STRUCTURE AND LIGAND BASED DRUG DESIGNING OF AMINO ACID HYDROXYLASE ENZYMES INVOLVED IN NEURODEGENERATION

A dissertation submitted in partial fulfillment for the award of the degree of

MASTER OF TECHNOLOGY

in

BIOINFORMATICS

by

ATISHA JAIN

03/BINF/2010

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(FORMERLY DELHI COLLEGE OF ENGINEERING)

2010-2012

Supervisors:

Dr. Yasha Hasija
Associate Head & Assistant Prof.
Department of Biotechnology

Dr. Monica Sharma
Assistant Prof.
Department of Biotechnology

Cosupervisor:

Dr. AnjaniTiwari
Scientist 'D'
D.C.R.S (INMAS)

CERTIFICATE

This is to certify that **Atisha jain,** M.Tech, Bioinformatics student of Delhi Technological University (formerly D.C.E) has undergone her 6 Months project work at Institute of Nuclear Medicine and Allied Science (INMAS), DRDO, Delhi.

The project work entitled **"structure and ligand based drug designing of amino acid hydroxylase enzymes involved in neurodegeneration"** embodies the bonafide work carried out under the guidance and direction of the undersigned.

Dr. Anjani Tiwari
Scientist 'D'
INMAS, DRDO

DECLARATION

I, Atisha Jain, hereby declare that the work entitled "**structure and ligand based drug designing of amino acid hydroxylase enzymes involved in neurodegeneration**" has truly been carried out by me under the guidance of Dr. Anjani. K. Tiwari, in I.N.M.A.S, D.R.D.O, Delhi and Dr. Monica Sharma, in D.T.U (formerly D.C.E), Delhi.

This project is a part of partial fulfillment of requirement for the degree of M.Tech in Bioinformatics. This is the original work and has not been submitted for any other degree in any other university.

Atisha Jain

ACKNOWLEDGEMENT

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. I seize this opportunity to express my indeptness and gratitude to **Prof. P. B. Sharma**, Vice Chancellor, D.T.U, Delhi and **Prof. Sager Maji**, H.O.D (department of biotechnology), D.T.U, Delhi for according me permission to carry out the research work

I would like to express my thanks to **Dr. A. K. Mishra** scientist 'F', Head of Department of Cyclotron and Radiopharmaceutical Sciences (DCRS), INMAS, for providing me excellent infrastructure facilities to carry out the present work.

I express my gratitude to **Dr. Anjani Tiwari** scientist 'D', DCRS, INMAS, for his constant encouragement, able guidance and moral support during the course of project work. I am greatly indebted for his valuable guidance. His vital and thought provoking discussions and professional approach were largely responsible for the finalization of the project.

Also, I would like to thank **Dr. Monica Sharma** for her noble and sincere guidance for the work undertaking. I am very thankful to her for her encouragement to bring out the real potentiality of mine.

I express my deep sense of thankfulness to **Dr. Yasha Hasija** whose enthusiastic zeal, cool mindedness boosted me for the successful completion of my work.

Last but not the least I would like to thank my family for their best wishes and colleagues for their constant support and encouragement.

Atisha Jain

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. ABSTRACT

The aromatic amino acid hydroxylases represent a family of structurally and functionally closely related enzymes like phenylalanine hydroxylases, tyrosine hydroxylases and tryptophan hydroxylases. All the enzymes catalyze key steps in important metabolic pathways by utilizing a cofactor known as tetrahydrobiopterin. The purpose of the present study is to undertake a molecular docking study for various analogues of the cofactor on these amino acid hydroxylases using GLIDE application of Maestro (a molecular modeling module of Schrodinger software) and to develop a pharmacophore model for the best screened compounds using PHASE application of Maestro (a molecular modeling module of Schrodinger software). The intermolecular hydrogen bonding interaction of the best-fit ligands are found to be associated with Arg270, Glu280, Thr278, Pro279, Gly346, Ser349, Glu353, Val379 and Fe425 amino acid residue at the phenylalanine hydroxylase receptor active site. Among all the observed interactions with similar binding pattern, ligand 6-(1,4-Dihydroxy-2-methyl-butyl)-2-methyl-7,8-dihydro-3H-pteridin-4-one (3 (f)) showed higher affinity with a glide score of -9.3. For tryptophan hydroxylase receptor active site, interactions are found to be associated with Tyr125, Leu236, Thr265, Pro266, Glu317, Gly333, Ser336, Ser337, Glu340, Thr367 and Thr368 and ligand 2,7-Diamino-6-(4-amino-1,2-dihydroxy-butyl)-7-hydroxy-7,8-dihydro-3H-pteridin-4-one (22 (f)) showed higher affinity with a glide score of -10.1. In addition, pharmacophore mapping studies were undertaken for 24 best screened compounds. Two pharmacophore models were developed, a pharmacophore with 2 H bond donors and 2 H bond acceptors and another with 2 H bond donors, 1 H bond acceptor and 1 aromatic ring. The pharmacophore hypotheses yielded a statistically significant 3-D QSAR model with correlation factors of 0.85 and 0.73 respectively for training and test set compounds. Also, the docking studies performed on mutant phenylalanine hydroxylase shows the therapeutic applications of these cofactor analogues.

2. INTRODUCTION

2.1 BIOLOGICAL INTRODUCTION

To understand the function of a protein, it is necessary to know the physiological system it belongs to. Phenylalanine hydroxylase (PheOH), tyrosine hydroxylase (TyrOH) and tryptophan hydroxylase (TrpOH), all the three enzymes belong to a family of tetrahydro-biopterin dependent aromatic amino acid hydroxylases [1]. All the enzymes catalyze key steps in important metabolic pathways [2]. Thus the enzymes share many physical, structural and catalytic properties.
PheOH catalyses the hydroxylation of phenylalanine to tyrosine (fig 1) and is the rate limiting step in the only pathway to catabolize phenylalanine [3]. TyrOH catalyses the hydroxylation of tyrosine to 3,4-dihydroxyphenylalanine (L-DOPA), which is the rate limiting step in the biosynthesis of catecholamine (fig 2) [4]. And TrpOH is the rate-limiting step in the biosynthesis of 5-hydroxytryptamine (serotonin) (fig 3), and catalyses the hydroxylation of tryptophan to 5-hydroxytryptophan [5].

a. Localization of enzymes

PheOH is mainly found in the liver [3]. TyrOH is present in the central nervous system (CNS), peripheral sympathetic neurons and the adrenal medulla [6]. TrpOH is involved in neurotransmitter synthesis and could be expected to be found in tissues of the nervous system. TrpOH activity has been detected in the brain, the pineal gland and also in the enteric neurons of the gut [7].

b. Biological relevance of enzymes

PheOH catalyzes the rate-limiting step in the breakdown of L-Phe to carbon dioxide and water. The hydroxylation of phenylalanine requires a heme iron and a non protein cofactor tetrahydrobiopterin (BH4) as it promotes the hydroxylation of phenylalanine to tyrosine. It acts as a monooxygenase, incorporating one molecule of oxygen into the amino acid substrate, while the other oxygen atom is reduced to water using BH_4 as the reductant. The active cofactor is the reduced tetrahydrobiopterin, enzyme dihydropteridine reductase catalyses its regeneration by the reduction of the dihydrobiopterin back to BH4 [8]. Phenylalanine (L-Phe) is hydroxylated in the *para* position to tyrosine (L-Tyr) by PheOH and L-Tyr is hydroxylated in the *meta* position by TyrOH to 3,4-dihydroxyphenylalanine (DOPA).

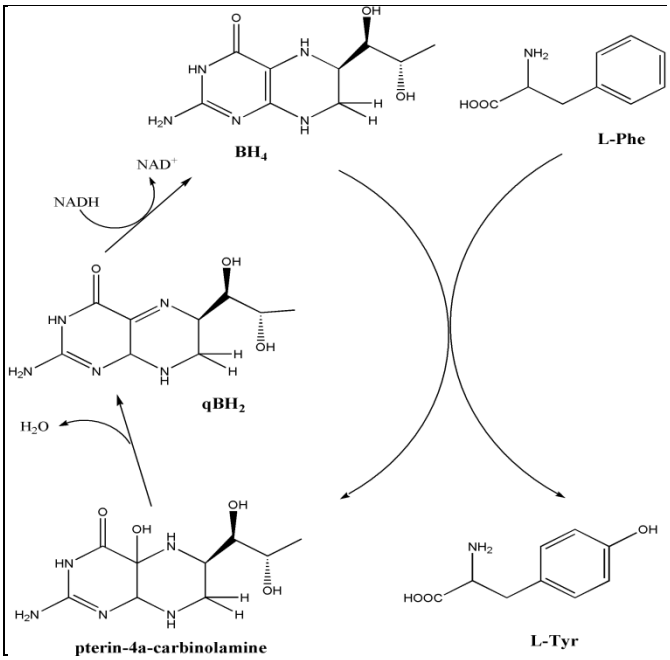


Fig1: Biochemistry of PheOH and reactivation of cofactor

Thus, when there is any defect in the enzymatic activity of PheOH, often due to due to mutations in the PAH gene, this causes hyperphenylalaninemia (HPA), therefore, blood phenylalanine levels increase above 20 times the normal concentration, accumulate in toxic amounts and results in metabolic disease known as phenylketonuria (PKU) [9]. Excessive phenylalanine is metabolized into phenylketones through a transaminase pathway with glutamate forming metabolites like phenyl acetate, phenylpyruvate and phenethylamine [10].
The oral administration of tetrahydrobiopterin can reduce blood levels of this amino acid in certain patients [11]. The company BioMarin Pharmaceutical has produced a tablet preparation of the compound sapropterin dihydrochloride (Kuvan), which is a form of tetrahydrobiopterin. Kuvan is the first drug that can help BH4-responsive PKU patients lower Phe levels to recommended ranges [12]. Some researchers and clinicians working with PKU are finding Kuvan a safe and effective addition to dietary treatment and beneficial to patients with PKU [13].

TyrOH (Tyrosine hydroxylase) catalyzes the conversion of the amino acid L-tyrosineto L-3,4-dihydroxyphenylalanine (L-DOPA). DOPA is a precursor for dopamine, which, in turn, is a precursor for the important neurotransmitters like nor epinephrine (nor adrenaline) and epinephrine (adrenaline) [14].

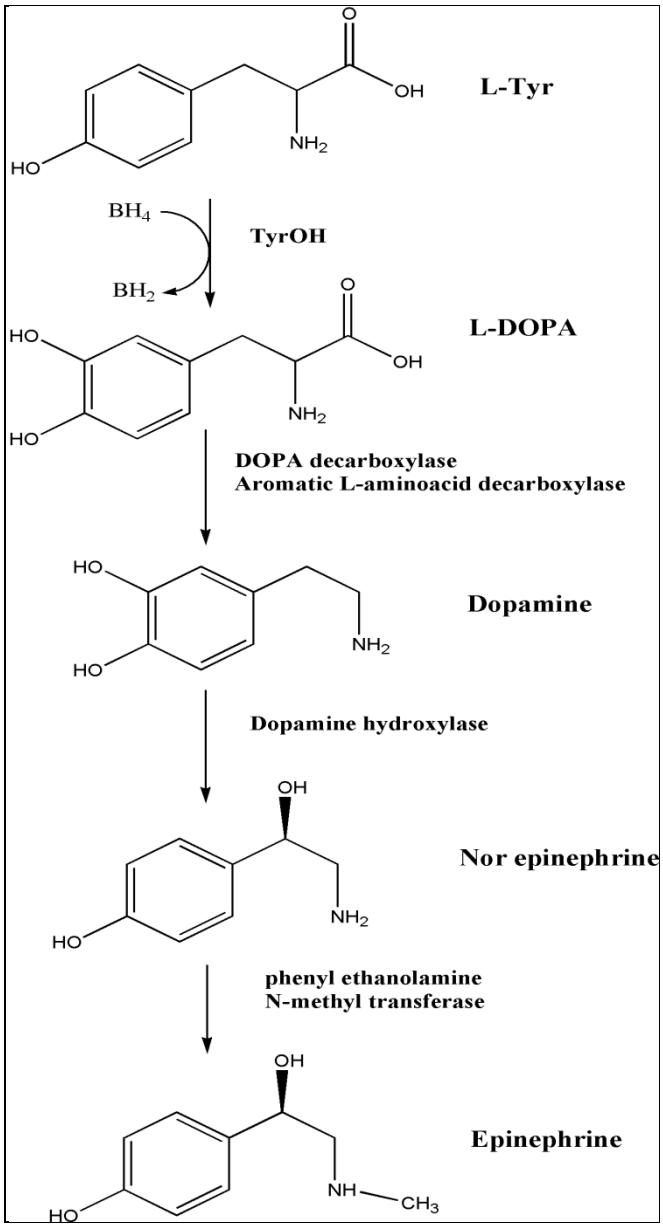


Fig2: Biosynthesis of neurotransmitters: dopamine, norepinephrine and epinephrine

Any alteration in the enzymatic activity of TyrOH leads to disorders such as Parkinson's disease and schizophrenia. Since TyrOH catalyzes the formation of L-DOPA, the rate-limiting step in the biosynthesis of dopamine, enzymatic deficiency leads to the degeneration of dopaminergic neurons in substantia nigra, leading to a reduction of stratial dopamine levels [15].

TrpOH (Tryptophan hydroxylase) oxidizes L-tryptophan to 5-hydroxy-L-tryptophan in the rate-determining step of serotonin biosynthesis. In humans, the stimulation of serotonin production by administration of tryptophan has an antidepressant effect and inhibition of tryptophan hydroxylase may precipitate depression [16].

Serotonin (5-hydroxytryptamine) is a hormone and neurotransmitter that serves regulatory purposes in the central nervous system (CNS) and in several peripheral organs. In the human brain, serotonin is involved in numerous physiological functions, including sleep, pain, appetite, sexual behaviour, and mood, and is the precursor of the pineal hormone melatonin. In addition to its role in the nervous system, serotonin is important for smooth muscle contraction, haemostasis, and intestinal function [17]. It has been implicated in a variety of physiological and pathological functions in CNS. Numerous studies have suggested associations between various neuropsychiatric disorders and genes that modulate central serotonergic neurotransmission, such as the 5-HT transporter, 5-HT receptors, and monoamine oxidases. Therefore, the brain 5-HT system is a major target for several psychiatric disorders such as tricyclic antidepressants, selective serotonin reuptake inhibitors, monoamine oxidase inhibitors, and psychostimulants. Hence any defect in the enzymatic activity of TrpOH leads to neuropsychiatric disorders [18].

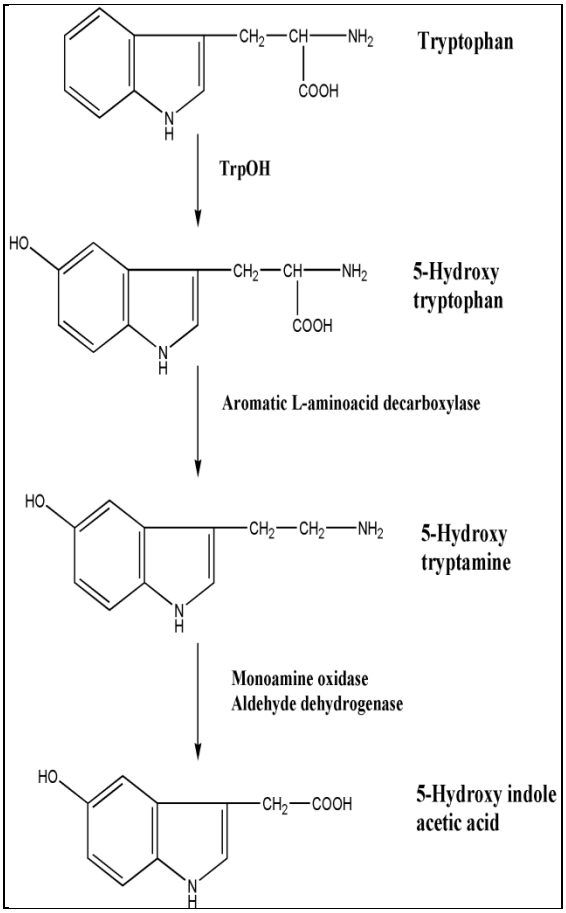


Fig3: Pathway for serotonin synthesis

c. Structural insight into the enzymes

All the three enzymes have a three domain structure: a regulatory N-terminal domain, a catalytic domain and a C-terminal domain. Also they have a common structural motif for the iron active site. The motif has been referred to as a 2-His-1-carboxylate facial triad [19]. The three endogenous iron ligands are completely conserved in all known amino acid hydroxylases. The remaining coordination sites are occupied by water molecules and are accessible to exogenous ligands [20]. All enzymes are activated by phosphorylation of Ser residues in the regulatory domain.

PheOH (pdb id 1DMW from http://www.rcsb.org):

The PheOH monomer (51.9 kDa) consists of three distinct domains: a regulatory N-terminal domain (residues 1-117), the catalytic domain (residues 118-427), and a C-terminal domain (residues 428-453).
The regulatory nature of the N-terminal domain is conferred by its structural flexibility [21]. Hydrogen/deuterium exchanges analysis indicates that allosteric binding of Phe globally alters the conformation of PheOH such that the active site is less occluded as the interface between the regulatory and catalytic domains is increasingly exposed to solvent [22].
The active site consists of an open and spacious pocket lined by hydrophobic residues. Three glutamic acid residues, two histidines, and a tyrosine are also present and critical for pterin- and iron-binding. The pterin binds in the second coordination sphere of the catalytic iron and interacts through several hydrogen bonds to two water molecules coordinated to the iron, as well as to the main chain carbonyl oxygens of Ala322, Gly247, and Leu249 and the main chain amide of Leu249 [16]. The N1 and N8 pterin atoms form hydrogen bonds to the amide backbone of Leu249 at distances of 3.3 and 2.8 Å, respectively. N2 also makes a strong hydrogen bond to the main chain carbonyl oxygen of Gly247 at a distance of 2.8 Å. The dihydroxypropyl side chain interacts with the main chain carbonyl oxygen of Ala322 through a strong hydrogen bond with O10 (2.8 Å) [23].
Fe(II) is coordinated by 3 water molecules (wat1, wat2, wat3), His285, His290, and Glu330, thereby showing an octahedral geometry. Two waters are located approximately equidistant between the iron and O4 of the pterin (Wat3 and Wat1), and the third is on the opposite side of the iron (Wat2), facing His290.Another water molecule is hydrogen bonded to N3 of the pterin ring at a distance of 2.7 Å (Wat4). This water molecule is further hydrogen bonded to Glu286, thus making important connections between this residue and the pterin (fig 4). Glu286 has been identified as a critical residue for pterin function [24].

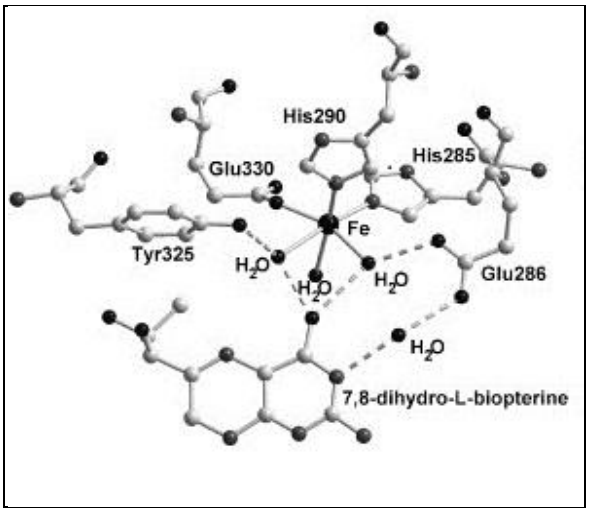


Fig4: Active site of PheOH (pdb id 1DMW) co-crystallized with BH_2

Regulation of the enzymatic activity is provided by Ser16. Some important conformational changes are seen in the active site upon pterin binding. The loop between residues 245 and 250 moves in the direction of the iron, and thus allows for several important hydrogen bonds to the pterin ring to be formed. The loop between residues 245 and 250 shows the largest displacements. The pterin ring forms an aromatic π-stacking interaction with Phe254, and Tyr325 contributes to the positioning of the pterin ring and its dihydroxypropyl side chain by hydrophobic interactions. This explains that the dihydroxypropyl side chain is very essential. Tyr325 also contributes to the correct positioning of the pterin, but has no direct function in the catalytic reaction. In the binary complex, the phenyl ring of Tyr325 establishes hydrophobic contacts with the C3 methyl group of the dihydroxypropyl side chain of the pterin, and thus contributes to the correct positioning of the pterin cofactor for catalysis.
There are several interactions between the pterin and the protein on the other side of the pterin molecule, pointing away from the iron. As mentioned above, this is the region of the protein that undergoes the largest movements upon pterin binding, and is thus involved in recognizing the pterin as a cofactor [25].
The pyrimidine portion of the pterin ring seems most essential for cofactor activity [26]. This can be explained by the strong hydrogen bonds between N2 and N1 and active site amino acid residues (Gly247 and Leu249).

TyrOH (pdb id 2TOH from http://www.rcsb.org):

The TyrOH monomer (51.9 kDa) consists of three distinct domains: a regulatory N-terminal domain (457-498 residues), the catalytic domain (188-456 residues), and a C-terminal domain.
The active site of TyrOH consists of a 17 Å deep cleft at the center of the catalytic domain basket. The cleft is lined primarily by four α-helices (residues 297-304, 329-340, 343-356, and 361-372) (fig 5). The iron is directly coordinated by the "2-his-1-carboxylate facial triad" analogous to PheOH [27].

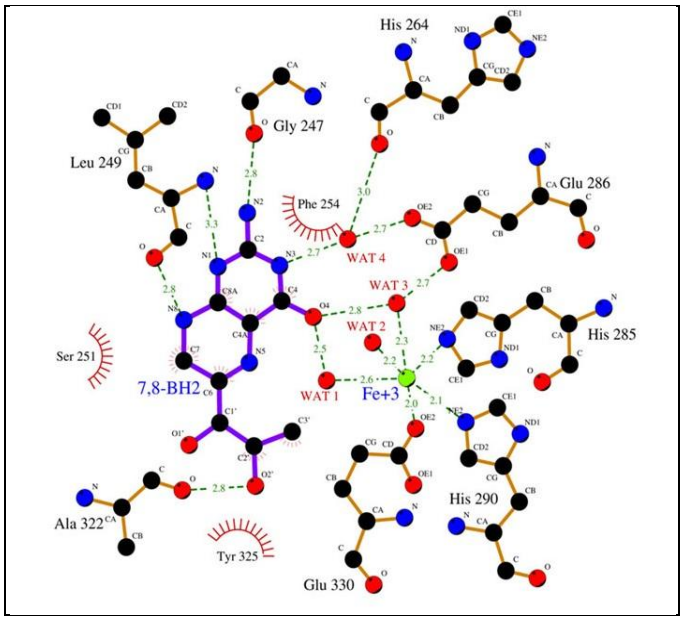


Fig5: Active site of TyrOH (pdb id 2TOH) co-crystallized with BH_2

The N8 pterin atom forms hydrogen bond to the amide backbone of Leu295 (3.1 Å). Pterin also forms hydrogen bonds from O-4 to Tyr371 and Glu376, from the C-1′ OH to the main-chain amides of Leu294 (3.2 Å) and Leu295 (3.2 Å), and from the C-2′ hydroxyl to an iron-coordinating water (3.3 Å). This hydroxyl is also 4.0 Å from the carboxyl of the conserved residue Glu332.

The pterin binds on one face of the large active-site cleft, forming an aromatic π-stacking interaction with Phe300. This phenylalanine residue of TyrOH is found to be hydroxylated in the meta position, most likely through an autocatalytic process, and to consequently form a hydrogen bond to the main-chain carbonyl of Gln310 which anchors Phe300 in the active site. The part of the pterin closest to the iron is the O-4 carbonyl oxygen at a distance of 3.6 Å. The iron is 5.6 Å from the pterin 4a carbon which is hydroxylated in the enzymatic reaction.

The carbonyl at position C-4 of the pterin ring forms two hydrogen bonds with active site residues. One is with the iron-binding residue Glu376 (3.1 Å), and the second is with the highly conserved residue Tyr371 (3.2 Å). Three additional hydrogen bonds are formed by the hydroxypropyl group attached to C-6 [28]. Tyrosine hydroxylase activity is increased in the short term by phosphorylation.

TrpOH (pdb id 1MLW from http://www.rcsb.org):

The structure of TrpOH is similar to TyrOH and PheOH. Few differences are observed in the position of amino acid residues. They are:

The hTrpOH active site consists of an approximately 9 Å deep and 10 Å wide cavity. Lining the active site channel are two loops (residues 263-269 and residues 363-372). The largest positional differences observed in the TrpOH structure, as compared to the rTyrOH and hPheOH structures, are found in these two loops close to the substrate binding region. The pterin π-stacks on Phe241 (3.8 Å) and Tyr235 (3.6 Å), forming hydrogen bonds to Gly234 [BH_2(NH_2)-Gly234(C=O) distance of 3.1 Å] and Leu236 [BH_2(N1)-Leu236(N) distance of 3.0 Å, BH_2(N8)- Leu236(C=O)

distance of 2.7 Å]. The pterin O4 atom is also hydrogen bonded to two of the three water molecules coordinated to the iron, with distances of 2.4 and 2.8 Å. The O4 atom of the pterin ring is 3.3 Å away from the third water molecule coordinated to the iron (fig 6).

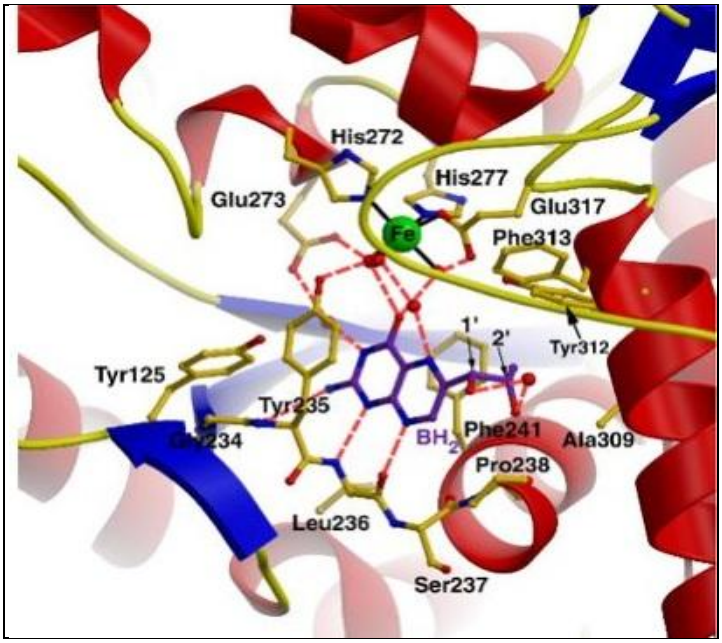


Fig6: Active site of TrpOH (pdb id 1MLW) co-crystallized with BH_2

Glu273 also forms two water-mediated hydrogen bonds to the pterin (one of which is the iron ligand Wat3) at the following distances:
$BH_2(NH_2)$-Wat4-Glu273 distances of 2.9 and 2.7 Å and BH_2(O4)-Wat3-Glu273 distance of 2.8 and 2.7 Å, respectively. The pterin N3 atom also forms a hydrogen bond to Wat4 (2.8 Å).
Catalytic Fe(III) atom is coordinated to His272 (N-Fe distance of 2.1 Å), His277 (N-Fe distance of 2.0 Å), one carboxyl oxygen atom of Glu317 (O-Fe distance of 2.4 Å) and 3 water molecules: Wat1 (axial to His272)-Fe distance of 2.2 Å, Wat2 (axial to His277)-Fe distance of 2.3 Å, and Wat3 (axial to Glu317)-Fe distance of 2.2 Å.
Pro238 is 3.6 Å away from the BH_2 dihydroxypropyl side chain 1′-OH atom. The dihydroxypropyl hydroxyl groups hydrogen bond to a water molecule at equivalent distances (2.8 Å), and the dihydroxypropyl side chain 2′-OH atom also has a long hydrogen bond (3.3 Å) to the carbonyl oxygen of Ala309 [29].

d. Mechanism of hydroxylation

The reactions are divided into two partial reactions, formation of the hydroxylating intermediate and oxygen transfer to the amino acid substrate. The reaction proceeds in 3 steps:

1. Formation of a Fe(II)-O-O-BH_4 bridge-

An iron dioxygen complex is initially formed and stabilized as a resonance hybrid of $Fe^{2+}O_2$ and $Fe^{3+}O_2^-$. The activated O_2 then attacks BH_4, forming a transition state characterized by charge

separation between the electron-deficient pterin ring and the electron-rich dioxygen species. The Fe(II)-O-O-BH4 bridge is subsequently formed [30].

2. Heterolytic cleavage of the O-O bond to yield the ferryl oxo hydroxylating intermediate Fe(IV)=O-

Once formed, the Fe(II)-O-O-BH_4 bridge is broken through heterolytic cleavage of the O-O bond to Fe(IV)=O and 4a-hydroxytetrahydrobiopterin; thus, molecular oxygen is the source of both oxygen atoms used to hydroxylate the pterin ring and phenylalanine (fig 7). Fe(IV)=O intermediate is added to phenylalanine in an electrophilic aromatic substitution reaction that reduces iron from the ferryl to the ferrous state. Also, the evidence strongly supports Fe(IV)=O as the hydroxylating intermediate [31].

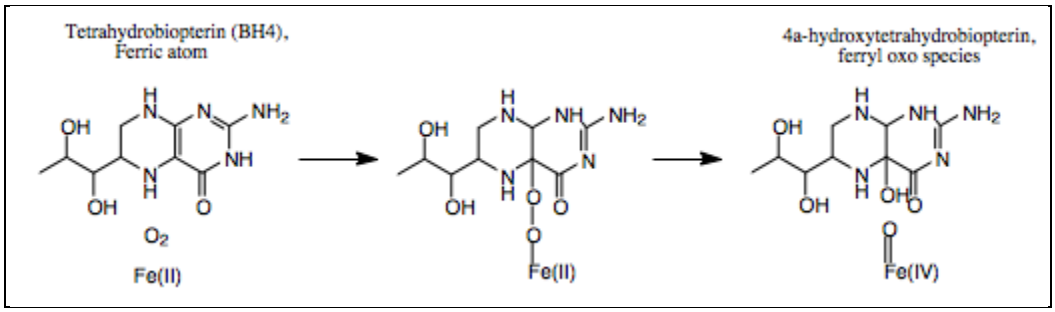


Fig7: Formation and cleavage of Fe(II)-O-O-BH4 bridge to form Fe(IV)=O

3. Attack on Fe(IV)=O to hydroxylate phenylalanine substrate to tyrosine.

The reaction instead proceeds through a cationic intermediate that requires Fe(IV)=O to be coordinated to a water ligand [32]. This cationic intermediate subsequently undergoes a 1,2-hydride NIH shift, yielding a dienone intermediate that then tautomerizes to form the tyrosine product (fig 8) [33].
The pterin cofactor is regenerated by hydration of the carbinolamine product of PheOH to quinonoid dihydrobiopterin (qBH_2), which is then reduced to BH_4.

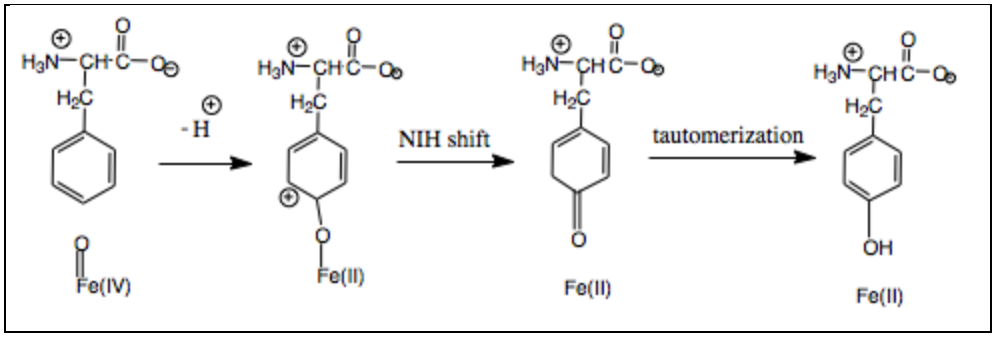


Fig8: Hydroxylation of phenylalanine to tyrosine

2.2 COMPUTATIONAL BACKGROUND

a. Rational Drug Design (RDD)

It is a process used in the biopharmaceutical industry to discover and develop new drug compounds. RDD uses a variety of computational methods to identify novel compounds, design compounds for selectivity, efficacy and safety, and develop compounds into clinical trial candidates. Depending on how much information is available about drug targets and potential drug compounds, these methods fall into several categories:
Structure-based drug design
Ligand-based drug design
De novo design
Homology modeling
Drugs work by interacting with target molecules (receptors) in our bodies and altering their activities in a way that is beneficial to our health. In some cases, the effect of a drug is to stimulate the activity of its target (an agonist) while in other cases the drug blocks the activity of its target (an antagonist). With the advent of greater understanding of bioinformatics & physiological mechanisms it is now possible to rationally design drugs. It is an inventive process of finding new medications based on the knowledge of the biological target. Rational drug design is a more focused approach, which uses information about the structure of a drug receptor or one of its natural ligands to identify or create candidate drugs. The three-dimensional structure of a protein can be determined using methods such as **X-ray crystallography or nuclear magnetic resonance spectroscopy.** The search for small molecules (drugs) that bind to the target begins by screening libraries of potential drug compounds. If the structure of the target is available, a virtual screen may be performed for candidate drugs. Ideally the candidate drug compounds should be "drug-like", that is they should possess properties (drug like properties) such as oral bioavailability, adequate chemical and metabolic stability, and minimal toxic effects [34].
This type of drug designing often relies on computers, so it is also known as Computer-aided drug design. It uses computational chemistry to discover, enhance, or study drugs and related biologically active molecules. The most fundamental goal is to predict whether a given molecule will bind to a target and if so, then how strongly. Molecular mechanics or molecular dynamics are most often used to predict the conformation of the small molecule and to model conformational changes in the biological target that may occur when the small molecule binds to it [35].

b. Molecular modeling

Molecular modeling is the science of representing molecular structures numerically and simulating their behavior with the equations of quantum and classical physics. It is also called computer aided molecular modeling (CAMM). It has proven to be a highly valuable tool for rational drug design. Main applications of CAMM come from structure-based drug design (if 3D information of the target molecule is available) and QSAR (Quantitative Structure Activity Relationship). By using molecular modeling, we will be able to design new and more potent drugs against diseases.

It is a field that is used to model and deduce the information of the system at atomic level. It includes all methodologies used in computational chemistry like computation of the energy of a molecular system, energy minimization, molecular dynamics etc. in turn, this knowledge is aimed at designing new active molecules that can be successfully used as drugs [36].
The role of molecular modeling in drug design has been divided into two separate paradigms, one focuses on the structure-activity problem in the absence of detailed, three-dimensional structural information about the receptor, and the other focuses on understanding the interactions in receptor-ligand complexes and using the known three-dimensional structure of the therapeutic target to design novel drugs [37].

c. Structure-based drug designing- from computer to clinic

We used Structure-based drug design (SBDD) to design and find ligands which are specific for certain known target receptors. SBDD is an iterative process in which macromolecular crystallography has been the predominate technique used to elucidate the three-dimensional structure of drug targets. Although both nucleic acids and proteins are potential drug targets, by far the majority of such targets are proteins. Given that many proteins undergo considerable conformational change upon ligand binding, it is important to design drugs based on the crystallographic structures of protein-ligand complexes, not the unliganded structure. It relies on the knowledge of the three dimensional structure of the biological target obtained through methods such as x-ray crystallography or NMR spectroscopy. It takes on a systematic approach to rationally designing drugs. Determining the structure is crucial because it provides key information about a protein's function and how it may possibly interact on the molecular level with other substances. This knowledge is very useful because it enables a rational and direct approach to developing drugs [38].
In SBDD, scientists use detailed knowledge of the active sites of protein targets associated with particular diseases to design synthetic compounds that fight the disease. The active site of an enzyme is the area into which a chemical or biological molecule fits to initiate a biochemical reaction. SBDD aims to create a molecule that will bind to the active site of a targeted enzyme, thereby preventing the normal chemical reaction and ultimately halting the progression of the disease.
Once a target is selected, researchers use x-ray crystallography to determine the precise three-dimensional molecular structure of the proteins. This structure serves as a blueprint for the drug design of a lead compound. The compounds are modeled for their fit in the active site of the target, considering both steric aspects (i.e., geometric shape) and functional group interactions, such as hydrogen bonding and hydrophobic interactions [39].

d. Structure determination by NMR and X-ray crystallography

A fundamental component of the SBDD protocol is the iterative structure determination process. As each new lead candidate is identified, a new complex structure is required. Clearly this aspect of the process is critically dependent on a very rapid determination of the protein–ligand complex. The two most common methods used to investigate molecular structures are X-ray crystallography (also called X-ray diffraction) and nuclear magnetic resonance (NMR) spectroscopy.
The knowledge of accurate molecular structures is a prerequisite for rational drug design and for structure based functional studies to aid the development of effective therapeutic agents and

drugs. Crystallography can provide the answers to many structure related questions, from global folds to atomic details of bonding. In contrast to NMR (which is a spectroscopic method), no size limitation exists for the molecule or complex to be studied. X-ray crystallography has historically been a major source for obtaining three-dimensional structures of protein–ligand complexes for the iterative drug design cycle. But the use of NMR for the structure elucidation of protein-ligand complexes is a relatively recent addition to the SBDD approach. Specifically, NMR requires extensive isotope labeling of the protein and may take six months to a year using standard methodology to determine a high-resolution structure for proteins <40 kDa. Conversely, X-ray crystallography routinely solves protein–ligand structures in weeks to months and in some cases as fast as a few days. Recent advances in probe technology, software development and NMR methodology show exciting promise in significantly reducing the time requirement to determine a protein structure by NMR [38].

After the structure has been determined, the first category of SBDD methods is about **"finding ligands"** for a given receptor, which is usually referred as **database searching**. In this case, a large number of potential ligand molecules are screened to find those fitting the binding pocket of the receptor. This method is usually referred as **ligand-based drug design**. The key advantage of database searching is that it saves synthetic effort to obtain new lead compounds.

Another category of structure-based drug design methods is about **"building ligands",** which is usually referred as receptor-based drug design. In this case, ligand molecules are built up within the constraints of the binding pocket by assembling small pieces in a stepwise manner. These pieces can be either individual atoms or molecular fragments. The key advantage of such a method is that novel structures, not contained in any database, can be suggested. This is also known as pharmacophore modeling [34].

e. Protein data bank (PDB)

The information regarding the protein-ligand complex can be obtained from Protein Data Bank (PDB). It is a repository for the 3-D structural data of large biological molecules, such as proteins and nucleic acids. The data is obtained by X-ray crystallography or NMR spectroscopy and submitted by biologists and biochemists from around the world. The data is freely available worldwide and are freely accessible on the Internet via the websites of its member organisations [40].

PDB was developed and managed by the Brookhaven National Laboratories but now it is managed & maintained by the RCSB (Research Collaboratory for Structural Bioinformatics). It is a non profit consortium of three organizations: Rutzers, The State University of New Jersey; the San Diego Supercomputer Center at the university of California, San Diego and the National Institute of Standards and Technology. Information on structures can be retrieved from the main PDB website at **http://www.pdb.org/** or one of its major sites **http://www.rcsb.org/pdb/mirrors.html.** PDB is a member of Worldwide Protein Data Bank, wwPDB. It curates and annotates PDB data according to agreed upon standards. The file format initially used by the PDB was called the PDB file format. Around 1996, the "macromolecular Crystallographic Information file" format, mmCIF, started to be phased in. An XML version of this format, called PDBML, was described in 2005. The structure files can be downloaded in any of these three formats. When the author has completed his/her deposition, PDB IDs are automatically assigned by the software. The 3-dimensional co-crystallized structure of the receptor with some ligand could be easily obtained from PDB and studied. Then the information can be further processed for protein preparation, ligand preparation, docking and finally

obtaining the docking score which tells how selective the ligand is for the receptor. Then this analysis can be used for synthesis of drugs [40].

f. Docking

Docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules. It is the Core of the target-based structure-based drug design for lead generation and optimization. Molecular docking is thought of as an optimization problem, which describes the "best-fit" orientation of a ligand that binds to a particular protein of interest. It is similar to *"lock-and-key"* model, where one is interested in finding the correct relative orientation of the *"key"* which will open up the *"lock"*. Thus the protein can be thought of as the "lock" and the ligand can be thought of as a "key".

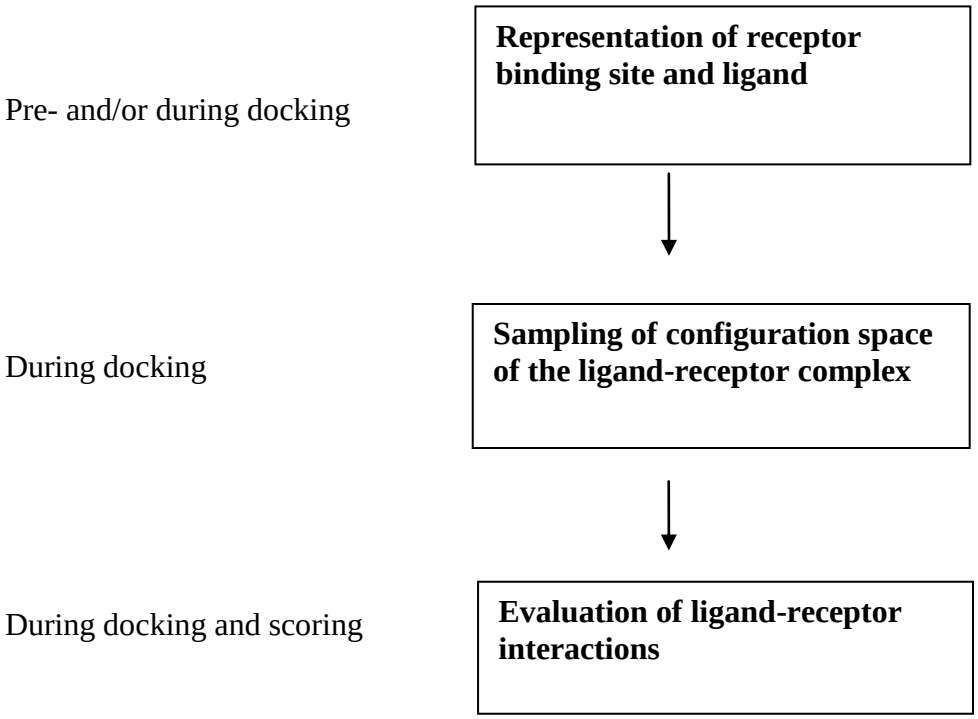


Fig9: Three Components of Docking

Docking is important as a binding interaction between a small molecule ligand and an enzyme protein may result in activation or inhibition of the enzyme. If the protein is a receptor, ligand binding may result in agonism or antagonism. Docking is most commonly used in the field of drug design — most drugs are small organic molecules, and docking may be applied to:

Hit identification – docking combined with a scoring function can be used to quickly screen large databases of potential drugs in silico to identify molecules that are likely to bind to protein target of interest

Lead optimization – docking can be used to predict in where and in which relative orientation a ligand binds to a protein (also referred to as the binding mode or pose). This information may in turn be used to design more potent and selective analogs [38].

GLIDE (grid-based ligand docking with energetics):

Glide has been designed to perform an exhaustive search of the positional, orientational, and conformational space available to the ligand. Glide uses a series of hierarchical filters to search for possible locations of the ligand in the active-site region of the receptor. The shape and properties of the receptor are represented on a grid by different sets of fields that provide progressively more accurate scoring of the ligand pose. The next step produces a set of initial ligand conformations. Given these ligand conformations, initial screens are performed over the entire phase space available to the ligand to locate promising ligand poses. Starting from the poses selected by the initial screening, the ligand is minimized in the field of the receptor using a standard molecular mechanics energy function (OPLS-AA force field [41]). Finally, the lowest-energy poses obtained in this fashion are subjected to a Monte Carlo procedure that examines nearby torsional minima and a GlideScore is generated for use in predicting binding affinity and rank-ordering ligands in database screens [42].

g. Pharmacophore modelling

The concept of pharmacophore was first introduced in 1909 by Paul Ehrlich, who defined the pharmacophore as 'a molecular framework that carries (phoros) the essential features responsible for a drug's (pharmacon) biological activity'. **Ligand-based pharmacophore modeling** has become a key computational strategy for facilitating drug discovery in the absence of a macromolecular target structure. It is usually carried out by extracting common chemical features from 3D structures of a set of known ligands representative of essential interactions between the ligands and a specific macromolecular target. In general, pharmacophore generation from multiple ligands (usually called training set compounds) involves two main steps:

1. Creating the conformational space for each ligand in the training set to represent conformational flexibility of ligands
2. Aligning the multiple ligands in the training set and determining the essential common chemical features to construct pharmacophore models [43].

Molecular alignment is the major challenging issue in ligand based pharmacophore modeling. The alignment methods can be classified into two categories in terms of their fundamental nature: point-based and property-based approaches [44].
The points (in the **point-based method**) can be further differentiated as atoms, fragments or chemical features. In point-based algorithms, pairs of atoms, fragments or chemical feature points are usually superimposed using a least squares fitting. The biggest limitation of these approaches is the need for predefined anchor points because the generation of these points can become problematic in the case of dissimilar ligands.
The **property-based algorithms** make use of molecular field descriptors, usually represented by sets of Gaussian functions, to generate alignments. The alignment optimization is carried out with some variant of similarity measure of the intermolecular overlap of the Gaussians as the objective function.

Another challenging problem lies in the practical task of proper selection of training set compounds. This problem, apparently being simple and non-technical, often confuses users, even

experienced ones. It has been demonstrated that the type of ligand molecules, the size of the dataset and its chemical diversity affect the final generated pharmacophore model considerably [45].

h. Pharmacophore based virtual screening

Once a **pharmacophore model** is generated by either the ligand based or the structure-based approach, it can be used for querying the 3D chemical database to search for potential ligands, which is so-called 'pharmacophore-based virtual screening' (VS). In the pharmacophore-based VS approach, a pharmacophore hypothesis is taken as a template. The purpose of screening is actually to find such molecules (hits) that have chemical features similar to those of the template. Some of these hits might be similar to known active compounds, but some others might be entirely novel in scaffold. Pharmacophore-based VS can be very time-consuming, especially in cases of screening large chemical databases with flexible molecules. After pharmacophore-based VS, the de novo design approach can be used to create completely novel candidate structures that conform to the requirements of a given pharmacophore [43].

In the present study, we have investigated a series of cofactor 7, 8-dihydro-biopterin analogues, which is found to act as a common cofactor for all the three enzymes under study: phenylalanine hydroxylases, tyrosine hydroxylases and tryptophan hydroxylases. We performed molecular docking studies of a series of cofactor analogues with these target proteins. Then, ligand based drug design approaches like **pharmacophore mapping and 3-D QSAR** to develop a pharmacophore model. Here, we describe the development of robust ligand based 3D pharmacophore hypotheses using pharmacophore alignment. The alignment obtained from pharmacophore points was used to derive an atom based 3D-QSAR model. In addition, the best screened compounds were docked with the mutant phenylalanine hydroxylase receptor.

3. REVIEW OF LITERATURE

Robert S**.** Phillips et al have examined the interaction of phenylalanine hydroxylase with phenylalanine, tetrahydropterin cofactors, and an activating phospholipid, lysophosphatidylcholine. They suggested that activation of phenylalanine hydroxylase results in a conformation change and the exposure of buried tryptophan(s) and possibly a cysteine residue [46].

Heidi Erlandsen et al presented the crystal structure of the dimeric catalytic domain (residues 117-424) of human phenylalanine hydroxylase (hPheOH), cocrystallized with various potent and well-known catechol inhibitors and refined at a resolution of 2.0 Å and suggested that the catechols bind by bidentate coordination to each iron in both subunits of the dimer through the catechol hydroxyl groups, forming a blue-green colored ligand-to-metal charge-transfer complex. Crystallographic comparison with the structurally related rat tyrosine hydroxylase binary complex with the oxidized cofactor 7,8-dihydrobiopterin revealed overlapping binding sites for the catechols and the cofactor, compatible with a competitive type of inhibition of the catechols versus BH4 [47].

The conformation and distances to the catalytic iron of both L-Phe and the cofactor analogue L-erythro-7,8-dihydrobiopterin (BH2) simultaneously bound to recombinant human PAH have been estimated by 1H NMR by Knut Teigen et al. They demonstrated that the pterin ring of BH2 π-stacks with Phe254, and the N3 and the amine group at C2 hydrogen bond with the carboxylic group of Glu286. The ring also establishes specific contacts with His264 and Leu249. The distance between the O4 atom of BH2 and the iron (2.6 Å) is compatible with coordination. Also the hydroxyl groups in the side-chain at C6 hydrogen bond with the carbonyl group of Ala322 and the hydroxyl group of Ser251, an interaction that seems to have implications for the regulation of the enzyme by substrate and cofactor. Some frequent mutations causing PKU are located at residues involved in substrate and cofactor binding. These are adequate for the intercalation of iron-coordinated molecular oxygen, in agreement with a mechanistic role of the iron moiety both in the binding and activation of dioxygen and in the hydroxylation reaction [48].

Torben Gjetting et al showed that the majority of *PAH* missense mutations impair enzyme activity by causing increased protein instability and aggregation. Naturally occurring N-terminal *PAH* mutations are distributed in a nonrandom pattern and cluster within residues 46–48 and 65–69, two motifs highly conserved in PDH. Their work suggested that impairment of phenylalanine-mediated activation of PAH may be an important disease-causing mechanism of some N-terminal *PAH* mutations, which may explain some well-documented genotype-phenotype discrepancies in PAH deficiency [49].

Phenylalanine hydroxylase from a bacterial source Chromobacterium violaceum phenylalanine hydroxylase (CvPheOH) has been structurally characterized by Joo Y. Kim at high resolution and compared to the human analog . The bacterial enzyme displayed higher activity and thermal melting temperature, and structurally, differences were observed in the N and C termini, and in a loop close to the active-site iron atom [50].

The crystal structures of the catalytic domain of human phenylalanine hydroxylase (hPheOH) in complex with the physiological cofactor 6(R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH4) and the substrate analogues 3-(2-thienyl)-L-alanine (THA) or L-norleucine (NLE) have also been determined at 2.0A ° resolution by Ole Andreas Andersen et al. Both structures demonstrate that substrate binding triggers structural changes throughout the entire protomer, including the

displacement of Tyr138 from a surface position to a buried position at the active site, with a maximum displacement of 20.7 A ° for its hydroxyl group. The carboxyl and amino groups of THA and NLE are positioned identically in the two structures, supporting the conclusion that these groups are of key importance in substrate binding, thus explaining the broad non-physiological substrate specificity observed for artificially activated forms of the enzyme [51].

Activity of tyrosine hydroxylase is regulated by feedback inhibition and inactivation by catecholamines, and activation by protein phosphorylation. The regulatory domain of tyrosine hydroxylase contains multiple serine (Ser) residues that are phosphorylated by a variety of protein kinases [52]. Long term regulation of tyrosine hydroxylase can also be mediated by phosphorylation mechanisms. Hormones (e.g glucocorticoids), drugs (e.gcocaine), or second messengers such as cAMP increase tyrosine hydroxylase transcription. Increase in tyrosine hydroxylase activity due to phosphorylation can be sustained by nicotine for up to 48 hours [53]. Reaction mechanisms for the conversion of tyrosine hydroxylase to an inactive/stable form by catecholamines, and activation of tyrosine hydroxylase by phosphorylation at Ser-40 were discussed by Hitoshi Fujisawa et al. [54]. Also, a study by S. Colette Daubner demonstrated the role of polypeptide loop in tyrosine hydroxylase (TyrH) whose homolog in phenylalanine hydroxylase (PheH) takes on a different conformation when substrates are bound has been studied using sitedirected mutagenesis. Mutagenesis of residues in the center of the loop resulted in alterations in the KM values for substrates, the Vmax value for DOPA synthesis, and the coupling of tetrahydropterin oxidation to tyrosine hydroxylation. The variant with the most altered KM value for 6-methyltetrahydropterin was TyrH F184A. The variants with the most affected Ktyr values were those with substitutions in the center of the loop, TyrH K183A, F184A, D185A, P186A and D187A. These variants also had the most reduced Vmax values for DOPA synthesis and thus play a dominant role in determining amino acid substrate specificity [55].

Using rational site-directed mutagenesis, G. C. T. Jiang determined that Tyr235 (Y235), within the active site of TPH, appears to be involved as a tryptophan substrate orienting residue [56]. A stable N-terminally truncated form of human TPH that includes the catalytic domain was prepared and characterized by Jeffrey McKinney et al. The conformation and distances to the catalytic non-heme iron of both L-Trp and the tetrahydrobiopterin cofactor analogue L-*erythro*-7,8-dihydrobiopterin (BH2) were also determined by using 1H NMR spectroscopy and the bound conformers of the substrate and the pterin were then docked into the modeled three-dimensional structure of TPH. Finally they concluded that L-Trp binds to the enzyme through interactions with Arg257, Ser336, His272, Phe318, and Phe313, and the ring of BH2 interacts mainly with Phe241 and Glu273. The distances between the hydroxylation sites at C5 in L-Trp and C4a in the pterin, i.e., 6.1 (0.4 Å), and from each of these sites to the iron, i.e., 4.1 (0.3 Å) and 4.4 (0.3 Å), respectively, were also in agreement with the formation of a transient iron-4a-peroxytetrahydropterin in the reaction, as proposed for the other hydroxylases [57]. Insight into the nature of oxygen activation in tryptophan hydroxylase has been obtained from density functional computations. Conformations of O2-bound intermediates have been studied with oxygen trans to glutamate and histidine, respectively by Kasper P. Jensen et al. The weaker trans influence of histidine has shown to give rise to a bent O2 coordination mode with O2 pointing towards the cofactor and a more activated O–O bond (1.33 Å) than in Oglu (1.30 Å). It has been shown that the cofactor can hydrogen bond to O2 and activate the O–O bond further (from 1.33

to 1.38 Å). The Ohis intermediate leads to a ferryl intermediate (Fhis) with an isomer shift of 0.34 mm/s, which is also consistent with the experimental value (0.25 mm/s) which was proposed to be the structure of the hydroxylating intermediate, with the tryptophan substrate well located for further reaction 3.5 Å from the ferryl group [58].

4. OBJECTIVE

The objective of the present study is:
(i) Screening of more effective ligand for in-vivo binding with amino acid hydroxylase enzymes.
(ii) Generation of a 3-D QSAR mathematical model to produce pharmacophore features for the best screened molecules.
(iii) Designing of a new molecular entity on the basis of docking and pharmacophores modeling to reduce the cost-effectiveness for designing drugs for neurodegeneration.

5. METHODOLOGY

1. To perform a docking screen, the first requirement is a structure of the protein of interest and to identify the active site of protein. So, the first step was to search for a suitable protein structure from the website of RCSB PDB by applying various parameters like resolution, bound ligand etc.
All the protein structures studied were downloaded in the PDB file format for the ease of study in the available software for molecular modeling known as the **SCHRODINGER SOFTWARE**.

2. Next step is to ensure the chemical correctness and to optimize the protein structures for use with Glide. This is done via **protein preparation wizard** application of **maestro**.

3. All the ligand structures were drawn in **CHEMDRAW** and the Schrödinger ligand preparation product **LigPrep** was used to prepare high quality, all atom 3D structures for large numbers of drug-like molecules, starting with 2D or 3D structures.

4. In order to predict the preferred orientation of one molecule to a second, when bound to each other and to predict the affinity of the binding, ligand-receptor docking was performed. The ligands were docked with the protein PDB files using GLIDE application of maestro. Thus, we could visualize the docking results in terms of Gscore, electrostatic, hydrophobic and hydrophilic interactions.

5. For pharmacophore perception, structure alignment and activity prediction, Pharmacophore modeling was carried out to generate a pharmacophore hypothesis using PHASE application of maestro.

6. Finally, to study the interactions between the compounds and the mutant phenylalanine hydroxylase, ligand-receptor docking was again performed with the mutant PDB ID.

Molecular modeling study

Molecular modeling investigations were carried out using Maestro software package. Coordinates for the protein structure were taken from the RCSB Protein Data Bank (PDB) and prepared using the Protein Preparation Wizard, which is part of the Maestro software package (Maestro, v9.2, Schrodinger, LLC, NewYork, NY).

1. Experimental procedure of Docking:

a. Dataset of compounds:

A dataset of 47 compounds consisting of different 7,8-dihydro-biopterin analogues were used in this study. Correct empirical structure of ligands were drawn (by using CHEM DRAW 7th VERSION) and converted into MOL-SD file.

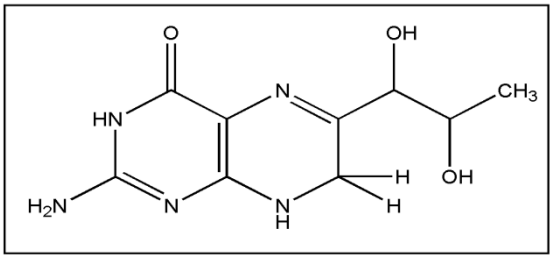


Fig10: Structure of BH_2

Since, it was found from the literature that tetrahydro-biopterin (BH_4) and its analogue 7,8-dihydro-biopterin are obligatory enzymatic cofactors for amino acid hydroxylases, so we kept the basic skeleton of the cofactor as it is and did some modifications at different positions. Figures show the structures of the cofactor with positions 1, 2, 3 and 4, 5, 6 respectively. Modifications in the structure of cofactor were done by substituting electron donating groups at these positions.

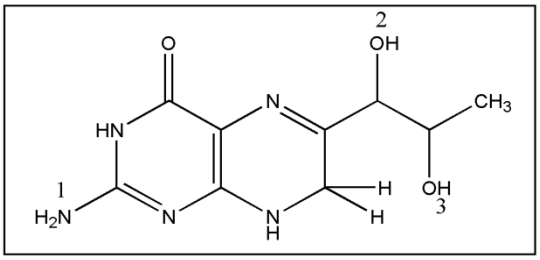


Fig11: Structure of BH_2 showing positions 1, 2 and 3

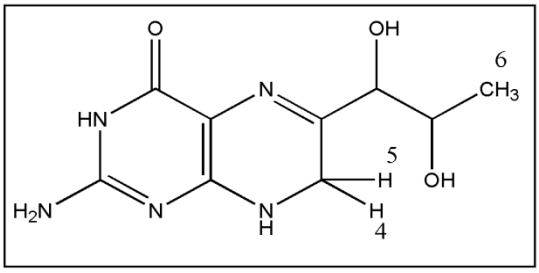


Fig12: Structure of BH_2 showing positions 4, 5 and 6

Prior to docking, a single low-energy 3D structure was generated for each ligand with the aid of LigPrep.

b. Preparation of Ligand:

Once imported the MOL file into the Maestro workspace, Ligprep were used for preparing the ligands. Thus this step left us with the ligand with a minimized ligand molecule ready to dock with the protein.

The Schrodinger package was used to prepare the protein and ligand data required for the docking. **LigPrep** is a robust collection of tools designed to prepare high quality, all-atom 3D structures for large numbers of drug-like molecules, starting with 2D or 3D structures. It is a utility of Schrodinger software suit that combines tools for generating 3D structures from 1D (Smiles) and 2D (SDF) representation, searching for tautomers, steric isomers and performs a geometry minimization of the ligands. LigPrep can also produce a number of structures from each input structure with various ionization states, tautomers, stereochemistries, and ring conformations, and eliminate molecules using various criteria including molecular weight or specified numbers and types of functional groups present.
The LigPrep process consists of a series of steps that perform conversions, apply corrections to the structures, generate variations on the structures, eliminate unwanted structures, and optimize the structures. This step removes unwanted molecules, add hydrogens, and minimize the ligand structure [59].

c. Preparation of the receptor:

PDB structures with PDB ID (1DMW, 1TDW, 2TOH) from www.rcsb.org were downloaded and refined, prepared using Schrodinger protein preparation wizard tool, which performs the steps including assigning of bond orders, addition of hydrogens, optimization of hydrogen bonds by flipping amino side chains, correction of charges, and minimization of the protein complex. All the bound water molecules, ligands and cofactors were removed (preprocess) from the proteins which were taken in .mae format. The tool neutralized the side chains that are not close to the binding cavity and do not participate in salt bridges.
This step is then followed by restrained minimization of co-crystallized complex, which reorients side chain hydroxyl groups and alleviates potential steric clashes.
For each structure, a brief relaxation was performed using an all-atom constrained minimization carried out with the Impact Refinement module (Impref). Vander-walls radius scaling was taken as default absolute value less than the specified cutoff. Hydrogen atoms were added to the receptor atoms and input partial charges were assigned to receptor. Thus after this step, we had refined, hydrogenated structures of the ligand – receptor complex.

d. Receptor grid generation:

The receptor grid can be set up and generated from the Receptor Grid Generation panel. This step is necessary because Ligand docking jobs cannot be performed until the receptor grids have been generated. Also, the receptor grid generation requires a "prepared" structure: an all-atom structure with appropriate bond orders and formal charges.

e. Ligand-Receptor Docking:

Docking experiments were carried out by using GLIDE (grid based ligand docking with energetic) and ligand docking program from calculations in extra precision (XP) mode. **GLIDE** stands for **Grid-based Ligand Docking with Energetics**. Glide searches for favorable interactions between one or more ligand molecules and a receptor molecule, usually a protein. Glide can be run in rigid or flexible docking modes. The combination of position and orientation of a ligand relative to the receptor, along with its conformation in flexible docking, is referred to

as a *ligand pose*. The ligand poses that Glide generates pass through a series of hierarchical filters that evaluate the ligand's interaction with the receptor.

The docking was initiated with putting specified receptor grid and prepared ligand molecule together. GLIDE was run in the flexible docking mode which automatically generates conformations for each input ligand. The ligand poses that GLIDE generates pass through a series of hierarchial filters that evaluate the ligand's interaction with the receptor. The initial filters test the spatial fit of the ligand to the defined active site, and examine the complentarity of ligand-receptor interactions using a grid based method pattern after the empirical chemscore function. Poses that pass their initial screens enter the final stage of the algorithm, which involves evaluation and minimization of a grid approximation to the OPLS-AA non-bonded ligand receptor interaction energy. Final scoring was then carried out on the energy minimized poses [41].
The docking pattern for the various ligands and the docking scores were collected using the show-table tool.

f. Those 25 compounds showing the higher binding affinity i.e. higher glide score (more than -7.5) with the protein are subjected to further modifications by extending the bond length at positions 1 to 6.

g. Docking was again performed between these modified ligands and all the 3 PDB IDs by repeating steps 2 and 5. Using the XP-visualiser of Glide, the interactions between the ligand and the receptor were studied and the corresponding images were saved.

2. Experimental procedure of pharmacophore modeling

Pharmacophore modeling was carried out using PHASE which is also a part of the Maestro software package (Maestro, v9.2, Schrodinger, LLC, NewYork, NY) for those 24 compounds which were found to have higher binding affinities.

a. **Generating conformers:**
This step Clean up the ligand structures, generate variations on stereochemistry or ionization state and generate sets of conformers for each ligand. Then we defined the pharm (active) set and the inactive set by setting the activity thresholds.

b. **Creating pharmacophore sites:**

For purposes of pharmacophore model development, each ligand structure is represented by a set of points in 3D space, which coincide with various chemical features that may facilitate non-covalent binding between the ligand and its target receptor. These are called pharmacophore sites. From the total 24 compounds, 20 were randomly chosen for training set and remaining 4 for test set by using "automated random selection" option present in PHASE software. PHASE provides six built-in types of pharmacophore features: hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobe (H), negative ionizable (N), positive ionizable (P), and aromatic ring (R). The pharmacophore model was developed using a set of pharmacophore features to generate sites for all compounds. For QSAR development, pharmacophore models of training set

molecules were placed into regular grid of cubes, with each cube allotted zero or more bits to account for the different type of pharmacophore features in the training set that occupy the cube. This representation gives rise to binary valued occupation patterns that can be used as independent variables to create Partial Least Square (PLS) factors 3-D QSAR models.

c. **Scoring hypotheses:**

The process of scoring with respect to actives is designed to filter out inappropriate pharmacophores. Hypotheses are assigned a score comprised of geometric and heuristic factors that can be weighted according to the user's preference. Each pharmacophore is treated temporarily as a reference in order to assign a score. Accordingly, all other non-reference pharmacophores are aligned, one-by-one, to the reference pharmacophore, and the quality the alignments are measured using two criteria:
(1) The root-mean-squared deviation (RMSD) in the site point positions and
(2) The average cosine of the angles formed by corresponding pairs of vector features (acceptors, donors and aromatic rings). These factors are combined with separate weights to yield a combined site + vector score for each non-reference pharmacophore i that's been aligned to the reference. After all pharmacophores have been treated as a reference, the one yielding the highest Reference_ Score is selected as the hypothesis.
The ligand that contributes the reference pharmacophore is referred to as the reference ligand for that hypothesis.

d. **Building 3D QSAR models**
If a sufficient number of molecules of varying activity are available, a 3D QSAR model can be developed for each hypothesis using training set structures that match the pharmacophore on three or more sites [60].
The data related to distance and angle measurements were studied, recorded and the corresponding images were saved.

3. Finally docking was again performed between these 24 compounds and mutant PDB ID (1TDW) using same procedure as in step1 and the interactions were studied using XP Visualiser.

6. RESULTS AND DISCUSSION

Since one of the main goal of drug discovery is the identification of small molecules that can act as lead molecules, we have investigated a series of cofactor 7, 8-dihydro-biopterin analogues, which is found to act as a common cofactor for all the three enzymes under study: phenylalanine hydroxylases, tyrosine hydroxylases and tryptophan hydroxylases. The study mainly focuses at identification of small molecules that exhibit high binding affinity for all the target proteins. For this purpose, the approach that we followed was the docking study of all the designed cofactor analogues via Glide, selection of the best suited candidates according to their binding affinity and then using those selected candidate molecules to prepare a pharmacophore model via Phase. [61]

1. Glide docking analysis

Glide performs grid-based ligand docking with energetics and searches for favourable interactions between one or more typically small ligand molecules and a typically larger receptor molecule, usually a protein. For each of the 3 amino acid hydroxylases: phenylalanine hydroxylases (1DMW), tyrosine hydroxylases (2TOH) and tryptophan hydroxylases (1TDW), it was found that a common cofactor analogue 7,8-dihydro-biopterin was bound in all the protein structures as obtained from protein data bank (PDB).
For each of the PDB IDs: 1DMW, 2TOH and 1MLW, a set of 47 compounds were assessed for binding affinity. So, a total of 141 compounds were docked into the binding sites of these enzymes.
They are as follows:

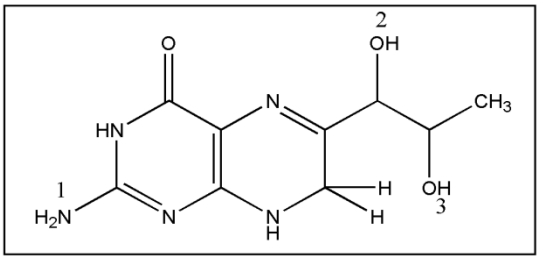


Fig11: Structure of BH_2 showing positions 1, 2 and 3

COMPOUND	POSITION 1	POSITION 2	POSITION 3	DOCKING SCORE
1.	**CH_3**	**NH_2**	**NH_2**	**-8.0**
2.	**NH_2**	**CH_3**	**NH_2**	**-7.6**
3.	NH_2	NH_2	CH_3	-6.6
4.	OH	OH	CH_3	-7.2
5.	OH	CH_3	OH	-7.2
6.	**CH_3**	**OH**	**OH**	**-8.0**
7.	NH_2	CH_3	CH_3	-6.4
8.	CH_3	NH_2	CH_3	-6.7
9.	CH_3	CH_3	NH_2	-6.8
10.	OH	NH_2	OH	-7.2
11.	**OH**	**OH**	**NH_2**	**-8.3**
12.	CH_3	CH_3	OH	-7.4
13.	CH_3	OH	CH_3	-7.4
14.	OH	CH_3	CH_3	-6.5
15.	**OH**	**NH_2**	**NH_2**	**-7.7**
16.	NH_2	OH	NH_2	-7.2
17.	NH_2	NH_2	OH	-7.2
18.	CH_3	OH	NH_2	-7.3
19.	CH_3	NH_2	OH	-7.2
20.	**OH**	**NH_2**	**CH_3**	**-8.2**
21.	OH	NH_2	CH_3	-6.6
22.	NH_2	OH	CH_3	-7.2
23.	NH_2	CH_3	OH	-7.0

Table1: Docking results of compounds with modified groups at different positions with PDB ID 1DMW

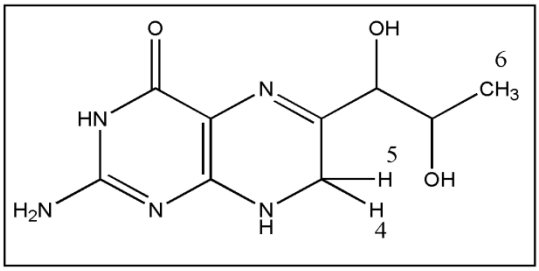


Fig12: Structure of BH_2 showing positions 4, 5 and 6

COMPOUND	POSITION 4	POSITION 5	POSITION 6	DOCKING SCORE
24.	CH_3	NH_2	NH_2	-6.6
25.	NH_2	CH_3	NH_2	-5.9
26.	NH_2	NH_2	CH_3	-6.0
27.	OH	OH	CH_3	-6.8
28.	OH	CH_3	OH	-6.3
29.	CH_3	OH	OH	-6.3
30.	NH_2	CH_3	CH_3	-6.1
31.	CH_3	NH_2	CH_3	-6.2

32.	CH_3	CH_3	NH_2	-6.2
33.	NH_2	OH	OH	-6.5
34.	OH	NH_2	OH	-6.8
35.	OH	OH	NH_2	-7.2
36.	CH_3	CH_3	OH	-6.6
37.	CH_3	OH	CH_3	-6.4
38.	OH	CH_3	CH_3	-6.4
39.	**OH**	**NH_2**	**NH_2**	**-7.9**
40.	NH_2	OH	NH_2	-6.5
41.	NH_2	NH_2	OH	-6.1
42.	CH_3	OH	NH_2	-6.7
43.	CH_3	NH_2	OH	-6.4
44.	OH	NH_2	CH_3	-6.7
45.	OH	NH_2	CH_3	-5.8
46.	NH_2	OH	CH_3	-6.1
47.	NH_2	CH_3	OH	-6.2

Table2: Docking results of compounds with modified groups at different positions with PDB ID 1DMW

COMPOUND	POSITION 1	POSITION 2	POSITION 3	DOCKING SCORE
48.	**CH_3**	**NH_2**	**NH_2**	**-8.3**
49.	NH_2	CH_3	NH_2	-7.4
50.	NH_2	NH_2	CH_3	-7.4
51.	**OH**	**OH**	**CH_3**	**-7.6**
52.	**OH**	**CH_3**	**OH**	**-7.6**
53.	**CH_3**	**OH**	**OH**	**-7.9**
54.	**NH_2**	**CH_3**	**CH_3**	**-7.6**
55.	**CH_3**	**NH_2**	**CH_3**	**-7.7**
56.	**CH_3**	**CH_3**	**NH_2**	**-7.7**
57.	OH	NH_2	OH	-6.9
58.	OH	OH	NH_2	-7.5
59.	**CH_3**	**CH_3**	**OH**	**-7.9**
60.	**CH_3**	**OH**	**CH_3**	**-8.0**
61.	**OH**	**CH_3**	**CH_3**	**-7.9**
62.	**OH**	**NH_2**	**NH_2**	**-10.2**
63.	NH_2	OH	NH_2	-6.8
64.	NH_2	NH_2	OH	-6.7
65.	CH_3	OH	NH_2	-6.9
66.	**CH_3**	**NH_2**	**OH**	**-7.7**
67.	OH	CH_3	NH_2	-7.4
68.	OH	NH_2	CH_3	-6.7
69.	NH_2	OH	CH_3	-7.5
70.	**NH_2**	**CH_3**	**OH**	**-7.6**

Table3: Docking results of compounds with modified groups at different positions with PDB ID 1MLW

COMPOUND	POSITION 4	POSITION 5	POSITION 6	DOCKING SCORE
71.	CH_3	NH_2	NH_2	-4.5
72.	NH_2	CH_3	NH_2	-4.5
73.	NH_2	NH_2	CH_3	-5.5
74.	OH	OH	CH_3	-6.3
75.	OH	CH_3	OH	-5.5
76.	CH_3	OH	OH	-5.5
77.	NH_2	CH_3	CH_3	-4.6
78.	CH_3	NH_2	CH_3	-5.4
79.	CH_3	CH_3	NH_2	-4.7
80.	NH_2	OH	OH	-5.1
81.	OH	NH_2	OH	-5.4
82.	**OH**	**OH**	**NH_2**	**-9.2**
83.	CH_3	CH_3	OH	-5.4
84.	CH_3	OH	CH_3	-4.8
85.	OH	CH_3	CH_3	-4.6
86.	**OH**	**NH_2**	**NH_2**	**-7.8**
87.	**NH_2**	**OH**	**NH_2**	**-7.8**
88.	NH_2	NH_2	OH	-4.8
89.	**CH_3**	**OH**	**NH_2**	**-7.9**
90.	CH_3	NH_2	OH	-5.2
91.	**OH**	**NH_2**	**CH_3**	**-7.8**
92.	OH	NH_2	CH_3	-4.7
93.	NH_2	OH	CH_3	-5.1
94.	NH_2	CH_3	OH	-5.1

Table4: Docking results of compounds with modified groups at different positions with PDB ID 1MLW

COMPOUND	POSITION 1	POSITION 2	POSITION 3	DOCKING SCORE
95.	CH_3	NH_2	NH_2	-7.2
96.	NH_2	CH_3	NH_2	-6.2
97.	NH_2	NH_2	CH_3	-6.6
98.	OH	OH	CH_3	-5.1
99.	OH	CH_3	OH	-5.0
100.	CH_3	OH	OH	-6.1
101.	NH_2	CH_3	CH_3	-4.7
102.	CH_3	NH_2	CH_3	-6.4
103.	CH_3	CH_3	NH_2	-6.1
104.	OH	NH_2	OH	-6.7
105.	OH	OH	NH_2	-6.5
106.	CH_3	CH_3	OH	-5.4
107.	CH_3	OH	CH_3	-5.3
108.	OH	CH_3	CH_3	-4.6
109.	OH	NH_2	NH_2	-7.0
110.	NH_2	OH	NH_2	-6.2
111.	NH_2	NH_2	OH	-6.7
112.	CH_3	OH	NH_2	-6.2
113.	CH_3	NH_2	OH	-6.9
114.	OH	CH_3	NH_2	-6.1
115.	OH	NH_2	CH_3	-6.2

116.	NH_2	OH	CH_3	-5.2
117.	NH_2	CH_3	OH	-5.1

Table5: Docking results of compounds with modified groups at different positions with PDB ID 2TOH

COMPOUND	POSITION 4	POSITION 5	POSITION 6	DOCKING SCORE
118.	CH_3	NH_2	NH_2	-6.8
119.	NH_2	CH_3	NH_2	-6.7
120.	NH_2	NH_2	CH_3	-5.1
121.	OH	OH	CH_3	-5.2
122.	OH	CH_3	OH	-5.2
123.	CH_3	OH	OH	-5.2
124.	NH_2	CH_3	CH_3	-5.6
125.	CH_3	NH_2	CH_3	-5.6
126.	CH_3	CH_3	NH_2	-5.6
127.	NH_2	OH	OH	-7.0
128.	OH	NH_2	OH	-6.6
129.	OH	OH	NH_2	-6.6
130.	CH_3	CH_3	OH	-4.6
131.	CH_3	OH	CH_3	-4.7
132.	OH	CH_3	CH_3	-4.8
133.	OH	NH_2	NH_2	-7.1
134.	NH_2	OH	NH_2	-7.0
135.	NH_2	NH_2	OH	-5.0
136.	CH_3	OH	NH_2	-6.3
137.	CH_3	NH_2	OH	-6.4
138.	OH	NH_2	CH_3	-6.3
139.	OH	NH_2	CH_3	-6.6
140.	NH_2	OH	CH_3	-6.3
141.	NH_2	CH_3	OH	-6.3

Table6: Docking results of compounds with modified groups at different positions with PDB ID 2TOH

All the 141 compounds showed a wide range of binding affinity with the different PDB IDs in terms of glide score (Kcal/mol). Those 25 compounds showing the higher binding affinity i.e. higher glide score (more than -7.5) with the protein are subjected to further modifications by extending the bond length at positions 1 to 6. This is because more negative the glide score, indicates better fitting to the receptor active sites [43]. List of interactions of these higher affinity compounds with the protein and the interacting residues is given in table 7 to table 10.

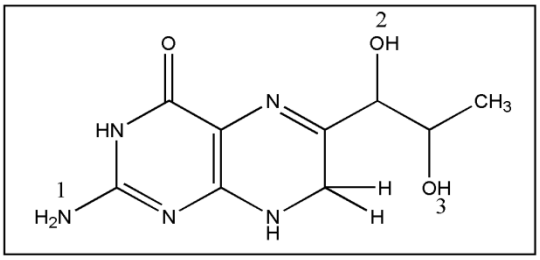


Fig11: Structure of BH_2 showing positions 1, 2 and 3

COMPOUND	POSITION 1	POSITION 2	POSITION 3	DOCKING SCORE	RESIDUES and DISTANCE (Å)						
					FE 425	GLU 280	THR 278	VAL 379	SER 349	GLU 353	ARG 270
1.	**CH_3**	**NH_2**	**NH_2**	**-8.0**							
a.	C_2H_5	NH_2	NH_2	-7.8	2.273	2.117 (NH_2)	1.823 (NH_2)	-	-	-	-
b.	C_3H_7	NH_2	NH_2	-8.4	2.263	-	1.911 (NH_3^+)	-	-	-	-
c.	CH_3	CH_2NH_2	NH_2	-8.3	2.245	-	2.231 (NH_3^+)	-	-	-	-
d.	CH_3	$C_2H_4NH_2$	NH_2	-8.8	2.332	-	1.830 (NH_3^+)	2.175 (NH_3^+)	-	-	-
e.	CH_3	NH_2	CH_2NH_2	-8.1	2.252	-	-	2.114 (NH_3^+)	-	-	-
f.	CH_3	NH_2	$C_2H_4NH_2$	-8.7	2.320	2.132 (NH_3^+)	-	-	-	2.086 (NH_3^+)	-
2.	**NH_2**	**CH_3**	**NH_2**	**-7.6**							
a.	CH_2NH_2	CH_3	NH_2	-7.3	2.237	-	1.955 (NH_2)	-	-	-	-
b.	$C_2H_4NH_2$	CH_3	NH_2	-7.8	2.243	-	2.128 (NH_2)	-	-	-	-
c.	NH_2	C_2H_5	NH_2	-7.0	2.237	-	1.933 (NH_2)	-	-	-	-
d.	NH_2	C_3H_7	NH_2	-7.7	2.237	-	1.829 (NH_2)	-	-	-	-
e.	NH_2	CH_3	CH_2NH_2	-7.4	2.237	-	-	-	2.369 (NH_2)	-	-
f.	NH_2	CH_3	$C_2H_4NH_2$	-8.4	2.237	-	-	-	1.800 (NH_2)	-	-
3.	**CH_3**	**OH**	**OH**	**-8.0**							
a.	C_2H_5	OH	OH	-8.0	2.238	1.889 (OH)	-	-	-	-	-
b.	C_3H_7	OH	OH	-8.1	2.238	2.100 (OH)	-	-	-	-	-
c.	CH_3	CH_2OH	OH	-8.4	2.259	-	2.018 (OH)	2.055 (OH)	-	-	-
d.	CH_3	C_2H_4OH	OH	-8.6	2.237	-	2.114 (OH)	-	2.123 (OH)	-	-
e.	CH_3	OH	CH_2OH	-8.6	**2.291**	1.825 (OH)	-	-	2.018 (OH)	-	-
f.	CH_3	OH	C_2H_4OH	-9.3	2.237	1.875 (OH)	-	-	2.038 (OH)	-	2.493 (OH)
4.	**OH**	**OH**	**NH_2**	**-8.3**							

a.	CH_2OH	OH	NH_2	-7.4	2.252	1.917 (OH)	-	-	-	-	-
b.	C_2H_4OH	OH	NH_2	-7.6	2.238	-	2.017 (NH_2)	-	-	-	-
c.	OH	CH_2OH	NH_2	-7.3	2.238	-	2.194 (NH_2)	-	-	-	-
d.	OH	C_2H_4OH	NH_2	-7.8	2.246	-	-	-	2.146 (OH)	-	-
e.	OH	OH	CH_2NH_2	-7.3	2.237	-	-	-	2.309 (NH_2)	-	-
f.	OH	OH	$C_2H_4NH_2$	-8.4	2.237	1.809 (OH)	-	-	1.954 (NH_2)	-	-
5.	**OH**	**NH_2**	**NH_2**	**-7.7**							
a.	CH_2OH	NH_2	NH_2	-7.7	2.259	-	-	1.694 (NH_3^+)	-	-	-
b.	C_2H_4OH	NH_2	NH_2	-8.1	2.251	2.005 (NH_2)	1.907 (NH_2)	-	-	-	-
c.	OH	CH_2NH_2	NH_2	-8.0	2.294	-	2.377 (NH_3^+)	2.078 (NH_3^+)	-	-	-
d.	OH	$C_2H_4NH_2$	NH_2	-9.1	2.253	2.438 (NH_3^+)	-	1.992 (NH_3^+)	-	-	-
e.	OH	NH_2	CH_2NH_2	-8.5	**2.151**	-	1.677 (NH_3^+)	-	-	-	-
f.	OH	NH_2	$C_2H_4NH_2$	-8.4	2.237	-	1.713 (NH_3^+)	-	-	-	-
6.	**OH**	**CH_3**	**NH_2**	**-8.2**							
a.	CH_2OH	CH_3	NH_2	-7.4	2.237	-	2.074 (NH_2)	-	-	-	-
b.	C_2H_4OH	CH_3	NH_2	-7.9	2.232	-	2.023 (NH_2)	-	-	-	-
c.	OH	C_2H_5	NH_2	-7.1	2.239	-	1.933 (NH_2)	-	-	-	-
d.	OH	C_3H_7	NH_2	-7.5	2.276	-	1.791 (NH_2)	-	-	-	-
e.	OH	CH_3	CH_2NH_2	-7.3	2.252	-	-	-	2.078 (NH_2)	-	-
f.	OH	CH_3	$C_2H_4NH_2$	-7.8	2.252	-	-	-	2.015 (NH_2)	-	-

Table7: Docking score of the interactions between different analogs and the residues with PDB ID 1DMW, along with their distances from the interacting groups

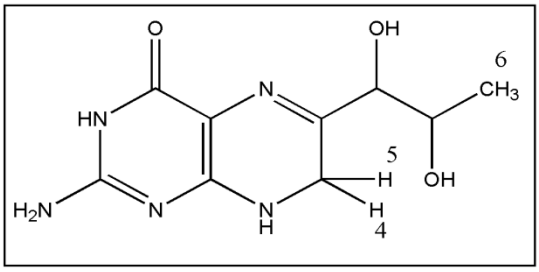


Fig12: Structure of BH_2 showing positions 4, 5 and 6

COMPOUND	POSITION 4	POSITION 5	POSITION 6	DOCKING SCORE	RESIDUES and DISTANCE (Å)						
					FE 425	SER 349	VAL 379	THR 278	GLU 280	GLY 346	PRO 279
7.	**OH**	**NH_2**	**NH_2**	**-7.9**							
a.	CH_2OH	NH_2	NH_2	-8.1	2.239	2.279 (NH_2)	-	-	-	2.137 (NH_2)	-
b.	C_2H_4OH	NH_2	NH_2	-8.4	2.245	-	2.332 (NH_2)	2.053 (OH)	2.163 (OH)	-	-
c.	OH	CH_2NH_2	NH_2	-7.9	2.237	-	-	2.021 (NH_2)	-	-	-
d.	OH	$C_2H_4NH_2$	NH_2	-8.2	2.243	-	2.197 (NH_2)	-	-	-	2.013 (NH_2)
e.	OH	NH_2	CH_2NH_2	-8.9	2.227	2.053 (NH_2)		-	2.098 (OH)	-	-
f.	OH	NH_2	$C_2H_4NH_2$	-9.1	2.264	1.875 (NH_2)	2.102 (NH_2)	-	-	-	-

Table8: Docking score of the interactions between different analogs and the residues with PDB ID 1DMW, along with their distances from the interacting groups

All the 42 compounds were found to show interactions with phenylalanine hydroxylase enzyme residues: Arg270, Glu280, Thr278, Pro279, Gly346, Ser349, Glu353, Val379 and Fe425. Fe (III) was found to be associated with 6 active site residues including water and also with the compound. The interacting distance of Fe (III) with the atom of different compounds was within a range of 2.151 to 2.291.

However table 7 suggests that compounds with substituted groups at positions 1, 2 and 3 interact with residues Glu353 and Arg270 whereas table8 suggests that compounds with substituted groups at positions 4, 5 and 6 interact with residues Gly346 and Pro279. Rest interacting residues are the same.

The docking poses of all the compounds represented that they bind in a very similar pattern with the active site of phenylalanine hydroxylase and the best results obtained with docking scores were -9.1 for 5 (d) and 7 (f). Compound 5 (d) show 2 hydrogen bond interactions, one with Glu280 with a distance of 2.438 (Å) and another with Val379 (Å) with a distance of 1.992 (fig 13).

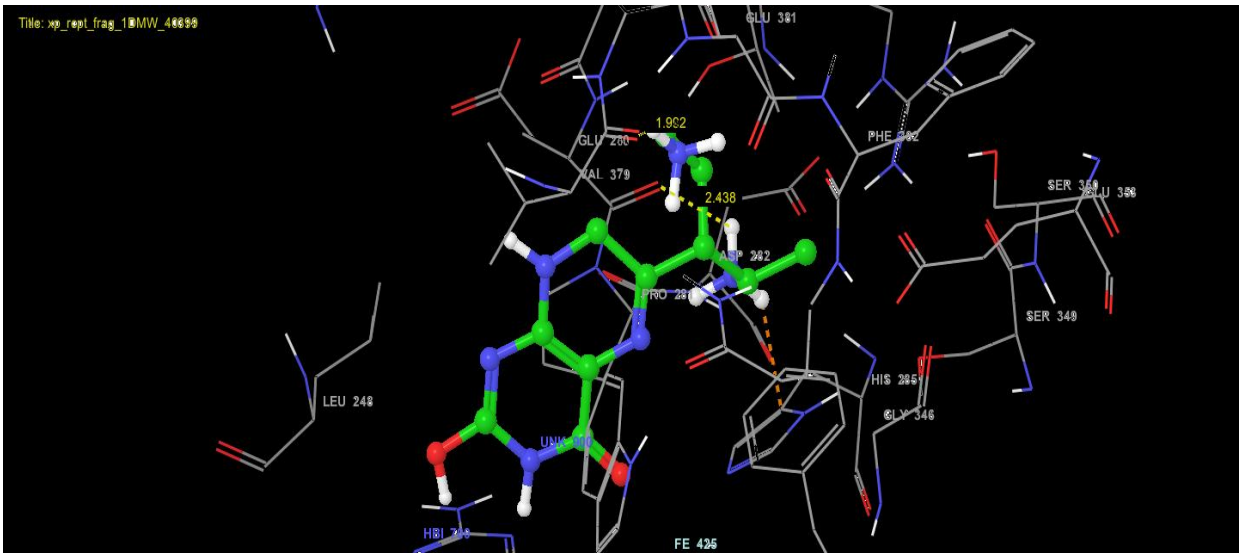
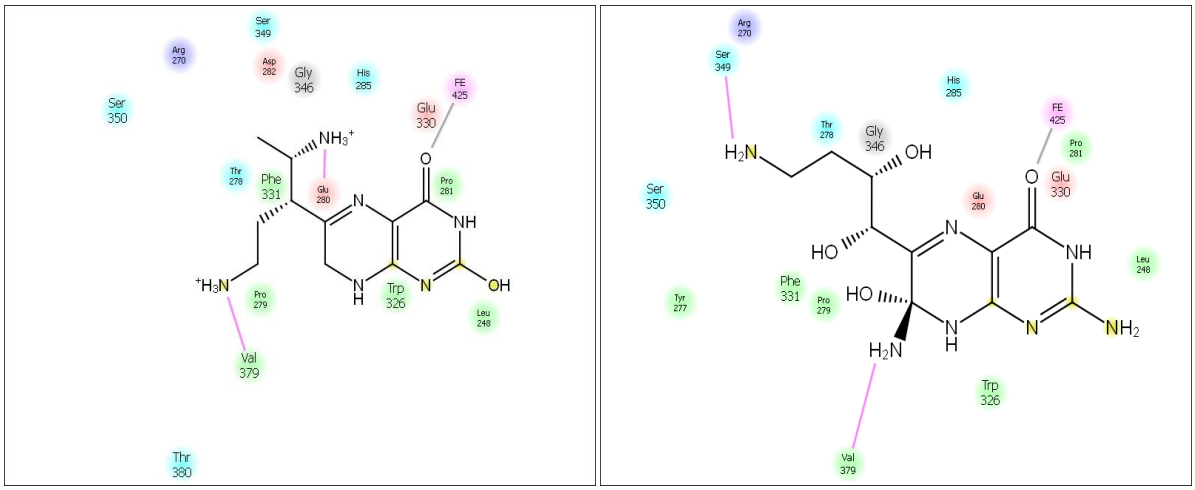


Fig13: Interactions between compound 5(d) and 1DMW

Compound 7 (f) also show 2 hydrogen bond interactions, one with Ser349 with a distance of 1.875 (Å) and another with Val379 with a distance of 2.102 (Å) (table8). Also, based on the glide scores it can be inferred that –OH, -NH_2, -NH_2 substitutions at positions 1, 2, 3 and 4, 5, 6 in compounds 5 (d) and 7 (f) respectively, with mainly long and branched chain increased the affinity. Further the presence of –NH_2 assisted in hydrogen bond interaction with Glu280, Ser349 and Val379 as shown in fig 14.



5 (d) 7 (f)

Fig14: Ligand interaction diagram of compounds 5(d) and 7(f) with pdb id 1DMW

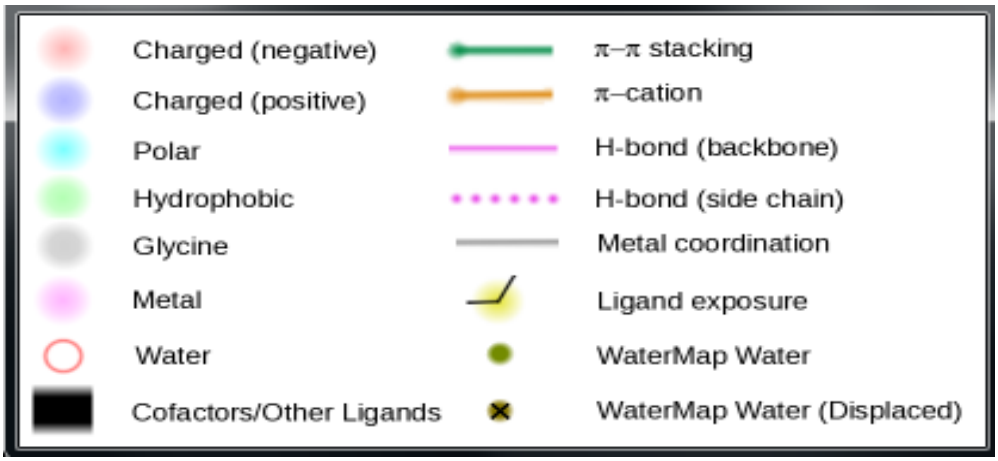
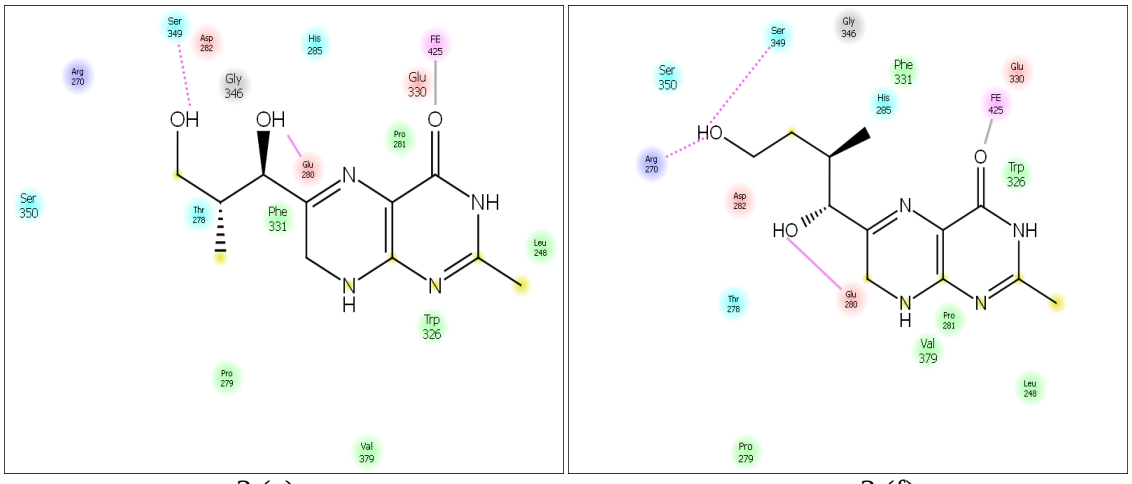


Fig15: Legend for ligand interaction diagram

S.NO	LIGANDS	DOCKING SCORE	Glu280 DISTANCE (Å)	Ser349 DISTANCE (Å)
1.	3 (e)	-8.6	1.825	2.018
2.	3 (f)	-9.3	1.875	2.038

Table9: Docking score and distances of the interacting residues of the compounds 3(e) and 3(f) with pdb id 1DMW



3 (e) 3 (f)

Fig16: Ligand interaction diagram of compounds 3(e) and 3(f) with pdb id 1DMW

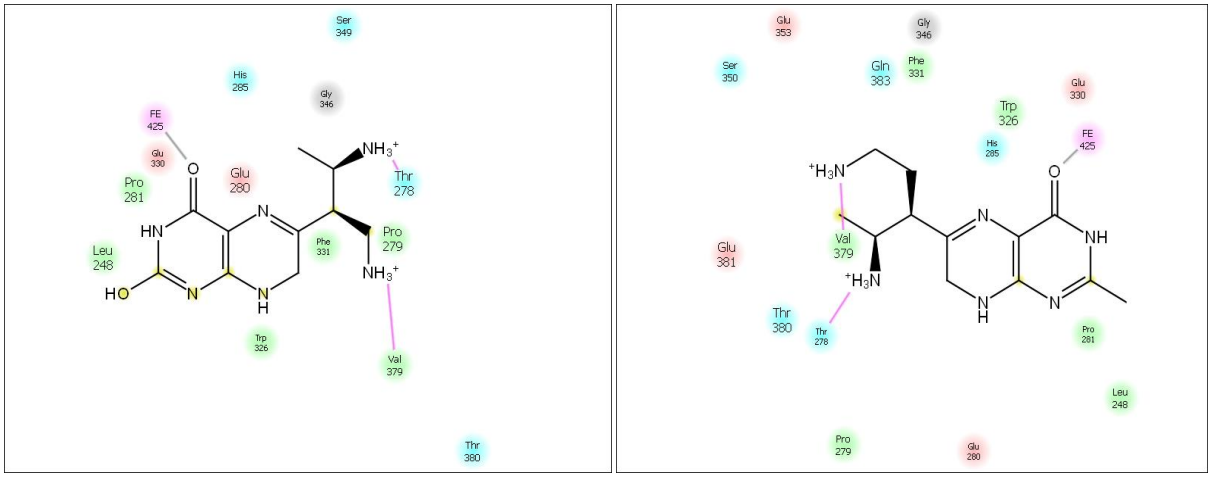
The table suggests that, the distance between H^+ atom of the compound 3 (e) and O^- atom of Ser349 is 2.018 and 2.038 for compound 3 (f) for a docking score of -8.6 and -9.3 respectively. It means that the docking score improved with an increase in distance between Ser349 and O^- atom of the compound.

Similarly, the distance between H^+ atom of the compound 3 (e) and O^- atom of Glu280 is 1.825 and 1.875 for compound 3 (f). So, the docking score again improved with an increase in distance between Glu280 and O^- atom of the compound.

This suggests that distance between O^- atom of Ser349 as well as Glu280 and H^+ atom of the compounds should be maximized in order to improve the binding affinity between the compounds and residues Ser349 and Glu280 of the protein as we know that the binding affinity improves with the improved docking score.

S.NO	LIGANDS	DOCKING SCORE	Val379 DISTANCE (Å)	Thr 278 DISTANCE (Å)
1.	5 (c)	-8.0	2.078	2.377
2.	1 (d)	-8.8	2.175	1.830

Table10: Docking score and distances of the interacting residues of the compounds 5(c) and 1(d) with pdb id 1DMW



5(c) 1(d)

Fig17: Ligand interaction diagram of compounds 5(c) and 1(d) with pdb id 1DMW

The table suggests that, the distance between H^+ atom of the compound 5 (c) and O^- atom of Val379 is 2.078 and 2.175 for compound 1 (d) for a docking score of -8.0 and -8.8 respectively. It means that the docking score improved with an increase in distance between Va379 and O^- atom of the compound.

Similarly, the distance between H^+ atom of the compound 5 (c) and O^- atom of Thr278 is 2.377 and 1.830 for compound 1 (d). So, the docking score improved with a decrease in distance between Thr278 and O^- atom of the compound.

This suggests that distance between O^- atom of Val379 and H^+ atom of the compounds should be maximized and distance between O^- atom of Thr278 and H^+ atom of the compounds should be minimized in order to improve the binding affinity between the compounds and residues Val379 and Thr278 of the protein as we know that the binding affinity improves with the improved docking score.

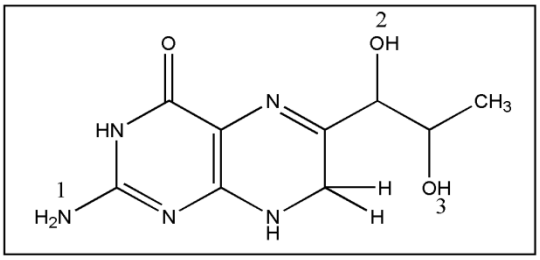


Fig11: Structure of BH_2 showing positions 1, 2 and 3

COMPOUND	POSITION 1	POSITION 2	POSITION 3	DOCKING SCORE	RESIDUES and DISTANCE (Å)							
					GLU 317	TYR 125	LEU 236	GLU 340	SER 337	THR 265	THR 368	PRO 266
8.	**CH_3**	**NH_2**	**NH_2**	**-8.3**								
a.	C_2H_5	NH_2	NH_2	-8.3	2.300 (NH)	2.385 (NH)	2.303 (NH_2)	-	-	-	-	-
b.	C_3H_7	NH_2	NH_2	-8.4	2.310 (NH)	2.392 (NH)	2.152 (NH_3^+)	-	-	-	-	-
c.	CH_3	CH_2NH_2	NH_2	-7.8	2.066 (NH)	-	-	-	-	-	-	-
d.	CH_3	$C_2H_4NH_2$	NH_2	-8.1	2.098 (NH)	-	2.200 (NH_3^+)	-	-	-	-	-
e.	CH_3	NH_2	CH_2NH_2	-7.8	1.965 (NH)	-	2.183 (NH_3^+)	-	-	-	-	-
f.	CH_3	NH_2	$C_2H_4NH_2$	-8.4	2.045 (NH)	-	1.713 (NH_3^+)	-	-	-	-	-
9.	**OH**	**OH**	**CH_3**	**-7.6**								
a.	CH_2OH	OH	CH_3	-8.4	2.153 (NH) 2.307 (OH)	2.483 (NH)	-	-	-	-	-	-
b.	C_2H_4OH	OH	CH_3	-6.0	-	2.200 (NH)	-	-	-	1.823 (OH)	-	-
c.	OH	CH_2OH	CH_3	-7.5	2.252 (NH)	2.441 (NH)	-	-	-	-	-	-
d.	OH	C_2H_4OH	CH_3	-6.3	-	2.110 (NH)	1.723 (OH)	-	-	-	-	-
e.	OH	OH	C_2H_5	-7.7	2.092 (NH)	-	-	-	-	-	-	-
f.	OH	OH	C_3H_7	-7.6	-	2.292 (OH)	-	-	-	-	-	-
10.	**OH**	**CH_3**	**OH**	**-7.6**								
a.	CH_2OH	CH_3	OH	-7.9	1.875 (OH)	-	-	1.818 (OH)	1.809 (OH)	1.803 (NH)	-	-
b.	C_2H_4OH	CH_3	OH	-6.6	-	2.178 (NH)	1.972 (OH)	-	-	2.031 (OH)	-	-
c.	OH	C_2H_5	OH	-7.1	-	-	-	-	-	-	-	-
d.	OH	C_3H_7	OH	-8.1	2.175 (NH)	2.486 (NH)	-	-	-	-	-	-
e.	OH	CH_3	CH_2O	-8.6	-	-	-	1.873	2.016	1.833	-	-

			H					(OH) 1.996 (NH)	(OH)	(NH)		
f.	OH	CH_3	CH_2OH	-8.1	1.989 (NH)	-	-	-	-	-	-	-
11.	**CH_3**	**OH**	**OH**	**-7.9**								
a.	C_2H_5	OH	OH	-8.2	2.198 (NH)	2.473 (NH)	2.259 (OH)	-	-	-	-	-
b.	C_3H_7	OH	OH	-8.1	2.426 (NH)	2.379 (NH)	2.105 (OH)	-	-	-	-	-
c.	CH_3	CH_2OH	OH	-8.5	2.119 (NH)	-	2.166 (OH)	-	-	-	-	-
d.	CH_3	C_2H_4OH	OH	-8.5	2.181 (NH)	-	1.711 (OH)	-	-	-	-	-
e.	CH_3	OH	CH_2OH	-8.8	2.145 (NH)	2.457 (NH)	2.080 (OH)	-	-	-	-	-
f.	CH_3	OH	CH_2OH	-8.9	2.108 (NH)	2.451 (NH)	-		-	-	-	-
12.	**NH_2**	**CH_3**	**CH_3**	**-7.6**								
a.	CH_2NH_2	CH_3	CH_3	-7.7	1.982 (NH)	-	-	-	-	-	-	-
b.	$C_2H_4NH_2$	CH_3	CH_3	-7.5	2.143 (NH)	2.104 (NH)	-	-	-	-	-	-
c.	NH_2	C_2H_5	CH_3	-7.8	2.091 (NH)	-	-	-	-	-	-	-
d.	NH_2	C_3H_7	CH_3	-7.9	2.045 (NH)	-	-	-	-	-	-	-
e.	NH_2	CH_3	C_2H_5	-7.9	2.235 (NH)	2.447 (NH)	-	-	-	-	-	-
f.	NH_2	CH_3	C_3H_7	-8.0	1.991 (NH)	-	-	-	-	-	-	-
13.	**CH_3**	**NH_2**	**CH_3**	**-7.7**								
a.	C_2H_5	NH_2	CH_3	-7.9	2.234 (NH)	2.454 (NH)	-	-	-	-	-	-
b.	C_3H_7	NH_2	CH_3	-8.3	2.273 (NH)	2.423 (NH)	-	-	-	-	-	-
c.	CH_3	CH_2NH_2	CH_3	-7.8	2.013 (NH)	-	-	-	-	-	-	-
d.	CH_3	$C_2H_4NH_2$	CH_3	-8.0	2.048 (NH)	-	-	-	-	-	-	-
e.	CH_3	NH_2	C_2H_5	-7.7	2.021 (NH)	-	-	-	-	-	-	-
f.	CH_3	NH_2	C_3H_7	-8.4	2.131 (NH)	2.466 (NH)	-	-	-	-	-	-
14.	**CH_3**	**CH_3**	**NH_2**	**-7.7**								
a.	C_2H_5	CH_3	NH_2	-7.9	2.299 (NH)	2.418 (NH)	-	-	-	-	-	-
b.	C_3H_7	CH_3	NH_2	-7.9	2.299 (NH)	2.418 (NH)	-	-	-	-	-	-
c.	CH_3	C_2H_5	NH_2	-7.8	2.195 (NH)	2.486 (NH)	-	-	-	-	-	-
d.	CH_3	C_3H_7	NH_2	-8.2	2.169 (NH)	2.486 (NH)	-	-	-	-	-	-
e.	CH_3	CH_3	CH_2NH_2	-8.0	1.995 (NH)	-	-	-	-	-	-	-

f.	CH_3	CH_3	$C_2H_4NH_2$	-8.4	2.032 (NH)	-	-	-	-	-	-	-
15.	**CH_3**	**CH_3**	**OH**	**-7.9**								
a.	C_2H_5	CH_3	OH	-8.8	2.094 (NH)	2.468 (NH)	2.213 (OH)	-	-	-	-	-
b.	C_3H_7	CH_3	OH	-8.6	2.424 (NH)	2.405 (NH)	2.036 (OH)	-	-	-	-	-
c.	CH_3	C_2H_5	OH	-8.2	2.206 (NH)	-	2.187 (OH)	-	-	-	-	-
d.	CH_3	C_3H_7	OH	-8.8	2.100 (NH)	2.478 (NH)	-	-	-	-	-	-
e.	CH_3	CH_3	CH_2OH	-8.5	1.995 (NH)	-	2.212 (OH)	-	-	-	-	-
f.	CH_3	CH_3	CH_2OH	-9.1	2.133 (NH)	2.484 (NH)	2.170 (OH)	-	-	-	-	-
16.	**CH_3**	**OH**	**CH_3**	**-8.0**								
a.	C_2H_5	OH	CH_3	-8.2	2.236 (NH)	2.425 (NH)	-	-	-	-	-	-
b.	C_3H_7	OH	CH_3	-8.7	2.229 (NH)	2.383 (NH)	-	-	-	-	-	-
c.	CH_3	CH_2OH	CH_3	-8.1	2.205 (NH)	2.460 (NH)	-	-	-	-	-	-
d.	CH_3	C_2H_4OH	CH_3	-8.1	2.158 (NH)	-	-	-	-	-	-	-
e.	CH_3	OH	C_2H_5	-7.9	2.041 (NH)	-	-	-	-	-	-	-
f.	CH_3	OH	C_3H_7	-8.1	1.972 (NH)	-	-	-	-	-	-	-
17.	**OH**	**CH_3**	**CH_3**	**-7.9**								
a.	CH_2OH	CH_3	CH_3	-7.4	-	-	-	1.837 (OH)	1.808 (OH)	1.837 (NH)	-	-
b.	C_2H_4OH	CH_3	CH_3	-8.6	2.054 (NH)	-	-	-	-	-	-	-
c.	OH	C_2H_5	CH_3	-7.7	2.155 (NH)	-	-	-	-	-	-	-
d.	OH	C_3H_7	CH_3	-7.9	2.006 (NH)	-	-	-	-	-	-	-
e.	OH	CH_3	C_2H_5	-8.0	2.035 (NH)	-	-	-	-	-	-	-
f.	OH	CH_3	C_3H_7	-8.1	1.989 (NH)	-	-	-	-	-	-	-
18.	**OH**	**NH_2**	**NH_2**	**-10.2**								
a.	CH_2OH	NH_2	NH_2	-7.8	2.216 (NH) 2.250 (OH)	-	2.263 (NH_2)	-	-	-	-	-
b.	C_2H_4OH	NH_2	NH_2	-8.3	2.260 (NH)	2.409 (NH)	2.205 (NH_2)	-	-	-	-	-
c.	OH	CH_2NH_2	NH_2	-9.3	-	1.947 (OH)	-	1.887 (NH_3^+)	-	-	-	2.073 (NH)
d.	OH	$C_2H_4NH_2$	NH_2	-8.4	-	2.079 (OH)	-	1.779 (NH_3^+)	-	2.203 (NH_3)	1.785 (NH_3)	2.093 (NH)
e.	OH	NH_2	CH_2NH_2	-7.6	-	2.169 (OH)	-	1.615 (NH_3^+)	-	-	-	1.872 (NH)

f.	OH	NH_2	$C_2H_4NH_2$	-7.5	2.007 (NH)	-	-	-	-	-	-	-
19.	**CH_3**	**NH_2**	**OH**	**-7.7**								
a.	C_2H_5	NH_2	OH	-7.8	2.042 (NH)	-	-	-	-	-	-	-
b.	C_3H_7	NH_2	OH	-8.1	2.080 (NH)	-	-	-	-	-	-	-
c.	CH_3	CH_2NH_2	OH	-7.6	-	-	1.970 (OH) 1.58 (NH_3^+)	1.946 (OH)	-	-	1.841 (NH_3^+)	2.054 (NH)
d.	CH_3	$C_2H_4NH_2$	OH	-7.3	2.332 (NH)	2.445 (NH)	2.243 (OH)	-	-	-	-	-
e.	CH_3	NH_2	CH_2OH	-8.4	2.150 (NH)	-	2.077 (OH)	-	-	-	-	-
f.	CH_3	NH_2	CH_2OH	-8.0	2.096 (NH)	-	-	-	-	-	-	-
20.	**NH_2**	**CH_3**	**OH**	**-7.6**								
a.	CH_2NH_2	CH_3	OH	-8.4	2.348 (NH)	2.381 (NH)	1.952 (OH)	-	-	-	-	-
b.	$C_2H_4NH_2$	CH_3	OH	-6.6	-	2.013 (NH)	2.089 (OH)	-	-		-	-
c.	NH_2	C_2H_5	OH	-8.4	2.285 (NH)	2.444 (NH)	-	-	-	-	-	-
d.	NH_2	C_3H_7	OH	-7.9	2.155 (NH)	2.465 (NH)	-	-	-	-	-	-
e.	NH_2	CH_3	CH_2OH	-8.4	2.237 (NH)	-	1.774 (OH)	-	-	-	-	-
f.	NH_2	CH_3	CH_2OH	-8.7	1.996 (NH)	-	1.833 (OH)	-	-	-	-	-

Table11: Docking score of the interactions between different analogs and the residues with PDB ID 1MLW, along with their distances from the interacting groups

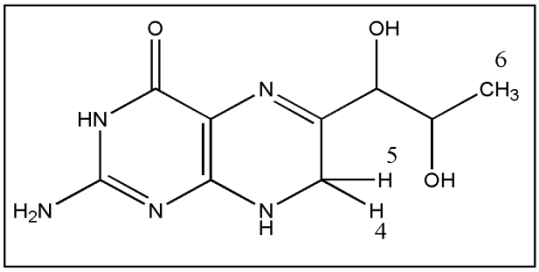


Fig12: Structure of BH_2 showing positions 4, 5 and 6

COMPOUND	POSITION 4	POSITION 5	POSITION 6	DOCKING SCORE	RESIDUES and DISTANCE (Å)									
					GLU 317	TYR 125	GLU 340	THR 265	THR 368	PRO 266	THR 367	GLY 333	SER 337	SER 336
21.	**OH**	**OH**	**NH_2**	**-9.2**										
a.	CH_2OH	OH	NH_2	-9.4	2.004 (NH_2)	-	-	1.863 (NH) 1.759 (OH)	2.185 (NH)	-	-	-	-	-
b.	C_2H_4OH	OH	NH_2	-7.7	1.890 (NH_2)	-	-	2.006 (NH) 1.899 (OH)	2.163 (NH)	-	-	-	-	-
c.	OH	CH_2OH	NH_2	-9.2	2.477 (NH_2)	2.323 (NH)	-	-	-	-	-	-	-	-
d.	OH	C_2H_4OH	NH_2	-8.3	2.181 (NH_2)	-	-	1.952 (NH) 2.133 (OH)	2.079 (NH)	-	-	1.961 (OH)	-	-
e.	OH	OH	CH_2NH_2	-9.1	-	2.445 (NH_2)	-	1.832 (OH)	2.108 (NH)	2.068 (NH)	1.985 2.167 (OH)	-	-	-
f.	OH	OH	$C_2H_4NH_2$	-9.1	-	2.092 (NH_2)	1.918 1.514 (OH)	1.820 (OH)	-	2.268 (NH)	-	-	2.029 (OH)	-
22.	**OH**	**NH_2**	**NH_2**	**-7.8**										
a.	CH_2OH	NH_2	NH_2	-6.8	-	-	-	1.772 (NH)	2.294 (NH)	-	-	-	-	-
b.	C_2H_4OH	NH_2	NH_2	-6.5	-	-	-	2.259 (OH) 1.762 (NH_2)	-	-	2.041 (OH)	1.833 (NH_2)	-	-
c.	OH	CH_2NH_2	NH_2	-7.8	-	2.058 (NH_2)	-	2.195 (NH)	2.350 (NH)	-	-	-	-	-
d.	OH	$C_2H_4NH_2$	NH_2	-6.8	2.158 (NH_2)	-	-	2.081 (NH)	2.220 (NH)	1.980 (OH)	-	-	-	-
e.	OH	NH_2	CH_2NH_2	-7.9	1.928 (NH_2)	-	1.821 (OH)	-	-	-	-	-	1.797 (OH)	-
f.	OH	NH_2	$C_2H_4NH_2$	-10.1	-	1.951 (NH_2)	1.906 (OH)	-	-	2.028 (NH)	-	-	1.895 (OH)	-
23.	**NH_2**	**OH**	**NH_2**	**-7.8**										
a.	CH_2NH_2	OH	NH_2	-8.0	-	2.054 (NH_2)	-	2.072 (NH)	2.298 (NH)	-	-	-	-	-
b.	$C_2H_4NH_2$	OH	NH_2	-6.9	1.866 (NH_2)	-	-	-	-	-	-	-	2.204 (NH_2)	1.947 (NH_2)
c.	NH_2	CH_2OH	NH_2	-9.5	2.011 (NH_2) 1.866 (OH)	-	-	2.253 1.912 (NH) 1.778 (OH)	2.112 (NH)	-	-	-	-	-
d.	NH_2	C_2H_4OH	NH_2	-7.1	2.315 (NH_2)	-	-	2.450 (NH) 1.908 (OH)	2.232 (NH)	1.959 (NH)	-	-	-	-
e.	NH_2	OH	CH_2NH_2	-7.9	-	-	1.975 (NH_2) 2.343	-	2.131 (OH)	-	2.177 (OH)	--	1.897 (OH)	-

							1.887 (OH)							
f.	NH_2	OH	$C_2H_4NH_2$	-7.9	-	2.080 (NH)	1.752 (NH_2)	2.134 (OH)	-	1.920 (NH)	-	-	1.925 (NH_2)	-
24.	**CH_3**	**OH**	**NH_2**	**-7.9**										
a.	C_2H_5	OH	NH_2	-7.0	-	-	-	1.923 (NH) 2.126 (OH)	2.039 (NH)	-	-	-	-	-
b.	C_3H_7	OH	NH_2	-8.7	-	1.831 (NH_2)	1.918 (NH_3^+) 1.812 (OH)	1.807 (NH)	1.726 (NH_3^+)	-	-	-	-	-
c.	CH_3	CH_2OH	NH_2	-7.5	2.004 (NH_3^+)	-	-	1.837 (NH) 1.803 (OH)	2.170 (NH)	-	-	-	-	-
d.	CH_3	C_2H_4OH	NH_2	-7.7	1.814 (OH)	-	-	1.815 (NH)	2.403 (NH)	1.869 (OH)	-	-	-	-
e.	CH_3	OH	CH_2NH_2	-7.4	1.791 (NH)	-	1.665 (NH_3^+) 1.616 (OH)	-	-	-	-	-	-	-
f.	CH_3	OH	$C_2H_4NH_2$	-7.3	-	-	1.860 (NH_3^+)	-	1.746 (NH_3^+)	2.179 (OH)	-	-	-	-
25.	**OH**	**CH_3**	**NH_2**	**-7.8**										
a.	CH_2OH	CH_3	NH_2	-7.8	2.270 (NH_2)	-	-	1.851 (NH) 1.949 (OH)	2.406 (NH)	-	-	-	-	-
b.	C_2H_4OH	CH_3	NH_2	-6.9	1.757 (OH)	1.793 (OH)	-	2.027 (NH)	2.309 (NH)	-	-	-	-	-
c.	OH	C_2H_5	NH_2	-7.9	1.973 (NH_2)	-	-	1.838 (NH)	2.433 (NH)	-	-	-	-	-
d.	OH	C_3H_7	NH_2	-7.8	1.970 (OH)	-	-	1.793 (NH)	2.392 (NH)	-	-	-	-	-
e.	OH	CH_3	CH_2NH_2	-7.7	-	2.093 (NH)	-	1.853 (NH)	2.147 (NH)	-	-	-	-	-
f.	OH	CH_3	$C_2H_4NH_2$	-7.3	-	-	-	1.777 (NH)	2.324 (NH)	-	-	-	-	-

Table12: Docking score of the interactions between different analogs and the residues with PDB ID 1MLW, along with their distances from the interacting groups

All the 108 compounds were found to show interactions with tryptophan hydroxylase enzyme residues: Tyr125, Leu236, Thr265, Pro266, Glu317, Gly333, Ser336, Ser337, Glu340, Thr367 and Thr368. Interactions with these residues were within a range:
However table12 suggests that compounds with substituted groups at positions 1, 2 and 3 interacted with residues and Thr367, Gly333 and Ser336 whereas table13 suggests that compounds with substituted groups at positions 4, 5 and 6 interacted with residue Leu236. Rest interacting residues were the same.
The docking poses of all the compounds show that they bind in a very similar pattern with the active site of tryptophan hydroxylase and the best results obtained with docking scores were -9.3 for compound 18 (c), -9.4 for compound 21 (a), -9.5 for compound 23 (c) and -10.1 for compound 22 (f).
Compound 18 (c) show 4 hydrogen bond interactions: 2 with Glu340, 1 with Tyr125 and 1 with Pro266 (fig 18).

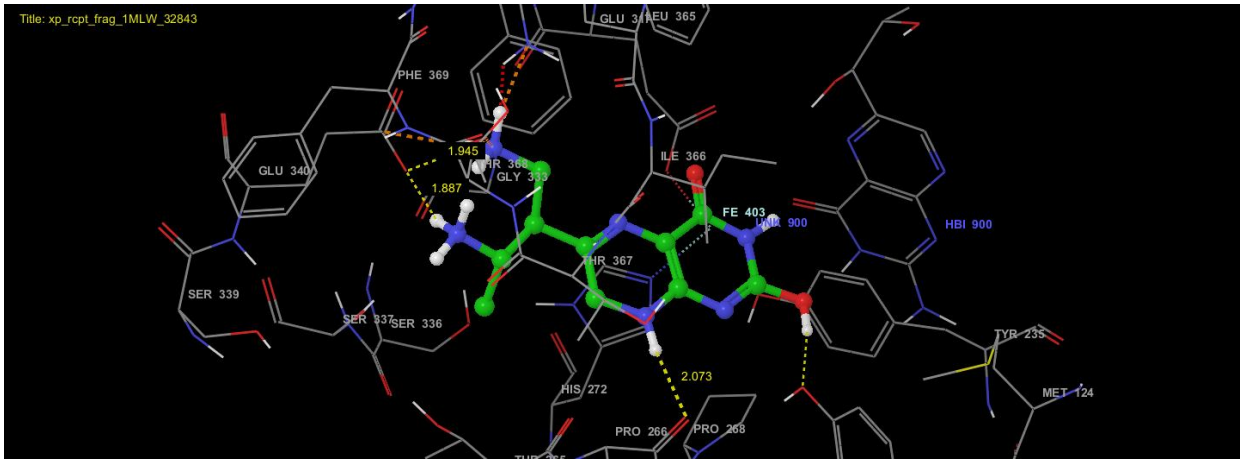


Fig18: Interactions between compound 18(c) and 1 MLW

Compound 21 (a) show 4 hydrogen bond interactions: 1 with Thr368, 3 with Thr265 and 1 with Glu317. Compound 23 (c) show 6 hydrogen bond interactions: 3 with Thr265, 2 with Glu317 and 1 with Thr368.
Compound 22 (f) show 6 hydrogen bond interactions: 2 with Glu340, remaining each with Ser337, Thr265, Pro266 and Tyr125 (fig 19).

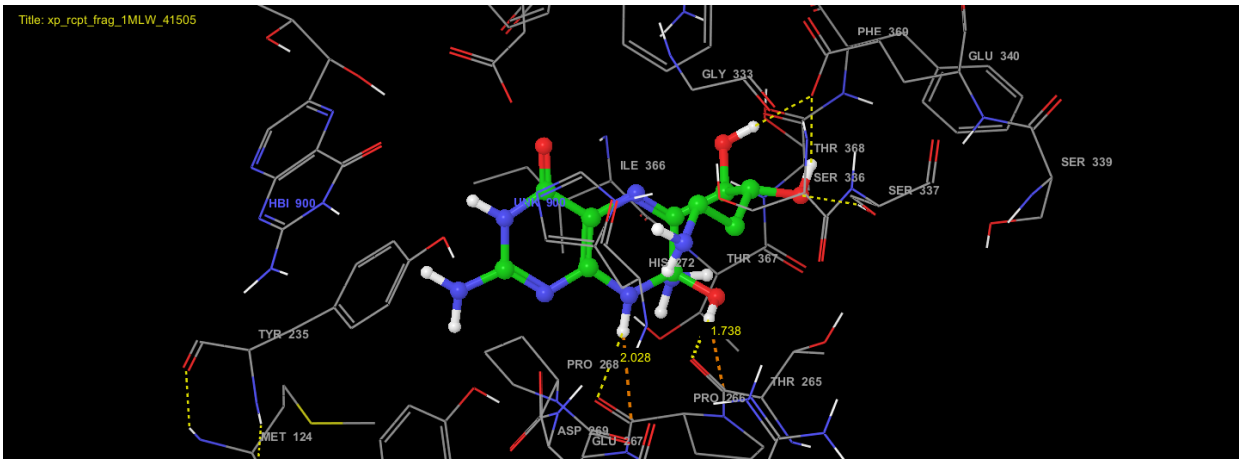
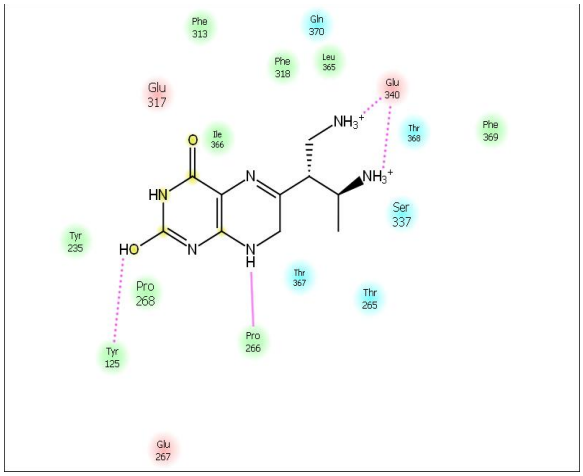
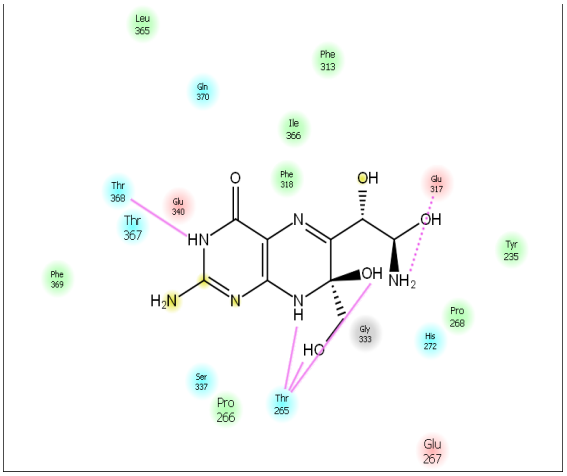


Fig19: Interactions between compound 22(f) and 1MLW

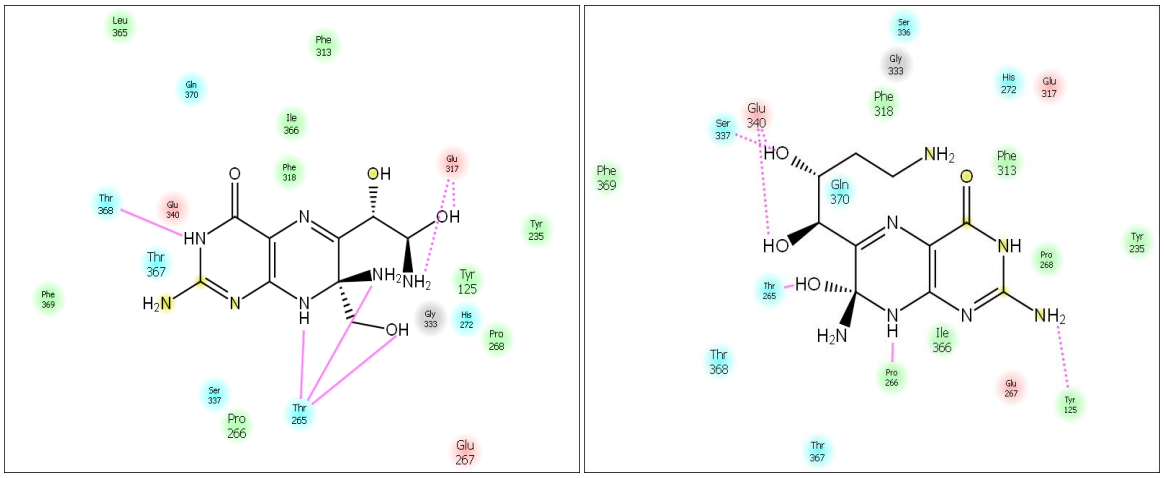
Also, based on the glide scores it can be inferred that –OH, -NH_2, -NH_2 substitutions at positions 1, 2, 3 and 4, 5, 6 in compounds 18 (c) and 22 (f) respectively; -OH, -OH, -NH_2 at positions 4, 5, 6 in compound 21 (a) and –NH_2, -OH, -NH_2 at same positions in compound 23 (c) with mainly long and branched chain increased the affinity. Further the presence of –NH_2 and –OH assisted in hydrogen bond interaction with Thr265, Thr368, Glu340, Glu317, Ser337, Tyr125 and Pro266 as shown in fig 20.



18 (c)



21 (a)

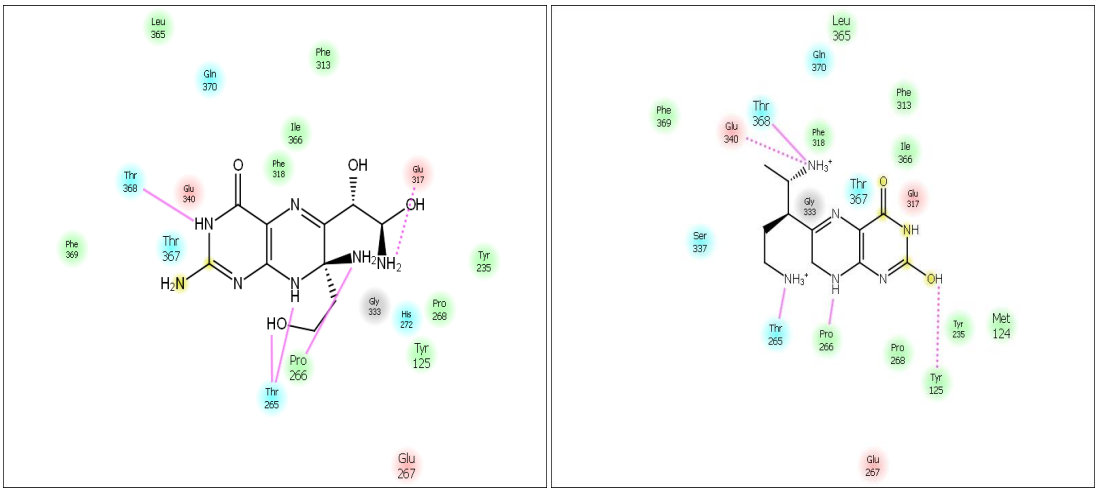


23 (c) 22 (f)

Fig20: Ligand interaction diagram of compounds 18(c), 21(a), 23(c) and 22(f) with pdb id 1MLW

S.NO	LIGANDS	DOCKING SCORE	Thr 265 DISTANCE (Å)	Pro 266 DISTANCE (Å)
1.	23 (d)	-7.1	2.450	1.959
2.	18 (d)	-8.4	2.203	2.093

Table13: Docking score and distances of the interacting residues of the compounds 23(d) and 18(d) with pdb id 1MLW



23(d) 18(d)

Fig21: Ligand interaction diagram of compounds 23(d) and 18(d) with pdb id 1MLW

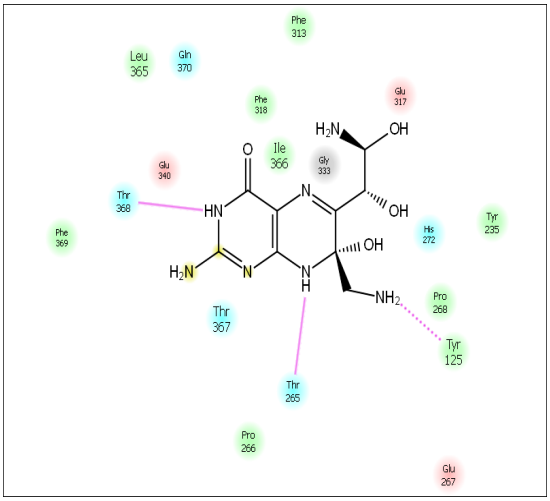
The table suggests that, the distance between H atom of the compound 23(d) and O atom of Thr265 is 2.450 and 2.203 for compound 18(d) for a docking score of -7.1 and -8.4 respectively. It means that the docking score improved with a decrease in distance between Thr265 and O atom of the compound.

However, the distance between H atom of the compound 23(d) and O atom of Pro266 is 1.959 and 2.093 for compound 18(d). So, the docking score improved with an increase in distance between Pro266 and O atom of the compound.

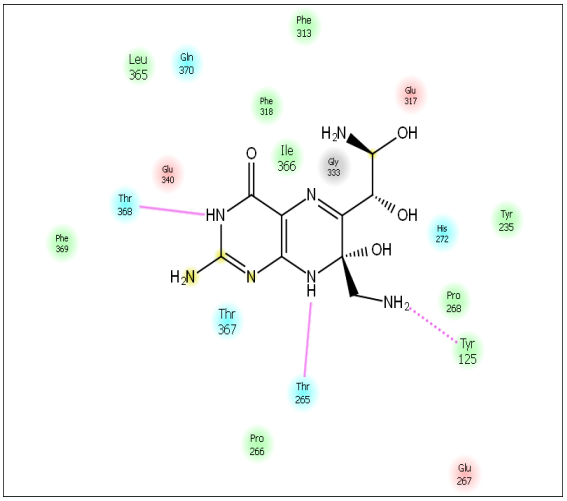
This suggests that distance between O atom of Thr265 should be minimized and distance between O atom Pro266 and H atom of the compounds should be maximized in order to improve the binding affinity between the compounds and residues Thr265 and Pro266 of the protein as we know that the binding affinity improves with the improved docking score.

S.NO	LIGANDS	DOCKING SCORE	Thr 265 DISTANCE (Å)	Tyr 125 DISTANCE (Å)
1.	22 (c)	-7.8	2.195	2.058
2.	23 (a)	-8.0	2.072	2.054
3.	24 (b)	-8.7	1.807	1.831

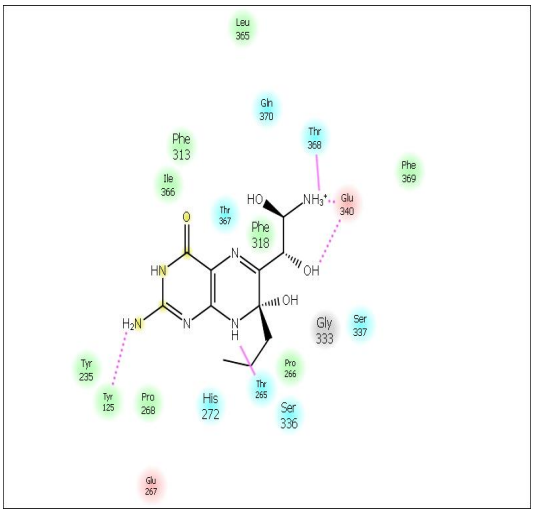
Table14: Docking score and distances of the interacting residues of the compounds 22(c), 23(a) and 24(b) with pdb id 1MLW



22 (c)



23(a)



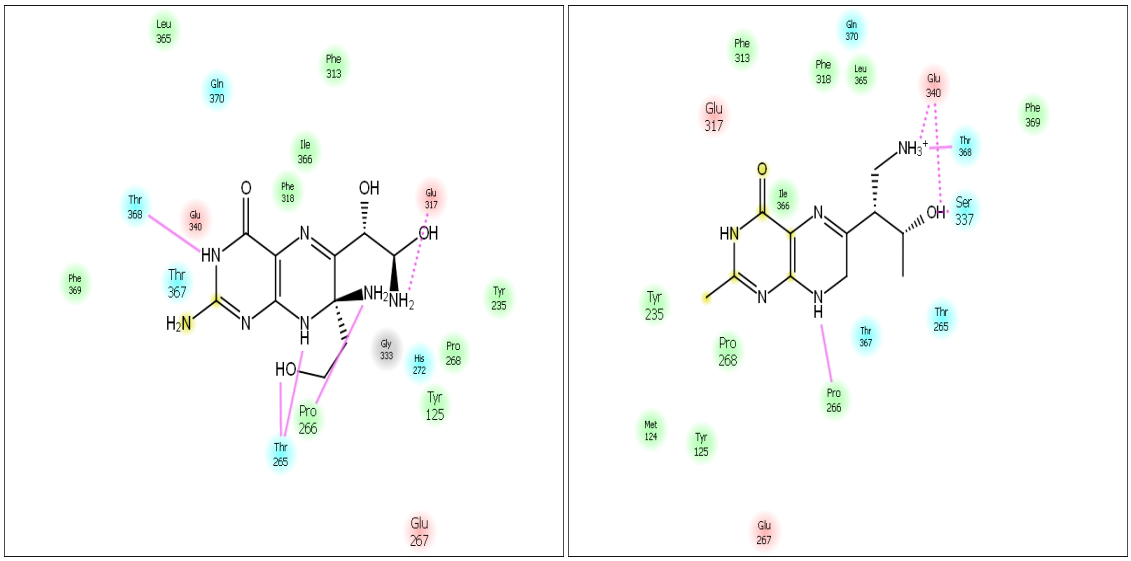
24(b)

Fig22: Ligand interaction diagram of compounds 22(c), 23(a) and 24(b) with pdb id 1MLW

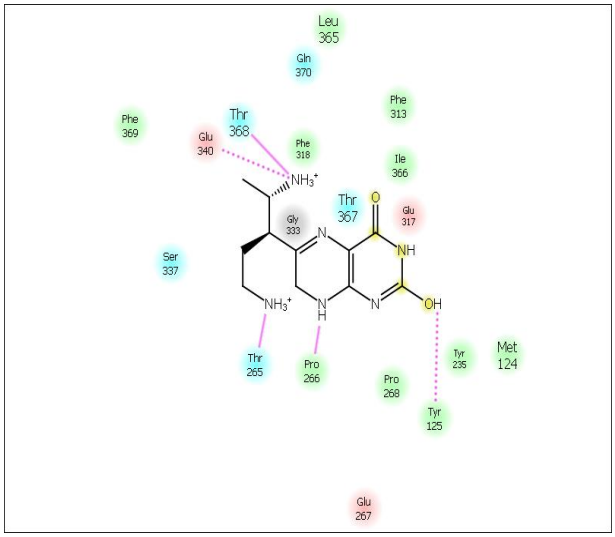
The table suggests that, the docking score improved with a decrease in distance between Thr265 and O atom of the compound. Similarly, the docking score improved with decrease in distance between Tyr125 and O atom of the compound.
Thus, the distance between O atom of Thr265 and Tyr125 should be minimized in order to improve the binding affinity between the compounds and residues Thr265 and Tyr125 of the protein as we know that the binding affinity improves with the improved docking score.

S.NO	LIGANDS	DOCKING SCORE	Thr 368 DISTANCE (Å)	Pro 266 DISTANCE (Å)
1.	23 (d)	-7.1	2.232	1.959
2.	19 (c)	-7.6	1.841	2.054
3.	18 (d)	-8.4	1.785	2.093

Table15: Docking score and distances of the interacting residues of the compounds 23 (d), 19 (c) and 18 (d) with pdb id 1MLW



23(d) 19(c)



18(d)

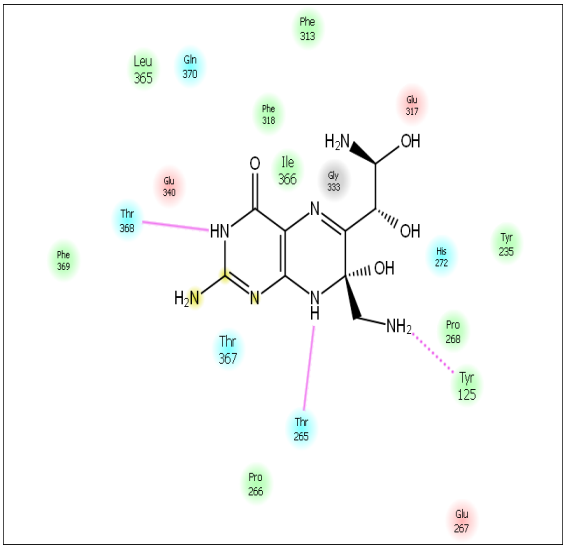
Fig23: Ligand interaction diagram of compounds 23(d), 19(c) and 18(d) with pdb id 1MLW

The table suggests that the docking score improved with a decrease in distance between Thr368 and O atom of the compound.

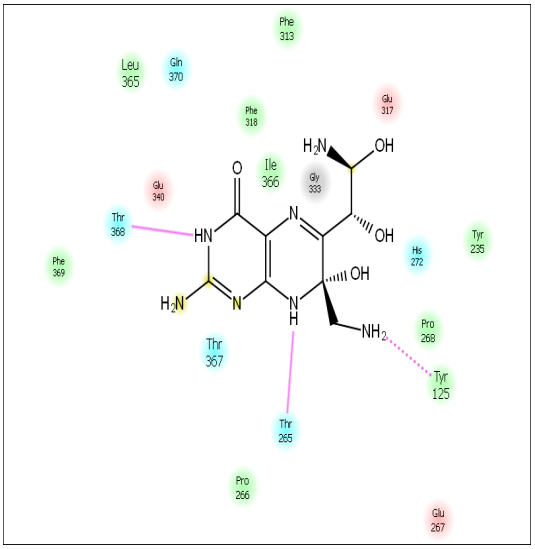
However, the docking score improved with an increase in distance between Pro266 and O atom of the compound. This suggests that distance between O atom of Thr368 should be minimized and distance between O atom Pro266 and H atom of the compounds should be maximized in order to improve the binding affinity between the compounds and residues Thr368 and Pro266 of the protein as we know that the binding affinity improves with the improved docking score.

S.NO	LIGANDS	DOCKING SCORE	Thr 368 DISTANCE (Å)	Tyr 125 DISTANCE (Å)
1.	22 (c)	-7.8	2.350	2.058
2.	23 (a)	-8.0	2.298	2.054
3.	24 (b)	-8.7	1.726	1.831

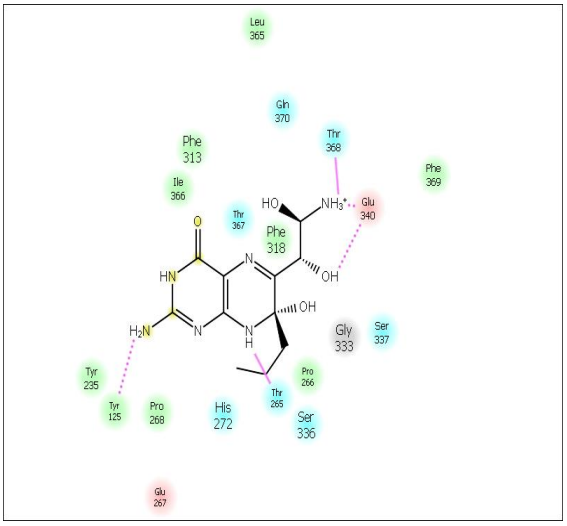
Table16: Docking score and distances of the interacting residues of the compounds 22 (c), 23 (a) and 24 (b) with pdb id 1MLW



22(c)



23(a)



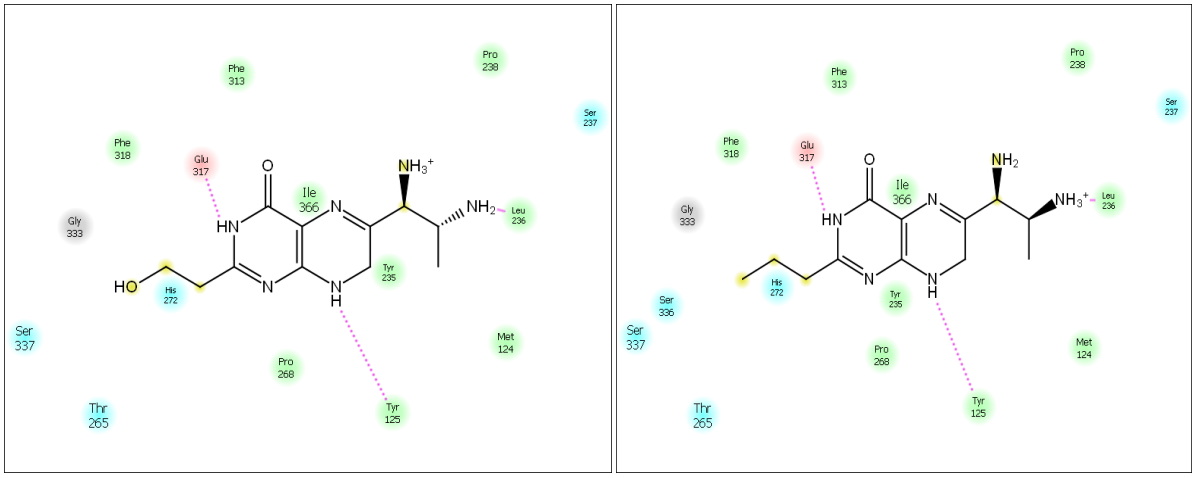
24(b)

Fig24: Ligand interaction diagram of compounds 22(c), 23(a) and 24(a) with pdb id 1MLW

The table suggests that, the docking score improved with a decrease in distance between Thr368 and O atom of the compound. Similarly, the docking score improved with decrease in distance between Tyr125 and O atom of the compound.
Thus, the distance between O atom of Thr368 and Tyr125 should be minimized in order to improve the binding affinity between the compounds and residues Thr368 and Tyr125 of the protein as we know that the binding affinity improves with the improved docking score.

S.NO	LIGANDS	DOCKING SCORE	Leu 236 DISTANCE (Å)	Tyr 125 DISTANCE (Å)
1.	18 (b)	-8.3	2.205	2.409
2.	8 (b)	-8.4	2.152	2.392

Table17: Docking score and distances of the interacting residues of the compounds 18(b) and 8(b) with pdb id 1MLW



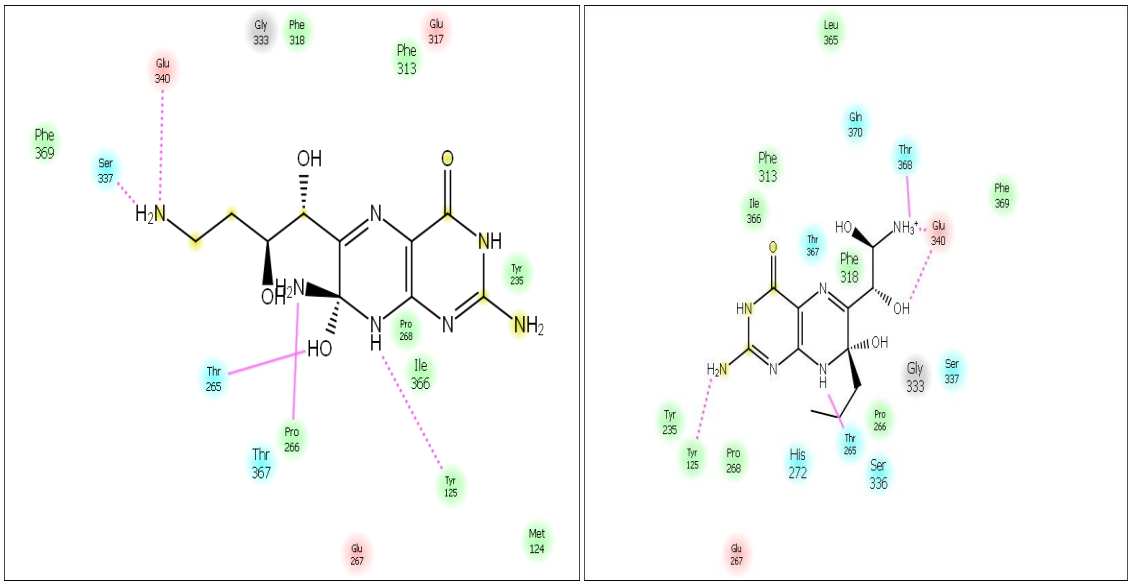
18(b) 8(b)

Fig25: Ligand interaction diagram of compounds 18(b) and 8(b) with pdb id 1MLW

The table suggests that, the docking score improved with a decrease in distance between Leu236 and O atom of the compound. Similarly, the docking score improved with decrease in distance between Tyr125 and O atom of the compound.
Thus, the distance between O atom of Leu236 and Tyr125 should be minimized in order to improve the binding affinity between the compounds and residues Leu236 and Tyr125 of the protein as we know that the binding affinity improves with the improved docking score.

S.NO	LIGANDS	DOCKING SCORE	Glu 340 DISTANCE (Å)	Tyr 125 DISTANCE (Å)
1.	23 (f)	-7.9	1.752	2.080
2.	24 (b)	-8.7	1.918	1.831

Table18: Docking score and distances of the interacting residues of the compounds 23(f) and 24(b) with pdb id 1MLW



23(f) 24(b)

Fig26: Ligand interaction diagram of compounds 23(f) and 24(b) with pdb id 1MLW

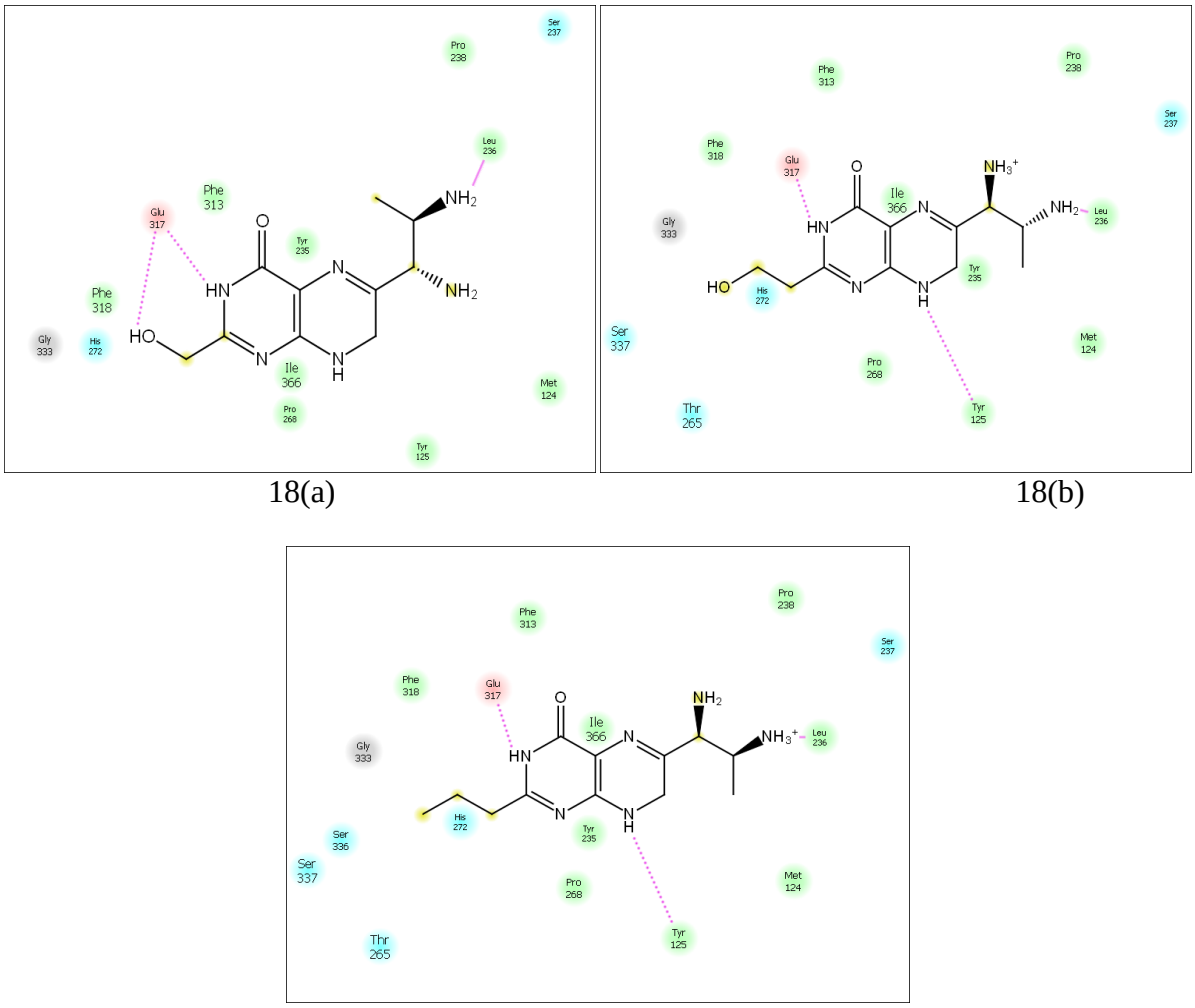
The table suggests that the docking score improved with a decrease in distance between Tyr125 and O atom of the compound.

However, the docking score improved with an increase in distance between Glu340 and O atom of the compound.

This suggests that distance between O atom of Tyr125should be minimized and distance between O atom Glu340 and H atom of the compounds should be maximized in order to improve the binding affinity between the compounds and residues Tyr125 and Glu340 of the protein as we know that the binding affinity improves with the improved docking score.

S.NO	LIGANDS	DOCKING SCORE	Leu 236 DISTANCE (Å)	Glu 317 DISTANCE (Å)
1.	18 (a)	-7.8	2.263	2.216
2.	18 (b)	-8.3	2.205	2.260
3.	8 (b)	-8.4	2.152	2.310

Table19: Docking score and distances of the interacting residues of the compounds 18(a), 18(b) and 8(b) with pdb id 1MLW



18(a) 18(b)

8(b)

Fig27: Ligand interaction diagram of compounds 18(a), 18(b) and 8(b) with pdb id 1MLW

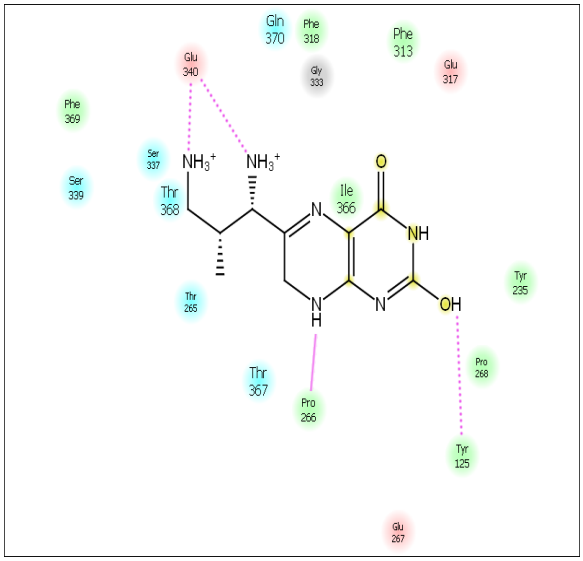
The table suggests that the docking score improved with a decrease in distance between Leu236 and O atom of the compound.

However, the docking score improved with an increase in distance between Glu317and O atom of the compound.

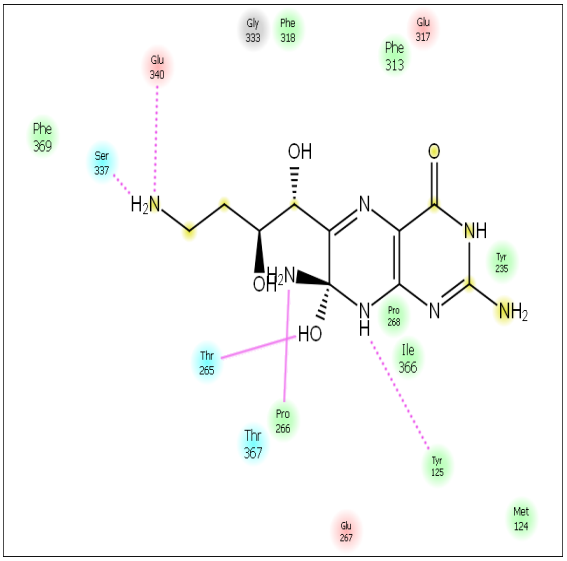
This suggests that distance between oxygen (O) atom of Leu236 should be minimized and distance between O atom Glu317 and H atom of the compounds should be maximized in order to improve the binding affinity between the compounds and residues Leu236 and Glu317 of the protein as we know that the binding affinity improves with the improved docking score.

S.NO	LIGANDS	DOCKING SCORE	Glu 340 DISTANCE (Å)	Pro 266 DISTANCE (Å)
1.	18 (e)	-7.6	1.615	1.872
2.	23 (f)	-7.9	1.752	1.920
3.	18 (d)	-8.4	1.779	2.093

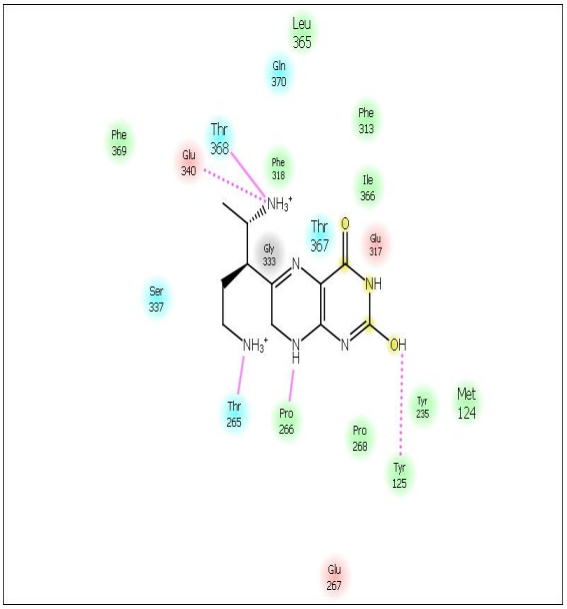
Table20: Docking score and distances of the interacting residues of the compounds 18(e), 23(f) and 18(d) with pdb id 1MLW



18(e)



23(f)



18(d)

Fig28: Ligand interaction diagram of compounds 18(e), 23(f) and 18(d) with pdb id 1MLW

The table suggests that, the docking score improved with an increase in distance between Glu340 and O atom of the compound. Similarly, the docking score improved with decrease in distance between Pro266 and O atom of the compound.
Thus, the distance between O atom of Glu340 and Pro266 should be minimized in order to improve the binding affinity between the compounds and residues Glu340 and Pro266 of the protein as we know that the binding affinity improves with the improved docking score.

2. Pharmacophore modeling analysis

From the docking analysis of these compounds with modifications at positions 1 to 6, we have screened out following 24 common compounds on the basis of their highest G scores and were inspected to identify a target pharmacophore:

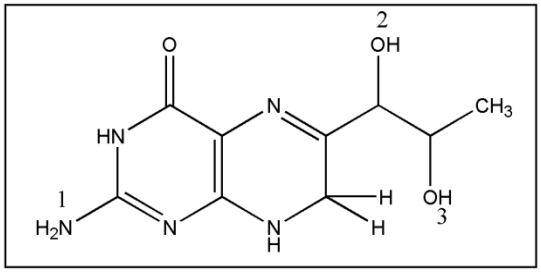


Fig11: Structure of BH_2 showing positions 1, 2 and 3

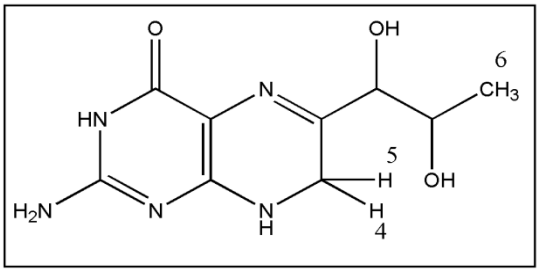


Fig12: Structure of BH_2 showing positions 4, 5 and 6

COMPOUND	POSITION 1	POSITION 2	POSITION 3	G SCORE	
				1DMW	1MLW
1.	C_2H_5	NH_2	NH_2	-7.7	-7.8
2.	C_3H_7	NH_2	NH_2	-8.1	-8.3
3.	CH_3	CH_2NH_2	NH_2	-8.0	-9.3
4.	CH_3	$C_2H_4NH_2$	NH_2	-9.1	-8.4
5.	CH_3	NH_2	CH_2NH_2	-8.5	-7.6
6.	CH_3	NH_2	$C_2H_4NH_2$	-8.4	-7.5
7.	C_2H_5	OH	OH	-7.8	-8.3
8.	C_3H_7	OH	OH	-8.4	-8.4
9.	CH_3	CH_2OH	OH	-8.3	-7.8
10.	CH_3	C_2H_4OH	OH	-8.8	-8.1
11	CH_3	OH	CH_2OH	-8.1	-7.8
12.	CH_3	OH	C_2H_4OH	-8.7	-8.4
13.	CH_2OH	NH_2	NH_2	-8.0	-8.2

14.	C_2H_4OH	NH_2	NH_2	-8.1	-8.1
15.	OH	CH_2NH_2	NH_2	-8.4	-8.5
16.	OH	$C_2H_4NH_2$	NH_2	-8.6	-8.5
17.	OH	NH_2	CH_2NH_2	-8.6	-8.8
18.	OH	NH_2	$C_2H_4NH_2$	-9.3	-8.9
COMPOUND	**POSITION 4**	**POSITION 5**	**POSITION 6**	**1DMW**	**1MLW**
19.	CH_2OH	NH_2	NH_2	-8.1	-6.8
20.	C_2H_4OH	NH_2	NH_2	-8.4	-6.5
21.	OH	CH_2NH_2	NH_2	-7.9	-7.8
22.	OH	$C_2H_4NH_2$	NH_2	-8.2	-6.8
23.	OH	NH_2	CH_2NH_2	-8.9	-7.9
24.	OH	NH_2	$C_2H_4NH_2$	-9.1	-10.1

Table21: List of compounds used to perform pharmacophores modeling along with their Gscore with PDB ID's 1DMW and 1MLW

To obtain the common pharmacophore features, the dataset were divided into active and inactive sets by using Gscore as activity of these molecules. Molecules with binding affinity values higher than 0.905 were considered to be active and those less than 0.905 were considered to be inactive. Hypotheses emerging from this process were subsequently scored and the hypotheses that survived the scoring process were used to build an atom-based QSAR model. A total of 387 four point hypotheses were obtained after completion of the scoring process. Focusing only on those pharmacophore models whose scores ranked in the top 1% of the actives, two hypotheses were selected.

Hypotheses 1:

Hypotheses 1 with survival-inactive score of 1.396 is shown in table 22. Pharmacophore features which were identified as common to all the 24 compounds were: A (H bond acceptor), D (H bond donator) as shown in Fig 30.

ID	**Survival**	**Survival - inactive**	**Post -hoc**	**Site**	**Vector**	**Volume**	**Selectivity**	**Energy**	**activity**
AADD.529	3.414	1.396	3.415	0.77	0.904	0.739	0.78	2.93	0.968

Table22: Details of the hypotheses (AADD)

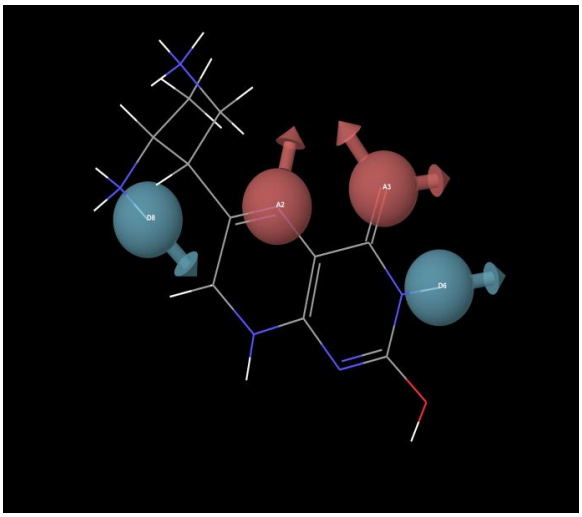


Fig29: PHASE hypotheses that yielded the common pharmacophore features (AADD)

Entry	Site1	Site2	Site3	Angle
AADD.529	A2	A3	D8	28.9
AADD.529	A2	A3	D6	114.9
AADD.529	D8	A3	D6	114.8
AADD.529	A3	A2	D8	125.8
AADD.529	A3	A2	D6	29.8
AADD.529	D8	A2	D6	123.7
AADD.529	A3	D8	A2	25.3
AADD.529	A3	D8	D6	19
AADD.529	A2	D8	D6	33.1
AADD.529	A3	D6	A2	35.3
AADD.529	A3	D6	D8	46.1
AADD.529	A2	D6	D8	23.2

Table23: Site measurements for angles for AADD

Entry	Site1	Site2	Distance
AADD.529	A3	A2	2.874
AADD.529	A3	D8	5.455
AADD.529	A3	D6	2.467
AADD.529	A2	D8	3.251
AADD.529	A2	D6	4.507
AADD.529	D8	D6	6.867

Table24: Site measurements for angles for AADD

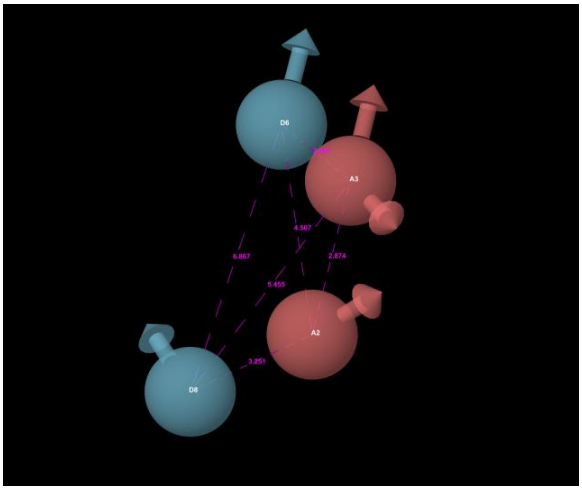


Fig30: Site measurements for QSAR model (AADD)

Then the QSAR model was built using 4 PLS factors by randomly selecting 20 compounds as training sets and rest 4 as test sets as shown in table 25.

Ligand Name	QSAR Set	Activity	# Factors	Predicted Activity	Pharm Set	Fitness
1.	test	0.892	1 2 3 4	0.92 0.93 0.92 0.91	inactive	1.15
2.	training	0.919	1 2 3 4	0.93 0.93 0.92 0.92	active	1.16
3.	training	0.968	1 2 3 4	0.95 0.94 0.94 0.95	active	3
4.	training	0.924	1 2 3 4	0.94 0.94 0.93 0.93	active	2.88
5.	training	0.881	1 2 3 4	0.89 0.89 0.88 0.88	inactive	2.42

6.	training	0.875	1 2 3 4	0.89 0.89 0.89 0.89	inactive	2.4
7.	test	0.892	1 2 3 4	0.92 0.93 0.92 0.91	inactive	1.16
8.	training	0.924	1 2 3 4	0.93 0.93 0.92 0.91	active	1.15
9.	training	0.919	1 2 3 4	0.95 0.94 0.94 0.94	active	2.99
10.	training	0.944	1 2 3 4	0.94 0.94 0.94 0.94	active	2.88
11.	training	0.908	1 2 3 4	0.90 0.91 0.91 0.92	active	2.42
12.	training	0.94	1 2 3 4	0.90 0.91 0.92 0.93	active	2.4
13.	training	0.903	1 2 3 4	0.92 0.90 0.91 0.91	inactive	2.68
14.	training	0.908	1 2 3 4	0.92 0.90 0.91 0.91	active	2.64
15.	training	0.924	1 2 3 4	0.92 0.91 0.92 0.92	active	2.66

16.	training	0.934	1 2 3 4	0.93 0.94 0.94 0.93	active	2.71
17.	training	0.934	1 2 3 4	0.93 0.95 0.95 0.94	active	2.38
18.	training	0.968	1 2 3 4	0.94 0.96 0.97 0.97	active	2.35
19.	training	0.908	1 2 3 4	0.91 0.91 0.90 0.91	active	2.38
20.	training	0.924	1 2 3 4	0.92 0.92 0.92 0.92	active	2.25
21.	test	0.898	1 2 3 4	0.92 0.92 0.92 0.91	inactive	2.3
22.	training	0.914	1 2 3 4	0.92 0.91 0.91 0.92	active	2.48
23.	training	0.949	1 2 3 4	0.95 0.95 0.95 0.95	active	2.7
24.	test	0.959	1 2 3 4	0.94 0.94 0.94 0.95	active	2.7

Table25: Alignment of hypotheses with other ligands (AADD)

(i) Analysis of 3D-QSAR Model: validation of pharmacophore modeling:

Reliable predictions can only come from statistically validated QSAR model. There are several statistical parameters such as Leave-n-out cross validation for training set (R^2), Leave-n-out cross validation for test set (Q^2), standard deviation (SD), root mean square error (RMSE), variance ratio (F) that can evaluate the robustness of a QSAR model as shown in table 26.

ID	# Factors	SD	R-squared	F	Stability	RMSE	Q-squared
AADD.529	1	0.0165	0.5551	22.5	0.7742	0.0272	0.0733
	2	0.0129	0.7419	24.4	0.6009	0.0279	0.0265
	3	0.012	0.791	20.2	0.7487	0.023	0.3375
	4	0.0108	0.8405	19.8	0.5659	0.0154	0.7017

Table26: QSAR results (AADD)

High R^2 is necessary but not sufficient condition for a predictive QSAR model. So, besides high R^2, best QSAR model should be chosen based on its predictive ability, so, the best model should have high Q^2 also [62]. Thus, hypotheses1 shows good R^2 value, predictive (Q^2) value and lowest RMSE value.

(ii) Analysis of scatter plot of 3-D QSAR Model

The established QSAR model gave high slope of regression lines. The performance of the QSAR model on the training and test set molecules is shown in fig 31.

The solid line in the test set plot indicates the hypothetical ‘‘best fit’’ line between the predicted and experimental binding affinity values, BA (Kcal). The scatter plot indicates a good correlation between the predicted and experimental affinities with r^2=0.85 and equation: y=0.74x + 0.24

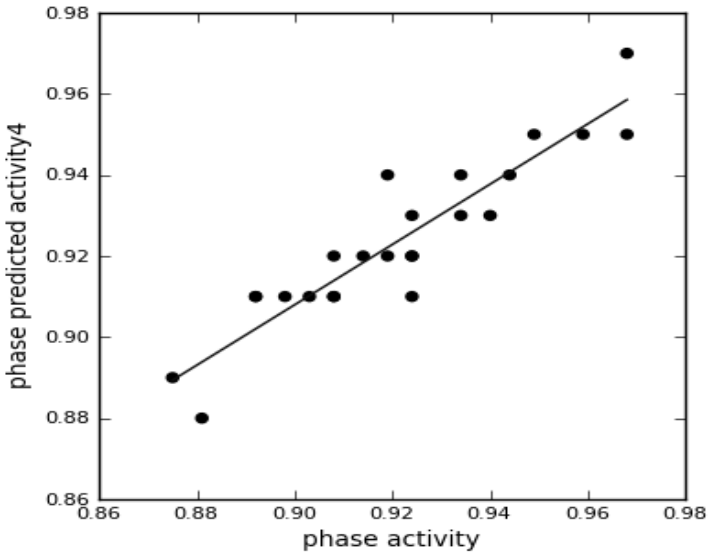


Fig31: Scatter plot for the QSAR model (AADD)

(iii) Analysis of atom based PHASE 3-D QSAR model

Additional insight into the binding activity can be gained by visualizing the QSAR model and identifying the most significant favorable and unfavorable features. The blue cubes are the features, important for the activity of the molecule and the red cubes are those indicating its potential for low activity as shown in fig 33. Thus, we can identify the features important for the interaction between ligands and their target protein. This is known as contribution map, and they allow identification of those positions that require a particular physiochemical property to enhance the bioactivity of a ligand.

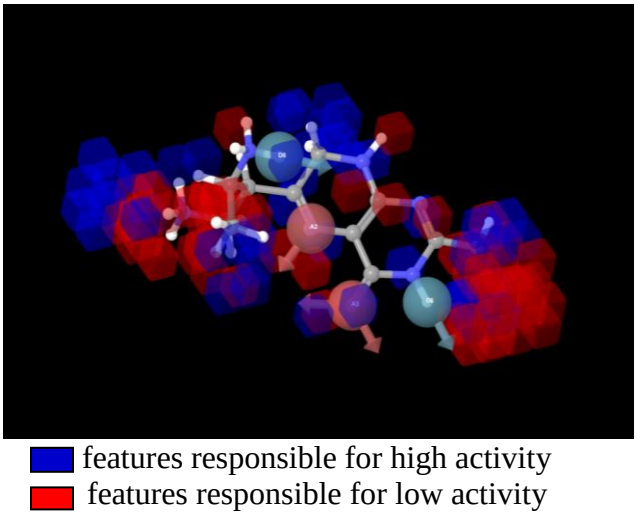


Fig32: QSAR model visualized in context of the best fit molecule in the training set (AADD)

Hypotheses 2:
Hypotheses 2 with survival-inactive score of 1.086 is shown in table 27. Pharmacophore features which were identified as common to all the 24 compounds were: A (H bond acceptor), D (H bond donator), R (aromatic ring) as shown in Fig 33.

ID	Survival	Survival - inactive	Post-hoc	Site	Vector	Volume	Selectivity	Energy	activity
AADR.5	2.916	1.086	2.916	0.54	0.758	0.623	0.99	1.661	0.949

Table27: Details of the hypotheses (AADR)

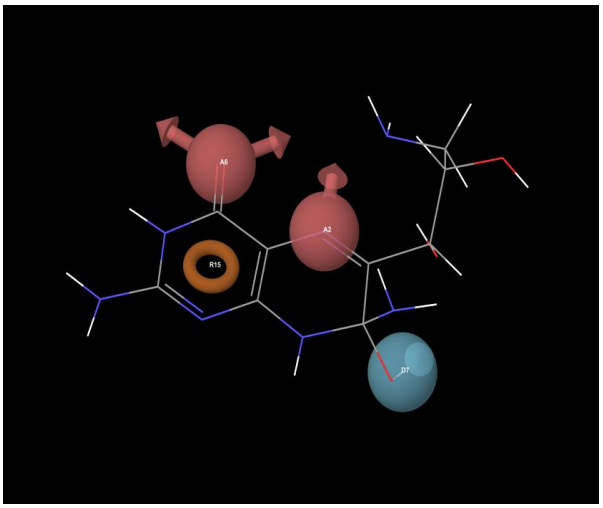


Fig33: PHASE hypotheses that yielded the common pharmacophore features (AADR)

Entry	Site1	Site2	Distance
AADR.5	A2	A6	2.879
AADR.5	A2	D7	3.339
AADR.5	A2	R15	2.8
AADR.5	A6	D7	5.963
AADR.5	A6	R15	2.669
AADR.5	D7	R15	4.934

Table28: Site measurements for distances for AADR

Entry	Site1	Site2	Site3	Angle
AADR.5	A6	A2	D7	146.9
AADR.5	A6	A2	R15	56
AADR.5	D7	A2	R15	106.7
AADR.5	A2	A6	D7	17.8
AADR.5	A2	A6	R15	60.5

AADR.5	D7	A6	R15	54.8
AADR.5	A2	D7	A6	15.3
AADR.5	A2	D7	R15	32.9
AADR.5	A6	D7	R15	26.2
AADR.5	A2	R15	A6	63.5
AADR.5	A2	R15	D7	40.4
AADR.5	A6	R15	D7	98.9

Table29: Site measurements for angles for AADR

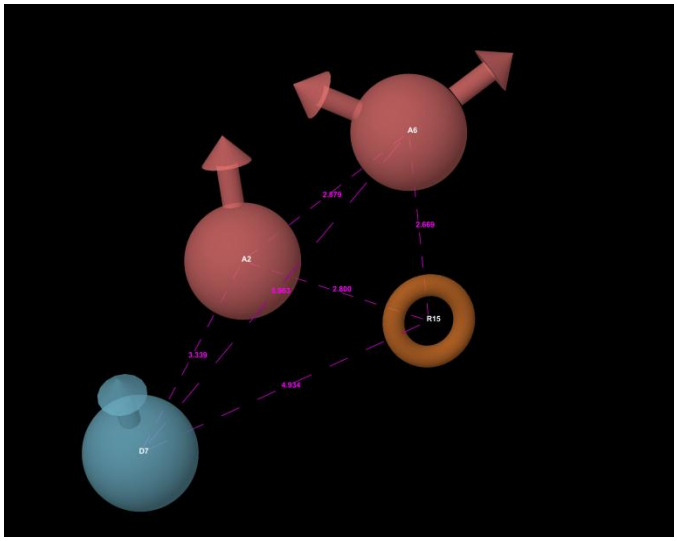


Fig34: Site measurements for QSAR model (AADR)

Then the QSAR model was built using 4 PLS factors by randomly selecting 20 compounds as training sets and rest 4 as test sets as shown in table 30.

Ligand Name	QSAR Set	Activity	# Factors	Predicted Activity	Pharm Set	Fitness
1.	test	0.892	1 2 3 4	0.92 0.91 0.92 0.92	inactive	1.41
2.	training	0.919	1 2 3 4	0.92 0.91 0.92 0.92	active	1.39

3.	training	0.968	1 2 3 4	0.94 0.94 0.94 0.95	active	1.83
4.	training	0.924	1 2 3 4	0.92 0.92 0.92 0.92	active	1.59
5.	training	0.881	1 2 3 4	0.88 0.89 0.89 0.89	inactive	1.99
6.	training	0.875	1 2 3 4	0.92 0.91 0.90 0.88	inactive	2.15
7.	test	0.892	1 2 3 4	0.92 0.93 0.92 0.92	inactive	1.39
8.	training	0.924	1 2 3 4	0.93 0.93 0.93 0.92	active	1.37
9.	training	0.919	1 2 3 4	0.94 0.94 0.94 0.95	active	1.83
10.	training	0.944	1 2 3 4	0.94 0.95 0.94 0.94	active	1.66
11.	training	0.908	1 2 3 4	0.88 0.90 0.90 0.90	active	1.99
12.	training	0.94	1 2 3 4	0.93 0.92 0.92 0.92	active	2.15

13.	training	0.903	1 2 3 4	0.91 0.90 0.91 0.90	inactive	2.09
14.	training	0.908	1 2 3 4	0.91 0.90 0.91 0.91	active	2.08
15.	training	0.924	1 2 3 4	0.93 0.93 0.93 0.93	active	1.9
16.	training	0.934	1 2 3 4	0.93 0.93 0.93 0.94	active	1.93
17.	training	0.934	1 2 3 4	0.93 0.93 0.94 0.94	active	2.08
18.	training	0.968	1 2 3 4	0.95 0.96 0.98 0.97	active	2.13
19.	training	0.908	1 2 3 4	0.92 0.91 0.90 0.91	active	2.53
20.	training	0.924	1 2 3 4	0.92 0.92 0.92 0.92	active	1.94
21.	test	0.898	1 2 3 4	0.92 0.91 0.91 0.91	inactive	1.97
22.	training	0.914	1 2 3 4	0.92 0.91 0.91 0.91	active	2.03

23.	training	0.949	1 2 3 4	0.94 0.95 0.95 0.95	active	3
24.	test	0.959	1 2 3 4	0.94 0.94 0.94 0.94	active	2.88

Table30: Alignment of hypotheses with other ligands (AADR)

(i) Analysis of 3D-QSAR Model: validation of pharmacophore modeling

Reliable predictions can only come from statistically validated QSAR model. There are several statistical parameters such as Leave-n-out cross validation for training set (R^2), Leave-n-out cross validation for test set (Q^2), standard deviation (SD), root mean square error (RMSE), variance ratio(F) that can be used to evaluate the robustness of a QSAR mode as shown in table 31.

ID	# Factors	SD	R-squared	F	Stability	RMSE	Q-squared
AADR.5	1	0.0168	0.5366	20.8	0.7038	0.0249.	0.2214
	2	0.0141	0.6896	18.9	0.4016	0.0227	0.352
	3	0.0127	0.7642	17.3	0.2699	0.0228	0.3507
	4	0.0116	0.8158	16.6	0.147	0.0211	0.4421

Table31: QSAR results (AADR)

High R^2 is necessary but not sufficient condition for a predictive QSAR model. So, besides high R^2, best QSAR model should be chosen based on its predictive ability, so, the best model should have high Q^2 also. Thus, hypotheses1 shows good R^2 value, predictive (Q^2) value and lowest RMSE value.

(ii) Analysis of scatter plot of 3-D QSAR Model

The established QSAR model gave high slope of regression lines. The performance of the QSAR model on the training and test set molecules is shown in fig 35.

The solid line in the test set plot indicates the hypothetical ‘‘best fit’’ line between the predicted and experimental binding affinity values, BA (Kcal). The scatter plot indicates a good correlation between the predicted and experimental affinities with r^2=0.73 and equation: y=0.71x + 0.27

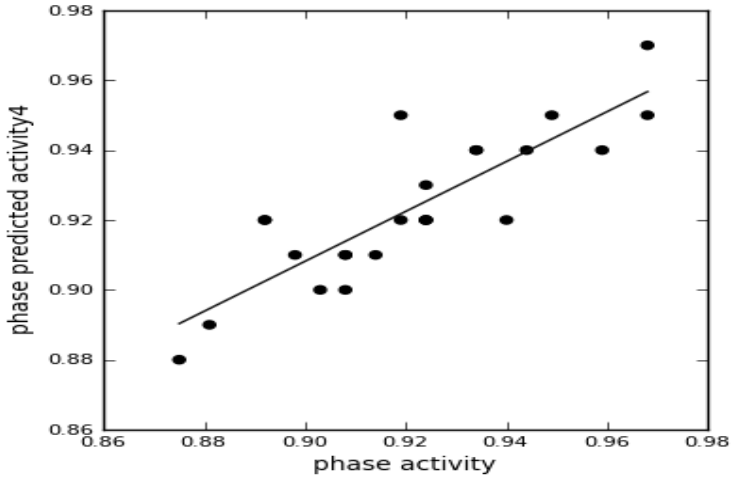


Fig35: Scatter plot for the QSAR model (AADR)

(iii) Analysis of atom based PHASE 3-D QSAR model

Additional insight into the binding activity can be gained by visualizing the QSAR model and identifying the most significant favorable and unfavorable features. The blue cubes are the features, important for the activity of the molecule and the red cubes are those indicating its potential for low activity as shown in fig 33. Thus, we can identify the features important for the interaction between ligands and their target protein. This is known as contribution map, and they allow identification of those positions that require a particular physiochemical property to enhance the bioactivity of a ligand.

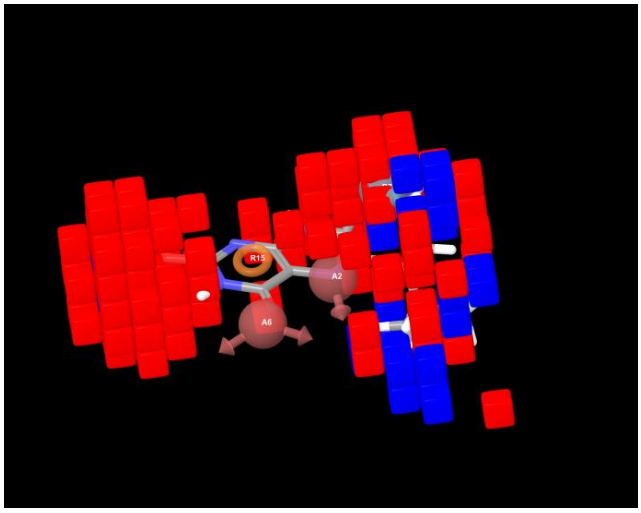


Fig36: QSAR model visualized in context of the best fit molecule in the training set (AADR)

Thus, the results indicated that QSAR models possessed a high accommodating capacity; they may be reliable for being used to predict the activities of new derivatives. Hence, both the hypotheses were retained for further studies.

Since the phamacophore models define the essential features of the molecules, thus they can be used to identify or search for novel ligands that will bind to the same receptor.

3. Therapeutic application of cofactor analogues

We have also considered a mutant phenylalanine hydroxylase (PDB ID 1TDW) in which Ala at position 313 is mutated with Thr in order to see the responsiveness of structurally modifies cofactor analogues. This is because, Phenylketonuria (PKU) patients having a subset of mutations in phenylalanine hydroxylase, have also shown normalization of blood phenylalanine level upon oral administration of cofactor. Cofactor treatment instead of a phenylalanine-restricted diet might be possible in many patients and would be expected to improve their quality of life substantially [63]. Interestingly, BH4 has been administered as a pharmacological treatment and has been shown to reduce blood levels of phenylalanine for a segment of PKU patients whose genotypes lead to some residual PAH activity but have no defect in BH4 synthesis or regeneration. Follow-up studies also suggest that in the case of certain PheOH mutants, excess BH4 acts as a pharmacological chaperone to stabilize mutant enzymes with disrupted tetramer assembly and increased sensitivity to proteolytic cleavage and aggregation [64].

So, finally we performed docking study of these 24 compounds with mutant phenylalanine hydroxylase (1TDW)

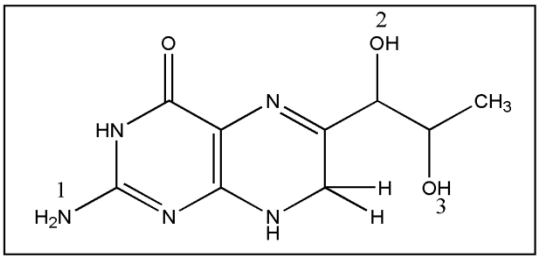


Fig11: Structure of BH_2 showing positions 1, 2 and 3

COMPOUND	POSITION 1	POSITION 2	POSITION 3	DOCKING SCORE	RESIDUES and DISTANCE (Å)								
					GLY 247 (NH_2)	LEU 249 (NH)	LEU 249 (N)	SER 251 (NH_3^+)	SER 251 (OH)	TYR 377 (NH_3^+)	TYR 377 (OH)	ALA 322 (NH_3^+)	ALA 322 (OH)
26.	**CH_3**	**NH_2**	**NH_2**	**-8.6**									
a.	C_2H_5	NH_2	NH_2	-8.5	-	1.842	2.225	-	-	-	-	-	-
b.	C_3H_7	NH_2	NH_2	-10.8	-	-	-	-	-	-	-	-	-
c.	CH_3	CH_2NH_2	NH_2	-9.8	-	-	-	-	-	-	-	-	-
d.	CH_3	$C_2H_4NH_2$	NH_2	-9.6	-	-	-	-	-	-	-	-	-
e.	CH_3	NH_2	CH_2NH_2	-7.4	-	-	-	-	-	-	-	-	-
f.	CH_3	NH_2	$C_2H_4NH_2$	-9.0	-	1.765	2.143	-	2.172	-	-	-	-
27.	**CH_3**	**OH**	**OH**	**-7.9**									
a.	C_2H_5	OH	OH	-8.4	-	1.759	2.171.	-	-	-	-	-	-
b.	C_3H_7	OH	OH	-8.8	-	1.752	2.185	-	-	-	-	-	-
c.	CH_3	CH_2OH	OH	-8.5	-	1.771	2.091	-	-	-	-	-	-
d.	CH_3	C_2H_4OH	OH	-9.3	-	1.812	2.097	-	-	-	-	-	-
e.	CH_3	OH	CH_2OH	-8.7	-	1.836	2.254	-	-	-	-	-	-
f.	CH_3	OH	C_2H_4OH	-9.8	-	1.720	2.076	-	-	-	-	-	-
28.	**OH**	**NH_2**	**NH_2**	**-8.9**									
a.	CH_2OH	NH_2	NH_2	-7.0	2.010	2.008	2.361	-	-	-	-	-	-
b.	C_2H_4OH	NH_2	NH_2	-6.0	-	2.095	2.465	-	-	-	-	2.136	-
c.	OH	CH_2NH_2	NH_2	-7.5	2.204	1.791	2.118	2.102	-	-	-	-	-
d.	OH	$C_2H_4NH_2$	NH_2	-9.6	1.884	1.863	2.127	-	-	-	-	-	-
e.	OH	NH_2	CH_2NH_2	-7.7	-	1.810	2.132	-	-	-	-	1.852	-
f.	OH	NH_2	$C_2H_4NH_2$	-8.5	1.929	1.808	2.078	-	-	-	-	-	-

Table32: Docking score of the interactions between different analogs and the residues with PDB ID 1TDW, along with their distances from the interacting groups

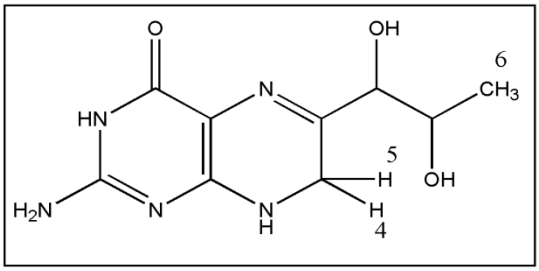


Fig12: Structure of BH_2 showing positions 4, 5 and 6

COMPOUND	POSITION 4	POSITION 5	POSITION 6	DOCKING SCORE	RESIDUES and DISTANCE (Å)								
					LEU 249 (N)	LEU 249 (NH)	GLY 247 (NH2)	SER 251 (NH3+)	SER 251 (OH)	ALA 322 (NH3+)	ALA 322 (OH)	TYR 377 (NH3+)	TYR 377 (OH)
29.	**OH**	**NH_2**	**NH_2**	**-8.0**									
a.	CH_2OH	NH_2	NH_2	-10.1	2.154	1.901	1.957	-	1.913	-	-	-	-
b.	C_2H_4OH	NH_2	NH_2	-9.9	2.098	1.924	1.894	-	2.068	-	-	-	-
c.	OH	CH_2NH_2	NH_2	-10.6	2.128	1.904	1.833	-	1.990	-	-	-	-
d.	OH	$C_2H_4NH_2$	NH_2	-11.1	2.085	2.047	-	-	1.862	-	-	-	-
e.	OH	NH_2	CH_2NH_2	-10.4	2.134	1.812	1.742	-	-	-	-	-	-
f.	OH	NH_2	$C_2H_4NH_2$	-10.7	2.106	1.802	1.807	-	2.434		-		-

Table33: Docking score of the interactions between different analogs and the residues with PDB ID 1TDW, along with their distances from the interacting groups

S.NO	POSITION 1	POSITION 2	POSITION 3	1DMW	1MLW	1TDW
1.	C_2H_5	NH_2	NH_2	-7.8	-8.3	**-8.5**
2.	C_3H_7	NH_2	NH_2	-8.4	-8.4	**-10.8**
3.	CH_3	CH_2NH_2	NH_2	-8.3	-7.8	**-9.8**
4.	CH_3	$C_2H_4NH_2$	NH_2	-8.8	-8.1	**-9.6**
5.	CH_3	NH_2	CH_2NH_2	-8.1	-7.8	-7.4
6.	CH_3	NH_2	$C_2H_4NH_2$	-8.7	-8.4	**-9.0**
7.	C_2H_5	OH	OH	-8.0	-8.2	**-8.4**
8.	C_3H_7	OH	OH	-8.1	-8.1	**-8.8**
9.	CH_3	CH_2OH	OH	-8.4	-8.5	**-8.5**
10.	CH_3	C_2H_4OH	OH	-8.6	-8.5	**-9.3**
11.	CH_3	OH	CH_2OH	-8.6	-8.8	-8.7
12.	CH_3	OH	C_2H_4OH	-9.3	-8.9	**-9.8**
13.	CH_2OH	NH_2	NH_2	-7.7	-7.8	-7.0
14.	C_2H_4OH	NH_2	NH_2	-8.1	-8.3	-6.0
15.	OH	CH_2NH_2	NH_2	-8.0	-9.3	-7.5
16.	OH	$C_2H_4NH_2$	NH_2	-9.1	-8.4	**-9.6**
17.	OH	NH_2	CH_2NH_2	-8.5	-7.6	-7.7
18.	OH	NH_2	$C_2H_4NH_2$	-8.4	-7.5	**-8.5**
S.NO	**POSITION 4**	**POSITION 5**	**POSITION 6**	**1DMW**	**1MLW**	**1TDW**
19.	CH_2OH	NH_2	NH_2	-8.1	-6.8	**-10.1**
20.	C_2H_4OH	NH_2	NH_2	-8.4	-6.5	**-9.9**
21.	OH	CH_2NH_2	NH_2	-7.9	-7.8	**-10.6**
22.	OH	$C_2H_4NH_2$	NH_2	-8.2	-6.8	**-11.1**
23.	OH	NH_2	CH_2NH_2	-8.9	-7.9	**-10.4**
24.	OH	NH_2	$C_2H_4NH_2$	-9.1	-10.1	**-10.7**

Table34: Comparison of docking scores of 24 compounds for all 3 PDB IDs (1DMW, 1MLW, 1TDW)

The table suggests that docking scores of 18 out of 24 compounds for PDB ID 1TDW were higher in comparison to 1DMW and 1MLW. This shows that PKU patients will also be responsive for these structurally modified analogues of the cofactor. Thus, we can propose that high-dose cofactor treatment may compensate for the decreased affinity of the mutant phenylalanine hydroxylase for the cofactor and can prove to be useful for the correction of impaired phenylalanine hydroxylation, already known as cofactor therapy [65].

7. CONCLUSION

The purpose of this study was to understand the mechanistic aspect of interaction between structurally modified cofactor (dihydrobiopterin) analogues with different amino acid hydroxylase enzymes like phenylalanine hydroxylase, tyrosine hydroxylase and tryptophan hydroxylase via combined approach of ligand-recpetor docking and pharmacophores modeling. Most of the newly designed molecules were found to show interactions with phenylalanine hydroxylase enzyme residues: Arg270, Glu280, Thr278, Pro279, Gly346, Ser349, Glu353, Val379 and Fe425 and the best results obtained with docking scores were -9.1 for 5 (d) and 7 (f). Similarly, Most of the newly designed molecules were found to show interactions with tryptophan hydroxylase enzyme residues: Tyr125, Leu236, Thr265, Pro266, Glu317, Gly333, Ser336, Ser337, Glu340, Thr367 and Thr368 and the best results obtained with docking scores were -9.3 for compound 18 (c), -9.4 for compound 21 (a), -9.5 for compound 23 (c) and -10.1 for compound 22 (f).

In addition, a pharmacophore modeling approach was applied to identify the essential features common to all the screened compounds on the basis of their Gscore. Different pharmacophores hypotheses of cofactor analogues of amino acid hydroxylases were developed using PHASE and alignment based on these pharmacophores hypothesis were used as input for the development of 3D-QSAR models. A four point pharmacophore hypothesis AADD and another pharmacophore hypothesis AADR with correlation coefficient of 0.85 and 0.73 respectively were associated with a 3D-QSAR model with good statistical significance and good predictive ability.

Further analysis of the docking results of these best screened compounds with the mutant phenylalanine hydoxylase enzyme suggested the therapeutic application of these compounds. Thus, it would be concluded that these compounds could be potential selective cofactors of amino acid hydroxylases that could be validated through some wet lab experimentally. The obtained model can also be employed for 3D search query to screen against compound libraries in order to identify some more scaffolds. And their backbone structure scaffold could serve as building blocks in designing cofactors for amino acid hydroxylases.

Hence, Structure based drug design (docking) and Ligand based drug design (pharmacophores modeling) play a very important role in drug designing and produces some exciting results if implemented together in order to strengthen drug designing. In summary, both SBDD and LBDD 3D-QSAR model presented in this study could be very useful for designing new molecules which can bind with the amino acid hydroxylase enzymes similarly as the cofactor.

8. FUTURE PERSPECTIVE

One of the main goals in drug discovery is the identification of innovative small molecular scaffolds exhibiting high binding affinity and selectivity for the target together with a reasonable absorption, distribution, metabolism, excretion and toxicity (ADMET) profile, lead and/or drug likeness. Such chemical entities are likely to be able to enter higher phases of the drug development process. Lipinski's rule of 5 is a rule of thumb to evaluate drug likeliness or to determine if a chemical compound with a certain pharmacological or biological activity has properties that would make it a likely orally active drug in humans. The rule describes molecular properties important for a drug's pharmacokinetics in the human body, including its ADMET. Thus, *in vitro* approaches can be used to investigate the ADMET properties of these newly designed chemical entities and to optimize selection of the most suitable candidates for drug development.

Thus, the QikProp program can be further used to obtain the ADMET properties of the analogues. It will predict both physically significant descriptors and pharmaceutically relevant properties.
Also, virtual screening on the basis of pharmacophore model can be done in order to propose various drug molecules, which have all the essential features. Then wet lab synthesis of these compounds as well as characterization via techniques like H^1 N.M.R and L.C.M.S can be done.

Further, the acceptability of the analogues can also be evaluated as this also plays an essential role for rational drug design. Since, drug safety evaluation is an important issue in new drug discovery. Given a low success rate of drug candidates, the compounds can also be tested for safety as detection of potential toxicity and side effect in early stages of drug development can potentially save money and time by focusing resources on only those drug leads and candidates that are likely to be safe to patients.

9. REFERENCES

[1] Kaufman. S “A new cofactor required for the enzymatic conversion of phenylalanine to tyrosine” J. Biol. Chem., 1958, vol.230 (2), pg 931–939.

[2] Kaufman. S “New tetrahydrobiopterin-dependent systems” Annual review of nutrition, 1993, vol.13, pg 261-286.

[3] Shiman. R “Folates and pterins” John wiley & sons: New York, 1985, vol. 2, pg 179-249.

[4] Kaufman. S “ Folates and pterins” John wiley & sons: New York,1985, vol.2, pg 251-352.

[5] Grahame-smith “Tryptophan hydroxylation in brain” Biochem biophys res commun, 1964, vol.16, pg586-92.

[6] Kaufman. S “Tyrosine hydroxylase” Adv enzymol relat areas, mol biol, 1995, vol.70, pg103-220.

[7] Walter. L, Eric. J, Albert. S “Tryptophan hydroxylation: measurement in pineal gland, brainstem, and carcinoid tumor” Science, 1967, vol.155, pg217-219.

[8] Erlandsen. H, Fusetti. F, Martinez. A “Crystal structure of the catalytic domain of human phenylalanine hydroxylase reveals the structural basis for phenylketonuria” Nat struct biol., 1997, vol.4, pg995-1000.

[9] Kaufman. S "A model of human phenylalanine metabolism in normal subjects and in phenylketonuric patients" Proc. natl. acad. sci. u.s.a, 1999, vol.96, pg3160-3164.

[10] Michals. K, Matalon. R "phenylalanine metabolites, attention span and hyperactivity"American journal of clinical nutrition, 1985, vol.42, pg361–365.

[11] Michals-Matalon. K "Sapropterin dihydrochloride, 6-R-L-erythro-5,6,7,8-tetrahydrobiopterin, in the treatment of phenylketonuria" Expert Opin Investig Drugs, 2008, vol.17, pg245–251.

[12] Burton. B. K, Grange. D. K, Milanowski. A, Vockley. G "The response of patients with phenylketonuria and elevated serum phenylalanine to treatment with oral sapropterin dihydrochloride (6R-tetrahydrobiopterin): a phase II, multicentre, open-label, screening study" Journal of Inherited Metabolic Disorders, 2007, vol.30, pg700–707.

[13] Lee. P, Treacy. E, Crombez. E et al. " Safety and efficacy of 22 weeks of treatment with sapropterin dihydrochloride in patients with phenylketonuria". Am J Med Genet, 2008, vol.146A, pg2851–2859.

[14] Daubner. S. C, Le. T, Wang. S "Tyrosine hydroxylase and regulation of dopamine synthesis" Archives of Biochemistry and Biophysics, 2011, vol.508, pg1–12.

[15] Goodwill. K, Sabatier. C, Marks. C, "crystal structure of tyrosine hydroxylase at 2.3 Å and its implications for inherited neurodegenerative diseases" Nature structural biology, 1997, vol. 4, pg578–585.

[16] Wang. L, Erlandsen. H, Haavik. J, Knappskog. P. M, Stevens. R. C "Three-dimensional structure of human tryptophan hydroxylase and its implications for the biosynthesis of the neurotransmitters serotonin and melatonin" Biochemistry, 2002, vol.41, pg12569–12574.

[17] Vanhoutte. P, "Serotonin and the vascular system" Review article, Raven press: New York, 1985, vol.31.

[18] Kazuhiro. N, Yuko. S "Late Developmental Stage-Specific Role of Tryptophan Hydroxylase 1 in Brain Serotonin Levels" The Journal of Neuroscience, 2006, vol.26, pg530 – 534

[19] Eric .L, Hegg "The 2-his-1-carboxylate facial triad — an emerging structural motif in mononuclear non-heme iron(ii) enzymes" , Eur. J. Biochem, 1997, vol.250, pg625–629.

[20] Teigen. K, McKinney. J. A, Haavik. J, Martínez. A "Selectivity and affinity determinants for ligand binding to the aromatic amino acid hydroxylases" Current medicinal chemistry, 2007, vol.14, pg455–467.

[21] Li. J, Dangott. L. J, Fitzpatrick. P. F "Regulation of phenylalanine hydroxylase: conformational changes upon phenylalanine binding detected by hydrogen/deuterium exchange and mass spectrometry" Biochemistry, 2010, vol.49, pg3327–35.

[22] Li. J, Ilangovan. U, Daubner. S. C, Hinck. A. P, "Direct evidence for a phenylalanine site in the regulatory domain of phenylalanine hydroxylase" Arch. Biochem. Biophys, 2011, vol.505, pg250–255.

[23] Flatmark. T, Stevens. R "Structural insight into the aromatic amino acid hydroxylases and their disease-related mutant forms" Chem. Rev, 1999, vol.8, pg 2137 2160.

[24] Erlandsen. H, Flatmark. T, Stevens. R "crystal structure and site-specific mutagenesis of pterin-bound human phenylalanine hydroxylase" Biochemistry, 2000, vol.39, pg2208-2217.

[25] Dickson. P. W, Jennings. I. G, Cotton. R. G "delineation of the catalytic core of phenylalanine hydroxylase and identification of glutamate 286 as a critical residue for pterin function" J. biol. Chem, 1994, vol.269, pg20369-20375.

[26] Kappock. T. J, Caradonna. J. P “pterin-dependent amino acid hydroxylases” Chem rev, 1996, vol.96, pg2659-2756.

[27] Ramsey. A. J, Daubner. S. C, Ehrlich. J. I, Fitzpatrick. P. F "Identification of iron ligands in tyrosine hydroxylase by mutagenesis of conserved histidinyl residues" Protein science, 1995, vol. 4, pg2082–2086.

[28] Kenneth. E, Goodwill, Christelle. S, Stevens. R “Crystal structure of tyrosine hydroxylase with bound cofactor analogue and iron at 2.3 å resolution: self-hydroxylation of phe300 and the pterin-binding site” Biochemistry, 1998, vol. 37.

[29] Wang. L, Erlandsen. H, Haavik. J, Stevens. R “Three-dimensional structure of human tryptophan hydroxylase and its implications for the biosynthesis of the neurotransmitters serotonin and melatonin” Biochemistry, 2002, vol. 41.

[30] Fitzpatrick. P. F "Mechanism of aromatic amino acid hydroxylation" Biochemistry, 2003, vol.42, pg14083–14091.

[31] Panay AJ, Lee M, Krebs C, Bollinger JM, Fitzpatrick PF "Evidence for a high-spin Fe(IV) species in the catalytic cycle of a bacterial phenylalanine hydroxylase". Biochemistry, 2011, vol.50, pg1928–1933.

[32] Olsson E, Martinez A, Teigen K, Jensen VR "Formation of the iron-oxo hydroxylating species in the catalytic cycle of aromatic amino acid hydroxylases" Chemistry, 2011, vol.17, pg3746–58.

[33] Pavon. J. A, Fitzpatrick. P. F "Insights into the catalytic mechanisms of phenylalanine and tryptophan hydroxylase from kinetic isotope effects on aromatic hydroxylation" Biochemistry, 2006, vol. 45, pg11030–11037.

[34] http://genome.wellcome.ac.uk/doc_WTD020912.html

[35] http://en.wikipedia.org/wiki/Drug_design

[36] John. T “Molecular Modeling in Drug Design” Journal of Medicinal Chemistry, 1988, vol.31, pg2229.

[37] Hornak. V, Dvorsky. R, Sturdik. E “Receptor - ligand Interaction and Molecular Modelling” Gen Physiol Biophys, 1999, vol.1 8, pg231-248.

[38] Donald J. Abraham, Burger's Medicinal Chemistry and Drug Discovery, sixth edition, vol.1.

[39] http://www.proxychem.com/sbdd.html\

[40] http://www.pdb.org/pdb/home/home.do
[41] Halgren. T. A, Murphy. R. B, Friesner. R. A, Beard. H. S, Frye. L. L, Pollard. W. T, Banks. J. L "Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening" J. Med. Chem, 2004, vol.47, pg1750-1759.
[42] Jorgensen. W. L, Maxwell. D. S, Tirado. R. J "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids" J. Am. Chem. Soc, 1996, vol.118, pg11225-11236.

[43] Simone. B "Pharmacophore modeling: A continuously evolving tool for computational drug design" New perspectives in medicinal chemistry, 2009, vol.1, pg13-23.

[44] Wolber. G " Molecule-pharmacophore superpositioning and pattern matching in computational drug design" Drug Discov. Today, 2008, vol.13, pg23–29.

[45] Dror. O "Predicting molecular interactions in silico. I. An updated guide to pharmacophore identification and its applications to drug design" Front. Med. Chem, 2006, vol.3, pg551–584.

[46] Robert. S. P, Michael. A. P, Kaufman. S " Spectroscopic Investigation of Ligand Interaction with Hepatic Phenylalanine Hydroxylase: Evidence for a Conformational Change Associated with Activation" Biochemistry, 1984, vol.23, pg3836-3842.

[47] Heidi. E, Torgeir. F, Raymond. C. S " Crystallographic Analysis of the Human Phenylalanine Hydroxylase Catalytic Domain with Bound Catechol Inhibitors at 2.0 Å Resolution" Biochemistry, 1998, vol.37, pg15638-15646.

[48] Knut. T, Nils. A, Aurora. M "The Structural Basis of the Recognition of Phenylalanine and Pterin Cofactors by Phenylalanine Hydroxylase: Implications for the Catalytic Mechanism" J. Mol. Biol., 1999, vol.294, pg807-823.

[49] Torben. G, Marie. P, Per. G "Missense Mutations in the N-Terminal Domain of Human Phenylalanine Hydroxylase Interfere with Binding of Regulatory Phenylalanine" Am. J. Hum. Genet, 2001, vol.68, pg1353–1360.

[50] Heidi. E, Joo. Y. K, Marianne. G "Structural Comparison of Bacterial and Human Irondependent Phenylalanine Hydroxylases: Similar Fold, Different Stability and Reaction Rates" J. Mol. Biol, 2002, vol.320, pg645–661.

[51] Ole. A, Anne J. S, Torgeir. F "2.0 A° Resolution Crystal Structures of the Ternary Complexes of Human Phenylalanine Hydroxylase Catalytic Domain with Tetrahydrobiopterin and 3-(2-Thienyl)-L-alanine or L-Norleucine: Substrate Specificity and Molecular Motions Related to Substrate Binding" J. Mol. Biol, 2003, vol.333, pg747–757.

[52] Fitzpatrick. P. F "Tetrahydropterin-Dependent Amino Acid Hydroxylases", Annual Review of Biochemistry, 1999, vol. 68, pg355–381.

[53] Bobrovskaya. L, Gilligan. C, Bolster. E. K, "Sustained phosphorylation of tyrosine hydroxylase at serine 40: A novel mechanism for maintenance of catecholamine synthesis" Journal of Neurochemistry, 2007, vol.100, pg479–489.

[54] Hitoshi. F, Sachiko. O "Regulatory mechanism of tyrosine hydroxylase activity" review, Biochemical and Biophysical Research Communications, 2005, vol.338, pg271–276.

[55] Daubner. S. C, James. T, Meredith. G "A Flexible Loop in Tyrosine Hydroxylase Controls Coupling of Amino Acid Hydroxylation to Tetrahydropterin Oxidation" J. Mol. Biol, 2006, vol.359, pg299–307.

[56] Jiang. G, Yohrling. G "Identification of Substrate Orienting and Phosphorylation Sites Within Tryptophan Hydroxylase Using Homology-based Molecular Modeling" J. Mol. Biol. 2000, vol.302, pg1005-1017.

[57] Jeffrey. M, Knut. T, Nils. A "Conformation of the Substrate and Pterin Cofactor Bound to Human Tryptophan Hydroxylase. Important Role of Phe313 in Substrate Specificity" Biochemistry, 2001, vol.40, pg15591-15601.

[58] Lærke. T. H, Kasper. P. J "Experimentally calibrated computational chemistry of tryptophan hydroxylase: Trans influence, hydrogen-bonding, and 18-electron rule govern O2-activation" Journal of Inorganic Biochemistry, 2010, vol.104, pg136–145.

[59] Ligprep 2.0 (2006) Schrodinger, LLC, New York, NY

[60] Steven L. D "PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results" J Comput Aided Mol Des, 2006, vol.20, pg647–671.

[61] Das. N, Dhanawat. M, Shrivastava. S. K "Benzoxazinones as Human Peroxisome Proliferator Activated Receptor Gamma (PPARγ) Agonists: A Docking Study Using Glide" Indian J Pharm Sci, 2011, vol.73, pg159–164.

[62] Tropsha. A "Best Practices for QSAR Model Development, Validation, and Exploitation" Molecular Informatics, 2010, vol.29, pg476-488.

[63] Ania. C. M "Tetrahydrobiopterin as an alternative treatment for mild phenylketonuria" N Engl J Med, 2002, vol. 347.

[64] Muntau. A. C, Gersting. S. W "Phenylketonuria as a model for protein misfolding diseases and for the development of next generation orphan drugs for patients with inborn errors of metabolism". J. Inherit. Metab. Dis, 2010, vol. 33, pg649–658.

[65] Smith. I, Hyland. K, Kendall. B "Clinical role of pteridine therapy in tetrahydrobiopterin Deficiency" J Inherit Metab Dis, 1985, vol.1, pg39-45.