0090-9556/07/3512-2177–2185$20.00
DRUG METABOLISM AND DISPOSITION
Vol. 35, No. 12
17939/3274869
DMD 35:2177–2185, 2007
*Printed in U.S.A.*




# A Shuffled CYP1A Library Shows Both Structural Integrity and Functional Diversity

Wayne A. Johnston, Weiliang Huang, James J. De Voss, Martin A. Hayes, and Elizabeth M.J. Gillam

*Physiology and Pharmacology, School of Biomedical Sciences (W.A.J., W.H., E.M.J.G.) and Chemistry, School of Molecular and Microbial Sciences (J.J.D.V.), University of Queensland, St. Lucia, Australia; and Discovery Drug Metabolism and Pharmacokinetics & Bioanalytical Chemistry Department, AstraZeneca R&D Mölndal, Mölndal, Sweden (M.A.H.)*

Received July 29, 2007; accepted September 5, 2007

**ABSTRACT:**

**The cytochrome P450 enzymes (P450s) that mediate mammalian xenobiotic metabolism are highly versatile monooxygenases, which show wide and overlapping substrate ranges but generally poor catalytic rates. Re-engineering of these P450s may enable the development of useful biocatalysts for industrial applications. In the current study, restriction enzyme-mediated DNA family shuffling was used to create a library from human CYP1A1 and CYP1A2. Among sequenced clones (four randomly selected and eight functional clones), 5.9 ± 2.3 crossovers and 1.5 ± 1.5 spontaneous mutations (mean ± S.D.) were detected per mutant. A high level of structural integrity as well as diverse functionality were found, with 53% of clones expressed at significant levels (>50 nM P450 hemoprotein) and 23% of clones showing activity on one or more of the following compounds: luciferin 6′-chloroethyl ether (luciferin-CEE), luciferin 6′-methyl ether (luciferin-ME), 6′-deoxyluciferin (luciferin-H), the ethylene glycol ester of luciferin 6′-methyl ether, 7-ethoxyresorufin, and *p*-nitrophenol (PNP). Different activity profiles were seen with higher specific activity on individual compounds (e.g., clone 22; 9 times the CYP1A1 specific activity toward luciferin-CEE), novel activities (e.g., clone 35; activity toward luciferin-H and PNP), and broadening of substrate range observed in particular clones (e.g., clone 9; activity toward both selective substrates luciferin-ME and luciferin-CEE as well as toward luciferin-H and PNP). In summary, forms were found with distinct and novel activity profiles, despite the relatively small number of mutants examined. In addition, the whole-cell metabolic assays described here provide simple, high-throughput methods useful for screening larger libraries.**

Enzymes are exquisitely specific biocatalysts, able to modify even complex organic chemicals with a high degree of stereo- and regioselectivity by virtue of their complex, stereoselective, three-dimensional structures. Directed or artificial evolution is being used to engineer enzymes with novel catalytic and physicochemical properties that can be used as biocatalysts for industrial applications. The development of enzyme libraries relies upon the ability to navigate a practically infinite sequence space of possible enzymes while identifying novel and desired characteristics in library clones. DNA family shuffling is a powerful technique that allows sparse sampling of a large region of the sequence space, whereas techniques based on random mutagenesis either explore more limited sequence space close to the parental starting point or are prohibitively massive in scope (Crameri et al., 1998).

Cytochrome P450 (P450) enzymes are ideal starting points for directed evolution. The need for organisms to clear a diversity of xenobiotics has given several mammalian P450s the capability to metabolize a wide range of foreign chemicals via monooxygenation, with considerable regiospecificity on even large and complex molecules (Gillam, 2005). Despite this diversity of activity, the sequence homology of P450 subfamilies (>55%) is sufficiently close to allow DNA family shuffling. Although human CYP1A1 and CYP1A2 share 74% similarity at the sequence level, their function and expression profile are significantly different. CYP1A1 is primarily present in lung and other extrahepatic tissues and functions to monooxygenate polycyclic hydrocarbons and other relatively planar xenobiotics. CYP1A2 is primarily hepatic, and its substrate range also includes relatively planar substrates. CYP1A2 catalyzes diverse reactions including the oxidation of caffeine, fluorescent resorufins, and heterocyclic aromatic amines. Both are implicated in the activation of carcinogens (reviewed in Guengerich, 2005).

Random mutagenesis has been used for the directed evolution of

This study was supported by Australian Research Council Discovery Project DP210635 (to E.M.J.G. and J.J.D.V.) and by Australian Research Council Linkage Project LP560595 (to E.M.J.G., J.J.D.V., and M.A.H.) with AstraZeneca R&D Mölndal, Sweden.

This study was presented in preliminary form at the 16th International Symposium on Microsomes and Drug Oxidations, Budapest, Hungary, September 3 to 7, 2006.

Article, publication date, and citation information can be found at http://dmd.aspetjournals.org.

doi:10.1124/dmd.107.017939.

**ABBREVIATIONS:** P450, cytochrome P450; hNPR, human NADPH-cytochrome P450 reductase; luciferin-CEE, luciferin 6′-chloroethyl ether; luciferin-ME, luciferin 6′-methyl ether; luciferin-H, 6′-deoxyluciferin; luciferin ME-EGE, the ethylene glycol ester of luciferin 6′-methyl ether; luciferin-BE, luciferin 6′-benzyl ether; PNP, *p*-nitrophenol; 7-ER, 7-ethoxyresorufin; 7-PR, 7-pentoxyresorufin; PCR, polymerase chain reaction; EROD, 7-ethoxyresorufin *O*-deethylase; PROD, 7-pentoxyresorufin *O*-dependentylase.

CYP3A4 and CYP2B enzymes to alter substrate specificity, thermostability, and solvent tolerance (Kumar et al., 2005, 2006a,b). In the CYP1A subfamily, the activity of CYP1A2 toward 7-methoxyresorufin has been increased 5-fold via mutagenesis (Kim and Guengerich, 2004). Additionally, segment-directed mutagenesis of CYP1A1 (in an amino acid region with sequence diversity between CYP1A1 and CYP1A2) generated mutants with activity toward CYP1A-selective substrates, but the resulting mutants generally showed relative activities toward various substrates that were similar to the parental CYP1A1 (Taly et al., 2007). A CYP1A library was also created by DNA family shuffling using DNase I-mediated fragmentation of human CYP1A1 and CYP1A2 and screened for naphthalene hydroxylation, yielding approximately 20% of active clones (Abécassis et al., 2000). Further characterization of such libraries revealed mutants with significantly altered profiles compared with those of CYP1A1 and CYP1A2 (Abécassis et al., 2003; Taly et al., 2007).

One aim of the current study was to use restriction enzyme-mediated DNA shuffling (Kikuchi et al., 1999; Huang et al., 2007) to generate a hybrid CYP1A1/1A2 library and characterize the structural integrity and functional diversity of the resulting mutants. The technique of DNA shuffling was originally reported, with DNase I used to generate fragments for recombination (Stemmer, 1994), but more recently, restriction enzymes have been used to perform the cleavage (Kikuchi et al., 1999; Huang et al., 2007). Although restriction enzymes carry the disadvantage that cleavage is nonrandom (Gillam, 2005), this modification appears to allow greater recovery of functional chimeras and less parental contamination (Kikuchi et al., 1999; Huang et al., 2007).

Although diversification of sequences is fundamental to the success of directed evolution experiments, equally important is the efficient navigation of sequence-function space by high-throughput screening. This task is not trivial because few substrates of interest generate colored or fluorescent products (Gillam, 2005), and downstream cell processing to extract enzyme is generally necessary. Previously, high-throughput screening of chimeric P450 activity in our laboratory has relied upon detection of indigo production in cultures or processed cell extracts (Gillam et al., 1999). However one of the aims of the current work was to apply new whole-cell activity assays using luminogenic substrates specific for the parent forms (and other P450s), as well as traditional fluorescence- and absorbance-based metabolic assays for P450 activity, in rapid screening of libraries. A secondary aim was to couple these activity measurements to improved methods of whole-cell P450 measurement by CO-difference spectroscopy to simplify overall library screening protocols.

## Materials and Methods

**Materials.** Bicistronic bacterial expression plasmids containing cDNAs for the parental P450 forms, CYP1A1 and CYP1A2, designated pCW′/P450/human NADPH-cytochrome P450 reductase (hNPR), were prepared as described previously (Parikh et al., 1997). Both plasmids had been modified to encode a hexa-His tag at the C terminus of the P450. The chaperone expression vector (pGro7) was obtained as a generous gift from Professor K. Nishihara (HSP Research Institute, Kyoto, Japan) (Nishihara et al., 1998). The pCW′/hNPR expression vector encoding hNPR alone was obtained from Professor F.P. Guengerich (Vanderbilt University, Nashville, TN). Enzymes for molecular biology were obtained from New England Biolabs Inc. (Beverly, MA). The *Escherichia coli* strain DH5α F'IQ used in expression trials was obtained from Invitrogen (Mulgrave, Australia).

P450-GLO kits for measurement of CYP1A activity [CYP1A1, luciferin 6′-chloroethyl ether (luciferin-CEE); CYP1A2, luciferin 6′-methyl ether (luciferin-ME)] were a generous gift from Dr. James Cali (Promega, Madison, WI). Additional P450-GLO kits providing 6′-deoxyluciferin (luciferin-H), the ethylene glycol ester of luciferin 6′-methyl ether (luciferin ME-EGE), and luciferin 6′-benzyl ether (luciferin-BE) were purchased directly from Promega. Substrates *p*-nitrophenol (PNP), 7-ethoxyresorufin (7-ER), and 7-pentoxyresorufin (7-PR) were sourced from Molecular Probes Inc. (Eugene, OR). Difco Bactotryptone, Bacto-Yeast Extract, and Bactopeptone for culture media were obtained from BD Biosciences (San Jose, CA). Other chemicals were obtained from local suppliers at the highest quality available.

**Construction of a Shuffled CYP1A1/1A2 Library.** Shuffling of the parental CYP1A1 and 1A2 sequences was done according to the general method described by Huang et al. (2007). Specifically, parental sequences were fragmented with two separate combinations of restriction enzymes: 1) NlaIII plus Eco0109 and (ii) BbvI plus MboII. Digestion was performed in 200 μl at 37°C (2.5 h) on 30 μg of total DNA, using 30 units of each enzyme.

Fragments were recovered by filtration through Microcon PCR units (Millipore Corporation, Billerica, MA), followed by ethanol/sodium acetate precipitation and then subjected to reassembly in a primerless PCR. Re-assembled shuffled clones were amplified using the following primers: forward, placed immediately upstream of and including the NdeI site and start codon (underlined), 5′ GGAAACAGGATCCATCGATGCTTAGGAGGTCATATG 3′; and reverse, placed downstream of the P450 cDNA, 150 bp into the hNPR open reading frame, 5′ GAGGTCAATGTCTGAATTTTGGTGAACTCGGGGAC 3′. The amplified product was ligated into the NdeI and XbaI sites of similarly cut pCW′/2C9/hNPR, replacing the CYP2C9 coding sequence, and ligation mixtures were used to transform DH5αF'IQ cells already containing the pGro7 chaperone vector. Sixty-four separate clones were isolated on selective media, cultured overnight in Luria-Bertani medium containing ampicillin (100 μg/ml) and chloramphenicol (20 μg/ml), and stored in 15% glycerol at −80°C.

**Sequence Analysis of Mutant Clones.** Four randomly selected clones and eight clones showing significant enzyme activity as measured in initial metabolic screening (see below) were subjected to sequencing. Plasmids were isolated from mutant clones using QIAGEN miniprep kits (QIAGEN, Doncaster, Australia). Automated sequencing was performed using the Big Dye Terminator kit (version 3.1) at the Australian Genome Research Facility (Brisbane node; University of Queensland, St. Lucia, Australia). Library sequences were analyzed by automatic comparison of clone bases with the corresponding base for each library parent (CYP1A1 or CYP1A2) in a sequence alignment. A mutation was flagged when the clone base did not match the corresponding base in either parental sequence. Input sequence alignments were determined using Vector NTI (Invitrogen).

**Expression of Shuffled CYP1A1/1A2 Library Mutants.** Library clones were coexpressed with hNPR as described previously, using bacterial chaperones GroES and GroEL expressed using the pGro7 vector (Gillam et al., 1993; Notley et al., 2002). Library clones (15% glycerol cultures, maintained at −80°C) were core sampled without thawing using sterile pipette tips, used to inoculate 1.0-ml starter cultures (Luria-Bertani media containing 100 μg/ml ampicillin and 20 μg/ml chloramphenicol in 24-well plates), and grown overnight at 37°C with shaking at 300 rpm.

Triplicate expression cultures were conducted in 24-well plates (0.4–1.4-ml fill volumes) in Terrific broth media containing ampicillin (100 μg/ml), chloramphenicol (100 μg/ml), and thiamine (1 mM), with trace elements (Bauer and Shiloach, 1974). Wells were inoculated with starter cultures (5% v/v) and incubated at 25°C for 5 h with shaking at 400 rpm before induction. Recombinant protein expression was then induced by the addition of arabinose (4 mg/ml), δ-aminolevulinic acid (δ-ala: 0.5 mM), and isopropyl-β-thiogalactopyranoside (1 mM). After induction, Breath-easy polyurethane membranes (Diversified Biotech, Boston, MA) were applied to plates to restrict oxygen tension in the cultures. Library expression was performed in three separate batches of clones, with three library clones (9, 22, and 64) re-expressed in the second batch for comparison purposes, and the parents (CYP1A1 and CYP1A2) and hNPR (as positive and negative controls, respectively) expressed in all batches.

Cultures were harvested between 23 and 94 h by sedimentation at 3300*g* in plates, followed by resuspension in 750 μl of whole-cell assay buffer (100 mM potassium phosphate buffer, pH 7.4, 6 mM magnesium acetate, 10 mM dextrose). Resuspended cultures (200 μl) were transferred to 96-well plates for whole-cell P450 measurement. The remainder of each resuspended culture was stored at 4°C and subsequently assayed for activity toward various compounds, as follows: luciferin-CEE dechloroethylation, luciferin-ME demethylation, luciferin-H hydroxylation, luciferin-BE debenzylation, demethylation of lucife-

rin ME-EGE, 7-ER *O*-deethylation (EROD), 7-PR *O*-dependentylation (PROD), and PNP *o*-hydroxylation.

**Whole-Cell P450 CO-Difference Spectroscopy.** P450 concentration in library clones was measured in whole cells by Fe[II].CO versus Fe[II] difference spectroscopy using the method of Otey (2003), with quantification using a Spectramax M2 spectrophotometer (Molecular Devices, Sunnyvale, CA). A detailed account of the optimization of P450 quantification in whole cells using cultures of recombinant mammalian P450s will be presented elsewhere (Johnston WA, Huang W, De Voss JJ, Hayes MA, Gillam EMJ, unpublished data).

**Library Screening Using P450-GLO Luminogenic Substrates.** P450-GLO kits were adapted for measurement of activity in whole cells containing expressed CYP1As as follows. Luminogenic substrates were added directly to 50-$\mu$l suspensions of cells in whole-cell assay buffer at suggested kit concentrations (luciferin-CEE, 30 $\mu$M; luciferin-ME; 100 $\mu$M; luciferin-H, 100 $\mu$M; luciferin ME-EGE, 10 $\mu$M) and incubated directly (5–20 min, 350 rpm, 25°C) in 96-well white plates (Corning Life Sciences, Acton, MA). Luminescence was subsequently developed by addition of 50 $\mu$l of luciferin detection reagent and incubation for 20 min at 25°C. Final luminescence levels were quantified using a Microbeta Trilux 1450 LSC counter (Perkin Elmer, Wellesley, MA) and standardized against 50 $\mu$l of 0.1 $\mu$M luciferin and substrate-only standards included on each plate. Linearity of luciferin-ME and luciferin-CEE turnover with cell loading and time for the assay conditions used was verified for parents CYP1A1 and CYP1A2.

**Other Library Screening Assays.** PNP hydroxylation was measured on whole cells by the method of Winters and Cederbaum (1992), with a loading of 5 pmol P450 (in whole cells) per 100-$\mu$l reaction in 96-well plates. Cultures were incubated for 2 h with 250 rpm agitation. Product formation was measured via absorbance at 515 nm using a Spectramax M2 plate reader (Molecular Devices).

The measurement of EROD activity was adapted from the method of Chang and Waxman (1997). Specifically, harvested cell suspensions were incubated with 5 $\mu$M 7-ER in a total volume of 100 $\mu$l in 96-well plates for 20 min at 25°C and with 350 rpm agitation. Resorufin production was quantified by fluorescence using an excitation wavelength of 530 nm and an emission wavelength of 582 nm in a Spectramax M2 plate reader. Calibration was performed using in-plate standards of 1.25, 2.5, and 5 $\mu$M resorufin, which were spiked with whole cells expressing hNPR but no P450, with biomass (o.d. 600) matched to the sample biomass concentrations. Initially endpoint determination of fluorescence was conducted. However, due to problems with background fluorescence noise of whole-cell suspensions, which generated false-positive readings, continuous fluorescence measurement was subsequently used, and plates were shaken between reads (30-s interval over 30 min total) within the plate reader. Initial activity was calculated from fluorescence versus time curves. PROD activity was measured using the same protocol and 5 $\mu$M 7-PR.

Linearity of whole-cell EROD assays with variable cell loading was verified for the library parents CYP1A1 and CYP1A2, with linearity apparent to approximately 4 pmol P450/100 $\mu$l reaction in whole cells. Screening was subsequently performed with approximately 1 pmol P450/100 $\mu$l reaction for library clones. Specific EROD activity was calculated after post hoc normalization using P450 concentrations derived from whole-cell CO-difference spectra recorded for the relevant library clones.

## Results

**Sequence Analysis.** A library of mutant clones was constructed by DNA family shuffling of human CYP1A1 and CYP1A2 using fragmentation by: 1) NlaIII plus Eco0109 and 2) BbvI plus MboII. Fragments were recovered by size-selective filtration before reassembly in a primerless PCR reaction. A portion of the library (64 clones) was sampled and archived. Four clones randomly selected from the library (3, 24, 33, and 47) were subjected to sequencing, as well as a further eight clones (8, 9, 22, 23, 28, 32, 35, and 38) that proved to be of interest from the screening studies (see below). All clones sequenced encoded full-length P450 sequences, and no premature stop codons were observed. Analysis (Fig. 1) showed an average of 5.9 ± 2.3 (mean ± S.D.) crossovers and 1.5 ± 1.5 (mean ± S.D.) spontaneous mutations per mutant.

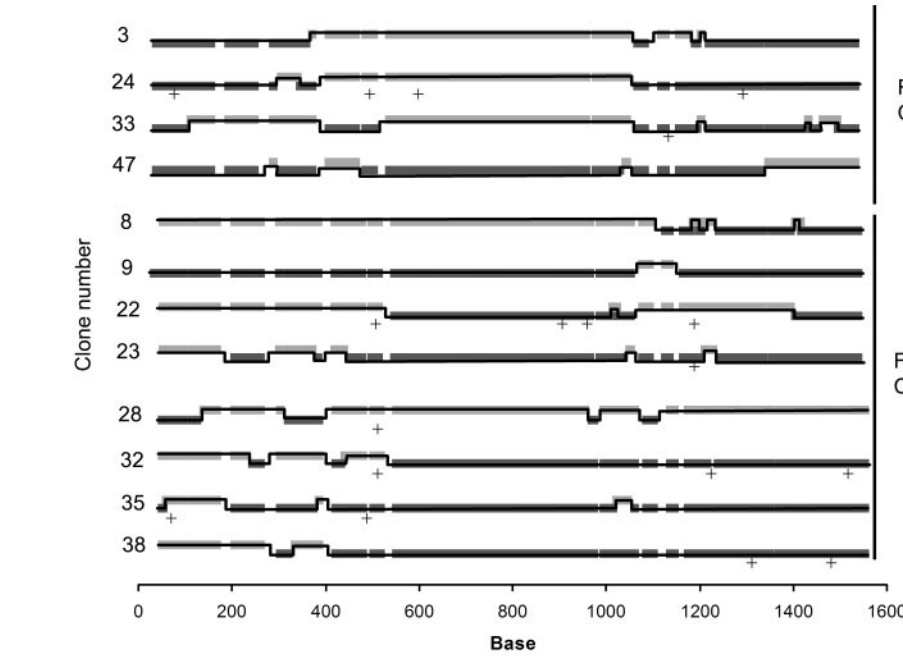
**Functional Characterization of CYP1A Shuffled Mutants.** The shuffled CYP1A library was investigated in three stages. First, the whole-cell P450 assay was used to establish optimal high-throughput culture techniques for the expression of the library parents CYP1A1 and CYP1A2. Second, the entire library was expressed and screened for P450 expression level, activity toward luciferin-ME, luciferin-CEE, and 7-ER using the optimized culture expression parameters determined in the first step. Third, a subset of clones demonstrating activity in the initial screening round were re-expressed and screened for activity toward luciferin ME-EGE, luciferin-H, 7-ER, and PNP.

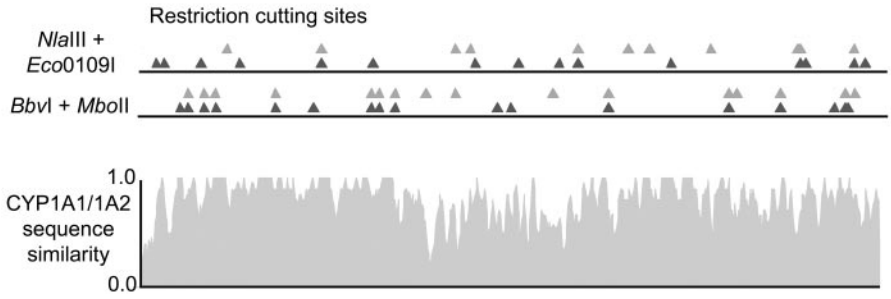
**Optimization of Screening Conditions for CYP1A Shuffled Mutants.** Expression was conducted in 24-well plates to facilitate multiple repeat measurements on the same replicate cultures. In Fig. 2, the effect of harvest time on P450 expression is presented. Time profiles for CYP1A1 and CYP1A2 were similar for both P450 expression level (detected by whole-cell Fe[II].CO versus Fe[II] spectra; Fig. 2A) and enzyme activity (detected as whole-cell activity on luciferin-CEE and luciferin-ME, respectively; Fig. 2B). The luminescent activity assays were more sensitive in detecting P450 at the initial stages of expression in both cases. Optimal harvest time was determined to be approximately 72 h for both CYP1A1 and CYP1A2, and this harvest time was used for subsequent screening of the library. However, this harvest time represents a "best guess"; it cannot be assumed that 72 h was optimal for all mutants, nor would this harvest time be necessarily optimal for other chimeric P450 libraries with different subfamily parents.

Figure 3 shows the effect of increasing initial 24-well fill volume (0.4–1.4 ml) on the expression level of CYP1A2 at 72 h of culture. Although the greatest final expressed P450 concentration was observed with the least well fill volume (Fig. 3A), total well P450 content was maximal at fill volumes greater than 1.0 ml (Fig. 3B). Because there was some indication of decreased enzyme activity at fill volumes greater than 1.0 ml (as measured by luciferin-ME demethylation; Fig. 3C), subsequent expressions were performed with a fill volume of 1.0 ml.

**Screening of Total CYP1A1/1A2 Library for Expression and Activity toward Luciferin-ME, Luciferin-CEE, and EROD.** Of the 64 mutant library clones, two clones (6 and 42) showed anomalous high P450 expression without activity on CYP1A1/1A2-selective substrates and proved upon diagnostic sequencing to be CYP2C9, which had been used to prepare the vector backbone in the final ligation step of DNA family shuffling. These clones are not considered further. The use of a P450 with strongly divergent properties to the library parents when preparing vector backbone in the DNA shuffling protocol (such as CYP2C9 to CYP1A1/1A2) enables easy detection of such contaminant clones at an early stage of screening.

Of the 62 clones remaining, 33 (53%) showed P450 expression above the detection limit for the whole-cell assay (50 nM). One clone (designated 50) did not grow in starter culture. The profiles of P450 expression level for the library clones as measured by whole-cell CO-difference spectrum are provided in Fig. 4A with parental expression levels provided for reference. There was no systematic decrease of clone harvest biomass as P450 production increased (e.g., due to metabolic loading by the foreign protein). Hence, the profile of specific clone P450 content (i.e., P450 yield per o.d. 600 as a measure of cell density) for the clones was very similar to that presented for total P450 yield in Fig. 4 and is therefore not shown. All clones were screened in whole cells against the luminogenic substrates luciferin-CEE (CYP1A1-selective) and luciferin-ME (CYP1A2-selective).







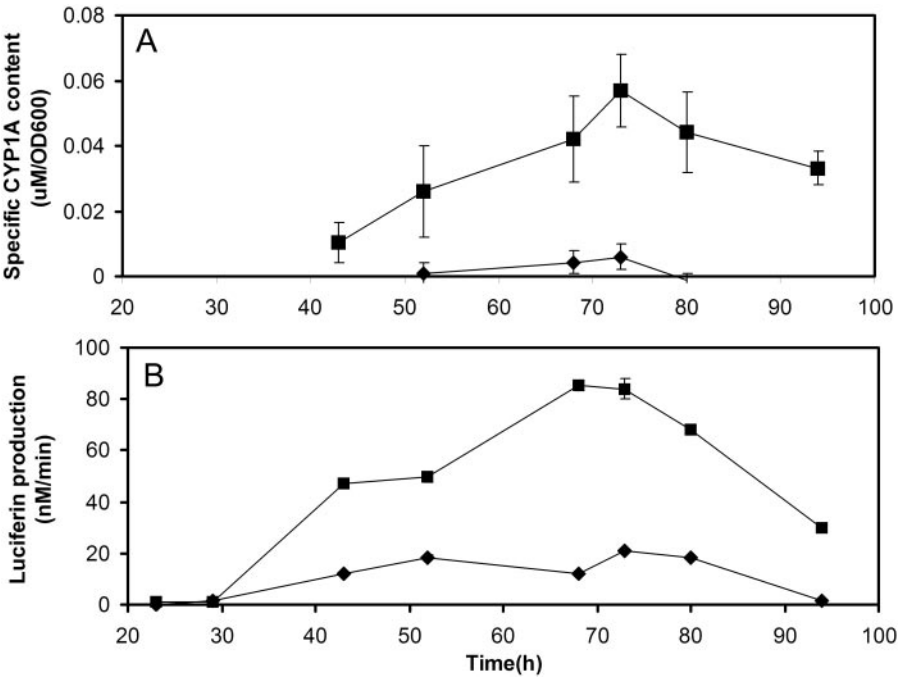



FIG. 1. Chimeragenesis in the CYP1A1/1A2 library. Library parents are depicted in light gray (CYP1A1) and dark gray (CYP1A2). Data squares represent 100% parental match at each base position. The solid line for each clone defines the library parents identified within the sequence, including crossovers. Where the location of crossovers cannot be exactly determined, they are placed approximately midway between defined parental matches. Black crosses show mutations, i.e., clone bases that do not match either parent. Over the 12 total clones that were sequenced, the crossover rate was 5.9 ± 2.3, and the spontaneous mutation rate was 1.5 ± 1.5 per sequence (mean ± S.D.). No significant difference was seen in crossover and mutation rate between the functional and randomly selected clones. Two clones sequenced (6 and 42) proved to be contaminant CYP2C9 introduced during the DNA shuffling and are not presented. Cleavage sites for restriction enzymes used for DNA shuffling are presented below main plot: light-gray triangles, cleavage sites in CYP1A1; and dark-gray triangles, cleavage sites in CYP1A2. Additionally, sequence similarity between CYP1A1 and CYP1A2 over the total base pair range is represented beneath the restriction digest sites (averaged over a 10-bp window).

FIG. 2. Time course of P450 expression (A) and activity (B) toward luciferin-ME (a selective substrate for CYP1A2) and luciferin-CEE (a selective substrate for CYP1A1) during expression of CYP1A1 (diamonds) and CYP1A2 (squares) in 24-well plates. Cultures were induced and semipermeable membranes were added at 5 h. Data represent the mean ± S.D. of triplicate well cultures harvested at each time point.

Specific activities (product formation per nanomole of P450) for library clones were calculated using whole-cell P450 concentrations (Fig. 4A) and normalized by percentage to the appropriate parental activity for each substrate (Fig. 4, B and C). All clones were tested for EROD activity using an endpoint assay that subsequently proved to be prone to false positives and background interference and hence was

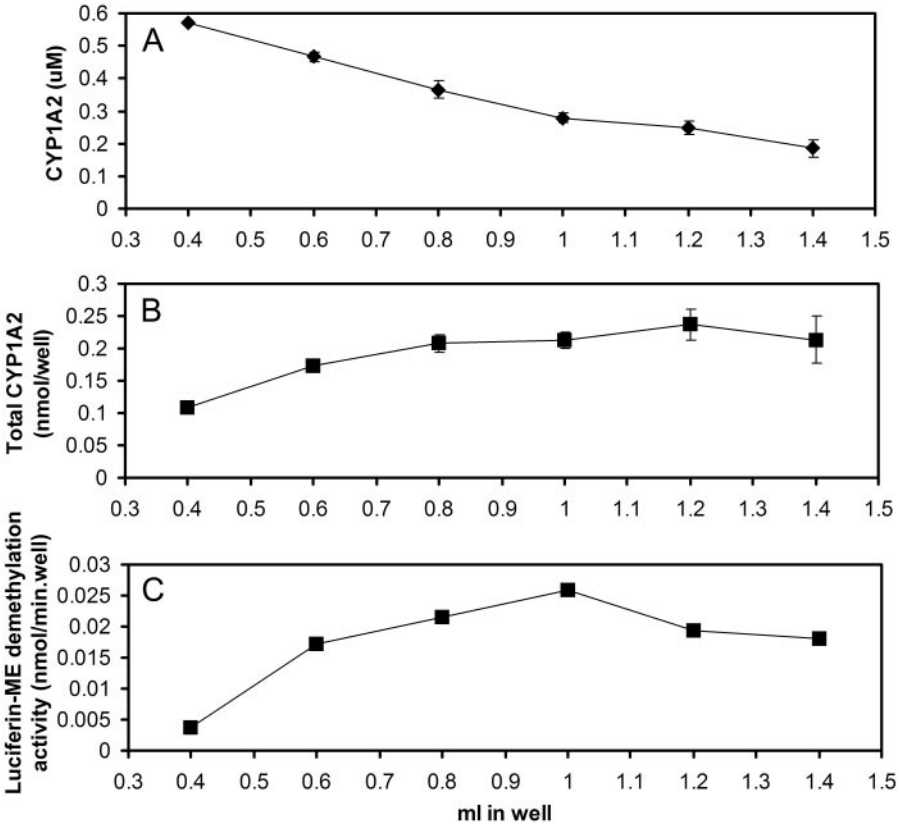



FIG. 3. Effect of changing initial well fill volume on harvest parameters: CYP1A2 final concentration (A), total amount of CYP1A2 per well (B), and total luciferin-ME activity in each well (C). Cultures were incubated in 24-well plates and harvested at 72 h.

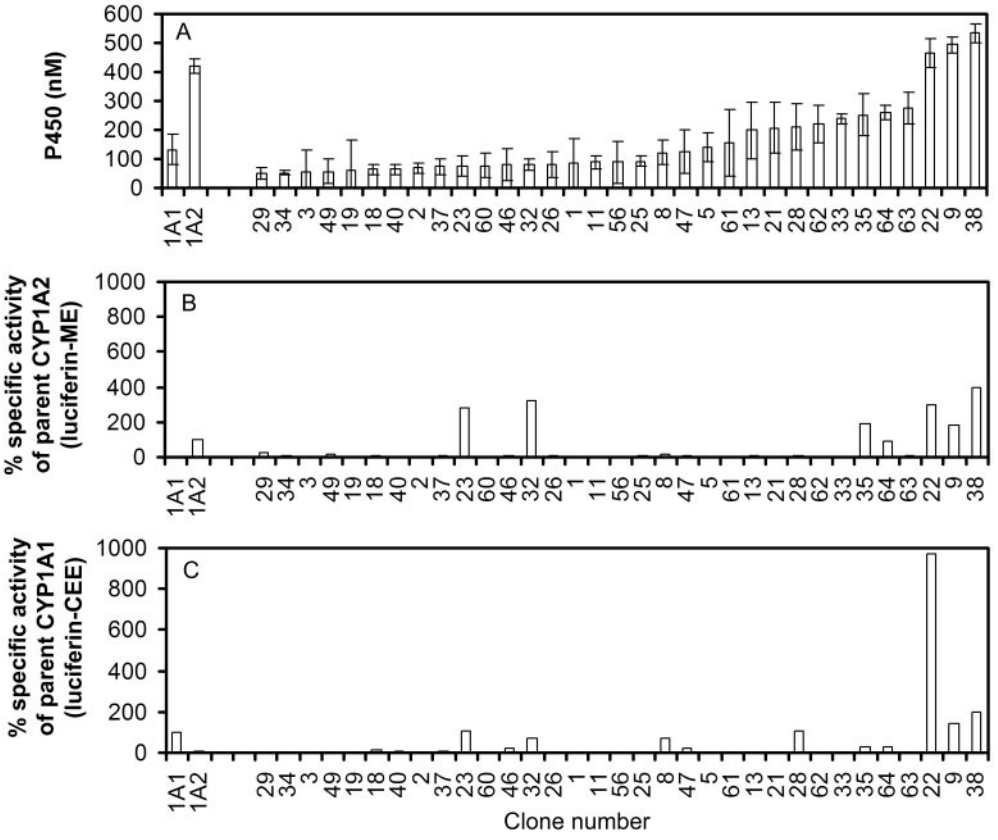



FIG. 4. Metabolic screening results of the entire shuffled CYP1A1/1A2 library (including parents). A, P450 expression; B, activity toward luciferin-ME (CYP1A2-selective); C, activity toward luciferin-CEE (CYP1A1-selective). Only clones with P450 expression exceeding the detection limit (50 nM) are plotted, with clones arranged in order of ascending P450 expression level. P450 data represent mean ± S.D. for $n = 3$ cultures incubated in 24-well plates and harvested at 72 h. Specific activities of clones toward luciferin derivatives were calculated with reference to P450 contents quantified by CO-difference spectra and then normalized to the activity of parent clones (CYP1A1 and CYP1A2, respectively, for luciferin-CEE and luciferin-ME).

considered qualitative only (data not supplied). However, this assay allowed preliminary detection of EROD activity in clones 8, 13, 19, 21, 22, 23, 25, 26, 28, 32, and 38, and these clones were included in further quantitative screening for EROD and another activities. The library was also screened against luciferin-BE and for PROD activity, with no clones showing activity.

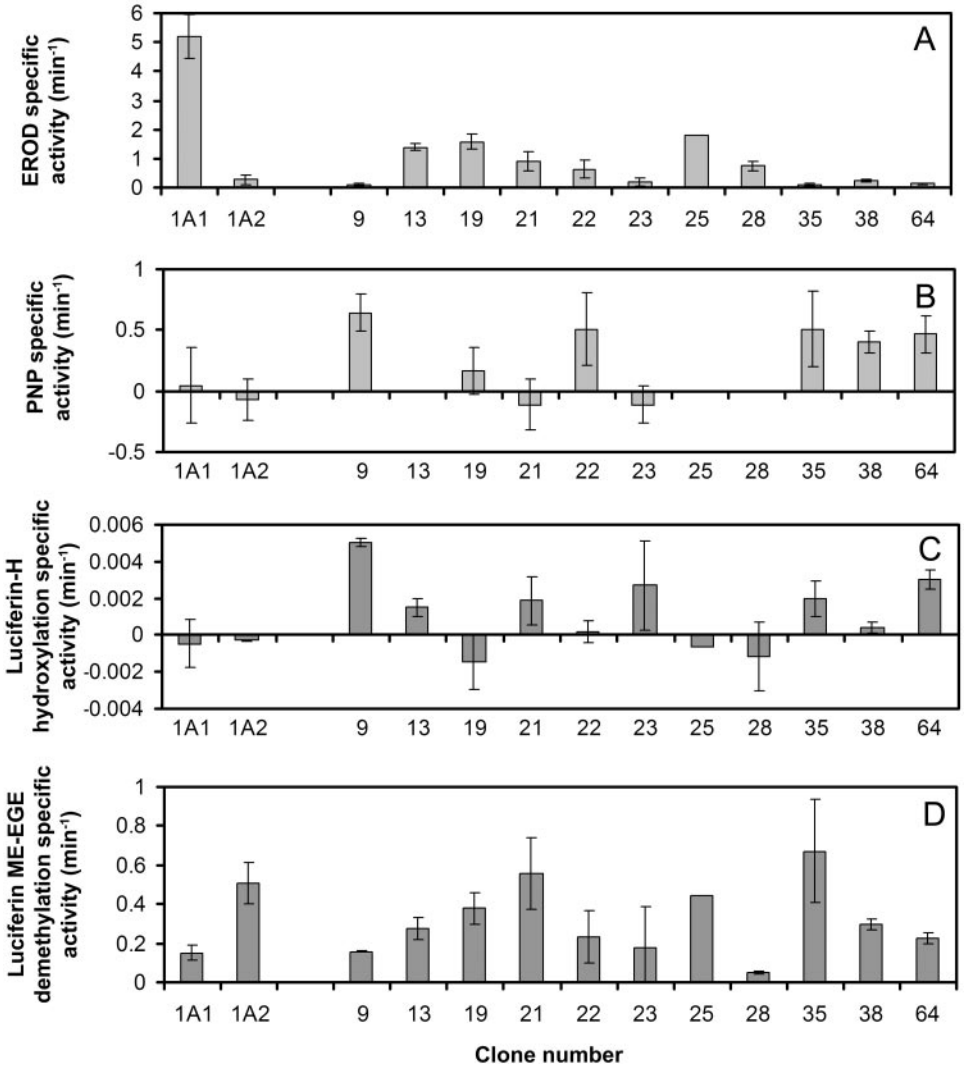



FIG. 5. Activities of selected clones in the shuffled CYP1A1/1A2 library (including parents). A, specific EROD activity; B, specific PNP hydroxylase activity; C, specific activity toward luciferin-H; D, specific activity toward luciferin ME-EGE. Only clones with P450 expression exceeding the detection limit (50 nM) are plotted. Data represent mean ± S.D. for $n = 3$ cultures incubated in 24-well plates. Specific activities of clones toward luciferin derivatives were calculated with reference to P450 contents quantified by CO-difference spectra.

Three clones with expression and P450-GLO activity in the first expression trial (9, 22, and 64) were included in the second expression round for comparative purposes. Overall results were comparable between repeat screening rounds.

**Screening of Selected Clones for Specific Activity toward 7-ER, PNP, Luciferin-H, and Luciferin ME-EGE.** From initial screening of the entire CYP1A1/1A2 library (Fig. 4), 14 clones (mutants 8, 9, 13, 19, 21, 22, 23, 25, 26, 28, 32, 35, 38, and 64) were selected with detectable activity toward at least one compound (i.e., toward luciferin-ME, luciferin-CEE, or 7-ER). These clones were screened for specific activity toward 7-ER, PNP, luciferin-H, and luciferin-ME-EGE (Fig. 5). Only those clones demonstrating P450 concentration in at least one replicate at or above the 50 nM detection limit are shown. Clones 8, 26, and 32 had detectable EROD activity, but P450 expression in all replicates was less than 50 nM; thus, specific activity levels could not be calculated. Similarly, clones 8 and 26 showed activity toward luciferin-H. For the remaining clones (9, 13, 19, 21, 22, 23, 25, 28, 35, 38, and 64), specific activities toward EROD and luciferin ME-EGE were approximately equal to or less than parental activities (Fig. 5, A and D). However, activities toward PNP and luciferin-H in some clones (Fig. 5, B and C) represent novel activities not present in either CYP1A1 or CYP1A2. A table listing specific activities for all mutants tested is available as on-line supplementary information.

## Discussion

The overall objective of the current work was to assess the technique of restriction enzyme-mediated DNA shuffling, combined with rapid methods of detecting both enzyme expression level and activity in whole cells, as a strategy for generating robust and diverse P450 libraries with desired properties. Sensitive measurement of enzyme concentration is critical for libraries using mammalian P450s because forms such as CYP1A1 express at levels many times lower than their bacterial counterparts (Yun et al., 2006). Optimal determination of low P450 levels in library clones necessitated improvements in the sensitivity of the assay and manipulation of culture conditions during expression to maximize P450 production. In initial aerobic flask expression studies, maximum whole-cell expression levels were approximately 200 nM for CYP1A2 and undetectable for CYP1A1 (data not shown). Limitation of oxygen tension and optimization of harvest time and well fill volume enhanced relative P450 concentrations, as previously noted in fermenter cultures (Vail et al., 2005). Although direct control of oxygen tension in multiwell plates is not a practical goal, a general reduction of oxygen tension for culture plates was possible in the current study by using semipermeable polyurethane membranes (e.g., Breath-easy; Diversified Biotech). Limiting oxygen tension also led to a marked reduction of the interfering 417- to 418-nm peak on CO-difference spectra, as will be discussed in detail

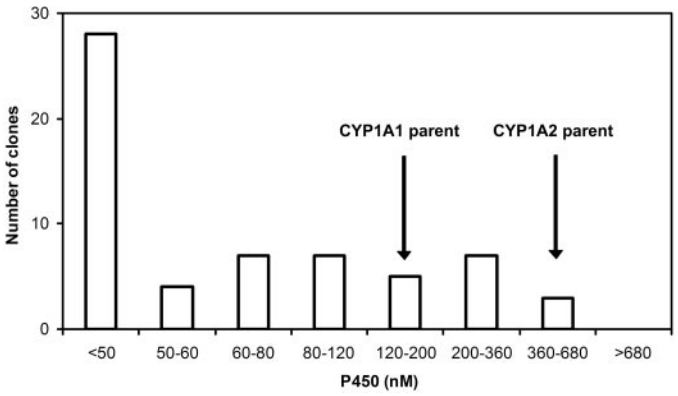



FIG. 6. Distribution of mean whole-cell P450 expression levels ($n = 3$) for CYP1A1/1A2 library clones. Arrows, mean expression levels of parent P450s. Histogram category lengths are logarithmic.

elsewhere (Johnston WA, Huang W, De Voss JJ, Hayes MA, Gillam EMJ, unpublished data). This peak has been attributed to the CO-bound complexes of the cellular terminal oxidases cytochrome *bo* and *bd* (Cheesman et al., 1993). These combined improvements enabled determination of CYP1A1 level in whole cells and thus obviated the need for downstream processing (i.e., subcellular fractionation to isolate membranes).

The restriction enzyme-mediated DNA shuffling process successfully produced a CYP1A1/1A2 library with significant sequence diversity. Although the CYP1A crossover rate (5.9 ± 2.3, mean ± S.D.) was somewhat lower than previously seen with a CYP2C library (9.1 ± 1.8, mean ± S.D.) (Huang et al., 2007), it is comparable with that of a previous CYP1A library (5.4 ± 2.2) created via DNaseI-mediated shuffling and recombination in yeast (Abécassis et al., 2000). The reduced crossover rate seen in the current study, as compared with that observed in the CYP2C library made with the same method, may be due to the paucity of recombinations in the middle of the mutant sequences (Huang et al., 2007). The lack of recombination sites between ~550 and ~950 bp was notable and corresponds to a region of lower sequence identity between the two parents (Fig. 1), in which recombination may be more difficult. In contrast, the degree of homology was relatively consistent throughout the length of the CYP2C parental sequences. An alternative explanation is that fragmentation of sequences was incomplete within this region due to the methylation sensitivity of two of the enzymes used (MboII and Eco0109I). However, an analysis of the restriction sites that would be predicted to be methylated showed that the small proportion of sites that would be affected by methylation lay outside the region showing no apparent crossovers.

Significant diversity was also seen in the P450 expression level among library clones. P450 expression levels in mutants were distributed around the parents, with an expected skewed distribution toward expression levels lower than both parents (Fig. 6). P450 levels above the detection limit (50 nM in whole cells) were seen in 53% of shuffled CYP1A clones (33 of 62). This compares well with proportions of folded enzyme observed (≈50%) in libraries of bacterial CYP102s with defined crossovers ($n = 7$) computationally optimized for stability and diversity with the SCHEMA algorithm (Otey et al., 2006) and is significantly higher than in previous 1A libraries prepared using DNaseI fragmentation (Abécassis et al., 2000; Taly et al., 2007). The high degree of functional integrity in this CYP1A1/1A2 library may be attributed to the fact that libraries created by DNA shuffling tend to show lower "kill rates" compared with those created by random mutagenesis since most amino acid changes that are introduced are already present in one of the successfully expressed library parents. Spontaneous mutations are more likely to cause nonfunctional enzyme due to premature stop codons or misfolding. This was apparent in a previous CYP1A1/1A2 library, where spontaneous mutation rate was significantly lower (8.3 ± 3.2) for functional when separated from nonfunctional clones (14.0 ± 4.2) (Abécassis et al., 2000). Hence, the low average spontaneous mutation rate in the current study (1.5 ± 1.5, mean ± S.D.) is advantageous for producing functional clones.

Selective substrates such as luciferin-CEE and 7-ER (both selective for CYP1A1) and luciferin-ME and luciferin-ME-EGE (selective for CYP1A2) are powerful tools for determining activity profiles for mutant clones that indicate their relative performance compared with library parents. Also useful are compounds (e.g., PNP and luciferin-H) toward which the parental forms show no significant activity. Examination of the activity profiles (Fig. 7) shows that the CYP1A library contains diverse functionality in activity, i.e., hybrid mutants with both CYP1A1 and CYP1A2 activity (clones 9, 19, 21, 22, 23, 64, and 38), superparental CYP1A1-like activity (clone 22), and superparental CYP1A2-like activity (clones 22, 23, and 35). Clone 22 represents an extreme case because it showed approximately 9 times the specific luciferin-CEE activity of CYP1A1, in addition to 3 times the specific luciferin-ME activity of CYP1A2 (Fig. 7). Additionally, clones were detected with novel activities (luciferin-H, clones 9, 23, 35, and 64; PNP, clones 9, 19, 35, 64, 22, and 38). For the purpose of second round shuffling, this indicates the presence of library clones suitable as new parents to improve both parental activity and broaden substrate specificity. Indeed, functional clones tended to combine both broadening of substrate range and enhancement of parental activity (Fig. 7). The diversity seen here in activity profiles was comparable with that seen in libraries derived from bacterial CYP102s with lower overall sequence identity, constructed using the SCHEMA algorithm to maximize functional protein recovery (Otey et al., 2004).

Detection of novel activity toward PNP and luciferin-H in the library indicates that although the genetic variation available through DNA family shuffling is limited to that already present in the parents (plus infrequent beneficial spontaneous mutations accumulated during shuffling and amino acid changes due to recombination within codons), this library still encodes proteins with a remarkable diversity of possible enzymatic activities. However, it was not possible to detect activity on relatively large compounds (luciferin-BE, 7-PR), suggesting that there are limits to the possible restructuring of the relatively planar CYP1A active site cavity (Sansen et al., 2007). Importantly, no library clone showed an activity profile characteristic of either parent (Fig. 7). This is consistent with the lack of library parental contamination inherent in the restriction enzyme-mediated shuffling protocol due to introduction of a size-selective filtration step to recover digested fragments as well as efficient fragmentation using restriction enzymes (Huang et al., 2007; Rosic et al., 2007). In comparison, at least 13.8% of the previous DNaseI-mediated shuffled 1A library appeared to show parental hybridization patterns, a value that may have been elevated due to linkage between probed segments (Abécassis et al., 2000).

Combining whole-cell assays for P450 and enzyme activity on various compounds proved to be useful in allowing rapid and technically straightforward screening of the 1A library. Importantly, such methods avoid the labor intensive fractionation of cells, as well as concomitant losses in enzyme content and activity. However, turnover rates reported were low when compared with typical values obtained with pure enzymes or subcellular fractions (Fig. 5). These whole-cell turnover values do not reflect $k_{cat}$ values since they will be confounded by the additional kinetic processes of substrate diffusion into and product diffusion out of the whole cells in suspension (both

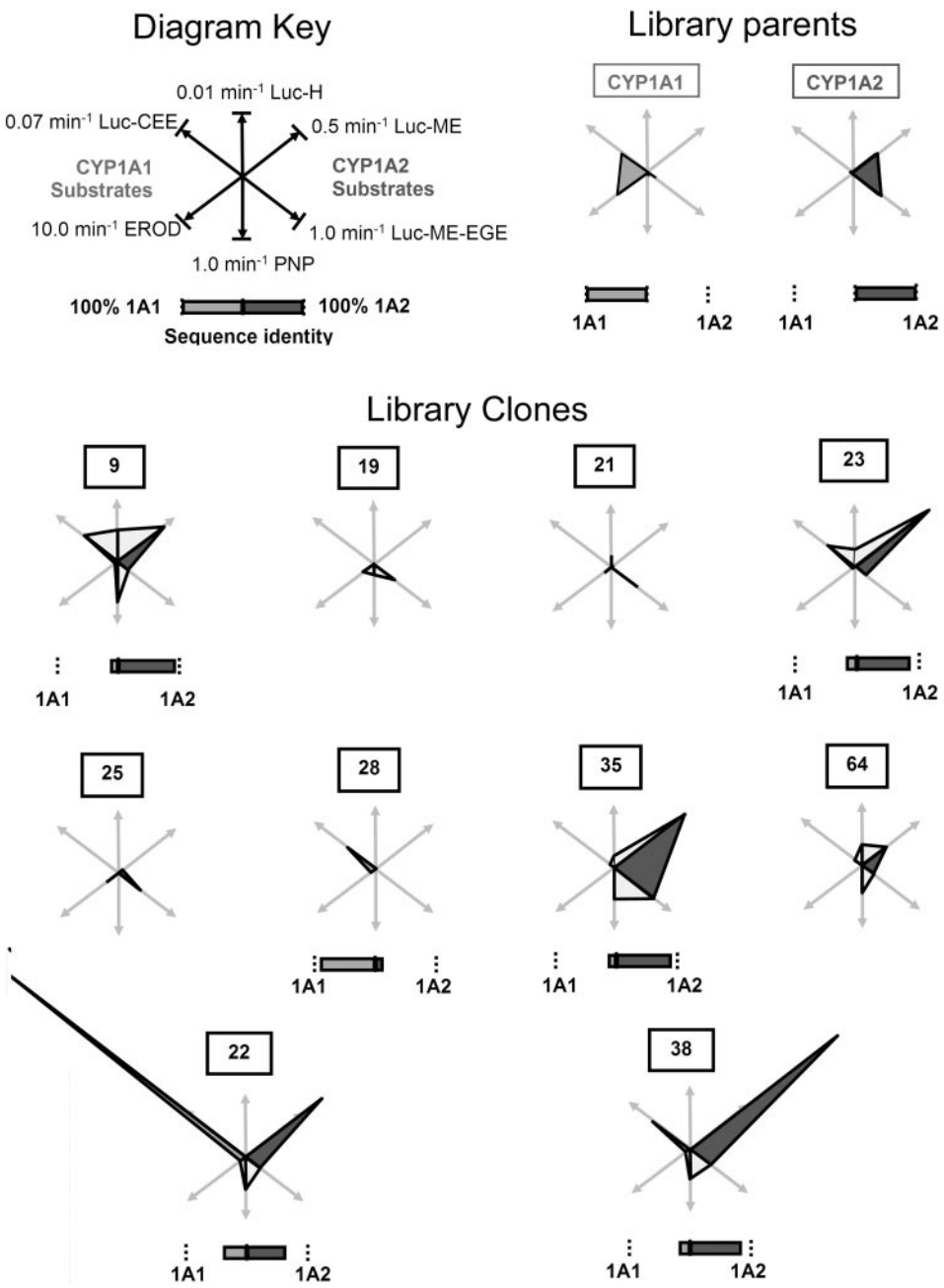



FIG. 7. Activity profiles for selected clones in the CYP1A1/1A2 library. Average specific activities normalized for P450 whole-cell expression level on various substrates are plotted. Substrates with activity biased toward each parent (luciferin-CEE and EROD for CYP1A1; luciferin-ME and luciferin ME-EGE for CYP1A2) are plotted on the left and right of each plot, respectively, as shown in the diagram key. Compounds for which neither parent showed significant activity (luciferin-H, PNP) are plotted vertically. The percentage nucleotide identity with each parent is represented beneath each plot for selected clones (light-gray bar, CYP1A1; dark-gray bar, CYP1A2).

reducing apparent $V_{max}$ values). Additionally, the internal cellular substrate concentration (i.e., that in contact with the enzyme) will be lower than the external assay substrate concentration (thus increasing apparent $K_m$). A detailed examination of these kinetic effects in whole-cell activity assays is beyond the scope of this article. However, for the purpose of routine, first tier screening, these effects can be considered constant across different clones, allowing useful comparisons of relative specific activity. In cases where a more detailed kinetic comparison is required (such as for quantitative, structure-function studies), such studies should be performed on enzymes isolated by subcellular fractionation.

The diversity of activity detected in the current CYP1A library was high considering the small sample size. The minimum sample size to capture all theoretical possible genetic diversity in the library (discounting that introduced by spontaneous mutations) can be estimated at $5.2 \times 10^{53}$ possible mutants by assuming recombination could occur at all possible points across all sampled sequences and accounting for recombination within codons (Boden M, Gillam EMJ, unpublished data). Hence, practical considerations determine the size of the sample that can be assessed in such a library. One advantage of increasing the size of the sample examined (besides finding additional candidates for further investigation and evolution) is to elucidate sequence-structure-function relationships. P450 structure-function relationships determined solely from the mutagenesis of putative substrate binding sites have been disadvantaged by the fact enzyme activity is affected by changes throughout the protein in backbone flexibility, domain motion, or interactions with redox partners (Kim and Guengerich, 2004). Hence, structure-function relationships based on assessing changes throughout the total sequence may yield further insights. In an extended CYP1A library (with sequence data for functional clones), functional clones could be grouped by similarities in activity profile (e.g., augmentation or attenuation of individual

activities), then compared computationally to detect sequence similarities. For example in the current CYP1A library, clone 28 can be considered to be a "damaged" CYP1A1 that has retained luciferin-CEE activity but lost EROD activity (Fig. 7). The sequence for clone 28 shows a small proportion (10.5%) of CYP1A2 sequence (in four separate sections) inserted into a majority CYP1A1 sequence (Fig. 1). Comparison with other clones with similar activity profile but different sequence in a larger library might allow resolution of exact residues causing this differential activity change.

In summary, even within a small sample of the CYP1A library, a remarkable diversity of expression level and catalytic activity profile was found after a single round of restriction enzyme-mediated DNA family shuffling. Combining whole-cell screening of both P450 content and activity presents a rapid and effective way of characterizing P450 libraries even when they contain poorly expressed forms. This combination of techniques represents a strategy for generating robust and diverse libraries of enzymes that can be used to develop tailored P450 catalysts with desired properties.

**Acknowledgments.** We thank Brenda L. Mooney for construction of the hexa-His-tagged version of the CYP1A1 expression vector, Mikael Boden for assistance in calculating the theoretical library size, James Cali for donating some of the luciferin substrates, and Dominic J.B. Hunter for proofreading of the manuscript.

## References

Abécassis V, Pompon D, and Truan G (2000) High efficiency family shuffling based on multi-step PCR and *in vivo* DNA recombination in yeast: statistical and functional analysis of a combinatorial library between human cytochrome P450 1A1 and 1A2. *Nucleic Acids Res* **28:**e88.

Abécassis V, Urban P, Aggerbeck L, Truan G, and Pompon D (2003) Exploration of natural and artificial sequence spaces: towards a functional remodeling of membrane-bound cytochrome P450. *Biocatal Biotransform* **21:**55–66.

Bauer S and Shiloach J (1974) Maximal exponential growth rate and yield of *E. coli* obtainable in a bench-scale fermentor. *Biotechnol Bioeng* **16:**933–941.

Chang TK and Waxman DJ (1997) Enzymatic analysis of cDNA expressed human CYP1A1, CYP1A2 and CYP1B1 with 7-ethoxyresorufin, in *Methods in Molecular Biology* (Phillips I and Shephard E eds), pp 103–109, Humana Press Inc., Totowa, NJ.

Cheesman MR, Watmough NJ, Pires CA, Turner R, Brittain T, Gennis RB, Greenwood C, and Thomson AJ (1993) Cytochrome *bo* from *Escherichia coli*: identification of heme ligands and reaction of the reduced enzyme with carbon monoxide. *Biochem J* **289:**709–718.

Crameri A, Raillard SA, Bermudez E, and Stemmer WPC (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391:**288–291.

Gillam EMJ, Baba T, Kim B-R, Ohmori S, and Guengerich FP (1993) Expression of modified human cytochrome P450 3A4 in *Escherichia coli* and purification and reconstitution of the enzyme. *Arch Biochem Biophys* **305:**123–131.

Gillam EMJ (2005) Exploring the potential of xenobiotic-metabolising enzymes as biocatalysts: evolving designer catalysts from polyfunctional cytochrome P450 enzymes. *Clin Exp Pharmacol Physiol* **32:**147–152.

Gillam EMJ, Aguinaldo AMA, Notley LM, Kim D, Mundkowski RG, Volkov AA, Arnold FH, Soucek P, DeVoss JJ, and Guengerich FP (1999) Formation of indigo by recombinant mammalian cytochrome P450. *Biochem Biophys Res Commun* **265:**469–472.

Guengerich FP (2005) Human cytochrome P450 enzymes, in *Cytochrome P450* (Ortiz de Montellano PR ed), pp 377–530, Kluwer Academic/Plenum Publishing, New York.

Huang W, Johnston WA, Hayes MA, De Voss JJ, and Gillam EMJ (2007) Extending the capabilities of nature's most versatile catalysts: a shuffled CYP2C library with a high degree of structural integrity and functional versatility. *Arch Biochem Biophys*, in press.

Kikuchi M, Ohnishi K, and Harayama S (1999) Novel family shuffling methods for the in vitro evolution of enzymes. *Gene* **236:**159–167.

Kim D and Guengerich FP (2004) Enhancement of 7-methoxyresorufin *O*-demethylation activity of human cytochrome P450 1A2 by molecular breeding. *Arch Biochem Biophys* **432:**102–108.

Kumar S, Chen CS, Waxman DJ, and Halpert JR (2005) Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates, including the anticancer prodrugs cyclophosphamide and ifosfamide. *J Biol Chem* **280:**19569–19575.

Kumar S, Liu H, and Halpert JR (2006a) Engineering of cytochrome P450 3A4 for enhanced peroxide-mediated substrate oxidation using directed evolution and site-directed mutagenesis. *Drug Metab Dispos* **34:**1958–1965.

Kumar S, Sun L, Liu H, Muralidhara BK, and Halpert JR (2006b) Engineering mammalian cytochrome P450 2B1 by directed evolution for enhanced catalytic tolerance to temperature and dimethyl sulfoxide. *Protein Eng Des Sel* **19:**547–554.

Nishihara K, Kanemori M, Kitagawa M, Yanagi H, and Yur T (1998) Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl Environ Microb* **64:**1694–1699.

Notley LM, de Wolf CJF, Wunsch RM, Lancaster RG, and Gillam EMJ (2002) Bioactivation of tamoxifen by recombinant human cytochrome P450 enzymes. *Chem Res Toxicol* **15:**614–622.

Otey CR (2003) High-throughput carbon monoxide binding assay for cytochromes P450, in *Methods in Molecular Biology* (Arnold FH and Georgiou G eds), pp 137–139, Humana Press, Totowa, NJ.

Otey CR, Landwehr M, Endelman JB, Hiraga K, Bloom JD, and Arnold FH (2006) Structure-guided recombination creates an artificial family of cytochromes P450. *PLoS Biol* **4:**e112.

Otey CR, Silberg JJ, Voigt CA, Endelman JB, Bandara G, and Arnold FH (2004) Functional evolution and structural conservation in chimeric cytochromes P450: calibrating a structure-guided approach. *Chem Biol* **11:**309–318.

Parikh A, Gillam EMJ, and Guengerich FP (1997) Drug metabolism by *Escherichia coli* expressing human cytochromes P450. *Nat Biotechnol* **15:**784–788.

Rosic NN, Huang W, Johnston WA, DeVoss JJ, and Gillam EMJ (2007) Extending the diversity of cytochrome P450 enzymes by DNA family shuffling. *Gene* **395:**40–48.

Sansen S, Yano JK, Reynald RL, Schoch GA, Griffin KJ, Stout CD, and Johnson EF (2007) Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2. *J Biol Chem* **282:**14348–14355.

Stemmer WPC (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370:**389–391.

Taly V, Urban P, Truan G, and Pompon D (2007) A combinatorial approach to substrate discrimination in the P450 CYP1A subfamily. *Biochim Biophys Acta* **1770:**446–457.

Vail RB, Homann MJ, Hanna I, and Zaks A (2005) Preparative synthesis of drug metabolites using human cytochrome P450s 3A4, 2C9 and 1A2 with NADPH-P450 reductase expressed in *Escherichia coli*. *J Ind Microbiol Biotechnol* **32:**67–74.

Winters DK and Cederbaum AI (1992) Expression of a catalytically active human cytochrome P450 2E1 in *Escherichia coli*. *Biochim Biophys Acta* **1156:**43–49.

Yun CH, Yim SK, Kim DH, and Ahn T (2006) Functional expression of human cytochrome P450 enzymes in *Escherichia coli*. *Curr Drug Metab* **7:**411–429.

**Address correspondence to:** Elizabeth M.J. Gillam, School of Biomedical Sciences, The University of Queensland, St. Lucia, Brisbane 4072, Australia. E-mail: e.gillam@uq.edu.au