RESEARCH

In-silico endogenic substrate prediction study of Cytochrome P450 1B1 (CYP1B1)

Rahul Vishvkarma

School of Biotechnology and Biosciences, Lovely Professional University, Phagwara, Punjab, India

*Corresponding author e-mail: rvishvkarma613@gmail.com

ABSTRACT

Cytochrome P450 1B1 (CYP1B1) belongs to cytochrome P450 monooxygenases superfamily and is one of the CYP1 family isoforms, which actively take part in the metabolism of various xenobiotics in the body. CYP1B1 is highly expressed in multiple human tissues, implied in various types of human cancers, but its intracellular function and role in various human pathologies are yet to be understood entirely. In this regard, research in the past two decades has yielded sufficient information about its structure, amino acid sequence, and interaction with various xenobiotics, including pro-carcinogens. One aspect that remains hard to pin down is the information regarding the endogenic substrates of this protein. The scope of the present study is to predict the endogenic substrates (metabolic products of either arachidonic acid or cholesterol) of CYP1B1 with high binding affinity characteristics. We performed *in-silico* screening for these endogenic compounds with CYP1B1 protein modes to find best binding affinity, atomic contact energy (ACE) and score values. A total of thirteen compounds showed the highest binding affinity with CYP1B1. We found that leukotriene E4, leukotriene A4, Deoxycholic acid, tetrahydrocortisone showed highest ACE values and other substrate binding characteristics, making them excellent candidates for validation by laboratory experiments.

KEYWORDS: Cytochrome P450, CYP1B1, Endogenic Substrate, Leukotrienes, Prostaglandins, Docking

Citation: Vishvkarma R. In-silico endogenic substrate prediction study of Cytochrome P450 1B1 (CYP1B1). Polymorphism. 2019;3:24-34.

INTRODUCTION

The human CYP1B1 protein belongs to the Cytochrome P450 heme-thiolate monoxygenases superfamily, which is involved in the metabolism of a variety of xenobiotic compounds (Nelson et al., 1996). The CYP450 enzymes are found in several life forms, including animals, plants, fungi, bacteria, archaea, protists and even viruses (Johnson and Stout, 2005). The cytochrome P450s actively take part in the metabolism of many xenobiotics, endogenic and natural compounds (Faiq et al., 2014). The biotransformation of xenobiotics is mainly done by CYPP450 subfamilies 1, 2, and 3, which convert the non-polar compounds into polar compounds so that they can efficiently excrete out of the body. Up to seventy-five per cent of the drugs are metabolized by CYPs (Guengrich et al., 2003). Cytochrome P450s in humans are also involved in the synthesis of steroids, cholesterol and other vital compounds, such as thromboxanes and leukotrienes. These compounds are metabolites of arachidonic acid (Rendric & Guengerich 2015; Guengerich, 2015 Omura, 1999). CYP1B1 has a role in the metabolic activation of pro-carcinogen compounds, such as polycyclic aromatic hydrocarbon (PAH) and cyclic amines/amide, which attacks DNA and cause mutation (Nebert et al., 2002). CYP1B1 is expressed in significant levels in the liver, lungs, heart, prostate, and parts of the brain. CYP1B1 is known to be overexpressed in lymphomas and tumours of the breast, ovary, prostate, and lungs (Hessel et al., 2013). Therefore, a potential role in tumorigenesis is also suggested. CYP1B1 also serves as a functional biomarker in the diagnosis of many types of cancers. Metabolic activation of endogenic compounds, e.g. 17β-estradiol (E2) to 4-hydroxy-17β-estradiol is postulated to be a crucial carcinogenic factor in breast cancer (Murray et al., 1997). CYP1 family genes are activated by transcription factor aryl hydrocarbon receptor (Ahr) to which different xenobiotics bind as ligands (Kesharwani et al., 2014). A recent study has suggested that *cyp1b1* generates reactive oxygen species, which results in hypertension and is responsible for cardiovascular diseases (Schmidt et al., 1996). The isoforms of the CYP family show varying substrate specificity. CYP1A1 and CYP1A2 pose approximately 70% sequence similarity and possess about 40% sequence similarity to CYP1B1 (Moorthy et al., 2015).

The endogenic substrate of CYP1B1 is not yet known. If the endogenic substrate is found, we can predict the biological consequences of CYP1B1 metabolite products and will be very helpful in understanding the pathogenesis of many diseases where CYP1B1 is implicated. So, In contrast to scientific approaches for finding CYP1B1 substrate specificity which mainly include catalytic assays, we tried the in-silico approach which employs various molecular modelling and docking methods. The significant advantage of in-silico approaches is that it allows prediction without experimental determination and is particularly favored in earlier stages of substrate prediction, though the accuracy of this approach is not high. The Molecular docking technique helps in understanding substrate conformation and orientation after the interaction of ligand with the active site of the protein. The binding reveals key groups and nearby atoms for interacting with the ligand. The docking process consists of continually posing and ranking the molecules in the vicinity of the active site. This process gives the binding score, area and the atomic combining energy or energy of interaction of the ligand-protein complexes. The docking process cannot precisely simulate the biological process since several assumptions are made by the scoring functions (Ekins et al., 2001). This method is not comprehensive in estimating the active site or substrate prediction but gives valuable information concerning binding affinity and steric effects. After docking, a detailed visual inspection is required to guarantee that the ligand and protein are undergoing a metabolic reaction. Thus, in this study, we employed in silico approaches to find out the best suitable endogenic compounds involved in the metabolism by CYP1B1. The endogenic

compounds considered for this study are all previously known endogenic compounds of the body. We emphasized in selecting all the endogenic compounds which originated from either cholesterol or arachidonic acid.

MATERIALS AND METHODS

The methodology included the finding of the best suitable substrate based on molecular docking by studying various parameters. The first step of this work consisted of in-silico approach by which we obtained the preliminary data about the substrates which may show the best binding to CYP1B1 and then we further moved to in-vitro spectral binding studies.

Ligand Selection

Initial steps involved the finding of the endogenic compounds which are involved in various human metabolic pathways. The list of endogenic compounds was prepared from the literature of the Cytochrome P450 substrates. These compounds are actively engaged in the multiple metabolic pathways. The compounds mainly included those which are metabolized by CYP1, CYP2 and CYP3 families of cytochromes. 2,3,4,5 tetramethoxystilbene, a known most potent inhibitor of CYP1B1, was considered as a reference standard for the entire study (Kim et al., 2002). Since inhibitors and substrates share very similar binding characteristics in the active site of the enzyme, this compound is an important parameter of analysis. The list of compounds studied is given below in Table 1.

Ligand Preparations

The ligands were downloaded from the PubChem server (https://pubchem.ncbi.nlm.nih.gov/) in SWF format. The file format is then converted into the PDB format by using software "Bebel" (http://openbabel.org/wiki/Main_Page) file converter because docking prediction can be made by using the same file format only.

Table 1: List of Endogenic compounds

S. No.	Compounds Name	Reference/PubMed Id
1	Aldosterone	PubMed Id 25465475
2	Allopregnanolone	PubMed Id 25678418
3	Arachidonic acid	PubMed Id 25595103
4	Bilirubin diglucuronide	PubMed Id 24837423
5	Bilirubin glucuronate	PubMed Id 24837423
6	Bilirubin	PubMed Id 24837423
7	Deoxycholic acid	PubMed Id 25678418
8	Estrone	PubMed Id 25678418
9	Hydrocortisone	PubMed Id 25465475
10	Dehydroepiandrosterone	PubMed Id 25678418
11	Leptin	PubMed Id 90470904
12	Leukotriene A4	PubMed Id 6313004

13	Leukotriene B4	PubMed Id 23362865
14	Leukotriene C4	PubMed Id 23362865
15	Leukotriene D4	PubMed Id 23362865
16	Leukotriene E4	PubMed Id 23362865
17	Melatonin	PubMed id 25658124
18	Pregnenolone	PubMed Id 25678418
19	Prostaglandins F2 β	PubMed Id 23362865
20	Prostaglandins A1	PubMed Id 23362865
21	Prostaglandins B1	PubMed Id 23362865
22	Prostaglandins D2	PubMed Id 23362865
23	Prostaglandins D3	PubMed Id 23362865
24	Prostaglandins E3	PubMed Id 23362865
25	Prostaglandins F1	PubMed Id 23362865
26	Prostaglandins F3 α	PubMed Id 23362865
27	Prostaglandins G2	PubMed Id 23362865
28	Prostaglandins H2	PubMed Id 23362865
29	Prostaglandins A2	PubMed Id 23362865
30	Testosterone	PubMed Id 1958510
31	Tetrahydrocortisol	PubMed Id 25678418
32	Tetrahydrocortisone	PubMed Id 25678418
33	Thromboxane A2	PubMed Id 24623680
34	Thromboxane A3	PubMed Id 24623680
35	Thromboxane B2	PubMed Id 24623680
36	Thromboxane B3	PubMed Id 24623680
37	17 beta- estradiol	PubMed Id 24777982
38	6-keto prostaglandin E1	PubMed Id 23362865
39	8-iso prostaglandin E2	PubMed Id 23362865
40	2,3,4,5-tetramethoxystelbene	PubMed Id 25613613

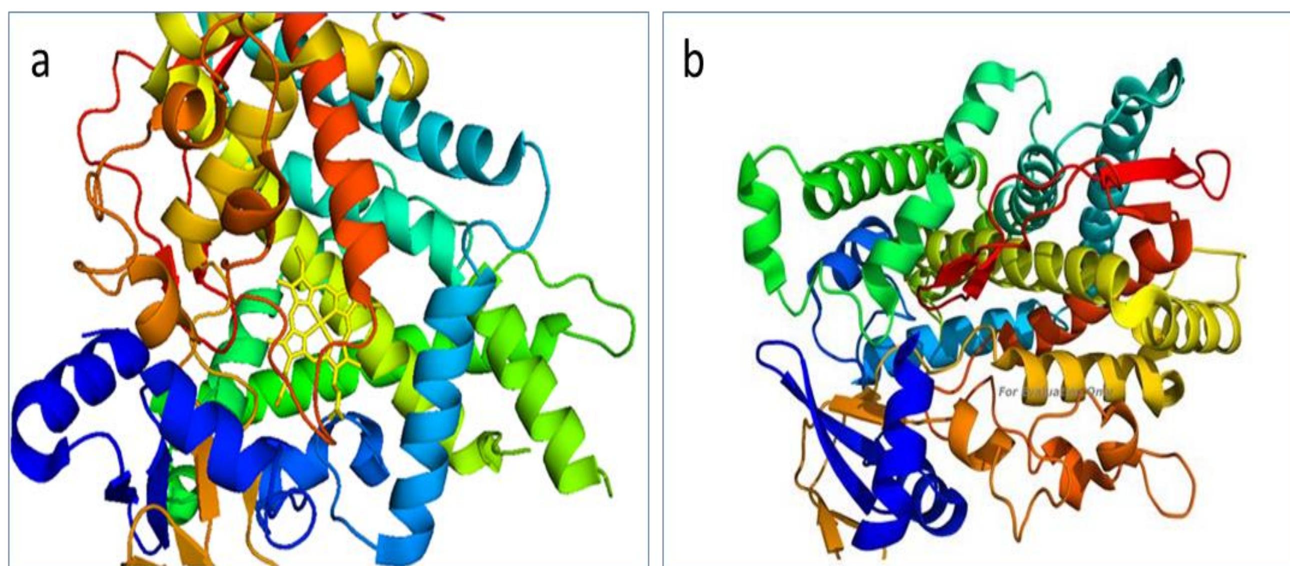

Figure 1. Three-Dimensional Structure of CYP1B (a) with Heme (b) Without Heme ((Image showing spectrum colour, the image generated through PyMol 3D viewer using CYP1B1 PDB file, ID 3PMO).

Receptor CYP1B1 protein modes

The receptor file for CYP1B1 was prepared by matching the parameters including the two different structure of the CYP1B1 (Original structure with Heme and Without Heme). These files were prepared by removing OH and Heme HETATM for showing different molecular binding (Figure 1).

Molecular Docking

CYP1B1 receptor files were used for docking with each endogenic substrate shown in Table 1. The docking was performed by using the online docking severs "Patchdoc" (https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php) and "Hex" server (http://hexserver.loria.fr/). The result of docking comprises the area of binding, score and ACE value. The "Patchdoc" calculates the logarithm around 6 A^0 and gives the Ace (Atomic contact energy) value. The MRLD (mean root square deviation) algorithm was used for docking (Duhovny et al., 2002; 2005).

The ACE value denotes the "Atomic contact energy"; higher the negative value shows high interaction between the ligand and receptor molecules.

Prediction of the active site

The active site construct is built by finding the amino acid residues which actively interact with ligand in close proximity. Applying "ClustalW", we found the sequence similarity between the amino acid residues, and hence, we can predict the active site. By finding the distance between ligand and its surrounding amino acid residues, we found the exact position of the ligand binding.

Results and discussion

Protein-Ligand interaction

All ligand molecules were docked with CYP1B1 in reference with 2,3,4,5 tetramethoxystilbene (with Heme and without Heme), and all atomic contact energy (ACE) and Score values were measured. The information is tabulated below (Table 2). Best ACE value comparisons are given in Figure 2.

The molecular docking with heme-containing CYP1B1 showed that the "leukotriene E4" gives the highest ACE value (-419.80) while leukotriene A4 (-346.92) and deoxycholic acid (-322.02) give relatively lower values. 6-ketoprostaglandin and 8-isoprostaglandin E2 provide similar values (-306.87

and -306.27). Tetrahydrocortisone and tetrahydrocortisol give the same values (-305.15). Out of all, four best interacting molecules are shown in figure 3.

Without heme CYP1B1 binding showed results with relatively higher ACE value. Leukotriene family compounds showed high binding affinity to the CYP1B1 and gave best binding ACE values. The ACE values of Leukotriene A4, C4, D4, and E4 were -269.99, -580.6, -347.18 and -453.58, respectively. Prostaglandin D3 (-329.52) and prostaglandin F2 beta (-347.24) gave satisfactory ACE values.

Table: 2 Endogenic substrates and their ACE and score values

S.No.	Endogenic Compounds	ACE (atomic contact energy) KJ/MOL		SCORE	
		With HEME	Without HEME	With HEME	Without HEME
1	17 beta estradiol	-228.71	-229.83	3478	3556
2	6-ketoprostaglandin E2	**-306.87**	-285.75	3974	4338
3	8-isoprostaglandins E2	**-306.87**	-285.75	3974	4338
4	Aldosterone	-245.35	-297.70	3856	3720
5	Allopergnanolone	-264.73	-289.63	3448	3618
6	Arachedonic acid	-283.75	-252.60	4092	4122
7	Bilirubin	-298.63	**-587.39**	4422	5134
8	Bilirubinglucuronate	-175.36	-270.01	5238	5848
9	Deoxycholic acid	**-322.02**	**-384.761**	3946	4068
10	Estrone	-230.81	-205.73	3548	3008
11	Hydrocortisone	-281.23	-279.56	3640	3862
12	Dehydroepiandrosterone	-245.71	-273.60	3550	3540
13	Leukotriene A4	**-346.92**	-269.22	4034	4994
14	Leukotriene B4	-166.66	-232.25	3978	4736
15	Leukotriene C4	-239.02	**-580.6**	4918	5920
16	Leukotriene D4	-269.97	**-347.18**	4538	5120
17	Leukotriene E4	**-419.80**	**-453.58**	4474	5482
18	Melatonin	-198.89	-266.54	3602	3468
19	Pregnanolone	-275.13	-275.49	3688	3734
20	Prostagladin D3	**-300.4**	**-329.52**	4188	4544
21	Prostaglandin E3	**-300.2**	**-300.62**	4182	4358
22	Prostaglandin A1	-283.76	**-333.39**	3772	4352
23	prostaglandin A2	**-305.62**	-263.93	4308	4346
24	Prostaglandin B1	-284.4	**-314.29**	3754	4238
25	Prostaglandin D2	-294.45	-308.33	4332	4326
26	Prostaglandin F1	-193.98	**-327.79**	3656	4436
27	Prostaglandin F2 beta	-286.07	**-347.24**	5016	4560
28	Prostaglandin F3 alpha	-193.98	**-327.79**	3656	4436
29	Prostaglandin G2	-280.01	-281.36	4388	4922
30	prostaglandin h2	-167	-277.45	4076	4906

31	Testosterone	-229.92	-229.92	3426	3552
32	Tetrahydrocortisol	-305.15	-271.21	3582	3542
33	Tetrahydrocortisone	-305.15	-271.21	3582	3542
34	Thromboxane A2	-287.53	-254.20	4608	4944
35	Thromboxane A3	-287.53	-254.20	4608	4944
36	Thromboxane B2	-284..29	-330.07	4364	4492
37	Thromboxane B3	-158.82	-270.84	3942	4334
38	2,3,4,5 tetramethoxystilbene	-259.23	-286.1	4211	4178

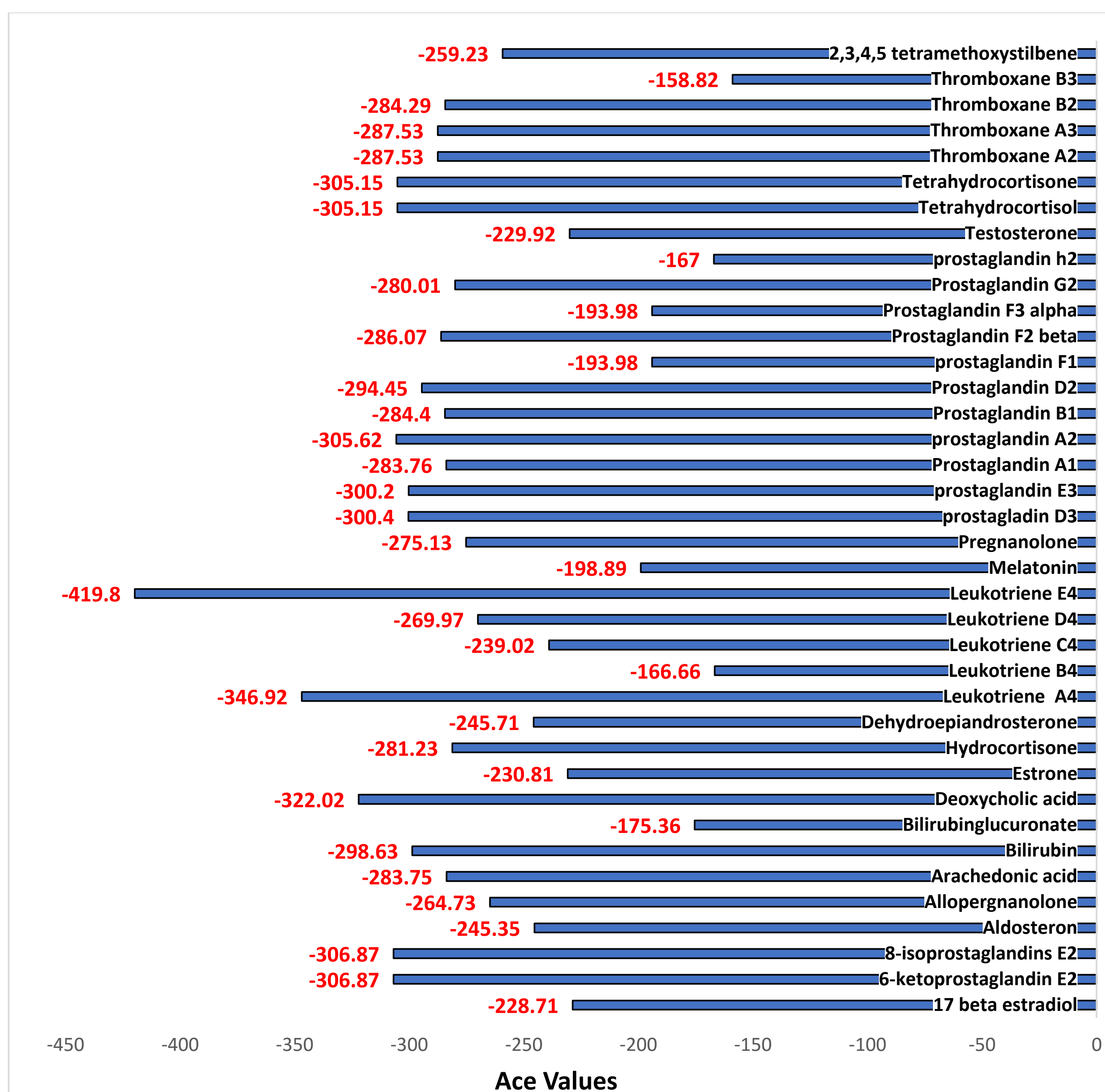

Figure 2. Graph showing the ACE values of heme-containing CYP1B1 with different endogenic compounds. The substrate name is shown in the bar.

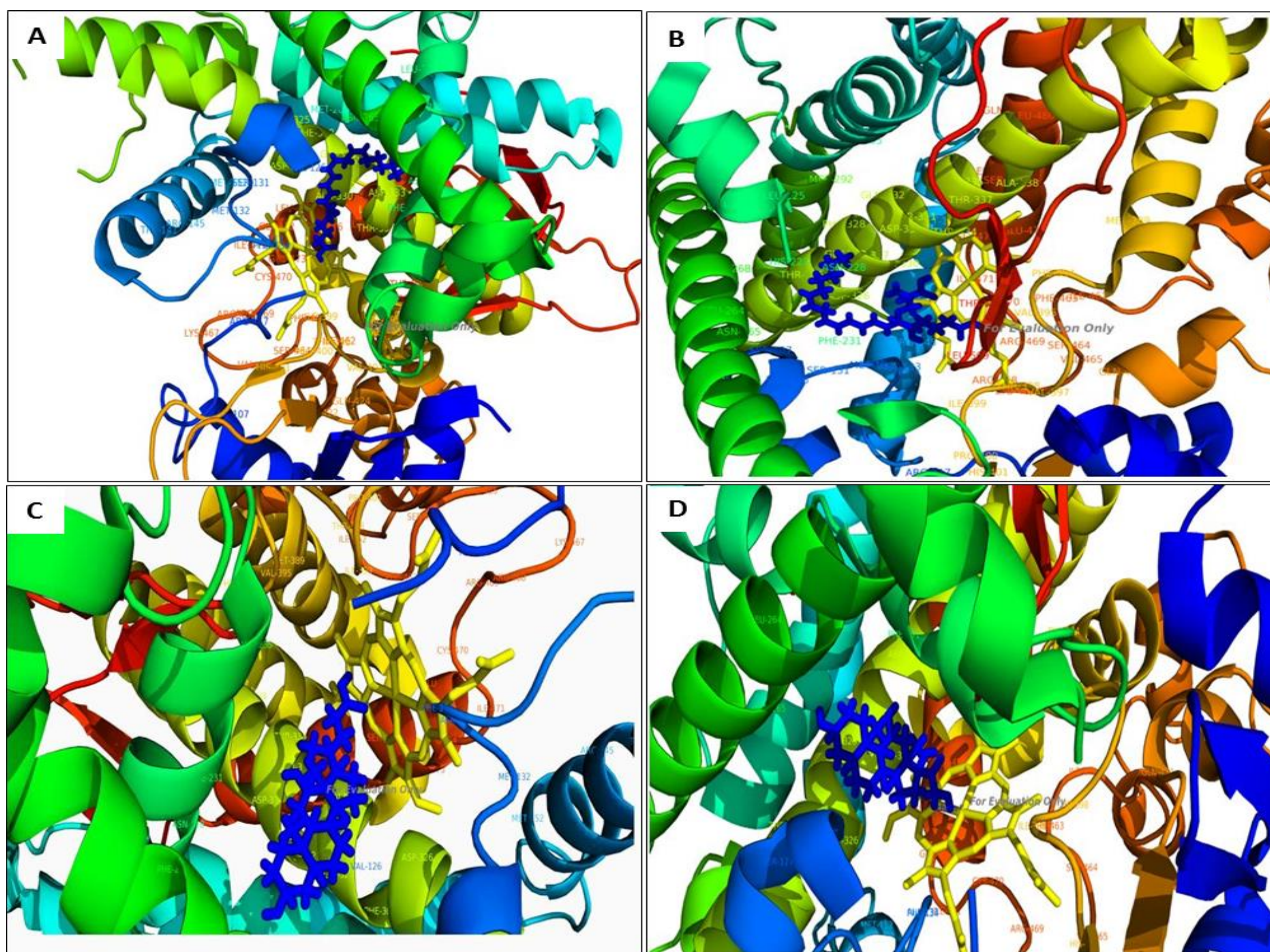

Figure 3. Figure showing 4 highest ACE value ligands (blue color) interacting with CYP1B1. (A) Leukotriene E4, (B) Leukotriene A4, (C) Deoxycholic acid, (D) Tetrahydrocortisol.

It was found that 13 substrates showed high ACE values when the heme was removed, these compounds showed ACE values higher than original CYP1B1 which contain heme. Out of these 13 substrates, Bilirubin, leukotriene C4, E4, and D4 showed highest ACE values. The rest of the substrate showed relatively less ACE values.

From the above results, we conclude that the leukotriene family compounds actively interact with CYP1B1 and give high ACE value than any other compounds (Figure 4). Therefore, leukotriene E4 maybe the most putative endogenic substrates for CYP1B1. Rest of the potential substrates such as deoxycholic acid, tetrahydrocortisol and tetrahydrocortisone showed relatively low ACE values.

Finally, we sorted out ten endogenic ligand molecules based on the ACE values, and one can further go for in-vitro spectral binding studies using rat hepatic microsomes for further validation of this preliminary finding.

Active site prediction

For the prediction of the active site of CYP1B1, ten best compounds showing high ACE value were taken and analysed in PyMol 3D viewer. We took the CYP1B1 amino acid sequence present in the close proximity (within $6A^0$) of ligand binding site for each ligand. Then sequence similarity was checked by ClustalW analysis and found that the first three amino acid residues of the receptor molecules were interacting with the ligand and these were valine (V), serine (S), asparagine (N) respectively, at number five, eleven and fifteen, phenylalanine (F) is

conserved. At thirteen and fourteen number, leucine (L) and arginine (N) are conserved. Threonine (T) and aspartic acid (D) are found in the vicinity of CYP1B1 binding site. Based on the residue analysis, it may be concluded that these residues are present with every ligand type (Figure 5).

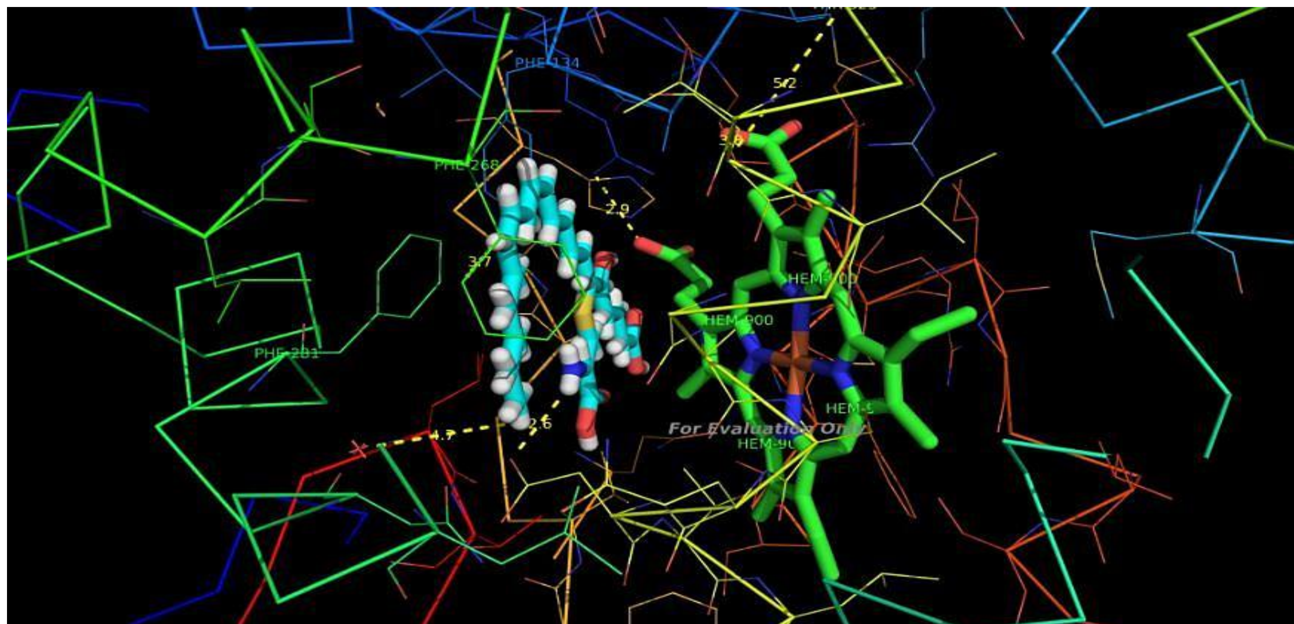

Figure 4. CYP1B1 active site interacting amino acid residues and their distance from the ligand Leukotriene E4, view in PyMOlL3D viewer.

```
VSSAFLHNEEFGLNFAITDIFGASQDTVTIIL-----  32  Prostaglandin D3
VSSAFLHNEEFGLNF--TDIFGASQDTVTIILT----  31  Prostataglandin E3
VSSAFLLHNEFGLNF--TDIG-ASQDTVIFCLTK---  31  Leukotriene A4
VS-AFLHNFGVFLNF--TDIFGASQDTVILT------  28  Dioxycholic acid
VSSAFN----FGLNF--TDIG-ASQDTICLT------  24  Tetrahydrocortisole
VSSAFN----FGLNF--TDIG-ASQDTICLT------  24  Tetrahydrocortisone
VS-AFH-NEEFGLNF--TDIG---ADTIL--------  22  Prosta glandin A2
VSAFLLSHNEFGLNF--TDGA---QDTVTIQIFSCLT  32  6-ketoprostaglandin
VSAFLLSHNEFGLNF--TDGA---QDTVTIQIFSCLT  32  8-isoprostaglandin
VSSAFLN---F-LNF--TDGA---QDTVILT------  22  Leukotriene E4
**   :      . ***  **        **:
```

Figure 5. Amino acids sequence similarity of CYP1B1 present within close proximity of ligand binding site. An * (asterisk) indicates positions which have a single, fully conserved residue. A : (colon) indicates conservation between groups of strongly similar properties. A . (period) indicates conservation between groups of weakly similar properties.

CONCLUSION

It is well known that CYP1B1 has significant roles in various biological functions of the body and is the cause of several cancers. The CYP1B1 is non-specific and shows a wide range of substrate specificities, and yet endogenic substrates of CYP1B1 remain unknown. Therefore, this study was conducted to narrow down the search of the endogenic

substrates of CYP1B1. In the preliminary steps, we screened various endogenic compounds in reference to CYP family substrates. In this search, we were able to find a few compounds with the lowest binding energy, best ACE values and scores. In conclusion, we found that leukotriene E4, leukotriene A4, Deoxycholic acid, tetrahydrocortisone showed highest ACE values and other substrate binding characteristics, making them excellent candidates for validation by laboratory experiments. This could finally help in narrowing the search for the elusive endogenic substrates of CYP1B1. Laboratory experiments need to check how these compounds bind with the CYP1B1 protein in reality.

Acknowledgements

Faculty members of the school of Biosciences and Biotechnology, Lovely Professional University is acknowledged for valuable support. The author performed this study under the guidance of Dr E. Naveen Prasad Reddy as project work for the completion of his Master's degree. The author acknowledges suggestions and advice from Prof. Chirag Chopra in molecular docking studies.

Conflict of interest statement

The author has declared that no competing or conflict of interests exists. The funders, if any, had no role in study design, writing of the manuscript and decision to publish.

Authors' contributions

RV conceived the idea under the mentorship of Dr. E. Naveen Prasad Reddy as project work for the completion of his Master's degree.

REFERENCES

Duhovny D, Nussinov R, Wolfson HJ. Efficient Unbound Docking of Rigid Molecules. In Gusfield et al., Ed. Proceedings of the 2'nd Workshop on Algorithms in Bioinformatics(WABI) Rome, Italy, Lecture Notes in Computer Science 2452, pp. 185-200, Springer Verlag, 2002

Ekins, S.; de Groot, M.J.; Jones, J.P. Pharmacophore and three-dimensional quantitative structure activity relationship methods for modeling cytochrome p450 active sites. Drug Metab. Dispos. 2001, 29, 936–944.

Faiq MA, Dada R, Sharma R, Saluja D, and Dada T (2014) CYP1B1: a unique gene with unique characteristics. Curr Drug Metab 15:893–914

Guengerich FP. Human cytochrome P450 enzymes. In: Cytochrome P450: Structure, Mechanism, and Biochemistry; Ortiz de Montellano PR, (Ed) Springer 2015; 4th ed, pp. 523-.

Hessel S, Lampen A, Seidel A. Polycyclic aromatic hydrocarbons in food—effluxof the conjugated biomarker 1-hydroxypyrene is mediated by Breast CancerResistance Protein (ABCG2) in human intestinal Caco-2 cells. Food ChemToxicol.2013 Dec;62:797-804.

http://hexserver.loria.fr/

http://openbabel.org/wiki/Main_Page

http://openbabel.org/wiki/Main_Page

https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php

https://pubchem.ncbi.nlm.nih.gov/

Johnson EF, Stout CD. Structural diversity of human xenobiotic-metabolizing cytochrome P450 monooxygenases. BiochemBiophys Res Commun. 2005 Dec 9;338 (1):331-6.

Kesharwani SS, Nandekar PP, Pragyan P, Sangamwar AT. Comparative proteomics among cytochrome p450 family 1 for differential substrate specificity. Protein J. 2014 Dec;33(6):536-48.

Kim S, Ko H, Park JE, Jung S, Lee SK, Chun YJ. Design, synthesis, and discovery of novel trans-stilbene analogues as potent and selective human cytochrome P450 1B1 inhibitors. Journal of medicinal chemistry. 2002 Jan 3;45(1):160-4.

Moorthy B, Chu C, Carlin DJ. Polycyclic aromatic hydrocarbons: from metabolism to lung cancer. Toxicol Sci. 2015 May;145(1):5-15

Murray GI, Taylor MC, McFadyen MCE, McKay JA, Greenlee WF, Burke MD & Melvin WT Tumor-specific expression of cytochrome P450 CYP1B1 (1997) Cancer Res 57: 3026-3031.

Nebert DW, Russell DW. Clinical importance of the cytochromes P450. Lancet. 2002 Oct 12;360(9340):1155-62.

Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC, Nebert DW. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics. 1996 Feb;6(1):1-42.

Omura T. Forty years of cytochrome P450. Biochem Biophys Res Commun 1999; 266(3): 690

Peter Guengerich F, Chun YJ, Kim D, Gillam EM, Shimada T. Cytochrome P450 1B1: a target for inhibition in anticarcinogenesis strategies. Mutat Res. 2003 Feb-Mar 523-524:173-82.

Rendic S, Guengerich FP. Survey of human oxidoreductases and cytochrome P450 enzymes involved in the

metabolism of xenobiotic and natural chemicals. Chem Res Toxicol 2015; 28(1): 38-42.

Schmidt JV, Su GH, Reddy JK, Simon MC, Bradfield CA. Characterization of amurine Ahr null allele: involvement of the Ah receptor in hepatic growth and development. ProcNatlAcadSci U S A. 1996 Jun 25;93(13):6731-6.

Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucl. Acids. Res. 33: W363-367, 2005.