

Discerning Relationships Among Human Cytochrome P450s by Computational Analyses

*Alok Juneja, Kulvinder Singh Saini and Rajeev Soni**

Department of Molecular Biotechnology, New Drug Discovery Research,
Ranbaxy Research Laboratories, Plot No. 20, Sector 18, Udyog Vihar
Industrial Area, Gurgaon, Haryana 122001, INDIA

Abstract

Cytochrome P450s are involved in the biotransformations and metabolism of wide variety of xenobiotics and are essential players in the metabolism of pharmaceuticals and their role in drug-drug interactions. Here the authors have attempted to reveal their structure-function and evolutionary relationships of six major human P450s (CYP3A4, CYP2D6, CYP2C9, CYP2C19, CYP1A2 and CYP2E1) using some of the current computational tools available. These six isoforms are responsible for metabolism of over 90% of drugs and xenobiotics in humans. The computational analyses performed corroborate literature findings and, in addition, have revealed information that could be exploited to explain the functional differences between these enzymes.

Keywords

CYP3 A4, Cytocromes, JEMBOSS, Crystal structures.

* To whom correspondence should be addressed.

Introduction

Cytochrome P450s have been classified as a family of enzymes depending on resemblance in their structure and function. These enzymes play a key role in phase I drug metabolism, display substrate/inhibitor selectivity and cause drug-drug interactions. In addition, existence of single-nucleotide polymorphisms (SNPs) among various P450s would lead to inter-individual variability in the metabolism and pharmacokinetic properties of drugs and their substrates.

Homology models based upon the crystal structures (both earlier available bacterial P450s, and recently available mammalian P450) for human CYP subtypes give a valuable indication of their structure-function relation (Ridderstrom *et al.* 2001; Danielson 2002; Lewis 2003; Venhorst *et al.* 2003). Relationship across different subfamilies of human P450s is very scarce, and cannot easily be portrayed and generalized. A high degree of protein sequence homology is insufficient to explain these subtle differences in the activity of these enzymes. Several authors (Lewis *et al.* 1998; Danielson 2002; Lewis 2003) published information related to structural and pharmacophoric features for classifying substrates and inhibitors for these enzymes, while any evidence that elucidates the functional relationship among human P450 enzymes would open the doors for novel research areas. In this context, tools from bioinformatics may provide some insights and would lead to rational drug design.

In this paper the authors have tried to evaluate the relationships among six major human CYPs, based on their protein sequence alone. Our results provide further insights about these CYPs and point to several questions that remain unanswered.

Materials and Methods

Computational Tools used in this study are given below. Following computational modules were used with default settings unless specified otherwise.

Jembooss (Java European Molecular Biology Open Software Suite)

<http://www.hgmp.mrc.ac.uk/Software/EMBOSS/Jembooss/>

- (1) Emma Multiple alignment program - interface to ClustalW program
- (2) Prettyplot Displays aligned sequences, with colouring and boxing
- (3) Distmat Creates a distance matrix from multiple alignments
- (4) Plotcon Plots the quality of conservation of a sequence alignment
- (5) Pepstats Protein statistics
- (6) showalign Displays a multiple sequence alignment

MEME (Multiple Expectation maximization for Motif Elicitation) (Bailey and Elkan 1994)

MAST (Motif alignment and search tool) (Bailey and Gribskov 1998)

Phylip (Phylogeny Interface Package);

<http://evolution.genetics.washington.edu/phylip.html>

Results and Discussion

Multiple Sequence Alignment Analysis

Fig. 1 shows the sequences of six human CYPs (CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP2E1, and CYP3A4) aligned using **JEMBOSS** module *Emma* and *PrettyPlot*. The scores of PSA (Pairwise Sequence Alignment) analyses suggest that the closest sequences amongst all the six are CYP2C9 and CYP2C19 (aligned score: 91), followed by CYP2C9 and CYP2E1 (aligned score: 56), CYP2C9 and CYP2D6 (aligned score: 33), CYP2C9 and CYP1A2 (aligned score: 11), CYP2C9 and CYP3A4 (aligned score: 8). These results are corroborated by **PHYLIP** and *DistMat*, which

suggest that all the CYPs are evolutionary related to each other with a varying degree of the similarity, as evident from the PSA score and the evolutionary distance calculated by DistMat module in the matrix format (data not shown). Deciphering of sequence-structure-function relation in any of these sequences would be beneficial in predicting relationship within CYPs.

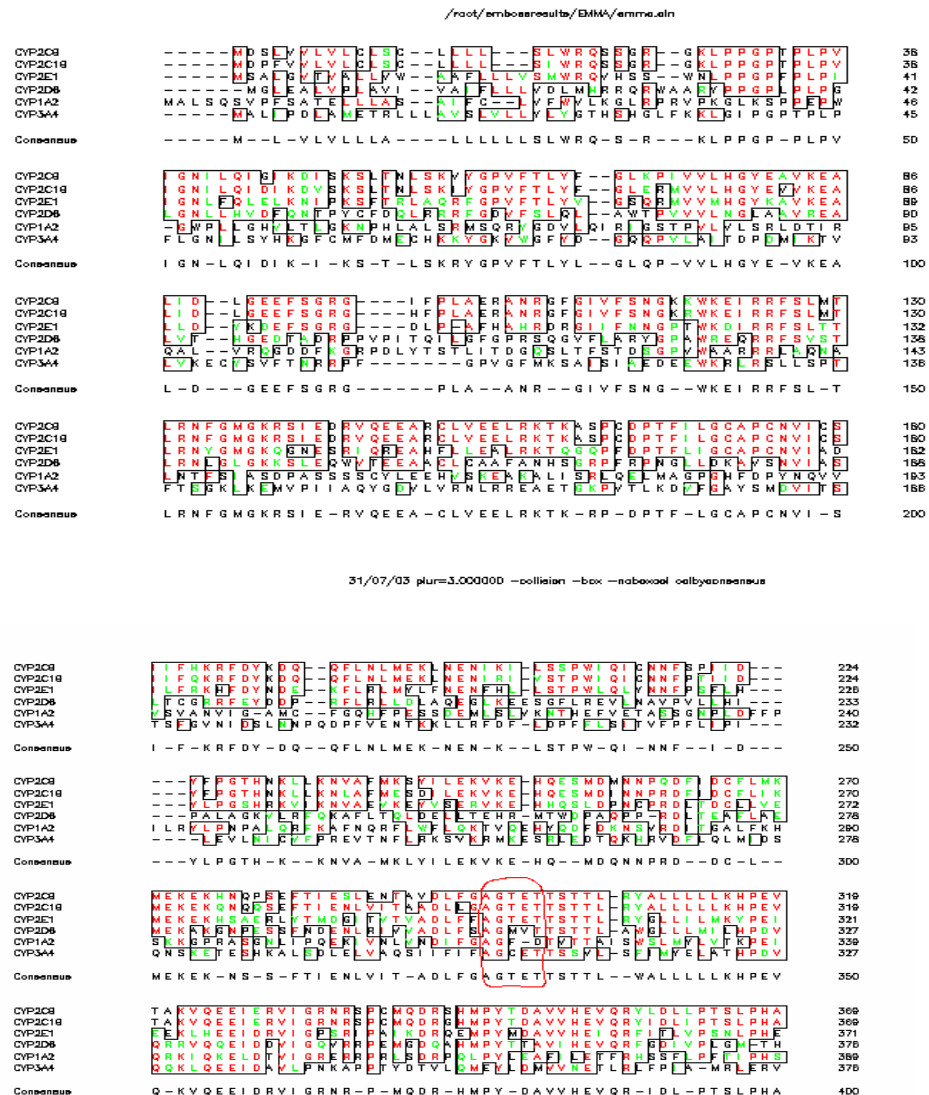


Fig. 1: MSA along with the Consensus sequence. The result is generated using 'PrettyPlot' module on the results of MSA obtained from 'Emma'. The Identical residues are shown in RED, Similar residues are in GREEN, and the other residues are in **BLACK**.

Motif Similarity and Dissimilarity

Motifs in the six CYPs were identified using *MEME* and interpreted using *MAST*. (Fig. 2). The program was prompted to find 8 motifs across these six CYPs. The result was 8 distinct motifs with cytochromes 2C9, 2C19 and 2E1 possessing all the 8 motifs whereas CYP2D6 lacking motif number 8. CYP3A4 and CYP1A2 lack motif 3, 6 and 8. This suggests that cytochromes 2C19, 2C9, 2E1 and 2D6 are similar to and different than CYP3A4 and CYP1A2, and this can probably be attributed to the absence of motif 3, 6 and 8 in CYP3A4 and CYP1A2. The literature on CYP3A4 (Anzenbacher and Anzenbacherova 2001) reveals that the active site constituent residue sequence from 360-375 (VNETLRLFP IAMRLER) is distributed in motif 1 from 360-365 and last residue i.e. 375 in motif 7; remaining residues from 366-374 are not the part of any of the motifs. Along with this the residues 210-211 are also not present in any motifs. It could mean that these non-motif sub-sequence differentiates the CYP3A4 from other CYPs.

The cytochromes 2C9, 2C19, 2E1 and 2D6 can be placed under one roof of functionality depending upon the presence of motifs. However, they differ in their substrate specificity. Two scenarios might exist: In the first case, active site constituent residues are not the part of any motif meaning different amino acids contributing to the respective active sites in the tertiary conformation. In the second case, some of the active site residues are the part of the different motifs, hence the orientation of these residues in a particular manner make them substrate specific. In both the cases, the active site differs from one another making them substrate specific, even if the enzymes are similar in functionality.

The conserved motifs represent different functionalities that the CYPs possess. Motif 2 with consensus FXXGXXXCXG at the carboxy-terminus of all the CYPs (encircled blue in Fig.1), forms the heme binding decapeptide

loop. The cysteine is most essential as this residue coordinates as the fifth ligand of the heme (Nebert *et al.* 1988; Danielson 2002). The well-conserved motif 4 having the consensus sequence (G/A)GX(D/E)T, near the center of protein (encircled red in Fig.1), has a terminal threonine that constitutes to the oxygen-binding pocket (Poulos *et al.*, 1987; Danielson 2002).

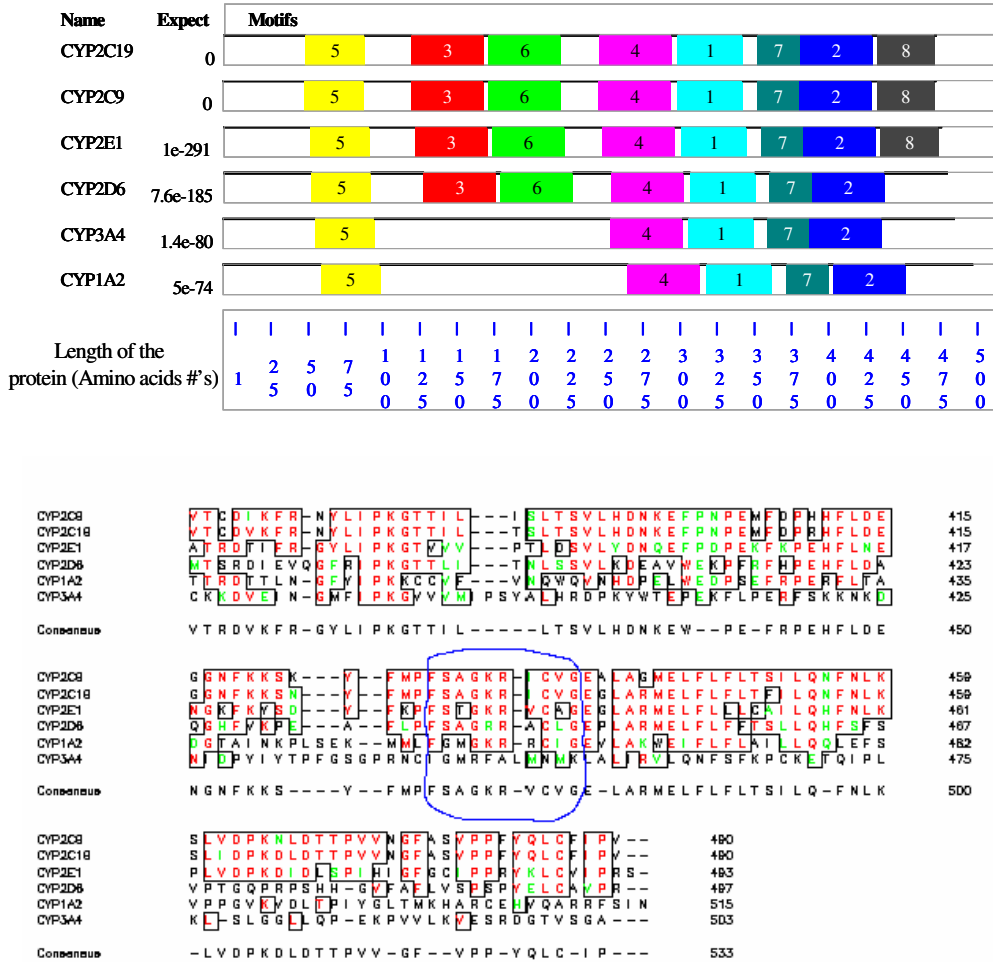


Fig. 2: Motifs found in six CYPs using MEME and MAST

CYPs Statistics

The result of *PepStats* module of Jemboss indicates that all the CYPs of interest have high Non Polar residues (1A2: 54.95%, 2C19: 55.30%, 2C9:

55.71%, 2D6: 59.35%, 2E1: 56.59% and 3A4: 56.85%). In addition, Leucine is the most favored amino acid in every CYP. 1A2 (L w.r.t. NP: 21.20%, L w.r.t. FL:11.65%), 2C19 (L w.r.t. NP:20.66%, L w.r.t. FL:11.42%), 2C9 (L w.r.t. NP:21.61%, L w.r.t. FL:12.04%), 2D6 (L w.r.t. NP:22.37%, L w.r.t. FL:13.27), 2E1 (L w.r.t. NP:21.14%, L w.r.t. FL:11.96%), 3A4 (L w.r.t. NP:20.62%, L w.r.t. FL:11.72%) where L = Leucine, NP = Non-Polar amino acid residues, FL = Full Length (in amino acids) of the particular CYP450.

An interesting question that arises from our study is the presence of Leucine in majority over the other and non-polar amino acids. The iso-electric point and charge value of all the six CYPs are different which might be playing a role in the functionality of CYPs.

Although human CYP2C9 and CYP2C19 are the closest CYPs, they are different in terms of the substrates they metabolize. The Isoelectric points of these two CYPs are similar while the charge value is different for CYP2C9 (5.5) and CYP2C19 (2.03). This might account for the differences between these enzymes (Smith *et al.* 1997; Ridderstrom 2001; de Groot, M.J. *et al.* 2002; Williams *et al.* 2003).

Active Site and 3D Structure Prediction

P450s active sites vary in their stability and flexibility. Spectroscopic experiments have revealed that CYP1A2 possesses the most stable conformation and inflexible active site, which can accommodate planar molecules of moderate volume and basicity (Anzenbacher *et al.* 1998a), whereas CYP3A4 has one that is flexible and denatures easily (Anzenbacher 2001). A review (Smith *et al.* 1997; Lewis *et al.* 1998; Anzenbacher 2001; Ekins *et al.*, 2001; Ridderstrom *et al.* 2001; Danielson 2002; de Groot and Ekins 2002; Lewis 2003) of the literature reveals that the active sites of all the CYPs are different and their properties may be governed by the nature of substrates they act upon. Does that mean the user can swap the active sites

with one another, thereby resulting in a CYP with the desired function (Shimoji *et al.* 1998)?

Crystal structures of P450s have further assisted the modeling of other forms via homology modeling. With this approach, significant numbers of human P450s have been modeled (Lewis *et al.* 1998). The membrane bound nature of CYP450s has slowed down the elucidation of their crystal structure. Recent elucidation of CYP2C9 crystal structural (Mancy *et al.* 1995; Jones *et al.* 1996; Ridderstrom *et al.* 2001; de Groot *et al.* 2002; Williams *et al.* 2003) has revealed so much information about the CYP2C19, which is evolutionary related to CYP2C9 to a greater extent. Even after possessing close similarity, the difference in the substrate selectivity of these enzymes could be attributed to the nature of the amino acids in their active site. CYP2C9 with large active site assisted by lack of significant conformational changes offers binding of additional small molecules to the active site (Williams *et al.* 2003).

Predictions and Unresolved Issues

Motif analysis from our study suggests that cytochromes 2C9, 2C19, 2D6, and 2E1 are similar to a great extent and are different than CYP1A2, and CYP3A4, and this could be due to the absence of motifs 3, 6 and 8. Certain points that are still untouched or not been exploited so far, are: Why Leucine and non-polar amino acids are most favored in CYPs; is it of evolutionary significance or is there a functional correlation? Is it possible for us to swap the active sites of one CYP with the other and observe a change in the functionality of the CYP?

Conclusions

The authors observed using the current computational tools available some interesting similarities and differences among the chosen six human CYP enzymes. Data analyses have generated some unanswered queries in understanding the sequence—function relationship that requires further

research. Since these enzymes are involved in the metabolism of majority of the marketed drugs, wet lab research augmented with computational approach is essential to understand fully the sequence-structure-function relationship of CYPs. This could help us not only in further elucidation of the CYPs in a better way, but also enhance modern day drug discovery.

Acknowledgements

The authors wish to thank Dr. Sreedhara R. Voleti for his valuable comments and critical reading of the manuscript.

References

Anzenbacher, P., Anzenbacherova, E. 2001. Cytochrome p450 and metabolism of xenobiotics. *Cell.Mol.Life Sci.*, 58: 737-747.

Anzenbacher, P., Bec, N., Hudecek, J., Lange, R., Anzenbacherova, E. 1998a. High conformational stability of cytochrome P450 1A2: evidence from UV absorption spectra. *Collection of Czechoslovak Chemical Communications*, 63: 441-448.

Bailey, T.L., Elkan, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, Menlo Park, California, AAAI Press.

Bailey, T.L., Gribskov, M. 1998. Combining evidence using p-values: application to sequence homology searches. *Bioinformatics*, 14: 48-54.

Danielson, P.B. 2002. The Cytochrome P450 Superfamily: Biochemistry, Evolution and Drug Metabolism in Humans. *Current Drug Metabolism*, 3: 561-597.

De Groot, M.J., Alex, A.A., Jones, B.C. 2002. Development of a combined protein and pharmacophore model for cytochrome P450 2C9. *J. Med. Chem.*, 45: 1983-1993.

De Groot, M.J., Ekins, S. 2002. Pharmacophore modeling of cytochromes P450. *Advanced Drug Delivery Reviews*, 54: 367-383.

Ekins, S., de Groot, M.J., Jones, J.P. 2001. Pharmacophore and three-dimensional quantitative structure activity relationship methods for modeling cytochrome p450 active sites. *Drug Metabolism and Disposition*, 29: 936-944.

Jones, B.C., Hawksworth, G., Horne, V.A., Newlands, A., Morsman, J., Tute, M.S., Smith, D.A. 1996. Putative active site template model for cytochrome P4502C9 tolbutamide hydroxylase. *Drug Metab. Dispos.*, **24**: 260-266.

Lewis, D.F. 2003. Essential requirements for substrate binding affinity and selectivity toward human CYP2 family enzymes. *Arch Biochem Biophys.*, 409: 32-44.

Lewis, D.F. 2003. P450 structures and oxidative metabolism of xenobiotics. *Pharmacogenomics*, 4: 387-395.

Lewis, D.F.V., Eddershaw, P.J., Dickens, M., Tarbit, M.H., Goldfarb, P.S. 1998. Structural determinants of P450 substrate specificity, binding affinity and catalytic rate. *Chem Biol Interactions*, 115: 175-199.

Mancy, A., Brotto, P., Dijols, S., Dansette, P.M., Mansuy, D. 1995. The substrate binding site of human liver cytochrome P450 2C9: an approach using designed tienilic acid derivatives and molecular modeling. *Biochemistry*, 34: 10365-10375.

Nebert, D.W., Jones, J.E., Owens, J., Puga, A. 1988. Evolution of the P450 gene superfamily. *Prog Clin Biol Res.*, 274: 557-576.

Poulos, T.L., Finzel, B.C., Howard, A.J. 1987. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.*, 195: 687-700.

Ridderstrom, M., Zamora, I., Fjellstrom, O., Andersson, T.B. 2001. Analysis of Selective Regions in the Active Sites of Human Cytochromes P450,2C8, 2C9, 2C18, and 2C19 Homology Models Using GRID/CPCA:. *J. Med. Chem.*, 44: 4072-4081.

Shimoji, M., Yin, H., Higgins, L., Jones, J.P. 1998. Design of a novel P450: a functional bacterial-human cytochrome P450 chimera. *Biochemistry*, 37: 8848-8852.

Smith,D.A., Ackland,M.J., Jones,B.C. 1997. Properties of cytochrome P450 isoenzymes and their substrates. *Drug Discovery Today*, 2: 479-486.

Venhorst, J., ter Laak, A.M., Commandeur, J.N., Funae, Y., Hiroi, T., Vermeulen, N.P. 2003. Homology modeling of rat and human cytochrome P450 2D CYP2D isoforms and computational rationalization of experimental ligand-binding specificities. *J Med Chem.*, 46: 74-86.

Williams, P.A., Cosme, J., Ward, A., Angove, H.C., Vinkovic, D.M., Jhoti, H. 2003. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *NATURE*, 424: 464-468.