Protein engineering of the cytochrome P450 monooxygenase from *Bacillus megaterium*

By Vlada B. Urlacher and Rolf D. Schmid

Institut of Technical Biochemistry, University of Stuttgart

Allmandring 31

D-70569 Stuttgart

Germany

Tel: +49-711-685-3192

Fax: +49-711-685-3196

e-mail: itbvkha@po.uni-stuttgart.de

e-mail: rolf.d.schmid@rus.uni-stuttgart.de

Table of contents

Introduction

The role and importance of cytochrome P450 enzymes (CYP) in drug development, biodegradation processes and biocatalysis has been widely acknowledged. P450 monooxygenases exhibit an extremely wide substrate spectrum which is the basis of their ability to activate or detoxify a large variety of target molecules. P450 monooxygenases have been isolated from bacteria, yeasts, insects, as well as mammalian and plant tissues. Currently, the enzyme family is one of the best known gene subfamilies with over 1000 characterized members (http://drnelson.utmem.edu/CytochromeP450.html).

Many studies have been dedicated to structural models of cytochrome P450 in order to improve our understanding of the mechanistic details of the enzymes,[1,2] their substrate specificity and their pronounced stereo- and regiospecificity.[3] In addition, homology modeling of mammalian P450s[4-6] and QSAR analyses using chemicals which are metabolized by P450s,[7] have added considerably to our understanding of the metabolic variations and functions of the enzyme.

Cytochrome P450 enzymes are of considerable interest to pharmaceutical and chemical industry and have thus become targets for protein engineering approaches. Protein engineering is generally defined as the modification of an enzyme by site-directed or random mutagenesis with the aim of altering its properties. Rational design requires a solid structural basis and profound knowledge of the catalytic mechanism of the enzyme which was provided by determining the structures of CYPs using X-ray crystallography at high-resolution. Ten of the twelve[8-17] crystallized cytochrome P450s are of prokaryotic origin and water-soluble. From a technical point of view, microbial P450s are easier to handle than P450 enzymes from plants and animals. They are not membrane-associated and exhibit a relatively high stability. Eukaryotic cytochrome P450 enzymes are membrane-associated proteins and are hence more difficult to crystallize. Currently, only the X-ray structures of two membrane-bound

mammalian P450s, rabbit CYP2C5[18] and human CYP2C9[19] are known. Models of other mammalian P450s were built based on the structure of CYP2C5[20] and its bacterial analogues.[21]

P450cam, the cytochrome P450 monooxygenase from *Pseudomonas putida,* is the best characterized microbial P450 enzyme.[1,22-25] In the last fifteen years, a large number of other soluble prokaryotic P450 enzymes have been identified, isolated, subcloned in *Escherichia coli*, overexpressed and characterized.[10-17]

Cytochrome P450 BM-3 from *Bacillus megaterium* is catalytically self-sufficient. It contains a P450-heme domain of 54 kDa and an FAD/FMN-reductase domain of 64kDa on a single polypeptide chain.[26] The enzyme catalyzes the subterminal oxidation of saturated and unsaturated fatty acids with a chain length of 12 to 20 carbons. High-resolution X-ray crystal structures are available for substrate-free,[9] palmitic acid-bound[27] and N-palmitoylglycine-bound[28] wild-type and mutant[29] P450 BM-3 enzymes. The structure resolved by NMR is also available.[30] The well-known structure, the availability of the CYP102A1 gene which encodes the protein, and the possibility of expressing the protein in *E. coli* have encouraged a number of research groups to undertake site-directed mutagenesis studies in order to identify key amino acids. Insights into the mechanisms of P450 BM-3 have been gained and transferred to eukaryotic P450 enzymes. Since P450 BM-3 is an excellent model for addressing questions on the wide substrate specificity of P450s in general and techniques involving the mutagenesis of P450 BM-3 have led to a variety of biocatalysts with features of industrial interests. This review will summarize the recent research on this particular P450 enzyme.

Isolation and expression of P450 BM-3

The CYP102 gene from *B. megaterium*, coding for a heme domain and an FAD/FMN reductase, was cloned employing Southern blot analysis.[31] Initially, protein expression was performed in *E. coli* cells harboring the recombinant plasmid (pUC13 derivative) under the control of the original *B. megaterium* promoter.[31] A 5 kb DNA fragment containing the CYP102 gene was sequenced. One open reading frame was found whose sequence could be matched to the amino acid sequences of P450 and reductase domains, generated by trypsin digestion.[32]

As the DNA sequence of CYP102A1 is known, it can be usually isolated directly from genomic DNA of *Bacillus megaterium* by polymerase chain reaction (PCR) using primers that introduce appropriate restriction sites upstream of the ATG (methionine) and downstream of the stop codon.

The overexpression of heterologous proteins in *E. coli* requires a strong promoter and an efficient ribosome-binding site at an optimal distance from the first methionine codon. The *lac* promotor was used to subclone CYP102A1 into pUC-derivatives.[31] *In vivo* promoter activity is efficiently controlled by the *lac* repressor protein which is encoded by the *lacI* gene that is located inside the *lac* operon. Transcription can be induced by isopropyl thio-β-D-galactosidase (IPTG). Other expression systems were also successfully used for the recombinant expression of P450 BM-3. Many mammalian cytochrome P450 enzymes could be expressed in *E. coli* using the pCWORI(+) vector (a derivative of plasmid pHSe5).[33-35] Therefore, the use of this vector was also tested with P450 BM-3.[36] The transcription/translation region of pCWORI(+) contains a *lacUV5* promoter and two copies of a *tac* promoter. The pGLW11 vector (a derivative of pK223), using a *tac* promoter was also successfully applied for the expression of P450 BM-3 in *E. coli*.[37]

In our laboratory two *E. coli* expression vectors were used for the recombinant expression of the P450 enzyme from *Bacillus megaterium*. Standard methods for the manipulation of DNA were employed.[38] The CYP102A1 gene was isolated by PCR using genomic DNA of *Bacillus megaterium* (DSM 32T) as a template, and primers introducing a *Bam*HI site at the N-terminus and an *Eco*RI site at the C-terminus of the gene. This fragment was used for the subsequent subcloning into the pCYTEXP1 vector, resulting in the plasmid pT-USC1BM3.[39] The pCYTEXP1 vector contains the bacteriophage tandem promoters P_R and P_L which are preceded by the *clts857* repressor gene, and the transcription terminator from the *fd* bacteriophage. Plasmid DNA was transformed using $CaCl_2$-treated *E. coli* DH5α cells.[38] Expression was induced by a temperature shift from 37° to 42° for 5 hours, yielding 300-350 nmol of CO-reactive P450 per liter of cell culture. When σ-aminolevulinic acid (1 m*M* final concentration) is added to the growing cells after induction a P450 yield of up to 500 nmol per liter broth can be reached. P450 BM-3 mutants were expressed in the same manner as a wild-type enzyme.

P450 BM-3 and its mutants could also be expressed fused to a His_6-tag in a pET system. The pET system has a strong bacteriophage T7-promoter and is one of the most popular expression systems. Under the control of a T7-promoter the target gene can only be transcribed by phage T7-DNA-polymerase which is only present in specially engineered *E. coli* strainssuch as BL21 or BL21 (DE3).

The gene encoding P450 BM-3 was amplified from pT-USC1BM-3[39] by PCR using primers which were specifically designed to facilitate the cloning of the gene into pET28a(+) vector between *Bam*HI *and Eco*RI restriction sites .[40] The initial cloning was performed in DH5α *E. coli* cells which have a high transformation efficiency and give an excellent plasmid yield. Subsequently, the gene was expressed in strain BL21 (DE3).

For the high-level expression of active P450 the reaction was induced by the addition of isopropyl thio-β-D-galactosidase (IPTG) to a final concentration of 0,5 m*M*, and the cells grown at 30° and 120 rpm. After 4-5 hours the cells were harvested by centrifugation (20 min, 6000 rpm, 4°). After cell disruption and centrifugation, 400-500 nmol of cytochrome oxidase-reactive P450 per liter of cell culture was obtained without the addition of σ-aminolevulinic acid prior to centrifugation.

The advantage of this expression system is the high expression level of heterelogous protein. Upon optimization of the expression protocol (lower incubation temperature: 25-30°, or lower concentration of the inductor IPTG), the number of inclusion bodies, which are usually observed for mammalian P450s in a pET system, can thus be reduced drastically. One should mention that stirring speed influences the expression level of active P450 BM-3 and hence is a parameter that must not be neglected.

The roles of key amino acids and active-site residues

There are several residues in the P450 BM-3 which are believed to be important for the catalytic reaction of the enzyme. The role of these residues and their manifold effects have been examined by structural analysis and molecular modeling followed by site-directed mutagenesis and functional characterization. The information gained adds to our understanding of the regio-, stereo- and chemoselectivity of the oxidation process of the enzyme, the way water accesses its active site, electron transport and its redox state.

The active site of P450 BM-3 consists of a long, hydrophobic channel, extending from the heme to the protein surface.[27] Comparison of substrate-free[9] and substrate-bound[27] crystal structures and site-directed mutagenesis studies points to the important role of the amino acid residues R47 and Y51 that are located at the entrance of the active center. These two positions interact with the carboxylate of the fatty acid and are thus crucial for the proper

positioning of the substrates. Mutagenesis experiments at residue R47 confirmed the important contribution the guanidinium group of arginine has on enzyme activity.[41,45] Although R47E, R47A[41,46,47] and R47G[47] mutants retained their activity towards C12-C16 fatty acids, their k_{cat}/Km values are 5-15-fold lower than those of the wild-type enzyme. The combination of R47L/Y51F increases the oxidation activity of phenanathrene, fluoranthene and pyrene up to 40-fold.[37] However, the substitution at position 51 has less impact on enzyme activity than that at position 47.[47]

The phenylalanine residue at position 87 is highly conserved. It is located in the active site of the protein and very important for the correct orientation of the fatty acid hydrocarbon chain.[30] Comparison of substrate-free and substrate-bound crystal structures of P450 BM-3 revealed a substantial conformational difference that is caused by the phenyl ring of phenylalanine.[9,27] Mutations of P450 BM-3 at position 87 can affect its activity, and stereo- or regioselectivity.[48] An unfavorable substitution at position F87 can lead to irreversible conformational changes during catalytic turnover which will then result in a decrease or loss of catalytic competence.[47] In some cases, the combination of a mutation at position 87 with other mutations, in and outside the active site, has revealed a surprisingly strong effect on substrate selectivity [37,49,50] and peroxygenase activity of the enzyme.[36]

It is assumed that the threonine residue at position 268, located in the distal I helix plays an important role in oxygen binding and activation. Enzymatic properties and the crystal structure of the heme domain of the W268A mutant were determined using sodium laurate as a substrate. The mutant exhibited slower rates of NADPH and oxygen consumption and much higher uncoupling rates of electron transfer and substrate hydroxylation.[29]

High-resolution crystal analysis with subsequent site-directed mutagenesis also helped clarify the position and important mechanistic role of the solvent. The crystal structure of a complex between P450 BM-3 and N-palmitoylglycine at a resolution of 1.65 Å revealed

features of the active site which had not been previously determined.[28] The binding of the substrate leads to a conformational change, resulting in a shift of the A264 carbonyl away from the heme iron. The pivotal water molecule in the active site is thus brought in close vicinity of the heme group and can fill the sixth coordination site of the heme iron.

Interactions of the monooxygenase and reductase domains and electron transport

Protein-protein interactions are of great importance in the cytochrome P450 system. They provide essential electron functions and control the rate-limiting electron transfer.

Potential roles of specific residues in the heme domain of P450 BM-3 as well as in the reductase domain relate to specific affinity interactions, electrostatic charge and the direct electronic coupling of redox centers. Earlier investigations proposed a similar docking of redox partners (ferredoxin or flavin reductase) in the prokaryotic and eukaryotic P450 systems at a proximal part of the heme. Involvement of both electrostatic and hydrophobic protein-protein interactions has been demonstrated.[9,10,12,51] However, no clear evidence exists for the direct influence of certain amino acid residues on the electron transfer between P450 and its redox partner.

The reductase domain of P450 BM-3 is very similar to the microsomal NADPH-dependent cytochrome P450 reductase. However, there is a significant difference in the reduction mechanism of the heme iron.[52] In the mammalian enzyme the fully reduced flavin (FMNH2) is the electron donor to the heme iron, whereas in the case of the P450 BM-3 reductase domain the 1-electron donor is reduced semiquinone (FMNH) . Based on the analysis of sequence similarities between the P450 BM-3 reductase and those flavoproteins whose three-dimensional structures were already known, binding sites for FMN, FAD and NADPH were suggested. The crystal structure of the complex between the heme and the FMN-binding domain (2.3 Å)[53] shows that the flavin domain is located at the proximal end of the heme

domain. The region between P382-Q387 is assumed to be involved in transferring electrons from the FMN to the heme iron (electron transfer pathway through the polypeptide chain). Site-directed mutagenesis was applied to validate this observation. Cysteine residues were introduced at positions 104 and 387 which were expected to be responsible for the interactions between the two domains. Since the position 372 is located on the other side of the heme domain and so can not be involved in the electron transfer, glutamate at this position was substituted through cysteine and served as a control. The cysteine residues were subsequently modified with a bulky sulfhydryl reagent (DC modification) in order to prevent close protein interactions between the FMN domain and heme domain. In addition, the cysteine residue at position 156 was substituted by alanine.

Procedure. Mutations E372C and Q387C were introduced in a first PCR with modified 5'-primers which carried the desired mutations. An oligonucleotide corresponding to the carboxyl terminus of the heme domain was used as a 3'-primer. The mutations L104C and C156S were introduced accordingly, but the 3'-primers were modified. In a second PCR, 0.2-0.5 kb fragments served as 5'- or 3'-primers. The 1.4 kb fragments were cloned into the pNEB vector. To create double and triple mutants, the corresponding regions of the gene were excised with endonucleases and exchanged.[53]

Spectral analyses, laser flash photolysis experiments and DC modification of the mutants have shown that C387 has no significant effect on the electron transport from the FMN to the heme iron. In contrast, it is the C104 residue and surrounding area which are most critical for the docking of the redox partner.

The tryptophan residue at position 96 in the heme domain was also assumed to have an effect on the electron transfer from the FMN to the heme.[54] Substitutions of this tryptophan residue by alanine, phenylalanine or tyrosine caused a lower heme content of P450 BM-3 while the levels of catalytic activity remained unaltered. This indicates that W96 plays a role

in the association of the heme's prosthetic group and is probably required for an efficient redox interaction between the heme and flavin domains.

Other residues such as W574, W536 and G570 of the reductase domain were ruled out as being crucial for FNM binding. Only the tryptophane at position 574 has a minor effect on the electron transport[55].

Redox potential of the P450 heme iron

The analysis of the structure of P450 BM-3 and the characterization of mutants demonstrated the important role of phenylalanine at position 393 in controlling the reduction potential of the P450 heme iron.[42] This position is one of very few highly conserved amino acid residues of P450 enzymes. Changing this residue obviously affects the catalytic properties of the enzyme. Unfortunately, no details are available which would explain this effect. F393Y, F393A and F393H mutants were expressed in *E. coli*. No effect of F393 on stabilising the heme was observed. The analysis of the oxidation products of myristic acid generated by the F393H mutant revealed the same product proportions as for the wild-type enzyme. Nevertheless, a lower turnover rate of the mutant was observed. The crystal structure of the mutant and spectroscopic analysis suggested the thermodynamic control of F393 over the heme iron: this position seems to establish the equilibrium between the rate of heme reduction and the rate at which the ferrous heme can bind and, subsequently, reduce molecular oxygen.[56]

Artificial P450 systems using P450 BM-3 domains

As a natural fusion protein, P450 BM-3 is the catalytically most efficient P450 enzyme that is currently known. P450 BM-3 does not require the additional expression and purification of the redox partners. Both P450 BM-3 domains were used to create artificial fusion proteins.

The aim of such projects was the creation of highly active and soluble P450 enzymes that are applicable on an industrial scale.

Non-physiological partners such as flavodoxin from *Desulfovibrio vulgaris* and the reductase domain of P450 BM-3 have been used for constructing such fusion proteins.[57] The artificial combination of *Desulfovibrio vulgaris* flavodoxin and the heme domain of P450 BM-3 was made possible through the PCR-based introduction of an additional *Nla*III restriction site at the 3'-terminus of the P450 BM-3 heme domain and at the 5'-end of the flavodoxin gene. The artificial genes were expressed in *E. coli.* Electrochemical experiments demonstrated the feasibility of this assembled protein.[57]

The solubility of microsomal P450 enzymes has also been addressed by fusing the human P450 2E1 gene with the P450 BM-3 reductase gene. The membrane-bound N-terminus of the human enzyme was replaced by that of the P450 BM-3 heme domain.[57] The P450 2E1 reductase of the artificial BM-3 protein was constructed in two steps. The first 80 N-terminal residues of P450 2E1 were replaced by the first 54 amino acid residues of the heme domain of P450 BM-3. The modified P450 2E1 was then fused with the reductase domain. After cloning into the pT7 vector, the fusion protein was expressed in a soluble form in BL21 (DE3) CL *E. coli* cells using the standard protocol and purified on DEAE sepharose. After ultracentrifugation of the cell lysate at 100000 g, the protein still remained in the soluble fraction.

Changing the properties of P450 BM-3 by mutagenesis

"Rational evolution"

Combination of computer-assisted protein modeling with methods of directed evolution allows to improve significantly the efficiency of the search for enzyme variant with new properties. In a procedure termed "rational evolution" (protein design combined with directed

evolution), substrate specificity could be shifted from fatty acids with 12 carbons to those with 10 and 8 carbons. In these experiments, the fatty acid pseudosubstrates 10- and 8-*p*-nitrophenoxycarboxylic acid were used.[44]

Substrate docking was examined on the basis of the crystallographically determined structure of the palmitate-bound P450 BM-3 which was obtained from the Protein Data Bank. A model of 8-*p*NCA was used as a substrate molecule. The chemical structure of 8-*p*NCA and the mutations in the binding pocket were deduced using the biopolymer tool of SYBYL. The substrate's C1-C4 atoms were placed in analogy of the C6-C9 atoms of palmitic acid. The C7 and C8 atoms of 8-*p*NCA were placed at a distance of 4Å and 3.6Å from the heme iron. This distance was previously determined by NMR for P450 BM-3-laurate and 12-bromolaurate complexes. The *p*-nitrophenoxy group was placed manually into the binding pocket. Eight sites of the binding pocket were selected for further saturation mutagenesis: V26, R47, S72, A74, F87, L188 and M354 (Fig.1).

Saturation mutagenesis and recombination of mutations

Saturation mutagenesis describes the substitution or insertion of codons encoding all possible amino acids at any predetermined position in a gene.

PCR was utilized to produce the first library of P450 BM-3 mutants rapidly. Mutagenic PCR was performed using a slightly modified protocol from Stratagene Quik-Change Kit which allows the introduction of the nucleotide exchange which is necessary to obtain all possible codons.

Mutant F87A was used as template DNA. After saturation mutagenesis, subsequent ligation into the pCYTEXP1 vector and protein expression in *E. coli* strain DH5*a*, the mutant library was tested with different *p*NCA substrates with a chain length of fewer than 15 carbons. Compared to standard assays that are based on the consumption of NADPH or oxygen, the

*p*NCA assay is substrate-specific and allows the detection of the reaction product *p*-nitrophenolate. ω-Hydroxylation of *p*NCA by P450 BM-3 first leads to the formation of an instable hemiacetal intermediate, which then dissociates into ω-oxocarboxylic acid and yellow *p*-nitrophenolate. A pH of 8.2 is preferable since 92% of all *p*-nitrophenolate is deprotonated at this pH and thus exhibits a yellow color. In addition, P450 BM-3 remains active. The spectral absorption of *p*-nitrophenolate was measured at 410 nm, the extinction coefficient being ε= 13200 $M^{-1}cm^{-1}$. As *p*NCA is not easily dissolved in water, small amounts (1%) of organic solvent (DMSO or acetone) should be added to the reaction mixture.[44]

Synthesis of pNCA. Synthesis of *p*NCA is a multiple-step process. In the first three steps, the corresponding ω-bromocarboxylic acids are esterified. This is followed by an *sn*2 reaction with sodium *p*-nitrophenolate. The last step, the hydrolysis of the esters, can be done chemically or enzymatically using a lipase.[39]

p-NCA assay in 1 ml scale. 8 µl of 6 nmol 10- or 8-*p*NCA or 12-*p*NCA, dissolved in DMSO (1%) are added to 892 µl Tris/HCl buffer (100 m*M*, pH 8.2) in 1 ml cuvette. The mixture is incubated for 5 min at RT before the reaction starts with the addition of 100 µl of an aqueous solution of 1 m*M* NADPH. The amount of *p*-nitrophenolate formed in the reaction is determined photometrically at 410 nm.[39]

p-NCA assay in the microtiterplates . 200 µl containing 18 nmol of 10- and 11-*p*NCA, 12 nmol 12-*p*NCA, or 10 nmol of 15-*p*NCA dissolved in 2 µl DMSO (acetone) is placed into a 96-well microtiterplate. The reaction is started by adding 20 µl of 1 m*M* NADPH solution to each well. The procedures are performed by an automated workstation. The amount of *p*-nitrophenolate formed in each well is determined with a microtiterplate reader.[39]

Mutants with the highest activity towards 8-*p*NCA and 10-*p*NCA were cultivated again and their specific activities calculated in terms of P450 concentrations. The nucleotide sequences of the mutants were analyzed in order to determine the respective mutation. Complete

structural and kinetic characterizations of the expressed and purified enzymes provide important information for the subsequent recombination experiments. As a result, the mutations V26T, R47F, S72G, A74G, F87A, L188K, M354T were selected. Recombination of corresponding positive mutations was carried out by subsequent site-directed mutagenesis All mutant enzymes had a full complement of flavin and heme which indicated that the mutations did not disrupt or change the tertiary structure. After each mutagenic step, a structural model of a new mutant was designed according to which the generalized effect on substrate selectivity was studied. F87A is regarded as a key substitution that affects substrate specificity of P450 BM-3 by altering the contact between substrate and heme iron. The combination of L188K and F87A results in the formation of a new carboxylate-binding site with increased activity towards 12- and 10-*p*NCA. The A74G substitution obviously plays an important role in altering the chain-length specificity of P450 BM-3. Residue A74 is located at the N-terminus of the α-helix B and the side chain of alanine interacts sterically with the neighboring residues. If alanine is replaced by glycine at this particular position and phenylalanin is substituted with alanine at position 87, this increases the size of the binding site and so enables the hydroxylation of shorter chain fatty acids with the bulky *p*-nitrophenoxy group. It is well known that R47 is important for an efficient catalysis since it builds a carboxylate-salt bridge with a fatty acid and contributes to its binding. The addition of R47F to the triple mutant increased the k_{cat} for 10-*p*NCA and 8-*p*NCA. With this mutant we sought to understand whether preventing the original carboxylate binding and thus an increased hydrophobicity at the entrance of the binding pocket enables the diffusion of shorter chain substrates further into the active site. It was noticed that the introduction of M354T, V26T or S72G mutations to (F87A)LAR decreased k_{cat} values for shorter chain acids. Active mutants also showed activity towards the free C-8 and C-10 fatty acids,[58] and also towards the unnatural substrates naphthalene, n-octane, and 8-methylquinoline.[59]

Activity against indole

The rational evolution of P450 BM-3 towards a catalyst for the hydroxylation of shorter-chain fatty acids also revealed some colonies that produced a blue, water-soluble pigment. The high performance thin-layer-chromatography detected a rapidly moving blue and slower moving red component. Mass spectrometry and NMR analysis confirmed the presence of indigo and indirubin in the mixture.[60]

Sequence analysis revealed that all mutants producing this pigment contained mutations at either one or several of the three positions A74, F87 or L188. A strategy for the subsequent site-specific randomization of each site, starting with the best mutant from the previous mutagenesis step was developed. The best mutant which had a 10-fold higher activity towards indole than the wild-type was the A74G, F87V, L188Q mutant.

Activity assay for indole hydroxylation. The mixture of indole solution in DMSO and enzyme in the appropriate amount of 0,1 *M* Tris/HCl buffer in a final volume of 900 µl is preincubated for 10 min at RT. The reaction is started by the addition of 50µl of 1 m*M* NADPH solution. The reaction is stopped after 2 min by adding 1,2 m*M* KOH. Indigo formation is spectrophotometrically determined at 670 nm (extinction coefficient $\varepsilon = 3{,}9\ M^{-1}cm^{-1}$). Formation of indirubin has only a weak effect on the absorption at 670 nm.[60]

Activity against polycyclic aromatic hydrocarbons

Several mutants of P450 BM-3 with activity against several polycyclic aromatic hydrocarbons were designed. Two hydrophobic substitutions - R47L and Y51F - increased the activity of this enzyme against phenanthrene, fluoranthene and pyrene up to 40-fold. After combination of these mutations with mutation A264G, PAH oxidation increased another 5-fold, with simultaneous enhancement of NADPH oxidation and coupling efficiency between NADPH oxidation and substrate hydroxylation. All of the above residues are located

in the substrate binding pocket of the enzyme. However, substitution of M354 and L437, also located within the active site, reduced PAH oxidation activity.[37]

Directed evolution

Directed evolution involves either random mutagenesis (e.g. error-prone PCR) or recombination of gene fragments (e.g. gene shuffling, staggered extension process etc.). It has proved very efficient in improving the enzymatic activity[36,61] and stability of P450 BM-3[62]. Directed evolution does not rely on the structural information about a certain protein. Instead, it can lead to mutants that help us understand the relationship between structure and function.

Error-prone PCR has been applied to alter the features of P450 BM-3. Error-prone PCR procedures are generally modifications of a standard PCR protocol which are aimed at increasing the natural error rate of the polymerase. A standard error-prone reaction mixture contains a higher concentration of $MgCl_2$ (6-7 m*M*) compared to basic PCR (1,5 m*M*), which is required for the stabilization of non-complimentary pairs. To increase the error-rates, either $MnCl_2$ (0-0,2 m*M*) can be added to the reaction mixture or unbalanced amounts of dNTPs. *Taq* polymerase, which is typically applied for error-prone PCR, has a high error rate of its own. However, there are several newly-constructed polymerases which allow increasing the number of variations in the mutants[63].

Procedure. Mutagenic PCR was performed on the heme domain of P450 BM-3 on the whole CYP102A1 gene or its mutant. The standard reaction mixture in 50 µl contains: 5 µl 10 x PCR buffer (100 m*M* Tris-HCl, pH 8.8 at 25°C; 500 m*M* KCl, 0.8% Nonidet P40 and 15 m*M* $MgCl_2$), 4 µl dNTPs mix in unbalanced concentrations (2,5 m*M* : 10 m*M* : 10 m*M* : 10 m*M*); forward and reverse primers introducing the restriction sites (the same as for gene amplification) (40 pmol each), template DNA (10 ng), $MnCl_2$ (0-0,2 m*M*).

PCR program, subsequent restriction of the PCR products and transformation of the plasmid are the same as in the standard PCR protocol. To find the appropriate mutagenesis rate, the percentage of inactive clones among mutants was estimated.

To take advantage of the powerful error-prone PCR, an efficient expression system (see above) is required. Therefore protein induction and isolation are performed in a 96-well microtiterplate scale.

The colonies are picked with a robot and inoculated into a microtiter plate containing 150µl Luria-Bertani (LB) supplemented with ampicillin (100 µg/ml). The plates are incubated overnight at 37°C on an orbital shaker at 200 rpm. These microtiter plates are also used as stock plates. Then, 25 µl of overnight culture is added to a 1.2 ml well-plate with square wells containing 375 µl LB medium with ampicillin (100 µg/ml). Cells are grown at 37°C and 200 rpm to an OD_{578} ~ 0,8 and then induced with 500 µM IPTG. After incubation at 30°C and 200 rpm for 12-16 hr after induction, the cells are harvested by centrifugation and the medium discarded.

The cell pellet is resuspended in lysobuffer (potassium phosphate buffer (50 m*M*, pH 7.5, containing 1 mg/ml lysozyme, 1 mg/ml DNAse, protease-inhibitor and 40 m*M* NaCl). After incubation at 4°C for 30 min the plates are centrifuged at 2500 g. The supernatant contains the active P450 BM-3 enzyme.

For activity measurements a rapid and reproducible assay is necessary that is sensitive to minute changes. In addition, it should be adapted to high-throughput screening. The standard method to determine P450 activity involves the measurement of the rate of NADPH consumption in the presence of substrate. This method was also adapted to a microtiter-plate scale.[64] However, the rate of NADPH consumption can be very low in the case of unusual substrates and the uncoupling between reductase and monooxygenase activity may be significant. To overcome this problem, assays with substrate analogs can be used. These can

include colorimetric assays for fatty acid hydroxylation (*p*-NCA assay)[65] (see above) or alkane hydroxylation (*p*-pnpane assay),[36] which are optimized for HTS. The screening for alkane oxidation activity is based on the use of a substrate analog which generates aldehyde and yellow *p*-nitrophenolate after terminal hydroxylation.

Arnold and colleagues have studied directed evolution for the conversion of P450 BM-3 into an efficient alkane hydroxylase.[50] After five mutagenesis rounds starting with the wild-type P450 BM-3 and subsequent screening for a better activity, one variant with eleven substitutions was found that could accept not only hexane and octane as a substrate, but also gaseous alkanes such as butane and propane. Surprisingly, some mutant were found which lead to either the (R)- or (S)-enantiomer products of alkane hydroxylation (A. Glieder, personal communication).

Another direction of research was to employ hydrogen peroxide instead as both an oxidant and an electron donor, thereby substituting NADPH.[61] Active peroxygenase[61] and some thermostable variants[62] were evolved, which could use hydrogen peroxide instead of oxygen for the hydroxylation of various substrates.

Conclusions

Protein engineering approaches are powerful tools with regard to elucidating the critical role of key residues of the P450 BM-3 monooxygenase. Site-directed and random mutageneses enable the manipulation of catalytic properties of this enzyme. Both methods are also suitable for investigating the electron transfer between the reductase and the heme domain. Although insights into different aspects of mechanism and catalytical functions of P450 enzymes have already been provided, many unresolved issues remain. The chemical activation of the substrate, physical aspects of protein-protein and protein-substrate interactions, reactions without oxygen insertion or multi-step reactions can be solved by

combining protein engineering and molecular modeling based on the huge natural variation of P450s in structure and sequence.

Note added in proof

Since the article has been submitted significant progress in protein engineering of P450 BM-3 has been achieved. The stability and activity of P450 BM-3 in the presence of different polar organic solvents was significantly increased using methods of directed evolution.[66] Using a combination of error-prone PCR and site-directed mutagenesis different P450 BM-3 variants were engineered, which are able to hydroxylate linear alkanes from C3 to C10 regio- and enantioselectively.[67]

Acknowledgments

We thank Juergen Pleiss and Stephan Tatzel who provided the model of CYP102A1 shown in Fig. 1. We wish to acknowledge financial support by EU grant QLRT 200000725 and by BASF AG, Ludwigshafen. R.D. Schmid acknowledges generous support by the Verband der Chemischen Industrie (VCI), Frankfurt.

References

[1] I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, and S. G. Sligar, *Science* **287,** 1615 (2000).

[2] M. Hata, Y. Hirano, T. Hoshino, and M. Tsuda, *J. Am. Chem. Soc*. **123,** 6410 (2001).

[3] D. A. Rock, B. N. Perkins, J. Wahlstrom, and J. P. Jones, *Arch. Biochem. Biophys.* **416,** 9 (2003).

[4] R. Dai, M. R. Pincus, and F. K. Friedman, *J. Protein. Chem*. **17,** 121 (1998).

[5] N. V. Belkina, M. Lisurek, A. S. Ivanov, and R. Bernhardt, *J. Inorg. Biochem.* **87,** 197 (2001).

[6] D. F. Lewis, *Xenobiotica* **32,** 305 (2002).

[7] D. F. Lewis, and M. Dickins, *J. Enzyme. Inhib.* **16,** 321 (2001).

[8] T. L. Poulos, B. C. Finzel, and A. J. Howard, *Biochemistry* **25,** 5314 (1986).

[9] K. G. Ravichandran, S. S. Boddupalli, C. A. Hasermann, J. A. Peterson, and J. Deisenhofer, *Science* **261,** 731 (1993).

[10] C. A. Hasemann, K. G. Ravichandran, J. A. Peterson, and J. Deisenhofer, *J. Mol. Biol.* **236,** 1169 (1994).

[11] K. Nakahara, H. Shoun, S. Adachi, T. Iizuka, and Y. Shiro, *J. Mol. Biol.* **239,** 158 (1994).

[12] J. R. Cupp-Vickery, and T. L. Poulos, *Nat. Struct. Biol.* **2,** 144 (1995).

[13] S. Y. Park, K. Yamane, S. Adachi, Y. Shiro, K. E. Weiss, and S. G. Sligar, *Acta Crystallogr. D Biol. Crystallogr.* **56,** 1173 (2000).

[14] L. M. Podust, Y. Kim, M. Arase, B. A. Neely, B. J. Beck, H. Bach, D. H. Sherman, D. C. Lamb, S. L. Kelly, and M. R. Waterman, *J. Biol. Chem.* **278,** 12214 (2003).

[15] L. M. Podust, T. L. Poulos, and M. R. Waterman, *Proc. Natl. Acad. Sci. U S A* **98,** 3068 (2001).

[16] J. K. Yano, F. Blasco, H. Li, R. D. Schmid, A. Henne, and T. L. Poulos, *J. Biol. Chem.* **278,** 608 (2003).

[17] S. Nagano, H. Li, H. Shimizu, C. Nishida, H. Ogura, P. R. Ortiz De Montellano, and T. L. Poulos, *J. Biol. Chem.* [Epub ahead of print] (2003).

[18] P. A. Williams, J. Cosme, V. Sridhar, E. F. Johnson, and D. E. Mcree, *J Inorg Biochem* **81,** 183 (2000).

[19] P. A. Williams, J. Cosme, V. Sridhar, E. F. Johnson, and D. E. Mcree, *Mol. Cell.* **5,** 121 (2000).

[20] D. F. Lewis, C. Sams, and G. D. Loizou, *J. Biochem. Mol. Toxicol.* **17,** 47 (2003).

[21] G. D. Szklarz, and J. R. Halpert, *Life Sci.* **61,** 2507 (1997).

[22] R. Davydov, T. M. Makris, V. Kofman, D. E. Werst, S. G. Sligar, and B. M. Hoffman, *J. Am. Chem. Soc.* **123,** 1403 (2001).

[23] C. F. Harford-Cross, A. B. Carmichael, F. K. Allan, P. A. England, D. A. Rouch, and L.-L. Wong, *Prot. Eng.* **13,** 121 (2000).

[24] K. J. French, D. A. Rock, J. I. Manchester, B. M. Goldstein, and J. P. Jones, *Arch. Biochem. Biophys.* **398,** 188 (2002).

[25] O. Sibbesen, J. J. De Voss, and P. R. Montellano, *J. Biol. Chem.* **271,** 22462 (1996).

[26] H. Y. Li, K. Darwish, and T. L. Poulos, *J. Biol. Chem.* **266,** 11909 (1991).

[27] H. Li, and T. L. Poulos, *Nat. Struct. Biol.* **4,** 140 (1997).

[28] D. C. Haines, D. R. Tomchick, M. Machius, and J. A. Peterson, *Biochemistry* **40,** 13456 (2001).

[29] H. Yeom, S. G. Sligar, H. Li, T. L. Poulos, and A. J. Fulco, *Biochemistry* **34,** 14733 (1995).

[30] C. F. Oliver, S. Modi, M. J. Sutcliffe, W. U. Primrose, L. Y. Lian, and G. C. Roberts, *Biochemistry* **36,** 1567 (1997).

[31] L. P. Wen, and A. J. Fulco, *J. Biol. Chem.* **262,** 6676 (1987).

[32] R. T. Ruettinger, L. P. Wen, and A. J. Fulco, *J. Biol. Chem.* **264,** 10987 (1989).

[33] H. J. Barnes, *Methods in Enzymology* **272,** 3 (1990).

[34] C. W. Fisher, D. L. Caudle, C. Martin-Wixtrom, L. C. Quattrochi, R. H. Tukey, M. R. Waterman, and R. W. Estabrook, *Faseb J.* **6,** 759 (1992).

[35] C. W. Fisher, M. S. Shet, and R. W. Estabrook, *Methods in Enzymology* **272,** 15 (1990).

[36] E. T. Farinas, Schwaneberg, U., Glieder, A., Arnold, F.H., *Advanced synthesis and catalysis* **343(6+7),** 601 (2001).

[37] A. B. Carmichael, and L. L. Wong, *Eur. J. Biochem.* **268,** 3117 (2001).

[38] J. Sambook, and D. W. Russell "Molecular Cloning." Cold Spring Harbor Laboratory Press, New York, 2001.

[39] U. Schwaneberg, C. Schmidt-Dannert, J. Schmitt, and R. D. Schmid, *Anal. Biochem.* **269,** 359 (1999).

[40] S. C. Maurer, H. Schulze, R. D. Schmid, and V. Urlacher, *Advanced Synthesis & Catalysis* **345,** 802 (2003).

[41] T. W. Ost, C. S. Miles, J. Murdoch, Y. Cheung, G. A. Reid, S. K. Chapman, and A. W. Munro, *FEBS Lett.* **486,** 173 (2000).

[42] T. W. Ost, C. S. Miles, A. W. Munro, J. Murdoch, G. A. Reid, and S. K. Chapman, *Biochemistry* **40,** 13421 (2001).

[43] T. A. Kunkel, *Proc. Natl. Acad. Sci. U S A* **82,** 488 (1985).

[44] Q. S. Li, U. Schwaneberg, M. Fischer, J. Schmitt, J. Pleiss, S. Lutz-Wahl, and R. D. Schmid, *Biochim. Biophys. Acta* **1545,** 114 (2001).

[45] C. F. Oliver, S. Modi, W. U. Primrose, L. Y. Lian, and G. C. Roberts, *Biochem. J.* **327 (Pt 2),** 537 (1997).

[46] L. A. Cowart, J. R. Falck, and J. H. Capdevila, *Arch. Biochem. Biophys.* **387,** 117 (2001).

[47] M. A. Noble, C. S. Miles, S. K. Chapman, D. A. Lysek, A. C. Mackay, G. A. Reid, R. P. Hanzlik, and A. W. Munro, *Biochem. J.* **339,** 371 (1999).

[48] S. Graham-Lorence, G. Truan, J. A. Peterson, J. R. Falck, S. Wei, C. Helvig, and J. H. Capdevila, *J. Biol. Chem.* **272,** 1127 (1997).

[49] Q. S. Li, J. Ogawa, R. D. Schmid, and S. Shimizu, *Appl. Environ. Microbiol.* **67,** 5735 (2001).

[50] A. Glieder, E. T. Farinas, and F. H. Arnold, *Nat. Biotechnol*. **20,** 1135 (2002).

[51] A. Muller, J. J. Muller, Y. A. Muller, H. Uhlmann, R. Bernhardt, and U. Heinemann, *Structure* **6,** 269 (1998).

[52] I. F. Sevrioukova, and J. A. Peterson, *Biochimie* **77,** 562 (1995).

[53] I. F. Sevrioukova, H. Li, H. Zhang, J. A. Peterson, and T. L. Poulos, *Proc. Natl. Acad. Sci. U S A* **96,** 1863 (1999).

[54] A. W. Munro, K. Malarkey, J. Mcknight, A. J. Thomson, S. M. Kelly, N. C. Price, J. G. Lindsay, J. R. Coggins, and J. S. Miles, *Biochem. J.* **303 (Pt 2),** 423 (1994).

[55] M. L. Klein, and A. J. Fulco, *J. Biol. Chem* **268,** 7553 (1993).

[56] T. W. Ost, A. W. Munro, C. G. Mowat, P. R. Taylor, A. Pesseguiero, A. J. Fulco, A. K. Cho, M. A. Cheesman, M. D. Walkinshaw, and S. K. Chapman, *Biochemistry* **40,** 13430 (2001).

[57] G. Gilardi, Y. T. Meharenna, G. E. Tsotsou, S. J. Sadeghi, M. Fairhead, and S. Giannini, *Biosens. Bioelectron.* **17,** 133 (2002).

[58] O. Lentz, Q.-S. Li, and U. Schwaneberg, Lutz-Wahl, S., Fischer, P., Schmid, R.D., *Journal of molecular catalysis B: Enzymatic* **15,** 123 (2001).

[59] D. Appel, S. Lutz-Wahl, P. Fischer, U. Schwaneberg, and R. D. Schmid, *J. Biotechnol.* **88,** 167 (2001).

[60] Q.-S. Li, U. Schwaneberg, P. Fischer, and R. D. Schmid, *Chem. Eur. J.* **6,** 1481 (2000).

[61] P. C. Cirino, and F. H. Arnold, *Angew. Chem. Int. Ed. Engl.* **42,** 3299 (2003).

[62] O. Salazar, P. C. Cirino, and F. H. Arnold, *Chembiochem.* **4,** 891 (2003).

[63] P. C. Cirino, K. M. Mayer, and D. Umeno, *Methods Mol. Biol* . **231,** 3 (2003).

[64] A. Glieder, and P. Meinhold, *Methods Mol. Biol* . **230,** 157 (2003).

[65] U. Schwaneberg, C. Otey, P. C. Cirino, E. Farinas, and F. H. Arnold, *J. Biomol. Screen.* **6,** 111 (2001).

[66] T. Seng Wong, F. H. Arnold, and U. Schwaneberg, *Biotechnol. Bioeng.* **85**, 351 (2004).

[67] M. W. Peters, P. Meinhold, A. Glieder, and F. H. Arnold, *J. Am. Chem. Soc.* **125**, 13442 (2003).

Figure legends

Fig. 1. P450 BM-3-palmitate complex model, showing the mutations that can affect the substrate specificity.