s was supplemented with trace element solution B containing sodium molybdate×2 H$_2$O (0.03 g L^{-1}), zinc sulphate×7 H$_2$O (4.4 g L^{-1}), iron sulfate×7 H$_2$O (5.5 g L^{-1}), copper sulfate×5 H$_2$O (0.08 g L^{-1}), manganese chloride×4 H$_2$O (0.18 g L^{-1}) and 1 mL sulfuric acid (97%). Different media for pre- and main cultivation were used for the following strains: *Mycobacterium aromaticivorans* (NL148), *Mycobacterium hassiacum* (middlebrock), *Salinispora arenicola* (bacto marine broth), *Cystobacter fuscus* (SP), *Actinomadura atramentaria* (GYM), *Colletotrichum fioriniae* (oatmeal). Recipes are attached in the supplementary information. Precultures of microbial strains were grown in 100 mL shake flasks (1:5 filling volume) for 3 days, 24–37 °C 180 rpm, 2.5 cm shaking diameters, main-cultures were grown in 500 mL shake flasks (1:12 filling volume) for 3 days, 24–37 °C 180 rpm, 2.5 cm shaking diameters. The screening was performed in 96-deep well plates (Enzyscreen, System Duetz®), 500 μL filling volume, 28 °C, 200 rpm (shaking diameter = 5 cm). Samples were taken after 48 h and 72 h of incubation. Reaction was stopped by addition of 300 μL isopropanol and centrifugation at 4,400 rpm, 4 °C and 10 min. All screening experiments were performed in duplicates.

Substrates

Substrates were dissolved in different solvents (methanol, dimethylsulfoxide and isopropanol). The stock concentration was 10 mg mL^{-1} for amodiaquine (AMO), cyclosporine (CYC), olivitolic acid (OA), ritonavir (RIT), rapamycin (RAP) and testosterone (TEST).

Stock concentration of haloperidol (HAL) was 5 mg mL^{-1} and 2 mg mL^{-1} for tamoxifen (TAM). All substrates were used in a 1:100 ratio resulting in final concentrations of 0.02 to 0.1 mg mL^{-1}, respectively.

Analysis of Screening Products

Samples were analyzed by LC-MS using a 1260 Infinity II LC system equipped with Diode Array Detector (1260 DAD HS) combined with an Agilent Technologies 6120 Single Quadrupole LC-MS (Agilent Technologies). As a stationary phase a Poroshell 120 EC–C18 column, 4.6×100 mm, 2.7 µm (Agilent Technologies) was used. The column temperature was set to 30 °C (analysis of ritonavir, amodiaquine, tamoxifen, testosterone, and haloperidol) or 40 °C (analysis of rapamycin and cyclosporine). The flow rate was 1 mL min^{-1}. As running buffer for the analysis of ritonavir, tamoxifen, testosterone, and haloperidol 0.1% formic acid (v/v, solvent A) and 100% acetonitrile (solvent B) were used with the following gradient: 0–3 min: 5 to 40% B, 3–9.5 min: 40 to 70% B, 9.5–10.5 min: 70 to 95% B, 10.5 to 13 min: 95% B. For the analysis of rapamycin 0.1% formic acid+5 mM ammonium acetate (v/v, solvent A) and 100% methanol (solvent B) were used as running buffer with the gradient: 1–3 min: 5 to 40% B, 3–16 min: 40 to 95%, 16–20 min: 95 to 100% B. For cyclosporine two different methods were used. Method 1: 0.1% formic acid (v/v, solvent A) and 100% acetonitrile (solvent B) with the gradient: 0–3 min: 5 to 64% B, 3–10 min: 64 to 95% B, 10–12 min: 95% B. Method 2: 0.1% formic acid+5 mM ammonium acetate (v/v, solvent A) and 100% methanol (solvent B) with the gradient: 1–3 min: 5 to 40% B, 3–16 min: 40 to 95%, 16–20 min: 95 to 100% B. As running buffer for the analysis of AMO 0.1% formic acid (v/v, solvent A) and 100% acetonitrile (solvent B) were used with two different gradients. Gradient 1: 0–6 min: 5 to 68% B, 6–10 min: 68 to 95% B, 10–13 min: 95% B. Gradient 2: 0–7 min: 5 to 23% B, 7–8 min: 23 to 100% B, 8–9 min: 100% B. The quadrupole was interfaced with electron spray ionization (ESI). Following ESI parameters were set: drying gas temperature: 350 °C, nebulizer pressure: 4 bar, drying gas flow: 12 L min^{-1}, capillary voltage: 3,000 V. Samples were detected in a positive scan mode in an m/z range of 100 to 1,300. Wavelengths for DAD detection were as follows: ritonavir (254 nm), amodiaquine (342 nm), tamoxifen (270 nm), testosterone (245 nm), haloperidol (245 nm), rapamycin (278 nm), and cyclosporine (278 nm).

BioLector® Screening Conditions

48-well flower plates were filled with 787 µL medium and inoculated with 13 µL of a preculture grown in 20 mL NL148 medium for three days. The media M9 medium (glucose (4.0 g L^{-1}), magnesium sulfate×7 H$_2$O (0.25 g L^{-1}), calcium chloride×2 H$_2$O (1.47×10^{-2} g L^{-1}), thiamin-HCl (0.34 g L^{-1}), proline (0.1 g L^{-1}), disodium hydrogen phosphate×2 H$_2$O (3.99 g L^{-1}), potassium dihydrogenphosphate (3.0 g L^{-1}), sodium chloride (0.5 g L^{-1}), ammonium chloride (1.0 g L^{-1})), oatmeal medium (oatmeal (30.0 g L^{-1}), iron sulfate×7 H$_2$O (1.0×10^{-6} g L^{-1}), manganese(II) chloride×4 H$_2$O (1.0×10^{-6} g L^{-1}), zinc sulfate×7 H$_2$O (1.0×10^{-6} g L^{-1})), NL148sb medium (see above), GYM Streptomyces medium (glucose (4.0 g L^{-1}), yeast extract (4.0 g L^{-1}), malt extract (10.0 g L^{-1})), CY medium (casitone (3.0 g L^{-1}), calcium chloride (1.36 g L^{-1}), yeast extract (1.0 g L^{-1})) and TB medium (tryptone (10.8 g L^{-1}), yeast extract (21.6 g L^{-1}), glycerol (4.5 g L^{-1}), potassium dihydrogenphosphate (2.32 g L^{-1}), dipotassium hydrogen phosphate (12.54 g L^{-1})) were utilized. Cultivation was performed with a shaking frequency of 1200 min^{-1} at 28 °C for 120 h. 8 µL of the substrate (ritonavir stock: 10 mg mL^{-1}) was added after 72 or 20 h. The cultivation time was 120 h or 169 h (determination of total ritonavir conversion).

Samples were taken directly (0 h) and 48 h after substrate addition or at 0 h, 48 h, 72 h, 148 h, 196 h after substrate addition. After harvesting, 300 µL sodium carbonate solution (0.1 M, pH 9.6) was added and samples were extracted three times with an equal amount of ethyl acetate. Darunavir (100 µM) was added as an internal standard prior to extraction. Samples were evaporated to dryness and afterwards resolved in 800 µL methanol. Samples were analyzed by LC-MS (see above). Ritonavir conversion was calculated from the integrated substrate peak at a retention time of 9.4 min in the DAD signal at an absorption of 254 nm.

Biotransformation of Ritonavir in a Preparative Scale

For preparative scale fermentation a laboratory fermenter KLF 3.1 L and the BioScada software for process parameter control from Bioengineering AG was used. 1.7 L of GYM medium and 170 µL antifoam 204 (Sigma-Aldrich) were in situ sterilized for 20 min, 121 °C. Medium was inoculated with 27.6 mL of a preculture (grown 3 days in 100 ml shake flask, 28 °C, 180 rpm). 5 mL samples were taken in duplicates after substrate addition, after 4 h, 7 h, 20 h, 24 h, 44 h and 48 h. Samples were centrifuged at 4000 g for 5 minutes at 4 °C. The cell fraction was used for cell dry weight determination. 800 µL of the supernatant was extracted as described above and analyzed via LC-MS. 1 mL of the supernatant was used for the analysis of sugar concentration via HPLC Hitachi LaChrom Elite®. The column Repro-Gel H 8×300 mm, 9 µm (Maisch). The flow rate was 1 mL min^{-1}, isocratic with 5 mM H$_2$SO$_4$. The column temperature was set to 40 °C. The injection volume of samples was 20 µL. For detection the RI detector L2490 (35 °C, polarity positive) was used. Analysis was performed with the software EZChrome Elite Client/Server (Scientific Software Inc.). Calibration curves (4 g L^{-1}, 2 g L^{-1}, 1 g L^{-1}, 0.5 g L^{-1}, 0.25 g L^{-1}, and 0.125 g L^{-1}) of maltose and glucose were recorded.

Characterization of Ritonavir Metabolites by LC-MS^2

LC-MS^2 was performed with a Nucleoshell RP18 column, 2.0× 100 mm, 2.7 µm, using an Agilent 1260 Infinity LC system combined with a compact quadrupole time of flight (Q-TOF) mass spectrometer (Bruker Daltonics). The flow rate was set to 0.4 mL min^{-1}. A gradient of 0.1% formic acid (v/v, solvent A) and 100% acetonitrile (solvent B) was used. Solvent gradient was 0–3 min: 5 to 40% B, 3–9.5 min: 40 to 70% B, 9.5–10.5 min: 70 to 95% B, 10.5 to 13 min: 95% B. The Q-TOF was interfaced with electron spray ionization (ESI). Following ESI parameters were set: drying gas temperature: 220 °C, nebulizer pressure: 4.8 bar, drying gas flow: 12 L min^{-1}, capillary voltage: 4,500 V, end plate offset: 500 V. Quadrupole was used for selection of precursor ions and subjected to collision induced dissociation (CID) with nitrogen (collision energy: 28 eV). Analysis of fragments was performed with Q-TOF in a range of m/z 100 to 800. CID fragments for the 13 metabolites and ritonavir were recorded by multiple reaction monitoring (MRM, Table 3).

Fractionation of Fermentation Broth via Preparative HPLC

The soluble phase was fragmented by reverse phase HPLC with a NucleoDur C18 HTEC column, 10×250 mm, 5 µm (Macherey Nagel). The column temperature was set to 30 °C. The flow rate was 2 mL min^{-1}. A gradient of 0.1% trifluoroacetic acid (TFA) (v/v, solvent A) and 100% acetonitrile (solvent B) was used for the analysis. Solvent gradient was: 0–5 min: 35 to 50% B, 5–25 min: 50 to 100% B, 25–28 min: 100% B. Fractions were taken: F1: 7.6–8.25 min, F2: 8.3–8.75 min, F3: 9.35–10.65 min, F4: 10.7–11.7 min, F5: 11.7–12.6 min, F6: 12.7–13.3 min, F7: 13.3–14.0 min, F8: 14.35–

Table 3. CID fragments determined for the 13 different metabolites by MRM. Shown are the parent ™ion and the generated fragments (28 eV). n.d. stands for not defined.

Metabolite	t_R [min]	m/z	Parent ion m/z	CID fragments m/z										
RIT	9.36	721	361	426	296	285	268	268	197	171	140	–	91	98
M1	6.28	582	582	426	157	–	268	–	129	–	–	–	91	98
M2	6.64	580	290	–	–	285	–	268	197	171	140	–	91	–
M3	8.13	707	354	–	–	–	–	–	183	–	140	–	91	98
M4-1	6.40	723	362	–	–	285	268	270	199	173	156	–	91	98
M4-2	6.52	723	362	–	–	285	268	270	199	173	156	–	91	98
M5-1	6.94	737	369	–	312	–	268	284	213	187	156	–	91	98
M5-2	7.16	737	369	–	312	–	268	284	213	187	156	155	91	98
M6-1	5.91	753	377	–	–	301	–	284	213	187	156	155	107	98
M6-2	6.21	753	377	–	–	285	268	300	229	203	172	155	91	98
M6-3	5.77	753	377	–	–	–	–	–	–	187	156	155	91	98
M7	5.21	596	596	–	312	285	268	–	–	187	156	155	91	98
M8-1	6.70	735	368	n.d.	–	–	–	–	–	–	–	–	–	–
M8-2	8.15	735	368	–	–	268	284	211	185	154	–	91	98

15.2 min, F9: 17.6–18.5 min. Elution of compounds was monitored with a DAD at 254 nm.

NMR Analysis

NMR spectra were recorded on a Bruker AV 600 or 700 Avance III HD system with methanol-d4 as solvent and internal standard. The solvent signals were referenced to δ_H 3.31 ppm and δ_C 49.0 ppm.

Ritonavir: ^{1}H NMR (600 MHz, [D4] MeOH): $\delta = 8.95$ (s, 1H; CH), 7.81 (s, 1H; CH), 7.21 (m, 4H; Ar–H), 7.15 (m, 6H, Ar–H), 7.14 (s, 1H; CH), 5.19 (s, 2H; CH$_2$), 4.55 (d, ^{3}JH,H = 16.2 Hz, 1H; CH$_2$), 4.47 (d, ^{3}JH,H = 16.2 Hz, 1H; CH$_2$), 4.30 (dddd, 1H; CH), 3.97 (ddd, 1H; CH), 3.94 (d, ^{3}JH,H = 7.5 Hz, 1H; CH), 3.69 (dt, ^{3}JH,H = 7.0 Hz, 2.0 Hz, 1H; CH), 3.30 (m, 1H; CH), 2.97 (s, 3H; CH$_3$), 2.81 (dd, ^{3}JH,H = 13.8 Hz, 5.7 Hz, 1H; CH$_2$), 2.77 (dd, 1H; CH$_2$), 2.76 (dd, 1H; CH$_2$), 2.75 (dd, ^{3}JH,H = 13.6 Hz, 8.2 Hz, 1H; CH$_2$), 1.93 (m, 1H; CH), 1.66 (ddd, 1H; CH$_2$), 1.59 (ddd, 1H; CH$_2$), 1.37 (d, ^{3}JH,H = 6.9 Hz, 6H; CH$_3$), 0.84 (d, ^{3}JH,H = 6.8 Hz, 3H; CH$_3$), 0.83 (d, ^{3}JH,H = 6.8 Hz, 3H; CH$_3$)

Ritonavir: ^{13}C NMR (600 MHz, [D4] MeOH): $\delta = 180.5$ (CCH), 174.1 (CO), 160.1 (CO), 158.0 (CO), 156.8 (CH), 153.7 (CCH$_2$), 143.6 (CH), 140.2 (CCH$_2$), 139.6 (CCH$_2$), 136.0 (CCH$_2$), 130.6 (CH), 130.4 (CH), 129.3 (CH), 129.2 (CH), 127.2 (CH), 127.2 (CH), 115.5 (CH), 70.9 (CH), 61.9 (CH), 58.9 (CH$_2$), 57.2 (CH), 49.9 (CH$_2$), 49.5 (CH), 41.7 (CH$_2$), 40.0 (CH$_2$), 39.3 (CH$_2$), 35.2 (CH$_3$), 34.4 (CH), 32.0 (CH), 23.6 (CH$_3$), 23.5 (CH$_3$), 20.0 (CH$_3$), 18.8 (CH$_3$)

M5-2: ^{1}H NMR (700 MHz, [D4] MeOH): $\delta = 9.12$ (s, 1H; CH), 7.88 (s, 1H; CH), 7.21 (m, 4H; Ar–H), 7.15 (m, 6H, Ar–H), 7.15 (s, 1H; CH), 5.18 (s, 2H; CH$_2$), 4.51 (s, 2H; CH$_2$), 4.30 (m, 1H; CH), 3.97 (m, 1H, CH), 3.94 (d, ^{3}JH,H = 7.6 Hz, 1H; CH), 3.70 (dt, 1H; CH), 2.97 (s, 3H; CH$_3$), 2.81 (dd, ^{3}JH,H = 16.0 Hz, 6.5 Hz, 1H; CH$_2$), 2.77 (dd, 1H; CH$_2$), 2.76 (dd, 1H; CH$_2$), 2.75 (dd, ^{3}JH,H = 13.6 Hz, 8.2 Hz, 1H; CH$_2$), 1.91 (m, 1H; CH), 1.67 (m, 1H; CH$_2$), 1.60 (m, 1H; CH$_2$), 1.60 (d, ^{3}JH,H = 3.5 Hz, 6H; CH$_3$), 0.85 (d, ^{3}JH,H = 6.8 Hz, 3H; CH$_3$), 0.82 (d, ^{3}JH,H = 6.8 Hz, 3H; CH$_3$)

M5-2: ^{13}C NMR (700 MHz, [D4] MeOH): $\delta = 183.6$ (CCH), 174.0 (CO), 160.1 (CO), 158.0 (CO), 157.4 (CH), 153.9 (CCH$_2$), 142.4 (CH), 140.1 (CCH$_2$), 139.6 (CCH$_2$), 136.7 (CCH$_2$), 130.6 (CH), 130.6 (CH), 130.4 (CH), 129.2 (CH), 129.2 (CH), 127.3 (CH), 127.2 (CH), 116.5 (CH), 74.0 (COH), 70.9 (CH), 61.9 (CH), 58.8 (CH$_2$), 57.2 (CH), 49.8 (CH$_2$), 49.5 (CH), 41.7 (CH$_2$), 40.0 (CH$_2$), 39.3 (CH$_2$), 35.2 (CH$_3$), 32.1 (CH), 30.9 (CH$_3$), 30.9 (CH$_3$), 20.1 (CH$_3$), 18.9 (CH$_3$)

M5-1: ^{1}H NMR (600 MHz, [D4] MeOH): $\delta = 9.07$ (s, 1H; CH), 7.87 (s, 1H; CH), 7.21 (m, 4H; Ar–H), 7.15 (m, 6H, Ar–H), 7.15 (s, 1H; CH), 5.19 (s, 2H; CH$_2$), 4.53 (s, 2H; CH$_2$), 4.30 (m, 1H; CH), 3.97 (m, 1H, CH), 3.94 (d, 1H; CH), 3.70 (dt, 1H; CH), 3.30 (m, 1H; CH), 2.97 (s, 3H; CH$_3$), 2.81 (dd, ^{3}JH,H = 13.9 Hz, 5.5 Hz, 1H; CH$_2$), 2.77 (m, 1H; CH$_2$), 2.76 (m, 1H; CH$_2$), 2.75 (m, 1H; CH$_2$), 1.91 (m, 1H; CH), 1.67 (m, 1H; CH$_2$), 1.60 (m, 1H; CH$_2$), 1.37 (s, 3H; CH$_3$), 0.85 (d, 3H; CH$_3$), 0.82 (d, ^{3}JH,H = 6.2 Hz, 3H; CH$_3$)

M5-1: ^{13}C NMR (600 MHz, [D4] MeOH): $\delta = 180.5$ (n.d., CCH), 174.1 (CO), 160.1 (CO), 158.0 (CO), 156.8 (n.d., CH), 153.3 (CCH$_2$), 143.6 (n.d., CH), 140.2 (CCH$_2$), 139.6 (CCH$_2$), 136.0 (n.d., CCH$_2$), 130.6 (CH), 130.4 (CH), 129.3 (CH), 129.2 (CH), 127.2 (CH), 127.2 (CH), 116.3 (CH), 70.9 (CH), 61.9 (CH), 58.9 (CH2), 57.2 (CH), 49.9 (CH$_2$), 49.5 (CH), 41.7 (CH$_2$), 40.0 (CH$_2$), 39.3 (CH$_2$), 35.2 (CH$_3$), 34.4 (n.d., CH), 32.0 (CH), 20.0 (CH$_3$), 18.8 (CH$_3$), 17.1 (CH$_3$)

M6-2: ^{1}H NMR (600 MHz, [D4] MeOH): $\delta = 9.12$ (n.d., 1H; CH), 7.88 (n.d., 1H; CH), 7.21 (m, 4H; Ar–H), 7.15 (m, 6H, Ar–H), 7.14 (s, 1H; CH), 5.21 (s, 2H; CH$_2$), 4.51 (s, 2H; CH$_2$), 4.30 (m, 1H; CH), 3.97 (m, 1H, CH), 3.94 (m, 1H; CH), 3.70 (m, 1H; CH$_2$), 3.54 (m, 1H; CH$_2$), 3.51 (m, 1H; CH$_2$), 2.97 (s, 3H; CH$_3$), 2.81 (m, 1H; CH$_2$), 2.77 (m, 1H; CH$_2$), 2.76 (m, 1H; CH$_2$), 2.75 (m, 1H; CH$_2$), 1.91 (m, 1H; CH), 1.67 (m, 1H; CH$_2$), 1.60 (m, 1H; CH$_2$), 1.14 (d, 3H; CH$_3$), 0.85 (d, 3H; CH$_3$), 0.83 (d, 3H; CH$_3$)

M6-2: ^{13}C NMR (600 MHz, [D4] MeOH): $\delta = 180.5$ (n.d., CCH), 174.1 (CO), 160.1 (CO), 158.0 (CO), 156.8 (n.d., CH), 153.6 (CCH$_2$), 143.6 (n.d., CH), 140.2 (CCH$_2$), 139.7 (CCH$_2$), 136.0 (n.d., CCH$_2$), 130.7 (CH), 130.6 (CH), 130.4 (CH), 129.3 (CH), 129.2 (CH), 127.3 (CH), 127.3 (CH), 115.5 (n.d., CH), 76.7 (COH), 73.9 (CH$_2$), 70.9 (CH), 61.9 (n.d., CH), 58.9 (n.d., CH$_2$), 57.2 (CH), 49.9 (CH$_2$), 49.5 (CH), 41.7 (CH$_2$), 40.0 (CH$_2$), 39.3(CH$_2$), 35.2 (CH$_3$), 32.1 (CH), 20.0 (CH$_3$), 18.9 (CH$_3$), 17.2 (CH$_3$)

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] R. Bernhardt, *J. Biotechnol.* **2006**, *124*, 128–145.
[2] F. T. Peters, C.-A. Dragan, D. R. Wilde, M. R. Meyer, J. Zapp, M. Bureik, H. H. Maurer, *Biochem. Pharmacol.* **2007**, *74*, 511–520.
[3] M. Hans, J. Laan, B. Meijrink, W. van Scheppingen, R. Kerkman, M. van den Berg, M. Kittelmann, A. Kuhn, A. Riepp, J. Kühnöl, A. Fredenhagen, L. Oberer, O. Ghisalba, S. Lütz, D. Mangan, T. Moody, D. Schmid, A. Osorio-Lozada, F. Özde. Ütkür, M. L. Thompson, in *Practical Methods for Biocatalysis and Biotransformations 2* (Eds.: J. Whittall, P. W. Sutton), John Wiley & Sons, Hoboken, **2012**, pp. 129–162.
[4] J. Genovino, S. Lütz, D. Sames, B. B. Touré, *J. Am. Chem. Soc.* **2013**, *135*, 12346–12352.
[5] O. Ghisalba, M. Kittelmann, in *Modern Biooxidation: Enzymes, Reactions and Applications* (Eds.: R. D. Schmid, V. B. Urlacher), Wiley-VCH, Weinheim, **2007**, pp. 211–232.
[6] A. W. Munro, D. G. Leys, K. J. McLean, K. R. Marshall, T. W. B. Ost, S. Daff, C. S. Miles, S. K. Chapman, D. A. Lysek, C. C. Moser, C. C. Page, P. L. Dutton, *Trends Biochem. Sci.* **2002**, *27*, 250–257.
[7] A. Rentmeister, T. R. Brown, C. D. Snow, M. N. Carbone, F. H. Arnold, *ChemCatChem* **2011**, *3*, 1065–1071.
[8] C.-H. Yun, K.-H. Kim, D.-H. Kim, H.-C. Jung, J.-G. Pan, *Trends Biotechnol.* **2007**, *25*, 289–298.
[9] A. M. Sawayama, M. M. Y. Chen, P. Kulanthaivel, M.-S. Kuo, H. Hemmerle, F. H. Arnold, *Chemistry* **2009**, *15*, 11723–11729.
[10] D. R. Nelson, *Hum. Genomics* **2009**, *4*, 59–65.
[11] M. Fischer, M. Knoll, D. Sirim, F. Wagner, S. Funke, J. Pleiss, *Bioinformatics* **2007**, *23*, 2015–2017.
[12] J. Rolf, K. Rosenthal, S. Lütz, *Catalysts* **2019**, *9*, 190.
[13] K. McLean, D. Leys, A. Munro, in *Cytochrome P450: Structure, Mechanism, and Biochemistry* (Ed.: P. R. Ortiz de Montellano), Springer US, Boston, **2015**, pp. 261–407.
[14] K. Ohta, H. Agematu, T. Yamada, K. Kaneko, T. Tsuchida, *J. Biosci. Bioeng.* **2005**, *99*, 390–395.
[15] A. Worsch, F. K. Eggimann, M. Girhard, C. J. von Bühler, F. Tieves, R. Czaja, A. Vogel, C. Grumaz, K. Sohn, S. Lütz, M. Kittelmann, V. B. Urlacher, *Biotechnol. Bioeng.* **2018**, *115*, 2156–2166.
[16] a) F. Chen, W. Chen, M.-K. Lee, J.-H. Yu, C. Jefcoate, S.-C. Kim, *Genome Biol.* **2014**, *6*, 1620–1634; b) J. Shin, J.-E. Kim, Y.-W. Lee, H. Son, *Toxins (Basel)* **2018**, *10*, 112.
[17] A. Fredenhagen, K. Schroer, H. Schröder, D. Hoepfner, M. Ligibel, L. Porchet Zemp, C. Radoch, E. Freund, A. Meishammer, *ChemBioChem* **2019**, *20*, 650–654.
[18] Ł. Gricman, C. Vogel, J. Pleiss, *Proteins Struct. Funct. Bioinf.* **2015**, *83*, 1593–1603.
[19] D. C. Lamb, M. R. Waterman, B. Zhao, *Expert Opin. Drug Metab. Toxicol.* **2013**, *9*, 1279–1294.
[20] M. Parvez, L. B. Qhanya, N. T. Mthakathi, I. K. R. Kgosiemang, H. D. Bamal, N. S. Pagadala, T. Xie, H. Yang, H. Chen, C. W. Theron, R. Monyaki, S. C. Raselemane, V. Salewe, B. L. Mongale, R. G. Matowane, S. M. H. Abdalla, W. I. Booi, M. van Wyk, D. Olivier, C. E. Boucher, D. R. Nelson, J. A. Tuszynski, J. M. Blackburn, J.-H. Yu, S. S. Mashele, W. Chen, K. Syed, *Sci. Rep.* **2016**, *6*, 33099.
[21] J. D. Rudolf, C.-Y. Chang, M. Ma, B. Shen, *Nat. Prod. Rep.* **2017**, *34*, 1141–1172.
[22] D. C. Lamb, B. Zhao, P. F. Guengerich, S. L. Kelly, M. R. Waterman, in *Streptomyces : molecular biology and biotechnology* (Ed.: P. J. Dyson), Caister Academic Press, Norfolk, **2011**.
[23] H. Ichinose, *Biol. Pharm. Bull.* **2012**, *35*, 833–837.
[24] V. Moktali, J. Park, N. D. Fedorova-Abrams, B. Park, J. Choi, Y.-H. Lee, S. Kang, *BMC Genomics* **2012**, *13*, 525.
[25] M. Ide, H. Ichinose, H. Wariishi, *Arch. Microbiol.* **2012**, *194*, 243–253.
[26] H. Ouellet, J. B. Johnston, P. R. Ortiz de Montellano, *Arch. Biochem. Biophys.* **2010**, *493*, 82–95.
[27] L. M. Senate, M. P. Tjatji, K. Pillay, W. Chen, N. M. Zondo, P. R. Syed, F. C. Mnguni, Z. E. Chiliza, H. D. Bamal, R. Karpoormath, T. Khoza, S. S. Mashele, J. M. Blackburn, J.-H. Yu, D. R. Nelson, K. Syed, *Sci. Rep.* **2019**, *9*, 3962.
[28] L.-H. Xu, S. Fushinobu, S. Takamatsu, T. Wakagi, H. Ikeda, H. Shoun, *J. Biol. Chem.* **2010**, *285*, 16844–16853.
[29] S. Lee, J. W. Park, S. R. Park, J. Seog Ahn, C. Choi, Y. Joon Yoon, *J. Microbiol. Biotechnol.* **2006**, *16(6)*, 974.
[30] N. Sawada, T. Sakaki, S. Yoneda, T. Kusudo, R. Shinkyo, M. Ohta, K. Inouye, *Biochem. Biophys. Res. Commun.* **2004**, *320*, 156–164.
[31] B. Črešnar, Š. Petrič, *Biochim. Biophys. Acta Proteins Proteomics* **2011**, *1814*, 29–35.
[32] T. Hasegawa, M. P. Lechevalier, H. A. Lechevalier, *Int. J. Syst. Bacteriol.* **1978**, 304–310.
[33] K. Watanabe, T. Okuda, K. Yokose, T. Furumai, H. B. Maruyama, *J. Antibiot.* **1983**, *63*, 321–324.
[34] T. Li, Y. Fan, K. Nambou, F. Hu, T. Imanaka, L. Wei, Q. Hua, *Appl. Biochem. Biotechnol.* **2015**, *175*, 2845–2856.
[35] a) M. T. Granfors, J.-S. Wang, L. I. Kajosaari, J. Laitila, P. J. Neuvonen, J. T. Backman, *Basic Clin. Pharmacol. Toxicol.* **2006**, *98*, 79–85; b) L. L. von Moltke, D. J. Greenblatt, J. M. Grassi, B. W. Granda, S. X. Duan, S. M. Fogelman, J. P. Daily, J. S. Harmatz, R. I. Shader, *J. Clin. Pharmacol.* **1998**, *38*, 106–111.
[36] a) G. N. Kumar, A. D. Rodrigues, A. M. Buko, J. F. Denissen, *J. Pharmacol. Exp. Ther.* **1996**, *277*, 423–431; b) J. F. Denissen, B. A. Grabowski, M. K. Johnson, A. M. Buko, D. J. Kempf, S. B. Thomas, B. W. Surber, *Drug Metab. Dispos.* **1997**, *25*, 489–489.
[37] F. Li, J. Lu, X. Ma, *Chem. Res. Toxicol.* **2011**, *24*, 2109–2114.
[38] R. N. Rao, B. Ramachandra, R. M. Vali, S. S. Raju, *J. Pharm. Biomed. Anal.* **2010**, *53*, 833–842.
[39] S. N. Williams, E. Dunham, C. A. Bradfield, in *Cytochrome P450: Structure, Mechanism, and Biochemistry* (Ed.: P. R. Ortiz de Montellano), Springer US, Boston, **2005**, pp. 323–346.
[40] L. M. Blank, B. E. Ebert, B. Bühler, A. Schmid, *Biotechnol. Bioeng.* **2008**, *100*, 1050–1065.
[41] M. Miura, K. Ito, M. Hayashi, M. Nakajima, T. Tanaka, S.-i. Ogura, *PLoS One* **2015**, *10*, 1–9.
[42] D. A. DeGoey, D. J. Grampovnik, W. J. Flosi, K. C. Marsh, X. C. Wang, L. L. Klein, K. F. McDaniel, Y. Liu, M. A. Long, W. M. Kati, A. Molla, D. J. Kempf, *J. Med. Chem.* **2009**, *52*, 2964–2970.