بسم الله الرحمن الرحيم

SUDAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATE STUDIES

A Study of QSAR, Docking and Molecular Modeling of some Dihydropyrimidine Derivatives as Anticancer Agents

دراسة العلاقة الكمية بين البنية والفعالية والترسية والنمذجة الجزيئية لبعض مشتقات ثنائي هيدروبيرميدين عواملا مضادة للسرطان

By

ESRA MOHAMMED KHAIR SALAH ELDEIN MOHAMMED

(BSc.Chemistry)

A thesis submitted in partial fulfillment of the requirements of Master degree in chemistry at Sudan University of Science and Technology.

Supervisor:

Prof .Dr. Ahmed Elsadig Mohammed Saeed

Chemistry Department

College of Science

March .2020

Approval Page

(To be completed after the college council approval)

Name of Candidate: Esra | Mohammed | Khair | Dalaheldeen

Thesis title: A Study of QSAR, Docking and Molecular Modeling of Some Dihydropyrimidine Derivatives as Anti cancer Agents

دراسة العلاقة الكمية بين البنية والفعالية والنمذجة الجزيئية لبعض مشتقات ثنائية هيدروبيريمدين عوامل مضادة للسرطان

Degree Examined for: M.Sc. Chem

Approved by:

1. External Examiner

Name: Dr. Mayeen Mohammed Al mahdi

Signature: ..... Date: 26-11-2020

2. Internal Examiner

Name: Dr. Mohammed Elbushra Hassan Ahmed

Signature: ..... Date: 26-11-2020

3. Supervisor

Name: Ahmed E.M. Saeed

Signature: ..... Date: 26-11-2020

A Study of QSAR, Docking and Molecular Modeling of some Dihydropyrimidine Derivatives as Anticancer Agents

دراسة العلاقة الكمية بين البنية والفعالية والترسية والنمذجة الجزيئية لبعض مشتقات ثنائي هيدروبيروميدين عواملاً مضادة للسرطان

قال الله تعالى

(واتقوا يوما ترجعون فيه إلى الله ثم توفى كل نفس ما كسبت وهم لايظلمون)

صدق الله العظيم

سورة البقرة الاية 281)

Dedication

This work is dedicated

To my father and my mother for always loving and supporting me

To my family

To my friends

Esra

ACKNOWLEDGMENT

First of all thanks for ALLAH for everything.

And my sincere thanks to my supervisor:

Prof. Dr. Ahmed Elsadig Mohammed Saeed, for suggesting the idea of this work and guidance during all stage of research.

My thanks to Tawassl Taj Elsir Hassan, Department of Chemistry, college of science, Sudan University of Science and Technology. I am deeply indebted to her for useful help

I am also grateful to any person help me in Sudan University of Science and Technology

Abstract

Dihydropyrimidine derivatives are very important in synthetic medicinal chemistry because of their wide biological ranges in chemo therapy and drugs; this importance led to consider the activity of newly designed of dihydropyrimidine derivatives as macrophage migration (MIF) enzyme inhibitor.

In this work the set of data was used to study quantitative structure activity relationship (QSAR) of dihydropyrimidine derivatives. The descriptors in two dimensional were removed because R^2<0.7. The models obtained in three dimensional can be used to predict the activity of newly designed dihydropyrimidine derivatives against colon cancer as macrophage migration enzyme factor. A highly descriptive and predictive QSAR model was obtained through calculation of independent descriptors using MOE2009.10 software. A training set composed of 13compounds and obtained by partial least squares (PLS) analysis resulted in a model displaying a squared correlation coefficient (R^2) of 0.889 (PIC_{50}=5.213+ 0.373 npr^2 -0.674MNDO-LUMO). Validation of this model was performed using method leave-one-out (LOO) giving Q^2 of 0.796 and R^2_{pre} of 0.979 for a test set of 3 compounds.

This model was used to predict the biological activity of 94 new designed dihydropyrimidine derivatives and CHS828 was used as a reference compound. From these compounds 71 compounds which had a higher biological activity than a reference compound, were selected to study their affinity to colon cancer.

Docking study for data set of 71 new designed compounds having high predicted activity to evaluate their interaction with protein of (MIF). In this study a number of compounds showed a goodness of interactions.

الخلاصة

مشتقات ثنائي هيدروبيرميدين مهمه جدا في الكيمياء الدوائية الصناعية بسبب نطاقها البيولوجي الواسع في العلاجات الكيمائيه والأدوية وقد أدت هذه الاهمية إلي دراسة نشاط مشتقات ثنائي هيدروبيرميدين المصممة كمثبطات انزيم هجرة البلاعم(MIF).

في هذه الدراسة مجموعة من البيانات استخدمت لإجراء دراسة كمَّية لعلاقة البنية بالفعالية QSAR لمشتقات ثنائي هيدروبيرميدين تم الحصول عليها من ورقة علمية منشورة تمتلك هيكل أساسي مماثل. تم الحصول على نموذج QSAR ذو وصفية وتنبؤية عالية من خلال حساب واصفات مستقلة باستخدام برنامج MOE2009.10 . تم تكوين مجموعة التدريب من 13 مركب بطريقة تحليل المربعات الصغرى(PLS) نموذجا يعرض مربع معامل الارتباط R مقداره0.889 -PIC_{50}=5.213+ 0.373 npr^2-0.674MNDO (LUMO).

تم إجراء التحقق من صحة هذا النموذج بإستخدام طريقة (LOO) وحسبت قيمة (Q^2) ومقدارها 0.796 وقيمة(R^2_{pred}) وقدرها0.979 لمجموعة الإختبار المكونة من 3 مركبات .

تم استخدام هذا النموذج للتنبؤ بالنشاط البيولوجي لـ 94 مشتق لثنائي هيدروبيرميدين الجديدة . من هذه المشتقات تم عزل71 لهم نشاط بيولوجي اعلى من المركب المرجعي.

تم تنفيذ دراسة الالتحام الجزيئي لكل من مركبات مجموعة البيانات التي تم تصميمها وتمتلك نشاط متوقع بيولوجي عاملي من أجل دراسة ارتباطها مع بروتين انزيم MIF خلال هذه الدراسة اظهرت بعض المركبات ارتباطا عاليا .

TABLE OF CONTENTS

CHAPTER TWO

Material and Methods

Appendixes

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ABBREVIATION	MENING
ASA-	Negative accessible surface area
ATP	adenosine tri phosphate
A^0	Bond length
b.p	boiling point
CADD	Computer aided drug design
CoMFA	Computer molecular field analysis
CoMSIA	Comparative molecular similarly index analysis
DHP	dihydropyridine
DHPM	dihydropyrimidine
DNA	deoxy ribonucleic acid
E	potential energy
FcasA	fractional negative accessible surface area
F	Fischers value
glob	molecular globularity
LOO	leave one out
Lys	Lysine
MLR	Multiple linear regression
MOE	Molecular operating enviroment
PLS	Partial least square
PCA	Principal component analysis comparative
QSAR	Quantitative structure activity relation ship
Q^2	Cross validation regression coefficient

R	correlation coefficient
R^2	square of correlation coefficient
RNA	Ribonucleic acid
Rgyr	radius of gyration
S	Free bending energy
SAR	Structure activity relation ship
SoMFA	Self organizing molecular field analysis
SEE	Error of estimation
SPSS	Statistical package for social science
Tyr	Tyrosine
HBV	hepatitis B virsus
MIF	Macrophage migration inhibitory factor
Vsurf	Hydropholic integy moment
Dens	Mass density
MINDO-LUMO	LUMO energy
Npr2	Normalized PMI ratio
PM3	Dipole moment
Eang	Angle bend energy
IC_{50}	Half maximal inhibitory concentration
PIC_{50}	Negative log of the IC_{50}
RMSE	Root mean square
F-value	Fischers value

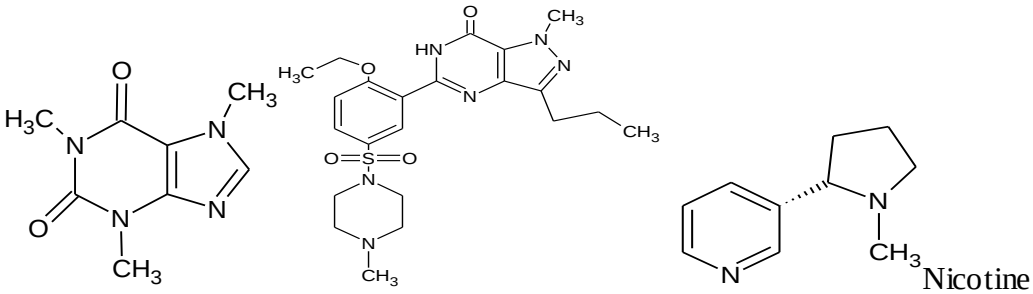
CHAPTER ONE

Introduction

1. Introduction:

1.1. Heterocyclics:

Heterocyclics are cyclic compounds in which one or more atoms of the ring are heteroatoms: O, N, S, P, etc. They are present in many biologically important molecules such as amino acids, nucleic acids and hormones. They are also indispensable components of pharmaceuticals and therapeutic drugs. Caffeine, sildenafil (the active Ingredient in Viagra), acyclovir (an antiviral agent), clopidogrel (an anti platelet agent) and nicotine, they all have heterocyclic systems (Fig1.1) (Rowlands, 2001).



Caffeine sidenafil

Figure 1.1: Examples of molecules containing heterocyclic systems

Some of heterocyclics are aromatic if they obey Hückel's rules of aromaticity:

a) The ring is planar.

b) There is a continuous conjugated system (all atoms of the ring are sp^2 hybridized).

c) The number of pi electrons is equal to 4 n +2, where n = 0, 1, 2, 3, etc.

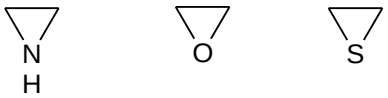
According to these rules, pyrrole, imidazole, pyridine, pyrimidine and purine, just to mention a few, are aromatic heterocyclic.

Classification of hetero cyclic:

-Three membered rings, four membered rings, five membered rings, six membered rings and poly cyclic.

Three-membered:

The three membered heterocycles containing single atom of nitrogen, oxygen, and sulfur-aziridine, oxirane, (or ethylene oxide) and thiirane, respectively-and their derivatives can all be prepared by nucleophilic reactions. Aziridine is formed by heating β-amino ethyl hydrogen sulfate with a base (Fig 1.2).

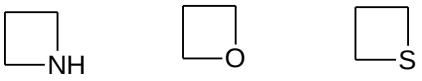


Aziridine oxirane thiirane

Figure 1.2: Three membered heterocycles

Four –membered:

Azetidine, oxetane and thietane – four membered rings containg respectively one nitrogen, oxygen, or sulfur atom are prepared by nucleophilic displacement reaction similar to those used to prepare corresponding three-membered rings.

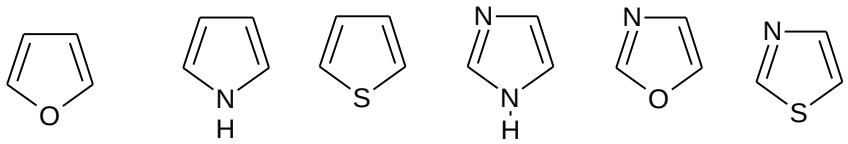


Azetidine oxetane thietane

Figure 1.3: Four membered heterocycles

Five-membered:

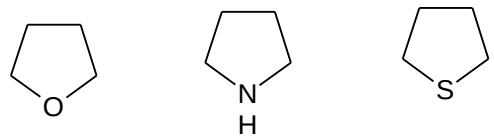
The parent aromatic compounds of this family pyrrole, furan and thiophene have the structures shown.



Furan Pyrrole Thiophene Imidazole Oxazole Thiazole

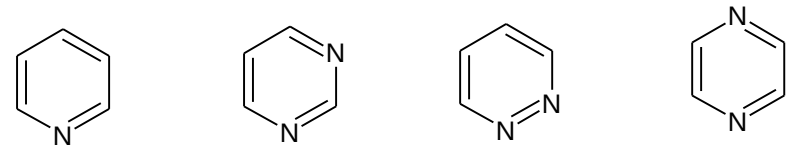
Figure 1.4: Five membered heterocycles

The saturated derivatives are called pyrrolidine, tetrahydrofuran, and thiolane, respectively.

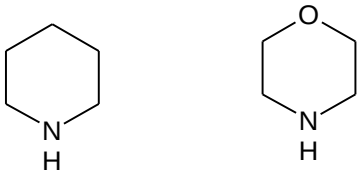


Tetrahydrofuran Pyrrolidine Thiolane

Six-membered:



Pyridine Pyrimidine Pyridazine Pyrazine

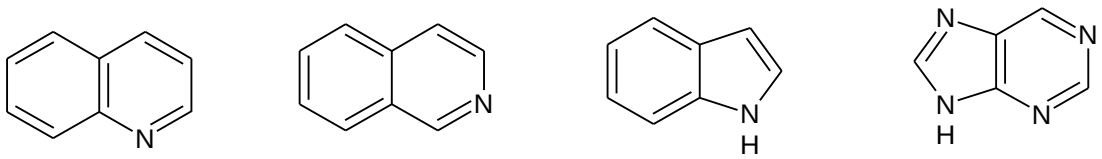


Piperidine Morpholine

Figure 1.5: Six membered heterocycles

Poly cyclic:

The bicyclic compounds made of pyrrole, furan, or thiophene ring fused to benzene ring are called indole (or isoindole), benzofuran, and benzothiophene, respectively.



Quinoline Isoquinoline Indole Purine

Figure 1.6: Poly cyclic compound

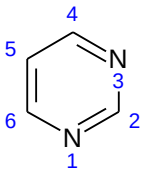
1.2. Pyrimidines:

1.2.1. Chemistry of pyrimidines

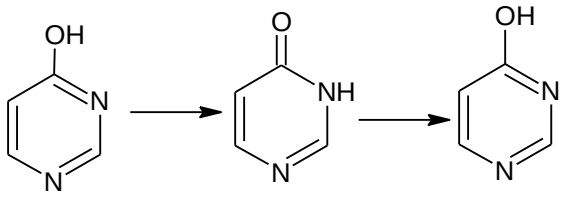
Pyrimidines are six-membered heterocyclic ring compounds composed of nitrogen and carbon. They are present throughout nature in various forms and are the building blocks of numerous natural compounds from antibiotics to vitamins and liposacharides.

The most commonly recognized pyrimidines are the bases of RNA and DNA, the most abundant being cytosine, thymine or uralic. The origin of the term pyrimidine dates back to 1884 when Pinner coined the term from a combination of the words pyridine and amidine because of the structural similarity to those compounds. Since these initial investigations hundreds of pyrimidine-containing compounds have been found in Biochemistry

Pyrimidine is the most important member of all the diazines as this ring system occurs widely in



living organism. Purines, uric acid, alkoxan, barbituric acid and mixture of anti- malarial and anti-bacterial also contain the pyrimidine ring. Since pyrimidine is symmetrical about the line passing C-2 and C-5, the position C-4 and C-6 are equivalent and so are N-1 and N-3 .When a hydroxy or amino group is present at the 2-,4-,6-, position than they are tautomeric with oxo and imino respectively as illustrated in fig(1.7).



Lactin form Lactam form

Figure 1.7: Lactin and lactam form

There are eight possible partially hydrogenated pyrimidines although it is not certain that all of them are stable to be isolated (Bansal, 2009).

1.2.2. Physical and spectroscopic properties:

Pyrimidine is colorless compound; it is weakly basic (pk_a1.3) as compared to pyridine (pk_a5.2) or imidazole (pk_a7.2).

The decrease in its basicity is due to the electron-withdrawing effect of the second nitrogen atom present in the ring. Moreover, the addition of the proton does not increase the probability for mesomerism and hence the resonance energy. Presence of alkyl groups, however, enhance the basicity, thus 4-methyl pyrimidine has pk_a2.0 while 4,6-dimethyl pyrimidine has a value of 2.8 the 2-and 4-aminopyrimidine are more basic with pk_a23.54 and 5.71 respectively. In these two compound more resonance structure are possible in the cation than in the neutral molecule as illustrated in fig (1.8).

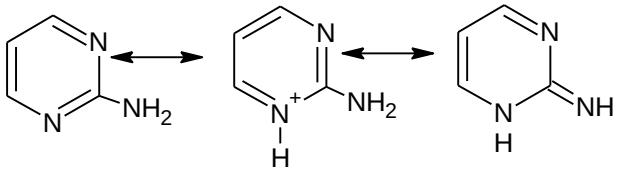


Figure 1.8: resonance structure of 2-aminopyrimidine

The close relationships of pyrimdine with benzene suggest the former is highly aromatic and the ring is virtually planner. The following canonical structures contribute to the resonance hybrid Fig (1.9). But pyrimidine ring is less aromatic compared to pyridine and benzene (Bansal, 2009).

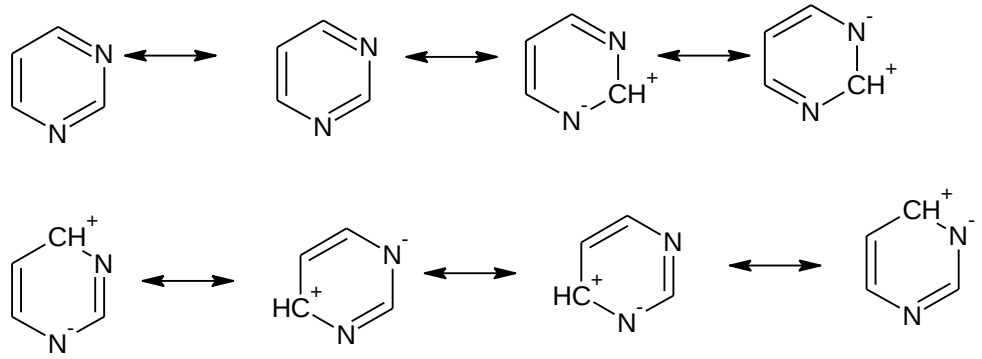


Fig 1.9: Canonical structures contributed to resonance hybride

This view is corroborated by the resonance energies which are benzene (36Kcal/mole), pyridine(31Kcal/mole) and pyrimdine (26Kcal/mole).

Pyrimdine have a dipole moment of 2.40D. The molecular dimensions of pyrimidine have not been determined (Bansal, 2009).

1.2.3. Synthetic method:

Pyrimidines have been prepared by a number of methods but the most important are those in which the ring is formed from two fragments which contribute the C-C-C and N-C-N atoms respectively.

1. **From malonic ester**: a simple synthesis of pyrimidine ring involves a condensation between a malonic ester and urea in the presence of base to yield barbituric acid fig (1.10). Modification of this method consists of using substituted malonic esters. Besides malonic esters, a series of other compounds such as β-keto acids or ester may be employed uracil, for instance, is obtained from a-formylacetic acid (produced in *situ* by decarboxylation of malic acid with conc. sulfuric and reaction of the β-keto acid with urea fig (1.11). Uracil can be converted to pyrimidine in the following steps:

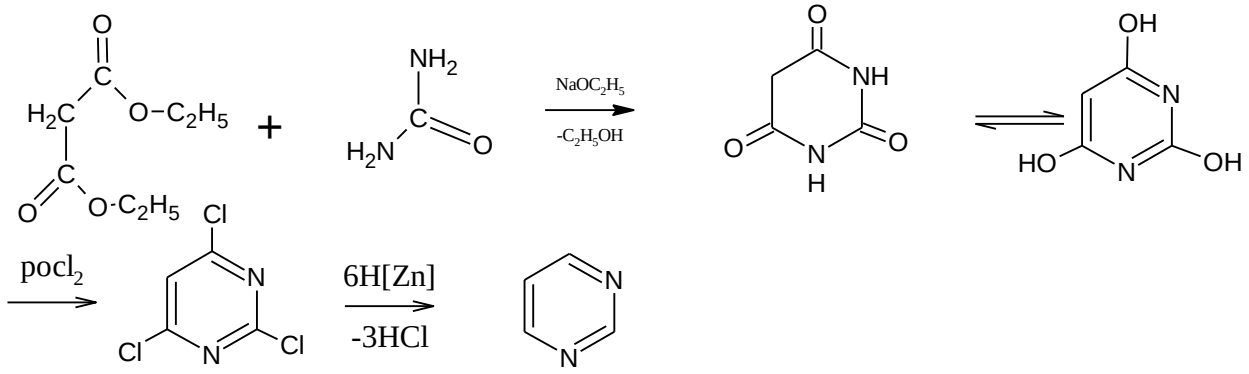
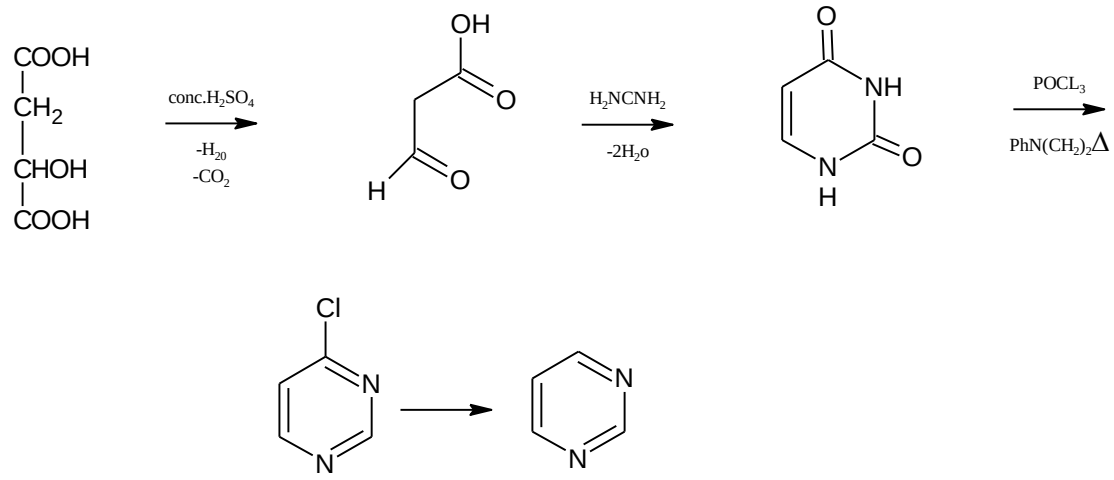


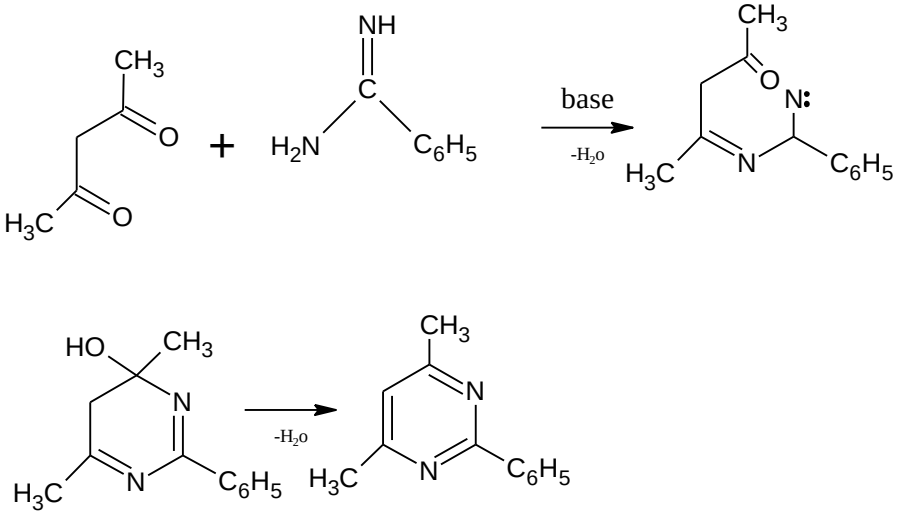
Figure 1.10: Synthesis of pyrimidine from malonic ester



78%

Figure 1.11: Synthesis of pyrimidine from uracil

Urea may be replaced by amidine, guanidine or thiourea and condensed with a 1,4-diketone. Thus acetyl acetone on reaction with benzamidine give 4,6-di methyl-2-phenyl pyrimidine.



A 1,4-diketone may be condensed with an aldehyde and ammonia to furnish pyrimidine derivatives (Rowlands,2001).

2. From ethyl cortonate: another useful method involves the condensation of amidines or urea with unsaturated compounds such as ethyl crotonate in the presence of a base fig (1.12).

Adihydropyrimidine is the initial product which is readily oxidized to the corresponding pyrimidine (Wu *et al.*, 2018).

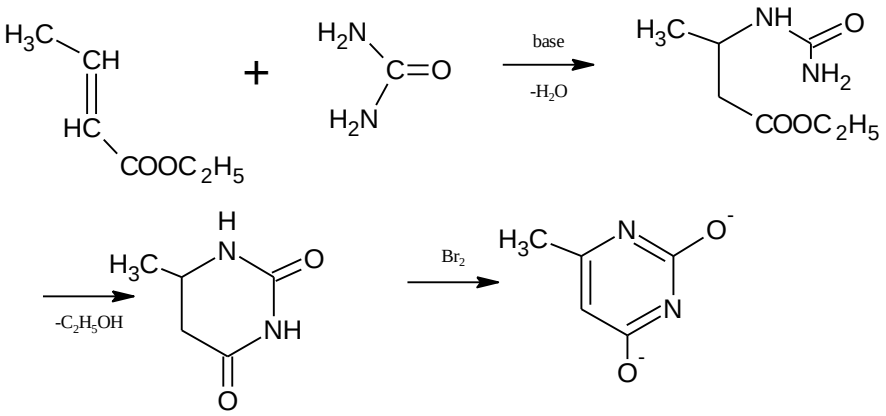


Figure 1.12: Synthesis of pyrimidine from ethyl cortonate

3-Pyrimidine itself can be obtained by the decarboxlation of pyrimidine-4, 6-dicarboxlic acid or by the dechlorination of 2,4-dihydropyrimidine.

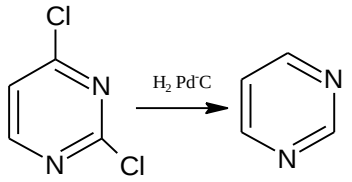


Figure 1.13: Synthesis of pyrimidine from dechlorination of 2,4-dihydropyrimidine

4- From α, β -Unsaturated ketones: An interesting reaction of simple α, β -unsaturated ketone with amidine to give pyrimidine. The initial product of this reaction is probably a dihydropyrimidine which is the readily oxidized by a stream of air to the corresponding pyrimidine. Benzamidine and β-benzoylstryene furnish 2, 4, 6-triphenyl pyrimidine (Khobragade *et al.*, 2010).

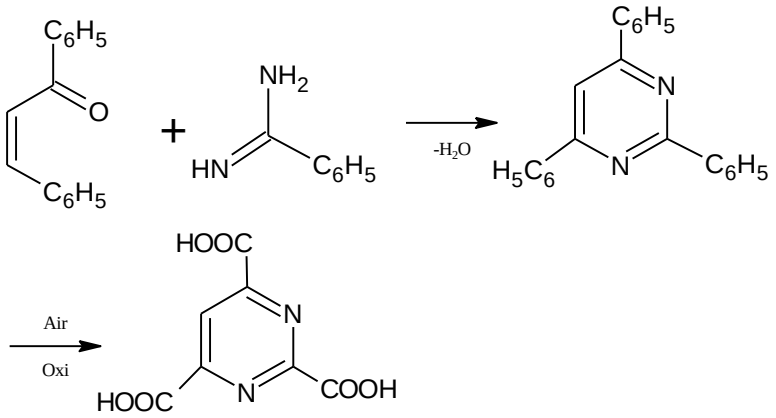
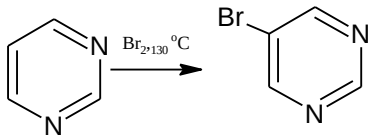


Figure 1.14: Synthesis of pyrimidine from α, β -Unsaturated ketones

1.2.4 Chemical reactions of pyrimidine:

1. **Reaction with acid:** pyrimidine though a weak base can be protonated in the presence of acids. Diprotonation unlike pyridine takes place in strong acids. Diprotonation is possible because the nitrogen atoms are not present in the adjacent position as in pyridazine.
2. **Electrophilic substitution:** pyrimidine is also resistant to electrophilic substitution. The attack at position 2, 4, and 6 is particularly retarded because of electron deficiency at the positions. The 5-position is also difficult to attack as it is influenced by the inductive effect of the two nitrogen atoms and this resembles position-3 in pyridine. Electrophilic substitution at position-5 is easy if one or more electron –releasing groups are present on the ring. The iodination of aminopyrimidine has been investigated. Bromination of pyrimidines yields 5-bromopyrimidine. The reaction is not simple electrophilic substitution rather proceed via a perbromide intermediate (Ingraham, 1981).



3. **Reaction with nucleophilic reagents:** The attack of a nucleophile takes place easily on the pyrimidine ring similar to pyridine, quinoline and isoquinoline. The positions susceptible to attack are 2, 4, and 6. Pyrimidine is stable in cold alkali but in boiling hydrazine it rearranges to pyrazole, via a ring opened intermediate (van *et al.*, 1974).

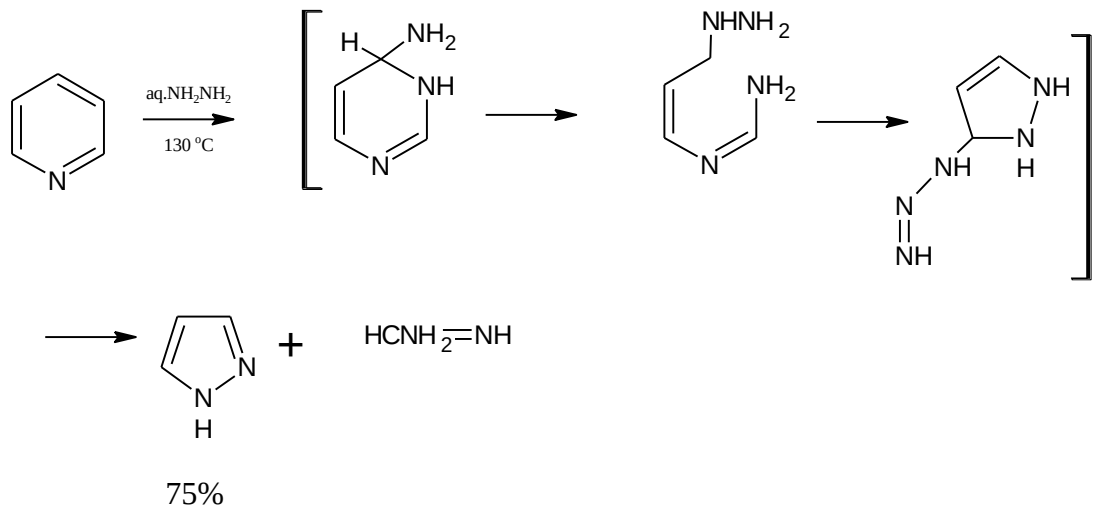


Figure 1.15: Reaction of pyrimidine with nucleophilic reagents

4. Reaction with Oxidizing and Reducing Agents: Pyrimidine gives a low yield of N-oxide on oxidation with a peracid. The ring is largely destroyed during N-oxide formation. But the alkylpyrimidines give a satisfactory yield of N-oxide. 4-Methyl pyrimidine with H_2O_2 yields 4-methyl pyrimidine N-oxide.

This can be converted back to pyrimidine by refluxing with phosphorous oxy chloride. Reissert addition reaction appears to be fairly common in pyrimidine N-oxides. 4-Methyoxy pyrimidine N-oxide on reaction with sodium cyanide and benzoyl chloride under alkaline conditions furnishes 2-cyano-4-methoxypyrimidine fig (1.16).

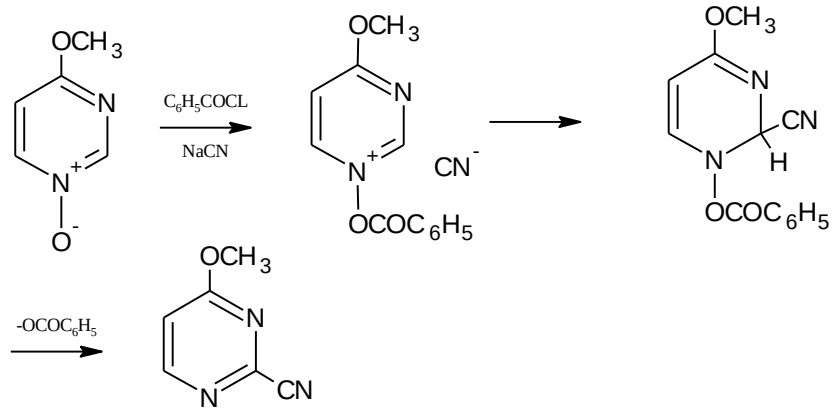


Figure 1.16: Reaction of pyrimidine with oxidizing and reducing agents

4. **Thio –claisen rearrangement:** Certain suitably substituted thiopyrimidines undergo the familiar thio-clasin rearrangement. Thus 3-methyl-4-allyl thiopyrimidine-2-one on heating yields 5-allyl-3-methyl-4-thiouracil fig (1.17) (Bansal, 2009).

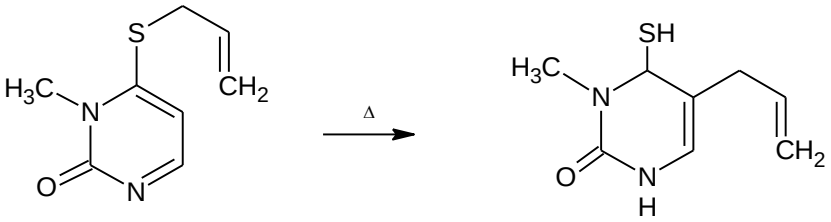
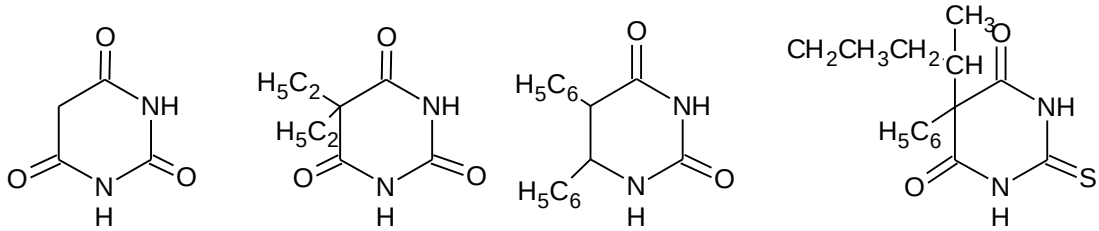


Figure 1.17: Thio-clasin rearrangement

1.2.5 Naturally occurring and biologically active pyrimidines:

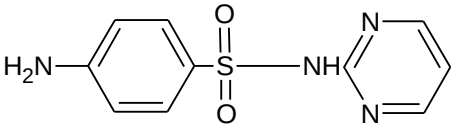
Pyrimidine itself is not found in nature but substituted pyrimidines and compound containing the pyrimidine ring are widely distributed in nature. Derivatives of barbituric acid, i.e., oxygenated pyrimidines are perhaps the most widely used in medicines, for example, Veronal, luminal are used as hypnotic while pentothal is used as an anaesthetic fig (1.18).



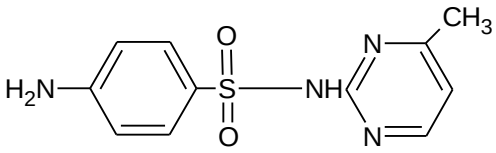
Barbituric acid Veronal Luminal Pentothal

Figure 1.18: Derivatives of barbituric acid

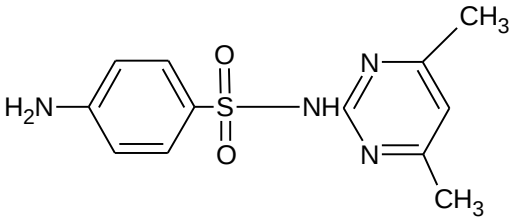
Several important sulfa drugs are pyrimidine derivatives namely sulfadiazine, sulfamerazine and sulfadimidine.



Sulfadiazine



Sulfamerazine

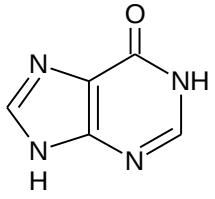


Sulfadimidine

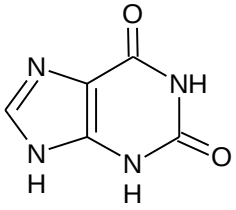
Eigure 1.19: Sulfa drugs of pyrimidine

Sulfadiazine is still widely used but the last two are no longer used for chemotherapy of infections. Three pyrimidines are of considerable biological importance because their relation to the nucleic acids, these are uracil, thymine and cytosine. The purine ring system obtained by the fusion of pyrimidine and imidazole unclei also is important because certain of its derivatives, in particular adenine and guanidine are building blocks of RNA and DNA.

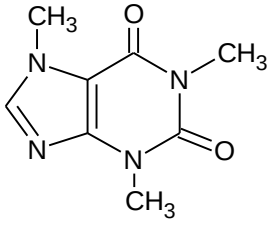
A Variety of natural products such as alkaloids also contain the pyrimidine ring system, these include hypoxanthine, and xanthine which occur in tea and caffeine and theophylline are the constituents of tea leaves. Theo bromine is found in cocoa beans fig(1.20).



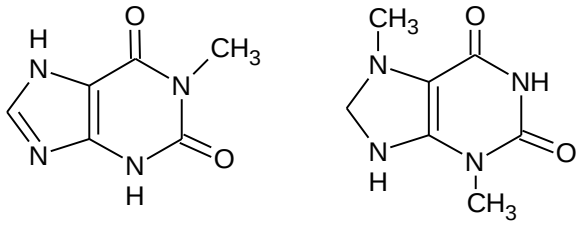
Hypoxanthine



xanthine



caffeine



Theophylline Theo bromine

Figure 1.20: Natural products contain the pyrimidine ring system

Folic acid: The pteridine ring system is also widely distributed in nature. The important growth factor folic acid, vitamin B_{10} is constructed of apteridine ring, p-amino benzoic acid and glutamic acid, i.e.pteroylglutamic acid. It is widely distributed and has been isolated from liver and yeast (Bansal, 2009).

1.3 Dihydropyrimidine derivatives:

1.3.1 Synthetic method:

Dihydropyrimidines consists of a six membered heterocyclic ring having two nitrogen atoms at one and three positions (Gazz *et al.*, 1893). The three-component reaction between aromatic aldehydes, urea and acetoacetic esters was discovered by the Italian chemist P. Biginelli in 1893. The products of this reaction is 3, 4-dihydropyrimidin-2(1*H*)-ones (DHPMs). For a long time the Biginelli reaction was not commonly used, but in the last 20–30 years this class of heterocyclic compounds received considerable attention as a type of privileged heterocyclic scaffolds with a significant pharmacological potential (Simurova and Maiboroda, 2017). It has been shown that some dihydropyrimidines exhibit high antiviral, anti-inflammatory, anticancer, and anti hypertensive activity (Ruijter, 2012) Therefore, the number of publications and patents devoted to the synthesis of DHPMs derivatives by means of the Biginelli reaction is growing every year. This can be explained by the simplicity of the synthetic procedure, the possibility to vary the starting reagents, catalysts, and solvents, as well as the possibility of introducing substituent that are easily converted into various functional groups. The rapid development of combinatorial chemistry also led to an increased interest in Biginelli reaction fig (1.21) (Simurova and Maiboroda, 2017).

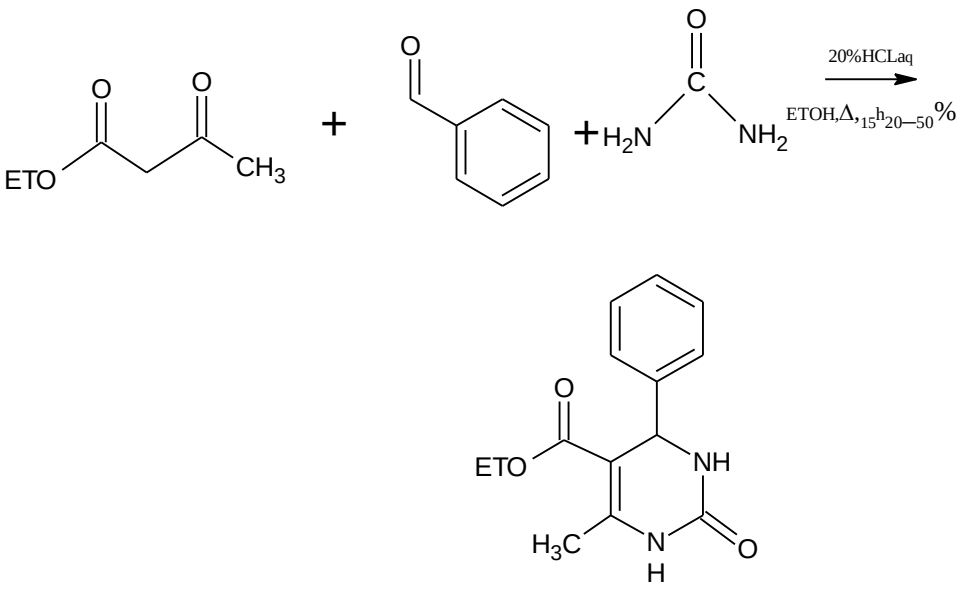


Figure 1.21: Biginelli reaction

1.3.2 Reaction conditions:

The classic method of carrying out the Biginelli reaction assumes one-pot condensation of ethyl acetoacetate, benzaldehyde and urea under strongly acidic conditions. The reaction proceeds with low yields and requires relatively long time (15–20 h).

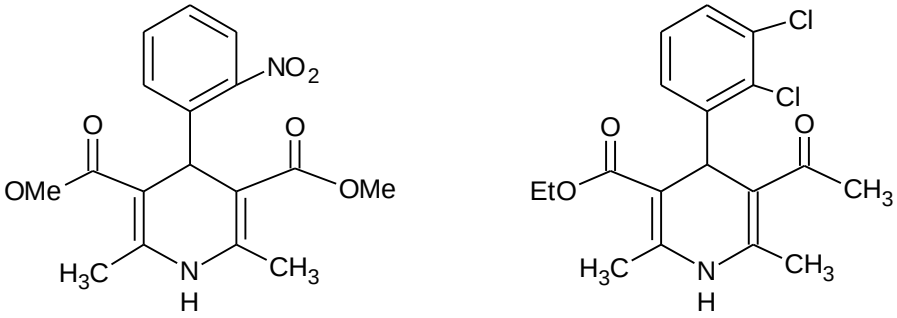
A significant number of works has been devoted to the optimization of reaction conditions in order to increase the yields of target DHPMs. The influence of solvents and catalysts on the yields of the target products obtained in Biginelli reaction has been studied (Alvim *et al.*, 2014). One approach is optimization of the solvents (acetic acid, aceto nitrile, THF, DMFA, etc.) and the selection of appropriate catalyst. (organic and inorganic acids, Lewis acids, ionic liquid in order to accelerate the reaction, experiments have been performed with microwave irradiation, infrared irradiation, as well as ultrasonication, thereby reducing the reaction time to few minutes and increasing yields up to 98% (Simurova and Maiboroda, 2017).

1.3.3 Biological activity of dihydropyrimidine derivatives:

Similar groups/structures often exhibit similar biological activities. However, they usually exhibit different potency. The traditional structure activity relationship (**SAR**) is a useful tool in the search for new drugs. However, SAR is usually determined by making minor changes to the

structure of the existing compound and assessing the effect on its biological activity (Kuntz *et al.*, 1982).

Biological aspect of Biginelli product i.e. dihydropyrimidines, intensive investigation is Carried out because they possess close resemblance of clinically used nicardipine, nifedipine etc fig (1.22). feiodipine and nifedipine which are analogues of Biginelli product further they had resemblance to marine natural alkaloids batzelladine B **10**. Again biologists & chemists synthesized modified Biginelli product scaffolds, which showed activities like ant proliferative, antiviral, antitumor, anti-inflammatory, antibacterial, antifungal and antitubercular activity.



Nifedipine feiodipine

Figure 1.22: Biological aspect of biginelli product

Antihypertensive agents, dihydropyrimidine derivatives as anti-hypertensive drug activity were showed in (Alam *et al.*, 2010), **anti filarial agents** were showed in (Singh *et al.*, 2008), **anti-HIV agents** were showed in (Matos *et al.*, 2008), **antitumor activity** were showed in (Kaan and Ulaganathan, 2010), **antioxidant Activity** were showed in (Stefani *et al.*, 2006; Ismaili *et al.*, 2008) and **anti cancer activity were** showed in (Eckhardt *et al.*, 2002; Lee et al., 2002; Lee, Y.S *et al.*, 2009; Rojo *et al.*, 2008; Anjos *et al.*, 2012; Gressler *et al.*, 2010; Jain *et al.*,2006; Patrawala *et al.*, 2005).

1.4 Modeling and informatics in drug design:

Modeling and informatics have become indispensable components of rational drug design. For the last few years, chemical analysis through molecular modeling has been very prominent in computer - aided drug design (CADD). But currently modeling and informatics are contributing in tandem toward CADD.

Modeling in drug design has two faces: modeling on the basis of knowledge of the drugs/leads/ Ligands often referred to as ligand - based design and modeling based on the structure of macromolecules often referred to as receptor - based modeling (or structure – based modeling). Computer - aided drug design is a topic of medicinal chemistry that can be employed to understand the properties of chemical species and the properties of bimolecular on the other. Information technology is playing a major role in decision making in pharmaceutical sciences. Storage, retrieval, and analysis of data of chemicals/biochemical of therapeutic interest are major components of pharmaco informatics. Quite often, the efforts based on modeling and informatics gets thoroughly integrated with each other, as in the case of virtual screening exercises. The molecular modeling methods that are in vogue in the fields of computational chemistry, computational biology, computational medicinal chemistry, and Pharmacoinformatics (Prasad *et al.*, 1998).

1.4.1 Computational chemistry:

Two - dimensional (2D) structure drawing and three - dimensional (3D) structure buildings are the important primary steps in computational chemistry for which several molecular visualization packages are available. The most popular of these are ChemDraw Ultra and Chem3D pro, which are a part of the ChemOffice suite of software packages, ACD/ChemSketch, Mol Suite, and many more of these kinds are other programs for the same purpose. Refinement has to be carried out on all the drawings and 3D structures so as to improve the chemical accuracy of the structure on the computer screen. Structure refinement based on heuristic Rules/cleanup procedures is a part of all these software packages. However, chemical Accuracy of the 3D structures still remains poor even after cleanup. Further refinement can be carried out by performing energy minimization using either molecular mechanical or quantum chemical procedures. By using these methods, the energy of a molecule can be estimated in any given state (Goossen, 2009).

Following this, with the help of first and second derivatives of energy, it can be ascertained whether the given computational state of the molecules belongs to a chemically acceptable state or not.

During this process, the molecular geometry gets modified to a more appropriate, chemically meaningful state – the entire procedure is known as geometry optimization.

The geometry optimized 3D structure is suitable for property estimation, descriptor calculation, conformational analysis, and finally for drug design exercise (Prasad *et al.*, 1998).

1.4.2 Energy Minimization and geometry optimization:

Drug molecules prefer to adopt equilibrium geometry in nature, that is, a geometry that possesses a stable 3D arrangement of atoms in the molecule. The 3D structure of a molecule built using a 3D builder does not represent a natural state; slight modifications are required to be made on the build 3D structure so that it represents the natural state. For this purpose, the following questions need to be addressed: (1) which minimal changes need to be made? (2) How much change needs to be made? (3) How does one know the representation at hand is the true representation of the Natural state? To provide answers to these questions depend on energy, because molecules prefer to exist in thermodynamically stable states. This implies that if the energy of any molecule can be minimized, the molecule is not in a stable state and thus the current representation of the molecule may not be the true representation of the natural state. This also implies that minimize the energy and the molecular structure in that energy minimum state probably represents a true natural state. Several methods of energy minimization have been developed by computational chemists, some of which are non derivative methods (simple method) but many of which are dependent on derivative methods (steepest descent, Newton – Raphson, conjugate gradient, variable metrics, etc.) and involve the estimation of the gradient of the potential energy curve. The entire procedure of geometry modification to reach an energy minimum state with almost null gradient is known as geometry optimization in terms of the structure of the molecule and energy minimization in terms of the energy of the molecule. All computational chemistry software packages are equipped with energy minimization methods of which a few incorporate energy minimization based on *a initio* methods while most include the semi empirical and molecular mechanics based energy minimization methods (Dolinšek *et al.*, 1998).

1.4.3 Computational biology:

Computational biology is a fast growing topic and it is really not practical to distinguish this topic from bioinformatics. Structure prediction of biomolecules (often referred to as "Structural bioinformatics") adopts many aspects of computational chemistry. For example, energy minimization of protein receptor structure is one important step in computational biology.

Molecular mechanics, molecular simulations, and molecular dynamics are employed in performing conformational analysis of macromolecules.

A rational drug design approach is very much dependent on the knowledge of receptor protein structures and is severely limited by the availability of target protein structure with experimentally determined 3D coordinates. Proteins exhibit four tiered organization: (1) primary structure defining the amino acid sequence, (2) secondary structure with helical and sheet folds, (3) tertiary structure defining the folding of secondary structure held by hydrogen bonds, and (4) quaternary structure involving non covalent association between two or more independent proteins (Holtje, 2003).

Methods for identifying the primary amino acid sequence in proteins are now well developed; however, this knowledge is not sufficient enough to understand the function of the proteins, the drug – receptor mutual recognition, and designing drugs.

Various experimental techniques like X - ray crystallography, nuclear magnetic resonance, and electron diffraction are available for determining the 3D coordinates of the protein structure; however, there are many limitations. It is not easy to crystallize proteins and even when we succeed, the crystal structure represents only a rigid state of the protein rather than a dynamic state. Thus, the reliability of the experimental data is not very high in biomolecules. Computational methods provide the alternative approach — although with equal uncertainty but at a greater speed.

The 3D structures of proteins are useful in performing molecular docking, *de novo* design, and receptor - based pharmacophore mappings (Weissig and Bourne, 2003; Bharatam *et al*., 2010).

1.4.4 Computational medicinal chemistry:

Representation of drug molecular structures can be handled using computational chemistry methods, whereas that of macromolecules can be handled using computational biology methods. However, finding the therapeutic potential of the chemical species and understanding the drug – receptor interactions *in silico* requires developed techniques of computational medicinal chemistry.

1.4.5 Molecular design software:

Molecular design software is software for molecular modeling, distinctive property of which is the presence of the special support for the developing the molecular model(Paul.J.Hall).In contrast to the usual molecular modeling programs such as the molecular dynamics and quantum chemistry programs such software directly supports the aspects related to the construction of molecular models:

- Molecular graphics.
- Interactive molecular drawing and conformational editing. `
- Building of polymeric molecules, crystals and solvated systems.
- Partial charges development.
- Geometry optimization.
- Support for the different aspects of force field development etc.

Molecule Editor:

A molecule editor is a computer program for creating and modifying representation of chemical structures. There are number of type of molecule editor. Molecular drawing program are used to generate two-dimensional (flat) representation of molecules and chemical reactions that can be used as illustration or for querying chemical databases .Three-dimensional molecule editor are used to build molecular models, usually as part of molecular modeling software packages. database molecular editors such as leather face, RECAP and molecule slicer allow large number of molecules to be modified automatically according to rules such as deprotonate carboxylic acids or break exocyclic bonds that can be specified by the user.

Most molecule editors use proprietary file formats, but most can read and write several file formats, including smiles, a short ASCII- representation of molecules. Files generated by two – or three –dimensional molecule editors can also be display by molecular graphics tools, which can be realized as a small applet to show molecules in web pages (Paul.J.Hall).

1.5 Quantitative structure – activity relationship (QSAR):

QSAR is a statistical approach that attempts to relate physical and chemical properties of molecules to their biological activities. This can be achieved by using easily calculatable

descriptors like molecular weight, number of rotatable bonds, and log *P*. Developments in physical organic chemistry over the years and contributions of Hammett and Taft in correlating the chemical activity to structure laid the basis for the development of the QSAR paradigm by Hansch and Fujita. Table (1.1) gives an overview of various QSAR approaches in practice. The 2D and 3D QSAR approaches are commonly used methods, but novel ideas are being implemented in terms of 4D – 6D QSAR. The increased dimensionality does not add any additional accuracy to the QSAR approach; for example, no claim is valid which states that the correlation developed using 3D descriptors is better than that based on descriptors.

TABLE 1.1 : Different Dimensions in QSAR:

1D QSAR: Affinity correlates with p *K* a, log *P*, etc.
2D QSAR: Affinity correlates with a structural pattern.
3D QSAR: Affinity correlates with the three - dimensional structure.
4D QSAR: Affinity correlates with multiple representations of ligand.
5D QSAR: Affinity correlates with multiple representations of induced - fit scenarios.
6D QSAR: Affinity correlates with multiple representations of solvation models.

1.5.1 2 D QSAR*:* Initially, 2D QSAR or the Hansch approach was in vogue, in which different kinds of descriptors from the 2D structural representations of molecules were correlated to biological activity. The basic concept behind 2D QSAR is that structural changes that affect biological properties are electronic, steric, and hydrophobic in nature. These properties can be described in terms of Hammett substituent and reaction constants, Verloop sterimol parameters, and hydrophobic constants. These types of descriptors are simple to calculate and allow for a relatively fast analysis (Dunbrac and Karplus, 1994).

Most 2D QSAR methods are based on graph theoretical indices. The graph theoretical descriptors, also called the molecular topological descriptors, are derived from the topology of a molecule, that is, the 2D molecular structure represented as graphs.

The electro topological state index (E - state) combines the information related to both the topological environment and the electronic character of each skeletal atom in a molecule. The constitutional descriptors are dependent on the constitution of a molecule and are numerical

descriptors, which include the number of hydrogen bond donors and acceptors, rotatable bonds, chiral centers, and molecular weight (1D). Apart from that, several indicator descriptors, which define whether or not a particular indicator is associated with a given molecule, are also found to be important in QSAR. The quantum chemical descriptors include the molecular orbital energies (HOMO, LUMO), charges, super delocalizabilities, atom – atom and molecular polarizabilities, dipole moments, total and binding energies, and heat of formation. These are 3D descriptors derived from the 3D structure of the molecule and are electronic in nature (Park and Levitt, 1995; Sternberg, 1996).

Statistical data analysis methods for QSAR development are used to identify the correlation between molecular descriptors and biological activity. This correlation may be linear or nonlinear and accordingly the methods may be divided into linear and nonlinear approaches. The linear approaches include simple linear regression, multiple linear regression (MLR), partial least squares (PLS), and genetic algorithm – partial least squares (GA - PLS). Simple linear regression develops a single descriptor linear equation to define the biological activity of the molecule.

MLR is a step ahead as it defines a multiple term linear equation. More than one term is correlated to the biological activity in a single equation. PLS, on the other hand, is a multivariate linear regression method that uses principal components instead of descriptors. Principal components are the variables found by principal component analysis (PCA), which summarize the information in the original descriptors (Randic, 1975; Hall *et al.*, 2003).

The aim of PLS is to find the direct correlation not between the descriptors and the biological activity but between the principal component and the activity. GA - PLS integrates genetic algorithms with the PLS approach (Hosoya, 1971).

The QSAR model developed by any statistical method has to be validated to confirm that it represents the true structure – activity relationship and is not a chance correlation. This may be done by various methods such as the leave one out and leave multiple out cross - validations and the bootstrap method. The randomization test is another validation approach used to confirm the adequacy of the training set.

Attaching chemical connotation to the developed statistical model is an important aspect. A successful QSAR model not only effectively predicts the activity of new species belonging to the same series but also should provide chemical clues for future improvement. This requirement, as

well as the recognition that the 3D representation of the chemicals gives more detailed information, led to the development of 3D QSAR (Bonchev, 1983) (Balaban, 1982).

1.5.2 3D QSAR*:* 3D QSAR methods are an extension of the traditional 2D QSAR approach, wherein the physicochemical descriptors are estimated from the 3D structures of the chemicals. Typically, properties like molecular volume, molecular shape, HOMO and LUMO energies, and ionization potential are the properties that can be calculated from the knowledge of the 3D coordinates of each and every atom of the molecules. When these descriptors of series of molecules can be correlated to the observed biological activity, 3D QSAR models can be developed. This approach is different from the traditional QSAR only in terms of the descriptor definition and, in a sense, is not really 3D in nature (Cramer, *et al*., 1988).

Molecular fields (electrostatic and steric), which can be estimated using probe - based sampling of 3D structure of molecules within a molecular lattice, can be correlated with the reported numeric values of biological activity Such methods proved to be much more informative as they provide differences in the fields as contour maps. The widely used CoMFA (comparative molecular field analysis) method is based on molecular field analysis and represents real 3D QSAR methods. Asimilar approach was adopted in developing modules like CoMSIA (comparative molecular similarity index analysis), SOMFA (self - organizing molecular field Analysis), and COMMA (comparative molecular moment analysis). Utilization and predictivity of CoMFA itself has improved sufficiently in accordance with the objectives to be achieved by it. Despite the formal differences between the various methodologies, any QSAR method must include some identifiers of chemical structures, reliably measured biological activities, and molecular descriptors. In 3D QSAR, alignment (3D superimposition) of the molecules is necessary to construct good models. The main problems encountered in 3D QSAR are related to Improper alignment of molecules, greater flexibility of the molecules, uncertainties about the bioactive conformation, and more than one binding mode of ligands (Klebe *et al*.,1994) (Robinson *et al*., 1999).

Alignment of 3D structures of molecules is carried out using RMS atoms alignment, moment's alignment, or field alignment. The relationship between the biological activity and the structural parameters can be obtained by multiple linear regressions or partial least squares analysis (Dunbrac and Karplus, 1994).

1.5.3 4D QSAR*:* 4D QSAR analysis developed by Vedani and colleagues incorporates the conformational alignment and pharmacophore degrees of freedom in the development of 3D QSAR models. It is used to create and screen against 3D - pharmacophore QSAR models and can be used in receptor - independent or receptor - dependent modes.

4D QSAR can be used as a CoMFA preprocessor to provide conformations and alignments; or in combination with CoMFA to combine the field descriptors of CoMFA with the grid cell occupancy descriptors (GCODs) of 4D QSAR to build a " best " model; or in addition to CoMFA because it treats multiple alignments, conformations, and embedded pharmacophores, which are limitations of CoMFA (Vedani *et al.*,2000).

1.5.4 5D QSAR*:* The 4D QSAR concept has been extended by an additional degree of freedom the fifth dimension allowing for multiple representations of the topology of the quasi - atomistic receptor surrogate. While this entity may be generated using up to six different induced - fit protocols, it has been demonstrated that the simulated evolution converges to a single model and that 5D QSAR, due to the fact that model selection may vary throughout the entire simulation, yields less biased results than 4D QSAR, where only a single induced - fit model can be evaluated at a time (software Quasar) (Vedani *et al.*, 2005).

1.5.5 *6D* QSAR*:* A recent extension of the Quasar concept to sixth dimension (6D QSAR) allows for the simultaneous consideration of different solvation models .This can be achieved explicitly by mapping parts of the surface area with solvent properties (position and size are optimized by the genetic algorithms) or implicitly (Vedani and Dobler, 2002). In Quasar, the binding energy is calculated as

$$E\text{binding} = E\text{ligand-receptor} - E\text{desolvation, ligand} - T\,\Delta S - E\text{internal strain} - E\text{induced fit}$$

1.6 Molecular Docking:

There are several possible conformations in which a ligand may bind to an active site, called the binding modes. Molecular docking involves a computational process of searching for a confirmation of the ligand that is able to fit both geometrically and energetically into the binding site of a protein. Docking calculations are required to predict the binding mode of new hypothetical compounds. The docking procedure consists of three interrelated components identification of the binding site, a search algorithm to effectively sample the search space (the set of possible ligand positions and conformations on the protein surface), and a scoring function. In most docking algorithms, the binding site must be predefined, so that the search space is

limited to a comparatively small region of the protein. The search algorithm effectively samples the search space of the ligand – protein complex. The scoring function used by the docking algorithm gives a ranking to the set of final solutions generated by the search.

The stable structures of a small molecule correspond to minima on the multidimensional energy surface, and different energy calculations are needed to identify the best candidate. Different forces that are involved in binding are electrostatic, electrodynamic, and steric forces and solvent related forces. The free energy of a particular conformation is equal to the solvated free energy at the minimum with a small entropy correction. All energy calculations are based on the assumption that the small molecule adopts a binding mode of lowest free energy within the binding site. The free energy of binding is the change in free energy that occurs upon binding and is given as

$\Delta G\text{binding} = G\text{complex} - (G\text{protein} + G\text{ligand})$

Where *G* complex is the energy of the complexed protein and ligand, *G* protein is the free energy of non interacting separated protein, and *G* ligand is the free energy of non interacting separated ligand (Vedani *et al.*, 2005).

The common search algorithms used for the conformational search, which provide a balance between the computational expense and the conformational search, include molecular dynamics, Monte Carlo methods, genetic algorithms, fragment - based methods, point complementary methods, distance geometry methods, tabu searches, and systematic searches .

1.6.1 DOCK: DOCK is a simple minimization program that generates many possible orientations of a ligand within a user selective region of the receptor. DOCK is a program for locating feasible binding orientations, given the structures of a "ligand" molecule and a "receptor" molecule. DOCK generates many orientations of one ligand and saves the best scoring orientation. The docking process is handled in four stages — ligand preparation, site characterization, scoring grid calculation, and finally docking. Site characterization is carried out by constructing site points, to map out the negative image of the active site, which are then used to construct orientations of the ligand. Scoring grid calculations are necessary to identify ligand orientations. The best scoring poses may be viewed using a molecular graphics program and the underlying chemistry may be analyzed (Leach and Leach, 2001).

There are many other widely used molecular docking software packages, like Flexi dock (based on genetic algorithm), Auto dock (based on Monte Carlo simulations and annealing), MCDOCK

(Monte Carlo simulations), FlexE (ensemble of protein structures to account for protein flexibility), and DREAM (to dock combinatorial libraries) (Kuntz *et al*., 1982).

1.7 Aim and objective of current study:

Dihydropyrimidine and its fused heterocyclic derivatives tested as potential anticancer agents, and important class of therapeutic agents. Various drugs containing pyrimidine nucleus were synthesized and used as anticancer agents like 5-fluorouracil (5-FU), tegafur and thioguanine .
The main objective of this study is to design and develop new dihydropyrimidine derivatives by using computational method.

The specific objectives of this study are to:

- Generate QSAR model with acceptable statistical parameters that can be used to predict the anti cancer activity of newly designed dihydropyrimidine derivatives by using dataset compounds.
- Design new dihydropyrimidine derivatives using the computer software to predict their biological activity by generated model.
- Select some of the training set compound which were posses higher biological activity than reference compound *N*-[6-(4-chlorophenoxy)hexyl]-1-(1-diazyn-1-ylidene)-*N*'-(pyridin-4-yl)methanediamine.
- Predict interaction of selected training set compound with receptor macrophage migration (MIF) enzyme protein through molecular docking study.
- Select some of the newly designed derivatives which were posses higher predicted biological activity than reference compound CHS 828.
- Predict interaction of new designed dihydropyrimidine derivatives with receptor (MIF) enzyme protein through molecular docking study.

CHAPTER TWO

Material & Methods

2 Material and methods:

2.1 materials and software:

2.1.1 Data set

The dihydropyrimidine derivatives (Table 2.1) collected from (Abdo, 2015) were used as data set. The biological activities (IC_{50}) of those compounds were converted to their negative logarthimic scale (pIC_{50} = -log IC_{50}). The data set comprised of 21 compounds which covered a wide range of IC_{50} value. Then the data set was reduced to 16 compounds related to their similar structure. All compounds were grouped randomly to training set and test set.

2.1.2 Software:

2.1.2.1 ACD labs software

(Version 12.01 run for windows) Specializes in software for small molecule chemistry. ACD labs provide enterprise solutions for analytical data handling and knowledge mangment, molecular property modeling and property-based design. ACD/labs were founded in 1994 as a private company. The intention was to monetize the experience of an international team of scientists specializing in quantitative structure-property relationships, such as NMR spectra and various physico-chemicals properties such as PKa. LogP, logD, boiling point, vapor pressure. Early on the development focused on predictors, chemical drawing and chemical naming.

2.1.2.2 Chem. Draw software

Chem Draw (verison 12.0, run for widows) is a molecule editor first developed in 1985 by David A. Evans and Stewart Rubenstein (later by the chemoinformatics company Cambridge soft). chemDraw, along with chem3D and chemfinder , is part of chemoffice suite programs and is available for Microsoft windows.

Feature of chemDraw chemical structure to name conversion, chemical name to structure conversion. NMR spectrum simulation, mass spectrum simulation, structure cleanup, an extensive collection of templates, including style templates for most major chemical journals. chemDraw can export to MOL, SDF, and SKC chemical file formats.

2.1.2.3 MOE Software

Molecular operating environment (Version 110, run for windows) is a drug discovery software platform that integrates visualization, modeling and simulations, as well as methodology development, in one package. MOE scientific applications are used by biologists, medicinal chemistry and computational chemists in pharmaceutical, biotechnology and academic research.

Main applications area in MOE include structure-based design, fragment–based design, pharmacophore discovery, medicinal chemistry applications , biologics applications, protein and antibody modeling, molecular modeling and simulations ,chemoinformatics &QSAR.

2.1.2. 4 SPSS software

SPSS (Version 16.020 run for windows) is a Windows based program that can be used to perform data entry and analysis and to create tables and graphs. SPSS is a capable of handling large amounts of data and can perform all of the analyses. SPSS is commercially distributed for data mangment and statistical analysis it helped revolutionize research practices in the social sciences, enabling research to conduct complex statistical analysis on their own. (Felix Frey University of Leipzig).

1.3 Methods

2.2.1 Preparation for QSAR modeling study

The biological activity of 16 dihydropyrimidine derivatives (dataset) in terms of IC_{50} **(**nM) converted to calculated pIC_{50} and divided in to training set and test set. The test set was selected by random selection and composed of 3 compounds and training set of 13 compounds. All compounds were drawn using ACD/lab and saved in mol.format (Table2.1).

2.2.1.1 Molecular modeling descriptors

The mol. files were opened by MOE 2009.10 software used minimized energy, the 220 different molecular descriptors in two and three dimensions were calculated .The descriptors in two dimensions were removed because the squared correlation coefficient r^2 (<0.7) showed in appendix A. The descriptors in three dimensions were decreased to 25 by

correlation matrix, the correlation of descriptor with each other and with pIC_{50} of the molecules were examined (Fig A.1).

7 descriptors were left in clouding as listed in table (2.2):

ASA- Negative accessible surface area.

dens. Mass density.

E.ang. Angle bend energy.

MINDO-LUMO. LUMO energy (eV).

npr^2 .Normalized PMI ratio (2) (pmi2/pmi3).

PM3.dipole .dipole moment.

Vsurf.IW1.hyrophoilic integy moment at -0.2.

2.2.1.2 Model development

The QSAR models were developed based on the partial least square method (PLS) using the descriptors in MOE as an independent variable and pIC_{50} as dependent variables by forward regression analysis. The quality of each regression model was evaluated using a squared correlation coefficient r^2 (>0.7) and root mean square error (RMSE) (Dearden *et al.*, 2009).

About 14 QSAR models were generated by using partial least square regression method (Table2.3).

2.2.1.3 Validation Model

Internal validation of the developed model was performed by using leave-one-out (LOO) cross-validation coefficient (*q*2) (Golbraikh and Tropsha, 2002), in which each compound in the training set was eliminated in the calculation of linear regression analysis. The value of q2 > 0.5 indicated the predictive ability of the developed model (Golbraikh *et al.* 2003; Tropsha *et al.*, 2003). In addition, Tropsha *et al.* (2003) explained that the internal cross validation should be accomplished by external validation using test set, which was represented by the value of

external cross validation coefficient (R^2pred). A model was considered to be valid if it possessed R^2pred value higher than 0.7.

The statistical reliability of the QSAR model was evaluated based on several criteria, i.e. squared correlation coefficient (R^2), Fischer's value for statistical significance (F), and standard error of estimation (SEE) (Dearden *et al*., 2009).

2.2.2 Modeling of new compound of dihydropyrimidine derivatives

From compounds **2a** which is most active compound (IC_{50} =49),**7c** and **9d** the set of new 94 dihydropyrimidine derivatives were designed using the computer software ACD/lab table(2.5) bond length and angle standardized by clean structure and then saved by (mol) format. These compounds were sketched to predict their activity against colon cancer as macrophage migration inhibitory enzyme (MIF).

2.2.2.1 Predict the biological activity of new designed dihydropyrimidine derivatives

The mol. files of 94 designed derivatives were opened by MOE 2009.10 software, energy minimized and the different 5 descriptors were calculated. The fit of model-1 was evaluated to predict the biological activity of new dihydropyrimidine derivatives in the term of pIC_{50} (Table2.5).

2.2.3 Molecular docking studies

Molecular docking studies of dihydropyrimidine derivatives with the protein receptor was carried out by using MOE 2009.10 software with distributes simulation from molecular operating environment to studies for the targets site prediction of human macrophage migration inhibitory (MIF) activity.

2.2.3.1 Preparation of ligands and protein

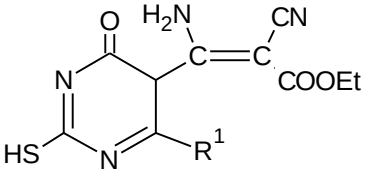
The mol. files of 71 ligands selected from 94 new dihydropyrimidine derivatives designed according to their higher predicted biological activity than CHS828(reference compound) were opened in MOE 2009.10. The 3-D protonated structures of ligands were energy minimized .Then ligands were saved in a molecular database (mdb) file for docking.

The crystal structure of macrophage migration inhibitory (MIF) was download from a protein data bank. The protein structure was prepared using Molecular operating environment (MOE) software make 3-D protonation and removed of unwanted molecules except for unique ligand. The prepared protein was saved in moe format.

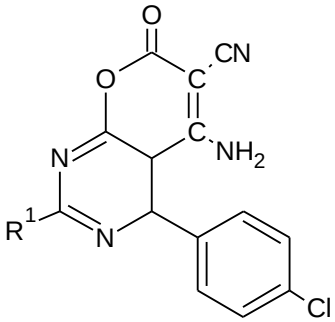
2.2.3.2 Analysis of docking

The ligand-protein interactions were presented in 2-dimensionals space by making use the MOE ligand interaction program. The binding energy (s), length of bonds between ligand and amino acids in receptor were listed in (Table 2.6) (Table 2.7).

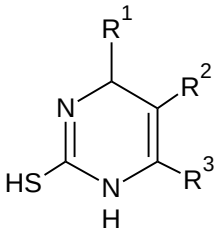
Table 2.1: Structures, experimental IC_{50} , experimental PIC_{50},predict PIC_{50} and residual values of 4,5-dihydro-2-mercapto-4-oxo-6-substituted Arylpyrimidine derivatives **2a**–**d** and their fused rings **4b**, **5b**, and also 1,4-dihydro-2-mercaptopyrimidine derivatives**7a**–**e**, **9a**–**e** used in training and test set for inhibit colon cancer.



(2a, 2b, 2c, 2d)

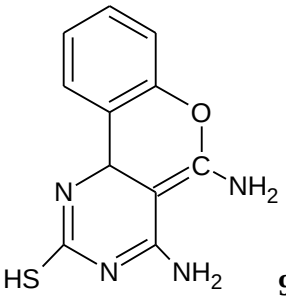


(4b, 5b)

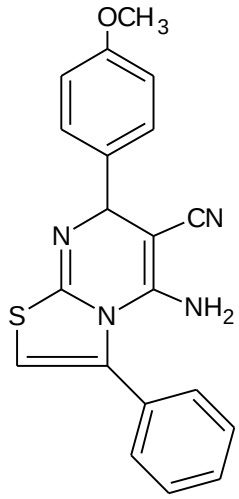


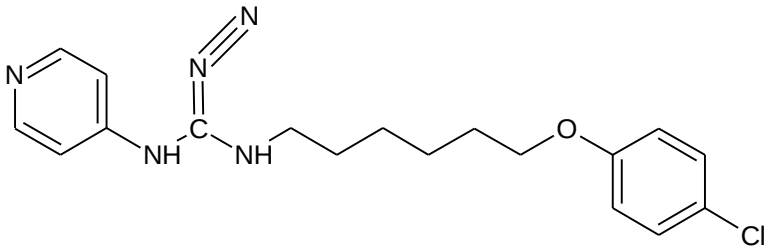
(7a, 7b, 7c, 7d, 7e, 9a, 9b, 9d)

NO	Compound	R^1	R^2	R^3	(IC_{50} inM) exp	PIC_{50} exp	Pred PIC_{50}	Residual
1	2a	C_6H_5	-	-	49	7.30	7.15	0.15
2	2b	4-CLC_6H_4	-	-	3129	5.50	5.39	0.11
3	2c	4-$OCH_3C_6H_5$	-	-	188	6.72	7.01	0.29
4	2d	2-furyl	-	-	126	6.89	6.85	0.04
5	4b	H_2NHN	-	-	608	6.21	6.37	-0.16
6	5b	PhHNHN	-	-	146	6.83	6.39	0.44
7	7a	C_6H_5	COOEt	CH_3	4840	5.31	5.48	-0.17
8	7b	4-CLC_6H_4	COOEt	CH_3	1640	5.78	5.67	0.11
9	7c[1]	4-$OCH_3C_6H_4$	COOEt	CH_3	2482	5.60	5.48	0.12
10	7d	2-furyl	COOEt	CH_3	1240	5.90	5.58	0.32
11	7e	C_6H_5	Fused ring	-	2294	5.63	5.63	0.00
12	9a	C_6H_5	CN	NH_2	2140	5.66	5.76	-0.1
13	9b[1]	4-CLC_6H_4	CN	NH_2	2146	5.67	6.08	-0.41
14	9d	2-furyl	CN	NH_2	3224	5.49	5.64	-0.15
15	9e	-	-	-	2690	5.57	5.80	-0.23
16	10[1]	-	-	-	1770	5.75	6.50	-0.75
17	CHS828	-			2315	5.63	6.44	-0.81



9e





10 **CHS 828(**reference compound**)**

Table 2.2: Value of chemical descriptors used in QSAR modeling of dihydropyrimidine derivatives data set:

No	Compound	npr^2	Eang	MINDO-LUMO	PM3-dipole	Vsurf.IW1	ASA-	dens
1	2a	0.9229	22.533	-1.6200	5.1996	8.7509	248.9072	1.1112
2	2b	0.6970	17.3671	-	3.3334	7.0914	301.9202	1.1703
3	2c	0.8907	28.7953	-1.5657	5.6784	5.7542	231.1059	1.1255
4	2d	0.7976	16.2212	-1.7493	7.4160	8.7906	259.5882	1.1661
5	4b	0.6801	31.3718	-1.5539	5.6561	8.2277	259.5959	1.1929
6	5b	0.7151	32.9552	-1.4349	5.9157	9.0718	334.4916	1.1356
7	7a	0.7174	10.0213	-0.0466	1.7812	0.0000	174.8729	1.0326
8	7b	0.7118	9.4842	-0.3500	2.7417	7.1748	224.2732	1.1080
9	7d	0.7346	10.2966	-0.1119	2.6334	7.2973	175.6205	1.0848
10	$7c^T$	0.7111	16.515	-0.0662	2.0261	6.1279	157.1701	1.0588
11	7e	0.6603	18.7678	-0.5279	3.3772	7.1870	193.5976	1.1389
12	9a	0.7278	13.02966	-0.4201	2.8840	7.2900	213.3961	1.0761
13	$9b^T$	0.8175	13.0635	-0.4935	3.5657	6.9216	259.9131	1.0599
14	9d	0.6824	13.2967	-0.4473	3.1892	7.3590	215.9130	1.1368
15	9e	0.6936	18.8869	-0.6357	4.5976	6.4904	175.9130	1.1482
16	10^T	0.9423	31.4653	-0.5532	4.2551	5.6941	267.4363	1.0593

T: Test set

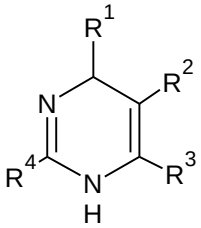
Table 2.3: Models with descriptors used in predict the biological activity of dihydropyrimidine derivatives dataset

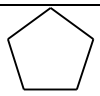
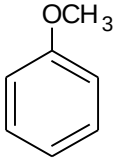
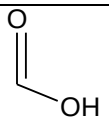
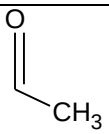
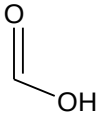
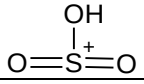
No	eq	R^2	RMSE
1	PIC_{50}=5.213+ 0.373 npr^2 -0.674MNDO-LUMO	0.889	0.209
2	PIC_{50}=3.257+0.507 npr^2 + 0.584 PM3-dipole	0.863	0.232
3	PIC_{50}=2.918+0.531 npr^2 + 0.522AM1dipole	0.807	0.276
4	PIC_{50}=8.705-0.142 Eang -1.000 MNDO-LUMO	0.806	0.266
5	PIC_{50}=1.346+0.719 npr^2 + 0.450 Vsurf.IW1	0.791	0.288
6	PIC_{50}=1.745+0.660 npr^2 + 0.460 $FCasA^-$	0.790	0.288
7	PIC_{50}=2.248+0.660 npr^2 + 0.449 Eang	0.780	0.295
8	PIC_{50}=0.697+0.714 npr^2 + 0.432ASA^-	0.774	0.294
9	PIC_{50}=1.804+0.805 npr^2 + 0.426 Estrain	0.770	0.301
10	PIC_{50}=-0.317+0.754 npr^2 + 0.403 rgyr	0.752	0.313
11	PIC_{50}=7.573+0.543 Eang + 0.614 glob	0.743	0.319
12	PIC_{50}=3.426+0.606 npr^2+ 0.410 E	0.732	0.326
13	PIC_{50}=-8.173+0.819 npr^2 + 0.360 dens	0.723	0.331
14	PIC_{50}=0.836+0.651 npr^2 + 0.359 ASA	0.706	0.341

Table 2.4: The statistical parameters for five models have greater R^2.

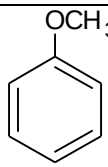
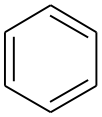
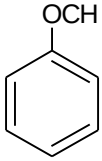
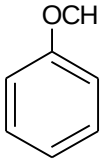
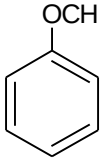
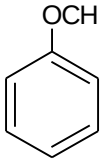
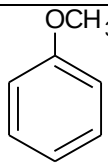
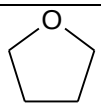
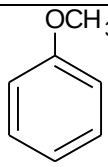
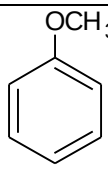
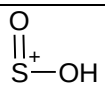
	No. of training	No. of test	R^2	Q^2	R^2_{pred}	RMSA	F-value	P-value	SEE
Equation (1)	13	3	0.889	0.796	0.979	0.209	34.535	0.0001	0.24
Equation (2)	13	3	0.863	0.742	0.992	0.232	31.629	0.0001	0.26
Equation (3)	13	3	0.807	0.644	0.886	0.276	20.981	0.0001	0.31
Equation (4)	13	3	0.806	0.672	0.886	0.266	18.749	0.001	0.32
Equation (5)	13	3	0.791	0.532	0.988	0.288	18.953	0.0001	0.33

Table 2.5: The value of chemical descriptors and predicted PIC_{50} values of new dihydropyrimidine derivatives designed

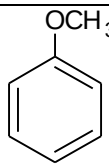
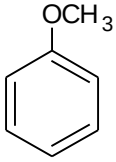
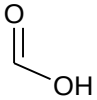
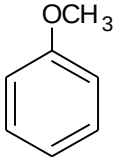
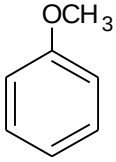
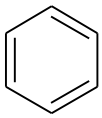
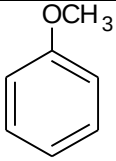
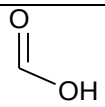
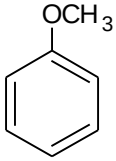
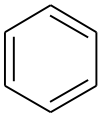
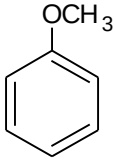
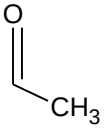
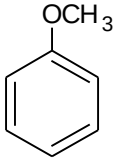
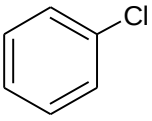


NO	Cpd	R^1	R^2	R^3	R^4	E-ang	MIDO LUMO	npr^2	Pm3-dipole	Vsurf –Iw1	Pred PIC_{50}
1	I		CN	NH_2	HS	18.9726	-0.3833	0.718	3.260	0.000	5.713
2	II		CN		HS	17.252	-0.7923	0.811	3.130	6.036	6.265
3	III		CN	NH_2	HS	13.3230	-0.4645	0.666	2.918	0.000	5.609
4	IV		CN	NH_2	HS	13.885	-0.5822	0.677	3.428	0.000	5.721
5	V		CN	NH_2	HS	13.382	-0.776	0.621	4.679	0.000	5.681
6	VI		CN	NH_2	HS	13.981	-2.367	0.656	5.865	0.000	6.839

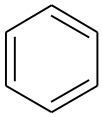
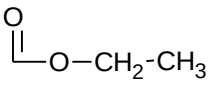
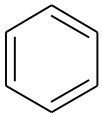
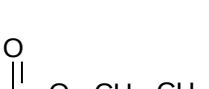
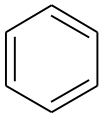
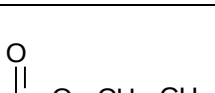
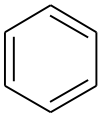
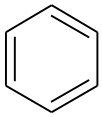
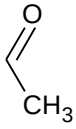
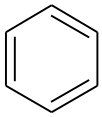
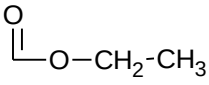
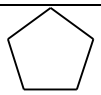
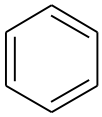
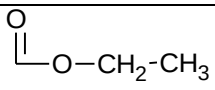
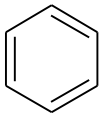
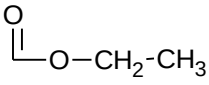
Continued:

7	VII		CN		HS	13.648	-0.35	0.828	3.578	5.929	6.030
8	VIII		CN	-	HS	12.996	-0.29	0.802	3.476	5.779	5.910
9	IX		CN	Cl	HS	13.646	-0.69	0.814	3.080	5.893	6.214
10	X		CN		HS	17.983	-0.38	0.779	3.553	5.878	5.900
11	XI		CN	NH_2	OH	19.380	-0.34	0.861	4.451	5.625	6.122
12	XII			NH_2	HS	23.388	0.189	0.676	4.027	5.720	5.209
13	XIII		OH	NH_2	HS	18.868	0.132	0.861	4.222	5.695	5.805
14	XIV			NH_2	HS	20.634	-0.32	0.740	1.768	5.570	5.745

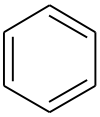
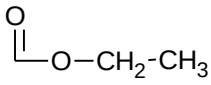
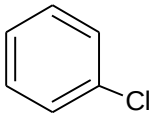
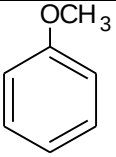
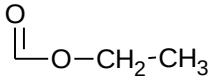
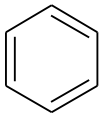
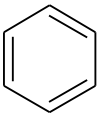
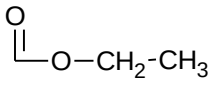
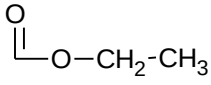
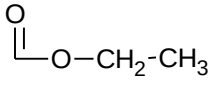
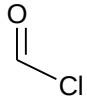
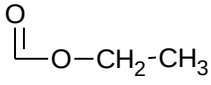
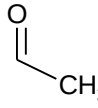
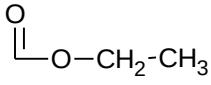
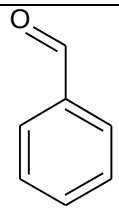
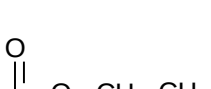
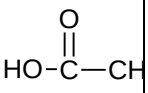
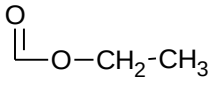
Continued:

15	XV		CN	NH_2	CH_3	19.786	-0.23	0.858	5.464	5.654	6.036
16	XVI		CN	NH_2		23.143	-0.55	0.793	3.393	6.230	6.054
17	XVII		CN	NH_2	-	19.068	-0.25	0.918	5.110	5.710	6.237
18	XVIII		CN	NH_2		21.512	-0.27	0.767	5.439	6.656	5.793
19	XIX			NH_2	HS	21.745	-0.12	0.785	2.005	5.568	5.744
20	XX			NH_2	HS	19.098	0.098	0.632	2.043	5.965	5.137
21	XXI			NH_2	HS	20.535	0.066	0.789	2.185	5.419	5.631
22	XXII			NH_2	HS	19.263	-0.36	0.636	2.026	5.949	5.451

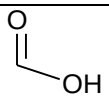
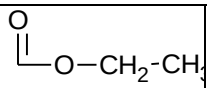
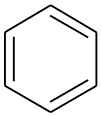
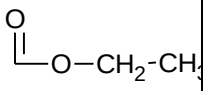
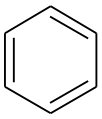
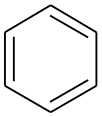
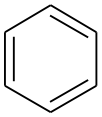
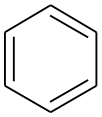
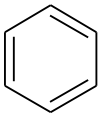
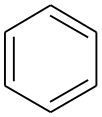
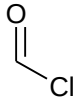
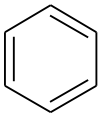
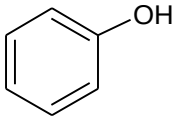
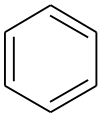
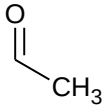
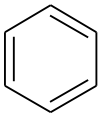
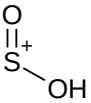
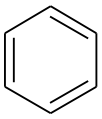
Continued:

23	XXIII			CH_3	HS	10.078	-0.127	0.717	1.017	0.00	5.54
24	XXIV			Cl	HS	9.701	-0.355	0.701	1.710	7.94	5.64
25	XXV				HS	10.483	-0.287	0.670	3.920	7.13	5.50
26	XXVI				HS	11.761	-0.339	0.653	0.932	8.08	5.48
27	XXVII				HS	14.537	0.017	0.689	2.223	7.11	5.36
28	XXVIII			CH_3	HS	14.641	-0.147	0.760	4.825	7.36	5.68
29	XXIX			CH_3	NH_2	13.141	-0.017	0.686	2.076	8.72	5.37

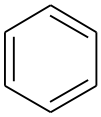
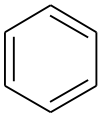
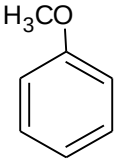
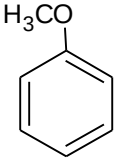
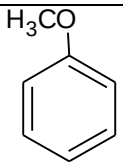
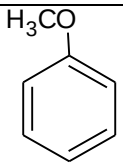
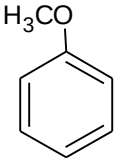
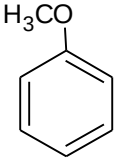
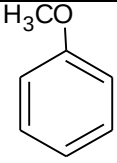
Continued:

30	XXX			CH_3		12.348	-0.58	0.874	3.449	7.74	6.319
31	XXXI			CH_3		18.047	-0.15	0.720	1.770	7.63	5.568
32	XXXII			CH_3	CH_3	10.327	0.086	0.675	2.861	0.00	5.272
33	XXXIII	OH		CH_3	HS	10.668	-0.12	0.862	3.406	7.30	5.97
34	XXXIV			CH_3	HS	16.334	-0.08	0.739	2.149	0.00	5.579
35	XXXV			CH_3	HS	10.038	-0.61	0.707	3.845	7.36	5.835
36	XXXVI			CH_3	HS	9.9718	-0.17	0.792	3.003	7.79	5.802
37	XXXVII			CH_3	HS	12.668	-0.16	0.618	3.124	7.07	5.268
38	XXXVIII			CH_3	HS	11.375	-0.40	0.698	2.730	7.21	5.67

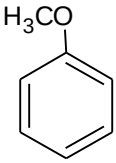
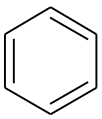
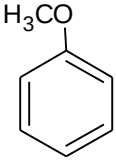
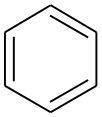
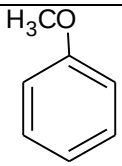
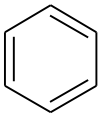
Continued:

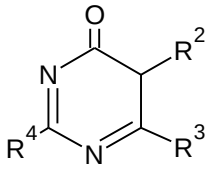
39	XXXIX			CH_3	HS	10.76	-0.53	0.769	4.076	7.50	5.96
40	XL			OH	HS	9.596	-0.168	0.711	0.881	0.000	5.551
41	XLI		OH	CH_3	HS	7.716	0.191	0.782	4.025	6.665	5.528
42	XLII		$H_3C—CH_3$	CH_3	HS	7.862	0.178	0.737	3.844	7.042	5.399
43	XLIII		$H_2C=CH_2$		HS	8.248	-1.166	0.728	6.072	7.271	6.262
44	XLIV		CH_3	CH_3	HS	7.362	0.204	0.792	3.468	6.449	5.548
45	XLV			CH_3	HS	9.915	-0.577	0.687	3.645	7.306	5.748
46	XLVI			CH_3	HS	8.232	0.124	0.727	2.055	6.754	5.407
47	XLVII			CH_3	HS	8.549	0.014	0.715	4.365	6.719	5.443
48	XLVIII			CH_3	HS	9.224	-0.438	0.633	2.496	7.253	5.493
49	XLIX		CN	CH_3	HS	7.299	-0.400	0.729	3.588	6.923	5.760

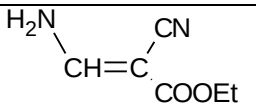
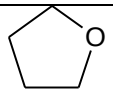
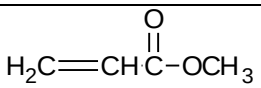
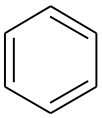
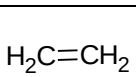
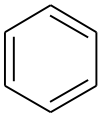
Continued:

50	L		-	CH_3	HS	6.452	0.205	0.836	3.315	6.489	5.68
51	LI			CH_3	HS	13.248	0.253	0.629	2.013	7.749	5.02
52	LII		CH_3	NH_2	-	19.195	0.223	0.964	2.785	5.729	6.05
53	LIII		SH	-	-	12.820	0.204	0.895	0.735	6.458	5.86
54	LIV		-	NH_2	-	17.979	0.171	0.983	3.00	5.801	6.14
55	LV		-	NH_2		18.701	-0.087	0.866	2.783	5.706	5.96
56	LVI		OH	NH_2		19.406	-0.123	0.832	2.300	5.779	5.88
57	LVII		OH	-		13.510	-0.101	0.867	2.324	5.498	5.97
58	LVIII		OH	-	-	12.716	0.297	0.931	2.390	6.092	5.90

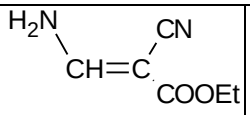
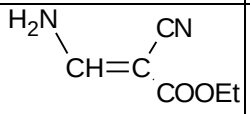
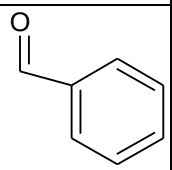
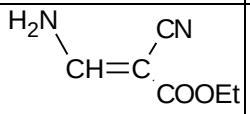
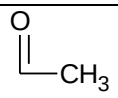
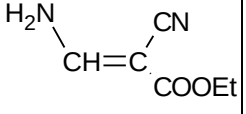
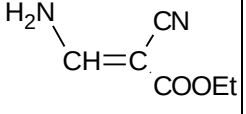
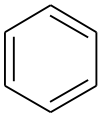
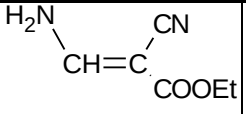
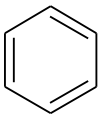
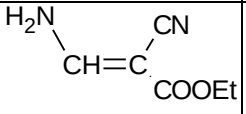
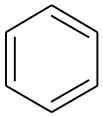
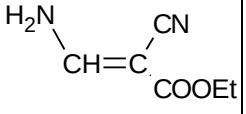
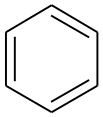
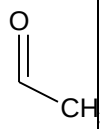
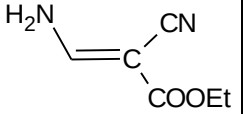
Continued:

59	LIX		OH		-	13.464	0.135	0.939	0.625	6.294	6.04
60	LX		-		-	12.901	-0.085	0.888	3.142	6.389	6.03
61	LXI		SH		-	13.638	-0.134	0.942	0.857	6.299	6.22

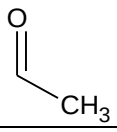
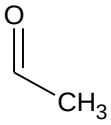
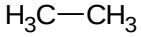
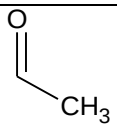
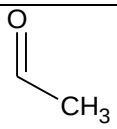
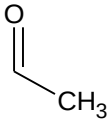
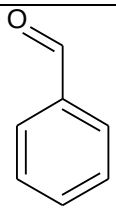
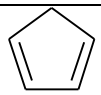
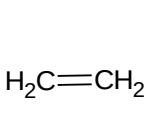
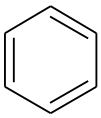


62	LXII	-			HS	20.567	-1.68	0.751	3.593	7.680	6.673
63	LXIII	-			HS	16.061	-1.28	0.680	4.394	7.722	6.197
64	LXIV	-			HS	826.07	-1.35	0.844	6.020	0.00	6.73
65	LXV	-			HS	8.248	-1.16	0.728	6.072	7.271	6.26

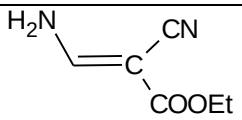
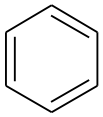
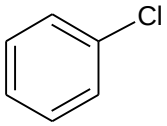
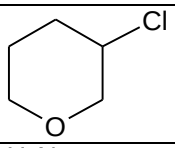
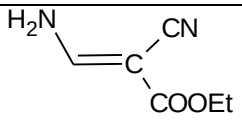
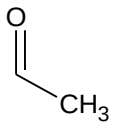
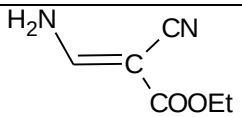
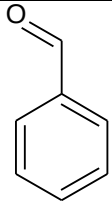
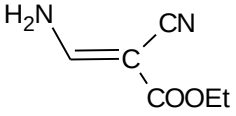
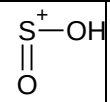
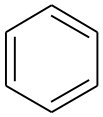
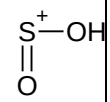
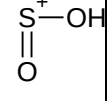
Continued:

66	LXVI	-			HS	16.338	-2.08	0.829	2.44	7.464	7.17
67	LXVII	-			HS	22.925	-1.76	0.813	6.722	7.493	6.91
68	LXVIII	-			HS	18.057	-1.75	0.701	3.797	7.829	6.57
69	LXIX	-		OH	HS	19.671	-1.53	0.786	4.121	8.629	6.67
70	LXX	-			Cl	22.107	-1.84	0.934	4.275	8.385	7.33
71	LXXI	-			H_3C—	17.761	-1.39	0.766	4.579	7.456	6.53
72	LXXII	-			HO—C	18.987	-	0.727	3.969	7.85	6.36
73	LXXIII	-				22.800	-1.712	0.945	3.904	9.35	7.27
74	LXXIV	CO		-	-	14.821	-1.536	0.771	4.562	2.741	6.63

Continued:

75	LXXV	-		-	OH	5.377	-0.520	0.881	3.460	2.409	6.29
76	LXXVI	-			OH	6.238	-0.528	0.713	3.727	2.515	5.79
77	LXXVII	-		Cl	OH	6.030	-0.984	0.743	3.818	2.607	6.18
78	LXXVIII	-		Cl	Cl	5.397	-1.207	0.811	4.082	2.212	6.53
79	LXXIX	-		-	-	4.509	-0.420	0.905	3.528	2.461	6.30
80	LXXX	-		-	-	6.301	-0.420	0.883	4.051	2.977	6.25
81	LXXXI	-		-	-	19.771	-0.302	0.857	2.329	2.70	6.08
82	LXXXII	-		-	CO	5.252	-0.198	0.872	5.457	2.13	6.05
83	LXXXIII	-			CO	8.038	-0.262	0.740	5.428	2.83	5.70

Continued:

84	LXXXIV	-		-	-	13.00	-0.849	0.707	1.974	2.90	5.99
85	LXXXV		-	-	-	4.112	-0.286	0.918	2.490	1.22	6.25
86	LXXXVI	-		-	-	4.985	-0.555	0.981	1.904	2.80	6.62
87	LXXXVII	-		-	-	7.525	-0.248	0.954	3.403	3.03	6.33
88	LXXXVIII	-			-	13.64	-1.109	0.778	2.293	2.84	6.37
89	LXXXIX	-			-	17.19	-1.033	0.644	3.644	2.74	5.91
90	XC	-		-		14.13	-1.243	0.859	1.602	2.86	6.70
91	XCI	-		-		5.587	-0.810	0.976	4.910	2.120	6.77
92	XCII	-	-	-		4.495	-0.873	0.819	4.767	2.218	6.34

Continued:

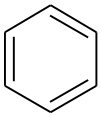
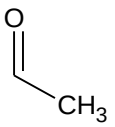
93	XCIII	-	-	-		5.235	-0.484	0.847	2.126	2.079	6.17
94	XCIV	-		-	-	4.509	-0.420	0.905	3.528	2.461	6.30

Table 2.6: Binding energy, type of bond interaction, amino acid interaction and bond length for docking some compounds of training set with protein (MIF) pocket as inhibitor of macrophage migration.

No	**Compounds**	**Free binding energy S**	**Type of bond interaction**	**Amino acid interaction**	**Bond Length (A^0)**
1	H I	-10.594	-	-	-
2	H II	-11.367	H-bond π –bond π –bond	Lys32 Phe113 Pro1	3.80 - -
3	H III	-13.390	-	-	-
4	H IV	-10.621	-	-	-
5	H V	-21.624	π -bond	Pro1	-
6	H VI	-12.098	π -bond	Pro 1	-

Table 2.7: Binding energy, type of bond interaction, amino acid interaction and bond length for docking new designed compounds with protein (4K9G) pocket as inhibitor of macrophage migration

No	Compounds	Free binding energy S	Type of bond interaction	Amino acid interaction	Bond Length (A^0)
1	I	-7.519	H-bond acceptor H-bond donor	Lys32 IIe64	2.63 1.84
2	II	-13.987	H-bond acceptor	Lys32	2.67 -
3	IV	-12.072	H-bond acceptor	Lys32	2.78
4	V	-9.502	H-bond donor	IIe64	2.15
5	VI	-11.337	H-bond acceptor H-bond acceptor	Lys32 Lys32	2.95 3.00
6	VII	-10.044	H-bond acceptor H-bond donor π –cation bond	Lys32 IIe64 Lys32	2.63 1.96 -
7	VIII	-14.238	H-bond acceptor π –cation bond	Lys32 Pro1	2.45 -
8	IX	-13.868	H-bond acceptor π-cation bond	Lys32 Pro1	2.43 -
9	X	-10.618	π -π bond π – cation bond	Phe112 Pro1	- -
10	XI	-10.475	No interaction	-	-
11	XIII	-13.989	π –cation bond	Pro1	-
12	XIV	-14.030	H-bond acceptor π –cation bond	Lys32 Pro1	2.40 -
13	XV	-14.059	H-bond donor π –cation bond π – π bond	Pro1 Lys32 Tyr36	2.37 - -
14	XVI	-19.403	H-bond acceptor H-bond acceptor π –cation bond	Lys32 Lys32 Lys32	2.14 2.20 -
15	XVII	-13.016	H-bond acceptor π –cation bond	Lys32 Pro1	2.49 -
16	XVIII	-10.368	H-bond acceptor H-bond donor π – π bond	Lys32 IIe64 Lys32	2.40 1.89 -
17	XIX	-14.394	H-bond acceptor H-bond donor π –cation bond π – π bond	Lys32 IIe64 Lys32 Tyr36	1.83 2.01 - -
18	LXII	-2.99	H-bond donor H-bond acceptor	Pro1 IIe64	2.30 2.00
19	LXIII	-14.637	No interaction	-	-
20	LXIV	-14.359	H-bond donor	Pro1	2.41
21	LXV	-13.358	No interaction	-	-

Continued:

22	LXVI	-8.300	No interaction	-	-
23	LXVII	-10.267	No interaction	-	-
24	LXVIII	-8.293	No interaction	-	-
25	LXXV	-12.813	No interaction	-	-
26	LXIX	-12.774	H-bond acceptor	Lys32	2.67
27	LXX	-9.949	H –bond donor	IIe64	2.05
28	LXXI	-11.175	H –bond donor	IIe64	2.21
29	LXXII	-12.398	No interaction	-	-
30	LXXIII	-11.609	H –bond acceptor	Pro1	1.94
31	XXVIII	-10.020	No interaction	-	-
32	XXX	-11.968	π –cation bond	Lys32	-
33	XXXIII	-16.870	H –bond donor H –bond acceptor	Pro1 Tyr36	1.96 3.08
34	XXXV	-14.860	No interaction	-	-
35	XXXVI	-11.603	No interaction	-	-
36	XXXVIII	-16.765	No interaction	-	-
37	XXXIX	-14.468	H –bond acceptor	Lys32	1.92
38	XLIII	-13.353	No interaction	-	-
39	XLV	-4.612	No interaction	-	-
40	XLIX	-6.972	H –bond acceptor H –bond acceptor	Lys32 IIe64	2.52 2.16
41	L	-14.489	π –cation bond	Pro1	-
42	LII	-13.516	H –bond donor	Pro1	2.35
43	LIII	-13.381	H –bond donor	Pro1	2.17
44	LIV	-14.358	H –bond acceptor π –cation bond π – π bond	Lys32 Pro1 Phe113	2.15 - -
45	LV	-18.017	π –cation	Pro1	-
46	LVI	-17.988	H –bond acceptor	IIe64	1.92
47	LVII	-17.623	H –bond acceptor H –bond acceptor π –cation bond π –cation bond	Lys32 Lys32 Lys32 Pro1	1.93 2.51 - -
48	LVIII	-15.265	H –bond π –bond	Lys32 Pro1	2.15 -
49	LIX	-15.644	H –bond acceptor π –cation bond π –cation bond	Lys32 Lys32 Pro1	1.88 - -
50	LX	-18.547	H –bond π –bond	Lys32 Phe113	2.08 -
51	LXI	-7.828	π –cation bond	Pro1	-
52	LXXIV	-10.039	No interaction	-	-

Continued:

53	LXXVI	-8.947	No interaction	-	
54	LXXVII	-12.953	H –bond donor	Pro1	2.31
55	LXXVIII	-11.107	H –bond acceptor	Lys32	1.92
56	LXXIX	-11.166	No interaction	-	-
57	LXXX	-15.648	H –bond acceptor π –cation bond	Lys32 Pro1	1.97 -
58	LXXXI	-12.128	No interaction	-	-
59	LXXXII	-11.981	No interaction	-	-
60	LXXXIII	-7.054	No interaction	-	-
61	LXXXIV	-10.484	H –bond donor	Pro1	2.36
62	LXXXV	-16.366	H –bond acceptor π –cation bond	Lys32 Lys32	2.20 -
63	LXXXVI	-15.066	π –cation bond	Pro1	-
64	LXXXVII	-12.844	H –bond donor	Pro1	2.82
65	LXXXVIII	-12.105	H –bond acceptor H -bondvacceptor	Lys32 Lys32	2.00 1.98
66	LXXXIX	-17.738	H –bond acceptor π –cation bond	Lys32 Pro1	1.78 -
67	XC	-17.206	No interaction	-	-
68	XCI	-20.468	H –bond acceptor H –bond acceptor H –bond donor H –bond donor π – π bond	Lys32 Lys32 IIe64 Pro1 Tyr36	2.58 2.05 2.10 2.07 -
69	XCII	-17.832	H –bond acceptor H –bond acceptor H –bond donor	Lys32 Lys32 Pro1	2.18 2.47 2.03
70	XCIII	-10.681	π –cation bond	Pro1	-
71	XCIV	-11.172	No interaction	-	-

CHAPTER THREE

DISCUSSION

3. DISCUSSION

3.1 QSAR Study

The data set of dihydropyrimidine derivatives was selected for QSAR from (Abdo, 2015) scientific paper 16 compounds were selected. The biological activity of each compound was changed into pIC_{50} and considered as the dependent variables. The values of descriptors were used as the independent variables the data set was divided into training set (13 compounds) and test set (3 compounds).

7 descriptors were selected (ASA- , dens, E.ang, MINDO-LUMO, $npr^{2,}$ pm3.dipole, Vsurf.IW1).

Where:

ASA- (negative accessible surface area) is the surface area of a biomolecule that is accessible to a solvent. ASA is often used when calculating the transfer free energy required to move a biomolecule from aqueous solvent to a non-polar solvent (Connolly, 1983).

Dens (mass density) is a quantitative expression of the amount of mass contained per unit volume .the unit is kilogram per meter cubed (Kg/m^3). Mass density can help to predict chances of corrosion of substance.

E.ang (Angle bend energy) angle bending between atoms that are geminal to each other (bonded to the same central atom).

MINDO LUMO (LUMO energy eV) is the lowest energy orbital that has the scope to accept electrons and hence it act as electron acceptor and characterizes the susceptibility of the molecule toward attack by nucleophilies.

npr^2 (Normalized PMI ratio (2)).

PM3.dipole (dipole moment) is measurement of the separation of two opposite electrical charges. In chemistry , dipole moment are applied to the distribution of electrons between two bonded atoms .The existence of a dipole moment is the different between polar and non polar bonds. Molecules with a net dipole moment are polar molecules (Dinnel *et al.*, 1993).

Vsurf-IW1 (hydropholic integy moment at -0.2) integy moments are vectors pointing from the center of mass to the center of hydropholic regions. High integy moments indicate a clear concentration of hydrated regions in only one part of the molecular surface , small moment indicate that the polar moieties are either close to the center of mass or they balance at opposite end of molecule, so that their resulting barycenter is close to the center of the molecule .

The training set consisting of 13 compounds was used to perform the correlation between the selected descriptors and PIC_{50} .The value of correlation coefficient for each pair of selected descriptors was examined. The greatest value of correlation coefficient (0.889) is that belonging to the pair of descriptors npr^2 and MIDO-LUMO.

The best models were obtained for predication of training set of dihydropyrimidine derivatives are given below:

PIC_{50}=5.213 + 0.373 npr^2 - 0.674MNDO-LUMO

(Model [1]) r^2= 0.889.

PIC_{50}=3.257+0.507 npr^2 + 0.584 PM3-dipole

(Model [2]) r^2= 0.863.

PIC_{50}=8.705-0.142 Eang -1.000 MNDO-LUMO

(Model [3]) r^2= 0.807.

Evaluation of the models with leave-one-out (LOO) cross-validation coefficient (q^2) showed that the difference between r^2 and q^2 values were less than 0.3, which indicated that the models were valid (Golbraikh and Tropsha, 2002). The statistical reliability of the QSAR model was evaluated based on several criteria, i.e. Fischer's value for statistical significance (F), adjusted squared correlation coefficient, standard error of estimation (SEE) root mean square root (RMSE) (Dearden *et al.*, 2009). F (F-test) or p-value shows value.

The plot showing the fit between observed and calculated activities for the training and test set compounds is given in figure 3.1-3.3 for the best model (model [1]).

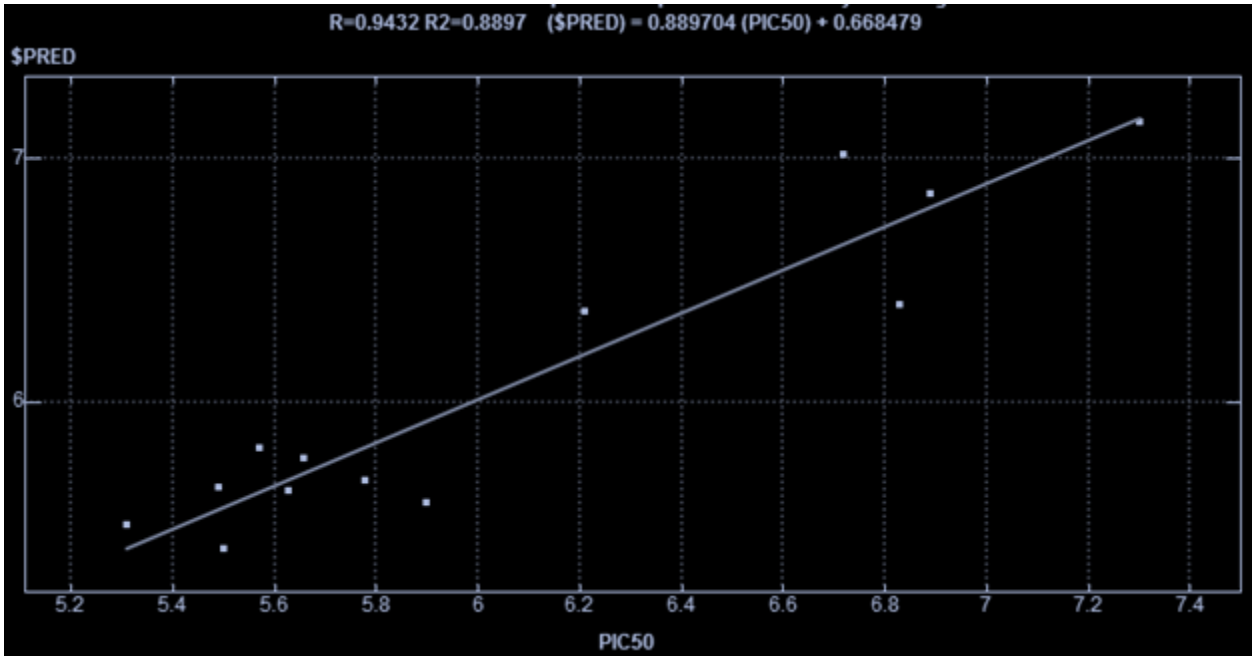


Figure 3.1: plot of predicted versus experimental pIC_{50} values for training set compouound **(model [1])**

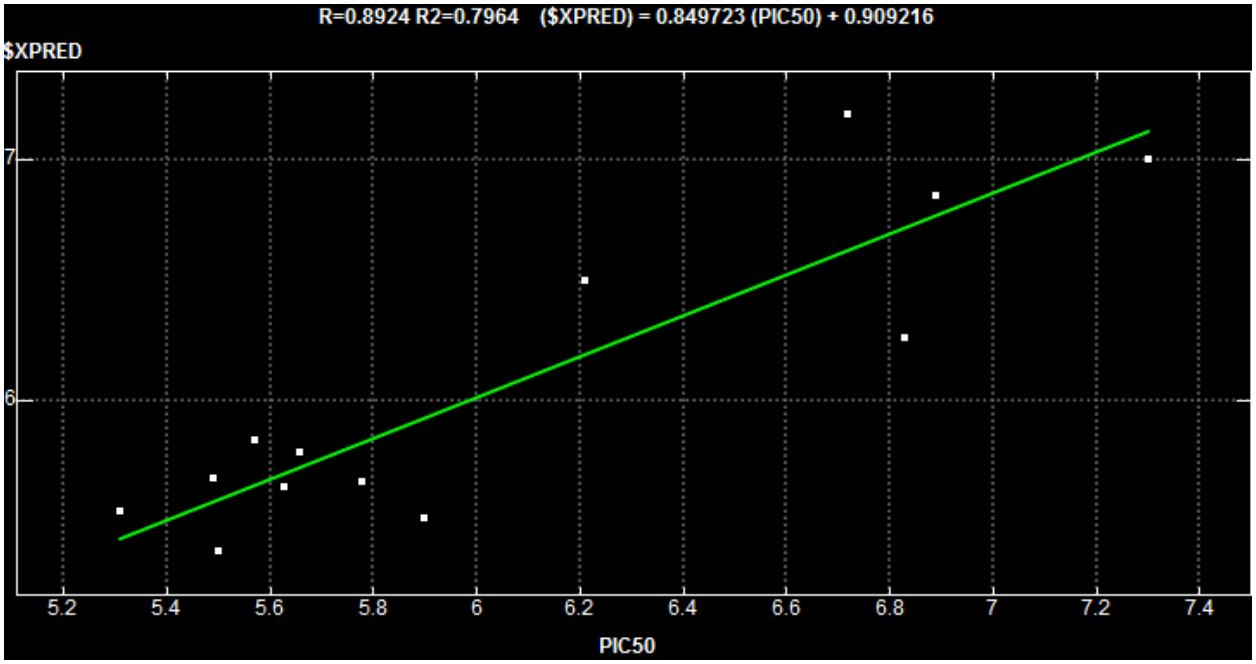


Figure 3.2: plot of predicted versus experimental pIC_{50} values of cross validation **(model [1])**

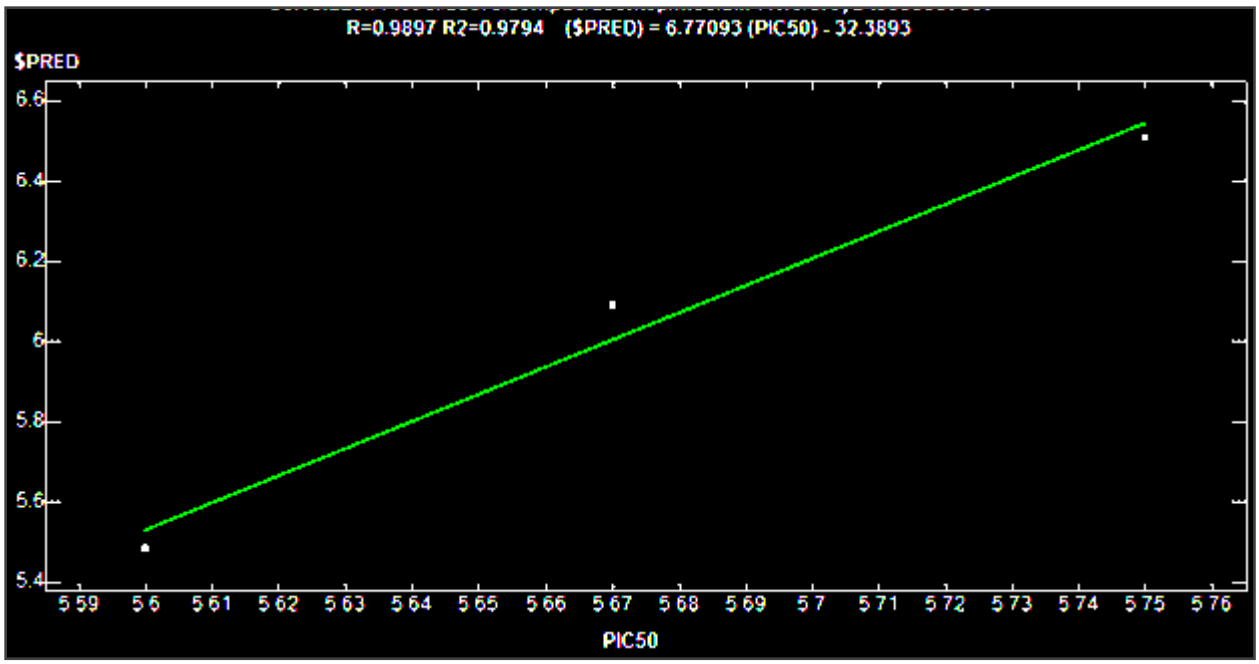


Figure 3.3: plot of predicted versus experimental pIC_{50} of test set compound **(model [1]).**

From data set which contained 16 compounds, 3 compounds were selected to design 94 compounds as new compounds of dihydropyrimidine derivatives and their descriptors were calculated. The model [1] obtained from collected data set was used to predicted the biological activity of a new dihydropyrimidine derivatives.

According to the biological activity of reference compound *N*-[6-(4-chlorophenoxy)hexyl]-1-(1-diazyn-1-ylidene)-*N*'-(pyridin-4-yl)methanediamine (CHS 828) in term of PIC_{50} (PIC_{50} = 5.63), 71 compounds were selected from a new designed of dihydropyrimidine derivatives showing a higher values of predicted biological activity than CHS828.

The higher PIC_{50} values of group A were summarized in 4 compounds (6.26 for II, 6.83 for V, 6.03 for VI, 6.21 for VIII).The higher PIC_{50} values of group B were summarized in 3 compounds (5.80 for XIII, 6.03 for II, 6.23 for IV). The higher PIC_{50} values of group C were summarized in 8 compounds (6.67 for LXII, 7.17 for LXVI, 6.19 for LXIII, 6.91 for LXVII, 6.73 for LXIV, 6.67 for LXIX, 7.33 for LXX, 7.27 for LXXIII). The higher PIC_{50} values of group D were summarized in 4 compounds (6.31 for XXX, 5.80 for XXXV and XXXXVI, 5.90 for XXXIX). The higher PIC_{50} values of group E were summarized in 2 compounds (5.74 for XLV and 6.26 for XLIII). The higher PIC_{50} values of group F were summarized in 6 compounds (6.05 for LII, 6.14 for LIV, 5.96 for LV and LVII, 6.03 for LX, 6.22 for LXI). The higher PIC_{50} values

of group G were summarized in 6 compounds (6.63 for LXXIV, 6.29 for LXXV, 6.53 for LXXVIII, 6.33 for LXXXVII, 6.70 for XL, 6.30 for XCIV).

3.2 docking study:

For some compounds of training set and new designed dihydropyrimidine derivatives

Macrophage migration inhibitory factor (MIF) also known as glycosylation- inhibiting factor is a protein that in humans and important regulator of innate immunity.

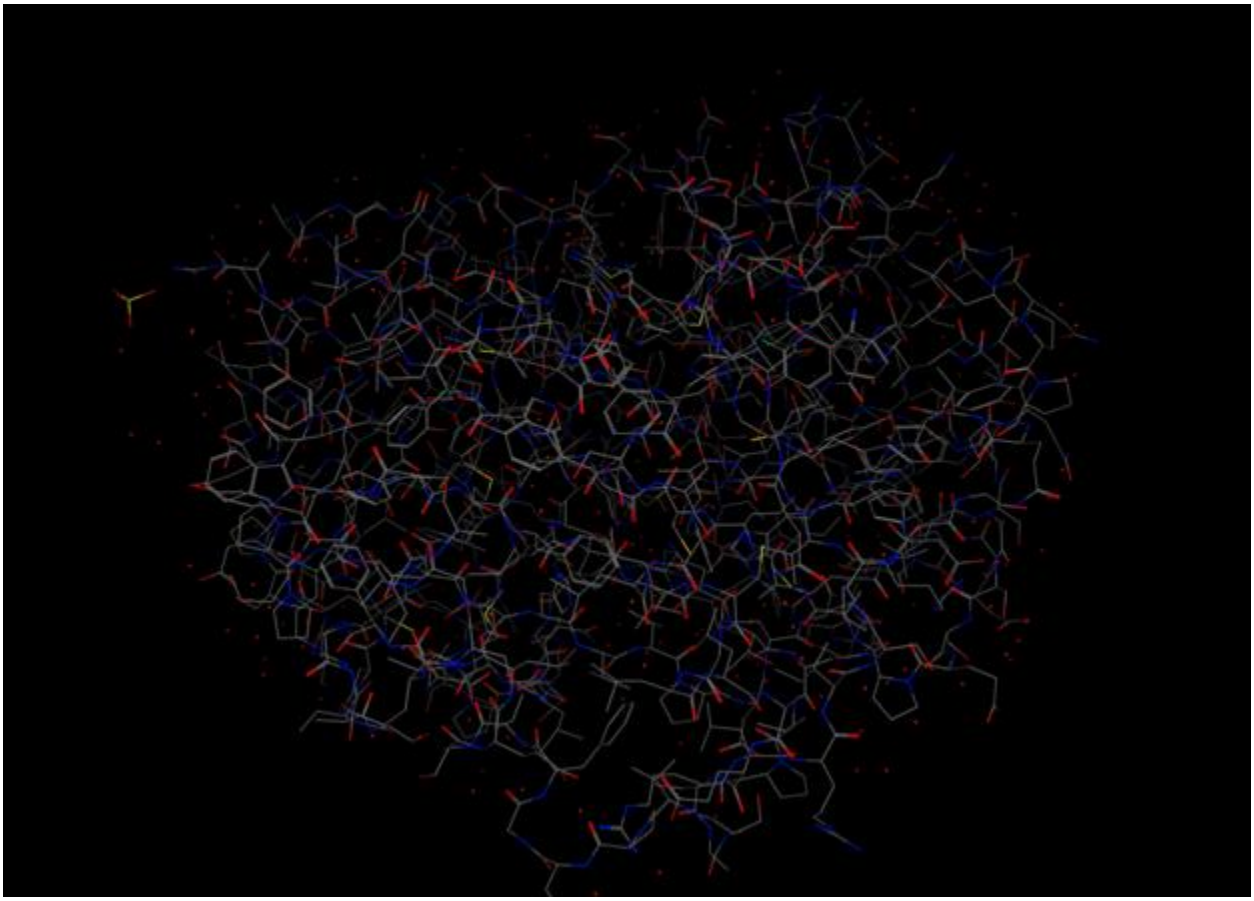
MIF is pleiotropic pro- inflammatory cytokine , which possesses a contributing role in cancer progression and metastasis and, thus, is now considered a promising anticancer drug target. Many MIF –inactivating strategies have proven successful in delaying cancer growth. (2014 INT J ONCOL.45)

Macrophage migration inhibitory factor assembles in to a trimer composed of three identical subunits .each of these monomers contain two antiparallel alpha helices and a four –standard beta sheet. The monomers surround a central channel with 3-fold rotational symmetry.

MIF contains two motifs with catalytic activity .The first is a 27 amino acid motif located at N-terminus function as a phenylpyruvate tautomerase that can catalyze the conversion of 2-carboxy-2,3-dihydroindole-5,6-quinone in to 5,6-dihydroindole-2-carboxylic acid.MIF also contains a cys-Ala-leu-cys catalytic site between residues 57 and60 that appears to function as disulfide reductase.

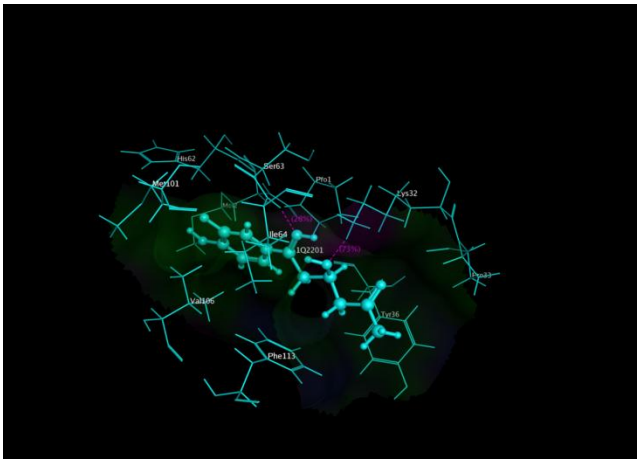
MIF is a potential drug target for sepsis, rheumatoid, arthritis and cancer and shows efficacy in melanoma and colon cancer models (Schroder et al; 2005).

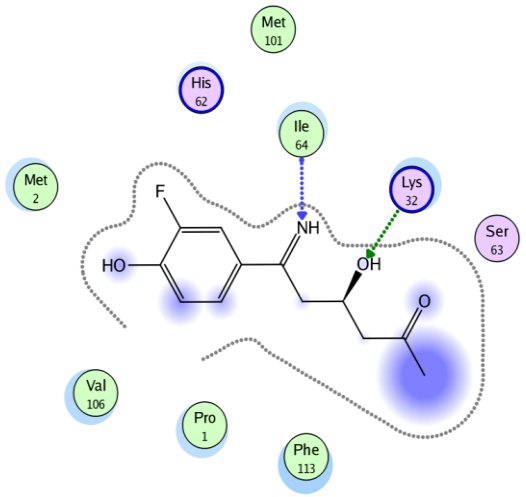
Figure (3.4) shows structure of macrophage migration inhibitory factor protein (MIF) that was imported from PDB.



The molecular docking produced the different docking conformation based on building energy. The variants with the minimal energy of enzyme –inhibitor complex were selected for studies of binding mode.

All the docked conformation for training set and new designed of dihydropyrimidine derivatives compounds were analyzed and it was found that the most favorable docking poses with maximum number of interaction were those ranked the highest, based on the minimal binding energy, computed as a negative value by the software as shown in table(2.6)(2.7).

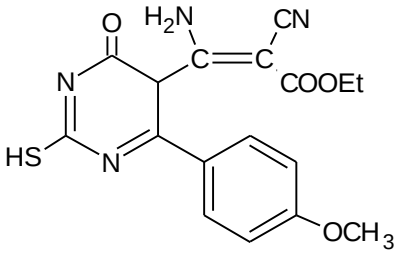




A B

Figure 3.5 :(A) Active sites of MIF protein, **(B)** ligand interaction with protein.

From training set compounds, 6 compounds were selected which had higher predicted PIC_{50} than reference compound (CHS 828), these ligands were prepared to docking with a protein 4K9G. From 6 compounds, 3 were docked with protein. Compound (H II) was showed the most favorable interactions.



H II

The binding free energy(s) of compound H II is -11.367 Kcal/mol and H II showing 2 π -bond interactions and one H bond –interaction. π - π bond between phenyl ring and amino group phe113 and π –cation bond between phenyl ring and amino group pro1. The H-bond acceptor between O of carboxylic and amino group Lys32 and bond length was 3.80.

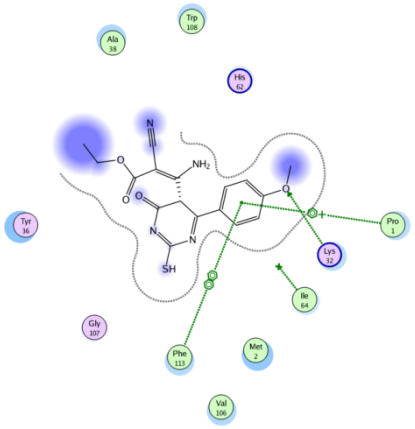
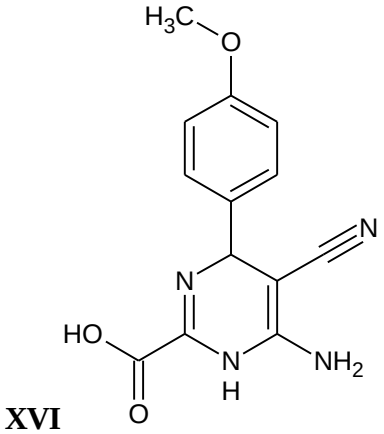


Figure 3.6: The 2D molecular interactions of compound H II with active site of MIF

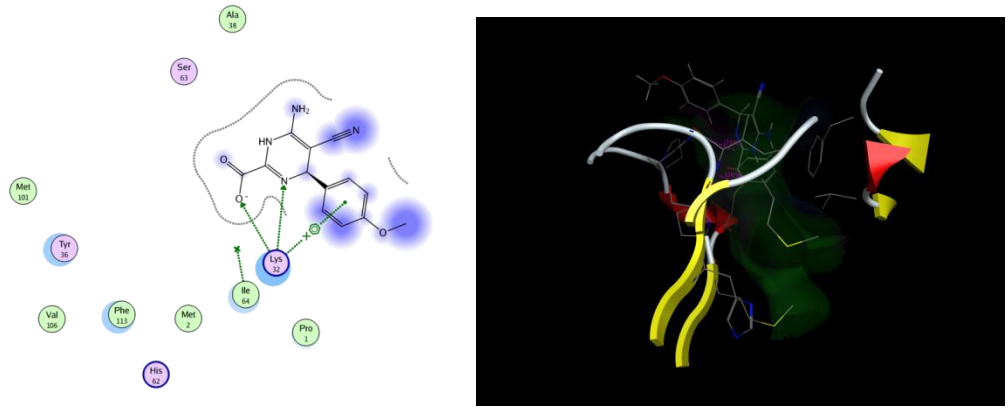
From 94 new designed derivatives, 71 derivatives were selected which had higher predicted PIC_{50} than reference compound (CHS 828), these ligands were prepared to docking with a protein 4K9G. From 71 compounds, 52 were docked with protein. 20 compounds were showed the most favorable interactions.

There were five compounds (XVI, XIX, LVII, XCI, and XCII) showed higher interactions.



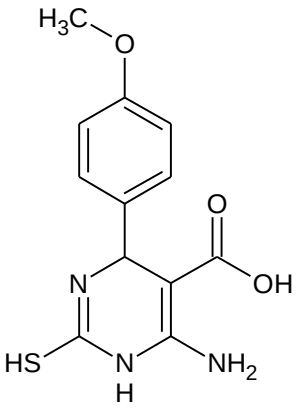
XVI

The binding free energy(s) of compound XVI is -19.403 Kcal/mol and XVI showing 2 H-bond interactions and one pi-interaction. Two H-bond acceptor between N of pyrimidine ring and O of carboxylic acid with amino group was in lys32 the bond length of 2H –bond was 2.14 and 2.70 respectively .while π –cation interaction between phenyl ring and amino group was in Lys 32.



A B

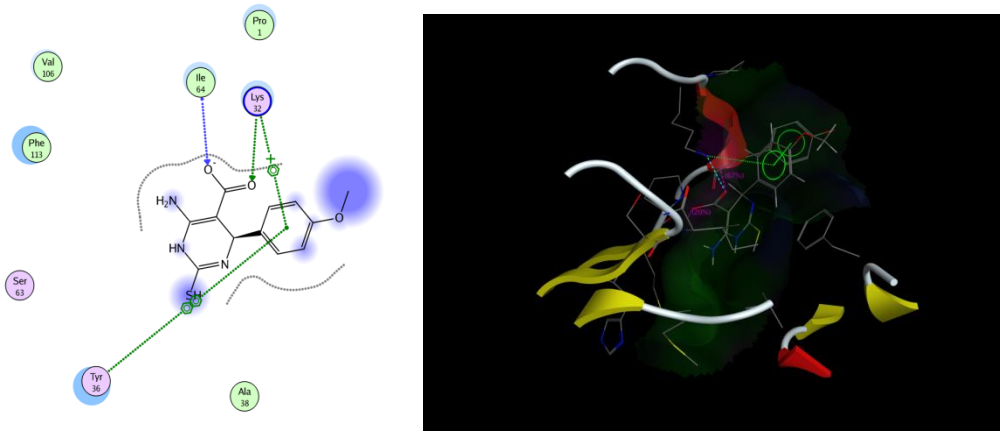
Figure 3.7: The 2D (A) and3D (B) Molecular interactions of compounds XVI with active site of MIF



XIX

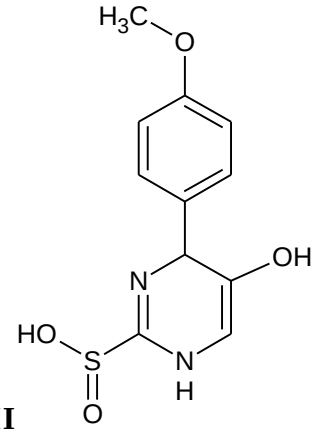
The binding free energy (s) of compound XIX is -14.394 Kcal/mol and XIX forming 2H-bond interactions and 2pi-interactions.

One of H-bond donor between O of carboxylic and amino group was in lle 64 with bond length 2.01, the other H-bond acceptor between O of carbonyl group and amino group was in Lys32 with bond length 1.83. While the one of π –cation interactions between phenyl group and amino groups Lys32 and the other $\pi - \pi$ interactions between phenl ring and amino group was in Tyr36.



A B

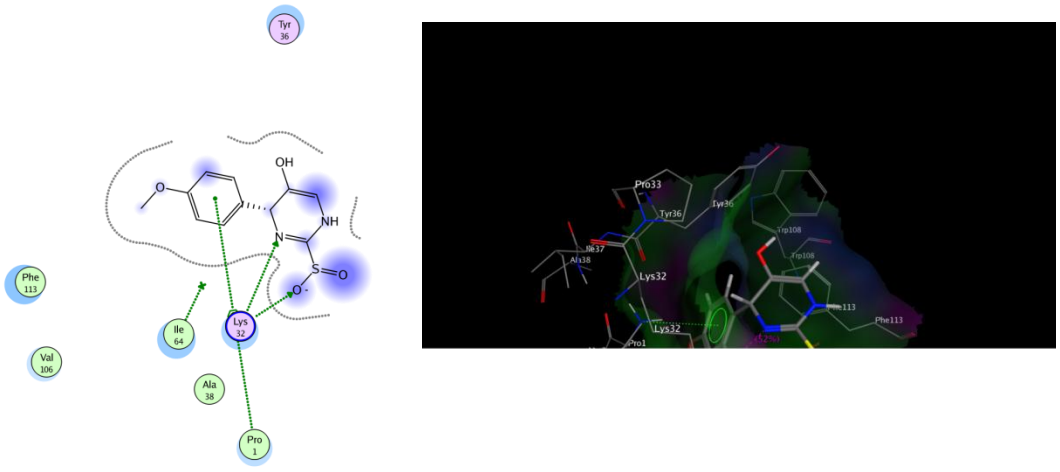
Figure 3.8: the 2D (A) and 3D (B) Molecular interactions of compounds XIX with active site of MIF



LVII

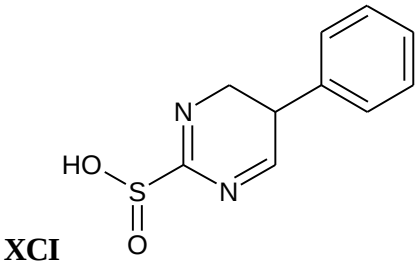
The binding free energy of compounds LVII -17.623 Kcal/mol and LVII forming 2 H-bond interactions and one pi-interaction.

Two H-bond acceptor between N of pyrimidine ring and O of sulfinic acid with amino group was in Lys32 .the bond length was 1.93 and 2.51 respectively.while π –cation interaction between phenyl group and amino group was in pro1.



A B

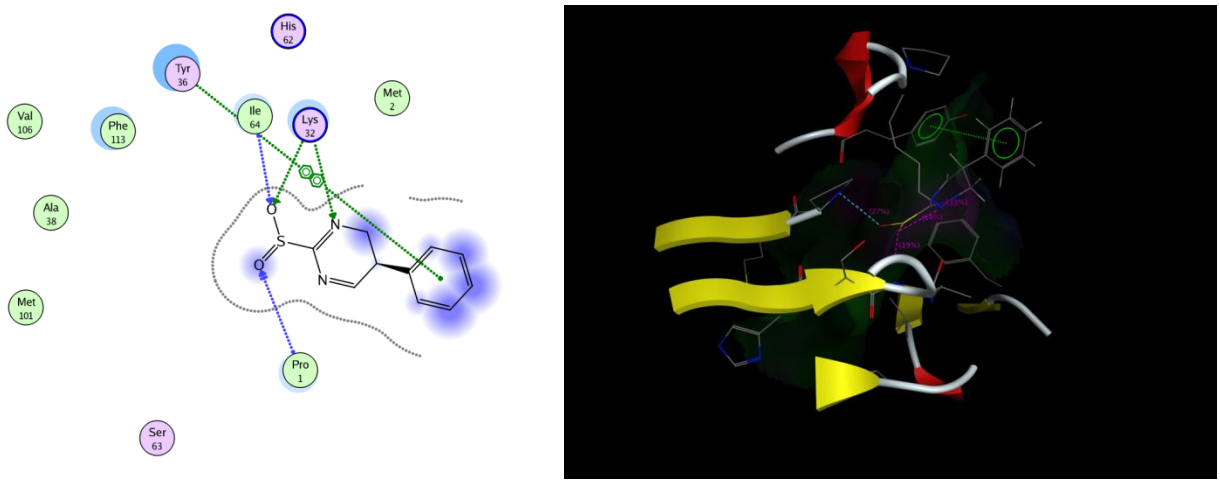
Figure 3.9: The 2D (A) and 3D (B) Molecular interactions of compound LVII with active site of MIF



XCI

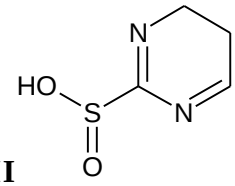
The binding free energy (s) of compound XCI is -20.468 Kcal/mol and XCI forming 4 H-bond interactions and one pi-interaction.

The 2 H-bonds acceptor between N of pyrimidine ring and O of sulfinic acid with amino group was in Lys32, bond length was 2, 58 and 2.05 respectively. The third H-bond donor between O of sulfonyl group and amino group was in lle64 the bond length was 2.10.the fourth H-bond donor between O of carbonyl and amino group was in pro1 the bond length was 2.07.while the π - π interaction between phenyl group and amino group was in Tyr36.



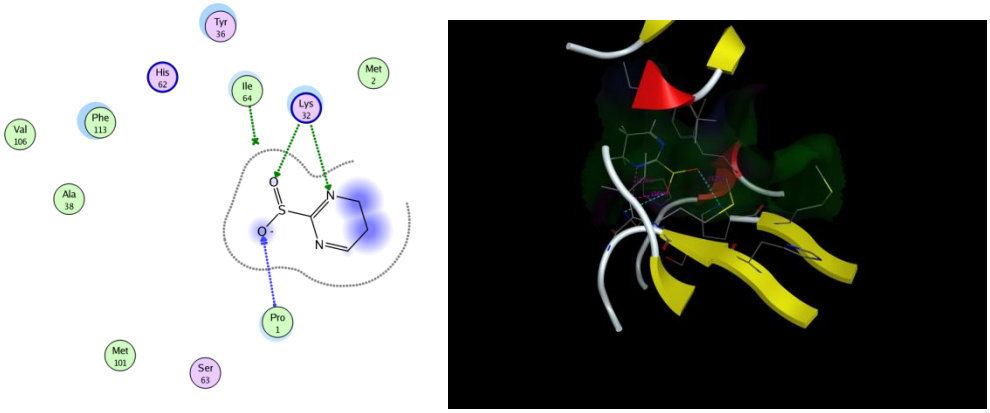
A B

Figure 3.10: The 2D (A) and 3D(B) Molecular interactions of compound XCI with active site of MIF



XCII

The binding free energy of compound XCII is -17.832 Kcal/mol and XCII forming 3H-bond interactions. The two H-bond acceptor between N of pyrimidine ring and O of sulfonyl group with amino group was in Lys32, third H-bond donor between O of sulfinic acid and amino group was in pro1. The bond length was 2.18, 2.47 and 2.03 respectively.



A B

Figure 3.11: The 2D and 3D Molecular interactions of compoundG XCII with active site of MIF

Comounds (VI, XV, XVIII, LIV, LIX) showed three interactions with an amino acid by H-bond interactions, pi-interactions or both.while compounds (II, VII, IX, X, XIV, XVII, LXII, LXIX, XXXIII, XLIX, LXXX, LXXXV, LXXXVIII and LXXXIX) showed two interaction with an amino acid by H-bond interactions , pi-interactions or both. while compounds (I, IV, V, VI, XIII, LXIV, LXX, LXXI, LXXII, XXX, XXXIX, LV, LVI, LXXVII, LXXVIII, LXXXIV, LXXXVI, LXXVII and XCII) showed one interaction with an amino acid by H-bond interaction or pi-interaction. All of the figures 2D models of the listed compound are shown in the appendix.

3.3 Retrosynthetic analysis:

Retrosynthetic analysis is a technique widely used by organic chemists to design synthetic routes to "target" molecules, where the target is transformed in to simpler precursor molecules until commercially available "starting" molecules are identified (corey *et al*, 1995).

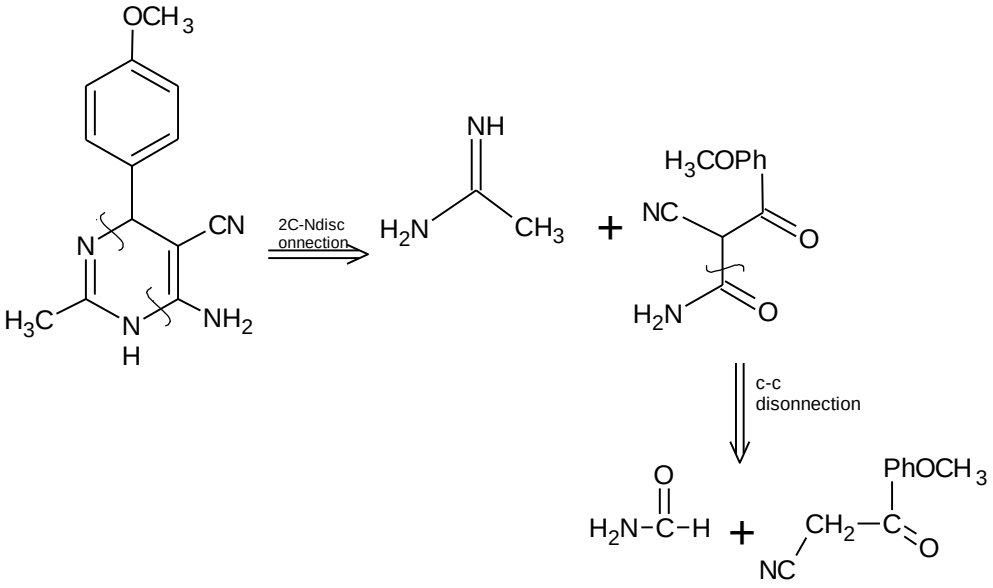


Figure 3.12: Retrosynthetic analysis of 6-amino-4-(4-methoxyphenyl)-2-methyl-1,4-dihydropyrimidine-5-carbonitrile (XV).

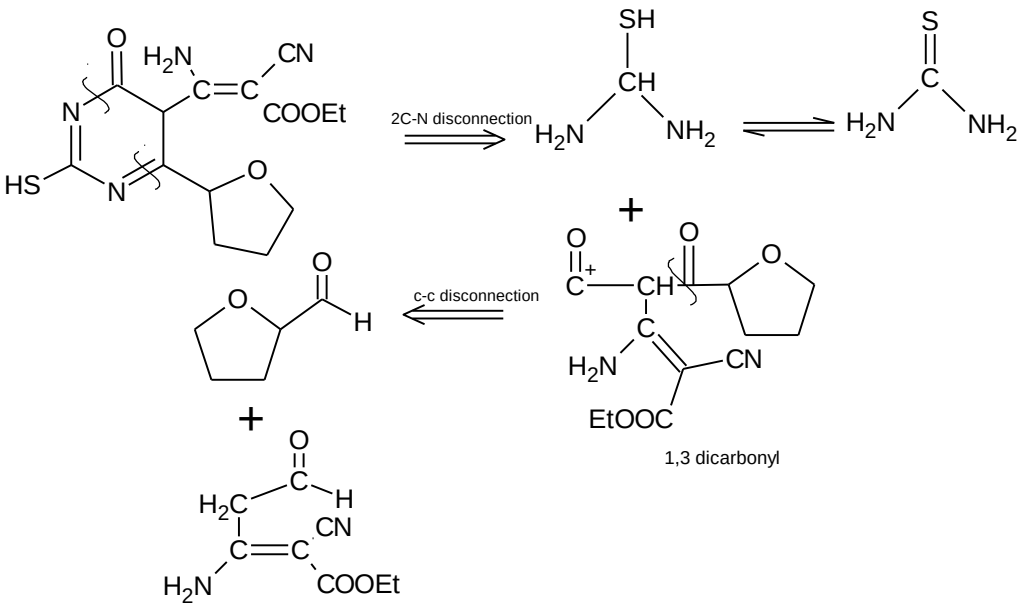


Figure 3.13: Retrosynthetic analysis of ethyl(2E)-3-amino-2-cyano-3-[4-oxo-2-sulfanyl-6-(tetrahydrofuran-2-yl)-4,5-dihydropyrimidin-5-yl]prop-2-enoate (LXII).

CHAPTER FOUR

CONCLUSION AND RECOMMENDATIONS

4. CONCLUSION AND RECOMMENDATIONS:

The following points could be concluded from this work:

- QSAR is a mathematical relationship that correlates chemical structure and biological activity for a series of compounds.
- In this study, descriptors were used to represent the dihydropyrimidine derivatives structures, which were taken as the independent variable, while the biological activities (PIC_{50}) were used as the dependent variable.
- PLS Analysis resulted in a QSAR model that was used to predict the activity of dihydropyrimidine derivatives.
- The correlation coefficient (r^2 <0.7) when the descriptor logP(o/w) is correlated against biological activity PIC_{50} of data set.
- From data set compounds there were 94 compounds designed as a new dihydropyrimidine derivatives, 71 compounds were selected according to their PIC_{50} higher than reference compound.
- Docking study of some compounds of training set showed interactions with active site of protein by H-bond interaction, pi-interaction or both.
- Docking study of new designed of dihydropyrimidine derivatives showed a good interaction with active site of protein by H-bond interaction, pi- interaction or both. Also the binding energy in some interactions was lower than other.
- Some of new derivatives showed a higher biological activity.
- The 2D models of dihydropyrimidine derivatives showed that the main residues in the active pocket of macrophage migration inhibitory factor.
- Docking is correlated strongly with QSAR results.
- QSAR and docking results of this work could be useful for other chemists working on the field of predicted biological activity as macrophage migration (MIF) enzyme inhibitor for designed or newly heterocyclic dihydropyrimidine synthesis.

CHAPTER FIVE

References

5. References:

× Abdo, N.Y.M., 2015. Synthesis and antitumor evaluation of novel dihydropyrimidine, thiazolo [3, 2-a] pyrimidine and pyrano [2, 3-d] pyrimidine derivatives. *Acta Chimica Slovenica*, *62*(1), pp.168-180.

× Alam, O., Khan, S.A., Siddiqui, N., Ahsan, W., Verma, S.P. and Gilani, S.J., 2010. Antihypertensive activity of newer 1, 4-dihydro-5-pyrimidine carboxamides: Synthesis and pharmacological evaluation. *European journal of medicinal chemistry*, *45*(11), pp.5113-5119.

× Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. and Clarke, M.F., 2003. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences*, *100*(7), pp.3983-3988.

× Alvim, H.G., Lima, T.B., de Oliveira, A.L., de Oliveira, H.C., Silva, F.M., Gozzo, F.C., Souza, R.Y., da Silva, W.A. and Neto, B.A., 2014. Facts, presumptions, and myths on the solvent-free and catalyst-free Biginelli reaction. What is catalysis for?. *The Journal of Organic Chemistry*, *79*(8), pp.3383-3397.

× Anjos, J.V.D., Srivastava, R.M., Costa-Silva, J.H., Scotti, L., Scotti, M.T., Wanderley, A.G., Leite, E.S., Melo, S.J.D. and Junior, F.J., 2012. Comparative computational studies of 3, 4-dihydro-2, 6-diaryl-4-oxo-pyrimidine-5-carbonitrile derivatives as potential antinociceptive agents. *Molecules*, *17*(1), pp.809-819.

× Balaban, A.T., 1982. Highly discriminating distance-based topological index. *Chemical physics letters*, *89*(5), pp.399-404.

× Bansal, R.K. ed., 2009. *Phosphorous Heterocycles I* (Vol. 20). Springer.

× Bharatam, P.V., Khanna, S. and Francis, S.M., 2010. Modeling and informatics in drug design. *Pharmaceutical Sciences Encyclopedia: Drug Discovery, Development, and Manufacturing*, pp.1-46.

× Bonchev, D., 1983. *Information theoretic indices for characterization of chemical structures* (No. 5). Research Studies Press.

× Bonnet, D. and Dick, J.E., 1997. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature medicine*, *3*(7), pp.730-737.

- Connolly, M.L., 1993. The molecular surface package. *Journal of molecular graphics*, *11*(2), pp.139-141.
- Cramer, R.D., Patterson, D.E. and Bunce, J.D., 1988. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*, *110*(18), pp.5959-5967.
- Darehkordi, A. and Hosseini, M.S., 2012. Montmorillonite modified as an efficient and environment friendly catalyst for one-pot synthesis of 3, 4-dihydropyrimidine-2 (1h) ones.*Iranian Journal of Materials Science and Engineering*, *9*(3), pp.49-57.
- Dearden, J.C., Cronin, M.T. and Kaiser, K.L., 2009. How not to develop a quantitative structure–activity or structure–property relationship (QSAR/QSPR). *SAR and QSAR in Environmental Research*, *20*(3-4), pp.241-266.
- Deres, K., Schröder, C.H., Paessens, A., Goldmann, S., Hacker, H.J., Weber, O., Krämer, T., Niewöhner, U., Pleiss, U., Stoltefuss, J. and Graef, E., 2003. Inhibition of hepatitis B virus replication by drug-induced depletion of nucleocapsids.*Science*, *299*(5608), pp.893-896.
- DINNEL, S., MATULEWSKI, K. and WILLIEMS, R., 1993. Coastal Ocean Modeling Workshop.
- Dolinšek, J., Bharatam, J., Dusseault, M. and Pintar, M.M., 1998. Two-dimensional NMR study of surface water dynamics in hydrated silica spheres. *Physical Review B*, *58*(11), p.7340.
- Dunbrack, R.L. and Karplus, M., 1994. Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nature structural biology*, *1*(5), pp.334-340.
- Eckhardt, S., 2002. Recent progress in the development of anticancer agents. *Current medicinal chemistry-anti-cancer agents*, *2*(3), pp.419-439.
- Ginestier, C., Hur, M.H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., Jacquemier, J., Viens, P., Kleer, C.G., Liu, S. and Schott, A., 2007. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell stem cell*, *1*(5), pp.555-567.
- Golbraikh, A. and Tropsha, A., 2002. Beware of q2!. *Journal of molecular graphics and modelling*, *20*(4), pp.269-276.

- Goossen, E., CHEM 443: Biochemistry Lab Spring 2009 Dr. Jeff Watson, Instructor.
- Gressler, V., Moura, S., Flores, A.F., Flores, D.C., Colepicolo, P. and Pinto, E., 2010. Antioxidant and antimicrobial properties of 2-(4, 5-dihydro-1H-pyrazol-1-yl)-pyrimidine and 1-carboxamidino-1H-pyrazole derivatives.*Journal of the Brazilian Chemical Society*, *21*(8), pp.1477-1483.
- Hall, L.M., Hall, L.H. and Kier, L.B., 2003. Modeling drug albumin binding affinity with e-state topological structure representation. *Journal of chemical information and computer sciences*, *43*(6), pp.2120-2128.
- Ho, M.M., Ng, A.V., Lam, S. and Hung, J.Y., 2007. Side population in human lung cancer cell lines and tumors is enriched with stem-like cancer cells. *Cancer research*, *67*(10), pp.4827-4833.
- Holtje, H., 2003. Molecular modelingBasic principles and applications.
- Hosoya, H., 1971. Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. *Bulletin of the Chemical Society of Japan*, *44*(9), pp.2332-2339.
- Ismaili, L., Nadaradjane, A., Nicod, L., Guyon, C., Xicluna, A., Robert, J.F. and Refouvelet, B., 2008. Synthesis and antioxidant activity evaluation of new hexahydropyrimido [5, 4-c] quinoline-2, 5-diones and 2-thioxohexahydropyrimido [5, 4-c] quinoline-5-ones obtained by Biginelli reaction in two steps. *European journal of medicinal chemistry*, *43*(6), pp.1270-1275.
- Jain, K. S., T. S. Chitre, P. B. Miniyar, M. K. Kathiravan, V. S. Bendre, V. S. Veer, S. R. Shahane, and C. J. Shishoo. "Biological and medicinal significance of pyrimidines." *Current science* (2006): 793-803.
- Kaan, H.Y.K. and Ulaganathan, V., 2010. Rath,; Prokopcová, OH; Dallinger, D.; Kappe, CO; Kozielski, F. *J. Med. Chem*,*53*, p.5676.
- Khobragade, C.N., Bodade, R.G., Dawane, B.S., Konda, S.G. and Khandare, N.T., 2010. Synthesis and biological activity of pyrazolo [3, 4-d] thiazolo [3, 2-a] pyrimidin-4-one derivatives: in silico approach. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *25*(5), pp.615-621.

- Klebe, G., Abraham, U. and Mietzner, T., 1994. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *Journal of medicinal chemistry*, *37*(24), pp.4130-4146.
- Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E., 1982. A geometric approach to macromolecule-ligand interactions. *Journal of molecular biology*, *161*(2), pp.269-288.
- Leach, A.R. and Leach, A.R., 2001. *Molecular modelling: principles and applications*. Pearson education.
- Lee, C.W., Hong, D.H., Han, S.B., Jung, S.H., Kim, H.C., Fine, R.L., Lee, S.H. and Kim, H.M., 2002. A novel stereo-selective sulfonylurea, 1-[1-(4-aminobenzoyl)-2, 3-dihydro-1H-indol-6-sulfonyl]-4-phenyl-imidazolidin-2-one, has antitumor efficacy in in vitro and in vivo tumor models. *Biochemical pharmacology*, *64*(3), pp.473-480.
- Lee, Y.S., Park, S.M., Kim, H.M., Park, S.K., Lee, K., Lee, C.W. and Kim, B.H., 2009. C5-Modified nucleosides exhibiting anticancer activity. *Bioorganic & medicinal chemistry letters*,*19*(16), pp.4688-4691.
- Lloyd, J., Finlay, H.J., Atwal, K., Kover, A., Prol, J., Yan, L., Bhandaru, R., Vaccaro, W., Huynh, T., Huang, C.S. and Conder, M., 2009. Dihydropyrazolopyrimidines containing benzimidazoles as KV1. 5 potassium channel antagonists. *Bioorganic & medicinal chemistry letters*, *19*(18), pp.5469-5473.
- Matos, L.H.S., Masson, F.T., Simeoni, L.A. and Homem-de-Mello, M., 2018. Biological activity of dihydropyrimidinone (DHPM) derivatives: A systematic review. *European Journal of Medicinal Chemistry*, *143*, pp.1779-1789.
- Nair, N. and Goodman, J.M., 1998. Genetic algorithms in conformational analysis. *Journal of Chemical Information and Computer Sciences*, *38*(2), pp.317-320.
- Park, B.H. and Levitt, M., 1995. The complexity and accuracy of discrete state models of protein structure. *Journal of molecular biology*, *249*(2), pp.493-507.
- Patrawala, L., Calhoun, T., Schneider-Broussard, R., Zhou, J., Claypool, K. and Tang, D.G., 2005. Side population is enriched in tumorigenic, stem-like cancer cells, whereas ABCG2+ and ABCG2− cancer cells are similarly tumorigenic. *Cancer research*, *65*(14), pp.6207-6219.
- Paul.J.Hall. Computational Genomics .

- Randic, M., 1975. Characterization of molecular branching. *Journal of the American Chemical Society*, *97*(23), pp.6609-6615.
- Robinson, D.D., Winn, P.J., Lyne, P.D. and Richards, W.G., 1999. Self-organizing molecular field analysis: A tool for structure− activity studies. *Journal of Medicinal Chemistry*, *42*(4), pp.573-583.
- Rojo, F., Albanell, J., Rovira, A., Corominas, J.M. and Manzarbeitia, F., 2008, November. Targeted therapies in breast cancer. In *Seminars in diagnostic pathology* (Vol. 25, No. 4, pp. 245-261). WB Saunders.
- Rowlands, G.J., 2001. Experimental Organic Chemistry Daniel R. Palleros. John Wiley & Sons: New York. 2000. 833 pp. $86.95. ISBN 0-471-28250-2.
- Ruijter, E., 2012. Drug Discovery Today: Technol. 2013, 10, 15− 20.(b) Slobbe, P.; Ruijter, E.; Orru, RVA. *Med. Chem. Commun*, *3*, pp.1189-1218.
- Schröder, Martin, and Randal J. Kaufman. "The mammalian unfolded protein response." *Annu. Rev. Biochem*. 74 (2005): 739-789.
- Shen, M., Xiao, Y., Golbraikh, A., Gombar, V.K. and Tropsha, A., 2003. Development and validation of k-nearest-neighbor QSPR models of metabolic stability of drug candidates. *Journal of medicinal chemistry*, *46*(14), pp.3013-3020.
- Simurova, N. and Maiboroda, O., 2017. Biginelli reaction–an effective method for the synthesis of dihydropyrimidine derivatives (microreview). *Chemistry of Heterocyclic Compounds*, *53*(4), pp.413-415.
- Singh, B.K., Mishra, M., Saxena, N., Yadav, G.P., Maulik, P.R., Sahoo, M.K., Gaur, R.L., Murthy, P.K. and Tripathi, R.P., 2008. Synthesis of 2-sulfanyl-6-methyl-1, 4-dihydropyrimidines as a new class of antifilarial agents.*European Journal of Medicinal Chemistry*, *43*(12), pp.2717-2723.
- Stefani, H.A., Oliveira, C.B., Almeida, R.B., Pereira, C.M., Braga, R.C., Cella, R., Borges, V.C., Savegnago, L. and Nogueira, C.W., 2006. Dihydropyrimidin-(2H)-ones obtained by ultrasound irradiation: a new class of potential antioxidant agents. *European journal of medicinal chemistry*, *41*(4), pp.513-518.
- Sternberg, M.J. ed., 1996. *Protein Structure Prediction: A Practical Approach: A Practical Approach*. Oxford University Press, USA.

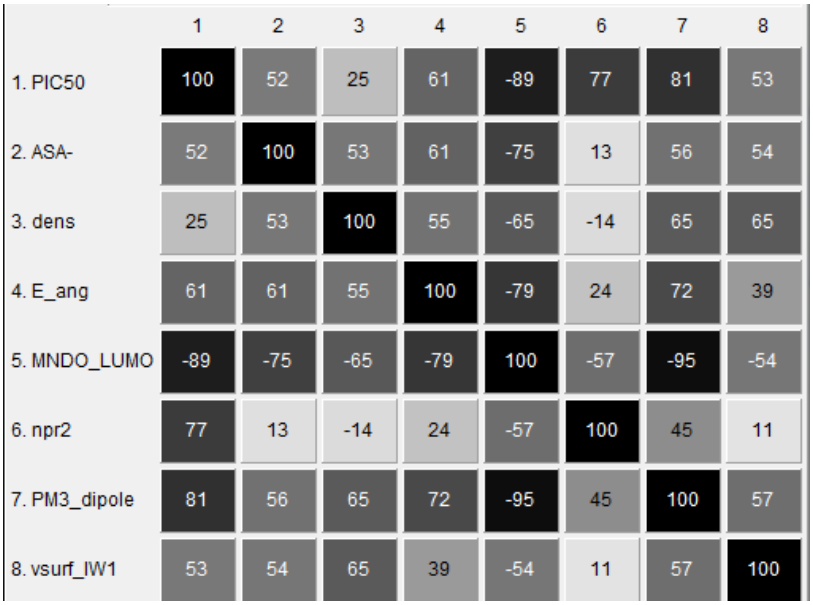
- Stray, S.J., Bourne, C.R., Punna, S., Lewis, W.G., Finn, M.G. and Zlotnick, A., 2005. A heteroaryldihydropyrimidine activates and can misdirect hepatitis B virus capsid assembly. *Proceedings of the National Academy of Sciences*,*102*(23), pp.8138-8143.
- Ultra, C.C., 8.0: http://www. cambridgesoft. com/products/family. cfm.
- van den Haak, H.J.W., 1981. *Synthesis and amination of naphthyridines* (Doctoral dissertation).
- van der Plas, H.C., Vollering, M.C., Jongejan, H. and Zuurdeeg, B., 1974. Ring transformations in reactions of heterocyclic compounds with nucleophiles (VIII). Conversion of pyrimidines into isoxazoles. *Recueil des Travaux Chimiques des Pays-Bas*, *93*(8), pp.225-227.
- Vedani, A. and Dobler, M., 2002. 5D-QSAR: the key for simulating induced fit?. *Journal of medicinal chemistry*, *45*(11), pp.2139-2149.
- Vedani, A., Briem, H., Dobler, M., Dollinger, H. and McMasters, D.R., 2000. Multiple-conformation and protonation-state representation in 4D-QSAR: the neurokinin-1 receptor system. *Journal of medicinal chemistry*, *43*(23), pp.4416-4427.
- Vedani, A., Dobler, M. and Lill, M.A., 2005. Combining protein modeling and 6D-QSAR. Simulating the binding of structurally diverse ligands to the estrogen receptor. *Journal of medicinal chemistry*, *48*(11), pp.3700-3703.
- Vedani, A., Dobler, M., Dollinger, H., Hasselbach, K.M., Birke, F. and Lill, M.A., 2005. Novel ligands for the chemokine receptor-3 (CCR3): a receptor-modeling study based on 5D-QSAR. *Journal of medicinal chemistry*, *48*(5), pp.1515-1527.
- Weissig, H. and Bourne, P.E. eds., 2003. *Structural Bioinformatics*. Wiley-Liss.
- Winquist, R.J., Furey, B.F. and Boucher, D.M., 2010. Cancer stem cells as the relevant biomass for drug discovery.*Current opinion in pharmacology*, *10*(4), pp.385-390.
- Wu, F.Y., Luo, Y. and Hu, C.B., 2018. Recent progress in the synthesis of thiazolo [3, 2-a] pyrimidine compounds. *MS&E*, *292*(1), p.012038.
- Zhu, X., Zhao, G., Zhou, X., Xu, X., Xia, G., Zheng, Z., Wang, L., Yang, X. and Li, S., 2010. 2, 4-Diaryl-4, 6, 7, 8-tetrahydroquinazolin-5 (1H)-one derivatives as anti-HBV agents targeting at capsid assembly. *Bioorganic & medicinal chemistry letters*, *20*(1), pp.299-301.

Appendix A

Table A.1: Models with descriptors in 2D dimensional used in predict the biological activity of dihydropyrimidine derivatives dataset

NO	eq	R^2	RMSE
1	PIC_{50}=7.376+4.44chi0-3.766weight	0.685	0.353
2	PIC_{50}=6.502-0.422logP(o/w)+0.732vdw-vol	0.599	0.399
3	PIC_{50}=6.644-0.300logP(o/w)+0.714chi1	0.591	0.409
4	PIC_{50}=7.245-1.507mr+2.18 chi0	0.576	0.410
5	PIC_{50}=7.486+0.659a-acc+0.134bpol	0.535	0.429
6	PIC_{50}=6.914+0.430a-acc+0.352chi1	0.560	0.418
7	PIC_{50}=6.638-0.456logP(o/w)-0.629logs	0.447	0.468
8	PIC_{50}=5.825+0.183logP(o/w)+0.753TPSA	0.430	0.475
9	PIC_{50}=5.604+0.988mr-0.349diameter	0.414	0.482
10	PIC_{50}=6.505-0.504logs+0.001nmol	0.254	0.544

Figure A.1: Correlation matrix of training set compounds:



Appendix B

Model interactions of designed dihydropyrimidine derivatives with macrophage migration inhibitory factor (MIF).

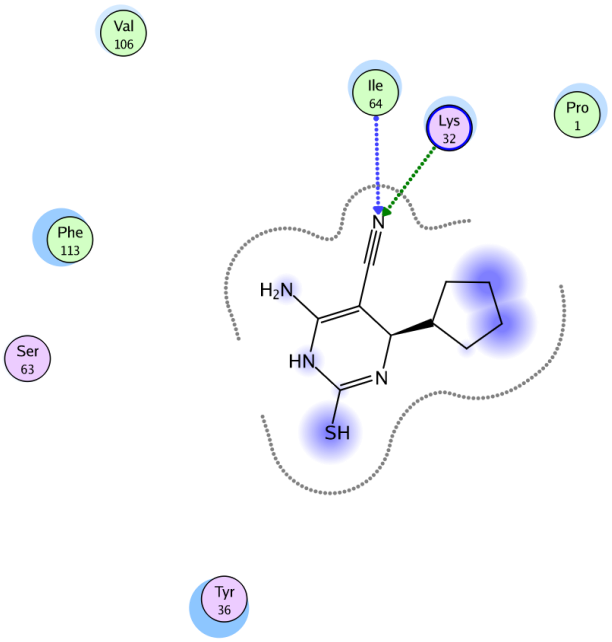


Figure B.1: Interactions of compound (I)

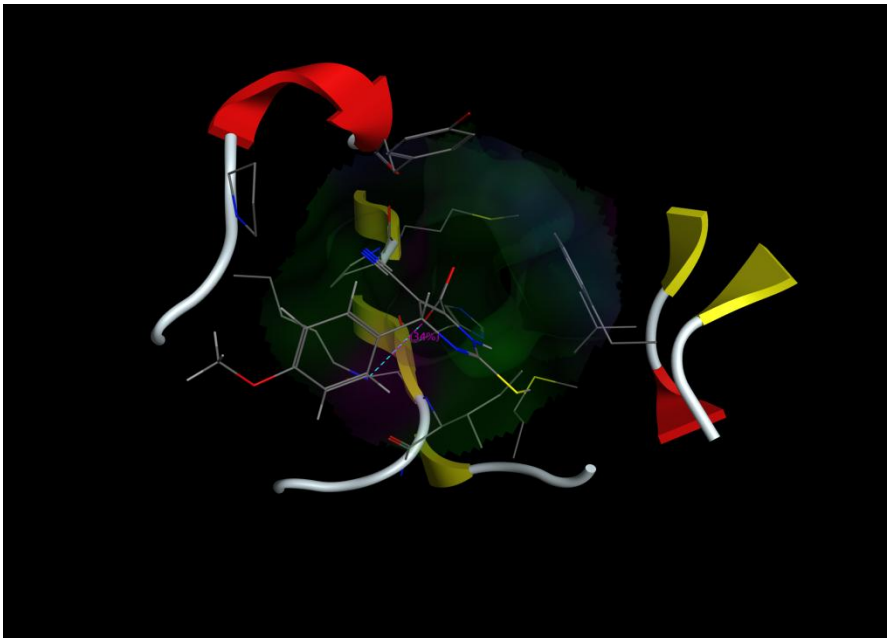


Figure B.2: Interactions of compound II

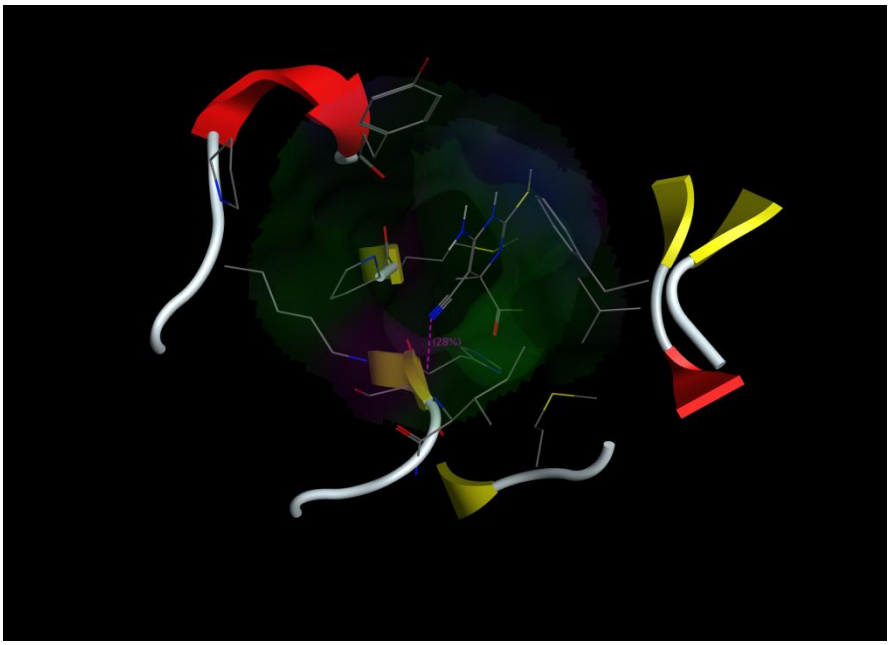


Figure B.3: Interactions of compound IV

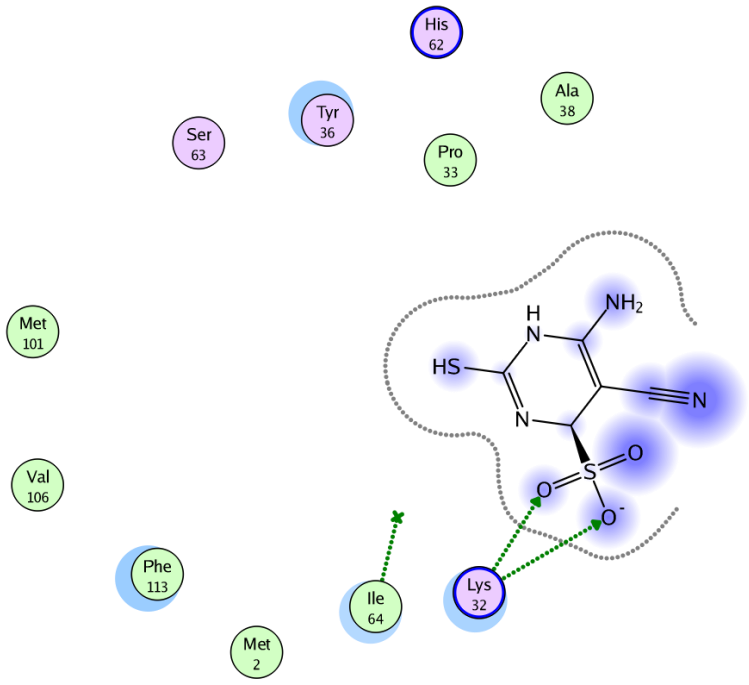


Figure B.4: interactions of compound V

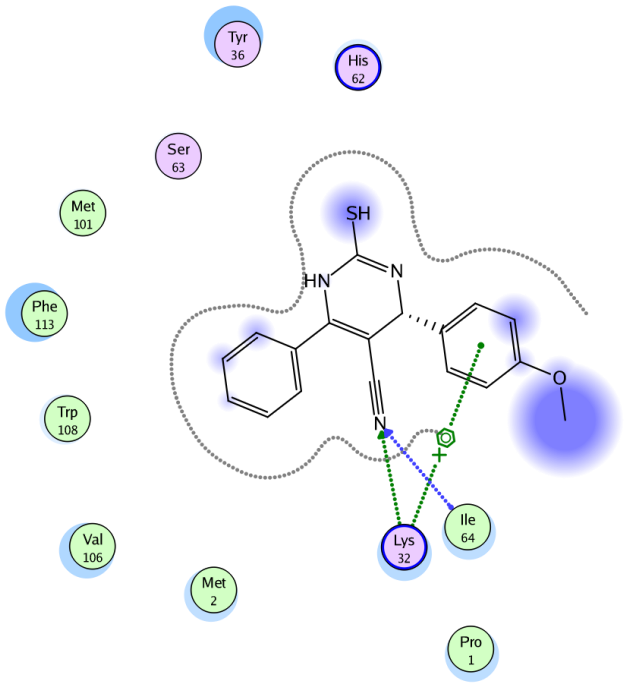


Figure B.5: Interactions of compound VII

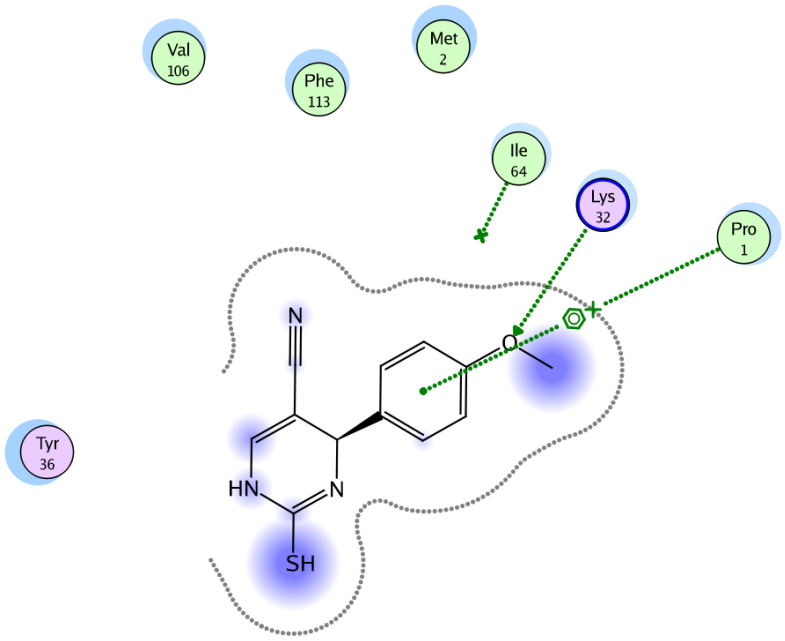


Figure B.6: Interactions of compound VIII

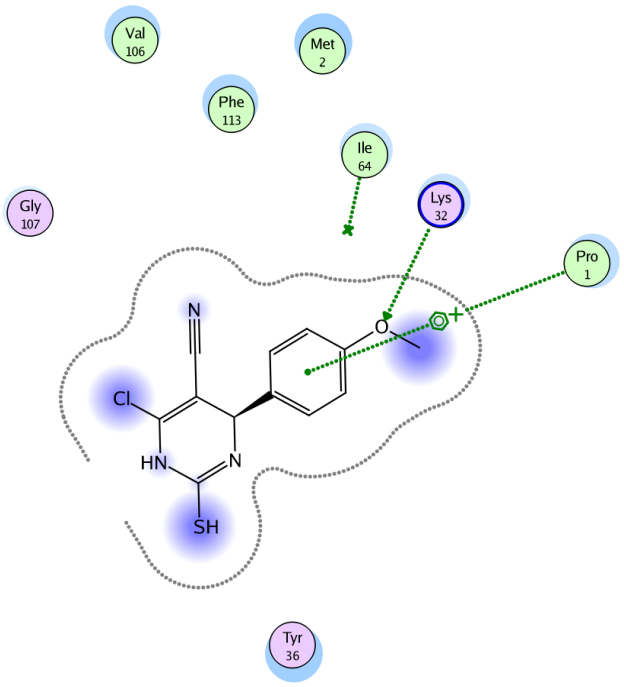


Figure B.7: Interactions of compound IX

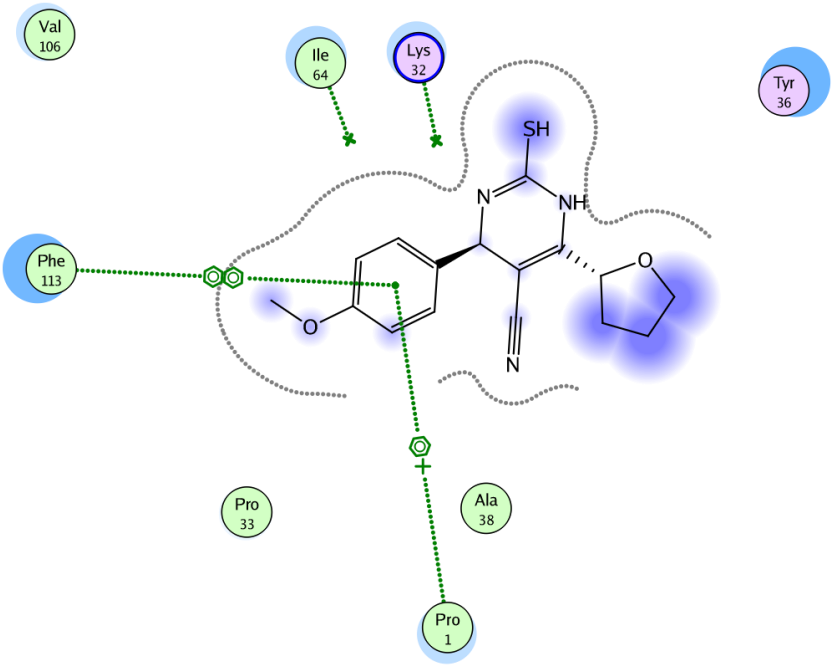


Figure B.8: Interactions of compound X

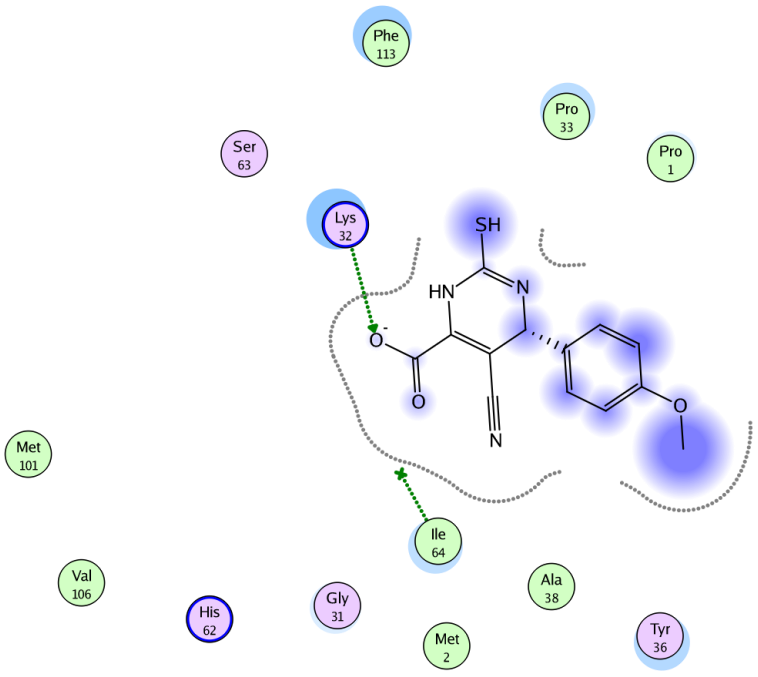


Figure B.9: Interactions of compound II

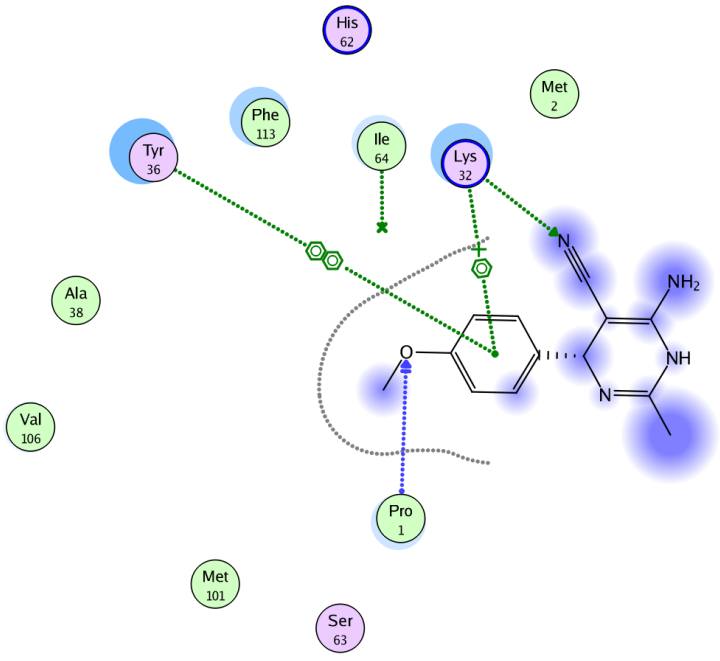


Figure B.10: Interactions of compound XV

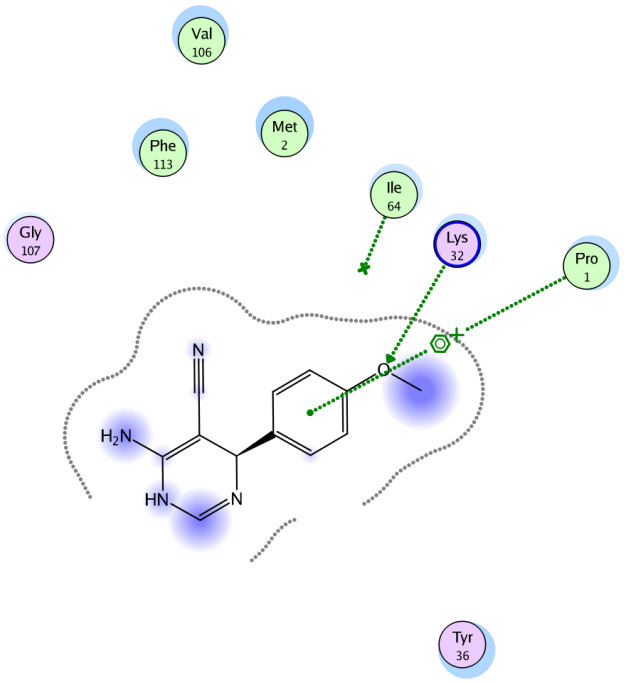


Figure B.11: Interactions of compound XVII

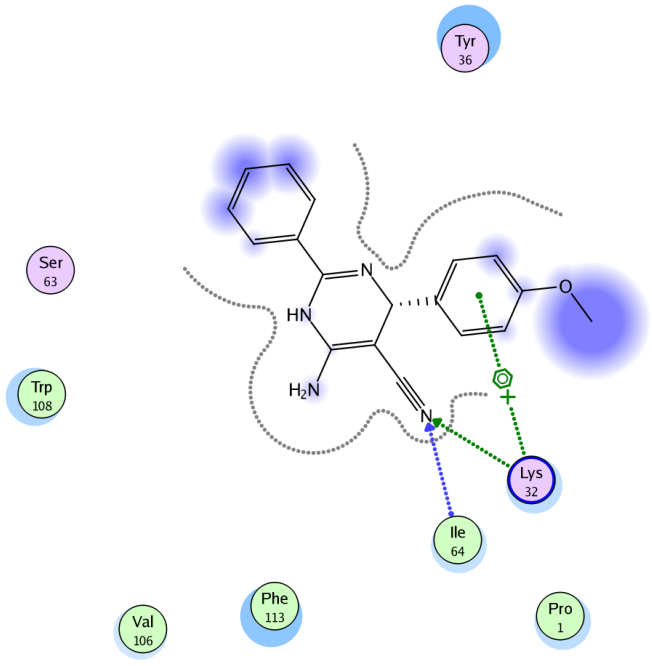


Figure B.12: Interactions of compound XIX

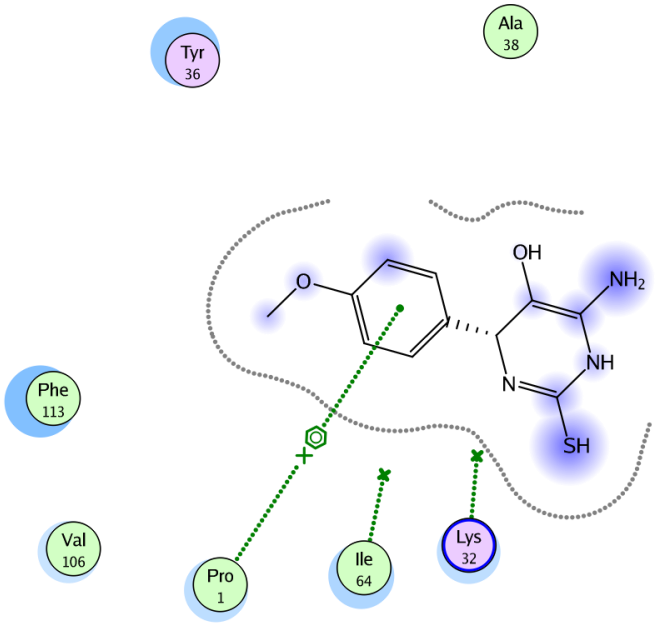


Figure B.13: Interactions of compound XIII

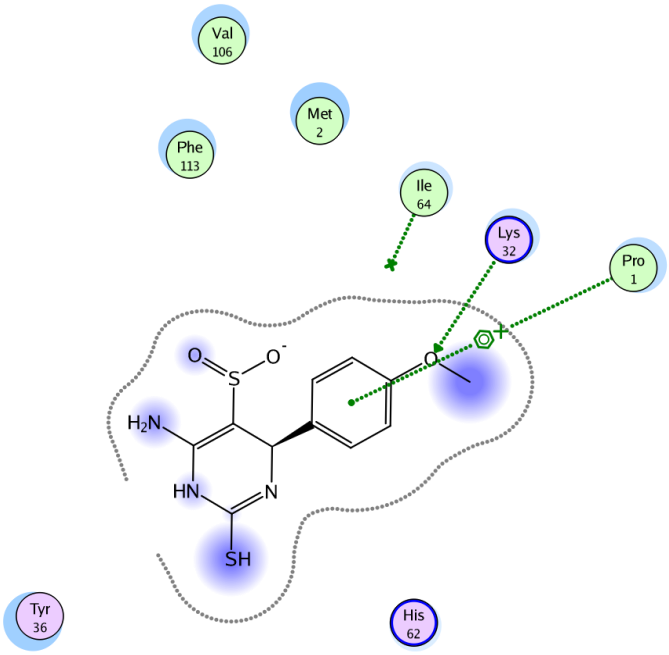


Figure B.14: Interactions of compound XIV

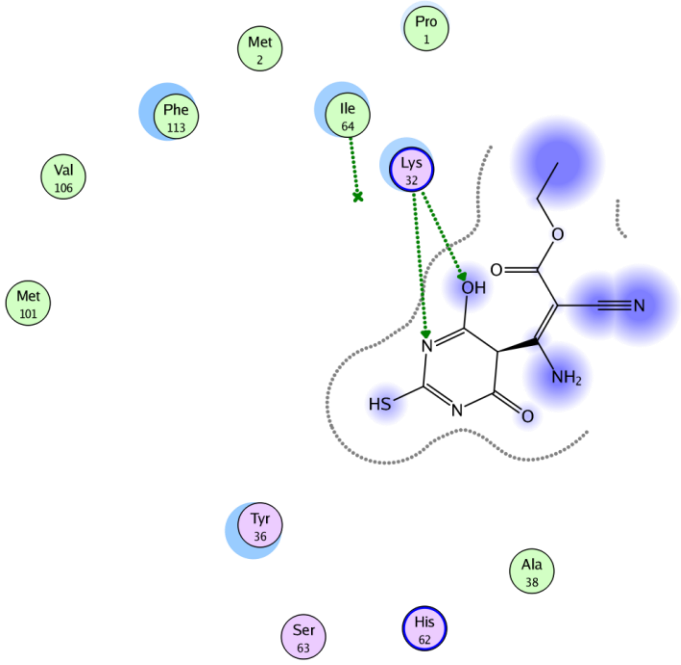


Figure B.15: Interactions of compound LXIX

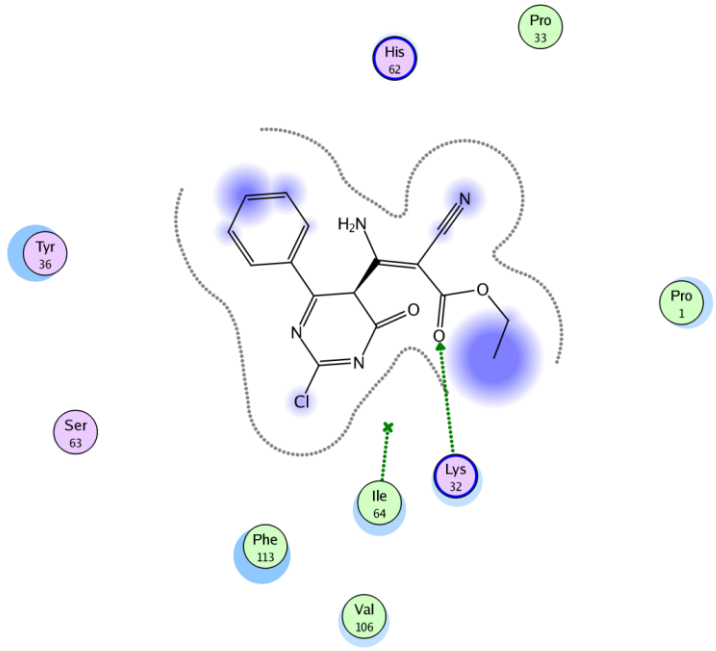


Figure B.16: Interactions of compound LXX

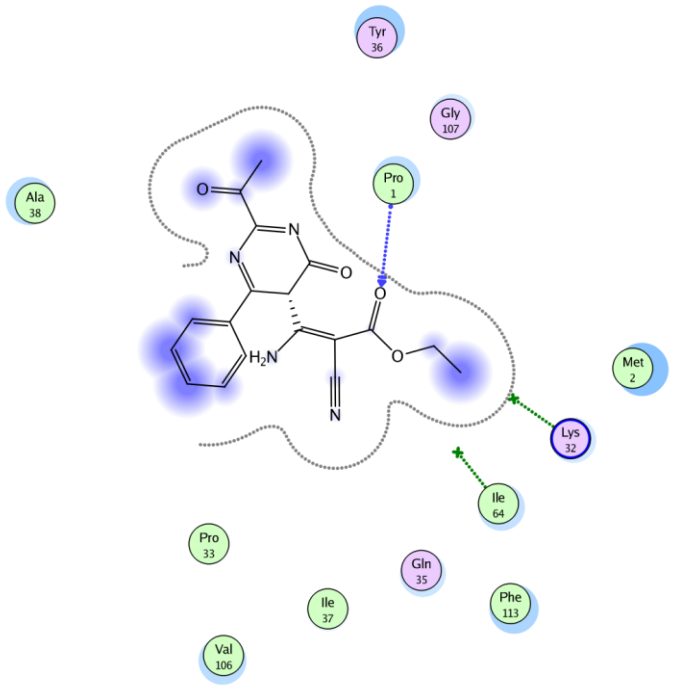


Figure B.17: Interactions of compound LXXIII

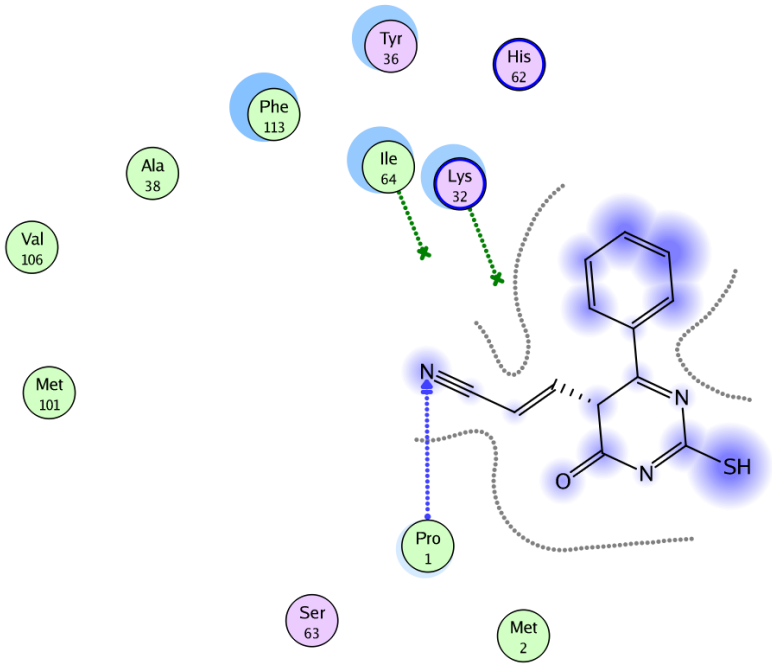


Figure B.18: Interactions of compound LXIV

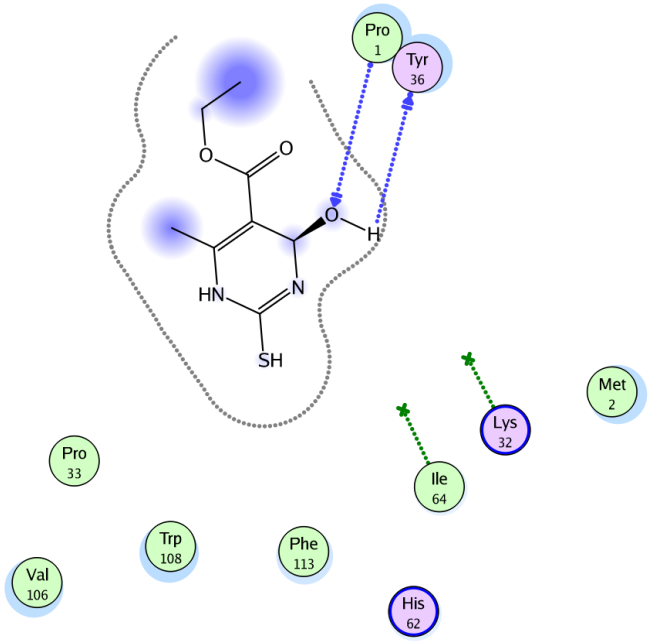


Figure B.19: Interactions of compound XXXIII

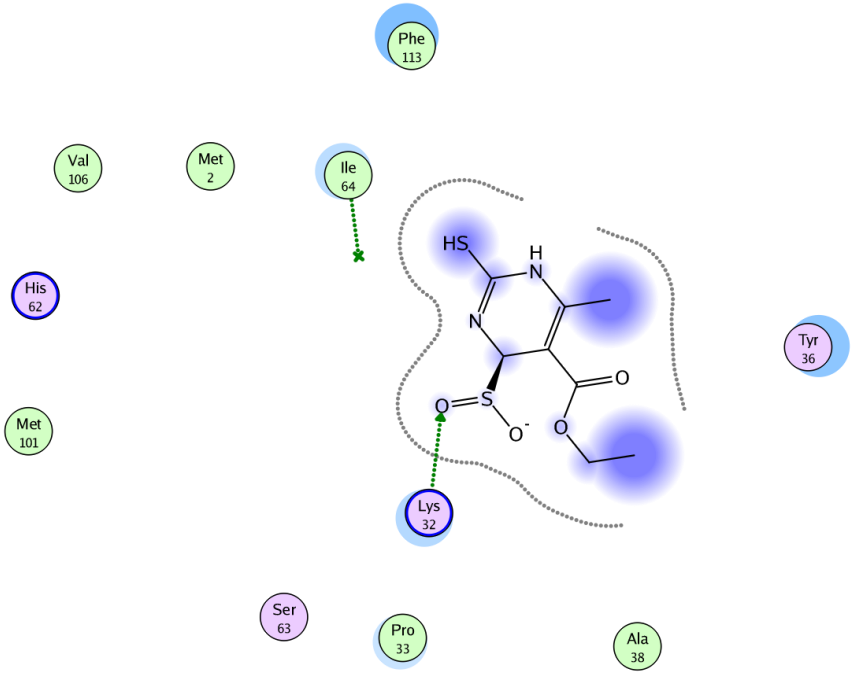


Figure B.20: Interactions of compound XXXIX

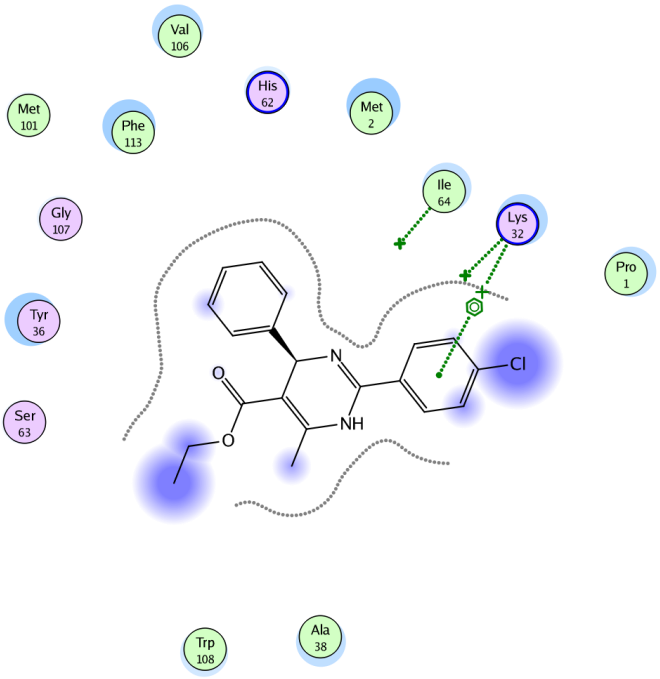


Figure B.21: Interactions of compound XX

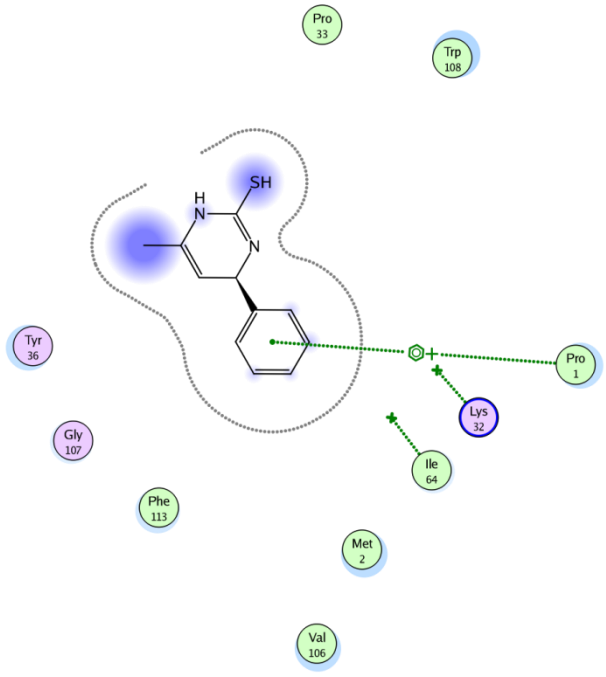


Figure B.22: Interactions of compound L

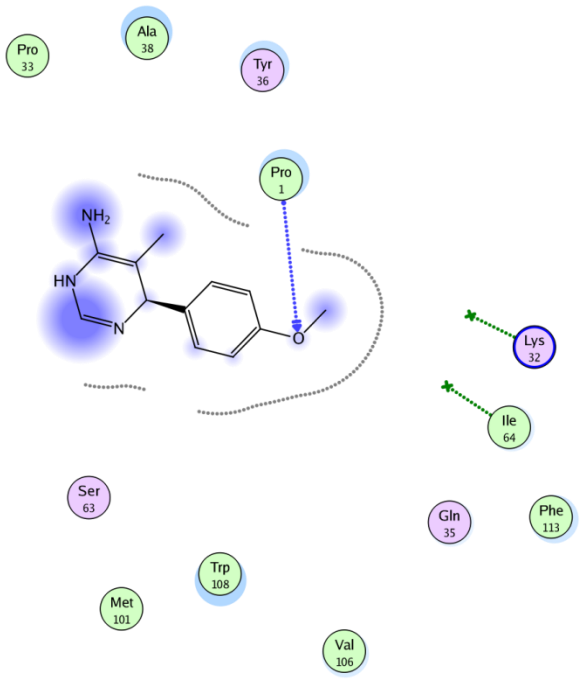


Figure B.23: Interactions of compound LII

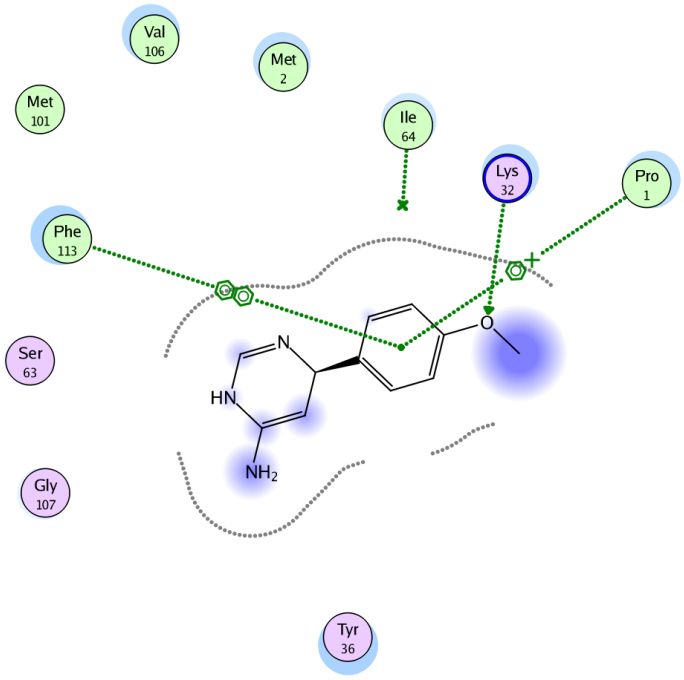


Figure B.24: Interactions of compound LIV

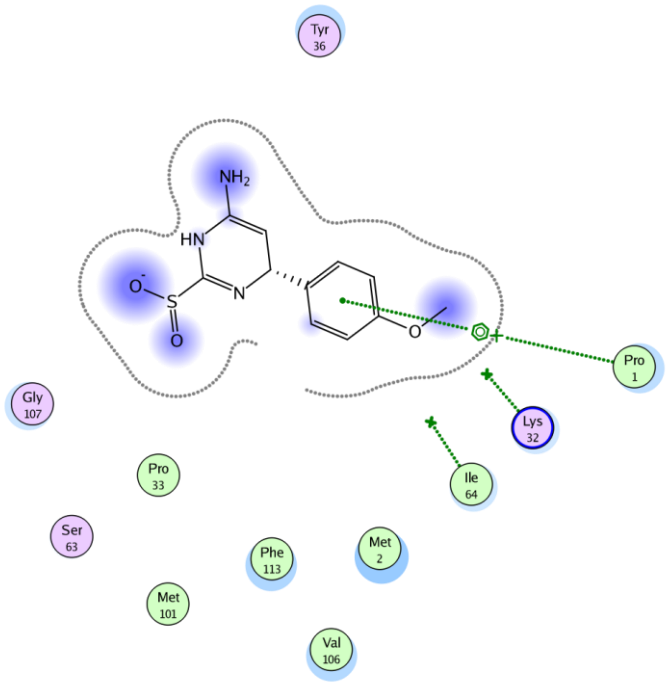


Figure B.25: interactions of compound LV

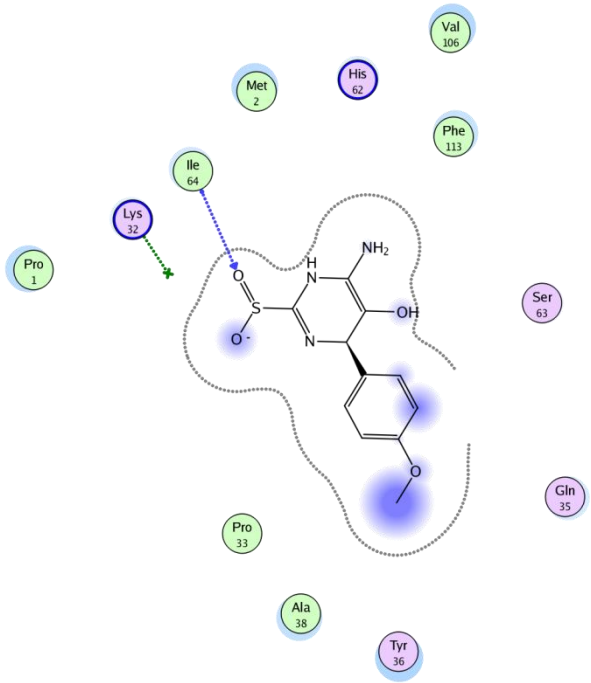


Figure B.26: interactions of compound LVI

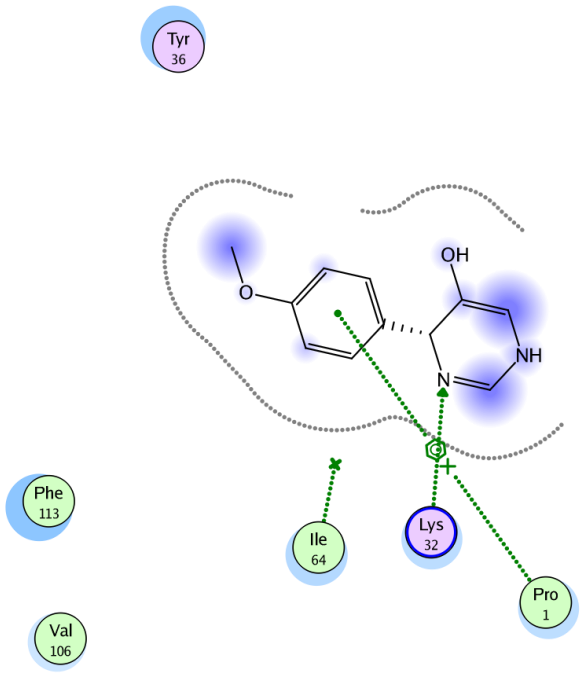


Figure B.27: Interactions of compound LVII

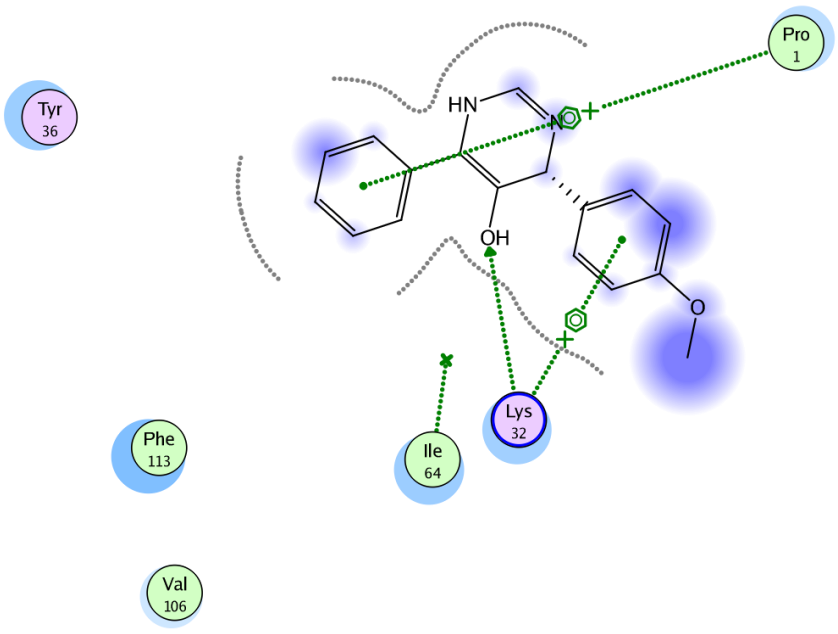


Figure B.28: Interactions of compound LIX

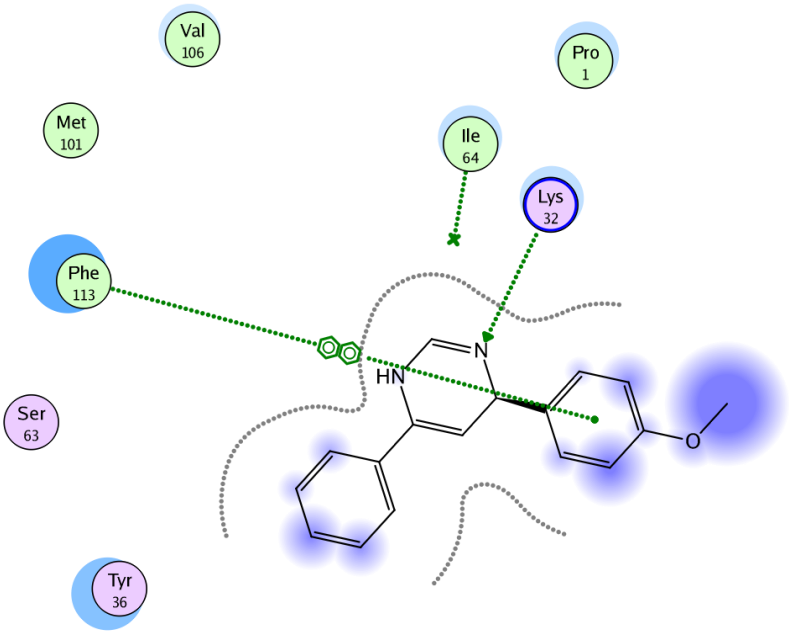


Figure B.29: Interactions of compound LX

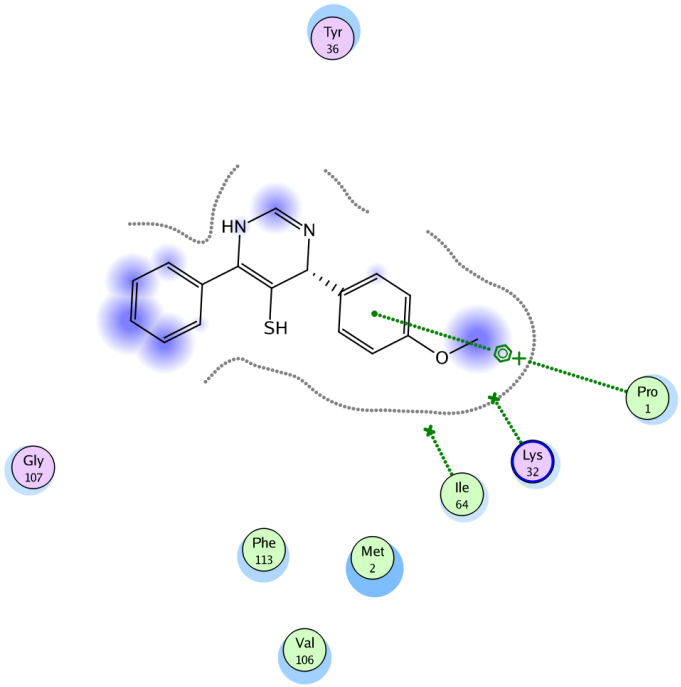


Figure B.30: Interactions of compound LXI

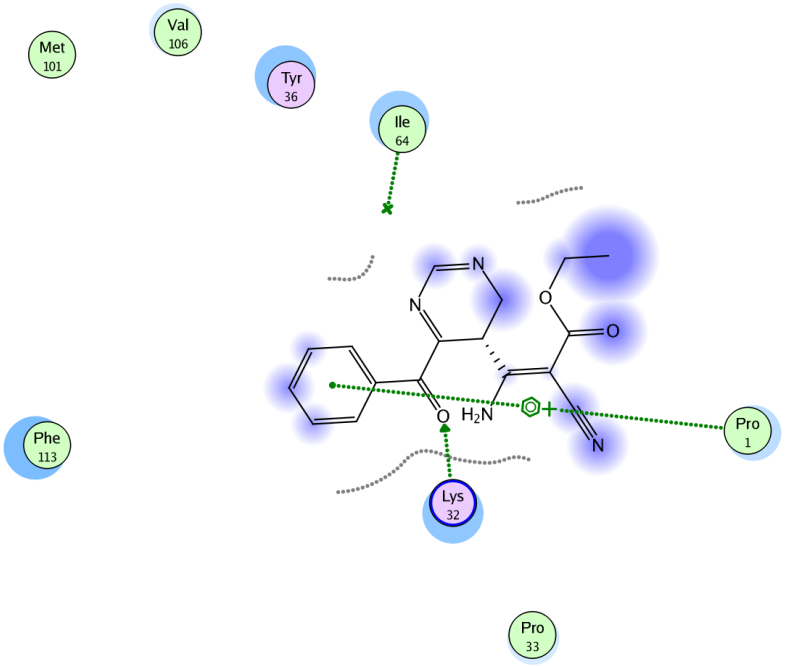


Figure B.31: Interactions of compound LXXXIX

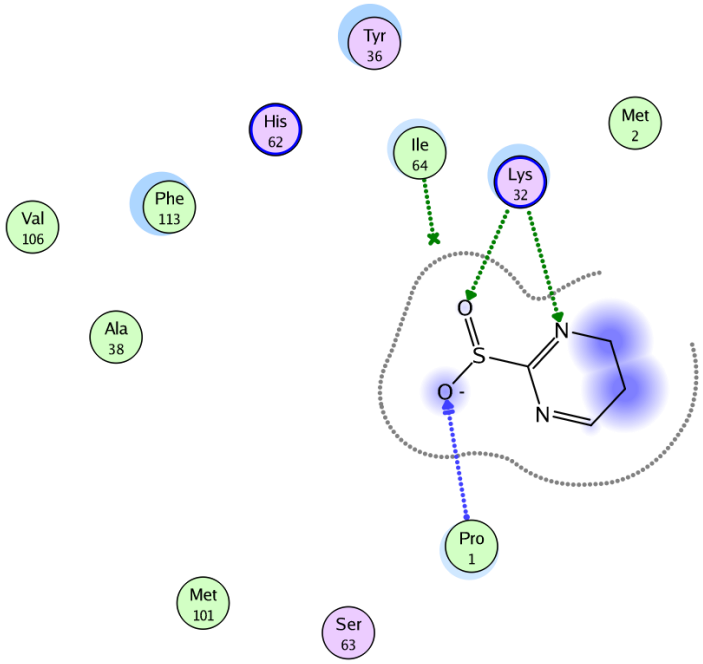


Figure B.32: Interactions of compound XCII

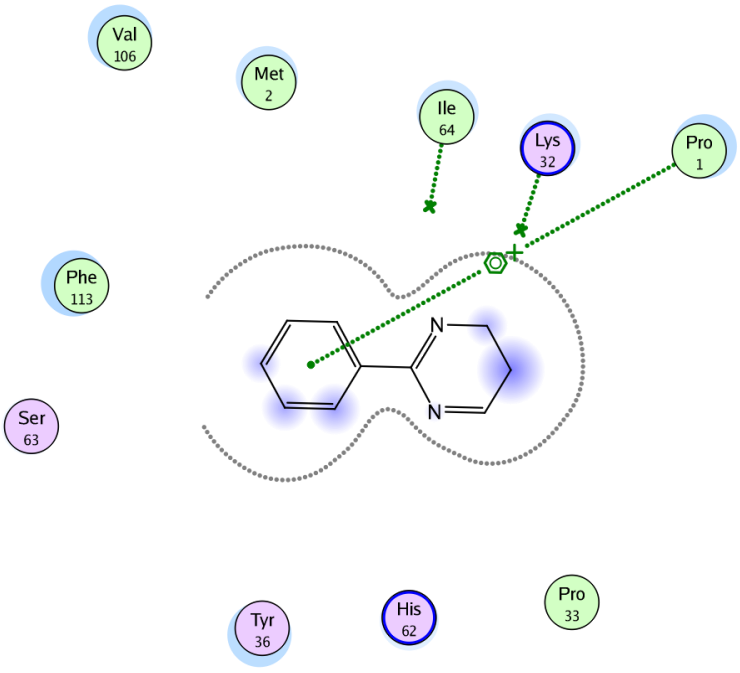


Figure B.33: Interactions of compound XCII

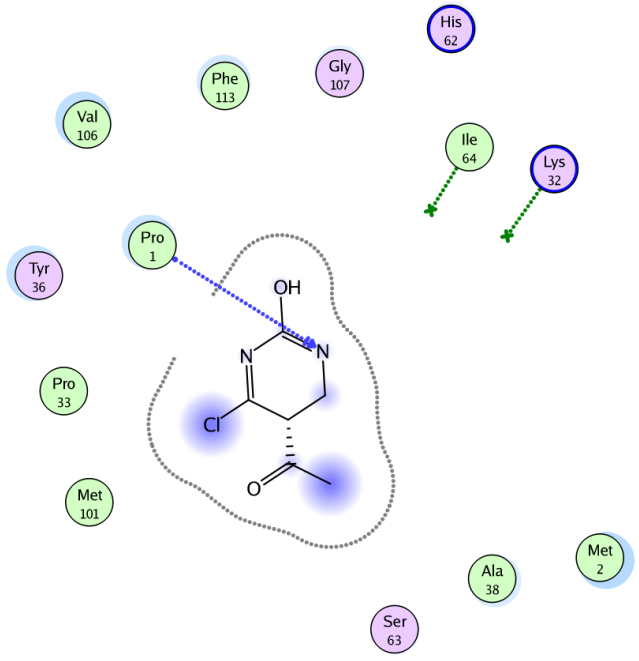


Figure B.34: Interactions of compound LXXVII

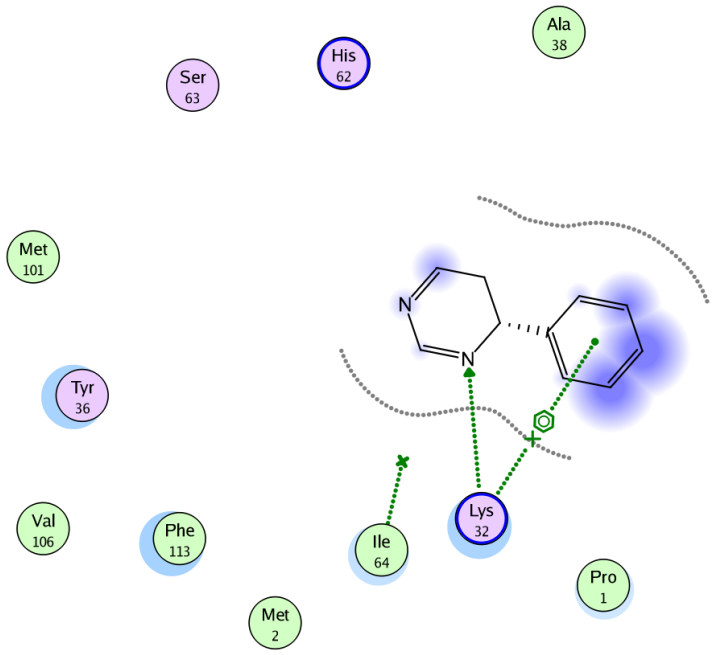


Figure B.35: Interactions of compound LXXXIV

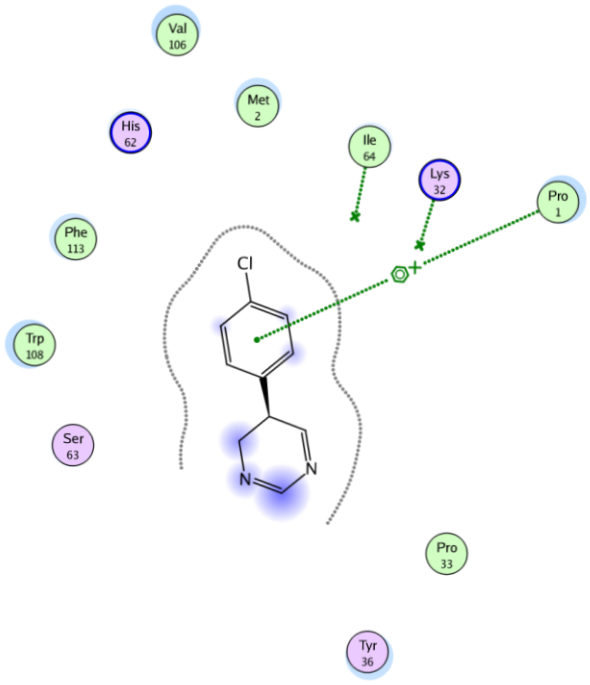


Figure B.36: Interactions of compound LXXXVI