# MOLECULAR MODELLING OF CYP2B6 BASED ON HOMOLOGY WITH THE CYP2C5 CRYSTAL STRUCTURE: ANALYSIS OF ENZYME-SUBSTRATE INTERACTIONS

David F.V. Lewis[1]*, Brian G. Lake[2], Maurice Dickins[3] and Peter S. Goldfarb[1]

*[1]School of Biomedical and Life Sciences, University of Surrey, Guildford, Surrey, [2]BIBRA International Limited, Carshalton, Surrey and [3]GlaxoSmithKline Research and Development Limited, Ware, Herts, UK*

**SUMMARY**

The results of homology modelling of CYP2B6 based on the CYP2C5 crystal structure is described in terms of substrates and inhibitors binding within the putative active site. In general these results are in agreement with currently available evidence from substrate metabolism, mode of inhibitor action and site-directed mutagenesis experiments within the CYP2B subfamily of enzymes. Consequently, the model based on the CYP2C5 template represents an advance on those models produced from bacterial P450s, such as CYP101 and CYP102. Quantitative Structure-Activity Relationships (QSARs) for substrates binding to CYP2B6 indicate a key role for hydrogen bonding, and lipophilic character, as determined by the log P parameter (where P is the octanol/water partition coefficient), is also of importance for explaining the variation in experimental binding affinity for CYP2B6 substrates. It is possible to estimate the binding energies for typical CYP2B6 substrates based on their properties and interactions with the enzyme, which show good concordance with experimental data in the form of apparent $K_m$ values.

* Author for correspondence:
David F.V. Lewis
School of Biomedical and Life Sciences
University of Surrey
Guildford, Surrey, GU2 7XH, UK
e-mail: d.lewis@surrey.ac.uk

**KEY WORDS**

cytochrome P450, xenobiotic metabolism and activation, modelling, CYP2 structural template

## INTRODUCTION

The cytochromes P450 (CYP) represent a superfamily of haem-thiolate enzymes, for which over 2,000 individual members are currently known, and have been found in most species ranging from bacteria to mammals (reviewed in /1-3/). Phase 1 oxidations of many drug substrates are mediated by P450s from families CYP1, CYP2 and CYP3 in mammalian species, including *Homo sapiens* /4/, with the CYP2 family being particularly important with respect to polymorphic drug metabolism in man /5-7/. Within this family, the CYP2B subfamily is known to be involved in the metabolism of drug substrates such as bupropion, mephenytoin, diazepam and benzphetamine /4,8/. The human enzyme, CYP2B6, specifically metabolizes 4-trifluoromethyl-7-ethoxycoumarin by O-deethylation, and recently thiotriethylenephosphoramide (thioTEPA) has been identified as a selective inhibitor of CYP2B6 /9/. Although not regarded as a major contributor to drug metabolism pathways in man, there is, nevertheless, interest in CYP2B6 substrate selectivity /8,10/ due to its key importance in certain cases such as buproprion where it is regarded as the main catalyst of compound metabolism, and Table 1 presents a summary of several CYP2B6 substrates. Following the publication of the CYP2C5 crystal structure /11/ there is an opportunity to utilize this as a template for homology modelling of other mammalian P450s, especially those within the CYP2 family /12/ in which the sequence homologies are relatively high (50-80%). In this respect, it can be anticipated that CYP2C5 represents an improved template compared to CYP102 which was employed in our previous investigations of CYP2B enzymes /13/. Consequently, we have employed the crystallographic coordinates of CYP2C5 in constructing a three-dimensional model of CYP2B6, the human CYP2B orthologue, such that selective substrate interactions can be investigated. A recent QSAR study of CYP2B6 substrates /10/ also utilized CYP2C5 as a template for modelling the CYP2B6 active site region, and the current study involves homology modelling of the entire enzyme, minimizing

the 3-D structure, and then investigating substrate interactions. However, QSAR studies have also been conducted and estimates of binding energy have been undertaken based on the key interactions between substrates and the CYP2B6 active site based on the modelling. This approach consolidates and reinforces the homology modelling of CYP2B6, a human enzyme which is associated with the Phase 1 metabolism of several drug substrates and other exogenous compounds /14/.

## METHODS

Figure 1 shows a multiple sequence alignment for CYP2 family enzymes including that of CYP2B6 and the template, CYP2C5. The alignment was generated using the GCG software package (Genetics Computer Group, Madison, Wisconsin) and a small degree of manual editing was employed in order to conserve secondary structural motifs (such as the α-helices) with respect to the template structure, although very little change was required. The location of residues which have been the subject of site-directed mutagenesis experiments across the CYP2 family have been emboldened in the multiple sequence alignment presented in Figure 1. The crystal structure of CYP2C5 (pdb code: 1dt6) was employed in the homology modelling of CYP2B6 based on the alignment shown in Figure 1. This involved residue replacement using the Sybyl Biopolymer package (Tripos Associates, St. Louis, Missouri) as required by the alignment shown in Figure 1, together with a very small number of deletions and insertions. Loop searching of the protein databank was employed for insertion of residues, although these were few in number and only involved up to three residues, and the raw model was then inspected for unfavourable steric interactions caused by the replacement process. Following initial manual refinement of the model to reduce unfavourable steric interactions, the entire structure was energy minimized using molecular mechanics (Tripos force field) to achieve a low energy conformation of satisfactory geometry in terms of allowable protein bond lengths and angles. The final model of CYP2B6 was then probed using selective substrates and inhibitors via interactive docking. This procedure was facilitated by the location of bound substrate in the CYP2C5 crystal structure /11/ and from information on site-directed mutagenesis of key residues likely to be involved in substrate binding for CYP2B enzymes (reviewed in /15,16/).

```
                membrane   anchor   region
                1          10         20          30          40
cyp2a1   ...MLDTGLL LVVILASLSV MLLVSLWQQ. KIRGRLPPGP TPLPFIGNYL
cyp2a4   ...MLTSGLL LVAAVAFLSV LVLMSVWKQR KLSGKLPPGP TPLPFVGNFL
cyp2a5   ...MLTSGLL LVAAVAFLSV LVLMSVWKQR KLSGKLPPGP TPLPFIGNFL
cyp2a6   ...MLASGML LVALLVCLTV MVLMSVWQQR KSKGKLPPGP TPLPFIGNYL
cyp2b1   .......MEPT ILLLLALLVG FLLLLVRGHP KSRGNFPPGP RPLPLLGNLL
cyp2b4   .......MEFS LLLLLAFLAG LLLLLFRGHP KAHGRLPPGP SPLPVLGNLL
cyp2b6   .......MELS VLLFLALLTG LLLLLVQRHP NTHDRLPPGP RPLPLLGNLL
cyp2c2   .......MDL VVVLGLCLSC LLLPSLWKQS HGGGKLPPGP TPFPILGNVL
cyp2c3   .......MDL LIILGICLSC VVLLSLWKKT HGKGKLPPGP TPLPVVGNLL
cyp2c4   .......MDP VAGLVLGLCC LLLLSLWKQN SGRGKLPPGP TPFPIIGNIL
cyp2c5   .......MDP VVVLVLGLCC LLLLSIWKQN SGRGKLPPGP TPFPIIGNIL
cyp2c8   .......MEP FVVLVLCLSF MLLFSLWRQS CRRRKLPPGP TPLPIIGNML
cyp2c9   .......MDS LVVLVLCLSC LLLLSLWRQS SGRGKLPPGP TPLPVIGNIL
cyp2c19  .......MDP FVVLVLCLSC LLLLSIWRQS SGRGKLPPGP TPLPVIGNIL
cyp2d1   MELLNGTGLW SMAIFTVIFI LLVDLMHRRH RWTSRYPPGP VPWPVLGNLL
cyp2d6   ...MGLEALV PLAVIVAIFL LLVDLMHRRQ RWAARYPPGP LPLPGLGNLL
cyp2e1r  ....MAVLGI TVALLGWMVI LLFISVWKQI HSSWNLPPGP FPLPIIGNLL
cyp2e1h  .....MSALGV TVALLVWAAF LLLVSMWRQV HSSWNLPPGP FPLPIIGNLF
cyp2f1   ......MDSI STAILLLLLA LVCLLLTLSS RDKGKLPPGP RPLSILGNLL
cyp2f2   ......MDGV STAILLLLLA VISLSLTFSS RDKGQLPPGP KPLPILGNLL

                < Ahelix >      |1-1|  | 1-2 |   <   B helix   >
                50         60         70          80          90
cyp2a1   QLNTKDVYSS ITQLSERYGP VFTIHLGPRR VVVLYGYDAV KEALVDQAEE
cyp2a4   QLNTEQMYNS LMKISQRYGP VFTIYLGSRR IVVLCGQEAV KEALVDQAEE
cyp2a5   QLNTEQMYNS LMKISQRYGP VFTIYLGPRR IVVLCGQEAV KEALVDQAEE
cyp2a6   QLNTEQMYNS LMKISERYGP VFTIHLGPRR VVVLCGHDAV REALVDQAEE
cyp2b1   QLDRGGLLNS FMQLREKYGD VFTVHLGPRP VVMLCGTDTI KEALVGQAED
cyp2b4   QMDRKGLLRS FLRLREKYGD VFTVYLGSRP VVVLCGTDAI REALVDQAEA
cyp2b6   QMDRRGLLKS FLRFREKYGD VFTVHLGPRP VVMLCGVEAI REALVDKAEA
cyp2c2   QLDFKDLSKS LTNLSKVYGP VFTVYLGMKP TVVVHGYEAV KEALVDLGHE
cyp2c3   QLETKDINKS LSMLAKEYGS IFTLYFGMKP AVVLYGYEGV IEALIYRGEE
cyp2c4   QIDVKDISKS LTKFSERYGP VFTVYLGMKP TVVLHGYKAV KEALVDLGEE
cyp2c5   QIDAKDISKS LTKFSECYGP VFTVYLGMKP TVVLHGYEAV KEALVDLGEE
cyp2c8   QIDVKDICKS FTNFSKVYGP VFTVYFGMNP IVVFHGYEAV KEALIDNGEE
cyp2c9   QIGIKDISKS LTNLSKVYGP VFTLYFGLKP IVVLHGYEAV KEALIDLGEE
cyp2c19  QIDIKDVSKS LTNLSKIYGP VFTLYFGLER MVVLHGYEVV KEALIDLGEE
cyp2d1   QVDLSNMPYS LYKLQHRYGD VFSLQKGWKP MVIVNRLKAV QEVLVTHGED
cyp2d6   HVDFQNTPYC FDQLRRRFGD VFSLQLAWTP VVVLNGLAAV REALVTHGED
cyp2e1r  QLDLKDIPKS FGRLAERFGP VFTVYLGSRR VVVLHGYKAV REMLLNHKNE
cyp2e1h  QLELKNIPKS FTRLAQRFGP VFTLYVGSQR MVVMHGYKAV KEALLDYKDE
cyp2f1   LLCSQDMLTS LTKLSKEYGS MYTVHLGPRR VVVLSGYQAV KEALVDQGEE
cyp2f2   QLRSQDLLTS LTKLSKEYGS VFTVYLGSRP VIVLSGYQTV KEALVDKGEE
```

**Fig. 1:** A multiple sequence alignment of CYP2 family proteins including that of CYP2B6 and of the crystallographic template, CYP2C5. The substrate recognition sites (SRS regions) are indicated as shaded boxes, together with the N-terminal membrane anchor peptide. Residues which have been probed using site-directed mutagenesis are emboldened, and the numbering system employed is that of the CYP2C5 enzyme.

```
                    SRS1                          <  C helix >
           100              110              120              130
cyp2a1  FSGRGEQATY NTLEKG...Y GVAFSS.GER AKQLRRLSIA TLRDFGVGKR
cyp2a4  FSGRGEQATE DWLFKG...Y GIAFSS.GER AKQLRSFSIA TLRDFGVGKR
cyp2a5  FSGRGEQATE DWLFKG...Y GVVFSS.GER AKQLRRFSIA TLRDFGVGKR
cyp2a6  FSGRGEQATE DWVFKG...Y GVVFSN.GER AKQLRRFSIA TLRDFGVGKR
cyp2b1  FSGRGTxxxE EPIFKE...Y GVIFAN.GER WKALRRFSLA TMRDFGMGKR
cyp2b4  FSGRGKIAVV DPIFQG...Y GVIFAN.GER WRALRRFSLA TMRDFGMGKR
cyp2b6  FSGRGKIAMV DPFFRG...V GVIFAN.GNR WKVLRRFSVT TMRDFGMGKR
cyp2c2  LSGRSRFLVT AKLNKG...F GVIFSN.GKR WTETRRFSLM TLRNFGMGKR
cyp2c3  FSGRGIFFVF DRVFKG...L GIVFSS.GEK WKETRRFSLT VLRNLGMGKK
cyp2c4  FAGRGHxxxx EKVNKG...L GIVFTN.ANT WKEMRRFSLM TLRNFGMGKR
cyp2c5  FAGTGSVPIL EKVSKG...I GIAFSN.AKT WKEMRRFSLM TLRNFGMGKR
cyp2c8  FSGRGNSPIS QRITKG...L GIISSN.GKR WKEIRRFSLT TLRNFGMGKR
cyp2c9  FSGRGIFFLA ERANRG...F GIVFSN.GKK WKEIRRFSLM TLRNFGMGKR
cyp2c19 FSGRGHxPLA ERANRG...F GIVFSN.GKR WKEIRRFSLM TLRNFGMGKR
cyp2d1  TADRPPVPIF KCLGVRPRSQ GVILASYGPE WREQRRFSVS TLRTFGMGKK
cyp2d6  TADRPPVPIT QILGFGPRSQ GVFLARYGPA WREQRRFSVS TLRNLGLGKK
cyp2e1r FSGRGEIPAF .REFKD...K GIIFNN.GPT WKDTRRFSLT TLRDYGMGKQ
cyp2e1h FSGRGDLPAF .HAHRD...R GIIFNN.GPT WKDIRRFSLT TLRNYGMGKQ
cyp2f1  FSGRGDYPAF FNFTKG...N GIAFSS.GDR WKVLRQFSIQ ILRNFGMGKR
cyp2f2  FSGRGAYPVF FNFTRG...N GIAFSD.GER WKILRRFSVQ ILRNFGMGKR

         <     D helix    >          < D'> < Ehelix >
        140        150        160          170        180
cyp2a1  GVEERILEEA GYLIKMLQGT CGAPIDPTIY LSKTVSNVIS SIVFGERFDY
cyp2a4  GIEERIQEEA GFLIDSFRKT NGAFIDPTFY LSRTVSNVIS SIVFGDRFDY
cyp2a5  GIEERIQEEA GFLIDSFRKT NGAFIDPTFY LSRTVSNVIS SIVFGDRFDY
cyp2a6  GIEERIQEEA GFLIDAHRGT GGANIDPTFF LSRTVSNVIS SIVFGDRFDY
cyp2b1  SVEERIQEEA QCLVEELRKS QGAPLDPTFL FQCITANIIC SIVFGERFDY
cyp2b4  SVEERIQEEA RCLVEELRKS KGALLDNTLL FHSITSNIIC SIVFGKRFDY
cyp2b6  SVEERIQEEA QCLIEELRKS KGALMDPTFL FQSITANIIC SIVFGKRFHY
cyp2c2  SIEERVQEEA HCLVEELRKT NASPCDPTFI LGAAPCNVIC SVIFQNRFDY
cyp2c3  TIEERIQEEA LCLIQALRKT NASPCDPTFL LFCVPCNVIC SVIFQNRFDY
cyp2c4  SIEDRVQEEA RCLVEELRKT NALPCDPTFI LGCAPCNVIC SVILHNRFDY
cyp2c5  SIEDRIQEEA RCLVEELRKT NASPCDPTFI LGCAPCNVIC SVIFHNRFDY
cyp2c8  SIEDRVQEEA HCLVEELRKT KASPCDPTFI LGCAPCNVIC SVVFQKRFDY
cyp2c9  SIEDRVQEEA RCLVEELRKT KASPCDPTFI LGCAPCNVIC SIIFHKRFDY
cyp2c19 SIEDRVQEEA RCLVEELRKT KASPCDPTFI LGCAPCNVIC SIIFQKRFDY
cyp2d1  SLEEWVTKEA GHLCDAFTAQ AGQSINPKAM LNKALCNVIA SLIFARRFEY
cyp2d6  SLEQWVTEEA ACLCAAFANH SGRPFRPNGL LDKAVSNVIA SLTCGRRFEY
cyp2e1r GNEDRIQKEA HFLLEELRKT QGQPFDPTFV IGCTPFNVIA KILFNDRFDY
cyp2e1h GNESRIQREA HFLLEALRKT QGQPFDPTFL IGCAPCNVIA DILFRKHFDY
cyp2f1  SIEERILEEG SFLLADVRKT EGEPFDPTFV LSRSVSNIIC SVLFGSRFDY
cyp2f2  SIEERILEEG SFLLEVLRKM EGKPFDPVFI LSRSVSNIIC SVVFGSRFDY
```

**Fig. 1** *cont.*

```
          < F helix SRS2 >     <  F' >          <       SRS3
         190        200        210        220        230
cyp2a1   EDTEFLSLLQ MMGQMNRFAA SPTGQLYDMF HSVMKYLPGP QQQIIKVTQK
cyp2a4   EDKEFLSLLR MMLGSLQFTA TSMGQVYEMF SSVMKHLPGP QQQAFKELQG
cyp2a5   EDKEFLSLLR MMLGSFQFTA TSMGQLYEMF SSVMKHLPGP QQQAFKELQG
cyp2a6   KDKEFLSLLR MMLGIFQFTS TSTGQLYEMF SSVMKHLPGP QQQAFQLLQG
cyp2b1   TDRQFLRLLE LFYRTFSLLS SFSSQVFEFF SGFLKYFPGA HRQISKNLQE
cyp2b4   KDPVFLRLLD LFFQSFSLIS SFSSQVFELF PGFLKHFPGT HRQIYRNLQE
cyp2b6   QDQEFLKMLN LFYQTFSLIS SVFGQLFELF SGFLKYFPGA HRQVYKNLQE
cyp2c2   TDQDFLSLMG KFNENFKILN SPWVQFCNCF PILFDYFPGS HRKAVKNIFY
cyp2c3   DDEKFKTLIK YTHENFELLG TPWIQLYNIF PILGHYLPGS HRQLFKNIDG
cyp2c4   KDEEFLKLME RLNENIRILS SPWLQVYNNF PALLDYFPGI HKTLLKNADY
cyp2c5   KDEEFLKLME SLNENVRILS SPWLQVYNNF PALLDYFPGI HKTLLKNADY
cyp2c8   KDQNFLTLMK RFNENFRILN SPWIQVCNNF PLLIDCFPGT HNKVLKNVAL
cyp2c9   KDQQFLNLME KLNENIKILS SPWIQICNNF SPIIDYFPGT HNKLLKNVAF
cyp2c19  KDQQFLNLME KLNENIRIVS TPWIQICNNF PTIIDYFPGT HNKLLKNLAF
cyp2d1   EDPYLIRMVK LVEESLTEVS GFIPEVLNTF PALLR.IPGL ADKVFQGQKT
cyp2d6   DDPRFLRILD LAQEGLKEES GFLREVLNAV PVLLH.IPAL AGKVLRFQKA
cyp2e1r  KDKQALRLMS LFNENFYLLS TPWLQVYNNF SNYLQYMPGS HRKVIKNVSE
cyp2e1h  NDEKFLRLMY LFNENFHLLS TPWLQLYNNF PSFLHYLPGS HRKVIKNVAE
cyp2f1   DDERLLTIIR LINDNFQIMS SPWGELYDIF PSLLDWVPGP HQRIFQNFKQ
cyp2f2   DDERLLTIIH FINDNFKIMS SPWGEMYNIF PSVLDWIPGP HKRLFRNFGG

            G helix   >          < H  >                <
         240        250        260        270        280
cyp2a1   LEDFMIEKVR QNHSTLDPN. SPRNFIDSFL IRMQEE.KNG NSEFHMKNLV
cyp2a4   LEDFITKKVE HNQRTLDPN. SPRDFIDSFL IRMLEEKKNP NTEFYMKNLV
cyp2a5   LEDFITKKVE HNQRTLDPN. SPRDFIDSFL IRMLEEKKNP NTEFYMKNLV
cyp2a6   LEDFIAKKVE HNQRTLDPN. SPRDFIDSFL IRMQEEEKNP NTEFYLKNLV
cyp2b1   ILDYIGHIVE KHRATLDPS. APRDFIDTYL LRMEKEKSNH HTEFHHENLM
cyp2b4   INTFIGQSVE KHRATLDPS. NPRDFIDVYL LRMEKDKSDP SSEFHHQNLI
cyp2b6   INAYIGHSVE KHRETLDPS. APKDLIDTYL LHMEKEKSNA HSEFSHQNLN
cyp2c2   VKNYITEQIK EHQKSLDIN. NPRDFIDCFL IKMEQEKCNQ QSEFTIENLL
cyp2c3   QIKFILEKVQ EHQESLDSN. NPRDFVDHFL IKMEKEKHKK QSEFTMDNLI
cyp2c4   TKNFIMEKVK EHQKLLDVN. NPRDFIDCFL IKMEKE...N NLEFTLGSLV
cyp2c5   IKNFIMEKVK EHEKLLDVN. NPRDFIDCFL IKMEQE...N NLEFTLESLV
cyp2c8   TRSYIREKVK EHQASLDVN. NPRDFIDCFL IKMEQEKDNQ KSEFNIENLV
cyp2c9   MKSYILEKVK EHQESMDMN. NPQDFIDCFL MKMEKEKHNQ PSEFTIESLE
cyp2c19  MESDILEKVK EHQESMDIN. NPRDFIDCFL IKMEKEKQNQ QNEFTIENLV
cyp2d1   FMALLDNLLA ENRTTWDPAQ PPRNLTDAFL AEVEKAKGNP ESSFNDENLR
cyp2d6   FLTQLDELLT EHRMTWDPAQ PPRDLTEAFL AEMEKAKGNP ESSFNDENLR
cyp2e1r  IKEYTLARVK EHHKSLDPS. CPRDFIDSLL IEMEKDKHST EPLYTLENIA
cyp2e1h  VKEYVSERVK EHHQSLDPN. CPRDLTDCLL VEMEKEKHSA ERLYTMDGIT
cyp2f1   LRDLIAHSVH DHQASLDPR. SPRDFIQCFL TKMAEEKEDP LSHFHMDTLL
cyp2f2   MKDLIARSVR EHQDSLDPN. SPRDFIDCFL TKMAQEKQDF LSHFNMDTLL
```

**Fig. 1** *cont.*

```
          SRS4    I helix              > <   J helix     >
            290         300         310         320         330
cyp2a1   MTTLSLFFAG SETVSSTLRY GFLLLMKHPD VEAKVHEEIE QVIGRNRQPQ
cyp2a4   LTTLNLFFAG TETVSTTLRY GFLLLMKYPD IEAKVHEEID RVIGRNRQPK
cyp2a5   LTTLNLFFAG TETVSTTLRY GFLLLMKHPD IEAKVHEEID RVIGRNRQPK
cyp2a6   MTTLNLFIGG TETVSTTLRY GFLLLMKHPE VEAKVHEEID RVIGKNRQPK
cyp2b1   ISLLVLFFAG TETSSTTLRY GFLLMLKYPH VAEKVQKEID QVIGSHRLPT
cyp2b4   LTVLSLFFAG TETTSTTLRY GFLLMLKYPH VTERVQKEIE QVIGSHRPPA
cyp2b6   LNTLSLFFAG TETTSTTLRY GFLLMLKYPH VAERVYREIE QVIGPHRPPE
cyp2c2   TTVSDVFMAG TETTSTTLRY GLLLLMKHPE VIAKVQEEIE RVIGRHRSPC
cyp2c3   TTIWDVFSAG TDTTSNTLKF ALLLLLKHPE ITAKVQEEIE HVIGRHRSPC
cyp2c4   IAVFDLFGAG TETTSTTLRY SLLLLLKHPE VAARVQEEIE RVIGRHRSPC
cyp2c5   IAVSDLFGAG TETTSTTLRY SLLLLLKHPE VAARVQEEIE RVIGRHRSPC
cyp2c8   CTVADLFVAG TETTSTTLRY GLLLLLKHPE VTAKVQEEID HVIGRHRSPC
cyp2c9   NTAVDLFGAG TETTSTTLRY ALLLLLKHPE VTAKVQEEIE RVIGRNRSPC
cyp2c19  ITAADLLGAG TETTSTTLRY ALLLLLKHPE VTAKVQEEIE RVIGRNRSPC
cyp2d1   MVVVDLFTAG MVTTATTLTW ALLLMILYPD VQRRVQQEID EVIGQVRCPE
cyp2d6   IVVADLFSAG MVTTSTTLAW GLLLMILHPD VQRRVQQEID DVIGQVRRPE
cyp2e1r  VTVADMFFAG TETTSTTLRY GLLILLKHPE IEEKLHEEID RVIGPSRMPS
cyp2e1h  VTVADLFFAG TETTSTTLRY GLLILMKYPE IEEKLHEEID RVIGPSRIPA
cyp2f1   MTTHNLLFGG TKTVSTTLHH AFLALMKYPK VQARVQEEID LVVGRARLPA
cyp2f2   MTTHNLLFGG TETVGTTLRH AFLILMKYPK VQARVQEEID RVVGRSRMPT

         < J'>  <     K helix    > SRS5    |1-4|              |1-3|
            340         350         360         370         380
cyp2a1   YEDHMKMPYT QAVINEIQRF SNLAPLGIPR RIIKNTTFRG FFLPKATDVF
cyp2a4   YEDRMKMPYT EAVIHEIQRF ADLIPMGLAR RVTKDTKFRD FLLPKGTEVF
cyp2a5   YEDRMKMPYT EAVIHEIQRF ADMIPMGLAR RVTKDTKFRD FLLPKGTEVF
cyp2a6   FEDRAKMPYM EAVIHEIQRF GDVIPMSLAR RVKKDTKFRD FFLPKGTEVY
cyp2b1   LDDRSKMPYT DAVIHEIQRF SDLVPIGVPH RVTKDTMFRG YLLPKNTEVY
cyp2b4   LDDRAKMPYT DAVIHEIQRL GDLIPFGVPH TVTKDTQFRG YVIPKNTEVF
cyp2b6   LHDRAKMPYT EAVIYEIQRF SDLLPMGVPH IVTQHTSFRG YIIPKDTEVF
cyp2c2   MQDRSRMPYT DATVHEIQRY INLIPNNVPH TTICNLKFRN YLIPKGTDVL
cyp2c3   SQDRSRMPYT DAVMHEIQRY VDLVPTSLPH AVTQDIEFNG YLIPKGTDII
cyp2c4   MQDRSHMPYT DAVIHEIQRF IDLLPTNLPH AVTRDVKFRN YFIPKGTDII
cyp2c5   MQDRSRMPYT DAVIHEIQRF IDLLPTNLPH AVTRDVRFRN YFIPKGTDII
cyp2c8   MQDRSHMPYT DAVVHEIQRY SDLVPTGVPH AVTTDTKFRN YLIPKGTTIM
cyp2c9   MQDRSHMPYT DAVVHEVQRY IDLLPTSLPH AVTCDIKFRN YLIPKGTTIL
cyp2c19  MQDRGHMPYT DAVVHEVQRY IDLIPTSLPH AVTCDVKFRN YLIPKGTTIL
cyp2d1   MTDQAHMPYT NAVIHEVQRF GDIAPLNLPR FTSCDIEVQD FVIPKGTTLI
cyp2d6   MGDQAHMPYT TAVIHEVQRF GDIVPLGVTH MTSRDIEVQG FRIPKGTTLI
cyp2e1r  VRDRVQMPYM DAVVHEIQRF IDLVPSNLPH EATRDTTFQG YVIPKGTVVI
cyp2e1h  IKDRQEMPYM DAVVHEIQRF TTLVPSNLPH EATRDTIFRG YLIPKGTVVV
cyp2f1   LKDRAAMPYT DAVIHEVQRF ADIIPMNLPH RVTRDTAFRG FLIPKGTDVI
cyp2f2   LEDRTSMPYT DAVIHEVQRF ADVIPMNLPH RVTRDTPFRG FLIPKGTDVI
```

**Fig. 1** *cont.*

```
          < K'>  <K">                                              <
            390        400        410        420        430
cyp2a1   PILGSLMTDP KFFPSPKDFD PQNFLDDKGQ LKKNAAFLPF STGKRFCLGD
cyp2a4   PMLGSVLKDP KFFSNPKDFN PKHFLDDKGQ FKKSDAFVPF SIGKRYCFGE
cyp2a5   PMLGSVLKDP KFFSNPKDFN PKHFLDDKGQ FKKNDAFVPF SIGKRYCFGE
cyp2a6   PMLGSVLRDP SFFSNPQDFN PQHFLNEKGQ FKKSDAFVPF SIGKRNCFGE
cyp2b1   PILSSALHDP QYFDHPDSFN PEHFLDANGA LKKSEAFMPF STGKRICLGE
cyp2b4   PVLSSALHDP RYFETPNTFN PGHFLDANGA LKRNEGFMPF SLGKRICLGE
cyp2b6   LILSTALHDP HYFEKPDAFN PDHFLDANGA LKKTEAFIPF SLGKRICLGE
cyp2c2   TSLSSVLHDD KEFPNPDRFD PGHFLDASGN FRKSDYFMPF STGKRVCVGE
cyp2c3   PSLTSVLYDD KEFPNPEKFD PGHFLDESGN FKKSDYFMPF STGKRACVGE
cyp2c4   TSLTSVLHDE KAFPNPKVFD PGHFLDESGN FKKSDYFMPF SAGKRMCVGE
cyp2c5   TSLTSVLHDE KAFPNPKVFD PGHFLDESGN FKKSDYFMPF SAGKRMCVGE
cyp2c8   ALLTSVLHDD KEFPNPNIFD PGHFLDKNGN FKKSDYFMPF SAGKRICAGE
cyp2c9   ISLTSVLHDN KEFPNPEMFD PHHFLDEGGN FKKSKYFMPF SAGKRICVGE
cyp2c19  TSLTSVLHDN KEFPNPEMFD PRHFLDEGGN FKKSNYFMPF SAGKRICVGE
cyp2d1   INLSSVLKDE TVWEKPHRFH PEHFLDAQGN FVKHEAFMPF SAGRRACLGE
cyp2d6   TNLSSVLKDE AVWEKPFRFH PEHFLDAQGH FVKPEAFLPF SAGRRACLGE
cyp2e1r  PTLDSLLYDK QEFPDPEKFK PEHFLNEEGK FKYSDYFKPF SAGKRVCVGE
cyp2e1h  PTLDSVLYDN QEFPDPEKFK PEHFLNENGK FKYSDYFKPF STGKRVCAGE
cyp2f1   TLLNTVHYDP SQFLTPQEFN PEHFLDANQS FKKSPAFMPF SAGRRLCLGE
cyp2f2   TLLNTVHYDS DQFKTPQEFN PEHFLDDNHS FKKSPAFMPF SAGRRLCLGE

            L helix     > |3-1|              SRS6          |3-2|
            440        450        460        470        480
cyp2a1   GLAKMELFLL LTTILQNFRF KFPMKLEDIN ESPKPLGFTR IIPKYTMSFMPI
cyp2a4   GLARMELFLF LTNIMQNFHF KSTQAPQDID VSPRIVGFVT IPPTYTMSFLSR
cyp2a5   GLARMELFLF LTNIMQNFHF KSTQAPQDID VSPRLVGFAT IPPTYTMSFLSR
cyp2a6   GLARMELFLF FTTVMQNFRL KSSQSPKDID VSPKHVGFAT IPRNYTMSFLPR
cyp2b1   GIARNELFLF FTTILQNFSV SSHLAPKDID LTPKESGIGK IPPTYQICFSAR
cyp2b4   GIARTELFLF FTTILQNFSI ASPVPPEDID LTPRESGVGN VPPSYQIRFLAR
cyp2b6   GIARAELFLF FTTILQNFSM ASPVAPEDID LTPQECGVGK IPPTYQIRFLPR
cyp2c2   ALARMELFLF LTAILQNFTP KPLVNPNNVD ENPFSSGIVR VPPLYRVSFIPV
cyp2c3   GLARMELFLL LTTILQHFTL KPLVDPKDID PTPVENGFVS VPPSYELCFVPV
cyp2c4   GLARMELFLF LTSILQNFKL QSLVEPKDLD ITAVVNGFVS VPPSYQLCFIPI
cyp2c5   GLARMELFLF LTSILQNFKL QSLVEPKDLD ITAVVNGFVS VPPSYQLCFIPI
cyp2c8   GLARMELFLF LTTILQNFNL KSVDDLKNLN TTAVTKGIVS IPPSYQICFPIV
cyp2c9   ALAGMELFLF LTSILQNFNL KSLVDPKNLD TTPVVNGFAS VPPFYQLCFIPV
cyp2c19  GLARMELFLF LTFILQNFNL KSLIDPKDLD TTPVVNGFAS VPPFYQLCFIPV
cyp2d1   PLARMELFLF FTCLLQRFSF .SVPVGQPRP STHGFFAFPV APLPYQLCAVVREQGL
cyp2d6   PLARMELFLF FTSLLQHFSF .SVPTGQPRP SHHGVFAFLV SPSPYELCAVPR
cyp2e1r  GLARMELFLL LSAILQHFNL KPLVDPEDID LRNITVGFGR VPPRYKLCVIPRS
cyp2e1h  GLARMELFLL LCAILQHFNL KPLVDPKDID LSPIHIGFGC IPPRYKLCVIPRS
cyp2f1   LLARMELFLY LTAILQSFSL QPLGAPEDID LTPLSSGLGN LPRPFQLCLRPR
cyp2f2   PLARMELFIY FTSILQNFTL QPLVDPEDID LTPLSSGLGN LPRPFQLCMHIR
```

**Fig. 1** *cont.*

## RESULTS AND DISCUSSION

Figure 2 shows how the substrate, 4-trifluoromethyl-7-ethoxycoumarin, is able to fit within the putative active site of CYP2B6 for O-deethylation. This orientation is achieved by a combination of π-π stacking interactions between the coumarin ring system and a phenylalanine residue within the haem environment, together with hydrogen bonding between the trifluoromethyl and carbonyl groups on the substrate and complementary amino acid residues in the haem pocket.

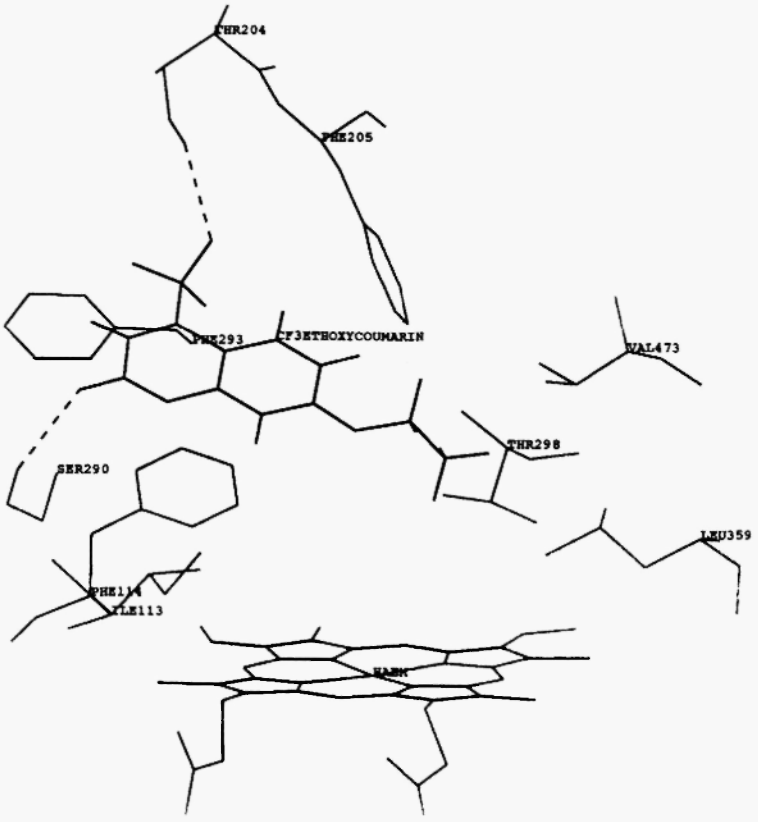


**Fig. 2:** An orientation of 4-trifluoromethyl-7-ethoxycoumarin within the putative active site of CYP2B6 showing the presence of hydrogen bonded interactions (dashed lines). Residues are numbered according to the alignment with CYP2C5 (see Figure 1 for details).

The distances between sites of metabolism and the haem iron atom in the CYP2B6-modelled substrate interactions are shown in Table 1. These values are consistent with typical distances observed for other P450 substrate-haem interactions in reported crystal structures (reviewed in /2/).

Substrates of CYP2B6 tend to be relatively lipophilic in character (average log P = 2.80 and possessing a range of log P values from 0.23 to 4.89) with the degree of lipophilicity being closely related to binding affinity (R = 0.98 for ten compounds investigated), as shown in Tables 2 and 3, respectively. CYP2B6 substrates tend to be either neutral or weakly basic in character with medium-sized molecules that are relatively non-planar in structure /13,17,18/. They also usually contain at least one hydrogen bond acceptor atom and at least one aromatic ring. The hydrogen bond acceptor atom should be an appropriate distance from the site of CYP2B6-mediated metabolism, however, in order that selectivity towards this enzyme is apparent. The relatively high affinity substrates, PNU 249173 and 4-trifluoromethyl-7-ethoxycoumarin, are outliers to the relationship between binding affinity and lipophilicity presented in Table 3. The reason for this is due to additional binding interactions (mainly hydrogen bonding) relative to the remaining CYP2B6 substrates. The relatively selective substrate, 4-trifluoromethyl-7-ethoxycoumarin, possesses a higher affinity for the enzyme than expected on the grounds of lipophilicity alone, and it is likely that this is due to the additional hydrogen bond-forming potential of the trifluoromethyl group. Molecular modelling of this substrate within the putative active site of CYP2B6 shows that two hydrogen bonds may be formed with nearby amino acid residues which serve to orientate the molecule for O-deethylation. Moreover, the fused aromatic ring system of coumarin provides a somewhat higher $\pi$-$\pi$ stacking interaction energy than for those substrates which only contain a single benzene ring. Another high affinity CYP2B6 substrate, 7-benzyloxyresorufin, also displays a high binding affinity for CYP2B6 (see Table 1) due to the presence of two aromatic ring systems and some additional hydrogen bond-forming potential; and Figure 3 shows how a molecular template of several CYP2B6 substrates (including 7-benzyloxyresorufin and 4-trifluoromethyl-7-ethoxycoumarin) can be superimposed within the enzyme's active site.

The location of hydrophobic groupings in the molecule is also of relevance to binding affinity and CYP2B6 selectivity, as has been

**TABLE 1**

Substrates of CYP2B6, metabolic data and distances between site of metabolism and haem iron

| Compound | Metabolic Pathway | $K_m$ (μM) | Reference | Distance (Å) |
|---|---|---|---|---|
| 1. 4-Trifluoromethyl-7-ethoxycoumarin | O-deethylation | 2.9 | Ekins and Wrighton /8/ | 3.906 |
| 2. 7-Benzyloxyresorufin | O-debenzylation | 1.28 | Ekins and Wrighton /8/ | 3.886 |
| 3. S-Mephenytoin | N-demethylation | 564 | Heyn *et al.* /24/ | 3.132 |
| 4. Benzphetamine | N-demethylation | 93.4 | Ekins and Wrighton /8/ | 4.634 |
| 5. Testosterone | 16□-hydroxylation | 50.5 | Ekins and Wrighton /8/ | 3.120 |
| 6. Methoxychlor | 3-hydroxylation | N/A | Dehal and Kupfer /25/ | 3.284 |
| 7. Bupropion | t-butyl oxidation | 107.5 | Ekins and Wrighton /8/ | 3.613 |
| 8. PNU 249173 (RP 73401)* | cyclopentyl hydroxylation | 19.6 | Stevens *et al.* /26/ | 4.876 |
| 9. Diazepam | N-demethylation | 113 | Ekins and Wrighton /8/; Yong *et al.* /27/ | N/D |
| 10. 7-Ethoxycoumarin | O-deethylation | 115 | Ekins and Wrighton /8/ | N/D |
| 11. 4-Chloromethyl-7-Ethoxycoumarin | O-deethylation | 33.7 | Ekins and Wrighton /8/ | N/D |
| 12. 3-Cyano-7-ethoxycoumarin | O-deethylation | 71.3 | Ekins and Wrighton /8/ | N/D |

* 3-Cyclopentyloxy-N-(3,5-dichloro-4-pyridyl)-4-methoxybenzamide.
N/A = data not available; N/D = not determined.
$K_m$ = Michaelis constant (μM) for substrate metabolism.

**TABLE 2**

Physicochemical properties of CYP2B6 substrates /22/

| Compound | log P | $pK_a$ | log $D_{7.4}$ | $M_r$ | $K_m$ (μM) | $\Delta G_{bind}$ (kcalmol$^{-1}$) |
|---|---|---|---|---|---|---|
| 1. 4-Trifluoromethyl EC | 3.31 | neutral | 3.31 | 242.21 | 2.9 | -7.849 |
| 2. 7-Benzyloxyresorufin | 4.75[c] | 10.22[b] | 1.98[c] | 303.33 | 1.28 | -8.3587 |
| 3. S-Mephenytoin | 1.74 | 8.1[b] | 0.96 | 218.28 | 564 | -4.6082 |
| 4. Benzphetamine | 2.27[c] | 6.6[b] | 2.21[c] | 239.35 | 93.4 | -5.7159 |
| 5. Testosterone | 3.32 | neutral | 3.32 | 288.43 | 50.5 | -6.0347 |
| 6. Methoxychlor | 4.30 | neutral | 4.30 | 345.65 | N.A. | N/A |
| 7. Bupropion | 2.54 | 8.35[b] | 1.54 | 239.77 | 107.5 | -5.6293 |
| 8. PNU 249173 | 4.89[c] | neutral | 4.89[c] | 381.28 | 22.5 | -6.5928 |
| 9. Diazepam | 2.86 | 3.7[b] | 2.82 | 284.80 | 113 | -5.5986 |
| 10. 7-Ethoxycoumarin (EC) | 2.30 | neutral | 2.30 | 190.21 | 115 | -5.5878 |
| 11. SM-12502 | 1.06 | 4.46[b] | 1.02 | 208.30 | 1767 | -3.9047 |
| 12. Antipyrine | 0.23 | neutral | 0.23 | 188.25 | 17700 | -2.4852 |

N/A = data not available; log P = logarithm of the octanol/water partition coefficient;
$pK_a$ = negative logarithm of the acid/base dissociation constant; log $D_{7.4}$ = logarithm of the distribution coefficient at pH 7.4;
$M_r$ = relative molecular mass; $K_m$ = Michaelis constant for substrate metabolism (μM); $\Delta G_{bind} = RT\ln K_m$; b = basic; c = calculated value;
SM-12502 = (+)-*cis*-3,5-dimethyl-2-(3-pyridyl)thiazolidin-4-one.
Substrates are either neutral or weakly basic with log P values lying within the range 1.74 to 4.89 (compounds 11 and 12 have been excluded on the grounds that they are poor substrates). The log P value for 4-trifluoromethyl-7-ethoxycoumarin lies close to the average for 10 compounds (log P average = 3.23). Basic substrates are thought to bind in the unionized form to CYP2B enzymes.

**TABLE 3**

Baseline lipophilicity relationship in CYP2B6 substrates

| Compound | $\Delta G_{part}$ | $\Delta G_{bind}$ |
|---|---|---|
| 1. 7-Benzyloxyresorufin | -6.7377 | -8.3587 |
| 2. S-Mephenytoin | -2.4681 | -4.6082 |
| 3. Benzphetamine | -3.2199 | -5.7159 |
| 4. Testosterone | -4.7093 | -6.0947 |
| 5. Bupropion | -3.6029 | -5.6293 |
| 6. Diazepam | -4.0568 | -5.5986 |
| 7. 7-Ethoxycourmarin (EC) | -3.0497 | -5.5878 |
| 8. SM-12502 | -1.5036 | -3.9047 |
| 9. Antipyrine | -0.3262 | -2.4852 |
| 10. 4-Chloromethyl EC | -4.1703 | -6.3439 |

SM-12502 = (+)-*cis*-3,5-dimethyl-2-(3-pyridyl)thiazolidin-4-one.
Correlation equation showing the baseline lipophilicity relationship for the ten compounds:
$\Delta G_{bind} = 0.861\ (\pm 0.065)\ \Delta G_{part} - 2.518$
n = 10; s = 0.3431; R = 0.9781; F = 176.7
n = number of compounds; s = standard error; R = correlation coefficient;
F = variance ratio
$\Delta G_{bind} = RT\ln K_m$ where $K_m$ is the Michaelis constant (μM);
$\Delta G_{part} = -RT\ln P$ where P is the octanol/water partition coefficient.
Note: Two substrates are outliers to this relationship, namely PNU 249173 and 4-trifluoro-methyl-7-ethoxycoumarin, due to the fact that they both appear to form additional hydrogen bonds to the enzyme. If an allowance is made for the hydrogen bond energy, then these two substrates lie on the line corresponding to the lipophilicity relationship described above.
Reference: /23/.

shown by Ekins and coworkers /19/. In fact, active site modelling of many CYP2B6 substrates shows that there are opportunities for hydrophobic interactions with various complementary residues lining the haem pocket, some of which have been probed using site-directed

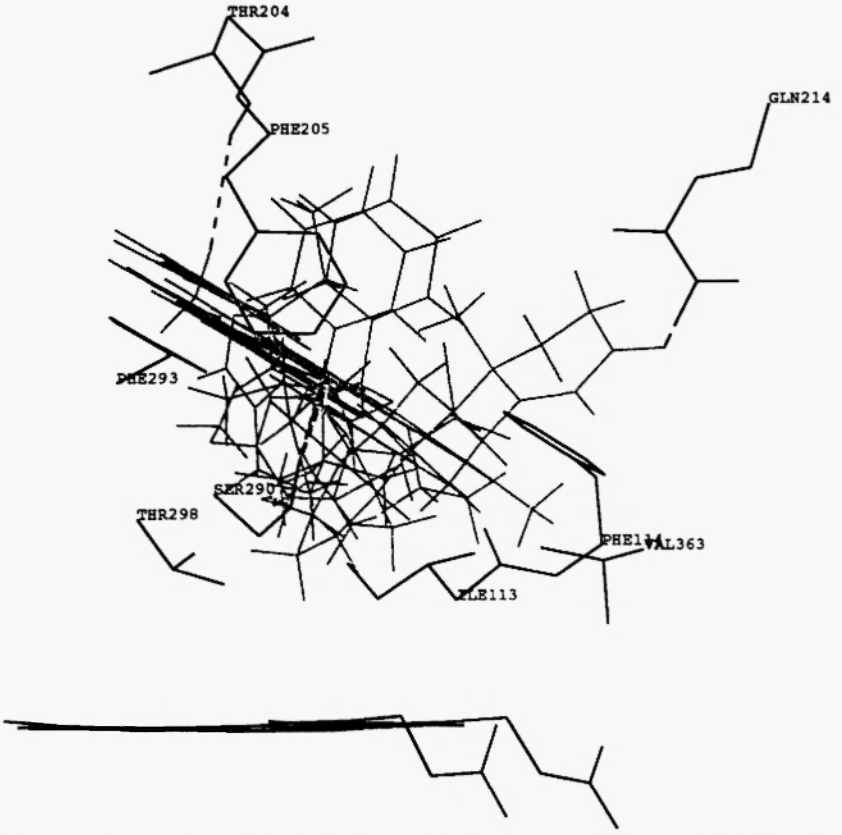


**Fig. 3:** A superimposed molecular template of CYP2B6 substrates fitted into the CYP2B6 active site showing a reinforcement of hydrogen-bonded interactions (dashed lines). Residues are numbered according to the alignment with CYP2C5 presented in Figure 1. A total of eight compounds are included and these correspond with substrates **1-8** listed in Table 1.

mutagenesis (reviewed in /10,16,15/). Table 4 provides a summary of site-directed mutagenesis information on CYP2B family enzymes which have been mapped within the substrate recognition site (SRS) regions /20/. Many of the residues identified via site-specific mutation in CYP2B subfamily enzymes correspond with those present in the putative active site of CYP2B6, and some of these represent substrate contact points in the current model, thus providing some degree of validation.

**TABLE 4**

Site-directed mutagenesis in the SRS* regions of CYP2B subfamily enzymes

| CYP | Change | 2C5 | SRS | Comments | References |
|---|---|---|---|---|---|
| 2B1 | V103A | I102 | 1 | in active site region | Kobayashi *et al.* /28/ |
| 2B4 | I114F | A113 | 1 | in active site | Szklarz *et al.* /29/ |
| 2B1 | I114A | A113 | 1 | in active site | Hasler *et al.* /30/ |
| 2B1 | F206L | V205 | 2 | in active site | He *et al.* /31/ |
| 2B1 | L209A | L208 | 2 | in active site | Szklarz *et al.* /32/ |
| 2B11 | D290I | I286 | 4 | near active site | Harlow and Halpert /33/; Harlow *et al.* /34/ |
| 2B4 | S294T | D290 | 4 | in active site | Szklarz *et al.* /29/ |
| 2B1 | T302S | T298 | 4 | conserved distal threonine | He *et al.* /35/ |
| 2B4 | T302A | T298 | 4 | conserved distal threonine | Vaz *et al.* /36/ |
| 2B1 | V363A | L359 | 5 | in active site | He *et al.* /31/ |
| 2B11 | I265V | T361 | 5 | in active site | Born *et al.* /37/ |
| 2B1 | V367A | L363 | 5 | in active site | He *et al.* /35/ |
| 2B1 | I477A | F473 | 6 | in active site | Szklarz *et al.* /32/ |
| 2B1 | G478A | V474 | 6 | in active site | Halpert and He /38/ |
| 2B1 | I480A | V476 | 6 | in active site | Szklarz *et al.* 1995 /32/ |

* Mutations in SPS3 were not found to be effective in altering catalytic activity and, consequently, there is some doubt regarding its role in substrate binding /32/.

General references to site-directed mutagenesis investigations in the CYP2 family include /2,15,39,40/.

From the results of QSAR analyses /21,22/ previously conducted on substrates of CYP2B6, on substrates and inhibitors of CYP2B4, and on CYP2B1 substrates, it would appear that either log P or the related quantity, molar volume, is involved in the correlations with binding affinity, inhibition and metabolic clearance. For CYP2B6 substrates, a good correlation was found between binding affinity and the number of active site hydrogen bonds ($N_{hb}$) together with the number of hydrogen bond donor/acceptor atoms ($HB_D/HB_A$) in the molecule. The relevant relationship is shown as follows:

$$\Delta G_{bind} = \underset{(\pm 0.758)}{3.985 N_{hb}^{site}} - \underset{(\pm 0.858)}{5.409 HB_D} - \underset{(\pm 0.226)}{1.922 HB_A} - 4.194$$

$$n = 10;\ s = 0.5524;\ R = 0.965;\ F = 26.96$$

where n is the number of compounds, s is the standard error, R is the correlation coefficient and F is the variance ratio. However, for CYP2B6 substrates, there is also a strong correlation between $\Delta G_{bind}$, the binding affinity obtained from $K_m$ values, and partitioning energy calculated for log P values, where P is the octanol-water partition coefficient, as presented in Table 3. Furthermore, similar expressions involving log P also exist for substrates and inhibitors of CYP2B4 and CYP2B1 /21,22/.

These findings tend to suggest that compound lipophilicity is of importance to substrate binding in the CYP2B subfamily, and it is interesting that there is a strong linear relationship between log P and $K_m$ for a representative number of CYP2B6 substrates (see Table 3) when these values are expressed as free energies. Such a lipophilicity relationship can be regarded, therefore, as a baseline for comparing the binding properties of various CYP2B6 substrates, and other expressions of a similar nature have been observed for other P450s and their substrates /25/. It is likely, therefore, that $\Delta G_{part}$, the partitioning energy, makes a significant contribution to overall binding affinity, and Table 5 indicates that this may indeed be the case. Using an additive expression for binding energy based on contributions from lipophilicity, hydrogen bonding and $\pi$-$\pi$ stacking, together with loss in rotatable bond energy, it is possible to obtain very reasonable

**TABLE 5**

Comparison between calculated and experimental binding affinities

| Compound | $\Delta G_{part}$ | $\Delta G_{hb}$ | $\Delta G_{\pi-\pi}$ | $\Delta G_{rot}$ | $\Delta G^{calc}_{bind}$ | $\Delta G^{expt}_{bind}$ |
|---|---|---|---|---|---|---|
| 1. 4-Trifluoromethyl EC | -4.6951 | -4 | -1.8 | 2.4 | -8.0951 | -7.8549 |
| 2. 7-Benzyloxyresorufin | -6.7377 | -2 | -1.8 | 1.8 | -8.7377 | -8.3587 |
| 3. S-Mephenytoin | -2.4681 | -2 | -1.8 | 1.2 | -4.1681 | -4.6082 |
| 4. Benzphetamine | -3.2199 | -2 | -1.8 | 1.8 | -5.2199 | -5.7159 |
| 5. Testosterone | -4.7093 | -2 | 0 | 0.6 | -6.1093 | -6.0947 |
| 6. 4-Chloromethyl EC | -4.1703 | -2 | -1.8 | 2.4 | -5.5703 | -6.3439 |
| 7. Bupropion | -3.6029 | -2 | -0.9 | 1.2 | -5.3029 | -5.6293 |
| 8. PNU 249173 | -1.5036 | -2 | -1.8 | 3.6 | -7.1363 | -6.5928 |
| 9. Diazepam | -4.0568 | -2 | -0.9 | 1.2 | -5.7568 | -5.5986 |
| 10. 7-Ethoxycoumarin (EC) | -3.0497 | -2 | -1.8 | 1.8 | -5.0497 | -5.5878 |
| 11. SM-12502 | -1.5036 | -2 | -0.9 | 0.6 | -3.8036 | 3.9047 |
| 12. Antipyrine | -0.3262 | -2 | -0.9 | 0.6 | -2.6262 | -2.4852 |

$\Delta G_{part} = -RT\ln P$ where P is the octanol/water partition coefficient; $\Delta G_{hb}$ = hydrogen bond energy (-2 kcalmol$^{-1}$ per hydrogen bond); $\Delta G_{\pi-\pi}$ = $\pi$-$\pi$ stacking energy (-0.9 kcalmol$^{-1}$ per $\pi$-$\pi$ stack); $\Delta G_{rot}$ = loss in bond rotational energy (+0.6 kcalmol$^{-1}$ per rotational bond frozen on binding); $\Delta G^{calc}_{bind} = \Delta G_{part} + \Delta G_{hb} + \Delta G_{\pi-\pi} + \Delta G_{rot}$; $\Delta G^{expt}_{bind} = RT\ln K_m$ where $K_m$ is the Michaelis constant ($\mu$M).

Correlation between experimental and calculated values for substrate binding affinity gives the following statistics:

$\Delta G^{expt}_{bind} = 1.008\ \Delta G^{calc}_{bind}$; n = 10; s = 0.3448; R = 0.9819; F = 215.0

estimates (R = 0.98) of binding affinity for CYP2B6 substrates, and this is presented in Table 5. Consequently, the homology model can be used to generate contributions to the overall enzyme-substrate binding affinity which agree closely with experimental findings.

## CONCLUSIONS

Homology modelling of CYP2B6 based on the CY2C5 crystal structure provides a good model of the enzyme for the investigation of substrate binding interactions. The results of these may be employed for the estimation of binding affinity and the findings show good agreement with experimental data.

## ACKNOWLEDGEMENTS

The financial support of GlaxoSmithKline Research and Development Ltd., Merck Sharp & Dohme Ltd. and the University of Surrey Foundation Fund is gratefully acknowledged by DFVL.

## REFERENCES

1. Lewis DFV. Cytochromes P450: Structure, Function and Mechanism. London: Taylor & Francis, 1996.
2. Lewis DFV. Guide to Cytochrome P450 Structure and Function. London: Taylor & Francis, 2001.
3. Guengerich FP. Cytochrome P450. In: Ioannides C, ed. Enzyme Systems that Metabolize Drugs and other Xenobiotics. New York: Wiley, 2002; 33-65.
4. Rendic S, DiCarlo FJ. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers and inhibitors. Drug Metab Rev1997; 29: 413-580.
5. Ingelman-Sundberg M. Implications of polymorphic cytochrome P450-dependent drug metabolism for drug development. Drug Metab Dispos 2001; 29: 570-573.
6. Ingelman-Sundberg M, Oscarson M, McLellan RA. Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. Trends Pharm Scie 1999; 20: 342-349.
7. Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapies. Science 1999; 286: 487-491.
8. Ekins S, Wrighton SA. The role of CYP2B6 in human xenobiotic metabolism. Drug Metab Rev 1999; 31: 719-754.

9. Rae JM, Soukhova NV, Flockhart DA, Desta Z. Triethyl-enethiophosphoramide is a specific inhibitor of cytochrome P450 2B6: implications for cyclophosphamide metabolism. Drug Metab Dispos 2002; 30: 525-530.
10. Wang Q, Halpert JR. Combined three-dimensional quantitative structure-activity relationship analysis of cytochrome P450 2B6 substrates and protein homology modeling. Drug Metab Dispos 2002; 30: 86-95.
11. Williams PA, Cosme J, Sridhar V, Johnson EF, McRee DE. Mammalian cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. Mol Cell 2000; 5: 121-131.
12. Lewis DFV. Homology modelling of human CYP2 family enzymes based on the CYP2C5 crystal structure. Xenobiotica 2002; 32: 305-323.
13. Lewis DFV, Lake BF, Dickins M, Eddershaw PJ, Tarbit MH, Goldfarb PS. Molecular modelling of CYP2B6, the human CYP2B isoform, by homology with the substrate-bound CYP102 crystal structure: evaluation of CYP2B6, CYP2B1 and CYP2B4 substrate characteristics. Xenobiotica 1999; 29: 361-393.
14. Rendic S. Summary of information on human CYP enzymes: human P450 metabolism data. Drug Metab Rev 2002; 34: 83-448.
15. Lewis DFV. The CYP2 family: models, mutants and interactions. Xenobiotica 1998; 28: 617-661.
16. Domanski TL, Halpert JR. Analysis of mammalian cytochrome P450 structure and function by site-directed mutagenesis. Curr Drug Metab 2001; 2: 117-137.
17. Lewis DFV. On the recognition of mammalian microsomal cytochrome P450 substrates and their characteristics. Biochem Pharmacol 2000; 60: 293-306.
18. Lewis DFV. Structural characteristics of human P450s involved in drug metabolism: QSARs and lipophilicity profiles. Toxicology 2000; 144: 197-203.
19. Ekins S, Bravi G, Ring BJ, Gillespie TA, Gillespie JS, VandenBranden M, Wrighton SA, Wikel JH. Three-dimensional quantitative structure activity relationship analyses of substrates for CYP2B6. J Pharmacol Exp Ther 1999; 288: 21-29.
20. Gotoh O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. J Biol Chem 1992; 267: 83-90.
21. Lewis DFV, Eddershaw PJ, Dickins M, Tarbit MH, Goldfarb PS. Structural determinants of P450 substrate specificity, binding affinity and catalytic rate. Chem-Biol Interact 1998; 115: 175-199.
22. Lewis DFV, Modi S, Dickins M. Structure-activity relationships for human cytochrome P450 substrates and inhibitors. Drug Metab Rev 2002; 34: 69-82.
23. Lewis DFV. Baseline lipophilicity relationships in human cytochromes P450 associated with drug metabolism. Drug Metab Rev 2003; 35: 1-18.
24. Heyn H, White RC, Stevens JC. Catalytic role of cytochrome P4502B6 in the N-demethylation of S-mephenytoin. Drug Metab Dispos 1996; 24: 948-954.
25. Dehal SS, Kupfer D. Metabolism of the proestrogenic pesticide methoxychlor by hepatic P450 monooxygenases in rats and humans. Drug Metab Dispos 1994; 22: 937-946.

26. Stevens J, Hauer M, Beck J, Voorman R. Further characterization of human liver CYP2B6 expression and catalytic function. Drug Metab Rev 2000; 32: 218.
27. Yong TJ, Shou M, Korzekwa KR, Gonzalez FJ, Gelboin HV, Yang SK. Role of cDNA-expressed human cytochromes P450 in the metabolism of diazepam. Biochem Pharmacol 1998; 55: 889-896.
28. Kobayashi Y, Strobel SM, Hopkins NE, Alworth WL, Halpert JR. Catalytic properties of an expressed cytochrome P450 2B1 from a Wistar-Kyoto rat liver cDNA library. Drug Metab Dispos 1998; 26: 1026-1030.
29. Szklarz GD, He YQ, Kedzie KM, Halpert JR, Burnett VL. Elucidation of amino acid residues critical for unique activities of rabbit P4502B5 using hybrid enzymes and reciprocal site-directed mutagenesis with rabbit cytochrome P450 2B4. Arch Biochem Biophys 1996; 327: 308-318.
30. Hasler JA, Harlow GR, Szklarz GD, John GH, Kedzie KM, Burnett VL, He YA, Kaminsky LS, Halpert JR. Site-directed mutagenesis of putative substrate recognition sites in cytochrome P450 2B11: importance of amino acid residues 114, 290 and 363 for substrate specificity. Mol Pharmacol 1994; 46: 338-345.
31. He Y, Luo Z, Klekotka PA, Burnett VL, Halpert JR. Structural determinants of cytochrome P450 2B1 specificity: evidence of five substrate recognition sites. Biochemistry 1994; 33: 4419-4424.
32. Szklarz GD, He YA, Halpert JR. Site-directed mutagenesis as a tool for molecular modelling of cytochrome P4502B1. Biochemistry 1995; 34: 14312-14322.
33. Harlow GR, Halpert JR. Mutagenesis study of Asp-290 in cytochrome P4502B11 using a fusion protein with rat NADPH-cytochrome P450 reductase. Arch Biochem Biophys 1996; 326: 85-92.
34. Harlow GR, He YA, Halpert JR. Functional interaction between amino acid residues 242 and 290 in cytochromes P450 2B1 and 2B11. Biochim Biophys Acta 1997; 1338: 259-266.
35. He YQ, He YA, Halpert JR. *Escherichia coli* expression of site-directed mutants of cytochrome P450 2B1 from 6 substrate recognition sites: substrate specificity and inhibitor selectivity studies. Chem Res Toxicol 1995; 8: 574-579.
36. Vaz ADN, Pernecky SJ, Raner GM, Coon MJ. Peroxo-iron and oxenoid-iron species as alternative oxygenating agents in cytochrome P450-catalyzed reactions: switching by threonine-302 to alanine mutagenesis of cytochrome P450 2B4. Proc Natl Acad Sci USA 1996; 93: 4644-4648.
37. Born SL, John GH, Harlow GR, Halpert JR. Characterization of the progesterone 21-hydroxylase activity of canine cytochrome P450 PBD-2/P450 2B11 through reconstitution, heterologous expression and site-directed mutagenesis. Drug Metab Dispos 1995; 23: 702-707.
38. Halpert JR, He Y. Engineering of cytochrome P450 2B1 specificity. J Biol Chem 1993; 268: 4453-4457.
39. Halpert JR. Structural basis of selective cytochrome P450 inhibitor. Annu Rev Pharmacol Toxicol 1995; 35: 29-53.

40. Waller SC, He YA, Harlow GR, He YQ, Mash EA, Halpert JR. 2,2′,3,3′,6,6′-Hexachlorobiphenyl hydroxylation by active site mutants of cytochrome P450 2B1 and 2B11. Chem Res Toxicol 1999; 12: 690-699.
41. Domanski TL, He Y-Q, Scott EE, Wang Q, Halpert JR. The role of cytochrome P450 2B1 substrate recognition site residues 115, 294, 297, 298 and 362 in the oxidation of steroids and 7-alkoxycoumarins. Arch Biochem Biophys 2001; 394: 21-28.