BIBLIOGRAPHIC INFORMATION SYSTEM

👁 **Vision:** Journal of Biomedical Research & Environmental Sciences main aim is to enhance the importance of science and technology to the scientific community and also to provide an equal opportunity to seek and share ideas to all our researchers and scientists without any barriers to develop their career and helping in their development of discovering the world.

ORIGINAL ARTICLE

Molecular Investigation of Cytochrome P-450 Enzymes Inductors

Terzulum G[1]* and Andrei Ivanovich K[2]

[1]M.Sc Biotechnology (in view), Engineering School of New Production Technologies, Tomsk Polytechnic University, Tomsk, Russia

[2]N.M. Kizhner Research Center, Tomsk Polytechnic University, Tomsk, Russia

Abstract

Cytochrome P-450 enzymes play a critical role in xenobiotic metabolism and homeostasis, with their induction offering therapeutic potential for liver diseases and metabolic disorders. Using molecular docking and statistical analyses, the research examines the binding affinities and structural features of 46 known inducers, classifying them into Type I (affecting apo-protein expression) and Type II (directly influencing the heme center). Linear Discriminant Analysis (LDA) achieved an 87% accuracy in classifying these inducers, with key interactions identified at amino acid residues Glu301, Phe115, and Ser294. Compounds such as 3,5-difluoromethylurea and 3,4-dichlorophenylmethylurea demonstrated strong binding, highlighting their potential as effective inducers. The study also developed Quantitative Structure-Activity Relationship (QSAR) models to predict binding constants, providing insights into the structural determinants of induction. These findings advance the understanding of cytochrome P-450 induction mechanisms and lay the groundwork for designing safer, more selective drugs.

Introduction

The industrialization and chemical development of crucial life industries over the years has exponentially increased circulation of foreign chemical factors in the environment. Living organisms therefore, are constantly exposed to these foreign chemicals - xenobiotics, which are unnatural participants in normally occurring biochemical reactions in the organism. Hundreds of thousands of these chemicals are metabolized by heme-containing enzymes belonging to the cytochrome P-450 superfamily which are present in all living things. Over 300 CYP isoforms have been identified, able to catalyse not less than 60 varying enzymatic reactions [1,2].

Cytochrome P-450 enzymes (CYPs) are heme-thiolate proteins that mediate the metabolism of a numerous endogenous and exogenous compounds, comprising drugs, toxins, hormones, and fatty acids. The name "cytochrome P-450 enzymes" is derived from its following characteristics: they are bound to the cell membrane (cyto) and contain heme pigment (chromium and P). When bound to carbon monoxide, these proteins give a spectrum with an absorption maximum around 450 nm [1,3].

Scan Me

Cytochrome P-450 catalyses aromatic and aliphatic hydroxylation, oxidative deformylation, reductive dehalogenation, dehydrogenation, reduction of arenes and amine oxides, epoxidation, dealkylation, deamination, denitrosation de-esterification, and β-splitting [3,4]. The diversity of cytochrome P-450 isoforms is regarded as a remarkable phenomenon in biochemistry. This is largely due to their ability to exist in multiple forms, allowing them to participate in the oxidation of numerous compounds. However, the exact number of chemicals metabolized by CYPs remains uncertain [4].

The cytochrome P-450 superfamily plays a crucial role in maintaining homeostasis, with alterations in the ratio of different isoforms often linked to various pathological conditions [2,5]. These structurally and functionally distinct iso-enzymes in mammals originate from a complex gene superfamily.

Aim of study: To study the molecular interactions between cytochrome P-450 (CYP) enzymes and low molecular weight compounds that can induce its activity in the body.

- To conduct a thorough review of the functional and structural characterization of cytochrome P 450 enzymes.

- To investigate the molecular interactions between cytochrome P-450 and low molecular weight inducers of the enzyme by docking method.

- To classify experimental inducers by statistical comparison of docking results.

Object and Research Methodology

Object of research

The object of this study is the human cytochrome P450 monooxygenase, which is involved in the metabolism of various endogenous substrates. The enzyme's binding site was assessed using molecular docking to evaluate its complementarity with several low-molecular-weight compounds known to act as inducers of P450 in rats. A key component of the binding site is the heme moiety (cofactor), which contains a chelated iron ion (Figures 1,2).

Methodology: molecular docking

A homology model of human CYP2C9 was utilized as the primary receptor structure, incorporating the
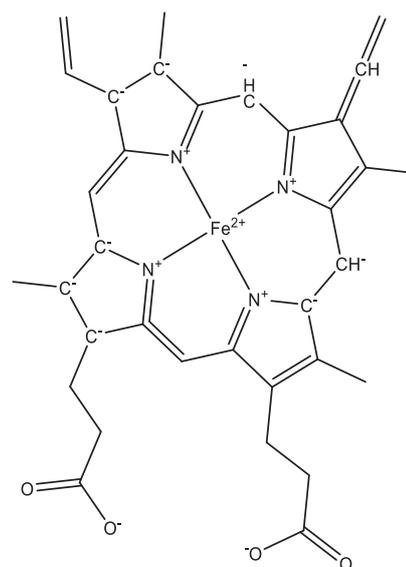
Figure 1 Schematic of the computational steps including protein preparation, ligand selection, docking, and interaction analysis for identifying CYP2C9 inhibitors.
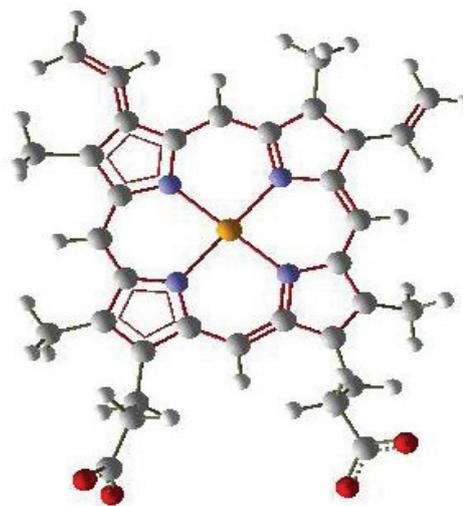
Figure 2 Cytochrome CYP2C9_HUMAN.

docked heme cofactor (HEM_492) and the active PHE_1 domain of hem synthase, modulated by the peptide agonist Ala-Arg-Cys-Cys-Glu-Gly-His-Ile-Leu-Lys-Phe-Pro-Ser-Thr-Val (sourced from UniProt as structure WT2). This model was chosen due to its relevance in representing the functional state of CYP2C9, including the heme-binding site critical for catalytic activity and induction-related interactions [5]. The receptor structures were imported into Molegro Virtual Docker 6.0 (MVD) for preparation and docking. Prior to docking, the receptor models

underwent validation checks within MVD to ensure completeness of the structures (e.g., no missing atoms or residues in the active site). The receptor structures were treated as rigid during docking, a standard approximation to reduce computational complexity while focusing on ligand flexibility, which is often sufficient for initial pose prediction in cytochrome P-450 studies [1].

Docking simulations were performed using Molegro Virtual Docker 6.0 (MVD), leveraging its MolDock scoring function, which evaluates binding interactions based on a combination of steric, electrostatic, hydrogen-bonding, and torsional energy terms [QiSoft, 2015]. The search space for docking was defined as a spherical region with a radius of 9 Å, centered on the geometric center of gravity of the bound peptide molecule (PHT_1 for HEM_492) within the CYP2C9 homology model. This radius was selected to encompass the active site, including the heme cofactor and adjacent residues known to mediate ligand binding, ensuring comprehensive exploration of potential interaction sites [1]. The MolDock scoring function was applied with a grid resolution of 0.3 Å to balance computational efficiency with accuracy in energy calculations. Conformational mobility of the ligands was accounted for by allowing flexibility in torsional angles, which were automatically detected and parameterized by MVD. For each ligand-receptor pair, 300 independent docking runs were conducted to enhance sampling robustness and reduce the likelihood of missing favorable poses. The receptor structures remained rigid during docking, consistent with the focus on ligand conformational adaptability in this initial study.

Linear Discriminant Analysis (LDA) was then performed to determine classification functions to recognize compounds as belonging to one or another inducer class (I or II). Such a task belongs to the Structure-Activity Relationship (SAR) analysis. In addition, within each of the two classes, we attempted to perform QSAR analysis (Quantitative Structure-Activity Relatonships) using literature data on dissociation constants of inductor-enzyme complexes [1].

Results and Discussions

For each of the 46 inducers investigated, the best docking position was selected for further analysis. The data were ordered according to the distance between ligand and receptor, partial values of the evaluation function corresponding to amino acid residues, and compound class (type of binding to cytochrome according to the pattern of spectral changes) figures 3-7.
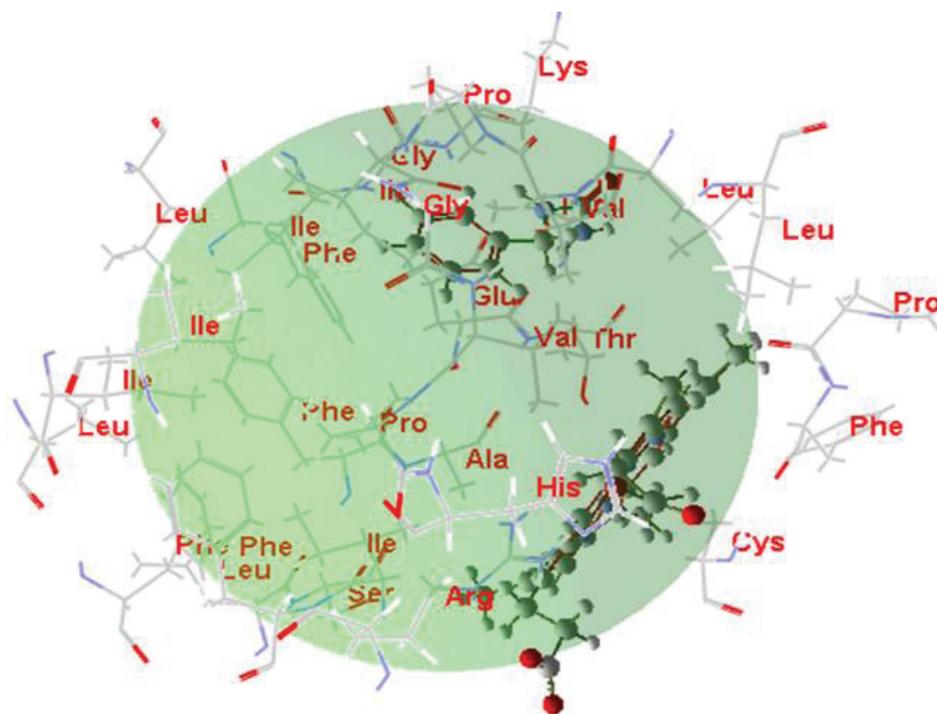

Figure 3 Spherical search region of docking positions, cofactor and surrounding amino acid residues.

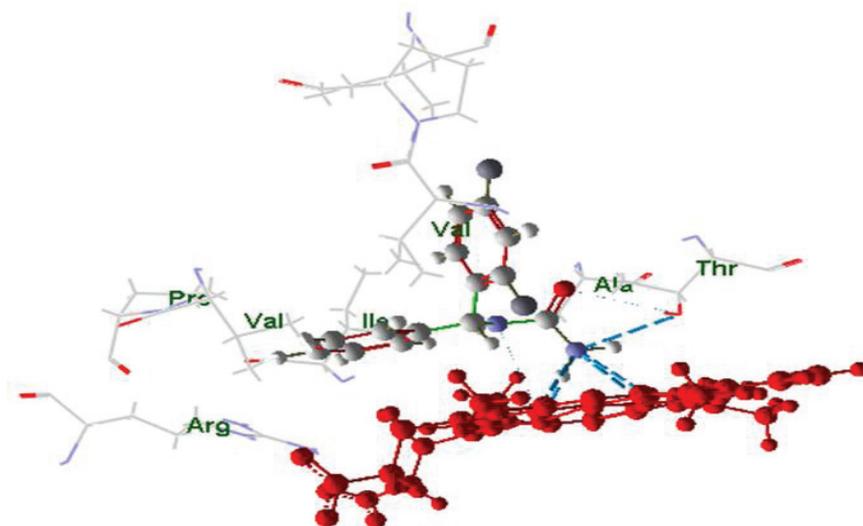**Figure 4** 3,5 Difluoromethylurea. Numerous hydrophobic contacts and hydrogen bonds are observed.
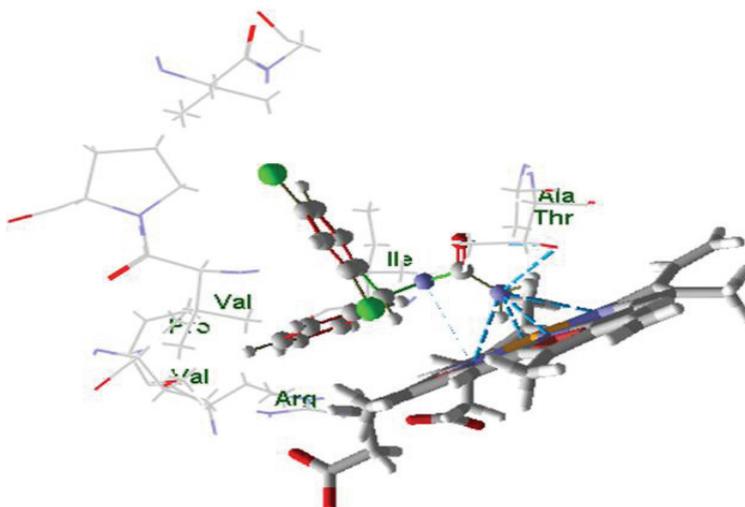
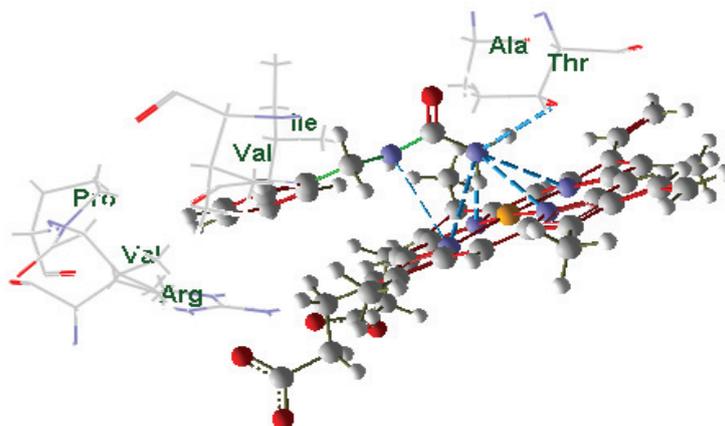

Figure 5 3,4 dichlorophenylmethylurea.

Figure 6 1 Benzylurea in docking position 1. Hydrogen bonds between ligand and receptor are shown.
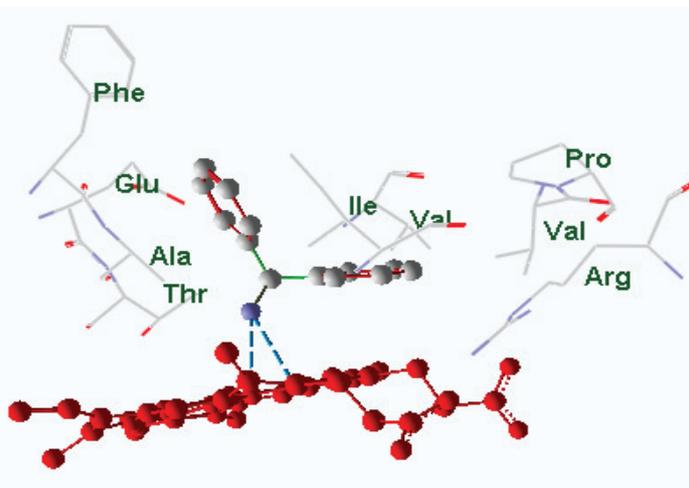
Figure 7 Docking position of benzhydrylamine demonstrating ligand-receptor interaction.

The table summarizes the partial values of the evaluation function at surrounding amino acid residues for each docking position obtained, the distance between ligand and receptor (CYP2C9), and the inducer class based on UV –vis spectral changes. Data from the table were used to construct SAR models using the LDA method in Statistica 6.0 program (Table 1).

Results of discriminatory and correlation analysis of data using statistica 6.0

The results of inductor classification by LDA method are presented in table 2. The table above shows that 4 out of 31 Class II compounds were misclassified based on the LDA model. Similarly, only 2 out of 15 class I compounds were grouped incorrectly. Overall,

the percentage of correct classifications is about 87%, which is a reasonably good result in molecular recognition of bioactive compounds of different groups [6]. The specific compounds that were incorrectly classified are listed in table 3. Linear discriminant analysis of the data makes it possible to identify, for example, 3,5-dibromophenylurea as a class II inducer instead of its experimental classification as a class I compound according to spectral changes (UV-Vis) upon complexation with cytochrome P-450. The same is true for 4-benzylhydrourea and 5,5-diphenylimidazolidin-2,4-dione, which have structural similarity to class II compounds.

Similarly, 3-chlorophenylmethylurea, 5-(12-diphenylethyl) urea and 1,3-diacetyl-5-ethyl-5-phenylpyrimidine, experimentally assigned to

Table 1: Dataset of docking results.

A1			f_x	Соединения															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Соединения	Тип изм. спектр	K⁵ моль/л	Расстояние (Ams)	Ala 298	Arg 73	Arg 98	Arg 125	Arg 308	Arg 434	Asp 361	Cys 436	Glu 105	Glu 218	Glu 301	Glu 387	Glu 439	Glu 445	Gly 299	Gly
2	Benzylurea	II	0.000475	**2.106**	-4.112	0	-6.613	0	0	0	0	-0.517	0	0	-0.867	0	0	0	-0.576	
3	3.4-Diodophenylurea	II	0.00000663	2.009	0.771	0	-4.033	0	0	0	0	-0.507	0	0	-5.346	0	0	0	-0.409	
4	3.4-Dimethylphenylurea	II	0.0000455	2.044	-2.889	0	-6.804	0	0	0	0	-0.483	0	0	-6.925	0	0	0	-0.871	
5	3.5-Difluorophenylurea	II	0.000118	2.007	-2.917	0	-6.565	0	0	0	0	-0.484	0	0	-6.9	0	0	0	-0.5117	
6	3.4-Dichlorophenylurea	II	0.000000474	1.931	-3.775	0	-6.404	0	0	0	0	-0.502	0	0	-7.327	0	0	0	-0.494	
7	3.5-Dibromophenylurea	I	0.000000453	1.895	-3.269	0	-6.348	0	0	0	0	-0.511	0	0	-7.306	0	0	0	-0.808	
8	3.4-Diodophenylmethylurea	II	0.000173	1.893	-3.306	0	-6.373	0	0	0	0	-0.512	0	0	-7.385	0	0	0	-0.811	
9	2.4-Dinitrophenylmethylurea	II	0.0000166	мар.17	-7.852	0	1.275	0	0	0	0	0	0	0	-12.064	0	0	0	0	
10	2.5-Difluorophenylmethylurea	II	0.000274	1.926	-2.916	0	-6.522	0	0	0	0	-0.504	0	0	6.522	0	0	0	-0.835	
11	2.5-Dichlorophenylmethylurea	II	0.0000851	2.038	-3.448	0	-6.332	0	0	0	0	-0.471	0	0	-6.097	0	0	0	-0.486	
12	2.5-bis(dimethyaminophenyl)m	II	0.000131	1.857	-8.683	0.324	-3.459	0.845	1.026	0.878	-0.725	-0.539	-0.317	-1.073	-13.125	-0.723	-0.312	-0.824	-0.91	
13	1-Phenylpropylurea	II	0.0000801	3.517	0	0	0	0	0	0	0	0	0	0	-4.798	0	0	0	0	
14	2.5-Dihydroxyphenylmethylure	II	0.00000589	1.906	-2.729	0	-6.598	0	0	0	0	-0.509	0	0	-5.975	0	0	0	-0.87	
15	2.5-Dinitrophenylmethylurea	II	0.00000885	3.167	-5.821	0	-1.436	0	0	0	0	0	0	0	-8.741	0	0	0		
16	3-Chlorophenylmethylurea	II	0.00000131	2.075	0	0	-6.612	0	0	0	0	-0.586	0	0	-7.659	0	0	0	-0.451	
17	2.3-(m-tolyl(p-tolyl)methyl)urea	II	0.0000175	2.167	-1.572	0	-6.377	0	0	0	0	-0.541	0	0	-7.557	0	0	0	-0.477	
18	2.4-diphenylmethane	II	0.0000377	4.091	-11.886	0	-5.861	0	0	0	0	0	0	0	-5.391	0	0	0	-0.41	
19	2.5-diphenylmethanamine	II	0.0000718	2.932	-13.062	0	-4.029	0	0	0	0	0	0	0	-6.109	0	0	0	-0.403	
20	N-benzhydrylformamide	II	0.0000355	3.462	3.849	0	-4.779	0	0	0	0	-0.402	0	0	-2.986	0	0	0	-0.333	
21	N-(benzhydrylcarbamoyl)aceta	II	0.0000733	4.986	-7.531	0	-3.159	0	0	0	0	0	0	0	-17.325	0	0	0	-0.311	
22	1.3-dibenzhydrylurea	II	0.0000455	3.867	-10.572	0	-5.905	0	0	0	0	0	0	0	-4.523	0	0	0	0	
23	(9H-fluoren-9-yl)urea	II	0.000116	2.008	4.231	0	-4.298	0	0	0	0	0	0	0	-1.619	0	0	0	-0.812	
24	(1-phenylbutyl)urea	II	0.000154	2.162	-3.359	0	-2.721	0	0	0	0	-0.464	0	0	-3.453	0	0	0	-0.382	
25	4-benzhydrylurea	I	0.00000231	2.016	-1.983	0	-6.895	0	0	0	0	-0.488	0	0	-4.793	0	0	0	-0.969	
26	5-(12-diphenylethyl)urea	II	0.00000631	2.169	-2.71	0	-6.955	0	0	0	0	-0.465	0	0	-3.03	0	0	0	-0.391	

Table 2: Classification matrix obtained by linear discriminant analysis.

Class of Inductor	Classification matric			
	Total compounds	I	II	% correct classification
II	31	4	27	87
I	15	13	2	86.6
Total	46	17	29	86.9

Table 3: List of compounds wrongly classified in the LDA SAR analysis.

Compound	Class	
	Experimental	Calculated LDA
3,5-Dibromophenylurea	I	II
3-Chlorophenylmethylurea	II	I
4-Benzhydrylurea	I	II
5-(12-diphenylethyl)urea	II	I
1,3-diacetyl-5-ethyl-5-phenylpyrimidine	II	I
5,5-diphenylimidazolidine-2,4-dione	I	II

class II, have more similarity to class I according to LDA model.

The most important amino acid groups

The SAR model we obtained can be considered quite satisfactory, since good molecular recognition rates (87% correct classifications, see above) were achieved on the basis of partial evaluation functions for only three amino acid residues (Table 4), which can thus be considered the most significant structural elements determining the differences between the classes of cytochrome P-450 inducers studied. Thus, the degree of interaction with Glu301, Phe115 and Ser294 residues mainly determines the differences between the two types of inducers studied.

Based on the data given in table 4, the classification functions for the compounds under study can be expressed by equations of the following form:

$$F_I = a^I_o + a^I_1 x_1 ... + a^I_n x_n$$

Using the above coefficients, the equation for classes I and II will be as follows:

$$F_I = -61,537 + 1,057 * Glu\ 301 - 0,044 * Phe\ 115 - 0,066 * Ser\ 294$$

$$F_{II} = -77,578 + 1,175 * Glu\ 301 - 0,019 * Phe\ 115 - 0,087 * Ser\ 294$$

According to the basic idea of the LDA method, classification based on these features can be performed as follows:

$$F_I > F_{II} \rightarrow \text{Класс I}$$

$$F_{II} > F_I \rightarrow \text{Класс II}$$

Thus, as a result of the performed studies, a rather easy-to-use classification model has been obtained, which makes it possible to predict whether inducers belong to class I and II. From the biological point of view, type II inducers tend to directly affect the heme-center of the enzyme, increasing its catalytic activity or affinity for the substrate. On the other hand, type I inducers affect mainly the apo-protein and lead to increased cytochrome expression (Table 5).

Leave One Out Cross Validation (LOOCV)

LOO is a data cross validation training model for machine learning. For prediction, each compound was removed from the sample one by one, and a modified LDA model was built for the remaining connections, which was then used to determine the class of the removed connection. Thus, each of the connections was "external" at prediction, i.e., not used in model building.

For three compounds, only the zero values in the table column (No variance) remained when removed, so predictions could only be made for 43 connections. Table 6 summarizes the classification matrix. The percentage of correct predictions was 76.7%. This is

Table 4: Amino acid residues most important for molecular recognition of inducers and the corresponding classification function coefficients.

Effect	Classification Function	
	I $p = 0.6739$	II $p = 0.3261$
Linear coefficient	-61.537	-77.578
Glutamine (Glu 301)	1.0573	1.1749
Phenylalanine (Phe 115)	-0.0443	-0.0193
Serine (Ser 294)	-0.0656	-0.0869

Table 5: Compounds that returned a "No variance" LOO result.

Compound	Inductor Type	
	Experimental	Predicted. LOO
2,5-bis(dimethaminophenyl)methylurea	II	0
1-phenylpropylurea	II	0
6-methoxy-1,3-dimethyl-5,5-diphenyldihydropyrimidine	II	0

TOPIC(S): PHARMACOLOGY ● METABOLISM

quite good for biological data, considering that these are predictions of "external" compounds that were not part of the corresponding training sample for the modified LDA model table 7.

Multiple Regression Analysis (QSAR)

Using the data set obtained by molecular docking (Table 1), it is possible to perform QSAR analysis within each class of the studied compounds, since experimentally determined binding constants with cytochrome P-450 are available for them. First of all, however, the lack of interactions with some amino acid residues for class I compounds should be noted. Thus, in contrast to class II, none of the class I compounds had partial interactions with some amino acids, such as Asp 361, His 369, Leu 365 and Ser 360. For class I, a regression QSAR model was obtained by stepwise regression analysis and presented in table 5.

Thus, the multiple regression equation for calculating the pKs values of cytochrome P-450 class I inducers, based on the results of QSAR analysis is as follows:

pKs = 8,5619 + 0,317 Glu 299 − 0,62 Thr 306 + 0,333 Leu 362 − 0,34 Val 363

Conclusions

A systematic literature review allowed us to identify known inducers of cytochrome P-450 in the rat liver monooxygenase system. In this work, molecular docking and statistical analysis were performed and a classification model was constructed to recognize type I and type II inducers. The main conclusions are summarized as follows:

Specific structural features of low molecular weight compounds responsible for hydrostatic interactions and hydrogen bonds are crucial for their affinity for CYP2C9. 1,3-dibenzhydryl urea and 6-(1,3,3-triphenylpropyl) urea showed strong binding interactions with affinity scores $\Delta G = -154.485$ and $\Delta G = -133$ respectively, and hydrogen bond energy of -1.02 kj/mol.

Previously, Saratov AS. et al [1] established a classification model for these compounds into type I and II inducers according to spectra changes in UV-vis from 390nm – over 400nm, however, our LDA model identified six of these inductors highlighted in table 3 to have been wrongly classified by his method. This observation may be due to unique electronic or steric properties not accounted for in UV-vis measurement.

The consistence of the data set highlights the amino acids Glu 301, Phe 115 and Ser 294 as crucial building blocks of bioactive compounds capable of inducing the secretion of monooxygenase cytochrome P-450 either at the heme-center or its apo-protein regions.

Mechanistically, all examined type I inductors of CYP2C9 share activity with these amino acids (Glu 301, Phe 115 and Ser 294) while the type II inductors have share one or two and combines reactivity around Thr 306, Leu 362 and Valine 363.

Further research is needed to explore the role of other cytochrome P-450 isoforms in maintaining homeostasis and disease progression for the development of targeted therapeutic approaches. Advanced computational and experimental methods, such as dynamic modelling and high-throughput screening, can improve the accuracy of predicting the specificity of potential inducers and accelerate the process of new drug development.

Finally, study deepens our understanding of cytochrome P-450 induction by low molecular weight compounds, highlighting the multiplicity and diverse chemical activity of the super enzyme. Molecular docking and statistical analysis provides a sound basis for the integration of artificial intelligence in life science research through machine learning

Table 6: Classification matrix obtained by LOO method.

Inductor Class	Classification Matrix			
	Total Compounds	I	II	% of accuracy
II	28	7	21	75
I	15	12	3	80
Total	43	19	24	76.7

Table 7: Characteristics of the regression model approximating pKs values for type I inductors.

n = 15	Regression summary of dependent variables. pK$_s$ R = 0,916; R^2 = 0,839		
	Coefficient	Standard deviation	p- level
Linear Coefficient	8.561860	2.169997	0.00275
Glu 299	0.317427	0.160451	0.076082
Thr 306	-0.624162	0.132600	0.00833
Leu 362	0.333369	0.132769	0.030865
Val 363	-0.341817	0.157533	0.055188
Standard error of approximation pKs: 0.452.			

to enhance precision medicine and future studies aimed at optimizing enzyme modulation for clinical benefits.

Conflict of Interests

Authors Andrey Ivanovich and Terzulum Gwaza have no financial interests to declare. Andrey Ivanovic is a professor and lead researcher at the N. M Keinzer Engineering School of New Production Technologies.

References

1. Saratikov AC, Fkhmedzhanov RR, Bakiaev AA, Khlebinkov AI, Novozheeva TI, Bystritsky EL. Regulators of enzymatic detoxification systems among nitrogen-containing compounds. Tomsk: Siberian Publishing House. 2002.

2. 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 Jan 20;28(1):38-42. doi: 10.1021/tx500444e. Epub 2014 Dec 19. PMID: 25485457; PMCID: PMC4303333.

3. Testa B, Pedretti A, Vistoli G. Reactions and enzymes in the metabolism of drugs and other xenobiotics. Drug Discov Today. 2012 Jun;17(11-12):549-60. doi: 10.1016/j.drudis.2012.01.017. Epub 2012 Jan 28. PMID: 22305937.

4. de Groot MJ. Designing better drugs: predicting cytochrome P450 metabolism. Drug Discov Today. 2006 Jul;11(13-14):601-6. doi: 10.1016/j.drudis.2006.05.001. PMID: 16793528.

5. Shimizu H, Park SY, Shiro Y, Adachi Si. X-ray structure of nitric oxide reductase (cytochrome P450nor) at atomic resolution. Acta Crystallogr D Biol Crystallogr. 2002 Jan;58(Pt 1):81-9. doi: 10.1107/s0907444901017383. Epub 2001 Dec 21. PMID: 11752781.

6. Almazroo OA, Miah MK, Venkataramanan R. Drug Metabolism in the Liver. Clin Liver Dis. 2017 Feb;21(1):1-20. doi: 10.1016/j.cld.2016.08.001. Epub 2016 Oct 15. PMID: 27842765.

7. Danielson PB. The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. Curr Drug Metab. 2002 Dec;3(6):561-97. doi: 10.2174/1389200023337054. PMID: 12369887.