J Biol Inorg Chem (2010) 15:159–174
DOI 10.1007/s00775-009-0580-y

ORIGINAL PAPER

Scanning chimeragenesis: the approach used to change the substrate selectivity of fatty acid monooxygenase CYP102A1 to that of terpene ω-hydroxylase CYP4C7

Chiung-Kuang J. Chen · Robert E. Berry · Tatjana Kh. Shokhireva · Marat B. Murataliev · Hongjung Zhang · F. Ann Walker

Received: 18 March 2009 / Accepted: 13 August 2009 / Published online: 30 August 2009

Abstract CYP102A1 is a highly active, water-soluble, bacterial monooxygenase enzyme that contains both substrate-binding heme and diflavin reductase subunits, both in a single polypeptide. Recently we developed a procedure which uses the known structure of the substrate-bound heme domain of CYP102A1 and its sequence homology with a cytochrome P450 of unknown structure, both of which react with a common substrate but produce different products, to create recombinant enzymes which have substrate selectivity different from that of CYP102A1, and produce the product of the enzyme of unknown structure. Insect CYP4C7, a terpene hydroxylase from the cockroach, was chosen as the cytochrome P450 of unknown structure, and farnesol was chosen as the substrate. CYP102A1 oxidizes farnesol to three products (2,3-epoxyfarnesol, 10,11-epoxyfarnesol, and 9-hydroxyfarnesol), whereas CYP4C7 produces 12-hydroxyfarnesol as the major product. In earlier work it was found that the chimera C(78-82,F87L) showed a change in substrate selectivity from fatty acids to farnesol, and was approximately sixfold more active than wild-type CYP102A1 (Chen et al. in J Biol Inorg Chem 13:813–824, 2008), but neither it nor any other earlier chimera produced 12-hydroxyfarnesol. In this work we added amino acid residues 327–332, to create six new full-length, functional chimeric proteins. Four of these, the most active of which was C(78-82,F87L,328-330), produce 12-hydroxyfarnesol as the major product, with approximately twofold increase in turnover number as compared with wild-type CYP102A1 toward farnesol. Methylfarnesoate was metabolized to 12-hydroxymethylfarnesoate (70%) and 10,11-epoxymethylfarnesoate (juvenile hormone III) (30%). The latter is metabolized to 65% 12-hydroxy-10,11-epoxymethylfarnesoate and 35% 15-hydroxy-10,11-epoxymethylfarnesoate. Substitution of residues 328–330, APA, by VPL was crucial to accomplishing this change in product.

Keywords Cytochrome · Heme · Molecular modeling · Nuclear magnetic resonance · Site-directed mutagenesis

Introduction

CYP102A1 (also called cytochrome P450 BM3) is a "self-sufficient" monooxygenase from *Bacillus megaterium* [1–7]. It is a bacterial class II [7] cytochrome P450 enzyme which contains a heme domain fused to a FAD/FMN diflavin reductase [8] domain in a single 119-kDa polypeptide [1, 4]. In the presence of NADPH and O_2, this soluble enzyme catalyzes hydroxylation of medium- and long-chain fatty acids such as lauric, myristic, palmitic, and stearic acids at the ω-1, ω-2, and ω-3 carbons [1, 2], and epoxidizes double bonds stereospecifically [9]. Its fusion of heme and reductase domains affords monooxygenation of saturated or unsaturated fatty acids with the highest catalytic activities of any of the P450 monooxygenases [3, 10]. CYP102A1 has attracted attention as a template for engineering new enzymes with modified regio- and stereospecificities and/or substrate selectivities for biotechnological and biomedical uses [9–22].

Several structures of the heme domain of CYP102A1 have been reported [23–28], two of which have substrate,

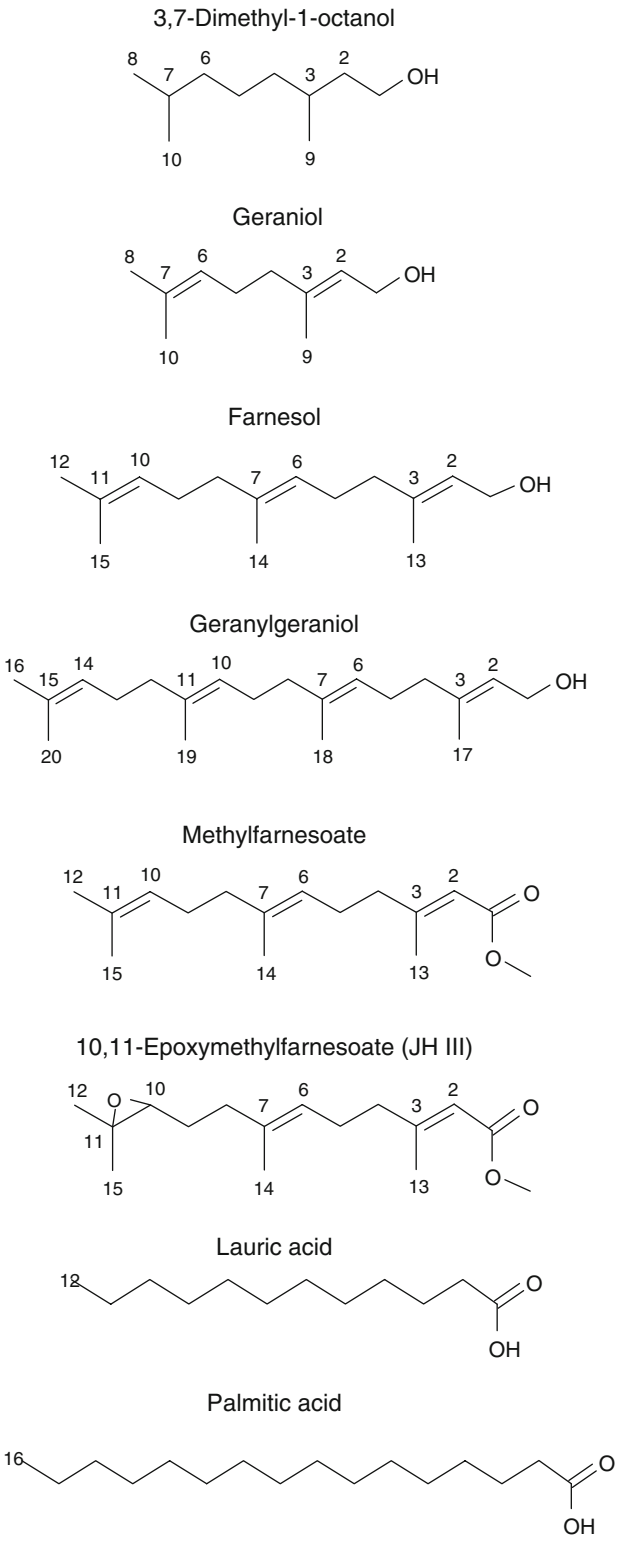
Electronic supplementary material The online version of this article (doi:10.1007/s00775-009-0580-y) contains supplementary material, which is available to authorized users.

C.-K. J. Chen · R. E. Berry · T. Kh. Shokhireva · M. B. Murataliev · H. Zhang · F. A. Walker (✉)
Department of Chemistry and Biochemistry,
University of Arizona,
Tucson, AZ, 85721-0041, USA
e-mail: awalker@u.arizona.edu

either palmitoleic acid or *N*-palmitolylglycine, bound to the ferriheme state of the enzyme [24, 26]. A puzzling feature of these structures is that the closest approach of the substrate to the heme iron is 7.75 Å [24, 26]. In agreement with this large distance, NMR paramagnetic relaxation studies of the substrate-bound form of the ferriheme domain of CYP102A1 [29] have shown that in the ferric form the substrate is much too far from the heme iron for the hydroxylation reaction to occur. However, upon one-electron reduction the substrate moves about 6 Å closer to the ferrous iron center [30], which indicates a change in the structure of the substrate-binding channel upon reduction of heme iron. The complete 3D structure of the substrate-bound ferrous form of the heme domain of CYP102A1 is not known.

Site-directed mutagenesis [5, 7, 10, 14, 31–50] and directed evolution [51–61] have been used to produce modified P450 enzymes with altered catalytic activities, regiospecificities, and stereospecificities; many of these have been single, double, or triple mutants of CYP102A1 [5, 7, 10, 27, 31–34, 37–44, 50]. Several "chimeras" of human cytochromes P450 have also been prepared, where these enzymes were fused with the reductase domain of CYP102A1 to create functional two-domain proteins [62, 63]. In our laboratory "scanning chimeragenesis" was developed to introduce peptide sequences from a P450 enzyme for which no crystal structure is available into homologous regions of CYP102A1, as a way of changing substrate selectivity and regiospecificity [64, 65]. The insect P450 enzyme CYP4C7, isolated from *Diploptera punctata* (the cockroach) [66], was selected for peptide sequence replacement because one of its substrates, farnesol (Structure 1), is also a substrate of CYP102A1, but the two enzymes produce different products [64, 66]. Farnesol is known as a precursor of 10,11-epoxymethylfarnesoate, also known as juvenile hormone III, which exists in cockroaches and mosquitoes [66–72]. The primary function of 10,11-epoxymethylfarnesoate is the control of insect development, metamorphosis, and reproduction [66, 68, 72]. CYP4C7 is believed to suppress the synthesis of 10,11-epoxymethylfarnesoate, and it metabolizes this hormone to 12-hydroxy-10,11-epoxymethylfarnesoate [66]. CYP4C7 also reacts with farnesol to produce 12-hydroxyfarnesol as the primary metabolite [66], whereas CYP102A1 oxidizes farnesol to a mixture of 2,3-epoxyfarnesol, 10,11-epoxyfarnesol, and 9-hydroxyfarnesol [64, 73]. The goal of this project is to produce an active enzyme, using the CYP102A1 scaffold, which catalyzes the oxidation of farnesol to 12-hydroxyfarnesol, as proof of principle of the "scanning chimeragenesis" approach.

In addition to its importance as a precursor of 10,11-epoxymethylfarnesoate in insects, farnesol has emerged as a potential mammalian signaling molecule implicated in regulation of a wide variety of cell functions. Initial studies suggested farnesol was a regulator of 3-hydroxy-3-methylglutaryl-CoA reductase activity [74–77]. Later studies implicated farnesol in regulation of tumor cell



Structure 1

proliferation and apoptosis [78–80] and regulation of vascular tone and blood pressure [81, 82]. It was also found that farnesol is a potent blocker of neuronal N-type voltage-gated Ca^{2+} channels, which suggested it may be involved in regulation of neurotransmission and thus brain function [83].

Farnesol has also been demonstrated to inhibit rabbit CYP2E1 in hepatic microsomes [84]. It was also shown that farnesol is metabolized to 12-hydroxyfarnesol by CYP2E1 [85], which in humans is a very important liver enzyme for metabolizing endogenous molecules such as both native *cis*-5,6-arachidonic acid [86] and also *trans*-5,6-arachidonic and epoxyeicosatrienoic acids [87], as well as more than 70 low molecular weight substrates [88–92]. CYP2E1 levels, and the toxicity of metabolites, vary markedly in response to alcohol consumption [93], diabetes [94], obesity [95], and fasting [96]. Activation of acetaminophen by CYP2E1 into strongly electrophilic *N*-acetyl-*p*-benzoquinone imine is the initial and major step responsible for acetaminophen hepatotoxicity [97, 98]. The structures of two heme ligand complexes of CYP2E1 have recently been reported [99], which will permit comparison of its binding pocket with a third-generation chimera of CYP102A1 and CYP4C7 of this study.

The "scanning chimeragenesis" experimental approach began by identifying individual structural elements of CYP102A1 which participate in recognition and binding to the typical substrates of that enzyme, and replacing them with the homologous peptides of CYP4C7 [64]. To identify the sites, the sequence of CYP102A1, for which two X-ray structures of the heme domain with substrate bound have been reported [24, 26], was aligned with that of CYP4C7, for which no structure is available, using the GAP program of the GCG package [100] and FASTA3 [101]. This alignment was used with the crystal structures [24, 26] to identify substrate-interacting residues of CYP102A1, and thus CYP4C7, to be considered for replacement in CYP102A1. Many of the residues within our chosen 5-Å distance from substrate atoms were found to be within the substrate recognition sites (SRSs) of CYP102A1. Residues inferred as being part of the six SRSs were identified by X-ray crystallography of substrate-bound forms of cytochromes P450, first of the CYP2 family [102], but now all CYP enzymes, including CYP102A1 [14]. SRS-1–SRS-6 are located along the primary sequences of CYP enzymes, and constitute about 16% of the total residues [102] (of the heme domain for CYP102A1). They occur in structurally homologous regions of the amino acid sequence, although there is little or no conservation of actual amino acid *identity* from one P450 to another.

A study of a series of SRS-1 (the β'-helix) chimeric proteins that used portions of residues 73–82 has been published [64]. It showed that whereas the full-length chimera, C(73-82), has very low activity, the shorter chimeras C(73-78), C(75-80), and C(78-82) have higher activities and change the regiospecificity of the enzyme to produce new metabolites that were identified as 10-hydroxypalmitic acid and 5-hydroxyfarnesol. Neither of these is produced by wild-type CYP102A1. We then extended this study to include residue 87 [65], which is phenylalanine for CYP102A1, but leucine for CYP4C7. We found that although *all* of the F87-mutated chimeric enzymes had altered regiospecificity compared with the wild-type CYP102A1, *most* also had altered regiospecificity compared with the first-generation chimeric proteins [64], and some also had strongly altered substrate selectivity. In particular, C(78-82,F87L) (where the numbers indicate the amino acids of CYP102A1 that were replaced with the homologous residues of CYP4C7) was most active toward farnesol, 5.7-fold more active than the wild-type enzyme. It produced one primary metabolite, which was identified by NMR spectroscopy as 10,11-epoxyfarnesol [65]. This chimera had very low activity toward lauric and palmitic acids, and thus its substrate selectivity was altered from medium- and long-chain fatty acids to farnesol. When farnesol was incubated with C(78-82,F87L) for 15 min, secondary metabolites were detected [10,11-epoxy-9-hydroxyfarnesol and (2,3),(10,11)-bisepoxyfarnesol], which were produced in a 1:1:4 ratio to the primary product, 10,11-epoxyfarnesol. But *none* of these F87-mutated chimeric enzymes produced ω-hydroxylated products with substrates palmitic acid or farnesol [65]. This study thus showed that although chimeragenesis of the 73–87 peptide of CYP102A1 is capable of changing substrate selectivity from fatty acids to farnesol, *it is not alone sufficient to produce the product of CYP4C7, 12-hydroxyfarnesol*. Thus, we have extended our study to other SRSs.

Partial sequence alignment of the heme domain of CYP102A1 with CYP4C7 is shown in Fig. 1, where amino acids found by X-ray crystallography to be within 5 Å of the substrates palmitoleic acid [24] and *N*-palmitolylglycine [26] are marked. Of the other SRSs, our preliminary evaluation, based on the structure of CYP102A1 in that region (β1–4 residues TAPAPS), and the larger sizes of the corresponding residues of CYP4C7 (VVPLIA), suggested that SRS-5 might be a good candidate for achieving the goal of producing 12-hydroxyfarnesol. Thus, the amino acid residues 327–332 and 328–330, along with amino acids 78–82 and F87A or F87L, were replaced with the homologous fragments of CYP4C7. Six chimeric proteins were expressed, purified, and characterized. Four chimeras produced 12-hydroxyfarnesol as the primary product. C(78-82,F87L,328-330) is the best, but C(328-330) alone is able to produce 12-hydroxyfarnesol, although it does not inhibit the oxidation of fatty acids.

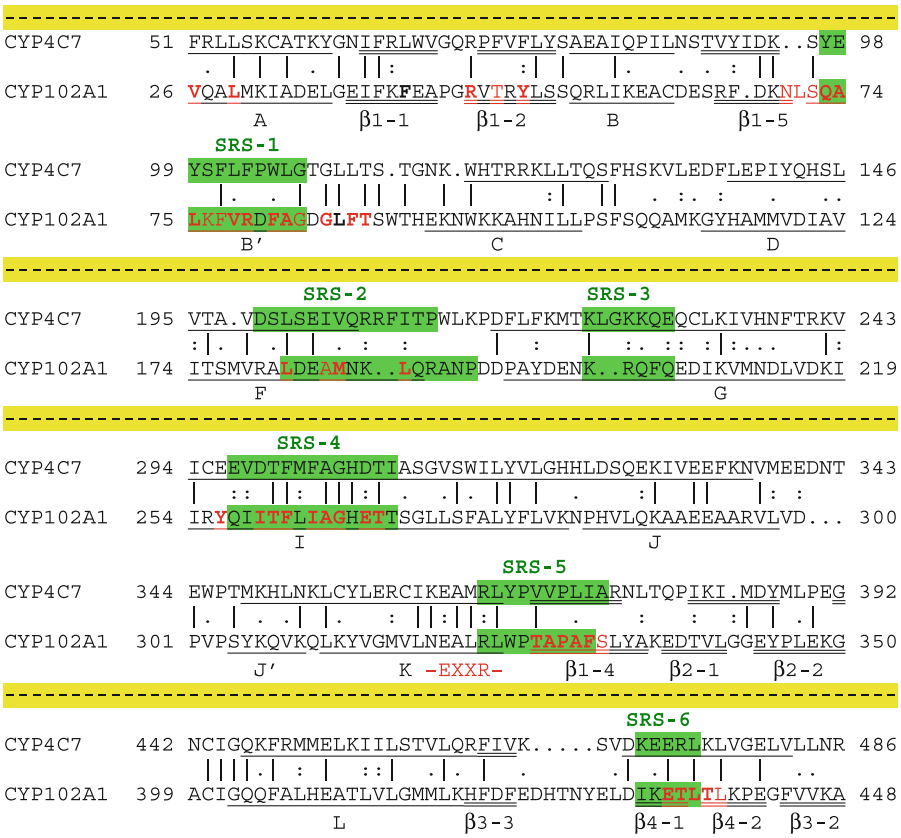


Fig. 1 Partial sequences of CYP4C7 and the heme domain of CYP102A1 showing the alignment generated with the GAP program of the GCG package [100]. Residues which map to within 5 Å of the bound palmitoleic acid in Protein Data Bank (PDB) file 1FAG [24] are shown in *bold*, and those which map to within 5 Å of the bound *N*-palmitolylglycine in PDB file 1JPZ [26] are shown in *red*. The α-helices are *underlined* and named, and *β*-sheets are *double underlined* and named. The substrate recognition sites (*SRS-1* to *SRS-6*) [102] are *shaded green* and labeled. Gaps in the sequences, between the substrate recognition sites, are signified by *dashed lines* highlighted in *yellow*. The –EXXR– motif [50] mentioned in "Discussion" is highlighted in *red* as part of the K helix

Materials and methods

All primers were obtained from Operon. The polymerase and restriction enzymes were ordered from Bio-Labs. Substrates lauric and palmitic acids, farnesol, geraniol, geranylgeraniol, and 3,7-dimethyl-1-octanol were purchased from Sigma-Aldrich. The 10,11-epoxymethylfarnesoate and methylfarnesoate were purchased from Echelon. The solvents used for all experiments were high performance liquid chromatography or higher grade. All instruments used in this project were the same as used previously [64, 65]. NMR spectra, including 1D ^{1}H and 2D correlation spectroscopy (COSY) and ^{1}H{^{13}C} heteronuclear multiple quantum correlation spectra, were obtained using a Bruker DRX-600 spectrometer operating at 600.03 MHz for the proton Larmor frequency. Spectra were collected at 25 °C with the proton chemical shifts referenced to residual solvent (CD_3OD) CD_2H. 2D NMR spectra were collected with 2,048 data points in t_2 and with 256–512 blocks in t_1 with 32–64 scans per block.

Mutant gene preparation, protein expression, and purification

The amino acid sequences of CYP102A1 and CYP4C7 for the regions encompassing SRS-1 and SRS-5, together with the amino acid replacements for the mutant proteins, are shown in Table 1. The mutant gene preparation, protein expression, and purification followed the protocols published previously [64, 103], using the primers shown in Table S1. Clones with the correct insert size for the section of the gene being mutated were sequenced several times to confirm the presence of the fragment of the sequence specific for CYP4C7. After insertion of the portion of the gene that had been mutated into the pbs-BM3 gene, the

Table 1 Amino acid substitutions of chimeric proteins

CYP102A1 (71-89)	**LSQALKFVRDFAGDGLFTS**
CYP4C7 (95-113)	**LSYEYSFLFPWLGTGLLTS**
C(78-82)	**LSQALKFLFPWLGDGLFTS**
C(78-82,F87A)	**LSQALKFLFPWLGDGLATS**
C(78-82,F87L)	**LSQALKFLFPWLGDGLLTS**
CYP102A1 (323-340)	**RLWPTAPAFSLYAKEDTV**
CYP4C7 (366-383)	**RLYPVVPLIARNLTQPIK**
C(327-332)	**RLWPVVPLIALYAKEDTV**
C(328-330)	**RLWPTVPLFSLYAKEDTV**

chimeric proteins were expressed in *Escherichia coli* as described previously [64, 65]. The expressed proteins were judged to be homogeneous by sodium dodecyl sulfate polyacrylamide gel electrophoresis. These proteins were preliminarily characterized by UV/vis spectroscopy [3, 5, 64, 65] and matrix-assisted laser desorption/ionization (MALDI) time of flight (TOF) mass spectrometry (MS), as described previously [65]. All proteins were also expressed, purified, and characterized in two separate batches to confirm the reproducibility of the protein preparation. The final purified proteins were extensively dialyzed against 50 mM potassium phosphate buffer (pH 7.6), concentrated, and then stored at −20 °C in 50% glycerol.

UV/vis spectroscopy

The resting state spectra were recorded over the wavelength range from 300 to 850 nm, as reported previously, and are the same as those of the first- and second-generation chimeras reported previously, i.e., the spectra were those typical of wild-type CYP102A1, including having flavin absorbance in the 450–525-nm range (Fig. S1a) [3, 5, 64, 65]. The CO-bound difference spectra of the dithionite-reduced proteins were measured, as reported previously [3, 5, 7], and are comparable to those of other cytochromes P450 [104, 105] (which lack the diflavin reductase domain). With use of argon-purged buffers [200 mM tris(hydroxymethyl)aminomethane (Tris)–HCl buffer at pH 7.7], a solution of each protein (1 μM in 3 mL) was reduced with 20 μL of a saturated solution of sodium dithionite (freshly prepared in the same buffer) and the spectrum was recorded. Then 20-μL aliquots of a saturated solution of CO in the same buffer (calculated from its solubility as approximately 800 μM) were added to the solution (at 25 °C) with mixing. The spectrum was recorded again following each addition until there was no further change in the spectrum (Fig. S1b) [3, 5, 7, 104, 105].

Enzyme activity assays for cytochrome *c* reduction and NADPH oxidation

To measure the reductase activity, the absorbance at 550 nm ($\varepsilon_{ox} = 8.9\ mM^{-1}\ cm^{-1}$, $\varepsilon_{red} = 29.9\ mM^{-1}\ cm^{-1}$ [5]) was followed for 1-mL samples containing 2 nM enzyme and 50 μM oxidized cytochrome *c* containing the NADPH-regenerating system (1.0 mM glucose 6-phosphate, 3.0 U mL^{-1} glucose 6-phosphate dehydrogenase, and 50 μM NADPH). This was repeated for a solution containing no enzyme, and the difference in rate is reported [5, 6, 103]. NADPH oxidation was measured in the presence of 100 μM nucleotide, 10 nM enzyme, 50 μM oxidized cytochrome *c*, and 200 μM laurate by following the decrease in absorbance at 340 nm and using an extinction coefficient of 6.22 $mM^{-1}\ cm^{-1}$ [103]. The enzyme activities with cytochrome *c*, NADPH, lauric acid, palmitic acid, farnesol, methylfarnesoate, and 10,11-epoxymethylfarnesoate are shown in Table 2.

Enzyme activity assays for farnesol, methylfarnesoate, 10,11-epoxymethylfarnesoate, lauric acid, and palmitic acid

All enzyme incubations and activity assays were carried out at 25 °C in 3-mL volumes (containing 100 mM Tris–HCl pH 7.6 buffer and the NADPH-regenerating system consisting of 3.0 mM glucose 6-phosphate, 3.0 U mL^{-1} glucose 6-phosphate dehydrogenase, and 100 μM NADPH). Farnesol, methylfarnesoate, 10,11-epoxymethylfarnesoate, and related substrates, lauric acid or palmitic acid (250 μM) was added and the solution was incubated with 1–10-μL aliquots (0.65–13.3 μM) of each enzyme separately for 1–30 min to measure the individual enzyme activities. The incubation time was chosen on the basis of the activity of each enzyme, so as to yield sufficient products for full characterization. The reactions were stopped by adding 200 μL of 3 N HCl per reaction volume of 3 mL. For lauric and palmitic acid assays, the reaction products were extracted with 6 mL of diethyl ether and were methylated with diazomethane (Sigma-Aldrich) to convert the carboxylic acid to its methyl ester. For farnesol, methylfarnesoate, 10,11-epoxymethylfarnesoate, and other related substrates, after the reaction had been stopped with 200 μL of 3 N HCl, the reaction products were extracted using a 200-μg C_{18} solid-phase extraction column (Honeywell Burdick and Jackson) and were eluted with acetonitrile. The products of the reactions were analyzed by gas chromatography (GC). The integrated areas of the substrate and product peaks were used to calculate the enzyme reaction rate with each substrate. Experimental errors were determined by averaging GC peak areas for three to five activity measurements for a given substrate.

Product isolation

A common GC diagnostic derivatization reagent, *N*,*N*-bis(trimethylsilyl)-2,2,2-trifluoroacetamide (BSTFA) (Sigma-Aldrich), was used to increase the volatility of the reaction products to improve their separation [64, 65, 106]. A 500-μL aliquot of BSTFA (5 mg mL^{-1} in dimethylformamide) was incubated overnight with the methylated fatty acid reaction mixture before separation by GC.

All of the farnesol, methylfarnesoate, 10,11-epoxymethylfarnesoate, geraniol ,and other substrate metabolites were dried with N_2 gas and dissolved in 200 μL acetonitrile. This concentrated solution of the metabolites was diluted with 100 mM Tris–HCl buffer at pH 7.6 and loaded

Table 2 Chimeric protein turnover numbers (nmol substrate min^{-1} $nmol^{-1}$ enzyme) with cytochrome *c*, NADPH, and substrates lauric acid, palmitic acid, farnesol, methylfarnesoate, and 10,11-epoxymethylfarnesoate

	Cytochrome *c*	NADPH	Lauric acid	Palmitic acid	Farnesol	Methylfarnesoate	10, 11-Epoxymethyl-farnesoate
CYP102A1	2,980 ± 230[a]	1,650 ± 120[a]	705 ± 35[a]	2,980 ± 140[a] 3,160 ± 60[b]	285 ± 22[a] 300 ± 8[b]	45 ± 5	ND
CYP4C7	–	–	–	–	4.1 ± 0.5[c]	2.3 ± 0.2	0.83 ± 0.27
C(78-82)	3,330 ± 270[a]	1,690 ± 110[a]	305 ± 25[a]	2,790 ± 160[a] 2,640 ± 340[b]	345 ± 32[a] 360 ± 5[b]	NM	NM
C(78-82,F87A)	3,150 ± 140[a]	1,410 ± 80[a]	6.9 ± 0.2[a]	23 ± 1[a]	268 ± 3[a]	106 ± 7	ND
C(78-82,F87L)	3,660 ± 270[a]	1,880 ± 140[a]	4.0 ± 0.1[a]	4.0 ± 0.1[a]	1,700 ± 70[a]	578 ± 37	ND
C(327-332)	2,690 ± 150	1,880 ± 140	40 ± 3	248 ± 18	41 ± 3	47 ± 2	ND
C(78-82,F87A,327-332)	1,980 ± 80	1,070 ± 150	1.3 ± 0.3	29 ± 4	31 ± 4	25 ± 3	ND
C(78-82,F87L,327-332)	3,280 ± 100	1,750 ± 110	1.0 ± 0.4	41 ± 3	32 ± 2	20 ± 1	ND
C(328-330)	3,230 ± 350	1,610 ± 300	135 ± 11	307 ± 35	346 ± 18	944 ± 66	ND
C(78-82,F87A,328-330)	2,640 ± 190	1,550 ± 130	ND	6.8 ± 2.8	80 ± 8	70 ± 7	ND
C(78-82,F87L,328-330)	2,770 ± 290	1,300 ± 120	1.7 ± 0.5	1.9 ± 0.1	567 ± 45	92 ± 2	113 ± 9

Substrate turnover numbers represent the amount of product formed by 1 μM enzyme reacting with 250 μM substrate per minute. The numbers are the average of six measurements, with standard deviation from two separate batches shown

ND no product detected, *NM* not measured

[a] Taken from [65]

[b] Taken from [64]

[c] Taken from [66]

onto a 200-μg C_{18} solid-phase extraction column. Epoxyfarnesols, geraniols, and epoxymethylfarnesoates were eluted with 2.6 mL of a methanol/water mixture (42:58 ratio) and hydroxyfarnesols, hydroxymethylfarnesoates, and all other hydroxyl derivatives were eluted with 3.0 mL of an acetonitrile/water mixture (47:53 ratio); to separate 9-hydroxyfarnesol and 12-hydroxyfarnesol, a solvent mixture of 40:60 acetonitrile/methanol was used. After the removal of solvents, the metabolites were dissolved in tetrahydrofuran for GC. For GC–MS the hydroxyl products were reacted with BSTFA as described for the fatty acid methyl esters.

Product characterization

The GC substrate and product peaks were identified using the retention times published previously [1, 2, 64, 65]. The temperature of the column oven was programmed as follows: initial temperature set to 100 °C, held for 2 min, then increased at a rate of 15 °C min^{-1} to 250 °C, and then held for 5 min. Trimethylsilyl (TMS) derivatives of hydroxypalmitic and hydroxylauric acid methyl esters were analyzed by GC–MS. Two groups of ions were used to identify reaction products. The first group was common to all of the hydroxylated fatty acid methyl ester derivatives, and are listed as A, B, C, and D, respectively, in Table S2 for lauric acid and Table S3 for palmitic acid. The second group includes fragmentation peaks differing by 14 mass units (CH_2), which showed the position of hydroxylation. The products of lauric acid hydroxylation by wild-type CYP102A1 were identified as 11-OH, 10-OH, 9-OH, 8-OH, 7-OH, and 6-OH, corresponding to ω-1 to ω-6 hydroxylation, and for palmitic acid hydroxylation the products were identified as 15-OH, 14-OH, 13-OH, 12-OH, 11-OH, and 10-OH, corresponding to ω-1 to ω-6 hydroxylation, with ω-1 to ω-3 being the major products in both cases.

For farnesol, methylfarnesoate, and all other related substrates, products were separated first by GC. All epoxides had very similar retention times and often could not be resolved, whereas the hydroxylation products had longer retention times than the epoxides, and increased retention times as the hydroxyl group was further displaced from the C-1 hydroxyl. Following partial separation of products by GC, the partially purified products were examined by 1D and 2D NMR spectroscopy in CD_3OD. All of the new farnesol and other terpene metabolites have been identified by ^1H NMR and GC–MS. The ions used to identify hydroxy reaction products of farnesol are given in Table S4, those for methylfarnesoate and 10,11-epoxymethylfarnesoate are given in Table S5, those for geraniol are given in Table S6, and those for 3,7 dimethyl-1-octanol are given in Table S7.

The NMR assignments for each new product are shown in Table S8. The product ratios with farnesol are shown in Table S9, those with methylfarnesoate are shown in Table S10, those with 10,11-epoxymethylfarnesoate are shown in Table S11, and those with geraniol are shown in Table S12. The relative turnover numbers for 3,7 dimethyl-1-octanol, geraniol, farnesol, and geranylgeraniol are given in Table S13.

Structural models and substrate docking

A detailed discussion of the computational modeling is presented in the electronic supplementary material. Models of the chimeras, docking of the substrates, and figures were produced using the AutoDock version 4.0 [107, 108], ADT4.0 software package in AutoDockTools (version 1.5.2, revision 2) [109], VMD (version 1.8.6) [110], SwissPDB [111] (version 4.0.1, including a version of the GROMOS 43B1 force field [112]), and Persistence of Vision Ray Tracer (POV-Ray version 3.6). The CYP102A1 X-ray structure [Protein Data Bank (PDB) code 1FAG] was used as the modeling template.

Results

Protein characterization

The absorption spectra of the mutant proteins of this study, like those reported previously [65], are essentially identical to those of wild-type CYP102A1, which suggests that the mutant proteins fold properly and contain all of the cofactors required for catalysis. As shown in Fig. S1 for two triple-chimera examples as compared with wild-type CYP102A1, the proteins have similar reduced CO difference spectra, with Soret bands at 450 nm. Thus, these are typical cytochrome P450 proteins, with no detectable P420 inactive form present [104].

As reported previously [65], MALDI-TOF measurements on CYP102A1 yielded a molecular mass of 117,638 *m*/*z* within an expected error of ±0.1% (about 118 *m*/*z*) in contrast to a calculated molecular mass of the protein of 117,666 Da. In this molecular mass range a ±0.1% error is to be expected (V.H. Wysocki, personal communication). The entire protein mass, including the heme and the flavins, was thus calculated on the basis of a molecular mass of 119,525 Da; molecular masses of the mutant proteins of this study were calculated on the same basis, by including the masses of the mutated residues in place of those of CYP102A1. The MALDI-TOF measurements of C(78-82), C(78-82,F87L), and C(78-82,F87L,328-330) are 117,850, 117,816, and 117,889 Da within ±0.1% error in each case. Although the ±0.1% error prevents differentiation of chimeras with different protein sequences, it shows that the two-domain structure is intact. The fact that the gene sequence of each chimera was checked multiple times during mutant gene production removes any doubt as to the protein sequences of the chimeric enzymes.

Turnover numbers of the mutant proteins with respect to reduction of cytochrome *c* and oxidation of NADPH and substrates laurate, palmitate, farnesol, methylfarnesoate, and 10,11-epoxymethylfarnesoate were measured as described in "Materials and methods," and are presented in Table 2. As can be seen, the cytochrome *c* reductase turnover numbers of most chimeric enzymes of this study are fairly similar to that of wild-type CYP102A1, and the NADPH oxidase turnover numbers are approximately half those of the cytochrome *c* reductase reactions. Thus, NADPH transfers electrons well through these enzymes to cytochrome *c*, with a stoichiometry of one NADPH per two cytochrome *c* molecules. Among the chimeric enzymes of this study, the cytochrome *c* reductase and NADPH oxidase turnover numbers of the chimera C(78-82,F87A,327-332) are both only about 60% those of wild-type CYP102A1 and the second-generation chimeras. The turnover numbers of this chimera toward substrates lauric and palmitic acid, farnesol, and methylfarnesoate are also very low, but so are those of the F87L counterpart, C(78-82,F87L,327-332), which has cytochrome *c* reductase and NADPH oxidase turnover numbers similar to those of wild-type CYP102A1.

However, with respect to the monooxygenase activity toward the substrates, the turnover numbers toward laurate and palmitate are markedly decreased for all of the chimeras of this study, and more so for the chimeras that involve replacements within both SRS-1 and SRS-5. The same is also true for the substrate farnesol, except for the increase in turnover number of C(328-330) (SRS-5 mutation only) and the triple chimera C(78-82,F87L,328-330) to a value nearly double (factor of 1.7) that of wild-type CYP102A1.

GC–MS identification of metabolites

Fatty acids (lauric and palmitic acids)

In our two previous publications we identified the major oxidation products of palmitic and lauric acid methyl esters as their TMS derivatives by GC–MS analysis [64, 65]. The mass fragmentation tables are shown in Tables S2 and S3. As previously reported, the products of palmitate modification by the chimeras were identified as ω-1 to ω-6 hydroxylated [64] palmitate, and the products of laurate modification were identified as ω-1 to ω-6 hydroxylated [65] laurate. No ω-hydroxylauric or ω-hydroxypalmitic

acid was detected as a product of any of the chimeric enzymes of this study.

Farnesol, (2E,6E)-methylfarnesoate, and (2E,6E)-10,11-epoxymethylfarnesoate

The single TMS derivatives of methylfarnesoate and 10,11-epoxymethylfarnesoate, and the double TMS derivatives of hydroxyfarnesols were analyzed by GC–MS. Farnesol, hydroxyfarnesols, methylfarnesoate, and 10,11-epoxymethylfarnesoate (Structure 1) have similar series of unique ions as presented in Tables S4 and S5. Treatment with BSTFA resulted in a characteristic series of *m/z* ions including, for example, for hydroxyfarnesols, *m/z* 73 $[(CH_3)_3Si]^+$, *m/z* 143 $[(TMSOCH_2)C(CH_3)CH]^+$, *m/z* 157 $[(TMSOCH_2)C(CH_3)CHCH_2]^+$, and *m/z* 382 $[M(TMSO)_2]^+$ as the total molecular weight. Since TMS derivatives have the tendency to fragment in the order primary > secondary > tertiary carbon, the results above provide evidence of hydroxyl group addition to farnesol. The position of hydroxyl addition was determined by ^{1}H NMR spectroscopy.

For (2*E*,6*E*)-methylfarnesoate and (2*E*,6*E*)-10,11-epoxymethylfarnesoate, similar sets of unique ions were observed, as listed in Table S5. The TMS derivatives of the two within the pair 12-hydroxymethylfarnesoate and 15-hydroxymethylfarnesoate, and also the two within the pair 12-hydroxy-10,11-epoxymethylfarnesoate and 15-hydroxy-10,11-epoxymethylfarnesoate each have the same GC–MS fragmentation patterns, respectively, but different GC retention times. On the basis of the 12-hydroxyfarnesol to farnesol GC retention time ratio, the major product of C(78-82,F87L,328-330) during methylfarnesoate and 10,11-epoxymethylfarnesoate oxidation was determined to be the 12-hydroxy isomer in both cases, the same as found by Helvig et al. [70] for CYP15A1.

Geraniol

The TMS derivatives of geraniol metabolites were analyzed by GC–MS. All unique ions of TMS derivatives of geraniol metabolites of the chimeric enzymes of this study are shown in Table S6. The results are consistent with reports of the GC–MS of the TMS derivatives of 10-hydroxygeraniol and 6,7-epoxygeraniol [31–33]. There were two hydroxygeraniol metabolites, both of whose TMS derivatives had the same GC–TOF-MS fragmentation patterns, but with different GC retention times. Since 12-hydroxyfarnesol had a longer GC retention time compared with 15-hydroxyfarnesol under the same GC conditions, the primary hydroxygeraniol was assumed to be 8-hydroxygeraniol, and the secondary hydroxygeraniol was assumed to be10-hydroxygeraniol. The position of hydroxyl addition was also confirmed by ^{1}H NMR spectroscopy.

3,7-Dimethyl-1-octanol

Reaction of the saturated analogue of geraniol, 3,7-dimethyl-1-octanol (Structure 1), with the chimeric enzymes of this study provided one metabolite, whose TMS derivative was analyzed by GC–MS. All unique ions of TMS derivatives of the metabolite of 3,7-dimethyl-1-octanol are shown in Table S7. The only metabolite was determined to be 6-hydroxy-3,7-dimethyl-1-octanol.

NMR spectroscopy of farnesol metabolites

GC–MS analysis showed that four of the new chimeras produced new farnesol metabolites that were shown by their retention times to be hydroxyfarnesols. Both SRS-5 single chimeras, C(327-332) and C(328-330), and both F87L triple chimeras, C(78-82,F87L,327-332) and C(78-82,F87L,328-330), showed the presence of 10,11-epoxyfarnesol (identified previously by NMR spectroscopy [64, 65]), and the same two hydroxyfarnesols, whose retention times were uniquely different from those of any of the previously identified hydroxyfarnesols [64, 65]. NMR analysis of the hydroxyfarnesol metabolites of C(78-82, F87L,327-332) is described herein; the metabolites of the other chimeric enzymes were analyzed in a like manner, and the products were found to be identical to those described below.

The hydroxyfarnesol fraction of metabolites of C(78-82,F87L,327-332) contained two compounds with GC intensity ratio A:B of approximately 3:2. The 1D ^{1}H NMR spectrum of the mixture is shown in Fig. 2b, and that of farnesol as a reference is shown in Fig. 2a. As can be seen, there are four multiplets in the region 5.0–5.5 ppm, with ratio 3:5:2:5, with the first two peaks partially overlapping. In this region farnesol has the resonances of C2–H, C6–H, and C10–H. In this NMR sample C2–H has the same chemical shift as that of farnesol itself, whereas C6–H is slightly shifted, C10–H has disappeared, and two multiplets, at 5.39 and 5.26 ppm with ratio 3:2, have appeared. The C1–H_2–OH methylene protons appeared at 4.08 ppm as a doublet for farnesol, and do not change their position and multiplicity in either of the products. There are two new singlets at 4.06 and 3.91 ppm that are characteristic of methylene protons on a carbon bearing a hydroxyl group. This means that the NMR sample contains two different primary hydroxyfarnesols in the ratio 3:2. There are three overlapping multiplets in the region 2.0–2.2 ppm, where C4–H_2, C5–H_2, C8–H_2, and C9–H_2 are observed for farnesol; the peak patterns of the new hydroxyfarnesols have a different appearance, with C9–H_2 and C8–H_2 peaks being shifted and overlapping with each other, and C5–H_2 and C4–H_2 remaining at the same positions as observed for farnesol.

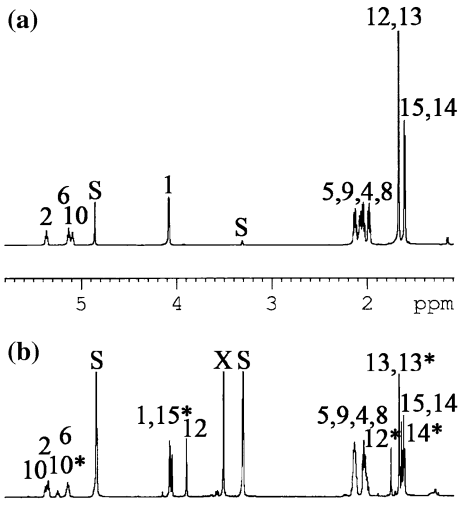
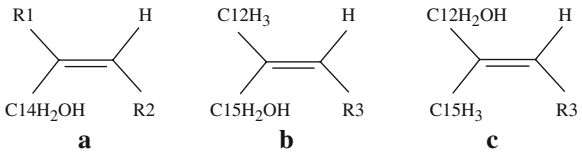


Fig. 2 The 1D ^{1}H NMR spectra of farnesol as a reference (**a**) and of the hydroxyfarnesol fraction of metabolites of C(78-82,F87L,327-332) (**b**). This fraction contained 12-hydroxyfarnesol and 15-hydroxyfarnesol with ratio approximately 3:2. Individual resonances of 15-hydroxyfarnesol are shown with *asterisks*. Spectra were recorded at 25 °C in CD_3OD. The appearance of two new peaks at 4.06 and 3.91 ppm is due to new –C*H*$_2$–OH groups as a result of 15-methyl and 12-methyl group oxidation, respectively. *S* solvent residual proton resonance, *X* unknown impurity. The chemical shifts of all resonances of these two new metabolites are presented in Table S8

The methyl resonances appear in the region 1.6–1.8 ppm. There are five different-intensity methyl peaks with ratios 2:5:3:3:2. The peak at 1.76 ppm is a doublet. The signal at 1.67 ppm from the $C13H_3$ methyl group has the same position in these two compounds as in farnesol. On the basis of peak intensities, the NMR proton resonances can be separated into two groups for compounds A and B: for A, 5.39, 5.36, 5.16, 4.08, 3.91, 2.14, 2.14, 2.04, 2.04, 1.67, 1.64, and 1.62 ppm; for B, 5.36, 5.26, 5.16, 4.08, 4.06, 2.14, 2.14, 2.04, 2.02, 1.76, 1.67, and 1.61 ppm.

In Fig. S2 the nuclear Overhauser effect spectroscopy correlations of the CH_3–CH_2C and CH_3–CH= regions of the NMR spectrum of farnesol are presented, to show what nuclear Overhauser effects can be observed in these systems. COSY and heteronuclear multiple quantum correlation spectra of the farnesol metabolites are presented in Fig. S3c and d, respectively. As can be seen in the COSY map of the mixture (Fig. S3c), there are through-bond couplings between protons at positions 1 and 2, 4 and 5, 5 and 6, 8 and 9, and 9 and 10.

The fact that there are two additional C*H*$_2$–OH singlets at 3.91 and 4.06 ppm suggests that methyl groups were oxidized. We excluded from consideration the $C13H_3$ group (Structure 1) because neither C2–H nor C1–H_2 changed its position after oxidation. Considering that the products are expected to be *trans,trans*-farnesol metabolites, in Structure 2 we present 14-hydroxyfarnesol,



Structure 2

15-hydroxyfarnesol, and 12-hydroxyfarnesol in the forms that show the positions of vicinal groups, as was done previously [113] for substituted alkenes. Here R1, R2, and R3 are the remaining parts of the molecule, (**a** C7=C6, **b** C11=C10, and **c** C11=C10).

As can be seen, structures **a** and **b** are similar to each other in having the single proton *trans* to CH_2OH, but structure **b** is closer to what is shown in [113], where R1, R2, and R3 are CH_3 in that work, but in our case R1, R2, and R3 are CH_2–C. For these two situations, we can expect the ^{13}C chemical shift to be close to $\delta_{CH_2OH} = 60.5$ ppm, $\delta_{C12H3} = 21.3$ ppm for structure **b** and there to be no effects on chemical shifts of other CH_3 groups for structure **a**. In the case of 12-hydroxyfarnesol (structure **c**), the ^{13}C chemical shift of C12–H_2OH is expected to be close to 68.3 ppm and for $C15H_3$ it should be around 13 ppm. On the basis of this analysis, we concluded that product A is 12-hydroxyfarnesol. To make a choice between 14-hydroxyfarnesol and 15-hydroxyfarnesol, it should be remembered that oxidation of the $C14H_3$ group should not affect the chemical shifts of the other three methyl groups, but we observed a very large shift that is typical for a methyl group in the α-position [113], which corresponds to the case for structure **b** (15-hydroxyfarnesol). Another fact that supports this assignment is that there is no sign of a C10–H resonance at 5.01 ppm, where it appears in farnesol. This means that the peak pattern in the case of 14-hydroxyfarnesol would be different from that observed. Thus, NMR data analysis allows us to assign compound A as 12-hydroxyfarnesol and compound B as 15-hydroxyfarnesol. The ^{1}H and ^{13}C chemical shifts of farnesol and 12-hydroxyfarnesol and 15-hydroxyfarnesol are shown in Table S8.

Discussion

On the basis of the superior catalytic activity of CYP102A1 among all known cytochromes P450 [1–7], this "complete enzyme" was selected as our experimental model to show how scanning chimeragenesis can be used to create enzymes with altered substrate selectivity and regiospecificity. CYP102A1 and CYP4C7, with farnesol as a common substrate, suggested that CYP4C7 would be a worthwhile target for proof of principle for scanning

chimeragenesis. Homologous peptide fragments of CYP4C7 were inserted into CYP102A1, with our final goal being to change the fatty acid hydroxylase CYP102A1 into a terpene hydroxylase with the same major product as shown by CYP4C7 (12-hydroxyfarnesol). This goal was achieved in three rounds of chimeragenesis. Surprisingly, only nine amino acids had to be changed to alter substrate selectivity and regiospecificity from fatty acids to farnesol and to produce the desired product, 12-hydroxyfarnesol, as the major metabolite. Only eight of these actually contact the substrate, since D80P is on the side of the helix pointing away from the substrate, at least in the homology model discussed below.

Since we could not be certain that the peptide fragment chosen would lie in exactly the same position in the chimera as it does in the P450 of unknown structure (because insertions/deletions in the primary sequence might translate or rotate the structural features somewhat), or that the substrate of that P450 would bind in the same way as it would to CYP102A1, in each generation of scanning chimeragenesis we chose several possible fragments or replacements in the region chosen (the "scanning" aspect). In the first generation, whereas C(73-82) had extremely low activity [64], which suggested that large-scale replacements would not fit within the CYP102A1 scaffold, C(73-78) and C(75-80) produced the three metabolites of wild-type CYP102A1, plus a new compound, 5-hydroxyfarnesol. Although it was the primary metabolite of both of these mutant enzymes, their turnover numbers were more than tenfold lower than those of wild-type CYP102A1. These two also shifted the regiospecificity toward the center of the fatty acid substrates [64]. In contrast, C(78-82) had turnover numbers similar to those of the wild-type enzyme, but produced no new products.

The F87 amino acid residue was selected as the second-generation replacement because of its position in the structures of the substrate-bound heme domain of CYP102A1, sandwiched between the substrate and the heme [23–26]. F87 had been the subject of several studies prior to our work, with F87G ,F87A, and F87V having been studied by several research groups [32–34, 38]. These point mutants were found to have major effects on substrate activity and regiospecificity, and it was reported that the F87A mutant catalyzed hydroxylation of both lauric and myristic acids almost exclusively at the ω-position [33]. However, in our study of the F87 point mutants [65], we found that although a small amount of ω-hydroxylauric acid was indeed produced, it was far from being a major reaction product, and ω-hydroxylaurate was not detected among the products of any of the other enzymes. In addition, when reacted with palmitate, ω-hydroxypalmitate was not detected for the F87A enzyme or for any of the other mutant enzymes [65].

With respect to farnesol metabolites, the point mutant F87A made 2,3-epoxyfarnesol, 6,7-epoxyfarnesol, 10,11-epoxyfarnesol, plus 9-hydroxyfarnesol, whereas the point mutant F87L made only 2,3-epoxyfarnesol,10,11-epoxyfarnesol, plus 9-hydroxyfarnesol [65]. In contrast, C(73-78,F87A) and C(78-82,F87A) produced 6,7-epoxyfarnesol as the only new product. 5-Hydroxyfarnesol remained the primary product for both C(73-78,F87A) and C(73-78,F87L), but these mutant enzymes no longer produced 9-hydroxyfarnesol. C(75-80, F87A) and C(75-80,F87L) no longer made 5-hydroxyfarnesol, unlike C(75-80) [64], and produced 10,11-epoxyfarnesol and 9-hydroxyfarnesol in a 3:1 ratio [65]. C(73-78,F87A), C(73-78,F87L), C(75-80,F87A), and C(75-80,F87L) remained low in turnover number toward farnesol. Thus, it was clear that the two earlier peptides in the β'-helix were *not* leading toward our goal. In contrast, C(78-82,F87A) changed the substrate selectivity toward farnesol and remained almost as active as wild-type CYP102A1 [65]. It also changed its regiospecificity by making 6,7-epoxyfarnesol, 10,11-epoxyfarnesol, and 9-hydroxyfarnesol in 1:1:1 ratio. But better yet, C(78-82,F87L) changed its substrate selectivity to farnesol and increased its activity by 5.7-fold over that of wild-type CYP102A1 [65]. Because of its high activity, it also became a 10,11-epoxyfarnesol epoxidase and hydroxylase by producing (2,3),(10,11)-bisepoxyfarnesol and 10,11-epoxy-9-hydroxyfarnesol as secondary metabolites. The results proved that F87 should be included in the β'-helix (SRS-1) scanning chimeragenesis, to cause alteration in substrate selectivity from fatty acids to farnesol [65].

CYP102A1 residues 327-332 (SRS-5) provided the third and final generation for scanning chimeragenesis in this system. Six chimeric proteins were created, C(327-332), C(328-330), C(78-82,F87A,327-332), C(78-82,F87L,327-332), C(78-82,F87A,328-330), and C(78-82,F87L,328-330). Figure 3 shows the relative amounts of each metabolite produced by these, as well as the wild-type proteins and the best F87-generation mutants. As can be seen, the chimeras C(78-82,F87A,327-332) and C(78-82,F87A,328-330) produced 6,7-epoxyfarnesol and 10,11-epoxyfarnesol in a 2:3 ratio as the primary products, similar to the C(78-82,F87A) ratio of 1:1. Neither of the F87A-containing chimeras produced any 12-hydroxyfarnesol or 15-hydroxyfarnesol. In contrast, C(327-332), C(328-330), C(78-82,F87L, 327-332), and C(78-82,F87L,328-330) all produced both 12-hydroxyfarnesol and 15-hydroxyfarnesol, with the 12-hydroxy isomer being the more abundant in all cases (Fig. 3). These four chimeric enzymes each produced three products, 10,11-epoxyfarnesol, 15-hydroxyfarnesol, and 12-hydroxyfarnesol, in ratios of 1:0.61:1.09, 1:1.99:3.27, 1:0.32:0.64 and 1:1.35:4.12, respectively. However, as shown in Fig. 4, the single-SRS chimeras C(327-332) and

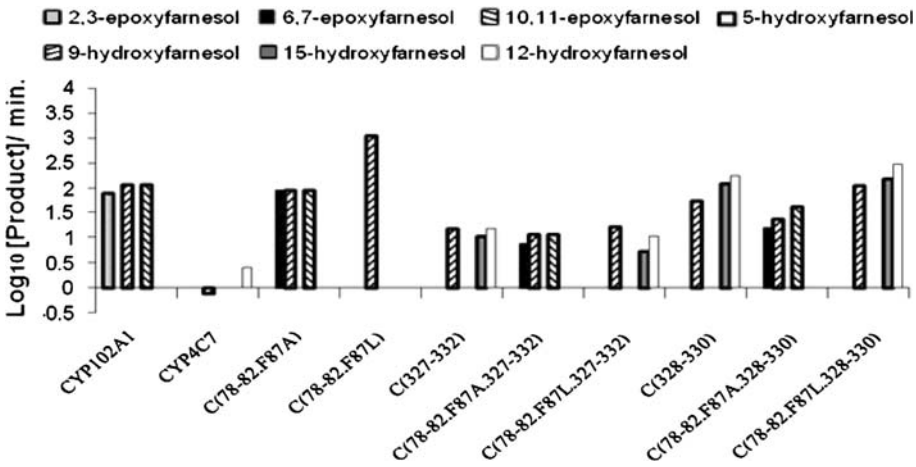


Fig. 3 The logarithm of the turnover number (concentration of individual hydroxy and epoxy products formed per minute) by 1 μM CYP102A1 or chimeric protein reacting with 250 μM farnesol to form hydroxyfarnesols and epoxyfarnesols

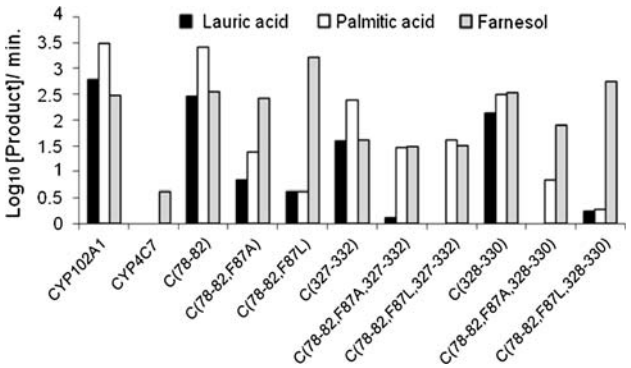


Fig. 4 The logarithm of the turnover number from Table 2 (total concentration of all hydroxy and epoxy products formed per minute) by 1 μM CYP102A1 or chimeric protein reacting with 250 μM lauric acid, palmitic acid, or farnesol

C(328-330) both retained relatively high activity toward lauric and palmitic acids, although lower than that of wild-type CYP102A1. In contrast, C(78-82,F87A,327-332) and C(78-82,F87L,327-332) showed very low turnover number toward lauric acid, and a bit of substrate selectivity change from fatty acids to farnesol, whereas C(78-82,F87L,328-330) made the substrate selectivity change strongly to farnesol. Furthermore, it produced 12-hydroxyfarnesol as the primary metabolite, with more than 100-fold greater turnover number than CYP4C7 and almost twofold greater turnover number than wild-type CYP102A1; C(78-82,F87L,328-330) is thus the best chimera we have found. These results are summarized in Scheme S1.

A328 is five amino acids after the conserved –EXXR– motif of SRS-5 of CYP102A1, which is shown by 6,379 CYP sequences investigated by Seifert and Pleiss [50]. Whereas most of the CYP sequences contain a hydrophobic residue at position 5 and a positively charged residue at position 10 or 11, CYP102A1 has alanine at position 5 and leucine and tyrosine at positions 10 and 11. However, CYP4C7 has valine at position 5 and arginine and asparagine at positions 10 and 11. The positively charged residue is expected to hydrogen-bond to the closest propionate of the heme [50], the 6-propionate in the case of CYP102A1. In CYP102A1 both leucine and tyrosine are quite far from the heme carboxylates. The most important residue of this group is that at position 5, which is larger for CYP4C7 (leucine) than for CYP102A1 (alanine). We find that the residue at position 7 after the –EXXR– motif is also important. This residue is alanine for CYP102A1 and leucine for CYP4C7, position 330. As predicted by Seifert and Pleiss [50], residues in this area "constrain the substrate orientations close to the activated oxygen and limit the accessibility to the active site. As a consequence, increasing the size of their side chains is expected to decrease the number of reaction products and thus to increase regioselectivity." This prediction was made in mid-2008, after we had completed our experimental work on SRS-5, and it is clear that their analysis of structures of a large number of CYP enzymes was very prophetic.

With regard to lauric acid and palmitic acid as compared with farnesol, it should be noted that wild-type CYP102A1 and even the first round of chimeras show nearly a factor of 10 lower activity toward laurate than toward palmitate, and the factor becomes much larger for second- and third-generation chimeras, to the point where, for the best third-generation chimera, C(78-82,F87L,328-330), has 2.5 orders of magnitude lower activity toward laurate than toward farnesol. Since the two are the same length, it is important to consider the difference between them. As a carboxylic acid, it was expected that the carboxyl groups of palmitoleate and *N*-palmitolylglycine would form ionic interactions with a positively charged group such as R47 [24, 34, 38]. However, in both the palmitoleate (1FAG) and the palmitolylglycine (1JPZ) structures, R47 is approximately 4.5 Å from the carboxyl group. In contrast, the

structure of CYP152A1 from *Bacillus subtilus* has the palmitoleic acid in "backwards," with the carboxyl end near the heme and forming a good hydrogen bond or ionic interaction to R242 (2.82 Å, PDB file 1IZO). Thus, although R47 may be involved in *steering* the fatty acid into the binding pocket [10], it is not clear what stabilizes the carboxyl end of the fatty acids bound to CYP102A1, or whether lauric acid can move the end of its hydrophobic tail in to the same extent as does palmoleic acid or *N*-palmitolylglycine, despite the shorter hydrocarbon chain. Surprisingly, neither substrate has any interactions with protein side chain atoms within 3 Å; the closest interaction is between the carboxylate and Y51-OH, at 3.08 Å (1FAG), but between the carboxamide carbonyl and the same side chain, at 2.56 Å for *N*-palmitolylglycine (1JPZ). But a major difference between lauric acid and farnesol is that farnesol likely does not bind in the extended form shown in Structure 1, as will be discussed further below.

All of the chimeric proteins of this study inhibited the production of 2,3-epoxyfarnesol, as did CYP4C7 [66]. And only C(78-82,F87A,327-332) and C(78-82,F87A,328-330) inherit the ability to make 6,7-epoxyfarnesol from C(78-82,F87A). Chimeric proteins containing F87L, where leucine is the residue present at that position in CYP4C7, a side chain that appears almost as bulky as the phenylalanine of CYP102A1, nevertheless are much better able to regulate the activity of the enzyme toward farnesol, and to inhibit production of 6,7-epoxyfarnesol. C(78-82,F87L, 327-332) has 15-fold lower farnesol turnover number compared with C(78-82,F87L,328-330). Interestingly, as mentioned above, the single-SRS chimeras C(327-332) and C(328-330) are able to produce 12-hydroxyfarnesol and 15-hydroxyfarnesol as major products, but without change in substrate selectivity. Only C(78-82,F87L,328-330) shows not only 12-hydroxyfarnesol as the major product, but also is able to reduce to a very low level the turnover numbers toward fatty acids.

From comparison of the oxidation of farnesol and its relatives, methylfarnesoate and 10,11-epoxymethylfarnesoate, by CYP102A1 and the chimeric proteins in Fig. S5, C(328-330) showed the highest rate of formation of products from methylfarnesoate among all substrates that were tested. Of these, 55% was 12-hydroxymethylfarnesoate and 45% was 10,11-epoxymethylfarnesoate. All of the chimeric proteins produced 10,11-epoxymethylfarnesoate and 10,11-epoxyfarnesol. C(78-82,F87L) produced 10,11-epoxymethylfarnesoate as the only metabolite of methylfarnesoate, with superior activity to CYP15A1 [70], whereas C(78-82,F87L,328-330) made 12-hydroxy-10,11-epoxymethylfarnesoate and its minor metabolite in a ratio of 1:0.43, similar to the ratio of metabolites of CYP4C7 of 1:0.54 (Table S11).

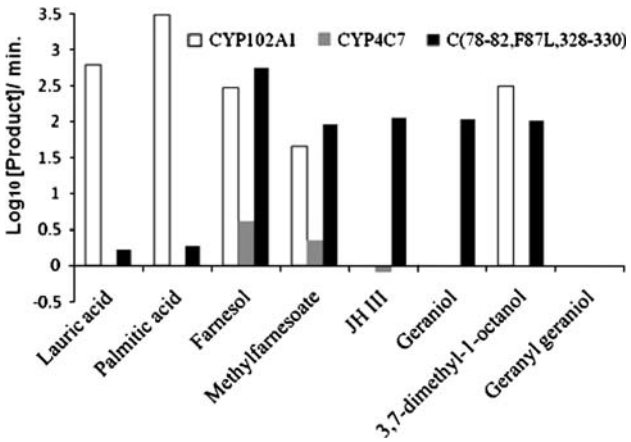


Fig. 5 The logarithm of the turnover number from Tables 2 and S13 (concentration of hydroxy and epoxy products formed per minute) by 1 μM CYP102A1 or chimeric protein reacting with 250 μM concentrations of each of the substrates of this study. Note that no detectable products were formed from geranylgeraniol as a substrate. *JH III* (juvenile hormone III (i.e., 10,11-epoxymethylfarnesoate)

In addition to farnesol, other related terpene alcohols, including the shorter geraniol and the longer geranylgeraniol (Structure 1), were also investigated as substrates of the chimeric enzymes of this study. The relative amounts of products obtained from metabolism of each substrate investigated in this study, by CYP102A1, CYP4C7, and the most active chimeric enzyme (C78-82,F87L,328-330), are summarized in Fig. 5. Table S12 shows that C(78-82, F87L), C(78-82,F87L,327-332), and C(78-82,F87L,328-330) acted as 6,7-epoxidases of geraniol. C(328-330) produced 6,7-epoxygeraniol, 8-hydroxygeraniol, and 10-hydroxygeraniol in a ratio of 1:2:0.69. C(327-328), C(328-330), C(78-82,F87L), C(78-82,F87L,327-332), and C(78-82,F87L,328-330) became 6-hydroxylases of 3,7-dimethyl-1-octanol, as did CYP102A1. The saturated analogue of geraniol, 3,7-dimethyl-1-octanol, produced only this single metabolite, as shown in Table S13. In contrast to the shorter terpene alcohols, the longer geranylgeraniol was essentially inactive (see Table S13) with all enzymes tested, including wild-type CYP102A1. At best, the geranylgeraniol activity of C(78-82,F87L) was 1% that of geraniol and 0.12% that of farnesol. Thus, the length of the terpene has a major effect.

In this work we have emphasized the role of SRS-5 of CYP102A1 (Fig. 1), of which residues 327–332 and 328–330 were investigated separately and in combination with the two most active chimeras from the second round of "scanning chimeragenesis," C(78-82,F87A) and C(78-82,F87L) [65], to create six new chimeric enzymes. Residues 328–330 of SRS-5, APA in CYP102A1 and VPL in our best chimera, are near the heme and F87, diagonally opposite the other SRS-1 residues, and thus form part of the substrate binding pocket over the heme as shown in

Fig. 6a with the long chain fatty acid palmitoleate bound. In Fig. 6b is shown the model structure of the chimera C(78-82,F87L,328-330), which shows a substantially smaller binding pocket, as would be expected for mutations in this chimera, which are mainly from smaller to larger residue side chains (three Ala to Leu or Val, a Val to Leu, and a Phe to Trp mutation). Of the mutations that result in smaller side chains, only one (F87L) increases the size of the binding pocket (R79F and D80P side chains do not point toward the substrate). It should be pointed out that the substrate farnesol, in Fig. 6b, is much closer to iron than is palmitoleate, Fig. 6a. The result of these mutations is that long chain fatty acid substrates cannot bind in the same conformation as that shown for palmitoleate bound to wild-type CYP102A1 (Fig. 6a), but shorter-chain substrates such as farnesol would bind comfortably, possibly as in the model shown in Fig. 6b. Unfavorable binding in a constricted pocket would provide a reasonable rationale for the much lower turnover number observed for turnover of palmitic acid to C(78-82, F87L,328-330), as compared with high turnover number observed with farnesol (Table 1). And although lauric acid is the same length as farnesol, the much lower turnover number observed for laurate suggests that this substrate has a preference for an extended conformation rather than the folded conformation of farnesol shown in the model in Fig. 6b.

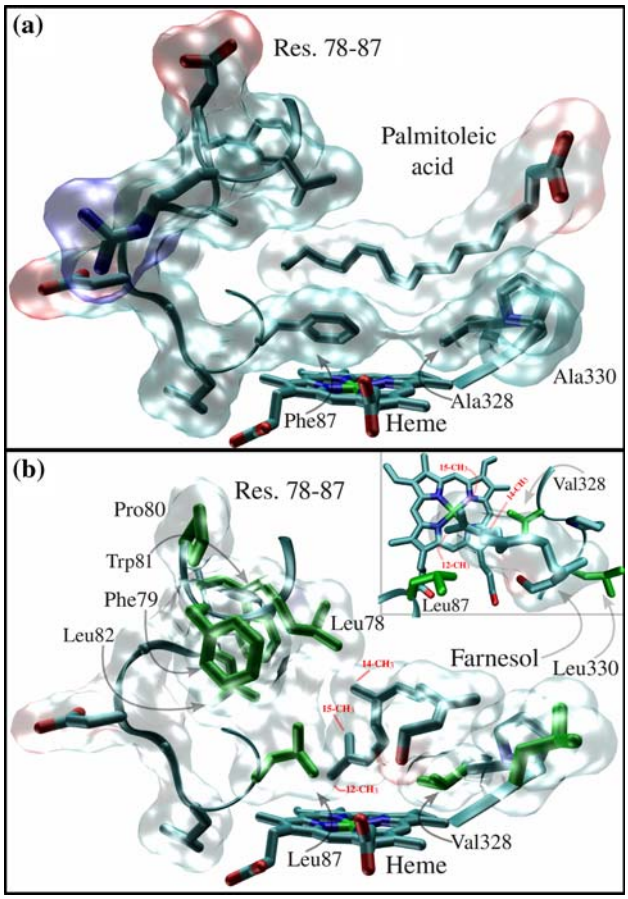


Fig. 6 Residues 78–87 and 328–330 of CYP102A1 and the chimeric enzyme C(78-82,F87L,328-330) substrate binding pocket, with the mutated residues of the chimera colored *green*. The substrate palmitoleate is shown bound to the wild-type enzyme (**a**) and the substrate farnesol is shown docking with the chimeric enzyme (**b**), with an *inset* view shown from above the heme and the substrate. The substrates and the amino acid side chains are shown in transparent space-filling mode. The wild-type structure of CYP102A1 is from the X-ray structure (PDB code 1FAG) and the chimera C(78-82,F87L,328-330) is an energy-minimized model (discussed in detail in the electronic supplementary material)

To rationalize the changes in the major products formed is even more challenging, because the crystal structure (and the model which is based on that structure) is of the resting state enzyme. As mentioned in "Introduction," NMR studies [30] have indicated that the conformation of the binding pocket of CYP102A1 changes upon reduction to the ferrous state, and that the change in conformation allows the fatty acid substrate to move some 6 Å closer to iron. Thus, the active state structure is not known and likely involves a number of different conformations for the substrate, in view of the number of different products produced by the wild-type enzyme. More details of the model and docking of substrates are provided in the electronic supplementary material. We infer from our experimental results that by changing the shape of the binding pocket from that in Fig. 6a to that in Fig. 6b, we have selected against long straight-chain fatty acids as substrates. We find that the activity of our best chimera, C(78-82,F87L,328-330), toward farnesol is nearly twice as great as that of wild-type CYP102A1. For transition from the resting state to the reduced enzyme, for which there is no structure available, we are left to conclude that just as the wild-type substrate binding pocket (with substrate relatively far from the iron active site) is able to undergo the changes in conformation necessary to bring the ω-1 to ω-3 carbons of the long-chain fatty acid within reactive reach of the ferryl heme site, the C(78-82,F87L,328-330) chimera is able to bring the $C12H_3$ of farnesol to that position so that 12-hydroxyfarnesol can form as the primary metabolite; however, the opposite $C15H_3$ and the neighboring 10,11 double bond are also able to react almost as favorably, as evidenced by the fact that 15-hydroxyfarnesol and 10,11-epoxyfarnesol are produced to about one half and one third the quantity of 12-hydroxyfarnesol, respectively (Fig. 5). All chimeras that contain the A328V and A330L mutations produce these three metabolites as major products, except F87A-containing mutants (Ala at position 87 is not native to CYP102A1 or CYP4C7), which all produce 9-hydroxyfarnesol and 6,7-epoxyfarnesol instead of 15-hydroxyfarnesol and 12-hydroxyfarnesol. Thus, during

enzyme turnover, the side chains of V328 (which in the model in Fig. 6b is partially between the heme and farnesol) and L330 protect the $C2H_2$ end of farnesol from reactive reach of the ferryl heme site.

In Fig. S10 we have provided the distances from the closest carbons of farnesol to the heme iron in the model shown in Fig. 6b; as can be seen, C-12 is only 3.75 Å from iron, whereas C-10 is 4.9 Å from iron. In contrast, C-9 is 6.2 Å from iron and C-15 is 5.6 Å from iron (distances not shown in Fig. S10), and farnesol would have to bind in a different conformation to bring C-15 close to iron, because of the 10,11 double bond. However, these distances may be misleading, considering the inexact science of homology modeling. Since the amount of 15-hydroxyfarnesol produced is substantial (21% for our best chimera), it is likely that there are alternative conformations for the substrate to that in Fig. 6b.

We also briefly investigated sections of the I-helix (SRS-4) as possible peptide fragments whose replacement might affect substrate selectivity and regiospecificity. We chose the region which involves those residues that are within 5 Å of the substrates palmitoleic acid and *N*-palmitolylglycine, I257–T268. Of the four chimeras prepared, C(255-263), C(259-274), C(259-263), and C(I259D;I263F), the two with the largest replacements failed to produce full-sized proteins, whereas the other two chimeric enzymes, C(259-263) and C(I259D;I263F), showed dramatic inhibition of catalytic turnover with laurate, palmitate, and farnesol as substrates [114]. Furthermore, no significant changes in the regiospecificity of oxidation of these substrates were detected [114]. These results suggest that for the substrates laurate, palmitate, and farnesol, the I-helix part of the catalytic site plays a mostly general structural role (because of its length and location above the 3-methyl, 4-vinyl edge of the heme), by generally blocking part of the distal pocket from the substrate, rather than one involving specific substrate binding interactions. Other mutants within the I-helix that have been reported by others include the A264E, A264H, and A264K mutants, which have been studied by Munro and coworkers [27, 42, 43]. It was shown that all three of those amino acid side chains can bind to iron, a demonstration of the close positioning of the I-helix to the heme. Another I-helix residue, T268 of CYP102Al, is the threonine present in the I-helix of many CYP enzymes. It is believed to be involved in stabilization of the oxy complex of the reduced enzyme, and possibly to participate in a proton shuttle mechanism [115–117]. The T268A mutant has been produced and shown to have much slower rates of O_2 and NADPH consumption, and its electron transfer is uncoupled from substrate hydroxylation, which indeed suggests the importance of this threonine to the mechanism of action of CYP102A1 [7].

In comparison with the chimeras created in this study, the crystal structure of CYP2E1, which also metabolizes farnesol to 12-hydroxyfarnesol [85], has a much smaller substrate binding pocket [99] than does CYP102A1 [24, 26], and although the structures are of two inhibitor complexes, those of 4-methylpyrazole (only six non-hydrogen atoms) and indazole (nine non-hydrogen atoms), which each bind to the heme through a heterocyclic nitrogen atom [99], it is unlikely that the substrate farnesol would require a significant expansion of the binding pocket. Emphasis was placed of the small size of the substrate binding pocket [99], but this has not yet been tested by crystallization and structure determination of the enzyme bound to any substrate. We aligned the protein sequence of CYP2E1 with that of CYP4C7 and with our third-round chimera C(78-82,F87L,328-330), as shown in Fig. S11. It is interesting to note that the residues of the β1–4 strand (SRS-5) of CYP4C7 which we mutated, VPL, are VPS for CYP2E1, with the valine in both cases being five residues after the –EXXR– motif [50]. Least-squares overlap of our farnesol-docked chimera C(78-82,F87L,328-330) with CYP2E1 (PDB file 3E4E) based on the heme position, using the program SwissPDB [111], but without any adjustments of the residues of CYP2E1, shows some similarity in the shape of the binding pocket, as shown in Fig. S12. It would be interesting to obtain a crystal structure of CYP2E1 bound to farnesol.

In conclusion, C(78-82,F87L) and C(78-82,F87L,328-330) clearly show the most complete change in substrate selectivity from fatty acids to farnesol, and both retain superior enzyme activity with respect to CYP102A1 approximately 5 times and approximately 2 times greater, respectively. The change in substrate selectivity is clearly shown in Fig. 4. In addition, C(78-82,F87L,328-330) produces 12-hydroxyfarnesol as the major metabolite, as does CYP4C7. All experimental results support the conclusion that scanning chimeragenesis has succeeded in changing bacterial monooxygenase CYP102A1 into an enzyme capable of carrying out the major reaction of insect terpenoid ω-hydroxylase CYP4C7, with a 100-fold increase in turnover number. The direct comparison of the logarithm of the product concentration per minute in Fig. 5 shows the dramatic difference between CYP102A1, CYP4C7, and C(78-82,F87L,328-330) with all substrates tested. On the basis of these results, scanning chimeragenesis clearly can be a useful method for producing new enzymatic products from CYP102A1. We also believe that scanning chimeragenesis could be used as a new systematic tool for changing substrate selectivity and regiospecificity among any two cytochromes P450 that have a common substrate, to study the interaction between enzymes and the substrate or to create new chimeric proteins for pharmaceutical and industrial uses [9–22].

Acknowledgments We thank Armand Fulco of the University of California Los Angeles for the original *E. coli* strain pbsBM-3 with the plasmid carrying the CYP102A1 gene, P. Shu of the MS Facility in the Department of Chemistry at the University of Arizona for GC–MS and GC–TOF-MS fragmentation analysis, and Mary T. Flores for assistance in preparing the NMR figures.

References

1. Miura Y, Fulco AJ (1974) Biochim Biophys Acta 249:1880–1888
2. Miura Y, Fulco AJ (1975) Biochim Biophys Acta 388:305–317
3. Narhi LO, Fulco AJ (1986) J Biol Chem 261:7160–7169
4. Fulco AJ (1991) Annu Rev Pharmacol Toxicol 31:177–203
5. Klein ML, Fulco AJ (1993) J Biol Chem 268:7553–7561
6. Klein ML, Fulco AJ (1994) Biochim Biophys Acta 1201:245–250
7. Yeom H, Sligar SG, Li H, Poulos TL, Fulco AJ (1995) Biochemistry 34:14733–14740
8. Murataliev MB, Feyereisen R, Walker FA (2004) Biochim Biophys Acta 1698:1–26
9. Capdevila JH, Wei S, Helvig C, Fulco JR, Belosludtsev Y, Truan G, Graham-Lorence SE, Peterson JA (1996) J Biol Chem 271:22663–22671
10. Munro AW, Leys DG, Mclean KJ, Marshall KR, Ost TW, Daff S, Miles CS, Chapman SK, Lysek DA, Moser CC, Page CC, Dutton PL (2002) Trends Biochem Sci 27:250–257
11. Gilardi G, Meharenna YT, Tsothou GE, Sadeghi SJ, Fairhead M, Giannini S (2002) Biosens Bioelectron 17:133–145
12. Fairhead M, Giannini S, Gillam EM, Gilardi G (2005) J Biol Inorg Chem 10:842–853
13. Dodhia VR, Fantuzzi A, Gilardi G (2006) J Biol Inorg Chem 11:903–916
14. Miles CS, Ost TW, Noble MA, Munro AW, Chapman SK (2000) Biochim Biophys Acta 1543:383–407
15. Li QS, Schwaneberg U, Fisher P, Schmitt J, Pleiss J, Lutz-Wahl S, Schmid RD (2001) Biochim Biophys Acta 1545:114–121
16. Li QS, Ogawa J, Schmid RD, Shimizu S (2001) Appl Environ Microbiol 67:5735–5739
17. Cowart LA, Falck JR, Capdevila JH (2001) Arch Biochem Biophys 387:117–124
18. Guengerich FP, MacDonald TL (1990) FASEB J 4:2453–2459
19. Guengerich FP, Shimada T (1991) Chem Res Toxicol 4:391–407
20. Guengerich FP (1995) Environ Health Perspect 103:25–28
21. Guengerich FP (2002) Nat Rev 1:359–366
22. Appel D, Lutz-Wahl S, Fischer P, Schwaneberg U, Schmid RD (2001) J Biotechnol 88:167–171
23. Li H, Poulos TL (1995) Acta Crystallogr D 51:21–32
24. Li H, Poulos TL (1997) Nat Struct Biol 4:140–146
25. Sevrioukova IF, Li H, Zhang H, Peterson JA, Poulos TL (1999) Proc Natl Acad Sci USA 99:1863–1868
26. Haines DC, Tomchick DR, Machius M, Peterson JA (2001) Biochemistry 40:13456–13465
27. Joyce MG, Girvan HM, Munro AW, Leys D (2004) J Biol Chem 279:23287–23293
28. Kuper J, Wong TS, Roccatano D, Wilmanns M, Schwaneberg U (2007) J Am Chem Soc 129:5786–5787
29. Modi S, Primrose WU, Boyle JMB, Gibson CF, Lian L-Y, Roberts GCK (1995) Biochemistry 34:8982–8988
30. Modi S, Sutcliffe MJ, Primrose WY, Lian L-Y, Roberts GCK (1996) Nat Struct Biol 3:414–417
31. Munro AW, Malarkey K, McKnight J, Thomson AJ, Kelly SM, Price NC, Lindsay JG, Coggins JR, Miles JS (1994) Biochem J 303:423–428
32. Graham-Lorence S, Truan G, Peterson JA, Falck JR, Wei S, Helvig C, Capdevila JH (1997) J Biol Chem 272:1127–1135
33. Oliver CF, Modi S, Sutcliffe MJ, Primrose WU, Lian LY, Roberts GC (1997) Biochemistry 36:1567–1572
34. Noble MA, Miles CS, Chapman SK, Lysek DA, Mackay AC, Reid GA, Hanzlik RP, Munro AW (1999) Biochem J 339:371–379
35. Loughran PA, Roman LJ, Aitken AE, Miller RT, Masters BSS (2000) Biochemistry 39:15110–15120
36. Kumar S, Halpert JR (2005) Biochem Biophys Res Commun 338:456–464
37. Ost TWB, Miles CS, Murdoch J, Cheung Y-F, Reid GA, Chapman SK, Munro AW (2000) FEBS Lett 486:173–177
38. Carmichael AB, Wong L-L (2001) Eur J Biochem 268:3117–3125
39. Ost TWB, Miles CS, Munro AW, Murdoch J, Reid GA, Chapman SK (2001) Biochemistry 46:13421–13429
40. Chen Z, Ost TWB, Schelvis JPM (2004) Biochemistry 43:1798–1808
41. Sulistyaningdyah WT, Ogawa J, Li Q-S, Shinkyo R, Sakaki T, Inouye K, Schmid RD, Shimizu S (2004) Biotechnol Lett 26:1857–1860
42. Girvan HM, Marshall KR, Lawson RJ, Leys D, Joyce MG, Clarkson J, Smith WE, Cheesman MR, Munro AW (2004) J Biol Chem 279:23274–23286
43. Girvan HM, Seward HE, Toogood HS, Cheesman MR, Leys D, Munro AW (2007) J Biol Chem 282:564–572
44. Neeli R, Roitel O, Scrutton NS, Munro AW (2005) J Biol Chem 280:17634–17644
45. Kumar S, Sun L, Liu H, Muralidhara BK, Halpert JR (2006) Protein Eng Des Sel 19:547–554
46. Kumar S, Liu H, Halpert JR (2006) Drug Metab Dispos 34:1958–1965
47. Sun L, Chen CS, Waxman DJ, Liu H, Halpert JR, Kumar S (2007) Arch Biochem Biophys 458:167–174
48. Zhao Y, Halpert JR (2007) Biochim Biophys Acta 1770:402–412
49. Stjernschantz E, van Vugt-Lussenburg BMA, Bonifacio A, de Beer SBA, van der Zwan G, Gooijer C, Commandeur JNM, Vermeulen NPE, Oostenbrink C (2008) Proteins Struct Funct Bioinform 71:336–352
50. Seifert A, Pleiss J (2009) Proteins Struct Funct Bioinform 74:1028–1035
51. Shao Z, Arnold FH (1996) Curr Opin Struct Biol 6:513–518
52. Joo H, Lin Z, Arnold FH (1999) Nature 399:670–673
53. Li QS, Schwaneberg U, Fisher P, Schmid RD (2000) Chem Eur J 6:1531–1536
54. Arnold FH, Wintrode PL, Miyazaki K, Gershenson A (2001) Trends Biochem Sci 26:100–106
55. Farinas ET, Schwaneberg U, Glieder A, Arnold FH (2001) Adv Synth Catal 343:601–606
56. Glieder A, Farinas ET, Arnold FH (2002) Nat Biotechnol 20:1135–1139
57. Salazar O, Cirino PC, Arnold FH (2003) Chembiochem 4:891–893
58. Meinhold P, Peters MW, Chen MM, Takahashi K, Arnold FH (2005) Chembiochem 6:1765–1768
59. Otey CR, Bandara G, Lalonde J, Takahashi K, Arnold FH (2006) Biotechnol Bioeng 93:494–499
60. Peters MW, Meinhold P, Glieder A, Arnold FH (2003) J Am Chem Soc 125:13442–13450
61. Kumar S, Chen CS, Waxman DJ, Halpert JR (2005) J Biol Chem 280:19569–19575

62. Fairhead M, Giannini S, Gillam EMJ, Gilardi G (2005) J Biol Inorg Chem 10:842–853
63. Dodhia VR, Fantuzzi A, Gilardi G (2006) J Biol Inorg Chem 11:903–916
64. Murataliev MB, Trinh LN, Moser LV, Bates RB, Feyereisen R, Walker FA (2004) Biochemistry 43:1771–1780
65. Chen C-KJ, Shokhireva TK, Berry RE, Zhang H, Walker FA (2008) J Biol Inorg Chem 13:813–824
66. Sutherland TD, Unnithan GC, Andersen JF, Evans PH, Murataliev MB, Szabo LZ, Mash EA, Feyereisen R (1998) Proc Natl Acad Sci USA 95:12884–12889
67. Baker FC, Mauchamp B, Tsai LW, Schoolery DA (1983) J Lipid Res 24:1586–1594
68. Bede J, Goodman WG, Tobe S (1999) Pure Appl Chem 70:1–9
69. Yu SJ (2000) Zool Stud 39:243–249
70. Helvig C, Koener JF, Unnithan GC, Feyereisen R (2004) Proc Natl Acad Sci USA 101:4024–4029
71. Raner GM, Muir AQ, Lowry CW, Davis BA (2002) Biochem Biophys Res Commun 293:1–6
72. Barchuk AR, Maleszka R, Simoes LP (2004) Insect Mol Biol 13:459–467
73. Boddupalli SS, Pramanik BC, Slaughter CA, Estabrook RW, Peterson JA (1992) Arch Biochem Biophys 292:20–28
74. Meigs TE, Simoni RD (1997) Arch Biochem Biophys 345:1–9
75. Hampton RY, Bhakta H (1997) Proc Natl Acad Sci USA 94:12944–12948
76. Correll CC, Ng L, Edwards PA (1994) J Biol Chem 269:17390–17393
77. Beigs TE, Roseman DS, Simoni RD (1996) J Biol Chem 271:7916–7922
78. Yazlovitskaya EM, Melnykovych G (1995) Cancer Lett 88:179–183
79. Haug JS, Goldner CM, Yazlovitskaya EM, Voziyan PA, Melnykovych G (1994) Biochim Biophys Acta 1223:133–140
80. Miquel KA, Pradines A, Terce F, Selmi S, Favre G (1998) J Biol Chem 273:26179–26186
81. Roullet JB, Luft UC, Xue H, McCarron DA (1997) J Biol Chem 272:32240–32246
82. Rollet JB, Xue H, Chapman J, McDougal PC, Roullet CM, McCarron DA (1996) J Clin Invest 97:2384–2390
83. Rollet JB, Spaetgens RL, Burlingame T, Feng ZP, Zamponi GW (1999) J Biol Chem 274:25439–25446
84. Raner GM, Muir AQ, Lowry CW, Davis BA (2002) Biochem Biophys Res Commun 293:1–6
85. DeBarber AE, Bleyle LA, Roullet J-BO, Koop DR (2004) Biochim Biophys Acta 1682:18–27
86. Laethem RM, Balazy M, Falck JR, Laethem CL, Koop DR (1993) J Biol Chem 268:12912–12918
87. Roy U, Joshua R, Stark RL, Balazy M (2005) Biochem J 390:719–727
88. Guengerich FP, Kim DH, Iwasaki M (1991) Chem Res Toxicol 4:168–179
89. Koop DR, Coon MJ (1986) Alcohol Clin Exp Res 10:445–495
90. Koop DR (1986) Mol Pharmacol 29:399–404
91. Tasaneeyakul W, Birkett DJ, McManus ME, Tassaneeyakul W, Veronese ME, Andersson T, Tukey RH, Miners JO (1994) Biochem Pharmacol 47:1767–1776
92. Peter R, Bocker R, Beaune PH, Iwasaki M, Guengerich FP, Yang CS (1990) Chem Res Toxicol 3:566–573
93. Nanji AA, Zhao S, Sadrzadeh SM, Dannenberg AJ, Tahan SR, Waxman DJ (1994) Alcohol Clin Exp Res 18:1280–1284
94. Barnett CR, Rudd S, Flatt PR, Ioannides C (1993) Biochem Pharmacol 45:313–319
95. Raucy JL, Lasker JM, Kraner JC, Salazar DE, Lieber CS, Corcoran GB (1991) Mol Pharmacol 39:275–280
96. Johansson J, Ekstrom G, Scholte B, Puzycki D, Iornvall H, Ingelman-Sundberg M (1988) Biochemistry 27:1925–1934
97. Patten CJ, Thomas PE, Guy RL, Lee M, Gonzalez FJ, Guengerich FP, Yang CS (1993) Chem Res Toxicol 6:511–518
98. Raucy JL, Lasker JM, Lieber CS, Black M (1989) Arch Biochem Biophys 271:270–283
99. Porubsky PR, Meneely KM, Scott EE (2008) J Biol Chem 283:33698–33707
100. http://www.accelrys.com/products/gcg/
101. Pearson WR (1996) Methods Enzymol 266:227–258
102. Gotoh O (1992) J Biol Chem 267:83–90
103. Murataliev MB, Feyereisen R (1996) Biochemistry 35:15029–15037
104. Omura T, Sato R (1964) J Biol Chem 239:2379–2385
105. O'Keeffe DH, Ebel RE, Peterson JA (1978) Methods Enzymol 52:151–157
106. Nicolaides N, Soukup VG, Ruth EC (1983) Biomed Mass Spectrom 10:441–449
107. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comput Chem 19:1639–1662
108. Morris GM, Huey R, Hart WE, Lindstrom W, Gillet A, Goodsell DS, Olson AJ (1991–2007) AutoDock 4.00(c). http://autodock.scripps.edu/
109. Sanner MF (1999) J Mol Graph Model 17:57–61
110. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14:33–38. http://www.ks.uiuc.edu/Research/vmd
111. Guex N, Peitsch MC (1997) Electrophoresis 18:2714–2723. http://www.expasy.org/spdbv/
112. van Gunsteren WF et al (1996) Biomolecular simulation: the GROMOS96 manual and user guide. Vdf Hochschulverlag ETHZ. http://www.igc.ethz.ch/GROMOS
113. Vogeli U, Von Philipsborn W (1975) Org Magn Res 7:617–627
114. Chen C-KJ (2007) PhD dissertation. University of Arizona
115. Imai M, Shimada H, Watanabe Y, Matsushima-Hibiya Y, Makino R, Koga H, Horiuchi T, Ishmura Y (1989) Proc Natl Acad Sci USA 86:7823–7827
116. Martinis SA, Atkins WM, Stayton PS, Sligar SG (1989) J Am Chem Soc 111:9252–9253
117. Raag R, Martinis SA, Sligar SG, Poulos TL (1991) Biochemistry 30:11420–11429