DOI: 10.1002/cbic.200800799

Rational Design of a Minimal and Highly Enriched CYP102A1 Mutant Library with Improved Regio-, Stereo- and Chemoselectivity

Alexander Seifert,[a] Sandra Vomund,[a] Katrin Grohmann,[b] Sebastian Kriening,[b] Vlada B. Urlacher,[a] Sabine Laschat,[b] and Jürgen Pleiss*[a]

A minimal CYP102A1 mutant library of only 24 variants plus wild type was constructed by combining five hydrophobic amino acids (alanine, valine, phenylalanine, leucine and isoleucine) in two positions. Both positions are located close to the centre of the haem group. The first, position 87, has been shown to mediate substrate specificity and regioselectivity in CYP102A1. The second hotspot, position 328, was predicted to interact with all substrates during oxidation and has previously been identified by systematic analysis of 31 crystal structures and 6300 sequences of cytochrome P450 monooxygenases. By systematically altering the size of the side chains, a broad range of binding site shapes was generated. All variants were functionally expressed in *E. coli*. The library was screened with four terpene substrates geranylacetone, nerylacetone, (4*R*)-limonene and (+)-valencene. Only three variants showed no activity towards all four terpenes, while eleven variants demonstrated either a strong shift or improved regio- or stereoselectivity during oxidation of at least one substrate as compared to CYP102A1 wild type.

Introduction

Cytochrome P450 monooxygenase CYP102A1 from *Bacillus megaterium* (also known as P450-BM3) is a widely used and well-investigated monooxygenase. It is one of the most promising cytochrome P450 monooxygenases (CYPs) for preparative synthesis, since it is a soluble fusion protein of the monooxygenase and a diflavin reductase[1] and is relatively stable under process conditions.[2] The wild type enzyme is a highly active fatty acid hydroxylase[3] and its substrate profile has been widened by mutagenesis.[4] While the wild type enzyme is unselective, mutants have been described with increased regio- and stereoselectivity for different substrates. Among others, position 87 which resides in the substrate recognition site 1[5] has a strong impact on substrate specificity and regioselectivity for different substrates.[6–9] The CYP102A1 crystal structure reveals that the residue in this position is involved in the formation of the substrate binding cavity close to the haem centre. Molecular dynamics simulations of the complex of CYP102A1 with α-pinene revealed that different residues in this position cause changes in regioselectivity by stabilising the unpolar substrate in different binding conformations.[10] A second hotspot for regioselectivity and substrate specificity in CYP102A1 (position 328) was identified by systematic analysis of 31 crystal structures and 6300 sequences.[11] This study predicted that in 98% of all CYP sequences the residue localised in position 5 after the highly conserved ExxR motif is located in the immediate vicinity of the haem group. Its side chain points towards the haem centre and therefore is expected to interact with all substrates during oxidation. The study further revealed that residues in this position are predominantly hydrophobic. Previously, position 328 in CYP102A1 has been shown to play an important role in enantio- and regioselectivity towards alkanes and alkenes.[12,13] A funnel like haem access region in human CYP2C9 has been previously shown by molecular dynamics simulations to restrict possible orientations of the substrate (*S*)-warfarin close to the haem centre, and hence to mediate regioselectivity.[14] Hence, the hotspot positions 87 and 328 in CYP102A1 are expected to influence the accessibility of the activated haem oxygen from opposite sides of the haem access channel (Figure 1). Due to their spatial closeness, we expect a cooperative effect of both positions on substrate access to the activated oxygen and therefore on activity, specificity and selectivity.

Four differently sized and shaped terpenes [(4*R*)-limonene (**6**), nerylacetone [(*Z*)-**1**], geranylacetone [(*E*)-**1**] and (+)-valencene (**11**)] were chosen to screen the library. Oxidation of each compound can lead to interesting and valuable products. (4*R*)-limonene (**6**), the major constituent of citrus peel essential oils, has been shown to be oxidised by CYP102A7 to racemic mixtures of (4*R*)-limonene-1,2-epoxide (**7**), (4*R*)-limonene-8,9-epoxide (**8**) and carveol (**9**).[15] It is further known to be regio- and stereoselectively converted by *Xanthobacter* sp. C20 to form (4*R*,8*R*)-limonene-8,9-epoxide (**8**).[16] It has been suggested that a CYP catalyses this oxidation reaction, however, the respective enzyme has not yet been isolated.

[a] *A. Seifert, S. Vomund, Priv.-Doz. Dr. V. B. Urlacher, Prof. Dr. J. Pleiss*
Institute of Technical Biochemistry, University of Stuttgart
Allmandring 31, 70569 Stuttgart (Germany)
Fax: (+49) 711-685-3196
E-mail: juergen.pleiss@itb.uni-stuttgart.de

[b] *K. Grohmann, S. Kriening, Prof. Dr. S. Laschat*
Institute of Organic Chemistry, University of Stuttgart
Pfaffenwaldring 55, 70569 Stuttgart (Germany)

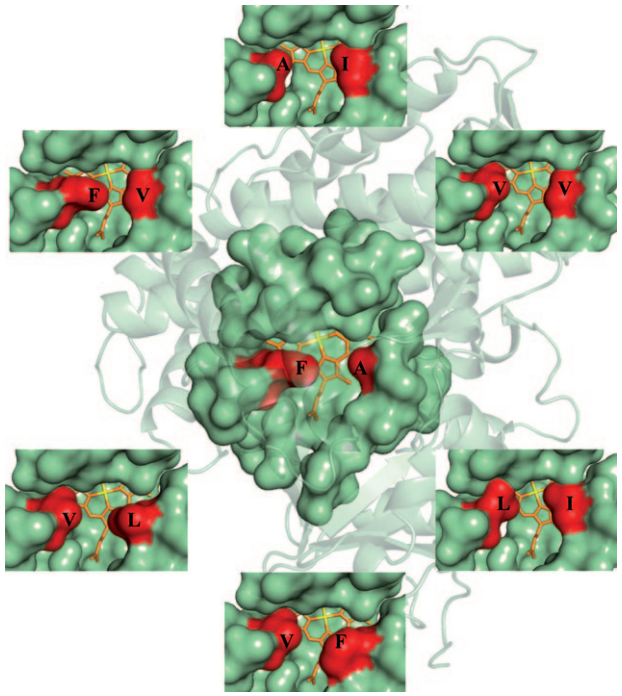
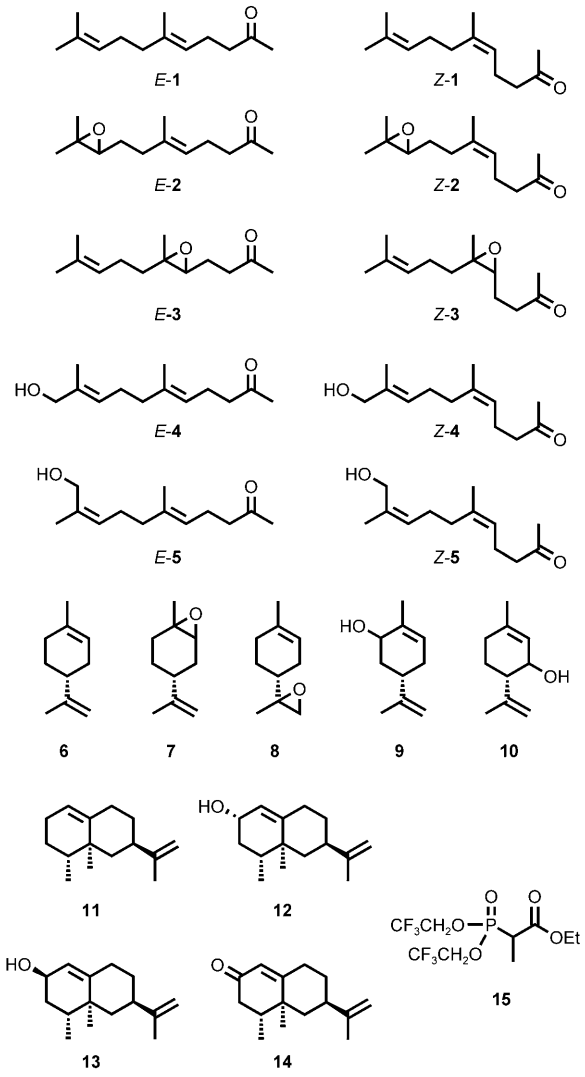


Figure 1. Shape of the substrate binding site in CYP102A1 wild type (centre) near to the active haem (yellow). The mutated positions 87 (left) and 328 (right) are shown in red. The high diversity of substrate binding site shapes of six highly selective variants A328V (FV), F87A/A328I (AI), F87V/A328V (VV), F87L/A328I (LI), F87V/A328F (VF) and F87V/A328L (VL) are visualised in the insets. Models are based on CYP102A1 crystal structure (PDB entry 1bu7 chain A) and were generated with the Pymol 0.99 program.[46]

(+)-Nootkatone (**14**), the most expensive aromatic of grapefruit, can be obtained by the selective oxidation of (+)-valencene (**11**), a cheap constituent of orange oil. Besides chemical oxidation, the biotransformation of (+)-valencene (**11**) by *Chlorella* and *Mucor* species has been reported.[17] Furthermore, cytochrome P450 monooxygenases P450cam, CYP102A1 and mutants thereof have been shown to oxidise (+)-valencene (**11**).[18] P450cam mutants showed relatively high regioselectivity for C2 oxidation (85%), however, activity was rather low. On the other hand, CYP102A1 variants showed higher activity towards (+)-valencene (**11**), but were unselective. Hence, CYP102A1 variants with higher selectivity are needed. Furthermore, CYP102A1 has also been shown to unselectively oxidise geranylacetone to three products.[19] In the same study a triple mutant (R47L/Y51F/F87V) was introduced that catalyses the epoxidation of geranylacetone [(*E*)-**1**] to 9,10-epoxygeranylacetone [(*E*)-**2**] with high regio- and stereoselectivity. The same mutant also oxidises nerylacetone [(*Z*)-**1**] to five different products. However, allylic alcohols in positions C11 and C12 of the two substrates have not (or only in minor amounts) been detected among the products of the corresponding terpenes. These hydroxy products are valuable starting materials for the total syntheses of natural products such as smenochromene D,[20] indole alkaloids,[21] pseudopteranes, furanocembranes,[22] brown algae-derived linear C18 terpenoids,[23] antitumor cembrane lactones crassin and isolobophytolide[24,25] and macrocyclic terpenoids humulene, flexibilene, helminthogermacrene and β-elemene[26–28] as well as cyclopropane-derived materials.[29,30] Thus, there is a need for CYP102A1 variants with changed chemoselectivity. By systematically combining five hydrophobic residues in both hotspot positions a minimal mutant library was constructed which can due to its small size be rapidly screened for variants with changed chemoselectivity and/or improved regio- and stereoselectivity.



Results

Terpene oxidation

A minimal library of 24 variants plus wild type was constructed by combination of five hydrophobic amino acids (alanine, valine, phenylalanine, leucine and isoleucine) in two positions (87 and 328), which have been previously identified as hotspots for selectivity. Altering the side chain size of these two amino acids drastically changes the shape of the substrate binding cavity in close vicinity of the haem group (Figure 1). The presence of the Soret-band at 450 nm upon measuring of CO-difference spectra indicated that all 24 variants and the wild type functionally incorporated the haem group during ex-

pression. The CYP102A1 wild type enzyme converted all terpenes [geranylacetone [(*E*)-**1**], nerylacetone [(*Z*)-**1**], (4*R*)-limonene (**6**) and (+)-valencene (**11**)], however with poor regioselectivity and activity. The screening of the mutant library revealed eleven variants that convert at least one terpene to a new valuable product and possess strongly increased selectivity as compared to the wild type enzyme (Table 1). While four single mutants in positions 87 or 328 and one double mutant showed improved or strongly shifted regio- or chemoselectivity towards the small acyclic terpenes geranyl- or nerylacetone [(*E*)-**1**, (*Z*)-**1**], amino acid substitution in both positions were necessary to improve selectivity towards the more bulky substrates (4*R*)-limonene (**6**) and (+)-valencene (**11**). Only three variants (A328F, F87I/A328F and F87A/A338V) did not convert any of the four terpenes. In the following we focus only on the variants with the highest regio- and stereoselectivity, strongly shifted product profile, and variants which produce new and interesting products.

Table 1. Comparison of terpene oxidation catalysed by CYP102A1 variants designed in this study. Black cells indicate variants that show the strongest shifts in regio- or chemoselectivity, or have the highest regio- or stereoselectivity as compared to wild type. Enzyme–substrate combinations with no or only minor improvement as compared to wild type are depicted in white. Grey cells indicate very little or no conversion of the respective substrates.

CYP102A1 variants	(4*R*)-(+)-limonene	geranyl acetone	neryl acetone	(+)-valencene
F87A/A328	□	□	□	□
F87A/A328I	□	□	□	■
F87A/A328L	□	□	□	▒
F87A/A328V	▒	▒	▒	▒
F87A/A328F	■	□	□	▒
F87I/A328	□	■	□	▒
F87I/A328I	□	□	▒	▒
F87I/A328L	□	▒	▒	▒
F87I/A328V	□	□	□	▒
F87I/A328F	▒	▒	▒	▒
F87L/A328	▒	■	▒	▒
F87L/A328I	□	□	□	■
F87L/A328L	□	□	□	▒
F87L/A328V	▒	□	▒	▒
F87L/A328F	▒	□	▒	▒
F87V/A328	□	■	□	□
F87V/A328I	□	□	□	■
F87V/A328L	□	■	□	▒
F87V/A328V	□	□	□	■
F87V/A328F	■	□	▒	▒
F87/A328[a]	□	□	□	□
F87/A328I	□	▒	▒	▒
F87/A328L	□	▒	▒	▒
F87/A328V	□	■	■	▒
F87/A328F	▒	▒	▒	▒

[a] CYP102A1 wild type.

Table 2. Comparison of activity of CYP102A1 wild type and the most selective variants (Table 3) towards (4*R*)-(+)-limonene (**6**), geranylacetone [(*E*)-**1**], nerylacetone [(*Z*)-**1**] and (+)-valencene (**11**). The substrate oxidation rate is defined as $a/(c{\cdot}t)$ [μmol·(μmol CYP)$^{-1}$·min^{-1}], with converted substrate a [μM], enzyme concentration c [μM] and time t to consume the NADPH [min].

CYP102A1 variants	Substrate	NADPH turnover [min^{-1}]	Oxidation rate [min^{-1}]	Conversion after 1 h [%]
wild type	**6**	112±7.3	54.7±4.6	48.7±1
F87A/A328F	**6**	88±12	60±7.4	68±1.3
F87V/A328F	**6**	277±8.3	181±4.2	65.3±1
wild type	(*E*)-**1**	478±38	324±17	67.9±2
F87V	(*E*)-**1**	133±6	83±3	62.6±2.3
F87I	(*E*)-**1**	64±9	33±4	51.1±2.6
F87L	(*E*)-**1**	17±4	6±1	38.2±1.3
A328V	(*E*)-**1**	1145±105	550±72	48.1±4.4
F87V/A328L	(*E*)-**1**	300±23	152±11	50±3.8
wild type	(*Z*)-**1**	301±27.4	155±6.9	51.6±2.6
A328V	(*Z*)-**1**	594±24.7	338±9.8	56.9±1
wild type	**11**	43±3.3	10±2	22.9±3.1
F87A/A328I	**11**	25.3±5.6	12.8±3.7	50.3±3.9
F87V/A328I	**11**	54.4±10.3	15.7±3.3	28.7±1.7
F87L/A328I	**11**	17.4±3.6	4.2±0.7	24.6±5.3
F87V/A328V	**11**	63.5±4.2	24.2±2.5	38.3±5.0

Best geranylacetone-converting variants

Geranylacetone [(*E*)-**1**] is converted by the CYP102A1 wild type enzyme to two products (Table 3). The main product is 9,10-epoxygeranylacetone (*E*)-**2** (54%). The second product at a retention time of 12.3 min could not be identified. 11-Hydroxy- and 12- hydroxygeranylacetone [(*E*)-**4**, (*E*)-**5**] are not among the products. The screening of our minimal library showed three single mutants (F87V, F87I and F87L) with a strong preference for the formation of 9,10-epoxygeranylacetone [(*E*)-**2**]. Variant F87L produced 9,10-epoxygeranylacetone [(*E*)-**2**] almost exclusively (97%). Furthermore, nine variants were, in contrast to the CYP102A1 wild type, able to hydroxylate geranylacetone at allylic positions C11 and C12. Two of these nine variants showed a strong shift in chemoselectivity. Variant A328 V produced the highest amount of 11-hydroxygeranylacetone [(*E*)-**4**, 37%; Table 3]. However, this variant was rather unselective, since it converted this acyclic terpene [(E)-1] to a total of three products including 12-hydroxygeranylacetone [(*E*)-**5**, 31%] and 9,10-epoxygeranylacetone [(*E*)-**2**, 30%]. Variant F87V/A328L in turn almost exclusively converted geranylacetone to two allylic alcohols, with a strong preference for position C12 (80%) over C11 (15%). To our knowledge, this is the first time a CYP is presented that oxidised geranylacetone [(E)-**1**] to allylic alcohols in position 11 and 12. Except for variant A328V, which showed 1.7-fold increased oxidation rate as compared to wild type, the remaining variants with improved selectivity were less active than the wild type (Table 2).

Best nerylacetone-converting variants

CYP102A1 wild type oxidised nerylacetone [(*Z*)-**1**] unselectively to more than 4 different products (Table 3). Two products were identified as epoxides 9,10-epoxynerylacetone [(*Z*)-**2**, 47%] and 5,6-epoxynerylacetone [(*Z*)-**3**, 11%]. Two minor products were allylic alcohols 11-hydroxy- and 12-hydroxynerylacetone [(*Z*)-**4**, (*Z*)-**5**, 11% and 1%, respectively]. Variant A328V showed a strong shift in chemoselectivity and a considerable improve-

Table 3. Observed products of A) (4*R*)-(+)-limonene (**6**), B) geranylacetone [(*E*)-**1**], C) nerylacetone [(*Z*)-**1**] and D) (+)-valencene (**11**), converted by CYP102A1 wild type and the most selective variants. t_R indicates the GC retention time (in min) of unknown products.

CYP102A1 variants	Products in [%]											
A)	**7**	**8**	**9**	**10**	**t_R 73.5**	**t_R 74.8**						
wild type	30.1 ± 0.3	6.7 ± 1.2	53.7 ± 1	9.4 ± 0.1								
F87A/A328F		94 ± 0.2				3.5 ± 0.6						
F87V/A328F	1.2 ± .3	97 ± 0.2			1.9 ± 0.1							
B)	**(*E*)-3**	**(*E*)-2**	**t_R 12.3**	**t_R 12.8**	**t_R 13.9**	**(*E*)-5**	**(*E*)-4**					
wild type		54 ± 0.2	46 ± 0.2									
F87V	0.9 ± 0.1	87.4 ± 0.4	6.9 ± 0.3	4.1 ± 0.1			0.8 ± 0.1					
F87I		84.9 ± 0.6	7.4 ± 0.1	1.0 ± 0.1	2.1 ± 0.2	3.0 ± 0.2	1.6 ± 0.2					
F87L		97.2 ± 0.2	2.8 ± 0.2									
A328V		30.1 ± 0.4	2.0 ± 0.2			31 ± 0.8	36.9 ± 0.2					
F87V/A328L		5.3 ± 0.2				80 ± 0.3	14.7 ± 0.4					
C)	**(*Z*)-3**	**t_R 8.7**	**t_R 9.2**	**t_R 9.7**	**(*Z*)-2**	**t_R 10.5**	**t_R 11.8**	**(*Z*)-5**	**t_R 13.9**	**(*Z*)-4**		
wild type	11.4 ± 0.4	0.9 ± 0.1	1.1 ± 0.1	1.7 ± 0.1	46.9 ± 0.6	3.2 ± 0.6	22.4 ± 0.1	1.1 ± 0.1		11.3 ± 0.6		
A328V	0.7 ± 0.1							33.5 ± 0.2	2.1 ± 0.3	63.7 ± 0.3		
D)	**t_R 9.1**	**t_R 9.9**	**12**	**13**	**t_R 10.4**	**t_R 10.6**	**14**	**t_R 12.2**	**t_R 12.4**	**t_R 12.6**	**t_R 13.4**	**t_R 13.6**
wild type	3.8 ± 0.1	18.5 ± 0.9	3.0 ± 0.2	26.3 ± 1				0.9 ± 0.5	10.5 ± 0.6	37.1 ± 1.8		
F87A/A328I		0.9 ± 0.4	54.8 ± 1.7	12.6 ± 0.3		0.7 ± 0.2	26.1 ± 0.9	0.5 ± 0.2	1.9 ± 0.1	1.0 ± 0.2	0.7 ± 0.1	0.9 ± 0.3
F87V/A328I		6.2 ± 0.2	13.7 ± 1.5	61.4 ± 1.2			13.6 ± 1.1	0.8 ± 0.6	1.2 ± 0.8	2.3 ± 0.1	0.2 ± 0.2	0.6 ± 0.6
F87L/A328I		2.0 ± 0.3	62.4 ± 0.2	22.2 ± 0.3	2.8 ± 0.8		10.7 ± 0.5					
F87V/A328V		8.9 ± 0.6	2.9 ± 0.4	79.5 ± 1			3.9 ± 0.3		1.3 ± 0.0	3.5 ± 0.5		

ment of regioselectivity. In contrast to the wild type enzyme this variant produces almost exclusively two allylic alcohols, with a marked preference for 11-hydroxynerylacetone [(*Z*)-**4**, 64%] as compared to 12-hydroxynerylacetone [(*Z*)-**5**, 34%]. To our knowledge, for the first time a P450 catalyst is presented that oxidised nerylacetone [(*Z*)-**1**] preferentially at allylic position 11 to form 11-hydroxynerylacetone [(*Z*)-**4**]. Variant A328V revealed 2.2-fold increased nerylacetone oxidation rate as compared to wild type (Table 2).

Best (4*R*)-limonene-converting variants

The CYP102A1 wild type converts (4*R*)-limonene **6** to four different products (racemic mixtures of (4*R*)-limonene-1,2-epoxide (**7**, 30%), (4*R*)-limonene-8,9-epoxide (**8**, 7%), isopiperitenol (**10,** 54%) and carveol (**9,** 9%) (Table 3). The screening of our minimal library revealed two variants (F87A/A328F and F87V/A328F) with a strong increase in regioselectivity. In contrast to the wild type enzyme both variants almost exclusively epoxidise at the C8-C9 double bond resulting in 94% and 97% (4*R*)-limonene-8,9-epoxide (**8**), respectively (Table 3). While F87A/A328F showed a slightly decreased NADPH turnover rate as compared to wild type, NADPH turnover in F87V/A328F was increased (Table 2). In total, considerably more (4*R*)-limonene (**6**) was converted to products by each of the two variants than by the wild type enzyme (Table 2).

Best (+)-valencene-converting variants

(+)-Valencene (**11**) is unselectively oxidised by CYP102A1 wild type enzyme to more than four products (Table 3). Only 29% of the products result from an oxidation at the C2 atom to (+)-nootkatol (**13**). (+)-Nootkatone (**14**), a possible product of further oxidation at C2, was not detected. The four CYP102A1 variants F87V/A328I, F87A/A328I, F87L/A328I and F87V/A328V of our minimal library revealed 89%, 94%, 95% and 86% preference for oxidation at the C2 atom, respectively. Thereof, variants F87A/A328I, F87V/A328I and F87L/A328I showed the highest ratio of (+)-nootkatone (**14**) with 26%, 14% and 11%, respectively, while F87V/A328V produced only 4% of (+)-nootkatone (**14**). These product distributions were measured in standard reactions containing the same amount of substrate and NADPH. An increase of the NADPH concentration led to increasing amounts of (+)-nootkatone (**14**) and decreasing amounts of (+)-nootkatol (**13**) for the four variants (up to 64% for variant F87A/A328I). This indicates the need for two oxidation steps to arrive at (+)-nootkatone starting from (+)-valencene (**11**). The first step constitutes the oxidation of (+)-valencene (**11**) to (+)-

nootkatol (**13**), which is then further oxidised by the CYP to (+)-nootkatone (**14**). This effect was smallest for variant F87V/A328V. The fact that this effect is dependent on the CYP variant indicates that the second oxidation step is catalyzed by the CYP rather than other *E. coli* enzymes such as alcohol dehydrogenases. This variant combines high regioselectivity and stereoselectivity. F87V/A328V produces 82% (+)-nootkatol (**13**) with its *trans*-isomer at an enantiomeric excess of 93%, while wild type produces only 29% (+)-nootkatol (**13**) with its *trans*-isomer at an enantiomeric excess of 79%. Variants F87V/A328V and F87V/A328I showed higher NADPH turnover rates than the wild type enzyme, while variants F87A/A328I and F87L/A328I showed decreased NADPH turnover rates (Table 2). For all four variants (+)-valencene (**11**) conversion was equal or higher than for the wild type. Thereof, variants F87A/A328I and F87V/A328V showed the highest (+)-valencene conversion (Table 3).

Discussion

CYP102A1 from *Bacillus megaterium* is one of the most promising monooxygenases for the application in preparative synthesis. Hence, we chose this enzyme as a platform for the generation of a versatile toolbox of oxidation catalysts. Directed evolution by sequential rounds of random gene mutagenesis has been successfully applied to improve enzyme properties.[42] All directed evolution approaches rely on an effective screening of huge combinatorial libraries. However, only a very small part of the theoretical library size can effectively be screened. Information about protein structure allows one to focus on a reduced number of positions leading to decreased library size. An approach that includes saturation mutagenesis at selected positions was previously applied to engineer CYP102A1 for enantioselective alkene epoxidation.[12] However, even in the presence of crystal structure information the number of potential substrate-interacting residues is quite high. Therefore, an exhaustive analysis of possible cooperative effects between the different positions is rarely feasible due to the large number of combinatorial possibilities. A stepwise improvement of properties can be achieved by iterative cycles of combinatorial active-site saturation mutagenesis, starting from different positions.[43] This strategy leads to a higher probability to identify cooperative effects between different positions.

The major limitations of all the mentioned strategies are: 1) only those combinations of mutations are found of which at least one of the underlying single mutants leads to improved properties; 2) for each substrate a suitable assay has to be developed to screen large mutant libraries efficiently; 3) the mutant libraries have to be at least partially reconstructed for each screening process. Hence, instead of generating new libraries for every new substrate we decided to generate a single minimal, highly enriched CYP102A1 mutant library which can be rapidly screened with different substrates. We focussed on two hotspot positions (87 and 328) which are located in the immediate vicinity of the activated oxygen and therefore are most likely in contact with every substrate in its reactive orientation, independent of its size and shape. The generation of all possible combinations of five hydrophobic amino acids in both positions led to a great variety of substrate binding cavity shapes in the immediate vicinity of the haem (Figure 1). The screening of this minimal library with four differently sized and shaped terpenes identified variants with shifted or increased regio-, stereo- and chemoselectivity. It is widely accepted that increased regio- and stereoselectivity is the result of a reduced number of substrate orientations close to the haem.[10,44,45] Here we show that the systematic variation of the size and shape of hydrophobic residues in the immediate vicinity of the haem can turn a preferentially epoxidating into an hydroxylating enzyme and vice versa; this provides evidence that not only regio- and stereoselectivity but also chemoselectivity of CYP102A1 can be altered by mutations that induce changes in substrate orientations close to the haem centre. Some of the oxidation products described here are valuable and interesting compounds [for example, (+)-nootkatone (**14**)] and/or have been shown for the first time to be produced by CYPs (for example, the hydroxy products **4** and **5** of geranyl- and nerylacetone).

Previously, CYP102A1 wild type and mutants in five positions (including position 87) were screened with geranyl- and nerylacetone [(*E*)-**1**, (*Z*)-**1**].[19] Geranylacetone [(E)-**1**] was converted by a triple mutant to 9,10-epoxygeranylacetone [(*E*)-**2**] with high regio- and stereoselectivity. The variant contained a substitution of phenyalanine in position 87 to valine. Consistently, our results show that substitutions of F87 by valine (but also by leucine and isoleucine) strongly increased the selectivity for the formation 9,10-epoxygeranylacetone [(*E*)-**2**]. However, the formation of hydroxy products in position C11 and C12 of the substrate was not catalysed by those mutants. Here we show that amino acid substitutions in position 328 extended the product spectra towards hydroxy products, while combined mutations in 87 and 328 led to a variant that almost exclusively produces hydroxy products with a pronounced regioselectivity for the C12 atom. The screening of our minimal library with neryl acetone [(*Z*)-**1**] showed that position 328 had the strongest impact on regio- and chemoselectivity upon oxidation of this substrate. Only one mutant revealed a considerable increase in selectivity towards its *Z*-isomer neryl acetone [(*Z*)-**1**], in contrast to geranyl acetone [(E)-**1**], although both substrates have a similar chemical reactivity.

(+)-Nootkatone (**14**) is a valuable oxidation product of (+)-valencene (**11**). Previously, a CYP102A1 mutant (R47L/Y51F/F87A) has been introduced which, in contrast to mutant R47L/Y51F and wild type, produces (+)-nootkatol (**13**) and (+)-nootkatone (**14**) in small amounts, however, the mutant also produced epoxides and other products.[18] This is in agreement with our results, which show that mutant F87A produces (+)-nootkatol (**13**) and (+)-nootkatone (**14**) among three additional oxidation products (C2-selectivity of 55%). This result confirms the important role of position 87 for substrate orientation close to the haem centre and hence for regioselectivity. However, double mutant F87A/A328I shows highly increased C2-regioselectivity (95%), which indicates a much stronger restriction of possible (+)-valencene (**11**) binding orientations by a combination of this mutation with a mutation in posi-

tion 328. Also variants F87V/A328I, F87L/A328I and F87V/A328V showed strongly increased regioselectivity, hence different combinations of amino acids in these two positions can have similar effects. Interestingly, the single mutants F87L, A328V and A328I showed no conversion of (+)-valencene (**11**); this indicates that the effects of mutations in position 87 and 328 are nonadditive but cooperative. This fact makes it impossible to predict the high selectivity of the double mutants based on the properties of the single mutants. An iterative approach would have found variants F87A/A328I, F87V/A328I and F87V/A328V, because the single mutants F87A and F87V have slightly increased C2 regioselectivity as compared to wild type (data not shown). However, variant F87L/A328I could not be identified by an iterative approach, since both underlying single mutants showed no (+)-valencene (**11**) conversion. Furthermore, the preference for C12 hydroxylation of variant F87V/A328L could not be found by an iterative approach, since the respective single mutant F87V preferentially formed epoxides and single mutant A328L was inactive towards geranylacetone.

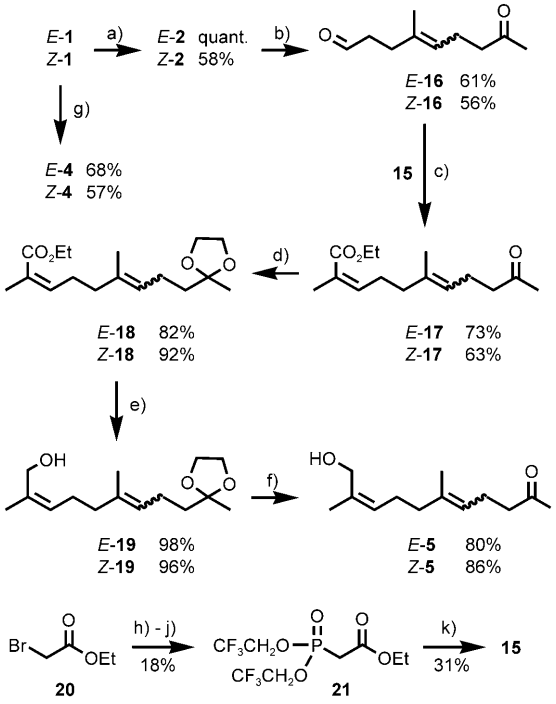
In general, more single mutants than double mutants showed improved regio- and chemoselectivity towards the small acyclic terpenes neryl- and geranylacetone [(*Z*)-**1**, (*E*)-**1**], while all variants with increased selectivity towards (4*R*)-limonene (**6**) and (+)-valencene (**11**) are double mutants. This demonstrates that combined mutations in both positions are required to improve selectivity towards the more bulky substrates (4*R*)-limonene (**6**) and (+)-valencene (**11**). It is instructive to analyse the three variants which were inactive towards all four terpene substrates. In case of variants F87/A328F and F87I/A328F, the amino acids side chains in the two positions are probably too bulky to permit haem access for the substrates used in this study. However, in variant F87A/A328V the haem is easily accessible for the terpene substrates used here, therefore it is surprising that this variant was inactive towards all four compounds. It would be interesting to verify whether more bulky substrates are converted by this variant.

Although it showed to be very effective to systematically combine only five hydrophobic amino acids in position 87 and 328, it might be promising to extent randomisation with serine, threonine and glycine since these residues are also observed in CYP sequences at a position corresponding to position 328 in CYP102A1.[11] The compounds used to screen our minimal library have different shape, size and polarity. Both hotspot positions also showed to influence selectivity of non-terpene substrates, hence we anticipate that this library is also a rich source for biocatalysts with improved selectivity towards non-terpene substrates. An ideal CYP catalyst would combine both a broad substrate specificity and high selectivity. Our results indicate that these requirements can not be fulfilled by one variant only, but rather by a collection of variants. Our minimal library is a first step towards such an ideal CYP catalyst, since it provides a small collection of active variants with improved selectivity for different substrates. We expect that our approach is generic and can be applied to other CYPs. Position 87 can be identified in all CYPs with crystal structure information as well as in homologous CYPs. Position 328 can be identified from sequence alone in almost all CYPs, owing to its conserved distance from the highly conserved ExxR motif.[11]

Experimental Section

Materials: Tryptone/peptone from caseine and yeast extract were purchased from Roth (Karlsruhe, Germany). NADPH tetrasodium salt was from Codexis (Jülich, Germany). 5,6-Epoxy- and 9,10-epoxygeranylacetone [(*E*)-**3**, (*E*)-**2**], 5,6-epoxy- and 9,10-epoxynerylacetone [(*Z*)-**3**, (*Z*)-**2**] were prepared as described previously,[19] (4*R*)-limonene-8,9-epoxide (**8**) was prepared as described elsewhere.[31] All other chemicals (except 11-hydroxy- and 12-hydroxy-geranylacetone [(*E*)-**4**, (*E*)-**5**], 11-hydroxy- and 12-hydroxynerylacetone [(*Z*)-**4**, (*Z*)-**5**)] used in this work were purchased from Fluka (Buchs, Switzerland) or Sigma (Deisenhofen, Germany) and were of analytical grade or higher.

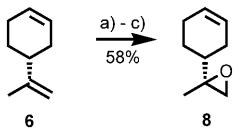
Synthesis of authentic samples: 11-Hydroxygeranylacetone [(*E*)-4] and 11-hydroxynerylacetone [(*Z*)-4] were prepared from geranylacetone [(E)-1] and nerylacetone (*Z*)-**1** respectively through selenium dioxide oxidation according to the method by McMurry,[25,28] which gave the desired products (*E*)-**4**, (*Z*)-**4** in 68% and 57% yield respectively (Scheme 1). The synthesis of 12-hydroxygeranylacetone [(*E*)-**5**] commenced with treatment of geranylacetone [(E)-1] with MCPBA in $CHCl_3$ at 0 °C to give the corresponding 9,10-epoxygeranylacetone [(*E*)-**2**] quantitatively.[15] Epoxide (*E*)-**2** was submitted to periodate cleavage under acidic conditions following the method by Mori and Kogen[32,33] and the aldehyde (*E*)-**16** was isolated in 61% yield. Horner–Emmons olefination of compound (*E*)-**16** using the Still-Gennari modification[34,35] with the phosphonate **15** in the presence of KHMDS and [18]crown-6 in THF at −78 °C



Scheme 1. a) MCPBA, $CHCl_3$, 0 °C; b) HIO_4, Et_2O, THF, 0 °C; c) **15**, KHMDS, [18]crown-6, THF, −78 °C, 1.5 h; d) $(TMSOCH_2)_2$, TMSOTf, CH_2Cl_2, −78 °C, 1 h; e) DIBAL, CH_2Cl_2, −78 °C, 1 h; f) acetone, H_2O, Amberlyst 15, RT, 20 h; g) SeO_2, *t*BuOOH, CH_2Cl_2, 0 °C, 5 h; h) $P(OEt)_3$, 150 °C, 5 h; i) PCl_5, 80 °C, 16 h; j) CF_3CH_2OH, toluene, *i*Pr_2NEt, 0 °C, 1 h; k) NaH, THF, 0 °C, then MeI, −10 °C, 1 h.

gave the α,β-unsaturated ester (*E*)-**17** in 73% yield. The Still-Gennari-phosphonate **15** was synthesized in four steps in an overall yield of 6%.[34,35] Sequential reaction of ethyl bromoacetate **20** with triethylphosphite at 155 °C for 5 h followed by treatment of the phosphonate with PCl_3 and subsequent addition of 2,2,2-trifluoroethanol yielded the unbranched Still–Gennari phosphonate **22** in 33%. Methylation of **22** with methyl iodide and sodium hydide in THF gave compound **15** in 31% yield. Protection of the carbonyl group in (*E*)-**17** was achieved by transacetalisation employing the method by Noyori and Hashimoto[36] with bis-1,2-(trimethylsilyloxy)ethane in the presence of TMSOTf in CH_2Cl_2 to yield the acetal (*E*)-**18** in 82%. After reduction of (*E*)-**18** with DIBAL in CH_2Cl_2 at −78 °C the allylic alcohol (*E*)-**19** was obtained in 98% yield. Subsequent deprotection with Amberlyst 15 ion exchange resin in water and acetone at room temperature gave the desired 12-hydroxygeranylacetone [(*E*)-**5**] in 80% yield. The synthesis of 12-hydroxynerylacetone [(*Z*)-**5**] proceeded in a similar fashion, giving the desired product (*Z*)-**5** in six steps and 16% overall yield from nerylacetone (*Z*)-**1**.

Following a method by Jones,[31] (4*R*)-limonene (**6**) was sequentially treated with NBS in MeOH at room temperature, followed by oxidation with MCPBA in CH_2Cl_2 and final treatment with zinc and NH_4Cl in ethanol to give (4*R*)-limonene-8,9-epoxide (**8**) in 58% overall yield (Scheme 2).



Scheme 2. a) NBS, MeOH, RT, 1 d; b) MCPBA, CH_2Cl_2, 0 °C to RT, 3 d; c) Zn, NH_4Cl, EtOH, RT, 1 d.

(*E,E*)-6,10-dimethyl-11-hydroxy-5,9-undecadien-2-one (11-hydroxygeranylacetone) (*E*-4): Geranylacetone [*E*-**1**] (200 mg, 1.03 mmol) was added to a solution of selenium dioxide (46 mg, 0.41 mmol) and *tert*-butyl hydroperoxide (0.64 mL, 5 M solution in nonane, 3.2 mmol) in CH_2Cl_2 at 0 °C. After stirring under nitrogen at 0 °C for 5 h, the mixture was diluted with ethyl acetate (15 mL) and washed successively with water (2×10 mL), saturated $NaHCO_3$ (10 mL), water (10 mL) and brine (10 mL). The solution was then dried and evaporated under reduced pressure. Chromatographic purification on silica gel (hexanes/ethyl acetate 7:3) afforded alcohol *E*-**4** (147 mg, 68%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.60$ (s, 3H), 1.62 (s, 3H), 2.01–2.05 (m, 2H), 2.12 (s, 3H), 2.23–2.35 (m, 2H), 2.44–2.50 (m, 1H), 3.99 (br, 2H), 5.06 (qt, $J = 1.3$, 7.2 Hz, 1H), 5.34 (qt, $J = 1.4$, 6.9 Hz, 1H); ^{13}C NMR (125 MHz, $CDCl_3$), $\delta = 13.7$ (CH_3), 15.9 (CH_3), 22.5 (CH_2), 25.9 (CH_2), 29.8 (1-C), 39.2 (CH_2), 43.7 (3-C), 68.8 (4−C), 122.9 (CH), 125.5 (CH), 134.9 (C_q), 135.9 (C_q), 209.3 ppm (2-C).

(*E,Z*)-6,10-Dimethyl-11-hydroxy-5,9-undecadien-2-one (11-hydroxynerylacetone) (*Z*-4): The alcohol *Z*-**4** was prepared as described for *E*-**4** from *Z*-**1** (200 mg, 1.03 mmol). Chromatographic purification on silica gel (hexanes/ethyl acetate 7:3) gave alcohol *E*-**4** (123 mg, 57%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.58$ (s, 3H), 1.64 (s, 3H), 2.10–2.12 (m, 2H), 2.14 (s, 3H), 2.21–2.33 (m, 2H), 2.42–2.53 (m, 1H), 4.00 (br, 2H), 5.04–5.11 (m, 1H) 5.39–5.44 (m, 1H); ^{13}C NMR (125 MHz, $CDCl_3$), $\delta = 13.6$ (CH_3), 22.4 (CH_2), 23.2 (CH_3), 25.7 (CH_2), 30.0 (1-C), 31.5 (CH_2), 43.9 (3-C), 68.6 (4-C), 123.7 (CH), 125.2 (CH), 135.1 (C_q), 136.1 (C_q), 209.2 ppm (2-C).

(*E*)-4-Methyl-8-oxonon-4-enal (*E*-16): A solution of epoxide *E*-**2** (750 mg, 3.57 mmol) in dry Et_2O (5 mL) was added to a solution of periodic acid dihydrate (976 mg, 4.28 mmol) in dry THF (20 mL) at 0 °C under N_2. After stirring for 0.5 h, saturated aqueous $NaHCO_3$ was added at 0 °C and the reaction mixture was stirred for another 15 min. The white suspension was filtered through a plug of silica and rinsed with Et_2O. The aqueous material was extracted with Et_2O. The combined organic layers were washed with H_2O (50 mL), saturated aqueous $NaHCO_3$ (50 mL) and brine (50 mL), dried and evaporated under reduced pressure. The crude product was purified by chromatography on silica gel (hexanes/ethyl acetate 5:1) to give the aldehyde *E*-**16** as a colorless oil (345 mg, 61%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.63$ (s, 3H), 2.13 (s, 3H), 2.22–2.35 (m, 4H), 2.42–2.55 (m, 4H), 5.11 (qt, $J = 1.3$, 7.1 Hz, 1H), 9.74 ppm (t, $J = 1.9$, 1H).

(*Z*)-4-Methyl-8-oxonon-4-enal (*Z*-16): The aldehyde *Z*-**2** was prepared as described for *E*-**16** from *Z*-**2** (1.70 g, 8.10 mmol). Chromatographic purification on silica gel (hexanes/ethyl acetate 6:1) gave the aldehyde ***Z*-16** (756 mg, 56%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta =$ 1.63 (s, 3H), 2.13 (s, 3H), 2.21–2.35 (m, 4H), 2.42–2.55 (m, 4H), 5.11 (qt, $J = 1.3$, 7.0 Hz, 1H), 9.75 ppm (t, $J = 1.9$, 1H).

(*2Z,6E*)-Ethyl-2,6-dimethyl-10-oxoundeca-2,6-dienoate (*E*-17): A solution of **15** (671 mg, 2.02 mmol) in dry THF (3 mL) under N_2 atmosphere was added to a solution of [18]crown-6 (1.6 g, 6.06 mmol) in dry THF (15 mL). The solution was cooled to −78 °C and KHMDS (4.5 mL, 0.5 M in toluene, 2.22 mmol) was added. After stirring for 0.5 h, a solution of the aldehyde *E*-**16** (340 mg, 2.02 mmol) in dry THF (7 mL) was added dropwise. The reaction mixture was maintained at −78 °C for 1 h, then quenched with aqueous NH_4Cl (60 mL). The mixture was extracted with Et_2O (8× 30 mL). The combined organic phases were washed with brine (20 mL), dried over $MgSO_4$ and concentrated in vacuo to give a pale yellow oil. Chromatographic purification on silica gel (hexanes/ethyl acetate 9:1) gave the unsaturated ethyl ester *E*-**17** as a colourless oil (338 mg, 66%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.30$ (t, $J = 7.1$ Hz, 3H), 1.61 (s, 3H), 1.88 (s, 3H), 2.03- 2.10 (m, 2H), 2.13 (s, 3H), 2.22–2.31 (m, 2H), 2.42–2.49 (m, 2H), 2.50–2.60 (m, 2H), 4.20 (q, $J = 7.1$ Hz, 2H), 5.06 (qt, $J = 1.3$, 7.1 Hz, 1H), 5.88 ppm (qt, $J =$ 1.5, 7.3 Hz, 1H).

(*2Z,6Z*)-Ethyl-2,6-dimethyl-10-oxoundeca-2,6-dienoate (*Z*-17): The ethyl ester ***Z*-17** was prepared as described for *E*-**17** from *Z*-**16** (1.01 g, 3.04 mmol). Chromatographic purification on silica gel (hexanes/ethyl acetate 5:1) gave the unsaturated ethyl ester *Z*-**17** (484 mg, 63%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.30$ (t, $J = 7.1$ Hz, 3H), 1.68 (s, 3H), 1.89 (s, 3H), 2.10–2.17 (m, 2H), 2.14 (s, 3H), 2.22–2.31 (m, 2H), 2.42–2.48 (m, 2H), 2.49–2.59 (m, 2H), 4.20 (q, $J =$ 7.2 Hz, 2H), 5.08–5.13 (m, 1H), 5.90 ppm (qt, $J = 1.5$, 7.5 Hz, 1H).

(*2Z,6E*)-Ethyl-2,6-dimethyl-9-(2-methyl-1,3-dioxolan-2-yl)nona-2,6-dienoate (*E*-18): A solution of bis-1,2-(trimethylsilyloxy)ethane (473 mg, 2.3 mmol) in dry CH_2Cl_2 (5 mL) was added to a solution of *E*-**17** (445 mg, 1.8 mmol) in dry CH_2Cl_2 (10 mL) under N_2 atmosphere. The solution was cooled to −78 °C and TMSOTf (6.4 μL, 35 μmol) was added dropwise. After 1 h, TLC control showed incomplete conversion of *E*-**17**. An additional portion of bis-1,2-(trimethylsilyloxy)ethane (40 mg, 0.19 mmol) was added to the reaction mixture. After stirring for 10 min, dry pyridine (4.5 mL) was added and stirred for another 3 min. The solvents were evaporated in vacuo to give a yellow oil. Chromatographic purification on silica gel (hexanes/ethyl acetate 15:1) yielded the protected unsaturated ethyl ester *E*-**18** as a colourless oil (427 mg, 82%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.30$ (t, $J = 7.2$ Hz, 3H), 1.33 (s, 3H), 1.61 (s, 3H), 1.63–1.70 (m, 2H), 1.87–1.90 (m, 3H), 2.03–2.14 (m, 4H), 2.51–2.60 (m, 2H), 3.91–3.97 (m, 4H), 4.20 (q, $J = 7.1$ Hz, 2H), 5.15 (qt, $J = 1.3$, 7.1 Hz, 1H), 5.90 ppm (qt, 1H, $J = 1.5$, 7.3 Hz).

(*2Z,6Z*)-Ethyl-2,6-dimethyl-9-(2-methyl-1,3-dioxolan-2-yl)nona-2,6-dienoate (*Z*-18): ***Z*-18** was prepared as described for *E*-**18** from *Z*-**17** (438 mg, 1.90 mmol). Chromatographic purification on silica

gel (hexanes/ethyl acetate 5:1) gave the unsaturated ethyl ester *Z*-**18** (520 mg, 92%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.30$ (t, $J = 7.2$ Hz, 3H), 1.32 (s, 3H), 1.54 (s, 3H), 1.67–1.70 (m, 2H), 1.87–1.90 (m, 3H), 2.03–2.17 (m, 4H), 2.50–2.59 (m, 2H), 3.91–3.97 (m, 4H), 4.22 (q, $J = 7.1$ Hz, 2H), 5.12–5.19 (m, 1H), 5.87–5.94 ppm (m, 1H).

(*2Z,6E*)-ethyl-2,6-dimethyl-9-(2-methyl-1,3-dioxolan-2-yl)nona-2,6-dien-1-ol (*E*-19): A solution of *E*-**18** (338 mg, 1.14 mmol) in dry CH_2Cl_2 (10 mL) was cooled to $-78\,^\circ$C in a N_2 atmosphere. DIBAL (2.5 mL, 1 M in hexane, 2.5 mmol) was added to this solution. After stirring for 1 h at $-78\,^\circ$C, a saturated aqueous potassium sodium tartrate solution (25 mL) was added. The reaction mixture was allowed to warm to room temperature and was stirred subsequently over night. The aqueous phase was extracted with Et_2O (5×20 mL). The combined organic layers were dried and evaporated under reduced pressure. The crude product was purified by chromatography on silica gel (hexanes/ethyl acetate 3:1) to give the alcohol *E*-**19** as a colorless oil (284 mg, 98%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta =$ 1.33 (s, 3H), 1.59–1.63 (m, 3H), 1.63–1.70 (m, 2H), 1.77–1.82 (m, 3H), 1.96–2.06 (m, 2H), 2.06–2.20 (m, 4H), 3.90–3.97 (m, 4H), 4.11 (dd, $J = 0.7$, 5.8 Hz, 2H), 5.13 (qt, $J = 1.3$, 7.0 Hz, 1H), 5.24–5.32 ppm (m, 1H).

(*2Z,6Z*)-Ethyl-2,6-dimethyl-9-(2-methyl-1,3-dioxolan-2-yl)nona-2,6-dien-1-ol (*Z*-19): *Z*-**19** was prepared as described for *E*-**19** from *Z*-**18** (338 mg, 1.14 mmol). Chromatographic purification on silica gel (hexanes/ethyl acetate 5:1) gave the unsaturated ethyl ester *Z*-**19** (278 mg, 96%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.32$ (s, 3H), 1.60–1.67 (m, 2H), 1.67–1.70 (m, 3H), 1.78–1.82 (m, 3H), 1.99–2.19 (m, 6H), 3.90–3.98 (m, 4H), 4.10 (dd, $J = 0.7$, 5.9 Hz, 2H), 5.13–5.21 (m, 1H), 5.27–5.35 ppm (m, 1H).

(*E,E*)-6,10-Dimethyl-12-hydroxy-5,9-undecadien-2-one (12-hydroxygeranylacetone) (*E*-5): Amberlyst 15 (200 mg) was added to a solution of *E*-**19** (280 mg, 1.1 mmol) in acetone/4% H_2O (25 mL) and the mixture was stirred for 20 h at room temperature. The ion exchange resin was filtered off, the filtrate was washed with acetone and successively concentrated in vacuo. The residue was dried azeotropically with toluene. The crude product was purified by chromatography on silica gel (hexanes/ethyl acetate 3:1) to give the unprotected alcohol *E*-**5**, a light yellow oil (184 mg, 80%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.42$ (s, 1H), 1.60–1.63 (m, 3H), 1.78–1.82 (m, 3H), 2.07–2.16 (m, 7H), 2.22–2.29 (m, 2H), 2.42–2.48 (m, 2H), 4.10 (d, $J = 0.7$ Hz, 2H), 5.07–5.14 (m, 1H), 5.26–5.33 (m, 1H); ^{13}C NMR (125 MHz, $CDCl_3$), $\delta = 16.0$ (CH_3), 21.3 (CH_3), 22.4 (CH_2), 26.0 (CH_2), 30.0 (1-C), 39.7 (CH_2), 43.7 (3-C), 61.5 (4-C), 123.1 (CH), 127.9 (CH), 134.6 (C_q), 135.9 (C_q), 209.4 (2-C) ppm; HRMS (ESI): *m/z* calcd for $C_{13}H_{22}O_2Na$: 233.1512; found: 233.1510 $[M+Na]^+$.

(*E,Z*)-6,10-Dimethyl-12-hydroxy-5,9-undecadien-2-one (12-hydroxynerylacetone) (*Z*-5): *Z*-**5** was prepared as described for *E*-**5** from *Z*-**19** (274 mg, 1.08 mmol). Chromatographic purification on silica gel (hexanes/ethyl acetate 3:1) gave the unprotected alcohol *Z*-**5** (195 mg, 86%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.43$ (s, 1H), 1.67–1.69 (m, 3H), 1.78–1.82 (m, 3H), 2.07–2.16 (m, 7H), 2.22–2.29 (m, 2H), 2.42–2.48 (m, 2H), 4.10 (d, $J = 0.7$ Hz, 2H), 5.07–5.14 (m, 1H), 5.26–5.33 (m, 1H); ^{13}C NMR (125 MHz, $CDCl_3$), $\delta = 21.4$ (CH_3), 22.2 (CH_3), 23.4 (CH_2), 30.1 (CH_2), 32.0 (1-C), 39.7 (CH_2), 43.8 (3-C), 61.6 (4-C), 123.9 (CH), 127.9 (CH), 134.9 (C_q), 136.1(C_q), 208.3 ppm (2-C), HRMS (ESI): *m/z* calcd for $C_{13}H_{22}O_2Na$: 233.1512; found: 233.1501 $[M+Na]^+$.

Ethyl-2-(bis(2,2,2-trifluoroethoxy)phosphoryl)acetate (21):[34,35] was prepared analogously to Still and Gennari[35] from ethyl-2-(diethoxyphosphoryl)acetate (13.23 g, 64.5 mmol, 1 equiv) in toluene (80 mL), CF_3CH_2OH (13.55 g, 135.5 mmol, 2.1 equiv) and iPr_2NEt in toluene (20 mL). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.30$ (t, $J = 7.2$ Hz, 3H), 3.16 (d, $J = 21.1$ Hz, 2H), 4.23 (q, $J = 7.2$ Hz, 2H), 4.52–4.40 ppm (m, 4H).

Ethyl-2-(bis(2,2,2-trifluoroethoxy)phosphoryl)propanoate (15): A suspension of NaH (840 mg, 60%, 21 mmol) in dry THF (15 mL) was cooled to 0 °C under N_2. A solution of **22** (6.8 g, 20.5 mmol) in dry THF (10 mL) was added to this suspension. The reaction mixture was allowed to warm to room temperature and was stirred for 3 h. The orange solution was cooled to $-10\,^\circ$C and a solution of MeI (1.4 mL, 21.7 mmol) was added. The reaction mixture was allowed to warm to 0 °C temperature, H_2O (30 mL) was added and the solution was stirred for 1 h. The reaction mixture was diluted with CH_2Cl_2 (100 mL) and brine (100 mL) was added. The aqueous phase was extracted with CH_2Cl_2 (2×50 mL). The combined organic layers were dried and evaporated under reduced pressure. The crude product was purified by chromatography on silica gel (hexanes/ethyl acetate 6:1) to give compound **15** as a pale orange oil (2.21 g, 31%). ^{1}H NMR (300 MHz, $CDCl_3$), $\delta = 1.30$ (t, $J = 7.1$ Hz, 3H), 1.52 (qd, $J = 7.1$, 19.3 Hz, 3H), 3.19 (qd, $J = 7.4$, 22.7 Hz, 1H), 4.24 (q, $J = 7.2$ Hz, 2H), 4.51–4.36 ppm (m, 4H).

Mutant expression: The 24 CYP102A1 single and double mutants plus wild type were heterologously expressed in *E. coli* as reported previously.[37] The pET22b and pET28+ derivatives were used. Following the manufacturer's protocol, the QuikChange™ site-directed mutagenesis kit from Stratagene (La Jolla, California, USA) was used to introduce site directed mutations.

Enzyme activity measurements: CO-difference spectra measurements were used to determine CYP concentrations as described elsewhere.[38] An extinction coefficient of 91 $mM^{-1}cm^{-1}$ was used for calculations. A NADPH oxidation assay was applied to determine the initial activity of the mutants towards nerylacetone [(*Z*)-**1**], geranylacetone [(E)-1], (4*R*)-limonene (**6**) and (+)-valencene (**11**). One mL final reaction mixture contained potassium phosphate buffer (50 mM, pH 7.5), DMSO (2%, *v/v*), substrate (0.2 mM) as well as CYP enzyme (0.5 μM). Reaction was started by adding NADPH solution (2 mM, 100 μL). The initial activity was determined from the absorption decrease at 340 nm, whereas the slope of the first 30 s was measured and an extinction coefficient of 6.22 $mM^{-1}cm^{-1}$ was applied.

GC-MS analyses: The aqueous reaction mixture was extracted twice with diethyl ether (or ethyl acetate in case of (+)-valencene) and the organic phase was dried over magnesium sulfate. Analysis of reaction products and conversion were performed on a Shimadzu QP2010 GC/MS with EI-ionisation, the GC was equipped with a FS-supreme-5 capillary column (length: 30 m, internal diameter: 0.25 mm, film thickness: 0.25 μm).

*Analysis of (4*R*)-limonene oxidation products*: The GC was programmed as follows: 40 °C, 1 min. iso; 1 °C min^{-1}. to 70 °C, 20 min. iso; 1 °C min^{-1}. to 100 °C; 10 °C min^{-1}. to 110 °C; 30 °C min^{-1}. to 250 °C; injector temperature 250 °C. The oxidation products carveol (**9**) and (4*R*)-limonene-8,9-epoxide (**8**) were identified by authentic samples. (4*R*)-limonene-1,2-epoxide (**7**) was identified by comparison of its characteristic mass fragmentation pattern with the NIST mass spectrometry data base.[39] Isopiperitenol was identified by comparison of the mass spectra to literature data.[40]

Analysis of neryl- and geranylacetone oxidation products: The GC was programmed as follows: 120 °C, 30 sec. iso; 5 °C min^{-1}. to 165 °C, 2 °C min^{-1}. to 185 °C, 30 °C min^{-1}. to 280 °C, 1 min. iso; injector temperature 250 °C. The oxidation products 11-hydroxy- and 12-hydroxygeranylacetone [(*E*)-**4**, (*E*)-**5**], 5,6-epoxy- and 9,10-epoxy-

geranylacetone [(*E*)-**3**, (*E*)-**2**], 11-hydroxy- and 12-hydroxynerylacetone [(*Z*)-**4**, (*Z*)-**5**], 5,6-epoxy- and 9,10-epoxy-nerylacetone [(*Z*)-**3**, (*Z*)-**2**] were identified using authentic samples.

Analysis of (+)-valencene oxidation products: The GC was programmed as follows: 150 °C, 4 min. iso; 10 °C min^{-1}. to 250 °C, 5 min iso; 50 °C min^{-1}. to 300 °C, 1 min. iso; injector temperature 250 °C. (+)-Nootkatone (**14**) was identified using an authentic sample and *cis*- and *trans*-nootkatol (**12**, **13**) were identified by comparison of the mass spectra to literature data.[41] To calculate the amount of converted substrate GC peak area ratios between substrate and an internal standard (geraniol for geranyl- and nerylacetone conversions; citronellal for (4*R*)-limonene conversions; (−)-carvone for (+)-valencene conversions) were measured, before and after reactions. With an enzyme concentration of c μM, *a* μM of converted substrate and a time of *t* minutes to consume the NADPH, the substrate oxidation rate is defined as $a/(c \cdot t)$ [μmol · (μmol CYP)$^{-1}$ · min^{-1}]. Regioselectivity was determined from the gas chromatograms by integrating the product peaks.

Abbreviations: CYP, cytochrome P450 monooxygenase

Acknowledgement

Financial support by the Deutsche Forschungsgemeinschaft (SFB 706), the Fonds der Chemischen Industrie and the Ministerium für Wissenschaft, Forschung und Kunst des Landes Baden-Württemberg (Landesgraduierten fellowship for K.G.) are gratefully acknowledged. We would like to thank Dr. Alexandrine Busch for initial syntheses of reference compounds.

Keywords: cytochrome P450 monooxygenase · enzymes · heme proteins · mutagenesis · terpenes

[1] L. O. Narhi, A. J. Fulco, *J. Biol. Chem.* **1987**, *262*, 6683–6690.
[2] K. Kühnel, S. C. Maurer, Y. Galeyeva, W. Frey, S. Laschat, V. B. Urlacher, *Adv. Synth. Catal.* **2007**, *349*, 1451–1461.
[3] 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.
[4] V. B. Urlacher, S. Lutz-Wahl, R. D. Schmid, *Appl. Microbiol. Biotechnol.* **2004**, *64*, 317–325.
[5] O. Gotoh, *J. Biol. Chem.* **1992**, *267*, 83–90.
[6] A. B. Carmichael, L. L. Wong, *Eur. J. Biochem.* **2001**, *268*, 3117–3125.
[7] S. Graham-Lorence, G. Truan, J. A. Peterson, J. R. Falck, S. Wei, C. Helvig, J. H. Capdevila, *J. Biol. Chem.* **1997**, *272*, 1127–1135.
[8] H. M. Li, L. H. Mei, V. B. Urlacher, R. D. Schmid, *Appl. Biochem. Biotechnol.* **2008**, *144*, 27–36.
[9] V. B. Urlacher, A. Makhsumkhanov, R. D. Schmid, *Appl. Microbiol. Biotechnol.* **2006**, *70*, 53–59.
[10] R. J. Branco, A. Seifert, M. Budde, V. B. Urlacher, M. J. Ramos, J. Pleiss, *Proteins Struct. Funct. Bioinf.* **2008**, *73*, 597–607.
[11] A. Seifert, J. Pleiss, *Proteins* **2008**, (DOI: 10.1002/prot.22242).
[12] T. Kubo, M. W. Peters, P. Meinhold, F. H. Arnold, *Chem. Eur. J.* **2006**, *12*, 1216–1220.
[13] P. Meinhold, M. W. Peters, A. Hartwick, A. R. Hernandez, F. H. Arnold, *Adv. Synth. Catal.* **2006**, *348*, 763–772.
[14] A. Seifert, S. Tatzel, R. D. Schmid, J. Pleiss, *Proteins Struct. Funct. Bioinf.* **2006**, *64*, 147–155.
[15] M. Dietrich, S. Eiben, C. Asta, T. A. Do, J. Pleiss, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2008**, *79*, 931–940.
[16] M. J. van der Werf, P. M. Keijzer, P. H. van der Schaft, *J. Biotechnol.* **2000**, *84*, 133–143.
[17] M. Furusawa, T. Hashimoto, Y. Noma, Y. Asakawa, *Chem. Pharm. Bull.* **2005**, *53*, 1513–1514.
[18] R. J. Sowden, S. Yasmin, N. H. Rees, S. G. Bell, L. L. Wong, *Org. Biomol. Chem.* **2005**, *3*, 57–64.
[19] Y. Watanabe, S. Laschat, M. Budde, O. Affolter, Y. Shimada, V. B. Urlacher, *Tetrahedron* **2007**, *63*, 9413–9422.
[20] M. Bruder, C. J. Moody, *Synlett* **2008**, 575–577.
[21] J. S. Clark, J. Myatt, L. Roberts, N. Walshe, *Synlett* **2005**, 697–699.
[22] J. A. Marshall, W. J. Dubay, *J. Org. Chem.* **1994**, *59*, 1703–1708.
[23] Y. Li, W. D. Li, Y. L. Li, *Synth. Commun.* **1994**, *24*, 117–121.
[24] J. E. Mcmurry, R. G. Dushin, *J. Am. Chem. Soc.* **1989**, *111*, 8928–8929.
[25] J. E. Mcmurry, R. G. Dushin, *J. Am. Chem. Soc.* **1990**, *112*, 6942–6949.
[26] J. E. Mcmurry, P. Kocovsky, *Tetrahedron Lett.* **1984**, *25*, 4187–4190.
[27] J. E. Mcmurry, P. Kocovsky, *Tetrahedron Lett.* **1985**, *26*, 2171–2172.
[28] J. E. Mcmurry, J. R. Matz, K. L. Kees, *Tetrahedron* **1987**, *43*, 5489–5498.
[29] A. B. Charette, H. Juteau, H. Lebel, D. Deschenes, *Tetrahedron Lett.* **1996**, *37*, 7925–7928.
[30] A. B. Charette, H. Juteau, H. Lebel, C. Molinaro, *J. Am. Chem. Soc.* **1998**, *120*, 11943–11952.
[31] Q. A. R. Almeida, J. Jones, *Synth. Commun.* **2005**, *35*, 1285–1290.
[32] S. Muto, Y. Nishimura, K. Mori, *Eur. J. Org. Chem.* **1999**, 2159–2165.
[33] K. Tago, M. Arai, H. Kogen, *J. Chem. Soc. Perkin Trans. 1* **2000**, 2073–2078.
[34] J. P. Malerich, D. Trauner, *J. Am. Chem. Soc.* **2003**, *125*, 9554–9555.
[35] W. C. Still, C. Gennari, *Tetrahedron Lett.* **1983**, *24*, 4405–4408.
[36] T. Tsunoda, M. Suzuki, R. Noyori, *Tetrahedron Lett.* **1980**, *21*, 1357–1358.
[37] S. C. Maurer, H. Schulze, R. D. Schmid, V. Urlacher, *Adv. Synth. Catal.* **2003**, *345*, 802–810.
[38] T. Omura, R. J. Sato, *J. Biol. Chem.* **1964**, *239*, 2370–2378.
[39] *NIST 05 Mass Spectral Library*, National Institute of Standards and Technology, Gaithersburg, Maryland, USA.
[40] J. Lücker, W. Schwab, M. C. R. Franssen, L. H. W. Van Der Plas, H. J. Bouwmeester, H. A. Verhoeven, *Plant J.* **2004**, *39*, 135–145.
[41] S. Takahashi, Y. S. Yeo, Y. X. Zhao, P. E. O'Maille, B. T. Greenhagen, J. P. Noel, R. M. Coates, J. Chappell, *J. Biol. Chem.* **2007**, *282*, 31744–31754.
[42] O. Kuchner, F. H. Arnold, *Trends Biotechnol.* **1997**, *15*, 523–530.
[43] M. T. Reetz, L. W. Wang, M. Bocola, *Angew. Chem.* **2006**, *118*, 1258–1263; *Angew. Chem. Int. Ed.* **2006**, *45*, 1236–1241.
[44] S. G. Bell, X. H. Chen, R. J. Sowden, F. Xu, J. N. Williams, L. L. Wong, Z. H. Rao, *J. Am. Chem. Soc.* **2003**, *125*, 705–714.
[45] R. Raag, T. L. Poulos, *Biochemistry* **1991**, *30*, 2674–2684.
[46] W. L. DeLano in *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA, USA, **2002**.

Received: December 2, 2008
Published online on February 16, 2009

