Functional characterization of cytochrome P450 2A6 (CYP2A6) allelic variants CYP2A6*15, CYP2A6*16, CYP2A6*21 and CYP2A6*22

Kai Hung Tiong, Beow Chin Yiap, Eng Lai Tan, Rusli Ismail, and Chin Eng Ong

School of Pharmacy and Health Sciences, International Medical University, 126, Jalan 19/155B, Bukit Jalil, 57000 Kuala Lumpur, Malaysia (K.H.T., B.C.Y., E.L.T.); Pharmacogenetics Research Group, Institute for Research in Molecular Medicine, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia (R.I.); and School of Medicine and Health Sciences, Monash University Sunway Campus, Jalan Lagoon Selatan, 46150 Bandar Sunway, Selangor, Malaysia (C.E.O.)
Running title: Functional characterization of CYP2A6 alleles

*Address correspondence to: Dr Chin Eng Ong, School of Medicine and Health Sciences, Monash University Sunway Campus, Jalan Lagoon Selatan, 46150 Bandar Sunway, Selangor, Malaysia. Phone: +6 03 55144918, Fax: +6 03 55146323, E-mail:
ong.chin.eng@med.monash.edu.my

Number of text pages: 26
Number of tables: 2
Number of figures: 5
Number of references: 37
Number of words in Abstract: 200
Number of words in Introduction: 778
Number of words in Discussion: 1,343

The abbreviations used are:
CYP, cytochrome P450
NADPH, nicotinamide adenine dinucleotide phosphate
OxR, NADPH-cytochrome P450 oxidoreductase
SRS, substrate recognition site
Abstract

Variation in CYP2A6 levels and activity can be attributed to genetic polymorphism and thus, functional characterization of allelic variants is necessary to define the importance of CYP2A6 polymorphism in human. The present study aimed to investigate the reported allele CYP2A6*15, CYP2A6*16, CYP2A6*21 and CYP2A6*22, in terms of the functional consequences of their mutations on the enzyme catalytic activity. Utilizing the wild type CYP2A6 cDNA as template, site-directed mutagenesis was carried out to introduce nucleotide changes encoding Lys194Glu substitution in CYP2A6*15, Arg203Ser substitution in CYP2A6*16, Lys476Arg substitution in CYP2A6*21 and concurrent Asp158Glu and Leu160Ile substitutions in CYP2A6*22. Upon sequence verification, the CYP2A6 wild type and mutant constructs were individually co-expressed with NADPH CYP-reductase in E. coli. Kinetic study using coumarin 7-hydroxylase assay indicated that CYP2A6*15 exhibited higher $V_{max}$ than the wild type while all mutant constructs, except for variant CYP2A6*16, exhibited higher $K_m$ values. Analysis of $V_{max}/K_m$ ratio revealed that all mutants demonstrated 0.85- to 1.05-fold difference from the wild type, with exception of variant CYP2A6*22 which only portrayed 39% of the wild-type intrinsic clearance. These data suggested that individuals carrying CYP2A6*22 allele are likely to have lower metabolism of CYP2A6 substrate than individuals expressing CYP2A6*15, CYP2A6*16, CYP2A6*21 and the wild-type.
Introduction

Human cytochrome P450 2A6 (CYP2A6) remains as one of the less well characterized members among the many known isoforms of human cytochrome P450s (Pelkonen et al., 2000). To date, it has been recognized as the major isoform involved specifically in the oxidative metabolism of nicotine, a major constituent in the tobacco smoke (Rossini et al., 2008). CYP2A6 also significantly contributes to the catalytic metabolism of clinically used drugs such as methoxyflurane, halothane, losigamone, letrozole, valproic acid, disulfiram, and fadrozole and activates some procarcinogens such as 4-methylnitrosoamo-1-(3-pyridyl)-1-butane and \(N\)-nitrosodiethylamine (Oscarson et al., 2001). Early studies of hepatic coumarin 7-hydroxylase activity have indicated pronounced interindividual differences in CYP2A6 expression levels and activity (Pelkonen et al., 1985; Camus et al., 1993; Shimada et al., 1996). High variations were detected with some livers found to be completely lacking the enzyme (Shimada et al., 1996). Interestingly, this variability in CYP2A6 shows important ethnic differences, with only 1% of Caucasians being reported as poor metabolisers (PMs), whereas up to 20% has been reported in Asians (Rautio et al., 1992; Shimada et al., 1996; Ingelman-Sundberg et al., 2007). Thus, unequivocal elucidation of all functional alleles and global genotyping for CYP2A6 is vital due to the distinctive role played by the isoform in the metabolism of various substrates, especially pharmacologically and toxicologically relevant compounds. In tandem with nicotine and other tobacco-specific carcinogens being established as high-affinity substrates for CYP2A6, much attention has focused on the toxicological and clinical significance of this highly polymorphic isoform in human.
To date, 33 allelic variants, designated as CYP2A6*2 to CYP2A6*37 have been identified (the list of allelic variants is available at http://www.imm.ki.se/CYPalleles/). Among the first few isolated allele, CYP2A6*2 has been reported to encode a protein with a Leu160His substitution which renders the enzyme inactive (Hadidi et al., 1997). CYP2A6*3 which was initially reported as a hybrid allele of multiple gene conversions with a pseudogene, CYP2A7, is now believed to be an artifact due to the shortcoming of the previously used PCR-based genotyping assay for CYP2A6*1, CYP2A6*2 and CYP2A6*3 (Oscarson et al., 1999). Alleles CYP2A6*4, CYP2A6*5 and CYP2A6*20 were also found to display abolished functional activity due to gene deletion or point mutation in their primary sequence. Twelve CYP2A6 alleles (*6, *7, *10, *11, *12, *17, *18, *19, *23, *24, *26 and *27) have been ascertained to cause reduction in enzyme activity, whereas genetic polymorphisms in the promoter region of CYP2A6 have been implicated for the decreased transcriptional activity observed in alleles CYP2A6*1D, CYP2A6*1H and CYP2A6*9 (Ariyoshi et al., 2001; Kitagawa et al., 2001; Daigo et al., 2002; Fukami et al., 2004; Fukami et al., 2005; Ho et al., 2008).

Functional significance of some alleles such as CYP2A6*15 (Lys194Glu), CYP2A6*16 (Arg203Ser), CYP2A6*21 (Lys476Arg), and CYP2A6*22 (Asp158Glu and Leu160Ile) has yet to be investigated in detail. The occurrence of these four alleles in human population was first revealed in 2002 and 2005 with variant CYP2A6*15 found in Korean and Japanese population at frequencies of approximately 1.2% and 1.5 to 2.2% respectively (Kiyotani et al., 2002; Nakajima et al., 2006). Conversely, frequency of CYP2A6*16 was prominent among Caucasians (0.3 to 3.6%) and African-Americans (1.7%) while this allele remained undetected in Asian population (Kiyotani et al., 2002;
Nakajima et al., 2006). Allelic variant CYP2A6*21 has higher occurrence as well in white subjects with 0.5 to 7.0% as compared to Chinese (3.4%) and black subjects (0.6%) whereas Japanese subjects remained unaffected (Haberl et al., 2005; Koudsi et al., 2006; Nakajima et al., 2006). Comprehensive evaluation of CYP2A6 polymorphic alleles by Nakajima and co-workers in four ethnic populations did not detect any occurrence of CYP2A6*22 within the populations studied (Nakajima et al., 2006). So far, the variant was only reported at low frequency (0.3%) among the Caucasians (Haberl et al., 2005). While no data is available for CYP2A6*15 and CYP2A6*16, haplotype analyses on CYP2A6*21 and CYP2A6*22 revealed that haplotypes carrying the two alleles occurred at low frequencies (at 0.6% and 0.3% respectively) in Caucasians (Haberl et al., 2005). Haplotype frequencies of these four alleles in other populations as well as their phenotypic association with protein level and activity, however, remain unknown at this stage.

Despite the existence of these four CYP2A6 variants in the human populations, functional characterisation of the manifested polymorphisms has yet to be determined in detail. With one or more amino acid mutations in their primary sequences, it is likely that these mutations would have certain degree of effects on their structural stability and catalytic activity. Structural and functional characterization of these polymorphic alleles is necessary because it contributes to our better understanding of the mutation consequences and hence aid in defining the pharmacological and toxicological importance of CYP2A6 polymorphism in human.
Materials and Methods

Materials

QuikChange® site-directed mutagenesis system was purchased from Stratagene (Cedar Creek, TX, USA) while endonuclease restriction enzymes were obtained from New England Biolabs (Beverly, MA, USA). Mouse anti-human cytochrome P450 CYP2A6 monoclonal antibody and goat anti-mouse IgG-HRP were purchased from Santa Cruz Biotechnology Incorporation (Santa Cruz, CA, USA). 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), isopropyl-b-D-thiogalactopyranoside (IPTG) and Tris base were acquired from Promega (Madison, WI, USA). *Escherichia coli* DH5α competent cells, oligonucleotide primers, Luria-Bertani and Teriffic broth media were purchased from Invitrogen Corporation (Carlsbad, CA, USA). All other chemicals and reagents were obtained from Sigma-Aldrich (St. Louis, MI, USA).

In vitro Site-directed Mutagenesis of CYP2A6 cDNA

Site-directed mutagenesis on CYP2A6 cDNA was performed using Stratagene’s QuikChange® site-directed mutagenesis system according to the manufacturer’s instructions (Stratagene, Cedar Creek, TX, USA). The procedure basically utilized a supercoiled double-stranded DNA (dsDNA) vector with an insert of interest and two synthetic oligonucleotide primers containing the desired mutation. The DNA template in use was a pCW-CYP2A6 vector which was previously constructed in our laboratory (unpublished data). It contained N-terminal sequence modification, P450 17α-derived MALLLAVF sequence as reported by Barnes and co-workers (Barnes et al., 1991) and
the full coding sequence of human *CYP2A6*. The sequence has 100% identity to that of the reported wild type *CYP2A6* in the GenBank (GenBank accession no: NM000762). Mutagenic primers which were exclusively designed for generating mutant alleles of *CYP2A6*\(^*15\), *CYP2A6*\(^*16\), *CYP2A6*\(^*21\) and *CYP2A6*\(^*22\) are listed in Table 1. Single nucleotide substitution of AAG to GAG, CGC to AGC and AAA to AGA on the CYP2A6 primary sequence had each contributed to the amino acid substitution of Lys194Glu, Arg203Ser and Lys476Arg respectively in the expressed bacterial membrane fragment of CYP2A6\(^*15\), CYP2A6\(^*16\) and CYP2A6\(^*21\) proteins. Similarly, primer possessing both GAC to GAG and CTC to ATC substitutions has permitted expression of mutant CYP2A6\(^*22\) with concurrent Asp158Glu and Leu160Ile substitutions on the primary sequence. Prior to full nucleotide sequencing of the entire cDNA coding frame of each clone, each mutant construct was subjected to restriction analyses by several endonucleases, *EcoRI*, *FspI*, *PmlI* and *Hpy99I*. The preliminary analysis with restriction endonucleases was necessary as indication that desired mutations have successfully taken place on the nucleotide strand of *CYP2A6* cDNA. The full nucleotide sequencing of all four mutant cDNAs was further verified by capillary based sequencing method which was outsourced to AITBIOTECH Pte Ltd, Singapore.

**Heterologous Expression of CYP2A6 Wild Type and Mutant Constructs in Bacterial Expression System**

*CYP2A6* plasmid constructs harboring the desired mutations, *pCW-CYP2A6*\(^*15\), *pCW-CYP2A6*\(^*16\), *pCW-CYP2A6*\(^*21\) and *pCW-CYP2A6*\(^*22\) were individually co-transformed into *E.coli* DH5\(\alpha\) cells together with *pACYC-OxR* plasmid, the essential NADPH-CYP
oxidoreductase (OxR) coenzyme. CYP2A6 and OxR protein expression in bacterial cells and subsequent membranes preparation were carried out as described previously (Singh et al., 2008). The membrane fragments of E.coli were stored at -80°C in a 1:1 mixture of pH 7.6 TES (100 mM Tris/ 0.5 mM EDTA/ 500 mM sucrose buffer) and ice-cold distilled water before analysed in the enzyme assay reaction.

**Coumarin 7-hydroxylase Assay**

Enzyme kinetic activity of the wild type and mutants of CYP2A6 were assessed by fluorescence-based coumarin 7-hydroxylase assay with slight modifications based a published protocol (Ghosal et al., 2003; Donato et al., 2004). Coumarin 7-hydroxylation was measured in reaction mixture consisted of 50 μg expressed CYP2A6, NADPH generating system (1.3 mM NADP, 3.5 mM glucose 6-phosphate, 2 IU glucose 6-phosphate dehydrogenase and 5 mM MgCl₂) and 0.313 - 40 μM coumarin in 100 mM Tris-HCl buffer (pH 7.5). Coumarin was dissolved in acetonitrile with the final concentration of the organic solvent in each incubation mixture at 1% (v/v) or less. Reactions were initiated by addition of 50 μg CYP2A6 protein after prewarming at 37°C in a metabolic shaker for 10 min and later terminated by 50 μl of 500 mM Tris base after 25 min of incubation at 37°C. Quantification of 7-hydroxycoumarin was performed immediately by using Tecan Infinite™ 200 series microplate reader (Männerdorf, Switzerland) at the excitation wavelength 365 nm and emission wavelength 450 nm. Standard curves of 7-hydroxycoumarin were constructed in the range of 15.63 - 2000 pmol/well, and metabolite formation rate was calculated based on the curves. All samples and standards were incubated in duplicate.
Kinetic Analysis

Enzyme kinetic data were analyzed by nonlinear least squares regression analysis software EZ-Fit™ (Perrella Scientific, USA) and the kinetic parameters, Michaelis-Menten constant (K_m) and maximum velocity (V_max) were determined over the substrate range studied. Statistical analyses were performed by the SPSS statistical program (SPSS Inc., USA).

Sodium Dodecyl Sulfate-polyacrylamide Gel Electrophoresis (SDS-PAGE) and immunoblotting of CYP2A6 proteins

Briefly, bacterial membrane fractions (50 μg) expressing wild type or mutant CYP2A6 proteins were heat inactivated and separated on a 10% polyacrylamide gel before transferred electrophoretically to a nitrocellulose membrane. The membrane was then treated with 5% non-fat skimmed milk in tris buffer saline (TBS) at room temperature for an hour. Immunoblotting was later performed using a 1:200 dilution of the mouse anti-human CYP2A6 monoclonal antibody at 37°C for 1 hour, followed by incubation with 1:3000 dilution of peroxidase-labeled goat anti-mouse immunoglobulin G (IgG) as the secondary antibody. The binding of the antibody to the expressed protein was detected using the 4-chloro-1-napthol developing solution.

Other Methods

Protein concentrations were measured according to Bradford method (Bradford, 1976). The CYP content of the membranes was determined by carbon monoxide difference
spectra (Omura and Sato, 1964). The level of P450 reductase was estimated in membranes using a spectrophotometric assay to measure cytochrome c reduction (Strobel and Dignam, 1978).
Results

Site-directed Mutagenesis of CYP2A6

Four CYP2A6 mutant alleles had been successfully generated in this study using Stratagene’s QuikChange® Site-Directed Mutagenesis system. These four mutants, CYP2A6*15, CYP2A6*16, CYP2A6*21 and CYP2A6*22 were constructed to harbor one or more nucleotide changes at positions as reported in literature (Kiyotani et al., 2002; Haberl et al., 2005; Mwenifumbo et al., 2008). Isolated digestions were performed with the designated restriction endonucleases, namely EcoRI for CYP2A6*15 (Figure 1A), FspI for CYP2A6*16 (Figure 1B), PmlI for CYP2A6*21 (Figure 1C) and Hpy99I for CYP2A6*22 (Figure 1D) respectively. All restriction reactions gave the expected band patterns, indicating that the desired mutations have successfully taken place on the nucleotide strand of CYP2A6 cDNAs. Moreover, nucleotide sequencing evidently showed that all mutant constructs possessed the desired nucleotide changes, precisely AAG to GAG substitution in CYP2A6*15 (Figure 2A), CGC to AGC substitution in CYP2A6*16 (Figure 2B), AAA to AGA substitution in CYP2A6*21 (Figure 2C) and both GAC to GAG and CTC to ATC substitutions in CYP2A6*22 (Figure 2D). Observations from the electropherograms have detected no other nucleotide changes in the whole stretch of these mutant cDNAs (data not shown). Both restriction and sequencing analyses have evidently verified that the desired mutations have been successfully incorporated on the constructs at the desired positions and were ready for protein expression of the mutants and subsequent functional analyses.
Expression of CYP2A6 Proteins in *Escherichia coli*

Bacterial expressions of CYP2A6 mutant proteins were determined accordingly via SDS-PAGE and immunoblotting. Immunodetection with monoclonal mouse anti-human CYP2A6 revealed the presence of a molecular mass with approximate 49 kDa in all the recombinant mutant proteins (Figure 3, lanes *15, *16, *21, *22) which was in accordance to the reported molecular weight of CYP2A6 (Maurice et al., 1991). The levels of CYP2A6 mutant protein expressions appeared to be comparable to the wild type CYP2A6, indicating consistency in the expression of these proteins in the bacterial membrane fragments. As predicted, absence of bands was noted in the control protein expressing original pCWori+ plasmid (Figure 3, lane C). Cytochrome c reductase activity was determined in all mutant clones and the wild type CYP2A6 as well. OxR protein expression level in all mutants revealed to be insignificantly different (data not shown). Reduced CO-difference spectroscopy was also carried out for the four mutant CYP2A6 proteins. By performing three independent spectral determinations, we found no substantial deviations in expression levels between the wild-type and mutant proteins. Figure 4 shows the typical spectra obtained for the wild type as well as that of CYP2A6*15. The other three mutants (CYP2A6*16, CYP2A6*21 and CYP2A6*22) also showed spectra of similar pattern (data not shown). These results have indicated that the amino-acid exchanges had no influence on protein stability or expression level.
Kinetic Study of the CYP2A6 Wild Type and Mutant Proteins

Enzyme catalytic activity was determined for CYP2A6 variant proteins by fluorescent metabolite detection assay as described previously. Each variant protein was measured at substrate concentration of 0.63 - 80.0 µM, of which kinetics for coumarin 7-hydroxylation was fitted to the Michaelis-Menten equation. Reproducibility of the data was confirmed with three independent determinations. The apparent $K_m$ values of variant CYP2A6*15, CYP2A6*16, CYP2A6*21 and CYP2A6*22 were summarized in Table 2. Multiple comparisons of $K_m$ values among the wild type and variants of CYP2A6 revealed the differences to be highly significant for some mutants. While no difference was observed between CYP2A6*16 and the wild type, the rest of the mutants, CYP2A6*15, CYP2A6*21 and CYP2A6*22, showed 1.5- to 2.9-fold higher values than the wild type. As for the $V_{max}$ values, all the variants appeared to show higher values than the wild type, with fold differences of 1.8, 1.05, 1.4 and 1.15 in CYP2A6*15, CYP2A6*16, CYP2A6*21 and CYP2A6*22 respectively. Except for CYP2A6*15, the differences for the mutants were however not statistically significant from the wild type. Analysis of $V_{max}/K_m$ ratios which represent the intrinsic clearance of the enzymes indicated that variants CYP2A6*15, CYP2A6*16 and CYP2A6*21 exhibited similar $V_{max}/K_m$ (0.85- to 1.05-fold difference) to that of the wild type. Interestingly, CYP2A6*22 showed significantly much lower $V_{max}/K_m$ (39% of the wild type) indicating compromised catalytic activity. The lower value was mainly due to significantly higher apparent $K_m$ value of this mutant and a much lesser degree of increase seen in the $V_{max}$. 

14
Genetic polymorphism of CYP2A6 is believed to be the major cause of interindividual variation in enzymatic activity for various CYP2A6 substrates and it is thereby critical to characterize the enzymatic properties caused by the allelic polymorphism (Kitagawa et al., 2001). Therefore, construction of mutant alleles, CYP2A6*15, CYP2A6*16, CYP2A6*21 and CYP2A6*22 was accomplished in the present study with the purpose of exploring the functional consequences of polymorphism in these alleles. Currently allelic frequencies of the four alleles have been characterized in different populations; the functional consequences of point mutations in these variants, particularly in relation to catalytic activity, have not been investigated thoroughly. A recent study on CYP2A6*21 demonstrated very little or negligible impact of the mutation of the allele on in vivo nicotine metabolism in Caucasian subjects (Koudsi et al., 2006). Studies on the other three alleles, either in vitro or in vivo, have generally been lacking and merit further exploration.

_E. coli_ was chosen as the expression system in the present study because it is one of the most commonly used hosts in carrying out structural, biophysical and kinetic studies for CYPs. A host of CYP isoforms have been heterologously expressed in _E. coli_ for study of structure-function relationships by site-directed mutagenesis, including those of CYP2A6 (Kim et al., 2005). It is conceivable that the bacterial system is not amendable for studying post-translational event of the expressed protein, and this can only be investigated using other host cells such as yeast and mammalian cells. Experiments using these cell lines are currently in progress in our laboratory.
Of the four CYP2A6 mutants investigated in this study, all the amino acid mutations were found not only in substrate recognition sites (SRSs) but also located in other region of the open reading frame. A single point mutation which causes Lys194Glu (K194E) substitution in variant CYP2A6*15, is located not within any particular SRS site but adjacent to Helix F which partially embraces SRS-2 (see Figure 5). Kinetics study of K194E mutation revealed that the single nucleotide substitution affecting relatively more on $K_m$ of coumarin 7-hydroxylation than the $V_{max}$ value (i.e. 2.1-fold increase in $K_m$ but only 1.8-fold increase in $V_{max}$). However, catalytic efficiency of variant CYP2A6*15 was only slightly lower in comparison to the wild type (0.85-fold of the wild type value). This indicates that carriers of CYP2A6*15 allele would not be expected to exhibit difference in their drug clearance for CYP2A6 substrates.

Variant CYP2A6*16 on the other hand exhibits a single amino acid mutation, Arg203Ser (R203S) which is located within the Helix F and directly positioned within the highly conserved SRS-2 of CYP2A family (see Figure 5). Thus, it would be possible that mutation in helix F or specifically in SRS-2 may affect the folding of neighbouring secondary structures of active site, hence affecting the protein stability as well as the local haem and substrate binding. Moreover, replacement of basic Arg with a small and neutral residue Ser, a non-conservative substitution, would be expected to alter morphology of the active site. Our results however indicated that Arg203Ser did not alter substrate binding (i.e. no change in the $K_m$) as well as catalytic capacity (i.e. no change in $V_{max}$) resulting in similar catalytic efficiency for CP2A6. From molecular modeling point of view, it has been reported that coumarin orientates for 7-hydroxylation via hydrogen-bonded contacts with Gln104 and His477 and is $\pi$ stacking
with Phe209 (Lewis et al., 1999). In addition, Thr212 could be involved in directing the access of coumarin to the binding site (Fukami et al., 2004). From such analyses, it is likely that Arg203 residue does not have direct bonding contact with coumarin molecule for 7-hydroxylation, hence no change was observed in the catalytic behaviour of this mutant.

Substitution of Lys476 with another strongly basic residue, Arg, in allele CYP2A6*21 did not affect catalytic efficiency when compared to the wild type enzyme (i.e. 0.91-fold change in $V_{\text{max}}/K_m$) even though the mutation is located in the SRS-6 of $\beta$-sheet 4 (see Figure 5). In other words, the subtle difference of catalytic efficiency between variant CYP2A6*21 and the wild type is not appreciable despite alteration in the highly-conserved region of SRS-6. This finding seems to suggest that residue changes at the 476 site does not disrupt the conformation structure of CYP2A6*21 protein as the local residue polarity remained unchanged. Our finding seems to be consistent with data obtained by Koudsi and colleagues (Koudsi et al., 2006) which concluded that CYP2A6*21 did not have detectable impact on in vivo metabolism of nicotine, another CYP2A6 substrate.

The fourth allelic construct, CYP2A6*22 has been generated to carry two point mutations in accordance to the reported sequence in the literature. Both Asp158Glu (D158E) and Leu160Ile (L160I) substitutions are located in the D-helix, which appears to be exterior to the putative active site of CYP2A6 (Figure 5). Mutational effect was however observed in coumarin 7-hydroxylation with catalytic efficiency of CYP2A6*22 declined to 39% of the wild type, mainly by affecting the coumarin $K_m$ value (2.9-fold increase in $K_m$ versus 1.15-fold increase in $V_{\text{max}}$). The effects of these substitutions...
suggest that structural elements outside the active site may as well play roles in changing the catalytic activity through a variety of changes proposed in the previous studies, such as blocking of substrate or product access channels, reductase binding and the motion of protein during conformational changes (Kim and Guengerich, 2004a; Kim and Guengerich, 2004b). With the elucidation of CYP2A6 crystal structure (Yano et al., 2005), this has given us better understanding that Asp158 and Leu160 reside on the surface of CYP2A6 enzyme and does not make direct contact with coumarin. It is possible that mutations in helix D at positions 158 and 160 may have affected the folding of the neighbouring secondary structures, hence the effect on substrate binding and catalytic efficiency. It is likely that Asp158 and Leu160 together with other residues in helix D in their native configuration are able to provide a folding motif that locks other neighbouring secondary structures for proper heme and substrate binding. The three-dimensional structure of CYP2A6 as revealed in the X-ray crystallography (Yano et al., 2005) illustrates that helix D is located adjacent to many other structures that form part of the substrate access and binding sites, including the helices B, C, F and F’ as well as B-C loop. Consequently, alteration at 158 and 160 positions may have disrupted this part of the protein, causing altered association and spatial repositioning of these associated structures. In other words, Asp158 and Leu160 may be involved in ‘long-range’ interactions that are transmitted to residues in the surrounding structures which contact the heme and substrates. This type of interaction is not unique and has been reported for other CYP isoforms. For example, Glu351, a residue located in helix K of CYP21A2, has been found to be important for heme binding even though it is located far from the heme moiety in the active site. Glu351Lys mutation introduced to the
isoform was shown to result in the loss of enzyme activity towards the substrate progesterone (Krone et al., 2005). Result from this study is consistent with our findings and indicate that residues not involved in substrate and heme binding may play important functional or structural role in CYP catalysis.

In summary, kinetic data attained from this project has indeed given us insights particularly on the functional consequences of genetic polymorphisms in the four selected alleles (CYP2A6*15, CYP2A6*16, CYP2A6*21 and CYP2A6*22). Kinetic analyses of these polymorphic alleles of CYP2A6 indicated that point mutations harbored in these variants have encoded enzymes which were metabolically active towards coumarin oxidation, with the exception of CYP2A6*22 which has reduced but not inactive metabolic activity. Individual carriers of homologous CYP2A6*22 allele would be expected to have decreased clearance of coumarin. However, it is unclear whether other substrates of CYP2A6 such as nicotine would also be affected by this polymorphism as it is known that catalytic activity of alleles may differ according to the substrates being investigated. Further study is currently undertaken in our laboratory to elucidate enzyme kinetics of this variant with other substrates.
Acknowledgements

We would like to express our thanks and gratitude to Professor John Miners (Flinders University, Adelaide, Australia) and Professor Donald Birkett (Johnson & Johnson Research Pty Ltd., Sydney, Australia) for their kind gifts of vectors pCWori+ and pACYC-OxR.
References


Footnotes

This work was supported by the Science and Technology Research Grant from the Malaysia Toray Science Foundation (269817-K) as well as the International Medical University Research Fund (IMU 112/2006) from the International Medical University, Kuala Lumpur, Malaysia.
Legends for figures

Figure 1. Restriction analyses of wild type pCW-CYP2A6 in comparison with the mutant constructs (A) CYP2A6*15 by using enzyme EcoRI; (B) CYP2A6*16 by using enzyme FspI; (C) CYP2A6*21 by using enzyme PmlI; and (D) CYP2A6*22 by using enzyme Hpy99I. Restriction pattern following digestion by (A) EcoRI: A fragment sized 6448 bp for pCW-CYP2A6*15 and non-cutter for pCW-CYP2A6; (B) FspI: Two fragments sized 4640 bp and 1808 bp for pCW-CYP2A6*16 and three fragments sized 2433 bp, 2207 bp and 1808 bp for pCW-CYP2A6; (C) PmlI: Non-cutter for pCW-CYP2A6*21 and a single fragment sized 6448 bp for pCW-CYP2A6; (D) Hpy99I: The presence of 3462 bp fragment in Hpy99I-digested pCW-CYP2A6*22 was distinctive from the wild type pCW-CYP2A6 which produced two fragments of 2297 bp and 1165 bp. Several bands of less than 1000 bp were observed from both digestions with sizes of 974 bp, 592 bp, 510 bp and 413 bp. Lane 1 is the molecular weight markers (2-log DNA ladder or 100bp DNA ladder, New England Biolabs, USA).

Figure 2. Electropherograms showing the differences in the nucleotide sequences of the wild type CYP2A6 (CYP2A6*1) in comparison to CYP2A6 mutants, (A) CYP2A6*15, (B) CYP2A6*16, (C) CYP2A6*21 and (D) CYP2A6*22 respectively. Arrows indicate the positions of nucleotide changes. Note that electropherograms of the sense-strand (5’ end to 3’ end) are shown.
**Figure 3.** Immunoblot analyses of CYP2A6 expression in DH5α cells. The *E.coli* membrane proteins (50 μg) were separated on a 12% SDS-polyacrylamide gel, transferred to a nitrocellulose membrane and reacted with CYP2A6 antibody. The membranes were developed using peroxidase-conjugated goat anti-mouse IgG. Ladder: molecular weight markers (prestained protein standards, Bio-Rad, USA); Control (C): membrane fractions of the control cell expressing pCWori+ plasmid; WT: membrane fractions expressing co-transformed OxR and wild type *pCW-CYP2A6* plasmids; *15, *16, *21, *22: membrane fractions co-expressing OxR and the variant proteins as indicated on the top.

**Figure 4.** Reduced CO-difference spectra of the wild type CYP2A6 (continuous line) and CYP2A6*15 (broken line) proteins expressed in DH5α cell membranes. The spectra were recorded in 0.1 M phosphate buffer (pH 7.4) containing 20% glycerol in accordance to a published protocol by Omura and Sato (1964).

**Figure 5.** The amino acid sequence of CYP2A6, with the initial alignment of amino acids highlighted in gray representing membrane anchor region. Amino acids shown in bold are in α-helical conformation, and those in β-strands are shown in italics. The six substrate recognitions sites (SRSs) of CYP2A6 are underlined, whereas amino acids which form the boundaries of the active site cavities are shaded gray. Boxed residues are amino acids that were substituted with the mutated residues as reported in the four CYP2A6 alleles in the present study (adapted from Lewis, 2003; Johnson and Stout, 2005; Kim et al., 2005).
Table 1: Specifically-designed oligonucleotide primers which have been used in the generation of *CYP2A6*<sup>15</sup>, *CYP2A6*<sup>16</sup>, *CYP2A6*<sup>21</sup> and *CYP2A6*<sup>22</sup> mutant cDNAs.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer Sequence&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2A6&lt;sup&gt;15&lt;/sup&gt;For</td>
<td>5' CGC TTT GAC TAT AGG GAC AAA GAA TTC CTG TCA CTG 3'</td>
</tr>
<tr>
<td></td>
<td><em>EcoRI</em></td>
</tr>
<tr>
<td>CYP2A6&lt;sup&gt;16&lt;/sup&gt;For</td>
<td>5' TTC CTG TCA CTG TTG AGC ATG ATG CTA GGA ATC 3'</td>
</tr>
<tr>
<td></td>
<td><em>FspI</em></td>
</tr>
<tr>
<td>CYP2A6&lt;sup&gt;21&lt;/sup&gt;For</td>
<td>5' GAC GTG TCC CCC AGA CAT GTG GGC TTT GCC 3'</td>
</tr>
<tr>
<td></td>
<td><em>PmlI</em></td>
</tr>
<tr>
<td>CYP2A6&lt;sup&gt;22&lt;/sup&gt;For</td>
<td>5' GGC TTC CTC ATC GAG GCC ATC CGG GGC ACT GGC 3'</td>
</tr>
<tr>
<td></td>
<td><em>Hpy99I</em></td>
</tr>
</tbody>
</table>

* The underlined regions denote restriction sites which have been created or abolished in the primer sequences.
* Sequences shown are those of the forward primers, the reverse primers contained the opposite sequences to those of the forward primers and annealed to the same sequences on opposite strands of the plasmid template.
Table 2: Kinetic data of wild type and mutants of CYP2A6 in coumarin 7-hydroxylation.

<table>
<thead>
<tr>
<th>CYP protein</th>
<th>$V_{max}$ (pmol/min/pmol CYP)</th>
<th>$K_m$ (μM)</th>
<th>$V_{max}/K_m$ (μl/min/pmol, x $10^{-2}$)</th>
<th>Fold</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2A6*1 (wild-type)</td>
<td>1.70 ± 0.14</td>
<td>10.61 ± 1.58</td>
<td>16.03 ± 2.74</td>
<td>1.00</td>
</tr>
<tr>
<td>CYP2A6*15</td>
<td>3.03 ± 0.14*</td>
<td>22.13 ± 1.29**</td>
<td>13.70 ± 1.01</td>
<td>0.85</td>
</tr>
<tr>
<td>CYP2A6*16</td>
<td>1.78 ± 0.12</td>
<td>10.53 ± 1.45</td>
<td>16.87 ± 2.59</td>
<td>1.05</td>
</tr>
<tr>
<td>CYP2A6*21</td>
<td>2.40 ± 0.42</td>
<td>16.39 ± 0.41**</td>
<td>14.65 ± 2.60</td>
<td>0.91</td>
</tr>
<tr>
<td>CYP2A6*22</td>
<td>1.96 ± 0.42</td>
<td>31.16 ± 2.22**</td>
<td>6.28 ± 1.42**</td>
<td>0.39</td>
</tr>
</tbody>
</table>

All kinetic values are shown as the means ± SD from three independent determinations. The last column represents the fold of change in $V_{max}/K_m$, relative to wild type. Asterisks represent degree of statistical differences between the mutants and the wild type according to one-Way ANOVA test: The mean difference is significant at 0.05 level ($p < 0.05$) and highly significant at the 0.01 level ($p < 0.01$). *$p < 0.05$ and **$p < 0.01$. 

This article has not been copyedited and formatted. The final version may differ from this version.
Figure 1A

(A) Marker pCW-CYP2A6 pCW-CYP2A6*15

3.0 kbp
1.5 kbp
1.0 kbp
0.5 kbp
Figure 1C

(C)

<table>
<thead>
<tr>
<th>Marker</th>
<th>pCW-CYP2A6</th>
<th>pCW-CYP2A6*21</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.0 kbp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5 kbp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0 kbp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5 kbp</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This article has not been copyedited and formatted. The final version may differ from this version.
Figure 2A

(A) CYP2A6*1

5' CTA TAA GGA CAA AGA GTT CCT 3'

5' CTA TAA GGA CAA AGA ATT CCT 3'

CYP2A6*15
Figure 2B

(B)  

CYP2A6*1  
5' GTC ACT GTT GCC CAT GAT GCT 3'  

CYP2A6*16  
5' GTC ACT GTT GAG CAT GAT GCT 3'
Figure 2C

(C) CYP2A6*1

5’ GTC CCC CAA ACA CGT GGG CTT 3’

(C) CYP2A6*21

5’ GTC CCC CAG ACA TGT GGG CTT 3’
Figure 2D

(D) CYP2A6*1

5’ CTC ATC GAC GCC CTC CGG GGC 3’

CYP2A6*22

5’ CTC ATC GAG GCC ATC CGG GGC 3’
Figure 3
Figure 4

Absorbance (A) vs. Wavelength (nm)
Figure 5

MLASGMLLVA LLVCLTVVML MSVWQQRKSK GKLPPGPTPL PFIGNYLQLN

αA

β1-1

β1-2

αB

51

TEQMYNSLMK ISERYGPVFT IHLGPRRVVV LCGHDAVREA LVDQAEFSG

β1-5

αB’

αC

αD

101

RGEQATEDNV FKGYGVVFSN GERAKQLRRF SIATLRDFGV GKRGIEERIQ

*22

αE

*15

αF

151

eEAGFLIDAL RGTGGANIDP TFFLSRTVSN VISSIVVGDR FDYK3KEFLS

*16

SRS2

αF’

αG’

SRS3

αG

201

LRRMMLGIFQ FTSTSTGQLY EMFSSVMKHL PGPQQAFQL LQGLEDFIAK

αH

SRS4

251

KVEHNQRTLĐ PNSPRDFIDS FLIRMQEEEK NPNTEFYLKN LVMTTLNLFL

αI

αJ

αJ’

301

GGTETVSTTL RYGFLLLMKH PEVEAKVHEE IDRVIDKNRQ PKFEDRAKMP

αK

SRS5

β1-3

β2-1

β2-2

β1-4

αK’

351

YMEA\text{IHEI}Q RFGDVP\text{IPMSL} ARRVKKDTKF R\text{DFPLK}GTE VYP\text{MGLS}VL

αL

401

DPSFFSNPQD FNPQHFLNEK GQFKKSDAFV PF\text{SIGKRNC}F GEGLARMELF

*21

β3-1

β4-1

SRS6

β4-2

β3-2

451

LFFTTVMQNF RLKSSQSPKD IDV\text{S}PH\text{HG}VGF ATIPRNYTMS FLPR