Molecular Modeling of Cytochrome P450 1a1 using the newly Crystallized Template Structure of CYP1A2-Human

Rehan Zafar, Fahed Parvaiz, Babar Aslam, Umar Niazi
NUST Center of Virology & Immunology
rehanzfr@gmail.com, mfahed85@gmail.com
BABARMED1@yahoo.com, saint_6671@yahoo.com

Abstract

Cytochrome family 1 enzymes play significant roles in carcinogenesis and xenobiotic detoxification. CYP1A1 is the P450 family 1 enzyme preferably expressed extrahepatically and participates extensively in monooxygenase activity which can either change the substrate to normal or carcinogenic metabolites, having the ability to initiateoncogenesis in lung and breast. Variegated structural properties evident in the prosites of available Cytochrome P450 (CYP) structures show versatility among CYP catalyzed reactions. In order to understand the CYP1A1 functions, hypothesized homology model has been constructed and characterization of the active site was performed by identifying important residues using docking studies and pharmacophore analysis. Model of CYP1A1-Human has been constructed using the available crystal structure of CYP1A2-Human. Active site and entry site of CYP1A1 was found to be more compact than CYP1A2. Difference of wild type CYP1A1 against its polymorphisms shows the role of mutations in the active site architecture, which explain that M2 and M4 mutations in CYP1A1 have no possible significant roles in the substrate binding and orientation for detoxification or carcinogenic activation. Different ligands including Alpha- naphthoflavone (ANF), Ethoxyresorufin, Theophylline, Tamoxifen, Ethanol, Phenacetin and Hesperetin were docked and reconfirm the ligand specific wet lab studies.

Key words: CYP1A1 , Homology Modelling, CYP1A2.

Introduction

Eighteen families of Cytochrome P450 and 43 subfamilies are ultimately translated from 57 genes of human with more than 58 pseudogenes (Nelson DR *et.al.,* 2004). They are the part of a multi-component electron transfer chains, called Microsomal Cytochrome P450-containing systems. P450s constitute multigenic superfamily of hemoproteins playing their biological roles as heme-thiolate monooxygenases enzymes in the synthesis and breakdown of endogenous compounds such as synthesis and metabolism of steroids and bile acids, metabolism of vitamin D and synthesis of cholesterol. They are also involved in detoxification of exogenous compounds such as hydrophobic xenobiotics (Nebert 1991) and activation of procarcinogenic to carcinogenic compounds, For example, the oxidation of benz(a)pyrene in cigarette smoke is catalysed by CYP1A1 to form BP-7,8-epoxide, which can be further oxidized by epoxide hydrolase to form 7,8-dihydrodiol (Shimada T and Oda Y 2001). Finally CYP1A1 catalyses this intermediate to form 7,8-dihydrodiol-9,10-epoxide, which is the ultimate carcinogen.(Beresford AP 1993; Guengerich FP 1995; Kawajiri K and Hayashi SI 1996). Such compounds form adducts with the DNA and proteins causing disruption of their normal physiological role (Rendic S and Di Carlo FJ 1997). [Some families of CYPs have the ability to metabolize multiple substrates, which accounts for their central role in drug to drug interactions.

In mammals, family 1 contains three well characterized P450s; CYP1A1, CYP1A2, and CYP1B1 involved extensively in the biotransformation of xenobiotics to more polar form for efficient excretion. Versatile nature of these enzymes is their capacity to oxidize multiple Polynuclear Aromatic Hydrocarbons (PAHs) (Yano JK, Hsu MH *et al.* 2005). Their induction is mediated by a ligand-activated transcription factor, the aryl hydrocarbon receptor having a basic-loop-helix PAS domain protein which binds to enhancers flanking the CYP1A1, CYP1A2 and CYP1B1 genes and stimulates their transcription.

CYP1A1 is also known as aryl hydrocarbon hydroxylase (AHH).. CYP1A1 is a P450 comprised of 512 amino acids. Its subcellular locations are endoplasmic reticulum and microsome membrane specifically expressed in lung, lymphocytes and placenta. Due to the specificity for tissues other than liver, CYP1A1 is considered to be extrahepatic. CYP1A1 mRNA levels are depressed by inflammatory cytokines (including interleukin-6 (Il-6), interleukin-1 alpha (Il-1 alpha), and tumor necrosis factor-alpha (TNF-alpha)) and growth factors (Barker CW *et.al.* 1992). A common reason for the repression of CYP1A1 could be the involvement of reactive oxygen species (ROS) (Morel Y and Barouki R 1998). Fluoroquinolones and macrolides can inhibit the expression of CYP1A1. CYP1A1 inhibitors like aryl hydrocarbon receptor (AhR) antagonist, resveratrol, in red wine can have anti-cancer nutrition activity (Casper RF *et.al.* 1999).

Several alleles have been identified of CYP1A1 gene include *1A1*1, 1A1*2A* (T>C; M1), *1A1*2B, 1A1*2C(A>G; M2), 1A1*3(T>C; M3), 1A1*4(C>A; M4) and 1A1*5 to 11* (Ingelman-Sundberg M *et.al.* 2001).*Translated products of some of which have high inducible AHH activity.*

Two polymorphisms i.e. *M1* and *M2* have been

Corresponding Author: Rehan Zafar

studied extensively in relation to cancer susceptibility and associated with high risk of smoking-induced lung cancer in Asians but not in Caucasians (Kawairi *et.al.* 1990). Recently *1A1*2B* and *M2* were found for causing increased risk of breast cancer (Moreno M *et.al.* 2008. These polymorphisms or genetic variability can be important contributors to inter individual preferences in drug biotransformation and different abilities to metabolize drugs, leading to extensive or slow metabolizers. But recent studies made an ambiguity for the mutation specific carcinogenesis activity of CYP1A1(Persson I *et.al.* 1997)

CYP1A1 have 73% amino acids sequence identity to CYP1A2, but both are susceptible to different substrates in terms of specifity and inhibition(Guengerich FP 1995; Kawajiri K and Hayashi SI 1996) CYP1A1 is involved in metabolising benz[a]pyrene and other polycyclic aromatic hydrocarbons (PAHs) to their toxic products (Guengerich FP 1995; Kawajiri K and Hayashi SI 1996; Shou M et.al. 1996), while CYP1A2 preferably oxidizes heterocyclic and aromatic amines (Hammons GJ *et.al.* 1997; Turesky RJ *et.al.* 1998).

The characterization of CYP1A1 structure will contribute to greater extent in understanding the enzyme function and modes of catalysis and may provide groundwork for the rational design of drugs and inhibitors.

CYP1A1 homology models based on several crystallographic templates, P450TERP, P450BM3, P450CAM and rabbit CYP2C5 have been reported (Iori F *et.al.* 2005; Szklarz GD and Paulsen MD 2002). Homology model dependent upon low sequence identity between CYP1A1 and the bacterial P450s or human CYPs 2C5, 2C8, 2C9 and CYP1A1 have also been reported (Lewis BC and Mackenzie PI 2007), but are unlikely to provide an accurate representation of the human CYP1A1 active-site.

This report attempts to provide the homology model and characterization of wild type human CYP1A1 and its M2 and M4 polymorphs. Different ligands including Alpha- naphthoflavone (ANF), Ethoxyresorufin, Theophylline, Tamoxifen, Ethanol, Phenacetin and Hesperetin were docked and reconfirmed for the ligand specific wet lab studies.

Experimental procedures

Homology modelling of CYP1A1

The complete sequence of the CYP1A1 was taken from UniProtKB/Swiss-Prot database (Accession # P04798) (Swiss-Prot Protein knowledgebase) (http://www.uniprot.org/). The protein sequence of CYP1A2 was 514aa long. The sequence of CYP1A1 protein was searched for reference sequence entries at NCBI (www.ncbi.nlm.nih.gov) using the algorithm PSI-BLAST from PDB database (Medha B and Aravind L 2007). Based on the distance in the evolutionary tree from CYP1A1 and the percentage sequence identity, the template structure having the PDB ID 2HI4 (CYP1A2) was chosen. Sequences of CYPs 1A1 & 1A2 were aligned using server based 3D-Coffee alignment program (O'Sullivan O 2004) and the alignment refined using RASCAL (Thompson JJD *et.al.* 2003). MODELLER release 9v7 (Eswar N *et.al.* 2006) was used for comparative modeling of the CYP1A1 structure based on the CYP1A2 template structure (PDB ID 2HI4).

Thirty two models for CYP1A1 were generated, and sixteen models of these models with lowest energies were clustered into 4 groups based on the root mean square distance (RMSD) between the corresponding residues in their structures using NMRCLUST (Kelley LA 1996). Representative 3 models from each cluster and outliers were selected for further analysis and, the free energy of the models, Ramachandran plot calculated using PROCHECK (Roman A *et.al.* 1993) and ERRAT (Colovos C and Yeates TO 1993) score were used as criterion for final model selection.

Docking of ligands:

Different ligands including Alpha-naphthoflavone (ANF), Ethoxyresorufin, Theophylline, Tamoxifen, Ethanol, Phenacetin and Hesperetin were docked in the heme neighbour. Each of these ligands was docked to generate fifty docking poses using GOLD (Abecasis GR and Cookson WO, 2000) within 15 Å of heme. The population size was 100, with a selection pressure of 1.1, 5 islands, niche size of 2, maximum operations were set to 100000 and the protein side chains were fixed. The resulting fifty docking poses for each ligand were ranked according to their corresponding binding score. The "correct" binding pose for each ligand, out of the fifty resulting poses, was chosen based on the position of the expected site of metabolism relative to the heme and the binding score. The results were then analyzed using PyMol (DeLano WL 2002) and SwissPdb Viewer (Guex N and Peitsch MC 1996).

Results and discussion

The phylogenetic tree of CYP1A1 and reference sequence entries in NCBI showed that CYP1A1 was closest to 2HI4 (CYP1A2-human), 2OJD (CYP2R1) and 2F9Q (CYP2D6); while sequence identity results showed that CYP1A1 has the highest percentage sequence identity with 2HI4 (CYP1A2) of 73%. Based on the distance in the evolutionary tree from CYP1A1 and the higher percentage sequence identity, the template structure chosen was 2HI4 (CYP1A2).

Figure 1 shows the ERRAT plot for the "best" modeled CYP1A1. It had an overall quality factor of 57.32 and had areas of high error at the N and C terminals. Additionally some areas of the active site also had a high error i.e. BC loop area (110 to 115aa), the area between the F and G helix (235 to 245aa), some residues in the I helix (315 to 320aa), Substrate recognition sites (Gotoh, 1992) (SRSs) i.e. SRS5 area (382-385aa) and SRS6 area (490 to 500aa).

The Ramachandran plots for the "best" CYP1A1 model showed good stereochemistry with 90% of residues for CYP1A1 in most favoured regions while only two residues were in disallowed regions (Ramachandran GGN *et.al.* 1963).

Human cytochrome P450 1A1 structure

The structure of modeled CYP1A1 exhibits some of the same structural properties of CYP1A2, with some structural divergence. Figure 2 shows the CYP1A1 sequence with the secondary structure elements labeled as arrows (sheets) and coils (helices)

Possible entry site

The active site of P450s is located inside the protein. The recognition of substrates thus sometimes requires dynamic changes in the protein to provide an access route as well as specific binding in the active site. Similarly an egress route through the protein provides a way for release of the product, which may be the rate-limiting step. The substrate

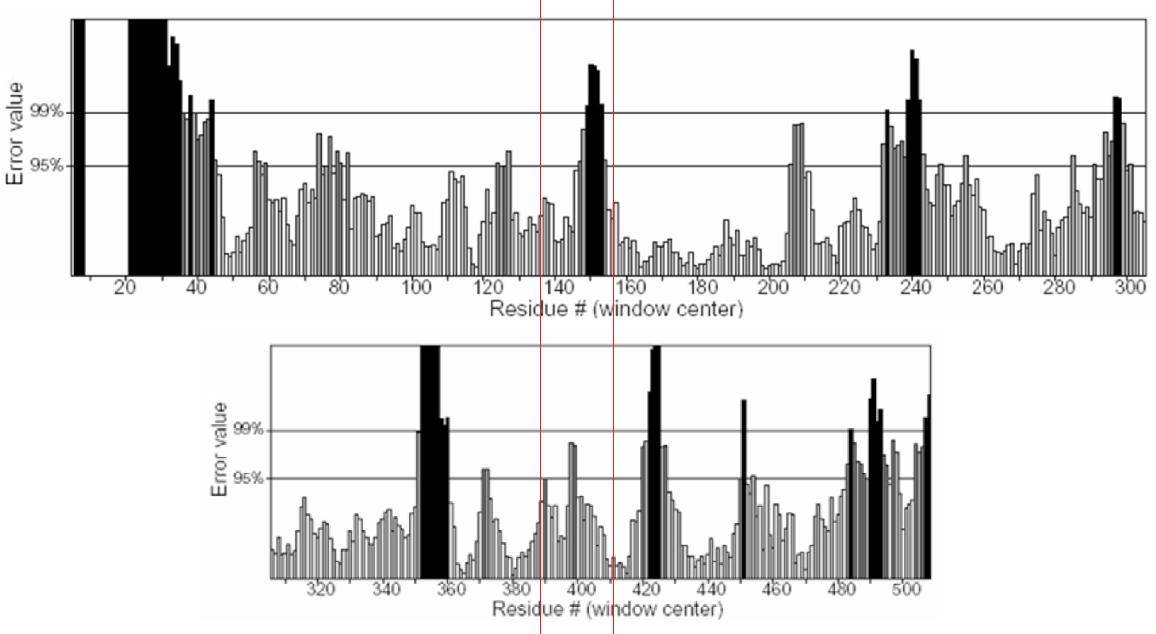

Fig. 1. ERRAT plot for CYP1A1. On the error axis (y axis), two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value. The regions with high error value (black lines) indicate areas of high energy or clashes in the structure.

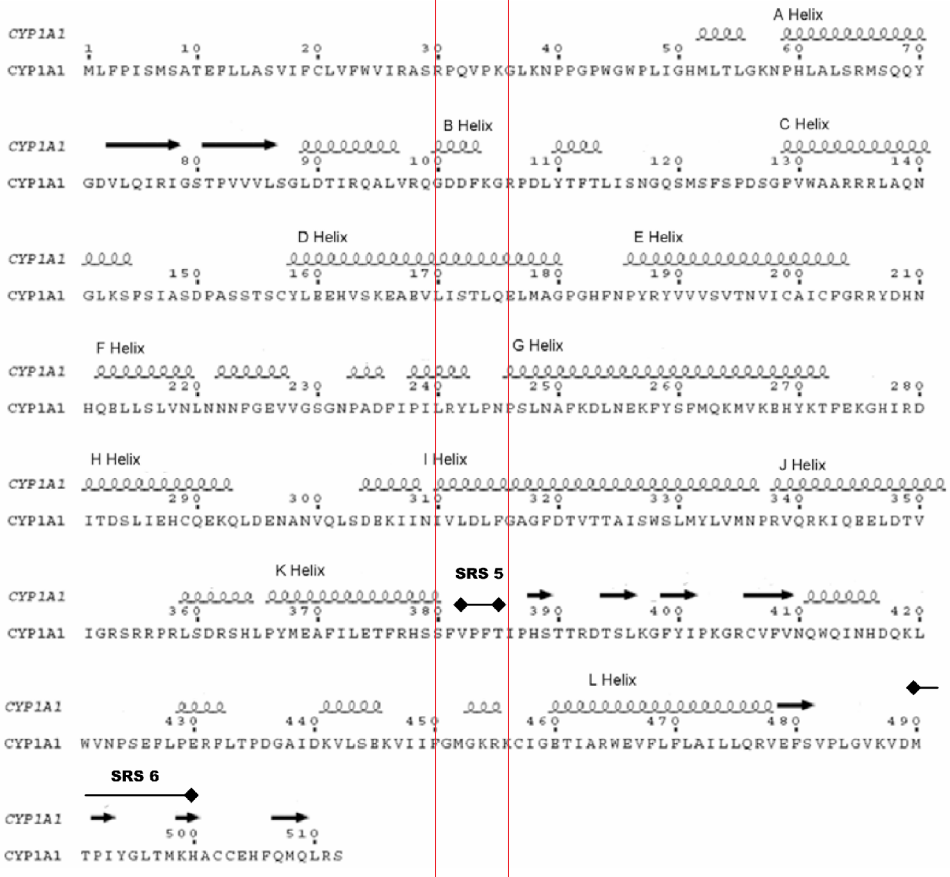

Fig. 2. Amino Acid sequence of CYP1A1 with the sheets labeled as arrows; the major helices are named and labeled as coils.

specificity of P450s can be influenced not only by the interactions made in the active site but also by accessibility to and from the active site, which highlights the importance of entrance and egress channels.

Figure 3 shows one possible entry channel into the active site. This was determined using the "Alpha site finder" utility in the program MOE (Molecular Operating Environment) (Chemical Computing group), and it corresponds to the entry channel for CYP2D6 (PDBID 2F9Q) (Rowland P *et.al.* 2006). The second potential channel appears to be between BC loop, G and I helices, similar to pw2c pathway identified by Rebecca Wade et al (Wade RC *et.al.* 2005). CYP1A2 crystal structure (PDBID 2HI4) does not show any entry channels. The entrance channel for our model appears to be lined by polar and charged residues from the F helix, I helix and E helix. The arrows show the possible entrance path. Asn219 from the F helix is replaced by Lys in CYP1A2, Asp in CYP2D6 and Ser in CYP1B1. Asn222 is replaced by His in CYP1A2, Gln in CYP2D6 and Glu in CYP1B1. Asn223 is only present in CYP1A1, while the CYPs 1B1, 1A2 and 2D6 have a Glu in this position. On the SRS6, Lys499 is conserved in CYPs 1A1, 1A2 and 1B1, while 2D6 has a SER in this position. Arg188 and Tyr187 from the E Helix also point into the entrance channel.

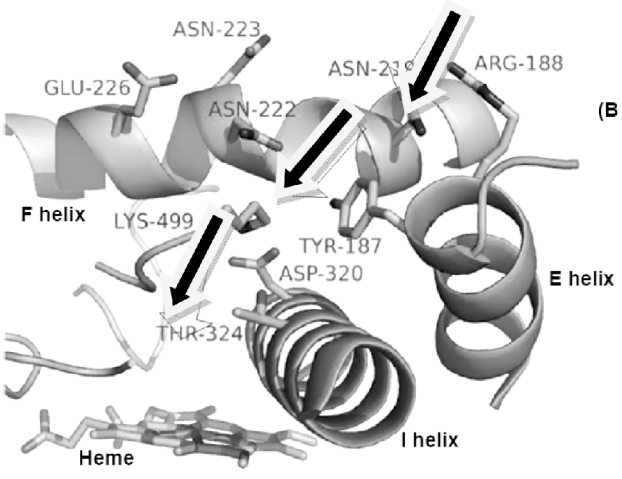

Fig. 3. Possible entry channel into the active site of CYP1A1 marked by arrows, surrounded by residues from the F helix, I helix and E helix.

Active site of CYP1A1

Active site of CYP1A1 is composed of BC loop (Figure 4), I helix (Figure 5), SRS 5 and 6 (Figure 6), F helix (Figure 7), G helix (Figure 8). The residues along the I Helix, BC Loop and SRS 5 area surround the heme, helping to place the ligand on top of the heme.

In the BC loop (Figure 4), Phe123 is conserved across a range of P450s and provides important pi stacking hydrophobic interactions placing the ligand on top of the heme. Ser122 and Ser120 can provide important hydrogen bonding interactions and Ile115 can provide hydrophobic interactions. Ser122 in CYP1A1 aligns with Thr124 in CYP1A2 (Data is not shown here). Thr124 in CYP1A2 is important in substrate binding. This was also observed in

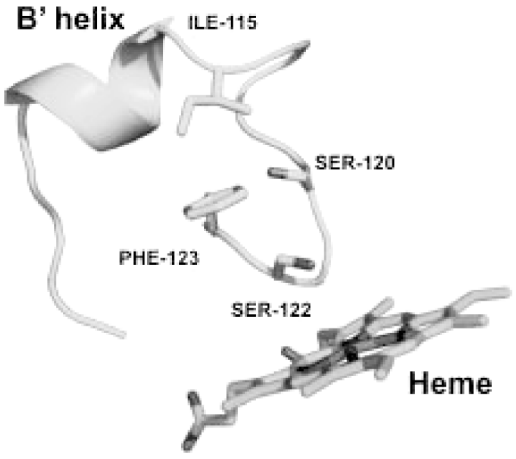

Fig. 4. Residues in the BC loop

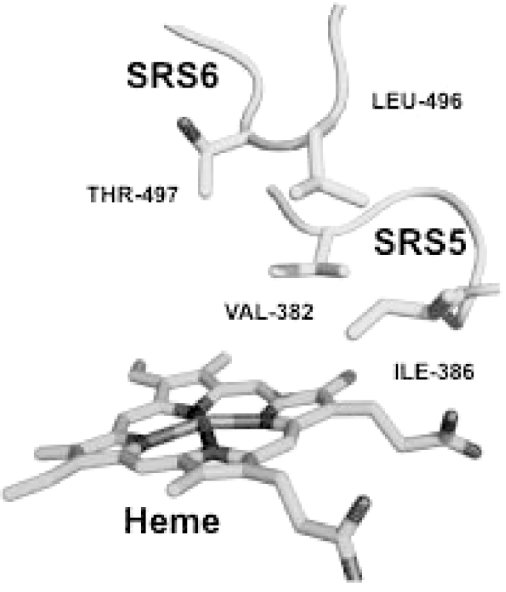

Fig. 5. Residues of I helix are oriented towards the active site in CYP1A1.

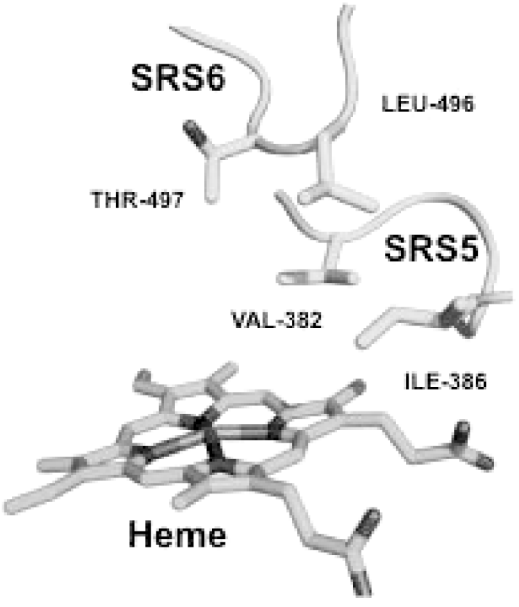

Fig. 6. Val382 and Ile386 (SRS5); Thr497 and Leu496 (SRS6) for CYP1A1 are shown.

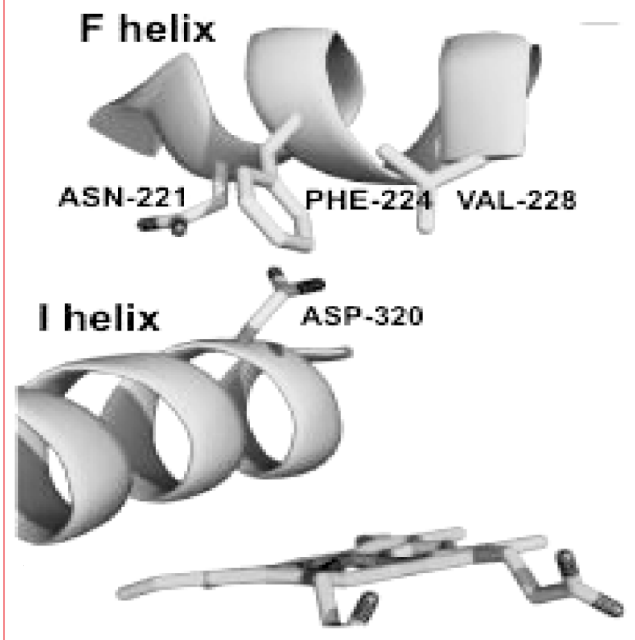

Fig. 7. Residues from F helix oriented towards the active site in CYP1A1. Asp320 (I helix) is also visible.

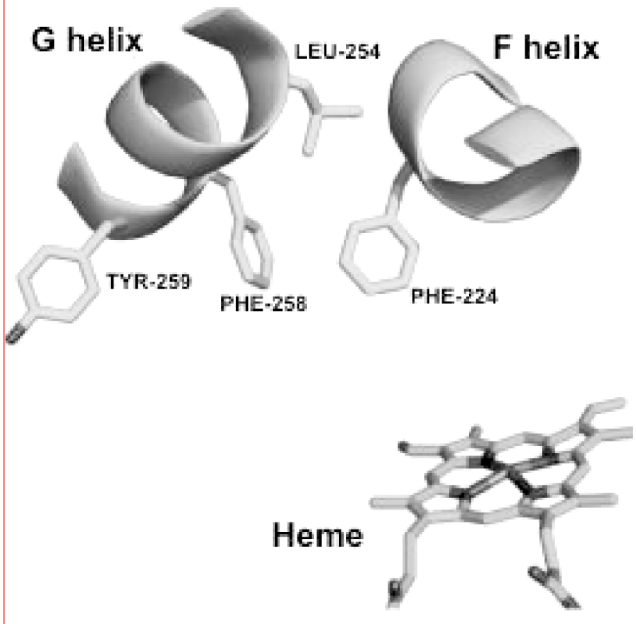

Fig. 8. Residues of G helix from CYP1A1 oriented into the active site. Phe224 from F helix is shown as well

equivalent CYP1A1 S122T mutant, which displayed significantly increased O-dealkylation activity for both 7-ethoxy and 7-methoxyresorufin over CYP1A1 wild-type (Sansen S *et. al.* 2007).

Asp313. Asp320, Ala317, Thr321 present in helix I (Figure 5) are conserved in CYPs 1A1 and 1A2.

Val382 and Ile386 (SRS5); Thr497 and Leu496 (SRS6) are shown in Figure 6. Modeling suggests that Val382 provides important hydrophobic interactions to position the ligand on top of the heme, e.g. in the case of ethoxyresorufin, the ethyl group forms hydrophobic interactions with this Val382 thus positioning the ligand on top of the heme for O-de-ethylation (EROD Activity: – Figure 12). The ethoxyresorufin-*O*-deethylase (EROD) assay is utilized for observing the induction of the CYP1A1 and is used as a biomarker for the exposure of xenobiotics that bind with to AhR (Burke M and Mayer R. 1974).

CYP1A2 has a Leu in the position of CYP1A1:Val382. Leu is a larger residue than Val, which is likely why wild-type CYP1A1 (Val-shorter residue) versus CYP1A2 shows a clear preference for 7-ethoxyresorufin versus 7-methoxyresorufin O-dealkylation. The reciprocal CYP1A1 V382L and CYP1A1 L382V mutants display interchanged specificities (Sansen SS *et.al.* 2007). Ile386 is conserved in both CYPs i.e 1A1, 1A2. Mutagenesis studies on Ile386 in CYP1A2 have shown to alter substrate specificity, e.g. I386T mutation showed a marked shift in catalytic specificity, with much reduced phenacetin O-de-ethylation but increased EROD activity (Zhou HH *et.al.* 2004). Val382, Ile386 and Leu496 appear to form an important hydrophobic network.

Phe224 present in F helix (Figure 7) provides important pi stacking interactions for the aromatic ligands, and is also conserved in CYP1A2. Asn221:1A1 (F helix) is important in determining the flexibility of CYP1A1 active site because in CYP1A2 this Asn is replaced by a Thr223 which leads to a hydrogen bond formation between Thr223 (F Helix) and Asp320 (I Helix), making the active site restricted (Sansen SS *et.al.* 2007).

In CYP1A1, Leu is present at the 254 position instead of Phe:1A2 (Figure 8). Phe258 is found conserved both in CYP1A1 and CYP1A2. Phe258 and Phe224 of F helix (Figure 8) together can form an aromatic network, sandwiching the ligand between them. Tyr259 (G helix) in CYP1A1 is present instead of a Leu (CYP1A2) in this position. It might be important, but in the current rotamer conformation it appears to be too far from the main active site cavity.

Amino acids substitutions in more than 5 polymorphisms of CYP1A1 characterize different biological responses; most of them are concerned with the activation of carcinogenic property of compounds. From previous studies, we came to know that in M2 polymorphs the Ile462 is mutated to the Val-462. The major oxidative routes of estrone and estradiol are 2- and 4-hydroxylation by cytochrome P450 2B1, 1A and 3A. The presence of polymorphs can ultimately generate 2-OH derivative of estradiol in excess which can ultimately fatal in breast cancers. The structure determination of Polymorphisms 1A1*2A (M1), 1A1*2B and 1A1*2C (M2) can thus be defended for their role in oncogenesis.

In M2, the difference of Ile-462-Val was thought as a major fact for the oncogenic biotransformation of metabolites. But Persson and collegues (Persson I *et.al.* 1997) observed that the Val-462 variant was not functionally important in the PAH induced cancerogenesis. Furthermore, Smart and Daly (Smart J and Daly AK 2000) reported that polymorphisms in the Ah receptor gene contributed more to CYP1A1 levels than did the polymorphism in the CYP1A1 protein. From our modeled CYP1A1 structure, it can be supposed that the mutation in Ile462Val would neither induce any conformational change

in the active site of CYP1A1 nor participate in the orientation of different substrates for catalysis.

Actually, the 462 position was on the downward side of chair like coordinates for active site (Figure 9). The heme group is present in y axis whereas the 462 position is with –y value in coordinates. (Figure 9). As the role of the CYP1A1, DNA sequence variants in cigarette smoking related cancer development is not yet resolved (Watanabe M 1998; Houlston RS 2000) but the model of CYP1A1 showed that the M2 polymorph has nothing to do with the active site disruption which can ultimately result in the oncogenesis.

Similarly the mutation Thr461Asn in M4 polymorph of CYP1A1 has shown that the effect of mutations should not affect the substrate binding in the active site and confirmed by the previous wet lab studies of these mutations and their substrate binding specificity or metabolisation.

Docking analysis

The single preferred orientation, observed for Alpha-naphthoflavone (ANF) binding with CYP1A2 is shown in Figure 10. ANF is metabolised by CYP1A1 to form ANF-5, 6-diol and ANF-5,6-oxide, and is docking in a pose to be metabolised at position 5 and 6 (Figure 11). The possible reasons for this difference between ANF binding in CYP1A1 and CYP1A2 could be explained by looking at s 10 and 11. CYP1A2 has a triad of Phe, while CYP1A1 is missing the Phe256, instead being replaced here by Leu254. Leu254 (CYP1A1) is about 5.6 Å away from the ANF, which makes it a very weak hydrophobic interaction; while Phe256 (CYP1A2) is about 4.2 Å away, thus having stronger hydrophobic interactions.

The second difference, which appears to be important is CYP1A1 having a Val382 on SRS5, which is smaller and creates more space for the triple ring of ANF to fit there. On the other hand CYP1A2 has Leu382 in this position, which is a longer residue. The triple ring of ANF will clash with Leu382 if it binds that way. These two differences seem to be the most obvious for ANF binding one way in CYP1A1 and the other way in CYP1A2.

The docking pose for ethoxyresorufin (Figure 12) shows a good position for O-de-ethylation to occur. Phe123 provides pi stacking interactions; Ser122 makes a hydrogen bond with the nitrogen while the ethyl group has hydrophobic interactions Val382, Ile386 and Leu495.

Theophylline (Rendic SS 2002) can be metabolised by CYP1A1 at positions 1, 3 or 8. Figure 13 shows the docking pose for theophylline in CYP1A1 model. It is in a good orientation for 1-de-methylation. Asp320 makes a hydrogen bond with the Nitrogen at position 7, while Ser122 donates a hydrogen bond to the Oxygen.

Tamoxifen is metabolised by CYP1A2 and CYP1A1 into N-desmethylTAM (Crewe HK *et.al.* 2002). Figure 14 shows the docking pose of tamoxifen into the CYP1A1 model. Phe224 and Phe258 sandwich the ring, providing strong hydrophobic interactions along with Leu254. Phe123 and Ile115 provide hydrophobic interactions along the BC loop. Close to the heme Ile386, Val382 and Leu496 form a triad of hydrophobic residues.

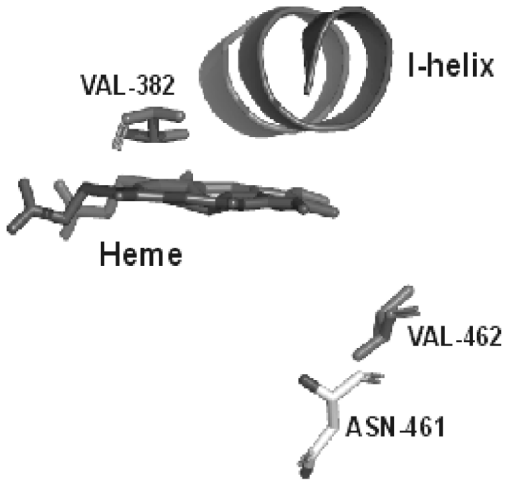

Fig. 9. Position of Asn-461 & Val-462 in M2 Polymorph of CYP1A1.

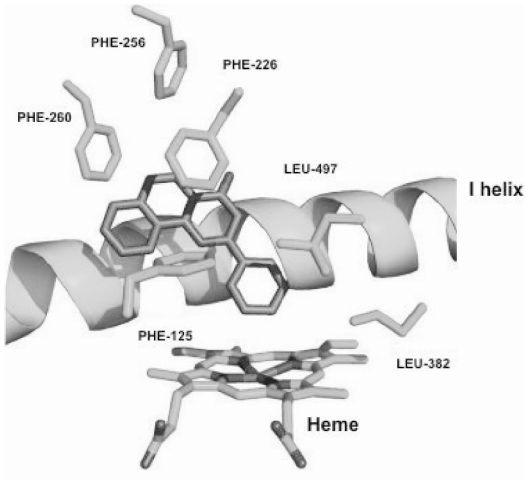

Fig. 10. ANF binding in CYP1A2 crystal structure 2HI4.

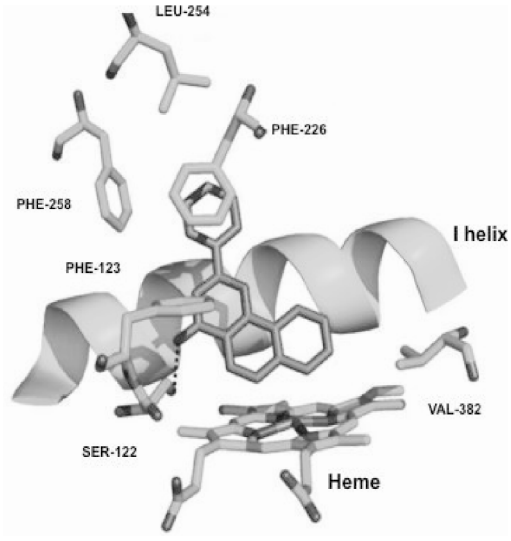

Fig. 11. ANF docking pose in the CYP1A1 model. Ser122 makes a hydrogen bond, (dashed line). Val382 (being shorter than Leu from 1A2) has more space for the triple ring of ANF to fit on top of the heme, thus placing positions 5 and 6 of ANF on top of the heme.

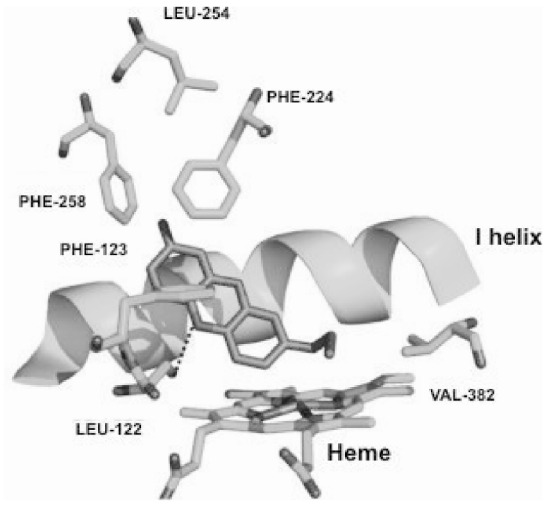

Fig. 12. Ethoxyresorufin docking into the CYP1A1 model. A possible hydrogen bond between Ser122 and Nitrogen is shown.

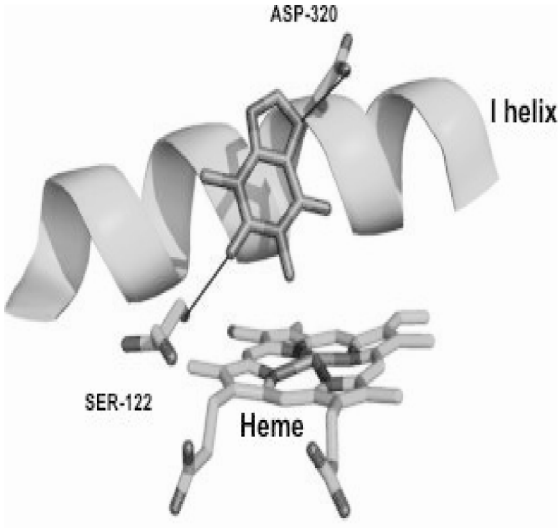

Fig. 13. Theophylline docking into CYP1A1 model in a good orientation for 1-de-methylation. Ser122 (BC loop) and Asp320 3(I helix) are involved in hydrogen bonding with Theophylline.

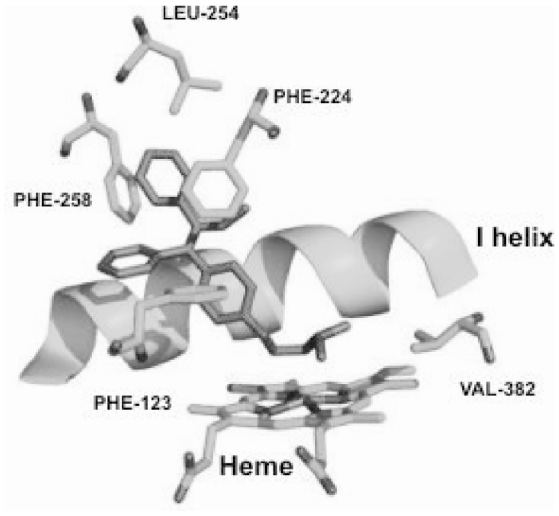

Fig. 14. Docking pose of tamoxifen into CYP1A1 model. Tamoxifen is in a good position for N-de-methylation.

Figures 15, 16 and 17 show the docking poses for some other ligands i.e. Ethanol, Phenacetin and Hesperetin (Rendic SS 2002). All three dock in good positions for metabolism to occur.

Extensive utilization of homology models to predict key amino acids that may have a role in interaction with the substrate can give an efficient throughput screening for the lead products. For Example, *In silico* docking Studies with CYP3A4 (Szklarz and Halpert, 1997; He YA *et.al.* 1997) showed important key residues which were then latterly confirmed by site-directed mutagenesis (He YA *et.al.* 1997; Domanski TL *et.al.* 1998). Our studies have identified some residues in the active site which are believed to be involved in substrate recognition and binding, including: Ile115, Ser122, Phe123, Phe224, Phe258, Tyr259, Asp313, Ala317, Asp320, Thr321, Val382, Ile386, Leu496 and Thr497. These residues provide an avenue to study CYP1A1 active site and its interaction with lead compounds.

Conclusion

A comparative model of CYP1A1 based on the structure of P450 1A2 built having 73% of sequence homology. Since, a profound comparative model always require optimum homologue template for defining its restraints, we presented a comparative model of CYP1A1 with the highest confidence in the accuracy of restraints in comparison to the earlier models which were modeled using the maximum limit of 36% homology. Such studies are commonly performed *in vitro* and *in silico* in the process of drug development to estimate the risk of drug interactions. Our CYP1A1 homology models give good insight into the CYP1A1 active site and its interaction with various substrates. Our studies have identified some residues in the active site which are believed to be involved in substrate recognition and binding, including: Ile115, Ser122, Phe123, Phe224, Phe258, Tyr259, Asp313, Ala317, Asp320, Thr321, Val382, Ile386, Leu496 and Thr497. These residues are identical to the residues identified in other studies (Szklarz GD and Paulsen MD 2002; Iori F *et.al.* 2005; Lewis BC *et.al.* 2007). The models also predict sites of metabolism for some of the ligands and binding positions of some ligands like Ethoxyresorufin which are similar to those predicted in other studies (Lewis BC *et.al.* 2007). Our model suggests the role of mutations to be ineffective for metabolisation of substrates and reconfirm the observations of Persson and collegues (Persson I *et.al.* 1997).The search for CYP1A1 inhibitors together with *in silico* models developed herein may provide useful information for the development of chemoprotectors and wet lab conformational studies.

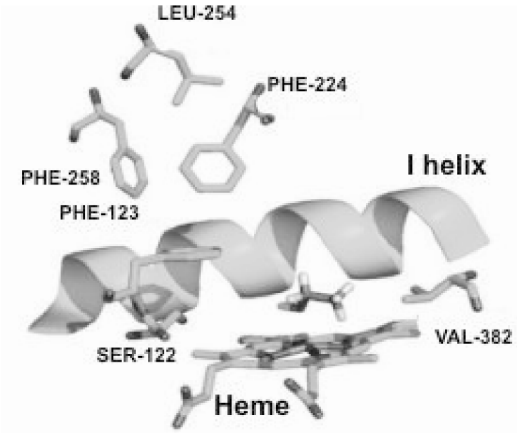

Fig. 15. Ethanol docking pose into CYP1A1 model. It is in a good position for oxidition to Acetaldehyde. The Hydrogens on Ethanol are shown for clarity.

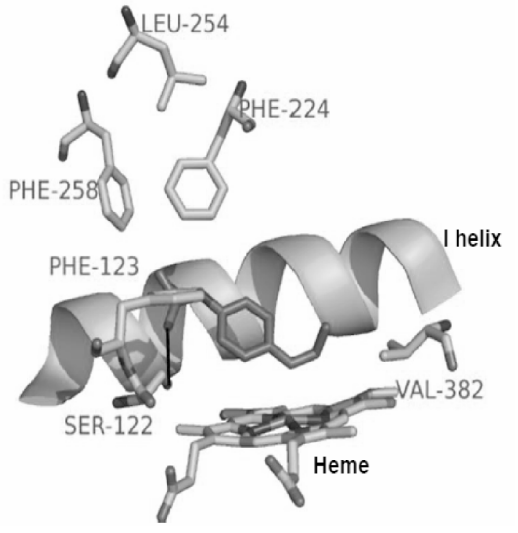

Fig. 16. Phenacetin docking pose in the CYP1A1 model. The oxygen is positioned on top of the heme for O de-ethylation to form Paracetamol. Ser122 forms a hydrogen bond and Val382 has hydrophobic interactions with the ethyl group to position the oxygen on top of the heme.

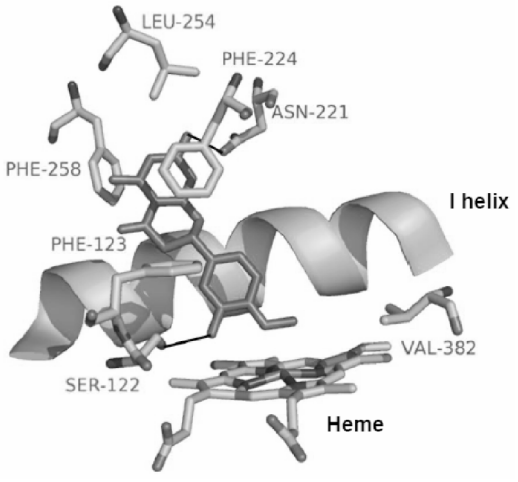

Fig. 17. Hesperetin docking pose in the CYP1A1 model. Val382 provides hydrophobic interactions with the methyl group, Ser122 and Asn221 make hydrogen bonds while Phe224 and Phe258 provide pi stacking interactions. These interactions place Hesperetin on top of the heme for O demethylation to occour.

REFERENCES

1. Abecasis GR and Cookson WO. 2000. GOLD--graphical overview of linkage disequilibrium. Bioinformatics, 16(2):182-183.
2. Barker CW, Pagan JB and Pasco DS. 1992. Interleukin-1β suppresses the induction of P450 1A1 and P450 1A2 mRNAs in isolated hepatocytes. Journal of Biological Chemistry, 267: 8050–8055
3. Beresford AP. 1993. CYP1A1: friend or foe? Drug Metabolism Reviews, 25 (4): 503–517.
4. Burke M and Mayer R. 1974. Ethoxyresorufin: direct fluometric assay of a microsomal O-dealkylation which is preferentially inducible by 3-methylchloranthrene. Drug Metabolism and Disposition, 2:583–588.
5. Casper RF, Quesne M, Rogers IM, Shirota T, Jolivet A, Milgrom E and Savouret JF .1999. Resveratrol has antagonist activity on the aryl hydrocarbon receptor: implications for prevention of dioxin toxicity. Molecular Pharmacology, 56:784-790.
6. Colovos C and Yeates TO.1993. Verification of protein structures — Patterns of nonbonded atomic interactions. Prot. Sci. 2, 1511–1519.
7. Crewe HK, Notley LM, Wunsch RM, Lennard MS, and Gillam EMJ. 2002. Metabolism of Tamoxifen by Recombinant Human Cytochrome P450 Enzymes: Formation of the 4-Hydroxy, 4'-Hydroxy and N-Desmethyl Metabolites and Isomerization of trans-4-Hydroxytamoxifen.Drug Metabolism and Disposition, 30(8): 869 - 874.
8. DeLano WL.2002. The PyMOL Molecular Graphics System. DeLano Scientific San Carlos, CA, USA. http://www.pymol.org. Accessed on 12-01-2010.
9. Domanski TL, Liu J, Harlow GR and Halpert JR. 1998. Analysis of four residues within substrate recognition site 4 of human cytochrome P450 3A4: role in steroid hydroxylase activity and α-naphthoflavone stimulation. Achieve of Biochemistry and Biophysics, 350:223-232.
10. Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, Pieper U, Sali A. 2006. Comparative Protein Structure Modeling With MODELLER. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30: 200.
11. Gotoh .1992. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences, Journal of Biological Chemistry, 267: 83–90.
12. Guengerich FP. 1995. Human cytochrome P450 enzymes, in Cytochrome P450: Structure, Mechanism and Biochemistry. Plenum Press, New York, 473–535.
13. Guex N and Peitsch MC.1996. Swiss-PdbViewer: A Fast and Easy-to-use PDB Viewer for Macintosh and PC. Protein Data Bank Quaterly Newsletter, 77: 7.
14. Hammons GJ, Milton D, Stepps K, Guengerich FP, Tukey RH, Kadlubar FF.1997. Metabolism of carcinogenic heterocyclic and aromatic amines by recombinant human cytochrome P450 enzymes. Carcinogenesis, 18: 851–854.

15. He YA, He YQ, Szklarz GD and Halpert JR. 1997. Identification of three key residues in substrate recognition site 5 of human cytochrome P450 3A4 by cassette and site-directed mutagenesis. Biochemistry, 36:8831-8839.
16. Hildebrand DC, Palleroni NJ, Hendson M, Toth J and Johnson JL. 1994. Pseudomonas avescens isolated from walnut blight cankers. Systematic Bacteriology, 44: 410–415.
17. Houlston RS. 2000. CYP1A1 polymorphisms and lung cancer risk: a meta-analysis. Pharmacogenetics, 10:105–114.
18. Iori F, da Fonseca R, João Ramos M and Menziani MC. 2005. Theoretical quantitative structure-activity relationships of flavone ligands interacting with cytochrome P450 1A1 and 1A2 isozymes. Bioorganic and Medicinal Chemistry, 13(14): 4366-4374.
19. Ingelman-Sundberg M, Oscarson M, Daly AK, Garte S, Nebert DW. 2001. Human cytochrome P-450 (CYP) genes: a web page for the nomenclature of alelles. Cancer Epidemiology, Biomarkers and Prevention, 10: 1307-1308.
20. Kawajiri K and Hayashi SI. 1996. The CYP1 family: Cytochromes P450-Metabolic and Toxicological Aspects. CRC Press, Boca Raton, Florida, 77–97.
21. Kawajiri K,Nakachi K,Imai K,Yoshii A,Shinoda N,Watanabe J. 1990. Identification of genetically high risk individuals to lung cancer by DNA polymorphisms of the cytochrome P450IA1 gene. FEBS Lett, 263: 131–133.
22. Kelley LA, Gardner SP and Sutcliffe MJ. 1996. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. Protein Engineering, 9: 1063-1065.
23. Lewis BC, Mackenzie PI, and Miners JO. 2007. Comparative homology modeling of human cytochrome P4501A1 (CYP1A1) and confirmation of residues involved in 7-ethoxyresorufin O-deethylation by site-directed mutagenesis and enzyme kinetic analysis. Archives of Biochemistry and Biophysics, 468(1):58-69.
24. Medha B and Aravind L.2007. PSI-BLAST Tutorial. Comparative Genomics. Methods in Molecular Biology, 395: 177-186.
25. Moreno M, Sarti E, Herrera N, Ramirez J, Robles V and Tapia R. 2008. Steroid metabolism gene CYP17, CYP1A1, 2B, CYP1A1, 2C and risk of breast cancer in Mexican women. Journal of Clinical Oncology, ASCO Annual Meeting Proceedings, 26(15S): 1524
26. Morel Y and Barouki R.1998. Down-regulation of cytochrome P450 1A1 gene promoter by oxidative stress: Critical contribution of nuclear factor 1. Journal of Biological Chemistry, 273:26969–26976.
27. Molecular Operating Environment - MOE. 2008, Chemical Computing Group Inc.
28. Nebert DW. 1991. Role of genetics and drug metabolism in human cancer risk. Mutation Research, 247(2):267-281.
29. Nelson DR, Zeldin DC, Hoffman SM, Maltais LJ, Wain HM and Nebert DW. 2004. Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. Pharmacogenetics, 14(1):1-18.
30. O'Sullivan O. 2004. Combining Protein Sequences and Structures within Multiple Sequence Alignments. Journal of Molecular Biology, 340: 385-395.
31. Persson I, Johansson I, Ingelman SM. 1997. In vitro kinetics of two human CYP1A1 variant enzymes suggested to be associated with interindividual differences in cancer susceptibility. Biochemical and Biophysical Research Communications, 231: 227–30.
32. Ramachandran GGN, Ramakrishnan CC, and Sasisekharan VV.1963. Stereochemistry of polypeptide chain configurations. Journal of Molecular Biology, 7: 95-99.
33. Rendic S and Di Carlo FJ. 1997. Human cytochrome P450 enzymes: A status report summarizing their reactions, substrates, inducers, and Inhibitors. Drug Metabolism Reviews, 29: 413-580.
34. Roman A. Laskowski, Malcolm W. MacArthur, David S. Moss and Janet M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography, 26, 283-291.
35. Rowland P, Blaney FE, Smyth MG, Jones JJ, Leydon VR, Oxbrow AK, Lewis CJ, Tennant MG,
36. Modi S, Eggleston DS, Chenery RJ and Bridges AM .2006. Crystal Structure of Human Cytochrome P450 2D6. Journal of Biological chemistry, 281:7614-7622.
37. Rendic S. 2002, Summary of information on human CYP enzymes: human P450 metabolism data. Drug Metabolism Review, 34: 83-448.
38. Sansen S, Yano JK, Reynald RL, Schoch GA, Griffin KJ, Stout CD and Johnson EF. 2007. Adaptations for the Oxidation of Polycyclic Aromatic Hydrocarbons Exhibited by the Structure of Human P450 1A2. Journal of Biological chemistry, 282: 14348-14355.
39. Schoch GA, Yano JK, Wester MR, Griffin KJ, Stout CD and Johnson EF. 2004. Biological Chemistry, 279: 9497–9503.
40. Scott EE, Wester MR, White M A, Chin CC and Stout CD. 2003. Proceedings of the National Academy of Sciences, USA, 100: 13196–13201.
41. Shimada, T, Oda Y, Gillam EM, Guengerich FP and Inoue K. 2001. Metabolic activation of polycyclic aromatic hydrocarbons and other procarcinogens by cytochromes P450 1A1 and P450 1B1 allelic variants and other human cytochromes P450 in Salmonella typhimurium NM2009. Drug Metabolism and Disposition, 29: 1176–1182.
42. Shou M, Gonzalez FJ and Gelboin HV. 1996. Stereoselective epoxidation and hydration at the K-region of polycyclic aromatic hydrocarbons by cDNA-expressed cytochromes P450 1A1, 1A2, and epoxide hydrolase. Biochemistry, 35:15807–15813.
43. Smart J, Daly AK. 2000. Variation in induced CYP1A1 levels: relationship to CYP1A1, Ah receptor and GSTM1 polymorphisms. Pharmacogenetics, 10: 11–24.

44. Swiss-Prot Protein knowledgebase, Swiss Institute of Bioinformatics, (http://www.expasy.org/sprot) accessed on 24-12-2009.
45. Szklarz GD and Halpert JR. 1997. Molecular modeling of cytochrome P450 3A4. Journal of Computer Aided Molecular Design, 11: 265-272.
46. Szklarz GD and Paulsen MD. 2002. Molecular modeling of cytochrome P450 1A1: enzymesubstrate interactions and substrate binding affinities. Biomolecular structural Dynamics, 20:155–162.
47. Szklarz GD, Halpert JR.1997b. Molecular modeling of cytochrome P450 3A4. Journal of Computer-Aided Molecular Design, 11:265–272.
48. Thompson JJD, Thierry JJC, and Poch OO. 2003. RASCAL: Rapid scanning and correction of multiple sequence alignments. Bioinformatics, 19: 1155.
49. Turesky RJ, Constable A, Richoz J, Varga N, Markovic J, Martin MV, Guengerich FP.1998. Activation of heterocyclic aromatic amines by rat and human liver microsomes and by purified rat and human cytochrome P450 1A2. Chemical Research in Toxicology, 11: 925–936.
50. Wade RC, Motiejunas D, Schleinkofer K, Sudarko, Winn PJ, Banerjee A, Kaniakin A and Jung C. 2005. Multiple molecular recognition mechanisms. Cytochrome P450—a case study. International Journal of Biochemistry, Biophysics and Molecular Biology, 1754:239–244.
51. Watanabe M. 1998. Polymorphic CYP genes and disease predisposition. Toxicology Letters, 102-103: 167–71.
52. Wester MR, Johnson EF, Marques-Soares C, Dijols S, Dansette PM, Mansuy D, Stout CD. 2003b. Structure of mammalian cytochrome P450 2C5 complexed with diclofenac at 2.1 Å resolution: evidence for an induced fit model of substrate binding. Biochemistry, 42:9335–9345.
53. Yannick M and Robert B. 2000. The repression of nuclear factor I/CCAAT transcription factor (NFI/CTF) transactivating domain by oxidative stress is mediated by a critical cysteine (Cys-427). Biochemistry, 348: 235–240.
54. Yano JK, Hsu MH, Griffin KJ, Stout CD, and Johnson EF. 2005. Structures of human microsomal cytochrome P450 2A6 complexed With coumarin and methoxsalen. Nature Structural And Molecular Biology, 12: 822–823.
55. Zhou H, Josephy PD, Kim D and Guengerich FP. 2004. Functional characterization of four allelic variants of human cytochrome P450 1A2. Archives of Biochemistry and Biophysics, 422: 23–30.

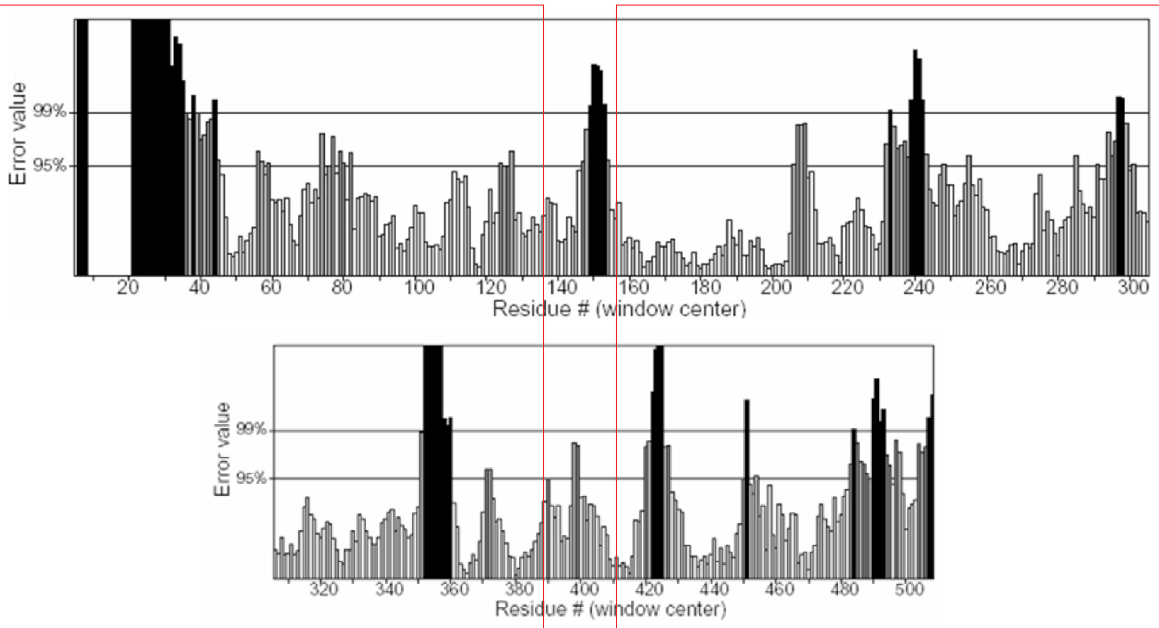

Fig. 1. ERRAT plot for CYP1A1. On the error axis (y axis), two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value. The regions with high error value (black lines) indicate areas of high energy or clashes in the structure.

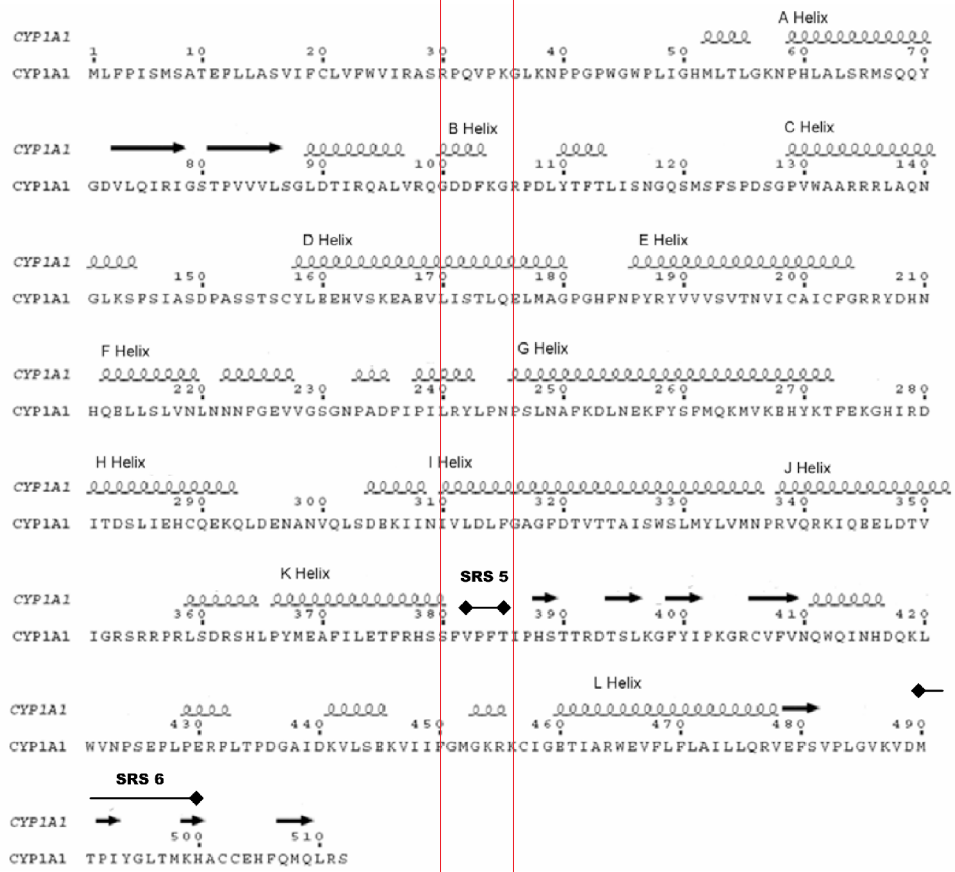

Fig. 2. Amino Acid sequence of CYP1A1 with the sheets labeled as arrows; the major helices are named and labeled as coils.

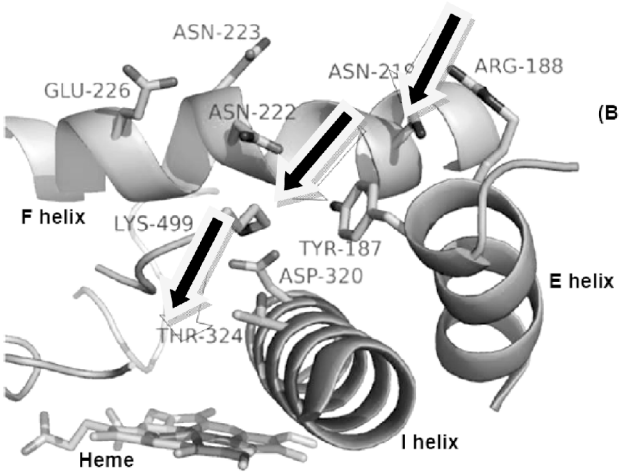

Fig. 3. Possible entry channel into the active site of CYP1A1 marked by arrows, surrounded by residues from the F helix, I helix and E helix.

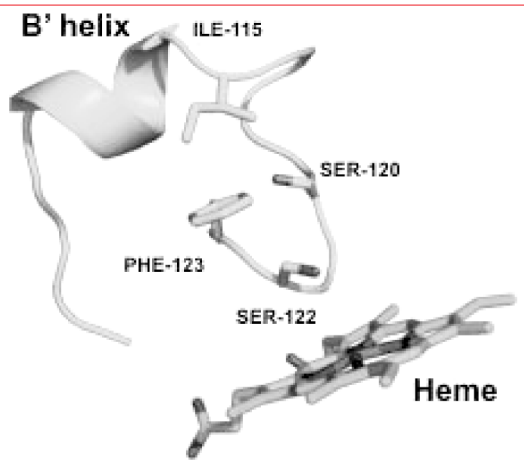

Fig. 4. Residues in the BC loop

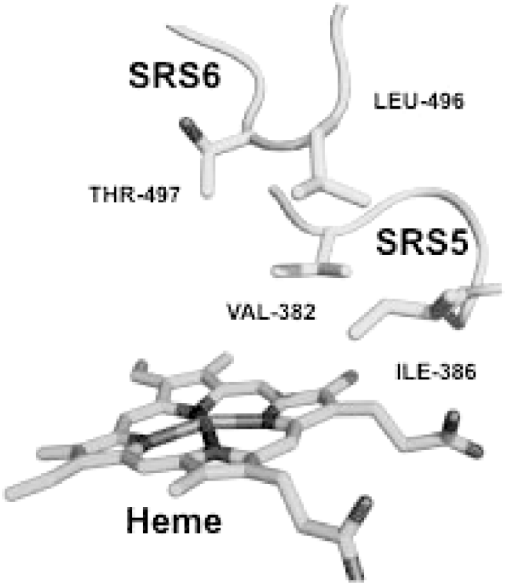

Fig. 5. Residues of I helix are oriented towards the active site in CYP1A1.

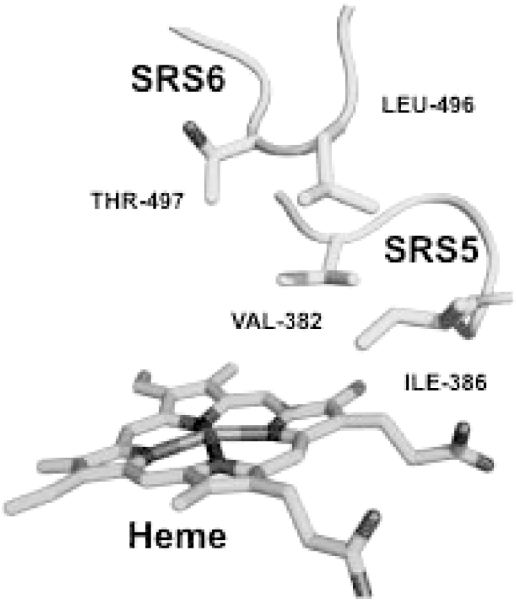

Fig. 6. Val382 and Ile386 (SRS5); Thr497 and Leu496 (SRS6) for CYP1A1 are shown.

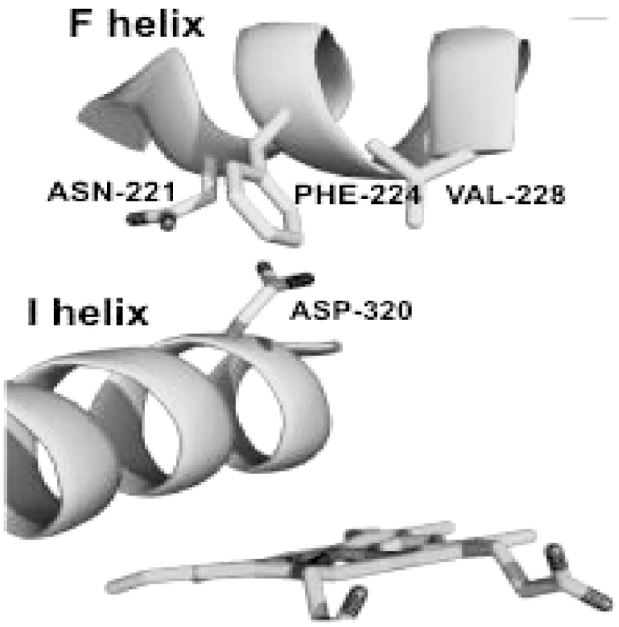

Fig. 7. Residues from F helix oriented towards the active site in CYP1A1. Asp320 (I helix) is also visible.

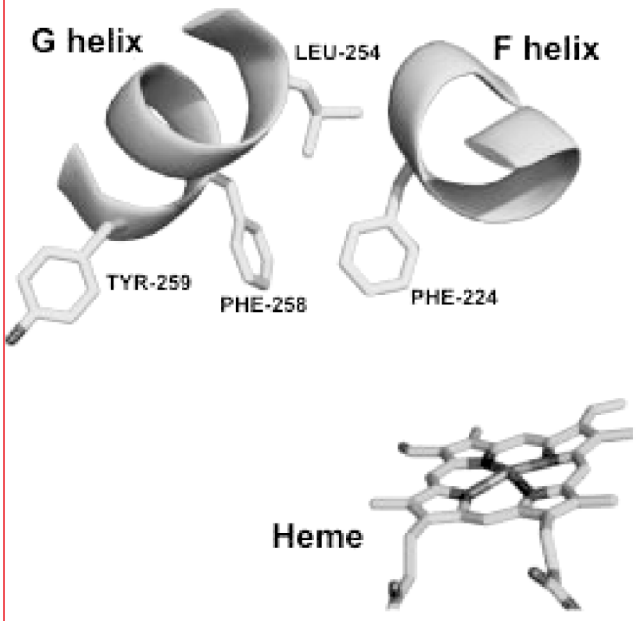

Fig. 8. Residues of G helix from CYP1A1 oriented into the active site. Phe224 from F helix is shown as well

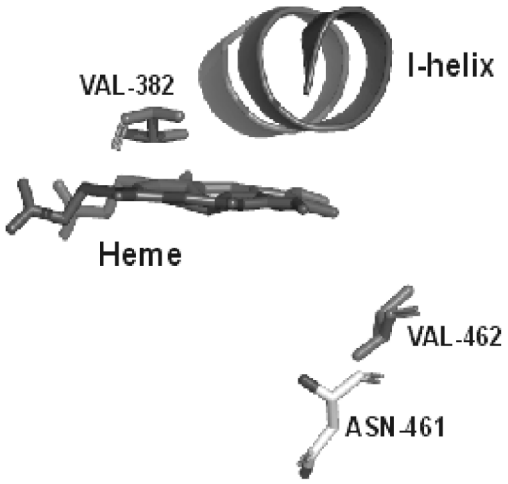

Fig. 9. Position of Asn-461 & Val-462 in M2 Polymorph of CYP1A1.

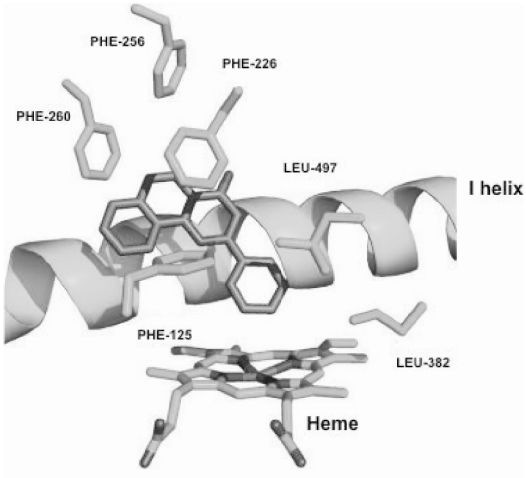

Fig. 10. ANF binding in CYP1A2 crystal structure 2HI4.

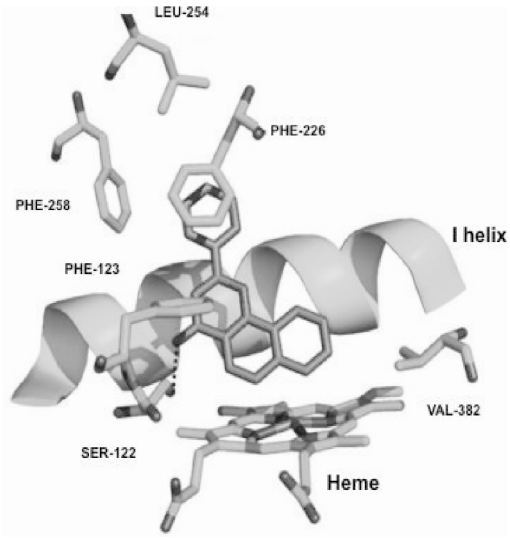

Fig. 11. ANF docking pose in the CYP1A1 model. Ser122 makes a hydrogen bond, (dashed line). Val382 (being shorter than Leu from 1A2) has more space for the triple ring of ANF to fit on top of the heme, thus placing positions 5 and 6 of ANF on top of the heme.

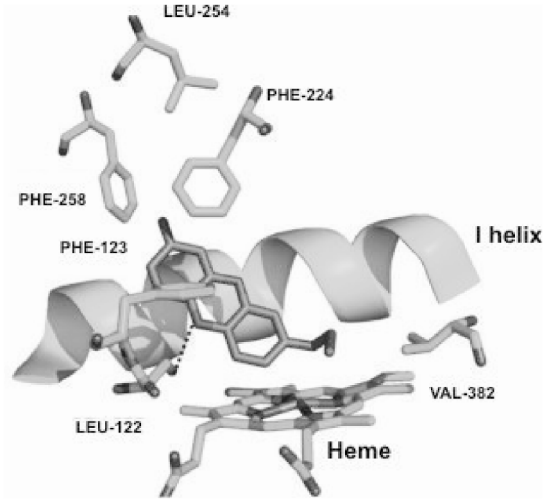

Fig. 12. Ethoxyresorufin docking into the CYP1A1 model. A possible hydrogen bond between Ser122 and Nitrogen is shown.

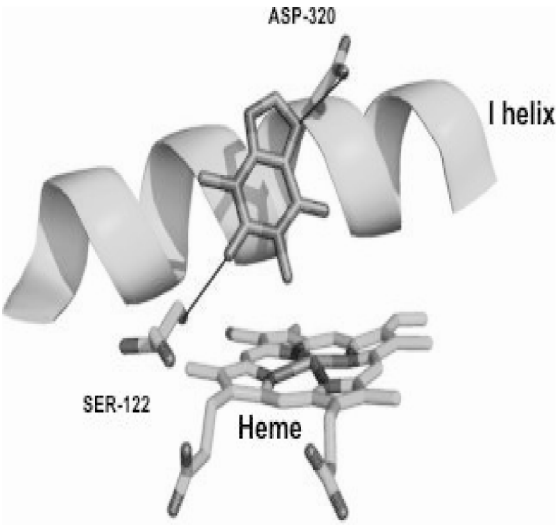

Fig. 13. Theophylline docking into CYP1A1 model in a good orientation for 1-de-methylation. Ser122 (BC loop) and Asp320 3(I helix) are involved in hydrogen bonding with Theophylline.

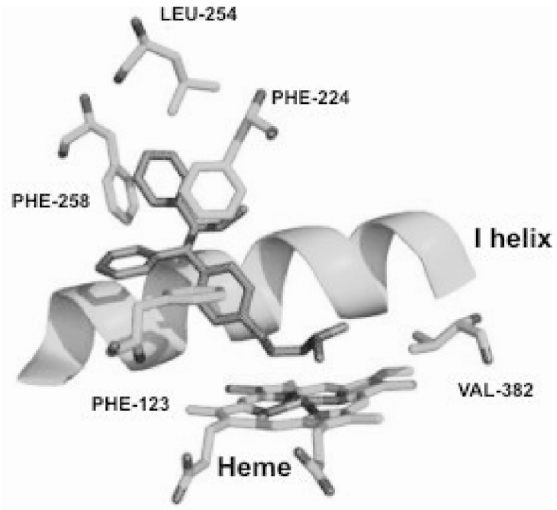

Fig. 14. Docking pose of tamoxifen into CYP1A1 model. Tamoxifen is in a good position for N-de-methylation.

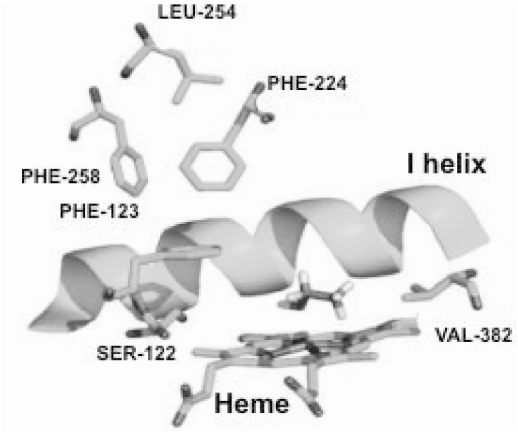

Fig. 15. Ethanol docking pose into CYP1A1 model. It is in a good position for oxidition to Acetaldehyde. The Hydrogens on Ethanol are shown for clarity.

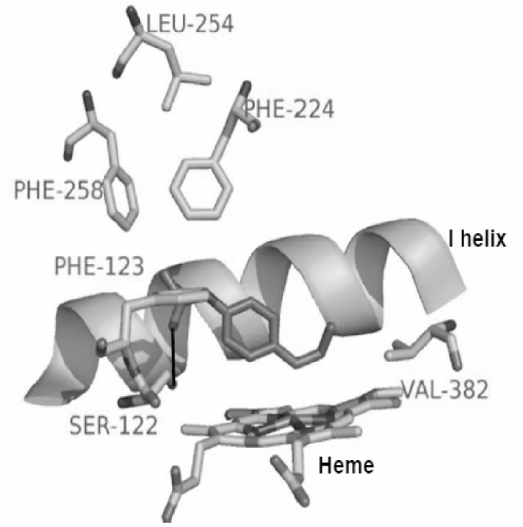

Fig. 16. Phenacetin docking pose in the CYP1A1 model. The oxygen is positioned on top of the heme for O de-ethylation to form Paracetamol. Ser122 forms a hydrogen bond and Val382 has hydrophobic interactions with the ethyl group to position the oxygen on top of the heme.

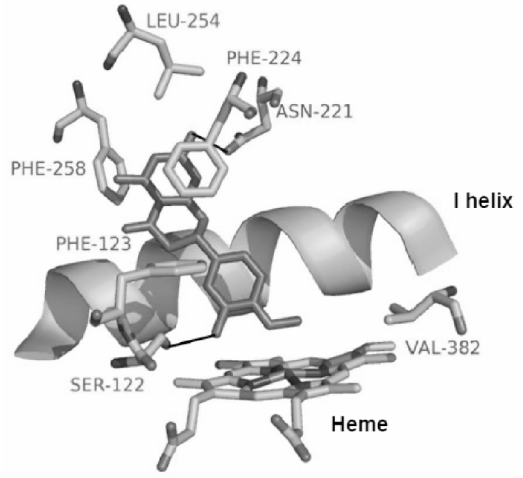

Fig. 17. Hesperetin docking pose in the CYP1A1 model. Val382 provides hydrophobic interactions with the methyl group, Ser122 and Asn221 make hydrogen bonds while Phe224 and Phe258 provide pi stacking interactions. These interactions place Hesperetin on top of the heme for O demethylation to occour.