Supporting information

Biotransformation, using Recombinant CYP450-expressing Baker's Yeast Cells, Identifies a Novel CYP2D6.10^{A122V} Variant Which is a Superior Metaboliser of Codeine to Morphine Than the Wild-type Enzyme

Ibidapo S. Williams,[†] Linda Gatchie,[†] Sandip B. Bharate[‡] and Bhabatosh Chaudhuri[†]*

[†]*CYP Design Ltd, The Innovation Centre, 49 Oxford Street, Leicester, LE1 5XY, UK*

[‡]*Medicinal Chemistry Division, CSIR-Indian Institute of Integrative Medicine, Canal Road, Jammu, India*

CONTENTS

Section S1. Comparisons of the amino acid sequences of the four CYP2D6 proteins

```
                        10        20        30        40        50        60
                         *         *         *         *         *         *
2D6-wt          MGLEALVPLAVIVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTPYCFDQ
2D6.1           MGLEALVPLAVIVAIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTPYCFDQ
2D6.10          MGLEALVPLAVIVAIFLLLVDLMHRRQRWAARYSPGPLPLPGLGNLLHVDFQNTPYCFDQ

                        70        80        90       100       110       120
                         *         *         *         *         *         *
2D6-wt          LRRRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQGVF
2D6.1           LRRRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQGVF
2D6.10          LRRRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQGVF

                       130       140       150       160       170       180
                         *         *         *         *         *         *
2D6-wt          LARYGPAWREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDK
2D6.1           LARYGPAWREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDK
2D6.10          LARYGPAWREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDK

                       190       200       210       220       230       240
                         *         *         *         *         *         *
2D6-wt          AVSNVIASLTCGRRFEYDDPRFLRLLDLAQEGLKEESGFLREVLNAVPVLLHIPALAGKV
2D6.1           AVSNVIASLTCGRRFEYDDPRFLRLLDLAQEGLKEESGFLREVLNAVPVLLHIPALAGKV
2D6.10          AVSNVIASLTCGRRFEYDDPRFLRLLDLAQEGLKEESGFLREVLNAVPVLLHIPALAGKV

                       250       260       270       280       290       300
                         *         *         *         *         *         *
2D6-wt          LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNDENLRIVVA
2D6.1           LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNDENLRIVVA
2D6.10          LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNDENLRIVVA

                       310       320       330       340       350       360
                         *         *         *         *         *         *
2D6-wt          DLFSAGMVTTSTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI
2D6.1           DLFSAGMVTTSTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI
2D6.10          DLFSAGMVTTSTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI

                       370       380       390       400       410       420
                         *         *         *         *         *         *
2D6-wt          HEVQRFGDIVPLGMTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRFHPEHF
2D6.1           HEVQRFGDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRFHPEHF
2D6.10          HEVQRFGDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRFHPEHF

                       430       440       450       460       470       480
                         *         *         *         *         *         *
2D6-wt          LDAQGHFVKPEAFLPFSAGRRACLGEPLARMELFLFFTSLLQHFSFSVPTGQPRPSHHGV
2D6.1           LDAQGHFVKPEAFLPFSAGRRACLGEPLARMELFLFFTSLLQHFSFSVPTGQPRPSHHGV
2D6.10          LDAQGHFVKPEAFLPFSAGRRACLGEPLARMELFLFFTSLLQHFSFSVPTGQPRPSHHGV

                       490
                         *
2D6-wt          FAFLVSPSPYELCAVPR
2D6.1           FAFLVSPSPYELCAVPR
2D6.10          FAFLVTPSPYELCAVPR
```

Figure S1A. Aligned protein sequences of CYP2D6-wt (2D6-wt; NCBI Accession No M20403 for DNA sequence; NCBI Accession No AAA52153 for debrisoquine 4-hydroxylase_Homo sapiens protein sequence;[10] CYP2D6.1

(2D6.1; NCBI Accession No NM_000106 for DNA sequence; NCBI Accession No NP_000097 for protein sequence (the NCBI database describes this as "cytochrome P450 2D6 isoform 1 [Homo sapiens]", i.e. CYP2D6.1;[29] and CYP2D6.10 (2D6.10; NCBI Accession No ABB01372), with dissimilar amino acid residues highlighted in red.

```
                    10        20        30        40        50        60
                     *         *         *         *         *         *
2D6.10          MGLEALVPLAVIVAIFLLLVDLMHRRQRWAARYSPGPLPLPGLGNLLHVDFQNTPYCFDQ
2D6-C           MGLEALVPLAVIVAIFLLLVDLMHRRQRWAARYSPGPLPLPGLGNLLHVDFQNTPYCFDQ

                    70        80        90       100       110       120
                     *         *         *         *         *         *
2D6.10          LRRRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQGVF
2D6-C           LRRRFGDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQGVF

                   130       140       150       160       170       180
                     *         *         *         *         *         *
2D6.10          LARYGPAWREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDK
2D6-C           LVRYGPAWREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDK

                   190       200       210       220       230       240
                     *         *         *         *         *         *
2D6.10          AVSNVIASLTCGRRFEYDDPRFLRLLDLAQEGLKEESGFLREVLNAVPVLLHIPALAGKV
2D6-C           AVSNVIASLTCGRRFEYDDPRFLRLLDLAQEGLKEESGFLREVLNAVPVLLHIPALAGKV

                   250       260       270       280       290       300
                     *         *         *         *         *         *
2D6.10          LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNDENLRIVVA
2D6-C           LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNDENLRIVVA

                   310       320       330       340       350       360
                     *         *         *         *         *         *
2D6.10          DLFSAGMVTTSTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI
2D6-C           DLFSAGMVTTSTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI

                   370       380       390       400       410       420
                     *         *         *         *         *         *
2D6.10          HEVQRFGDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRFHPEHF
2D6-C           HEVQRFGDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRFHPEHF

                   430       440       450       460       470       480
                     *         *         *         *         *         *
2D6.10          LDAQGHFVKPEAFLPFSAGRRACLGEPLARMELFLFFTSLLQHFSFSVPTGQPRPSHHGV
2D6-C           LDAQGHFVKPEAFLPFSAGRRACLGEPLARMELFLFFTSLLQHFSFSVPTGQPRPSHHGV

                   490
                     *
2D6.10          FAFLVTPSPYELCAVPR
2D6-C           FAFLVTPSPYELCAVPR
```

Figure S1B. Aligned protein sequences of CYPD6.10 (2D6.10; NCBI Accession No ABB01372) and CYP2D6-C (2D6-C), with the dissimilar amino acid residue, V122 in CYP2D6-C, highlighted in red.

Section S2. A basic plasmid, YIpADH2S, used for cloning of the four human CYP2D6 alleles

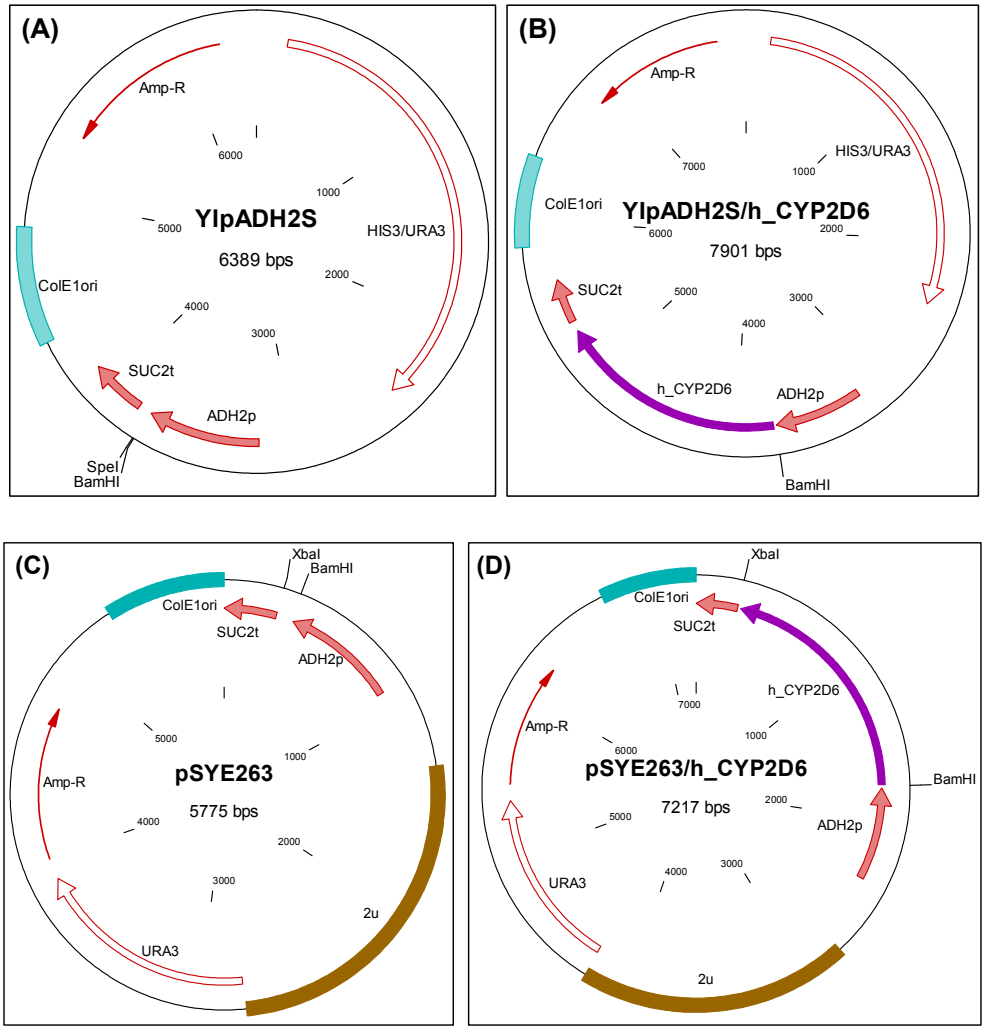


Figure S2. (A) Represents a basic plasmid, YIpADH2S, used for cloning of the four human CYP2D6 alleles that would allow integration of a CYP2D6 gene expression cassette into the yeast genome of the strain YY7 at the *HIS3* or *URA3* chromosomal locus. It contains the ethanol-inducible yeast ADH2 promoter (ADH2p) and the SUC2 transcription termination signal (SUC2t). The human CYP2D6 alleles were cloned as *BamHI-Xba*I fragments at the *BamH*I, *Spe*I sites of an integration plasmid that bears either the *HIS3* or *URA3* genes as an auxotrophic marker. (B) Shows an integration plasmid after cloning of a CYP2D6 allele downstream of the ADH2p in the plasmid, YIpADH2S. (C) Shows a basic episomal, 2µ-plasmid, pSYE263, that bears a functional *URA3* gene. The human CYP2D6 alleles were cloned as *BamHI-Xba*I fragments at the *BamH*I, *Xba*I sites of pSYE263. (D) Shows an episomal plasmid after cloning of a CYP2D6 allele downstream of the ADH2p in the plasmid, pSYE263. The term 'h_CYP2D6' represents any one of the four CYP2D6 alleles chosen for this study.

Section S3. Comparison of CYP2D6-C enzyme activities expressed from an integrated copy and an episomal plasmid, using EOMCC as a substrate. Figure S1 shows a comparison of CYP2D6 enzyme activities obtained from yeast strains containing expression cassettes for the CYP2D6-C allele integrated at the *HIS3* chromosomal locus and borne on an episomal plasmid.

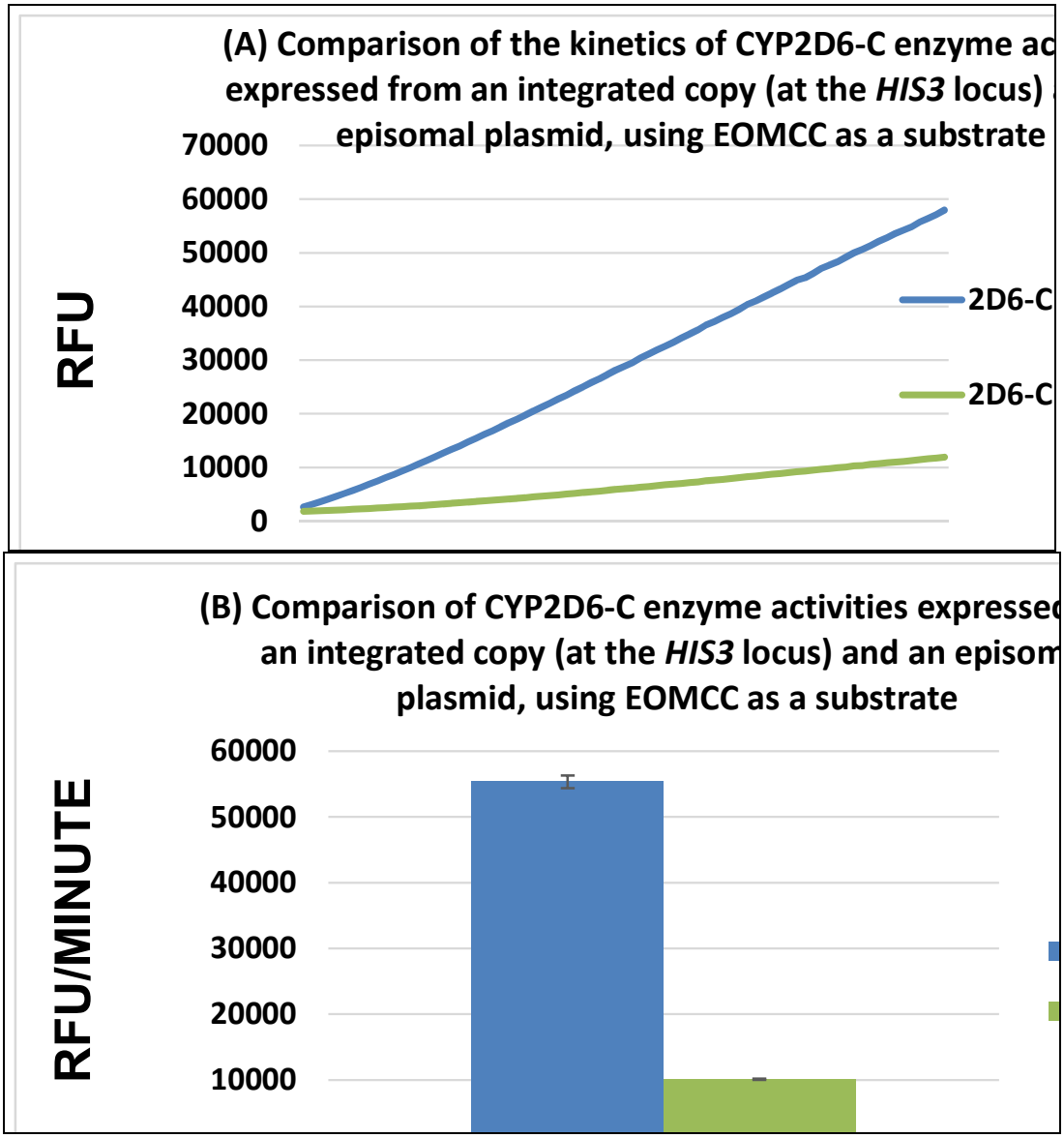


Figure S3. The graph (A) compares the kinetics of enzyme activities of the CYP2D6-C (CYP2D6.10^{A122V}; 2D6-C) allele produced in the strain YY7 from gene expression cassettes either (i) integrated at the *HIS3* chromosomal locus or (ii) borne on an episomal plasmid (2μ). The kinetics of enzyme activity, present in ~1x 10^6 cells, was followed over a time course of 76 min. The concentration of the fluorogenic substrate, EOMCC, used for each assay was 2 μM. The amount of fluorescent product, *7-HCC*, formed was monitored at each time point using a fluorescent plate reader. The graphs represent the average of results obtained from three independent experiments. The bar plot (B) mirrors the fluorescence values in the graphs in (A), at time point 74 min. The data represent mean ± S.D. of three independent experiments. 'RFU' represents relative fluorescence units.

Section S4. Comparison of CYP2D6.10 enzyme activities expressed from an integrated copy and an episomal plasmid, using EOMCC as a substrate. Figure S2 shows a comparison of CYP2D6 enzyme activities obtained from yeast strains containing expression cassettes for the CYP2D6.10 allele integrated at the *HIS3* chromosomal locus and borne on an episomal plasmid.

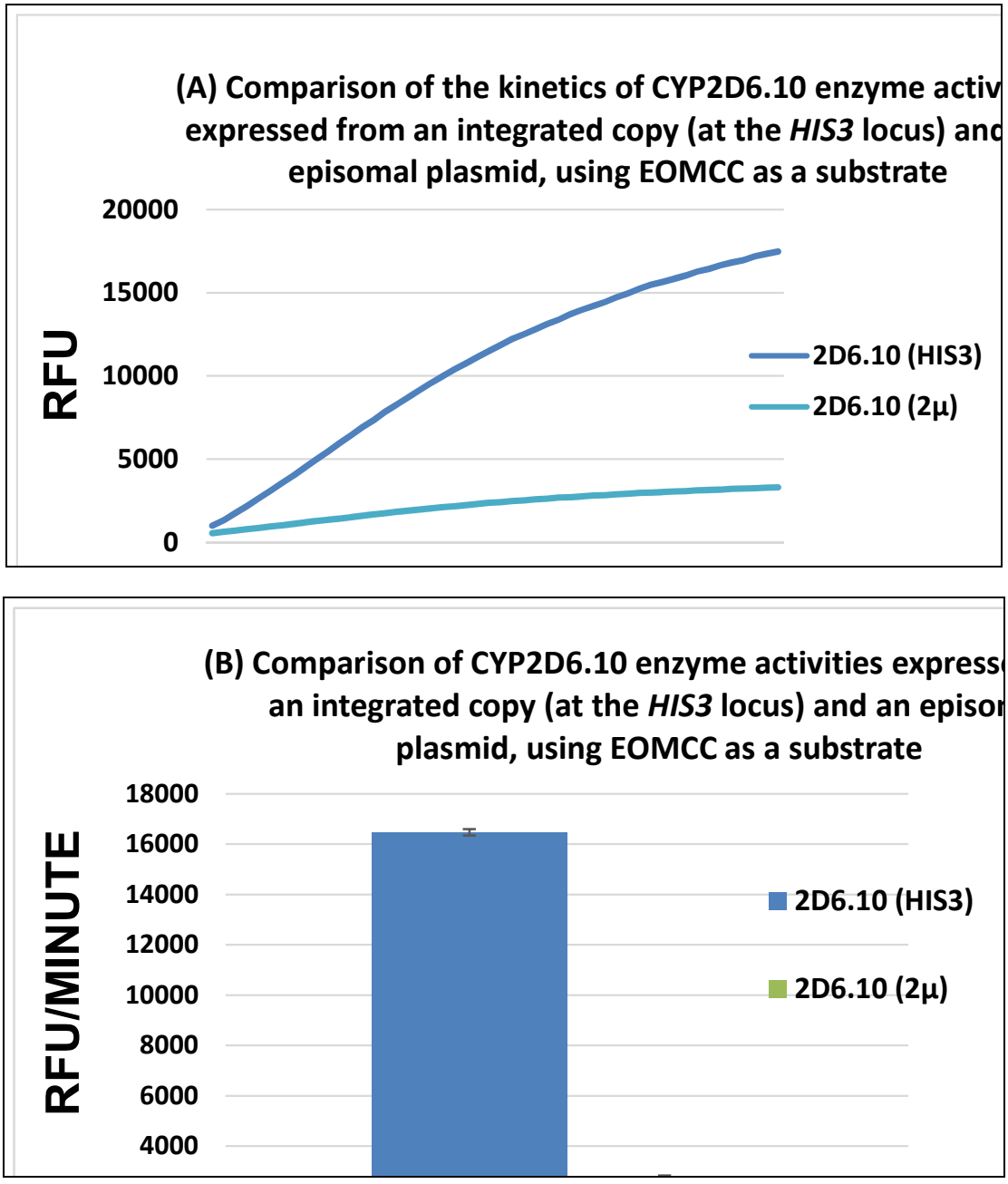


Figure S4. The graph (A) compares the kinetics of enzyme activities of the CYP2D6.10 (2D6.10) allele produced in the strain YY7 from gene expression cassettes either (i) integrated at the *HIS3* chromosomal locus or (ii) borne on an episomal plasmid (2µ). The kinetics of enzyme activity, present in ~1x 10^6 cells, was followed over a time course of 49 min. The concentration of the fluorogenic substrate, EOMCC, used for each assay was 2 µM. The amount of fluorescent product, *7-HCC*, formed was monitored at each time point using a fluorescent plate reader. The graphs represent the average of results obtained from three independent experiments. The bar plot (B) mirrors the fluorescence values in the graphs in (A), at time point 48 min. The data represent mean ± S.D. of three independent experiments. 'RFU' represents relative fluorescence units.

Section S5. TLC images of biotransformation experiment

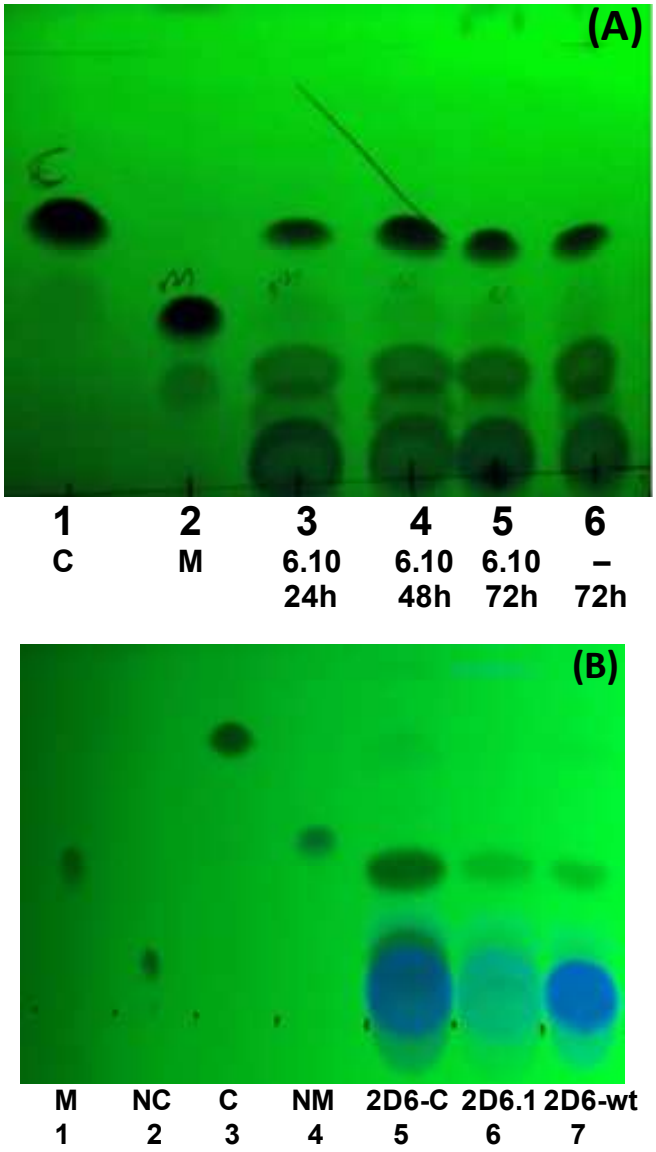


Figure S5. TLC analysis of biotransformation reactions of codeine using yeast-expressed CYP2D6 enzymes. (A) Biotransformation of codeine using cells expressing CYP2D6.10 (6.10). Lane 1, codeine (standard); lane 2, morphine (standard); lane 3, codeine + CYP2D6.10 cells incubated for 24 h; lane 4; codeine + CYP2D6.10 cells incubated for 48 h; lane 5, codeine + CYP2D6.10 cells incubated for 72 h; lane 6, codeine + empty plasmid-containing cells incubated for 72 h. (B) Lane 1, morphine (standard); lane 2, norcodeine (standard); lane 3, codeine (standard); lane 4, normorphine (standard); lane 5, codeine + CYP2D6-C (CYP2D6.10^{A122V}) cells incubated for 72 h; lane 6, codeine + CYP2D6.1 (CYP2D7^{V374}) cells incubated for 72 h; lane 7, codeine + CYP2D6-wt (CYP2D7^{M374}) cells incubated for 72 h. Solvent system used for both TLCs was $CHCl_3$:MeOH:NH_3 = 36:1:0.6. Abbreviations: M, morphine; NC, norcodeine; C, codeine; NM, normorphine.

Section S6. Molecular modeling of thioridazine with CYP2D6 (4WNW)

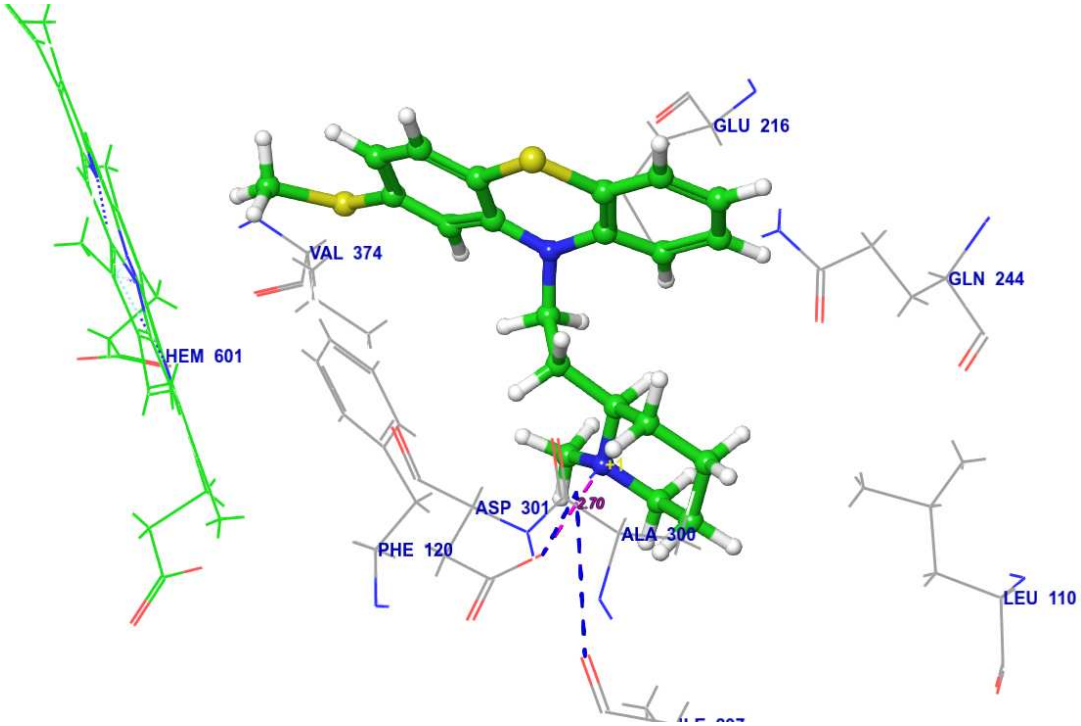


Figure S6a. The interaction pattern of docked thioridazine with CYP2D6 (4WNW)

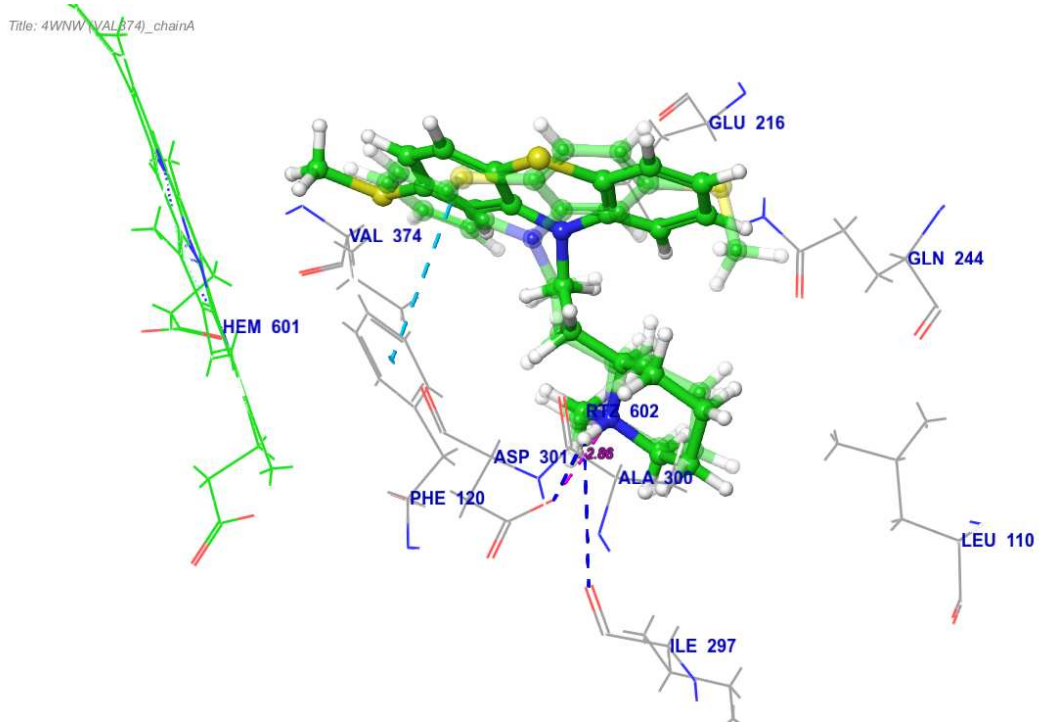


Figure S6b. The interaction pattern of thioridazine ligand from co-crystallized protein (4WNW)

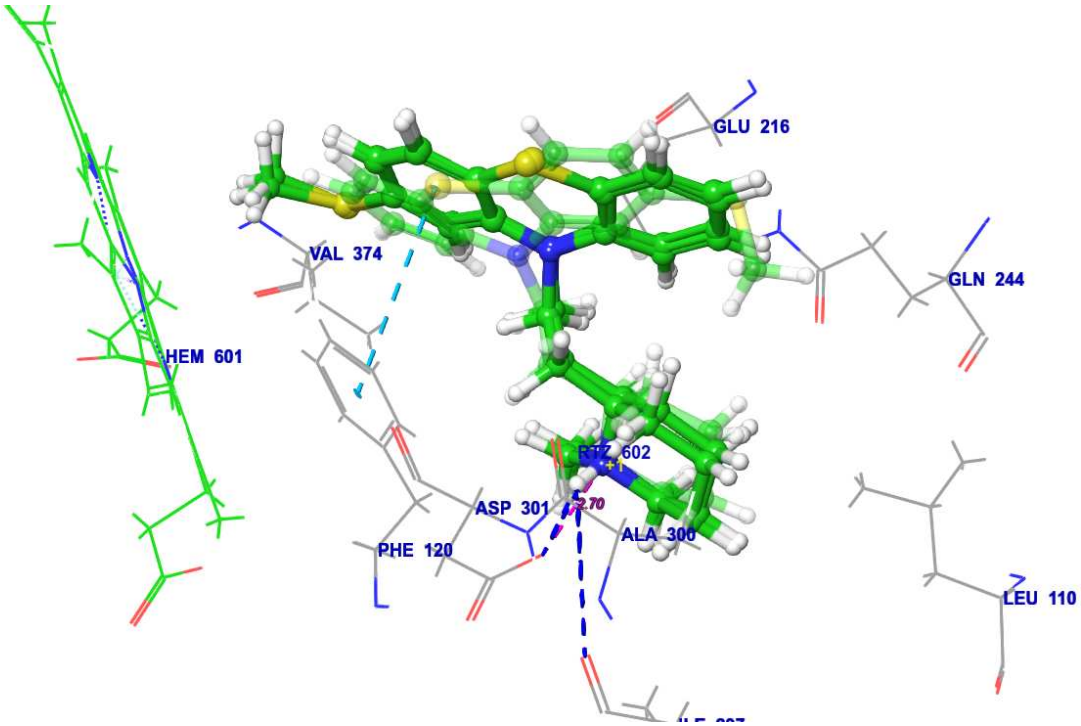
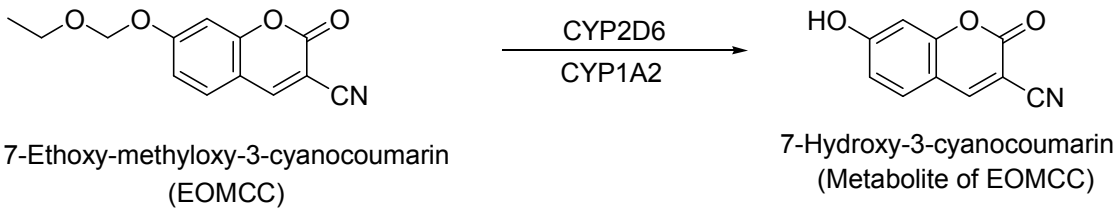


Figure S6c. The overlay image of ligand from co-crystallized protein and the docked thioridazine with CYP2D6 (4WNW)

Section S7. Molecular modeling of EOMCC substrate with CYP2D6 (4WNW) wild-type and 2D6-C



7-Ethoxy-methyloxy-3-cyanocoumarin (EOMCC)

7-Hydroxy-3-cyanocoumarin (Metabolite of EOMCC)

Figure S7a. The conversion of EOMCC to its hydroxy metabolite via CYP2D6/ CYP1A2

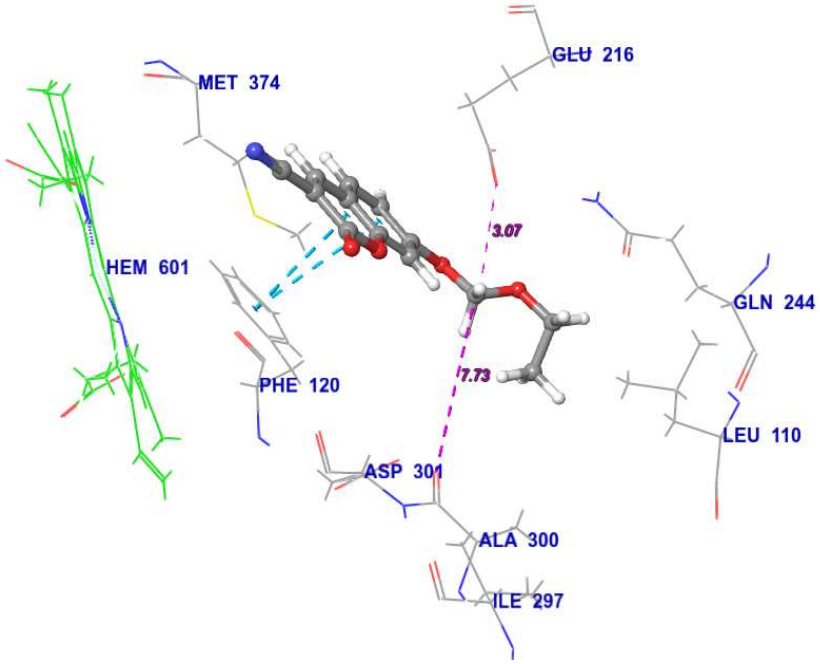


Figure S7b. The interaction pattern of EOMCC ligand with CYP2D6-wild (Met 374).

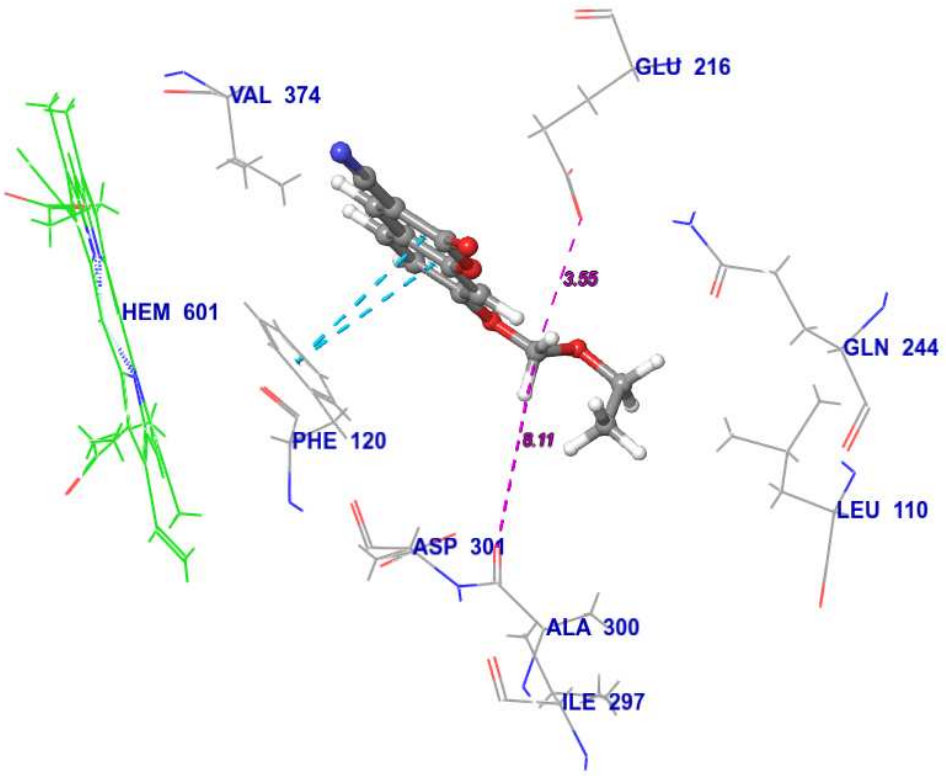


Figure S7c. The interaction pattern of EOMCC ligand with CYP2D-C.