Functional interpretation and structural insights of *Arabidopsis lyrata* cytochrome P450 CYP71A13 involved in auxin synthesis

Vijaya Kumari, Vijay Kumar & Tek Chand Bhalla*

Department of Biotechnology, Himachal Pradesh University, Shimla-171005; Tek Chand Bhalla – Email: bhallatc@rediffmail.com; *Corresponding author

Abstract:
Cytochrome P450 CYP71A13 of Arabidopsis lyrata is a heme protein involved in biosynthesis of indole-3-acetonitrile which leads to the formation of indolyl-3-acetic acid. It catalyzes a unique reaction: formation of a carbon-nitrogen triple bond and dehydration of indolyl-3-acetaldoxime. Homology model of this 57 kDa polypeptide revealed that the heme existed between H-helix and J- helix in the hydrophobic pocket, although both helixes are involved in catalytic activity, where Gly305 and Thr308, 311 of H- helix were involved in its stabilization. The substrate indole-3-acetaldoxime was tightly fitted into the substrate pocket with the aromatic ring being surrounded by amino acid residues creating a hydrophobic environment. The smaller size of the substrate binding pocket in cytochrome P450 CYP71A13 was due to the bulkiness of the two amino acid residues Phe182 and Trp315 pointing into the substrate binding cavity. The apparent role of the heme in cytochrome P450 CYP71A13 was to tether the substrate in the catalysis by indole-3-acetaldoxime dehydratase. Since the crystal structure of cytochrome P450 CYP71A13 has not yet been solved, the modeled structure revealed mechanism of substrate recognition and catalysis.

Keywords: Cytochrome P450 CYP71A13, indolyl-3-acetic acid and homology model.

Background:
Indole-3-acetic acid (IAA) is the main endogenous auxin in higher plants which plays role in plant growth including initiation and emergence of lateral roots, patterning of the root apical meristem, gravitropism, root elongation and development [1-4]. Plants and some plant pathogens can produce IAA to modulate plant growth. Plants can produce IAA from both Trp-dependent and Trp-independent pathways [5-6]. In Trp-dependent pathway, cytochromes CYP79B2 and CYP79B3 of Arabidopsis have been shown to metabolize tryptophan to indole-3-acetaldoxime **(Figure 1)** which is a precursor for the production of indole-3-acetonitrile in IAA synthesis as well as the precursor for the thiohydroximates in glucosinolate biosynthesis [7-9]. Cytochrome P450 CYP71A13 is involved in the conversion of indole-3-acetaldoxime to indole-3-acetonitrile which is subsequently converted to IAA by nitrilases [10, 11-29]. Cytochrome P450 CYP71A13 is a heme

protein with protoheme as the prosthetic group to catalyze a unique reaction: formation of a carbon-nitrogen triple bond and dehydration of indolyl-3-acetaldoxime. Therefore, it is of interest to analyze the predicted structural model of cytochrome P450 CYP71A13 from Arabidopsis lyrata for further docking procedures with suitable ligands.

Methodology:
Amino acid sequence analysis, template searching and sequence alignment
Amino acid sequences of all known cytochrome P450 CYP71A13 was retrieved from Expert Protein Analysis System (http://www.expasy.ch) i.e. the proteomic server of Swiss Institute of Bioinformatics [12] and retrieved amino acid sequences were aligned using multiple alignment fast fourier transform (MAFFT version 6). A protein BLAST search was performed against PDB (Protein Databank) to retrieve the

corresponding template for the enzyme **[13]**. A BLAST search of the Protein Databank demonstrated strongest similarity with human cytochrome P450 PDB code 3QM4 **[14]**. This similar protein was used as template for modeling, after alignment with the MAFFT program **[15]**
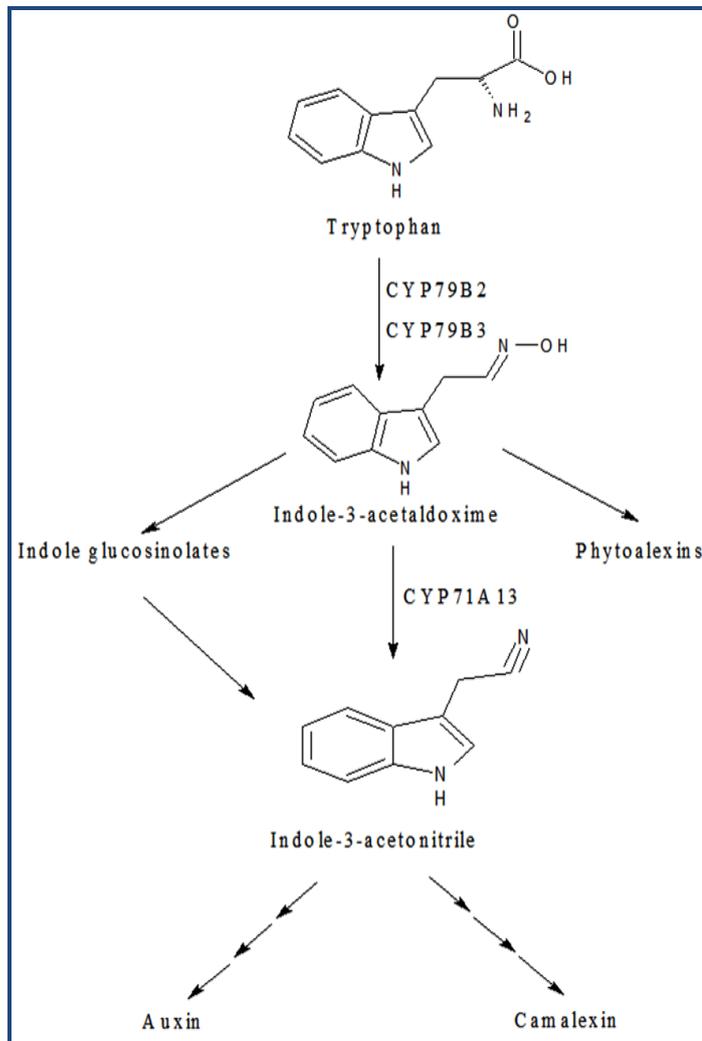

Figure 1: Biosynthesis of auxin and related indolic compounds. In Arabidopsis tryptophan is converted to indole-3-acetaldoxime a precursor of auxin and indole glucosinolates by CYP79B2 and CYP79B3. Indole-3-acetaldoxime is also an intermediate of other phytoalexins cyclobrassinin, brassilexin and spirobrassinin in related cruciferous plants (29). In the biosynthesis of auxin, indole-3-acetaldoxime is dehydrated to indole-3-acetonitrile by CYP71A13, which is then hydrolysed to auxin by nitrilase.

Homology modeling

The model building of cytochrome P450 CYP71A13 using human cytochrome P450 (PDB code 3QM4) was achieved using MODELLER 9.11 program **[16, 17]**. The model was built including the coordinates of heme group from 3QM4 during the modeling. The energy of the model was optimized by several rounds of energy minimizations with Swiss-PdbViewer **[18]**, monitored manually to avoid major rearrangements of the structure. Bond lengths and angles were verified and optimized by consultation of Ramachandran maps followed by energy minimizations.

Figure 2: Sequence alignment of human cytochrome P450 (P10635) and Arabidopsis lyrata cytochrome P450. The alignment was generated using the program MAFFT program (15). Red tubes denote α-helices and cyan β-sheets. Residues shown on a green background are involved in heme anchoring and residues on a magenta background are within five angstroms of the docked substrate.
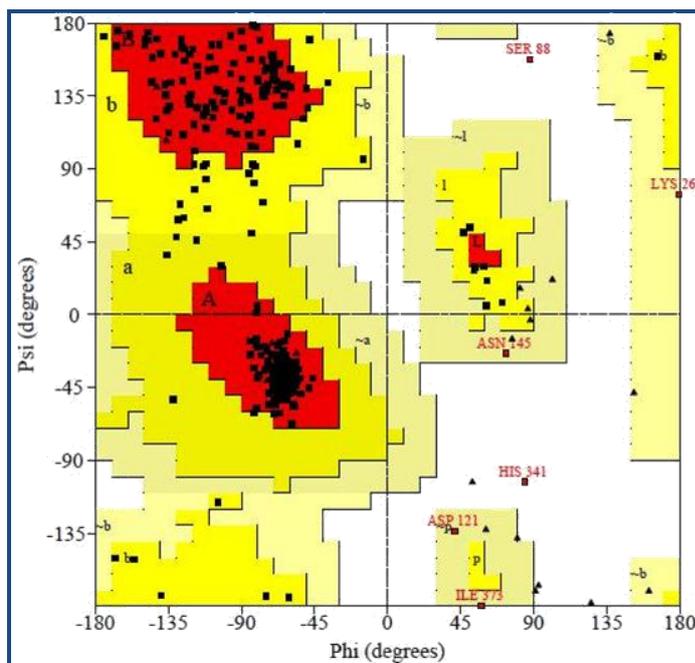
Figure 3: Ramachandran plot of φ-ψ distribution of modeled cytochrome P450 CYP71A13 produced by PROCHECK after homology modeling and energy minimization. (A, B, L) most favoured regions; (a, b, l, p) additional allowed region; (-a, -b, -l, -p) generously allowed region; white areas are disallowed regions.

BIOINFORMATION
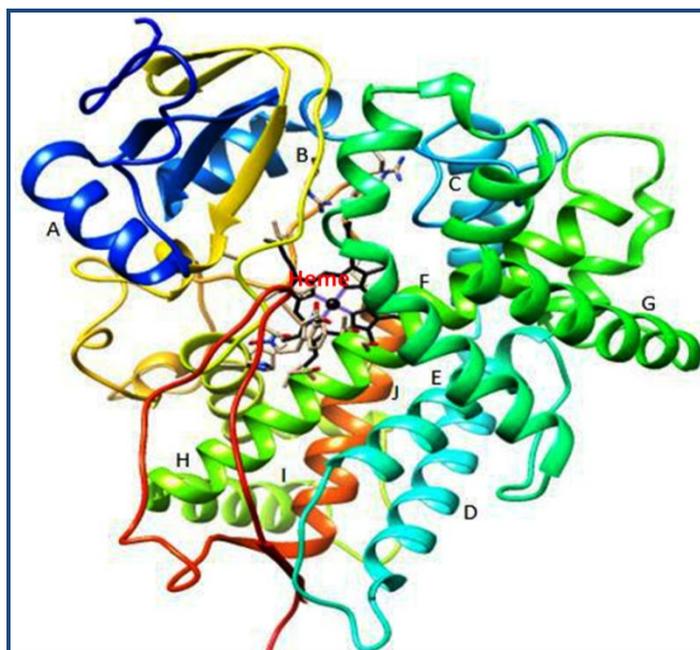
Figure 4: Overview of the modeled cytochrome P450 CYP71A13 structure. The overall structure of cytochrome P450 CYP71A13 was coloured blue to red from N-terminus to C-terminus. The letters indicate the name of the helices.

Physiochemical properties analysis
Different physiochemical properties of human cytochrome P450 [14], cytochrome P450 CYP71A13 of A. lyrata [19] and cytochrome P450 CYP71E1 of Sorgum vulgare [20] were predicted **Table 1 (see supplementary material)** using proteomic server ProtParam [21]. The pI was predicted using pK values of amino acids described in Bjellqvist method [22] and molecular weight was calculated by the addition of average isotopic masses of amino acids in the protein and the average isotopic mass of one water molecule. The extinction coefficient of cytochrome P450 CYP71A13 and human cytochrome P450 was calculated using the Edelhoch method [23] and aliphatic index was calculated according to Ikai [24] method. The GRAVY value was calculated [21] displaying the hydropathicity of a protein.

Substrate docking
Three dimensional structure of indolyl-3-acetaldoxime was prepared using ProDrg server [25] and docked in the model of cytochrome P450 CYP71A13 using the Autodock 4.1 program [26, 27] which uses a genetic algorithm as a global optimizer combined with energy minimization as a local search method. The active site was defined as a collection of amino acid residues enclosed within a sphere of 5 Å (1 Å=0.1 nm) radius centered on the catalytic triad of Arg84, His233 and Ser127. Autotors program of Autodock was used to define torsional degrees of freedom in the ligand. The minimization was achieved using Lamarckian genetic algorithm (LGA). The grid size was set to 40 x 46 x 40 points, centered on the heme in the crystal structure of the complex. Each docking run consisted of 100 independent docks with 1000 iteration cycles. A random start was used to generate the substrate position within the docking box. The substrate orientation giving the lowest interaction energy was chosen for additional rounds of docking.

Results & discussion:
The alignment used for building the homology model is shown in **Figure 2** the amino acid sequence identity of cytochrome P450 CYP71A13 was 27% to 3QM4. Based on sequence alignment as an input, the MODELLER 9.11 software generated a large number of spatial restraints from the template structure and constructed a molecular model of cytochrome P450 CYP71A13. The possible applications of generated model depend mainly on the quality of the obtained model. The quality of generated model of cytochrome P450 CYP71A13 was evaluated by PROCHECK [28] which indicated that more than 98.5% of residues φ-ψ angles were in favoured and additional allowed region of Ramachandran plot (**Figure 3**). Again it indicated that the finally obtained 3D model of cytochrome P450 CYP71A13 was satisfactory with respect to Ramachandran plot where only two residues were in disallowed region and four (1.0%) were in generously allowed regions. Residues located in unfavourable region were far from substrate binding domain indicating that these residues may not affect the ligand protein binding simulation. Analysis of physiochemical properties revealed that cytochrome P450 CYP71A13 had higher content of negatively and positively charged amino acids (62, 69 respectively) resulting in increased hydrophilic character **Table 1 (see supplementary material).** The values of pI showed that plant cytochrome P450 CYP71A13 and cytochrome P450 CYP71E1 was stable at higher pH as compared to human cytochrome P450. The final models of cytochrome P450 CYP71A13 contained α+β structure consisting of ten major α-helices that form the conserved 3D fold of CYP proteins and seven β- sheets (**Figure 4**). The heme existed between H-helix and J- helix in the hydrophobic pocket where Gly305 and Thr308, 311 of H- helix were involved in its stabilization (**Figure 5D**). Analysis of the surface of the structural model of cytochrome P450 CYP71A13 identified surface exposed hydrophilic domains in the region positioned near the N-terminal region while a hydrophobic core was present near the C-terminal region (**Figure 5A).** The hydrophobic core was formed by Leu218, 370, 364, 452, Phe438, Pro446, Ile448 and Ala451 amino acids near the C-terminal region. This region seemed to be most likely the binding site for substrate.

Cytochrome P450 CYP71A13 converts indole-3-acetaldoxime to indole-3-acetonitrile. The substrate indole-3-acetaldoxime was docked into the substrate binding pocket so that the atoms involved in the initial dehydration reaction were positioned above the heme. The smaller size of the substrate binding pocket in cytochrome P450 CYP71A13 was due to the bulkiness of the two amino acid residues Phe182 and Trp315 pointing into the substrate binding cavity. Indole group of the oxime was positioned slightly closer to the heme iron and O-atom towards the Ser369 and His367 respectively with a final distance of 2.1Å and 4.3 Å (**Figure 5B).** The substrate indole-3-acetaldoxime was tightly fitted into the substrate pocket with the aromatic ring being surrounded by residues creating a hydrophobic environment. Leu312, 370 and Pro368 from the C-terminal were positioned favourably for interaction with the indole group of the substrates in the proximal end of the pocket, while Ser369 in the heme distal end of the substrate pocket was positioned to interact with the leaving group of indole-3-acetaldoxime (**Figure 5C).** The apparent role of the heme in cytochrome P450 CYP71A13 was to tether the

substrate in the catalysis by indole-3-acetaldoxime dehydratase. This homology model revealed that hydrogen bonding between the OH group of indole-3-acetaldoxime and the side chains of Ser369 and His367 controls the specific orientation of the heme-bound substrate suitable for the elimination of the OH group of indole-3-acetaldoxime and that these residues and the heme created a prefixed site for substrate recognition and binding **(Figure 5C).**
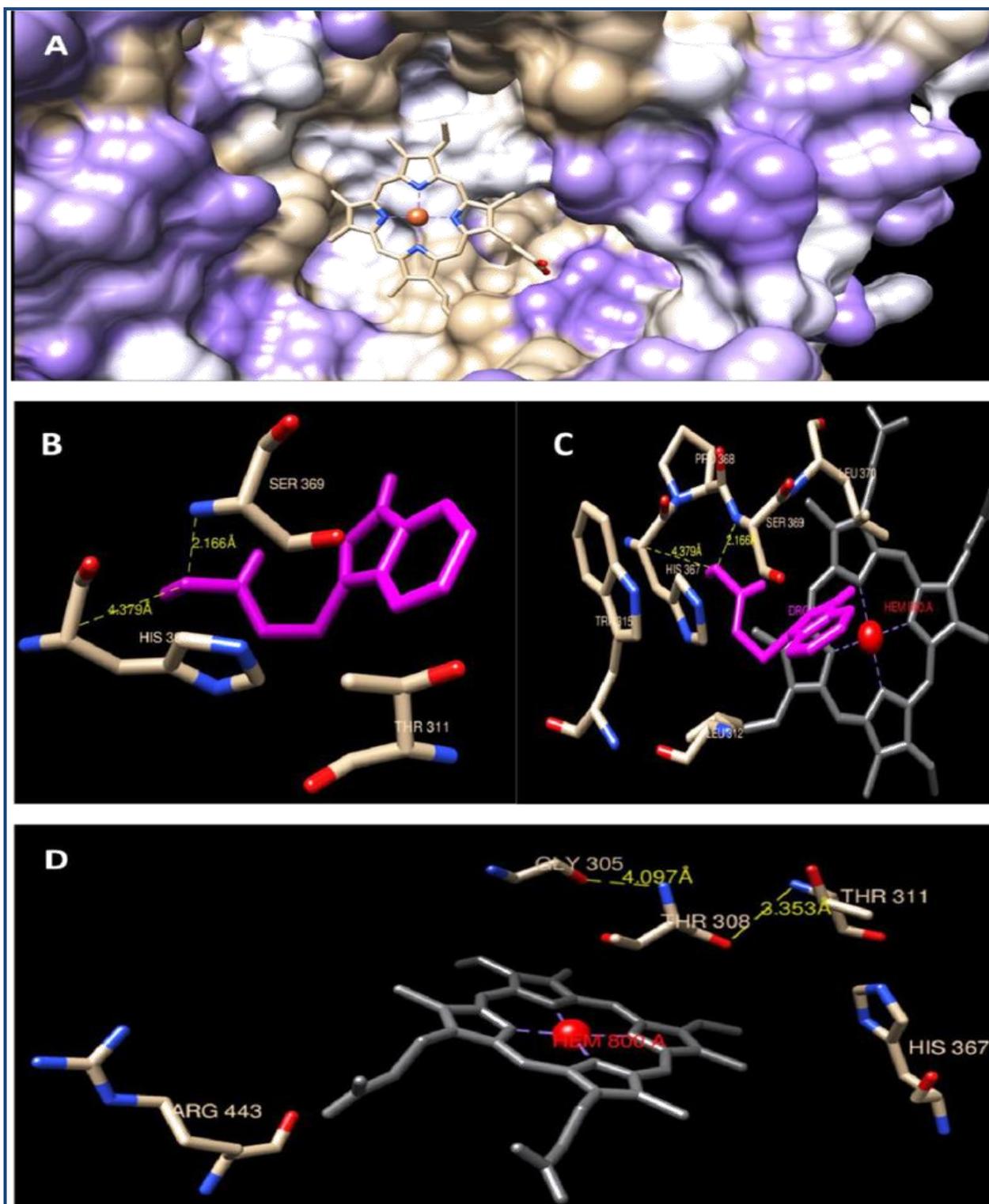
Figure 5: Active site of cytochrome P450 CYP71A13 from *Arabidopsis lyrata* **A)** Surface representation of amino acid hydrophobicity in cytochrome P450 CYP71A13 with a color scale that varies from blue to white, representing dodger blue for the most hydrophilic, to white for the most hydrophobic residues; **B** and **C** show features of the substrate indole-3-acetaldoxime binding pocket of cytochrome P450 CYP71A13 and D represents amino acids involved in stabilization of heme in cytochrome P450 CYP71A

BIOINFORMATION

Conclusion:
We describe structural model of Arabidopsis lyrata cytochrome P450 CYP71A13 involved in auxin synthesis generated using molecular modeling techniques. The model was further docked with indole-3-acetaldoxime to explore structural features and binding mechanism. Analysis of active site and binding specificities of indole-3 acetaldoxime to the enzyme show two amino acid residues (His367 and Ser369) for catalytic activity and substrate binding. The model complex provide insights for the understanding of mechanism of substrate recognition and its conversion to corresponding nitrile by cytochrome P450 CYP71A31.

References:
[1] Muday GK, *J Plant Growth Regul*. 2001 **20**: 226 [PMID: 12033223]
[2] Casimiro I *et al. Trends Plant Sci*. 2003 **4**: 165 [PMID: 12711228]
[3] Friml J, *Curr Opin Plant Biol*. 2003 **6**: 7 [PMID: 12495745]
[4] Grebe M, *Bioessays* 2004 **26**: 719 [PMID: 15221854]
[5] Normanly J *et al. Plant Physiol*. 1995 **107**: 323 [PMID: 12228361]
[6] Bartel B, *Ann Rev Plant Physiol Plant Mol Biol*. 1997 **48**: 51 [PMID: 15012256]
[7] Normanly J & Bartel B, *Curr Opin Plant Biol*. 1999 **2**: 207 [PMID: 10375566]
[8] Hull AK *et al. Proc Natl Acad Sci U S A*. 2000 **97**: 2379 [PMID: 10681464]
[9] Mikkelsen MD *et al. J Biol Chem*. 2000 **275**: 33712 [PMID: 10922360]
[10] Nafisi M *et al. Plant Cell*. 2007 **19**: 2039 [PMID: 17573535]
[11] Mano Y & Nemoto K, *J Exp Bot*. 2012 **63**: 2853 [PMID: 22447967]
[12] Schneider M *et al. Plant Physiol Biochem*. 2005 **42**: 1013 [PMID: 15707838]
[13] Altschul SF *et al. J Mol Biol*. 1990 **215**: 403 [PMID: 2231712]
[14] Wang A *et al. J Biol Chem* 2012 **287**: 10834 [PMID: 22308038]
[15] Katoh K *et al. Nucleic Acids Res*. 2002 **30**: 3059 [PMID: 12136088]
[16] Eswar N *et al. Curr Protoc Bioinformatics*. 2006 **5**: 5 [PMID: 18428767]
[17] Sali A & Blundell TL, *J Mol Biol*. 1993 **234**: 779 [PMID: 8254673]
[18] Johansson MU *et al. BMC Bioinformatics* 2012 **13**: 173 [PMID: 22823337]
[19] Hu TT *et al. Nat Genet*. 2011 **43**: 476 [PMID: 21478890]
[20] Bak S *et al. Plant Mol Biol*. 1998 **36**: 393 [PMID: 9484480]
[21] Kyte J & Doolittle RF, *J Mol Biol*. 1982 **157**: 105 [PMID: 7108955]
[22] Bjellqvist B *et al. Electrophoresis*. 1994 **15**: 529 [PMID: 8055880]
[23] Edelhoch H, *Biochemistry*. 1967 **6**: 1948 [PMID: 6049437]
[24] Ikai A, *J Biochem* 1980 **6**: 1895 [PMID: 7462208]
[25] Schuttelkopf AW & van Aalten DM, *Acta Crystallogr D Biol Crystallogr*. 2004 **60**: 1355 [PMID: 15272157]
[26] Morris GM *et al. J Comput Chem*. 1998 **19**: 1639
[27] Goodsell DS *et al. J Mol Recognit*. 1996 **9**: 1 [PMID: 8723313]
[28] Laskowski RA *et al. J App Cryst*. 2004 **26**: 283
[29] Pedras MS *et al. Phytochemistry*. 2000 **53**: 161 [PMID: 10680168]

BIOINFORMATION

Supplementary material:

Table 1: Analysis of physiochemical properties of human cytochrome P450, cytochrome P450 CYP71A13 and cytochrome P450 CYP71E1

Parameters	*Homo sapiens* cytochrome P450	*Arabidopsis lyrata* cytochrome P450 CYP71A13	*Sorghum vulgare* cytochrome P450 CYP71E1
Uniprot accession number	P10635	D7LC87	O48958
Theoretical pI	6.77	8.94	8.76
Negatively chargedresidues (Asp + Glu)	53	62	60
Positively charged residues	51	69	66
Extinction coefficient (M^{-1}cm^{-1})at 280nm	47815	51910	55850
Instability index	44.63	38.66	36.81
Aliphatic index	95.15	99.01	89.83
GRAVY	-0.031	-0.179	-0.147