Expansion of Substrate Specificity of Cytochrome P450 2A6 by Random and Site-directed Mutagenesis

Zhong-Liu Wu, Larissa M. Podust, and F. Peter Guengerich

From the Department of Biochemistry and the Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146

The natural product indole is a substrate for cytochrome P450 2A6. Mutagenesis of P450 2A6 was done to expand its capability in the oxidation of bulky substituted indole compounds, which are not substrates for the wild-type enzyme or the double mutant L240C/N297Q, as determined in our previous work (Wu, Z.-L., Aryal, P., Lozac'h, O., Meijer, L., and Guengerich, F. P. (2005) Chem. Biodivers. 2, 51–65). Error-prone PCR and site-directed mutagenesis led to the identification of two critical amino acid residue changes (N297Q and I300V) that achieve the purpose. The new mutant (N297Q/I300V) was able to oxidize both 4- and 5-benzyl-oxy(OBzl)indoles to form colored products. Both changes were required for oxidation of these bulky substrates. The colored product derived from 5-OBzl-indole was mainly 5,5'-di-OBzl-indirubin, whereas the dominant blue dye isolated upon incubations with 4-OBzl-indole was neither an indigo nor an indirubin. Two-dimensional NMR experiments led to assignment of the structure as 4-OBzl-2-(4'-OBzl-1',7'-dihydro-7'-oxo-6'H-indol-6'-ylidene)indolin-3-one, in which a pyrrole ring and a benzene ring are connected with a double bond instead of the pyrrole-pyrrole connection of other indigoinds. Monomeric oxidation products were also isolated and characterized; three phenols (4-OBzl-1H-indol-5-ol, 4-OBzl-1H-indol-6-ol, and 4-OBzl-1H-indol-7-ol) and one quinone (4-OBzl-1H-indole-6,7-dione, the postulated immediate precursor of the final blue dye) were identified. The results are interpreted in the context of a crystal structure of a P450 2A6-coumarin complex. The I300Q change opens an additional pocket to accommodate the OBzl bulk. The N2297Q change is postulated to generate a hydrogen bond between Glu and the substrate oxygen. Thus, the substrate specificity of P450 2A6 was expanded, and new products were obtained in this study.

Microsomal P450 enzymes (also termed “heme thiolate P450” (1)) are well known for their remarkable capabilities in the catalysis of diverse oxygenation reactions (2–6). These enzymes have been studied mostly as the principal catalysts involved in sterol synthesis, drug metabolism, and xenobiotic disposition. The use of P450 enzymes in the area of biocatalysis and fine chemical production is still largely unexploited (7). Random mutagenesis is one of the main approaches in terms of developing P450 enzymes with new functions and can also be used to enhance the knowledge of structure-function relationships of these enzymes and thus provide more information for drug development.

Random mutagenesis and molecular breeding approaches can be very useful when there is no crystal structure available to perform rational design, which is still the case for most P450 enzymes today. Furthermore, even rational designs based on crystal structures do have limitations because the enzymes possess both rigidity and flexibility. Predictions are even more complex in the case of oxidoreductase-catalyzed reactions when several factors are involved in the catalytic cycle. In addition to the active site (where substrate is bound) that is obvious in the crystal structure, other residues in the protein may also play important roles. With the availability of various high throughput screening methods and robotic systems, random mutagenesis has become a widely used tool in protein engineering both to improve catalytic efficiency and to investigate structure-function relationships (8). For P450 enzymes, random mutagenesis and molecular breeding are emerging areas. Bacterial P450 102A1 has been used as a starting point to develop catalysts that can hydroxylate alkane in a regio- and enantioselective manner (9) or that have special features (10). This laboratory has developed P450 1A2 using several strategies to improve the catalytic efficiency toward several substrates (11–13).

Molecular breeding of P450 2A6 was done in this laboratory (14), and a double mutant (L240C/N297Q) was selected from libraries constructed by random mutagenesis using randomized primers, each covering four positions of one of the substrate recognition site regions (2), and a staggered extension process method (15). The screening was based on the metabolism of indole, a substrate of P450 2A6 (16, 17) that can be hydroxylated to a product (indoxyl) that dimerizes to the blue compound indigo. The activity of the selected mutant was enhanced toward indole as well as some substituted indoles. This mutant has been used successfully to produce various substituted indigoinds with enhanced activities as potent kinase inhibitors (18, 19).

In the course of previous work on the biotransformation of 45 indole compounds catalyzed by the L240C/N297Q mutant, we found that this mutant showed limitations in using some substrates, among which were those bearing bulky groups (19). In the x-ray crystal structure of the P450 2A6-coumarin complex recently determined by Johnson and coworkers (20), the active site of P450 2A6 contains large aromatic residues that reduce the volume of the substrate-binding site. The active site is ~6-fold smaller than that of P450 2C8 (Protein Data Bank code 1PQ2) (21), consistent with the size of typical substrates such as coumarin (22), nicotine (23, 24), and indole.

We are interested in expanding the substrate specificity of P450 2A6.

1 To whom correspondence should be addressed: Dept. of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, 638 Robinson Research Bldg., 23rd and Pierce Aves., Nashville, TN 37232-0146. Tel.: 615-322-2261; Fax: 615-322-3141; E-mail: f.guengerich@vanderbilt.edu.

2 The abbreviations used are: P450, cytochrome P450; OBzl, OBn, and OCH3Ph, benzyl-oxy; HPLC, high performance liquid chromatography; APCC, atmospheric pressure chemical ionization; NOESY, nuclear Overhauser effect correlation spectroscopy; HMBc, heteronuclear multiple bond correlation; CIGAR, constant time inverse-detected gradient accordounced-resolved long-range; WT, wild-type; MS, mass spectrometry.

Received for publication, July 27, 2005, and in revised form, September 8, 2005. Published, JBC Papers in Press, October 7, 2005, DOI 10.1074/jbc.M508182200
using mutagenesis methods to augment the active pocket so that it can accommodate indoles with bulky substitution groups to increase the diversity of libraries of products generated for use as chemosats for screening of biological activity (19). In this work, error-prone PCR was used to construct randomized libraries, and a colorimetric colony-based method was used in the screening of mutants. Two critical amino acid residue changes that allow oxidation of both 4- and 5-benzyl-oxyl(Obz1) indoles to form colored products were identified. The results are interpreted in the context of the x-ray crystal structure (20).

**EXPERIMENTAL PROCEDURES**

**Chemicals**—Indole, 4-chloroindole, and 4- and 5-Obzl-indoles were purchased from Aldrich or Lancaster Synthesis (Windham, NH). All other reagents and solvents were obtained from general commercial suppliers and used without further purification.

**Spectroscopy**—UV-visible spectra were recorded either with a Hewlett-Packard 1040A diode array detector (on-line with HPLC) or in Me2SO or H2O using a Cary 14/OLIS or Aminco DW2a/OLIS spectrophotometer for quantitative studies. Mass spectra were obtained in the Vanderbilt facility either on-line with HPLC or via direct infusion using a Finnigan TSQ 7000 triple quadrupole mass spectrometer (Thermo Electron Corp., San Jose, CA) equipped with a standard API-1 atmospheric pressure chemical ionization (APCI) source in the positive or negative ion mode. N2 was used as a sheath gas (50 p.s.i.). The vaporizer temperature was set to 500 °C, and the corona current was maintained at 5 μA. The capillary was set at 220 °C and 25 V (or −25 V in the negative mode). The lens tube voltage was set to 80 V (or to −30 V in tandem mass spectrometry experiments. Data acquisition and spectral analysis were conducted with Finnigan ICIS software on a Digital Equipment Alpha workstation.

All NMR experiments were carried out in the Vanderbilt facility at 298 K with Bruker Avance™ DRX 500- and 400-MHz instruments. Samples were dissolved in d6-Me2SO or CD3CN, and semiquantitative estimates of the amounts of indigoids were made by adding known amounts of CH3CO2H and referencing to this internal –CH3 signal (21). The tube lens voltage was set to 80 V (or to −30 V in tandem mass spectrometry experiments. Data acquisition and spectral analysis were conducted with Finnigan ICIS software on a Digital Equipment Alpha workstation.

Construction of P450 2A6 Libraries—The open reading frame region of the P450 2A6 gene (1.5 kb) was mutated using a low fidelity PCR method and amplified in a 50-μL PCR mixture containing 250 ng of pCW2A6bc-L240C/N297Q or pCW2A6bc-WT plasmid, 120 ng of forward primer 5′-TAGGAGGTCATATGGCTGT-3′ and reverse primer 5′-ATTCTTAGACGGGAGCTT-3′, 2.5 units of Mutazyme® DNA polymerase (Stratagene, La Jolla, CA), 0.2 mM each dNTP, and 10 X Mutazyme® reaction buffer (Stratagene). The 1.5-kb amplified PCR library fragment was purified by agarose gel electrophoresis and cloned into the pCW2A6bc vector using the NdeI and XbaI restriction sites. The ligation mixture was transformed into E. coli DH10B ultracompetent cells (Invitrogen) by electroporation. The library plasmid DNA was purified using a QIAprep® miniprep kit (Qiagen Inc.).

**Phenotypic Selection Based on Color Formation**—The bicycronic library DNA was transformed into the E. coli trna− strain by heat shock at 41 °C, plated on large bioassay Q tray plates, and allowed to grow at 30 °C for 2 days. Individual colonies with blue color were picked and grown in 3 ml of LB medium containing ampicillin (100 μg/ml) as starter cultures for 15 h at 32 °C (see Fig. 1). Plasmid DNA was prepared from the cultures using a QIAprep® miniprep kit. Verification of the new mutants was performed by inoculating the starter cultures into 1 ml of Terrific Broth expression medium (1%, v/v) fortified with 1 mM 5-Obzl-indole on 24-well plates. The plates were then incubated at 30 °C for 30 h with gyration shacking at 250 rpm. Clones able to transform 5-Obzl-indole to form blue cultures were selected.

**Nucleotide Sequence Analysis**—Sequencing of the plasmid DNA in the clones that produced blue cultures (following verification) was performed in the Vanderbilt facility using an Applied Biosystems Model 3700 fluorescence sequencing unit with a Taq dye terminator kit (Applied Biosystems, Foster City, CA). The mutated sequences were verified by comparing the sequences of the sense and antisense strands (entire open reading frame).

**Construction of Site-directed Mutants**—Site-directed mutagenesis was conducted using the QuikChange® mutagenesis kit (Stratagene) following the supplier’s protocol. The primers were designed so that the mutation position was close to the middle with −10 to −15 bases of template complementary sequence on both sides and with a melting temperature of ≥75 °C (supplemental data, Table 1S). For practical purpose, a Bluescript SK(+) vector of the selected P450 2A6 mutant DNA (Bn1; −4.5 kb) was first constructed and subjected to site-directed mutagenesis (see Fig. 1). The 1.5-kb mutated fragment was subcloned into the pCW2A6bc vector (~8.5 kb) using the NdeI and XbaI restriction sites. The ligation mixture was transformed into E. coli, and positive clones were confirmed following purification of plasmids. The constructed site-directed mutant plasmids were verified by sequencing both strands of DNA (see above).

**Whole Cell Assay for Oxidation of Indoles**—Cells were incubated at 28 °C for 14 h with gyration shacking at 200 rpm. Cell pellets containing...
Expansion of Substrate Specificity of P450 2A6

expressed mutant P450 2A6 and NADPH-cytochrome P450 reductase were harvested from 500 ml of Terrific Broth expression medium of the E. coli trnA- strain by centrifugation at 5000 × g for 20 min, washed with 200 ml of M9 minimal medium (28), resuspended in 100 ml of M9 minimal medium, and used directly (after measuring the P450 concentration). The reaction mixtures consisted of 200 pmol of P450 (whole cells), 100 mM potassium phosphate buffer (pH 7.4), and varying concentrations of substrate in a total volume of 1 ml. The reactions were carried out at 37 °C in 24-well plates with gyratory shaking at 280 rpm and quenched by the addition of 0.1 ml of 0.5 M aqueous NaOH after 30 min (for indole and 4-chloroindole) or after 100 min (for 4- and 5-OBzI-indoles). The formation of colored products (indigo and indirubin) was measured by recording absorbance spectra in the visible region, and the estimated concentrations were calculated based on the extinction coefficient of the major product (subsequently determined). Estimates of $k_{cat}$ and $K_m$ were made using nonlinear regression with GraphPad Prism software.

Preparation of Membranes and Purification of Recombinant WT P450 2A6 and Mutants—Bacterial inner membranes containing mutant P450 2A6 and NADPH-cytochrome P450 reductase were isolated from 500 ml of Terrific Broth expression medium of the E. coli trnA- strain. The pellets were resuspended in a final volume of 10 ml of 50 mM Tris acetate (pH 7.4) containing 250 mM sucrose and 0.25 mM P450 2A6 and Mutants.

The coordinates of the P450 2A6 mutant were built by introducing substitutions into the WT P450 2A6 coordinates. The course of experiments is briefly described. WT P450 2A6 was subjected to random mutagenesis and a staggered extension process (StEP) (3), with selection using enhancement of the ability to convert indole to blue indigo. The result was the double mutant L240C/N297Q (DM) (14). This mutant served as the starting material in this study. The cDNA was subjected to error-prone PCR, and the resulting proteins were screened for their ability to convert 5-OBzI (OBzI-indole) to colored products. The result was a mutant (Bn1) that was found to have a total of five mutations relative to WT P450 2A6 (I140M, L240C, N297Q, I300V, and I366V). Site-directed mutagenesis was used to produce all combinations of these substitutions and to establish which are necessary to allow P450 2A6 to oxidize the bulky 4- and 5-OBzI-indoles. The result was Bn1/125, with only two substitutions (N297Q and I300V). SRS, substrate recognition site.

FIGURE 1. Scheme for preparation and analysis of P450 2A6 mutants. The course of experiments is briefly described. WT P450 2A6 was subjected to random mutagenesis and a staggered extension process (StEP) (3), with selection using enhancement of the ability to convert indole to blue indigo. The result was the double mutant L240C/N297Q (DM) (14). This mutant served as the starting material in this study. The cDNA was subjected to error-prone PCR, and the resulting proteins were screened for their ability to convert 5-OBzI (OBzI-indole) to colored products. The result was a mutant (Bn1) that was found to have a total of five mutations relative to WT P450 2A6 (I140M, L240C, N297Q, I300V, and I366V). Site-directed mutagenesis was used to produce all combinations of these substitutions and to establish which are necessary to allow P450 2A6 to oxidize the bulky 4- and 5-OBzI-indoles. The result was Bn1/125, with only two substitutions (N297Q and I300V). SRS, substrate recognition site.

5 h at 30 °C with gyroratory shaking at 250 rpm. At that time, the substrate (4- or 5-OBzI-indole) in Me2SO solution (1 M stocks) was added to the medium (0.5 ml), and shaking was continued for 22 h at 30 °C.

The pelleted cells and supernatants were separated by centrifugation at 5000 × g for 30 min. The supernatant was extracted with ethyl acetate (3 × 200 ml) and dried with Na2SO4. The pellet (from 500 ml of bacterial culture) was stirred with 500 ml of (CH3)2CO and filtered through Celite® (18, 19). The ethyl acetate and (CH3)2CO extracts were combined and concentrated in vacuo, and the resulting solids were fractionated by semipreparative HPLC. HPLC was done on a Beckman Ultrasphere column (10 × 250 mm, 5 μm) with a CH3CN/H2O mobile phase at a flow rate of 3 ml/min with 40% (v/v) CH3CN (in H2O) for 10 min, a 40–60% CH3CN linear gradient over 20 min, a 60–80% CH3CN linear gradient over 10 min, and a 80–100% CH3CN linear gradient over 5 min. To separate compounds 5 and 6, HPLC conditions of 30% (v/v) CH3CN (in H2O) for 10 min and a 30–55% CH3CN linear gradient over 20 min were used with the same column. Detection was at 254 or 600 nm (single wavelength) or with the on-line HPLC/ diode array detector system. Peaks were collected manually, and CH3CN was removed with a rotary evaporator. The remaining aqueous phase was lyophilized to dryness. In analytical systems, a Beckman Ultrasphere HPLC column (4.6 × 250 mm, 5 μm) was used with a CH3CN/H2O or CH3OH/H2O mobile phase at a flow rate of 1 ml/min.

Substrate Docking and Modeling of the P450 2A6 N297Q/I300V Mutant—The coordinates of the P450 2A6 N297Q/I300V mutant were built by introducing substitutions into the WT P450 2A6 coordinates.
## TABLE ONE
Steady-state kinetic parameters for WT P450 2A6 and mutants in the biotransformation of indoles

**DM, double mutant.**

| P450 2A6 Mutations | Indole | 4-Cl-indole | 5-OBzl-indole | 4-OBzl-indole |
|--------------------|--------|-------------|--------------|--------------|
|                    | $k_{cat}$ | $K_m$ | $k_{cat}/K_m$ | $k_{cat}$ | $K_m$ | $k_{cat}/K_m$ | $k_{cat}$ | $K_m$ | $k_{cat}/K_m$ | $k_{cat}$ | $K_m$ | $k_{cat}/K_m$ |
| WT None            | 1.7 ± 0.1 | 157 ± 11 | 0.011 ± 0.001 | 1.2 ± 0.1 | 55 ± 5 | 0.022 ± 0.002 | — — | — — | — — | — — | — — |
| DM L240C/N297Q     | 27 ± 4 | 222 ± 66 | 0.12 ± 0.04 | 15 ± 2 | 215 ± 51 | 0.070 ± 0.019 | — — | — — | — — | — — | — — |
| Bnl V140M/L240C/N297Q/I300V/I366V | 12 ± 1 | 152 ± 23 | 0.079 ± 0.013 | 14 ± 1 | 186 ± 32 | 0.075 ± 0.014 | 0.28 ± 0.02 | 51 ± 13 | 0.0055 ± 0.0015 | 0.45 ± 0.01 | 47 ± 6 | 0.0096 ± 0.0012 |
| Bnl1/1b L240C/N297Q/I300V/I366V | 9.3 ± 1.0 | 149 ± 35 | 0.062 ± 0.016 | 7.0 ± 0.5 | 83 ± 16 | 0.084 ± 0.017 | 0.38 ± 0.02 | 43 ± 10 | 0.0088 ± 0.0021 | 0.48 ± 0.03 | 42 ± 12 | 0.011 ± 0.003 |
| Bnl1/2 V140M/N297Q/I300V/I366V | 5.2 ± 0.6 | 159 ± 38 | 0.033 ± 0.009 | 3.6 ± 0.3 | 87 ± 19 | 0.041 ± 0.009 | 0.50 ± 0.02 | 55 ± 7 | 0.0091 ± 0.0012 | 0.91 ± 0.09 | 108 ± 25 | 0.0084 ± 0.0021 |
| Bnl1/3 V140M/L240C/I300V/I366V | 2.8 ± 0.2 | 75 ± 14 | 0.037 ± 0.007 | 6.3 ± 0.5 | 276 ± 34 | 0.023 ± 0.003 | — — | — — | — — | — — | — — |
| Bnl1/4 V140M/L240C/N297Q/I300V/I366V | 131 ± 69 | 1769 ± 1055 | 0.074 ± 0.059 | 51 ± 19 | 1212 ± 552 | 0.042 ± 0.025 | — — | — — | — — | — — | — — |
| Bnl1/5 V140M/L240C/N297Q/I300V | 8.8 ± 0.9 | 96 ± 26 | 0.092 ± 0.027 | 3.4 ± 0.2 | 74 ± 15 | 0.046 ± 0.010 | 0.60 ± 0.04 | 54 ± 1 | 0.011 ± 0.001 | 1.5 ± 0.2 | 139 ± 36 | 0.011 ± 0.003 |
| Bnl1/12 N297Q/I300V/I366V | 8.7 ± 0.4 | 164 ± 17 | 0.053 ± 0.006 | 54 ± 58 | 3039 ± 3544 | 0.018 ± 0.028 | 1.7 ± 0.2 | 285 ± 63 | 0.0060 ± 0.0015 | 1.2 ± 0.1 | 122 ± 27 | 0.0098 ± 0.0023 |
| Bnl1/15 L240C/N297Q/I300V | 11.9 ± 0.6 | 62 ± 10 | 0.19 ± 0.03 | 43 ± 21 | 1620 ± 893 | 0.027 ± 0.020 | 3.0 ± 0.5 | 418 ± 98 | 0.0072 ± 0.0021 | 4.8 ± 0.6 | 486 ± 86 | 0.0099 ± 0.0021 |
| Bnl1/41 L240C/N297Q/I366V | 6.7 ± 0.6 | 129 ± 26 | 0.052 ± 0.011 | 14 ± 3 | 304 ± 98 | 0.046 ± 0.018 | — — | — — | — — | — — | — — |
| Bnl1/45 V140M/L240C/N297Q/I366V | 64.0 ± 4.0 | 71 ± 14 | 0.090 ± 0.019 | 8.0 ± 0.9 | 126 ± 33 | 0.063 ± 0.018 | — — | — — | — — | — — | — — |
| Bnl1/52 V140M/N297Q/I300V | 6.5 ± 0.3 | 56 ± 10 | 0.12 ± 0.02 | 8.9 ± 1.5 | 383 ± 97 | 0.023 ± 0.007 | 1.6 ± 0.2 | 288 ± 50 | 0.0056 ± 0.0012 | 5.3 ± 2.3 | 864 ± 472 | 0.0061 ± 0.0043 |
| Bnl1/125 N297Q/I300V | 18 ± 3 | 196 ± 59 | 0.092 ± 0.032 | 14 ± 3 | 652 ± 233 | 0.021 ± 0.011 | 1.7 ± 0.2 | 201 ± 45 | 0.0085 ± 0.0021 | 3.8 ± 0.6 | 449 ± 98 | 0.0085 ± 0.0023 |
| Bnl1/1253 I300V | 1.9 ± 0.1 | 145 ± 25 | 0.013 ± 0.002 | 1.9 ± 0.2 | 174 ± 37 | 0.011 ± 0.003 | — — | — — | — — | — — | — — |
| Bnl1/1254 N297Q | 13 ± 2 | 280 ± 59 | 0.046 ± 0.010 | 8.3 ± 0.9 | 260 ± 47 | 0.032 ± 0.007 | — — | — — | — — | — — | — — |

* —, no activity was detected.

b The numbers 1–5 indicate the positions that were recovered from the Bnl mutant corresponding to V140M, L240C, N297Q, I300V, and I366V in sequence.
FIGURE 2. Steady-state kinetics of oxidation of indole compounds by P450 2A6 mutants. A, 4-chloroindole as substrate; B, 4-OBzl-indole as substrate; C, 5-OBzl(OBn)-indole as substrate; D, 4-OBzl-indole as substrate. Each point presented is the mean ± S.D. (range) of duplicate assays. DM, double mutant L240C/N297Q.

(20) in silico using interactive graphics techniques in the program O (31), followed by manual docking of each of the two substrates into the binding site cavity. A rotamer of Gln capable of donating a hydrogen bond to the substrate bound in the active site was selected. Subsequent energy minimization (conjugate gradient method) was performed using a CNS software routine with no experimental energy term (32). Energy minimization resulted in molecular models of P450 2A6 with bound substrates having a root mean square deviation of 0.55 Å for C-α atoms compared with the original structure. Subsequently, ligand coordinates were removed, and protein coordinates were used as a rigid receptor for docking of these same substrates using the FlexX Suite module of Tripos SYBYL software (33), including the pharmacophore constraint of Gln297 in the enzyme active site as a hydrogen bond donor.

RESULTS

Library Screening—Two randomized libraries were constructed using error-prone PCR following the manufacturer’s instructions (Stratagene) to adjust the mutation rate to zero to three mutations/1000 bp using either WT P450 2A6 or the double mutant L240C/N297Q as a template (Fig. 1). The screening method was a convenient colorimetric colony-based method that allows rapid screening of thousands of colonies by simply plating them on LB expression plates fortified with the tryptophanase-negative strain (instead of Terrific Broth to decrease the background color. A trypto-

Colored Products Generated from 4-Chloroindole and 4- and 5-OBzl-indoles

For the small substrates indole and 4-chloroindole, all of these mutants (as well as WT P450 2A6) had catalytic activity; the original double mutant (L240C/N297Q) was one of the most efficient in terms of catalytic efficiencies. No dramatic enhancement of activity was found for indole or 4-chloroindole oxidation catalyzed by the new mutants, which was not surprising because the screening method used in this work focused specifically on 5-OBzl-indole biotransformation.

The mutations at the 3-position (N297Q) and 4-position (I300V) (Fig. 1) are critical in gaining activity for hydroxylation of bulky substituted substrates (TABLE ONE). Mutants containing both of these mutations (Bn1/1, Bn1/2, Bn1/5, etc.) all showed activity in colored product formation from bulky substituted substrates, whereas mutants lacking either of these (Bn1/3, Bn1/4, Bn1/41, etc.) showed no activity. Thus, a new double mutant, P450 2A6 N297Q/I300V (Bn1/125), was found to be the simplest mutant able to oxidize both 4- and 5-OBzl-indoles to form colored products. Only small differences were observed in catalytic efficiencies among these gain-of-activity mutants, although the $k_{cat}$ values do differ, with the Bn1 mutant having the lowest $k_{cat}$. 

L240C/N297Q/I300V/I366V by sequence analysis, with three more mutation positions added to the template DNA.

Critical Residue Identification and Kinetic Parameters for WT P450 2A6 and Mutants—To determine which mutation positions were critical in gaining the ability to oxidize 5-OBzl-indole, 13 site-directed mutants were constructed from the Bn1 mutant, with one to four mutation sites recovered (Fig. 1 and TABLE ONE, Bn1/1 to Bn1/125). The hydroxylation of indole (the native substrate), 4-chloroindole (bearing a small substituent), and 4- and 5-OBzl-indoles (bearing bulky substrates) was measured based on the formation of colored dimers of the substrates (16). The low solubility of the 4- and 5-OBzl-indole substrates in H2O made vigorous shaking a necessity to increase diffusion; therefore, whole cells containing WT and mutant P450 2A6 were used instead of membranes or purified enzymes, which were more fragile under these conditions. The amounts of colored products were calculated based on their extinction coefficients, which were obtained either from the literature (indigo) (35) or from the isolation, identification, and quantification of the products as described under “Characterization of Colored Products Generated from 4-Chloroindole and 4- and 5-OBzl-indoles” (see below).
Steady-state kinetic plots of the biotransformation of indole compounds catalyzed by Bn1/125, Bn1, Bn1/1253, Bn1/1254, the double mutant L240C/N297Q, and WT P450 2A6 are shown in Fig. 2. Convenient as the colorimetric method is, the caveat exists that those mutants that showed no color derived from 4- or 5-OBzl-indole might catalyze the oxidation of the substrates to products without absorbance in the visible range. To address this possibility, HPLC profiles (monitored at 254 nm) were analyzed with crude extracts from the biotransformation of 4- and 5-OBzl-indoles catalyzed by WT P450 2A6, the original double mutant L240C/N297Q, and from Bn1/125, enlarged HPLC profile of cell extracts from Bn1/125, and photographs showing the colored cell cultures. mAU, milli-absorbance units.

**Substrate Binding Affinities of WT P450 2A6 and Mutants**—The binding affinities of WT P450 2A6 and some mutants for indole and 5-OBzl-indole were estimated (TABLE TWO). In general, there was no correlation between the catalytic activities and the apparent $K_d$ values. WT P450 2A6 had the lowest catalytic efficiency toward indole among the enzymes tested; however, this enzyme showed the tightest binding of indole (apparent $K_d$ of 13 $\mu$M). For 5-OBzl-indole, which was the substrate for the Bn1 and Bn1/1 enzymes but not for WT P450 2A6, the $K_d$ values are extremely high for Bn1 (apparent $K_d$ of 710 $\mu$M) and Bn1/1 (apparent $K_d$ of 414 $\mu$M), but very low for the WT enzyme (apparent $K_d$ of 8 $\mu$M). These discrepancies suggest that some of the apparent binding may be nonproductive.

**Coumarin 7-Hydroxylation Activity**—Analysis of the rates of coumarin 7-hydroxylation with membranes prepared from WT P450 2A6 and the double mutants L240C/N297Q and N297Q/I300V (Bn1/125) indi-
cated that the $k_{cat}$ for both mutants was increased ∼5-fold over that for the WT enzyme, whereas the catalytic efficiency indicated by $k_{cat} / K_m$ was not enhanced due to an increase in $K_m$ (TABLE THREE).

Characterization of Colored Products Generated from 4-Chloroindole and 4- and 5-OBzl-indoles—The colored products from cell cultures fortified with indole, 4-chloroindole, and 4- and 5-OBzl-indoles were analyzed by HPLC at a wavelength of 600 nm. The main product derived from indole was indigo (compound 1), as established by comparison with standard material. 4-Chloroindole yielded 4′,4′-dichloroindigo (compound 2) as a single peak (Fig. 4). 5-OBzl-indole yielded mainly 5,5′-di-OBzl-indirubin (compound 3) (Fig. 4), an expected dimer based on our previous work (18, 19). The main product of 4-OBzl-indole was originally expected to be an indigo because the substituent group at the 3′-position makes formation of an indirubin unfavorable due to its interference with the carbonyl group at the 3′-position (19, 36). However, the dominant blue dye isolated upon incubations with 4-OBzl-indole was neither indigo nor indirubin, although it had the same molecular mass as that of the indigoids and a UV-visible spectrum similar to that of indigo. 1H NMR clearly showed an unsymmetrical structure with one group of two doublets and one triplet (hydrogens from a 4-substituted benzene ring) plus another group of a singlet and two triplets with small coupling constants (hydrogens at the 2- and 3-positions of the pyrrole ring), suggesting that a pyrrole ring and a benzene ring are connected with a double bond instead of a pyrrole-pyrole connection. 1H-H COSY, 1H-1H NOESY, 1H-C HMBCC, and 1H-C CIGAR-HMBC NMR experiments (supplemental data, Figs. 15–55) were used to assign the structure as 4-OBzl-2-′(4′-OBzl-1′,7′-dihydro-7′-oxo-6′-H-indol-6′-ylidene)indolin-3-one (compound 4) (Fig. 4).

The formation of this compound was as reproducible as that of the other indigoids. The yield of this dye from cultures (13 μmol/liter of culture) is at the same level as that of 5,5′-di-OBzl-indirubin (9 μmol/liter of culture) by the same mutant, P450 2A6 N297Q/I300V (Bn1/125), and similar to that reported in our previous work on 6-nitroindole and methyl indole-5-carboxylate converted by P450 2A6 L240C/
Peaks with qualifying mass spectra and reasonable abundance were collected by preparative HPLC and analyzed in detail. Compounds 5 and 6 were originally eluted together ($t_R = 15.4$ min), but separated under modified elution conditions ($t_R = 24.9$ and $26.7$ min, respectively) (see "Large Scale Biotransformation and Product Separation" under "Experimental Procedures"). Analysis of the metabolites derived from 4-OBzI-indole showed that the benzene ring was hydroxylated at all three positions (positions 5–7). Three phenols, 4-OBzI-1H-indol-5-ol (compound 8), 4-OBzI-1H-indol-6-ol (compound 9), and 4-OBzI-1H-indol-7-ol (compound 10), and one quinone, 4-OBzI-1H-indole-6,7-dione (compound 6; which could be the immediate precursor of the final blue dye), were isolated, and their structures were established by MS, $^1$H NMR, $^1$H-1H COSY, and $^1$H-1H NOESY methods (supplemental data, Figs. 6S–9S). Two other products derived from hydroxylation on the pyrrole ring, 4-OBzI-indolin-2-one (compound 7) and 4-OBzI-indoline-2,3-dione (compound 5), were also isolated, as expected (Fig. 5). The characterization data for these eight compounds are included in supplemental "Analytical Data for Monomeric Metabolites Derived from 4-OBn Indole and 5-OBn Indole" (where "OBn" is benzyloxy).

Although there were several dimers detected by liquid chromatography/tandem mass spectrometry and $^1$H NMR, the amounts were insufficient for full characterization. 4-OBzI-indolin-2-one (compound 7) and 5-OBzI-indolin-2-one (compound 11) appear to be the primary oxidation products formed from each corresponding substituted indole, consistent with our previous study of indole oxidation (17).

Postulated Mechanism for 4-OBzI-indole Biotransformation—On the basis of the above results and previous work (17), we propose that the mechanism involved in the formation of compound 4 follows the course shown in Fig. 6. The substrate 4-OBzI-indole is first hydroxylated at one of the five free positions to form compounds 7–10 and an indoxyl (compound 13a), which undergoes further oxidation to form isatin (compound 5) via a tautomer (compound 13b) and an intermediate compound (compound 13c). Accordingly, the primary compounds 9 and 10 can be oxidized to give compound 6, possibly via radical intermediates (compounds 9a and 10a, respectively). Alternatively, oxidation of compounds 9 and 10 to catechols can be proposed, followed by 2-electron oxidation to the o-quinone (compound 6). The final blue dimer (compound 4) then could be the result of combinational reaction of compounds 13c and 6 or compounds 10a and 5.

Substrate Docking and Modeling of the P450 2A6 N297Q/I300V Mutant—Analysis of the active site in the crystal structure of the P450 2A6-coumarin complex indicates that coumarin occupies a cavity of ~359 Å$^3$. The single mutation I300V, despite elimination of only one methylene group in Val compared with Leu, results in expansion of the active site up to 359 Å$^3$. The longer side chain of glutamine presumably enables a new hydrogen bond with the oxygen atom in the substrate.
Expansion of Substrate Specificity of P450 2A6

Although exhaustive systematic screening of the WT P450 2A6 and P450 2A6 L240C/N297Q (the original double mutant) libraries to cover 95% possible changes (38) was not performed, this strategy used here was successful. Searching for gain-of-activity mutants, for which a low frequency is expected, using the colony-based method described here should be considered more of a selection than a screening process. All blue colonies definitely indicate a gain-of-activity mutant. Any incorrect identification resulted from the uncertainty in reading the color as “blue” due to the disturbance of background color and the brown color of the colonies, which made the second confirmation necessary.

From the results of site-directed mutagenesis, we know that one of the critical changes is the N297Q mutation, an existing mutation in the original double mutant enzyme. This mutant (L240C/N297Q) had been selected previously in a molecular breeding experiment (14) that was based on the use of randomized oligonucleotides (39) targeted to the substrate recognition site regions expected to comprise the active site (2). The nucleotide change is from AAC to CAA, a complex double transversion that would not likely be achieved by error-prone PCR. This could be the main reason why no gain-of-activity mutants were found in the WT P450 2A6 library.

Analysis of the N297Q and I300V substitutions in silico using the available crystal structure revealed that, although the I300V mutation adds ~83 Å³ to the wild-type binding site, resulting in 440 Å³ of total space, it is not enough to accommodate either of the OBzI-substituted indoles used in this study (which would require a volume of ~600 Å³). Presumably substrate-induced fit adjustment to the enzyme active site is required to complement the favorable steric and electronic effects of both mutations as deduced from the substrate-docking model. Alternatively, the active site of the N297Q/I300V mutant may have changed enough from that of the WT enzyme to expand the space, although we prefer the former hypothesis regarding induced fit.

The compound 4 colored dimer appears to be novel in that no such structure has been reported. In our previous work, ~20 indigoids were isolated from P450 2A6-catalyzed reactions (18, 19). All of them have parent rings of indigo or indirubin, in which both positions of the pyr-
role ring (positions 2 and 3) are oxidized and one of the positions is connected to another monomer, leaving the benzene ring intact, consistent with the reactivity of the indole. Indole is a π-excessive heteroaromatic, which means that the π-electron densities on carbon atoms of the pyrrole ring are greater than those of the benzene, and the pyrrole ring should be more readily oxidized. The most important resonance structures of indole maintain an unbroken benzene ring and bear the positive charge on nitrogen and the negative charge on the C-3 atom. The 3-position of the indole ring is highly reactive toward electrophilic reagents (40). Although it has been reported that indole can be hydroxylated at the 5- and 6-positions in P450-catalyzed oxidation systems, the amounts of these products are limited, and they apparently do not undergo further oxidation to form dimers (17, 41). The intermediate 4-OBzl-1H-indole-6,7-dione (compound 6) does have analogs in nature, e.g. the redox cofactors in methyamine dehydrogenase (tryptophan tryptophylquinone) (42) and amine dehydrogenase (cysteine tryptophylquinone) (43), which are also 6,7-diones derived from L-tryptophan and formed during post-translational modification. How the two quinone oxygens are incorporated into tryptophan tryptophylquinone is still unclear, although it is believed that other enzymes, probably monoxygenases, are required for the biogenesis instead of a self-processing mechanism (44). Tryptophan tryptophylquinone could be considered as an indole dimer as well in that the pyrrole ring (C-2) of Trp108 is connected to the benzene ring (C-4) of Trp108 with a single bond. Another report on the oxidation of indole derivatives includes structures involving formation of dimers with the benzene-benzene ring connected; in this case, however, the 3-position of pyrrole was protected and the benzene ring was dioxidized in advance (45).

The isolation of six intermediates upon incubation with 4-OBzl-indole, especially those derived from hydroxylation on the benzene ring, not only supports the formation of compound 4, but also clearly indicates that the substrate is accommodated inside the active site with

FIGURE 8. Stereo views of 4- and 5-OBzl-indoles docked into the active site of the P450 2A6 N297Q/I300V mutant. Substrates are shown in yellow in the ball-and-stick mode. A, binding of 5-OBzl-indole. Coumarin (from the original P450 2A6 crystal structure) is shown in green. B and C, binding of 4-OBzl-indole in two different orientations. The figure was generated using the program SETOR.
variable orientations. On the other hand, the gain-of-activity mutants retained their normal activities (producing indigo or indirubin), like WT P450 2A6, with the smaller indole compounds tested (indole and 4-chloroindole), which are presumed to be more flexible in the active site. The lack of oxidation of the benzene ring of indole by the mutant might be explained partially by the electron-donating property of the OBzI group. The 4-OBzI substituent may activate the para-position to be hydroxylated to compound 8, which contributes to the whole process (Fig. 6). In addition, the rigidity of substrate docking of a bulky substrate compared with a substrate of smaller size may also contribute to the increased chance of hydroxylation on the benzene ring, as long as the substrate can be accommodated properly inside the active site with the benzene ring presented to the active oxygen species.

In conclusion, we have isolated new mutants of P450 2A6 with expanded substrate specificity toward bulky substituted indoles from a randomized library constructed by error-prone PCR and identified the critical residues that contribute to the effect by site-directed mutagenesis. The products (dimers and monomeric oxidation products) were compared with a substrate of smaller size may also contribute to the expansion of Substrate Specificity of P450 2A6.

Acknowledgments—We thank E. F. Johnson and J. K. Yano for sharing the assistance in preparation of the manuscript.

REFERENCES

1. Palmer, G., and Reedijk, J. (1992) J. Biol. Chem. 267, 665–677
2. Gotoh, O. (1992) J. Biol. Chem. 267, 83–90
3. Zhao, H., Giver, L., Shao, Z., Affholter, J. A., and Arnold, F. H. (1998) Nat. Biotechnol. 16, 258–261
4. Ortiz de Montellano, P. R. (ed) (2005) Cytochrome P450 Structure, Mechanism, and Biochemistry, 3rd Ed., Kluwer Academic/Plenum Publishers, New York
5. Guengerich, F. P. (2001) Chem. Res. Toxicol. 14, 611–650
6. Guengerich, F. P. (2005) in Cytochrome P450 Structure, Mechanism, and Biochemistry (Ortiz de Montellano, P. R., ed) 3rd Ed., pp. 377–531, Kluwer Academic/Plenum Publishers, New York
7. Guengerich, F. P. (2002) Nat. Rev. Drug Discov. 1, 359–366
8. Arnold, F. H., and Georgiou, G. (2003) Methods Mol. Biol. 230
9. Peters, M. W., Meinhold, P., Glieder, A., and Arnold, F. H. (2003) J. Am. Chem. Soc. 125, 13442–13450
10. Seng Wong, T., Arnold, F. H., and Schwaneberg, U. (2004) Biotechnol. Bioeng. 85, 351–358
11. Parikh, A., Josephy, P. D., and Guengerich, F. P. (1999) Biochemistry 38, 5283–5289
12. Yun, C.-H., Miller, G. P., and Guengerich, F. P. (2000) Biochemistry 39, 11319–11329
13. Kim, D., and Guengerich, F. P. (2004) Biochemistry 43, 981–988
14. Nakamura, K., Martin, M. V., and Guengerich, F. P. (2001) Arch. Biochem. Biophys. 395, 25–31
15. Stemmer, W. P. (1994) Nature 370, 389–391
16. Gillam, E. J., Aguinaldo, A. M., Notley, L. M., Kim, D., Mundkowski, R. G., Volkov, A., Arnold, F. H., Soucek, P., Devoss, J. J., and Guengerich, F. P. (1999) Biochem. Biophys. Res. Commun. 265, 469–472
17. Gillam, E. J. M., Notley, L. M., Cai, H., Devoss, J. J., and Guengerich, F. P. (2000) Biochemistry 39, 13817–13824
18. Guengerich, F. P., Sorrells, J. L., Schmitt, S., Krauser, J. A., Aryal, P., and Meijer, L. (2004) J. Med. Chem. 47, 3236–3241
19. Wu, Z.-L, Aryal, P., Ozog, O., Meijer, L., and Guengerich, F. P. (2005) J. Biol. Chem. 280, 13442–13450
20. Sambrook, J., and Russell, D. W. (2001) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21. Schoch, G. A., Yano, J. K., Wester, M. R., Griffin, K. J., Stout, C. D., and Johnson, E. F. (2005) Nat. Struct. Mol. Biol. 12, 822–823
22. Hecht, S. S., Hochalter, J. B., Villalta, P. W., and Murphy, S. E. (2000) Proc. Natl. Acad. Sci. U.S.A. 97, 12493–12497
23. Source, P. (1999) Arch. Biochem. Biophys. 370, 190–200
24. Parikh, A., Gillam, E. M. J., and Guengerich, F. P. (1997) Nat. Biotechnol. 15, 784–788
25. Soucek, P. (1999) J. Biol. Chem. 274, 9497–9503
26. Parikh, A., Gillam, E. M. J., and Guengerich, F. P. (1997) Nat. Biotechnol. 15, 784–788
27. Sandhu, P., Baba, T., and Guengerich, F. P. (1993) Arch. Biochem. Biophys. 306, 443–450
28. Schenken, J. R., Remmer, H., and Estabrook, R. W. (1967) Mol. Pharmacol. 3, 113–123
29. Guengerich, F. P. (2001) in Principles and Methods of Toxicology (Hayes, A. W., ed) 4th Ed., pp. 1625–1687, Taylor & Francis, Philadelphia
30. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
31. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grissle-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Milges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
32. Rarey, M., Kramer, B., Lengauer, T., and Klebe, G. (1996) J. Mol. Biol. 261, 470–489
33. Liu, Q. S., Schwaneberg, U., Fischer, P., and Schneid, R. D. (2000) Chem. Eur. J. 6, 1531–1536
34. Polychronopoulos, P., Magiatis, P., Skaltsounis, A. L., Miyranthopoulos, V., Mikros, E., Tarricone, A., Musacchio, A., Roe, S. M., Pearl, L., Leost, M., Greengard, P., and Meijer, L. (2004) J. Med. Chem. 47, 935–946
35. Steinmetz, J., and Wyman, G. M. (1956) Proc. Natl. Acad. Sci. U.S.A. 53, 957–963
36. Cirino, P. C., Mayer, K. M., and Umeno, D. (2003) Mol. Pharmacol. 611–650
37. Tanizawa, K., Ikeda, T., Kano, K., and Mathews, F. S. (2001) J. Mol. Biol. 267, 469–472
38. Ikeda, T., Ikeda, T., Kano, K., Tanizawa, K., and Mathews, F. S. (2001) J. Biol. Chem. 276, 265, 41428–14273
39. Pearson, A. R., De La Mora-Rey, T., Graichen, M. E., Wang, Y., Jones, L. H., Marimanikumpul, S., Agger, S. A., Grimsrud, P. A., Davidson, V. L., and Wilmot, C. M. (2004) Biochemistry 43, 5494–5502
40. Tabatabaie, T., and Dryhurst, G. (1992) J. Med. Chem. 35, 2261–2274