Engineered Bacterial Mimics of Human Drug Metabolizing Enzyme CYP2C9

Andrea Rentmeister, Tristan R. Brown, Christopher D. Snow, Martina N. Carbone, and Frances H. Arnold

Simple and universal methods for the preparation of human drug metabolites are required to produce quantities sufficient for their characterization and toxicity testing. Synthetic chemistry lacks general catalysts for selective oxidation of unactivated C–H bonds, a transformation that plays a key role in metabolism; bioconversions using P450 enzymes have emerged as a powerful alternative. Variants of P450BM3 from Bacillus megaterium act on diverse substrates, including drugs. Acidic substrates, such as the compounds metabolized by CYP2C9, which is one of three main hepatic human P450s, are not accepted by P450BM3 variants engineered to date. Herein, we report bacterial mimics of CYP2C9, which are active on two widely administered drugs, naproxen and ibuprofen, that are CYP2C9 substrates in vivo. These P450BM3 variants can also act on desmethylnaproxen, the human metabolite of naproxen, and convert it to the 1,4-naphthoquinone derivative. We analyzed the crystal structure of the heme domain of an early intermediate in the directed-evolution experiment. The active site mutation, L75R, which initially conferred activity on charged substrates, dramatically increased structural flexibility in the B’-helix. This increased flexibility, which was accompanied by a dramatic decrease in enzyme stability, may contribute to the variant’s ability to accept a broader range of substrates.

Introduction

Metabolites play a significant role in the pharmacology and toxicity of a drug. Most drugs are substrates for the hepatic P450 oxidases. Commonly, a compound is rendered less toxic, with notable exceptions. Metabolites can impose their own pharmacological, toxicological, and physiological effects, which lead to complications that have resulted in the withdrawal of several drugs from the market in recent years, including iproniazid, troglitazone, benoxaprofen, phenfluramine, and pemoline. For prodrugs, the metabolites are the efficacious species. Because of the potential pharmaceutical activity and toxicity of drug metabolites, major metabolites must also be tested in ADME (absorption, distribution, metabolism, and excretion) studies. Testing is especially important if metabolites are present at high levels in plasma and tissues, and if human metabolites are not formed in animal models. Common reactions in drug metabolism involve regioselective oxidation of unactivated C–H bonds, a transformation which currently lacks a general counterpart in synthetic organic chemistry. Therefore, drug metabolites usually cannot be synthesized directly from the drug itself and require a new synthetic strategy. Direct preparation by using P450 enzymes is an attractive alternative. Hepatic microsomes are a source for many different P450 activities, but their availability is limited, and their enzyme expression levels often depend on prior induction and exposure to xenobiotics. These limitations severely restrict their application for preparative synthesis of metabolites. Heterologous expression of human CYPs could be a solution. Unfortunately, human P450s are membrane-bound and expressed at low levels, require two-component electron-transfer systems, and exhibit low activity and stability. Although many bacteria and fungi also perform a mammalian-like metabolism, many of these strains are uncharacterized and grow poorly in culture. A potential solution is to clone the P450 genes from these strains into host organisms that are more easily cultivated and then use these host organisms to produce relevant drug metabolites.

The bacterial P450 monoxygenase, P450BM3, a fatty acid hydroxylase isolated from Bacillus megaterium, has a covalent link between its hydroxylase and diflavin reductase domains, is expressed at high levels in Escherichia coli, and exhibits very fast reaction rates (thousands of turnovers per minute) on favored substrates. P450BM3 can be engineered to accept a variety of substrates, including drugs. We, and others, have demonstrated that P450BM3 variants are capable of producing human metabolites of propanolol, verapamil, astemizole, 7-ethoxycoumarin, chloroxazone, amodiaquine, dextromethorphan, buspirone, and 3,4-methylenedioxymethyl-amphetamine (MDMA).
To date, most of the drugs that can be metabolized by using P450variants are substrates of human CYP3A4, CYP2E1, and CYP2D6. CYP3A4 and CYP2C9 are the two main hepatic P450s, constituting approximately 50% of the total hepatic P450 protein,[13] and contribute to the metabolism of more than 60% of all drugs, based on available literature.[14] CYP2C9 alone is responsible for approximately 15% of known P450 conversions and has a preference for substrates with weak acidity.[15] Many commonly used nonsteroidal anti-inflammatory drugs (NSAIDs), such as those of the profen family (naproxen, ibuprofen, flurbiprofen, ketoprofen) and diclofenac, have a carboxyl moiety and are substrates of CYP2C9.[16] There has been no report on a P450 variant exhibiting good activity on these charged molecules.[7,11]

We aimed to address this shortcoming and extend the activity of P450 variants to weakly acidic drugs that fall into the substrate spectrum of human CYP2C9. Substrates chosen for this study include two widely applied NSAIDs, naproxen 1 and ibuprofen 5 (Scheme 1). We previously reported a variant of the P450 heme domain with very low, but detectable, activity on naproxen.[17] This variant, termed 13C9, was found by using random mutagenesis and screening of 5H6, a thermostabilized, H2O2-driven hydroxylase (heme domain "peroxxygenase" without the reductase domain). With respect to 5H6, variant 13C9 carried a single mutation at L75R, a position that has been shown to affect activity on short-chain carboxylic acids.[18] We had hypothesized that R75, in the substrate-binding pocket, allowed 13C9 to bind naproxen by providing a compensating positive charge for the naproxen carboxylate group.[19] It was possible that the R75 in helix B interacted directly with the carboxylate group, but there was no experimental evidence to support this.

The goals of the current study were 1) to increase enzyme turnover to a level that would enable preparative-scale reactions[19] (TTN of at least 1000, with TTN = total turnover number = mol product per mol enzyme catalyst) by directed evolution of 13C9 and 2) to elucidate the structural basis for the broad tolerance for negatively-charged substrates caused by the L75R mutation. Therefore, we determined the crystal structure of 22A3, an early variant in the directed-evolution experiment carrying the L75R mutation and still showing broad substrate tolerance.

Results and Discussion

Directed evolution to generate a CYP2C9 functional mimic

Analysis using HPLC and HPLC–MS confirmed that the human metabolite desmethylnaproxen 2 was the main product of the bioconversion of naproxen in the presence of the parent P450 variant 13C9 and peroxide (Figure 1, m/z = 215, TTN 8). Three aromatic hydroxylation products (m/z = 246, TTN 1 each) were also found. Wild-type P450 showed no activity on naproxen (Figure 1).

To improve the TTN of 13C9 for desmethylnaproxen, we first reattached the reductase domain to obtain the efficient holoenzyme of P450. Rebuilding the monooxygenase from the heme domain peroxxygenase has been shown to improve the hydroxylase activity of P450 peroxxygenase variants[20] and increases the lifetime of the enzyme by substituting O2, NADPH, and a regeneration system for the peroxide.

The reconstituted monooxygenase 13C9R1 indeed showed improved activity for the production of desmethyl naproxen (TTN 15, Figure 1) and produced a mixture of hydroxylated side products. 13C9R1 was subjected to five rounds of directed evolution toward higher yields of desmethylnaproxen. To test for improved demethylation activity, we used the purpald screen, which detects the formaldehyde released upon hydroxylation at the methoxy group.[21]

We used error-prone polymerase chain reaction (PCR) to randomly mutate the DNA encoding the heme domain, with mutation rates between 1.5 and 6 nucleotides (nt) per gene, corresponding to an average of 0.3–3 amino acid substitutions (see the Supporting Information, Table S2). Variants exhibiting higher demethylation activity usually contained 1–2 amino acid substitutions with respect to their parent (Table 1, Supporting Information, Table S3). In one round (starting from 16G2), we did not find any improved variant upon screening approximately 3000 clones, but repeating the screen with a new random mutant library gave improved variant W7D8. Figure 1B shows the course of the evolution, in terms of the relative activity in the purpald screen and the TTNs of bioconversions by using purified enzymes. Final variant X3H1 carried six mutations with respect to 13C9R1 and exhibited ≥1000 TTN for the conversion of naproxen to desmethyl naproxen, corresponding to a more than 60-fold improvement over the recon-
activity from the purple screening assay. Mutations relative to the preceding generation are indicated. Middle: TTN from bioconversions using purified enzyme. TTN for naproxen (gray bars) increases with TTN on ibuprofen (black bars) for variants 13C9R1 to W7D8. Variant X3H1 shows increased activity on naproxen, but reduced activity on ibuprofen, indicating specialization on approximate. Bottom: Previously reported that P450 products (profen with X3H1; total ion count for indicated masses shows ibuprofen (7)) by human P450 CYP2C9 and CYP2C8 (Scheme 1). We presented in Table 1. The first variant 22A3, which differed from 13C9R1 by a single F162I substitution, showed increased activity and TTN on both, naproxen and ibuprofen. Increased activity on ibuprofen correlated with improvements on naproxen for variants 13C9R1 to W7D8 (Figure 1B). Only M185K in variant X3H1 increased the activity on naproxen, but not on ibuprofen.

Most mutations are deleterious, which is also true for mutations that increase activity. The thermostability was measured using the T50 value. During the course of directed evolution, T50 decreased from 54 to 45.5 °C for 16G2 before increasing again to 48.7 °C (Figure 1). As a consequence of its decreased stability, variant 16G2 catalyzed fewer TTNs than its parent 22A3, even though it was more active in the screen (Figure 1). Notably, the T50 nadir corresponded to the point at which it became hard to detect improved mutants (only 1 improved variant identified in 6000 clones), which is anecdotal evidence in support of the observation made by Bloom et al. that a T50 of 47 °C is a minimum threshold for P450 variants to accept further mutations under these experimental conditions.23 Subsequent variant W7D8 showed improved activity, thermostability, and TTN. The most active variants, X3H1 and W7D8, allowed conversion of the NSAIDs naproxen and ibuprofen with TTNs of 1000 and 600, respectively.

As the variants catalyzed more turnovers on naproxen, a second product began to appear in the bioconversions. The new product had an absorption spectrum very different from at positions 1 and 2,23 but ibuprofen itself was not a substrate. We now found that the activity on ibuprofen was concomitant with acquisition of activity on naproxen in the initial rounds of directed evolution. Bioconversion of ibuprofen yielded one major (6) and one minor (7) hydroxylation product (Figure 1A). Interestingly, the hydroxylation at position 1 of the ibuprofen methyl ester was not observed, presumably because it would require further burial of the charged group within the hydrophobic binding channel. By using the P450 variant W7D8, we can produce the human drug metabolite, 2-hydroxyibuprofen 6, on a preparative scale with 96% isolated yield (Scheme 1).

Conversions and TTNs of the different P450 variants for ibuprofen, as well as selectivities in the form of product ratios for 2'- versus 3'-hydroxylation, are presented in Table 1. The first variant 22A3, which differed from 13C9R1 by a single F162I substitution, showed increased activity and TTN on both, naproxen and ibuprofen. Increased activity on ibuprofen correlated with improvements on naproxen for variants 13C9R1 to W7D8 (Figure 1B). Only M185K in variant X3H1 increased the activity on naproxen, but not on ibuprofen.

In the bioconversions, we also detected activity of the evolved P450 variants on ibuprofen 5, an NSAID commonly administered for its anti-inflammatory, analgesic, and antipyretic properties. Ibuprofen is hydroxylated at positions 2 (6) and 3 (7) by human P450s CYP2C9 and CYP2C8 (Scheme 1). We previously reported that P450 variants were active on the methylester of ibuprofen and isolated products hydroxylated

Table 1. Evolved P450 variants and their mutations and characteristics. Mutations listed are with respect to the variant on the preceding row, except for 13C9R1. Conditions: substrate (3 mM), P450 variant (0.033 mol%), regeneration system, 0 °C, shaken overnight.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Mutation(s)</th>
<th>Naproxen (1) Conv. to 2</th>
<th>TTN</th>
<th>Ibuprofen (5) Total conv. Ratio 6/7</th>
<th>TTN for 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13C9R1</td>
<td></td>
<td>&lt;1</td>
<td>15</td>
<td>4</td>
<td>1:1</td>
</tr>
<tr>
<td>22A3</td>
<td>F162I</td>
<td>3</td>
<td>90</td>
<td>8</td>
<td>3:1</td>
</tr>
<tr>
<td>16G2</td>
<td>Q109L, E140G</td>
<td>1</td>
<td>30</td>
<td>3</td>
<td>1:5</td>
</tr>
<tr>
<td>W7D8</td>
<td>V286A, R398H</td>
<td>20</td>
<td>650</td>
<td>20</td>
<td>35:1</td>
</tr>
<tr>
<td>X3H1</td>
<td>M185K</td>
<td>35</td>
<td>1010</td>
<td>10</td>
<td>7:1</td>
</tr>
</tbody>
</table>

[a] See the Supporting Information.
that of naproxen and desmethylnaproxen (Supporting Information, Figure S1). The concentration of the new product strongly correlated with the activity on naproxen. The progression of the reaction, and the analysis of the reaction starting from desmethylnaproxen, suggested that the new product was formed in a consecutive rather than a parallel reaction (see the Supporting Information, Figure S1). Isolation and characterization by using HPLC–MS and 13C, and 2D-NMR spectroscopy revealed that the product was (S)-2-(6-hydroxy-5,8-dioxo-5,8-dihydronaphthalene-2-yl)propanoic acid 4.

We propose that desmethylnaproxen was hydroxylated at position 5 (Scheme 1, 3), as in the Elbs-oxidation of β-naphthol with persulfate,[24] and that the phenolic product was further oxidized to quinone by air[25] or superoxide.[26] Alternative paths for the oxidation of 2-hydroxynaphthalene to a 2-hydroxy-1,4-naphthoquinone have been described.[25, 27] The naproxen bioconversion could be controlled to produce 82% (43% isolated yield) desmethylnaproxen, because the aromatic hydroxylation is significantly slower than the demethylation.

We also tested the variants for activity on 1-napthalene acetic acid (NAA; Figure 2A, 8), which is structurally related to naproxen. NAA is a plant hormone in the auxin family and an ingredient in many commercial plant rooting horticultural products. Parent 13C9R1 displayed a low, but detectable activity on 8, which was increased slightly in the early intermediate 22A3. Starting with 16G2, however, variants lost their activity on 8, specializing instead on 1 and 5. Variants with increased activity on NAA were found as the libraries were screened on NAA, and these variants, 56C4 and S31A1, had mutations different from those in 16G2. Activity on NAA was further improved by recombination and error-prone PCR. Figure 2B illustrates the specialization of variants screened for activity on 1 and 8. Early intermediate 22A3 is the last common ancestor with improved activity on all three substrates.

We thus extended the drug metabolite repertoire of P450BM3 variants to include weakly acidic substrates of human CYP2C9. Mutations R398H and M185K exhibited the largest activity improvements. M185K was the only mutation that increased activity on naproxen, but not ibuprofen. R398H was a surprising mutation, because R398 is highly conserved and stabilizes the heme propionate through bidentate hydrogen bonding with the propionate side chain of the heme.[28] Therefore, the mutation of R398 to histidine in W7D8 is likely to influence the conformation of the heme propionates or even the overall heme orientation, although we conjecture that H398 retains one hydrogen bond to the propionate as shown in the modeled histidine rotamer (Supporting Information, Figure S3). Another potential disruptive effect of R398H would be a conformational change of the loop following the B helix, caused by the loss of hydrogen bonds to the backbone carbonyls of L86 and S89. Finally, the proximity of R398H to previously accumulated mutations L75R and H100R (Figure S3) suggests possible interactions mediated by a conformational change of the heme or the loop following the B helix.

Structure of an early intermediate in the directed evolution

We crystallized the heme domains of the early intermediates 13C9 and 22A3 to better understand the effect of the L75R mutation in P450BM3 and to explore the structural basis of their broadened substrate ranges. Specifically, we wanted to determine whether the resulting R75 in P450BM3 played a role similar to R108 in human CYP2C9, in which the arginine side chain formed hydrogen bonds directly with the carboxyl group of the substrate, as shown in the X-ray structure of the complex with flurbiprofen (Figure 3A).[29] Overall, the P450BM3 variant 22A3 carries 18 mutations in the heme domain and is, to the best of our knowledge, the most highly mutated P450BM3 variant, for which the structure has been determined to date. Seven of these mutations were accumulated during directed evolution for increased thermostability (highlighted in red in the Supporting Information, Figure S2). Four mutations (L75R, F87A, H100R, F107L) were located within 8/p'3x of the heme; among them is the key L75R mutation. The other three mutations originated from previous directed evolution for an increased peroxynase activity.[30] Most mutations are far (>8 Å) from the active site.

We prepared 22A3 crystals that diffracted to 3.1 Å (Supporting Information, Table S5). The unit cell contained two copies of the protein (chain A and chain B). The destabilizing influence of the L75R mutation[31] was immediately apparent from the 22A3 crystal structure, as it increased local flexibility. For chain A, the entire B helix containing the mutation was unresolved. For chain B, helix B and R75 were resolved, but with exceptionally high B factors (Figure 4) for the entire region, suggesting that the region was not well determined. This usually reflects imperfect order in the crystal, indicating high flexibility, for example, in less structured regions. Introduction of R75 into the predominantly hydrophobic active site of P450BM3...
provided the basis for novel albeit rather nonspecific acceptance of negatively charged substrates. L75R, however, disrupted the local structure, resulting in a reduced stability of the protein and an increased flexibility of the B’ helix close to the active site. This flexibility may contribute to the broad substrate tolerance\cite{31} in 13C9 and 22A3. Subsequent rounds of directed evolution revealed mutations that improved accommodation of these drastic changes, resulting in increased thermostability as well as specificity for the target substrate.

**Modeling naproxen into 22A3**

Human CYP2C9 has a preference for weakly acidic ligands,\cite{15} and its active-site residue R108 forms hydrogen bonds with the anionic ligand moiety (Figure 3A).\cite{29,32} In the substrate-binding pocket of P450BM3, R75 could play a similar role in the interaction of the bacterial P450 with naproxen.\cite{17} To assess the role of the L75R mutation, we modeled potential transition state conformations for the naproxen hydroxylation in the 22A3 active site by using the ROSETTA energy function.\cite{33} After fixing the bond lengths and angles for Fe–O–H–R\cite{34}, we exhaustively searched a large number of discrete substrate orientations. Potential docking positions that did not clash with the protein backbone were further optimized by allowing the sidechains of the protein to relax.\cite{35} For each resulting model, we calculated the ROSETTA score. The most favorable conformation of the transition state of naproxen placed the carboxylate moiety toward the entrance of the binding channel. In this general orientation, substrate binding displaced R75 and L437 (Figure 3B) and the R75 guanidinium group could not form a direct hydrogen bond to naproxen, but instead remained associated with the heme propionate.

Thus, we propose that R75 in P450BM3 plays a role different from that of R108 in CYP2C9. Instead of forming direct hydrogen bonds, R75 seems to compensate the negative charge in a more general way. By introducing favorable electrostatic interactions, the buried positive charge of R75 decreases the penalty for desolvating naproxen. In contrast, modeling suggests that K185 could interact directly with naproxen in the transition state, depending on the exact position of helix F.

**Conclusions**

Drug metabolites can impose their own pharmacological, toxicological, and physiological effects, leading to complications that have resulted in the withdrawal of several drugs from the market.

---

**Figure 3.** Charge compensation in the human P450 CYP2C9 and its bacterial mimic from P450BM3. A) CYP2C9 (pdb 1RO9, purple) in a complex with flurbiprofen. R108 (cyan) formed hydrogen bonds with the carboxyl group of the propionate side chain. Hydroxylation occurs at the 4’-position pointing toward the heme. B) Naproxen conformers were docked to the 22A3 crystal structure (white backbone) and positioned in a transition state geometry relative to the heme (blue spheres) of a hypothetical compound I oxygen (red sphere). The most favorable naproxen poses (orange) induced relaxation of L437 and R75 (green). In all cases, the favorable naproxen poses extended along the binding channel, and R75 was unable to form a direct hydrogen bond to naproxen.

**Figure 4.** Flexibility induced in helix B’ of the P450BM3 variant 22A3. Chain B of pdb3QI8 is colored according to the B factor (least flexible to most flexible: blue–green–yellow–red). Helix B’, harboring the L75R mutation, was highly flexible and showed the highest B factor (red).
market. Most drugs are substrates for the hepatic P450 oxidases. In vivo CYP2C9 has a preference for substrates with weak acidity like many NSAIDs. Up to now, there has been no reported P450 variant exhibiting good activity on these charged molecules. In the current study, we report two P450 variants, X3H1 and W7DB, which allow conversion of the NSAIDs naproxen and ibuprofen with TTNs of 1000 and 600, respectively, and can be used to produce known and novel drug metabolites on a preparative scale.

From crystal structure analysis and modeling, we conclude that the mode of substrate binding in the early, promiscuous P450 variants differs from that of human CYP2C9. The negative charge and drastic increase in flexibility in the neighborhood of L75R may be the structural basis for the broadened substrate tolerance. This basal activity subsequently served as a starting point for further directed evolution. Such a progression corresponded to previously observed enzyme ‘promiscuity’, that appeared during early rounds of directed evolution with low activities, and that is followed by respecialization at higher activities. However, the stability-function tradeoff is not obligatory, as was apparent during later directed evolution rounds in which mutations were found that increased both stability and activity. Although such mutations are generally rare, we found them repeatedly when screening near the apparent stability threshold, where stabilizing mutations can positively affect activity integrated over the screening period.

Experimental Section

All experimental details and protocols are described in the Supporting Information. Data collection and refinement statistics of the crystal structure are also appended. Coordinates and structure factors have been deposited in the Protein Data Bank (www.pdb.org) with accession number 3Q18.

Acknowledgements

We thank Dr. J.T. Kaiser and P. Nikolovski for their excellent support in high-throughput crystallization, data collection, and crystallographic questions. We acknowledge M. Shahgholi and Dr. D. VanderVelde for assistance with LC–MS, HRMS, and 2D-NMR spectroscopy, respectively. The authors acknowledge support from the Department of Energy BES (grant no. DE-FG02-06ER15762), the National Institutes of Health ARRA (grant no. R01 GM068664-05A1Z), and the Army Research Office (grant no. W911NF0810227). A.R. thanks the DFG for a postdoctoral research fellowship. We also acknowledge the Gordon and Betty Moore Foundation for the support of the Molecular Observatory at Caltech. Operations at SSRRL are supported by the US DOE and NIH.

Keywords: biocatalysis · biomimetic synthesis · directed evolution · enzymes · metabolism


Received: December 19, 2010
Published online on April 21, 2011